IC 2023

The Institution of Engineering and Technology

IC 2023

The 6th International Conference on Innovative Computing

PROCEEDING

FEBRUARY 01, 2023

SINGAPORE & ONLINE EVENT

The 6th international Conference on Innovative Computing (IC 2023)

Co-located Conferences

The International Workshop on Technique for Language and Literature Information Modeling (LIM 2023)
The 7th International Conference on Big-data, IoT, Cloud computing Technologies and Applications (BICTA 2023)

IC 2023
Proceeding

Singapore & Online Event
February 1, 2023

Organized by

Frontier Computing Conference Group

Sponsors

IET

IET Taipei Local Network

Message from Organizing Committees

The International Conference on Innovative Computing (IC 2023) will be held in Singapore, 1 - 3, February 2023. We can finally hold a face-to-face conference after three years' COVID-19 pandemic. This event is the 6th event of the conference series, in which fruitful results can be found in IC2015 (Xiamen, China), IC2016 (Taichung, Taiwan), IC2020 (Ho Chi Minh, Vietnam), IC2021 (Online), and IC2022 (Online). Each event brings researchers worldwide together to have exciting and fruitful discussions as well as future collaborations. This conference series aims at providing an open forum to reach a comprehensive understanding of the recent advances and emergence of innovative computing in information technology, science, and engineering.

There are two international workshops and international conferences are jointly operated with IC2023 at the same time and place, i.e., Workshop on Technique for Language and Literature Information Modeling (LIM 2023), and The 7th International Conference on Big-data, IoT, Cloud computing Technologies and Applications (BICTA 2023), which are organized by FC conference group and Korean Institute of Information Technology, Korea Institute of information technology and innovation (KIITI) and SIEC Korea Chapter.

The papers accepted for inclusion in the conference proceeding primarily cover the topics: networking and communications, embedded system, soft computing, social network analysis, security and privacy, optics communication, ubiquitous, artificial intelligence, and pervasive computing. Many papers have shown their great academic potential and value and indicate promising directions of research in the focused realm of this conference series. We believe that the presentations of these accepted papers will be more exciting than the papers themselves, and lead to creative and innovative applications. We hope that the attendees (and readers as well) will find these results useful and inspiring to your field of specialization and future research.

On behalf of the organizing committee, we would like to thank the members of the organizing and the program committees, the authors, and the speakers for their dedication and contributions that make this conference possible. We appreciate the contributions of these experts and scholars to enrich our IC2023. We would like to thank and welcome all participants to IC2023. We also sincerely hope that all participants from overseas enjoy the technical discussions at the conference, build a strong friendship, and establish ties for future collaborations.

We send our sincere appreciation to the authors for their valuable contributions and the other participants of this conference. The conference would not have been possible without their support. Appreciates are also due to the many experts who contributed to making the event a success in Singapore.

IC2023 Organizing Committees
FC Conference Group
Korean Institute of Information Technology
Korea Institute of Information Technology and Innovation
SIEC Korea Chapter

February 2023

Organizing Committees

■ General Chair

Jen-Shiun Chiang, Tamkang University, Taiwan
Jing-Ming Guo, National Taiwan University of Science and Technology, Taiwan

■ Program Chairs

Chao-Tung Yang, Tunghai University, Taiwan
Jia Wei Chang, National Taichung University of Science and Technology, Taiwan
Hai Jiang, Arkansas State University, USA
Pedro, Peris López, Carlos III University of Madrid, Spain
Zhou Rui, Lanzhou University, China
Dmitry Novikov, Institute of Control Sciences V. A. Trapeznikov, Academy of Sciences, Russia
Daniel Shapiro, Clockrr Inc., Canada
Mahdi Zamani, Yale Univesity, USA

■ Workshop Chairs

Carmen Camara, Technical University of Madrid, Spain
Shih-Nung Chen, Asia University, Taiwan
Young-Ae Jung, Sun Moon University, Korea
Sujata Pandey, Amity University Uttar Pradesh, India
Jun Shen, University of Wollongong, Australia
Chih-Chuan Yeh, Overseas Chinese University, Taiwan

■ Special Session Chairs

Kuan-Chou Lai, National Taichung University of Education, Taiwan
Jenn-Wei Lin, Fu-Jen University, Taiwan
Xinghua Sun, Hebei North University, China
Chengjiu Yin, Kobe University, Japan
Xiaokang Zhou, Shiga University, Japan
Yishui Zhu, Chang'an University, China
Azhar Imran Mudassir, Air University, Islamabad, Paksitan

■ Local Arrangement Chair

YenJou Wang, Waseda University, Japan

■ Publicity Chairs

Soumya Banerjee, Birla Institute of Technology, India
Jindrich Kodl, Authorised expert in security of information systems, Czech Republic
Min-Feng Lee, National Museum of Natural Science, Taiwan
Poonphon Suesaowaluk, Assumption University of Thailand, Thailand
Shing-Chern You, National Taipei University of Technology, Taiwan
Linjing Wei, Gansu Agricultural University, China
Jun-Hong Shen, Asia University, Taiwan
Ching-Ta Lu, Asia University, Taiwan
Goldina Ghosh, Indian Institute of Information Technology, India

Table of Content

FCGSM: Fast Conjugate Gradient Sign Method for Adversarial Attack on Image Classification

Xiaoyan Xia[1], Wei Xue[1,2,3(✉)], Pengcheng Wan[4], Hui Zhang[1]
Xinyu Wang[1], and Zhiting Zhang[1]

[1] School of Computer Science and Technology, Anhui University of Technology, Maanshan 243032, China
xuewei@ahut.edu.cn
[2] Anhui Engineering Research Center for Intelligent Applications and Security of Industrial Internet, Maanshan 243032, China
[3] Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei 230088, China
[4] National Key Laboratory of Science and Technology on Automatic Target Recognition, National University of Defense Technology, Changsha 410073, China

Abstract. Deep neural network is sensitive to adversarial samples that crafted by adding imperceptible perturbations to original images, and many methods of generating adversarial samples have emerged. Although existing methods based on gradient direction have good attack performance, some ill-conditioned issues may reduce their performance on occasion. In this paper, we propose a novel attack method based on three-terms conjugate gradient direction, which is effectively for improving this limitation, and its is named as fast conjugate gradient sign method (FCGSM). The proposed method FCGSM can jump from the local maximum during the process of finding the maximum value of loss function, thus generating more adversarial samples than the SOTA methods APGD and ACG. Experiments conducted on two benchmark datasets show that the FCGSM works well in attacking deep neural network-based classification models.

Keywords: adversarial machine learning, deep learning, conjugate gradient, adversarial attack, adversarial training

1 Introduction

Deep neural networks (DNNs) have shown the tremendous capacity and ability in making a good progress in the filed of computer vision. However, it is also demonstrated that DNNs are highly vulnerable to adversarial samples [18, 4], which are manufactured by adding small-imperceptible perturbations on input and make a model output incorrect classification. Plenty of methods in generating adversarial samples have been proposed since it helps to evaluate the vulnerability of models and enhance the robustness of various DNN algorithms by adversarial training [15, 5]. Moreover, It is important for improving the robustness of models by adversarial training to learn how to generate adversarial samples with better transferability [10, 19].

With full access to the knowledge structure of a model including the composition and parameters, most methods of generating adversarial samples can successfully attack the transparent model. This type of attack is known as a white-box attack, including optimization-based methods such as box-constrained L-BFGS [18], C&W [1], and gradient-based methods such as fast gradient sign method (FGSM) [4], iterative-FGSM (IFGSM) [9], projected gradient descent (PGD) [11], which use the steepest gradient to update the sign. In addition, some attack methods utilize the current and past gradient information to determine the next update, such as momentum iterative-FGSM (MI-FGSM) [3] and Auto-PGD (APGD) [2]. However, the steepest descent method may be inefficiently attack the deep learning models due to the fact that the convergence speed is relatively slow and the objective function of adversarial attack is highly nonconvex. To solve the above challenges, [20] applies the conjugate gradient (CG) method to generate adversarial samples. Although the traditional conjugate gradient method has some improvement in the accuracy of calculation and convergence of the objective function, for nonlinear objective functions, sometimes infinite cycling away from the optimal solution.

To this end, in this paper we propose a new adversarial attack algorithm, fast conjugate gradient sign method (FCGSM), based on a three-terms CG direction with adaptive stepsize selection strategy. In summary, we make the following contriutions:

1. We propose an effectual adversarial machine learning algorithm that based on the fast gradient sign method and auto CG attack method, which possesses the ability to search more diverse direction and generate more diverse adversarial samples;
2. We further use the obtained adversarial samples to execute the adversarial training to improve the robustness of the classification models that based on DNNs, and the corresponding experiments demonstrate adversarial training is an effective security defensive mechanism.

2 Proposed Method

In this section, we present the proposed adversarial attack method. In order to better describe our approach, we first give a brief review of the FGSM method and the CG method.

2.1 FGSM method

FGSM is to generate an adversarial sample x^{adv} by stacking the original image x with variations that are consistent with the direction of gradients. Suppose that J is the object function, we can use it to compute the current gradient and then obtain the perturbation $\eta = \epsilon sign(\nabla_x J(x, y))$ constrained by $\| x^{adv} - x \|_\infty \leq \epsilon$, where $\epsilon > 0$ is an artificial parameter. Subsequently, the adversarial example generated by FGSM can be presented by $x^{adv} = x + \eta$.

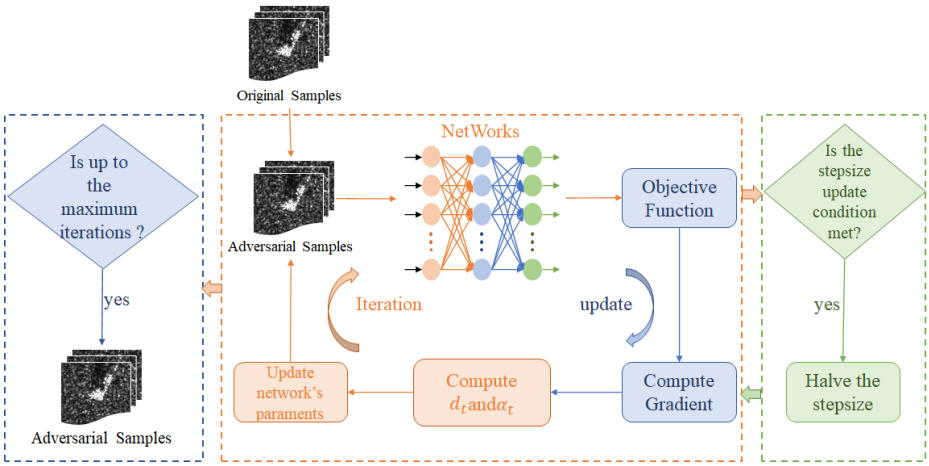


Fig. 1: Flowchart of FCGSM.

2.2 CG method

CG method is a very efficient optimization algorithm. Consider the minimization problem $\min_{x\in R^n} f(x)$, where the objective function f is differentiable, then given an initial point x_0, a CG method generates a sequence $\{x_t\}$ by $x_{t+1} = x_t + \alpha_t d_t$, where α_t is the stepsize usually obtained by a line search, and the search direction d_t is computed by $d_k = -\nabla f(x_t) + \beta_t d_{t-1}$ with $d_0 = -\nabla f(x_0)$. Here, β_t is the CG update parameter. Some well-known formulas for β_t are β_t^{HS}, β_t^{PR}, $\beta_k t^{DY}$, et al., see [6] for details.

2.3 FCGSM method

Based on the theories mentioned above, we now describe the proposed FCGSM method. The flowchart of FCGSM is shown in Fig. 1. Note that adversarial attack can be formulated as a maximization optimziation problem. Consider the problem $\max_{x\in R^n} f(x)$, where f is a continuous. Given the initial point x_0 and the initial search direction $d_0 = \nabla f(x_0)$, we then update x_t^{adv} (i.e., adversarial sample) at the t-th iteration with d_t as follows

$$x_{t+1}^{adv} = x_t^{adv} + \alpha_t \cdot sign(d_t), \tag{1}$$

and

$$d_t = \nabla f(x_t) - \beta_t d_{t-1} + \gamma_t y_{t-1}, \tag{2}$$

where d_t is the so-called three-terms CG direction. We set $\beta_t = \beta_t^{PR} (= \frac{(g_t)^T y_t}{(g_{t-1})^T g_{t-1}})$ and $\gamma_t = \frac{(g_t)^T d_{t-1}}{(g_{t-1})^T g_{t-1}}$, where $g_t = \nabla f(x_t)$, $y_t = g_t - g_{t-1}$.

For the stepsize α_t, we calculate it according to the following two conditions proposed in [2]: (1) $\sum_{i=w_{j-1}}^{w_j} N < \rho \cdot (w_j - w_{j-1})$, (2) $\eta^{(w_{j-1})} = \eta^{(w_j)}$ and $f_{\max}^{(w_{j-1})} = f_{\max}^{(w_j)}$, where N indicates the count of the cases for which $f(x_{t+1} > f(x_t))$ holds and $f_{\max}^{(w_j)}$ is the highest objective value in the w_j iterations.

In summary, FCGSM generates the adversarial samples by

$$x_{t+1}^{adv} = Cilp_{x,\epsilon}\left\{x_t^{adv} + \alpha_t \cdot sign(d_t)\right\}, \tag{3}$$

where $Cilp_{x,\epsilon}$ means that x_t^{adv} has be cliped into the ϵ-neighbourhood of the original sample x at each iteration to control the perturbation amplitude.

3 Experiments

In this section, we present comparison experiments with ACG [20] and APGD [2] to show the feasibility and efficiency of the proposed method.

3.1 Experimental setup

Datasets and Models. We choose sixe classification models (VGG-11, VGG-13, and VGG-16; ResNet-18, ResNet-34, and ResNet-50) and two benchmark datasets (MSTAR and CIFAR-10).

Hyperparameter setting. We set the maximum perturbation $\epsilon = 8/255$, the initial stepsize $\eta^{(0)} = 0.01$, the stepsize selection parameter $\rho = 0.75$, and the maximum number of iterations $T = 100$.

Choice of loss function. ACG and APGD use the CW loss proposed in [1] and DLR loss proposed in [2], respectively. For FCGSM, we choose the cross-entropy loss as the objective function.

Evaluation metrics. We adopt the evaluation metrics based on accuracy, attack success rate (ASR for short), and the diversity which is described as the Euclidean norm of two successive adversarial samples, where the ASR is defined as $ASR = \frac{\text{accuracy before attacking}-\text{accuracy after attacking}}{\text{accuracy before attacking}}$.

3.2 Analysis of comparison results

Table 1 reports the time of generating adversarial samples on MSTAR dataset for VGG-16 by using three attack methods. The time is calculated from the first sample started to be attacked to the end of the attack on the last example. We can see that the generating time of FCGSM is the lowest.

Table 1: Time of generating adversarial samples on MSTAR dataset.

CPU	RAM	GPU	VGG-16	APGD	ACG	FCGSM
Intel(R) Xeon(R)	24GB	NVIDIA GeForce	ASR	95.46	98.15	98.48
Silver 4314 × 4		RTX 3090 × 2	time	**5m46s**	**3m24s**	**3m22s**

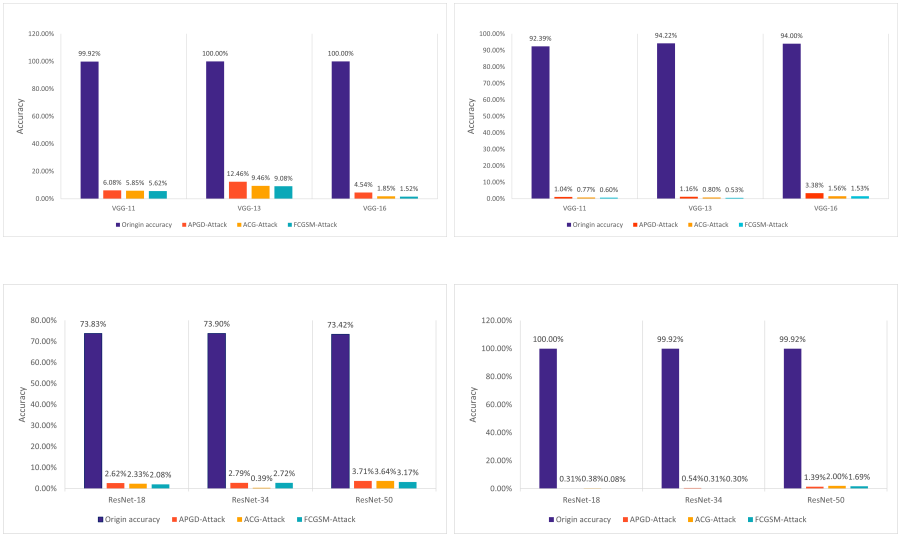


Fig. 2: Comparision results of accuracy on the original sample set and adversarial sample set. The left is on MSTAR, and the right is on CIFAR-10.

Table 2: The ASR of FCGSM, ACG and APGD for attacking the trained models. The highest ASR is in bold, and the second is underlined. diff is the difference between bold and underlined.

MSTAR	Attack Success Rate (%)			
Architecture	APGD	ACG	FCGSM	diff
VGG-11	93.92	<u>94.15</u>	**94.38**	0.23
VGG-13	87.54	<u>90.54</u>	**90.92**	0.38
VGG-16	95.46	<u>98.15</u>	**98.48**	0.33
ResNet-18	97.38	<u>97.67</u>	**97.92**	0.25
ResNet-34	97.21	**99.61**	<u>97.28</u>	2.33
ResNet-50	96.29	<u>96.36</u>	**96.83**	0.47
CIFAR-10	Attack Success Rate (%)			
Architecture	APGD	ACG	FCGSM	diff
VGG-11	98.96	<u>99.23</u>	**99.40**	0.17
VGG-13	98.84	<u>99.20</u>	**99.47**	0.27
VGG-16	96.62	<u>98.44</u>	**98.47**	0.03
ResNet-18	<u>99.69</u>	99.62	**99.92**	0.23
ResNet-34	99.46	<u>99.69</u>	**99.70**	0.01
ResNet-50	**98.61**	98.00	<u>98.31</u>	0.30

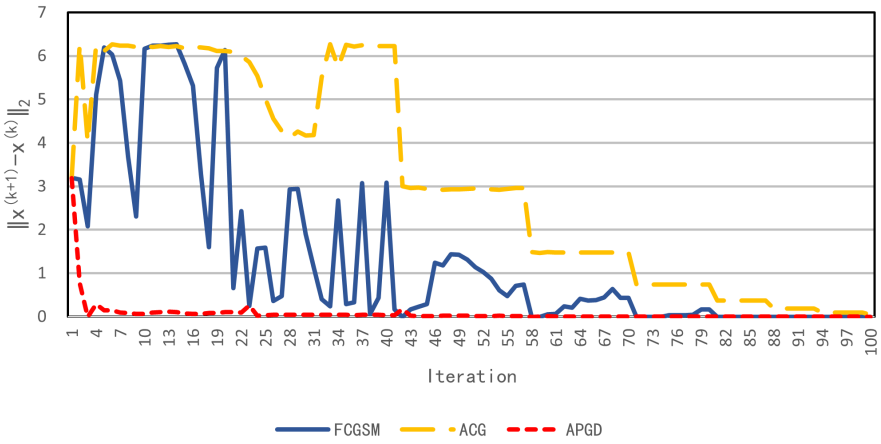


Fig. 3: Comparison of the diversity of search direction by three methods.

Fig. 2 shows comparison results of six models trained on MSTAR and CIFAR-10 after being attacked by FCGSM as well as ACG and APGD. The highest is the original accuracy, and it can be seen that the accuracy value decreases significantly after attacking. Table 2 reports the ASR results, and overall, FCGSM has a higher ASR than other two methods in all scenarios.

We further examine the diversity of search direction by FCGSM. From the Fig. 3, we can see that the amount of perturbation between two points fluctuates widely, which indicates that FCGSM possesses the ability to search more diverse direction and generate more diverse adversarial samples.

3.3 Adversarial training

In order to improve the robustness of the three classification models, we conduct the adversarial training by using the oabtined adversarial samples. Specifically, we divide the adversarial samples on VGG-16 generated by FCGSM into two parts. 70% of these are put into the training set for the adversarial training, and 30% of the adversarial samples generated from ACG and APDG are simultaneously chosen as test set to calculate the classification accuracy. We selecte VGG-16 as the adversarial training model and validate the effect on MSTAR and CIFAR-10. Tables 3 and 4 show that the robustness of the model after adversarial training is significantly improved, while the model after adversarial training using the adversarial samples generated by FCGSM has good defense against the attacks of the adversarial samples generated by ACG and APGD.

Table 3: Adversarial training results on MSTAR.

MSTAR	Accuracy		
Architecture	APGD	ACG	FCGSM
VGG-16	34.53	21.70	33.76
RobustVGG-16	40.92	42.71	40.67

Table 4: Adversarial training results on CIFAR-10.

CIFAR-10	Accuracy		
Architecture	APGD	ACG	FCGSM
VGG-16	8.04	9.13	10.07
RobustVGG-16	44.82	46.56	46.02

4 Conclusion

In this paper, we proposed a three-terms conjugate gradient direction-based adversarial attack method, which has more diverse search ability to improve the attack performance. Experimental results verified the validity and feasibility of the proposed method, and in the future work we will apply this method to attack deep learning detection models.

Acknowledgements. This work was supported in part by the Anhui Provincial Natural Science Foundation (Grant No. 2208085MF168), the Program for Synergy Innovation in the Anhui Higher Education Institutions of China (Grant No. GXXT-2022-052), and the College Students' Innovation and Entrepreneurship Training Programs (Grant Nos. 202210360079, 202110360079, and S202110360291).

References

1. Carlini, N., Wagner, D. A.: Towards Evaluating the Robustness of Neural Networks. In: IEEE Symposium on Security and Privacy, pp. 39-57 (2017)
2. Croce, F., Hein, M.: Reliable evaluation of adversarial robustness with an ensemble of diverse parameter-free attacks. In: International Conference on Machine Learning, pp. 2206-2216 (2020)
3. Dong, Y., Liao, F., Pang, T., Su, H., Zhu, J., Hu, X., Li, J.: Boosting Adversarial Attacks with Momentum. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 9185-9193 (2018)
4. Goodfellow, I. J., Shlens, J., Szegedy, C.: Explaining and Harnessing Adversarial Examples. In: International Conference on Learning Representations, (2015)

5. Gowal, S., Qin, C., Uesato, J., Mann, T., Kohli, P.: Uncovering the Limits of Adversarial Training against Norm-Bounded Adversarial Examples. arXiv:2010.03593v3, (2021)
6. Hager, W. W., Zhang, H.: Algorithm 851: CG_DESCENT, a conjugate gradient method with guaranteed descent. ACM Transactions on Mathematical Software, 32, 113-137 (2006)
7. Ibrahim, A., Shareef, S.: Modified Conjugate Gradient Method For Training Neural Networks Based On Logisting Mapping. Journal of University of Duhok, 22(1), 45-51 (2019)
8. Krizhevsky, A., Sutskever, I., Hinton, G.: ImageNet Classification with Deep Convolutional Neural Networks. Communications of the ACM, 60(6), 84-90 (2017)
9. Kurakin, A., Goodfellow, I., Bengio, S.: Adversarial Machine Learning at Scale. In: International Conference on Learning Representations, (2017)
10. Liu, Z., Liu, Q., Liu, T., Xu, N., Lin, X., Wang, Y., Wen, W.: Feature Distillation: DNN-Oriented JPEG Compression Against Adversarial Examples. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 860-868 (2019)
11. Madry, A., Makelov, A., Schmidt, L., Tsipras, D., Vladu, A.: Towards Deep Learning Models Resistant to Adversarial Attacks. In: International Conference on Learning Representations, (2018)
12. Ma, G., Lin, H., Jin, W., Han, D.: Two Modified Conjugate Gradient Methods For Unconstrained Optimization With Applications In Image Restoration Problems. Journal of Applied Mathematics and Computing, 68, 4733-4758 (2022)
13. Papernot, N., Mcdaniel, P., Goodfellow, I., Jha, S., Celik, Z. B., Swami, A.: Practical Black-Box Attacks against Machine Learning. In: ACM on Asia Conference on Computer and Communications Security, pp. 506-519 (2017)
14. Simonyan, K., Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition. In: International Conference on Learning Representations, (2015)
15. Song, C., He, K., Lin, J., Wang, L., Hopcroft, J. E.: Robust Local Features for Improving the Generalization of Adversarial Training. In: International Conference on Learning Representations, (2020)
16. Sun, J., Zhang, J.: Global Convergence of Conjugate Gradient Methods without Line Search. Annals of Operations Research, 103, 161-173, (2001)
17. Szegedy, C., Wei, L., Jia, Y., Sermanet, P., Reed, S. E., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going Deeper with Convolutions. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-9 (2015)
18. Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., Fergus, R.: Intriguing Properties of Neural Networks. In: International Conference on Learning Representations, (2014)
19. Xie, C., Zhang, Z., Zhou, Y., Bai, S., Yuille, A. L.: Improving Transferability of Adversarial Examples With Input Diversity. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 2730-2739 (2019)
20. Yamamura, K., Sato, H., Tateiwa, N., Hata, N., Mitsutake, T., Oe, T., Ishikura, H., Fujisawa, K.: Diversified Adversarial Attacks based on Conjugate Gradient Method. In: International Conference on Machine Learning, pp. 24872-24894 (2022)
21. Zhang, L., Zhou, W., Li, D.-H.: A descent modified Polak–Ribiere–Polyak conjugate gradient method and its global convergence. IMA Journal of Numerical Analysis, 26(4), 629-640 (2006)

A lightweight network for detecting small targets in the air

Jiaxin Li [1], Hui Li[1], Ting Yong[2,3,*]
Xingyu Hou [4]

[1] Institute of System Engineering

{lihui940216, sc_lijiaxin}@qq.com

[2] The Institute of North Electronic Equipment.

Beijing, China,

[3] National Key Laboratory of Science and Technology on Information System Security,

Beijing, China,

25185703@qq.com

[4] Zhengzhou Xinda Institute of Advanced Technology,

Zhengzhou, China,

1029588176@qq.com

Abstract. Fast and accurate detection and identification of small airborne targets are of great importance, to security in the air. Unmanned aerial vehicle detection algorithms are mostly deployed on edge devices, and a yolov5-based aerial target lightweight detector is proposed by compressing channels and network cropping for the limited resource characteristics on edge devices. Firstly, the shallow cross-stage partial module is extended and optimized when designing the feature extraction network to maximize the use of shallow features. Secondly, the network is cropped to reduce the number of down-sampling, which makes the computation faster. Finally, the pyramid network used for feature fusion is simplified by modifying from two upsampling operations and two downsampling operations to only one upsampling operation. On the homemade dataset, the proposed Yolo-mini achieves 94.44% mean average accuracy on the test set and the Giga floating-point operations per second of the model is only 3.2, which achieves a better balance of accuracy and computation compared to other lightweight algorithms.

Keywords: Object Detection; UAV; Small Object; Neural Networks; Deep Learning.

1 Introduction

With the rapid development of the unmanned aerial vehicle (UAV) industrial industry, UAVs are widely used in various industries, such as industry, urban management, sports, peacekeeping, transportation, power cruising, agriculture plant protection, express delivery and disaster rescue, and other scenarios in which UAVs are used [1,2]. Also in the military battlefield, drones are frequently seen as weapons [3,4]. Despite attracting widespread attention in different civilian and commercial applications, there is no doubt that drones pose a threat to airspace security and may endanger people and property. Drones are also likely to be used for nefarious purposes, such as this collecting data from private areas, tracking people alive vehicles as spies, remote bugging, carrying explosives for unpredictable terrorist attacks in public places, etc. Therefore, the development of drone countermeasure systems is crucial [5–7].

There are many methods for UAV detection based on video images, such as Faster-RCNN[8], SSD[9], YOLO[10], etc. These algorithms have achieved good results, but detection accuracy and detection speed are a pair of oxymorons, and these algorithms have their advantages and disadvantages, high detection accuracy means complex network structure, and complex network structure means limited detection speed.

To balance the detection speed of the model, a series of lightweight network structures have been proposed in the industry [11–14]. Widodo Budiharto et al. Constructed a detection model using Mobile Net and SSD to implement a fast detection algorithm with an accuracy that meets the practical requirements[15]. MobileNet is a lightweight network structure for edge devices. Sheng Yuan et al. proposed a lightweight network structure based on yolov5. They used the proposed CI network structure and then pruned the Neck structure. The method performs well on a bit of a stone detection task[16]. Haiying Liu et al. proposed an improved feature fusion method based on PANet and BiFPN, which effectively improves the detection of small objects[17]. These algorithms guarantee the model accuracy to meet the demand while ensuring the model size, and these models are relatively small in computation and suitable for deployment at the edge.

UAV countermeasure systems are mostly applied to edge devices. With the rapid development of UAVs, the detection accuracy and detection speed of UAV detection algorithms need to be further improved. Yolov5 is a more classical target detection algorithm with not bad detection accuracy and detection speed. In this paper, a lightweight network based on the yolov5 is proposed to detect small targets in the air. The backbone network is trimmed and optimized to improve the speed of the algorithm and reduce the model size without losing the accuracy of the algorithm, which is easy to deploy on embedded devices. The contributions of this paper are as follows.

1. A model tailoring idea for small targets is proposed.
2. The cross-stage partial (CSP) module is optimized for better expressiveness.
3. A detection algorithm for small targets in the air is proposed, which outperforms the original model in terms of speed and accuracy.

4. The model can identify the detected object in the images and mark the object's bounding box by joining the results across the regions.

2 Materials and Methods

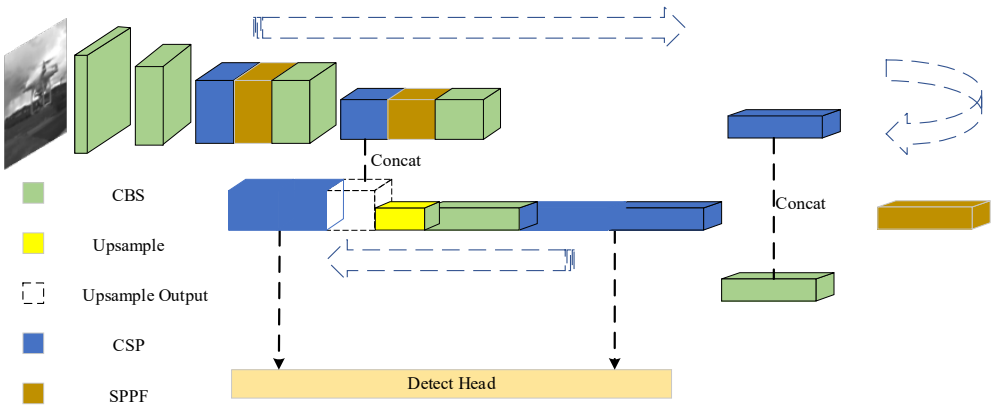


Fig. 1. Yolo-mini network

2.1 Improving the YOLOV5s network structure

The aerial targets are generally acquired using ground equipment, and generally, the target in the image or video is small and belongs to small object detection. The original yolov5 backbone network contains a large number of down-sampling operations, and the feature map size after multiple down-sampling is small, which is not conducive to small object detection. In this paper, firstly, we reduce the number of down-sampling, which makes the computation faster. Secondly, the spatial pyramid pooling-fast (SPPF) module is added to the CSP module to improve the feature expression capability. Fig. 1 gives the network structure of the improved Yolo-mini network, which contains four down-sampling, one up-sampling, and two CONCAT operations, fully fusing features of different sizes and different channel numbers to achieve multiscale feature fusion. The overall network is built by the extended CSP module and CBS (Conv, batch-norming, and silo-activation) module, which does not contain complex operations and is easy to deploy.

As shown in Fig. 1, the network structure of Yolo-mini consists of three stages:

- Backbone (Down-sampling) stage: As shown in the first row of Fig. 1, the main role of this stage is to extract features and down-sample the images to reduce the computational effort. The CBS module represents the standard convolution, normalization, and activation function operations, and is mainly used for down-sampling. SPPF module contains four pooling sizes of 1×1, 5×5, 9×9, and 13×13, which are later fused by standard convolution. This stage expands the channel dimension of the feature map while down-sampling, as shown in Fig. 1. After the down-sampling stage,

the feature map becomes slender from flat, and the feature information is concentrated in the channel dimension, which is easy to use for subsequent prediction.

- Neck (Up-sampling) stage: The main role of this stage is feature fusion, which fuses features of different dimension sizes to improve the expression of the features. As the second row of Fig. 1, after this stage, the feature map becomes flat from slender. This stage mainly uses modules such as the CSP module, CBS module, CONCAT module, and up-sampling module. The feature maps of different stages are stacked by CONCAT operation, and then the stacked feature maps are fused using the CSP module.
- Head stage: This stage is mainly for detection and classification based on the extracted feature maps. This part is consistent with the original yolov5.

2.1.1 CSP extension module

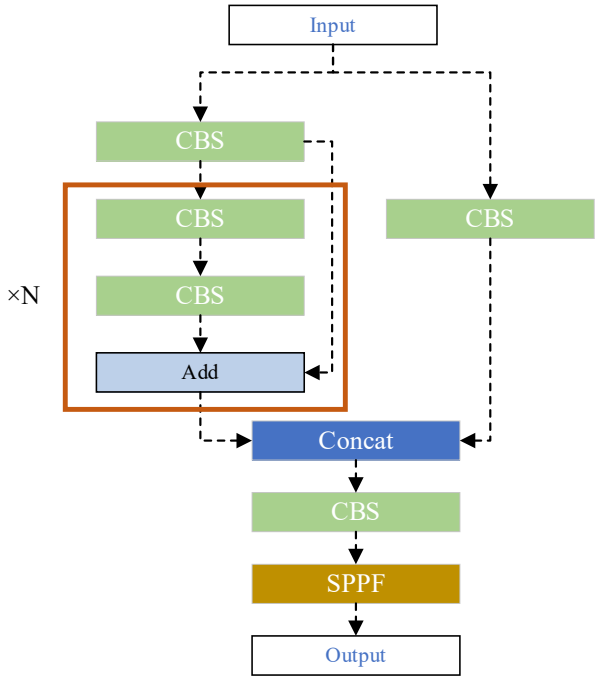


Fig. 2. Extended CSP module. The SPPF module is added after the last convolutional layer.

With the gradual deepening of the network level, the convolutional neural network can extract the semantic information of the high-level features better, but the resolution of the high-level feature maps is lower. In contrast, the resolution of the feature maps is higher at the shallow level, while the semantic information of features extracted from the shallow network is weaker. For an object with fewer and weaker features in the image, deep convolution can lead to difficult extraction or even loss of object features. To maximize the extraction of features that facilitate the detection of a weak object in

the airborne, it is necessary to make full use of the high-resolution features of the convolutional neural network at the shallow layer. Therefore, in the feature extraction stage, we extend the thickness of the CSP module in the shallow feature extraction process. Through stepwise feedback iteration, the object features in the feature map can be fully extracted, and multi-feature extraction from shallow to deep layers can be achieved. Moreover, in deepening the CSP module in the whole feature extraction network by controlling the width and depth factors, we only extend the thickness of the CSP module to extract shallow features. This enhances the ability to extract shallow feature information without increasing the size of the network model and the complexity of the algorithm, which facilitates the detection of a weak object in images. In addition, the CSP structure divides the feature mapping into two branches for extracting features and then merges them, which can achieve a richer combination of gradients while reducing the computational effort.

The structure of the extended CSP module is given in Fig. 2. The red box is the dynamic expansion port, which will repeat N operations. Different CSPs have different N values. In general, a larger N value can increase the network depth and extract better features, but it will increase the computational effort. In the original yolov5, the N values of first these three CSPs in the backbone stage are [1,2,3], which gives less attention to the shallow features and more attention to the deep features. To give more attention to the shallow features, we modify the N values of the first three CSP modules in the backbone to [3,2,1].

The SPPF module in yolov5 is an improved version of the SPP module, which draws on the idea of spatial pyramids and enables the fusion of local and global features through the SPP module, enriching the expressiveness of the feature map and facilitating the detection of large differences in target size in the image to be detected, so it has a great improvement on the accuracy of detection. In the original yolov5 network structure, only an SPPF module is finally employed in the backbone network. To improve the expression of shallow features, as in Fig. 2, we add an SPPF module after each CSP module to enhance feature extraction.

2.1.2 Network Trimming

The original backbone network of the yolov5 network contains a large number of downsampling operations, and the feature map size after multiple downsampling is small, which is not good for small object detection. As in Fig. 1, the improved backbone network is given. Firstly, the 5 downsampling operations are reduced to 4 downsampling operations, which prevents the features from being too small and unfavorable to small object detection. The size of the input image is 640×640, then the minimum feature map of the original yolov5 backbone network is 20×20, and the minimum feature map of the cropped backbone network is 40×40.

The pyramid network structure was also modified. Replacing the original two up-samples operations with only one upsampling operation, only retaining the 40×40 and 80×80 scales of output while adjusting to two anchors and two outputs.

3 Results

3.1 Introduction to the data set

The machine learning model should be trained on a set of annotated images with markers to detect and identify small targets in the air. There are three main sources of the dataset, one source is a public dataset, one source is public images collected online, and one source is real data. Among them, the public dataset is from [18], and the original data format is .mat format, which is converted to the applicable Yolo format after script processing is applied. The dataset covers several types of UAVs, civil aircraft, helicopters, and several species of birds of prey. Factors affecting the object detection results, such as foreground occlusion, smoke, target size, imaging angle of view, and color, were also considered. The dataset was scaled to the video source ratio to ensure the best detection results. The dataset was divided into a training set, test set, and validation set according to the ratio of 6:3:1. Then the sample data in the dataset were labeled using the image labeling software called LabelImg. There was a total of 218,173 images in the dataset, and each image containing at least one target frame.

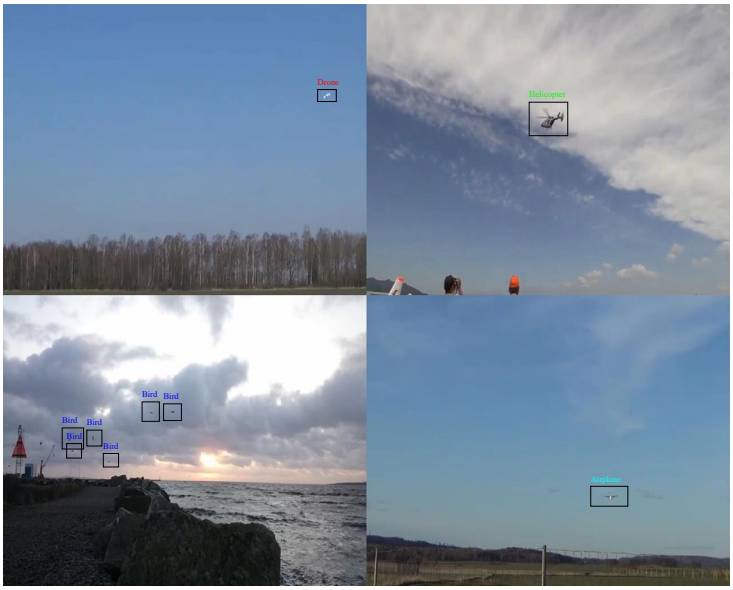
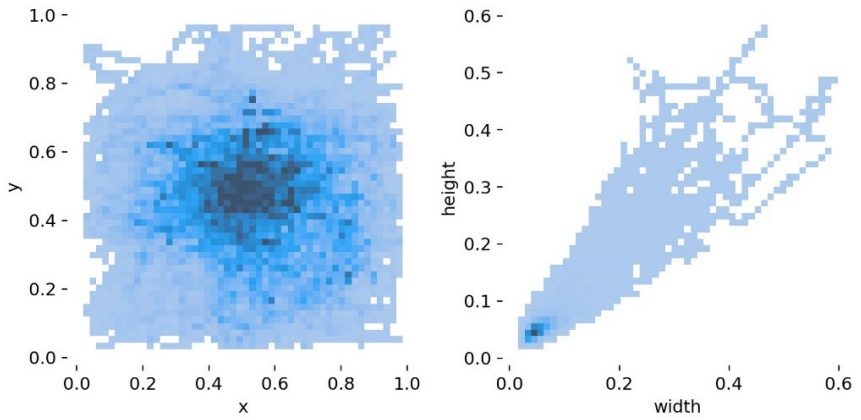


Fig. 3. Sample images from the dataset

A sample dataset is given in Fig. 3. The whole dataset is all aerial targets and the target size is small. Since it is an aerial target, the background is mostly the sky. Some of the targets have a white surface, which is close to the color of the clouds and poses a great challenge to detect. In addition, due to the difference in flight altitude of the four categories of targets: airplane, bird, UAV, and helicopter, the dataset does not contain multiple categories in the same image, in other words, a picture contains only one category of targets.



(a) (b)

Fig. 4. Distribution of the sizes and locations of targets in images of the dataset

In the Yolo series network, the target is described using the four dimensions of the target box, which correspond to [x, y, width, height]. [X, y] is the coordinates of the center point of the target box, and [w, h] is the width and height of the target box. Figure 4 depicts the size and position distribution of the targets in the dataset image, where [x, y, width, height] are all normalized to between [0,1]. Figure 4(a) depicts the position statistics of the centroids of all target boxes in the dataset. It can be seen that the target distribution covers all positions of the image, and most of the targets are concentrated in the middle position of the image. The analysis concludes that the target position distribution in the dataset is close to the normal distribution and is relatively comprehensive. Figure 4(b) depicts the width and height statistics of all target frames in the dataset. It can be seen that the width-to-height ratio of the target boxes is relatively balanced and close to square, and the width-to-height distribution of most of the targets is between [0.0,0.1], indicating that most of the targets in the dataset belong to small targets.

3.2 Experiment Introduction

The computer configuration for the experiments is as follows: 8 GB NVIDIA RTX3090 graphics processing unit (GPU), 16 GB main memory, 1.297 GHz CPU, and SSD hard disk. We used the original yolov5 model weights on the coco dataset (keeping only the uncropped part) to accelerate the training. To run the Yolo-mini process on this GPU, training was performed using cuda11.1 and cudnn8.0. the code was implemented in torch, using torch version 1.7.1. All experiments were trained for 300 epochs, and the first 3 epochs were hot started. We used various data enhancement techniques and set parameters (e.g., rotation, translation, scaling, and other parameters) to enable the model to generate various images from a single image to enrich the given dataset.

The training loss plot is given in Fig. 5, and it can be seen that the loss decreases relatively fast in the first 10 Epochs, after which the trend of train/box_loss and

Val/box_loss also slowly decrease, indicating that the training is effective. Train/obj_loss and Val/obj_ The trend lines of loss almost overlap, which indicates that the model can discriminate the background and target well. Train/cls_loss keeps decreasing, but Val/cls_loss even shows an increasing trend, but the overall level remains low.

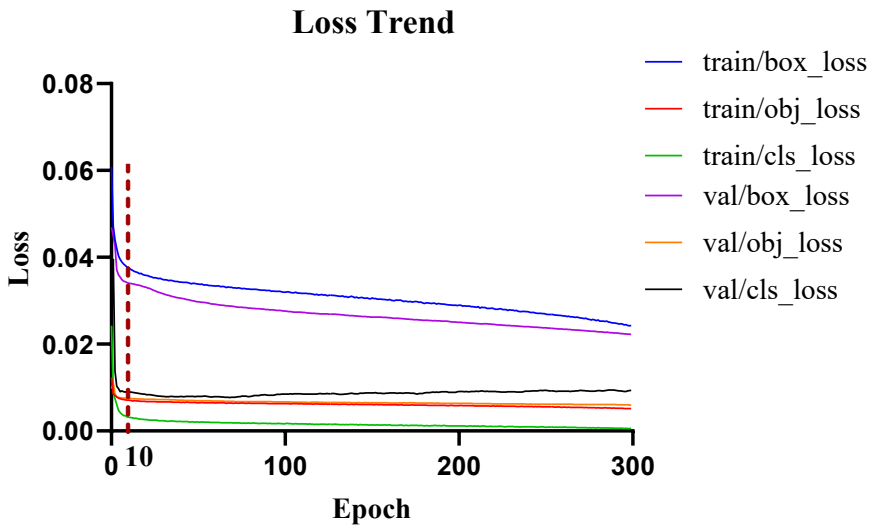


Fig. 5. Training loss trend.

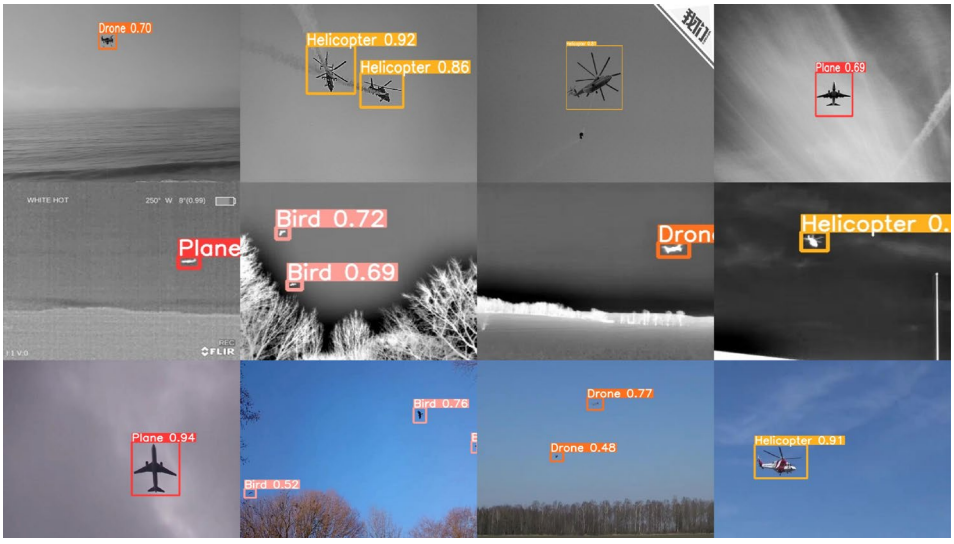


Fig. 6. The detection results on the test dataset

The detection results are given in Fig. 6, with several typical targets selected for display. To facilitate the display, all images are scaled to 480*640 size. As in the second row of Fig. 6, the model can detect well for some very small targets, which shows that our model is effective for small targets. Most of the target frames in Fig. 6 are close to the outer rectangle of the target, which indicates that the regression of our target frames is very good and close to the ideal effect. Fig. 6 contains two types of images, infrared and visible, which indicates that our model can support the detection of both types of images.

3.3 Ablation experiments

To investigate the impact of our improvements on the model, we have separately investigated the N values for the first three CSP modules of the backbone network in Fig. 1. The first three N values are [1,2,3] by default, and the three N values are [3,2,1] after taking the inverse. Plus whether to add the SPPF module after the CSP module, one has four combinations. Withsppf_inver is our improved model with N value taking inverse and CSP module adding SPPF two improvements. Withsppf and Withoutsppf_inver are controlled experiments, which add only N values taking inverse or CSP module adding SPPF. Withoutsppf is the original model without any improvements added.

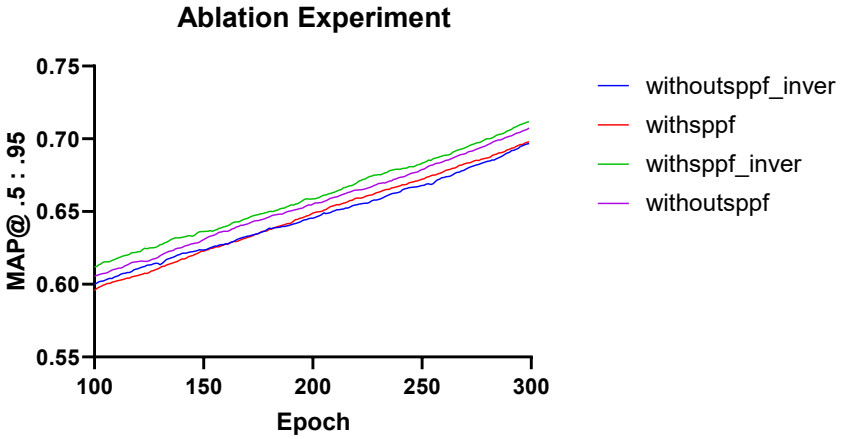


Fig. 7. Ablation experiment results

As in Fig. 7, the results of the four comparison experiments are given. The horizontal coordinate is the number of epochs trained and the vertical coordinate is mAP@.5:.95. To facilitate the presentation of the results, we have taken, the results from 100 to 300 Epochs. From Fig. 7, we can see that Withsppf_inver achieves the best results, which shows that our improvement is effective. In addition, Withsppf and Withoutsppf_inver also work better than Withoutsppf (the original version), respectively, indicating that

individual improvement is also useful for the model. We analyze the reasons for the usefulness of the improvements: since the backbone network of the model is continuously down-sampling, the N value is taken inverse to make the model focus more on shallow features, which improves the detection of small targets. On the other hand, since the backbone network in Fig. 1 is constantly down-sampled, we inverse the N value, which improves the computational effort. The SPPF module can integrate features with different granularity, which improves the feature representation, so it also has a role in map improvement.

3.4 Comparison with classical lightweight object detection models

Table 1. Comparison with other models.

Network	Recall/%	Precision/%	mAp.5/%	mAP_0.5:0.95%	Gflops
Mobilenet[19]	92.98	86.03	89.52	46.93	6.4
Yolov7n[20]	96.35	94.59	94.98	58.25	13.2
Shffule[21]	91.96	81.59	85.12	40.30	1.6
YOLO-mini	95.57	93.06	94.44	71.17	3.2

In the field of object detection, there are many classic lightweight models, and to verify the effectiveness of our proposed method. We have selected three lightweight models for comparison. All network inputs were used with 640*640 inputs and trained with 300 epochs to compare the effect on the test set.

As shown in Table 1, the proposed Yolo-mini achieved the highest mAP_0.5:0.95% by 71.71 compared to the other models. Compared to Mobilenet, all accuracy performance achieved a lead, while the model size is half smaller. Compared to the yolov7n model, the precision is comparable to that of the yolov7n model with a model size of 1/4 of its size. Compared to Shffule, the model is twice as large, but the performance improvement is large enough to make these additional computations worthwhile.

4 Conclusions

To overcome the shortcomings of image detection of small targets, a lightweight detection model Yolo-mini is proposed for small air targets such as UAVs, flying birds, helicopters, and planes. The mAP of the model reaches 94.4%, and the Gflops of the model is only 3.2. In this paper, the network structure is firstly cropped to detect small targets, and only the feature maps with the larger resolution are retained, and then a series of optimizations such as order adjustment and module expansion is carried out for the backbone network. Through a series of comparative experiments, it is found that our model has advantages in the same volume model structure.

References

1. Azar, A.T.; Koubaa, A.; Ali Mohamed, N.; Ibrahim, H.A.; Ibrahim, Z.F.; Kazim, M.; Ammar, A.; Benjdira, B.; Khamis, A.M.; Hameed, I.A.; et al. Drone Deep Reinforcement Learning: A Review. *Electronics* **2021**, *10*, 999.
2. Sandino, J.; Vanegas, F.; Maire, F.; Caccetta, P.; Sanderson, C.; Gonzalez, F. UAV Framework for Autonomous Onboard Navigation and People/Object Detection in Cluttered Indoor Environments. *Remote Sensing* **2020**, *12*, 3386.
3. Udeanu, G.; Dobrescu, A.; Oltean, M. Unmanned Aerial Vehicle in Military Operations. *Sci. Res. Educ. Air Force* **2016**, *18*, 199–206.
4. Pedrozo, S. Swiss Military Drones and the Border Space: A Critical Study of the Surveillance Exercised by Border Guards. *Geographica Helvetica* **2017**, *72*, 97–107.
5. Restas, A.; others Drone Applications for Supporting Disaster Management. *World Journal of Engineering and Technology* **2015**, *3*, 316.
6. Gallacher, D. Drone Applications for Environmental Management in Urban Spaces: A Review. *International Journal of Sustainable Land Use and Urban Planning* **2016**, *3*.
7. Lee, S.; Choi, Y. Reviews of Unmanned Aerial Vehicle (Drone) Technology Trends and Its Applications in the Mining Industry. *Geosystem Engineering* **2016**, *19*, 197–204.
8. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-Cnn: Towards Real-Time Object Detection with Region Proposal Networks. *Advances in neural information processing systems* **2015**, *28*.
9. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y.; Berg, A.C. Ssd: Single Shot Multibox Detector. In Proceedings of the European conference on computer vision; Springer, 2016; pp. 21–37.
10. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition; 2016; pp. 779–788.
11. Liu, Y.; Zhang, X.-Y.; Bian, J.-W.; Zhang, L.; Cheng, M.-M. SAMNet: Stereoscopically Attentive Multi-Scale Network for Lightweight Salient Object Detection. *IEEE Transactions on Image Processing* **2021**, *30*, 3804–3814.
12. Wieczorek, M.; Si\lka, J.; Woźniak, M.; Garg, S.; Hassan, M.M. Lightweight Convolutional Neural Network Model for Human Face Detection in Risk Situations. *IEEE Transactions on Industrial Informatics* **2021**, *18*, 4820–4829.
13. Du, X.; Song, L.; Lv, Y.; Qiu, S. A Lightweight Military Target Detection Algorithm Based on Improved YOLOv5. *Electronics* **2022**, *11*, 3263, doi:10.3390/electronics11203263.
14. Yu, J.; Zhou, G.; Zhou, S.; Qin, M. A Fast and Lightweight Detection Network for Multi-Scale SAR Ship Detection under Complex Backgrounds. *Remote Sensing* **2021**, *14*, 31.
15. Budiharto, W.; Gunawan, A.A.; Suroso, J.S.; Chowanda, A.; Patrik, A.; Utama, G. Fast Object Detection for Quadcopter Drone Using Deep Learning. In Proceedings

of the 2018 3rd international conference on computer and communication systems (ICCCS); IEEE, 2018; pp. 192–195.
16. Yuan, S.; Du, Y.; Liu, M.; Yue, S.; Li, B.; Zhang, H. YOLOv5-Ytiny: A Miniature Aggregate Detection and Classification Model. *Electronics* **2022**, *11*, 1743, doi:10.3390/electronics11111743.
17. Liu, H.; Sun, F.; Gu, J.; Deng, L. SF-YOLOv5: A Lightweight Small Object Detection Algorithm Based on Improved Feature Fusion Mode. *Sensors* **2022**, *22*, 5817.
18. Svanström, F.; Alonso-Fernandez, F.; Englund, C. A Dataset for Multi-Sensor Drone Detection. *Data in Brief* **2021**, *39*, 107521.
19. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv preprint arXiv:1704.04861* **2017**.
20. Wang, C.-Y.; Bochkovskiy, A.; Liao, H.-Y.M. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. *arXiv preprint arXiv:2207.02696* **2022**.
21. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. Shufflenet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition; 2018; pp. 6848–6856.

Exploring the Effects of using Different Audio Lengths in Transfer Learning for Sound Recognition

Jia-Wei Chang, Zhong-Yun Hu* and Jason C. Hung

National Taichung University of Science and Technology, Taichung City, Taiwan
e871223eeee@gmail.com

Abstract. Sound recognition is a challenging task because of the complexity of sound and the excessive noises in environments. This study aims to discuss the influence of different sound lengths on the accuracy of model training. Therefore, this study used LeNet, a simple model with few parameters, and adopted the de-sign of aver-age pooling to enable the proposed models to receive audio of any length. In this study, we verified the preliminary feasibility of transfer learning by LeNet from short-to-long and long-to-short audio, and then further used Res-Net and ResNet-18 instead of LeNet. In experiments, we used the ESC-10 dataset for training models and validated their performance via the self-collected chain-saw-audio da-taset. The results show that (a) the models trained with different audio lengths (1s, 3s, and 5s) have accuracy from 74%~78%, 74%~77%, and 79%~83% on the self-collected dataset. (b) The generalization of the previous models is signifi-cantly improved by transfer learning, the models achieved 85.28%, 88.67%, and 91.8% of accuracy. (c) In transfer learning, the model learned from short-to-long audios can achieve better results than that learned from long-to-short audios, es-pecially being differed 14% of accuracy on 5s chainsaw-audios. (d) the models with lower complexity, i.e. LeNet and ResNet, have higher benefits using transfer learning from short-to-long audios and even perform better than the complex model, RestNet-18.

Keywords: Voice Recognition, Environmental Sound Classification, Chainsaw Sound Recognition, Transfer Learning.

1 Introduction

In recent years, the awareness of environmental protection has gradually gained attention among the general public. In addition to natural disasters such as wildfires and landslides, the issues related to forests are the prevention of human factors. Among the human-made events, the most destructive to forests is illegal logging. It is not enough to protect a forest through monitors or manual patrols. However, through sound monitoring, the cost of deploying protective nets can be reduced, and the occurrence of illegal logging incidents can be responded more immediately. This task belongs to the category of Environmental Sound Classification (ESC). Before performing the task of environmental sound classification, the sound needs to be preprocessed. There are many preprocessing methods, and many different features can be obtained for the model to use. . Zero-crossing rate (Zhang and Kuo, 2001), wavelet features (Valero and Alías, 2012), Mel cepstral coefficient (MFCC) (Uzkent et al., 2012). At present,

machine learning and deep learning have been widely used in environmental sound classification tasks. Support Vector Machine (SVM) (Chu et al., 2009; Piczak, 2015b), Random Forest Classifier (Random Forest Classifier, RF) (Piczak, 2015b), Gaussian Mixture Model (GMM) (Piczak, 2015b; Dhanalakshmi et al., 2011) are all classic machine learning methods. But the initial training of machine learning is time-consuming and costly. Without sufficient data, it is difficult to train a usable model. In recent years, deep learning techniques have been well applied to extract high-discriminative features from sound signals to perform environmental sound classification. Extracting useful features and still maintaining good generalization ability for subtle sounds makes deep learning the preferred method for environmental sound classification. The difference between environmental sound classification and speech recognition tasks is that the sounds to be recognized in environmental sound classification are usually scattered, and the spectrograms converted from the same type of sound may show considerable gaps, and the sound pattern can be continuous. , irregular, instantaneous, and most of them will contain many noisy or silent frames, and the length of the sound we input may also be different. And if the model needs to be applied to small monitoring equipment in the forest, there are large restrictions on the size, complexity and length of the sound of the model, so this paper wants to study in a simple model, so this paper Try to change the number of seconds of audio during training for transfer learning (Zhuang et al., 2020; Liao et al., 2021; Hung and Chang., 2021) to study the sensitivity of the model to the judgment of chainsaw sounds outside the training set Spend. The remaining chapters of this study are organized as follows: Section 2 describes the work related to the original architecture of the environmental sound model used in this study, Section 3 introduces the data set used in this study and explains the methodology used in this study, and Section 4 presents the research The comparison of model training results and the results of the model's ability to judge sounds other than audio in the data set, Chapter 5 discusses the experimental results, and Chapter 6 summarizes the results of this research.

2 Related Work

This section presents related work on ambient sound classification using deep learning-based models.

2.1 Apply spectrogram to CNN

Since two-dimensional features can be obtained after converting the audio into a spectrogram, the spectrogram has always been a favorite preprocessing method for deep learning models of sound. After the advent of CNN (Piczak, 2015a), it was the first time to propose a 2D-CNN that uses spectrogram features as input and executes ESC. According to the research results, compared with machine learning models such as SVM, RF, and GMM, PiczakCNN significantly improves the accuracy of identification. , Inspired by PiczakCNN, more and more people input spectrograms into different CNN models, and some people have combined pre-trained networks to get excellent

results (such as GoogleNet (Szegedy et al., 2015) and AlexNet (Krizhevsky et al., 2012)).

2.2 LeNet-5

The model used in this study is referenced from LeNet-5 (LeCun et al., 2015) and some modifications have been made to meet the requirements of training tasks. In the 1990s, due to the development of algorithms such as SVM, the development of deep learning has been greatly affected. hinder. But LeCun et al. (LeCun et al., 2015) persevered and still worked hard in this field. In 1998, LeCun proposed the LeNet-5 network to solve the problem of handwriting recognition. LeNet-5 is known as the "Hello Word" of convolutional neural networks, which is enough to see the importance of this paper. The model has 7 layers in total, including 3 convolutional layers, 2 average pooling layers, and 2 fully connected layers.

2.3 ResNet

One of the models used in this study uses the Residual Block (He et al., 2016) design. Compared with the current network, the network at that time was very shallow. The reason for this is that the deeper network at that time was easier to fail to train, which made the deeper network sometimes bring worse results. The residual learning proposed by ResNet simply makes deep networks easier to train, and also opens up the era of various ultra-deep networks. Figure 1 shows the network design of the Residual Block. An additional line is used for cross-layer connection, so that the features after the volume base can additionally retain the original features. The overall network can achieve the existing deep features while retaining the original features to avoid losing too much message.

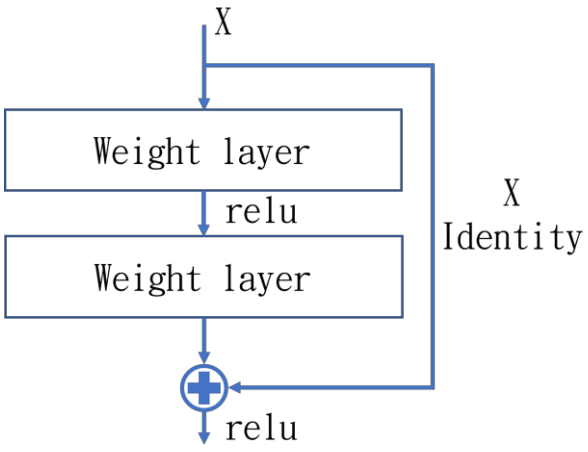


Fig. 1. Residual Block Architecture Diagram

2.4 AdaptiveAvgPool2d

Before proceeding with the scheme that I want to study in this article, we need to find a way to enable the model to input sound information of different dimensions, so AdaptiveAvgPool2d is added to the original model after the convolution layer and before the flattening layer. The concept of this method is similar to global pooling. Layer (Global Average Pooling, GAP) (Lin et al., 2013), so that the model can input sound data of different dimensions. Normal average pooling needs to calculate the window and pace by itself, but AdaptiveAvgPool2d can only input the data dimension you want to output Size, it will automatically calculate the window and pace so that the output format meets the model requirements.

2.5 Transfer Learning

In some fields, the marking of labels is expensive, resulting in insufficient training data, and it is easy to cause the trained model to overfit, that is, the generalization ability of the data outside the training data is insufficient, resulting in the model having no practical value. Migration learning There are two commonly used methods, feature extraction and fine-tuning. Feature extraction refers to using the pre-trained model as the part of data feature extraction to extract useful features for the target task. Fine-tuning technology is to use the model trained for the original task and The parameters are applied to the target training task, so that the target training task can have a better initial gradient position for training, which can achieve faster convergence and increase accuracy. Migration learning has achieved good results in past research, so this study The proposed model training method is based on fine-tuning technology, and it is studied whether it is possible to improve the accuracy by changing the length of the input audio for transfer training without changing the complexity of the model.

3 Method

3.1 DataSet

- ESC-10 Dataset(Piczak et al., 2015b):

The ESC-10 data set is a subset of the ESC-50 data set, which contains 400 labeled collections of indoor and outdoor environmental recordings. It is suitable for benchmarking methods for environmental sound classification. The audio in this data set is composed of 5-second long records , the sampling rate is 44100Hz, and is classified into 10 categories on average, one of which is chainsaw sound, and each category has 40 audios. The labels in this data set have been pre-arranged 5-fold for cross-validation to ensure the same Fragments of the original source file are contained in the same fold.

- Chainsaw Dataset:

The sound clips collected in this study including the sound of chainsaws do not contain any audio data from ESC-10 and ESC-50. The audio data set is composed of 5-second

long records with a sampling rate of 44100Hz. Special mention The problem is that these clips are not all clean chainsaw sounds, to simulate the noise that would be present in real life when judgment is required.

3.2 Architecture for training models in different seconds

There are four sections in this chapter. The first section explains the data preprocessing, the second section introduces the details of the model implementation and the parameter setting of the experiment, the third section explains the experimental design, and the fourth section explains the experimental environment and hyperparameter settings.

3.2.1 data preprocessing

Spectrogram features are extracted from a given acoustic signal. The sampling rate is 44100Hz, the frame shift is set to 512, the window length is 2048, the number of filters is 128, the highest frequency is 22050, and the lowest frequency is 20. In this study, the Libroas library in Python (McFee et al., 2015) was used to extract the spectral signal. Since this study wants to train the model and perform migration training through different seconds, it wants to extract three different types of short, medium and long The length is used to make a difference, so I chose 1 second, 3 seconds, and 5 seconds. The converted feature sizes are (128,87,1), (128,259,1), (128,431,1) respectively. Since the sounds are all 5-second segments, the data will increase by 5 times and 3 times when extracting 1-second and 3-second sounds, respectively. Figure 2 shows the method of sound segment extraction. In this experiment, two model training methods are designed: (1) the normal ESC-10 label, the label is classified from 0 to 9, a total of 10 labels, and (2) all the sound labels other than the chainsaw sound are set to 0 , and the chainsaw sound label is set to 1 for binary classification.

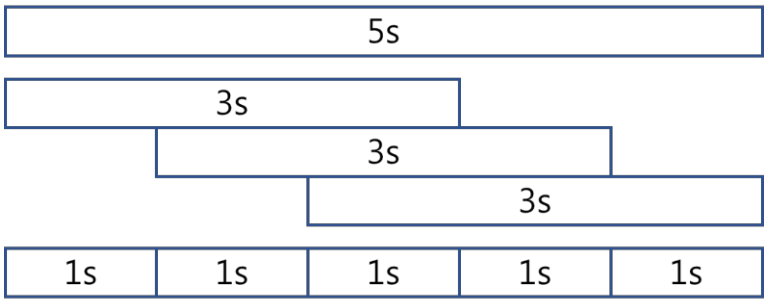


Fig. 2. Schematic diagram of extraction of different sound lengths

3.2.2 Model implementation details and experimental parameter settings.

Figures 3, 4, and 5 show the model architecture used in this study.
The relevant parameters of the model in Figure 3 are as follows.

- A1: The input size is (1, 128, W), and W is the width converted into a spectrogram for 1 second, 3 seconds, and 5 seconds.
- A2: It is a 2D-CNN convolutional layer, the number of input channels is 1, the number of output channels is 16, the kernel_size is 5, the stride is 1, and the padding is 0.
- A3: It is a maximum pooling layer, the kernel_size is 2, the stride is 2, and the padding is 0.
- A4: It is a 2D-CNN convolutional layer, the number of input channels is 16, the number of output channels is 32, the kernel_size is 5, the stride is 1, and the padding is 0.
- A5: It is a maximum pooling layer, the kernel_size is 2, the stride is 2, and the padding is 0.
- A6: It is a binary adaptive average pooling layer, which performs adaptive average pooling according to the direction of the channel, and the output is one with an array length of 32.
- A7: It is a fully connected layer with 120 nodes.
- A8: It is a fully connected layer with 84 nodes.
- A9: It is a fully connected layer, and the number of nodes is 10 or 2, depending on the experimental project.

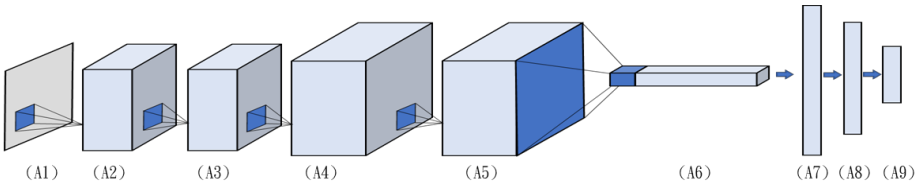


Fig. 3. Model architecture diagram of LeNet combined with Global Average Pooling

The relevant parameters of the model in Figure 4 are as follows.

- A1: The input size is (1, 128, W), and W is the width after converting 1 second, 3 seconds, and 5 seconds into a spectrogram.
- A2: It is a 2D-CNN convolutional layer, the number of input channels is 1, the number of output channels is 64, the kernel_size is 7, the stride is 2, and the padding is 3.
- A3: It is a maximum pooling layer, kernel_size is 2, stride is 2, padding is 1, followed by a BatchNorm2d layer and a Relu layer.
- A4: It is a 2D-CNN convolutional layer, the number of input channels is 64, the number of output channels is 128, the kernel_size is 5, the stride is 1, and the padding is 0.
- A5: It is a 2D-CNN convolutional layer, the number of input channels is 128, the number of output channels is 128, the kernel_size is 5, the stride is 1, and the padding is 0.

- A6: It is a binary adaptive average pooling layer, which performs adaptive average pooling according to the direction of the channel, and the output is one with an array length of 128.
- A7: It is a fully connected layer with 64 nodes.
- A8: It is a fully connected layer with 10 nodes.

A4 and A5 are followed by a maximum pooling layer with kernel_size of 2, stride of 2, padding of 0 and a Relu layer.

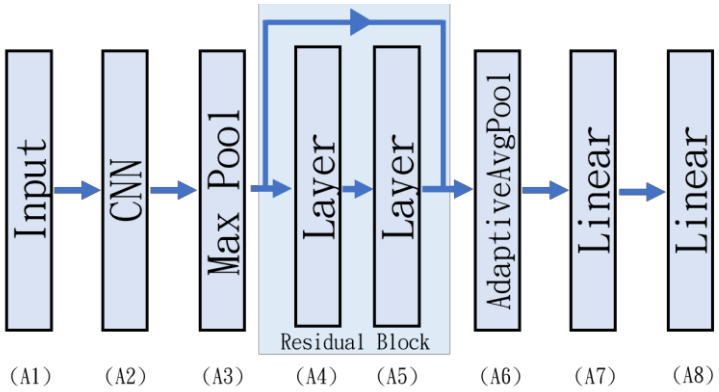


Fig. 4. Model architecture diagram of single-layer Residual Block combined with Global Average Pooling

The relevant parameters of the model in Figure 5 are as follows.

- A1: The input size is (1, 128, W), and W is the width after converting 1 second, 3 seconds, and 5 seconds into a spectrogram.
- A2: It is a 2D-CNN convolutional layer, the number of input channels is 1, the number of output channels is 64, the kernel_size is 7, the stride is 2, and the padding is 3.
- A3: It is a maximum pooling layer, kernel_size is 2, stride is 2, padding is 1, followed by a BatchNorm2d layer and a Relu layer.
- A4: It is a 2D-CNN convolution layer, the number of input channels is 64, the number of output channels is 64, kernel_size is 5, stride is 1, padding is 0, followed by a maximum pooling layer, kernel_size is 2, stride is 2, The padding is 0 and a Relu layer.
- A5: It is a 2D-CNN convolutional layer, the number of input channels is 64, the number of output channels is 64, the kernel_size is 5, the stride is 1, and the padding is 0.
- A6: It is a 2D-CNN convolution layer, the number of input channels is 64, the number of output channels is 128, the kernel_size is 5, the stride is 1, and the padding is 0.

- A7: It is a 2D-CNN convolution layer, the number of input channels is 128, the number of output channels is 128, the kernel_size is 5, the stride is 1, and the padding is 0.
- A8: It is a 2D-CNN convolutional layer, the number of input channels is 128, the number of output channels is 256, the kernel_size is 5, the stride is 1, and the padding is 0.
- A9: It is a 2D-CNN convolutional layer, the number of input channels is 256, the number of output channels is 256, the kernel_size is 5, the stride is 1, and the padding is 0.
- A10: It is a 2D-CNN convolutional layer, the number of input channels is 256, the number of output channels is 512, the kernel_size is 5, the stride is 1, and the padding is 0.
- A11: It is a 2D-CNN convolutional layer, the number of input channels is 512, the number of output channels is 512, the kernel_size is 5, the stride is 1, and the padding is 0.
- A12: It is a binary adaptive average pooling layer, which performs an adaptive average pooling layer according to the direction of the channel, and the output is one with an array length of 512.
- A13: It is a fully connected layer with 64 nodes.
- A14: It is a fully connected layer with 10 nodes.

A4 to A11 are followed by a maximum pooling layer with a kernel_size of 2, a stride of 2, padding of 0 and a Relu layer.

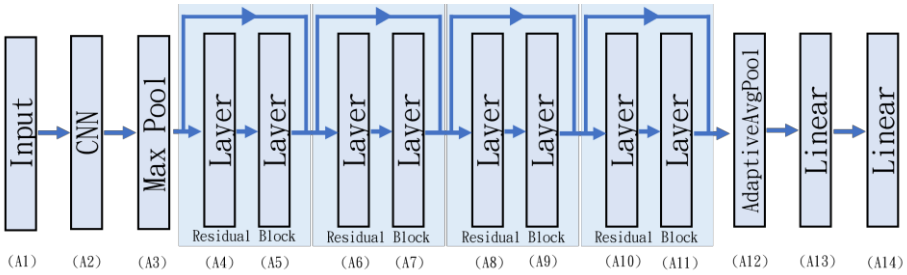


Fig. 5. Model architecture diagram of ResNet-18 combined with Global Average Pooling

3.2.3 Experimental design

In this study, the ESC-10 data sets of different lengths (1 second, 3 seconds, 5 seconds) were used to train the model, and the experiment was divided into three stages sequentially. In the first stage, we used two different label methods to train separately Figure 3 model, and uniformly used the 1-second, 3-second and 5-second sound files of the chainsaw sound data set collected by this research to verify the recognition ability of the chainsaw sound. In the second stage, we will perform migration training on the model in Figure 3 with the label method that performed better in the first stage, and compare it with the model that has not undergone migration training second, 3 second

and 5 second sound files to verify the recognition ability of the chainsaw sound. In the third stage, we will use the models in Figure 4 and Figure 5 to train in 1 second, 3 seconds, and 5 seconds, respectively, and transfer learning for training and comparison. The comparison benchmark uses the ESC-10 test set, and the three-stage model training is unified. Use the ESC-10 data set to perform 5-Fold cross-validation.

3.2.4 Experimental environment and hyperparameter settings

All models were developed and run on an NVIDIA GeForce RTX3060 6G GPU with 8GB RAM and the proposed experimental method was developed using the open source Pytorch 1.12 library running Python on the Windows 10 operating system. The batch size is 64, the Adam optimizer (Kingma and Ba, 2014) is used for optimization, the learning rate is 0.0002, and the learning rate is reduced to half of the original every 10 times of training, and the loss function method is CrossEntropy Loss (Zhang and Sabuncu, 2018), with a total of 30 training times. The part of transfer training is to repeatedly train the model with audio data of different lengths with the above parameters.

4 Experimental results

There are three stages of experimental results in this chapter. The first stage is the result of the normal training of the model in Figure 3 using two different label data, and the second stage is the result of the transfer training of the model in Figure 3 with the label method that performed better in the first stage , the third stage is the result of the normal training of the models in Figure 4 and Figure 5 and the transfer training.

4.1 First stage

Figure 6 and Figure 7 are the results of the first stage of the experiment. The bars in the figure represent the average accuracy of 5-fold verification. The high point of the error bar on the long bar is the highest accuracy in 5-fold, and the low point is 5-fold lowest accuracy.

4.1.1 Accuracy of model trained on binary classification

Figure 6 shows the judgment accuracy of the model using binary classification as the final result after feeding different number of seconds. There are 3 models, which are trained in 1 second, 3 seconds, and 5 seconds respectively, and are fed in 1 second, 3 seconds, and 5 seconds for prediction. The three models have 53.52%, 60.04%, and 53.68% accuracy for the 1-second test audio, which are higher than 49.47%, 55.4%, 44.87% for the 3-second test audio, and 44% and 50% for the 5-second test audio , 39.8%. As a result, the binary classification model has the highest accuracy in predicting the sound of chainsaws in 1 second.

10

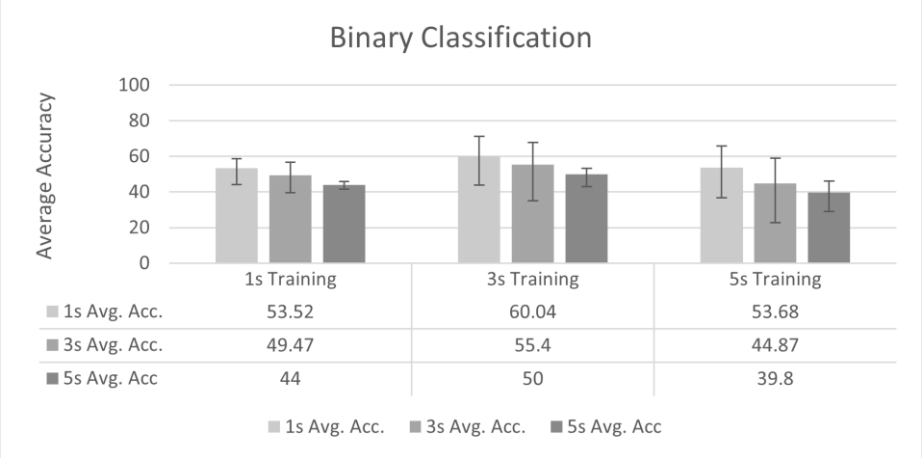


Fig. 6. Accuracy histogram after training with binary classification, the model was trained separately for 1 second, 3 seconds, and 5 seconds, and the accuracy was tested using the test audio for 1 second, 3 seconds, and 5 seconds.

4.1.2 The accuracy of the model after training according to the ESC-10 classification

Figure 7 shows the judgment accuracy of the model using ESC-10 classification as the final result after training with different numbers of seconds. There are 3 models, which are trained in 1 second, 3 seconds, and 5 seconds respectively, and are fed into 1 second, 3 seconds, and 5 seconds for prediction. The three models have 74.16%, 74.16%, and 79.32% accuracy for the 1-second test audio, 78.2%, 76.13%, and 83.53% accuracy for the 3-second test audio, and 78.8%, 77.6% for the 5-second test audio. %, 83.2% accuracy, the result is that no matter how many seconds the model is trained, the accuracy of the model's prediction for the test data of different seconds is similar, but it can be seen that the accuracy of the model trained for 5 seconds is higher. For the test data, the prediction accuracy of feeding longer seconds is generally higher, and all the prediction accuracy ratios are much higher than using the binary prediction method.

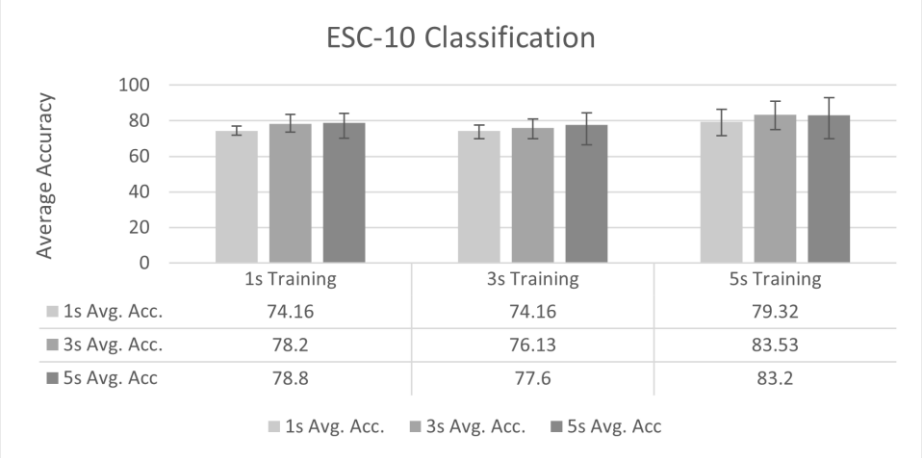


Fig. 7. Accuracy histogram after training using ESC-10 classification, the model is trained separately for 1 second, 3 seconds and 5 seconds, and the accuracy is tested using the test audio of 1 second, 3 seconds and 5 seconds.

4.2 Second stage

Figure 8 shows the results of the second stage of the experiment. The bars in the figure represent the average accuracy of 5-fold verification. The high point of the error bar on the bar is the highest accuracy in 5-fold, and the low point is the lowest accuracy in 5-fold. Spend. Figure 9 shows the error range of the model training in Figure 7 and Figure 8.

4.2.1 The accuracy of the model after using transfer training

Figure 8 shows the accuracy of the model after migration training. The model is trained in (1 second → 3 seconds → 5 seconds) and (5 seconds → 3 seconds → 1 second) seconds, and both are fed for 1 second , 3 seconds, 5 seconds for prediction. The two models have 85.28% and 74.52% accuracy for the 1-second test audio, 88.67% and 77.06% accuracy for the 3-second test audio, and 91.8% and 77.8% accuracy for the 5-second test audio. Compared with the model with the best performance in Figure 7, comparing the test data of the model trained with 5 seconds, it can be found that the model trained with the number of seconds from small to large (1 second → 3 seconds → 5 seconds) is at 5 seconds The accuracy of the second-second audio increased by 8.6%, the 3-second audio increased by 5.14%, and the 1-second audio increased by 5.96%. The model trained from large to small (5 seconds → 3 seconds → 1 second) showed poor results. And there is no improvement in the accuracy at each second compared to any of the models in Figure 4.

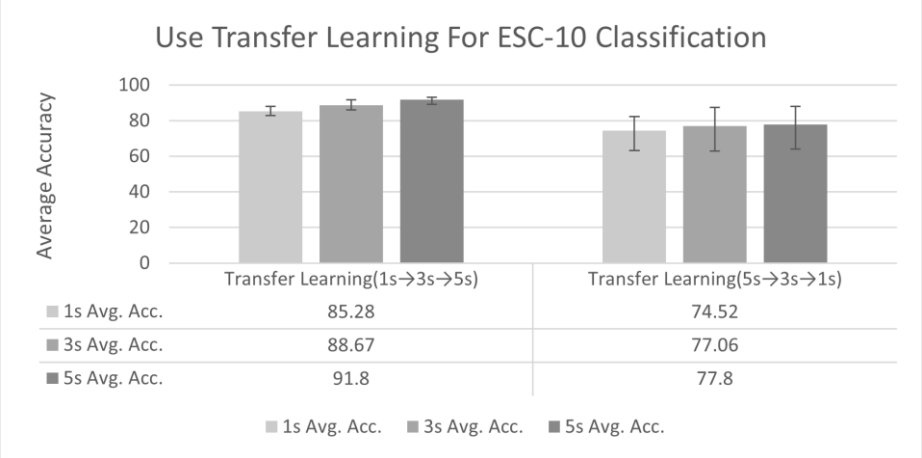


Fig. 8. The accuracy histogram of ESC-10 classification using the model after transfer learning. The model performs transfer learning in the order of (1 second → 3 seconds → 5 seconds) and (5 seconds → 3 seconds → 1 second) , and use 1 second, 3 second and 5 second test audio for accuracy test.

4.2.2 The prediction range of K-Fold after the model uses migration training

Figure 9 shows the full range of predicted values for K-Fold after the model uses migration training. The calculation method is to subtract the value with the highest predicted accuracy of K-Fold from the value with the lowest predicted accuracy. The first three models are models that have not undergone transfer learning and only use a single number of seconds (1 second, 3 seconds, 5 seconds) for training. The 1-second prediction accuracy values are 5%, 7.66%, 15%, and 3 The second prediction accuracy ranges are 9.8%, 11%, and 16%, respectively, and the 5-second prediction accuracy ranges are 13.8%, 18%, and 23%, respectively. The latter is the accuracy of the two models after using migration training. The model Respectively (1 second → 3 seconds → 5 seconds) for training, and (5 seconds → 3 seconds → 1 second) seconds, the prediction accuracy range of 1 second is 5.2%, 19.2%, respectively, and the prediction accuracy of 3 seconds is full The distances are 5.67% and 24.32% respectively, and the full range of 5-second prediction accuracy is 4% and 2.4%. From the training results, it can be seen that (1 second → 3 seconds → 5 seconds) after migration training can be effectively Reduce the full distance of the predicted value, and from the accuracy of Figure 4 and Figure 5, it can be seen that the accuracy can be greatly improved while reducing the full distance. For the 5-second test data, the accuracy is increased by 8.6% and reduced. Up to 19% range, 5.14% accuracy improvement and 12.33% range reduction for 3-second test data, 5.96% accuracy improvement and 8.6% range reduction for 1-second test data , studies have shown that the migration-trained model increases the performance of extracting audio features, increases the generalization of the model, and improves the accuracy of model judgment.

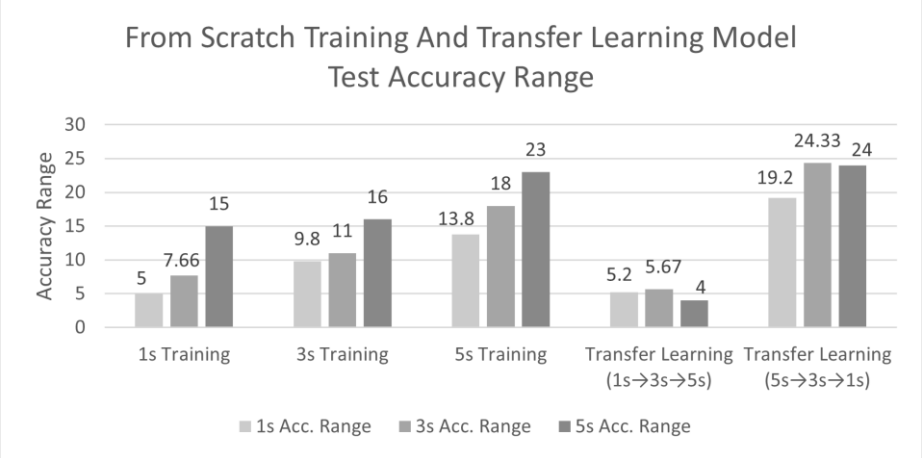


Fig. 9. The range histogram of the prediction accuracy of the ESC-10 classification K-Fold using the From Scratch and transfer learning training models respectively. The models are trained separately in 1 second, 3 seconds and 5 seconds and (1 second → 3 seconds → 5 seconds), (5 seconds → 3 seconds → 1 second) for transfer learning, and use 1 second, 3 seconds, and 5 seconds of test audio to perform prediction value range calculation. For K-Fold's prediction accuracy range, the range calculation method is to subtract the value of the lowest prediction accuracy from the value of the highest prediction accuracy in K-Fold.

4.3 Third stage

Figure 10 is the training result of the model in Figure 4, and Figure 11 is the training result of the model in Figure 5. The bars represent the average accuracy of 5-fold verification. The high point of the error line on the bar in the figure is the highest accuracy in 5-fold. The low point is the lowest accuracy in 5-fold.

4.3.1 The results of training using the ResNet model

Figure 10 shows that the model in Figure 4 is trained separately in 1 second, 3 seconds, and 5 seconds, and transfer learning is performed in the order of (1 second → 3 seconds → 5 seconds), (5 seconds → 3 seconds → 1 second), 1 Second, 3 second, and 5 second ESC-10 test sets for accuracy testing. The five models have 77.3%, 7a2.75%, 58.65%, 77.65%, and 78.15% accuracy for the 1-second test audio, and 83.25%, 82.58%, 70.17%, 87.5%, and 84.92% accuracy for the 3-second test audio The five-second test audio has an accuracy of 82.25%, 85%, 79.25%, 90.75%, and 82%, respectively. In the separate training part, it can be seen that the accuracy of the 5-second training is poor. The 1-second and 3-second training have their own advantages and disadvantages in the test audio of different seconds. In the transfer training part, it can be found that the number of seconds is increased from small to large. (1 second → 3 seconds → 5 seconds) the model trained from large to small (5 seconds → 3 seconds → 1 second) showed better results than the model trained from small to large (1 second → 3 seconds

→ 5 seconds) compared with the model test results of other training methods, except that the test accuracy of 1 second is slightly lost by 0.5% compared with the highest accuracy, and the test accuracy of 3 seconds is the second highest compared with the second highest. was 2.58% more accurate for, and the five-second test was 5.75% more accurate than the next best.

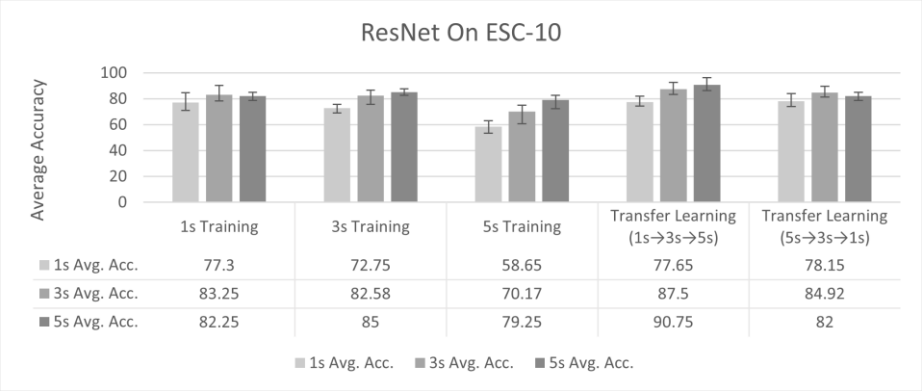


Fig. 10. The prediction accuracy histogram of the ESC-10 classification K-Fold using From Scratch and transfer learning training the model in Figure 4, the model is trained separately in 1 second, 3 seconds and 5 seconds and (1 second → 3 seconds → 5 seconds), (5 seconds → 3 seconds → 1 second) for transfer learning, and the ESC-10 test sets for 1 second, 3 seconds, and 5 seconds for accuracy testing.

4.3.2 The results of training using the ResNet-18 model

Figure 11 shows that the model in Figure 4 is trained separately in 1 second, 3 seconds, and 5 seconds, and transfer learning is performed in the order of (1 second → 3 seconds → 5 seconds), (5 seconds → 3 seconds → 1 second), 1 Second, 3 second, and 5 second ESC-10 test sets for accuracy testing. The five models have 75.45%, 75.1%, 66.25%, 71.85%, and 76.25% accuracy for the 1-second test audio, and 79.58%, 85.42%, 80.25%, 82.58%, and 83.83% accuracy for the 3-second test audio The five-second test audio has an accuracy of 80.25%, 87.75%, 86.75%, 87.5%, and 84.5%, respectively. In the part of separate training, it can be seen that the accuracy of 5-second training and 1-second test is poor, and the rest have their own advantages and disadvantages in the test audio of different seconds. In the part of transfer training, it can be seen that the accuracy of the model has not been improved or even It also reduces accuracy compared to training on seconds alone.

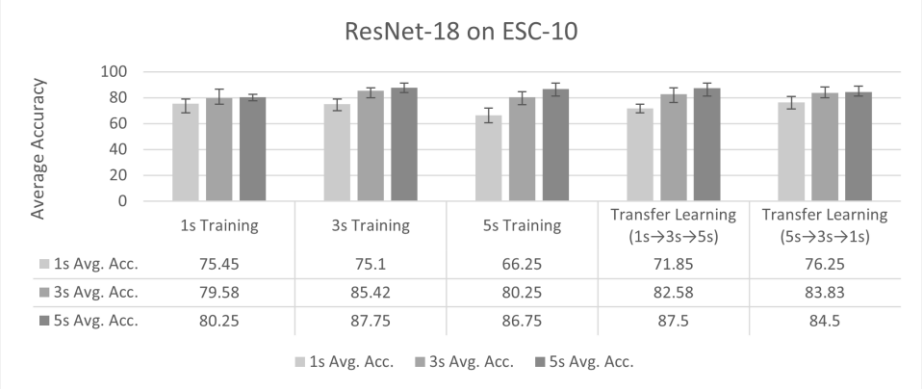


Fig. 11. The prediction accuracy histogram of the ESC-10 classification K-Fold using From Scratch and transfer learning training the model in Figure 4, the model is trained separately in 1 second, 3 seconds and 5 seconds and (1 second → 3 seconds → 5 seconds), (5 seconds → 3 seconds → 1 second) for transfer learning, and the ESC-10 test sets for 1 second, 3 seconds, and 5 seconds for accuracy testing.

5 Discussions

From the experimental results, it can be seen that the model trained by binary classification is not effective. It is speculated that the data distribution during training is too skewed because the data fed in is not the sound of the chainsaw and the audio ratio of the sound of the chainsaw is 9. Compared with 1, the generalization ability of the model is reduced. The model trained with the normal ESC-10 label has considerable accuracy even if the audio is different from the number of seconds used for training. It is interesting to see that the model is given a longer number of seconds No matter how many seconds the model takes to train, the accuracy of prediction is generally high. However, it can be seen that the accuracy of the two training methods is quite different in 5-fold. Good models classified by ESC-10 have the most difference in highest and lowest accuracy when inputting longer seconds to make predictions. In the model using migration training, the training is performed in increments of seconds, and it can be seen that its judgment accuracy is higher in 5 seconds, and it is also improved in 3 seconds and 1 second. It can be seen that all judgments can be seen in training in descending seconds. The accuracy has not improved and the error has become larger. The 1-second prediction accuracy ranges are 5.2% and 19.2%, the 3-second prediction accuracy ranges are 5.67%, 24.32%, and the 5-second prediction accuracy ranges are respectively 4%, 2.4%, so the research shows that if you want to use different lengths of audio for migration training, you can get better results from short audio training to long audio, but if you train from long audio to short audio, it will not improve. The accuracy of the judgment and the error of the model are improved. The difference between the highest accuracy and the lowest accuracy is significantly smaller from short audio training to long audio training. Therefore, correct transfer learning is indeed beneficial to the recognition of chainsaw sounds. After using migration training for different models, it

can be found that for the three models used in this research, one feature can be found, that is, the higher the complexity of the model, the lower the benefit can be obtained after using migration training, and even the complexity can be seen. The accuracy of the lower model after migration training is higher than the accuracy of the model with high complexity. Here are three possible explanations to be verified: (a) There is a marginal benefit in the volume base of the sound spectrogram. Overly complex models lead to poor accuracy. (b) Migration training methods of different lengths can improve the simple model better, so it can be used as a method to improve the accuracy of simple models. (c) Whether it is due to insufficient training data or too simple classification tasks that lead to overfitting of complex models.

6 Conclusion

This study raises a question to make the model accept the audio that is different from the length of the training audio for prediction and explore the help of transfer learning for this model. The comparison of different training seconds for the generalization ability of the model, the experiment compares two different The labeling method and the further use of transfer learning for training on the better-performing labeling method, the results prove that the model can effectively generalize and have good accuracy for audio that is different from the length of the training audio. In transfer learning On the one hand, it can be seen that this training method can effectively improve the generalization ability and accuracy of the model. Two different models were additionally tested and compared with the data to verify the performance of transfer learning and to propose the extension direction of this research.

References

1. Chu, S., Narayanan, S., & Kuo, C. C. J. (2009). Environmental sound recognition with time–frequency audio features. IEEE Transactions on Audio, Speech, and Language Processing, 17(6), 1142-1158.
2. Dhanalakshmi, P., Palanivel, S., & Ramalingam, V. (2011). Classification of audio signals using AANN and GMM. Applied soft computing, 11(1), 716-723.
3. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
4. Hung, J. C., & Chang, J. W. (2021). Multi-level transfer learning for improving the performance of deep neural networks: theory and practice from the tasks of facial emotion recognition and named entity recognition. Applied Soft Computing, 109, 107491.
5. Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
6. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25.
7. LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.

8. Liao, J. Y., Lin, Y. H., Lin, K. C., & Chang, J. W. (2021, December). 以遷移學習改善深度神經網路模型於中文歌詞情緒辨識 (Using Transfer Learning to Improve Deep Neural Networks for Lyrics Emotion Recognition in Chinese). In International Journal of Computational Linguistics & Chinese Language Processing, Volume 26, Number 2, December 2021.
9. Lin, M., Chen, Q., & Yan, S. (2013). Network in network. arXiv preprint arXiv:1312.4400.
10. McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E., & Nieto, O. (2015, July). librosa: Audio and music signal analysis in python. In Proceedings of the 14th python in science conference (Vol. 8, pp. 18-25).
11. Piczak, K. J. (2015a, September). Environmental sound classification with convolutional neural networks. In 2015 IEEE 25th international workshop on machine learning for signal processing (MLSP) (pp. 1-6). IEEE.
12. Piczak, K. J. (2015b, October). ESC: Dataset for environmental sound classification. In Proceedings of the 23rd ACM international conference on Multimedia (pp. 1015-1018).
13. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).
14. Uzkent, B., Barkana, B. D., & Cevikalp, H. (2012). Non-speech environmental sound classification using SVMs with a new set of features. International Journal of Innovative Computing, Information and Control, 8(5), 3511-3524.
15. Valero, X., & Alías, F. (2012, August). Gammatone wavelet features for sound classification in surveillance applications. In 2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO) (pp. 1658-1662). IEEE.
16. Zhang, T., & Kuo, C. C. J. (2001). Audio content analysis for online audiovisual data segmentation and classification. IEEE Transactions on speech and audio processing, 9(4), 441-457.
17. Zhang, Z., & Sabuncu, M. (2018). Generalized cross entropy loss for training deep neural networks with noisy labels. Advances in neural information processing systems, 31.
18. Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., ... & He, Q. (2020). A comprehensive survey on transfer learning. Proceedings of the IEEE, 109(1), 43-76.

Mobile Robot Controller Design Using Deep Learning

Jyun-Yu Jhang [1,2]

[1] Department of Computer Science and Information Engineering, National Taichung University of Science and Technology, Taichung 404, Taiwan
[2] College of Intelligence, National Taichung University of Science and Technology, Taichung 404, Taiwan

Abstract. This paper presents a real-time navigation control system based on lidar sensing in unknown environments. The input is the distance between the robot and the wall from the lidar sensor, and the output is the steering angle of the robot, so as to navigate to the destination without collision in various unknown environments. The experimental results show that the navigation control system developed in the simulated and actual environments can effectively assist the Ackerman robot to complete the navigation task in unknown environments.

Keywords: Deep Learning Controller, Navigation, Mobile Robot.

1 Introduction

Due to labor shortages and increased labor costs, many factories are transitioning to fully automated mechanization, and one of the key technologies is autonomous mobile robots. Ideally, autonomous mobile robots can move or work in real-world environments without human operation, but unknown environments and uncertain dynamic obstacles make this work difficult [1]. If the problem of unknown environment and dynamic obstacles can be solved, the technology of autonomous mobile robot can also be applied to self-driving cars [2] or handling large objects [3-4]. The navigation control of autonomous mobile robots can be divided into two technologies, including goal finding and obstacle avoidance, where these technologies are built on the robot with a robust controller. In this paper, a navigation control method is designed for the autonomous Ackerman robot in unknown environments. The proposed system consists of a behavioral controller to control the Ackerman robot for obstacle avoidance or heading toward the goal without global map information.

2 Methods

Fig. 1 presents a training environment measuring 11 m × 8 m. To allow mobile robots to encounter different environments, the training environment includes straight lines, corners, and right-angled corners.

2

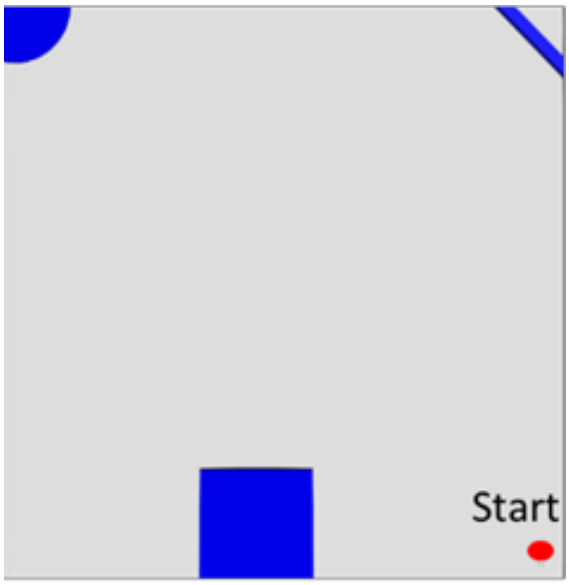


Fig. 1. Mobile robot training environment

If an obstacle is detected in front of the autonomous Ackerman robot, the behavior controller will switch to wall-following mode to assist the robot to walk along the object until the robot leaves the obstacle. To achieve this behavior, a fuzzy controller with wall following function, namely Wall-Following Fuzzy Controller (WFFC), is designed. Fig. 2 present the system flow of wall-following mode. First, the lidar sensor detects the distance to obstacles around the robot. Then the distance information is used as the input of the controller and the output is the steering angle of the robot.

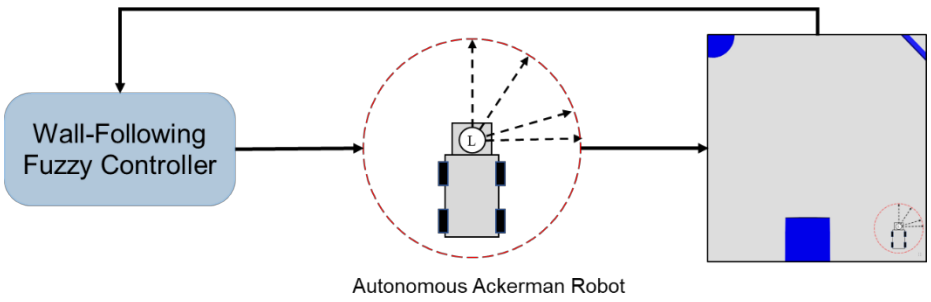


Fig. 2. wall-following mode

3 Experimental Result

The autonomous Ackerman robot used in this paper is an independent research and development. The robot uses a Velodyne Puck (VLP-16) lidar sensor to scan for surrounding obstacles and an edge-embedded device NVIDIA Jetson AGX Xavier (AGX) for real-time data processing. The sensing distance range of VLP-16 is 50cm~5m, and its horizontal measurement angle is 360°. AGX uses ubuntu 16.04 and Robot Operating System (ROS) to drive the robot's motor system. Through control commands, control the movement speed and turning angle of the robot. In addition, the robot chassis structure adopts Ackerman architecture to move more smoothly when handling heavy objects and rough terrain.

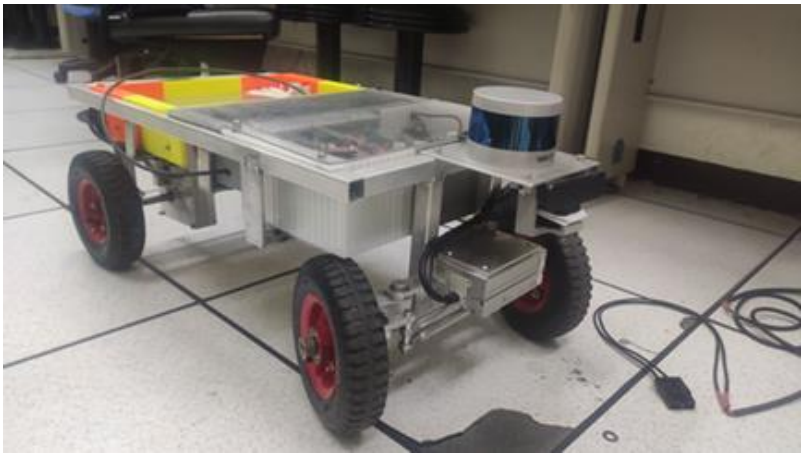


Fig. 3. Ackerman robot

To verify the performance of the proposed navigation control method, the testing environments were shown in Fig. 4. Fig. 4. show a simple concave environment, four clasp obstacles respectively.

4

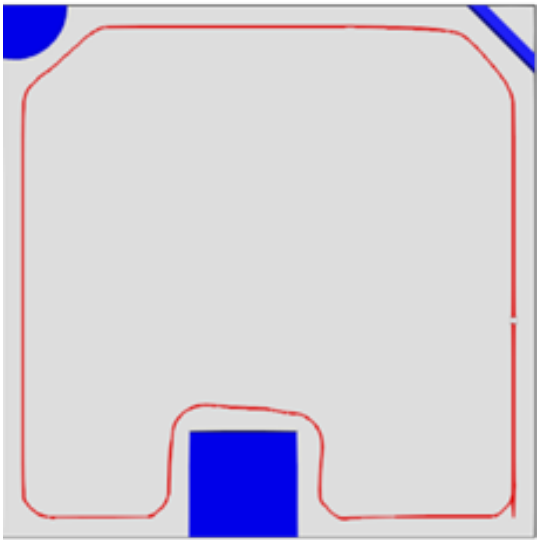


Fig. 4. The move paths of a mobile robot

4 Conclusions

This study proposes an effective navigation control method in an unknown environment. Experimental results reveal that the performance of the proposed method in terms of wall following and navigation control is more efficient than that of other methods in unknown environments.

References

1. X. Kan, H. Teng and K. Karydis, "Online Exploration and Coverage Planning in Unknown Obstacle-Cluttered Environments," in IEEE Robotics and Automation Letters, vol. 5, no. 4, pp. 5969-5976, Oct. 2020.
2. H. B. Said, R. Marie, J. Stéphant and O. Labbani-Igbida, "Skeleton-Based Visual Servoing in Unknown Environments," in IEEE/ASME Transactions on Mechatronics, vol. 23, no. 6, pp. 2750-2761, Dec. 2018.
3. M. Tsuru, A. Escande, A. Tanguy, K. Chappellet and K. Harad, "Online Object Searching by a Humanoid Robot in an Unknown Environment," in IEEE Robotics and Automation Letters, vol. 6, no. 2, pp. 2862-2869, April 2021.
4. O. Liu, S. Yuan and Z. Li, "A Survey on Sensor Technologies for Unmanned Ground Vehicles," 2020 3rd International Conference on Unmanned Systems (ICUS), 2020, pp. 638-645.

An Artificial Intelligence Camera System to Check Worker Personal Protective Equipment before Entering Risk Areas

Watthanaphong Muanme[1], Sawat Pararach[2] and Phisan Kaeprapha[3]

[1] Thammasat School of Engineering Thammasat University, Rangsit Campus Khlong Luang, Pathum Thani – 12120, Bangkok, Thailand
[2,3] Thammasat School of Engineering Thammasat University, Rangsit Campus Khlong Luang, Pathum Thani – 12120, Pathum Thani, Thailand
Wattanaphong@outlook.co.th

Abstract. Factories overlooking the need to check personal protective equipment (PPE) before workers enter risk areas is a factor contributing to injuries. To assist in resolving such issues, we are investigating and designing a safety management system (SMS) using artificial intelligence (AI) camera technology to detect PPE devices with a You Only Look Once (Yolo) deep learning algorithm. The technology checks and displays real-time operator PPE equipment inspections. The AI camera system tabulates and processes data from the extant database for accurate detection. The-camera can detect and notify about the following PPE devices: reflective clothing, helmets, safety goggles, and safety gloves. A warning will be displayed on the screen when PPE devices are not worn to check that each worker is wearing the correct PPE devices before entering risk areas. The model AI camera system can also operate in conjunction with automatic doors to prevent entry to a risk area and has been designed for use in industrial plant or job site safety management. Additionally, tests have shown that using this equipment increases the incidence of wearing PPE on entry to risk areas.

Keywords: Artificial intelligence (AI), Safety management (SMS), System design.

1 Introduction

Researchers studied the root causes of problems arising from the neglect of joint safety checks for workers who do not wear PPE [1]. The designed solutionhas been combined with modern industrial technology to develop a safety management system to make workers aware of the importance of the dangers that will arise during work that do not wear protective equipment. Some factories have neglected to check workers readiness checks, such as checking equipment, workers safety, and availability. Currently, some factories have a poor safety culture and underestimate the importance of less safe work practices, such as unsafe work due to employees neglecting to wear PPE and a lack of PPE checks before entering the facility. Working with risks that can cause accidents every time, such as an accident from hand cut workpieces from not wearing safety gloves, fire splashes in the eyes from not wearing safety glasses [2], or pieces of workpieces on high ground falling on the head when not wearing safety helmets, etc.

The researcher aim is to design a working system for detecting each type of device and to classify each device as follows: reflective vests, helmets, safety glasses, safety gloves. This detection uses a camera to detect PPE devices in real time [3][4], checking workers before entering work, This research uses safety management principles and AI technology systems, and Image Processing pro-working with The Yolov3 algorithms to help improve the process [5] of PPE checks for industrial facilities [5] or on-site to help reduce the

problem of neglect of personal protective equipment inspections and to reduce accidents caused by not wearing protective equipment [7]. This work builds on and improves upon works that have gone before, improvements have come from technological and hardware improvements allowing more accurate detection of more objects.

2 Related Works

2.1 Computer Vision and Machine Learning

Image processing and computer vision [8] can be applied management instead of humans without bias. It depends on the training information they receive. They can provide accurate and fast approval of entry into the safe zone. In the future, the computer system may extend to the reporting of violation. Image Processing [9] or Computer Vision is the processing of a learning algorithm and verifying the results to improve the outcome such that for a given task it can become better at making decisions and analyzing results. Images are processed using a set of algorithms known as YOLO that clearly defines the area to be searched for objects of potential interest to be classified. The system is trained on a set of images to classify target images [10]. In this system the targets are items of PPE: - Safety glasses, Hard hat, Reflective vest, and Gloves. The greater the number of images of PPE the more accurate the system should become.

3 Proposed System

3.1 The Process of Problem Analysis

Researcher uses the concept of finding the cause of safe work using the fishbone diagram analysis technique, which will be used to analyze [11] the true cause of the problem with the main idea of the heading of fish and any factors that contribute to the important issues that need to be determined to determine the root cause. (See Fig.1)

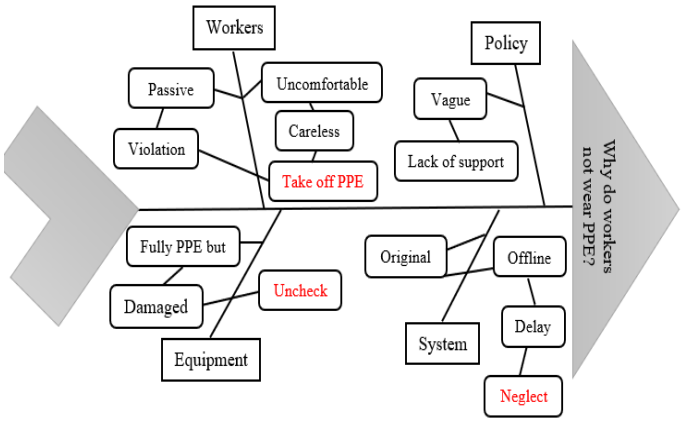


Fig. 1. Fishbone diagram analysis.

3.2 System of PPE Detection Algorithms

From Object Detection with Image Processing and learning, PPE Detection by Yolov3, the system uses a database of images of selected PPE to test images taken from the camera against. The more images and conditions that the systems test the greater the accuracy, according to the learning principles. An area of future work may be to attempt to improve training by using human assist training.

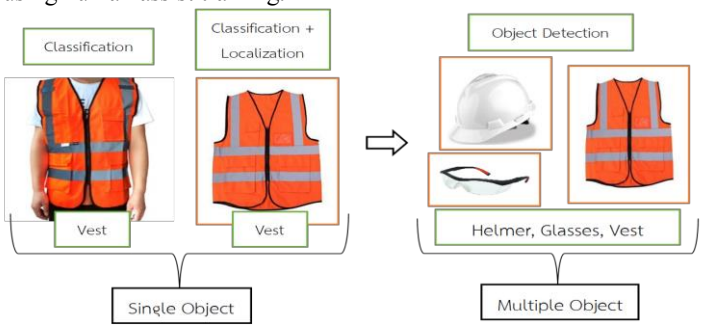


Fig. 2. Algorithms for object detection.

3.3 PPE Detections Algorithms by Yolov3

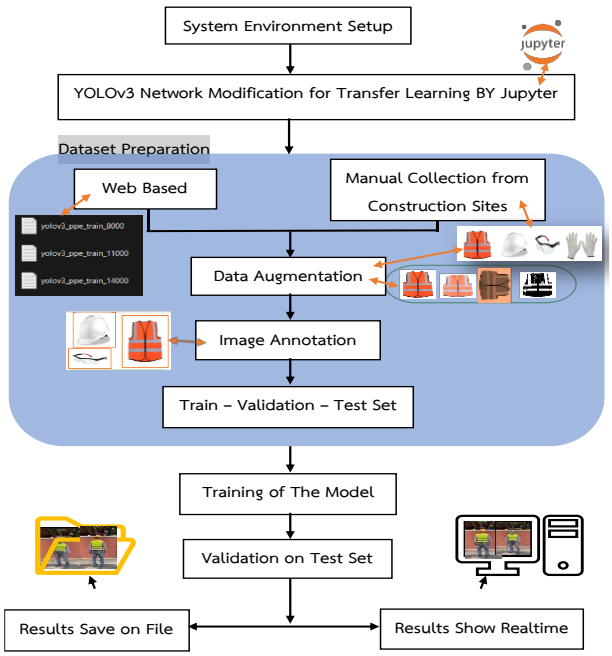


Fig. 3. Chart AI algorithms for PPE detection.

4. System Design and Experiment

In this study, the researchers propose an AI camera system for checking PPE by combining the concept of image processing with safety management. to solve problems arising

from neglect to wear the correct PPE or check before entering the risk work area. Using a digital camera to record video images in real time on the central computer screen, to show whether the workers is wearing PPE or not, rough digital processing with a theory called "Object Detection," the system can detect PPE devices and identify the type of safety device.

4.1 The Design of Systems Work

The work will include a central computer for analyzing the processed data and sending commands through the camera, and the AI camera system will work continuously in real time, with the door opening and closing in response to the specified program's commands. Before going to work, for workers to check by categorizing each type of equipment as follows: reflective vest, safety helmet, safety goggles, and safety gloves, The researcher has designed a system so that PPE equipment detection can be divided into 2 cases as follows this below.

- Case 1: The worker is wearing all required PPE, entry is allowed.
  Case 2: Some or all the of the PPE is missing or not worn correctly, entry is not allowed.

TABLE 1. In case of detection PPE

Object	Vest	Helmet	Glasses	Gloves
Present	Vest	Helmet	Glasses	Gloves
Not Present	No-Vest	No-Helmet	No-Glasses	No-Gloves

The system design of the camera-based detection system to check PPE, to make it easier to understand. (see in Fig. 4,5).

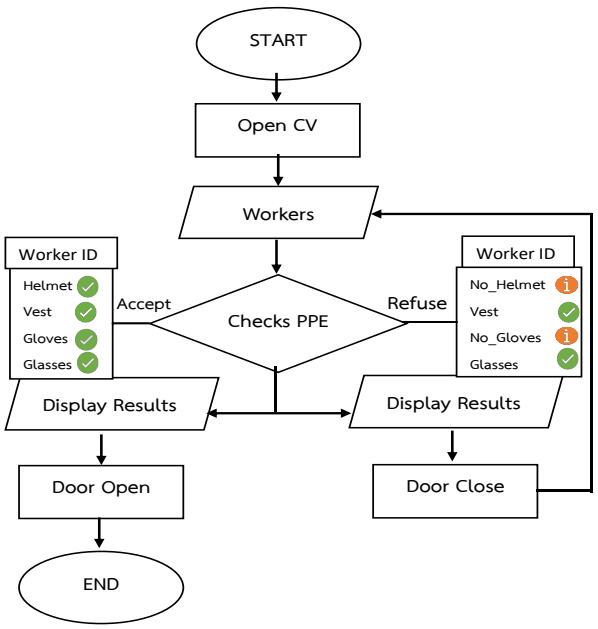


Fig. 4. PPE device detection system visual chart.

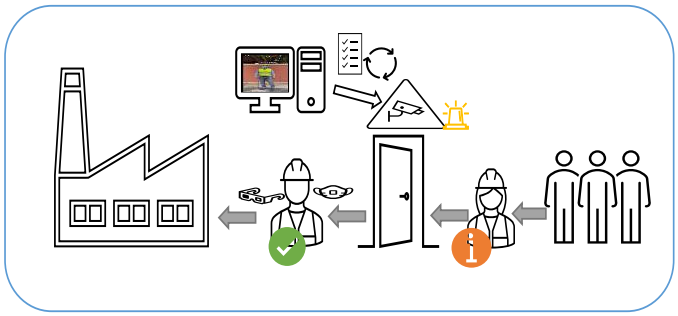


Fig. 5. PPE device detection system model.

The researcher uses a central computer to process the images and execute the commands. The program was designed to carry out the analysis, process and operate the cameras and doors. It is essential to use a central computer that is detached from the normal computer because it must have a relatively high processing power specification for real time performance to be achieved. (As in Fig. 6,7). The camera receives commands from the central computer to record detections in real time, detects PPE devices and transmits the resulting data values to the display. The door either opens or closes according to the main program command. It opens if all required PPE detected otherwise remains close
and wait for the next command from the central computer.

4.2 PPE Detection System Trials

A test of the PPE detection system when wearing five devices is (as shown in Fig. 3). A split-level test using a database with 5,000, 10,000 and 15,000 image training algorithms (Fig. 4) is used in real-time worker image detection training, compared to datasets with each level of training for Intersection Over Union (IOU) measurements of PPE detection performance.
The researchers set the standard value of the probability that each device object can be detected to have an accurate detection (IOU) value of between 0.7 and 1. If there is an (IOU) value below 0.7 this will prevent the system from detecting that type of device likewise, the closer the (IOU) value is to 1, the higher the detection accuracy. (See Fig. 6,7)

- Detection experiment when not wearing PPE

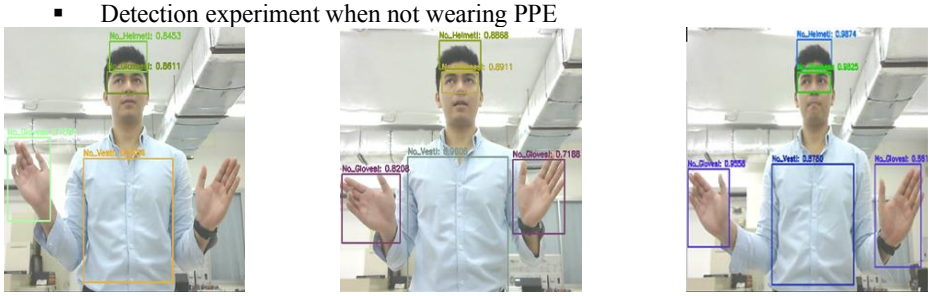


Fig. 6. The result is detected when not wearing PPE.

TABLE 2. Image training for PPE detection an accuracy (IOU) when not wearing PPE.

	Accuracy Intersetion Over Union (IOU)					
Image train	No_Helmet	No_Glasses	No_Vest	No_Gloves(L)	No_Gloves(R)	AVE.
5,000	0.8453	0.8611	0.7454	0	0.7301	0.63638
10,000	0.8868	0.8911	0.9808	0.7188	0.8208	0.85966
15,000	0.9874	0.9825	0.878	0.861	0.9558	0.93294

- Detection experiment when wearing PPE

Fig. 7. The result is detected when wearing PPE.

TABLE 3. Image training for PPE detection an accuracy (IOU) when wearing PPE.

	Accuracy Intersetion Over Union (IOU)					
Image train	Helmet	Glasses	Vest	Gloves(L)	Gloves(R)	AVE.
5,000	0.9973	0.7098	0.9982	0	0.888	0.71866
10,000	0.9976	0.8083	0.9473	0.8516	0.9552	0.912
15,000	0.9991	0.7868	0.9918	0.9967	0.9808	0.95104

From training 5,000, 10,000 and 15,000 images in order for the camera system to learn and remember to detect PPE, there will be differences in the IOU values as a measure of detection assessment. In our testing, 15,000 image trained detection AI camera systems yielded a better detection IOU than 5,000 and 10,000image training, training with more than 10,000 or more images should be sufficient for this camera system to detect PPE accurately.

The experiment confirmed that the AI camera system was able to detect PPE devices and was able to identify each type of personal protective equipment (PPE). Each area will have different protection tools, so the AI camera system can be adjusted to the work of that area.

5. Conclusion

Constraints on this system; currently the system will accurately detect PPE if the target is standing in a fixed pose at less than 150 cm from the camera.
At a Test site with 78 employees the results obtained with the system were as follows:

TABLE 4. Test site PPE results

	No PPE Worn	%	Partial PPE worn	%	Full PPE Worn	%
Prior to installation	22	28.21	30	38.46	26	33.33
After Installation	0	0	6	7.69	72	92.31

This research developed an AI camera system for monitoring personal safety. It distinguishes the security of each type of device in real time, eliminating problems caused by a lack of monitoring. Inspect each area within the factory or on the job site before entering work to speed up batch inspections. The table 4 shows that the system can detect and differentiate each PPE, being worn by employees in real time detection through the camera system.

6. Further Improvement

The real time training assisted by human operator to provide potential improvement in either accuracy or speed of detection.

Possibility of improvements in either accuracy or speed of detection can be obtained by having the candidate at variable distances and in fixed or free poses in front of the camera, this may limit the use of the system in some modes.

Peripheral connection of the system to door controls, ID card reader or Facial recognition to be able to report workers who exceed a pre-determined number of exceptions.

Trial the system for use with workers in high-risk areas who may be at risk of removing items of safety equipment.

References

1. Balamurugan Balakreshnan, Grant Richards, Gaurav Nanda, Huacho Mao, Ragu Athinarayanan and Joseph Zaccaria. (2020). PPE Compliance Detection using Artificial Intelligence in Learning Factories, 10th Conference on Learning Factories, CLF2020, PP 2-5.
2. Lifang Zhou, Hui Zhao and Jiaxu Leng, MTCNet: Multi-task collaboration network for rotation-invariance face detection. Pattern Recognition, PP 3-10, Nov 2021.
3. Xincong Yang, Yantao Yu, Sara Shirowzhan, Samad sepasgozar and Heng Li. (2020). Automated PPE-Tool pair check system for construction safety using smart IoT, Journal of Building Engineering, PP 2-10.
4. Ayatullah Faruk Mollah, Nabamita Majumder, Subhadip Basu and Mita Nasipuri, "Design of an Optical Character Recognition System for Camera based Handheld Devices", IJCSI International Journal of Computer Science Issues, Volume: 8, July-2011.
5. Rafael C. Gonzalez and Richard E. Woods. A textbook on “Digital Image Processing”Publications of Pearson, Second Edition,2002
6. Dzyubachyk O, Niessen W, Meijering E: Advanced Level – Set Based Multiple - Cell Segmentation and Tracking in Time - Lapse Fluorescence Microscopy Images. In IEEE International Symposium on Biomedical Imaging: From Nano to Macro Edited by: Olivo- Marin JC, Bloch I, Laine A. IEEE, Piscataway, NJ; 2008:185-188.
7. Quang-Huy Tran, Thi-Lan Le and Si-Hong Hoang. (2020). A fully automated vision-based system for real-time personal protective detection and monitoring, ResearchGate, PP 2-6.

8. L. Torres, "Is there any hope for face recognition?" in Proc. of the 5th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2004). Lisboa, Portugal, 2004.
9. Venkata Santosh Kumar Delhi, R. Sankarlal and Albert Thomas. (2020). Detection of Personal Protective Equipment (PPE) Compliance on Construction Site Using Computer Vision Based Deep Learning Techniques, Frontiers journal, PP1-8.
10. G.N. SRINIVASAN, Dr. SHOBHA G,"Segmentation Techniques for Target Recognition". International Journal of Computers and Communication, Issue 3, Volume 1, 2007
11. Jiss Kuruvilla, Anjali Sankar, Dhanya Sukumaran,"A Study on image analysis of Myristica fragrans for Automatic Harvesting"IOSR Journal of Computer Engineering (IOSR-JCE) ISSN: 2278-0661.p-ISSN: 2278-8727PP50-55

Applying 5PKC-based Skeleton Partition Strategy into Spatio-Temporal Graph Convolution Networks for Fitness Action Recognition

Jia-Wei Chang and Hao-Ran Liu*

National Taichung University of Science and Technology, Taichung City, Taiwan
kn880701@gmail.com

Abstract. With the rise of health awareness, people's demand for fitness has gradually increased. However, improper exercise may easily cause damage to the body. It would be possible to avoid wrong actions if automatic action recognition can detect and judge the human motion of exercises. Therefore, we aim to grasp the user's fitness status through human action recognition. However, most human action recognition mostly uses CNN-based models to process images, which may introduce unnecessary noise other than the human body from the background. To address this problem, we use the Spatio-Temporal Graph Convolutional Network (ST-GCN) as the backbone and take skeleton data as input to learn skeleton relationships. To further improve the accuracy, we propose a novel partition strategy based on Five Primary Kinetic Chains (5PKC) to explore the skeleton partition status and then enrich the skeleton relationships. Finally, the proposed method with 9 ST-GCN blocks that integrated the proposed partition strategy achieved 99.5% of accuracy which outperforms the model using 9 ST-GCN blocks with 84.5%.

Keywords: Action recognition, Fitness, ST-GCN, Five Primary Kinetic Chains

1 Introduction

In recent years, fitness has become increasingly popular around the world, and gyms have gradually increased. However, when exercising, improper exercise can easily cause damage to the body. Although there are fitness trainers who can assist, there are still some dangers that may be overlooked due to environmental or human factors. For gym owners, it is very important to improve the safety of the gym and reduce operating costs. Therefore, if automatic human action recognition can be used to continuously detect and judge the user's motion, it can not only effectively avoid human negligence and improve safety, also effectively reduce personnel costs. Even the user can use it at home, allowing the user to analyze whether the movement is qualified or not when exercising at home. Human action recognition has been widely used in multimedia computing, such as intelligent surveillance, virtual reality, and human–machine interaction. Although there have been many advances in the research of human action recognition in recent years, the high complexity and variability of human motion make the recognition accuracy and efficiency still have much room for improvement.

Most of the existing human action recognition models are based on images and consider the background. However, the same action will show completely different results in different illumination, viewing angles, and backgrounds. Most of the existing models are based on images and take into account factors such as background, which makes it easy to introduce unnecessary noise when performing action recognition. In order to deal with these noises, these models need to improve ability of modeling the change of background, but the processing will also increase burden on the models. Although some people reduce the impact of background noise by converting the image into a depth, thermal view, the effect of removing background noise is still limited. And the human body is a deformable object with a high degree of freedom, rather than a fixed shape. This makes it difficult to capture human body.

In the related works of human action recognition, deep neural networks [1] have become the main tool for this task. In recent years, it has been proposed to use skeleton-based temporal CNN or RNN for action recognition. By rearranging structured data, the human skeleton data is represented as a vector sequence to adapt to the neural network. Representative works include [2][3][4][5][6][7][8][9][10]. However, since the skeleton is essentially a non-Euclidean graph, directly sending the coordinates of the skeleton to the network cannot effectively analyze the structural information of the skeleton data [11].

To address this problem, Graph Convolutional Networks (GCN) [12][13][10] have been applied to skeleton-based action recognition because it can efficiently analyze the structural information of skeleton data. Spatio-temporal Graph Convolutional Network (ST-GCN) [14] is one of the representative works in skeleton-based action recognition. By using joints as nodes and connections between joints as edges, an undirected spatio-temporal graph is constructed, and a partition strategy is designed according to distance and spatial configuration, and the graph convolution operation is performed based on this. This method has also been proven to effectively improve the effect of human action recognition. Many variants derived from ST-GCN [14] have achieved excellent results [15][16][17][18][19], and ST-GCN [14] has also become the One of the most used frameworks for tasks.

In this paper, we propose a spatio-temporal GCN-based skeleton classification and scoring network. Based on the spatio-temporal graph convolutional network (ST-GCN) [14], it can effectively explore the distribution relationship between joints and joints through spatial configuration. The angle extends the relationship between the joints into successive frames. Then multiple ST-GCN layers are stacked to jointly transfer joints information in space and time. Finally, the action is classified by the score from the loss function.

2 Related Works

2.1 Skeleton-based Action Recognition

Human action recognition is based on the human body, which means that human skeleton is the most important basis for action recognition. Analysis of bones can also reduce the complexity of action recognition. Past methods usually rearrange the skele-

ton data into a grid-like structure or sequence of coordinate vectors and send it to CNN [4][6][7][9] or RNN [2][1][8] architecture. However, as stated in [11], the skeleton is a non-Euclidean graph, and the spatial subdivision of the skeleton cannot be effectively analyzed using CNN and RNN. With the development of GCN, GCN has been widely used in skeleton-based action recognition [14][15][20][21], because the spatial subdivision of skeleton can be effectively analyzed through GCN. The ST-GCN proposed by Yan et al. [14] first applied GCN to skeleton-based action recognition. It not only captures the relationship of joints in space, but also extends the relationship between joints to the concept of time, thereby capturing the relationship between joints in space and time between each consecutive frame.

2.2 Graph Convolutional Networks

Graph Convolutional Networks (GCN) generalize convolution to graph-structured data, applying them to irregular data such as interpersonal social networks and biological data. Graph convolutional networks can be divided into two types, namely in the spectral domain [13][10][22][23][24] or the spatial domain [20][25][26][27][28][29] to transfer node features. The former considers graph convolution from a spectral point of view, using the Fourier transform of the graph to operate in the spectral domain, while the latter obtains the information of its neighboring nodes directly on the node. To improve the performance of GCN, someone introduced attention mechanism in GCN [28][29].

3 Methodology

3.1 Human Skeleton Graph Construction

In action recognition, the skeleton graph uses joint as nodes and bones as edges to form a human body topology. There are 18 nodes in the human body topology, as shown in Fig.1, which are left hip, right hip, left knee, right knee, left ankle, right ankle, left foot, right foot, left shoulder, right shoulder, left elbow, right elbow, left wrist, right wrist, the left eye, the right eye, the head, and a dynamic center of gravity. Dynamic center is calculated by adding and averaging the coordinates of all nodes. The spatio-temporal skeleton is based on the input video, and each frame is converted into a skeleton and stacked in time series.

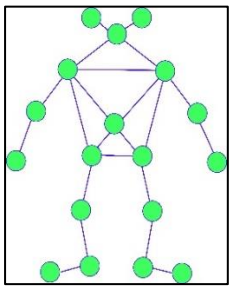


Fig. 1. Human Skeleton Graph

Suppose a given skeleton sequence has T frames, each with N joints. According to the human skeleton diagram in Fig.1, we construct an undirected graph and define it as $G = (V, E)$, where the joints are defined as $V = \{v_i \in \mathbb{R}^C | i = 1, \cdots, N\}$, N represents the number of joints, C is the dimension of joint features, all adjacent joints of joint v_i are denoted as $N(v_i)$, and bones are defined as $E = \{v_i v_j | (i, j) \in e\}$, and e represents a set of edges formed by every two adjacent joints.

3.2 Spatial Graph Convolution

To aggregate joint information from adjacent joints into each joint using graph convolution operation, as shown in Fig.2, we define the joint adjacency matrix $A \in \mathbb{R}^{N \times N}$ according to the skeleton graph, if e is Existing $A_{ij} = 1$, otherwise $A_{ij} = 0$. Since the human skeleton graph is an undirected graph, A is a symmetric matrix.

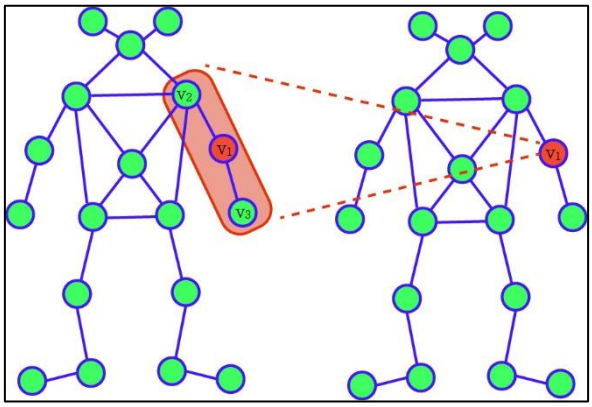


Fig. 2. Aggregate Adjacent Joint Information to each Joint

Next, for the graph convolution operation, we denote the adjacent region with 1-distance from v_i as $N(v_i)$, and the entire skeleton feature is defined as $V \in \mathbb{R}^{N \times C}$. The formula for taking V and A as the input of the spatial GCN is as follows:

$$V^{l+1} = \sigma(\bar{A} V^l W^l) \tag{1}$$

$V^l \in \mathbb{R}^{N \times C_l}$ is the skeleton feature of the l-th layer, $V^{l+1} \in \mathbb{R}^{N \times C_{l+1}}$ is the skeleton feature of the l+1th layer, C_l and C_{l+1} are the channel numbers of the l-th layer and the l+1th layer, respectively. $W^l \in \mathbb{R}^{C_l \times C_{l+1}}$ is the training weight of the l-th layer. $\bar{A} = \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}}$ is the normalized adjacency matrix, and $\tilde{A} = A + I$ is the adjacency matrix that increases the identity matrix to keep the senior features. $\tilde{D} \in \mathbb{R}^{N \times N}$ is the number of nodes matrix. Eq. (1) updates the features of each node according to the weighted average of the adjacent node features, and transforms the number of channels in each layer separately by the weights.

3.3 Spatio-Temporal Graph Convolution

The above introduction only considers the space in the graph convolutional network, which is only suitable for static images, but because we want to take a movie as input, we convert the movie into a spatio-temporal skeleton, as shown in Fig.3. According to [6], we extend the spatial GCN to the spatio-temporal GCN by redefining the adjacent positions of nodes. This means that the i-th node-adjacent node range $N(v_i)$ not only spatially adjacent nodes, but also contains the same joints on consecutive T frames. To this end, we redefine the skeleton feature V to the original space plus the dimension of time T as $V \in \mathbb{R}^{N\times C\times T}$.

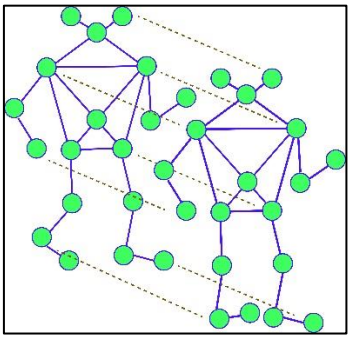


Fig. 3. Spatio-Temporal Graph for Human Skeleton

3.4 Skeleton Partition Strategy with Five Primary Kinetic Chains

We follow the spatial connection configuration of the skeleton in [14] and add other different connection methods. In the human skeleton, the human body is a structure composed of interconnected joints. In the process of exerting force, these joints that are responsible for transmitting force along the way are connected to each other, and they are connected together like a "chain" to form a kinetic chain, and the kinetic chain is the path of power output. Therefore, we refer to the Five Primary Kinetic Chains (5PKC) systems mentioned by Joseph in [30]. The functions by different joints are different, and they follow a certain logical distribution. Different kinetic chains exist as a single entity but also depend on each other to create a balanced and efficient movement, so we added it to the spatial connection configuration to extract more useful information from the human skeleton.

In a normal connection, the adjacent area of a node can be defined as Fig.2. The neighbors of node v_1 are $\{v_2, v_3\}$. And our redefined connection configuration is shown in Fig.4. The spatial connection configuration of Fig.4(a) is to divide each node into centrifugal groups that are farther from the center of gravity in each joint itself, according to the distance of other nodes in the adjacent area of each node and our custom dynamic center of gravity. and the centripetal groups in the adjacent regions that are closer to the center of gravity. Then there are the five primary kinetic chains systems mentioned by Joseph in [30], in which only the other four except the posterior oblique sling (POS) are used because the skeleton does not distinguish between front and rear. The deep longitudinal sling (DLS) Fig. 4(b), the anterior oblique

sling (AOS) Fig. 4(c), the lateral sling (LS) Fig. 4(d), the intrinsic (IS) Fig. 4(e), through these four connection configurations, the potential information of each joint in the skeleton can be analyzed more effectively.
And we also re-divide the original adjacent matrix A into four sub-matrices according to the above definition, and redefine Eq. (1) as Eq. (2):

$$V^{l+1} = \sigma\left(\sum_{i=1}^{4} \bar{A}_i V^l W_i^l\right) \tag{2}$$

Here i is the index of the sub-matrix, $\bar{A}_i$ is the i-th sub-matrix separated from the adjacent matrix A, and W_i^l is the trainable weight of the i-th sub-matrix.

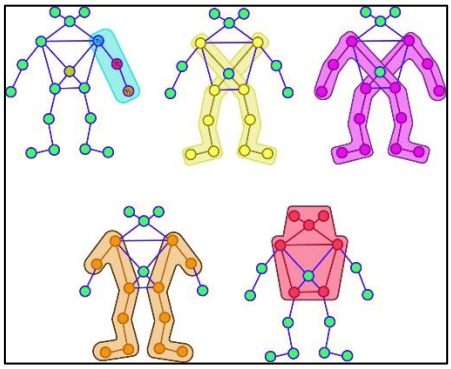


Fig. 4. Skeleton Partition Strategy

3.5 Network Architecture

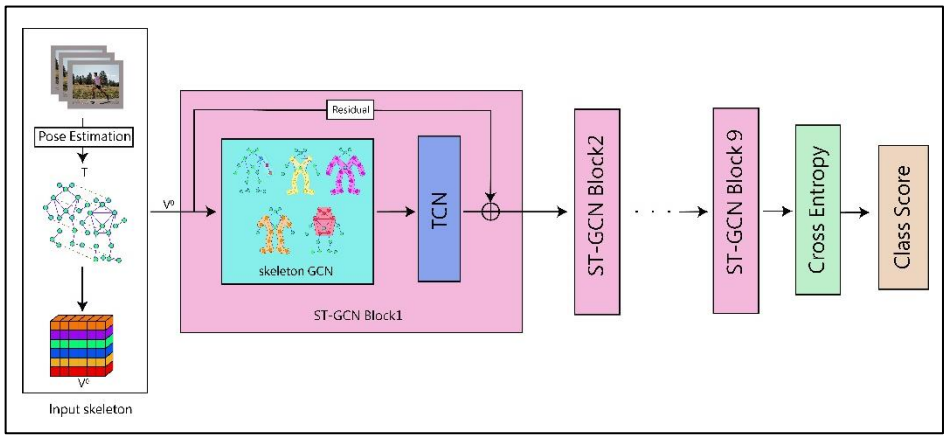


Fig. 5. Network Architecture of ST-GCN with 5PKC

Our network architecture is shown in Fig.5. We use a spatio-temporal graph convolution network based on ST-GCN [14], which takes the human skeleton converted from the video as input V^0, and then passes V^0 through Multiple stacked ST-GCN modules

process the spatio-temporal relationship of joints, and each ST-CGN is composed of GCN and TCN. Then use the Cross Entropy loss function to obtain the loss score, such as Eq. (3):

$$L = -\frac{1}{M}\sum_i \sum_{m=1}^{M} log\left(y_{im} log\left(P(Y_m|X_i)\right)\right) \quad (3)$$

M is the number of categories, $y_{im} \in \{0,1\}$ is the sign function, which is used to indicate whether the category of the input sample i is equal to the real category m, if they are equal, it is equal to 1, otherwise it is 0. $P(Y_m|X_i)$ represents the probability that the input sample X_i belongs to the real sample Y_m. The loss score can be obtained by Eq. (3).

4 Experimental Results

The dataset used in our experiments is InfiniteRep [31], which is suitable for detecting fitness and extracting human skeletons. We compare our proposed model with ST-GCN and judge the accuracy for classification. All experiments are performed on the PyTorch, a deep learning package. The initial learning rate is set to 0.1 and running for 50 epochs, the learning rate is reduced by a factor of 10 every 10 epochs. Batch size is set to 64. Each input consists of a multi-frame skeleton. The model consists of 9 ST-GCN layers. The number of channels is 64, 64, 64, 128, 128, 128, 256, 256, 256, respectively.

4.1 Datasets

InfiniteRep [31] is a synthetic dataset for fitness and physiotherapy (PT), which mainly consists of performing everyday fitness activities and repeating them multiple times, which are then converted into 3D joint points. This data set has ten categories, arm raise, bird dog, curl, fly, leg raise, overhead press, push up, squat, bicycle crunch, and superman. A total of there are 1,000 action data, and each will repeat the action 5 to 10 times in 7 indoor scenes.

4.2 Testing of ST-GCN blocks

We conduct an ablation study on the choice of the number of ST-GCN blocks, and we test the effect of one to nine blocks on the accuracy with ST-GCN plus dynamic center of gravity and five kinetic chains. The output channels of the 11 ST-GCN are 64, 64, 64, 128, 128, 128, 256, 256, 256, 512, and 512, respectively. We sequentially increase the number of ST-GCN, and the results are shown in Fig.6. As the number of ST-GCN blocks increases, the classification accuracy gradually increases. The best accuracy of 99.5% is achieved when the number of ST-GCN block is 9, and then the accuracy starts to decrease as the number of stacks increases. It shows that stacking too many ST-GCN blocks may lead to overfitting and lower accuracy.

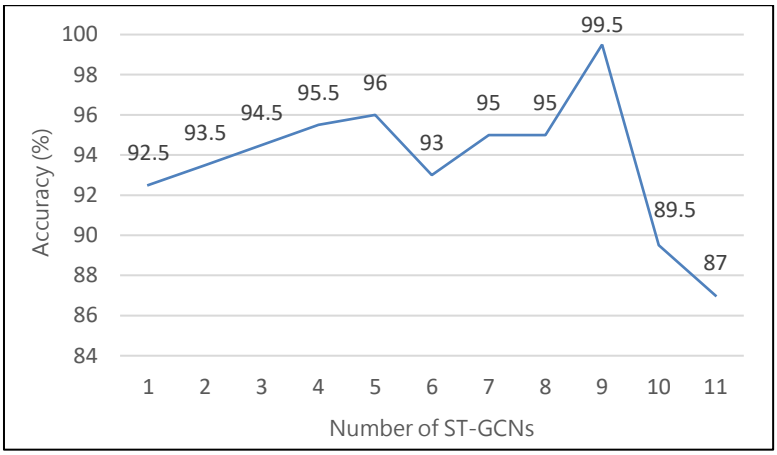


Fig. 6. Performance on ST-GCN blocks.

4.3 Partition Strategy Tests

In this section, we compare our model with the original ST-GCN model on InfiniteR-ep, and the results are shown in Table 1. The results show that adding a dynamic center of gravity to the skeleton can improve the accuracy compared to the original ST-GCN. Compared with ST-GCN, adding a dynamic center of gravity to the head can better capture the changes in motion. After adding five primary kinetic chains (5PKC) to the connection configuration, the accuracy has been improved, which proves that adding 5PKC can effectively transmit useful information between joints.

Models with different Partition Strategy	**Accuracy**
ST-GCN with 9 blocks	84.5%
ST-GCN with 9 blocks + Dynamic Center	93.5%
ST-GCN with 9 blocks + Dynamic Center + 5PKC	99.5%

Table 1. Ablation Comparisons of ST-GCN Models

5 Conclusions

Based on ST-GCN, this paper adds a dynamic center of gravity in the skeleton and five primary kinetic chains in the connection configuration. The proposed model can achieve abnormal action recognition, effectively process spatial relationships of the skeleton, and better capture the changes of motion. Experiments on the fitness action dataset show that the ST-GCN with the proposed partition strategy achieves classification accuracy of 99.5%, better than the original ST-GCN of 84.5%, proving the effectiveness of the proposed method.

References

1. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature 521(7553), pp.436-444 (2015).
2. Shahroudy, A, Liu, J, Ng, T-T, Wang, G.: Ntu rgb+ d: a large scale dataset for 3d human activity analysis. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1010-1019. (2016)
3. Liu, J., Shahroudy, A., Xu, D., Wang, G.: Spatio-temporal LSTM with trust gates for 3D human action recognition. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9907, pp. 816–833. Springer, Cham (2016).
4. Li, C., Hou, Y., Wang, P., Li, W.: Joint Distance Maps Based Action Recognition With Convolutional Neural Networks. In: IEEE Signal Processing Letters 24, pp. 624-628 (2017)
5. Ke, Q., Bennamoun, M., An, S., Sohel, F., Boussaid, F.: A New Representation of Skeleton Sequences for 3D Action Recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3288-3297 (2017).
6. Li, C., Zhong, Q., Xie, D., Pu, S.: Co-occurrence Feature Learning from Skeleton Data for Action Recognition and Detection with Hierarchical Aggregation. In: IJCAI, 2018, pp. 1–8. (2018)
7. Liu, M., Liu, H., Chen, C.: Enhanced skeleton visualization for view invariant human action recognition. Pattern Recognition.In: PR, pp. 346-362 (2017)
8. Song, S., Lan, C., Xing, J., Zeng, W., Liu, J.: An end-to-end spatio-temporal attention model for human action recognition from skeleton data. In: Proceedings of the AAAI conference on artificial intelligence, vol. 1 (2017)
9. Zhang, P., Lan, C., Xing, J., Zeng, W., Xue, J., Zheng, N.: View adaptive recurrent neural networks for high performance human action recognition from skeleton data. In: Proceedings of the IEEE international conference on computer vision, pp. 2117-2126 (2017)
10. Kim, T.S., Reiter, A.: Interpretable 3D Human Action Analysis with Temporal Convolutional Networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 1623-1631. (2017)
11. Li, M., Chen, S., Chen, X., Zhang, Y., Wang, Y., Tian, Q.: Actional-Structural Graph Convolutional Networks for Skeleton-based Action Recognition. (2019)
12. Liu, K., Gao, L., Khan, N.M., Qi, L., Guan, L.: A Vertex-Edge Graph Convolutional Network for Skeleton-Based Action Recognition. In: 2020 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1-5 (2020)
13. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)
14. Yan, S., Xiong, Y., Lin, D.: Spatial temporal graph convolutional networks for skeleton-based action recognition. In: Thirty-Second AAAI Conference on Artificial Intelligence (2018)
15. Defferrard, M., Bresson, X., Vandergheynst, P.: Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering. (2016)
16. Li, C., Cui, Z., Zheng, W., Xu, C., Yang, J.: Spatio-temporal graph convolution for skeleton based action recognition. In: Proceedings of AAAI, pp. 3482–3489 (2018)
17. Li, B., Li, X., Zhang, Z., Wu, F.: Spatio-Temporal Graph Routing for Skeleton-Based Action Recognition. Proceedings of the AAAI Conference on Artificial Intelligence, pp. 8561-8568 (2019)
18. Gao, X., Hu, W., Tang, J., Liu, J., Guo, Z.: Optimized skeleton-based action recognition via sparsified graph regression. In: Proceedings of the 27th ACM International Conference on Multimedia, pp. 601-610 (2019)

19. Shi, L., Zhang, Y., Cheng, J., Lu, H.: Two-Stream Adaptive Graph Convolutional Networks for Skeleton-Based Action Recognition. (2018)
20. Monti, F., Boscaini, D., Masci, J., Rodolà, E., Svoboda, J., Bronstein, M.M.: Geometric deep learning on graphs and manifolds using mixture model CNNs. (2016)
21. Peng, W., Hong, X., Zhao, G.: Tripool: Graph triplet pooling for 3D skeleton-based ac-tion recognition. Pattern Recognition (2021)
22. Peng, W., Shi, J., Zhao, G.: Spatial Temporal Graph Deconvolutional Network for Skeleton-Based Human Action Recognition. IEEE Signal Processing Letters, pp. 244-248 (2021)
23. Henaff, M., Bruna, J., LeCun, Y.: Deep Convolutional Networks on Graph-Structured Data. (2015)
24. Duvenaud, D., Maclaurin, D., Aguilera-Iparraguirre, J., Gómez-Bombarelli, R., Hirzel, T., Aspuru-Guzik, A., Adams, R.P.: Convolutional Networks on Graphs for Learning Molecular Fingerprints. (2015)
25. Li, Y., Tarlow, D., Brockschmidt, M., Zemel, R.: Gated Graph Sequence Neural Networks. (2015)
26. Bruna, J., Zaremba, W., Szlam, A., LeCun, Y.: Spectral Networks and Locally Con-nected Networks on Graphs. (2013)
27. Zoph, B., Le, Q.V.: Neural Architecture Search with Reinforcement Learning. (2016)
28. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., Bengio, Y.: Graph Attention Networks. (2017)
29. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I.: Attention Is All You Need. (2017)
30. Joseph Schwartz.: The 5 Primary Kinetic Chains (5PKC). https://dna-assessment.com/the-master-template/ (2016). Accessed 20 Aug 2016
31. Weitz, A., Colucci, L., Primas, S., Bent, B.: InfiniteForm: A synthetic, minimal bias dataset for fitness applications. (2021).

A Skeletal Sequence-Based Method for Assessing Motor Coordination in Children

Zitong Pei[1], Wenai Song[1], Nanbing Zhao[1], Zhiyu Chen[1], Wenbo Cui[1], Yi Lei[2], Yanjie Chen[3], Qing Wang[4,5]

[1] School of Software Engineering, North University of China, Taiyuan 038507, China
[2] Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China
[3] Department of Children's Health Care Centre, Beijing Children's Hospital, Capital Medical University, Beijing 100045, China
[4]Department of Automation, Tsinghua University, Beijing 100084, China
[5] Pharmacovigilance Research Center for information technology and Data Science, Cross-strait Tsinghua Research Institute, Xiamen 361015, China
qing.wang@tsinghua.edu.cn

Abstract. Children's motor coordination is an important component of physical fitness test for young children. The development of children's motor coordination occurs throughout children's motor development, and it is not only limited by the maturity of children's neurodevelopment, but also plays an important role in promoting children's neurodevelopment. This study proposes an automatic assessment method based on deep learning to improve assessment efficiency and reduce costs. The method combines human posture estimation, similarity calculation and time series feature extraction for the assessment of children's movements. The results showed that the accuracy rate and redundancy rate of the fine action coin toss keyframes finding algorithm are 89.8% and 7.5%; the accuracy rate of the dynamic action standing long jump keyframes finding algorithm is 74.3%; the accuracy rate, precision rate and recall rate of the fine action coin toss assessment algorithm are 74.4%, 72.5% and 90.0%; the accuracy rate, precision rate and recall rate of the static action single-leg balance assessment algorithm are 87.1%, 73.5% and 87.1% ; the accuracy rate, precision rate and recall rate of the dynamic action standing long jump assessment algorithm are 71.6%, 73.5% and 71.4%; the results of the three actions generally matched the expert assessment results. The method provides a good auxiliary tool for determining the motor development level of young children, and provides a good technical support for achieving the goal of "promoting the early comprehensive development of young children through actions".

Keywords: Human Posture Estimation, Similarity, Action Assessment, Time Series Features.

1 Introduction

The development of children's movement is very important in early life [1]. The development of children's movements is not only the achievement of milestone projects, but also the identification of development risks behind children through movement development, and the earlier the risk is identified, the higher the value is. Many children with developmental disorders have certain problems in their motor development, such as children with autism will have somatization movements, children with developmental retardation will have body use disorders, children with developmental coordination disorders will have clumsy movements, and so on [2,3]. Developmental coordination disorder is mainly characterized by clumsy movement and poor physical coordination. At present, the incidence is about 6% ~ 8% [4].

At present, MABC-2[5] test is used to evaluate children's motor coordination. The test method is an international test method for children's sports coordination development level, which is suitable for children aged 3~6 years. At present, the assessment of the development of children's movement in China mainly adopts the form of scale, mainly based on standardized guidelines, toolbox, etc. The assessment standard is interpreted manually, which may produce certain deviation. There will be certain differences in the judgment of children by different personnel, different institutions and different environments. The assessment of intervention effect will also have a certain impact.

The automatic assessment of the development of children's movement using artificial intelligence is mainly based on key technologies such as human target detection, human keypoints recognition, motion capture, pose estimation and fine statistical measurement. These technologies are currently mostly used in film entertainment and sports video analysis. According to the report released by Fior Markets in 2018, the global motion capture market is expected to grow from $163.2 million in 2018 to $261.7 million in 2026, the CAGR is 8.13% in the forecast period 2019-2026. With the development of global health care and sports industry and the improvement of people's attention to sports health, the demand and scale of motion capture related industries are growing driven by market applications in many fields such as health and sports, and the market share is also increasing [6].

This study attempts to implement the research results in the related fields of human pose estimation and movement assessment into the diagnosis system of children's developmental dyskinesia. If the system can be applied to the actual auxiliary examination, it can improve the diagnosis and treatment ability of areas with insufficient medical ability, timely intervene and treat children with the disease, so as not to affect their normal development. Based on the existing manual design features and depth features, this paper explores the assessment method to optimize motion assessment, and puts forward a motion assessment method for developmental coordination disorders.

2 Methods

The motion assessment method uses the following methods: keypoints detection, keyframes detection and similarity calculation [7]. The process is based on cropping and formatting the video, keypoints detection and output of the skeletal sequences for the test video, data processing and extraction of the keyframes, similarity calculation between the skeletal sequences of the test action keyframes and the skeletal sequences of the standard action keyframes, and finally a composite score where the scored parts of each action are weighted and summed up using different weights.

2.1 Keypoints Detection

In this paper, we use MediaPipe algorithm to detect keypoints. In the cross platform artificial intelligence work pipeline framework MediaPipe [8], Google launched the body pose attention function BlazePose with the latest technology, which can instantly and accurately locate the keypoints of body pose on the mobile phone. Now the standard model for focusing on pose is based on the COCO topology, but it mainly depends on the powerful computing power of desktop computers. The human pose sensing method BlazePose released by Google now uses machine learning to infer 33 2D feature points of human body, as shown in Fig.1. In addition to being more accurate than the coco topology, BlazePose can use the CPU of the mobile device to make real-time speculation. In addition to pose, BlazePose can also pay attention to facial expression and hand pose at the same time.

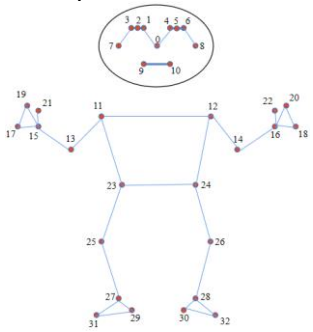


Fig. 1. Topological graph.

Fig. 2. Coin action keyframes.

Google showed the application scenarios of BlazePose, including squats and push up. The application can automatically count user data, verify technology and train quality. Therefore, the human pose estimation based on MediaPipe algorithm proposed in this paper can meet the basic needs of children's pose estimation.

2.2 Assessment Methods

After obtaining the preprocessed skeleton sequence, we can compare it with a preset template motion skeleton sequence, obtain the assessment results of motion quality

through the comparison with the standard motion, then calculate the similarity of keypoints between the corresponding coordinate points to measure the matching error between the two skeleton sequences, and then score the quality assessment score.

OKS (Object Keypoints Similarity). In the human body keypoints assessment task, the quality of the keypoints obtained by the network is not calculated only by simple Euclidean distance, but by adding a certain scale to calculate the similarity between the two points. This indicator is mainly used in multi-person pose estimation tasks. However, the original calculation method does not take the importance of each keypoints into account. We introduce a penalty factor σ. Different penalty factors are set according to the importance of the keypoints. The greater the penalty factor, the greater the impact caused by the deviation of the keypoints.

$$\mathrm{OKS_p} = \frac{\sum_i e^{\frac{-d_p^2}{2S_p^2 k_i^2}} \delta\left(v_{pi}>0\right)}{\sum_i \delta\left(v_{pi}>0\right)} \tag{1}$$

In the above formula: The prepresents the ID of a person in the growth truth. The p^i indicates the ID of the keypoints. δ is 1 at $v_{pi} > 0$ and 0 otherwise, meaning that only visible keypoints are calculated. The S_p represents the square root of the area occupied by this person, which is calculated according to the box of people in the ground truth. The k_i represents the normalization factor of the i-th bone point, which is obtained by calculating the standard deviation of all ground truth in the existing data set, reflecting the impact of the current bone point on the whole. The larger the value, the worse the standard effect of this point in China in the whole data set; The smaller the value, the better the annotation effect of this point in the whole data set. The d_p represents the Euclidean distance between the detection keypoints and the standard keypoints.

According to the clinician's instructions, each subject's movement is judged as a composite, so that each movement is weighted and summed using different weights, and when the weighted sum of the scores is greater than 0.5, the movement is considered normal. The composite score is calculated by means of equation (2).

$$S = \sum_{i=1}^{n} We_i * S_i \tag{2}$$

Of which, n is the number of scored parts, We_i is the weight of the i-th score, and S_i is the score of the ith score.

Fine movements. For the fine movements, a coin toss (6 tosses with the right hand as the dominant hand) is selected as a case study. The coin action is subdivided into four parts: scoring the similarity of keyframes, scoring the duration, scoring the standard deviation of the right hand keypoints coordinates of keyframes and scoring the complexity of the right hand keypoints coordinates waveform graph time series. The keyframes are extracted by setting the prominence of the peaks, the minimum height of the peaks, and the minimum horizontal distance between adjacent peaks of the right hand keypoints coordinate time series waveform graph [9]. The keyframes are shown in Fig. 2, and Fig. 3 shows the number of keyframes for all data. According to Fig. 4, we can observe that the peak wave tip is for the action of throwing in the coin. The similarity scoring of keyframes is based on the similarity calculation between the keypoints coordinates of the keyframes of the test action and the keyframes of the standard action according to equation (1). According to the specific situation of collecting the

coin throwing action, avoiding the inaccuracy of keypoints identification caused by the obscuration of the table and the judging criteria of the coin action, we adopt the upper itself keypoints for the similarity calculation. Time duration scoring is based on the Movement Assessment Battery for Children-Second Edition (MABC-2). Time series complexity is scored by calculating the degree of confusion between the front and back parts of the keypoints coordinate waveform. The final assessment is based on equation (2) and the scoring results for all data are shown in Fig. 5.

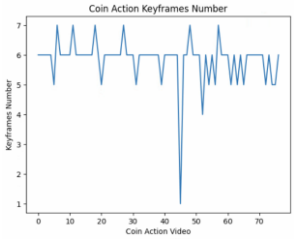


Fig. 3. the number of keyframes for coin action.

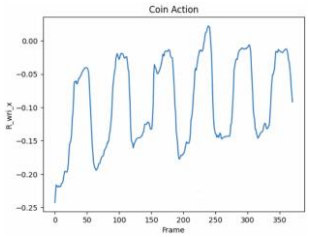


Fig. 4. X coordinate of the right hand side of the coin action.

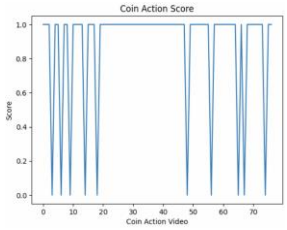


Fig. 5. Coin action scoring results.

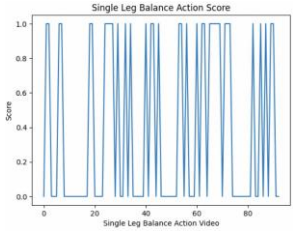


Fig. 6. Single leg balance action scoring results.

Static movements. For the static movements, a single-leg balance movement is chosen as a case study. The single-leg balancing movement is subdivided into three parts: the similarity score of keyframes, the duration score and the standard deviation of keypoints coordinates. According to the assessment rules, every frame in the standard movement video is immovable and identical, so we choose one frame from the standard movement video as the standard frame. The scoring of keyframes is a similarity calculation between the keypoints sequence of each frame of the test action video and the keypoints sequence of the standard frame based on equation (1). Duration scoring is based on the Movement Assessment Battery for Children-Second Edition (MABC-2). The standard deviation score is calculated by calculating the standard deviation of the coordinates of each keypoints and the resulting standard deviation data is used to indicate the stability of the single-leg balance movement. The final assessment is based on equation (2) and the scoring results for all data are shown in Fig. 6.

Dynamic movements. For the dynamic movements, the standing long jump is chosen as a case study. The vertical jump is refined into five keyframes for the similarity calculation to be assessed. Based on the clinician's guidance, we select five keyframes of the standing long jump movement by combining the angle of the joint, the coordinates of the keypoints and the positions of the keyframes that have been derived, as shown in Fig. 7. The five keyframes of the test movement and the five keyframes of the standard movement are calculated according to equation (1). The final assessment is based on equation (2) and the scoring results for all data are shown in Fig. 8.

Fig. 7. Standing long jump keyframes

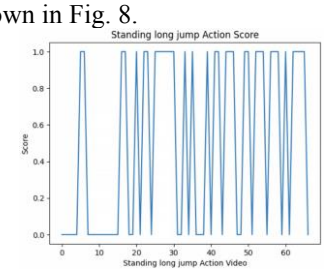


Fig. 8. Standing long jump action scoring results

3 Experimental Results and Discussion

The medical action video data used in this paper is not a public dataset and is provided by the collaborators - the Digital Medical Health Engineering Research Centre of Tsinghua University Institute of Information Technology and Beijing Children's Hospital. Medical data is sensitive and is used for scientific research to strictly protect patient privacy. Once the collection is complete, the clinician will do a diagnostic assessment of the subjects, complete the labelling of the data to separate the normal samples from the abnormal samples, and label the movement of the subject samples as positive abnormal as well.

We designed experiments to compare the effectiveness of OpenPose and MediaPipe in detecting the three movements of coin, single leg balance and standing long jump, as well as the effectiveness of the three movement assessment methods mentioned above [10,11,12].

3.1 Keypoints Detection

In order to try to avoid the problem of loss of detection values due to overlap of some keypoints of the human body in 2D images, the data acquisition criteria are developed with requirements on the angle of the shot to ensure the integrity of the action but also to avoid overlap as much as possible and to minimize the loss of some joint points caused by overlap, which could affect subsequent analysis and processing. We used OpenPose and MediaPipe to estimate human posture for the three movements of coin toss, single-leg balance and standing long jump respectively, as shown in Fig. 9. And the recognition rate calculation of each action by OpenPose and MediaPipe is derived by equation (3).

(a) (b)

Fig. 9. Keypoints detection chart. (a)OpenPose Detection(b) MediaPipe Detection

$$\text{Recognition Ratio} = \frac{X}{X+Y} \tag{3}$$

In the formula Recall Ratio is the recognition degree, X is the number of keypoints detected and Y is the number of keypoints undetected.

Table 1. OpenPose and MediaPipe recognition effects.

	Coin toss		Single leg balance		Standing long jump		Average	
Recognition algorithms	Velocity(fms)	Recognition ratio (%)	Velocity(fms)	Recognition ratio (%)	Velocity(fms)	Recognition ratio (%)	Velocity(fms)	Recognition ratio (%)
OpenPose	12	76.8	12	98.5	12	74.7	12	83.3
MediaPipe	8	99.9	7	100	7	98.8	7.3	99.5

Table 1 shows that the recognition rate of MediaPipe is higher than that of OpenPose . The MediaPipe model is slow due to its top-down detection method, while the OpenPose model is fast due to its bottom-up detection method. OpenPose and MediaPipe are comparable in terms of single-person recognition accuracy, MediaPipe outperforms OpenPose overall due to the high level of timeliness and accuracy required by the human motion recognition algorithm. The MediaPipe model is chosen for skeletal keypoints detection based on relevant performance comparisons and usage context.

3.2 Action Assessment

A control group is made based on the clinician's artificially selected keyframes images and assessment results. The data collected for the three movements of coin toss, single leg balance and standing long jump are processed separately using specific keyframes extraction algorithms and assessment algorithms, and the results obtained are analyzed in combination with the control group. The redundancy rate, accuracy rate, precision rate and check-all rate are used as assessment metrics to test the keyframes and assess the method performance.

$$Fa = \frac{n}{m} \tag{4}$$

$$Fr = \frac{M-n}{M} \tag{5}$$

The formula Fa is the accuracy, n is the accuracy value and m is the number of frames extracted by the clinician;Fr is the redundancy and M is the number of frames extracted.

Table 2. Keyframes extraction results for each action.

Type of action	Coin toss	Standing long jump
Extracting the correct number of frames	418	299
Extracting frames	452	402
Specified number of frames to be extracted	462	402
Accuracy (%)	89.8	74.3
Redundancy (%)	7.5	-

$$Accuracy\ Ratio = \frac{TP+TN}{TP+FP+TN+FN} \tag{6}$$

$$Precision\ Ratio = \frac{TP}{TP+FP} \tag{7}$$

$$Recall\ Ratio = \frac{TP}{TP+FN} \tag{8}$$

In the formula, Accuracy Ratio is the accuracy ratio, Precision Ratio is the precision ratio and Recall Ratio is the recall ratio; TP is a positive sample with a positive prediction result; TN is a negative sample with a negative prediction result; FP is a positive sample with a negative prediction result; and FN is a negative sample with a positive prediction result.

Table 3. Analysis of action assessment results.

Type of action	Coin	Single leg balance	Standing long jump
Recall ratio (%)	90.0	87.1	71.4
Precision ratio (%)	72.5	73.5	73.5
Accuracy ratio (%)	*74.4*	87.1	71.6

According to Table 2 respectively, the keyframes extraction accuracy for the standing long jump movement is 74.3%, and the keyframes extraction accuracy for the coin movement is 89.8%, with a redundancy of 7.5%, which achieved a relatively good keyframes extraction effect. According to Table 3, the recall ratio, precision ratio and accuracy ratio for the coin action assessment algorithm are 90%, 72.5% and 74.4%; the recall ratio, precision ratio and accuracy ratio for the single leg balance action assessment algorithm are 87.1%, 73.5% and 87.1%; the recall ratio, precision ratio and accuracy ratio for the standing long jump action assessment algorithm are 71.4%, 73.5% and 71.6%. The assessment methods for all three movements were generally consistent with the physician's assessment.

4 Conclusion

In this paper, a combination of human pose estimation, similarity calculation and time series feature extraction methods is used to achieve movement assessment of children. Different methods are used for data pre-processing, keyframes extraction and movement assessment for the coin, single leg balance and standing long jump movements,

and the results of keyframes selection and movement assessment are approximately the same as our manual judgement results. At a later stage, when more and more data are available, deep learning can be used to train models for relevant movement assessment, which can be more accurate and convenient and will make a great contribution to the development of movement assessment for children.

Acknowledgements. This work was supported by the National Key R&D Program of China (2020YFC2006702, 2020YFC2005503)

References

1. Chen Y J, Wang H, Liang A M.: Study on the relationship between children's social skills and developmental coordination disorder[J]. Chinese Journal of Reproductive Health, 2021, 32 (04): 311-314.
2. Dong Y G.: Study on the relationship between motor development and physical health level of children in grades 1-3 in Beijing[D]. Capital Institute of Physical Education, 2021.
3. Chen Y J, Wang H, Liang A M.: Study on the relationship between children's physical evaluation and coordination disorder evaluation index[J]. Chinese Journal of Child Health Care, 2021, 29 (05): 542-544 + 549.
4. Wu D, Tang J L.: Diagnosis and treatment of developmental motor coordination disorder[J]. Chinese Journal of Rehabilitation Medicine, 2020, 35:513-516.
5. Hua J, Wu Z C, Meng W., et al.: Preliminary analysis on the application validity of child developmental coordination disorder assessment tool in China[J]. Chinese Journal of Child Health Care, (2010-07): 28-31.
6. Zhu Z P.: Analysis on age characteristics of standing and jumping learning of preschool children[D]. Overseas Chinese University, 2020.
7. Wang J, Qiu K, Peng H., et al.: "AI Coach: Deep Human Pose Estimation and Analysis for Personalized Athletic Training Assistance[C]. Proceedings of the 27th ACM International Conference on Multimedia. 2019: 374-382.
8. Bazarevsky V, Grishchenko I, Raveendran K., et al.: BlazePose: On-device real-time body pose tracking[J]. arXiv preprint arXiv:2006. 10204, 2020.
9. Wang Ziyi.: Research on time series classification methods based on feature extraction[D]. Nanjing: Nanjing University, 2019.
10. Guan C.: Realtime multi-person 2d pose estimation using shufflenet[J]. 2019 14th International Conference on Computer Science & Education (ICCSE). IEEE, 2019: 17-21.
11. KEHAN CHEN.: Sitting Posture Recognition Based on OpenPose[J]. IOP Conference Series: Materials Science and Engineering, 2019, 677(3):032057.
12. YOHEI OKUGAWA, MASAO KUBO, HIROSHI SATO., et al.: Evaluation for the Synchronization of the Parade with OpenPose[J]. Journal of Robotics, Networking and Artificial Life, 2019, 6(3):162.

A Big Data based Learning Model from Student Questionnaire

Hwa-Young Jeong

Humanitas College, Kyung Hee University, 1 Hoegi-dong, Dongdaemun-gu, Seoul, 130-701, Republic of Korea
hyjeong@khu.ac.kr

Abstract. The results of analyzing students' requirements become very important data for teachers. Because it is a very important criterion for improving learning satisfaction while increasing students' learning effects. In this paper, students' requirements are analyzed and what subjects or contents students want are investigated. The purpose of this paper is to construct a learning model that reflects the needs of students.

Keywords: Big Data, Word Cloud, Student's Needs, Learning Model.

1 Introduction

Until the late 21st century, general teaching methods were mainly based on frontal teaching in classrooms. The use of digital tools has the effect of increasing learning efficiency, but the method of education insisted on traditional classroom lectures [3]. After COVID-19 pandemic, university education was changed and operated from offline to online, and many students became familiar with online education [1]. Teleconferencing platforms such as Zoom have become hugely popular with the spread of online education due to COVID-19 pandemic. This has become a catalyst for many institutions to seek traditional education methods as various educational methods that can conduct synchronous or asynchronous remote education [3]. In online education, learners' activities and learning data were analyzed and processed using a LMS (Learning Management System). The large amount of learning data extracted from the LMS platform provided basic information for both teachers and students that could help improve learning satisfaction and educational goals [1]. In addition, students' learning satisfaction is generally obtained through a post-learning survey, and most of them analyze and use it when students answer multiple-choice questions written by teachers. However, the multiple-choice questions given by the teacher are not sufficient data for the student satisfaction survey, and the results of the student's data analysis may vary depending on the teacher's intention. For this purpose, it is necessary to analyze students' learning satisfaction and their needs through big data analysis when presented as a subjective question, not a multiple-choice question, and students freely describe their opinions and submit answers to the questionnaire.

NLP (Natural Language Processing) is natural human language and communication variables such as voice, text, audio, and video, interpretations and applications [2] We call NLP as non-formal data. Text mining that is one of technique in Big data is well known a process to extract meaning words from the data and analyze them.

This paper aims to make a learning model using analysis of students' needs. For this process, students get a subjective questionnaire and they write their opinion freely without any form and select items. Their answers is used to analyze what they want to study or learn. Consequently, this research shows a frequency of their opinions that can be the students' needs from their study.

2 Related Works

2.1 Big Data

Big data is a technology that can handle unstructured data, unlike databases that only dealt with structured data. Of course, this can perform both processing of collecting and analyzing unstructured and structured data.

It also guarantees both volatility, speed, volume, diversity and integrity. In big data, data is collected in various ways, such as a web browser and a mobile web with various data formats. In the existing analysis method, different formats of unstructured and structured data could not be managed, but both big data are possible. In big data, Hadoop is cost-efficient, scalable, and enables fast and flexible parallelism. It also uses the Hadoop framework for big data analysis because it provides availability, resilient properties, security, and authentication. It is also well known for its open-source software architecture, which includes processing and storage. The part to be stored is HDFS (Hadoop Distributed File System), and the part to be processed is MapReduce. D.K. Jain et al. [4] depicted the Hadoop architecture as shown in Fig. 1.

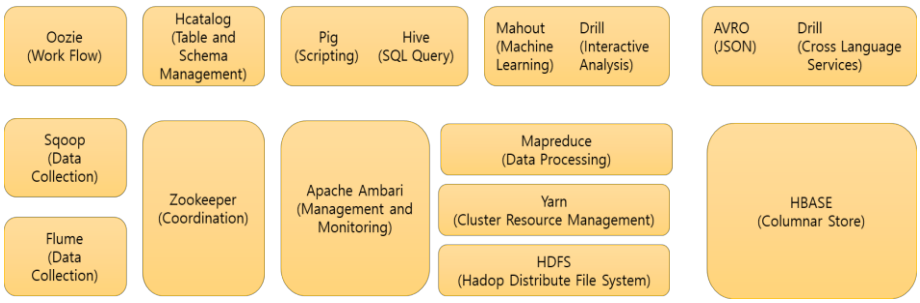


Fig. 1. Hadoop architecture by D.K. Jain et al.

Hadoop, first run by Doug Cutting and Mike Caparella in 2005, is a Java-based open-source software. Providing a distributed framework for processing and managing big data is known as Hadoop's advantage. Therefore, using Hadoop can manipulate a large amount of data. In addition, the Hadoop Distributed File System (HDFS) can store MapReduce for the purpose of the process. MapReduce is used to analyze

and generate a wide range of datasets using distributed and parallel processing methods in clusters [4].

Big data is being used in various fields. In particular, techniques for analyzing the frequency of words in unstructured data to show importance and visualize them are widely known. Figure 2 shows an example of the results analyzed by word cloud of big data for CC (Cognitive Computing) [5].

Fig. 2. A sample visualization by analyze word cloud of Big Data.

2.2 Learning Model

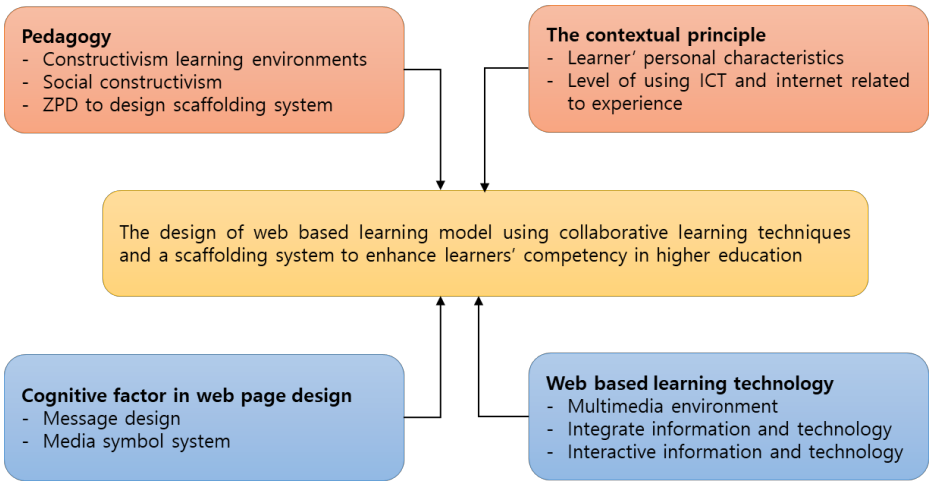


Fig. 3. The web based learning model by K. Deejring

The e-learning system promotes learning as teachers and students interact in a non-face-to-face manner, and learning materials and evaluations are also conducted online. This approach can be a good solution to accommodating an exponentially growing number of students and their curriculum, but only popular large-scale online curriculums can cause problems such as student crowding, student performance, failure and departure [6]. K. Deejring depicted e-learning model as shown in Fig. 3 [7].

3 Data Analysis for Students' Needs

3.1 Survey of Students' Requirements for their Class

For this study, a survey was conducted on 40 general students who did not major in IT in the liberal arts subject of K University in Seoul. The survey items are as follows.

Question: *Feel free to describe any topic or content you want to hear in the lecture. Or if there is a subject that you want to make a lecture in general culture, please write it with the reason.*

The comments as below represent one of the student's answers among the results of the survey.

Answer: *In my opinion, the topic we need in the Fourth Industrial Revolution is 'human ennui among the negative aspects of liberal utopia'. In the 4th Industrial Revolution, as each of us could access information or technology so easily, it destroyed the feudal system of the past and transformed it into a nation by the public, giving everyone a chance to politics by democracy and capitalism. As educational opportunities diversified, it began to turn into a world where everyone creates opportunities through education. But we must think about whether we will be happy. I get everything I want and come across, but I feel despondent. The reason is that in the era of the Fourth Industrial Revolution, we have it whenever we want, so we work hard to achieve something and do not feel the joy, satisfaction, and fulfillment that comes from it. Therefore, we maximize convenience and efficiency, but we may not be able to enjoy our mental satisfaction and happiness.*

3.2 Frequency from the Survey

Students' survey results are unstructured data written freely. If you analyze the frequency of words using the word cloud of big data, it is shown as shown in the Fig 4. The results of this analysis were all processed students' opinions, and they were not able to extract actually important words from the students' answers, but only the frequency was processed from the entire content.

Fig. 4. The result of word cloud

Fig. 5. The result of word cloud filtering meaningless word

Fig 5 shows the results except for meaningless words. What these results show is that students want subjects related to technology and are paying a lot of attention to preparing for the future. In particular, if you look at topics such as education, development, paradigm, artificial, computer, big data, metabus, and program, it can be seen that students want lectures related to this.

4 Learning Model

A learning model as shown in the Fig 6 is constructed based on the results of student requirement analysis.

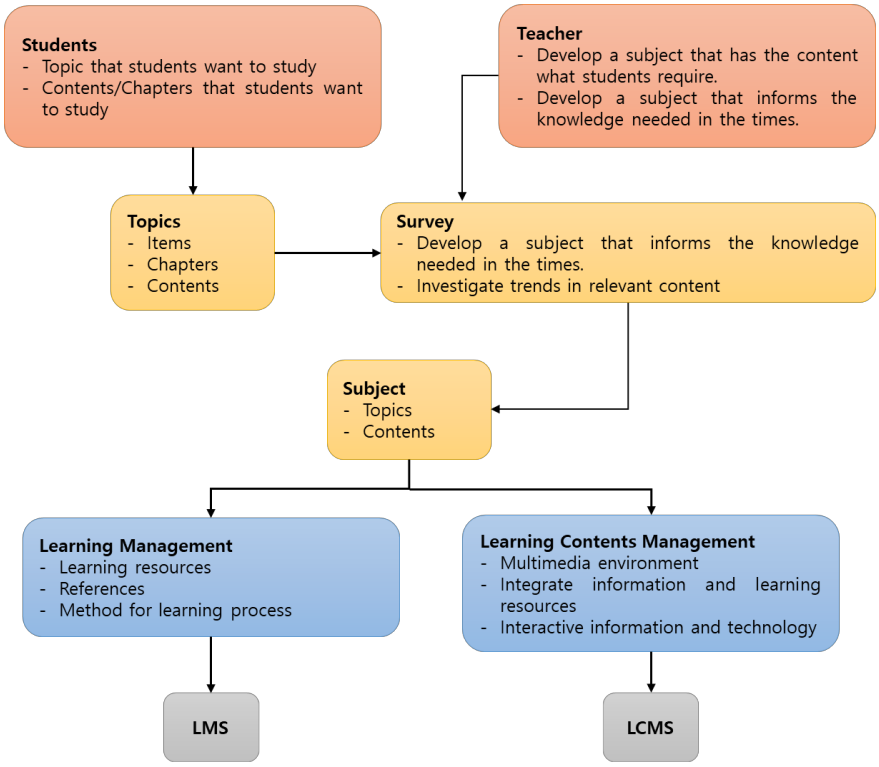


Fig. 6. A learning model that reflects student needs

5 Conclusion

This paper analyzed what students wanted to learn in class and applied it to the process of developing learning subjects. In order to know the topics what students want to study, the research was applied survey with unstructured data. In the survey result, we can see the topics that students' needs such as education, development, paradigm, artificial, computer, big data, metabus, and program. The results showed

that students wanted to be able to develop more technical subjects and take them. Therefore, this paper shows the learning model that applied their requirement. Each course indicated how the teacher should reflect it and develop a learning subject after analyzing the students' requirements.

Acknowledgement

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2020S1A5A2A01042497)

References

1. M. Cantabella et al., Analysis of student behavior in learning management systems through a Big Data framework, Future Generation Computer Systems 90, 262–272 (2019).
2. S. Bhattacharjee et al., Business and government applications of text mining & Natural Language Processing (NLP) for societal benefit: Introduction to the special issue on text mining & NLP, Decision Support Systems 162, 113867 (2022).
3. Y. Yorkovsky and I. Levenberg, Distance learning in science and mathematics - Advantages and disadvantages based on pre-service teachers' experience, Teaching and Teacher Education 120, 103883 (2022).
4. D.K. Jain et al., An Intelligent Cognitive-Inspired Computing with Big Data Analytics Framework for Sentiment Analysis and Classification, Information Processing and Management 59, 102758 (2022).
5. A.G. Sreedevi et al., Application of cognitive computing in healthcare, cybersecurity, big data and IoT: A literature review, Information Processing and Management 59, 102888 (2022).
6. F. Neves et al., Chapter 8. Assisted education: Using predictive model to avoid school dropout in e-learning system, Intelligent Systems and Learning Data Analytics in Online Education, 153-178 (2021).
7. K. Deejring, The design of web-based learning model using collaborative learning techniques and a scaffolding system to enhance learners' competency in higher education, Procedia - Social and Behavioral Sciences 116, 436-441 (2014).

A Comparative Study of Female Image in "Eouyadam" and "Yojaejii"

She Shaoshuo[1], Young-Hoon An[2] and Hwa-Young Jeong[3]

[1, 2] Department of Korean Language and Literature, Kyung Hee University, 26, Kyungheedae-ro, Dongdaemun-gu, Seoul, 02447, Republic of Korea
[3] Humanitas College, Kyung Hee University, 26, Kyungheedae-ro, Dongdaemun-gu, Seoul, 02447, Republic of Korea
sheshaoshuo@naver.com[1], yhnahn@khu.ac.kr[2], hyjeong@khu.ac.kr[3]

Abstract. In the new era, women are becoming more colorful in society and at home as women pursue equal rights and status with men and pay more and more attention to unique styles and attractions, a sign of self-awareness that has taken a step further in women's social history development.The report shows the diversity and progress of women's images in the 17th and 18th centuries, focusing on the female stories of Korean writer Yoo Mong-in's "Eouyadam" and Chinese writer Pu Song-ling's "Yojaejii“

Keywords: Eouyadam, Yojaejii, the female image, comparative study

1 Introduction

In the 21st century, many people value mental needs and human rights have been the focus in Korea. In particular, women with low social status are trying to have proper rights and equal social status with men because they have been oppressed by feudalism since ancient times. It can be felt that the female aspects of modern society are becoming rich and colorful. Not only Korea, but also China. This may show progress in the times. However, such progress did not appear suddenly, but the promotion of history and the change of society are inevitable products that have worked and appeared together. Therefore, when understanding the aspects of women in ancient times, the character and aspects of modern women are helpful in understanding the overall change and progress. It recognizes the aspects of women in ancient times, and through the specific characters of classical literature, it is easier to understand the aspects of women at that time, and the situation of women of that time can be seen from the overall perspective. Therefore, the text will study various female aspects by organizing and comparing female characters that appeared in the book, focusing on Yoo Mong-in's "Eouyadam" and Pyo Song-ryeong's "Yojaejii" in the early Qing Dynasty of China.

Yoo Mong-in is a literary man in the middle of the Joseon Dynasty, and the representative work is "Eouyadam", and Posongryeong is a famous literary man in the early Qing Dynasty of China, and the representative work is "Yojaejii." The two writers have similar backgrounds, who live at the end of feudal society and are deeply involved in Confucianism and Cheng Zhu's Philosophy the same era. In the male

perspective of this social background, these psychological activities are well represented through female characters who appeared in literary works, so the female aspects of the two works, "Eouyadam" and "Yojaejii", have something in common. Although the social environment is similar, there are differences in the female aspects created because the personal experiences of Yoo Mong-in and Posong-ryeong are different and their social status is different.

The text organizes and compares information on Yoo Mong-in and the conscription decree in a table so that personal history can be seen more clearly. In addition, information on the two works "Eooyadam" and "Yojaejii" was simply organized, and classified and organized according to the aspects of female characters that appeared in the work. The female aspect created by Yoo Mong-in and Posong-ryeong can be found through information contrast that summarizes the causes and reasons that have commonalities and differences. In addition, it is hoped that we can understand the aspects of modern women by looking at the appearance and aspects of women living in feudal society.

2 Related works

Kim Jin-sun revealed that women are in a desperate social position by exploring the relationship between women and women in "Eouyadam" and analyzing women's self-awareness and efforts to escape secular ideas. Hyun Hye-kyung tried to identify the characteristics and meaning of the various and rich shapes of women's life in "The Shape and Meaning of the Shape of Women's Life" in "Eouyadam". Gao Hai-lui revealed that women's social status is underground and they have to rely on men even after death according to the analysis of the aspects of the maiden ghost in "Eouyadam", Ko Sook-hee briefly analyzed the status and aspects of women's marriage in the traditional feudalism in the work and simply divided them into two types: "traditional women" and "progressive women.". Lee Soo-yeon studied and analyzed the types of female characters in the love story in "Yojaejii", and explored and studied four types of women according to their characteristics: "Current wife-in-law model," "Affectionatement-seeking type," "Chongmyeongjae daughter type," and "Great martial arts type". Jeon Soon-nam classified women who appeared mainly in the feudal ethics system into three types: "feudal female shape," "anti-feudal female shape," and "complex female shape," and examined women's psychology and survival under the feudal system. Wang Meng "Comparison of Night Talk in the Late Joseon Dynasty and the Women's Talk in the Novels of the Qing Dynasty" in the paper systematically sorted out and analyzed the women's talks introduced in the five late Korean night talk collections and the Qing Dynasty Notes Novel Collection. In the above study, a meaningful and rewarding study was conducted on the female aspects of "Eouyadam" and "Yojaejii". However, no detailed comparative study was conducted on the two works. Therefore, the text used big data and conducted a simple comparative study of "Eouyadam" and "Yojaejii", so that the commonalities and differences between the two works could be more intuitive and detailed.

3 Comparison of Yoo Mong-in and Po Song-ryeong's Personal History

Since literature originates from life and the writer's career has a great influence on the literary work he creates, it is necessary to understand the author's personal career first to understand the literary work.

First, let's look at the personal history of Yoo Mong-in and Po Song-ryeong.

Table 1. A list of personal details of Yoo Mong-in and Po Song-ryeong

	Yoo Mong-in	Po Song-ryeong
life time	1559-1623	1640－1715
pseudonym	Ngmun, Eudang, Ganjae, Mukhoja	Yuseon, Yucheon, Yucheon Geosa

Personal History

Age	Year	Event	Year	Event
1	1559	Born in Goheung, Jeollanam-do (noble family)	1640	Born in Shandongseong jinan Chicheon (merchant family)
18			1657	Married Miss Liu
19			1658	He passed the childbirth testand became a doctor's student. a disciple of Si Yun-jang
21			1660	Fail the imperial examination
23			1662	The first son was born.
24	1582	Pass the entrance examina-tion for a student	1663	Fail the imperial examination
26			1665	Work as a governess
30			1670	Work as personal secretary
31	1589	The first place in the imperial examination	1671	Resign from secretarial position The second son was born.

4

32			1672	Visited Laoshan Mountain Fail the imperial examination
34	1592	War Responsible for Diplomatic Operations		
35	1593	Became the prince's teacher.	1675	Fail the imperial examination The third child was born.
40			1679	A Preliminary Completion of "Yojaejii"
41			1680	Mother's dead.
48			1687	Fail the imperial examination
51	1609	Third Mission to the Ming Dynasty	1690	Fail the imperial examination
54	1612	Served as Yejochampan and Ijochampan.		
55	1622	A compilation of "Eouyadam"		
56	1623	Resign from office, Be sentenced to death		
63			1702	Fail the imperial examination
72			1711	Took the imperial examination and became a preparatory student.
74			1713	Wife is dead.
76			1715	Dead.

Through this table, more clearly, the personal history of Yoo Mong-in and Posong-ryeong is compared. As you know, passing past exams and serving in government posts in ancient times was the only way for writers to participate, so both Yoo Mong-in and Po Song-ryeong took the past exams. However, because of this, the fate of Yoo Mong-in and Posong-ryeong unfolded completely differently. Yoo Mong-in passed

the examination for the first birthplace and Jinsashi at the age of 24, and passed the examination as a manor in the department of Jeunggwang Literature at the age of 31. It was followed by a smooth entry into government office, which became important. On the other hand, at the age of 19, Fosongryeong, a Cheongdae literate, took the examination for his younger brother's poem and received the first prize, became a doctor's student and a disciple of Si Yoon-jang. For a 19-year-old child, this was a very high evaluation and good result, but after that, he took the exam eight times in the past for more than 40 years, but failed and passed the examination until the age of 72 and became a craftsman.

As mentioned earlier, in ancient times, the past was the only way for writers. Yoo Mong-in is from an aristocratic family and is a great-great-grandchild of Yu Yi, and his grandfather is Yu Chung-gwan, Sagan. And my father is Yu Taeng, a housewife, and my mother is the daughter of Cham Bong-min. Yoo Mong-in, a native of an aristocratic family, passed the examination in the past and became a government official, and his job was smooth, so he had no worries about living.

However, the conscription is the opposite of Yoo Mong-in. Fosongryeong was from a scholar's family, and both her great-grandfather passed the floodgates, but her grandfather failed to pass the floodgates, so her family began to decline, and her father also failed to pass the floodgates, so she made a lot of money by doing business. However, in the middle age, as he believed in Buddhism and stopped doing business, the situation gradually became difficult. Po Song-ryeong was born into this family. When the family situation became difficult, Fosongryeong's father took the role of a teacher and taught Fosongryeong's knowledge, and from an early age, Fosongryeong listened to his father's merchant thoughts and Buddhist ideas.

Due to poor family circumstances, Posongryeong had no worries about living in her childhood, but because she devoted her body and mind to creating literature in the past, she had to worry about her livelihood because she had no income source, passed the exam several times, and had to feed her wife and children, and eventually became a Seodang teacher.

In addition, this table is prepared based on age comparison, and if you look at the table, the difference in fate between Yoo Mong-in and Posong-ryeong may seem more intuitive. First of all, Yoo Mong-in and Posong-ryeong both took past tests at the age of 24, but the results were completely different, and Yoo Mong-in passed the test, but Posong-ryeong failed. At the age of 35, Yoo Mong-in had already become a government official, but Posong-ryeong failed again in the past. In addition, at the age of 51, Yoo Mong-in was dispatched to the Ming Dynasty as a Seongjeolsa and a private teacher, but Posongryeong still failed while fighting the past system.

Through this clear contrast, it can be seen that the difference between the fate of Yumongin and Posongryeong is very large. The difference in fate has become one of the many reasons why the two expressed their thoughts through the female aspects that appeared in literary works.

On the other hand, as shown in the conclusion of the Yumongin, Yumongin was a noble and had no financial worries, but he was involved in partisan battles and had a lot of heartache and despondency. Posongryeong suffered from difficulties in life as she failed the past exams several times, and her mind was filled with disappointment

and resentment for the world. Based on this, it is possible to understand that the thoughts and emotions contained in the two literary works have many things in common.

4 Comparison of "Eouyadam" and "Yojaejii"

In this part, information on the two works is organized and contrasted from various angles, such as the period of creation, the background of the times, and mainstream social ideas. In addition, the aspects of women appearing in the work are classified and prepared by organizing the number of copies. Through this table, you can examine the characteristics of women in the mid-Joseon Dynasty and early Qing Dynasty, and feel the diversity of ancient women in the late feudal society and the progressiveness of the early implementation of women.

Table 2. Female Characters in "Eouyadam" and "Yojaejii"

Work	Eouyadam	Yojaejii
the year of creation	1618-1622	About 1672-1710
Time	the mid-Joseon Period	the early Qing Dynasty
mainstream social thought	Confucianism, Jeongjuri, Jeongjuri	
bibliography	Wan Zongqi Edition, Stone Pillow Publishing, 2006.	Chinese bookshop, 2015.
Total number of articles	522	491
Number of female content	68	181

	Eouyadam									
	human beings					inhumanity total				
female image	virtuous woman	clever woman	woman of strong character	chivalrous woman	defile one's chastity	bad wife	Other types	fairy	Foxes and ghosts.	
Number	1	16	6	4	12	1	20	1	7	68

specific gravity	1%	24%	9%	6%	18%	1%	29%	1%	11%	100%
	Yojaejii									
Number	11	15	11	4	14	8	22	22	70	181
specific gravity	6%	8%	9%	2%	7%	4%	12%	12%	38%	100%

The creation period of "Eouyadam" and "Yojaejii" was in the 17th and 18th centuries, the last feudal period in the history of both Korea and China, and social economy, ideology, and culture changed significantly at the end of feudal society. Social-led ideas were still studying abroad, but Jeongjuri and Yangmyeonghak gradually became mainstream ideas.

At a time when social ideas clashed, new changes began to appear in the lives of women who were weighed down by the feudal social system, and self-consciousness became more awakened and women's patterns diversified.

Yoo Mong-in and Posong-ryeong depicted various female aspects in the work, and according to their characteristics, women in the work were classified into types with several social representations. In other words, the human part contains Hyunbuckam, Wisdom Story, Yeolnyeo Story, Heopnyeo Story, Destruction Story, and Akcheom Story, and Sinseondam, Fox, and Ghost Story in the second-class part. Through various female aspects, the diversity and abundance of female aspects of society at the time can be seen. This can also be said to be a common feature of the female aspects created by Yoo Mong-in and Posong-ryeong.

Another common feature is that looking at this table, it can be seen that wisdom accounts for a large proportion of the human female aspect. This is deeply related to the times. At the end of the feudal period, Yoo Mong-in and Posong-ryeong, who were greatly influenced by Confucian ideas and Jeong Ju-ri ideas, represented the character and aspects that women should have as women in the eyes of most men at that time. In feudal society, men's thoughts that women are pretty, nice, wise, good at housework, managing large families, and always being able to help their husbands with good strategies when needed were well expressed through women of Wisdom.

In addition to the traditional female aspects such as virtuous woman, clever woman, and woman of strong character, there were also female aspects such as chivalrous woman, defile one's chastity, and bad wife, which deviate from the traditional female aspect and rebel against the oppression of women's human rights and nature in the feudal society. This aspect of women can be seen as a progressive awakening of women's will to traditional feudal society.

On the other hand, this part of Foxes and ghosts stories deserves attention. Foxes and ghosts stories accounted for 11% in Eouyadam, but 38% in "Yojaejii". This is also the biggest difference that can be found through this table. Exploring the cause is related to the writer's origin and personal career. It is more urgent and necessary to express feelings of dissatisfaction, anger, and disappointment about reality that have

long been built through surreal characters such as foxes and ghosts, even though they were born in poverty and failed in the past.

5 Conclusion

In this text, Yoo Mong-in and Posong-ryeong also worked on the women's aspects in "Eouyadam" and "Yojaejii" which were composed of two works, briefly organized and studied in a table using digital. In addition, the commonalities and differences of female aspects were analyzed by combining the author's personal history and the background of the times. In particular, in Chapter 3, the personal history of Yoo Mong-in and Po song-ryeong, the creative background of the work, and the female images of the works were organized in a table using big data, so that comparative research on the female aspects of both Korea and China could be conducted more coherently and intuitively. In addition, after calculating the number and ratio by clearly listing various female aspects shown in "Eooyadam" and "Yojaeji," you can further highlight the diversity and abundance of female aspects and feel the progressiveness of women's times.

References

1. Yoo Mong-in, "Eooh Yadam", stone pillow (2006)
2. Posongryeong, Eocheonji Remarks, Yojaejii, Beijing: China Bookstore (2015)
3. Ko Sook-hee, "The Study of Women in Yojaejii ", Master's thesis at Sookmyung Women's University (1995)
4. Lee Soo-yeon, "A Study on the Types of Female Characters in Yojaejii Love Story", Kyunghee University's Master's Degree thesis (2005)
5. Jeon Soon-nam, "The Study of Women in Yojaejii, Master's thesis at Yeungnam University (2011)
6. Kim Jin-sun, “A Study on the Existence of Women in Yadamjip”, A Master's thesis at Kyung Hee University (2006)
7. Gao Hai-lui, "The Descriptive Patterns and Features of Ghosts in Eouyadam”, Master's thesis at Pusan National University (2020)
8. Wang Mong, "Comparison of Yadam in the Late Joseon Dynasty and Women's Stories in Written Fiction in Qing Dynasty", Korea University's Ph.D. thesis, (2017)
9. Hye-kyung Hyun, "The Pattern and Meaning of the Shape of Women's Life in Eouyadam”, Korean Classical Women's Literature Research, Korean Classical Women's Literature Society (2004)

A Study on Data Mining for Type of Korean Painting Poetry

Park Haeyoung[1][0000-0002-0161-8465], An Younghoon[2][0000-0002-3012-9724], Jeong Hwayoung[3][0000-0002-5017-934X]

[1, 3] Humanitas College, Kyung Hee University, 26, Kyungheedae-ro, Dongdaemun-gu, Seoul, Republic of Korea
[2] Department of Korean Language and Literature, Kyung Hee University, 26, Kyungheedae-ro, Dongdaemun-gu, Seoul, Republic of Korea

hy000p@khu.ac.kr
yhnahn@khu.ac.kr
hyjeong@khu.ac.kr

Abstract. This study introduced the work of analyzing the meaning by using the computer information processing method. We extracted the painting poetry of Mukjukdo (Bamboo Paintings) in the early Joseon Dynasty from the *Hanguk Munjip Chonggan* (Korean Literary Collections in Classical Chinese) of *Hanguk Gojeon Jonghap DB* (Korea Classics DB). Through the data mining method, we divided the types by extracting and cataloging the painting poetry from the Korean classical literature. Then, the scope was narrowed down to painting poetry of Mukjukdo in the early Joseon Dynasty, and the text was analyzed in units of syllable corpus.

Keywords: Data mining, Painting poetry, Bamboo paintings, Extracting, Cataloguing, Type Classification

1 Introduction

This study introduced the work of analyzing the meaning by using the computer information processing method. We extracted the painting poetry of Mukjukdo (Bamboo Paintings) in the early Joseon Dynasty from the *Hanguk Munjip Chonggan* (Korean Literary Collections in Classical Chinese) of *Hanguk Gojeon Jonghap DB* (Korea Classics DB). This is one of the basic tasks of research on Hansi (poems in the Chinese style), which has subdivided types of Hansi according to their material. This study presents expression techniques and meanings objectively by converting them into numerical values by using a computer information processing method. Traditional research on Hansi has mainly relied on researchers' intuition, however, this digital data analysis, which is a quantitative method, can complement the traditional method.

2 Research Methods

2.1 Data

The basic data for this study is *Hanguk Munjip Chonggan*; the Database of Hanguk Gojeon Beonyeokwon (Institute for the Translation of Korean Classics). It is called "Hanguk Gojeon Jonghap DB". Hanguk Gojeon Beonyeokwon organizes and translates Korean classics into this database and then digitally converts them and discloses them to the public. This data is provided in XML format in the form of openAPI[1]. This data is categorized by author, book name, style, title, original text, year of publication, etc. Based on this, the painting data from the early Joseon Dynasty were extracted and analyzed.

2.2 Data Mining Techniques

Data mining is a technique of extracting useful information by analyzing statistical patterns, rules, and relationships in large amounts of data[2]. Today, it is used in various fields such as computerization, statistics, and management. We also intend to apply this technology to the study of Korean classical liteture. Currently, 142 anthologies from the early Joseon Dynasty are included in the *Hanguk Munjip Chonggan*. The primary data mining is to select the painting poetry based on this database and to classify them by their type. The secondary mining is to analyze the text by narrowing the scope[3] to painting poetry of Mukjukdo in the early Joseon Dynasty.

3 Result

3.1 Extracting and Cataloguing

We chose the poetry from the style category at Hanguk Gojeon Jonghap DB, and selected poems with characters in the title, such as ‘do (picture)’, ‘hwa (painting)’, ‘je (mention)’, ‘muk (ink)’, ‘hoe (drawing)’, ‘sa (drawing)’, ‘byeong (folding screen), ‘jok (hanging scroll)’, ‘cheop (album)’, and ‘chuk (scroll)’. After that, we compare the title and the content of each poem to determine whether the poem was painting poetry or not. Through this, it was found that there are a total of 842 poetry poems currently included in *Hanguk Munjip Chonggan*.

3.2 Type Classification

Based on the previously extracted lists, the painting poetry of the early Joseon Dynasty was classified by type based on the subject matter of the painting[4]. This classification is possible because the subject matter of the picture is presented as the name of each object before and after 'do' or 'hwa', which usually means a picture. For example, the titles of most painting poems of Mukjukdo appear in expressions such as “Mukjukdo” or“Jejuk (reciting a bamboo)”. The figure below is the result of categorizing

Mukjukdo painting poems using the TOPIC MAP based on the list of painting poems in the early Joseon Dynasty and the types of paintings.

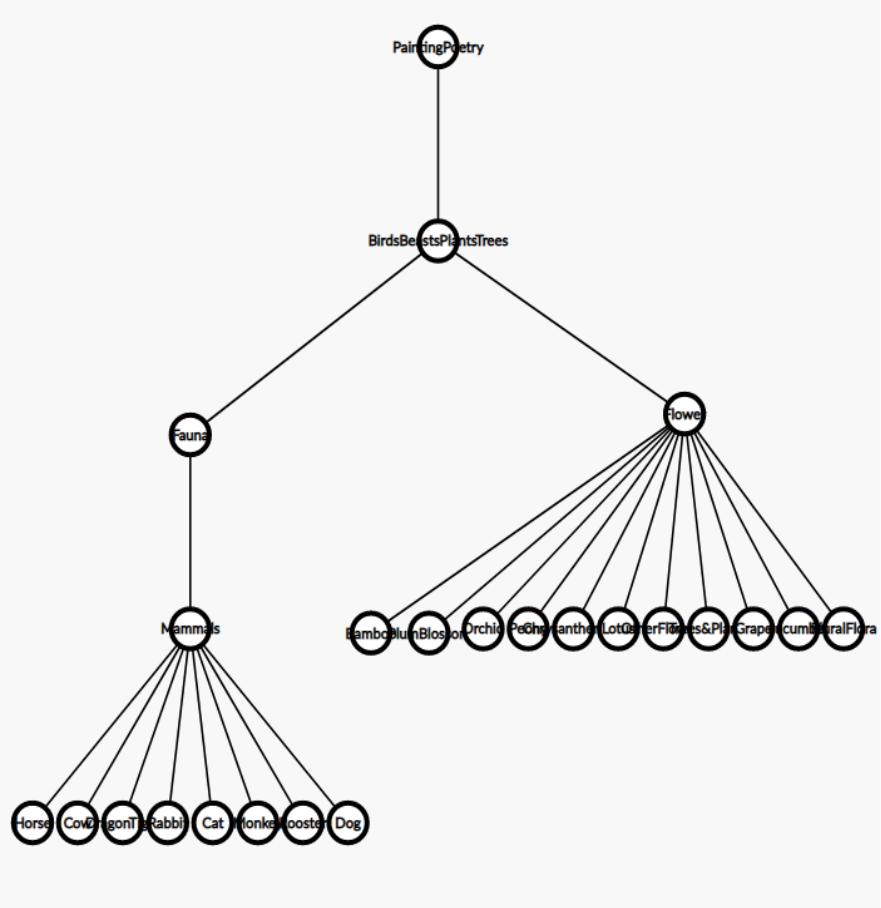


Fig. 1. A Type classification of Mukjukdo painting poetry using TOPIC MAP

3.3 Syllable Unit Segments and Statistics

Next, we analyzed the text data of the painting poetry of Mukjukdo in the early Joseon Dynasty. In order to interpret the poem[5], it is convenient to divide it into morphemes, the smallest unit of meaning. Since Hansi is composed of Chinese characters, it can be divided into morphemes up to one syllable unit. In addition, the units of these syllables are basically combined into five and seven words to form a row. Therefore, we set one row as the basic unit. Among the original text of painting poetry, o'eon (five-character) was divided from 1 to 5, and chil'eon (seven-character) was divided from 1 to 7 syllables. In succession, per line, o'eon generated five corpus of one syllable, two syllables, three syllables, four syllables, and five syllables, and chil'eon generated

seven corpus of one syllable, two syllables, three syllables, four syllables, five syllables, six syllables, and seven syllables.

Table 1. The corpus unit of o'eon and chil'eon poem

If 1 line of o'eon poem is assumed to "1 2 3 4 5", the corpus unit	If 1 line of chil'eon poem is assumed to "1 2 3 4 5 6 7", the corpus unit
1 syllable corpus(5): 1/2/3/4/5	1 syllable corpus (7): 1/2/3/4/5/6/7
2 syllable corpus(4): 12/23/34/45	2 syllable corpus (6): 12/23/34/45/56/67
3 syllable corpus(3): 123/234/345	3 syllable corpus (5): 123/234/345/456/567
4 syllable corpus(2): 1234/2345	4 syllable corpus (4): 1234/2345/3456/4567
5 syllable corpus(1): 12345	5 syllable corpus (3): 12345/23456/34567

The following is the result of dividing the original data of Mukjukdo painting poetry by syllables and calculating the frequency. This is the figure measured using the JAVA program to see how many syllable corpus units are repeated in the actual painting poetry text data. The numerical value of 1-5 syllable corpus is the sum of the data for each syllable of o'eon and chil'eon, and the numerical values of 6-7 syllables is for chil'eon only. In these values, only words from 1 to 3 syllables were actually valid, and all 4-7 syllables were meaningless because the word combination was under twice.

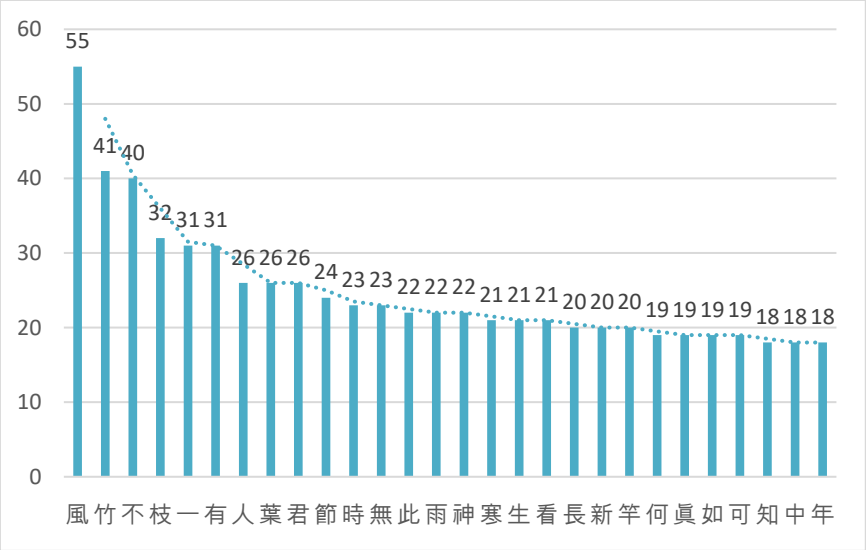


Fig. 2. The frequency of one syllable corpus in Mukjukdo painting poetry

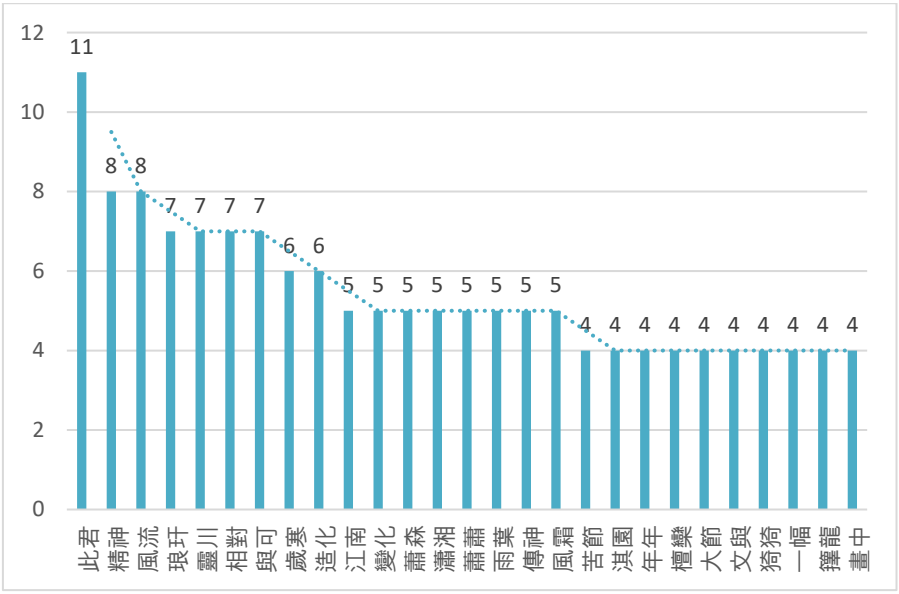


Fig. 3. The frequency of two syllable corpus in Mukjukdo painting poetry

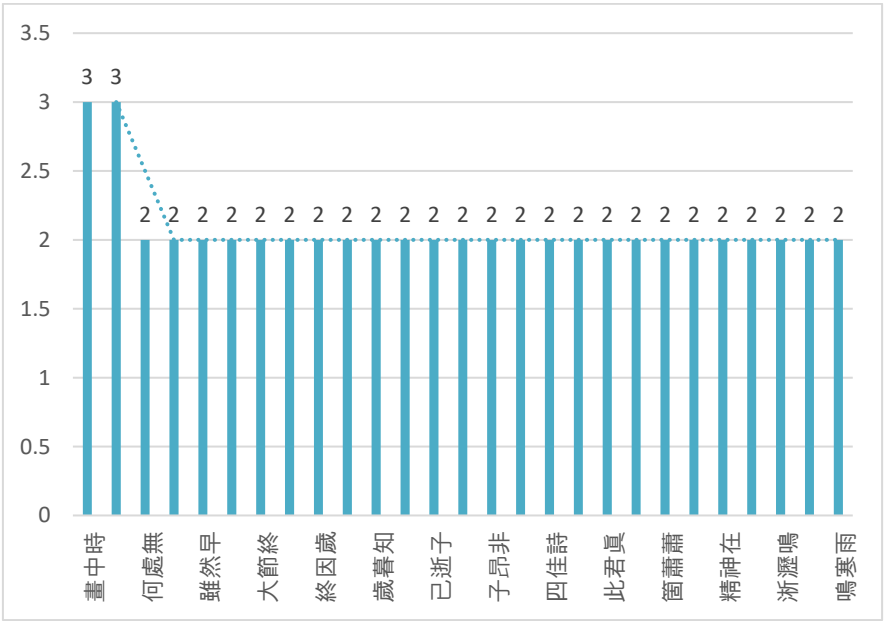


Fig. 4. The frequency of three syllable corpus in Mukjukdo painting poetry

3.4 Analysis of the Meaning of Corpus

The most commonly used one-syllable word in painting poetry of Mukjukdo is “poong (wind: 55 times)”. “poong” is a combination of words in the order of “poong ryu (a taste for the arts: 8 times)”, “poong sang (wind and frost: 5 times)”, “Chun poong(the spring breeze: 3 times)” etc in 2 syllables. Poong ryu is presented when referring to the wonderful scenery of bamboo. Pung sang is wind and frost, which gives bamboo trials and tribulations. Chun poong appears when describing a situation in which bamboo shoots sprout.

Next is “juk (a bamboo: 41 times)”, but "juk" refers to bamboo itself, so it is not a discriminating result. And “bul (no: 40 times)”, which belongs to an adverb that represents negativity in Chinese characters, is not very meaningful in single syllables, but further emphasizes its meaning when combined with other words. And “ji (Branch: 32 times)”, which is connected in two syllables: “poong ji (a branch swaying in the wind: 2 times)", “so ji(a slender branch: 2 times). It excludes the numerical value that does not form a special meaning, such as juk, bul, ji.

The next significant figure is “goon (a man of virtue: 26 times)”. “goon” is a combination of words in the order of “cha goon (this man of virtue: 11 times)”, “goon jin (a man of virtue is truly~ : 3 times)”, “goon dok (a man of virtue alone~ : 2 times)” etc in 2 syllables. And “jeol” is 24 times, it means principles. “jeol” is a combination of words in the order of “go jeol(a distressed principles: 4 times)”, “no jeol (mature principles: 2 times)”, “jik jeol (a straight principles: 2 times)” , etc., in 2 syllables. The jeol represents the principles of bamboo, and the attitude of keeping faith firmly is matched with "go" and "no”, to reinforce the unchanging properties.

4 Conclusion

The data figures of corpus by syllable in Mukjukdo painting poetry are in line the symbolic meanings of bamboo in the early Joseon Dynasty. The symbolic meaning of Mukjukdo[6] is as follows. As an object, a bamboo was personified as the ideal existence of a god dragon and a Junzi. As an event, the virtual space of bamboo about the bamboo of Two queens and Qu Yuan was realized as the sorrowful and unworldly image.

Through this, the measurement of the frequency by dividing the original text of Mukjukdo painting poetry by syllables and setting it as a corpus composition unit is meaningful in revealing the symbolic meaning of bamboo in painting poetry. Conversely, the symbols of objects in various types of painting poetry can be demonstrated and shown with an objective indicator of language.

References

1. Lee Byoungchan, Min Kyoungju: A Study on Visualization of the Analysis between the Collectionss of Korean Literature in Korea Classic DB: Hanguk Gojeon Beonyeokwon :national culture 57, 5-32, (2021).

2. ko.wikipedia.org/wiki/Data Mining, last accessed 2022/12/01
3. Kim Dongkeon, Jeong Hwayoung: A Study of Computational Literature Analysis based Classification for a Pairwise Comparison by Contents Similarity in a section of Tokkijeon, 'Fish Tribe Conference': The Korea Contents Association: The Journal of the Korea Contents Association 22, 15-25, (2022).
4. Park Haeyoung: A study on the painting poetry of the former part of the Joseon Dynasty: Kyunghee University Graduate school, 19-34, (2021).
5. Lee Byoungchan: A Study on the Construction and utilization of Korean Chinese poetry corpus: geun-yeoghanmunhaghoe: hanmunhaknonjib 53, 153-177, (2019).
6. Park Haeyoung: The symbol of the painting poetry about a bamboo painting in the former part of Joseon Dynasty: Society of Yol-Sang Academ: Yeol-sang Journal of Classical Studies 73, 123-151, (2021).

A phonetic investigation of Korean monophthong vowels by Vietnamese female speakers

Juhee Lee

Kyung Hee University, Republic of Korea

Abstract. In this study, we discuss the phonetic characteristics of Korean monophthong vowels produced by Vietnamese female speakers at the beginner's level. Unlike Korean female speakers, the quantitative results for /e/ and /ɛ/ show that there was a statistically significant difference in F1 values ($p < 0.05$) as well as F2 values ($p < 0.05$) in the production of vowels by Vietnamese learners of Korean. Therefore, the difference in height between the two vowels was contrastive by Vietnamese learners of Korean differently from native Korean female speakers. On the other hand, the back vowels /u/ and /o/ produced by female Vietnamese learners of Korean had a statistically significant difference in both F1 and F2 ($p < 0.05$) values. Therefore, female Vietnamese learners' pronunciation was contrastive by the difference in tongue height (F1) as well as the position of the tongue (F2) in the production of Korean rounded back vowels /u/ and /o/, whereas these two vowels were contrastive only by the difference in tongue position (F2) in native Korean female speakers. Therefore, all pairs of monophthong vowels produced by Vietnamese female speakers formed in opposition to each other, while Korean native speakers formed a seven-vowel system due to the merger of /e/ and /ɛ/. Moreover, Vietnamese female speaker's Euclidean distance between /u/ and /o/ was shorter (88.4) than Korean female speaker's (146.5). Thus, it can be also argued that the pronunciation of the Korean vowels /u/ and /o/ produced by Vietnamese female learners of Korean are considerably different to the production of Korean rounded back vowels.

Keywords: formants, Korean vowels, Vietnamese, beginner level, phonetics.

1 Introduction

In this paper, we discuss the acoustic and phonetic characteristics of Korean monophthong vowels produced by northern Vietnamese female learners of Korean at the beginner's level, and we identify the differences compared to those of Korean female speakers. Since Vietnam has a long topography from north to south, the pronunciation of vowels varies across the north, central, and south, so the experiment was conducted based on the northern dialect. The monophthong vowels of Vietnamese are nine vowels in the northern Hanoi dialect, which is considered as standard (Han 1966, Kirby 2011, Đào and Nguyễn 2018). Let us consider the following in Figure 1:

2

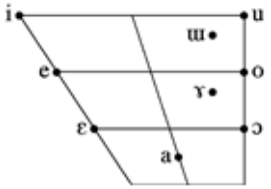


Figure 1. Vietnam monophthong vowels (Kirby 2011: 384)

In Figure 1, Vietnamese vowels between /e/ and /ɛ/ are distinguishable, while Korean forms a seven-vowel system (/i, e, ɯ, ʌ, a, u, o/) that is based on the current spoken language due to the merger of /e/ and /ɛ/ (Yoon and Kang 2014, Lee et al. 2016, Kang and Kong 2016). Thus, we discuss the phonetic characteristics of Korean monophthong vowels produced by Vietnamese female speakers at the early stage of learning.

2 Methods

2.1 Participants

A total of 22 subjects participated in the experiment: 12 Vietnamese female learners of Korean (beginner level: less than 6 months of learning experience; average age: 21.4) from the northern part of Vietnam, including Hanoi. All of them were attending a Korean language school at the universities in Seoul. 10 Korean female speakers (average age: 21.1) who indicated their birthplaces and residences were in Seoul and Gyeonggi. All of them are students attending universities in Seoul. A predetermined honorarium was paid to all subjects.

2.2 Procedure

The words used in the experiment were 24 nonce words in Korean with two syllables composed of eight monophthong vowels. Due to the merging of /e/ and /ɛ/, the monophthong vowels of Seoul Korean form a seven-vowel system based on the production level (Yoon and Kang 2014, Lee et al. 2016, Kang and Kong 2016). However, we included nonce words with eight vowels to examine how Vietnamese speakers pronounced /e/ and /ɛ/, respectively, and how they differed from subjects who speak Korean as their mother tongue. The nonce words in Table 1 are in the form of "V_1+CV_2" so that eight monophthong vowels (V_1) are the word-initial position, while the following consonants are Korean plosives /p, t, k/.

Table 1. Nonce words.

	_ka	_ta	_pa
i	ika	ita	ipa
e	eka	eta	epa
ɛ	ɛka	ɛta	ɛpa
ɯ	ɯka	ɯta	ɯpa
ʌ	ʌka	ʌta	ʌpa

a	aka	ata	apa
u	uka	uta	upa
o	oka	ota	opa

The experiment was conducted in a manner that the subjects repeatedly read the 24 nonce words presented in Table 1 three times, and the words were read with the carrier sentence 'This is _____'. In this way, 72 pieces of data were collected per subject, and the total number of data obtained through each group of Korean (10 people × 24 words × 3 repetitions) and Vietnamese (12 people × 24 words × 3 repetitions) was 1,584. The recording was performed using an LG Gram laptop, which is included in the public software Praat version 6.2.23. A SONY ECM-LV1 pin microphone was attached to the upper body of the subject and connected to the laptop. The recording took place in a quiet space on the campus without noise.

For the data analysis, using Praat version 6.2.23, we checked the spectrogram of all vowels and set the section with the least variation in formant as the stable section. Since all the subjects were women, the maximum formant value was set to 5,500 (Hz), the number of formants was set to five, and the window was set to 25 (ms). The measurement of the formant value was calculated by finding the mid points of the stable section found in the collection. Through this process, an analysis was conducted on a total of 1,584 tokens. Formant values were summarized using Excel, and statistical analysis was performed using IBM's statistical analysis program SPSS 26 to understand the statistical significance.

3 Results

3.1 Production

To discuss the acoustic and phonetic characteristics of Korean monophthong vowels produced by Korean female speakers, we present the formant values in Table 2. To judge the significant differences among Korean vowels, the results of repeated-measures ANOVA followed by Bonferroni post-hoc analysis are reflected.

Table 2. Korean female speakers estimated mean in F1/ F2 (Hz)

Vowel (F1)	Estimated mean	Standard error	95 % Confidence interval	
			minimum	maximum
i	433.5	7.1	419.5	447.5
e	611.1	4.0	603.3	619.0
ɛ	618.8	4.1	610.6	627.0
ɯ	449.2	5.5	438.3	460.1
ʌ	603.1	5.8	591.5	614.7
a	931.0	6.8	917.6	944.4
u	420.0	4.8	410.5	429.5
o	418.0	4.7	408.6	427.4

4

Vowel (F2)	Estimated mean	Standard errors	95 % confidence interval	
			minimum	maximum
i	2759.7	61.9	2636.8	2882.6
e	2210.9	44.0	2123.7	2298.2
ɛ	2229.6	43.1	2144.0	2315.2
ɯ	1565.5	29.7	1506.5	1624.4
ʌ	932.3	11.1	910.3	954.3
a	1424.0	14.6	1395.0	1453.0
u	854.2	15.7	822.9	885.4
o	715.2	11.8	691.8	738.6

In Table 2, in F1 values, there is no significant difference between /e/ and /ɛ/ ($p > 0.05$) and /u/ and /o/ ($p > 0.05$). On the other hand, in F2 values, there is no significant difference between /e/ and /ɛ/ ($p > 0.05$), but contrast is maintained between /u/ and /o/ ($p < 0.05$) in the position of the tongue in F2 values. Based on the estimated mean in Table 2, we present the Korean female speakers' vowel production in Figure 2.

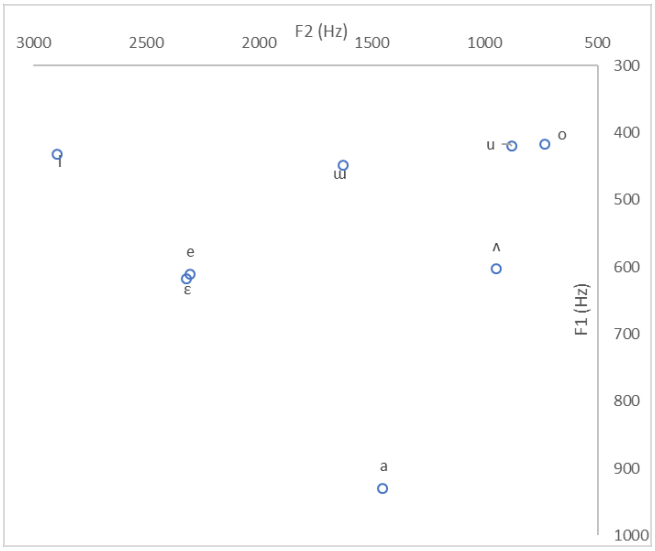


Figure 2. Korean female speakers Koran vowel production (Hz)

In Figure 2, the vowel diagram shows that the traditional Korean phonological category (the so-called eight-vowel system) does not match the production of Korean female speakers. However, unlike in Korean, the Vietnamese vowel system distinguishes front vowels between /e/ and /ɛ/, as shown in Figure 1. It is unclear whether Vietnamese female speakers pronounce these two vowels differently, or simply merge them in the same way as Korean speakers. Thus, we investigated the acoustic and phonetic characteristics of Korean monophthong vowels produced by Vietnamese female learners of Korean at the beginner's level. The results of repeated-measures ANOVA followed by Bonferroni post-hoc analysis are also reflected in this analysis.

Table 3. Vietnam female speakers (beginner) F1/ F2 values (Hz)

Vowel (F1)	Estimated mean	Standard errors	95 % confidence interval	
			minimum	maximum
i	425.6	5.8	414.0	437.1
e	524.4	7.9	508.7	540.0
ɛ	573.8	9.1	555.8	591.8
ɯ	446.5	6.2	434.3	458.7
ʌ	708.7	11.6	685.7	731.7
a	912.7	9.5	893.9	931.5
u	404.5	5.8	393.0	416.0
o	480.4	8.5	463.6	497.2
Vowel (F2)	Estimated mean	Standard errors	95 % confidence interval	
			minimum	maximum
i	2875.0	14.5	2846.1	2903.8
e	2569.7	18.6	2532.7	2606.6
ɛ	2618.3	18.4	2581.8	2654.8
ɯ	1565.6	15.6	1534.6	1596.6
ʌ	1179.0	9.501	1160.2	1197.9
a	1665.3	12.0	1641.4	1689.1
u	814.1	9.7	794.9	833.3
o	859.5	8.4	842.8	876.2

Unlike Korean female vowel production, the quantitative analysis of F1 and F2 values in Table 3 shows that all eight vowels are categorically significant in F1 and F2 values ($p < 0.05$). Based on the estimated mean in Table 3, we present the diagram for Vietnamese female learners' vowel production.

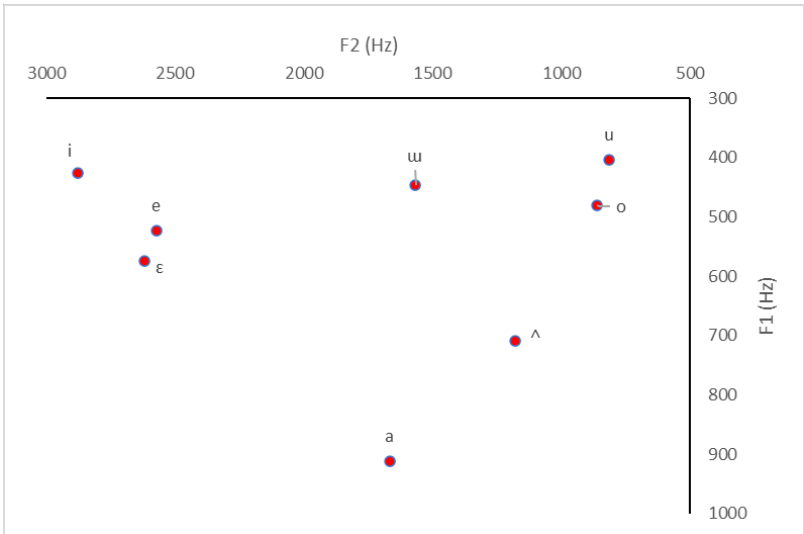


Figure 3. Korean vowel production of Vietnamese female learners' (beginner) (Hz)

In Figure 3, the distance between each vowel is statistically contrastive among Vietnamese female speakers. Therefore, we argue that the native (Vietnamese) phonological category, as shown in Figure 1, is reflected in the production of Korean vowels.

3.2 Euclidean distance

Based on the analysis of the F1 and F2 values, the distance between the two points corresponding to each vowel was obtained by applying the following formula to mathematically determine the distance between the vowels. Euclidean distance is a technique for calculating the distance between two points, and when two points have coordinates of (P_1, P_2, P_3…P_n) and (Q_1, Q_2, Q_3…Q_n), the distance between two points is expressed by the Euclidean distance formula as follows.

$$d(p,q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2}$$

The distance between two points (two paired vowels) can be calculated using the formula to measure the two distances. The Euclidean distance between each vowel is shown as in Table 4.

Table 4. KF/ VF (beginner) speakers' Euclidean distance for Korean vowels

KF	ΔF1	ΔF2	ED	**VF**	ΔF1	ΔF2	ED
i-e	170.5	589.7	613.9	i-e	98.4	305.0	320.5
e-ɛ	**7.7**	**18.2**	**19.8**	**e-ɛ**	**49.8**	**48.3**	**69.4**
ɛ-a	338.9	872.1	926.0	ɛ-a	338.9	953.0	1011.5
i-ɯ	15.7	1267.7	1267.8	i-ɯ	20.9	1309.4	1309.6
ɯ-u	29.2	746.7	747.3	ɯ-u	42.0	751.5	752.7
ɯ-o	31.2	893.2	893.7	ɯ-o	33.9	706.1	999.2
ɯ-ʌ	153.9	676.4	693.7	ɯ-ʌ	262.2	386.6	467.1
ɯ-a	175.9	481.8	512.9	ɯ-a	466.2	99.7	457.2
u-o	**2.0**	**146.5**	**146.5**	**u-o**	**75.9**	**45.4**	**88.4**
o-ʌ	185.1	216.8	285.1	o-ʌ	228.3	319.5	392.7
ʌ-a	327.9	500.5	598.3	ʌ-a	204.0	486.3	527.4

In Korean speaker's production, the Euclidean distance between the vowels /e/ and /ɛ/ was shorter (/e : ɛ/, 19.8) than Vietnamese speaker's production. The difference in tongue height (F1) between these vowels was not significant for Korean female speakers ($p > 0.05$), while these two vowel categories were statistically contrastive in Vietnamese learners' production ($p < 0.05$). For the case of the rounded back vowel

/u/ and /o/, Korean production of these vowels' distance is short in F1 (/u : o /, 2.0) but the distance was maintained in F2 (/u : o/, 146.5). Thus, the Euclidean distance between /u/ and /o/ is 146.5. However, the Euclidean distance for the /u : o/ in Vietnamese learners' production was shorter than the Korean realization of /u : o/.

4 Discussion

The results of quantitative formants for /e/ and /ɛ/ vowels show that there was a statistically significant difference in F1 as well as in F2 values ($p < 0.05$) in the production of Vietnamese learners of Korean. Therefore, the difference in height between the two vowels was contrastive by Vietnamese speakers differently from native Korean female speakers. The back vowels /u/ and /o/ produced by female Vietnamese speakers had a statistically significant difference in F1 value ($p < 0.05$), but there was no statistically significant difference in F2 value ($p > 0.05$). Therefore, female Vietnamese speakers' production was contrastive by the difference in tongue height (F1), whereas the distance was maintained only by the difference in tongue position (F2) in native Korean speakers. Therefore, all pairs of monophthong vowels produced by Vietnamese female speakers formed an opposition to each other. Interestingly, Vietnamese female speakers' Euclidean distance between /u/ and /o/ was shorter (88.4) than Korean female speakers (146.5). Thus, statistical formant analysis concluded that the production of the Korean vowels /u/ and /o/ produced by Vietnamese female learners have considerable phonetic differences compared Korean speakers, as shown in Figure 4.

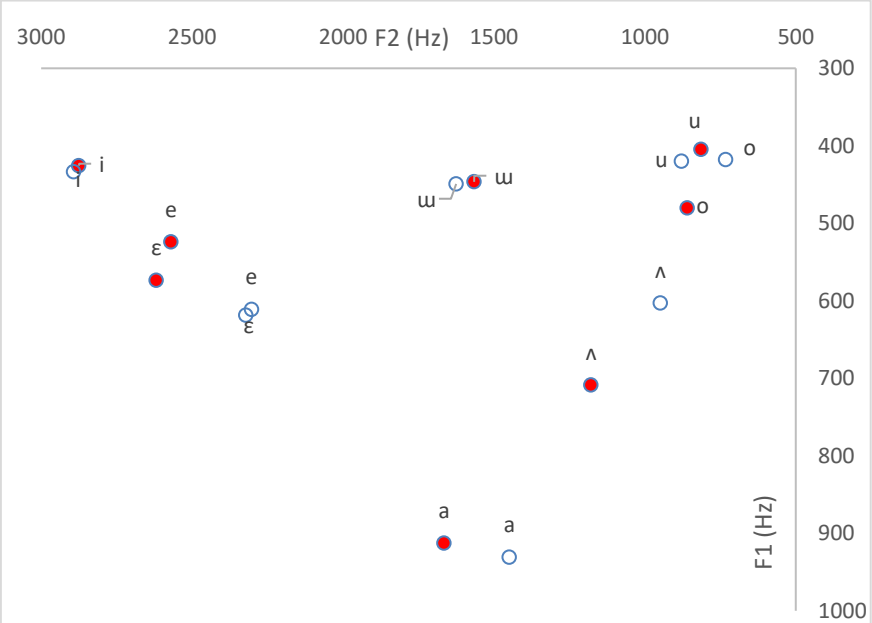


Figure 4. Korean vowels produced by KF (white dot) and VF (beginner, red dot)

8

References

1. Đào, Đích Mục and Anh-Thư T. Nguyễn. “L1 Korean vocalic transfer in adult L2 Korean learners’production of Vietnamese monophthong vowels,” Asian-Pacific Journal of Second and Foreign Language Education 3(13), 1-20 (2018).
2. Han, Mieko S. “Vietnamese vowels (Studies in the Phonology of Asian Languages 4),” Los Angeles, CA: University of Southern California, Acoustic Phonetics Research Laboratory. (1966).
3. Kang, Jieun and Eun Jong Kong. “Static and dynamic spectral properties of the monophthong vowels in Seoul Korean,” Phonetics and Speech Sciences 8(4), 39-47 (2016).
4. Kirby, James P “Illustration of the IPA: Vietnamese (Hanoi Vietnamese),” Journal of the International Phonetic Association 41(3), 381-392 (2011).
5. Lee, Juhee, Kyuchul Yoon, and Koonhyuk Byun. “A study of vowel shifts in Seoul Korean,” The Journal of Studies in Language 31(4), 979-998 (2016).
6. Park, See-Gyoon and Ji-Young Kim. “A study on the analysis of the L1-L2 similarity between Korean and Vietnamese Monophthongs,” Korean Journal of Linguistics 42(4), 691-716 (2017).
7. Thompson, Laurence C. A Vietnamese grammar. Seattle, WA: University of Washington. (1965)
8. Yoon, Tae-Jin and Yoonjung Kang. “Monophthong analysis on a large-scale speech corpus of read-style Korean,” Phonetics and Speech Sciences 6(3), 139-145 (2014).

Korean Causal Connective Expressions in a Cross-linguistic and Cultural Perspective

Sujeong Choi[1][0000-0002-0424-880X] and Sinhye Nam[2][0000-0002-7177-0235]

[1] KDI School of Public Policy and Management, 263 Namsejong-ro, Sejong-si, 30149, Republic of Korea
[2] Kyung Hee University, 26, Kyungheedae-ro, Dongdaemun-gu, Seoul, 02447, Republic of Korea
schoi@kdis.ac.kr, namsh@khu.ac.kr

Abstract. This study aims to explore the reason why there are various causal connective expressions in Korean in comparison with other languages from a cultural perspective. In this study, Korean causal connective expressions are linguistically analyzed from the cross-linguistic perspective with English expressions in the Korean-English Parallel Corpora. Then, the differences between Korean causal connective expressions and the corresponding English expressions are interpreted from a cultural point of view. The findings of this study are as follows. First, there are various causal connective expressions implying a negative meaning in Korean, and negative nuances can be indirectly delivered through grammatical expressions. Second, in Korean the causal connective expression that includes the conjecture meaning is much more frequently used compared to English. The phenomenon of using a lot of guessing expressions in Korean can be interpreted as Korean speakers exhibiting a tendency to express their thoughts or opinions mildly and indirectly rather than strongly and clearly, and this is also related to showing politeness toward the listener. Third, there is the causal connective expression implying that it is one of various reasons in Korean. It implies that Korean speakers intend to avoid conclusive expressions by emphasizing that it is one of several reasons rather than concluding a single reason. This can also be viewed as a way to keep from expressing one's intentions too strongly and avoid causing the other person to lose face.

Keywords: Cross-linguistics, Causal Connective Expressions, Interpretive Ethno-grammar

1 Introduction

Communication competence encompasses not only linguistic and grammatical competence, but also cultural competence. In order to communicate properly in a particular language society, speakers must not only understand basic linguistic structures such as vocabulary, phonemes, and grammar, but also know how to speak appropriately in a given social and cultural context (Saville Troike, 2003). This is because a language is

shaped by the culture. Depending on the culture, there may be differences in the way of expressing the same subject in each language.

For example, the Hanunoo language of the Philippines has dozens of words for different kinds of *rice* (Conklin, 1957) and Russian has several different everyday words for different kinds of *friends* rather than one basic everyday word like the term *friend* used in English (Wierzbicka, 1997), and this phenomenon is not limited to words. Compared to other languages, in Korean, there are a lot of connective expressions that express a cause or reason. This phenomenon can be analyzed along the same lines as the phenomenon in which *rice* in the Hanunoo and *friends* in Russian appear significantly more than in other languages. In other words, it can be said that the expressions that indicate the cause or reason appear very diverse due to the cultural characteristics of Koreans' tendencies or communication.

This study aims to explore the reason why there are various causal connective expressions in Korean from a cultural point of view. For this, first, Korean causal connective expressions are linguistically analyzed from the cross-linguistic perspective with English expressions. Then, the differences between Korean causal connective expressions and the corresponding English expressions are interpreted from a cultural point of view. If speakers are aware of the reasons for the variety of expressions representing the cause or reason in Korean from a linguistic cultural perspective, it will be possible to help foreign learners avoid misuse of language in a cultural context and achieve successful cross-cultural communication.

2 Korean Causal Connective Expressions

Choi (2022) analyzed the grammar items from 10 classes of Korean textbooks and 3 classes of grammar books to categorize them based on their meanings. As a result, a total of 68 semantic categories of grammar items were presented and a list of synonymous grammar items was organized for each category. It was found that the meaning category which has the greatest number of synonymous grammar items was the [cause] category. According to Choi (2022), there are 24 synonymous grammar items expressing the cause or reason in Korean: *-a/eoseo, -(eu)nikka, -(eu)meuro, -gi ttaemune, -neurago, -neun barame*, *-gilrae, -gie, -deoni, -at/eotdeoni, -(eu)n/neun mankeum, -a/eo gajigo,* and so on. These all have a common meaning, which is the cause or reason for the following clause in the sentence. However, they differ in terms of syntactic conditions, contextual formality, or semantic features.

For example, in Korean textbooks, the following causal grammar items are provided with descriptions as they have the additional semantic features compared to the basic/neutral causal grammar items '*-a/eoseo, -(eu)nikka, -(eu)meuro, -gi ttaemune'*.

(1) Ast-(eun) tase, Vst-neun tase: It is a negative expression of cause or reason.
(2) *Ast/Vst-(eun) nameoji*: It is used when some action or situation in the first clause becomes worse, and it leads to a negative result in the following clause.
(3) *Vst-neurago*: It indicates a cause, reason, or purpose getting a negative result.
(4) *Vst-neun barame*: It indicates a cause or reason. The preceding situation negatively affects the following action.

(5) *Vst-neun tonge*: It indicates a cause or reason that caused the negative situation or result of the following clause.
(6) *Ast/Vst-a/eoseo geureonji*: It is used when the previous action or situation seems to be the cause and reason for what follows but cannot be determined for sure.
(7) *Ast/Vst-go haeseo*: This is used to show that the preceding clause is one of many reasons for the content in the following clause.

As underlined above, the causal connective expressions in (1) to (7) have additional semantic features: the grammar items in (1) to (5) contain [+negative] features, the grammar item in (6) has [+uncertainty] features, and the grammar item in (7) implies [+plurality]. Like this, Korean causal connective expressions are not only able to express the cause or reason for the following clause, but they can also imply additional semantic features or nuances.

As such, it seems there are numerous grammar items expressing cause or reason in Korean. At this point, the question of whether this phenomenon exists in other languages as well, or whether this is a characteristic unique to Korean can be raised.

In order to explore the above question in this study, seven causal connective expressions that contain additional semantic features are compared with the corresponding English expressions. In addition, through these analysis results, we attempt to interpret the reason why there are many causal connective expressions in Korean from a cultural perspective.

3 Data and Methodology

The data of this study was collected by the Korean-English parallel corpora developed by *AI-Hub*. *AI Hub* is an AI integration platform operated by Korea's Ministry of Science and ICT and the National Information Society Agency. The Korean-English parallel corpora were released as a part of the data construction project for artificial intelligence learning. The Korean-English parallel corpora consist of three styles of corpus: literary, colloquial, and colloquial conversation, covering 1,600,000 sentences.

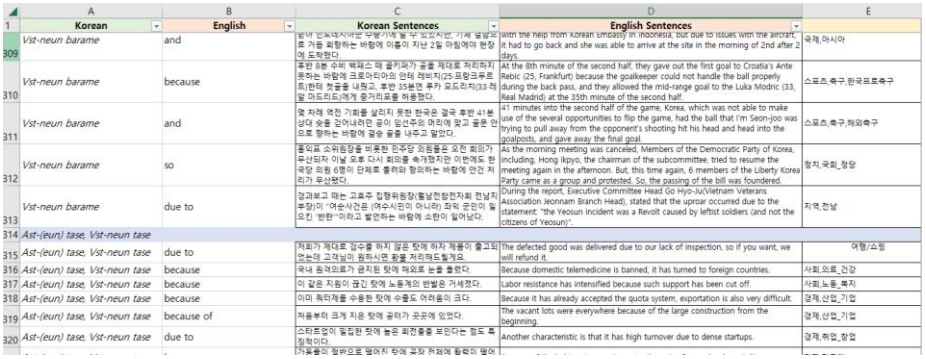
In this study, 200,000 sentences (2,658,545 words) from the news article corpus in the literary style and 100,000 sentences (779,541 words) from the conversation corpus in the colloquial style were analyzed as follows.

Table 1. Information of Korean-English Parallel Corpora

No.	Style	Context	The Number of Sentences/Words	
1	Literary	News Articles	200,000 sentences (2,658,545 words)	300,000 sentences (3,438,086 words)
2	Colloquial (Conversation)	Conversation in a meeting, shopping, school, restaurant, etc.	100,000 sentences (779,541 words)	

A total of 300,000 sentences (3,438,086 words) in the Korean-English parallel corpora were examined in order to explore the corresponding English expressions with Korean reason/causal connective expressions.

4



	A	B	C	D	E
1	Korean	English	Korean Sentences	English Sentences	
309	*Vst-neun barame*	and	발이 인도네시아로 우방기에 돌 수 있었지만, 기체 결함으로 다시 가는 회항하는 바람에 이틀이 지난 2일 아침에야 현장에 도착했다.	With the help from Korean Embassy in Indonesia, but due to issues with the aircraft, it had to go back and she was able to arrive at the site in the morning of 2nd after 2 days.	국제_아시아
310	*Vst-neun barame*	because	후반 8분 수비 백패스 때 골키퍼가 공을 제대로 처리하지 못하는 바람에 크로아티아의 안테 레비치(25·프랑크푸르트)한테 찬스를 내줬고, 후반 35분엔 루카 모드리치(33·레알 마드리드)에게 중거리포를 허용했다.	At the 8th minute of the second half, they gave out the first goal to Croatia's Ante Rebic (25, Frankfurt) because the goalkeeper could not handle the ball properly during the back pass, and they allowed the mid-range goal to the Luka Modric (33, Real Madrid) at the 35th minute of the second half.	스포츠_축구,한국프로축구
311	*Vst-neun barame*	and	몇 차례 역전 기회를 살리지 못한 한국은 결국 후반 41분 상대 슛을 걷어내려던 공이 임선주의 머리에 맞고 골문 안으로 향하는 바람에 결승 골을 내주고 말았다.	41 minutes into the second half of the game, Korea, which was not able to make use of the several opportunities to flip the game, had the ball that Im Seon-joo was trying to pull away from the opponent's shooting hit his head and head into the goalposts, and gave away the final goal.	스포츠_축구,해외축구
312	*Vst-neun barame*	so	돌연료 소위원장을 비롯한 민주당 의원들은 오전 회의가 무산되자 이날 오후 다시 회의를 속개했지만 이번에도 한국당 의원 6명이 단체로 출석을 항의하는 바람에 안건 처리가 무산됐다.	As the morning meeting was canceled, Members of the Democratic Party of Korea, including, Hong Ikpyo, the chairman of the subcommittee, tried to resume the meeting again in the afternoon. But, this time again, 6 members of the Liberty Korea Party came as a group and protested. So, the passing of the bill was foundered.	정치_국회_정당
313	*Vst-neun barame*	due to	경과보고 때는 고흥주 집행위원장(월남전참전자회 전남지부장)이 "여순사건은 (여수시민이 아니라) 좌익 군인이 일으킨 '반란'이라고 발언하는 바람에 소란이 일어났다.	During the report, Executive Committee Head Go Hyo-Ju(Vietnam Veterans Association Jeonnam Branch Head), stated that the uproar occurred due to the statement: "the Yeosun incident was a Revolt caused by leftist soldiers (and not the citizens of Yeosun)".	지역_전남
314	*Ast-(eun) tase, Vst-neun tase*				
315	*Ast-(eun) tase, Vst-neun tase*	due to	저희가 제대로 검수를 하지 않은 탓에 하자 제품이 출고되었는데 고객님이 원하시면 환불 처리해드릴게요.	The defected good was delivered due to our lack of inspection, so if you want, we will refund it.	여행/쇼핑
316	*Ast-(eun) tase, Vst-neun tase*	because	국내 원격의료가 금지된 탓에 해외로 눈을 돌린다.	Because domestic telemedicine is banned, it has turned to foreign countries.	사회_의료_건강
317	*Ast-(eun) tase, Vst-neun tase*	because	이 같은 지원이 끊긴 탓에 노동계의 반발은 거세졌다.	Labor resistance has intensified because such support has been cut off.	사회_노동_복지
318	*Ast-(eun) tase, Vst-neun tase*	because	이미 쿼터제를 수용한 탓에 수출도 어려움이 크다.	Because it has already accepted the quota system, exportation is also very difficult.	경제_산업_기업
319	*Ast-(eun) tase, Vst-neun tase*	because of	처음부터 크게 지은 탓에 공터가 곳곳에 있었다.	The vacant lots were everywhere because of the large construction from the beginning.	경제_산업_기업
320	*Ast-(eun) tase, Vst-neun tase*	due to	스타트업이 밀집한 탓에 높은 회전율을 보인다는 점도 특징적이다.	Another characteristic is that it has high turnover due to dense startups.	경제_취업_창업

Fig. 1. An Example of Data Analysis

As shown in **Figure 1**, 479 sentences using the seven Korean causal connective expressions were searched from the Korean-English parallel corpora and the corresponding English expressions for each Korean causal connective expression were extracted from the English sentences.

After examining all the corresponding English expressions, they were categorized into two groups: reason/causal expressions and non-reason/causal expressions. If the English expressions are used to indicate the reason or cause, they were included in the reason/causal expressions. In contrast, if the English expressions are not used to indicate the reason or cause, they were included in the non-reason/causal expressions.

4 Korean Causal Connective Expressions in a Cross-linguistic Perspective

Table 2 shows the frequency and rate of Korean causal connective expressions and corresponding English expressions in the Korean-English parallel corpora.

Table 2. Korean Causal Connective Expressions and Corresponding English Expressions

No.	Korean	N	English		N	%
1	Ast-*(eun) tase*, Vst-*neun tase*	145	Reason/causal expressions	because (of)	57	39.31
				due to	32	22.07
				as	25	17.24
				since	14	9.66
				so	2	1.38
				by	1	0.69
				for	1	0.69
				therefore	1	0.69
			Non-reason/causal expressions	cause	4	2.76
				after	2	1.38
				and	1	0.69
				as a result	1	0.69

				result in	1	0.69
				lead	1	0.69
				so … that	1	0.69
				while	1	0.69
2	Ast/Vst-*(eun) nameoji*	24	Reason/causal expressions	because	5	20.83
				as	3	12.50
				for	2	8.33
			Non-reason/causal expressions	and	5	20.83
				so ... that	3	12.50
				∅	5	20.83
				by ~ing	1	4.17
3	Vst-*neurago*	13	Reason/causal expressions	because	4	30.77
				due to	1	7.69
				from	1	7.69
				as	1	7.69
				with	1	7.69
				~ing	2	15.38
			Non-reason/causal expressions	and	1	7.69
				bring	1	7.69
				on	1	7.69
4	Vst-*neun barame*	153	Reason/causal expressions	because	61	39.87
				so	24	15.69
				as	14	9.15
				since	11	7.19
				due to	6	3.92
				from	1	0.65
			Non-reason/causal expressions	and	12	7.84
				∅	9	5.88
				cause	6	3.92
				, which	2	1.31
				after	2	1.31
				when	2	1.31
				drive	1	0.65
				disrupt	1	0.65
				get	1	0.65
5	Vst-*neun tonge*	4	Reason/causal expressions	since	1	25
				due to	1	25
			Non-reason/causal expressions	and	1	25
				confuse	1	25
6	Ast/Vst-*a/eoseo geureonji*	124	Reason/causal expressions	because	31	24.80
				maybe (it's) because	28	22.40
				so	13	10.40
				probably because	11	8.80
				perhaps because	7	5.60

6

				since	7	5.60
				due to	3	2.40
				maybe that's why	2	1.60
				maybe the reason	2	1.60
				not sure if it is because	2	1.60
				with	2	1.60
				guess it's because	1	0.80
				perhaps ... so	1	0.80
				perhaps due to	1	0.80
				whether it's from	1	0.80
			Non-reason/causal expressions	∅	8	6.40
				and	2	1.60
				as to whether	1	0.80
				no wonder	1	0.80
				so ... that	1	0.80
7	Ast/Vst-*go haeseo*	9	Reason/causal expressions	because	2	22.22
				so	2	22.22
				since	1	11.11
				due to	1	11.11
				as	1	11.11
			Non-reason/causal expressions	∅	2	22.22

As described in Chapter 2, the Korean causal connective expressions (1) to (5) connote [+negative] meaning, the expression in (6) has [+uncertainty] meaning, and the expression in (7) contains [+plurality] meaning additionally. We analyzed whether these additional semantic qualities appear in the corresponding expressions in English, and the results are examined by each additional semantic quality below.

4.1 [+Negative] feature in causal connective expressions

The English expressions that correspond with the Korean causal connective expressions *Ast-(eun) tase, Vst-neun tase, Ast/Vst-(eun) nameoji, Vst-neurago, Vst-neun barame, Vst-neun tonge*, which imply the [+negative] feature, were analyzed. As a result, in English there were no causal expressions that have the [+negative] feature noticed in Korean expressions, and only basic/neutral causal expressions such as *because, due to, as, since* were used as the corresponding expressions.

In English, it was found that there was a large tendency to use direct negative vocabulary when trying to connote a negative meaning rather than implying such negative meaning through grammar items. In that, in English, words with negative meanings are used directly when expressing negative intentions, whereas in Korean, grammatical expressions implying the negative meaning are used somewhat indirectly. It means in Korean, even if the speaker does not use direct vocabulary to express negative intentions, negative nuances can be indirectly delivered through grammatical expressions. The phenomenon of speakers indirectly expressing their intentions in Korean can also

be found when they express their thoughts or opinions with conjecture expressions such as *–(eu)n/neun/(eu)l geot gat-* or when they express their plans with *–(eu)lkka ha-, -(eu)lkka sip-,* and so on.

4.2 [+Uncertainty] feature in causal connective expression

The English expressions that correspond with the Korean causal connective expression *Ast/Vst-a/eoseo geureonji*, which implies the [+uncertainty] feature, were analyzed. In English, conjecture expressions such as *maybe, probably, and perhaps* were not included in as many as 50% of the corresponding English expressions with *Ast/Vst-a/eoseo geureonji* in the reason/causal expression category. In that, cases involving speculation in the causal expressions appeared much more in Korean.

As mentioned above, this coincides with the phenomenon of using a lot of guessing expressions in Korean. In Korean, when expressing a cause or reason, there is a strong intention to express it mildly rather than strongly and clearly, and this is also related to showing politeness toward the listener.

4.3 [+Plurality] feature in causal connective expression

The English expressions that correspond with the Korean causal connective expression *Ast/Vst-go haeseo*, which implies the [+plurality] of the reason, were analyzed. In Korean, this expression shows that the preceding clause is one of many reasons for the content in the following clause. However, in English, this connotative meaning was not expressed. Unlike English, the causal expression implying that it is one of various reasons frequently appears in Korean because it implies that the speaker intends to avoid conclusive expressions by emphasizing that it is one of several reasons rather than concluding a single reason. This can also be said to be a way to keep from expressing one's intentions too strongly. This is used in a context where the speaker wants to avoid definitive reasons: when the speaker has to refuse the other's request or suggestion, the speaker dooes not want to cause the other person to lose face by giving the connotative meaning that there are many reasons for being unable to accept the request or suggestion, or when the speaker does not want to explicitly express their actual reasons.

4.4 Additional Findings

Through the analysis, a few findings can be discussed further.

First, the corresponding English expressions with Korean causal expressions were mostly one of four expressions such as *because, since, due to, and as,* and it shows that in English various grammar items expressing the cause or reason are not used as much as in Korean.

Second, even though there are some different causal expressions in English, their difference is the degree of formality rather than semantic features. However, in Korean, there are various causal expressions, and their difference can be explained in many aspects: formality, syntactic condition, and semantic features.

Third, there are cases in which Korean causal expressions do not correspond with the non-reason/causal expressions in English. This indicates that causal expressions are

used more frequently in Korean, and sometimes the causal expression is used for other intentions such as supporting their thoughts rather than expressing the actual cause or reason in a logical context.

5 Conclusion

It is generally recognized that languages differ in the amount – and kind – of attention given to different aspects of reality through their lexical systems: Arabic has numerous words for *sand*, Hanunoo for *rice*, and so on (Wierzbicka, 2002). This study considered that it also applies to grammatical expressions and not only to the lexical system. Thus, this study focused on the phenomenon that there are various causal connective expressions in Korean in comparison with other languages from a cultural perspective.

As a result of the linguistic analysis from the cross-linguistic perspective between Korean causal connective expressions and corresponding English expressions, it was found that the following three aspects were characteristics unique to Korean and these were interpreted in a cultural perspective.

First, there are various causal connective expressions implying a negative meaning in Korean, whereas there are no such expressions in English. In Korean, the negative nuance can be indirectly delivered through grammatical expressions.

Second, in Korean the causal connective expression that includes the conjecture meaning is much more frequently used compared to English. As mentioned above, the phenomenon of using a lot of guessing expressions in Korean can be interpreted as Korean speakers exhibiting a tendency to express their thoughts or opinions mildly and indirectly rather than strongly and clearly, and this is also related to showing politeness toward the listener.

Third, there is the causal connective expression implying that it is one of various reasons in Korean, whereas no such expression is used in English. It implies that Korean speakers intend to avoid conclusive expressions by emphasizing that it is one of several reasons rather than concluding a single reason. This can also be said to be a way to keep from expressing one's intentions too strongly and avoid causing another person to lose face.

The results of this study from the linguistic-cultural perspective will be able to help foreign learners choose the proper causal connective expressions so that they can express their intentions with a cultural understanding and avoid misuse of language in a cultural context.

References

1. Choi, Sujeong, A Study of the Meaning-Based Categorization of Grammar Items for Synonymous Grammar Education of Korean Language, Doctoral dissertation, Seoul: Yonsei University (2022).
2. Conklin, Harold, Hanunoo Agriculture. Rome: Food and Agriculture Organization of the United Nations (1957).

3. Saville Troike, Muriel, The Ethnography of Communication: An Introduction. 3rd edn. Oxford: Blackwell (2003).
4. Wierzbicka, Anna, Understanding Cultures through Their Key Words: English, Russian, Polish, German, Japanese. New York: Oxford University Press (1997).
5. Wierzbicka, Anna, English Causative Constructions, In N.J. Enfield (Ed.), Ethnosyntax, New York: Oxford University Press (2002).

Smart Farm Management System Using Humidity Meter

Yuseung Shin[1] and Jaeyun Jeong[2]

Kyunghee High School, Seoul, 02447, Korea

[2] Hankuk University of Foreign Studies, Seoul. 17, 02450, Korea

Abstract. With the recent development of IoT technology, farmers can enjoy convenient and practical lives with smart farms created by combining agriculture and IoT technology. In this paper, we introduce the characteristics of plants and explain the direction of Beacon devices and smart devices through AAP. When managing smart farm moisture using a hygrometer, it is useful for promoting plant growth as well as saving water.

Keywords: Smart Farm, IoT, ICT

1 Introduction

The development of the Internet of Things (IoT) has made people enjoy a more comfortable life. Since then, the combination of agriculture and the Internet of Things has allowed farmers to enjoy a convenient and practical life with smart farms.
Although cultivation kits are being released as personal smart farms, there is a limit that plants are not compatible with various pots, and the types of plants are limited to vegetables, and temperature and humidity light should be correlated in the actual environment. To solve this problem, moisture can be measured with existing moisture sensors and weight sensors, but measurement errors and weight sensors are unstable due to plant growth, so an automated humidity control algorithm is needed with development of humidity sensors and beacons. In this paper, we propose a method to provide the appropriate humidity of plants using a humidity sensor.

2 Related

This section introduces existing smart farm and beacon technologies and explains the characteristics of plants and the plants introduced by Korea.

2.1 Definition of Beacon

Beacon is a Bluetooth protocol-based NFC device. Beacon's wireless communication has recently been in the spotlight as a near-field communication technology due to many advances such as low power, miniaturization, life extension, and increased reception distance, without the pairing process that had to be done to

connect between the two devices using Bluetooth. In addition, the maximum communication distance is relatively long at about 50m, and sophisticated location can be identified indoors. Beacons classify certain objects with beacons as UUID values and transmit signals to users without a separate pairing procedure for each close-range section using RSSI (Received Signal Strength Indicator) to individuals with smartphones at low cost. The beacon transmitter periodically signals its UUID and RSSI values, and when a person with a smartphone comes within the reach of this signal, the smartphone recognizes it and sends signal information to the server. [3]

2.2 Implementing Beacons

There are Starbucks siren orders, hospital appointments, and mobile payments for medical expenses using APP, but the service is not working well in some places in the hospital due to battery consumption problems, but the problem is expected to be solved in the future. There is also a disadvantage of weak security.

2.3 Smart Farm

It is a system created by the fusion of precision agriculture and ICT technology that emerged in the 1980s, and a system that collects data on plant growth and environment and helps decision-making is called a smart farm. It uses crop data collected through satellites, weather information, and environmental information collected using various sensors [1]

2.4 Smart Farm Trends

According to the Korea Institute for Science and Technology Jobs, industrial trends by smart farm country are spreading to areas such as distribution and consumption of smart farms in Korea, but so far, agricultural production has been the core. It is believed that it is concentrated in the monitoring and control stages, and developing optimized algorithms using big data and automation technologies related to robots are currently in the R&D stage Currently, the smart farm system applied to our farms remains at the level of opening and closing of cultivation facilities (insulation cover, ceiling, curtain, ventilator, sprinkler, fluid, hot air, etc.) through smart media based on environmental information (temperature, humidity, CO2, illumination, etc.). In the future, it is required to develop a growth optimal environment setting model for precise crop management by growth stage based on cultivation growth information and to develop a specialized model for diagnosis of crop physiological disorders and pests.

The Netherlands is a representative smart farm-using country, and although its land area is only 1/2 of that of Korea, it has become the world's second-largest exporter of agricultural products through the introduction of ICT. The Netherlands is a representative horticultural country, and 99% of all greenhouses are glass greenhouses, and various sensors and control solutions have been developed based on decades of accumulated big data and experience optimized for the cultivation

environment. Through these agricultural ICT technologies, production and quality optimization will be planned, and Priva, a leading Dutch company, is producing the world's best greenhouse environment control system and exporting it to countries around the world.

The U.S. is attempting to use not only IoT but also nanotechnology and robot technology for agriculture in earnest. In the case of Google, it is trying to develop an artificial intelligence decision support system technology that helps spread seeds, fertilizers, and pesticides by collecting big data on soil, moisture, and crop health.

In Japan, companies such as IBM, NEC, Fujitsu, and NTT provide various services by incorporating ICT technology into the agricultural field.

Examples of Japan are IBM's agricultural product history tracking service, NEC's M2M-based growth environment monitoring and logistics service, and Fujitsu's agricultural management cloud service system.

Israel is a leader in monitoring the growing environment and automatically measures crop growth information such as crop size, stem change, and leaf temperature, and predicts accurate yields by automatically adjusting water supply cycles and water supply, especially, the development of crop stress sensors has increased production by more than 40% [2]

2.5 Plants

The current status of inflow-oriented plants in Korea and their generative characteristics the distribution of origin of 114 species of inflow-oriented plants is shown in Figure 1. There were 17 species of plants native to North and South America, accounting for 14.9% of the total. Next, 15 species of plants native to Africa and Asia each accounted for 13.3%. In addition, there were 14 species of plants native to North America and 11 species in South America, 42 species native to North and South America, accounting for 36.8% of the total. Therefore, thorough quarantine should be carried out because seeds of imported plants are most likely to be mixed or adhered to agricultural products imported from North and South America, Africa, and Asia. And there were nine species of plants native to the Mediterranean coast. Therefore, the nine species were distributed on three continents: Europe, Africa, and Asia. It was included in the top 100 malignant weeds designated by IUCN and was designated as an introductory plant in Korea, but some of them are native to tropical regions, so they cannot survive even if they enter Korea. Although it is judged that plants of this inflow should be excluded, even if some tropical regions are native, Jeju Island has a tropical climate due to global warming, suggesting the possibility of survival.

3 Smart Farm Management System Using Humidity Meter

This section presents prior research and the direction in which Beacon devices and smart devices configure smart farm systems through (APP) apps.

3.1 System Configuration Diagram

4

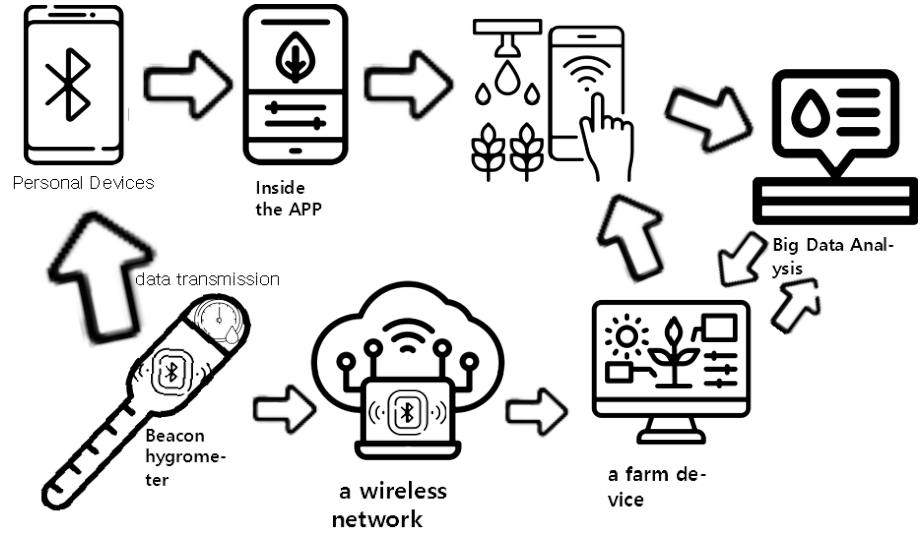


Fig. 1. System Configuration Diagram

After connecting the Beacon device built into the hygrometer and the smart device (smartphone) through the (APP) app, farm use is presented at startup. Users can choose plant types by presenting a list of plants, register photos and names, and finish setting up Wi-Fi after connecting the mobile device and the humidity sensor using a beacon in the process of adding them. For farms, help connect the farm device to the sensor.

The hygrometer settings are as follows.

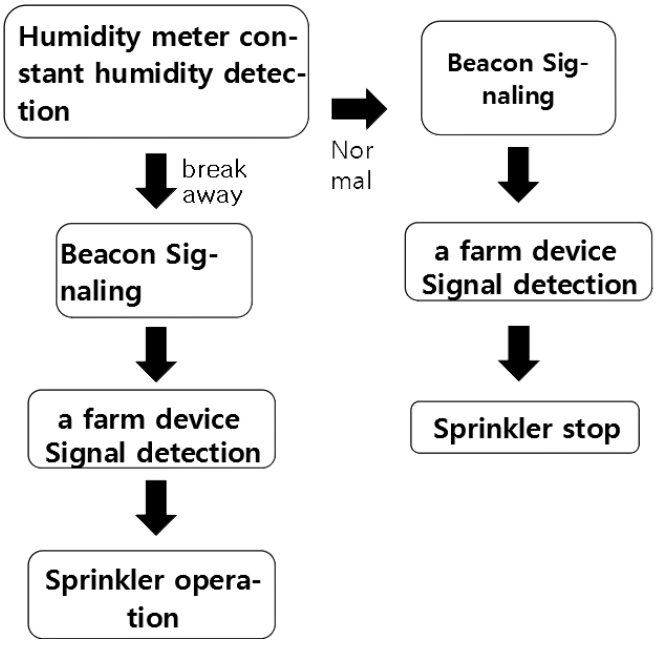


Fig. 2. The hygrometer settings

3.2 APP Internal

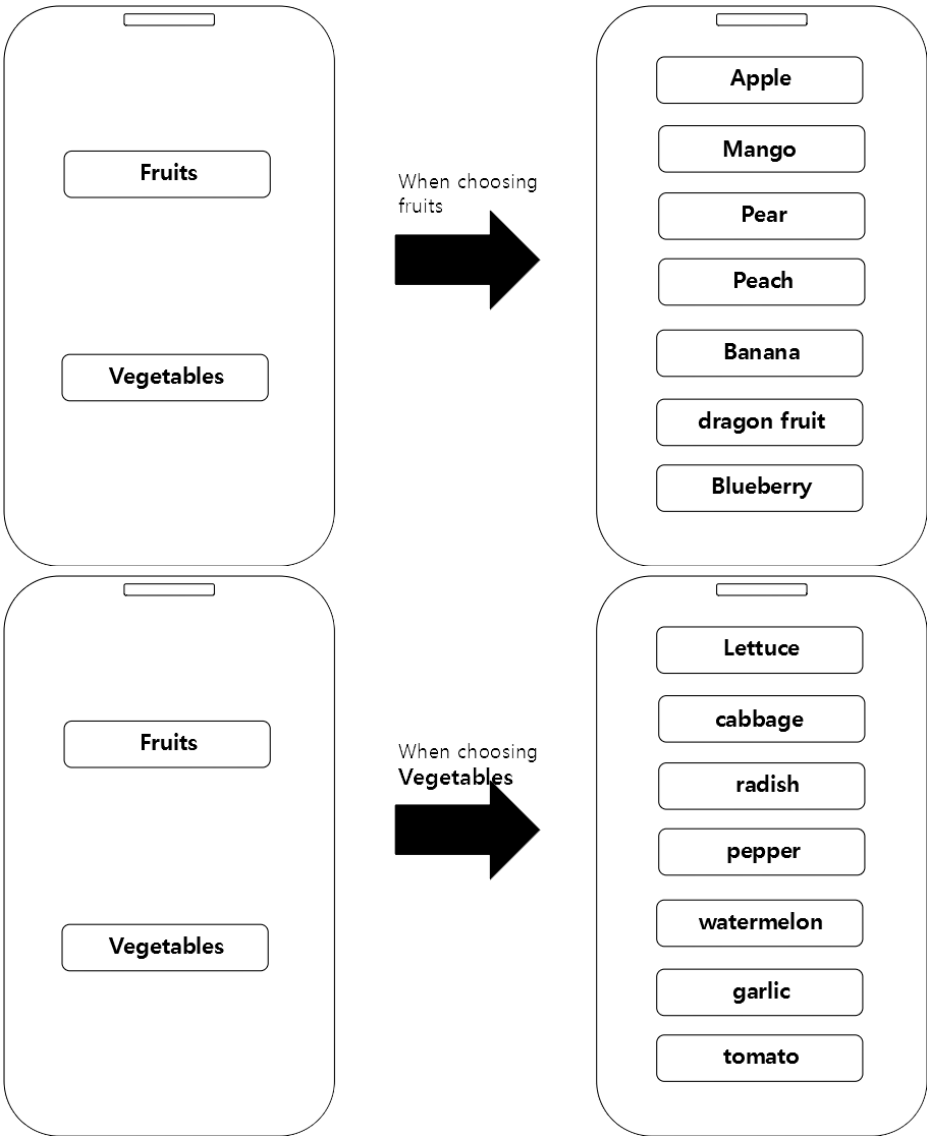


Fig. 3. device Screen

After connecting the personal device and the hygrometer through APP, fruits, vegetables, and fruits are presented, and when the user selects fruits, the fruit type is presented, and even if the vegetables are selected, the vegetable type is presented. When the user selects the type of fruit or vegetable, set the appropriate humidity on the hygrometer

There is a 'farm type' installation method so that the humidity controller can be applied in various places.

4. Conclusion

In this paper, we limited the humidity measurement system using soil humidity sensors that secure the limitations of plant types applied to existing smart farms and increase utilization and efficiency in smart farms. The system may expand the scope of application of existing smart farms such as various types of flower pots, vinyl houses, and open fields using various materials. Also, due to global warming, fruit production in Korea is changing little by little by little. It can also be applied to tropical fruits and plants such as mangoes and apple mangoes, which are tropical fruits grown on Jeju Island, suggesting higher viability. By implementing a humidity meter using Beacon and implementing an (APP) app, it presents a direction to grow various types of plants and fruits beyond smart farms, where the types of plants are currently limited to vegetables.

References

1. Solhee Park, JinSung Cho, Seok Hyun Eom: Smart Farm water management system using weight sensors. Proceedings of the Korean Information Science Society Conference, 1731-1733 (2022).
2. https://www.bioin.or.kr/InnoDS/data/upload/tech/CF87C038-7017-11A5-6A54-5581930C2D2F.pdf
3. Jihoon Park__SoYeon Lee_JUNGMINWOO_Kim Dae-Young: Fire Evacuation System Using Beacon for Hard-of-Hearning People. JKICS, 319-330 (2022)

A Study of OSMU for Henan Seolheon's works

Zhao Wenxuan[1], Young-Hoon An[2], Hwa-Young Jeong[3]

[1, 2] Department of Korean Language and Literature, Kyung Hee University, 26, Kyungheedae-ro, Dongdaemun-gu, Seoul, 02447, Republic of Korea
[3] Humanitas College, Kyung Hee University, 26, Kyungheedae-ro, Dongdaemun-gu, Seoul, 02447, Republic of Korea

munseon981011@naver.com[1], yhnahn@khu.ac.kr[2], hyjeong@khu.ac.kr[3]

Abstract. The purpose of this paper was to study the contents of the work theory of the famous Korean ancient female writer Henan Seol-heon. When introducing classical works or classical poets to modern people, they are looking for content such as more fun and easier to receive, and storytelling. When referring to the keywords of classical literature and female poets, the name Henan Seolheon appears a lot. It needs to know the works of female poets who have a great position in classical literature. During the Joseon Dynasty, when Heo Nan-seolheon lived, most of the creators of literary works were male writers, claiming Confucian ideas. Due to Confucianism, women of that era were restricted in various fields such as status, recognition, freedom, and study. In this paper, we investigate the contents using the current works of Heo Nan-seol-heon and investigate the big direction of how to promote Heo Nan-seol-heon and how to proceed with the contents. In this era, almost everything is related to the database, but classical works seem to be difficult to relate to the database. In addition, it seems that to achieve this, we must support technology on many levels.

Keywords: Heo Nanseolheon, content using classical literature, works by Heo Nanseolheon.

1 Introduction

Heo Nan-seol-heon was a female poet, painter, writer, and government official in the mid-Joseon Period. In an era when women had no name, Henan Seol-heon made his own name. Her real name is Heo Cho-hee, and it is passed down as Heo Ok-hye. The pen name is Nan Seolheon, and the ruler is Gyeongbeon. Both fathers and children of Heo Nan-seol-heon's family were excellent in writing, and people in the world called Heo's five sentences (Heo-yeop, Heo-sung, Heo-bong, Heo Nan-seol-heon, and Heo Gyun), but considering the Confucian society at the time, they were relatively generous to women, and they were able to study Chinese characters. Figure 1 is the standard image of Henan Seol-heon. This painting was created by artist Son Yeon-chil in 1997. The painting is now in the collection of the National Museum of Modern and Contemporary Art in Korea.

When of Heo Nan-seol-heon was 8 years old, she was called a prodigy and famous among scholars after building the Gwanghanjeonbaek Okru Sangryangmun. At the

age of 15, she married Kim Seong-rip of Andong Kim's family. The people who influenced Heo Nan-seol-heon the most while learning writing was her brother Ha-gok and her teacher Son-gok. While learning poetry from Songok, Heo Nan-seol-heon even accepted it as his person. Songok's dissatisfaction, sense of defiance, arrogance, and rudeness in knowing the world would have been in line with Gyosan and Nanseolheon. Misfortunes are continuously encountered in Henan Seol-heon's family while her marriage is not smooth and her relationship with her mother-in-law is not good. Her father Heo Yeop died in 1580, and she had a son and a daughter as children, both of whom died at a young age due to an epidemic. Henan Seol-heon says that she died in 1589 at the young age of 27 because her family declined in the middle, and her father, brother, and her children died one after another, and she was under a lot of pressure and stress from her mother-in-law.

Fig. 1. Henan Seol-heon, the picture if from Namuwiki.com

The 16th century was a time when Confucian ethics were strictly applied throughout politics, economy, and culture. At that time, a society in which people demanded strict moral ethics and closed allowed women to admire the unreal world and find desires that could not be solved in a fictional world.

Although the period when Heo Nan-seol-heon lived was a period of great development in the literary and artistic aspects, the political turmoil of the Joseon Dynasty was at that time. From the 15th century to the 17th century, data on women's songs in the early and mid-Joseon Periods are mainly concentrated on sijo, and the writers' class is also centered on the kisaeng class. In comparison, not only are there very few female writers in the upper class but there is also a problem with the credibility of the author. However, the literary activities of the female class were active in the late Joseon Dynasty, centering on the Gyubang lyrics 1) .In the late 16th century, especially during the reign of King Seonjo, the literary atmosphere was so strong that it was called "Mureungseongse," and Seongrihak had Hwang Jin-deok, Song Soon, Imje, Jeong Cheol, Park In-no, and Sinheum, Jang Yu, Lee Jeong-gu, Seo Yang Sa-eon, and Hanho 2). It can be said that the period when Heonnan Seolheon lived was the most prosperous period of Joseon literature.

2 Related works

Until now, papers on poet Heo Nan-seol-heon can be largely divided into parts. First, the study of life and poetry in poets, i.e., work theory and writer theory, the second is the study of the poet's work and domestic and international poet's comparison, the third is the study of the problem of work belonging to works, the fourth is the study of translation problems of poetry by Heo Nan-seol-heon, and the fifth is the study of women's ideas in poet's works.

There are only a few content and storytelling papers on Henan Seol-heon's work. The main contents are as follows.

Lee Hyuk-jin and Shin Ae-kyung presented A Study on the Directions of Utilization for Cultural Tourism Contents of Gangneung City in Gangwon Province - Focused on Specific Historical Figures and Places- 3) Among the papers, Gangneung-si, Gangwon-do, was presented as a case area, and the direction of exploration and use of cultural tourism contents centered on geography, tourism resource status, and historical figures. In addition, the purpose was to promote Gangneung-si through historical figures such as Kim Si-seup, Sin Saimdang, Yulgok Yi-i, Heo Gyun and Heo Nan-seol-heon of the Joseon Dynasty, and related places.

In Kang Myung-ye's ‘A method of reality correspondence and storytelling of Heo Nanseolheon and Yoon Heesoon ‘4) In preparation for Heo Nan-seol-heon and Yoon Hee-soon, who are believed to have many things in common because they are marginalized and foreigners in the background of the times, especially local (Gangwon-do), they reviewed their world view, self-response, and writing patterns, and even briefly promoted the storytelling as an appendix.

In the thesis of Kim Hee-sook and Jang Woo-kwon's ‘ A Study on the Content and Composition of Digital Character Archive in Works and Subjects: Female Writers in the mid of the Joseon Dynasty 5)’ The purpose of this study was to explore the contents and composition of works and subject-type digital character archives for the works of Shin Saimdang, Heo Nan-seol-heon, and Song Deok-bong among female writers in the mid-Joseon Dynasty.

Park Yong-jae's ‘A Study on the Extensiveness of Cultural Contents in Hernanseulhen Poetry 6)’ The paper studied the cultural background and storytelling method of the creation of the play "Dream Journey to the Peach Blossom" through the medium of Henan Seolheon, and the expansion of the poem into cultural contents.

Shin Soo-yeon's ‘Analysis of storytelling elements of the memorial spaces for Korean female artists 7)’ The thesis focused on the feminist perspective, which has recently become a hot topic in the cultural world. Among them, Heo Nan-seol-heon's example was seen as a change in the perception of oppression imposed on women in history.

4

3 3 Contents related to the research and work of Henan Seol-heon

Looking at Korean domestic papers, the contents of alternative studies are shown in the following table for the study of Henan Seol-heon's works. Among the DBPIA.co.kr papers, it is written focusing on the results that come out by setting the keyword 'Henan Seol-heon'.

Table 1. Previous research papers related to Henan Seol-heon

Number	Year	author and thesis name	Major Research Directions
1	2021	LIM MIJUNG, Reconsideration on the Materials of Heo Nanseolheon's Poems	complementary work
2	2021	Lee Cheol-hui, A critical investigation into the authorship of two proses in Nanseolheonsijip, allegedly written by Heo Nanseolheon	the question of quieting one's work
3	2018	Jeong Soyeon, Diglossia of Literature in the Middle Ages and Literacy Education: -Hwang Jini and Heo Nanseolheon in the 16th century-	a contrast study
4	2017	Yunhyeji, The Depressed Mood in Poetry by Female Writers from Ancient Korea and China - focusing on Huh-Nanseolhun and Wang-Fengxian	a contrast study
5	2017	Yun Inhuyn, Heonanseolheon's Consciousness through her Chinese Poems	the theory of works
6	2016	Lee, Hwa-hyung , A Study on the Consciousness of "Subject and Liberty" in HuhNanseolHeon's Life and Literature	the theory of works
7	2016	Park Hyun Kyu , Study on the Selected Edition of Heo Nanseolheon's Nanseol sihan Compiled by Heo Gyun in 1597	the problem of ripening/distribution
8	2016	Kang-myeonghye , Heo Nanseolheon's Yousun Poems and the Poems' Color Aesthetics - Focusing on Comparing Characteristics Color aesthetic, with other Youson poems, China and Joseon dynasty	a work theory/comparative study

9	2015	Yu Yukrye ,A Study on Nanseolheon Heo Romantic Love and Yearning Feeling Poems	the theory of works
10	2015	Han Seonggeum , Speculation on Chinese Poetry Written by Women from Noble Families in the 16th Century and the Expressive Aspects Used - Chinese Poetry by Song Duk-Bong and Huh Nanseolhun-	Comparative Research & Theory of Works
11	2014	Son Aenghwa , The study of Unfortunate consciousness that appears in Yuseonsa by Heonanseolheon - On the basis of poetic-word statistics and analysis	the theory of works
12	2014	Kang Minkyoung, The study on the time images in Yusun literature of Heonanseolheon	the theory of works
13	2030	Yi Dongha, Fictionalization of the Noble Women's Life during the Chosun Dynasty	writer's theory
14	1990	Lee Sanglan, A Comparative Study of HuLansulhun and Emily Dickinson -A Long Night Journey to the “Mother’s space”-	comparative study
15	1980	Jangjin, A Study on Heo Nan-seolheon's poems	The Theory of Writers and Works

In the table above, we can see the research history of Heo Nan-seol-heon in Korea. Most of them write papers on writer theory, work theory, comparative research, and work acquisition problems. In addition, many books about Henan Seol-heon's works are now included in Korea. He always studies the works of Heo Nan-seol-heon, focusing on "Nansol-heon Poetry," which Heo Gyun, Heo Nan-seol-heon's younger brother, edited. "Nanseolheonsi" was edited by Heo Gyun in 1608 and contained 210 poems in total. Until now, Heo Nan-seol-heon's content paintings have been produced in the form of dance, music drama, ballet, and musicals, but there are only a few works except for special performances related to the Pyeongchang-dong Mirror Olympics. The format used as the content is shown in the following table.

Fig. 2.. Utilization of content by Heo Nan-seol-heon

Time	content/work	Content utilized
2014.02.24	About the life and work of Heo Nanseolheon	documentary drama/Gangneung MBC
2016.08.20	The works "Kyuwon" and "Gamwoo"	Chamjak Dance (Gangneung Wonju University, Haerang Cultural Center)
2016.12.23	About the life and work of Heo	Music Drama (Gangneung

	Nanseolheon	Wonju University, Haerang Cultural Center)
2017.05.05	"Gamwoo", "Dream Journey to the Peach Blossom Land"	Ballet (CJ Towol Theater, Arts Center)
2018 PyeongChang Winter Olym-pics	About the life and work of Heo Nanseolheon	special performance
-	"Dream Journey to the Peach Blossom Land"	Musical
-	-	Heo Gyun and Heo Nanseolheon Memorial Hall (Gangneung)

Fig. 2. MBC documentary drama "Henan Seol-heon". Koo Hye-sun, a picture that appeared in an article titled "The 24th Broadcast," which released a still of the documentary drama "Heonan Seol-heon."

Figure 2 is in 2014, actress Koo Hye-sun filmed, made, and acted in the MBC documentary drama "Henan Seol-heon". Heo Nan-seol-heon, played by actress Koo Hye-sun in 2014, is a female literary scholar who is easier and more understood by the public. If you watch the video rather than the book, you can learn about Heo Nan-seol-heon's life more simply and interestingly.

Fig. 3. "Heonan Seolheon's Musical poster by Naver.com

Figure 3 shows poster of a musical play about Henan Seol-heon. In this format, it is introduced to the public by Henan Seol-heon. Since the number of spectators is also large, the method of combining classical and contemporary content through this can say a successful word.

Fig. 4. 'Heonan Seolheon' Memorial Hall image picture by Naver.com

Figure 4 is Ballet created by Heo Nan-seol-heon, using "Gamwoo" and "Dream Journey to the Peach Blossom Land". Kang Hyo-hyung will present 55-minute works under the themes of Heo Nan-seol-heon's poems "Gamwoo" and "Mongyu Gwangsangshansi." In the first half, he expressed Henan Seol-heon's warm and happy time through "Gamwoo," and in the second half, he expressed his painful and sad later life through "Mongyu Gwangsangshansi."

Heo Nan-seol-heon's life can be largely divided into two parts. The first half was when she was at home, and the days before the breakup were favored by her family, and unlike the women of the time, she was a woman from a prestigious family who could learn letters or literature. Before marriage, Heo Nan-seol-heon was a girl who lived without any worries. On the other hand, Heo Nan-seol-heon's poetry changed greatly due to her unhappy life after marriage, conflicts with her husband, her son's early death, family misfortune, and these causes.

4 Conclusion

Only those who are interested in classical literature and scholars who have studied classical literature have read a lot. Most modern people are familiar with classical works. In addition, classical works are recorded in Chinese characters, not in Korean, and young people who use only Hangul today use longer to understand the works.

Following the contents of Heo Nan-seol-heon's work mentioned in Chapter 3, it is possible to create various contents by re-interpreting and quoting the original work of Heo Nan-seol-heon and receiving them more easily from young people and foreign-

ers. Classical literature works are easily received by modern people by mixing content or storytelling methods. It seems that it is the current trend to create more interesting content after using the media than reading the original text. Even if this is not easy to realize, the combination of classical literature and content, and the combination of classical literature and database can be said to be a future trend.

References

1. Kim Eung-kyo, "Memory of Yoon Dong-ju - March 2017: Yoon Dong-ju's Perception of Yoon Dong-ju: Yoon Dong-ju Research 10", "Research of Korean Literature", 62, Korean Literature Research Society, 2017, pp. 320-323.
2. Ryu Eun-young, "Narrative and Storytelling: From Literature to Cultural Content", "Humanities Content", 14, Humanities Contents Society, 2009, p. 231.
3. Han Hye-won, "Academic Expansion and Convergence of Literature", 46 Korean Language Research, Korean Language Literature Society, 2013, 424-426.
4. Yoon Dong-ju, "The Complete Collection of Yoon Dong-ju in the Original Contrast, Sky, Wind, Stars, and Poetry", Jeong Hyun-jong et al., Yonsei University Press, 2004, p. 325.
5. Kim Woo-chang, "The Situation of Writers under Japanese Imperialism", "Poet of a Poor Age", Minumsa, 1977, p. 55.
6. Jeong Hwa-young, Ko In-hwan, "Application and Design of Contemporary Literature Content Ontology Using Topic Maps", Volume 10, No. 6, Korea Digital Policy Society, 2012, pp. 215-216.

Organizational Layout and Optimization Model of Agricultural Logistics Industry Based on Ant Colony Algorithm

Jingjun Shu(✉)

Wuhan Business University, Wuhan, Hubei, China
(✉)Corresponding author: hubeiwuhansjj@163.com

Abstract: With the rapid rise of rural e-commerce and the steady increase in the level of rural distribution demand, due to the weak rural logistics infrastructure and imperfect distribution infrastructure and distribution system, the traditional conditions of rural logistics and distribution include uneven resource allocation, information asymmetry, and enterprises. Inability to communicate with each other, lack of knowledge sharing network and other issues. The lack of system integration and distribution has led to problems such as low efficiency, high cost, and poor performance at the logistics end, which have not been effectively solved in rural areas for a long time. Based on the ant colony algorithm, this paper studies the organizational layout and optimization model of the agricultural logistics industry. This paper analyzes the cost of rural express logistics distribution, and analyzes the design goals and principles of the optimization model. The experimental results show that the scale of rural netizens and the rural Internet penetration rate have continued to increase in the past five years, reaching 293 million and 44.71% of the population and popularization respectively. With the rapid popularization of the Internet in rural areas, rural logistics products have high expectations and great potential.

Keywords: Ant Colony Algorithm, Agricultural Logistics, Industrial Organization Layout, Optimization Model

1 Introduction

In the process of implementing the rural revitalization strategy, rural express logistics undertakes the important historical mission of activating the rural economy. With the rapid rise of rural e-commerce and the steady increase in the level of rural distribution demand, it has become a key link in the "connecting the past" in rural economic and social development. An important indicator to measure the quality of online shopping is the speed of logistics distribution, and the speed of logistics distribution is determined by the pros and cons of logistics route selection. A fast logistics route selection scheme can give customers a good shopping experience and reduce The overall cost of an e-commerce platform or my country's logistics and distribution [1-2].

In the relevant research, Tadi mentioned that logistics is the main means to effectively realize the flow of people, goods and information in the rural tourism supply chain (RTSC) and improve the competitiveness of tourism products [3]. Logistics provides material and non-material basis for rural tourism services. The

author analyzes the key issues and structures of RTSC, constructs the logistics structure of agritourism, and analyzes specific fields from the perspective of logistics processes, processes and activities. Vakhidov et al. proposed a model for calculating the braking parameters of transport and technical agricultural machinery equipped with ultra-low pressure wheels [4]. The difference between this model and the previous model is that its output parameter is not the braking efficiency, but the time difference between the front and rear axles locked. The results show that satisfying the advance locking condition of the front axle ensures the stability of the tractor movement during emergency braking, which has a positive impact on road traffic safety.

The main purpose of this paper is to study the organizational layout and optimization model of the agricultural logistics industry based on the ant colony algorithm. This paper analyzes the cost of rural express logistics distribution, and analyzes the design goals and principles of the optimization model. In this paper, the ant colony algorithm is used to solve the VRP problem, and the parameters are set based on the current status of the system. The research is of great significance for improving the process of rural logistics. In practice, the rural logistics distribution model developed in this paper provides a model for the development of rural logistics distribution; the control measures obtained from the study have important reference value for promoting rural logistics and policy making.

2 Design Research

2.1 Analysis of The Cost of Rural Express Logistics Distribution

(1) High transportation cost

The rural express logistics infrastructure is weak, its transportation organization is unreasonable, the network layout and rural residents are widely distributed, and the number of express parcels in the same area is unstable, which leads to roundabout transportation, repeated transportation and empty vehicle transportation of delivery vehicles to the countryside. At the same time, the unloaded rate of delivery vehicles to the countryside is high, resulting in high transportation costs for express delivery to the countryside [5-6].

(2) High transit costs

The rural distribution area is wide, but the rural express logistics distribution system is imperfect compared to the city, and a professional and stable distribution system has not been formed. Usually, it needs to go through multiple transfers to reach the terminal distribution network or agent. The operation leads to an increase in the transfer cost and the storage and transportation cost during the transfer process; and the problems such as the prolonged distribution time caused by this increase the time cost of distribution.

(3) The delivery cost is high

Due to the small and scattered demand for rural end distribution, the establishment of distribution outlets will increase the cost of outlet construction. In practice, most companies are unwilling and unable to set up rural distribution outlets, resulting in reduced “last mile” distribution efficiency in rural distribution; Or affected by other factors, it is often difficult to complete a one-time pickup, which leads to an increase

in the cost of secondary distribution.

(4) High cost of operation and implementation

Restricted by the objective environment in rural areas, the resources of rural express logistics and distribution are limited, the hardware foundation is poor, and the "poor, narrow, and weak" rural roads in remote areas lead to low road accessibility and high vehicle depreciation costs; To achieve interconnection, the phenomenon of waste of distribution resources is serious.

2.2 Design Goals and Principles

The design goal of this system is to improve the parts that do not conform to the business process based on the existing system functions of the distribution center. At the same time, an intelligent dispatching module is added to realize the system automatically dispatching vehicles and planning the distribution path, so as to provide the distribution center with the business process and operation. Simple information management, vehicle scheduling services [7-8].

In order to achieve the system goals, the system should follow the following principles when designing:

(1) Security principle

There is a large amount of enterprise internal information stored in the system, so information security is very important. The vehicle dispatching platform of the distribution center is used in the intranet of the enterprise, and the data transmission is also connected with the internal system of the enterprise, and the security is relatively strong. When the user logs in to the system and performs operations, the user's identity should be verified, and the user's password should be irreversibly encrypted, and the setting of multiple input wrong passwords to lock access should be activated to prevent software attacks that crack the password. The system administrator should clean up the user information of the resigned staff in a timely manner to prevent the information from being stolen.

(2) The principle of reliability

The system needs to work 24 hours a day to process orders from various e-commerce platforms at any time. If there is a failure, it can be recovered within 12 hours, and the backup data can be used to ensure the normal operation of the system [9-10].

(3) The principle of scalability

With the continuous improvement of the business scale, the business scale of the distribution company will also continue to expand. To meet business requirements at the same time, the design of the system should enable integration between small and functional units.

(4) The principle of portability

There are many urban distribution centers in logistics enterprises, and the system should be designed to be as universal as possible, so that the system can be used in local distribution centers after briefly modifying some parameters or adding or subtracting business modules.

(5) The principle of ease of use

The purpose of using the system is to improve work efficiency, and the vehicle dispatchers and managers faced by the system are usually not computer professionals. Therefore, the system needs to provide a user-friendly operation interface, minimize

manual operations, and facilitate learning and use [11-12].

2.3 Algorithm Operation Process

The steps to solve the VRP problem using the ant colony algorithm are as follows:

(1) Parameter initialization

m is the number of insects, α is the main pheromone factor, β is the heuristic activity, ρ is the vaporized pheromone, Q is the total amount of pheromone released, and n is the maximum number.

(2) Constructing the solution space

All insects are placed in a distribution center, and each insect selects the next distribution point for distribution based on the concentration of "pheromone". The calculation process is:

$$p_{ij}^{k}=\begin{cases}\dfrac{[\tau_{ij}(t)]^{\alpha}\cdot[\eta_{ij}(t)]^{\beta}}{\sum_{s\in I_k}[\tau_{ij}(t)]^{\alpha}\cdot[\eta_{ij}(t)]^{\beta}}, s\in I_k \\ 0, s\notin I_k\end{cases} \tag{1}$$

Where $\tau_{ij}(t)$ is the pheromone concentration. At the beginning of the analysis, the pheromone concentration is the same between receptor sites, so assuming $\tau_{ij}(0)=0$, is a set of receptors that do not transmit k, a heuristic function calculated by Equation (2).

$$\eta_{ij}(t)=\frac{1}{\sqrt{(x_i-x_j)^2+(y_i-y_j)^2}} \tag{2}$$

The larger the number of heuristic activities, the higher the probability of insect selection until all insects have completed the delivery of all collection points and returned to the distribution center.

(3) Update pheromone

After the search is completed, the path length of each insect is calculated, the shortest path in the current iteration is recorded, and the pheromone concentration TAry(i,j) between each receiving point is updated according to formula (3):

$$\begin{cases}\tau_{ij}(t+1)=(1-RHO)\tau_{ij}(t)+\Delta\tau_{ij} \\ \Delta\tau_{ij}=\sum_{k=1}^{n}\Delta\tau_{ij}^{k}\end{cases} \tag{3}$$

$\Delta\tau_{ij}^{k}$ represents the pheromone concentration, k is the number of animals, and i, j are the receiving points.

The calculation formula (4) of the antcyclesystem model is as follows:

$$\Delta\tau_{ij}^{k}\begin{cases}Q/Length_{ij}, \text{The kth ant visits j from the receiving point i} \\ 0, \text{other}\end{cases} \tag{4}$$

(4) Judging termination conditions

When the number of iterations reaches the preset maximum number of iterations, stop the work and obtain the optimal solution; otherwise, delete the insect path record and resume the second step.

3 Experimental Study

3.1 System Status Analysis and Parameter Setting

Through investigation, it is found that the distribution center has a complete logistics information management system, and the informatization level is higher than the industry average. First, the existing system of the distribution center will be analyzed. If the system calls the algorithm proposed above for vehicle allocation, the following parameters are required:

(1) Address number

(2) The latitude and longitude coordinates of the delivery address

(3) The total volume of goods to be delivered at the receiving point

(4) Available fleets, number of available franchise vehicles and corresponding models

However, during the investigation, it was found that the data management functions of the existing system have the following defects:

(1) The address library is not fully utilized

According to the survey, the system is embedded with a GIS system, which can automatically obtain the information of provinces, cities, districts, counties, and latitude and longitude of the receiving address and mark it on the map. However, in the actual car distribution process, the address database information is only used to divide the order area according to the administrative district and county where the delivery address is located, and the rest of the data is not fully utilized.

(2) Special orders cannot be marked through the system

In the process of allocating vehicles, the dispatcher is sometimes required to handle orders with special needs alone, but such orders cannot be identified by the system, and can only be obtained by the dispatcher recording the order number through manual inquiry, and corresponding processing. Therefore, the system cannot judge the order entering the automatic vehicle distribution process (normal order), and thus cannot calculate the total volume of the goods to be delivered at the corresponding receiving point.

(3) The vehicle management is chaotic and the information is not updated in time

The distribution center carrier is divided into two categories: fleet and franchised vehicles, but the system has loopholes in the management of the two. First of all, the basic information is incomplete, and the system cannot reflect the actual docking situation of the fleet and the franchised vehicles. Secondly, the system cannot reflect the real-time status of the vehicle, and cannot know through the system whether the participating vehicle is currently on the way or idle, and whether the fleet can undertake the delivery task. The above information is obtained by the dispatcher through offline inquiry. Therefore, the system cannot provide the available fleet, the number of available franchise vehicles and the corresponding model.

The above problems all lead to the inability of the system to provide corresponding

parameters for automatic vehicle allocation, and are also the key problems to be solved in the following system analysis and design.

3.2 Functional Module Requirements

According to the business requirements and role analysis of the distribution center, the vehicle scheduling platform should include the following six functional modules, namely: login module, user management, static data management, order management, vehicle scheduling, and document management. The specific functional requirements are shown in the following figure 1:

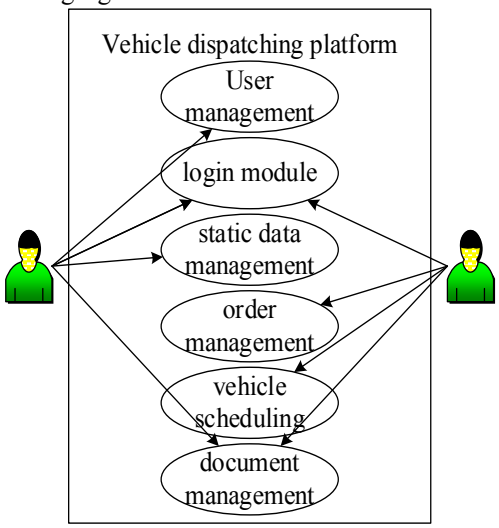


Figure 1. The overall functional module requirements of the vehicle dispatching system

(1) User management

User management is mainly used to record and verify the identity of the system user, and the user needs to log in with the user name and password. The account number and initial password are uniformly assigned by the system administrator, and the user can change the password after logging in for the first time.

(2) Static data management

Static data management is mainly used to record data with high frequency and low update frequency in the system, including fleet information, franchise vehicle information and address database (for B2B business, customers are relatively fixed, so addresses are regarded as static data). The administrator can add, modify, delete and query the data of the joined car, fleet and address database. In order to ensure the stability of the system operation, the vehicle dispatcher can only query static data as required, and cannot perform other operations.

(3) Order management

The orders of the distribution center come from major e-commerce websites, which are uniformly processed by the order processing system and converted into standard formats and then directly imported into the vehicle dispatching platform. Therefore, there is no need to do anything with the generated order unless there are special circumstances. If there are special circumstances (such as expedited delivery, etc.), the

system administrator and the vehicle dispatcher can mark the order specially, and the dispatcher will arrange it separately when arranging the vehicle delivery.

(4) Vehicle scheduling

The vehicle dispatching module is the core of the whole platform and is only operated by the vehicle dispatcher. The vehicle dispatcher needs to manually allocate vehicles for the specially marked orders according to the original operation process. For ordinary orders, no redundant operations are required, and the platform automatically allocates the delivery vehicles according to the algorithm and issues the rush orders. After the order grab is over, the system will automatically generate a dispatch order, a delivery order and a delivery order, and the dispatcher can confirm and print it after confirming that it is correct.

(5) Document management

Document management is mainly used to record the execution of the order. After the driver completes the delivery work and returns the delivery order, the system will confirm it, forming a closed-loop operation. The dispatcher can query the carrier and delivery status of the order according to the document management.

3.3 Non-Functional Requirements

(1) Response speed requirements

Since the distribution center has strict requirements on delivery timeliness, the response time of the system should be fast, and it should not take more than 10 minutes from order input to completion of vehicle distribution.

(2) Input and output requirements

The system input mainly comes from two aspects, one is the order information entered by the order management system and the order grabbing result returned by the order grabbing system, and the other is the administrator and operator, which requires a system interface and a friendly input interface.

The system output mainly includes printing documents, sending orders for grabbing orders and sending order distribution results to the WMS system, so the system needs to have a system interface and connect to a printing device.

4 Experiment Analysis

4.1 Characteristics of Rural Express Logistics and Distribution

It is mainly reflected in the huge market potential and the booming of rural e-commerce. The following are the statistics on the scale and popularization of rural netizens in the past five years as shown in Table 1:

Table 1. Statistics on the scale of rural netizens and the Internet penetration rate in the past five years

years	1	2	3	4	5
Scale of rural netizens (100 million people)	2.11	2.27	2.55	2.71	2.93
Internet penetration rate in rural areas (%)	36.10%	37.23%	39.54%	41.21%	44.71%

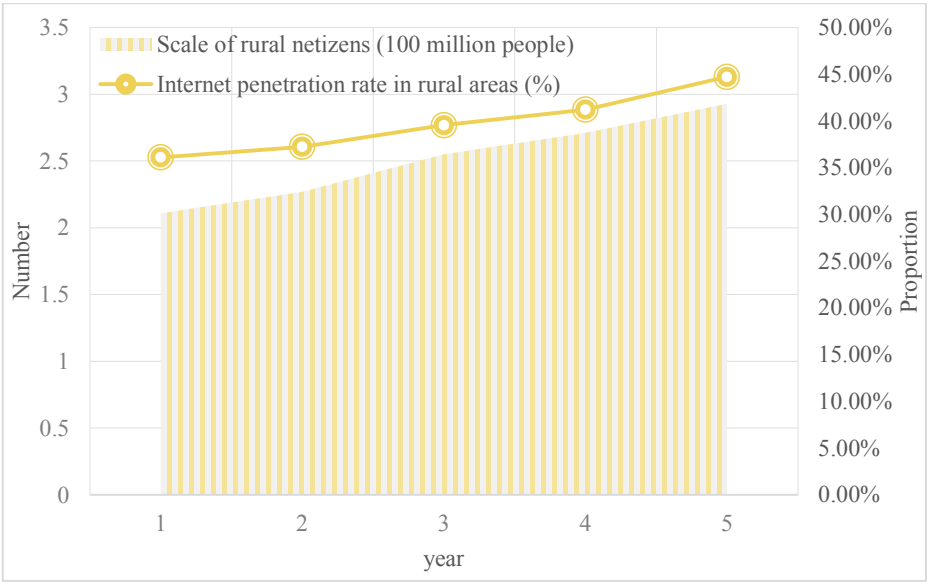


Figure 2. Analysis of the scale of rural netizens and the Internet penetration rate in the past five years

Analysis of Figure 2 shows that the scale of rural netizens and the rural Internet penetration rate have continued to increase in the past five years, reaching 293 million and 44.71% of the population and popularization respectively; The use of retail, whether it is rural online retail sales and market share, is also increasing year by year, and rural logistics products have high expectations and great potential.

4.2 Delivery Service Issues

(1) Problems occur from time to time

Due to the low level of specialization in the logistics distribution of rural express terminals and the uneven quality of human distribution, component problems often occur in the distribution of rural logistics terminals. The parcels of express parcels are damaged, the quantity of goods is in short supply or even lost. The percentage of problem pieces in the survey is as follows Table 2 and Figure 3 .

Table 2. Proportion of problem parts of express shipments and statistics of complaint handling of problem parts

	express problem			attitude towards results		
problems and attitudes	lost	damaged	none	dissatisfied	generally	satisfy
proportion	5.9%	26.2%	67.9%	33.7%	39.2%	26.1%

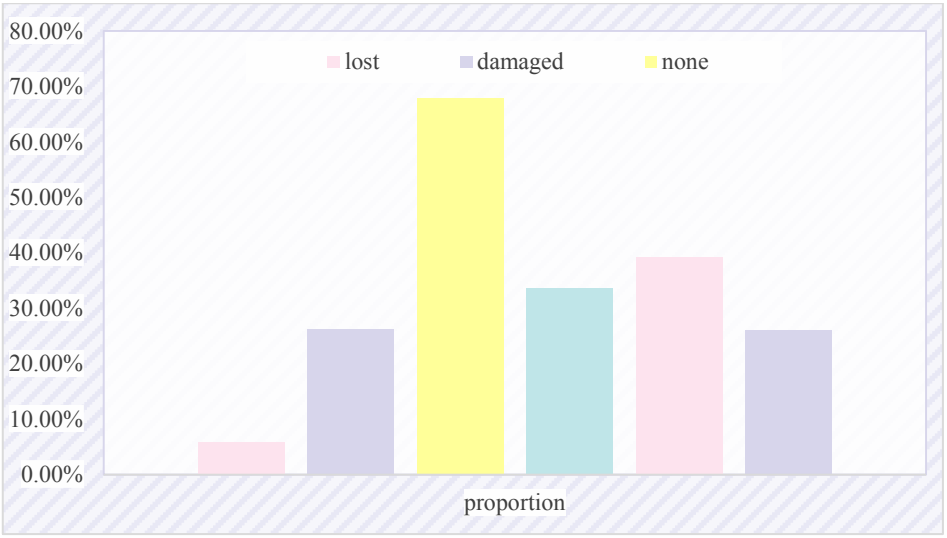


Figure 3. Analysis of the proportion of problem pieces of express shipments and the results of complaint handling of problem pieces

Due to the large number of delivery links at the end of rural express logistics, ineffective transit links and long delivery time, it is difficult to divide the boundaries of responsibilities in the delivery process. Complaint handling issues with low satisfaction.

5 Conclusions

Rural express logistics is a concept of reduced regional logistics. Rural express logistics is to serve the vast number of rural residents in rural areas, including cargo handling, packaging, storage, sorting, distribution, delivery, distribution processing and information services. and a series of logistics activities to meet the fast demand of rural residents for the growing production and living materials. With the popularity of online shopping, more and more people choose to buy the goods they need online, but the speed of receiving the goods will directly affect the user's choice direction. The quality of the logistics path selection is an important factor in determining the delivery speed of the goods. How to establish an optimized logistics distribution route selection scheme is an important issue.

Acknowledgements

The National Social Science Fund of China:Research on the coordinated development mode of agricultural industry system based on agricultural logistics park (2018BJY138)

References

1. Olkhova M, Roslavtsev D, Matviichuk O, et al. City Delivery Routes Planning Based on the Ant Colony Algorithm. *Science & Technique,* 2020, 19(4):356-362.
2. Bhutta M, Ahmad M. Secure Identification, Traceability and Real-time Tracking of Agricultural Food Supply during Transportation Using Internet of Things. *IEEE Access*, 2021, PP(99):1-1.
3. Tadi S, Veljovi M. Logistics of Rural Tourism. *International Journal for Traffic and Transport Engineering*, 2020, 10(3):323-350.
4. Vakhidov U S, Kurkin A A, Levshunov L S, et al. Ensuring the Stability of Agricultural Transport and Technological Machines Equipped with Ultra-Low Pressure Tires during Braking. *Engineering Technologies and Systems*, 2020, 30(4):609-623.
5. Radovi S. Innovative solution in the modeling of food distribution channels as a factor of successful organization of agricultural production. *Poljoprivredna Tehnika*, 2020, 45(3): 38-43.
6. Antara M, Sumarniasih M S. Featured Food Commodities For Food Security Support In Bali Province, Indonesia. *Agricultural Social Economic Journal*, 2020, 20(2):147-158.
7. Agholor A I, Nkosi M. Sustainable Water Conservation Practices and Challenges among Smallholder Farmers in Enyibe Ermelo Mpumalanga Province, South Africa. *Journal of Agricultural Extension*, 2020, 24(2):112-123.
8. Barakat A, Saad N A, Hammad M A. Key Performance Indicators of Cold Supply Chain Practices in Agriculture Sector Empirical Study on the Egyptian Export Companies. *The Open Industrial & Manufacturing Engineering Journal*, 2020, 14(10):1002-1006.
9. Kupalova H I, Goncharenko N V. State Stimulation of the Development of Organic Crop Production in Ukraine. *The Problems of Economy*, 2020, 2(44):144-152.
10. Medeiros D L, Kiperstok A C, Nascimento F, et al. Human urine management in resource-based sanitation: water-energy-nutrient nexus, energy demand and economic performance. *Sustainable Production and Consumption*, 2021, 26(-):988-998.
11. Akn Y, Elen B, Elen M F, et al. Agriculture And Pandemic: How Should Turkish Agriculture Change After COVID-19?. *EJONS International Journal of Mathematic Engineering and Natural Sciences*, 2020, 4(16):904-914.
12. Nguyen T D, Nguyenquang T, Venkatadri U, et al. Mathematical Programming Models for Fresh Fruit Supply Chain Optimization: A Review of the Literature and Emerging Trends. *AgriEngineering*, 2021, 3(3):519-541.

Application of Image Recognition in Equipment Monitoring

Haidong Zou(✉), Shaoqiang Yang and Wei Wu

China Satellite Marine Tracking and Control Department, Jiangyin, Jiangsu, China
(✉)Corresponding author: abigkun@163.com

Abstract. To solve the incomplete remote monitoring status of equipments, video streaming is used to monitor the indicators of equipment in the front panel. With the help of image recognition technology, the working status of equipments can be automatically get. By using methods such as image binarization, grayscale processing, positioning and calculation, the status of the equipment indicators are analyzed. The corresponding working status of equipments is automatically determined to further improve the equipment status monitoring. The effectiveness of image recognition technology in equipment status monitoring is verified through practical tests in this paper.

Keywords: Image Recognition, Statement Monitoring, Automatic, Video

1 Introduction

Due to the inadequate remote monitoring data of some devices, it is difficult for the managers to judge the working status of this device, which in turn affects the verdict on the working status of the whole system. Such problems occur from time to time, which is not conducive to the unattended demand of the machine room. Take a certain type of inverter as an example, its remote monitoring data does not show the working status of the equipment, but only provides the working parameter settings of the equipment. In case of equipment failure, it requires management personnel to dispose of it on site, which seriously affects the efficiency of problem disposal. However, the content of the panel indicator of the device is relatively rich, including information such as power indication, this vibration alarm, alarm storage, etc. From the field working status of the indicator of the device panel, you can basically judge the working status of the device quickly. If the panel indicator field work status, in the form of data communication to remote management personnel, can greatly enhance the system fault disposal efficiency.

Given the richness of current video surveillance means and the development of target recognition technology based on video surveillance, its application to practical problems can improve the processing efficiency [1]. In this paper, we intend to use video surveillance to monitor the status of equipment panel instructions in real time, and then realize the remote monitoring function of the working status of some equipment through automated technologies such as image recognition and data processing. This helps to improve the efficiency of equipment management and to provide technical support for the realization of unmanned server rooms.

2 System Architecture

The system is implemented by a combination of hardware and software. The status indicator information of the device panel is collected through a webcam and transmitted to the computer at the remote end. Then, through software programming, the recognition area of the captured image is cut out and processed as much as possible without missing key information. Then the status indicators in the cut image are segmented by edge recognition and other technologies, and the indicator working status library is established and the indicator status judgment threshold is calculated. Finally, the current working status of the device is determined by combining the status information of all indicators. The specific implementation architecture is shown in Figure 1.

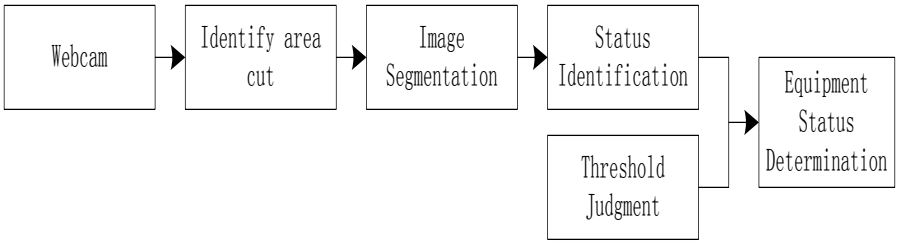


Figure 1. Architecture of the system

3 System Key Technology

Based on image recognition technology, the equipment working panel status monitoring system needs to solve the technical problems of equipment panel area recognition and panel indicator area segmentation, indicator status recognition and equipment working status determination, and so on.

3.1 Identification Area Determination

By installing identification tags on the equipment identification area, it is easy to quickly determine the identification area[2-3]. For fixed video surveillance, the corresponding equipment panel is relatively fixed, so the recorded equipment working panel image is also fixed. That is, the relative position of the equipment panel in a monitoring picture is fixed. After the precise measurement of the acquired image, we can get the relatively accurate equipment working panel area, and remove the irrelevant area to improve the efficiency of subsequent image processing.

3.2 Image Segmentation

There are many indicators in the working panel of the equipment, so it is necessary to segment them according to the work requirements to improve the image calculation speed. Fortunately, the shape of the indicators of the equipment panel in this system is relatively regular, usually mainly round indicators, and the boundary of the area between indicators is relatively obvious. According to the rules and characteristics of image processing, this paper adopts the image segmentation method based on edge detection to realize the image segmentation of panel indicators [3-5].

3.3 Image State Recognition

The state of the device panel indicator is only bright and off, but its color distribution is richer, commonly used are red, green, yellow, white, orange, etc.. To accurately determine the state of the indicator, it is necessary to first determine the value of each color indicator in the current lighting situation in the off state as the reference value 1, and then take the value of each color indicator in the on state as the reference value 2, with the difference between the two reference values to determine the threshold value[6]. The subsequent judgment of the working status of all indicators is based on the judgment threshold determined by their respective positions. The basic algorithm is as follows:

a. Determine the reference value of the indicator off state 1. For the N images acquired, calculate the mean values ri, gi, and bi for each indicator in the area R, G, and B channels that are off, and then calculate the mean values rm, gm, and bm for these three channels as the reference value 1.

$$ri=mean(mean(alpha\_r(pos\_on))) \tag{1}$$

$$gi=mean(mean(alpha\_g(pos\_on))) \tag{2}$$

$$bi=mean(mean(alpha\_b(pos\_on))) \tag{3}$$

$$(rm, gm, bm) = \frac{1}{N}\sum_{i=0}^{N-1} (ri, gi, bi) \tag{4}$$

b. Determine the reference value for the indicator on state 2. for the N images acquired, calculate the mean values Ri, Gi, Bi for each indicator in the area R, G, B channels that are off, and then calculate the mean values Rm, Gm, Bm for these three channels as the reference value 2. since the calculation method is the same, this step can continue to use the four calculation formulas in the above step, only the distinction needs to be made for the off state indicator light value.

c. Determine the judgment thresholds (Rt,Gt,Bt). The difficulty of this step is that each indicator due to color differences and the impact of various possible relationships at the scene, resulting in their respective judgment thresholds have a large difference, so in determining the threshold value, the need to have tolerance considerations. The respective calculated base value 1 and base value 2 can be used as the final qualitative judgment indicator by taking the median value of both for the judgment threshold.

$$(R\,t, G\,t, B\,t) = mean\,((rm, gm, bm),(Rm, Gm, Bm)) \tag{5}$$

d. The result is obtained by comparing the current state value with the judgment threshold. If the current value is greater than the judgment threshold, the indicator will be turned on, otherwise it will be turned off.

$$state = ((R,G,B) > (\mathrm{Rt},\mathrm{Gt},\mathrm{Bt})\,?1:0) \tag{6}$$

3.4 Equipment Status Determination

The determination of the operating status of the device requires a comprehensive

judgment combined with the operating status values of each indicator as described above. The basic working logic of the program is shown in Figure 2.

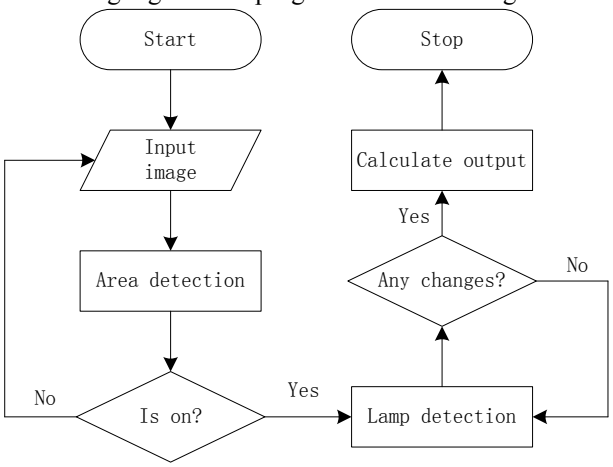


Figure 2. Flow chart for determining the working status of the equipment

After the program is started, it will ask for the input of the picture to be processed, and after the detection of the positioning of the relevant working area in the picture is completed, the recognition program is started to judge whether the equipment is already in working state[7-9]. If the result is judged to be true, it enters the monitoring stage of the equipment status indicator, focusing on whether there is any abnormality in each key status indicator, and then outputs the status determination result of the equipment in this way[10].

4 System Test

Taking an inverter as an example, the site uses a webcam from Beiqingshitongwith an effective pixel of 3 million. By calculating the working indicator area of the device and then cutting the picture, the working area of the device panel indicator is obtained, as shown in Figure 3(a). As can be seen from the figure, the power indication, remote control and internal reference source indicators are green[9-10], while the local vibration alarm and alarm storage indicators are red, and the color of each indicator differs slightly due to its location, so it needs to be segmented. After using the edge detection method, the segmentation effect is shown in Figure 3(b).

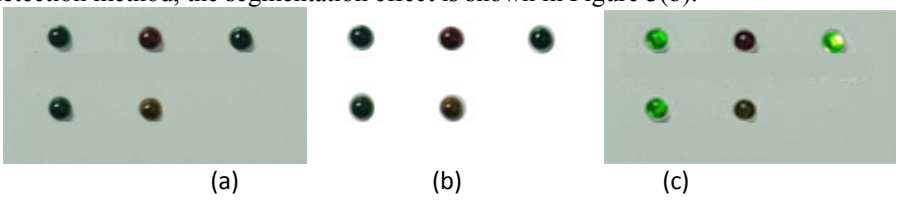


Figure 3. (a)Working area map (b)Splitting effect map (c)Test map

As Figure 3(a) is the device shutdown status indicator display, after taking the value of each indicator separately, we can get its respective base value 1. Similarly, take the value of each indicator after it is lit, and then calculate the respective base value 2, and then use it to calculate the judgment threshold. For example, if the "Benzene Alarm" indicator has a reference value of 55 and a reference value of 125, then the threshold value for determining its operating status is 90.

When the test chart is shown in Figure 3(c), after processing and calculation, the system comes up with the current device status: the equipment is on and working normally.It is in line with the actual situation.

5 Conclusions

By using of image recognition technology, with the status recognition of the indicatorin front of equipment panel, the status of the equipment can be checked. It can help the management to effectively dispose of the equipment problems in a timely manner and improve the system operation and maintenance efficiency. The system employs frame-splitting processing of the video images captured by the webcam to cut the effective positions in the images in order to improve the efficiency of subsequent image processing. An edge detection algorithm is used to extract the position of each indicator, and then calculate the respective state value separately, compare it with the corresponding judgment threshold, and finally determine the working status of the device. The method has good generality and has a certain degree of generalization. Subsequently, the research on state monitoring in different environments and under different lighting influences should be enhanced to improve the applicability of the system.

References

1. Fagundes L G, Santos R. Development of Computer Graphics and Digital Image ProcessingApplications on the iPhone, *Graphics, Patterns and Images Tutorials (SIBGRAPI-T)*, 2010 23rdSIBGRAPI Conference on. IEEE Computer Society, 2010.
2. Goyal A, Meenpal T. Patch-Based Dual-Tree Complex Wavelet Transform for Kinship Recognition. *IEEE Transactions on Image Processing*, 2021, 30:191-206.
3. Chakraborty, S., Chatterjee, A.&Goswami, S. K.(2014). A dual-tree complex wavelet transform-based approach for recognition of power system transients. *Expert Systems*, 32(1), 132–140.
4. Park J,Low C Y,Andrew B. Divergent Angular Representation for Open Set Image Recognition. *IEEE Transactions on Image Processing*, 2022, 31.
5. Chakraborty, et al. "A dual-tree complex wavelet transform-based approach for recognition of power system transients." *Expert systems: The international journal of knowledge engineering 32.1*(2015):132-140.
6. Dubnov Y A. The Feature Selection Method Based on a Probabilistic Approach and a Cross-Entropy Metric for the Image Recognition Problem. *Scientific and Technical Information Processing*, 2022, 48(6):430-435.
7. Keys R G. Cubic convolution interpolation for digital image processing. *IEEE Transactions on Acoustics*, Speech, and Signal Processing, 2003, 29.
8. Chris Solomon, Toby Breckon. Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab. John Wiley&Sons, Ltd,2011.

9. Mahmood A,Khan A Q,Mustafa G, et al. Remote Fault-Tolerant Control for Industrial Smart Surveillance System. *Mathematical Problems in Engineering*, 2021.
10. Varma N, Epstein A E, IrimpenA, et al. Efficacy and safety of automatic remote monitoring for implantable cardioverter-defibrillator follow-up: the Lumos-T Safely Reduces Routine Office Device Follow-up (TRUST) trial. *Circulation*, 2010, 122(4):325-332.

Stability of Marine Physics Detection Sensor Based on Artificial Intelligence Technology

Xiran Liu(✉)

Changchun College of Electronic Technology, Changchun, Jilin 130114, China
(✉)Corresponding author: liuxiran1122@163.com

Abstract. In order to maintain maritime safety, marine physical exploration plays an important role in it. For the specific application of marine physical detection, marine monitoring sensors are deployed in the extremely complex and variable marine environment to realize real-time monitoring of the ocean. Therefore, the combination of artificial intelligence technology is of great significance to the design and stability research of marine physical detection sensors. The purpose of this article is to study the stability of marine physical detection sensors based on artificial intelligence technology. This article introduces the functional modules and software system of the sensor, and analyzes the stored data of each sensor. This article tests the stability of the entire marine physical detection sensor, simulates the actual environment to measure temperature and electrode information, compares and analyzes the experimentally measured data, and draws a response conclusion. Experimental test results show that during 1-30 minutes, the temperature measured by the sensor fluctuates between 4.70031℃-4.69890℃, and the resolution of the temperature detection module can reach at least five decimal places. It can be seen that the performance of the sensor is stable, and the measurement accuracy basically meets the requirements of use.

Keywords: Ocean Exploration, Seabed Observation Network, Sensor Stability, Artificial Intelligence

1 Introduction

Ocean development is inseparable from the development of ocean exploration technology [1-2]. The marine industry is rising day by day, and the research on marine physical exploration is also receiving more and more attention [3-4]. Among them, the ocean magnetotelluric method reflects the distribution of subsea media through electrical parameters, which can provide more valuable information [5-6]. However, the marine environment is complex and changeable, and the seawater has great interference to electromagnetic signals, making it more difficult to obtain valuable signals [7-8]. Therefore, it is of great significance to develop a high-performance marine physical detection sensor and complete its performance stability test.

Regarding the research of ocean exploration, many scholars at home and abroad have conducted multi-directional and in-depth discussions on it. For example, Liu CH uses optical video image processing technology to intelligently identify and classify weak targets on the sea and non-ocean waves [9]; Bell K proposed a detection scheme

that combines coarse and fine detection of ship targets [10]; Wagner combines radar digital signal processing with machine learning to realize an efficient algorithm for exploring marine targets [11]. It can be seen that since the development of ocean exploration technology, the scientific development of its related technologies has been concerned by the majority of researchers. Therefore, this article combined with artificial intelligence technology is of great significance to the subject of marine physical detection sensor stability research.

This article aims to study the stability of marine physical detection sensors based on artificial intelligence technology. This article first introduces the functional modules and software of the sensor, including the processing module, sensor module, positioning module and other hardware. Then the stability of the sensor is tested. The experimental test results verify that the sensor has stable performance, and the measurement accuracy basically meets the requirements of use.

2 Stability of Marine Physical Detection Sensors Based on Artificial Intelligence Technology

2.1 Marine Physical Detection Sensor Hardware and Functional Modules Based on Artificial Intelligence Technology

(1) Processing module

The processing module includes functions such as power management, distributed processing and storage. The power management optimizes the sensor node in design, extends its life cycle, and manages the power supply. The main controller of the processing module adopts the CC2531 processing chip, which has the characteristics of low power consumption. Distributed processing is the processing of data. When the collected data suddenly becomes larger or smaller, the data will be collected many times by itself, and then the average value will be taken. The function of storage is to save certain parameters and key data in the sensor node.

(2) Sensor module

The sensors in this study, whether they are analog sensors or digital sensors, can be connected to the node through a common interface. The design of the universal port is to design the analog interface, serial digital interface and parallel digital interface into a universal module in the interface [12].

1) Positioning module

In marine physical exploration, the sensor will drift due to the interference of weather and natural factors. Therefore, the position of the sensor node must be known in advance during deployment, that is, the relative position of the sensor node is entered into the node or the ID number record of the node is translated into a valid position. This can effectively reduce the power consumption generated by the sensor node's own positioning, reduce the cost, and make the positioning accuracy more accurate. Therefore, in this module, the GPS/mobile base station is omitted.

2) Power management

The power supply mode of the sensor node is: a rechargeable lithium battery is combined with a solar battery. When the solar energy is sufficient, it will supply power to the sensor node and charge the lithium battery at the same time. When solar

power is insufficient, it is powered by a rechargeable battery. The power management in the processor reads the energy of the power module at intervals to determine the power status. This design method can effectively solve the problem of energy limitation, greatly extend the life cycle, reduce the frequency of manual maintenance, and save labor.

3) Serial communication circuit

Using RS485 transmission technology, the data is output from the serial port and transmitted to the deck via the ready-made RS232 to RS485 module, and then sent to the client software via the RS485 to RS232 module.

(3) Scheme design of temperature detection module

This design uses two-wire molybdenum resistors to form a Wheatstone bridge circuit. The voltage reference chip AD780 of this system is used as the power supply of the bridge, so that the current through the uranium resistor is not more than 1mA. The ADS1256 analog-to-digital converter is selected, which has the advantages of low noise, high resolution, high performance, high precision bits with built-in gain, and perfect self-correction and system correction functions.

(4) Scheme design of in-depth inspection module

The rated power supply current of the selected pressure sensor is 0.5~2mA, and the constant current of the designed constant current source is a typical value of 1.5mA. Because the pressure sensor has a built-in Wheatstone bridge, no additional circuit is required. The analog-to-digital converter is also used It is a high-precision 24-bit A/D analog-to-digital converter ADS1256.

(5) Magnetic strength detection module

The magnetic strength detection module mainly includes a constant voltage source circuit, a magnetoresistive sensor, and an A/D mode converter with built-in gain. The constant voltage source module and the analog-to-digital conversion circuit all use the same device model in the temperature and depth detection module. The resistance sensor is a three-axis magnetoresistive sensor, which can measure the parameters of geomagnetic intensity.

2.2 Software Design

(1) Data storage

When data is stored, the experiment-related information and sensor voltage signal data transmitted from the acquisition module are respectively saved. The file format of the data stored in the SD card is TXT, and each file is named by time. Each time the file is opened for writing, the address pointer stays at the end of the last written file. Each write operation sequentially writes a 150-bit array, storing 60-bit GPS information, 48-bit attitude sensor data, 36-bit AD conversion data and 6 newline escape characters.

(2) Data browsing

The main function of this module is to read the files saved by the data storage module. Through the channel selection button, the data of different channels can be displayed in the waveform graph. At the same time, two cursors are set in the waveform graph. By dragging the two cursors to select the head value and the tail value as the selected interval, the waveform in the specified interval can be cut out to enlarge the display, and the relevant information of the waveform can be obtained, such as maximum, minimum, mean and variance.

(3) Magnetic sensor data processing

The magnetic sensor has a zero adjustment resistance in each measurement direction. After power on, the AD differential channel is connected to the direct differential output of the magnetic sensor to set the zero adjustment, that is, it is considered that the magnetic field strength is zero in the current state, and the magnetic substance appears the change in the magnetic field caused by time is the measured value. Due to the working environment of the buoy, the energy detector is used to determine the threshold change for the Z axis and XY axis, and the change is recorded as 1, and the corresponding azimuth angle is calculated. At the same time, the vector sum is calculated to eliminate the energy change caused by its own rotation. The magnetic sensor data processing process occurs after the data is transmitted via the wireless transceiver module.

(4) Vector hydrophone data processing

The working principle of the vector hydrophone is that when a sound wave reaches the sensitive structure, the plastic cilia cylinder will vibrate and resonate with the acoustic signal, and the cilia cylinder will slightly swing in four directions, which in turn drives the cantilever beam below to deform. Due to the piezoelectric effect, the resistance value on the cantilever beam will also change, so that the underwater acoustic signal is converted into an electrical signal.

This study selects the power spectrum analysis method to process the hydrophone data. The process is as follows:

First perform Fourier transform on the signal, then square the modulus of the obtained amplitude spectrum, and then divide by the duration to estimate the power spectrum of the signal, as shown in formulas (1) and (2):

$$X(\mathrm{e}^{j\omega}) = \sum_{n=0}^{N-1} x(n) e^{-jwn} \tag{1}$$

$$P(\omega) = \frac{1}{N} \left| X(e^{j\omega}) \right|^2 \tag{2}$$

In the formula, $x(n)$ represents the time domain vector of the signal, $X(\mathrm{e}^{j\omega})$ represents the frequency domain of the signal, and $P(\omega)$ represents the power spectrum.

3 Experimental Research Design

3.1 Experimental Equipment and Environment

Signal acquisition instrument: DAQ2010 multifunctional data acquisition card;
Experimental environment: refrigerator

3.2 Experimental Project

(1) Experiment 1: Electrode sensitivity test

In a normal temperature environment, keep the distance between the two electrodes constant, and pass signals with amplitudes of 5mV and 30mv into the water tank with a frequency of 1Hz. Use Ag/AgCl electrodes to detect this signal and display the

output result on an oscilloscope. This experiment uses two eDAQ potentiostats, one as a signal generator and the other as an oscilloscope.

(2) Experiment 2: Temperature measurement and debugging experiment

In the refrigerator constant temperature experiment environment, place the container full of water in the refrigerator for a whole day. After the water temperature is consistent with the temperature in the refrigerator, put the hardware circuit in the refrigerator, and then put the sensor in the water for temperature measurement experiment. After 30 minutes, take out the hardware circuit and record the collected data.

4 Analysis of Experimental Results

4.1 Sensor Electrode Stability

Experiment 1 was performed 8 times. Table 1 shows the results of the signals detected by the electrodes in different signal amplitudes. It can be seen that the signals detected by the electrodes are relatively stable in the environment of 1Hz and 5mV; in the environment of 1Hz, 30mV, the signal detected by the electrode fluctuates up and down.

Table 1. Electrode detection signal

Experiment number	1Hz,5mV	1Hz,30mV
1	150	198
2	147	210
3	149	225
4	151	212
5	150	200
6	144	214
7	149	229
8	147	224

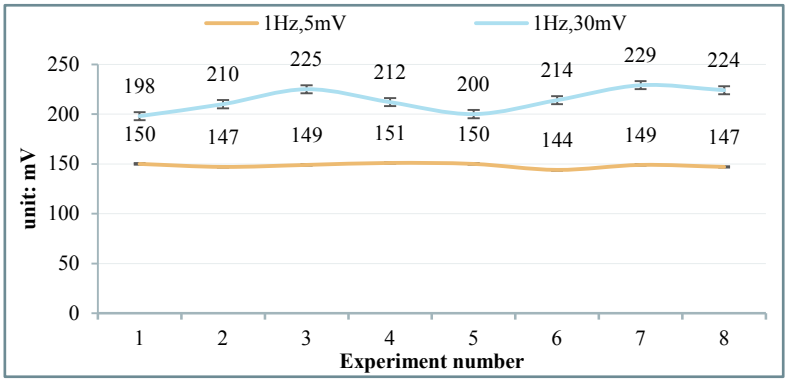


Figure 1. Electrode detection signal

It can be found from Figure 1 that due to the limitation of the test instrument, the

signal generated by the signal generator is interfered by the power frequency signal. Under the environment of 1Hz and 30mV, the signal detected by the electrode has a DC drift of 200mV. However, no matter the amplitude or phase, the detection signal has no distortion. Therefore, the sensor electrode proposed in this study can sensitively detect the strength of the signal, and can be used to detect the abundant electric field signals existing in seawater.

4.2 Data Analysis of Temperature Measurement and Debugging

In the second experiment, the unreasonable data generated by manual operation was eliminated, and the results are shown in Table 2: during 1-30 minutes, the temperature measured by the sensor fluctuates between 4.70031°C and 4.69890°C. The temperature data in Table 2 shows that the resolution of the temperature detection module can reach at least five decimal places, which meets the design requirements of the system.

Table 2. Temperature data measured within 30 minutes (℃)

time	temperature	time	temperature	time	temperature
1	4.70031	11	4.69921	21	4.69899
2	4.70054	12	4.69918	22	4.69904
3	4.70054	13	4.69914	23	4.69910
4	4.70056	14	4.69909	24	4.69921
5	4.69983	15	4.69904	25	4.69842
6	4.69951	16	4.69892	26	4.69947
7	4.69924	17	4.69894	27	4.69951
8	4.69930	18	4.69890	28	4.69964
9	4.69914	19	4.69891	29	4.69971
10	4.69920	20	4.69893	30	4.69997

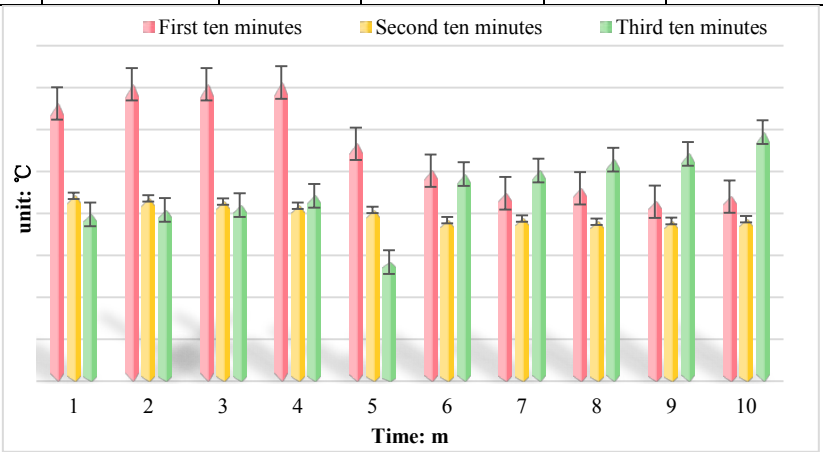


Figure 2. Temperature data measured within 30 minutes (℃)

It can be seen from Figure 2 that in the first ten minutes, the temperature data collected changes less than the curve, and the curve changes in a wave shape. This is affected by the working mechanism of the refrigerator, and its work is intermittent.

When the temperature reaches the set value, the refrigerator stops cooling. After that, the temperature in the refrigerator rises until the refrigeration work restarts. At the same time, due to the larger specific heat capacity of water, the temperature change of the water body is smaller.

5 Conclusion

With the development of artificial intelligence technology, sensor monitoring technology combined with artificial intelligence technology is a new driving force for the development of this field. Due to the unique characteristics of the marine environment, it is necessary to study marine physical detection sensors based on artificial intelligence technology suitable for the marine environment. Through research, this paper has completed the following tasks: introduced the functional modules and software system of the sensor, analyzed the stored data of each sensor; tested the stability of the entire ocean physical detection sensor, and verified the stable performance of the sensor. The characteristics and measurement accuracy also meet the requirements of use.

References

1. Glover A G. Abyssal fauna of the UK-1 polymetallic nodule exploration area, Clarion-Clipperton Zone, central Pacific Ocean: Cnidaria. *Biodiversity Data Journal*, 2016, 4(4):e9277.
2. German C R, Petersen S, MD Hannington. Hydrothermal exploration of mid-ocean ridges: Where might the largest sulfide deposits be forming?. *Chemical Geology*, 2016, 420(1):114-126.
3. Wilson W H, Gilg I C, Moniruzzaman M, et al. Genomic exploration of individual giant ocean viruses. *The ISME Journal*, 2017, 11(8):1736.
4. Lunine J I. Ocean Worlds Exploration. *Acta Astronautica*, 2016, 131(FEB.):123-130.
5. Ballard R D, Leonardi A P. New Frontiers in Ocean Exploration The E/V Nautilus and NOAA Ship Okeanos Explorer 2015 Field Season Epilogue. *Oceanography*, 2016, 29(1Suppl.):76-77.
6. Pawlenko N. Technology and Ocean Exploration. *Oceanography*, 2019, 32(1SUPPL): 100-100.
7. Valette-Silver N, Cantelas F, Beaverson C, et al. Sponsored Projects: NOAA's Office of Ocean Exploration and Research Introduction. *Oceanography*, 2019, 32(1SUPPL): 118-118.
8. Gaffney, Paul. Ocean Exploration: A Supply-Demand Mismatch. *Marine Technology Society journal*, 2016, 50(6):8-9.
9. Liu C H, Huang X, Xie T N, et al. Exploration of cultivable fungal communities in deep coal‐bearing sediments from 1.3 to 2.5 km below the ocean floor. *Environmental Microbiology*, 2017, 19(2):803-818.
10. Bell K, Copeland A, Chow J S, et al. All Hands on Deck: The 2018 National Ocean Exploration Forum. *Oceanography*, 2019, 32(1SUPPL):4-5.
11. Wagner, Katie. Ocean Exploration Celebration. *Oceanography*, 2018, 31(1 Suppl.): 52-52.
12. Voss J, Pomponi S. Cooperative Institute for Ocean Exploration, Research, and Technology. *Oceanography*, 2019, 32(1SUPPL):119-119.

Optimization System of Microbial Test on Account of Genetic Algorithm

Mingming Shao(✉)

Xi'an Medical University, Xi'an, Shaanxi, China
(✉)Corresponding author: smmxayxy@yeah.net

Abstract. Microorganism is an important part of geochemical cycle and plays an irreplaceable role in ecosystem. Optimization of microbial assay is very important. In this paper, genetic algorithm is used to optimize the microbial test. Using the operating mechanism of genetic algorithm, that is, imitating the basic laws of nature, carrying out natural selection and survival of the fittest, using this principle to treat the detection of microbial detection optimization. Through the natural selection and survival of the fittest, genetic algorithm weight adjustment, so as to achieve more accuracy of the test. By referring to the mathematical formulas (1) and (2) in Part 3 of this paper, the requirements of determining the definition of microbial detection can be achieved by initializing the population of microorganisms and analyzing the global convergence of the samples that meet the standards and do not meet the standards. This paper studies the knowledge of microbial test optimization system based on genetic algorithm, and describes the methods and principles of microbial test. The results show that the optimization effect of microbial test is improved significantly by the optimization system based on genetic algorithm.

Keywords: Genetic Algorithm, Microbial Test, Test Optimization, Optimization System

1 Introduction

Because microorganisms are ubiquitous in the air, land and water, the inspection results of microorganisms not only represent the quality of the product itself, but also reflect the sanitary conditions of the product processing environment, the health of the processing personnel, the safety of the product transportation and the reasonable conditions of storage. In addition, microorganisms are highly adaptable and easily mutated, and sometimes the mutated individual will have biochemical reaction characteristics completely different from the original individual, so the qualitative test results are also very important. The optimization system of microbial test based on genetic algorithm is beneficial to the optimization treatment of microbial test.

As for the research of genetic algorithm, many scholars at home and abroad have studied it. In foreign studies, Ortiz S proposed a genetic algorithm. Compared with existing path planning methods, the proposed path planning method has many advantages, combining sliding mode control with classical simultaneous localization and mapping (SLAM) method. This combination can overcome the bounded uncertainty problem in SLAM [1]. Et. Proposed a new fine-grained sentiment analysis

model combining convolutional neural network and random forest classifier. The continuous Word bag (CBOW) model is used for vectorizing text input. The most important features are extracted by convolutional neural network (CNN). The extracted features are used for emotion classification by random forest (RF) classifier [2]. Al-obaidi MA proposed an optimization framework based on species conservation genetic algorithm (SCGA) to optimize process design and operational parameters. In order to enable readers to have a deeper understanding of the process, the effects of membrane design parameters on xylenol retention rate, water recovery rate and specific energy consumption level under two different process conditions were studied [3].

In today's society, computer technology continues to develop rapidly, and computers have become necessary equipment for every scientific research institution. Many product inspection items are more or less began to introduce computer software to assist or replace manual operation[4-5]. However, manual operation is still used in the microbiological testing using the microbiological testing methods formulated by the Ministry of Health. The identification results are obtained through multi-step operation and the judgment results are compared with the manual standards. Of so result issue often time is long, return easy occurrence error, because this artificial judgement result begins to be challenged greatly. The optimization system of microbial test based on genetic algorithm promotes the efficiency and accuracy of microbial test.

2 Design and Exploration of Microbial Test Optimization System on Account of Genetic Algorithm

2.1 Genetic Algorithm

Genetic algorithm is one of the important algorithms, whose basic principle is to imitate the basic laws of nature, natural selection and survival of the fittest, and optimize the algorithm through these two laws[6-7]. Genetic algorithm has good application in many fields, but the application requirement is very low. It is a very effective global optimization algorithm with good adaptability when solving fuzziness of data.

The basic elements include genetic operation, coding mode, parameter selection and fitness function[8-9]. The basic solution process of this algorithm is as follows, refer to Figure 1:

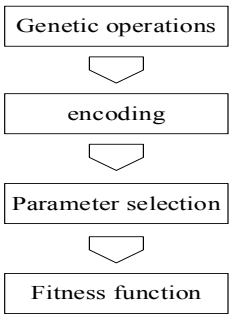


Figure 1. The basic building blocks of genetic algorithms

1) Genetic algorithm firstly transcodes data and performs binary codes, which are like gene fragments and constitute elements of genetic algorithm[10-11]. These elements are grouped into algorithmic populations by certain rules, and the process is like survival of the fittest.

2) Perform genetic manipulation on these gene fragments

3) To deal with the individual, using the selection strategy;

4) After algorithm iteration, a population will be initialized, and the global optimal solution is formed at this time.

When genetic algorithm performs iterative optimization, it must be modeled, which is a complex network model. In this paper, multiple analysis strategies are applied to the complex network model so that the inhomogeneity can be defined more accurately.

Two classical network topology models were widely used in the early stage of complex network research, as follows:

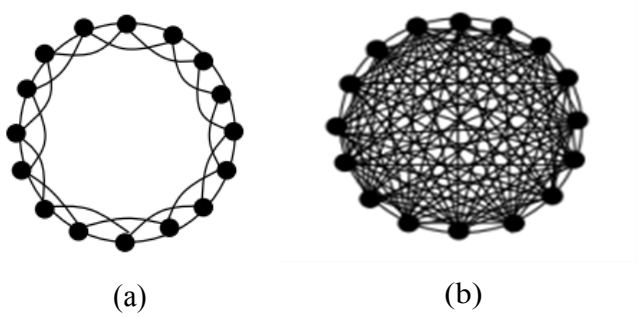


Figure 2. The basic building blocks of genetic algorithms

1) As shown in Figure 2(a), in the network model, each node only establishes edge relation with its adjacent nodes, and each node has the same number of edge.

2) As shown in Figure 2(b), in the network model, any two nodes have established edge relations, so this model is conducive to information exchange between nodes.

2.2 Optimization System of Microbial Test Based on Genetic Algorithm

Microbiological test optimization system, first of all, to conduct microbial test results analysis[12-13]. The management system needs to meet the laboratory requirements based on microbial testing methods.

(1) User demand analysis

The analysis and management system of microbial test results is mainly managed by the sample receiver or adoption personnel, the inspection personnel manage the sample test results, and other authorized personnel manage the test conclusions[14-15].

(2) Functional requirement analysis

Microbial inspection conclusion need to analyze the test results can be, usually a sample need many steps of operation, and each step will get a result step by step, finally according to these results comprehensive analysis to determine the final conclusion step by step, so the microbial inspection conclusion analysis of the workload is bigger, the staff to come to the conclusion that often requires repeated

comparison standard Therefore, it is very important to make an analysis system of microbial test results to liberate labor force.

(3) Feature requirement analysis

A system with dual functions of analysis and management of inspection results is required. Especially for arbitration inspection institutions, it is very important to issue inspection conclusions quickly and accurately[16]. In addition, the data confidentiality of the inspection conclusion is very high, and the arbitration inspection generally requires more than 3 years to keep files. Once the data is leaked, it sometimes not only damages the interests of the prosecution, but even causes social chaos. For enterprises, microbial test data not only reflect the quality of products themselves, but also reflect the environmental quality of factories and warehouses, so it belongs to the category of trade secrets. Therefore, the data security of microbial test results analysis and management system is better.

Microorganisms need to be tested before the test results can be obtained, and only after the test results are analyzed and judged can conclusions be obtained. The whole inspection process, result determination and data management must be carried out in accordance with laboratory regulations. According to the business process, the central laboratory adopts the management mode of sampling and separation. After the sample is registered and processed in the sample room, the sample taker takes it to the laboratory. After the samples are tested in the laboratory, the inspection personnel shall draw the inspection conclusion and form the inspection report, which shall be approved by the technical director, the laboratory director and the center leaders in turn and then form an official document to be sent to the person/institution being tested.

3 Exploring The Effect of Optimization System for Microbial Testing on Account of Genetic Algorithm

According to the analysis of functional requirements, the system is divided into five subsystems: system management, standard management, inspection results analysis and inspection conclusion management,as shown in Figure 3.

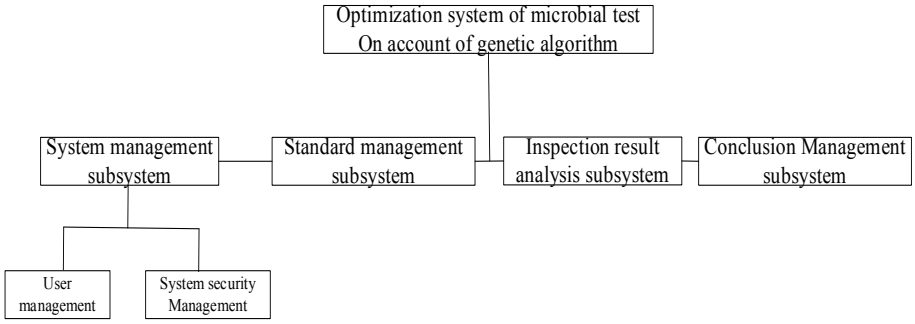


Figure 3. The basic building blocks of genetic algorithms

(1) System management subsystem

The system management subsystem mainly manages the user information and the

database of the system to ensure the security of data. The system management subsystem mainly includes two functional modules: user management and system security management.

① User Management

User information is designed to add, delete, modify, permission Settings, in order to unified management of users. The user information is unified input by the system administrator, who takes the real name, sets the login password, sets the permissions according to the department, and the permissions are set according to the four subsystems.

② System security management

The backup and recovery function of the existing database is designed to prevent the loss and damage of the database caused by human and non-human factors.

(2) Standard management subsystem

The standard management subsystem is mainly used to input product standards and provide judgment basis for analyzing test results. To ensure the authenticity and effectiveness of the input standard, it is operated by the inspection personnel. Standard management has designed the input, modification and deletion of standard information. Standard information mainly includes: standard name, standard number, internal control code, product category, standard value, release time, implementation time, status.

(3) Test result analysis subsystem

The inspection result analysis subsystem is mainly to analyze and judge the results obtained from the inspection of the sample input by the sample management system. Process: according to the management requirements of our center, the inspection personnel input the inspection results, and other personnel have no right to carry out this operation.

(4) Test conclusion Management subsystem

The inspection conclusion management subsystem mainly realizes the issue of inspection reports to the inspection results after analysis, and queries and prints related reports according to different requirements. At present, according to the common inquiry methods of our center, it can be divided into: inquiry by product category, inquiry by task source, inquiry by inspection item and inquiry by inspection conclusion. The report is printed in a fixed format according to the query mode.

$z \in x$ is the population with a size of N, and the population fitness can be calculated according to Equation (1) :

$$f(x) = \max_{z \in x} \{ f(z) \} \tag{1}$$

For any initial population B(0), if

$$\lim_{t \to \infty} p\{ f(B(t)) = y_h \} = 1 \tag{2}$$

It indicates that the algorithm has global convergence, where P (*) represents the probability of occurrence of event *, t represents the t-generation population, and y_h means that the set is divided into H subsets.

4 Investigation and Analysis of Optimization System for Microbial Testing on Account of Genetic Algorithm

This software chooses Windows XP as the design platform, uses Visual Basic 6.0 as the design language, and uses Microsoft SQL Server 2000 as the database.

Test method:

Database arbitrary CRUD operations and execution of their respective SQL queries. The database CRUD operation refers to:

C: Create: Creates a user.

R: Retrieve -- Performs the retrieve view operation.

U: Update -- Updates database information.

D: Delete: Deletes the database.

Test results:

Database main code is not empty; The outer code is equal to the corresponding main code or is null; The construction of data type, length and index is reasonable to meet the requirements of data and database integrity. All access methods and processes can operate as designed without data damage.

As shown in Figure 4, the system test data is displayed in the system test checklist. The first row of The table contains Number and up to Standard, and The first column contains The test cases. The test case contains four kinds of data, namely, Design Total number of test case sets (DS), The Number of test case sets passed completely (NP), Number of failed test case sets (NF) and Set of test cases to be tested (SCB). The table means that the Design total number of test case sets (DS) is 205 times, and the number of test case sets passed (NP) is 189 times. The Number of failed test case sets (NF) is 16, and the Set of test cases to be tested (SCB) is 0.

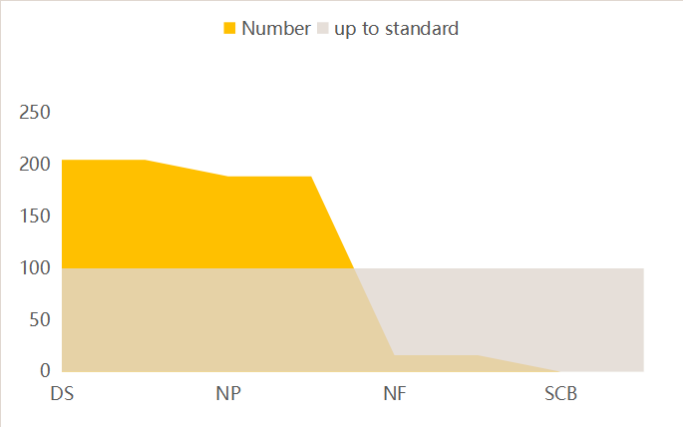


Figure 4. Test quantity chart

As shown in Figure 4, In the figure, the Design Total number of test case sets (DS) and the number of fully passed test case sets (NP) were 205 times and 189 times, respectively, far exceeding the up to standard line. The test results show that the optimization system of microbial test based on genetic algorithm is very effective.

The data show that the optimization system of microbial test based on genetic

algorithm has high performance in the optimization of microbial test.

5 Conclusions

This article through to the microbiological determine trival, fees, only a few expensive analysis instrument science problems were discussed, think development suitable for microbial detection methods formulated by the ministry of health of microbial test results analysis and management system to manage digital analysis and test results is necessary. This can not only reduce the working pressure of inspectors, but also can systematically manage the sample information and test results, but also can quickly, accurately and selectively query the test conclusion, can accelerate the pace of office automation center to a certain extent. The optimization system of microbial test based on genetic algorithm is beneficial to improve the efficiency and quality of microbial test.

Acknowledgments

2021 Shaanxi Undergraduate and Higher Continuing Education Teaching Reform Research Project (Project No. 21BY137); 2021 University-Industry Collaborative Education Program of Ministry of Education (Project No. 202102650012); 2021 Annual General Project of Shaanxi Provincial Education Science "14th Five-Year Plan" (Project Approval No. SGH21Y0262)

References

1. Ortiz S, Yu W. Autonomous navigation in unknown environment using sliding mode SLAM and genetic algorithm. *Intelligence & Robotics*, 2021, 1(2):131-150.
2. Et. A. Genetic Algorithm Based Hybrid Model Of convolutional Neural Network And Random Forest Classifier For Sentiment Classification. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 2021, 12(2):3216-3223.
3. Al-Obaidi M A, A Ruiz-García, Hassan G, et al. Model Based Simulation and Genetic Algorithm Based Optimisation of Spiral Wound Membrane RO Process for Improved Dimethylphenol Rejection from Wastewater. *Membranes*, 2021, 11(8):595-595.
4. Dolezel P, Holik F, Merta J, et al. Optimization of a Depiction Procedure for an Artificial Intelligence-Based Network Protection System Using a Genetic Algorithm. *Applied Sciences*, 2021, 11(5):2012-2012.
5. S Güler, Yenikaya S. Analysis of shielding effectiveness by optimizing aperture dimensions of a rectangular enclosure with genetic algorithm. *Turkish Journal of Electrical Engineering and Computer Sciences*, 2021, 29(2):1015-1028.
6. Nayana. Electric Vehicle Charging with Battery Scheduling and Multicriteria Optimization using Genetic Algorithm. *Journal of Electrical Engineering and Automation*, 2021, 2(3):123-128.
7. Khosravian P, Emadi S, Mirjalily G, et al. QoS-aware service composition based on context-free grammar and skyline in service function chaining using genetic algorithm. *PeerJ Computer Science*, 2021, 7(4):e603-e603.
8. Davoudi K, Thulasiraman P. Evolving convolutional neural network parameters through the genetic algorithm for the breast cancer classification problem. *SIMULATION:*

Transactions of The Society for Modeling and Simulation International, 2021, 97(8): 511-527.
9. Batayneh W, Bataineh A, Jaradat M A. Intelligent Adaptive Fuzzy Logic Genetic Algorithm Controller for Anti-Lock Braking System. *International Review on Modelling and Simulations*, 2021, 14(1):44-44.
10. Abidjan,cte d'ivoir eE cole supérieure africaine des TIC,Abidjan,Cote d'Ivoir eU niversité Felix Houphout Boigny, Abidjan,Cote d'Ivoire. Genetic Algorithm for the Pick-up and Delivery Problem with Time Window by Multi-Compartment Vehicles. *Transport and Telecommunication Journal*, 2021, 22(3):343-352.
11. Dayal S, In R, Gruplarnn R, et al. Genetic Algorithm-Based Approach for Constructing Product Groups in Make-to-Order Production Environments. *Yönetim ve Ekonomi Celal Bayar Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 2021, 20(2):259-259.
12. Ajmi N, Helali A, Lorenz P, et al. MWCSGA—Multi Weight Chicken Swarm Based Genetic Algorithm for Energy Efficient Clustered Wireless Sensor Network. *Sensors*, 2021, 21(3):791-791.
13. Indian A, Bhatia K. An Approach to Recognize Handwritten Hindi Characters Using Substantial Zernike Moments With Genetic Algorithm. *International Journal of Computer Vision and Image Processing*, 2021, 11(2):66-81.
14. Vadivoo N S, Usha B, Sudha K. A comprehensive review of current microbiological detection methods of SARS CoV-2. *Indian Journal of Microbiology Research*, 2020, 7(3):113-124.
15. Pal M, Bulcha M R, Banu M G, et al. Emerging Role of Biosensors for Detection of Foodborne Pathogens. *American Journal of Microbiological Research*, 2021, 9(3):92-95.
16. Widodo E, Pranibilan A, Ardilla Y, et al. Effect on encapsulated liquid smoke in combination with formic acid on intestinal development and microbial counts in broiler. *IOP Conference Series Earth and Environmental Science*, 2021, 788(1):012185-012185.

The Application of Virtual Reality Technology in Ophthalmology

Jingying Wang(✉), Yi Zhang

Chongqing Medical and Pharmaceutical College, Chongqing, China
(✉)Corresponding author: cqyygz@yeah.net

Abstract. Virtual reality (VR) is changing the way we perceive and interact with various digital information, so that many scenes can place users in an ideal visual sensory environment through head mounted devices. In recent years, the research of VR in the field of ophthalmology is mainly reflected in clinical application and teaching. This paper analyzes the relevant research in the fields of vision training and amblyopia treatment, myopia prevention and control, eye adjustment and convergence function, strabismus diagnosis, ophthalmic surgery assistance and ophthalmic teaching, and introduces the application of VR technology in the field of Ophthalmology.

Keywords: Virtual reality technology, Ophthalmology, Medical treatment, Amblyopia, Myopia.

1 Introduction

With the continuous development and progress of computer hardware and software, computer human-computer interaction interface technology based on computer image, multimedia and multi-sensor has also developed rapidly. VR and augmented reality (VR) technology are more and more applied and studied in the fields of entertainment, medical treatment, education and so on. With the innovation of technology, VR equipment has been paid more and more attention in the clinical and teaching fields of Ophthalmology.

2 Overview of VR and Ophthalmology

VR uses computer simulation to generate a virtual world in three-dimensional space, which provides users with visual and other sensory simulation, so that users seem to experience their environment and can observe things in three-dimensional space in real time and without restrictions. When the user moves the position, the computer can immediately carry out complex calculation and transmit the accurate three-dimensional world image back to make the user feel telepresence [1], as shown in Figure 1

The three-dimensional display of VR is based on the principle of binocular parallax and realized by means of head mounted display equipment. From the perspective of technology, VR system has three basic characteristics: Immersion interaction conception. The details are as follows: (1) immersion: it means that the user is in a

completely virtual environment from the first perspective, rather than watching from the third perspective like other 3D display devices, so he has a stronger sense of scene; (2) Interactivity: it means that users can interact with the built virtual environment, such as performing surgical operations on a virtual human body; (3) Imaginative: with the help of those conventional unreachable or abstract scenes, users can be in any environment, so as to expand their vision and imagination. For example, viewing the anatomical structure of the eyeball from the inside of the eyeball, following the atrial flow to feel the aqueous circulation, and displaying the complex visual path from different angles [2].

VR technology can use head dynamic instrument, eye vision sensor, hand touch sensor and so on to generate feedback of simulated operation information in virtual space in real time, to improve the user's experience of the reality of three-dimensional space.

Fig.1. VR and ophthalmology

3 Application of VR in Ophthalmic Clinical Field

3.1 Research in the Field of Myopia

Previous studies found that watching VR stereo video can simulate far and near vision activities, so as to train ciliary muscle function and relieve ciliary muscle spasm, so as to alleviate visual fatigue and the progress of myopia. Ha et al. [3] found that wearing VR equipment for 30 minutes will briefly lead to the progression of myopia, but this effect can be completely recovered after 40 minutes. In recent years, it has been reported that VR equipment may be used to control myopia. Turnbull and Phillips [4] found that the diopter and binocular visual function (such as stereopsis and adjustment amplitude) of human eyes have no significant change after wearing VR equipment, while the choroidal thickness of human eyes will become thicker. Choroidal thickening may be related to myopic defocus, so it may delay the development of myopia. From the findings of basic research, in the animal model of myopia, hyperopia defocus can accelerate the development of myopia, and myopia defocus can slow down the development of myopia. At the same time, increasing outdoor activities can slow down the occurrence of myopia. For example, children can effectively slow down the

occurrence of myopia by moving for 3 hours under the light intensity of > 10000 illuminance every day. For the reasons why outdoor sports can alleviate myopia, there are two aspects recognized internationally: one is that high-intensity light promotes dopamine secretion and then delays the development of myopia; Second, because high-intensity light can induce the pupil to shrink, and then increase the depth of field, so as to improve the visual blur, so as to delay the emergence of myopia. Therefore, in the next step, it is possible to better control the defocus of the surrounding retina through eye tracking technology, fixation point rendering technology and focal plane display technology in VR equipment, and then combined with VR equipment to control the brightness and spectral components, so as to control the progress of myopia, as shown in Figure 2.

Fig 2. Application of VR in myopia

3.2 Research in the Field Of Eye Regulation And Convergence (Divergence) Function

At present, many scholars have proposed that when wearing VR equipment, the inconsistency between accommodation and convergence may lead to functional eye diseases such as visual fatigue, dry eye, transient accommodative strabismus, video terminal syndrome and so on. Mohamed Elias et al. [5] wore VR glasses to 34 young people. They measured the binocular adjustment force and convergence and divergence function before wearing and 30 minutes after wearing. They found that the use of VR equipment will lead to the advance of eye adjustment. At the same time, the ratio of accommodation convergence / accommodation (AC / a) will be reduced, and the binocular convergence and divergence function will be weakened. They also found that if VR equipment is used to see virtual close range for too long, it will lead to mild exotropia, As shown in Figure 3. Godinez et al. [6] compared and studied the different reactions of 20 young people (aged 18 ~ 24) to wearing VR equipment and traditional computer display. It was found that VR equipment would lead to the increase of Bo direction blur point (near and far vision) and the slight increase of accommodation amplitude in the examination of convergence and dispersion range, but the difference was not statistically significant. Yoon et al. [7] also found that after wearing VR equipment for 30 minutes, although the ocular diopter will not change, the never point

of convergence (NPC) and never point of accommodation (NPA) will increase. Although the impact of VR use on visual function is not clear, the impact on human eye adjustment function and convergence and dispersion function after wearing VR equipment is still an important direction of its safety detection in the future.

Fig 3. Role of Vr In Ocular Accommodation and Dispersion

3.3 Research in the Field of Strabismus

Compared with amblyopia, VR is rarely used in strabismus research. In 2018, Thomsen et al. [8] found that after 6 months of training for 25 patients with intermittent exotropia (5 adults and 20 children), their strabismus degree decreased or disappeared, stereopsis was established, and there was no change in diopter degree. Miao et al. [9] found that VR equipment can better evaluate the degree of ocular strabismus, which is basically consistent with the diagnosis results of doctors and affirmed the accuracy and effectiveness of VR equipment by comparing the diagnosis of ocular strabismus of 17 different patients (5 orthosis and 12 exotropia) by VR equipment and doctors.

Fig 4. The role of VR in strabismus

At the same time, the research of moon et al. [10] found that VR training can improve the clinical diagnosis skills of ophthalmologists for esotropia and exotropia in a short time, and affirmed the effectiveness and convenience of VR application, as shown in Figure 4. Therefore, VR equipment is expected to be applied to the auxiliary

diagnosis of strabismus in the future.

4 The Role of Visual Rehabilitation Training in Children with Visual Impairment

Visual impairment includes blindness and amblyopia. In the early stage of children's visual development, active and correct amblyopia treatment will produce good results. However, children in this period have poor cognitive ability, so it is difficult to cooperate with and adhere to the traditional therapy with monotonous and long training cycle. Therefore, in recent years, ophthalmologists have tried to find a new VR treatment method that can not only stimulate children's interest in training, but also improve the treatment effect.

4.1 Amblyopia Treatment Based on VR

VR technicians from the University of Nottingham and ophthalmologists from Queen's Medical Center have developed an interactive "binocular processing system" to provide interactive 3D games and videos for children with amblyopia. Research shows that the system can provide a relaxed and pleasant treatment method, which can enable children to obtain ideal curative effect in a short time. Chinese ophthalmologists have also made similar explorations and developed the "vision enhancement" system software, which integrates amblyopia treatment with virtual scenes, as shown in Figure 5.

Fig 5. Application of VR in amblyopia treatment

The system adopts a variety of stimulation modes. On the one hand, it improves the visual acuity of amblyopia and makes up for the shortcomings of traditional therapy; On the other hand, help children establish normal binocular visual function and promote their visual function and healthy development of body and mind. According to the clinical report of Mian Yao, the system software of "increasing visual energy" has the advantages of strong pertinence, easy operation, diversification, and children's

willingness to accept. Its training effect is better than that of traditional therapy, especially for ametropic amblyopia and mild amblyopia. When using the system software to treat amblyopia children of different ages, it is pointed out that children need to have certain hand eye coordination ability because they need to control the mouse by hand in the training process, Therefore, too young children are not suitable for using the system. In another study, they emphasized the early detection and treatment of amblyopia, and suggested that qualified families use the software for training as soon as possible.

4.2 Research on the Types of Visual Function Defects Based on VR

The traditional types of visual impairment are divided into three types: ametropic amblyopia, anisometropic amblyopia and strabismus amblyopia based on the examination of visual acuity chart, and are divided into three grades: mild, moderate, and severe. Using the "children's vision and intelligent VR database system based on perceptual learning", 323 children with amblyopia were examined for visual function defects. According to the types of visual information processing defects, amblyopia was divided into "low-level visual function defect", "high-noise visual function defect" and "high-level visual function defect", and a good distinction effect was obtained. When diagnosing amblyopia children, we should increase the evaluation of their visual status on the basis of measuring their visual acuity level with the traditional visual acuity chart, and take this as the basis for targeted treatment to repair their visual dysfunction, as shown in Figure 6.

Fig 6. VR is used for visual function adjustment

5 Research on VR in Ophthalmology Teaching

Using VR technology to build a simulation system of normal human eye anatomical structure, the created image has both three-dimensional and realistic feeling. At the same time, it can also rotate, zoom in, zoom in, zoom out, etc., which can more intuitively observe the internal structure of the eyeball. At the same time, the research of Jin et al. [11] also pointed out that using VR technology can build various three-dimensional scenes, and then simulate the symptoms and signs of various ophthalmic diseases, such as visual blur, visual object deformation, visual field defect, etc., which is helpful to assist the teaching of students' ophthalmology courses. As long-term use of VR equipment may lead to visual fatigue, improving VR technology

and equipment to reduce students' visual fatigue after use is an important direction of development in the future.

In the future, VR technology has a wide application prospect in the teaching of simulated ophthalmic diseases. It can also be used to establish a standardized patient database for ophthalmic teaching and assessment. The eye Si (vrmagic, Germany) surgical simulator is most used in ophthalmic surgery teaching. This simulator can simulate three-dimensional images in surgery under the microscope, simulate and train cataract and vitreous surgery. It has the advantages of simple and controllable use, high degree of simulation and reverse operation. It can significantly improve the technical level of ophthalmic inpatients in cataract surgery, especially capsulorhexis and anti-shaking, and has a significant correlation with the actual operation. The surgical simulation system can also support the training of vitreous surgery. Through this system, users can carry out basic intraocular micromanipulation training, such as vitrectomy, intraocular laser, posterior vitrectomy, stripping of internal limiting membrane and so on. Through training, surgical skills can be improved to varying degrees, but whether it can be successfully converted to real patient surgery remains to be further studied.

6 Conclusions

With the rapid development of modern medicine science and technology, more and more medical technology achievements benefit mankind. However, in the field of ophthalmic medicine education, students have some problems in the process of learning ophthalmic medicine, such as boring theoretical knowledge, shortage of experimental sites, unsatisfactory practical operation and so on. From the perspective of Ophthalmology, this paper studies the application of VR in the fields of vision training and amblyopia treatment, myopia prevention and control, eye regulation and convergence function, strabismus diagnosis, ophthalmic surgery assistance and ophthalmic teaching. At the same time, it expounds the advantages of VR technology in physics and cognition and puts forward suggestions on the application of VR technology in this field, so as to provide more references for the effective implementation of ophthalmic medical education.

Acknowledgments

School-level project: “Integration of socialist economic construction, political construction, cultural construction and social construction, integration of learning and innovation” The Construction of Optometric Talent Training Path of Exploration and Practice YGZCG1905; school-level key discipline project: optometric medicine ygz2021302.

References

1. Mitrousia V., Giotakos O. Virtual reality therapy in anxiety disorders. *Psychiatriki*, 2016, 27(4):276-286.

2. Pratorius M., Burgbacher U., Valkov D., et al. Sensing thumb-to-finger taps for symbolic input in VR environments. *IEEE Comput Graph Appl*, 2019，20-24.
3. Ha S.G., Na K.H., Kweon I.J., et al. Effects of head-mounted display on the oculomotor system and refractive error in normal adolescents. *J Pediatr Ophthalmol Strabismus*, 2016, 53(4):238-245.
4. Turnbull P., Phillips J.R. Ocul V.R. effects of virtual reality headset weVR in young adults. *Sci Rep*, 2017, 7(1):16172.
5. Mohamed Elias Z., Batumalai U.M., Azmi A. Virtual reality games on accommodation and convergence. *Appl Ergon*, 2019, 81:102879.
6. Godinez A., Harb E.N., Grimes J., et al. Oculomotor changes after sustained Virtual Reality use. *Investigative ophthalmology & visual science*, 2019, 60(9):5924.
7. Yoon H.J., Kim J., Park S.W., et al. Influence of virtual reality on visual parameters: immersive versus non-immersive mode. *BMC Ophthalmol*, 2020, 20(1):200.
8. Thomsen AS, Smith P., Subhi Y., et al. High correlation between performance on a virtual-reality simulator and real-life cataract surgery. *Acta Ophthalmol*, 2017, 95(3): 307-311.
9. MiaoY., Jeon J.Y., Park G., et al. Virtual reality-based measurement of ocular deviation in strabismus. *Comput Methods Programs Biomed*, 2020, 185:105132.
10. Moon H.S., Yoon H.J., Park S.W., et al. Usefulness of virtual reality-based training to diagnose strabismus. *Sci Rep*, 2021, 11(1):5891.
11. Jin B, Ai Z, Rasmussen M. Simulation of eye disease in virtual reality. *Conf Proc IEEE Eng Med Biol Soc*, 2005, 2005:5128-5131.

Research on the Application of BIM Technology in the Whole Process Cost Management of Construction Project

Kangyan Zeng, Zhen Wen(✉)

School of Civil Engineering, Chongqing College Architecture and Technology, Chongqing 401331, China,
(✉)Corresponding author: Wendy198787@163.com

Abstract. Cost management, with the information technology used frequently, is the most important link in the process of engineering construction. It is, however, no longer able to adapt to the trend of information technology development by using traditional work methods. The whole process cost focuses on the whole construction process and the overall interests of the project. BIM technology, as an electronic information modeling, provides an efficient information exchange platform, on which the cost management work can be connected at any stages in series with the result of repetitive work reduced and work efficiency increased. The improvement of work efficiency and way can be realized by applying the suitable BIM software to each stage of cost management can improve work efficiency. The paper studies specifically the integration of the whole process cost management on construction engineering. There is a commercial residential project in the case part in the use of relevant BIM software to realize the systematic collaborative management for the whole process of engineering cost. On the basis of the transfer of the cost data to the whole process of construction by using the BIM software, it concludes the route and method of the whole process cost management in the use of relevant software.

Keywords: BIM Technology, Whole Process, Whole Process Cost Management

1 Introduction

With the increasingly development of the scale and the output value in recent years, the construction market has reached a high proportion of GDP in China, and even up to 12 percent in 2021[1]. Nowadays, the development of new technologies is rapidly in all walks of life, so the construction technology has been updated and improved to a certain extent. The efficiency on the construction technology, nevertheless, is still at low level, because of the properties of the products and the trend of industry development. There are problems in the work, including every link in the construction process fails to transmit information efficiently, and participants of the construction work fail to cooperate with each other, and the repetitive work is emerged at each stage. It will not only lead to low efficiency in the construction process, but also cause serious waste of resources[2].

BIM technology is an effective way to improve the informatization of the

construction industry. It turns the construction process into a modern industrialized production model of an assembly line, in which BIM forms a digital production line with a powerful data in all related software and models through a shared work platform, upgrading the production process of construction products to a modern industrial-grade production model. At this stage, the cost management methods of construction engineering are neither able to adapt to the current trend of informatization development, nor to control construction investment effectively. The paper effectively integrates BIM technology into the cost management of all aspects of engineering construction based on a comprehensive and systematic analysis of the whole process engineering cost management and BIM technology [3].

2 Hole Process Cost Management of Construction Project

The whole process cost management of construction engineering includes total cost management, whole lifecycle management, total factor cost management, and whole process cost management. The whole process cost management place emphasis on the work of the cost management which should be involved in advance. It runs through the whole range of construction engineering with the starting point of the construction engineering planning and decision-making stages to the end point of the completion[4].

2.1 Cost management in Each Stage

First of all, the investor of the construction project needs to take such factors as the amount of capital to be invested, opportunity cost, and actual technical level into consideration so as to make a decision on the scale and usage of the project at the investment decision-making stage. Project cost management staff should grasp the investment of the project as a whole and prepare investment estimation [5].

The design stage is the key point for the actual formation of the project cost and the most effective control. At this stage, the design unit often calculates the project cost according to its proposed construction scheme or construction drawings.

The project price calculated in the bidding stage is not only the basis for both parties to determine the contract price, but also the basis for settlement between both parties in the later stage [6].

The project construction stage is the central link in the whole project life cycle of the construction unit and the contractor. The contractor completes an actual building through the construction process, and most of the project investment will be spent at this stage. This stage is the formation process of the actual project cost[7].

The completion settlement of the project is the process in which both parties calculate all the completed construction products and pay the project price according to the specific provisions of the construction contract after the completion acceptance is qualified. The completion settlement received by the construction party is generally composed of the contract price plus or minus the adjustment amount recognized by Party A and Party B as well as the deduction of project progress payment and quality warranty deposit paid during the construction process [8].

2.2 Difficulties in the Implementation of the Whole Process Cost

(1) It is difficult to gather professionals to support the whole process cost management at the full stages of the construction process with the result of the small scale in the cost industry.

(2) The key point of cost management focuses on the project pricing business. The quantity surveyors always take the measurement and valuation, rather than focusing on the value management of the entire life cycle of the construction project[9].

(3) There is the situation with slow update speed for the measurement and valuation basis is not fast enough in the whole process of cost management. There is insufficient amount of similar engineering cost data in the estimation and budgetary estimation stage.

(4) At present, the work mode of most quantity surveyors is to take the measurement and valuation in the use of the project cost software and quota data set by government, which is apt to fail to be in accordance with the actual situation[10].

Firstly, on the basis of the combination of the above-mentioned difficulties in project cost management, the quantity surveyors need to use electronic information tools to improve the accuracy and efficiency of engineering measurement; secondly, the information platform should be built for idea exchanges between different professional staff at each stage to record and analyze the whole process cost data; finally, a record carrier of engineering cost data which is practical and reliable is needed to analyze and store the engineering cost data in a structured manner[11].

3 Bim technology is Applied in the Whole Process Cost to do Bim Fusion Analysis

3.1 Investment Decision-making Stage

The application of BIM includes initial modeling, model maintenance, cost estimation, etc. in the project planning stage. According to the existing data, the current 2D drawings are imported into software with BIM technology to build a 3D modeling. Generally speaking, it is the initial project modeling created the early stage. The investment estimation is taken in the use of the BIM technology with a powerful information statistics function based on this modeling. At this stage, relatively accurate engineering quantities can be obtained according to the model, and the further calculation can be taken on the installation costs of the building. At the same time, the project cost data can be used to weigh the pros and cons of different schemes, compare and optimize the schedule, so as to prepare and provide an important basis for project decision making.

3.2 Design Stage

In the past, the drawings were made by different designers with different majors such as civil engineering, water and electricity, and fire-fighting pipelines. Conflicts and collisions as well as size deviation is easy to occurs between different majors and different views of the same major. The designers, auditors and other parties are unable to completely find and correct the unreasonable points, even if they spend a lot of energy to check and compare the drawings. These conflicts are manifested in the construction process, which has caused great uncertainties to the cost management, even quality and safety of the project, and result in an increase in costs. When it comes

to the establishment of 3D models, the collaborative design of various professional designers, and the visual analysis of different professional components adopted in the process of the design, the conflicts caused by the drawings will be resolved in time, and the interactive check can reduce errors in the design.

The various dimensional information provided by the BIM model will also simplify the calculation of the engineering quantity in the design stage. It can be directly calculated for the engineering quantity in the use of the BIM model. The data of each component in the model is related to the calculation process of the engineering quantity. When the components in the model are changed, the engineering quantity will also be updated, so that the engineering cost data can be updated in real time. In the design stage, the cost personnel can use the BIM technology to greatly shorten the time for calculating the project quantity, realize the rapid and accurate preparation of the project estimate, and can also discover some conflicting problems that were only discovered during the construction in advance, and reduce the later engineering changes.

3.3 Tendering and Biding Stage

For the tenderer, BIM can truly provide the engineering entity information required in the calculation of the engineering quantity to automate the calculation, improve the accuracy of the calculation, and allow the cost staff to change from repetitive calculation work to thinking and controlling the factors that affect the price of the project, a more scientific budget can be prepared. If bidders want to have their own bidding data, they need to introduce BIM to quickly calculate and fully store the consumption standards during the construction process. Through reuse or rapid establishment of 3D models, fast and accurate calculation of engineering quantities will no longer be a problem. In addition, the bidder can use the 3D design model to quickly locate the structural information of heavy and difficult areas, determine and adjust the construction plan according to the actual situation of the project, correctly evaluate the difficulty of the project, and make accurate quotations.

3.4 Construction Stage

The key point of construction units on costs management is the management of the project construction costs. The cost targets are mostly compared with the unit price and amount of the signed contract. Generally speaking, the post-event analysis is been taken. There is a lack of cost control in the process. The construction schedule only contains the size of the project and the completion time information, instead of changing the project plan and actual completion; most of the construction schedule of the project department is determined by the sophisticated construction management personnel. There will be deviations between the engineering quantity and the amount of labor, materials, and machinery resources calculated by the project manager and the actual value, with result of the increasing on the engineering cost in the actual construction process. It is usually more accurate for the construction schedule preparation completed by BIM technology. The time information is added to form a 4D model in the use of the 3D model of BIM. The resource consumption required by each construction process and construction node can be accurately calculated, and then cost information is added to form a 4D model. The 5D model of the project, using quota consumption data, etc., accurately calculates the number of labor and construction machinery required for each construction process. In order to prevent insufficient

resource input, it is feasible to add progress information, cost information, and construction organization information into the 3D model to calculate the consumption of people, materials, and machines for the entire project.

3.5 Project Acceptance Stage

The settlement data of the project can be obtained by collecting and arranging the project information and data during the design and construction of the construction project. Using BIM technology to collect the information of the building in a complete and structured manner, the completion and settlement of the project can be quickly counted. It can quickly compare and calculate with the contract price, and finally form an accurate settlement price.

4 CASE

We take a commercial real estate project as an example. According to the whole process cost management process, we adopt the bill of quantities valuation method to calculate the engineering cost of the civil works. In the process, BIMMAKE, a kind of Glodon modeling software, is used to establish BIM model, Glodon GTJ is used to calculate engineering quantity, Glodon cloud pricing platform (GCCP) is used to calculate list unit price and project cost, and Glodon BIM5D is used to take construction simulation. At first, Glodon BIM software is used to build a 3D model in this project. The established model is imported into the Glodon GTJ2018 to calculate the construction project volume. Next, the Glodon pricing platform is used to apply the list quota, so as to, on the one hand, avoid the data loss caused by the REVIT model in the process import procedure or the errors caused by manual copying of CAD drawings, on the other hand, it is, in a large extend, to reduce the workload of the cost engineers for modeling.

Here are the specific work of in the use of BIM and related softwares to carry out the whole process cost management of the case project: at the planning stage, the total investment of the project should be estimated; at the design stage, it is mainly for the preparation of budget estimates; at the tendering and bidding stage, the bidding control price is prepared; in the mid-construction settlement, the engineering quantity calculation and engineering change control are carried out; in the completion settlement, the engineering quantity is calculated and the claim management is carried out; and finally the cost data, the extraction and preservation of the cost target are completed.

In the investment decision-making stage, Glodon software is used for 3D modeling, on which the engineering volume is quickly calculated, and then the Glodon Index Network is used to query and check investment estimation indicators, which is quickly and accurately achieved on the investment estimation documents. This will greatly improve the accuracy of estimation and provide an accurate data source for subsequent cost management work.

At the design stage, the exact engineering quantities should be calculated on the basis of the designed construction drawings, and the cost of the project should be calculated in the use of the current bill of quantities valuation specifications and local quotas. At this stage, BIM software can be used to perform some direct conflict checks of various disciplines. For example, when we check the collision of the

drainage pipes, it can be judged whether the engineering pipes collide with the frame beams; when it comes to every view, we would wonder and check that the discrepancies between the structural drawings and the architectural drawings. At the design stage, the mistakes in the drawings should be corrected as much as possible, so as to avoid the occurrence of rework and changes during the later construction, thereby avoiding the increase of the engineering cost.

In the bidding stage, the tenderer needs to use the BIM model to quickly calculate the quantities when preparing the cost documents. The software has built-in list specifications to form a complete bill of quantities; the calculation rules of list and quota have been set. There is no need for cost personnel to remember the calculation rules. The software will automatically deduct according to the drawing of component elements and use them at the same time. The quantities of two calculation rules can be obtained from the same model; the software provides multiple engineering quantity codes, which can be combined and extracted freely; we can use the Glodon cloud pricing platform to calculate the unit price of the bill of quantities, take the fee, summarize and calculate the bidding control price.

The bidder adopts the three-dimensional model provided by the tenderer to calculate the quantities faster and uses the pricing software to prepare the bidding price. The BIM model established by the bidder at this stage can be imported into Glodon BIM-5D software to prepare the schedule, and carry out engineering change cost, monthly settlement and quarterly settlement at the construction stage.

The cost management to be carried out in the construction stage includes change management, process payment management and progress management. The premise of using Glodon's change software to record the design change is to have a model file approved by both Party A and Party B to draw the changed components on the basis. It is convenient to use the change software to see the increase or decrease of the changed parts. At present, the quantities of the general list are settled according to the facts, or the changed parts can be drawn directly. The actual quantities can be counted in the progress settlement or completion settlement. Process payment management and progress management can be realized through Glodon BIM-5D. During settlement, Party A and Party B shall calculate the actual quantities on time according to the comprehensive unit price or price adjustment method signed in the unit price contract. The calculation of quantities can use the three-dimensional model that has been used in the process to find differences. After modifying the components, you can update yourself and related quantities. With regard to the change of the comprehensive unit price in the list of quantities, the price of materials can be calculated by using Glodon Assistant to calculate the weighted average of the monthly information price or market price. The comprehensive unit price of the list can be calculated automatically by using the pricing software and then the settlement documents can be prepared according to the contract.

In this process, the same three-dimensional model has been used for data flow, which can reduce the modeling time of cost personnel in each stage and avoid data loss and error caused by repeated modeling.

5 Conclusions

This paper introduces BIM Technology to realize the whole process cost management of construction engineering, and mainly obtains the following research results:

(1) Through qualitative analysis, it proves that BIM Technology has the characteristics of simulation and visualization, which can greatly improve the speed and accuracy of cost personnel in calculating quantities, shorten the time of calculating project cost and provide an effective and advanced working method for cost management.

(2) Through quantitative analysis, the initial BIM model is established by using BIM make software, which transforms the traditional two-dimensional drawing into the three-dimensional physical drawing of what you see is what you get, strengthening the intuitiveness of the drawing and easy to understand and find design errors; the BIM calculation model is established by using Glodon GTJ2021, which realizes the rapid and accurate calculation of quantities and can correlate the design change with the calculation results of quantities in real time. After the change, the quantities can be calculated and counted quickly.

(3) This paper analyzes the BIM software used in each stage of project construction and the use process, methods and important functions of the software, which provides practical experience for similar projects to use BIM Technology for cost management in the later stage. Using BIM Technology can greatly improve the efficiency and accuracy of cost management.

From the perspective of cost management in the whole process of construction engineering, the application of BIM should focus on the overall construction process rather than just considering a certain stage. BIM model should be continuously transferred to the whole process of cost management in order to achieve the best use effect.

References

1. Reza Mohajeri BorjeGhaleh, Javad Majrouhi Sardroud. Approaching Industrialization of Buildings and Integrated Construction Using Building Information Modeling. *Procedia Engineering*, 164. 2016
2. Chen L J, Luo H. A BIM-based construction quality management model and its applications. *Automation in Construction*, 46:64-73. 2014
3. Raimundo F. Dos Santos Jr., Chang-Tien Lu. Geography Markup Language (GML). *Encyclopedia of GIS*, 2016.
4. Jiang Yi. Research on Construction Project Cost Management. *Fujian Building Materials*. 01, 111–113. 2021
5. Li Guo jun. Present Situation and Countermeasures of the Construction Project Whole Process. *Chinese architectural decoration*. 01, 160–161. 2021
6. Zhang Qi. Application Research of BIM Technology in Project Cost Management. *Neijiang technology*. 42(05), 59-60. 2021
7. Cao Yingjie. Practice of BIM Technology in Construction Project Management. *Residential and Real Estate*. 24, 123. 2020
8. Liu Peiyuan. Present Situation and Development Suggestion of Whole Process Cost Management. *Residential and Real Estate*. 24, 41-42.2020
9. Carmine Cavalliere, Guido Raffaele Dell' Osso, Fausto Favia, Marco Lovicario.

BIM-based Assessment Metrics for the Functional Flexibility of Building Designs. *Automation in Construction*, 107. 2019
10. Xiaojuan Li, Chen Wang, Ali Alashwal, Shilpi Bora. Game Analysis on Prefabricated Building Evolution Based on Dynamic Revenue Risks in China. *Journal of Cleaner Production*, 267. 2020
11. Petar Kochovski, Vlado Stankovski. Supporting Smart Construction with Dependable edge Computing Infrastructures and Applications J. *Automation in Construction*, 85:182-192. 2018

Modal Parameter Identification of Bridge Structure based on Hybrid Genetic Algorithm

Rong Hu(✉)

School of Road Bridge and Architecture, Chongqing Vocational College of Transportation, Chongqing 402247, China
(✉)Corresponding author: mingtiandaban@163.com

Abstract. Bridges play an irreplaceable role in the structure of modern transportation system and play an important pivotal role in the development of politics, economy and culture. In the development of bridge engineering, the modal parameter identification(MPI) of bridge structure is particularly important. Therefore, this paper studies and analyzes the MPI of bridge structure based on hybrid genetic algorithm(HGA). Firstly, the identification method of bridge structural modal parameters and the identification content of bridge construction parameters are briefly analyzed, and the HGA is proposed. It is analyzed that the HGA mainly plays the role of fitting and Optimization in the identification of bridge structural modal parameters; Finally, based on the monitoring project of a Provincial Railway temporary bridge, combined with the finite element theory analysis of ANSYS, the genetic algorithm is applied to MPI by combining signal filtering and random decrement method. The test results show that the minimum frequency error is 1.93%, the maximum error is 9.33%, and the first three frequency errors are within 6%. When the genetic algorithm is applied to MPI, the modal order determination problem has a great impact on the results of parameter identification, The feasibility and effectiveness of HGA applied to bridge structure MPI are verified.

Keywords: Hybrid Genetic Algorithm, Bridge Structure, Modal Parameters, Parameter Identification

1 Introduction

In the process of bridge construction, in addition to considering the influence of non-uniformity of materials on structural stress, climate humidity, temperature and other uncertain factors also need to be considered. In addition, the construction method adopted is generally multi process and multi-stage construction. With the progress of construction, these factors often make the displacement and internal force of each construction stage gradually deviate from the theoretical value. The parameters adopted in the design, such as the rigidity of the cradle, the dead weight of the structure, the modulus of elasticity of materials, the shrinkage and creep coefficient of concrete and the temporary construction load, will be different from those in the actual project, which will make the state of the actual structure in each construction stage different from the theoretical calculation. Therefore, the main parameters of the bridge structure should be calculated according to the measured data during the construction process,

and then the modified parameters should be fed back to the actual construction control calculation. In order to ensure the quality and safety of bridge construction, the identification of bridge parameters in the construction stage is indispensable.

MPI of bridge structure based on HGA has been studied and analyzed by many scholars at home and abroad. Matsubaram proposed a method for identifying the parameters of passenger car tires based on the three-dimensional flexible ring model. This method can identify the modal parameters through experimental modal analysis, and compare the model parameters with the modal parameters by using the model calculation. The recalculated results using the model parameters show a good correlation with the experimental results [1]. Schfletr proposed a new off-line optimization method to solve the coverage path planning problem. For grid based environment representation, a new HGA is proposed, which uses turning start point and backtracking spiral algorithm for local search. The calculation results show that, compared with the traditional method, the path improvement rate of HGA is as high as 38.4%, and it has the same adaptability to different starting positions in the environment [2].

In this paper, a HGA is proposed by combining genetic algorithm with random decrement method and signal filtering. The MPI method based on HGA is discussed and analyzed. The performance of HGA (HGA) in avoiding the trap of local optimization and finding the global optimal solution is studied; How to identify the modal parameters of linear time invariant structural system by using the optimization function of genetic algorithm is discussed. The application of MPI based on HGA in engineering practice is discussed. Through the analysis of various analog signals and measured bridge signals, it is proved that the bridge structure MPI method based on hybrid algorithm proposed in this paper can process the bridge test signals in various environments and identify the bridge structure modal parameter information [3-4].

2 MPI of Bridge Structure

2.1 Bridge Structure MPI Method

Frequency domain method: most frequency domain identification methods are based on fast Fourier transform. They have the advantages of mature theory, simple operation, fast identification speed and high identification accuracy. However, due to its limited frequency resolution, it is not enough to decouple the dense modes. The principle that the structural frequency response function has a maximum value at the natural frequency is used to identify the natural frequency. This method does not need to set parameters, and has the advantages of convenient operation and fast recognition speed. However, the dense modes cannot be decoupled, and the mode shapes cannot be obtained directly. Instead, the working deflection line shape is used to approximately replace the mode shapes [5].

Frequency domain decomposition method: the frequency domain decomposition method is an extension of the peak picking method. This method has certain anti noise ability and high identification accuracy, but its decoupling ability of low-frequency dense modes is still not high.

Polynomial fitting method: polynomial fitting method generally carries out high-order polynomial fitting for each frequency response function, and then uses

some form of averaging to obtain the overall modal parameters of the structure. But this method is easy to lead to ill conditioned matrix, and can not get high-precision fitting. In order to solve this problem, an orthogonal polynomial fitting method is proposed to improve the accuracy of modal identification.

Time domain method: the time domain method directly uses the system response time history signal to identify the modal parameters, and does not need to use Fourier transform to transform the signal into the frequency domain for analysis. Therefore, there is no problem of frequency resolution. However, the time domain method is difficult to determine the system order, sensitive to noise and prone to false modes. The time domain method uses the response data obtained by the random decrement method or the natural excitation technique to establish the mathematical model of the characteristic matrix equation, and uses the relationship between the system modal frequency, modal damping and the eigenvalue of the characteristic matrix to solve the modal parameters.

The random decrement method eliminates the structural response caused by random load through the sample averaging method, so as to convert the random response signal into a free attenuation signal [6]. The natural excitation technology uses the cross-correlation function of the response signals of two arbitrary measuring points on the structure to have a similar mathematical expression with the impulse response function under impulse excitation, so the cross-correlation function is used to replace the impulse response function.

Random subspace method: the random subspace identification method is a completely data-driven parameter identification algorithm. It does not need to obtain the free attenuation signal or impulse response function of the structure through random decrement method and natural excitation technology, and has certain anti-interference ability to noise. Stochastic subspace algorithm has been widely used in engineering because of its clear concept, perfect theory and easy programming. Generally, the stability diagram method or singular value entropy method is used to judge the system order, and the stability diagram method can also help eliminate false modes and improve the identification accuracy [7-8].

Time frequency domain method: since both frequency domain method and time domain method assume that the test process is a stationary random process, it cannot meet the requirements of non-stationary signal and time-varying system parameter identification. Modern time-frequency analysis methods provide a means to analyze non-stationary signals. In addition, these methods also have excellent low-frequency dense mode decoupling ability, which is of great significance to practical projects, especially long-span bridge structures.

Wavelet transform method: the basic process of the wavelet transform identification method of modal parameters based on environmental excitation is as follows: firstly, the random decrement method or natural excitation technology is used to preprocess the structural response under environmental excitation to obtain the free attenuation signal or impulse response function; Then the wavelet base is constructed by using the appropriate mother wavelet, and the processed signal is transformed by wavelet to obtain the time-frequency distribution of wavelet coefficients, on which the wavelet ridge is extracted; Finally, the modal parameters of the structure are extracted from the wavelet coefficients of the wavelet ridge. This method has excellent anti noise ability, low-frequency dense mode decoupling ability and analysis

ability for non-stationary signals [9]. However, this method still needs to be further improved, such as the extraction of wavelet ridge, the elimination of endpoint effect and the design of optimal wavelet basis function.

2.2 Parameter Identification During Bridge Construction

Parameter identification content: parameter identification is to first determine the bridge structural parameters that have a great impact on the bridge response, then based on the error between the measured response data and the theoretical calculation data, and finally feed back the actual structural parameters to the construction control calculation, so as to timely adjust the theoretical values required for the bridge construction in the next stage. For the identification of bridge structural parameters, the main structural parameters causing the structural state deviation must be determined by some analysis method, and then the appropriate parameter identification theory or method should be used to identify the structural parameters [10]. For the general bridge structure, the main structural parameters refer to the factors that can significantly cause the change of the bridge structure state.

MPI of vertical Bridge

Signal processing: through the preliminary analysis of the test data, it can be seen that the vertical sensors arranged at the north side span of the bridge have failed, the test results are not ideal, and the test data are unavailable. The data collected by the vertical sensors at other positions are ideal, which can be used for the analysis and calculation of MPI. Due to the large span of the main span, the vibration amplitude under environmental excitation is large, and the sensor has obvious perception of vibration, so the reliability of the collected vibration data is high. For the long-span bridge constructed by phased cantilever, the numerical analysis and construction control during construction play an irreplaceable role in the smooth construction, and the parameter identification is the difficulty and focus of numerical analysis and construction control [11]. The parameter identification in the bridge construction stage is to first analyze the main parameters that have a great impact on the bridge structure state through the parameter sensitivity analysis, and then use certain methods to estimate the error between the actual parameters and the theoretical parameters according to the error between the measured data and the theoretical data during the bridge construction, so as to identify the bridge parameters in the actual construction state, and use the identified parameters to guide the subsequent construction stage, Finally, the bridge completion state of the structure is consistent with the ideal bridge completion state [12].

3 Frequency Domain MPI based on HGA

As a new optimization method, HGA (HGA) is attractive for its excellent computational performance and remarkable application effect. The combination of genetic algorithm and computer technology has created a new research field, and constantly infiltrated into other fields to give full play to its excellent performance.

HGA (HGA) mainly plays the role of fitting and Optimization in MPI of bridge structures. The vibration of the multi degree of freedom system is assumed to be the superposition of multiple impulse responses. Through the random decrement technology and signal filtering technology, the free attenuation signal is fitted with the

determined impulse response function. When the signal contains fewer frequency components, the easier the fitting optimization is and the more accurate the result is. Therefore, before using the HGA, filtering the high-frequency noise components in the signal can effectively improve the accuracy of the recognition results. At the same time, when using HGA, the setting of parameters plays an important role in the accuracy of the results.

Because the solution obtained by HGA always makes the objective function tend to the minimum value when optimizing problems, it is necessary to transform the objective function when using this algorithm. The frequency response function model of the structural system is:

$$K_{lp}(\gamma; e, \delta, \gamma_r) = \sum_{r=1}^{N} \frac{1}{e_r(1-(\gamma/\gamma_r)^2 + 2\delta\gamma/\gamma_r)} \tag{1}$$

er, δ r, γ R (r=1,2,... N) is the modal parameter to be identified. $\hat{K}_{lp}$ is the measured frequency response function, and the theoretical frequency response function is KL, P. the identification problem is transformed into minimizing the difference between $\hat{K}_{lp}$ and KL, P. Namely:

$$\min X = \min \sum_{i=1}^{T} [\hat{K}_{lp}(\gamma_i) - \sum_{r=1}^{N} \frac{1}{e_r(1-(\gamma_i/\gamma_r)^2 + 2\delta_r\gamma_i/\gamma_r)}]^2 \tag{2}$$

The fitness function is:

$$\begin{aligned} j &= j_{\max} - X \\ &= j_{\max} - \sum_{i=1}^{T} [\hat{K}_{lp}(\gamma_i) - \sum_{r=1}^{N} \frac{1}{e_r(1-(\gamma_i/\gamma_r)^2 + 2\delta_r\gamma_i/\gamma_r)}]^2 \end{aligned} \tag{3}$$

JMax is a known quantity set before identification to ensure j>0.

4 MPI of Bridge Structure based on HGA

Based on the monitoring project of a Provincial Railway temporary bridge, this paper studies the application of MPI based on HGA in engineering. Through the detailed project overview, the risks existing in the project construction are understood, which reflects the necessity of monitoring. Through the improvement of time domain MPI method based on modal decomposition, the identification method suitable for engineering practice is obtained, so as to improve the accuracy of identification results.

In order to monitor and evaluate the health status of the bridge during its operation, a health monitoring system was designed during the construction of the bridge. A variety of sensors were installed at the main positions of the bridge girder, tower, stay cable, etc. if the annual data were analyzed and processed, the workload would be huge. Therefore, the data volume of a day with ideal test data quality was selected as

the analysis object, with a total of 24 time history files, The sampling time of each time history data is 3600s and the sampling frequency is 20Hz.

4.1 Finite Element Theoretical Analysis based on ANSYS

ANSYS large-scale general finite element software is widely used in structural engineering, bridge engineering, geotechnical engineering, water conservancy engineering and other fields because of its powerful function and versatility. In order to master the dynamic characteristics of the temporary railway bridge, the finite element theoretical analysis of the temporary railway bridge is carried out by using the finite element software ANSYS, so as to compare with the identification value of the measured signal.

The bridge deck is made of in-situ reinforced concrete continuous slab, and the pier body is made of angle steel lattice column pier. The first five natural frequencies and vibration modes of the temporary bridge before horizontal and vertical bending are calculated by subspace iteration method. See Table 1.

Table 1. Natural frequency and mode shape of temporary bridge in transverse and vertical bending

Transverse bending		Vertical bend
stage	Frequency /HZ	Frequency /HZ
1	1.752	7.455
2	1.877	7.462
3	2.271	8.116
4	3.381	8.121
5	7.134	8.190

4.2 MPI based on Genetic Algorithm

Signal preprocessing: the measured signal is generated by environmental excitation. The vibration signal under environmental excitation is used to identify the modal parameters. Data preprocessing is required to make the signal conform to the form required by the time domain identification method. Usually, the random decrement method is used to extract the free vibration signal, or the next method is used to take the cross-correlation function as the time domain identification input data. Firstly, the signal is filtered and denoised, and then the free vibration response signal is extracted from the original signal by random decrement method, and the modal parameters of the preprocessed signal are identified.

After preprocessing the measured signal, the free attenuation vibration response of each channel is obtained, and the input data required by the real-time domain identification method is obtained. A genetic algorithm parameter identification program based on MATLAB is developed to identify the modal parameters of horizontal and vertical measured signals respectively. The parameters of HGA are set as follows: population size 600, initial range [1:9], crossover probability 0.95, mutation probability 0.015, iterative evolution times 100, stop criterion using maximum evolution times, and coding method using binary coding. See Table 2 and figure 1 and Figure 2 for the identification results of damping ratio and ANSYS theoretical values.

Table 2. Ratio of transverse bending MPI result to ANSYS theoretical value

stage	Natural frequency (Hz)			Damping ratio (%)
	GA Identification value	ANSYS Theoretical value	relative error(%)	GA Identification value
1	1.720	1.754	1.92	14.9078
2	1.932	1.877	2.91	13.8950
3	2.155	2.269	5.08	11.8640
4	3.223	3.381	4.59	5.7279
5	6.466	7.133	9.32	2.8179

Based on time domain method and frequency domain method, this paper applies genetic algorithm to MPI of single degree of freedom and multi degree of freedom simulation signals. The results show that the maximum error of frequency and damping ratio is 1.84% and 3.1% respectively without noise. When the noise is 20%, the maximum frequency error is 24.04% of the maximum damping ratio error. It can be seen that the frequency identification accuracy is high and the damping ratio identification accuracy is relatively low. When using genetic algorithm to identify modal parameters, it has strong anti noise ability, which reflects its strong robustness. However, the algorithm has many parameters and needs more debugging in use to minimize the error.

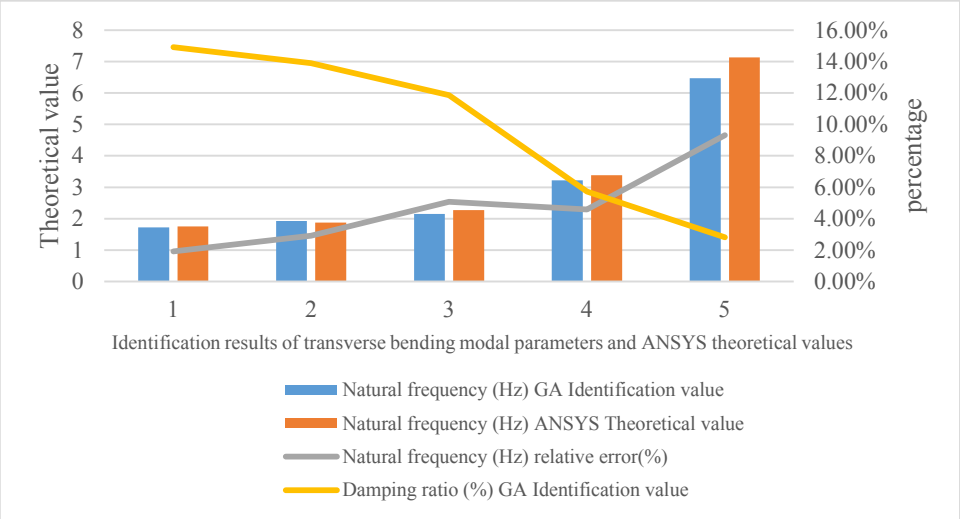


Figure 1. Identification results of transverse bending modal parameters and ANSYS theoretical values

Comparing Fig. 1 and Fig. 2, it can be found that in the first five natural frequencies, the natural frequency of vertical bending is larger than that of horizontal bending, and the damping ratio of vertical bending is smaller than that of horizontal bending. This is mainly because the lattice columns are dense, and the transverse stiffness is smaller than the vertical stiffness, making the temporary bridge more prone to transverse vibration. Comparing the identification value of HGA with the theoretical value of ANSYS, it can be found that the relative error increases with the

increase of order. Using the optimization function of genetic algorithm to identify the natural frequency, the identification accuracy of low-order frequency is higher than that of high-order frequency. As the actual structural stiffness is less than the modeling stiffness, the identification result is also less than the ANSYS theoretical calculation result.

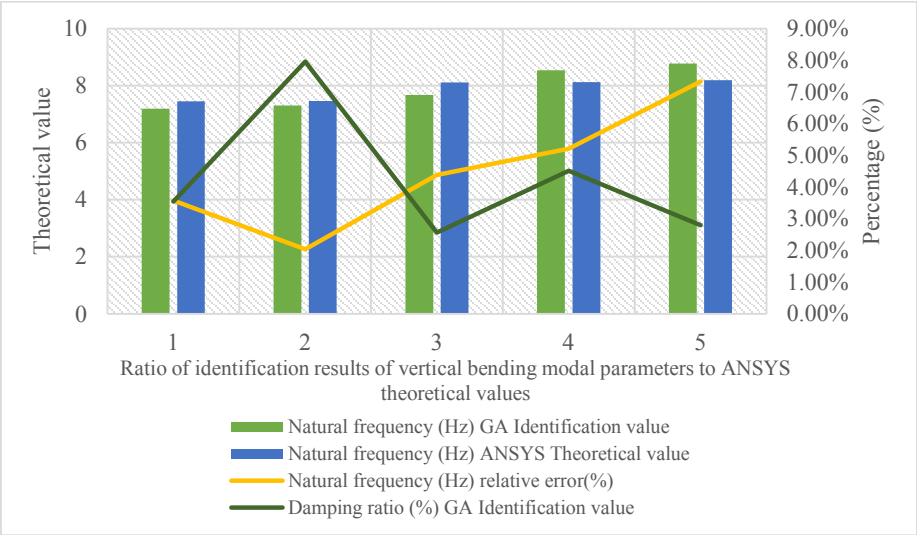


Figure 2. Ratio of identification results of vertical bending modal parameters to ANSYS theoretical values

In this paper, by combining signal filtering and random decrement method, genetic algorithm is applied to MPI, that is, HGA is applied to bridge MPI. It can be seen from Fig. 1 and Fig. 2 that the minimum frequency error is 1.93%, the maximum error is 9.33%, and the first three frequency errors are within 6%, but the error increases with the increase of modal order, so the accuracy problem when used to identify high-order modes is worth considering. Genetic algorithm is applied to MPI, and modal order determination has a great impact on the results of parameter identification, which shows the feasibility and effectiveness of HGA applied to bridge structural MPI.

5 Conclusions

At present, many scholars at home and abroad have done a lot of research on MPI methods and achieved rich results. However, each identification method has certain limitations. It is particularly important to apply a new method to MPI to overcome the limitations of existing methods. For this purpose, this paper combines genetic algorithm with random decrement technology and signal de-noising technology to propose a bridge structure MPI method based on HGA. Although some achievements have been made, there are still many shortcomings worth further study: MPI is based on test signal analysis, so the identification results are greatly affected by the quality of test

data, The quality of the data even directly affects the development of the identification work. How to process the signal to extract the effective information of the structure in the case of weak vibration signal and general test data quality still needs further research; In this paper, due to the limited resources, the research object is a single bridge type. Both the model bridge and the actual bridge are cable-stayed bridges. Whether other types of bridges can also achieve good identification results needs further research and verification.

References

1. Matsubara M, Kawamura S. Parameter Identification of a Three-dimensional Flexible Ring-based Model of a Tire Using Experimental Modal Analysis. *International Journal of Automotive Engineering*, 2019, 10(2):133-138.
2. Schfle T R, Mitschke M, Uchiyama N. Generation of Optimal Coverage Paths for Mobile Robots Using HGA. *Journal of Robotics and Mechatronics,* 2021, 33(1):11-23.
3. Watanabe S, Keyaki T, Naito N, et al. Automatic Identification Method for Natural Frequency of Bridge Piers by Microtremor Measurement at Both Sides on Top of Pier. *Quarterly Report of RTRI*, 2020, 61(2):103-108.
4. Silva M S, Neves F A. Modal identification of Bridge 44 of the Carajás Railroad and numerical modeling using the finite element method. *Revista IBRACON de Estruturas e Materiais*, 2020, 13(1):39-68.
5. Fan L, Liu X, Cai G P. Dynamic modeling and modal parameters identification of satellite with large-scale membrane antenna. *Advances in space research*, 2019, 63(12):4046-4057.
6. J Naranjo-Pérez, JF Jiménez-Alonso, A Sáez. Parameter identification of the dynamic Winkler soil–structure interaction model using a hybrid unscented Kalman filter–multi-objective harmony search algorithm. *Advances in Structural Engineering*, 2020, 23(12):2653-2668.
7. Omidalizarandi M, Herrmann R, Kargoll B, et al. A validated robust and automatic procedure for vibration analysis of bridge structures using MEMS accelerometers. *Journal of Applied Geodesy*, 2020, 14(3):327-354.
8. Ahmad M, Kumar N, Kumari R. A HGA Approach To Solve Inverse Kinematics Of A Mechanical Manipulator. *International Journal of Scientific & Technology Research*, 2019, 8(9):1777-1782.
9. Mathur A. Hybrid Combination of Error Back Propagation and Genetic Algorithm for Text Document Clustering. *International Journal of Computer Trends and Technology*, 2020, 68(11):64-68.
10. Khalaf J A, Majeed A A, Aldlemy M S, et al. Hybridized Deep Learning Model for Perfobond Rib Shear Strength Connector Prediction. *Complexity*, 2021, 2021(8):1-21.
11. Shrividya G. Application of HGA for Successful CS-MRI Reconstruction. *Journal of Advanced Research in Dynamical and Control Systems*, 2020, 12(3):408-414.
12. Sun L, Xu Y. MPI and finite element model updating of a long-span aqueduct structure based on ambient excitation. *Journal of Vibroengineering,* 2020, 22(3):896-908.

Overload Damage Detection Method of Motor Car Axle Based on Neural Network Algorithm

Pin Xia(✉)

College of Intelligent Manufacturing and Automobil, Chongqing Vocational College of Transportation, Chongqing 402247, China
(✉) Corresponding author: xarp@163.com

Abstract. Neural network is a new theoretical model. It has the ability of parallel processing. It can classify, define and optimize information and knowledge by simulating biological neural system. In this paper, a typical nonlinear deformation and damage monitoring method is trained by BP algorithm, which is based on neural network to detect the axle overload strength. Firstly, the fatigue response characteristics of the corresponding working conditions (such as low speed) under different stress states on the axle when the method is running in the motor car are studied by experimental method. Secondly, the actual working environment is simulated as the process of high-speed driving through the design model, and the detection degree of axle overload damage in this scenario is tested by the model. Finally, the test results show that the running time of the motor car axle overload damage detection model based on neural network algorithm is relatively short, and the delay time is also relatively short. The probability of checking the overload damage is basically more than 90%, which shows that the motor car axle overload damage detection rate of this model is very high and can meet the needs of users.

Keywords: Neural Network Algorithm, Motor Car Axle, Overload Load Loss and Damage Detection

1 Introduction

With the rapid development of social economy and technology, highway traffic plays a more and more important role in cities, and traffic accidents are also increasing [1-2]. Therefore, it is particularly urgent to evaluate the safety of vehicles. In order to minimize the personal and property losses and maximize the driving speed, it is necessary to develop an action efficiency that can accurately predict the accident probability, and take corresponding measures in time to ensure that the personal and property will not be damaged. Neural network is a nonlinear system formed by a large number of neurons through simulation. It has the advantages of good approximation performance and strong fault tolerance, and is widely used in the field of traffic safety [3-4].

Many scholars at home and abroad have done relevant research on neural networks. Neural network is a new computer-aided system. It has been widely used in the field of biological intelligence and human brain, and has been widely used in various scientific and engineering designs [5-6]. The research on sports injury detection

technology started early in foreign countries. The United States, Germany and other developed countries have begun to use this technology for fault diagnosis and maintenance. American scholars have proposed artificial neuron simulation method to predict the damage degree of vehicle axles. Japanese scholars have developed a nonlinear finite element simulation software based on BP algorithm - fuzzy bases and artificial neural network to deal with the stress distribution and size change law of the top of the car model, and optimize the model on the computer [7-8]]. There are also some mature companies in China that are developing integrated intelligent vehicle component detection methods and research work based on artificial neural network (annr), BP algorithm and other artificial intelligence systems, and have achieved some results. The above research has laid the research foundation for this paper.

Neural network is a nonlinear system analysis method, which has the characteristics of high parallelism and good robustness. It is widely used in solving complex problems. In this paper, the intelligent traffic monitoring platform is modeled based on the principle of neural network algorithm. Firstly, the overload damage detection technology and working process of intelligent axle are introduced. Then the linear crack initiation mechanism is established based on BP algorithm under a certain working condition, and the corresponding diagnosis model and method are proposed. Finally, using the research results, the influence factors of different parameters on the deformation characteristics, contact stress distribution and crack propagation of vehicle journal are analyzed.

2 Discussion on Overload Damage Detection Method of Motor Car Axle Based on Neural Network Algorithm

2.1 Overload Measurement Method of Motor Car Axle

The overload load detection method of vehicle axle is mainly based on artificial neural network, which designs and learns the structure of cerebral cortex by simulating human brain neurons and external signal stimulation, so that it has better anti fatigue and strong robustness. It also includes the direct contact method [9-10]. The test is to calculate whether the deformation occurs at the corresponding parts by manually collecting the radial tensile stress, rotation angle and other data at different positions of the upper body. However, this method can only obtain a point strain diagram with a direction parallel to the centerline of the wheel axis, and can not obtain the displacement curves and corresponding angle values of all cut-in points on the centerline diagram in the axis top plane. Because the traditional manual measuring equipment has certain limitations in on-line vehicle monitoring, and its work efficiency is also low, it requires a lot of manpower to complete data collection and other operations. At the same time, manual ranging can not meet the requirements of real-time dynamic monitoring and the defects of slow data processing speed and low accuracy. It is also common for vehicles to suffer from axle overload damage caused by various factors during driving Random occurrence.

2.2 Influence of Axle Overload on Motor Car

Axle damage refers to the deformation of vehicle body caused by external force during driving, resulting in the bending of vehicle body surface or interior, wheel locking

(depression), roll and tire wear. What affects the overload fatigue life of the axle is that the changes of its main braking performance and structural parameters interfere with the test results to a great extent [11]. When there is a certain error between the vehicle motion track and the actual situation, the measured value will deviate, resulting in inaccurate measurement. During the running process of the motor car, the axle is subjected to the force between the wheel and the track, resulting in bending deformation, torsion and compression. When the vehicle body is damaged by the ground applied to the vehicle (such as the front wheel) and the steering linkage (or the rear wheel), the vibration waveform will be distorted and the vehicle body will shake or roll over. At the same time, under the driving state of the vehicle body, due to the gravity of the vehicle itself, the axle will also be bent, deformed, twisted and compressed due to excessive external force.

2.3 Factors Affecting Overload of Motor Car Axle

The main factors affecting vehicle axle overload are: (1) operating conditions. This includes driving speed, number of stops, etc. In practical work, the requirements for the bearing force and stiffness of the frame are different under different working conditions. At high speed, the braking pressure is large and the deceleration is slow. At low speed, the vehicle speed is fast but the braking distance is long and there are some obstacles to restrict its normal movement or, if it is necessary to reduce the range, the detection method must be used to evaluate and determine whether there is necessary to leave enough clearance between the vehicle axle and the rail to ensure safety. (2) Load characteristics. The vehicle is subject to a variety of forces during driving, mainly gravity, wind, etc., and will also be affected by various power sources in different directions and angles. Therefore, the bearing capacity of the axle to the track is different and variable. At the same time, considering the friction resistance between the vehicle body and the bridge wall and the structural stiffness problems, the deformation of the vehicle wheels may cause the vertical vibration, bending, deflection or even fracture failure of the vehicle body, resulting in serious consequences such as vehicle safety accidents or frequent traffic accidents.

2.4 Neural Network Algorithm

Neural network is a new and widely used information processing model. It simplifies and parallelizes the functions of neurons in the human brain connecting with the outside world, so that it can achieve the maximum performance goal of human brain's cognitive ability when solving complex problems. It is designed to simulate the structure of biological nervous system. When processed in the computer, the input and output signals are connected with different types of neurons. Through the storage of neural information and learning rules to achieve the automatic adaptability of the artificial system. BP routing protocol can be divided into three layers: perception layer, hidden layer and application service layer. It has a strong self-learning, self-learning and automatic adaptation system. This network is a mathematical algorithm model. By adjusting the connection weights and other parameters between a large number of internal neurons, it can intelligently learn and train the input and output sample data, mine the potential relationship between input and output, and have the ability to calculate and predict new samples to obtain the prediction results. Neural algorithm can be used to predict the changing trend of unknown parameters (inputs) in the model

without any external factors to modify the model parameters, so as to obtain the optimal results.

Input information Xi and threshold in neuron expression θ K constitutes a linear combination, so the threshold can be regarded as a specific input information, then input x0= − 1, and the corresponding weight wk0=k to obtain:

$$u_k = \sum_{i=0}^{m} w_{ki} x_i \tag{1}$$

There are two kinds of nodes in the network: input node and calculation node. The input node only receives signals, and the calculation node is the unit neuron. So the final mathematical expression of unit neuron is:

$$y_k = \varphi\left(\sum_{i=0}^{m} w_{ki} x_i \right) \tag{2}$$

The single-layer neural network, which puts many neurons on the same computing level, is just an output layer, which can solve the linear separable problem well, but it can not deal with the nonlinear separable problem at the same time. Therefore, a single neuron can be regarded as a multi input and single output system, while a single-layer neural network can be regarded as a multi input and multi output system, but their working mechanisms are not much different in essence.

3 Experimental Process of Overload Damage Detection Method of Motor Car Axle Based on Neural Network Algorithm

3.1 Process of Axle Overload Damage Detection Method Based on Neural Network Algorithm

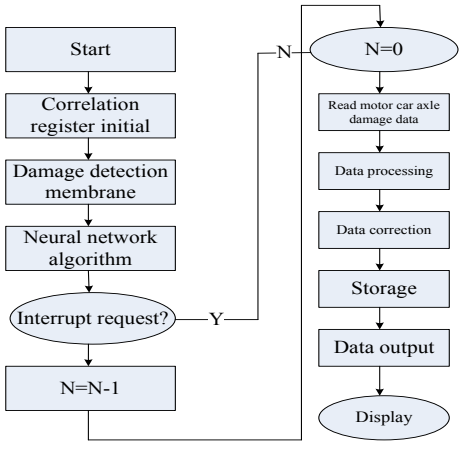


Figure 1. Axle overload damage detection method and flow

It can be seen from Figure 1 that according to the training and simulation results of neural network, combined with the actual working conditions, the detection method of axle overload damage has been deeply studied, and it is concluded that the neural network model based on BP algorithm has good anti noise ability when the factors such as the motion track, initial state and node position in the vehicle change. The intelligent traffic safety system based on BP algorithm can effectively prevent traffic accidents. The system is composed of several subsystems. Firstly, judge whether each subsystem has fault (i.e. whether it is an accident) through expert experience. Secondly, predict the axle overload damage. During the detection process, the axle will be slightly damaged due to various factors on the body surface, such as temperature, humidity, etc. According to the neural network model, the deformation of vehicle body under different working conditions is analyzed. It is mainly described by establishing the relative position relationship between the corresponding nodes between the grid and the vehicle contact area, and then using the geometric coordinate system of the grid and the contact surface to convert it into a standard state vector, and calculate the corresponding input and output values.

3.2 Test Steps for Overload Damage Detection of Motor Car Axle Based on Neural Network Algorithm

The basic idea of neural network algorithm is to classify the nodes in the system layer by layer through input and output neurons, and aggregate the data sets of different types, sizes and attributes according to certain rules to form a parallel processing function with strong adaptability, rich information and good global optimization ability that can be combined with other topological structures. When the artificial neural network algorithm is applied in fault diagnosis, it should be determined according to the damage model and system state, and the methods used in different cases will be different. Therefore, in order to ensure the training success rate and data processing effect. First, initialize the vehicle body. That is, the vehicle body starts to learn and complete layer by layer from static to motion and from motion to stop. Secondly, the starting point of each iteration is the input and output current sampling value at the nodes of each part on the axle. After the corresponding damage model is established, it is necessary to collect the initial state and operating environment of the system. The physical model in the process of vehicle driving is established according to the kinematics theory. Secondly, the vehicle speed, acceleration and other parameter values are obtained through the training set and used as the prediction basis. The input and output are pre estimated by BP neural network algorithm and the error signal correction is calculated and analyzed. Finally, the axle overload damage detection task is realized.

4 Experimental Analysis of Overload Damage Detection Method of Motor Car Axle Based on Neural Network Algorithm

Detection and Analysis of Overload Damage of Motor Car Axle Based on Neural Network Algorithm

Table 1 shows the performance test results of the neural network algorithm.

Table 1. Performance test results

Number of tests	Intelligent algorithm	System operation time (s)	System delay time (s)	Damage detection rate
1	Artificial neural algorithm	3	2	91%
2	Artificial neural algorithm	5	1	94%
3	Artificial neural algorithm	3	2	93%
4	Artificial neural algorithm	2	2	96%
5	Artificial neural algorithm	4	3	93%

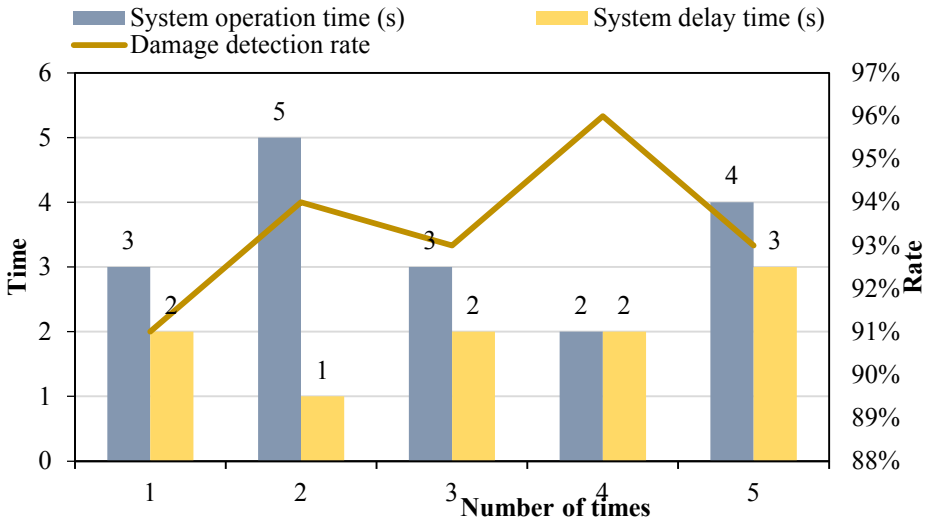


Figure 2. System test

This paper mainly studies the basic principles and related theories of neural network algorithm, and combines the actual motor car axle overload damage detection experiment to complete the intelligent traffic accident early warning, vehicle safety protection and rescue based on BP neural network. After the whole vehicle simulation platform is built, it is necessary to judge whether there is a fault according to the actual working conditions, and then compare the collected data with the operating conditions of the standard sample vehicle in the system to determine the detection method and performance indicators under the overload state of the motor car, and diagnose and evaluate different working conditions. It can be seen from Figure 2 that the detection model of motor car axle overload damage based on neural network algorithm has a short running time and a short delay time. The probability of checking the overload damage is basically above 90%, which shows that the detection rate of motor car axle overload damage of this model is very high and can meet the needs of users.

5 Conclusions

Neural network is a new and large-scale application field. It plays an important role in solving complex engineering problems and improving system performance. It is especially suitable for dynamic characteristic analysis under some nonlinear or uncertain working conditions. With the development and wide application of artificial intelligence technology and computer soft science, and the deepening understanding of neural network theory, an intelligent detection method for axle overload damage diagnosis based on artificial neuron is proposed. This paper analyzes and summarizes the neural network detection method of axle over strength damage. Firstly, the three-dimensional model is established and the body structure characteristics and vehicle driving conditions are modeled. Secondly, the nonlinear design unit is constructed by using BP neural network theory and kinematics equations to improve the identification accuracy. Finally, the fatigue response and strain degradation of the system under different types are verified by experiments, which has good application prospects and practical significance.

Acknowledgements

This work is partially supported by the Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJZD-K202105701)

References

1. Ahmed H. Janabi, Triantafyllos Kanakis, Mark Johnson:Convolutional Neural Network Based Algorithm for Early Warning Proactive System Security in Software Defined Networks. *IEEE Access*. 10: 14301-14310 (2022).
2. Ahmed H. Janabi, Triantafyllos Kanakis, Mark Johnson:Convolutional Neural Network Based Algorithm for Early Warning Proactive System Security in Software Defined Networks. *IEEE Access* 10: 14301-14310 (2022).
3. Febryan Setiawan, An-Bang Liu, Che-Wei Lin:Development of Neuro-Degenerative Diseases' Gait Classification Algorithm Using Convolutional Neural Network and Wavelet Coherence Spectrogram of Gait Synchronization. *IEEE Access* 10: 38137-38153 (2022).
4. Y. K. Bharath:Griffiths' Variable Learning Rate Online Sequential Learning Algorithm for Feed-Forward Neural Networks. Autom. Control. *Comput. Sci*. 56(2): 160-165 (2022).
5. Parameshwaran Ramalingam, Abolfazl Mehbodniya, Julian Webber, Mohammad Shabaz, Gopalakrishnan Lakshminarayanan:Telemetry Data Compression Algorithm Using Balanced Recurrent Neural Network and Deep Learning. Comput. *Intell. Neurosci*. 2022: 4886586:1-4886586:10 (2022).
6. I. Kalphana, T. Kesavamurthy:Convolutional Neural Network Auto Encoder Channel Estimation Algorithm in MIMO-OFDM System. *Comput. Syst. Sci. Eng*. 41(1): 171-185 (2022).
7. Sergi Abadal, Akshay Jain, Robert Guirado, Jorge López-Alonso, Eduard Alarcón:Computing Graph Neural Networks: A Survey from Algorithms to Accelerators. *ACM Comput. Surv*. 54(9): 191:1-191:38 (2022).
8. Soodeh Hosseini, Ali Emamali Nezhad, Hossein Seilani:Botnet detection using negative selection algorithm, convolution neural network and classification methods. *Evol. Syst*. 13(1): 101-115 (2022).

9. Iman Shafieenejad, Elham Dehghan Rouzi, Jamshid Sardari, Mohammad Siami Araghi, Amirhosein Esmaeili, Shervin Zahedi: Fuzzy logic, neural-fuzzy network and honey bees algorithm to develop the swarm motion of aerial robots. *Evol. Syst.* 13(2): 319-330 (2022).
10. Dhaya Ramakrishnan, Kanthavel Radhakrishnan: Applying deep convolutional neural network (DCNN) algorithm in the cloud autonomous vehicles traffic model. *Int. Arab J. Inf. Technol*. 19(2): 186-194 (2022).
11. S. Deepika, S. Senthil:Credit card fraud detection using moth-flame earth worm optimisation algorithm-based deep belief neural network. Int. J. Electron. *Secur. Digit. Forensics* 14(1): 53-75 (2022).
12. Amrit Kaur Bhullar, Ranjit Kaur, Swati Sondhi: Modified neural network algorithm based robust design of AVR system using the Kharitonov theorem. *Int. J. Intell. Syst*. 37(2): 1339-1370 (2022).
13. Armin Salimi-Badr, Mohammad Mehdi Ebadzadeh: A novel learning algorithm based on computing the rules' desired outputs of a TSK fuzzy neural network with non-separable fuzzy rules. *Neurocomputing* 470: 139-153 (2022).