Received: 28 October 2021 | Revised: 15 December 2021 | Accepted: 21 December 2021 | Published online: 11 January 2022

Journal of Computational and Cognitive Engineering
2022, Vol. 1(2) 83–87
DOI: 10.47852/bonviewJCCE19522514
BON VIEW PUBLISHING

RESEARCH ARTICLE

Research on Microvideo Character Perception and Recognition Based on Target Detection Technology

Yue Lei[1,*]

[1]*Gingko College of Hospitality Management, China*

Abstract: With the rapid development of the Internet short-video platform, microvideo is increasingly attracting users because of its small traffic and time period. However, the content of microvideo on the Internet is complex, with illegal information. This research is an improved algorithm based on target detection technology, aiming at the characteristics of difficult character recognition in microvideo. The algorithm combines Gabor wavelet transform algorithm and 2D principal component analysis (PCA) algorithm and makes full use of video perception technology to realize character recognition in microvideo. Through MATLAB simulation analysis, it can be seen that the accuracy of the optimization algorithm proposed in this paper is 86.34% and the recognition time is 34.28s. Compared with the traditional PCA algorithm and artificial neural network (ANN) algorithm, the optimization algorithm proposed in this paper has better recognition rate and recognition efficiency.

Keywords: target detection, microvideo, face recognition, Gabor, 2DPCA algorithm

1. Introduction

Microvideo character recognition is conducive to maintaining the security of the Internet and has special significance for national security control, counterterrorism, and pornography prohibition (Shi et al., 2019; Su et al., 2018; Tueller et al., 2020; Yu et al., 2021). Since the development of the Internet, various short-video platforms have been developing rapidly. If only the content in microvideo is identified manually, the efficiency will be very low in the face of massive video (Li et al., 2018; Li et al., 2020; Rui et al., 2018; Sakulchit et al., 2019). With the continuous development of modern biometric recognition technology, the application range of face recognition technology is becoming wider. Microvideo character recognition based on target detection technology can achieve more flexible face recognition and target detection, overcome the influence of complex background environment and other influencing factors in video shooting, and solve the situation that the face may be blocked (Liu, 2019; LoSardo, 2019 Matsumura et al., 2021; Xue, 2019; Zhao et al., 2018). This research tries to propose a mature microvideo character recognition technology, which can be applied to practical needs.

Dutta et al. (2021) proposed a mathematical model of face image decomposition and introduced a genetic algorithm to optimize and improve the model, which can effectively remove redundant eigenvalues and improve face recognition performance. Face recognition will be affected by many external factors, so multiobjective optimization methods can be adopted to realize effective face recognition, while multiobjective genetic programming algorithms can realize multiobjective combinatorial optimization (Bi et al., 2021; Wu, 2021; Wu et al., 2021). Ranade and Anand (2021) proposed a color face recognition method for different color spaces, which has good accuracy and efficiency. At present, the face recognition system has been widely used in various identity recognition application systems. People have conducted in-depth and extensive research on various stages of face recognition, including face feature capture, feature generation, face correction, etc. (Juneja & Rana, 2021).

Compared with the traditional microvideo character recognition technology, the proposed technology has the following innovations: (1) it can realize the processing of microchannel small video under complex background, including the influence of different lighting, climate conditions, and other influencing factors, and can obtain good recognition results. (2) By improving the traditional principal component analysis (PCA) algorithm, a bidirectional 2DPCA algorithm is obtained, which makes it have better face recognition effect. (3) Finally, a technology that can realize face detection, tracking, and recognition is obtained.

2. Relevant Theoretical Basis

Image processing: Image processing technology refers to the use of computer technology to process the target image information according to the established target, including image enhancement, image segmentation, image recognition, image description, and other technologies.

Gabor wavelet transform: Gabor is a linear filter. Its main purpose is to extract edge features. Similar to the human visual system, it is often used for texture feature analysis and extraction and is not sensitive to illumination changes.

*Corresponding author: Yue Lei, Gingko College of Hospitality Management, China. Email: leiyuebess@126.com.

Image feature extraction: Image feature extraction is an advanced stage of digital image processing. It is a transition process from digital image to image recognition and collects some features or characteristics in the image.

2DPCA algorithm: Its full name is two-dimensional face recognition feature extraction algorithm, which is an improvement on the traditional PCA feature algorithm. It operates the face image itself, and there is no need to process the image into column vectors before calculation.

Face feature classification: Face feature classification is one of the key technologies for face feature extraction. The features of face recognition mainly include geometric features, model features, statistical features, and neural network features. The features of face can be confirmed and recognized using the recognition algorithm.

3. Image Feature Representation Based on Gabor Wavelet Transform

Gabor wavelet transform algorithm uses time and frequency to describe the process of time function and solve the Gabor coefficient. Gabor transform cannot be transformed after the window size is calculated. Therefore, this method has no zoom characteristics and cannot carry out multiscale signal analysis (Kasthuri et al., 2019; Xu et al., 2019; Zhang et al., 2020). Therefore, on this basis, scholars proposed two-dimensional Gabor wavelet transform, which can not only maintain the multiscale characteristics of wavelet transform, but also have directionality and locality. The expression of twodimensional Gabor wavelet filter is

$$\psi_j(\vec{x}) = \frac{\left\|\vec{k}_j\right\|^2}{\sigma^2}\exp\left(-\frac{\left\|\vec{k}_j\right\|^2\|\vec{x}\|^2}{2\sigma^2}\right)\left(\exp\left(i\vec{k}_j\vec{x}\right) - \exp\left(-\frac{\sigma^2}{2}\right)\right) \tag{1}$$

In equation (2), $\frac{\|\vec{k}_j\|^2}{\sigma^2}$ is the attenuation energy spectrum compensation function, and $\exp\left(ik_j\vec{x}\right)$ is the complex plane wave, its real part is $\cos\left(k_jx\right)$ and its imaginary part is $\sin\left(k_jx\right)$. The effect of wavelet transform can be maintained through $\exp\left(ik_j\vec{x}\right)$ function, which will not be affected by the gray value and illumination change of the image. The feature of $\vec{x}$-neighborhood image can be extracted using the change of the two-dimensional Gabor wavelet, the corresponding Gabor wavelet can be obtained, and the Gaussian filter bank can be obtained by ψ_j.

Let the face image set be $I(\vec{x})$, the corresponding value on the image domain is $\vec{x}$, the Gabor face is $O_{u,v}(\vec{x})$, and its expression is

$$O_{u,v}(\vec{x}) = I(\vec{x}) * \psi_j(\vec{x}) \tag{2}$$

Further, let the scale and direction of Gabor filter be reflected as $\vec{k}_j$, and the size of two-dimensional Gabor wavelet is σ, which is taken as $\frac{\pi}{2}$. They jointly determine the size of two-dimensional Gabor wavelet. u,v in equation (2) determines the sampling direction and scale, respectively, and the expression of $\vec{k}_j$ is

$$\vec{k}_j = \begin{pmatrix} k_{jx} \\ k_{jy} \end{pmatrix} = \begin{pmatrix} k_v \cos\varphi_u \\ k_v \sin\varphi_u \end{pmatrix} \tag{3}$$

In equation (3), k_v is taken as $2^{\frac{v+2}{2}}$ and φ_u is taken as $u\frac{\pi}{8}$.

In the process of face image recognition, Gabor convolution operation may occur in the process of operation, resulting in the adverse impact of edge domain oscillation on matching recognition. Therefore, it is necessary to remove the linear features in the process of Gabor transform, but the corresponding amplitude must be retained to reflect the energy characteristics of image region and realize face image recognition.

4. Face Image Feature Extraction Technology Based on Bidirectional 2DPCA Algorithm

In face recognition, the extraction of face features should be calculated by instruments or algorithms, so as to obtain the original features of face image. However, considering that the dimension of the original features is too large, it is necessary to reduce the dimension of the original feature data, so as to realize feature extraction and improve the efficiency of the algorithm (Benitez-Garcia et al., 2019; Liao & Gu, 2019; Zhang et al., 2020). The resulting features are quadratic features. The result of feature extraction has a direct impact on the final face recognition, so the representative and small redundancy face features must be retained. This time, PCA is introduced for face image feature extraction. Due to the large amount of calculations and relatively low accuracy of the traditional PCA algorithm, the bidirectional 2DPCA algorithm optimized for many times is selected. Although the algorithm will be applied to more coefficients, considering the bidirectional operation of rows and columns, it can greatly improve the face recognition rate and reduce the face reconstruction coefficients.

Let the training sample set be X_i, the average image matrix be u, and the row vectors of x_i and u on row j are $x_i^{(j)}$,$u^{(j)}$, respectively, so

$$x_i = \left[\left(x_i^{(1)}\right)^T\left(x_i^{(2)}\right)^T\cdots\left(x_i^{(m)}\right)^T\right]^T \tag{4}$$

$$u = \left[\left(u^{(1)}\right)^T\left(u^{(2)}\right)^T\cdots\left(u^{(m)}\right)^T\right]^T \tag{5}$$

Let the image covariance moment in the row direction be G, and its expression is

$$G = \frac{1}{M}\sum_{i=1}^{M}\sum_{j=1}^{m}\left(x_i^{(j)} - u^{(j)}\right)^T\left(x_i^{(j)} - u^{(j)}\right) \tag{6}$$

As above, the column vectors of x_i and u set on column j are $x_i^{(j)}$, $u^{(j)}$, respectively, and the calculation formula of image covariance moment G in the column direction is

$$G = \frac{1}{M}\sum_{i=1}^{M}\sum_{j=1}^{m}\left(x_i^{(j)} - u^{(j)}\right)\left(x_i^{(j)} - u^{(j)}\right)^T \tag{7}$$

Let the orthogonal matrix of the column vector be $Z \in R^{m*q}$ and the input sample image matrix be X, and its projection in the Z direction can obtain a $q*n$ matrix. The expression is

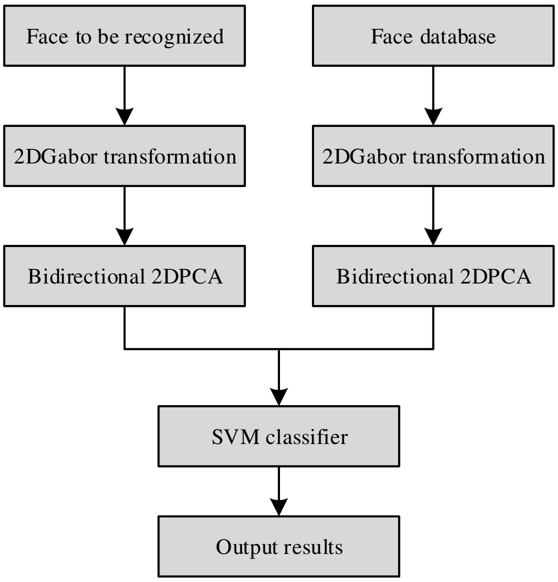
$$B = Z^TX \tag{8}$$

Let the judgment criterion be $J(Z)$, and its expression is

$$J(Z) = trace\left\{Z^TE\left[X - EX(X - EX)^T\right]Z\right\} \tag{9}$$

Thus, the projection matrix Z_{opt} of the optimal eigenvector can be formed, and its expression is

Figure 1
Figure recognition flowchart in microvideo



$$Z_{opt} = \left[z_1, z_2, \cdots, z_q \right] \quad (10)$$

The face recognition features based on bidirectional 2DPCA algorithm are compared and matched with the extracted features in the face database, and the matched images are obtained through the classifier. The overall process is shown in Figure 1.

5. Face Feature Classification

After the extraction of face features in wechat video, face features need to be classified in order to recognize accurately (Yuan et al., 2019; Zhi & Liu, 2019). This time, support vector machine (SVM) is selected to deal with a large number of image problems. Its working principle is to construct an optimal classification surface, which can distinguish different samples and maximize the linear distance between different samples. The classification diagram is shown in Figure 2.

In Figure 2, two different training samples, square and circular, are separated by three straight lines, where H represents the optimal classification line of all samples, H_1, H_2 passes through the point closest to the optimal classification line, and the distance between them is the classification interval. Through the optimal classification line, SVM can be classified with high accuracy.

Let the normal vector of the classification surface be w, the sample set be (x_i, y_i), and the linear discriminant function of the high-dimensional space be

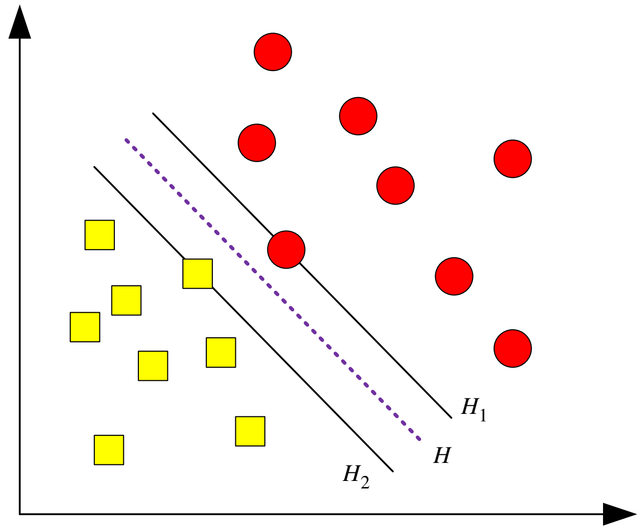
$$f(x) = w \cdot x + b \quad (11)$$

Normalize it to make it conform to $|f(x)| \geq 1$. In order to achieve accurate differentiation of all samples, it is necessary to meet the following requirements:

$$y_i[(w \cdot x_i) + b] - 1 \geq 0, \ i = 1,2,\cdots,n \quad (12)$$

For nonlinear problems, it is necessary to introduce a relaxation factor for classification. When the number of samples is very large, it is difficult

Figure 2
Optimal classification diagram based on SVM



to obtain a good classification effect even if the relaxation factor is introduced. Therefore, it is necessary to transform the extracted image feature vector to the high Viterbi eigenspace through the kernel function and transform the nonlinear problem into a linear problem in the high-dimensional space. Its schematic diagram is shown in Figure 3.

The final output of SVM is the linear combination of support vectors of the input image data information. With the increase in the number of training samples, the support vector will also increase, which will greatly increase the amount of calculation of SVM discriminant function. Therefore, it should be properly selected when selecting training samples.

In the actual face feature classification and recognition, because it is necessary to deal with multiple classification problems at the same time, when designing the classifier, it is necessary to transform a variety of problems into two types of problems as much as possible to facilitate the processing of SVM. This processing method mainly includes "one-to-one" or "one-to-many" classification strategies. The "one-to-one" classification strategy refers to the pair-wise matching of different types of samples. Using this strategy, even the training of large samples will not affect the decision-making efficiency. When using the "one-to-many" classification strategy, it is mainly used when the number of samples is small, which can obtain better training speed and classification effect.

There are four samples S_1, S_2, S_3, S_4, and an SVM classifier is constructed for the samples:

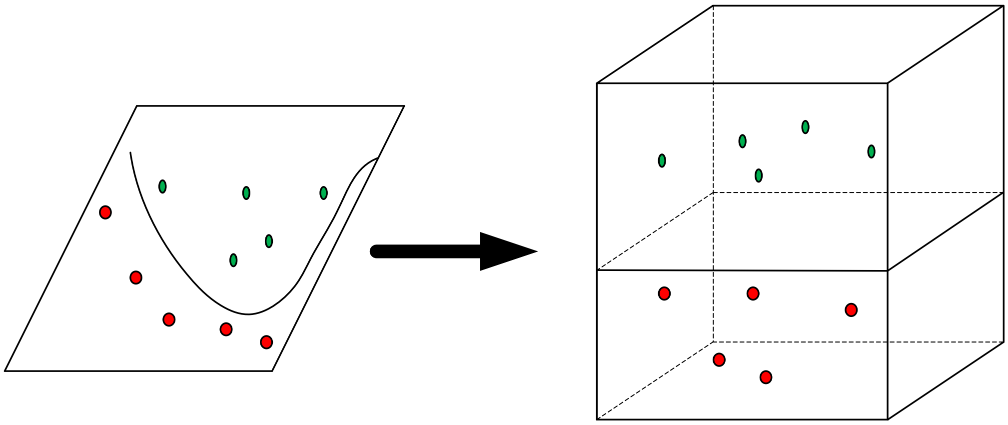
$$f_i(i = 1,2,3,4) \quad (13)$$

Each classifier will use the features of one of the four types of samples for training. The class label is displayed as "+1," and the other sample feature labels are defined as "−1" negative set. During training, if no sample feature response is "+1," it will be divided into corresponding types. If the corresponding response is "−1," it will be divided into new samples.

6. Simulation Analysis

In order to verify and analyze the performance of the method proposed in this paper, 40 people are selected to shoot this time, and each person shoots 10 microvideo clips, a total of 400 microvideo clips within 10s. 1600 microvideo clips related to characters were

Figure 3
Schematic diagram of transforming nonlinear problems into linear problems in high-dimensional space



randomly selected from the database, and 400 captured microvideos were put into it, with a total of 2000 microvideo clips.

The improved algorithm in this paper is compared with the traditional PCA algorithm and artificial neural network (ANN) algorithm, and the recognition efficiency and accuracy of the algorithm for characters in microvideo are compared and analyzed respectively. The results are shown in Figures 4 and 5.

Through the analysis of Figures 4 and 5, it can be seen that in the comparison of accuracy and recognition efficiency of the three algorithms, the optimized and improved algorithm in this paper has better performance. In Figure 4, as the number of recognized images increases, the recognition rate gradually decreases. When the number of video images reaches 400, the optimized and improved algorithm in this paper improves by 33.77% and 11.20%, respectively, compared with the traditional PCA algorithm and ANN algorithm. In Figure 5, with the increase in the number of video images, the algorithm time also gradually increases. When it reaches 400, the optimization algorithm proposed in this paper takes the least time, which is 36.64% and 11.08% lower than the traditional PCA algorithm and ANN algorithm, respectively. The comprehensive analysis shows that the 2DPCA algorithm proposed in this paper has better performance and better application effect in the research of microvideo character recognition.

Compared with the traditional PCA algorithm, 2DPCA algorithm can make up for some of its shortcomings, and 2DGABOR change can effectively reduce the feature dimension, classify, and recognize the face images in microvideo. Through

Figure 4
Comparative analysis of recognition accuracy of three algorithms

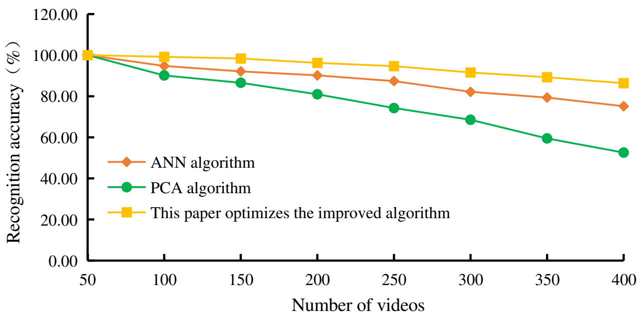
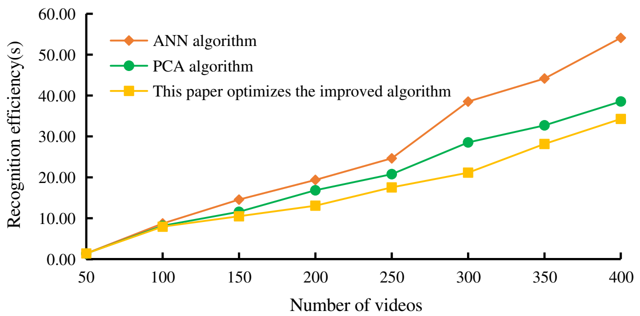


Figure 5
Comparative analysis of recognition efficiency of three algorithms



the comparative analysis of the two algorithms, it can be clearly seen that the improvement effect of this algorithm is very significant. In the actual simulation analysis process, even the face in the complex background can be well recognized, and the fusion algorithm has a very high recognition rate. Face recognition using PCA algorithm not only takes a long time, but also the recognition rate is low. 2DPCA algorithm can improve the efficiency of image recognition by constructing covariance matrix.

7. Conclusion

Microvideo character recognition can effectively combat illegal acts on the Internet. Therefore, a fusion algorithm based on 2DPCA and the two-dimensional Gabor change method is proposed to realize character recognition, and the input image features are effectively classified by SVM classifier. In order to verify the effectiveness of the optimized and improved recognition method, through the simulation analysis of 400 microvideos, the recognition efficiency and accuracy of traditional PCA algorithm and ANN algorithm are compared. Through simulation analysis, it can be seen that compared with PCA algorithm and ANN algorithm, the optimized and improved algorithm proposed in this paper improves the recognition accuracy by 33.77%, 11.20%, and the algorithm efficiency by 36.64% and 11.08%. Overall, the proposed method has good microvideo character recognition effect and good application value. The proposed wechat small video face recognition technology can effectively detect, track, and recognize faces, especially in complex environments. However, in future research, it is also necessary to improve the detection rate when there are some occlusions or other influencing factors.

Conflicts of Interest

The author declares that he has no conflicts of interest to this work.

References

Benitez-Garcia, G., Nakano-Miyatake, M., Olivares-Mercado, J., Perez-Meana, H., Sanchez-Perez, G., & Toscano-Medina, K. (2019). A low complexity face recognition scheme based on down sampled local binary patterns. *The International Arab Journal of Information Technology*, *16*(3), 338–347.

Bi, Y., Xue, B., & Zhang, M. (2021). Multi-objective genetic programming for feature learning in face recognition. *Applied Soft Computing*, *103*(4), 107152. https://dx.doi.org/10.1016/j.asoc.2021.107152

Dutta, K., Bhattacharjee, D., Nasipuri, M., & Krejcar, O. (2021). Complement component face space for 3D face recognition from range images. *Applied Intelligence*, *51*(10), 2500–2517. https://doi.org/10.1007/s10489-020-02012-8

Juneja, K., & Rana, C. (2021). An extensive study on traditional-to-recent transformation on face recognition system. *Wireless Personal Communications*, *118*, 3075–3128. https://doi.org/10.1007/s11277-021-08170-3

Kasthuri, A., Suruliandi, A., & Raja, S. P. (2019). Gabor oriented local order features based deep learning for face annotation. *International Journal of Wavelets, Multiresolution and Information Processing*, *17*(5), 1950032. https://dx.doi.org/10.1142/S0219691319500322

Li, L., Mu, X., Li, S., & Peng, H. (2020). A review of face recognition technology. *IEEE Access*, *8*, 139110–139120. https://dx.doi.org/10.1109/ACCESS.2020.3011028

Li, L., Peng, Y., Qiu, G., Sun, Z., & Liu, S. (2018). A survey of virtual sample generation technology for face recognition. *Artificial Intelligence Review*, *50*(6), 1–20. https://dx.doi.org/10.1007/s10462-016-9537-z

Liao, M., & Gu, X. (2019). Face recognition algorithm based on feature descriptor and weighted linear sparse representation. *IET Image Processing*, *13*(12), 2281–2293. https://dx.doi.org/10.1049/iet-ipr.2018.5263

Liu, C. (2019). A survey of virtual sample generation technology for face recognition. *Computing Reviews*, *60*(5), 220–221. https://doi.org/10.1007/s10462-016-9537-z

LoSardo, A. (2019). Faceoff: The fight for privacy in American Public Schools in the wake of facial recognition technology. *Seton Hall Legislative Journal*, *44*(2), 373.

Matsumura, H., Taketomi, T., & Kato, H. (2021). Impact of facial contour compensation on self-recognition in face-swapping technology. *Multimedia Tools and Applications*, *80*(3), 7727–7748. https://dx.doi.org/10.1007/s11042-020-09866-7

Ranade, S. K., & Anand, S. (2021). Color face recognition using normalized-discriminant hybrid color space and quaternion moment vector features. *Multimedia Tools and Applications*, *80*(7), 10797–10820. https://dx.doi.org/10.1007/s11042-020-10244-6

Rui, C., Lima, S., & Rocha, Á. (2018). Rocha. Application of face recognition technology based on CA algorithm in intelligent residential property management. *Journal of Intelligent & Fuzzy Systems*, *35*(3), 2909–2919. https://dx.doi.org/10.3233/JIFS-169646

Sakulchit, T., Kuzeljevic, B., & Goldman, R. D. (2019). Evaluation of digital face recognition technology for pain assessment in young children. *The Clinical Journal of Pain*, *35*(1), 18–22. https://dx.doi.org/10.1097/AJP.0000000000000659

Shi, Y., Li, J., Yin, Y., Xi, B., & Li, Y. (2019). Hyperspectral target detection with macro-micro feature extracted by 3-D residual autoencoder. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(12), 4907–4919. https://dx.doi.org/10.1109/JSTARS.2019.2939833

Su, N., Chen, X., Jian, G., Mou, X., & Liu, N. (2018). Detection and classification of maritime target with micro-motion based on CNNs. *Journal of Radars*, *7*(5), 565–574. https://dx.doi.org/10.12000/JR18077

Tueller, P., Kastner, R., & Diamant, R. (2020). Target detection using features for sonar images. *IET Radar Sonar & Navigation*, *14*(12), 1940–1949. https://dx.doi.org/10.1049/iet-rsn.2020.0224

Wu, F., Jing, X. Y., Feng, Y., Ji, Y. M., & Wang, R. (2021). Spectrum-aware discriminative deep feature learning for multi-spectral face recognition. *Pattern Recognition*, *111*(2), 107632. https://doi.org/10.1016/j.patcog.2020.107632

Wu, G. L. (2021). Masked face recognition algorithm for a contactless distribution cabinet. *Mathematical Problems in Engineering*, *2021*(2), 1–11. https://doi.org/10.1155/2021/5591020

Xu, Y., Du, Q., Li, W., & Younan, N. H. (2019). Efficient probabilistic collaborative representation-based classifier for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, *16*(11), 1746–1750. https://dx.doi.org/10.1109/LGRS.2019.2906839

Xue, S. (2019). Face database security information verification based on recognition technology. *International Journal of Network Security*, *21*(4), 601–606.

Yu, H., Sharma, A., & Sharma, P. (2021). Adaptive strategy for sports video moving target detection and tracking technology based on mean shift algorithm. *International Journal of System Assurance Engineering and Management*, 1–11. https://doi.org/10.1007/s13198-021-01128-5

Yuan, X., Liu, Z., Miao, Z., Zhao, Z., Zhou, F., & Song, Y. (2019) Fault diagnosis of analog circuits based on IH-PSO optimized support vector machine. *IEEE Access*, *7*, 137945–137958. https://dx.doi.org/10.1109/ACCESS.2019.2943071

Zhang, Q., Li, H., Li, M., & Ding, L. (2020). Feature extraction of face image based on LBP and 2-D Gabor wavelet transform. *Mathematical Biosciences and Engineering*, *17*(2), 1578–1592. https://dx.doi.org/10.3934/mbe.2020082

Zhao, J., Zhang, C., Zhang, S., Lu, T., Su, W., & Jia, J. (2018). Pre-detection and dual-dictionary sparse representation based face recognition algorithm in non-sufficient training samples. *Journal of Systems Engineering and Electronics*, *29*(1), 200–206. https://dx.doi.org/10.21629/JSEE.2018.01.20

Zhi, H., & Liu, S. (2019). Face recognition based on genetic algorithm. *Journal of Visual Communication & Image Representation*, *58*, 495–502. https://dx.doi.org/10.1016/j.jvcir.2018.12.012

How to Cite: Lei, Y. (2022). Research on Microvideo Character Perception and Recognition Based on Target Detection Technology. *Journal of Computational and Cognitive Engineering* 1(2), 83–87, https://doi.org/10.47852/bonviewJCCE19522514