A Siamese-network-based Facial Recognition System

Chih-Yung Chen,[1] Huang-Chu Huang,[2]
Jyun-Cheng Jheng,[3] and Rey-Chue Hwang[3*]

[1]Program of Artificial Intelligence and Mechatronics, National Pingtung University of Science and Technology,
No. 1, Shuefu Road, Neipu, Pingtung 912301, Taiwan
No. 1, Sec. 1, Syuecheng Rd., Dashu District, Kaohsiung City 84001, Taiwan
[2]Department of Telecommunication Engineering, National Kaohsiung University of Science and Technology,
No. 142, Haijhuan Rd., Nanzih District, Kaohsiung City 81157, Taiwan
[3]Department of Electrical Engineering, I-Shou University,
No. 1, Sec. 1, Syuecheng Rd., Dashu District, Kaohsiung City 84001, Taiwan

In this paper, we introduce a facial recognition system comprising two key components: face detection and localization, and facial recognition. For face detection and localization, the RetinaFace method is employed to accurately identify facial regions within images and to separate them from intricate backgrounds, thus facilitating facial detection based on isolated facial imagery. In the domain of facial recognition, we address the limitations of conventional convolutional neural networks (CNNs), which are typically constrained to recognizing known categories. To overcome this limitation, in our study, we leverage a Siamese network rooted in metric learning as the central architecture for facial recognition. The primary objective of this architecture is to acquire image features. It operates by minimizing the feature distance between similar images and maximizing the feature distance between dissimilar ones. Consequently, images can be directly fed into the Siamese network to extract corresponding features, followed by similarity calculation to ascertain their presence within the database. Diverging from the conventional approach of directly classifying individuals using models, we significantly inhibit the need for model retraining owing to personnel changes in the differentiation of members and nonmembers. Furthermore, the model does not increase in size with the growth of the personnel dataset. The study outcomes demonstrate that the attained average values for accuracy, recall rate, precision, and F1-Score all surpass 96%. These results robustly demonstrate the feasibility and superior performance of this approach.

1. Introduction

In recent years, rapid advancements in artificial intelligence (AI) have led to its widespread integration into various industrial sectors encompassing technology, manufacturing, finance, services, and healthcare, among others. Prominent applications include license plate recognition, autonomous vehicles, AI-driven stock prediction, breast cancer ultrasound diagnostics, AI-

augmented classification, soil and crop monitoring, and predictive analysis. As technology evolves, the safeguarding of personal identity has garnered substantial attention, emphasizing the importance of information security management.

The exploration of biometric identification methods has gained momentum in recent years, encompassing techniques such as face recognition, retinal recognition, fingerprint recognition, voice recognition, and palmprint recognition. Each technique offers distinct advantages, rendering counterfeiting challenging. Among these techniques, face recognition stands out as a particularly natural and user-friendly method, prompting its significant exploration in the domain of identity authentication.

Addressing the limitations of face recognition and enhancing its discrimination accuracy and speed constitute a prominent and ongoing research endeavor. Face recognition involves extracting discernible facial features from images and comparing them against the features of all faces in the database. This comprehensive process encompasses stages from database setup, face detection, and image preprocessing to feature extraction, identification, and verification. The adoption of face recognition has added value to numerous industrial applications, introducing a novel identity validation mechanism and mitigating resource expenditure in areas such as access control, security management, mobile phone facial unlocking, personnel attendance tracking, health management, and facial payment authentication.

Given the rapid evolution of face recognition technology, diverse face recognition systems have been proposed. In recent times, the COVID-19 pandemic has prompted increased hesitancy toward physical contact with public objects. Face recognition technology emerges as a solution to this predicament. By leveraging camera-based face detection, it minimizes the need for physical contact during identity verification, thereby supporting epidemic prevention measures. In this research, we endeavor to establish an access control system grounded in face recognition technology, exemplifying its relevance and potential for various industrial and societal contexts.

Currently, there is a plethora of research endeavors on exploring the application of face recognition systems, encompassing methodologies such as principal component analysis (PCA),[1–5] facial landmark detection,[5–7] support vector machine (SVM),[8–11] and convolutional neural networks (CNNs).[12–16]

PCA is a classic technique for feature extraction and data representation widely applied in the image recognition and computer vision domains. It primarily operates by identifying a linear transformation matrix that is then used for dimensionality reduction. Consider a set of N face images, where the feature vectors of these images are represented as $\{x_1, x_2, x_3, ..., x_N\}$, and each face feature has a size of n. PCA aims to use an $n \times m$ matrix A to condense the original n-dimensional features into new m-dimensional parameters, as shown in the equation below, with $m \leq n$.

$$r_j = A^T x_j, \, j = 1, 2, 3, ..., N \tag{1}$$

Through dimensionality reduction, the face features $\{x_1, x_2, x_3, ..., x_N\}$ of each image are transformed, yielding a fresh set of features $\{r_1, r_2, r_3, ..., r_N\}$ with dimensions lower than those of their pre-transformation counterparts.

PCA's advantage lies in its ability to use the aforementioned outcomes to represent the distinctions between various face images using lower-dimensional features. It captures the most pertinent features of each individual. However, as noted in the literature,[3] PCA also exhibits certain limitations. When the dimension of face images is exceedingly high or the number of images is substantial, the calculation of features, average vectors, and scatter matrices becomes intricate. Additionally, incorporating or modifying a face image necessitates the recalculation of all associated features, average vectors, and scatter matrices, which can be cumbersome and time-consuming.

SVM, a supervised learning algorithm grounded in statistical learning principles, has emerged as a robust solution for addressing classification problems. SVM leverages the concept of empirical risk minimization from statistics to identify a hyperplane that effectively segregates disparate data categories. The primary objective is to ascertain a decision boundary that maximizes the margin between two distinct categories.

In the context of SVM training, a prevalent approach is the one-against-rest training strategy.[10] This strategy involves employing N classifiers for categorization when N categories are present within the dataset. Each classifier exclusively determines its corresponding category, as depicted in Fig. 1. If the jth classifier determines that a given facial image pertains to an alternative category, the similarity between the facial image and the jth facial category is deemed to be 0%. Conversely, if the jth classifier identifies the facial image as belonging to the jth category, it assigns a similarity score of S_j%.

Upon the completion of classification by all the classifiers, the SVM system ascertains the facial image's category on the basis of the highest similarity score across all outputs. As elucidated by Xiao and Li,[8] this approach demonstrates superior exclusivity compared with alternative SVM methods and effectively accommodates the processing of non-database images. Nevertheless, an inherent drawback arises when integrating a new category. In such instances, a corresponding classifier must be introduced. Consequently, as the number of categories increases, the system's dimensions expand, prolonging the computation time. Thus, this approach is deemed inappropriate for the system envisioned within the scope of this research.

The CNN algorithm has been extensively applied in contemporary neural networks and offers robust capabilities in image recognition and natural language processing. CNN comprises essential components, including convolutional, pooling, and fully connected layers. Image features are predominantly extracted through the convolutional and pooling layers, while the
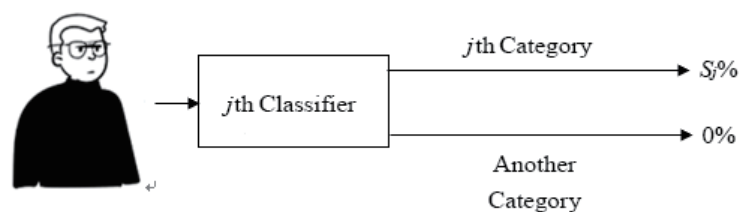

Fig. 1.   (Color online) Schematic depiction of the jth classifier within the one-against-rest training strategy.

fully connected layer outputs category probabilities. The highest probability category is ultimately designated as the judgment result. However, a challenge arises when CNNs are employed in classification applications. Specifically, they are limited to classifying only known categories, rendering them incapable of making accurate distinctions for unknown categories. In such instances, when confronted with this situation, CNNs tend to categorize the input into the known category with the highest probability.[13]

Figure 2 illustrates an access control system utilizing CNN for face recognition, wherein the three names on the right denote recognized categories within the system. If the individual to be identified belongs to one of these known categories, CNN calculates probabilities for these three individuals and accurately classifies the person on the basis of their corresponding probabilities. However, if the individual does not belong to any of the known categories, CNN still classifies them as the individual with the highest probability among the three known categories. Consequently, this situation enables individuals from unfamiliar, non-known categories to successfully bypass access control, significantly compromising system security. As the access control system necessitates the capability to identify strangers, employing CNN for facial recognition proves inadequate and unsuitable as an identification methodology.

2.   Siamese NN

Facial recognition is fundamentally a matching problem involving two images. However, the aforementioned facial recognition techniques typically handle each facial image independently, which deviates from intuitive processing. Facial recognition should ideally involve parallel processing. For instance, given two facial images, the task is to ascertain whether these two images belong to the same individual. This is achieved through a facial recognition system that extracts features from each facial image separately. The recognizer then evaluates the similarity of the two individuals' facial images to determine whether they correspond to the same person. This approach differs from the use of classifiers in previous methods. Classifiers take a facial image as input and categorize it into known categories, whereas recognizers differentiate between the similarity and dissimilarity of facial features, specifically evaluating facial resemblance and distinctiveness.
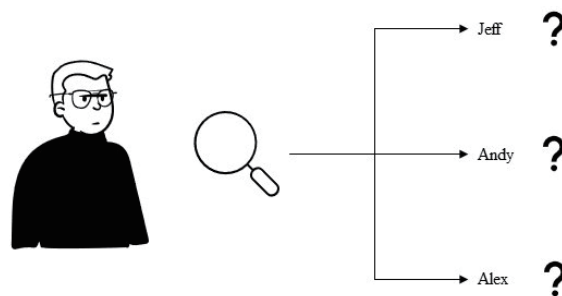

Fig. 2.   Schematic diagram of CNN processing unknown categories.

Currently, numerous NN applications incorporate metric learning algorithms.[17,18] Metric learning, also referred to as similarity learning, finds utility in training Siamese NN models through methodologies outlined in the literature.[19,20] The underlying objective is to enhance training by minimizing feature distances within the same category while expanding feature distances between distinct categories. This facilitates the computation of image similarities, thereby permitting the establishment of a threshold value to discriminate between known and unknown categories. This training approach exclusively refines the similarity function rather than the entire model, substantially curtailing computational demands compared with conventional methods and significantly enhancing processing speed. This research is aimed at establishing an access control system based on face recognition technology so as to widely meet the security needs of industrial, commercial, and a variety of settings. The inherent merits of the Siamese NN align with the aims of this investigation, thus rendering it an appropriate model framework for the intended facial recognition study.

3.   Facial Recognition System

A comprehensive facial recognition system necessitates the integration of two core components: face detection and face recognition. When an image is detected and fed into the facial recognition system, the system is tasked with precisely identifying the facial region within the image and subsequently isolating it through a preprocessing procedure. After preprocessing, the image's distinctive features are extracted, facilitating the subsequent facial recognition task. Constructing an integrated system encompasses a sequence of essential steps encompassing the acquisition of training data, facial detection, random pair labeling, image preprocessing, NN training, and model assessment and validation. The holistic process of the facial recognition model is visually outlined in Fig. 3, showing the interplay of these critical stages.

3.1   Data collection and face detection

The approach employed in collecting training data for this research involves employing OpenCV to control the camera for image capture, followed by utilizing facial detection to isolate the facial segment within the captured image. Within the domain of facial detection, the
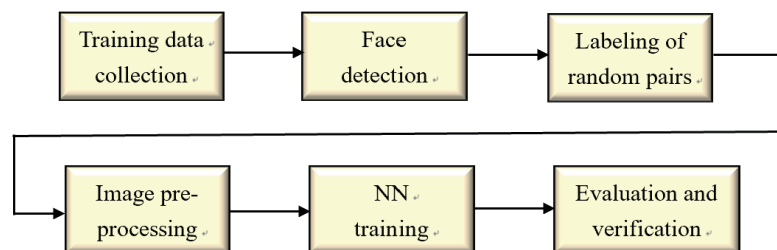

Fig. 3.   (Color online) Flowchart for constructing a facial recognition system.

RetinaFace technique is adopted.[21] This method leverages the feature pyramid network (FPN) for the extraction of multiscale features, coupled with the context module of the secure shell protocol (SSH) algorithm for the precise localization of five crucial facial landmarks, that is, the centers of the eyes, the nose, and the mouth corners. Subsequently, the facial region is meticulously extracted from the image, as visually depicted in Fig. 4.

3.2 Labeling of random pairs

To facilitate effective and expeditious differentiation within NN during its training phase, the processed facial data undergoes pairing and subsequent labeling. Specifically, identical individuals are assigned a label of 1, denoting a positive sample, whereas distinct individuals are assigned a label of 0, signifying a negative sample. This categorization is visually depicted in Fig. 5. By employing this structured labeling mechanism, the proficiently trained neural model becomes adept at accurately determining whether two provided images correspond to the same individual.

3.3 NN model and training

Both the Siamese NN and the CNN utilize convolutional and pooling layers to extract features from images. However, a pivotal distinction surfaces: CNNs struggle to recognize unfamiliar categories and require a substantial amount of training data for comprehensive learning. In contrast, Siamese networks compute the image similarity through weight
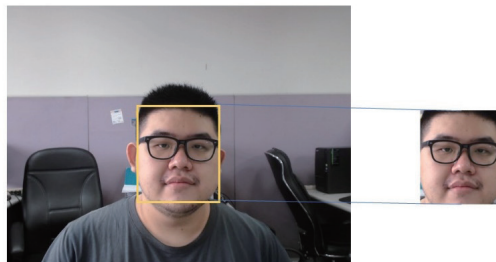

Fig. 4.   (Color online) Facial extraction utilizing RetinaFace.

Fig. 5.   (Color online) Labeling of random pairs.

calculations, as illustrated in Fig. 6. Following evaluation, the similarity values are juxtaposed with a predetermined threshold to determine whether the image belongs to a recognized or unfamiliar category. Siamese NNs do not require a large amount of training data like CNNs; achieving high accuracy is possible with minimal training data. This approach endows the model with flexibility, avoiding the need for retraining the entire model when adding a new recognizer.

The employed Siamese NN model in this investigation comprises eight convolutional layers, three pooling layers, and a solitary fully connected layer. As portrayed in Fig. 7, all eight convolutional layers within this model implement padding techniques, with a convolution kernel dimension of 3 × 3 and a stride of 1. The activation function applied across all layers is the leaky rectified linear unit (LeakyReLU). Concerning the pooling layers, max pooling is utilized three times to curtail the feature map's parameter count to achieve training acceleration and anti-noise effects. Subsequently, the sigmoid function serves as the activation function for the ultimate fully connected layer.

In the training of the NN model, the facial images acquired in the study were uniformly resized to 48 × 48 dimensions. Following the processes of pairing and labeling, all images were
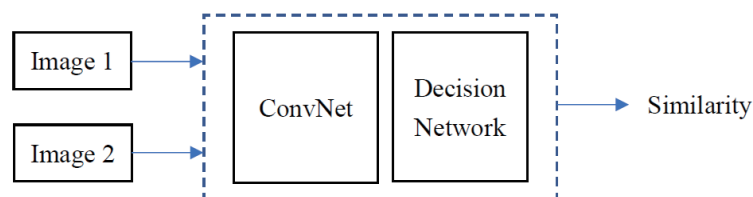

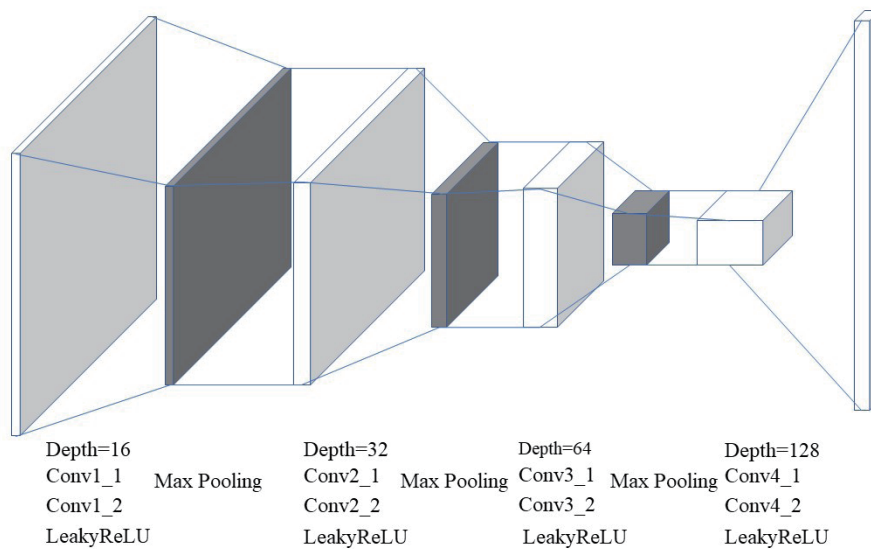Fig. 6. (Color online) Schematic diagram of a Siamese NN.

Fig. 7. (Color online) Siamese NN model.

input into the network model for feature learning. The entire model training procedure is as follows. Each set of 48 × 48 images, after undergoing convolution operations in Conv1_1 and Conv1_2 as illustrated in Fig. 7, generates feature maps with dimensions of 48 × 48 × 16. After parameter reduction through pooling layers, the feature map dimensions contract to 24 × 24 × 16. It is observable that after convolution, the number of feature maps doubles, while after the pooling layer processing, their size diminishes to half. Consequently, after the eighth convolutional layer, the feature map dimensions amount to 6 × 6 × 128. Finally, upon flattening, they are connected to the fully connected layer, facilitating the output of the image's features.

3.4 Loss function and weight update

The loss function plays a vital role during the training of neural networks and serves as a critical indicator to assess the effectiveness of a network model's learning performance. In this research, the employed loss function for training the Siamese NN in the face recognition system is contrastive loss. Its computation is given by Eq. (2), where m represents the margin value ($m > 0$) and D_w signifies the Euclidean distance calculated as Eq. (3). Here, G_w denotes the output of the Siamese NN.

$$L = (d)\frac{1}{2}(D_w)^2 + (1-d)\frac{1}{2}\{\max(0, m - D_w)\}^2 \tag{2}$$

$$D_w = \sqrt{\{\{G_w(x_1) - G_w(x_2)\}^2\}} \tag{3}$$

This loss function effectively captures the degree of match between paired data in the network. When the training data corresponds to the same individual (i.e., $d = 1$), the loss value is governed by $(D_w)^2$ in Eq. (2). For similar training data, a larger Euclidean distance corresponds to a larger loss value. On the other hand, when the training data represents different individuals (i.e., $d = 0$), the loss value is governed by $\{\max(0, m - D_w)\}^2$ in Eq. (2). In this case, for dissimilar training data, if the Euclidean distance is very small and less than the margin value m, the loss value increases. Conversely, if the Euclidean distance is sufficiently large and greater than the margin value m, the loss value is 0. The weight adjustment in this research employs the adaptive moment estimation (Adam) algorithm, a commonly used learning rule in deep learning. This method efficiently computes weight updates and adapts well to large-scale datasets. Each update is performed within a defined range of learning rates, contributing to smoother weight updates.

4. Experiment and Results

Once a complete facial recognition system is established, before the actual deployment, it is essential to assess the recognition performance of the system. In this section, we will outline the evaluation metrics employed in this study, the experimental training and testing procedures, the simulation methodology, and the results obtained from the simulations.

4.1 Evaluation index

To evaluate the effectiveness of the Siamese NN model, four parameters were employed to demonstrate the results of actual testing. These four parameters are defined as follows: true positive (*TP*), representing the count of instances where the actual result is positive and the model's prediction is also positive; true negative (*TN*), representing the count of instances where the actual result is negative and the model's prediction is also negative; false positive (*FP*), representing the count of instances where the actual result is negative while the model predicts positive; and false negative (*FN*), representing the count of instances where the actual result is positive but the model predicts negative.

For the assessment of accuracy, the following four evaluation metrics were employed to evaluate the model's performance:

Accuracy, defined as the ratio of the number of correctly predicted samples to the total number of samples, is calculated as

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \tag{4}$$

Precision, defined as the ratio of the count of instances where the actual result is positive and the predicted result is also positive to the total count of instances predicted as positive, measures the proportion of correctly predicted positives among all predicted positives. Its calculation is shown as

$$Precision = \frac{TP}{TP + FP}. \tag{5}$$

Recall, defined as the ratio of the count of instances where the actual result is positive and the predicted result is also positive to the total count of instances with an actual positive result, measures the proportion of correctly predicted positives among all actual positives. It is calculated as

$$Recall = \frac{TP}{TP + FN}. \tag{6}$$

*F*1-Score, the harmonic mean of precision and recall, is calculated as

$$F1\text{-}Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}. \tag{7}$$

Among the four evaluation metrics, accuracy assesses the overall prediction performance on the entire sample set, while precision, recall, and *F*1-Score focus on the prediction performance of each individual class. To comprehensively evaluate the prediction performance across all

classes, we adopt the macro-average approach. In this approach, the evaluation metrics (precision, recall, and *F*1-Score) for different classes are summed and then averaged. The calculation is shown in Eqs. (8)–(10), where *N* represents the total number of classes, *P* stands for precision, *R* signifies recall, and *F* denotes *F*1-Score.

$$Macro - Precision = \frac{1}{N} \sum_{i=1}^{N} P_i \tag{8}$$

$$Macro - Recall = \frac{1}{N} \sum_{i=1}^{N} R_i \tag{9}$$

$$Macro - F\text{1-Score} = \frac{1}{N} \sum_{i=1}^{N} F_i \tag{10}$$

4.2 Training and testing data

In this study, the experimental dataset used is divided into two categories. The first type of face dataset is a self-captured dataset of male and female subjects with relatively minimal noise. It consists of a total of 48 individuals, comprising 6576 face images. The captured images of each individual exhibit variations in shooting distance, facial expression, and other factors, with no fixed patterns. Additionally, for all individuals, the images captured from different angles were prepared, as illustrated in Fig. 8. The second type of face dataset is drawn from the MS-Celeb-1M dataset, which contains more noise. It includes 48 individuals, with a total of 3080 face images.

4.3 Simulations

In our study, all experimental simulations were divided into two main phases: the training phase and the testing and evaluation phase. To enhance the credibility of the results, we employed the *K*-fold cross-validation approach for comparison. A total of five rounds of experiments were conducted. The process is illustrated in Fig. 9, where the simulated dataset is divided into five subsets. In each round, one subset is used as unseen testing data for the model, while the remaining four subsets are used for training purposes.

The datasets used for experimental simulations are presented in Table 1, including 48 self-captured images and 48 images sourced from the MS-Celeb-1M dataset. For the testing phase, 10 individuals were randomly selected from the self-captured images as testing data, while the remaining 78 individuals were used as training data.

The training and testing accuracies for each round in the experiments are presented in Table 2. The average testing accuracy is 96.92%. For the training and testing data across five rounds, the statistical values of macro-recall, macro-precision, and macro-*F*1-Score are presented in Tables 3, 4, and 5, respectively. The average values for the testing data are a macro-recall of 97.13%, a macro-precision of 99.46%, and a macro-*F*1-Score of 98.20%.

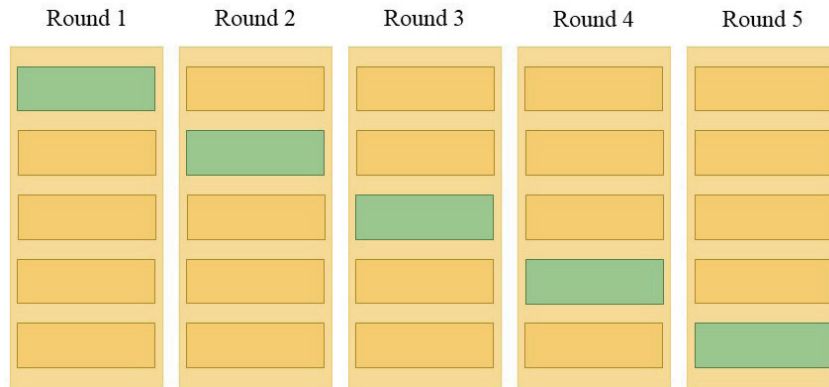Fig. 8. (Color online) Schematic diagram of face captured from different angles.

Fig. 9. (Color online) *K*-fold approach.

Table 1
Experimental dataset.

	Training data	Testing data
Self-captured images	38	10
MS-Celeb-1M images	48	0
Total images	86	10

Table 2
Accuracies of training and testing for each round.

	Accuracy	
	Training (%)	Testing (%)
Round 1	97.38	96.58
Round 2	98.52	97.09
Round 3	91.97	99.78
Round 4	97.66	97.13
Round 5	98.89	94.03
Average	96.88	96.92

Table 3
Macro-recall values of training and testing for each round.

	Macro-recall	
	Training (%)	Testing (%)
Round 1	97.66	96.85
Round 2	98.57	98.02
Round 3	93.54	99.70
Round 4	97.27	97.33
Round 5	98.88	93.74
Average	97.18	97.13

Table 4
Macro-precision values of training and testing for each round.

	Macro-precision	
	Training (%)	Testing (%)
Round 1	98.29	99.68
Round 2	98.38	98.01
Round 3	94.19	100.00
Round 4	98.44	99.90
Round 5	99.49	99.72
Average	97.76	99.46

Table 5
Macro-F1-Scores of training and testing for each round.

	Macro-F1-Score	
	Training (%)	Testing (%)
Round 1	97.83	98.19
Round 2	98.41	97.90
Round 3	92.69	99.85
Round 4	97.73	98.54
Round 5	99.16	96.50
Average	97.16	98.20

5. Conclusions

In the realm of image recognition classification problems, CNNs exhibit impressive processing capabilities. However, they are confined to situations where all images to be classified must belong to known categories. When faced with images from unknown categories, CNNs struggle to provide effective recognition. Given that the recognition system developed in this study is aimed not only at distinguishing people among known members on the basis of their facial images but also at discerning strangers, a metric-learning-based Siamese NN was chosen as the primary algorithm for the facial recognition system.

The results of the simulation experiments demonstrate that the Siamese NN, after undergoing training, accurately recognizes and classifies previously unseen test data. Additionally, when confronted with facial images of non-database strangers, it aptly identifies them as nonmembers. Since this architecture primarily learns image features, its methodology involves bringing similar image features closer together while pushing apart features of dissimilar images. This approach enables direct image input to the Siamese network to obtain features, followed by similarity calculations to determine membership within the database. Furthermore, when the need arises to add a new member, their facial image can be directly incorporated into the database. Similarly, reducing the number of members can be accomplished by removing their facial images from the database. This approach minimizes the necessity for retraining the model after personnel changes, enhancing its flexibility in management aspects.

The simulated results of the study demonstrate an average classification accuracy of 96.92%, an average recall rate of 97.13%, an average precision rate of 99.46%, and an average F1-Score of 98.20%. These outcomes indicate highly accurate performance in recognition. However, it is

important to note that our experiments were conducted in specific scenarios. The system's recognition accuracy might potentially decrease when faced with changes in scenes or lighting conditions. These aspects present a direction for future research and investigation.

Acknowledgments

References

1   Y. Zhou, S. Cao, D. Wen, H. Zhang, and L. Zhao: Proc. 2011 Int. Conf. Electronics, Communications and Control (IEEE, 2011) 2964. https://doi.org/10.1109/ICECC.2011.6066375
2   X. F. Zhang and X. M. Ren: Proc. 2011 Int. Conf. Multimedia Technology (IEEE 2011) 934. https://doi.org/10.1109/ICMT.2011.6002199
3   A. B. Xu, X. Jin, Y. G. Jiang, and P. Guo: Proc. 18th Int. Conf. Pattern Recognition (IEEE, 2006) 481. https://doi.org/10.1109/ICPR.2006.395
4   M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski: IEEE Trans. Neural Networks **13** (2002) 1450. https://doi.org/10.1109/TNN.2002.804287
5   Y. Gizatdinova and V. Surakka: Proc. 14th Int. Conf. Image Analysis and Processing (IEEE 2007) 419. https://doi.org/10.1109/ICIAP.2007.4362814
6   S. L. Happy and A. Routray: IEEE Trans. Affective Comput. **6** (2014) 1. https://doi.org/10.1109/TAFFC.2014.2386334
7   T. Terada, Y. Chen, and R. Kimura: Proc. 2018 14th Int. Conf. Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD) (IEEE, 2019) 390. https://doi.org/10.1109/FSKD.2018.8687254
8   X. Xiao and L. Li: Proc. 2008 Int. Conf. Computer Science and Software Engineering (IEEE 2008) 907. https://doi.org/10.1109/CSSE.2008.869
9   W. F. Cao, L. Li, and X. L. Lv: Proc. 2007 Int. Conf. Machine Learning and Cybernetics (IEEE, 2007) 2869. https://doi.org/10.1109/ICMLC.2007.4370637
10   S. Shu, J. Luo, B. Sun, and N. Yu: Proc. 2010 5th Int. Conf. Computer Science and Education (IEEE, 2010) 1045. https://doi.org/10.1109/ICCSE.2010.5593408
11   M. M. Sani, S. A. Samad, and K. A. Ishak: Proc. 2012 IEEE 8th Int. Coll. Signal Processing and Its Applications (IEEE, 2012) 48. https://doi.org/10.1109/CSPA.2012.6194689
12   T. Ahmed, P. Das, M. F. Ali, and M. F. Mahmud: Proc. 2020 11th Int. Conf. Computing, Communication and Networking Technologies (ICCCNT) (IEEE, 2020) 1. https://doi.org/10.1109/ICCCNT49239.2020.9225688
13   X. Jiang and W. W. Gao: Proc. 2019 3rd Int. Conf. Electronic Information Technology and Computer Engineering (EITCE) (IEEE, 2019) 2007. https://doi.org/10.1109/EITCE47263.2019.9095075
14   K. Tao, Y. He, and C. Chen: Proc. 2019 Chinese Automation Congress (CAC) (IEEE, 2019) 5403. https://doi.org/10.1109/CAC48633.2019.8996236
15   Y. Zhiqi: Proc. 2021 IEEE 5th Conf. Advanced Information Technology, Electronic and Automation Control (IAEAC) (IEEE, 2021) 2530. https://doi.org/10.1109/IAEAC50856.2021.9390856
16   Y. Zhou, H. Ni, F. Ren, and X. Kang: Proc. 2019 IEEE Int. Conf. Mechatronics and Automation (IEEE, 2019) 1091. https://doi.org/10.1109/ICMA.2019.8816192
17   G. J. Liu, H. Chen, L. Y. Wang, and D. Zhu: Proc. 2020 Int. Conf. Machine Learning and Cybernetics (ICMLC) (IEEE, 2021) 219. https://doi.org/10.1109/ICMLC51923.2020.9469547
18   S. Yue and S. Ozawa: Proc. 2012 11th Int. Conf. Machine Learning and Applications (IEEE, 2013) 671. https://doi.org/10.1109/ICMLA.2012.125
19   P. Rama Devi, R. Yashashvini, G. Navyadhara, and M. Ruchitha: Proc. 2021 2nd Int. Conf. Emerging Technology (INCET) (IEEE, 2021) 1. https://doi.org/10.1109/INCET51464.2021.9456199
20   S. Zhao, W. Li, Q. Du, and Q. Ran: Proc. IGARSS 2018 – 2018 IEEE Int. Geoscience and Remote Sensing Symp. (IEEE, 2018) 2567. https://doi.org/10.1109/IGARSS.2018.8519286
21   S. E. Sitepu, G. Jati, M. R. Alhamidi, W. Caesarendra, and W. Jatmiko: Proc. 2021 6th Int. Workshop Big Data and Information Security (IWBIS) (IEEE, 2021) 81. https://doi.org/10.1109/IWBIS53353.2021.9631848

About the Authors

Chih-Yung Chen received his B.S and M.S. degrees in computer science and information engineering from Tamkang University in 1998 and 2000, respectively, and his Ph.D. degree in electrical engineering from I-Shou University in 2008. Currently, he is an associate professor at the Program of Artificial Intelligence and Mechatronics, National Pingtung University of Science and Technology, Taiwan. His research interests are in the areas of artificial intelligence, autonomous robot control, database systems, and software engineering. (mikechen@mail.npust.edu.tw)

Huang-Chu Huang received his Ph.D. degree in electrical engineering from National Sun Yat-Sen University, Taiwan, in 2001. Currently, he is a professor in the Electronic Communication Department, National Kaohsiung University of Science and Technology, Kaohsiung City, Taiwan. His research interests are in the areas of control, power signal prediction, and neural network applications. (h4530@nkust.edu.tw)

Jyun-Cheng Jheng received his B.S. degree and M.S. degrees from the Electrical Engineering Department of I-Shou University, Taiwan, in 2020 and 2022, respectively. His research interests are in artificial intelligence and signal processing. (wl102547@gmail.com)

Rey-Chue Hwang received his Ph.D. degree in electrical engineering from Southern Methodist University, Dallas, TX, in 1993. Currently, he is a full professor in the Electrical Engineering Department, I-Shou University, Taiwan, R.O.C. Dr. Hwang has published more than 300 papers in various journals and conferences in the areas of artificial intelligence systems, signal processing, and fuzzy control. He is now a fellow of IET and a senior member of IEEE. He chartered the IEEE CIS Chapter, Tainan Section, and served as the co-chair and chair from 2004 to 2009. (rchwang@isu.edu.tw)