Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

SRTTU JECEI

Research paper

A Siamese Network-Based Xception for Face Recognition

Ali Habibi*, *Mahlagha Afrasiabi*[*], *Moniba Chaparian

Department of Computer Engineering, Faculty of Electrical and Computer Engineering, Hamedan University Of Technology, Hamedan, Iran.

Article Info

Article History:
Received 23 February 2025
Reviewed 13 April 2025
Revised 10 May 2025
Accepted 06 June 2025

Keywords:
Face recognition
Convolutional neural networks
Siamese network
Transfer learning
Small-sample dataset

[*]Corresponding Author's Email Address: *m.afrasiabi@hut.ac.ir*

Abstract

Background and Objectives: Facial recognition technology has become a reliable solution for access control, augmenting traditional biometric methods. It primarily focuses on two core tasks: face verification, which determines whether two images belong to the same individual, and face identification, which matches a face to a database. However, facial recognition still faces critical challenges, such as variations in pose, illumination, facial expressions, image noise, and limited training samples per subject.

Method: This study employs a Siamese network based on the Xception architecture within a transfer learning framework to perform one-shot face verification. The model is trained to compare image pairs rather than classify them individually, using deep feature extraction and Euclidean distance measurement, optimized through a contrastive loss function.

Results: The proposed model achieves high verification accuracy on benchmark datasets, reaching 97.6% on the Labeled Faces in the Wild (LFW) dataset and 96.25% on the Olivetti Research Laboratory (ORL) dataset. These results demonstrate the model's robustness and generalizability across datasets with diverse facial characteristics and limited training data.

Conclusion: Our findings indicate that the Siamese-Xception architecture is a robust and effective approach for facial verification, particularly in low-data scenarios. This method offers a practical, scalable solution for real-world facial recognition systems, maintaining high accuracy despite data constraints.

How to cite this paper:

A. Habibi, M. Afrasiabi, M. Chaparian, "A siamese network-based Xception for face recognition," J. Electr. Comput. Eng. Innovations, 14(1): 33-40, 2026.

DOI: 10.22061/jecei.2025.11780.834

URL: https://jecei.sru.ac.ir/article_2347.html

Introduction

Facial recognition technology has become an indispensable tool for identity verification across various sectors, including security operations, financial transactions, public safety, and daily interactions. This technology fundamentally divides into two main tasks: Face Verification and Face Identification [1]

A probe image is compared with a single gallery image to confirm whether they depict the same individual. In contrast, Face Identification requires comparing a probe image against multiple gallery images to determine the specific identity of the individual [2]. These tasks rely on analyzing and matching facial features, a process complicated by numerous factors. Variability in lighting conditions, facial expressions, head poses, makeup, hairstyles, and aging significantly impact facial recognition accuracy [3]. Additionally, the challenge is exacerbated when individuals share similar features, such as twins, family members, or even strangers with notable resemblances [4].

Face recognition systems face a wide range of challenges, especially in uncontrolled environments. Variations in illumination, pose, facial expressions, and occlusions (e.g., masks, glasses) can distort facial features and hinder recognition accuracy. Aging, low-resolution images, and plastic surgery further complicate the extraction and matching of features. In real-world scenarios, like surveillance, images are often captured under poor conditions—blur, motion, inconsistent lighting—that amplify these issues [5]. Moreover, intra-class similarity (e.g., twins) and inter-class variability increase classification difficulty. Addressing these challenges requires robust models capable of generalizing across diverse inputs. Recent research trends focus on attention mechanisms, 3D modeling, and domain adaptation to enhance model resilience and real-world performance.

Recent advancements in deep learning and computational power have led to the widespread adoption of Convolutional Neural Networks (CNNs) in facial recognition and other computer vision tasks [6]. CNNs have revolutionized the field by enabling more accurate image classification [7], object detection [8], and image retrieval [9] through hierarchical feature learning from extensive datasets. The effectiveness of CNNs, however, is highly dependent on the availability of diverse and representative training samples [10]. In practical scenarios, limited samples per class can hinder the efficiency and accuracy of facial recognition systems [11].

In the past few years, research has increasingly focused on enhancing model generalization and robustness under challenging conditions. Advanced techniques such as Siamese networks, triplet loss architectures, attention mechanisms, and domain adaptation methods have been introduced to overcome issues related to limited data and variations in real-world environments [11].

To address these challenges, our study introduces a novel approach that utilizes a network design featuring two identical CNNs and incorporates transfer learning techniques [12]. This approach processes pairs of images to extract and compare their features, determining whether they represent the same individual. The integration of transfer learning enhances the model's performance, particularly when trained on smaller datasets, thereby improving recognition accuracy [13].

The structure of this paper is as follows: Section 2 provides a comprehensive review of related research in the field, while Section 3 outlines the proposed methodology and network design. Section 4 presents the results and evaluation of our approach, and Section 5 concludes with a summary of findings and recommendations for future research.

Related Work

In recent years, Convolutional Neural Network (CNN)-based algorithms have achieved remarkable advancements in the field of facial recognition and verification [14]. These developments have significantly enhanced the accuracy and robustness of facial recognition systems. For instance, researchers in [15] utilized the Weighted PCA-EFMNet deep learning framework to address challenges such as variations in facial expressions, head poses, lighting conditions, and occlusions. This approach focused on feature extraction to improve the system’s ability to handle these common issues.

Further innovation is evident in the study presented in [16], where a novel verification technique employing an Auto Encoder with Class Sparsity Supervised Encoding (CSSE) was introduced. This method facilitates the acquisition of feature representations from labeled training datasets by enforcing class sparsity, thereby improving the system's ability to discern between different individuals even with limited data.

Another significant contribution to the field is detailed in [17], which proposed a part-based learning strategy for facial verification. This approach utilizes a Convolutional Fusion Network (CFN) to extract feature representations from facial images by focusing on different parts of the face. This method enhances the system’s ability to recognize faces by analyzing and integrating features from various facial regions.

In addition, [18] introduced a new technique that employs Layer Blocks (DLB) to improve the alignment of face image pairs. This technique accelerates the face verification process by enhancing the accuracy of face pair alignment, which is crucial for effective verification.

In [19], significant strides were made by developing a CNN integrated into a Siamese network architecture. Their model achieved a notable 94.8% recognition success rate by training on the limited LFW dataset. This approach demonstrated the effectiveness of using Siamese networks for face recognition, particularly with constrained data.

Building on these advancements, another study employed a similar network structure but enhanced it with transfer learning using the VGG-16 model, which was pre-trained on the ImageNet dataset. This approach achieved an impressive 95.2% accuracy [20], illustrating the benefits of leveraging pre-trained models to boost performance.

Recent research using the LFW dataset has focused on utilizing the face mesh technique to detect and recognize images with a Deep Neural Net model and 3D display, achieving 94.23% accuracy with this approach [21].

In [22], a Siamese CNN model is employed for one-shot facial recognition. By comparing image pairs and applying data augmentation, the method achieves 96% accuracy on the LFW dataset, outperforming traditional face recognition algorithms.

In the survey [23], facial recognition performance is compared across different methods using the LFW and ORL datasets. The study highlights that ORL is more suitable for controlled testing, while LFW better represents real-world scenarios, though it poses greater challenges for recognition systems.

In [5], the authors review the expanding applications of face recognition in sectors such as healthcare, security, and education, emphasizing the importance of image preprocessing and deep learning techniques. They evaluate various CNN models on the FER and LFW datasets, concluding that the LFW dataset yields better performance due to its ability to handle challenges posed by uncontrolled environments.

Expanding prior research, our work introduces a Siamese network architecture based on the Xception model, integrating depth-wise separable convolutions to enhance feature extraction and improve verification accuracy. Experimental results on the ORL and LFW datasets confirm the superior performance and robustness of the proposed approach under diverse conditions.

Proposed Approach

A Siamese network [24] is a specialized neural network architecture that consists of two parallel subnetworks. These subnetworks share the same weights and parameters, allowing them to process inputs consistently. Each subnetwork receives an input, such as an image, and its outputs are combined to produce predictions. The core concept of the Siamese network is to learn a representation of the data that facilitates comparison between inputs. By comparing the feature representations extracted from the parallel networks, the Siamese architecture can effectively measure the similarity or dissimilarity between inputs [13].

In contrast to traditional CNN classifiers that predict class labels, Siamese networks are designed to learn a similarity function, making them particularly well-suited for tasks such as one-shot learning and identity verification. Furthermore, Siamese networks require fewer examples to train compared to standard CNN classifiers, enhancing their efficiency. By focusing on learning feature distances rather than absolute class labels, they are more robust to variations in pose, lighting, and occlusions, providing superior performance in comparison to conventional classification-based methods [25].

Fig. 1 illustrates the architecture of a Siamese network, highlighting its parallel structure and shared parameters [26].

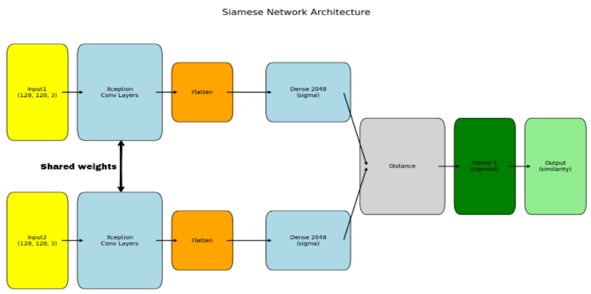

Fig. 1: A diagram for Siamese network architecture, where input1 and input2 are a pair of input images, after using a convolutional neural network to vectorize the inputs, calculating the distance between the output results, and determining the similarity of images.

A. Transfer Learning in Deep Learning

Transfer learning [27] is a pivotal technique in deep learning, leveraging pre-existing models to address new but related challenges. This method involves applying models that have been previously trained on large and diverse datasets to solve different problems with limited data. By utilizing these pre-trained models, transfer learning can accelerate the development process and enhance performance, particularly when dealing with tasks similar to those for which the original model was trained. This approach capitalizes on the knowledge and features learned from prior models, allowing for more effective solutions in new contexts.

B. Pre-Trained Xception Model

The Xception model [28] developed by François Chollet, is a state-of-the-art convolutional neural network known for its exceptional performance. It achieves a top-five test accuracy of 94.5% on the ImageNet dataset, which comprises 1000 classes and over 1.4 million labeled images. The architecture of Xception is notable for its depth and complexity, featuring 36 convolutional layers with 3×3 kernel sizes, 14 pooling layers with 2×2 kernel sizes, and three fully connected layers. The fully connected layers consist of 2048 neurons in the first two layers and a single neuron in the final layer. This architecture enables Xception to capture intricate patterns and features from images, making it highly suitable for transfer learning applications in facial recognition tasks [29].

The strengths of Xception go beyond its architecture. Compared to other CNN models, Xception enhances efficiency and reduces computational costs while maintaining high accuracy by replacing standard convolutions with depth-wise separable convolutions. This architectural modification not only optimizes performance but also significantly reduces the number of parameters, making the model less prone to overfitting [30].

C. Face Recognition Using Siamese-Xception

In our approach, we utilize a pre-trained Xception model for feature extraction within a Siamese network architecture and subsequently fine-tune it. Fig. 2 illustrates the Flowchart of the proposed facial recognition system.

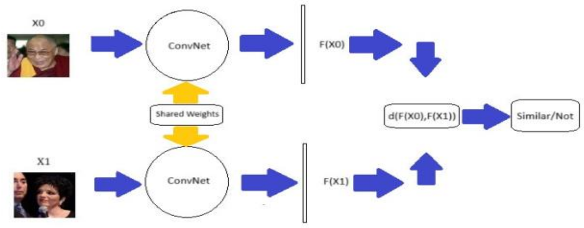

Fig. 2: Flowchart of the proposed facial recognition model based on a Siamese network with an Xception backbone.

Initially, we freeze all convolutional layers of the Xception model except for the last seven layers in the Exit Flow block. These layers consist of three separable convolutional layers, one pooling layer, and one global average pooling layer. Freezing these layers ensures that their weights remain unchanged during training, allowing only the weights in the Exit Flow block to be updated. To this setup, we add a flattened layer followed by two fully connected layers—one with 2048 neurons and the other with a single neuron—both utilizing the sigmoid activation function [31] (refer to Fig. 3 for visualization).

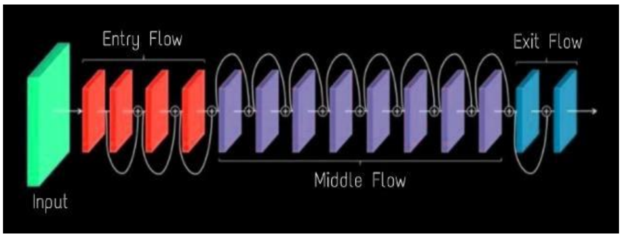

Fig. 3: Xception network architecture.

As the fully connected layers are initialized randomly, there is a risk that adjusting their weights could disrupt the learned features from the pre-trained layers. To mitigate this, we set the network input size to 128 × 128 × 3 pixels. During training, we employ the contrastive loss function [32] to refine the model's ability to differentiate between similar and dissimilar image pairs. The contrastive loss function works by minimizing the squared Euclidean distance for similar pairs of images and maximizing this distance for dissimilar pairs, effectively pulling similar samples closer and pushing dissimilar samples apart. The contrastive loss is computed using the formula:

$$L = (1 - Y)\frac{1}{2}D^2 + (Y)\frac{1}{2}\{max(0 \cdot m - D)\}^2 \qquad (1)$$

where (y) is a binary label indicating whether the image pairs are similar (y = 0) or dissimilar (y = 1). The margin (m) is a positive threshold, set to 1 in this research, beyond which dissimilar pairs do not impact the loss function. (D) represents the Euclidean distance between the feature vectors of the image pairs generated by the Siamese network, calculated as:

$$D = \|f(x_0) - f(x_1)\|^2 \qquad (2)$$

here, $f(x_0)$ and $f(x_1)$ are the feature vectors of the images x_0 and x_1, respectively, each with 512 dimensions. If the Euclidean distance (D) between the output vectors is less than 0.5, the model determines the image pair to be similar (D = 0). Conversely, if the distance is 0.5 or greater, the images are considered dissimilar (D = 1). Given the binary nature of the labels, the model's predicted values should match these labels, ensuring a direct comparison between predicted and actual values. This process ultimately yields the final prediction by the model.

$$D: \{0 => d < 0.5 \quad 1 => d \geq 0.5\} \qquad (3)$$

The pseudocode of the proposed method is illustrated in Fig. 4:

```
Siamese-Xception Based Face Recognition
---------------------------------------------------------------
Input: Dataset of image pairs {(Image_A, Image_B, label)}, where label ∈ {0, 1}
Output: Trained model and Similarity Score (S) during inference

1. Load the pre-trained Xception model (trained on ImageNet)
2. Freeze all layers except the last 7 layers in the Exit Flow block
3. Append a Flatten layer, a Dense(2048) layer, and a Dense(1, sigmoid) layer to form the Siamese head
4. For each training epoch:
   For each image pair (Image_A, Image_B, label) in the training dataset:
     a. Resize each image to 128 × 128 × 3
     b. Normalize pixel values to range [0, 1]
     c. Pass Image_A and Image_B through the shared Xception model to get feature vectors F_A and F_B
     d. Compute Euclidean distance D between F_A and F_B:
          D = sqrt(Σ (F_A[i] - F_B[i])²), for i = 1 to 512
     e. Compute contrastive loss:
          If label = 0 (similar): Loss = D²
          If label = 1 (dissimilar): Loss = max(0, margin − D)²
     f. Backpropagate the loss to update the trainable layers
5. After training, for each test image pair:
   a. Compute D and output similarity score S = sigmoid(D)
   b. If S < threshold (e.g., 0.5), classify as "similar", else "dissimilar"
```

Fig. 4: Pseudocode of the proposed Siamese-Xception-based facial recognition approach.

By only a few pictures. Notably, all images have dimensions of 250*250 pixels.

Results and Discussion

In our facial recognition system, the Siamese network leverages pairs of images along with corresponding labels for the learning process. This approach requires careful curation of training data to meet the system's specifications. To generate these image pairs, we follow a structured methodology as outlined in prior studies. When two images represent the same individual, they are categorized as "similar." Conversely, if the images depict different individuals, they are labeled as "different".

To evaluate the performance of the proposed method, two well-known datasets were utilized: Labeled Faces in the Wild (LFW) and the ORL (Olivetti Research Laboratory Face Database).

The Labeled Faces in the Wild (LFW) [33] dataset contains 15,000 images of faces collected from various individuals under unconstrained conditions. The images were captured in real-world environments, featuring variations in lighting, pose, and facial expressions, making the dataset highly challenging and suitable for robust facial analysis (Refer to Fig. 5 for illustration).

Fig. 5: Some samples of the LFW dataset.

The Olivetti Research Laboratory (ORL) [34] dataset consists of 400 grayscale images of 40 individuals, with 10 images per person. The images exhibit variations in facial expressions (such as open/closed eyes, smiling/not smiling) and pose (slight tilts and rotations), while maintaining a relatively controlled environment in terms of lighting and background. This dataset is widely used for benchmarking face recognition and classification algorithms.

For each dataset, a randomized split was performed to ensure rigorous training and evaluation, with 80% of the data assigned to the training set and the remaining 20% reserved for testing and accuracy evaluation. This balanced partitioning facilitates the model's ability to effectively discriminate between similar and dissimilar faces, a capability essential for enhancing recognition performance. Representative examples of the image pairs generated for a single subject are illustrated in Fig. 6.

The experiments were conducted on Google Colab, using an Nvidia A100 GPU with 16 GB of memory. The deep learning models were implemented using the Keras framework. The study was performed on the LFW and ORL datasets, where a batch size of 256 was employed, and the ADAM optimizer with a learning rate of α = 0.01.was used.

Fig. 6: Samples of similar and dissimilar pairs.

The accuracy on the LFW dataset was evaluated and presented in Table 1, while the results on the ORL dataset were reported in Table 2.

Table 1: Face Recognition Rate in Different Methods with LFW Dataset

Method	Accuracy
RF+LDA [23]	67.96
KNN+LDA [23]	69.77
NB+ICA [23]	72.35
LR+ICA [23]	78.55
MLP+PCA [23]	80.88
SVM+PCA [23]	83.54
CNN+STN [5]	86.3
CFN+APEM [18]	87.50±1.57
DLB [19]	88.50
PSI-CNN [5]	88.7
L-CSSE+KSRC [16]	92.02
Face Mesh [21]	94.23
SiameseFace [17]	94.80
Weighted PCA-Efmnet [27]	95.00±0.71
Siamese-VGG [17]	95.62±0.42
Siamese+CNN [22]	96
Siamese-Xception(proposed method)	97.66±0.21

A comparative analysis against various recent methods indicates that our proposed Siamese network-based Xception model consistently achieves the highest accuracy, highlighting its robustness and effectiveness for face recognition tasks under both controlled and unconstrained conditions.

Table 2: Face Recognition Rate in Different Methods with ORL Dataset

Method	Accuracy
NB+ICA [23]	86.25
KNN+ICA [23]	87.50
RF+PCA [23]	93.75
MLP+LDA [23]	93.75
SVM+PCA [23]	93.75
LR+ICA [23]	93.75
Siamese-Xception(proposed method)	96.25

The comparison between the LFW and ORL datasets reveals significant differences in the performance of facial recognition methods, particularly when evaluating the Siamese network-based Xception model. For the ORL dataset, the proposed Siamese-Xception model achieved an impressive accuracy of 96.25%, outperforming traditional machine learning models and CNN-based methods. On the other hand, when tested on the more challenging LFW dataset, the same model achieved a higher accuracy of 97.66%. This indicates that the Siamese-Xception model performs exceptionally well in both controlled and real-world conditions, demonstrating its robustness and generalization ability. In comparison to other recent methods, the proposed model consistently outperforms existing techniques, including Siamese-VGG, which reached 95.62% on LFW. These results highlight the effectiveness of the Siamese network-based Xception architecture, particularly in handling variations in facial features across different datasets. The model's success on both datasets underscores its potential for real-world applications in identity verification and face recognition tasks, with notable improvements in accuracy compared to previous approaches.

An analysis of the model's performance, presented in Fig. 7 and Fig. 8, highlights the training and validation dynamics for the LFW and ORL datasets, respectively. Fig. 7 depicts the progression of accuracy and loss across epochs for the LFW dataset, while Fig. 8 illustrates the corresponding trends for the ORL dataset.

In both cases, the accuracy shows a consistent upward trajectory, indicating the model's capability to effectively learn from the training data while maintaining strong generalization to unseen samples. Although a slight gap between training and validation accuracy appears midway through training, this discrepancy remains modest, and it gradually stabilizes during later epochs, suggesting that the model successfully mitigates severe overfitting. Minor fluctuations observed in the validation loss and accuracy are negligible and do not disrupt the overall convergence trend.

These findings collectively demonstrate that the proposed Siamese-Xception model achieves a well-balanced performance between underfittings and overfitting [35], delivering robust and reliable results across both controlled (ORL) and unconstrained (LFW) face recognition environments.

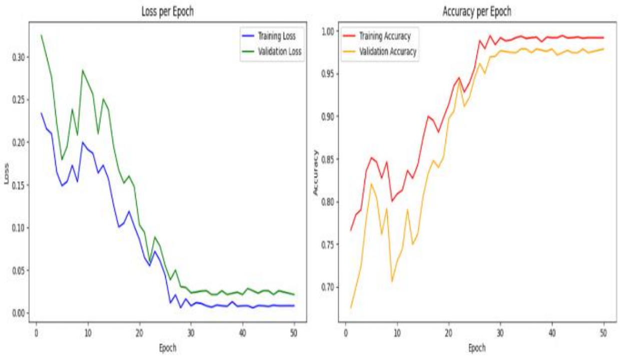

Fig. 7: Performance of the proposed model based on accuracy and loss per epoch on the LFW dataset.

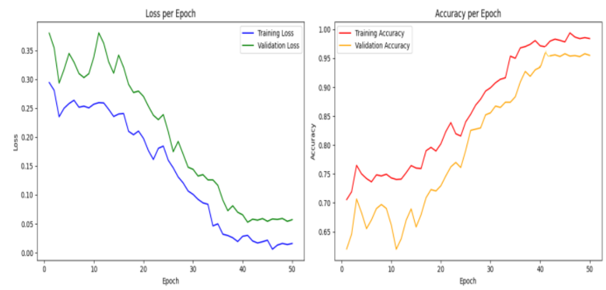

Fig. 8: Performance of the proposed model based on accuracy and loss per epoch on the ORL dataset.

Conclusion and Future Works

In this study, a Siamese network architecture based on transfer learning with Xception was employed to address the challenges of face verification in limited-data scenarios. The model demonstrated strong capability in extracting deep facial features, achieving competitive accuracy rates of 97.6% on the LFW dataset and 96.25% on the ORL dataset, which confirms its robustness across datasets with varying characteristics.

Despite these promising results, certain limitations remain. The model's performance is inherently dependent on the quality and variability of the training data, and further improvements could be realized by expanding the dataset and incorporating additional evaluation metrics beyond accuracy.

For future work, integrating advanced techniques such as triplet loss and more sophisticated data augmentation strategies is recommended to enhance feature learning and model generalization further. Exploring architectures that better capture both low-level and high-level facial features could also contribute to higher verification precision. These directions hold potential for developing even more reliable and efficient face recognition systems for practical applications.

Author Contributions

The first author was responsible for developing and implementing the core algorithms and code. The second author reviewed the code and revised the manuscript for technical accuracy and clarity. The third author contributed to writing the manuscript and organizing its structure. All authors reviewed and approved the final version of the manuscript.

Acknowledgment

We would like to express our gratitude to the editors and reviewers of JECEI for their valuable feedback and support during the review process. This work was entirely self-supported, with no financial support from any external agency.

Funding

This research received no external funding.

Conflict of Interest

The authors declare that there is no potential conflict of interest regarding the publication of this work. Additionally, the authors confirm that all ethical issues, including plagiarism, informed consent, misconduct, data fabrication and falsification, double publication or submission, and redundancy, have been fully addressed.

References

[1] E. Solomon, A. Woubie, E. S. Emiru, "Deep learning based face recognition method using siamese network," arXiv preprint arXiv:2312.14001, 2023.

[2] D. Bansal, B. Gupta, S. Gupta, A. Anand, Sumit, A. Sagar, "Facial recognition advancements with siamese networks: A comprehensive survey," in Proc. International Conference on Computation of Artificial Intelligence & Machine Learning: 29-41, 2024.

[3] J. J. Valero-Mas, A. J. Gallego, J. R. Rico-Juan, "An overview of ensemble and feature learning in few-shot image classification using siamese networks," Multimedia Tools Appl., 83: 19929-19952, 2024.

[4] G. Guo, N. Zhang, "A survey on deep learning based face recognition," Comput. vision image understanding, 189: 102805, 2019.

[5] R. S. C. Kavita, U. Chhillar, "Analysing face recognition system based on convolution neural networks for FER and LFW datasets," in Proc. National Conference on Computational Intelligence and Data Science, 2023.

[6] B. Amirgaliyev, M. Mussabek, T. Rakhimzhanova, A. Zhumadillayeva, "A review of machine learning and deep learning methods for person detection, tracking and Identification, and face recognition with applications," Sensors, 25: 1410, 2025.

[7] W. Rawat, Z. Wang, "Deep convolutional neural networks for image classification: A comprehensive review," Neural Comput., 29: 2352-2449, 2017.

[8] J. Han, D. Zhang, G. Cheng, N. Liu, D. Xu, "Advanced deep-learning techniques for salient and category-specific object detection: a survey," IEEE Signal Process. Mag., 35: 84-100, 2018.

[9] L. Zheng, Y. Yang, Q. Tian, "SIFT meets CNN: A decade survey of instance retrieval," IEEE Trans. Pattern Anal. Mach. Intell., 40: 1224-1244, 2017.

[10] A. Mumuni, F. Mumuni, "Automated data processing and feature engineering for deep learning and big data applications: A survey," J.Inf. Intell., 3(2): 113-153, 2025.

[11] Y. Li, C. P. Chen, T. Zhang, "A survey on siamese network: Methodologies, applications, and opportunities," IEEE Trans. Artif. Intell., 3: 994-1014, 2022.

[12] J. Maggu, E. Chouzenoux, G. Chierchia, A. Majumdar, "Convolutional transform learning," in Proc. Neural Information Processing: 25th International Conference, ICONIP 2018, Siem Reap, Part III 25: 162-174, 2018.

[13] A. Nandy, S. Haldar, S. Banerjee, S. Mitra, "A survey on applications of siamese neural networks in computer vision," in Proc. 2020 International Conference for Emerging Technology (INCET): 1-5, 2020.

[14] M. Wang, W. Deng, "Deep face recognition: A survey," Neurocomputing, 429: 215-244, 2021.

[15] B. Ameur, M. Belahcene, S. Masmoudi, A. B. Hamida, "Weighted PCA-EFMNet: A deep learning network for face verification in the wild," in Proc. 2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP): 1-6, 2018.

[16] A. Majumdar, R. Singh, M. Vatsa, "Face verification via class sparsity based supervised encoding," IEEE Trans. Pattern Anal. Mach. Inte., 39: 1273-1280, 2016.

[17] C. Xiong, L. Liu, X. Zhao, S. Yan, T. K. Kim, "Convolutional fusion network for face verification in the wild," IEEE Trans. Circuits Syst. Video Technol., 26: 517-528, 2015.

[18] S. C. Chong, A. B. J. Teoh, T. S. Ong, "Unconstrained face verification with a dual-layer block-based metric learning," Multimedia Tools Appl., 76: 1703-1719, 2017.

[19] J. Zhang, X. Jin, Y. Liu, A. K. Sangaiah, J. Wang, "Small sample face recognition algorithm based on novel Siamese network," J. Inf. Process. Syst., 14: 1464-1479, 2018.

[20] M. Heidari, K. Fouladi-Ghaleh, "Using siamese networks with transfer learning for face recognition on small-samples datasets," in Proc. 2020 International Conference on Machine Vision and Image Processing (MVIP): 1-4, 2020.

[21] S. Hangaragi, T. Singh, "Face detection and recognition using face mesh and deep neural network," Procedia Comput. Sci., 218: 741-749, 2023.

[22] B. Desai, S. K. Satapathy, "Facial recognition using siamese neural network and data augmentation techniques," in Proc. 2024 2nd World Conference on Communication & Computing (WCONF): 1-6, 2024.

[23] M. Mosadag, M. A. Mohammed, K. Bashir, A. Mubark, "A comparative analysis for the performance of LFW and ORL databases in facial recognition," Conclusions Eng., 1(1): 58-63, 2025.

[24] D. Chicco, "Siamese neural networks: An overview," Artificial Neural Networks: 73-94, 2021.

[25] L. Lu, J. S. Kang, F. Meng, M. Yu, "Non-intrusive load identification based on retrainable siamese network," Sensors, 24: 2562, 2024.

[26] S. Javed, M. Danelljan, F. S. Khan, M. H. Khan, M. Felsberg, J. Matas, "Visual object tracking with discriminative filters and siamese networks: a survey and outlook," IEEE Trans. Pattern Anal. Mach. Intell., 45: 6552-6574, 2022.

[27] M. Iman, H. R. Arabnia, K. Rasheed, "A review of deep transfer learning and recent advancements," Technologies, 11: 40, 2023.

[28] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in Proc. IEEE Conference on Computer Vision and Pattern Recognition: 1251-1258, 2017.

[29] S. Li, H. Qu, X. Dong, B. Dang, H. Zang, Y. Gong, "Leveraging deep learning and xception architecture for high-accuracy mri classification in alzheimer diagnosis," arXiv preprint arXiv:2403.16212, 2024.

[30] P. Joshi, V. Nivethitha, "Deep fake image detection using xception architecture," in Proc. 2024 5th International Conference on Recent Trends in Computer Science and Technology (ICRTCST): 533-537, 2024.

[31] G. Kaur, N. Sharma, D. Upadhyay, M. Singh, R. Gupta, "Brain tumour disease detection in MRI images using xception model," in Proc. 2024 3rd International Conference for Innovation in Technology (INOCON): 1-5, 2024.

[32] F. Wang, H. Liu, "Understanding the behaviour of contrastive loss," in Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition: 2495-2504, 2021.

[33] G. B. Huang, M. Mattar, T. Berg, E. Learned-Miller, "Labeled faces in the wild: A database forstudying face recognition in unconstrained environments," in Workshop on Faces in'Real-Life'Images: Detection, Alignment, and Recognition, 2008.

[34] F. S. Samaria, A. C. Harter, "Parameterisation of a stochastic model for human face identification," in Proc. 1994 IEEE Workshop on Applications of Computer Vision: 138-142, 1994.

[35] S. Hashemifar, A. Marefat, J. H. Joloudari, H. Hassanpour, "Enhancing face recognition with latent space data augmentation and facial posture reconstruction," Expert Syst. Appl., 238: 122266, 2024.

Biographies

Ali Habibi was born in Hamedan, Iran, in 2000. He received his B.Sc. degree in Computer Engineering from the University of Hamedan in 2023. He is passionate about data and works as a freelance data scientist and researcher. His research interests include computer vision, data engineering, large language models, and machine learning.

- Email: alihabibi.education@gmail.com
- ORCID: 0009-0006-9742-3967
- Web of Science Researcher ID: NA
- Scopus Author ID:NA
- Homepage: NA

Mahlagha Afrasiabi was born in 1985 in Hamedan, Iran, where her early education was completed before higher academic pursuits were undertaken. A Bachelor's degree in Computer Engineering was obtained from Bu-Ali Sina University, Hamedan, followed by a Master's degree and subsequently a Ph.D. in the same field, during which extensive research was conducted. In 2020, a faculty position was offered to her at Hamedan University of Technology, where she has been employed as an Assistant Professor. Over the years, her research activities have been primarily focused on image processing, signal processing, machine learning, and deep learning, through which significant contributions have been made. Numerous research papers have been authored and published in well-regarded journals and conferences, and active collaborations with researchers in related fields have been established to further scientific advancements.

- Email: m.afrasiabi@hut.ac.ir
- ORCID: 0000-0001-9472-4453
- Web of Science Researcher ID: AAA-7717-2021
- Scopus Author ID: NA
- Homepage: https://ce.hut.ac.ir/~m.afrasiabi

Moniba Chaparian was born in Tehran, Iran, in 2001. She received her B.Sc. degree in Computer Engineering from Hamedan University of Technology in 2023. She is passionate about Human-Computer Interaction and currently works as a freelance software engineer and researcher. Her research interests include computer vision, machine learning, HCI, and user experience design.

- Email: Moniba.chaparian@gmail.com
- ORCID: : 0009-0005-3245-6023
- Web of Science Researcher ID: NA
- Scopus Author ID:NA
- Homepage: NA