Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 10, NCREHAASTE - 2026

A Face Recognition Technique Based on Multi-Scale Facial Component Analysis with Convolutional Neural Networks

Vimal Kumar * and Jeetu Rani
Assistant Professor,
Department of Computer Science & Engineering,
School of Computer Science & Applications, IFTM University, Moradabad, 244102, India

***Corresponding Author Email:** vimalchauhan.it@gmail.com

Abstract*—Face recognition remains a challenging problem for real-time applications because of the unpredictable nature of parameters such as illumination, pose, expression, scale, low resolution, and partial occlusion of the face. Among these, low-resolution face recognition and partial face recognition (occlusion) remain major open challenges that directly affect the performance of access-control, authentication, and surveillance systems. This paper proposes a hybrid face recognition technique that combines holistic and structural (component-based) information in the feature extraction stage, uses Principal Component Analysis (PCA) or Linear Discriminant Analysis (LDA) for dimensionality reduction, and employs a Convolutional Neural Network (CNN) for classification. Facial components—eyes, nose, mouth, and the remaining facial region—are extracted and down-sampled with different resolution ratios according to their discriminative significance, then combined into a single multi-scale feature vector. Two recognition schemes are evaluated: in the first, the multi-scale feature vector is fed directly into a CNN classifier; in the second, the feature vector is projected onto a PCA or LDA feature (Eigen/Fisher) space to obtain a weight vector, which is then classified using the CNN. The proposed methods are evaluated on the ORL and FERET face databases for full-face and partial-face recognition at multiple resolutions. Experimental results show recognition accuracies of up to 99% with SIFT+CNN and 98.29% with LBP/CANNY+CNN feature combinations, with only a 6–7% drop in accuracy when using facial components instead of the full face, and a further 1–2% improvement when PCA/LDA weight vectors are used instead of raw feature vectors as CNN input.

Index Terms*—Face recognition, Convolutional Neural Network (CNN), Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), multi-scale facial components, low-resolution face recognition, partial face recognition, HOG, SIFT, Gabor filter.*

I. NTRODUCTION

Face recognition is a technique within computer vision that aims to identify or verify an individual from an image or a video clip, and it has attracted sustained interest from pattern-recognition researchers over the past few decades. It is employed across many domains including robotics, biometric security, drowsiness and fatigue detection, and criminal detection from video surveillance. The recognition pipeline typically begins with face detection and extraction from an image, followed by feature extraction using techniques such as the Histogram of Oriented Gradients (HOG), the Scale Invariant Feature Transform (SIFT), the Gabor filter, and the Canny edge detector. The final stage compares the extracted representation against a database using classifiers such as Principal Component Analysis (PCA), k-Nearest Neighbors (KNN), and Support Vector Machines (SVM).

Deep learning, particularly Convolutional Neural Networks (CNNs), has become dominant in recent face recognition research owing to its robustness in feature extraction and its ability to perform a very large number of computations on a single image region. However, real-world performance is undermined by variation in pose, illumination, expression, low resolution, and occlusion. Two of these—low-resolution face recognition and partial (occluded) face recognition—remain particularly difficult and are the central focus of this work.

Face recognition methods are broadly grouped into holistic methods, which use the entire face image (e.g., Eigenfaces via PCA, Fisherfaces via LDA), and component-based methods, which use local facial features such as the eyes, nose, and mouth. Holistic methods perform well for frontal faces but are sensitive to pose variation because global features are highly receptive to translation and rotation. Component-based methods compensate for pose changes by allowing a flexible geometrical relationship among facial components, but reliable automatic extraction and validation of components without human assistance remains a challenge.

Human observers do not weight all facial regions equally when recognizing a face; the eyes, for instance, are believed to carry more discriminative information than the cheeks or forehead. Motivated by this observation, this paper proposes a hybrid, multi-scale face recognition technique that combines componential cues (eyes, nose, mouth) with holistic (whole-face) information, applies PCA/LDA for

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 10, NCREHAASTE - 2026

dimensionality reduction, and uses a CNN for final classification.

A. Low-Resolution Face Recognition

Low-resolution face recognition (LRFR) is critical in video surveillance, where the distance between camera and subject often yields faces far smaller than the resolutions conventionally used for recognition. Limited pixel detail, combined with uncontrolled pose and illumination, makes feature extraction and matching considerably harder. To improve recognition across resolutions, this work applies PCA and LDA dimensional-reduction techniques to hybrid features extracted from pre-processed face images, and down-samples facial components—eyes, nose, mouth, and the whole face—at different resolution ratios. Down-sampling helps LDA in particular, since the between-class and within-class scatter matrices can be estimated more reliably when the dimensionality of the face vector is reduced relative to the number of training samples, mitigating the Small Sample Size (SSS) problem.

B. Hybrid Approach

The proposed hybrid approach selects four observers on the face—the two eyes, the nose, the mouth, and the remaining facial region (assuming eye coordinates are known)—and passes the non-eye patches through a Gaussian low-pass filter to reduce noise. These patches are then combined into a single feature vector. More emphasis (higher sampling ratio, i.e., less down-sampling) is placed on the eyes, followed by the nose, the mouth, and finally the remaining region, which is down-sampled most aggressively since it is considered least discriminative. Concretely, sampling ratios of 1:1 (eyes), 1:2 (nose), 1:4 (mouth), and 1:8 (remaining face/whole face) are used. This scheme reduces the dimensionality of the original image data while allowing differential emphasis on the most discriminative facial regions.

C. Partial Face Recognition

Partial face recognition addresses scenarios where a full frontal face is unavailable—for example, due to occlusion from clothing, masks, or non-cooperative subjects. In this study, 'partial' refers to a part or combination of parts of the frontal face (eye, nose, mouth). Both proposed recognition methods (direct CNN classification and PCA/LDA-projected CNN classification) are evaluated on partial as well as full face images, and performance is compared across eyes-only, nose-and-mouth, and nose-mouth-eyes combinations.

D. Objectives of the Research

The objectives of this study are to: (1) pre-process face images through average filtering, histogram equalization, and resizing; (2) evaluate face recognition performance using the two proposed methods with partial face components (eyes, nose, mouth); (3) identify the most significant facial component among eyes, nose, mouth, and the remaining face region; and (4) evaluate recognition performance using PCA/LDA weight vectors of multi-scale facial components across different resolutions, ultimately suggesting a suitable method for recognizing faces at varying resolutions and levels of occlusion.

II. RELATED WORK

Automatic Face Recognition (AFR) is inherently difficult because only a limited number of training images per subject are typically available, and training and test images are often captured under different conditions. Holistic approaches such as Eigenfaces (PCA) [34] and Fisherfaces (LDA) [33] represent a face image as a linear combination of basis vectors in a lower-dimensional 'face space,' reducing dimensionality while retaining the most significant variance in the data. PCA is fast and simple but is not optimal for discrimination since it does not use between-class information, whereas LDA explicitly maximizes the ratio of between-class to within-class scatter, giving better class separability at the cost of requiring larger training samples and being harder to estimate reliably in high-dimensional spaces.

Component-based methods extract and combine local features such as eyes, nose, and mouth, and were shown in modular Eigenspace / Eigenfeature work to outperform whole-face Eigenfaces in lower-order feature spaces [48]. Hybrid approaches combining holistic and component information have subsequently been proposed to draw on the complementary strengths of both paradigms.

On the deep-learning side, several works report high accuracy by combining classical descriptors with CNNs: a DNN-based system for noisy/obscured faces reported 92.3% accuracy; a CNN operating on the covariance of Gabor wavelet responses improved robustness to poor lighting, blur, and low resolution; a system combining 2D-PCA and Local Binary Patterns (LBP) with a CNN achieved over 90% accuracy on large datasets; and a one-dimensional deep CNN (1D-DCNN) trained on LDA-derived one-dimensional facial features reported 100% accuracy on its evaluation set. Table I summarizes representative hybrid face recognition results reported in the literature on the ORL database; Fig. 1 plots the same comparison against the accuracy achieved by the proposed method.

More recent work continues to target the same two open problems addressed in this paper—low resolution and occlusion—using deeper architectures. Suharjito and Puspita [19] extended a Trunk-Branch Ensemble CNN (TBE-CNN) for still-to-video low-resolution face recognition, combining pre-processing with complementary trunk and branch feature representations. For occlusion, Tsai et al. [20] integrated a

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 10, NCREHAASTE - 2026

Single Shot Multi-Box Detector with a pre-trained deep CNN for partially occluded face recognition, while He et al. [21] proposed locality-aware channel-wise dropout to suppress the contribution of occluded feature channels during CNN training. Neto et al. [22] showed that CNNs trained for masked face recognition generalize reasonably well to general occluded face recognition, and Eman et al. [23] combined pretrained mask detection/segmentation with robust PCA (RPCA) and a KNN classifier to separate occluded from non-occluded facial regions before matching—an RPCA-based strategy conceptually related to the PCA/LDA weighting used in the present work. A 2025 scoping review of 266 CNN-based face recognition studies published between 2019 and 2023 found that only about 7% evaluated occluded conditions alongside clean data, underscoring that partial/occluded face recognition—the focus of Sections III-C and IV of this paper—remains comparatively under-studied relative to unoccluded recognition [18].

TABLE I
Performance Comparison of Hybrid Face Recognition Methods (ORL Database)

Method	% Recognition
Down-sampling + low-resolution FR	94.0
Scale-robust feature extraction (100×100 / 12×14)	90.0 / 80.0
Weighted modular PCA	87.0
Hybrid (structural + holistic)	90.3
Hybrid model using facial components	88.57
PCA, LDA + Neural Network	95.80
Low-resolution neural-network-based FR	94.50

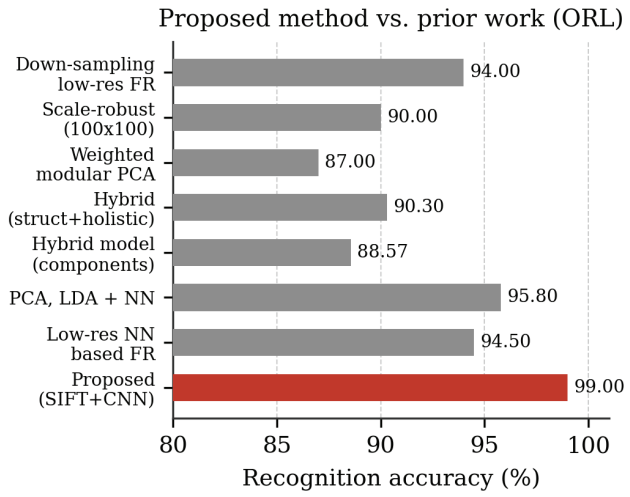


Fig. 1. Recognition accuracy of the proposed method (SIFT+CNN) compared against representative hybrid methods reported in the literature on the ORL database.

III. METHODOLOGY

The proposed system follows four key stages: (1) pre-processing of the face image, (2) multi-scale feature extraction, (3) classification using a CNN, and (4) accuracy-rate evaluation. Each stage is described below.

A. Pre-Processing

Each input face image is cropped to the face region using fixed corner points determined from the database, so that all training and test images are cropped consistently. Illumination normalization adjusts high and low illumination levels and normalizes the image to zero mean and unit standard deviation. Images are then resized using bi-cubic interpolation, which produces smoother output than other interpolation methods and reduces the computational complexity of the subsequent PCA/LDA and CNN stages. Finally, a 5×5 mean (averaging) low-pass filter is applied to reduce noise introduced by down-sampling.

B. Multi-Scale Feature Extraction

Four facial components—left eye, right eye, nose, and mouth—are cropped from the pre-processed, normalized face image using fixed corner points (MATLAB's imcrop function), and the remaining facial region is obtained by removing these patches from the whole face. Each component is down-sampled at a resolution ratio reflecting its discriminative significance: eyes at 1:1, nose at 1:2, mouth at 1:4, and the remaining region/whole face at 1:8 (higher ratios, e.g. 1:16, are also evaluated). For a sample image of size 192×128, the pre-processed face is reduced to a 32×32 (1024-element) representation; the left- and right-eye patches (12×7 each) contribute 84 elements each; the nose patch (10×8) contributes 40 elements; the mouth patch (4×14) contributes 14 elements; and the remaining region (32×32) contributes 128 elements. These patches are scanned in lexicographic order and concatenated into a single 350×1 feature column vector [L, R, N, M, Re].

C. Feature Descriptors

In addition to raw pixel-based multi-scale patches, four complementary descriptors are evaluated for feature extraction:

1) Scale Invariant Feature Transform (SIFT): SIFT detects keypoints through Difference-of-Gaussians (DoG) scale-space extrema, localizes stable candidate keypoints by eliminating low-contrast responses, assigns a dominant orientation to each keypoint, and finally computes a distinctive local descriptor. Detection was performed on 112×92 ORL images using three layers per octave, a contrast threshold of 0.04, an edge-filtering threshold of 10, and a Gaussian sigma of 1.6.

2) Histogram of Oriented Gradients (HOG): HOG represents local shape and appearance via the distribution of gradient orientations. The gradient magnitude and orientation

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 10, NCREHAASTE - 2026

are computed per pixel, and histograms are accumulated over 8×8-pixel cells using nine orientation bins, with block-wise (2×2 cell) normalization.

3) Canny Edge/Contour Detector: The Canny detector (Canny, 1986) computes image gradients using the Sobel operator and then tracks strong edges to construct a contour map of the face.

4) Gabor Filter: The Gabor filter is a band-pass convolution filter combining a Gaussian envelope with a sinusoidal carrier, used here for texture analysis and descriptor extraction across selected frequency bands.

D. CNN Classifier Architecture

The classifier is a CNN consisting of three convolutional layers with 6, 16, and 64 filters respectively—5×5 kernels for the first two layers and a 3×3 kernel for the third—each using a ReLU activation. Each convolutional layer is followed by a 2×2 max-pooling layer that halves the spatial dimensions. The resulting feature maps are flattened and passed through two fully connected layers; the first uses ReLU activation followed by dropout (rate 0.5) to mitigate overfitting, and the second uses Softmax or Sigmoid activation for final classification. Network inputs are 48×48 grayscale images. Training uses the categorical cross-entropy loss function, with the Adam, Adamax, RMSprop, and SGD optimizers evaluated for comparison.

E. Dimensionality Reduction: PCA and LDA

In the holistic branch of the proposed system, PCA (the Karhunen–Loève transform) is used to project the high-dimensional face-vector space (e.g., 10,304 dimensions for a 112×92 image) onto a lower-dimensional 'face space' formed from the eigenvectors of the training-image covariance matrix with the largest eigenvalues, reducing redundancy while retaining the most informative variance. LDA (Fisherfaces) instead seeks a projection that maximizes the ratio of the between-class scatter matrix Sb to the within-class scatter matrix Sw, thereby grouping same-class samples closely while separating different classes—a criterion directly aligned with the classification objective, unlike PCA's reconstruction-oriented criterion. In the proposed pipeline, the multi-scale feature vector formed in Section III-B is projected onto the PCA or LDA feature space to yield a compact weight vector, which is then supplied as input to the CNN classifier.

F. Face Databases

Experiments are conducted on two standard databases. The ORL (Olivetti Research Laboratory) database contains 400 images of 40 subjects (10 images per subject), with variation in pose, expression, lighting, and presence/absence of spectacles. The FERET database is used with up to 500 frontal images of 100 subjects (5 variations per subject), captured under variable pose, expression, illumination, and eyewear across different photo sessions.

G. Algorithmic Steps

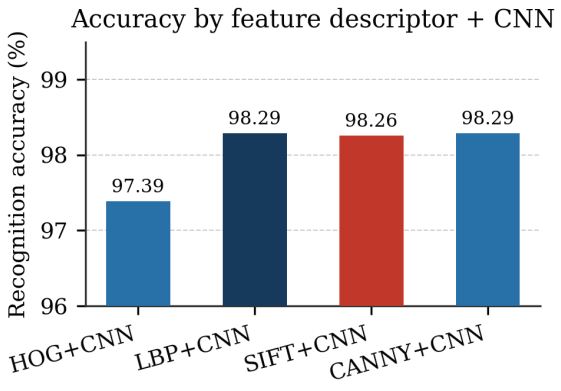
The complete proposed algorithm proceeds as follows: (1) pre-process and normalize the input face image; (2) resize the image using bi-cubic interpolation; (3) extract the face ellipse using a standard mask; (4) crop the eye, nose, and mouth patches using fixed corner points, held consistent across the training and test sets; (5) down-sample each component according to its resolution ratio; (6) extract the remaining facial region and down-sample it (or the whole face) at a higher ratio; (7) concatenate the component patches in lexicographic order into a single column feature vector; (8) either feed this vector directly to the CNN classifier (Method 1), or project it onto the PCA/LDA feature space to obtain a weight vector before CNN classification (Method 2). During testing, the same steps are applied to the probe image, and its resulting (weight) vector is matched against the trained CNN without retraining the network.

IV. RESULTS AND DISCUSSION

Experiments were conducted on both the ORL database (400 images, 40 subjects) and the FERET database (frontal images, 100 subjects), with ORL results averaged over 10 independent runs using different training/testing splits, and FERET results computed for a single run. Recognition performance was evaluated with respect to face-image resolution, component down-sampling ratio, training-set composition, number of hidden-layer nodes, and number of principal components.

TABLE II
Recognition Accuracy by Feature-Descriptor + CNN Combination

Feature + Classifier	Accuracy (%)
HOG + CNN	97.39
LBP + CNN	98.29
SIFT + CNN	98.26
CANNY + CNN	98.29



(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 10, NCREHAASTE - 2026

Fig. 2. Recognition accuracy for the four evaluated feature-descriptor + CNN combinations (Table II).

Table II reports recognition accuracy for four feature-descriptor/CNN combinations; LBP+CNN and CANNY+CNN each achieve 98.29%, closely followed by SIFT+CNN at 98.26% and HOG+CNN at 97.39%. Under further optimizer comparison, SIFT+CNN combined with the Softmax activation function and Adam/Adamax/RMSprop optimizers achieved a peak accuracy of up to 99%. Overall training-model accuracy reached 98.29%.

Recognition performance on the FERET database was consistently higher than on ORL, attributable to the smaller degree of facial-expression variation in FERET relative to ORL. Across both databases, using facial components alone (excluding the full face) resulted in only a 6–7% drop in recognition accuracy compared to full-face recognition, confirming that the eyes, nose, and mouth jointly retain most of the discriminative information in a face. Furthermore, replacing the raw multi-scale feature vector with its PCA/LDA (Eigen/Fisher) weight-vector projection before CNN classification improved recognition accuracy by a further 1–2%, confirming the value of the two-stage dimensionality-reduction-plus-classification pipeline.

TABLE III
Recognition (%) vs. Component Down-Sampling Ratio, No Expression Variation

Set	L.Eye	R.Eye	Nose	Mouth	Rem.	ORL(%)	FERET(%)
1	1:1	1:1	1:2	1:4	1:8	96.22	97.00
2	1:1	1:1	1:4	1:8	1:16	99.34	99.00
3	1:2	1:1	1:4	1:8	1:8	98.21	98.00
4	1:2	1:2	1:8	1:8	1:16	93.85	97.00
5	1:2	1:2	1:2	1:4	1:8	95.68	96.00
6	1:1	1:1	1:2	1:8	1:16	97.87	98.87

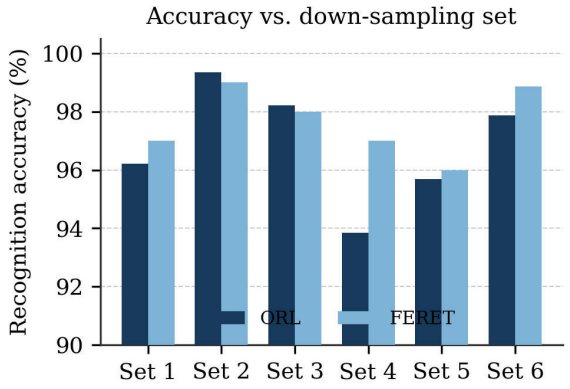


Fig. 3. ORL vs. FERET recognition accuracy across the six component down-sampling sets in Table III.

Table III (whole-face recognition percentages, values based on component down-sampling combinations across the ORL and FERET databases) shows that recognition accuracy is highest when the eyes retain full resolution (1:1) while the nose, mouth, and remaining region are progressively down-sampled—consistent with the assumption that the eyes carry the most discriminative information. Set 2 (eyes 1:1, nose 1:4, mouth 1:8, remaining 1:16) achieves the best overall performance on both databases (99.34% ORL, 99% FERET), while aggressively down-sampling the eyes themselves (Set 4) degrades performance noticeably.

Additional trends observed during CNN-architecture tuning: recognition accuracy increases with the number of hidden-layer nodes, since a larger weight-matrix dimension eases training, but accuracy eventually declines beyond a certain size due to increased computational complexity; the best-performing configuration was a 40:50:40 CNN structure. Similarly, accuracy improves as the number of principal components increases up to roughly 30 (about 15% of a 200-face gallery) and then levels off, declining again once the component count exceeds 50.

V. CONCLUSION

This paper presented a hybrid face recognition technique that combines holistic and component-based (eyes, nose, mouth) information through multi-scale down-sampling, PCA/LDA-based dimensionality reduction, and CNN-based classification. Two recognition schemes were evaluated—direct CNN classification of the multi-scale feature vector, and CNN classification of its PCA/LDA-projected weight vector—across the ORL and FERET databases, at multiple resolutions and under partial-face (occluded) conditions. The results show that SIFT+CNN achieves the best overall accuracy of up to 99%, that facial components alone yield recognition rates only 6–7% below full-face recognition, and that PCA/LDA projection prior to CNN classification improves accuracy by a further 1–2% over the direct-feature approach. The face recognition rate was observed to increase with image resolution and level off beyond roughly 56×40 pixels, and to increase with the number of hidden-layer nodes and principal components up to identifiable optimal points before declining due to increased model complexity. These findings support the use of eye-weighted, multi-scale, hybrid feature representations for robust face recognition under resolution and occlusion constraints, with clear applicability to access control, authentication, and video-surveillance systems.

VI. FUTURE SCOPE

Future work can target: (1) improved accuracy under more extreme illumination, pose, and occlusion conditions, potentially by incorporating channel-wise occlusion-aware dropout or attention mechanisms of the kind proposed in [21] and [22] directly into the CNN classifier used here; (2)

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 10, NCREHAASTE - 2026

optimization for real-time deployment in access-control, surveillance, and security applications; (3) multi-modal biometric fusion with iris, fingerprint, or voice recognition to strengthen overall security; and (4) evaluation on larger and more diverse databases to assess generalization beyond ORL and FERET, including masked and unconstrained-occlusion benchmarks consistent with recent masked-face-recognition literature [23]. Practical deployment must also address data-privacy concerns inherent to biometric data collection, robustness against spoofing attacks (e.g., photographs or masks), and continued sensitivity to lighting and pose variability, which remain open challenges for large-scale, responsible deployment of face recognition systems.

REFERENCES

[1] K. K. Kumar, Y. Kasiviswanadham, D. V. S. N. V. Indira, P. Priyanka Palesetti, and Ch. V. Bhargavi, "Criminal face identification system using deep learning algorithm multi-task cascade neural network (MTCNN)," Materials Today: Proceedings, 2021.

[2] P. Wang, P. Wang, and E. Fan, "Violence detection and face recognition based on deep learning," Pattern Recognition Letters, vol. 142, pp. 20-24, 2021.

[3] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91-110, 2004.

[4] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Proc. IEEE CVPR, 2005, vol. I, pp. 886-893.

[5] S. Fischer, G. Cristóbal, and R. Redondo, "Sparse overcomplete Gabor wavelet representation based on local competitions," IEEE Trans. Image Processing, vol. 15, no. 2, pp. 265-272, 2006.

[6] M. Turk and A. Pentland, "Eigenfaces for recognition," Journal of Cognitive Neuroscience, vol. 3, pp. 71-86, 1991.

[7] P. Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711-720, 1997.

[8] A. Pentland, B. Moghaddam, and T. Starner, "View-based and modular eigenspaces for face recognition," in Proc. IEEE CVPR, 1994, pp. 84-91.

[9] M. Kirby and L. Sirovich, "Application of the Karhunen-Loève procedure for the characterisation of human faces," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 12, pp. 831-835, 1990.

[10] J. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Face recognition using LDA-based algorithms," IEEE Trans. Neural Networks, vol. 14, no. 1, pp. 195-200, 2003.

[11] C. Li, Y. Huang, W. Huang, and F. Qin, "Learning features from covariance matrix of Gabor wavelet for face recognition under adverse conditions," Pattern Recognition, vol. 119, 2021.

[12] Y. Zhu and Y. Jiang, "Optimization of face recognition algorithm based on deep learning multi feature fusion driven by big data," Image and Vision Computing, vol. 104, 2020.

[13] J. Mohammed Sahan, E. I. Abbas, and Z. M. Abood, "A facial recognition using a combination of a novel one-dimension deep CNN and LDA," Materials Today: Proceedings, 2021.

[14] M. Savvides, R. Abiantun, J. Heo, C. Xie, and B. V. K. Vijayakumar, "Partial & holistic face recognition on FRGC-II data using support vector machine," in Proc. CVPR Workshop, 2006, p. 48.

[15] S. Gutta and H. Wechsler, "Partial faces for face recognition: Left vs. right," in Proc. CAIP, Springer LNCS 2756, 2003, pp. 630-637.

[16] A. S. Tolba, A. H. El-Baz, and A. A. El-Harby, "Face recognition: A literature review," International Journal of Signal Processing, vol. 2, no. 2, 2006.

[17] Research and Markets, "Global Facial Recognition Market Report 2018," 2018.

[18] A. Nemavhola, S. Viriri, and C. Chibaya, "A scoping review of literature on deep learning techniques for face recognition," Human Behavior and Emerging Technologies, vol. 2025, Art. no. 5979728, 2025.

[19] Suharjito and A. D. Puspita, "Face recognition on low resolution face image with TBE-CNN architecture," Advances in Science, Technology and Engineering Systems Journal, vol. 5, no. 2, pp. 730-738, 2020.

[20] A.-C. Tsai, O. Yang-Yen, W.-C. Wu, and J.-F. Wang, "Integrated single shot multi-box detector and efficient pre-trained deep convolutional neural network for partially occluded face recognition system," IEEE Access, vol. 9, pp. 164148-164158, 2021.

[21] M. He, J. Zhang, S. Shan, X. Liu, Z. Wu, and X. Chen, "Locality-aware channel-wise dropout for occluded face recognition," IEEE Transactions on Image Processing, vol. 31, pp. 788-798, 2021.

[22] P. C. Neto, J. Ribeiro Pinto, F. Boutros, N. Damer, A. F. Sequeira, and J. S. Cardoso, "Beyond masks: On the generalization of masked face recognition models to occluded face recognition," IEEE Access, vol. 10, pp. 86222-86233, 2022.

[23] M. Eman, T. M. Mahmoud, M. M. Ibrahim, and T. Abd El-Hafeez, "Innovative hybrid approach for masked face recognition using pretrained mask detection and segmentation, robust PCA, and KNN classifier," Sensors, vol. 23, no. 15, Art. no. 6727, 2023.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)