

## Article

# Multi-Resolution Discrete Cosine Transform Fusion Technique Face Recognition Model

Bader M. AlFawwaz <sup>1,\*</sup>, Atallah AL-Shatnawi <sup>1</sup>, Faisal Al-Saqqar <sup>2</sup> and Mohammad Nusir <sup>3</sup>

<sup>1</sup> Department of Information Systems, Prince Hussein Bin Abdullah Faculty for Information Technology, Al al-Bayt University, Mafraq 25113, Jordan; atallah@aabu.edu.jo

<sup>2</sup> Department of Computer Science Department, Prince Hussein Bin Abdullah Faculty for Information Technology, Al al-Bayt University, Mafraq 25113, Jordan; faisalss@aabu.edu.jo

<sup>3</sup> CBM Integrated Software Inc. (CBMIS), San Diego, CA 92101, USA; amashatnawi@yahoo.com

\* Correspondence: bm\_alfawwaz@aabu.edu.jo; Tel.: +962-6297000 (ext. 3399)

**Abstract:** This work presents a Multi-Resolution Discrete Cosine Transform (MDCT) fusion technique Fusion Feature-Level Face Recognition Model (FFLFRM) comprising face detection, feature extraction, feature fusion, and face classification. It detects core facial characteristics as well as local and global features utilizing Local Binary Pattern (LBP) and Principal Component Analysis (PCA) extraction. MDCT fusion technique was applied, followed by Artificial Neural Network (ANN) classification. Model testing used 10,000 faces derived from the Olivetti Research Laboratory (ORL) library. Model performance was evaluated in comparison with three state-of-the-art models depending on Frequency Partition (FP), Laplacian Pyramid (LP) and Covariance Intersection (CI) fusion techniques, in terms of image features (low-resolution issues and occlusion) and facial characteristics (pose, and expression per se and in relation to illumination). The MDCT-based model yielded promising recognition results, with a 97.70% accuracy demonstrating effectiveness and robustness for challenges. Furthermore, this work proved that the MDCT method used by the proposed FFLFRM is simpler, faster, and more accurate than the Discrete Fourier Transform (DFT), Fast Fourier Transform (FFT) and Discrete Wavelet Transform (DWT). As well as that it is an effective method for facial real-life applications.

**Keywords:** feature fusion; face recognition; Laplacian Pyramid; multi-resolution discrete cosine transform



**Citation:** AlFawwaz, B.M.; AL-Shatnawi, A.; Al-Saqqar, F.; Nusir, M. Multi-Resolution Discrete Cosine Transform Fusion Technique Face Recognition Model. *Data* **2022**, *7*, 80. <https://doi.org/10.3390/data7060080>

Academic Editor: Giuseppe Ciaburro

Received: 14 April 2022

Accepted: 11 June 2022

Published: 15 June 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Since its beginnings in the 1970s, research on face recognition has increased rapidly, particularly since the revolution in ICT and digital photography and scanning technologies since the 1990s [1]. A Face Recognition System (FRS) can deploy automated technologies to check faces according to stored and verified features. On the most fundamental level, FRSs determine and predict the extent of similarity between two faces (i.e., between two images of faces) to enable verification of a person's identity; on the elemental level, it takes a photo of a person seeking verification and compares the photo against a pre-verified stored (library) image. This rudimentary overview includes numerous phases of varying complexity in practical applications. When an FRS is confronted by a face, it must first detect that this target object is indeed a face. During this basic process of determining a face (e.g., using algorithmic calculations based on major facial characteristics and physical distances etc. within the facial topography), the FRS can undertake initial attempts to identify the particular face, by checking whether it matches any known (stored) faces in accessible databases [2].

FRS has diverse applications in many fields requiring identity verification and person location, such as finding missing children or criminals in crowds, enabling employee access to appropriate areas within a workplace, and calling up customer or patient records from governmental or healthcare databases [3,4]. FRS usually involves detection and

pre-processing of the face being processed, followed by feature extraction, and then face recognition [5].

In the initial phase, the face image is typically improved by various enhancements to eliminate noise within the image (i.e., superfluous data), in order to hone in on precise facial features of interest and utility for identification purposes [5]. Features extracted can be either “local” (e.g., the mouth, nose, eyes, etc.) or global (such as the facial topography and relative locations of local features) [6,7]. Features extracted by these methods are subsequently categorized using machine learning classifiers, such as Artificial Neural Network (ANN), K-nearest neighbour, and Support Vector Machine (SVM) [8,9].

Facial recognition can struggle to define effective and discriminative facial descriptors due to great variability in light levels, pose, expression, image resolution, and partial occlusion, among other potential issues [10–12]. Most modern FRSs deploy single features, while complex and multifarious face recognition tasks required more modalities of facial features to comprehensively classify facial image data [10–12]. Consequently, improvements in information fusion during feature extraction and processing are sorely required for FRS industrial solutions.

Face recognition data fusion is executed at the feature or decision level [13,14]. Feature-level methods combine multiple feature sets in a unified, fused set, which is subsequently processed with a normal classifier. Decision-level techniques combine numerous characteristic-related classifiers to generate a stronger aggregate classifier [10,15,16]. Fusion at the feature level has simple training requirements for the mono-phase learning process for the combined feature vector, and it can detect and utilize multiple features’ correlations during an initial period. It is necessary that fused features are in a uniform format [14,16].

Utilizing such techniques; this research presents an FFLFRM model rooted in Multi-Resolution Discrete Cosine Transform (MDCT) technique for feature fusion. In this research, the MDCT technique is selected for feature fusion due to its effectiveness in enhancing the fused image resolutions [17]. Local and global facial features were extracted via MDCT, followed by Principal Component Analysis (PCA) and Local Binary Pattern (LBP). Subsequently, the feature vectors that had been fused underwent classification with the use of Multi-Layer Perceptron (MLP) ANN. Following this, the developed FFLFRM was executed with 10,000 greyscale images for evaluation. Test images were sourced from the Olivetti Research Laboratory (ORL) facial image database. Model performance was contrasted with three advanced facial recognition models, depending on the fusion techniques of Covariance Intersection (CI), Frequency Partition (FP), and Laplacian Pyramid (LP). The comparison considered the issues of expression, illumination, low image resolution, occlusion, and pose.

This study demonstrates the coherent integration of local and global methods of feature extraction to generate sound conclusions. The particular contributions of this study are to propose an MDCT fusion-based FLFRM; evaluate its performance with MLP ANN in terms of expression, illumination, low resolution of face images, occlusion, and pose with 10,000 facial images from the ORL database; and contrast the developed model’s performance with that of three advanced fusion models based on CI, FP, and LP [18–20]. The proposed MDCT fusion-based FLFRM has a significant impact on real-time face recognition applications, hence is able to analyse more sophisticated facial characteristics and categorize images/faces according to higher-order identity markets. In addition, it can be applied in many fields requiring identity verification and personal location, such as finding missing children or criminals in crowds and calling up customer or patient records from governmental or healthcare databases.

The following section review literature related to techniques and models for facial recognition. Section 3 presents the FFLFRM developed in this study, followed by a presentation and discussion of the experimental results. Finally, Section 5 concludes the paper and identifies areas for future research.

## 2. Literature Review

The integration of data from multiple images in a single image is the crux of image fusion [21]. The resultant image ought to incorporate detailed information and improve applicability for observers (perceiving the image) than the original image/object. In terms of facial recognition, the decision or feature levels can be the key dimension in which this occurs [13]. Decision-level image fusion has been explored by many studies, whereby individual classifiers are scored in relation to individual local features extracted from images [10,22–30], based on whose integration final decisions are made [14,16]. This usually entails the combination of classifiers' output scores [14,16]. LBP, Gabor, and pixel scores were fused in [10], with post-processing normalization. The same local descriptors were used in [31] to fuse variable LDA-depend on one-shot similarity scores and were deployed with further Gabor features by Wolf et al. [32]. Hellinger, ranking-based, one-shot, and two-shot distances are demonstrably capable of attaining highly efficient classification [14].

Feature-level fusion begins by collating the features extracted within a single feature vector, which is subsequently handed on to the classifier [14]. LBP technique was generalized for texture classification in [33] with variances and pixel intensities derived from local patches. Extensive experimentation with three sophisticated texture datasets (KTH-TIPS2b, Outex, and CURET) revealed that the developed model achieved the best classification for KTH-TIPS2b data, with outcomes comparable to advanced existing solutions for CURET, due to including LBP variants within a joint histogram approach.

In [34] a novel image set-matching method was developed comprising strong facial region descriptors depending on local features, numerous exemplar, and sub-space metrics for comparing related facial regions, and joint learning of more discriminative facial regions when ascertaining optimal weight mixing to combine metrics. MOBIO, LFW, and PIE face datasets were used in experiments which determined that the algorithm significantly outperformed comparator techniques, including the Local Principal Angle and the Kernel Affine Hull techniques.

In [35] a novel descriptor for re-identification purposes was developed, utilizing advanced Fisher vectors with a basic attribute vector of pixel coordinates whose per-pixel intensity was calibrated in the ETHZ and VIPeR two-person re-identification benchmarks. The descriptor attained impressive effectiveness for the studied datasets. To obtain a global image representation, the local descriptors were turned into Fisher Vectors before pooling. Consequently, local descriptors encoded by the Fisher Vector (LDFV) were experimentally confirmed.

Yuan et al. [36] used Local Phase Quantization (LPQ) and LBP to devise an FRS whereby the facial image is segmented into different zones; these are subjected to LBP operator analysis for feature detection, while the LPQ operator determines related frequency area features. This LBP-LPQ hybrid FRS presents an improved feature vector for face description. AR and YALE face databases were used for experimental investigations which demonstrated that the method has better facial recognition accuracy than individual methods.

In [37] a new method using Local Ternary Patterns (LTP) and LBP descriptors for facial image representation was developed, with feature similarity selection and classification algorithm to improve recognition. The facial image is initially divided into smaller zones which are used to draw LTP and LBP histograms, which are subsequently collated within a single feature vector. Experimental testing with the ORL database and Extended Yale Face Database B affirmed the impressive performance of the algorithm.

Gu and Liu [38] developed a new LBP feature extraction method with encoded features and texture data, defined by Gabor wavelet features, edges, and colour features. The process begins by extracting feature pixels from the target image to form a binary image, then a distance-vector field is generated with distance vector calculation between each pixel and the nearest feature pixels within the binary image. Experiments using eye detection with FERET and BioID datasets revealed the suitability of the technique (FLBP), which achieved more accurate localization of the eye centre compared to other tested models.

Li [39] obtained local SIFT and LBP features from densely sampled and multi-scale image spots. They trained a Gaussian mixture model (GMM) connection each feature with the corresponding position, to obtain all of the training set facial images' spatial distribution. To verify facial identification, SVM was trained on the vector to calibrate vector variance related to all feature pairs to determine whether a tracked faces matched. They proposed a joint Bayesian adaptation technique to calibrate the general (universally trained) GMM and model pose differences in target faces, which consistently improved face verification accuracy. They demonstrated that their method significantly outperformed alternative models for the YouTube video face dataset's most highly constricted protocol and the Labeled Face in the Wild (LFW) dataset.

Vu's [40] novel facial image description technique, Patterns of Oriented Edge Magnitudes (POEM), considers links between orientations and gradient magnitudes of numerous local image structures. Whitened PCA dimensionality reduction method was applied for POEM- and POD-based images to attain compact and additional discriminative face descriptors. An experimental investigation with numerous common benchmarks was conducted, including the LF and FERET data sets' non-frontal and frontal images. The outcomes indicated that POEM achieved higher efficiency than alternative methods, with greater simplicity and more powerful performance.

Tan and Triggs [10] developed an FRS with feature-level fusion by extracting two feature sets with Gabor wavelet descriptors and LBP local appearance. The joined feature vector was tested with the Kernel Discriminative Common Vector technique to indicate discriminant non-linear recognition features. Performance outcomes were tested with numerous face databases, including FERET, FRGC 1.0.4, and FRGC 2.0.4.

Mirza [41] explored the fusion of global and local features for gender classification. LBP was utilized to extract local features, supported by two-dimensional DCT, while global feature extraction used PCA and Discrete Cosine Transform (DCT). The suggested system was tested extensively using the FERET dataset, for which it achieved a recognition efficiency of 98.16%.

Yan et al. [42–46] suggested a Multi-feature Fusion and Decomposition (MFD) model based on the multi-head attention and the backbone network for age-invariant face recognition. The proposed model reduces the intra-class variants by learning the discriminative, efficient, and robust features. CACD and CACD-VS databases were used for experimental investigations which demonstrated that the suggested model has better facial recognition accuracy than state-of-the-art models.

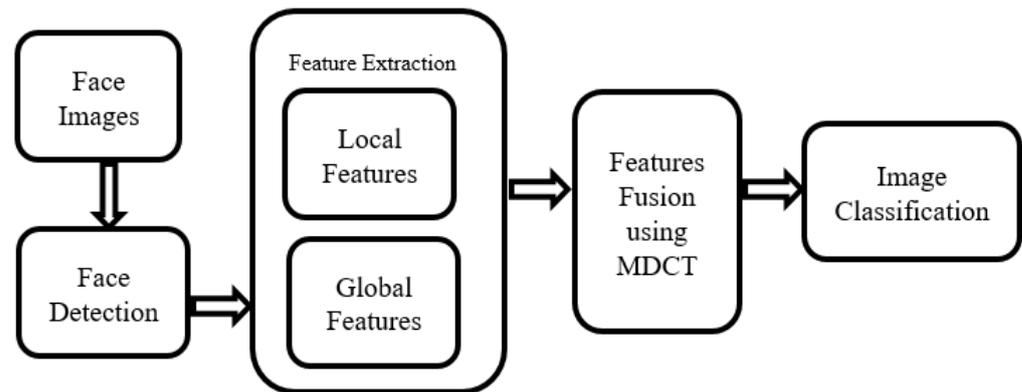
Nusir [18] developed an FRS using the FP method and feature-level fusion to collate local and global features with LBP and PCA. Experimental work on facial images from the ORL database revealed that the developed method achieved improved face recognition efficiency and was more powerful than individual LBP- and PCA-based methods.

Qing Guo et al. [47] integrated the expectation-maximization (EM) algorithm with the covariance intersection (CI) principle in a new image fusion approach using data source cross-correlation, thereby producing estimates with accuracy and consistency due to the use of convex combinations. In practical applications covariance information is usually unknown, thus EM helps by providing a maximum likelihood estimate (MLE) for the covariance matrix.

Al-Shatnawi et al. [19] used Laplacian Pyramid (LP) for facial recognition with feature-level fusion, combining local and global features utilizing LBP and PCA. Testing with ORL database facial images with MLP NN revealed that the developed model produced more efficient face recognition results than models depending on LBP and PCA techniques alone in terms of facial expression, illumination, and occlusion challenging contexts. El-Bashir et al. [20] also produced a feature-level fusion-based face recognition model based on the CI technique that they experimentally evaluated using ORL database facial images with MLP NN, and they also reported the improved effectiveness of their developed model.

### 3. Proposed Model (FFLFRM)

This study proposes an MDCT fusion-based FFLFRM to detect faces and surfaces, extract and fuse features, and classify faces with MLP ANN (Figure 1). Initial face detection based on core features (e.g., eyes, etc.) uses the Haar-cascade face detection technique. Local and global feature extraction utilizing LBP and PCA enables MDCT fusion. Subsequently, fused feature vectors are input in the MLP ANN for facial classification. The proposed FFLFRM architecture is displayed in Figure 1, and its steps are described below.



**Figure 1.** MDCT-Based FFLFRM Architecture.

#### 3.1. Face Detection with Haar-Cascade

The Haar-cascade face detection method identifies the core facial features (i.e., mouth, nose, and eyes) depending on appearance [2,12,48]. Haar-like facial features are used, unlike pixel analysis techniques [49]. Haar-like features comprise related features' quasi-rectangular shapes representing target features [49]. Extraction typically deploys integral image methods, Adaptive Boosting (AdaBoost), and attentional cascade [50]. This research deployed Wang's [50] modified Viola-Jones Haar-cascade face detection system to detect four face feature patches: mouth, nose, eyes, and face (general).

#### 3.2. Global Facial Feature Extraction with Principal Component Analysis (PCA)

Feature extraction is the most essential stage in FRS, which determines its ultimate effectiveness, seeking to efficiently represent the facial image based on characteristics (features) extracted globally or locally using PCA and LBP, prior to MDCT fusion of the features [12,51]. Traditional statistical linear transform PCA is widely used for pattern [52] and face recognition [53]. PCA extracts features statistically by identifying global features for fusion with local ones extracted using MSCT [54]. PCA is commonly used for feature extraction to limit image feature dimensionality. It begins by determining the data matrix mean, then it calculates covariance, eigenvalues, and eigenvectors [55]. PCA identifies the space indicating the maximal difference of studied data, determining the low-dimensional space (PCA space ( $W$ )) used for data transformation ( $X = \{x_1, x_2, \dots, x_N\}$ ) from a higher- to lower-dimensional space, where  $N$  is the number of samples, and  $x_i$  expresses the  $i_{th}$  observation, sample, or pattern [56].

#### 3.3. Local Facial Feature Extraction with LBP

Texture analysis is the underlying technique of LBP local statistical face feature extraction [57], which fuses locally extracted features with global features extracted with MDCT. The central pixel within a  $3 \times 3$  pixel block is used to determine feature values based on the pixel threshold. The complete facial image is then rendered as a feature vector in decimal values [18,58].

### 3.4. Feature Fusion with Multi-Resolution Discrete Cosine Transform (MRDCT)

Naidu’s [17] MDCT-based image fusion of extracted local and global facial features deployed feature vectors subsequently used by MLP ANN. As per the 1D DCT method, MDCT-based fusion separates the target image into columns and rows, which are then processed to generate 1D vector data; the data vector subsequently undergoes DCT. MDCT decomposes the vector to the maximum decomposition available ( $2^7 - 2^9$ ) by deployed DCT, to generate image enhancement. DCT decomposition of the vector results in High Frequency (HF) and Low Frequency (LF) coefficients; the former has minimal image data, while the latter encompasses the prerequisite material to generate the processed, fused image. The LF data is thus sent to Inverse DCT to obtain new vector data, which enables further decomposition for the next level [17].

$$Z_k(u) = DCT(z_k(x)), x, u = 0, 1, \dots, \frac{MN}{2^{k-1}} - 1 \tag{1}$$

$$ZL_k(u) = Z_k(u), u = 0, 1, \dots, 0.5 \frac{MN}{2^{k-1}} - 1 \tag{2}$$

$$ZH_k(u) = Z_k(u), u = 0.5 \frac{MN}{2^{k-1}}, 0.5 \frac{MN}{2^{k-1}} + 1, \dots, MN - 1 \tag{3}$$

$$Z_{k+1}(x) = IDCT(ZL_k(u)) \tag{4}$$

where  $k$  is the multi-resolution decomposition level; note that  $z_k(x) = Z(x)$  at  $k = 1$  level.

Image  $z_1(x, y)$  and  $z_2(x, y)$  integration by MDCT is undertaken as described below.

1: Convert images  $z_{k1}(x, y)$  and  $z_{k2}(x, y)$  to 1D vectors  $z_{k1}(x)$  and  $z_{k2}(x)$ .

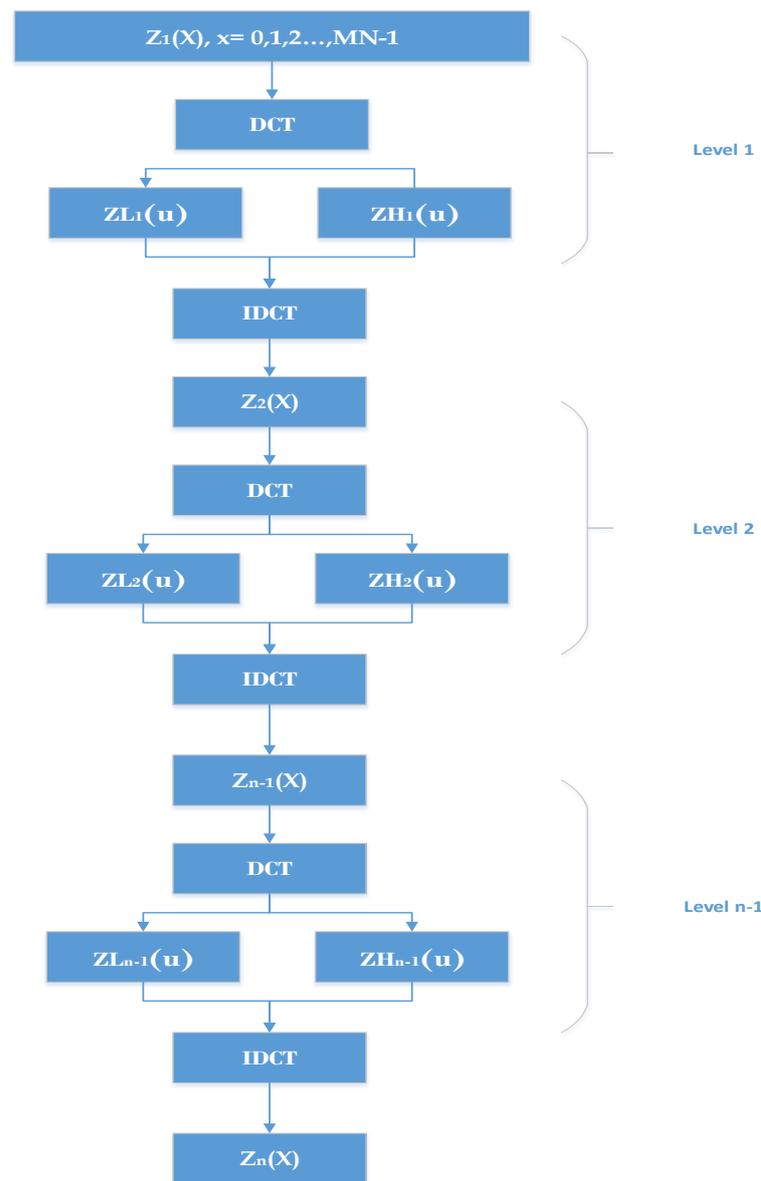
2: Find the DCT coefficients  $z_{k1}(u)$  and  $z_{k2}(u)$  using the previous two vectors  $z_{k1}(x)$  and  $z_{k2}(x)$ . The fusion rule of MDCT coefficients are:

$$ZL_{kf}(u) = \frac{1}{2}(ZL_{k1}(u) + ZL_{k2}(u)), u = 0, 1, \dots, 0.5 \frac{MN}{2^{k-1}} - 1. k = K. \tag{5}$$

$$ZH_{kf}(u) = \begin{cases} ZH_{k1}(u) & \text{if } |ZH_{k1}(u)| \geq |ZH_{k2}(u)| \\ ZH_{k2}(u) & \text{if } |ZH_{k1}(u)| < |ZH_{k2}(u)| \end{cases}, u = 0.5 \frac{MN}{2^{k-1}}, 0.5 \frac{MN}{2^{k-1}} + 1, \dots, \frac{MN}{2^{k-1}} - 1 - 1. k = K, K - 1, \dots, 1 \tag{6}$$

3: Obtain the fused image utilizing Equations (1) through (4). Figure 2 displays the MDCT diagram [17].

This study used FFLRM with the MDCT fusion technique for the fusion of extracted features from the PCA and the LBP. PCA-derived extracted features undergo maximum decomposition in MDCT ( $2^7 - 2^9$ ). Features extracted from the LBP are maximally decomposed by MDCT ( $2^7 - 2^9$ ) to attain a higher-resolution image (than the original). Following LBP and PCA feature decomposition, the MDCT deploys Inverse DCT at each decomposition level to generate the fused image.



**Figure 2.** Multi-Resolution image decomposition using 1D DCT [17].

### 3.5. Face Recognition with Artificial Neural Network (ANN)

ANN is an effective classification tool widely applied in prediction, pattern classification, and approximation activities [5,59–61]. In terms of FRS, multi-layer perceptron ANN (MLP ANN) was used in this paper for fused facial images' facial recognition.

## 4. Experimental Results and Discussion

The effective face recognition model has to ideally deal with a number of challenges. It mainly contains expression changes, illumination, images with low resolution, occlusion, and pose [2]. To validate the proposed FFLFRM based on MDCT model effectiveness, it has been compared with three state-of-the-art models that were developed based on FP [18], LP [19], and CI [20] features fusion techniques. The proposed FFLFRM is based on the MDCT technique, and three state-of-the-art models were run using MATLAB@2015a programming language on a PC with an Intel Core i7 processor (2.40 GHz, 8 GB RAM).

The proposed FFLFRM and three state-of-the-art models were tested and evaluated using MLP ANN for 10,000 face images derived from the ORL database based on the expression changes, illumination, and images with low resolution, occlusion, and pose chal-

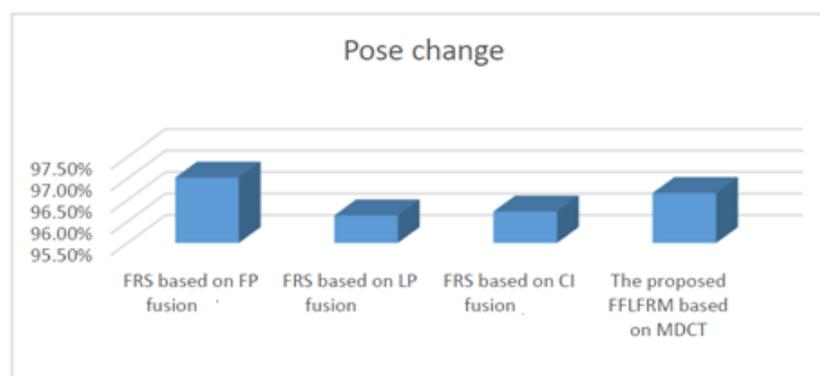
lenges. The comparative evaluation results of their classification efficiencies are presented and summarized in Table 1 and are elaborated on and discussed in more depth below.

**Table 1.** Performance evaluation results of the proposed FFLFRM and the three state-of-the-art models' classification efficiency.

	Pose Change	Illumination Change	Expression Change	Low-Resolution Images	Occlusion
FRS based on FP fusion [18]	97.02%	96.47%	97.73%	96.99%	96.18%
FRS based on LP fusion [19]	96.14%	97.03%	98.2%	96.5%	96.2%
FRS based on CI fusion [20]	96.23%	96.89%	97.68%	96.1%	96.84%
FFLFRM based on MDCT	96.66%	97.07%	97.70%	97.11%	96.87%

#### 4.1. Pose Change

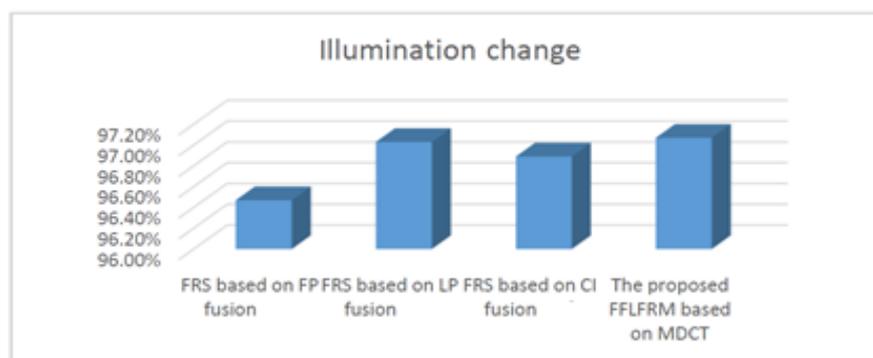
Changing angles of capture (i.e., of cameras) and changing positions of photo subjects (i.e., human faces) during image capture result in changes in pose, which alter facial geometry (in the captured images) [19,36]. This can result in inaccurate renderings of facial characteristics and thus reduce image recognition accuracy. The comparative evaluation results of the proposed FFLFRM and the three state-of-the-art models for pose change are presented in Figure 3. It can be noticed that the proposed FFLFRM achieved 96.66% efficiency in classification, which is lower than one model (the FR-based model, with 97.02%), but slightly higher than the other two using LP and CI (with 96.14% and 96.23%). These results indicate the acceptability of the proposed FFLFRM's effectiveness in facial recognition in the context of changing poses.



**Figure 3.** Performance evaluation results of the proposed FFLFRM and the three state-of-the-art models the pose change.

#### 4.2. Illumination Change

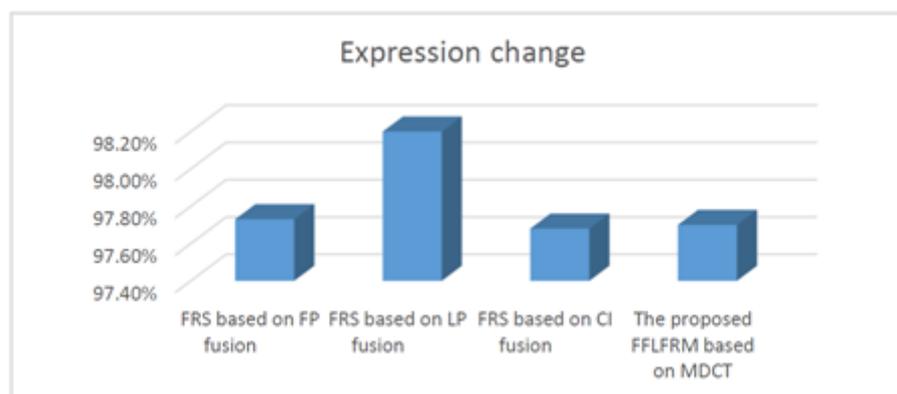
Altering lighting conditions result in illumination changes, which can significantly alter the facial appearance, thereby undermining accuracy in FRS [19,36]. The comparative evaluation results for the proposed FFLFRM and the three state-of-the-art models for illumination change are displayed in Figure 4. It highlights the proposed FFLFRM classification efficiency of 97.07%, outperforming the other models (which had 96.47%, 97.03%, and 96.89% efficiency), indicating its effectiveness for face recognition with illumination change.



**Figure 4.** Performance evaluation results of the proposed FFLFRM and the three state-of-the-art models—illumination change.

#### 4.3. Expression Change

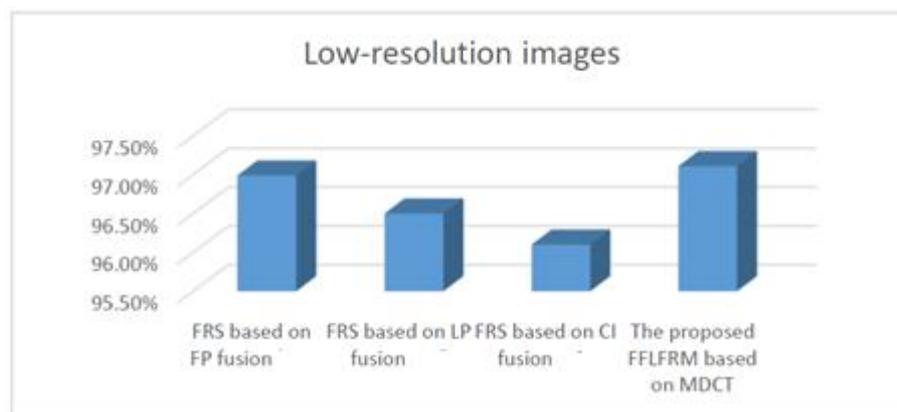
Varying facial expressions are fundamental to human communication, entailing significant changes in facial features, with obvious implications for FRS [19,36]. Comparative evaluation results for performance between the proposed FFLFRM and the three state-of-the-art models for change in facial expression are shown in Figure 5. It can be seen that the FFLFRM had the highest classification efficiency (97.70%). The classification efficiencies of the three other models were 97.73%, 98.02%, and 97.68%. The performance of the proposed FFLFRM is better than that of the model depending on IC, but only lower than that of the models depending on LP and FR. This approves the effectiveness of the proposed FFLFRM in face recognition under the condition of expression change.



**Figure 5.** Performance evaluation results of the proposed FFLFRM and the three state-of-the-art models change in face expression.

#### 4.4. Low-Resolution Images

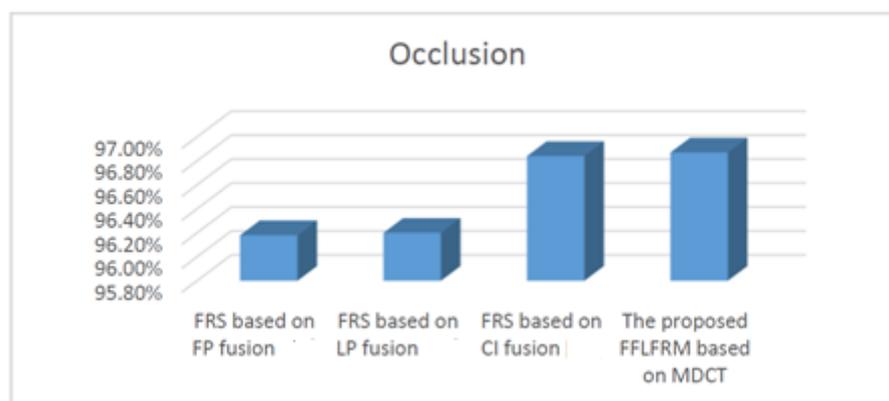
The resolution of an image of a face is affected by various contextual factors, including ambient conditions during image capture (particularly illumination, as discussed above), and the technical specifications and abilities of the camera used to capture the image [19,36]. Low-resolution images generally undermine the accuracy of FRS. Comparative evaluation results for performance between the proposed FFLFRM and the three state-of-the-art models for image resolution are shown in Figure 6. It can be seen that the proposed FFLFRM achieved a classification efficiency of 97.11%, outperforming the other models' classification efficiency rates (96.99%, 96.5%, and 96.1%). This indicates the proposed FFLFRM effectiveness for face recognition with low-resolution images.



**Figure 6.** Performance evaluation results of the proposed FFLFRM and the three state-of-the-art models' image resolution.

#### 4.5. Occlusion Challenge

The complete or partial covering of the face results in occlusion, which hinders feature extraction in FRS [19,36]. Comparative evaluation results for performance between the proposed FFLFRM and the three state-of-the-art models for occlusion are shown in Figure 7. It can be seen that the proposed FFLFRM attained the highest classification efficiency with 96.87%, compared to the other models (with 96.18%, 96.2%, and 96.48%), which affirms the proposed FFLFRM effectiveness for occlusion-condition face recognition.



**Figure 7.** Performance evaluation results of proposed FFLFRM and three state-of-the-art models' occlusion.

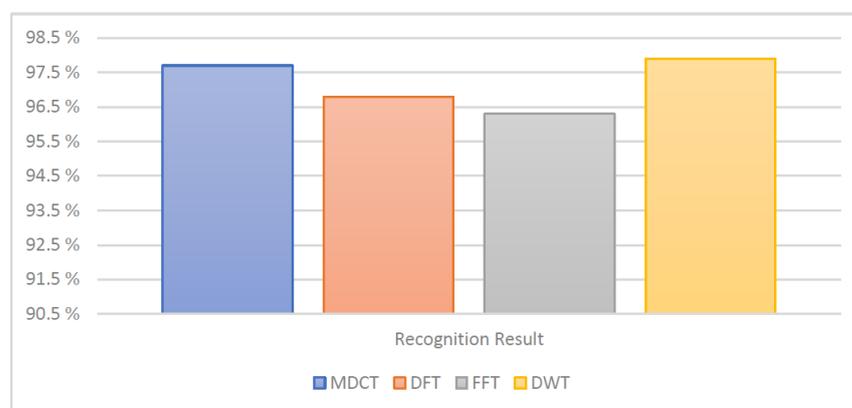
Thus, based on the abovementioned five discussed challenges, the proposed FFLFRM based on MDCT model classification efficiency achieved the best results outperforming the other three state-of-the-art models in terms of illumination change, dealing with low-resolution images, and working effectively with occlusion-condition. But in terms of changing poses challenges, the proposed model achieved promising classification results higher than the two state-of-the-art based on LP and CI models. Furthermore, it achieved better classification efficiency compared to the CI state-of-the-art model under the condition of expression change. These results indicate the acceptability of the proposed FFLFRM effectiveness in facial recognition in the context of the above-mentioned challenges. Consequently, it proves the proposed FFLFRM based on the MDCT model effectiveness compared to the three state-of-the-art models (i.e., Frequency Partition (FP), Laplacian Pyramid (LP), and Covariance Intersection (CI)).

Moreover, to validate the proposed FFLFRM based on MDCT model effectiveness compared to various transformation methods such as Discrete Fourier Transform (DFT), Fast Fourier Transform (FFT) and Discrete Wavelet Transform (DWT), we have tested

and evaluated the proposed FFLFRM and three abovementioned transformations methods using MLP ANN for 10,000 face images derived from the ORL database based on classification accuracy and execution time during the training and testing phases. The comparative evaluation results of the proposed FFLFRM based on MDCT and the three abovementioned methods of classification efficiencies are presented and summarized in Table 2, and graphically shown in Figure 8. As well as the executed time of the FFLFRM model using the four transformation methods (i.e., MDCT, DFT, FFT, and DW) during the training and testing phases using the ORL dataset is presented in Table 3.

**Table 2.** Performance evaluation results of the proposed FFLFRM based on MDCT and the three transformation methods classification efficiency.

Methods	MDCT	DFT	FFT	DWT
Recognition results	97.7%	96.8%	96.3%	97.5%



**Figure 8.** Performance evaluation results of the proposed FFLFRM using MDCT, DFT, FFT, and DW fusion methods tested on the ORL database using MLP ANN.

**Table 3.** The proposed FFLFRM based on MDCT and the three transformation methods execution time on the ORL dataset.

Methods/Time	MDCT	DFT	FFT	DWT
Training	$2.7 \times 10^{-5}$	$1.2 \times 10^{-4}$	$2.7 \times 10^{-4}$	$1.9 \times 10^{-2}$
Testing	$3.4 \times 10^{-6}$	$4.3 \times 10^{-4}$	$3.9 \times 10^{-4}$	$2.00 \times 10^{-3}$

It is noticed that in Table 2, the proposed FFLFRM model using the MDCT method achieved the best classification accuracy outperforming the other three transformation methods (i.e., DFT, FFT, and DW) when tested on the ORL database using MLP ANN. Furthermore, Table 3 proves that the proposed FFLFRM model using the MDCT method is faster than the other three abovementioned transformation methods, and thus proves the effectiveness of the MDCT as a feature fusion technique for fiscal real-life applications.

## 5. Conclusions and Future Works

This research has presented an FFLFRM rooted in the MDCT fusion technique with face detection, feature extraction and fusion (using the MDCT method), and face classification (utilizing the MLP ANN) stages. Facial characteristics such as eyes and mouth etc. were initially detected by the Haar-cascade technique, followed by local and global feature extraction with PCA and LBP for the MDCT fusion technique. Consequently, MLP ANN applied facial classification based on fused feature vector inputs. Testing of the developed model's performance in terms of facial recognition involved comparative analysis with the accurate classification performance of three state-of-the-art models for fusion-level

face recognition, utilizing the FP, LP, and CI fusion techniques. Testing was conducted to evaluate models' performance with 10,000 images from the ORL database utilizing the MLP ANN conditions. The proposed model achieved the following classification efficiency levels for the studied conditions: expression change (97.70%), illumination change (97.07%), low image resolution (97.11%), occlusion (96.87), and pose change (96.66%).

The proposed FFLFRM based on MDCT model classification efficiency achieved the best results outperforming the other three state-of-the-art models in terms of illumination change, dealing with low-resolution images, and working effectively with occlusion-condition. In terms of dealing with low-resolution images, the proposed FFLFRM based on the MDCT model produced (97.11%) classification accuracy, while the other three state-of-the-art models produced (96.99%, 96.5% and 96.1%) classification accuracies respectively. In terms of working effectively with occlusion the proposed FFLFRM based on the MDCT model produced (96.87%) classification accuracy, while the other three state-of-the-art models produced (96.18%, 96.2% and 96.2%) classification accuracies respectively. In terms of illumination change the proposed FFLFRM based on the MDCT model produced (97.07%) classification accuracy, while the other three state-of-the-art models produced (96.47%, 97.03% and 96.89%) classification accuracies respectively. But, dealing with changes poses challenges, the proposed model achieved promising classification results that are higher than the two state-of-the-art based on LP and CI models. Hence it produced (96.66%) classification accuracy, while the other three state-of-the-art models produced (97.02%, 96.14% and 96.23%) classification accuracies respectively. Furthermore, it achieved better classification efficiency compared to the CI state-of-the-art model under the condition of expression change. Hence it produced (97.7%) classification accuracy, while the other three state-of-the-art models produced (97.73%, 98.2% and 97.68%) classification accuracies respectively. These results indicate the acceptability of the proposed FFLFRM effectiveness in facial recognition in the context of the abovementioned challenges. Consequently, it proves the proposed FFLFRM based on the MDCT model effectiveness compared to the three state-of-the-art models (i.e., Frequency Partition (FP), Laplacian Pyramid (LP), and Covariance Intersection (CI)).

Furthermore, the proposed FFLFRM based on MDCT model effectiveness is verified and compared with Discrete Fourier Transform (DFT), Fast Fourier Transform (FFT) and Discrete Wavelet Transform (DWT) transformation methods, in terms of classification accuracy and execution time during the training and testing phases. The results proved that the proposed FFLFRM model using the MDCT method achieved the best classification accuracy compared to the other three transformation methods when tested on the ORL database using MLP ANN. Hence, it produced (97.7%) classification accuracy, whereas the other three methods produced (96.8%, 96.3% and 97.5%) classification accuracies respectively. The proposed FFLFRM based on the MDCT model is faster than the other three transformation methods in both the training and testing phases. Therefore, this paper concluded that MDCT is simpler, faster, and more accurate than the DFT, FFT, and DWT as well as, and it is an effective method for facial real-life applications.

In future research, it is planned to evaluate the FFLFRM's performance with alternative classifiers, including SVM and HMM, and to compare decision-level fusion performance. Further research can explore the fusion of various global and local feature extraction techniques with the deployment of MDCT. Furthermore, it will be interesting to test and evaluate the proposed model using other large-scale face recognition datasets.

**Author Contributions:** Conceptualization, B.M.A., A.A.-S., F.A.-S. and M.N.; methodology, B.M.A., A.A.-S., F.A.-S. and M.N.; software, B.M.A., A.A.-S., F.A.-S. and M.N.; validation, B.M.A., A.A.-S., F.A.-S. and M.N.; formal analysis, B.M.A., A.A.-S., F.A.-S. and M.N.; investigation, B.M.A., A.A.-S., F.A.-S. and M.N.; resources, B.M.A.; data curation, B.M.A., A.A.-S., F.A.-S. and M.N.; writing—original draft preparation, B.M.A., A.A.-S., F.A.-S. and M.N.; writing—review and editing, B.M.A., A.A.-S., F.A.-S. and M.N.; visualization, B.M.A., A.A.-S., F.A.-S. and M.N.; supervision, B.M.A.; project administration, B.M.A.; All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Publicly available dataset “The Database of Faces” by AT & T Laboratories Cambridge was analyzed in this study. This data can be found at <https://cam-orl.co.uk/facedatabase.html> (accessed on 10 June 2022).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Huang, J.; Yuen, P.C.; Lai, J.-H.; Li, C.-H. Face Recognition Using Local and Global Features. *EURASIP J. Adv. Signal Process.* **2004**, *2004*, 870582. [CrossRef]
2. de Carrera, P.; Marques, I. Face Recognition Algorithms. Master’s Thesis, Universidad Euskal Herriko, Leioa-Biscay, Spain, 2010.
3. Taskiran, M.; Kahraman, N.; Erdem, C.E. Face recognition: Past, Present, And Future (A Review). *Digit. Signal Process.* **2020**, *106*, 102809. [CrossRef]
4. Pavlović, M.; Stojanović, B.; Petrović, R.; Stanković, S. Fusion of visual and thermal imagery for illumination invariant face recognition system. In Proceedings of the 2018 14th Symposium on Neural Networks and Applications (NEUREL), Belgrade, Serbia, 20 November 2018. [CrossRef]
5. Al-Allaf, O. Review of face detection systems based artificial neural networks algorithms. *arXiv* **2014**, arXiv:1404.1292. [CrossRef]
6. Ding, H. Combining 2D Facial Texture and 3D Face Morphology for Estimating People’s Soft Biometrics and Recognizing Facial Expressions. Ph.D. Thesis, Université de Lyon, Lyon, France, 2016.
7. Nguyen, H. Contributions to Facial Feature Extraction for Face Recognition. Ph.D. Thesis, Université de Grenoble, Saint-Martin-d’Hères, France, 2014.
8. Dhriti, M.; Kaur, M. K-nearest neighbor classification approach for face and fingerprint at feature level fusion. *Int. J. Comput. Appl.* **2012**, *60*, 13–17. [CrossRef]
9. Le, T.H. Applying Artificial Neural Networks for Face Recognition. *Adv. Artif. Neural Syst.* **2011**, *2011*, 673016. [CrossRef]
10. Štruc, V.; Gros, J.; Dobrišek, S.; Pavešić, N. Exploiting representation plurality for robust and efficient face recognition. In Proceedings of the 22nd International Electrotechnical and Computer Science Conference (ERK’13), Portoroz, Slovenia, 16–19 September 2013; pp. 121–124.
11. Tan, X.; Triggs, B. Fusing Gabor and LBP Feature Sets for Kernel-Based Face Recognition. In *International Workshop on Analysis and Modeling of Faces and Gestures*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 235–249. [CrossRef]
12. Zhao, W.; Chellappa, R.; Phillips, P.; Rosenfeld, A. Face recognition: A literature survey. *ACM Comput. Surv. (CSUR)* **2003**, *35*, 399–458. [CrossRef]
13. Jagalingam, P.; Hegde, A. Pixel level image fusion—A review on various techniques. In Proceedings of the 3rd World Conference on Applied Sciences, Engineering and Technology, Kathmandu, Nepal, 27–29 September 2014.
14. Wang, H.; Hu, J.; Deng, W. Face Feature Extraction: A Complete Review. *IEEE Access* **2017**, *6*, 6001–6039. [CrossRef]
15. Kittler, J.; Hatef, M.; Duin, R.; Matas, J. On combining classifiers. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20*, 226–239. [CrossRef]
16. Singh, M.; Singh, R.; Ross, A. A comprehensive overview of biometric fusion. *Inf. Fusion* **2019**, *52*, 187–205. [CrossRef]
17. Naidu, V.P.S. Novel image fusion techniques using DCT. *Int. J. Comput. Sci. Bus. Inform.* **2013**, *5*, 1–18.
18. Nusir, M. Face Recognition using Local Binary Pattern and Principle Component Analysis. Master’s Thesis, Al al-Bayt University, Al-Mafraq, Jordan, 2018.
19. AL-Shatnawi, A.; Al-Saqqar, F.; El-Bashir, M.; Nusir, M. Face Recognition Model based on the Laplacian Pyramid Fusion Technique. *Int. J. Adv. Soft Comput. Its Appl.* **2021**, *13*, 27–46.
20. El-Bashir, M.S.; AL-Shatnawi, A.M.; Al-Saqqar, F.; Nusir, M.I. Face Recognition Model Based on Covariance Intersection Fusion for Interactive devices. *World Comput. Sci. Inf. Technol. J.* **2021**, *11*, 5–12.
21. Haghighat, M.; Aghagolzadeh, A.; Seyedarabi, H. Multi-focus image fusion for visual sensor networks in DCT domain. *Comput. Electr. Eng.* **2011**, *37*, 789–797. [CrossRef]
22. Liu, Z.; Liu, C. *Robust Face Recognition Using Color Information*. *International Conference on Biometrics*; Springer: New York, NY, USA, 2009; pp. 122–131. [CrossRef]
23. Pinto, N.; DiCarlo, J.; Cox, D. How far can you get with a modern face recognition test set using only simple features. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 2591–2598.
24. Chan, C.-H.; Kittler, J.; Tahir, M.A. Kernel Fusion of Multiple Histogram Descriptors for Robust Face Recognition. In Proceedings of the Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR), Izmir, Turkey, 18–20 August 2010; Springer: New York, NY, USA, 2010; Volume 6218, pp. 718–727. [CrossRef]

25. Chan, C.H.; Tahir, M.A.; Kittler, J.; Pietikainen, M. Multiscale Local Phase Quantization for Robust Component-Based Face Recognition Using Kernel Fusion of Multiple Descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *35*, 1164–1177. [[CrossRef](#)]
26. Arashloo, S.; Kittler, J. Class-specific kernel fusion of multiple descriptors for face verification using multi scale bi-narised statistical image features. *IEEE Trans. Inf. Forensics Secur.* **2014**, *9*, 2100–2109. [[CrossRef](#)]
27. Hu, J.; Lu, J.; Tan, Y.-P. Discriminative Deep Metric Learning for Face Verification in the Wild. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 1875–1882. [[CrossRef](#)]
28. Ding, C.; Xu, C.; Tao, D. Multi-Task Pose-Invariant Face Recognition. *IEEE Trans. Image Process.* **2015**, *24*, 980–993. [[CrossRef](#)]
29. Nikan, S.; Ahmadi, M. Local gradient-based illumination invariant face recognition using local phase quantization and multi-resolution local binary pattern fusion. *IET Image Process.* **2015**, *9*, 12–21. [[CrossRef](#)]
30. Zhang, X.; Mahoor, M.; Mavadati, S. Facial expression recognition using lp-norm MKL multiclass SVM. *Mach. Vis. Appl.* **2015**, *26*, 467–483. [[CrossRef](#)]
31. Taigman, Y.; Wolf, L.; Hassner, T. Multiple One-Shots for Utilizing Class Label Information. In Proceedings of the British Machine Conference Rama Chellappa, College Park, MD, USA, 7–10 September 2009. [[CrossRef](#)]
32. Wolf, L.; Hassner, T.; Taigman, Y. Similarity Scores Based on Background Samples. In *Asian Conference on Computer Vision*; Springer: New York, NY, USA, 2010; pp. 88–97. [[CrossRef](#)]
33. Liu, L.; Zhao, L.; Long, Y.; Kuang, G.; Fieguth, P. Extended local binary patterns for texture classification. *Image Vis. Comput.* **2012**, *30*, 86–99. [[CrossRef](#)]
34. Sanderson, C.; Harandi, M.T.; Wong, Y.; Lovell, B.C. Combined Learning of Salient Local Descriptors and Distance Metrics for Image Set Face Verification. In Proceedings of the 2012 IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance, Beijing, China, 18–21 September 2012; pp. 294–299. [[CrossRef](#)]
35. Ma, B.; Su, Y.; Jurie, F. Local Descriptors Encoded by Fisher Vectors for Person Re-identification. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 413–422. [[CrossRef](#)]
36. Yuan, B.; Cao, H.; Chu, J. Combining Local Binary Pattern and Local Phase Quantization for Face Recognition. In Proceedings of the 2012 International Symposium on Biometrics and Security Technologies, Taipei, Taiwan, 26–29 March 2012; pp. 51–53. [[CrossRef](#)]
37. Tran, C.K.; Lee, T.F.; Chang, L.; Chao, P.J. Face Description with Local Binary Patterns and Local Ternary Patterns: Improving Face Recognition Performance Using Similarity Feature-Based Selection and Classification Algorithm. In Proceedings of the 2014 International Symposium on Computer, Consumer and Control, Taichung, Taiwan, 10–12 June 2014; pp. 520–524. [[CrossRef](#)]
38. Gu, J.; Liu, C. Feature local binary patterns with application to eye detection. *Neurocomputing* **2013**, *113*, 138–152. [[CrossRef](#)]
39. Li, H.; Hua, G.; Lin, Z.; Brandt, J.; Yang, J. Probabilistic Elastic Matching for Pose Variant Face Verification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 23–28 June 2013; pp. 3499–3506. [[CrossRef](#)]
40. Vu, N.-S. Exploring Patterns of Gradient Orientations and Magnitudes for Face Recognition. *IEEE Trans. Inf. Forensics Secur.* **2012**, *8*, 295–304. [[CrossRef](#)]
41. Mirza, A.M.; Hussain, M.; Almuzaini, H.; Muhammad, G.; Aboalsamh, H.; Bebis, G. Gender Recognition Using Fusion of Local and Global Facial Features. In *International Symposium on Visual Computing*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8034, pp. 493–502. [[CrossRef](#)]
42. Yan, C.; Gong, B.; Wei, Y.; Gao, Y. Deep Multi-View Enhancement Hashing for Image Retrieval. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *43*, 1445–1451. [[CrossRef](#)] [[PubMed](#)]
43. Yan, C.; Li, Z.; Zhang, Y.; Liu, Y.; Ji, X.; Zhang, Y. Depth image denoising using nuclear norm and learning graph model. *ACM Trans. Multimed. Comput. Commun. Appl.* **2020**, *16*, 1–17. [[CrossRef](#)]
44. Yan, C.; Hao, Y.; Li, L.; Yin, J.; Liu, A.; Mao, Z.; Chen, Z.; Gao, X. Task-Adaptive Attention for Image Captioning. *IEEE Trans. Circuits Syst. Video Technol.* **2021**, *32*, 43–51. [[CrossRef](#)]
45. Yan, C.; Teng, T.; Liu, Y.; Zhang, Y.; Wang, H.; Ji, X. Precise No-Reference Image Quality Evaluation Based on Distortion Identification. *ACM Trans. Multimed. Comput. Commun. Appl.* **2021**, *17*, 1–21. [[CrossRef](#)]
46. Yan, C.; Meng, L.; Li, L.; Zhang, J.; Wang, Z.; Yin, J.; Zhang, J.; Sun, Y.; Zheng, B. Age-Invariant Face Recognition by Multi-Feature Fusion and Decomposition with Self-Attention. *ACM Trans. Multimed. Comput. Commun. Appl.* **2021**, *18*, 1–18. [[CrossRef](#)]
47. Guo, Q.; Chen, S.; Leung, H.; Liu, S. Covariance intersection based image fusion technique with application to pansharpening in remote sensing. *Inf. Sci.* **2010**, *180*, 3434–3443. [[CrossRef](#)]
48. Viola, P.; Jones, M. Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, Kauai, HI, USA, 8–14 December 2001. [[CrossRef](#)]
49. Viola, P.; Jones, M. Robust real-time face detection. *Int. J. Comput.* **2004**, *57*, 137–154. [[CrossRef](#)]
50. Wang, Y.-Q. An Analysis of the Viola-Jones Face Detection Algorithm. *Image Process. Line* **2014**, *4*, 128–148. [[CrossRef](#)]
51. Jafri, R.; Arabnia, H.R. A Survey of Face Recognition Techniques. *J. Inf. Process. Syst.* **2009**, *5*, 41–68. [[CrossRef](#)]
52. Al-Saqqar, F.; Al-Shatnawi, A.M.; Al-Diabat, M.; Aloun, M. Handwritten Arabic Text Recognition using Principal Component Analysis and Support Vector Machines. *Int. J. Adv. Comput. Sci. Appl.* **2019**, *10*, 1–6. [[CrossRef](#)]
53. Bansal, A.; Mehta, K.; Arora, S. Face Recognition using PCA & LDA Algorithms. In Proceedings of the Second International Conference on ACCT, Rohtak, India, 7–8 January 2012; pp. 251–254.
54. Pearson, K. On Lines and Planes of Closest Fit to Systems of Points in Space. *Philos. Mag.* **1901**, *2*, 559–572. [[CrossRef](#)]

55. Balola, A.; Shaout, A. Hybrid Arabic Handwritten Character Recognition Using PCA and ANFIS. In Proceedings of the International Arab Conference on Information Technology, Beni-Mellal, Morocco, 6–8 December 2016.
56. Tharwat, A. Principal component analysis—A tutorial. *Int. J. Appl. Pattern Recognit.* **2016**, *3*, 197–240. [[CrossRef](#)]
57. Ojala, T.; Pietikäinen, M.; Harwood, D. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognit.* **1996**, *29*, 51–59. [[CrossRef](#)]
58. Al-Shatnawi, A.; Al-Saqqar, F.; Alhusban, S. A Holistic Model for Recognition of Handwritten Arabic Text Based on the Local Binary Pattern Technique. *Int. J. Interact. Mob. Technol.* **2020**, *14*, 20–34. [[CrossRef](#)]
59. Gardner, M.; Dorling, S. Artificial neural networks (the multilayer perceptron) a review of applications in the atmospheric sciences. *Atmos. Environ.* **1998**, *32*, 2627–2636. [[CrossRef](#)]
60. Al-Shatnawi, A.M.; Al-Saqqar, F.; Souri, A. Arabic Handwritten Word Recognition Based on Stationary Wavelet Transform Technique using Machine Learning. *ACM Trans. Asian Low-Resour. Lang. Inf. Process.* **2022**, *21*, 43. [[CrossRef](#)]
61. Prasetyo, M.L.; Wibowo, A.T.; Ridwan, M.; Milad, M.K.; Arifin, S.; Izzuddin, M.A.; Setyowati, R.D.N.; Ernawan, F. Face Recognition Using the Convolutional Neural Network for Barrier Gate System. *Int. J. Interact. Mob. Technol.* **2021**, *15*, 138–153. [[CrossRef](#)]