Noname manuscript No.
(will be inserted by the editor)

Efficient Human Face Recognition in Real-Life Applications using the Discrete Wavelet Transformation (HFRDWT)

Saddam M. Eragi *[1] · Fatma BenSaid [2] · Adel M. Alimi [2,3]

Received: date / Accepted: date

Abstract Human Face receives major attention and acquires most of the efforts of the research and studies of Machine Learning in detection and recognition. In real-life applications, the problem of quick and rapid recognition of the Human Face is always challenging researchers to come out with powerful and reliable techniques. In this paper, we proposed a new human face recognition system using the Discrete Wavelet Transformation named HFRDWT. The proposed system showed that the use of Wavelet Transformation along with the Convolutional Neural Network to represent the features of an image had significantly reduced the face recognition time, which makes it useful in real-life areas, especially in public and crowded places. The Approximation coefficient of the Discrete Wavelet Transformation played the dominant role in our system by reducing the raw image resolution to a quarter while maintaining the high level of accuracy rate that the raw image had. Results on ORL, Japanese Female Facial Expression, extended Cohn-Kanade, Labeled Faces in the Wild datasets, and our new Sudanese Labeled Faces in the Wild dataset showed that our system obtained the least recognition timing (average of 24 milliseconds for training and 8 milliseconds for testing) and acceptable high recognition rate (average of 98%) compared to the other systems.

Saddam M. Eragi
E-mail: saddam.m.eragi@regim.usf.tn

Fatma BenSaid
E-mail: fatma.bensaid@regim.usf.tn

Adel M. Alimi
E-mail: adel.alimi@regim.usf.tn

[1]College of Computer Science and Information Technology, Sudan University of Science and Technology, Sudan ·
[2]REGIM-Lab.: REsearch Groups in Intelligent Machines, University of Sfax, National Engineering School of Sfax (ENIS), BP 1173, Sfax, 3038, Tunisia ·
[3]Department of Electrical and Electronic Engineering Science, Faculty of Engineering and the Built Environment, University of Johannesburg, South Africa .

Keywords Human face recognition · Discrete Wavelet Transformation · Approximation coefficient · Convolutional Neural Network · Face recognition time

1 Introduction

When it comes to humans' recognition, faces play an an essential role in confirming a person's identity. It receives major attention and acquires most of the efforts of the researches and studies of Machine Learning (ML) in detecting and recognizing human faces. This is because faces are easily obtained in real-life conditions without users' interaction [45]. On the contrary to other biometrics, face detection and recognition do not need the collaboration of the subjects to work; in other words, properly installed systems in train stations, airports, and other public places can identify subjects without even being noticed, other biometrics do not have this ability of mass identification. As a result of that, the used techniques and methodologies and the designed recognition systems are getting more accurate and less time day after day. Discrete Wavelet Transformation (DWT) is among those powerful and reliable techniques; it achieves excellent performance and has its wide share in enriching this field of knowledge by giving well-designed solutions [44], whether on its own or by conjunction with other tools [22].

The problem of face recognition can be defined as follows: provided that still or video images are given to you, how could you identify or verify individuals in those images [1]. The solution is that easy: by extracting features and comparing patterns from a predefined database of faces. Some face recognition algorithms can identify facial features by obtaining landmarks from a face image, such as eyes, noses, and lips, and detect the distance between them in order to encode the discriminative information as a consolidated features vector [60]. These algorithms include - but are not limited to -: Local Binary Patterns (LBP) [24], Haar features Transform [50], Histogram of oriented gradients [19] (HOG), Principal Component Analysis (PCA) [27], Scale Invariant Feature Transform (SIFT) [8], and Speeded Up Robust Features (SURF) [59]. Regarding classification algorithms, one might use many of them; for example, Deep Neural Networks, Support Vector Machines (SVM), and decision trees [63].

Deep Learning techniques are among the recent advances in face detection, feature extraction, and model classification underlining the prominence of hierarchical feed-forward methods and offering a good opportunity to build models that adapt across different obstacles in the face recognition field [17, 21, 41, 49, 58, 30]. Though their performance was very efficient, recent researchers usually tended to combine two or more techniques [28, 17] to achieve an accurate recognition rate or to enhance processing timing.

This paper aimed to address the face recognition processing problem by proposing a novel rapid recognition system. The contributions of this work are as follows:

- Proposal for a new human face recognition system named Human Face Recognition system using the DWT (HFRDWT). The proposed system applied the Discrete Wavelet Transformation technique in reducing the face image resolution to a quarter which provides robust feature representation.
- HFRDWT system resolved the problem of human face recognition by minimizing the recognition time that can be an essential part of various biometric surveillance and security systems.
- Collection of a new Sudanese faces dataset called Sudanese Labeled Faces in the Wild (SuLFiW), which could contribute to the research community of the face recognition field. To the best of our knowledge, this is the only dataset of Sudanese public face images that is available for research on face detection and recognition.

Extensive experimental studies showed that our proposed system can get better performance compared to other time reduction methods.

The remainder of this paper was organized as follows. The second section presented a review of the most known human face recognition systems. The next section described in detail our proposed HFR system focusing on detection, representation, transformation, and recognition techniques used in this paper. In the fourth section, we introduced our new Sudanese Labeled Faces in the Wild (SuLFiW) and its main features and characteristics. Then, we presented our experimental results in section five to underline the robustness of our system compared to state-of-the-art works. Discussion and summarization of the obtained results were shown in section six, and finally, the future work was presented in the conclusion section.

2 Related works

In this section, we started by presenting a review of the concept of human face recognition and the systems that shared the same techniques as our proposed system. We discussed specifically the Convolutional Neural Network (CNN) and Discrete Wavelet Transformation (DWT) methodologies that were used in the human face recognition field. And finally, we gave an overview of the state-of-the-art works that aimed to reduce the computation time of face recognition in tangible measurements.

2.1 Review on Human Face Recognition Systems

Face recognition is a way of identifying or confirming an individual's identity using a predefined faces database [40]. Figure 1 shows the general architecture of any human face recognition system which is composed of five steps: image preprocessing, face detection, feature extraction, feature selection, and feature classification.

At the beginning, several techniques are used to preprocess image data, as examples, image resizing [65], converting images to gray-scale [23], and image

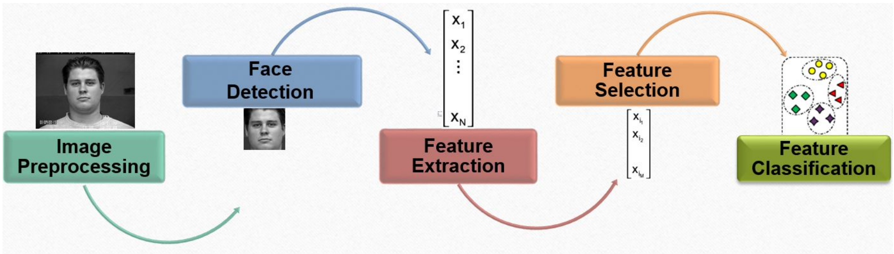

Fig. 1 General Face Recognition System

augmentation [46]. Face detection techniques focus on localizing and extracting the face region which is considered the first step to face recognition.

For feature extractor models, there are many types and techniques, such as CNN, SIFT, and STIP [18]. Among CNNs architectures, Alexnet [29] has always proven its effectiveness as a feature descriptor in ImageNet competition. Its significant features are the simplicity of the layout compared to the complicated modern architectures, very computationally efficient, and quick to train [3]. We discussed recent works on deep learning and CNNs next in this section.

Eventually, the ultimate goal of all face recognition systems is to categorize a huge number of data into different classes [35], in fewer words, classifying the features.

2.2 Convolutional Neural Networks (CNNs)

Neural networks and similar techniques are convenient for achieving high performance in term of accuracy and robustness [2]. CNNs have been applied successfully to a wide range of applications including the face recognition field, but though their performance was very efficient, recent researchers usually tended to combine two or more techniques to achieve an accurate recognition rate or to enhance processing timing as our system proposed.

Many works combined the CNNs with other techniques such as, Gabor Wavelet filters [15], Haar Cascade [32], and Boltzman machine [54] to produce the optimum accuracy by learning the visual features directly from the face image. In 2020, Fredj *et al.* [7] built a face recognition framework using aggressive data augmentation in an unconstrained environment. Another work by Deng *et al.* [12] proposed an addition Angular Margin Loss to obtain features that were highly discriminative. Their system (ArcFace) considered four types of Geodesic Discrete Constraint (GDC). On the other hand, Wu *et al.* [61] and Leng *et al.* [33] proposed another approach of feature discrimination based on spectrum correlation analysis deep network (IDICN) and Discrete Cosine Transform (DCT), respectively, to obtain the best discrimination effect.

High dimensional deep local representation is dsigned by Peng *et al.* [47] who extracted and concatenated deep features via CNN. In the same approach, Sun *et al.* [55] constructed two very deep CNN architectures (Deep ID3) for face recognition by rebuilding stacked Convolution and inception layers. An ensemble of the proposed two architectures achieved a state-of-the-art recognition rate in an unconstrained environment database. Furthermore, Alkhaldi *et al.* [48] proposed a mathematical model to select a suitable CNN architecture from the experimental results on (VggFace2) dataset.

2.3 Discrete Wavelet Transformation (DWT)

Recently, wavelet transformations have become very widespread, and new interest is rising in this topic. The main factor is the advantage of relevant feature extraction, reduction of computational time, and increased recognition accuracy rate.

The first attempts using Wavelet Transformation in Human Face Recognition were done by Garcia *et al.* [56] in 2004. The proposed method suggested using Approximation coefficients Matrices which were obtained by three levels of Wavelet decomposition and Fisher's linear discriminant for face recognition.

Lahaw *et al.* [31] suggested a combination of DWT, Independent Component Analysis (ICA), Principal Component Analysis (PCA), and Support Vector Machine (SVM) to use the most important information in the face representation, extraction, and classification.

Khan *et al.* [26] developed a method based on Particle Swarm Optimization (PSO). It utilized the features extracted from the texture and wavelet domain.

In the last two subsections, we have presented a systematic study of deep learning and Wavelet Transformation systems effects on face recognition performance. We observed that the studied models scored a 100% significant recognition rate on some datasets, but did not touch the computational complexity these models suffer from. For this reason, given proper architecture and transformation configurations, a deep learning model can perform remarkably in reducing the computational timing of face recognition systems.

2.4 Review on computational reduction techniques based HFR systems

Khalajzadeh *et al.* [25] proposed a hierarchal structure-based CNN to provide robust information processing. This method reduced the number of free parameters and the training time without compromising the high test data accuracy rate.

Contributions in this field also focused on a powerful face representation to achieve high accuracy for human recognition rate. Deep Face Recognition with a deep presentation named VIPLFace Net was used in [36] which had 10-layer deep CNN, seven Convolutional layers, and three fully-connected layers. Another contribution in this area was done by [34] random projection for feature extraction of biometrics. It projects the image matrix from high-dimensional space to low-dimensional space which led to a reduction in storage and computational cost. Shallow CNN was utilized by [42] to investigate the face region segments effect on the face recognition rate. The paper demonstrated that using the shallow CNN reduced the recognition accuracy rate but with less computational time required.

Divya *et al.* [9] headed the same way as [42] and tested the full face and the upper face by using Hidden Markov Model (HMM) and Singular Value Decomposition (SVD) Coefficients. The proposed recognition system was tested on two benchmark datasets and showed a remarkable reduction in computational complexity and memory consumption.

Table 1 shows details of the state-of-the-art works and the related HFR systems sorted by earliest first. It depicts the steps of these systems to detect, represent, and recognize human faces.

Although these works achieved good results in reducing the computational time of human face recognition, much more work on improving the dimensionality reduction is required. That's why we proposed a new system that utilizes the DWT capabilities in feature reduction to minimize the recognition processing time.

3 Proposed Human Face Recognition (HFRDWT) System

In this paper, we proposed a new HFR system using the DWT in order to reduce the processing time of human face recognition. This section described all the steps required to form the proposed system named HFRDWT and a summary of the differences between the proposed HFRDWT versions with regard to the different objectives that each version aimed to serve.

3.1 Description of the HFRDWT

We proposed a new HFR system utilizing the features and the abilities of the DWT and the CNN to design three variants: HFRDWT_V1, HFRDWT_V2, and HFRDWT_V3. Each variant has different architecture starting from face detection and ending with face classification and recognition. Figure 2 depicts the three variants' architectures of the proposed system which consists of five steps sequentially. The aim of the proposed designs of the three variants is to enhance the recognition rate and reduce the processing time. The Image Preprocessing step is not mentioned below because in the HFRDWT system we started the work with the raw facial images in the datasets. The following subsections described all the different parts of our HFRDWT system.

Table 1 The related work in detail.

Reference	System	System Steps
Arya *et al.* [4] 2019	Noise-robust low-resolution SIFT	-Extraction: SIFT -Classification: Euclidean distance
Ben *et al.* [7] 2021	CNN + aggressive data augmentation	-Detection: multi-task cascaded CNN -Extraction: CNN -Classification: Softmax + center loss function
Choudhary *et al.* [9] 2021	Singular Value Decomposition (SVD) + Hidden Markov Model	-Detection: Viola and Jones -Extraction: SVD -Classification: Hidden Markov Model
Cuculo *et al.* [10] 2019	Sparsity-Driven Sub-Dictionary Learning Using Deep Features (SSLD)	-Extraction: VGGface CNN -Selection: LDA -Classification: k-LiMapS Classifier
Divya *et al.* [14] 2019	DWT + Statistical Features	-Transformation: DWT -Extraction: Statistical Features -Classification: ANN
Dumitrescu *et al.* [15] 2019	Neural Networks + Global Gabor Features	-Extraction: Gabor Features -Classification: ANN + KNN
Ghazal *et al.* [16] 2020	Curvelets, Invariant moments features and SVM	-Extraction: invariant moments features -Classification: SVM
Khalajzadeh *et al.* [25] 2013	Hierarchical Structure CNN	-Extraction: CNN -Classification: Softmax
Khan *et al.* [26] 2018	LBP + DWT + PSO + DFT	-Detection: VIPL Toolkit -Extraction: Deep CNN -Classification: Softmax
Lahaw *et al.* [31] 2018	DWT+PCA+SVM	-Preprocessing: 2D DWT -Extraction: PCA, ICA or LDA -Classification: SVM
Liu *et al.* [36] 2017	VIPLFaceNet	-Detection: Viola and Jones -Transformation: LBP+DWT -Selection: PSO -Classification: Euclidean distance classifier
Melaugh *et al.* [42] 2019	CNN	-Detection: Viola and Jones -Extraction: CNN -Classification: Softmax
Mulyono *et al.* [43] 2019	PCA-eigenface	-Extraction: PCA -Classification: Euclidean distance
Senthilkumar *et al.* [52] 2020	Bag Of Visual Words (BOVW)	-Extraction: SIFT -Classification: SVM
Sun *et al.* [53] 2020	Fuzzy Convex-Concave Partition + Local Gradient Number Pattern (FCCP_LGNP)	-Extraction: LGNP -Classification: kNN
Xu *et al.* [62] 2018	Projection Neural Network-Based iterative method (PNNBIM)	-Detection: multitask cascaded CNNs -Extraction: Facenet -Selection: PCA
Yee *et al.* [64] 2020	Laplacian Completed Local Ternary Pattern	-Extraction: Laplacian filter (CLTP) -Classification: LapCLTP
Zhang *et al.* [66] 2018	Face Patches strategy	-Detection: funnel-structured (FuSt) cascade schema -Extraction: CNN -Classification: CNN

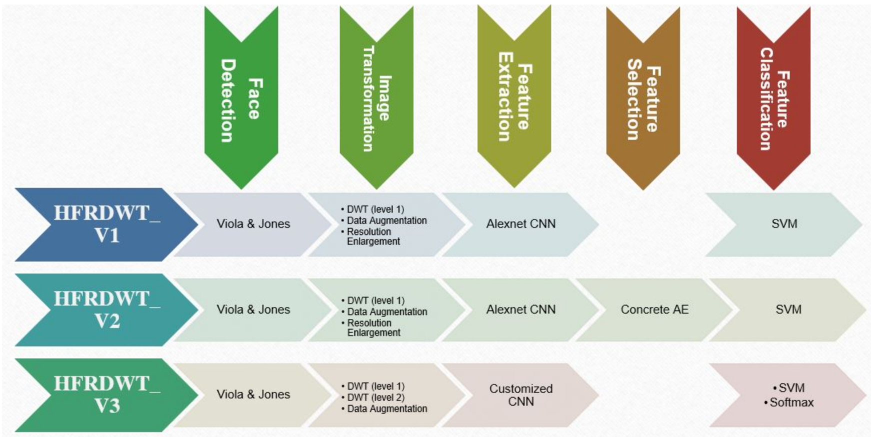

Fig. 2 Overview of the HFRDWT system

3.1.1 Face Detection

Firstly, we used Viola and Jones technique [57] in the proposed system, as a face detector. It generates images that contain only the person's face and removes all unnecessary appendages to facilitate the image recognition process. Figure 3 presents the implementation of Viola and Jones on two sample faces from LFW and SuLFiW faces datasets.

Fig. 3 Implementation of Viola and Jones on two sample faces: (a) LFW, (b) SuLFiW faces datasets

3.1.2 Image Transformation

For the purpose of filtering and compressing the detected faces images, Discrete Wavelet Transform (DWT) was used in the proposed HFRDWT system as an analog filter that captures both frequency and location information [5] showing great efficiency in such tasks. Moreover, Applying the DWT as a descriptor reduced the resolution of the detected faces images to a quarter by dividing it into four coefficients, Approximation, Diagonal, Horizontal, and Vertical [11] which can be expressed by the following equation:

$$F(a,b) = \int_{-\infty}^{\infty} f(x)\psi^*_{(a,b)}(x)dx \tag{1}$$

where the $*$ is the complex conjugate symbol and ψ is a function that is chosen arbitrarily if it obeys certain rule [13]. Hence, DWT produced four more image datasets, one for each coefficient of DWT. Figure 4 demonstrates a sample of the resulting images after applying DWT on the detected faces of JAFFE dataset.

The easiest and most common way to reduce overfitting on images is to enlarge the dataset artificially using a label preservation transformation (data augmentation) that allows to produce transformed images from the original image with very little computation by flipping and mirroring images horizontally. Another way of transformation technique used in this work is the resolution enlargement, the raw and the four DWT image datasets were scaled to $227 * 227$ to fit the resolution of the pre-trained Alexnet Fully Connected layer.

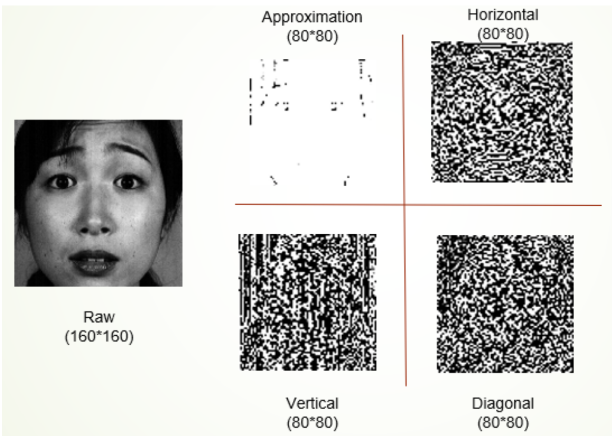

Fig. 4 Applying DWT on a sample of JAFFE dataset

3.1.3 Feature Extraction

The raw images as well as the four wavelet-transformed images were fed to two types of CNNs to produce five feature vectors. The first is the pre-trained Alexnet and the second one is a customized CNN with four convolutional layers at the beginning and two fully connected layers at the end. The contributions of this paper are not depending on the selection of the best CNN architecture, instead, the DWT is the major factor that has the dominant role in reducing the processing timing. Therefore, the customized CNN was selected due to its simplicity and efficiency in achieving good results in face recognition rate.

3.1.4 Feature Selection

In this paper, selecting the most important features for the face recognition process was done by using the Auto Encoder technique, specifically the Concrete Autoencoder as described in [6]. The Concrete Autoencoder is a comprehensive differentiable method for selecting global features, which effectively selects part of the most useful features and at the same time learns a neural network to reconstruct the input data from those selected features.

3.1.5 Feature Classification

For the purpose of recognizing human face images, two techniques were used in the HFRDWT system: SVM and Softmax to facilitate the provision of a decision weight on each tested image of the five image datasets. We used other tools to enhance the accuracy rate: Top-k accuracy and Majority voting. Top-k accuracy method computes the number of times where the correct label is among the top k of the classifier distribution. Its goal is to produce improved results or reduce the likelihood of an unfortunate selection of a poor one. A majority Voting is an integrated machine learning model that combines predictions from several other models. In the case of classification, the predictions for each label are aggregated and the label with the majority vote is predicted.

3.2 Summary of the different HFRDWT architectures

In this section, we presented a comparison between the different HFRDWT architectures that were proposed in this paper and summarized in Table 2. Some techniques are common and have been used for all architectures of the HFRDWT system including Viola and Jones, the DWT (level 1), Data Augmentation, and the SVM classifier. Some other techniques are specific for just one architecture or two, such as, the Concrete Autoencoder is just utilized in HFRDWT_V2; while the Resolution Enlargement and Alexnet CNN are utilized in the HFRDWT_V1 and HFRDWT_V2; and finally, the DWT (level 2), the customized CNN, and Softmax classifier are used specifically in HFRDWT_V3.

Table 2 HFRDWT variants.

Architectures	**HFRDWT_V1**	**HFRDWT_V2**	**HFRDWT_V3**
Design Objective	Recognition Rate Enhancement		Processing Time Reduction
System Steps:			
1. Face Detection	Viola and Jones		
2. Image Transformation	DWT (level 1)		DWT (level 1) and (level 2)
	Data Augmentation		
	Resolution Enlargement		No Resolution Enlargement
3. Feature Extraction	Alexnet CNN		Customized CNN
4. Feature Selection	None	Concrete AE	None
5. Feature Classification	SVM		SVM & Softmax
Accuracy Enhancement	Majority Voting		top-k

4 Sudanese Labeled Faces in the Wild (SuLFiW) dataset

This section described the process of creating a new Sudanese dataset for face expression detection and recognition purposes named SuLFiW.

Sudanese faces are characterized by their wide variety of skin colors, shapes, and ethnicities, which makes a great deal of challenge in detection and recognition systems.

The objective of SuLFiW is to bridge this gap and provide a consistent and comprehensive dataset for significant practical analysis. SuLFiW dataset will focus mainly on facial expression recognition which is also useful for identity recognition via face detection and it will be helpful for gender classification systems too. On the other hand, we will not take into consideration any illumination or model posing change.

Fig. 5 Sample images from SuLFiW dataset

SuLFiW dataset includes a set of frontal human faces of Sudanese people who are representing many categories in Sudanese Society which were collected from Google Images as shown in Figure 5. Table 3 demonstrated the different categories that comprise the SuLFiW dataset. It consisted of 600 persons of both genders (with the majority of Males models) aged between (20-90) years old. Seven standard facial expressions existed in the SuLFiW dataset, these expressions were: Neutral, Anger, Disgust, Fear, Happy, Sad, and Surprise. The image itself consisted of all head parts. We followed the same construction and labeling process of LFW dataset creators in [20].

The SuLFiW advantages are:

- The first Sudanese dataset to provide facial expressions.
- The first Sudanese dataset to provide gender classification.
- The first Sudanese dataset to provide such a wide variety of categories representing all the Sudanese society.
- Non-restricted environment for analysis.

Table 3 SuLFiW dataset categories.

Number	**Category**	**No. of models**
1	Politicians	88
2	Scientists	18
3	Religious	25
4	Sportsmen	54
5	Poets	81
6	Writers	49
7	Businessmen	12
8	Actors	53
9	Media	74
10	Artists	146
Total		**600**

5 Experimental results

In this section, we presented our experimental testbed, compared systems, and performance evaluation parameters. Then we described all the experiments that were conducted in the HFRDWT system in order to verify its efficiency compared to other related works.

5.1 Experimental testbed and Compared systems

Experimental testbed: We conducted several experiments on a variety of benchmark datasets. The HFRDWT system was evaluated by testing it on various size benchmark frontal faces datasets and on the SuLFiW dataset.

ORL [51]: The ORL Database of Faces is a set of face images taken between 1992 and 1994. The database was used for a face recognition project carried out in the Cambridge University Engineering Department. All the images were taken against a homogeneous dark background with the subjects in an upright frontal position.

JAFFE [38, 39]: The JAFFE database consists of 123 images with size 256*256 posed by 10 Japanese females and photographed with constant illumination and unified background. From the JAFFE database, we considered all 213 images.

LFW [20]: The unconstrained environment is considered harder than many other constrained datasets. LFW has become the standard benchmark for facial recognition in wild performance evaluation in the last decade. Extensive work has been done to leverage the limits of accuracy on it, however, the impressive progress has been obtained when using modern deep learning techniques. For data augmentation purpose, we produced more images by horizontally flipping and mirroring images in the training set.

CK+ [37]: This database consists of 123 models aged between (18-50) years of which the females are the most with 69% and with different ethnic groups. Those set of images is digitized to 640*490 pixels with 8-bit depth. We took all 123 models

SuLFiW: The New Sudanese dataset for face recognition, expression detection, and gender classification as presented in Section 4. SuLFiW consists of 600 persons and 5400+ gray and color images which are realistic social media photos that contain persons in diverse poses, activities, and events.

Table 4 shows the list of datasets that were used to evaluate the HFRDWT system with more details on our experimental testbed. The datasets have been ordered from the oldest up to the latest with a clear vision of the number of images used for the training and testing sets. Figure 6 displays samples from all the datasets used in this paper.

Table 4 Benchmark datasets.

Dataset	#Subjects	Size	Training Sz	Testing Sz	Type	Resolution	Year
ORL	41	410	328	82	Gray	$70 * 80$	1994
JAFFE	10	213	173	40	Gray	$256 * 256$	1997
LFW	5,749	13,233	10,586	2,647	Color	$250 * 250$	2007
CK+	121	10,260	8,254	2,006	Gray+Color	$640 * 490$	2010
SuLFiW	600	5,448	4,358	1,090	Gray+Color	$250 * 250$	2020

Compared Systems: In our experiments, the HFRDWT system was compared to some state-of-the-art HFR systems which the aim was to reduce hu-

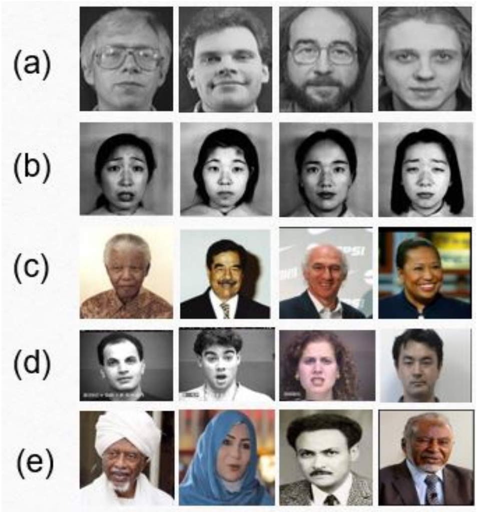

Fig. 6 Sample images from all datasets: (a) ORL, (b) JAFFE, (c) LFW, (d) CK+, (e) SuLFiW faces datasets

man face time processing [25, 36, 42, 9]. We compared the HFRDWT also with other previous works that used the DWT technique [26, 31, 14]. Table 5 shows details of the HFR compared systems. The objectives of all these systems were reducing human processing time or/and enhancing recognition rate.

Performance Evaluation: Inclusive practical experiments have been conducted to discover the most favorable human face recognition system that yields the best results.

The first two HFRDWT architectures started with cropping the face images to 227∗227 in order to match the prerequisite resolution of Alexnet's Fully Connected layer (FC), and if the input image is grayscale, it was converted to RGB by replicating the single channel to obtain a 3-channel RGB image. The third architecture cropped the faces but kept the same resolution of the raw images. Then applying the Discrete Wavelet Transformation (DWT) to the detected images of faces is necessary to produce four more image datasets, one for each coefficient, that is, Approximation, Diagonal, Horizontal, and Vertical.

The CNN is strong enough to extract sufficient information by using only row pixel data. Therefore, we employed CNN as our base network and a fully connected layer is then used to extract features from the input facial images.

The (DWT+CNN) hybrid system made a significant improvement by first applying encoder-decoder on the features with different values for learning rate, regularization, and hidden neurons, and then feeding the result on labeled data to an SVM or Softmax classifier. This idea was inspired by [14] who suggested a hybrid model to improve the performance of CNN on some benchmark Datasets.

Table 5 HFR Compared Systems.

Reference	System	Dataset/s
Choudhary *et al.* [9] 2021	Singular Value Decomposition (SVD) + Hidden Markov Model	ORL CK+
Divya *et al.* [14] 2019	DWT + Statistical Features	ORL JAFFE
Khalajzadeh *et al.* [25] 2013	Hierarchical Structure CNN	ORL JAFFE
Khan *et al.* [26] 2018	LBP + DWT + PSO + DFT	JAFFE CK+
Lahaw *et al.* [31] 2018	DWT+PCA+SVM	ORL
Liu *et al.* [36] 2017	VIPLFaceNet	LFW
Melaugh *et al.* [42] 2019	CNN	JAFFE CK+

A multi-class SVM and Softmax were trained using 80% of the images and the remaining images for testing. The training and testing process was repeated 10 times giving a 10-fold cross-validation recognition rate. Finally, measuring the average time for training and testing for each dataset were calculated and recorded. To gain better performance in recognizing faces, majority voting and top-k accuracy techniques were used in the system. When using such powerful enhancement techniques, the face recognition rates were up to 100% in some benchmark datasets.

It is worth mentioning that the first two architectures of the HFRDWT system had achieved high face recognition rates on the benchmark datasets. On the other hand, the third architecture (HFRDWT_V3) reduced the features dimensionality which leads subsequently to a reduction of the training and testing process while maintaining good recognition rates, which outperformed the state-of-the-art systems with different structures and data augmentation techniques.

We took the advantage of the deep learning features in Python to implement the aforementioned system.

5.2 Experiment I: Evaluation of the proposed architectures on JAFFE dataset

In this part, we evaluated the HFRDWT system with the three variants on the JAFFE dataset. Figure 7 shows a comparison between the accuracy rates of the different proposed architectures of the HFRDWT system on the five image sets (Raw, Approximation, Diagonal, and Vertical sets) of the JAFFE dataset, and highlights that the third variant of the HFRDWT system (HFRDWT_V3) along with the DWT Approximation coefficient and the Softmax classifier had the best performance compared to the other variants. We plotted the Receiver Operating Characteristic (ROC) Curve by calculating the False Acceptance Rate (FAR) and False Rejection Rate (FRR) for many different thresholds to assess the efficiency of the third variant. Figure 8 showed the power of the HFRDWT_V3 in classifying and recognizing the data comparing the other variants. The effect of the DWT technique on the processing time is very remarkable in Figure 9, especially when calculating the training and the testing time separately. The training and testing time were reduced by 26% and 12% respectively when utilizing the DWT Approximation coefficient. That is why we adopted the HFRDWT_V3 system to test it on the remaining datasets: ORL, CK+, LFW, and SuLFiW.

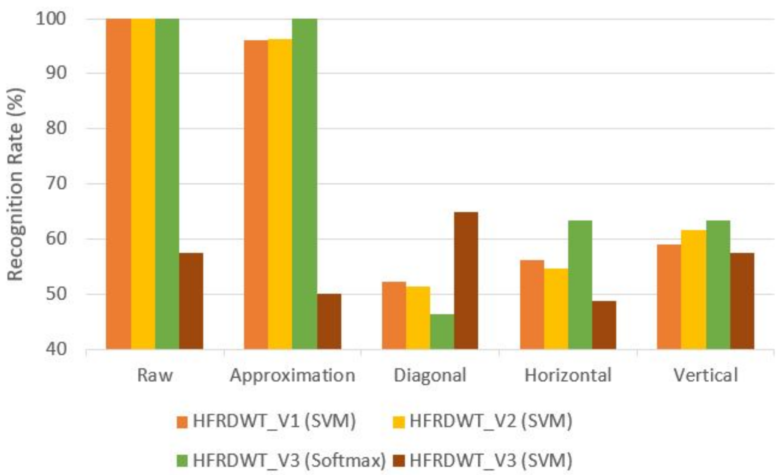

Fig. 7 Evaluation of the Recognition Rates of all HFRDWT Architectures on JAFFE dataset

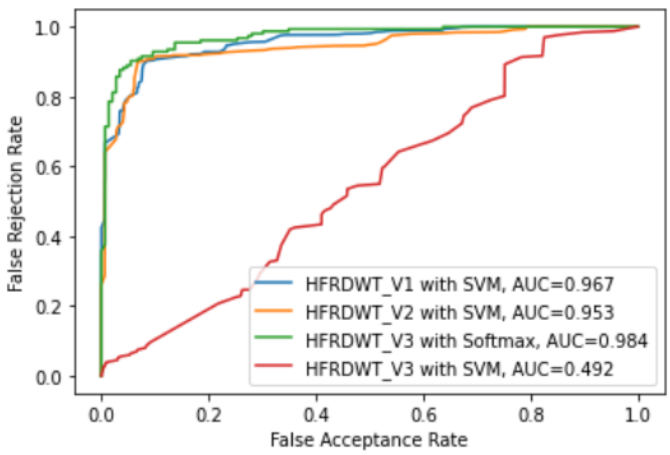

Fig. 8 The ROC curve and AUC of HFRDWT_V1, HFRDWT_V2, and HFRDWT_V3 on JAFFE dataset

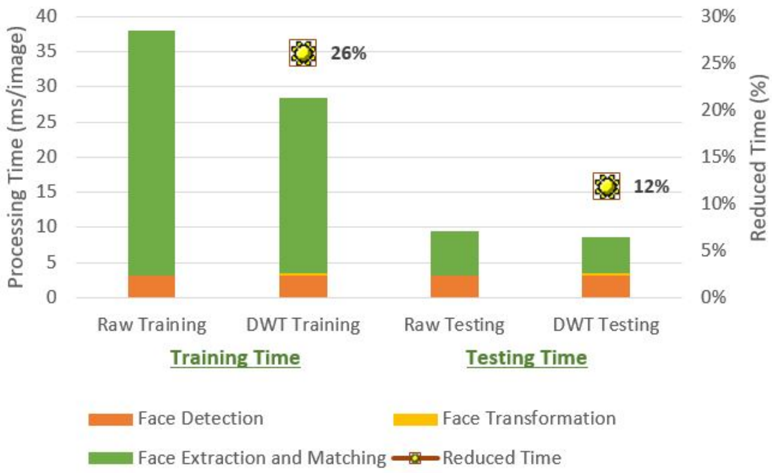

Fig. 9 The Processing Time of HFRDWT_V3 on JAFFE dataset

5.3 Experiment II: Evaluation of HFRDWT_V3 on JAFFE, ORL, CK+, LFW, and SuLFiW datasets compared with other state-of-the-art works

In this part, we implemented the HFRDWT_V3 system on all datasets in the testbed, i.e. JAFFE, ORL, CK+, LFW, and SuLFiW datasets. The application of HFRDWT_V3 was evaluated on two levels of the DWT in all the datasets as shown in Figure 10 and Figure 11. In spite of the fact that the resolution in the first DWT (Wav (1)) is minimized, the recognition rates were almost close to the recognition rates of the raw face images. On the other hand, the second level DWT (Wav (2)) recognition rate was very low compared to the raw and the first DWT (Wav (1)) recognition rates. Contrariwise, Figure 12 depicted the efficiency of (Wav (2)) in reducing the processing time of training and testing for all datasets compared to (Wav (1)).

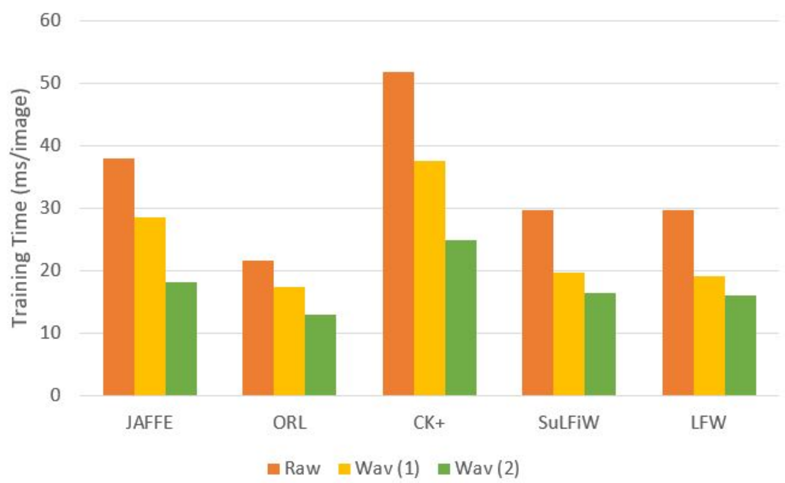

Fig. 10 Training Time of HFRDWT_V3 with two DWT levels on JAFFE, ORL, CK+, SuLFiW, and LFW datasets

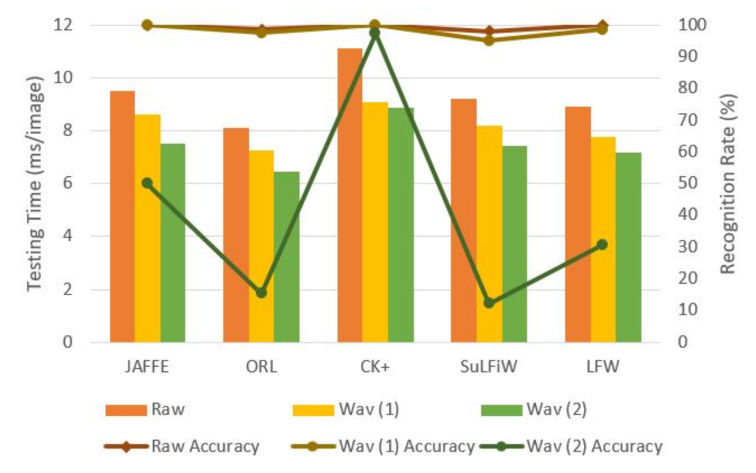

Fig. 11 Testing Time and Accuracy Rate of HFRDWT_V3 with two DWT levels on JAFFE, ORL, CK+, SuLFiW, and LFW datasets

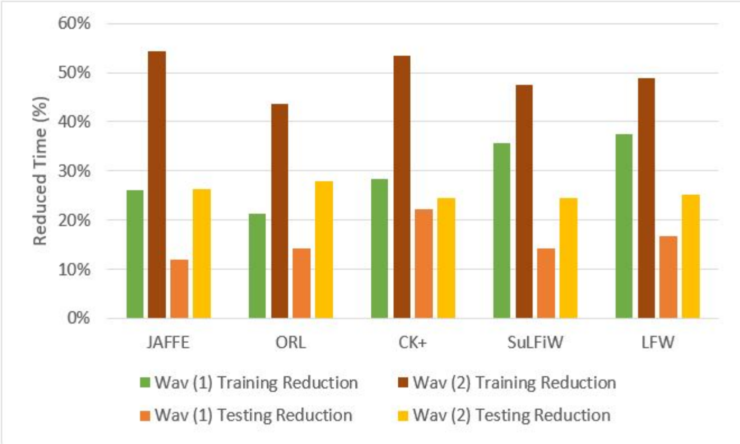

Fig. 12 The Reduction in processing Time of HFRDWT_V3 with two DWT levels on JAFFE, ORL, CK+, LFW, and SuLFiW datasets

Many works had discussed the issue of enhancing the Face Recognition rate and the processing time related to that. Table 6 verified the efficiency of our HFRDWT_V3 system in recognizing the human face images and in minimizing the computational time needed to process and recognize these images.

Table 6 Comparative study between HFRDWT_V3 and 7 HFR systems on JAFFE, ORL, CK+, and LFW.
AR: Accuracy Rate; TrnT: Training Time; TesT: Testing Time

	JAFFE			ORL			CK+			LFW		
System	**AR (%)**	**TrnT (ms/img)**	**TesT (ms/img)**	**AR (%)**	**TrnT (ms/img)**	**TesT (ms/img)**	**AR (%)**	**TrnT (ms/img)**	**TesT (ms/img)**	**AR (%)**	**TrnT (ms/img)**	**TesT (ms/img)**
Choudhary *et al.* [9] 2021				99.25	320	150	100.00	500	60			
Divya *et al.* [14] 2019	**100.00**			**99.50**								
Khalajzadeh *et al.* [25] 2013	**100.00**	193		97.50	292							
Khan *et al.* [26] 2018	98.80						98.60					
Lahaw *et al.* [31] 2018	**100.00**	193		96.00	292							
Liu *et al.* [36] 2017										**98.60**	150	
Melaugh *et al.* [42] 2019	76.56	91					87.32	161				
HFRDWT 2022	**100.00**	**28**	**9**	97.42	**17**	**7**	**100.00**	**38**	**9**	98.48	**19**	**8**

6 Discussion

Human face Recognition is a challenging field. Each year new systems are developed, published, and perhaps commercially distributed. In our work, we have shown that the using of a simple customized CNN with the Approximation Wavelet Transformation could lead to a huge enhancement in the process of the recognition time. This enhancement which reached to 40% of the raw images timing in some datasets is vital for real-life applications that need every millisecond of processing time. The enhancement in the processing time which is supported by the improvement in the recognition rate was verified by plotting the ROC curve. Another finding in this paper is the non-significance of the second level of the DWT in the HFR rate although it saves more than half of the time. One of the reasons of this degradation might be the very low resolution these images had, and the other reason was the noise that occurred when transforming an already Approximation transformed images that wiped the raw images.

7 Conclusion

Face recognition is a difficult task in computer vision; most systems need a large number of labeled examples per class for training purposes. Therefore, constructing a robust invariant system is an essential step toward finding a solution to face recognition problems. The HFRDWT system resulted in an efficient model for feature representation and classification using a few number of labeled examples even in an unconstrained environment. It is noteworthy that though we have paid some cost in the training phase to extract and represent the features, that leads to a fast, accurate, and integrated model for recognition. Taking advantage of the combination of deep learning techniques, hand-engineered features, and ensemble methods, extensive experiments have been carried out on four various sizes well-known benchmark faces datasets, namely, ORL, JAFFE, CK+, and LFW. The experiments demonstrated that our hybrid system improved the performance of extracting facial characteristics and the efficiency of face classification, and achieved state-of-the-art performance under unconstrained environments, pose variations, and heavy occlusions. With regard to the recognition rate, the proposed system obtained 100% on JAFFE and CK+ while getting very good and competitive results on the others. On the other hand, the training and testing time of the images belonging to these datasets are very low compared to the other state-of-the-art works with the average of (24) and (8) milliseconds respectively.

We have also introduced the new Sudanese Faces in the Wild dataset (SuLFiW), which is the first dataset to cover the Sudanese faces' expressions and genders in the wild, its objective is to bridge that gap and provide a consistent and comprehensive dataset for significant practical analysis.

For this approach, we have trained all the models from scratch, which leads to an undesirable increment in processing time. Therefore, more enhancements need to be carried out in the system to avoid this limitation; a good example of it is using the Transfer Learning technique. Moreover, in the future work, we are considering the extension of our approach to cover color images as well as grayscale images. Another extension would be the implementation of the system on facial emotion recognition, age, gender prediction, and other demographic soft information recognition whose performance suffers from uneven distribution of training data. This will permit us to explore the efficiency of our hybrid system in tackling much more challenging tasks in the visual recognition field.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The research leading to these results has received funding from the Tunisian Ministry of Higher Education and Scientific Research under the grant agreement number LR11ES48. Appreciation sound also goes to Mrs. Rawaa Wadie, without her help SuLFiW faces dataset would not see the light.

References

1. Adjabi I, Ouahabi A, Benzaoui A, Jacques S (2021) Multi-block color-binarized statistical images for single-sample face recognition. Sensors 21(3):728
2. Ali W, Tian W, Din SU, Iradukunda D, Khan AA (2021) Classical and modern face recognition approaches: a complete review. Multimedia Tools and Applications 80(3):4825–4880
3. Arsenovic M, Karanovic M, Sladojevic S, Anderla A, Stefanovic D (2019) Solving current limitations of deep learning based approaches for plant disease detection. Symmetry 11(7):939
4. Arya K, Rajput SS, Upadhyay S (2019) Noise-robust low-resolution face recognition using sift features. In: Computational Intelligence: Theories, Applications and Future Directions-Volume II, Springer, pp 645–655
5. Aslam W, Khurshid K, Khan AA (2016) Optimized image scaling using dwt and different interpolation techniques. International Journal of Advanced Computer Science and Applications 7(6):294–300
6. Balın MF, Abid A, Zou J (2019) Concrete autoencoders: Differentiable feature selection and reconstruction. In: International conference on machine learning, PMLR, pp 444–453

7. Ben Fredj H, Bouguezzi S, Souani C (2021) Face recognition in unconstrained environment with cnn. The Visual Computer 37(2):217–226
8. Chen Y, Xu W, Zuo J, Yang K (2019) The fire recognition algorithm using dynamic feature fusion and iv-svm classifier. Cluster Computing 22(3):7665–7675
9. Choudhary P, Tripathi A, Singh AK, Vashist PC (2021) Upper half face recognition using hidden markov model and singular value decomposition coefficients. In: Advances in Computational Intelligence and Communication Technology, Springer, pp 549–561
10. Cuculo V, D'Amelio A, Grossi G, Lanzarotti R, Lin J (2019) Robust single-sample face recognition by sparsity-driven sub-dictionary learning using deep features. Sensors 19(1):146
11. Dashtdar M, Dashtdar M (2019) Fault location in the transmission network using a discrete wavelet transform. American Journal of Electrical and Computer Engineering 3(1):30–37
12. Deng J, Guo J, Xue N, Zafeiriou S (2019) Arcface: Additive angular margin loss for deep face recognition. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 4690–4699
13. Deokar S, Waghmare L (2014) Integrated dwt–fft approach for detection and classification of power quality disturbances. International Journal of Electrical Power & Energy Systems 61:594–605
14. Divya A, Raja K, Venugopal K (2019) Sorting pixels based face recognition using discrete wavelet transform and statistical features. In: 2019 3rd International Conference on Imaging, Signal Processing and Communication (ICISPC), IEEE, pp 150–154
15. Dumitrescu CM, Dumitrache I (2019) Combining neural networks and global gabor features in a hybrid face recognition system. In: 2019 22nd International Conference on Control Systems and Computer Science (CSCS), IEEE, pp 216–222
16. Ghazal MT, Abdullah K (2020) Face recognition based on curvelets, invariant moments features and svm. Telkomnika 18(2):733–739
17. Guo G, Zhang N (2019) A survey on deep learning based face recognition. Computer vision and image understanding 189:102805
18. He Q, He B, Zhang Y, Fang H (2019) Multimedia based fast face recognition algorithm of speed up robust features. Multimedia Tools and Applications 78(17):24035–24045
19. Hosseini-Fard E, Roshandel-Kahoo A, Soleimani-Monfared M, Khayer K, Ahmadi-Fard AR (2022) Automatic seismic image segmentation by introducing a novel strategy in histogram of oriented gradients. Journal of Petroleum Science and Engineering 209:109971
20. Huang GB, Ramesh M, Berg T, Learned-Miller E (2007) Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Tech. Rep. 07-49, University of Massachusetts, Amherst
21. Huang Y, Wang Y, Tai Y, Liu X, Shen P, Li S, Li J, Huang F (2020) Curricularface: adaptive curriculum learning loss for deep face recognition. In: proceedings of the IEEE/CVF conference on computer vision and pattern

recognition, pp 5901–5910
22. Jarraya I, BenSaid F, Ouarda W, Pal U, Alimi AM (2022) A new convolutional neural network based on a sparse convolutional layer for animal face detection. Multimedia Tools and Applications pp 1–34
23. Karolin M, Meyyappan T (2022) Visual secret share creation with grayscale image converted to rgb images using zigzag scanning algorithm. In: Proceedings of the International Conference on Paradigms of Communication, Computing and Data Sciences, Springer, pp 735–742
24. Kayhan N, Fekri-Ershad S (2021) Content based image retrieval based on weighted fusion of texture and color features derived from modified local binary patterns and local neighborhood difference patterns. Multimedia Tools and Applications 80(21):32763–32790
25. Khalajzadeh H, Mansouri M, Teshnehlab M (2013) Hierarchical structure based convolutional neural network for face recognition. International Journal of Computational Intelligence and Applications 12(03):1350018
26. Khan SA, Ishtiaq M, Nazir M, Shaheen M (2018) Face recognition under varying expressions and illumination using particle swarm optimization. Journal of computational science 28:94–100
27. Kherif F, Latypova A (2020) Principal component analysis. In: Machine Learning, Elsevier, pp 209–225
28. Kortli Y, Jridi M, Al Falou A, Atri M (2020) Face recognition systems: A survey. Sensors 20(2):342
29. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems 25
30. Kumar A, Kaur A, Kumar M (2019) Face detection techniques: a review. Artificial Intelligence Review 52(2):927–948
31. Lahaw ZB, Essaidani D, Seddik H (2018) Robust face recognition approaches using pca, ica, lda based on dwt, and svm algorithms. In: 2018 41st International Conference on Telecommunications and Signal Processing (TSP), IEEE, pp 1–5
32. Lal PV, Srilakshmi U, Venkateswarlu D (2022) Face recognition using deep learning xception cnn method. Journal of Theoretical and Applied Information Technology 100(2)
33. Leng L, Zhang J, Xu J, Khan MK, Alghathbar K (2010) Dynamic weighted discrimination power analysis in dct domain for face and palmprint recognition. In: 2010 international conference on information and communication technology convergence (ICTC), IEEE, pp 467–471
34. Leng L, Zhang J, Chen G, Khan MK, Alghathbar K (2011) Two-directional two-dimensional random projection and its variations for face and palmprint recognition. In: Computational Science and Its Applications-ICCSA 2011: International Conference, Santander, Spain, June 20-23, 2011. Proceedings, Part V 11, Springer, pp 458–470
35. Lin SH, Kung SY, Lin LJ (1997) Face recognition/detection by probabilistic decision-based neural network. IEEE transactions on neural networks 8(1):114–132

36. Liu X, Kan M, Wu W, Shan S, Chen X (2017) Viplfacenet: an open source deep face recognition sdk. Frontiers of Computer Science 11(2):208–218
37. Lucey P, Cohn JF, Kanade T, Saragih J, Ambadar Z, Matthews I (2010) The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In: 2010 ieee computer society conference on computer vision and pattern recognition-workshops, IEEE, pp 94–101
38. Lyons M, Akamatsu S, Kamachi M, Gyoba J (1998) Coding facial expressions with gabor wavelets. In: Proceedings Third IEEE international conference on automatic face and gesture recognition, IEEE, pp 200–205
39. Lyons MJ (2021) " excavating ai" re-excavated: Debunking a fallacious account of the jaffe dataset. arXiv preprint arXiv:210713998
40. Mady H, Hilles SM (2018) Face recognition and detection using random forest and combination of lbp and hog features. In: 2018 international conference on smart computing and electronic enterprise (ICSCEE), IEEE, pp 1–7
41. Masi I, Wu Y, Hassner T, Natarajan P (2018) Deep face recognition: A survey. In: 2018 31st SIBGRAPI conference on graphics, patterns and images (SIBGRAPI), IEEE, pp 471–478
42. Melaugh R, Siddique N, Coleman S, Yogarajah P (2019) Facial expression recognition on partial facial sections. In: 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA), IEEE, pp 193–197
43. Mulyono IUW, Susanto A, Rachmawanto EH, Fahmi A, et al. (2019) Performance analysis of face recognition using eigenface approach. In: 2019 International Seminar on Application for Technology of Information and Communication (iSemantic), IEEE, pp 1–5
44. Nath S, Naskar R (2021) Automated image splicing detection using deep cnn-learned features and ann-based classifier. Signal, Image and Video Processing 15(7):1601–1608
45. Palma TA, Garcia-Marques L (2021) Does repetition always make perfect? differential effects of repetition on learning of own-race and other-race faces. Basic and Applied Social Psychology 43(2):90–109
46. Pei Z, Xu H, Zhang Y, Guo M, Yang YH (2019) Face recognition via deep learning using data augmentation based on orthogonal experiments. Electronics 8(10):1088
47. Peng C, Wang N, Li J, Gao X (2019) Re-ranking high-dimensional deep local representation for nir-vis face recognition. IEEE Transactions on Image Processing 28(9):4553–4565
48. Qusay AK, Motlak H, Al-Libawy H (2020) A new rule to constrain convolution neural network architecture in face recognition system. In: Journal of Physics: Conference Series, IOP Publishing, vol 1530, p 012125
49. Ranjan R, Sankaranarayanan S, Bansal A, Bodla N, Chen JC, Patel VM, Castillo CD, Chellappa R (2018) Deep learning for understanding faces: Machines may be just as good, or better, than humans. IEEE Signal Processing Magazine 35(1):66–83

50. Raveendra K, Ravi J (2021) Performance evaluation of face recognition system by concatenation of spatial and transformation domain features. International Journal of Computer Network and Information Security 13(1):47–60
51. Samaria FS, Harter AC (1994) Parameterisation of a stochastic model for human face identification. In: Proceedings of 1994 IEEE workshop on applications of computer vision, IEEE, pp 138–142
52. Senthilkumar R, Gnanamurthy RK (2020) A proposed method for the improvement in biometric facial image recognition using document-based classification. The Journal of Supercomputing 76(6):4476–4494
53. Sun J, Lv Y, Tang C, Sima H, Wu X (2020) Face recognition based on local gradient number pattern and fuzzy convex-concave partition. IEEE Access 8:35777–35791
54. Sun Y, Wang X, Tang X (2013) Hybrid deep learning for face verification. In: Proceedings of the IEEE international conference on computer vision, pp 1489–1496
55. Sun Y, Liang D, Wang X, Tang X (2015) Deepid3: Face recognition with very deep neural networks. arXiv preprint arXiv:150200873
56. Tenllado C, García C, Prieto M, Piñuel L, Tirado F (2004) Exploiting multilevel parallelism within modern microprocessors: Dwt as a case study. In: International Conference on High Performance Computing for Computational Science, Springer, pp 556–568
57. Viola P, Jones M (2001) 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, Ieee, vol 1, pp I–I
58. Wang M, Deng W (2021) Deep face recognition: A survey. Neurocomputing 429:215–244
59. Wang R, Shi Y, Cao W (2019) Ga-surf: A new speeded-up robust feature extraction algorithm for multispectral images based on geometric algebra. Pattern Recognition Letters 127:11–17
60. Wood E, Baltrušaitis T, Hewitt C, Dziadzio S, Cashman TJ, Shotton J (2021) Fake it till you make it: Face analysis in the wild using synthetic data alone. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp 3681–3691
61. Wu F, Jing XY, Dong X, Hu R, Yue D, Wang L, Ji YM, Wang R, Chen G (2018) Intraspectrum discrimination and interspectrum correlation analysis deep network for multispectral face recognition. IEEE transactions on cybernetics 50(3):1009–1022
62. Xu B, Liu Q, Huang T (2018) A discrete-time projection neural network for sparse signal reconstruction with application to face recognition. IEEE transactions on neural networks and learning systems 30(1):151–162
63. Yaman MA, Subasi A, Rattay F (2018) Comparison of random subspace and voting ensemble machine learning methods for face recognition. Symmetry 10(11):651

64. Yee SY, Rassem TH, Mohammed MF, Awang S (2020) Face recognition using laplacian completed local ternary pattern (lapcltp). In: Advances in electronics engineering, Springer, pp 315–327
65. Zeng D, Zhao F, Ge S, Shen W (2019) Fast cascade face detection with pyramid network. Pattern Recognition Letters 119:180–186
66. Zhang Y, Shang K, Wang J, Li N, Zhang MM (2018) Patch strategy for deep face recognition. IET Image Processing 12(5):819–825