Computers, Materials & Continua
DOI:10.32604/cmc.2021.015466
Article

Tech Science Press

Real-Time Multimodal Biometric Authentication of Human Using Face Feature Analysis

Rohit Srivastava[1], Ravi Tomar[1], Ashutosh Sharma[2], Gaurav Dhiman[3], Naveen Chilamkurti[4] and Byung-Gyu Kim[5,*]

[1]University of Petroleum and Energy Studies, Dehradun, India
[2]Lovely Professional University, India
[3]Government Bikram College of Commerce, Patiala, 147001, Punjab, India
[4]La Trobe University, Melbourne, Australia
[5]Sookmyung Women's University, Seoul, 04310, Korea
[*]Corresponding Author: Byung-Gyu Kim. Email: bg.kim@sookmyung.ac.kr
Received: 23 November 2020; Accepted: 08 January 2021

Abstract: As multimedia data sharing increases, data security in mobile devices and its mechanism can be seen as critical. Biometrics combines the physiological and behavioral qualities of an individual to validate their character in real-time. Humans incorporate physiological attributes like a fingerprint, face, iris, palm print, finger knuckle print, Deoxyribonucleic Acid (DNA), and behavioral qualities like walk, voice, mark, or keystroke. The main goal of this paper is to design a robust framework for automatic face recognition. Scale Invariant Feature Transform (SIFT) and Speeded-up Robust Features (SURF) are employed for face recognition. Also, we propose a modified Gabor Wavelet Transform for SIFT/SURF (GWT-SIFT/GWT-SURF) to increase the recognition accuracy of human faces. The proposed scheme is composed of three steps. First, the entropy of the image is removed using Discrete Wavelet Transform (DWT). Second, the computational complexity of the SIFT/SURF is reduced. Third, the accuracy is increased for authentication by the proposed GWT-SIFT/GWT-SURF algorithm. A comparative analysis of the proposed scheme is done on real-time Olivetti Research Laboratory (ORL) and Poznan University of Technology (PUT) databases. When compared to the traditional SIFT/SURF methods, we verify that the GWT-SIFT achieves the better accuracy of 99.32% and the better approach is the GWT-SURF as the run time of the GWT-SURF for 100 images is 3.4 seconds when compared to the GWT-SIFT which has a run time of 4.9 seconds for 100 images.

Keywords: Biometrics; real-time multimodal biometrics; real-time face recognition; feature analysis

1 Introduction and Motivation

Mobile communication is being developed with an exponential growth of mobile users, how to protect mobile networks against various attacks has become a big challenge. As multimedia data

This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

sharing increases, data security in mobile devices and its mechanism to certify it is critical. To this end, there are various authentication methods. Human biometric authentication is considered the most powerful. Especially, the face recognition is widely employed in authentication system as the hardware technology has been developed so fast.

The rapid increase in the evolution of data and technology demands a secure mechanism to access it. Almost IT-enabled systems need access security in the form of passwords, USB keys, tokens, biometric. The biometric way of authentication has gained a lot of attention because of its inherent characteristics like unique fingerprints, unique iris, unique face sentiments and so on. Among all features of an individual used in biometric systems, face attributes act as a significant value in distinguishing individuals and indicating the sentiments in the face image. The human capacity to perceive faces is momentous.

This capacity remains against changes in visual conditions, for example, outward appearances, age, as well as changes in glasses, and facial hair. The face is a critical subject in applications such as security frameworks, visa control, mugshot matching (photographs are taken from the face), observation and control of access, individual recognition. For this, an information base containing facial photographs for a gathering of individuals is utilized. Real-time face recognition is a representative one amongst the most well-known biometric frameworks. Some of the key features for using face recognition are:

- The face recognition is non-intrusive and requires no physical interaction on behalf of the user.
- Authentication can be done passively without any participation of the user, which is further beneficial for security purposes.
- It provides efficient and accurate results. Besides, no expert is required for comparison interpretation.
- Users can use a single image multiple times without having to re-enroll multiple times.

With these advantages of face recognition, it is highly preferred. However, certain problems arise due to several features of the face such as:

- Dimensionality issue arises in face image and the face, patterns are variable.
- The face is not rigid which causes various deformations in face analysis.
- The face is processed as a 2D structure. But in real, it is a 3D structure with motion, which causes the hypothesis to fail.

To resolve the issue of dimensionality different techniques are applied utilizing different optimization processes like Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Fisher Discriminant Analysis (FDA), and Independent Component Analysis (ICA). Other techniques are based on the modeling of shape and texture information like Local Feature Analysis (LFA), Active Appearance Model (AAM) and Gabor Wavelet Transforms (GWT), two-dimensional fisher's linear discriminate (2DFLD). Several methods based on multi-scale filtering with Gabor kernels are incorporated for mapping of 3D face images into its two-dimensional representation.

Scale Invariant Feature Transform (SIFT) has emerged as a cutting-edge methodology in general object recognition. It can capture the main gray-level features of an object's view using local patterns extracted from a scale-space decomposition of the image. In this respect, the SIFT approach is similar to the Local Binary Patterns method, with the difference of producing a view-invariant representation of the extracted 2D patterns. Despite having all the major benefits, the computing time of the SIFT is high which limits the speed in live image applications. The SURF

provides all the benefits of the SIFT but at lower computing time. The SURF feature vector is also smaller in size, which helps in faster classification. To utilize all these advantages of the SIFT and the SURF, the proposed work comprises of the SIFT and the SURF both to generate the feature vectors for template matching. Also, the Discrete Wavelet Transform (DWT) and Gabor Wavelet Transform (GWT) can be employed to get a higher accuracy rate.

Because of its higher acceptance rate, analysts have created different calculations for face recognition. Additionally, it has been accepted as the best real time technique for the usage of images and their investigation. Different techniques and calculations like Principal Component Analysis (PCA), Local Binary Pattern (LBP) [1] and Triplet half-band Wavelet Frequency Band (TWFB) [2] calculations have been reported for face recognition purposes [3]. The framework design for the face-based analysis includes the template matching on a higher accuracy rate to identify the best matching. The matching is done based on the distance between the ear and nose, the distance between the two ears, etc. Based on these calculations, the face template is generated and stored in the database for matching purpose [4]. The face feature template, which is to be stored in the real-time database, should be free from any type of entropy.

The real-time face authentication is usually performed in two different ways: the first way is the appearance-based analysis and is connected to the entire face picture; the second way is the segment based which uses the straight connections between the facial features, for example, eyes, mouth, and nose. The feature-based methodologies use some extra facial points and describe them by applying a bank of channels.

In this work, two appearance-based techniques are examined to separate features from face pictures. For the real-time face recognition framework, the total face or shape, the position of eyes, nose, lips, or ear are used to verify a person. One of the problems usually experienced is that the framework requires the face to be imaged in a closed environment, which is not accurate in all the conditions as an ideal methodology is not available all the time. Additionally, indistinguishable twins will have the same features and it will be hard to separate them. Two different algorithms of the SIFT and the SURF are employed for generating feature descriptors. The feature descriptors generated from the SIFT are given as an input to the SURF descriptors to increase the authentication rate by reducing the Gaussian scale-space deviation.

The rest of the paper is organized as follows: Section 2 discuss the literature review and related work in the face detection and authentication area. Further Section 3 explains the proposed methodology with the existing techniques used to enable proposed work, followed by Section 4 focused on results and discussion. The conclusion will be discussed with establishing the future work in Section 5.

2 Related Work

Real time face recognition has received generous consideration from specialists in biometrics, PC vision, design recognition, and intellectual brain science networks for security, man-machine correspondence, content-based picture recovery, and picture/video coding. We introduce a portion of the noteworthy explores for face recognition.

Ramesha et al. [5] presented a novel system that joins the multi-resolution feature of the discrete wavelet change (DWT) with the nearby connections of the facial structures communicated through the auxiliary concealed Segmental Hidden Markov Model (SHMM). A scope of wavelet channels, for example, Haar, biorthogonal 9/7, and Coiflet just as Gabor filter, have been employed so as to scan for the best execution. The SHMMs played out a careful probabilistic

investigation of any successive example by uncovering the two internal and external structures at the same time. That was accomplished by means of the idea of nearby structures presented by the SHMMs. In this manner, the long-go reliance issue natural to conventional Hidden Markov Models (HMMs) [6] has been radically decreased. The SHMMs have not recently been connected to the issue of face distinguishing proof. The detailed outcomes have demonstrated that the SHMM was superior to the conventional concealed Markov with 73% of increment in precision.

A face recognition strategy utilizing fake insusceptible systems dependent on Principal Component Analysis (PCA) has been suggested by Snelick et al. [7]. Since the PCA abstracts main eigenvectors of the picture to get the best component depiction, thus it can lessen the quantity of contributions of resistant systems [8]. Based on this, the picture information of the diminished measurements was contributed to make safe system classifiers. In this way, the antibodies of the resistant systems are streamlined for utilizing hereditary calculations. Woo et al. [9] exhibited neighborhood Gabor XOR designs which encode the Gabor stage by utilizing the nearby XOR design administrator. At that point, they presented square-based Fisher's straight discriminate (BFLD) to decrease the dimensionality of the proposed descriptor and in the meantime, improve its discriminative power by Amato et al. [10]. By utilizing BFLD, authors combined nearby examples of Gabor extent and stage for face recognition. Specifically, relative exploratory investigations are performed of distinctive nearby Gabor designs.

Zhang et al. [11] have proposed a shape coordinating based-face recognition framework, which utilized "shape" for an Indentity (ID) of appearances. The possibility of utilizing form coordinating for human face identification was exhibited through trial examination. The upside of utilizing form coordinating is that the structure of the face was unequivocally spoken to in its depiction alongside its algorithmic and computational effortlessness that makes it appropriate for equipment usage. The info shape was coordinated with enrolled form utilizing straightforward coordinating calculations. The proposed calculation was tried on the Bio ID face database and the recognition rate was observed to be 100%. The reliable client identification framework was reported by Huang et al. [12]. Biometric offers a solid answer for specific parts of personality by perceiving the people dependent on their inalienable physical. The multimodal biometric individual check was being increased much prevalence lately according as they beat unimodal individual checks. This paper showed an individual check framework utilizing discourse and face information. In [13], Mulder et al. proposed a shrewd multimodal biometric confirmation framework for physical access control, in viewpoint of the combination of iris, face, and fingerprint features. Feature vectors were made freely to inquiry pictures. This framework was intended to suit implanted answers for high security get to in inescapable situations utilizing biometric features.

The Expectation-Maximization (EM) evaluated calculation has been proposed for score level information combination by Hong et al. [14]. These biometric frameworks for individual validation and identification depended on physiological or social features that were normally particular. Multi-biometric frameworks, which can solidify data from various biometric sources, are picking up since they were ready to overcome some limitations. Reenactment demonstrated that Finite Mixture Modal (FMM) was very compelling in demonstrating the authentic and impostor score densities on 2005 multi-biometric database which was dependent on face and discourse modalities.

Chaudhary et al. [15] have displayed a multimodal biometric recognition framework coordinating palm print, fingerprint and face dependent on score level combination. The element vectors were removed autonomously from the pre-handled pictures of a palm print, fingerprint impression and face. The element vectors of inquiry pictures were contrasted independently. Also, the enlistment layouts which were taken and put away during database planning for each biometric

characteristic individually. These individual standardized scores were lastly consolidated into an all-out score by aggregate guideline, which was gone to the choice module which announces the individual as veritable or a fraud. Yu et al. [16] introduced a novel four-picture encryption dependent on the quaternion Fresnel changes (QFST), which the computer created a multi-dimensional image. In this method, the two-dimensional (2D) Logistic-changed Sine map (LASM) was added. In their work to treat the four pictures comprehensively, two kinds of changes were performed, and a quaternion lattice was determined.

In [17], a novel visual saliency-guided picture recovery strategy has been invented by Wang et al. In their work, they introduced a multi-include combination worldview of pictures. For grouping, the intricacy of the pictures, intellectual burden based-characterization, and psychological based-characterization was made. Scanty strategic relapse was coordinated to play out the model. Devpriya et al. [18] introduced a novel hash-based multifaceted secure shared confirmation scheme. It incorporates numerical hashing properties, endorsements, nonce esteems, customary client ids, and secret phrase systems that oppose MITM assaults, replay assaults, and fraud assaults. The execution was done in Microsoft Azure to get the results.

Konga et al. [19] proposed a shape coordinating-based face recognition framework, which utilized "shape" for the ID of appearances. The possibility of utilizing form coordinating for human face identification was exhibited through trial examination. The upside of utilizing form coordinating is that the structure of the face was unequivocally spoken to in its depiction alongside its algorithmic and computational effortlessness. The info shape was coordinated with the enrolled form utilizing straightforward coordinating calculations. In [20], Alsmirat et al. investigated to address the need to determine the optimum ratio of the fingerprint image compression to ensure the fingerprint identification system's high accuracy. They proved that the used software functioned perfectly until a compression ratio of (30–40%) of the raw images, but any higher ratio would negatively affect the accuracy of the used system. Also, the computerized picture preprocessing was utilized to refine out clamor, upgraded the picture, converted to the twofold picture, and find the reference point by Telgad et al. [21]. For twofold pictures, Katz's calculation was used to appraise the fractal measurement (FD) from two measurements (2D) pictures. Biometric attributes were extricated as fractal designs utilizing Weierstrass cosine work (WCF) with various FDs.

A fingerprint catch comprises of contacting or rolling a finger onto an inflexible detecting surface was depicted by Satya et al. [22]. In their demonstration, the shape was mis-conducted because the versatile skin of the finger. The amount and bearing of the weight connected by the client, the skin conditions, and the projection of a sporadic 3D object (the finger) onto a 2D level plane may present mutilations, commotion, and irregularities on the captured fingerprint picture. Due to these negative impacts, the portrayal of a similar fingerprint impression changed each time when the finger was put on the sensor plate. As a result, a new way to deal with catching fingerprints has been proposed. Their methodology alluded to as contactless or contactless fingerprinting attempts to defeat the above-referred to issues.

To combine some features, Raju et al. [23] suggested a novel face recognition strategy, which misused both worldwide and nearby discriminative features. In their strategy, worldwide features were extricated from the entire face pictures by keeping the low-recurrence coefficients of Fourier change, For nearby element extraction, Gabor wavelets have been used for considering their natural significance. From that point forward, Fisher's direct discrimination (FLD) was independently connected to the worldwide Fourier features and every neighborhood fix of Gabor feature. Along these lines, various FLD [24] classifiers were acquired in each typifying diverse facial confirmations for face recognition. Every one of these classifiers are consolidated to shape

a various leveled troupe classifier. They assessed their proposed strategy utilizing two substantial scale face databases: FERET and FRGC variant 2.0. Tests demonstrated that the aftereffect of the proposed strategy was astonishingly superior to the best-known outcomes with a similar assessment convention.

Azeem et al. [25] have reported an effective neighborhood appearance feature extraction strategy dependent on Steerable Perjuring (S–P) wavelet change for face recognition. Data was extricated by figuring the insights of each sub-square gotten by separating S–P sub-groups. The acquired neighborhood features of each sub-band were consolidated at the component and chosen dimension to improve face recognition accuracy. The reason for their paper was to investigate the helpfulness of S–P as a feature extraction strategy for face recognition. In [26], Diverse multi-goals changes, wavelet (DWT), Gabor, curvelet, and contourlet were likewise looked at against the square-based S–P technique. Tested results on ORL, Yale, Essex, and FERET face databases showed that the proposed technique could give a superior portrayal of the class data. It also acquired a lot higher recognition correctness's in genuine world circumstances incorporating changes in posture, articulation, and light.

To separates the best possible features straightforwardly from picture, Kumar et al. [26] introduced a novel calculation for picture feature extraction, to be specific, the two-dimensional territory saving projections (2DLPP). The results on the Poly-U palm print database demonstrated the viability of the proposed scheme. a wavelet-based palm print check approach was proposed Niu et al. [27]. The unmistakable wavelet area features, for example, sub-band vitality circulation, histogram, and co-event features have been extracted from the palm prints adequately because of coefficient annoyances brought about by translational as well as rotational varieties in palm prints. In this work, a novel approach which was named as "versatile decoration of sub-bands," was introduced to capture the spatially limited vitality appropriation dependent on the spread of important lines. Yao et al. [28] depicted a novel recognition technique for high-goals palm print. The fundamental commitments of the proposed technique include: (1) Use of different features especially, thickness and vital lines for palmprint recognition to altogether enhance the coordinating execution of the regular calculation. (2) Design of a quality-based and versatile introduction field estimation calculation, which is superior to the current calculation in the event of areas with a substantial number of wrinkles. (3) Use of a novel combination plot for identification application which performs superior to ordinary combination strategies. For example, there are a weighted total guideline, SVMs, or Neyman–Pearson rule. Also, they investigated the discriminative intensity of various element mixes and found that thickness was useful for palmprint recognition.

Rani et al. [29] proposed a component level combination approach for enhancing the effectiveness of palmprint identification. Different curved Gabor channels with various introductions were utilized to extricate the stage data on a palm print picture. Then, they were converted by a combination principle to deliver a solitary include called the Fusion Code. The likeness of two Fusion Codes was estimated by their standardized hamming separation. A dynamic edge was utilized for the last choices. In [30], a dynamic weighted discrimination power analysis (DWDPA) has been suggested to enhance the DP of the selected discrete cosine transform coefficients (DCTCs) without pre-masking window by Leng et al. The DCTCs were adaptively selected according to their discrimination power values (DPVs). More DCTCs with higher DP have been preserved. The selected coefficients were normalized and dynamic weighted according to their DPVs. Based on this, it does not need to optimize the shape and size of pre-masking window.

Also, Leng et al. [31] proposed multi-instance palmprint feature fusion based on the features of left and right palmprints with two-dimensional discrete cosine transform (2D-DCT) to constitute a dual-source space. Since more discriminative coefficients can be preserved and retrieved with discrimination power analysis (DPA) from dual-source space, the accuracy performance is improved. To balance the conflict between security and verification performance, two cancelable palmprint coding schemes such as PalmHash Code and PalmPhasor Code, have been proposed by Leng et al. [32]. In order to reduce computational complexity and storage cost, the cancelable palmprint coding frameworks were extended from one dimension to two dimensions.

Although various methods and approaches have been reported in the above, there is still room for improvement by feature fusion. In this study, we suggest a novel face recognition scheme using the DWT on spatial features of the SIFT and SURF in face image.

3 Proposed Methodology

An important thing of feature extraction is the dimensionality reduction by choosing some distinct features that can represent the face image with less distortion. The main goal of the proposed work is to check how the recognition performance can be increased in less time. Therefore, the image transformation approach is used as a pre-processing before the feature extraction stage. The SIFT and SURF algorithms are used for feature extraction. Before extracting feature, the input images are transformed using the GWT. After the GWT transformation, eight sub-band images are generated. The SIFT and SURF are applied for feature extraction on all the eight sub-bands to generate a score. These scores are then fused using score-level fusion and the decision is made based on the best score. The approach is tested on two different datasets: ORL and PUT for Magnitude and Phase. The tested results show the better performance of GWT over 1-scale and 2-scale DWTs when tested on both the SIFT and SURF approaches. The proposed approach is compared with various other algorithms based on both accuracy and run-time, which also depicts that the proposed GWT-SURF method is the best when both accuracy and run-time are taken together. To get a more descriptive view, all the methodologies will be discussed separately.

3.1 Scale Invariant Feature Transform (SIFT)

The initial step of the SIFT algorithm is to build the scale-space representation. Initially, a Gaussian scale-space is generated based on comparison of the pixel values. As shown in Fig. 1, each pixel in the 4×4 local region is compared to all the neighboring pixels. The maximum and minimum pixel value is obtained after the comparison. The pixel value which has some entropy is simply discarded. The feature vectors are generated based on the maximum and minimum values [33]. The complete region of the image is then divided into sub-bands and a histogram is generated for each band which is merged to generate a final histogram. Then the final histogram is normalized to a single unit vector as a descriptor.

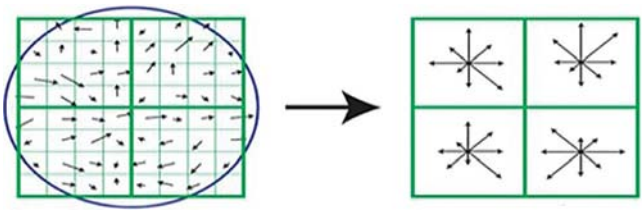


Figure 1: Key descriptor array generation for SIFT

3.2 Speeded up Robust Features

The SURF is an efficient method for key-point detection and descriptor construction. Feature descriptors are calculated by the Hessian matrix approximation in Fig. 2. Key points are localized in a 3×3 matrix [34]. To remove the effects of illumination and rotation, the orientation of the region of interest is computed based on the information point around the square image region which yields a feature vector of 64 bits.

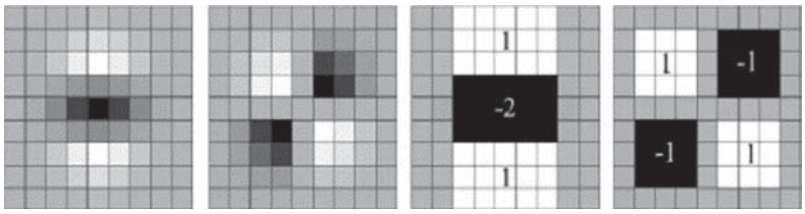


Figure 2: SURF matrix filters for Gaussian scale-space

Appropriate Gaussian kernel may affect on the accuracy of recognition. The total count of all reactions inside a sliding introduction window of 60° evaluates the principle introduction. The vigorous part is that necessary pictures can be utilized to discover wavelet reaction at any scale.

3.3 Wavelet Transform

2D Discrete Wavelet Transform (2D-DWT) and Gabor Wavelet Transform (GWT) are generally utilized as tunable channels for identifying and extracting feature data from the picture. Apart from the extraction, an invariance to enlightenment property makes them fit to catch the transformed data of the pixels [35]. A Gabor wavelet channel is a Gaussian bit capacity regulated by a sinusoidal plane wave as shown in Eq. (1):

$$\varphi_g(u, v) = \exp\left(-\pi^2\left(\frac{(u' - f)^2}{\alpha^2} + \frac{v'^2}{\beta^2}\right)\right), \tag{1}$$

where

$$u' = ucos\theta + vsin\theta,$$
$$v' = v\cos\theta - u\sin\theta. \tag{2}$$

Here, f is a frequency component, (u, v) is a pixel position, α and β are the standard deviations in two dimensions, respectively.

Gabor wavelets have different features and properties that could be utilized in various ways and applications. A unique feature of Gabor wavelets is directional selectivity. With this element, we can arrange Gabor wavelets in any ideal way.

The 2D-DWT of a flag is performed by rehashing the 2D examination channel bank on the low pass sub picture [36]. Here, four sub-images are utilized rather than one in the preparation of each scale. There are three wavelets, which are related to the 2D wavelet change [37]. Reiteration of the separating and destruction process on low-pass component yields various dimensions (scales).

In Fig. 3, the 2D-DWT changes the connected structure on a face picture and yields four diverse sub-band pictures, to be specific: approximate (average), vertical, horizontal, and diagonal components.



Figure 3: Box filter analysis for SURF

The proposed methodology contains subparts of the bands in completing the recreations. Each picture is changed utilizing DWT or GWT [38]. Two methods are proposed utilizing the SIFT and the SURF. In the first methodology, the SURF or the SIFT is utilized as an element extraction calculation, yet before extricating features, input face pictures are changed utilizing the DWT. The DWT produces four distinctive sub-band pictures in particular: approximate (average), vertical, horizontal, and diagonal [39]. Figs. 4 and 5 demonstrate 1-scale changes of input face pictures utilizing the SIFT and SURF. Key point recognition and portrayal are performed on the yield sub-band pictures utilizing the SIFT and SIFT characterized as the DWT-SIFT and the DWT-SURF.

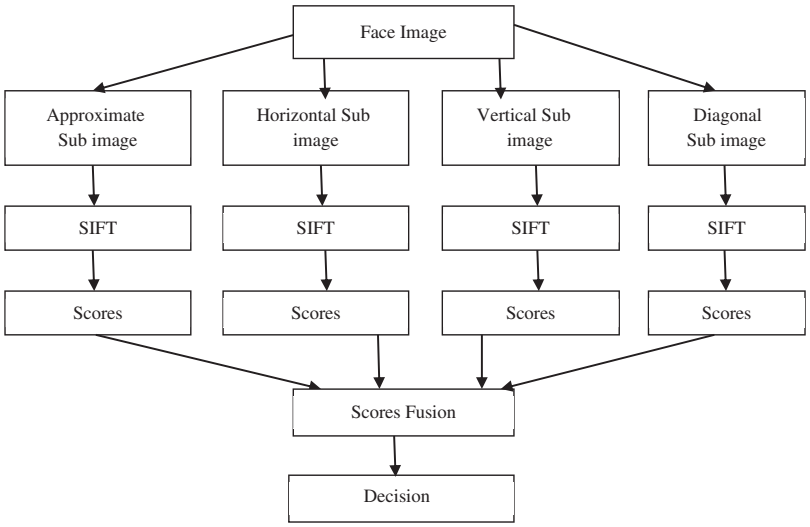


Figure 4: Sub-image characterization in the SIFT with the DWT

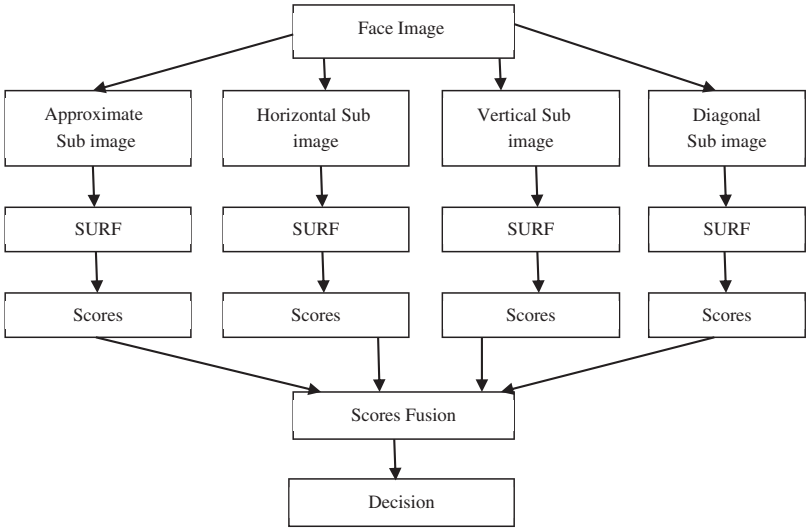


Figure 5: Sub-image characterization in the SURF with the DWT

All key point features that are extricated from the SURF, will be examined. At this point, each relating feature of key points will be compared to get a score (that characterizes the quantity of the coordinated key points). The summation of scores is done. Finally, the choice will be made according to the obtained score which will characterize if a subject is authorized or not.

3.4 Procedure of the Proposed Method

As shown in Fig. 6, the procedure of the proposed methodology can be described as:

- Input the Face Image to the processed.
- Filter the image with the GWT to detect and extract the orientation information from the image. Besides, the phase information of pixels and texture of the image is also identified.
- Eight different sub-bands of the processed image is generated by the repetition of filtering
- Normalize all the feature descriptors and then apply the SIFT/SURF of all the eight sub-band images.
- After applying the SIFT/SURF on all the feature descriptors, scores are generated for individual sub-image. These scores are fused in a single score by Score-level fusion.
- The best match score is computed and then stored in the database for authenticating humans.
- Repeat the process for all the input test images and then train the images accordingly.
- The database is updated with the best match score.
- The user's face is compared with the best match template stored in the database. The decision is made based on the best match threshold. For example, if the matching distance is less than the threshold (T) (between 70–80% match), the user is said to be Genuine and accepted. Otherwise, it will be rejected.

  Let t be a test picture and g be a reference picture. A vector $k(Xt,\ Yt)$ is applicable if:

$$k\,(X_t, Y_t) = \sqrt{(X_t - X_g)^2 + (Y_t - Y_g)^2} < T, \tag{3}$$

  where X and Y are directions of the element vector focuses. The cosine comparability is utilized for computing the vector correlation.

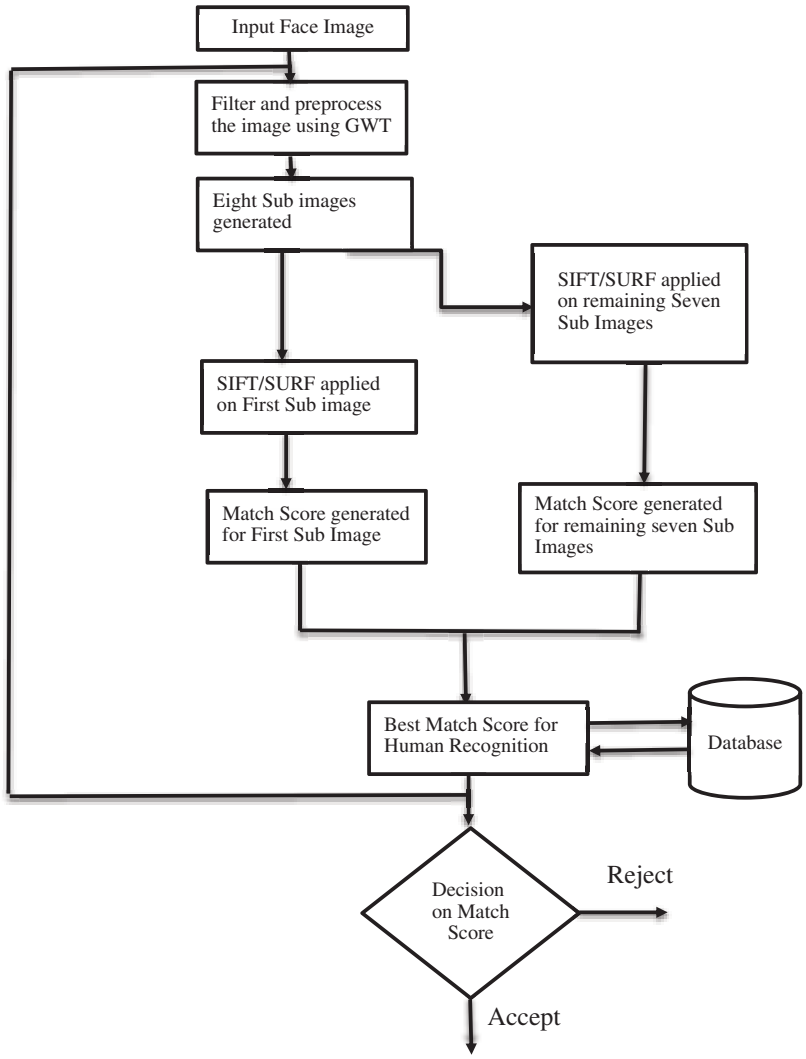


Figure 6: Flow process of the proposed algorithm

The process is applied for both the algorithms. For example, with the SIFT and SURF, the results are computed. The results of both approaches are tested on ORL and PUT databases. After the experimental analysis, it is concluded that the GWT-SURF is the most effective for real-time biometric face authentication of humans with an accuracy of 98.75% and 3.4 seconds run time for 100 sample images, because the proposed algorithm utilizes the frequency feature including spatial features of the SIFT and SURF.

4 Results and Discussion

For the investigations on each dataset, 5 randomly picked face pictures are considered as the display (train) set, and the rest of the 5 face pictures are considered as the test set. There was no covering between pictures from the display and test sets. Subjects in the two databases have 10 face pictures with various conditions such as extraordinary light, present appearance, and so forth. Every one of the pictures in the test set are analyzed against the pictures in the data set. Then the outcomes and scores are fused together before an ultimate conclusion is made.

The proposed methodologies were tried against traditional schemes. The matching score of the SIFT and SURF descriptors from various sub-bands were intertwined before a final choice at each examination.

4.1 Performance Analysis

Two different datasets are used for performance analysis: Olivetti Research Laboratory (ORL) [37] and Poznan University of Technology (PUT) face databases [40]. In the ORL database, there are 40 distinct subjects (people), 10 pictures for each subject, a sum of 400 face pictures. For the vast majority of the subjects, the face pictures were recorded at light change, time fluctuation, face subtleties (glasses/no glasses), face demeanors (open/shut eyes, grinning/not grinning), and head presents (pivot and tilting up to 20°).

Most of the face pictures were recorded against a dim normal foundation. In the PUT face database, diverse datasets have been taken under different lighting conditions in a controlled environment to check the veracity of the subjects. The subjects consisted of the data from an age group of 5 years to 65 years including 200 male, 355 female and children datasets which are taken in different lighting conditions. For most of the subjects, the pictures were recorded with various face appearances, enlightenment and head presents. The database supplies extra data including square shapes containing face, nose, eyes, and mouth parts.

4.1.1 Results with ORL Database

The general recognition rate utilizing an alternate number of subjects have appeared in Figs. 7 and 8, individually. It was seen that the period of complex GWT sub-band pictures did not work appropriately with the SIFT in Fig. 7. By combining the size and features, the propose scheme can improve the recognition accuracy significantly compared with the SIFT calculation.

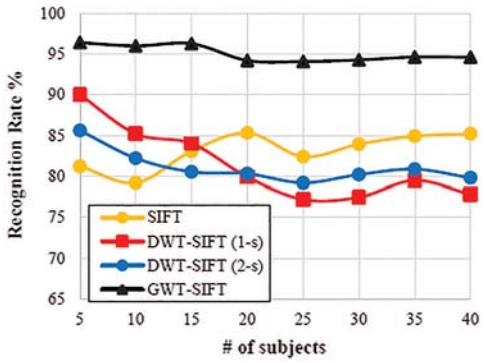


Figure 7: Comparative analysis of the SIFT features on ORL database

It can be seen from Fig. 7 that the recognition rate of the proposed GWT is better than single SIFT when applied in 1-scale and 2-scales separately. The recognition rate of single SIFT applied was obtained upto 85% whereas the SIFT with the GWT yielded a recognition rate of 95%. Also, Fig. 8 depicts the outcome of the ORL database when the SURF, DWT SURF, and GWT is applied, respectively.

In Fig. 8, we can see that the DWT yields the same outcome for both the dimensions even if the picture is having illumination entropy. This means that illumination change does not affect the feature generation. The outcome clearly shows that the GWT-SURF can yield the best recognition

rate as compared to other schemes, because the proposed algorithm utilizes the frequency feature including spatial features of the SIFT and SURF.

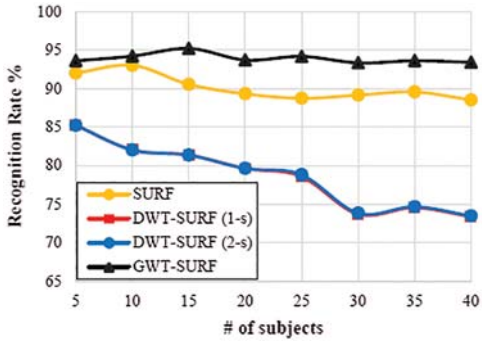


Figure 8: Comparative analysis of the SURF features on ORL database

At the first stage (5 subjects), the recognition rate was 90.09% utilizing an alternate number of people from 5 to 40 subjects when the SURF was employed as shown in Fig. 8. The execution time and recognition rate of all calculations in these investigations were diminished when the number of subjects expanded. To utilize distinctive change of channels, 1-scale change was connected to face pictures. The execution of the proposed methodology was bad enough because, after the change of face pictures, the SURF did not separate enough features to depict them. The rough sub-band pictures had zero SURF that focused on any picture in the exhibition or test sets. Along these lines, the remaining sub-band pictures (flat, vertical, and diagonal) of the DWT were utilized in the investigations.

The accuracy of the SIFT utilizing the ORL database was 83.15%. The proposed methodology utilizing 1-size DWT-SIFT achieved to ~7% when comparing to the SIFT calculation. Also, the normal distinction between 1-scale and 2-scale of SIFT-DWT was upto 2% utilizing the ORL face pictures. While applying an alternate number of subjects on the proposed methodology, the GWT-SURF beats and shows better outcomes compared with the DWT-SURF and the SURF method.

In checking the execution time, the proposed methodologies were diverse when compared with the past outcomes based on the GWT rather than the DWT as a change was occurred on the face pictures. In the 1-scale change, the GWT yielded sub-band pictures in a structure, which makes SURF nonfunctional in removing features. In this case, the proposed methodologies performed well. The recognition accuracy of the proposed methodologies utilizing Magnitude and Phase of changed pictures was achieved by a factor of ~5% higher than the customary SURF algorithm. When the number of subjects is expanding, the recognition execution of the proposed methodologies diminishes less compared with the SURF calculation. From this result, we can see that the proposed methodologies can produce a higher recognition rate of upto 12% than customary SIFT method.

Results for the SIFT, the SURF, and the proposed methodologies with the ORL database are illustrated in Tabs. 1 and 2. The recognition rate is calculated based on single SURF and then Phase and Magnitude are applied along with the SURF separately. After that, the combination of Phase and Magnitude is applied along with the GWT and the rate is computed. From these

results, it can be seen that the combination of Phase and Magnitude has yielded better results compared to the single one.

Table 1: SURF with magnitude and phase applied on ORL database

# of subjects	SURF (%)	GWT mag-SURF (%)	GWT phase-SURF (%)	GWT (mag + phase)-SURF (%)
100	93.00	90.20	92.60	97.20
200	90.53	91.73	93.60	95.20
500	89.30	90.60	92.50	93.70
1000	88.72	91.28	92.88	94.16
1500	89.13	88.73	92.33	95.33
2000	89.54	89.31	92.00	95.06

Table 2: SIFT with magnitude and phase applied on ORL database

# of subjects	SIFT (%)	GWT mag-SIFT (%)	GWT phase-SIFT (%)	GWT (mag + phase)-SIFT (%)
100	79.20	89.12	94.60	97.00
200	83.06	90.73	96.93	96.27
500	85.30	91.62	95.80	94.20
1000	82.40	90.18	94.72	94.08
1500	83.93	89.73	94.40	94.27
2000	84.91	90.31	93.89	94.63

4.1.2 Results with PUT Database

The general recognition execution rate utilizing an alternate number of subjects have been appeared in Figs. 9 and 10. In this analysis, 5 to 100 unique subjects from the PUT database were tried. The proposed GWT gave ~13% distinction in the recognition rate compared with traditional SIFT method. The average recognition rate of the SIFT was 83.15%. From this result, the general recognition rate of the SIFT by applying the GWT was about 95.06%. We observed that the GWT module with the blend of Magnitude and Phase could efficiently improve the recognition accuracy compared to the SIFT method. The general recognition rate for the various number of subjects can be seen in Fig. 9.

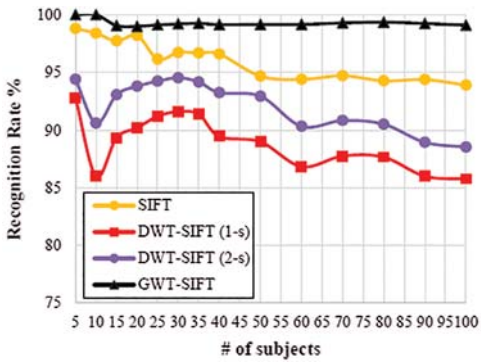


Figure 9: Comparative analysis of the SIFT features for PUT database

From Fig. 10, the subsequent normal recognition rate was measured by 84.84% utilizing only SURF algorithm. The recognition accuracy of the SURF algorithm was diminished when the quantity of subjects increased. In 1-scale (dimension) change, the accuracy was bad because the SURF algorithm was not ready to extract robust features to portray face pictures.

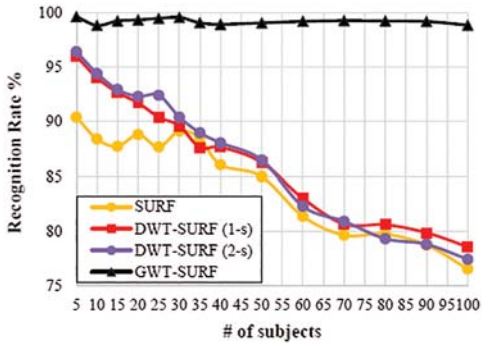


Figure 10: Comparative analysis of the SURF features on the PUT database

Results for the SIFT, the SURF, and the proposed methodologies utilizing the PUT database are shown in Tabs. 3 and 4. The recognition ratio of the proposed methodologies with the Phase of changed pictures was improved by upto 3% higher than the traditional SURF method. Also, the recognition rate of the proposed methodologies diminished very gradually compared with the SURF method as the number of subjects rapidly increased.

Table 3: The SURF with magnitude and phase applied on PUT database

# of subjects	SURF (%)	GWT mag-SURF (%)	GWT phase-SURF (%)	GWT (mag + phase)-SURF (%)
100	88.40	98.80	98.20	98.75
200	88.80	96.30	97.10	97.00
500	89.13	95.53	95.60	95.27
1000	86.05	94.90	93.60	94.20
1500	81.38	93.16	94.62	95.13
2000	79.76	95.20	94.83	95.36

Table 4: The SIFT with magnitude and phase applied on PUT database

# of subjects	SIFT (%)	GWT mag-SIFT (%)	GWT phase-SIFT (%)	GWT (mag + phase)-SIFT (%)
100	98.40	98.80	97.00	99.32
200	98.20	96.10	92.60	96.00
500	96.73	95.27	94.00	96.20
1000	96.60	95.25	95.80	95.15
1500	94.40	94.27	96.15	94.16
2000	94.28	94.39	94.69	96.36

4.2 Comparative Analysis

For more comparative analysis, the proposed method is compared with different algorithms: PCA, LDA, FDA, LBP, and conventional SIFT and SURF. The comparison is done in terms of two parameters: Accuracy and Run-time (Tabs. 5 and 6).

Table 5: Accuracy comparison on PUT database

Algorithms	Accuracy (%)
PCA	81.34
LDA	89.87
FDA	97.50
LFA	89.61
SIFT	93.42
GWT-SIFT	99.32
SURF	92.50
GWT-SURF	98.75

Table 6: Run-time analysis on PUT database

Algorithms	Run time (Seconds)
PCA	1.8
LDA	1.5
FDA	7.2
LFA	6.9
SIFT	3.7
GWT-SIFT	4.9
SURF	5.3
GWT-SURF	3.4

For the PUT database, the recognition accuracy is shown in Tab. 5 for the proposed method and various other face recognition algorithms including the conventional SIFT and SURF methods. The analysis has been done for 100 face images. From the result, we can see that the DWT-SIFT gives the best accuracy than all other algorithms. In terms of run-time, the proposed GWT-SURF achieved the shortest processing time as 3.4 seconds in Tab. 6.

5 Conclusion

In the paper, a score-level fusion methodology has been proposed with the SIFT and the SURF along with the addition of the GWT. The proposed method gave the better accuracy on a consumed run time. The proposed method was tested on two different databases: ORL and PUT. After the experimental analysis, the accuracy of the GWT-SIFT was more than the GWT-SURF when tested on just 100 images. However, when considering the results of both accuracy analysis and run-time analysis at same time, the GWT-SURF turned out to be more effective. Despite the GWT-SIFT giving better accuracy of 99.32% than the GWT-SURF scheme, the GWT-SURF with 3.4 seconds was the best approach as compared to the GWT-SIFT which has a run-time of

4.9 seconds for 100 images. When the sample amount was increased to 2000 images, the accuracy become a constant equal upto 96% which concludes that the proposed approach is also effective in the case of larger datasets. From this comparative analysis, we can conclude that the proposed approach is better in the case of real-time face recognition and vary large number of samples. The proposed scheme can be applied to mobile authentication service using smartphone camera which is connected to the cloud server.

The major advantage of this approach is that the accuracy of recognition is not affected by entropy factors such as distortion and illumination. Once the template is generated properly, the framework will be able to authenticate the human with a higher accuracy. The disadvantage is that the process of generation of the template and the consumed time for finding the best matching depends on the processing speed of the system. This run-time (consumed time) can be further improved in the future with the help of Gaussian filters which can extract the feature data in much less time. Also, the limitation of this method is that for processing, the ideal image size for the best accuracy should be 300 dots per inch (dpi). Thus, the increase of input image size may reduce the accuracy of recognition. The increase in the sample size can depend on the environment and the number of templates to be processed.

Acknowledgement: Authors thank those who contributed to write this article and give some valuable comments.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] B. Kranthikiran and P. Pulicherla, "Face detection and recognition for use in campus surveillance," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 3, pp. 2901–2913, 2020.

[2] S. Kumar, S. Singh and J. Kumar, "A study on face recognition techniques with age and gender classification," *Proc. IEEE Int. Conf. on Computing, Communication and Automation*, vol. 2017, no. 1, pp. 1001–1006, 2017.

[3] J. Fan and T. W. S. Chow, "Exactly robust kernel principal component analysis," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 3, pp. 749–761, 2020.

[4] Q. Zhang, M. Zhang, T. Chen, Z. Sun and Y. Ma *et al.,* "Recent advances in convolutional neural network acceleration," *Neurocomputing*, vol. 323, no. 61176031, pp. 37–51, 2019.

[5] R. Ramesha and K. Raja, "Execution evaluation of face recognition in light of DWT and DT-CWT utilizing multi-matching classifiers," *International Journal on Computational Intelligence and Communication Network*, vol. 601, pp. 601–605, 2011.

[6] M. S. Saleem, M. J. Khan, K. Khurshid and M. S. Hanif, "Crowd density estimation in still images using multiple local features and boosting regression ensemble," *Neural Computing & Applications*, vol. 32, pp. 16445–16454, 2019.

[7] R. Snelick, U. Uludag, A. Mink, M. Indovina and A. Jain, "Large-scale evaluation of multimodal biometric authentication using cutting edge systems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 3, pp. 450–455, 2015.

[8] M. J. Khan, A. Yousaf, A. Abbas and K. Khurshid, "Deep learning for automated forgery detection in hyperspectral document images," *Journal of Electronic Imaging*, vol. 27, no. 5, pp. 1, 2018.

[9] U. Woo and C. Seng, "An audit of biometric innovation alongside patterns and prospects," *Pattern Recognition*, vol. 47, no. 7, pp. 2673–2688, 2014.

[10] G. Amato, F. Falchi, C. Gennaro, F. V. Massoli, N. Passalis *et al.,* "Face verification and recognition for digital forensics and information security," in *Proc. of the 2019 7th Int. Symp. on Digital Forensics and Security*, Barcelos, Portugal, Piscataway, NJ, USA: IEEE, pp. 1–6, 2019.

[11] J. Zhang, C. Tsoi and S. Lo, "Scale invariant feature transform flow trajectory approach with applications to human action recognition," in *Proc. Int. Joint Conf. on Neural Networks*, Beijing, China, pp. 1197, 2014.

[12] J. Huang, Su. Kehua, J. El-De, T. Hu and J. Li, "An MPCA/LDA based dimensionality reduction algorithm for face recognition," *Universal Journal on Mathematical Problems in Engineering*, vol. 2014, pp. 1–12, 2014.

[13] M. Mulder, S. Bethard and M. Moens, "A survey on the application of recurrent neural networks to statistical language modeling," *IEEE Transactions on Computer Speech & Language*, vol. 30, no. 1, pp. 61, 2015.

[14] A. Jain, L. Hong and S. Pankanti, "Biometric identification," *Communications of the ACM*, vol. 43, no. 2, pp. 90–98, 2000.

[15] S. Chaudhary and R. Nath, "A multimodal biometric recognition framework based on the fusion of palmprint, fingerprint, and face," in *Proc. Int. Conf. on Advances in Recent Technologies in Communication and Computing*, Kottayam, Kerala, pp. 596–600, 2016.

[16] C. Yu, J. Li, X. Li, X. Ren and B. B. Gupta, "Four-image encryption scheme based on quaternion Fresnel transform, chaos, and computer-generated hologram," *Multimedia Tools and Applications*, vol. 77, no. 4, pp. 4585–4608, 2018.

[17] H. Wang, Z. Li, Y. Li, B. B. Gupta and C. Choi, "Visual saliency guided complex image retrieval," *Pattern Recognition Letters*, vol. 130, pp. 64–72, 2020.

[18] R. Telgad, P. Deshmukh and A. Siddiqui, "Combination approach to score level fusion for multimodal biometric system by using face and fingerprint," in *Proc. Int. Conf. on Recent Advances and Innovations in Engineering*, Pune, 2014.

[19] A. Konga, D. Zhanga and M. Kamelb, "Palm print identification using feature level fusion," *Pattern Recognition*, vol. 39, no. 4, pp. 478–487, 2016.

[20] M. Alsmirat, F. Alalem, M. Al-Ayyoub, Y. Jararweh and B. B. Gupta, "Impact of digital fingerprint image quality on the fingerprint recognition accuracy," *Multimedia Tools and Applications*, vol. 78, no. 4, pp. 1–40, 2015.

[21] R. Telgad, P. Deshmukh and A. Siddiqui, "Combination approach to score level fusion for multimodal biometric system by using face and fingerprint," in *Proc. Int. Conf. on Recent Advances and Innovations in Engineering*, Pune, 2014.

[22] V. Satya and C. Saravanan, "Performance analysis of various fusion methods in multimodal biometric," in *Proc. IEEE Int. Conf. on Computational and Characterization Techniques in Engineering & Sciences*, Lucknow, India, pp. 5–8, 2018.

[23] A. S. Raju and V. Udayashankara, "A survey on unimodal, multimodal biometrics, and its fusion techniques," *International Journal of Engineering & Technology*, vol. 7, no. 4, pp. 689–695, 2018.

[24] I. Ahmad, M. Ilyas, M. Isa and R. Ngadiran, "Information Fusion of Face and Palmprint Multimodal Biometrics," in *Proc. IEEE Conf. in Information and Forensics*, Malaysia, 2014.

[25] A. Azeem, M. Sharif, J. Shah and M. Raza, "Hexagonal scale-invariant element change (H-SIFT) for facial element extraction," *Journal of Applied Research and Technology*, vol. 13, no. 3, pp. 402–408, 2015.

[26] A. Kumar, D. Wong, C. Shen and A. Jain, "Personal verification using palmprint and hand geometry biometric," in *Proc. of Audio and Video-Based Biometric Individual Authentication*, Mysore, pp. 668–678, 2013.

[27] B. Niu, Q. Yang, Simon Chi Keung Shiu and Sankar K. Pal, "Two-dimensional Laplacian faces method for face recognition," *Pattern Recognition*, vol. 41, no. 10, pp. 3237–3243, 2008.

[28] Y. Yao, X. Jing and X. Wong, "Face and palm print feature level fusion for single example biometrics acknowledgment," *Neurocomputing*, vol. 70, no. 7, pp. 1582–1586, 2007.

[29] P. E. Rani and R. Shanmugalakshmi, "Personal identification system based on palmprint," *Journal of Applied Sciences*, vol. 14, no. 18, pp. 2032–2039, 2014.

[30] L. Leng, J. Zhang, J. Xu, K. Khan and K. Alghathbar, "Dynamic weighted discrimination power analysis: A novel approach for face and palmprint recognition in DCT domain," *International Journal of Physical Sciences*, vol. 5, no. 17, pp. 467–471, 2010.

[31] L. Leng, M. Li, C. Kim and X. Bi, "Dual-source discrimination power analysis for multi-instance contactless palmprint recognition," *Multimedia Tools and Applications*, vol. 76, no. 1, pp. 333–354, 2017.

[32] L. Leng and J. Zhang, "PalmHash code *vs*. PalmPhasor code," *Neurocomputing*, vol. 108, pp. 1–12, 2013.

[33] A. K. Jaber and I. Abdel-Qader, "A hybrid feature extraction framework for face recognition," *International Journal of Handheld Computing Research*, vol. 8, no. 1, pp. 81–94, 2017.

[34] S. Kakarwal, R. R. Deshmukh and B. Ambedkar, "Hybrid feature extraction technique for face recognition," *International Journal of Advanced Computer Science and Applications*, vol. 3, no. 2, pp. 60–64, 2012.

[35] R. Fu, D. Wang, D. Li and Z. Luo, "University classroom attendance based on deep learning," in *Proc. 10th Int. Conf. on Intelligent Computation Technology and Automation*, Changsha, China, pp. 128–131, 2017.

[36] M. Arsenovic, S. Sladojevic, A. Andela and D. Stefanovic, "FaceTime—Deep learning-based face recognition attendance system, in *Proc. IEEE 15th Int. Symp. on Intelligent Systems and Informatics*, Subotica, Serbia, pp. 53–58, 2017.

[37] K. Walker, S. Kanade and V. Jadhav, "A multi-modal and multi-algorithmic biometric system combining iris and face," in *Proc. Int. Conf. on Information Processing*, Pune, India, pp. 496–501, 2015.

[38] Q. Zhang, M. Zhang, T. Chen, Z. Sun, Y. Ma *et al.,* "Recent advances in convolutional neural network acceleration," *Neurocomputing*, vol. 323, no. 61176031, pp. 37–51, 2019.

[39] M. S. Saleem, M. J. Khan, K. Khurshid and M. S. Hanif, "Crowd density estimation in still images using multiple local features and boosting regression ensemble," *Neural Computing & Applications*, vol. 32, no. 21, pp. 16445–16454, 2020.

[40] B. Anand and P. K. Shah, "Facer recognition utilizing SURF features and SVM classifier," *International Journal of Hardware Engineering Research*, vol. 8, no. 1, pp. 1–8, 2016.