Innovative Computing Review (ICR)
Volume 2 Issue 2, Fall 2022
ISSN$_{(P)}$: 2791-0024 ISSN$_{(E)}$: 2791-0032
Homepage: https://journals.umt.edu.pk/index.php/ICR

Article QR

Title: **Feature Based Techniques for a Face Recognition using Supervised Learning Algorithms based on Fixed Monocular Video Camera**

Author (s): Isra Anwar, Syed Farooq Ali, Jameel Ahmad, Sumaira Zafar, Malik Tahir Hassan

Affiliation (s): University of Management and Technology, Lahore, Pakistan

DOI: https://doi.org/10.32350.icr.22.05

History: Received: September 9, 2022, Revised: November 29, 2022, Accepted: December 10, 2022

Citation: I. Anwar, S. F. Ali, J. Ahmad, S. Zafar, and M. T. Hassan, "Feature based techniques for a face recognition using supervised learning algorithms based on fixed monocular video camera," *UMT Artif. Intell. Rev.,* vol. 2, no. 2, pp. 82-110, 2022, doi: https://doi.org/10.32350.icr.22.05

Conflict of Interest: Author(s) declared no conflict of interest

A publication of
School of Systems and Technology
University of Management and Technology, Lahore, Pakistan

Feature Based Techniques for Face Recognition using Supervised Learning Algorithms Based on Fixed Monocular Video Camera

Isra Anwar[1], Syed Farooq Ali [2], Jameel Ahmad[3], Sumaira Zafar[4*], and Malik Tahir Hassan[5]

School of Systems and Technology, University of Management and Technology, Lahore, Pakistan

***Abstract*- Automatic face recognition has ample significance in biometric research. Recent decades have witnessed enormous growth in this research area. Face-based identification is always considered more expedient as compared to other biometric authentications owing to its uniqueness and wide acceptance. The major contribution of this work is twofold; firstly, it comprises an extension of manual thresholding feature-based face recognition approach to an automatic feature-based supervised learning face recognition. Secondly, various new feature sets are proposed and tested on several classifiers for 2, 3, 4, and 5 persons. In addition, the use of slope features of facial components, such as the nose, right eye, left eye, and lips along with other conventional features for face recognition is also a unique contribution of this research. Multiple experiments were performed on the UMT face database. The results demonstrated a comparison of 5 different sets of feature-based approaches on 7 classifiers using the metrics of time efficiency and accuracy. They also depicted that the proposed approaches achieve a percentage accuracy of up to 95.5%.**

***Index Terms*- face identification, face recognition, geometric features, supervised learning, Support Vector Machine (SVM)**

I. Introduction

Despite the fact that numerous algorithms have been proposed in the literature, face recognition in unconstrained scenes remains a challenging problem due to changes in appearance caused by the variations in pose and expression. Still, face recognition has achieved mammoth popularity during the last few decades in computer vision, pattern recognition, neural networks, neuroscience, cognitive science, image processing, psychology, and physiology for good reasons.

The major reasons for its wide acceptance are its simplicity, security, and authenticity. Face authentication works regardless of

* Corresponding Author: sumairazafar.sz@gmail.com

any user input in the form of passwords, PIN codes, plastic cards, keys, tokens, and smart card authentication, which makes it simple. As far as security is concerned, all other methods which involve user input are no match for face recognition, as user input in any aforementioned form always comes with the risk of its misuse. This is exemplified by the use of passwords and PINs, which can be guessed or stolen.

Moreover, the solicitude of enhanced physical security systems has enjoyed rapid growth in recent years due to a dramatic increase in the crime rate. Face recognition plays a pivotal role in this regard. It is widely used in numerous applications, such as automatic attendance systems, physical access control, surveillance control, people tagging, gaming, image search, and many more. This work proposes several feature-based approaches of face recognition containing some new features, namely slope table, along with random projection and regional properties.

There are three major contributions of the current study. Firstly, it proposes a new approach for face recognition which extends the manual thresholding system [1] to a supervised learning algorithm. Secondly, five different sets of new features are suggested for feature matching. Thirdly, a detailed comparison of these approaches is provided using 7 existing classifiers on a dataset of 2, 3, 4, and 5 persons.

The proposed work uses the slopes of different fiducial landmarks of facial components (nose, right eye, left eye, and lips) to build a slope table that is merged with other geometric and holistic features. Finally, features are fed to the classifiers for training and matching. The rest of the paper is organized as follows. In Section 2, a review of the related literature is presented. Section 3 gives an overview of the data set used for experiments. It also incorporates the proposed framework. Experimental evaluation is presented in Section 4. Finally, conclusions are drawn in Section 5.

II. Literature Review

Face recognition has been an active area of research for the last few decades. Initially, a semi-automated system was proposed for face recognition and only major features, such as ears, eyes, mouth, and nose were used for this purpose [2]. The algorithms of face recognition can be divided into two broad categories, namely feature-based techniques and holistic approaches [3]. Feature-based approaches initially identify and

extract the unique facial features and then compute geometric relationships among these distinctive features. Ultimately, some standard statistical pattern recognition techniques are applied to these computed geometric features for matching.

Liu et al. [4] proposed a groundbreaking Gabor-Fisher Classifier (GFC) for face identification with 100% accuracy in face recognition using only 62 features. Gabor-Fisher Classifier is vigorous to illumination and facial expression variations. Its feasibility was tested on the FERET database containing 600 FERET frontal facial images, corresponding to 200 subjects, captured in manifold illuminations and distinct facial expressions. Although, these features have been categorized as the most successful face representations, which are highly rated dimensions for rapid extraction and accurate results.

Shen et al. [5] proposed a framework of Gabor wavelets and general discriminant analysis (GDA) which extracts features from the whole face image. This algorithm was also tested on the same FERET database, as well as on the BANCA face database, with a 97.5% recognition rate and 5.96% verification error rate, respectively. Afterwards, in some variations of these approaches, Gabor features were replaced by a graph-matching strategy [6]–[8] and the histograms of oriented gradients [9]. Campadelli et al. [10] proposed an elastic bunch graph-based technique with automatic fiducial point localization. This technique computes 16 facial fiducial points and the head pose. Their proposed technique normalizes the image and characterizes facial fiducial points with its jets vector to extract the peculiar texture around jets. Finally, they measured the similarity among manifold jets to recognize the image. This approach was tested on a database of 2500 face foreground images, yielding an accuracy of 93% for face recognition. They claimed to obtain the same performance as the existing elastic bunch graph approaches with manual graph placement.

The discriminant elastic graph matching (DEGM) algorithm by Zafeiriou et al. [7] uses discriminant techniques at all phases of elastic graph matching for face verification. This conventional EGM was extended to a generalized EGM (G-EGM) in order to provide improved performance on globally warped faces [8]. It improved the robustness of node descriptors to globally misaligned faces and introduced warping-compensated

edges, with an incredible accuracy rate for face identification. Despite great accuracy, the technique has not gained adequate attention in many suitable real-time applications due to its expensive computational cost. Albiol et al. [9] applied HOG descriptors instead of Gabor features, which are more robust to illumination changes and insignificant displacements. As a result, higher accuracy was achieved for face verification in comparison to the existing Gabor-EBGM based approaches.

Yang and Zhang [10] employed sparse representation-based classification (SRC) for face recognition. This algorithm codes the input testing images as a sparse linear combination of training samples. Local Gabor features of images were used for SRC and an algorithm was introduced to compute the associated Gabor occlusion dictionary which handles occluded face images with much higher recognition rates. Moreover, the computational cost of sparse coding was intensely reduced by compacting the occlusion dictionary. The major contribution of this work was to achieve high accuracy against the variations in lighting, expression, pose, and occlusion.

Nevertheless, Gabor transformation is considered impractical for numerous real-time face recognition applications owing to its unusual time and space complexity. Yang et al. proposed a new and swift technique called monogenic binary coding for local feature extraction to combat this issue [11]. In this scheme, the original signal is decomposed into three harmonious elements, namely amplitude, orientation, and phase. Monogenic variation and features are encoded in each local region and pixel, respectively

Another variation of Gabor transformation was proposed by Zhenhua, et al. which deals with inter-person and intra-person variations in face images by using the robustness of ordinal measures, along with the distinctiveness of Gabor features [12]. The above authors also computed the histogram and initially generated feature vectors by concatenating statistical distributions, which were further refined by linear discriminant analysis. Finally, face recognition classifier was trained by practicing cascade learning and greedy block selection methods.

In holistic approaches, global representations of the entire face are used for recognition instead of using some selective local features, such

as Eigenfaces [13], principal component analysis (PCA), linear discriminant analysis (LDA) [14], and independent component analysis-based approaches. These approaches take all images as matrices of intensity values. For every input image, a direct correlation comparison of this image is performed with all other faces in the database for face recognition. Holistic approaches can be further subdivided into two main categories, that is, statistical approaches and artificial intelligence (AI) based approaches. The simplest approach in statistical approaches uses intensity values of all facial pixels and performs correlation comparisons with all faces stored in the database. This approach is considered impractical because it is not only computationally expensive but works only in a restricted view, such as equal illumination, face orientation, and size [13]. Another shortcoming of such approaches is that it usually classifies data in a space of very high dimensionality. However, several other approaches have been proposed to combat the dimensionality problem by reducing the statistical dimensionality and procuring only meaningful feature dimensions for comparisons. Lu et al. also presented an unsupervised approach for automatic feature learning from raw pixels of facial images to learn hierarchical feature representation, instead of using Gabor features or any other conventional descriptor [14]. They used multiple feature dictionaries to represent different physical characteristics of the unique face region and computed multiple related features from them.

Sirovich et al. [15] represented facial images economically by utilizing PCA [16], [17]. They used Eigen-pictures coordinate space to represent a face efficiently and reconstructed that face using a small collection of Eigen-pictures and their projections. Jian et al. [18] introduced the Spectro face method with 98% accuracy on Yale and Olivetti face databases. This method combines the Fourier transform and wavelet transform and proves that low-frequency images, decomposed using the wavelet transform, are less sensitive to the variations in facial expression. This method uses transformed images, invariant to scale, translation, and on-the-plane rotation. The outcome of the study showed that wavelet transform delivered robust representations in illumination variations and captured substantial facial features with low computational complexity.

Following these considerations, an efficient scheme for face

recognition by applying wavelet sub-band representation was proposed by Zhang et al. [19]. Their method used two-dimensional (2D) wavelet sub-band coefficients to represent the face images. They also used a personalized classification method for recognition based on kernel associative memory (KAM) models. This method was tested on three standard face identification datasets, namely XM2VTS, FERET, and ORL. The results showed that this method outperformed the existing approaches in terms of accuracy. They also utilised KAM in combination with the Gabor wavelet network to construct a unified structure of Gabor wavelet associative memory (GWAM) [20]. This method was tested using three famous databases and reported a very high performance as compared to other techniques.

Kwak et al. [21] systematically formed an enhancement of the generic independent component analysis called FICA by supplementing it with Fisher LDA and presenting it along with its underlying architecture. This method improved well-separated classification rates significantly and also dealt with illumination and facial expression invariants.

It is generally accepted that low-resolution representation of face images reduces the performance of face recognition. To combat this issue, Huang et al. [22] proposed a super-resolution approach that maps coherent features, nonlinearly. It improves higher recognition of the nearest classifiers for the recognition of low-resolution face image. Initially, coherent subspaces are constructed by applying the canonical correlation analysis among low-resolution face images and the features of high-resolution based on PCA. Then, radial basis functions are used to build the nonlinear mapping between high and low-resolution features. Lastly, super-resolved coherent features are constructed and fed to a simple NN classifier for face identification.

The computational complexity of these techniques was further reduced by Azad et al. [23]. In their proposed face recognition scheme, instead of taking all pixels of the whole image only the face was identified and PCA technique was applied to that particular region in order to extract the vector features. Then, multiclass support vector machine (SVM) classifiers were applied to these features for identification and classification. This technique achieved a 98.45% accuracy rate on the FEI database.

Zhang et al. also presented a novel and efficient low-resolution face recognition algorithm known as coupled marginal discriminant mappings (CMDM) [24]. CMDM makes data points in the low and the original high-resolution features and compels them into a unified space. It constructs samples closely from the same class and separately from a distinct class with a great margin. This method avoids the dimensional mismatch problem and fills the data gap of different resolutions. CMDM achieved outstanding performance on the AR and FERET face databases as compared to other up-to-date lower solution face recognition techniques.

In 2016, Yihang et al. [25] proposed an approach that deals with multimodal face and ear recognition in unconstrained scenes by using the fusion of both spherical depth map and the spherical texture map characteristics. They initially used sparse representation for recognition and then the final results were refined using Bayesian decision-level fusion. Recently, Yongjie et al. [26] investigated low-resolution face recognition with one image per person. A cluster-based regularized simultaneous discriminant analysis (C-RSDA) method was proposed based on SDA. C-RSDA computes cluster-based scatter matrices from unsupervised clustering. The between-class scatter matrices are regularized with inter-cluster scatter matrices and within-class matrices with intra-cluster scatter matrices. As this approach allows exploiting more variations from the narrow training samples, so, the overfitting and singularity problems are dealt effectively. The effectiveness of this algorithm has been proven by extensive testing on low-resolution face images captured in both controlled and uncontrolled environments.

Feature-based approaches using classifiers show promising results and the research community remains convinced that these learning-based techniques can provide better results in the future. This paper is an extension of the existing manual threshold feature-based approach to automatic supervised learning-based approach for face recognition [1]. In this work, different sets of new features are suggested for feature matching. Each approach was tested on seven existing classifiers and the results are reported in Section 5. These classifiers include neural network (Multilayer Perceptron (MLP)), Naive Bayes, J48, SVM, AdaBoost, Decision Tree, Decision Table, and Nearest-Neighbor-like Algorithm (NNge) [27]–[29]. Features calculated in approach 5

outperformed all other approaches in terms of accuracy with more than 90% true positive results. Sagonas et al. [30] explored the problem of joints and facial contaminated data. They proposed the JIVE and Robust JIVE techniques where 1-norm was employed on data. They performed extensive 2D and 3D face age progression experiments where they found the RJIVE method as the most reliable technique, among others.

Wang et al. [31] offered a recurrent face aging (RFA) GRU framework with triple layers. GRU triple layer better determines face identification than bi-layer with a combination of RNNs. However, during their testing, there was a lack of age input in their experiments which may be catered to in future work. Singhal and Srivastava [32] carried out a survey in which different databases and approaches were explored to rectify the problems faced during the face identification process. They explored many facial databases, real-time images, and videos, so that effective machine-learning approaches could be applied to improve facial identification processes. Zhu et al. [33] proposed a cascaded CNN approach instead of using traditional methods. They used a face profiling approach to use abundant samples during training and a cost function OWPDC was also introduced to improve the parameters' priority. Mollahosseini et al. [34] trained neural networks with multiple scenarios to differentiate their performance between labelled and non-labelled data. As per their observation, noise estimation may apply to a few expressions.

Hang et al. proposed an approach in which detection traced faces in a frame. Where, face is aligned according to the canonical view to crop it with the normal size of the pixels [35].

Govind et al. observed that maximum performance percentage could be gained by using masked faces in training classifiers. Based on empirical analysis, it can be improved with new values and variables [36]. Marcus et al. also conducted an analysis on facial recognition application [37].

III. Methodology

The proposed approaches were evaluated on a self-generated UMT face dataset[1] consisting of 5 subjects. The UMT face database (DB) contains 150 face images, with

[1]https://sites.google.com/site/farooq1us/dataset

30 face images of each subject marking the variations in pose and expression. This data was used with 10-fold cross-validation. Fig. 1 and Fig. 2 show different frames that contain various expressions and different poses.

Fig. 1. Key frames from UMT database

BoRMaN was used to detect face images and compute the coordinates of 22 facial points which act as inputs to the proposed algorithm [38]. Table I provides a vivid description of the dataset. The current authors assumed a single camera setup to capture static colored images of 500*375 resolutions with various poses and expressions, although illumination remained constant.

In this study, various new feature sets are proposed and used with existing classifiers for face recognition, instead of using a manual threshold system for matching [1]. The proposed framework works in two major phases, namely training and testing. Each phase consists of a sequence of four steps. The initial three steps of both phases are similar, starting from the identification and localization of the face region in the

2D face image. The second step is the extraction of distinctive facial components, such as mouth, eyes, nose, as well as other facial fiducial points calculated by using BoRMAN. In the third step the geometrical relationship among these facial points is estimated. Thus, a vector of geometric features from the input facial image is shown below.

Fig. 2. UMT database with pose and expression variations

$$FV(I_i) = FeaturesCalculation(IF(I_i)) \quad (1)$$

If (I_i) is the face pixels of the fifth training image for Approach 1, Approach 2, and Approach 3, FV (I_i) is the feature vector of the fifth training image. The proposed methodology uses five different approaches to calculate features and these features are used with 7 classifiers including Neural Network (Multilayer Perceptron (MLP)), Naive Bayes, SVM, J48, AdaBoost, Decision Tree, Decision Table, and Nearest-Neighbor-like Algorithm (NNge) [39], [40]. In the final step of the training phase, these extracted features are employed by the classifiers to train it as shown below.

$$Trained\ Classifier = Classifier(FV(I_i) \quad (2)$$

where $1 \leq i \leq N$, $N = 5060$

Table I
UMT Face Dataset Description

Properties	Description
No. of subjects	5
No. of images/Videos	150
Videos/Static	Static
Multiple face	Single
Gray Scale/Color	Color
Face Pose/Various	Various Pose
Pose Facial/Various	Various Expression
Expression Illumination	Constant Illumination
Resolution	500*375

On the other hand, in the final step of the face recognition phase, the feature vector (FV) of a given test image a test has employed by the trained classifier for matching. The trained classifier categorizes the image into 6 categories, namely Person 1, Person 2, Person 3, Person 4, Person 5, and no match.

$$Category_{(1<j<6)} = Trained\ Classifier\ (FV(I_{test})) \quad (3)$$

Section 5 delineates the discussion regarding the results of all classifiers. All aforementioned strides are well explained in the sub-sections below. Fig. 3 shows the block diagram of the proposed algorithm.

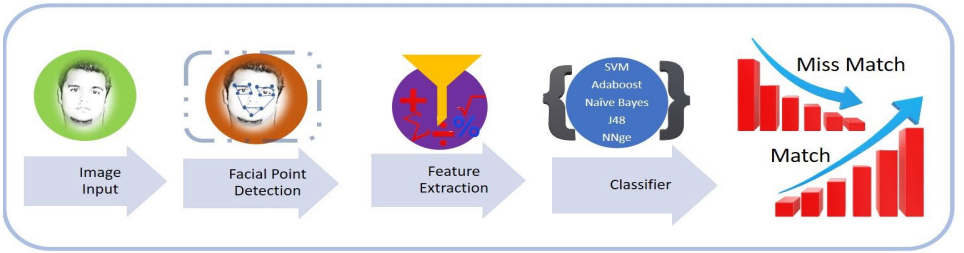

Fig. 3. Block diagram of proposed framework

A. Facial Fiducial Points Detection

The proposed framework employed BoRMaN to detect face images and extract the facial fiducial points for each frame. BoRMaN iteratively uses support vector regression (SVR) and local appearance-based features to provide an initial prediction of 22 points of facial components, including nose, lips, chin, mouth, eyebrows, and eyes. The distribution of these 22 facial fiducial points is shown in Fig. 4. Ten points are used to mark both eyes, three collectively for the nose, four for lips, four points represent eyebrows, and one point is used to spot the chin. Then, it computes

ratios, angles, and areas of different invariant features using these facial fiducial points and formulates several approaches to build various novel sets of features used in the testing and training of classifiers.

Fig 4. Facial fiducial points are detected and shown on the facial image at the right side by applying BoRMaN on the left image

B. Features Reckoning

Facial fiducial points computed in the previous phase were used to calculate different features, using several approaches given to the classifiers for training and testing. Identification labels of these features were passed as input to a classifier for training and testing purposes. Face recognition is based on the output of the classifiers. If the values of the features determine the difference as below the threshold during the testing phase, then as per the matching frame, the classifiers are automatically classified. The proposed five different methods to extract features from 2D frames are explained below. The first four techniques differ by feature collection, whereas the fifth approach is a combination of Approach 4 in combination with the slope table.

1) Approach 1: Random projection features (RPF)

This is a holistic feature extraction approach that is time efficient and based on features extracted from the facial image and sparse matrix. Initially, features are extracted effectively using a sparse measurement matrix and then multiplied with the image features. This model employs random projection which preserves the structure of an image. A random matrix Z $\epsilon Z^{(\mathrm{n}*\mathrm{m})}$ whose rows have unit length projects data from the

high-dimensional image space $x \epsilon Z^m$ to a low-dimensional space $v \epsilon Z^n, v = Zx$. Fig 6 illustrates the feature-based face recognition model using random projection.

$$z_{i,j} = \begin{cases} if\ val < 0\ then\ val = -1 \\ if\ val > 0\ then\ val = 1 \\ if\ val = 0\ then\ val = 0 \end{cases} \quad (4)$$

In this work, a sparse random measurement matrix adopted efficient dimensionality reduction with entries, as shown in Eq. 4.

2) Approach 2: Histogram of oriented gradient features (HOGF)

This appearance and shape-based approach calculates the features based on the gradients of the given facial image. The gradient is a generalization of the usual concept of derivative to the function of two variables of the given image represented as $F(f_x, f_y)$, whose components are derivatives in horizontal and vertical directions. It is, thus, a vector-valued function. Initially, the gradient vector of the given image is calculated as shown in Eq. 5.

$$\nabla F(x,y) = \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}\right) \quad (5)$$

The gradient vector was used to compute a magnitude of gradients using Eq. 6.

$$M = \sqrt{(f_x^2 + f_y^2)} \quad (6)$$

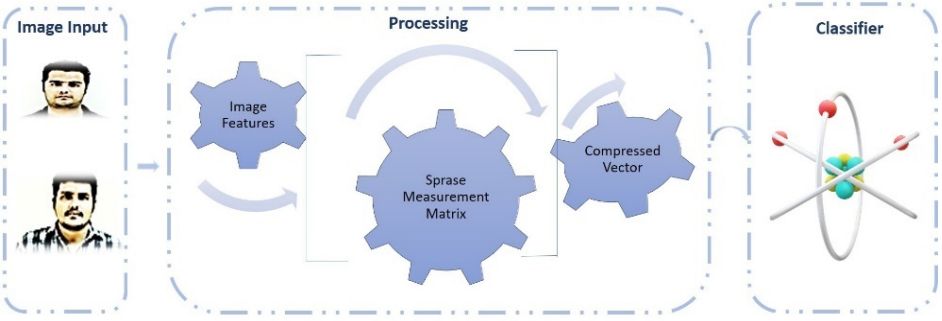

Fig. 5. Explaining random projection that converts image features into the compressed vector using sparse measurement matrix

In the next step, a histogram of this magnitude is calculated to divide the entire range of values into a series of intervals. This reduces the feature vector size based on frequency. The histogram of magnitude is also considered in stating distinctive values for each image. The overall process of feature calculation using a

histogram of the oriented gradient is illustrated in Fig 6.

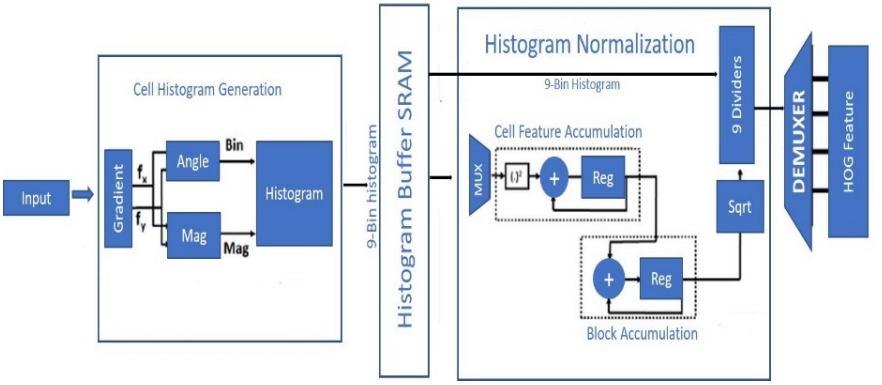

Fig. 6. Feature extraction by computing histogram of oriented gradients using a facial image as input

3) Approach 3: regional properties features (RePF)

In this approach, the regional properties template is applied over image pixel values. A regional property template is a set of different scalar regional properties, such as area, Euler number, eccentricity, extent, major axis length, orientation, solidity, Equiv-diameter, max intensity, minor axis length, mean intensity, and minor intensity. Where, area identifies the actual number of pixels in the region and the Euler number specifies the total number of objects in the image subtracting the total number of holes. The proposed regional property template uses 8-connectivity to compute the Euler number measurement. Equiv-Diameter determines the diameter of a circle with the same area as the region and is calculated using Eq. 7.

$$\sqrt{4\left(\frac{Area}{\pi}\right)} \tag{7}$$

The eccentricity property is used to measure the eccentricity of the ellipse that has the same second moments as the region. Eccentricity is the ratio between the major axis length and the distance between the foci of the ellipse. Its value is between 0 and 1. In the regional properties template, the ratio of pixels in the region to pixels in the total bounding box is computed using Eq. 8.

$$Extent = \frac{Area}{Area(boundingbox)} \tag{8}$$

The major and minor length axis specifies the length of the major and minor axes of the ellipse that have the same normalized second central moments as the region, respectively. Orientation states the angle between the x-axis and the major axis of the ellipse that has the same second moments as the region. The solidity of the region is calculated by the proportion of pixels in the convex hull, as shown below in Eq. 9.

$$Solidity = \frac{Area}{ConvexArea} \quad (9)$$

Fig. 7. Feature extraction through regional properties

4) Approach 4: Ratio and angles of facial image (RAF)

In this approach, a geometric feature set is constructed a combination of the ratios of length to width, angles among triangles from the facial points, and the ratios of the areas of triangles as given by BoRMaN. After the detection of these facial points, the ratio of length to width of lips and eyes, the ratio of the area of triangles, and the angle of different triangles are computed for face recognition. The details of these features are presented in this section.

The ratios of length to width of these points (left and right eyes and upper and lower lips) are computed instead of using the length or width of lips, left eye, and right eye to ensure that the scale is invariant. The length of the left eye and right eye or the width of the left eye and right eye alone could be taken as a feature but could not be considered as a scale constant. This is why the ratio of the length and width of lips and eyes (both left and right) is considered as a feature to calculate its scale constant, as shown in Eq. 10.

$$Ratio = Lips_{length}/Lips_{width} \quad (10)$$

Where, lip length is calculated using the Euclidean distance formula between points 5 (x_5, y_5) and 6 (x_6, y_6) marked on Fig 8(a), as shown in Eq. 10.

$$Lips_{length} = \sqrt{((x_6 - x_5)^2 + (y_6 - y_5)^2} \quad (11)$$

Similarly, lip width is also calculated using the Euclidean distance formula between points 7 (x_7; y_7) and 8 (x_8; y_8) marked in Fig 8(a), as shown in Eq. 12.

$$Lips_{width} = \sqrt{((x_6 - x_5)^2 + (y_6 - y_5)^2} \quad (12)$$

The ratio of length to width of the left eye in Fig. 8(a) is computed with eye length and width. Eye length and width are calculated by the Euclidean distance formula using points 1 (x_1; y_1), 2 (x_2; y_2) and 3 (x_3; y_3), 4 (x_4; y_4), respectively. The ratio of length to width of the right eye is also computed in a similar manner.

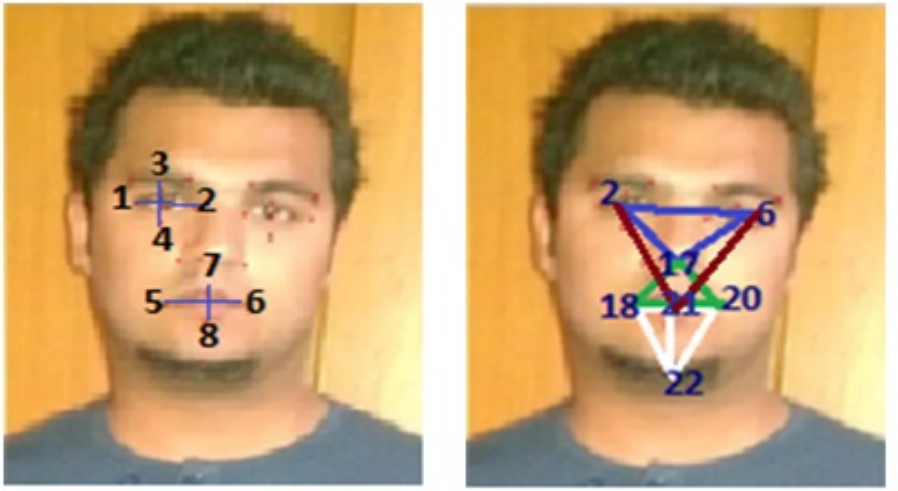

Fig. 8. Ratios of angles of triangles

The proposed method uses the ratios of the areas of triangles as ratio is a scaled invariant. The three triangles, marked in Fig. 8 (b), are used to calculate the ratios of the triangles Δ 2; 6; 17; Δ 2, 6, 21 and Δ 18, 20, 17, using the formula given in Eq. 13.

$$Ratio_{AreaOfTriangle} = \frac{Area_{Triangle1}}{Area_{Triangle1}} \quad (13)$$

The areas of triangles are computed by using Hero's formula, as shown in Eq.14.

$$Area_{Triangle} = \sqrt{(s(s-a)(s-b)(s-c))} \quad (14)$$

Where, a, b, c is the length of the sides of the triangle computed by Euclidean distance and s is the average length of the three sides of the triangle calculated in Eq. 15.

$$s = \frac{(a+b+c)}{3} \quad (15)$$

The values of different triangles' angles marked in Fig. 8 b are used as features for face recognition. These angles are calculated using the

fundamental cosine formula described below in Eqs. 16-18.

$$\alpha = cos^{-1}((b^2 + c^2 - a^2)/2bc \quad (16)$$

$$\beta = cos^{-1}((a^2 + c^2 - b^2)/2ac \quad (17)$$

$$\gamma = cos^{-1}((a^2 + b^2 - c^2)/2ab \quad (18)$$

The angles of eight triangles (< 2,6,17, < 2,6,21, < 2,6,22, <17,20,22, < 6,2,17, < 6,2,21, < 6,2,22 and < 17,18,22) among the points marked in Fig. 8(a, b) are computed using the formulas in Eqs. [16-18].

5) Approach 5: Ratios, angles, and slopes of facial image (RASF)

In this approach, the same geometric features are used as in RAF discussed in Section 3.2.4. However, this approach also uses the slope features of a facial image along with the ratios of length to width, area, and angles. Ratios and angles are computed using Eqs. [13-18]. For slope features, the proposed framework introduces a groundbreaking technique, where a slope table is used as a feature along with the other multiple features for face recognition. A slope table is constructed by calculating the slopes of the different fiducial landmarks of facial components (left eye, right eye, nose, and lips). Initially, centroids of each facial component (obtained from BoRMaN) are computed. Then, Euclidean distance is computed by using these centroids and the rest of the points. Finally, slope tables are computed for the left and right eyes by calculating the slope of each point of the left eye and right eye, respectively. Similarly, slope tables of lips and nose are also computed by employing their respective fiducial points. Slope computation is shown in Fig. 9. The centroid of a facial component is calculated using Eq. 19.

$$C = \left(\frac{\Sigma x_i}{n}, \frac{\Sigma y_i}{n}\right) \quad (19)$$

In the above equation, (X_i, Y_i) are the two coordinates of the fifth point of a facial component. P_x and P_y are Euclidean distances of the boundary point to the centroid along the x-axis and y-axis, respectively. P represents the line from the boundary point to the centroid, whereas P_x is the x-component and P_y is the y-component of the line P. P_x and P_y are used to compute the slope as shown below.

$$m = \tan\theta = \frac{P_y}{P_x} \quad (20)$$

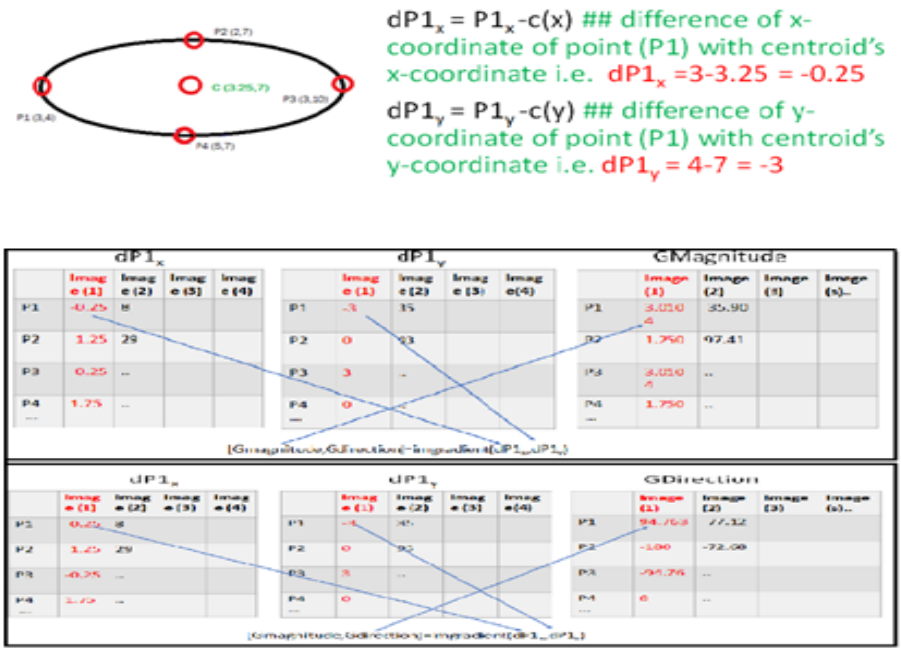

Fig. 9. Slope table computation

Several classifiers are trained and tested using these facial features.

C. Training and Testing

All the features calculated in Section 4.2 were employed as input to seven classifiers to train them. Then, the features of the test image were fed to trained classifiers for face matching. These classifiers included SVM, AdaBoost, Nearest-Neighbor-like algorithm (NNge), Naive Bayes, Decision Table, J48, and Neural Network (Multilayer Perceptron (MLP)) [38]. In the matching phase, classifiers automatically categorized the input image into 6 categories in the extreme case (for 5-person testing) as person 1, 2, 3, 4, and 5, or not matched. Similarly, for the 4-person problem, classifiers categorized into 5 categories as person 1, 2, 3, and 4, or not matched, and so on.

IV. Results and Discussion

The proposed approach was tested on seven classifiers for accuracy and scalability on a wide variety of images using five features-based approaches. The results, in terms of percentage accuracy (90% percentage split as 90:10 training: testing respectively) and time efficiency, are shown in Table II and Table III, respectively.

Table II and Table III show the results for features computed using all the above elaborated approaches applied on different numbers of persons. To test these approaches, the data set was divided into sub-parts. Hence, all the approaches were tested on 2 persons, then 3, 4, and 5 persons, which helped to understand the dimensionality and generalization of the current approaches.

Table II depicts the performance in terms of percentage accuracy of Approach 5 (RASF) using NNge as an effective approach than the rest of the four techniques and six classifiers for 2 persons. However, for persons 3 and 5, RASF with SVM classifiers performs better. For person 4, the best approach remains RASF but with MLP classifiers. Moreover, the average percentage accuracy of Approach 4 (RAF) using all classifiers for persons 2 and 5 is remarkably better than the rest of the approaches, while the average percentage accuracy of RASF for persons 3 and 4 is best among all approaches. Hence, it can be summarized that RAF and RASF show a clear generalized behavior that outperforms all the other approaches, irrespective of the number of persons. The possible reasons for the better performance of RAF and RASF are that both of these approaches use geometrical features and the motion vector-based interpolation technique. The performance of RASF is best because of the addition of a slope feature in the features of the RAF approach. So, it can be concluded that the features of the slope table are working as very useful face recognition features, achieving the highest accuracy in all the sub-parts of the data set.

The average percentage accuracy of classifiers based on all approaches shows that the Support Vector Machine (SVM) classifiers outperform the rest of the classifiers including Neural Network (Multilayer Perceptron (MLP)), Decision Table, AdaBoost, Naive Bayes, J48 and NNge. It is because SVM has a regularization parameter which helps in avoiding overfitting and uses the kernel trick, which builds expert knowledge about the problem via engineering the kernel. The second and third-best average percentage accuracy are shown by a neural network (Multilayer Perceptron (MLP)) and Naive Bayes classifiers, respectively (see Table II). The grounds of improved percentage accuracy of these classifiers must be the ability to handle the missing values which is known as completeness property. This property assures the scrutinizing of all possible

combinations of condition values. Similarly, RAF showed the second-best percentage accuracy for 3- and 4-person problems and RASF reported second-best results for 2- and 5-person problems, on average, of all classifiers.

Fig 10 (a-e) shows the pictorial representations of the results of multiple applied approaches namely Random Projection (RPF), Histogram of Oriented Gradients (HOGF), Regional Properties (RePF), Ratio and Angles (RAF), and Ratio, Angles, and Slope (RASF) in terms of percentage accuracy for 2-persons, 3-persons, 4-persons, and 5-persons problems.

Table II

Results of Random Projection (RPF), Histogram of Oriented Gradients (HOGF), Regional Properties (RePF), Ratio and Angles (RAF), and Ratio, Angles, and Slope in terms of Percentage Accuracy along with 10-fold Cross Validation

	Person	SVM	AdaBoost	NNge	N-B	D-Table	J48	MLP	Average
	1	71.6	75	71.6	76	78	78	71.6	74.5
	2	68.8	67.5	60.7	63.7	75	76.7	68	68.6
	3	56.3	61.5	58.6	59.6	68	70.2	64	62.6
RPF	4	52.8	58.7	51.6	52.8	64	68.6	56	57.7
	1	80.3	81.6	81.6	79.7	81.7	85	81.6	81.6
	2	78.5	76.6	76.5	70.8	72.2	72.2	72.2	74.1
	3	75.5	64.7	62.5	68.5	68	66.7	67.7	67.6
HOGF	4	63.3	58	58	62.7	64	58	59.3	60.4
	1	86.6	81.7	80	81.67	83.3	80	81.6	82.1
	2	79	76.6	73.5	74.5	68.9	78	73.3	74.8
	3	63.8	68.5	68.6	67.5	55.8	68	68.7	65.8
RePF	4	58.6	60	58.3	62.7	53.3	61	56.6	58.6
	1	90.6	91.6	91.6	95	91.6	81.6	93.3	90.7
	2	91.1	86.6	93.3	94.4	78.8	86.6	94.4	89.3
	3	85	75.8	85	84.1	67.5	75	88.3	80.1
RAF	4	90.6	63.6	82.6	82	62	68	91.3	77.1
	1	90.3	91.6	93.3	91.6	91.6	81.6	88.3	89.7
	2	95.5	86.6	91.1	87.7	82.8	86.6	94.4	89.3
	3	90	75.7	88.6	78.3	70.3	79.6	90.8	81.9
RASF	4	92	68.8	79.3	75.3	63.6	68	90.6	76.8
Average		78	73.5	75.3	75.4	72	74.4	77.6	

Naive Bayes=N-B, Decision table=D-Table

Table 3 shows the results for the time taken in building the model for the features computed and tested using the above mentioned approaches. It is clear from the results that Naive Bayes is the most time-efficient approach and AdaBoost is the second most time-efficient classifier. On the other hand, neural network (MLP) consumes the most time out of all the classifiers, whereas SVM overtakes all the other classifiers by achieving the highest accuracy and does not compromise time efficiency either.

Table III
Time taken to build a model for Random Projection (RPF), Histogram of Oriented Gradients (HOGF), Regional Properties (RePF), Ratio and Angles (RAF), and Ratio, Angles, and Slope (RASF) with Different Classifies

	Person	SVM	AdaBoost	NNge	N-B	D-Table	J48	MLP
RPF	1	10	30	10	0	40	10	3680
	2	30	10	30	0	40	40	5330
	3	40	10	30	0	60	50	7390
	4	80	10	60	10	80	80	10050
HOGF	1	10	10	10	0	10	0	90
	2	20	0	0	0	10	0	150
	3	40	0	10	0	20	10	620
	4	50	0	0	0	20	0	350
RePF	1	10	10	10	10	40	10	2150
	2	20	30	10	0	60	10	3080
	3	30	10	30	0	60	20	4090
	4	60	10	40	0	110	20	5830
RAF	1	10	30	10	0	60	10	4470
	2	20	50	20	0	120	10	6440
	3	40	10	40	0	160	30	8760
	4	70	10	60	0	270	50	9940
RASF	1	60	130	220	40	330	80	2190
	2	30	40	20	0	80	20	3030
	3	40	10	30	0	100	20	4110
	4	70	10	30	0	140	80	5540
Average		37	21	33.5	**3**	90.5	27.5	4364.5

Naïve Bayes=N-B, Decision table=D-Table

Fig. 11(a-e) shows the pictorial representations of the results of the approaches used namely Random Projection (RPF), Histogram of Oriented Gradients (HOGF), Regional Properties (RePF), and

Ratio and Angles (RAF), Ratio, Angles, and Slope (RASF) in terms of execution time for 2-persons, 3-persons, 4-persons, and 5-persons problems on the seven aforementioned classifiers.

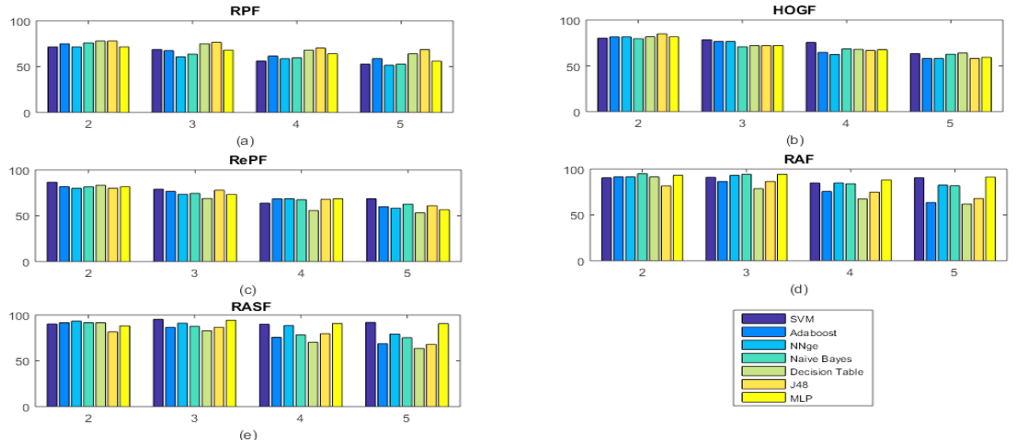

Fig. 10. Graph showing the results of Random Projection (RPF), Histogram of Oriented Gradients (HOGF), Regional Properties (RePF), Ratio and Angles (RAF), Ratio, Angles, and Slope (RASF) in terms of percentage accuracy with 10-fold cross-validation

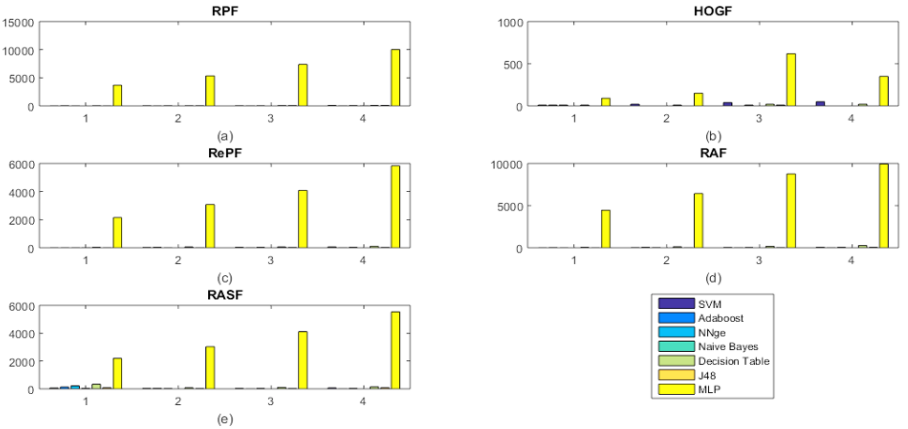

Fig. 11. Graph showing the pictorial representations of the results of Random Projection (RPF), Histogram of Oriented Gradients (HOGF), Regional Properties (RPF), Ratio and Angles (RAF), and Ratio, Angles, and Slope (RASF) in terms of execution time for 2-persons, 3-persons, 4-persons, and 5-persons problems on the seven aforementioned classifiers

A. Conclusion and Future Work

The proposed framework experiments with various features for face identification problems and suggests the use of BoRMaN and slopes. As is clearly shown by the results that feature-based face identification using BoRMaN and the slopes is the best approach among all the other discussed approaches. BoRMaN returns 22 constant facial features used to calculate features like ratios and angles. This framework is scale-invariant, however, it does not perform well in dim light. In the future, a system shoud be developed that also works in gloomy light. Besides other extensions, the availability of a large dataset repository is an important issue. Hence, more repositories should also be developed.

Acknowledgements

We would like to thank Touqir Attique, Mohsin Ijaz, Khawaja Ubaid Ur Rehman, Junaid Jabbar Faizi, Hassaam Zahid, Muhammad Maaz Aslam, Wasiq Maqsood who are voluntarily responsible for the data generation of this paper. Our gratitude goes to all the anonymous reviewers for their effective input. This research work was held up by the National ICT RD under grant no. NICTDF/NGIRI/2013-14/Crops/2; and University of Management Technology, Lahore, Pakistan.

References

[1] I. Anwar, S. Nawaz, G. Kibria, S. F. Ali, M. T. Hassan, and J.-B. Kim, “Feature based face recognition using slopes,” in *2014 Int. Conf. Cont., Automation Inform. Sci.*, IEEE, Gwangju, Korea (South), Dec. 2–5, 2014, pp. 200–205, doi: https://doi.org/10.1109/ICCAIS.2014.7020558

[2] W. W. Bledsoe, “Man-machine facial recognition,” *Rep. PRi,* vol. 22, 1966.

[3] S. B. Ahmed, S. F. Ali, J. Ahmad, M. Adnan, and M. M. Fraz, “On the frontiers of pose invariant face recognition: A review,” *Artif. Intell. Rev.*, vol. 53, no. 4, pp. 2571–2634, Apr. 2020, doi: https://doi.org/10.1007/s10462-019-09742-3

[4] C. Liu and H. Wechsler, “Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition,” *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 467–476, Apr. 2002, doi: https://doi.org/10.1109/TIP.2002.999679

[5] L. Shen, L. Bai, and M. Fairhurst, “Gabor wavelets and general discriminant analysis for

face identification and verification," *Image Vision Comput.*, vol. 25, no. 5, pp. 553–563, May 2007, doi: https://doi.org/10.1016/j.imavis.2006.05.002

[6] S. Arca, P. Campadelli, and R. Lanzarotti, "A face recognition system based on local feature analysis," in *Int. Conf. Audio- and Video-Based Biomet. Person Authentica*. Springer, 2003, pp. 182–189.

[7] S. Zafeiriou, A. Tefas, and I. Pitas, "The discriminant elastic graph matching algorithm applied to frontal face verification," Pattern Recognition, vol. 40, no. 6, pp. 2798–2810, 2007, doi: https://doi.org/10.1016/j.patcog.2007.01.026

[8] H. Shin, S.-D. Kim, and H.-C. Choi, "Generalized elastic graph matching for face recognition," *Pattern Recog. Lett.*, vol. 28, no. 7, pp. 1077– 1082, 2007, doi: https://doi.org/10.1016/j.patrec.2007.01.003

[9] A. Albiol, D. Monzo, A. Martin, J. Sastre, and A. Albiol, "Face recognition using hog–ebgm," *Pattern Recog. Lett.*, vol. 29, no. 8, pp. 1537–1543, July 2008, doi: https://doi.org/10.1016/j.patrec.2008.03.017

[10] M. Yang and L. Zhang, "Gabor feature based sparse representation for face recognition with gabor occlusion dictionary," in *Eur. Conf. Comput. Vision*. Berlin, Heidelberg, Springer, 2010, pp. 448–461, doi: https://doi.org/10.1007/978-3-642-15567-3_33

[11] M. Yang, L. Zhang, S. C.-K. Shiu, and D. Zhang, "Monogenic binary coding: An efficient local feature extraction approach to face recognition," *IEEE Trans. Inform. Forens. Secur.*, vol. 7, no. 9, pp. 1738–1751, Sep. 2012, doi: https://doi.org/10.1109/TIFS.2012.2217332

[12] Z. Chai, Z. Sun, H. Mendez-Vazquez, R. He, and T. Tan, "Gabor ordinal measures for face recognition," *IEEE Trans. Inform. Forens. Secur.*, vol. 9, no. 10, pp. 14–26, Nov. 2014, doi: https://doi.org/10.1109/TIFS.2013.2290064

[13] R. J. Baron, "Mechanisms of human facial recognition," *Int. J. Man-Machine Stud.*, vol. 15, no. 13, pp. 137–178, Aug. 1981, doi: https://doi.org/10.1016/S0020-7373(81)80001-6

[14] S. Satonkar Suhas, B. Kurhe Ajay, and B. Prakash Khanale,

"Face recognition using principal component analysis and linear discriminant analysis on holistic approach in facial images database," *Int. Organ. Sci. Res.*, vol. 2, no. 12, pp. 15–23, 2012.

[15] L. Sirovich and M. Kirby, "Low-dimensional procedure for the characterization of human faces," *J. Opti. Soc. Am. A*, vol. 4, no. 14, pp. 519–524, 1987, doi: https://doi.org/10.1364/JOSAA.4.000519

[16] A. K. Jain and R. C. Dubes, *Algorithms for clustering data*. Prentice-Hall, Inc., 1988.

[17] J. H. Lai, P. C. Yuen, and G. C. Feng, "Face recognition using holistic fourier invariant features," *Pattern Recog.*, vol. 34, no. 17, pp. 95– 109, 2001, doi: https://doi.org/10.1016/S0031-3203(99)00200-9

[18] B.-L. Zhang, H. Zhang, and S. S. Ge, "Face recognition by applying wavelet subband representation and kernel associative memory," *IEEE Trans. Neural Netw.*, vol. 15, no. 19, pp. 166–177, 2004, doi: https://doi.org/10.1109/TNN.2003.820673

[19] H. Zhang, B. Zhang, W. Huang, and Q. Tian, "Gabor wavelet associative memory for face recognition," *IEEE Trans. Neural Netw.*, vol. 16, no. 20, pp. 275–278, 2005, doi: https://doi.org/10.1109/TNN.2004.841811

[20] K.-C. Kwak and W. Pedrycz, "Face recognition using an enhanced independent component analysis approach," *IEEE Trans. Neural Netw*, vol. 18, no. 21, pp. 530–541, Mar. 2007, doi: https://doi.org/10.1109/TNN.2006.885436

[21] H. Huang and H. He, "Super-resolution method for face recognition using nonlinear mappings on coherent features," *IEEE Trans. Neural Netw.*, vol. 22, no. 22, pp. 121–130, Nov. 2011, doi: https://doi.org/10.1109/TNN.2010.2089470

[22] R. Azad, B. Azad et al., "Optimized method for real-time face recognition system based on pca and multiclass support vector machine," *Adv. Comput. Sci.*, vol. 2, no. 23, pp. 126–132, 2013.

[23] P. Zhang, X. Ben, W. Jiang, R. Yan, and Y. Zhang, "Coupled marginal discriminant mappings for low-resolution face recognition," *Optik-Int. J. Light Electron Optics*, vol. 126, no. 24, pp. 4352–4357, Dec. 2015, doi:

https://doi.org/10.1016/j.ijleo.2015.08.138

[24] Y. Li, Z. Mu, and T. Zhang, "Local feature extraction and recognition under expression variations based on multimodal face and ear spherical map," *9th Int. Cong. Image Signal Process., BioMedical Eng. Informat. (CISP-BMEI)*, Datong, China, 15–17 Oct. 2016, pp. 286–290. https://doi.org/10.1109/CISP-BMEI.2016.7852723

[25] Y. Chu, T. Ahmad, G. Bebis, and L. Zhao, "Low-resolution face recognition with single sample per person," *Signal Process.*, vol. 141, pp. 144–157, 2017, doi: https://doi.org/10.1016/j.sigpro.2017.05.012

[26] S. Aftab, S. F. Ali, A. Mahmood, and U. Suleman, "A boosting framework for human posture recognition using spatio-temporal features along with radon transform," *Multimed. Tools Appl.,* vol. 18, pp. 42325–42351, Aug. 2022, doi: https://doi.org/10.1007/s11042-022-13536-1

[27] S. F. Ali, R. Khan, A. Mahmood, M. T. Hassan, and M. Jeon, "Using temporal covariance of motion and geometric features via boosting for human fall detection," *Sensors*, vol. 18, no. 6, Art. no. 1918, 2018, doi: https://doi.org/10.3390/s18061918

[28] M. Kirby and L. Sirovich, "Application of the karhunen-loeve procedure for the characterization of human faces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 25, pp. 103–108, Jan. 1990, doi: https://doi.org/10.1109/34.41390

[29] C. Sagonas, E. Ververas, Y. Panagakis, and S. Zafeiriou, "Recovering joint and individual components in facial data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 26, pp. 2668–2681, Dec. 2017, doi: https://doi.org/10.1109/TPAMI.2017.2784421

[30] W. Wang, Y. Yan, Z. Cui, J. Feng, S. Yan, and N. Sebe, "Recurrent face aging with hierarchical autoregressive memory," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 27, pp. 654–668, Feb. 2018, doi: https://doi.org/10.1109/TPAMI.2018.2803166

[31] P. Singhal, P. K. Srivastava, A. K. Tiwari, and R. K. Shukla, "A survey: Approaches to facial detection and recognition with machine learning techniques,"

in *Proc. Second Doct. Symp. Comput. Intell.*, vol. 40, no. 28, 2021, pp. 654–668, doi: https://doi.org/10.1007/978-981-16-3346-1_9

[32] X. Zhu, X. Liu, Z. Lei, and S. Z. Li, "Face alignment in full pose range: A 3d total solution," *IEEE Trans. Pattern Anal. Mach. Intell.*,vol. 41, no. 29, pp. 78–92, Nov. 2017, doi: https://doi.org/10.1109/TPAMI.2017.2778152

[33] A. Mollahosseini, B. Hasani, M. J. Salvador, H. Abdollahi, D. Chan, and M. H. Mahoor, "Facial expression recognition from world wild web," in *Proc. IEEE Conf. Comput. Vision Pattern Recog. Workshops*, vol. 42, no. 29, 2016, pp. 78–92.

[34] M. Valstar, B. Martinez, X. Binefa, and M. Pantic, "Facial point detection using boosted regression and graph models," in *Comput. Vision Pattern Recog., 2010 IEEE Conf. IEEE*, June 13–18, 2010, pp. 2729–2736, doi: https://doi.org/10.1109/CVPR.2010.5539996

[35] S. F. Ali, A. S. Aslam, M. J. Awan, A. Yasin, and R. Damaˇseviˇcius, "Pose estimation of driver's head panning based on interpolation and motion vectors under a boosting framework," *Appl. Sci.*, vol. 11, no. 24, Art. no. 11600, Dec. 2021, doi: https://doi.org/10.3390/app112411600

[36] S. F. Ali and M. T. Hassan, "Feature based techniques for a driver's distraction detection using supervised learning algorithms based on fixed monocular video camera," *KSII Trans. Internet Info. Syst.*, vol. 12, no. 8, pp. 3820–3841, Aug. 2018.

[37] S. F. Ali, M. Muaz, A. Fatima, F. Idrees, and N. Nazar, "Human fall detection," in *Inmic. IEEE*, 2013, pp. 101–105.

[38] H. Du, H. Shi, D. Zeng, X.-P. Zhang, and T. Mei, "The elements of end-to end deep face recognition: A survey of recent advances," *ACM Comput. Surveys (CSUR)*, vol. 54, no. 10s, pp. 1–42, 2022, doi: https://doi.org/10.1145/3507902

[39] G. Jeevan, G. C. Zacharias, M. S. Nair, and J. Rajan, "An empirical study of the impact of masks on face recognition," Pattern Recognition, vol. 122, Art. no. 108308, Feb. 2022, doi: https://doi.org/10.1016/j.patcog.2021.108308

[40] M. Smith and S. Miller, "The ethical application of biometric facial recognition

technology," *Ai & Society*, vol. 37, no. 1, pp. 167–175, Apr. 2022, doi: https://doi.org/10.1007/s00146-021-01199-9

[41]