Indonesian Journal of Electrical Engineering and Computer Science
Vol. 27, No. 2, August 2022, pp. 820~831
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v27.i2.pp820-831

Gender and race classification using geodesic distance measurement

Zahraa Shahad Marzoog[1], Ashraf Dhannon Hasan[2], Hawraa Hassan Abbas[1,2]
[1]School of Engineering, Kerbala University, Kerbala, Iraq
[2]Computer Center, Kerbala University, Kerbala, Iraq

Article Info

Article history:

Received Dec 18, 2021
Revised Jun 18, 2022
Accepted Jun 25, 2022

Keywords:

Ethnicity classification
Face recognition
Gender classification
Geodesic distance
Soft biometric

ABSTRACT

Gender and ethnicity classifications are a long-standing challenge in the face recognition's field. They are key-demographic traits of individuals and applied in real-world applications such as biometric and demographic research, human-computer interaction (HCI), law enforcement and online advertisements. Thus, many methods have been proposed to address gender or/and race classifications and achieved various accuracies. This research improves race and gender classification by employing a geodesic path algorithm to extract discriminative features of both gender and ethnicity. PCA is also utilized for dimensionality reduction of Gender-feature and race-feature matrices. KNN and SVM are used to classify the extracted feature. This research was tested on the face recognition technology (FERET) dataset, with classification results demonstrating high-level performance (100%) in distinguishing gender and ethnicity.

Corresponding Author:

Zahraa Shahad Marzoog
School of Engineering, Kerbala University
Karbala 56001, Iraq
Email: zahraa.marzoog@uokerbala.edu.iq

1. INTRODUCTION

Human identification utilizes various features such as a thumbprint, iris, voice, and face to identify humans uniquely. However, facial features are a highly dependable metric for human identification. The human face possesses demographic factors such as gender, race, and age [1].

The demographic classification based on ethnicity and gender has recently evolved extraordinarily and exposed different applications, such as biometrics, surveillance monitoring, forensic art, disease diagnosis, etc. Race and gender information belongs to "soft biometrics" that supplies vital information of the identityof individuals [2]. Furthermore, it can enhance face recognition's performance by narrowing down the research space and improving results of identification performance. Thus, gender and ethnicity recognition systems are currently employed to afford smart services in different systems. They have been applied in systems, such as transport stations, police stations, airports, colleges, and clinics that require a high degree of security and accuracy system. In security applications, gender and ethnicity estimation are also used in social insurance structures, scholarly investigation and examination, and data recovery to convey customers to different gender and race classifications [3]. Much research in various fields (such as computer vision) has studied the ethnicity and sexual discriminative features in human faces by integrating strategies for processing, comprehension and extracting data from facial pictures. Generally, different studies have focused on one demographic factor (race, gender, or age) or combined two or more in classification.

Different modalities have been utilized in gender and race classifications, containing hand shape, iris, gait, and face. However, most state-of-the-art (SOA) studies concentrate on the human face's modality to detect gender and race. Thus, our method has used face images since they offer more beneficial information for ethnicity and gender classification than other modalities [4]. Our research aims to classify individuals' faces depending on both their ethnicity and gender. The geodesic path algorithm is used to extract discriminative features of 209 samples from the face recognition technology (FERET) dataset that includes different races and genders. In this work, we have conducted two main experiments. The first experiment works on classifying facial samples according to their gender(male/female), while the second experiment concentrates on ethnicity classification, in which the model classifies face samples into three ethnic groups (Asian, African, and American). Support vector machine (SVM) technique and k-nearest neighbour (KNN) are utilized to classify the extracted feature to the corrected gender and race. Thus, an automatic classification model is proposed to classify input face images according to their race and gender. To summarize, the important contributions of our research are as follows: i) We developed a new race-gender classification model based on a geodesic path algorithm as discriminative feature extractor and SVM and KNN algorithms as classifiers and ii) we conducted experiments on gender and race classification using the FERET dataset and obtained superior performance compared with SOA studies.

The remainder of this paper is as follows: section 2 shows related works on race and ethnicity classification. In section 3, the proposed method is explained in detail, including utilizing geodesic path technique in extracting features, PCA for dimensionality reduction and SVM and KNN for classification. The experiments' results and the comparison with related studies on both gender and race are discussed in section 4. Section 5 includes the conclusion of this study.

2. RELATED WORK

Various research recently proposed to address the race classification problem based on face recognition. These studies have employed statistics and mathematical algorithms for discriminating features extraction. Lu proposed a model to classify race by examining the face at different scales [5]. It applied linear discriminate analysis (LDA) to facial samples to improve the classification result. Although the accuracy was 96.3%, it only classified face images into two classes (Asian and Non-Asian). Manesh *et al.* [6] suggested a model that utilized the golden ratio mask by applying decision-making rules on various facial regions. The extracted gober features were classified using the SVM technique. It also was classified the face samples into two classes: Asian and non-Asian classes, with the accuracy being 98%. Guo et al. proposed a model that predicted race and gender based on canonical correlation analysis technique (CCA) [7]. The accuracy results were 99% and 98% for both race and gender classification. Some methods have employed a specific facial region in race classification. Lyle *et al.* [8] proposed a method that applied local binary pattern (LBP) feature-based technique to extract texture features from the periocular region. Similar to Manesh's study, it classified face samples collected from the FRGC dataset into two classes: Asian and non-Asian classes, with the accuracy result being 93%. Xei *et al.* [9] facial colour dependent feature combined with Kernel class-based feature analysis for race classification. It focused on the periorbital region and employed facial colour-dependent features and filtered responses to collect suitable features for race classification. It classified MBGC and Mugshot datasets into three ethnic groups (Asian, African, and Caucasian) and achieved higher efficiency results than previous studies (96.5% and 96.3%, respectively). Hosoi *et al.* [10] suggested a method that collaborated the gabor wavelet features and retina sampling to extract features that are classified using the SVM technique. The accuracy of face samples classification into three ethnic groups (Asian, European, and African) are 96.3%, 93.1% and 94.3%, respectively. Roomi *et al.* [11] suggested a race classification model based on the viola-jones algorithm. In this method, features from different facial regions, which include skin colour, lip colour, and forehead area, are extracted. It classified FERET and yale datasets into three groups (caucasian, negroid, mongolian), with an accuracy result being 81%.

Some proposed methods recently concentrated on deep learning in tackling the ethnicity classification problem. Baig *et al.* [12] suggested an approach that depended on (convolutional neural network) CNN to classify faces into two categories (Asian and non-Asian). This method integrates important facial features (such as colour, surface, skin) with other secondary features to effectively classify facial images. However, it achieves an accuracy of 84.91%. Khan *et al.* [13] proposed a race classification model that utilized a deep convolution neural network to create a face segmentation structure. It extracted deep features from seven various classes and constructed probability maps for each class. It utilized the probabilistic classification approach on different datasets (VMER, VNFaces, CAS-PEAL and FERET), and achieved different accuracies (93.2%, 92%, 99.2, 100%, consequently). Masood *et al.* [14] suggested a race classification method using two techniques: convolution neural network (VGGNet) and artificial neural network (ANN). This method classified three ethnic groups from the FERET dataset: Mongolian, Caucasian and Negro. It extracted geometric features and colour characteristics from facial images to be classified. However, it achieves an accuracy of 82.4% when using ANN and 98.6% when deploying CNN. Vo *et al.* [15] proposed a new race recognition framework (RRF) that contains two models: race recognition based on deep convolution neural network (RRCNN), and race recognition based on deep learning architecture VGG (RRVGG). This method utilized the VNFaces dataset to test the performance of the two models. The RRCNN achieved 88.64%, while RRVGG attained a slightly better performance (88.87%). These models applied on different datasets that include various ethnic groups (such as Chinese, Japanese or Brazilian), they have not achieved more than 91%. Greco *et al.* [16] suggested a new dataset named VMER, that includes four ethnic groups (Asian Indian, East Asian, Caucasian Latin, and African American). Then, they utilized four different convolution neural networks, namely ResNet 50, MobileNet V2, VGG-16 and VGG-face, to classify the proposed dataset. They achieved similar performance (ranging between 93.1% and 94.1%). Heng *et al.* [17] suggested a hybrid classification method that took advantage of CNN and the discriminative features extracted by a neural network. It combined convolution neural network classification results and an image ranking engine to benefit from the fitting of features between query and images in the dataset. Then the hybrid feature vectors were classified using SVM and achieved 95.2%.

Different research has also addressed gender classification using various techniques because of its importance in human identification. Some of them applied their proposed methods for both ethnicity and gender classification, such as in [6]–[8]. However, they achieved 94%, 93% for gender classification, respectively. Gutta *et al.* [18] suggested hybrid classification architectures, which include an ensemble of decision tree (DT) and radial basis function (RBF). It utilized the FERET dataset and achieved 96% on gender classification. Moghaddam and Yang [19] suggested a method that employed nonlinear Support Vector Machine SVM on low-resolution facial images, which achieved outstanding results (97%) compared to traditional classification techniques such as Nearest neighbour and Fisher Linear Discriminant technique. Buchala *et al.* [20] addressed gender classification using different face regions, including nose, eyes, and full face. It utilized three datasets: AR, BioID and FERET. It extracted two face regions and utilized midpoints of the mouth and eyes. It employed PCA for dimensionality reduction and SVM for classification and realized the best accuracy performance of 85.5%. Singh *et al.* [21] employed two feature extraction techniques: histogram oriented gradient (HOG) and local binary pattern (LBP). It also used the Haar cascade to detect face region from image, and SVM to classify gender. The best accuracy achieved in this method is 95.5%. Bekhouche *et al.* [4] suggested a method that depended on extracting features using multi-level local phase quantization from face images. SVM technique was used to classify the gender. Balci and Atalay [22] proposed a gender classification method that applied pruning schema to multi_layer perceptron (MLP) which utilized eigenvector coefficients created by PCA. The classification performance result was 92%. Abdelkader and Griffin [23] presented a new method that matched N face regions against M face images to form a normalized feature vector by utilizing the facelt algorithm. Karhunen-Loeve transform was used for dimensionality reduction and SVM and FLD for classification, with an accuracy of 94.2%. Makinen and Raisamo [24] compared four gender classification algorithms and four automatic alignment techniques with manually aligned and nonaligned faces. It demonstrated that resizing face image size after or before alignment could affect the classification accuracy. It applied the SVM technique to the image size of 36x36 pixels, achieving the best accuracy of 84%. Yang and Ai [25] suggested an approach that is used the local binary patterns histograms (LBPH) feature and used Chi-Square distance of sample for LBPH feature as a confidence measurement for classification. It employed the AdaBoost technique for gender classification and achieved an accuracy of 93%. Abbas *et al.* [26] proposed a new 3D geometric descriptor approach to effectively analyse gender depending on geodesic path technique. The geodesic paths between facial landmarks determined curvature features, which is used to classify the Caucasian teenagers' gender and achieved 87.3%.

Some research employed deep learning techniques to address the gender recognition problem in recent years. Agbo-Ajala and Viriri [27] suggested a model based on a convolution neural network to extract deep features from real-life facial images to be classified then to the correct gender. A dropout and augmentation regularization of data was also adopted to reduce overfitting's risk. They utilized the OIU-Audience dataset and achieved an accuracy of 89.7% in classifying gender. Khalifa *et al.* [28] suggested a gender-recognition method based on iris after segmenting this region from background face images utilizing the graph cutting segmentation method. The model contained three convolution layers to extract features and three fully connected layers to classify images. It applied the proposed model on a dataset that includes 3000 images separated equally into males and females and achieved 98.88%. Haider *et al.* [29] proposed a gender classification method based on deep learning techniques, namely "Deepgender". It applied a convolutional neural network that includes four convolutional layers, two fully connected layers, three max-pooling layers, and one multinominal logistic regression layer. The method combined two datasets (FEI and CAS-PEAL) and applied them to pre-process technique to achieve an accuracy of 98%. Duan *et al.* [30] also proposed agender and age classification based on CNNs as a features extractor and external learning machine (ELM) as a classifier, making up a hybrid CNN-ELM algorithm to accomplish the task of gender and age classification. The CNN includes contrast normalization layers, max-pooling layers, and convolutional layers, and connects to ELM structure to classify face images. They tested the model using MORPH-II and Adience Benchmark, with the best performance achieved was 88.2%. Tilki *et al.* [31] proposed a gender classification model based on deep learning techniques. It employed AlexNet and a proposed CNN to classify the gender of the kaggle dataset that includes 2,500 male and 2,500 female face images. The created CNN includes convolution layers, dense and flattened layers and RELU and max-pooling layers. It achieved better accuracy when using the proposed CNN (92.4%) compared with AlexNet (90.5%). Sumi *et al.* [32] also suggested a new CNN model for feature extraction from face images and gender classification. They passed face images through convolution layers, RELU and max-pooling layer to extract features. In the classification part, the extracted features were submitted to the fully connected layer and classifier. This model utilized k-fold cross-validation to optimize the performance and achieved 97.44% and 90% using KAGGLE and Nottingham scan datasets. Dhomne *et al.* [33] also employed a deep convolution neural network in gender classification. particularly VGGNet architecturewas utilized to predict the gender of celebrities' faces dataset and achieved higher accuracy (95%) compared with other CNNs techniques. In this paper, we have proposed a novel gender-race classification method that has applied a geodesic path algorithm for discriminative feature extraction of the FERET dataset that contains various ethnic groups of both genders. It achieved higher performance in classifying both gender and ethnicity in comparison with SOA methods mentioned above.

3. THE PROPOSED MODEL

The proposed race-gender model has conducted several steps during experimental testing and development as shown in Figure 1. The face images are preprocessed to detect the face region which is converted to grayscale image before being applied to the system. Then, discriminative features are extracted and represented as face vectors using geodesic distance technique before applying PCA to reduce the dimensions of face vectors. Finally, SVM and KNN techniques have been applied to classify these face vectors according to their gender or race.

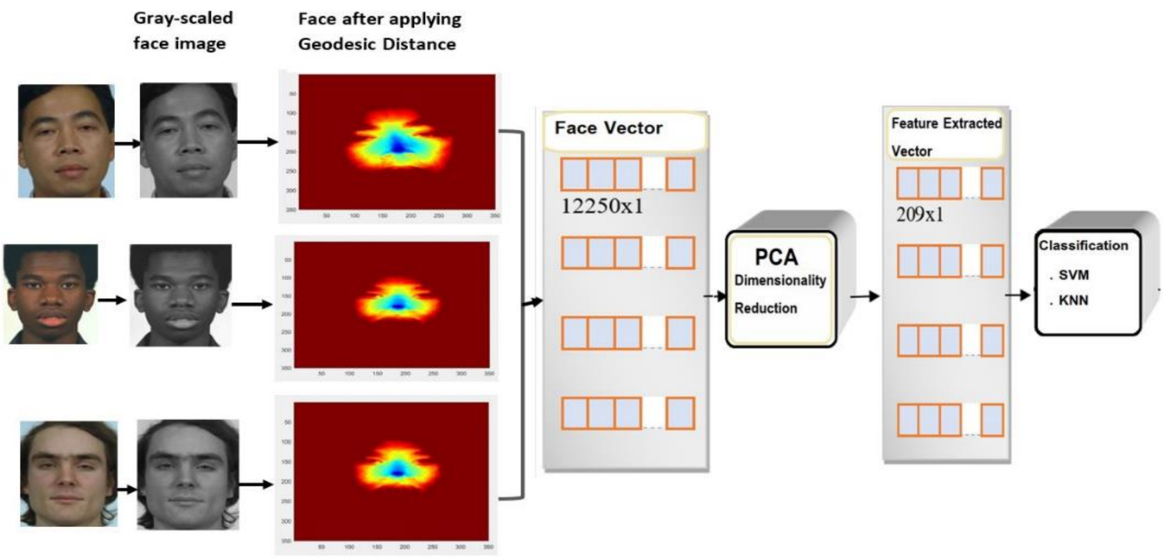

Figure 1. Steps of race and gender classification model

3.1. Dataset

Different dataset has been employed for testing and classifying the demographic features such as gender and ethnicity. In this work, the FERET dataset was utilized for the development and testing of the proposed classifier. This dataset is collected by the national institute of standards and technology (NIST) and utilized by 460 research. This research has derived a new dataset from the FERET dataset, which included 209 facial photos of different ethnic groups. The collected dataset composes 102 and 107 males and females, respectively, and three ethnic groups: Asian, African, and Caucasian. The male facial images include 34, 33, 35 of African, Asian, and Caucasian, respectively. The female subset includes 34 African, 30 Asians, and 43 caucasians.

3.2. Pre-processing algorithm

Pre-processing step is important to prepare the data to be fed to the proposed model. Converting coloured images to grayscale is a part of data preparation. Then, Face detection technique has been utilized to centralize and extract the face region from background for all 2D grayscale-images in the chosen dataset.

3.3. Geodesic path technique

To apply the geodesic path technique on pre-processed face images, it should select a reference point for each image. This point was selected and located on the tip of the nose of each face in this work. It is also known as fiducial point, P0. Figure 2 expresses the steps of detecting the fiducial point (nose tip).

Figure 2. Pre-processing and applying geodesic distance

3.4. Applying geodesic path technique

The geodesic path $G_{P0,\ PN}$ represents the shortest line between the reference point P0, the tip of nose in this work, and another point PN of a 2D face surface, as seen in (1).

$$GP0, P \ = MinimumPath\ (P0, PN) \tag{1}$$

This research considered Dijkstra's algorithm to calculate the geodesic path and distance between the referenced point and other points on the face surface. In this algorithm, the following parameters are defined:

- Distance matrix [D], in which $D(P_0) = 0$ for the referenced point; $D(P_N)=\infty$ for facial surface's other points. The distance matrix's values are updated as follows: for each new point $P_{new.}$
- If $D(P_{old})$+weight $(P_{new,}\ P_{old})<D(P_{new})$, where weight $(P_{new,}\ P_{old})$ is obtained from the value of the P_{new} and $P_{old,}$ then it has to obtain a new minimum path and update the value of $D(P_{new})$ to a new minimal point.
- Otherwise, $D(P_{new})$ has not been updated.

In the end, the algorithm has visited all points on the facial surface and obtained the shortest path ($G_{P0,\ P}$) from the source image. Figure 3 illustrates the processed face image after applying the geodesic path technique to face the image. By repeating the geodesic paths' computations between reference point P0 and all other points PN on the facial surface of the dataset, a high-dimensional matrix H is constructed of all geodesic paths GP, as seen in (2).

$$[\mathrm{H}] = \begin{bmatrix} GP11 & \cdots & GP1m \\ \vdots & \ddots & \vdots \\ GPn1 & \cdots & GPnm \end{bmatrix} \tag{2}$$

The above geodesic path matrix is not suitable for practical purposes since it is in a high dimensional space. In this work, face images are transmitted into 350×350 matrices and generate consequently the same size after applying geodesic path matrices. This problem is projected by applying a dimensional reduction technique to reduce the matrices dimensions to lower-dimensional space. The dimensionality reduction technique used in this work is principal component analysis (PCA).

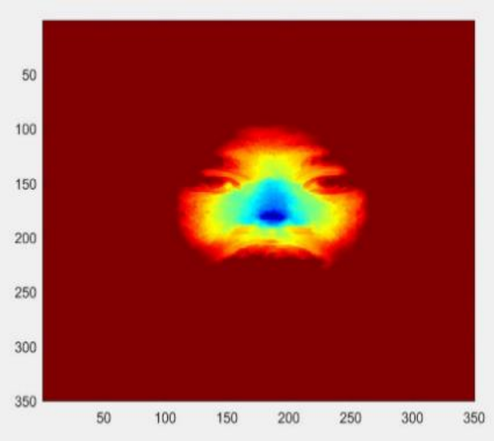

Figure 3. Face image after applying geodesic path technique

3.5. Principal component analysis (PCA)

PCA is a technique commonly utilized for dimensionality reduction purposes [34]. This procedure removes unimportant and repeated data and concentrated on significant features. This algorithm represents the original data in a lower dimension space than the original data while reducing information loss. For applying the PCA algorithm, a face vector of size 122,500×1 is created from each face matrix, with the dataset size entered to PCA being 12,250×209. The output matrix of PCA was 209×209, with each column demonstrating a face vector. This new matrix is now ready for classification.

3.6. Classification

The next step is to classify the extracted features of all facial images in the dataset. The dataset is divided into a training set to train the proposed model and testing to forecast the test data based on what the model learned from the training set. The classification techniques used in this work are KNN and SVM.

3.6.1. K-nearest neighbour technique (kNN)

K-NN has been repeatedly employed in various classification and regression problems due to its simplicity and efficiency. Consequently, it is one of the top ten data mining algorithms. The kNN is a supervised machine learning approach that predicts testing set labels depending on the k most similar training data in the feature space [35]. kNN algorithm considers two significant parameters: the k value's selection and the distance measurement. The value of k is tuned to achieve the optimum classification results. Measuring the distance require a suitable distance function, which is Euclidean distance in this research.

3.6.2. Support vector machine (SVM)

Another supervised machine learning algorithm to project the classification problems is SVM. It creates decision boundaries to gain an optimal hyperplane, that helps in separating the feature space into different labelled classes. As it is a non-linear classification problem, SVM utilizes kernel function instead of a linear one [36]. The penalty rate is another SVM's parameter that can be efficient when misclassification in the training set happens by strictly splitting data of different classes.

3.7. Statistical analysis

Analysis of variance (ANOVA) is a statistical test used for identifying variances in group means or samples between different models. It is an essential mechanism utilized by researchers in various studies, e.g., medical applications and clinical diagnosis, to compare between approaches and models. In this research, we employed ANOVA to verify if the differences between our method and the former models in gender and race classifications were important or not. Therefore, ANOVA statistically analyzed the efficiency of geodesic distance in feature extraction of face images. The ANOVA's basic terms are explained as follows:

- Grand mean: ANOVA technique used two mean types, the grand mean obtained by calculating all observation's mean and the different distinct sample means.
- Hypothesis: it is a suggestion for an unsolved problem depending on forcing a specific argument which is probably rejected by experiments and observation. ANOVA commonly used an Alternative and null hypothesis. The alternative hypothesis is rejected, and the null hypothesis is supported when sample

means are equal and no distinguishable difference among them. On the other hand, the alternative hypothesis is accepted in case of a noticeable difference between samples means.
- Between-Group Variability: it means variations between allocations in each group as separate groups have various values.
- Within-Group Variability: it denotes the differences within each group and does not consider any interface between samples.

4. RESULTS AND DISCUSSION

In this work, two main experiments have been conducted. The first experiment concentrates on classifying the face samples according to the gender (male or female). The second experiment composes classifying the samples according to their ethnicity into three ethnic groups: African, Asian, and Caucasian.

4.1. Gender-based classification

This experiment concentrated on classifying the dataset according to the gender of individuals (male or female). The dataset was labelled as “male” and “female”, with 209 samples totally, 102 male and 107 females. The classification performance is evaluated using a confusion matric, which has four basic terms: i) true negative (TN) and true positive (TP): these parameters represent the numbers of males and females that are correctly classified in their true gender and ii) False-positive (FP) and false-negative (FN): denote how many males and females were incorrectly estimated as a different gender. The classification accuracy is calculated in (3).

$$Accuracy = \frac{(TN+TP)}{N} \times 100\% \tag{3}$$

Where N is the total number of both gender in the dataset.

The misclassification rate (error rate) refers to the number of samples which are misclassified as different labels. It is defined in (4). The classification results utilizing KNN and SVM allowed both genders to be accurately estimated. Figure 4 shows the result of confusion matrix which demonstrates that all 107 female samples and 102 male samples are accurately predicted. This means that the accuracy rate was 100%, while the misclassification rate was 0%.

$$Error\ rate = \frac{(FN+FP)}{N} \times 100\% \tag{4}$$

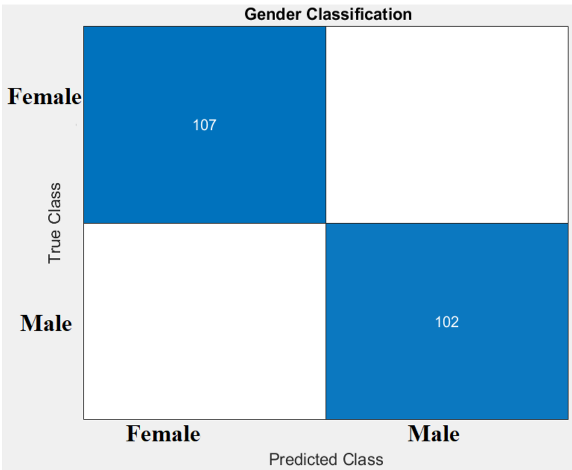

Figure 4. Confusion matrix after applying SVM and KNN classification techniques to classify gender

4.2. Race-based classification

This experiment focused on classifying the dataset based on their ethnicity. As mentioned before, the dataset includes three ethnic groups. The dataset was consequently labelled as “Asian”, “African”, and “Caucasian”, with 63 Asian, 68 African and 78 Caucasian. The confusion matrix has also been employed to evaluate the classification performance. As there are three classes in this experiment, the confusion matrix is 3x3, with each column representing the "true classes" and each row referring to the "predicted classes". The matrix's diagonal elements signify the true values (TP) of each class. Figure 5 presents the three-class confusion matrix with classes (1: Caucasian, 2: Asian, 3: African). For example, TP1 represents the true positive (TP) of class 1 or "Caucasian", and F21 and F31 are the samples of class 1 or "Caucasian" that are misclassified as 2 and 3 ("Asian" and "African", respectively). Thus, the false-negative (FN) of class 1 or Caucasian is the sum of F21 and F31, which represents all samples of class Caucasian incorrectly predicted as Asian or African. Basically, FN (False negative) is the sum of errors in the row while FP (False positive) of any anticipated class is the sum of a column related to that class label. The accuracy is calculated depending on true positive values of all classes as seen in (5).

$$Accuracy\_Race = \frac{(TP_1+TP_2+TP_3)}{N} \times 100\% \quad (5)$$

where 1 is Asian, 2 is African, 3 is Caucasian, African, and Asian, and N is number of samples.

The result of classification of both SVM and KNN illustrates that all samples are correctly predicted in their ethnic groups. The confusion matrix in Figure 6 shows that 68 Africans, 63 Asians and 78 Caucasians are accurately predicted. This means that the accuracy rate is 100% and the error rate is zero.

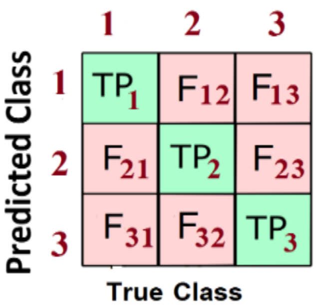

Figure 5. Three-class confusion matrix

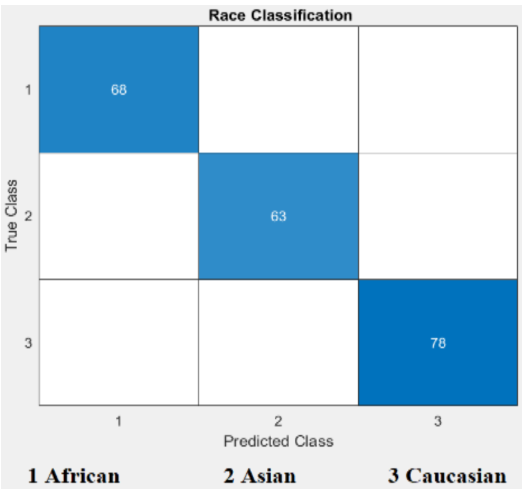

Figure 6. Confusion matrix after applying SVM and KNN classification techniques to classify ethnicity

4.3. Discussion

Our proposed method has achieved higher performance than previous studies that addressed gender, ethnicity or both problems. The first experiment, which aimed to classify the gender of selected face samples of FERET, achieved very high performance (about 100%) compared with SOA studies. Some methods, such as in [6]–[8], were proposed to classify both gender and ethnicity and adopted various techniques for feature extraction and classification. For example, in [7], Gober features were extracted from different regions of the face, while Lyle *et al.* [8] focused on extracting features from the periocular region only using LBP feature-based to be classified then. Buchala *et al.* [20] also utilized various face regions, including eyes and mouth, to extract two regions' features before classification using SVM; However, they attained less performance than our method. Many studies recently concentrated on deep learning techniques for gender classification purposes, as shown in recent studies [27]–[33]. Khalifa *et al.* [28], the study focused on extracting deep features from the iris region using CNN proposed for this study and achieved better performance compared with other studies (about 98.8%). Haider *et al.* [29] proposedthe 'DeepGender' technique based on CNN, which includes four convolution layers, three max-pooling layers and two fully connected layers to classify gender and produced 98%. Compared to our study, the geodesic path technique demonstrates higher performance in extracting suitable features for gender classification. Similarly, [27], [30] also proposed a new CNN structure for gender classification but realized lower accuracies (89.7% and 88.2%). Some studies used popular CNN techniques such as VGG in [33] and AlexNet in [31] but fulfilled lower accuracies than ours (95.5% and 90.5%, respectively), see Table 1.

The second experiment conducted an ethnicity classification problem, and the comparison with related works on race classification shows the superiority of the geodesic path technique in race classification over SAO studies. Xie *et al.* [9], the researchers focused on the periorbital region to extract discriminative features utilizing facial colour-based features and filtered responses. However, it achieved 95.6% in the best-case scenario. In comparison, the study proposed in [11] extracted features from many regions (such as lip colour, forehead region) using the Viola-Jones algorithm. However, the classification into three groups (Caucasian, Negroid, Mongolian) resulted in only 81%. Hosoi *et al.* [10] proposed a new technique to extract race-based features by collaborating Gabor wavelet features and retina sampling, but it fulfilled 94.33%. Deep learning also has been increasingly employed in race classification. Baig *et al.* [12] proposed a new CNN, which included a combination of convolutional layers and subsampling layers, to integrate skin, colour, and facial surface features. It classified face images into Asian and Non-Asia and achieved 84.91%, which is far from our method's performance. Khan *et al.* [13] also proposed a DCNN which includes four convolutional layers and four max-pooling layers followed by two fully connected layers. It achieved 93.2%, 92%, 99.2 and 100%, respectively, when classifying VMER, VNFaces, CAS-PEAL and FERET datasets. Although it achieved an accuracy of 100% in the FERET dataset, it classified data into just two classes (Asian and Non-Asian). Masood *et al.* [14] compare ANN and CNN and demonstrated better performance (98.6%) when utilizing the VGGNet convolution neural network, which is still less than our result. Vo *et al.* [15] also compared two CNNs models: RRCNN and RRVGG. However, the best performance, which was achieved when testing on different datasets was only 91%. Greco *et al.* [16] utilized four well-known CNNs: VGG16, MobileNet, ResNet 50, VGG-face. Nevertheless, they could not achieve more than 94.1%. Heng *et al.* [17] combined the image ranking engine with the classification. Result of CNN. However, utilizing hybrid feature vectors into classification attained lower accuracy (95%) than applying the geodesic technique to extract features, see Table 2.

Table 1. Comparison between our proposed method and related methods on gender classification

Method	Accuracy result
Guo and Mu [7]	98%
Manesh *et al.* [6]	94%
Lyle *et al.* [8]	93%
Gutta *et al.* [18]	96%
Moghaddam *et al.* [19]	97%
Buchala *et al.* [20]	85.5%
Singh *et al.* [21]	95.5%
Bekhouche *et al.* [4]	88.8
Balci and Atalay [22]	92%
Abdelkader and Griffin [23]	85%
Makenen *et al.* [24]	84%
Yang *et al.* [25]	93%
Agbo-Ajala and Viriri [27]	89.7%
Khalifa *et al.* [28]	98.88%
Haider et al. [29]	98%
Duan *et al.* [30]	88.2%
Tilki *et al.* [31]	92.4%
Sumi *et al.* [32]	97.44%
Dhomne *et al.* [33]	95%
Proposed method	100%
Method	Accuracy result

Table 2. Comparison between our proposed method and related methods on race classification

Method	Accuracy result
Lu and Jain [5]	96.3%
Manesh *et al.* [6]	98%
Guo and Mu [7]	99%
Lyle *et al.* [8]	93%
Xei *et al.* [9]	96.5%
Hosoi *et al.* [10]	94.33%
Roomi *et al.* [11]	81%
Biag *et al.* [12]	84.91%
Khan *et al.* [13]	93.2, 92, 99.2, 100
Masood *et al.* [14]	98.6%
Vo and Le [15]	91%
Greco *et al.* [16]	94.1%
Heng *et al.* [17]	95.2%
Proposed method	100%

4.4. ANOVA results

The results of the race and gender-based classification models were presented to ANOVA software to verify if utilizing geodesic path technique for feature extraction can achieve classification improvement. In the case of ethnicity classification, the accuracies of nineteen previous models have been used (see Table 3). In terms of gender classification, thirteen prior models were compared with our method. MS represents the mean square error, and F value refers to the ratio of between-group to within-group variability. P-value represents the differences' probability arising by random chance. The enhancement of the proposed model is important and improbable to happen by chance if the P-value is less than 0.05. As seen in Table 3, the P-value is about 0.0001 in the race classification model and almost zero in the gender classification model; this signifies that the improvement attained using geodesic path technique as feature extractor is important.

Table 3. ANOVA result of gender and race classification

	F	P-Value	MS
Previous gender-based models against our model	46.28	5.95517e-08	505.379
Previous race-based models against our model	20.62	0.0001	294.841

5. CONCLUSION

This work proposed a new model that predicts both gender and ethnicity problems. This model utilized the geodesic path technique to acquire suitable features for better discrimination of gender and race groups. PCA reduced the dimensionality of the extracted features' matrix without information's loss. SVM and KNN are the classification techniques being used to identify gender (male and female) and distinguish between three race groups Asian, Caucasian, and African. In this paper, different SOA studies about race and gender classifications have been presented. The experimental work shows that the proposed model demonstrates the highest performance compared with the related studies with all samples in the dataset being correctly anticipated.

REFERENCES

[1] A. Swaminathan, M. Chaba, D. K. Sharma, and Y. Chaba, "Gender classification using facial embeddings: a novel approach," *Procedia Comput. Sci.*, vol. 167, pp. 2634–2642, 2020, doi: 10.1016/j.procs.2020.03.342.

[2] C. Zhang, H. Ding, Y. Shang, Z. Shao, and X. Fu, "Gender classification based on multiscale facial fusion feature," *Math. Probl. Eng.*, vol. 2018, pp. 1–6, Nov. 2018, doi: 10.1155/2018/1924151.

[3] S. A. Khan, M. Ahmad, M. Nazir, and N. Riaz, "A comparative analysis of gender classification techniques," *Int. J. Bio-Science Bio-Technology*, vol. 5, no. 4, pp. 223–243, 2013.

[4] S. E. Bekhouche, A. Ouafi, A. Benlamoudi, A. Taleb-Ahmed, and A. Hadid, "Facial age estimation and gender classification using multi level local phase quantization," in *2015 3rd International Conference on Control, Engineering & Information Technology (CEIT)*, May 2015, pp. 1–4. doi: 10.1109/CEIT.2015.7233141.

[5] X. Lu and A. K. Jain, "Ethnicity identification from face images," in *Proocedings of SPIE*, 2004, doi: 10.1117/12.542847.

[6] F. S. Manesh, M. Ghahramani, and Y. P. Tan, "Facial part displacement effect on template-based gender and ethnicity classification," in *2010 11th International Conference on Control Automation Robotics & Vision*, Dec. 2010, pp. 1644–1649. doi: 10.1109/ICARCV.2010.5707882.

[7] G. Guo and G. Mu, "A framework for joint estimation of age, gender and ethnicity on a large database," *Image Vis. Comput.*, vol. 32, no. 10, pp. 761–770, Oct. 2014, doi: 10.1016/j.imavis.2014.04.011.

[8] J. R. Lyle, P. E. Miller, S. J. Pundlik, and D. L. Woodard, "Soft biometric classification using periocular region features," in *2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, Sep. 2010, pp. 1–7. doi: 10.1109/BTAS.2010.5634537.

[9] Y. Xie, K. Luu, and M. Savvides, "A robust approach to facial ethnicity classification on large scale face databases," in *2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, Sep. 2012, pp. 143–149. doi: 10.1109/BTAS.2012.6374569.

[10] S. Hosoi, E. Takikawa, and M. Kawade, "Ethnicity estimation with facial images," in *Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings.*, pp. 195–200. doi: 10.1109/AFGR.2004.1301530.

[11] S. M. M. Roomi, S. L. Virasundarii, S. Selvamegala, S. Jeevanandham, and D. Hariharasudhan, "Race classification based on facial features," in *2011 Third National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics*, Dec. 2011, pp. 54–57. doi: 10.1109/NCVPRIPG.2011.19.

[12] T. I. Baig *et al.*, "Classification of human face: asian and non-asian people," in *2019 International Conference on Innovative Computing (ICIC)*, Nov. 2019, pp. 1–6. doi: 10.1109/ICIC48496.2019.8966721.

[13] K. Khan, R. U. Khan, J. Ali, I. Uddin, S. Khan, and B. Roh, "Race classification using deep learning," *Comput. Mater. Contin.*, vol. 68, no. 3, pp. 3483–3498, 2021, doi: 10.32604/cmc.2021.016535.

[14] S. Masood, S. Gupta, A. Wajid, S. Gupta, and M. Ahmed, "Prediction of human ethnicity from facial images using neural networks," *Data Engineering and Intelligent Computing*, 2018, pp. 217–226. doi: 10.1007/978-981-10-3223-3_20.

[15] T. Vo, T. Nguyen, and C. Le, "Race recognition using deep convolutional neural networks," *Symmetry (Basel).*, vol. 10, no. 11, p. 564, Nov. 2018, doi: 10.3390/sym10110564.

[16] A. Greco, G. Percannella, M. Vento, and V. Vigilante, "Benchmarking deep network architectures for ethnicity recognition using a new large face dataset," *Mach. Vis. Appl.*, vol. 31, no. 7–8, p. 67, Nov. 2020, doi: 10.1007/s00138-020-01123-z.

[17] Z. Heng, M. Dipu, and K. H. Yap, "Hybrid supervised deep learning for ethnicity classification using face images," in *Proc. - IEEE Int. Symp. Circuits Syst.*, vol. 2018-May, 2018, doi: 10.1109/ISCAS.2018.8351370.

[18] S. Gutta, H. Wechsler, and P. J. Phillips, "Gender and ethnic classification of face images," in *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 194–199. doi: 10.1109/AFGR.1998.670948.

[19] B. Moghaddam and M. Yang, "Learning gender with support faces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 707–711, May 2002, doi: 10.1109/34.1000244.

[20] S. Buchala, N. Davey, R. J. Frank, T. M. Gale, M. J. Loomes, and W. Kanargard, "Gender classification of face images: the role of global and feature-based information," 2004, pp. 763–768. doi: 10.1007/978-3-540-30499-9_117.

[21] V. Singh, V. Shokeen, and B. Singh, "Comparison of feature extraction algorithms for gender classification from face images," *Int. J. Eng. Res. Technol.*, vol. 2, no. 5, pp. 1313–1318, 2013.

[22] K. Balci and V. Atalay, "PCA for gender estimation: which eigenvectors contribute?," in *Object recognition supported by user interaction for service robots*, vol. 3, pp. 363–366. doi: 10.1109/ICPR.2002.1047869.

[23] C. BenAbdelkader and P. Griffin, "A local region-based approach to gender classification from face images," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops*, 2005, vol. 3, pp. 52–52. doi: 10.1109/CVPR.2005.388.

[24] E. Makinen and R. Raisamo, "Evaluation of gender classification methods with automatically detected and aligned faces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 3, pp. 541–547, Mar. 2008, doi: 10.1109/TPAMI.2007.70800.

[25] Z. Yang and H. Ai, "Demographic classification with local binary patterns," in *Advances in Biometrics*, 2007, pp. 464–473. doi: 10.1007/978-3-540-74549-5_49.

[26] H. Abbas, Y. Hicks, D. Marshall, A. I. Zhurov, and S. Richmond, "A 3D morphometric perspective for facial gender analysis and classification using geodesic path curvature features," *Comput. Vis. Media*, vol. 4, no. 1, pp. 17–32, Mar. 2018, doi: 10.1007/s41095-017-0097-1.

[27] O. Agbo-Ajala and S. Viriri, "Face-based age and gender classification using deep learning model," 2020, pp. 125–137. doi: 10.1007/978-3-030-39770-8_10.

[28] N. Khalifa, M. Taha, A. Hassanien, and H. Mohamed, "Deep iris: deep learning for gender classification through iris patterns," *Acta Inform. Medica*, vol. 27, no. 2, p. 96, 2019, doi: 10.5455/aim.2019.27.96-102.

[29] K. Z. Haider, K. R. Malik, S. Khalid, T. Nawaz, and S. Jabbar, "Deepgender: real-time gender classification using deep learning for smartphones," *J. Real-Time Image Processing*, vol. 16, no. 1, pp. 15–29, Feb. 2019, doi: 10.1007/s11554-017-0714-3.

[30] M. Duan, K. Li, C. Yang, and K. Li, "A hybrid deep learning CNN–ELM for age and gender classification," *Neurocomputing*, vol. 275, pp. 448–461, Jan. 2018, doi: 10.1016/j.neucom.2017.08.062.

[31] S. Tilki, H. B. Dogru, A. A. Hameed, A. Jamil, and J. Rasheed, "Gender classification using deep learning techniques," *Manchester J. Artif. Intell. Appl. Scienes*, vol. 2, no. 2, pp. 126–131, 2021, [Online]. Available: https://mjaias.co.uk/mj-en/article/view/31.

[32] T. A. Sumi, M. S. Hossain, R. U. Islam, and K. Andersson, "Human gender detection from facial images using convolution neural network," in *Communications in Computer and Information Science*, 2021, pp. 188–203. doi: 10.1007/978-3-030-82269-9_15.

[33] A. Dhomne, R. Kumar, and V. Bhan, "Gender recognition through face using deep learning," *Procedia Comput. Sci.*, vol. 132, pp. 2–10, 2018, doi: 10.1016/j.procs.2018.05.053.

[34] S. M. Shaharudin, N. Ahmad, and S. M. C. M. Nor, "A modified correlation in principal component analysis for torrential rainfall patterns identification," *IAES Int. J. Artif. Intell.*, vol. 9, no. 4, p. 655, Dec. 2020, doi: 10.11591/ijai.v9.i4.pp655-661.

[35] S. Zhang, X. Li, M. Zong, X. Zhu, and D. Cheng, "Learning k for kNN classification," *ACM Trans. Intell. Syst. Technol.*, vol. 8, no. 3, pp. 1–19, Apr. 2017, doi: 10.1145/2990508.

[36] I. Wirasati, Z. Rustam, J. E. Aurelia, S. Hartini, and G. S. Saragih, "Comparison some of kernel functions with support vector machines classifier for thalassemia dataset," *IAES Int. J. Artif. Intell.*, vol. 10, no. 2, p. 430, Jun. 2021, doi: 10.11591/ijai.v10.i2.pp430-437.

BIOGRAPHIES OF AUTHORS

Zahraa Shahad Marzoog received the B.Sc. degree in computer and software engineering from Al- Mustansiriya University, Iraq, the M.Sc. degree in software engineering (management) from the University of Queensland, Australia. Now, she served as a lecturer in Electrical and Electronic Engineering department at Karbala university. Her research interest includes image processing, pattern recognition, machine learning, and deep learning. She can be contacted at email: zahraa.marzoog@uokerbala.edu.iq.

Ashraf Dhannoon Hasan Asst lecture of Engineering at University of Kerbala. He received his B.Sc. degree in computer engineering from Al_Hussain University, Iraq and M.Sc. degree in computer engineering from Ferdowsi University, Iran. His research interests include image processing, computer network design, computer hardware design, computer vision. He can be contacted at email: ashraf.dh@uokerbala.edu.iq.

Prof. Dr. Hawraa Hassan Abbas Doctor of Engineering. Professor at University of Kerbala a Ph.D. from Cardiff University/UK. She received her B.Sc. degree in computer engineeringfrom Baghdad University, Iraq and M.Sc. degree in computer engineering also from Baghdad University, Iraq. Her research interests include 3D face modeling, classification of facial traits, image processing, computer network design, genetic associations, computer vision. She can be contacted at email: hawraa.h@uokerbala.edu.iq.