A machine learning approach for ethnic classification: the British Pakistani face

Jilani, Shelina Khalid; Ugail, Hassan; Bukar, Ali M.; Logan, Andrew; Munshi, Tasnim

A Machine Learning Approach for Ethnic Classification: The British Pakistani Face

S. K. Jilani*, H. Ugail, A. M. Bukar
Centre for Visual Computing
University of Bradford
Bradford, UK.
S.Jilani@bradford.ac.uk

A. Logan
Optometry and Vision Science
University of Bradford
Bradford, UK.

T. Munshi
School of Chemistry
University of Lincoln
Lincoln, UK.

Abstract— Ethnicity is one of the most salient clues to face identity. Analysis of ethnicity-specific facial data is a challenging problem and predominantly carried out using computer-based algorithms. Current published literature focusses on the use of frontal face images. We addressed the challenge of binary (British Pakistani or other ethnicity) ethnicity classification using profile facial images. The proposed framework is based on the extraction of geometric features using 10 anthropometric facial landmarks, within a purpose-built, novel database of 135 multi-ethnic and multi-racial subjects and a total of 675 face images. Image dimensionality was reduced using Principle Component Analysis and Partial Least Square Regression. Classification was performed using Linear Support Vector Machine. The results of this framework are promising with 71.11% ethnic classification accuracy using a PCA algorithm + SVM as a classifier, and 76.03% using PLS algorithm + SVM as a classifier.

Keywords— Ethnicity; Pakistani; Classification; Face; Machine Learning .

I. Introduction

The human face includes a range of cues which form part of an individual's identity. Accordingly, faces are important for recognition and identification of specific individuals. In the context of a crime, this presents the following important challenge: how can the perpetrator's face be accurately extracted and described from the memory of a witness? One approach is to rely upon an artist to replicate the face based on verbal descriptions from witnesses. Having a witness verbally describe the complex spatial configuration of the perpetrator's face, often after a period, limits the amount of face information which can be captured within a schematic representation [1]. Further, previous studies suggest that there is significant variability in the recall of unfamiliar faces (e.g. gender difference) and considerable differences in face feature encoding strategies [2] [3].

In recent years, the use of demographic information such as age, gender and ethnicity [4] [5] [6] for facial image classification, within the machine-learning framework, has drawn increased research attention. Practical applications range from fields such as policing (law enforcement) and disaster victim identification (DVI) [7] to biometrics [8]. Significant advances have been made in feature extraction from front view facial images for ethnicity assignment. Ethnic classification based on profile images, however, is limited to profile silhouettes [9] [10].

While there have been previous investigations of South Asian faces, there is a paucity of evidence regarding the British Pakistani face. To bridge this gap, we present our approach for a two-class (British Pakistani vs other ethnicity) classification task. To our knowledge, this is the first attempt to use profile facial images for classification of ethnicity.

The proposed approach for ethnicity classification from profile images consists of three fundamental components: *(i)* geometric feature based extraction using marked landmarks, *(ii)* Dimensionality reduction using Principal Component Analysis (PCA) and Partial Least Square Regression (PLS) and *(iii)* demographic (ethnicity) classification using Support Vector Machine (SVM) algorithm as a binary, two-class procedure. (See figure 1)
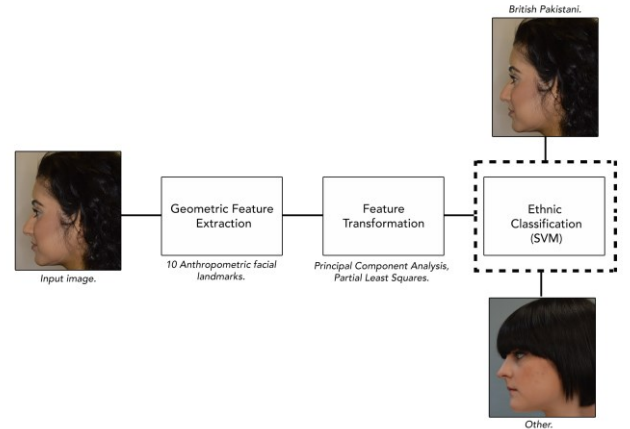
Fig. 1: Proposed framework for binary ethnicity classification task of the British Pakistani face from profile images.

Our aim is to develop a system for accurate ethnic classification of the British Pakistani face based on profile-view face images. Importantly, we have generated an original database, which is specific to our objectives and consists of participants from diverse racial and ethnic backgrounds.

II. Methodology

Geometric feature-based extraction is based on the detection of facial features such as the eyes, nose, mouth and chin.

Properties of the detected features plus their relation to one another (such as distances and angles) are used as the primary descriptors of the face. This method is a pre-processing step and removes distracting variance from a dataset. This enables the classifier to perform efficiently. Features of interest within each facial image are extracted using landmarks, which are positioned on the boundary of important components of the face.

Our approach to this form of face representation is via the use of salient anthropometric landmarks such as; Trichion *(tr)*, Glabella *(g)*, Nasion *(n)*, Pronasale *(prn)*, Subnasale *(sn)*, Labiale Superius *(ls)*, Stomion *(sto)*, Labiale inferius *(li)*, Sublabiale *(sl)* and Pogonion *(pog)*. Each shape within the training set is represented by a 2-dimensional vector, which subsequently represents both x and y coordinates of each marked landmark (x_i, y_i) [11].

Given an n set of 2 dimensional landmarks, the face can be represented by a single linear equation,

$$Fp = (x_1, x_2,...,x_n, y_1, y_2,...,y_n)^T \qquad (1)$$

Fig. 2: Geometric feature based annotation of 10 anthropometric facial landmarks from profile images (highlighted with blue crosses) *(Participant consent has been obtained).*

Dimensionality Reduction

To discard redundant data, reduce computation time and retain the most significant features that capture ethnic variations amongst each facial image, we utilised two different dimensionality reduction techniques. This feature compression was achieved by two methods: Principal Component Analysis (PCA) [11] and Partial Least Square regression (PLS) [12]. Principal Component Analysis seeks to reduce the dimensionality of the data by locating the Principal Components (PCs) of the original variables, which include the largest variance, and can be classed as an unsupervised model.

Partial Least Square regression (PLS) is a supervised dimensionality reduction algorithm that captures the variation that exists between the dependent and independent variables [13]. Although, like PCA, PLS regression is more powerful in its regression abilities because it searches for components that capture the highest degree of variance in x, in addition to the direction which best describes x and y. Whereas, PCA works

to find the direction of highest variance only in x, that subsequently result in the principal components (PCs) best describing x [11] [18].

Classification

Having used geometric feature based extraction to highlight the pertinent areas of the face, followed by the subsequent reduction in the dimensionality of the facial features, we enter the third stage of our proposed framework. At this stage, we employ a Support Vector Machine (SVM) as a linear classifier. Previous studies suggest that SVM is successful as a binary classifier and operates by defining an optimum separating hyperplane between two classes of data [13]. In our work, the two classes are British Pakistani and other ethnicity. The algorithm classifies data by defining an optimum separating hyperplane between two classes of the data, which typically involves solving an optimization problem.

Given a training set;

$$\left\{ (x_i,\, y_i) \right\}_{1 \le i \le n},\, x_i \in \mathbb{R}^d,\, y_i \in \{+1,\, -1\}, \qquad (2)$$

SVM finds the OSH by solving

$$\begin{cases} \min_{w,b} & \frac{1}{2} \parallel w \parallel^2 \\ \text{with} & y_i.(w.x+b) \ge 1 \end{cases} \qquad (3)$$

for $i = 1 \ldots n$ are the observations, where the weights w and b the bias is learned during training. The classifier is learned such that, $y_i \in \{+1,\, -1\}$, $+1$ denotes facial images labelled as British Pakistani and -1 denotes other ethnicity images. Thus, using a Support Vector Machine algorithm, an optimum separating hyperplane was computed, which separated the classes into the two categories.

III. EXPERIMENTAL DESIGN

Facial Image Database

While there is a wealth of face image databases available, such as the Glasgow Unfamiliar Face Database [14], SCFace Database [15] and others, they are not suitable to address our aims. As a result, we created a database of facial images of students (mean age: 35 years old ± 17) from the University of Bradford. This study was granted approval from the Ethics Committee of the University of Bradford.

Through a bank of three cameras, the capture system takes five images from differing angles; front view, right view at 45 degrees, right profile, left view at 45 degrees and left profile (see figure 3). For image capture participants were asked to maintain a neutral facial expression.

Fig. 3: Graphic of The Halo (capture system) at the University of Bradford provided in collaboration with Acumè Forensic, UK.

Images are taken using three 18 MP (megapixel) UI-3590LE cameras with a Tamron lens. Participants are illuminated using a total of six LED panels mounted onto the camera banks. And those wearing glasses were asked to remove them.

The database consists of 5 colour images for each of the 135 participants, totaling to 675 images (figure 4). However, the current study uses profile images only. Participants varied in ethnic and racial backgrounds and included Black (Nigerian), Caucasian, Middle Eastern (Egyptian, Kurdish and Jordanian) and South Asian (Bengali, Gujarati and Pakistani) (see figure 4). To reduce computational time the original dimensions of the images (4000 x 6016 pixel) were cropped to 1500 x 2100 pixels.

Fig. 4: Example of facial images captured with The Halo and used for ethnic classification. *(Participant consent has been obtained).*

Ethnic classification

In our experiments to achieve a two-class ethnicity classification, the number of principal components k, were chosen by selecting the eigenvectors that accounted for 95% of the variance. According to the curve seen as part of Figure 5, k_{pca} is equivalent to 10. In comparison, for our second set of experiments, the number of PLS latent vectors k_{pls} were selected via cross validation.

To evaluate the accuracy of the ethnicity classification algorithm, Receiver Operating Characteristics (ROC) Curve was generated. The curve provides a graphical comparison of the true positive and false positive rates (TPR and FPR) at different thresholds. It was generated to demonstrate the

performance of our binary classifier and evaluate its accuracy [16]. A key advantage of using ROC curve analysis is that the optimal point of interest (the best trade-off between sensitivity and specificity) can be chosen for the investigated variable. TPR and FPR can be calculated as,

$$TPR = \frac{True\ Positive}{True\ Positive + False\ Positive} \qquad (4)$$

$$FPR = \frac{True\ Negative}{True\ Negative + False\ Negative} \qquad (5)$$

According to our ROC curve (figure 6) the experiment with PLS outperformed the experiment with PCA, demonstrating its strength as an algorithm. This is illustrated by the position of the PLS curve: it is shifted to the left of the red curve for PCA.
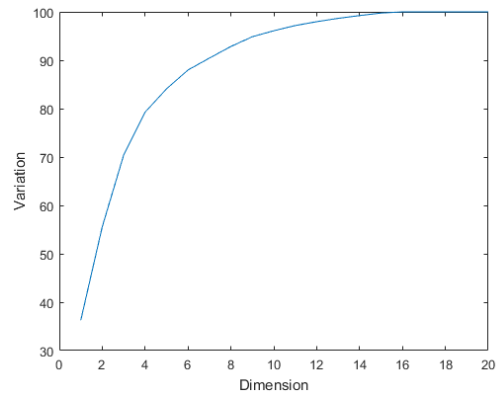

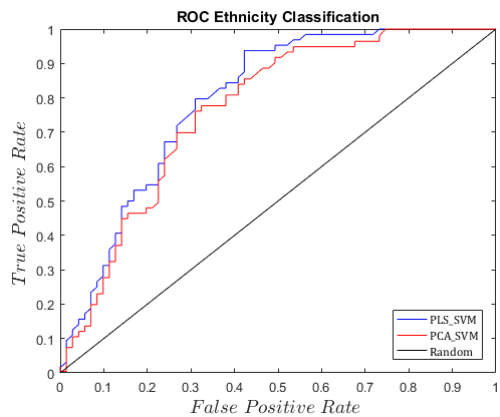**Fig. 5:** Curve to show the variance for k_{pca}.

Fig. 6: Receiver Operating Characteristics (ROC) Curve for PLS and PCA experiment.

IV. RESULTS

We carried out a two-class ethnicity classification task using profile facial images. As a preprocessing step, geometric feature-based extraction was carried out using 10 anthropometric facial landmarks. This was followed by dimensionality reduction with both Principal Component

Analysis (PCA) and Partial Least Square (PLS) regression. We investigated PCA because previous reports highlight its reliability in classifying facial images by race [17].

Support Vector Machine (SVM) approach is highly regarded as an algorithm because of its utility as a separator. Since the ethnic classification task was binary, the selected classifier was SVM. To evaluate the algorithm, we generated a Receiver Operating Characteristics (ROC) Curve to distinguish PCA and PLS results.

The classification results from the two experiments are shown as part of table 1. We calculated percentages of the total number of test images, which were accurately classified as either British Pakistani or another ethnicity. Tariq, Hu and Huang (2009) also investigated ethnicity classification, [9] hence a comparison is shown as part of table 2.

TABLE I.   ETHNICITY CLASSIFICATION USING PCA AND PLS.

	Principal Component Analysis	Partial Least Squares
Percentage Accuracy (%)	71.11%	76.03%

TABLE II.   COMPARISON OF RESULTS ON ETHNICITY CLASSIFICATION.

	Our Work	Tariq et al., (2009)
Percentage Accuracy (%)	**PCA** - 71.11% **PLS** - 76.03%	71.66%.

Our work is distinct since we have used discriminative facial landmarks to assist classification, whereas silhouetted face profiles were used by Tariq et al. (2009). Moreover, we used a Support Vector Machine (SVM) algorithm for classification, while k-nearest neighbors algorithm (KNN) was employed by Tariq et al. (2009). The choice of classifier may have impacted the achieved results since SVM is far greater in binary classification than KNN. And while the reported results of Tariq et al. (2009) are lower when compared to ours, they have been narrated as averages. [9]. For a fairer comparison of results, it would be appropriate to request their testing dataset.

V. CONCLUSION

This paper has investigated ethnicity classification based on profile facial images of a specific demographic (the British Pakistani face). We examined different machine learning frameworks for the two-class ethnic classification task: PCA and PLS. Both approaches achieved promising results (71.11% and 76.03%). Moreover, we have generated a novel facial image database. Owing to the limited availability of research into ethnicity classification, specifically for British Pakistani facial images, we propose further work with the use of Linear Discriminant Analysis (LDA). To establish what, if any, discriminatory features are present. We hypothesize that LDA may lead to improved ethnicity classification accuracy. Additionally, we propose to test the different facial image perspectives for accurate ethnicity classification.

REFERENCES

[1] H. D. Ellis, J. W. Shepherd, and G. M. Davies, "The deterioration of verbal descriptions of faces over different delay intervals," *Journal of Police Science and Administration,* vol. 8, no. 1, pp. 101-6, 1980.

[2] J. K. Hall, S. B. Hutton, and M. J. Morgan, "Sex differences in scanning faces: Does attention to the eyes explain female superiority in facial expression recognition?," *Cognition & Emotion,* vol. 24, no. 4, pp. 629-637, 2010.

[3] Y. Fu, G. Guo, and T. S. Huang, "Age synthesis and estimation via faces: A survey," *IEEE transactions on pattern analysis and machine intelligence,* vol. 32, no. 11, pp. 1955-1976, 2010.

[4] G. Guo, and G. Mu, "Joint estimation of age, gender and ethnicity: CCA vs. PLS." pp. 1-6.

[5] X. Lu, and A. K. Jain, "Ethnicity identification from face images." pp. 114-123.

[6] S. Hosoi, E. Takikawa, and M. Kawade, "Ethnicity estimation with facial images." pp. 195-200.

[7] N. Chhaya, and T. Oates, "Joint inference of soft biometric features." pp. 466-471.

[8] A. Dantcheva, C. Velardo, A. D'angelo, and J.-L. Dugelay, "Bag of soft biometrics for person identification," *Multimedia Tools and Applications,* vol. 51, no. 2, pp. 739-777, 2011.

[9] U. Tariq, Y. Hu, and T. S. Huang, "Gender and ethnicity identification from silhouetted face profiles." pp. 2441-2444.

[10] N. Davidenko, N. Witthoft, and J. Winawer, "Gender aftereffects in face silhouettes reveal face-specific mechanisms," *Visual Cognition,* vol. 16, no. 1, pp. 99-103, 2008.

[11] I. K. Fodor, "A survey of dimension reduction techniques," *Center for Applied Scientific Computing, Lawrence Livermore National Laboratory,* vol. 9, pp. 1-18, 2002.

[12] C. F. Dormann, J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J. R. G. Marquéz, B. Gruber, B. Lafourcade, and P. J. Leitão, "Collinearity: a review of methods to deal with it and a simulation study evaluating their performance," *Ecography,* vol. 36, no. 1, pp. 27-46, 2013.

[13] G. Chierchia, N. Pustelnik, and J. C. Pesquet, "Random primal-dual proximal iterations for sparse multiclass SVM." pp. 1-6.

[14] "York FaceVar Lab," http://www.facevar.com/downloads .

[15] P. Tome, J. Fierrez, R. Vera-Rodriguez, and D. Ramos, "Identification using face regions: Application and assessment in forensic scenarios," *Forensic science international,* vol. 233, no. 1, pp. 75-83, 2013.

[16] J. Lever, M. Krzywinski, and N. Altman, "Points of Significance: Classification evaluation," *Nature Methods,* vol. 13, no. 8, pp. 603-604, 2016.

[17] P. J. Phillips, F. Jiang, A. Narvekar, J. Ayyad, and A. J. O'Toole, "An other-race effect for face recognition algorithms," *ACM Transactions on Applied Perception (TAP),* vol. 8, no. 2, pp. 14, 2011.

[18] A. M. Bukar, H. Ugail, and D. Connah, "Automatic age and gender classification using supervised appearance model," *Journal of Electronic Imaging,* vol. 25, no. 6, pp. 061605-061605, 2016.