International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

A Profound Multitask System for Gender Identification face recognition, Confront Discovery, Point of interest Localization, and Head Position Estimation Hyperface

Mrunal Padalkar1, a , Tejal Shirsat1, b , Snehal Vekhande1, c , and Dr.Ekta Sarda 2, d
1 Department of Computer Engineering, Ramrao Adik Institute of Technology, D Y Patil Deemed to be University, Navi Mumbai, India
2 Asst Prof, Department of Computer Engineering, Ramrao Adik Institute of Technology, D Y Patil Deemed to be University, Navi Mumbai, India.

Abstract

Machine learning is a technology that has risen in its usage and popularity in the last few years. A huge number of people from around the world are learning this technology and putting the knowledge to various use. Machine learning algorithms are capable of learning from the provided data with high accuracy. Even though a significant amount of research has been conducted on face recognition, the integrated model of face recognition, landmark localization, head posture estimation, and gender identification that is capable of high accuracy and speed has not yet been investigated. As a result, we have developed a face recognition system that can make predictions about photos that are comparable to those made by humans. The principal component analysis PCA and the SVM were used here to accomplish facial recognition. In feature extraction, to reduce the dimensionality of large datasets, principal component analysis is performed. After the data have been preprocessed, they are entered into the SVM classifier to be used for image classification. The study of this is done via visualization, and it is used to measure the effectiveness of the model. This face recognition algorithm has an accuracy of at least 80% when it comes to classifying people's portraits. The findings of the experiments show that the suggested technique can successfully identify faces since it employs a feature-based algorithm that combines PCA classification and SVM detection.

Keywords: - Feature Extraction, SVM, Dimensionality reduction PCA, Posture Detection.

Introduction

Face recognition, which has many applications including image analysis and interpretation, is one of the most successful and interesting uses of computer vision. In this field, it has taken decades of research, increased access to high-quality data sets, and a significant increase in computer power to achieve accuracy that is close to that of a person. Higher-order computer vision tasks, such as facial identification and facial expression analysis, depend heavily on accurately detecting landmarks inside face images[8,9,10]. It's not just computer scientists that are curious about this issue; neuroscientists and psychologists are as well. Developments in computer vision will help psychologists and neuroscientists better understand the human brain. As information technology evolves, so does the need for increased safety and security. A lot of research has gone into facial recognition for security concerns. Using Principal Component Analysis and Support Vector Machines, we created a face recognition model in this study.

PCA is used to reduce the number of dimensions for recognizing faces[9]. Every image in the training set is modelled using PCA as a linear combination of unique weighted eigenvectors. Eigenfaces from principal component analysis have been employed in our facial recognition pre-processing. We went with PCA since it's a flexible method that can be use with many different kinds of information. It works well with massive data sets. As its name implies, PCA is only used to convert a high-dimensional data space to a lower-dimensional one.

Using PCA and SVM, we were able to achieve accurate facial recognition. This study examined the use of principal component analysis to address the issue of face recognition and proposed an eigenfaces-based system. The Python implementation of the technique has been tested on the Labeled Faces picture in the Wild - dataset. With an accuracy of 80% or better, our Face Recognition Model can assign labels to pictures of people.

- To be able to detect faces without much knowledge of facial features.
- To detect faces from static images.
- To illustrate the importance of principle components to identify an image.
- To implement fast and easy to understand face detection model.
- To give maximum accuracy.

In recent years, machine learning and artificial intelligence have gone ahead by leaps and bounds. One such application of this technology is hyperface. Hyperface is a model which combines powerful techniques from machine learning and artificial intelligence for face recognition. Hyperface has garnered widespread attention for their uses in various fields. Even though there has been a lot of research on face recognition, the best way to combine face recognition, landmark localization, head pose estimation, and gender recognition is still not known. Face recognition is easy and natural for people, but it's not so simple for a machine. Face recognition should be done on a computer in this digital world. This would help make the process faster and more accurate. This leads us to study how machine learning can be used to recognise faces. So, it gives us a reason to work on this project.

Literature Survey

Wu, H., Zhang, K., and Tian, G. have used feature fusion to improve CNN architecture in their paper Detection of faces and posture estimation simultaneously utilizing a cascade of convolutional neural networks [10]. First, they used cascaded networks to find faces quickly and accurately. In the last step, they estimated the poses of the last fine faces. Then, add multitask learning to each CNN architecture on its own.

In their paper Real-time HyperFace: Multitask Face Analysis using Region Proposal Network Vaibhav Gandhi and Aditya Verma made a face recognition system that works in real time by using R-CNN, random forest, and a multi-task function[2,7]. The middle layers of a convolution neural network have been joined together. Face recognition gets an AUC score of 0.87, and gender recognition gets a score of 0.97.

R Ranjan, V.M. Patel, and R. Chellappa proposed a single CNN model that can find all of the above features at the same time[12]. They start with image classification, where the CNN's features are spread out in a hierarchical way. Then, the training is done on the dataset, which is made up of pictures from the real world[1]. During the testing phase, they use a forward-pass through the hyperface network to guess the task labels.

In the paper "Multi-view Face Detection," written by Farfade, S. S., Saberian, M. J., and Li, L. J., Deep Convolutional Neural Networks: [3,10] has used deep convolutional neural networks to classify faces from multiple views and pull out their features. By making the architecture simpler[17], they have cut down on the complexity. Because of this, the model is very good at telling people apart.

Face Detection, Pose Estimation, and Landmark Localization in the Wild [5] by X. Zhu and D. Ramanan is based on a mix of trees with a shared pool of parts. They have modelled every feature of the face as a separate part and use global mixtures to show changes in the shape of the face[13]. The benefit of this model is how easy it is to code for elasticity and three-dimensional structure.

In the paper Face Facial Landmark Detection And Application [9], Kostiantyn Khabarlak and Larysa Koriashkina did research on the above topic. They found that having the algorithm do more than one thing makes it more reliable.

Multi-view Face Detection Using Deep Convolutional Neural Networks is not feasible to detect occluded and rotated faces. Face Detection, Pose Estimation, and Landmark Localization in the Wild is only trained with 900 faces[14]. Real Time face Detection is difficult to Achieve because of Slow performance of the DEL System. HyperFace: A Deep Multi-Task Learning Framework for Face Detection, Landmark Localization, Pose Estimation, and Gender Recognition Contain Small Distorted and blurry Images which result in low detection. Face Facial Landmark Detection and Application takes more time than usual. Algorithm performs slower than Real time. Viola Jones Cannot Operate On 45 Degree Angle. Lightning has an effect in this Algorithm.Multiple Detection of same Face is a Problem.

Proposed System

Building a machine learning based Facial recognition model for the given static image as input and to process and identify the name of person that the image belongs to from the stored database. In this digital world, though it is a very simple and intuitive task for humans to recognize faces it is quite difficult for machine to recognize them[6]. Therefore, this model will help the machine to recognize people with 80% or higher accuracy. Dimensionality reduction is accomplished using principle component analysis, and picture classification is handled by a Support vector machine. The dataset used to train this model is "Labeled faces in the wild."

The methodology works like any other machine learning problem. The processing will start from Preprocessing, Feature Extraction Dimension Reduction and Classification. For each image (in training and testing) all these tasks will be performed as shown in fig 3.1

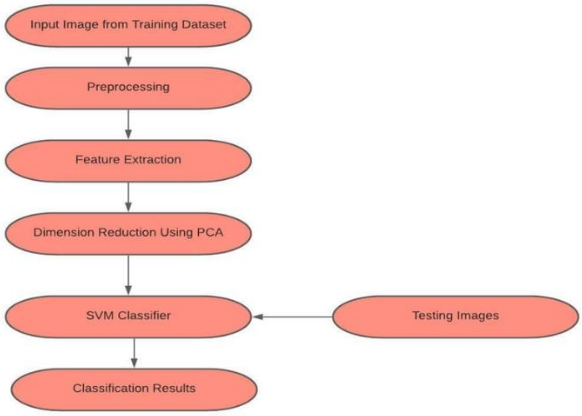


Fig 3.1 Proposed Methodology

In this paper, utilizing the Labelled Faces in the Wild - dataset, we created a face recognition system used for PCA and SVM. Dimensional simplification is goal for a Principal Component Analysis Dimensionality is reduced by projecting the training data into a compact feature space, which is done with the help of Eigenvalues and Eigenvectors. As mentioned earlier we are using Labeled Faces in the wild dataset. This dataset contains face photographs for studying face recognition problems. its dataset contains almost 14,000 images for different looks on the Internet. There is a name for every single face in the collection. Each person in the dataset has several photos in various poses and human gestures.

1. To make future processing of a picture
easier, a technique called "preprocessing" is conducted on the data. Assuring that all resource data is presented in a standard, usable manner is the job of preprocessing. Since our data consists of photographs, it is crucial that they all have at least one common dimension; otherwise, it would be impractical to work with the dataset. The images may also have different levels of illumination[15]. We standardize both features of all images. That is why we normalize the sizes of all the images in the collection and move them all into the same place. Light is another issue that grayscale conversion helps us tackle. After these two adjustments, however, each image is still stored as a 2D array in our system. Each picture is turned into a tuple where each index represents the total number of items in the tuple.

2. Feature Extraction is nothing but a procedure where we decide the targeted data which we help us identify each image uniquely and act as unique feature of that image. These features will help us classify each image into various classes and recognize a particular person from his image[16]. Image is nothing but a combination of pixels, every image is represented as number of pixels present in height and number of pixels present in the width of the image (ex. 50px x 37px). Each pixel is nothing but a number in a scale 0-255. We have total 1280, images in our dataset and each of size (50px x 37px) that means to uniquely identify an image all these pixels play an important role. These are features of that image. That means for 1280 images each image has 50x37 = 1850 features. Which is a huge number. We simply cannot take 1850 features for each image to recognize it, there is a great need to reduce these features, but while reduction process we hope not to lose information. That is the reason of performing dimensionality reduction.

3. Reducing the number of features, extracting the most relevant information, and discarding the rest is what's known as dimensionality reduction. In this approach, there will be no data loss or diminished functionality. The most common method for accomplishing Dimensionality reduction is known as Principal Component Analysis.

Every picture will be turned into a vector first.

We then take the mean of these facial vectors and remove it from each individual vector. After that, we'll turn each face vector into a matrix. The covariance matrix will be shorn down to its essential elements.

The eigenvalues and eigenvectors of this covariance matrix are determined by this method.

The k greatest eigenvalues' corresponding eigenvectors are chosen.

Each face vector is then represented as a linear combination of the best K eigen vectors.

4. At this point, we have simplified representations of all images. Using a Support Vector Machine, we assign categories to these pictures. SVM stands for support vector machine and is a straightforward supervised learning technique for both classification and regression. Primarily, it serves as a useful tool for organizing information. SVM Essentially, it locates a hyper-plane that divides the data into distinct groups. When using SVM, each dataset item is plotted on an N-dimensional space, N denotes the total number of characteristics. The best possible hyperplane dividing the information. Since our data contains seven distinct categories, the SVM classifier we've trained with the Training Images will attempt to assign each new image to one of those categories.

5. Lastly, we put our trained model to the test by attempting to predict new (Test) data. To train on an Unknown Face y, we must first preprocess the face to make it

Figure 3.3 Design the workflow of the system

symmetrical and of the same size as the training face. e take the mean of the eigenvalues of the face and remove the face. The linear combination of Eigenfaces is then projected back into Eigenspace and a vector is produced. Subtracting the training vector from the training images yields the least distance between the training field and the testing field. Furthermore, distance influences whether a particular face is recognized.

The Design of System represented in figure 3.3 depicts the journey of each image from testing dataset. First Removal

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

of Mean Face takes place then the image is converted into face vector. This face vector is Normalized and projected into Eigen Space. This gives the Weight Vector of that image.

The Comparison of input images weight vector and all the weight vectors of training set is done by classifier that will enable the Support Vector Classifier to classify the input image in one of the seven classes and finally if the image is classified in one the class means it is recognized otherwise it is not known to the model.

Implementation

Stage 1:

We will import all the libraries necessary for processing.

Stage 2:

We will first load the dataset as NumPy a

array by using fetch_lfw_people function from sklearn.dataset. we will observe data structure and classes present in the data, and

list down the classes and the size of data Grayscale pictures (pixel values between 0 and 255)

The Variable Explorer shows that there are 1850 features over a height of 50px and width of 37px in a total of 1288 samples (images).

Stage 3:

Analytical PCA Here are some things to do: Whenever we are presented with a new face y that we don't know anything about, we must first perform some pre-processing to ensure that it is centred in the picture and has the same proportions as the training face. It's time to take the average eigenvalues of the face and remove them from the face. For the linear combination of eigenfaces, we now project the normalised vector into eigenspace. The vector is derived from the projection. To find the smallest gap between the training and test vectors, we begin by subtracting the vector acquired in the preceding step from the training image. If deviation is less than the threshold, the face is recognised from the training image; otherwise, no matches are found in the training set.

Stage 4:

To keep track of the array of data in the dataset, we'll designate it as X. The 1288 samples in X are accompanied with 1850 characteristics apiece. We shall establish the names for the categories to which each picture belongs, or labels. The goal label for each image is denoted by y. The seven names of the persons to be identified are stored in a variable called "name," which provides more definition for the label. The numerical representation of "target" is the letter Y. All the parameters are listed below.

Stage 5:

Now that our data is all preprocessed let us divide them in training and testing images. We have training to testing ratio as 75:25.

Stage 6:

Now the dimensionality of each image in training and testing is 50x37 which are the features of each image, we will reduce the dimensionality of each image in training and testing by calling pca transform function on them.so now the featured reduced from 1850 to 50.

Stage 7:

Classification is the next step. An SVM Classification Model will be Trained. SVMs are a form of supervised machine learning technique that may be used for classification

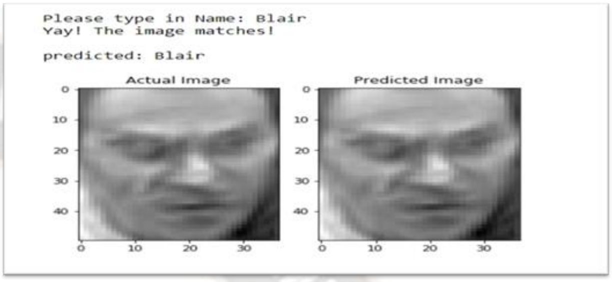


Figure 4.1 The actual image and the predicted image.

and/or regression and are very simple to build. It is most often used for classification, although it may be quite helpful for

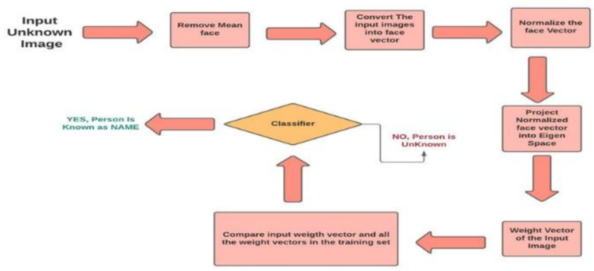


Figure 4.2 Classification report of the system

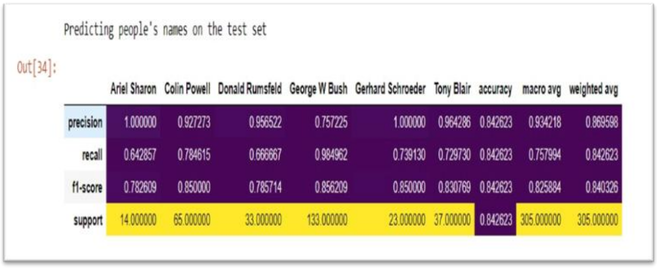
regression in specific situations. To divide the data into distinct categories, SVM locates a hyper-plane. This hyper-plane is equivalent to a straight line in a 2-dimensional space.

SVM involves projecting each dataset item into an N-dimensional space, where N is the

total number of features or characteristics.

Proceed to identifying the best hyperplane

for data partitioning.



Predicting people's names on the test set

Out[34]:

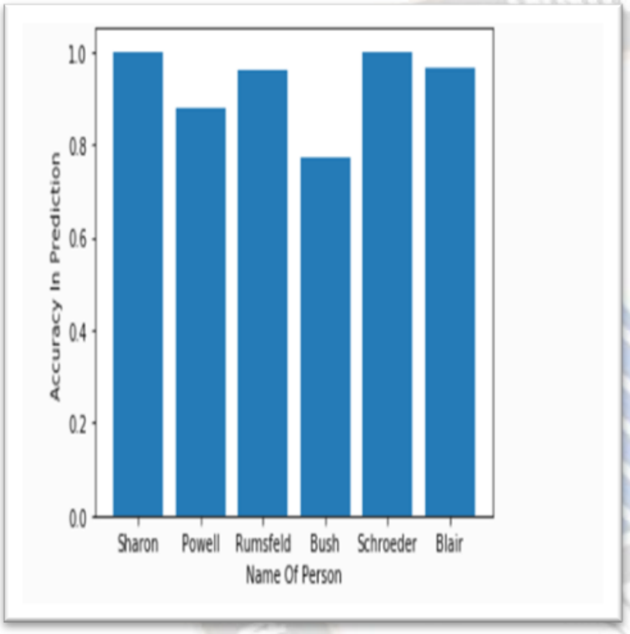
	Ariel Sharon	Colin Powell	Donald Rumsfeld	George W Bush	Gerhard Schroeder	Tony Blair	accuracy	macro avg	weighted avg
precision	1.000000	0.927273	0.956522	0.757225	1.000000	0.964286	0.842623	0.934218	0.869598
recall	0.642857	0.784615	0.666667	0.984962	0.739130	0.729730	0.842623	0.757994	0.842623
f1-score	0.782609	0.850000	0.785714	0.856209	0.850000	0.830769	0.842623	0.825884	0.840326
support	14.000000	65.000000	33.000000	133.000000	23.000000	37.000000	0.842623	305.000000	305.000000

Figure 4.3 Data visualization of the classification model

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Result and Analysis

Initially, we evaluated our face recognition model (HyperFace) on the Labeled-faces in the Wild dataset actual image and predicted image are shown in figure 4.1. Though there are many images of one person like smiling, with and without spectacles, with different orientations of the face, the algorithm correctly identifies the person's face with the correct name of the person in the image. classification report that shows the accuracy, recall, and other "goodness" characteristics of the categorization. This classifier is right on many photos, demonstrating the simplicity of this knowledge model. This model's overall accuracy is estimated to be approximately 84%.



To quantify the effectiveness of our model, as shown in figure 4.2 We created a

Figure 4.3 gives a visual summary of the classification done on the dataset giving the Figure 4.4: Predicted image with different principal component values. clear idea of the classification model.

The model was further tested on input sample apart from dataset. Figure 4.4 depicts the predicted image with different principal component value which shows that the unit-pixels of the predicted image can be set depending on the clarity required with the accuracy is 92%.

Conclusion and Future Work

Thus, we now know how face can be detected using machine learning to get a near human accuracy. The fundamental component of this is face recognition since Hyperface is a deep multi-task learning model that incorporates face detection, landmark localization, head posture estimation, and gender recognition. We used PCA and SVM to create the facial recognition system. SVM is used for classification while PCA is used for feature extraction. The method was tested and implemented in Python using the image labelled Faces in the Wild Dataset as well as additional input images. The developed technique is universal and effective with all kinds of images. The tests are carried out on PNG and JPEG images of various people in various postures. A classification report is also generated to assess the model's effectiveness. This demonstrated that this technique provided extremely good face categorization with an accuracy of roughly 84%, despite its limitations due to image size changes. As a result, the eigenface method offers a realistic solution that is ideally suited to the challenge of face identification. It is quick, simple, and has been proved to operate effectively in a restricted context. This section discusses the Hyperface model. Our project is only a humble venture to satisfy the needs to recognize faces from given dataset and classify the recognized images and to plot the photograph with eigenfaces. This project shall prove to be satisfying for serving as a face recognition system. The objective of this project is to provide a fast and effective face recognition algorithm and should be updated regularly as the project progresses. We will be working on a real-time face recognition module soon. The system can further be extended to encompass other aspects. The system may be expanded to include other features. We worked on some still images for this project, but soon, we hope to build a system that employs a video camera and works with real-time face recognition. We utilised the Labelled Faces in the Wild - dataset in this case. We want to tackle the problem of reliably detecting faces of varying sizes. We will compare the performance of the PCA-based method to all other face-recognition algorithms now in use.

Reference

[1] Ranjan, R., Patel, V. M., & Chellappa, R. (2017). Hyperface: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition. IEEE transactions on pattern analysis and machine intelligence, 41(1), 121-135.

[2] D. Chen, S. Ren, Y. Wei, X. Cao, and J. Sun. Joint cascade face detection and alignment. In European Conference on Computer Vision, pages 109–122. Springer, 2014.

[3] Farfade, S. S., Saberian, M. J., & Li, L. J. (2015, June). Multi-view face detection using deep convolutional neural networks. In Proceedings of the 5th ACM on International Conference on Multimedia Retrieval (pp. 643-650).

[4] M. Kostinger, P. Wohlhart, P. M. Roth, and H. Bischof. Annotated ¨ facial landmarks in the wild: A large-scale, real-world database for facial landmark localization. In Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on, pages 2144–2151. IEEE, 2011.

[5] X. Zhu and D. Ramanan. Facedpl: Detection, pose estimation, and landmark localization in the wild. preprint, 1(2):6, 2015.

[6] Zhu, X., & Ramanan, D. (2012, June). Face detection, pose estimation, and landmark localization in the wild. In 2012 IEEE conference on computer vision and pattern recognition (pp. 2879-2886). IEEE.]

[7] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker. Multi-pie. Image and Vision Computing, 2010.

[8] Vu, H. T., Nhat, M. V. M., & Le, B. (2015, October). An efficient model for simultaneous face detection, pose estimation and landmark localisation. In 2015 Seventh International Conference on Knowledge and Systems Engineering (KSE) (pp. 13-18). IEEE

[9] Kostiantyn s. Khabarlak ,Larysa s. Koriashkina.(2014) Face Facial Landmark Detection And Application

[10] Wu, H., Zhang, K., & Tian, G. (2018). Simultaneous face detection and pose estimation using convolutional neural network cascade. IEEE Access, 6, 49563-49575.

[11]Viola, P., & Jones, M. (2001, December). Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001 (Vol. 1, pp. I-I). IEEE.

[12]] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.

[13] X. Cao, Y. Wei, F. Wen, and J. Sun. Face alignment by explicit shape regression. International Journal of Computer Vision, 107(2):177–190, 2014.

[14] R. Caruana. Multitask learning. In Learning to learn, pages 95–133. Springer, 1998.

[15] D. Chen, S. Ren, Y. Wei, X. Cao, and J. Sun. Joint cascade face detection and alignment. In D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, editors, European Conference on Computer Vision, volume 8694, pages 109–122. 2014.

[16] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models—their training and application. Comput. Vis. Image Underst., 61(1):38–59, Jan. 1995.

[17] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, pages 248–255. IEEE, 2009.