International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 116-122 www.ijaem.net ISSN: 2395-5252

A Comprehensive Overview of Face detection and Face recognition Methods, Techniques and Algorithms

Aniedu A.N[1], Nwokoye C.S[2],Nzeamalu C. A[3],Onwubualili B. C[4],Okafor K.H[5], UcheO.C[6],Avah S.E[7]

Submitted: 25-01-2022 Revised: 05-02-2022 Accepted: 08-02-2022

ABSTRACT- Face detection is a supervised learning (binary classification task) technique used to verify that a captured image contains a face or not. Here there is a label (target) which is compared with the test dataset. A validation set is used to tune the model using the hyper-parameter. Since every human face is unique, complex, and multidimensional and is affected by factors such as age, occlusion, lighting, expressions, and viewpoint, this makes face recognition often difficult. The remedy is to reconstruct a face using the sum of the weights of its eigenfaces[1]. The face can be recognized using the principal component analysis (PCA) method which is a dimensionality reduction method, an unsupervised Machine Learning method. In order words, the model parameter is often used to determine how well the model recognized a face.

Keywords- Face detection, face recognition, principal component analysis, supervised learning, unsupervised learning.

I. INTRODUCTION

Supervised learning is a method of training a model by providing a dataset for which there is already a set of correct answers called labels. The model learns by analyzing patterns with each record and comparing the patterns with the label set. The supervised learning method has two categories which are regression and classification task. The regression task is basically about prediction while classification task seeks to compare two variables and compare it with the label set in order to make a decision. In the classification task, they are of two categories which are binary and multi-classification. Application of binary classification is in image classification, that is: looking at an image and classify it as “face” or “non-face”. Basically for face detection, the supervised learning method can be applied. Unsupervised learning is a method of training a model without the label set but instead the model learns by looking at distinct patterns within the data and making observations.[2]It is applied where the goal is basically to reveal patterns. They are of two categories which are clustering and dimensionality reduction. Dimensionality reduction is the process of simplifying a dataset by eliminating redundant or irrelevant features. Clustering in brief is arranging the dataset in clusters where each cluster is characterized by a contained set of data points and a cluster centroid. A cluster centroid is the mean average of all the data points that the cluster contains across all features. Dimensionality reduction is of two categories which are feature selection and feature extraction. Feature extraction is basically extracting the face features using the eigenspace projection method. This method is simply the vectors that capture the distribution of the face images confined around the entire image locations. Those vectors that capture the distribution must satisfy the eigenvectors of the covariance matrix that are consistent with the initial face images. The eigenface is the eigenvector derived from the principal component analysis.

II. LITERATURE REVIEW

Face recognition is a ubiquitous task. In the early 1970s, face recognition was regarded as fictional until 1971 when Goldstein and et al [5], publishingon “Identification of human faces” revolutionized the area of facial recognition. The first crude attempt proposed 21 subjective facial features such as hair, skin color, lip thickness, etc to identify a face in a photograph.

The drawback of Goldstein approach is that the 21 subjective facial features were manually computed using unsupervised computation and classification.

In 1987, Sirovich and Kirby seminal publication on “A low-dimensional procedure for the characterization of human faces” [6] and Turk and

Pentland 1991 publication on “Face recognition using Eigenfaces” [7] demonstrates a standard linear algebra method for dimensionality reduction termed principal component analysis (PCA) which was used to identify a face using a feature vector less than 100 dimension. Also these principal components also called eigenvectors or eigenfaces can be used to reconstruct faces from the initial dataset. This means that a face can be represented and also identified as a linear combination of the eigenfaces. Feature-based methods such as local binary patterns for facial recognition have also been published and are still applicable to real-world applications.

The Eigenface algorithm applies PCA in constructing a low dimensional representation of the face images.

According to Sirovich and Kirby [7], an eigenface is a face represented as a linear combination of its eigenvectors.

To identify a face, the Euclidean distance is calculated between the eigenface representations and taking the face to be identified as a k-nearest neighbor (K-NN) classification task. A K-NN is simply choosing a value of k and looking at its neighbors, the neighbor with the highest number of vote wins.In 2004, Ahonen et al. [9] presented a paper on “Face recognition with local binary patterns (lbp)” which uses a method of dividing a face image into a 7*7 grid of equally sized cells. Dividing the image into cells introduces some locality into the final feature vector. Additionally some cells are weighted such that they add more weight to the overall representation. Cells in the corner carry less identifying characteristics compared to the cells in the grid which are eyes, nose and lips. But the actual identification is performed by K-NN classification using the distance between the query image and the dataset of labeled faces.

III METHODS AND MATERIALS

Face Recognition System (FRS) has four subsystem which are face detection, face preprocessing, feature extraction and finally face recognition. This is shown in Fig 1.

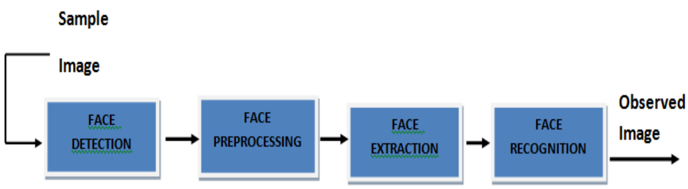

Fig 1:Block Diagram of a Face Recognition System

A. Face Detection

Face detection is a subset of image classification whereby a face is detected in an image using the label (target) provided to the model during training. The target normally contains features such as eyes, nose, mouth, etc. The supervised learning model usually contains three dataset namely; training, validation and the test dataset. The datasets ratio here may be60%, 20% and 20% respectively or as the case may be. The label (target) is removed from the training dataset and kept in the test dataset to prevent over fitting the model. The process of transforming a raw data into features that are useful is called feature engineering. After training the model, if the model under fits then the model is tuned in the validation dataset using the hyper-parameter which can be configured manually. The model is tuned till it can detect a face from the sample image. Face detection uses a binary classification method because it basically has two tasks which are to determine if it is a face or not a face. If the model is to determine if a face has pimple or dimple then this is a multi-label, multiclass task. There are three techniques which can be used to detect the face in the sample image. They are: knowledge-based, feature invariant, and appearance based methods. Knowledge based method use rules to describe a face, feature invariant method uses feature extraction to detect a face, while appearance based method uses machine learning approaches like discriminate functions to learn attributes of facial and non-facial images so as to detect a face.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 116-122 www.ijaem.net ISSN: 2395-5252

Face Detection Algorithm
Start;
Training Image; Validation Image; Test Image;
Test Image= selectpart (Training Image);
Convert (Test Image);
Compare (Test Image, Validation Image);
if (true){
printf (‘Face detected’);
Else printf (‘Face not detected’);
End;
The preceding algorithm can be expressed using a flowchart as seen in fig 2.

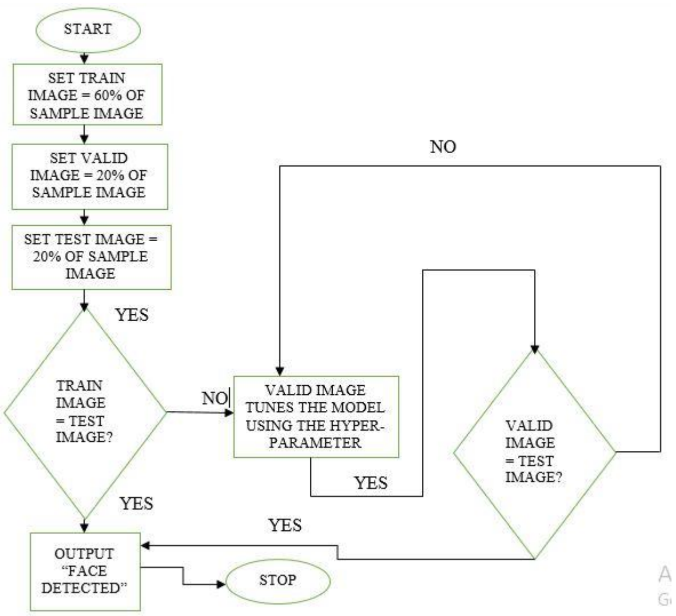

Figure 2: Flowchart for face detection

B. Face Preprocessing

After detecting a face, the next step is facepreprocessing. Face preprocessing removes the factors that causes face variability such as unsteady lightening conditions, image shadows, compression noise, etc. Face preprocessing increases the efficiency of face detection. Dimensionality reduction is a process of simplifying a dataset by eliminating redundant or irrelevant features [3].

Dimensionality reduction has two categories which is feature selection and feature extraction

The feature extraction method used here is the principal component analysis (PCA) method [2]. In the PCA method, also known as the eigenspace projection [3], arevectors that capture the distribution of the face images confined around the entire image locations. Those vectors that capture the distribution

must satisfy the eigenvectors of the covariance matrix that are consistent with the initial face images. The eigenface is the eigenvector derived from the principal component analysis.
Mathematically,
$Var = \Sigma |(r - p)^T . k|^2$
Where vector k_{min} is a unit vector used to minimize the variance (Var) function and $k_{max}$is a unit vector used to maximize the variance (Var) function
$Var = \Sigma k^T(r - p)(r - p)^T . k$
$k^T [\Sigma(r - p)(r - p)^T]. k$
$k^T Zk$ where $Z = \Sigma(r - p)(r - p)^T$
$a = r - p, \; c = (r - p)^T$
Where Z = Covariance matrix of the initial face images (if divided by the number of faces)
p = mean of the face images.
r = training set of face images ($r_1, r_2, r_3, ...,rn$).
a = mean subtracted matrix vector.
c = transpose of the mean subtracted matrix vector.
k = the eigenvector and the eigenvalues of Z.
$k_{max}$is the eigenvector of Z with the largest eigenvalue.
$k_{min}$is the eigenvector of Z with the smallest eigenvalue.
Assuming the initial face images is N-dimensional, the eigenvectors of Z expands to a new coordinate system and hence the eigenvector with the largest eigenvalue has the most variation among the training vectors r while the eigenvector with the smallest eigenvalue has the least variation.
Compressing the training data by using the top few eigenvectors results to selecting a linear subspace and representing its points on a hyper-plane. The top few eigenvectors are known as the principal components.
Assuming the face to be recognized is in a point of high dimensional space, L*M is a point in the H^{LM}, hence It's possible to define vectors in that space.
The face to be recognized is a subspace of a set of images. Hence, fitting a hyper-plane to the unknown face is achieved by using a vector span of $k_1, k_2, v_3, ..., k_n$. If the face is recognized then $r \approx p+c_1k_1+c_2k_2+...+c_wk_w$.
PCA is used to extract the eigenvectors of Z given the span vectors $k_1, k_2, k_3, ..., k_n$.
Each of these span vectors points to a face space.
If $|k| - (p + c_1k_1 + c_2k_2 + ... + c_wk_w) <$ decision boundary (threshold) then the face can be recognized.
PCA is very similar to the Karhunen–Loève Transform (KLT)used in signal processing which is derived as the orthogonal transform with the basis $a = [a_1, a_2, a_3, ..., a_n)$ that any $k \leq N$ minimizes the average L_2 reconstruction error for the data points r,where $L_2$is referred to as ridge. Ridge addresses the case of multi-co-linearity i.e. having multiple prediction values that varies together strongly. It is only suitable for datasets with large features.
While face detection is a supervised learning method, face recognition is an unsupervised learning method. The unsupervised method consists of only the training and the test dataset and their ratio here is 75% and 25% respectively. The main concept here is that the model doesn't have a label (target/set of correct answers) but instead the model looks at the distinct patterns within the face images and makes observations.
The observed image is the identified image which the model derived simply by looking at the distinct patterns within the face images (training and the test images).

C. Face Recognition

Pseudocode of Face Recognition

I. Start
II. Acquire the training Image.
III. Calculate eigenfaces from training image keeping only N images that correspond to k_{max} .These N images define the face space.
IV. Calculate the eigenspace projection in this N-dimensional space for each detected face images by projecting the face images onto this face space.
V. Recognize the new face images
VI. For the detected face images, calculate a set of weights based on the N eigenfaces by projecting each of the detected face images onto each of the eigenfaces.
VII. Determine whether the model can recognize the detected face images or not by checking whether its Euclidean distance with other Face images is less than the decision boundary or not.
VIII. Stop.

Algorithm of Face Recognition

The preceding pseudo code can be expressed as the following algorithm, and flowcharch (Fig. 3)
Start
Training Image; Test Image;
Test Image = selectpart (Training Image);
Compare (Training Image, Test Image);
if (true){
printf ('Face Recognized');
Else printf ('Face not Recognized');
End

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 116-122 www.ijaem.net ISSN: 2395-5252

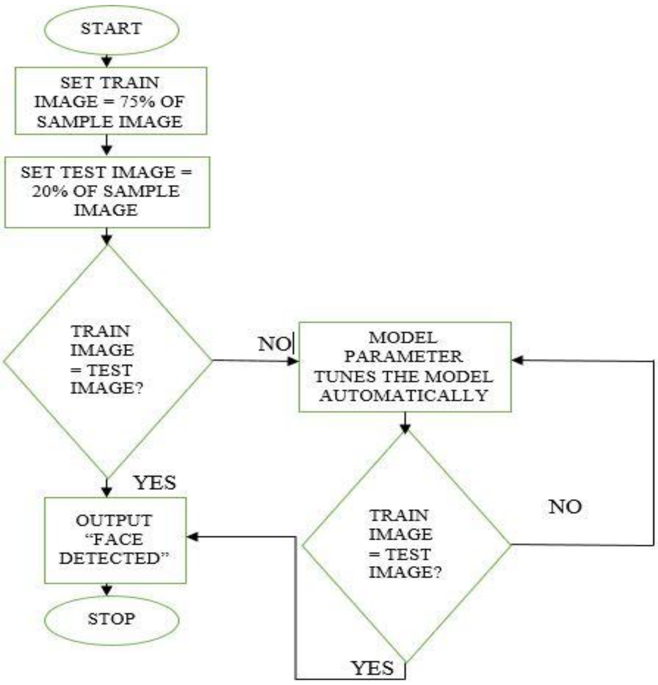

Figure 3: Flowchart of Face Recognition

IV. PERFORMANCE METRICS

For the FRS, the following performance metrics was adopted:

Accuracy

Number of correct prediction as a portion out of all predictions.

Precision

Number of correct positive predictions as a portion out of all positive instances.

Recall

Number of correct positive predictions as a portion out of all relevant instances.

Given a sample image of 20 the confusion matrices are shown in in table 1 and table 2;

Table 1: Confusion Matrix for Face Detection

		Actual	
		Face Detected	No Face Detected
Predicted	Face Detected	5	8
	No Face Detected	12	16

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Where TP = True Positive = 5, TN = True Negative = 16, FP = False Positive = 8, FN = False Negative = 12

$$\text{Accuracy} = \frac{5 + 16}{5 + 16 + 8 + 12} = 0.51 = 51\%$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Precision} = \frac{5}{5 + 8} = 0.38 = 38\%$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{Recall} = \frac{5}{5 + 12} = 0.29 = 29\%$$

Table 2: Confusion Matrix for Face Recognition

		Actual	
		Face Recognized	Face Not Recognized
Predicted	Face Recognized	10	6
	Face Not Recognized	4	13

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Where TP = True Positive = 5, TN = True Negative = 16, FP = False Positive = 8, FN = False Negative = 12

$$\text{Accuracy} = \frac{10 + 13}{10 + 13 + 6 + 4} = 0.70 = 70\%$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Precision} = \frac{10}{10 + 6} = 0.63 = 63\%$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{Recall} = \frac{10}{10 + 4} = 0.71 = 71\%$$

The model detects the sample images with average accuracy because the number of sample images is too small. Accuracy increases as the number of sample image increases. The model is less precise in detecting a face because there are still some factors (occlusion, viewpoint, etc) contained in the sample image which makes it difficult for the model to detect the sample image accurately. But after preprocessing and extracting the relevant features, the model's precision increases. Same goes for the recall.

V. CONCLUSION

In recent times, face recognition and face detection has become popular and its applications become ubiquitous in the field of Artificial Intelligence and Machine learning. In this paper, attempt has been made to define face detection and recognition, and to examine some of the techniques and methods often adopted in its application. Also shown are the algorithms deployed in face detection and face recognition and the performance matrix employed in determining the level of accuracy (or not) of face detection/recognition solutions and deployments.

REFERENCES

[1] Marshall Robinson, Matthew Escarra, Jon Krueger, Doug Kochelek (2004). Face Recognition using Eigenfaces

[2] Azubuike N. Aniedu, Hyacinth C. Inyiama, Augustine C. O. Azubogu, Sandra C. Nwokoye.Pattern Recognition using Support VectorMachines as a Solution for Non-Technical Losses in Electricity Distribution Industry.Vol.7 No.2, March (2021).

[3] Manisha M. Kasar1, Debnath Bhattacharyya1 and Tai-hoon Kim.Face Recognition Using Neural Network: A Review. International Journal of Security and Its Applications Vol. 10, No. 3 (2016), pp.81-100.

[4] Abdelfatah Aref Tamimi1, Omaima N. A. AL-Allaf2, Mohammad A. Alia.Eigen Faces and Principle Component Analysis for Face Recognition Systems: A Comparative Study. International Journal of Computers & Technology Vol 14. No. 4.

[5] Manisha Urkude, Dr. S.B. Kishor and Dr. S.G. Naranje. Face Recogntion Technique (A Literature Survey on Fcae Recognition and insight on Machine recognition using Software Textbook.

[6] Goldstein et al. Identification of human faces (1971).

[7] Kirby, L. S. (Mar 1987). Low-Dimensional Procedure fortheCharacterization of Human Faces. Optical Society of America A, 4,

[8] M. Turk, A. Pentland, Eigen Faces for Recognition, Journal OfCognitive Neuroscience, Vol. 3, No. 1, 1991, Pp. 71-86.

[9] Ahonen et al. Face recognition with local binary patterns (2004).