Effective Usage of Support Vector Machine in Face Detection

Y. Md. Riyazuddin, S. Mahaboob Basha, K. Krishna Reddy, S. Naseera Banu

Abstract: With the rapid growth in Technology in terms of multimedia contents such as Biometrics, Facial recognition etc. Facial detection got much attention over the past few years. Face recognition describes a biometric technology that attempts to establish an identity. In this paper, I would like to review about a facial recognition system using machine learning especially, using support vector machines. In any case, point of this exploration is to give extensive writing survey over face acknowledgment alongside its applications. Furthermore, after top to bottom conversation, a portion of the significant discoveries are given in end.

Keywords: Face recognition; Machine Learning; Image classification; feature extraction; pattern recognition ;SVM; Haar cascade classifier;

I. INTRODUCTION

Face acknowledgment is one among the most wailed advancements in the field of AI. As of late, the utilization cases for this innovation have widened from explicit observation applications in government security frameworks to more extensive applications over different businesses in such undertakings as client recognizable proof and authentication, health and advertising. In fact, the facial recognition market is projected to grow to USD 7.76 billion by 2022, at a Compound Annual Growth Rate (CAGR) of 13.9% (MarketsAndMarkets). In the early barely any years a few papers have been distributed on face discovery in the network which talks about various strategy like neural system, edge indicators andmany more. Before, numerous analysts and architects have structured diverse reason explicit and application explicit detectors.The principle objective of this sort of classifiers was to accomplish a high discovery rate alongside low computational expense. These sort of indicators, for example, corner finders, Haar Cascade classifier are for the most part utilize basic and quick classifiers that reject the most well-known negative examples and afterward they utilize logically increasingly complex classifiers to manage the more troublesome and odd negative samples.And ultimately, facial acknowledgment has surfaced in online life applications on stages, for example, Facebook which recommend clients to label companions who have been distinguished in pictures. Plainly there are numerous applications the utilizations for facial acknowledgment systems.In general the means to accomplish this are the accompanying: face identification, include extraction and in conclusion preparing a model.

∗Dr. Y. Md. Riyazuddin, Department of CSE, GITAM University, Hyderabad, India. E: mail: riyazymd@gmail.com
Dr. S. Mahaboob Basha, Department of CSE, MITS, Madanapalle, India. E: mail: smbasha.phd@gmail.com
Dr. K. Krishna Reddy, Department of Physics, Y V University, Kadapa, India. E: mail: krishna.kkreddy@gmail.com
S. Naseera Banu, Department of CSE, VTU,Belgavi, India. E: mail: Mail: snaseerabanu@gmail.com

II. LITERATURE REVIEW

Right now talk about different ideas like skin edge location, morphological administrators, and bolster vector machine utilized by noticeable creators of face discovery and a geometric face model is framed with the identification of eyes performed utilizing the Haar Cascade Classifier, while nose recognition has been utilized as a reaffirmation system alongside the eyes. Afterward, HOG (Histogram of Oriented Gradients) highlights are separated from enormous quantities of facial pictures to be utilized as a major aspect of the acknowledgment instrument[1]. These HOG highlights are then marked together for a face/client and a Support Vector Machine (SVM) model is prepared to anticipate faces that are encouraged into the framework.

III. SUPPORT VECTOR MACHINES

A Support Vector Machine (SVM) is a discriminative classifier officially characterized by an isolating hyper plane i.e., with a given named preparing information (administered learning); the calculation yields are ideal hyper plane which classifies new models. In 2-D space this hyper plane is a line partitioning a plane in two sections which is known as division of each class on either side.
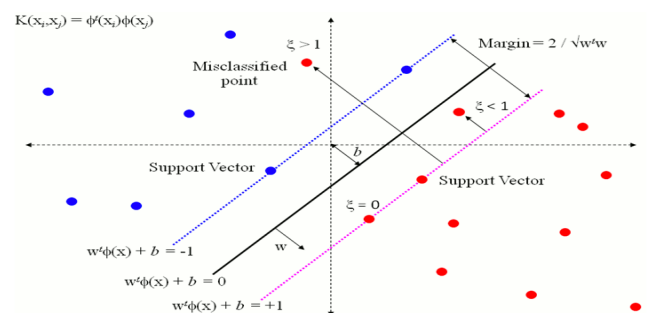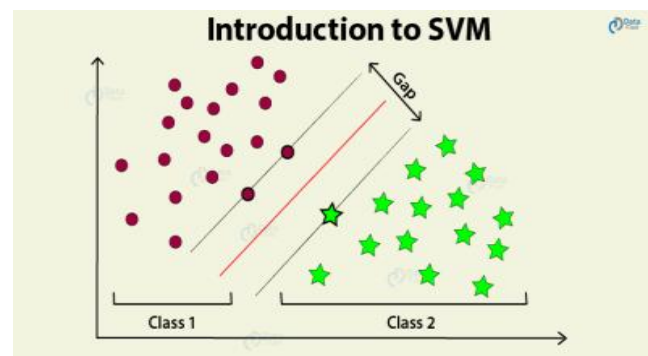

Figure 1: A separable classification toy problem: separate balls from diamonds.

The optimal hyper plane is orthogonal to the shortest line connecting the convex hulls of the two classes (dotted), and intersects it half way. There is a weight vector w and a threshold bsuch that

$yi \cdot ((w \cdot xi) + b) > 0$. Rescaling w and bsuch that the point(s) closest to the hyperplane satisfy $|(w \cdot xi) + b| = 1$, we obtain a form **(w,b)** of the hyperplane with $yi.((w \cdot xi) + b) \geq 1$. Note that the margin, measured perpendicularly to the hyperplane, equals **2/|| w ||**. To maximize the margin, we thus have to minimize |w| subject to $yi ((w \cdot xi) + b) \geq 1$.

Hyperplane classifiers

To design learning algorithms, we thus must come up with a class of functions whose capacity can be computed. SV classifiers are based on the class of hyperplanes $(w \cdot x) + b = 0$ $w \in RN, b \in R$

corresponding to decision functions

f(x) = sign((w·x) + b)

We can show that the ideal hyperplane, characterized as the one with the maximal edge of partition between the two classes (see Figure 1), has the least limit. It tends to be remarkably built by taking care of an obliged quadratic enhancement issue whose arrangement w has an expansionw =∑vixi as far as a subset of preparing designs that lie on the edge (see Figure 1). These preparation designs, called bolster vectors, convey all significant data about the order issue. Discarding the subtleties of the computations, there is only one vital property of the calculation that we have to stress: both the quadratic programming issue and an official choice capacity rely just upon dab items between designs. This is exactly what lets us sum up to the nonlinear case[3].

How does a Haar Cascade Classifier works?

Just like any other form of biometric identification, face recognition requires samples to be collected, identified, extracted with necessary (features) information, and stored for recognition. The entire face recognition solution is divided into following major modules:

1. Face Capture
2. Face Train
3. Face Recognition

Face Detection is the fundamental step in any of the operations carried out in the face recognition process. The Haar Feature-based Cascade Classifier is a widely used mechanism for detecting faces[1]. In order to train a classifier to detect faces, two large sets of images are formed, with one set containing images with faces, and the other set without. These images are then used to generate classifier models. The classifier is generated by extracting Haar features from the positive and negative images.

A Haar classifier is based on an object detection framework proposed by Paul Viola and Michael Jones in their paper, "Rapid object detection using boosted cascade of simple features." A single classifier is trained using each feature shown in the illustration below. However, a single classifier alone does not produce high accuracy, and so multiple such classifiers are cascaded. The final classifier formed is a weighted sum of weak classifiers. Using this method, the classifier provides classification accuracy of more than 95%.

(a) Edge Features

(b) Line Features

(c) Four-rectangle features
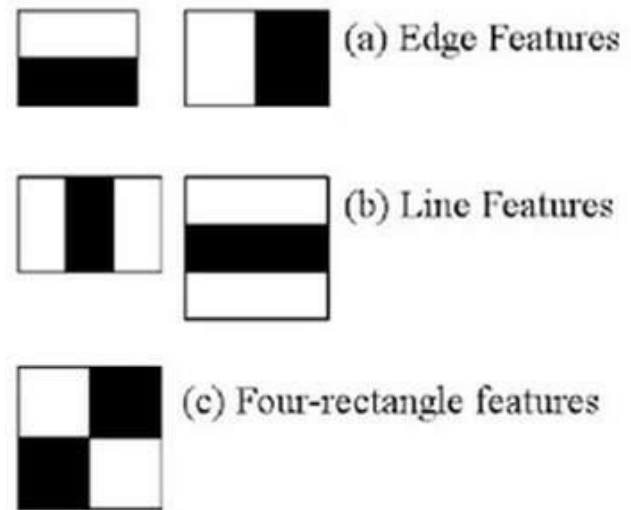
In the model beneath, pictures are utilized as convolutional portions to remove highlights, where each element is a worth acquired by subtracting the summation of pixels under the dark square shape from the summation of pixels under the white square shape.
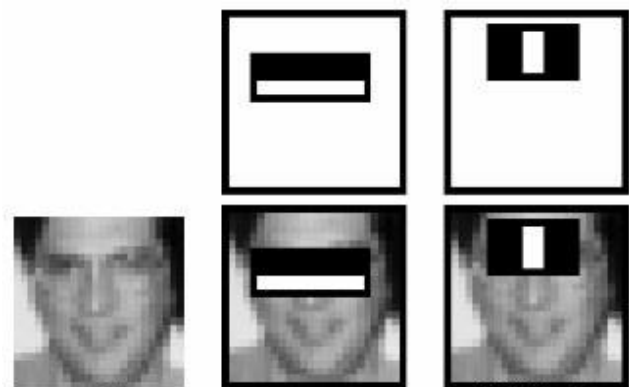
The image above provides examples of two features: edge and line. Edge features efficaciously map the face attribute; i.e., the eye region is darker than any other part of the face. The line feature maps the nose attribute; i.e., the vertical nose line on the face is lighter than the sides[3]. Since, individually, any of these features cannot classify the pattern accurately, they are used in a cascade; hence the name of Haar Feature-based Cascade Classifiers.

Now, let us consider the three major modules of face recognition in detail as discussed below.

Face Capture

The very first step in face recognition is to collect face samples. This is carried out in three basic steps as follows:

1. Detect the face
2. Crop the cardinal section of the face.
3. Save the face image.

The detection of the face is achieved using Haar Feature-based Cascade Classifiers, as discussed in the previous section. Typically, the accuracy of face recognition is highly dependent on the quality and variety of the sample images[4].

The variety of sample images can be obtained by capturing multiple images with multiple facial expressions for the same face.

Once the face is detected, it can be cropped and stored as a sample image for analysis. The ubiquitous use of rectangles to bind regions in an image introduces a superfluous section of the cropped head image. Thus, rectangle-shaped bounded faces obtained using Haar Cascade Classifiers contain insignificant data such as the area surrounding the neck, ears, hair, etc[5]. This can be mitigated using a geometric face model, which is formed using the geometric relationship between the various features within a face, including eyes, nose, and mouth.

How does a geometric face model works?

To form a geometric face model, a pair of eyes is typically considered as the first feature to be located within the image. Ideally, any of the features can be used as a starting point to form a face model, but starting with the location of eyes produces a face model with higher accuracy. In some cases, the location of the nose is used to determine the face model. However, the eyes are typically considered as a primary starting feature, while the nose is considered as a secondary starting feature for situations when the eyes are not located or are partially occluded.
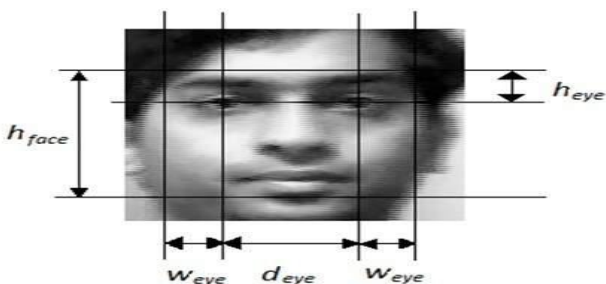
Geometric Face Model using the Eyes:

From the coordinates for the centre of both the eyes, the necessary section (features) of the face is obtained using the equations shown below:

$$h_{face} = K_f\, d_{eye} \quad (eq.\ 1)$$
$$h_{eye} = K_e\, h_{face} \quad (eq.\ 2)$$
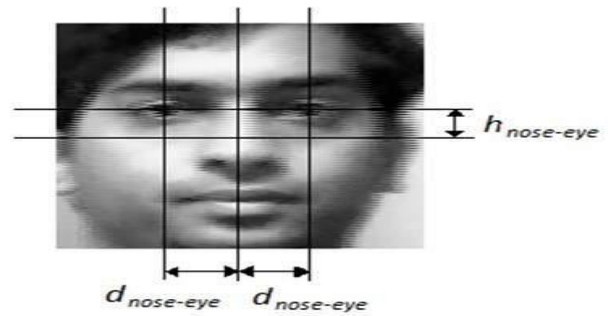$$w_{eye} = K_{we}\, h_{face} \quad (eq.\ 3)$$

Geometric Face Model using Nose:

Using the coordinates for the centre of the nose, the coordinates for the centre of both the eyes are obtained using the equations shown below. Further, cropping of a necessary section of the face is obtained using equations (eq.) 1, 2, and 3 (mentioned above).

$$d_{nose-eye} = Nx - K_{wne}Nx \quad (eq.\ 4)$$
$$h_{nose-eye} = Ny/K_{hne} \quad (eq.\ 5)$$

To improve the recognition accuracy, faces cropped with dimensions less than 256 x 256 are discarded during the face capture process. Additionally, face regions deviate significantly with respect to the direction of the source light. To mitigate this, histogram equalization is performed on the cropped face image. This reduces asymmetries formed within the face due to uneven lighting.

Face Train

In this stage, features from images associated with each person are gathered. Later, a complete set of information from all of the stored images, isolated per person as a single SVM label, is trained to generate an SVM model.

**Why are Support Vector Machines (SVMs)?**

Support vector machines (SVMs) are supervised machine learning models that divide and classify data.
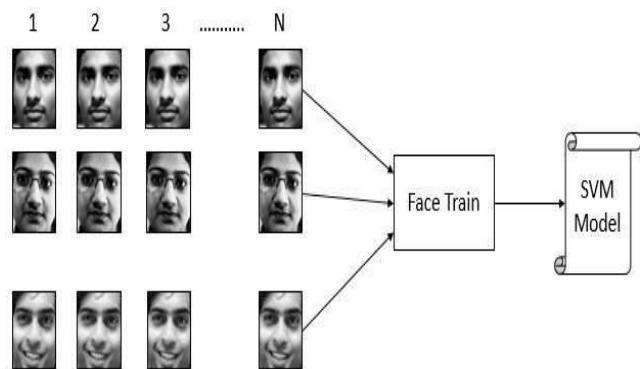
SVMs are widely used for applications such as face detection, classification of images, handwriting recognition, etc[2]. A SVM model can be considered as a point space wherein different classes are disengaged utilizing hyper planes.

What is a Histogram of Oriented Gradients (HOG)?

A HOG is a component descriptor for the most part utilized for object identification. Swines are generally known for their utilization in person on foot identification. A HOG depends on the property of items inside a picture to have the circulation of force inclinations or edge bearings. Angles are determined inside a picture for every square. A square is considered as a pixel network in which slopes are comprised from the greatness and course of progress in the forces of the pixel inside the square.
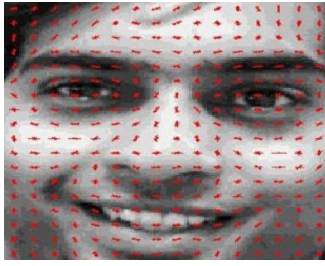
In the present model, all the face test pictures of an individual are bolstered to the element descriptor extraction calculation; i.e., a HOG. The descriptors are inclination vectors created per pixel of the picture. The angle for every pixel comprises of size and course, determined utilizing the accompanying formulae:

$$g = \sqrt{g_x^2 + g_y^2}$$

$$\theta = \arctan \frac{g_y}{g_x}$$

In the present model, Gx and Gy are separately the flat and vertical segments of the adjustment in the pixel force. A window size of 128 x 144 is utilized for face pictures since it coordinates the general viewpoint proportion of human countenances. The descriptors are determined over squares of pixels with 8 x 8 measurements. These descriptor esteems for every pixel more than 8 x 8 squares are quantized into 9 receptacles, where each container speaks to a directional edge of inclination and incentive in that canister, which is the summation of the sizes of all pixels with a similar point. Further, the histogram is then standardized over a 16 x 16 square size, which implies four squares of 8 x 8 are standardized together to limit light conditions. This component mitigates the precision drop because of an adjustment in light. The SVM model is prepared utilizing various HOG vectors for different countenances.

Face Recognition

The recognition of a face in a video sequence is split into three primary tasks: Face Detection, Face Prediction, and Face Tracking. The tasks performed in the Face Capture program are performed during face recognition as well. To perceive the face acquired, a vector of HOG highlights of the face is removed. This vector is then utilized in the SVM model to decide a coordinating score for the information vector with every one of the names[4]. The SVM restores the mark with the greatest score, which speaks to the certainty to the nearest coordinate inside the prepared face information.
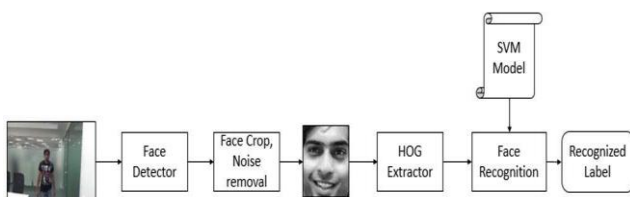
The task of calculating matching scores is exceptionally heavy to compute. Hence, once detected and identified, the labelled face in an image needs to be tracked to reduce the computation in future frames until the face eventually disappears from the video. Of all the available trackers, the Camshift tracking algorithm is used since it produces the best results with faces.

IV. RESULT ANALYSIS

The proposed calculation was prepared and assessed on the dataset of around 100 pictures containing 250 face pictures. This dataset was work from my assortment of photos and some arbitrary pictures from web. The test pictures comprised of pictures with various lumination condition – evening time, daytime and mix of them. The picture positions adequate to the calculation are jpeg, png, bmp, and so on. The dataset comprise of pictures of size running from 400x320 to 2000x1800. In the event that the size of the picture is more than 2000x1800, at that point it would make issue in handling the picture

Two parameters are defined to measure the success of algorithm. Error Rate (ER) which is defined as number of false detection in the image divides by the total number of detections (face and non-face)

Error Rate (ER) = $\frac{\text{No: of False Detections}}{\text{Total No: of Detections}}$ x 100%

The false detection means those objects that are identified as face but are not face. The total number of detections is the summation of face detected and non-face object detected Face Detection Success Rate (FDSR). It is defined as the number of faces detected correctly over the total number of faces.

FDSR = $\frac{\text{No: of Face Detected}}{\text{Total No: of Faces}}$ x 100%

ER: 7.5%; FDSR: 88.8%

Comparison of Algorithms

SVM and Detector Algorithm

SVM: ER: 7.5%; FDSR: 88.8%

Detector: ER: 25.79%; FDSR: 84.92%

V. CONCLUSION AND FUTURE DIRECTIONS

Right now have displayed face identification calculation utilizing the skin edge recognition, facial element extraction, Haar course arrangement and utilizing the idea of Histogram of situated slopes. After these pre-handling stages, the calculation uses the exceptionally influential idea of Support Vector Machine (SVM) to group the picture into face and non-face locale. As the Human face acknowledgment has a significant job as impact of present day reconnaissance and security applications.With the assistance of HOG and SVM models, one can accomplish elite levels in perceiving human faces and examining facial highlights, even in scenes containing complex backgrounds.Further, it is proposed that for acknowledgment of video pictures, YouTube Faces could be investigated for assessment.

REFERENCES

1. N.S. Abouzakhar, G.A. Manson, Evaluation of intelligent intrusion detection models, J. Digital Evidence 3(1).
2. C.-C. Chang, C.-W. Hsu, C.-J. LinThe analysis of decomposition methods for support vector machines
   IEEE Trans. Neural Networks, 11 (4) (2000), pp. 1003-1008
3. N. Cristianini, J. Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press, Cambridge, UK, 2000,
4. T. Eitrich, B. Lang, Parallel tuning of support vector machine learning parameters for large and unbalanced data sets, Preprint BUW-SC 2005/7, University of Wuppertal (May 2005).
5. G.A. Gray, T.G. Kolda, APPSPACK 4.0: asynchronous parallel pattern search for derivative-free optimization, Sandia Report SAND2004-6391, Sandia National Laboratories, Livermore, CA, August 2004
6. S. Hettich, C.L. Blake, C.J. Merz, UCI Repository of machine learning databases, 1998

AUTHORS PROFILE

Dr. Riyazuddin Y. Md, completed his PhD in CSE,His Research Areas is Networks and Security,Machine Learning, Neural Networks.He has published good no of Publications in Reputed Journals. He is Reviewer Member for Reputed Journals.

Dr. S. Mahaboob Basha, completed his PhD in CSE,He is Expert in Java Prograaming and Having 27 Years of Experience in Teaching and Research Field and his Research Areas is Machine Learning,.He has published good number of Publications in Reputed Journals

Dr. K. Krishna University, completed his PhD in Physics,His Research Areas is Atmospheric Physics and Artificila Neural Networks. CSE,He is Expert in Java Prograaming and Having 27 Years of Experience in Teaching and Research Field He has published good no of Publications in Reputed Journals.

S. Naseera Banu, completed her B.Tech in CSE from VTU, Belgavi. Her Research Areas is Networks and Security,Machine Learning, Neural Networks .She has published good no of Publications in Reputed Journals