International Journal of Engineering Technology, Management and Applied Sciences

www.ijetmas.com April 2017, Volume 5, Issue 4, ISSN 2349-4476

Effective Face Detection: A Comprehensive Survey

Prarthana Patel
PG Student
LJ Institute of Engineering and Technology
Ahmedabad, India

Dr. A.C.Suthar
Director
LJ Institute of Engineering and Technology
Ahmedabad, India

Abstract

In growing fast world facial recognition is quiet challenging as there are variety of faces in the world and also complexity of noises and backgrounds. Face detection is one of the key areas under research. It has number of applications and uses. Many methods and algorithms are put forward. Face recognition is one of the key territories of research. It has various applications and employments. Numerous strategies and algorithm are formed. We can also develop many intelligent applications which may provide security and identity.

We propose a review paper on facial Detection in which certain algorithms that are two stages Convolution Neural Network (CNN) and Support-Vector Machine(SVM) are basically used for feature classification and the Convolution Neural Network (CNN) is used feature extraction and by using these algorithm will try to provide accurate and effective Face Detection.

Key words

Face Detection, Convolutional Neural Network(CNN),Support Vector Machine(SVM).

INTRODUCTION

With the swift increase of computational technology and availability of modern sensing, analysis and rendering equipment and technologies, computers are becoming more and more quick-witted technology. Many research works, projects and commercial products have demonstrated the capacity for a computer to interact with human in a natural way by looking at people through lenses and hearing to people through wearable device, understanding those inputs, and then reacting to people in a friendly manner. Face detection is one of the basic techniques that licence such unrefined human- computer interaction (HCI).Identity verification is one of the stepping stone to all facial analysis algorithms, including face alignment, face modelling, face relighting, face recognition, face verification and authentication, head pose tracking, facial expression tracking/recognition, gender/age recognition, and many more. Only when face verification can be properly understood by computers, then only we will begin to truly understand people's thoughts and intentions.

Given any example picture, the principle objective of face recognition is to figure out if or not there are any countenances in the picture and, if there is, give back the face location and degree of each face. While this appears as a major task for human beings, it is quit challenging task for computers. It has been one of the tops considered research points in the previous few decades. The trouble related with face identification can be credited to numerous varieties in scale, area, introduction (in-plane pivot), posture (out-of-plane revolution), outward appearance, lighting conditions, impediments, and so forth.

STEPS FOR FACIAL DETECTION

1) Pre-processing: A pre-processing module finds the eye position and deals with the encompassing lighting condition and color difference. To start with, face in a scene must be recognized. Once the face is identified, it must be limited and standardization process might be required to bring the measurements of the live facial example in arrangement with the one on the layout or template. Some facial acknowledgment approaches utilize the entire face while others focus on facial parts or potentially regions(such as lips, eyes and so on).

The presence of the face can change impressively amid discourse and because of outward appearances. Specifically, the mouth is subjected to central changes. It is carried out in two parts, one in which less then 30% is covered is placed in one category and the other more then 30% is placed in another category.

2) Feature Extraction: Feature Extraction begins from an underlying arrangement of measured information and assembles determined qualities (highlights) expected to be useful and non-excess, encouraging the ensuing learning and speculation steps, and now and again prompting to better human elucidation. Feature extraction is identified with dimensionality lessening. feature extraction includes reducing the measure of assets required to depict a vast set of information. When performing an analysis of complex information one of the real issues originates from the number of variables included. Investigation with countless by and large requires a lot of memory and computation power, likewise, it might make a Classification algorithm overfit to preparing tests and sum up inadequately to new specimens. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy.

3) Feature Classification: The classification includes a wide scope of choice theoretic ways to deal with the ID of images (or parts thereof). All classification algorithms depend on the presumption that the picture being referred to delineates at least one elements and that each of these elements belongs with one of the particular and exclusive classes. In my project feature classification is carried out by two stages Convolution Neural Network (CNN) and Support-Vector Machine (SVM).

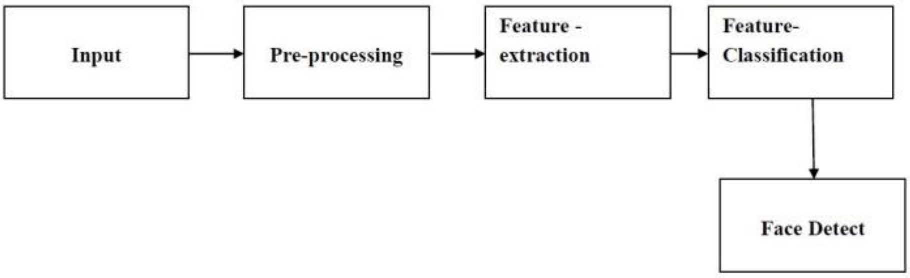


Figure 1: Block Diagram of Face Detection

ALGORITHM

A. Convolution Neural Network (CNN):-

Convolutional neural system (CNN), initially proposed by LeCun, is a neural system demonstrate with three key engineering thoughts: nearby responsive fields, weight sharing, and sub-testing in the spatial area. The system is intended for the acknowledgment of two-dimensional visual examples. Convolutional neural network has many strengths. First, feature extraction and classification are integrated into one structure and are fully adaptive. Second, the network extracts 2-D image features at increasing dyadic scales. Third, it is relatively invariant to geometric, local distortions in the image. CNN has been used for in several applications including hand-written digit recognition, face detection, and face recognition.

a) Network architecture

Convolutional neural networks are designed to process two-dimensional (2-D) image. A CNN consists of three main types of layers:

(i) convolution layers,

(ii) sub-sampling layers, and

(iii) an output layer.

Network layers are arranged in a loop forward structure: every convolution layer is trailed by a sub-sampling layer, and the last convolution layer is trailed by the output layer The convolution and sub-inspecting layers are considered as 2-D layers, though the output layer is considered as a 1-D layer. In CNN, every 2-D layer has a few planes. A plane comprises of neurons that are arranged in a 2-D exhibit. The output of a plane is known as a component delineate.

In a convolutional layer, each plane is associated with at least one element maps of the previous layer. An association is related with a convolution veil, which is a 2-D matrix of movable passages called weights. Each plane first registers the convolution between its 2-D sources of info and its convolution veils. The convolution yields are summed together and after that additional with an adjustable scalar, known as an inclination term. At last, an activation function is connected to the outcome to acquire the plane's output. The plane yield is a 2-D grid called a feature map; this name emerges in light of the fact that every convolution yield demonstrates the nearness of a visual element at a given pixel area. A convolution layer produces at least one element maps. Each element guide is then associated with precisely one plane in the following sub-sampling layer. sub-sampling layer has an indistinguishable number of planes from the previous convolution layer. A subsampling plane divides its 2-D input into non-overlapping blocks of size 2 × 2 pixels. For each block, the sum of four pixels is calculated; this sum is multiplied by an adjustable weight before being added to a bias term. The result is passed through an activation function to produce an output for the 2 × 2 block. Clearly, each sub-sampling plane reduces its input size by half, along each dimension. A feature map in a sub-inspecting layer is associated with at least one planes in the following convolution layer.

In the last convolution layer, each plane is associated with precisely one preceding feature highlight outline. This layer utilizes convolution covers that have the very same size as its info feature maps. Therefore, each plane in the last convolution layer will produce one scalar output. The outputs from all planes in this layer are then connected to the output layer.

The output layer, all in all, can be built from sigmoidal neurons or radial basis function (RBF) neurons. Here, we will concentrate on utilizing sigmoidal neurons for the output layer. The yields of this layer are considered as the network outputs.

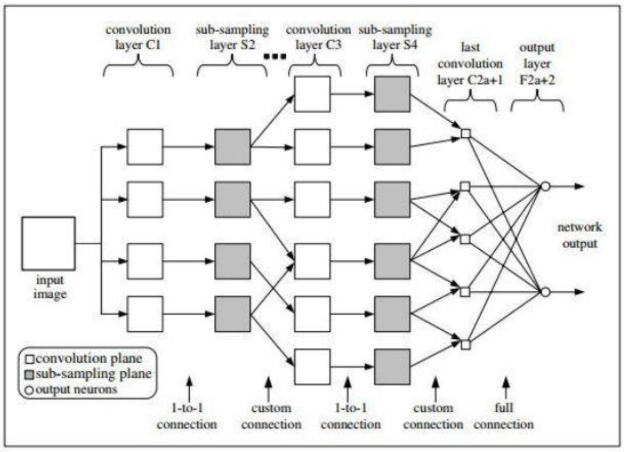


Figure 2. Layers in CNN

B.SUPPORT VECTOR MACHINE(SVM):-

SVMs introduced in COLT-92 by Boser, Guyon & Vapnik. Became rather popular since. Theoretically well motivated algorithm: developed from Statistical Learning Theory (Vapnik & Chervonenkis) since the 60s. Empirically good performance: successful applications in many fields (bioinformatics, text, image recognition, . . .).Centralized website: www.kernel-machines.org.Several textbooks, e.g. ”An introduction to Support Vector Machines” by Cristianini and Shawe-Taylor is one. A large and diverse community work on them: from machine learning, optimization, statistics, neural networks, functional analysis, etc.

In machine learning, support vector machines (SVMs, additionally support vector networks[1]) are supervised learning models with related learning algorithms that investigate information utilized for classification and regression analysis. Given an arrangement of preparing cases, each set apart as having a place with either of two classifications, an SVM algorithm assembles a model that doles out new cases to one classification or the other, making it a non-probabilistic binary straight classifier. An SVM model is a portrayal of the cases as focused in space, mapped so that the cases of the different classifications are separated by a reasonable gap that is as wide as could be expected under the circumstances. New cases are then mapped into that same space and anticipated to have a place with a class in light of which side of the gap they fall.

Notwithstanding performing linear classification, SVMs can proficiently play out a non-linear classification utilizing what is known as the kernel trick, verifiably mapping their contributions to high-dimensional element spaces.

At the point when information is not named, supervised learning is unrealistic, and an unsupervised learning methodology is required, which endeavours to discover common bunching of the information to gatherings, and after that guide new information to these shaped gatherings. The clustering algorithm which gives a change to the support vector machines is called support vector clustering and is frequently utilized as a part of mechanical applications either when information are not marked or when just a few information are named as a pre-processing for a characterization pass.

SVMs are useful in content and hypertext classification as their application can altogether decrease the requirement for marked training instances in both the standard inductive and transductive settings. Classification of pictures can likewise be performed utilizing SVMs.

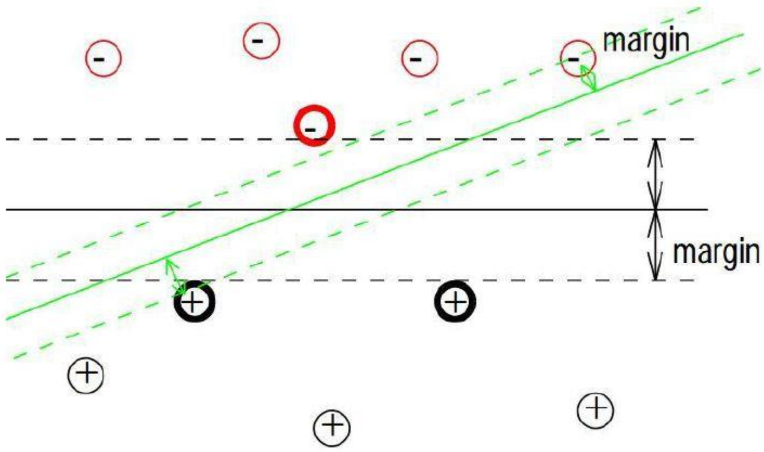


Figure 3: Positive and negatives in SVMs

LITERATURE SURVEY

Eran Eidinger et.al[1] This paper concerns the estimation of facial properties to be specific, age and gender from pictures of confronts obtained in testing, in the wild conditions. This issue has gotten far less consideration than the related issue of face acknowledgment. Here, They address this issue by making the given commitments. In the first place, in reply to one of the key issues of age estimation investigate nonappearance of information we offer a one of a kind informational collection of face pictures, named for age and sex, gained by advanced cells and other cell phones, and transferred without manual separating to online picture repositories.They demonstrate the pictures in our gathering to be more testing than those offered by other face-photograph benchmarks. Second, we portray the dropout-bolster vector machine approach utilized by our framework for face trait estimation, keeping in mind the end goal to abstain from overfitting.

Detection rate:

- Age Detection rate =95.3%
- Gender Detection rate =88.6%

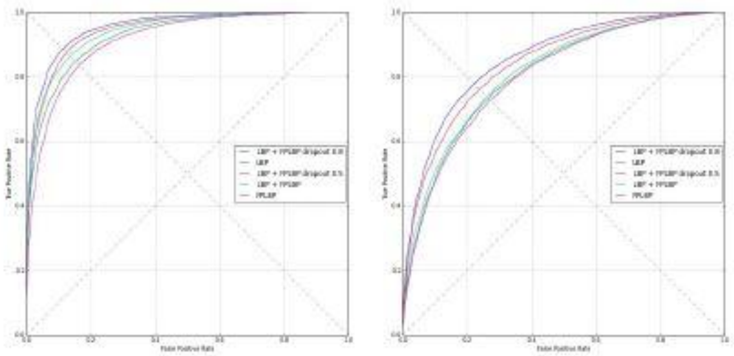


Fig 4.1 ROC curve of gender estimation results(1)

Anna Gruebler et.al [2] proposed the design of a wearable device that reads positive facial expressions using physiological signals. They first break down facial morphology in 3 measurements and facial electromyographic motions on various facial areas and demonstrate that we can identify electromyographic signals with high amplitude on territories of low facial versatility in favor of the face, which is related to ones got from anodes on conventional surface electromyographic catching positions on top of facial muscles on the front of the face. They utilize a multi-ascribe decision-making technique to discover sufficient anode positions in favor of face to catch these signs. the proposed gadget utilize independent analysis and an artificial neural system to examine them and accomplish a high outward appearance acknowledgment rate in favor of the face. Detection rate of positive signal is 98%

(a) Smiling (b) Frowning

Figure 4.2 Reading of Positive signals(2)

Gong Cheng, et.al [3] proposed paper a novel and effective approach to learn a rotation-invariant CNN (RICNN) model for advancing the performance of object detection, which is achieved by introducing and learning a new rotation-invariant layer on the basis of the existing CNN architectures. Be that as it may, not the same as the training of customary CNN models that alone optimizes the multinomial strategic relapse objective, our RICNN model is prepared by improving another objective function by means of forcing a regularization requirement, which unequivocally implements the component portrayals of the preparation tests previously, then after the fact turning to be mapped near each other, henceforth accomplishing rotation invariance. To facilitate training, we first train the rotation-invariant layer and then domain-specifically fine-tune the whole RICNN. Detcetion rate is 69.8%-72.23%

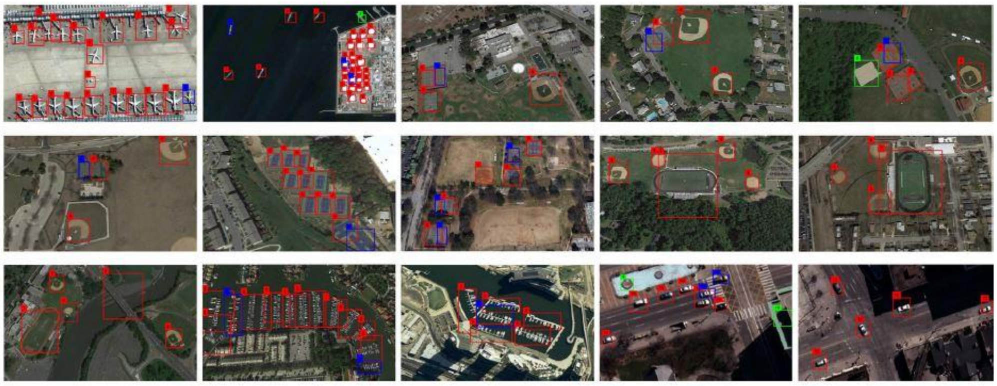


Figure 4.3 no of object detection with proposed approach(3)

Yanwei Pang et.al [4] proposed work develops a distributed object detection framework (DOD) by making the best use of spatial-temporal correlation, where the process of feature extraction and classification is distributed in the current frame and several previous frames. In each framework, only subfeature vectors are extracted and the response of partial linear classifier is computed. To reduce the dimension of traditional block-based histograms of oriented gradients (BHOG) feature vector, this paper proposes a cell-based HOG (CHOG) algorithm, where the features in one cell are not shared with overlapping blocks**.** Detcetion rate Day time=98.5% Night time =97.1%

Figure 4.4 Examples of successful human detection [4]

Anirban Dasgupta et.al [5] proposed a robust real-time embedded platform to monitor the loss of attention of the driver during day and night driving conditions. The percentage of eye closure has been used to indicate the alertness level. In this approach, the face is detected using Haar-like features and is tracked using a Kalman filter. The eyes are recognized utilizing principal component analysis amid daytime and utilizing block local-binary-pattern features amid night time. At last, the eye state is named open or shut utilizing support vector machines. In-plane and off-plane revolutions of the driver's face have been repaid utilizing affine transformation and perspective transformation individually. Detection rate is 72.04%

Fig 4.5 Training images for eye state classification using SVM. [5]

RESULTS

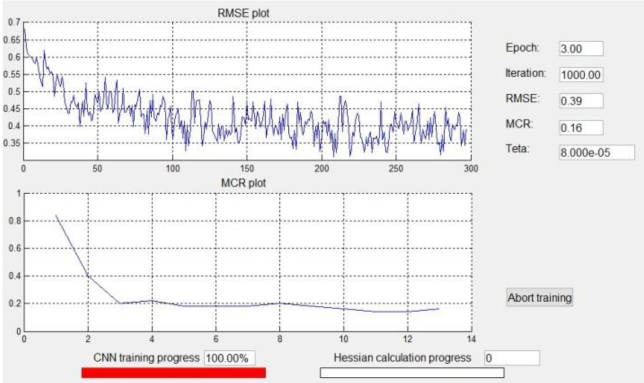
From the above literature, basic training of CNN is shown in the graph where,

Epoch=3

Iteration=300

RMSE=0.39

MSU=0.16



CONCLUSION

Face recognition technology has come a long way in the last twenty years. In this paper, we have talked about the commonly used face detection methods. Every single of these methods symbolizes the value and utility for different applications. Since these methods are progressive, a growing number of advancements are made every day to achieve appropriate and true face diagnosis. For application such as security and identity confirmation. Thus identifying faces that are covered with spectacles, scarf are also considered into consideration using 2 stage CNN and SVM. thus getting effective and accurate face detection. Nevertheless, latest face recognition systems will have widespread program in smart environments -- where computers and machines are more like helpful assistants.

FUTURE ASPECTS

From the study of numerous methods, one can build a system that can detects which uses the two stage CNN for feature extraction and SVM for feature classification and can be implemented and more accurate and effective and also contains images that are covered into their dataset.

REFERENCES

[1] Eran Eidinger, Roee Enbar, and Tal Hassner," Age and Gender Estimation of Unfiltered Faces"IEEE Transactions on Information Forensics and Security, Vol.9,No.12,December 2014.

[2] Anna Gruebler, Member, IEEE and Kenji Suzuki," Design of a Wearable Device for Reading Positive Expressions from Facial EMG Signals"IEEE Transactions on Affective Computing Vol. 5, No. 3, July-September 2014

[3] Gong Cheng, Peicheng Zhou, and Junwei Han," Learning Rotation-Invariant Convolutional Neural Networks for Object Detection in VHR Optical Remote Sensing Images "IEEE Transaction on Geoscience and Remote Sensing,2016.

[4]. Yanwei Pang, Kun Zhang, Yuan Yuan and Kongqiao Wang," Distributed Object Detection With Linear SVMs"IEEE Transactions on Cybernetics vol.44, No.11, November 2014.

[5] Anirban Dasgupta, Anjith George, S. L. Happy, and Aurobinda Routray ,"A Vision-Based System for Monitoring the Loss of Attention in Automotive Drivers"IEEE Transaction on Intelligent Transportation System,Vol.14,No.4,December 2013.

[6] Zheng-Hai Huang,Wen-Jaun,Jun Wang,Ting Zhang,"Face recognition based on pixel-level and feature –level fusion of the top-level's Wavelet sub-bands" Elsevier,Information fusion 2015.

[7] Kenny Hong,Stephan K. Chalup,Robert A.R.King,"Affective Visual Perception Using Machine Pareidlia of Facial Expression"IEEE Transaction on Affective Computing, Vol.5,N0.4,October-December 2014.

[8] Yashika Katyal, SuhasbV Alu, Shipra Dwivedi, Menaka R,"EEG Signal and Video Analysis Based Depression Indication" IEEE International Conference on Advanced Communication Control and Computing Technologies 2014.

[9] Mostafa K. Abd El Meguid and Martin D. Levine," Fully Automated Recognition of Spontaneous Facial Expressions in Videos Using Random Forest Classifiers" IEEE Transaction on Affective Computing,Vol.5,No.2,Aprl-June 2014.

[10] Amer AI-Rahayfeh, Miad Faezipour," Enhanced Eye Gaze Direction Classification Using a Combination of Face Detection, CRT and SVM" IEEE Conference 2014.

[11] Javier Galbally,Sebastien Marcel,Julian Fierrez,"Image Quality Assesment for Fake biometric Detection:Application to Iris,Fingerprint and Face recognition"IEEE Transaction on Image Processing, Vol.23,No.2,Feburary 2014.

[12] Allen Y. Yang,Zihan Zhou,Arvind Ganesh Balasubramanian,"Fast l1-Minmization Algorithm for Robust Face Recognition"IEEE Transactions on Image Processing, Vol.22,No.8,August 2013.

[13] Bingbing Ni, Zheng Song, and Shuicheng Yan," Web Image and Video Mining Towards Universal andRobust Age Estimator" IEEE Transactions on Multimedia, Vol13,No.6,December 2011.

[14] Jae Hyun Oh and Nojun Kwak," Recognition of a Driver's Gaze for
Vehicle Headlamp Control" " IEEE Transactions on Vehicular Technology,Vol.61,No.5,June 2012.

[15] Fayin Li and Harry Wechsler,"Open Set Face Recognition Using Transduction"IEEE Transaction on Pattern Analysis and Machine Inetlligence,Vol.27,No.11,November2015.

[16] Tolga Inan and Ugur Halici,"3-D Face Recogniton With Local Shape Descriptors"IEEE Transactions on Information Forensics and Security,Vol.7,No.2,April 2012.

[17] Pandu Rangarao Devarakota, Marta Castillo-Franco, Romuald Ginhoux," 3-D-Skeleton-Based Head Detection and Tracking Using Range Images" IEEE Transactions on Vehicular Technology,Vol.58,No.9,June 2012.

[18] Laurindo Britto Neto, Felipe Grijalva,Vanessa Regina," A Kinect-Based Wearable Face Recognition System to Aid Visually Impaired Users"IEEE Transaction on Human-Machine Systems 2016