International Journal of Engineering and Technical Research (IJETR)
ISSN: 2321-0869 (O) 2454-4698 (P), Volume-7, Issue-5, May 2017

One-to-Many Face Recognition with Bilinear CNNs: A Survey

Sonal Mishra, Mr. Pratyush Tripathi

Abstract*— Face recognition from the real data, capture images, sensor images and database images is challenging problem due to the wide variation of face appearances, illumination effect and the complexity of the image background. Face recognition is one of the most effective and relevant applications of image processing and biometric systems. In this paper we are discussing the face recognition methods, algorithms proposed by many researchers using artificial neural networks (ANN) which have been used in the field of image processing and pattern recognition. How ANN will used for the face recognition system and how it is effective than another methods will also discuss in this paper. There are many ANN proposed methods which give overview face recognition using ANN. Therefore, this research includes a general review of face detection studies and systems which based on different ANN approaches and algorithms. The strengths and limitations of these literature studies and systems were included, and also the performance analysis of different ANN approach and algorithm is analyzing in this research study. In the past two decades, Shape detection has been proven as the most interesting research field. We present a novel approach to measuring the similarities between shapes and exploit it for object recognition. The measurement of similarity preceded by solving correspondences between points on two shapes as well as by using the correspondences to estimate an aligning transform. In this paper, we are going to describe some important aspects of face detection, which are very much useful in many applications like face recognition, facial expression recognition, face tracking, facial feature extraction, gender classification, identification system, document control and access control, clustering, biometric science, human computer interaction (HCI) system, digital cosmetics and many more. Before that, I would like to focus on some well-known face detection techniques and then some feature detection methods because until and unless we extract the important features (eyes, nose, mouth) from a face we won't be able to uniquely identify the person properly. The dissimilarity between the two shapes is computed as some of matching error between corresponding points. We treat recognition in a nearest neighbor classification framework as the problem of finding the stored prototype shape that is maximally similar to that in the image.

***Index Terms*—Face Recognition, Biometric, Image Processing, Pattern Recognition, Artificial Neural Network**

I. INTRODUCTION

The Humans have recognition capabilities that are unparalleled in the modern computing era. Humans have always had the a ability to recognize and distinguish between faces features . Recently the computers have been shown the same ability. In the mid of 1960s, scientists began work on using the computer to recognize human faces, since that time the facial recognition software has come a long way, the governments and private companies start to used the face recognition system . A lot of changes are being made to enhance scientist's capabilities in this field. Many researchers belief that face recognition need to be about face detection, the prior step to face recognition the accurate detection of human faces in arbitrary scenes, is the most important process involved. Human faces are remarkably similar in global properties, including size, aspect ratio, and location of main features, but can vary considerably in details across individuals, gender, race, or due to facial expression, most traditional face recognition systems attempt to achieve a low recognition error rate, implicitly assuming that the losses of all misclassifications are the same (Shlizerman, 2011). Facial recognition software in any given episode, the security department at the fictional. Montecito Hotel and Casino uses its video surveillance system to pull an image of a card counter, thief or blacklisted individual, then runs that image through the database to find a match and identify the person. In 2001, the Tampa Police Department installed police cameras equipped with facial recognition technology in their City nightlife district in an attempt to cut down on crime in the area, the system failed to do the job, and it was scrapped in 2003 due to ineffectiveness (Patra and Das. 2008)

Sonal Mishra, Department of Electronics & Communication Engineering, M.Tech Scholar, Kanpur Institute of Technology, Kanpur, India.

Mr. Pratyush Tripathi, Associate Professor, Department of Electronics & Communication Engineering, Kanpur Institute of Technology, Kanpur, India.

II. FACE RECOGNITION SYSTEM

Every Biometric system has four main features which are shown in Figure 1: face Detection, preprocessing, Feature Extraction, and Face Recognition.

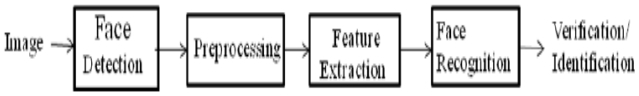

Figure 1 Architecture of Face Recognition System

As Figure 1 shows the first task of the face recognition system is capturing image by video, camera or from the database and this image is given to the further step of face recognition system that is discuss in this section:

Face Detection: The main function of this step is to detect the face from capture image or the selected image from the database. This face detection process actually verifies that weather the given image has face image or not, after detecting the face this output will be further given to the pre-processing step.

Pre-processing: This step is working as the pre-processing for face recognition, In this step the unwanted noise, blur, varying lightening condition, shadowing effects can be remove using pre-processing techniques .once we have fine smooth face image then it will be used for the feature extraction process.

Feature Extraction: Face can be extracted using feature extraction algorithm. Extractions are performed to do information packing, dimension reduction, salience extraction, and noise cleaning. After this step, a face patch is usually transformed into a vector with fixed dimension or a set of fiducial points and their corresponding locations.

Face Recognition: Face recognition can be largely classified into two different classes of approaches, the local feature-based method and the global feature-based method. The Human faces can be characterized both on the basis of local as well as of global features global features are easier to capture they are generally less discriminative than localized features local features on the face can be highly discriminative, but may suffer for local changes in the facial appearance or partial face occlusion. Now a day's face recognition system is recognize the face using multiple-views of faces, these Multi-view face recognition techniques has proposed by some authors for detecting each view of face such as left, right, front, top, and bottom *etc*.

III. Face Detection Technique

A large number of representation techniques are available for face detection, including Knowledge-based, Feature invariant based, Template matching method, Appearance-based methods, Part-based methods.

Feature Invariant Approaches to find structural features that exist even when the pose, viewpoint, or lighting conditions vary, and then use these to locate faces. To distinguish from the knowledge-based methods, the feature invariant approaches start at feature extraction process and face candidates finding, and later verify each candidate by spatial relations among these features, while the knowledge-based methods usually exploit information of the whole image and are sensitive to complicated backgrounds and other factors. Readers could find more works in [3][4][5]. Face detection based on color information, random labeled graph matching fall in this category.

Knowledge-based methods encode human knowledge [6] of what constitutes a typical face. Usually, the rules capture the relationships between facial features. These methods are designed mainly for face localization, which aims to determine
the image position of a single face.

Template matching methods standard patterns of a face are stored to describe the face as a whole or the facial feature separately. The correlations between an input image and the stored pattern are computed for detection. These methods have been used for both face localization and detection. Deformable template matching [7] falls in this category, where the template of faces is deformable according to some defined rules and constraints.

Appearance-based methods in contrast to template matching, the models (or templates) are learned from a set of training images, which should capture the representative variability of facial appearance. These learned models are then used for detection. More significant techniques are included in [1][8]. Examples of such type of methods are view-based face detection

Part-based methods with the development of the graphical model framework and the point of interest detection such as the difference of Gaussian detector [9] (used in the SIFT detector) and the Hessian affine detector, the part-based method recently attracts more attention. Some well-known approaches like face detection based on the generative model framework, component based face detection based on the SVM classifier falls into this category.

IV. Neural Network Technique

Neural network is a very powerful and robust classification technique which can be used for predicting not only for the known data, but also for the unknown data. It works well for both linear and non linear separable dataset. NN has been used in many areas such as interpreting visual scenes, speech recognition, face recognition, finger print recognition, iris recognition *etc.*

An ANN is composed of a network of artificial neurons also known as "nodes". These nodes are connected to each other, and the strength of their connections to one another is assigned a value based on their strength: inhibition (maximum being -1.0) or excitation (maximum being +1.0).

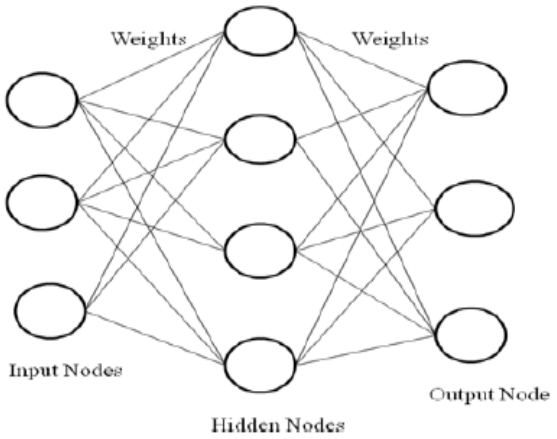

Figure 2 Artificial Neural Network

Single Layer Feed Forward Network: A neural network in which the input layer of source nodes is connected to an output layer of neurons but not vice-versa is known as single feed-forward or acyclic network. In single layer network, „single layer" refers to the output layer of computation nodes as shown in Figure 3

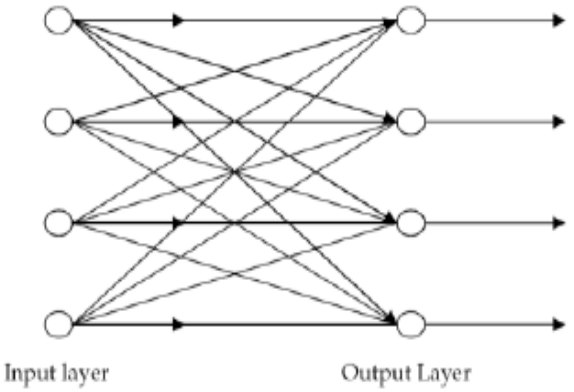

Figure 3 Single Layer Feed Forward Network

Multilayer Feed Forward Network: This type of network consists of one or more hidden layers, whose computation nodes are called hidden neurons or hidden units. The function of hidden neurons is to interact between the external input and network. The source nodes in input layer of network supply the input signal to neurons in the second layer of 1st hidden

layer. The output signals of 2nd layer are used as inputs to the third layer and so on. The set of output signals of the neurons in the output layer of network constitutes the overall response of network to the activation pattern supplied by source nodes in the input first layer [2].

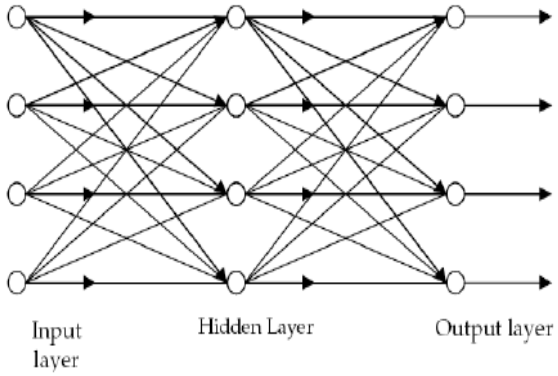

Figure 4 Multilayer Feed Forward Network

Recurrent Network: A feed forward neural network having one or more hidden layers with at least one feedback loop is known as recurrent network as shown in Figure 5. The feedback may be a self feedback, *i.e*., where output of neuron is given back to its own input. Sometimes, feedback loops involve the use of unit delay elements, which results in nonlinear dynamic behavior, assuming that neural network contains non linear units. There are various other types of networks like; delta-bar-delta, Hopfield, vector quantization, counter propagation, probabilistic, Hamming, Boltzman, bidirectional\ associative memory, spacio-temporal pattern, adaptive resonance, self organizing map, recirculation *etc* [2].

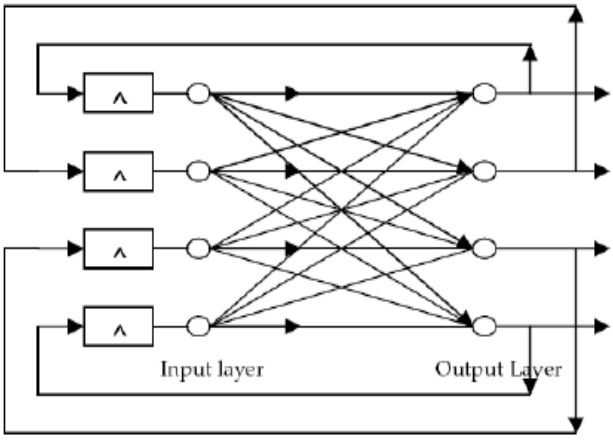

Figure 5 Recurrent Connected Network

ANN has been classified in different types like single layer feed forward Neural Network, Multilayer Feed Forward Neural Network, and Recurrent Network which is discuss here. The many researchers have used these types of ANN for Face recognition and Different they had proposed different algorithms for face recognition.

V. Convolutional Neural Network (CNNs)

Convolutional neural networks (CNNs) are composed of a hierarchy of units containing a convolution, pooling and non-linear layer. In recent years deep CNNs typically consisting of the order of 10 or so such units and trained on massive labeled datasets such as ImageNet have yielded generic features that are applicable in a number of recognition tasks ranging from image classification [13], object detection [6], semantic segmentation [8] to texture recognition [3]. In the domain of fine-grained recognition, such as identifying the breed of a dog, species of a bird, or the model of a car, these architectures, when combined with detectors that localize various parts of the object, have also yielded state-of-the-art results. Without the part localization, CNNs typically don't perform as well since there is a tremendous variation in appearance due to different poses that instances of a category can be in. This pose variation overwhelms the subtle differences across categories, a phenomenon typical also in the face recognition problem. However, the drawback of these approaches is that they require (a) manual annotation of parts which can be time-consuming, (b) the detection of parts which can be computationally expensive.

VI. Bilinear CNNs

Aruni RoyChowdhury Tsung-Yu Lin Subhranshu Maji Erik Learned-Miller [9] has proposed Bilinear CNN method for face identification which has shown dramatic performance gains on certain fine-grained recognition problems it bridges the gap between the texture models and part-based CNN models. It consists of two CNNs whose convolutional - layer outputs are multiplied using outer product at each location of the image. The resulting bilinear feature is placed across the image resulting in an order less descriptor for the entire image. This vector can be normalized to provide additional invariance. If one of the feature extractors was a part detector and the other computed local features, the resulting bilinear vector can model the representations of a part-based model. On the other hand, the bilinear vector also resembles the computations of a Fisher vector, where the local features are combined with the soft membership to a set of cluster centers using an outer product.

VII. Conclusion

This paper includes a summary review of literature studies related to face recognition systems based on ANNs. In this paper we are discussed different architecture, approach, algorithms, methods, database for training or testing images and performance measure of face recognition system were used in each study. Every researcher has their own approach for recognizing face from database or from video many researches has try to solve the problems associated with earlier proposed method but still there are some advantages and limitations in these discussed methods.

.

References

[1] A Krenker, J Bešter and A Kos," Introduction to the Artificial Neural Networks", Edited Kenji Suzuki, Published by InTech,, Janeza Trdine, Croatia, (**2011**), pp 3-18.

[2] Er. P Kumar, Er.P Sharma, "ARTIFICIAL NEURAL NETWORKS- A Study", International Journal of Emerging Engineering Research and Technology, vol. 2, no. 2, (**2014**), pp. 143-148.

[3] N Jindal, V Kumar," Enhanced Face Recognition Algorithm using PCA with Artificial Neural Networks", International Journal of Advanced

Research in Computer Science and Software Engineering, vol 3, no. 6, (**2013**), pp. 864-872.
[4] S. Branson, G. V. Horn, S. Belongie, and P. Perona. Bird species categorization using pose normalized deep convolutional nets. In Proc. BMVC, 2014.
[5] J.-C. Chen, V. M. Patel, and R. Chellappa. Unconstrained face verification using deep CNN features. CoRR, abs/1508.01722, 2015.
[6] M. Cimpoi, S. Maji, and A. Vedaldi. Deep filter banks for texture recognition and description. In Proc. CVPR, 2015.
[7] G. Csurka, C. R. Dance, L. Dan, J. Willamowski, and C. Bray. Visual categorization with bags of keypoints. In Proc. ECCV Workshop on Stat. Learn. in Comp. Vision, 2004.
[8] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei- Fei. ImageNet: A large-scale hierarchical image database. In Proc. CVPR, 2009.
[9] R. B. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proc. CVPR, 2014.
[10] P. Grother and M. Ngan. Face recognition vendor test (FRVT): Performance of face identification algorithms. In NIST Interagency report 8009, 2014.
[11] B. Hariharan, P. Arbel´aez, R. Girshick, and J. Malik. Simultaneous detection and segmentation. In Proc. ECCV, 2014.
[12] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical report, 07-49, University of Massachusetts, Amherst, 2007.
[13] H. J´egou, M. Douze, C. Schmid, and P. P´erez. Aggregating local descriptors into a compact image representation. In Proc. CVPR, 2010.
[14] I. Kemelmacher-Shlizerman and R. Basri. 3D face reconstruction from a single image using a single reference face shape. PAMI, 2011.
[15] B. F. Klare, B. Klein, E. Taborsky, A. Blanton, J. Cheney, K. Allen, P. Grother, A. Mah, and A. K. Jain. Pushing the frontiers of unconstrained face detection and recognition: IARPA Janus Benchmark A. In Proc. CVPR, 2015.

Sonal Mishra, Department of Electronics & Communication Engineering, M.Tech Scholar, Kanpur Institute of Technology, Kanpur, India.

Mr. Pratyush Tripathi, Associate Professor, Department of Electronics & Communication Engineering, Kanpur Institute of Technology, Kanpur, India.