IJournals: International Journal of Software & Hardware Research in Engineering
ISSN-2347-4890
Volume 5 Issue 9 September, 2017

A Survey on Face Recognition using Convolution Neural Network

Ms. Priya Tambi
M.Tech Scholar
Sri Aurobindo Institute of Technology
Indore, India

Prof. Sarika Jain
Assistant Professor
Sri Aurobindo Institute of Technology
Indore, India

Abstract

Face Recognition is a currently developing technology with multiple real life applications. Numerous algorithms and techniques have been developed for improving the performance of face recognition. Recently Deep learning has been highly explored for computer vision applications. Convolutional Neural Networks can be used in order to extract relevant facial features. These featurnisallowed to compare faces between them in an efficient way. The system can be trained to recognize a set of people. We provide various applications that can be developed that make use of this Face Recognition technology.

Keywords—Deep Learning (DL); Face Recognition (FR);

IJournals

I. INTRODUCTION

In the last few years, with the growth of security-oriented applications such as access control, card identification, security monitoring the importance of face recognition has been increased. Face recognition (FR) remains an actively studied topic in computer vision community and pattern recognition in comparing with other biometric recognition.

In recent years, deep learning methods, especially the deep convolutional neural networks (CNN), has achieved remarkable successes in various computer vision tasks, ranging from image classification to object detection and semantic segmentation, etc. In contrast to traditional computer vision approaches, deep learning methods avoid the hand-crafted design pipeline and have dominated many well-known benchmark evaluations, such as ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [1]. Along with the popularity of deep learning in computer vision, a surge of research attention has been emerging to explore deep learning for resolving face detection and recognition tasks.

II. NEURAL NETWORKS

A. Artificial Neural Network (ANN)

The neural network is a system of interconnected artificial “neurons” thatexchange messages between each other. The connections have numeric weights that are tuned during the training process, so that a properly trained network willrespond correctly when presented with an image or pattern to recognize. The network consists of multiple layers of feature-detecting “neurons”.[2] Each layer has many neurons that respond to different combinations of inputs from theprevious layers. As shown in Figure 1, the layers are built up so that the first layer detects a set of primitive patterns in the input, the second layer detects patterns of patterns, the third layer detects patterns of those patterns, and so on.

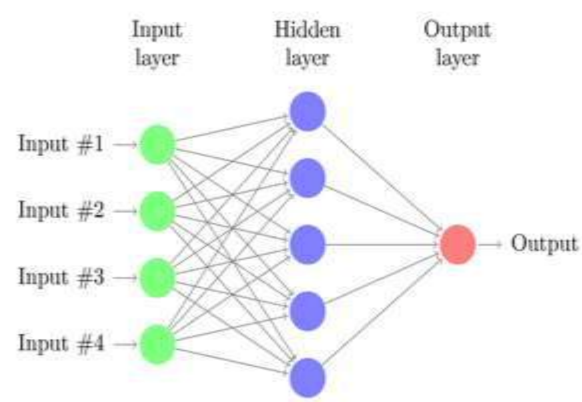

Figure 1. Artificial Neural Network

B. Convolution Neural Network (CNN)

Convolutional neural networks (CNNs) are the current state-of-the-art model architecture for image classification tasks. CNNs apply a series of filters to the raw pixel data of an image to extract and learn higher-level features, which the model can then use for classification. CNNs contains three components:

- **Convolutional layers**, which apply a specified number of convolution filters to the image. For each subregion, the layer performs a set of mathematical operations to produce a single value in the output feature map. Convolutional layers then typically apply a ReLU activation function to the output to introduce nonlinearities into the model.
- **Pooling layers**, which downsample the image data extracted by the convolutional layers to reduce the dimensionality of the feature map in order to decrease processing time. A commonly used pooling algorithm is max pooling, which extracts subregions of the feature map (e.g., 2x2-pixel tiles), keeps their maximum value, and discards all other values.
- **Dense (fully connected) layers**, which perform classification on the features extracted by the convolutional layers and downsampled by the pooling layers. In a dense layer, every node in the layer is connected to every node in the preceding layer.

Typically, a CNN is composed of a stack of convolutional modules that perform feature extraction. Each module consists of a conventionallayer, followed by a pooling layer. The last convolutional module is followed by one or more dense layers that perform classification. The final dense layer in a CNN contains a single node for each target class in the model (all the possible classes the model may predict), with a softmax activation function to generate a value between 0–1 for each node (the sum of all these softmax values is equal to 1). We can interpret the softmax values for a given image as relative measurements of how likely it is that the image falls into each target class.

Convolutional neural networks obtain a relatively good result in recognizing handwritten numeral . In this task, too much additional adaptation work is not needed. There are two aspects of adjustment: the transmutation of input image size of it and the matching of end product building block numbers. The facts display convolutional neural networks can also be used in more different recognition tasks[3]. Convolutional neural networks aim to use spatial information between the pixels of an image. Using this architecture makes convolutional networks fast to train. CNN is an efficient recognition algorithm which is widely used in pattern recognition and image processing [4].

III. Working of CNN

Convolutional Neural Networks combine three architectural ideas to ensure some degree of shift, scale, and distortion invariance: local receptive fields, shared weights, and spatial or temporal sub-sampling [5]. The network is usually trained like a standard neural network by back propagation.A convolutional layer is used to extract features from local receptive fields in the preceding layer. In order to extract different types of local features, a convolutional layer is organized in planes of neurons called feature maps which are responsible to detect a specific feature. In a network with a 5×5 convolution kernel each unit has 25 inputs connected to a 5×5 area in the previous layer, which is the local receptive field. A trainable weight is assigned to each connection, but all units of one feature map share the same weights. This feature which allows reducing the number of trainable parameters is called weight sharing technique and is applied in all CNN layers.

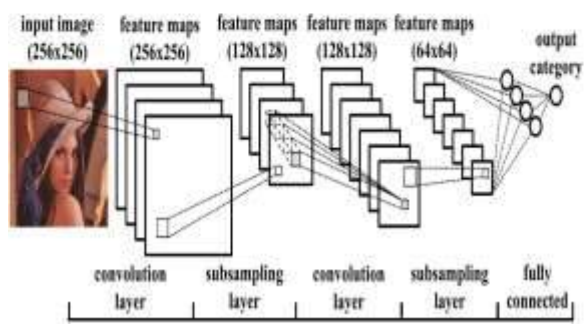

Figure 2. CNN Architecture

The previous figure showcases a 3D CNN architecture for face recognition. This architecture consists of one hardwired layer, three convolution layers, two subsampling layers, and one full connection layer. Detailed descriptions are mentioned in working section.

IV. Related Work

Deep convolutional neural community (CNN) has produced great advances in solving complex vision identification problems during recent years, inclusive of exceptional successes in recognizing natural images. There has been a lot of lookup on the use of deep CNN

IJournals

on a variety of difficult machine learning tasks [6].Recently deep Convolutional Neural Networks have been effectively utilized in many computer vision tasks and achieved promising results. So some works have introduced thedeep learning into face anti-spoofing. However, most approachesjust use the final fully-connected layer to distinguish the actual andfake faces. Inspired by the idea of each convolutional kernel canbe regarded as a part filter, we extract the deep partial featuresfrom the convolutional neural network (CNN) to distinguishthe real and fake faces.

Bong et. al[7]. Recently, proposed a method for face recognition using multi-scale convolution layer blocks and triplets of faces in unconstrained environments. We use the ensemble of deep convolution neural networks trained on differently scaled and aligned face images. This extracts low dimensional but high-level abstraction and discriminative features for face recognition. With these features, employed the jointly Bayesian model and transfer learning which adapts the knowledge trained from the source domain to target domain. Proposed method achieves 98.33% pair-wise verification accuracy on the LFW dataset.

Chen. et al. presented a dynamically reconfigurable hardware model for Convolutional Neural Network (CNN)[8]. The modular prototyping system is based on XILINX FPGAs and is capable of emulating hardware implementations of CNN for the task of face recognition. The system emulate the complex structure of CNN with exploitation of a small chip area by using the property of reconfiguration. A speedup of about 88 is achieved with FPGA modules of 50 MHz compared to a software implementation on a state of the art personal computerfor typical applications of CNN.

V. DATASETS

A. ORL Face Dataset

Database of Faces, (formerly 'The ORL Database of Faces'), contains a set of face images taken between April 1992 and April 1994 at the lab. The database was used in the context of a face recognition project carried out in collaboration with the Speech, Vision and Robotics Group of the Cambridge University Engineering Department.There are ten different images of each of 40 distinct subjects. For some subjects, the images were taken at different times, varying the lighting, facial expressions (open / closed eyes, smiling / not smiling) and facial details (glasses / no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement). The size of each image is 92x112 pixels, with 256 grey levels per pixel. The images are organised in 40 directories (one for each subject), which have names of the form sX, where X indicates the subject number (between 1 and 40).

Sample image set is shown in figure.3

Figure 3. Sample images from ORL Dataset

B. Yale Face Dataset

A close relationship exists between the advancement of face recognition algorithms and the availability of face databases varying factors that affect facial appearance in a controlled manner.Yle database Contains 165 grayscale images in GIF format of 15 individuals. There are 11 images per subject, one per different facial expression or configuration: center-light, w/glasses, happy, left-light, w/no glasses, normal, right-light, sad, sleepy, surprised, and wink.[9]

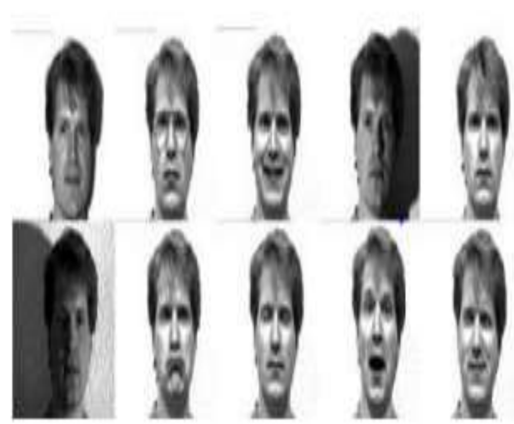

Figure 4. Sample images frm Yale face database

VI. CONCLUSION

The approaches are evaluated in this work aims at producing a more precise face classifier by means of processing.It is based on recognizing the faces from images using the convolution neural network.Since the data influence in the learned information, a good way to improve the performance of recognition methods is to use existing CNN model and face dataset were used to compare recognition performance in approaches proposed in this paper. We will apply variation in CNN layered architecture for better results when the classifier can also be customized as per our need.

References:

[1]. Alex Krizhevsky, IlyaSutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks.In Advances in neural information processing systems, pages 1097–1105, 2012.

[2]. M. Jaderberg, K. Simonyan, A. Vedaldi, and A. Zisserman. Synthetic data and artificial neural networks for natural scene text recognition.In Proc. Adv. Neural Inf. Process.Syst., 2014.

[3] Han, Xiaofeng, and Yan Li Li. "The Application Of Convolution Neural Networks In Handwritten Numeral Recognition." International Journal of Database Theory and Application Vol.8, No.3 (2015), pp.367-376.

[4] Stutz, D.: 'Understanding convolutional neural networks'. Seminar Report, FakultätfürMathematik, Informatik und Naturwissenschaften Lehr- und ForschungsgebietInformatik VIII Computer Vision, 2014.

[5] LeCun Y., Bottou L., Bengio Y., Haffner P. (1998), 'Gradient-based learning applied to document recognition', Proc. IEEE, 86(11):2278–2324.

[6] Hannun A, Case C, Casper J, et al. Deep speech: Scaling up end-to-end speech recognition [J]. arXiv preprint arXiv:1412.5567, 2014.

[7] Bong-Nam Kang, Yonghyun Kim and D. Kim, "Deep convolution neural network with stacks of multi-scale convolutional layer block using triplet of faces for face recognition in the wild," 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Budapest, 2016, pp. 004460-004465.

[8] L. Chen, X. Guo and C. Geng, "Human face recognition based on adaptive deep Convolution Neural Network," 2016 35th Chinese Control Conference (CCC), Chengdu, 2016, pp. 6967-6970.

[9] P. N. Bellhumer, J. Hespanha, and D. Kriegman.Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence, Special Issue on Face Recognition, 17(7):711--720, 1997.

IJournals