Fusion of CNN and LBP-HOG Features for Face Detection

Gopika G Das

Department of Computer Applications, Sree Narayana Guru Institute of Science and Technology, Manjali, N.Paravur, Ernakulam, Kerala, India

***Abstract*— Face recognition is widely used in security based applications. Even mobile phones and other such gadgets consider face as one of the most secure biometric application. It is necessary that the biometric authentication system needs to prevent sophisticated spoofing challenges. Advantages of deep learning, LBP-HOG and convolutional neural network are used in spoof detection.**

***Keywords*— Convolutional neural network (CNN), LBP-HOG, fake face detection, face detection, face liveliness.**

I. INTRODUCTION

Face anti-spoofing , as a security measure for face recognition system, are widely used due to the diversity of spoofing types. In some spoofing attacks no obvious visual cues are available to pick the genuine face images. Therefore much more generalized and discriminative features for face anti-spoofing, such as LBP and HOG are employed. These are called handcrafted features because they are designed manually. Also the features learned from convolutional neural networks are able to catch more discriminative signal in a data-driven manner. The success rate of deep convolutional neural network in the field of image classification and object recognition has attracted researchers to utilize these multi-layer end-to-end learning architectures to perform a variety of tasks. CNNs consist of many convolutional layers, followed by fully connected layers, to produce a probability distribution for the training classes.

The combination of feature fusion set of LBP and HOG are fed into Support Vector Machine (SVM) for classification process. For detection multiple feature vectors have been proposed like HOG, local binary pattern and haar like feature. One of the most popular classifier due to its efficiency and performance is SVM. The features extracted are used for training the model and this trained model is used for decision making. Convolutional neural network has achieved great success in the field of computer vision. Deep residual network is a successful improvement that breaks the limitation when trained with those containing hundreds of layers that work on extremely large dataset with thousands of classes and millions of samples, residual network is not easy to train comparing with CNN. Therefore, it is ideal to use a mixed convolutional layer followed by a residual layer can obtain the advantage of both kinds of architectures.

A. Architecture:

CNN's are highly layered structural neural network, most of which has the basic function layer including convolutional layers, pooling layers and a classification layer. CNN's are made of neurons that have learnable weights and biases. Each neurons receives some input, performs a dot product and optionally follows it with a non-linearity. CNN consist of several layers. Each layer takes a multidimensional array of numbers as input and produce another multidimensional array of numbers as output. When classifying images, the input to the first layer is the input image, while the output of the final layer is a set of different categories.

The framework contains 11 layers in total, including 3 convolutional layers , 3 residual layers, 2 pooling layers, 2 fully connected layer and 1 flatten layer. In residual layer, the building block contains one layer. The batches are fed into a convolutional layer, which is followed by a residual layer, and both of them contain 8 feature maps. Output of the residual layer is fed into second convolutional layer . The pooling layer is then implemented with a stride of 2 pixels. The output of the second network is send to another set of convolutional block with the same parameter as the first two. Another max-pooling layer is implemented with the same configuration as the first one before the output is fed to the flatten layer, which is followed by two connected layer, each of which has 512 neurons. The softmax technique is used after the second fully connected layer. The ReLU activation function is implemented after all the convolutional layers, residual layers and first fully connected layers. It is suitable for forward/backward propagation because of its concise form. The ReLU function, which holds the feature of non-saturating, non-linearity, is reported to run several times faster during the training process.

B. Training and Testing:

In training phase, HOG and LBP features are extracted from the preprocessed positive and negative samples, and then apply the averaging method to make them robust, these featuresare fused to make a single feature descriptor which is used with SVM train model of SVM tool to train the system. In the testing phase, detection is performed with sliding window method to extract the combination of HOG and LBP features over the test image. The feature vector obtained through the application of LBP are captured in histograms. The radius parameter used is set to 4 (pixels) and the number of points to consider are set to 14. The resulting histogram will be of 1182 dimensions for each images. Therefore, these features can be used directly for training the classifiers. Image feature extraction using HOG is then applied. 18 x 18 cell size and 1 x 1 block size are used for computing the features. The output of the HOG feature extraction is a histogram with 1224 dimensions (bins) for each image, which can be used for training classifiers.

The face detection process is done using Haar Cascading classifiers. It may be considered as a funnel wherein every region of an image is analyzed using a set of classifier called Haar Features that act as a funnel called the Haar cascade. The classifier at the top of the cascade are extremely far and have extremely ow false negative rate that immediately remove regions of an image that do not contain a face.

C. Local Binary Pattern

A particular case of texture spectrum model and a type of visual descriptor that is used for classification is known as local binary pattern (LBP). An LBP feature vector will divide the examined window into cells. For each pixel in a cell it compares the center pixel value with the 8 neighboring pixel value either clockwise or counter clockwise. If the centre pixel value is greater than the neighboring pixel value then replace it with 0 otherwise replace it with 1. This value is then converted in to decimal value for more convenience. Then compute the histogram over the cell according to the frequency of each number. This histogram is the LBP feature vector of the image. For more convenience normalize the histogram and then concatenate the normalized histograms of all cells. This will give the feature vector for the entire window. Fig 1 LBP Operator LBP is one of the methods in texture classification. The LBP codes obtained from the image are collected in the form of a histogram. The classification is then performed by computing and measuring the histogram similarities. The LBP operator denote pixels of an image by using 3x3 neighbourhood. This 3x3 neighbourhood is also known as matrix. In the matrix form rach of the pixel consists a value and this value can be varied depending upon the image and pixel quality.

II. CONCLUSIONS

Face recognition system based on CNN's have become the standard due to the significant accuracy improvement achieved over other types of methods. Face recognition system are part of facial image processing applications and their significance as a research area are increasing recently. In convolutional neural network, especially for large scale image classification the representation depth is beneficial for the classification accuracy. The performance of LBP operator along with convolutional neural network for face detection is studied. CNN architectures are capable of learning powerful features and robust to details of the connectivity of the architectures. The method presented in this review article contributes a method for face recognition based on convolutional neural network and local binary pattern by capturing the necessary facial charecteristics.

ACKNOWLEDGMENT

I gratefully acknowledge my guide for the support to complete this work and Wikipedia for the correct information source.

REFERENCES

[1]. Zi Wang, Chengcheng Li, and Huiri shao (April 2018), "Eye Recognition With Mixed Convolutional and Residual Network (MiCore-Net)", in IEEE Access, DOI 10.1109/ACCESS.2018.2812208.

[2]. Timo Ahonen, Abdenour Hadid, and Matti Pietikainen (2004). Face Recognition with Local Binary Patterns: Machine Vision Group, Infotech Oulu.

[3]. Snehal Humne and Prachi Sorte (2018). A Review on Face Recognition using Local Binary Pattern Algorithm: M.E. Dept. of I.T., R.M.D. School of Engineering, Savitribai Phule Pune University, Pune, Maharashtra, India Asst. Professor, Dept. of I.T., R.M.D. School of Engineering, Savitribai Phule Pune University, Pune, Maharashtra,India.

[4]. Rui Shao, Xiangyuan Lan, Pong C. Yuen (2017). Deep Convolutional Dynamic Texture Learning with Adaptive Channeldiscriminability for 3D Mask Face Anti-spoofing: Department of Computer Science, Hong Kong Baptist University.

[5]. Abhilasha A Patil, Lakshmi Maka, Abhilasha A Patil and Lakshmi Maka (2015). User Recognition Based on Face using Local Binary Pattern (LBP) with Artificial Neural Network (ANN) : International Journal of Ethics in Engineering & Management Education.