Fake Image Identification Using Machine Learning

Mr.A.Venkatrami Reddy [1] and Dr.K.Durga Prasad [2]
[1]Assistant Professor, Department of Information Technology, MREC (A), Hyderabad-500100
[2]Assoc. Professor, Department of Information Technology, MREC (A), Hyderabad-500100

ABSTRACT

Now-a-days biometric systems are useful in recognizing person's identity but criminals change their appearance in behaviour and psychological to deceive recognition system. To overcome from this problem we are using new technique called Deep Texture Features extraction from images and then building train machine learning model using CNN (Convolution Neural Networks) algorithm. This technique refers as LBPNet or NLBPNet as this technique heavily dependent on features extraction using LBP (Local Binary Pattern) algorithm. In this project we are designing LBP Based machine learning Convolution Neural Network called LBPNET to detect fake face images. Here first we will extract LBP from images and then train LBP descriptor images with Convolution Neural Network to generate training model. Whenever we upload new test image then that test image will be applied on training model to detect whether test image contains fake image or non-fake image. Below we can see some details on LBP. Local binary patterns (LBP) is a type of visual descriptor used for classification in computer vision and is a simple yet very efficient texture operator which labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. Due to its discriminative power and computational simplicity, LBP texture operator has become a popular approach in various applications. It can be seen as a unifying approach to the traditionally divergent statistical and structural models of texture analysis. Perhaps the most important property of the LBP operator in real-world applications is its robustness to monotonic gray-scale changes caused, for example, by illumination variations. Another important property is its computational simplicity, which makes it possible to analyze images in challenging real-time settings.

Keywords: *Fake Image, ML, Local binary patterns, CNN and LBPNet*

1. Introduction

Recently, the generative model based on deep learning such as the generative adversarial net (GAN) is widely used to synthesize the photo-realistic partial or whole content of the image and video. Furthermore, recent research of GANs such as progressive growth of GANs (PGGAN)[1] and BigGAN could be used to synthesize a highly photo-realistic image or video so that the human cannot 20 recognize whether the image is fake or not in the limited time. In general, the generative applications can be used to perform the image translation tasks [3]. However, it may lead to a serious problem once the fake or synthesized image is improperly used on social network or platform. For instance, cycleGAN is used to synthesize the fake face image in a pornography video [4]. Furthermore, GANs may be used to create a speech video with the synthesized facial content of any famous politician, causing severe problems on the society, political, and commercial activities. Therefore, an effective fake face image detection technique is desired. In this paper, we have extended our previous study associated with paper ID #1062 to effectively and efficiently address these issues.

In traditional image forgery detection approach, two types of forensics scheme are

widely used: active schemes and passive schemes. With the active schemes, the externally additive signal (i.e., watermark) will be embedded in the source image without visual artifacts. In order to identify whether the image has tampered or not, the watermark extraction process will be performed on the target image to restore the watermark [6]. The extracted watermark image can be used to localize or detect the tampered regions in the target image. However, there is no "source image" for the generated images by GANs such that the active image forgery detector cannot be used to extract the watermark image. The second one-passive image forgery detector–uses the statistical information in the source image that will be highly consistency between different images. With this property, the intrinsic statistical information can be used to detect the fake region in the image [7][8]. However, the passive image forgery detector cannot be used to identify the fake image generated by GANs since they are synthesized from the low-dimensional random vector. Nothing change in the generated image by GANs because the fake image is not modified from its original image

2. Background Study

Intuitively, we can adopt the deep neural network to detect the fake image generated by GAN. Recently, there are some studies that investigate a deep learning-based approach for fake image detection in a supervised way. In other words, fake image detection can be treated as a binary classification problem (i.e., fake or real image). For example, the convolution neural network (CNN) network is used to learn the fake image detector [9]. In [10], the performance of the fake face image detection can be further improved by adopting the most advanced CNN–Xception network [11]. However, there are many GANs proposed year by year. For example, recently proposed GANs such as [1][12][13][14][15][16][3][2] can be used to produce the photo-realistic images. It is hard and very time-consuming to collect all training samples of all GANs. In addition, such a supervised learning strategy will tend to learn the discriminative features for a fake image generated by each GAN's. In this situation, the learned detector may not be effective for the fake image generated by another new GAN excluded in the training phase.

In order to meet the massive requirement of the fake image detection for GANs-based generator, we propose novel network architecture with a pairwise learning approach, called common fake feature network (CFFN). Based on our previous approach [5], it is clear that the pairwise learning approach can overcome the shortcomings of the supervised learning-based CNN such as methods in [9][10]. In this paper, we further introduce a novel network architecture combining with pairwise learning to improve the performance of the fake image detection. To verify the effectiveness of the proposed method, we apply the proposed deep fake detector (DeepFD) to identify both fake face and generic image. The primary contributions of the proposed method are two-fold:

• We propose a fake face image detector based on the novel CFFN consisting of several dense blocks to improve the representative power of the fake image.

• The pairwise learning approach is first introduced to improve the generalization property of the proposed DeepFD.

3. Methodology

The LBP feature vector, in its simplest form, is created in the following manner:

Divide the examined window into cells (e.g. 16x16 pixels for each cell). For each pixel in a cell, compare the pixel to each of its 8 neighbors (on its left-top, left-middle, left-bottom, right-top, etc.). Follow the pixels along a circle, i.e. clockwise or counter-clockwise. Where the center pixel's value is greater than the neighbor's value, write "0". Otherwise, write "1". This gives an 8-digit binary number (which is usually converted to decimal for convenience). Compute the histogram, over the cell, of the frequency of each "number" occurring (i.e., each combination of which pixels are smaller and which are greater than the center). This histogram can be seen as a 256-dimensional feature vector.

Optionally normalize the histogram.

Concatenate (normalized) histograms of all cells. This gives a feature vector for the entire window. The feature vector can now be processed using the Support vector machine, extreme learning machines, or some other machine learning algorithm to classify images. Such classifiers can be used for face recognition or texture analysis.

Now-a-days biometric systems are useful in recognizing person’s identity but criminals change their appearance in behaviour and psychological to deceive recognition system. To overcome from this problem we are using new technique called Deep Texture Features extraction from images and then building train machine learning model using CNN (Convolution Neural Networks) algorithm. This technique refers as LBPNet or NLBPNet as this technique heavily dependent on features extraction using LBP (Local Binary Pattern) algorithm.

In this, we are designing LBP Based machine learning Convolution Neural Network called LBPNET to detect fake face images. Here first we will extract LBP from images and then train LBP descriptor images with Convolution Neural Network to generate training model. Whenever we upload new test image then that test image will be applied on training model to detect whether test image contains fake image or non-fake image. Below we can see some details on LBP.

Local binary patterns (LBP) is a type of visual descriptor used for classification in computer vision and is a simple yet very efficient texture operator which labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. Due to its discriminative power and computational simplicity, LBP texture operator has become a popular approach in various applications. It can be seen as a unifying approach to the traditionally divergent statistical and structural models of texture analysis. Perhaps the most important property of the LBP operator in real-world applications is its robustness to monotonic gray-scale changes caused, for example, by illumination variations. Another important property is its computational simplicity, which makes it possible to analyze images in challenging real-time settings.

The LBP feature vector, in its simplest form, is created in the following manner:

Divide the examined window into cells (e.g. 16x16 pixels for each cell). For each pixel in a cell, compare the pixel to each of its 8 neighbors (on its left-top, left-middle, left-bottom, right-top, etc.). Follow the pixels along a circle, i.e. clockwise or counter-clockwise. Where the center pixel's value is greater than the neighbor's value, write "0". Otherwise, write "1". This gives an 8-digit binary number (which is usually converted to decimal for convenience).

Compute the histogram, over the cell, of the frequency of each "number" occurring (i.e., each combination of which pixels are smaller

and which are greater than the center). This histogram can be seen as a 256-dimensional feature vector. Optionally normalize the histogram. Concatenate (normalized) histograms of all cells. This gives a feature vector for the entire window.

The feature vector can now be processed using the Support vector machine, extreme learning machines, or some other machine learning algorithm to classify images. Such classifiers can be used for face recognition or texture analysis.

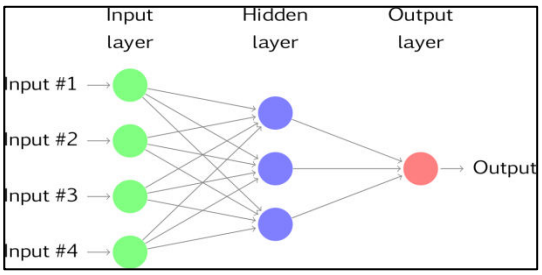
A useful extension to the original operator is the so-called uniform pattern,[8] which can be used to reduce the length of the feature vector and implement a simple rotation invariant descriptor. This idea is motivated by the fact that some binary patterns occur more commonly in texture images than others. A local binary pattern is called uniform if the binary pattern contains at most two 0-1 or 1-0 transitions. For example, 00010000 (2 transitions) is a uniform pattern, but 01010100 (6 transitions) is not. In the computation of the LBP histogram, the histogram has a separate bin for every uniform pattern, and all non-uniform patterns are assigned to a single bin. Using uniform patterns, the length of the feature vector for a single cell reduces from 256 to 59. The 58 uniform binary patterns correspond to the integers 0, 1, 2, 3, 4, 6, 7, 8, 12, 14, 15, 16, 24, 28, 30, 31, 32, 48, 56, 60, 62, 63, 64, 96, 112, 120, 124, 126, 127, 128, 129, 131, 135, 143, 159, 191, 192, 193, 195, 199, 207, 223, 224, 225, 227, 231, 239, 240, 241, 243, 247, 248, 249, 251, 252, 253, 254 and 255.

CNN working procedure

To demonstrate how to build a convolutional neural network based image classifier, we shall build a 6 layer neural network that will identify and separate one image from other. This network that we shall build is a very small network that we can run on a CPU as well. Traditional neural networks that are very good at doing image classification have many more parameters and take a lot of time if trained on normal CPU. However, our objective is to show how to build a real-world convolutional neural network using TENSORFLOW.

Neural Networks are essentially mathematical models to solve an optimization problem. They are made of neurons, the basic computation unit of neural networks. A neuron takes an input (say x), do some computation on it (say: multiply it with a variable w and adds another variable b) to produce a value (say; z= wx + b). This value is passed to a non-linear function called activation function (f) to produce the final output (activation) of a neuron. There are many kinds of activation functions. One of the popular activation function is Sigmoid. The neuron which uses sigmoid function as an activation function will be called sigmoid neuron. Depending on the activation functions, neurons are named and there are many kinds of them like RELU, TanH.

If you stack neurons in a single line, it's called a layer; which is the next building block of neural networks. See below image with layers

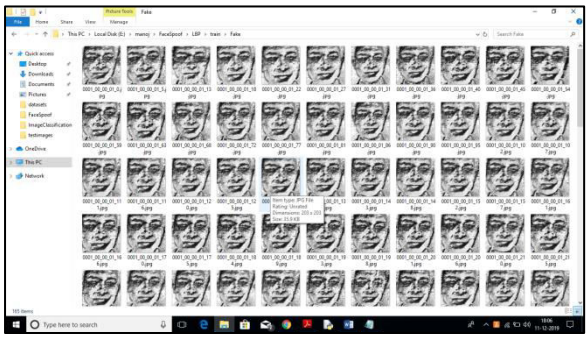


To predict image class multiple layers operate on each other to get best match layer and this process continues till no more improvement left.

4. Results Analysis

a. Dataset Details:

In this paper author has used NUAA Photograph Imposter (fake) Database with images obtained from real and fake faces. We also used images and convert that image into LBP format. Below are some images from LBP folder

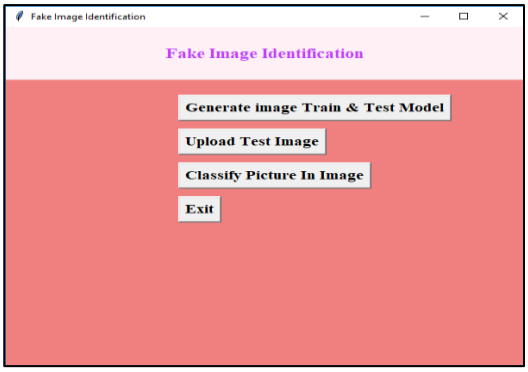


All this fake and real images you can see inside 'LBP/train' folder.

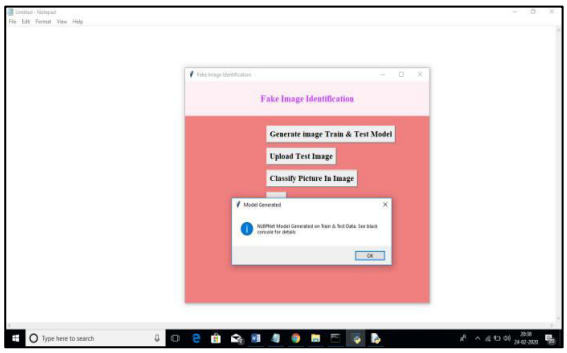
This paper consists of following modules:

1) Generate NLBPNet Train & Test Model: in this module we will read all LBP images from LBP folder and then train CNN model with all those images.
2) Upload Test Image: In this module we will upload test image from 'testimages' folder. Application will read this image and then extract Deep Textures Features from this image using LBP algorithm.
3) Classify Picture In Image: This module apply test image on CNN train model to predict whether test image contains spoof or non-spoof face.

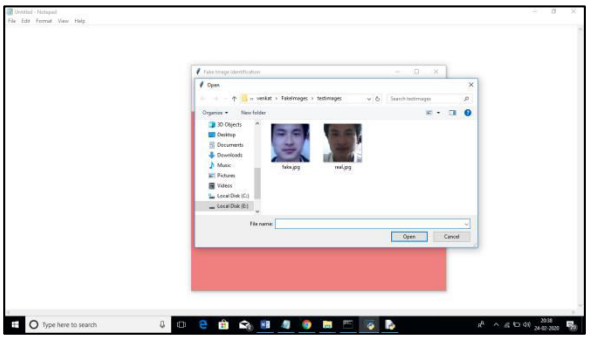
To run this project double click on 'run.bat' file to get below screen



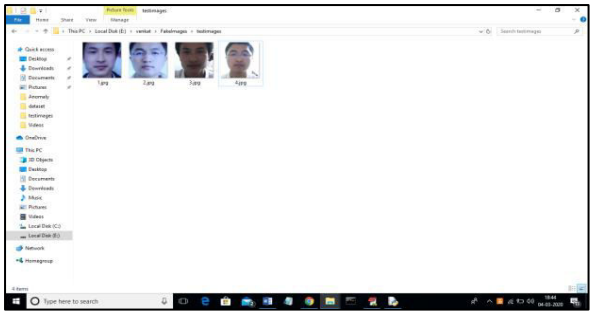
In above screen click on 'Generate Image Train & Test Model' button to generate CNN model using LBP images contains inside LBP folder.



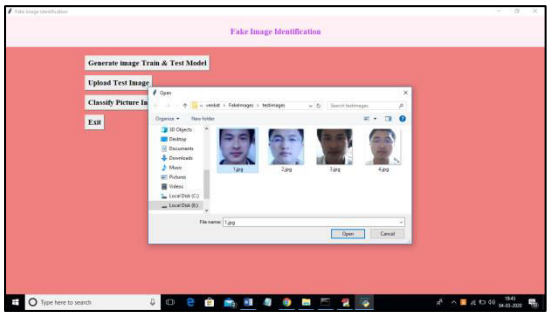
In above screen we can see CNN LBPNET model generated. Now click on 'Upload Test Image' button to upload test image



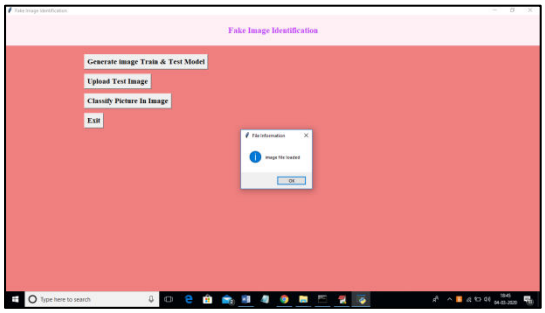
In above screen we can see two faces are there from same person but in different appearances. For simplicity I gave image name as fake and real to test whether application can detect it or not. In above screen I am uploading fake image and then click on 'Classify Picture In Image' button to get below result



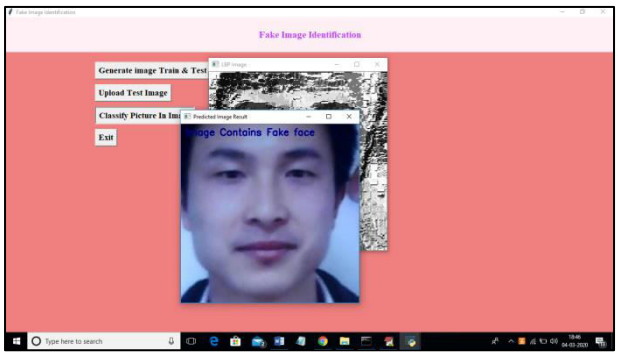
In above screen we can see all real face will have normal light and in fake faces peoples will try some editing to avoid detection but this application will detect whether face is real or fake



In above screen I am uploading 1.jpg and after upload click on open button to get below screen



And now click on 'classify Picture in Image' to get below details



In above screen we are getting result as image contains Fake face. Similarly u can try other images also. If u want to try new images then u need to send those new images to us so we will make CNN model to familiar with new images so it can detect those images also.

5. Conclusion

In this paper, we have proposed a novel common fake feature network based the pair-wise learning, to detect the fake face/general images generated by state-of-the-art GANs successfully. The proposed CFFN can be used to learn the middle- and high-level and discriminative fake feature by aggregating the cross-layer feature representations into the last fully connected layers. The proposed pair wise learning can be used to improve the performance of fake image detection further. With the proposed pairwise learning, the proposed fake image detector should be able to have the ability to identify the fake image generated by a new GAN. Our experimental results demonstrated that the proposed method outperforms other state-of-the-art schemes in terms of precision and recall rate.

References

1. Karras, T.; Aila, T.; Laine, S.; Lehtinen, J. Progressive growing of gans for improved quality, stability, and variation. arXiv Preprint, arXiv:1710.10196 2017. 256
2. Brock, A.; Donahue, J.; Simonyan, K. Large scale gan training for high fidelity natural image synthesis. arXiv Preprint, arXiv:1809.11096 2018.
3. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent 259 adversarial networks. arXiv Preprint, 2017.
4. AI can now create fake porn, making revenge porn even more complicated,. http://theconversation.com/ai-can-now-create-fake-porn-making-revenge-porn-even-more-complicated-92267, 262 2018.
5. Hsu, C.; Lee, C.; Zhuang, Y. Learning to detect fake face images in the Wild. 2018 International Symposium 264 on Computer, Consumer and Control (IS3C), 2018, pp. 388–391. doi:10.1109/IS3C.2018.00104.
6. H.T. Chang, C.C. Hsu, C.Y.a.D.S. Image authentication with tampering localization based on watermark 266 embedding in wavelet domain. Optical Engineering 2009, 48, 057002.
7. Hsu, C.C.; Hung, T.Y.; Lin, C.W.; Hsu, C.T. Video forgery detection using correlation of noise residue.

Proc. of the IEEE Workshop on Multimedia Signal Processing. IEEE, 2008, pp. 170–174.
8. Farid, H. Image forgery detection. IEEE Signal Processing Magazine 2009, 26, 16–25.
9. Huaxiao Mo, B.C.; Luo, W. Fake Faces Identification via Convolutional Neural Network. Proc. of the ACM Workshop on Information Hiding and Multimedia Security. ACM, 2018, pp. 43–47.
10. Marra, F.; Gragnaniello, D.; Cozzolino, D.; Verdoliva, L. Detection of GAN-Generated Fake Images over Social Networks. Proc. of the IEEE Conference on Multimedia Information Processing and Retrieval, 2018, 274 pp. 384–389. doi:10.1109/MIPR.2018.00084.
11. Chollet, F. Xception: Deep learning with depthwise separable convolutions. Proc. of the IEEE conference on 276 Computer Vision and Pattern Recognition 2017, pp. 1610–02357.