Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2019)
IEEE Xplore Part Number: CFP19K34-ART; ISBN: 978-1-5386-8113-8

Hybrid Deep Learning Approach For Face Spoofing Detection

Shilpa S
Department Of ECE
TKM College of Engineering
Kollam, Kerala, India
sshilpa656@gmail.com

Sajeena A
Department Of ECE
TKM College of Engineering
Kollam, Kerala, India
sajeena@tkmce.ac.in

***Abstract*—For access control in different applications and in their authentication, Face Recognition plays a vital role replacing the traditional password methods. To accurately simulate the details of physical and physiological users, crime experts are developing techniques; these are known as spoofing attacks. Along with the traditional biometric techniques robust counter measures should be introduced to prevent such thefts. For the work, deep features are extracted from the image with the help of the modified CNN (Wavelet CNN). A stacked auto encoder is introduced for spatiality reduction. Along with the hybridisation, type of spoof attack is also detected in terms of printed attack or camera attack. In order to increase the detection rate, a score based prediction is also performed presenting accurate results in finding the type of attack.**

Keywords ----- Face Recognition; Local Binary Pattern (LBP); Stacked Auto encoder; Wavelet Convolutional Neural Networks (modified CNN); Spoofing Detection.

I. INTRODUCTION

Deep learning is an unsupervised feature hierarchy learning method. The method helps in realising the extraction of feature, after the reconstruction of the original data. The neural system with multiple layers of autoencoders with output to be taken into next autoencoder is called a stacked autoencoder.

Multiple points of attacks, in the security system nowadays are exploited by the criminals. Most of the attacks in the biometric system are in form of tricking the sensor for imitating the original one, as the inner fact of the application is not required.

Due to the vast advantages such as asset cost is low, convenience, and demand and tolerability by user and also being easily available to the surrounding, including wireless ones, face is considered as the main biometric trait.

However but with just a common printed photograph obtained from worldwide network, the face recognition system, despite of all advantages suffer mostly the spoofing attack. To detect the correct type of attack a hybrid model with wavelet CNN and stacked autoencoder is proposed which provide a high clarified output.

II.EXSISTING SPOOFING DETECTION METHODS

To detect a fake face, the traditional approach extricates manmade aspect factors from images. These texture features extraction includes comparison of estimate of element of grayscale image with potency of its neighbours obtained from neighbourhood mode {P, R}, (P dictates the quota of neighbours to be considered and R is radius of the neighbourhood) and assigning a new value based on such all comparisons. A LBP descriptor follows such a process [2].

In traditional method, a histogram graph is generated based on rates from LBP based image and a tally of known images and their individual graphs are taken [3]. The Support Vector Machine (SVM) [7], predict the order of face (real or fake), after training it. Other techniques include dividing the image into bits and initiate graph for each bit and at last concatenate it. Multistage form of LBP (Multiscale Local Binary Patterns - MLBP) is an important man-made method for the face spoofing detection [9].

To characterise it, by varying the neighbourhood values of the LBP descriptor multiple man-made graphs from given image is generated. Other proposed system is rooted on dynamic local ternary patterns (DLTP) [4]. For classifying faces as real or synthetic one, DLTP compares the central pixel with its neighbours by using three labels instead of two which are used in traditional ones, which is the modified form of LBP. A LBP-CNN network is introduced [1]. The binary patterns are generated and introduced as an input to the network

III. STACKED AUTOENCODER

A very synthetic neural system for autonomous learning of economical encoding is what it deals with. This is done with the aim of spatiality reduction by illustrating a group of knowledge to the network .this process is called encoding and is illustrated in the figure 1. A group of such autoencoder when stacked

together result in a network called stacked autoencoder[12].

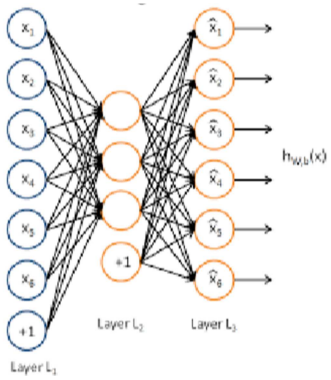


Figure 1: Structure of Autoencoder

An input stratum, concealed stratum and an output stratum are the three stratums present in an autencoder. The test image is given to the input stratum, compression of the data is done in the concealed stratum and reconstruction is done in the output stratum to get the result $h_{wb}(x)$ at out which is close to given data.

To find the important structures within the given file, during training, imposing of the dearth, on the concealed units is done. By manually zeroing or throughout coaching almost the few strongest hidden unit activations, we can achieve dearth by extra terms within the loss.

It will be either 1 or -1.When its one, the system is activated and once it's -1 the system isn't activated.

IV.CONVOLUTIONAL NEURAL NETWORKS

Deep learning architectures are linked through stratums, where variety of riddles are clinked to the given data, originally flattened images results in system known as Convolutional Neural Network (CNN).The out of given stratum serves as opener to above stratum until the head of constructed model is reached[6].

Other than the utterly bridged networks, CNNs produce a streamlined model, of which basic cell called neurons, of invariable stratums ply to allot variables, validating coherent understanding. On the head, for categorisation, with convolutional and sampling operations, stratum with cell perfectly linked is encompassed.

A flattened image is taken and in each stratum, a collection of convolutional riddles (kernels) are pertained, acquiring variety of modes of given image. A set of scale and paraphrased steadiness and reduction of data being handled is obtained when sampling operations called pooling is performed.

On the head of the topology, a high level original image representation is obtained, which is more robust than the original raw image pixel representation.

V. WAVELET CONVOLUTIONAL NEURAL NETWORK

For image classification and other image techniques in this network, two major approaches (Spectral and Spatial approaches) are launched and among different algorithms, implemented, the convolutional neural network (CNN), have recently achieved very significant improvement in several difficult tasks

CNN are considered to be a primarily abstraction approach, since it uses a method that pictures directly within the abstraction domain. Given that abstraction and spectral approaches output possess to have a totally different characteristics, to include a spectral approach in CNNs, it will be attention grabbing. Figure 2 represents the structure for a wavelet CNN [11].

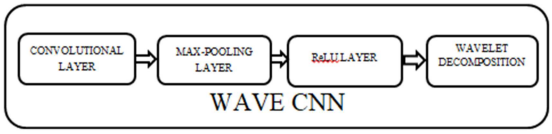


Figure 2: Wavelet CNN

A completely unique CNN design called wavelet CNN is proposed. A model is created, which comprises of a unique combination of multiple decomposition and CNNs. Our vision is that a CNN will be viewed as a restricted multiple decomposition. Based on this, through a wavelet transform we cover the missing part of multiple decomposition output, thus introduced in the form of an additional component in the topology.

For image process tasks, spectral data (wavelet information) is good and vital, that can be achieved through this wavelet CNN rather than a standard CNN. On image annotation and texture classification, evaluation of the vibrant performance of wavelet CNN is done.

On experimentations, wave CNNs are capable of producing more accurate results than the existing models in each tasks by using very few parameters than standard CNNs.

V1.SYSTEM MODEL

Novel well-referenced network model of CNN architecture, with Wave CNN and stacked auto encoder as key elements is introduced. The input is given to both stacked autoencoder and wave CNN.

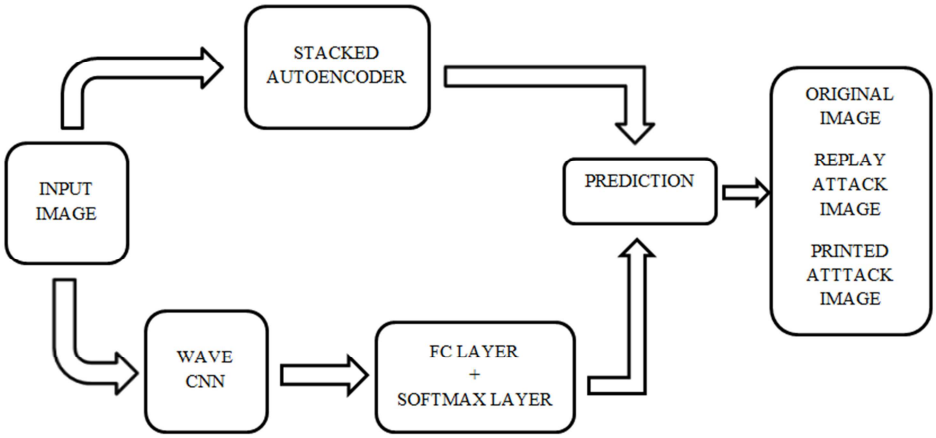


Figure 3: Block diagram

After processing, a prediction from the output of stacked autoencoder and modified CNN is obtained which is given further for the detection output.

The first stratum of network incorporates steps as follows: the convolution operation convolves the kernel value with the given scores, without finding the LBP scores of the image before convolution that is; the convolution is performed on original grayscale scores rather than on the transformed LBP.

Since it inherits the potentiality of exalting face spoofing ideas from the model, improves a lot the results of the proposed deep neural network in a bottomless and more sturdy construction, dealing with the high-level elements cultivated from the preparing data.

Data is given for convolution where wavelet decomposition of image take place .The wavelet decomposition in sense is the decomposition of image into four LL, LH, HL, HH components where L and H represents the high and low components of the image data. For further operation at pooling layer, data is forwarded.

The convolutional and pooling operations, the two mentioned operations are revised in the second stratum of the network without LBP calculation, using cores having the magnitude and steps of 5 and 1 and of 2 and 2. After the second stratum, we use 50 streamlined feature maps with size 13x13 as shown in figure 2.

The input is also provided to the stacked autoencoder where the data is compressed at the encoding side which is represented at the concealed stratum. The output of the concealed stratum is given to next autoencoder. From the stacked autoencoder, a number of values are generated. Based on the values from stacked encoder and modified CNN a score based prediction is done.

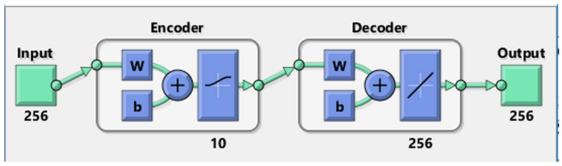


Figure 4: Structure of an Autoencoder

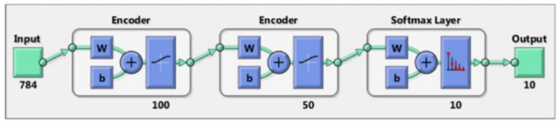


Figure 5: Structure of Stacked Autoencoder

A score based prediction is a process of generating values based on a trained machine learning model ,gives some new input data ,values or scores that are created can represent predicted values for the future.

Fully connected (FC) layer and Rectified Linear (ReLU) layer is present at the top of the architecture, in which the ReLU stratum activate by doing an inner product with the 13x13 edifices and by healing the data obtained, transferring negative rates, using equation 1:

$$\mathrm{ReLU}(x) = \max(0:x) \qquad (1)$$

where ‘x’ is the trained aggregate of signals from the later stratum of network which are valued in the 13x13 feature maps.

Before giving to the softmax layer and applying softmax function for activation, the utterly bridged stratum introduce three basic cells utterly

bidged to the units of ReLU layer where inner product operations are done.

In the softmax layer the detection result is obtained in form of whether the spoofing attacks is an original one or printed or relay attack. Corresponding to the input image given and comparison from the trained images the output is obtained.

VII. RESEARCH AND ANALYSIS

The advanced network is performed on the images from the NUAA Photograph Imposter Database, with images acquired from the collection of real dataset and the spoofed dataset. The dataset includes around 3,491 images as training images with 9,123 images as the test images .the dataset introduces these images from individuals in the name of peer group and gender and performed through different record fields making the dataset very realistic in nature.

A grayscale image with size 256x256 is taken as an input image. As shown in figure 6 a bank of around 70 test images are trained for testing. During testing the input image is compared with these test images.

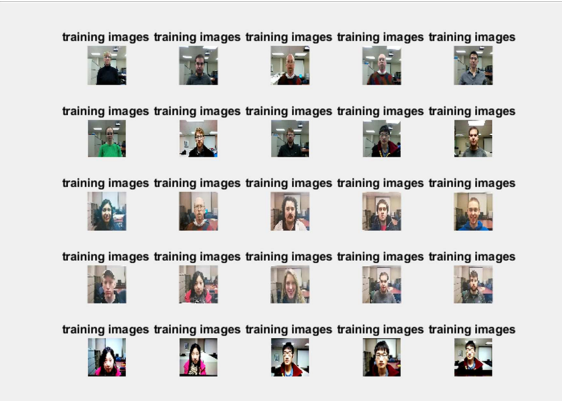


Figure 6: Folder of Trained Images

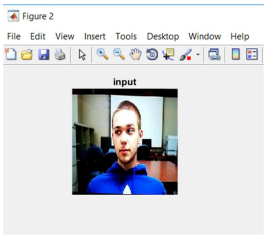
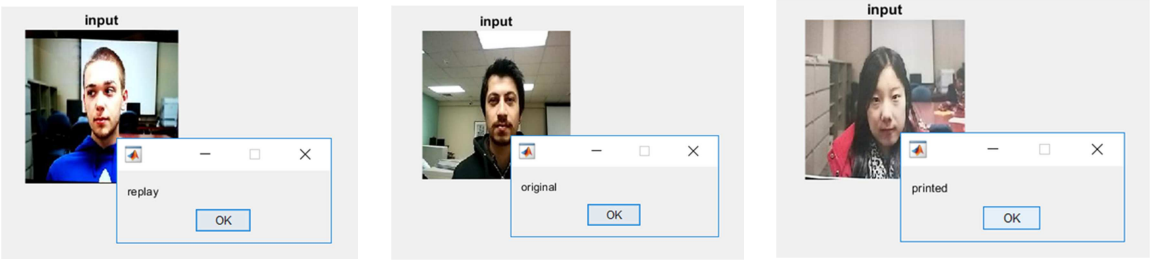


Figure 7: Input image and the LBP operation performed image



(a) (b) (c)

Figure 8: Output images showing the type of attack (a) replay attack (b) original (c) Printed attack

By passing through number of iterations, there will be comparison of given image (shown in figure 7) with that of trained image. The networks are trained for 200 iterations. The stacked autoencoder used includes hidden stratum of size 10. We had set the values of parameters as follows; L2 weight regulariser to 0.001 with sparsity regulariser to 4 and sparsity proportion to 0.05. Features are extracted from both hidden stratums of two autoencoders and then softmax layer for classification is trained.

We had stacked all the stratums together. We then defined the architecture for convolutional neural network and trained it. Testing is done to obtain the output, as a result the type of attack done in input is obtained in the message box. Figure 8 represents the output with corresponding input images defining the type of attack (original, replay or printed attacked image)

Many matrices, done using the NUAA dataset images evaluate the performance of previous work and modified work. The metrics include the very ROC (Receiver Operating Characteristics) curve and Accuracy rate (Acc).

By considering the initial normalised images, we expand the training set (magnifying its size) and perform the operations on images to obtain the relay, printed and original images.

Figure 9 dictates the curve relating to True Acceptance and False Acceptance rate, which is the

Receiver Operating Characteristic curve (ROC curve).

True acceptance rate is greater than False Acceptance rate, which shows a good ROC curve. ROC and AUC are the two vital degree units of measurements.

The probability curve and degree of separability is represented in the ROC curve .To distinguish between classes, the model helps out which is shown through the curve.

A comparison of false acceptance rate and true acceptance rate is shown in the ROC Curve. Thus it is also called as a relative operating characteristic curve.

For a better approach, ROC curve should be higher. As can be noticed, the advanced network eclipsed the prime approach, when compared with the existing systems. Based on this performance curve the output performance is measured on with the images taken from dataset .results show that proposed system is accurate in finding the class of attack.

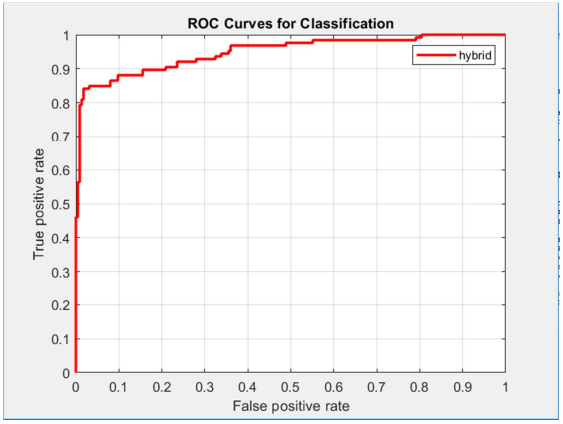


Figure 9: ROC Curve.

As the iteration increases the accuracy rate increases from 97.33% to about 99.78%, showing a great response in finding the correct type of spoofing attack with the wavelet features.

VIII. CONCLUSION

We give a completely unique CNN design which contains a spectral analysis into CNNs. We reformulated the convolution and pooling layers in CNNs into a generalized form wavelet decomposition of image with stacked encoding. We demonstrated that our model achieves better accuracy for classification. The hybrid operation produced a well-defined output when compared with previous model of using the LBP image with convolutional neural network The ROC curve also explains the same showing great accuracy in finding the correct type of spoofing done while giving an input image to the system.

REFERENCES

[1] Gustavo Botelho de Souza Daniel Felipe da Silva Santos,Rafael Goncalves Pires Aparecido Nilceu Marana,and Jo˜ao Paulo Papa,"Deep Texture Features for Robust Face Spoofing Detection "IEEE Transaction on Circuits and Systems .vol .64, no. 12,December 2017.

[2] T.Ojala, M. Pietik¨ ainen ,and T .M¨ aenp¨a¨a, "Multiresolution grayscale and rotation invariant texture classification with local binary patterns," IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 7, pp. 971-987, 2016.

[3] J. M¨ a¨att¨a, A. Hadid, and M. Pietik¨ainen, "Face spoofing detection from single images using micro texture analysis,"in Proc.of International Joint Conference on Biometrics, 2014.

[4] S. Parveen, S. M. S. Ahmad, N. H. Abbas, W. A. W. Adnan, M. Hanafi, and N. Naeem, "Face liveness detection using dynamic local ternary pattern (DLTP)," Computers, vol. 5, no. 2, 2016

[5] Menotti, G. Chiachia, A. Pinto, W. R. Schwartz, H. Pedrini, A. X. Falc˜ao, and A. Rocha, "Deep representations for iris, face, and fingerprint spoofing attack detection," IEEE Transactions on Information Forensics and Security, vol. 10, no. 4, pp. 864-879, 2015.

[6] Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Proc. of Neural Information Processing Systems, 2012, pp. 1-9.

[7] Cortes and V. N. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273297, 2005.

[8] K. Jain, S. Pankanti, S. Prabhakar, L. Hong, A. Ross, and J. L. Wayman, "Biometrics: a grand challenge," in Proc. of International Conference on Pattern Recognition, 2004, pp. 935-942.

[9] S. R. Arashloo, J. Kittler, and W. Christmas, "Face spoofing detection based on multiple descriptor fusion using multiscale dynamic binarized statistical image features," IEEE Transactions on Information Forensics and Security, vol. 10, no. 11, pp. 2396-2407, 2015.

[10] L. Deng and Y. Dong, "Deep learning: methods and applications," Found. and Trends in Signal Process., vol. 7, no. 3-4, pp. 197-387, 2014.

[11] K. Chellapilla, S. Puri, and P. Simard, "High performance convolutional neural networks for document processing," in Proc. of International Workshop on Frontiers in Hadwriting Recognition, 2006.

[12] Yahia Saeed, Jiwoong Kim, Lewis Westfall, and Ning Yang, " Handwritten Digit Recognition Using Stacked Autoencoders''Proceedings of Student-Faculty Research Day, CSIS, Pace University,2017