Liveness Detection Using OpenCV

D S Bhavani[1], P Bhavan Kumar[2]
[1] Professor, Dept of Computer Science and Engineering
[2]Dept of Computer Science and Engineering
[1, 2] Mahatma Gandhi Institute of Technology Hyderabad, Telangana, India – 500075

Abstract- *Face recognition is a widely used biometric approach. Face recognition technology has developed rapidly in recent years, and it is more direct, user friendly and convenient compared to other methods. But face recognition systems are vulnerable to spoof attacks made by non-real faces. It is an easy way to spoof face recognition systems by facial pictures such as portrait photographs. A secure system needs Liveness detection in order to guard against such spoofing. In this work, face liveness detection approaches are categorized based on the various types of techniques used for liveness detection. This categorization helps understanding different spoof attacks scenarios and their relation to the developed solutions. A review of the latest works regarding face liveness detection works is presented. The main aim is to provide a simple path for the future development of novel and more secured face liveness detection approach.*

Keywords- OpenCV, Deep Learning.

I. INTRODUCTION

The public has immense need for security measures against spoof attack. Biometrics is the fastest growing segment of such security industry. Some of the familiar techniques for identification are facial recognition, fingerprint recognition, handwriting verification, hand geometry, retinal and iris scanner. Among these techniques, the one which has developed rapidly in recent years is face recognition technology and it is more direct, user friendly and convenient compared to other methods. Therefore, it has been applied to various security systems [1]. But, in general, face recognition algorithms are not able to differentiate 'live' face from 'not live' face which is a major security issue. It is an easy way to spoof face recognition systems by facial pictures such as portrait photographs. In order to guard against such spoofing, a secure system needs liveness detection.

Biometrics is the technology of establishing the identity of an individual based on the physical or behavioral attributes of the person [3]. The importance of biometrics in modern society has been strengthened by the need for large-scale identity management systems whose functionality depends on the accurate deduction of an individual's identity on the framework of various applications. Some examples of these applications include sharing networked computer resources, granting access to nuclear facilities, performing remote financial transactions, or boarding a commercial flight [4]. The main task of a security system is the verification of an individual's identity. The primary reason for this is to prevent impostors from accessing protected resources. General techniques for security purposes are passwords or ID cards mechanisms, but these techniques of identity can easily be lost, hampered, or may be stolen thereby undermine the intended security. With the help of physical and biological properties of human beings, a biometric system can offer more security for a security system. But in face recognition, approaches are very much limited to deal with this problem.

Liveness is the act of differentiating the feature space into live and non-living. Imposters will try to introduce many spoofed biometrics into system. With the help of liveness detection, the performance of a biometric system will improve. It is an important and issue which determines the trustworthiness of biometric system security against spoofing. In face recognition, the usual attack methods may be classified into several categories [5]. The classification is based on what verification proof is provided to face verification system, such as a stolen photo, stolen face photos, recorded video, 3D face models with the abilities of blinking and lip moving, with various expressions and so on. Anti-spoof problem should be well solved before face recognition systems could be widely applied in daily life [6].

II. LITERATURE SURVEY

In face recognition systems, the face features of authenticated system are extracted to develop a system. But these systems may recognize the person from photos. This is a system failure. Hence to improve the face recognition system, face liveness detection plays vital role. Proposed method frame is extracted from input stream. After that, image is converted into gray scale, because the input to haar cascade is gray scale image. Haar cascade load into the system for detecting the frontal face. Now different features are collected from extracted face image. The SVM, KNN algorithm shows 77.41% and 97.69% accuracy. The proposed system is implemented using OpenCV library. In future, the deep learning algorithm can be implemented to improve the accuracy and to minimize the dependency on the feature engineering. Image should be capture on different light conditions for better accuracy. In future work can be carried out on efficient method to capture quality image which will help to have more accurate output as well as test cases should be analyzed.

Developed by Paul Viola and Michael Jones, the Viola-Jones algorithm is an object-recognition framework that allows the detection of image features in real-time. Despite being an outdated framework, Viola-Jones is quite powerful, and its application has proven to be exceptionally notable in real-time face detection. Viola-Jones was designed for frontal faces, so it is able to detect frontal the best rather than faces looking sideways, upwards or downwards. The Viola-Jones algorithm first detects the face on the grayscale image and then finds the location on the colored image. Viola-Jones outlines a box and searches for a face within the box. It is essentially searching for these haar-like features, which will be explained later. The box moves a step to the right after going through every tile in the picture. In this case, used a large box size and taken large steps for demonstration, but in general, you can change the box size and step size according to your needs. With smaller steps, a few boxes detect face-like features and the data of all of those boxes put together, helps the algorithm determine where the face is.

The study implements an eye-blink detection-based face liveness detection algorithm to thwart photo attacks. The algorithm works in real time through a webcam and displays the person’s name only if they blinked. In layman’s terms, the program runs as follows, detect faces in each frame generated by the webcam, for each detected face, detect eyes, for each detected eyes, detect if eyes are open or closed, if at some point it was detected that the eyes were open then closed then open, we conclude the person has blinked and the program displays its name. The system would work well to distinguish between known faces and unknown faces so that only authorized persons have access. Nonetheless, it would be easy for an ill-intentioned person to enter by only showing an authorized person’s photo.

The technique of anti-spoof problem as a binary classification problem was introduced by Tal. The key approach which the authors have used is that a real human face is different from a face in a photo. A real face is a 3D object while a photo is 2D by itself. The surface roughness of a photo and a real face is different. The authors presented a real-time and non-intrusive method to address this based on individual images from a generic web camera. The task is being formulated as a binary classification problem, in which, however, the distribution of positive and negative are largely overlapping in the input space, and a suitable representation space is found to be of great importance. Using the Lambert Ian model, they proposed two strategies to extract the essential information about different surface properties of a live human face or a photograph, in terms of latent samples. Based on these, two new extensions to the sparse logistic regression model were employed which allow quick and accurate spoof detection.

Here, liveness detection approaches are categorized based on the type of liveness indicator used to assist the liveness detection of faces. Three main types of indicators were mainly used: motion, texture, and life sign. This work helps to generate robust and highly secure liveness detection method. Methods are based on motion, texture and life sign of user which is used for face liveness detection and discussed some advantages and disadvantages which makes easy to select method of working for future work.

As Face Recognition (FR) technology becomes more mature and commercially available in the market, many different anti-spoofing techniques have been recently developed to enhance the security, reliability, and effectiveness of FR systems. As a part of anti-spoofing techniques, face liveness detection plays an important role to make FR systems be more secured from various attacks. In this paper, we propose a novel method for face liveness detection by using focus, which is one of camera functions. In order to identify fake faces (e.g. 2D pictures), this approach utilizes the variation of pixel values by focusing between two images sequentially taken in different focuses. The experimental result shows that this focus-based approach is a new method that can significantly increase the level of difficulty of spoof attacks, which is a way to improve the security of FR systems.

This paper proposes a single image-based face liveness detection method for discriminating 2-D paper masks from the live faces. Still images taken from live faces and 2-D paper masks were found to bear the differences in terms of shape and detailedness. In order to effectively employ such differences, we exploit frequency and texture information by using power spectrum and Local Binary Pattern (LBP), respectively. In the experiments, three liveness detectors utilizing the power spectrum, LBP, and fusion of the two were trained and tested with two databases which consist of images taken from live and four types of 2-D paper masks. One database was acquired from a web camera while the other was from the camera on the automated teller machine. Experimental results show that the proposed methods can efficiently classify 2-D paper masks and live faces.

III. METHODOLOGY

A. System Architecture

As shown in Figure 1, a system Architecture consists of system components and the sub systems developed that will work together to implement overall system. The proposed model will be identifying the real and the spoofed images as the output.

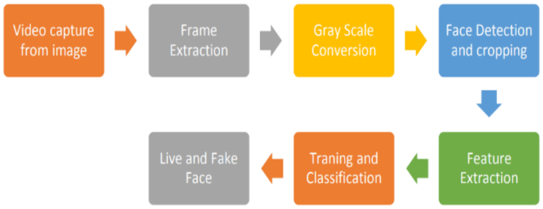

Figure 1: System Architecture

Initially a video will be taken as an input. Frame Extraction will be done simultaneously, Then the colour image will be converted to Grey Scale image because if we use the coloured images then the complexity of the proposed model will be increased, so that the coloured images are converted into grey scale images to reduce the complexity of the proposed system. The part of the image that the face is containing will be detected and cropped. Only the part of the face is extracted, and the other part of the image will be removed, this is known as Feature Extraction.

Then the model will be built, and the model will be trained based on the dataset that is containing real and spoofed images. Finally, after completion of training the model, then the model is going to predict the real and spoofed images.

I. Data Preprocessing Phase

In the proposed system deep down the system architecture can be divided into two parts i.e., data preprocessing and training the model. In the first step data preprocessing, a video will be taken as the input and then the object detection algorithm Voila Jones is used. Voila Jones is the object detection algorithm but we are extending this algorithm to detect the face. Before detecting a face, the image is converted into grayscale, since it is easier to work with and there’s lesser data to process. The Viola-Jones algorithm first detects the face on the grayscale image and then finds the location on the coloured image. The part of the image that the face is containing will be cropped as shown in Figure 2. Finally the part of the image that is containing face will be saved as the Region of Intrest.

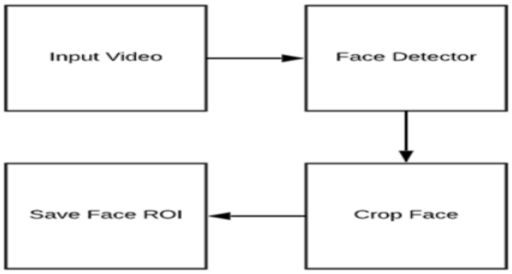

Figure 2 : Data Preprocessing Phase

II. Training Phase

In Training phase, a CNN model will be built to recognize the real and spoofed images as shown in Figure 3. This CNN model will be used with the Python script inorder to predict the output(i.e., distinguish between real and spoofed images). Liveness net is the CNN model build to perform liveness detection action.

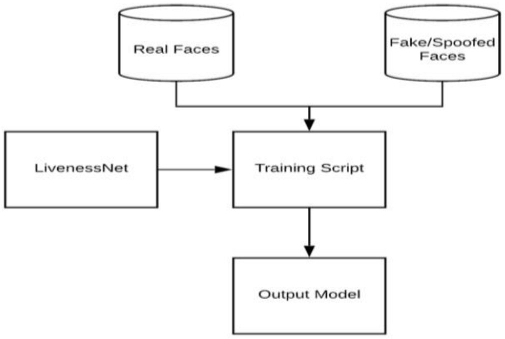

Figure 3 : Training Phase

B. VOILA JONES ALGORITHM

Viola-Jones was designed for frontal faces, so it can detect frontal the best rather than faces looking sideways, upwards or downwards. It compares how close the real World scenario is to the ideal Haar-like feature. The algorithm makes use of integral images to reduce effort while comparing two regions.Training classifiers are used to set thresholds above which a certain area of the face will be considered a Haar feature. This is done by converting the image into a 24x24 image and once these features are found, are magnified and their proportions are changed. It uses Haar-Like Features to detect parts of a face. There are 3 types of Haar-like features that Viola and Jones identified in their research:

- Edge features
- Line-features
- Four-sided features

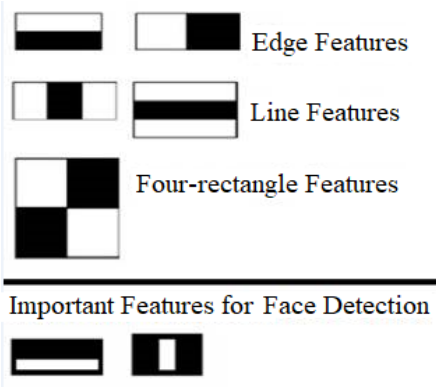

Figure 4 : Haar-Like Features

These features help the machine understand what the image is. Imagine what the edge of a table would look like on a b&w image. One side will be lighter than the other, creating that edge like b&w feature as shown in Figure 4.

Voila Jones algorithm extracts the part of the face which are required for recognizing real and spoofed images. For instance consider the part of the face i.e., eyebrows it will be consisting of two sides like one is brighter side and the part between the eyebrows and the eye will be considered as the dark side in which it will be in 2 – dimension it will be considered as the edge features as shown in Figure 5. Coming to the nose it will be considered as 3 – dimensional because it will be consisting of two dark sides and one brighter side. This is how Haar like features are extracted from the image.

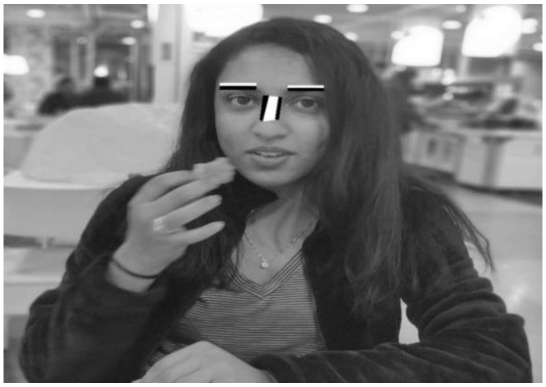

Figure 5 : Face recognized by algorithm

C. EYE ASPECT RATIO

The model showed inaccuracies when it was trained with videos of a person who belonged to one ethnicity, but was asked to predict the liveness of a person of another ethnicity.Hence, in order to increase its accuracy, some other features were also considered alongside the result of the CNN. These features were to be added in both the result of the CNN and this additional feature would have to be positive in order for it to show the person as live. The concepts that we used was Eye-Aspect Ratio or EAR is based on the concept of facial landmarks. These are used to detect particular parts of the face, (the eyes) in this case.

Detecting facial landmarks is a subset of the shape prediction problem. Given an input image (and normally an ROI that specifies the object of interest), a shape predictor attempts to localize key points of interest along the shape. It is a two step process:

1. Detecting the face
2. Detecting the key features in the face ROI (Region of Interest)

$$\mathrm{EAR} = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|}$$

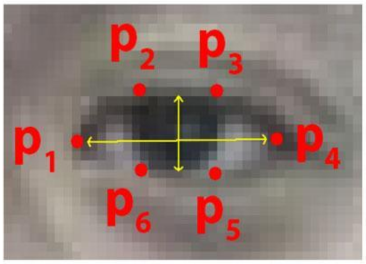

Figure 6: Eye Aspect Ratio

By using the above formula as shown in Figure 6, Eye Aspect Ratio will be calculated. Basically EAR will be used for eye blink detection. The ratio will be different for different cases like when the eye is open the EAR ratio will be different and when the eye is closed the EAR ratio will be different. To over come the ethnicity issue using voila jones we will be using this EAR ratio. So that there will be no inaccuracies.

D. CONVOLUTIONAL NEURAL NETWORK

A **Convolutional Neural Network (ConvNet/CNN)** is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNetshave the ability to learn these filters/characteristics. The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlap to cover the entire visual area. The role of the ConvNet is to reduce the images into a form which is easier to process, without losing features which are critical for getting a good prediction. This is important when we are to design an architecture which is not only good at learning features but also is scalable to massive datasets. The objective of the Convolution Operation is to **extract the high-level features** such as edges, from the input image. ConvNets need not be limited to only one Convolutional Layer. Conventionally, the first ConvLayer is responsible for capturing the Low-Level features such as edges, color, gradient orientation, etc.

It can be divided into three parts:

- Extracting regions of interest (images) from videos (two videos – fake and real) that are provided by the user frame wise and storing them as two different training datasets.
- Using these datasets to train a Convolutional Neural Network to identify whether the frames provided by the live video camera are real or fake based on the training it received (binary classification problem).
- Starting a live video stream and analysing it frame wise with the help of the trained CNN.

IV. RESULTS

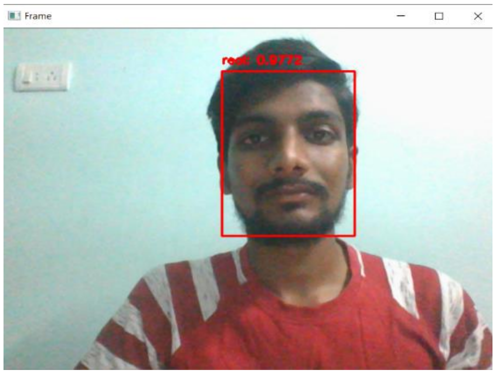

Figure 7 : Model detecting the face as Real

The model will first access the webcam and detect the part of the image that is conatining the face part. Other than face, other part of the image will be ignored. As shown in Figure 7, my face is predicted as Real

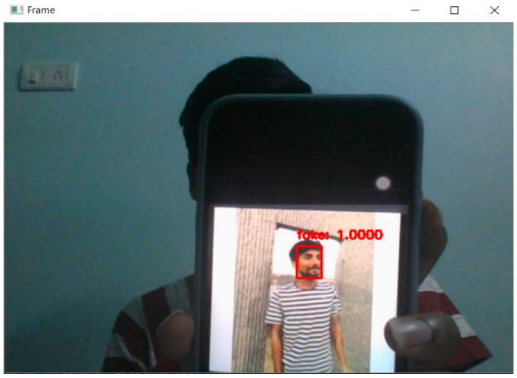

Figure 8 : Model detecting the face as Fake

As show in the Figure 8, I have kept the pic from my phone infront of the web cam. The model is detecting the image as Fake. Since the image is viewed from the mobile, it is not real. Hence the model is predicting the face as Fake.

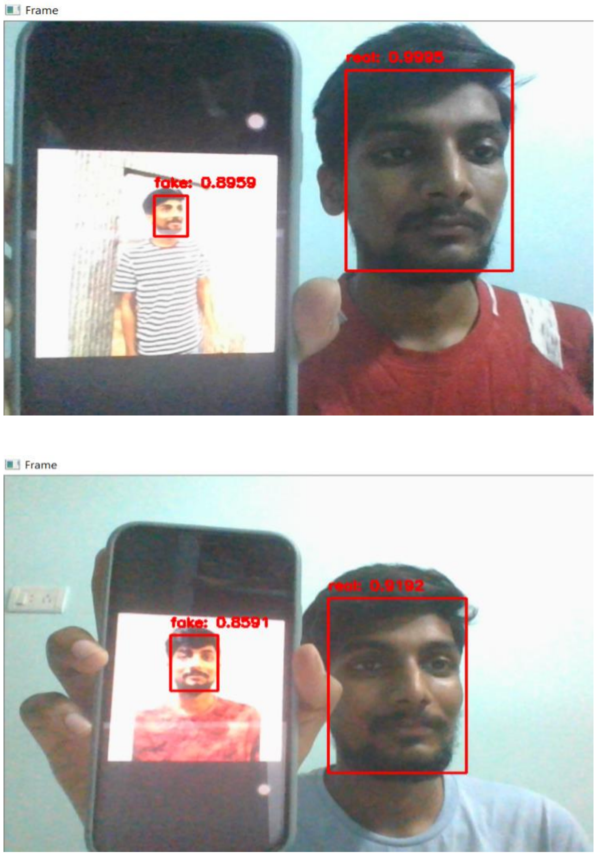

Figure 9 : Model is predicting the Real and Fake faces

As shown in Figure 9, I have used two different pics from the mobile. The model is predicting the real and spoofed(fake) images parallely. I have kept my face and the photo from my mobile in the frame. Then my model is predicting the Real(my face) and Spoofed(mobile pic) images parallely.

V. CONCLUSION

As the technology are blooming with emerging trends the model architecture had been modified to utilize 3D kernels rather than the standard 2D filters, enabling the model to include a temporal component for liveness detection. Finally, implementation of liveness detection is done using OpenCV’s keras module. Based on the results, the liveness detection model is performing efficiently.Difficulty which often present in an applied machine learning scenario is to define a problem accurately. It directly affects the type of algorithms a machine learning engineer may deploy, be it is a supervised or unsupervised learning, or a deep learning neural network.

VI. FUTURE SCOPE

As for future work, the aim is to improve sequential learning adaptive capability; thus, the deep learning has capability to continuously learn while at the same time doing verification (learning by doing). Transfer learning approach based on existing models to perform activity Liveness Detection on large-scale data may be a potential working direction. It may be used in various fields like let us take an example of police field. There are some databases which contains crucial information. There will be only limited access to that database. The higher authority people only be given access. Then some hacker will be using the official person face in the photo, and he will try to gain the access to that databases. Which may lead to data leak threat. The hacker may misuse this information. In those cases we can use this model and detect whether the image or the pic that is used to gain the access is real or spoofed. So that we can overcome those issues which are mentioned above.

REFERENCES

[1] Sanjay Ganorkar, SupriyaRajankar and Gaurav Rajpurohit, “Face Liveness Detection using Machine Learning”, in International Journal of Scientific and Technology Research Volume 8, Issue 09, September 2019.

[2] Rohan Gupta, “Breaking Down Facial Recognition: The Viola-Jones Algorithm,” in Towards Data Science, August 2019.

[3] Yi-Qing. Wang, "An Analysis of the Viola-Jones Face Detection Algorithm", Image Processing Online, vol. 4, pp. 128-148, 2014.

[4] Jordan Van Eetveldt, “Real-time face liveness detection with Python, Keras and OpenCV,” Towards Data Science, March 4 2019.

[5] D.P. Garud and S.S. Agrawal, "A Review: Face Liveness Detection”, International JouranalOf Advanced Research in Computer and Communication Engineering, vol. 5, no. 1, 2016.

[6] Saptarshi Chakraborty and Dhrubajyoti Das, “An overview of Face Liveness Detection,” in International Journal on Information Theory (IJIT), Vol.3, No.2, April 2014.

[7] G. Kim, S.Eum, J. K. Suhr, D. I. Kim, K. R. Park, and J, Kim, “ Face liveness detection based on texture and frequency analyses”, 5th IAPR International Conference on Biometrics (ICB), New Delhi, India. pp. 67-72, March 2012.

[8] J. Maatta, A. Hadid, M. Pietikainen, “Face Spoofing Detection From Single images Using MicroTexture Analysis”, Proc. International Joint Conference on Biometrics (UCB 2011), Washington, D.C., USA.

[9] Sooyeon Kim, Sunjin Yu, Kwangtaek Kim, Yuseok Ban, Sangyoun Lee, “Face liveness detection using variable focusing, Biometrics (ICB)”, 2013 International Conference on, On page(s): 1 – 6, 2013.

AUTHORS PROFILE

MsD S Bhavani1, B.Tech(CSE), M.Tech(CS), pursuing her Ph.D. in KoneruLakshmaiah Education Foundation(KLH Deemed to be University) at Hyderabad. She has teaching experience of 16 years. Currently working as Assistant Professor in the Department of Computer Science and Engineering, Mahatma Gandhi Institute of Technology, Hyderabad. She has various research papers published in the International Journals of repute. Her research area consists of Network Security, Internet of Things, Digital Forensics, Machine Learning, AI etc.

Correspondence Address:
Mahatma Gandhi Institute of Technology,
Chaitanya Bharathi Post, Gandipet,
RangaReddy District, Hyderabad-500075

Mr. P Bhavan Kumar2, 17261A0542, a Final year Student of Bachelor’s in Engineering in the field of Computer Science at Mahatma Gandhi Institute of Technology, Hyderabad. His areas of interest include Web Technologies, Data Mining, and Machine learning.