Anti-Spoofing Face lock with NFT

J Component Project

Winter Semester 2022

Rupin Patel 19BCB0030
Ishan Bhardwaj 19BCI0194
Aditya Singh 19BCE2242

B.Tech. Computer Science and Engineering

School of Computer Science and Engineering

Vellore Institute of Technology

Vellore

April , 2022

Abstract

Motivation:
In an increasingly digital world, protecting confidential information from hackers and unauthorized individuals is becoming more difficult and the need for robust security is paramount. As a result, Biometric spoofing is a growing concern as biometric traits are vulnerable to attacks. So, to counter such miscellaneous threats to biometric security, we are going to try and develop technology such that biometric spoofing isn't a concern.

Aim:
Our aim is to build an MFT- Vault with Facial anti-spoofing verification which will preventing false facial verification by using a photo, video, mask or a different substitute for an authorized person's face such as print attacks, replay/video / 3D mask attack etc. Deep learning and convolutional neural network (CNN) are additional solutions that can help with anti spoofing. Thinking of anti-spoofing as a binary classification problem when exploring technologies. We may train the CNN to recognize which are real photos and which are spoofed. And it should work. Not only facial biometric scanner user will also experience NFT authentication provided by Ethereum. We will also try to include VGG16 instead of traditional face detection technique. So that we can get more accuracy at higher frame rates also.

Methodology:
Biometric spoofing is the ability to fool a biometric system into recognizing a fake user as a genuine user by means of presenting a synthetic forged version of the original biometric trait to the sensor. Specific countermeasures that allow biometric system to detect fake artifacts and to reject them need to be developed. Our main goal is to provide an overview of different anti-spoofing techniques used in the now emerging field of anti-spoofing with special attention to face modality. For CNN we are using VGG16 to get better accuracy and type featuring.

Expected Outcomes:
To securely store data, which can only be accessed using facial recognition, and make it more air tight by developing our own countermeasures against biometric face spoofing.

Introduction

The performance of face detection and recognition systems have improved drastically in the last few years. Therefore, this innovation is currently considered as a developed system and is used in numerous real-world applications from banking security to smart house systems and device authentication. However, several studies show that this kind of system suffers from vulnerabilities to fake face spoofing attacks, a disadvantage that may restrict their use in many real-time scenarios. Indeed, it is a very tough task to protect against spoofs based on a static photo of a face, while the most effort of the present face recognition study has been focused on the "image matching" part of the system without noticing whether the corresponding face is live or fake. Many face liveness detection techniques have been proposed to restrain the face recognition systems against this kind of occurrences. These techniques have shown good performances on the existing face presentation attack databases. Besides, their performances deteriorate radically under real-world variations (e.g., illumination and camera device variations).

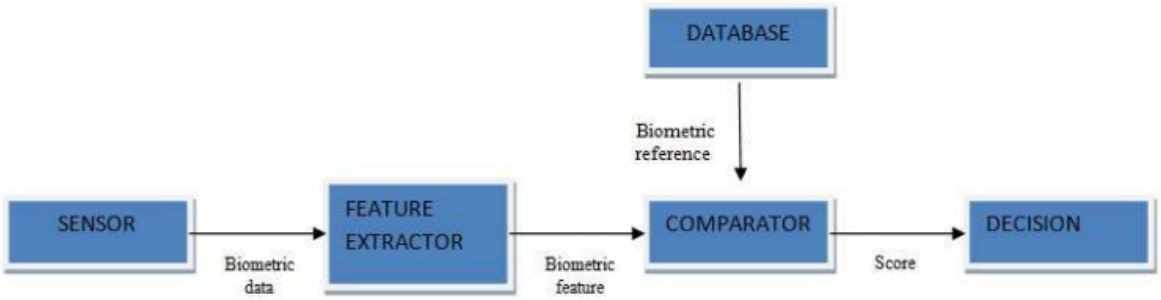


Fig-1 Sensor level Technique

Face biometrics is the second largest biometric used, with fingerprint being the first. Hence, it is more open to spoofing attacks or direct (presentation) attacks in which intruders use synthetically produced artifact or try to mimic the behavior of genuine users, to fraudulently gain access to the biometric system. Certain countermeasures have to be implemented in the form of anti spoofing methods in order to make biometric verification more secure. An anti spoofing technique is normally acknowledged to be any procedure, which can consequently recognize genuine biometric attributes displayed to the sensor from fake biometric characteristics.

Anti-spoofing techniques and what are the advantages and disadvantages of different techniques:
Sensor Level Technique: Usually referred to in the literature by the term hardware-based techniques. These methods add some specific device to the sensor in order to detect particular properties of a living trait (e.g., facial thermogram, blood pressure, fingerprint sweat, or specific reflection properties of the eye).It measures one of three characteristics, namely:

>Intrinsic properties of a living body - which could include properties like physical, electrical, spectral or visual properties.

> Involuntary signals of a living body eg. blood pressure, perspiration, electric heart signals

> Responses to external stimuli, also referred to as challenge-response methods, which requires the cooperation from the user as these responses are based on detecting voluntary (behavioral) or involuntary (reflex reactions) to an external signal.
Eg. When light is switched on the pupil contracts (reflex), or the head moves following a random path determined by the system (behavioral). Multibiometric anti spoofing is based on the assumption that the blending of various biometrics will decrease the vulnerability to assaults, as, in principle, producing multiple fake characteristics is more difficult than generating an individual fake characteristic. Based on this assumption, multimodal approaches fuse different Modalities.

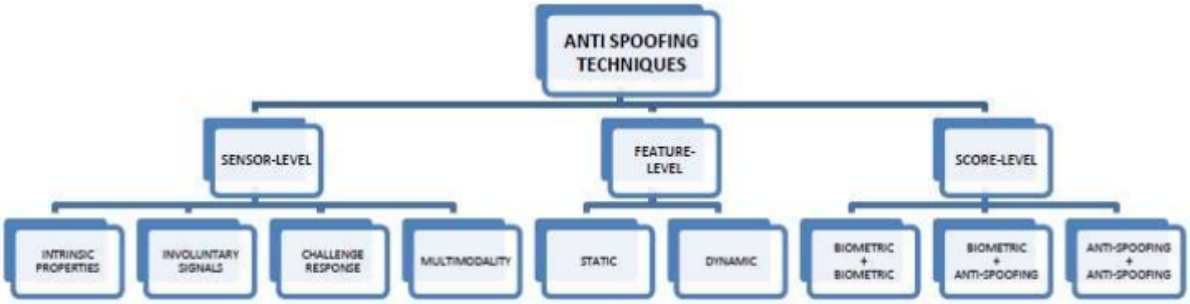


Fig-2 Types of methods to detect spoof faces

The strategy is using complementary traits for eg. Fingerprint and finger veins, this strategy requires additional hardware devices, therefore, these techniques may be included in the sensor-level group of anti-spoofing methods. The above assumption of fooling a multibiometric system has already been shown to be untrue as, in many cases, bypassing just one of the unimodal subsystems is enough to gain access to the complete application. Hence, multi biometry by itself does not necessarily guarantee a higher level of protection against spoofing attacks.

Feature-Level Technique:
Otherwise referred to as software-based techniques, here, the biometric data is acquired with a standard sensor and the distinction between fake and real faces is software based. Under Software based techniques there are two methods for anti spoofing - static and dynamic. Static features may present some degradation in performance but are still preferred over dynamic techniques because they are faster and less intrusive as they require less cooperation from the user. Static anti spoofing methods work on single images while dynamic anti spoofing methods work on video sequence. In feature level technique, multimodality can be implemented. From just one single high resolution image of a face, both face and iris recognition can be performed. It not only detects spoofing attacks but it also is capable of detecting other types of illegal break-in attempts. For eg. Feature level techniques protect the system against the injection of reconstructed or synthetic samples . The advantages of Feature-level dynamic are - It has a high

accuracy level. It exploits spatial and temporal features in a video sequence. It is known to be very effective against photo attacks. The disadvantages are – Cannot be used in single image scenario instances. It is comparatively slow. Accuracy is lost against video attacks. The advantages of Feature-level static are – It can not only be used with a video sequence but also can be used for single images. Faster when compared to Feature level dynamic technique. It is totally transparent to the user. The disadvantages are – It is based only on image spatial information which reduces the accuracy.

Score Level Technique:

Recently, a third group of protection methods which falls out of the traditional two-type classification (software- and hardware-based), has started to be analyzed in the field of fingerprint anti-spoofing. These protection techniques, much less common than the previous two categories, are focused on the study of biometric systems at score-level in order to propose fusion strategies that increase their resistance against spoofing attempts. Due to their limited performance, they are designed as supplementary measures to the sensor-level and feature-level techniques presented above, and are usually integrated in the matcher. The scores to be combined may come from:

i) two or more unimodal biometric modules;

ii) unimodal biometric modules and anti-spoofing techniques; or

iii) only results from anti-spoofing modules.

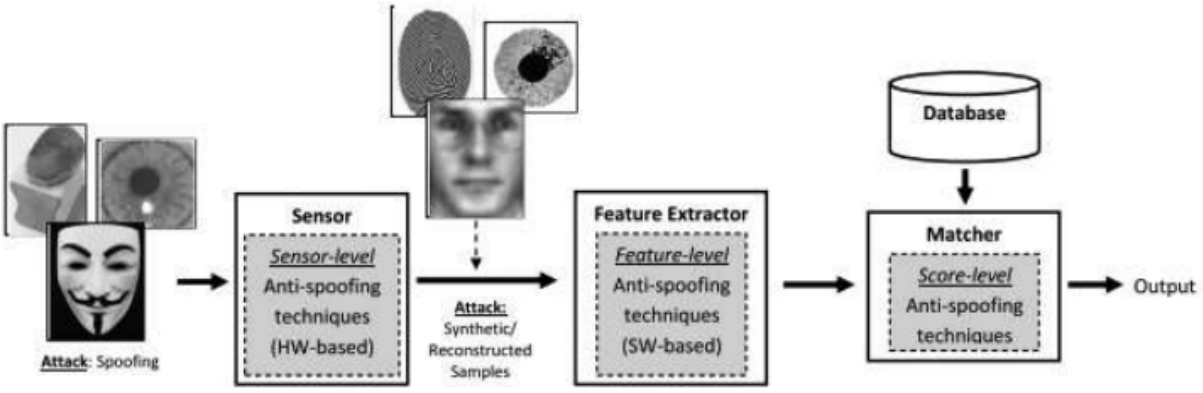


Fig-3 Score level Architecture

The disadvantages are – It is generally slower. Higher level of cooperation is required from the user. It is expensive due to the additional hardware that is required to process the biometric traits.

Literature Survey / Related Works

S. no.	Heading	About	Link
1	A Performance Evaluation of Convolutional Neural Networks for Face Anti Spoofing	The performance evaluation experimental setup in terms of the face anti-spoofing framework using CNN, hyperparameter settings, evaluation criteria and face spoofing database used.	https://ieeexplore.ieee.org/document/8852422
2	FACE ANTI-SPOOFING BASED ON MULTI-LAYER DOMAIN ADAPTATION	Maximum Mean Discrepancy (MMD) to multi-layer network distribution adaptation, which improves the generalization ability of the model. To further improve the performance of face anti-spoofing detection, we fuse the low-level features with the high-level features of convolutional neural networks for face anti spoofing detection.	https://ieeexplore.ieee.org/document/8795006
3	MULTIMODAL FACE SPOOFING DETECTION VIA RGB-D IMAGES	This research presents a novel multimodal face anti-spoofing method, which makes full use of available information on RGB-D images and no manually chosen regions are needed.	https://ieeexplore.ieee.org/document/8545849
4	3D Convolutional Neural Network Based on Face Anti-Spoofing	2017 2nd International Conference on Multimedia and Image Processing.	https://ieeexplore.ieee.org/document/8221060
5	FACE ANTI-SPOOFING BASED ON COLOR TEXTURE ANALYSIS	Analyze the joint color-texture information from the luminance and the chrominance channels using a color local binary pattern descriptor. More specifically, the feature histograms are extracted from each image band separately. Extensive experiments on two benchmark datasets, namely CASIA face anti spoofing and ReplayAttack databases, showed	https://ieeexplore.ieee.org/document/7351280

S. no.	Heading	About	Link
		excellent results compared to the state-of-the-art.	
6	Face Anti-Spoofing With Deep Neural Network Distillation	Capture spoofing-specific information and train a discriminative deep neural network on the application specific domain. Extensive experiments validate the effectiveness of the proposed scheme in face anti spoofing setups.	https://ieeexplore.ieee.org/document/9115256
7	Eyeblink-based Anti-Spoofing in Face Recognition from a Generic Web Camera	Formulate blink detection as inference in an undirected conditional graphical framework, and are able to learn compact and efficient observation and transition potentials from data. For the purpose of quick and accurate recognition of the blink behavior, eye closity, an easily-computed discriminative measure derived from the adaptive boosting algorithm, is developed, and then smoothly embedded into the conditional model.	https://ieeexplore.ieee.org/document/4409068
8	An Improved Face Liveness Detection Algorithm Based on Deep Convolutional Neural Network	Based on the traditional VGG-11 model, we propose an improved deep convolutional neural network which can accurately detect the face liveness of a single face image. Firstly, the training data set is enhanced by some methods such as random rotation, random brightness and saturation adjustment, which can improve the generalization ability of the network.	https://ieeexplore.ieee.org/document/9601431
9	Detection of 3D Mask in 2D Face Recognition System Using DWT and	Introduced a new approach to detect presence of 3D mask based face anti-spoofing using	https://ieeexplore.ieee.org/document/8644807

S. no.	Heading	About	Link
	LBP	frequency and texture based feature descriptors. The proposed approach extracts Local Binary Pattern based texture features from discrete wavelet transformed images.	
10	Application of Active Learning in Face Liveness detection	X. Zhao, "Application of Active Learning in Face Liveness detection," 2021 2nd International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE).	https://ieeexplore.ieee.org/document/9696139
11	An Improved Face Liveness Detection Algorithm Based on Deep Convolutional Neural Network	Improved VGG network named NA-VGG to detect DeepFake face image, which was based on image noise and image augmentation.	https://ieeexplore.ieee.org/document/9189596
12	A Personalized Spatial-Temporal Cold Pain Intensity Estimation Model Based on Facial Expression	Three different architectures (Inception V3, VGG-LSTM, and Convolutional LSTM) were used to estimate three intensities of cold pain: No pain, Moderate pain, and Severe Pain. Architectures with Sequential information were compared with single-frame architecture, showing the importance of spatial temporal information on pain estimation.	https://ieeexplore.ieee.org/document/9553012
13	A Complete Design Flow for Mapping CNN onto Embedded FPGA	Investigate state-of-the-art CNN models and CNN based applications. Requirements on memory, computation and the flexibility of the system are summarized for mapping CNN on embedded FPGAs. Based on these requirements, we propose Angel-Eye, a programmable and flexible CNN accelerator architecture, together with data quantization strategy and	https://ieeexplore.ieee.org/document/7930521

S. no.	Heading	About	Link
		compilation tool.	
14	Fusion of Multi-Intensity Image for Deep Learning-Based Human and Face Detection	Use these multi-intensity illuminated IR videos to evaluate several widely used object detectors, i.e., SSD, YOLO, Faster R-CNN, and Mask R-CNN, by analyzing the effective range of different illumination intensities. By including a tracking scheme, as well as developing a new fusion method for different illumination intensities to improve the performance, the proposed approach may serve as a new benchmark of face and object detection for a wide range of distances.	https://ieeexplore.ieee.org/document/9682685
15	A Real-Time Object Detection Method for Constrained Environments	YOLO is one of the state-of-the-art DNN-based object detection approaches with good performance both on speed and accuracy and Tiny-YOLO-V3 is its latest variant with a small model that can run on embedded devices. In this paper, Tinier-YOLO, which originated from Tiny-YOLO-V3, is proposed to further shrink the model size while achieving improved detection accuracy and real-time performance.	https://ieeexplore.ieee.org/document/8941141
16	A High-Throughput and Power-Efficient FPGA Implementation of YOLO CNN for Object Detection	This paper presents a Tera-OPS streaming hardware accelerator implementing a you-only-look-once (YOLO) CNN. The parameters of the YOLO CNN are retrained and quantized with the PASCAL VOC data set using binary weight and flexible low-bit	https://ieeexplore.ieee.org/document/8678682

S. no.	Heading	About	Link
		activation. The binary weight enables storing the entire network model in block RAMs of a field-programmable gate array (FPGA) to reduce off-chip accesses aggressively and, thereby, achieve significant performance enhancement.	
17	Focusing on Small Target and Occluded Object Detection	Y. Li, S. Li, H. Du, L. Chen, D. Zhang and Y. Li, "YOLO-ACN: Focusing on Small Target and Occluded Object Detection," in IEEE Access, vol. 8, pp. 227288-227303, 2020, doi: 10.1109/ACCESS.2020.3046515.	https://ieeexplore.ieee.org/document/8678682
18	Smart Attendance System Based On Face Recognition Techniques	Structure is based on a modern high-precision face detection algorithm using the YOLO v4 (You Only Look Once). The Yolo v4 superior using a single GPU achieving a high speed in detecting objects compared with other models that require to use many GPUs. Using the Darknet-53 layers for face detection we get an accuracy of 100%.	https://ieeexplore.ieee.org/document/9515760
19	A Web-Based Application for Identifying Objects In Images: Object Recognition Software	H. UCUZAL, A. G. İ. BALIKÇI ÇİÇEK, A. G. A. K. ARSLAN and C. ÇOLAK, "A Web-Based Application for Identifying Objects In Images: Object Recognition Software," 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), 2019, pp. 1-5, doi: 10.1109/ISMSIT.2019.8932735.	
20	A Fast and Accurate Unconstrained Face	Propose a method to address challenges in unconstrained face	https://ieeexplore.ieee.org/document/71306

S. no.	Heading	About	Link
	Detector	detection, such as arbitrary pose variations and occlusions. First, a new image feature called Normalized Pixel Difference (NPD) is proposed. NPD feature is computed as the difference to sum ratio between two pixel values, inspired by the Weber Fraction in experimental psychology. The new feature is scale invariant, bounded, and is able to reconstruct the original image.	26
21	A Fast and Accurate System for Face Detection, Identification, and Verification	Propose a novel face detector, Deep Pyramid Single Shot Face Detector (DPSSD), which is fast and detects faces with large scale variations (especially tiny faces).	https://ieeexplore.ieee.org/document/8680708
22	Face Spoofing Detection Through Visual Codebooks of Spectral Temporal Cubes	Introduce a low-cost and software-based method for detecting spoofing attempts in face recognition systems. Hypothesis is that during acquisition there will be inevitable artifacts left behind in the recaptured biometric samples allowing us to create a discriminative signature of the video generated by the biometric sensor.	https://ieeexplore.ieee.org/document/7185398
23	Refinement Neural Network for High Performance Face Detection	Present a single-shot refinement face detector namely RefineFace to achieve high performance. Specifically, it consists of five modules: Selective Two-step Regression (STR), Selective Two-step Classification (STC), Scale-aware Margin Loss (SML), Feature Supervision Module (FSM) and Receptive Field Enhancement (RFE). To enhance	https://ieeexplore.ieee.org/document/9099607

S. no.	Heading	About	Link
		the regression ability for high location accuracy, STR coarsely adjusts locations and sizes of anchors from high level detection layers to provide better initialization for subsequent regressor.	
24	Face Detection in Real Time Live Video Using Yolo Algorithm Based on Vgg16 Convolutional Neural Network	This paper intends to combine the YOLO (You Only Look Once) algorithm with the VGG16 pre-trained convolutional neural network to propose an improvement for face detection systems. Experimental results show that the proposed method has detected the test image set with over 95 % of average precision. Also, our proposed method considerably increased face detection speed in real-time live video. The experiment of this work was using the Image Processing Toolbox and the Deep Learning Toolbox in MATLAB.	https://ieeexplore.ieee.org/document/9446291
25	Face Expression Classification using Squeeze-Excitation based VGG16 Network	An experiment was carried out to realize an improved image classification neural network based on VGG16. The input features were firstly squeezed to c dimension, then weighted and excited by a softmax layer, finally scaled to the original dimension. Compared to the ordinary VGG network, it could capture the features in the channel dimension, therefore improving the accuracy and the convergence speed. It would be constructive to use other image datasets with photos having different illumination conditions and clarity. Experiments on the	https://ieeexplore.ieee.org/document/9712817

S. no.	Heading	About	Link
		FER2013 dataset validate the effectiveness of our method.	

Overall Architecture

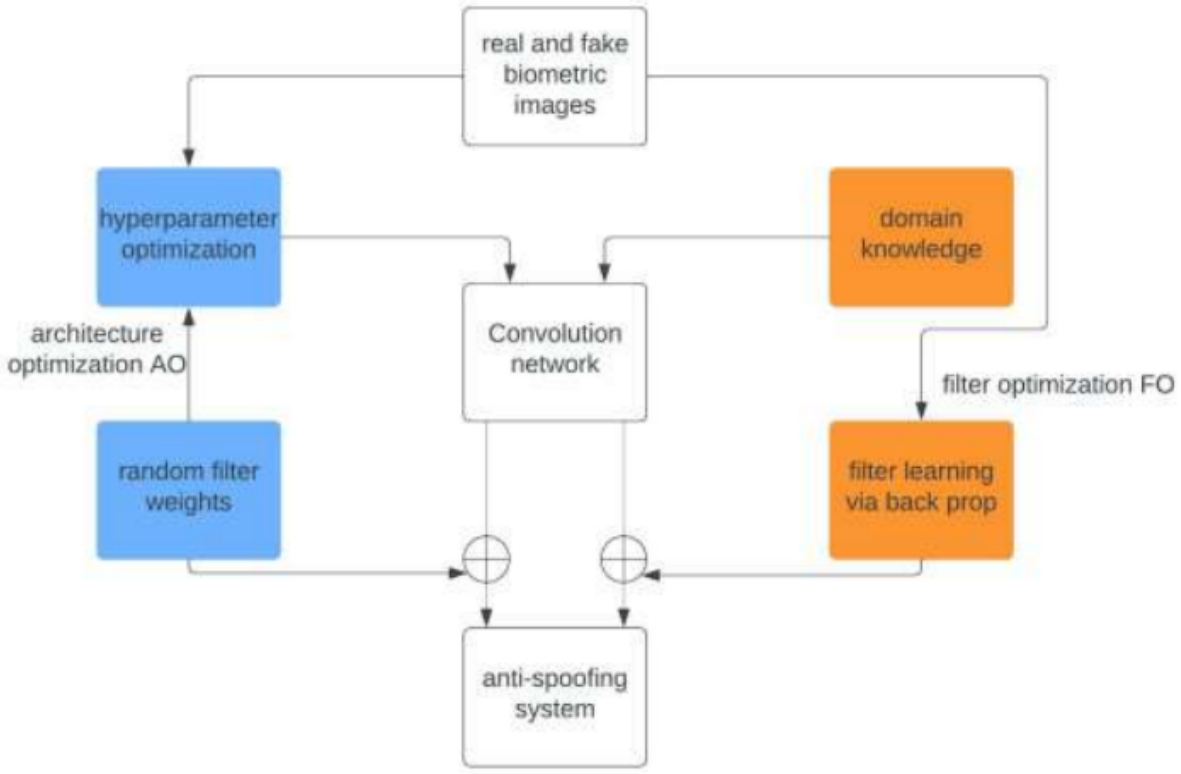


Fig-4 Image Data Flow diagram

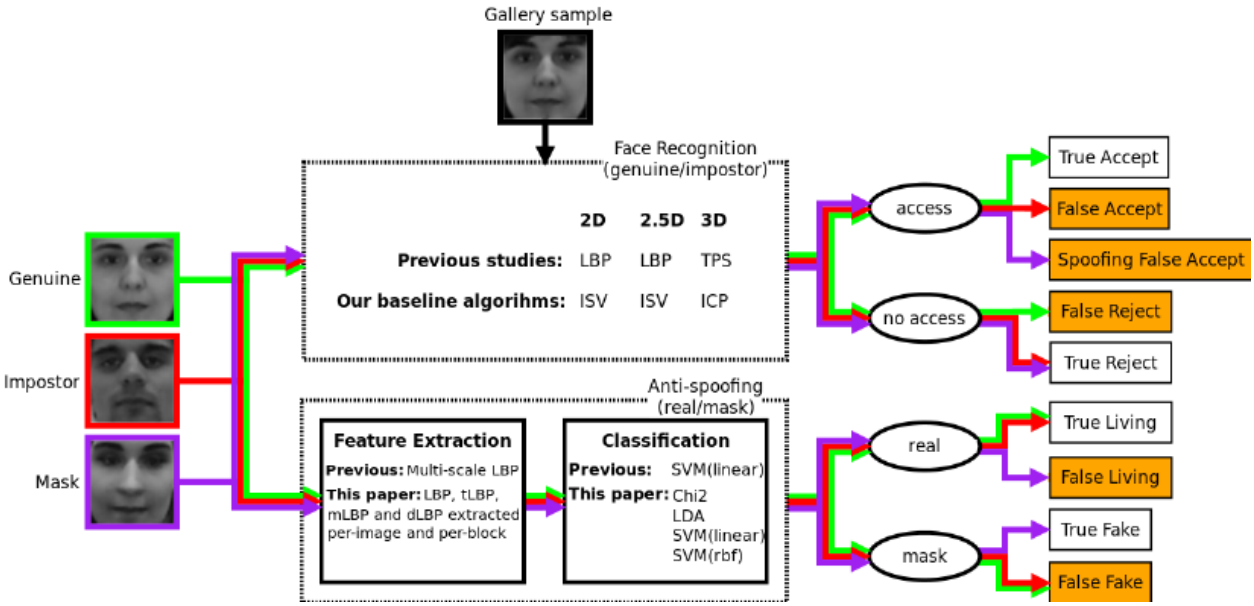


Fig-5 feature extraction mechanism

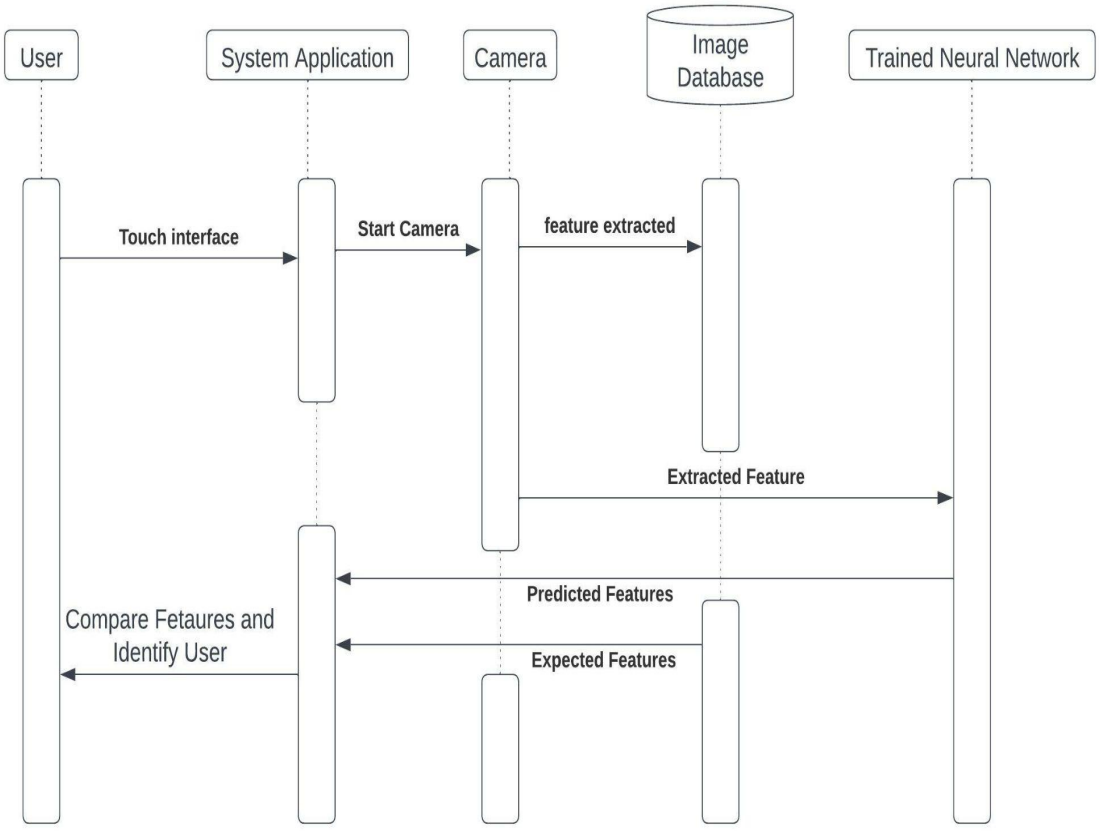


Fig-6 Sequence diagram of System

Proposed NFT Authentication Architecture:

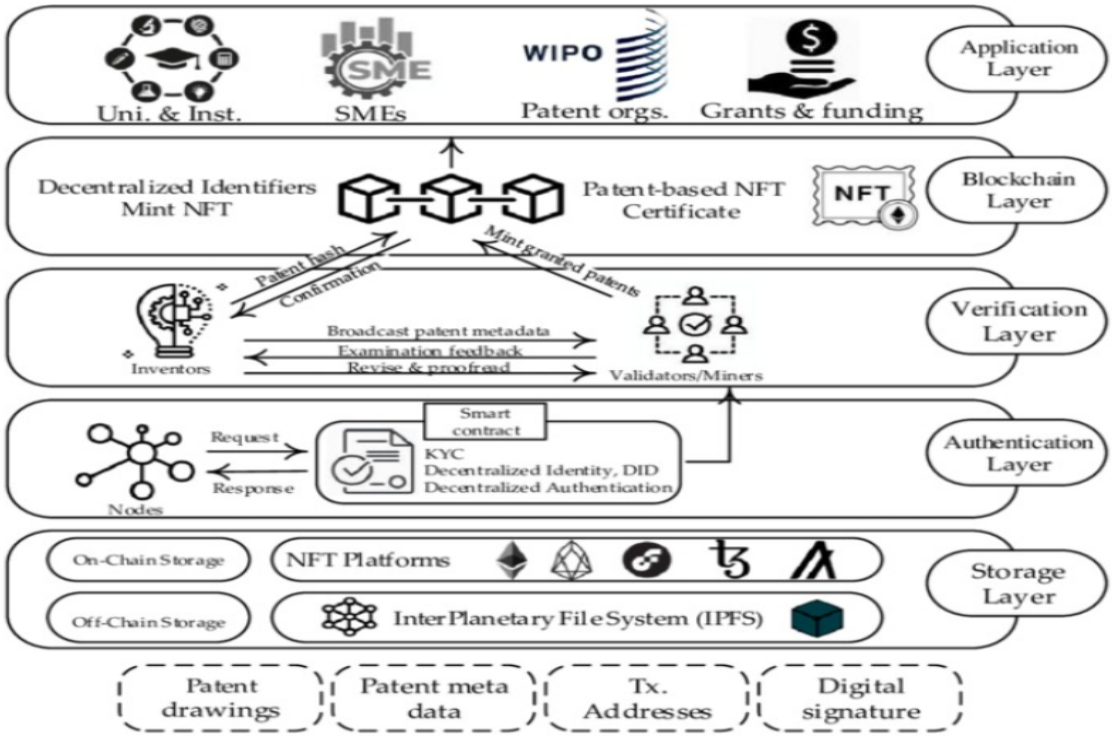


Fig-7[13] NFT Layers

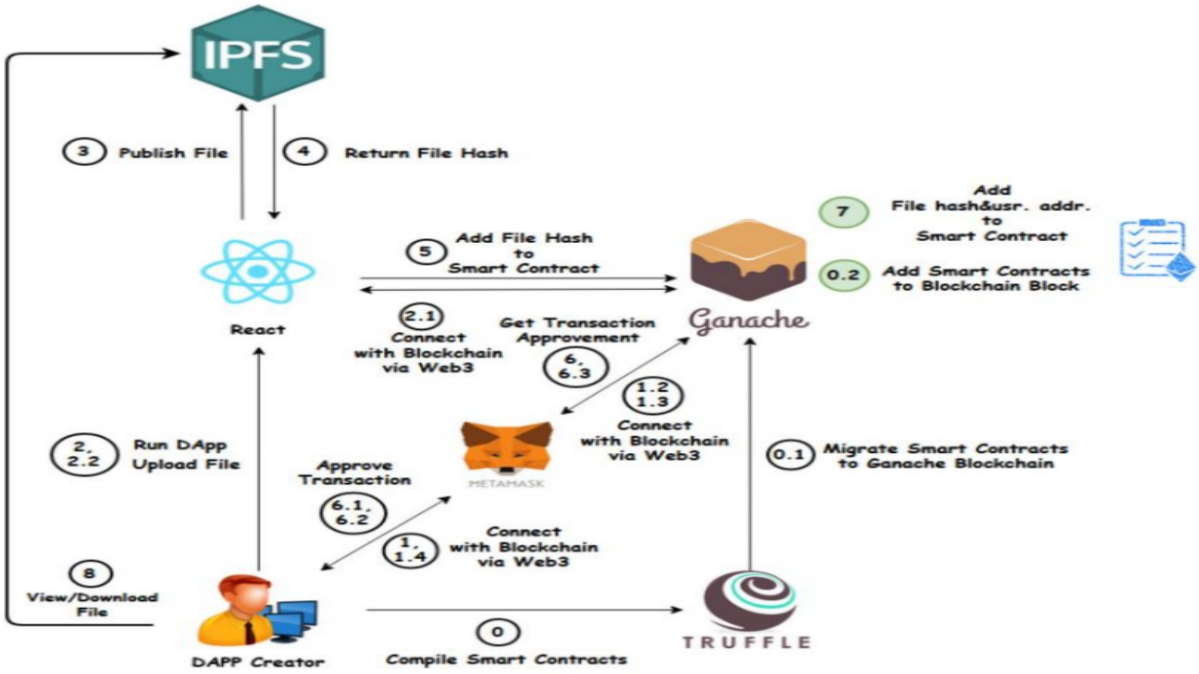


Fig-8 NFT deployments and smart contracts

Proposed Methodology

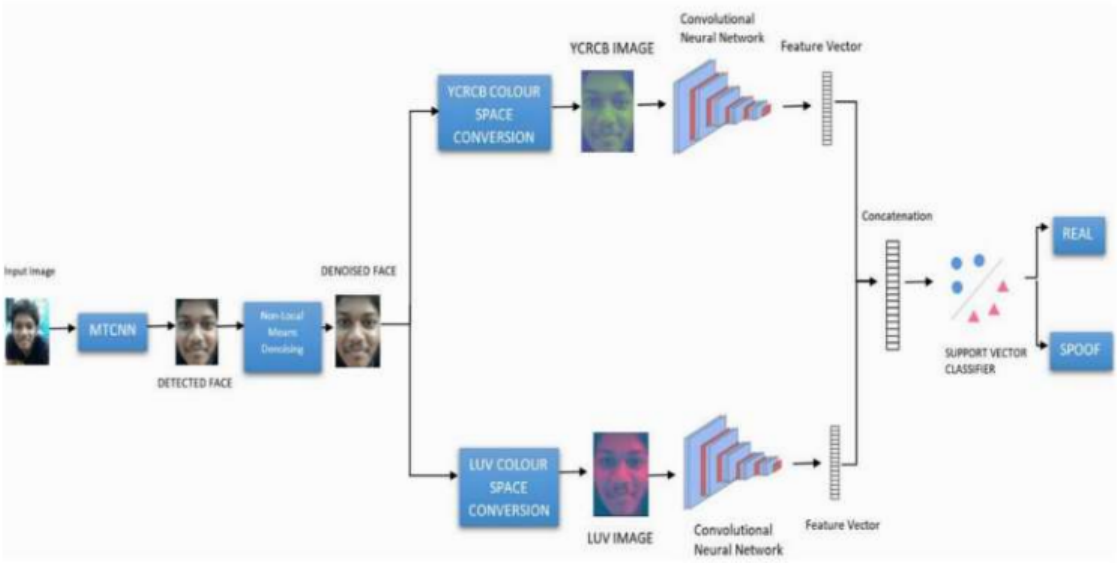
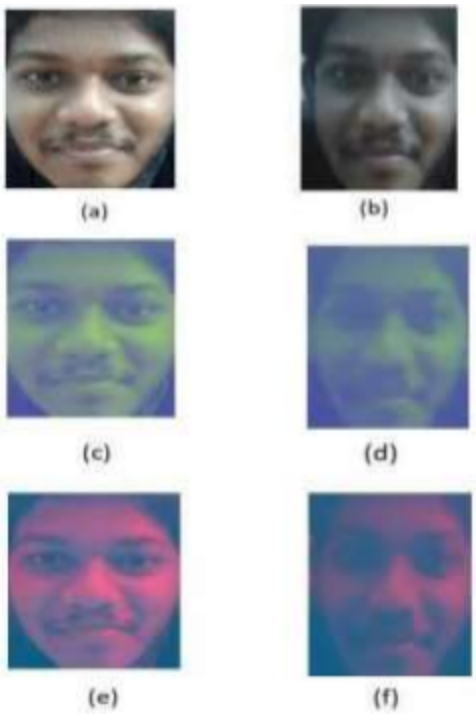


Fig-9 Applying filters for more type featuring

Face detection has been performed using MTCNN (Multi-Task Cascaded Convolutional Neural Network) . MTCNN comprises of three networks corresponding to each step respectively. Initially, an input image is passed through P-net for prediction of possible face positions and their bounding boxes. The respective output consists of large number of false positives. Hence the output is passed through R-net for regression of bounding boxes which eliminate false positive and improves accuracy. Further the output of R-net is passed through Onet for refinement of bounding boxes. The detected face is then denoised using Non-Local means denoising algorithm which replaces the target pixel based on mean value of all pixels in an image and similarity of all pixels with target pixel. Then the denoised face is then converted into YCbCr and CIELUV colour space. In YCbCr, Y represents the luma component of the image which is highly sensitive to human eye, Cb represents the chroma blue component of the image and Cr represents the chroma red components of the image. These chroma components are not very sensitive to human eye. This is the simplest anti spoofing face detection model therefore to increases performance and more accuracy we can add VGG16 Model or YOLO(You Only Live Once).



(a)Real Image (b) Spoof Image (c) Real Image in YCbCr color model (d) Spoof Image in YCbCr color model (e) Real Image in CIELUV color model (f) Spoof Image in CIELUV color model.

VGG-16 Architecture:-

Fig-10 VGG16 Model Summary

VGG16 improves accuracy by doing more type features. Further the output of R-net is passed through Onet for refinement of bounding boxes. The detected face is then denoised using a Non-Local means denoising algorithm which replaces the target pixel based on the mean value of all pixels in an image and the similarity of all pixels with the target pixel. Then the denoised face is then converted into YCbCr and CIELUV color space. In YCbCr, Y represents the luma component of the image which is highly sensitive to the human eye, Cb represents the chroma blue component of the image and Cr represents the chroma red components of the image. These chroma components are not very sensitive to the human eye.

As we can see, the vgg-16 model consists of extensive (DEEP) layers of Convolution coupled with MaxPooling where ever required before going into the traditional set of hidden inter-connected (DENSE) layers The input to the cov1 layer is of fixed size 224 x 224 RGB image. The image is passed through a stack of convolutional (conv.) layers, where the filters were used with a very small receptive field: 3×3 (which is the smallest size to capture the notion of left/right, up/down, center). In one of the configurations, it also utilizes 1×1 convolution filters, which can be seen as a linear transformation of the input channels (followed by non-linearity). The convolution stride is fixed to 1 pixel; the spatial padding of conv. layer input is such that the spatial resolution is preserved after convolution, i.e. the padding is 1- pixel for 3×3 conv. Layers. Spatial pooling is carried out by five max-pooling layers, which follow some of the conv. Layers (not all the conv. layers are followed by max-pooling). Max-pooling is performed over a 2×2 pixel window, with stride 2.

YOLO Architecture:-

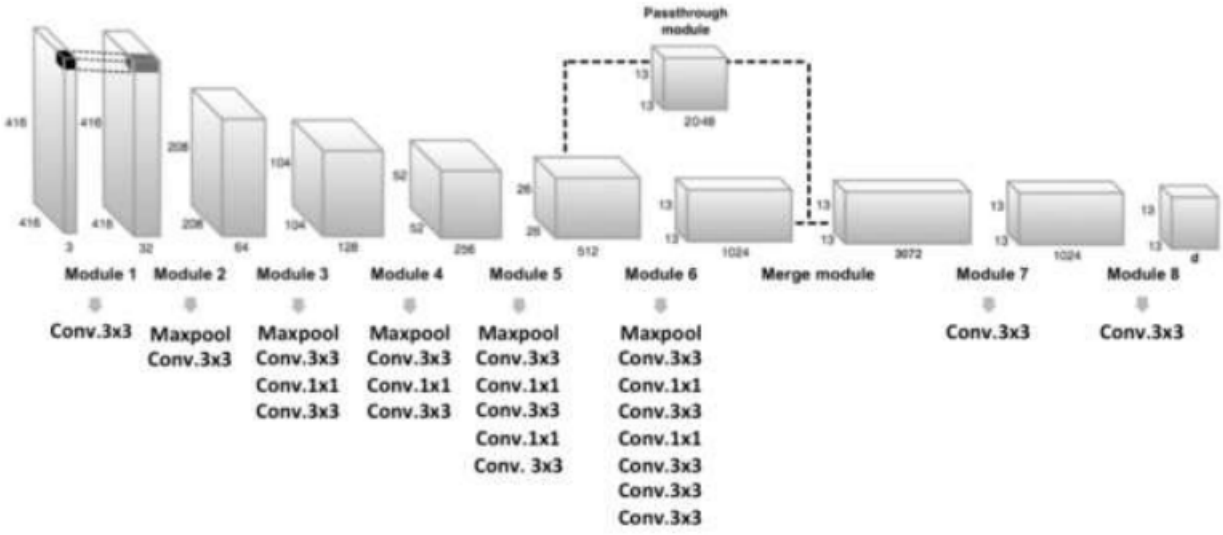


Fig-11 YOLO Architecture Layers

General YOLO architecture takes an image as input and resizes it to 448*448 by keeping the aspect ratio the same and performing padding. This image is then passed on the CNN network. This model has 24 convolution layers, 4 max-pooling layers followed by 2 fully connected layers. For the reduction of the number of layers (Channels), we use 1*1 convolution that is followed by 3*3 convolution. Notice that the last layer of YOLO predicts a cuboidal output. This is done by generating (1, 1470) from the final fully connected layer and reshaping it to size (7, 7, 30). This architecture uses Leaky ReLU as its activation function in the whole architecture except the layer where it uses the linear activation function. The definition of Leaky ReLU can be found here.Batch normalization also helps regularize the model. With batch normalization, we can remove dropout from the model without overfitting it. The simple YOLO has mAP of 63.4%

when trained on VOC 2007 and 2012, the Fast YOLO which is almost 3x faster in result generation has mAP of 52%. This is lower than the best Fast R-CNN model achieved (71% mAP) and also the R-CNN achieved (66% mAP). However, it beats other real-time detectors such as (DPMv5 33% mAP) on accuracy.

Benefits of YOLO:
Process frames at the rate of 45 fps (larger network) to 150 fps(smaller network) which is better than real-time. The network is able to generalize the image better.

Disadvantages of YOLO:
Comparatively low recall and more localization error compared to Faster R_CNN. Struggles to detect close objects because each grid can propose only 2 bounding boxes. Struggles to detect small objects.

Advantages of using NFTs for authentication:-

For accessing digital accounts, having a unique and secure certificate of authenticity is of enormous value, as the ability to forge documents and information on the internet is very easy. Thus, the use of NFT provides security to the buyer.

Also, being a non-fungible token, this technology contains unique information, which makes it different from any other NFT, with a very simple verification facility. All this contributes to rendering the creation of counterfeit assets and their subsequent circulation in the market useless.

Each NFT can be traced back to its origin or creator, as the code of an NFT carries the form of its creator, giving the possibility to authenticate the token on any browser or platform. It is a decentralized verification method that does not require any entity to host the NFT.

A final beneficial aspect is that NFT is not dependent on third party platforms as it is based on Blockchain. However, for now, there is a dependency of this technology on the marketplaces where NFTs are created and acquired.

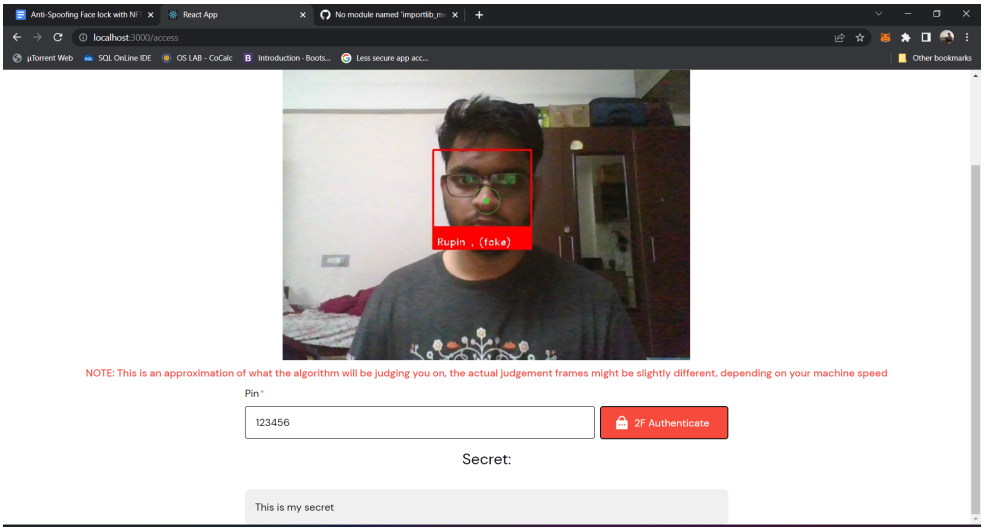
Results:-

After training all the models we tested them and compare them in terms of accuracy and performance. We also compare the feasibility of this models, in real time applications.

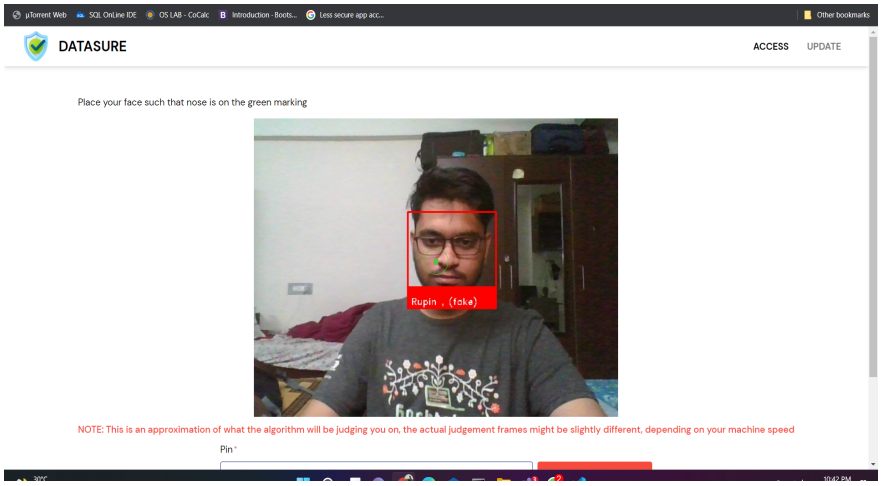
Meso_net model:

Mesonet's method of collecting deepfake data,they acquired deep fake image data from popular deep fake video platforms that are online. According to the September 2019 report called the state of deepfakes conducted by deep trace labs,96 percent of deepfake media is pornographic. Deepfake pornography is among the worst abuses in the world of deepfakes and besides being a significant social issue this also complicates the technical side of our deepfake classifier. This is why Mesonet was trained on real data acquired from TV and movie sources that offer a great variety of facial expressions and settings. However, the deepfake data acquired was from popular deepfake platforms on the internet sources that are overwhelming dominated by pornographic content therefore it's expected that the model took advantage of a data artifact the statistical reality that deepfakes tend to be pornographic and reals tend to be non-pornographic and used it as a heuristic for its predictions. The model makes predictions under real conditions of diffusion on the internet which explains their use of popular deepfake platforms including pornographic websites. One wonders then whether we can neutralize this effect and force the model to recognize deep fakes without the aid of statistical accidents by using some different data. Since deepfake data is limited in supply and overwhelmingly pornographic. We could acquire our real data from pornographic sites as well rather than just TV shows and movies and unlike deepfake data pornographic is relatively easy to find on the internet. We examined a model designed to identify deepfake images called Mesonet and we implemented it to explore how it works in doing so.

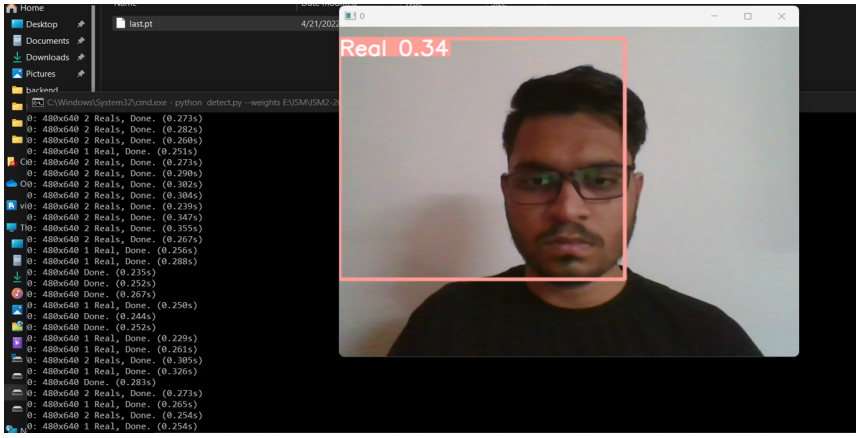
Mesonet:



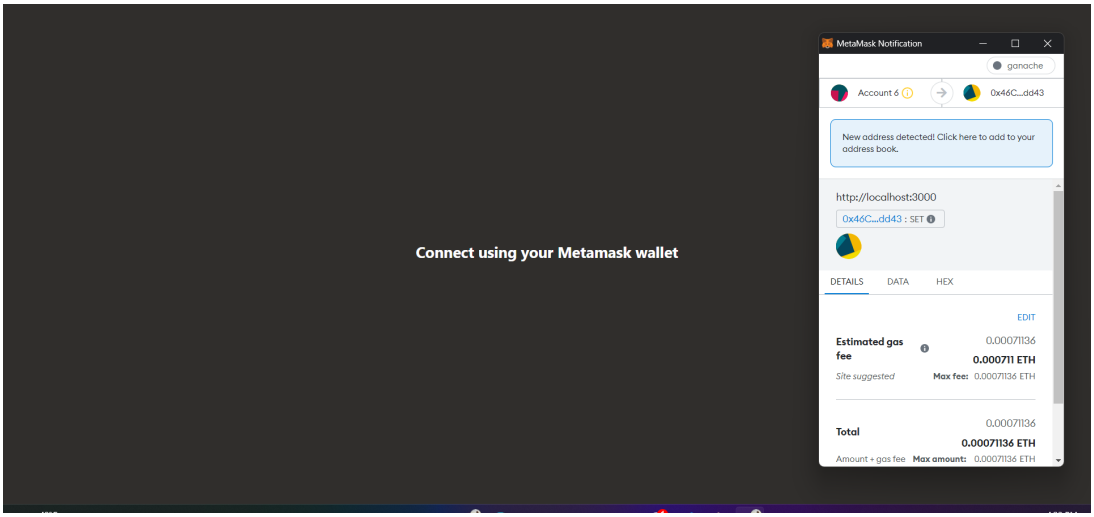
VGG16:-



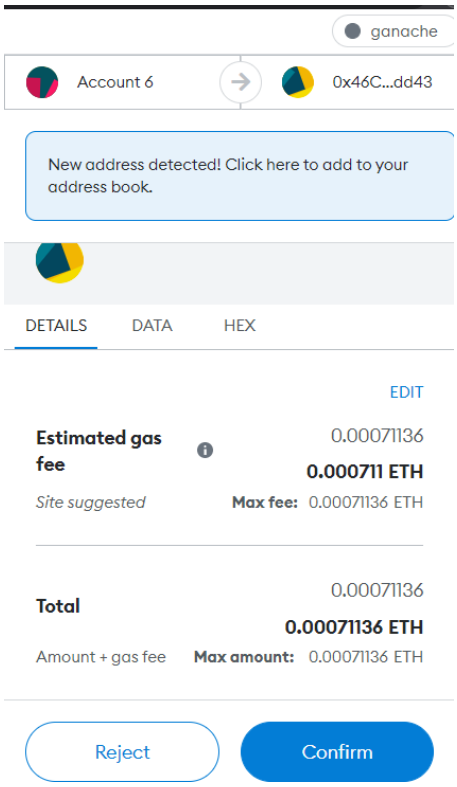
YOLOv5:-



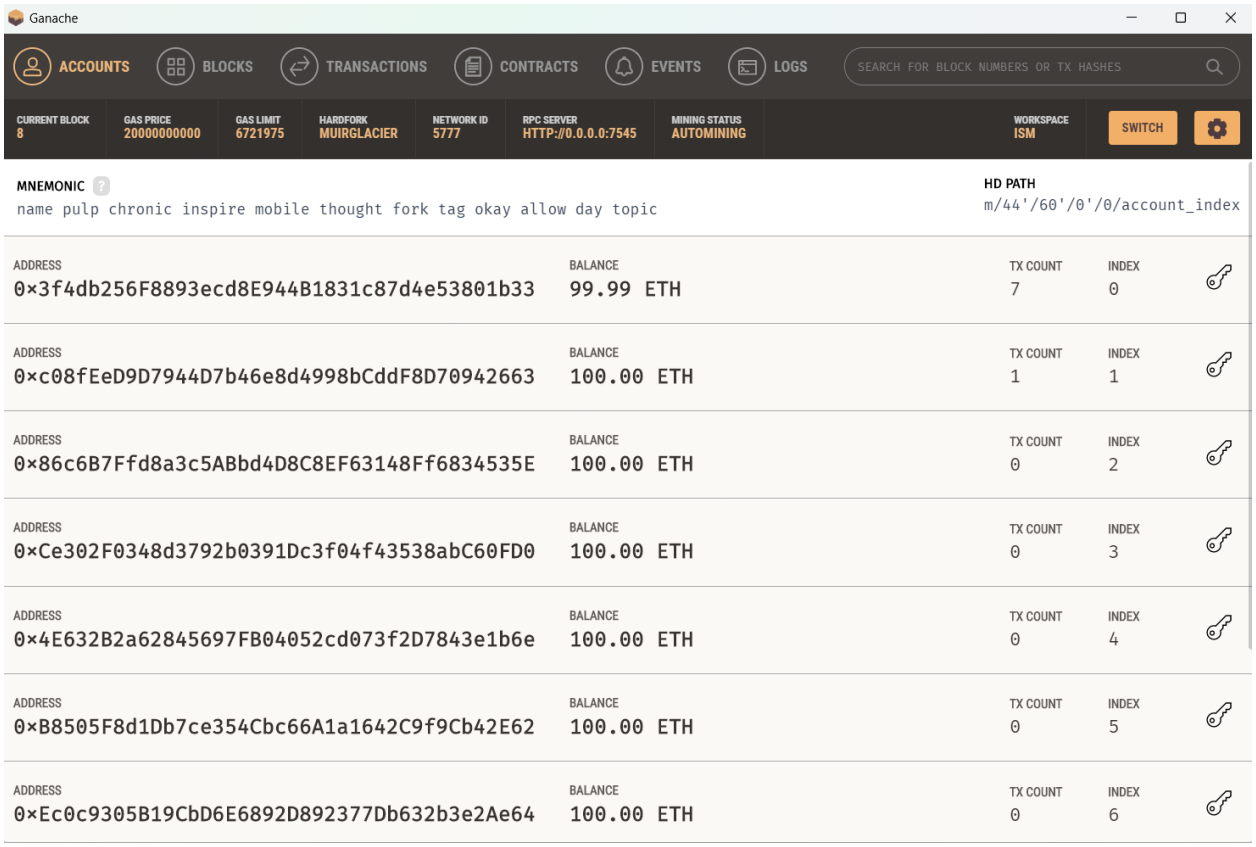
NFT:-



Wallet for transaction and fee receipt with some gas value as (tax):-



Ganache for accessing private keys and seeing account statements:-



Analysis:-

We Compare Mesonet with VGG16 and found that Mesonet is one of the balanced model we found, it was giving accuracy at better framerates, which means it was still efficient in terms of speed but if we talk about VGG16, we achieved a good accuracy and since VGG16 is a deep convolution model there is more type featuring due to which the performance of detecting and recognizing decrease but still it gives higher accuracy ,but the framerate is not good.

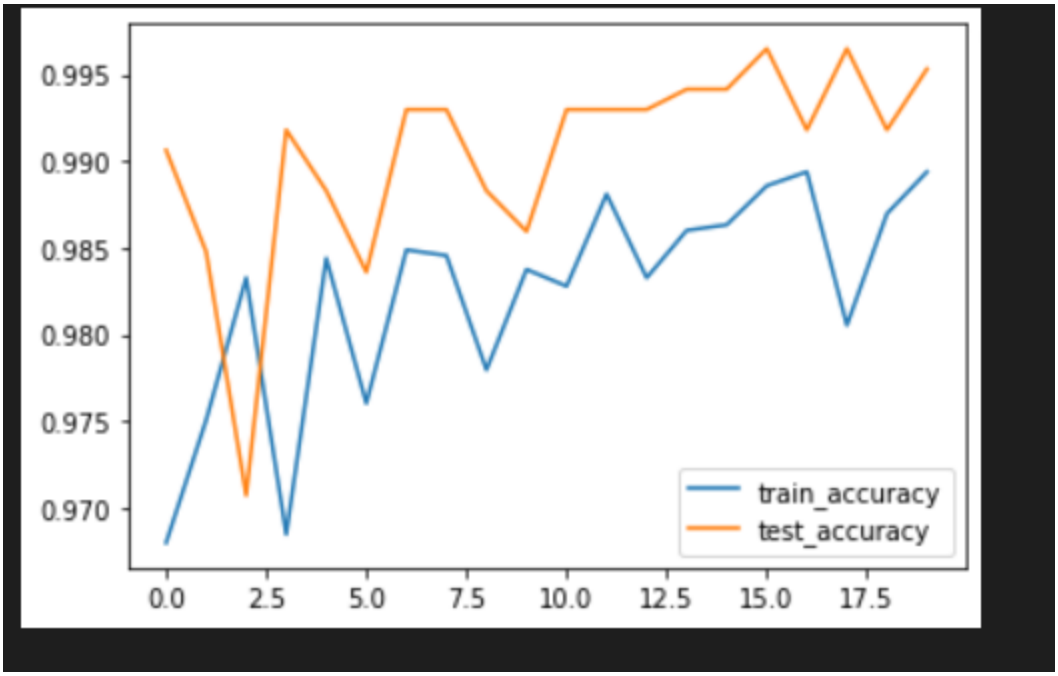


Fig-10 Training vs Validation

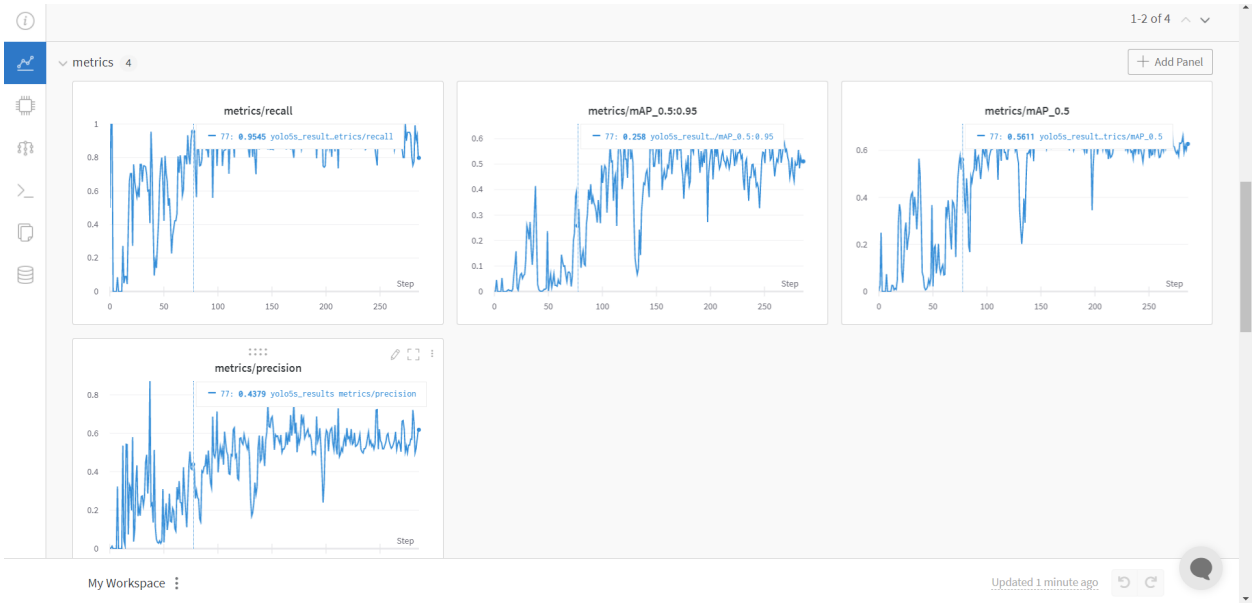
Here X-axis is Epochs and Y-axis is accuracy

```
    train_loss,train_acc=model.evaluate(x_train)

196/196 [==============================] - 68s 342ms/step - loss: 0.0095 - accuracy: 0.9978

    test_loss,test_acc=model.evaluate(x_test)

27/27 [==============================] - 8s 288ms/step - loss: 0.0111 - accuracy: 0.9953
```

Fig-11 Maximum training and validation accuracy

Since there is less difference in training accuracy and validation accuracy and overall it is above 95% we can say that model is relevant in terms of accuracy.

At last we compare VGG16 with YOLOv5.



This graphs are used to study precision , recall and mAP as Y-axis against number of epochs as X-axis

Fig-12 This graphs are used to study training loss , validation loss as Y-axis against number of epochs as X-axis

YOLO gives better performance in terms of speed but still it lacks behind in precision while recognizing.But YOLO was still able to detect face at higher frame rate but with less accuracy. Comparatively low recall and more localization error compared to Faster R_CNN. Struggles to detect close objects because each grid can propose only 2 bounding boxes. Struggles to detect small objects.

Best model to give precision at higher frame rate was meso_net ,in terms of accuracy VGG16 was one which gave maximum accuracy.

Conclusion and Future Work:-

Deepfakes is a synthetic face generation technology that uses a GAN. Due to the continuous advancement of video processing technology and the improvement in quality of videos, deepfake detection becomes more difficult. The current methods for detecting deep forgery can identify signals such as borders, shadows and uniform artifacts or double eyebrows, but the technology that produces deep forgeries is developing rapidly, so we must devote ourselves to building more technologies and detection tools. One of the biggest limitations when a video is tampered with is the lack of real data set that can be used to test new detection technologies, so Facebook is using paid actors to commission the first data set of this type for AI community use. It is part of the DFDC, that was also established in cooperation with AWS, Microsoft, the AI Media Integrity Association Committee and academia. The mission of DFDC, to encourage researchers all around the world to develop fresh cutting-edge technologies,which will enable them to spot counterfeit tampered media. We hope that by helping the AI community come together, we can foresee the challenges of this emerging technology. The research community is committed to developing the deepfake detection algorithm to solve the disturbing deepfake problem, and has published a number of research results. This report implements Mesonet using Meso 4 to make predictions on image data. We will be using the Meso 4 model Trained on the deepfake Data set. We will examine four sets of images-correctly classified deepfakes, correctly classified reals, misclassified deepfakes, misclassified reals.

Obviously, a war is brewing between the people who use advanced machine learning to generate advanced deep forgeries and those who try to detect them. Because authenticity provides a secure environment, deepfake technology must be integrated by applying a real layer across the Internet to provide trust metrics in social media and other networks. The detection of distorted vision content has become a hot topic in the scientific community. We will continue to study successful defense strategies. We will explore the new network model to more accurately discover the content of Deepfake , which may continue to be part of our project.

NFTs have provided an extra layer of security in authentication. Each new user on signing up mints an unique NFT for their wallet. On login, the system verifies if the wallet address of the user has the NFT that is linked to the wallet using smart contracts. Only the users who have NFTs minted can proceed further, otherwise the access is denied.

References:-

[1]-https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.researchgate.net%2Ffigure%2FProposed-Architecture-for-Presenting-NFT-Based-Patent_fig1_355101127&psig=AOvVaw2sWLxQjX8kan-8_Pm8XpbX&ust=1650625196550000&source=images&cd=vfe&ved=0CAwQjRxqFwoTCIipzZ6ApfcCFQAAAAAdAAAAABAe

[2] Brian Dolhansky, Russ Howes, Ben Pflaum, Nicole Baram, and Cristian Canton Ferrer. The deepfake detec- tion challenge (DFDC)preview dataset.

[3] "facebook deepfake detection dataset," https://deepfakedetectionchallenge.ai/

[4] X. Zhao and Y. -W. Si, "NFTCert: NFT-Based Certificates With Online Payment Gateway," 2021 IEEE International Conference on Blockchain (Blockchain), 2021, pp. 538-543, doi: 10.1109/Blockchain53845.2021.00081. : https://ieeexplore.ieee.org/abstract/document/9680582

[5] I. Chingovska, A. Anjos and S. Marcel, "On the effectiveness of local binary patterns in face anti-spoofing," 2012 BIOSIG - Proceedings of the International Conference of Biometrics Special Interest Group (BIOSIG), 2012, pp. 1-7. : https://ieeexplore.ieee.org/abstract/document/6313548

[6] Liveness Detection for Embedded Face Recognition System : https://www.researchgate.net/profile/Sung-Jung-23/publication/242522706_Liveness_Detection_for_Embedded_Face_Recognition_System/links/00b7d534b86e17aadc000000/Liveness-Detection-for-Embedded-Face-Recognition-System.pdf

[7] Understanding Security Issues in the NFT Ecosystem : https://arxiv.org/abs/2111.08893

[8] Z. Boulkenafet, J. Komulainen and A. Hadid, "Face anti-spoofing based on color texture analysis," 2015 IEEE International Conference on Image Processing (ICIP), 2015, pp. 2636-2640, doi: 10.1109/ICIP.2015.7351280. : https://ieeexplore.ieee.org/abstract/document/7351280

[9] S. A. Nazeer, N. Omar and M. Khalid, "Face Recognition System using Artificial Neural Networks Approach," 2007 International Conference on Signal Processing, Communications and Networking, 2007, pp. 420-425, doi: 10.1109/ICSCN.2007.350774. : https://ieeexplore.ieee.org/abstract/document/4156656

[10] Face and feature finding for a face recognition system : http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.15.2445

[11] B. J. Boom, G. M. Beumer, L. J. Spreeuwers and R. N. J. Veldhuis, "The Effect of Image Resolution on the Performance of a Face Recognition System," 2006 9th International Conference on Control,

Automation, Robotics and Vision, 2006, pp. 1-6, doi: 10.1109/ICARCV.2006.345480. : https://ieeexplore.ieee.org/abstract/document/4150409

Appendix:-

The work was divided equally among all the three members.
Rupin Patel:- Execution of Artificial Intelligences model and create backend.
Aditya Singh:- Creation of Frontend react app, database storage and documentation.
Ishan Bhardwaj:- Creation of NFTs, Smart Contracts and integration with Blockchain