CHAPTER 1 INTRODUCTION

Fake images have become a central problem in the last few years, especially after the advent of the so-called deep fakes, i.e., fake images manipulated with the help of powerful and easy to-use deep learning tools, such as auto encoders (AE) or generative adversarial networks (GAN). With this technology, creating realistic manipulated media assets may be very easy, provided one can access large amounts of data. Applications include photography, video-games, virtual reality, and may soon expand to movie productions. The same technology, can be used for malicious purposes, like creating fake adult images to blackmail people, or building fake-news campaigns to manipulate the public opinion. In the long run, it may also reduce trust in journalism, including serious and reliable sources. Figure shows some popular deep fakes circulating on the internet. These fakes are easy to spot since they were generated for fun and involve well-known actors and politicians in unlikely situations. In addition, on the web it is usually possible to retrieve both the original and the manipulated version, removing any doubt about authenticity.

FIGURE 1.1 REAL IMAGES AND THEIR RESPECTIVE EDITED IMAGES

CHAPTER 2 PROBLEM DEFINITION

This aim of the project is to design analysis methods for visual media integrity verification i.e detection of images which are manipulated. The main emphasis of the project is to identify emerging phenomenon of deep fakes, fake media created .The deep learning tools and modern data-driven forensics methods adopted to detect the fake images and media and protect from security threats .

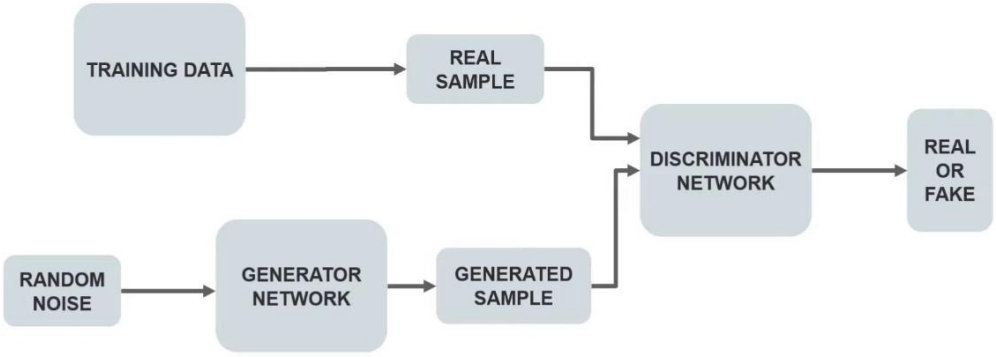
CHAPTER 3 LITERATURE SERVEY

To carry out the project work ,the literature survey is performed to identify the advantages and loopholes in the earlier method designed and implemented by various researchers which is listed below.

Paper title	**Authors & year**	**Inference**
Learning to Detect Fake Face Images in the Wild.	Chih-Chung Hsu, ChiaYen Lee and Yi-Xiu Zhuang. **2019**	This Paper gives the concept of discriminators and it's bond with GAN.
Fake Colorized Image Detection.	Yuanfang Guo Xiaochun Cao and Wei Zhang. **2018**	This Paper holds concept of contrast disturbance.
Image Montage for constructing Photorealistic-virtual world from different realistic Images.	Mizuki Tachibana and Tadahiro Fujimoto. **2018**	This Paper has the insights for random real time scenes and how to montage them to detect.
Deep Learning for Deepfakes Creation and Detection: A Survey	Thanh Thi Nguyen, Cuong M. Nguyen. Tien Dung Nguyen **2019**	This Paper had the insights of deep fake image creation and how GAN are involved in this.

1. Learning to Detect Fake Face Images in the Wild

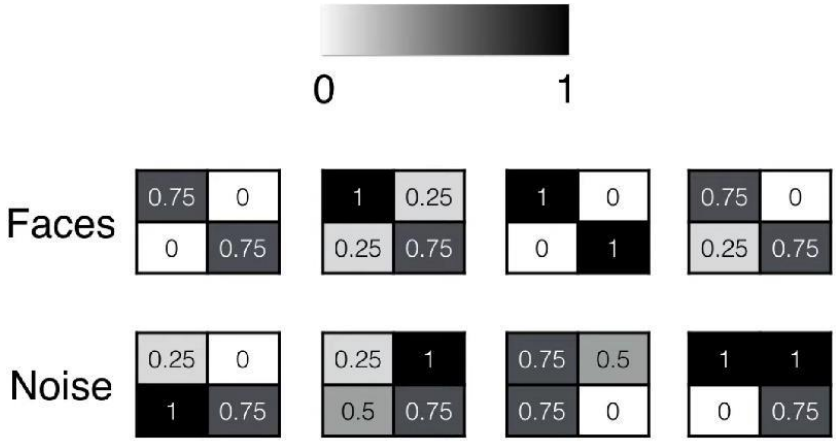
This paper addresses the issue encountered for image forgery detection, there are two different categories in the traditional approach: 1) extrinsic feature and 2) intrinsic feature. First forgery detection will embed external unique signals into the original images (e.g. digital watermarking). Then, the received image can be verified to determine that it is a forgery by comparing extracted watermark and original watermark. The second strategy discover the intrinsic and the invariant features of original images . The forgery image are detected by checking the statistical properties of the extracted intrinsic feature from the received images. Due to the tampered operations performed on the images will change the intrinsic feature. Contrastive loss learns such joint features from heterogeneous training images by introducing the pairwise information so that the DeepFD should be able to effectively distinguish any fake image generated by any GAN. Besides, the proposed DeepFD can further localize unrealistic details of the fake image based on a fully convolutional architecture. The proposed DeepFD discussed in this paper was able to localize unrealistic details of the fake image and follow the localized regions to improve the DeepFD .



CONCEPT OF DEEP FORGERY DISCRIMINATOR

2. Fake Colorized Image Detection

An emerging image editing technique is colorization, in which grayscale images are colorized with realistic colors. Unfortunately, this technique may also be intentionally applied to certain images to confound object recognition algorithms. Before this paper was published, there was no forensic technique invented using color images. According to observation, color images, which are generated from state-of-the-art methods, possess statistical differences for the hue and saturation channels as compared to natural images. Besides, we also observe statistical inconsistencies in the dark and bright channels, because the colorization process will inevitably affect the dark and bright channel values. Based on our observations, i.e., potential traces in the hue, saturation, dark, and bright channels, we propose two simple yet effective detection methods for fake colorized images: Histogram-based fake colorized image detection and feature encoding-based fake colorized image detection. Experimental results demonstrate that both proposed methods exhibit a decent performance.



3. Image Montage for Construction of Photorealistic Virtual World from Different Real Scene Images

An image montage technique creates a photorealistic output image without artifacts by synthesizing multiple input images. This paper propose an image montage method to create a photorealistic virtual scene that seems as if it exists in a real world by synthesizing real scene images captured in different places. The images to synthesize are appropriately selected from a large database of different scene images. Images in the database were captured in the same place by moving a camera to translate its 3D position and rotate its viewing direction. The method consists of obtaining similar image pairs such that two images in each pair have similar regions, as if they are same scene part that shifts its position in respective images by camera motion. Then, from the pairs, appropriate successive images are selected and arranged in a 3D virtual space such that every image is similar to both adjacent images in their overlapping regions to fake a camera motion of a walk-through in the space, which results in constructing a photorealistic virtual world. The similarity in the overlapping region is evaluated using gist and color histogram. The selection of the successive images is done by solving an optimization problem using a graph of images to search for appropriate closed paths, each of which satisfies a restriction of a 3D space. A final virtual scene image is synthesized by blending the overlapping regions of adjacent images in a desired path.

4. Deep Learning for Deepfakes Creation and Detection: A Survey

This paper analyzes that the underlying mechanism for deep fake creation is deep learning models such as autoencoders and generative adversarial networks, which have been applied widely in the computer vision domain. These models are used to examine facial expressions and movements of a person and synthesize facial images of another person making analogous expressions and movements. Deepfake methods normally require a large amount of image and video data to train models to create photo-realistic images and videos. A deepfake detection method using convolutional neural network (CNN) and long short-term memory (LSTM) to extract temporal features of a given video sequence, which are represented via the sequence descriptor. The detection network consisting of fully-connected layers is employed to take the sequence descriptor as input and calculate probabilities of the frame sequence belonging to either authentic or deepfake class. Another research direction is to integrate detection methods into distribution platforms such as social media to increase its effectiveness in dealing with the widespread impact of deepfakes. The screening or filtering mechanism using effective detection methods can be implemented on these platforms to ease deepfakes detection.

	image_id	lefteye_x	lefteye_y	righteye_x	righteye_y	nose_x	nose_y	leftmouth_x	leftmouth_y	rightmouth_x	rightmouth_y
0	000001.jpg	69	109	106	113	77	142	73	152	108	154
1	000002.jpg	69	110	107	112	81	135	70	151	108	153
2	000003.jpg	76	112	104	106	108	128	74	156	98	158
3	000004.jpg	72	113	108	108	101	138	71	155	101	151
4	000005.jpg	66	114	112	112	86	119	71	147	104	150

CHAPTER 4 PROJECT DESCRIPTION

4.1 PROJECT DESIGN

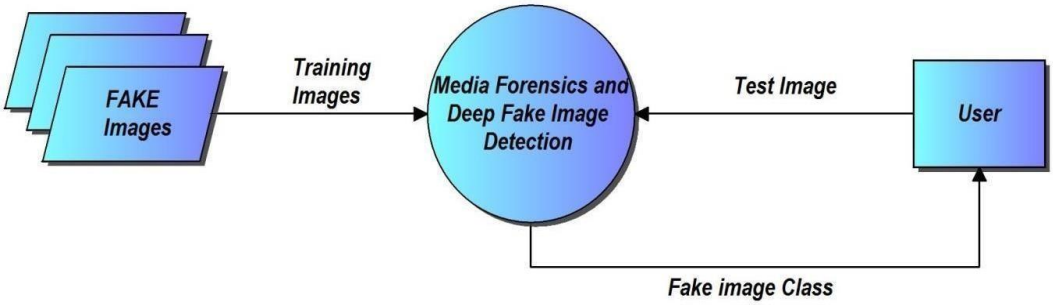


Figure 4.1(a) Level 0 Block Diagram

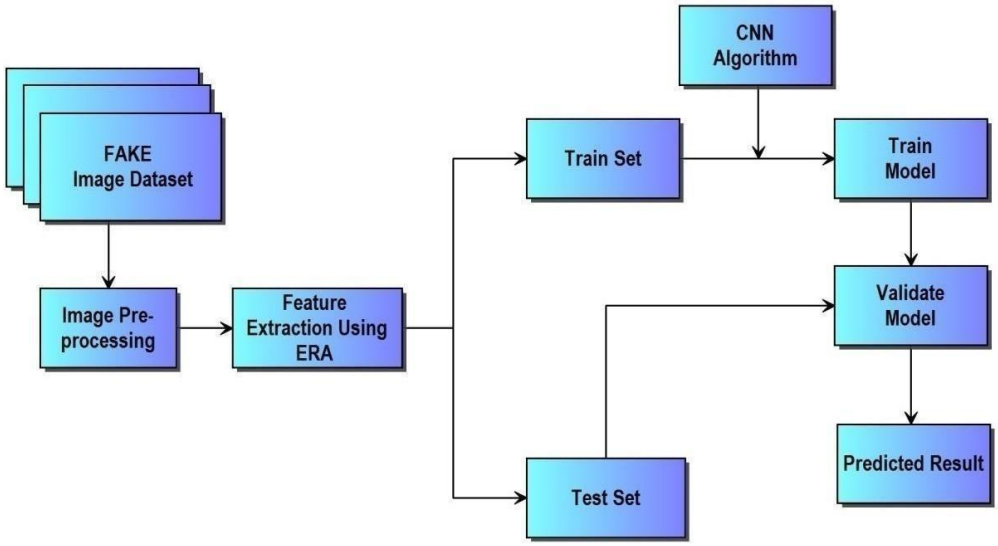


Figure 4.1(b) Level 1 Block Diagram

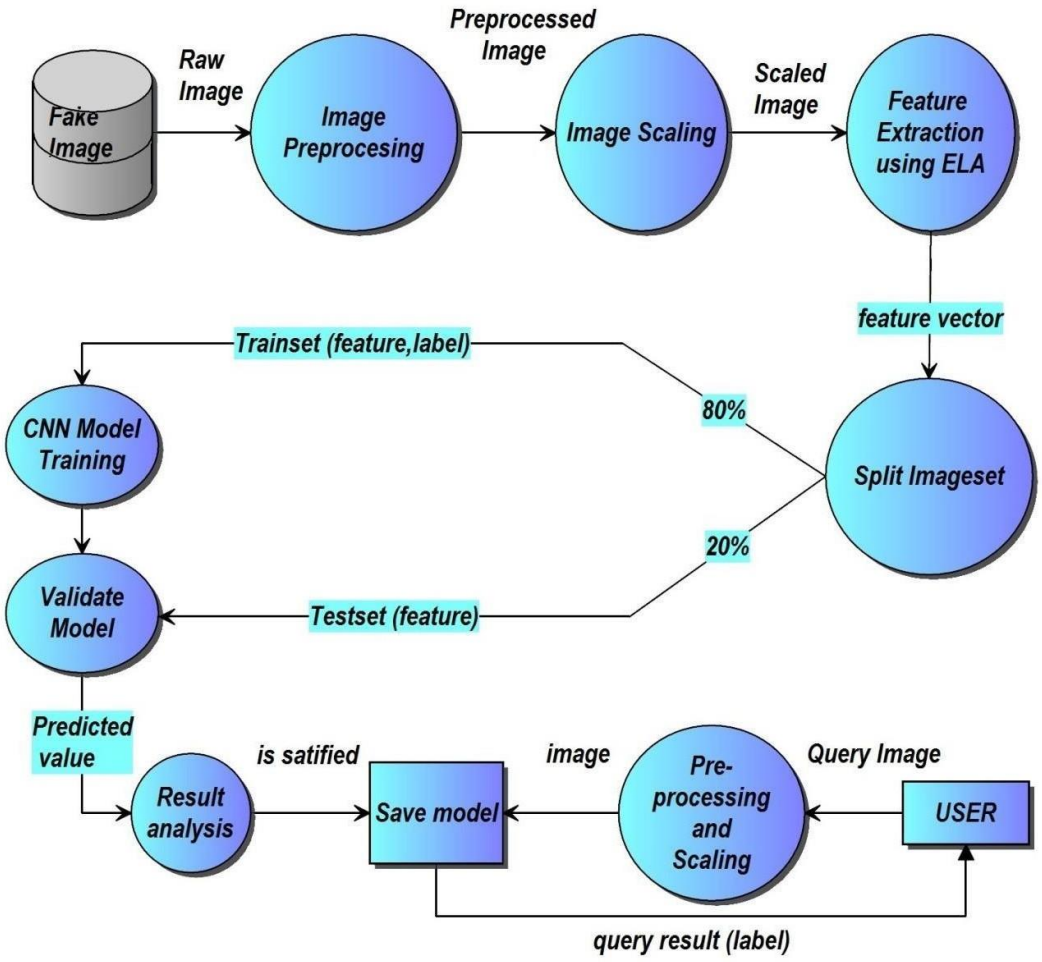


Figure 4.1(c) Level 2 Block Diagram

CHAPTER 5 REQUIREMENTS

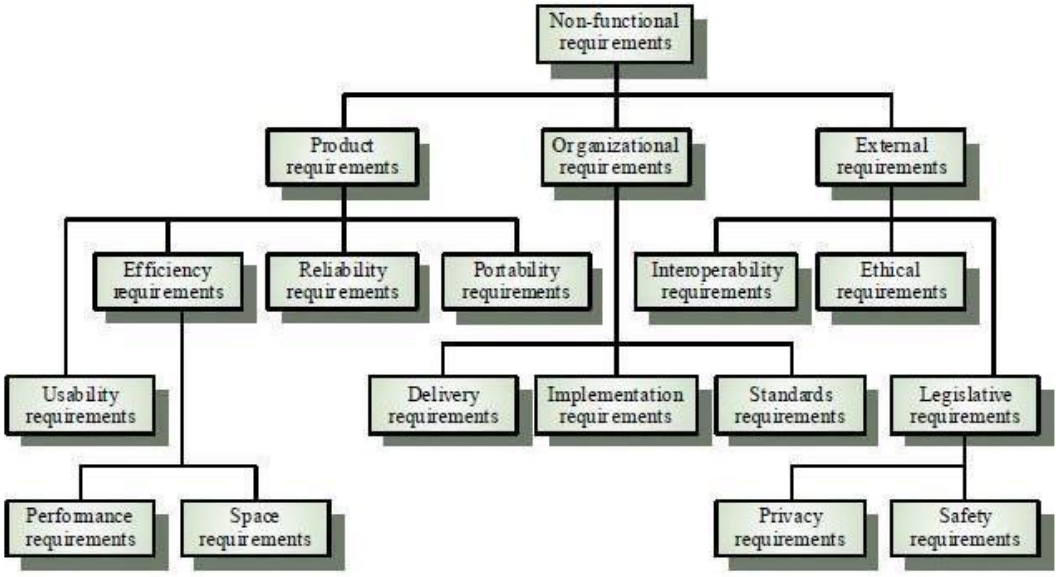
5.1 FUNCTIONAL REQUIREMENTS

This system must be developed using Python programming version 3.0 and above. Integrated development unit must be Jupyter Notebook which has to run on the Anaconda Framework. The functional requirements of the project are one of the most important aspects, the main functional requirement of the system is given below.

- This system has to able to classify the Real and Fake images using CNN.
- There should be two notebooks, one notebook for Training the Dataset and another for Testing the data that means to predict the sample input classification.
- The Training module should able to read the training dataset from the system directory and Using ELA we are convert all images then convert the images into numerical data which are stored in Numpy arrays.
- Once datasets were read 80% of the data has to allotted for training process and remaining 20% of data is allotted for testing process.
- Once training dataset were prepared then it has to enter into CNN process.
- In CNN following layers has to be executed one by one, Convolution Layer, Rectified Linear Unit and Max Polling.
- The above step has to be repeated until training dataset accuracy reached.
- Once training dataset accuracy reached then fully connected layer process has to start.
- Once training process is completed, testing process has to be performed and confusion matrix has to be printed.
- This system has to achieve minimum 70% accuracy

5.2 NON-FUNCTIONAL REQUIREMENT

Non-functional requirements describe how a system must behave and establish constraints of its functionality. This type of requirements is also known as the system's *quality attributes*. Attributes such as performance, security, usability, compatibility are not the feature of the system, they are a required characteristic. They are "developing" properties that emerge from the whole arrangement and hence we can't compose a particular line of code to execute them. Any attributes required by the customer are described by the specification. We must include only those requirements that are appropriate for our project.



CHAPTER 6 METHODOLOGY

Data collection

The Machine learning needs, lots of data to be collected and then design a model. When acquiring the data, be sure to have enough features (aspect of data that can help for a prediction, like the surface of the house to predict its price) populated to train correctly in learning model. In general, lot of data is needed.

The primary data collected from the online sources remains in the raw form of statements, digits and qualitative terms. The raw data contains error, omissions and inconsistencies. It requires corrections after careful scrutinizing the completed questionnaires. The following steps are involved in the processing of primary data. A huge volume of raw data collected through field survey needs to be grouped for similar details of individual responses.

Data Pre-Processing

Data Pre-processing is a technique that is used to convert the raw data into a clean data set. In other words, whenever the data is gathered from different sources it is collected in raw format which is not feasible for the analysis. Therefore, certain steps are executed to convert the data into a small clean data set. This technique is performed before the execution of Iterative Analysis. The set of steps is known as Data Pre-processing. It includes -

- Data Cleaning
- Data Integration

- Data Transformation
- Data Reduction

Data Pre-processing is necessary because of the presence of unformatted real-world data. Mostly real-world data is composed of -

- **Inaccurate data (missing data) -** There are many reasons for missing data such as data is not continuously collected, a mistake in data entry, technical problems with biometrics and much more.
- **The presence of noisy data (erroneous data and outliers) -** The reasons for the existence of noisy data could be a technological problem of gadget that gathers data, a human mistake during data entry and much more.
- **Inconsistent data -** The presence of inconsistencies are due to the reasons such that existence of duplication within data, human data entry, containing mistakes in codes or names, i.e., violation of data constraints and much more.

Error Level Analysis:

Error Level Analysis is a forensic method to identify portions of an image with a different level of compression. The technique could be used to determine if a picture has been digitally modified. They result in poor quality compressed images. Error Level Analysis (ELA) permits identifying areas within an image that are at different compression levels. With JPEG images, the entire picture should be roughly at the same level. If a section of the image is at a significantly different error level, then it is likely to indicate a digital modification.

Data Preparation and Model Construction

Many a times, people first split their dataset into 2 —Train and Test. After this, they keep aside the Test set, and randomly choose X% of their Train dataset to be the actual Train set and the remaining (100-X) % to be the Validation set, where X is a fixed number (say 80%), the model is then iteratively trained and validated on these different sets. So we will follow the same method to prepare data for training and testing phase.

The proposed model is built by using Convolutional neural network. Convolutional neural networks (CNN) are a special architecture of artificial neural networks, proposed by Yann LeCun in 1988. CNN uses some features of the visual cortex. After preprocessing step, the neural network is designed which is composed of 3 convolutional layer with 2 x 2 maxpooling.

Max-pooling is a technique used to reduce the dimensions of an image by taking the maximum pixel value of a grid. This also helps reduce over fitting and makes the model more generic. After maxpool layer , 2 fully connected layers are added. Since the input of fully connected layers should be two dimensional, and the output of convolution layer is four dimensional, flattening layer is added between them. At the very end of the fully connected layers is a softmax layer.

Model Training

After model construction it is time for model training**.** Artificial convolutional neural network is builded that can recognize images. The dataset is splitted into train and test dataset. Finally the Convolutional Neural Network model is trained using training dataset.

Model Testing and Evaluation

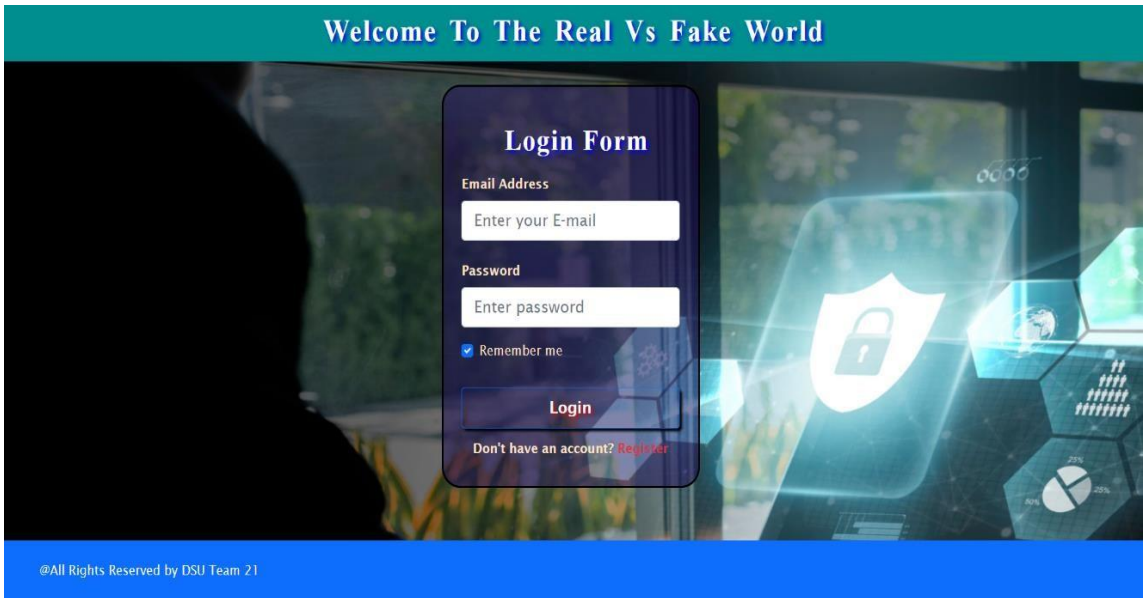
The model is trained by using training dataset. After this, it is possible to carry out the final step which is Model Testing**.** During this phase a test set of data is loaded. This data set has never been seen by the model and therefore its true accuracy will be verified. Finally, the saved model can be used in the real world. The name of this phase is model evaluation. This means that the model can be used to evaluate new data.

CHAPTER 7 EXPERIMENTATION

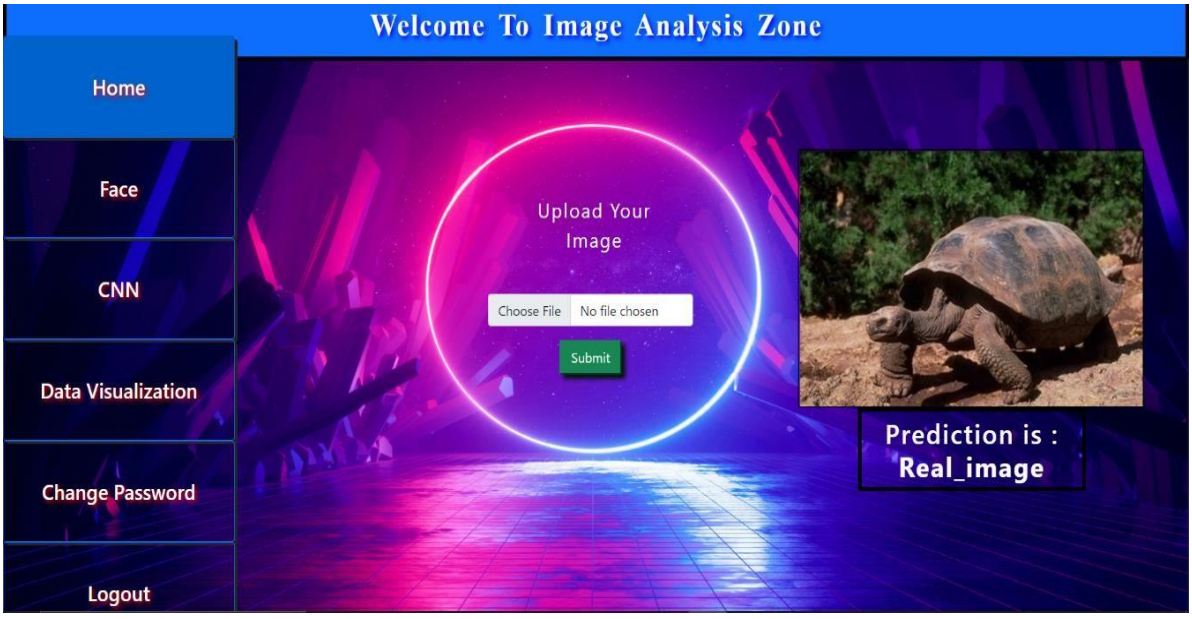
There are various techniques used by earlier researcher for detecting fake images but accuracy obtained by these models were less in case of face images using error level analysis . So the proposed method was incorporated with Gabor Filter which was especially used for face images detection to gain maximum accuracy .

CHAPTER 8 TESTING AND RESULTS

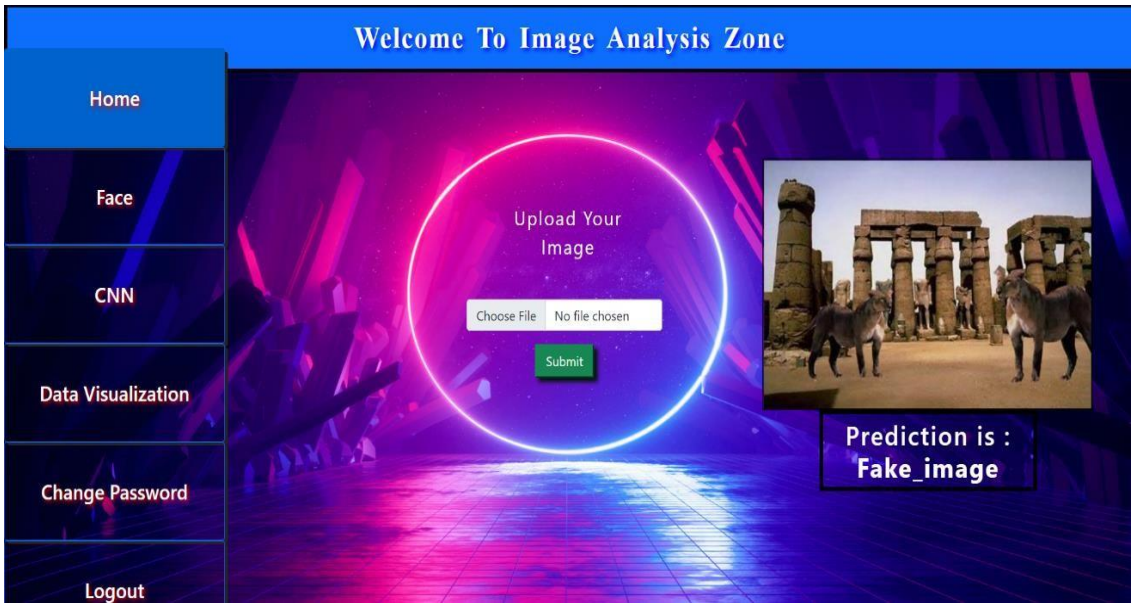
8.1 Results with Error Level Analysis



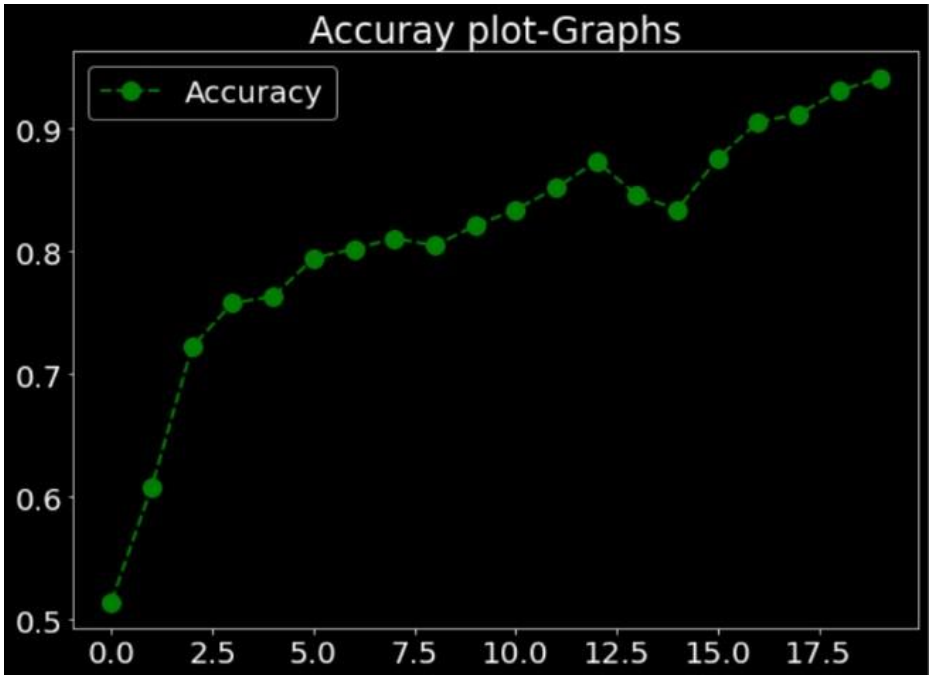
8(a) Login Page



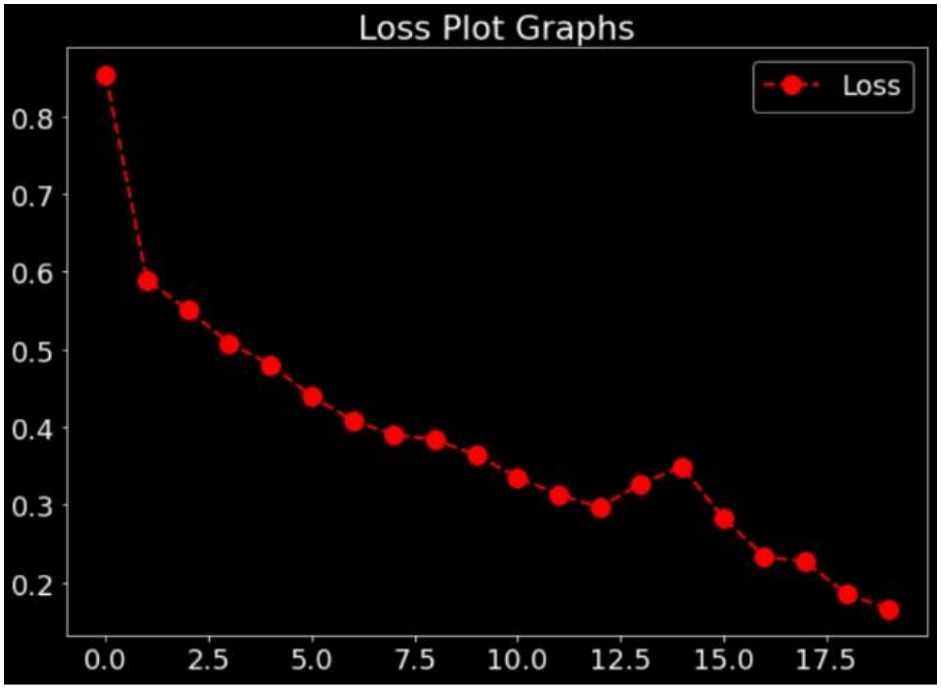
8.1(a) Real Image Prediction



8.1(b) Fake Image Prediction



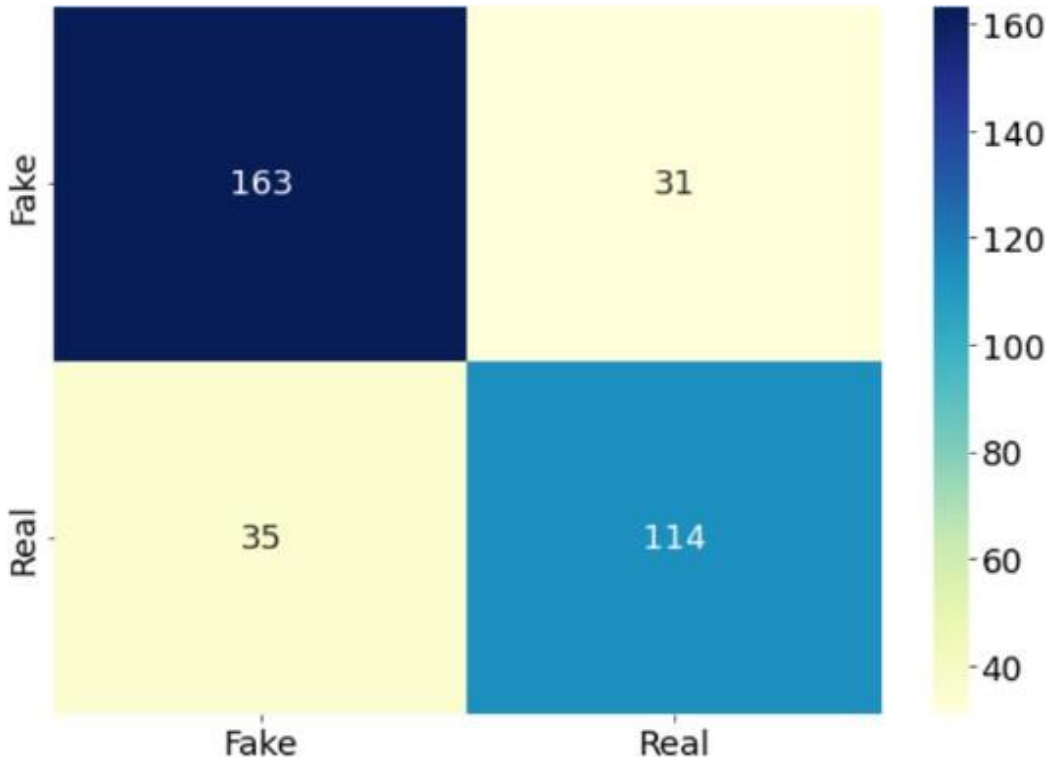
8.1 (c) Accuracy Graph for Error Level Analysis



8.1(d) Loss Graph for Error Level Analysis

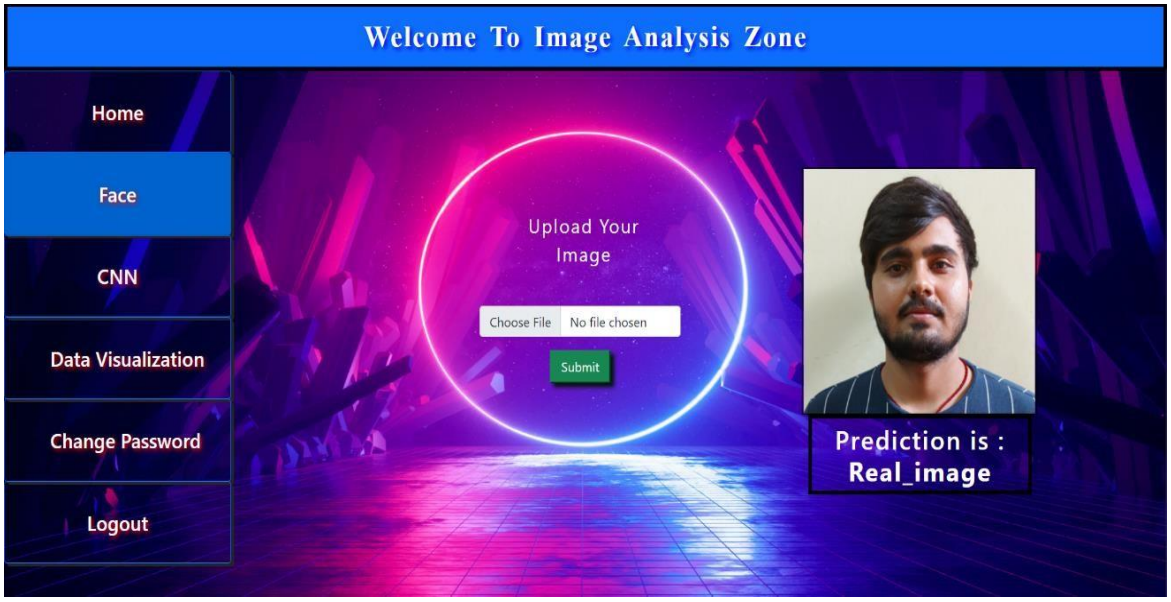
	precision	recall	f1-score	support
0	0.82	0.84	0.83	194
1	0.79	0.77	0.78	149
accuracy			0.81	343
macro avg	0.80	0.80	0.80	343
weighted avg	0.81	0.81	0.81	343

8.1(e) Classification Report

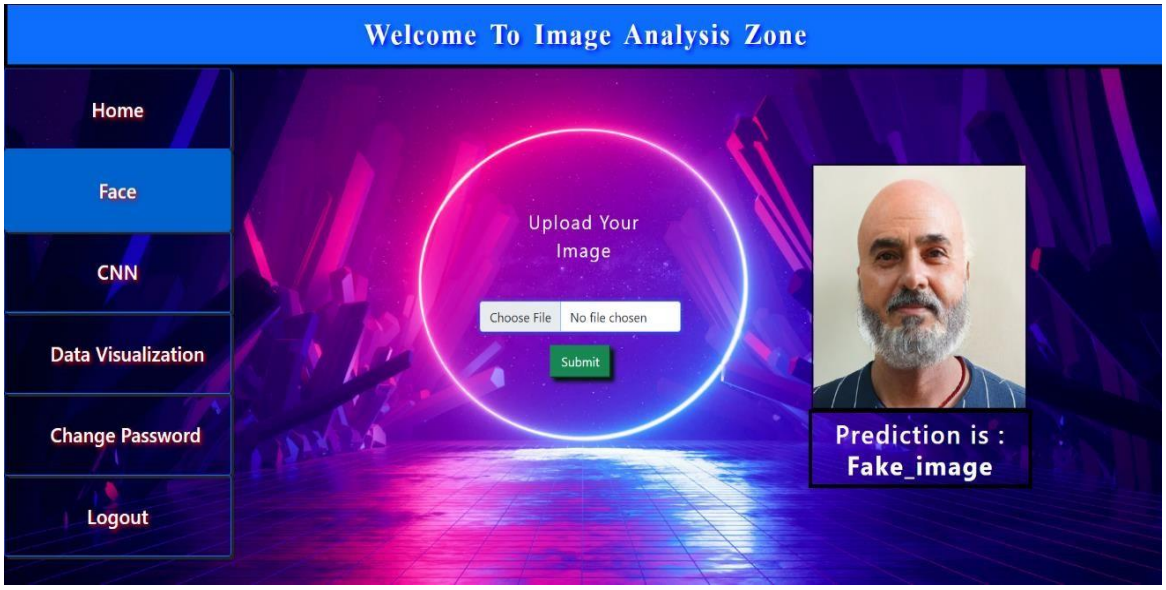


8.1(f) Confusion Matrix

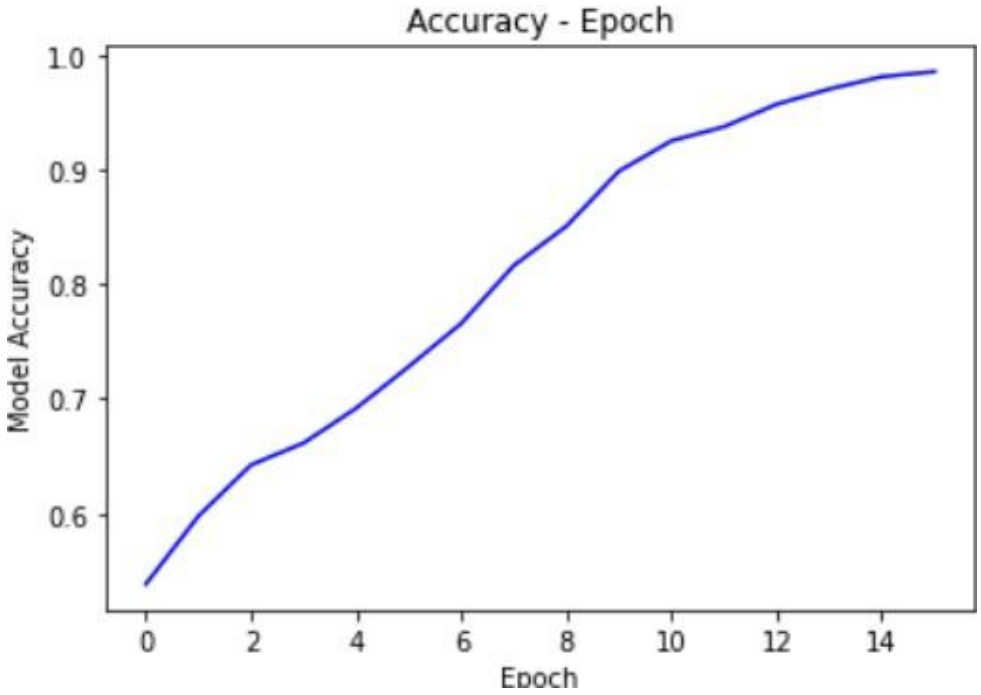
8.2 Results With Gabor Filter



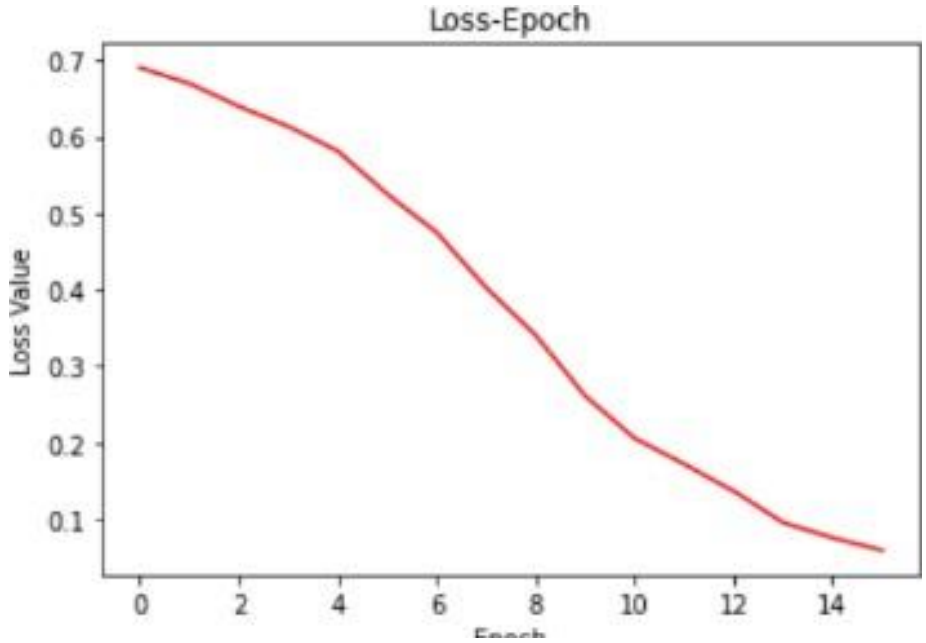
8.2(a) Real Image Prediction



8.2(b) Fake Image Prediction



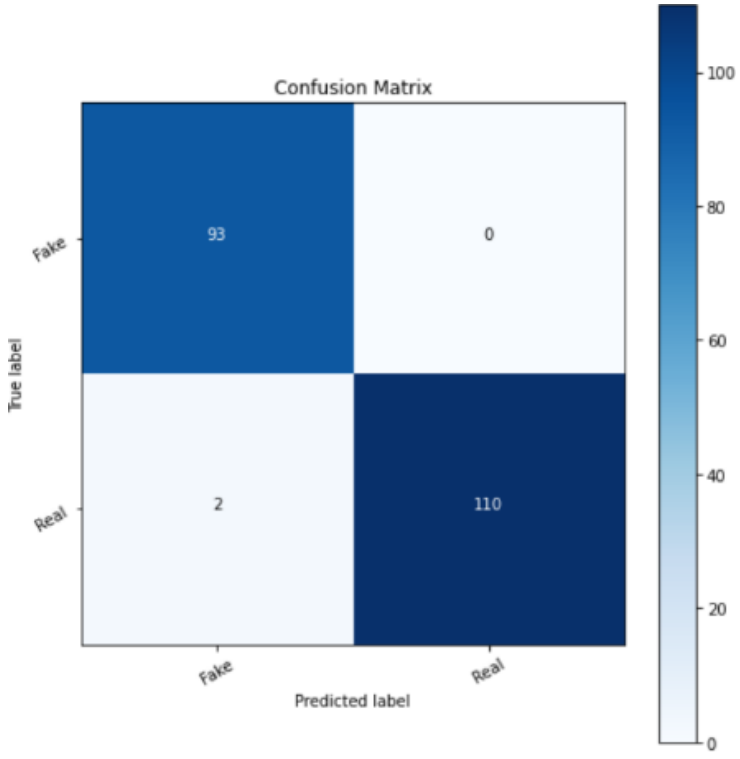
8.2(c) Accuracy Graph for Gabor Filter



8.2(d) Loss Graph for Gabor Filter

	precision	recall	f1-score	support
0	0.98	1.00	0.99	93
1	1.00	0.98	0.99	112
accuracy			0.99	205
macro avg	0.99	0.99	0.99	205
weighted avg	0.99	0.99	0.99	205

8.2(e) Classification Report



8.2(f) Confusion Matrix

CERTIFICATE OF PARTICIPATION

Dayananda Sagar University

School of Engineering, Kudlu Gate , Bengaluru - 560068

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Certificate of Participation

This is to certify that Mr./Ms. ISHA DHANYA has successfully participated in the Project Exhibition **'Project Expo-2k22'** organized by the Department of Computer Science & Engineering, School of Engineering, Dayananda Sagar University, Bengaluru on 3rd June, 2022

Dr. Meenakshi Malhotra	Dr. Girisha G S	Dr. A. Srinivas
Associate Professor and Project Exhibition Coordinator	**Professor and Chairman**	**Professor and Dean SOE**
Dept. of CSE, DSU, Bengaluru	**Dept. of CSE, DSU, Bengaluru**	**DSU, Bengaluru**

Dayananda Sagar University

School of Engineering, Kudlu Gate , Bengaluru - 560068

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Certificate of Participation

This is to certify that Mr./Ms. KUMARI RITIKA RANI has successfully participated in the Project Exhibition **'Project Expo-2k22'** organized by the Department of Computer Science & Engineering, School of Engineering, Dayananda Sagar University, Bengaluru on 3rd June, 2022

Dr. Meenakshi Malhotra	Dr. Girisha G S	Dr. A. Srinivas
Associate Professor and Project Exhibition Coordinator	**Professor and Chairman**	**Professor and Dean SOE**
Dept. of CSE, DSU, Bengaluru	**Dept. of CSE, DSU, Bengaluru**	**DSU, Bengaluru**

Dayananda Sagar University

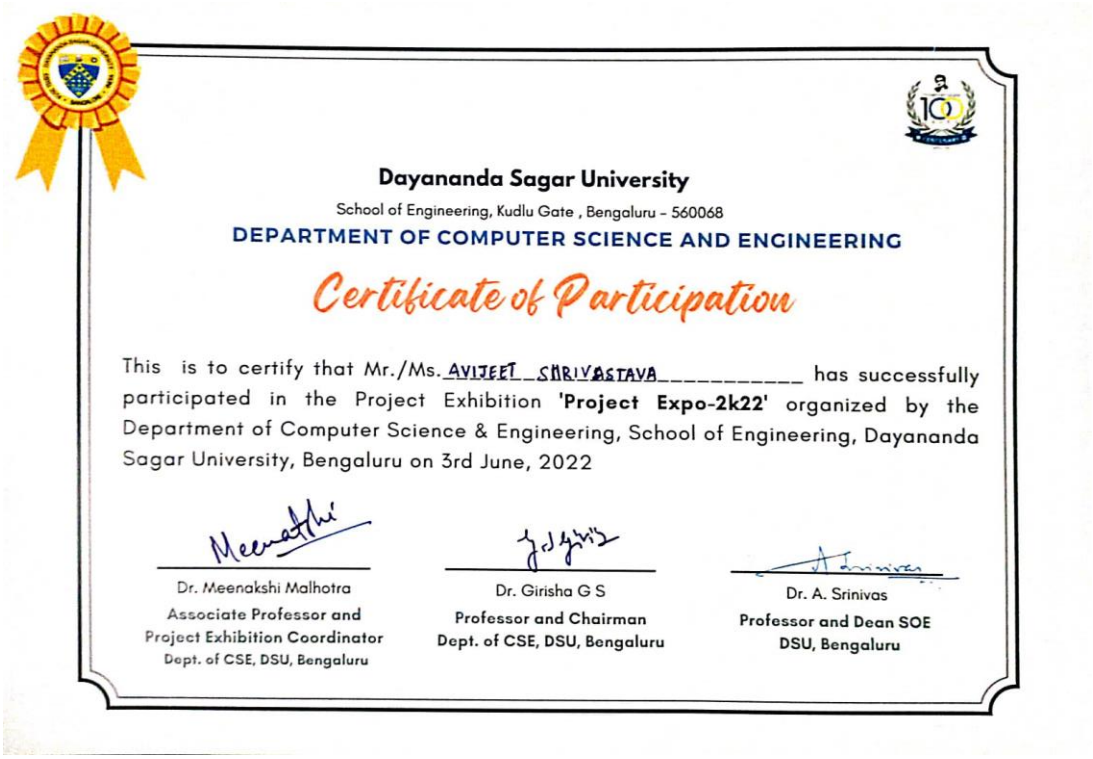
School of Engineering, Kudlu Gate , Bengaluru - 560068

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Certificate of Participation

This is to certify that Mr./Ms. ABHISHEK KUMAR has successfully participated in the Project Exhibition **'Project Expo-2k22'** organized by the Department of Computer Science & Engineering, School of Engineering, Dayananda Sagar University, Bengaluru on 3rd June, 2022

Dr. Meenakshi Malhotra	Dr. Girisha G S	Dr. A. Srinivas
Associate Professor and Project Exhibition Coordinator	Professor and Chairman	Professor and Dean SOE
Dept. of CSE, DSU, Bengaluru	Dept. of CSE, DSU, Bengaluru	DSU, Bengaluru



Dayananda Sagar University

School of Engineering, Kudlu Gate , Bengaluru - 560068

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Certificate of Participation

This is to certify that Mr./Ms. AVIJEET SRIVASTAVA has successfully participated in the Project Exhibition **'Project Expo-2k22'** organized by the Department of Computer Science & Engineering, School of Engineering, Dayananda Sagar University, Bengaluru on 3rd June, 2022

Dr. Meenakshi Malhotra	Dr. Girisha G S	Dr. A. Srinivas
Associate Professor and Project Exhibition Coordinator	Professor and Chairman	Professor and Dean SOE
Dept. of CSE, DSU, Bengaluru	Dept. of CSE, DSU, Bengaluru	DSU, Bengaluru

Dayananda Sagar University

School of Engineering, Kudlu Gate , Bengaluru - 560068

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Certificate of Participation

This is to certify that Mr./Ms. AKSHAT KUMAR has successfully participated in the Project Exhibition **'Project Expo-2k22'** organized by the Department of Computer Science & Engineering, School of Engineering, Dayananda Sagar University, Bengaluru on 3rd June, 2022

Dr. Meenakshi Malhotra
Associate Professor and
Project Exhibition Coordinator
Dept. of CSE, DSU, Bengaluru

Dr. Girisha G S
Professor and Chairman
Dept. of CSE, DSU, Bengaluru

Dr. A. Srinivas
Professor and Dean SOE
DSU, Bengaluru

REFERENCES

[1] A. Piva, “An overview on image forensics,” ISRN Signal Processing, pp. 1–22, 2012.

[2] E. Kee, J. O’Brien, and H. Farid, “Exposing photo manipulation with inconsistent shadows,” ACM Transactions on Graphics, vol. 32, no. 3, pp. 28–58, 2013.

[3] T. de Carvalho, C. Riess, E. Angelopoulou, H. Pedrini, and A. Rocha, “Exposing digital image forgeries by illumination color classification,” IEEE Trans. Inf. Forensics Security, vol. 8, no. 7, pp. 1182–1194, 2013

[4] Y. Wu, W. Abd-Almageed, and P. Natarajan, “Deep matching and validation network: An end-to-end solution to constrained image splicing localization and detection,” in ACM International Conference on Multimedia, 2017, pp. 1480–1502.

[5] Y. Lui, X. Zhu, X. Zhao, and Y. Cao, “Adversarial learning for constrained image splicing detection and localization based on atrous convolution,” IEEE Trans. Inf. Forensics Security, vol. 14, no. 10, pp. 2551–2566, 2019.

[6] A. Radford, et al.. "Unsupervised representation learning with deep convolutional generative adversarial networks," arXiv preprint arXiv:1511.06434, 2015.

[7] Gulrajani, Ishaan, et al. "Improved training of wasserstein gans," Advances in Neural Information Processing Systems. 2017