IRIS TEMPLATE ATTACK DETECTION USING MACHINE LEARNING AND DEEP LEARNING METHODS

MASTER OF SCIENCE IN INFORMATION TECHNOLOGY

SUBMITTED BY

AYSHA A

20PIT003

Under the Guidance of

Dr. D. SHANMUGAPRIYA, M.Sc., M.Phil., Ph.D., SET.,

Assistant Professor and Head,

Department of Information Technology

AVINASHILINGAM INSTITUTE FOR HOME SCIENCE AND HIGHER EDUCATION FOR WOMEN

SCHOOL OF PHYSICAL SCIENCES AND COMPUTATIONAL SCIENCES

DEPARTMENT OF INFORMATION TECHNOLOGY

COIMBATORE-641043

MAY 2022

ABSTRACT

The iris is a guarded, outwardly visible function that maintains its genomic structure during the entire adulthood. It's a worthy choice to be used as a biometric for identifying persons due to these characteristics. Each people's iris is distinct. Nevertheless Even fraternal identical twins and a person's left and right irises have distinct features. The chances of discovering two persons with the same iris patterns are estimated to be one in 1052. As biometric identification systems become more common, an attacker's incentive to stage a system compromise grows, as does the requirement to assure system security and integrity.

The Objective is to find the iris template attack in iris template of each user that is being stored on the background. A mix of multiple pre-processing and classification algorithms are being involved and used to this suggested project such as Eye Detection, Iris Detection, Morphological Operations, Edge Detection using Contour and Iris Segmentation. The template further undergoes possible Template Attack to create the attacked template image.

The model is being built using deep learning technique namely Convolutional Neural Network (CNN) without max pooling which provides 97.50% Accuracy and CNN with max pooling gives 100% Accuracy. In Machine Learning (ML) techniques, logistic regression is applied to classify and detect the attacked template from genuine iris template and it gives 90% accuracy.

Keywords: ***CNN, Digital Image Processing, Iris, Iris Attack Detection, Logistic Regression, Template Attack.***

1. INTRODUCTION

The chapter 1 consists of a brief discussion about the general introduction to the proposed system, the problem statement, motivation and justification and its objectives.

The iris is a delicate, round structured and a colored part of eye in the human eye that is located between the photoreceptors (the cornea and the lens). The pupil is a circular opening that is perforated near to the iris's center. The iris' job is to manage the intensity of light that reaches the pupil via the sphincter and dilator muscles that control pupil size. Iris patterns will not be identical in one-egged twins or future clones of an individual.

The iris is a unique internal organ since it is adequately protected from environmental harm by the eyelid and cornea. The most accurate and quick biometric authentication technology available today is Iris Recognition. But, it suffers from various forms of attacks that may affect the template in some way or the other. And when the Iris template is compromised, the whole biometric system will collapse and there is a high chance of heavy damage for the information stored in the system.

Thus, it is essential to come up with a good solution to this challenge that will secure the biometric system's safety and security. The goal is to create a baseline solution to detect iris template attack and to classify whether it is a template attack or a genuine iris sample according to various performance evaluation metrics discussed in the following sessions.

1.5 PROBLEM STATEMENT

To identify a template attack in each user's iris template that is saved in the background.

1.6 OBJECTIVE

To provide a baseline solution for detecting iris template attacks and classifying whether they are attacks or genuine iris samples based on a variety of performance assessment parameters.

1.7 MOTIVATION AND JUSTIFICATION

It is essential to come up with a good solution to this challenge that will secure the biometric system's safety and security.

Our top aim is to detect the assault before it damages the iris biometric system and template.

To give a baseline answer to the problem, a mix of multiple pre-processing and classification algorithms are being developed and used to this suggested project.

SUMMARY

The current chapter depicts an outline of the proposed project such as its Problem Statement, Objective, Motivation and justification, as well as the composite introduction to the general biometric system, The Iris and its comparison to other biometric systems. The next chapter shows the background study done to develop this project.

3. METHODOLOGY

The proposed system is a framework created for template attack detection and classification. The entire methodology is divided into five modules, namely, image acquisition, image preprocessing, template generation, model building using machine learning and deep learning and performance evaluation to evaluate the model performance.

Figure 1 shows the overview of the proposed methodology.

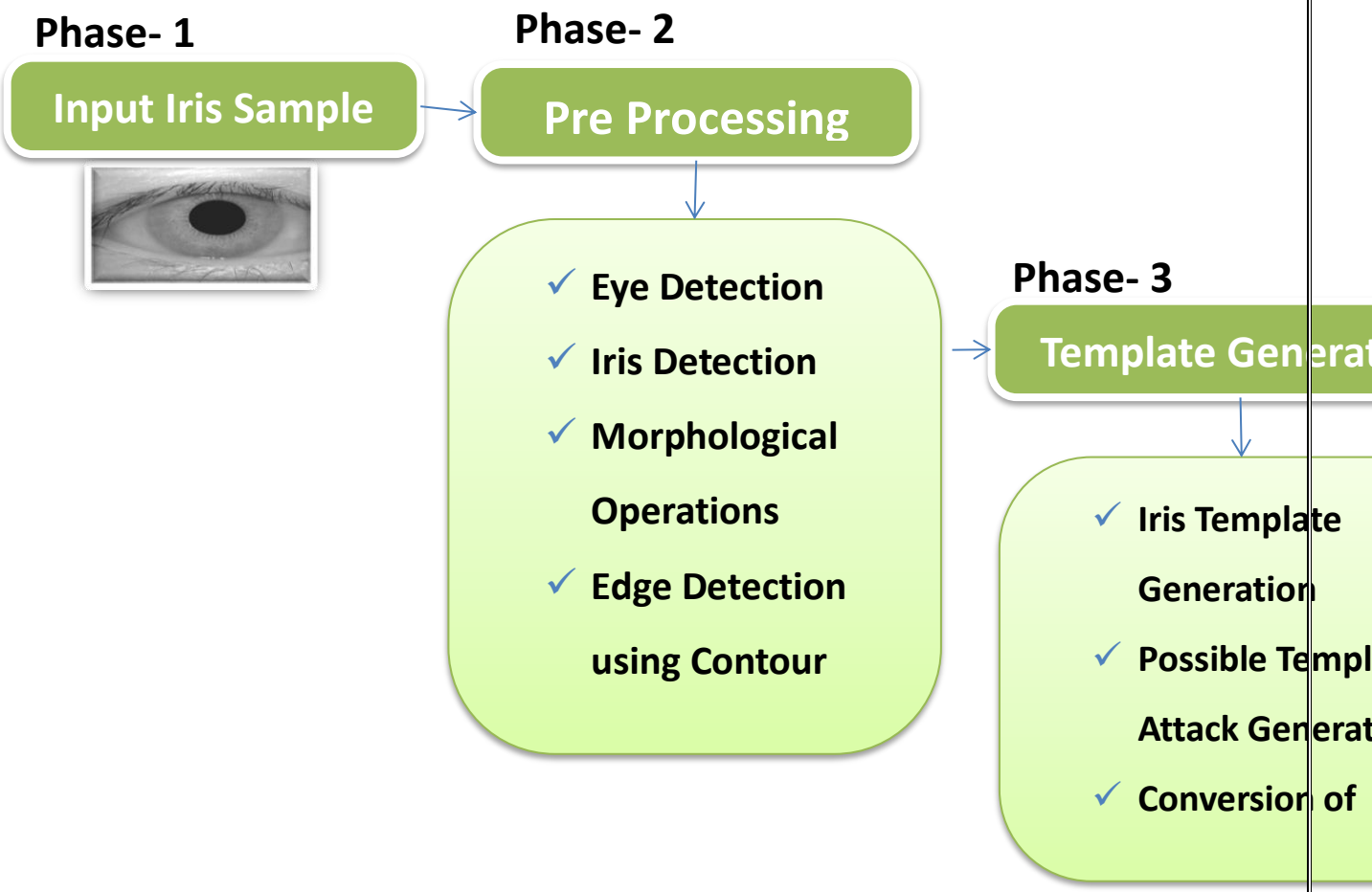


Figure 1. Methodology Overview

The entire methodology is divided into five modules, namely, image acquisition, image preprocessing, template generation, model building using machine learning and deep learning and performance evaluation to evaluate the performance.

3.1 PHASE 1: IMAGE ACQUISITION

Iris identification starts with isolating the real iris area in a digitized eye picture. The top and lower sections of the iris area are generally obscured by the eyelids and eyelashes. The imaging quality of ocular pictures determines the success of segmentation. In order to recognize the authentic user based on iris patterns, images from the CASIA iris database are utilized as input images.

3.2 PHASE 2: IMAGE PRE-PROCESSING

The preprocessing stage is crucial to generate the template and it also increase the performance of an iris pattern attack detection system, because data misrepresented as iris motif information would damage the generated biometric templates, leads to poor user recognition. Phase 2 is divided into five steps explained in the following Figure 2:

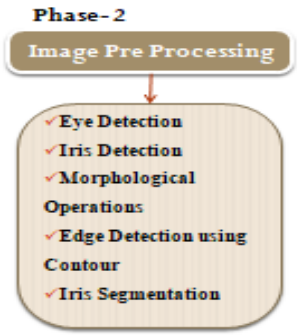


Figure 2: Steps involved in Image Pre-processing

3.2.1 Eye Detection

Casia iris images are required for this procedure. The ocular traits will be captured in positive photos. The characteristics of eyes' will then be retrieved from the photos. OpenCV has a training approach for detecting eyeballs in images called the Haar Cascade model. One must first import OpenCV before loading the relevant XML files. These XML files provide a collection of features that will be applied to the image. The eyes are identified using the later function detect Multi Scale, which returns a rectangle for the observed eyes.

3.2.2 Iris Detection

This step represents the second process where it is responsible for iris detection. The model takes the images detected in the last phase and it ensures that there is an iris inside the eyes the image will pass to the next step if it passes this step. It highlights the iris using Hough circle based on the specified iris threshold value.

3.2.3 Morphological Operations

Morphological processing refers to a group of indiscriminate operations which copes with the shape or morphological of feature representation. Here, Operations such as, Thresholding, Opening and Closing are being suggested in this proposed project for better feature extraction and classification.

3.2.4 Edge and Contour Detection

Edges appear at locations in a picture where the brightness values vary greatly, and as a result, they frequently show the edges, or occluding limits, of the objects in the scene. Large luminance shifts, on the other hand, might correlate to surface marks on objects.

Edge detection with contour in computer vision has historically been accomplished using a sequential filter to estimate a first or second component product to convolve the stream.

3.2.5 Iris Segmentation using Hough Transform

The Circle Hough Transform (CHT) is a core background subtraction strategy for recognising spheres in faulty images in digital image processing. The circles are created by "voting" in the Hough parameter space after selecting extreme points in an aggregate array.

3.3 PHASE 3: TEMPLATE GENERATION

A template-based method may be useful for templates with few characteristics or when the majority of the template picture serves as the matching image. The following Figure 3 shows the steps involved in the Template generation phase:

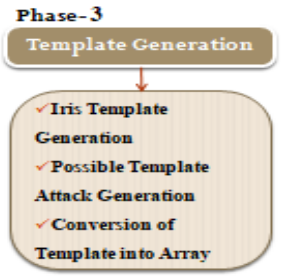


Figure 3: Steps in Template Generation Phase

Template-based matching may necessitate a high amount of sampling points, the sample size can be reduced by lowering the quality of the searching and pattern images by the same factor and doing sample. So, the template is generated for genuine iris image gathered from the above mentioned pre-processing techniques and stored it as genuine iris template. The possible attack in done in some samples of generate iris template and stored it as attacked iris template.

3.3.1 Converting Template into an Array Format

In this step, it's used to transform an image to a linear array, which is subsequently saved as a Data frame. It is accomplished by applying PIL and numpy library. The two csv file is generated for both the genuine iris template and attacked iris template.

3.4 PHASE 4: MODEL BUILDING

In this phase, the model building is done using the csv file gathered from the previous phase. The csv file is trained using machine learning model (Logistic Regression) and deep learning model (CNN) to classify and detect the iris template attack.

Logistic Regression

In the binary classification task, logistic regression is a guided learning approach. It is used to calculate or predict the probability of an occurring binary (yes/no) event. This is known as binary classification as there are only two potential answers to this question: yes or no.

The sigmoid function converts the expected values into probabilities. Its value cannot exceed tht limit of 0 and 1, and results in an "S"-shaped curve.

CNN

A Convolutional Neural Network (CNN) is a type of neural network that is used to process images. A digital picture is a binary representation of visual data. Our brain examines a large amount of info the moment we see a picture. Every neuron seems to have its own receptive field, which is linked to other neurons to cover the entire visual field. Each neuron in the human vision system can really only react to stimuli in a restricted area known as the receptive field, so each neuron in a CNN can only analyse data in its receptive field. Comparatively simple patterns (lines, curves, etc.) being recognized first in the layers, followed by more intricate patterns (faces, objects, etc.).

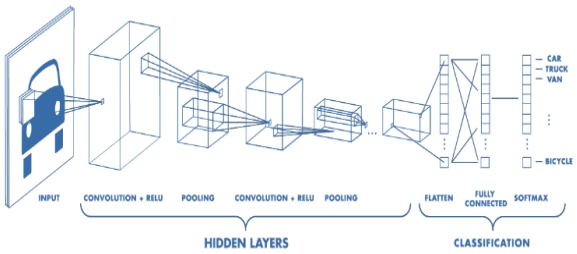


Figure 4. Basic Architecture of CNN

3.5 PHASE 5: PERFORMANCE EVALUATION

Performance measurements are used to examine the performance of the machine learning model. Any endeavour will necessitate a review of machine learning models or algorithms. A number of performance evaluation measures can be employed to put a model to the test.

The classification here is whether the image is genuine template or attacked template. The following are some popular words to be aware of:

- True positives (TP) are predictions that turn out to be true.
- False positives (FP) are when a positive outcome is predicted but the outcome is really negative.
- True negatives (TN) are predicted negatives that turn out to be true.
- False negatives (FN) are predictions that turn out to be true.

Accuracy

Although this is the most typically used indicator for assessing a model, it is not a consistent indicator of its performance. The scenario grows considerably worse when the courses are uneven.

$$\frac{TP + TN}{TP + FP + TN + FN} \quad \textbf{\{1\}}$$

True Positive Rate/Recall/Sensitivity

Favorable events expressed as a proportion of positive linear events. As a result, the portion (TP + FN) represents the collection's total amount of useful cases.

$$\frac{TP}{TP + FN} \quad \textbf{\{2\}}$$

Specificity

Critical events expressed as a proportion of total critical incidents. As a consequence, the component (TN + FP) represents the collection's total set of negative instances.

$$\frac{TN}{TN + FP} \quad \textbf{\{3\}}$$

F1 score

A symmetrical average exists between precision and recall. This component takes into account both, thus the higher the F1 score, the better. lowering the numerator lowers the final F1 score significantly. So, if the anticipated wins are true positives (accuracy), and the model does not overlook positives and predicts them as negatives, it scores well enough in the F1 Evaluation metrics (recall).

Using the above-mentioned performance metrics, the applied deep learning model and machine learning model is evaluated to compare its performance and conclude the best from it.

4. RESULTS AND DISSCUSSION

Figure 5 explains the iris dataset that contains the iris information of each user. It consists of 756 iris images from 108 Individuals grouped into 108 folders.

PHASE 1: IMAGE ACQUISITION

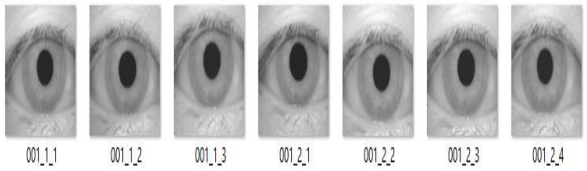


Figure 5: Sample of CASIA iris dataset

PHASE 2: IMAGE PREPROCESSING

Figure 6 shows the outcome of cascade classifier that is used to detect the eyes using CASIA iris dataset.

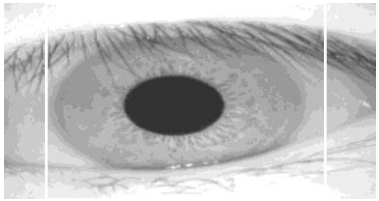


Figure 6: Eye Detection using Cascade Classifier

Figure 7 shows that the specific iris part is detected using Hough circles based on the threshold value.

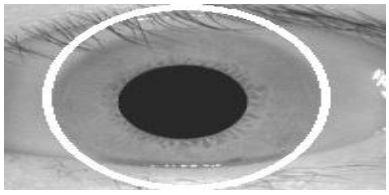


Figure 7: Iris detection using Hough circles

Figure 8 explains the morphological operation such as erosion and dilation of iris detected gray scale image.

(a) Erosion **(b) Dilation**

Figure 8: Erosion and Dilation in Gary Scale Image

Figure 9 and 10 explains the contour based edge detection using gray scale image

Figure 9: Contour based Edge Detection in Gray Scale Image

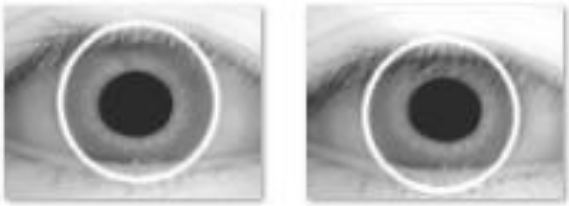


Figure 10: Outcome of Contour based Edge Detection

Figure 11 displays the segmented part of the iris using Hough transform.

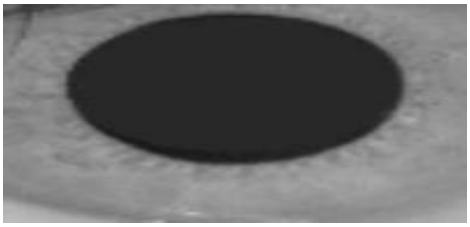


Figure11: Iris segmentation using Hough Transformation

PHASE 3: TEMPLATE GENERATION

The pre-processed image of user eye is now considered as genuine template. Now, the attack is performed in some genuine template by modifying the image. It is considered as attacked template.

Figure 12 shows the attacked template from the genuine template image.

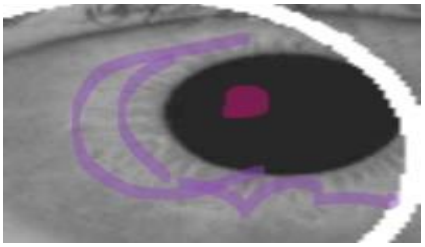
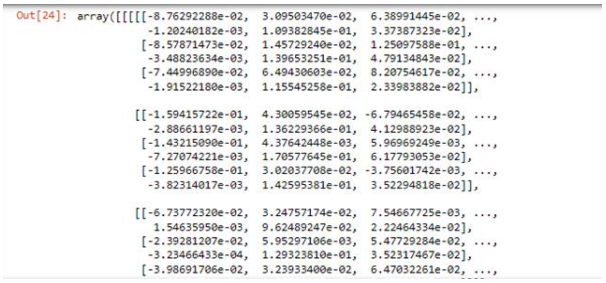


Figure 12: Attacked template from the genuine template image.



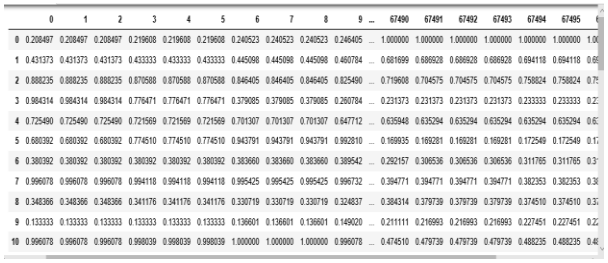
```
Out[24]: array([[[[[-8.76292288e-02,  3.09503470e-02,  6.38991445e-02, ...,
                   -1.20240182e-03,  1.09382845e-01,  3.37387323e-02],
                  [-8.57871473e-02,  1.45729240e-02,  1.25097588e-01, ...,
                   -3.48823634e-03,  1.39653251e-01,  4.79134843e-02],
                  [-7.44996890e-02,  6.49430603e-02,  8.20754617e-02, ...,
                   -1.91522180e-03,  1.15545258e-01,  2.33983882e-02]],

                 [[-1.59415722e-01,  4.30059545e-02, -6.79465458e-02, ...,
                   -2.88661197e-03,  1.36229366e-01,  4.12988923e-02],
                  [-1.43215090e-01,  4.37642448e-03,  5.96969249e-03, ...,
                   -7.27074221e-03,  1.70577645e-01,  6.17793053e-02],
                  [-1.25966758e-01,  3.02037708e-02, -3.75601742e-03, ...,
                   -3.82314017e-03,  1.42595381e-01,  3.52294818e-02]],

                 [[-6.73772320e-02,  3.24757174e-02,  7.54667725e-03, ...,
                    1.54635950e-03,  9.62489247e-02,  2.22464334e-02],
                  [-2.39281207e-02,  5.95297106e-03,  5.47729284e-02, ...,
                   -3.23466433e-04,  1.29323810e-01,  3.52317467e-02],
                  [-3.98691706e-02,  3.23933400e-02,  6.47032261e-02, ...,
```

Figure 13: Both template image into array format

Figure 14 displays the template information in table format.



	0	1	2	3	4	5	6	7	8	9	...	67490	67491	67492	67493	67494	67495	6
0	0.208497	0.208497	0.208497	0.219608	0.219608	0.219608	0.240523	0.240523	0.240523	0.246405	...	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.0
1	0.431373	0.431373	0.431373	0.433333	0.433333	0.433333	0.445098	0.445098	0.445098	0.460784	...	0.681699	0.686928	0.686928	0.686928	0.694118	0.694118	0.6
2	0.888235	0.888235	0.888235	0.870588	0.870588	0.870588	0.846405	0.846405	0.846405	0.825490	...	0.719608	0.704575	0.704575	0.704575	0.758824	0.758824	0.7
3	0.984314	0.984314	0.984314	0.776471	0.776471	0.776471	0.379085	0.379085	0.379085	0.260784	...	0.231373	0.231373	0.231373	0.231373	0.233333	0.233333	0.2
4	0.725490	0.725490	0.725490	0.721569	0.721569	0.721569	0.701307	0.701307	0.701307	0.647712	...	0.635948	0.635294	0.635294	0.635294	0.635294	0.635294	0.6
5	0.680392	0.680392	0.680392	0.774510	0.774510	0.774510	0.943791	0.943791	0.943791	0.992810	...	0.169935	0.169281	0.169281	0.169281	0.172549	0.172549	0.1
6	0.380392	0.380392	0.380392	0.380392	0.380392	0.380392	0.383660	0.383660	0.383660	0.389542	...	0.292157	0.306536	0.306536	0.306536	0.311765	0.311765	0.3
7	0.996078	0.996078	0.996078	0.994118	0.994118	0.994118	0.995425	0.995425	0.995425	0.996732	...	0.394771	0.394771	0.394771	0.394771	0.382353	0.382353	0.3
8	0.348366	0.348366	0.348366	0.341176	0.341176	0.341176	0.330719	0.330719	0.330719	0.324837	...	0.384314	0.379739	0.379739	0.379739	0.374510	0.374510	0.3
9	0.133333	0.133333	0.133333	0.133333	0.133333	0.133333	0.136601	0.136601	0.136601	0.149020	...	0.211111	0.216993	0.216993	0.216993	0.227451	0.227451	0.2
10	0.996078	0.996078	0.996078	0.998039	0.998039	0.998039	1.000000	1.000000	1.000000	0.996078	...	0.474510	0.479739	0.479739	0.479739	0.488235	0.488235	0.4

Figure 14: Template information in table format

PHASE 4: MODEL BUILDING

In this phase, the template dataset is split in the ratio of 80-20 for train and test data. Figure 15.1 and 15.2 shows the model accuracy and loss in the CNN model with and without max pooling.

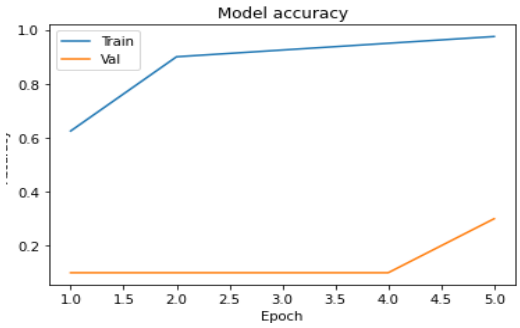


Figure 15.1 CNN model accuracy without max pooling

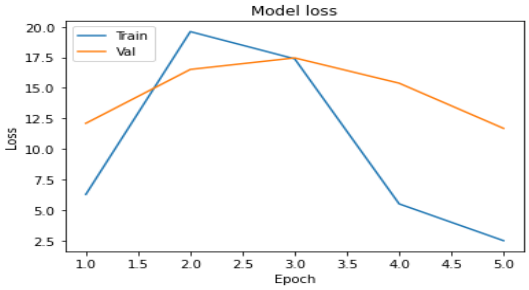


Figure 15.2 CNN model loss without max pooling

From figure 15.1 and 15.2, it is observed that the CNN model achieves maximum of 97.5% accuracy with minimal loss to detect and classify the template attack and genuine attack.

Figure 16.1 and 16.2 shows the model accuracy and loss while training the CNN model with max pooling.

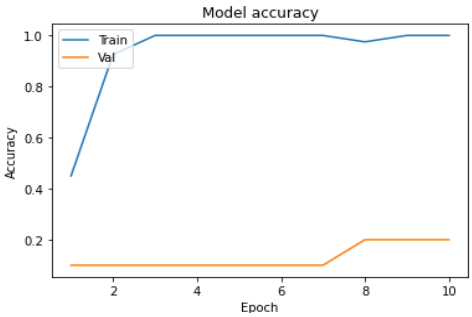


Figure 16.1 CNN model accuracy using train data with max pooling

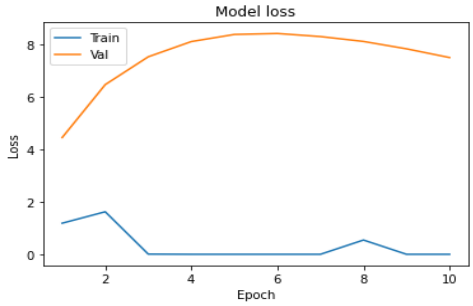


Figure 16.2 CNN model loss using train data with max pooling

From figure 16.1 and 16.2, it is observed that the CNN model achieves maximum of 100% accuracy with minimal loss to detect and classify the template attack and genuine attack with max pooling.

Figure 17.1 shows the confusion matrix and 17.2 shows other evaluation metrics of the LR model.

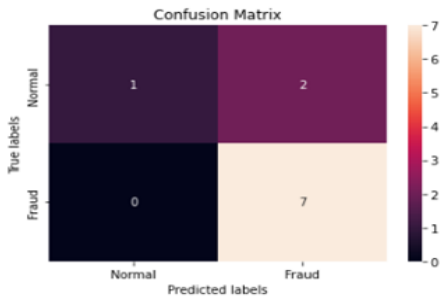


Figure 17.1 Confusion matrix to evaluate LR Model

```
Confusion Matrix:
  [[1 1]
 [0 8]]

 Recall: 1.0

 Accuracy Score: 0.9

 Precision Score: 0.8888888888888888

 F1 Score: 0.9411764705882353
```

Figure 17.2 Other Metrics to evaluate LR Model

It is observed that the LR model is able to detect the entire attack template but it fails to detect one genuine template and achieves the precision score of 89%. It achieves 90% accuracy and 100% recall with 94.11% F1 score.

PHASE 5: PERFORMANCE EVALUATION

Figure 18.1 shows the performance evaluation of CNN with and without max pooling layer and Figure 18.2 shows the performance evaluation of LR to detect and classify the attack template and genuine template.

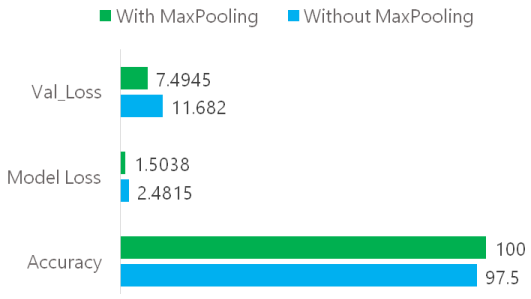


Figure 18.1: Performance Evaluation of CNN.

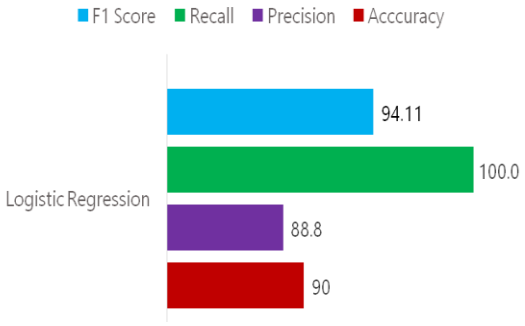


Figure 18.2: Performance Evaluation of LR.

5. CONCLUSION AND FUTURE SCOPE

The attack template and genuine template is detected and classified in CASIA iris V1 dataset. This image dataset successfully pre-processed using pre-processing techniques such as eye detection using cascade classifier, iris detection using Hough circle, morphological operation, edge contour detection and iris segmentation using Hough transform. This segmented image for CASIA iris image is considered as genuine template. The attack is successfully performed in genuine template and stored it as attacked template models namely CNN and LR are successfully trained using train and test data of both templates. CNN without Max pooling provides 97.5% accuracy while CNN with Max Pooling provides 100% accuracy. The Logistic Regression Model also provides a better accuracy of 90% with precision score of 88%, Recall Score of 100% and F1 score as 94.11%. It is observed that both the model performs better and achieves higher accuracy to detect and classify the genuine template and attack template. It is concluded that ML technique using LR and DL technique using CNN are explored successfully and obtains higher accuracy in detecting and classifying the genuine template and attacked template in CASIA iris V1 images.

In future, other algorithms in DL such as, Self Organizing Maps, Generative Adversarial Networks, etc.., can be explored for detecting and classifying the genuine template and attacked template using CASIA iris dataset.

6. REFERENCES

[1] B. E. D. A. D. Fatima, A. D. J. O. U. D. J. Réda and B. O. U. S. A. H. B. A. Nassima. "Studies of the Robustness of a Transformation-Based Multi-Biometric Template Schemes Protection," Int. J. Com. Dig. Sys. 2022, 11(1).

[2] M. Gupta and P. Sehgal. "HsIrisNet: Histogram based Iris recognition to allay replay and template attack using deep learning perspective," Pattern Recognition and Image Analysis. 2020, 30(4), pp. 786-794.

[3] J. McGrath, K. W. Bowyer and A. Czajka, "Open source presentation attack detection baseline for iris recognition," arXiv preprint arXiv:1809.10172. 2018.

[4] V. K. Sinha, A. K. Gupta and R. Khanna. "Detection of Fake Iris by using Frame Difference and Reflection Ratio."

[5] R. Sujitha and N. Lalithamani. "Counter Measures for Indirect Attack for Iris based Biometric Authentication," Indian Journal of Science and Technology. 2016, 9(19), pp. 1-7.

[6] M. Gomez-Barrero, J. Galbally, P. Tome and J. Fierrez. "On the vulnerability of iris-based systems to a software attack based on a genetic algorithm," Iberoamerican Congress on Pattern Recognition Springer, Berlin, Heidelberg. 2012, September, . pp. 114-121.

[7] C. Rathgeb and A. Uhl. "Statistical attack against iris-biometric fuzzy commitment schemes," CVPR 2011 WORKSHOPS. IEEE. 2011, June. pp. 23-30.

[8] J. Daugman. "How iris recognition works The essential guide to image processing" Academic Press. 2009, pp. 715-739.

[9] S. Sanderson and J. Erbetta. "Authentication for secure environments based on iris scanning technology," IEE Colloquium on Visual Biometrics. 2000.

[10] E. Wolff. "Anatomy of the Eye and Orbit," 7th edition. H. K. Lewis & Co. LTD. 1976.

[11] R. Wildes. "Iris recognition: an emerging biometric technology," Proceedings of the IEEE. 1997, 85(9).

[12] J. Daugman. “Biometric personal identification system based on iris analysis,” United States Patent, 1994, Patent Number: 5,291,56.

[13] J. Daugman. “High confidence visual recognition of persons by a test of statistical independence,” IEEE Transactions on Pattern Analysis and Machine Intelligence. 1993, 15(11).

[14] R. Wildes, J. Asmuth, G. Green, S. Hsu, R. Kolczynski, J. Matey and S. McBride. “A system for automated iris recognition,” Proceedings IEEE Workshop on Applications of Computer Vision. Sarasota, FL. 1994, pp. 121-128.

[15].https://www.cs.auckland.ac.nz/courses/compsci773s1c/lectures/ImageProcessing-html/topic4.htm

7. ACKNOWLEDGEMENT

I would like to acknowledge the help rendered by Center for Cyber Intelligence, DST – CURIE – AI Sponsored Phase II for providing the laboratory facilities to execute my project.