DEPARTMENT OF COMPUTER SCIENCE ENGINEERING,

SCHOOL OF ENGINEERING AND TECHNOLOGY,

SHARDA UNIVERSITY, GREATER NOIDA

Detetion Of Forged Currency Notes Using Machine Learning Algorithms

A project submitted
in partial fulfillment of the requirements for the degree of
Bachelor of Technology in Computer Science and Engineering

by

MADHAV SHARMA (2018009090)

GAURAV JOSHI (2018004746)

ABHISHEK SINGH (201801272)

Supervised by:

Dr. Vivek Sharma, Associate Professor

May, 2022

CERTIFICATE

This is to certify that the report entitled **"Detection Of Forged Currency Notes Using Machine Learning Algorithms"** submitted by "Mr. MADHAV SHARMA (2018009090), Mr. GAURAV JOSHI (2018004746), Mr. ABHISHEK SINGH (2018012726)" to "Sharda University", towards the fulfillment of requirements of the degree of **"Bachelor of Technology"** is record of bonafide final year Project work carried out by him in the "Department of Computer Science and Engineering, School of Engineering and Technology, Sharda University".

The results/findings contained in this Project have not been submitted in part or full to any other University/Institute for award of any other Degree/Diploma.

Signature of Supervisor

Name : Dr. Vivek Sharma

Designation : Associate Professor (CSE)

Signature of Head of Department

Name : Prof. (Dr.) Nitin Rakesh

(Office seal)

Place:

Date:

Signature of External Examiner
Date :

ACKNOWLEDGEMENT

A major project provides an excellent opportunity for learning and self-development. We consider our self honored and blessed to have plethora of wonderful people lead us through in completion of this project.

First and foremost we would like to thank Dr. Nitin Rakesh, HOD, CSE who gave us an opportunity to undertake this project.

My grateful thanks to **Dr.Vivek Sharma** for his guidance in our project work.

Dr. Vivek Sharma, who in spite of being extraordinarily busy with academics, took his time out to listen, guide and keep us at the correct path. We do not know where we would have been without his assistance.

CSE department monitored our progress and arranged all facilities to make life easier. We choose this moment to acknowledge their contribution gratefully.

Name and signature of Students

MADHAV SHARMA (2018009090)

GAURAV JOSHI (2018004746)

ABHISHEK SINGH (2018012726)

ABSTRACT

Currency is a medium of exchange either to purchase or sell the goods and services. In today's world, With enhanced technology, counterfeit currency has become a concerning threat. Counterfeit currency leads to harm the country's economy.

Severe implications are such as artificial inflation, terror funding. These concerning causes destabilize the real value of money and inflation surges over its sky. Moreover, due to terror-related activities the social-harmony gets disturbed, this can be termed as "Economic Terrorism".

The proposed methodology in this paper uses Random Forest & SVM Machine Learning Algorithms to classify forged currency notes The algorithms are evaluated using metrics Accuracy, Confusion Matrix and Classification Report. The results show that SVM has outperformed Random Forest by achieving an accuracy of 99.63%.

Keywords: Counterfeit Currency, Economic Terrorism, Machine Learning, Random Forest, SVM, Accuracy, Confusion Matrix, Classification Report.

Contents

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION AND LITERATURE SURVEY

Currency is a medium of exchange, each country has its own currency which is legally sanctioned by the respective Government. A Currency circulates within a country either for purchasing or selling the goods. Each currency has its own area of acceptance which means currency of one country can't be accepted in another country. Few examples of various currencies are : Indian Rupee, US Dollar, British Pound, Japanese Yen.

Due to the advancements in printing and scanning technology, currency counterfeiting continues to be a growing concern across nations. The cultivation and circulation of forged currency notes is a heinous crime. Banks and governments work hard every day to stop fraudsters, but some are more clever than others. This illegal activity of producing and using forged currencies can be termed as "Economic Terrorism".

1.1 MOTIVATION

The primary aim is to identify the forged currency notes using the Machine Learning Algorithms, which will automatically eradicate issues such as destabilization of sanctioned currency, artificial inflation and most vital issue that is Terror Funding.

Circulation of Forged currency notes badly affect the economy of a country.

These causes explained below :

- Value of real money gets deflated.
- Inflation surges top of its sky.
- Highly dangerous implication of it is increase in terror funding.
- Destabilizes the harmony among the individuals.

Being a classification problem supervised machine learning algorithm(s) are being used for providing a robust solution to solve this problem.

1.2 OVERVIEW

In India, In the last year, the quantity of fake 500 rupee notes has climbed by 31% **[1]**according to the (NCRB) I.e. National Crime Records Bureau. In 2019, forged currency worth around ₹25.3 crore was seized. In April, 2021, fake currency worth ₹1.8 crore seized in Kochi**[2]** and ₹26 lakh was recovered in Assam**[3]**. In the year of 2020 highest ever Fake Indian Currency Notes (FICN) seized worth Rs 92.17 crore, according to National Crime Records Bureau (NCRB)**[4]**.

Looking throughout the world, Mexican peso, British pound, Euros are the most counterfeited currencies in the world**[5]**.

Due to this illegal practice of Forged currencies, the value of real money gets affected, Inflation grows steeply. And, the biggest menace is terror funding which destabilizes the security of a nation**[6]**.

1.2.1 IMPLICATIONS OF FORGED CURRENCY

- Irregular & unwanted surge in supply side of currency.
- Increase in Inflation.
- Reduction in value of real money.
- Cultivation of highly dangerous activity such as
    - Money Laundering.
    - Terror Funding.
- Destabilizes the socio and economic balance in the society.

1.2.2 LITERATURE SURVEY

A literature survey is a summary of previously published works on a certain subject. It is intended to offer the author and the audience a broad picture of what has already been published well about the topic being discussed.

1. **"Review on Detection of Fake Currency using Image processing Techniques"**

    - The problem of counterfeit notes has impacted many countries and India also suffers the same.
    - Printing the currency with forged features has now become easier with the enhanced technology.
    - Discussed results of techniques like "Calculation of Mean intensity of RGB channels, K-NN Technique, HSV image and UML".
    - Keywords: Fake Currency, Detection Methods using Image Processing

2. **"Fake currency detection using Image processing"**

    - Fake currency has become the most important problem in the market.
    - As the technology is growing, the production of fake currency has been surged, resulting into the degradation of country's economy.
    - In this paper, the proposed method utilizes OpenCV for recognizing the note as fake or original.
    - Keywords : Currency, Image processing, Greyscale, Edge Detection

3. **"Indian Fake Currency Detection Using Computer Vision"**

- To identify the fake notes a computer vision approach is used.
- To extract the features from a fake note ORB (Oriented FAST and Rotated BRIEF) and Brute-Force matcher approach are used.
- Keywords: Feature Extraction, ORB algorithm, Edge Detection, Computer Vision, OpenCV.

4. **"A Hybrid Fake Banknote Detection model using OCR, Face Recognition and Hough Features"**

- It has become a big issue nowadays, due enhancement of fake currency machines.
- There are many fields to overcome this issue i.e. image processing, neural network, computer vision etc.
- In this paper, the proposed methodologies are OCR(Optical Character recognition), Face Recognition and Hough transformation algorithms.
- In Face Recognition model Mean and Standard Deviation (MSD) was chosen for building model, it achieved 93.33% accuracy
- Keywords: Digital image processing, Face recognition, Hough Transformation, Optical Character Recognition

5. **"Counterfeit Currency Detection using Deep Convolutional Neural Network"**

- Recent reports show that demonetization led to all time high inflow of fake notes into banks.
- It is very important to detect these fake notes. The techniques are image preprocessing techniques.
- In this paper Deep Learning Convolutional Neural Network(CNN) model is proposed. The motive of this model is to identify the fake notes.
- The testing accuracy obtained is about 85.6%, training and the validation accuracy were 98.57% and 96.55% respectively
- Keywords: Fake currency, Deep Convolution Neural Network, demonetization.

6. **"An Efficient Technique For Detection of Fake Currency"**

- Researchers have been using the machine learning algorithms to identify the fake currency by using mathematical formulation.
- To identify the patterns of fake currency lots of methods are being used edge detection, segmentation, feature extractions.
- PCA algorithm has been used.
- Keyword: Eigenvalues, Eigenvectors, Fake Currency Matching Techniques.

7. **"Detection of Counterfeit Indian Currency Note Using Image Processing"**

- Countries are suffering from the enhancement of fake currency. So it becomes very important to identify the fake notes.
- To identify whether a note is fake or real image processing technique is used and implemented in MATLAB.
- Keywords: Canny Edge, Image Processing, Image Segmentation, Morphological Processing, Mean Intensity.

8. **"Ethiopian Banknote Recognition and Fake Detection Using Support Vector Machine"**

- It is very important task to categorize the banknotes of different nations because of the enhancement of fake notes in today's economy.
- To recognize the fake notes there are many systems that contain image processing techniques like image acquisition, image preprocessing, extract features and SVC(Support Vector Classifier which is a machine learning classification algorithm).The accuracy is 93% for fake note detection
- Keywords: Currency recognition, fake note detection, local binary pattern and support vector machine.

9. **"Analysis of Counterfeit Currency Detection Techniques for Classification Model"**

- Fake Currency is a threat and it impacts the growth. Making face currency considered to be a crime. Researchers have analyzed many already been incorporated techniques to recognize the face currency to overcome this threat.
- For this problem there are various solutions in terms of hardware techniques. Image preprocessing, machine learning methods.
- Keywords: fake currency, classification model, statistical techniques, machine learning, Logistic Regression, Linear Discriminant Analysis, feature extraction, detection techniques, currency features, image preprocessing, Indian rupees bank notes.

10. **"Bogus Currency Authorization using HSV Techniques"**

- Indian government has introduced a new Rs 500 and Rs 2000, to reduce fund illegal activity in India But still the fake notes are circulated in the society.
- The main idea is to detect these bogus notes of fake currency.
- Edge Detection techniques are used to detect the strip lines from the currency. HSV techniques are used to saturate the value of an input image.
- Keywords: Demonetization, illegal activity, fake or bogus currency, strip lines, edge detection techniques, counterfeit.

1.3 EXPECTED OUTCOME

A well-organized research has been conducted covering the previous work done for providing the novel solution to counter this menace of counterfeiting of currencies.

Majority of the previous work done focused on providing their solution based on Image Processing Technique and one paper proposed the work based on CNN. All the solutions were oriented on identifying only one or max two currency notes only.

In our Research work, this limitation is being addressed. The novelty is that the proposed solution has the ability to identify any Indian currency notes.

Our collective efforts under the guidance of Dr. Vivek Sharma helped us to write Research Paper as an expected outcome.

1.4 GANTT CHART

A Gantt chart is indeed a kind of bar chart typically portrays a project schedule. The Gantt chart was invented by Henry Gantt.

In Gantt chart, the vertical axis displays the tasks to be completed, and the horizontal axis provides the time intervals.

Constituents elements in our Gantt chart are :

- Planning
- Research
- Finalizing Dataset & Algorithm
- Implementation
- Research Paper Witting
- Report Witting

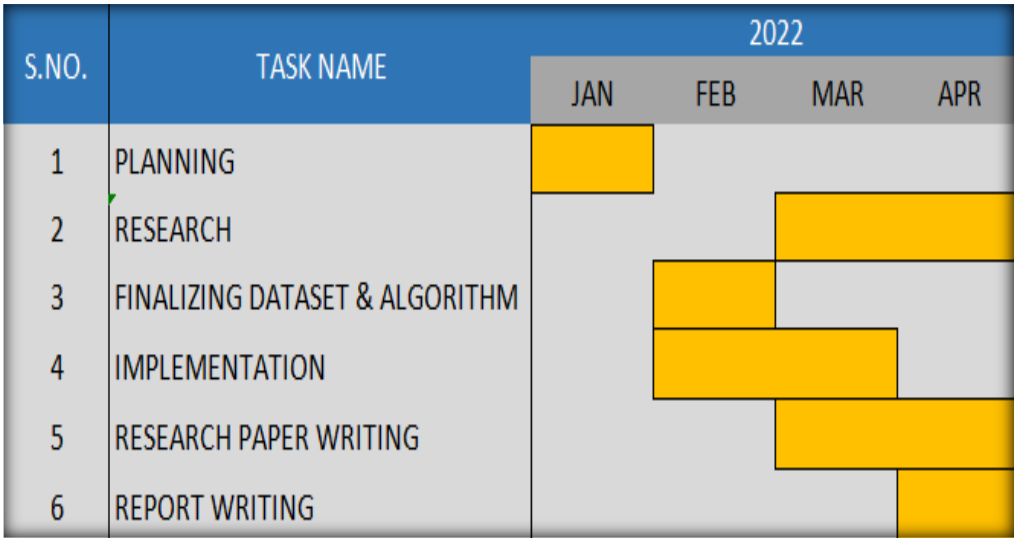

Fig i : Gantt Chart

1.4.1 PLANNING

The process of contemplating about the tasks required to accomplish a desired outcome is known as planning. A very first step to achieve any desired goal/ output.

1.4.2 RESEARCH

Research is a process of understanding and analyzing the prior work done to address the underlying problem.

1.4.3 FINALIZING DATASET & ALGORITHM

Incorporating the dataset over the algorithm helps for training the model built over that algorithm. Moreover, for testing & validation purposes dataset is segregated.

A tabular dateset consisting independent and dependent features being used. Being a classification problem, Supervised Machine Learning Algorithms are utilized.

1.4.4 IMPLEMENTATION

Execution of the plan which means practically performing the plan by incorporating the dataset, performing transformational operations over it, splitting it for training the algorithm.

1.4.5 RESEARCH PAPER WRITING

A research paper is indeed a piece of academic writing that offers very detailed study, evaluation, or interpretation of a single topic based on empirical facts.

Research paper is an art to represent the work done to other people. This helps other to understand the proposed methodology, performance metrics used, etc.

1.4.6 REPORT WRITING

A report is a comprehensive document containing all aspects of the research subject.

It includes sections such as :

- Introduction to the problem
- Motivation & Overview
- What are the existing researches
- Proposed methodology
- Performance metrics for Evaluating the techniques
- Conclusion & Future Scope

1.5 POSSIBLE RISKS.

Counterfeiting currency literally means “cultivation of fake currency notes” which is illegal. The main objective of counterfeiting currency seems to be deceive others or perhaps to cheat the system.

Major concerning risks of Forged currency notes are :

- Value of real money gets deflated.
- An artificial Inflation gets cultivated.
- Rapid Increase in Terror funding & Money-laundering, which further will cause issues such as
    - Terror Funding, destabilizes the society, which is a security threat.
    - Surge in Drug-Trafficking due to Money Laundering.
- Creation of dis-balance in the economy, as the supply side of notes will shoot up.

1.6 SPECIFICATIONS REQUIRED

The specifications that are required to execute the plan into real life by practically implementing it by utilizing some requirements.

1.6.1 HARDWARE REQUIREMENTS

Table.1. Minimum Requirements

Operating System	**Windows 10/11**
Processor	**Intel i3**
RAM	**4 GB RAM**
Hard Disk	**500 GB**
Internet Connection	**10 Mb/s**

Table.2. Recommended Requirements

Operating System	**Windows 10/11**
Processor	**Intel i5**
RAM	**8 GB RAM**
SSD	**128 GB**
Hard Disk	**1 TB**
GPU	**NVIDIA GEFORCE GTX 1050Ti**
Internet Connection	**50 Mb/s**

1.6.1 SOFTWARE REQUIREMENTS

1.6.1.1 WORKING ENVIRONMENT

Table.3. Used Working Environment

S.no	WORKING ENVIRONMENT
1.	JUPYTER NOTEBOOK

1.6.1.2 PROGRAMMING LANGUAGE

Table.4. Used Programming Language

S.no	PROGRAMMING LANGUAGE
1.	PYTHON

1.6.1.3 REQUIRED LIBRARIES

Table.5. Libraries Required

S.no	LIBRARIES
1.	PANDAS
2.	NUMPY
3.	SKLEARN

1.6.1.4 ALGORITHMS USED

Table.6. Used Algorithms

S.no	ALGORITHMS
1.	SUPPORT VECTOR MACHINE
2.	RANDOM FOREST

1.6.2 DATASET

- Tabular-data has been utilised..
- The dataset contains 1372 records.
- Features are extracted using Wavelet Transform Tool.
- Dataset contains 4 features
    a. Entropy
    b. Kurtosis
    c. Variance
    d. Skewness

CHAPTER 2: METHODOLOGY

2.1 SYSTEM VIEW

A view defining the systematic execution of the system. It is composed of several blocks having unique functionality, to maintain the efficiency of the system work while the execution to achieve the desired goal or output.

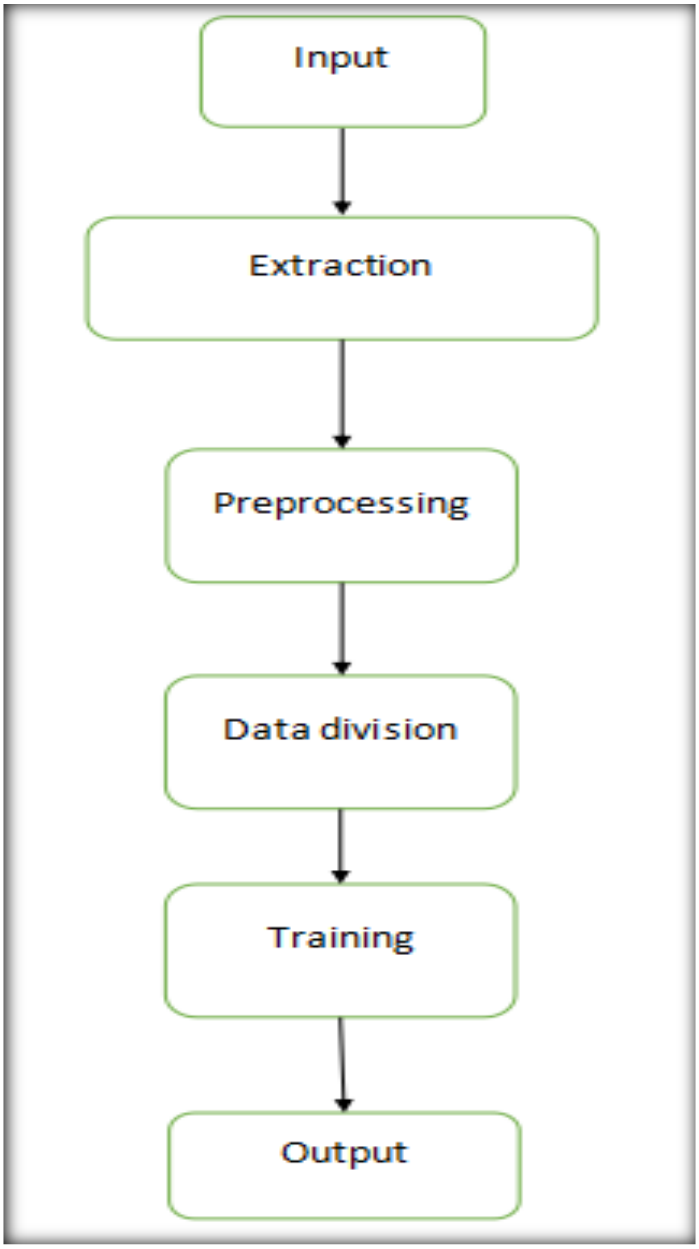

Fig ii : Pictorial form of System View

2.1.1 INPUT

An ubiquitous block in all the system is an "INPUT" block. For igniting the pipeline of system input of any information or data is a vital step.

2.1.2 EXTRACTION

It is a process of extracting the details. Post receiving the input,extraction of data from the input image is a crucial task.

2.1.3 PREPROCESSING

Preprocessing is indeed a technique for enhancing the quality of data. Utilising the exact raw data isn't considered as a best practice. Instead, preprocessing stage considered to be placed before any processing on the data.

Preprocessing includes :

- Cleansing of Data
- Integration of Data
- Transformation of Data
- Reduction of Data

2.1.3.1 CLEANSING OF DATA

Cleansing of Data can be applied over the extracted data to remove unwanted noise. Moreover, It helps to correct Data inconsistencies

2.1.3.2 INTEGRATION OF DATA

Data integration is used for merging the data from various multiple sources to make the robust dataset.

2.1.3.3 TRANSFORMATION OF DATA

Data transformation means to standardize the data by setting the value of mean to "0" and by setting the value of standard deviation to "1", the attributes in the dataset get rescaled.

The primary reason to perform standardization is to bring down all the existing features to one common scale.

2.1.3.4 REDUCTION OF DATA

Data Reduction can be used for reducing the size of data by aggregating, eliminating the redundant features.

2.1.4 DATA DIVISION

Post preprocessing the data, the splitting the data into test dataset and train dataset is a must step to be followed for Training and Testing purposes.

2.1.5 TRAINING

An algorithm needs data to get training. The training dataset gets incorporated over the algorithm, which helps to train the system.

2.1.6 OUTPUT

The last block of the system is "OUTPUT" block, post the successfully trained system is now ready to provide the desired output.

2.2 SYSTEM COMPONENTS & FUNCTIONALITIES

2.2.1 INITIALIZATION

Collecting the data is considered as an input, is a very first step to kickoff the operation of the system. Getting an image of a Indian currency note is an input, will further go into the next steps of the system's pipeline.

2.2.2 FEATURE EXTRACTION

An image is composed of square boxes called pixels. Pixel values indeed depict the brightness or intensity of the pixel.

A machine stores an image in a matrix form which is of numbered-form. A Black & White image get stored in 2-Dimensional Matrix, these images range between 0 to 255 pixel values, where 0 means black & 255 means white. In the case of coloured images, these are stored in 3-Dimensional Matrix, RGB channels are the main reason for it.

Using NumPy, Extraction of the values of an image and convert them into an array so that machine can understand the image now in the form of array(composed of the pixel values).

With the help of NumPy, Extraction of features such as :

- Entropy
- Kurtosis
- Skewness
- Variance

are possible.

2.2.2.1 ENTROPY

Entropy of an image depicts the degree of randomness An image is converted to grayscale to caculate the entropy of an image. Entropy value is used as it provides better comparison of the image details

```
from scipy.stats import entropy
```

Fig iii : Importing entropy

2.2.2.2 KURTOSIS

Measure of Tailedness is known as Kurtosis. It is a measure of whether the data is flat or peaked relative to the normal distribution.

```
from scipy.stats import kurtosis
```

Fig iv : Importing Kurtosis

2.2.2.3 SKEWNESS

Skewness is a measure of symmetry. A negative skewness means negatively skewed histogram. In Negative skewness left tail is longer. And, In Positive skewness, it is just the opposite

```
from scipy.stats import skew
```

Fig v : Importing Skweness

2.2.2.4 VARIANCE

Variance, tells the spread of the pixel values. To measure variability variance is used. More spread in the data means larger the variance and vice-versa.

```
np.var
```

Fig vi : Using Variance

2.2.3 PREPROCESSING

The process of preparing raw data so that it could be used in a machine learning model. A vital step before splitting the dataset. Generally, Real-World data consists noise, missing values, unscaled data, that can't be directly used to build models based on different algorithms of machine learning

Converting an RGB image into Greyscale image can be termed as a part of pre-processing.
Scaling the data using Standard-Scalar library in python, scales all the features of the dataset, where the value of mean is 0 and value of standard deviation is 1.

Fig vii : Importing Standard Scaler

2.2.4 SPLITTING DATASET

Post preprocessing the data, splitting of the dataset is the next main step. The preprocessed dataset gets divided into Training & Testing dataset.

Fig viii : Importing train test split

Examples for train-test split of the Dataset :

- Train: 85%, Test: 25%
- Train: 75%, Test: 25%
- Train: 70%, Test: 30%

Table.7. Splitting the Data

Training	Testing
80%	20%

2.2.5 TRAINING OF MODEL

Training becomes the crucial tasks now, after splitting the dataset. The Train Dataset incorporated over the machine learning algorithm, so that the model based on the algorithm can be trained.

Algorithms used are :

- Random Forest
- Support Vector Machine

2.2.6 CLASSIFICATION

Supervised Machine Learning Algorithm can be classified into :

- Regression
- Classification

If the Dependent attribute or Output variable contains “Continuous Values”, then it is termed as a Regression Problem.

But when the Dependent attribute or Output variable contains “Categorical Values”, then it is termed as a Classification Problem.

In the dataset, the Dependent attribute named “Class” contains two types of values

- 0
- 1

Where 0 means Forged Currency note, 1 means Real currency note

The values are categorical, Hence it is a classification problem,

Being a classification problem, the last block I.e. “OUTPUT” functionality is to classify whether the provided currency note is

- Forged (0) ,or
- Real (1)

2.3 DATA

- The used dataset is a public dataset**.**
- The data set contains information about the features of Indian currency note with class.

- The features are extracted using Wavelet Transform Tool.
- The total number of records are 1372 and there are 5 features present in the dataset.

S.NO	Features	Type	Datatype
1	*Variance*	*Independent feature*	*Numeric*
2	*Skewness*	*Independent feature*	*Numeric*
3	*Curtosis*	*Independent feature*	*Numeric*
4	*Entropy*	*Independent feature*	*Numeric*
5	*Class*	*Dependent feature*	*Boolean*

Fig ix : About Dataset

CHAPTER 3: DESIGN CRITERIA

3.1 SYSTEM DESIGN

A process of defining the elements of a system such as :

- Architecture
- Components

Is known as System Design

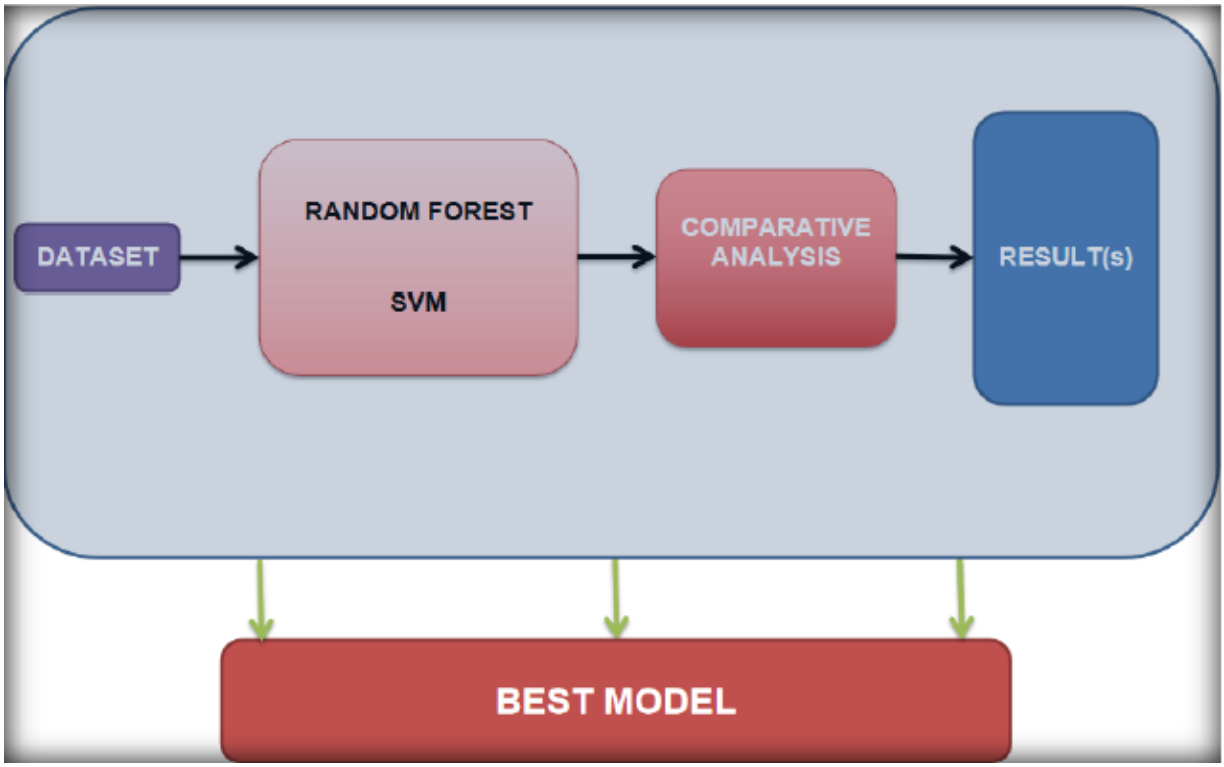

Fig x : Design of System

The Proposed System Design is composed of 4 blocks :

- Dataset
- Algorithms
- Comparative Analysis
- Result

3.1.1 DATASET

It is the 1st step towards the implementation of system design. The Dataset contains 5 features :

- Variance
- Entropy
- Kurtosis
- Skewness
- Class

The Independent attributes are :

- Variance
- Entropy
- Kurtosis
- Skewness

The Dependent attribute is :

- Class

3.1.2 ALGORITHMS

To identify whether the currency not is Real or Forged, Classifiers are required.
Since it a classification problem, Supervised Machine Learning Algorithm(S) are used :

- RANDOM FOREST
- SVM

3.1.2.1 RANDOM FOREST

A supervised machine learning algorithm that can be used for both regression and classification problems.

An important asepct about it is that, In the case of Classification, Random Forest has the ability to handle the dataset containing categorical variables, and in the case of Regression, Random Forest can handle the dataset consisting continuous variables.

This algorithm is developed by Leo Breiman, it is an ensemble learning based approach.

Ensemble Learning, A scheme of machine learning used for boosting the accuracy by combining multiple models and then solves the same problem.

Random sampling helps in predicting more accurate and better generalized results.

Random Forest generates a tree based structure containing several Decision Trees in the architecture, and uses Voting system for the Final output, which means the majority result from all the sub-results given by different Decision Tress is considered as the Ultimate result.

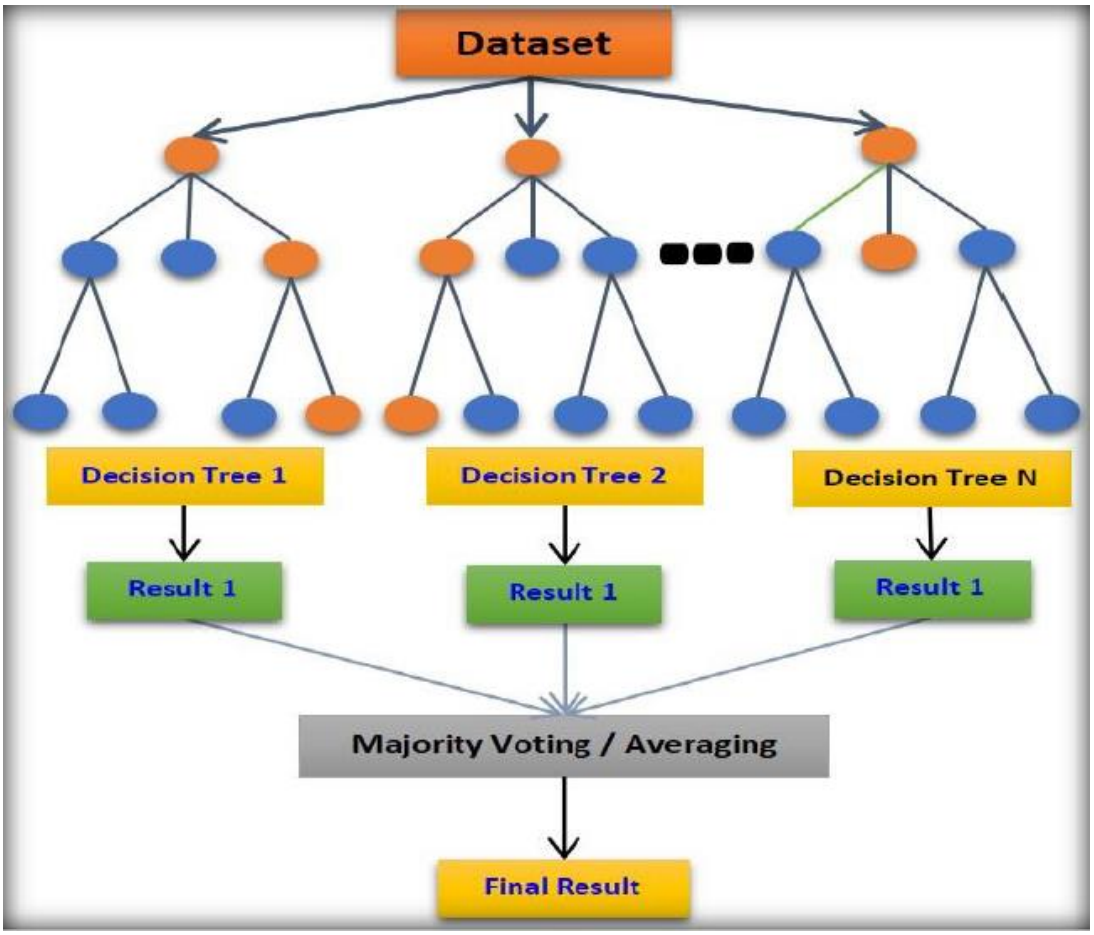

Fig xi : Depiction of Random Forest

3.1.2.2 SUPPORT VECTOR MACHINE

SVM or Support Vector Machine, A supervised machine learning algorithm that can be used for both regression and classification problems.

SVM is proposed by Vapnik, it is one of the most well-known Supervised Machine Learning Algorithm. It is having the concept of non-linear kernels to establish decision boundary for the data I.e. that is nonlinear in nature.

Mostly, for the classification problems SVM-C is used. This algorithm generates decision boundary that separates n-dimensional space into classes/groups, that distinctly classifies the data points. Hyper plane gets created by using the extreme points/vectors.

These extreme points/vectors are known as support vector, based on which this algorithm is named as Support Vector Machine.

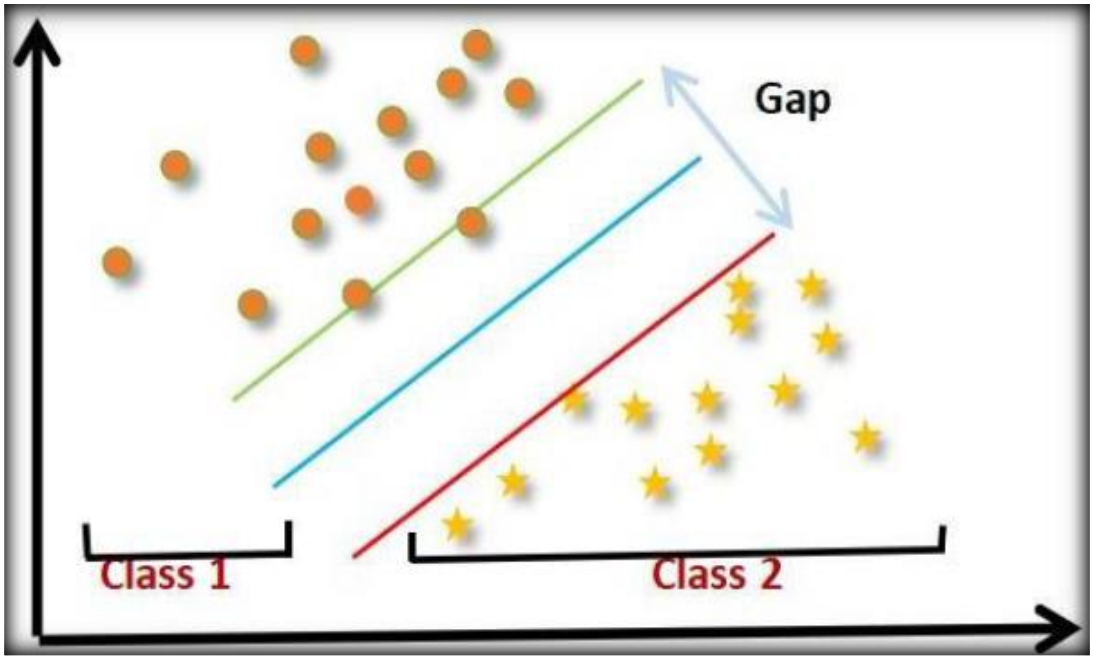

Fig xii : Depiction of SVM

3.1.3 COMPARATIVE ANALYSIS:

In comparative analysis the performance of the used algorithms are being compared based on the performance metrics.

Performance Metrics used are :

- Accuracy
- Confusion matrix
- Classification report

3.1.4 RESULT

Based on the performance of the algorithms, which are being compared on the above mentioned metrics, the best algorithm will be taken out.

CHAPTER 4: DEVELOPMENT & IMPLEMENTATION

4.1 DEVELOPMENTAL FEASIBILITY

To execute the built system efficiently, feasibility of specifications are the priority for the development of the proposed system.

An **Operating System** hence become the basic and the most vital requirement. It is used to provide the user a feasibility to harvest the capability of Hardware using software.

Processor on the other hand manages the various operations in an efficient manner.

Random-Access-Memory (RAM), is the short term memory of computer, which handles the ongoing applications and tasks. It helps to execute the ongoing instruction at a high speed that helps to reduce the time taken to completer the operation.

SSD & HDD are storage devices

SSD is :

- Fast
- Have short lifespan
- Expensive
- Non-mechanical
- Shock Resistant

HDD is :

- Slower than SSD
- Have longer lifespan
- Moderate Cost
- Have moving parts
- It is Fragile

To incorporate the proposed work, working environment is the requirement

- **WORKING ENVIRONMENT** is Jupyter Notebook

PROGRAMMING LANGUAGE, It is a language that is designed to communicate desired instructions to machine.

- **Python** has been utilised as programming language. Guido Van Rossum is considered as the father of Python.

PYTHON is :

- High Level Language.
- Interpreted Language
- Rich Libraries support
- Object Oriented
- Strong community base
- Simple to use & understand

LIBRARIES used are :

- **PANDAS**
  - Built on top of the Python programming language.
  - Flexible
  - Powerful
  - Open source
  - Easy to use data analysis and manipulation tool
  - Large community support

```
import pandas as pd
```

Fig xiii : Importing pandas

- **NUMPY**
  - In 2005, Travis Oliphant created NumPy .
  - NumPy means Numerical Python
  - Very fast computation, as most of the parts are written in C++ or C.
  - Used for working with Arrays
  - Open Source
  - Large community support

Fig xiv : Importing numpy

- **SKLEARN**
  - David Cournapeau developed it.
  - Mostly written in Python
  - For better performance, some of Sklearn's core algorithms are written in Cython
  - Built upon SciPy, NumPy and Matplotlib.
  - For machine learning, Scikit-learn(or Sklearn) is the most useful library..
  - Includes a number of useful machine learning and statistical modelling techniques, such as classification and regression.
  - Robust community support
  - Pre-installed Datasets also available

Fig xv : Importing sklearn

- **ALGORITHMS used are :**
  - **SUPPORT VECTOR MACHINE**
  - **RANDOM FOREST**

Table.8. Requirements

Operating System	**Windows 10/11**	**Yes Feasible**
Processor	**Intel i5**	**Yes Feasible**
RAM	**8 GB RAM**	**Yes Feasible**
SSD	**128 GB**	**Yes Feasible**
Hard Disk	**1 TB**	**Yes Feasible**
Internet Connection	**50 Mb/s**	**Yes Feasible**
Working Environment	**Jupyter Notebook**	**Yes Feasible**
Programming Language	**Python**	**Yes Feasible**
Libraries	**Pandas**	**Yes Feasible**
	Numpy	**Yes Feasible**
	Sklearn	**Yes Feasible**
Algorithms	**Support Vector Machine**	**Yes Feasible**
	Random Forest	**Yes Feasible**

4.2 IMPLEMENTATION SPECIFICATIONS

4.2.1 HARDWARE REQUIREMENTS

Table.9. Hardware Requirements

Operating System	**Windows 10/11**
Processor	**Intel i5**
RAM	**8 GB RAM**
SSD	**128 GB**
Hard Disk	**1 TB**
Internet Connection	**50 Mb/s**

4.2.2 SOFTWARE REQUIREMENTS

4.2.2.1 WORKING ENVIRONMENT

Table.10. Jupyter Notebook as Working Environment

S.no	**WORKING ENVIRONMENT**
1.	**JUPYTER NOTEBOOK**

4.2.2.2 PROGRAMMING LANGUAGE

Table.11. Python as Programming Language

S.no	**PROGRAMMING LANGUAGE**
1.	**PYTHON**

4.2.2.3 REQUIRED LIBRARIES

Table.12. Used Libraries

S.no	LIBRARIES
1.	PANDAS
2.	NUMPY
3.	SKLEARN

4.2.2.3 REQUIRED ALGORITHMS

Table.13. Used Algorithms

S.no	ALGORITHMS
1.	RANDOM FOREST
2.	SUPPORT VECTOR MACHINE

CHAPTER 5 : IMPLEMENTATION AND RESULT

5.1 PERFORMANCE METRICS

Post Implementing the algorithms, by exposing the dataset to them, It is critical to determine the robust and effective algorithm among them.

Performance metrics are the key factors to evaluate the algorithms.

Performance metrics used are :

- Accuracy
- Confusion Matrix
- Classification Report

5.1.1 ACCURACY

- Accuracy is one of the evaluating metric for the classification based models.
- It can also be defined as “No. of predictions that are correct out of the Total no. of predictions.”

- **Equation :**

$$\frac{(TP + TN)}{(TP + TN + FP + FN)} \qquad (5.1)$$

Where,

TP : True Positive

TN : True Negative

FP : False Positive

FN : False Negative

```
from sklearn.metrics import accuracy_score
```

Fig xvi : Importing Accuracy Score

5.1.2 CONFUSION MATRIX

- The other name of Confusion Matrix is Error Matrix.
- It is used to visualize the performance of an algorithm.
- Performance evaluator of Classification Problem.
- A table having 4 blocks, combination of Actual Values and Predicted Values.
- Moreover, Precision, Sensitivity, Recall & Specificity can be calculated

		Actual Values		
		Positive	Negative	
Predicted Values	Positive	TP	FP	Precision
	Negative	FN	TN	
		Sensitivity	Specificity	Accuracy

Fig xvii : Example of Confusion Matrix

5.1.2.1 TRUE POSITIVE

The class which is actually true, and model also predicts true

Example : “The student is actually Pass, and the model also predicted Pass”.

5.1.2.2 TRUE NEGATIVE

The class which is actually false, also model also predicts false

Example : “The student is actually Fail, and the model also predicted Fail”.

5.1.2.3 FALSE POSITIVE

The class which is actually false, but the model predicts true

Example : “The student is actually Fail, but the model predicted Pass”.

5.1.2.4 FALSE NEGATIVE

The class which is actually true, but the model predicts false

Example : “The student is actually Pass, but the model predicted Fail”.

5.1.3 CLASSIFICATION REPORT

Classification Report is a performance evaluating metrics for Classification Problem.

It is report in a tabular format consisting of :

- Precision
- Recall
- F1 Score

5.1.3.1 PRECISION

- It can be defined as the ratio of True Positive(TP) to the summation of True Positive(TP) and False Positive(FP).
- **Equation :**

$$\frac{\mathbf{TP}}{\mathbf{(TP + FP)}} \qquad \mathbf{(5.2)}$$

5.1.3.2 RECALL

- It can be defined as the ratio of True Positive(TP) to the summation of True Positive(TP) and False Negative(FN).
- **Equation :**

$$\frac{\mathbf{TP}}{\mathbf{(TP + FN)}} \qquad \mathbf{(5.3)}$$

5.1.3.3 F1 SCORE

- F1Score is the weighted harmonic mean of recall and precision.
- The closer the F1 score is to 1.0, the better the model's expected performance will be.

- **Equation :**

$$\frac{2*(\text{Recall} * \text{Precision})}{(\text{Precision}+\text{Recall})} \qquad (5.4)$$

5.2 RESULT

After completing the performance evaluation based on the above defined performance metrics.
The best algorithm among Random Forest & SVM can be picked out here.

5.2.1 ACCURACY

SVM achieved 99.63% accuracy which completely outscores Random Forest, which has achieved 98.54%.

Random Forest	SVM
98.54 %	99.63 %

Fig xviii : Accuracy of RF & SVM

5.2.2 CONFUSION MATRIX

This section is depicting the confusion matrix of both Random Forest & SVM in sequential order.

5.2.2.1 RANDOM FOREST

		Actual Values		
		Positive	Negative	
Predicted Values	Positive	*155*	*2*	*Precision = 98.7*
	Negative	*2*	*116*	
		Sensitivity= 98.7	*Specificity= 98.3*	*Accuracy = 98.5*

Fig xix : Confusion Matrix of Random Forest

5.2.2.2 SVM

		Actual Values		
		Positive	Negative	
Predicted Values	Positive	156	1	Precision = 99.3
	Negative	0	118	
		Sensitivity= 100	Specificity= 99.1	Accuracy = 99.6

Fig xx : Confusion Matrix of SVM

5.2.3 CLASSIFICATION REPORT

This section is depicting the classification report of Random Forest & SVM in sequential order.

5.2.3.1 RANDOM FOREST

	Precision	Recall	F1-score	Support
0	0.99	0.99	0.99	157
1	0.98	0.98	0.98	118
Accuracy			0.99	275
Macroavg	0.99	0.99	0.99	275
Weighted-avg	0.99	0.99	0.99	275

Fig xxi : Classification Report of Random Forest

5.2.3.2 SVM

	Precision	Recall	F1-score	Support
0	1.00	0.99	1.00	157
1	0.99	1.00	1.00	118
Accuracy			1.00	275
Macroavg	1.00	1.00	1.00	275
Weighted-avg	1.00	1.00	1.00	275

Fig xxii : Classification Report of SVM

CHAPTER 6 : CONCLUSION & FUTURE IMPROVEMENTS

Currency is legally sanctioned by the government, which we use it in our daily routine as a medium of exchange either for selling or purchasing the goods and services. Due to enhanced technology, the creation and circulation of forged currency notes have been increased, leading to artificial inflation and deflation of the real value of money.

Moreover, it is used for terror-financing, which results into disturbance in a society. To overcome such adverse effects of Forged currency, Random Forest and SVM (Supervised)Machine Learning Algorithms are used to detect the fake currency notes based using its features.

The performance metrics to evaluate Random Forest and SVM are Accuracy, Confusion Metrics and Classification Report. The results show that SVM outscored Random Forest by achieving an accuracy of 99.63%.

6.1 PERFORMANCE ESTIMATION

The performances of both the used algorithm(s) are quite impressive. The results clearly articulate that SVM by achieving 99.63% of accuracy has outscored Random Forest, which has achieved 98.54%.

6.2 USABILITY OF SYSTEM

The research that has been conducted in this paper provides the novel solution to identify Forged currency note.

Bank(s), Govt. Agencies can utilise this solution which will provide an assistance to them for identifying the fake notes.

Moreover, a common citizen can utilize this proposed solution.

6.3 FUTURE IMPROVEMENTS

A more complex & large records consisting dataset can be used in future improve our proposed methodology by making it more robust and hence provide us better results.

Moreover, incorporating Deep Learning based models over the complex dataset (non linear in nature) will help to improve the fake currency detection system’s robustness and effectiveness.

REFERENCES

[1] https://www.businessinsider.in/finance/banks/news/the-number-of-fake-500-notes-in-india-increased-by-31- in-the-last-one-year/articleshow/83030750.cms.

[2] https://timesofindia.indiatimes.com/city/kochi/kochi-police-ats-seize-rs1-8-crore-fake-currency-from- coimbatore/articleshow/82194907.cms

https://www.eastmojo.com/assam/2021/04/13/assam-fake-currency-notes-worth-rs-26-lakh-seized-in-chabua/

[4] https://www.newindianexpress.com/nation/2021/sep/17/fake-note-circulation-touched-record-high-in-2020- 2359895.html

[5] https://www.nbarizona.com/blog/tips-to-avoid-counterfeit-currencies-/

[6] https://www.fatf-gafi.org/publications/methodsandtrends/documents/money-laundering-terrorist-financing- counterfeit-currency.html

[7] S. V. Viraktamath, K. Tallur, R. Bhadavankar and Vidya, "Review on Detection of Fake Currency using Image processing Techniques," 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 2021, pp. 865-870, doi: 10.1109/ICICCS51141.2021.9432111.

[8] L. Latha, B. Raajshree and D. Nivetha, "Fake currency detection using Image processing," 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), 2021, pp. 1-5, doi: 10.1109/ICAECA52838.2021.9675592.

[9] D. Kumar and S. Chauhan, "Indian fake Currency Detection using computer vision", International Research Journal of Engineering and Technology, vol. 7, no. 5, pp. 2870-2874, 2020.

[10] A. Zarin and J. Uddin, "A Hybrid Fake Banknote Detection Model using OCR, Face Recognition and Hough Features," 2019 Cybersecurity and Cyberforensics Conference (CCC), 2019, pp. 91-95, doi: 10.1109/CCC.2019.000-3.

[11] K. Kamble, A. Bhansali, P. Satalgaonkar and S. Alagundgi, "Counterfeit Currency Detection using Deep Convolutional Neural Network," 2019 IEEE Pune Section International Conference (PuneCon), 2019, pp. 1-4, doi: 10.1109/PuneCon46936.2019.9105683.

[12] V. Saxena and Snehlata, "An Efficient Technique for Detection of Fake Currency", International Journal of Recent Technology and Engineering, vol. 8, no. 3, pp. 1298-1305, 2019.

[13] V. Sharan and A. Kaur, "Detection of Counterfeit Indian Currency Note using Image Processing", International Journal of Engineering and Advanced Technology, vol. 9, no. 1, pp. 2440-2447, 2019.

[14] E. Ayalew Tessfaw, B. Ramani and T. Kebede Bahiru, "Ethiopian Banknote Recognition and Fake Detection Using Support Vector Machine," 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), 2018, pp. 1354-1359, doi: 10.1109/ICICCT.2018.8473013.

[15] A. Upadhyaya, V. Shokeen and G. Srivastava, "Analysis of Counterfeit Currency Detection Techniques for Classification Model," 2018 4th International Conference on Computing Communication and Automation (ICCCA), 2018, pp. 1-6, doi: 10.1109/CCAA.2018.8777704.

[16] P. Ponishjino, K. Antony, S. Kumar and S. JebaKumar, "Bogus currency authorization using HSV techniques," 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), 2017, pp. 179-183, doi: 10.1109/ICECA.2017.8203667.

[17] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi and S. Homayouni, "Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 6308-6325, 2020, doi: 10.1109/JSTARS.2020.3026724.

[18] Qi, Y. (2012). Random forest for bioinformatics. In Ensemble machine learning (pp. 307-323). Springer, Boston, MA.

[19] F. Alamdar, F. Sheykh Mohammadi and A. Amiri, "Twin Bounded Weighted Relaxed Support Vector Machines," in IEEE Access, vol. 7, pp. 22260-22275, 2019, doi: 10.1109/ACCESS.2019.2897891.

[20] https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms- 934a444fca47

[21] https://www.kaggle.com/ritesaluja/bank-note-authentication-uci-data/version/1.