A HYBRID TECHNIQUE FOR FAKE CURRENCY DETECTION

Thesis submitted in partial fulfilment of the requirements for the award of degree of

Master of Engineering
in
Computer Science and Engineering

Submitted By
Anmol Sharan Singh
(Roll No. 801732002)

Under the supervision of:
Ms. Rajanpreet Kaur Chahal
Lecturer

COMPUTER SCIENCE AND ENGINEERING DEPARTMENT
THAPAR INSTITUTE OF ENGINEERING AND TECHNOLOGY
PATIALA – 147004
JULY 2019

CERTIFICATE

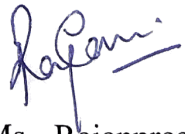
I hereby certify that the work which is being presented in the thesis entitled, **"*Hybrid Technique for Fake Currency Detection*"**, in partial fulfillment of the requirements for the award of degree of Master of Engineering in *Computer Science and Engineering* submitted in Computer Science and Engineering Department of Thapar Institute of Engineering and Technology, Patiala, is an authentic record of my own work carried out under the supervision of *Ms. Rajanpreet Kaur Chahal* and refers other researcher's work which are duly listed in the reference section.

The matter presented in the thesis has not been submitted for award of any other degree of this or any other University.

Signature:

Anmol Sharan Singh

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge.



Ms. Rajanpreet Kaur Chahal

Lecturer,

TIET, Patiala

ACKNOWLEDGEMENTS

The successful completion of any task would be incomplete without acknowledging the people who made it all possible for me and whose wise guidance and encouragement help me securing success.

At first i would like to acknowledge the omnipotent God who gave me the strength and courage to overcome all obstacles and showed me the silver lining in the dark clouds. With all the gratitude and heartiest regard, I express my sincerest feeling of indebtedness to my guide **Ms. Rajanpreet Kaur Chahal,** Lecturer, Computer Science and Engineering Department, Thapar University Patiala for their always positive attitude, invaluable cooperation, keen interest, constant encouragement and above all their blessings. I am grateful to **Dr. Maninder Singh,** Head of the Computer Science and Engineering Department. He has been a source of inspiration for the completion of this thesis. I would also like to express my gratitude to **Dr.S.S.Bhatia,** Senior Professor and Dean of academic affairs in the university for making provision of sate of the art infrastructure such as library, computer labs with high speed internet facility and valuable hardware.

Last but not the least I would like to thank my loving parents without their love and support i would have not completed my thesis. I am grateful for their support.

Anmol Sharan Singh

ABSTRACT

Indian currency consists in many different forms such as coins, banknotes etc. Fake or counterfeit currency notes are a major problem worldwide and India is also effected by the counterfeit banknotes. The Indian government is well aware of this threat and has started taking counter effective measures. Recently demonetisation of 1000 and 500 Rupee currency notes have been done and the 1000 Rupee banknote was replaced by 2000 Rupee banknote. The newly released banknotes are much more safe and hard to reproduce as compared to their old versions. Every denomination of Indian banknotes such as 10,100, 200, 500 etc has been renewed with better security features to insure more security but after all these measures fake Indian currency notes are being produced both locally and across the border, which is a serious problem as it helps in deteriorating economy of the our country. In this paper our main aim is to create a hybrid approach consisting of Digital Image Processing, Feature Extraction of Indian currency notes. Then clustering is performed followed by classification and comparing them with fake examples. Which will help us in detecting the fake Indian currency notes and hence further stopping the circulation of these counterfeits in our country.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Image Processing

Processing of image consists of performing various different things on a Digital Image, in order to extract some useful information from it or to get an enhanced version of image. It is basically a signal processing type in which an input image is taken as input and we can get various different features and characteristics from that input image or an image as our desired output. Popularity of image processing is increasing as it is more commonly used in various different technological fields. It is a unique and important area of research for budding scholars in field of computer science. Image processing can be done by performing these three important steps:

- Importing images with the help of various different image acquisitioning tools.
- Manipulating and analyzing the image to get desired features from the image
- Processing the output which consist of a different image or report of the image analysis.

Processing of image consist of two different methods namely, analogue and Digital Image Processing. In analogue image processing various different hard copies such as photographs, notes and printouts can be evaluated by image analysts by using fundamental techniques of interpretation and various visual techniques. On the other hand digital image processing is done using different techniques that helps in manipulating the digital images by using various different computers softwares. Digital image processing requires various different phases for all types of data which given as input, these three phases are image pre-processing, image enhancement, and displaying of information which is extracted. Nowadays Digital Image Processing is used more as compared to analogue image processing which is done at a smaller scale. Digital revolution has favoured Digital Image Processing over analogue image processing.

1.1.1 Digital Image Processing

Digital Image Processing is basically processing of a digital image with the help of computerised tools and other softwares. In other words we can also say that Digital Image Processing uses various different computer algorithms on a given image in order to get an enhanced image or extracted useful features and information which can be further used for research purposes. Image is the main object which is used in Digital Image Processing. Image is basically defined as a two dimensional function F(p,q), where p and q of the function F are just spatial coordinates, where amplitude of F at a given pair of coordinates (p,q) is known as intensity of the given image at a given point and when these amplitude values of F and spatial coordinates (p,q) are finite then it is known as a digital image.

So basically an image is defined by a two dimensional array arranged in a manner that it consist of rows and columns and should be finite. Digital image generally consist of finite elements and each of those elements have a unique value on a given location. These unique element's are known as pictoral elements, pixels, features and image elements. The word "Pixel" is very much used to express the various elements of a Digital Image.

Various types of image

- **Binary Image:** This type of image as its name tells, contains only two different pixel elements which are 0 and 1, where 1 refers to white part and 0 refers to black. It is also called Monochrome image.
- **Black and White Image:** It is the image which only consist of white and black colour and that's why it is known as Black and White Image.

- **8bit Colour Format:** It's also known as Grayscale Image format. It is the most popular image format used in image processing. It consist of 256 different shades of colours in it. In this format 0 is for black colour, 127 is for gray color and 255 is for white color.

- **16bit Colour Format:** It's also popularly known as High Color format. In this colour image format there are 65,536 different colours present. The color distribution is not same as Grayscale Image format. The High Color format is mainly divided into three different color format's namely Red, Green and Blue hence it is also famously known as RGB color format.

1.1.2 Phases in Digital Image Processing

- **Image Acquisition:** In this phase the given Digital Image is accquired with help of various sensors. Then scaling is performed on the acquired digital image. It is then followed by a color conversion such as RGB to Grayscale or vice versa.
- **Image Enhancement:** In this phase of Digital Image Processing, hidden details from the given picture is extracted and these can be subjective. It is one of the most basic and most happening areas of Image Processing.
- **Image Restoration:** Image restoration works on the mathematical model, degradation of the given Digital image or probabilistic model. It also deals with appealing of a given image but it is generally objective.
- **Colour Image Processing:** It is basically linked with psuedocolour Image Processing and full colour Image Processing. In this phase colour models are applied to Digital Image Processing.
- **Multi-resolution and Wavelets Processing**: It basically represents the Digital Images in various degrees or wavelets and consist of Multi Scale Approximation (MSA) of the given images which is further used for designing method for many different practically relevant Discrete Wavelet Transforms (DWT) and other techniques.

- **Image Compression:** It is generally the process of converting or encoding a given image in such a way that it consumes less space than its original image. Various different image compression techniques are used which reduces the size of an image file without degrading the quality of the image to a greater extent.
- **Morphological Processing:** It consist of dealing with various different tools that are used for extracting components of the given image, these components are used for representation & description of shapes. It uses various Image Processing functions and processes various images based on their shapes.
- **Object Detection and Recognition:** In this phase real world objects are searched in the given image, by performing feature extraction and learning algorithms. It also consist of algorithm that assigns a unique label to an object which is based on its descriptor which has been generated after Feature Extraction.

1.2 Feature Extraction

Feature Extraction is an essential part of Image Pre-Processing, Pattern Recognition and Machine Learning. In extraction of features an initial step is performed in which measured data is taken and its features are derived from the data, this derived data is informative and non-redundant, it facilitates the subsequent generalisation of steps and learning and in some cases it is even better than the human interpretations. Dimensionality reduction is also related to extraction of features. So, if the given input data from a dataset to a algorithm is very huge to be processed by the algorithm and it is also suspected to be unnecessary e.g: similarity in measurement of both meters, feet, or repetitiveness of pixels present in an image, then it is changed into a smaller set of features which are called feature vector. Discovering a simple subset of initial features is known as feature selection, the newly selected features can contain unique and admissible information or characteristics from the given data. Now with these reduced representations of the initial complete data and now the desired tasks can be done efficiently without using the complete initial data[1].

Feature Extraction basically consists of decreasing the number of resources required to describe a given large dataset. The huge number of variables involved from a huge dataset causes many different problems while analysing the complex data. It requires large amount of memory space and computational power to analyse a large number of variables, also it can cause overfiting of a classification algorithm on training samples and new samples can be generalised very poorly. Feature Extraction is also known as a method of generating combinations of the variables so that we can get around these problems and also describing the given input data with efficiency and accuracy. Many different Machine Learning data scientists think that proper optimisation of extraction of feature is required for effective construction of model[2].

One of the very useful area of application of Feature Extraction is that of Digital Image Processing. Various different techniques are used to extract and to find unique portions, shapes or features of a given digital video or image. Optical Character Recognition is the most highly sorted field for application of Feature Extraction.

1.2.1 Applications of Feature Extraction

- **Edge detection:** Extraction of various features is extensively used in edge detection and further extraction, the main aim of Edge Detection is to identify the points of a given image in which image brightness changed very sharply and where there are discontinuities. These points are termed edges and are extracted by using various feature extraction techniques.
- **Blob detection:** Feature Extraction is also used in Blob detection and extraction, Blob detection aims at detecting the various regions of a digitalised input image that have various different properties, such as colour and brightness which are then compared to all the surrounding areas of the image. Blob is basically a region or area of a digital image which consist of constant properties and these points are extracted using Feature Extraction techniques.

- **Ridge detection:** Feature Extraction is used in ridge detection, interior of elongated objects in the image domain are detected in this and after computer vision analysis are extracted using extraction of various features.
- **Scale-invariant feature transform:** In this Feature Extraction method it plays a major role as the whole algorithm is basically a feature extraction technique in which the given image is blurred using gaussian blur function followed by LoG (Laplacian of Gaussian) approximations, then key features of image or keypoints are detected and these features are then extracted a descriptor vector is formed after successfully performing feature extraction. It is one of the most famous feature extraction techniques used in image processing and mainly used for fake currency detection.

1.3 Speeded Up Robust Features (SURF)

SURF or speeded up robust features is one of the most used local feature detecting and descriptor technique. It is mainly used in various tasks like classification, object recognition, 3-D reconstruction or image registration. It is mainly based on the scale-invariant feature transform SIFT descriptor. In the normal version of SURF technique, it is many times faster than SIFT and is more efficient against various different image transforms as compared to SIFT.

In order to obtain the various interest points the SURF method uses an integer based approximation technique in which the determinant of Hessian matrix and a Blob detector is used which is then computed with 3 different integer operations using a single integral precomputed image. The sum which is obtained of the Haar wavelet response around a given point of interest mainly comprises SURF's feature descriptor. Integral image can also be used in order to compute these descriptors. SURF descriptors are used to recognise and locate objects, to reconstruct 3D scenes, recognise people or faces or to track various objects and also to extract various different points of interest. It was first

presented by Herbert Bay in 2006 European Conference on Computer Vision. An application of this algorithm has since been patented in the United States of America[3]. A different version of SURF is known as U-SURF, is very much faster than the normal version in computing and is better suited for applications, where the camera is more or less horizontal as its is not invariant to image rotation. The given Digital Image is transformed into its coordinates and then a multiple resolution pyramid technique is used to copy the original input image with Laplacian or Gaussian pyramid shape and in order to obtain an image which is generally the same size but has a lesser bandwidth. Then we get a blurring effect on the input image known as scale space as it ensures that the various points of interest remain scale invariant. The SURF technique is based similar principals as SIFT but each step is a little bit different. SURF has mainly three important steps which are interest point detection followed by local neighbourhood description and descriptor generation, after that feature or descriptor matching is done.

1.3.1 Interest point detection

In SURF approximation of Gaussian smoothing is done by using square shaped filters. Generally the integral image is being used and then filtering of the given image with a square can be done at a faster rate. A sum of input Digital Image is taken within a rectangle and is calculated swiftly by using the integral image and this is done only if evaluations are performed on all four corners of the given rectangle. It basically uses a Blob detecting mechanism which is further dependant on a matrix to spot various points of interest and the determinant is calculated of the Hessian Matrix and is used for measuring the change around the local areas of a point and then the various points are chosen where the value of determinant is maximum. As compared to the Hessian Laplacian detector used by Mikolajczyk and Schmid, SURF also uses the determinant of the Hessian matrix for choosing the scale, which was also done by Lindeberg. At different scales many different interest points can be found, because the searching for the correspondences always requires different images for comparison so that they can be

viewed at different scales. So in various other feature detection techniques, image pyramid is usually realised as the scale space. Images are smoothed by applying the Gaussian filter repeatedly, these are sub sampled in order to get to the higher level of the pyramid. So several floors are calculated with different measures of the mask.

The given scale space is then divided into different octaves, where the series of response maps which cover a doubling of the scale are referred as octave. The least level of the scale space is obtained as the output of the 9×9 filters and is also known as initial scale layer. Hence, unlike old methods, in SURF scale spaces are implemented by applying different box filters having various sizes and the scale spaces are analysed by scaling up the size of the filter instead of step by step reducing the given size of the image. The scale space interpolation is very important in as the difference in scales between the initial layers of every octave is generally very big.

1.3.2 Descriptor generation and Matching

The main aim of descriptor is to give robust, efficient, description of the extracted image features for example at any given point of interest it can be described as the distribution of pixels within the neighbourhood of that point. Descriptors are thus normally computed locally, so we always obtain a unique description of every point of interest which was identified previously. Dimensionality of the generated descriptor is directly proportional to the impact on the point-matching accuracy and computational complexity. A small descriptor can give too many false positives as it can be more sturdy against variations, but it will not offer proper discrimination.

The initial steps basically consists of fixing a given regenerable orientation which is obtained on the basis of information from the point of interest's surrounded circular region. Then a square shaped region is created in order to extract the SURF descriptors from it and align the selected orientation. The Haar Wavelet can be used to find the

proper orientation of the interest points and hence the rotational invariance is achieved. After finding the desired descriptors from a set of images then we can compare and match them, hence matching pairs of descriptors can be found.

1.4 Cluster Analysis

Clustering or Cluster analysis is known as grouping of given set containing various objects in such a way these objects having a similar group or a cluster are near to each other as compared to the other groups or cluster's. It is one of the most important task in the field exploratory mining of data and is also common technique of statistical analysis of data. It is used in different areas of expertise such as machine learning and data mining, information retrieval, recognition of patterns, data compression, bioinformatics, computer graphics and image analysis. Clustering analysis does not consist of one single algorithm but it consist of a task that is solved. Clustering is easily done by using various different techniques that can be different to each other significantly in their own definition of way of creating the clusters and how efficiently it can be found but will form efficient clusters of similar data. The word cluster cannot be easily defined that is why there are so many unique and different Clustering algorithms[4].

The definition of cluster is different in different algorithms and its properties also vary significantly. There are many different types of clustering models present. One of the few famous model is the centroid model, in this the cluster is represented by a single mean vector e.g. K-means Clustering. Connectivity models are also very popular such as hierarchal clustering in which clusters are created on the basis of distance connectivity. there can be some overlapping in these two, but there are also some higher distinction clustering in which every object belong to one cluster only.

1.4.1 K-means Clustering

K-means Clustering is basically a type of vector quantization method which originated from signal processing, it is mainly used in data mining for cluster analysis. K-means Clustering aims at partition of given observations into k cluster's in which each observation is a member of that cluster with which it has the nearest mean value, it serves as a prototype for the cluster. Voronoi cells are created as a result of separation of the given data space. k-means problem is categorised as computationally difficult which means it comes under the category of NP-Hard but, new and efficient heuristic algorithms connect swiftly to a local optimum and these are generally same as the expectation maximisation algorithms, which are used for Gaussian distributions by using a unique iterative refinement approach which is employed with the help of both Gaussian mixture modeling and k-means. These two techniques mainly use the cluster centres in order to model the data, but k-means clustering basically finds clusters of similar spatial extent efficiently, while different shapes of cluster can be allowed by the Expectation Maximisation Mechanism. K-means has a very loose relationship to the k-nearest neighbor classifier which is a popular Machine Learning technique used for Classification and is often confused with k-means due to similar names. For feature learning k-means clustering has been efficiently used, its also used as a dictionary learning step in both semi-supervised learning and unsupervised learning[5]. The approach behind this is that first a k-means based cluster representation is trained by using the given input data which can be both labeled or non-labeled. Then, the input datum is projected into a feature space which is new and it is followed by an encoding function, like thresholded product of the matrix of the given data with its locations of centroid is used and then distance from the datum to each centroid is computed, and acts as a simple indicating function of the closest computed centroid,[5][6] or a smooth transformation of the distance[7]. Also the transformation of the given sample cluster distance via a Gaussian RBF, gives us the radial basis function network's hidden layer[8].

1.5 Bag of Words Model

Bag of Words Model is normally implemented in information retrieval, Natural Language Processing. It is just a simplified representation of data. In Bag of Words Model, a text or image is represented as the bag of visual words with no regard to grammar, word order and feature placement but keeps the multiplicity intact. Bag of Visual Words Model is implemented in computer vision and can also be implemented in Classification of Images. In this feature of image are treated as word or vocabulary. It is also used in classification of documents. Bag of words model basically creates a sparse vector based on occurrences on count of words. In other words it consist of a sparse histogram of the generated vocabulary. It is also used in computer vision, the number of times a word has occurred in a given vocabulary of local image features is stored in th form of vector and is known as bag of visual words.

1.5.1 Feature Selection

After the process of detection of features, the given image is evaluated with numerous local patches. These patches are represented numerical vectors which are generally known as feature descriptors and it is the main task of different feature representation methods is to represent the patches properly. An efficient descriptor must always have full power to handle the intensity, rotation variations, scale variations and various other variations at a given situation. SIFT is a very famous descriptor[9]. SIFT transforms each patch to vector. After this there is no importance of the order of different vectors.

1.5.2 Codebook Generation

It is the last step in the Bag of Visual Words Model its main function is to convert the represented vector patches into codewords such as words, which therefore further helps in producing a codebook which is a type of dictionary of words. Now several similar features or patches can be considered as a representative of any given codeword. One of the most simplest technique is to perform cluster analysis over all the vectors preferably k-means[10]. The various centres of the newly learned clusters are defined as codewords of the codebook. The number of different clusters represent is the size of the generated codebook . Now each part of the given input image is a unique codeword.

1.6 Support Vector Machine

Support vector machine or SVM is very famous among supervised learning models in the field of machine learning, it is associated with learning algorithms which analyse the given information or data and is further used in regression and classification analysis. SVM basically constructs a set of different hyperplanes, in a high dimensional or infinite space, which can be used for classification purposes, regression or other tasks like outliers detection. If a training set of examples, is given in which the data is marked and belonging to one or another categories of the given two, then a SVM technique basically generates a model which assigns newer examples to various different categories therefor it is non probabilistic and binary and linear classifier. SVM model basically is a rendition of many points in the space, these points are mapped in such a way that there is a clear gap between unique examples which belong to separate different categories, this gap is made as big as possible. Now newer examples being are mapped onto this same space therefore depending on which side of gap they fall their category is predicted. SVM's does not only perform linear classification it also efficiently performs non-linear classification. When unlabelled data is given, Supervised Learning cannot be performed so Unsupervised Learning approach is required. Unsupervised Learning basically tries to

look for natural clustering of given data into various groups and it is followed by new data mapping into the formed groups.

1.6.1 Applications of SVM

SVM is a boundary value analysis technique and can be implemented successfully to solve real world problems such as:

- Text and Hypertext Categorization can done with the help of SVM, as application of SVM can very much reduce the need for a labeled training data instances in transductive and normal inductive settings. Some methods and techniques for shallow semantic parsing are based on SVMs[11].
- Image Classification is also performed using SVM. Various experimental analysis and results show that higher search accuracy is achieved while using support vector machine models as compared to other old query refinement schemes which use four to six rounds of feedback. It is also true for Image Segmentation systems, as the systems which include modified versions of SVM use the privileged approach which were suggested by Vapnik[12][13].
- Handwriting Recognition and hand written characters can be easily recognised by using support vector machine[14].
- SVM is also used in fake currency detection techniques using it to classify and aid help in texture analysis.
- The SVM algorithm can also be applied in various different sciences fields specially biological sciences. SVM has been used to classify different types of proteins and percentage of the compounds were classified correctly. Various Permutation tests which are based on SVM weights have been used as a mechanism for interpretation of various SVM models.[15][16]

CHAPTER 2
LITERATURE SURVEY

Yadav et al. (2011) developed a methodology to detect fake currency notes automatically. The technique basically used a feature extraction technique clubbed with HSV color space for image processing. Authors successfully implemented a fake note detection unit in MATLAB [17].

Devendran B et al. (2015) designed a system which helped in identifying the Indian currency notes and it also check's if the given note valid or fake thus help in differentiating between the real notes and fake notes. Unique currency features like see througth register, Identification mark or see through symbol, high security thread, Governor's signature, Micro-lettering, year of printing all were successfully extracted, features were extracted with the help of SIFT technique.[18]

Albert et al. (2016) used different dictionary-based approaches which were mainly: SIFT-K-SVD, BoW and SCSPM and compared them with each other to create a dictionary and after that to find the a unique sparse linear combination of atoms in the created dictionary that matches with the signal. A proper dictionary D is given, and the above mentioned methods and techniques are used to acquire a set of coefficients y which further feeds a SVM classifier. The various different atoms are evaluated and discriminated and by this way the classifier detects the among the texture images to find the counterfeits [19]. ‘

Liu and Shimamur (2014) proposed an approach in which L P analysis of a cross correlation sequence between zero crossing waves and its speech signal. The proposed approach performes speech analysis in a white and noisy environment simulation results showed that proposed approach has shown far better results as compared to existing [20].

Hassanpour et al. (2007) used various characteristics of different paper based currencies including colour, texture and size of the given currency in feature recognition. Histograms were used and many different colours in a paper currency were compared and computed with respect to a reference piece of paper currency and then Markov chain was applied to model textures of the various paper currencies randomly. This approach has been used for recognition of paper currencies from many different countries [21].

Saini et al. (2015) tested shifts in system-created documents, incorporating documents with plane and printed foundations. In this manipulation of the documents was done with the assistance of picture processing software applications including Paint and Adobe Photoshop. The documents were analyzed for any alterations and results were obtained. The results showed that may features were connected with picture manipulation [22].

Santhanam et al. (2013) proposed a method which identifies fake currency. It consist of polarisation concept, holographic detection and image processing techniques to detect fake currency on given physical properties of the given currency instead of the existing methodology based on its chemical properties. The given scheme is implemented using mathematical label, lab view simulation and experimental verification via automation in order to enhance the reliability and the accuracy in efficient manner [23].

Vishnu and Omman (2014) proposed an approach to identify features like numeral, shape, colors etc. In order to reduce dimensions a Principal component analysis (PCA) has been used and also a similarity based classifier was trained to predict the sample test and results were also supported by creating models by implementing various different classifiers. These were implemented using a software names WEKA and the testing was done using unseen samples were not considered as a method extraction of features. This result showed that 100% accuracy was shown by numeral results with all the families of currencies [24].

Jain and Vijay (2013) presented an image processing technique which extracted features from the given currency and the extracted ROI has been used along with pattern matching techniques and various neural networks. Authors used different pixel level for the denomination of different notes. The various pattern recognition techniques and the neural network matching techniques were used to match the given currency of the paper notes succesfully [25].

Sandeep Kaur et al.(2017) proposed a technique in which indian currency is taken and each currency has their unique identification irrespective matched properly. The proposed match various features of the currency note to its denomination check the note is legitimately of note. All these unique features are extracted using SIFT and matching of features is done using nearest neighbour classifier (NNC). Previously SIFT technique with nearest neighbour classifier has been used to detect fake currency. but in this proposed technique SIFT technique has been used with DWT tool. and Bayesian classifier (BC).[27]

CHAPTER 3

PROBLEM STATEMENT

Fake currency circulation has increased very rapidly in our country, everyday new cases are shown by our media, which involve people printing fake currency locally or people caught with huge amounts of fake currency notes smuggled in from other countries. Flooding of fake currency notes is very bad for our country's economy. There have been many cases reported in which people are getting "Children's Bank of India" notes and "Churan Label" notes from the ATM. These notes look very similar to the original ones but have various wrong markings and wordings that gave them away, ATM machines are getting fooled and people are notoriously depositing fake notes in the machines and causing a huge problem to other people.

Presently fake currency detection is been done efficiently, there are many different ways by which fake currency can be detected. One of the most famous method is feature extraction and detection by using various techniques such as SIFT, SURF, edge detection ,blob detection etc. These features which are extracted are then stored in different ways and are matched with the testing notes in order to check whether they are fake or not. Our main focus is to create a hybrid approach so that we can detect fake notes from the real one on the basis of feature extraction and classification which is much more safe and efficient as compared to other methods such as normal eye test. Our Indian currency notes have security features in built such as when tried to photocopy a note, the micro-lettering cannot be copied properly and see through register will completely vanish under the photo copying condition. Hence these features provide a little level of security but still fake currency notes are being generated and their detection is needed to be done at a higher rate and also it needs to be more efficient.

3.1 Research Methodology

Our proposed approach is explained in the following steps:

- A set of Indian currency notes of different denominations are taken and their digital images are stored.
- The acquired digital image is then pre processed and feature extraction is performed with the help of Speeded Up Robust Features (SURF).
- After that we get the feature descriptors from the input images.
- Then Bag of Feature approach is applied and the feature descriptors are clustered using k-means clustering and each cluster's center represents a unique feature or unique visual word as a part of vocabulary created.
- Then a multiple class support-vector machine classifier is trained by using a newly formed vocabulary and the input image.
- After that a test image is chosen from both categories real and fake currency notes and the test image is evaluated against the given classifer.
- A graphical user interface is created in MATLAB in order to select the type of denomination and to check whether the given note is real or fake and it will also sjow the accuracy also.

3.2 Research Objectives

- Fake currency is a major problem nowadays and our main objective is to propose a hybrid approach for fake currency detection.
- The proposed hybrid approach will use Feature Extraction, Bag of Visual Words model and classification techniques in order to check the authenticity of a given note.
- The model will be trained by using proper Indian currency notes and then the accuracy of the given input note will be checked with the trained model and then it will be decided that it is fake or not.

CHAPTER 4
IMPLEMENTATION

4.1 Creating DataSet

At first Indian currency notes of different denominations mainly 500 and 2000 Rupee Indian note were taken. Many different notes of the above mentioned denominations were taken and their serial numbers were noted. In order to check properly we also took photocopied versions of the above mentioned denominations as they are a type of fake currency, many different fake notes such as of children bank of India notes and churan label notes were also take as they are readily available and are also a type of fake currency. Then pictures of all the collected dataset is taken under similar lighting conditions and fixed camera settings using a iPhone X camera. These pictures are then stored for further processing.

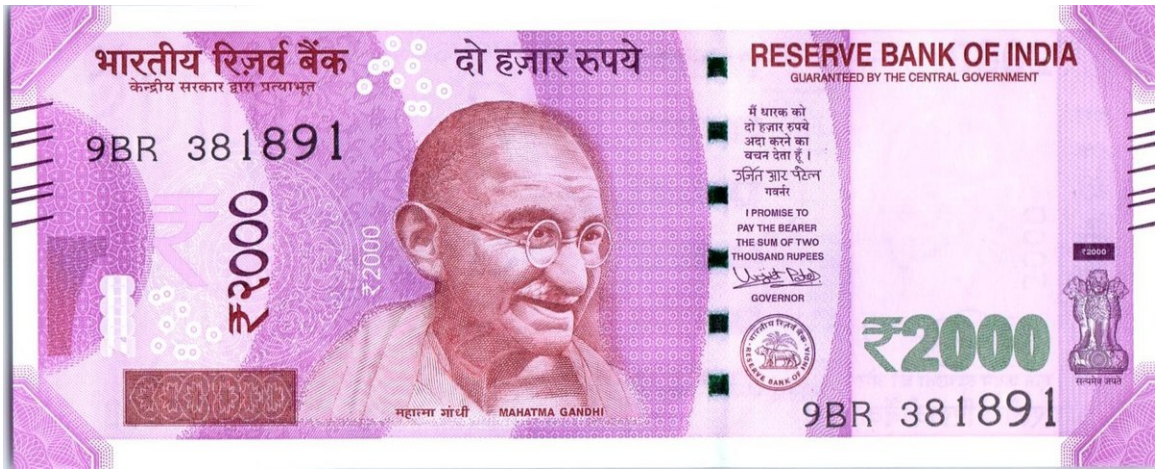


Fig 4.1 Indian 2000 Rupee note

4.2 Feature Extraction

The testing images are stored separately, the genuine currency images are stored in one folder and the fake currency images are stored in the other folder. Then images are processed and feature extraction takes place with the help of Speeded Up Robust features. It is applied in MATLAB. The feature point locations are selected using the Grid method. SURF generally consist of two different steps mainly feature detection and feature extraction. After getting the feature point locations SURF features are extracted. It's feature detector is generally based on Hessian matrix and it is also a multi scale space theory.because the Hessian matrix has good accuracy and performance. In a given image L, x = (x, y) is normally the point given and the Hessian matrix H(x, σ) in x at a given scale σ, where L_{xx} (x, σ) is also the convolution result of the given second order derivation of Gaussian filter, it can be define as

$$\mathbf{H(x, \sigma)} = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{yx}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix}$$

Fig 4.2 A Hessian Matrix

SURF basically creates a stack and sampling is done for larger levels of the pyramid therefore we get images of similar resolution. Since integral images are used so calculation of rectangular filters of box is also allowed in sustained time and after that keypoints are located and after that orientation of keypoints is done using wavelet responses, it is done in both vertical and horizontal direction and is done for a given neighbourhood size and proper gaussian weight is also applied to it and then they are plotted. A sliding window is used and sum of responses is calculated and then the orientation is found. Then the these keypoints are extracted into feature descriptors and is stored.

4.3 Bag of Words Approach

After obtaining feature descriptors a Bag of Words Approach is used and then k means clustering technique is applied at the feature descriptor which are extracted from the images using the statistical and machine learning toolbox of MATLAB[26].The k-means technique iteratively groups various extracted descriptors into *k* mutual exclusive cluster's and now the resulting cluster's are very small and are separated by same characteristics and features. Every cluster center represents a distinct feature, or unique visual word, hence a visual codebook or a dictionary is created. which can now be used to tell that how many visual words are now present in our input images and their frequency of occurrence.

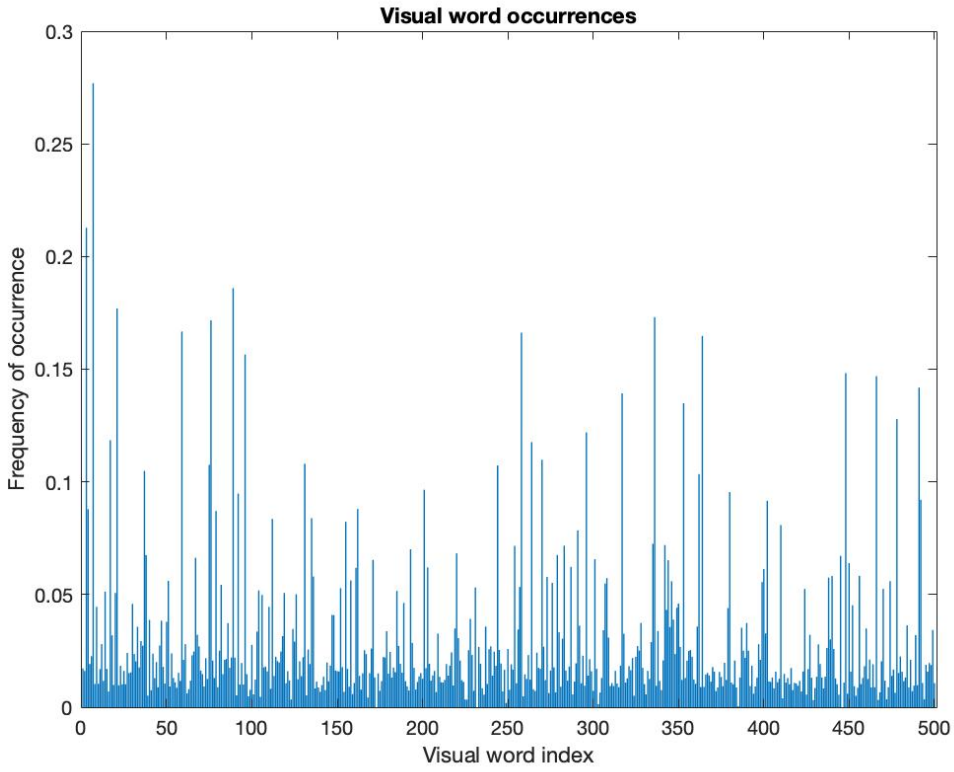


Fig 4.3 Histogram of the Visual Dictionary

4.4 Classification and Matching

After the creation of the visual dictionary or bag we now train a classifier. A binary support-vector machine classifier which is an image classifier, it takes the bag of words which is returned by the earlier steps to encode the images of the given set of images into a histogram consisting of visual words. This histogram of these words is now used to train a classifier. In the end accuracy of the predicted model is examined by using the testings images and it predicts that whether it is fake currency note or a real currency note.

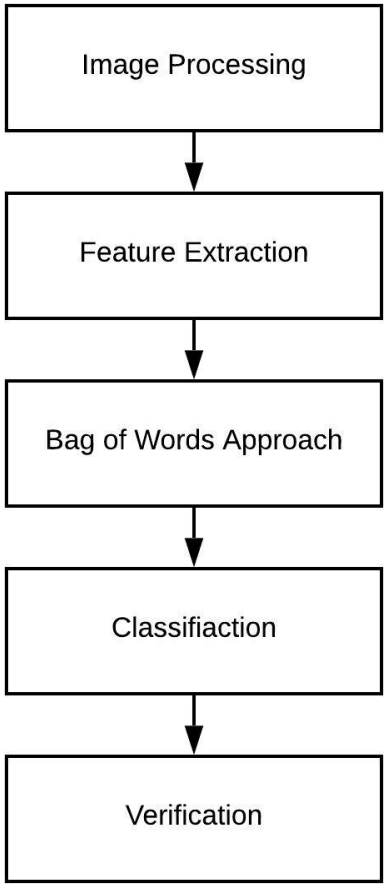


Fig 4.4 Work Flow Diagram

4.5 GUI Implementation

A Graphical User Interface is created using MATLAB guide feature, which create dynamic applications that can run in real time. A set of buttons is used which help in choosing different denominations for fake currency detection, bunch of static text places are used which provide realtime important information such as the score which are shown positive instead of negative. An image display is also placed in bottom which shows the different images of test notes which are selected. All of the above mentioned buttons, text spaces and display boxes are placed inside a main panel which comprises of the boundary of the whole GUI.

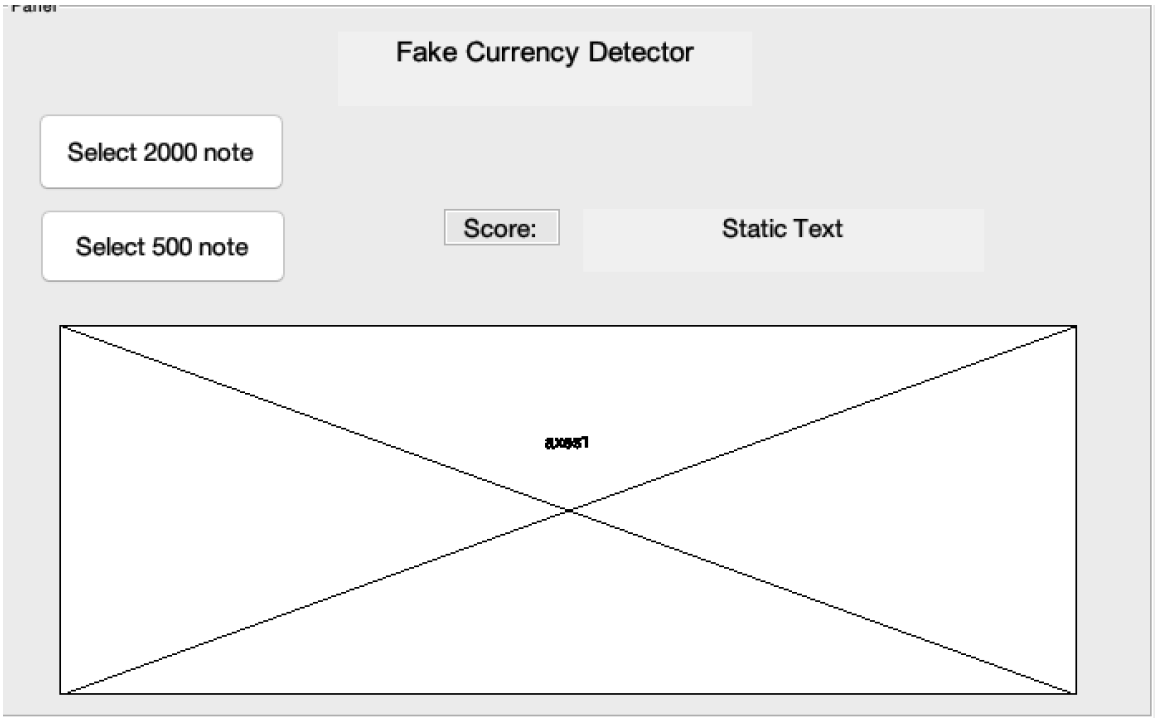


Fig 4.5 Graphical User Interface

CHAPTER 5
EXPERIMENTAL RESULTS

Many different Indian currency notes of denomination 500 and 2000 were tested and their accuracy was predicted, fake currency notes (photocopied notes and children's bank) notes were also tested and their accuracy is predicted. These were selected using the graphical user interface created using MATLAB. All the experimental results are shown below.

1. 2000 denomination (S.No:4ML926128)

Score: -.952243

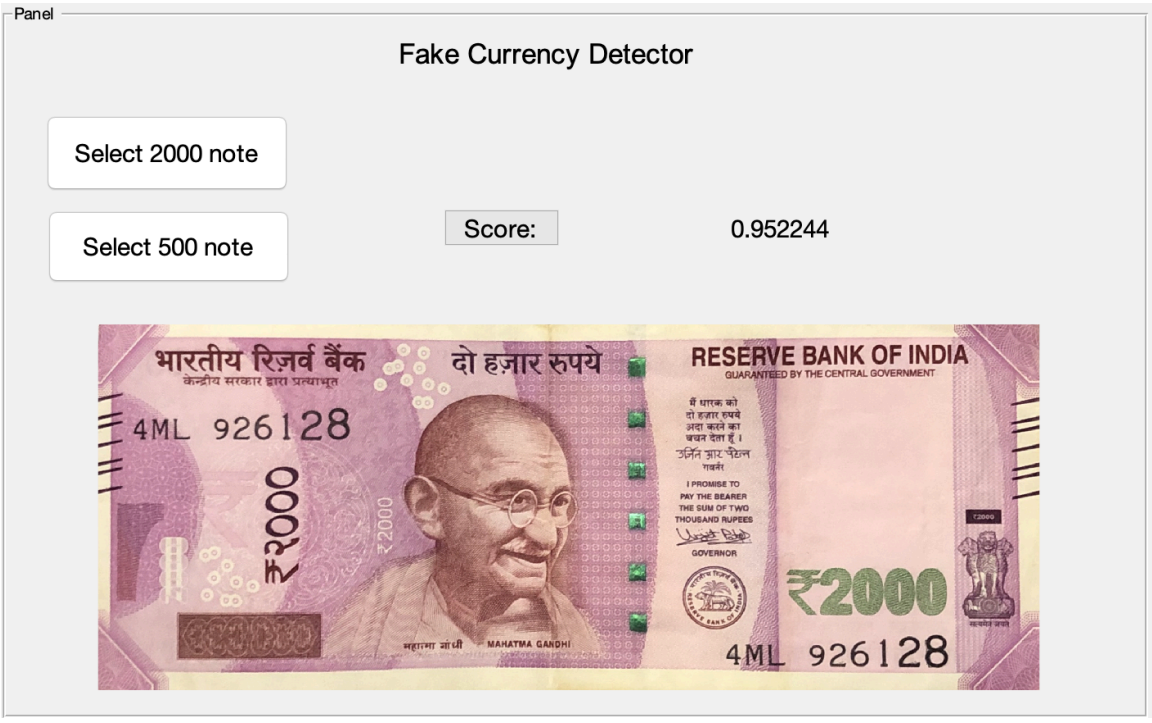


Fig 5.1(a) Analysis of 2000 rupee currency note using proposed approach

2. 2000 denomination (S.No:7BE680951)

Score: -92.8382

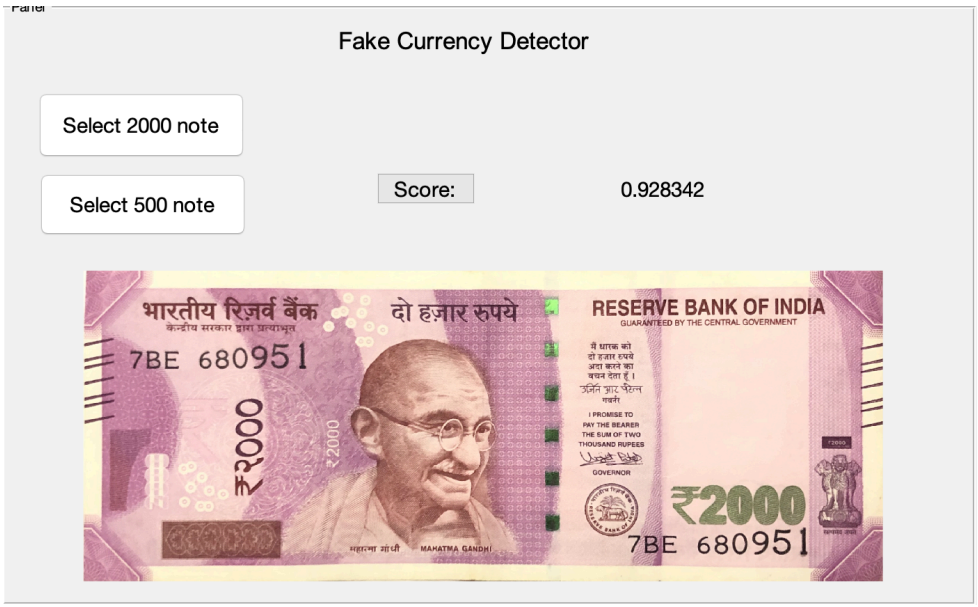


Fig 5.1(b) Analysis of 2000 rupee currency note using proposed approach

3. 2000 denomination(S.No.3EQ247262)

Score: -.902944

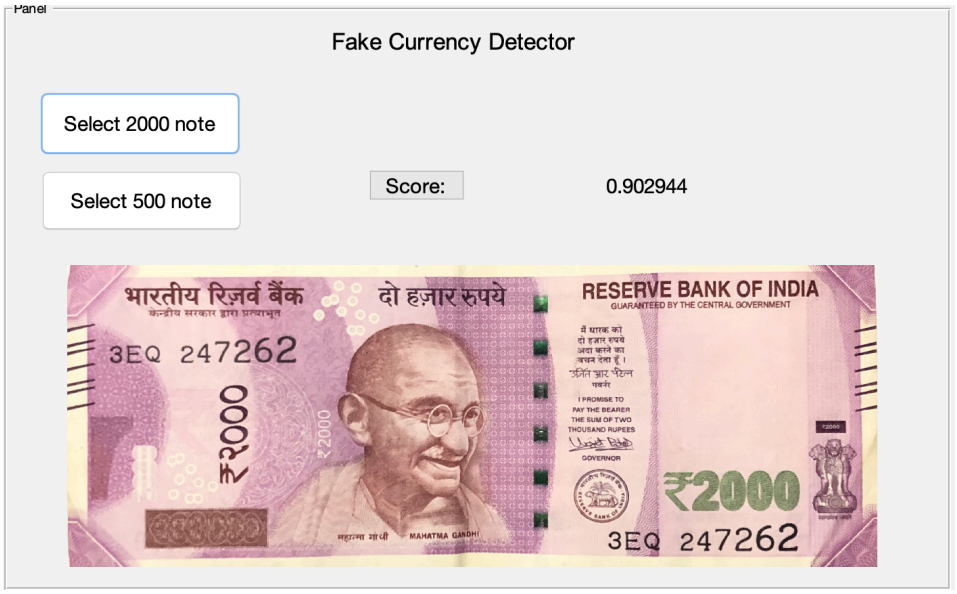


Fig 5.1(c) Analysis of 2000 rupee currency note using proposed approach

4. 2000 denomination(Photocopied S.No 0DA271418)

Score: -.515964

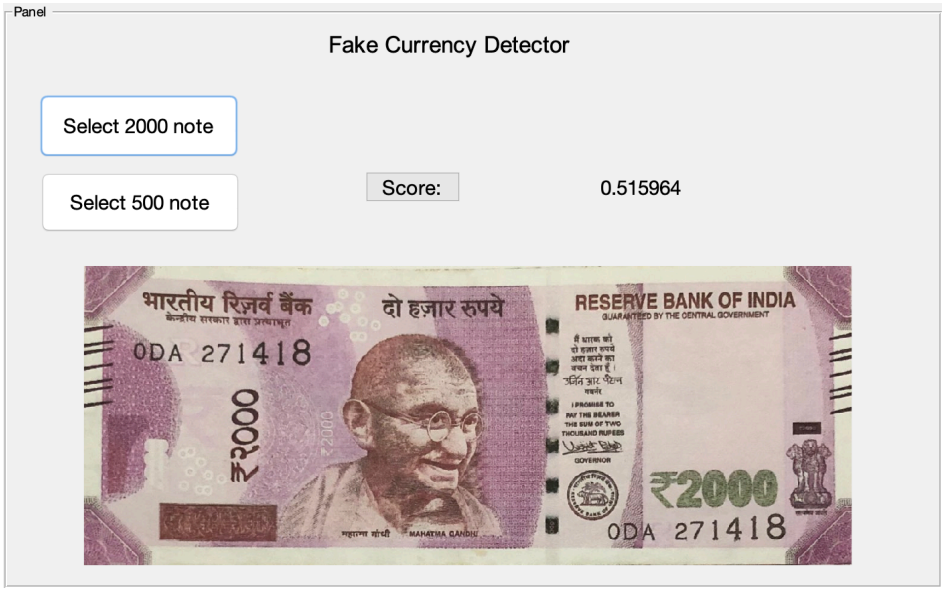


Fig 5.1(d) Analysis of 2000 rupee currency note using proposed approach

5. 2000 denomination(Photocopied S.No 4ML926128)

Score: -.388902

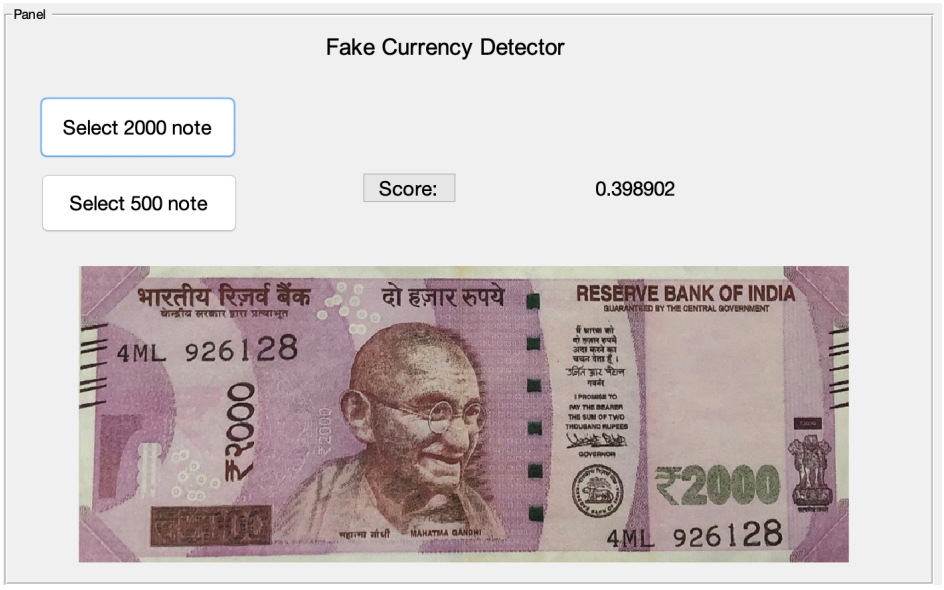


Fig 5.1(e) Analysis of 2000 rupee currency note using proposed approach

6. 2000 denomination(Children's Bank of India)

Score : -.487776

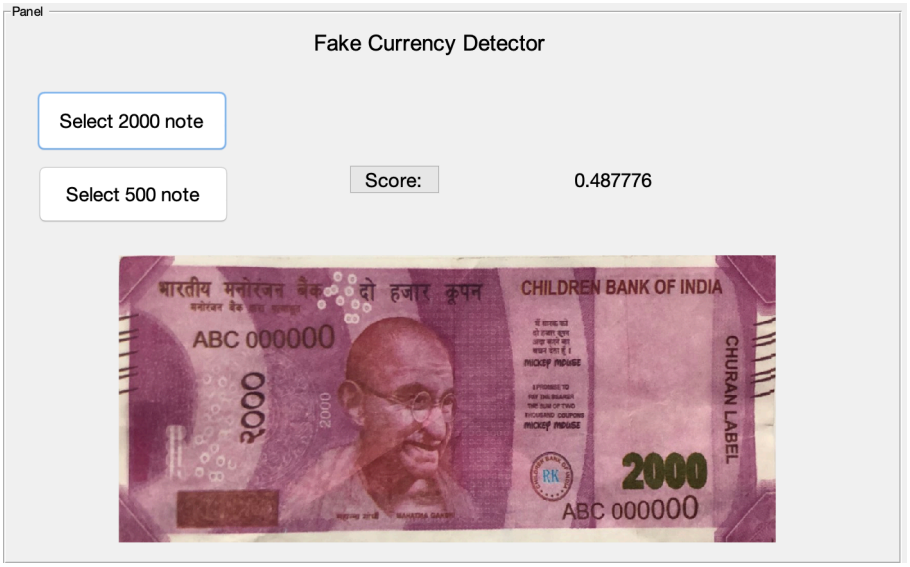


Fig 5.1(f) Analysis of 2000 rupee currency note using proposed approach

7. 500 denomination(S.No 8KU340345)

Score: -.806534

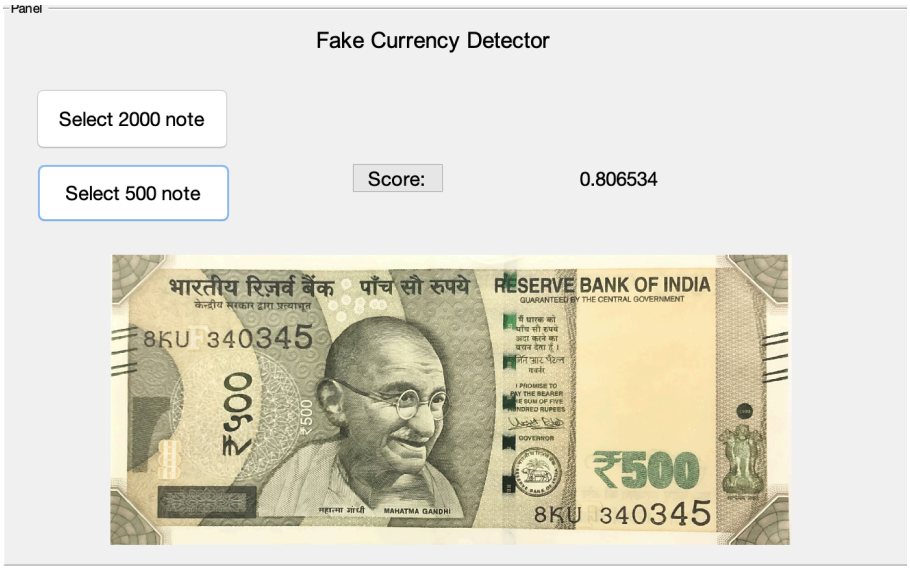


Fig 5.2(a) Analysis of 500 rupee currency note using proposed approach

8. 500 denomination(S.No 9LA*930192)

Score: -.808188

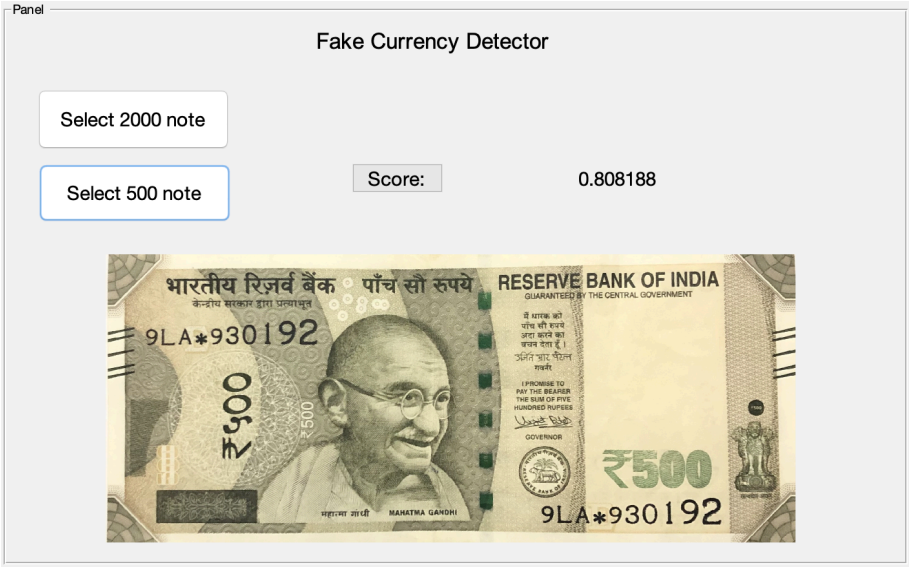


Fig 5.2(b) Analysis of 500 rupee currency note using proposed approach

9. 500 denomination(S.No 8KU340351)

Score: -.894912

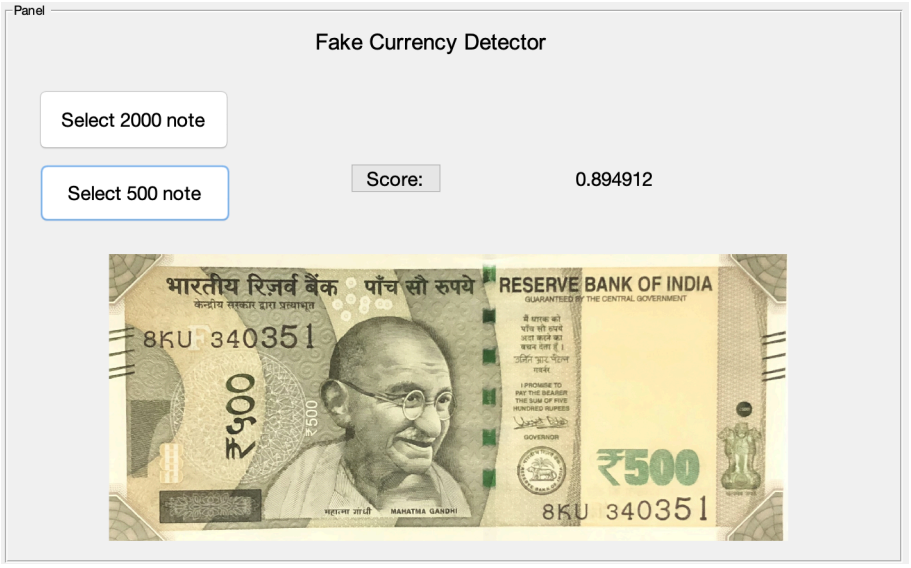


Fig 5.2(c) Analysis of 500 rupee currency note using proposed approach

10. 500 denomination(Photocopied S.No 8KU340352)

Score: -.451904

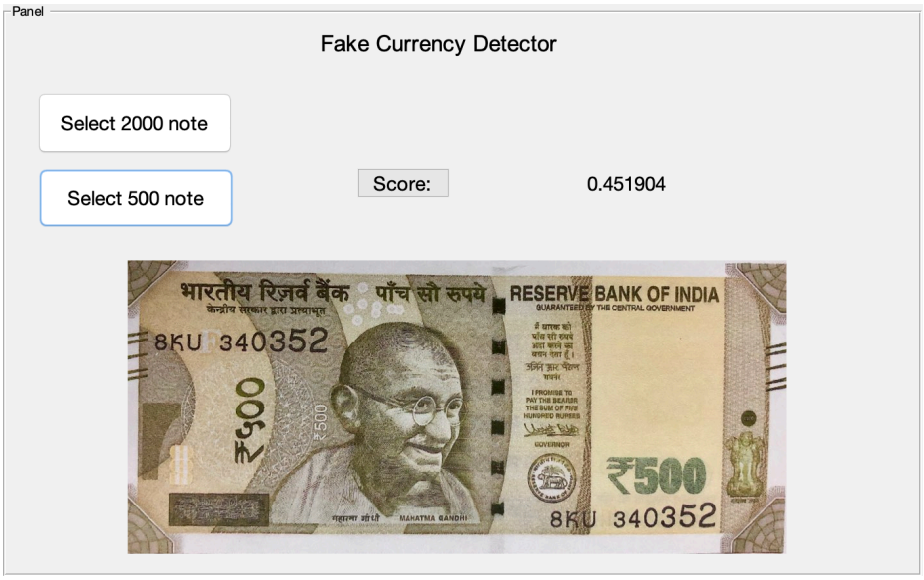


Fig 5.2(d) Analysis of 500 rupee currency note using proposed approach

11. 500 denomination(Photocopied S.No 8KU340353)

Score: -.473194

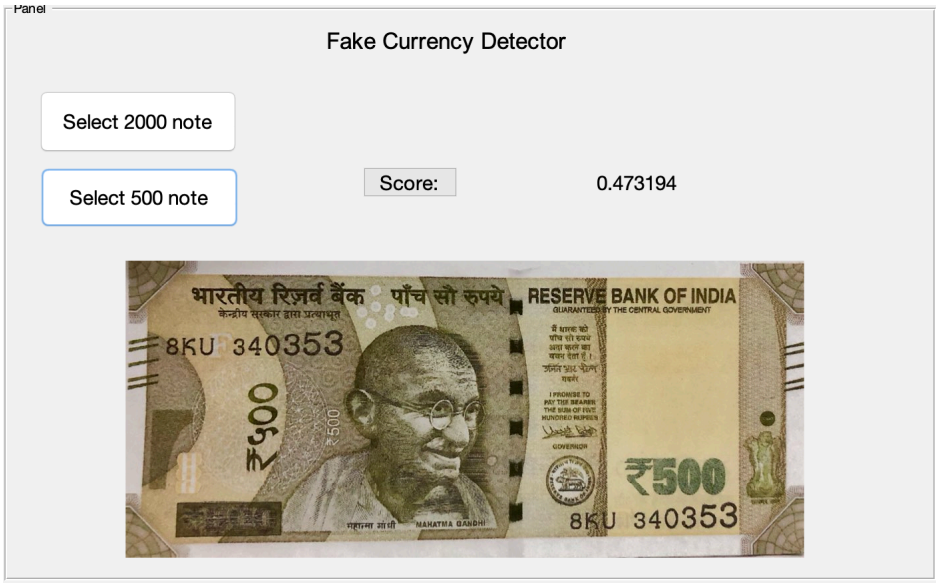


Fig 5.2(e) Analysis of 500 rupee currency note using proposed approach

CHAPTER 6
CONCLUSIONS AND FUTURE SCOPE

We have shown a hybrid approach with which fake currency notes can be detected by applying feature extraction and using classification. We found out that the score of genuine Indian Currency notes of both 2000 and 500 denomination notes differs from Fake and counterfeit notes .Finding a fake currency note can be difficult if the given real currency note is not in proper shape or in proper condition and can be marked as fake currency note, therefore the input note should be in proper condition in order to correct detection. Various different lighting conditions also play a role in fake currency detection as the images which were taken in perfect lighting condition of a genuine note were detected as real and when then same note's image was taken in bag lighting condition then it was marked as fake as it did not matched the trained example properly.

Future Scope :-

- Our proposed approach can further be improved in order to get better results from the input image of the currency notes.
- The lighting condition of the input note image also have a major part in correct detection of note which can be futher improved.
- Quality of the image is a major factor as images of genuine notes from a not so good camera can result in improper detection which can be further improved.

REFRENCES

[1].Alpaydin, Ethem *Introduction to Machine Learning*. London: The MIT Press. p. 110. (2010).

[2].Reality AI Blog, "Its all about the features," September 2017 Available - https://reality.ai/it-is-all-about-the-features/.

[3].Ryuji Funayama, Hiromichi Yanagihara, Luc Van Gool, Tinne Tuytelaars, Herbert Bay, “Robust Interest Point Detector and Descriptor” U.S.Patent US2009238460 September 2009.

[4].Estivill-Castro, Vladimir (20 June 2002). "Why so many clustering algorithms – A Position Paper". ACM SIGKDD Explorations Newsletter. 4 (1): 65–75.

[5].Coates, Adam; Ng, Andrew Y. (2012). "Learning feature representations with k-means" In Montavon, G.; Orr, G. B.; Müller, K.-R. (eds.). Neural Networks: Tricks of the Trade. Springer. 2012.

[6].Csurka, Gabriella; Dance, Christopher C.; Fan, Lixin; Willamowski, Jutta; Bray, Cédric (2004). Visual categorization with bags of keypoints (PDF). ECCV Workshop on Statistical Learning in Computer Vision. 2004.

[7].Coates, Adam; Lee, Honglak; Ng, Andrew Y. (2011). An analysis of single-layer networks in unsupervised feature learning (PDF). International Conference on Artificial Intelligence and Statistics (AISTATS). 2011.

[8].Schwenker, Friedhelm; Kestler, Hans A.; Palm, Günther. "Three learning phases for radial-basis-function networks". Neural Networks. 14 (4–5):pp. 439–458. 2001.

[9].Vidal-Naquet; Ullman "Object recognition with informative features and linear classification" International Conference on Computer Vision IEEE 2003.

[10]. T. Leung; J. Malik (2001). "Representing and recognizing the visual appearance of materials using three-dimensional textons" (PDF). International Journal of Computer Vision. 43 (1): 29–44.

[11].Pradhan, Sameer S., et al. "Shallow semantic parsing using support vector machines." Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics: HLT-NAACL 2004. 2004.

[12].Vapnik, Vladimir N.: Invited Speaker. IPMU Information Processing and Management 2014.

[13.] Barghout, Lauren. "Spatial-Taxon Information Granules as Used in Iterative Fuzzy-Decision-Making for Image Segmentation". Granular Computing and Decision-Making. Springer International Publishing, pp.285–318. 2015.

[14]. DeCoste, Dennis (2002). "Training Invariant Support Vector Machines" (PDF). Machine Learning. 46: 161–190.

[15].Gaonkar, Bilwaj; Davatzikos, Christos; "Analytic estimation of statistical significance maps for support vector machine based multi-variate image analysis and classification". 2013.

[16].Cuingnet, Rémi; Rosso, Charlotte; Chupin, Marie; Lehéricy, Stéphane; Dormont, Didier; Benali, Habib; Samson, Yves; and Colliot, Olivier; "Spatial regularization of SVM for the detection of diffusion alterations associated with stroke outcome", Medical Image Analysis, 2011, 15 (5): 729–737.

[17]. B. P. Yadav, C. S. Patil, R. R. Karhe, P. H. Patil , “An automatic recognition of fake Indian paper currency note using MATLAB”, International Journal of Engineering Science and Innovative Technology, Vol. 3, Issue 4, pp. 560-566. 2014.

[18].Kavya B R, Devendran B. “Indian Currency Detection and Denomination using SIFT”, International Journal of Science, Engineering and Technology Research (IJSETR), Volume 4, Issue 6, June 2015 .

[19]Albert Berenguel, Oriol Ramos Terrades, Josep Llado´s and Cristina Canero(2016), “Banknote counterfeit detection through background texture printing analysis” IEEE 2016.

[20]. L. Liu, T. Shimamura, “Linear prediction analysis of cross correlation sequence for voiced speech”, Signal and Information Processing Association Annual Summit and Conference, pp. 1-4. 2014.

[21]. H. Hassanpour, A. Yaseri, G. Ardeshiri (2007), “Feature extraction for paper currency recognition”, IEEE International Symposium on Signal Processing and its Applications, pp. 1-4. 2007.

[22]. K. Saini, S. Kaur, “Forensic examination of computer-manipulated documents using image processing techniques”, International Association of Law and Forensic Science, pp. 1-7. 2015.

[23]. K. Santhanam, S. Sekaran, S. Vaikundam, A. M. Kumarasamy, “Counterfeit currency detection technique using image processing, polarization principle and holographic technique”, International Conference on Computational Intelligence, Modeling and Simulation, pp. 231-235. 2013.

[24]. R. Vishnu, B. Omman, “Principal features for Indian currency recognition”, Annual IEEE India Conference (INDICON). 2014.

[25]. V. K. Jain, R. Vijay, “Indian currency denomination identification using image processing technique”, International Journal of Computer Science and Information Technologies, Vol. 4, pp. 126-128. 2013.

[26]. MATLAB Statistical and Machine Learning Toolbox Release 2017b, The MathWorks,Inc., Natrick, Massachusetts, United States.

[27].Sandeep Kaur, Seema Baghla, S Kumar .“Enhancement of Sift algorithm to check authenticity of Indian Currency”, International Journal of Computational Intelligence Research ISSN 0973-1873 Volume 13, Number 5, pp. 945-953. 2017.

APPENDIX A

PLAGIARISM REPORT

A Hybrid Technique for Fake Currency Detection

ORIGINALITY REPORT

7%	6%	4%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	en.m.wikipedia.org Internet Source	3%
2	en.wikipedia.org Internet Source	2%
3	"Palmprint recognition using SURF features", Electric Electronic and Control Engineering, 2015. Publication	1%
4	Carrillo, Justin T., Christopher T. Goodin, and Alex E. Baylot. "NIR sensitivity analysis with the VANE", Infrared Imaging Systems Design Analysis Modeling and Testing XXVII, 2016. Publication	<1%
5	research.ijcaonline.org Internet Source	<1%
6	G. Rohith Reddy, T. Sai Jahnavi, D. Rushali, B. Anil Kumar. "Bliss BOT for pharmaceutical inspection", 2017 International Conference on Trends in Electronics and Informatics (ICEI), 2017	<1%

7	H. S. Prasantha, H. L. Shashidhara, K. N. B. Murthy. "Chapter 25 Fast Computation of Image Scaling Algorithms Using Frequency Domain Approach", Springer Science and Business Media LLC, 2013 Publication	<1%
8	www.ijracse.com Internet Source	<1%
9	Manish Yadav, Swati Yadav, Dilip Sharma. "Image denoising using orthonormal wavelet transform with stein unbiased risk estimator", 2014 IEEE Students' Conference on Electrical, Electronics and Computer Science, 2014 Publication	<1%
10	www.cis.unisa.edu.au Internet Source	<1%
11	Singaraju Jyothi, Dadala Mary Mamatha. "chapter 39 Image Processing", IGI Global, 2018 Publication	<1%

APPENDIX B
LIST OF PUBLICATIONS