Dissertation on

Design and Development of Rotation Invariant Bank Note Identification System of Multiple Countries using Pattern Recognition

Thesis submitted towards partial fulfilment of the requirements for the degree of

Master of Technology in IT (Courseware Engineering)

Submitted by

Arnabjit Choudhury

CLASS ROLL NO.: 002030402015
EXAMINATION ROLL NO.: M4CWE22015
UNIVERSITY REGISTRATION NO.: 154497 of 2020-21

Under the guidance of
Prof. Ranjan Parekh
School of Education Technology
Jadavpur University

Course affiliated to
Faculty of Engineering and Technology
Jadavpur University
Kolkata-700032
India

2022

M.Tech. IT (Courseware Engineering)
Course affiliated to
Faculty of Engineering and Technology
Jadavpur University
Kolkata, India

CERTIFICATE OF RECOMMENDATION

This is to certify that the thesis entitled "Design and Development of Rotation Invariant Bank Note Identification System of Multiple Countries using Pattern Recognition" is a bonafide work carried out by Arnabjit Choudhury under our supervision and guidance for partial fulfillment of the requirements for the degree of Master of Technology in IT (Courseware Engineering) in School of Education Technology, during the academic session 2021-2022.

R. Parekh. 24/6/2022

SUPERVISOR
School of Education Technology
Jadavpur University,
Kolkata-700 032

Professor
School of Education Technology
Jadavpur University
Kolkata - 700 032

R. Parekh. 24/6/2022

DIRECTOR
School of Education Technology
Jadavpur University,
Kolkata-700 032

Director
School of Education Technology
Jadavpur University
Kolkata - 700 032

Subrata Chakraborty

DEAN - FISLM 24/06/2022
Jadavpur University,
Kolkata-700 032

Dean
Faculty of Interdisciplinary Studies, Law & Management
Jadavpur University, Kolkata-700032

M.Tech. IT (Courseware Engineering)
Course affiliated to
Faculty of Engineering and Technology
Jadavpur University
Kolkata, India

<u>CERTIFICATE OF APPROVAL</u> **

This foregoing thesis is hereby approved as a credible study of an engineering subject carried out and presented in a manner satisfactory to warranty its acceptance as a prerequisite to the degree for which it has been submitted. It is understood that by this approval the undersigned do not endorse or approve any statement made or opinion expressed or conclusion drawn therein but approve the thesis only for purpose for which it has been submitted.

Committee of final examination for evaluation of Thesis

** Only in case the thesis is approved.

DECLARATION OF ORIGINALITY AND COMPLIANCE OF ACADEMIC ETHICS

I hereby declare that this thesis contains literature survey and original research work by the undersigned candidate, as part of his **Master of Technology in IT (Courseware Engineering)** studies.

All information in this document has been obtained and presented in accordance with academic rules and ethical conduct.

I also declare that, as required by this rule and conduct, I have fully cited and referenced all materials and results that are not original to this work.

NAME: ARNABJIT CHOUDHURY

EXAMINATION ROLL NUMBER: M4CWE22015

THESIS TITLE: DESIGN AND DEVELOPMENT OF ROTATION INVARIANT BANK NOTE IDENTIFICATION SYSTEM OF MULTIPLE COUNTRIES USING PATTERN RECOGNITION

SIGNATURE: Arnabjit Choudhury DATE: 24/06/2022

Acknowledgement

I feel fortunate while presenting this dissertation at School of Education Technology, Jadavpur University, Kolkata, in the partial fulfilment of the requirement for the degree of Master of Technology in IT (Courseware Engineering).

I hereby take this opportunity to express warm thanks from the core of my heart to my guide and mentor Prof. Ranjan Parekh, Director of School of Education Technology for being so kind and supportive throughout the course period. His advice and guidance have paved the path of success of this thesis. He has encouraged me and has guided me with all possible suggestions and constructive criticism. I will remain grateful to him, for my entire life.

I would like to express my warm thanks to Prof. Matangini Chattopadhyay for her support and advices. I would also like to thank Dr. Saswati Mukherjee and Mr. Joydeep Mukherjee.

My thanks and appreciation goes to my classmates from M.Tech IT (Courseware Engineering) and MMD. I do wish to thank all the departmental support staffs and everyone else who has different contributions in this dissertation.

Finally, my special gratitude to my family who stood for me at all my odds and have supported me well enough to achieve this height.

Date: 24/06/2022

Place: KOLKATA

Arnabjit Choudhury

ARNABJIT CHOUDHURY
Class Roll No.: **002030402015**
Examination Roll No.: **M4CWE22015**
Registration No.: **154497 of 2020-21**
M.Tech IT (Courseware Engineering)
School of Education Technology
Jadavpur University
Kolkata-700032

Design and Development of Rotation Invariant Bank Note Identification System of Multiple Countries using Pattern Recognition

Contents

List of Figures

List of Tables

Executive Summary

Any kind of money that is in use or circulation as a means of trade, such as coins and banknotes, is standardised as a currency, usually issued by a government. There are 180+ currencies being used as a mode of payment. In this study, out of many cases, three of the most popular cases are being considered. The first one being identification of the country from a given banknote from both its obverse and reverse side. The second one is recognition of the denomination of the banknote. And last but not the least, is to identify the currency from any orientation. This study highlights different challenges confronted in above referred to cases and pursuits to solve them. The comparative study shows how various features like dimensions of different banknotes, colour moments, texture features, LBP, SURF, ORB, features generated from template matching and uses various classifiers to make the system efficient. It is derived that template matching is the ideal way if the dataset of currency is limited and if there are currencies present in the dataset which features are very similar to each other. A high correlation coefficient between the input image and the template image stored in the training dataset is computed to get the perfect recognition of the currency.

CHAPTER 1

INTRODUCTION

1.1 Overview

Whether or not we pull out paper bills or swipe a credit card, most of the transactions we interact in everyday use currency. Certainly, global economies depend on money to function. currency refers to banknotes or coins which might be in circulation. However, when examining the entire money supply, currency is merely one aspect of the financial economy and a very minor part of it. It offers a universal repository of value that various members of society can use right now. Whatever shape it takes, all forms of currency share the same fundamental objectives. By expanding the market for a variety of items, it aids in promoting financial activity. Additionally, it enables customers to save money and subsequently meet future demands. It is widely accepted that Currency "consists only of an agreement within a community to use something as a medium exchange" and that a community no longer refers to a few hundred in a town but a few billion all spread out across the globe. According to the World Factbook, CIA, [12] there are one hundred eighty currencies diagnosed as legal tender in nations that are members of the United Nations (UN), observer nations, partially recognised or unrecognised states, and their dependencies. However, aside from the currencies with fixed exchange rate, there are only one hundred thirty currencies that are impartial to a currency basket. Each of those currencies differs significantly in features including size, colour and texture. To get such features we need to process the image of the currency. In recent times, a number of studies is going on in the field of image processing and recognition or classification. Image processing involves converting an image into a digital format and carrying out specific procedures to extract some actionable data from it. In general, the image processing system treats all images as 2D signals. whilst making use of certain predetermined signal processing methods. Pattern recognition is an application of image processing where measurements of various patterns are being taken around the objects in the image. Here substantially pattern identification task has been dealt with. Image classification is a sub part of pattern recognition, wherein purpose is not only recognizing a single image from a collection of images but classify them into classes wherein they belong to. For example, a train database may contain images of flags of various countries. Each country is represented by their flags with different illumination and orientation. The test set, i.e., where the flags, which requires to be labelled are kept, may contain 5 different flags. These 5 flags may or may not belong to the existing classes of the train set. The system's task will be to classify the unknown image of a flag into unknown category and send a message to the user, telling its origin is unknown and labelling the other flags, which belongs to one of those existing classes correctly. So, suppose out of 5 flags, first one belongs to India, second one belongs to China, third one to USA, fourth one to UK and fifth one to Pakistan. The train database contains the images of flags excluding Pakistan. So, the result will tell that fifth one is unknown while other images will be matched to its closest counterpart and if the system is 100% accurate, the test images of these respective flags will be matched to the train images of the same, hence correctly labelled and classified.

This study is not concerned with identifying flag images, rather it is concerned with identifying and classifying various currencies with different denominations from both of its obverse and reverse side given in any angular orientation as these fields of study have many applications and an active area of research these days.

In this study, the above-mentioned cases are dealt with. In this section the advantages and challenges faced by the systems are discussed briefly alongside problem statement and objective of the thesis.

1.2 Applications

Currency image recognition is the process where a test image of banknote is matched to one of the x (in this case x = 196) classes of currencies which are trained with their unique images during training phase with preferably many train samples encompassing variations found in them. In this way, the test currency image is labelled and hence if there exists number of test samples consisting different currencies, the system classifies each of them to their unique classes based on what the system has learnt from the training phase. This is known as currency image classification. As there are huge number of currencies considering different denominations from both front and the back side of a banknote, arranging them manually is very difficult and prone to errors. The applications of different currency classifications are, listed as follows.

I. It takes a lot of time and effort for a human to rapidly and accurately recognise each banknote. Therefore, a highly effective automated system that aids in banknote recognition is widely desired for the future.

II. Currency recognition from any of its sides, i.e., the obverse side and the reverse side eliminates the limitation of the recognition of a banknote from a particular face.

III. Currency exchange machines can use the system to automatically identify the unit of the currency along with their denominations and can count more efficiently if both faces of banknotes are being recognized.

IV. Currency can be identified from any of the orientation it is placed, whether it is placed at an angle or even placed upside down. So, the system can be used any handheld smart device to identify a banknote and use it accordingly.

1.3 Challenges

Currency image classification has got several challenges which are described briefly as follows.

I. The main challenge for currency recognition is to perceive currency note in a bunch of currency notes where there is a possibility of locating different denomination note.

II. Although different countries have different currencies, there are many banknotes which are very similar to each other specially by colour and size. So, dealing with such situation is a real hurdle in multiple currency recognition process.

III. While a banknote is taken from any orientation, there is a probability that the system may fail to recognise it even after the system is trained with the said currencies due to the difference in angular coordination.

IV. A robust system should be able to identify worn, blurry, defaced and even damaged note during circulation by human being. This is one of the toughest challenges a currency recognition system face.

V. Another hurdle is to tackle the illumination problem while process a image of a banknote. If there is lack of proper light, it is really difficult to extract certain features from an image.

1.4 Problem Statement

Recognition and classification of various currencies of multiple countries and multiple denominations taken from obverse and reverse side of a banknote at any angular orientation is the intention.

1.5 Objectives

The objectives are as follows.

I. Study of existing techniques of currency recognition & classification and improve on them to address their limitations.

II. Finding the best feature for multiple currency recognition by a single system.

III. Recognition of the currency from any angular orientation to make the system use in any handheld smart device.

IV. Recognition of the currency from any of its reverse or obverse side.

1.6. Organization of the Thesis

Section 1 presents an introduction and overview of this thesis. Works which have been done previously are mentioned in Section 2. Section 3 represents the proposed approach. Section 4 and 5 presents experimental results and analysis, respectively. Section 6 is the conclusion.

CHAPTER 2

LITERATURE SURVEY

Multiple currency recognition is the active area of research. Over time many works have used different machine learning techniques to categorise currency images. Garkoti [5] et. al extracted the image of Indian Currency of various denomination from an acquired image. The ROI (Region Of Interest) is extracted by detecting the edge of the banknote by using canny edge detection. After removing noise, the image is resized and converted to grayscale image from RGB image. Then the grayscale image is being matched with selected templates like latent image, watermark, security thread and identification mark. This work does not deal with an input image taken from any orientation and also ample lighting condition should be maintained. This work also deals with currency of one nation only and have not mentioned whether the system will recognise the currency from its reverse side.

Sodhi [1] et. al proposed a methodology that deals with a total of 20 denominations of 3 different currencies. They worked with features like size or area. Further the currencies are classified with the Help of Decision Tree, Fine-KNN and SVM which gives an accuracy of 51%, 65.6% and 56.8% respectively. The accuracy of the recognition should be increased. Also, if the variation of currency is increased only one parameter like size or area is not sufficient to recognize the currencies correctly. Another limitation of the methodology is to have a certain size or area of a currency, a certain aspect ratio should be maintained, so if the aspect ratio of an image is differed from that mentioned aspect ratio, the size or area of the currency would be changed and might fail to identify a currency accurately.

Kalpna Gautam [4] proposed a system that employs LBP (Local Binary Pattern) and Euclidean Distance for matching with the test currency and Principal Component Analysis (PCA) for training purpose. The database should be enough large for the training purpose. Future scope of the system includes currency recognition for more than one country and currency taken from various orientation.

Kamal [9] et. al extracted features from Indian currencies using Bag of Words, Histogram of Oriented Gradients, Colour Comparison using delta-E metric, Template Matching. The system achieved an accuracy of 92% - 100%.

Akter [3] et. al proposed a system that extracts Hue, Saturation, Value from a RGB image of a Bangladeshi currency and also included edge and grey-level co-occurrence matrix as extracted features. Euclidean Distance is used to identify the test images. The system achieved an overall accuracy of 81.27%.

HSV colour space is one of the colour spaces used by human for determining colours. It is an alternative representation of the RGB colour space. The HSV illustration models how colours appear beneath light. A colour with the highest HSV value is analogous to shining a white light on a coloured item. For an instance, flashing a strong white light on a red object makes the object appear red but brighter and more lighted, but shining a low light on a red object makes the object appear darker and less vivid). RGB image is converted to HSV to get respective H, S and V values.

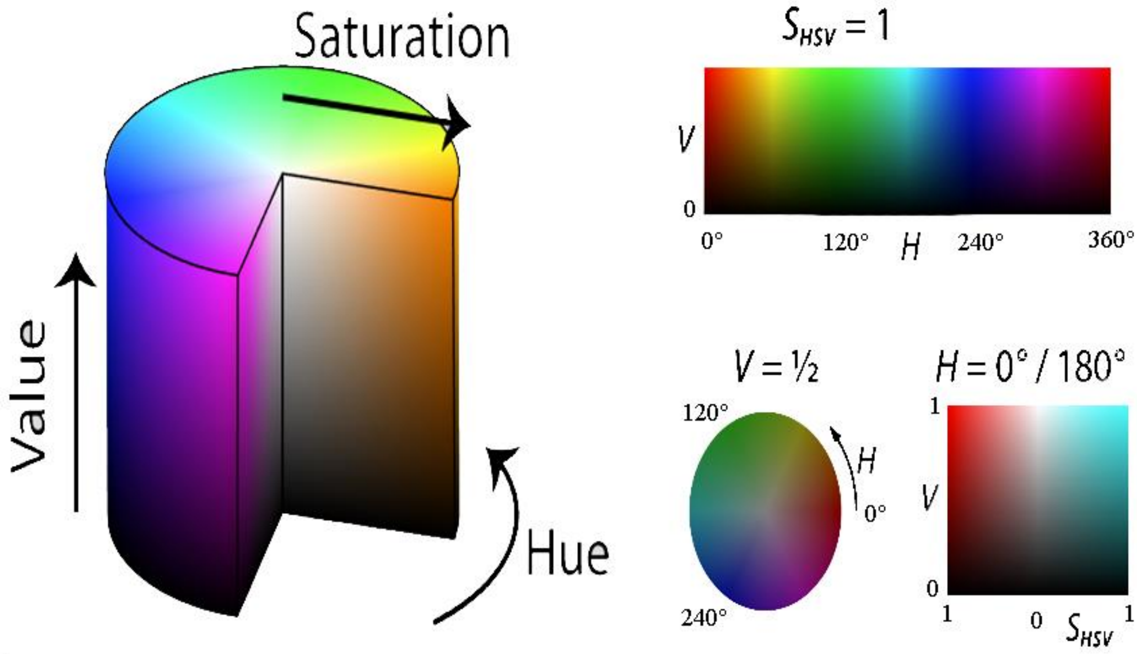

Figure 2.1: HSV colour space

Here,

$$H = Hue = cos - 1[(R - ½G - ½B)/\sqrt{R^2} + G^2 + B - RG - RB - GB]$$

$if\ G \geq B,$

or

$$H = 360 - cos - 1[(R - ½G - ½B)/\sqrt{R^2} + G^2 + B^2 - RG - RB - GB]$$

$if\ B > G.$

$$V = M/255$$
$$S = 1 - m/M \quad if\ M > 0$$
$$S = 0 \quad if\ M = 0$$

$where,$

$$M = max\{R, G, B\}$$
$$m = min\{R, G, B\}$$

For example,

a RGB image has a pixel of Red channel intensity 229, Green channel intensity 227 and Blue channel intensity 214. If the image is converted into HSV, the H value will be 52º, S value will be 0.0655 and V value will be 0.8980. In MATLAB, by **rgb2hsv** command the H value will be brought into 0 to 1, i.e., for this particular pixel H value will be 52/360 = 0.144.

For texture analysis, the **Gray Level Co-occurrence Matrix (GLCM)** is used. It is a technique for statistical texture analysis of the second order. In this method, two pixels are taken into account at once; these two pixels are known as the reference pixel and the neighbour pixel. Before computing the GLCM, the reference and neighbouring pixels have a certain spatial relationship established. The neighbour pixel could be, for instance, one pixel to the left of the current pixel, two pixels above, or three pixels diagonally (towards one of the directions of North-East, North-West, South-East, or South-West) from the reference pixel.

A GLCM of size (Range of Intensities x Range of Intensities) is constructed once a spatial relationship has been established, with all indices initialised to 0. A 256x256 GLCM, for instance, will be present in an 8-bit single channel image.

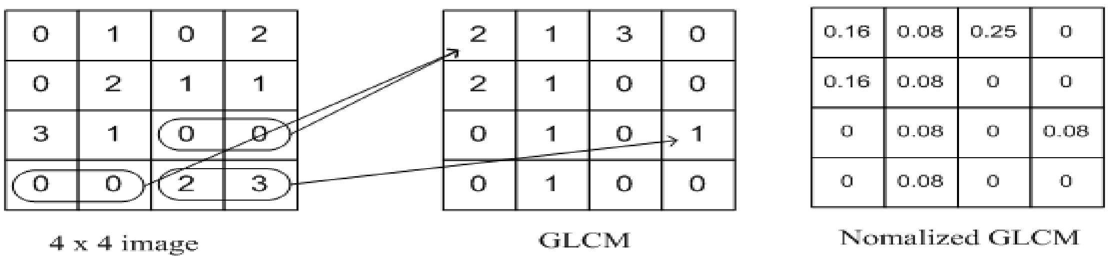

Figure 2.2: Example of Gray Level Co-occurrence Matrix

Each entry of the $GLCM[i, j]$ contains the count of the appearances of that pair of intensities in the image with the specified spatial relationship.

The matrix can be normalised so that each cell expresses the likelihood of that pair of intensities appearing in the image and added to its transpose to make it symmetrical.

The texture attributes from the matrix can then be calculated to represent the textures in the image once the GLCM has been computed.

The texture properties are as follows:

$$G_E = \sum_{i,j=0}^{N-1} \left(P_{ij}\right)^2 \qquad \dots (15)$$

$$G_{En} = \sum_{i,j=0}^{N-1} -ln\left(P_{ij}\right)P_{ij} \qquad \dots (16)$$

$$G_C = \sum_{i,j=0}^{N-1} P_{ij}(i-j)^2 \qquad \dots (17)$$

$$G_H = \sum_{i,j=0}^{N-1} \frac{P_{ij}}{1+(i-j)^2} \qquad \dots (18)$$

$$G_{Cor} = \sum_{i,j=0}^{N-1} P_{ij} \frac{(i-\mu)(j-\mu)}{\sigma^2} \qquad \dots (19)$$

$$G_S = sgn(A)|A|^{1/3} \quad \ldots (20)$$

$$G_P = sgn(B)|B|^{1/4} \quad \ldots (21)$$

where,

G_E = Energy G_C = Contrast G_{Cor} = Correlation

G_{En} = Entropy G_H = Homogeneity G_S = Shade

G_P = Prominence

P_{ij} = Element i,j of the normalized symmetrical GLCM

N = Number of gray levels in the image

μ = the GLCM mean (being an estimate of the intensity of all pixels in the relationship that contributed to the GLCM), calculated as:

$$\mu = \sum_{i,j=0}^{N-1} i\, P_{ij}$$

σ^2 = the variance of the intensities of all reference pixels in the relationships that contributed to the GLCM, calculated as:

$$\sigma^2 = \sum_{i,j=0}^{N-1} P_{ij}(i-\mu)^2$$

$$A = \sum_{i,j=0}^{N-1} \frac{P_{ij}\,(i+j-2\mu)^3}{\sigma^3(\sqrt{2(1+C)})^3} \quad \ldots (22)$$

C = The Correlation feature

sgn(x) = Sign of a real number

x= -1 for $x < 0$

x= 0 for $x = 0$

x= 1 for $x > 0$

$$B = \sum_{i,j=0}^{N-1} \frac{P_{ij}\ (i + j - 2\mu)^4}{4\sigma^4 (1 + C)^2} \quad \ldots (23)$$

Narra [6] et. al proposed a method that extracts security features by segmenting the Indian currency images of various denominations by Chan Vese segmentation. They have considered the security features and colour moments to construct a combined feature. They have achieved an accuracy of 82.7% by SVM.

Colour moments are a type of metric that describes how colours are distributed in an image. Colour moments are typically utilised for colour indexing tasks as features in image retrieval applications, with the goal of comparing how similar images are based on colour. A similar image is found and retrieved by comparing one image to a database of digital images with pre-computed attributes. Each image comparison generates a similarity score, with the lower the number, the more similar the two photos are thought to be.

Colour moments are insensitive to scaling and rotation. Colour moments are a good feature to use in changing lighting circumstances because they encapsulate both shape and colour information, but they can't manage occlusion very well. For any colour model, colour moments can be calculated. Per channel, three colour moments are calculated (e.g., 9 moments if the colour model is RGB and 12 moments if the colour model is CMYK). Colour moments are computed in the same way as probability distribution moments are computed.

Some of the colour moments are:

- **Mean:**

  The average colour in the image might be understood as the first colour moment. It can be calculated by using the following formula

$$E_i = \sum_{j=1}^{N} \frac{1}{N} p_{ij} \qquad \dots(1)$$

  where N denotes the image's pixel count and p_{ij} is the value of the image's j-th pixel in the i-th colour channel.

- **Standard Deviation:**

  The standard deviation is the second colour moment, which is calculated by taking the square root of the colour distribution's variance.

$$\sigma_i = \sqrt{\left(\frac{1}{N}\sum_{j=1}^{N}(p_{ij} - E_i)^2\right)} \qquad \dots(2)$$

Where E_i denotes the mean value, or first colour moment, for the i-th colour channel of the image.

- **Skewness:**
  The skewness is the third colour moment. It determines how asymmetric the colour distribution is and so provides information about the distribution's shape. Skewness can be computed with the following formula:

$$s_i = \sqrt[3]{\left(\frac{1}{N}\sum_{j=1}^{N}(p_{ij} - E_i)^3\right)} \quad \dots (3)$$

Colour moments are most commonly used in **colour indexing**. Images can be indexed, with the computed colour moments stored in the index. The colour moments of the image of interest will also be computed if someone has a specific image and wants to find similar photos in the database. The following function will then be used to generate a similarity score between the image of interest and all of the other photos in the database:

$$d_{mom}(H, I) = \sum_{i=1}^{r} w_{i1}|E_i^1 - E_i^2| + w_{i2}|\sigma_i^1 - \sigma_i^2| + w_{i3}|s_i^1 - s_i^2| \quad \dots (4)$$

where,

The colour distributions of the two examined images are H and I, respectively.

The channel index is i while the total number of channels is r.

E_i^1 and E_i^2 are the first order moments calculated for the image distributions.

σ_i^1 and σ_i^2 are the second order moments calculated for the image distributions.

s_i^1 and s_i^2 are the third order moments calculated for the image distributions.

The user specifies the weights w_{i1}, w_{i2} and w_{i3} for each of the three colour moments.

Finally, the database's photos will be ranked based on the obtained similarity score to the image of interest and the database images with the lowest $d_{mom}(H, I)$ value should be fetched. As the index contains no information regarding the correlation between the colour channels, a retrieval based on $d_{mom}(H, I)$ may result in false positives.

Abburu [11] et. al proposed a method which is able to identify 20 different currencies. The system suggests considering two steps to identify country of origin of the currency and identify denomination of the currency. After de-noising and resizing, they divide the currencies in three group on the basis of empty region. The RGB image of a currency converted into Binary image and proportion of black pixels to white pixels is calculated. After dividing the image into 3 blocks, it is examined by its size ratio, color and template matching. After the country has been determined, K-means clustering and OCR is used to determine the denomination. The system achieved an average computational time of 5.304 seconds but it did not consider different orientations.

Shahbaj Khan [8] et.al explain about a methodology that used basic image processing techniques to verify various cash notes. To make pre-processing of the acquired image easier, the image is transformed from RGB to Grayscale. The edges are identified using the Sobel operator, and the image is segmented using edge-based segmentation. The ORB feature detector has been used to extract features. Micro-lettering, security thread, intaglio printing and other features have been used. It may confront a variety of obstacles, including old notes, worn notes, and poor image quality. The attributes are compared to those of the original cash which is templated in the dataset. To obtain the output, template matching is conducted.

Oriented FAST and Rotated BRIEF (ORB) is essentially a combination of a BRIEF descriptor and a FAST key point detector with numerous tweaks to improve performance. Prior to applying Harris corner measure to determine the top N points among them, it employs FAST to identify crucial spots. Pyramid is also used to create multiscale features. FAST does not compute the orientation, which is one issue. It calculates the patch's intensity-weighted centroid with the corner at the centre. The orientation is determined by the vector's direction from this corner point to the centroid. Moments are computed with x and y that should be in a circular zone of radius r, where r is the size of the patch, in order to increase the rotation invariance.

Now, ORB uses BRIEF descriptors for creating descriptors. However, it is well recognised that BRIEF struggles with rotation. Therefore, ORB controls BRIEF in accordance with the positioning of key points. A (2 x n) matrix S is created that holds the coordinates of these pixels for any feature set of n binary tests at location (x_i, y_i). The S is then rotated to produce the steered (rotated) version S_θ utilising the orientation of the patch, as the rotation matrix.

In order to create a lookup database of pre-calculated BRIEF patterns, ORB individualises the angle in increments of 2π/30 (12°). The appropriate collection of points S_θ will be used to compute its description as long as the key point orientation remains constant across views.

Each bit feature in BRIEF has a significant variance and a mean that is close to 0.5. However, it loses this characteristic and becomes more dispersed when it is directed along a key point direction. A feature is more discriminatory when it has a high variance, since it responds differentially to inputs. The tests' lack of correlation is another desirable quality because each test will then contribute to the outcome. To address all issues, ORB does a greedy search across all feasible binary tests to discover the ones that have both large variance and means near to 0.5, as well as being uncorrelated.

Multi-probe LSH, an advancement over conventional LSH, is utilised for descriptor matching.

Sushma R G [10] et.al proposed that the silver bromide thread on the note can be used to identify it. The genuine note has the words 'RBI' written in both English and Hindi. In this study, an image is captured using the smartphone application IP Webcam. This application makes it possible to obtain a high-resolution snapshot without using any additional light sources. Image pre-processing techniques include filtering, histogram equalisation, and image conversion. By reducing noise, median filtering optimizes the outcome. The image's global contrast is increased using histogram equalisation. There are two forms of conversion: RGB to Grayscale and RGB to YCbCr. Segmentation is a technique for extracting the region of interest from a currency note image. Threshold is an important factor in feature extraction.

Suresh [2] et.al explains that the rectification of distortion, deterioration, and noise removal are all part of image restoration. Interpolation is a method for performing operations such as zooming, shrinking, rotating, and making geometric repairs. To produce a single value output, RGB colour is converted to grayscale. Edge detection aids in the detection of object boundaries within photographs. The goal of segmentation is to make the image representation simpler. Image segmentation algorithms are often based on image intensity value features such as discontinuity and similarity. Currency recognition and verification are aided by an extracted characteristic of the currency image.

Bhavani [7] et.al discussed Banknote detection using SURF, a unique feature extraction technique that combines both an interest point detector and a descriptor. This system includes the following steps: picture acquisition, pre-processing, feature extraction, classification, and outcome. The system will extract the test image's feature and compare it to the database's feature set. Image segmentation and feature extraction are performed by SURF. The region of interest is used to extract features. The obtained result demonstrates the efficacy of the SURF approach as well as currency recognition. This system activates precise money recognition. It shows the results of picture recognition by comparing them to the attributes of template images.

A quick and reliable technique for local, similarity-invariant representation and identification of photos is **Speeded-up Robust Features (SURF)** [14]. The algorithm of SURF consists of two steps:

- **Feature Extraction**
- **Feature Description**

- **Feature Extraction:**

  A relatively simple Hessian matrix approximation is used in the interest point detection method.

  The Integral Image is a simple and effective tool for calculating the sum of values (pixel values) in an image or a rectangular area of a grid (the given image). It can also be used to obtain the mean intensity of a collection of photos.

They make it possible to compute box type convolution filters quickly. The total of all pixels in the input picture I inside a rectangle region created by the origin and x, is represented by the entry of an integral image I_{Σ} (x) at a point x = x = $(x, y)^T$.

$$I_{\Sigma}(\mathbf{x}) = \sum_{i=0}^{i \le x} \sum_{j=0}^{j \le y} I(i,j) \qquad \ldots (11)$$

With I_{Σ} calculated, any upright, rectangular area's intensity sum can be calculated with just four additions., irrespective of its size.

As a substitute of utilising a separate measure for picking the location and the scale, i.e., the Hessian-Laplace detector, SURF depends on the determinant of the Hessian matrix for both. The Hessian of a pixel will be:

$$H\,(f\,(x,y)) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} \qquad \ldots (12)$$

For adjustment with any scale, the picture is filtered by a Gaussian kernel, therefore if a point X = (x, y) is considered, the Hessian matrix H (x, σ) in x at scale σ is specified as:

$$\mathcal{H}(x,\sigma) = \begin{bmatrix} L_{xx}(\mathrm{x},\sigma) & L_{xy}(\mathrm{x},\sigma) \\ L_{xy}(\mathrm{x},\sigma) & L_{yy}(\mathrm{x},\sigma) \end{bmatrix} \quad \ldots (13)$$

where the convolution of the Gaussian second order derivative with the image I at position x is represented by $L_{xx}\ (x,\sigma)$, and the same holds for $L_{xy}\ (x,\sigma)$ and $L_{yy}\ (x,\sigma)$.

To find the determinant of the Hessian matrix, convolution is needed to be performed with a Gaussian kernel first, then the second-order derivative. Box filters allow SURF to further approximate (both convolution and the second-order derivative). These approximate second-order Gaussian derivatives and can be analysed using integral pictures at a very minimal computing cost, regardless of size, which is one of the reasons why SURF is quick.

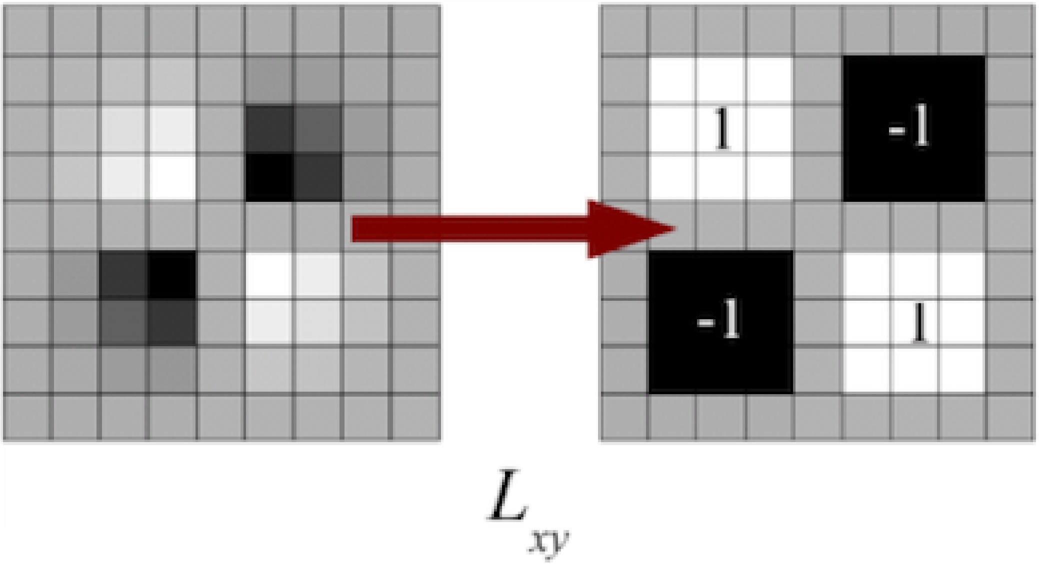

Figure 2.3: Gaussian partial derivative in xy [14]

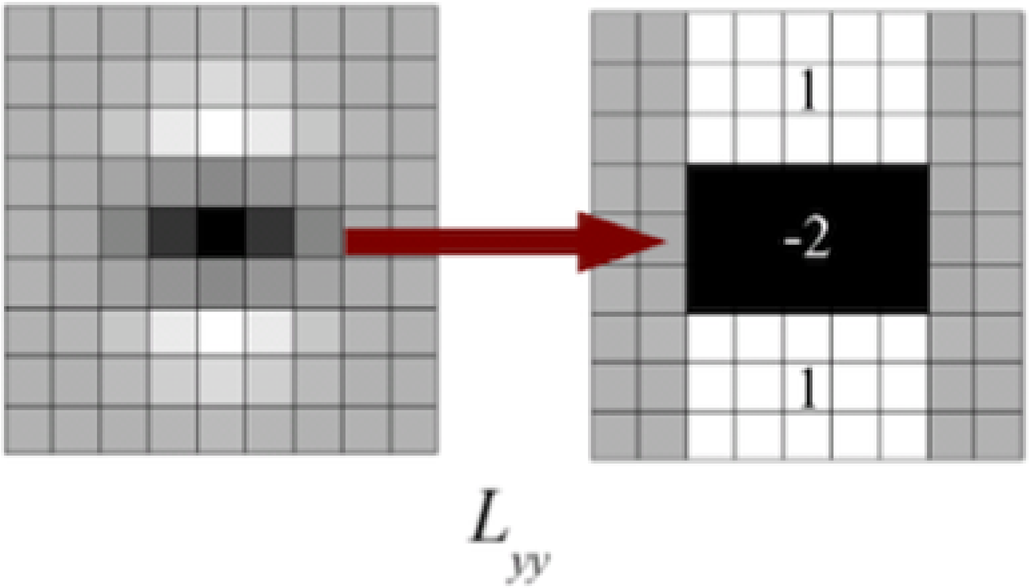

Figure 2.4: Gaussian partial derivative in y [14]

The photos above show 9 × 9 box filters that estimate Gaussian second order derivatives with = 1.2. If the estimations are denoted by D_{xx}, D_{yy}, and D_{xy}, then the determinant of the Hessian will be approximated as:

$$\det(\mathcal{H}_{approx}) = D_{xx}D_{yy} - (\omega D_{xy})^2 \qquad \ldots (14)$$

where, $\omega = 0.9$ (as per Bay's assumption)

Image pyramids are frequently used to implement scale spaces. To achieve a high degree of the pyramid, the pictures are successively smoothened with a Gaussian and then sub-sampled. Because of the use of box filters and integral images, SURF is able to apply filters of any size at exactly the same speed immediately on the original image and even in parallel, as opposed to having to repeat the process each time the output of a previous filtering layer is examined. As a result, rather than iteratively shrinking the image size, the scale space is examined by upscaling the filter size (9×9 → 15×15 → 21×21 → 27×27 and so on). Therefore, for each additional octave, the filter size growth is doubled at the same time as the sample intervals for the extraction of the interest points(σ) can also be doubled, enabling the filter to be scaled up at constant cost. A non-maximum suppression in a 3 × 3 × 3 neighbourhood is used to locate focus points in the photo and over scales.

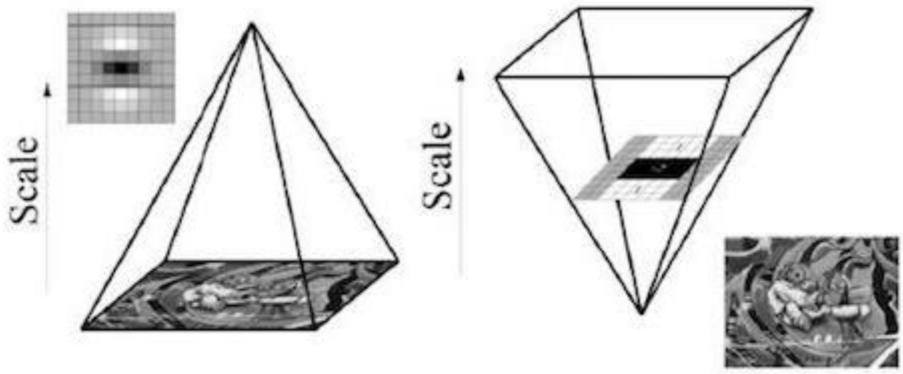

Figure 2.5: Rather than continuously decreasing the size of the image (left), the usage of integral images permits the up-scaling of the filter at constant cost (right) [14]

- **Feature Description:**

A descriptor's purpose is to offer a distinctive and thorough description of an image feature, for instance by characterising the intensity distribution of the pixels close to the place of interest. Thus, the majority of descriptors are computed locally. As a result, a definition is obtained for each previously noted point of interest.

The descriptor's dimensionality directly impacts both how difficult it is to compute and how reliable point matches are. A brief description might be more resistant to alterations in appearance, but it might not provide enough discriminating, leading to an excessive number of false positives.

Establishing a propagable orientation based on data from a circular surrounding the interest point is the first stage. The

SURF descriptor is then extracted from a square region that has been constructed and is aligned to the chosen orientation.

The orientation of the point of interest must be determined in order to perform a rotational invariance. In a $6s$ neighbourhood surrounding the point of interest, where s is the scale at which the point of interest was identified, the Haar wavelet responses in both the x- and y-directions are computed. The acquired results are then represented as locations in a two-dimensional plane, with the horizontal output in the abscissa and the vertical output in the ordinate, weighted by a Gaussian function centred at the location of interest. By summing together all replies inside a moving orientation window of size $\pi/3$, the governing orientation is estimated. Within the window, the horizontal and vertical outputs are combined. A regional orientation vector is then produced by adding the two outputs. The direction of the point of interest is defined by the longest such vector. To achieve the necessary equilibrium among robustness and angular resolution, the dimension of the moving window is a parameter that must be carefully determined.

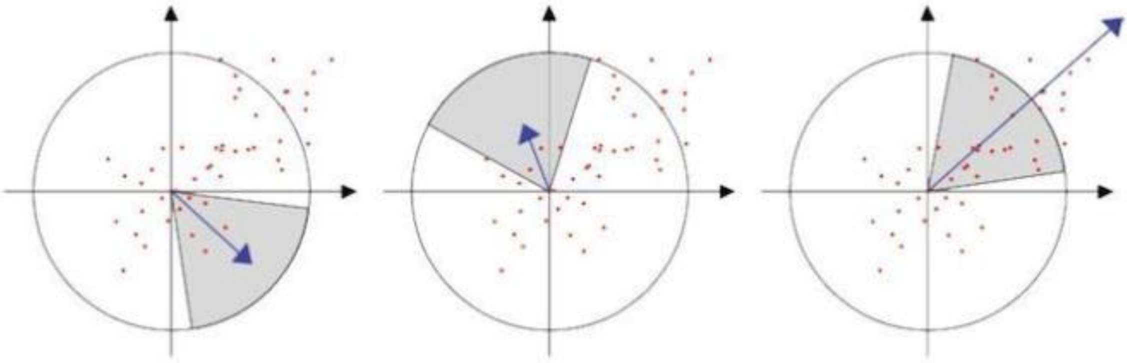

Figure 2.6: Orientation Assignment [14]

Descriptors must be derived when orientation has been established. The very first stage is creating a square region that is centred on the interest point and is aligned with the orientation determined in the previous phase. This window is 20s in size.

The area is then divided into $(4x4)$ smaller, equal-sized square sub-regions. A few straightforward features at (5×5) evenly spaced sample points are computed for each sub-region. dx represents the horizontal Haar wavelet output. while dy denotes the vertical Haar wavelet output (filter size 2s). The outputs dx and dy are first scaled with a Gaussian (σ = 3.3s) centred at the key point to boost robustness against geometric anomalies and localization mistakes.

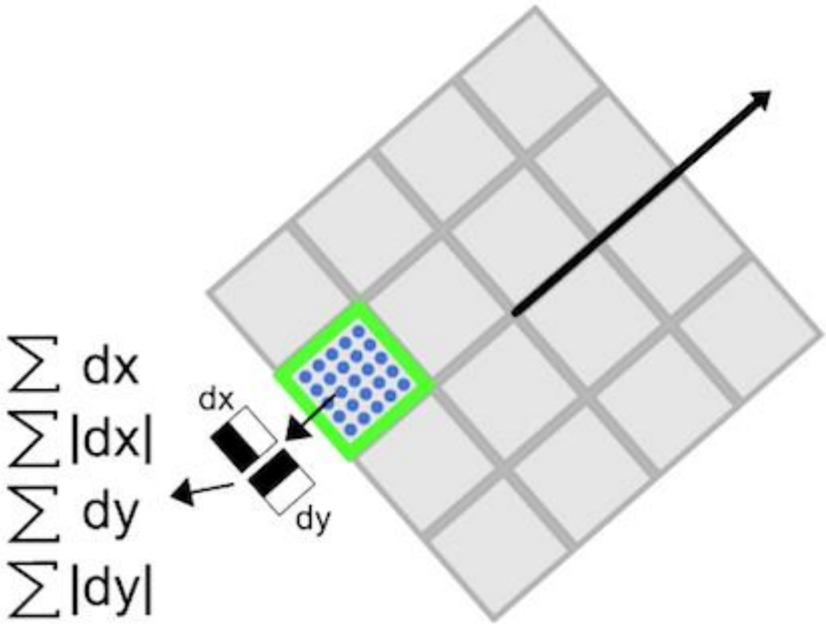

Figure 2.7: Descriptor Components [14]

Then, a first series of components to the feature vector are formed by adding up the wavelet outputs dx and dy across each sub-region. The aggregate of the absolute values of the outputs |dx| and |dy| is also retrieved to add information about the polarity of the intensity variations. As a result, each sub-underlying region's intensity structure V = (∑ dx, ∑ dy, ∑|dx|, ∑|dy|) has a four-dimensional descriptor vector. This produces a descriptor vector of length 64 for each $(4x4)$ sub-regions. SURF is quicker than SIFT because the descriptor in SIFT is a 128-D vector. Matching pairings can be determined by comparing the descriptions collected from various pictures.

S.A. Bhavani [15] and Xu et. al [16] in their respective system designs use SIFT as a feature to extract key points and uses similarity measurement to get to a result.

Scale Invariant Feature Transform (SIFT) [13] is invariance to image scale and rotation. It is often used when different orientation is taken into account. The algorithm of SIFT is discussed below.

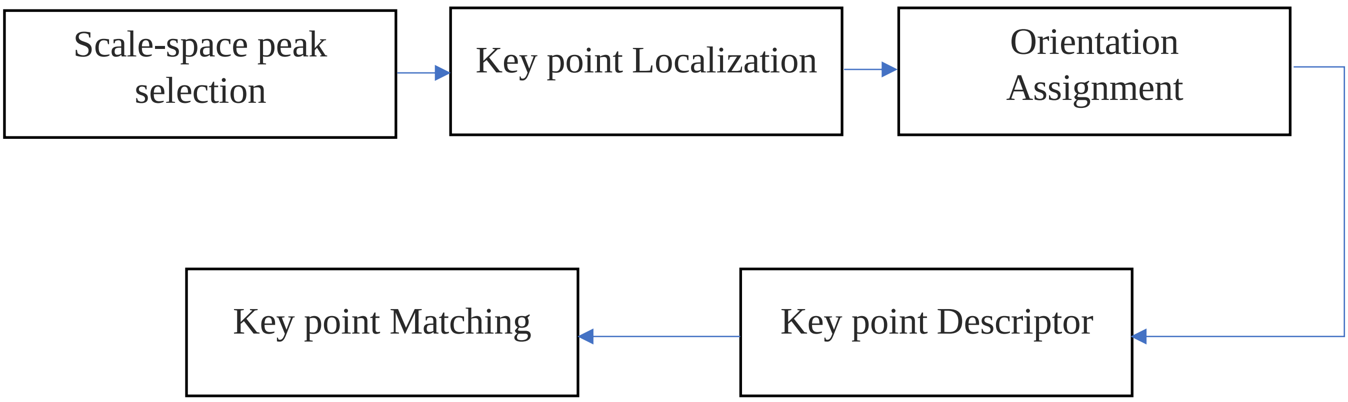

Figure 2.8: Algorithm of Scale Invariant Feature Transform

- **Scale-space peak selection:**

  The scale space of an image is a function $L(x, y, \sigma)$ obtained by convoluting a Gaussian kernel (Blurring) with the input picture at various sizes. The number of octaves and scale used in scale-space is determined by the size of the original image. As a result, numerous octaves of the original image are created. The image size of each octave is half that of the previous octave.

  Using the Gaussian Blur operator, images are gradually blurred across an octave. Blurring is defined as the convolution of the Gaussian operator and the picture in mathematics. Each pixel in a Gaussian blur is subjected to a certain operator. As a result, the image is blurred. Blurred image is expressed by:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad \text{... (5)}$$

Here, the Gaussian Blur operator is G, and the image is I. The position coordinates are x, y, and the scale parameter is σ.

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad \text{... (6)}$$

The blurred photos are now utilised to create a new set of images known as the Difference of Gaussians (DoG). These DoG illustrations are fantastic for identifying noteworthy key points in a photograph. The difference of Gaussian blurring of a picture with two different σ, let's say σ and *k*σ, yields the difference of Gaussian. In the Gaussian Pyramid, this technique is repeated for each octave of the image. It is represented in below image:

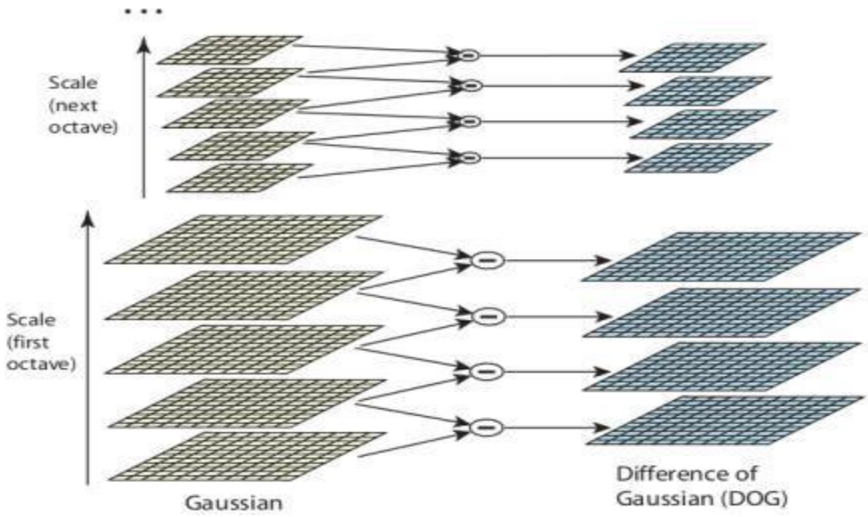

Figure 2.9: Difference of Gaussian [13]

The Difference of Gaussians can then be used to derive scale invariant Laplacian of Gaussian approximations.

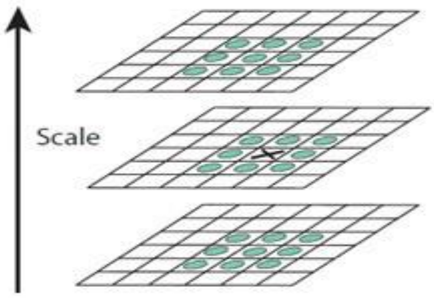

Figure 2.10: Extraction of Local Extrema [13]

Now comparison between each point and its 8 neighbours at the same level as well as 9 neighbours at the above level and 9 neighbours at the below level (total 26 pixels) is being made. It's a potential key point if it's a local extrema. It basically indicates that that scale best represents that key point.

- **Key point Localization:**

  The key points developed in the previous stage result in a large number of key points. Some of these are positioned along an edge, or there isn't enough contrast in them. They aren't as valuable as features in both circumstances. As a result, these points should no longer be considered as key points. As a result, edge features should be eliminated. Checking the intensities of low contrast features should be enough to delete them.

  To acquire a more precise location of extrema, a Taylor series expansion of scale-space is utilised, and if the intensity at this extrema is less than a threshold value (let's say 0.03), it is discarded. Edges have a stronger response in DoG, so they must be removed as well. A 2x2 Hessian matrix (H) has been used to compute the principal curvature.

  - **Reject flats:**
    - $|D(\hat{x})| < \mathbf{0.03}$
  - **Reject edges:**

$$\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad \ldots(7)$$

$$Tr(\boldsymbol{H}) = D_{xx} + D_{yy} = \alpha + \beta \quad \ldots(8)$$

$$Det(\boldsymbol{H}) = D_{xx}D_{yy} - (D_{xx})^2 = \alpha\beta \quad \ldots(9)$$

Where α is the eigenvalue with larger magnitude and β is the smaller one.

$$\frac{\mathrm{Tr}(\mathbf{H})^2}{\mathrm{Det}(\mathbf{H})} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(r\beta + \beta)^2}{r\beta^2} = \frac{(r+1)^2}{r} \quad \text{... (10)}$$

Where, r = $^{\alpha}/_{\beta}$

So, α = rβ

$^{(r+1)^2}/_{r}$ is at its minimum value, when the 2 eigenvalues are equal.

- r < **10**

- **Orientation Assignment:**

  After generating legitimate key points, detection of the orientation is the next job to be done. To get it done, a 16 x 16 square window is taken around detected features. Then edge orientation of each pixel has been computed. Weak edges generated from these calculations is discarded by thresholding the gradient magnitude. Then a histogram of remaining strong edge orientations is made.

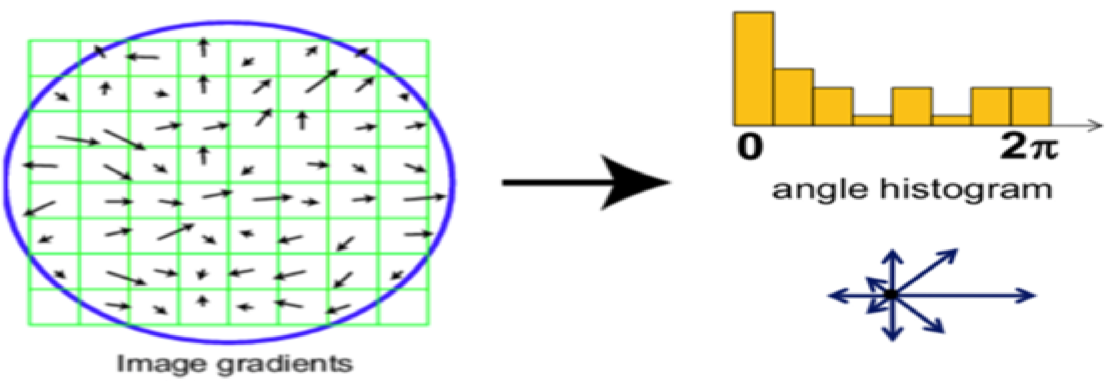

Figure 2.11: Orientation of key point assignment [13]

The orientations of gradient of the sample points inside an area surrounding the key point form an orientation histogram. The 360-degree range of orientations is covered by 36 bins in the orientation histogram. Every sample included to the histogram distribution is weighted by the gradient magnitude and even a Gaussian-weighted circular frame with a scale that is 1.5 times the scale of the key point.

- **Key point Descriptor:**
  Each key point has a scale, orientation and location at this point. The next stage is to develop a descriptor for each key point in the local image region that is extremely distinct and as invariant as possible to alterations like variations in viewpoint and illumination.

  A 4×4 spatial grid of gradient angle histograms is employed to accomplish this. The grid's dimensions are determined by the feature point scale, and the grid is centred on the feature point and rotated to the key point's orientation. An angle histogram divided into 8 is present in each of the spatial bins (128=4×4×8). The scale-space is used to construct the image gradient magnitude and angle.

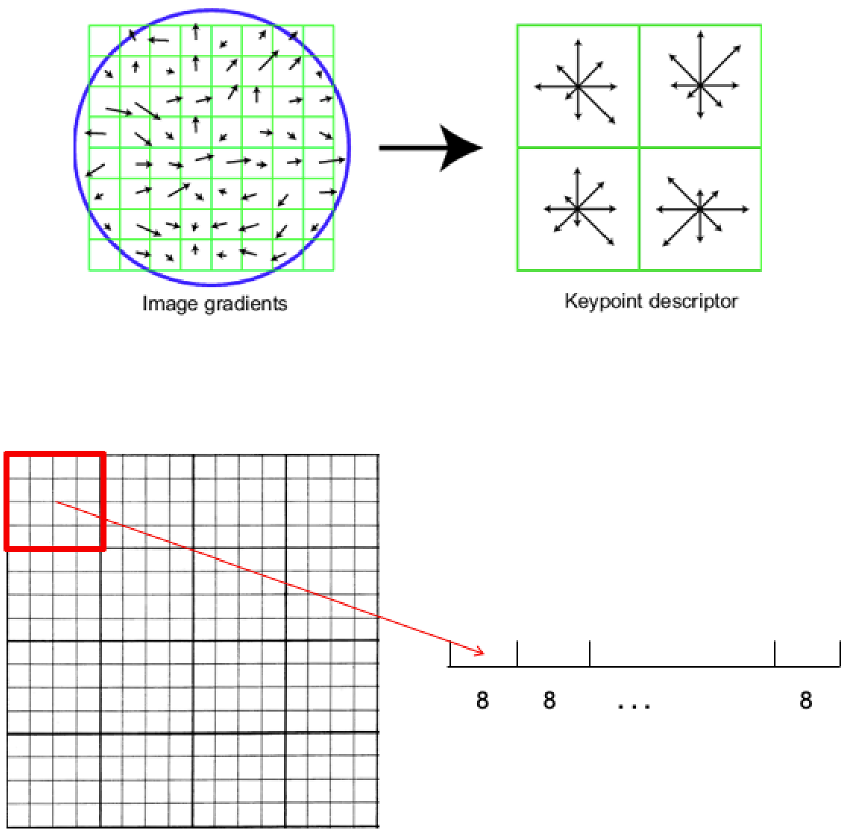

Figure 2.12: 128-dimensional key point descriptor generation [13]

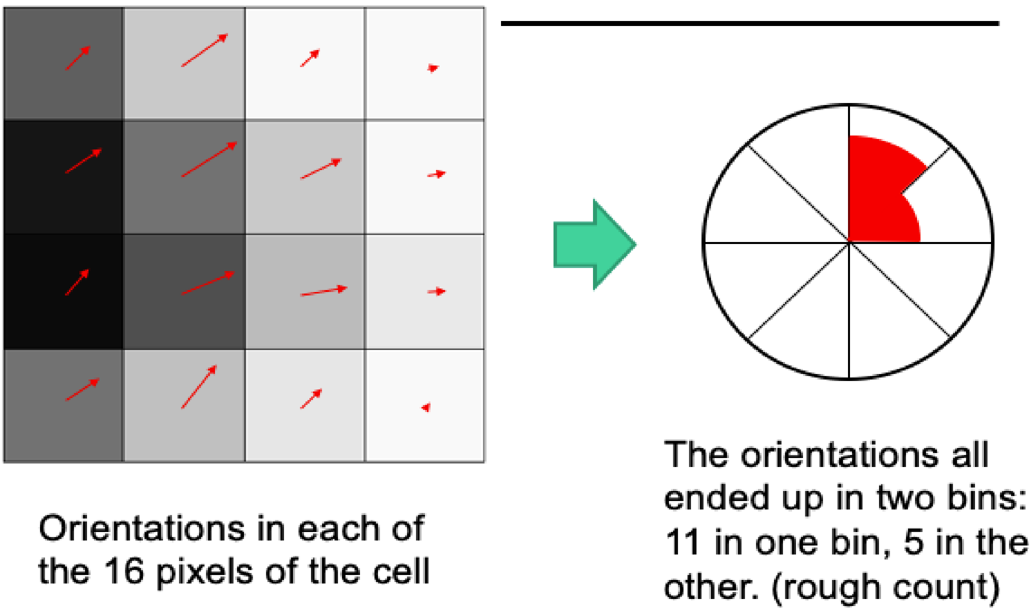

Figure 2.13: 2-bin orientation [13]

Rotation dependence: Gradient orientations are used in the feature vector. Clearly, everything changes when the image is turned. The orientations of all gradients shift as well. The rotation of the key point is removed from each orientation to establish rotation independence. As a result, each gradient orientation is relative to the orientation of the key point.

Illumination dependence: Illumination independence can be attained if numbers can be thresholded against huge numbers. As a result, any integer bigger than 0.2 (out of 128) gets transformed to 0.2. To obtain an illumination independent feature vector, the final feature vector is normalised one more time.

- **Key point Matching:**

  By locating their closest neighbours, key points across two photos are matched.

CHAPTER 3

PROPOSED APPROACH

3.1 Motivation

Image recognition is the processing of the image seen by the system (computer) in such a way that by analysing the digital statistics recording, it is feasible to classify the perceived objects so as to make a verdict. Authentication of the testing data is a recognition problem, when it is compared against the train data saved in the train database, with which the mechanism has been modelled. The perception behind different banknotes recognition of various nations is the same. In this study determination of banknotes of different countries along with their denomination or value have been dealt with. The proposed methodology can recognize the banknotes taken from any orientation. More than a hundred and eighty currencies are available worldwide and the demand for an automated system associated with currencies has been growing dramatically lately. The creation of structures that aid in detecting and recognising different currency notes has been significant due to the need for designing systems that recognise notes without human interaction for numerous exceptional uses. This endeavour is extremely challenging, though, because to the unique characteristics of each note and the security considerations associated with various currencies. Various systems have been proposed earlier that consider various features to recognize banknotes of different denomination of a single nation. However, those methods are not proven to be efficient enough when they deal with various currencies of multiple nations taken from any orientation. The existing methods cause incorrect acceptance, incorrect rejection and incorrect matching while the orientation of the given currency is taken into account. To design an adequate robust system, these hurdles must be curtailed. Therefore, the objective of this study is to deal with the aforementioned challenges and find an efficient solution to these problems so that these technologies can be applied in real world.

3.2 System Components

In order to distinguish the currency based on three different parameters (country of the currency, denomination of the currency and reverse/obverse side of the currency), the system must have following components.

1. Template Database.
2. Classification Module.

1. <u>Template Database:</u>

In this case, the template database is comprised of 198 different templates for reverse and obverse side of various currencies of 16 different countries.

2. Classification Module:

This module may be observed as the decision-making module of a recognition system. In this stage the components of template database are compared against the input data or input image. The templates are compared against the input image and maximum similarity is being found to label it as a representative of a particular class.

3.3 Block Diagram

The block diagram of currency recognition system is given below.

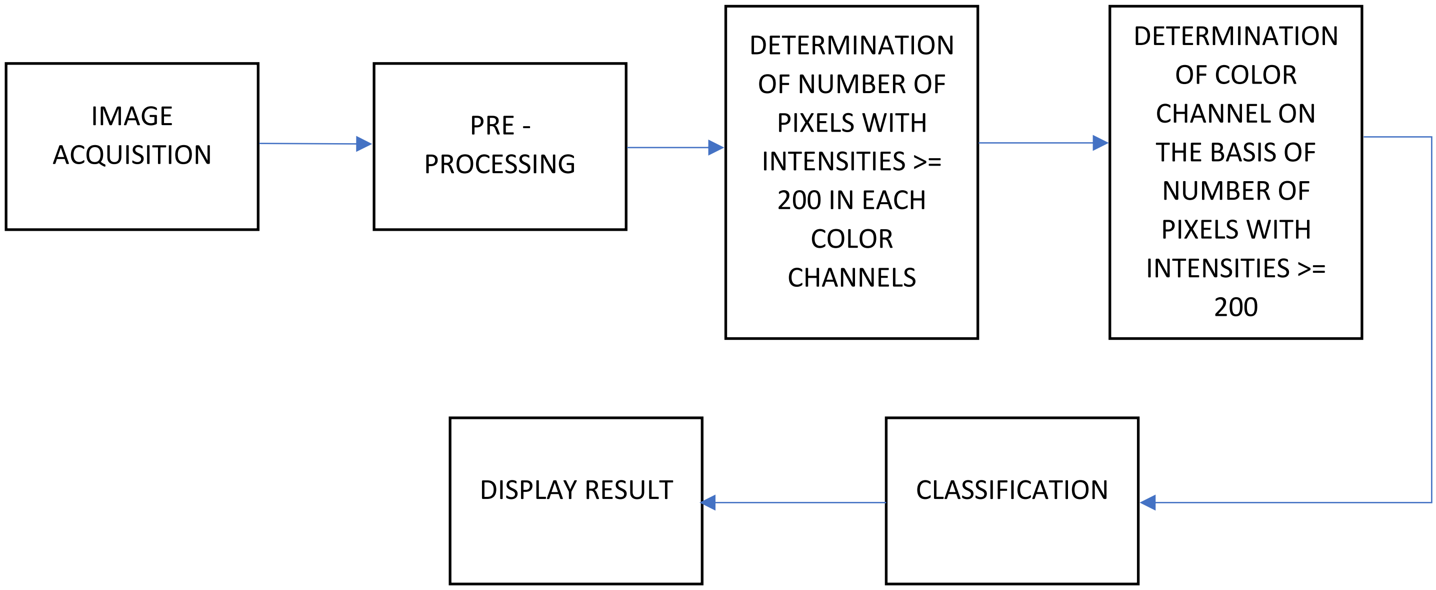

Figure 3.1: Generalised Block Diagram

A detailed flowchart of the system is shown here.

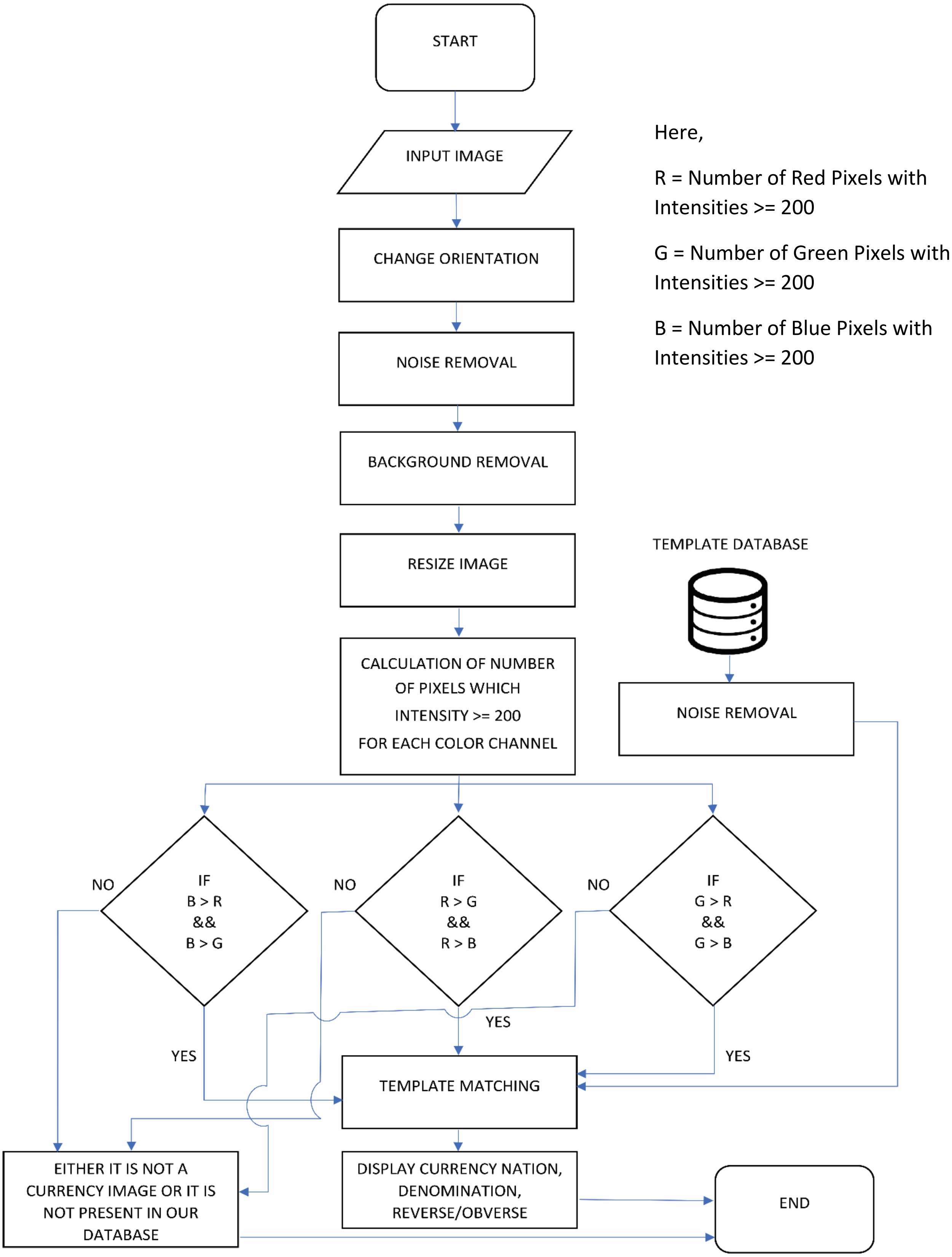

Figure 3.2: Flowchart of the proposed method

Image acquisition phase, the banknote is being acquired. If we acquire the banknote in some angular orientation other than 0°, the background of the banknote must be black. The banknote acquired must be as crinkle free as possible so that we can process the image in further steps more effectively.

The **pre-processing phase** is one of the most important phases in currency recognition process. Algorithm for the pre-processing phase is as follows:

1. The image must be divided into 3 colour channels, i.e., Red channel, Green Channel and Blue Channel.

2. By converting the colour image to a binary image, the image's angular orientation should be verified. Here, the colour image is turned into a binary image by assigning the value 1 (white) to every input pixel with a brightness greater than 0.1 and assigning the value 0 to every other pixel (black).

3. each colour channels should be rotated by the calculated angle such that the angle between the currency image and the X-axis of the image be 0°.

4. The image of the banknote must be free from any unnecessary noise. To de-noise the image, median filter is used. By applying median filter, edges are preserved while removing the noise. So, median filter is applied in each colour channel.

For example, figuring out a pixel neighbourhood’s median value. As can be seen, the median value of 124 replaces the centre pixel value of 150 because it is more indicative of the surrounding pixels. This area is a 3 by 3 square neighbourhood. Greater smoothing will be seen in larger neighbourhoods.

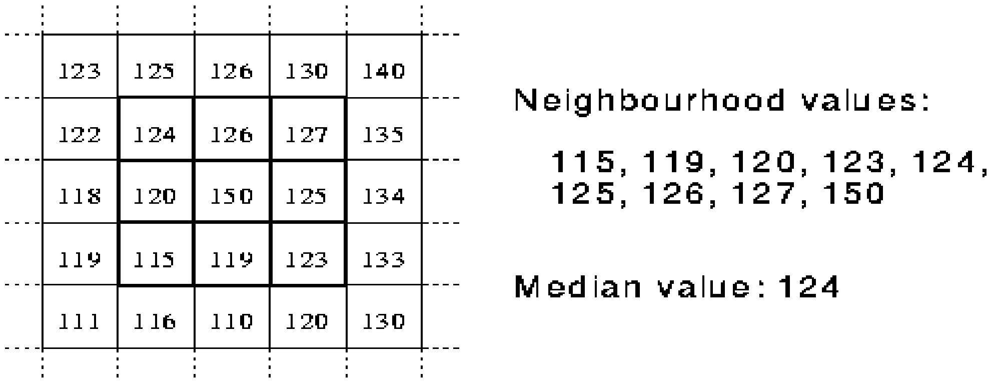

Figure 3.3: Median Filter

5. The RGB image is again concatenated by merging all three-color channels.

6. Due to change in the orientation, a black background is produced. To make the resultant RGB image free from the black background, it is again converted to HSV image.

7. A mask has been generated by thresholding maximum and minimum pixel values of each channel histograms.

8. To crop only the banknote image from the resultant RGB image, the mask has been multiplied with the resultant RGB image.

9. The cropped RGB image is resized to 510 pixels x 720 pixels to ease the further calculations.

Feature Extraction is the most imperative phase of in any recognition problem. Template matching technique is used to extract the features from a given currency.

Finding portions of an image that resemble a template is possible with the use of a technique called **template matching**. A template is a little graphic with certain characteristics. Finding the template in an image is the aim of template matching. Where the image structure fits the mask structure created using the template, where large image values are multiplied by large mask values, the cross-correlation output will be at its highest.

Typically, template matching is carried out by selecting a portion of the search image to serve as a template. Let's assume that the search image is represented by the notation $S\ (x, y)$, where (x, y) stands for the coordinates of each pixel in the search image. The template supposed to be $\boldsymbol{T}\ (\boldsymbol{x_t}, \boldsymbol{y_t})$, with $(\boldsymbol{x_t}, \boldsymbol{y_t})$ standing for the coordinates of each pixel in the template. Following that, the origin or focal point of the template $\boldsymbol{T}\ (\boldsymbol{x_t}, \boldsymbol{y_t})$ moved across each $(\boldsymbol{x}, \boldsymbol{y})$ point in the search image and over the whole region covered by the template, the product sum between the coefficients in in $\boldsymbol{S}\ (\boldsymbol{x}, \boldsymbol{y})$ and $\boldsymbol{T}\ (\boldsymbol{x_t}, \boldsymbol{y_t})$ is determined. The location with the highest correlation value is regarded to be the optimum position after taking into account all feasible positions for the template with respect to the search image. The template is known as a filter mask, and the technique is occasionally referred to as linear spatial filtering.

Comparing the pixel intensities using the Sum of absolute differences can be used to solve translation issues on photos when utilising template matching. In the event that a pixel in the search picture with coordinates $(\boldsymbol{x_s}, \boldsymbol{y_s})$ has intensity $\boldsymbol{I_s}(\boldsymbol{x_s}, \boldsymbol{y_s})$ and a pixel in the template with coordinates $(\boldsymbol{x_t}, \boldsymbol{y_t})$ has intensity $\boldsymbol{I_t}(\boldsymbol{x_t}, \boldsymbol{y_t})$, the absolute difference in the pixel intensities is to be referred as:

$$\boldsymbol{Diff}(\boldsymbol{x_s}, \boldsymbol{y_s}, \boldsymbol{x_t}, \boldsymbol{y_t}) = |\boldsymbol{I_s}(\boldsymbol{x_s}, \boldsymbol{y_s}) - \boldsymbol{I_t}(\boldsymbol{x_t}, \boldsymbol{y_t})| \quad \ldots.(24)$$

After we calculate the difference, the sum of absolute differences is computed as follows:

$$Sum\ of\ Absolute\ Difference\ (x, y) = \sum_{i=0}^{T_{\text{rows}}} \sum_{j=0}^{T_{\text{cols}}} \text{Diff}(x + i, y + j, i, j) \quad \ldots(25)$$

The concept of cycling through each pixel inside the search image to translate it into the template's origin and calculating the Sum of Absolute Difference which is mathematically expressed as follows:

$$\sum_{x=0}^{S_{\text{rows}}} \sum_{y=0}^{S_{\text{cols}}} Sum\ of\ Absolute\ Difference\ (x, y) \quad \ldots(26)$$

$\boldsymbol{S_{cols}}$ and $\boldsymbol{S_{rows}}$ signify the columns and the rows of the search image and $\boldsymbol{T_{cols}}$ and $\boldsymbol{T_{rows}}$ signify the columns and the rows of the template image respectively. In this method the lowest Sum of Absolute Difference gives the estimate for the optimum position of template within the query image. The method is simple to implement and understand, but it is one of the slowest methods.

Amongst the aforementioned features, Template Matching is selected as the feature for final system design for the following reasons:

- The accuracy discovered for this feature in comparison to others are a much better.
- The overlapping between different currency images belonging to different countries are least in this case and lesser than even SURF.
- Dissimilar currencies of outside the database often is identified as currency present in the Known Dataset while using SURF, SIFT and ORB as the descriptor of a key point generated from a query currency image sometimes meet the taken threshold value generated for a class in train dataset. So, discrepancies arise.
- While taking Colour Moments into consideration, in spite of being scaling and rotation invariant, colour moments differ from the trained colour moments under influence of different illumination. So, it is not very fruitful to use colour moments as feature.
- Another concern while taking colour moments is if the classes to be detected are huge and diverse, there is a probability that the colour moments for 2 classes fall in a

very close range, thus there is a chance of overlapping problem.

- If dataset is not large enough, SURF fails to identify class accurately.

Classification phase starts with classifying the images into 3 initial classes based on their number of red pixels, green pixels and blue pixels intensities those are greater than or equal to a particular intensity level. Then it is compared against selected templates to get the correlation values.

Cross-correlation calculations have been made in the spatial or frequency domains to obtain correlation values. The distance metric (squared Euclidean distance) serves as the driving force behind the employment of cross-correlation for template matching.

$$d^2{}_{f,t}(u,v) = \sum_{x,y}[f(x,y) - t(x-u,y-v)]^2 \qquad \ldots (27)$$

where, the image is represented by f, and the total is over x, y beneath the window containing the feature t situated at u,v. In the expression d^2,

$$d^2{}_{f,t}(u,v) = \sum_{x,y}[f^2(x,y) - 2f(x,y)t(x-u,y-v) + t^2(x-u,y-v)] \quad \ldots (28)$$

the term $\sum t^2(x-u,y-v)$ is constant. If another term $\sum f^2(x,y)$ is approximately constant, then the remaining cross-correlation term is

$$c(u,v) = \sum_{x,y} f(x,y)t(x-u,y-v) \qquad \ldots (29)$$

which is a measure of the similarity between the image and the feature.

The use of (29) for template matching has a number of drawbacks.

1. Matching using (29) may fail if the image energy $\sum f^2(x,y)$ fluctuates with position. As an illustration, the feature's correlation with a region in the image that perfectly matches, can be lower than its correlation with a luminous spot.
2. The range of $c(u,v)$ is dependent on the size of the feature.
3. Since the illumination conditions fluctuate throughout the image series, Eq. (29) is not invariant to changes in image amplitude.

By converting the image and feature vectors to unit length and producing a cosine-like correlation coefficient, the correlation coefficient gets around these problems.

$$\gamma(u,v) = \frac{\sum_{x,y}\left[f(x,y) - \bar{f}_{u,v}\right]\left[t(x-u,y-v) - \bar{t}\right]}{\left\{\sum_{x,y}\left[f(x,y) - \bar{f}_{u,v}\right]^2 \sum_{x,y}\left[t(x-u,y-v) - \bar{t}\right]^2\right\}^{0.5}} \quad \ldots (30)$$

where $\bar{t}$ is the mean of the feature and $\bar{f}_{u,v}$ is the mean of $f(x,y)$ in the region under the feature. (30) is referred to as normalized cross-correlation.

If the correlation value is greater than or equal to a defined threshold value, then it is classified into that particular class.

CHAPTER 4

EXPERIMENTATIONS AND RESULTS

This section, which contains all the findings and graphs that have brought us to our conclusion, represents a significant portion of this study. This section assisted us in selecting the best feature and classifier for the system. The experiments are carried out in MATLAB environment. MATLAB R2021a is the version used.

4.1 Dataset for Currency Recognition

The most early and one of the pivotal steps of designing a system is the collection of relevant, solid, reliable dataset required for the process. In this case, the dataset for currency recognition is collected from Numista, one of the largest numismatic catalogues in the world. A total 98 different denominations of various currencies to be considered to design the system. As the system can recognise the currency both from the front and the back, a total of 196 classes is considered here. To make it rotation invariant, both the front and the back side of the currencies of each denomination are rotated from 1^{o} to 360^{o}. The sample dataset is shown below.

Figure 4.1: 196 images of front and back side of the currencies

Figure 4.2: some of the currency images which are having orientation other than 0º

4.2 Experimentations

Five sets of experiments have been performed to evaluate the performance of the proposed system. 10 different images of a currency are selected in random orientation for each set of experiment.

Experiment 1:

The first experiment is conducted where number of green pixel intensity greater than 200 is the highest among all the red, green and blue pixels' intensity higher than 200. 10 images of 50 USD from reverse have been taken. The images taken are in different orientations. We can notice from the bar graph shown in Figure 4.3, "usd 50 back (111)", "usd 50 back (158)", "usd 50 back (185)", "usd 50 back (195)", "usd 50 back (264)" are giving maximum correlation coefficient with the upside-down template and not giving a good enough correlation coefficient with the template taken to identify the said currency. From this observation, it can be said that the aforementioned images are identified while their cropped images are being oriented in an upside-down manner. On contrary, "usd 50 back (1)", "usd 50 back (300)", "usd 50 back (342)", "usd 50 back (55)" and "usd 50 back (79)" are giving a maximum correlation coefficient with the template taken to identify the said currency while not giving a good enough correlation coefficient with the upside-down template. Hence, it is identified that their cropped images are being oriented in a straight way (0^{o} orientation with the X-axis). Here, in this case, the threshold value of the correlation coefficient is taken as 0.8 to reject any image that generates any correlation coefficient less than 0.8.

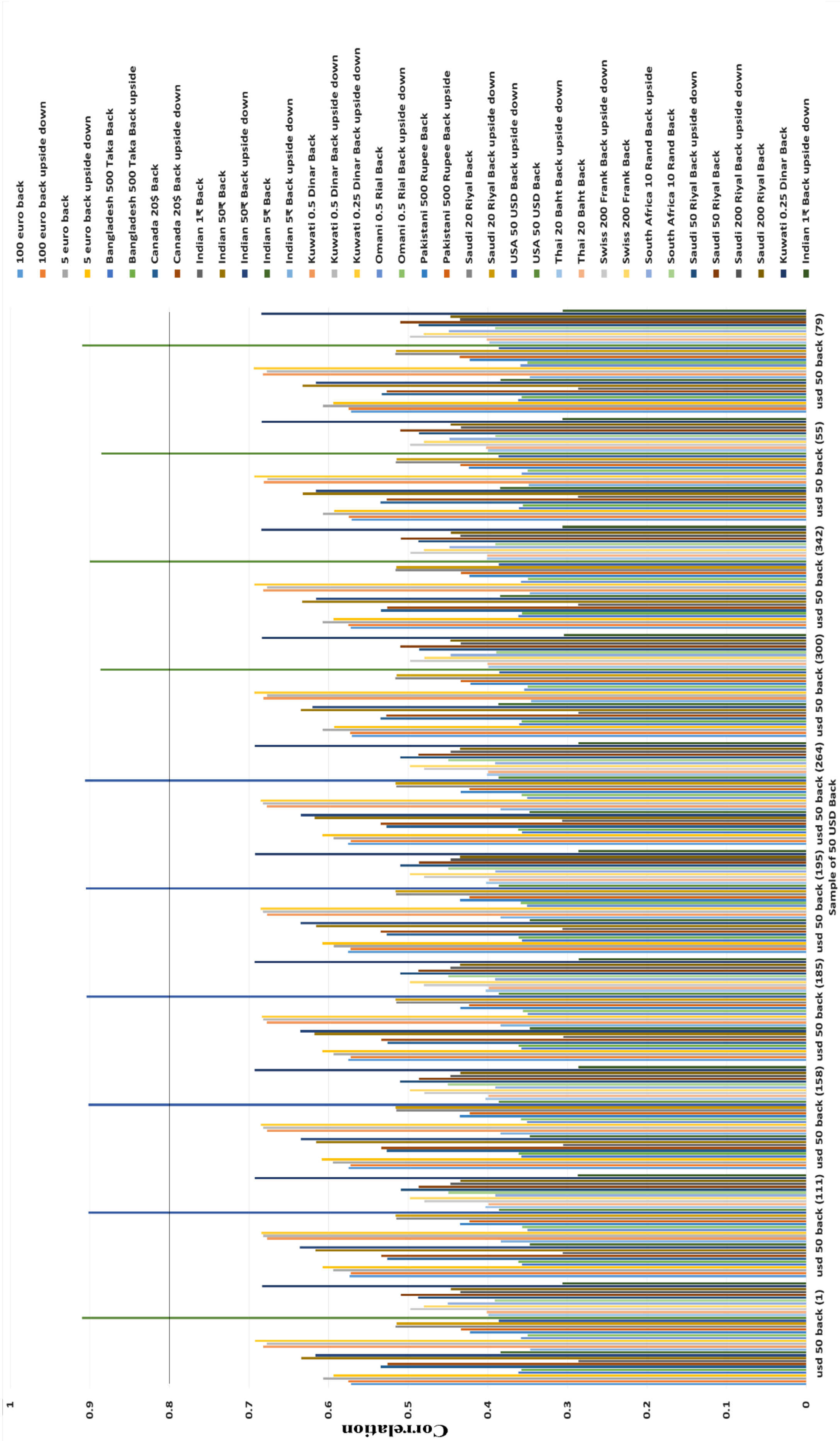

Figure 4.3: Correlation values of 10 test samples of 50 USD Reverse

Figure 4.4: Template matching with 50 USD reverse side test image

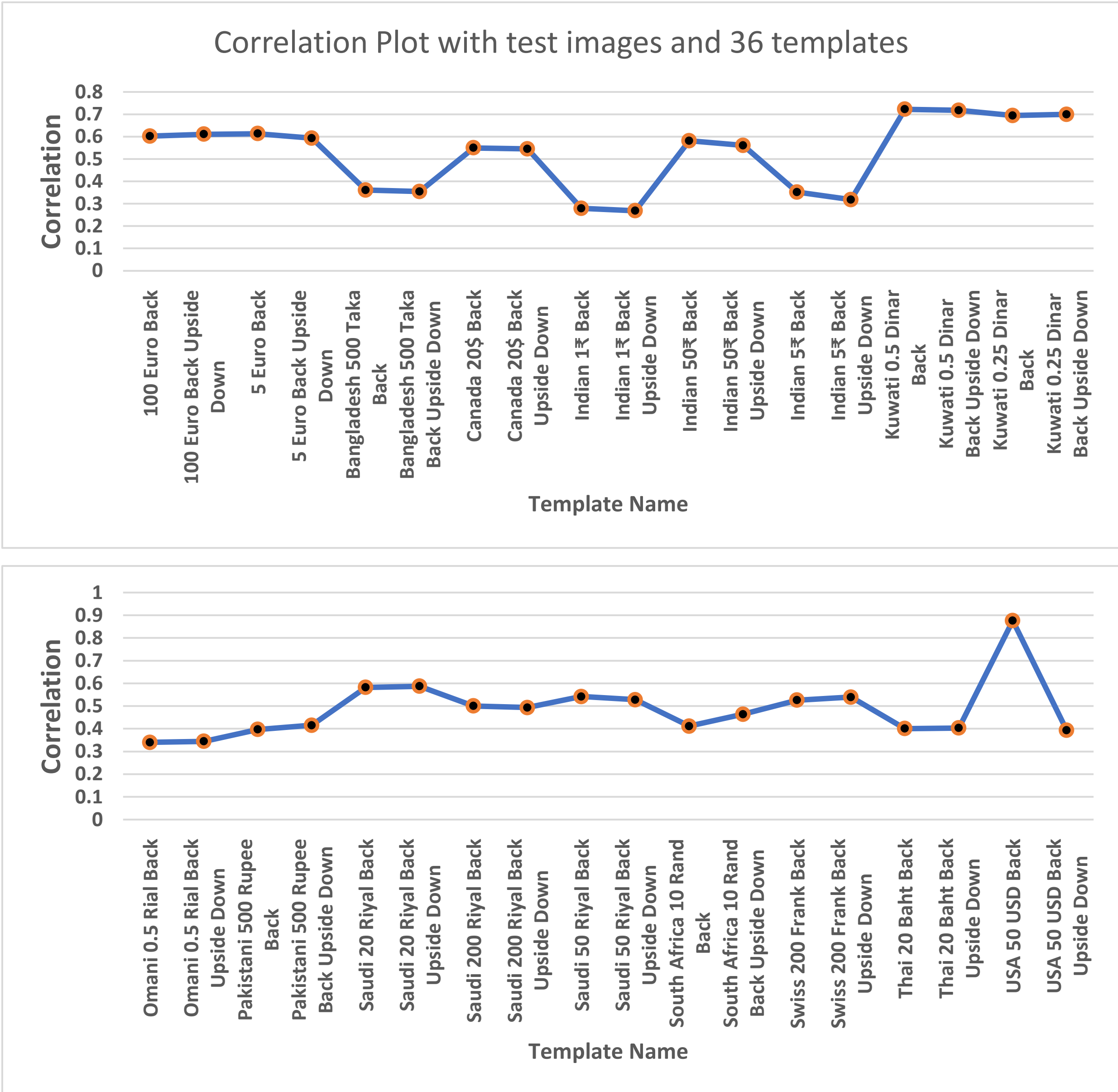

Figure 4.5: Correlation of 50 USD template with the test image

From the Figure 4.5, it is observed that the correlation of 50 USD template with the test image giving the highest correlation coefficient which is higher than the threshold value of 0.8. So, the test image is identified as a 50 USD Banknote (Back).

Experiment 2:

The second experiment is conducted where number of blue pixel intensity greater than 200 is the highest among all the red, green and blue pixels' intensity higher than 200. 10 images of 50 Baht from reverse have been taken. The images taken are in different orientations. We can notice from the bar graph shown in Figure 4.6, "Thai 50 Baht Back (110)", "Thai 50 Baht Back (152)", "Thai 50 Baht Back (203)", "Thai 50 Baht Back (216)", "Thai 50 Baht Back (263)" are giving maximum correlation coefficient with the upside-down template and not giving a good enough correlation coefficient with the template taken to identify the said currency. From this observation, it can be said that the aforementioned images are identified while their cropped images are being oriented in an upside-down manner. On contrary, "Thai 50 Baht Back (10)", "Thai 50 Baht Back (284)", "Thai 50 Baht Back (334)", "Thai 50 Baht Back (60)" and "Thai 50 Baht Back (88)" are giving a maximum correlation coefficient with the template taken to identify the said currency while not giving a good enough correlation coefficient with the upside-down template. Hence, it is identified that their cropped images are being oriented in a straight way (0^{o} orientation with the X-axis). Here, in this case, the threshold value of the correlation coefficient is taken as 0.7 to reject any image that generates any correlation coefficient less than 0.7.

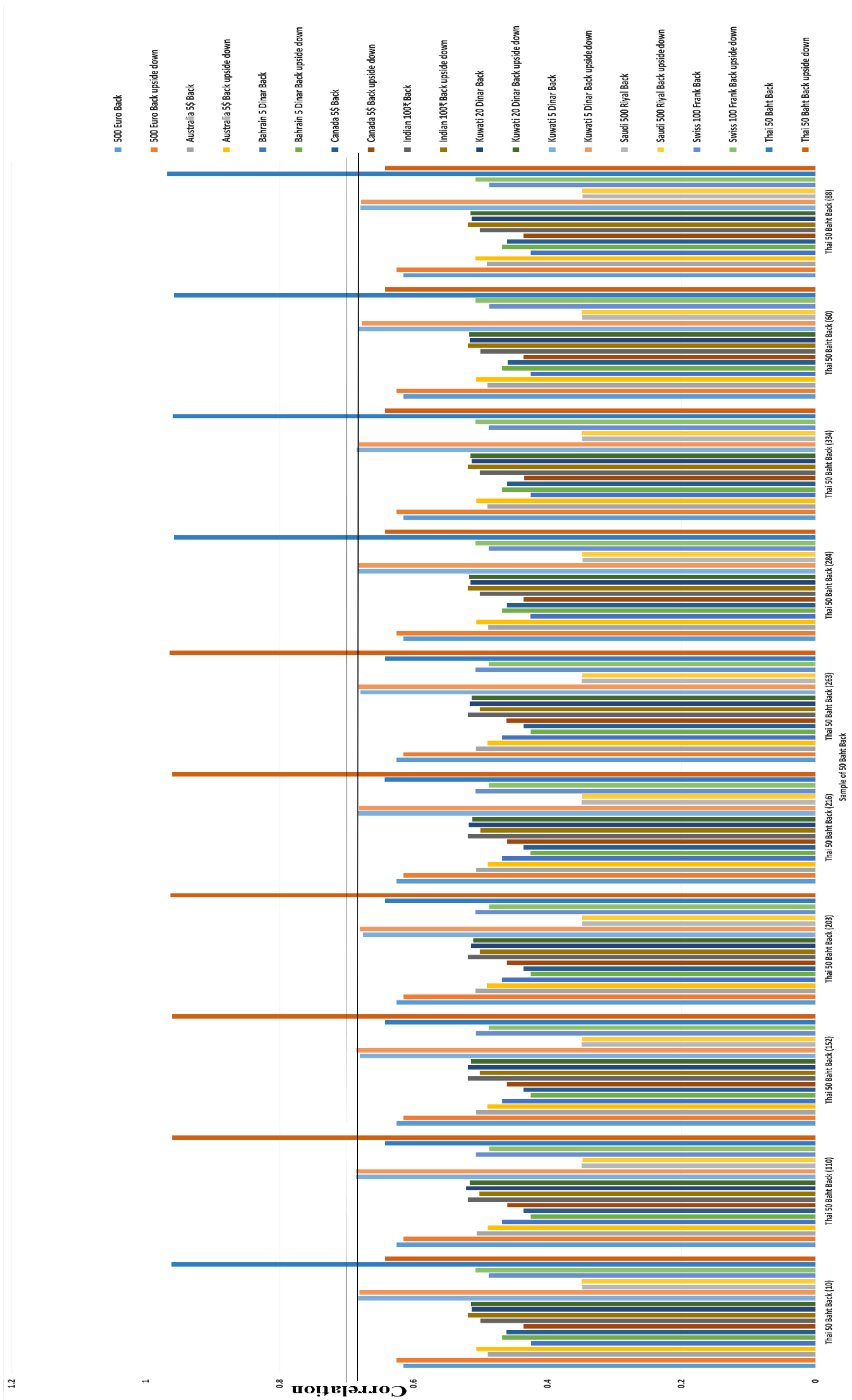

Figure 4.6: Correlation values of 10 test samples of 50 Baht Reverse

Figure 4.7: Template matching with 50 Baht reverse side test image

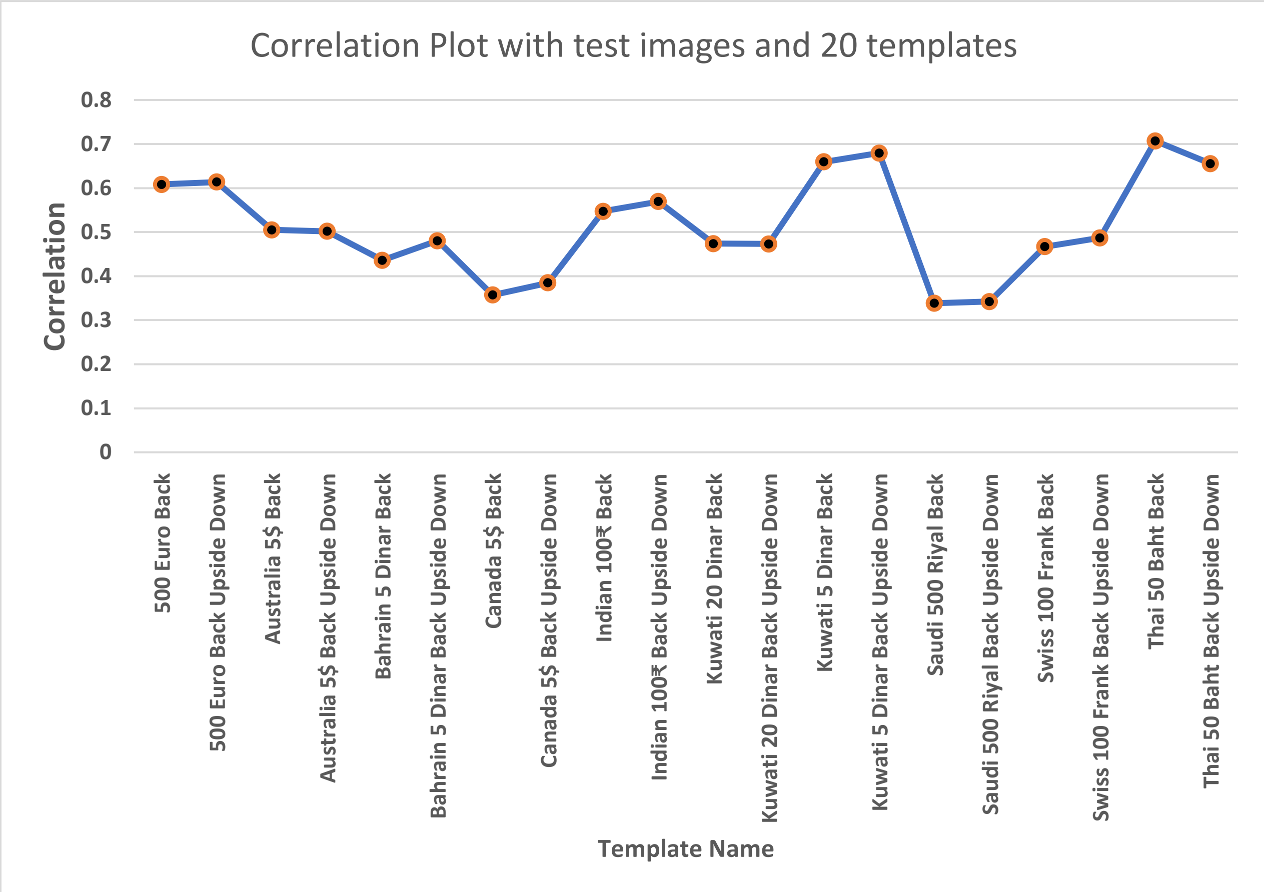

Figure 4.8: Correlation of 50 Baht template with the test image

From the Figure 4.8, it is observed that the correlation of 50 Baht template with the test image giving the highest correlation coefficient which is higher than the threshold value of 0.7. So, the test image is identified as a 50 Baht Banknote (Back).

Experiment 3:

The Third experiment is conducted where number of red pixel intensity greater than 200 is the highest among all the red, green and blue pixels' intensity higher than 200. 10 images of Bangladeshi 50 Taka from reverse have been taken. The images taken are in different orientations. We can notice from the bar graph shown in Figure 4.9, "Bangladesh 50 Taka Back (124)", "Bangladesh 50 Taka Back (138)", "Bangladesh 50 Taka Back (144)", "Bangladesh 50 Taka Back (222)", "Bangladesh 50 Taka Back (95)" are giving maximum correlation coefficient with the upside-down template and not giving a good enough correlation coefficient with the template taken to identify the said currency. From this observation, it can be said that the aforementioned images are identified while their cropped images are being oriented in an upside-down manner. On contrary, "Bangladesh 50 Taka Back (17)", "Bangladesh 50 Taka Back (293)", "Bangladesh 50 Taka Back (341)", "Bangladesh 50 Taka Back (40)" and "Bangladesh 50 Taka Back (73)" are giving a maximum correlation coefficient with the template taken to identify the said currency while not giving a good enough correlation coefficient with the upside-down template. Hence, it is identified that their cropped images are being oriented in a straight way (0^o orientation with the X-axis). Here, in this case, the threshold value of the correlation coefficient is taken as 0.7 to reject any image that generates any correlation coefficient less than 0.7.

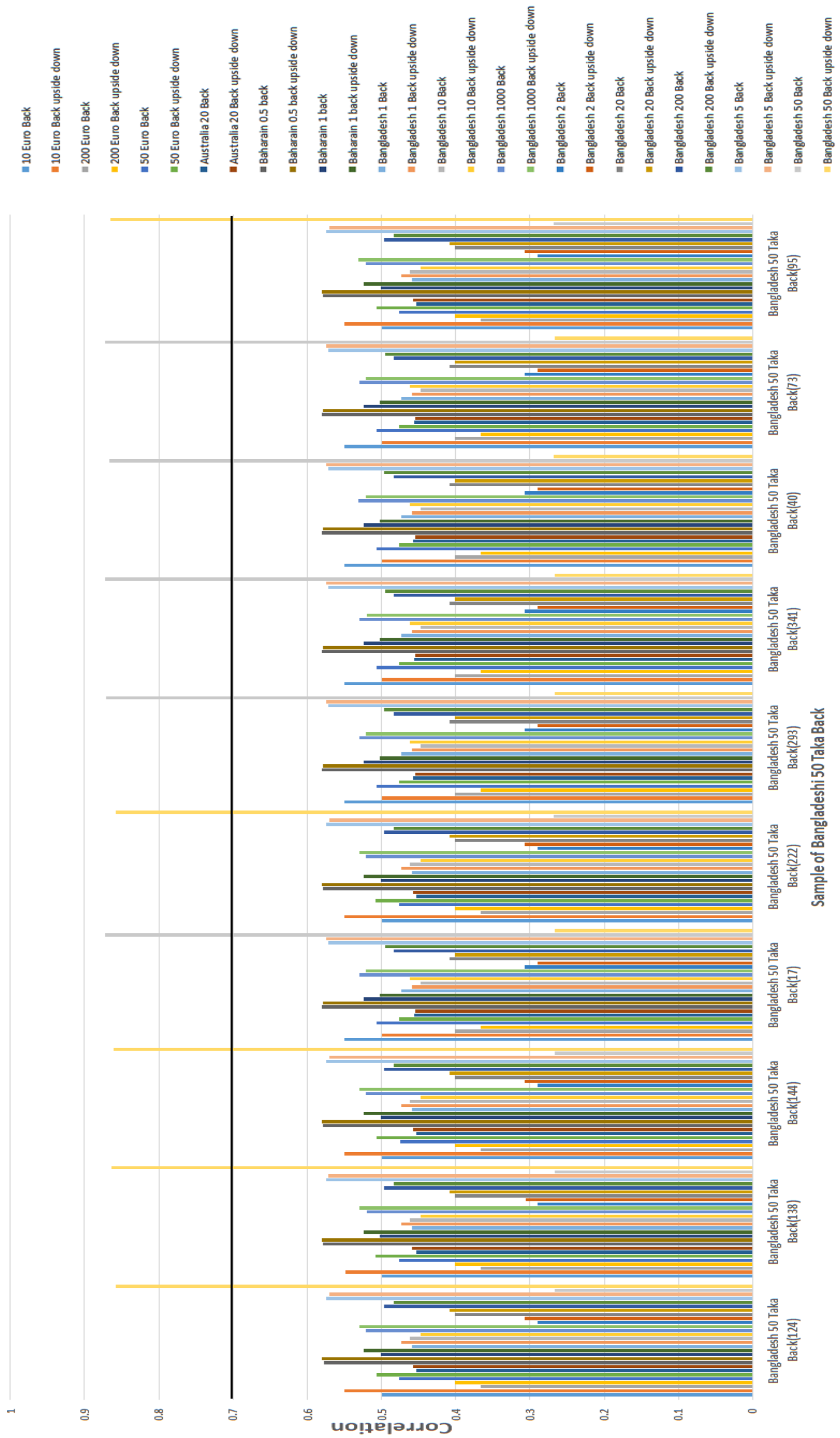

Figure 4.9: Correlation values of 10 test samples of 50 Taka Reverse

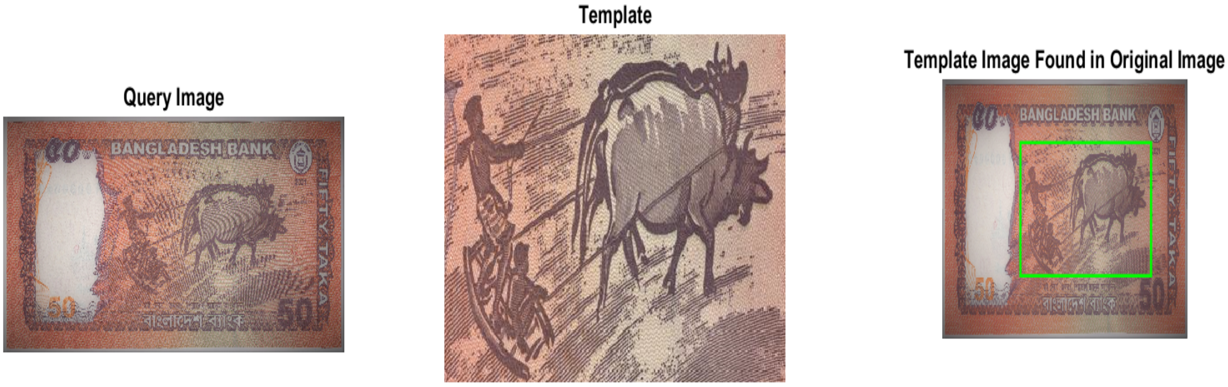

Figure 4.10: Template matching with 50 Taka reverse side test image

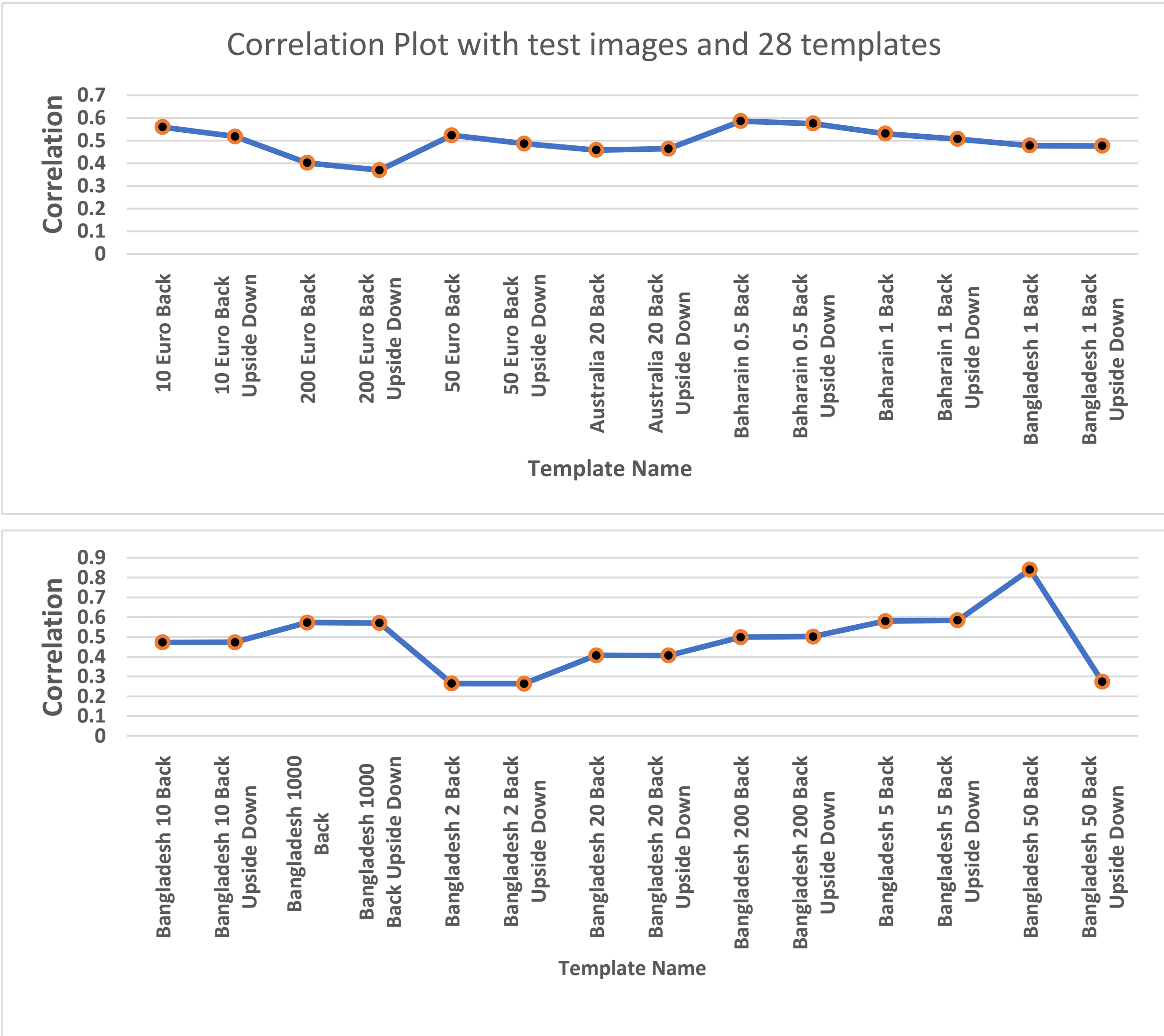

Figure 4.11: Correlation of 50 Taka template with the test image

From the Figure 4.11, it is observed that the correlation of 50 Taka template with the test image giving the highest correlation coefficient which is higher than the threshold value of 0.7. So, the test image is identified as a 50 Taka Banknote (Back).

Experiment 4:

This experiment is done further extend the scope of the system. Due to illumination problem the count of red, green and blue pixels might be altered. So, to handle this problem if the system fails to recognise a currency image from the given criteria, it checks all the templates. For example, an image of Indian 50 Rupees note captured from camera is taken, which was giving number of blue pixels greater than 200 intensity as the highest among all three colour channels which intensity is greater than 200. The proposed system assumes that Indian 50 Rupees banknote should have greater number of green pixels which are greater than 200. So, to handle this exceptional case, a block is provided where all the templates of the database will be presented to correlate with. There the 50 Rupee note is successfully identified.

Figure 4.12: Template matching with Indian 50 Rupee reverse side test image

Experiment 5:

This experiment is done with some Indian 100 Rupee notes with some marks on the template area and in absence of proper lighting conditions. While most of the test images are identified with the template made for the Indian 100 Rupee note, 2 of the images are identified as unknown. This implies that bank notes with dirt and bank note images taken in absence of ample lighting condition have some influences on the result. It is found that the Figure 4.13 has a correlation coefficient of 0.5852 and unable to identify the Indian 100 Rupee banknote.

Figure 4.13: Template matching with an unclean Indian 100 Rupee reverse side test image in presence of shadow

While the system failed to identify the Indian 100 rupee note on Figure 4.13, It identifies the Indian 100 Rupee note used in Figure 4.14 successfully.

Figure 4.14: Template matching with an unclean Indian 100 Rupee reverse side test image in presence of shadow

Figure 4.14 also shows an Indian 100 Rupee bank note that is not taken in a proper lighting condition and it also has some thing written on the place where the template is to be matched. But the amount of both shadow and griminess is much more in the Figure 4.13. So, the system can handle both shadow and filthiness over the template zone up to an extent.

The Confusion matrix is shown here to understand the accuracy of the identification system.

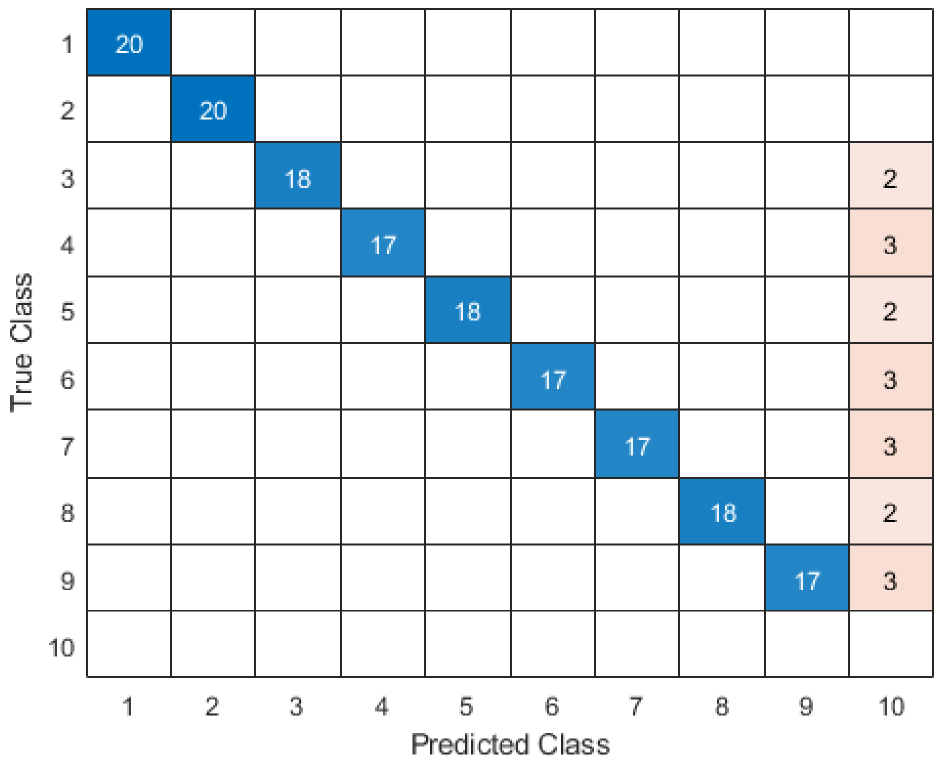

Figure 4.15: Confusion Matrix 1

Here, Class 1 is INR 1, Class 2 is INR 2, Class 3 is INR 5, Class 4 is INR 10 old banknote, Class 5 is INR 10 new banknote, Class 6 is INR 20, Class 7 is INR 50, Class 8 is INR 100, Class 9 is INR 200, Class 10 is Unidentified or Unknown. Class 1 and Class 2 is currency images of INR

1 AND INR 2, which are currency images taken from internet and does not suffer much shadow or illumination problem. On contrary Class 3 to Class 9 are various Indian banknotes that are taken by camera and also have some taken from the internet. Some of the currency images which are taken by camera suffer some issues with shadow and some of them are filthy. That's why some of them are identified as Class 10, i.e., as unidentified or unknown.

The next confusion matrix is of various other currencies of different denominations, most of which are taken from internet due to unavailability in real life. They are identified with an accuracy of 100%.

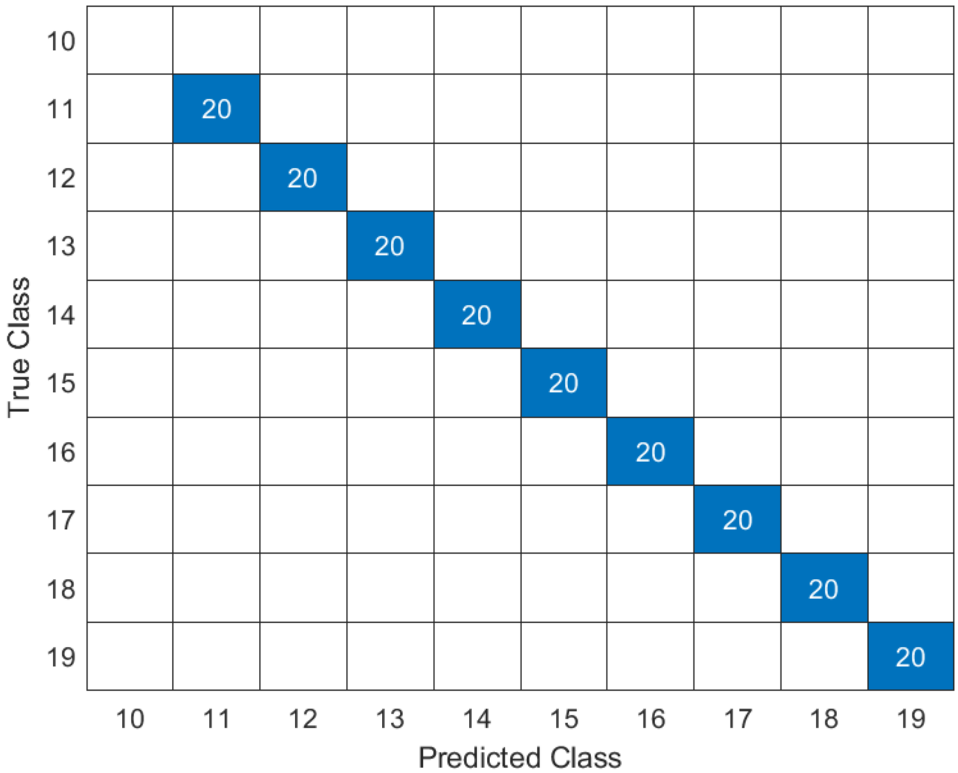

Figure 4.16: Confusion Matrix 1

Here, Class 10 is the Unidentified Banknotes, the number of which is 0.

Class 11 is 5 Euro, Class 12 is Australian 5 Dollars, Class 13 is Bahrain 5 Dinar, Class 14 is Bangladeshi 1 Taka, Class 15 is Bangladeshi 1000 Taka, Class 16 is Canadian 10 Dollars, Class 17 is Canadian 5 Dollars, Class 18 is Saudi 20 Riyal, Class 19 is Swiss 100 Frank. All other currencies also give 100% accuracy when those are tried from different orientations. The overall accuracy of the system is 99.08%.

CHAPTER 5

COMPARATIVE ANALYSIS

In this section, the results from the previous chapter and the results from the existing methodologies are discussed and conclusion is drawn.

5.1 Comparative Analysis 1

[5] proposed a system which deals with multiple templates matching along with mean of RGB values of a banknote image to identify the denominations. The system only works with Indian currency. While working with multiple currencies increase in number of templates to detect one single currency increases the time complexity multifold. Another hurdle is to detect the denominations by mean values of RGB. While working with multiple currencies, there is a huge probability to get almost same value for more than one denomination for various currencies. On top of that, if there is a highly illuminated picture than that of the training picture or a picture taken in absence of proper lights, a huge probability is that the RGB values either stays towards 255 in the histogram or be under 100 in the histogram for the aforementioned cases respectively. In these cases, mean values of the RGB channels differs highly, restraining the system work in any situation. The system does not

deal with orientation of the banknotes; thus, it is highly dependable on a particular orientation. Otherwise, the system fails to identify the currency even if the templates are present in the template database for the query currency image.

The proposed methodology does not deal with a single orientation, thus make the system rotation invariant. It also deals with a single template to identify a single denomination of a particular currency, making it achieve higher correlation value even if there is a difference in illumination in the template image and the query image. The system does not deal with any colour moments directly, but it takes the higher intensity values of every channel to make a comparison within their numbers and then treat them accordingly. For such treatment, the system does not need to match each template with the query currency photo, but a selected templates coming under the selected condition. Thus, the time complexity is handled as much as possible.

The system proposed by [5] achieved an accuracy of 100% while taking Indian currency into consideration. The proposed system achieved an accuracy of 100% while taking 16 different currencies into consideration.

5.2 Comparative Analysis 2

[1] proposed a methodology that involves with 20 different denominations from 3 different currencies i.e., USD, INR and EUR. The methodology simply takes size or area of a banknote as a feature. There are various currencies all around the globe which has similar or even exactly same area or size. So, to differentiate currencies on the basis of its size cannot solve the problem if the system deals with a greater variation of currencies. Dealing with 3 different currencies and taking size as a feature, the system gives a highest accuracy of 65.6%. So, dealing with 16 different currencies would make the accuracy considerably low.

The proposed methodology does not take size or area of a banknote as a feature. So, the accuracy improved which nullifies the discrepancy arose due to same size for different denominations of multiple currencies.

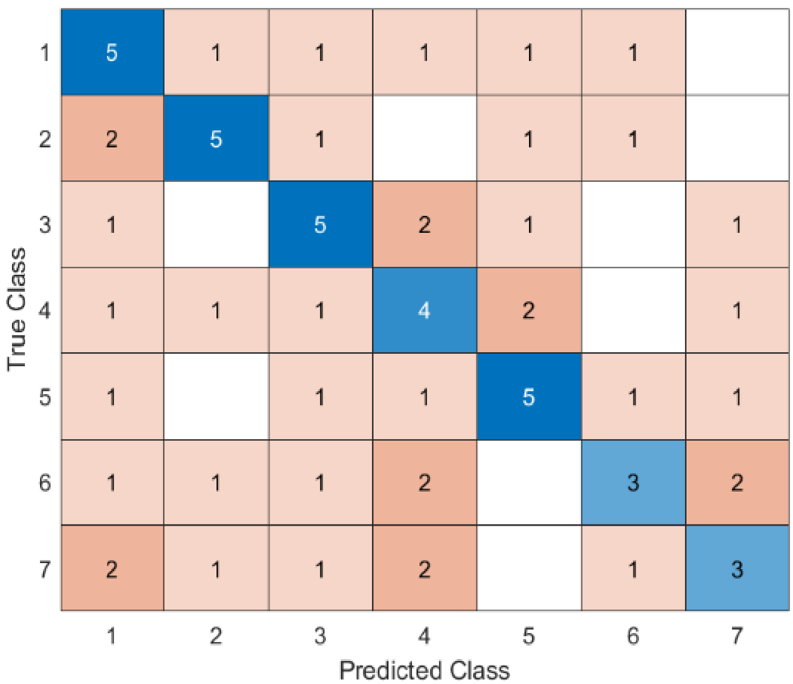

Figure 5.1: Confusion Matrix generated from the Dataset with help of the methodology suggested by [1]

Here, Class 1 is 5 Euro, Class 2 is 10 Euro, Class 3 is 20 Euro, Class 4 is 20 USD, Class 5 is 10 USD, Class 6 is 20 INR, Class 7 is 5 INR. The images are taken from different distances, thus creating various dimensions for the same class and creating a discrepancy in identification.

5.3 Comparative Analysis 3

[11] proposed a system that deals with 20 different types of currencies. The system differentiates the currencies on the basis of empty region on its first level. The empty region is detected by converting the currency image into a binary image taking adapting thresholding into account. While testing, it is found that the recognition of empty region is highly influenced by illumination and noise. If the currency image has higher amount of noise present in it, the empty region calculation goes wrong and it is treated differently on the second level of identification. Thus, the identification goes wrong. Although the system deals with 20 different types of currencies, it only deals with obverse side of any currency, it failed to identify the reverse side of the currency. The system does not take different orientation into account; thus, it fails terribly when the query banknote has some orientations other than 0º.

The proposed system does not take empty regions into account to differentiate between various currencies. So, the discrepancy due to different empty regions influenced by noise is avoided. It does not take size ratio or colour into account as proposed by [11], which helps the accuracy to be improved from 93.3% to 100%.

The System proposed by [11] takes the number of black pixels to number of white pixels ratio of the centre block among the three equally divided blocks as less than 2% to identify the Japanese Yen while the we detect that the centre block of the Japanese Yen has a ratio of black pixels to white pixels as 0.2382, 0.3209, 0.1147 and 0.2012 for different denominations if the bank note is taken from the obverse side with adaptive thresholding. So, if we take 0.3209 as the highest value of the centre block and 0.1147 to the lowest value of the centre block for identifying Japanese Yen, it gives us the following confusion matrix:

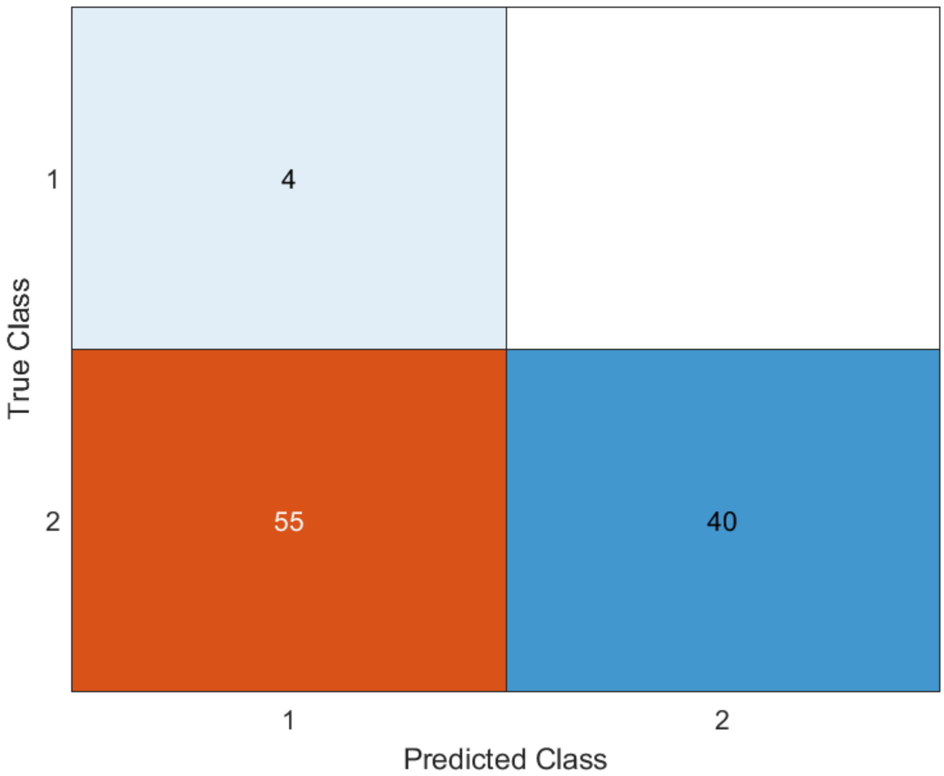

Figure 5.2: Confusion Matrix generated from the Dataset with the help of first phase of the methodology suggested by [11]

Here. Class 1 is Japanese Yen and Class 2 is not the Japanese Yen. While the dataset consists only 4 Japanese Yen taken from obverse, the condition of the first step suggested by [11] cause 59 currency images to identify as a Japanese Yen.

5.4 Comparative Analysis 4

[3] proposed a methodology that deals with Bangladeshi Taka of different denominations where the system uses hue, saturation and value of a RGB image, the edge features and the texture features. While working with a single type of currency, this method performs pretty well, but when it comes to identify multiple denominations of multiple currencies, these features does not work well. As HSV values are highly influenced by Value i.e., brightness of the image, so the feature does not work well if the Value is low. Another concern with the system is while using texture features, when the distance is taken as 1 it leads to reflect the degree of correlation between adjacent pixels (i.e., short range neighbourhood connectivity). On the other hand, increasing the distance value leads to reflect the degree of correlation between distant pixels. Now presence of noise influence GLCM values significantly, so using these values as feature create disparity in identify banknote images of different currencies.

The proposed system does not take HSV or texture as a feature. So, making the system more accurate while working with numerous currencies.

5.5 Comparative Analysis 5

[7] proposed a methodology that employs SURF to recognise Indian Rupee of different denominations. The algorithm of SURF generates different key points and creates descriptors for each key points. This gives a higher accuracy while working with a single type of currency. While working with multiple currencies, it gives error in recognising the key points as it tries to find the descriptors of the key points and working with several currencies create almost same descriptor for some key points. That generates misidentification and creates a higher percentage match even after the currencies belong from different classes. Thus, SURF is not accurate when working with multiple number of currencies. Another requirement for SURF is, to get better accuracy with this feature, a system needs huge dataset so that the key points generated for a class can create unique descriptors. While working with multiple currencies of different countries, it is really very hard to accumulate a huge set of training dataset. It is another major reason why SURF is avoided to be used as a feature.

An accuracy based comparative analysis is shown in table 5.1.

Table 5.1: Comparative Analysis

Ref. No.	**Currency Origin**	**Method of Feature Recognition Used**	**Accuracy Obtained (%)**
[5]	India	Template Matching	100
[1]	Multiple Currencies (3)	Size or Area, Fine-KNN	65.60
[11]	Multiple Currencies (20)	Empty Region Detection, Size Ratio, Colour, Text Extraction, Template Matching	93.30
[3]	Bangladesh	HSV features, Edge features, GLCM, Euclidean Distance	81.27
[7]	India	SURF, Euclidean Distance	96.42
Proposed System	Multiple Currencies (16)	RGB values, Template Matching	99.08

5.6 Contributions and Improvements over earlier approaches

The purpose of research is to improve on previous works while also contributing something new that can be used in real-world situations. Improvements to existing approaches, as well as new concepts have been made in this study for easing the process and making real-time applications more convenient.

I. The proposed system deals with any orientation of the currency image. Thus, it makes the system rotation invariant, which approximately every previously occurred system were unable to deal with.

II. The proposed system can identify the banknote both from the obverse and reverse side. This makes the system more robust to identify a currency.

III. [11] proposed a system that recognises the currencies on the basis of empty region at its preliminary stage, which may vary with proper lighting condition. Ignoring this criterion make the proposed system work more accurately in low light conditions

IV. The proposed approach is better than [1] in terms of accuracy. Only size ratio or area as a feature is not enough to identify multiple currencies by a single system.

V. Although SURF and SIFT is rotation invariant, to extract exact unique key points and a strong enough descriptor, a system needs to have a huge dataset for training purpose. When dealing with multiple currencies of various denominations, it is very tough to gather a huge enough dataset to identify properly with help of SURF and SIFT. The proposed approach is made rotation invariant without help of these features, thus making the data requirement lower.

VI. The proposed system avoids using colour moments and texture moments which creates discrepancy while working with multiple number of currencies of different denominations.

CHAPTER 6
CONCLUSIONS AND FUTURE SCOPES

Following conclusions can be drawn after analysing the experiments and conducting an in-depth study.

6.1 Conclusions

1. There are several features present in an image to identify it like colour moments, SURF, SIFT, LBP, ORB and so on. Of them Template Matching is the most effective way when there are several subclasses present in a single class. Here, the system is designed to identify 16 different currencies from both the reverse and obverse side of the image.

2. Although the accuracy of the proposed system is 100%, the time taken to achieve this accuracy is not optimised.

3. The proposed system even after using template matching gives a result that is rotation invariant. To make the result rotation invariant, the system does not use Rotation and scaling invariant features like SIFT and SURF. Thus, it reduces the requirement of a huge dataset drastically.

4. The proposed system is scaling invariant. Although it works with different scales, it is recommended to use big enough images to avoid discrepancies arose by noise.

6.2 Future Scopes

The future scope related to this study is discussed below.

I. The system is designed for 16 most valuable currencies. It can be extended more to identify more currencies and make it more robust.

II. The System is not optimes for time. To optimise the time with same accuracy, the system can use the algorithm 2 of the proposed method by [11]. The algorithm states to divide the currency image into multiple divisions and check the templates only for that particular divisions, but not for the whole image. If this divide and conquer method applied with this algorithm, the time complexity should be optimised more.

III. The system works with RGB colour space to divide it certain groups, which should be improved by using LAB colour space because a and b value of LAB colour space changes very less in comparison with red, green and blue intensity values when illumination is considered.

IV. The system can be implemented in handheld device and further provide with an audio output to make it usable for visually impaired person.

V. The system can be implemented in currency exchange machines to make the system more automated.

As a result, the preceding points summarise the study's future scope. As can be seen, one of the most prominent and essential research domains in currency recognition have been addressed and enhanced in this study, but there is still room for development, as described above.

REFERENCES

[1] G. Singh Sodhi and J. Singh Sodhi, "A Robust Invariant Image-Based Paper-Currency Recognition Based on F-kNN," 2021 International Conference on Intelligent Technology, System and Service for Internet of Everything (ITSS-IoE), 2021, Sana'a, Yemen, pp. 1-6, doi: 10.1109/ITSS-IoE53029.2021.9615287.

[2] Ingulkar Ashwini Suresh and Prof. P. P. Narwade, "Indian Currency Recognition and Verification using Image Processing", International Research Journal of Engineering and Technology, vol. 3, issue 6, June 2016, pp. 87-91.

[3] J. Akter, M. K. Hossen and M. S. A. Chowdhury, "Bangladeshi Paper Currency Recognition System Using Supervised Learning," 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2), 2018, University of Rajshahi, Rajshahi, pp. 1-4, doi: 10.1109/IC4ME2.2018.8465595.

[4] K. Gautam, "Indian Currency Detection using Image Recognition Technique," 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA), 2020, Gunupur, India, pp. 1-5, doi: 10.1109/ICCSEA49143.2020.9132955.

[5] P. Garkoti, P. Mishra, N. Rakesh, Payal, M. Kaur and P. Nand, "Indian Currency Recognition System Using Image Processing Techniques," 2022 9th International Conference on Computing for Sustainable Global Development (INDIACom), 2022, New Delhi, India, pp. 628-631, doi: 10.23919/INDIACom54597.2022.9763134.

[6] P. Narra and J. S. Kirar, "Indian Currency Classification and Fake Note Identification using Feature Ensemble Approach," 2021 International Conference on Computational Performance Evaluation (ComPE), 2021, North -Eastern Hill University, Shillong, pp. 022-029, doi: 10.1109/ComPE53109.2021.9752109.

[7] R Bhavani and A Karthikeyan, "A Novel Method for Banknote Recognition System", Indian Journal of Computer Science and Engineering (IJCSE), vol. 2, issue 4, pp. 165-167, April 2014.

[8] Shahbaj Khan, Yogesh Vadekar, Shardul Sant and Mahesh Jadhav, "Analysis and Recognition of Currency Notes", International Engineering Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, vol. 2, issue 2, 2016, pp. 479-483.

[9] S. Kamal, S. S. Chawla, N. Goel and B. Raman, "Feature extraction and identification of Indian currency notes," 2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG),

2015, Patna, India, pp. 1-4, doi: 10.1109/NCVPRIPG.2015.7490005.

[10] Sushma R G, Nandeesha K, Shreeharsha S, Srinidhi G and Vijay Raj Gopal R, “Indian Currency Note Denomination Recognition and Validation”, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, vol. 5, special issue 6, July 2016, pp. 58-62.

[11] V. Abburu, S. Gupta, S. R. Rimitha, M. Mulimani and S. G. Koolagudi, "Currency recognition system using image processing," 2017 Tenth International Conference on Contemporary Computing (IC3), 2017, Noida, India, pp. 1-6, doi: 10.1109/IC3.2017.8284300.

[12] World Factbook, CIA https://web.archive.org/web/20150215024949/https://www.cia.gov/library/publications/the-world-factbook//fields/2076.html

[13] Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision **60,** 91–110 (2004). https://doi.org/10.1023/B:VISI.0000029664.99615.94

[14] Bay, H., Tuytelaars, T., & Van Gool, L. (2006). SURF: Speeded Up Robust Features. Lecture Notes in Computer Science, 404–417. doi:10.1007/11744023_32

[15] S A Bhavani. Currency Recognition using SIFT. International Journal of Computer Applications, NY, USA, 167(9):15-20, June 2017.

[16] J. Xu, G. Yang, Y. Liu and J. Zhong, "Coin Recognition Method Based on SIFT Algorithm," 2017 4th International Conference on Information Science and Control Engineering (ICISCE), 2017, Changsha, Hunan, China, pp. 229-233, doi: 10.1109/ICISCE.2017.57.