Detection of Fake Currency using Machine Learning Technique

Akash Rana , Avanish Kumar , Shivam Kumar Jha
(1713101484) (1713101079) (1713101541)
17SCSE101500 17SCSE101086 17SCSE101564
B.Tech(CSE) B.Tech(CSE) B.Tech(CSE)
8th Sem 8th Sem 8th Sem

ABSTRACT

Malpractising has always been a serious challenge that resulted in a serious problem in society. The automation in technology creates a more copied currency that is entirely spread, resulting in reducing the economic growth of the country. The note detection is compulsory, and also necessary to be very consistent and reliable. The paper currency identification depends upon a number of steps, including edge detection, feature extraction, image segmentation, grayscale conversion, and comparison of images. This paper also consists of a literature survey consisting of different methodologies for detection. The review to detect malpractice concludes that whenever we apply some efficient preprocessing and feature extraction techniques, it helps in improving the algorithm as well as the detection system. Machine Learning techniques help in building tools that is required and necessary for the research work, and we can make computer learning design, implementation, and methods to have a difference between fake and genuine currency.

It is going to utilize pattern recognition and image processing learning and analyzing methods for distinguishing features. The paper is performed using anaconda, an integrated development environment of python, and then training is done on the datasets to identify the currency, it provides an advanced features of data analysis as well as an excellent visualization.
1.1 Introduction

Machine Learning techniques help in building applications that support in detection of currency, through automated system and algorithm. Machine Learning is going to use pattern recognition and image processing for analyzing the real characteristics. The aim of this work is to create a paradigm which can be supervised with the help of related set theory so that it can be further beneficial in detecting feigned datasets with a very few categorizing bugs.

Therefore another name referred as categorizing model grouped as data, consisting of attributes and labels for the bills referring as fake or genuine. Moreover it identifies decision boundaries which separate samples of two classes.

Figure 1. Indian Currency Recognition
We firstly take out data from images that rooted out from an original and copied banknote. In regard to digitization, we use camera for printing in terms of inspection. The size of all images have 300x 300 pixels, this is because of the distance concerning lens and grayscale pictures having targeted some dots per inch to achieve. Properties and characteristics are explored and taken out from images in support of transformation concerning wavelet.

Some methodologies that came from wavelet transformation are variance, skewness, kurtosis, entropy of pictures and class relevant to currency. Among these some are continuous in showing the features of a banknote and some ie. the last one actually the real sign of the banknote, indicates 1 as original and 0 for duplicate. The set theory has 1100 samples, 600 samples for feigned notes, remaining 500 copies for original currency.

This work basically aims at observing images taken as input that anonymously based on properties taken out after wavelet transformation and also on the problem based on machine learning. Therefore the processes are continuing with transformation of images.

We check their feasibility by testing the data set in respect of set theory, which will be visible but not mentioned in the model, for achieving accuracy. Set of data is given, by

using their properties we able to design supervised learning model and using that we classify whether the currency is fake or real. For the values that are not mentioned in the data set, we can search them by applying properties and methodologies.

As some of the properties are continuous in nature, normalization is applied to make the data set in the scope of 0 to 1. We cannot neglect any properties having outliers or falsify. In case of applying supervised learning, normalization helps in treating the properties equally and consistently. Other thing is, to establish a benchmark model.

It acts as a naïve classifier and also helps in detection of notes. Few metrics are introduced focusing on feasibility and overall classification.
\[ F_\beta = \left(1 + \beta^2\right) \frac{\text{precision} \cdot \text{recall}}{(\beta \cdot \text{precision}) + \text{recall}} \cdot \]

2. Literature Survey

Detection of fake currency aims at utilizing machine learning techniques using large sets of data as inputs. KNN method is a useful method that is used for analysis. These types of analysis can be very useful for deriving implicit information.

In this chapter, we review the recent available literature on classification and clustering of sequential data. This chapter also presents some of the major application areas of detection of currency. We are using three different types of algorithm KNN method, gradient method and support vector along with the concept of machine learning which helps in detection of currency at a large scale and also make India free from such malpractice.

Ying Li Tian in paper [1], identification of fake note is done for blind through image processing using segmentation. It extracts various features of currency notes with the help of MATLAB software. This enhances simplicity and high performance speed. Li Liu et al In paper [2], detection of fake note is possible using deep learning using SVM and FNN (Feed Forward Neural Network). FNN also uses for verification. It uses max pool operation, suppose the image it extracted, then that image would go through augmentation process and then annotation, these enable database creation, then we input image through real time through transfer learning by alex-network, after that it will go for feature extraction where there the comparison is done between real time and database features. At last on the basis of that it predicts whether the currency is fake or real. Mrs Monali Patil and Prof Jayant Adhikari[3], duplicacy of currency is a vulnerable threat on economy and it is now a common phenomenon due to advanced technology and laser printer, to get rid of this, some methods are processed. Detection of currency is possible through register, watermarking, optically variable ink, security thread, fluorescence, latent image and identification mark. This is really beneficial for banking sector. Author Bo tang, StevanKay in paper [4] describe a novel shape feature using angle distance method. Automated feature selection and automated feature reduction approach for input size detection and text categorization. Mirza and Ninda in paper [5], uses digital image processing technique, for this sobel operator is used with magnitude to make comparison between input image and genuine image and this going to help in identification and also helps many commercial sector.

Nayana Susane Jose in this paper [6] made an android based application for blind people for detection of fake currency, he also proposed denominations for impaired people.

This system mainly termed as android based currency recognition system. Mohammad H Alyshayeji in paper[7], uses bit-plane slicing technique with the help of edge detector algorithm for identification of currency. In paper [8], for roman coins, spatially local coding method is used. Along with that it also uses traditional rigid spatial structure model such as spatial pyramid.
2. Methodology

2.1. Data Preprocessing

Data Preprocessing helps to bring data into good shape, remove any unfunctionality, or modify any properties, before converting it into an algorithm. From the data visualization section, we observed that the dataset does not have any missing value. All the input features are continuous and have different range, so we need to perform normalization to scale all features in the range of 0-1.

2.2 Normalization

We use the following formula to normalize our features.

\[ X_{\text{norm}} = \frac{X_{\text{current}} - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \]

The first five samples after normalization are following table 1.
Table 1. Samples after Normalization

<table>
<thead>
<tr>
<th></th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.789224</td>
<td>0.871699</td>
<td>0.106813</td>
<td>0.786951</td>
</tr>
<tr>
<td>1</td>
<td>0.531528</td>
<td>0.348662</td>
<td>0.310605</td>
<td>0.450440</td>
</tr>
<tr>
<td>2</td>
<td>0.786629</td>
<td>0.810825</td>
<td>0.054924</td>
<td>0.687365</td>
</tr>
<tr>
<td>3</td>
<td>0.757105</td>
<td>0.839456</td>
<td>0.424665</td>
<td>0.736528</td>
</tr>
</tbody>
</table>

### 2.3 Algorithms and Techniques

1. Support Vector Classifier
2. Gradient Boosting Classifier
3. K-Nearest Neighbors Classifier

#### 2.3.1 Support Vector Classifier

In this paper, we shall propose a method by applying the approach of support vector machine to detect feigned notes from the original ones. SVM has a better ability of generalization and higher performance ability regarding pattern recognition as well as classification. DWT (discrete wavelength transformation) is used to reduce the input scale of SVM. It is helpful in high dimensional spaces and also works well with non-linear decision boundaries, in case if it is correct, kernel is used. Many kernel functions are evaluated for the decision function.

#### 2.3.2 Gradient Boosting Classifier

It is used for classification tasks. Machine Learning algorithm considered features as an input that we use to calculate an output value. Gradient Boosting Classifier is interpretable and explainable. It has properties like flexibility, customizability, robustness etc. It also uses machine learning techniques for regression problems i.e. decision tree.

#### 2.3.2 K-Nearest Neighbors Classifier

K-Nearest Neighbors algorithm is a very simple algorithm which is easy to implement, it is robust, multifaceted, and versatile. For Complex classifier such as SVM (Support Vector Machine) and ANN (Artificial Neural Network), it uses benchmark model and the data consisting of many attributes name which is used for classifying images. It uses distance function as well.
3. Results and Discussion

In this paper, we used three algorithms which uses machine learning techniques for classification and regression. Benchmodel, f-beta score evaluation metrices and accuracy we use for training our model. This also helps in analyzing which properties provide the most predictive power. The algorithm mainly used is Support vector, gradient boosting, k-nearest neighbor.

We evaluated the performance of these classifiers using accuracy and f-beta score evaluation metrics. We trained our classifier model on various sizes of training data, to analyze how it affects the prediction scores. Prediction scores were obtained for test data, and for a subset of training data.

We also calculated the time that is taken for training and prediction. Finally, we also refined the selected model using grid search.

Table 2. Train set use for training scores

<table>
<thead>
<tr>
<th></th>
<th>SVC</th>
<th>Gradient Boost</th>
<th>K-Neighbors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train set accuracy</td>
<td>0.9867</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Test set accuracy</td>
<td>0.98</td>
<td>0.9945</td>
<td>0.9982</td>
</tr>
<tr>
<td>Train set f-score</td>
<td>0.9941</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Test set f-score</td>
<td>0.9914</td>
<td>0.9953</td>
<td>0.9992</td>
</tr>
</tbody>
</table>

4. Conclusion

All counts and proven result, it is amazing to notice that algorithm such as support vector, gradient boost and k-nearest neighbor use efficiently the matching learning techniques which beneficially help in classification and regression process. It is noticeable that KNN is a very simple algorithm which is easy to implement, it is robust, multifaceted, and versatile. For Complex classifier such as SVM(Support Vector Machine) and ANN (Artificial Neural Network), it uses benchmark model and the data consisting of many attributes name which is used for classifying images.

It uses distance function as well. Also In this paper, we shall propose a method by applying the approach of support vector machine to detect feigned notes from the original ones. SVM has a better ability of generalization and higher performance ability regarding pattern recognition as well as classification. It is one of the necessary steps for financial implications. According to my store in perception I would like to say, all algorithm and methodology are robust in its performance in respect of training set size and its parameters. The data also play a vital role and we cannot say that detection is only possible with the help of algorithm only.
5. References


