Journal of Science and Technology
ISSN: 2456-5660 Volume 06,Special Issue 01, August 2021,

www.jst.org.in *DOI: https://doi.org/10.46243/jst.2021.v6.i04.pp285-290*

BANK NOTE AUTHENTICATION AND CLASSIFICATION USING ADVANCED MACHINE ALGORITHMS

Prof. Alpana D. Sonje[1], Prof. Dr. Harsha Patil[2], *Prof. Dr. D. M. Sonje [3]
[1]*ACBCS, Nasik, Maharashtra, India*
[2] *ACBCS, Nasik, Maharashtra, India*
[3]*R H Sapat COE. Mgmt Studies and Research, Nasik, Maharashtra, India*
[1]sonje.alpana@gmail.com, [2]harsha.kun.patil@gmail.com,[3]deepaksonje123@gmail.com

To Cite this Article
*Prof. Alpana D. Sonje[1], Prof. Dr. Harsha Patil[2], *Prof. Dr. D. M. Sonje "BANK NOTE AUTHENTICATION AND CLASSIFICATION USING ADVANCED MACHINE ALGORITHMS", Journal of Science and Technology, Vol. 06, Special Issue 01, August 2021, pp285-290*

Article Info
Received: 15.07.2021 Revised: 24.07.2021 Accepted:10.08.2021 Published: 16.08.2021

Abstract*: The currency coins or banknotes used by any country to perform economic activities in the global market should be always genuine. However, some of the miscreants in the society provoke forged notes into the bazaar which bears a resemblance to exactly the genuine note. It is stiff with human naked eye to notify the difference in between these two because they have a lot of analogous facial appearance. Hence, there is a call for of competent validation system which informs accurately whether the note in the transaction is authentic or not. This paper proposes machine learning based methodology to classify the fake and genuine notes. Accordingly, exhaustive testing has been performed using Multi Layer Perceptron Neural Network (MLPNN), Naïve Bayes and Random Forest (RF) Algorithms on the standard data set. The testing results show that RF and MLPNN algorithms gives comparable results in terms of accuracy and other performance measures as compared to any other algorithms. The complete training and testing of the dataset is performed in WEKA software.*

Keywords: *Banknote Authentication; Naïve Bayes; Random Forest (RF).*

1. INTRODUCTION

Currency is the most vital assets of any country. Currencies are printed in various denominations and also in the form of banknotes and coins. Usually bank notes are assigned with high denomination such as notes of Rs. 500 and Rs. 2000 than coins of Rs. 5 and Rs. 10 in India. As Today's market is completely volatile and competitive, to sustain in this global market, some miscreants in the society provoke fake notes into the market. This affects on the overall economy and financial growth of the country. To resolve such sensitive issue, it becomes essential to introduce some mechanism which can detect and isolate the fake note from the bazaar. There are numerous banknote recognition apparatus and techniques are existing which can detect fake bank notes, but till there is scope for improvement as there are limitations on the number of notes for screening. With the naked eye, it is difficult to recognize the genuine and fake note especially when these are in a bundle. These notes are also having analogous facial appearance and features. Hence, it becomes the motivation to the researchers to develop the system which classify and identify whether the particular set of banknotes are authentic or not. Many soft computing approaches have been employed to enhance the accuracy for detection and segregation of fake notes.

Akanksha Upadhyaya et al. have described the general security features used in the various denomination notes in India. The authors have also discussed the literature review on the counterfeit currency detection methods which were composed of different areas of machine learning and image processing. The outcome of the paper is to suggest different methods by different researchers in terms of considering particular characteristics or feature extraction from denomination values of particular country's currency [1].

Sangwook Baek et al. detected the fake banknotes using low resolution multispectral images. Low cost solution is provided based on image processing in which along with the RGB images, three additional infrared images are used for analysis. The authors have claimed very high classification accuracy for genuine banknotes and also for detecting counterfeit banknotes using neural network, likelihood test method and parallel piped classifier on Indian current notes and USD banknotes [2]. The deep learning method is also suggested to detect fake note from genuine notes [3].

Chhotu Kumar et al. used various machine learning methods and observed that decision tree and Multi layer percerptron are more effective techniques for authentication of banknotes. Authors tested the results on hold-out and cross validated data and claiming satisfactory results [4]. The random forest classifier is also proposed for classifying the genuine and fake notes [5].
The authors, proposes two dimensionality reduction approaches with Back Proliferation Neural Network (BPNN) for improving the classification accuracy [6]. Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA) are utilized and accomplished that proposed techniques gives better results than basic BPNN algorithm.
Mrutunjay Singh et al. proposed the robust cost-effective and solution for counterfeit currency detection using image processing techniques which could be deployed for mobile applications. Support Vector Machine (SVM), K- means algorithm and Histogram of Oriented Gradients (HOG) method is used to classify fake and genuine note. The result is presented with the help of some performance measures with overall accuracy [7].
Furthermore, another research employed for the banknote recognition based on the mobile application. The research is two-fold in which first part recognizes the denomination of the banknote and the counterpart executes the authenticity of the banknote. Authors are claiming the better results for cascaded KNN approach than individual KNN approach [8].
The core theme of the paper is to generate a novel and efficient technique based on advanced machine algorithms for classifying the fake and genuine notes. In this scheme, standard data set is used in which statistical features were obtained from images of different number of banknotes. These features are supplied as input to the various algorithms and the performance is observed. In machine algorithms, accuracy and many other performance indices are important to evaluate the performance of the classifier. Consequently, accuracy, root mean squared error (RMSE), root mean squared error (RMSE), mean absolute error (MAE), relative absolute error (RAE), root mean squared error (RMSE), false positive rate (FP), true positive rate (TP), recall, precision, and F-measure are estimated in the research [9].
The remaining part of the paper is structured as follows. In section 2, the generation of dataset and extraction of features is presented. The theoretical information of various classifier algorithms used in the research is briefed in section 3. Results and discussion are elaborated in section 4 and concluding remarks are given in the last section.

2. DATASET DESCRIPTION AND FEATURE EXTRACTION

In this paper, authors are detecting and classifying the various genuine and fake bank notes. The standard dataset from UCI machine learning repository has been used [10]. It has five attributes in which initial four are input attributes which are the statistical features obtained from the bank note and last one is target value. The resolution of the images obtained from industrial camera has 400 x 400 pixels. Wavelet transform is processed on the images to extract the novel features.
In the study, the images of 1372 different denomination notes are considered. Out of which, the genuine class has 762 and fake class 610 cases, respectively. The output class assigned binary values namely 1 & 0, where 1 represents authentic class and 0 represents forged class.
Every statistical feature represents its own distinctive characteristics which is useful to classify the fake and genuine notes. The variance or mean are used to express the probability density function (PDF). These terms are correlated to the spread of a distribution of data around its mean value. The value of skewness factor is close to zero in case if the distribution is close to balanced about its mean. Positive and negative values of skewness factor represents that the distribution has a longer tail to the left and right of the mean respectively. Kurtosis is also another parameter which indicates the proportion of samples which diverge from the mean by a minute value compared with those which diverge by a big value. It is not sensitive to Gaussian noise. Fundamentally, it measures the flatness of a distribution. In case of normally distributed data, the magnitude of kurtosis is zero, where as negative for flatter top and positive for sharper peak than normal distribution. Image entropy is the final feature which tells about the amount of information which must be coded by a compression algorithm.

3. ADVANCED MACHINE LEARNING ALGORITHMS

The numbers of popular machine algorithms are suggested for detecting and classifying the genuine and fake note. For training and testing the datasets, popular WEKA tool (Ver. 3.8.0) is used for number of researchers [12]. WEKA is well known data mining and machine learning tool which is very commonly used in both industrial and academic researcher studies [13]. In this paper, following advanced machine algorithms are used for bank note authentication and classification problem.

3.1 Random Forest Algorithm [11-12,14-15]

The principle of random forest (RF) algorithm is dependent on the collection of trees for regression and classification. In this case, combination of tree predictors and each tree depends on the values of independently sampled random vector with the same distribution for all trees in the forest. The performance of RF classifier is more accurate for large number of trees. The individual training set is accomplished using bagging method. The bagging means extracts a selected quantity from a training set randomly to enhance classification, and regression models as per the proposed stability and classification accuracy. Due to bagging process, variance decreases and it also avoids over-fitting. The overall result is obtained by accumulating the scores of component predictors on each class and then selecting the successor class in the form of the number of votes to it. The errors of classifiers also function of individual tree strength and the correlation between tree strengths.

3.2 Multi Layer Perceptron Neural Network Algorithm (MLPNN) [11-12]

During the training process, networks are fitted to the data by learning algorithms. Supervised learning has been used in the paper in which four features are used as inputs and there are two outputs. MLPNN algorithm is a feed forward network which maps the groups of input data into a set of fitting outputs. MLPNN consists of several layers of nodes in a directed graph, with every layer fully connected to the subsequent node. Other than the input nodes, each node is a neuron with a nonlinear activation function. This algorithm uses a supervised learning method called back propagation for training the network.

In this algorithm, two Sigmoid activation functions are used and are represented by Eq. (1),

$$y(v_i) = \tanh(v_i), \quad y(v_i) = \left(1 + e^{v_i}\right)^{-1} \tag{1}$$

Hyperbolic tangent function is used in the first part in which it varies from -1 to 1, and the second case, logistic function is used, is similar in contour but varies from 0 to 1. Here, y_i = output of the i^{th} node, v_i = weighted sum of the input synapses

The back propagation is used to optimize the error. The error in output node j in the n^{th} data point is calculated using Eq. 2

$$e_j(n) = y_j(n) - d_j(n) \tag{2}$$

where d is the desired value and y is actual the value produced by the algorithm.

The error is optimized in the entire output given by Eq. (3),

$$\varepsilon(n) = \frac{1}{2}\sum e_j^2(n) \tag{3}$$

3.3 Naïve Bayes Algorithm

If the features are not related each other in every class then Naïve Bayes Algorithm can be used. This classifier is working on determining $P\left(X/Y\right)$, i.e. probability of X given that Y is occurred. There are two stages of classifier in which first is training stage and later is prediction stage. The parameters of probability distribution are evaluated in case training stage provided that the features are conditionally not related to each other. For unknown test data and new data, later stage estimates the posterior probability belonging to the each class. Then it classifies as per maximum posterior probability. The functions like kernel, normal, multinomial, multivariate multinomial distributions are supported in Naïve Bayes Algorithm.

4. RESULT AND DISCUSSION

The standard data set from UCI machine learning repository has been taken for testing purpose. In data mining process, performance of the classifier can be studied on hold-out and cross validated (CV) data. In

Hold-out method, data set is divided into two parts. First part of the dataset is used for training and remaining is used as test data set. The dataset is divided into number subsets of same size in cross validation. Normally, 5 fold or 10 fold validation is used.

In the paper, the performance of the algorithms is checked on the dataset produced from different 1372 bank notes as mentioned earlier. In the hold-out method, 80% data is selected for training and 20% is used for testing for which the performance of the algorithm is checked. In the CV method, the algorithm is tested on 10 fold cross validation data. The results of the classifiers are compared and shown in Table 1-3. The performance measures such as F-measure, precision, recall, true positive rate (TP) and false positive rate (FP) are listed in Table 2. The confusion matrix which hints the about the number of correct and false results is presented in Table 3.

From Table 1, it is observed that the overall accuracy of the RF algorithm and MLPPN algorithm is 99.27% and 99.63% respectively, which is very high as compared to Naïve Bayes algorithm in which only 83.94% accuracy is obtained. The different errors like MAE, RMSE, %RAE, and %RRSE are also evaluated and giving satisfactory values in case RF algorithm and MLPPN algorithm as compared to Naïve Bayes Algorithm. The additional performance parameters like TP rate, FP rate, precision, recall and F-measure are also compared as shown in Table 2. The obtained results show the values of these measures are much closed to the expected values in case of RF and MLPNN algorithm as compared to Naïve Bayes algorithm. The confusion matrix is represented in Table 3. In this case also, RF algorithm and MLPNN algorithm is showing more accurate results and these classifiers has been more suitable for the detecting any fake note in currency.

Table no 1. Comparison of Various Machine Algorithms

Parameter	**RF Algorithm**		**MLPNN Algorithm**		**Naïve Bayes Algorithm**	
	10 Fold C. V.	**Hold Out (80:20)**	**10 Fold C. V.**	**Hold Out (80:20)**	**10 Fold C. V.**	**Hold Out (80:20)**
%Accuracy	99.1983 %	99.2701%	99.9271 %	99.635%	84.2566 %	83.9416%
MAE	0.018	0.0194	0.0048	0.0101	0.1894	0.1927
RMSE	0.076	0.0783	0.0333	0.0659	0.3238	0.3298
% RAE	3.6467 %	3.8726	0.98 %	2.0318 %	38.3423 %	38.7909 %
% RRSE	15.2882 %	15.5028 %	6.6974 %	13.1583 %	65.1596 %	65.8595 %

Table 2. Detailed Accuracy by Class

Parameter	**RF Algorithm**		**MLPNN Algorithm**		**Naïve Bayes Algorithm**	
	Hold Out	CV	Hold Out	CV	Hold Out	CV
TP Rate	0.993	0.992	0.996	0.999	0.843	0.839
FP Rate	0.007	0.007	0.004	0.001	0.167	0.164
Precision	0.993	0.992	0.996	0.999	0.843	0.840
Recall	0.993	0.992	0.996	0.999	0.843	0.839
F-Measure	0.993	0.992	0.996	0.999	0.842	0.839

Table 3. Confusion matrix

Classified Output	**RF Algorithm**				**MLPNN Algorithm**				**Naïve Bayes Algorithm**			
	Hold Out		**CV**		**Hold Out**		**CV**		**Hold Out**		**CV**	
	G	**F**	**G**	**F**	**G**	**F**	**G**	**F**	**G**	**F**	**G**	**F**
G	142	2	754	8	144	0	762	0	125	19	672	90
F	2	128	3	607	1	129	1	609	25	105	126	484

G Genuine F Fake

5. CONCLUSION

The authors have proposed novel and accurate approach for fake bank note detection and classification based on advanced machine algorithms. The standard dataset of UCI laboratory is used for research in the paper. Four statistical features obtained from the images of the notes are used as the input attributes to the algorithm and one output attribute is used which predicts the condition of the note. Exhaustive testing has been conducted with various machine algorithms available in WEKA environment and the best performance result of each algorithm is presented in the result and discussion section. It has been observed from the obtained results that RF algorithm and MLPNN algorithm performance is more accurate than the Naïve Bayes algorithm. The additional performance parameters like TP, FP, Precision, recall and F-measure also improved significantly in case of these algorithms. In the case of cross validated data and for hold-out data sets, the accuracy is almost constant. In the future, the performance of the proposed algorithms can be improved by using the combination of various algorithms.

REFERENCES

[1]	Akanksha Upadhyaya, et al. (2018), "Counter Feit Detection Techniques", IEEE International Conference on Cloud Computing, Data Science & Engineering, pp. 394-398.
[2]	Sangwook Baek, et al. (2018), "Detection of counterfeit banknotes using multispectral images", Digital Signal Processing, vol .8, pp.294-304.
[3]	Ms. Megha Jadhav, et al. (2019), "Currency Identification and Forged Banknote Detection using Deep Learning", IEEE International Conference on Trends and Advances in Engineering and Technology (ICITAET) pp. 178-183.
[4]	Chhotu Kumar, Anil Kumar Dudyala (2015), "Bank Note Authentication Using Decision Tree rules and Machine Learning Techniques", IEEE International Conference on Advances in Computer Engineering and Applications (ICACE-2015), pp. 310-314.
[5]	Rishabh Jaiswal, Suhani Jaiswal (2019), " Banknote Authentication using Random Forest Classifier", Journals for International Shodh in Engineering and Technology. vol. 4, no.4, pp.1-4.
[6]	Ravi Kumar G, Nagamani K (2018), "Banknote Authentication System Utilizing Deep Neural Network With PCA And LDA Machine Learning Techniques", International Journal of Recent Scientific Research, vol. 9, no. 12(D), pp. 30036-30038.
[7]	Mrutunjay Singh et al. (2018), "Image Processing based Detection of Counterfeit Indian Bank Notes" , IEEE International Conference on Computing Communication and Network Technologies.
[8]	Tamarafinide V. Dittimi, et al. (2018), "Mobile App for Detection of Counterfeit Banknotes", Springer International Publishing AG, part of Springer Nature, pp. 156–168.
[9]	Harsha Patil, R. S. Thakur (2018), "A semantic approach for text document clustering using frequent item sets and Word Net" , International Journal of Engineering & Technology, vol. 7, no. 29, pp. 102-105.
[10]	UCI machine learning repository, https://archive.ics.uci.edu/ml/machine-learning data bases/00267/data_ banknote_authentication.txt
[11]	Deepak Sonje et al. (2017), "A Novel Approach for Multi Class Fault Diagnosis in Induction Machine Based on Statistical Time Features and Random Forest Classifier", IOP Conf. Ser.: Mater. Sci. Eng.
[12]	http://www.cs.waikato.ac.nz/me/weka
[13]	Frank, E, Hall, M, Trigg, L, Holmes, G, witten, I. H. (2004), "Data Mining in Bioinformatics using Weka".
[14]	Deepak Sonje et al. (2017), "An Innovative Approach for Multiple Faults Detection in Induction Motor Using Statistical Time Measures and Random Forest Classifier", International Journal of Control Theory and Applications, vol.10, no.6, pp. 317-325.
[15]	Raj Kumar Patel, V.K. Giri (2016), "Feature selection and classification of mechanical fault of an induction motor using random forest classifier", Elsevier Engineering and Material

	Sciences, vol.8, 334-337.