CNN based framework for identifying the Indian currency denomination for physically challenged people

P.Selvi Rajendran¹, Dr. T.P. Anithaashri²

¹Hindustan Institute of Technology and Science, Chennai
²Institute of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai - 602 105, India.
Email: selvir@hindustanuniv.ac.in

Abstract: One of the premier issues confronting the visual hindered individual is money, acknowledgment especially for cash. Be that as it may, the outwardly weakened individual may not think about the estimation of cash, and they endure part in cash trade related issues in their normal life. To address this issue, we have built up a framework for acknowledgment of money, notes, which might be the helpful device for an outwardly debilitated individual. Investigation and trials were done on the money informational collection, which encouraged CNN dependent on the key highlights, for example, watermarks, pictures printed on cash, esteemed composed as words and numbers and the total cash. This paper deals with the utilization of Convolutional Neural Networks (CNNs) for solving this society issues and investigations about the exhibition and evaluation of different CNN models. Here, Alexnet, Googlenet, and Vgg16 models have been considered for assessment. All the models were adjusted as far as preparing and testing the individuals of data sets. Among these three models, Alexnet accomplished better in preparing fulfillment, Vgg16 model indicated the better execution and accomplished 100%, Google net arrives at 88% as far as productivity.

1. Introduction

According to the Indian authorities, the big proportion of visually impaired people became higher [1]. About one hundred and sixty-five people were visually impaired in line with lakh individuals. Blind people were eighty-two percent of them, and low vision was 18 percent [1]. The latest smartphone development makes it possible to identify currency with an attractive one. In this research work we apply the convolutional neural network models such as Alexnet, Googlenet and VGG16 to detect the value of the currency notes.

In the present situation, the reputation for currency denomination is becoming a vibrant topic for scientists in multiple apps such as blind person currency recognition and fake currency detection for normal people and apps linked to transactions. An important aspect of our normal life is money transactions. Each nation has a distinct currency layout and value. The distinctive skills, colorations, denominations, and the global value for their currency have been retained by each individual county. The denomination can be acknowledged for foreign cash without trouble, but from the real currency, it will become difficult to interpret a fake currency. Particularly suffer in cash exchanges for the visually impaired people. They may not now be ready to properly comprehend one-of - a-kind denominations and are often betrayed by distinct individuals.

We speak to an application that distinguishes cash by means of an application and dispatches the result by means of sound frameworks. One of the essential issues confronting outwardly weakened individuals
are the failure to see paper monetary forms because of the guess of the paper surface and span between excellent monetary and standards. Consequently, this present framework's capacity is to build up a disposition to taking care of this issue so as to cause dazzle people to feel safe and take care of the financial issues

Regarding picture inclusion, there are two critical fields of cash acknowledgment (i) scanner-based methodology and (ii) camera-based methodology. The scanner-based system takes a picture of the report all in all (like a scanner). These frameworks are proper for money counter offices. At the same time with camera-based methodology catching the paper utilizing a versatile camera that can be part of the report. The scanner-based technique is tended to by the most extreme related writing work[ 2-5]. It is affirmed to permit clients to catch any portion of the currency note by their advanced smartphone and permit the framework to inform the expense of the cash to visually disabled persons for the utilization of unmistakable weakness of blind individuals.

Right now, based Indian paper money is educated to analyze the utilization of profound learning models that have the picture preparing instruments that could make taking care of time extremely quick with the right exactness. The proposed structures are fit for taking care of in part and under unmistakable lighting conditions caught records.

Utilizing distinctive convolution neural system models, camera-based Indian paper cash is recognized on this exploration work, making preparation time extremely short with alluring exactness contrasted with the previous techniques referenced in the writing study. The recommended procedure can treat the incompletely and under unmistakable light conditions caught fiscal reports.

The rest of the paper is organized in the following way; Section 2 presents work of writing connected to consider research on cash recognition and versatile applications. The subtleties of the proposed framework and techniques are discussed in Section 3. The details of the data set and the collection of currency and coin data set describes in section 4. In section 5 explains popular CNN models available for image processing. Results and analysis are discussed in section 7.

2. Related Works

This section discusses works associated with literature. In recent years, various bank currency identification strategies or algorithms have been used to recognize and classify banknotes from various nations.

**Computer vision:** The writers acknowledge currencies based on computer vision in their approach [6]. They intended a scheme for recognizing four distinct currencies using extraction methods of characteristics. The main characteristic considered were currency texture, color, and forms. For classification, they used Artificial Neural Network. The median level of accuracy was 93.84%.

**SIFT Algorithm:** Iyad and their team created a visually impaired portable currency identification scheme and used Jordanian currencies as a test dataset [2]. Their strategy is based on the algorithm of
the invariant transform function (SIFT) scale [10]. It attained the accuracy level of 71% to recognize the currency and 25% for the coin type of currency.

**Radial Basis Function Network:** The author has created a portable paper currency identification scheme in [15,19] and has been tested on Saudi Arabia documents. The currencies were recognized here based on picture characteristics and correlations between two pictures. For classification, it utilizes the Radial Basis Function Network. The scheme is 95.37 percent accurate for Normal Non-Tilted Images, 91.65 percent for Noisy Non-Tilted Images, and 87.5 percent for Tilted Images.

**Matching Template:** Sungwook and their team have created an algorithm [4,18] which can recognize currencies from different countries. This technique identifies currencies based on the distinct currency note size and the multi templates correlation matching value. They achieved 100% precision in classifying and identifying the currencies of five distinct countries.

**Image processing:** Neoura has proposed the strategy which used image processing characteristics to define the currency [7,8,17]. Image segmentation, equalization, region of interest (ROI) extraction and then matching the image based on template, correlations between the images taken and the database dataset were the primary operations applied to the picture on their research work. This strategy has accomplished the precision of 89 percent in recognizing the currency. The information set for the Egyptian currency was used to test the scheme.

**LBP Algorithm:** Junfang’s approach identified the currency based on the characters extracted from the image [14,16] and used the Local Binary Pattern approach to process the image. It achieved the precision of 92 percent.

In this paper, convolution neural networks are used for the classification of bank currency notes, labeling the boxes into which the input image is divided into in the network, in this case 11 X11 layers are used. Each cell has the probability of being one of the 5 or 7 classes of Indian banknotes and Coins, respectively, and the network also performs a segmentation of the object or banknote, enclosing it with a bounding box in the image.

### 3. Proposed Approach

This work presents a system to detect and recognize banknotes from different denominations using convolutional neural networks with deep training. It involves the following processing steps.

1. The proposed convolution neural network is trained using large collection of images taken in the different lighting condition such as natural and artificial lighting effects, with variations in the lighting and different rotation. Here, \((x, y)\) coordinates represents the predicted bounding box position for the given input image and \((\hat{x}, \hat{y})\) represents the actual or original positions from the training data.

2. To deal with the error in deviation among bounding box and actual marked place, instead of calculating the width and height directly the square root of the predicted bounding box width and height is calculated.

\[
\sum_{i=0}^{s} \sum_{j=0}^{B} \prod_{ij}^{obj} \left( C_i - C_{\hat{i}} \right)^2 + \lambda_{\text{noobj}} \sum_{i=0}^{s} \sum_{j=0}^{B} \prod_{ij}^{obj} \left( C_i - C_{\hat{i}} \right)^2 
\]  

(1)  

Loss function -1
The parameter $\lambda$ represents here to utilize the initial segment contrastingly and to assign the weight portions of the misfortune capacities. This is important to expand the soundness of the model. The most noteworthy punishment is for arranging forecasts ($\lambda_{\text{coord}} = 5$) and the least for certainty expectations when no item is available ($\lambda_{\text{noobj}} = 0.5$).

3. Loss related with the certainty score for each bounding box indicator is computed. $C$ is the certainty score and $^\wedge C$ is the Crossing point over Union of the anticipated bounding box with the ground truth. $1_{\text{obj}}$ is rise to one when there's a question within the cell, and something else. $1_{\text{noobj}}$ is the opposite.

4. Another portion of the loss function is the object classification loss. The term $1_{\text{noobj}}$ is utilized to penalize a classification error when no object is show within the cell.

$$\sum_{i=0}^{S^2} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2$$

Loss function -II

![Figure 1 Pre-trained Model of the Convolutional Neural Network.](image)

The following step was carried out during CNN Training.

1. First, the CNN model should be pre-trained with the primary 20 convolutional layers. The ImageNet 1000-class dataset has been used for training with a input measure of 224x224
2. In the next step, input resolution of the image has been increased to 448x448
3. The complete network has been trained for the rate of parameter setting are almost 135 epochs employing a batch measure of 64, momentum of 0.9 and decay of 0.0005
4. Learning rate plan: for the primary epochs, the learning rate was gradually raised from 0.001 to 0.01. Train for almost 75 epochs and after that begin diminishing it.
5. Use information expansion with arbitrary scaling and interpretations, and arbitrarily altering presentation and saturation.

This framework also follows the same method of using the rectified linear units (ReLUs) through all of the CNN structures, and the completely-connected layers. This activation feature is popularly utilized in CNN because it works as a rectifier that lets in the simplest advantageous values to skip and enables to enhance the training pace as well as keep away from the gradient-vanishing, as compared to sigmoid function.
Figure. 1 suggests the model of the network architecture used for the recognition of Indian banknotes. The network used to categorize banknote denominations consists of nine convolutional layers accompanied by 2X2 Maxpool layers, 3 fully-connected layers and an exit or detection layer. This structure is based totally on and trained in the darknet framework. The memory used for every photograph is about 34.6 MB and the Parameters for each photograph for ahead propagation by myself is about 46.5 million.

The output of this version is a 7 X 7 X 15 tensor for the Indian banknote community and a 7 X7 X 17 tensor for the Indian coin network. This tensor can be considered a 7 X 7 grid representing the input picture, where each tensor cell may have the two field definition. Every container carries the parameters of x, y, width, top, and the confidentiality of items belonging to the lessons.

The network was trained with the hardware components of NVIDIA Geforce GT 740 SC graphics card (2GB VRAM), set up in a laptop with a i7 three.50 GH central processor with 8GB in RAM for approximately Five days.

4. Data Set Acquisition

The different categories of Indian currencies differs in value estimation and color usage, separated from the quality of printing, material used for printing and other which makes for simple visual distinguishing proof. In any case, for the visually disabled person, the content and color will not give the assistance at all and measure can lead to disarray since of the comparable measurements of the different coins. There are no benchmark datasets are accessible, due to this, the modern dataset has been made which comprises of 150 pictures of the currency note and coins. The 20% of the dataset is utilized as the test set for fine-tuned CNNs models, i.e. Alexnet, Googlenet, and Vgg16. All the dataset contains pictures of the front and the back sides of the Indian cash notes: 5, 10, 20,50, 100,200,500 and 1000 rupees, in expansion to 1, 2, 5, 10, 25 paise, 50 paise and 1 rupee coins, 10 pictures have collected of each denomination. Figure 2 and 3 shows the picture of the dataset.

![Sample data of Indian Rupees](image)

The above Figure 2. Shows different pictures of Indian currency dataset which includes the backside and front side of the various Indian currency denomination images. All the pictures have been taken using high configured cameras, with high resolutions like VGA, 1.3 megapixels (MP), 2 MP and 5 MP.
are the default settings of the cameras. For each cycle, there are 4 different half folds and 2 full-length setups was made. For each group we consider at slightest 12 different currencies, over 6 distinctive indoor situations and 7 different open air situations, whereas collecting the dataset. This presents numerous varieties in brightening, foundation and posture within the dataset. The dataset contains pictures of both modern and old currency, as well as currency with writes on them.

![Sample Indian coins](image)

**Figure.3** Sample Indian coins

### 5. Popular CNN Models

Generally, A Convolutional Neural Network prepare (CNN) fashions is the mixture of multi-layer neural structures, and it's far committed to recognizing the image highlights of pixel making utilize the shape of their pixels. The slicing edge and normally utilized CNN fashions are AlexNet, VGG16/19, GoogLeNet, and ResNet. In this section, the most pertinent models to recognize the currency, AlexNet, GoogLeNet, and VGG 16 have been taken into consideration for getting ready and testing the facts.

#### 5.1 Alexnet

AlexNet is a form of convolutional neural network and it is already well trained in the millions of various categories images from the ImageNet database [1]. This network structure have become one of the first deep neural networks to drive ImageNet magnificence accuracy through a big stride in evaluation. Its architecture framework composed of five convolutional layers accompanied through three completely connected layers, as depicted in the figure 1.

This type of network will have 8 deep layers and it may classify the input images into 1000 of categories, which include keyboard, mouse, pencil, and plenty of animals. As an end result, the Alexnet has observed out wealthy function representations for an extensive kind of input images. The network has a picture input duration of 227 by 227.

This Alexnet shape has changed into one of the initial level deep neural networks to classify the image net class accuracy by way of the use of a big stride in evaluation to traditional methodologies. It's far protected five convolutional layers followed by using 3 dedicated layers, as demonstrated in the figure 4. The depth of the networks is eight layers deep and may classify photographs into one of the thousand classes, in conjunction with keyboard, mouse, pencil, and lots of animals. As a result, the network version has educated and determined the rich feature illustration of a large variety of images. The length of the architecture is 227 by 227.
5.2 Google Net

This network also has the similar shape of CNN version and it used a CNN inspired by way of LeNet however executed a completely unique detail that is dubbed an inception module. It used batch normalization, image distortions and RMSprop. This module is based on numerous very small convolutions as a way to substantially lessen the quantity of parameters. Their form consisted of a 22 layer deep CNN but decreased the large form of parameters from 60 million (alexnet) to four million.

5.3 VGG Net

This model is made up of sixteen convolutional layers and may be very appealing due to its keeping the very uniform structure of some of the layers. Just like AlexNet, this network additionally has most effective 3x3 convolutions, however masses of filters. The weight configuration and parameter values of the VGGNet is publicly available for research purpose and has been utilized in lots of various applications and challenges as a baseline characteristic extractor. However, VGGNet consists of 138 million parameters, which may be hard to address.
5.4 ResNet-50

ResNet-50 is also the kind of convolutional neural network that is trained on more than 1,000,000 pics from the ImageNet database [1]. This type about 50 layers deep and can classify images into one thousand item classes, which incorporates keyboard, mouse, pencil, and masses of animals. As a result, the network has found out wealthy characteristic representations for a big variety of pictures. The network has a photograph input duration of 224 by 224.

5.5 Fine-tuning

The closing three layers of the CNNs which are already pre trained are retained as a fixed characteristic extractor for our dataset. The first step in switch mastering is mainly for replacing the last 3 layers in the pre-educated CNN with a hard and fast layer that may classify our targeted 11 instructions.

6. Results

This segment gives points of interest around preparing and testing the three fine-tuned models. To begin with, Figure 7 appears to be performance-based precision for preparing Vgg16, Alexnet, and Googlenet separately. It clearly appeared that Vgg16 illustrates the most extreme execution compared to Alexnet and Googlenet. To differentiate, Vgg16 required a longer time to total the preparation. In terms of learning speed, Googlenet shows up to be speedier than Alexnet and Vgg16 as outlined in Figure 7. This figure gives more subtle elements approximately the fine-tuned CNNs misfortune work.

![Classification performance of AlexNet and GoogLeNet trained](image)

**Figure.7** Classification performance of AlexNet and GoogLeNet trained

6.1 Experimentation

In this section, we first introduce the experiment procedure after which, talk the outcomes to prove the robustness and sensitivity of our approach. In this connection, the android utility has been designed for visually affected customers, there may be no symbols or buttons and no guide preparations needed. The application starts perusing the database of the spare cash banknotes at that point and begins taking pictures of the fronted scene. Once both flat edges are recognized a vibration effect takes place to tell the customer about the cash detection. After a while, a sound articulates the proper esteem in english is
listen. Figure 8 appears to be the application running on Samsung GT-N800 tablet testing Rs.2000/- cash paper.

6.2 System Evaluation

This developed framework has been assessed using Matlab software tools based on accuracy measurement. The formula which is used to evaluate the system are

\[
\text{Accuracy} = \frac{\text{True}}{\text{True} + \text{False}} \times 100
\]

In the above equation, true represents both genuine positive and genuine negative comes about whereas untrue shows the inverse. Table 1 appears to be the accuracy of our framework calculated for the 120 test data pictures (20 tests for each category). The main observation from table 1. are the highest accuracy obtained is 96% and it achieves to recognize the currency value 20 and 1000. The lowest accuracy comes about for the 50 and 500 rupees. It’s gathered that this value was brought about due to the truth that within the 50 currency paper, the beat right corner has exceptionally few points of interest which may cause the blame proportion. The normal running time for Matlab framework is 10 seconds whereas in android framework is 12 seconds.

| ₹    | 5  | 10 | 20 | 50 | 100 | 200 | 500 | 1000 |
|------|----|----|----|----|-----|-----|-----|------|
| Quantity | 50 | 50 | 50 | 50 | 50  | 50  | 50  | 50   |
| True   | 46 | 47 | 48 | 44 | 46  | 47  | 44  | 48   |
| False  | 4  | 3  | 2  | 6  | 4   | 3   | 6   | 2    |
| Accuracy(%) | 92 | 94 | 96 | 88 | 92  | 94  | 88  | 96   |

7. Conclusion

In this paper, Indian cash recognition framework has been proposed. The proposed framework begins with capturing still pictures. Basic picture handling strategies like thresholding, noise evacuation, histogram equalization and division are utilized to extricate the ROI and encourage the layout coordinating method. Correlation based format coordinating is utilized after that to discover the ROI within the dataset pictures. This system has been developed in Matlab and OpenCV library beneath Android stage as well. The Matlab framework considers offline captured pictures whereas the Android one was planned to coordinate visual disabled clients. Live video capturing is considered for the Android application. This paper furnished a proof of the utilization of the idea transfer getting to know for Banknote recognition. Indian currency papers have been used for training and testing to tune the network models, namely Alexnet, Googlenet model, Vgg16 confirmed an terrific performance compared to Alexnet and Googlenet. While Vgg16 achieved the 100% accuracy whilst it is ninety five percent achieved by Alexnet and Googlenet attains 88% of accuracy.

In future, after making the final model which already trained with datasets, an app will be made and embed the version offline in order that a blind individual utilizing this app gains an interface to the net on every occasion they should apprehend the notes they're carrying. This framework could be improved as to use the classification technique to examine the authentic or foreign money notes on a low end versatile smartphone for visually disabled humans and inform clients through voice notice in local language. In destiny it could be expanded to apprehend foreign foreign money.
References

[1] Noura A, Semary Sondos M. Fadl, Magda S. Essa, Ahmed F. Gad, "Currency recognition system for visually impaired: Egyptian banknote as a study case", 5th International Conference on Information & Communication Technology and Accessibility (ICTA), 2015.

[2] Bruce Poon, Md Ashraful Amin, "Recognizing Bangladeshi Currency for Visually Impaired", Conference Paper in Communications in Computer and Information Science, July 2014, DOI: 10.1007/978-3-662-45652-1_14

[3] Snehal Saraf, Vrushali Sindhikar, Ankita Sonawane, Shamali Thakare, "Currency Recognition System For Visually Impaired", IJARIE-ISSN(O), vol. 3, no. 2, pp. 2395-4396, 2017.

[4] Ahmed Yousry, Mazen Selim, Mohamed Taha, "Currency Recognition System for Blind people using ORB Algorithm", International Arab Journal of e-Technology, vol. 5, no. 1, pp. 34-40, January 2018.

[5] Mahesh Pawar, Ketan Desai, Sanjay Gupta, Trupti Chavan, Kirti Mhamunkar, "Object Detection and Currency Recognition Using CNN", Cikitsi Journal For Multidisciplinary Research, Vol. 6, no. 4, pp. 357-362, April 2019.

[6] Ali Abd Almisreb, Mohd. A. Saleh, "Transfer Learning Utilization for Banknote Recognition: a Comparative Study Based on Bosnian Currency", Southeast Europe Journal of Soft Computing, vol. 8 no. 1, pp. 28-31, March 2018.

[7] Srivastava, N., Geoffrey Hinton, A. K., Sutskever, I., Salakhutdinov, R. Dropout, "A simple way to prevent neural networks from overfitting", Journal of Machine Learning Research, Vol. 15, pp. 1929–1958, 2014.

[8] Tong, C., Lian, Y., Qi, J., Xie, Z., "A novel classification algorithm for new and used banknotes. Mobile Networks and Applications", vol. 22, no 3, pp. 395–404, 2017.

[9] V. D. Hoang and H. T. Vo, “Hybrid discriminative models for banknote recognition and anti-counterfeit,” NICS 2018 - Proceedings of 2018 5th NAFOSTED Conference on Information and Computer Science, 2019.

[10] V. Abburu, S. Gupta, S. R. Rimitha, M. Mulimani, and S. G. Koolagudi, “Currency recognition system using image processing,” 10th International Conference on Contemporary Computing, IC3 2017, 2018.

[11] Y. Liu, J. He, and M. Li, “New and Old Banknote Recognition Based on Convolutional Neural Network,” ICoMS 2018: Proceedings of the 2018 International Conference on Mathematics and Statistics, pp. 92–97, July 2018.

[12] J. M. Jiao, J. He, and B. Zhang, “Folding Paper Currency Recognition and Research Based on Convolution Neural Network,” International Conference on Advances in Computing, Communications and Informatics, ICACCI 2018, 2018.

[13] H. C. Shin et al., “Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning,” IEEE Trans. Med. Imaging, 2016.

[14] T. D. Pham, D. E. Lee, and K. R. Park, “Multinational banknote classification based on visible light line sensor and convolutional neural network,” Sensors (Switzerland), vol. 17, no. 7, 2017.

[15] P. Selvi Rajendran, "Virtual Bulletin Board using Man-Machine Interface (MMI) for Authorized Users”, Indian Journal of Science and Technology, vol. 12, no. 34, September 2019.

[16] P. Selvi Rajendran, "AREDAI: Augmented Reality based Educational Artificial Intelligence System, International Journal of Recent Technology and Engineering (IJRTE)", vol. 8, no 1, May 2019.

[17] P. Selvi Rajendran, "Virtual Information Kiosk Using Augmented Reality for Easy Shopping", International Journal of Pure and Applied Mathematics (IJPAM), Vol. 118, no. 20, pp. 985-994, 2018.
[18] H. Hassanpour, Payam Masoumifarahabadi,” Using Hidden Markov Models for paper currency recognition”, Expert Systems with Applications, vol. 36, no. 6, pp. 10105-10111, August 2009.

[19] Manish Shankar Kaushik, Aditya Bihar Kandali,” Recognition of facial expressions extracting salient features using local binary patterns and histogram of oriented gradients”, International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), 2017.