International Journal of Future Generation Communication and Networking
Vol. 13, No. 1s, (2020), pp. 24- 31

Artificial Intelligence Digital Assistant For Visually Impaired People

Mr. Vishal Patil[1]
Assistant Professor, Computer Engineering Department, K. J. Somaiya Institute of Engineering and Information Technology, Mumbai, India

Mr. Yogesh Jadhav[2],Dr. DeepaParasar[3]

Assistant Professor, Computer Engineering Department, Amity University, Mumbai, India

Mr.Ajay Sirsat[4]

Assistant Professor, Computer Engineering Department, Sent John College Of Engineering and Management,, Mumbai, India

Abstract

The World Health Organization states that, in 2017, 253 million people face vision impairment: 36 million are blind and 217 million have moderate to severe problem for vision impairment .In India currently around 12 million blind people are there, which puts India home to one-third of the world's blind population which is not a good number. We want to assist visually impaired people living in India by using the camera of an Android phone to capture images, identify them and provide information related to its which would improve their day-to-day routine. The application includes currency detection, product recognition using barcode and text reader. The application is trained using Convolutional Neural Network (CNN) to recognize the currency and by applying transfer learning over the model VGG 16 and Keras framework (Python). Barcode is detected using Mobile Vision API provided by Google and the information about the product is fetched using other public API's. Similarly, Mobile Vision API is used to read text from an image clicked by a user. The app is gesture-based, and the detected results are converted to speech, thus making it easier for the unsighted person to use.
Keywords: *AI, CNN, VGG, Currency, Vision.*

1. Introduction

With the advancement in technology, we can help visually impaired people in many ways. Lot of research work have been done in the development of assistive technologies for visually impaired people. But there are some limitations that are yet to be solved. Most of proposed solutions till date uses costly hardware to improve the performance which makes the system bulky and costly. Some applications have good accuracy but the interface has buttons, which makes it less useful for unsighted people while some proposed solutions are technically infeasible. Also, these mentioned systems are not based on Indian demographics.

Today, Android phones are available at affordable rates and easily available. Our solution to the above problem is using the camera of Android phone to provide assistance in the form of Indian Currency Detection and Product Recognition. The proposed idea can also be extended in banking sector where currency can be detected in ATMs, e-banking, etc. Here we propose a digital assistant for visually impaired people which is an Android application that detects Indian Currency and Products and also reads text from an image.

The paper is structured as follows. In section II we have presented a literature review of different research work and previously built applications, in section III we describe the Overall System Design, in section IV Implementation of the system is presented, in section V shows the results, section VI covers the future scope, section VII mentions we finally conclude.

ISSN: 2233-7857 IJFGCN

2. Related Work

There are many research works and implementations done till now for providing assistance to visually impaired people in various ways such as currency detection, product recognition, obstacle detection and object detection. The following are some of the recent works done in this field as well as for the unsighted people.

Ch.Ratna Jyothi et.al. in their research work[2],proposed a system which recognizes 4different paper currencies. How- ever, their system is not concerned with verification of the validity of the paper currencies (i.e. verifying that the paper currency is genuine and not faked[2]. The proposed approach involves image acquisition, image preprocessing using a median filter, extraction of features like the color mean intensity of the image (in pixels), color variance, the color skewness of the image area, image edge orientation, the image gray level co-occurrence matrix, and finally classification using Artificial Neural Networks[2]. Initially, their feed forward back propagation neural network is trained with 100 samples of each currency type[2]. Average recognition rate was seen as 93.84%. The processing time and accuracy of the system can be further improved [2].This work is in progress as a subsequent work together with the issue of considering dimensionality reduction with genetic algorithms.[2016ICEEOT] [2].

Noura A. Semary et. al. have presented a simple currency recognition system applied on Egyptian bank note for visually impaired people[1]. The paper discusses both Scanner-based and Camera-base techniques [1]. The basic techniques that they have utilized in the proposed system are image foreground segmentation, histogram enhancement, the region of interest (ROI) extraction and lastly, template matching based onthe cross-correlation between the image captured and the data set[1]. Their experimental results show that the proposed method can recognize currency with 89% accuracy and short time. They carried out the experiments on the Matlab R2012a running on 4 GB RAM and 2.30 GHz processor machine and then transferred to Android platform[1]. They generated the training data by scanning the paper notes. Testing was conducted on images captured using the camera of a mobile phone of different resolutions [1]. The android application was designed for visually impaired users, so there were no icons or buttons and no manual configurations required. An average accuracy of 90% is obtained for the system[2015 ICTA][1][4].

Khin Nyein NyeinHlaing et. al. describe a system to classify 5 classes of currency notes of Myanmar using a k-NN classifier algorithm. The proposed model is based on a textual feature such as Gray Level Co-occurrence Matrix (GLCM). Their recognition system is composed of four different parts. First is the skew correction of the rotated image. The captured image is then preprocessed in the second step and the third parties extracting its features by using GLCM. The last one is recognition, in which the main thing is the k-Nearest Neighbor Classifier. They generated the dataset by scanning and it includes 500 banknotes images of 5 different denominations, each containing 100 distinct images. Maximum performance of 99.2% was achieved with Euclidean distance measure and k being 1 for the k-NN classifier. Their paper also compared its performance with GLCM being used with ANN and FNN, and the proposed method with k-NN works much better than the other two. The processing time is around 0.1 seconds which is good and can be applied for real-time applications.[2016ACDT][5].

Neeru Rathee et. al. in [6] have discussed the solution to the issue of Fake currency detection in India. Image processing algorithms have been adopted to extract the features such as security thread, intaglio printing (RBI logo) and identification mark, which are adopted as security features for Indian currency. To make the system more accurate and robust, they fused the total score of all the three features to differentiate between real and fake currencies. First, they created a database of a number of authentic Indian notes of different denominations. Then they extracted their features, converted them into their binary equivalents and then calculated their mean square error. Keeping

ISSN: 2233-7857 IJFGCN

that value as a reference, they compared the features of the sample notes and observed that the fake notes showed a greater variance from the calculated mean square error. The extracted features were used for fake currency detection. The final decision was taken by fusing the decisive score of the individual features. The fake currency detection accuracy of the proposed system is 100%.This system can be applied for fake currency detection for electronic currency exchange and money deposit using ATM[2016INDIACom].

Shagufta Md. Rafique Bagwan et. al. have proposed an application[7] that recognizes direction of maximum intensity, major colors in any image and objects in an image. For color detection, it uses 'H' component of HSV after converting RGB to HSV, since H describes the actual color. For direction of maximum intensity, as the intensity of brightness goes on increasing, the frequency of sound signal generated will go on increasing. They used a hybrid algorithm, in which the Euclidean Distance measures and Artificial Neural Networks are used together for object recognition. All the results were conveyed using text to speech. The results were obtained with an accuracy of 97.5% and they also compared their results with SIFT object recognition algorithm and their hybrid algorithm works better [2015IC4].

Nitchaya Jethjarurach et. al. have proposed a system to detect barcodes for Thai people and then obtain information about the product using the barcode number from a database. There exists some standard barcode IDs such as UPC, EAN, etc. First they capture the image, preprocess it and obtain a better image of barcode. They have used Hough Line Transform in OpenCV to obtain the barcode from an image and ZXing library to convert barcode image to barcode number. After obtaining the number, it is fed to a database as a search query and the information obtained against that number is presented as output in vocalized format. In case of any failure, voice notifications are used [2014 ICISA][8].

Akshay Salil Arora et. al. in their system[9] have used infrared sensors and ultrasonic range finder circuit to detect hurdles for visually impaired people. They have used Blue-tooth and GPS technology incorporated in an application to provide voice assistance at any location and SMS alerts to registered mobile numbers. For the proposed system, they have used 8951 Microcontroller Unit, an Obstacle Detection Unit, Bluetooth Module link, Voice IC interface unit. Current flows through limiting resistors of a base of two transistors which magnetize the relay coils when the obstacle is detected which then activate the vibration motor and buzzer. The system detects hurdles at a distance of up to 180 cm. They have used Google maps to detect the location and send alert messages to registered numbers[2017 ICNTE].

Hao Jiang et. al. in [10] have used Computer Vision, Optical Character Recognition(OCR), and Text-to-Speech(TTS) to automatically recognize texts and signs in the environment and help the blind users navigate. The system first uses OCR on the image captured to identify the sections of image containing text and signs and then recognizing what that block contains and it returns the raw text. This text is forwarded to an algorithm which uses Context Free Grammar that corrects the raw text and recovers missing details if there is any. Finally, the text is converted to speech using TTS.

Microsoft has a research project "Seeing AI"[11] that recognizes products, currency, reads a document, detect emotions, detects color, objects, and handwritings for visually impaired people. It is quite expensive as compared to other applications built but has much better performance.

3. Proposed System

We have developed a system that assists visually impaired people in two ways: Currency Value detection and Product Recognition. The user should have an android mobile phone with android version greater than 4.0. The image is captured using the android phone camera and then it is processed according

ISSN: 2233-7857 IJFGCN

to the module selected. The system is easy to use and the interface is based on gestures. All the results are announced in speech, thus making it feasible for the unsighted user to use. Figure 1 shows the system workflow of the application.

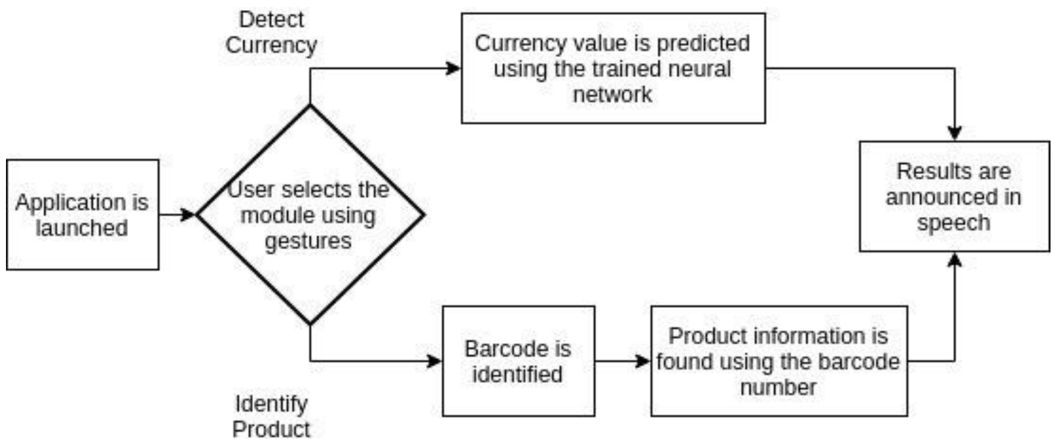

Figure 1. System Workflow

3.1. Gesture Based Interface

One of the major and common problems with the previously existing applications is, although the accuracy of the Model is good; they have a poor user interface. A visually impaired person has difficulty in seeing which makes it obvious that if the application is a button-driven, it would fail miserably in spite of better performance as the visually impaired user won't be able to use it. So, button based or touch based interfaces degrade the performance of such applications. To overcome this, we came up with a gesture based interface which is easy to use and does not require much interaction with the user. A user has to swipe to select the module that has to be used. A right swipe moves the user to next module, whereas a left swipe brings back the user to the previous module. Also,the results are announced in audio format. This is simpler to use than any touch based interface as the target user of the application is visually impaired and can not interact much with the application.

3.2. Currency Value Detection

This module detects the currency value by clicking the image of the note and predicting the value using the trained neural network. We have used Keras framework and applied transfer learning over pre-trained VGG 16 model. The net- work is trained for 8 denominations of Indian Currency(Rs. 10,20,50,100,500,2000) and old 500 and 1000 rupee notes which are concluded as fake notes as they are not in use. The training dataset consists of 850 notes of each denomination and validation dataset contains 150 notes of each denomination. We trained the model on the augmented dataset that produced 94% validation accuracy and 99% training accuracy.The predicted value is announced in speech.

3.3. Barcode Detection

This module recognizes the product by clicking the image of the product and identifying the barcode number and then getting the product information using the barcode number and publicly available API's. We have used Google Vision API to detect the barcode number. It detects the following categories of barcode: EAN-13, EAN-8, UPC-A, UPC-E, Code-39, Code-93, Code-128, ITF, Codabar, QR Code, Data Matrix, PDF-417, AZTEC. The API returns the barcode number obtained from the image. The API is real time and provides results within milliseconds. This returned barcode value is passed to other API which provides information about the product using barcode numbers. The product information is announced in speech.

ISSN: 2233-7857 IJFGCN

3.3. Text Reader

This module reads all the text from an image line by line. We have used Google Vision API to read the text. The module works by clicking an image on a single tap. This image is then processed to find the text written. The results obtained are then announced in speech line by line.

4. Implementation

To predict the currency from an image, we have used Convolutional Neural Networks (CNN). VGG16 is a pre- trained model that has trained weights for images of the Imagenet dataset. We have applied the concept of transfer learning and changed the last layer of the VGG16 model and trained it for our purpose. The training was done using a 16GB Ram. We have removed the last fully connected layer and treated the rest of the CNN as a fixed feature extractor for our dataset. This generates new bottleneck features that include the newly updated weights. There are 8 classes, thus 8 output nodes. The test image is then given to this newly trained network and it then predicts one of the 8 classes of Indian Currency. To generate the dataset of 1000 notes we used augmentation techniques which are rotation, sheer and zoom. To augment the data, we have used Augment or library.

We tried many models like VGG16, VGG19, Xception to see which model produces better accuracy on both, preprocessed data and non preprocessed data. Preprocessed data is the data we generated using Augmentor and non preprocessed data is the original data captured using amobile phone camera and scraping Google images. We found the following training accuracies shown in table I. We concluded thatVGG16 gives the best results on preprocessed data.

Table 1. Table of Training Accuracy Of Various Models

	VGG16	VGG19	Xception
Preprocessed	98.61	96.37	95.02
Non preprocessed	92.40	91.28	89.73

To connect the trained model with android application, we built an API using Django framework. The application uses the camera of the mobile phone to click the image of note to be predicted. This image is sent to the Django server where the model resides, through an API. The image is processed and the predicted value is sent back to the application in JSON format. The obtained result is converted in speech format .For barcode detection and text reader we have used Google Mobile Vision API. The Mobile Vision API provides a framework for finding objects in photos and video.

The Barcode API detects barcodes in real-time, on device, in any orientation. A start Camera Source method is used to start the android device camera. When the camera starts, user can place a product with barcode aligned in front of the camera. User can now tap on the screen so that image is captured. This image is then searched for barcode using the API library and the barcode in the image is identified. This barcode number is sent to Semantics3 UPC API which returns information about the product, given an input which is aUPC of any product.

Text recognition is the process of detecting text in images and video streams and recognizing the text contained there in. Once the text is detected, the recognizer then determines the actual text present in the content detected and segments it into lines and words. The same camera source as that of the barcode

ISSN: 2233-7857 IJFGCN

reader is used. After the image is captured it detects the text in the image and converts the text to speech format using the text To Speech method.

5. Experimental Results

We tried many models like VGG16, VGG19, Xception to see which model produces better accuracy on both, preprocessed data and non preprocessed data. We found the following training accuracies shown in table I. We concluded thatVGG16gives the best results on preprocessed data.

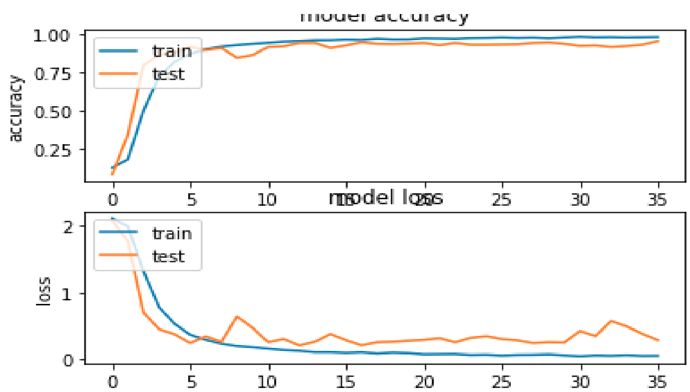

Figure 2. Validation Accuracy and Error Loss of VGG16 model

Figure 3. Currency Value Prediction

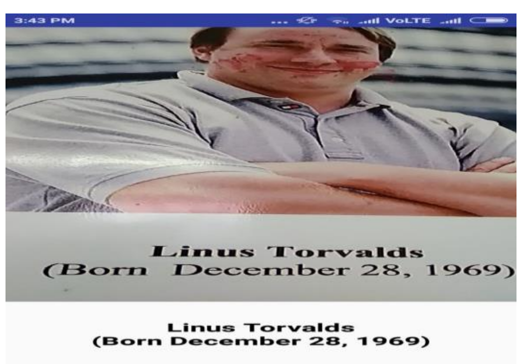

Fig 4.Text Reader

ISSN: 2233-7857 IJFGCN

6. Conclusions

The application gives good results for currency detection with an accuracy of 94%. It detects the bar code number correctly and fetches the information from publicly avail- able databases. The text reader module detects all the text accurately. The application is compatible with mobile phones having android version greater than or equal to 5.0(JellyBean). Neural networks produces better results than conventional machine learning algorithms and the accuracy can be further increased by increasing the dataset size and better preprocessing. We observed that augmented data generates better accuracy results than original data. Internet connection is required to connect to API and it takes about 2-3 seconds toget the results from the server.

7. Future Scope

More modules such as navigation, object recognition can be added to the application making it a great tool that could assist the visually impaired people in so many ways making their life a bit simpler. Also for currency recognition, detecting whether a note is fake or not could be added using security features of the note.

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" , 2015 5th International Conference on Information & Communication Technology andAccessibility (ICTA), 2015.

[2] Ch. RatnaJyothi, Y. K. SundaraKrishna, V. SrinivasaRao. "Paper currency recognition for color images based on Artificial Neural Network" , 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), 2016

[3] http://www.who.int/mediacentre/factsheets/fs282/en/ , October 2017

[4] www.hindustantimes.com/india-news/india-to-change-definition-of-blindness-reduce-number-of-blind-by-4-million/story-

[5] HxHKeH3XpfPBEtSr2moerO.html

[6] Ch.RatnaJyothi, Dr.Y.K.SundaraKrishna, Dr.V.SrinivasaRao,"Paper Currency recognition for color images based on Artificial Neural Network", 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT).

[7] Noura A. Semary, Sondos M. Fadl, Magda S. Essa, Ahmed F.Gad4,"Currency Recognition System for Visually Impaired: Egyp- tian Banknote as a Study Case", 2015 5th International Conference on Information & Communication Technology and Accessibility (ICTA).

[8] KhinNyeinNyeinHlaing, AnilkumarKothalilGopalakrish- nan,"Myanmar Paper Currency Recognition Using GLCM and k-NN", 2016 Second Asian Conference on Defence Technology (ACDT).

[9] NeeruRathee, ArunKadian, RajatSachdeva, VijulDalel, YatinJaie, "Feature fusion for fake Indian currency detection", 2016 3rd Interna- tional Conference on Computing for Sustainable Global Development (INDIACom)

[10] ShaguftaMd.RafiqueBagwan, Prof. L.J.Sankpal, "VisualPal: A Mo- bile App for Object Recognition for the Visually Impaired", IEEE International Conference on Computer, Communication and Control (IC4-2015)

[11] NitchayaJethjarurach, YachaiLimpiyakorn, "Mobile Product Barcode Reader for Thai Blinds", 2014 International Conference on Informa- tion Science & Applications (ICISA)

[12] AkshaySalilArora, VishakhaGaikwad, "Blind Aid Stick: Hurdle Recognition, Simulated Perception, Android Integrated Voice Based Cooperation via GPS Along With Panic Alert System.", 2017 Interna- tional Conference on Nascent Technologies in the Engineering

Field (ICNTE-2017)

[13] Hao Jiang, Thomas Gonnot, Won-Jae Yi and JafarSaniie, ”Computer Vision and Text Recognition for Assisting Visually Impaired People using Android Smartphone”, 2017 IEEE International Conference on Electro Information Technology (EIT)

[14] https://www.microsoft.com/en-us/seeing-ai/

ISSN: XXXX-XXXX IJXX