# Designing Mobile Application for visually impaired and blind persons

**Ranjeet Patil**
**Student, Department of Information Technology**
Ramarao Adik Institute of Technology University ,*DY Patil Deemed to be UniversityMaharashtra, India*

**Rahul Modi**
**Student, Department of Information Technology**
Ramarao Adik Institute of Technology University ,*DY Patil Deemed to be University ,Maharashtra, India*

**Avinash Parandekar**
**Student, Department of Information Technology**
Ramarao Adik Institute of Technology University,*DY Patil Deemed to be University,Maharashtra, India*

**Ms. Jyoti B. Deone**
**Asst. Prof., Department of Information Technology**
Ramarao Adik Institute of Technology University,*DY Patil Deemed to be University , Maharashtra, India*

***Abstract*— Application for Visually Impaired person is an application which will recognize objects in the scene or video and say it's description out loud, it will also read documents in front of the camera and it will also help navigate the person through obstacles, bill payments and travel in day-to-day life. Life of Visually impaired person is hard as they cannot see objects in their surroundings properly so to help them our application will guide them through their life on day-to-day basis. Our application will tell them what they are looking at by saying it out loud. With Advancement of technology in fields of AI and ML a lot of time and personnel is being used to create a solution for Visually impaired person, a lot of applications exists in market which will convert a video or a scene into a speech or would navigate a person through daily obstacles, but these applications are not solving more than one problem. Our application also converts a given text into regional or person specific language. More eLearning and edutainment components can be added to the mobile application. People with visual impairments (VI), such as those who have poor vision or are entirely blind, can use the app to learn about their surroundings.**

***Keywords— VI, AI, Speech-to-text, Object Recognition.***

## I. INTRODUCTION



Visual impairment is a serious handicap for visually impaired people. According to the World Health Organization (WHO), there are approximately 270.5 million individuals worldwide who are visually impaired, with 49.1 million being blind and 221.1 million having low vision [1]. These individuals frequently find it difficult to carry out their day-to-day activities. Like reading and writing, sending messages, navigating on the roads etc [2]. The currently existing system uses IOT devices which are often expensive and majorly focused on only navigating. Systems which are software based are works on only one domain like e-book reading text-speech. The proposed system uses Artificial Intelligence to assist the visually impaired people which is all based on voice command [3,4]. It captures the surrounding with use of mobile camera and with the help of object detection algorithm it recognizes the object and describes them in an audio. It uses voice synthesis to read e-books for the visually impaired and natural language and Text-to-Speech to convert documents/soft copies of books to speech. [5].

## II. LITERATURE SURVEY

Smartphones have become an unavoidable part of modern life. It has a number of features, including calling, messaging, and the ability to interact with others via social media. The number of smartphone users worldwide has risen to 7.26 billion, implying that 91.62 percent of the global population now possesses one [6]. This is to be expected, given that smartphones and tablets include mobile apps that provide several benefits and assist us in a variety of ways. We can, for example, use smartphone apps to make payments or even place food orders. We can also use a smartphone app to check our electricity bills and pay them directly. However, it is only useful for normal individuals, and the majority of its applications are restricted towards the visually impaired. Visually impaired people confront numerous challenges that are taken for granted by sighted people. We recognise items in our daily lives as sighted users. Without eyesight, finding an object in an unknown area is nearly impossible. In any case, visually impaired persons rely on their memories to locate their items, and they become irritated if the object is misplaced or taken. Visually impaired or blind persons are unable to visualise the item in front of them, making daily living difficult for them. They won't be able to see what the thing is even if they touch it because some objects are unfamiliar to them, and even if they can recognise the object after touching it, it will take time and the result will not always be right. People who are blind or visually challenged confront challenges that we are not aware of because we have eyes that can see. The majority of sighted people believe that blind persons can only see darkness. They believe that someone who is blind has the same experience as someone who is blind and walks around with their eyes closed. Nonetheless, it isn't. Somebody can be visually impaired and keep a specific degree of vision. For outwardly hindered individuals, their eyes should work at 20/200 vision or more regrettable This implies that their eyes are 10% as solid as those of an individual with practically no visual hindrance [7]. Outwardly debilitated individuals face troubles exploring outside the spaces they are normal to. Essentially, actual

development is probably the greatest test for blind individuals. Voyaging or in any event, strolling down a jam-packed road can be trying for blind individuals. In addition, they should recollect the area of pretty much any snag or item in their home climate, for example, tables, seats and bed which ought not be moved suddenly and without notification ahead of time to keep away from mishaps [8]. Many outwardly disabled individuals simply need everybody to see them as ordinary individuals, not as individuals with inability [9]. They could feel as offended by overcompensating thoughtfulness. For example, when they request directions, assist them with going across the road and take them to their objective or counting cash for their benefit when they are purchasing something [10]. In any event, getting something, they dropped and proposing to convey their things is totally seen as overcompensating graciousness, which many visually impaired individuals believe is discourteous. Thus, we need to ask first assuming that they need assistance, and recognize it assuming they say no [11]. Other than that, outwardly weakened has limit profession choices as visual impairment influences their capacity to perform many work capacities and assignments. This definitely influences their abundance and confidence [12,13]. Visual impairment likewise may create some issues assuming they are participating in assignments outside like games. This in the long run will influence their capacity of to associate and meet new individuals, which will prompt enthusiastic and emotional well-being issues.

## III. METHODOLOGY

The Figure 1 and Figure 2 are of a Pie Chart and a Line Graph respectively, which show how many people use Android on day-to-day bases [14]. Among these people are also the Visually Impaired. Android applications are free to develop. Developers are not supposed to pay any testing or membership fees. The Second Picture of the Line Graph shows the growth in the users preferring Android as Operating System rather than other Operating Systems [15,16].

According to both diagrams Android is one of the most used OS around the world according to the pie chart 44.52% of the user use android and according to the line graph since 2009 Android has seen an exponential growth where as other Operating Systems have seen decline in their market shares [17].

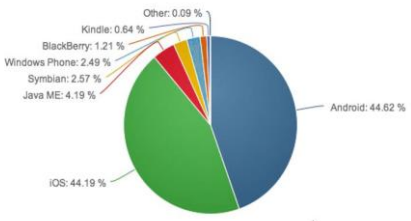



Fig. 1: Pie Chart showing usage different OS.

**The proposed system includes 4 steps:**

**A. Mobile Application:** The mobile application is the front-end of the project which will be used to access all the functions provided to the user. When the mobile application is on the user just has to give voice command to capture the screen and the mobile application will return thee voice output.

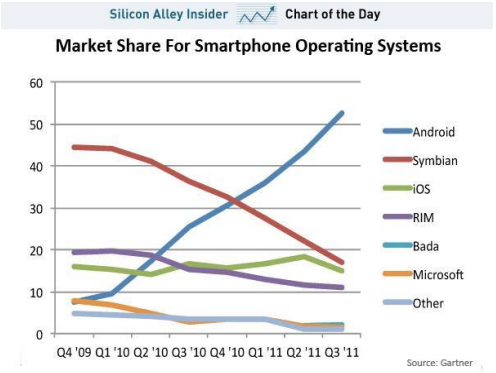



Fig. 2: Market Share for Smartphone OS

**B. Image-to-Speech:** Text-to-speech is one of the most important features of the Application in which the screen captured by the mobile camera will be converted into a speech will be said out loud by the voice assistant. The Language of the speech is determined by the language chosen by the user it can convert text to language that user wants to hear it in. e.g., English text can be converted into Hindi or any other language. Figure 3 shows the flowchart for text recognition in phases of giving input image, Text Detection, Detection Post-Proce, Text Recognition, Voice Assistant.

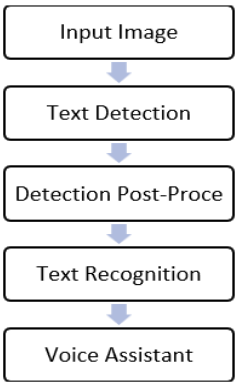



Fig. 3: Flowchart for Text recognition.

**C. Image description:** In this part the image captured by the mobile camera will have an object which will first get recognized by the system and then the available description is said out loud by the voice assistant. This is an important feature as it will help the user to navigate through roads, paying bills, walking in park or knowing about the place they are visiting.

**D. Knowledge Base**: To gather semantic information about identified labels, an ontological knowledge base was used (i.e., terms). WordNet, or Indo WordNet, is employed

as a basic knowledge base in the proposed mobile application [18], together with other developed ontologies. The programme may be set up to provide possible explanations for specified phrases. For example, the definition of chair is "a separate seat for one person, often having a back and four legs" [19,20]. This informative explanation will assist the user in comprehending his or her surroundings.

## IV. EXPERIMENTAL RESULTS

In this experiment we have given 80-100 input images for all the currencies and for object detection we used TensorFlow and dataset from Kaggle with 1000+ different objects. This data set will help model understand and learn how to recognize different objects when deployed. When the model is deployed it will detect the object in real time. The same application has Currency recognition and image text-to-speech, in currency recognition we have made a model using TensorFlow the model has input of different bills which help in developing a model that will detect the currency and its value and then return a voice command to the user to notify its value.

```
def get_string(img_path):
    img = cv2.imread(img_path)
    img = cv2.cvtColor(img,
cv2.COLOR_BGR2GRAY)
    # Apply dilation and erosion to remove
some noise
    kernel = np.ones((1, 1), np.uint8)
    img = cv2.dilate(img, kernel, iterations=1)
    img = cv2.erode(img, kernel, iterations=1)
    cv2.imwrite(src_path + "removed_noise.png", img)
    img = cv2.adaptiveThreshold(img, 255,
cv2.ADAPTIVE_THRESH_GAUSSIAN_C,
cv2.THRESH_BINARY,31, 2)
   cv2.imwrite(src_path + "thres.png", img)
   result =
pytesseract.image_to_string(Image.open(src_
path + "thres.png"), lang='eng')
    return result
```

Fig. 4: Function for Image Text-to-Speech

In Figure 4 code is given for the function that performs the conversion of image having text to speech. This will help user in reading books, bills and newspaper in day to day they just have to capture an image of the content and the application will return a voice output.

We have created a function called get_string with a input parameter img_path, the device will give img_path as input which will then be used to get the image and convert it into a voice.

### A. Data Sets:

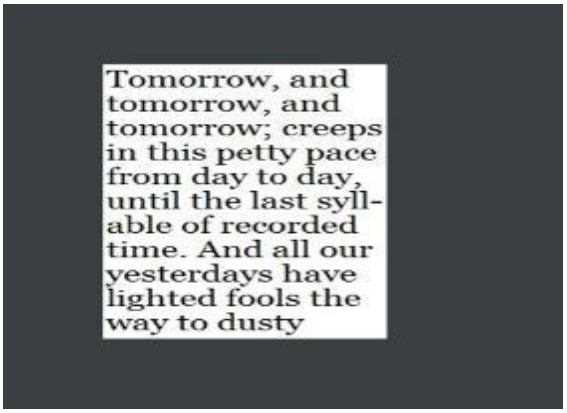
To our Knowledge data sets for object and currency recognition are not easy find , but for text recognition which is then converted to speech is easily available on sites like Kaggle .For currency recognition model we had picked 80 – 100 different images of each currency these pictures where then given to the model to go through and learn from them , these images help the model / machine to develop a understanding that will help in generating better results when the model is used for experimental or practical purposes . These images are different than each other such that all these images are capturing same model but at different angel as this application is being made for people who cannot see perfectly this model/machine has to recognize right currency with a given amount of data which can vary a lot.

Data that is fed to the text to speech converter are images of text from book written and printed which are used to make the model familiar with how the user is going to give the input.

Another data set that is provided to the model is voice for the currencies with images voices are also given to the application which are then used as the output when the currency is recognised.

Figure 5 and Figure 6 are subsets of dataset provided by us to make the model.


Fig. 5: shows some of our data subsets for currency recognition.




[1]


[2]

Fig. 6: Shows some of data subsets for speech to text conversation.

**B. Experimental Setup:**

As a measure of recognition performance, we utilise the widely used overall accuracy (OA), which is defined as:

OA = Number of correctly recognized objects / Number of real objects.

To obtain a tenable acknowledgement result, we randomly divide each informative index into a prepared subset and a test subset, and repeat this activity several times. The OA is processed for each split, and the mean and standard deviation of the 10 OAs are calculated as the final acknowledgement result. We set the proportions of the prepared subset to 20% and half for each informative index, while the proportions of the test subset to 80% and half.

For Object recognition we have used YOLOv3 algorithm this is one of the fastest and accurate algorithms for object detection, using this algorithm we can give a real time output to the user whenever they need to navigate through any Road, Obstacles and any other day to day activity [19].

**C. Results:**

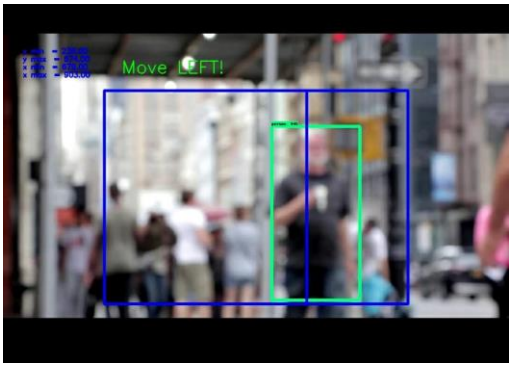



Fig. 7: Shows result of Object Detection.

Figure 7 is the result of the object detection algorithm which is detecting a human (person in picture) telling the user to move left. This image is an example of how we expect our application to work.

Figure 8 shows the result of the currency detection with this voice command will be given to the user.

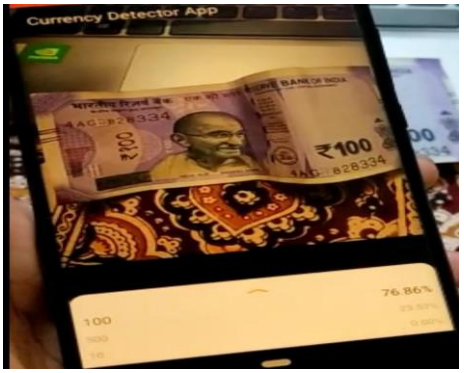



Fig. 8: Shows result of Note Detection example.

## CONCLUSION

In this paper, we are purposing an application in which the user with visual impaired Ness to live their life with less difficulties. From this application the person using it can read text written on a book, newspaper or any other paper the function we have built will help us in giving a output to the users input as soon as possible. Similarly, with respect to currency recognition our model will give a response to the user about what kind of currency and value of the note they are holding this will help them in paying bills and in money transactions. The third part will help the user in navigating through different parts of roads and in finding things in their surroundings, the third part is object detection which gives voice output about what user is seeing. Visual Impaired person have lots of challenges in their day-to-day life, this application can help them with few of these challenges. With advancement in Bioengineering, Artificial Intelligence and IOT we can expect in future to have gadgets that will not only be fast but will be more accurate. This application as per our knowledge gives solution for three different problems faced by visually impaired person that no other application does there are application that solve a single problem but not all three.

## REFERENCES

[1]. Akhilesh A. Panchal, Shrugal Varde, M.S. Panse," Character

Detection and Recognition System for Visually Impaired People", IEEE International Conference on Recent Trends in Electronics Information Communication Technology, May 20-21, 2016, India.

[2]. Shubham Melvin Felix, Sumer Kumar, and A. Veeramuthu "A Smart Personal AI Assistant for Visually Impaired People" IEEE International Conference on Recent Trends in Electronics Information Communication Technology, May 20-21, 2019, India

[3]. Dr. K. Srinivasan1 , V.R. Azhaguramyaa2 "Internet of Things (IoT) based Object Recognition Technologies " IEEE International Conference on Recent Trends in Electronics Information Communication Technology,, 2018, India

[4]. A. G. Karkar, J. M. AlJaam, and A. Mahmood, "Illustrate it! an arabic multimedia text-to-picture m-learning system," IEEE Access, vol. 5, pp. 12 777–12 787, 2017.

[5]. S. Shadi, S. Hadi, M. A. Nazari, and W. Hardt, "Outdoor navigation for visually impaired based on deep learning," CEUR Workshop Proc., vol. 2514, pp. 397–406, 2019.

[6]. S. Kavya, M. Khanum, K. Madhu, P. Gowda, and A. S. Manek, "Recognizant : An Android Application for VisuallyImpaired People," vol. 4, no. 2, 2019.

[7]. Lee, S. Jung, and H. Song, "CNN-Based drug recognition and braille embosser system for the blind," J. Comput. Sci. Eng., vol. 12, no. 4, pp. 149–156, 2018.

[8]. N. Mohapatra, "13 Things You May Not Know About Blind People Are You Reaching Your Full Potential ?," pp. 1–11, 2019.

[9]. Shahed Anzarus Sabab, Md. Hamjajul Ashmafee, "Blind Reader: An Intelligent Assistant for Blind", 19th International Conference on Computer and Information Technology, December 18-20, 2016, North South University, Dhaka, Bangladesh.

[10]. O. Penatti, K. Nogueira, and J. Santos, "Do deep features generalize from everyday objects to remote sensing and aerial scenes domains," in Proc. CVPR, 2015, pp. 44–51.

[11]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Proc. NIPS, 2012, pp. 1106– 1114.

[12]. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in Proc. ICLR, 2015, pp. 1–13

[13]. M. Cimpoi, S. Maji, and A. Vedaldi, "Deep filter banks for texture recognition and segmentation," in Proc. CVPR, 2015, pp. 3828-383.

[14]. F. Perronnin, J. Sanchez, and T. Mensink, "Improving the fisher kernel for large-scale image classification," in Proc. EEC, 2010, pp. 143-156.

[15]. K. Rainey, S. Parameswaran, J. Harguess, and J. Stastny, "Vessel classification in overhead satellite imagery using learned dictionaries," in Proc. SPIE, 2012, pp. 1-12.

[16]. U. Kanjir, H. Greidanus, and K. Ostir, "Vessel detection and classification from spaceborne optical images: A literature survey," Remote Sens. Environ., vol. 207, pp. 1-26, 2018

[17]. Z. Liu, W. Dai, M. Z. Win, "Mercury: An infrastructure-free system for network localization and navigation", IEEE Trans. Mobile Comput., vol. 17, no. 5, pp. 1119-1133, May 2018.

[18]. Vision impairment and blindness: Fact sheet", World Health Org., 2017,

[19]. Bitao Jiang, Xiaobin Li, Lu Yin , Wenzhen Yun , Shengjin Wang-" Object Recognition in Remote Sensing ImagesUsing Combined Deep Features" 978-1-5386-6243-4/©2019 IEEE

[20]. Rucha Doiphode, Mayuri Ganore, Ashwini Garud, Tejaswini Ghuge and Parminder Kaur (Guide)" Be My Eyes: Android Voice Application for Visually Impaired People "unpublished.