



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

SMART AID FOR VISUALLY IMPAIRED PEOPLE

T. Dhiliphan Rajkumar Department of Computer Science and Engineering

*Kalasalingam Academy of Research and Education
Krishnankoil, Virudhunagar, India*

K. Sanjeev Karthik Department of Computer Science and Engineering

*Kalasalingam Academy of Research and Education
Krishnankoil, Virudhunagar, India*

D. Anand Prakash Department of Computer Science and Engineering

*Kalasalingam Academy of Research and Education
Krishnankoil, Virudhunagar, India*

J. Saravana Kumar Department of Computer Science and Engineering

*Kalasalingam Academy of Research and Education
Krishnankoil, Virudhunagar, India*

Abstract - Speech and text is that the primary vehicle for human correspondences. individuals needs vision to get to the information in an extremely text. Anyway individuals who have low vision can accumulate data from voice. it's controls the Cameras and Bluetooth headsets which go about as an associate between the framework & consequently the client. Optical Character Recognitions or OCR is executed during this venture to allow characters which are perused out by the framework via a Bluetooth. The camera is mounted on exhibitions, it catches a full perspective on the paper into the framework. Additionally, when the camera takes the depiction of the paper. It's guaranteed that there are acceptable lighting conditions. The substance on the paper should be scripted in English and be of fine text dimension. When of this condition are meet the framework snaps the picture, measures it & in the event that it perceives the substance composed on the paper it'll declare on the Bluetooth headset speaker that the substance on the paper has been effectively processed. And next it stands up the substance that was changed over in to message design inside the framework from handling the picture of the paper. this can be utilize full for dazzle people groups to peruse the prescriptions and book successfully.

1. INTRODUCTION

There are 285 million individuals are indicated by the World Health Organization (WHO) are assessed to be outwardly hindered worldwide among which 90% live in agricultural nations & 45 million visually impaired people overall. Despite the fact that there are many existing answers for the issue of helping people who are oblivious in regards to peruse, anyway none of them give a perusing encounter that its any dimensions matches that of the discovered population. A visually impaired can peruse record simply by tapping words which is then perceptibly introduced through content to discourse motor. "Daze Reader" – created for contact gadgets which is easy to use and powerful intuitive framework for visionless or low vision individuals.

OCR is finished from Text extraction from an image. OCR helps with making scrutinizing devices for ostensibly handicapped individuals and advances including media transmission. In this structure the change of text to voice yield is by e-Speak count. It is a TextTo-Speech (TTS) system which changes over substance in the direction of talk. The fake making of human talk is known as talk mix. The talk synthesizer can be executed in programming or a hardware thing. The stage used for this aim is known as a talk synthesizer.

2. LITERATURE SURVEY

2.1. A Visually Impaired Stereo Image Processing System

The above research offers a system that supports blind navigation by combining stereo vision, image processing, and a sonification mechanism. A wearable computer, stereo cameras as vision sensors, and stereo earbuds are all part of the designed

system, which is moulded into a helmet.

Disadvantages

- The gear required is large and bulky;
- Musical stereo sound is essential for the blind to grasp the scene in front of them (No use of voice commands).

2.2 Blind Path Obstacle Detector using Smartphone Camera and Line Laser Emitter.

This study makes extensive use of two items: a mobile camera and a laser. The laser and the mobile are kept at a safe distance from each other. The image from the camera is collected, and the laser is observed as well. The distance is calculated using the static distance and the angle between the laser point and the camera.

Disadvantage

- The distance between the camera and the laser must be constant;
- On a shiny surface, the laser strength may drop, making it ineffective.

2.3 Shortest Path Based Trained Indoor Smart Jacket Navigation System for Visually Impaired Person.

The seeking, way-finding, course following, and impediment avoiding modules are the essential segments of a route framework to aid visually impaired individuals walk to their destination safely in an interior climate, although it remains a moving errand to consider obstruction avoiding throughout course following, as the interior climate is perplexing, changing, and maybe with hazards.. To solve this problem, we present a novel strategy that employs a revolutionary objective-choosing methodology to guide customers to the goal while also assisting them in overcoming deterrents. The proposed route gadget has been tested on a variety of people and has proven to be successful on interior route duties. A live video examination has been activated to detect human countenances and an ultrasonic sensor has been activated to detect human countenances.

3. OBJECTIVE

The main objective of the project is to produce a low cost and efficient Smart Aid for Visually Impaired People which makes them to read or analyze any documents by converting the text to voice format.

4. PROJECT PLAN

This project plans to develop an entire portable aid for blind pedestrians and cater to problems in existing systems efficiently. The system designed will make out an object or obstacle using ultrasonic sensors and provides audio instructions for suggestions. It also gives information about people looking or waving to them using face recognition. It includes authorizing a different login option in an exceedingly paired device in order that quite one blind person can use the identical device. To update the suggestions into database on their own without doctors help by employing a printed suggestion.

5. SYSTEM DESIGN

5.1 Existing System

PC Vision & its administrations are the center of industry modification which is accepted as industry 4.0. For

robotizing the procedure, messages installed in pictures that are viewed as pretty much practically as great wellspring of information about this item. Perusing text from regular pictures keeps on being the difficult issue on account of the confounded foundation, size and space varieties, sporadic courses of action of writings. Location and Recognition are the chief phases of perusing writings inside nature. In past number of years, numerous scientists have given numerous strategy to perceiving writings in pictures. This strategies have fine outcomes on even messages just however not on sporadic plans of writings. This paper are principally particular on profound studying model for narrative acknowledgment in pictures (DL-TRI). The model tends to different instances of bended and viewpoint textual styles and difficult to recognize because of complex foundation.

Performance:

This model are often employed in different industries for mechanization purpose. As this method works alright on non-identical use cases like document reading, car number plate identification, consumer products identification, natural scene text reading etc., many procedures is mechanized & linked to each other. For text detection, this model can play a very dominant role for implementing end-to-end identification model, because it has good recognition accuracy. this technique relies on CNN and LSTM architecture. It outperforms than B. shiet.al , C.Y. Lee et. al. ,Jaderberg et al., 2016, M. Jaderberg et al. models for specific datasets. However, some challenging cases are 2- line texts, symbols, texts in many languages than English. Also, the model has not addressed words including symbols that take other character (most featured) at the place of symbol. The model works particularly for West Germanic language word only.

Disadvantages of Existing System

- CNN has given the Many frameworks .
- Processor should be higher configuration.
- CNN has problem of incompleteness and its mostly computationally costly because it has to take a large database for training.

5.2 Proposed System

Discourse and text are the foremost vehicle for human correspondence. A personal are requirements vision to induce to the knowledge during a book. Anyway those who have helpless vision which is able to assemble data from voice.

This venture has been worked around PC. it's controlling the peripherals like Camera & Bluetooth headset which approach as an combine between the framework and during this manner the client.

Optical Character Recognition or OCR is executed during that undertaking to acknowledge characters that were perused out by the framework through a Bluetooth.

Additionally, while the camera takes the preview of the paper, it's guaranteed that there are acceptable torching conditions.

The substance on the paper should be communicated in English and be of fine text dimension . When of those conditions are meet the framework snaps the image, measures it & on the off chance that it perceives the substance composed on the paper .it'll declare on the Bluetooth headset speaker that the substance on the paper has been effectively treated.

After this it stands up the substance that was changed over in to message design inside the framework from handling the

image of the paper.

Working:

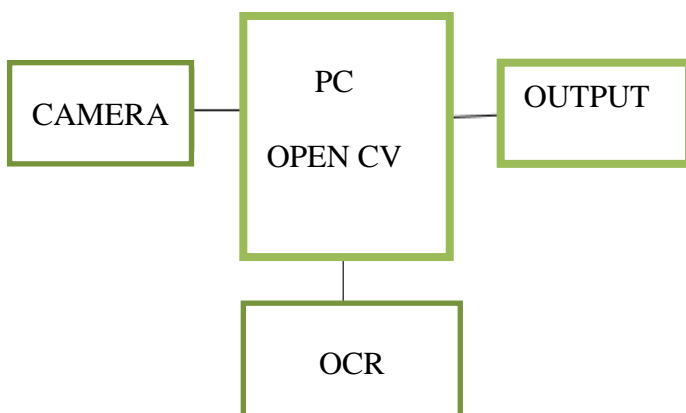
The proposed technique is to assist outwardly hindered individual in perusing the content available on the content marks, print up from notes & items as a camera based upon assistive content per user. An actualized thought are requires text acknowledgment from picture gripped by camera on display & perceives that content utilizing OCR. Change like the perceived record to voice yield by OCR calculation. The framework is pleasant for transportability. the flexibility are permits the client to carry the gadget anyplace and should use whenever. A model was created which utilizes the camera in display & Raspberry pi that projects during this genuine time. This framework catches that sting & examines the existence of text inside the casing. On the off chance that a personality is found by the camera the client are going to be educated that picture with some content was distinguished. Subsequently if the client needed to tune up to or to induce a handle on about the substance inside the image. The caught picture is first changed over to dark scale then sifted utilizing a Gaussian channel to downsize the clamor inside the images. Here versatile Gaussian pointing (or thresholding) is used to cut back the commotion inside the picture. A sifted picture are changed over to double. The binarised picture is trimmed particular the bits of the picture with none characters are eliminated. An edited edge is stacked to the Tesseract OCR to accomplished text acknowledgment. The yield of the Tesseract OCR are PC record which are the contribution of the e-Speak. It makes a easy sign just like the report given from the knowledge. the straightforward sign are created by the e-Speak is then given to an earphone to encourage the sound yield.

Advantages Of Proposed System:

- Higher Productivity
- Cost Reduction
- High Accuracy
- Increased Storage Space
- Superior Data Security
- 100% Text-searchable Documents
- Massively Improves Customer Service
- Makes Documents Editable.

6.PROJECT DELIVERABLES

6.1 System Design



6.2 Camera Interface on Python

Capture a recorded video, you would like into form the VideoCapture device. Its conflicts are often further the gadget file or the name of the recorded video.

Gadget file is barely the amount into work out which video camera. Regularly 1 camera are connected (as of my condition). Therefore, I basically pass 0 (or -1).

You'll select the 2nd camera by passing 1, etc. From that time forward, you'll be able to capture outline by-outline. Yet, regarding to the top, recollectto distribute the capture.

```

import numpy as np
import cv2

cap = cv2.VideoCapture(0)

while(True):
    # Capture outline by-outline
    ret, outline = cap.read()

    # Our procedure on the sting come here
    dark = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)

    # Display the next edge
    cv2.imshow('frame',gray)

    on the off chance that cv2.waitKey(1) and 0xFF == ord('q'):
        break

    # When everything done, discharge the
    capturecap.release()cv2.destroyAllWindows()
  
```

6.3 Transformation of Picture to Text Using OCR Tool

Tesseract is an open source-OCR engine. Its expects the data may that combined picture as well as voluntary polygonal substance zone described. The underlying advance may be a related part assessment wherein system of the portions is taken care of. By the assessment of the settling of styles, it's definitely not hard to acknowledge in reverse message & recollect its exactly upon schedule at dull upon white substance. At this point, outlines occur amassed, basically beside settling to masses. Masses is composed to text lines & therefore the lines & areas are dismembered for fixed pitch or comparing content. Grade across the road is employed to seek out text lines. This lines are shattered into words unmistakably in keeping with such a personality scattering. Fixed pitch text is severed rapidly by character cells. The cells are examined for connected letters & if it's found, by then it's divided. Nature of saw satisfaction is affirmed. If clarity is not adequate that substance is passed into associator. The organizer differentiates every apparent message & planning data. There word affirmation is completed by pondering assurance & rating.

6.4 Tranformation of Text to Voice Using E-SPEAK And PYTTSX3

Regular substance to talk change is done using eSpeak PYTTSX3 which might be the TTS system. The forgery making of human talk are perceived as talk mixture. Talk PC or talk manufacturer is employed therefore & might occur completed in programming or gear things. The limit of entire words or sentences for unequivocal customer territories, thinks about incredible yield. to make a "designed" voice yield a manufacturer might be use to append the model of the voiced

plot & other human vocal credits.

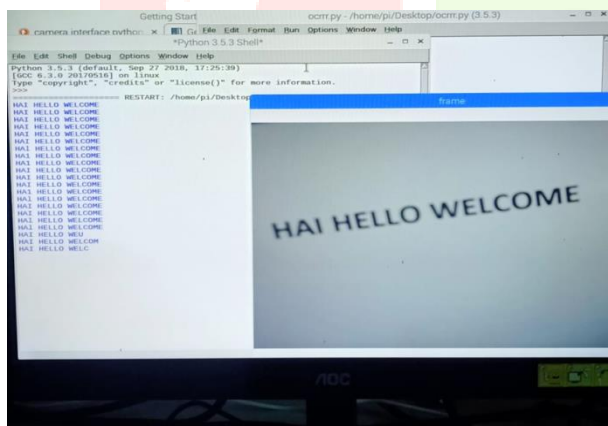
A TTS (or "engine") is framed out into 2 areas a front-end & a back-end. The front-end has 2 huge tasks, the standardization & spoken record of text. Standardization, pre-dealing with, or tokenization of text is that the difference in content containing pictures like shortenings and numbers into equivalent figured out words. The front-end around then designates phonetic record to each word. The metrics units such as stipulations, sentences, & articulations are stepped & apportioned. Text-to-phonemenol change are the route toward dispensing phonetic records to words. The yield from the front-end is a significant semantic depiction from the Phonetic records & hence the speech rthym information. The back-end plays out the limit of a manufacturer. There are meaningful phonetic depiction into sound change are the refined using that side. The principal engaging segment of the talk mix structure are ease & cognizance. The yield looks as though human talk whatever depicts the ease, the yield is ease with the clearness of appreciation. Talk association systems regularly attempt to enlarge both normal & understandability whatever are the characteristics of an ideal talk manufacturer.

6.5 Image to Text

We will probably change over a given book picture into a line of text, saving it to a record and to tune in to what exactly's composed inside the picture through sound.

For this, we'd prefer to import a few Libraries

Pytesseract(Python-tesseract) : it's an visual character acknowledgment (OCR) instrument in-order-to python supported by google.pyttsx3 : it's a disconnected cross-stage Text-to-Speech library Python Imaging Library (PIL) : It adds picture preparing abilities to your Python interprete.



6.6 Text to Speech

pyttsx3 might be a book to- discussion change information center in Python. In difference to elective libraries, it works disengaged & its applicable with both Python 2 & 3. An administration summons the pyttsx3.init() industrial facility capacity into asked the connection to the pyttsx3. Motor occurrence.Its the truly simple to employ device which tranposes over the launched text to conversation.

The pyttsx3 module upholds 2 voices initially are female & hence the second are male which was given by "sapi5" for windows.

It validates 3 TTS motors :

sapi5 – SAPI5 on Windows

nsss – NSSpeechSynthesizer on Mac OS X

espeak – eSpeak on all other platforms

```
ocrr final.py - D:/FINAL AI PROJECTS/ocr text/ocrr final.py (3.6.5)
File Edit Format Run Options Window Help
import cv2
from PIL import Image
import pytesseract
import pyttsx3

pytesseract.pytesseract_tesseract_cmd = 'C:/Users/Yeara/AppData/Local/Tesseract-OCR/tesseract.exe'
en = pyttsx3.init()
en.setProperty('rate', 80)
cap = cv2.VideoCapture(0)

while True:
    #print("ok")
    ret, frame = cap.read()
    #frame = cv2.imread("1.png")
    cv2.imshow('frame', frame)
    cv2.imwrite("1.jpg", frame)
    img = image.open("1.jpg")
    text = pytesseract.image_to_string(img, config="")
    print(text)

    en.say(text)
    en.runAndWait()

    if cv2.waitKey(1) & 0xFF == ord('q'):
        break

cap.release()
cv2.destroyAllWindows()
```

7. FUNDAMENTAL DESIGN KNOWLEDGE

7.1 OCR

Optical characters acknowledgment, or OCR, might be a strategy for changing over an examined pictures into messages. At the point when a page is filtered, it's normally put away as a touch planned move into the TIF designs. At the point when the photos is shown inside the screens, we will understand it. ... The PCs doesn't perceive any "words" inside the image. OCR or Optical Characters Recognition might be a product applications included with certain HP scanners. Generally, archives examined into a PCs are saved as PDFs and it will be perused uniquely with a PC. OCR permits a client during to examine records and save them to a PC, yet additionally to have the option to alter the archives.

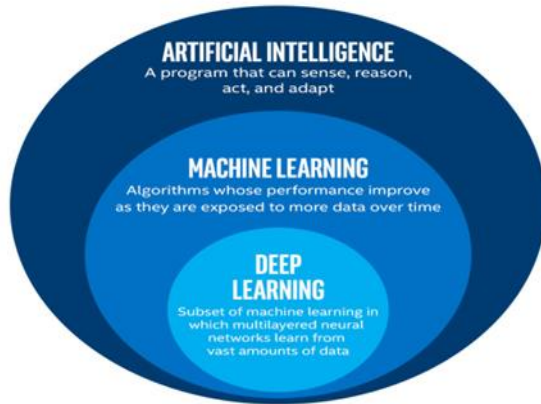


7.2 Artificial Intelligence

In programming, (AI), to a great extent called machine understanding, is information appeared by machines, rather than the trademark information appeared by individuals. Driving AI composed correspondence portray the area on the grounds that the examination of "clever trained professionals".

Man-made consciousness was set up as an educational request in 1955, and inside the years since has experienced some deluges of positive intuition followed by dissatisfaction and furthermore the insufficiency of financing (known as an "man-made knowledge winter"), followed by new procedures, accomplishment and revived sponsoring. These sub-fields rely on particular thoughts, for example, explicit goals (for instance "mechanical innovation" or "computer based intelligence"). cognizance decline as a result of absence of rest or fae a n falls

asl



7.3 Deep Learning

Deep learning are the essential for the most comprehensive group of AI techniques depends upon substitute neural organizations with portrayal Studying. Learning can be managed, semi-administered or autonomous.

Deep learning models, for instance, Deep neuronal organizations, Deep persuasion organizations, repetitive neuronal organizations & complexity neuronal organizations has been appealed to meadows as well as PC vision, discourse acknowledgment, characteristic language controlling, sound recognition, non-formal community separating, machine explication, biocomputing, drug scheme, clinical visualization exploration, information review & table entertainment projects.

Every most correctly, heartfelt studying substructures have the considerable credit task way (CAP) profundity. It is the series of interchanges from contribution into yield. However most analysts concur that profound learning includes CAP profundity higher than 2. Profound models (CAP > 2) can separate preferred highlights over shallow models and subsequently, additional layers help in learning the highlights adequately.

It also calculations are the adapted into solo studying assignments. These are important advantage on that grounds this untagged intelligence are most plentiful than named intelligence. It also modifies consequently, without pre-determined intelligence unambiguous encoded by software engineers.



To acquire the control upon the probability of deep learning, envisage the family, accompanied by the baby & caretakers. The baby clears objects are accompanied by his little finger & consistently says the word 'feline.' If it is people are concerned about his schooling, they resume authorizing him know 'Indeed, that is a feline' or 'No, that isn't a feline.' The small child, where it counts, doesn't have the foggiest idea why he can say it is a feline or not.

7.4 Open CV

OpenCV-Python is that the library of a Python during ties intended to unwind PC dreams issues. ... OpenCV-Python utilizes a Numpy, which could be an exceptionally modern library for the numerical tasks with the MATLAB-style grammar. All of this exhibit formation are interchanged over

to with against Number clusters.

OpenCV (Open Source Computer Vision) is that the library of a programming capacities mostly designed for the continuous PC vision. In basic dialects it's library utilized for the picture Processing. it's mostly acclimated do all the activity related with the photos.



Human eyes are provides the plenty of the informations based to the what they see. Machines is smoothened to the seeing everything, transfigure the innovations to the numbers & stores at the evocation. Here will the question arises how they computer convert the photographs into numbers. therefore the answer is that may the pixel values is employed to the convert images into numbers. A pixel is that the smaller unit of the digital images or graphics that be displayed and represented during a alphanumeric display device.

OpenCV is accessible free of cost.

Since theyOpenCV library was the written by C/C++, so it's quit fast. Now they'll be utilized by Python.

It required less RAM to the usage, it used maybe of 60-70 MB. Computer Vision is portables are OpenCV and may run on the any device that may run on C.

7.5 PYTESSERACT

Pytesseract permits us we to arrange the apply Tesseract OCR motors beside the setting into banners whatever they interchanges the route inside with whatever the picture are explored for the characters. There are 3 primary banners acclimated the designing a Tesseract OCR as language (- l), OCR Engine Mode (- - oem) and Page Segmentation Mode (- - psm).

Alongside during the backslide English dialects, Tesseract upholds numerous different dialects includings Hindi, Turkish, French and so on We were just be utilizing English here, yet you'll have the option to download the prepared information out of proper github page & join it up to a your bundle to concede different dialects.

Its likewise we conceivable into the perceive at least two distinctive language from the indistinguishable pictures can we This language are prepared by banner - l, to line it into a language utilize that codes in addition to the banner, instance for English it'll be - l eng, where eng is that the code for English.

7.6 PYTTTSX3

Pytttsx for python3 [Offline text to discourse for python3]

Pytttsx might be a decent book to discourse the transformations library in python yet it's composed distinctly inside the python2 untillnow ! Indeed, even some decent lot of googling didn't assist much with asking a tts library viable with Python3.

There was nonetheless, one library gTTS which works impeccably is python3 yet it needs in light of the fact that the web associations with the work since it depends in google to prompt the sound data. But Pytttsx is absolutely to offlines and works seamlessly and has various tts-motors support. The codes during this repos are somewhat altered form of the pytttsx modules of python 2.x and might be a clone from westonpace's repo. the point of making this repo is to help individuals who need to claim a disconnected tts lib for Python3 and don't have any desire to port it from python2 to python3 themselves.

This undertaking has been dead for over certain years now and pip introduce pytttsx doesn't download the fixed python3 form of the library. So I settled on a choice to go ahead with my own personal archive and spread the fixed version.

8. CONCLUSION & FUTURE WORK

8.1 Conclusion

The outcomes got to the technique depicted over the represented to the figures beneath. It exhibits the photos catch utilizing with camera on exhibition. It exhibits the prepared image which is that the given to tesseract OCR motor to separating the content from to the photos. It appears the yields via tesseract OCR motors. An exactness might that improved by the utilizing a HD auto mover camera.

The MATLAB are supplanted with Open CV and it's outcomes when at quick preparing. Open CV is the most current apparatus for image prepared was more bankrolling library than this MATLAB. A gadget has contains of a camera inaugurated by that scenes. The processor is use small & might be keep of in interior the pocket to the current customer. The wired associations are provided with during that camera to fast access. Force supplies gave that support causes in to figure for the gadget for around 6 to eight hours. By these highlights with the gadgets turning out to be straightforward, solid and more clients agreeable.

8.2 Future Work

The proposed framework are regularly improved through the extension of during inconsistent parts. Increases of the GPS to this frameworks will authorize the customer to encourage the headings to that give informations with respect to introduce area to the client. Additionally the gadgets are frequently utilized for face acknowledgment. Outwardly debilitated individual who needn't to figure individuals. He can recognize them on the grounds that the camera catching their countenances. GSM module might be attached to the current framework to the execute with push. In the event that the client is in a tough situation, at that point it could be utilize the capture to the glance for aiding by get off the arrangement to some predetermined versatile numbers that may expand an security of visually impaired persons. The apparatus could make the higher outcome if some composition is given to externally exhausted people to we provide with article discovery highlight to the visual narrator, it could recognize objects that are mainly used by the superficially paralysed people groups. Discerning objects like pecuniary standards, tickets, visa cards, numbers or subtleties in PDA & so on could make the lifetime of visually impaired persons plainer. Distinguishing evidence of traffic lights, sign sheets & other land impression might be useful in tripping. Bluetooth office are added in order into dispose of the wired association are joining the exhibition.

9. REFERENCES

- [1] Ms.Rupali, D Dharmale, Dr. P.V. Ingole, "Text Detection and Recognition with Speech Output for Visually Challenged Person", vol. 5, Issue 1, January 2016
- [2] Nagaraja, L., et al. "Vision based text recognition using raspberry PI." National Conference on Power Systems, Industrial Automation (NCPSIA 2015).
- [3] Rajkumar N, Anand M.G, Barathiraja N, "Portable Camera Based Product Label Reading For Blind People.",IJETT, Vol. 10 Number 11 - Apr 2014.
- [4] Boris Epshtein, EyalOfek, Yonatan Wexler, "Detecting Text in Natural Scenes with Stroke Width Transform."
- [5] Ezaki, Nobuo, et al. "Improved text-detection methods for a camera-based text reading system for visually impaired persons." Eighth International Conference on Document Analysis and Recognition (ICDAR'05). IEEE, 2005.
- [6] Ray Smith, "An Overview of the Tesseract OCR Engine."
- [7] Chucai Yi, Yingli Tian and Aries Arditi, "Portable Camera-Based Assistive Text and Product Label Reading from Hand-Held Objects for Visually impaired persons," IEEE/ASME Transactions on Mechatronics, Vol. 19, No. 3, pp. 808, June 2014.
- [8] Sherine Sebastian and Priya S., "Text Detection and Recognition from Images as an Aid to Visually impaired persons Accessing Unfamiliar Environments," Asian Research Publishing Network (ARPN) Journal of Engineering and Applied Sciences, ISSN 1819-6608, Vol. 10, No. 17, September 2015.