International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 06 | June 2020 www.irjet.net p-ISSN: 2395-0072

DEVELOPMENT OF REAL TIME TEXT RECOGNITION SYSTEM FOR VISUALLAY IMPAIRED PEOPLE

B. Hemanth Nag[1], Vadigalla Nagamani[2], Tharun Kumar Amarapalli[3], Rampalli Sandhya[4], Velduti Venu Venkatesh[5]

[1]Assistant Professor, [2,3,4,5]Students
[1-5]Electronics and Communication Engineering, Lendi Institute of Engineering and Technology, Visakhapatnam, India

---

Abstract -The visually impaired people faced many challenges in the day by day life, including these challenges reading the printed texts are often not well understood. This is based on a prototype which helps the user to listen to the contents of the text images in English. The Blind people have methods like Braille, which was introduced as an option for studying engraved text. But it has some issues. The main idea is the development of a system for dictating text for the blind people. The blind people can read the text from the image without taking the help of the human. The Image is captured by the camera module which was connected to the Raspberry Pi. On the captured image different operations are performed by using the open cv. From that image text is extracted by using the OCR and that text is converted into speech.

***Key Words*: Raspberry Pi, Open CV, OCR**

1. INTRODUCTION

Many people do not lose the eye and some people have been blind since childhood. According to the survey conducted by World Health Organization in 2015, some 246 million individuals worldwide are outwardly disabled and 39 million are blind. Many types of investigations have been carried out to solve these problems. Previously, Louis Braille was a French educator and inventor of Braille. This script is a hepatic that is it is difficult to understand by the layman and users must learn it before they understand it.

There have been many advances in this area to help visually impaired to read without many difficulties. The technologies which are already existed use a similar approach as mentioned in this paper, but they have certain drawbacks. Firstly, there is no complex background in the input images taken in previous works have, i.e. on a plain white sheet the test inputs are printed. It is easy to extract text from the image without pre processing, but such approach is not such useful in a real time system. And also in this method the word are not directly recognized, it uses segmentation of characters for recognition, the characters are read out as individual characters. This gives the unsatisfactory audio output to the user.

The Braille which was already existed have the complexity that it requires the text need to be translated into Braille literature for the visually impaired people. As the Braille literature is complex, if the book or document is needed to translated into Braille, it consumes more time and also expensive process. So it is difficult to translate into Braille. So to made easy the process of reading for the visually impaired people, this prototype is used. Using this prototype, text information which was in the captured image can be converted into its equivalent audio format.

As one of the most significant difficulty for a visually impaired person is to read. By developing camera based applications that combine computer vision tools with other existing beneficial products such as Optical Character Recognition (OCR) system, which is feasible to assist the blind person in the recent development in phones, computers, and availability of digital cameras.

2. PROPOSED SYSTEM

The aim of this is to help the blind people to read the text which was present in the image which was captured. This was done by using the methodology which was proposed. For the recognition of the text which was present in the image is done by using the tesseract ocr. But the efficient text recognition is done by using the algorithm called EAST (An Efficient and Accurate Scene Text Detector).

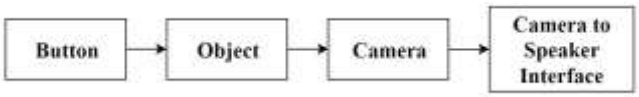


Figure-1: Block Diagram of Transmitter Section

In the transmission section we use the button which is used to control the camera. That is when the button is pressed the camera will capture the image. Here the object is nothing but the object whose image is going to capture. The camera to speaker interface is the Raspberry Pi, where all the operations are done on the image to get the output which is the audio. This is about the transmission section.

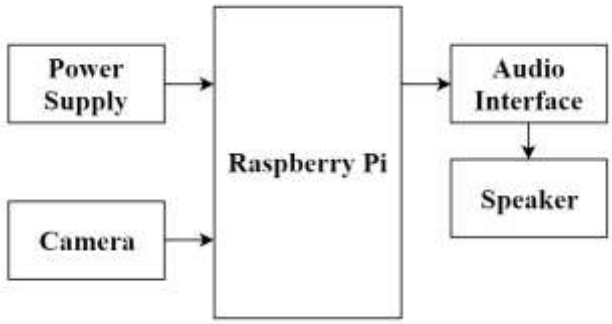


Figure-2: Block Diagram of Receiver Section

In the receiver section the main module used is Raspberry Pi, which is act as the heart to the proposed method. After the capturing of the image by the camera from the transmission section, that image is sent to the Raspberry Pi where the different operations are performed like image resizing, image segmentation, image edge detection etc. Later text is recognition is done(Which will be explained in the methodology). After that by using the speck audio output is extracted, that audio is listened by the speaker but in this project earphones. All these are supplied by the power supply.

3. METHODOLOGY

In this proposed system the methodology which is used:

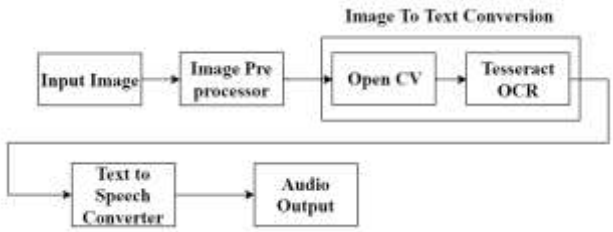


Figure-3: Image to Text Conversion using Tesseract OCR and Raspberry Pi

For the methodology different steps are performed to get the required output. They are

- Input image acquisition
- Image pre processing
- Image to text conversion
- Text to speech conversion
- Finally audio output

3.1 INPUT IMAGE:

The button which is connected is used to operate(control) the camera. By using the camera the input image is captured. That captured image is processed further to get the output.

3.2 IMAGE PREPROCESSING:

The image preprocessing steps are performed to get the efficient output. The fundamental steps are performed on the image, they are image acquisition, image enhancement, image restoration etc.

3.3 IMAGE TO TEXT CONVERSION:

Before going to convert the image to text on the image some operations are performed.

The different image processing operations are done internally by Tesseract (using the Leptonica library) before doing the actual OCR. It generally may do a very good job of this, but there will certainly be cases where it isn't good enough, due to this the result in a significant reduction in accuracy.

Tesseract is an optical character recognition engine for various operating systems. In tesseract ocr image processing uses the configuration variable Tess edit write images to true (or using convexified get. images) when running Tesseract. After done the process if the resulting Tess input. if file looks problematic, some of *these image processing operations before passing the image to Tesseract* done.

3.4 INVERT IMAGES:

The inverting of the image (dark background and light text) without any problem is handled by the tesseract version 3.05, for 4.x version the vice versa will be done that is dark text on light background.

3.5 RESCALING:

Tesseract resizes the best quality images having DPI of at least 300 dpi. It is beneficial of having this range of dpi.

The test for Optimal image resolution with suggestion for optimal Height of capital letter in pixels which is interesting is done by "Willus Dotkom".

3.4 NOISE REMOVAL:

The image may contain some noises like random variation of colour brighteness. Due o that it is difficult to find the text, so it is more difficult to read. These types of noises are not removed by Tesseract I the binarization step, due to this the accuracy rate may decrease.

3.5 DILATION AND EROSION:

Recognition of details and recognition accuracy are effected by bold charcters and thin characters (especially those with Serifs). There are many other image processing programs allow Dilation and Erosion of characters which are oppose the common background to grow in size.

Erosion technique is used to compensate heavy ink bleeding from historical documents. It is also used to shrink characters back to their normal glyph structure.

The extra bold historical fonts by reducing the lower threshold values are explained by the GIMP's Value Propagate filter.

3.6 R0TATION/DESKEWING:

When the image is scanned when it is not straight, a skewed image is formed. If the page is skewed the quality of the Tesseract's line segmentation will reduces obviously due to this the OCR output will automatically decreases. To reduce this error the page image is rotated horizontally.

3.7 SCANNING BORDER REMOVAL:

The pages which have been scanned will have the dark borders around them. Due to this extra characters will be picked up. Highly when it vary in shape and gradation.

3.8 MISSING BORDER:

If the image will have only text area without any extra area for the OCR, then the tesseract will have the problem.

3.9 TRANSPARENCY/ALPHA CHANNEL:

Some image formats can have an alpha-channel for providing a transparency feature. The image which is used before in the tesseract, Tesseract 3.0x wants that user need to remove the alpha channels.This can be done with the command like "ImageMagick":

Convert input.png -alpha off output.png

The alpha channel which was mentioned in the previous will be removed by using the Tesseract 4.00 leptonica function pixRemoveAlpha(): the alpha components are removed by blending it with background. There are some problems like(e.g. OCR of movie subtitles),to avoid these type of problems the user need to remove this alpha channel by themselves like preprocess the iage by inverting image colors).

So by using the tesseract ocr the text will be extracted from the image, this text extraction is efficiently done by using an algorithm called EAST(An Efficient and Accurate Scene Text Detector).

3.10 TEXT TO SPEECH CONVERTER:

The text which was extracted from the image is converted to the audio by using the different modules in the raspberry Pi. In this project eSpeak is used for the audio extraction from the text.

3.11 ESPEAK:

In this system the open source software speech synthesizer for English other languages is needed, for this purpose eSpeak is used for Linux and Windows. This eSpeak is simple, easy to use and support other languages. The advantage of the using the eSpeak is it does not use the internet. So by using this eSpeak module we need to get the audio output.

Finally the extracted audio output from the image is listened by the earphones.

3.12 ALGORITHAM WHICH WAS USED:

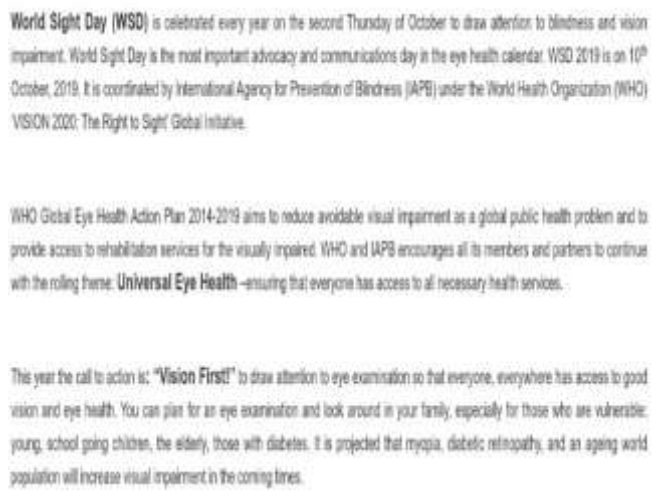
The algorithm which was used will use the key component is a neural network model, this model will directly predict the existence of the text and the full geometries from full image.

The neural model is a fully convolutional which is adapted for the text detection that outputs dense per-pixel predictions of words or text lines. By using this intermediate steps like candidate proposal, text region formation and word partition will be eliminated. Finally the post-processing steps only include thresholding and NMS on predicted geometric shapes. This type of the detector is named as EAST, since it is an Efficient and Accuracy Scene Text detection

4. RESULTS AND ANALYSIS

From the raspberry Pi connected to the computer to getting the audio output which was listening is done by using this proposed system. All these operations are controlled by using the single button which was operated by the visually impaired people. After the establishment of all the connections, Raspbian operating system is logged in. For the pi the external memory is needed for that the SD card is used. That SD card is booted and fixed to its respective slot. After the completion of all installations and updations, the image is captured by the camera. That image is processed in the pi by using the libraries which was discussed previously that is open cv, tesseract ocr, EAST, eSpeak. The processes is done by the python program. The input image which was taken is processed by using both tesseract and EAST algorithm for the efficient algorithm.

A. Output Using Tesseract:



World Sight Day (WSD) is celebrated every year on the second Thursday of October to draw attention to blindness and vision impairment. World Sight Day is the most important advocacy and communications day in the eye health calendar. WSD 2019 is on 10th October, 2019. It is coordinated by International Agency for Prevention of Blindness (IAPB) under the World Health Organization (WHO) VISION 2020: The Right to Sight' Global Initiative.

WHO Global Eye Health Action Plan 2014-2019 aims to reduce avoidable visual impairment as a global public health problem and to provide access to rehabilitation services for the visually impaired. WHO and IAPB encourages all its members and partners to continue with the rolling theme: **Universal Eye Health** –ensuring that everyone has access to all necessary health services.

This year the call to action is: **"Vision First!"** to draw attention to eye examination so that everyone, everywhere has access to good vision and eye health. You can plan for an eye examination and look around in your family, especially for those who are vulnerable: young, school going children, the elderly, those with diabetes. It is projected that myopia, diabetic retinopathy, and an ageing world population will increase visual impairment in the coming times.

Figure-4: Case Study – 1 with 203 words in the image

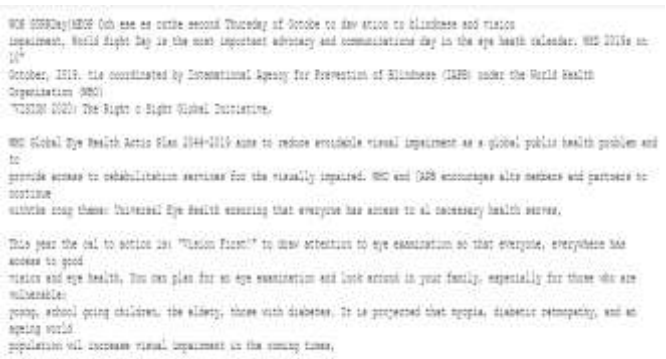


Figure-5: Screenshot of Text Document obtained as output containing 198 words

So, the percentage of errors can be calculated as below.

(%)Error = Expected Value – Actual Value X 100

Actual Value

= 203-198 X 100

198

= 2.53 %

The text which was extracted from the above image is converted into audio, which will be played that screen short is shown in the below image.

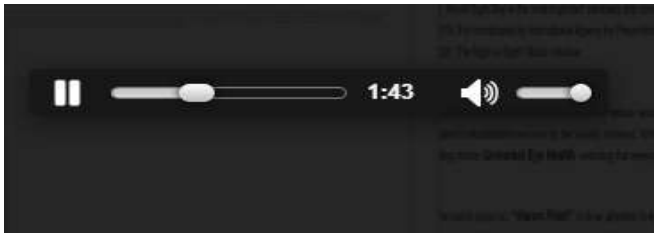


Figure-6: Screenshot of Audio File Output of the text extracted from input image

Feb 15, 2017

PSLV-C37 Successfully Launches 104 Satellites in a Single Flight

In its thirty ninth flight (PSLV-C37), ISRO's Polar Satellite Launch Vehicle successfully launched the 714 kg Cartosat-2 Series Satellite along with 103 co-passenger satellites today morning (February 15, 2017) from Satish Dhawan Space Centre SHAR, Sriharikota. This is the thirty eighth consecutive successful mission of PSLV. The total weight of all the 104 satellites carried on-board PSLV-C37 was 1378 kg.

PSLV-C37 lifted off at 0928 hrs (9:28 am) IST, as planned, from the First Launch Pad. After a flight of 16 minutes 48 seconds, the satellites achieved a polar Sun Synchronous Orbit of 506 km inclined at an angle of 97.46 degree to the equator (very close to the intended orbit) and in the succeeding 12 minutes, all the 104 satellites successfully separated from the PSLV fourth stage in a predetermined sequence beginning with Cartosat-2 series satellite, followed by INS-1 and INS-2. The total number of Indian satellites launched by PSLV now stands at 46.

After separation, the two solar arrays of Cartosat-2 series satellite were deployed automatically and ISRO's Telemetry, Tracking and Command Network (ISTRAC) at Bangalore took over the control of the satellite. In the coming days, the satellite will be brought to its final operational configuration following which it will begin to provide remote sensing services using its panchromatic (black and white) and multispectral (colour) cameras.

Of the 103 co-passenger satellites carried by PSLV-C37, two – ISRO Nano Satellite-1 (INS-1) weighing 8.4 kg and INS-2 weighing 9.7 kg – are technology demonstration satellites from India.

The remaining 101 co-passenger satellites carried were international customer satellites from USA (96), The Netherlands (1), Switzerland (1), Israel (1), Kazakhstan (1) and UAE (1).

With today's successful launch, the total number of customer satellites from abroad launched by India's workhorse launch vehicle PSLV has reached 180.

Figure-7: Case Study – 2 with 305 words in the image

Figure-8: Screenshot of Text Document obtained as output containing 301 words

So, the percentage of errors can be calculated as below.

(%)Error = Expected Value – Actual Value X 100

Actual Value

= 305-301 X 100

301

= 1.33%

The text which was extracted from the above image is converted into audio, which will be played that screen short is shown in the below image.

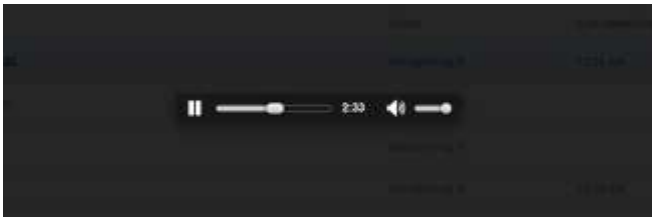
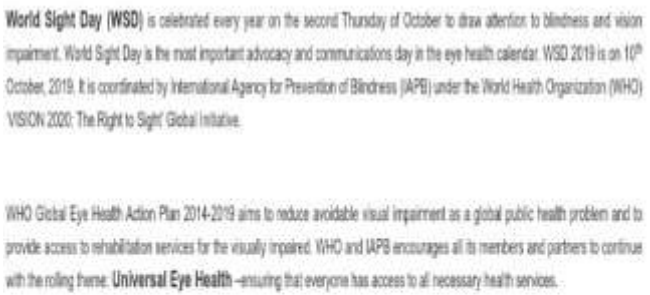


Figure-9: Screenshot of Audio File Output of the text extracted from input image

B. Outputs Using EAST Algorithm

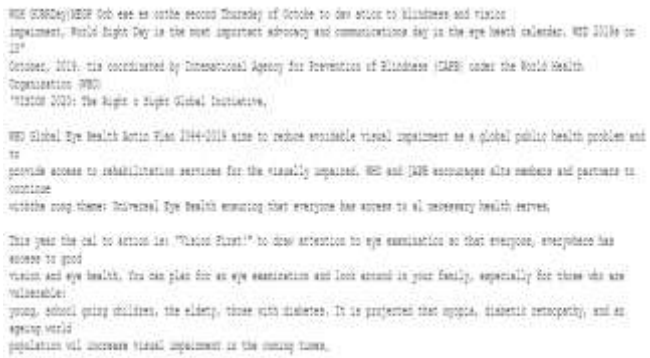


World Sight Day (WSD) is celebrated every year on the second Thursday of October to draw attention to blindness and vision impairment. World Sight Day is the most important advocacy and communications day in the eye health calendar. WSD 2019 is on 10th October, 2019. It is coordinated by International Agency for Prevention of Blindness (IAPB) under the World Health Organization (WHO) VISION 2020: The Right to Sight' Global Initiative.

WHO Global Eye Health Action Plan 2014-2019 aims to reduce avoidable visual impairment as a global public health problem and to provide access to rehabilitation services for the visually impaired. WHO and IAPB encourages all its members and partners to continue with the rolling theme: Universal Eye Health –ensuring that everyone has access to all necessary health services.

This year the call to action is: "Vision First!" to draw attention to eye examination so that everyone, everywhere has access to good vision and eye health. You can plan for an eye examination and look around in your family, especially for those who are vulnerable: young, school going children, the elderly, those with diabetes. It is projected that myopia, diabetic retinopathy, and an ageing world population will increase visual impairment in the coming times.

Figure-10: Case Study – 1 with 203 words in the image



World Sight Day (WSD) is celebrated every year on the second Thursday of October to draw attention to blindness and vision impairment. World Sight Day is the most important advocacy and communications day in the eye health calendar. WSD 2019 is on 10th October, 2019. It is coordinated by International Agency for Prevention of Blindness (IAPB) under the World Health Organization (WHO) 'VISION 2020: The Right to Sight' Global Initiative.

WHO Global Eye Health Action Plan 2014-2019 aims to reduce avoidable visual impairment as a global public health problem and to provide access to rehabilitation services for the visually impaired. WHO and IAPB encourages all its members and partners to continue with the rolling theme: Universal Eye Health ensuring that everyone has access to all necessary health services.

This year the call to action is: "Vision First!" to draw attention to eye examination so that everyone, everywhere has access to good vision and eye health. You can plan for an eye examination and look around in your family, especially for those who are vulnerable: young, school going children, the elderly, those with diabetes. It is projected that myopia, diabetic retinopathy, and an ageing world population will increase visual impairment in the coming times.

Figure-11: Screenshot of Text Document obtained as output containing 200 words

So, the percentage of errors can be calculated as below.

(%)Error = $\frac{\text{Expected Value – Actual Value}}{\text{Actual Value}}$ X 100

= $\frac{203-200}{200}$ X 100

= 1.50%

Figure-12: Screenshot of Audio File Output of the text extracted from input image

Feb 15, 2017

PSLV-C37 Successfully Launches 104 Satellites in a Single Flight

In its thirty ninth flight (PSLV-C37), ISRO's Polar Satellite Launch Vehicle successfully launched the 714 kg Cartosat-2 Series Satellite along with 103 co-passenger satellites today morning (February 15, 2017) from Satish Dhawan Space Centre SHAR, Sriharikota. This is the thirty eighth consecutively successful mission of PSLV. The total weight of all the 104 satellites carried on-board PSLV-C37 was 1378 kg.

PSLV-C37 lifted off at 0928 hrs (9:28 am) IST, as planned, from the First Launch Pad. After a flight of 16 minutes 48 seconds, the satellites achieved a polar Sun Synchronous Orbit of 506 km inclined at an angle of 97.46 degree to the equator (very close to the intended orbit) and in the succeeding 12 minutes, all the 104 satellites successfully separated from the PSLV fourth stage in a predetermined sequence beginning with Cartosat-2 series satellite, followed by INS-1 and INS-2. The total number of Indian satellites launched by PSLV now stands at 46.

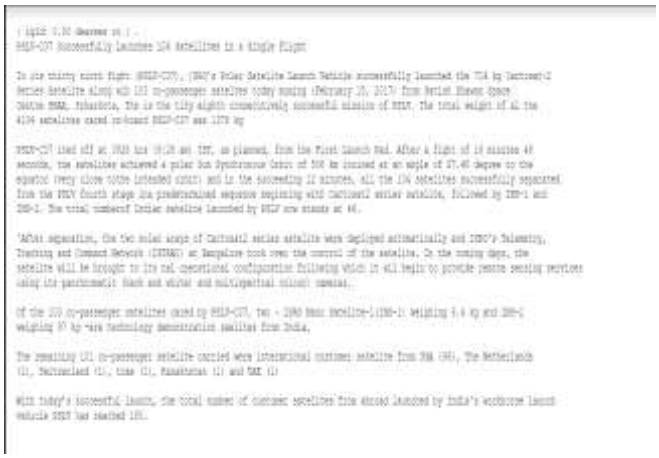
After separation, the two solar arrays of Cartosat-2 series satellite were deployed automatically and ISRO's Telemetry, Tracking and Command Network (ISTRAC) at Bangalore took over the control of the satellite. In the coming days, the satellite will be brought to its final operational configuration following which it will begin to provide remote sensing services using its panchromatic (black and white) and multispectral (colour) cameras.

Of the 103 co-passenger satellites carried by PSLV-C37, two – ISRO Nano Satellite-1 (INS-1) weighing 8.4 kg and INS-2 weighing 9.7 kg – are technology demonstration satellites from India.

The remaining 101 co-passenger satellites carried were international customer satellites from USA (96), The Netherlands (1), Switzerland (1), Israel (1), Kazakhstan (1) and UAE (1).

With today's successful launch, the total number of customer satellites from abroad launched by India's workhorse launch vehicle PSLV has reached 180.

Figure-13: Case Study – 2 with 305 words in the image



Feb 15, 2017

PSLV-C37 Successfully Launches 104 Satellites in a Single Flight

In its thirty ninth flight (PSLV-C37), ISRO's Polar Satellite Launch Vehicle successfully launched the 714 kg Cartosat-2 Series Satellite along with 103 co-passenger satellites today morning (February 15, 2017) from Satish Dhawan Space Centre SHAR, Sriharikota. This is the thirty eighth consecutively successful mission of PSLV. The total weight of all the 104 satellites carried on-board PSLV-C37 was 1378 kg.

PSLV-C37 lifted off at 0928 hrs (9:28 am) IST, as planned, from the First Launch Pad. After a flight of 16 minutes 48 seconds, the satellites achieved a polar Sun Synchronous Orbit of 506 km inclined at an angle of 97.46 degree to the equator (very close to the intended orbit) and in the succeeding 12 minutes, all the 104 satellites successfully separated from the PSLV fourth stage in a predetermined sequence beginning with Cartosat-2 series satellite, followed by INS-1 and INS-2. The total number of Indian satellites launched by PSLV now stands at 46.

After separation, the two solar arrays of Cartosat-2 series satellite were deployed automatically and ISRO's Telemetry, Tracking and Command Network (ISTRAC) at Bangalore took over the control of the satellite. In the coming days, the satellite will be brought to its final operational configuration following which it will begin to provide remote sensing services using its panchromatic (black and white) and multispectral (colour) cameras.

Of the 103 co-passenger satellites carried by PSLV-C37, two – ISRO Nano Satellite-1 (INS-1) weighing 8.4 kg and INS-2 weighing 9.7 kg – are technology demonstration satellites from India.

The remaining 101 co-passenger satellites carried were international customer satellites from USA (96), The Netherlands (1), Switzerland (1), Israel (1), Kazakhstan (1) and UAE (1).

With today's successful launch, the total number of customer satellites from abroad launched by India's workhorse launch vehicle PSLV has reached 180.

Figure-14: Screenshot of Text Document obtained as output containing 303 words

So, the percentage of errors can be calculated as below.

(%)Error = $\frac{\text{Expected Value – Actual Value}}{\text{Actual Value}}$ X 100

= 305-303 X 100

303

= 0.66%

Figure-15: Screenshot of Audio File Output of the text extracted from input image

5. CONCLUSION

Finally the implementation of the real time text recognition for visually impaired people is implemented. This development is done by using the key modules like open cv, Tesseract ocr,espeak. The algorithm which was used is EAST. The text will be extracted from the image by the tesseract,but it will be efficiently done by the EAST algorithm. By using the eSpeak the audio output is extracted. This one will helpful for the visually challenged people. This one is applied to many images and found that it successfully does its conversion.

6. FUTURE SCOPE

Up to this proposed one only if the images will have the text then it will be converted, if not then it simply said that it does not have the text. Further it will be improved as the image which was captured need to be recognized like it is an animal or like fruit. This is the future scope.

REFERENCES

[1].Christo Liambas,Miltiadis Saratzidis(2016),Autonomous OCR Dictating System for Blind People,in IEEE 2016 Global Humanitarian Technology Conference.

[2].Dr. I S Akila , Akshaya B, Deepthi S, Sivadharshini P(2018),A Text Reader for the Visually Impaired using Raspberry Pi, in Proceedings of the Second International Conference on Computing Methodologies and Communication (ICCMC 2018) IEEE Conference Record.

[3]. Ani R, Effy Maria(2017),Smart Specs: Voice Assisted Text Reading system for

Visually Impaired Persons Using TTS Method, in IEEE International Conference on Innovations in Green Energy and Healthcare Technologies(ICIGEHT'17).

[4]. Shalini Sonth, Jagadish S Kallimani (2017),OCR Based Facilitator for the Visually Challenged, in 2017 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT).

[5]. Rajesh M, Bindhu K. Rajan(2017),TEXT RECOGNITION AND FACE DETECTION

AID FOR VISUALLY IMPAIRED PERSON USING RASPBERRY PI, in 2017 International Conference on circuits Power and Computing Technologies [ICCPCT].

[6]. V. Mareeswari, Vijayan Ramaraj, G. Uma Maheswari, R. Sujatha, Preethi. E(2019),Accessibility Using Human Face, Object and Text Recognition for Visually Impaired People, in International Journal of Innovative Technology and Exploring Engineering (IJITEE).

[7]. G Nagalakshmi, K Maham, P Muthukrishnammal, Ganesan P, and T Jerry Alexander(2016), Raspberry Pi Based Optical Character Recognition for Smart and Intelligent Book Reader,in Research Journal of Pharmaceutical, Biological and Chemical Sciences.

[8].Deep Learning based Text Recognition (OCR) using Tesseract and OpenCV _ Learn OpenCV.mhtml.

BIOGRAPHIES

B. Hemanth Nag
Lendi Institute of Engineering and Technology, Dept. of ECE
JNTU Kakinada
Vizianagaram, Andhra Pradesh, India

Vadigalla Nagamani
Lendi Institute of Engineering and Technology,Dept.ofECE
JNTUKakinada
Vizianagaram, Andhra Pradesh, India

Tharun Kumar Amarapalli
Lendi Institute of Engineering and Technology, Dept. of ECE
JNTU Kakinada
Vizianagaram, Andhra Pradesh, India

RampalliSandhya
Lendi Institute of Engineering and Technology, Dept. of ECE
JNTU Kakinada Vizianagaram, Andhra Pradesh, India

Velduti Venu Venkatesh
Lendi Institute of Engineering and Technology, Dept. of ECE
JNTU Kakinada
Vizianagaram, Andhra Pradesh, India