International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 02, 2020
ISSN: 1475-7192

TEXT PICTURE TRANSLATION INTO SPEECH IN DESIRED LANGUAGE

[1]Anugna akkala,[2]S. Prem kumar

ABSTRACT-- *Text to speech is a technology through which computer is made to speak. If a study is performed like listening or reading and which is more effective definitely answer is listening, because people will understand more by listening than reading. In many research(s) it is proved that listening improves imagination power. The main aim is to provide speech output in real time for blind people when they gave printed text as input. This text to speech requires two technologies; Optical Character Recognition (OCR) and text to speech conversion. The former technique is used for extracting text from image. The later technology converts the extracted text into speech. In any reading assistant system, the first and important function is text information extraction and this is the integral part of OCR. To recognize words OCR is used.*

Key words--

I INTRODUCTION

This OCR system is able to recognize words, sentences and characters without any mistakes. High rate of recognition is possible with OCR that undergoes electronic conversion and gives computer readable text from photograph images of printed and type written. Text to speech conversion system is an aid to blind persons and makes them to participate actively. The hardcopy of a book is converted by text to speech conversion system into another audio format such as into mp3. The system can be used for blind people effectively. It follows certain steps while converting an image with text into speech output. They are (i) Getting the input image (ii) Reading that input image (iii) Processing the image (iv) Image is converted into text (v) Text to speech conversion (vi) Final output as speech in desired language.

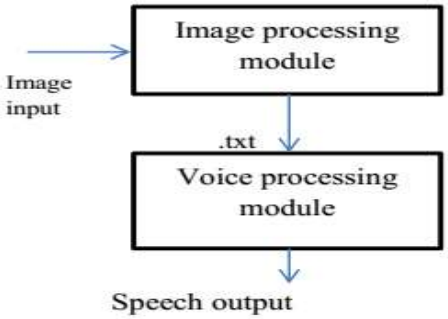


Figure 1: text to speech conversion

[1] *Student, , Department of ECE, Saveetha school of engineering, Chennai, India, Email:anugnareddyar@gmail.com*

[2] *Student, Department of ECE, Saveetha school of engineering, Chennai, India*

Received: 23 Dec 2019 | Revised: 05 Jan 2020 | Accepted: 27 Feb 2020

The books, magazines and newspapers whatever want to read can be made as a scanned document and can hear by using this text to speech convertor. So, no need of carrying heavy books everywhere. Since the output is voice the user can easily hear it even doing some other work. So, time is utilized. OCR involves certain steps for text extraction from image. The first step in OCR is acquiring the image and reading it. Then convert the image into its gray scale. These steps are called as pre-processing.

The data would not be affected when it is converted from gray to binary. For the gray image, a threshold is computed first and then using that threshold black and white image is formed. The situation of white area is between the black. Next boundary of that image is to be founded and image is cropped till the edge. Then characters in the image are extracted and that letters are resized according to size of template. Feature extraction and classification are very important in OCR. Thus the image is converted into speech. It recognizes both small letters and capital letters. It can also recognize numbers. The input character size should be 12 pt. Minimum tilt of the text line from vertical is 40 to 50.

Image magick software is used for converting image into binary image. This is used for image manipulation and is an open source tool. While taking the input image it will take a delay of 7 seconds. A python code is written for extraction of text from input image, then to translate that text in four languages (Tamil, Telugu, Hindi, Kannada) and finally to get speech output in the same 4 languages.

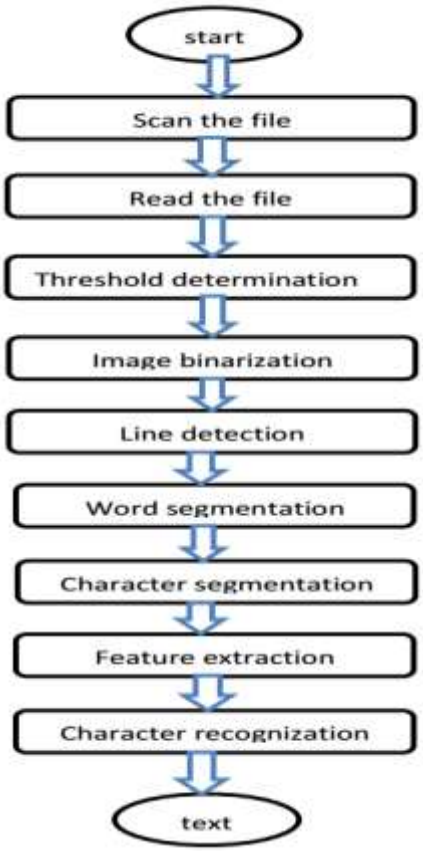


Figure 2: Image processing module

Received: 23 Dec 2019 | Revised: 05 Jan 2020 | Accepted: 27 Feb 2020

Python is an object oriented language. It has extensive standard libraries which to integrated with other languages. . This system not only converts text to speech, it is also gives speech output in desired language. The system is mainly designed for blind people since they cannot read the text.

In speech synthesis, the human voice is produced artificially and the device is called as speech synthesizer. It converts normal text into speech. The output from any speech synthesizer can be used for both software and hardware process. The speech is created from the stored database which is collection of information. If the speech output produced is clearly understood then the quality of speech synthesizer is high. Text to speech engine has two parts front end and back end. The former converts abbreviations and numbers into written out words and this process is called text normalization.

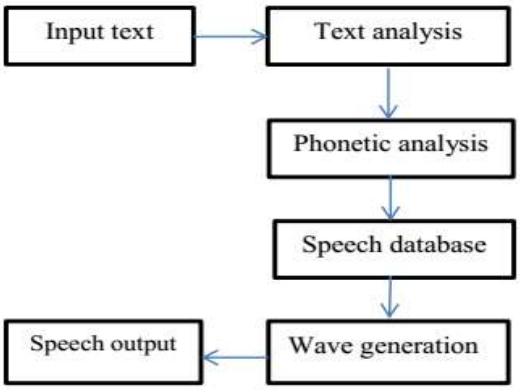


Figure 3: Voice processing module

Then it will allocate phonetic transcriptions to each word and marks and divides the text into classes, sentences and phases. This process is called grapheme to phoneme conversion. This is formed as speech output. The main aim of this project is getting output in our desired language. Google text to speech application is used for getting output in desired language. It is developed by Google it reads the text that is available and changes that text in other languages. This is called machine translation. At initial stage machine translation performs only some small substitution of words from one language to other language. But this one word translation cannot give good output if input is a whole phrase. So, human translation process is done for proper sentence formation. It follows certain steps first that is decoding the meaning of input text and then recode it in target language. By using this system we can able to get output in many languages.

II RESULTS


Saveetha Institute of Medical And Technical Sciences is a Private university located in Chennai, Tamil Nadu, India. It has eight disciplines of studies: Dental College, School of Management, School of Law, School of Engineering, School of Physiotherapy, School of Nursing and Medical College.[1] The first three disciplines are in Poonamalle while the rest are in Thandalam. Saveetha Engineering College is an Anna University-affiliated institution. Admissions are done through Class 12th Indian board examinations.

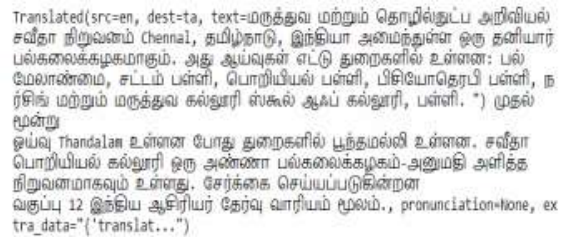
Figure 4: Input image

Received: 23 Dec 2019 | Revised: 05 Jan 2020 | Accepted: 27 Feb 2020

```
filename=str(input("Enter Image Name: "))
img=filename
image = Image.open(img, mode='r')
txt=image_to_string(image,lang='eng')
print(txt)
```

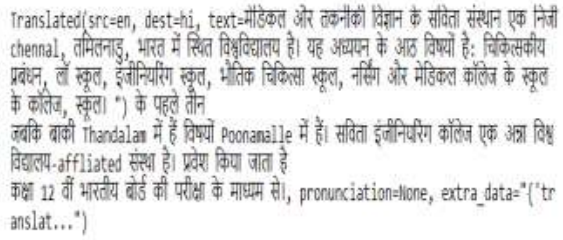
```
Enter Image Name: saveethainput.jpg
Saveetha Institute of Medical And Technical Sciences is a Private university
located in Chennai, Tamil Nadu, India. It has eight disciplines of studies: D
ental
College, School of Management, School of Law, School of Engineering, School o
f Physiotherapy, School of Nursing and Medical College.") The first three
disciplines are in Poonamalle while the rest are in Thandalam. Saveetha Engin
eering College is an Anna University-affiliated institution. Admissions are do
ne
through Class 12th Indian board examinations.
```

Figure 5: Code for text extraction from input image



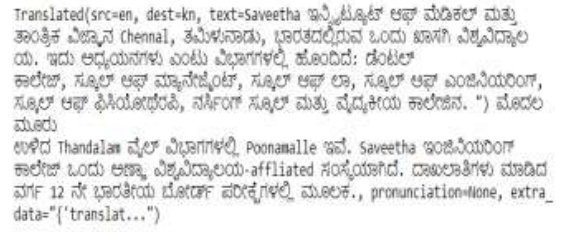
Translated(src=en, dest=ta, text=மருத்துவ மற்றும் தொழில்நுட்ப அறிவியல் சவீதா நிறுவனம் Chennal, தமிழ்நாடு, இந்தியா அமைந்துள்ள ஒரு தனியார் பல்கலைக்கழகமாகும். அது ஆய்வுகள் எட்டு துறைகளில் உள்ளன: பல் மேலாண்மை, சட்டம் பள்ளி, பொறியியல் பள்ளி, பிசியோதெரபி பள்ளி, நர்சிங் மற்றும் மருத்துவ கல்லூரி ஸ்கூல் ஆஃப் கல்லூரி, பள்ளி. ") முதல் மூன்று
ஒய்வு Thandalam உள்ளன போது துறைகளில் பூந்தமல்லி உள்ளன. சவீதா பொறியியல் கல்லூரி ஒரு அண்ணா பல்கலைக்கழகம்-அனுமதி அளித்த நிறுவனமாகவும் உள்ளது. சேர்க்கை செய்யப்படுகின்றன
வகுப்பு 12 இந்திய ஆசிரியர் தேர்வு வாரியம் மூலம்., pronunciation=None, extra_data="{'translat...")

Figure 6: Output in Tamil language



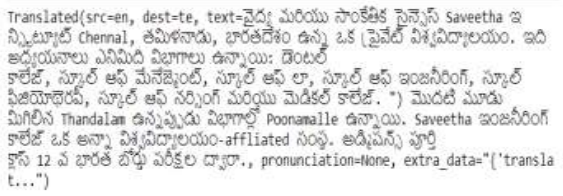
Translated(src=en, dest=hi, text=मेडिकल और तकनीकी विज्ञान के सविता संस्थान एक निजी chennal, तमिलनाडु, भारत में स्थित विश्वविद्यालय है। यह अध्ययन के आठ विषयों है: चिकित्सकीय प्रबंधन, लॉ स्कूल, इंजीनियरिंग स्कूल, भौतिक चिकित्सा स्कूल, नर्सिंग और मेडिकल कॉलेज के स्कूल के कॉलेज, स्कूल। ") के पहले तीन
जबकि बाकी Thandalam में हैं विषयों Poonamalle में हैं। सविता इंजीनियरिंग कॉलेज एक अन्ना विश्वविद्यालय-affliated संस्था है। प्रवेश किया जाता है
कक्षा 12 वीं भारतीय बोर्ड की परीक्षा के माध्यम से।, pronunciation=None, extra_data="{'translat...")

Figure 7: Output in Hindi language



Translated(src=en, dest=kn, text=Saveetha ಇನ್ಸ್ಟಿಟ್ಯೂಟ್ ಆಫ್ ಮೆಡಿಕಲ್ ಮತ್ತು ತಾಂತ್ರಿಕ ವಿಜ್ಞಾನ Chennal, ತಮಿಳುನಾಡು, ಭಾರತದಲ್ಲಿರುವ ಒಂದು ಖಾಸಗಿ ವಿಶ್ವವಿದ್ಯಾಲಯ. ಇದು ಅಧ್ಯಯನಗಳು ಎಂಟು ವಿಭಾಗಗಳಲ್ಲಿ ಹೊಂದಿದೆ: ಡೆಂಟಲ್
ಕಾಲೇಜ್, ಸ್ಕೂಲ್ ಆಫ್ ಮ್ಯಾನೇಜ್ಮೆಂಟ್, ಸ್ಕೂಲ್ ಆಫ್ ಲಾ, ಸ್ಕೂಲ್ ಆಫ್ ಎಂಜಿನಿಯರಿಂಗ್, ಸ್ಕೂಲ್ ಆಫ್ ಫಿಸಿಯೋಥೆರಪಿ, ನರ್ಸಿಂಗ್ ಸ್ಕೂಲ್ ಮತ್ತು ಮೆಡಿಕಲ್ ಕಾಲೇಜಿನ. ") ಮೊದಲ ಮೂರು
ಉಳಿದ Thandalam ಮೇಲ್ ವಿಭಾಗಗಳಲ್ಲಿ Poonamalle ಇವೆ. Saveetha ಇಂಜಿನಿಯರಿಂಗ್ ಕಾಲೇಜ್ ಒಂದು ಅಣ್ಣಾ ವಿಶ್ವವಿದ್ಯಾಲಯ-affliated ಸಂಸ್ಥೆಯಾಗಿದೆ. ದಾಖಲಾತಿಗಳು ಮಾಡಿದ ಮಗ 12 ನೇ ಭಾರತೀಯ ಬೋರ್ಡ್ ಪರೀಕ್ಷೆಗಳಲ್ಲಿ ಮೂಲಕ., pronunciation=None, extra_data="{'translat...")

Figure 8: Output in Kannada language



Translated(src=en, dest=te, text=వైద్య మరియు సాంకేతిక సైన్సెస్ Saveetha ఇన్స్టిట్యూట్ Chennal, తమిళనాడు, భారతదేశం ఉన్న ఒక ప్రైవేట్ విశ్వవిద్యాలయం. ఇది అధ్యయనాలు ఎనిమిది విభాగాలు ఉన్నాయి: డెంటల్
కాలేజ్, స్కూల్ ఆఫ్ మేనేజ్మెంట్, స్కూల్ ఆఫ్ లా, స్కూల్ ఆఫ్ ఇంజనీరింగ్, స్కూల్ ఫిజియోథెరపీ, స్కూల్ ఆఫ్ నర్సింగ్ మరియు మెడికల్ కాలేజ్. ") మొదటి మూడు
మిగిలిన Thandalam ఉన్నప్పుడు విభాగాల్లో Poonamalle ఉన్నాయి. Saveetha ఇంజనీరింగ్ కాలేజ్ ఒక అన్నా విశ్వవిద్యాలయం-affliated సంస్థ. అడ్మిషన్స్ పూర్తి
క్లాస్ 12 వ భారత బోర్డు పరీక్షల ద్వారా., pronunciation=None, extra_data="{'translat...")

Figure 9: Output in Telugu language

```
obj=gTTS(text=t,lang='ta',slow=False)
obj.save('Tamil.mp3')
playsound('Tamil.mp3')

obj1=gTTS(text=t1,lang='te',slow=False)
obj1.save('Telugu.mp3')
playsound('Telugu.mp3')

obj2=gTTS(text=t2,lang='hi',slow=False)
obj2.save('Hindi.mp3')
playsound('Hindi.mp3')

obj3=gTTS(text=t3,lang='ka',slow=False)
obj3.save('Kanadam.mp3')
playsound('Kanadam.mp3')
```

Figure 10: Code for speech output in 4 languages

III CONCLUSION

The system text picture translation into speech in desired language is very useful to blind people and people cannot understand English. This system operation is very simple. Not only Tamil, Telugu, Hindi and Kannada languages other languages output is also possible. Just we need to change the code and run the code. This is low cost and highly efficient.

REFERENCES

1. R. Mithe, S. Indalkar and N. Divekar. ” Optical Character Recognition" International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878(2013),Volume-2, Issue-1,.
2. ARCHANA A. and SHINDE, D. “Text Preprocessing and Text Segmentation for OCR”. International Journal of Computer Science Engineering and Technology, pp. (2012)810-812.
3. Nikisha jariwala, bankin patel “A system for the conversion of digital gujarati text to speech for visually impaired people.” Research gate january 2018.
4. Tapas kumar patra, biplabpatra, puspanjali mohapatra “Text to Speech Conversion with Phonematic Concatenation” International Journal of electronics communication and computer technology volume 2 (2012).
5. Anna Pribilava, Jiripribil “non-linear frequency scale mapping for voice conversion in text to speech system with cepstral description” www. Elsever.com august 2006.
6. Sanket munot, akshay patil, utkarsha kandale “image to speech conversion website” ,IJSART vol 5, issue 9.
7. Runmayee Patil, Ramesh Kagalkar “A review on conversion of image to text as well as speech using edge detection and image segmentation .International journal of science and research(IJSR) ”(2014)volume 3 issue 11..

8. Sagar patil, Mayuri phonde, siddarth prajapati, saraga rane " Multilingual speech and text recognition and translation using image." international journal of engineering research and technology(IJERT),(2016) volume 5.
9. Sangramsing, Monica mundada, santhosh gaikwad, bharti gawali N"performance evaluation of speech synthesis techniques for english language." springer science +business media Singapore (2016).
10. D.B.K.kamesh, s. nazma, J.K.R. sastry and s.venkateswarlu " camera based text to speech conversion , obstacle and currency detection for blind people." International journal of science and technology, (2016)volume 9(30).
11. Chaw su thu thu " implementation of text to speech conversion" international journal of engineering research and technology,(2014) vol 3, issue 3.
12. Burton L. levin "text to speech conversion system" September 28, 2010.
13. V. Ajantha devi, Dr. S. Santhosh baboo "Embedded OCR on Tamil text image using raspberry pi", international journal of computer science trends and technology". (2014)Volume 2.
14. Rithika.h, B. Nithya santhoshini, "Image text to speech conversion in desired language by translating with raspberry pi", IEEE, 2016
15. Saravanan, N. (2013). "Hand Geometry Recognition based on optimized K-means Clustering and Segmentation Algorithm". International Journal of MC Square Scientific Research, 5(1), 11-14.
16. Maharaja, D., & Shaby, M. (2017). "Empirical Wavelet Transform and GLCM Features Based Glaucoma Classification from Fundus Image". International Journal of MC Square Scientific Research, 9(1), 78-85.

Received: 23 Dec 2019 | Revised: 05 Jan 2020 | Accepted: 27 Feb 2020