COPY RIGHT

ELSEVIER
SSRN

Paper Authors

R S M Lakshmi Patibandla, Madupalli Manoj, Alapati Jagadeesh, Bathina Sasidhar

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

Automatic Transcriptor for Intellectually Disabled People

R S M Lakshmi Patibandla*, Madupalli Manoj, Alapati Jagadeesh, Bathina Sasidhar
*Department of IT & CA,
School of Computing and Informatics,
VFSTR Deemed to be University, Guntur, AP, India
patibandla.lakshmi@gmail.com

Abstract

Our technology's major objective is to create voice recognition systems for those with physical limitations. Recently, we are increasingly using a large number of individual devices for communication purposes. Developing a person's talent is one of the positive trends. The main objective of our technique is to develop voice granting systems to those who have physical constraints. The text can come from the audio in the audio file or from another type of language. As a result, you can read books aloud, speak instead of writing, as well as extract text from images.

This program is used by users who have disabilities such as deafness, blindness, or other impairments as well as those who find typing to be difficult, inconvenient, or impossible. You can benefit. His GUI/voice commands in this file are shown in the easy-to-use Python code for this project using Visual Studio.

Keywords — Python, Visual Studio Code, Voice Recognition, Versatile, Graphical User Interface

Introduction

The main goal is to create an Automatic Transcriptor for Intellectually Disabled people. The audiobook system only reads text, Portable Document Format, or captions. Audio to Text, Text to Audio, Portable Document Format to Audio, Image to Audio conversions is also a part of this system. You can input audio with voice-to-text and output text with text-to-text. Portable Document Format to Voice, on the other hand, can take an image or Portable Document Format file as input and produce an output in audio format. Generates output in audio format for each word in an image, or takes input as text and generates output in audio format. People who're deaf or blind, have little literacy, or who can study however now no longer communicate can gain from this, and it's going to additionally unfasten up it slow to finish different chores. If they communicate approximately being capable of realize their personal voice, human beings might be greater fascinated. Anyone can get right of entry to it and watch it from anywhere. It is straightforward to convert it into other kinds way to this utility. This will in all likelihood be the first-class and maximum user-pleasant net utility with inside the future.

Modules

• Text-to-Voice (TTS): The main purpose of the TTS engine was to convert the provided input text into voice. Finding out how the text is pronounced can be helpful if users are unable to pronounce certain text. It is also recommended for those who are not good at speaking.

• Portable Document Format to Voice: You can read books in Portable Document Format using the Portable Document Format to Audio module.

• Picture-to-Voice (p2S): This module takes input from the user in the form of pictures and then converts the text of the pictures into voice.

• Voice-to-Text: This feature converts spoken words into text.

• Google's AIVA-It tool is recommended for creating voice-controlled individual workers that can perform different works.

Literature Review

Prior to starting the Design, here we carried out the survey for finding out the number of individuals were and were not interested in using audiobooks. As a result, we came to the conclusion that most People are drawn to audiobooks. I did some research to find out why people love and don't like audiobooks.

According to Nikhat Parveen et al. Most people find listening more comfortable than reading.

This is especially useful for people who understand the language but are illiterate. Deaf and blind people can also greatly benefit from this audio book system. Because you can understand and feel the text better when you listen to it than when you read it. It is beneficial for both educators and students. According to Devidutta Dash et al. Teachers these days are moving from speaking with chalk to teaching with touch. The practice of independent learning has taken root in students. He doesn't like listening to audiobooks because some people think it reduces his ability to read. Some people think that to improve their language skills, they should read books instead of listening to them. Hasan U. Zaman and others say Some people enjoy not being able to read because of literacy issues, physical limitations, or the fact that they are capable of doing other things while listening to a book. Others argue that you don't have to always have a book with you when an unexpected interest strikes. Research shows Automatic Transcriptor has a format having more benefits than drawbacks.

Methods

This project uses Python(py) entirely. I have used Visual Studio code. Our Design is completely connected through internet. Probably, there are 4 modules in all: photo to speech and Text, Portable Document Format to speech, Text to voice, and Text to voice.

VOICE TECHNOLOGY	BRAIN TECHNOLOGY
Voice Activation	Voice Biometrics
Automatic Speech Recognition (ASR)	Dialog Management
(Teach-To-Speech (TTS)	Natural Language Understanding (NLU)
	Named Entity Recognition NER)

TABLE 1: Technology to create artificial intelligence systems that speak to people in their native language.

Text to Voice

TTS can be supported with text-to-voice. By clicking the submit button on your laptop or desktop device, you can transfer text into an audio file.

Text-to-Voice converts text into a phoneme representation and assigns a phonetic transcription to each word. This can be converted to sound waves as output. The resulting audio resembles comprehensible human voice. Text input is processed through a number of blocks to produce audio output. In the first stage, phrases are decomposed into short words and units are decomposed into phonemes according to their pronunciation.
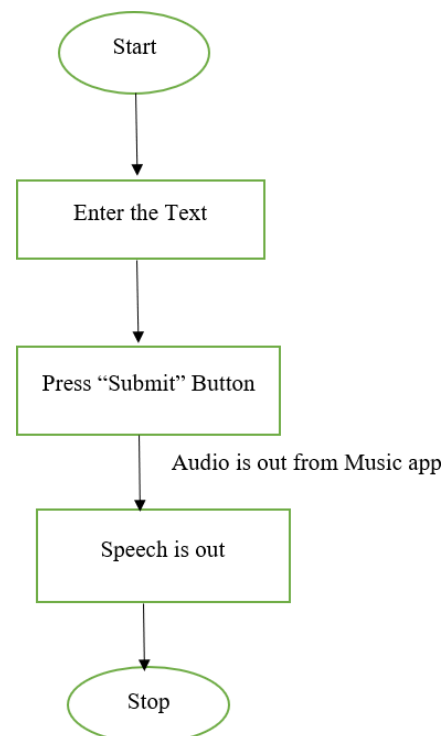

Fig.1. Text message to Voice

Portable Document Format-to-Voice

Portable Document Format to voice and vice versa is useful if you want to hear something rather than read it.

There is a screen with an option to upload files at the end of the run, but it only accepts Portable Document Format oriented documents. Here you have to click on the input file choice and select the appropriate Portable Document Format file and upload it. Click the Send button and the Portable Document Format file will be converted to an audio file. If the selected document is not in Portable Document Format, you will receive the warning message "Please select only PORTABLE DOCUMENT FORMAT file". There is also a quit button.
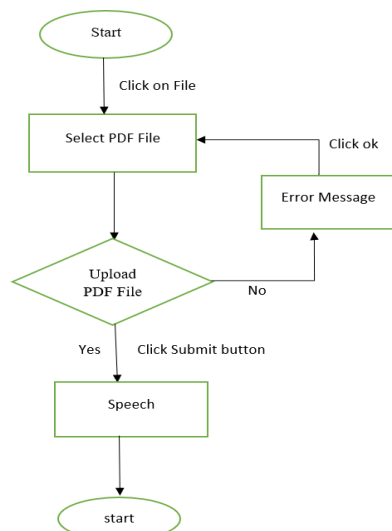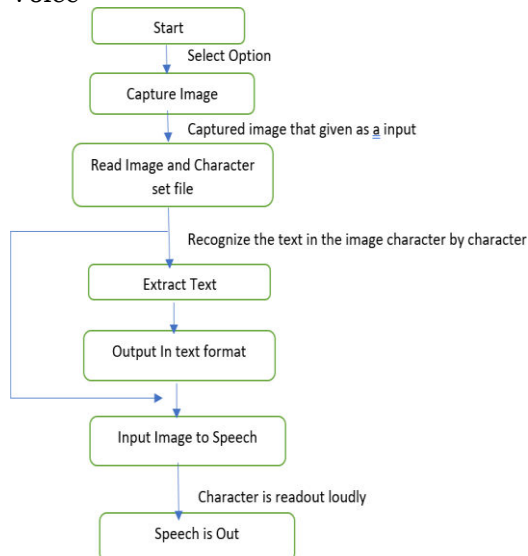

Fig. 2. Portable Document Format to Voice

Fig. 3. Image to Text and Voice

Image to Voice

Our aim was to convert a text-image into the text-string, put it in a file, & use audio for hearing what was written there. For this we need to import a lot of libraries. You can choose from three methods to complete the run. The first process is image-to-voice, which identifies the characters in the photo and converts them to text. Another process, known as image-to-voice conversion, identifies the words in the image and converts them to voice. Text-to-voice conversion is the third action, and termination is the fourth action. As a result, it runs a process and outputs results based on the user's selections.

Pytesseract is a variant of Python tesseract. This is an optical character recognition program built in Python and supported by Google. pyttsx3: it provides A text to speech inbuilt library which helps to work on all platforms & can supports offline. PIL, Enhances the imaging capabilities of the Python interpreter. Google Translate This Python package using the Google Translate API is free. Text can be translated into any language. Hindi, Japanese and Russian are just a few examples. The only requirement is that Google Trans speaks the target language. Further, the module pyttsx3 helps to speak the terminology it understands.

Voice to Text

It uses Google's voice recognition. One of the most user-friendly voice recognition systems is Google. Gets the output of the command will be displayed on the screen as "Speak Now" at the end of the module execution. Then you have to enter the voice command. It will be converted into a text format. In the absence of a spoken command, the message "An unknown error has occurred" is displayed.
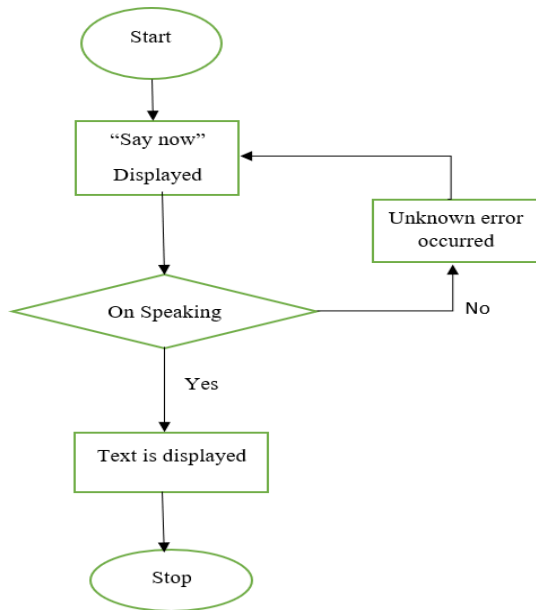
International Journal for Innovative Engineering and Management Research
A Peer Reviewed Open Access International Journal
www.ijiemr.org

Fig. 4. Voice to Text

Work Flow:

Two programs can communicate with each other through a programming interface. Ultimately, the API acts as a messenger, forwarding requests to the appropriate sender and returning results as responses. At most cases, this process involves creating human understandable messages using standard natural language processing techniques. This can be viewed as generating extraction test results for ongoing mixed media recording recovery tasks, such as programmatic descriptions and extraction of material from photographs, audio, and video.
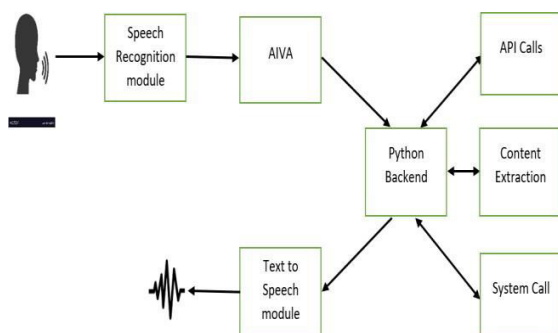

Fig.5.WorkFlow

Results

These form the modules of the system. Upload text from the Text-to-Voice module to the following websites: After clicking the send button, you will receive the entered content in audio format. You can upload any Portable Document Format file using the File button in the Portable Document Format to Voice module. If you

select Send, the output will be in mp3 format. The picture-to-Text module has four operations: picture-to-Voice, Image-to-simple Text, simple Text-to-Voice & End. Choosing any of these given options provides text or audio output.

Text to Voice

Fig. 6. Text to Voice

Portable Document Format to Voice

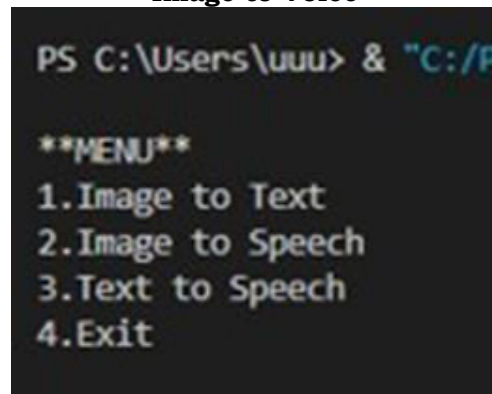Fig. 7. Portable Document Format to Voice

Image to Voice

Fig. 8. Image-to-Voice
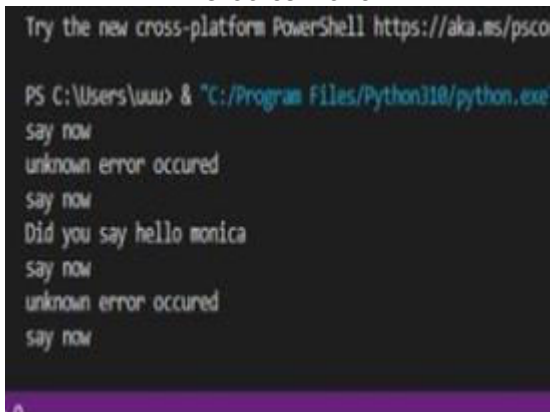
Voice to Text

Fig. 9. Voice to Text

Further Research

Using our current techniques, we claim that it is very useful for people with physical disabilities, such as blind and deaf people, and people who are illiterate, mute, or who can perform many tasks at the same time. increase. We currently offer 4 modules: Image to Audio, Simple-Text to Audio, Portable Document Format to Audio, and Audio-to-Text. Both reading and listening to books benefit. We are also planning to improve this undertaking by including some additional elements that will help upcoming future generations to learn. This format conversion has a more generational impact, as you can see in the example of Alexa and other modern technologies.

Further, we plan to update the image to Voice module to read image descriptions in addition to text and add a few more features.

A component such as text in an audio file. We will build a virtual library system in the future to achieve our goals. We plan to develop more modules that allow customers to upload all kinds of books & get audio files as output. This will help you acquire knowledge as quickly as possible.

Conclusion

We have developed four modules so far for our automatic transcriptor for intellectually disabled persons to improve fluency for people with physical disabilities. Each module has characteristic properties. The text-to-voice feature helps you read typed content aloud, spell words accurately, and deafen.

The Voice-to-Text engine helps convert spoken words into written words. Deaf people can benefit and readers can read better. This Image to Voice module converts both plain text and text in images to voice. Full Portable Document Format files are converted to voice using Portable Document Format to Voice. This benefits less skilled multitaskers as well as those with disabilities. The proposed solution can be utilised as a multilingual programme in the next days after incorporating a few extra modules and functionalities, allowing users to easily use of software in the local tongue. Due to this, it extracts text from images, so you can listen instead of reading, and speak instead of writing. In daily life, this system is very useful.

References

[1] Library Audiobook System Using Speech Recognition. Nikhat Parveen, Priyanka CH.Ruchitha Y.Geeteeka Y.Varni Priya

[2] Speech Recognition using Android. Bhushan Mokal, Sahil Patil, Aniket Kale, Prof. Archana Arudkar in 2020 https://www.irjet.net/archives/V7/i2/IRJETV7I2628.pdf

[3] Ayushi Trivedi, Navya Pant, Pinal Shah, Simran Sonik and Supriya Agrawal Department of Computer Science,NMIMS University, Mumbai, India. Corresponding Author: Navya Pant. Speech to text and text to speech recognition systems-A review Artificialintelligence(AI),sometimes called machine intelligence. https://www.iosrjournals.org/iosrjce/papers/Vol20-issue2/Version-1/E2002013643.pdf.

[4] Huang,J.,Zhou,M.andYang,D.,2007,January.Extracting Chatbo Knowledge from Online Discussion Forums.In IJCAI(Vol-7,pp.423-428

[5] CMUSphnix Basic concepts of speech - Speech Recognition process". http://cmusphinx.sourceforge.netlwiki/tutorialconcepts Hasan U. Zaman, Saif Mahmood, Sadat Hossain, Iftekharul Islam Shovon, Python Based Portable Virtual Text Reader

[6] Dept. Electrical and Computer Engineering, North South University, Dhaka, Bangladesh

[7] B.Marr,The Amazing Ways Google Uses Deep Learning AI. CortanaI ntelligence.

Google Assistant, AppleSiri

[8] Varish, N., Parveen, N, et.al, Image Retrieval Scheme Using Quantized Bins of Color Image Components and Adaptive Tetrolet Transform, IEEE Access 2020, 8, pp. 117639- 117665, 9121956

[9] Parveen, N., Roy, A., Sai Sandesh, D., Sai Srinivasulu, J.Y.P.R., Srikanth, N., Human computer interaction through hand gesture recognition technology, International Journal of Scientific and Technology Research, 2020, 9(4), pp. 505-513

[10] Gayathri .S , Porkodi Venkatesh , Pushpapriya Prem kumar on Voice Assistant for Visually Impaired in 2019, https://ijesc.org/upload/40664e91149af 2618afd09aaf1fca8f 8.Voice%20Assistant%20for%20Visual ly%20Impaired%20(2).pdf

[11] Speech Recognition Bhuvan Taneja, Jones C J, Rohan Tanwar HMRITM (GGSIP University), Delhi, India,2021

[12] Fryer, L.K. and Carpenter, R., 2006. Bots as language learning tools. Language Learning &Technology

[13] J Kiran, N Parveen, Holistic Review of Software Testing and Challenges, International Journal Of Innovative Technology and Exploring Engineering (IJITEE) , Volume.8, Issue 7, Page No pp.1506- 1521, June 201.

[14] Sreeram, G., Pradeep, S., Rao, K.S., Raju, B.D., Nikhat, P. Moving ridge neuronal espionage network simulation for reticulum invasion sensing, International Journal of Pervasive Computing and Communications, 2020.

[15] Rayan spring, RyujiTabuchi, Assessing the practicality of using an automatic speech recognition tool to teach English pronunciation online,2021.

[16] Dash, Devidutta, Arun Agarwal, Kabita Agarwal, and Gourav Misra. "Post Catastrophe Fallouts and Challenges to Swim to Safety." Journal of Information Technology 3, no. 01 (2021): 12-17.