Image-To-Text-To-Speech Convertor App

Omprakash Yadav¹, Swarali Gaonkar², Vailantina Fernandes³

¹*Professor, Computer Engineering ,Xavier Institute Of Engineering ,Mumbai, Maharashtra.*
²*Student,Computer Engineering ,Xavier Institute Of Engineering ,Mumbai, Maharashtra.*
³*Student ,Computer Engineering ,Xavier Institute Of Engineering ,Mumbai, Maharashtra.*

---***---

Abstract -Now a days the digital storage is most popular over paper storage. The information are scanned and hold on in style of image files. To retrieve a picture from massive information, text recognition is finished. the information in this image are often in any language and conjointly written. Image process is finished to extract text and people texts are reborn to audio format so as to avoid ambiguity in written information files because the handwriting of an individual is troublesome to know. There are few automatic ways in machine learning algorithms that didn't offer correct results. during this pre-processing the input image victimization Long STM in repeated Neural Network (RNN), a deep learning algorithmic rule is finished with addiction to it, Optical Character Recognition (OCR) uses OTSU's methodology for image binarization and segmentation then converts texts into audio format with higher accuracy and clarity. But in this project app we have use Cordova to build the system.

Key Words:Camera plugin , Mobile OCR plugin , Text to speech plugin.

1.INTRODUCTION

This document shows the suggested format and appearance of a manuscript prepared for SPIE journals. Accepted papers will be professionally typeset. This template is intended to be a tool to improve manuscript clarity for the reviewers. The final layout of the typeset paper will not match this template Apache Cordova is open source text file mobile development framework. It permits you to use normal internet technologies - HTML5, CSS3, and JavaScript for cross-platform development. Applications execute at intervals wrappers targeted to every platform, and place confidence in standards-compliant API bindings to access every device's capabilities like sensors, data, network standing, etc.

Use Apache Cordova if you are:

- a mobile developer and need to increase application across quite one platform, while not having to re-implement it with every platform's language and power set.

- a web developer and need to deploy an online app that is packaged for distribution in varied app store portals.

- a mobile developer inquisitive about compounding native application elements with a WebView (special browser window) which will access device-level Apis, or if you would like to develop a plugin interface between native and WebView elements layout.

Speech recognition is computer generated simulation of personage speech. it's accustomed translate written info into aural info wherever it's a lot of convenient, particularly for mobile applications like voice-enabled e-mail and unified electronic communication. As a developer that uses Javascript to form hybrid applications, we tend to assume that you simply apprehend the Speech Synthesis API obtainable in most of the online browsers. The usage of this API is basically simple, but it is not obtainable at intervals a cordova application, thus you wish to resort to the native API of the device through a Cordova plugin.

The ensuing applications ar hybrid, which means that they're neither actually native mobile application nor strictly Web-based. they're not native as a result of all layout rendering is completed via internet views rather than the platform's native UI framework. they're not internet apps as a result of they're prepackaged as apps for distribution and have access to native device arthropod genus. compounding native and hybrid code snippets has been attainable.

Plugins area unit associate degree integral a part of the Cordova system. they supply associate degree interface for Cordova and native parts to speak with one another and bindings to straightforward device.Apache Cordova project maintains a group of plugins known as the Core Plugins. These core plugins give your approach to ingress device accomplishment like battery, camera, contacts, etc.

In addition to the core plugins, there area unit many third-party plugins which offer extra bindings to options not essentially offered on all platforms. you'll be able to seek for Cordova plugins exploitation plugin search or npm. you'll be able to additionally develop your own plugins, as delineated within the Plugin Development Guide. Plugins could also be necessary, as an example, to speak between Cordova and custom native parts.

NOTE: after you produce a Cordova project it doesn't have any plugins gift. this can be the new default behavior. Any plugins you need, even the core plugins, should be expressly more.

Cordova doesn't give any UI widgets or MV* frameworks. Cordova provides solely the runtime within which those will execute. If you want to use UI widgets associate degreed/or an MV* framework, you may got to choose those and embody them in your application.

Installing Cordova:Cordova command-line runs on Node.js and is offered on NPM. Follow platform specific guides to put in further platform dependencies. Open a prompt or Terminal, and sort

npm install -g cordova.

Create a project :Create a blank Cordova project victimization the command-line tool. Navigate to the directory wherever

you would like produce|to make|to form} your project and kind cordovacreate .

For a whole set of choices, sort cordova facilitate produce.

Fig : Cordova installation and project creation
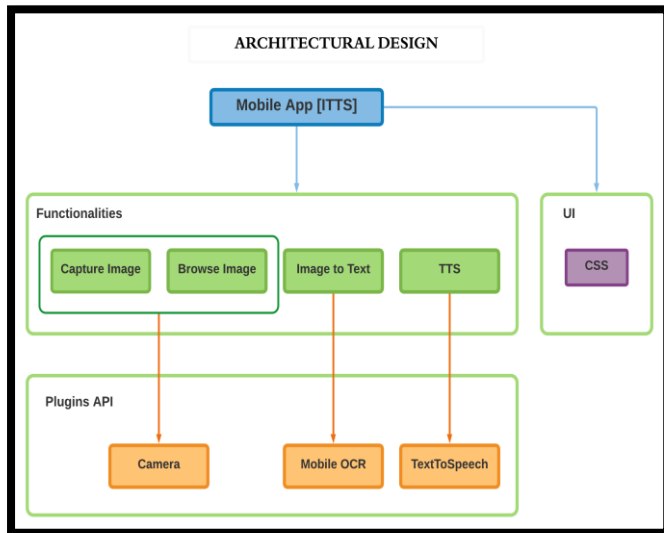
2. WORKING OF METHODOLOGY

Fig : Architectural design

Camera Plugin :

This plugin defines a worldwide navigator.camera object, that provides associate API for taking photos and for selecting pictures from the system's image library.Although the article is hooked up to the worldwide scoped navigator, it's not on the market till when the deviceready event.

Mobile OCR Plugins :

This plugin defines a world textocr object, that provides associate degree methodology that accepts image uri or base64 inputs. If some text was detected within the image, this text are came back as a string. The imageuri or base64 are often send to the plugin exploitation any another plugin like cordova-plugin-camera or cordova-plugin-document-scanner. though the thing is connected to the worldwide scoped window, it's not offered till when the deviceready event.

Text to speech Plugin :

The onfulfilled callback will be called when the speech finishes, and the onrejected callback (Windows Phone only) will be called when an error occurs.

Fig : Adding Plugins to Project

Build android app :

Fig : platform add android

Fig : Build app

Command to build the android platform :

- cordova platform add android

to install all android plugins

Command to build android app :

- cordova build

this will enable the app to run in mobile phone
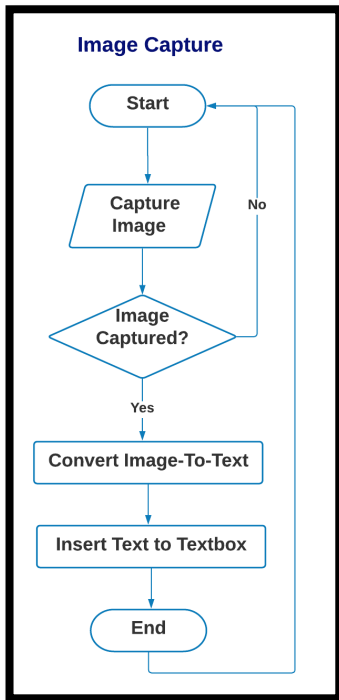
3. FLOWCHART OF THE APP

Fig : Flowchart of image capture

If we select camera icon the image capture function will be called firstly image is be captured it will ask the user to check the capture image if user want to change the image then user can again capture the image else it will move to next stage. The captured image will be converted into text and it will get stored into the textbox provided on the layout of the design.

Fig : Flowchart of browse image

If we select gallery icon then browse image will be called firstly it will redirect the user to gallery then it will allow the user to browse the image from the gallery if image is selected then it will convert the image to text form and paste it in the textbox to the app layout. If user has not selected the image then it will give the alert that image is not selectedand display the outcome in the textbox of the design.
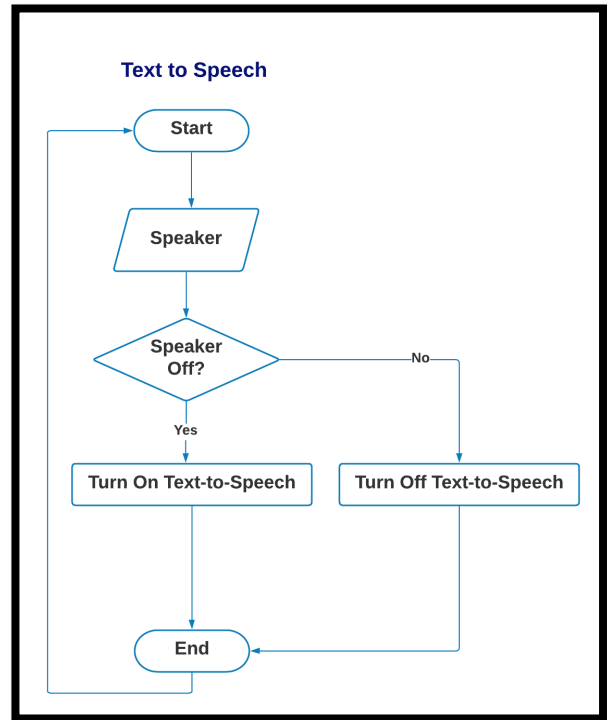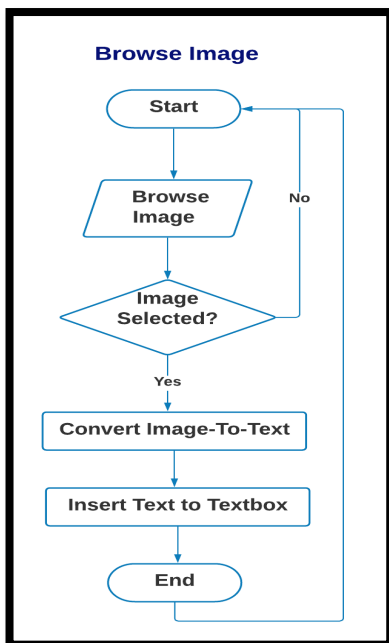

Fig :Flowchart of text to speech

If user select mic icon then speaker will be called to read the text present in the textbox of the layout. Initially if user click on mic icon then default mic will be turn on and start reading the text after turning on mic is user what to turn off the mic or to stop the audio user can again click on mic icon to turn off the mic.
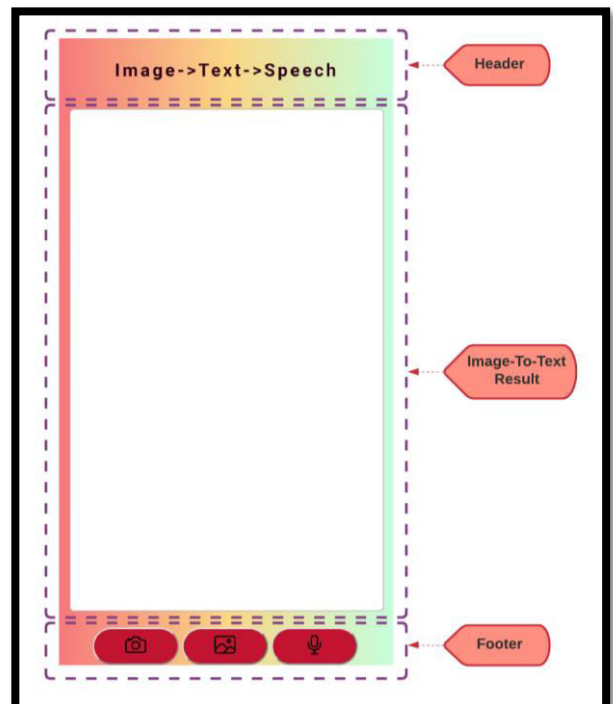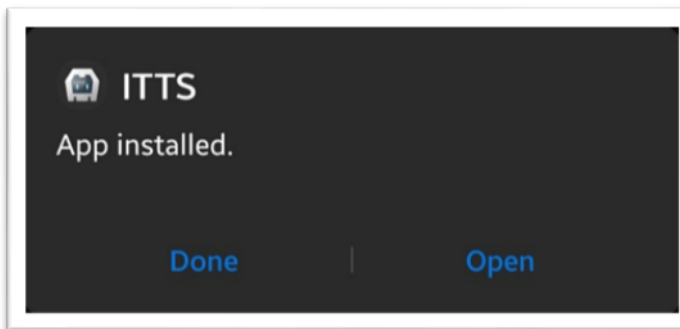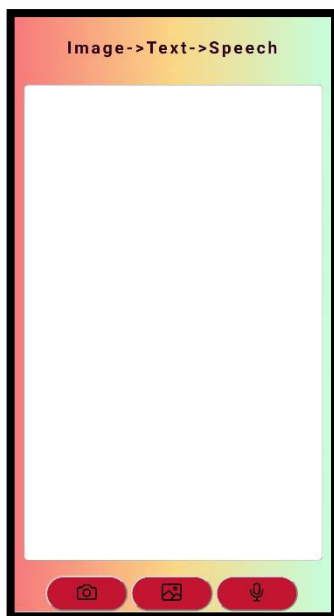
Layout of the App :
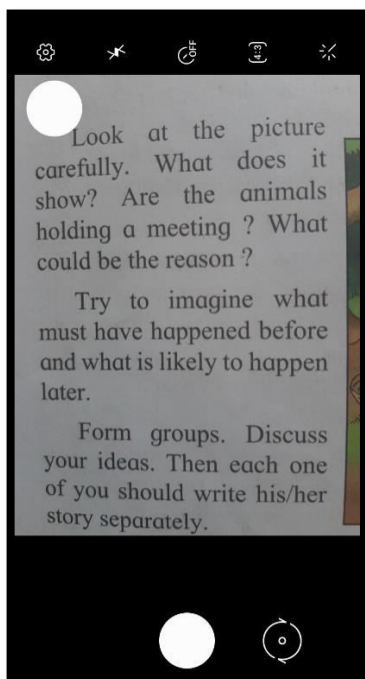

Fig : Interface Designing of the App
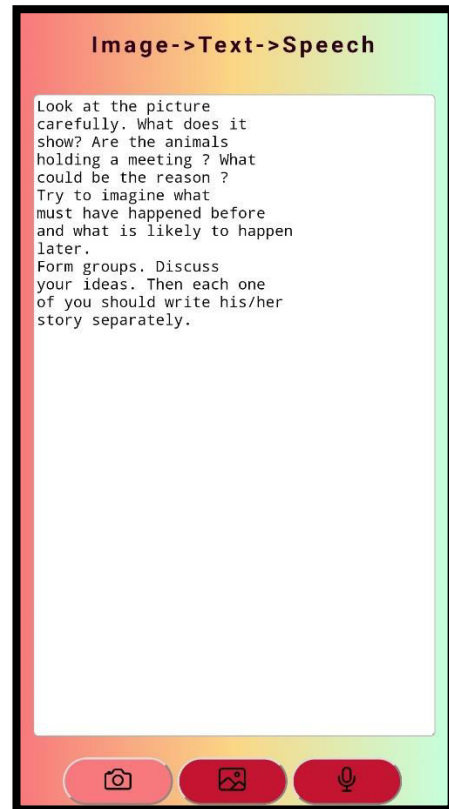
4. RESULT
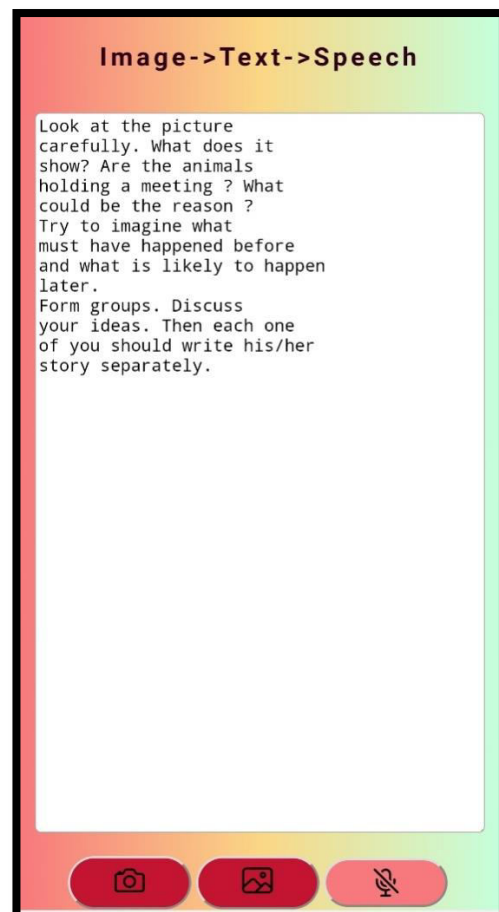
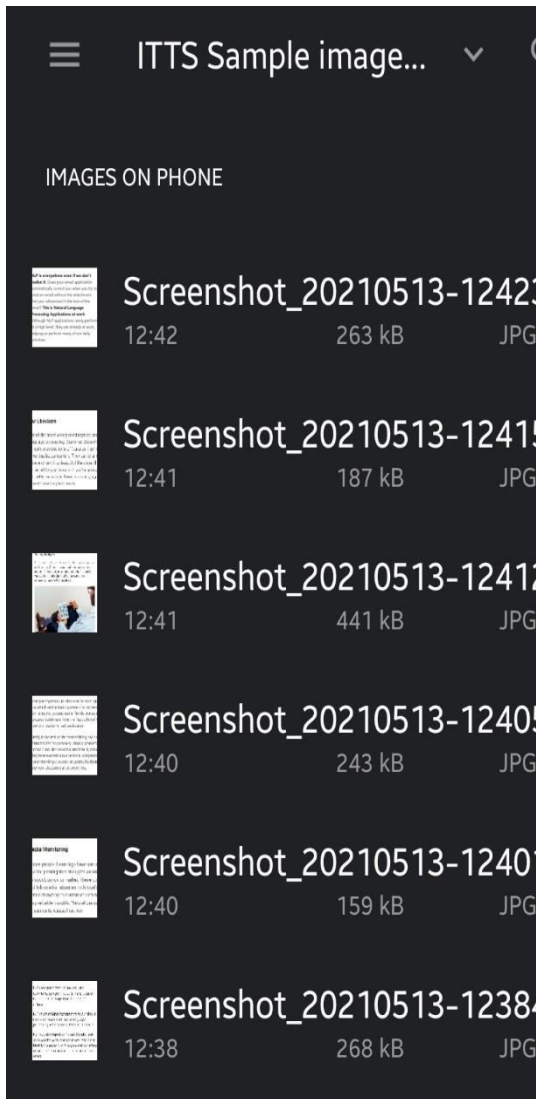- App installation on phone

- Main Screen

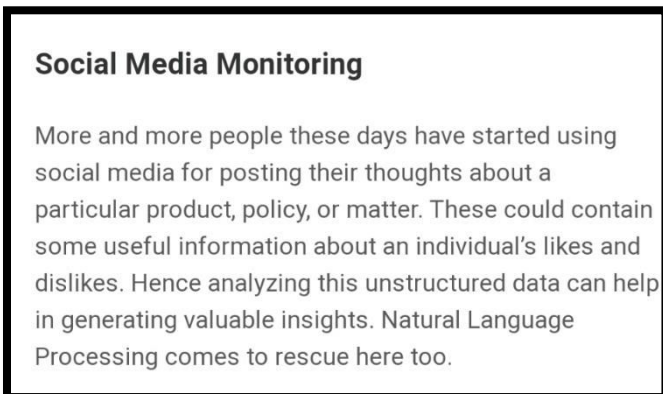- Capturing the image:

- Extracting Text from captured image

- Speaking the text from captured image

- Browsing the image :

- Extracting Text from browsed image

- Selecting the image from gallery

- Speaking the text frombrowsed image

- Turning off the mic Stopping TTS

3. CONCLUSIONS

Android currently talks, and then will your apps. bear in mind that so as for synthesized speech to be intelligible, you would like to match the language you decide on to it of the text to synthesize. Text-to-speech will assist you push your app in new directions. whether or not you utilize TTS to assist users with disabilities, to modify the utilization of your application whereas trying removed from the screen, or just to create it cool, we tend to hope you will get pleasure from this new feature. Text to speech synthesis may be a apace growing side of computer technology and is more and {more} enjoying a more important role within the manner we tend to move with the system and interfaces across a range of platforms. we've known the various operations and processes concerned in text to speech synthesis. we've additionally developed a awfully straightforward and attractive graphical programme that permits the user to type in his/her text provided within the text field within the application.

In future, we tend to arrange to create efforts to create engines for localized Nigerian language thus on create text to speech technology additional accessible to a wider vary of Nigerians. Another space of additional work is that the implementation of a text to speech system on other platforms, like telecommunication systems, ATM machines, video games and the other platforms wherever text to speech technology would be another advantage and increase functionality.

4. ACKNOWLEDGEMENT

5. REFERENCES

1. Lemmetty, S., 1999. Review of Speech Syn1thesis Technology. Masters Dissertation, Helsinki University Of Technology.
2. Dutoit, T., 1993. High quality text-to-speech synthesis of the French language. Doctoral dissertation, Faculte Polytechnique de Mons.
3. Suendermann, D., Höge, H., and Black, A., 2010. Challenges in Speech Synthesis. Chen, F., Jokinen, K., (eds.), Speech Technology, Springer Science + Business Media LLC.
4. Allen, J., Hunnicutt, M. S., Klatt D., 1987. From Text to Speech: The MITalk system. Cambridge University Press.
5. Rubin, P., Baer, T., and Mermelstein, P., 1981. An articulatory synthesizer for perceptual research. Journal of the Acoustical Society of America 70: 321–328.
6. Santen, J.P.H., Sproat, R. W., Olive, J.P., and Hirschberg, J., 1997. Progress in Speech Synthesis. Springer.
7. van Santen, J.P.H., 1994. Assignment of segmental duration in text-to-speech synthesis. Computer Speech & Language, Volume 8, Issue 2, Pages 95–128
8. Wasala, A., Weerasinghe R. , and Gamage, K., 2006, Sinhala Grapheme-to-Phoneme Conversion and Rules for Schwaepenthesis. Proceedings of the COLING/ACL 2006 Main Conference Poster Sessions, Sydney, Australia, pp. 890-897.
9. Lamel, L.F., Gauvain, J.L., Prouts, B., Bouhier, C., and Boesch, R., 1993. Generation and Synthesis of Broadcast Messages, Proceedings ESCA-NATO Workshop and Applications of Speech Technology.
10. van Truc, T., Le Quang, P., van Thuyen, V., Hieu, L.T., Tuan, N.M., and Hung P.D., 2013. Vietnamese Synthesis System, Capstone Project Document, FPT UNIVERSITY.
11. Black, A.W., 2002. Perfect synthesis for all of the people all of the time. IEEE TTS Workshop.
12. Kominek, J., and Black, A.W., 2003. CMU ARCTIC databases for speech synthesis. CMU-LTI-03-177. Language Technologies Institute, School of Computer Science, Carnegie Mellon University.

13. Zhang, J., 2004. Language Generation and Speech Synthesis in Dialogues for Language Learning. Masters Dissertation, Massachusetts Institute of Technology.

14. Dutoit, T., Pagel, V., Pierret, N., Bataille, F., van der Vrecken, O., 1996. The MBROLA Project: Towards a set of high quality speech synthesizers of use for noncommercial purposes. ICSLP Proceedings.

15. Text-to-speech (TTS) Overview. In Voice RSS Website. Retrieved February 21, 2014, from http://www.voicerss.org/tts/

16. Text-to-speech technology: In Linguatec Language Technology Website. Retrieved February 21, 2014, fromhttp://www.linguatec.net/products/tts/information/technology

17. Ngugi, K., Okelo-Odongo, W., and Wagacha, P. W., 2005. Swahili Text-To-Speech System. African Journal of Science and Technology (AJST) Science and Engineering Series Vol. 6, No. 1, pp. 80 – 89.

18. Mohanan, S., Salkar, S., Naik, G., Dessai, N.B., and Naik, S., 2012. Text To Speech Synthesizer for Konkani Language. International Conference on Computing and Control Engineering (ICCCE 2012), 12 & 13 April, ISBN 978-1-4675-2248-9.

19. Swathi, G., Mai, C. K., and Babu, B. R., 2013. Speech Synthesis System for Telugu Language. International Journal of Computer Applications (0975 – 8887), Volume 81 – No5.

20. Dutoit, T., 1997. High-Quality Text-to-Speech Synthesis:An Overview. Journal Of Electrical And Electronics Engineering Australia 17, 25-36.