**PROJECT REPORT**



# **GROUP: 09**

# **PROJECT NAME: TEXT RECOGNITION AND SPEECH FROM PICTURE**

# **COURSE: CSE 323**

# **SECTION: 06**

# **SUBMITTED TO: RASHED MAZUMDER Sir**

# **SUBMITTED BY:**

|  |  |
| --- | --- |
| **NAME** | **ID** |
| **MOHAMMAD MOSHFIQUE UDDIN** | **1812082042** |
| **ROHIT ROY** | **1811270042** |
| **MD. RAUFUL ISLAM RONY** | **1522041642** |

# **INDEX**

|  |  |
| --- | --- |
| **TOPIC** | **PAGE NUMBER** |
| **INTRODUCTION** | **3** |
| **OVERALL DESCRIPTION** | **4** |
| **SYSTEM FEATURES** | **5** |
| **NON-FUNCTIONAL REQUIREMENTS** | **6** |
| **DIAGRAM** | **7** |

# **INTRODUCTION**

* In this project we have created an android application named “Text Recognition and Speech from Picture”.
* We basically created a text converter into speech app following detecting text from images.
* Text recognition is the process of detecting text in images and recognizing the text contained therein.
* Once detected, the recognizer then determines the actual text in each block and segments it into lines and words. The Text API detects text in real time, on device.
* After detecting the text, the recognizer then converts the text into speech.
* It is also called Optical Character Recognitions (OCR).
* The method in which images of typed, hand written or printed text converted into machine encoded text through electronic conversion is Optical Character Recognition (OCR).

**PURPOSES, SERVICES:**

* This app will detect text from Images.
* After detecting the text, it will display the text.
* Then it converts the text into speech.
* Our main goal is to detect text from any kind of images.
* Some times we need some text to type which is in an image so using this app we can easily capture the text from that image.

# **OVERALL DESCRIPTION**

* We use Text Recognizer API from the google vision.
* To use this API the required libraries are implemented in the Gradle (App) file.
* Once implemented in the Gradle file, it makes the required APIs available in the Java code.
* In layout, it creates a separate constraint layout for surface view.
* This surface view is used to project the camera focus area.
* In the end, in this App, it also converts the read text into Speech using the Text To Speech API.
* When we open the app and press the “start capture” button then the camera starts and start detecting the text from the captured image.
* After detecting it shows the text and at last using the Text To Speech API, it converts the

text into speech.

**OPERATING ENVIRONMENT AND ATMOSPHERE:**

**Build IN:** Android Studio

**ANDROID VERSION: Android Jelly Bean 4.1**

**MINIMUM SDK: 16**

**OPERATING SYSTEM: Android**

# **SYSTEM FEATURES**

We used to google vision API which is available from Play services to implement the Optical Character Recognition in our Android App. Behind the scenes, It uses Text Recognizer API from the google vision. To use this API the required libraries are implemented in the Gradle (App) file. Once implemented in the Gradle file, it makes the required APIs available in the Java code.

The main function of our code is textRecognizer(). It opens the camera as a preview in 1080 x 1024 resolution for clear image quality to recognize the text. This is previewed in an android widget called SurfacePreview. After we set up the initial settings for camera and view, we do the actual work in textRecognizer.setProcessor(new Detector. Processor<TextBlock>() When we call this function, we process by detecting the text from the image captured by the camera, the texts are then detected and processed in arrays of strings. Since this is a lot of processing and heavy load, we run a java thread to handle the process where we call the function resultObtained(). This function takes those strings and with the help of the API mentioned earlier, takes the text and converts to speech in the API function textToSpeech.speak(stringResult, TextToSpeech.QUEUE\_FLUSH, null, null);

**PRONUNCIATION OF TEXT**

**Start Capturing**

**TEXT RECOGNIZER API**

**TEXT DETECTED**

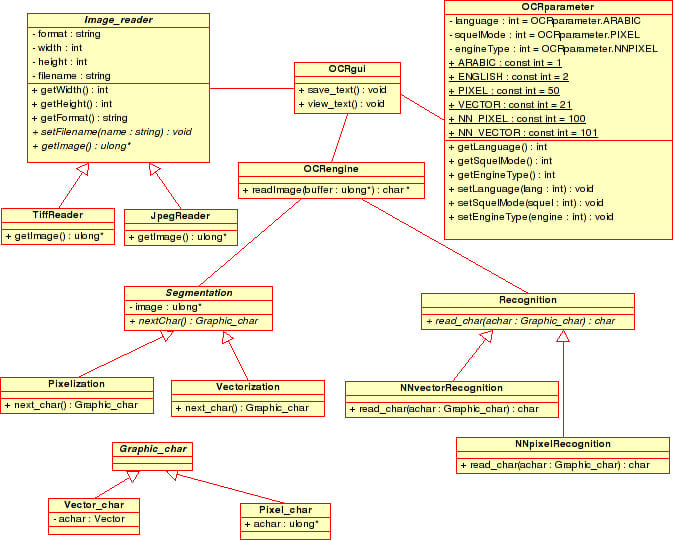
**TEXT TO SPEECH API**

# NON-FUNCTIONAL REQUIREMENTS

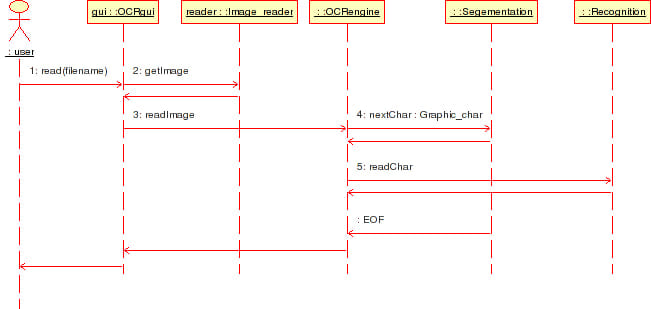
**PERFORMANCE:**

* **This app is useful for those who have optical issues.**
* **By using this app, they can hear what text is written.**
* **Anyone can easily capture text from images.**
* **This app is flexible and easy to use.**
* **We can hear the sound by using headphones also.**
* **This is an amateur level app so it has many limitations.**
* **It is hard to detect a line containing more than 3/4 words for this app.**
* **It only can detect English words.**
* **It needs clear picture for detecting text.**
* **This app can only detect real time images text.**

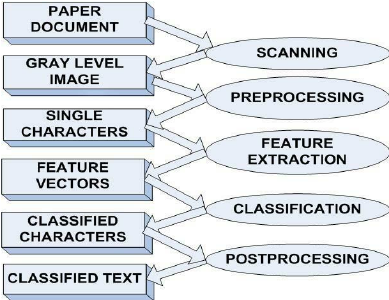
**CLASS DIAGRAM:**



**SEQUENCE DIAGRAM:**



**Feature Implementation:**



**Training and Testing:**

