MAT JOURNALS
Journal of Optical Communication Electronics
Volume 4 Issue 3

Camlens – an Innovative Android Phone Application to Empower the Blind and Visually Impaired in Reading any Kind of Printed Text in Real-Time using Opencv, Optical Character Recognition and Text-To-Speech

[1]Mr. Sahil Sachdeva, [2]Ms. Akshita Sachdeva, [3]Mr. Yash Bakshi, [4]Prof. Manoj Kumar
[1,3]Student, [4]Professor
[1,3,4]Department of Computer Science & Technology, Manav Rachna University, Faridabad, Haryana, India
[2]Social Innovator & Software Developer, Digital Impact Square, Maharashtra, India
Email: *[1]ss.sahil.sachdeva@gmail.com,[2] akshita.sachdeva1994@yahoo.com, [3]yashbakshi2015@gmail.com, [4]manoj@mru.edu.in*
DOI: https://doi.org/10.5281/zenodo.1451742

Abstract

CamLens is an innovative mobile phone application with the vision to empower the blind and visually impaired people in reading any kind of printed text. CamLens intelligently scans the frame inside the camera view to automatically select the exact edges of the document containing the text in real-time; crop the file from the four selected edges and alternate the perspective transformation of the cropped photographto obtain an image containing the scanned document. The application then applies Optical Character Recognition (OCR) and (TTS) conversion to extract the text from the image and to convert the extracted text into audio output respectively which may be listened by the blind and visually impaired person the usage of earphones.The genesis of the research comes from the fact that the three edges of a page of the book are easier to find with lesser possibilities of finding the fourth edge precisely which, in turn, hampers the cropping accuracy as well as the OCR and TTS output.

Keywords: *Blind, Visually-Impaired, OpenCV, Optical Character Recognition, Text-To-Speech, Printed Text, Android Phone Application*

INTRODUCTION

285 million people are estimated to be visually impaired worldwide: 39 million are blind and 246 have low vision and 90% of these live in low-income settings. They confront a number of visual challenges everyday – from reading the label on a frozen dinner to figuring out if they're at the right bus stop.

The advancements in technology in terms of Image Processing, Optical Character Recognition and text-to-speech paved way for the development of CamLens, the innovative mobile phone application that allows reading of any kind of printed text with 90-95% accuracy rates.

CamLens is an innovative mobile phone application with the vision to empower the blind and visually impaired people in reading any kind of printed text.

CamLens intelligently scans the body inside the camera view to robotically select the exact edges of the report containing the text in actual-time; crop the file as a result from the 4 selected edges and trade the attitude transformation of the cropped picture to attain an image containing document. The application then applies Optical Character Recognition (OCR) and text-to-speech (TTS) conversion to extract the textual content from the photo and to transform the extracted text into audio output

respectively which can be listened through the blind and visually impaired consumer the usage of earphones.The genesis of the research comes from the fact that the three edges of a page of the book are easier to find with lesser possibilities of finding the fourth edge precisely which, in turn, hampers the cropping accuracy as well as the OCR and TTS output.

OCR is the mechanical or electronic conversion of pix of typed, handwritten or published text into device-encoded textual content, whether from a scanned file, a photo of a record, a scene-image (for instance the text on signs and symptoms and billboards in a panorama image) or from subtitle textual content superimposed on an picture (for example from a television broadcast).broadly used as a form of statistics access from printed paper

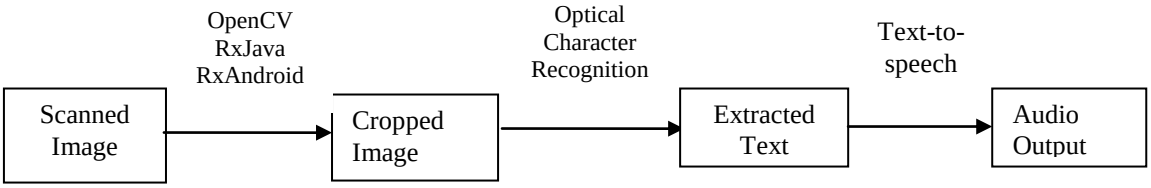


Fig: 1. *Framework*

Text to speech, abbreviated as TTS, is a shape of speech synthesis that converts text into spoken voice output. text to speech structures had been first advanced to useful resource the visually impaired by using providing a computer-generated spoken voice that might "study" text to the person.TTS ought to not be stressed with voice reaction structures. Voice reaction systems are confined to synthesizing sentences that include only phrases which have been predetermined through the gadget. TTS structures, in contrast, are theoretically capable of "reading" any string of text characters to shape authentic sentences.

BACKGROUND

285 million people are estimated to be visually impaired worldwide: 39 million are blind and 246 have low vision and 90% of these live in low-income settings. They confront a number of visual challenges everyday – From reading the label on a frozen dinner to figuring out if they're at the right bus stop. The advancements in technology in terms of Image Processing, Optical Character Recognition and text-to-speech paved way for the development of CamLens, the innovative mobile phone application that allows reading of any kid of printed text with 90-95% accuracy rates.

COMPARATIVE STUDY

Before coming up with CamLens, I performed a detailed scrutiny of the existing technologies that offer text-scanning capabilities such as Microsoft Office Lens & Cam Scanner mobile phone applications. Results of the detailed comparative study are derived after careful scrutiny and involve figuring out the pros & cons of the app based on personal user-experience as well as through blogs & articles comparing Microsoft Office Lens with other Scanning apps. Though simple & easy to use, Microsoft Office Lens couldn't beat the filters & 'control in the hands of user' features of Cam Scanner app. And more importantly, none of them suited the needs of a blind user. So, I decided to offer the existing functionality of Microsoft Office Lens accompanied by the 'filters' & 'control in the hands of user' features of Cam Scanner & started

with the development of 'CamLens' (Cam Scanner + Office Lens).

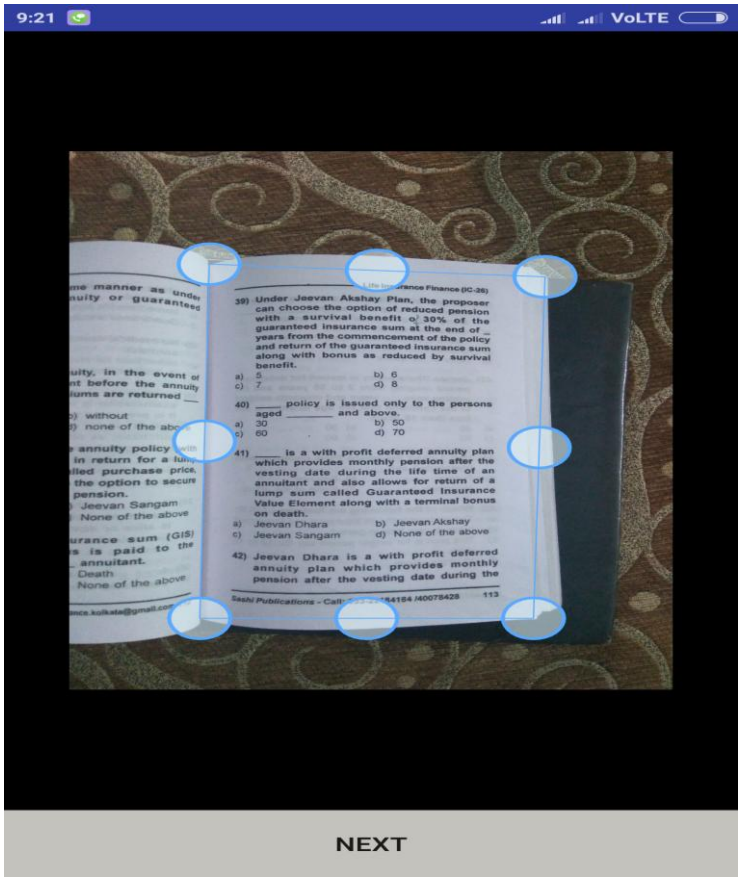
PROPOSED WORK

CamLens is an innovative android mobile phone application that intelligently scans the pages of a book to automatically and accurately detect and crop the 4 edges of the page in real-time using OpenCV, RxJava and RxAndroid libraries for Android. It, further, applies perspective transformation on the scanned image to get an image that acts as the input for the Optical Character Recognition Algorithm.

The text-to-speech algorithm is then applied on the text extracted using Optical Character Recognition to generate an audio output which is fed to the blind and visually impaired user's ears using earphones.

Development Phases:

Application



Detecting Edges & Cropping

Using OpenCV 2.4.11, RxJava 1.0.4 & xAndroid 0.25.0 to intelligently and automatically select the exact edges and to crop the document accordingly from the selected 4 edges and change the respective transformation of the cropped image.

Capturing Image

Using Camera2 API for Android.

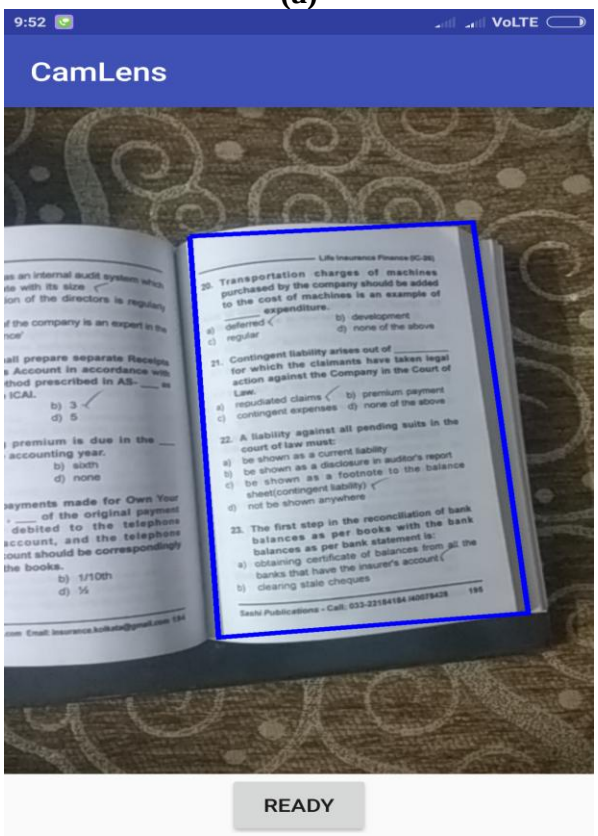
Offering user-control

Using touch-inputs to alter the intelligently and automatically detected 'Region of Interest' of the image on the whim of the user.

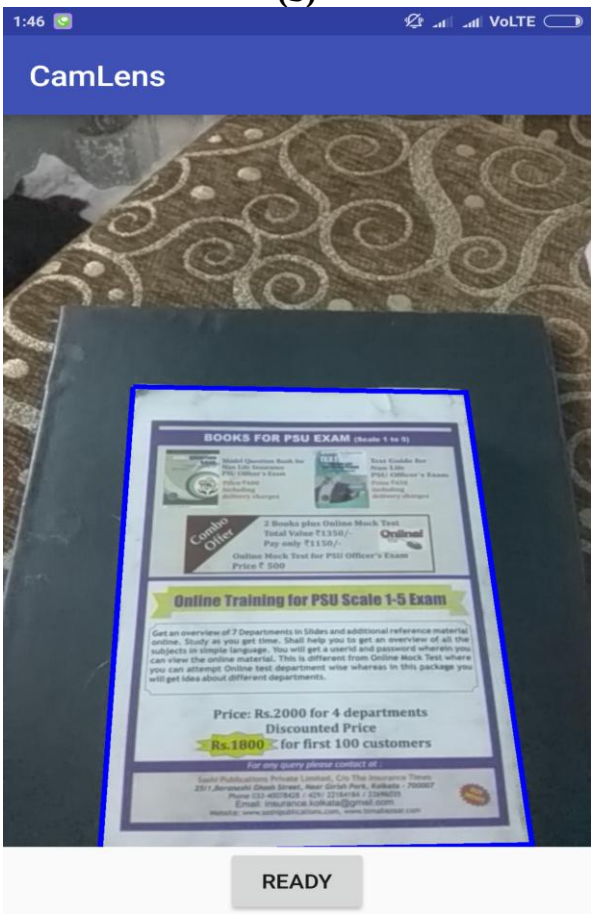
Offering Filters

Offering Magic Color, Gray Mode and Black & White filters for the scanned image.Saving the image to Phone Memory.

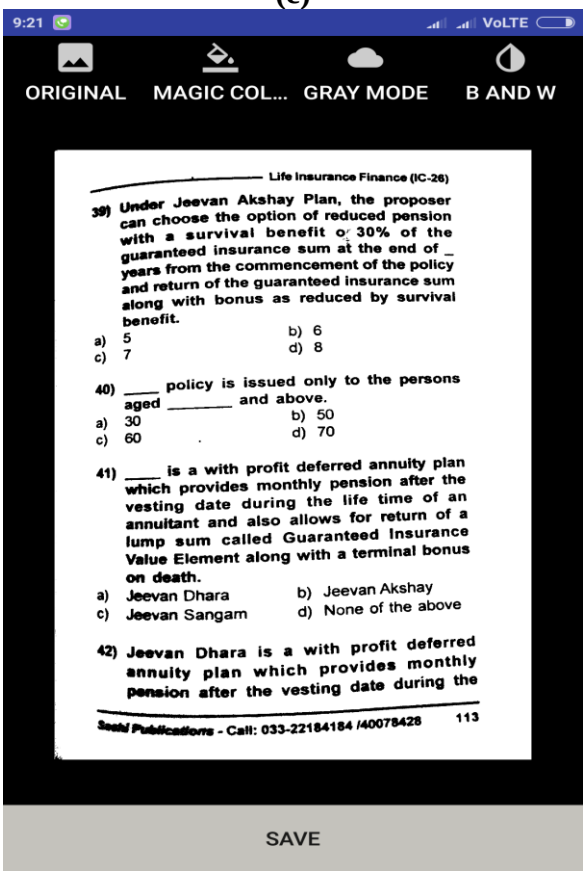
(a)



(b)



(c)



(d)

Fig: 2. *(a),(b),(c),(d). Screenshots of Cam-lens*

Algorithms

```
private Observable<MatData>
detectRect(MatData mataData) {
return Observable.just(mataData)
.concatMap
(OpenCVHelper::getMonochromeMat)
.concatMap(OpenCVHelper::getContours
Mat)
.concatMap(OpenCVHelper::getPath);}
private static <T>
Observable.Transformer<T, T>
mainAsync() {
return obs ->
obs.subscribeOn(Schedulers.newThread())
.observeOn(AndroidSchedulers.mainThrea
d());
}
private static Mat getEdge(Mat oriMat) {
long now = System.currentTimeMillis();
Mat sobelX = new Mat();
Mat sobelY = new Mat();
Mat destination = new Mat(oriMat.rows(),
oriMat.cols(), oriMat.type());
Imgproc.cvtColor(oriMat, destination,
Imgproc.COLOR_RGBA2GRAY);
Imgproc.Sobel(destination, sobelX,
CvType.CV_16S, 1, 0);
Imgproc.Sobel(destination, sobelY,
CvType.CV_16S, 0, 1);
Mat absX = new Mat();
Mat absY = new Mat();
Core.convertScaleAbs(sobelX, absX);
Core.convertScaleAbs(sobelY, absY);
Mat result = new Mat();
Core.addWeighted(absX, 0.5, absY, 0.5, 0,
result);
Log.v(TAG, "getEdge time:" +
(System.currentTimeMillis() - now));
return result;
}
```

RESULT ANALYSIS

In preliminary testing, our algorithm increased the real-time edge detection accuracy to 94.2% and Optical Character Recognition accuracy to 92.3%.

CONCLUSIONS

CamLens is an innovative android mobile phone application that intelligently scans the pages of a book to automatically and accurately detect and crop the 4 edges of the page in real-time using OpenCV, RxJava and RxAndroid libraries for Android. It, further, applies perspective transformation on the scanned image to get an image that acts as the input for the Optical Character Recognition Algorithm. The text-to-speech algorithm is then applied on the text extracted using Optical Character Recognition to generate an audio output which is fed to the blind and visually impaired user's ears using earphones.

REFERENCES

1. https://github.com/ReactiveX/RxJava
2. https://github.com/ReactiveX/RxAndroid
3. https://sourceforge.net/projects/opencvlibrary/files/opencv-android/2.4.11/
4. https://github.com/iyotetsuya/RectangleDetectio

Cite as

Mr. Sahil Sachdeva, Ms. Akshita Sachdeva, Mr. Yash Bakshi, & Prof. Manoj Kumar. (2018). Camlens – an Innovative Android Phone Application to Empower the Blind and Visually Impaired in Reading any Kind of Printed Text in Real-Time using Opencv, Optical Character Recognition and Text-To-Speech. Journal of Optical Communication Electronics, 4(3), 5–10. http://doi.org/10.5281/zenodo.1451742