****

**MINI-PROJECT REPORT ON**

**“Extraction of text from Images +Text To Speech”**

***BY***

NANDINI SAHU

SIDDHI SHEWALE

ANUSHKA GIRISH

ANIKET MORE

***Under the Guidance of:***

**DR. PD PATIL**

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

**MIT SCHOOL OF ENGINEERING**

Loni Kalbhor Pune

**MIT SCHOOL OF ENGINEERING**

DEPARTMENT OF COMPUTER SCIENCE

***CERTIFICATE***



**This is to certify that Mini -Project report entitled**

**“Extraction of text from Images+text to speech”**

**SUBMITTED BY**

NANDINI SAHU MITU21BTCS0344

SIDDHI SHEWALE MITU21BTCS0613

ANUSHKA GIRISH MITU21BTCS0104

ANIKET MORE MITU21BTCS0082

is a record of Bonafide work carried out by them, under my guidance ,in partial fulfilment of the requirement for the Second Year Of Engineering(CS)

at MIT School Of Engineering,Pune Under MIT Art Design & Technology

University.

Date:

**DR. PD Patil DR. RajneeshKaur Sachdeo**

**Guide Dean Engineering**

Department of CSE Head ,Department of CSE

**DECLARATION**

We, the team members

NANDINI SAHU MITU21BTCS0344

SIDDHI SHEWALE MITU21BTCS0613 ANUSHKA GIRISH MITU21BTCS0104

ANIKET MORE MITU21BTCS0082

Hereby declare that the project work incorporated in the present project entitled **“Extraction of text from Images+text to audio** **”** is original work. This work (in part or in full) has not been submitted to any University for the award or a Degree or a Diploma.We have properly acknowledged the material collected from secondary sources wherever required.We solely own the responsibility for the originality of the entire content.

Date: Name & Signature of the Team Members

Nandini Sahu

Siddhi Shewale

Anushka Girish

Aniket More

Name and Signature of Guide

Seal/Stamp of the college

**ACKNOWLEDGEMENT**

We would like to express our deep gratitude to our project guide

**Dr. P D PATIL**  for his guidance with unsurpassed knowledge and immense encouragement. We also thank our project coordinator, Department of Computer Scienc. We are grateful to the Head of the Department, Computer Science and Engineering, for providing us with the required facilities for the completion of the project work.

We are very much thankful to the Principal and Management, for their encouragement and cooperation to carry out this work. We express our thanks to all teaching faculty of Department of CSE, whose suggestions during reviews helped us in accomplishment of our project. We would like to thank all non-teaching staff of the Department of CSE for providing great assistance in accomplishment of our project.

We would like to thank our parents, friends, and classmates for their encouragement throughout our project period. At last but not the least, we thank everyone for supporting us directly or indirectly in completing this project successfully.

**ABSTRACT:**

**PROBLEM STATEMENT: -**

Text is a natural extension of our language. It was an inevitable step we took in our development along with other technologies. I*mage text can be found in captured images, scanned documents, magazines, Newspapers, posters, etc.* Now-a-days, there is a growing demand for the software systems to recognize characters in computer when information is scanned through paper documents as we know that there are number of historical, mythological books and newspapers which are in printed format. Day by day due to atmospheric changes or due to improper handling they get damaged. Therefore, nowadays there is a huge demand in storing the information available in these paper documents into a computer storage disk and then later reusing this information by searching process. Our motive is to Extract the text from the image and then Organize the text to convert into s0 so that humans can easily understand and store the organized data in any form. Extraction of texts can be achieved using Optical Character Recognition (OCR) to analyze text. The algorithm deals with taking scanned copy of a document as an input and extract text from image into text format using OCR, Otsu’s algorithm for segmentation, Noise removal from the image and the final step is to Organize the text by recognizing the pattern of particular text formats. Extraction of text from scanned copies of documents and text images and displaying its related information using various image processing techniques and converting that into useful information for organizations to reduce the work of printed document verification manually.

**METHODOLOGY:**

Text data present in images contain useful information for automatic explanation, indexing, and structuring of images. Extraction of this information involves detection, localization, tracking, extraction, enhancement, and recognition of the text from a given image.

The following proposed method are used here in this project to extract text from the desired images, they are OCR, Otsu’s method. Where in the OCR methodology, the process we inculcate is the conversion of texted or text imposed image to the proper text format when and ever any text representation is found at any region of the input document or an image and also this method goes through it’s pre-processing t echniques which are De-skewing, Despeckling, Binarisation, Line removal, Layout analysis, Line and script recognition and Character Isolation. Where in Otsu method, the process is to perform automatic image thresholding. In the simplest form, the algorithm returns a single intensity threshold that separate pixels into two classes, foreground and background. This threshold is determined by minimizing intra-class intensity variance, or equivalently, by maximizing inter-class variance.

After going through this whole process, the final step in this project is to align the necessary data extracted from the input in a constructed manner which is the most desired and the actual information that will be helpful in hat whole document or an image for which we train the system to get the text only from a specific targeted positions of the input image.

After converting image to text ,the text would be converted into speech.Text to Speech conversion system Methodology:- 1. Speech synthesis techniques will be used in order to get the naturalness quality in the synthetic speech. 2. The process of the English language can be Used as the basic unit for speech synthesis. 3. Speech database for the English language will be developed using phoneme.4. The input text will be separated into English Phoneme. 5. Phonemes will be searched in the database and corresponding phonemes sounds will be Concatenated to generate synthesized output Speech

**Contents**

1. **INTRODUCTION**

1.1 Images and types of images

1.2 Images as matrix

1.3 Image processing

1.4 Image segmentation

1.5 Thresholding

1.6 Edge detection

1.7 Tesseract OCR

1.8 Speech

1. **LITERATURE SURVEY**

2.1 Existing methods

2.2 Alogorith for text extraction from typed image

1. **METHODOLOGY**

3.1 System Architecture

3.2 Proposed Algorithm

3.2.1 Image Preprocessing

3.2.2 OCR

3.3 Organising text

3.4 Conversion to speech

1. **CODE AND OUTPUT**
2. **CONCLUSION**
3. **FUTURE SCOPE AND REFERENCES**

**1. INTRODUCTION**

People deal with chunks of information daily. Much data nowadays is sent in the form of images or scanned documents. Many situations require editing of data sent in the form of images. Thus, we need tools to convert these documents into editable forms. Hence, recognition of texts from printed/typed text images is an important task. The purpose of the paper is to identify methods to convert printed text images into editable documents. The primary objective is to extract characters from printed or typed text images. Nowadays, there are various techniques and algorithms available to do the same. Optical character recognition is one of the popular techniques used for recognizing texts. It is a technique that is used to convert scanned images of printed or handwritten text into character streams that can be read by machines and then can be edited for further functionalities.

A lot of research has already been done in this area. To recognize characters, techniques like Otsu’s method have been used for image segmentation and Hough transform for skew detection. For detection of characters, pattern recognition techniques are used.

In order to improve the accuracy of the proposed algorithm, document image enhancement is necessary.To automatically improve quality of the document images for OCR processing, they used quality measures and image restoration techniques. Small speckle factor, touching character factor, white speckle factor, broken character factor and font size factor were the quality measures used by them. The system took the document image as an input, characterized it using the quality measures. In this paper, an algorithm to convert scanned text images into editable format with maximum efficiency is proposed. The proposed algorithm uses techniques like Otsu’s method and Hough transform with some modification to enhance images, perform skew removal operations on images and extract text lines efficiently to detect and recognize characters. Despite using simplest techniques, the proposed algorithm showed good results against rotation and scaling after some modification. The proposed algorithm deals with English alphabets (a-z, A-Z) and numerals (0-9).

**1.1 IMAGE**

An image is defined as a two-dimensional function, F(x,y), where x and y are spatial coordinates, and the amplitude of at any pair of coordinates (x,y) is called the intensity of that image at that point. When x, y, and amplitude values of F are finite, we call it a digital image. In other words, an image can be defined by a two-dimensional array specifically arranged in rows and columns.

Digital Image is composed of a finite number of elements, each of which elements have a particular value at a particular location.A Pixel is most widely used to denote the elements of a Digital Image.

**TYPES OF IMAGES**

**1) Binary Image:** The binary image as its name suggests, contains only two pixel elements i.e. 0 & 1, where 0 refers to black and 1 refers to white

**2) Black and White Image:** The image which consists of only black and white color is called BLACK AND WHITE IMAGE.

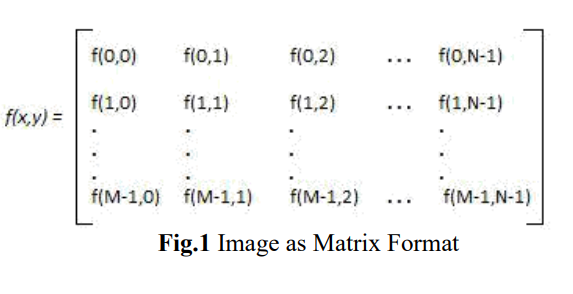
**3) 8 bit Color Format:** It is the most famous image format. It has 256 different shades of colors in it and commonly known as Grayscale Image. In this format, 0 stands for Black, and 255 stands for white, and 127 stands for gray.

**4) 16 bit Color Format:** It is a color image format. It has 65,536 different colors in it. It is also known as High Color Format. In this format the distribution of color is not the same as Grayscale image.

A 16 bit format is actually divided into three further formats which are Red, Green and Blue. That famous RGB format

**1.2 IMAGE AS A MATRIX**

As we know, images are represented in rows and columns we have the following syntax in which images are represented:



The right side of this equation is a digital image by definition. Every element of this matrix is called image element, picture element, or pixel.

**1.3 IMAGE PROCESSING**

Image processing is a method to perform some operations on an image, in order to get an enhanced image and or to extract some useful information from it.

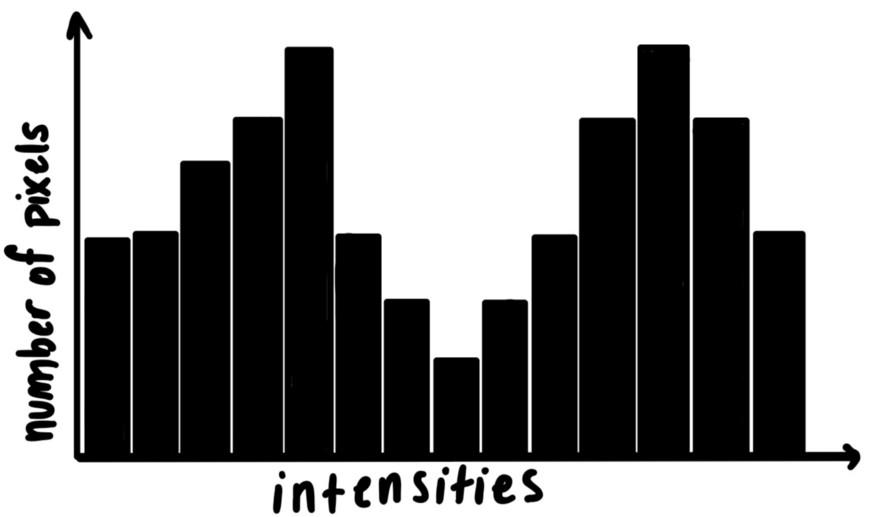
If we talk about the basic definition of image processing then “Image processing is the analysis and manipulation of a digitized image, especially in order to improve its quality”. In another word An image is nothing more than a two-dimensional matrix (3- D in case of coloured images) which is defined by the mathematical function f(x, y) at any point is giving the pixel value at that point of an image, the pixel value describes how bright that pixel is, and what colour it should be.

Image processing is basically signal processing in which input is an image and output is image or characteristics according to requirement associated with that image.

**1.4 Image Segmentation:**

In Image processing, segmentation is often the first step to pre-process images to extract objects of interest for further analysis. Segmentation techniques can be generally categorized into two frameworks, edge-based and region based approaches.The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyse. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label

**1.5 Thresholding — how to generate a binary image**

A binary image is obtained from a grey-scale image by following a process of information abstraction. **Thresholding** is the main techniques used at this stage. The main goal of thresholding is to extract the foreground from the background******

This is a binary image for example ;)

The initial step is to plot a **histogram** of the intensities of a grey-scale image. A threshold intensity is selected. Pixels that fall below this threshold are labelled ‘**white’** or **1** andthe ones that fall above are labelled **black**, or **0**.

In the best case scenario, the threshold that divides black from white will be easily identified by finding the local minima between the two highest peaks.

**1.7Conversion to speech:**

Optical character Recognition (OCR) is a process that converts scanned or printed text images , handwritten 1 text into editable text for further processing. is paper has presented a robust approach for text extraction and converting it to speech. Testing of device was done on raspberry pi platform. Raspy is initially connected to the internet through VLAN. e so ware is installed using command lines. Following steps are to be followed: 1. setup is to download the installation script, 2. Second step is to convert it to executable form and 3. last step starts the script which does the rest of the installation work

**1.8 EDGE DETECTION**

[Edge detection](https://www.mathworks.com/products/image.html) is an image processing technique for finding the boundaries of objects within images. It works by detecting discontinuities in brightness. Edge detection is used for [image segmentation](https://www.mathworks.com/discovery/image-segmentation.html) and data extraction in areas such as image processing, [computer vision](https://www.mathworks.com/discovery/computer-vision.html), and machine vision.

It is a widely used technique in digital image processing like pattern recognition,image morphology ,feature extraction

Edge detection allows users to observe the features of an image for a significant change in the gray level. This texture indicating the end of one region in the image and the beginning of another. It reduces the amount of data in an image and preserves the structural properties of an image.

**1.9 Tesseract OCR**

Tesseract is an open source text recognition (OCR) Engine,It can be used directly, or (for programmers) using an API to extract printed text from images. It supports a wide variety of languages.

Tesseract is compatible with many programming languages. It can be used in conjunction with an external text detector to recognize text from an image of a single text line.  ****

To recognize an image containing a single character, they typically use a Convolutional Neural Network (CNN).

*It is a multi-stage process where the steps are differentiated below:*

1. Word Finding

2. Line Finding

3. Character Classification

Word finding was done by organizing text lines into binary large objects, and the lines and regions are analyzed for fixed pitch or proportional text. Text lines are broken into words differently according to the kind of character spacing. Recognition then proceeds as a two-pass process. In the first pass, an attempt is made to recognize each word in turn.

Each word that is satisfactory is passed to an adaptive classifier as training data. The adaptive classifier then gets a chance to more accurately recognize text lower down the page.

**Advantage of Tesseract OCR**

Tesseract works best when there are extremely clean segmentations of the foreground text from the background.

**Limitations of Tesseract OCR**

1. Doesn't do well with images affected by artifacts including partial occlusion, distorted perspective, and complex background.
2. It is not capable of recognizing handwriting.
3. It is not always good at analyzing the natural reading order of documents. For example, it may fail to recognize that a document contains two columns, and may try to join text across columns.
4. Poor quality scans may produce poor quality OCR.
5. It does not expose information about what font family text belongs to.

***LITERATURE SURVEY***

Text Extraction from image is concerned with extracting the relevant text data from a collection of images. Recent studies in the field of image processing show a great amount of interest in content retrieval from images. This content can be in the form of objects, color, texture, shape as well as the relationships between them. The semantic information provided by an image can be useful for content based image retrieval, as well as for indexing and classification purposes.

Since the text data can be embedded in an image in different font styles, sizes, orientations, colors, and against a complex background, the problem of extracting the candidate text region becomes a challenging one. Current Optical Character Recognition (OCR) techniques can only handle text against a plain monochrome background and cannot extract text from a complex or textured background

. A Text Information Extraction system receives an input in the form of a still image or a sequence of images. The images can be in grayscale or color, compressed or uncompressed, and the text in the images may move or may not. The problem arises due to TIE system can be divided into the following sub-problems:

(i) Detection

(ii) Localization

(iii) Tracking

(iv) Extraction and Enhancement

(v) Recognition (OCR).

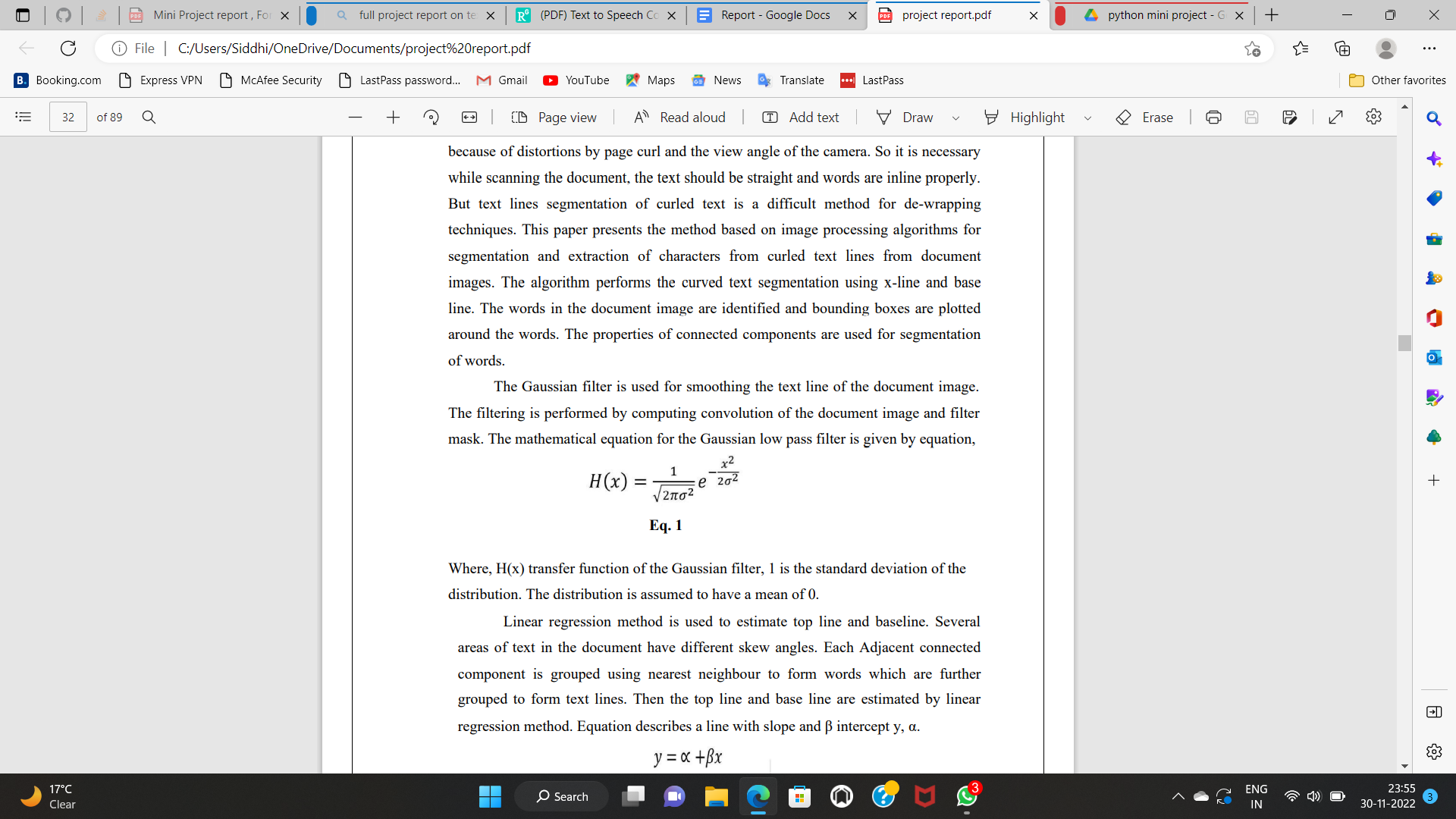
Text detection refers to the determination of the presence of text in a given sequence of images. Text localization is the process of determining the location of text in the image and generating bounding boxes around the text. Text tracking is performed to reduce the processing time for text localization and to maintain the integrity of position across adjacent frames. Although the precise location of text in an image can be indicated by bounding boxes, the text needs to be segmented from the background to facilitate its recognition. That means, the extracted text image has to be converted into a binary image and enhanced before it is fed into an OCR engine. Text extraction is the stage where the text components are segmented from the background. Text Enhancement of the extracted text components is required because the text region 23 usually has low resolution and is horizontal to noise. Thereafter, the extracted text images can be transformed into plain text using OCR technology.

**2.1 Existing Methods for Generating Text Extraction**

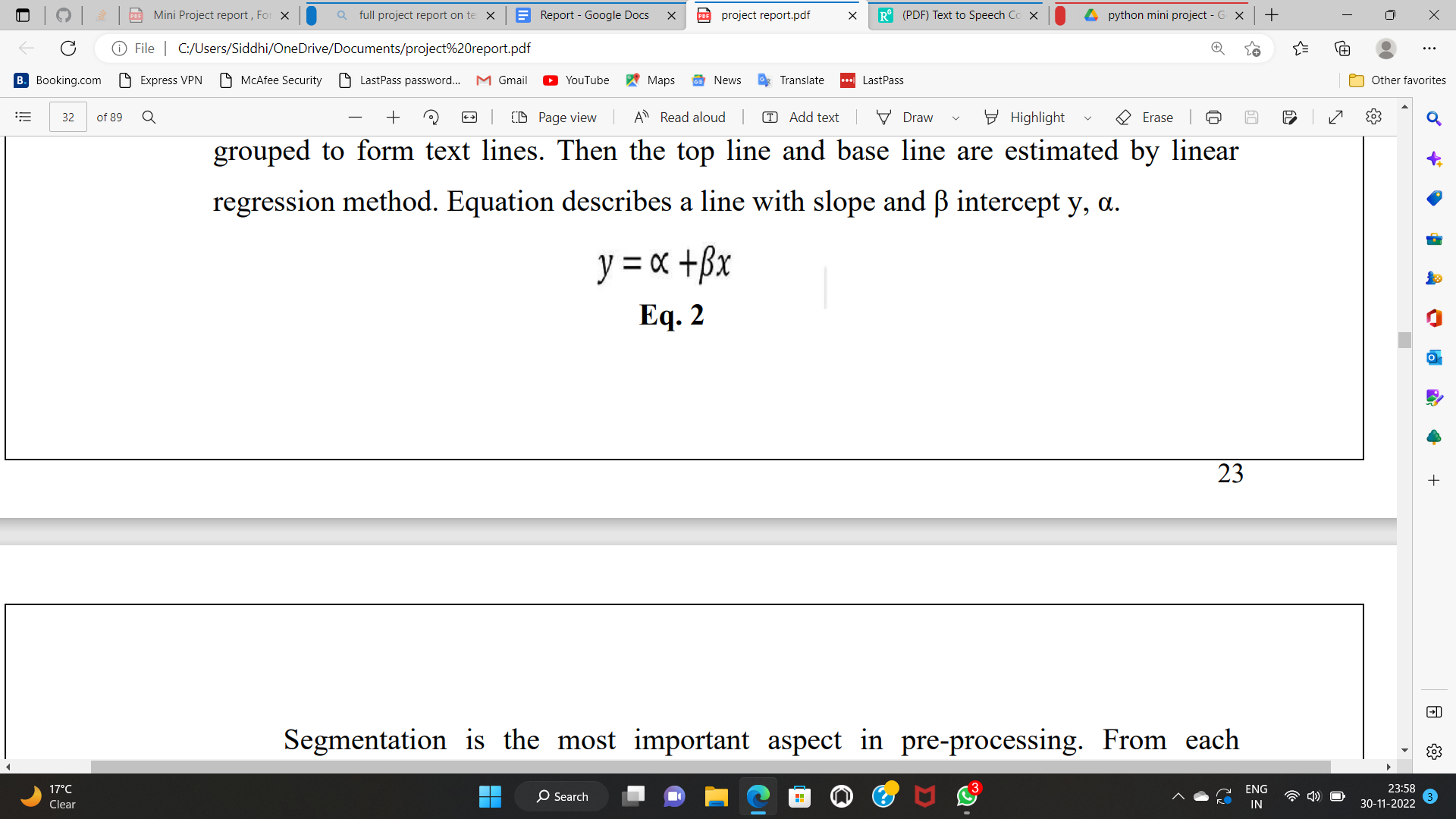
3.1.1 Segmentation and Extraction of Text from Curved Text Lines using Image Processing Approach.

The camera uploads images containing text are having curved text lines because of distortions by page curl and the view angle of the camera. So it is necessary while scanning the document, the text should be straight and words are inline properly. But text lines segmentation of curled text is a difficult method for de-wrapping techniques. This paper presents the method based on image processing algorithms for segmentation and extraction of characters from curled text lines from document images.

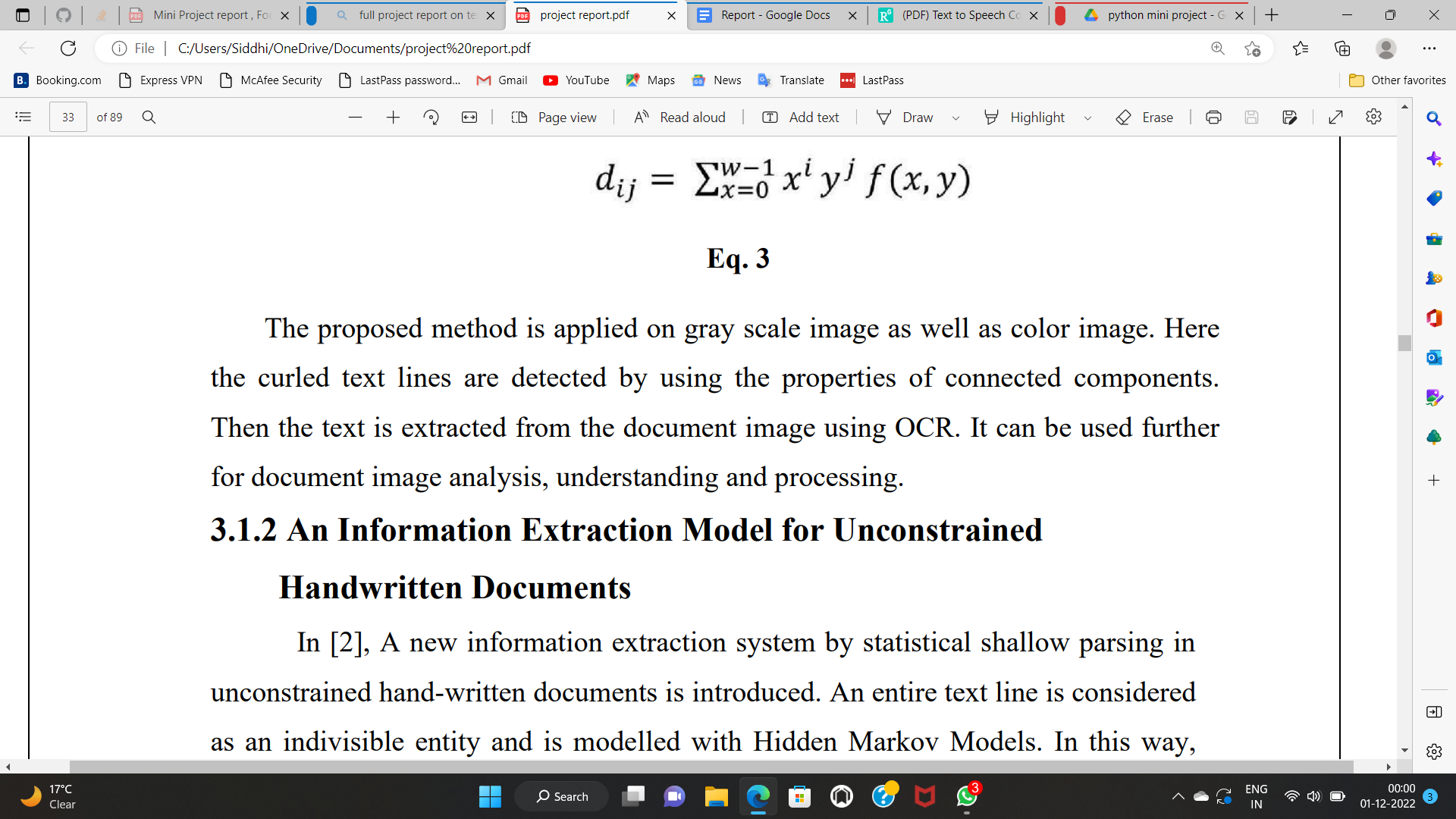
The algorithm performs the curved text segmentation using x-line and base line. The words in the document image are identified and bounding boxes are plotted around the words. The properties of connected components are used for segmentation of words. The Gaussian filter is used for smoothing the text line of the document image. The filtering is performed by computing convolution of the document image and filter mask. The mathematical equation for the Gaussian low pass filter is given by equation,



Where, H(x) transfer function of the Gaussian filter, 1 is the standard deviation of the distribution. The distribution is assumed to have a mean of 0. Linear regression method is used to estimate top line and baseline. Several areas of text in the document have different skew angles. Each Adjacent connected component is grouped using nearest neighbour to form words which are further grouped to form text lines. Then the top line and base line are estimated by linear regression method. Equation describes a line with slope and β intercept y, α.



Segmentation is the most important aspect in pre-processing. From each individual character it extracts the features for character recognition. Moments that make the process of recognizing an object scale, translation, and rotation invariant can be used for character recognition. It is represented by an equation.



**2.2 An Algorithmic Approach for Text Recognition from printed/typed Text Images**

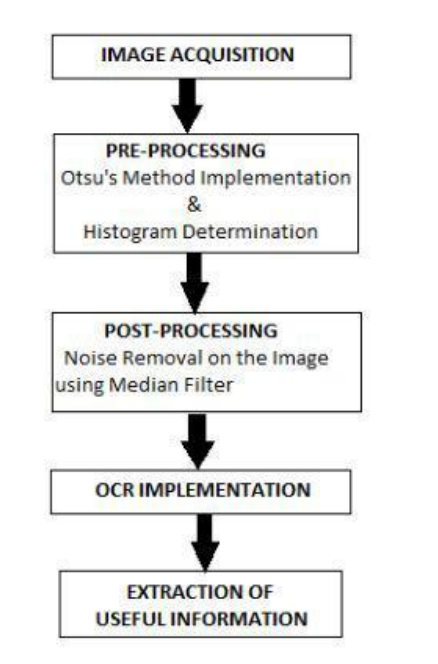
Extraction of texts from scanned copies of documents and text images is an important task in the recent scenario. Optical Character Recognition(OCR) is used to analyse text in images. The proposed algorithm deals with taking a scanned copy of a document as an input and extract texts from the image into a text format using Otsu’s algorithm for segmentation and Hough transform method for skew detection. The system was confined to recognize English alphabets (A-Z, a-z) and numerals (0-9). OCR technique has been implemented to recognize characters. Validation tests were done on screenshots of typed texts and images of scanned documents from Internet sources. Experimental results indicate that the proposed algorithm is able to recognize alphabets written in Verdana font style with size 14 and a showed good results with rotated images.

The proposed algorithm is successfully able to recognize characters from text images with an average accuracy of 93%.

***METHODOLOGY***

**3.1 SYSTEM ARCHITECTURE**

The proposed algorithm is represented by a flow diagram. Each of the input image is pre-processed; Pre-processing is done in various steps. The first step is to take the image from the source and convert into black and white image i.e convert it to grayscale image using global OTSU’s method. In the next step, the output from the first step might contain some noise or distortion. So, to remove such noise we use image denoising techniques. In the third step, a histogram is developed for the noiseless image which helps us in extracting the text form the noiseless image. Now the Tesseract OCR tool is used for extracting the text by dividing the image into edges and each line is checked whether any text is present on the line or not. All the text is extracted from the image. The final step is to organize the text into human understandable format using an algorithm which is described below.



**3.2 PROPOSED ALGORITHM**

In order to organize text from an image, the following steps need to be followed:

1. pre-processing of the image.
2. Histogram generation
3. OCR
4. Organise the text

***3.2.1 Image Pre-Processing and Image Enhancement***

**3.2.1.1 Image Acquisition:** In image processing, image acquisition is done by retrieving an image from a dataset for processing.It is the first step in the workflow sequence because, without an image no processing is possible.

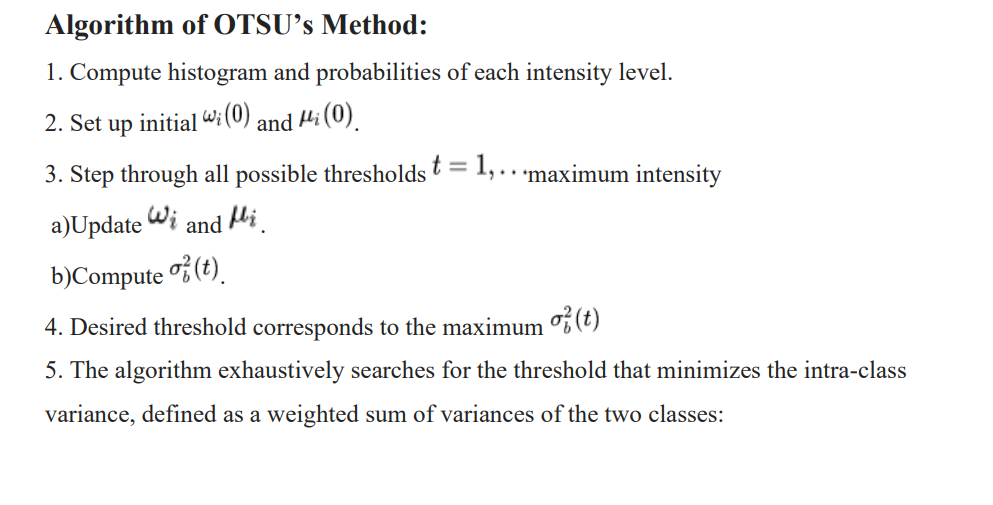
**3.2.1.2 Convert the Image from One Color Space to Another** : There are more than 150 color-space conversion methods available in OpenCV. For color conversion, we use the function cv2.cvtColor(input\_image, flag) where the flag determines the type of conversion. For example, to convert an image from RGB to Gray conversion we use the flag as cv2.COLOR\_BGR2GRAY.

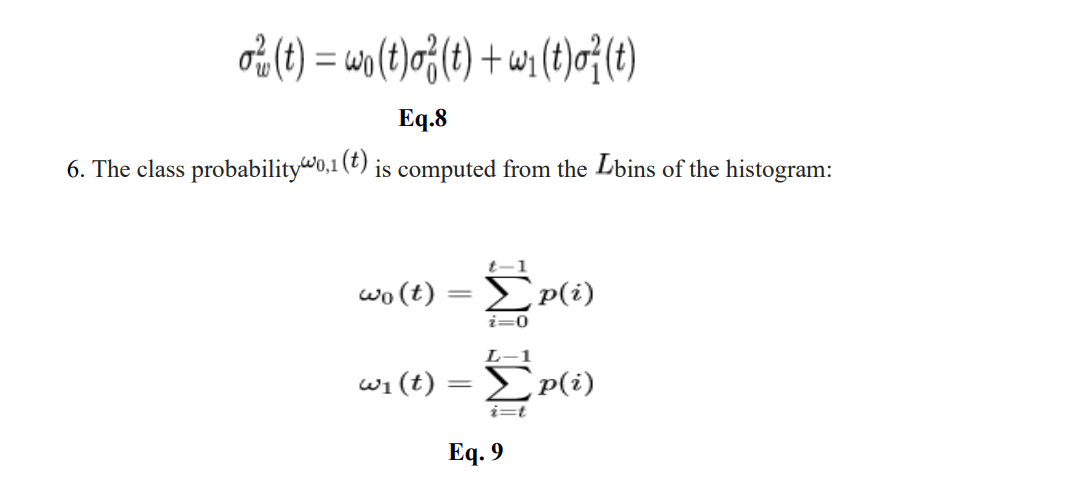
**3.2.1.3 Image Binarization using OTSU's Method:** As a segmentation technique, Otsu’s method is widely used in document binarization, and computer vision. In many cases Otsu’s method is used as a preprocessing technique to segment an image for further processing such as feature analysis and quantification. Otsu’s method searches for a threshold that minimizes the intra-class variances of the segmented image and can achieve good results when the histogram of the original image has two distinct peaks, one belongs to the background, and the other belongs to the foreground or the signal.

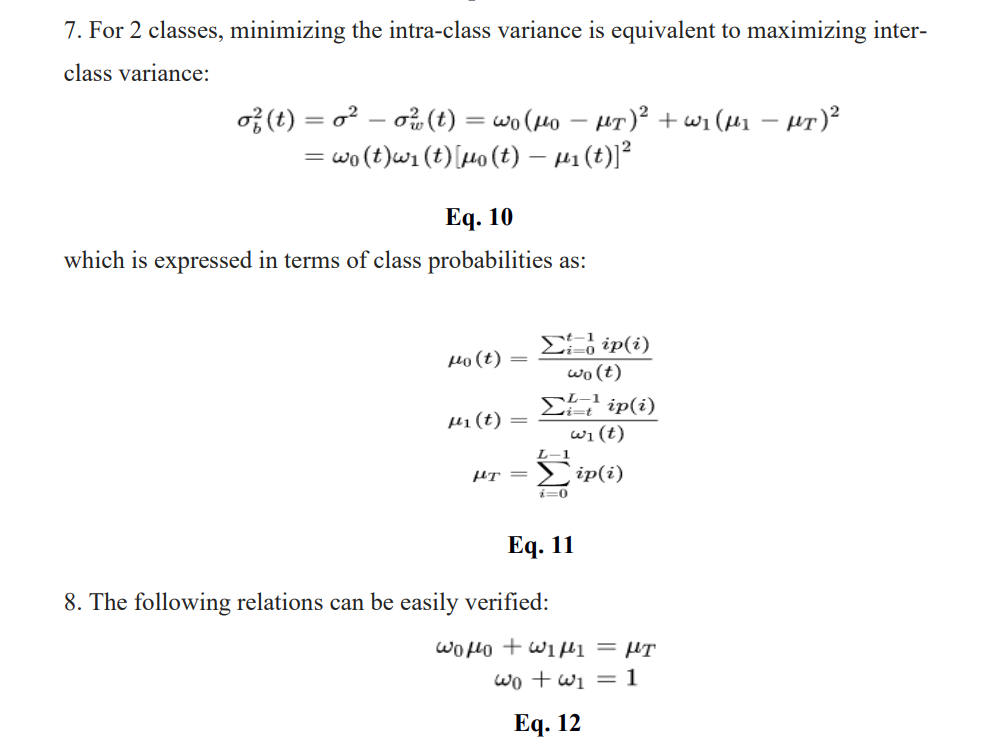
The threshold determined by Otsu’s method is more profoundly determined by the class that has the larger variance, be it the background or the foreground. As such, Otsu’s method may create suboptimal results when the histogram of the image has more than two peaks or if one of the classes has a large variance. Otsu's thresholding method involves iterating through all the possible threshold values and calculating a measure of spread for the pixel levels each side of the threshold, i.e. the pixels that either fall in foreground or background.

The aim is to find the threshold value where the sum of foreground and background spreads is at its minimum.

Thresholding is the simplest method of image segmentation. It is a non-linear operation that converts a grey-scale image into a binary image where the two levels are assigned to pixels that are below or above the specified threshold value. In OpenCV, we use cv2.threshold() function: cv2.threshold(src, thresh, maxval, type[dst])





**3.2.2 OCR (Optical Character Recognition)**

Optical character recognition or optical character reader (OCR) is the mechanical or electronic conversion of images of typed, handwritten or printed text into machine encoded text.

The Techniques involved in the OCR are listed below:

**1) Pre-processing:**

**a) *De-skew****:* To make lines of text perfectly horizontal or vertical.

**b) *De-speckle:*** Remove positive and negative spots, smoothing edges.

**c) *Binarisation:*** Convert an image from color or grayscale to black-and-white.

**d) *Line removal:*** Cleans up non-glyph boxes and lines

**e) *Layout analysis:*** Identifies columns, paragraphs, captions, etc. as distinct blocks.

**f) *Line and word detection:***Establishes baseline for word and character shapes, separates words if necessary.

**g) *Script recognition*:** The script may change at the level of the words and hence, identification of the script is necessary.

**h) *Character isolation:***For per-character OCR, multiple characters that are connected due to image artifacts must be separated; single characters that are broken into multiple pieces due to artifacts must be connected.

**2) Character recognition:** There are two basic types of core OCR algorithm:

**a) *“Matrix matching”*** involves comparing an image to a stored glyph on a pixel by-pixel basis; it is also known as "pattern matching".

**b) *“Feature extraction”*** decomposes glyphs into "features" like lines, closed loops, line direction, and line intersections. Nearest neighbour classifiers such as the k-nearest neighbours algorithm are used to compare image features with stored glyph features and choose the nearest match.

**3) *Post-processing:***

OCR accuracy can be increased if the output is constrained by a list of words that are allowed to occur in a document .

**4) *Application-specific optimizations*:**

The major OCR technology providers began to tweak OCR systems to deal more efficiently with specific types of input. Beyond an application-specific lexicon, better performance may be had by taking into account business rules, standard expression, or rich information contained in color images.

**3.3 Organizing the Text**

An algorithm is used for organizing the extracted text. The steps involved in the algorithm is as follows:

**Step1:** All the Extracted text is first analysed, then each and every text has a particular format i.e pattern.

**Step2:** Identify the pattern for each and every line of the text occurred. Consider the Name, for identifying the text as in the Name all the characters should be alphabets but not numbers. Similarly, For identifying the text as phone number it should contain 10 digits, then it is named as phone number or contact number. **Step3:** All the useful information is organized in the similar pattern recognition format. The rest is unaltered/ not considered.

**Step4:** Output of the organized text is printed.

**3.4 Conversion to Speech:**

Image processing module captures image, converting the image into text. Voice processing mod- ule changes the text into sound and processes it with specific physical characteristics so that the sound can be understood.

The block diagram of Text- To-Speech device,

1 st block is an image processing module, where OCR converts .jpg to .txt form.

2 nd is voice processing module which converts .txt to speech.

Block diagram of text-to-speech device. the block diagram of Text-To-Speech device, 1 st block is image processing module, where OCR converts .jpg to .txt form. 2 nd is voice processing module which converts .txt to speech. OCR is important element in this module.

**CODE AND**

**OUTPUT**

**CODE:**

import cv2

import pytesseract

from PIL import Image

from gtts import gTTS

from playsound import playsound

pytesseract.pytesseract.tesseract\_cmd = r"C:\Program Files\Tesseract-OCR\tesseract.exe"

img = cv2.imread("img1.png") # to read image

img = cv2.cvtColor(img, cv2.COLOR\_BGR2RGB) # to convert into rgb

# cv2.imshow("Image",img) #to open image

img\_h, img\_w, img\_c = img.shape #shape of image

# boxes = pytesseract.image\_to\_boxes(img)

# for box in boxes.splitlines():

# box = box.split(" ")

# #print(box)

# x, y, w, h = [int(x) for x in box[1:5]]

# #print(x,y,w,h)

# cv2.rectangle(img, (x, img\_h -y), (w, img\_h -h),(0,0,255),1)

# cv2.putText(img, box[0] , (x, img\_h - y +10), cv2.FONT\_HERSHEY\_COMPLEX, 0.5,(0,0,255),1)

boxes = pytesseract.image\_to\_data(img)

sentence = ""

for i, box in enumerate(boxes.splitlines()):

#print(box)

if i == 0:

continue

box = box.split()

#print(box)

if len(box)==12:

x, y, w, h = [int(x) for x in box[6:10]]

#print(x, y, w, h)

cv2.rectangle(img, (x,y), (x+w, y+h),(0,0,255),1)

cv2.putText(img, box[11], (x, y+30), cv2.FONT\_HERSHEY\_COMPLEX, 0.5,(0,0,255),1)

sentence += box[11]+" "

cv2.imshow("Image",img) #to open image

print(sentence) #to print the text on image

#text to sound

sound = gTTS(text = sentence, lang= 'en')

sound.save("sound.mp3")

playsound("sound.mp3")

#exit

if cv2.waitKey(0) & 0XFF == ord('q') : #to close image if q is pressed

cv2.destroyAllWindows()

**OUTPUT:**

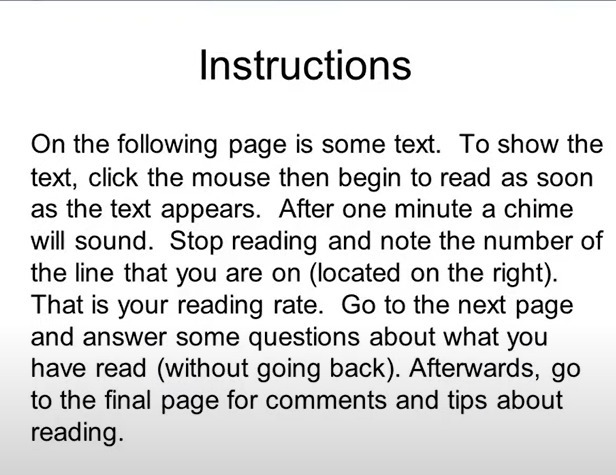
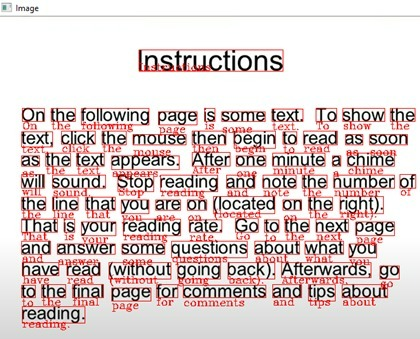
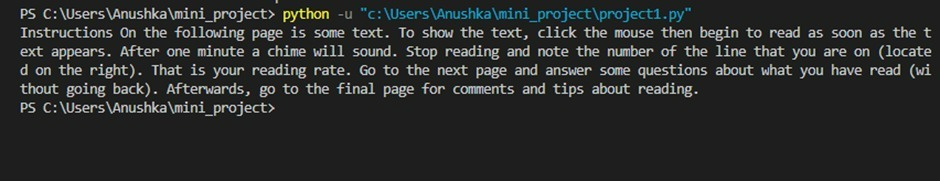
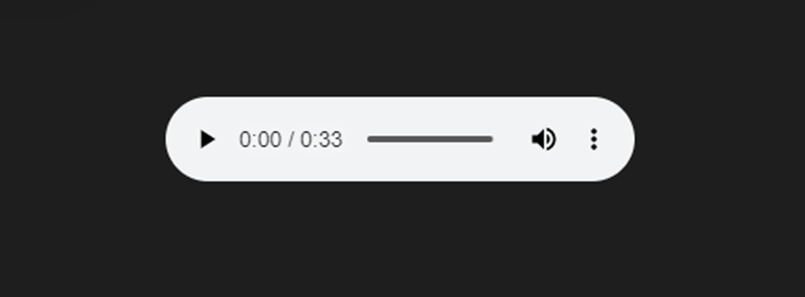


image:

image

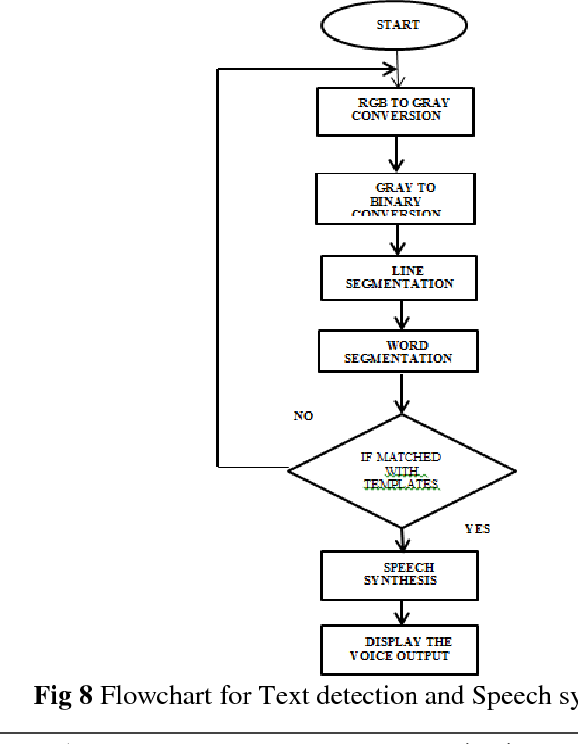
text





Speech

**Flowchart:**

****

**CONCLUSION**

In this system, we have implemented a system in which the text can be extracted from the scanned image and organized into a specific aligned format. The proposed algorithm has also shown accurate results and was also able to reduce noise from images to a good extent. This cannot work efficiently if the image format is not in the same alignment as described in the system. OCR used in this system cannot work properly when the image contains handwritten texts. The results are accurate enough for the specified format

**FUTURE WORK &**

**REFERENCES**

**FUTURE WORK**

As the proposed system can only organize text for some of the image documents such as identification proofs and a specific application form, this can be extended to other image documents as per user request and test it on well-designed datasets for improving the accuracy of the system.

**REFERENCES**

[1] Shejwal, M. A., & Bharkad, S. D. (2017). Segmentation and extraction of text from curved text lines using image processing approach. 2017 International Conference on Information, Communication, Instrumentation and Control.

[2] Thomas, S., Chatelain, C., Heutte, L., & Paquet, T. (2010). An Information Extraction Model for Unconstrained Handwritten Documents. 2010 20th International Conference on Pattern Recognition.

[3] Kochi, T., & Saitoh, T. (1999). User-defined template for identifying document type and extracting information from documents. Proceedings of the Fifth International Conference on Document Analysis and Recognition. ICDAR ’99.

[4] Agrawal, N., & Kaur, A. (2018). An Algorithmic Approach for Text Recognition from Printed/Typed Text Images. 2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence).

[6] Zhen-Long BAI and Qiang HUO. An Approach to Extracting the Target Text Line from a Document Image Captured by a Pen Scanner.

[7] K.N. Natei, J. Viradiya, S. Sasikumar. Extracting Text from Image Document and Displaying Its Related Information.

[9] Puthipong Sthitpattanapongsa and Thitiwan Srinark. A Two-stage Otsu’s Thresholding Based Method on a 2D Histogram.