GENERATION OF LECTURE NOTES USING HANDWRITING AND SPEECH RECOGNITION

Project Report

Submitted in fulfillment of the requirements for the degree of

B.E. (Information Technology)

By

Maclyn Afonso (Seat No: 0362)
Poornima Santosh Kole (Seat No: 0388)
Krupesh Ramesh Davajekar (Seat No: 0372)

Guide: Prof. Soniya Usgaonkar
Co-Guide: Prof. Mario Pinto

Information Technology Department
Goa College of Engineering
(Government of Goa)
Goa University
(2016-2017)

DISSERTATION APPROVAL SHEET

This is to certify that

Maclyn Afonso (Seat No: 0362)
Poornima Santosh Kole (Seat No: 0388)
Krupesh Ramesh Davajekar (Seat No: 0372)

Have been admitted to the candidacy of B.E. (Information Technology) in December/January 2016/17 and they have undertaken the project entitled "GENERATION OF LECTURE NOTES USING HANDWRITING AND SPEECH RECOGNITION" which is approved for the degree of B.E. (Information Technology) under Goa University as it is found satisfactory.

Guide(s)

Prof. Soniya Usgaonkar Guide Assistant Professor Dept of I.T, GEC	Prof. Mario Pinto Co-Guide Assistant Professor Dept of I.T, GEC

Dr. Anant J. Naik
Head of Department
Information Technology
Goa College of Engineering
Farmagudi

Examiners

(Internal Examiner) Name: Designation:	(External Examiner) Name: Designation:

ACKNOWLEDGEMENT

It is our privilege to express our sincerest regards to our project guide Prof. Soniya Usgaonkar and co-guide Prof. Mario Pinto for their valuable inputs, able guidance, encouragement, whole-hearted cooperation and constructive.
We deeply express our sincere thanks to our Head of Department Dr. Anant Naik for encouraging and allowing us to present the project on the topic “Generation of lecture notes using handwriting and speech recognition” at our Department premises.
We take this opportunity to thank all our lecturers who have directly or indirectly helped in our project. We pay our respect and love to our parents and friends for their love and encouragement.
Last but not the least, we express our thanks to each of our team members for all their support and cooperation.

ABSTRACT

When conducting a lecture, the main aim of the lecturer is imparting knowledge. However it has been observed that teachers mainly focus on getting the students complete their books by merely copying notes from a board. This wastes a lot of time and hence can compromise the goal of education.
We attempt to reduce the amount of time spent by the students in copying notes by automatically translating text on the blackboard into digital notes. Hence the teacher can spend more time on explanation rather than waiting for students to complete copying. This can also be done by using smart boards, but the cost of acquiring such expensive hardware will be burdensome to many. Hence the proposed system is designed with cost and quality in mind. The proposed system for generation of lecture notes using handwriting and speech recognition will be able to incorporate new methodologies/algorithms if required. However we aim to achieve the results with the basic available algorithms. Various hardware and software upgrades and optimizations can be added in the future to achieve a full smart classroom environment.

CONTENTS

CHAPTER **Page No.**

Acknowledgement……I
Abstract……...II
List of Figures……IV
List of Abbreviations……..V

1: INTRODUCTION……..……..9
1.1 Problem Definition.....……....9
1.2 Introduction to Project…….9

2: LITERATURE REVIEW…….....10
2.1 Background……....10
2.2 Analysis of Papers……10
2.3 Approaches to Project……11
2.4 Algorithms and Solution.....……11
2.5 Applications of Project ……12

3: SYSTEM REQUIREMENT AND SPECIFICATIONS…….....17
3.1 Software Requirements…….17
3.2 Hardware Requirements……17

4: DESIGN……..……18
4.1 Data Flow Diagram.……...18
4.2 Detail Process Flow……...19

5: IMPLEMENTATION……...20

6: CONCLUSION……..25

7: FUTURE SCOPE……..26

8: BIBLIOGRAPHY…….27

LIST OF FIGURES

Figure	Name	Page No.
Fig 2.4.1	Proposed System Architecture	4
Fig 4.2.1	Proposed System Process Flow	9
Fig 5.1	System Setup	10
Fig 5.2	Initial Input	10
Fig 5.3	Binarization Output	11
Fig 5.4	Line Segmentation Output	11
Fig 5.5	Word Segmentation Output	12
Fig 5.6	Observation Table	13

LIST OF ABBREVIATIONS

ABBREVIATION	EXPANSION
PiCam	Raspberry Pi camera module
Rpi	Raspberry Pi circuit board
USB	Universal Serial Bus
VGA	Video graphics Accelerator
KNN	K-Nearest Neighbor

CHAPTER 1

INTRODUCTION

1.1 PROBLEM DEFINITION

Generation of lecture notes using handwriting and speech recognition is an approach that provides a way to capture notes and discussions in a class environment, and make them available to students online whenever needed. The current methodology wherein students manually take down notes wastes a lot of time. Also, the lecture is just a one time activity, students cannot go back and review the lectures when they need an explanation of some topic that was previously discussed. The proposed system solves this problem by its online functionalities. The use of smart boards is a growing aspect. But due to the cost of such a system, it is not available to all. The proposed system can overcome this cost barrier by a large margin. Moreover, it is also easily re-locatable unlike the smart board setup. This system is targeted towards those lectures where majority of the notes are delivered through writing on boards and power point presentations. Hence the major users of this system will be lower level institutions.

1.2 INTRODUCTION TO PROJECT

This project forms a part of the idea of smart classrooms where in learning can be made an interactive and rapid process. The proposed system helps in reducing the time wasted by students while copying notes from the board by automatically capturing the notes and uploading it onto a server from where it will be accessible by all concerned students. Students will now be able to keep up with the discussions even if they were unable to attend the lectures. This idea will also shift the lectures focus towards preparing good content for the lecture rather than just preparing notes.

CHAPTER 2

LITERATURE REVIEW

2.1 BACKGROUND

In today's ever growing and emerging fields of computer science, one of the major field of research that still has great work to be done is handwritten text recognition. Since, computer plays an important role in various small and large scale industries for automation, various departments in today's mode of computerization uses handwritten text conversion mechanisms. Some of the obvious application that contains the auto indexing and processing of archived documents, self introducing forms, envelopes, hand written notes, ancients hand written text requires the use of handwritten text recognition techniques. Off-line character recognition is a mechanism of text recognition in which handwritten characters are recognized from papers, messages. In order to perform the offline handwritten text recognition, the text goes through various algorithms. The input text image is recognized by the recognition engine in the form of binary image or grayscale image.

2.2 ANALYSIS OF PAPERS

Off-Line Handwritten Character Recognition using Features Extracted from Binarization Technique by Amit Choudhary, Rahul Rishi, Savita Ahlawat states that the choice of pattern classifier and the technique used to extract the features are the main factors to judge the recognition accuracy and the capability of an Optical Character Recognition (OCR) system. The main focus of this work is to extract features obtained by binarization technique for recognition of handwritten characters of English language. The recognition of handwritten character images have been done by using multi-layered feed forward artificial neural network as a classifier. Some pre-processing techniques such as thinning, foreground and background noise removal, cropping and size normalization etc. are also employed to pre-process the character images before their classification. Very promising results are achieved when binarization features and the multilayer feed forward neural network classifier is used to recognize the off-line cursive handwritten characters.

A New Character Segmentation Approach for Off-Line Cursive Handwritten Words by Amit Choudhary, Rahul Rishi, Savita Ahlawat - This paper proposes a new vertical segmentation algorithm in which the segmentation points are located after thinning the word image to get the stroke width of a single pixel. The knowledge of shape and geometry of English characters is used in the segmentation process to detect ligatures. The proposed segmentation approach was tested on a local benchmark database and high segmentation accuracy was achieved.

An Improved Method for Handwritten Document Analysis using Segmentation, Baseline Recognition and Writing Pressure Detection by Abhishek Bal, and Rajib Saha - Handwritten document analysis is a scientific technique for identifying and understanding the personality of a writer through the strokes and patterns revealed by writer's handwriting. This research proposed an off-line handwritten document analysis through segmentation, skew recognition and writing pressure detection for cursive handwritten document. The proposed segmentation method is based on modified horizontal and vertical projection that can segment the text lines and words even if the presence of overlapped and multi-skewed text lines.

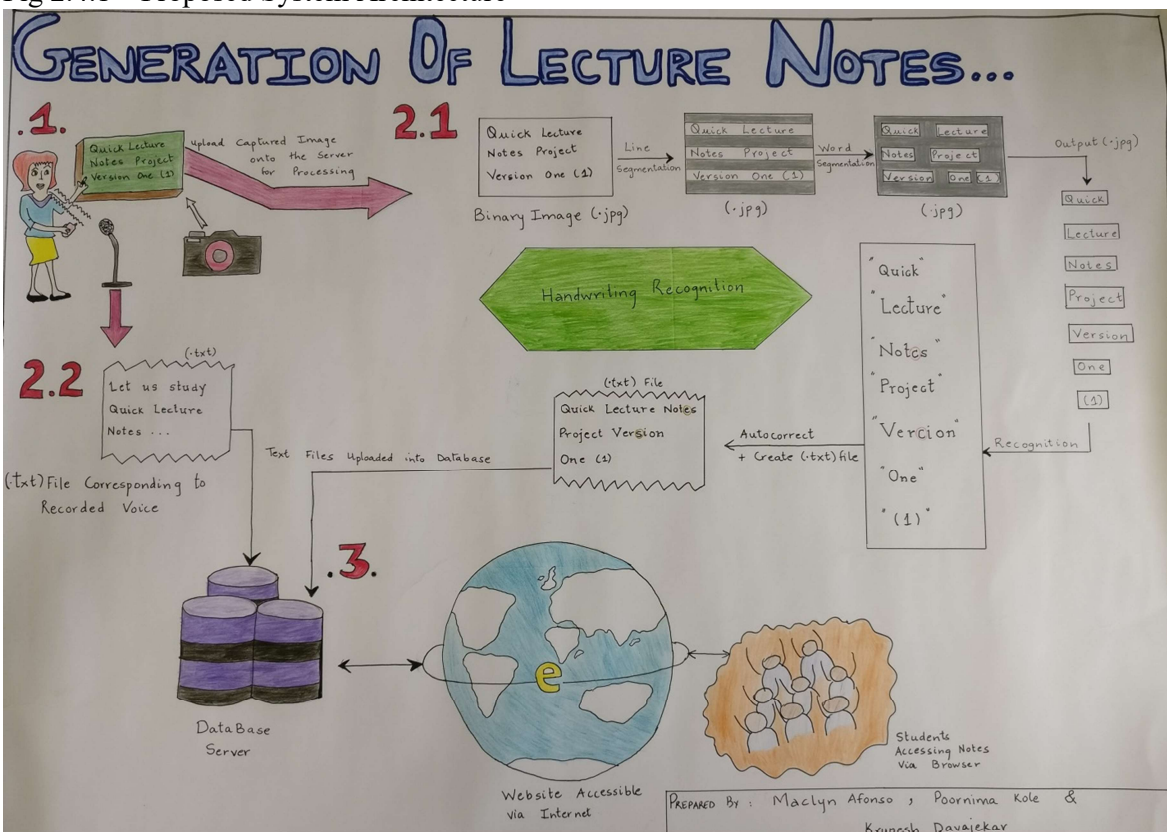
Handwritten Document into Digitized Text Using Segmentation Algorithm by D.Kavitha and P.Shamini - This paper presents an innovative application that can convert cursive handwriting document into digital text using intelligent word recognition (IWR).Intelligent Word Recognition uses artificial intelligence to recognize whole words in a document instead of individual characters. Handwriting Conversion helps preserve history by making information searchable, easily and reportable without the need for human labor. It can be applied for digitizing old manuscripts which helps in preservation of historical data.

2.3 APPROACHES TO PROJECT

Through this project we attempt to provide an interactive system for recording lecture notes. The system will have two main sources of input, first is the handwritten notes and second is speech. Students will be able to view the digitized notes online and also listen to the explanations whenever required. Also lecture schedules will be managed in order to organize the notes as per the respective subjects. Additional information sources could be provided as attachments in the form of links to pages or documents. Also, the decision to use of the system as note generator or simple video lecture recorder will be left to the lecturer.

2.4 ALGORITHMS AND SOLUTIONS

Fig 2.4.1 – Proposed System Architecture



Binarization Algorithm(OTSU)

For each potential threshold T:
Separate the pixels into two clusters based on threshold.
Find the mean of each cluster.
Square the difference between the means.
Multiply the number of pixels in one cluster times the number in the other.
If math from previous section is used then :
Compute histogram and probabilities of each intensity level.
Set initial Qi(0) and Ui(0).
Step through all possible threshold max intensity.
Update Qi and Ui and Compute Sigma_sq(t).
Desired Threshold corresponds to max.

Algorithm for Line Segmentation

1. Line segmentation algorithm is carried out by horizontal projection as follows

2. Read a handwritten document image as a multi-dimensional array.
3. Check the image is a binary image or not. If binary image then store it into a 2-d array IMG[ ][ ] with size M×N and go to Step 4,otherwise go to Step 3.

4. Convert the image to binary image and store into a 2-d array IMG[ ][ ].
5. Construct the horizontal projection histogram of the image IMG[ ][ ] and store into a 2-d array HPH[ ][ ].
6. Measure the height, starting row position and ending row position of each horizontally rising section of horizontal projection histogram image and store into 3–d array LH[ ][ ][ ] sequentially.
7. Count the number of rising section by counting the rows of the 3-d array LH[ ][ ][ ]. Then measure the threshold (Ti) value by calculating average height of rising sections from the 3-d array LH[ ][ ][ ].
8. Select each rising section from 3-d array LH[ ][ ][ ] and check the height of that rising section is less than the threshold or not. If yes then this rising sections is not considered as a line and go to Step 9, otherwise rising section is treated as a line and go to Step 8.
9. Find the rising section's starting and ending rows number from the array LH[ ][ ][ ]. Let starting and ending row are r1 and r2 respectively. Extract the line segment between r1 and r2 from the original binary image denoted by IMG[ ][ ].
10. Go to Step 7 for next rising sections till all rising section are not under consideration, otherwise go to next Step.
11. End

Algorithm for Word Segmentation

1. Read a segmented binary line as 2-d binary image LN[ ][ ].
2. Construct the vertical projection histogram of the line LN[ ][ ] and store into a 2-d array LVP[][].
3. From the vertical projection histogram (LVP[ ][ ]), measures width of each inter-word and intra-word gaps and store the width into 1-d array GAPSW[ ].
4. Count total number gaps as TGP by calculating the size of GAPSW[ ]. Add width of all gaps by adding the elements of GAPSW[ ] and store into TWD.
5. Calculate the threshold (Ti) as follows:
6. Ti = TWD/ TGP
7. In this equation Ti is the threshold value denoting average width of inter-word gaps, TWD denotes total width of all gaps and TGP denotes the total number of gaps.
8. For each i(1<=i<=sizeof(GAPSW[ ])), if GAPSW[i]>=Ti then this gaps is treated as interword gaps, otherwise gaps is treated as an intra-word gaps. Depending on inter-word gaps width, words are segmented from the line.
9. End

Character Classifier – K-Nearest Neighbor(KNN)

• In pattern recognition, the k-nearest neighbor algorithm is a method to classify objects based on nearest training sets in the feature space.
• Using the concept of majority voting of neighbors, an object is classified with being assigned to the class most common amongst its k nearest neighbors, where k is a positive integer (typically small).

• If k = 1, then the object is simply assigned to the class of its nearest neighbor.
• Given a training set D and a test object x = (x', y'), the algorithm computes the distance (or similarity) between z and all the training objects (x, y) ∈ D to determine its nearest neighbor list, Dz. (x is the data of a training object, while y is its class. Likewise, x' is the data of the test object and y' is its class). Once the nearest-neighbor list is obtained, the test object is classified based on the majority class of its nearest neighbors.

2.5 APPLICATIONS OF PROJECT

Here are several applications of the proposed project:

→ Automatic conversion of handwriting to text.

→ Allows for speech recording as lecture explanations.

→ Lecture content can reviewed as many time as possible.

→ Avoid wastage of time in copying notes manually.

→ Easily upgradable as and when new technology becomes available.

CHAPTER 3

SYSTEM REQUIREMENT AND SPECIFICATIONS

3.1 SOFTWARE REQUIREMENTS

The PiCam can be accessed via Rpi using Python library. The mobile application notifications are handled via Google Cloud Messaging service. Hence, for pre-3.0 devices, this requires users to set up their Google account on their mobile devices. A Google account is not a requirement on devices running Android 4.0.4 or higher. Image processing can be done on the server using OpenCV graphics library.

3.2 HARDWARE REQUIREMENTS

The system's hardware consists of a raspberry pi board with a PiCam mounted over it with USB microphone and an SD Card to store the Operating System and installed applications. The following points are to be noted regarding the camera.

→ It is vulnerable to static electricity (it is advisable to do a sink tap before handling the device)

→ The module has a 5 megapixel fixed-focus camera that supports 1080p30, 720p60 and VGA90 video modes, as well as stills capture.

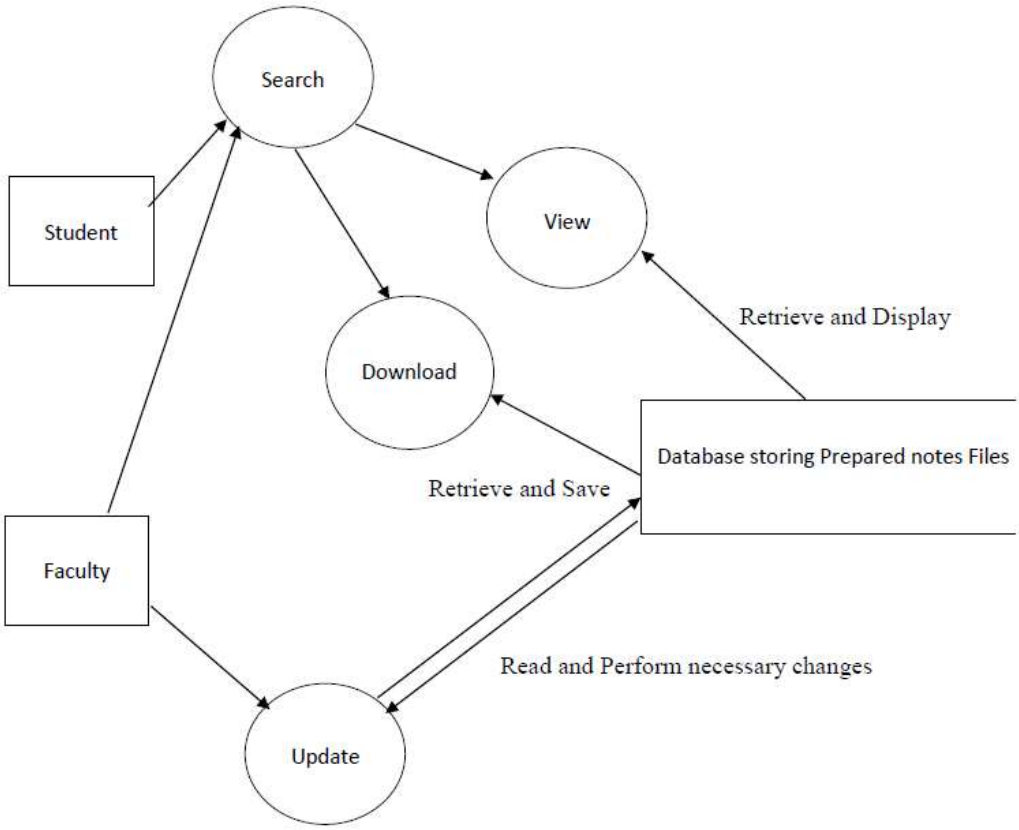
→ It attaches via a 15cm ribbon cable to the CSI port on the Raspberry Pi.

→ The blue face of the ribbon cable faces the Ethernet connector.

CHAPTER 4

DESIGN

4.1 DATA FLOW DIAGRAM



4.2 DETAIL PROCESS FLOW

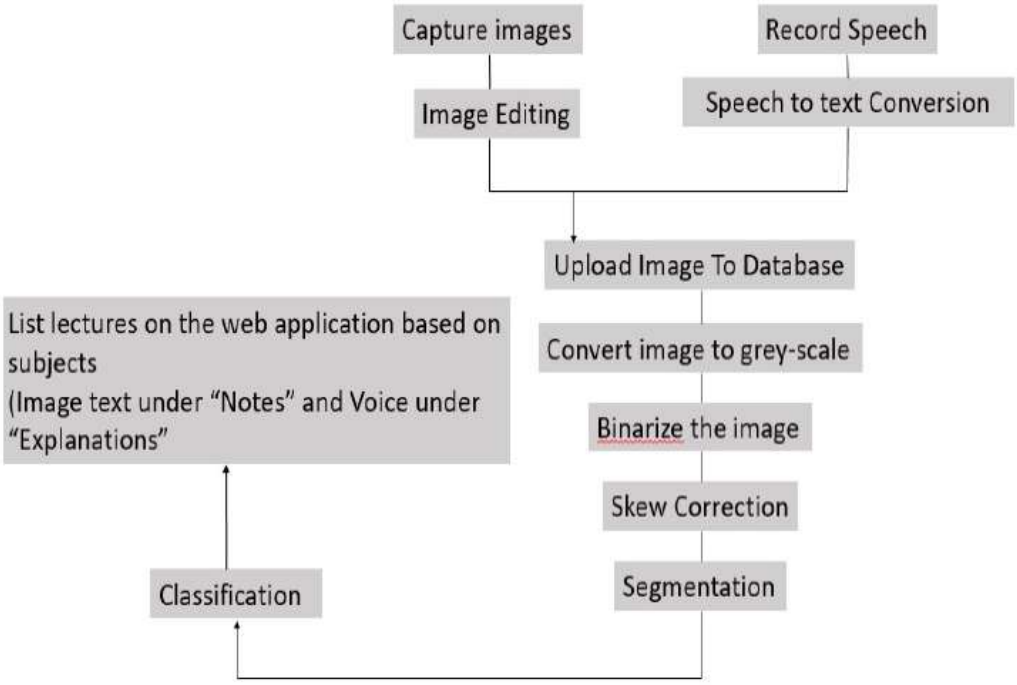


Fig 4.2.1 – Proposed System Process Flow

CHAPTER 5

IMPLEMENTATION

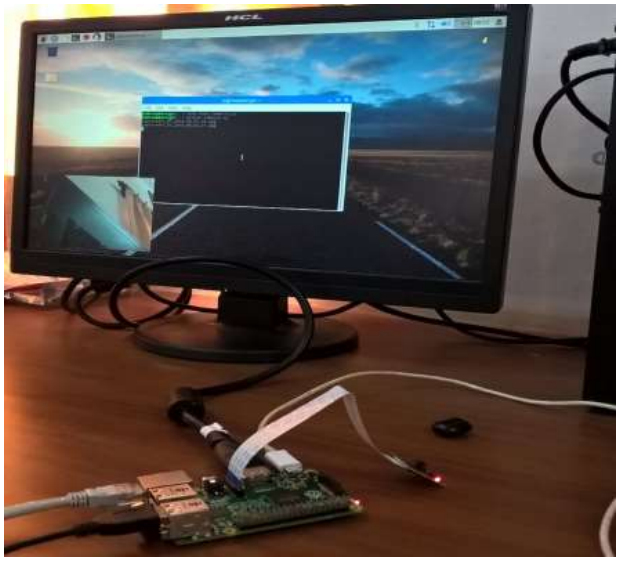


Fig 5.1 : System Setup

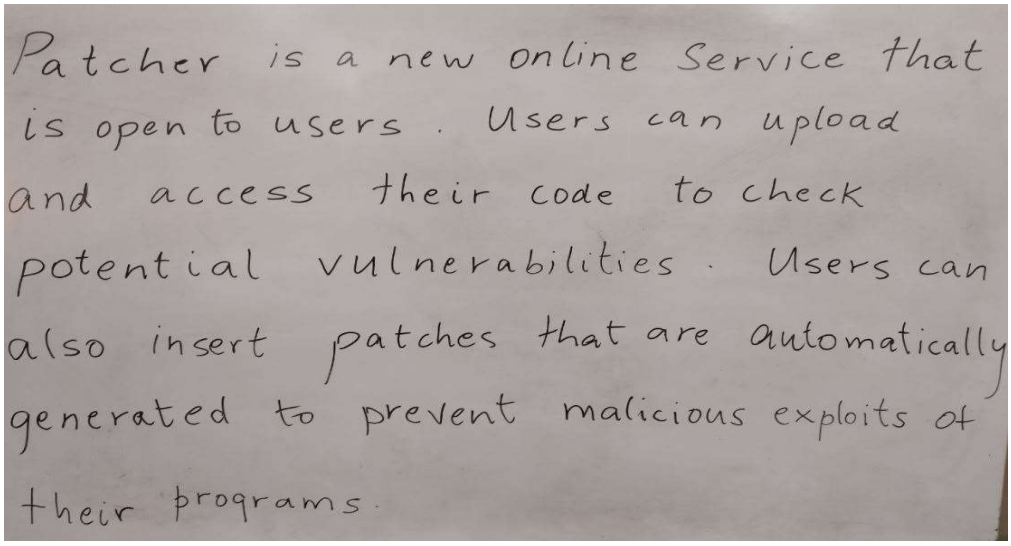


Fig 5.2 : Initial Input

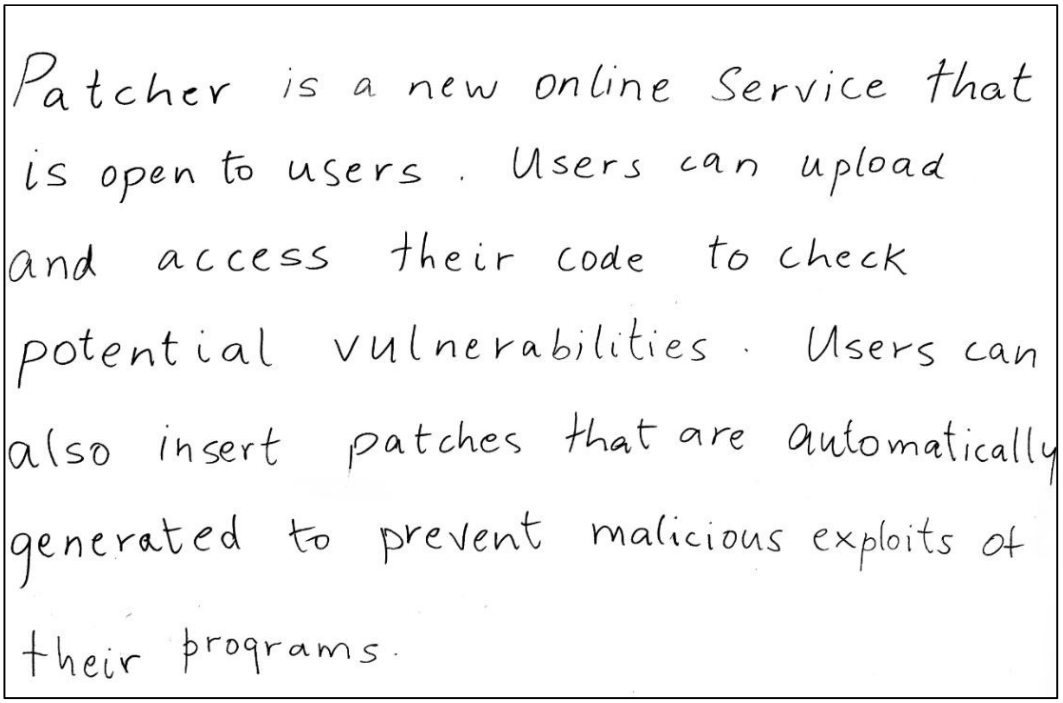


Fig 5.3 : Binarization Output

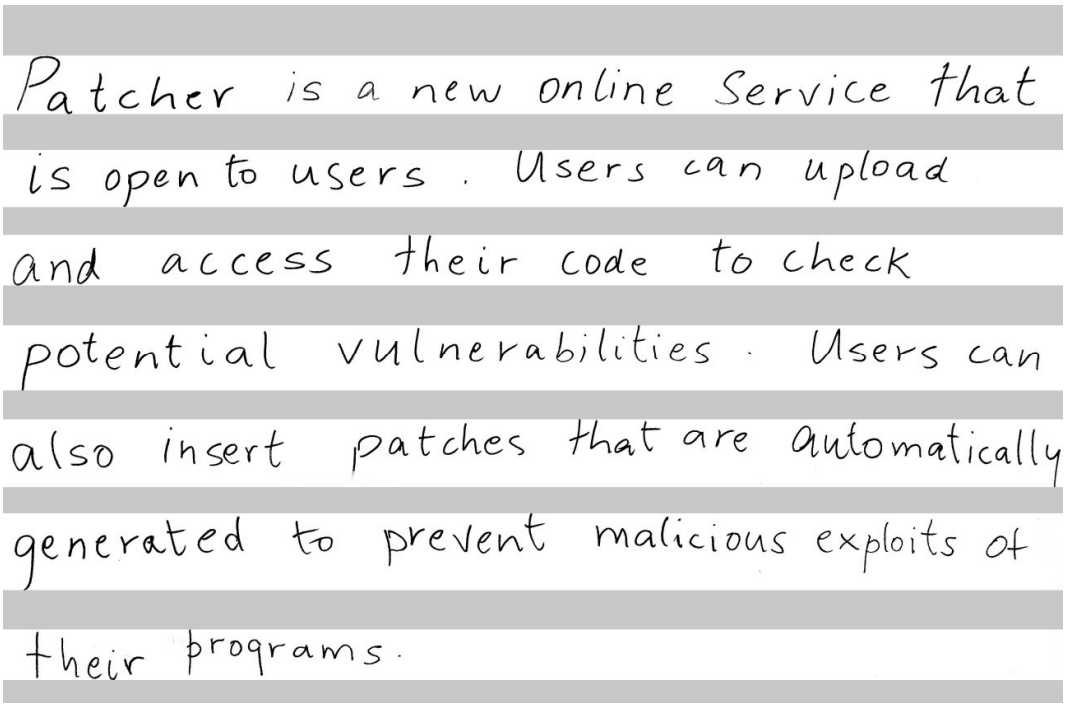


Fig 5.4 : Line Segmentation Output

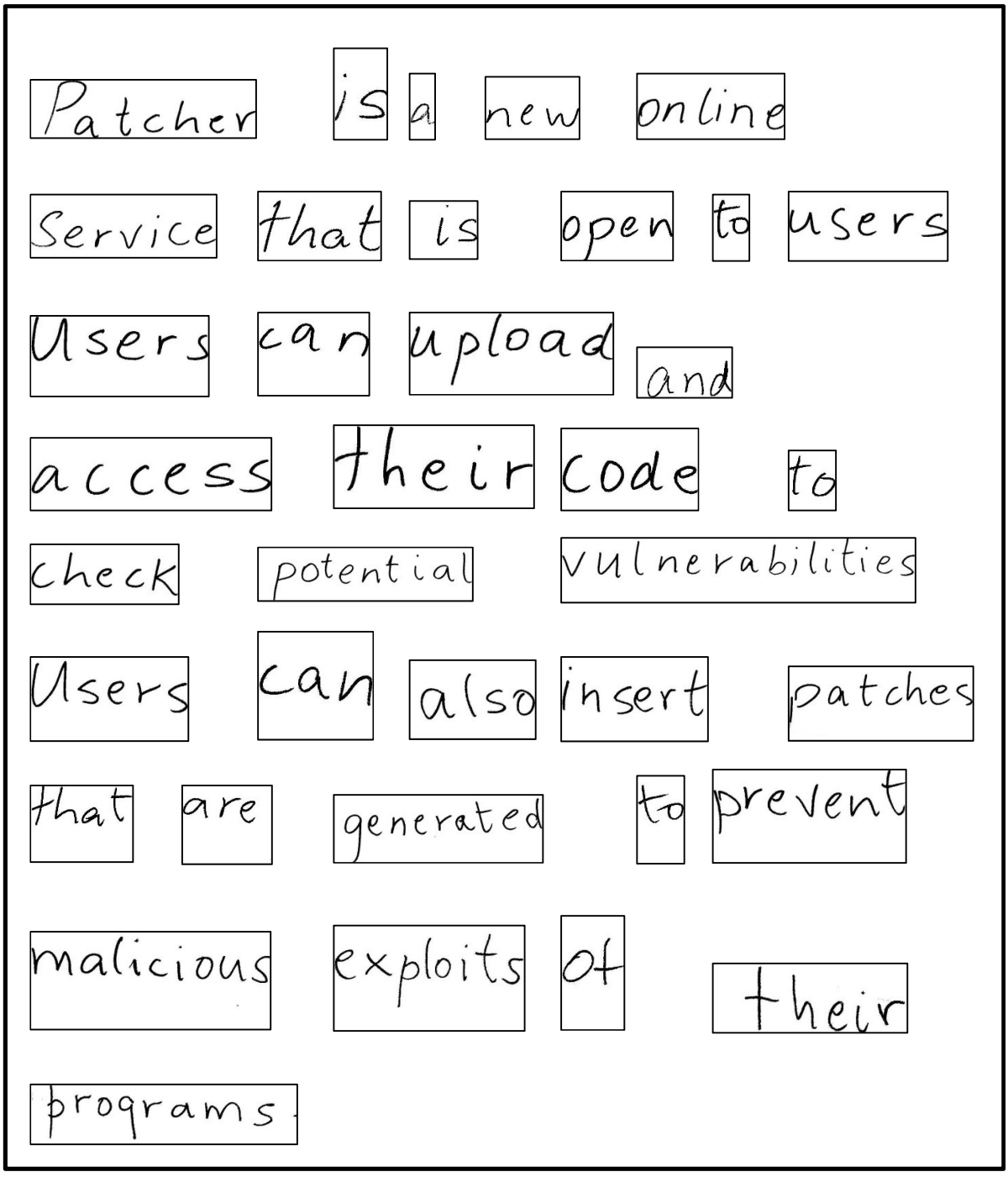


Fig 5.5 : Word Segmentation Output

Total Number of Samples = 20	
Number of Samples	**Outcome**
9	PERFECT
6	VERY GOOD
3	GOOD
2	POOR

Fig 5.6 : Observation Table

CONCLUSION

A system has been proposed to generate lecture notes using handwriting recognition and speech analysis. The proposed system has six processing modules: image preprocessing module that includes conversion to grayscale and then binarization, segmentation module that performs line, word and finally character segmentation, recognition module that uses KNN classification to detect character, speech module to convert speech to text, web module to make content accessible via the internet and finally the mobile application to manage the system via android mobile.
Also, the image preprocessing module has been implemented successfully using live inputs.

FUTURE SCOPE

The proposed system has ample amount of future scope. Some of the suggestion for future improvements are:
Use of Neural network classifier for better character recognition compared to KNN.
Semantic analysis of prepared notes using natural language processing techniques to obtain grammatically correct notes. Also, certain keywords can be searched in the text to add more useful information to the notes from other sources such as the internet or text book PDFs.
Providing for live streaming of lectures over the internet for those who are unable to attend lectures due to genuine reasons.

BIBLIOGRAPHY

[1] Off-Line Handwritten Character Recognition using Features Extracted from Binarization Technique by Amit Choudhary, Rahul Rishi, Savita Ahlawat.

[2] Cursive character segmentation using neural network techniques by Michael Blumenstein School of Information and Communication Technology, Griffith University, Gold Coast campus, PMB 50 Gold Coast Mail Centre, Queensland 9726, Australia.

[3] A New Character Segmentation Approach for Off-Line Cursive Handwritten Words by Amit Choudhary, Rahul Rishi, Savita Ahlawat.

[4] Enhanced Approach For Handwritten Text Recognition Using Neural Network by Aman Puri, Kamlesh Lakhwani Suresh Gyan Vihar University, Jaipur.

[5] Line And Word Segmentation of Handwritten Documents by G. Louloudis, B. Gatos, I. Pratikakis, C. Halatsis.

[6] A Simplified Method for Handwritten Character Recognition from Document Image by Mohammad Imrul Jubair and Prianka Banik Lecturer Department of CSE BGC Trust University Bangladesh.

[7] An Improved Method for Handwritten Document Analysis using Segmentation, Baseline Recognition and Writing Pressure Detection by Abhishek Bal and Rajib Saha.

[8] Handwritten Document into Digitized Text Using Segmentation Algorithm by D.kavitha, P.Shamini Department of Computer Applications, Easwari Engineering College, Chennai, Tamil nadu.