SMART NOTES MAKER FOR VISUALLY OR HEARING IMPAIRED STUDENTS

PROJECT REPORT FOR
CREATIVE AND INNOVATIVE PROJECT

Submitted by

YOGEESWAR.S (2019503573)
VIMAL RAAJ.M (2019503571)
KAVINKUMAR.S (2019503022)

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

MADRAS INSTITUTE OF TECHNOLOGY, CHROMEPET

ANNA UNIVERSITY, CHENNAI 600 044

JUNE 2022

ANNA UNIVERSITY

BONAFIDE CERTIFICATE

Certified that this project report titled "**SMART NOTES MAKER FOR VISUALLY OR HEARING IMPAIRED STUDENTS**" is the bonafide work of "**YOGEESWAR.S, VIMAL RAAJ.M & KAVINKUMAR.S**" who carried out the project work under my supervision.

SIGNATURE OF HOD

Dr. JAYASHREE P
Professor & Head

Department of Computer Technology
Madras Institute of Technology Campus,
Anna University,
Chromepet, Chennai – 600 044

SIGNATURE OF SUPERVISOR

Dr. JAYASHREE P
Professor
SUPERVISOR
Department of Computer Technology
Madras Institute of Technology Campus,
Anna University,
Chromepet, Chennai – 600 044

ABSTRACT

Actively taking notes during class can help focus and better understand main concepts. In many classes, one may be asked to watch an instructional video before a class discussion as good note-taking will improve one's active listening, comprehension of material, and retention. Taking notes on both synchronous and asynchronous material will help one better remember what one hears and sees, but this action of note making can be challenging for students with hearing or visual impairments. Students with hearing impairments cannot process the audio of oral lectures similarly students with visual impairments cannot easily translate the teacher’s speech into text and also cannot see text written on the blackboard. In order to remove such problems a system is designed which can be incorporated in classrooms so that the lectures communicated to the students are understood without any barrier. This project aims to create a product which will help visually or hearing impaired students for effective note making by providing a live captioning system for hearing impaired students and also extracting and converting the text present the black board into Braille characters for helping visually impaired students for note making.

ACKNOWLEDGEMENT

We take this humble opportunity to thank Dean, MIT Campus, Anna University, **Dr. J. Prakash** for providing all the lab facilities in pursuit of this project.

Undertaking this project has helped us learn a lot, and we would like to express our sincere gratitude towards our supervisor **Dr. Jayashree P**, Head, Department of Computer Technology, MIT, Anna University, whose kind guidance and proper directions helped shape this project perfectly. The timely feedback from the supervisor was much instrumental in successful completion of the project work carried out. Constant encouragement to implement the project to the fullest of our potential led us through, during the difficult phase of the project, and assisted in the successful completion of the proposed project work.

We acknowledge the efforts and feedback of the panel members **Dr. Jayashree P**, **Mr. S. Yoganand**, **Ms. R. Keerthana**, in reviewing our work, providing constant valuable comments and encouraging us to view the different aspects of the project in successful implementation of the project.

We thank all the teaching and non-teaching members of the Department of Computer Technology, MIT, Anna University, seniors, juniors and friends whose appreciable help, either directly or indirectly helped in the smooth completion of the Creative and Innovative Project in a limited time frame and made the demanding time joyful but always efficient.

S Yogeeswar

M Vimal Raaj

S Kavinkumar

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 MACHINE LEARNING

Machine learning (ML) is a field of inquiry devoted to understanding and building methods that 'learn', that is, methods that leverage data to improve performance on some set of tasks. It is seen as a part of artificial intelligence. Machine learning algorithms build a model based on sample data, known as training data, in order to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of applications, such as in medicine, email filtering, speech recognition, and computer vision, where it is difficult or unfeasible to develop conventional algorithms to perform the needed tasks. A subset of machine learning is closely related to computational statistics, which focuses on making predictions using computers, but not all machine learning is statistical learning. The study of mathematical optimization delivers methods, theory and application domains to the field of machine learning. Data mining is a related field of study, focusing on exploratory data analysis through unsupervised learning. Some implementations of machine learning use data and neural networks in a way that mimics the working of a biological brain.

1.2 DEEP LEARNING

Deep learning also known as deep structured learning is part of a broader family of machine learning methods based on artificial neural networks with representation learning. Learning can be supervised, semi-supervised or unsupervised. Deep-

learning architectures such as deep neural networks, deep belief networks, deep reinforcement learning, recurrent neural networks and convolutional neural networks have been applied to fields including computer vision, speech recognition, natural language processing, machine translation, bioinformatics, drug design, medical image analysis, climate science, material inspection and board game programs, where they have produced results comparable to and in some cases surpassing human expert performance.

1.3 CONVOLUTIONAL NEURAL NETWORKS

In deep learning, a convolutional neural network (CNN) or ConvNet is a class of artificial neural network (ANN), most commonly applied to analyze visual imagery. CNNs are regularized versions of multilayer perceptrons. Multilayer perceptrons usually mean fully connected networks, that is, each neuron in one layer is connected to all neurons in the next layer. The full connectivity of these networks makes them prone to overfitting data. Typical ways of regularization, or preventing overfitting, include penalizing parameters during training or trimming connectivity. CNNs take a different approach towards regularization, they take advantage of the hierarchical pattern in data and assemble patterns of increasing complexity using smaller and simpler patterns embossed in their filters. Therefore, on a scale of connectivity and complexity, CNNs are on the lower extreme.

1.4 BRAILLE

Braille is a tactile writing system used by people who are visually impaired, including people who are blind, deafblind or who have low vision. It can be read either on embossed paper or by using refreshable braille displays that connect to

computers and smartphone devices. Braille can be written using a slate and stylus, a braille writer, an electronic braille notetaker or with the use of a computer connected to a braille embosser. Braille characters are formed using a combination of six raised dots arranged in a 3 × 2 matrix, called the braille cell. The number and arrangement of these dots distinguishes one character from another. Since the various braille alphabets originated as transcription codes for printed writing, the mappings (sets of character designations) vary from language to language, and even within one; in English Braille there are 3 levels of braille,

1. Uncontracted braille a letter-by-letter transcription used for basic literacy
2. contracted braille an addition of abbreviations and contractions used as a space-saving mechanism
3. Grade 3 various non-standardized personal stenography that is less commonly used.

1.5 OBJECTIVE

The main objective of the project is to help visually or hearing impaired students in note taking for widely used English language, and one other Indian language this is done through,

1. Capture images from black board and audio of lecturer with the help of camera and microphone connected through raspberry Pi.
2. Process captured image and segment characters and recognizes the characters thereby extracting the text in the blackboard image for two languages English and Tamil.
3. Convert the audio captured into text for two languages English and Tamil.
4. Convert the text to Braille characters and store it in a file.

CHAPTER 2

LITERATURE SURVEY

Kaliappan, A. V. and et al. in [9] have designed the combination of convolution neural network (CNN) and weighted contribution of feature point-based methods,Whichare used to recognize the Tamil alphabet in this paper. Also various CNN configuration and feature point methods BRISK,ORB,KAZE are put into test.In the feature point based method K-nearest neighbor, decisiontree, regression and multi-layer perceptron classifiers are used. So the best point based configuration and CNN configuration is combined in this hybrid approach. The KAZE MLP is found to be the best feature point based classifier and CNN 15 is the best classifier.By getting 95.6% accuracy the hybrid model performed better than CNN and Point based methods.k means clustering is the algorithm used In this approach, evaluation has focused on 12 classes of Tamil vowels only.

Gnanasivam P. and et al. in[7] proposed a convolutional neural network model is used for converting digitized handwritten tamil words to readable format. Initially the input image needs to be preprocessed. Preprocess includes image resizing, formatting etc.Then using filters of convolutional layers, the features such as corners, edges are extracted. It uses handwritten data sets to train the module. Also using the Google cloud translation API the audio is created from the regnoised characters by the module. Algorithm used is CNN.Only certain Tamil characters are used to train the model so not all Tamil characters are recognized.

Hossain, M. T. and et al. in [8] recognized the Bangla handwritten words. This

paper proposes a method in which first the images are preprocessed and characters are isolated by finding the “matra” and then segmenting each character from it. 3 zones are created. The upper zone is the one above the “matra” and the lower zone is created by making pixels white to a certain thickness beyond the minimum character height below the upper zone. after these preprocessing a temporary decision about each character is made and final decision is made by combining all 3 zones.Results show that there is an 82% accuracy using CNN algorithm. Accuracy for consonants are less the consonants without the vowel modifiers also show better results than the consonants with the vowel modifiers Main reason for this is to managing the vowel modifier containing consonants with separate logic.

Kowsalyaand et al. in[10] focused to increase the efficiency of Tamil character recognition. Various image processing techniques are applied in order to finally get each character separately and then the character is recognized using a modified neural network and weights are updated using elephant herding optimization.On the basis of these metrics the proposed method is evaluated and the results were also compared to existing classifiers. The comparative results showed that the proposed MNN method provides better accuracy results than classifiers.Algorithm used is the Modified Neural Network using elephant herding optimization.Accuracy is less when compared to other existing works for Tamil character recognition.Not all Tamil characters are recognized with this model as only few are used to train the model.

Chung, J.and et al. in[3] proposed a less computationally expensive framework for recognizing the full page handwritten text. Text localization is done first by identifying to find out the handwritten passage and then line segmentation-conversion of words to lines using clustering algorithm. Next Text recognition

takes place which also has 2 stages first Handwriting recognition which is done using CNN and bi-LSTM, CNN is useful for extracting features from images and downsampled images. These features are given as input to the bidirectional LSTM and it generates the output then Language modeling the noisy input is given,it denoises and produces the candidate strings(using beam search algorithm).This methodology uses less time and memory and it has CER of 8.50. Algorithm used is the combination of Convolutional Neural Network and Long Short Term Memory.Several short words, typically less than three characters, are not detected properly. Only the English language is considered.

Dixit, U. D.and et al. in[5] focus to achieve the accurate word image recognition for handwritten word using SVM and K-NN classification methods. Each word in the input is segmented into characters of images. Histogram of Oriented Gradients (HOG) features of image character dataset is used to train the classifiers. The method used in this paper consist of pre-processing where the noise in the image is removed using median equation. Followed by segmentation of the word happens then extraction of feature for segmented characters is done using HOG. Using the obtained features classification is done by K-NN (k=1) and SVM algorithm. Finally combining all the identified characters the word is recognized.Using the HOG features combined with K-NN or SVM classifiers, SVM gives a good recognition rate of 75%.Algorithm used is the Histogram of Oriented Gradients (HOG),K-NN,SVM.The recognition approach is done only for words and not for the entire document. Also the recognition rate is less.

Ayyadevaraand et al. in [1] recognized the handwritten character is recognized using a combination of three different feature extraction techniques and MLP BP the MLP neural network using LevenbergMarquardt algorithm and the CNN.The

three feature extraction methods used here are gradient, geometric, zone based hybrid feature extraction. These techniques are applied individually as well as in the combination form. The results obtained implies that the feature extraction technique will affect the accuracy of character recognition. The combined feature extraction technique gives the best accuracy and CNN gives best accuracy rate compared to the other neural networks such as MLP BP and MLP LM.

Prakash and et al. in [16]usedthe Convolutional Neural Network for the classification of image and text can be extracted by using Tesseract.The CNN is used to overcome the overfitting problem.The character pattern recognition and line recognition can be done by using the LSTM.The CNN gives better accuracy compared to the SVM.

Mishra, P.and et al. in [12]In this paper,the deep Convolutional Neural Network architecture is proposed to recognise the isolated tamil characters in offline mode.The proposed network contains 2 convolutional layer followed by Max pooling layer and 2 fully connected layer followed by an output layer.124 (unique symbols) classes considered for classification. Achieved training accuracy is 88.2% and the achieved testing accuracy is 71.1%.The main disadvantage found was that highly similar symbols were misclassified regularly that leads to the reduction of testing accuracy.

Goria.and et al. in [6]recognized the handwritten text by using the OCR and convergence of CNN and RNN.The Handwritten Text Recognition divided into two parts: 1. Character Recognition 2.WordRecognition. Input image consists of a sentence that will have several words. Torecognise a particular word, you need to identify each character in the word. The letters will be used to predict the word. To

implement the handwritten character recognition, two kinds of Neural Network model is used: 1. 5 layer of Convolutional Neural Network (CNN)- to extract features from the image . 2. 2 layers of Recurrent Neural Network(RNN) – to determine the word .CER and WER acts as performance metrics to judge the accuracy of the model.

Adrian MLOISE and et al. in [13] developed a system which takes the computer written text as input and outputs the converted braille to a device which is used by blind to read embedded text.Where a FSM finite state machine is implemented using a software approach.The finite state machines states are considered with corresponding output. Which is then implemented in a software approach as code.The system is divided into two units hardware and software. In which the hardware contains a microcontroller and a usb connected to the computer. When a key is pressed in the keyboard its ascii value is stored in the buffer and its sent to the finite state machine which sends the signal to the 6 needles that are touched by the blink to read.The software will control the 6 pins activation according the finite state machine result and the code build in by the concepts mentioned above.The system is tested against various inputs from alphabets to special characters, which gave a braille result with a delay of 2 sec between each characters.

Mainkar and et al. in [11] implemented an android application which captures an image of the hand written text and in return an editable text document as an output. Main objective is to bring the conversion in offline and mobile application.This is done by 5 steps which are image acquisition, preprocessing,segmentation,feature extraction and postprocessing.In the image acquisition step the image is captured using the mobile application camera and in a gray scale.Its followed by the preprocessing on the gray scale image that are

binarization,skewing,normalization and thinning.Then the segmentation of the each individual characters is done from the sentence and then in words. After the segmentation the OCR does the feature extraction where the characters are recognized according to their slant,height and curve.The classified characters are then written to a text document and are stored in the mobile application.This stored application can be edited latter also can be viewed. It gives a significant result of the scanned image as a text.

Deepa and et al. in [4] used the Convolutional Neural Network for the classification of image and text can be extracted by using Tesseract.The CNN is used to overcome the overfitting problem.The character pattern recognition and line recognition can be done by using the LSTM.The CNN gives better accuracy compared to the SVM.

Sarika and et al. in [18] gives the description of OCR and the CNN layers ,architecture and its implementation. OCR is used to convert the text in an image to text format .OCR has different phases like digitization,pre processing,segmentation,feature extraction.CNN is used for recognizing the handwritten telugu characters.CNN layers are convolutional layer ,pooling layer,ReLu layer and fully connected layer.The CNN architectures discussed in this paper are LeNet,AlexNet and ZFNet. VGG-16 architecture which consists of one input layer, one output layer and other six layers is used and trained with telugu character dataset here. Survey made on several CNN techniques which are used for classification of handwritten character recognition.

Vaidya and et al.in [20] developed the system for recognising the handwritten character based on image segmentation.The convolutional neural network model is

used here.The Neural network is trained using a tensorflow and the image processing operations are carried out by open source library OpenCV.The steps involved in the process are preprocessing,conversion of gray scale ,thresholding and image segmentation.Finally the recognition of characters happens through the neural network.Limitations of this paper are the recognition of cursive handwriting is not possible,this system only supports for the English language and the recognition of letters and digits are possible whereas the recognition of special symbols is not possible.

Vinjit and et al. in [21] reviewed the task involved in identification of the handwritten characters and Digitizing manually written text.It is known as Handwritten Character Recognition.The various techniques have been discussed to recognize different handwritings.The exact recognition of letters is directly depends on the nature and quality of the material.Choosing the relevant characteristic feature extraction and classification techniques is the key to get good recognition rate. This paper will help to know the methodologies used in the field of handwritten character recognition till now along with their advantages, limitations and accuracy rate. Similar Characters like ‘1’ and ‘7’, ‘4’ and ‘9’ are difficult to recognise.To overcome this larger training data set and better neural network designs are to be used.Although there are several techniques proposed but none of the methods is fully perfect and still improvement have to be done in the system.

The related works were surveyed and the limitations were identified to propose an organised work for our research, as described in the following elaborative headings. Keepings the challenges face by the previous researches, the model to be developed has been carefully designed, which would be cost effective as well.

CHAPTER 3

PROPOSED WORK

3.1 METHODOLOGY WITH ARCHITECTURE DIAGRAM

Lecturers try to communicate a lesson or a concept with the aid of blackboard but this becomes a problem for visually challenged students so conversion of text to braille can overcome their problem, similarly hearing impaired students find it difficult to lip read in classrooms so having a live caption system will be of very big use.

The proposed system uses various image processing techniques to preprocess the image and separate each character which then is feed into machine learning or deep learning models such as decision tree, SVM, K-NN, CNN to classify the character to recognize and form a complete sentence this is further then converted into a braille format file as this is helpful for visually challenged students on the other hand audio of the lecturer is captured and converted to text using text to speech API and is displayed using an application which helps students having hearing problems.

Figure 3.1 represents the architecture diagram of the proposed system, it is clear from the diagram that the black board image and lecturer's audio is captured using raspberry Pi connected with camera and microphone, these are the inputs to the remaining. Text from the image is extracted using deep learning models which is then converted into Braille characters and stored in a file which can be later

accessed, audio captures is converted into text in respective language which is then in real time is displayed in an application.

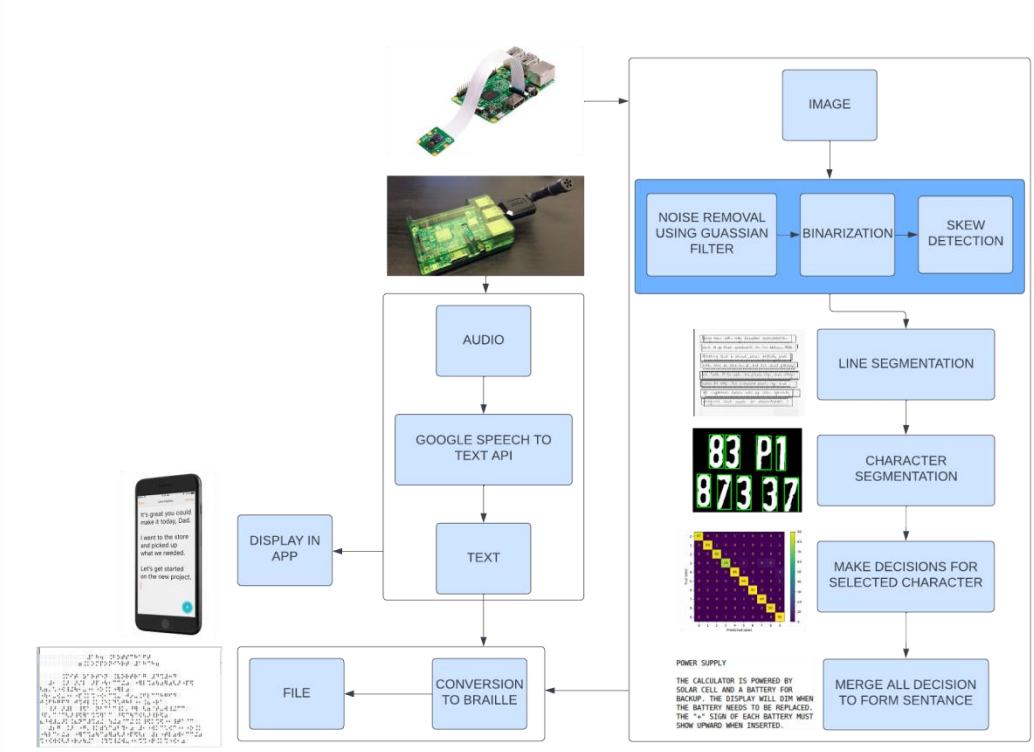

Fig. 3.1 – Architecture Diagram

3.2 MODULES

From figure 3.1 it is evident that the proposed system is divided into 4 main modules which are listed below,

I. Image processing and text recognition module
II. Speech recognition module
III. Text to Braille conversion module
IV. Application

3.2.1 IMAGE PROCESSING AND TEXT RECOGNITION MODULE

In this module text is recognized and extracted from the blackboard image, this is done by first preprocessing the image such as removing noise and unwanted details using Gaussian filter and then transforming gray scale image to black and white image using binarization technique, after this skew detection is done in our proposed system which aims to align an image before processing because text segmentation and recognition requires properly aligned next lines.

After this image preprocessing process, it is passed on to segment each line and then identify each character in each word. These characters are recognized using classification algorithms which are then put together to form an entire word.

3.2.2 SPEECH RECOGNITION MODULE

This module helps the hearing-impaired students in note taking. The audio signals produced by the lecturer are taken as an input and converted into text using an application programming interface (API) based upon the language used, this text is further sent to a created application for real time captioning and to the Braille conversion module which is helpful for the visually impaired students notes making.

3.2.3 TEXT TO BRAILLE CONVERSION MODULE

This module takes the input from speech recognition module or image processing and text recognition module outputs, which are the text documents of the converted speech or image processed. This text is converted into Braille text and gives a text

document as an output. This can be used by the visually impaired students for reading the notes.

3.2.4 APPLICATION

The application is used as an interface for the hearing-impaired students for real time captioning of the lecturers' lessons. This takes the input from the speech recognition module and the output of this module is the live recognized text. Links for the Braille format file which is generated will also be provided which can be further used by visually impaired students.

CHAPTER 4

IMPLEMENTATION

4.1 TOOLS REQUIRED

4.1.1 Raspberry Pi

The Raspberry Pi is a very cheap computer that runs Linux, but it also provides a set of GPIO (general purpose input/output) pins, allowing it to control electronic components for physical computing and explore the Internet of Things. The Raspberry Pi operates in the open-source ecosystem its main supported operating system, Pi OS, is open source and runs a suite of open-source software.

4.1.2 OpenCV

OpenCV is the huge open-source library for computer vision, machine learning, and image processing, it can be used for processing images and videos to identify objects, faces, or even handwriting of a human. When integrated with various libraries, such as NumPy, it is capable of processing the OpenCV array structure for analysis to Identify image patterns and its various features.

4.1.3 Tensorflow

TensorFlow is a free and open-source software library for machine learning and artificial intelligence. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks.

4.1.4 ScikitLearn

Scikit-learn also known as Sklearn is the most useful and robust library for machine learning in Python. It provides a selection of efficient tools for machine learning and statistical modeling including classification, regression, clustering and dimensionality reduction via a consistent interface in Python.

4.1.5 Google Colab

Colab notebooks are Jupyter notebooks that run in the cloud and are highly integrated with Google Drive, making them easy to set up, access, and share. notebooks that are created can be simultaneously edited by other team members. Colab allows to write and execute arbitrary python code through the browser, and is especially well suited to machine learning, data analysis.

4.2 IMAGE PROCESSING AND TEXT RECOGNITION MODULE

This module performs two main functionality and are performed in below order one after another,

1. Image processing
2. Character recognition

Final output of this module is a string which is the text extracted from black board image given as input to this module. The below algorithm explains how the text is extracted from the blackboard image given as an input. First characters are extracted from input image which is then passed to CNN models to predict which are then combined to form entire sentence.

ALGORITHM:

1. Process the input image
    1.1. Remove noise from the image
    1.2. Convert the image into a binary image
    1.3. If skew present correct the angle
    1.4. Find the lines and segment it
    1.5. wehFor each of the line segment it into words
    1.6. For each of the word segment it into characters
2. Load corresponding machine learning model based on the language selected
3. Predict the character for the given character
4. Append the predicted character to a string
5. Repeat from step 2 till all the characters are predicted

4.2.1 IMAGE PROCESSING

For better character recognition, noise in the input image is removed and the image is binarized, this is the process of taking a grayscale image and converting it to black-and-white, essentially reducing the information contained within the image from 256 shades of gray to 2 colors black and white, a binary image. Morphological changes to the image are performed such as dilation of image which expands foreground pixels that is used to find the lines based on finding the contours. Each of the lines so found is cut out from the original image and is again dilated in such a way that each word segmented by finding contours, characters from these words are found in the same way. This segmented character is resized to 28x28 px and is sent to the trained model to predict the character.

4.2.2 CHARACTER RECOGNITION

With the help of two deep learning models based on convolution neural networks one to recognize English characters, digits and another to recognize Tamil characters, is used to recognize character images segmented by using image processing techniques.

Model to recognize English characters and digits is trained using Extended Modified National Institute of Standards and Technology (EMNIST) dataset which is a set of handwritten character digits in 28x28 pixel image format, train set consists about 112,800 images and test set consist about 18,800 with about 47 characters in both. The characters are labeled from 0 to 46, The model trained with this dataset is able to recognize all the characters present in table 4.1.

Table 4.1- English characters recognized

0	1	2	3	4	5	6
7	8	9	A	B	C	D
E	F	G	H	I	J	K
L	M	N	O	P	Q	R
S	T	U	V	W	X	Y
Z	a	b	d	e	f	g
h	n	q	r	t		

A CNN model is created with the architecture specified in Fig. 4.1, the model is compiled after setting a few attributes such as optimizers, batch size. Adam optimizer is used with an initial learning rate of 0.001 after each 20 epochs learning rate is reduced by 2.5 of existing learning rate, batch size of 32 was set. Different hyper parameters such as stride, padding, and depth are used at each layer which tuned to build a better model, stride value of 2 was taken. The training was done for 100 epochs, after each epoch validation is done on the validation part, after tuning the model with optimum accuracy is saved. The saved model is tested with the test set.

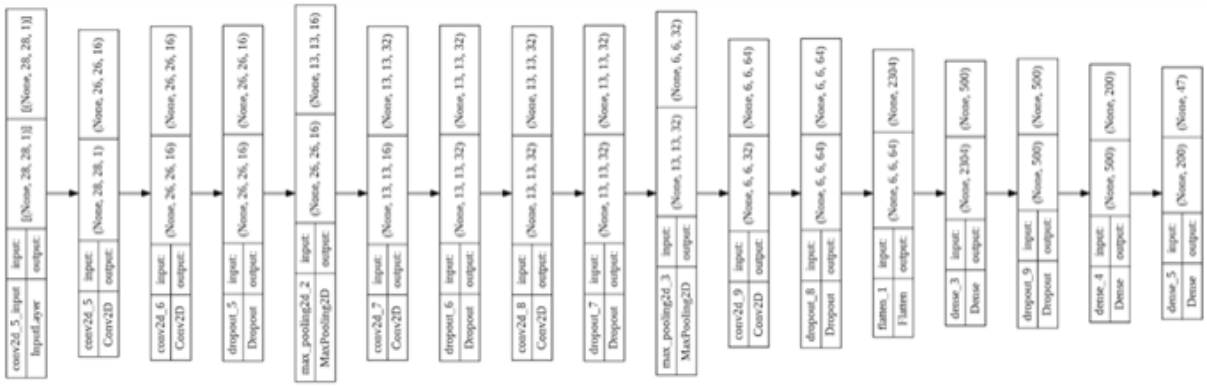

Fig. 4.1- English character recognition model architecture

Model to recognize Tamil characters is trained using hpl-tamil-iso-char-offline, the dataset contains approximately 500 isolated samples each of 156 Tamil characters written by native Tamil writers. The images are bi-level images with the background being white and foreground in black. Preprocessing included inverting the background and foreground color and then separating the images to respective character classes, the images are of varying sizes so all of them are converted into 28x28 px images. From each of the character classes 400 images are taken for training and 100 images are taken for testing. 20 percent of the training dataset is

taken as a validation dataset. CNN Model trained using these is able to recognize all the characters present in table 4.2.

Table 4.2- Tamil characters recognized

அ	ஆ	இ	ஈ	உ	ஊ	எ	ஏ	ஐ	ஒ	ஓ	ஔ
க	ங	ச	ஞ	ட	ண	த	ந	ப	ம	ய	ர
ல	வ	ழ	ள	ற	ன	ஜ	ஹ	க்ஷ	கி	ஙி	சி
ஞி	டி	ணி	தி	நி	பி	மி	யி	ரி	லி	வி	ழி
ளி	றி	னி	ஸி	ஷி	ஜி	ஹி	க்ஷி	கீ	ஙீ	சீ	ஞீ
டீ	ணீ	தீ	நீ	பீ	மீ	யீ	ரீ	லீ	வீ	ழீ	ளீ
றீ	னீ	ஸீ	ஷீ	ஜீ	ஹீ	க்ஷீ	கு	ஙு	சு	ஞு	டு
ணு	து	நு	பு	மு	யு	ரு	லு	வு	ழு	ளு	று
னு	ஸு	ஷு	ஜு	ஹு	க்ஷு	ஶ்ரீ	ஸூ	ஷூ	ஜூ	ஹூ	ஷூ
கூ	ஙூ	சூ	ஞூ	டூ	ணூ	தூ	நூ	பூ	மூ	யூ	ரூ
லூ	வூ	ழூ	ளூ	றூ	னூ	ா	ெ	ே	ை	க்	ங்
ச்	ஞ்	ட்	ண்	த்	ந்	ப்	ம்	ய்	ர்	ல்	வ்
ழ்	ள்	ற்	ன்	ஸ்	ஷ்	ஜ்	ஹ்	க்ஷ்	ஃ	ஸ	ஷ

A CNN model is created with the architecture specified in Fig. 4.2, the model is compiled after setting a few attributes such as optimizers, batch size. Adam optimizer is used with an initial learning rate of 0.001 after each 20 epochs learning rate is reduced by 2.5 of existing learning rate. The training was done for 100 epochs, after each epoch validation is done on the validation part, after tuning the model with optimum accuracy is saved. The saved model is tested with the test set which contains about 100 images from each class.

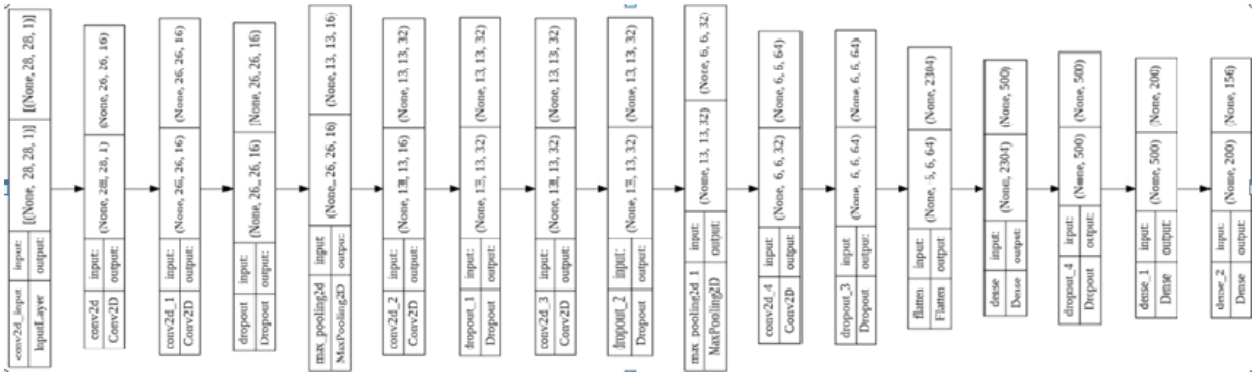

Fig. 4.2 - Tamil character recognition model architecture

4.3 SPEECH RECOGNITION MODULE

In this module we have used Google API Speech Recognition of languages. The language consists of Tamil and English. In the Speech Recognition module Microphone used as input. The output is a document containing English or Tamil recognised words. This is obtained by using Google API for both Tamil and English language. The audio is passed to the Google recognizer, which returns a recognised text which is written to a file. End condition for this program is “class over” for English and “நன்றி” for Tamil, when this condition becomes true recognition is stopped. The recognised text is sent to an application, which displays

the text in real time and converts the text into a braille conversion module for the conversion.

4.3.1 ALGORITHM:

4.2.1.1 ENGLISH:

1. Import the Speech Recognition Module.
2. Initialize the Recognizer.
3. Create a file named with Timestamp.
4. Take the microphone as Source.
5. Listen to the Source.
6. Send the audio to recognize_google
   - 6.1. Text = r.recognize_google(audio)
   - 6.2. If text equals to “class over” end the process
   - 6.3. else repeat from Step 4.
7. Write the text into the file.

4.2.1.2 TAMIL:

1. Import the Speech Recognition Module.
2. Initialize the Recognizer.
3. Create a file named with Timestamp.
4. Take the microphone as Source.
5. Listen to the Source.
6. Send the audio to recognize_google
   - 6.1. Text = r.recognize_google(audio, language=”ta-IN”)
   - 6.2. If text equals to “நன்றி” end the process
   - 6.3. else repeat from Step 4
7. Write the text into the file.

4.4 TEXT TO BRAILLE CONVERSION

The text which is obtained from the speech and text recognition module is converted to grade 2 Braille where contractions are used in addition to rules of grade 1 Braille where each character is translated to Braille, this done using the following algorithm which is implemented using python. In this module a dictionary is used for storing English characters which contains 26 alphabets, 31 contractions, 22 punctuations, 10 numbers and their corresponding Braille ASCII values. A similar dictionary for Tamil which contains 247 Tamil characters and their corresponding Braille ASCII values which is used as a Map.

ALGORITHM:

1. Split the text into words based on the whitespace and newline characters.
2. For each word, first handle the numbers by changing the numbers to Braille Unicode and also including the
3. Handle capital letters.
4. Trim the word, as words may contain punctuations and are saved as shaving.
5. Translate the words.
    5.1. Check to see if trimmed words can be contracted.
    5.2. Translate the characters that are still alphabets.
    5.3. Translate the shavings.
6. Write the returned Braille characters into the file.

4.5 APPLICATION

The mobile application is used to display text produced from the speech recognition module which helps hearing impaired students to understand lecturers' voices, and the voice notes converted to text are stored in a file for further reference and can be accessed through this application. The text from the document is read and printed on the website.

CHAPTER 5

RESULTS AND ANALYSIS

The implemented modules are tested against various scenarios, and the results have been documented. The smart notes maker that would help visually and hearing impaired students in an effective manner, and provides comfort to all students in terms of note making is capable,

1. Images from black board and audio of lecturer with the help of camera and microphone connected through raspberry Pi.
2. The captured image is processed and characters are segmented and are recognized thereby extracting the text in the blackboard image for two languages English and Tamil.

   The model created for recognizing 156 Tamil handwritten characters has a training accuracy of 0.9574, validation accuracy of 0.9305 and accuracy with test data is 0.9397. Below figures 5.1 and 5.2 shows accuracy and loss of training and validation respectively, from these it can be seen that the model is a good fit as train and validation lines almost overlap.

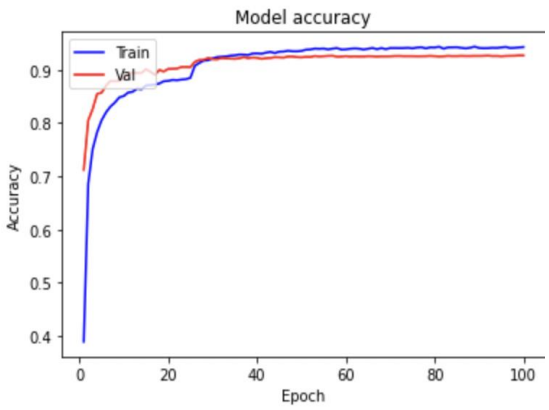

Fig. 5.1 - Model accuracy graph

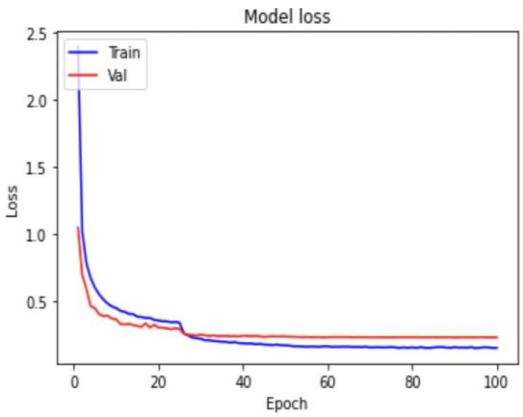

Fig. 5.2 - Model loss graph

Various machine learning models such as decision tree, support vector machine for recognizing English alphabets and digits were tried and achieved an accuracy of 0.6157 and 0.8284 but the model created using CNN shows a better training accuracy of 0.9202, validation accuracy of 0.9002 and accuracy with test data is 0.8976. Below figures 5.3 and 5.4 shows accuracy and loss of training and validation respectively.

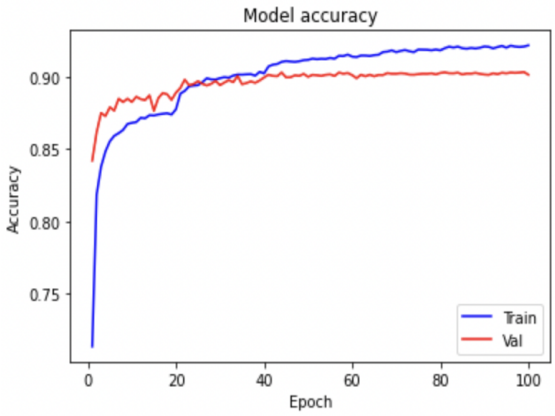

Fig. 5.3- Model accuracy graph

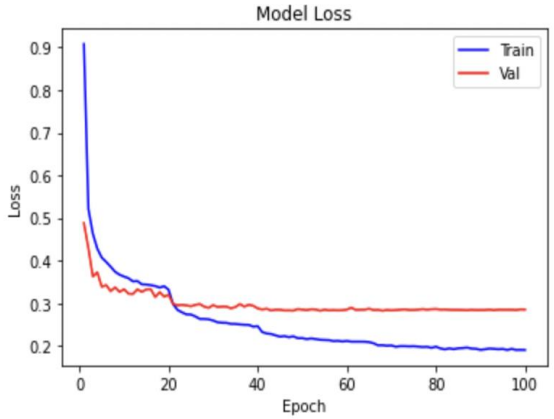

Fig. 5.4 - Model loss graph

3. Convert the audio captured into text for two languages English and Tamil. The audio captured is converted to text and is displayed in an web application for both languages English and Tamil.

Figure 5.5 shows the stored data for live caption Tamil text which was captured using microphone and converted using speech to text module, so is the Figure 5.6 for English language. These files are used to display live captions in the application, further these files are sent to text braille conversion module to get the braille files.

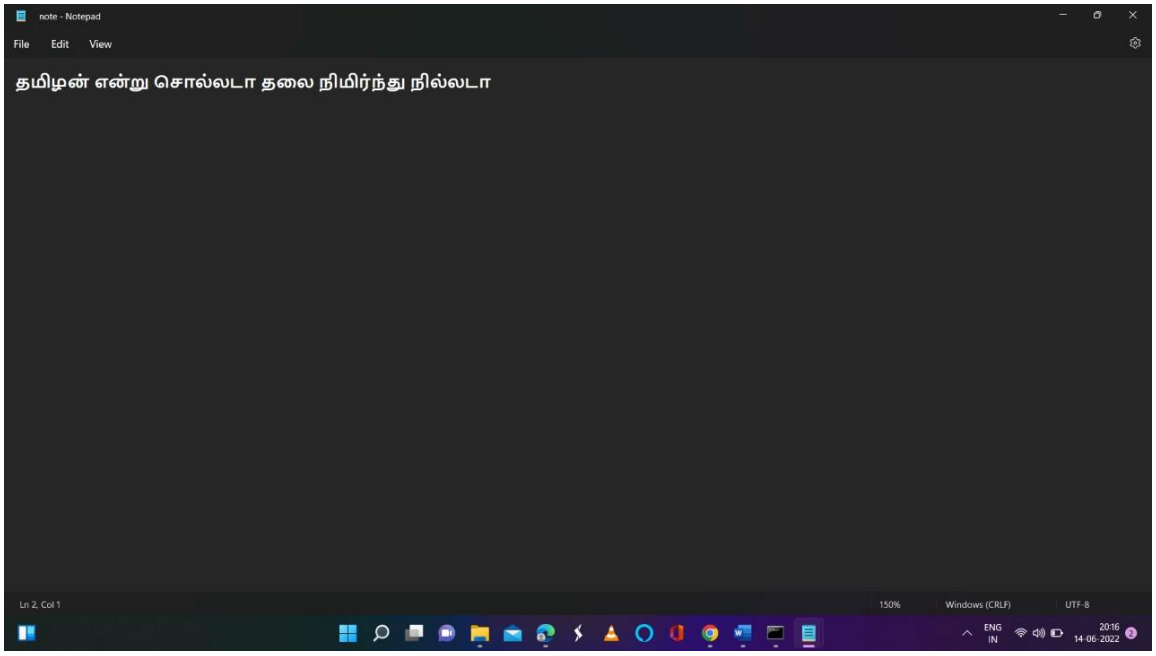

Fig. 5.5 - Tamil live caption stored data

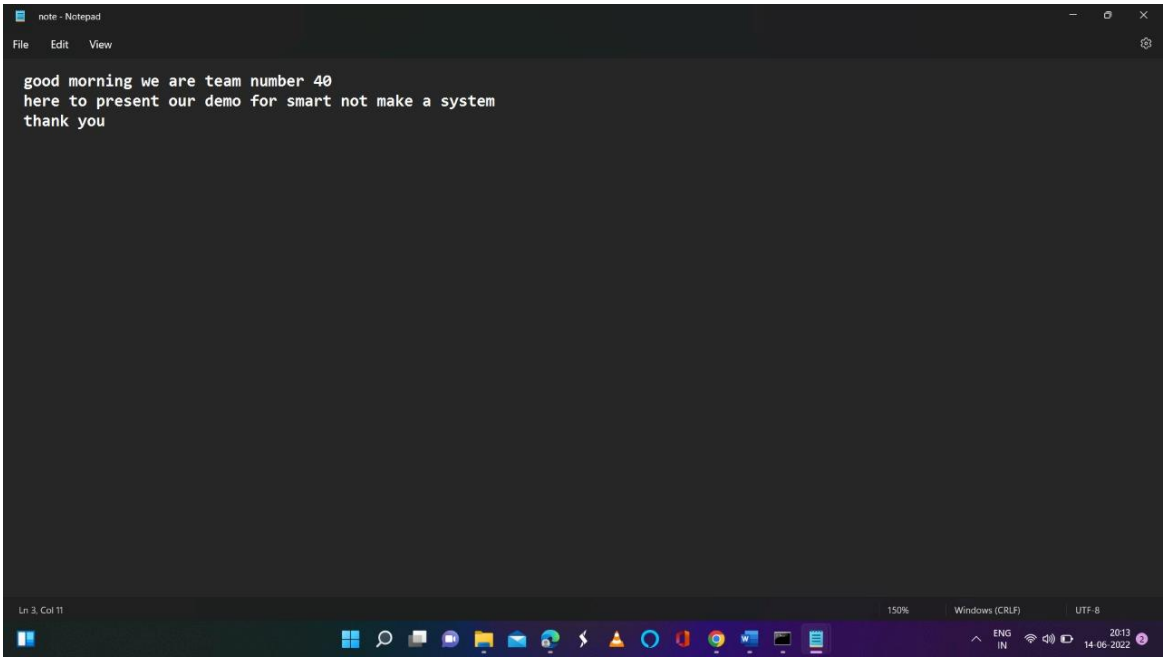

Fig. 5.6 - English live caption stored data

4. Text is converted into Braille characters and is stored in a file using the text to Braille conversion module, which can be accessed by the students in future.

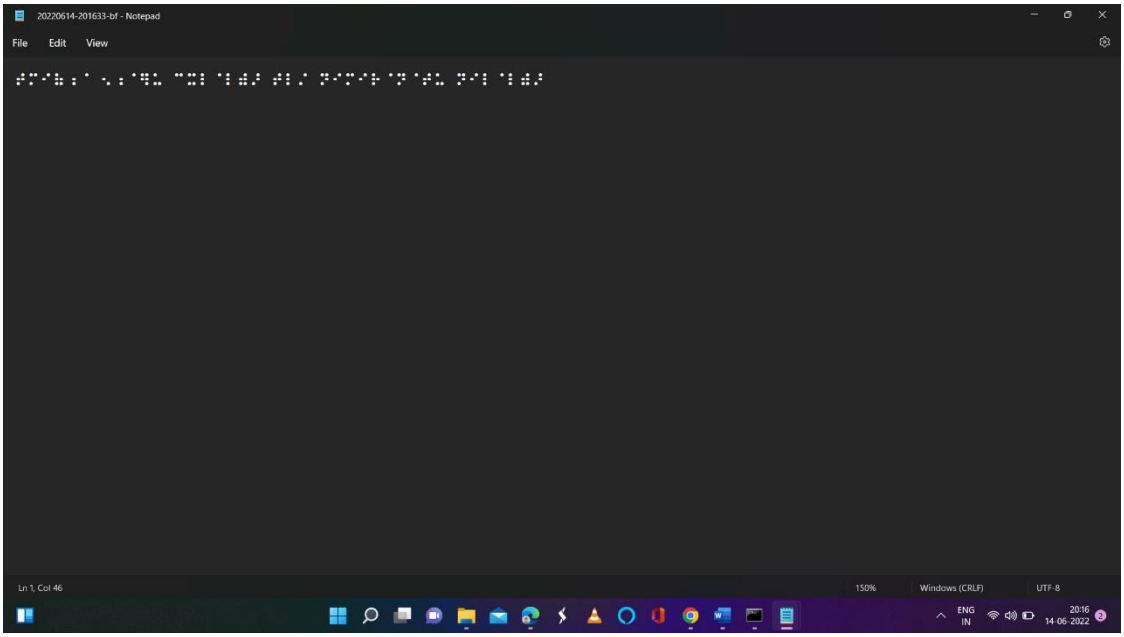

Fig. 5.7 - Tamil Braille converted data

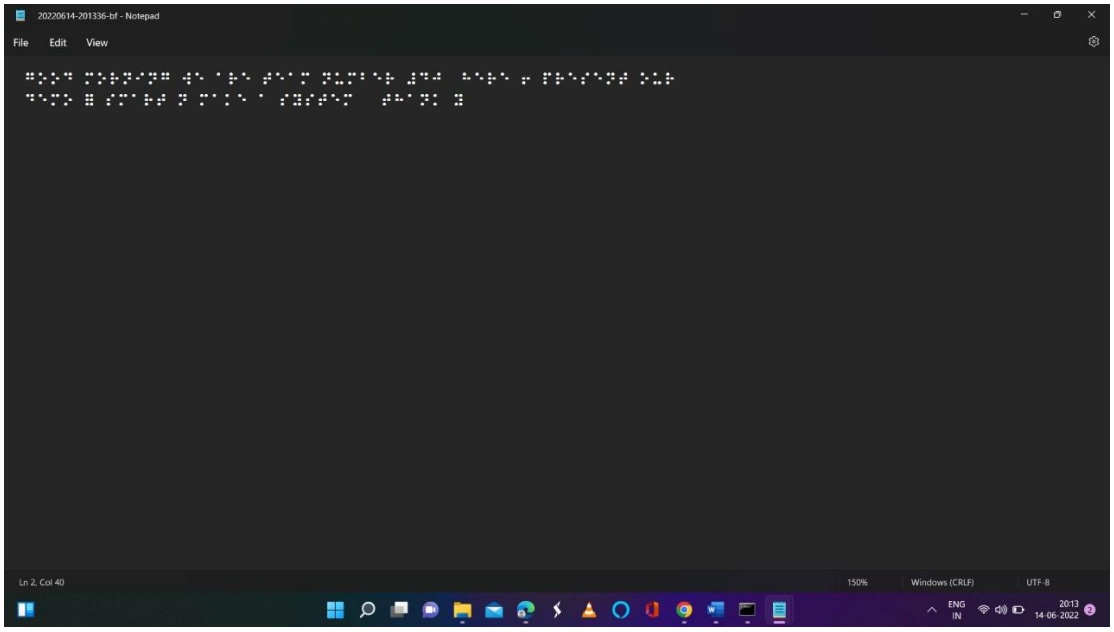

Fig. 5.8 - English Braille converted data

Figures 5.7 and 5.8 shows the file which is created using the Braille conversion module for Tamil and English language respectively. These files can be used for printing or can be given as an input to refreshable braille devices which is used for reading Braille characters.

5. Application for viewing live captioning is created which is represented by figures 5.9. This application is compatible for both Tamil and English language.

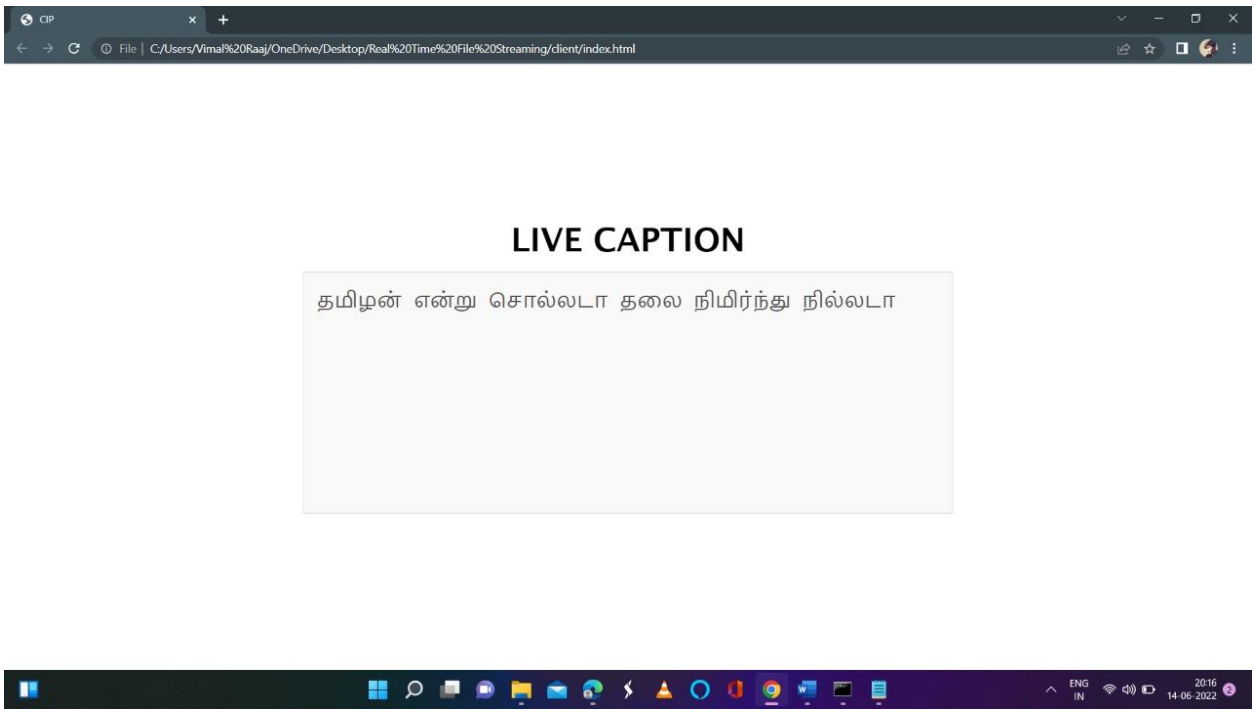

fig. 5.9 – application for live caption Tamil

CHAPTER 6

FUTURE WORKS

Some of the future works would be as follows,

1. Create a single machine learning or deep neural network which can extract text from images of most of the languages widely used and to create the same for recognizing speech.
2. Making the model to learn mathematical and scientific symbols will enable the note maker to be more utilitarian.
3. Introduce a way to convert images drawn on the black board in a useful way for visually impaired students.
4. Create methods such that the important or summary of the lecture is only stored instead of the entire lecture.

CHAPTER 7

CONCLUSION

Moving towards an equal opportunity for education for all, researches prove that visually and hearing impaired students find it difficult to take notes. In order to remove such problems a system is designed which can be incorporated in classrooms so that the lectures communicated to the students are understood without any barrier. The implemented modules would aid the visually impaired people, and further development of the modules described would enable automated note taking via live captioning that would prove helpful for students in order to understand concepts and focus more during lectures, as note making cannot be easily achieved by visually or hearing impaired students the product created would help by providing a live captioning system which converts lecturers' audio into text and is displayed in an android application for hearing impaired students and also converts the content written in blackboard or projected in a screen into a braille format file which can later be used for printing which would help visually impaired students.

REFERENCES

[1] Ayyadevara, S. A., Teja, P. N. V. S. R., & Rajesh Kumar, M. (2018). *Handwritten Character Recognition Using Unique Feature Extraction Technique. 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT).* pp.1239-1243.

[2] Al-Mahmud, AsnuvaTanvin, SaziaRahman.(2021). *Handwritten English Character and Digit Recognition.International Conference on Electronics, Communications and Information Technology (ICECIT)*, pp. 1-4

[3] Chung, J., &Delteil, T. (2019). *A Computationally Efficient Pipeline Approach to Full Page Offline Handwritten Text Recognition.International Conference on Document Analysis and Recognition Workshops (ICDARW).* doi:10.1109/icdarw.2019.40078, pp. 35-40.

[4] Deepa, R., &Lalwani, K. N. (2019). *Image Classification and Text Extraction using Machine Learning. 3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA).* doi:10.1109/iceca.2019.8821936, pp. 680-684.

[5] Dixit, U. D., Hiraskar, R., Purohit, R., &Shivanagutti, S. (2020). *Recognition Of Handwritten Word Images. IEEE Bangalore Humanitarian Technology Conference (B-HTC).* doi:10.1109/b-htc50970.2020.9297853.

[6] Gorai, M., & Nene, M. J. (2020). *Layout and Text Extraction from Document Images using Neural Networks.5th International Conference on Communication and Electronics Systems (ICCES).* doi:10.1109/icces48766.2020.9137859, pp.1107-1112.

[7] Gnanasivam P., Bharath G., Karthikeyan V., and Dhivya V. (2021). *Handwritten Tamil Character Recognition using Convolutional Neural*

Network. -Sixth International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET). doi:10.1109/wispnet51692.2021.9419451, pp.84-88.

[8] Hossain, M. T., Hasan, M. W., & Das, A. K. (2021). *Bangla Handwritten Word Recognition System Using Convolutional Neural Network. 15th International Conference on Ubiquitous Information Management and Communication (IMCOM).* doi:10.1109/imcom51814.2021.9377410, pp. 1-8.

[9] Kaliappan, A. V., & Chapman, D. (2020). *Hybrid Classification for Handwritten Character Recognition of a Subset of the Tamil Alphabet.-6th IEEE Congress on Information Science and Technology (CiSt).* doi:10.1109/cist49399.2021.9357190, pp.167-172.

[10] Kowsalya, S., &Periasamy, P. S. (2019). *Recognition of Tamil handwritten character using modified neural network with aid of elephant herding optimization. Multimedia Tools and Applications.* doi:10.1007/s11042-019-7624-2, pp. 25043-25061.

[11] Mainkar, V. V., Katkar, J. A., Upade, A. B., &Pednekar, P. R. (2020). *Handwritten Character Recognition to Obtain Editable Text. International Conference on Electronics and Sustainable Communication Systems (ICESC).* doi:10.1109/icesc48915.2020.9155786, pp.599-602.

[12] Mishra, P., Pai, P., Patel, M., & Sonkusare, R. (2020). *Extraction of Information from Handwriting using Optical Character recognition and Neural Networks.4th International Conference on Electronics, Communication and Aerospace Technology (ICECA).* doi:10.1109/iceca49313.2020.9297418, pp.1328-1333.

[13] Mloise, A., Bucur, G., & Popescu, C. (2017). *Automatic system for text to Braille conversion. 9th International Conference on Electronics, Computers and Artificial Intelligence (ECAI).* doi:10.1109/ecai.2017.8166391, pp. 1-6

[14] NagulUlaganathan, Rohith J, Sri Aravind S, Abhinav A S, Vijayakumar V, Ramanathan. (2020). *Isolated Handwritten Tamil Character Recognition using Convolutional Neural Networks. 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS).* doi: 10.1109/ICISS49785.2020.9315945, pp.383-390.

[15] M. A. Pragathi,K. Priyadarshini, S. Saveetha, A. Shavar Banu, K. O. Mohammed Aarif.(2019).*Handwritten Tamil Character Recognition UsingDeep Learning.International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN)*.doi: 10.1109/ViTECoN.2019.8899614, pp. 1-5

[16] Prakash, A. A., & Preethi, S. (2018). *Isolated Offline Tamil Handwritten Character Recognition Using Deep Convolutional Neural Network. International Conference on Intelligent Computing and Communication for Smart World (I2C2SW),* pp. 278-281

[17] Rhea Sawant, Prabhav, Pranjal Shrivastava, PritiShahane, R Harikrishnan.(2021).*Text to Braille Conversion System. International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES),* pp. 1-5.

[18] Sarika, N., Sirisala, N., &Velpuru, M. S. (2021). *CNN based Optical Character Recognition and Applications. 6th International Conference on Inventive Computation Technologies (ICICT),* pp.666-672.

[19] TasleemKausar, Sajjad Manzoor, AdeebaKausar, YunLu, MuhammadWasif. (2021). *Deep Learning Strategy for Braille Character Recognition*. Vol. *9,* pp.169357-169371.

[20] Vaidya, Rohan; Trivedi, Darshan,Satra, Sagar,Pimpale, Prof. Mrunalini (2018). *Handwritten Character Recognition Using Deep-Learning. IEEE Second International Conference on Inventive Communication and Computational Technologies (ICICCT), pp.772-775*.

[21] Vinjit, B. M., Bhojak, M. K., Kumar, S., &Chalak, G. (2020). *A Review on Handwritten Character Recognition Methods and Techniques. International Conference on Communication and Signal Processing (ICCSP)*, pp.1224-1228.