Deep Learning based Bangla Voice to Braille Character Conversion System.

by

Syed Bayes Iqbal
18101198
Riazul Islam Rifat
18101236
Md Akif Hussain
18101073
Simon Biswas
18101543

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science and Engineering

Department of Computer Science and Engineering
Brac University
May 2022

© 2022. Brac University
All rights reserved.

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Syed Bayes Iqbal 18101198	Riazul Islam Rifat 18101236
Md Akif Hussain 18101073	Simon Biswas 18101543

Approval

The thesis/project titled “Deep Learning based Bangla Voice to Braille Character Conversion System. ” submitted by

1. Syed Bayes Iqbal (18101198)

2. Riazul Islam Rifat (18101236)

3. Md Akif Hussain (18101073)

4. Simon Biswas (18101543)

Of Spring, 2022 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science and Engineering on May 20, 2022.

Examining Committee:

Supervisor:
(Member)

Md. Golam Rabiul Alam, PhD
Associate Professor
Department of Computer Science and Engineering
Brac University

Co-Supervisor:
(Member)

Md. Tanzim Reza
Lecturer
Department of Computer Science and Engineering
Brac University

Thesis Coordinator:
(Member)

Md. Golam Rabiul Alam, PhD
Associate Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi, PhD
Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Ethics Statement

Hereby, we, the members, consciously assure that the following is fulfilled for the manuscript, "Deep Learning based Bangla Voice to Braille Character Conversion System."

1. The contents of this paper are unique work of the writers' and it has not been published before.

2. Only the authors' own research and analysis is presented in the work with utmost accuracy.

3. Contributions by the co-authors and co-researchers are rightly acknowledged in the study.

4. All sources are appropriately mentioned (correct citation). Text that is literally copied must be identified as such by using quote marks and providing suitable reference.

5. The authors actively participated and put in effort leading to the article and any public responsibility related to it's content will be accepted.

Violations of the Ethical Statement standards may have serious repercussions. We agree with the preceding declarations and certify that this submission adheres to BRAC University's rules.

Abstract

From the very beginning of the 20th century, there has been a significant development in the education system using modern technology. However, this rapid development has very little scope to facilitate the education system for the people without sight. Though braille is a method for sight disabled people it has not been incorporated that much with the modern education system. On the other hand, many papers and works have been done on it but a maximum of these is about the conversion of a text character to a braille pattern with considerable accuracy so far no work has been done that can work with voice to convert a character into braille pattern for Bangla language. This paper introduces a development that will help the visually impaired people to start their very basic education with a technology that is capable of recognizing a Bangla character uttered by a person and can convert that character into braille pattern so that using a finger touch one can recognize the corresponding braille for that character. This voice recognition will be done using the Visual Geometry Group (VGG-16) model of Convolutional Neural Network (CNN) and Arduino Uno, LCD 16x4 display will be used to generate the braille pattern for the recognized character. The model is trained with 50 epochs and we achieved 98% of train accuracy and after evaluating test data we gained 92.42% test accuracy.

Keywords: Convolutional neural network(CNN), Braille, Spectrogram, Classification, VGG-16, Arduino Uno, LCD 16x4 display

Acknowledgement

To begin, glory be to Allah, who ensured that our thesis was completed on schedule and without serious setbacks.

We would like to offer our heartfelt appreciation to Md. Golam Rabiul Alam sir and Md Tanzim Reza sir, our distinguished mentors and advisors for their advice and assistance during our work.

Finally, we'd want to express our gratitude to our most deserving parents. We might not have been able to accomplish this work without their relentless support. We are approaching graduating as a result of their goodwill and efforts.

Table of Contents

List of Figures

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

α Alpha

β Beta

ϵ Epsilon

$ASCII$ American Standard Code For Information Interchange

DFA Deterministic Finite Automata

$EBCS$ Expert Braille Communicating System

GDP Gross domestic product

GSM Global System for Mobile communication

GTD Graphic tactile display

GUI Graphical user interface

HOG Histogram of Oriented Gradients

KNN K-Nearest Neighbor

LCE Later liquid crystalline elastomer

ML Machine Learning

RBD Refreshable braille display

SVC Support Vector Classifier

SVM Support Vector Machine

USD United States dollar

Chapter 1

Introduction

Bangla is one of the most spoken languages in the world, more than 300 million people use Bangla as their first language. And Bangla is the 2nd most spoken language in the Indian subcontinent. This language is recognized as the International Mother Language Day by UNESCO to commemorate the sacrifice made by the Bengali people in 1952 during the movement to establish Bangla as the State language. Bangla is the mother tongue for the Bangladeshi people and also the people of Kolkata from India use Bangla as their first language. Hence, here Bangla language plays a very significant role in the academic purpose. Though it's easy to start learning Bangla characters from an early age for physically stable people, it's a very challenging matter for visually impaired people. And here comes the braille system for visually impaired people to overcome the challenges of learning.

Braille is a unique reading/writing system for visually impaired people. It was introduced by Louis Braille in 1829. Before that Charles Barbier introduced the term of night writing system with the help of 12 raised dots. From that 12 raised dots of the night writing system, Louis Braille reduced it to 6 raised dots so that one can feel a letter by 1 finger touch. And this reading/writing system using 6 raised dots is known as the Braille System nowadays.

A single braille cell contains 6 raised/flat dots, 3 dots in right and 3 in the left. A single braille cell can form 64 different patterns (including empty patterns). Each dot pattern represents a letter, a combination of letters, numbers, or any symbols. And dot size, cell spacing, cell size are uniform throughout. Sometimes numeric numbers and alphabets may have the same cell pattern. A user can differentiate between similar numbers and alphabet patterns using two different patterns named 'Number Follows' and 'Capital follows'. A different language has different types of braille patterns. For different Bangla basic characters, there are in total 50 braille patterns.

However, in the modern education system braille is not incorporated that much as a result visually impaired people are lagging behind at the very early phase of their academic life. Recently voice to text conversion using different types of Machine Learning algorithms has become very popular. Here voice to braille character pattern conversion using deep learning like voice to text conversion will be a great opportunity for visually impaired people to overcome their challenges in the case of character learning at the early phase of education.

1.1 Research Problem

Braille is a reading and writing technique that is commonly used by blind people. Due to a scarcity of textbooks and other reading materials written in Bangla Braille, visually handicapped people in Bangladesh are denied access to education and contemporary technology. This contributes to the deepening of the knowledge divide and disparities in society. To make a significant difference in this situation, the model that we are going to develop is for the visually impaired people by which they can start their initial education on braille patterns. This model is going to convert the Bangla character from the recorded audio to its corresponding braille pattern. This model will focus on braille education. The main goal of this model is how a visually impaired person can learn braille patterns. For our work first, we will collect the dataset which consists of recorded audio of all Bangla alphabets, then using voice recognition the recorded Bangla characters will be identified. This voice recognition will be done using the Visual Geometry Group (Vgg16) model of Convolutional Neural Network (CNN) and Arduino Uno, LCD 16x4 display will be used to generate the braille pattern for the recognized character.

1.2 Research Objectives

This research aims for the development for the visually impaired people to start their basic education with a technology which will be capable to recognize a Bangla character uttered by a person and after that, the character can be converted into braille pattern so that using the touch of fingers a visually impaired person can understand the corresponding braille pattern of that character. The objectives of this research are as follows:

1. Understand how the braille pattern is helping visually impaired people.
2. Develop a technology on braille education for visually impaired people.
3. To evaluate the technology developed for visually impaired people.
4. To help the visually impaired people for initiating their primary education.
5. To offer recommendations on improving the model.

Chapter 2

Literature Review

Research works have proposed the conversion of handwritten Bangla characters into braille models in real time using image processing and machine learning [1]. This work is done in two parts. First, an image of a handwritten Bangla character is captured, then image processing concepts using OpenCV and machine learning are applied to capture and process the image, and then recognize the character. The recognized character is then transmitted to a device consisting of Atmega2560 chipset (Arduino Mega), servo motor and LCD screen which will process the data and finally display it in braille in real time. using a servo motor and will display the corresponding character on the LCD. This work is done with accuracy 93-94%.

The research work [2] proposed a method in which it coverts the scanned Hindi Handwritten text and converts it into printable text. This conversion is done by utilizing HOG features with SVM classifier. Afterwards the printable text in mapped to Braille by making use of UTF-8 code. In Hindi, the letters in a word are connected by an overhead line. To segment the word into characters it is required to remove the overhead line. This is achieved by adding pixels row-wise. After that, the segmented characters are then recognized using the SVM classifier and converted into a binary image with white text and black background. The classifier is trained on Hindi consonants and the test is done on printed and handwritten words not having matras (modifiers) and it has given approximately 94.5% accuracy which is considered very good.

The work proposed [3] a method for developing an inexpensive portable communication device designed to convert all Braille input into alphanumeric text format. It consists of 2 parts, which is a braille to text converter and a wrist vibrator. First, for short distance communication, Braille keyboard is used as input and LCD display as output respectively. Then for long distance communication, GSM module is interfaced to send message to remote receiver. A Braille to Text converter consisting of a Braille keyboard, LCD display, GSM module and ZigBee transmitter is interfaced with the PIC 16F877A. Here the vibration motor is connected to the Arduino Nano microcontroller through a relay and as soon as it is energized, the relay will form a closed circuit that makes the motor vibrate. This project is really important because it supports long distance communication.

This work [4] demonstrates a computational model employing Deterministic Finite

Automata (DFA) based on the research and analysis, which may be extended to an automated system for Bangla text to Braille translation. The term "deterministic" refers to the fact that the automaton can only transition to one state from its present state on each input. This computational model has 4 modules including Bangla Text, Rule Engine, Translator, Braille Text. In the Bangla Text module, there will be Bangla alphabets, spaces, and formatting characters. On the other hand, The Rule engine's job is to produce tokens and their related rule numbers from Bangla text. Moreover, a look-up table containing character mapping and rules is kept by the translator. The Mapping Table shows how to map Bangla and Braille characters. The number of the rule and the actions that go with it are listed in the Rules Table. Rule Engine sends a token and some rules to the Translator. Finally, the comparable Bangla Braille symbols of Bangla alphabets, numerals, punctuation, and formatting characters make up Braille text. Also, ASCII must be used for the Braille text.

This work [5] consists of a GUI, Motor Control Module, Motor Vibration Module. So, it depicts the computer's GUI as an input device that converts text into characters and then into Unicode. In the motor control module, they used Arduino Uno Rev3 and the program was saved into Arduino's memory. The motor vibration module consists of 6 motors that are connected to the 6 fingers of a person. According to the participants, a vibration time of 2 seconds was preferable. The Unicode is obtained by Arduino, and the output pins are set. It also gives feedback to the computer on the state of the vibration motors at the same time. The simulation is then displayed in the GUI. At the same time vibration occurs in the motor vibration module which indicates the braille pattern for blind people.

Paper [6] mainly works with Graphic tactile displays (GTDs) Refreshable braille displays (RBDs) discuss their kernel technologies about the actuators along with the latch structures. This paper works with the present situation of Graphic tactile displays (GTDs) Refreshable braille display (RBDs), it explores the common technologies used in GTDs predicts developing cost-effective displays in the near future. Graphic tactile displays (GTDs) were first made using piezoelectric actuators, later permanent magnet was used to make GTDs after those electroactive polymers were used to make them, finally, GTDs were made using pneumatic actuators. On the other hand, Refreshable braille displays (RBDs) were also first made using piezoelectric actuators. Later liquid crystalline elastomers (LCE) were used to make RBDs, afterwards, a dielectric elastomeric actuator was also used to make them. Even with more passage of time terpolymer liquid gas bubbles were used to make RBDs. Afterward, tiny solenoids were used to make RBDs finally RBDs were made of ferromagnetic pins. These RBDs have some design requirements, for instance, the tactile dots should be densely arranged, the dots need to have a high refresh rate, proper Braille dot latching, larger dot tactile force the device must be safe economic. 3 major actuators can be used for tactile display according to this paper those are pneumatic or fluid actuators, electro-sensitive deformable material actuators the last one is electromagnetic actuators. The mechanism of latching of the tactile display can mainly be of two types those are semi latching structure full latching structure. Among them, the full latching structure is a better choice as it has a larger dot tactile force. The performance of work of this tactile display is determined by the actuators latch structures. Some of the areas where future research can be done are- ensuring dense dot arrangement, large dot tactile force high dot refresh frequency. And ensuring all these at an economic price also is a big challenge. Moreover, making the displays durable, portable efficient in terms of power consumption is also a field that needs work.

Paper [7] mainly focuses on a method that a scaled document of Braille to a text file. This text file could be read a by the people using a computer. All three languages such as English, Hindi Tamil were considered while research was taking place. This paper suggests a method where the braille documents are preprocessed so that the dots are enhanced the noises are reduced. In this method at first, the braille cells are segmented. Then the dots of the braille cells are extracted. After segmentation and extraction, the cells are converted to number sequence and then this number sequence are mapped to the respective alphabets. Finally, a speech synthesizer speaks the text that was converted earlier. Moreover, this paper also suggests another method where a number pad keyboard types the Braille character. These typed characters are mapped to the respective alphabets spoken out. One of the limitations of this paper that can be worked on in the future is finding methods to remove dots noises during preprocessing. These dots noises are of the same size as the Braille dots which affects the accuracy of the converted text. Another major limitation that can be worked on is making the document usable even if there is a tilt while scanning.

Expert Braille Communicating System (EBCS) has been an important area of research in recent decades. Researchers have continuously tried to develop the Braille system over time. In 2008, Tirthankar Dasgupta introduced a system to translate Indian texts into Braille format, the motive behind creating such a system was to bridge the gap between a blind person and a blind people. However, this system requires a braille keyboard which is quite difficult to use for the visually impaired [8].

An automated speech recognition tool was made by Melissa Ramírez in the year 2016. This tool helped children who can't see in learning braille. Their research came up with a training algorithm which was new for a device. After testing the algorithm the device showed 89% positive approximation. This device inputs data using a microphone and gives output in a text format [9].

Chapter 3

Methodology

We have done our research work by maintaining a sequence. First, we have collected datasets. For this purpose, we have recorded audio of Bangla alphabets from different people regardless of their gender, age, etc. We have collected recorded audio of Bangla alphabets from children, young men, young women, teenage people, elder people, etc. so that we can train our machine learning algorithm properly. After collecting data, we started processing our datasets which includes audio to spectrum conversion, classification, etc. It is leading our way to testing out datasets but before that, we need to train our algorithm. After training, we tested our algorithms.

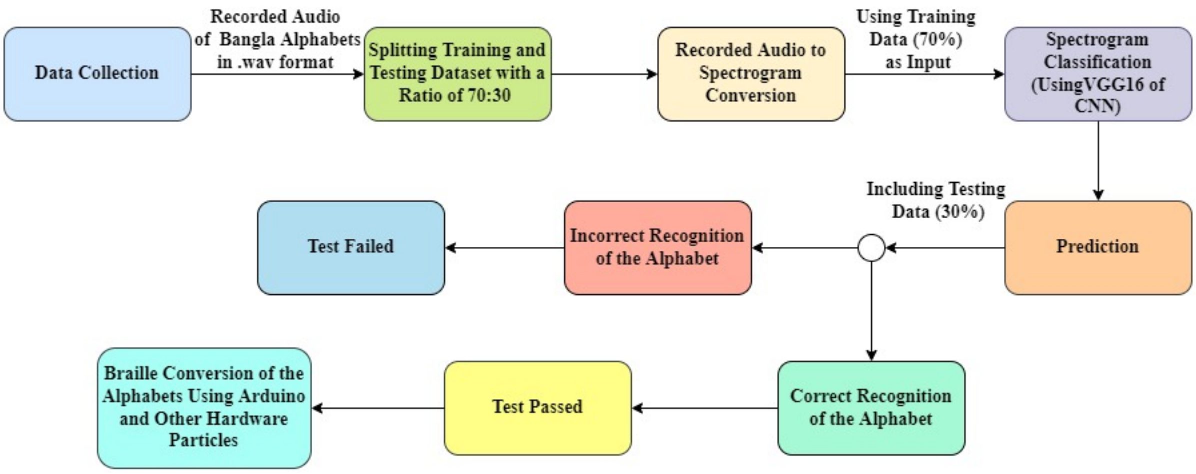

Figure 3.1: The work flow of our proposed model.

In short, we have designed a process that takes inputs from our datasets and processes the input data and after that, it will produce predictions in two fields: “correctly recognized” or “incorrectly recognized”. Here, “correctly recognized” means that our algorithm is recognizing Bangla alphabets properly. Then, we will work on those alphabets which are recognized correctly. We will use Arduino, LCD Monitor and other hardware particles to print the alphabet in Braille language pattern.

Data preprocessing is a really important part of our proposed work. It consists of three major steps:

1. Input data preprocessing: If data encoding is needed, it is done in these steps. In short, these steps format the input data in a way so that processing can be easily done.
2. Processing: In this stage, we are doing clustering and classification of our datasets. By using CNN, we are doing our classification part. Building a decision tree according to the dataset is part of it.
3. Prediction: This step gives us results that are found after testing the datasets. After that, we will work with Braille’s conversion of the alphabet.

We are using 70% of our dataset for training and 30% used for testing the datasets. Finally, we get the accuracy based on whether our system can recognize the Bangla alphabets correctly or not. We are trying with more and more datasets to improve accuracy. Taking the alphabets that are recognized properly we will do our Braille conversion part.

3.1 Data Collection

For our data collection, we have taken voice samples from people around us using voice recorder software. Our data mainly consists of 11 Bangla Vowels (Shorborno) we have taken each vowel 10 times. 11 Bangla Vowels (Shorborno)*10 times each, which means we have collected in total 110 voice samples from an individual person. We have collected 25 sets of data which contain almost 2600+ audio records. Among the voices we have collected so far, we have samples of both males and females, alongside male female samples we have also collected voice data of children. The voice that we have collected will be converted to Braille using Convolutional Neural Network (CNN). The data that we have collected were in m4a format, we have converted the collected data format to wav for ease of our use.

The below pie chart is used to give an insight into our collected data. As we can see in the pie chart about 50 percent of our collected voices comprise males. The second highest voices are the ones taken from females the least amount of voices were taken from children.

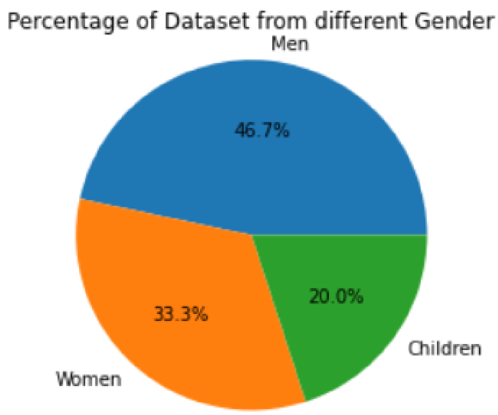

Figure 3.2: Percentage of dataset from different gender

3.2 Data preprocessing

The pictorial representation below is used to visualize the number of recordings for each command from a single dataset among the 23 datasets that we are working with. Here, the x-axis of the graph represents wav format files of 11 Bangla vowels (Shorborno) the y axis represents the number of recordings.

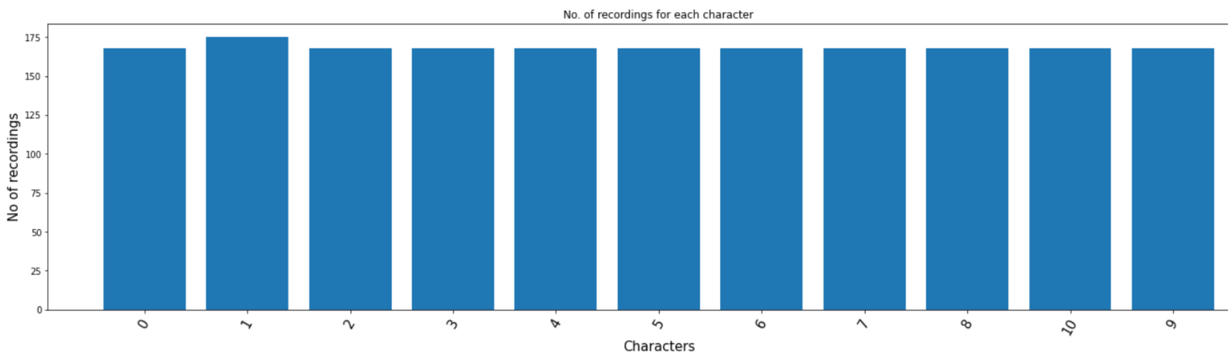

Figure 3.3: Data visualization.

In the image below we tried to give a visualization of the duration of recordings. Here in the image, our x-axis represents time in seconds the y-axis represents the number of datasets. For our dataset, we have used 2500+ wav files. If we take a closer look at the graph we can see in 2.7 seconds we get almost 500 characters from our dataset which is the highest of all the values. However, the lowest value is getting almost 10 characters in 4.5 seconds.

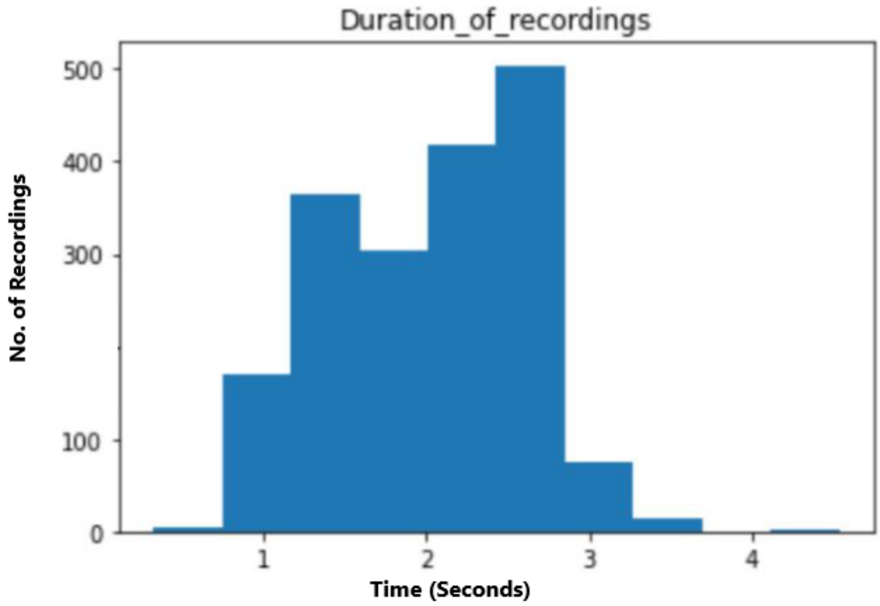

Figure 3.4: Duration of recordings (x-axis time in seconds(s), y-axis no. of recordings).

Next, we imported some libraries which are os, librosa, IPython.display, matplotlib.pyplot, numpy, scipy.io, warnings. After that, we have accessed the directory where a huge number of datasets are being kept. In data preprocessing, first we have converted all audio data to wav format from m4a format. Then we have labeled data following a naming format $character\_person\_charNum$ and moved train and test data to separate folders in a ratio of 7:3.

In the next part of data preprocessing, we have generated frequency spectrum for all individual data. To do this we have accessed wav files using oslistdir and created two sperate lists for train and test data. Then we have generated frequency spectrum (spectrogram) for all the individual data of those lists and saved all the train and test spectrograms in separate directories for further classification. We have used a sample rate of 16000 to generate frequency spectrum.

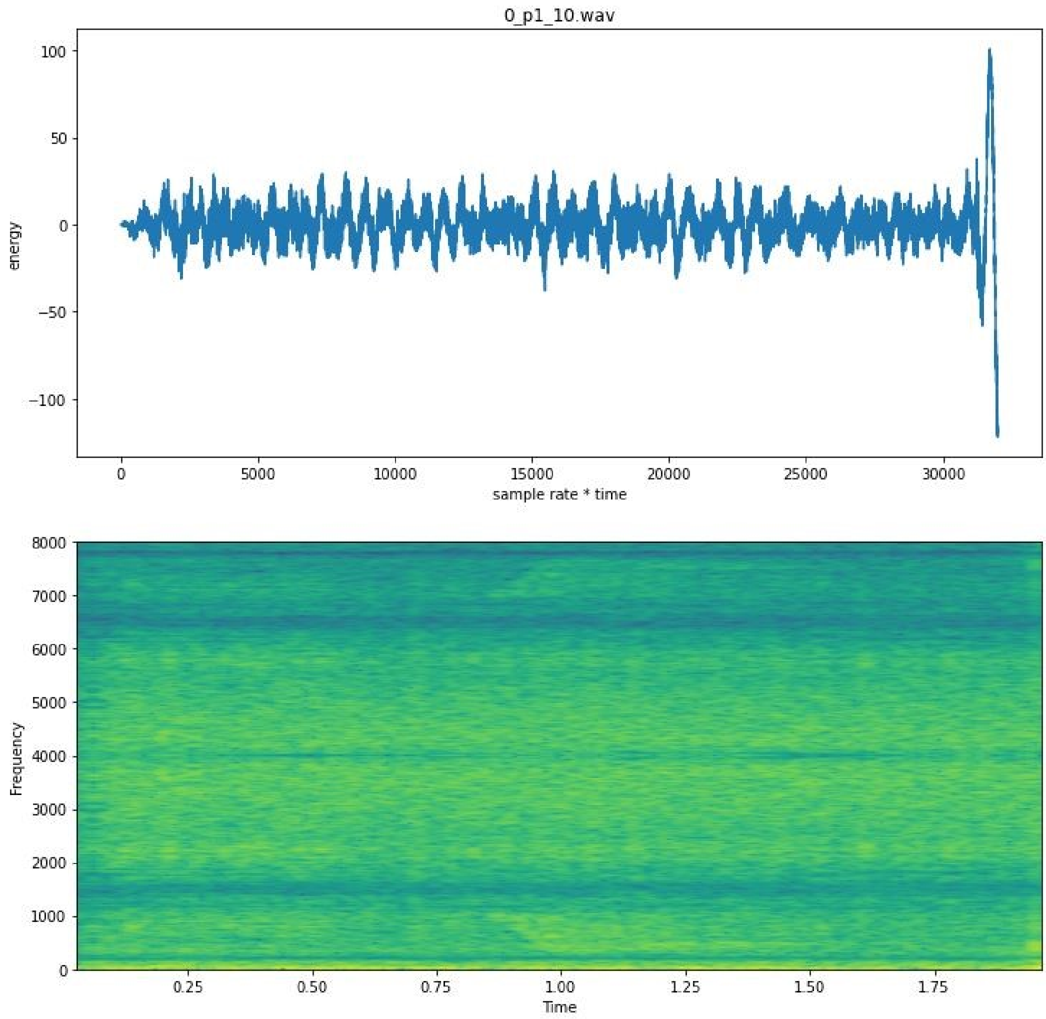

Figure 3.5: Waveform and frequency spectrum

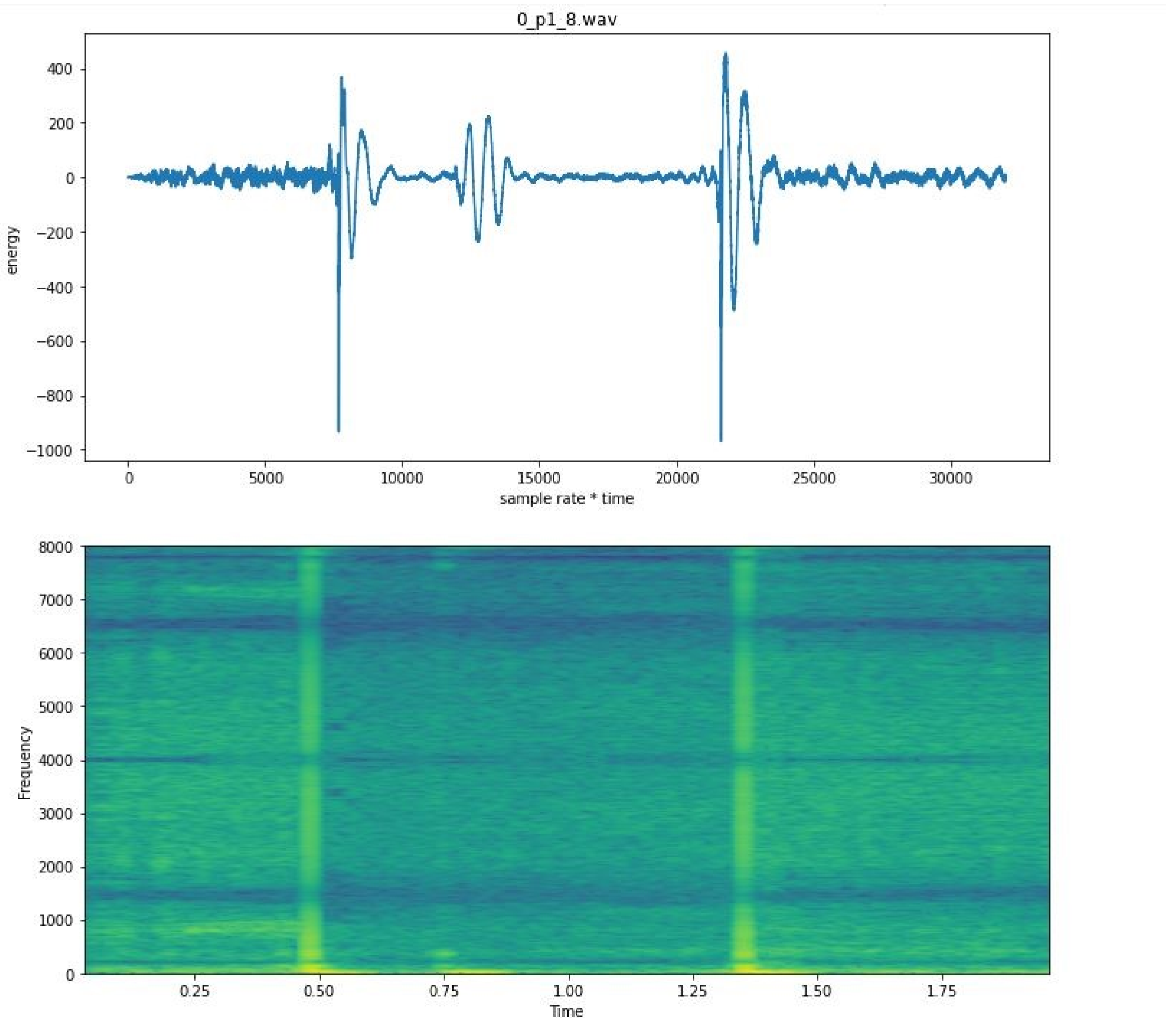

Figure 3.6: Waveform and frequency spectrum

Sample of plotted waveforms and frequency spectrums for different wav files. In waveform (sample rate * time) is in x-axis and energy is in y-axis. And for frequency spectrums we used time and frequency in x-axis and y-axis accordingly are shown figure

3.2.1 VGG-16

VGG 16, or Visual Geometry Group, is another well-known network [10] that was established in Oxford. It earned the (ILSVRC) for localization in 2014. There are 16 layers in the network, each of which can learn parameters. The architecture is similar to AlexNet, except it only employs 3x3 filters and is significantly more complex.

Figure 3.7: Visual description of VGG16

3.3 Spectrogram classification

For spectrogram classification we mainly worked with our image folder that we got in our previous part of work. We have used VGG16 for our spectrogram classification which is a convolutional neural network model. VGG16 is a specific convolutional network designed for image classification, detection and localization.

To train our model via VGG16 we have considered the initial input shape 256*256*3. We have set *include_top* = 'False' as we fitted the model on our own problem and set weights = 'imagenet' as we need to load the pre-trained weights of VGG16. Then we have used one of the imported library named ImageDataGenerator with it's different functions like rescale, shift, flip to import data with labels easily into the model. Then, using a batch size of 32 and an input shape of 256*256, we built an ImageDataGenerator object for both train and test data and passed the folder containing train data to the object generator train and similarly passed the folder containing test data to the object generator train. The folder structure of *generator_train* and *generator_train* we got is as follows:

```
Found 1701 images belonging to 11 classes.
Found 726 images belonging to 11 classes.
```

Figure 3.8: Folder structure of *generator_train* and *generator_test*

We got 1701 images belonging to 11 classes for *generator_train* and 726 images belonging to 11 classes for *generator_train*.

After initializing the model different convolution layer, layers were executed and they are Conv1D, Conv2D, Conv3D and Maxpooling.
Then we flattened the layers of our output model to pass the data to dense layer. We have used RELU (Rectified Linear Unit) activation for the dense layer of 1024 units so that we stopped forwarding negative values through the network. Because of we have 11 classes to predict, we employed an 11-unit dense layer with softmax activation at the end. Based on the model's confidence in which class the images belong to, the softmax layer will output a value between 0 and 1. After the creation of softmax layer the model is finally prepared.
After that, we used Adam optimizer to assemble the model and train it to the global minima. If we get caught in local minima while training, the Adam optimizer will assist us get out of it and reach global minima. We will also specify the learning rate lr = 0.00000e of the optimizer.
Now, the model is able for the spectrogram images classification. A sample of our classification is shown below:

```
{'class_0': 0,
 'class_1': 1,
 'class_10': 2,
 'class_2': 3,
 'class_3': 4,
 'class_4': 5,
 'class_5': 6,
 'class_6': 7,
 'class_7': 8,
 'class_8': 9,
 'class_9': 10}
```

Figure 3.9: Classification of images

3.4 Prediction

After getting accuracy , we mainly focused on generating output vector of our training model. For example, if we upload a spectrogram image of আ in our model then an output vector will be returned as an array like [0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0]. Only the index number 1 will be 1 as আ belongs to class 1 after the spectrogram classification . Along with the array the class number of that image will be shown also. Then we used this output vector as an input array in our hardware part for generating the braille pattern.

```
Saving 1_p5_10.png to 1_p5_10 (1).png
Output vector:  [0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0]
The character belongs to class 1 so it is আ
```

Figure 3.10: Output array and character

Chapter 4

Simulation and Test Bed

4.1 Simulation

We have imported the LiquidCrystal library in our arduino code as it allows an Arduino/Genuino board to control LiquidCrystal displays (LCDs) . In our arduino code, we need to mention which pins are connected with arduino and LCD 16x4 display by using LiquidCrystal lcd(5,6,7,8,9,10).This means 5,6,7,8,9,10 no. pins of Arduino UNO is connected with LCD 16x4 display.So, we can understand the importance the value of LiquidCrystal library over here.

We had to use a .hex file which was generated from the .ino file of our arduino code. We gave the .hex file of arduino code in our arduino uno as a source code.

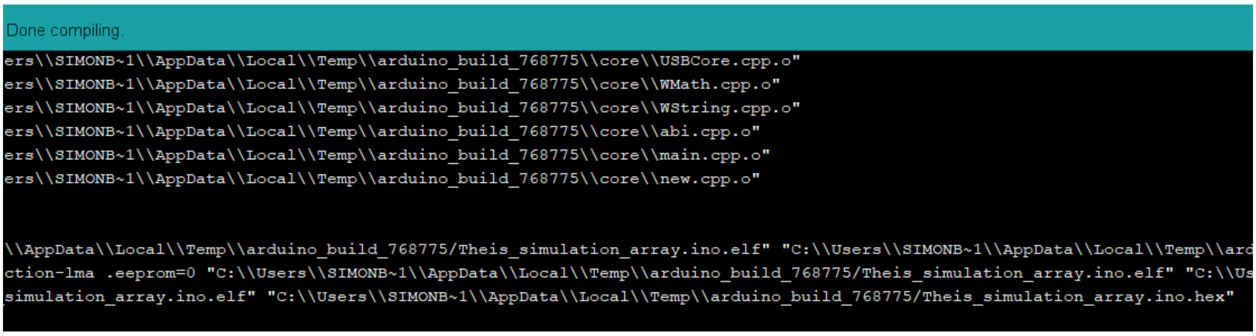

Figure 4.1: Successful compilation of arduino code and generating .hex file.

We are getting an array from our prediction code file which was generated from google colab. After that, we gave the array as input in our arduino code. In the array , only one index value will be 1 and all others will be 0 at a time. We have written our arduino code (.ino file) in such a way so that LCD can display only the braille pattern for which the index is 1.

There are many pins of LCD 16x4 displays and Arduino UNO boards. We connected both together by using male wires. The connection between Arduino UNO and LCD

16x4 display are given below:

Here, Pin no. 5 of Arduino is connected with RS of LCD 16x4 display. Then, we connected Pin no.6 with the Enable pin of the LCD. After that, we connected 4 digital pins of the 16x4 LCD with Arduino. D4,D5,D6,D7 are connected with arduino pins 7,8,9,10 sequentially. We connected these pins by using male wires.

The D0,D1,D2,D3 pins of the LCD 16x4 display are not used over here. VSS and VEE pins of the LCD 16x4 display are connected with Pot-Hg. The analog pins of arduino uno are not used over here.

Pot-Hg is connected with both Arduino UNO and LCD 16x4 display. Pot-Hg is also connected with power source and ground in proteus.

For example, we got an array from our prediction code [0,1,0,0,0,0,0,0,0,0,0]. Here, only the value of index number one is 1 and others are 0. So, for this array , in our LCD display we will get the braille pattern of আ as index number one which is representing the bangla shoroborno- আ .

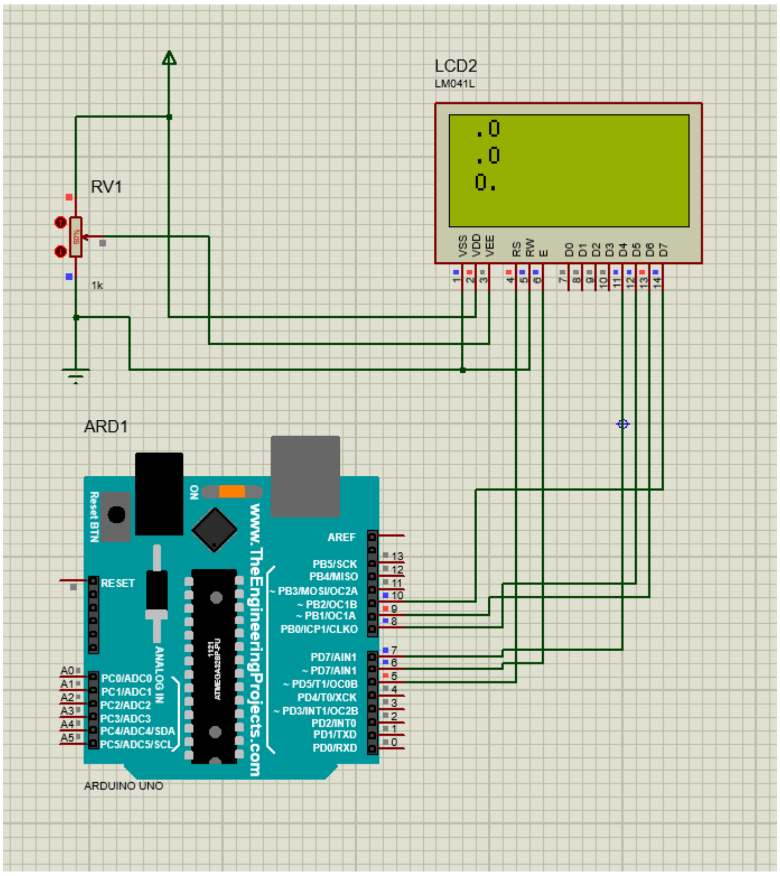

Figure 4.2: Printing the braille pattern of আ

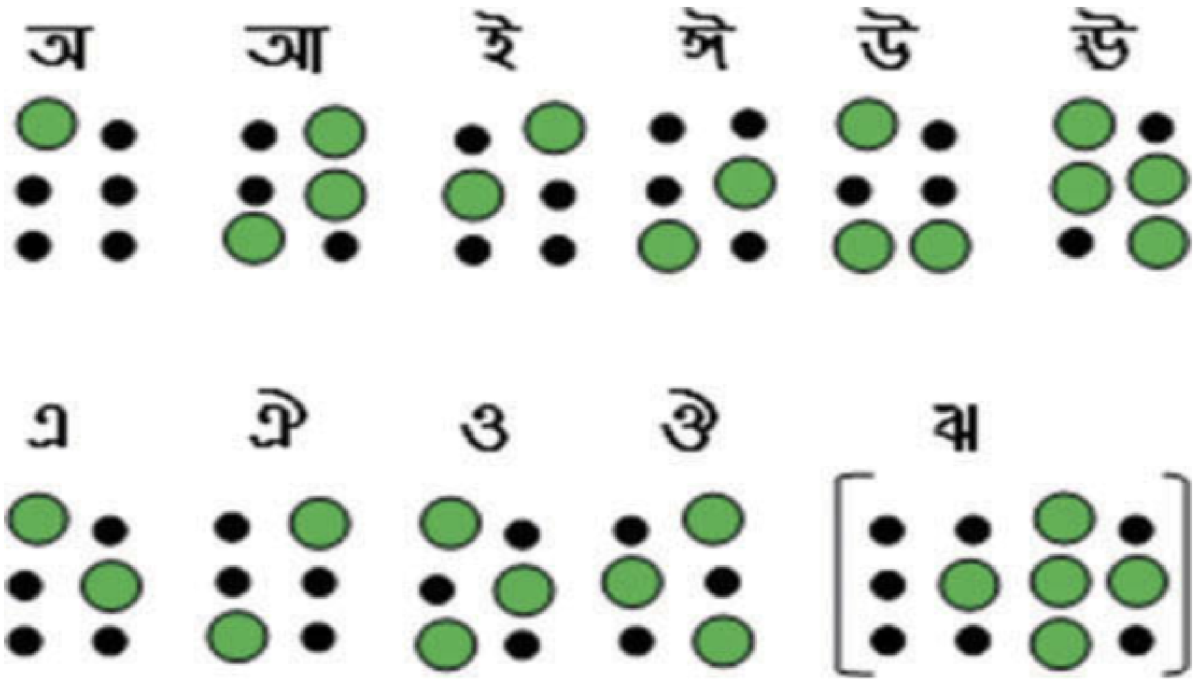

Figure 4.3: Bangla vowel braille pattern.

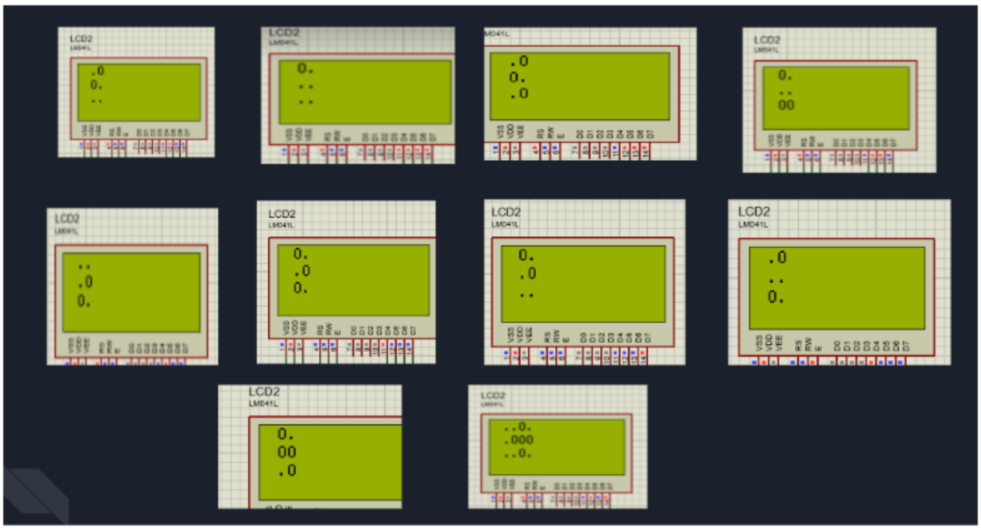

Figure 4.4: All other Bangla vowel braille patterns

4.2 Test Bed Implementation

4.2.1 Arduino UNO

Arduino UNO is a programmable open-source microcontroller board. It is really popular for doing different kinds of electrical projects. Besides, as it has low cost in comparison to Raspberry Pi , it is frequently used by tech lovers around the world. The Arduino name came from a bar in Ivrea, Italy and it was invented by Massimo Banzi. [11]

There are many important parts of Arduino UNO and we will describe them shortly. First of all, ATmega328 is called the brain of Arduino UNO. Here, the program is stored which is generated by arduino code's .hex file. The .hex file generated from the .ino file is used as source code in Arduino UNO. Then, there are some ground pins on the board.

Pulse Width Modulation(PWM) pins are one of the most important parts of the board. There are 6 PWM pins in the board and they are used mainly to control brightness of the LED and the speed of the servo motor, DC motor etc.

Digital pins are frequently used when we work with Arduino Uno [12]. There are 14 digital (0-13) I/O pins available on the board and they are used to connect the board with external electrical components.

Arduino UNO has 6 analog pins also. There are many electrical devices which give us input in analog form and these can read the analog signals easily and convert them into digital signals. [13]

3.3V pin used to supply 3.3V and 5V pin is used to supply 5V power. Vin pin is used for supplying voltage to the board.

To set an external reference voltage, Analog Reference(AREF) pin is used. After that, in order to reset the code loaded into the board, we need to press the reset button and this button will take the entire board into an initial state.

USB Interface is used to connect the board with the computer and to upload the Arduino programs. DC Power Jack is used when we want to power up the board with a power supply.

Power LED [14] is an important part of Arduino UNO as it confirms to us that the board is connected with power supply. Arduino UNO also supports external micro SD cards for supporting more information.

10(SS), 11(MOSI), 12(MISO), 13(SCK) pins of Arduino UNO are part of SPI(Serial Peripheral Interface).

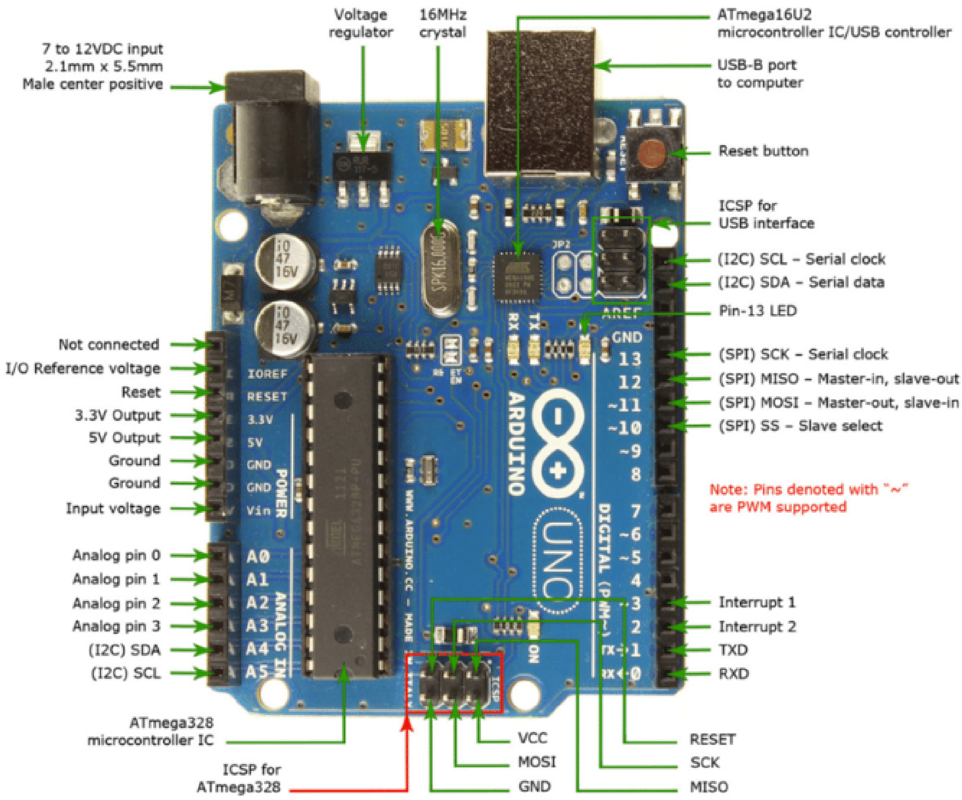

Figure 4.5: Description of Arduino UNO.

TX pin is used to transmit the serial data. RX pin is used to receive pin serial data.

4.2.2 LCD 16x4(LM041L) Display

LCD means Liquid Crystal Display. LCD is an electronically modulated optical device that uses the light-modulating properties of liquid crystals [15] combined with polarizers.There are many versions of LCDs for example 16x2,20x4, 16x4 etc.

LCD 16x4(LM041L) has 4 rows and 16 columns [24]. We used a 16x4 display in our thesis work as Bangla character braille pattern has 3 rows . So, we can't use a 16x2 display over here as it contains 2 rows.

There are many pins in the LCD 16x4 display. Vss is used to give a connection to the ground of the arduino uno board. Vdd pin is used for supplying 5V in LCD 16x4 display. RW pin is used for read/write. E pin is the enable pin. D0-D7 pins are digital pins used to connect with arduino. A is anode pin and K is cathode pin.

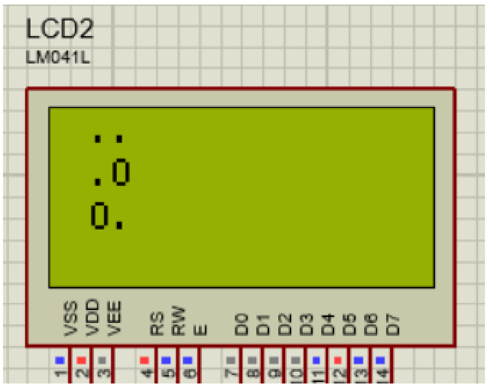

Figure 4.6: LCD 16x4(LM041L) display.

4.2.3 Pot-Hg

Pot-hg is actually a potentiometer in proteus which is used to control the resistance during the simulation of proteus. We can control the value of resistance easily by adding it. Pot-Hg needs to be connected with a power source(5V) and ground. Then, it also needs to be connected with an arduino uno board. Besides, we needed to connect it with an LCD 16x4 display for our thesis work.

This component is very effective to see the changes instantly when the value of resistance is changing . We can increase or decrease the value of resistance easily by clicking the upper or lower button of the Pot-Hg.

First of all, we used proteus for doing our simulation part. Proteus is a proprietary software tool suite used primarily for electronic design automation. In the proteus file, we took an arduino uno, Pot-Hg(resistor),LCD 16x4 display(LM041L) , ground, power source etc. for doing our simulation part.

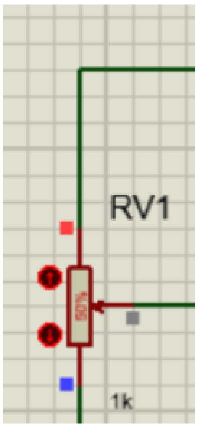

Figure 4.7: Pot-Hg

Chapter 5

Results and Discussion

For our thesis work we basically used the VGG16 model of Convolutional Neural Network (CNN). We have trained the model for 50 epochs and for training the used image size was 256*256. Here we achieved a validation accuracy of 98% and validation loss of 5%.

Then we evaluated the test accuracy and it was around 92.42%, here loss was around 30%.

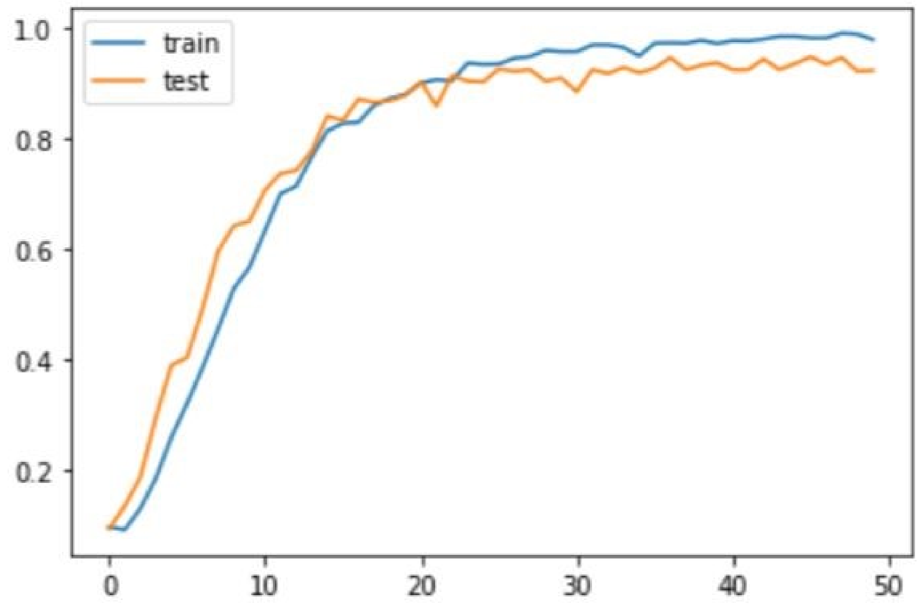

Figure 5.1: Train and Test Accuracy

Here the blue line represents the train accuracy which is 98% and the orange line represents the test accuracy which is around 92.42%.

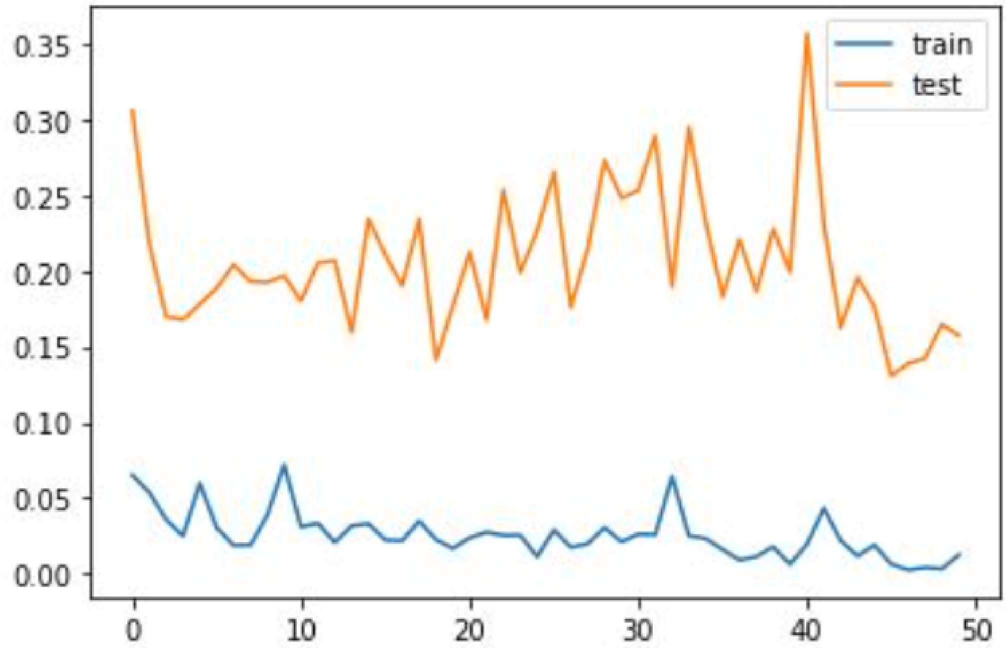

Figure 5.2: Loss of Train and Test

Here in the above picture the blue line represents the validation loss of train which is near about 5% and the orange line represents the loss of test accuracy which is around 35%

Chapter 6

Conclusion

Visually impaired people suffer a lot in their day-to-day life. This suffering becomes more pathetic if they can't get enough opportunities to be educated. Though they face many problems, the power of education can decrease their suffering. We want to develop a system so that visually impaired people of any age can learn from the basics. As our proposed method will print the Braille alphabet in a sheet after getting voice command of the alphabet, anyone can easily learn the alphabets which will lead the way in their future to learn more. We will train our system properly then hardware implementation using Arduino Uno. will be done. It will create interest among visually impaired people to learn properly. Many works have been done before for Braille language but we are focusing on Braille character conversion from voice. Finally, our research is focusing on helping visually impaired people so that with the power of knowledge they can prove their creativity in the future.

6.1 Future Works

In the future, we want to extend our works. In addition, we will collect more data sets of Bangla consonants from men, women, and children. Moreover, We will try to make our hardware part more useful and we will use a metallic pen. This metal pen will be connected to an Arduino and will print Bangla braille patterns on paper when a person will utter a Bangla character in front of the PC or laptop. Also, a visually impaired person will be able to touch this paper and feel the braille pattern which will be easier for them to learn Bangla braille characters.

Finally, we want to do some work for the betterment of society. We want education for all and if our initial future work plans executes properly, we will think about working with Bangla sentences as well in the future.

Bibliography

[1] T. R. Abir, T. S. B. Ahmed, M. Rahman, and S. Jafreen, *Handwritten Bangla character recognition to braille pattern conversion using image processing and machine learning (Doctoral dissertation)*. 2018.

[2] V. Jha and K. Parvathi, "Braille transliteration of hindi handwritten texts using machine learning for character recognition," *Int J Sci Technol Res*, vol. 8, no. 10, pp. 1188–1193, 2019.

[3] S. Ramachandran, N. Rajan, K. N. Pallavi, J. Subashree, S. Suchithra, and B. Sonal, "Communication device for the visual and hearing impaired persons to convert braille characters to english text," in *2021 International Conference on Emerging Smart Computing and Informatics (ESCI)*, Pune, India: IEEE, Mar. 2021.

[4] F. M. Mahbub-Ul-Islam, A. I. Khan, M. O. Gani, S. Azam, and S. A. Hossain, "Computational model for bangla text to braille translation," in *the 2nd International Conference on Computer Processing of*, Bangla, 2011.

[5] A. Hazra and M. M. Hoque, "Braille gloves: An intelligent hand-glove to generate bengali braille characters for visually impaired people," in *2019 IEEE Region 10 Symposium (TENSYMP)*, Kolkata, India: IEEE, Jun. 2019.

[6] W. Yang, J. Huang, R. Wang, W. Zhang, H. Liu, and J. Xiao, "A survey on tactile displays for visually impaired people," en, *IEEE Trans. Haptics*, vol. 14, no. 4, pp. 712–721, Oct. 2021.

[7] S. Padmavathi, M. K. S. S, S. S. Reddy, and D. Meenakshy, "Conversion of braille to text in english, hindi and tamil languages," Jul. 2013. arXiv: 1307.2997 `[cs.CV]`.

[8] T. Dasgupta and A. Basu, "A speech enabled indian language text to braille transliteration system," in *2009 International Conference on Information and Communication Technologies and Development (ICTD)*, Doha, Qatar: IEEE, Apr. 2009.

[9] M. Ram´lrez, M. Sotaquirá, A. De La, E. Cruz, G. Maria, and A. Avellaneda, "An automatic speech recognition system for helping visually impaired children to learn braille," in *2016 XXI Symposium on Signal Processing*, Bucaramanga, 2016, pp. 1–4.

[10] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," Sep. 2014. arXiv: 1409.1556 `[cs.CV]`.

[11] L. Louis, "Working principle of arduino and using it as a tool for study and research," *International Journal of Control, Automation, Communication and Systems*, vol. 1, no. 2, pp. 21–29, Apr. 2016.

[12] Y. A. Badamasi, "The working principle of an arduino," in *2014 11th International Conference on Electronics, Computer and Computation (ICECCO)*, Abuja, Nigeria: IEEE, Sep. 2014.

[13] L. Louis, "Working principle of arduino and using it as a tool for study and research," *International Journal of Control, Automation, Communication and Systems*, vol. 1, no. 2, pp. 21–29, Apr. 2016.

[14] F. M. Steranka, J. Bhat, D. Collins, *et al.*, "High power LEDs - technology status and market applications," en, *Physica Status Solidi A Appl. Res.*, vol. 194, no. 2, pp. 380–388, Dec. 2002.

[15] S.-W. Lee, "Intelligent liquid crystal display (i-LCD) for next generation television applications," *IEEE trans. consum. electron.*, vol. 53, no. 4, pp. 1247–1253, Nov. 2007.

[16] *Braille patterns*, en, https://unicode-table.com/en/blocks/braille-patterns/, Accessed: 2022-5-21.

[17] *Bangla to braille translator*, en, https://wecapable.com/braille-translator/bangla-braille/, Accessed: 2022-5-21.

[18] D. A. Kent, *What Is Braille?* en, ser. Overcoming Barriers. Enslow Elementary, Jan. 2013.

[19] The Editors of Encyclopedia Britannica, "Bengali language," in *Encyclopedia Britannica*. Jul. 2017.

[20] B. Guven, *Speech recognition in Real-Time using python*, en, https://towardsdatascience.com/speech-recognition-in-real-time-using-python-fbbd62e6ff9d, Accessed: 2022-5-21, Dec. 2021.

[21] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," en, *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

[22] T. Guo, J. Dong, H. Li, and Y. Gao, "Simple convolutional neural network on image classification," in *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)(*, Beijing, China: IEEE, Mar. 2017.

[23] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE Inst. Electr. Electron. Eng.*, vol. 86, no. 11, pp. 2278–2324, 1998.

[24] R. Hecht-Nielsen, "Harcourt brace & co," *Neural networks for perception*, vol. 2, pp. 593–605, 1992.

[25] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Computer Vision – ECCV 2014*, ser. Lecture notes in computer science, Cham: Springer International Publishing, 2014, pp. 818–833.

[26] C. Szegedy, W. Liu, Y. Jia, *et al.*, "Going deeper with convolutionals," *Co RR*, 2014.

[27] A. Waibel, "Phoneme recognition using time-delay neural networks. acoustics, speech and signal processing," *IEEE Transactions on*, vol. 37, no. 3, pp. 328–339, 1989.