Design and Development of Smart Braille Printer Using Convolutional Neural Networks

Tanmay R Nandurge[1], Aniket C Patil[2], Virag S Minajagi[3], Tejas D Nigave[4]

[1]B.Tech Student, Mechanical Engineering, DKTE Society's Textile and Engineering Institute, Maharashtra, India
[2]B.Tech Student, Mechanical Engineering, DKTE Society's Textile and Engineering Institute, Maharashtra, India
[3]B.Tech Student, Mechanical Engineering, DKTE Society's Textile and Engineering Institute, Maharashtra, India
[4]B.Tech Student, Mechanical Engineering, DKTE Society's Textile and Engineering Institute, Maharashtra, India

---***---

ABSTRACT-*Life can be experienced in a beautiful way by human beings, through the gift of five senses. But a minor population of human beings is underprivileged to experience life to full extent due to disability. A communication barrier exists between the people with different disabilities such as the blind and the deaf, dumb and various others. Convolutional Neural Networks could be used to design a cost-effective system, which could be afforded by almost anyone, to overcome the existing shortcomings. The paper includes design and development of economical and smart braille printer for integration of the problems faced by people with disabilities. The hand gestures in sign language were used as input for the web cam, followed by conversion of images to text and finally text to braille using Convolutional Neural Networks. Tensorflow, an open-source software for machine learning, was used in order to convert sign language to text. The braille output was given as an input to the dot matrix printer, which in turn embossed the braille on the paper.*

KEY WORDS*:* **Image processing, CNN, Canny Edge Detection, Sign-to-text conversion, text-to-braille conversion.**

1. INTRODUCTION

Approximately four percent of the world's population, which accounts to 250 million people, is suffering from visual impairment and another 40 million are blind [1]. The only source for complete literacy of blind people are braille scripts. Braille is tactile writing system. It consists of rectangular blocks called cells that contains bumps called raised dots. The number of dots and their arrangement inside the cell distinguish one character from another [2]. There are approximately 72 million deaf people around the world [3]. Sign language is used as a mode of communication by the deaf and dumb to communicate with normal population and vice versa.

There is hardly any device available, which can develop a smooth communication between deaf-dumb and blind people. Some of the devices available such as the braille embosser, which converts the voice input and written documents into braille scripts, are too expensive to be afforded by an individual. A typical braille embosser costs approximately in the range of 50,000 INR TO 1,30,000 INR [4]. While the cheaper devices cannot reduce the existing gap between deaf-dumb and blind people. To vanquish the problems associated with current systems, a smart braille printer system was developed.

Fig-1: Braille Embosser

This paper proposes a smart braille printer which uses sign language hand gestures as an input at the camera, then converts the images captured through camera into text and finally converts text into braille. The braille output is then used as an input for the printer which embosses the braille translation onto the paper. The conversion of sign language to text was done by using Tensorflow, an open-source Artificial Intelligence frame work. Canny Edge, an edge detection algorithm, was used in order to recognize the Edges of the images captured by the camera. Once image is converted to text a python script invokes at the background, which converts the text to braille. The converted Braille script then embosses on the paper by a Modified Dot Matrix printer which reduces the cost of the entire system drastically.

The proposed system is smart as it has greater number of inputs such as text and hand gestures of sign language and the system is more user friendly. The additional benefit of the system is that it has a great speed compared to some other proposed systems. Finally, it is affordable by almost any individual with a decent financial background.

2. METHODOLOGY

The prolific aim of this paper is to use the hand gestures in sign language as an input for the system, then image processing was used in order to convert the input images into text. Convolutional Neural Networks were used to convert sign language to text. The web camera captured the hand gestures used in sign language the captured image this in turn activates the code inherited in the software and initializes the code. The images captured were converted into text using Tensorflow framework. Then once the output is available at the command prompt a python code starts at the background which converts the text module into braille language. The converted text was then sent to the printing software which turns on the printer and the printer prints the output.

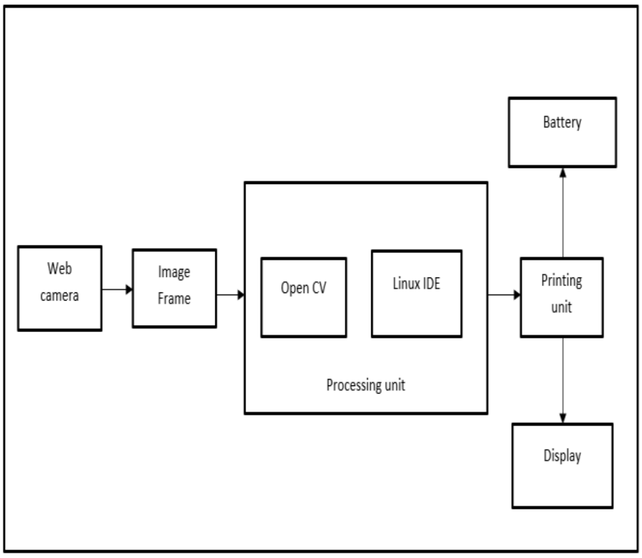

Fig-2: System Block Diagram

2.1 Problems with the existing systems

Several smart printers are available, but with a limited input range. Only text and voice can be given as an input for the system which then converts the given input into sign language. Therefore, a smart braille printer was needed to develop to increase the input range, which would bring advantage to various users. The available systems are less economic and could hardly be afforded by the people with humble background. The typical braille printers which convert the voice into braille, the braille embosser, costs INR 50000 and above [4]. It is clear that it could not be afforded by most of the individual customers. The current systems available produce a lot of noise while the braille output is being printed onto the paper. The repeated noise may cause irritation to the user and the user may feel uncomfortable to use it. Also, the speed of the various printers in the market is considerably slow as compared to the proposed system. In today's fast changing world, one cannot compromise on the speed of the mechanisms. There are hardly any mechanisms or technology which reduces the communication barrier between the deaf and the blind people.

2.2 Selection of suitable input method

The preliminary basic inputs for any type of system are of three types voice, visual and text. Among the three inputs available visual input was specifically suitable in order to achieve the aim of the project and it was chosen for its prolific advantage over various other inputs. The accuracy of the mechanical system is significantly high when compared to the visual or speech recognition systems. The complexity of the mechanical systems and voice recognition systems is high as compared to the visual recognition systems. In visual recognition systems the patterns required to train the model are easily recognizable which makes it considerably less complex than other systems. While on the other hand in the automatic speech recognition systems it is difficult to get accurate results due to various reasons. Firstly, there is a vast amount of variation that occurs during pronunciation of a specific word. Secondly, when the number of words increases, we end up comparing with a very large set of models, which is computationally not feasible. There is another problem of finding enough data to train these models. Finally, the visual recognition systems are much more feasible in terms of cost as compared to mechanical and voice recognition systems.

Table-1: Matrix Diagram for suitable input

Input Method	Accuracy	Intricacy	Time	Cost
Mechanical	High	High	High	Mid
Speech	Low	Mid	Mid	High
Visual	Mid	Low	Low	Low

3. IMAGE PROCESSING

Image processing plays a crucial role in the conversion of sign language into the text and text back into the braille output. The TensorFlow framework was used in order to recognize the images captured by the web camera. Deep learning was the only way to accurately predict output which have an accuracy of between 80 to 100. As traditional systems use the features that have manual filters. Then the models were trained using the Neural Networks. The Convolutional Neural Network was specifically chosen over the Artificial Neural Networks due to its specific advantages that were useful in this system development. The general advantage of Convolutional Neural Networks over the Artificial Neural Networks is the complexity of the process increases as the computational data increases.

3.1 Convolutional Neural Networks (CNN)

Convolutional Neural Network (CNN) is a class of deep neural networks which is mostly used to do image recognition, image classification, object detection, etc. It consists of mainly three layers the convolutional layer, which extracts the main edges corners of the and important features from the given image [5]. Secondly, the pooling layer, which distinguishes the Artificial Neural Networks (ANN) from the CNN, has the purpose to reduce the computational power required by the computer this is done by extracting more dominant features form the convoluted image. The final layer is fully connected layer where the flattening of the image into column vector is done.

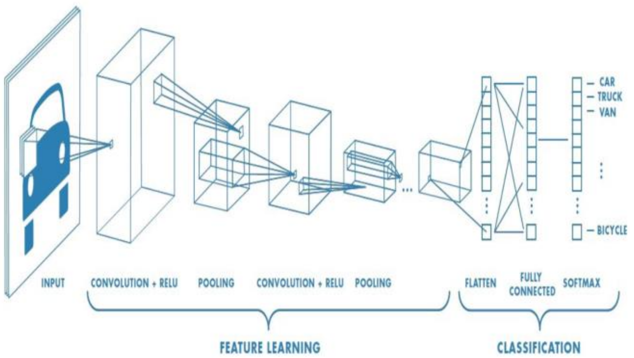

Fig-3: CNN Architecture

3.2 Image equitization and processing

The images were captured by the Logitech 720p camera at 30fps. The captured images were the sent as an input to the convolutional layer the edges of the input images were tracked in order to reduce the complexity of the images. In order to achieve this the kernel which is a set of learnable parameters was used to have a dot product with the matrix which is the restricted portion of the image captured. The Guassian filter or kernel of size (2k+1) is given by [6]

$$H_{ij} = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{\left(i-(k+1)\right)^2+\left(j-(k+1)\right)^2}{2\sigma^2}\right);1\le i,j\le(2k+1)$$

Fig-4: Kernel filter formulae

Then max pooling was done on the convoluted images in order to reduce the complexity and this max pooling operates on each feature map independently. Then the flattening of the image was done into a single column vector. Logistic regression was used in order to train the model.

The image captured first is converted from RGB to the grayscale image and the algorithm used was the method of weighted method which is more accurate. In this the rgb is converted to grayscale using the wavelengths of the red green and blue colors individually and summing them up.

Grayscale = 0. 299R+0.587G+0.114B [7]

Fig-5(A): RGB image

Fig-5(B): Grayscale image

Canny Edge detection algorithm was used due to its various advantages over the other techniques of edge detection such as Sobel and various others [8].

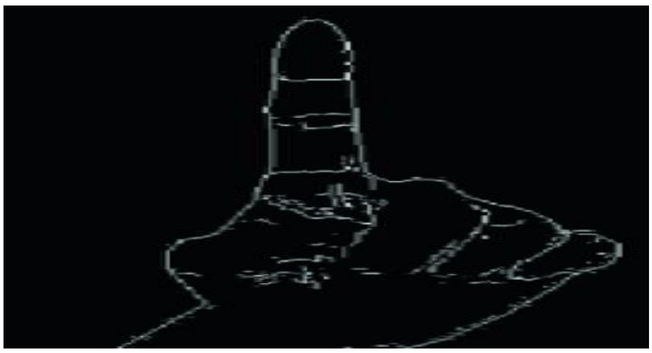

Fig-5(C): Canny edge detected image

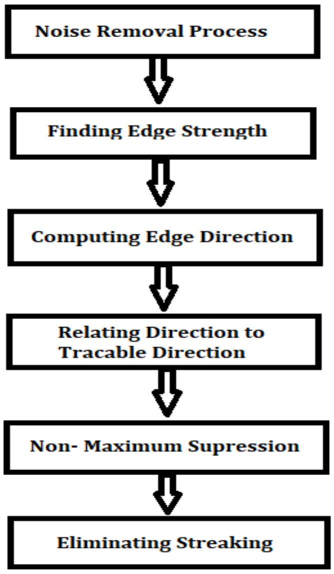

Fig-6: Steps in Canny Edge detection

In the initial stage of image processing for detection of sign language the label maps were created. Then the Tensorflow records were created.

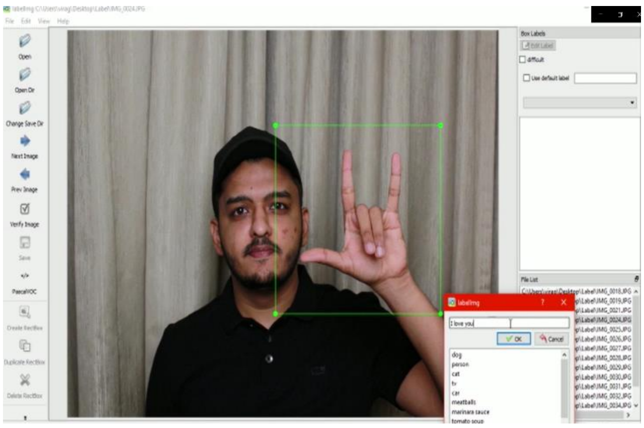

Fig-7: Image Labelling

Thereafter, the pretrained models which were trained before were used and the model configuration were copied to the training folder [9]. Finally, the configuration was updated in order to run the transfer learning. Then the models were trained and loaded from the checkpoint. The output was obtained on the command prompt which in turn will be used as an input for the python script.

3.3 Text to Braille conversion

A python code was used in order to implement the text to braille conversion. In this the python script reads the input from the command prompt, which is a text, and converts the text into braille. In the python script the mapping of several ASCII characters was done to various alphabets and few of the general words in the sign language. The code recognizes the text and matches it with the ASCII characters and gives the desired output in the result module.

3.4 Printer modification and output

A dot matrix printer to braille embosser was modified by Removing the ink ribbon o Modifying the platen of printer in this paper we chose to go with modification of Dot matrix printer [10]. In the printer the ribbon was removed to enable direct contact of the printer head pins with the paper. Moreover, to avoid direct contact of print head pins with the platen, the platen was covered in a Cushioned Foam Sleeve. These modifications enabled the printer to emulate a Braille Embosser. This module prints the braille text file. Once the converted braille appears in the result module then the system directs the output towards the printer software which in turn prints the output on a thick braille paper.

4. RESULTS

1. Sign to text conversion:

Real time image was fed to the computer using a web camera. And a real time sign language detector was designed which detects the image and displays the results in real time. The system was built using Tensorflow framework.

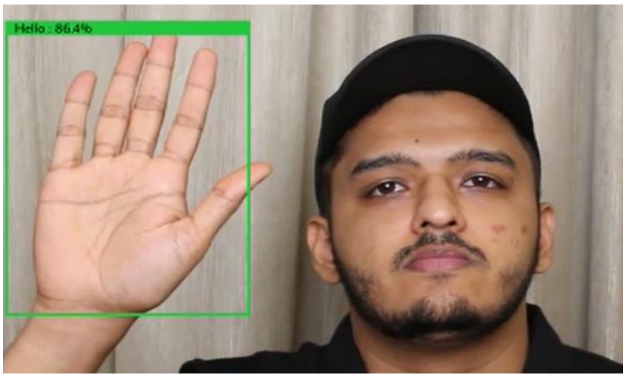

Fig-7(A): Real Time Detection

Fig-7(B): Sign to text conversion

2.Text to Braille conversion:

After the real time image is detected the python code invokes automatically and converts the text input to the braille output. In the python code the mapping of the ASCII characters was done to the various alphabets of the English Language.

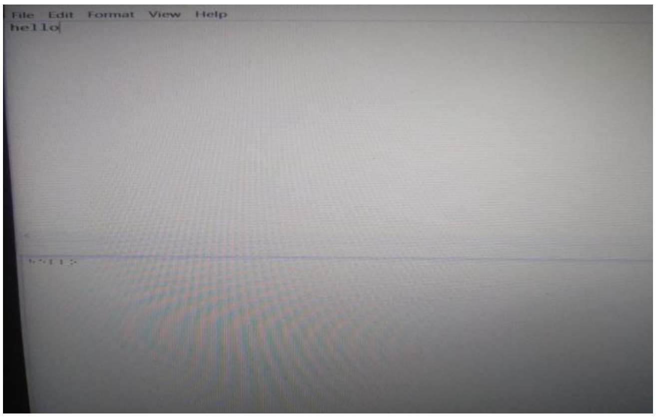

Fig-8: Text to braille conversion

3.Braille printer:

The software of the braille printer receives the converted braille output from the python script result module. The braille embosser, the dot matrix printer, embosses the received braille input on a braille paper.

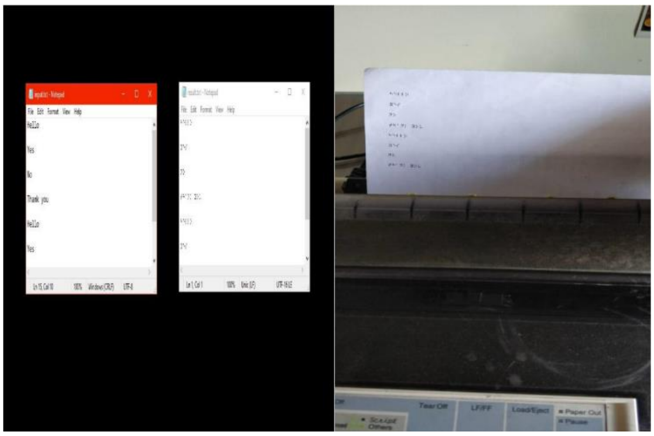

Fig-9-: Embossed Braille

5. CONCLUSIONS

The system which could be beneficial for the reduction of communication barrier between deaf, dumb and blind people was developed. This system could be used by blind people in order to increase their comfort and literacy. The proposed system overcame various barriers in the systems that were developed before. The speed of printing was improved and was recorded 10 w/s, but in this project the speed recorded was 30 w/s this increased the efficiency of the braille printer. Multiple inputs such as voice and text were added to the systems that were developed before. The cost of the braille printer was dramatically reduced by the use of the dot matrix printer. This play's and will play a critical role in the improvement of life of the people with disability. Convolutional Neural network played a critical role in the implementation of the project. The images were captured and converted to grey scale and HSV filter was applied which dramatically reduced the noise which was available during the capturing the image. Due to the image processing by CNN the moving images were detected and real-time sign language detection system was developed. Simple python code was used in order to covert text to braille.

6. FUTURE SCOPE

The current system recognizes and coverts only the American Sign Language, in future work all the other sign languages around the world could also be given as an input. Moreover, the computational time of the software could also be faster as this existing system detects with less speed. Thirdly, the voice output could also be provided to the device so as to make the dialogue between the deaf, dumb and blind people less hectic and time consuming. The feature of portability could also be introduced in the system by use of certain microprocessors or microcontrollers such as raspberry pi or Arduino. In future work parallelizing an edge-detection algorithms, provides better performance results for image-processing applications.

7. REFRENCES

1) University of Bath, "How Blind People See the World", 2021, Retrieved from https://www.bath.ac.uk/case-studies/how-blind-people-see-the-world/
2) Ariruna Dasgupta, Dibyabiva Seth, Anindita Ghosh, Ashok Nath, "Real Time Sign Language to Braille Interfacing System", 7th International Conference on Communication Systems and Network Technologies, IEEE, Nov. 2017, pp. 371-375, DOI: 10.1109/CSNT.2017.8418569.
3) World Federation of the Deaf, "International Day of Sign Languages 2021", Sept. 2021, Retrieved from https://www.business-standard.com/about/when-is-international-day-of-sign-languages
4) Ashwin P.S., Karthik B., Karthik D., Senthamizh Selvi R., "Speech Synchronised Brail Printer", International Journal of Computer Science & Communication Networks, vol. 7(2), Apr.-May 2017, pp. 21-25.
5) Anirudha Ghosh, Abu Sufian, Farhana Sultana, Amlan Chakrabarti, Debashis De, "Fundamental Concepts of Convolutional Neural Network", ResearchGate, Jan. 2020, pp. 1-40, DOI:10.1007/978-3-030-32644-9_36.

6) Dr. Rohini Deshpande, Y Divya Gnana Preetham, Vishal S., I. Krishna Jayanth, Srivatsa S., “Conversion of Sign Language to Text, Speech and Braille Code Using Image Processing”, International Journal of Advance Science and Technology, vol. 29, pp. 3713-3717.
7) Christopher Kanan, Garrison W. Cottrell, “Color-to-Grayscale: Does the Method Matter in Image Recognition?”, Plos One, vol. 7, Jan. 2012, pp. 1-7.
8) Raman Maini, Dr. Himanshu Aggarwal, ”Study and Comparison of Various Image Edge Detection Techniques”, International Journal of Image Processing (IJIP), vol. 3. Issue. 1, Feb. 2021, pp. 1-12.
9) Aurélien Géron, “Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow”, O'Reilly Media, Inc., Sept. 2019.
10) G. Singh, P. Kumar, “Improved Braille Printer”, Indian Patent: 3331/DEL/2016, November 2013.