Full research paper

BIDIRECTIONAL BRAILLE-SPEECH COMMUNICATION SYSTEM FOR DEAFBLIND STUDENTS

Ana Paula Pérez-Aguirre
Iván Arturo Morales-Pérez
Jorge Allan Gómez-Mercado
Rodrigo Alberto Gutiérrez-Martínez
Iván Matehuala-Moran
Rubén Fuentes-Alvarez✉

School of Engineering and Sciences,
Tecnologico de Monterrey, Mexico

✉ joru.fua@tec.mx

ABSTRACT

Deaf-blindness is a type of dual disability wherein visual and auditory capabilities are significantly impaired. Special communication methods have been developed for the deaf-blind community. Yet, these methods require that both people involved have prior knowledge and training to successfully communicate, limiting deaf-blind people's social interactions, particularly in academic settings. This paper describes the development of a device that enables two-way communication between a everely deaf-blind user and a hearing person with no prior knowledge of Braille and no additional intermediaries. A Convolutional Neural Network (CNN) scheme for speech recognition was designed and implemented along with the development of an algorithm capable of developing both text-to-speech and Finger-Braille-to-text conversion. Lastly, a system integration via 3D modeling and additive manufacturing was carried out to deliver a functional prototype. The resulting device aims to allow deaf-blind students to send and receive information entirely in finger Braille, using buttons and vibrotactile feedback. In contrast, the hearing tutor receives auditory messages and speaks to reply, making the educational experience as familiar as possible for both parties. Users testing the device achieved an average typing accuracy of over 95% and demonstrated an understanding of commands transmitted through the device's components.

KEYWORDS

Accessibility, assistive technology, Braille, Convolutional Neural Networks, deaf-blindness, education

HOW TO CITE

Pérez-Aguirre A. P., Morales-Pérez I. A., Gómez-Mercado J. A., Gutiérrez-Martínez R. A., Matehuala-Moran I., Fuentes-Alvarez R. (2024) 'Bidirectional Braille-speech Communication System for Deafblind Students', *Journal on Efficiency and Responsibility in Education and Science*, vol. 17, no. 2, pp. 164-177. http://dx.doi.org/10.7160/eriesj.2024.170206

Article history
Received
September 26, 2023
Received in revised form
Februrary 27, 2024
Accepted
May 14, 2024
Available on-line
June 30, 2024

Highlights

- *Creating a Convolutional Neural Network to recognize the Spanish alphabet and its translation to Braille language.*
- *A working prototype capable of sustaining two-way communication through Finger Braille and Voice Recognition for its usage in educational environments.*
- *A Braille translation algorithm adapted to the Spanish alphabet to cater to the deaf-blind population of Mexico.*
- *A communication method for the deaf-blind requires no prior Braille or Sign Language knowledge from the hearing party.*

INTRODUCTION

In Mexico, 4.9% of the population has a disability, be it physical, sensory, psychosocial, intellectual, or multiple (INEGI, 2020). Out of all disabled people in the country, 6.5% (466,178 people) have deaf-blindness (CONADIS, 2018), a type of multiple disability wherein both visual and auditory capabilities are significantly impaired (Ask Larsen and Damen, 2014), which creates a barrier for interaction, education and integration to society. Deaf-blindness can occur for several reasons, the most common one being Usher Syndrome (Ayton et al., 2023). The most prevalent types of this syndrome are defined by partial or profound deafness at birth and a progressive loss of vision due to retinitis pigmentosa, which generally starts in childhood or adolescence (Ayton et al., 2023).

Patients with Usher Syndrome often get cochlear implants (CIs) to restore hearing, with reported good outcomes when done early in life (Ayton et al., 2023). However, this raises ethical concerns about parents deciding for their children without their consent. Schulz et al. (2023) address concerns about autonomy, freedom, identity, participation, and justice regarding CIs, with parents of deaf children reporting a lack of information and a bias towards implantation when discussing options with their children's doctors. Some CI users express feeling limited by the implant in that certain activities were off-bounds due to risks of electrical shock or damage to the implant (Schulz et al., 2023). Besides these physical factors, Smolen and Paul (2023) describe the impact of CIs on identity. They state that "deaf" and "Deaf" (capitalized) are distinct terms that refer to the physical inability to hear and the ‘culturally distinct group, who typically use a visual language and are said to have a “Deaf Identity”’ (Smolen and Paul, 2023), respectively (the authors will henceforth use “deaf” and “deaf-blindness” to refer exclusively to the sensory impairments, but “Deaf,” “Deaf-blindness” and “d/Deaf” or “d/Deaf-blindness” to acknowledge the cultural aspect of the conditions). With Deaf Culture being so prevalent, by giving patients CIs, doctors often inadvertently give deaf children a conflicting identity, being audio-logically deaf but able to hear thanks to technology. The authors of this paper present these issues not as an argument for why CIs should not be offered to d/Deaf-blind patients but as an explanation for why they are not always chosen. This work will focus primarily on severe to profoundly deaf-blind people who, for any reason, do not have CIs.

Castiglione and Möller (2022) state that human communication generally relies on “far senses”: sight and hearing. Only when these senses are impaired do the “near senses” (i.e., taste, touch, and smell) gain importance in this aspect. Methods of communication with and between people with deaf-blindness can vary depending on the extent and type of their visual and/or hearing impairment and their personal preferences, having different options to communicate such as spoken languages, sign languages, tactile sign language, deaf-blind manual alphabet, tadoma, deaf-blind block alphabet, and finger Braille (Hersh, 2013).

The first two communication methods are only viable if the d/Deaf-blind person has enough residual visual and/or hearing capabilities, which is not always true. In individuals with Usher syndrome, once adequate methods may not always be so. In the words of a deaf-blind man who acquired visual impairment later in life: “I have become more isolated because of the vision impairment now during the last 15 years. […] when you have difficulties, you can no longer use lip reading” (Turunen-Taheri et al., 2023). Therefore, when visual or hearing stimuli are not viable, people with deaf-blindness and those who interact with them must learn another communication scheme.

Braille is a tactile writing system and alternative communication method that uses six dots to represent individual letters, punctuation marks, numbers, and more. These dots are numbered from top to bottom and from left to right, each having two states: raised or flat, allowing for 64 possible combinations (Blenkhorn, 1995). Some of these are reserved for “modifiers,” changing how the following character is interpreted. For example, a character with raised 3rd, 4th, 5th, and 6th dots (⠼) is, in several languages, considered a number indicator, meaning that the following character will be interpreted as a digit. Similarly, a character with dots 4 and 6 raised (⠨) is used to capitalize the following letter in Spanish Braille. Modifiers can maximize the efficiency of a six-dotted Braille cell by re-signifying Braille patterns systematically instead of assigning a new combination to each character. However, there are contexts where these 64 combinations (plus modifiers) are not enough, in this case, a fourth row can be added to the Braille cell to allow for up to 256 unique characters.

The Braille typewriter or Brailler is an Assistive Technology device that works like a conventional typewriter, except that it has only nine keys. Six of these represent each of the six dots in a Braille cell; another key functions as a spacebar, and the remaining two help navigate the text (Moore and Murray, 2001). Pressing the keys causes a mechanism to hit the paper, raising the corresponding dots. Figure 1 shows a Perkins Brailler, one of the most common Braille typewriters.

In Figure 1, keys 1-6 represent the six dots that make up the Braille cell. Keys 1 and 4 are pressed by the index fingers of the left and right hand, respectively. Keys 2 and 5 correspond to the middle fingers, and keys 3 and 6 to the ring fingers. Key 7 is the spacebar, and key 8 is used to jump to the next line. Key 9 is the Backspace key used to go back to the previous character and correct any possible mistakes. The number 10 is the carriage lever and is not considered a key. It marks the position of the current character and is responsible for announcing the end of the line by making a sound when it approaches the end of the page (Morgan et al., 2011).

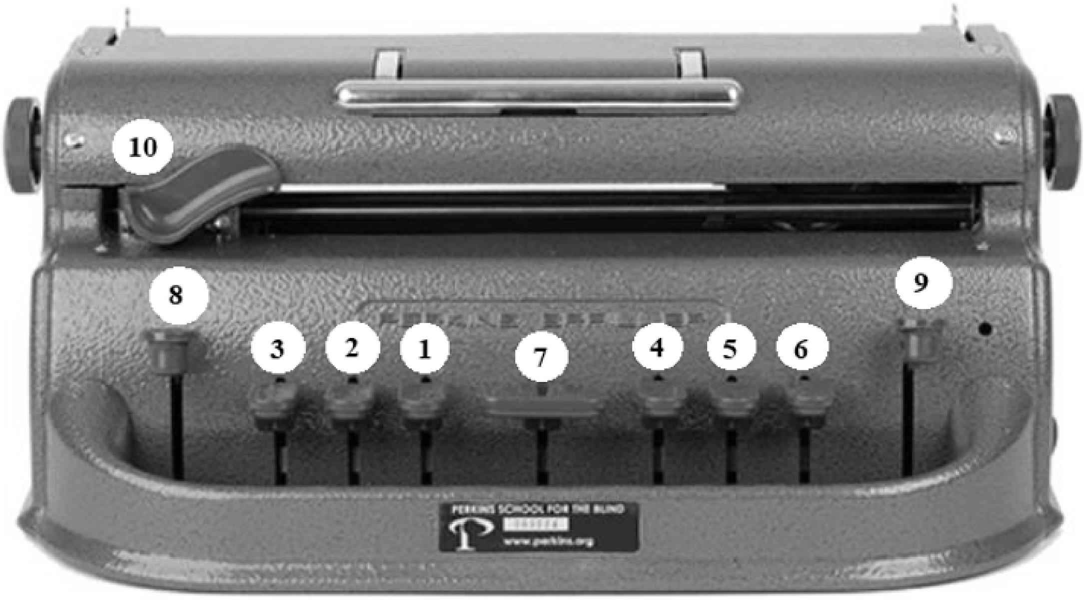


Figure 1: Perkins Brailler with its keys numbered.

As its name suggests, the Finger Braille system is based on traditional Braille. It consists of placing someone's index, middle, and ring fingers of both hands-on top of the same fingers of the person who is to receive the message. It follows the same logic as the Brailler, with each finger representing a dot in the cell, making up the Braille alphabet so that characters can be formed by tapping the fingers of the receiver (Ding, 2012). However, this system requires that both people know the Braille alphabet and its adaptation as Finger Braille, forcing them to interact with laymen via interpreters. Each of the methods listed above has advantages and disadvantages, like the deaf-blind block alphabet being simpler and easier to learn (Hersh, 2013) but much slower than the other methods. For example, tactile sign language is more difficult to master but is faster than finger Braille. Finger Braille can be learned faster than the deaf-blind manual alphabet, as it imitates the layout of a Braille typewriter (Hersh, 2013).

With the advent of novel technological paradigms, communication aids have been developed to further include people with sensory disabilities. Zdravkova et al. (2022) explored several applications of Artificial Intelligence (AI) in assistive technologies for communication and learning for different disabilities. For the d/Deaf-blind, they mention text-to-Braille translators with tactile displays, in which the Braille displays utilize AI for Optical Character Recognition, a Braille-to-text algorithm where a CNN converts a Braille line to text after image segmentation and optical Braille recognition; and Deep Neural Networks for text-to-Speech. Although the rest of their work focuses on other disabilities (i.e., dyslexia, Functional Speech Disorder, or blindness and deafness alone), the applications mentioned are cutting-edge for the use of AI to support people with sensory disabilities.

While developments have been made in other countries to facilitate the inclusion of the d/Deaf-blind, this is hardly the case in Mexico. For example, in the United States, several relay services are available for long-distance communication with a d/Deaf-blind user (NFB, 2014), and the National Center on Deaf-Blindness recognizes 52 State Deaf-Blind Projects (NCDB, 2013). In contrast, in 2008, Mexico's National Commission for Preventing Discrimination (CONAPRED) stated there was only one association, ASOMAS, in the whole country offering education, stimulation, and rehabilitation to children with d/Deaf-blindness or other "multiple challenges" (CONAPRED, 2008). Put into perspective, the USA has at least 52 times as many specialized projects for the d/Deaf-blind as Mexico (NCDB, 2013), but not nearly as many times its population. Bowen and Probst (2023) sustain the claim that teachers must have a specific skill set to effectively work with d/Deaf students with additional disabilities (which includes d/Deaf-blind children). Still, not enough teachers are qualified to do so given the growing number of d/Deaf and otherwise disabled students, and the contrasting decline in available preparation programs, even in the United States.

Adding to the requirements Bowen and Probst (2023) established for educating d/Deaf students with multiple disabilities, Olayi et al. (2023) suggest that for proper inclusion in the classroom, deaf-blind students should be provided with a special educator with Braille and (tactile) sign language skills, a teacher committed to serving those with special needs, a braillist, a sign language interpreter, and a caretaker. McKinney (2022) described a success story where the academic needs of a deaf-blind student, Ivey, were met through different communication methods, namely English-based signs, tactile symbols, Braille, and some spoken English. The student also had a one-on-one intervener (someone who specializes in aiding the deaf-blind with several tasks and is well-versed in alternative communication methods) to aid her in the classroom. The author highlights her own lack of experience and qualification for teaching the d/Deaf-blind before getting to teach Ivey, despite having years of experience with Individualized Education Programs and multiple degrees in special education. She also emphasizes the fact that having multiple disabilities generates unique conditions not experienced by those with only one of them, which calls for specialized services rather than ones that address the vision or hearing needs of the child (McKinney, 2022) solely.

A deaf-blind college student who participated in a study conducted in Ghana by Dogbe and Anku (2024) mentioned feeling like she had to carry her Braille machine everywhere for people who could not sign to her because although tactile sign language is quicker and more convenient for her, not enough people knew how to sign. An assistant lecturer in the same study stated that "including the deaf-blind learner in higher education warrants some essential resources, and it is time-consuming. The student needs assistive devices such as a braille machine and refreshable braille display, tape recorders, and note takers" (Dogbe and Anku 2014:128), further highlighting the special requirements that including a deaf-blind student in the classroom implies. All these examples indicate that teaching a d/Deaf-blind student requires plenty of resources and personalized services and that learners might benefit from having as many tools and communication methods available as possible.

Riccobono and Morrow (2022) conducted a survey of State Deaf-blind projects in the USA, which pointed out how the mere existence of these institutions did not guarantee that students would get access to a qualified intervener since there are not enough intervenors to meet the demand for their services. The authors mention that barriers keeping potential interveners from earning their qualifications include a severe lack of financial incentives, the long and tedious process of completing a portfolio, the general lack of interest in special education, and the retention of interveners due to the same reasons (Riccobono and Morrow, 2022). It is important to remember that this is the case in a country where several resources exist for the d/Deaf-blind. In Mexico, beyond a simple lack of specialized institutions or qualified professionals, a bulletin by the Mexico City Human Rights Commission (CDHCM, 2020) stated that there is not even specific data about deaf-blindness in Mexico other than what can be inferred from the intersection between data on blindness and deafness. This represents another obstacle for those seeking to assess and solve this issue.

In a study conducted in South Africa, which has similar struggles as Mexico regarding the lack of resources for the d/Deaf-blind, Manga and Masuku (2020) found that educators of d/Deaf-blind children struggled with teaching them due to under-preparedness and lack of support. In the words of one of the teachers interviewed, "It is a barrier because even though they can't see, they can't talk, they can't hear, you still have to teach them sign language… But they won't understand" (Manga and Masuku, 2020). Similarly, d/Deaf-blind participants of a study done in Spain by Rodríguez-Jiménez et al. (2022) said that educational centers lacked the appropriate resources to cater to their needs, using oral language when teaching, which hindered their access to curricular content.

The lack of specialized tools enabling d/Deaf-blind children to access education remains a global problem due to their high costs and specific training requirements. This issue extends beyond the demographic of d/Deaf-blind students, affecting those with various other disabilities. A 2018 study conducted in the Czech Republic highlighted this problem, revealing that students with disabilities face a 1.5 times higher chance of experiencing unsuccessful completion of university studies (Mazouch et al., 2018). Ideally, d/Deaf-blind children should have access to specialized materials and teachers with adequate preparation as Teachers of the Deaf With Disabilities. Still, particularly in the context of Mexico, the existence of infrastructure for d/Deaf-blind children is alarmingly poor. According to 2018 data on the Mexican education system, Deaf-blind children between 5 and 17 years old have a 33% reduced probability of school enrollment than children with no disabilities (WFDB, 2024). Mexico also had one of the highest gaps (15%) between the percentage of deaf-blind children enrolled in school and that of children with other disabilities (WFDB, 2024), indicating a need for solutions specific to deaf-blind children in the country. The lack of said solutions is a potential factor for school dropout of children with the condition, alongside the inexistence of communication tools such as Braille typewriters for Spanish-speaking students and professors.

An international survey carried out by the World Federation of the Deaf-blind, or WFDB (2023), found that the primary school net attendance of children with deaf-blindness was 20%, but 75% for non-disabled children overall. Moreover, even for 20% of deaf-blind children who do attend primary school, it is a challenge to measure the quality of the education received. The same survey found that deaf-blind children were about three times less likely to be classified as "developmentally on-track" as per the World Childhood Index; that is, only 20% of deaf-blind children fit the literacy-numeracy, physical functioning, social-emotional development, and learning criteria expected of their age group. These numbers were also found to correlate to the country's income, ranging from 40% deaf-blind children who are developmentally on-track in upper-middle-income countries to only 13% in lower-income ones (WFDB, 2023).

Thus, the development of a low-cost, intuitive tool based on AI techniques and additive manufacturing that does not require previous knowledge from non-disabled people (and is adapted to the way the 64 combinations in 6-dot Braille are used in Spanish, including characteristic letters such as ñ, á, é, í, ó, ú and ü) will be particularly useful to further the inclusion of the d/Deaf-blind in a country lacking resources and visibility for the condition. This paper offers an alternative to bridge the gulf between what is needed and what is available to d/Deaf-blind students and facilitate their integration in Mexico into non-specialized schools through the development of a bidirectional Braille-speech communication system based on the design of a CNN for speech recognition. The development of an algorithm capable of performing both text-to-speech and Finger-Braille-to-text conversion is presented as well, allowing for real-time communication with no intermediaries to serve as a bidirectional tool for deaf-blind students.

MATERIALS AND METHODS

Several stages were followed in the development of the proposed device, which will be described in the following subsections. First, a Signal Acquisition process was required to train an AI algorithm to enable the device to recognize speech. The acquired signals needed to be processed to extract the pronunciation variations existing in Spanish-speaking persons. Then, a Braille-Speech Conversion Model encompassing both the speech recognition algorithm and the Braille-text conversion code was generated. Afterward, the complete system was integrated into a physical prototype fabricated with additive manufacturing techniques.

Signal Acquisition

To train the CNN designed for speech recognition, it was necessary to build a database containing audio samples. Due to time constraints, the scope of this database was limited to audios of single letters. Therefore, each of the 24 volunteers was asked to slowly recite the Spanish alphabet, and a sample of the noise in the room at the moment was recorded as well. All recordings were done in a single, empty room using a condenser microphone and a commercial digital audio workstation. Prior to the acquisition of each volunteer's speech data, they were asked to sign a consent form for the use of their voice and perform a couple of tests (namely saying the letters "f," "j," and "p") to adjust the gain and noise gate settings to ensure that no clipping would occur, and no unnecessary noise would be recorded. The sample rate was set to 44.1kHz throughout the entire process. Personally identifiable information such as name, age, or gender was not saved along with the samples, making all audio anonymous. Such information, however, was collected separately for analytic purposes to determine whether any demographics were underrepresented in the data set. Race was not assessed as a factor.

Signal Processing

The raw data collected was edited to last exactly two seconds per letter, with the useful data roughly in the middle of each clip. Each letter was then exported as a separate audio file, with the name of the letter contained and the number of the volunteer to facilitate postprocessing. Since the data collected was insufficient to properly train a neural network (it amounted to 22.4 minutes of significant data once the pauses were cut out), the original 672 audio clips were modified in three different ways as a data augmentation method.

The first way involved pitch change by raising or lowering the frequency of the audio by two semitones; the second consisted of changing the speed to 120% and 70%, and finally, the third was achieved by performing a convolution of each audio with the room impulse response (RIR) of six different spaces. Added to the original files, this yielded a total of 7,392 samples or approximately four hours of audio data. Pitch and speed change required the use of the library "librosa" (McFee et al., 2015), while the last type of data augmentation utilized the room impulse response database generated by MIT's *Computational Analysis Lab (Traer* and McDermott, 2016). However, only six RIRs were extracted from this library: the inside of a car, an office, a hallway, a cafeteria, an atrium, and a theater. Performing the convolution of the audios with these RIRs yielded new files that simulated the effect of recording in these different spaces. The new files were labeled accordingly, and their spectrogram was obtained through a Fast Fourier Transform. Given how CNNs excel at classifying images, the team opted for using images of the spectrograms instead of using the audio directly.

Database Description

After the data augmentation process, the final database consisted of 7,392 images. Figure 2 shows an example of the letters represented in Fourier space, plotting frequency against time. These images were then divided into folders according to their class. These included the 27 letters of the Spanish alphabet as well as a sample of the recording room with no speech produced, intended to help the algorithm discern between useful data and background noise. A holdout validation process was carried out, so the samples were split into three additional folders: train, validation, and test, according to the volunteer number on the sample. 70% of the data was reserved for training, 25% for validation, and the remaining 5% for testing the CNN.

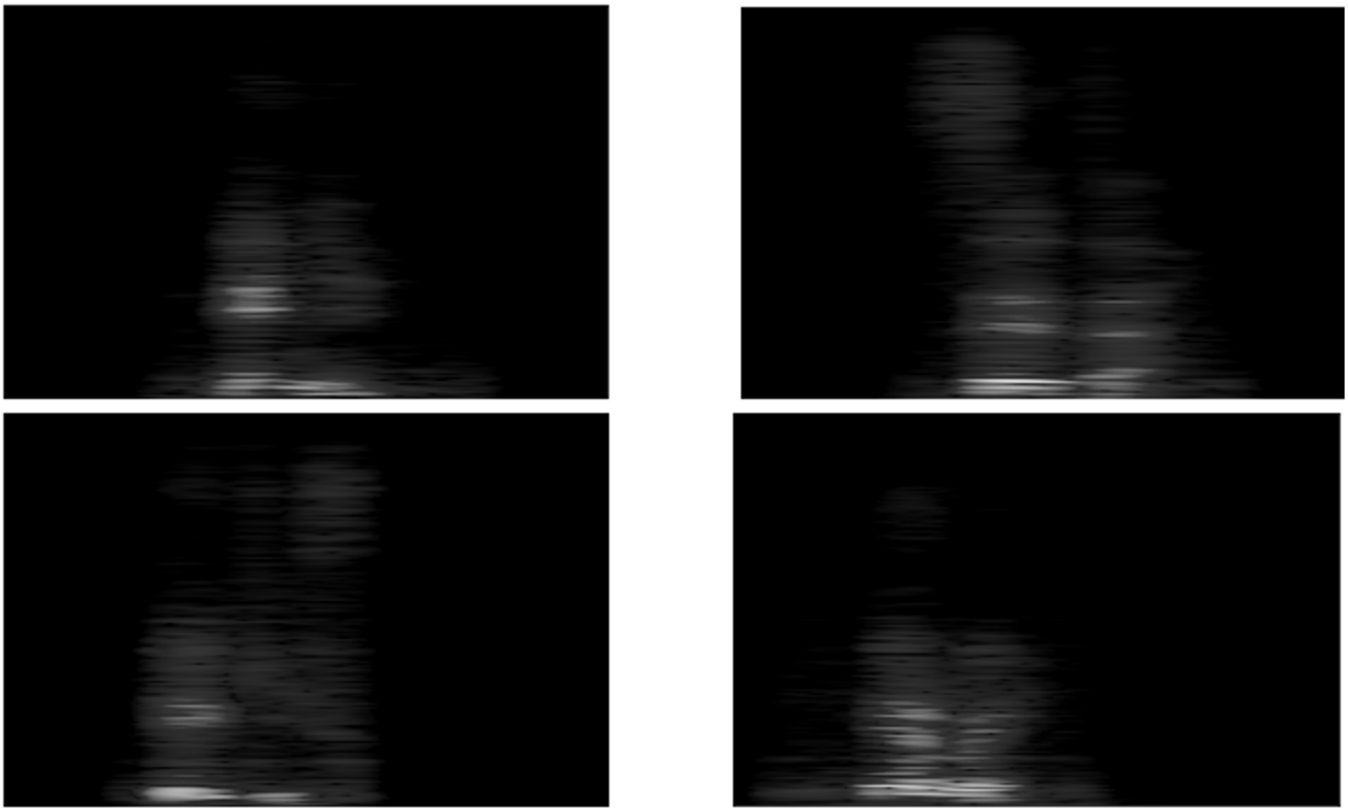


Figure 2: Spectrogram of letters Ñ (top left), Z (top right), X (bottom left) and R (bottom right)

Physical Prototype

As for the device itself, the 3D model was created using Computer Aided Design software and was then printed in polylactic acid (PLA) filament. The design is essentially a Braille keyboard with added coin vibration motors and hand rests. This design (depicted in Figure 3) was chosen since the layout of the keys closely resembles that of a Braille typewriter, aiming to aid the d/Deaf-blind user through familiarity. The hand rests added to allow the hand to adopt a comfortable position when touching the Braille keys while simultaneously keeping the intermediate phalanges of the index, middle, and ring fingers in contact with the coin motors, which vibrate in response to voltage changes to transmit Braille the way the user would receive it when communicating in Finger Braille. Key design elements were considered for its development, utilizing computer-aided design and additive manufacturing to generate the prototype shown.

In Figure 3, keys 1-6 are for writing Braille, as the Perkins machine shown in Figure 1. The key labeled "M" activates the microphone for Speech-to-Braille translation; the key labeled "T" is used to indicate the end of a Braille character, and the "S" key is used to initiate typing and mark the Braille message as complete before performing Braille-Speech conversion. Finally, elements "a" through "b" are the coin motors that transmit the message in Braille to the deaf-blind person.

Other relevant materials used besides PLA were push buttons, coin vibration motors, and polyester foam to help isolate the vibrations of each motor.

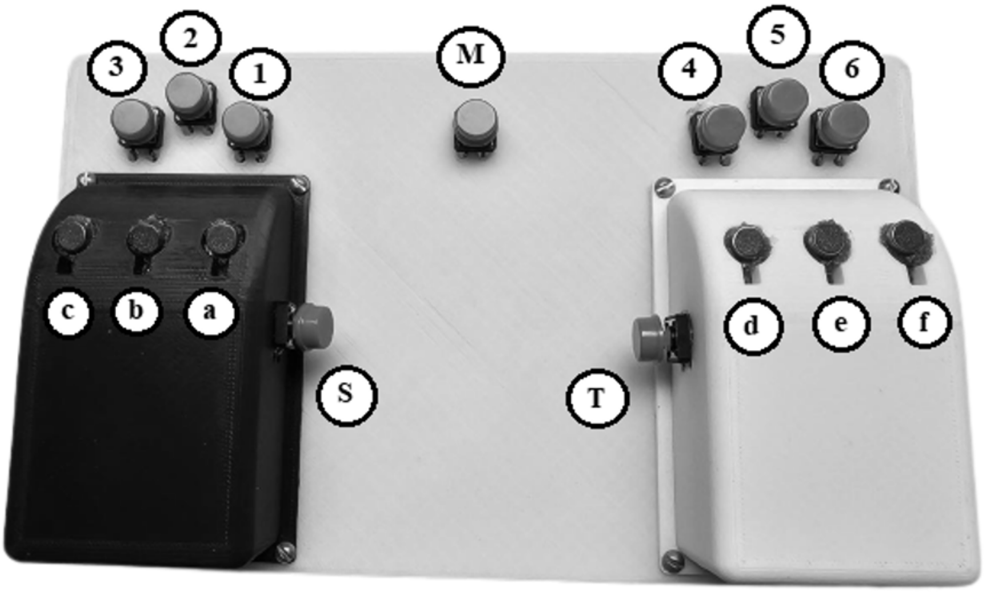


Figure 3: Fully assembled prototype

Braille-speech Conversion Model

To allow for bidirectional communication, both translation pathways were developed separately. Prior to the making of the physical prototype, the Braille-to-text pathway was developed using a computer keyboard as Braille input, likening certain keys to each of the six dots in a Braille cell. Then, since several keys were pressed simultaneously, they were rearranged alphabetically. At that point, a Python dictionary allowed the program to decode the Braille message and join it into a single string. Text-to-Speech was then implemented to convey the message to the hearing party. Once the physical device was ready, the input was simply changed to receive a Serial message sent by Arduino, which controlled the physical buttons and actuators. Three "control buttons" were implemented in the device to facilitate processing: one for activating the microphone, one for marking the end of a character, and one for activating typing mode and sending finger Braille messages. This enabled the code to run smoothly and without interruptions.

As for the speech to Braille pathway, a CNN was developed and trained using the dataset described above. Training is performed to obtain adequate *weights* between a neuron (the smallest processing unit within the CNN) and its inputs. Weights denote the strength of the link between an input and a certain neuron (Dongare, Kharde, and Kachare, 2012), where inhibitory connections are represented by negative weights and excitatory connections by positive ones. Weights are randomly initialized, after which an iterative process of analyzing the loss and updating the weights begins.

The loss value was obtained using categorical cross-entropy, which can be defined as shown in Equation 1 (Rusiecki, 2019):

$$E_{CC} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C}\left(p_{ic}loglog\left(y_{ic}\right)\right) \qquad (1)$$

In equation 1, E_{CC} is the Categorical Cross Entropy loss, N is the number of label/value pairs in the training dataset, y_{ic} is the output, and C is the number of categories (Rusiecki, 2019). To improve the performance of CNNs, a backpropagation algorithm can help reduce errors by sending feedback to the previous layer. Mathematically, this process can be described by Equation 2 (Li et al., 2012).

$$x_{k+1} = x_k - \eta_k g_k \qquad (2)$$

where *xk* is a matrix containing current weights and thresholds, η_k is the model's learning rate, and g_k is the current function's gradient (Li et al., 2012). Kingma and Ba's (2015) optimizer, Adam, was selected to optimize the learning process. Adam stands for Adaptive Moment Estimation; it estimates the previous momentum and gradient to update the learning parameters in every iteration. It also adapts the learning rate (essentially the magnitude of the changes made to the model with each iteration) for every parameter, depending on the momentum and gradient. To estimate the momentum, Adam uses Equations 3 and 4.

$$m_t = \beta_1 m_{t-1} + \left(1-\beta_1\right) g_t \qquad (3)$$

$$y_t = \beta_2 v_{t-1} + \left(1-\beta_2\right) g_t^2 \qquad (4)$$

where m_t and v_t are moving averages, g_t is the gradient of the current mini-batch, and β_1 and β_2 are hyperparameters with values of 0.9 and 0.999, respectively. The network architecture consists of four convolution layers of different dimensions with a 10x10 filter; these layers have a Rectified Linear Unit (ReLU) activation function, which eliminates all the negative values, preserving only the positive ones. The equation of ReLU is shown in Equation 5 (He, 2018).

$$f\left(x\right) = \left(0, x\right) \qquad (5)$$

Each convolution layer is followed by a *max pooling* layer that obtains the main characteristics of the image by calculating the maximum value of every image matrix. After this filtering stage, a *flattened* layer reshapes the image into a string. Afterward, a dense layer of 512 neurons with a ReLU is the last stage before the output layer, which, a dense layer of 28 neurons with a softmax activation function to obtain 28 categories corresponding to every letter in the Spanish alphabet plus a "silence" category. The *softmax* equation is shown in Equation 6 (Banerjee et al., 2021).

$$softmax(z)_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \qquad (6)$$

where z_i represents the input vector, and zj contains the values from the neuron output layer. The softmax activation function calculates the relative probabilities of each class, so the input will be categorized as the class with the highest probability. The proposed CNN structure is depicted in Figure 4.

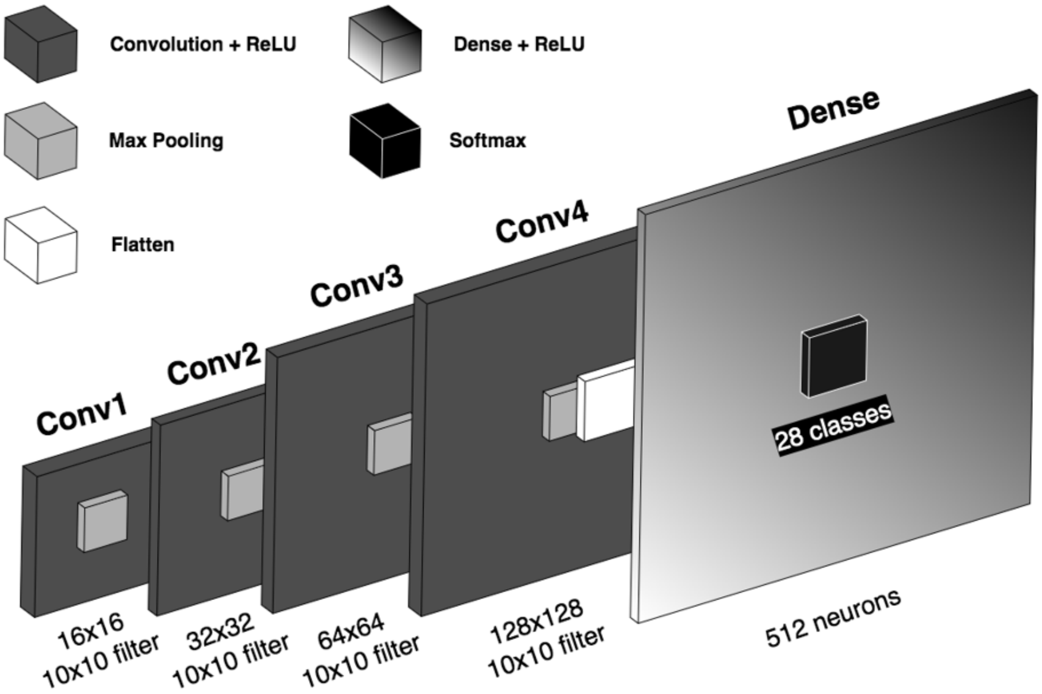


Figure 4: Architecture of the CNN

This algorithm delivered the spoken input as a String, which was then processed in Python. It used a dictionary like the one used for Braille to text but converted characters to the numbers of their corresponding Braille dot patterns. The Braille message was then sent to an Arduino board through a Serial port to deliver the instruction to activate or deactivate each vibration motor.

Mini-batch gradient descent was chosen over other methods for training the algorithm since it performed better over numerous training sessions. The final weights for the model were obtained with 150 epochs, which was empirically determined as the optimal number given that it allowed the model to converge without overfitting the weights to the training data. 5 steps were used per epoch, with a batch size of 17 samples for training and 6 samples for validation.

For this process, the training sample (images that the algorithm has never worked with) is run through the network and multiplied by the weights to obtain the highest output node (the category or class). It is then compared to the validation sample (images whose class the network already knows), and the weights are adjusted to improve the prediction. This process is repeated in every epoch until the prediction has the required accuracy.

The functioning of the device is shown in Figure 5. An interaction begins with the deaf-blind user typing in Braille (number 1 in Figure 5). Afterwards, the Braille message is translated and played on built-in speakers (number 2). Then, the hearing person replies by pressing the microphone activation button and speaking (3). Lastly, the message is translated by the CNN and text to Braille algorithms before being transmitted to the user via integrated vibration motors (4). This cycle is simply an example of how an interaction could occur, but Braille typing and spoken input can occur in any order if needed.

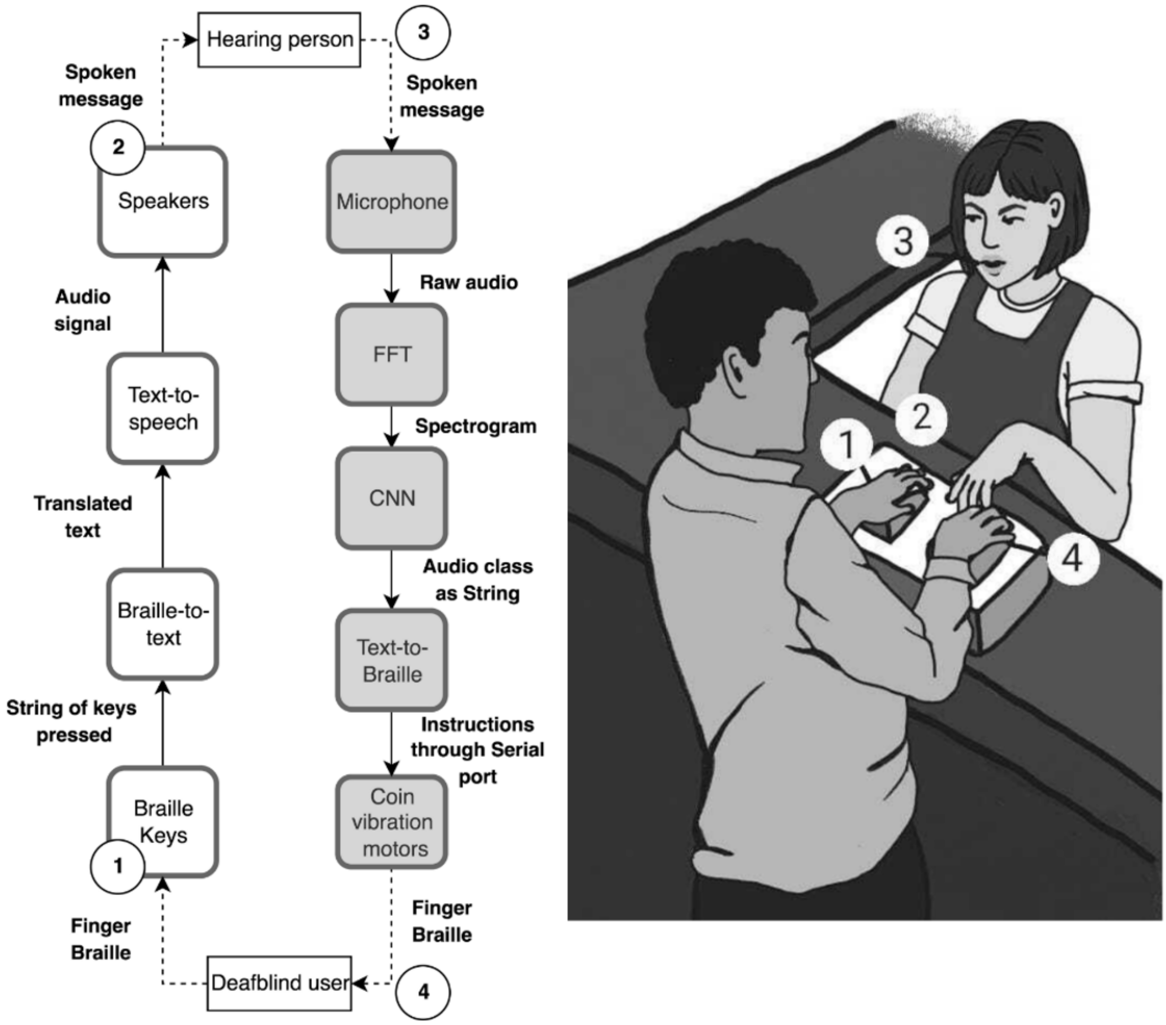


Figure 5: Diagram of device usage

RESULTS

Before assessing the performance of the device as a whole, the results of the CNN generated by the authors are presented. The accuracy graph of this model is shown in Figure 6, and its corresponding loss graph in Figure 7.

The model's accuracy with the training dataset approached 80%, while the classes of testing data (which the model had not learned from) were predicted with an accuracy of 60%. For reference, the team also tested Google Cloud Speech-to-Text. Since the authors' CNN was not designed to translate full sentences but only individual letters, Google's Speech-to-Text was tested only for its ability to recognize individual letters. The AI was used to predict the class of letters in the Spanish alphabet a total of 135 times, having a 60.74% accuracy, with 16.3% incorrect predictions and 22.96% instances of failure to classify speech at all. It must be noted that most of this

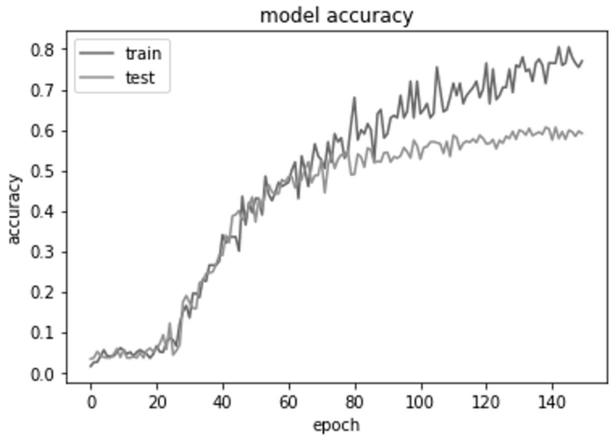


Figure 6: Accuracy graphed during training of the CNN

algorithm's mistakes stemmed from its being mainly designed to identify words instead of isolated letters. Therefore, when certain letters had the same sound as a word (for example, the Spanish pronunciation for the letter "c" and the word "se," "z" and "seta," or "d" and "de"), Google's algorithm often picked the word over the letter. Letters that did not have a homophone were identified correctly more often.

Three different trials were performed to test the device itself and the integration of its translation software. The first measured the performance of random subjects (first-time users) who were taught a few Braille letters and then asked to identify them using the device. This assessed the device's text-to-Braille algorithm and the adequacy of the Braille transmission system. The second trial tested the comprehension capacity of a deaf-blind user through the device, testing the Braille transmission portion of the algorithm. The third trial evaluated the written accuracy of a user, evaluating the ease of use of the physical prototype as well as the Braille-to-text communication pathway.

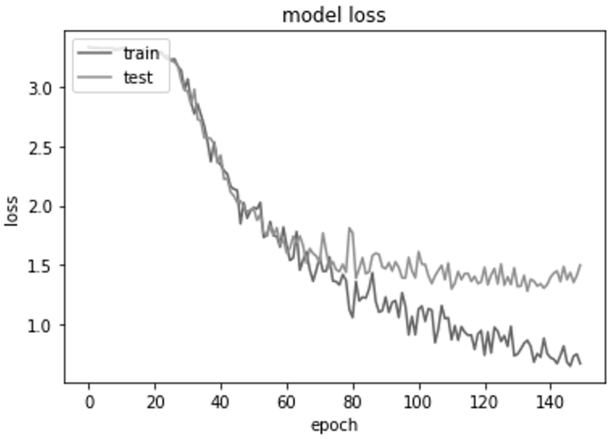


Figure 7: Loss graphed during training of the CNN

First Trial – Usability

20 subjects, including teachers, students, and parents at Tecnológico de Monterrey, were randomly chosen to learn 4 letters in Braille: a, b, c, and d. Age and gender varied among subjects but were not registered. All participants were sighted and hearing to facilitate this trial stage and guarantee that the subjects had no prior experience with Finger Braille. After approximately 2 minutes of learning the letters, they were told they would feel 4 letters in random order on the device and that they may or may not appear more than once. They were instructed to say the letters they perceived as they felt the vibrations. The point of this trial was to prove that a person without any previous experience with Finger Braille or Assistive Technology could easily adapt to the device and that the Braille transmission method used was adequate. From the trial, 3 out of 20 subjects made 1 mistake, which could be caused by the lack of concentration of the subject towards feeling the vibration on their phalanges. One of the subjects who made a mistake reported not being able to place his fingers in the required position due to his hands being too large for the device's design. He was asked to repeat the procedure, this time placing his fingertips on the coin vibration motors instead of his phalanges, and he got them all correct (the second attempt was not reported as part of the 20 samples). This suggests that the design and measurements of the hand rests are a possible area of opportunity instead of the algorithm itself or the idea of using vibration motors for this application. Apart from the mistakes, the other 17 subjects, with a perfect score, demonstrate that the ability to read Braille from the device can be easily learned. Deaf-blind people may have an advantage in learning how to use it since they are more used to employing their sense of touch to gather information, and many will already have at least a basic knowledge of Braille.

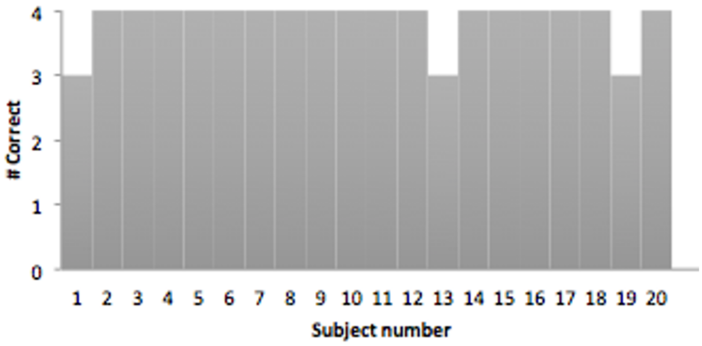


Figure 8: Number of letters identified correctly per subject

Using the Wald method (described by Equation 7) and counting each letter as a separate test (adding up to 80 binary instances in total), the confidence interval of this sample is 0.963±0.042 at a confidence level of 95% (Montgomery, 1991).

$$CI = \hat{p} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \qquad (7)$$

where CI is the confidence interval, "$\hat{p}$" is the point estimator of the parameter "p" and equals the number of nonconforming observations over the total of the sample, "n" is the size of the sample, and $Z_{\alpha/2}$ is the Z value at half of the tolerance α (which was equal to 5 given the confidence level).

Second Trial – Communication

The second trial aimed to demonstrate effective communication between a user with full visual and hearing capabilities and a deaf-blind user. In this case, the non-disabled user says a command out loud to the device to convert speech to text and text to Braille. After the conversion of the sentence, the "deaf-blind" user (since the team was unable to find a d/Deaf-blind volunteer, this consisted of a hearing and sighted person wearing a blindfold and noise-canceling headphones to nullify possible auditory or visual stimuli) complies with the command given. This trial was performed 20 times with different kinds of simple, short commands; in all of them, the command was understood and followed exactly. This trial demonstrates that phrases can be communicated by the sender to the "deaf-blind" user, verifying the integration of all systems in the prototype and the viability of communication.

Commands were issued in Spanish, but their English translation is presented in Table 1, along with the result.

Command	Result
Touch your nose	Success
Raise your left hand	Success
Smile	Success
Turn to the right	Success
Raise a finger	Success
Touch your mouth	Success
Nod	Success
Stand up	Success
Point down	Success
Touch your ear	Success
Touch the headphones	Success
Clap	Success
Make a fist	Success
Touch your forehead	Success
Spin in place	Success
Touch the table	Success
Lift your pinky	Success
Touch your right shoulder	Success
Take off the headphones	Success

Table 1: Commands and responses obtained in the first trial

The subject was not asked whether they correctly identified every individual letter, but had they not, it appears the context was sufficient to allow them to understand and respond to the commands.

Third Trial – Accuracy

The final trial was to test the potential typing accuracy of a deaf-blind user. A pangram (a phrase that contains all the letters in a given alphabet) was dictated, and the times a mistake happened were counted. The pangram in Spanish was: "*Zumba la wifi yoyo jala con gofres él extraño virus huirá pero los kilos quedarán*". For this last trial, the objective was to detect if the device's hardware or software had any noticeable flaws during translation, as well as verify if it was easy to use for someone who already knew Braille. Two users familiar with Braille were asked to type the pangram word by word, including written accents and the uppercase Z, using no abbreviations. Accuracy was measured by calculating the number of Braille dots per word and comparing the total to the dots typed correctly by the user. Both present and absent dots were counted, making up a total of 6 points per character. For example, letters h (⠓) and r (⠗) differ by one point (dot 3), but letters s (⠎) and l (⠇) differ by two: dot 1 and dot 4. Each letter was counted as six dots, except for the initial Z, which requires an additional modifier (6 more dots) to become a capital letter.

User 1 was 95.833% accurate, and user 2 was 98.529%, out of a total of 408 dots in the phrase. Figure 9 compares the number of points gotten per word by each user, along with the total points for reference.

To further facilitate understanding of how mistakes occurred and how they were measured, Table 2 shows the specific Braille patterns expected per word and those written by each user. The number of points deducted (labeled "dots missed") is also presented next to each user's Braille output. Additionally, characters where the user made a mistake are underlined. In cases where a full character was omitted, 6 points were deducted, and the space was left blank in the table.

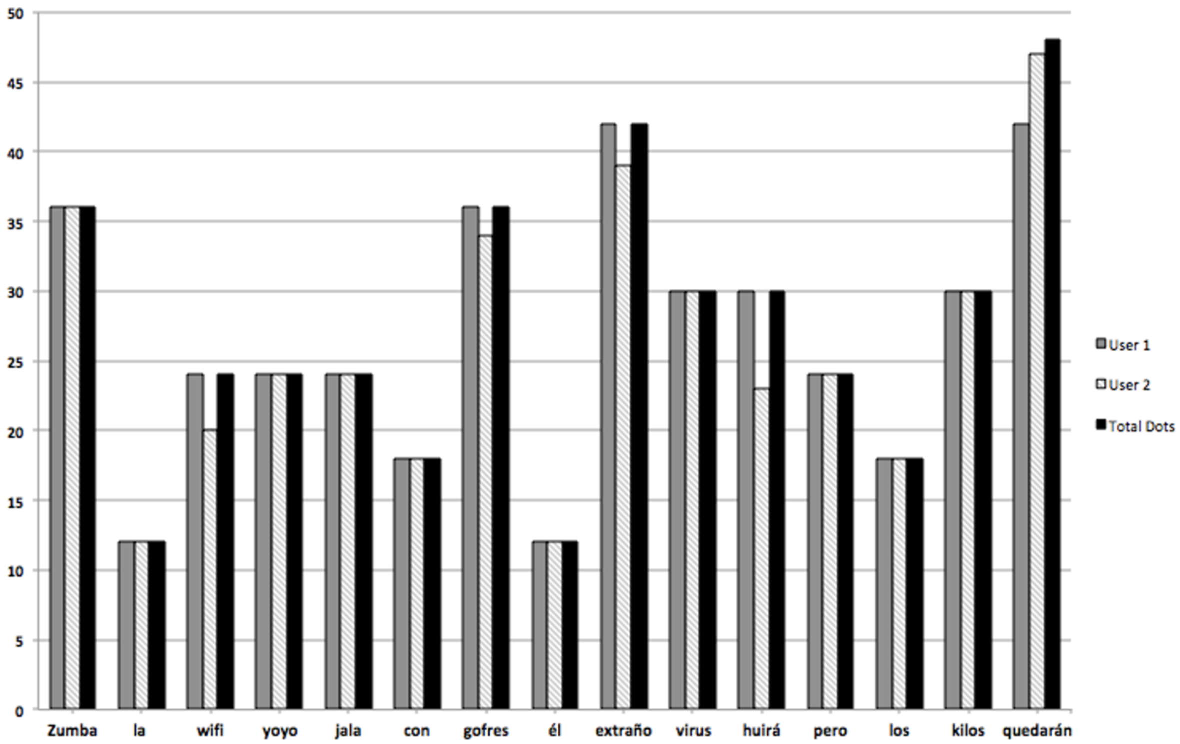


Figure 9: Correct dots per word by the user

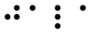
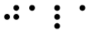
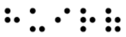
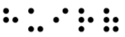
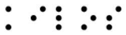
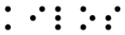
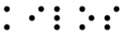
Word	Expected Braille	User 1 Braille	Dots missed	User 2 Braille	Dots Missed
Zumba			0		0
la			0		0
wifi			4		0
yoyo			0		0
jala			0		0
con			0		0
gofres			2		0
él			0		0
extraño			3		0
virus			0		0
huirá			7		0
pero			0		0
los			0		0
kilos			0		0
quedarán			1		6

Table 2: Detailed Braille output comparison for the third trial

DISCUSSION

As explained in the results section, the prototype successfully helped users without any Braille background to quickly learn and understand different letters. This trial proved that the device could be potentially used by people in the process of becoming deaf-blind to facilitate the learning of Finger Braille. The device could also be used by d/Deaf-blind students, as the second trial proved that users could understand simple commands without visual or auditory stimuli, using a neural network that translates full sentences. The final trial demonstrated that users with some Finger Braille background could write full sentences and all the Spanish alphabet letters with great accuracy, which means that deaf-blind people can use the device to establish communication with people without sensory disability.

One of the limitations of this study was the team's inability to find d/Deaf-blind volunteers. As stated in previous sections, Mexico severely lacks infrastructure in this regard, which made it hard to know where to contact potential users in the first place, so the very problem hindered the project it attempts to mitigate. The next step for this project is to test the prototype with d/Deaf-blind people to consider the cultural and experiential aspects of Deaf-blindness as a Culturally Deaf-blind user would experience the use of the device differently due to their unique experiences with Deaf-blind culture.

Another possible limitation of adopting this prototype in schools is that the students must keep their hands on the device to avoid missing out on information during the lesson. This could hinder taking notes. One way to solve this situation is to adapt the prototype to wearable gloves using haptic technology to allow the d/Deaf-blind user to fully use their hands. Alternatively, the device could be adapted to allow for simultaneous rather than alternating speech-to-Braille reception and Braille-to-Text conversion, connecting the system to a laptop to store the notes without forcing the d/Deaf-blind student to stop "listening" through the vibration motors when switching to a note-taking method.

Compared to existing technologies, an area in which the device is lacking is coverage since this project focuses on severe to profoundly deaf and blind users. Estimates place the deaf-blind population at up to 15 million people worldwide (Kassem et al., 2022), but as previously stated, deaf-blindness exists in a spectrum of visual and auditory impairment, and some deaf-blind people have CIs, which somewhat re-establish hearing as a possible communication pathway. These users could benefit from technologies directed at solely blind individuals, but the potential impact of implementing CIs or disregarding the existing degree of hearing impairment should not be overlooked.

Liu et al. (2023) proposed a wearable Braille typing system utilizing triboelectric nanogenerators (TENG) to bypass the need for an additional power supply, which could be a downside to the current project, given the long consecutive periods during which a d/Deaf-blind student attending a regular school would have to use the device proposed in this work. Substituting the push buttons for TENGs or otherwise reducing the system's energy requirements could help extend battery life to minimize this problem.

Vincent et al. (2021) implemented a method of face-to-face communication between d/Deaf-blind users and a hearing and sighted layperson using a Focus 40 Blue 5th gen Braille display connected to a smartphone. The applications of their method and the present project are quite similar, differing mainly because the cost of the Focus 40 Blue alone exceeds USD 3,000 (Vincent et al., 2021), making it an inaccessible option for many schools and families in Mexico.

The trials also helped to identify areas for improvement. For example, in the first trial, one of the subjects could not feel the vibration motors because they were not in direct contact with their phalanges, so there is still room for improvement in the design of the hand rests or, as stated before, adapt the prototype to a wearable device to ensure continuous contact. The neural network also has room for improvement to increase the prediction rate of the alphabet letters. As mentioned in the results section, the prototype can integrate third-party AI algorithms such as Google Cloud speech recognition. This could be very helpful in classrooms for d/Deaf-blind students to follow the lesson in real-time, but the downside is that they would have to keep their hands on the device throughout the lesson.

There are additional areas in which improvements are being considered: design, communication process, and portability. The tests comparing both speech recognition algorithms suggest two possibilities: the device, paired with a commercial algorithm like Google's, can guarantee a high accuracy for communication. However, it might be advisable to develop a more sophisticated algorithm specializing in identifying individual letters for the application of individuals learning Braille on their own with the aid of the device. This could be achieved by generating a more comprehensive database to train the CNN presented in this work to recognize different types of voices and accents in multiple environments.

CONCLUSION

The inclusion of minorities is a complicated issue. This is especially true for particularly small communities, such as the Deaf-blind. The smaller a community is, the fewer financial and social incentives companies and governments will see to develop solutions tailored to their needs, which is why they often go unaddressed. The designed device represents an additional tool for this group as well as people without severe sensory impairments since anyone can use the device. It can therefore aid those who are gradually becoming deaf-blind to learn finger Braille during their transition into deaf-blindness; be a real-time communication tool to allow profoundly deaf-blind users to gain more independence; grant d/Deaf-blind students access to hearing and sighted institutions; or make it easier for people without sensory impairment (including intervenors, family members, friends, public servants, healthcare professionals and anyone with interest in accessibility) to learn Finger Braille. Despite the limitations described in prior sections, the device showed great translation accuracy and ease of use, which indicates its viability for the applications above.

In the future, the device could be tested against existing assistive technology for solely d/Deaf users, as a form of automatic subtitles to what is being said in a hearing classroom, or as a form of Alternative Communication for non-verbal autistic people and other sectors who have trouble speaking. The physical disposition of the device can also make it easier to type than a traditional keyboard for those with limited mobility since all keys are designed to be in contact with the corresponding fingertip in a resting hand position. In other words, small variations to the device's design and algorithm could address the problems of vastly different sectors in future projects. Still, it is imperative to get direct feedback from d/Deaf-blind users and educators of the d/Deaf-blind to validate any future changes to the device to optimize it for its original purpose.

ACKNOWLEDGMENTS

We thank Dr. Alejandro Santos Díaz for providing guidance, Dr. Adrián Octavio Ramírez Morales for providing feedback and reference materials, and all volunteers for participating in the project's development. We also thank Tecnológico de Monterrey for lending us their laboratories and support.

REFERENCES

Ask Larsen, F. and Damen, S. (2014) 'Definitions of deaf-blindness and congenital deaf-blindness', *Research in Developmental Disabilities*, Vol. 35, No. 10, pp. 2568–2576. https://dx.doi.org/10.1016/j.ridd.2014.05.029

Ayton, L., Galvin, K., Johansen, L., O'Hare, F. and Shepard, E. (2023) 'Awareness of Usher syndrome and the need for multidisciplinary care: A cross-occupational survey of Allied Health Clinicians', *Journal of Multidisciplinary Healthcare*, Vol. 16, pp. 1927–1936. https://dx.doi.org/10.2147/jmdh.s411306

Banerjee, K., Prasad, V., Gupta, R., Vyat, K., Anushree H. and Mishra, B. (2021) 'Exploring alternatives to Softmax function', *Proceedings of the 2nd International Conference on Deep Learning Theory and Applications*, *(DeLTA, 2021)*, Paris, Vol. 1, pp. 81–86.

Blenkhorn, P. (1995) 'A system for converting braille into print', *IEEE Transactions on Rehabilitation Engineering*, Vol. 3, No. 2, pp. 215–221. https://dx.doi.org/https://doi.org/10.1109/86.392366

Bowen, S.K. and Probst, K.M. (2023) 'Deaf and hard of hearing students with disabilities: An evolving landscape', *Education Sciences*, Vol. 13, No. 7, pp. 752–767. https://dx.doi.org/10.3390/educsci13070752

Castiglione, A. and Möller, C. (2022) 'Usher syndrome', *Audiology Research*, Vol. 12, No. 1, pp. 42–65. https://dx.doi.org/10.3390/audiolres12010005

CDHCM (2020) It is necessary to consider measures and strategies that include deaf-blind people in the new normal | CDHCM, [Online], Available: https://cdhcm.org.mx/2020/06/cdhcm-es-necesario-que-en-la-nueva-normalidad-se-contemplen-medidas-y-estrategias-que-incluyan-a-personas-sordociegas/ [19 Feb 2024].

CONADIS (2018) Deaf-blindness in México: Facts for June the 27th, International Deaf-blindness Day | *Government of Mexico*, [Online], Available at: https://www.gob.mx/conadis/es/articulos/la-sordoceguera-en-/mexico-datos-por-el-27-de-junio-dia-internacional-de-la-/sordoceguera?idiom=es [9 Jul 2023].

CONAPRED (2008) 'Una Esperanza Para Ángel' [Hope for Ángel] | *National Council for Preventing Discrimination,* [Online], Available: https://www.conapred.org.mx/index.php?contenido=noticias&id=1854&id_opcion=&op=447 [10 Jul 2023].

Ding, X. (ed.) (2012) *Advances in Character Recognition: Finger braille recognition system*, Beijing: IntechOpen.

Dogbe, D.S.Q. and Anku, F.K. (2024) 'Prospects and challenges of educating a deaf-blind student in a university in Ghana', *European Journal of Education Studies*, Vol. 11, No. 1, pp. 118–134. http://dx.doi.org/10.46827/ejes.v11i1.5157

Dongare, A., Kharde, R. and Kachare, A. (2012) 'Introduction to artificial neural network', *International Journal of Engineering and Innovative Technology*, Vol. 2, No. 1, pp. 189–194.

He, J., Li, H., Xu, J. and Zheng, C. (2018) 'Relu deep neural networks and linear finite elements', *Journal of Computational Mathematics*, Vol. 38, No. 3, pp. 502–527. https://dx.doi.org/10.4208/jcm.1901-m2018-0160

Hersh, M. (2013) 'Deaf-blind people, Communication, independence, and isolation', *Journal of Deaf Studies and Deaf Education*, Vol. 18, No. 4, pp. 446–463. https://dx.doi.org/10.1093/deafed/ent022

INEGI (2020) [Population with disability, with limitations for daily activity and some mental problem or condition, by federated state and quinquennial age groups according to sex | *Interactive-generic tabulations*, [Online], Available: https://www.inegi.org.mx/app/tabulados/interactivos/?pxq=Discapacidad_Discapacidad_01_29827fe7-b1cd-4bd2-81d6-9d08bda47df8 [9 Jul 2023].

Kassem, K., Caramazza, P., Mitchell, K., Miller, M., Emadi, A. and Faccio, D. (2022) 'Real-time scene monitoring for deaf-blind people', *Sensors*, Vol. 22, No. 19, p. 7136. https://dx.doi.org/10.3390/s22197136

Kingma, D.P. and Ba, J. (2015) 'ADAM: A method for stochastic optimization', *Proceedings of the 3rd International Conference for Learning Representations (ICLR, 2015)*, San Diego.

Li, J., Cheng, J.-H., Shi, J.-Y. and Huang, F. (2012) 'Brief introduction of back propagation (bp) neural network algorithm and its improvement', *Advances in Computer Science and Information Engineering*, Vol. 169, pp. 553–558. https://dx.doi.org/10.1007/978-3-642-30223-7_87

Liu, W., Yu, W., Li, K., Zhou, S., Wang, Q. and Yu, H. (2023) 'Enhancing blind-dumb assistance through a self-powered tactile sensor-based braille typing system', *Nano Energy*, Vol. 116, p. 108795. https://dx.doi.org/10.1016/j.nanoen.2023.108795

Manga, T. and Masuku, K.P. (2020) 'Challenges of teaching the deaf-blind learner in an education setting in Johannesburg: Experiences of educators and assistant educators', *South African Journal of Communication Disorders*, Vol. 67, No. 1. https://dx.doi.org/10.4102/sajcd.v67i1.649

Mazouch, P., Ptáčková, V., Fischer, J. and Hulík, V. (2018) 'Students who have unsuccessfully studied in the past – analysis of causes', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 11, No. 3, pp. 65–71. https://dx.doi:10.7160/eriesj.2018.110303

McFee, B., McVicar, M., Raffel, C., Liang, D., Nieto, O., Moore, J., Ellis, D., Repetto, D., Viktorin, P., Santos, J.F. and Holovaty, A. (2015) 'Librosa: v0.4.0.', Zenodo. https://dx.doi.org/10.5281/zenodo.18369

McKinney, E. (2022) 'Meeting the Needs of a Student with Deaf-blindness in the Academic Classroom: A Model That Works', *Visual Impairment and Deaf-blind Education Quarterly*, Vol. 67, No. 4, pp. 17–43.

Montgomery, D.C. (1991) *Introduction to Statistical Quality Control*, 2nd Edition, New York: John Wiley and Sons.

Moore, C. and Murray, I. (2001) 'An electronic design of a low-cost braille typewriter', *Proceedings of The Seventh Australian and New Zealand Intelligent Information Systems Conference (ANZIIS, 2001),* Adelaide, pp. 153–157.

Morgan, D.S., Carhart, D.A., Dmitriyev, S., Gunderson, B.J., Patadia, S., Pulik, L. and Yanz, L.A. (2011) *Braille Writer: US Patent 20090274505A1*, [Online], Available: https://patents.google.com/patent/US20090274505A1/en [16 Sep 2023].

NCDB (2013) *State Deaf-Blind projects, State Deaf-Blind Projects | National Center on Deaf-Blindness,* [Online], Available: https://www.nationaldb.org/state-deaf-blind-projects/ [11 Jul 2023].

NFB (2021) *Deaf-blind Resources, Deaf-blind Resources | National Federation of the Blind*, [Online], Available: https://nfb.org/resources/deaf-blind-resources [11 Jul 2023].

Olayi, J.E., Igba, I.U. and Ewa, J.A. (2023) 'Communication strategies for teaching deaf-blind learners in an inclusive classroom', *Special needs education from the lens of interdisciplinary dialogue: A Festschrift in honour of Prof. Emeka D. Ozoji,* Vol. 1, No. 1, pp. 269–278.

Riccobono, D. and Morrow, S.M. (2022) 'State Deaf-blind Projects: National Survey Results', *Visual Impairment and Deaf-blind Education Quarterly,* Vol. 67, No. 4, pp. 109–125.

Rodríguez-Jiménez, M., Pérez-Jorge, D., Puerta-Araña, I. and Ariño-Mateo, E. (2022) 'Quality of life in Deaf-blind people and its effect on the processes of educational adaptation and social inclusion in Canary Islands, Spain', *Education Sciences*, Vol. 12, No. 7, p. 490. https://dx.doi.org/10.3390/educsci12070490

Rusiecki, A. (2019), 'Trimmed categorical cross-entropy for deep learning with label noise', *Electronic Letters*, Vol. 55, No. 6, pp. 319–320

Schulz, S., Harzheim, L., Hübner, C., Lorke, M., Jünger, S. and Woopen, C. (2023) 'Patient-centered empirical research on ethically relevant psychosocial and cultural aspects of cochlear, glaucoma and cardiovascular implants – a scoping review', *BMC Medical Ethics*, Vol. 24, No. 1. https://dx.doi.org/10.1186/s12910-023-00945-6

Smolen, E.R. and Paul, P.V. (2023) 'Perspectives on identity and d/deaf and hard-of-hearing students', *Education Sciences*, Vol. 13, No. 8, p. 782. https://dx.doi.org/10.3390/educsci13080782

Traer, J. and McDermott, J.H. (2016) Statistics of natural reverberation enable perceptual separation of sound and space | *IR Survey (Traer and McDermott)*, [Online], Available: https://mcdermottlab.mit.edu/Reverb/IR_Survey.html [17 Sep 2023].

Turunen-Taheri, S., Hagerman, A., Hellström, S., Skjönsberg, Å. and Backenroth, G. (2023) 'Combined severe-to-profound hearing and vision impairment—experiences of daily life and need of support, an interview study', *PLOS ONE*, Vol. 18, No. 6, p. e0280709. https://dx.doi.org/10.1371/journal.pone.0280709

Vincent, C., Wittich, W., Bergeron, F., Hotton, M. and Achou, B. (2021) 'Shopping when you are Deaf-blind: A pre-technology test of new methods for face-to-face communication—deaf-blindness and face-to-face communication', *Societies*, Vol. 11, No. 4, p. 131. https://dx.doi.org/10.3390/soc11040131

WFDB (2024) *World Federation of the Deaf-blind | Personas con sordoceguera y educación*, [Online], Available: https://wfdb.eu/es/wfdb-report-2018/deaf-blindness-and-education/ [19 February 2024].

WFDB (2023) *Second global report on the situation of persons with deaf-blindness | World Federation of the Deaf-blind,* [Online], Available: https://wfdb.eu/wp-content/uploads/2023/03/ENG_WFDB-2nd-Global-Report_-FINAL-V6.pdf [19 Feb 2024].

Zdravkova, K., Krasniqi, V., Dalipi, F. and Ferati, M. (2022) 'Cutting-edge communication and learning assistive technologies for disabled children: An artificial intelligence perspective', *Frontiers in Artificial Intelligence*, Vol. 5, p. 970430. https://dx.doi.org/10.3389/frai.2022.970430