Technology for Special Needs Education

Team: Speak Up!

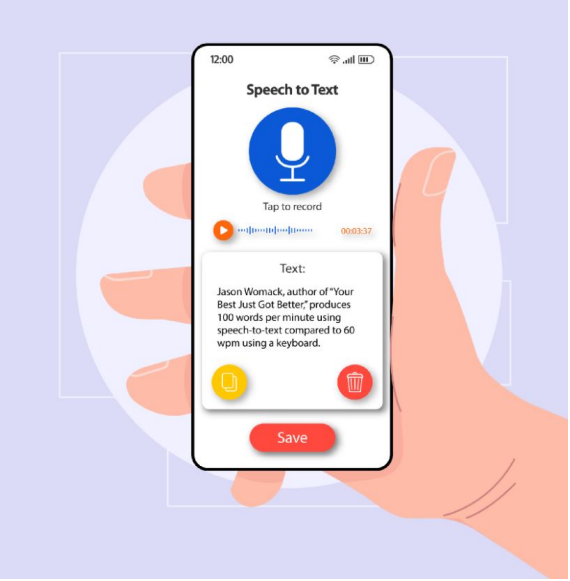
A step towards making communication easier…

Team Members

Chirag Ghosh (20CS10020)
Sarita Singh (20CS10053)
Jatin Gupta (20CS10087)
Manami Mondal (20CS10033)
Nikhil Saraswat (20CS10039)
Amit Kumar (20CS30003)

Voice to Text Converter

Speech-to-text tools save time, which is helpful for students who might forget their ideas once they try to write or students who struggle with getting any words on the page at all, feeling unable to transfer their thoughts. For some, this was because of the intimidation of writing academically, with spelling and grammar anxieties prohibiting them from starting. For others, the time taken to write out initial thoughts caused them to forget later conclusions and analyses, given the lack of immediacy in writing.

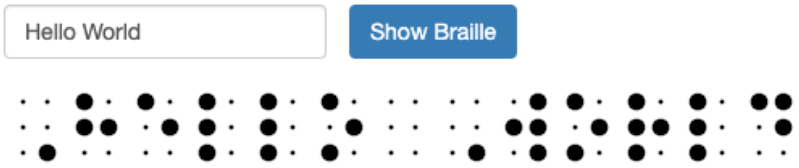


This assistive device aids an individual in communicating with someone who does not know how to use sign language. This application converts texts to an audio file and it allows individuals to type out the message that he/she wants to convey and
have the device read it aloud.

Computers and other electronic devices not only have the ability to convert spoken words into text but they can read text out loud. This technology can help people with vision impairment and learning disabilities in several ways. For example, if students have trouble comprehending words on a computer screen, they can use a software program or hardware device to convert text into audible words.
The voice to text is an integral part of our software used in implementing the Voice-to-Sign language feature.

Text to Braille Converter

Braille is a crucial and lively form of communication, education, and community building for persons who are blind or visually impaired.The direct attention kids require to achieve correctness and fluency in Braille is frequently difficult to deliver with a limited number of skilled teachers. There is a scarcity of skilled teachers who can teach Braille. There aren't enough tools to make learning Braille as easy as learning print.
Also, How do Deaf-Blind People Communicate? We need a text to Braille or Voice to Braille converter. This can be a tiny, portable gadget that allows them to converse with individuals who can see or speak. The gadget will feature a braille display. By putting his or her fingertips on the braille display, the deaf-blind person can read the written text.
So, we are developing a Web application which can convert Text to Braille (We are not developing the hardware model for it). Our project covers the software part of such a hardware prototype.



General features which we are going to implement :

1. Text to Braille - by general known algorithm to convert text to grade 2 braille (given below)
2. Braille to text by general known algorithm to convert grade 2 braille to text

Map which we have used :
Letters:
a ⠁ b ⠃ c ⠉ d ⠙ e ⠑ f ⠋ g ⠛ h ⠓

i ⠊ j ⠚ k ⠅ l ⠇ m ⠍ n ⠝ o ⠕ p ⠏
q ⠟ r ⠗ s ⠎ t ⠞ u ⠥ v ⠧ w ⠺ x ⠭
y ⠽ z ⠵

Contraction:
⠯ and ⠵ as ⠃ but ⠉ can ⠙ do
⠑ every ⠿ for ⠋ from ⠛ go ⠓ have
⠦ his ⠔ in ⠭ it ⠚ just ⠅ knowledge
⠇ like ⠍ more ⠝ not ⠷ of ⠏ people
⠟ quite ⠗ rather ⠎ so ⠞ that ⠮ the
⠖ to ⠥ us ⠧ very ⠴ was ⠺ will
⠾ with ⠽ you

Punctuation:
! ⠖ # ⠼ ' ⠄ (⠶) ⠶
, ⠂ - ⠤ . ⠲ / ⠌ : ⠒
; ⠆ ? ⠦ ⠤ – ⠤ – ⠤
— ⠤ – ⠤ — ⠤ — ⠤ ’ ⠄
“ ⠦ ” ⠴ … ⠲⠲⠲

Numbers (must proceed after the ⠼ escape code):
0 ⠚ 1 ⠁ 2 ⠃ 3 ⠉ 4 ⠙
5 ⠑ 6 ⠋ 7 ⠛ 8 ⠓ 9 ⠊

The Algorithm for Translating Alphabet Based Text to Grade 2 Braille:

1. Split up the text into words by dividing them based on whitespace characters.
    i. Whitespace includes spaces (' ') and new lines ('\n')
2. For each word, handle the numbers first.
    i. Numbers in braille use the same symbols as the first 10 letters of the alphabet.

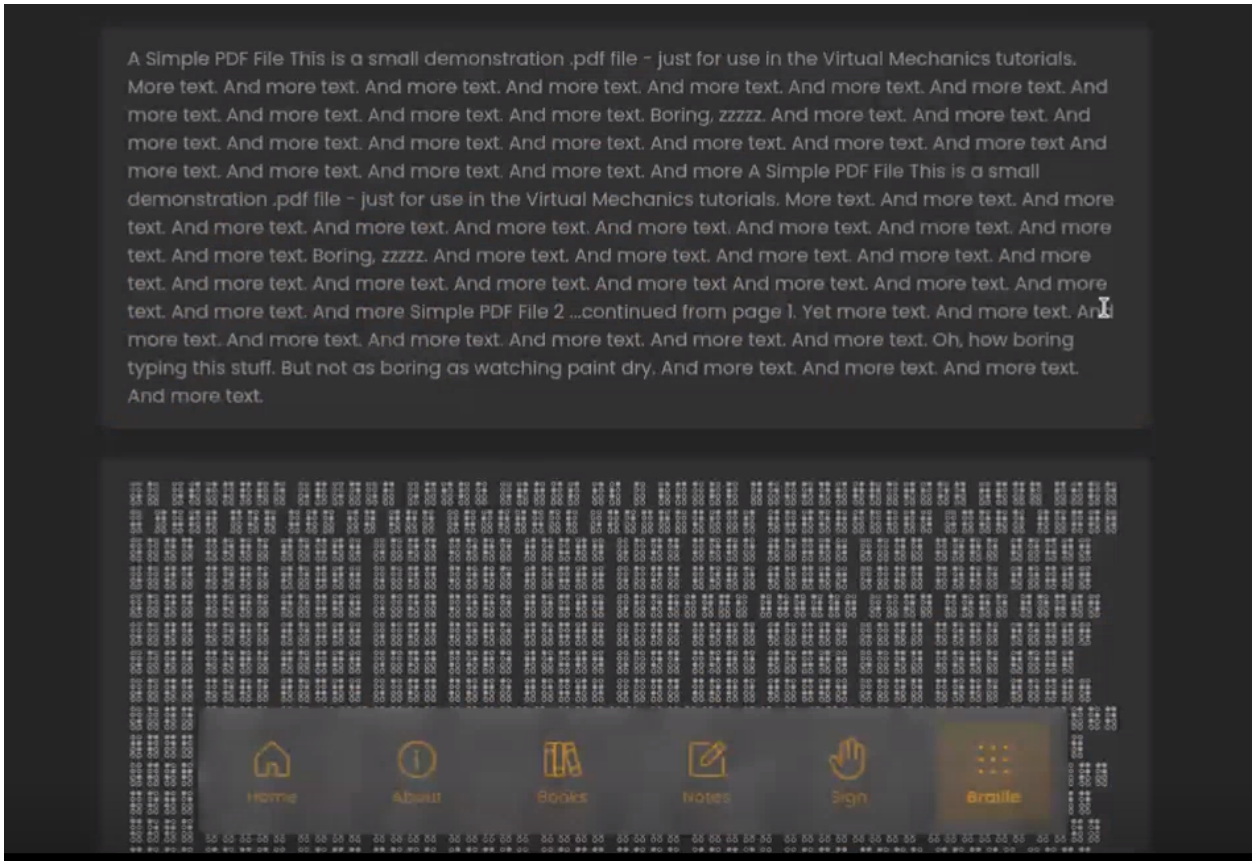
        - The number '7' and the letter 'g' are both represented by '⠛'.
        - To differentiate between numbers and letters, an escape code (⠼) is placed before groups of numbers.
        - Therefore '7' is actually '⠼⠛' whereas 'g' is only '⠛'.
    ii. In this step, only the numbers are dealt with, so there will be a mix of both braille and Alphabet symbols.
        - Example: "123-456-JUNK" becomes "⠼⠁⠃⠉-⠼⠙⠑⠋-JUNK"
3. Handle the capitals.
    i. Similarly to numbers in braille, capital letters need an escape code (⠠).
    ii. The escape code (⠠) is added to the beginning of each capital letter and the letter is changed to lowercase.
        - Example 1: "⠼⠁⠃⠉-⠼⠙⠑⠋-JUNK" becomes "⠼⠁⠃⠉-⠼⠙⠑⠋- ⠠j ⠠u ⠠n ⠠k". The dashes still remain.
        - Example 2: "Sweet" becomes " ⠠sweet". The non-capital letters remain untouched.
4. Trim the word.
    i. Sometimes the words extracted contain punctuation attached to them such as commas or brackets.
    ii. Words need to be trimmed so that they can be converted to contractions.
        - Example: The word "the" is represented by a single braille symbol (⠮).
        - If the word "the" has punctuation around it ("the!") then it will not be interpreted correctly.
        - This is also why capitals are converted to lowercase in step 3 because "The" would not work either.
    iii. The characters that are trimmed off are called "shavings".
        - Example: In the word "!where?", the shavings are "!?" and the trimmed word is "where".
5. Build the translation.
    i. Check to see if the trimmed word can be contracted.

- This includes common words like "the", "in", "you" etc...

ii. Translate the remaining characters that are still alphabetic.

iii. Translate the shavings (this will mostly just be punctuation).
   - Exceptions to be mindful of:
     - There is no braille symbol for a generic quote (")
     - There is only open quotation (“) and closed quotation (”).
     - Therefore we must keep track of what the last quotation was to translate it correctly.

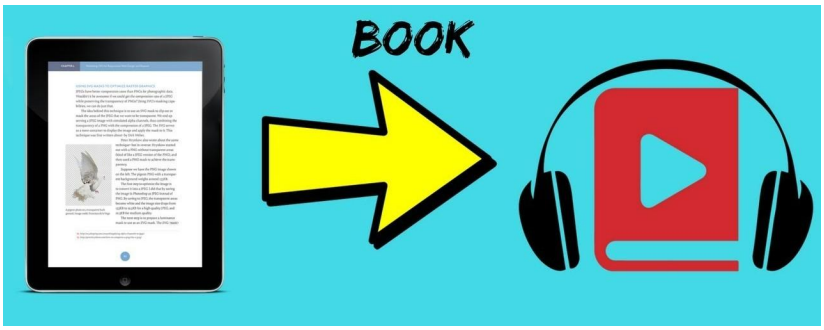
Example→



(Snapshot of text-to-braille feature of our software)

Book Narrator

Since as we all know books play a crucial role in the development of a human being. Books play an essential role in every student's life by introducing them to a world of imagination, providing knowledge of the outside world, improving their reading, writing and speaking skills as well as boosting memory and intelligence.The importance of books in our life cannot be undermined for they not only help in broadening our horizons but also act as doorways to connecting us with the world around us.



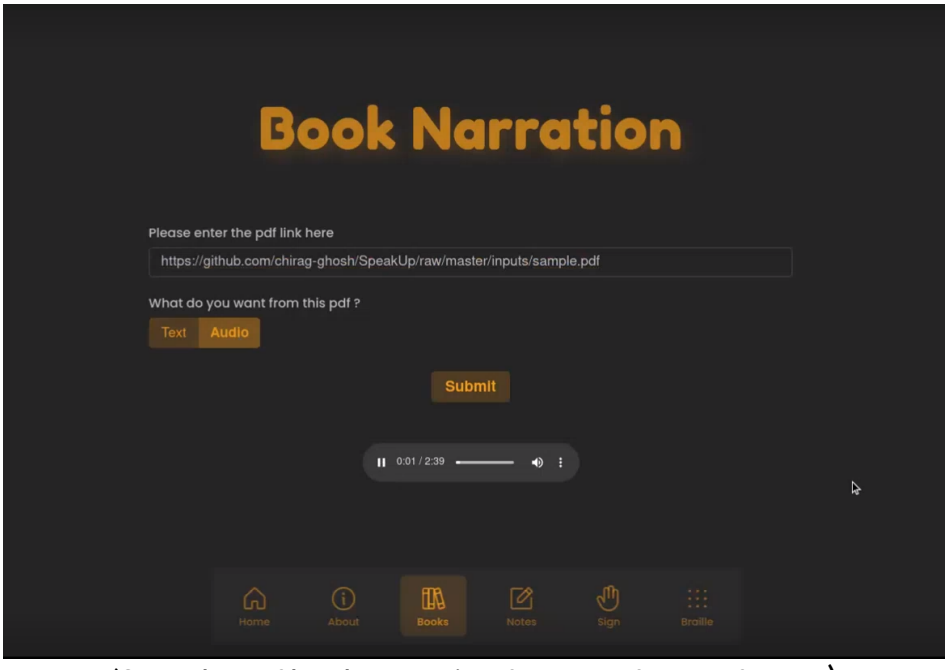
But due to visual impairment many people who are book-lovers face difficulties reading books unlike a normal person. Audio-books might help such people a lot who have visual impairments. By using audio-books, reading books is more like story-telling, which makes it easy for a person to understand the book content. Audio-books are not only helpful for a person with visual-impairment but this is also useful for a normal person. Audio books can help a child not only gain knowledge but can also boost their self-confidence and independence.

Audio books for those who are visually impaired help people with vision loss or low vision by enriching their knowledge and providing information on different subjects through the means of a book that is in the audible format. This tool is helpful, however, in reading longer passages.
Pairing books (in print or braille) with Audio Supported Learning is a great strategy to build listening skills.

For developing a book narrator, we can develop a feature which takes the book (pdf) as input, parses it and then converts it to audio and narrates it. For this feature, we need to obtain text from a pdf, which can be done using a pdf parsing tool. Sometimes in a book there are also some images (screenshots of some text) which also need to be converted to text, so using Optical Character Recognition (OCR) (Optical Character Recognition (OCR) is the process that converts an image of text into a machine-readable text format.).

We also need a tool to stitch the audio files of each page to make a single audio file of the whole book, which can be done easily using audio merging tools. We can use the PyPDF python module to parse pdfs of books. We can use OCR and audio APIs of Azure Services for OCR and audio requirements.

In this way, integrating all these modules in a python file, we can develop a program which takes a pdf file to convert it to its audio version. And this feature can be integrated in any application (in our project we have integrated it in a web-based application which provides a graphical interface for users to interact with it, in this way, it smoothens the book narration).

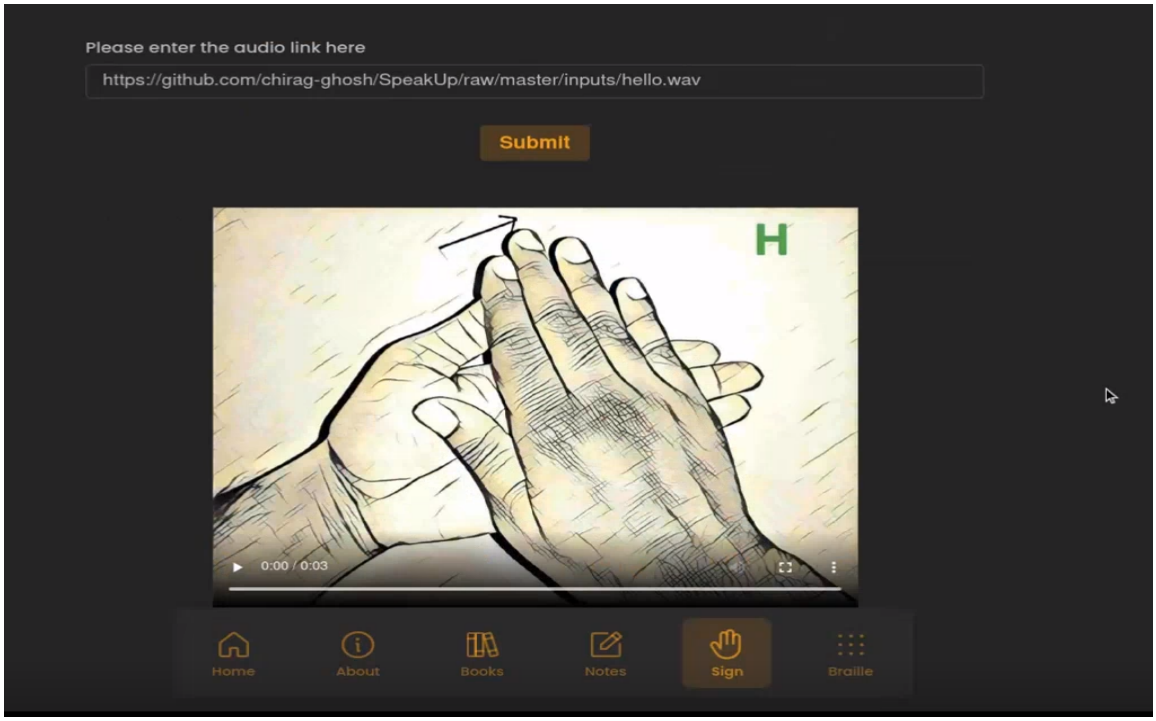


(Snapshot of book narration feature of our software)

Audio to Sign Language:

Sign language is one of the oldest and most natural forms of language for communication. Using the method described before, we can convert any audio to text. Using this generated text, we can replace the letters with the corresponding sign language and then join these images to make a video depicting the information through sign language. This will be useful to provide information access and services to deaf people in Indian sign language.

This is based on converting the audio signals received to text using speech to text api (python modules). After that we make use of the semantics of Natural Language Processing in order to break down the text into smaller understandable pieces with the help of machine learning. We have used the data sets of predefined sign language as our input so that the software can use artificial Intelligence to display the converted audio into the sign language.



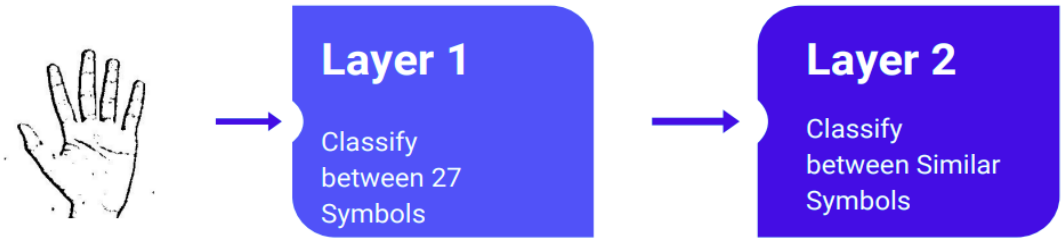
(Snapshot of audio-to-sign language feature of our software)

Sign Language to Text:

Communication is the process of exchange of thoughts and messages through speech, signaling, behavior, and visual cues.The only way for Deaf and Dumb people to communicate is through sign language. These people make use of their hands to express different gestures to express their ideas with other people. This nonverbal communication is called sign language. But the majority of people do not know sign language and interpreters are very hard to find.

Hence we have planned to develop a real-time method using neural networks for Sign language. Hence our sign language recognition system has been implemented to recognize 26 gestures from the Indian Sign Language by hand gesture recognition system in order to generate text. The signs are captured by using a webcam. This will provide an opportunity for people with hearing and speaking disabilities to communicate with non-signing people without the need of an interpreter.

We used a Vision based approach. In our approach, a computer camera will be the input device for observing the information of hands or fingers. The approach will require only a camera, thus realizing a natural interaction between humans and computers without the use of any extra devices. These systems tend to complement biological vision by describing 7 artificial vision systems that are implemented in software and/or hardware.



We implemented a total of 27 symbols which includes 26 symbols(A-Z) and blank Indian Sign Language (ISL). Now in order to predict the final symbol of the user we will use two layers of algorithm.

In the first layer, Algorithm Layer 1 we will apply a gaussian blur filter and threshold to the frame taken with opencv to get the processed image after feature extraction. This processed image will be passed to the CNN (Convolution Neural Network) model for prediction and if a letter is detected for more than 50 frames then the letter is printed and taken into consideration for forming the word. We will consider the space between words using the blank symbol.

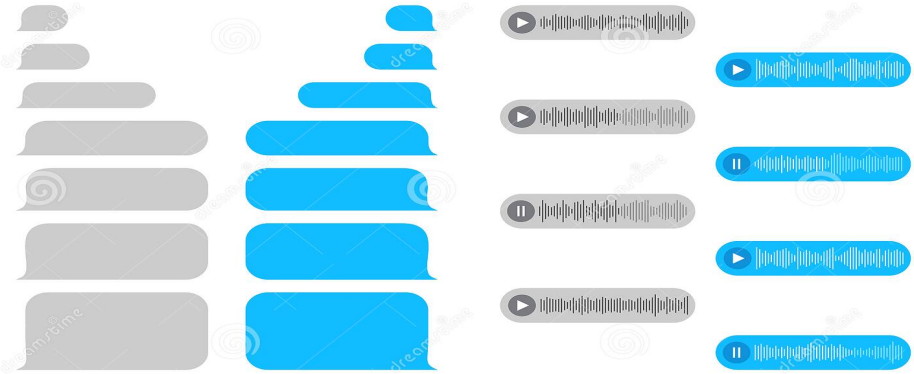
In the second layer, Algorithm Layer 2 we will detect various sets of symbols which show similar results on getting detected. Using the data we will then classify between those sets using classifiers made for those sets. We convert sign language to text, aiding those unable to comprehend sign language.

Instant Message 4 all

Chatting and messaging is one of the most popular methods of communication nowadays. They can be much more than ways to communicate. Apart from communicating, one can learn new perspectives and skills via messaging and chatting with people. Communication is a great way to socially interact with one another and improve the overall personality.

People with certain disabilities find it difficult to message and chat with other people and because of this, they are not able to use this amazing facility to communicate. They might feel lonely at some time and may want to chat with someone, maybe a stranger. People with visual impairment cannot read and write text. For them, chatting can be done via audio files. Audio files can also be used by people who find it difficult to read and write. People with hearing impairment cannot listen to these audio files. So they can stick to the traditional text communication for chatting.

We thought of solving this problem and creating a feature in our software that will act as messaging software for everyone. In this software, specially-abled ones can also have a real-time lag-free chat with the other person. The software will have features that will help people with certain disabilities to chat. This feature will be developed using a socket.io server. Along with this, we'll be using Google cloud speech API for incorporating text-to-speech and speech-to-text in our software.



For the implementation, we established a socket.io server for client-server communication. This server provides real-time communication. After getting this server ready, the users will join this server for a chat.
The user has the option to write the message or record the message.
If the input message is in the form of audio, then we use the google cloud speech API to convert the speech to text. If we have the audio message in text, then we'll be using the text-to-speech feature to convert the text message to audio. So, now we will have the input message in both audio formats as well as text formats.
The socket.io server will now publish this input message to another user in both the audio format as well as the text format. The other user can then read or listen to this message and respond accordingly. This acts as an all-in-one real-time lag free chatting feature.

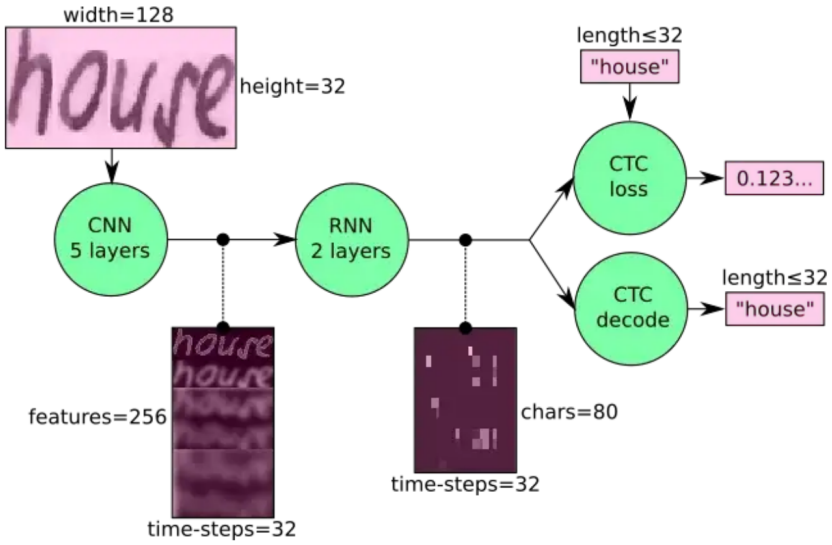
Handwriting Reader

This feature lets you scan handwritten documents or short notes that can then be fed to our book reading feature to be read aloud, i.e. it takes in handwritten notes and processes a text-based document that can be read and further processed as required.

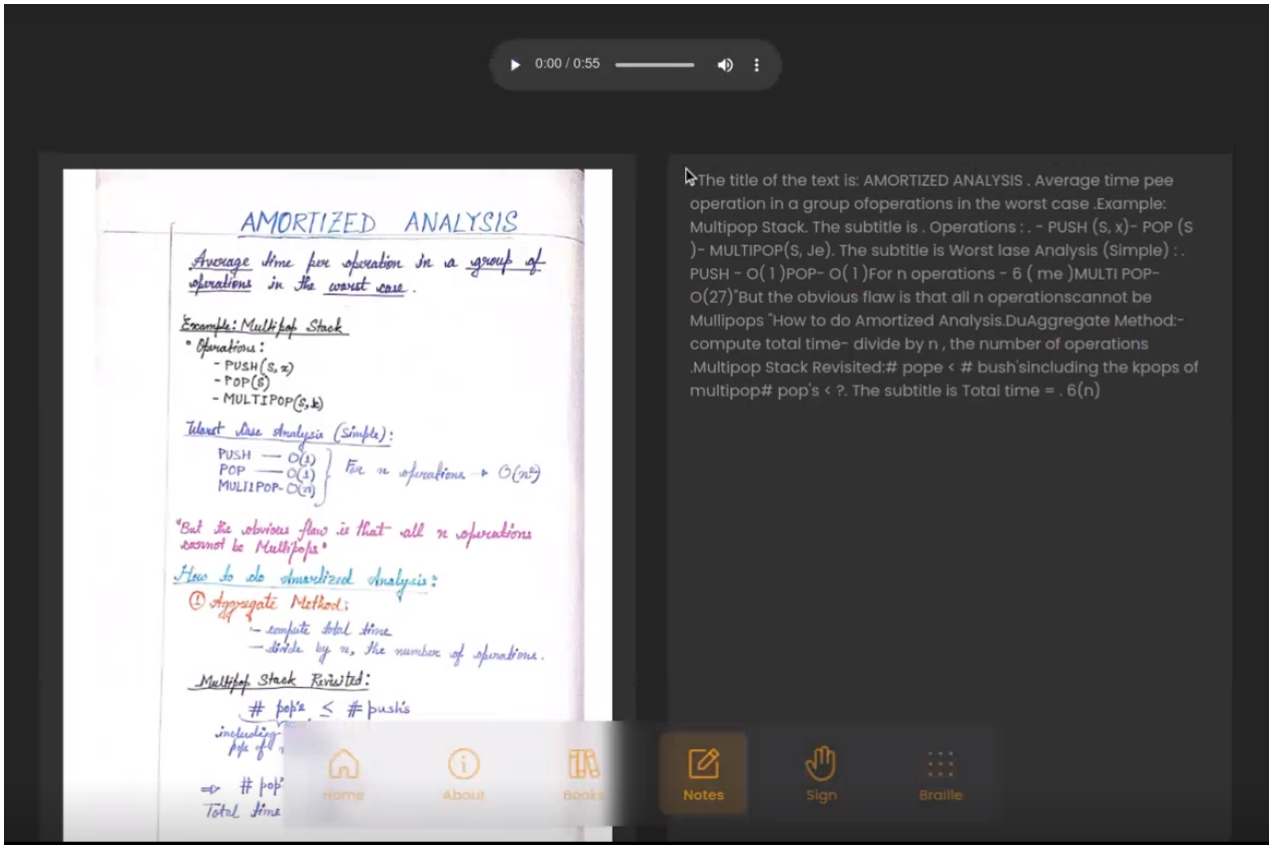
The model takes images of single words or text lines (multiple words) as input and outputs the recognized text. Hence, this helps the people who are visually impaired or having dyslexia to understand the content of handwritten documents.

The text is recognized on character-level, therefore words or texts not contained in the training data can be recognized too (as long as the individual characters get correctly classified). The popular Long Short-Term Memory (LSTM) implementation of RNNs is used, as it is able to propagate information through longer distances and provides more robust training-characteristics than vanilla RNN. The word beam search decoder is used instead of the two decoders shipped with TF. Words are constrained to those contained in a dictionary, but arbitrary non-word character strings (numbers, punctuation marks) can still be recognized.

A Neural Network (NN) which is trained on word-images from the IAM dataset will be built. As the input layer (and therefore also all the other layers) can be kept small for word-images, NN-training is feasible on the CPU (of course, a GPU would be better). It consists of convolutional NN (CNN) layers, recurrent NN (RNN) layers and a final Connectionist Temporal Classification (CTC) layer. The NN consists of 5 CNN and 2 RNN layers.



Overview of the NN operations (green) and the data flow through the NN (pink)



(Snapshot of handwriting feature of our software)