Smart Image and Voice to Braille Captioning and Input Assistant for the Visually Impaired

AMRITA
VISHWA VIDYAPEETHAM
DEEMED TO BE UNIVERSITY

Jaswanth Kumar N (BL.EN.U4EAC22023)
K Supriya (BL.EN.U4EAC22024)
R Kavin Kumar (BL.EN.U4EAC22043)
Dept. Electronics and Computer Engineering
Amrita School of Engineering Bangalore

Content

- ❑ Problem Statement
- ❑ Objective
- ❑ Literature Survey
- ❑ Design
- ❑ Simulation Results
- ❑ Software Implementation
- ❑ Hardware Implementation
- ❑ Analysis / Interpretation of Results
- ❑ Conclusion and Future Scope
- ❑ References

Problem Statement

- Braille is a crucial medium for visually impaired individuals to read and write, enabling access to education, communication, and independence.
- Traditional braille books are bulky, expensive, and require significant effort to produce, limiting accessibility.
- **Social Impact:** Promotes inclusivity by enhancing access to education, communication, and job opportunities for visually impaired individuals.
- **Technological Impact:** Leverages advanced electromechanical principles and 3D printing to create cost-effective Braille cells.
- **Economic Impact:** Reduces the cost of Braille display devices, making them more accessible to a wider audience.
- It eliminate disparities in education, empower and promote inclusion and promote global technology cooperation
- The given solution aligns with **SDG 4.5, SDG 10.2, SDG 17.8.**

Objective

To develop a cost-effective, portable, and scalable refreshable Braille display, called the Braille Pad, which enhances digital accessibility for visually impaired individuals by enabling real-time Braille reading and seamless interaction with other digital devices and also include visual and auditory inputs.

Sub-Objectives:

- Simulation and designing of a simplified braille cell module.
- Development of the chainable brain cells.
- Implement and Demonstrate the usage of multiple braille cells.
- Image captioning machine learning model
- Audio translation model
- Design and development of 3D components
- Inter connection of keypad and braille cell

Literature Survey

Paper Title	Problems Addressed	Methodology	Outcomes	Research Gap
[1] Gupta et al. Design and Implementation of Arduino-based Refreshable Braille Display Controller (2016) (Journal)	Design and implementation of a refreshable Braille display controller using Arduino for the visually impaired.	Developed an Arduino-based embedded system to transliterate ASCII text into Braille using six-dot combinations. Features multi-lingual support (English & Devanagari) and a stand-alone battery-powered device.	Low-cost, portable, and efficient refreshable Braille display with transliteration capabilities. Improves accessibility for the visually impaired.	Lacks real-time text scanning and OCR integration. No support for dynamic interaction with printed text.
[2] Rahimi et al. Printed Texts Tracking and Following for a Finger-Wearable Electro-Braille System Through Opto-Electrotactile Feedback (2021) (Journal)	Real-time printed text recognition and tracking for a finger-wearable Electro-Braille system.	Developed an opto-electrotactile feedback mechanism combined with Rapid Optical Character Recognition (R-OCR) to allow users to follow printed text by tracking their finger movement. Utilizes electro-tactile stimulation instead of traditional mechanical Braille displays.	Demonstrated improved accuracy in text tracking, allowing visually impaired users to maintain a 2mm proximity to text while scanning. Provides real-time feedback for better interaction.	Lacks extensive testing with real visually impaired users. Limited hardware scalability and no multi-lingual text recognition.
[3] Gozon et al. A Digitally Automated Text to Braille Device for the Visually-Impaired (2019) (Conference)	Development of a digitally automated text-to-Braille device to address the decline in Braille literacy due to bulky printed materials.	Uses an Arduino-based microcontroller to convert digital text into Braille via solenoid actuators. Implements Unified English Braille (UEB) coding and 3D-printed components for cost efficiency.	Successfully translates digital text into Braille with a reading speed of 180 words per minute. More portable and accessible than traditional refreshable Braille displays.	Only supports Unified English Braille, lacks adaptability for other Braille systems (e.g., Bharati Braille). No integration with Optical Character Recognition (OCR) for printed text conversion.

Literature Survey

Paper Title	Problems Addressed	Methodology	Outcomes	Research Gap
[4] Braille to Text and Speech for Cecity Persons. (Journal) (2015)	Lack of access to digital tools for visually impaired; communication and learning difficulties.	Uses Braille keypad as input to FPGA (Spartan 3); converts Braille to English text and audio using decoding logic; output on LCD and via voice chip.	Successfully converts Braille to both text and speech; uses minimal hardware; provides real-time feedback.	Limited to English language; lacks volume/speed control in audio; no support for pause/resume functionality.
[5] Analysis and Evaluation of Braille to Text Conversion Methods. (Journal) (2020)	Lack of standard, user-friendly Braille input systems on mobile/touch devices; usability issues with existing apps.	Comparative analysis of 38+ systems using scanned input and touch-screen input; deep learning (SDAE), OCR, gesture-based UIs, and haptic/audio feedback.	Provides usability metrics and performance comparison; defines new evaluation framework; suggests interaction models for future systems.	High learning curves, screen-dependency, lack of unified standard; most systems burden user instead of simplifying interaction.
[6] Electrostatic Smart Textiles for Braille-To-Speech Translation. (Journal) (2024)	Need for integrated, comfortable Braille-to-speech translation systems.	Developed fabric-based electrostatic transducers that function as both tactile sensors and loudspeakers; integrated with machine learning algorithms.	Achieved 99.09% accuracy in translating Braille alphabet and 97.08% for 40 commonly used words; offers a wearable solution.	Limited vocabulary size; scalability to full language translation not addressed.

Design - System Architecture

Inputs

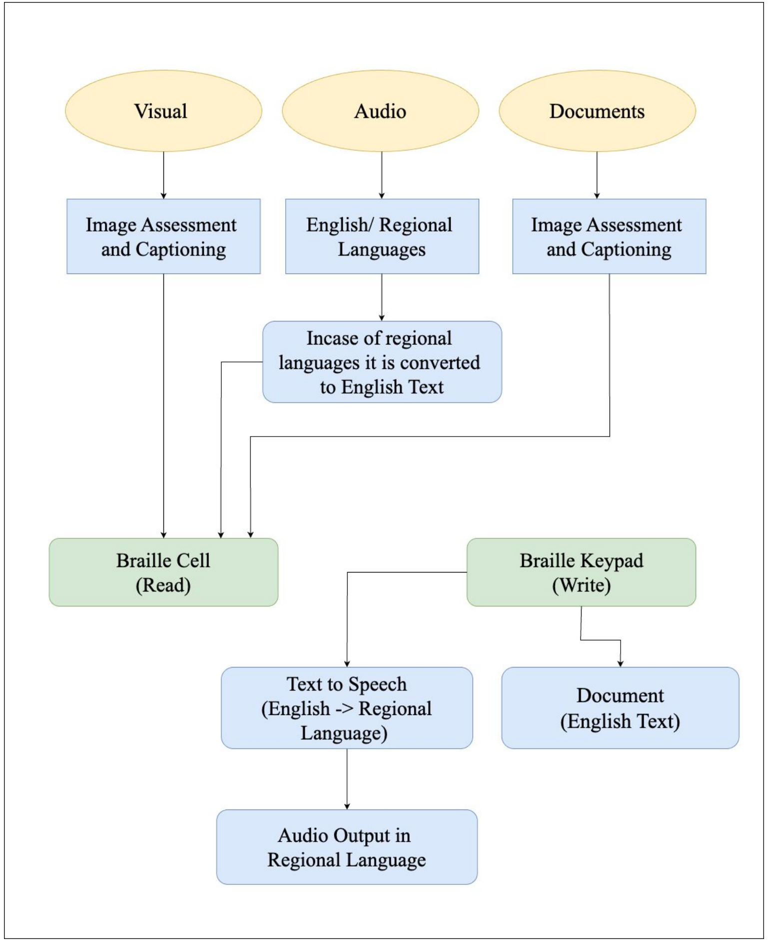
- **Visual**: Image assessment & captioning
- **Audio**: English/Regional language speech
- **Documents**: Text extraction via image processing

Processing

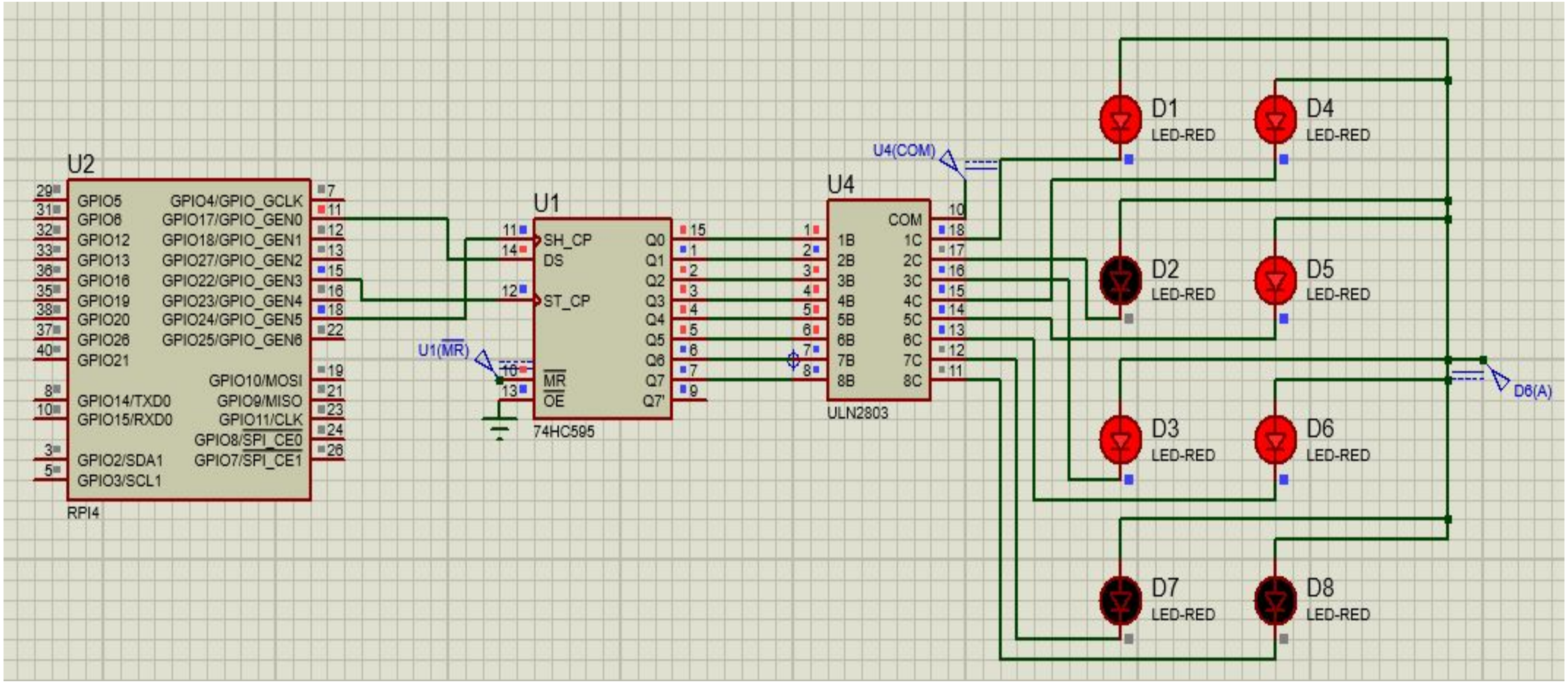
- Regional audio → English text (via ASR + Translation)
- Image/Document → Captions
- Text → Braille (Read/Write)
- Text → Speech (English to Regional)

Outputs

- **Braille Cell**: Tactile reading
- **Braille Keypad**: Document writing
- **Audio**: Regional language speech
- **Document**: English text output

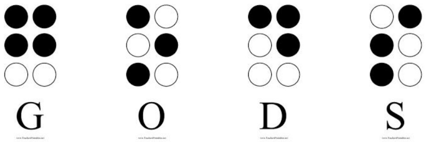


Simulation

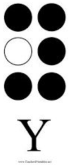


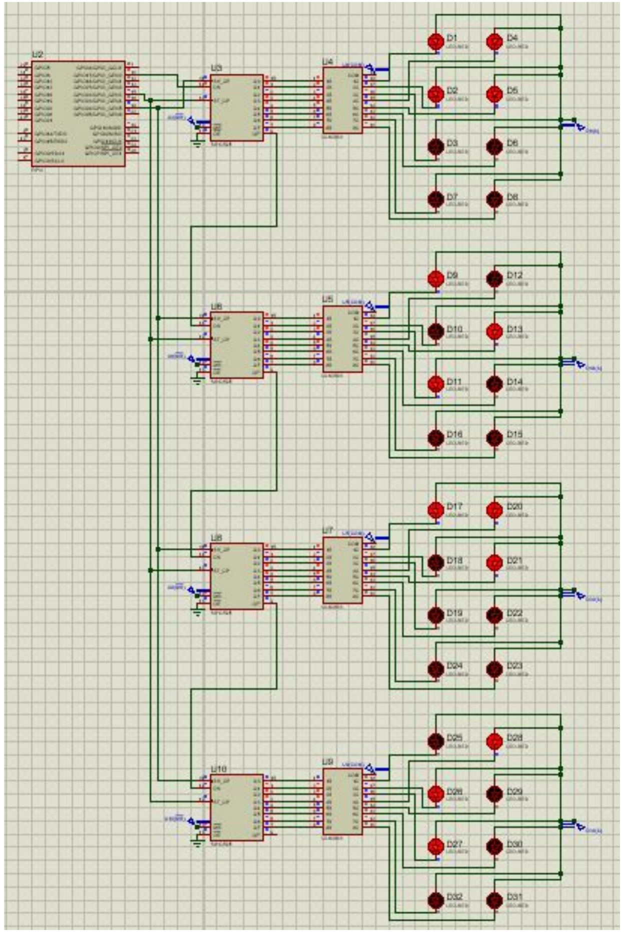
U2
RPI4
GPIO5
GPIO6
GPIO12
GPIO13
GPIO16
GPIO19
GPIO20
GPIO26
GPIO21
GPIO14/TXD0
GPIO15/RXD0
GPIO2/SDA1
GPIO3/SCL1
GPIO4/GPIO_GCLK
GPIO17/GPIO_GEN0
GPIO18/GPIO_GEN1
GPIO27/GPIO_GEN2
GPIO22/GPIO_GEN3
GPIO23/GPIO_GEN4
GPIO24/GPIO_GEN5
GPIO25/GPIO_GEN6
GPIO10/MOSI
GPIO9/MISO
GPIO11/CLK
GPIO8/SPI_CE0
GPIO7/SPI_CE1
U1
74HC595
SH_CP
DS
ST_CP
MR
OE
Q0
Q1
Q2
Q3
Q4
Q5
Q6
Q7
Q7'
U1(MR)
U4
ULN2803
U4(COM)
COM
1B
2B
3B
4B
5B
6B
7B
8B
1C
2C
3C
4C
5C
6C
7C
8C
D1
D2
D3
D4
D5
D6
D7
D8
LED-RED
D6(A)

Simulation

- Led Glow  the word - GODS (4 cells)

- LED Glowing for the alphabet - Y (1 cell)



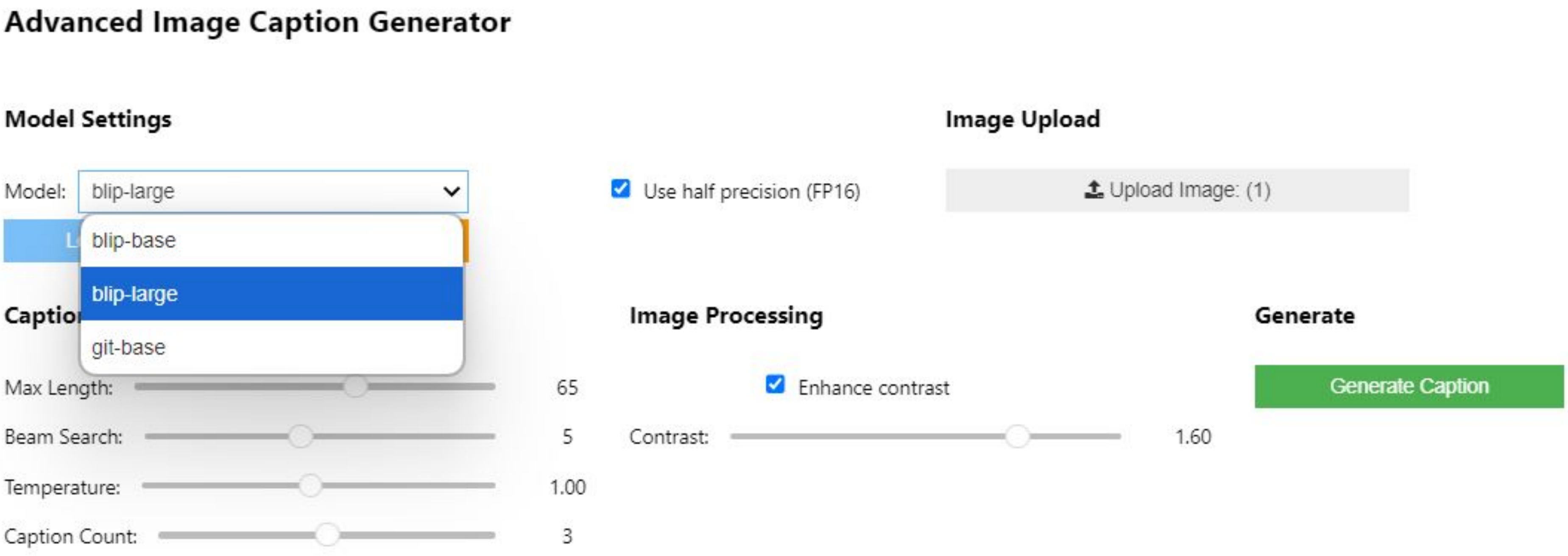


Software Implementation

Tools, Models and Techniques Used

- Python
- TensorFlow
- Keras
- Transformers (Hugging Face)
- Flask / Gradio
- Jupyter Notebook
- MarianMT (Helsinki-NLP)
- **Tokenizer** (associated)
- Seq2Seq Architecture

Software Implementation



Software Implementation

[INFO] 09:19:56 - image_captioner - Processing image...
[INFO] 09:19:56 - image_captioner - Processed image:

[INFO] 09:19:59 - image_captioner - Caption generation took 2.12 seconds

Generated Captions:

Caption 1: there is a blurry photo of a woman standing on a street

Caption 2: there is a blurry picture of a woman standing on a street

Caption 3: there is a blurry image of a woman standing on a street

[INFO] 09:19:59 - image_captioner - Caption saved to file

Software Implementation

Results

[INFO] 09:22:26 - image_captioner - Processing image...
[INFO] 09:22:26 - image_captioner - Processed image:

[INFO] 09:22:26 - image_captioner - Caption generation took 0.51 seconds

Generated Captions:

Caption 1: there is a white house sitting on a small island in the middle of the ocean

Caption 2: there is a white house on a small island in the middle of the ocean

Caption 3: there is a white house sitting on a small island in the middle of the water

[INFO] 09:22:26 - image_captioner - Caption saved to file

Software Implementation

🎙️ Whisper Translate - Bidirectional

1️⃣ Upload audio in any language to translate to English

```
FileUpload(value=(), accept='.wav,.mp3,.m4a,.flac', description='Upload Audio')
```

2️⃣ After detection, enter English text to translate back and hear it

Uploaded: Hindi_Voice.wav

0:00 / 0:07

🌐 **Translating to English...**

📝 **English Translation:**

The weather is very good today. It rained yesterday, but now it is hot. I hope it will rain at night.

🔍 **Detected Language:** hindi (hindi)

```
Textarea(value='', description='English:', layout=Layout(height='100px', width='100%'), placeholder='Enter Eng…

Button(button_style='primary', description='Translate & Speak', icon='microphone', style=ButtonStyle())
```

🌐 **Translating to hindi...**

📝 **hindi Translation:**

हां, आप सही कह रहे हैं, लेकिन अब फिर से बारिश होने लगी है, तो यह अच्छा है।

```
/tmp/ipykernel_3985/2869371396.py:136: DeprecationWarning: Due to a bug, this method doesn't actually stream the
  response.stream_to_file(temp_audio_file.name)
```

🔊 **Generated Audio:**

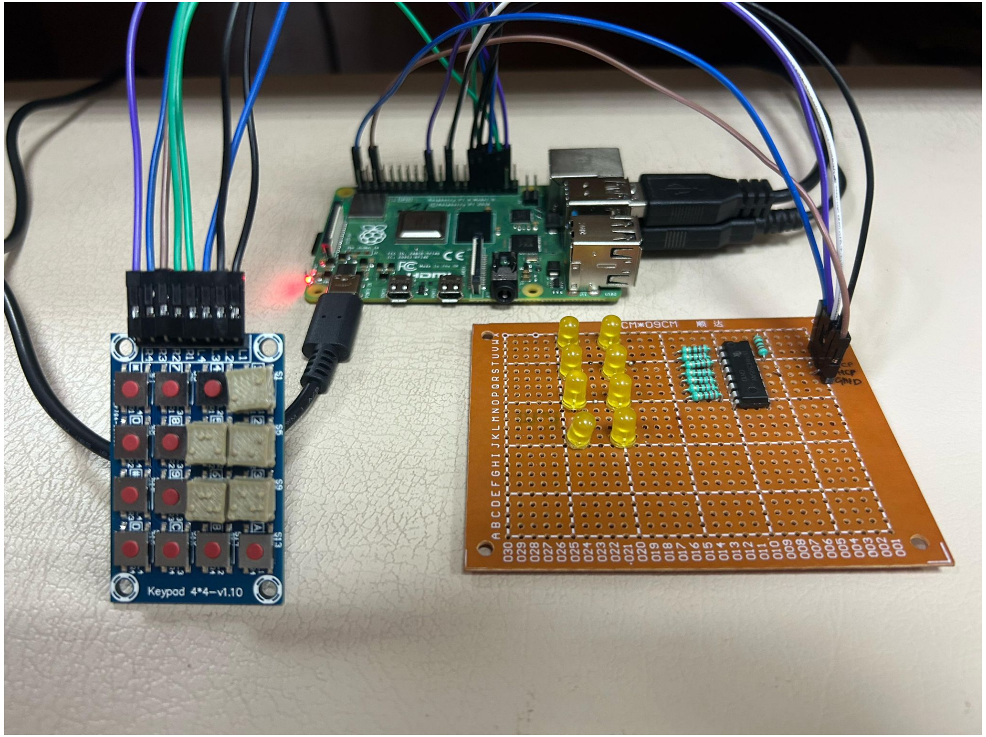
0:00 / 0:05

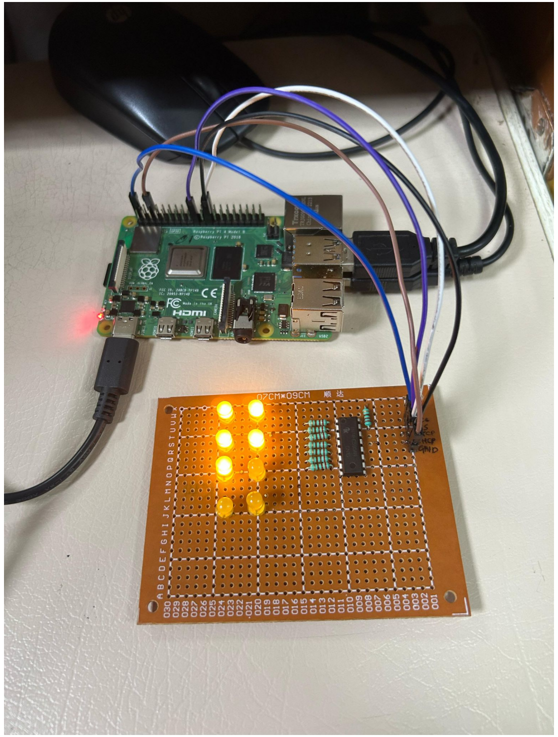
Hardware Implementation

Components Used

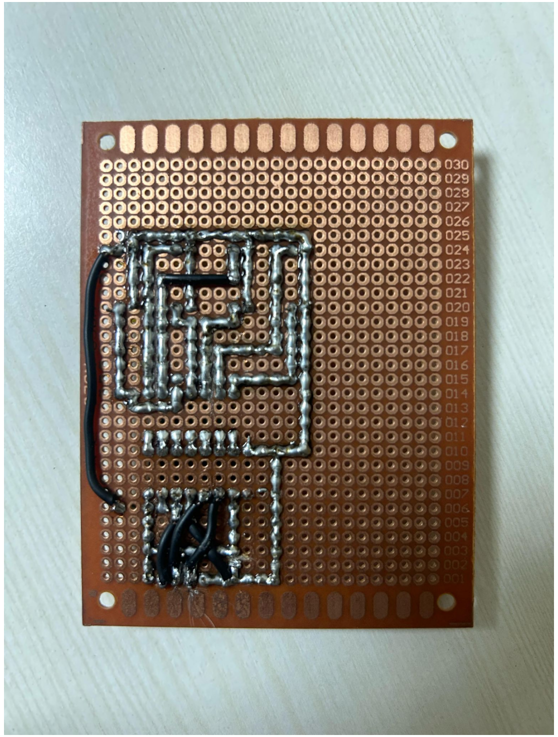
- Raspberry Pi 4
- 74HC595 8-Bit Shift Register
- ULN2803 Transistor Array
- Raspberry Pi Cam Module 3
- 220K-Ohm Resistors
- LED
- 4x4 Matrix 16-Button Keypad Module
- Jumper Wires
- PCB
- 3D Printed Keycaps
- 3D Printed Actuation Mechanism (Proof of concept)

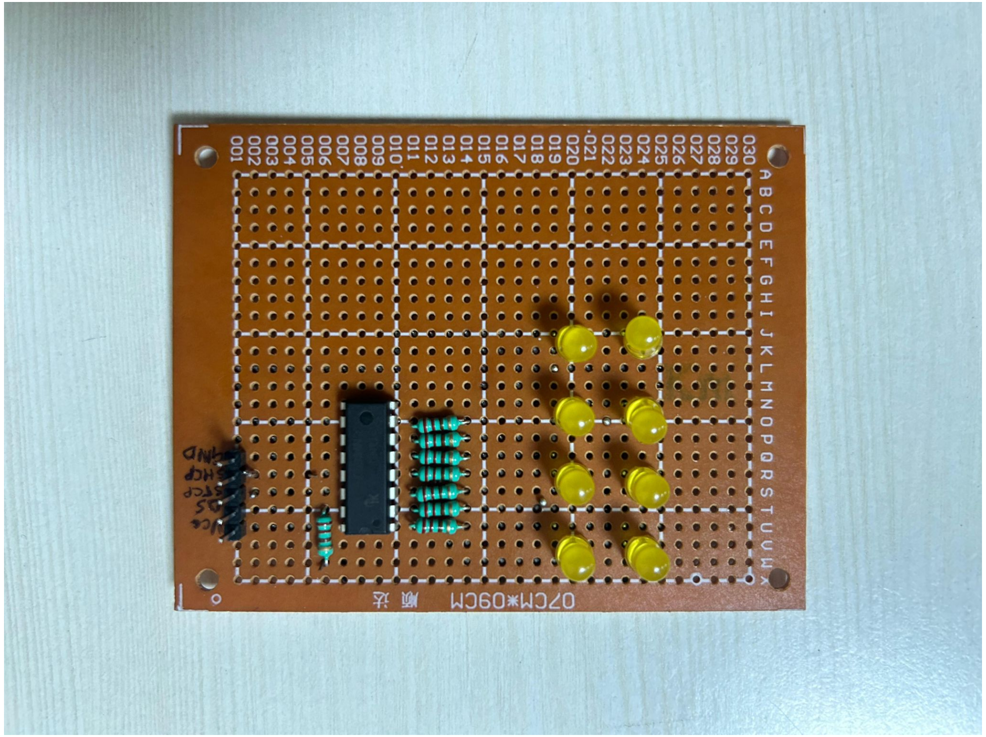
Hardware Implementation



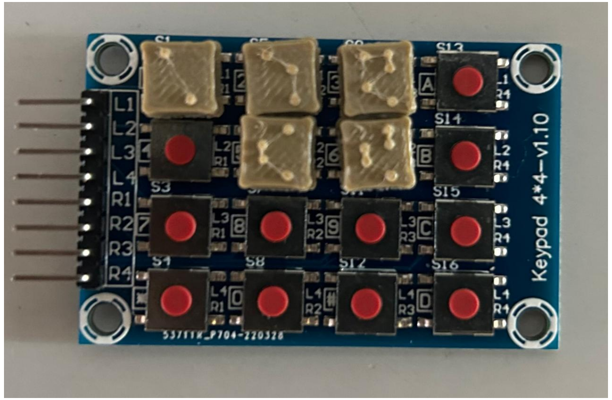


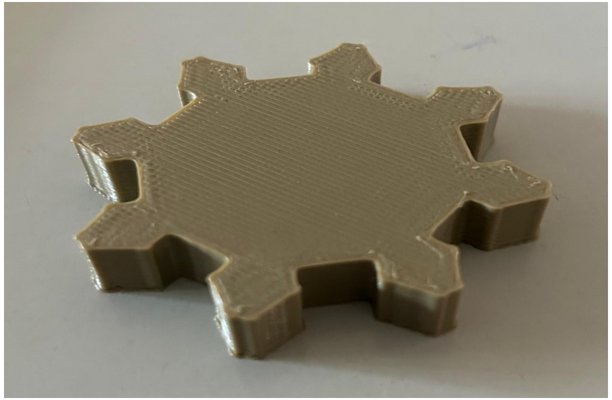
Hardware Implementation

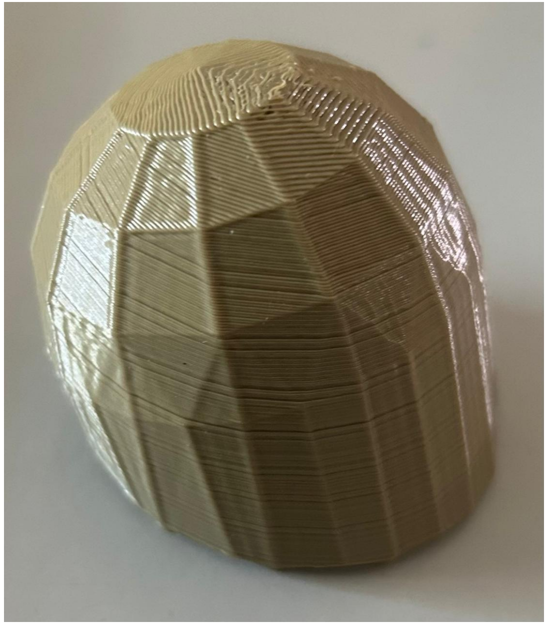


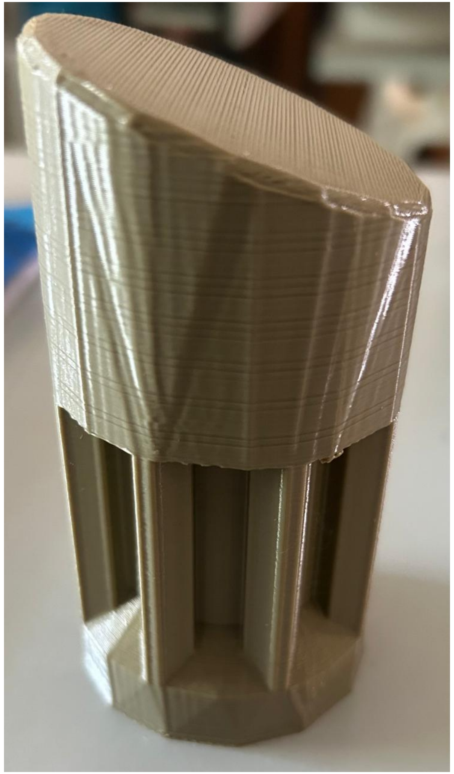


Hardware Implementation









Analysis / Interpretation of Results

Simulation Results:

1. **Successful Braille Cell Activation:**
   The simulation demonstrated accurate movement of the Braille pins to create the desired Braille patterns.
2. **Response Time Analysis:**
   The response time of the pins was within the expected range, ensuring smooth and error-free transitions.
3. **Voltage and Current Stability:**
   The system maintained consistent voltage and current levels, preventing overheating or component failure.

Image Caption Model Results:

1. **The Image is Successfully Processed and Caption is generated:**
   For the given test images, the outputs were generated successfully and are apt.
2. **Efficiency:**
   To maintain efficiency, various models have been developed to manage system requirements.
3. **Scalability:**
   Given that the models can be uploaded onto raspberry pi 5 which can further be developed into a standalone device.

Analysis / Interpretation of Results

Audio Translation Model:

1. **Successful Translation of Regional Languages:**
   The model is tested and works appropriately for all input regional languages.
2. **Bidirectional Translation:**
   The reply inputs were successfully converted back to english.
3. **Auditory Voicing:**
   The model is able to successfully read out the translated text.

Braille Cell and Keypad:

1. **Reliable Text Sequencing:**
   The text is being sequenced properly and the loop ends are clear and structured.
2. **The Keypad:**
   The only drawback noticed is the bounceback due to the tactile switch module, else the system works as expected.
3. **Interconnection:**
   The key presses are correctly being displayed on the braille cell.

Conclusion and Future Scope

Conclusion:

- Successfully implemented an AI-based remote meeting transcription system using Raspberry Pi 4 and Whisper.
- Achieved real-time, accurate speech-to-text transcription with multilingual support.
- The system is portable, cost-effective, and suitable for academic, professional, and assistive use cases.

Future Scope:

- Integrate cloud storage and real-time sharing for collaborative transcription.
- Enhance speaker diarization and keyword tagging for better meeting analytics.
- Expand capabilities with sentiment analysis and meeting summarization using LLMs.

References

[1] Gupta, R., Singh, P. K., & Bhanot, S. (2016). Design and implementation of Arduino based refreshable braille display controller. Indian Journal of Science and Technology, 9(33), 1-6.

[2] Rahimi, M., Shen, Y., Liu, Z., & Jiang, F. (2021). Printed Texts Tracking and Following for a Finger-Wearable Electro-Braille System Through Opto-electrotactile Feedback. arXiv preprint arXiv:2109.02385.

[3] Gozon, J., Li, A., Odulio, A., Sadurlo, J., Asaba, H., & Ong, C. (n.d.). A Digitally Automated Text to Braille Device for the Visually-Impaired.

[4] Chitte, P. P., Pimpalkar, Y. A., Nair, P. B., & Thombe, S. A. (2015). Braille to text and speech for cecity persons. International Journal of Research in Engineering and Technology, 4(01), 263-268.

[5] Shokat, S., Riaz, R., Rizvi, S. S., Khan, K., Riaz, F., & Kwon, S. J. (2020). Analysis and evaluation of braille to text conversion methods. Mobile Information Systems, 2020(1), 3461651.

[6] Li, Zhaoyang & Liu, Zhe & Xu, Sumei & Zhang, Kaijun & Zhao, Dazhe & Pi, Yucong & Guan, Xiao & Peng, Zhengchun & Zhong, Qize & Zhong, Junwen. (2024). Electrostatic Smart Textiles for Braille‐To‐Speech Translation. Advanced Materials. 36. e2313518. 10.1002/adma.202313518.

[7] Chen, H., Tao, W., Liu, C., Shen, Q., Wu, Y., Ruan, L., & Yang, W. (2023). A novel refreshable braille display based on the layered electromagnetic driving mechanism of braille dots. IEEE Transactions on Haptics, 16(1), 96-105.

[8]

[9] Rajkumar, E., & Doshi, S. (2021, April). Refreshable braille module using CAM actuated mechanism. In IOP Conference Series: Materials Science and Engineering (Vol. 1123, No. 1, p. 012028). IOP Publishing.

[10] Chen, D., Zhang, Y., Hu, X., Chen, G., Fang, Y., Chen, X., ... & Song, A. (2024). Development and Evaluation of Refreshable Braille Display and Active Touch-reading System for Digital Reading of the Visually Impaired. IEEE Transactions on Neural Systems and Rehabilitation Engineering.