

ASL - Alphabet Image Recognition

Team members:

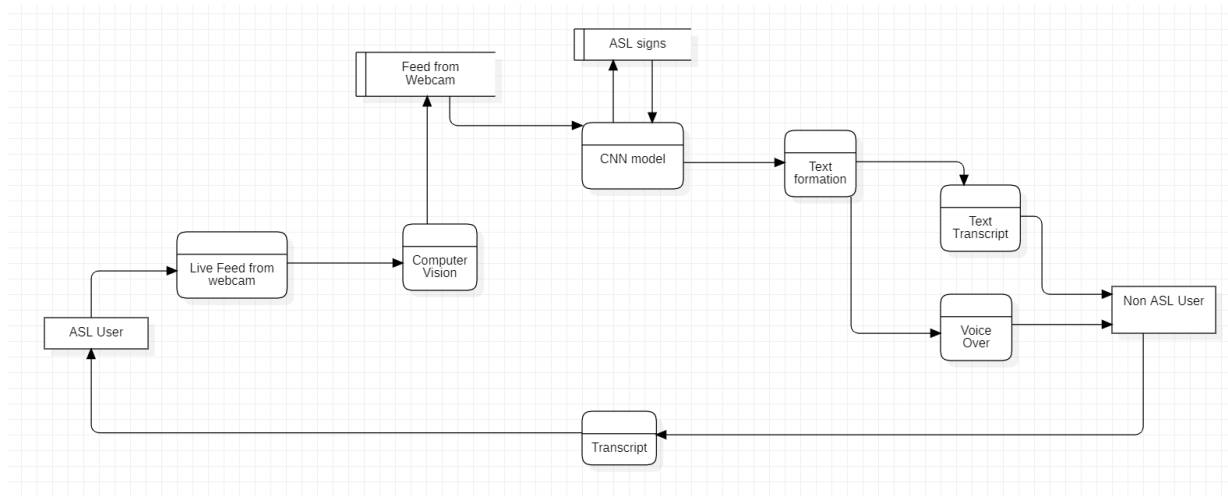
- Ananth Anand P B
- Adwyth Sumesh
- Joshua Jose

Proposed Solution

| S No. | Parameter | Description |
|-------|---------------------------|--|
| 1 | Problem Statement | How can we create a system that recognizes and translates ASL hand signs into text or speech, bridging the communication gap between deaf and hearing individuals? |
| 2 | Idea/Solution description | We're developing a functional API that utilizes a Convolutional Neural Network (CNN) trained on an ASL sign dataset and OpenCV for real-time video processing. This API will enable the recognition and translation of American Sign Language (ASL) hand signs into text or speech. |
| 3 | Novelty/ Uniqueness | Real-Time ASL Recognition: The ability to recognize ASL signs in real-time for instant communication. Accessibility: Making ASL communication accessible to both deaf and hearing individuals. Integration with Various Apps: The option to integrate your technology into different applications, including meeting platforms. Cutting-Edge Technology: Use of OpenCV and CNN for accurate ASL recognition. Text and Speech Output: Providing both text and speech outputs for user convenience. |

| | | |
|---|--------------------------------------|--|
| | | Subtitle Display: Enhancing accessibility for the deaf community with text subtitles. |
| 4 | Social Impact/ Customer satisfaction | <p>Better Communication: Fills the gap between ASL and non-ASL speakers for improved understanding.</p> <p>Inclusivity: Offers access for the deaf community, fostering equality and participation.</p> <p>Empowerment: Empowers ASL users by broadening their communication opportunities.</p> |
| 5 | Business Model | <p>We've opted for a licensing model as it suits our specialized ASL recognition technology. Licensing allows us to maintain ownership while generating upfront and potential ongoing revenue through license agreements. This approach works best for targeted industries and exclusive partnerships, offering revenue flexibility and IP protection.</p> |
| 6 | Scalability | <p>Integration with Video Conference Platforms:</p> <ul style="list-style-type: none"> • Widespread Adoption • Real-Time Communication • Global Reach <p>Expansion to Other Sign Languages:</p> <ul style="list-style-type: none"> • Increased Market • Diverse User Base • Continuous Growth • Inclusivity <p>Education and E-Learning:</p> <ul style="list-style-type: none"> • Applying the technology to education and e-learning platforms further enhances scalability and accessibility. |

Solution Architecture

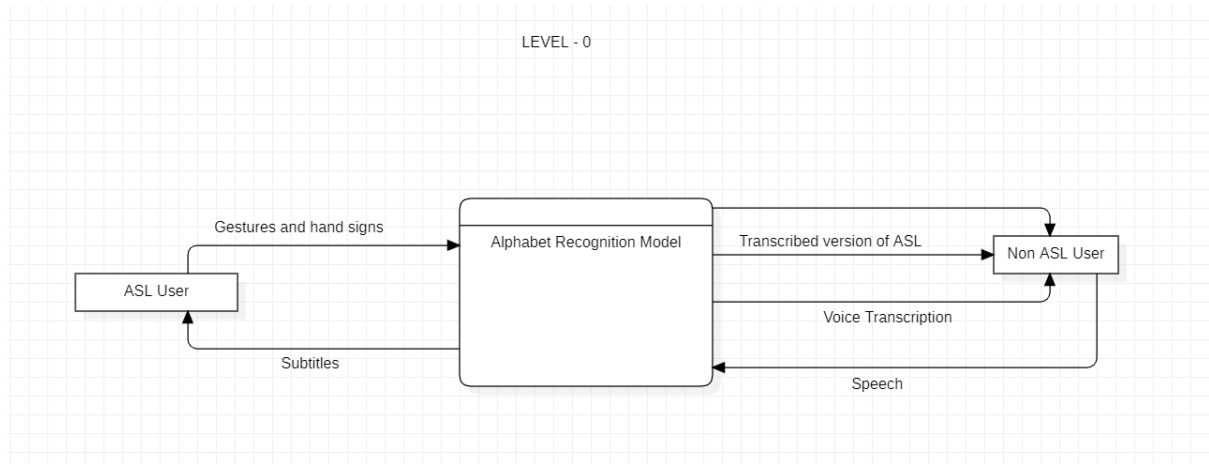


- The project starts with the ASL user signing the alphabet using American Sign Language.
- The camera captures a live video feed of the ASL user's hand gestures as they sign the alphabet.
- It is then processed by OpenCV which performs image preprocessing to enhance quality, detects the hand's location and boundaries, tracks hand movements, and segments the gestures from the background. Operating in real-time, OpenCV ensures that the live video feed is optimized for subsequent stages, such as ASL sign recognition by the CNN model, and facilitates the accurate interpretation of ASL signs in real-time.
- Next comes the CNN model (Convolutional Neural Network) which extracts key features from processed ASL sign images, employs convolutional layers to detect patterns, utilizes pooling layers to reduce data dimensionality, utilizes fully connected layers for mapping features to specific ASL sign categories, applies Softmax activation to generate probability scores, and ultimately predicts the ASL sign with the highest probability. This predicted sign corresponds to the user's sign and is the model's output, providing a textual representation of the ASL gesture.
- The generated text is transformed into speech using text-to-speech (TTS) technology.
- The output, in the form of spoken words, is made available to the non-ASL user who can understand the communication without knowing ASL.

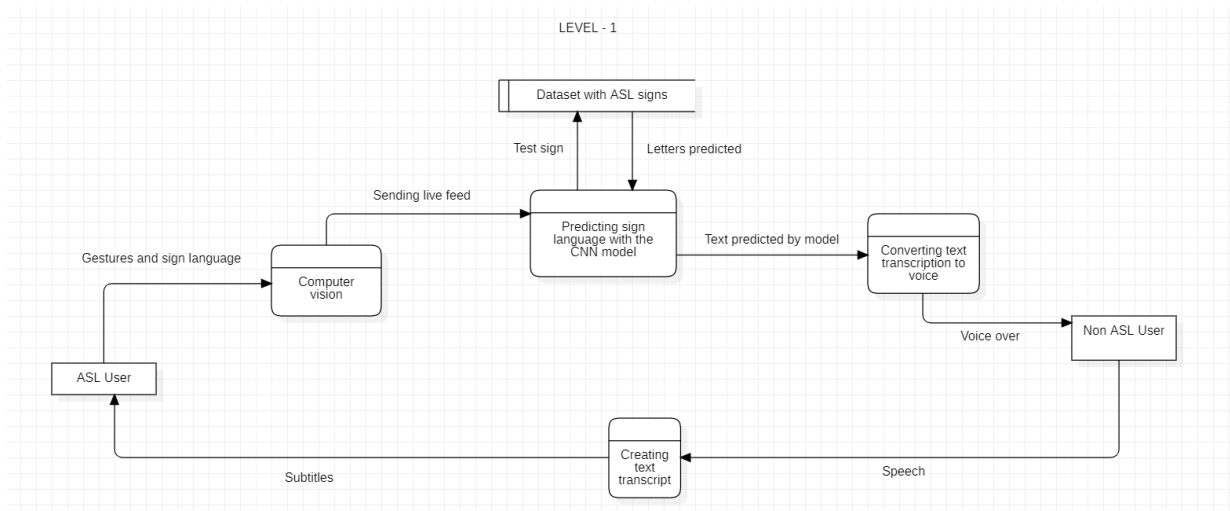
- To aid the deaf person, a text subtitle of the spoken words is displayed, ensuring they can comprehend the conversation effectively.

Data Flow Diagram

Level 0:



Level 1:



Level 2:

