Sign Language to Text

Team 8

Faculty of Computers and Information

Eng/ Amjad Dife

Team 8

- Aya Abdellatif Mohamed Ahmed - 162021077
- Pavly Samuel Atef Samuel - 162021080
- John Ashraf Helal Ajeeb - 162021093

Project Description

Sign language to text translation

Demo

Figure 1: Demo

Github Repository –Video Demo

Contribution

- Data is collected manually via a script
- Data is passed to Mediapipe for landmark detection
- Landmarks are passed to a custom-made model
- Model is used via a script to test the model

Data

- The dataset is created by a custom script.
- The dataset used to train the model consists of 241 images.
- Image format is converted to RGB.
- Image dimensions: captured at the default webcam resolution
- Data Augmentation: performed various data augmentation techniques (flipping, rotation, zooming)

Project architecture

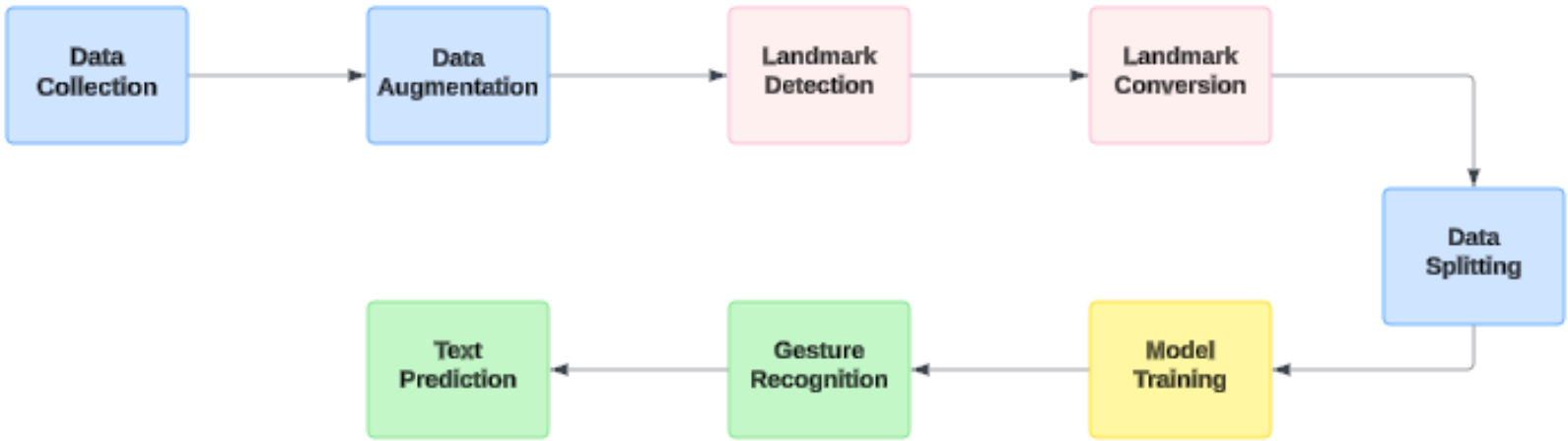

Figure 2: Project Architecture

Methods

- The model was built and trained using Keras.
- Integrates MediaPipe for hand landmark detection.
- Hyperparameters:
  - optimizer: adam
  - Loss Function: Categorical crossentropy
  - Metric: Accuracy
  - Epochs: 200

Methods: Layers

- Input Layer: Takes a 3D array of shape (21, 3, 1), representing 21 hand landmarks with 3 coordinates (x, y, z) each, and a single channel
- Hidden Layers:
  - Flatten Layer: Flattens the 3D input into a 1D vector of size (21 * 3 = 63).
  - Dense Layer 1: 256 neurons with ReLU activation for non-linearity.
  - Dropout Layer 1: 25
  - Dense Layer 2: 128 neurons with ReLU activation.
  - Dropout Layer 2: 25
  - Dense Layer 3: 64 neurons with ReLU activation.
  - Dropout Layer 3: 25
  - Dense Layer 4: 32 neurons with ReLU activation.
  - Dropout Layer 4: 25
- Output Layer: Dense layer with 9 neurons and softmax activation for multi-class classification.

Methods: Training

- Epochs: Trained for 200 epochs
- Early Stopping: Uses an EarlyStopping to prevent overfitting.

Results

- The measure is accuracy
  - The model achieves 99.1% accuracy on the test data.

Thanks!