Sign Language Detection

Content

Introduction

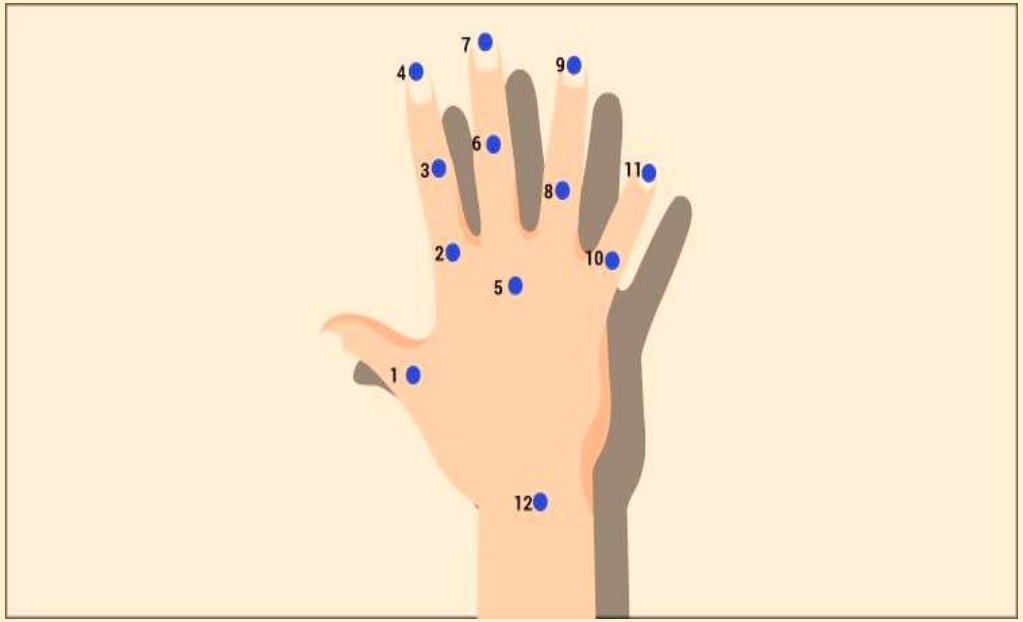
- Sign language is a language used as a manual communication method used by people who are deaf , mute, etc.
- Hand gesture is one of the main methods used in this language for non-verbal communication.
- In this project, we will be using the ISL (Indian Sign Language) dataset from Kaggle as well as ISLTranslate, which is a dataset containing frames and videos containing sign language, visual language, fingerspelling, and facial expressions in Indian Sign Language.
- The model will use a Deep Learning architecture that is efficient in Image recognition (Convolutional Neural Network Architecture).
- Using this model, we will train the model to recognize hand gestures and movement of hands with the dataset acquired.
- Once the model can successfully classify and recognize the images, it will generate English text according to Sign Language, which will make communication with mute and deaf people easy.

Technical Concepts (Algorithms) used

Using a CNN (Convolutional Neural Network) model for the classification of Indian Sign Languages involves several technical concepts and algorithms. CNN is a deep learning architecture known for its ability to process grid-like data such as images and videos effectively. When applying CNN architecture for Indian Sign Language classification, we encountered the following technical concepts and algorithms:

- **CNN Architecture**: Convolutional Neural Network (CNN) architecture is designed for processing and analysing computer vision tasks. It utilizes various layers to train very deep neural networks effectively on image and video dataset.

- **Convolutional Layers**: These layers automatically extract features from hand gesture images. They're crucial for recognizing patterns in images of the hand, helping the model identify signs language and classify them correctly.



- **Pooling layers**: These layers downgrade the dimensions of the image input by reducing its size. This helps in reducing the computational usage and focuses on important functions while retaining its features.

- **Fully Connected Layers:** These layers connect to every neuron in one layer with another layer, enabling the network to classify and predict based on the learned weights and features

- **Evaluation Metrics**: Metrics such as accuracy, precision, recall, F1-score, and AUC-ROC are used to assess the model's performance, ensuring its ability to correctly identify Hand Gestures.

Problem Statement

Sign language is a manual type of communication commonly used by deaf and mute people. It is not a universal language, so many deaf/mute people from different regions speak different sign languages. So, this project aims to improve the communication between deaf/mute people from different areas and those who cannot understand sign language. We are using deep learning methods which can improve the classification accuracy of sign language gestures.

LITERATURE SURVEY - 1

Title: Hand Gesture Recognition Based on Computer Vision: A Review of Techniques

Journal: Journal of Imaging

Methodology:

- The study looks at studies that investigate hand gestures as a nonverbal communication method in a variety of domains, including medical applications, robot control, human-computer interface (HCI), home automation, and communication for the deaf and silent.
- It groups the literature according to several methods, such as computer vision and instrumented sensor technology.
- Additionally, the article classifies hand gestures according to their posture, dynamic/static nature, or hybrid forms.

Research Gap:

- The present study largely ignores real-world healthcare applications in favour of computer applications, sign language, and virtual environment engagement.
- The majority of studies place more emphasis on developing algorithms and improving frameworks than on actually implementing healthcare practices, which indicates a large research vacuum in this area.

LITERATURE SURVEY - 2

Title: Hand Gesture Recognition: A Literature Review

Journal: International Journal of Artificial Intelligence & Applications

Methodology:

- The literature discusses techniques including orientation histogram for features representation, fuzzy c-means clustering, neural networks (NN), and hidden markov models (HMM).
- In particular, HMM techniques perform well in dynamic gestures, especially in robot control scenarios.
- Neural networks play a crucial role in hand form recognition as classifiers. In gesture recognition systems, methods for extracting features—such as algorithms for capturing hand shape—are essential.

Research gap:

- Their uses have been well documented in the literature, there are still plenty of unanswered questions regarding the real-world applications of these technologies, particularly in healthcare settings.
- Moreover, although the article address current recognition methods, a thorough assessment and comparison of these systems in practical healthcare settings is lacking.

LITERATURE SURVEY - 3

Title: An Exploration into Human–Computer Interaction: Hand Gesture Recognition Management in a Challenging Environment

Journal: SpringLink

Methodology:

- The methodology entails a methodical examination of pertinent literature to pinpoint important developments, strategies, and obstacles in hand gesture detection and human-computer interaction.
- After that, an image dataset is chosen for analysis, and then picture enhancement and segmentation procedures are carried out.
- By isolating the main image from its backdrop, converting colour spaces, and lowering background noise, these methods help to improve the quality of raw photographs.
- The next phase is using machine learning methods, namely Convolutional Neural Networks (CNN), to recognize hand gestures and learn attributes.

Research Gap:

- In order to provide a fair and impartial model, the study emphasises the significance of contrasting analytical and discriminatory inclinations.

LITERATURE SURVEY - 4

Title: Real-Time Hand Gesture Recognition Using Fine-Tuned Convolutional Neural Network

Journal: Sensors

Methodology:

- The research paper highlights the benefits and drawbacks of different sensors for the development of HGR systems through a methodological comparison.
- Hand regions are identified and resized using image enhancement and segmentation algorithms to match the input sizes of pre-trained Convolutional Neural Networks (CNNs).
- Hand region segmentation is achieved by using maximum-area-based filtering algorithms and depth thresholding.
- Additionally, the study uses a score-level fusion strategy that combines, using a sum-ruled-based fusion procedure, the normalised score vectors from two fine-tuned CNNs.

Research Gap:

- There are important concerns, aspects like user comfort, real-time performance, and adaptation to changing contexts are not fully covered.
- Furthermore, the research leaves gaps in comprehending the larger ramifications and possible social repercussions of HGR systems because it focuses exclusively on technical techniques. To close these gaps and create all-encompassing HGR systems that take into account both real-world usability issues and technological constraints, more study is required.

Dataset and input format

➔ **Input Format:**

Hand Gesture Image Data:

High-resolution Hand Gesture images should make up the majority of the dataset. A variety of alphabets and numbers which are labeled with hand gestures averaging about 1200 images must be covered by these photographs

Image Labels:

For the supervised learning approach to work, accurate alphabets and numeric labels are essential. The precise labels that each hand gesture image depicts will be identified by a clear label. To enable model training, a reliable and consistent labeling procedure will be adopted.

Objective

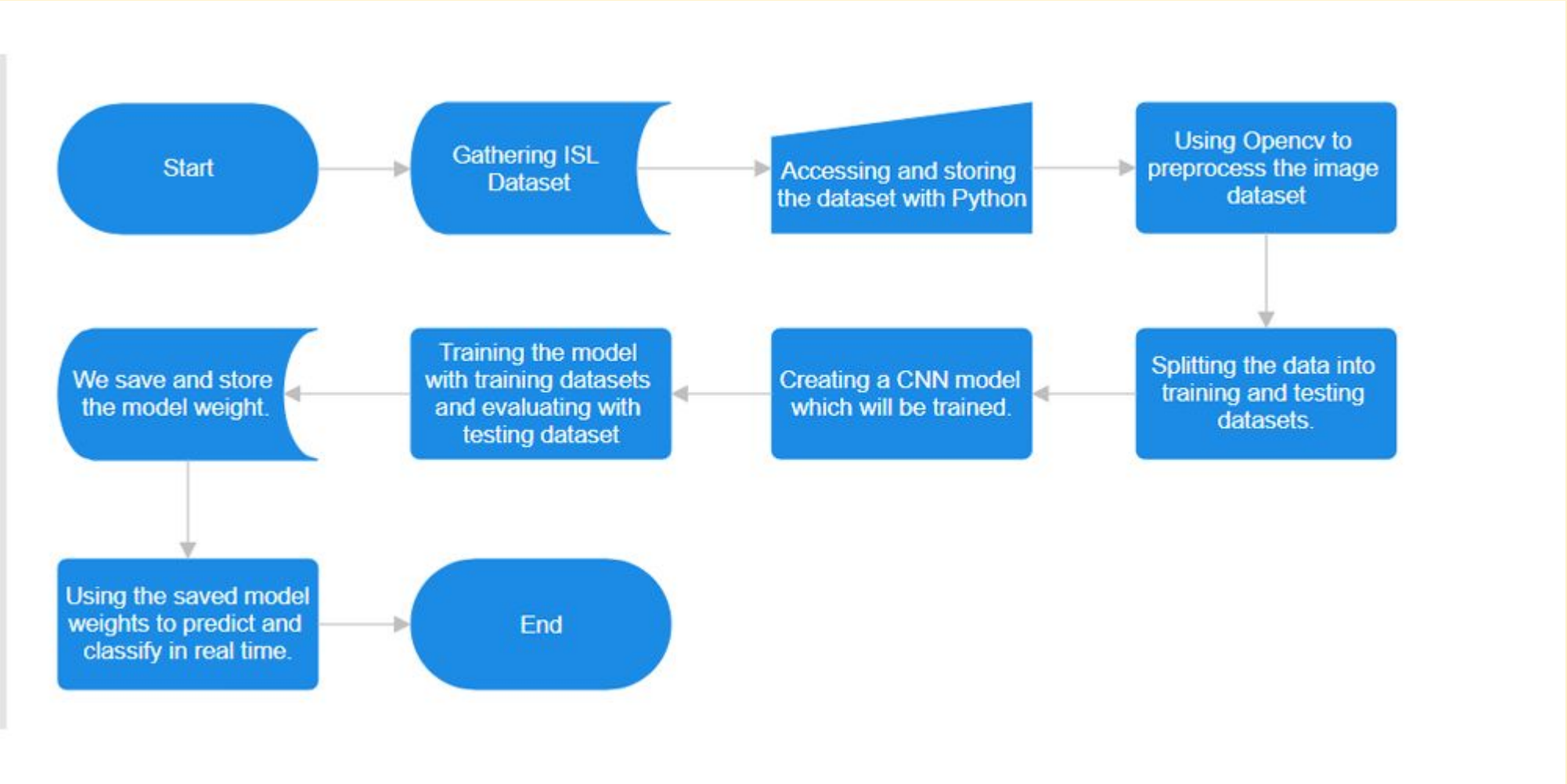
Main Objective

- To detect and classify the hand gesture used for sign language with high accuracy and precision.

Sub Objective

- To use the trained model to detect and classify images.

Steps:-


Start
Gathering ISL Dataset
Accessing and storing the dataset with Python
Using Opencv to preprocess the image dataset
Splitting the data into training and testing datasets.
Creating a CNN model which will be trained.
Training the model with training datasets and evaluating with testing dataset
We save and store the model weight.
Using the saved model weights to predict and classify in real time.
End

Working of the Model

1.Character Mapping

- A map_characters dictionary is created to link model output indices (0 to 34) to specific characters or numbers .This will allow us to translate model predictions into readable gestures.

2.Image Input :An image of a hand gesture is taken as input.

3.Edge Detection (Preprocessing Step)

The edge_detection function enhances the outline of the hand, making the key parts of the gesture stand out and easier for the model to interpret.

- **Blurring**: Gaussian blur is applied to the image to reduce noise, ensuring that the model isn't distracted by small, irrelevant details in the background.
- **Adaptive Thresholding**: A binary threshold is applied to highlight the edges of the hand. This emphasizes the shape and contours of the gesture while removing non-essential details.

4.Image Preprocessing and Model Input Formatting

- **Grayscale Conversion**: The image is converted to grayscale, simplifying it to a single color channel, which makes it easier for the model to process.
- **Resizing**: The grayscale image is resized to 64x64 pixels to match the input size expected by the CNN model.
- **Normalization**: The pixel values are scaled between 0 and 1, stabilizing the training and prediction processes.
- **Reshaping**: The preprocessed image is reshaped adding batch and channel dimensions so it fits the model's expected input shape.

5.Model Prediction

- After preprocessing, the image is fed into the pre-trained CNN model, which processes the image through its layers:
  - **Convolutional Layers** extract features by detecting edges, shapes, and unique patterns(Marking of landmark points) that characterize each gesture.
  - **Pooling Layers** reduce the spatial dimensions, helping the model focus on prominent features while reducing computational complexity.
  - **Fully Connected (Dense) Layers** interpret the extracted features to classify the gesture.
- The **softmax output layer** provides probabilities for each of the 35 classes (0-34).

6.Prediction Interpretation and Character Mapping

- The class with the highest probability is selected as the predicted class index.
- This class index is then mapped to a specific character or number using the map_characters dictionary, translating the numeric prediction into a readable gesture label (e.g., "A", "B", "1").

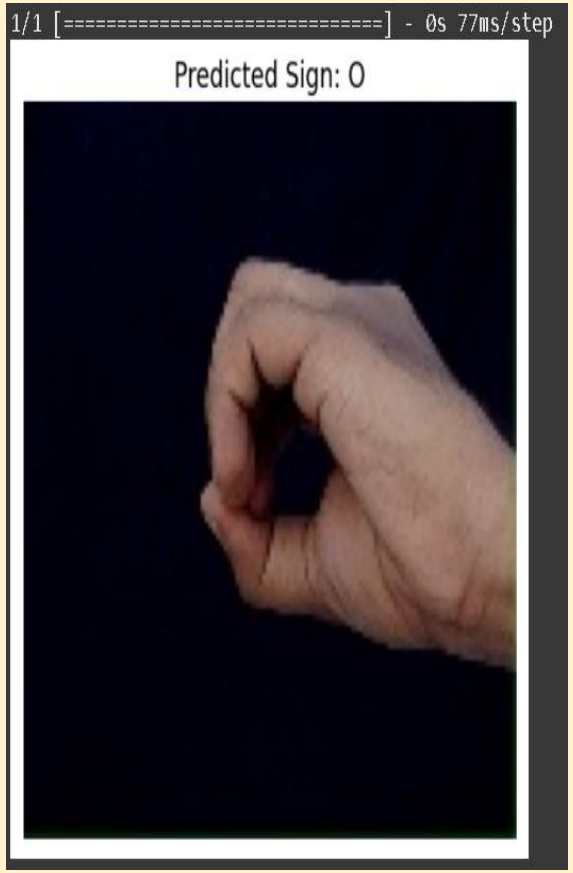
7.Output Display

- Finally, the input image is displayed alongside the predicted character or number, allowing the user to see both the gesture and the model’s prediction.

Results


1/1 0s 218ms/step
Predicted Sign: 8


1/1 [==========] - 0s 75ms/step
Predicted Sign: H


1/1 [==========] - 0s 77ms/step
Predicted Sign: O


1/1 [==========] - 0s 74ms/step
Predicted Sign: X

Conclusion

- The sign language detection model using CNN effectively classifies and recognizes gestures in Indian Sign Language (ISL).

- The project utilized an ISL dataset from Kaggle and ISLTranslate, achieving gesture recognition with good accuracy.

- This model demonstrates potential to improve communication for deaf and mute individuals.

- Future work could focus on:
- Expanding the variety of gestures covered.
- Enhancing real-world usability and accuracy.

References

- https://www.google.com/search?q=hand+gesture+recognition+cnn+image&tbm=isch&hl=en-GB&tbs=rimg%3ACdJN_1KyjlLpaYcaramXggfJWsgIRCgIIABAAOgQIARAAVfCqAD_1AAgDYAgDgAgA#imgrc=AUONS3HNoRhd3M
- Oudah, Munir, Ali Al-Naji, and Javaan Chahl. 2020. "Hand Gesture Recognition Based on Computer Vision: A Review of Techniques" Journal of Imaging 6, no. 8: 73. Available at: https://www.mdpi.com/2313-433X/6/8/73 [1]
- Chang, V., Eniola, R.O., Golightly, L. et al. An Exploration into Human–Computer Interaction: Hand Gesture Recognition Management in a Challenging Environment. SN COMPUT. SCI. 4, 441 (2023). https://doi.org/10.1007/s42979-023-01751-y [3]
- Sahoo JP, Prakash AJ, Pławiak P, Samantray S. Real-Time Hand Gesture Recognition Using Fine-Tuned Convolutional Neural Network. Sensors (Basel). 2022 Jan 18;22(3):706. doi:10.3390/s22030706. PMID: 35161453; PMCID: PMC8840381. [4]
- Oudah M, Al-Naji A, Chahl J. Hand Gesture Recognition Based on Computer Vision: A Review of Techniques. J Imaging. 2020 Jul 23;6(8):73. doi: 10.3390/jimaging6080073 PMID: 34460688; PMCID: PMC8321080. [6]

Thank You.