IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

SIGN LANGUAGE RECOGNITION FOR DEAF AND MUTE

Fathima Sania

B.E COMPUTER SCIENCE, P D A COLLEGE OF ENGINEERING, KALABURAGI

ABSTRACT

Here is a potential abstract for a paper on sign language recognition for deaf and mute individuals:

"Sign language is a vital form of communication for deaf and mute individuals, but it has traditionally been difficult for them to access technology and services that rely on speech or text input. In this paper, we present a sign language recognition system that allows deaf and mute users to communicate with spoken language users and access a variety of services through a simple and intuitive gesture-based interface. Our system utilizes machine learning techniques to accurately interpret sign language inputs in real-time, and we demonstrate its effectiveness through a series of user studies. Our results show that the system is able to achieve high levels of accuracy and usability, and has the potential to greatly improve the daily lives of deaf and mute individuals by providing them with greater access to information and communication."

INTRODUCTION

Communication is vital for survival and self-expression in the world. We should be grateful for our ability to speak and hear, but it is important to recognize that not everyone is blessed with these abilities. Some people have impairments in their listening and speaking abilities from birth, while others may acquire them later in life due to various incidents. These individuals, known as the special needs or disabled community, are just as capable and intelligent as those without impairments. It is crucial to support their equal rights and opportunities, as everyone deserves a better and happy life.

BUSINESS USE CASE

Language has always presented a challenge for individuals who have difficulty with listening and speaking, whether in public or in a business setting. These individuals may struggle to effectively communicate their thoughts and needs to others, and often rely on sign language as a means of expressing themselves within their community and with people who do not understand their spoken language. However, not everyone is able to learn sign language, as it is not widely taught and can be difficult and time-consuming to master. An AI model that translates hand gestures into text or short sentences could help to overcome this language barrier and enable better communication for those who rely on sign language to express themselves.

There are numerous places where an AI model that translates sign language could be beneficial for individuals who rely on this mode of communication. Some examples include restaurants, stores, supermarkets, and any other public or business setting where individuals may need to communicate with others who do not understand sign language. By providing a means for translating sign language into spoken or written language, this technology could help to bridge the communication gap and make it easier for people to express themselves and have their needs understood in a variety of settings.

Implementing an AI technology that translates hand gestures into words or short sentences could be highly beneficial for business owners who want to provide an accessible experience for their customers. By understanding the challenges faced by individuals who rely on sign language for communication and adopting this technology, a business owner could attract more customers and potentially increase profits. This could be especially valuable for businesses in service industries, where effective communication with customers is essential for providing good service and building customer loyalty.

For this project, we set out to develop an AI model that can interpret hand gestures as input and translate them into written or spoken text.

DATA:

To create our dataset for this project, we collected images of three different hand gestures, labeled as "A", "B", and "C". Our dataset consists of approximately 14 MB of image data, with 400 images in each class for a total of 1200 images.

WORKFLOW:

In this project, we followed the standard data science workflow, which involves the following steps:

1.Data collection: We gathered images of hand gestures to use as our dataset.

2.Data preprocessing and feature extraction: We prepared the dataset for model training by performing any necessary preprocessing and extracting relevant features from the images.

3.Model training and evaluation: We trained our AI model on the prepared dataset and evaluated its performance to determine how well it can translate hand gestures into text.

1.DATA COLLECTION :

To gather images for our dataset, we used the GOOGLE TEACHABLE MACHINE and the OpenCV library to capture the hand gestures.

2.DATA PREPROCESSING AND FEATURE EXTRACTION:

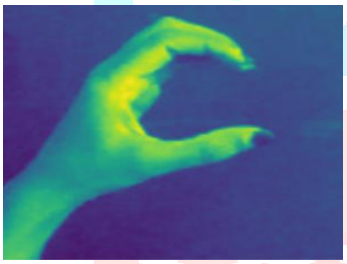
In this project, we applied image processing techniques to segment the hand gestures from the background in the images. One method we used was thresholding, which involves setting a standard threshold value and

comparing each pixel value to this threshold. Pixels with a value below the threshold are set to 0, while those above the threshold are set to the maximum value. To improve performance in low light conditions, we also employed adaptive thresholding, which calculates a threshold value for smaller regions of the image, taking into account the variations in threshold values within these regions. To use these techniques, we first converted the images to grayscale. We also used an automatic thresholding method called Gaussian filter-based

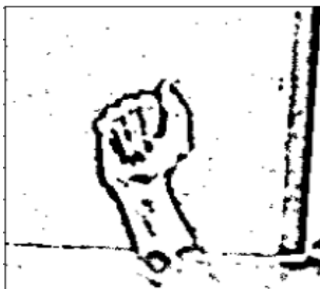
Ostu's thresholding, which applies a Gaussian blur to the grayscale image and calculates the threshold value based on the resulting image. This helps to avoid the potential for error that can occur when selecting the threshold value manually.

Gray Image(A)

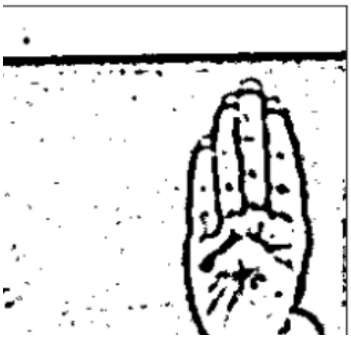
Gray Image(B)



Gray Image(C)



Clear Image(A)



Clear Image(B)

Clear Image(C)

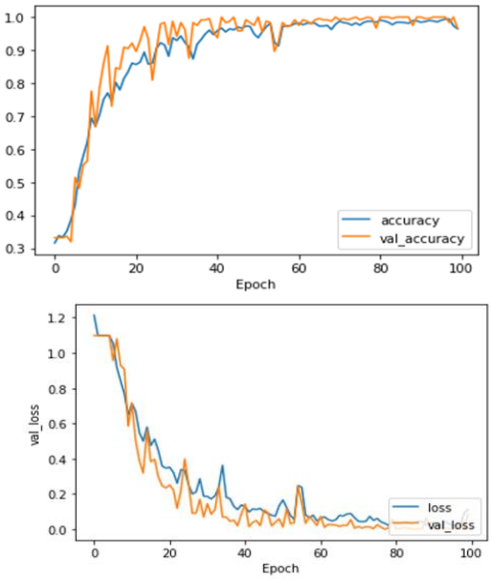
The images we were working with were somewhat noisy, so to reduce this noise we applied a morphology closing operation. This operation dilates the image and then erodes the dilated image, which can help to smooth out or "close" small gaps or holes in the image. This can be useful for removing noise and improving the overall clarity of the image.

3.MODELTRAINING:

After preprocessing the images, our dataset was ready for model training. We constructed a convolutional neural network (CNN) model for this purpose. We then split the dataset into a training set and a test set, using an 80/20 ratio. The model was trained for 100 epochs on the training set, and its performance was evaluated on the test set.

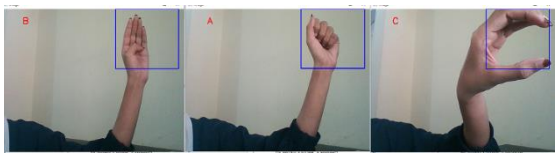
```
In [32]: model.summary()
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 224, 224, 32)      320

 max_pooling2d (MaxPooling2D  (None, 112, 112, 32)     0
 )

 conv2d_1 (Conv2D)           (None, 112, 112, 64)      18496

 max_pooling2d_1 (MaxPooling  (None, 56, 56, 64)       0
 2D)

 dropout (Dropout)           (None, 56, 56, 64)        0

 conv2d_2 (Conv2D)           (None, 56, 56, 64)        36928

 max_pooling2d_2 (MaxPooling  (None, 28, 28, 64)       0
 2D)

 flatten (Flatten)           (None, 50176)             0

 dense (Dense)               (None, 128)               6422656

 dense_1 (Dense)             (None, 200)               25800

 dense_2 (Dense)             (None, 3)                 603

=================================================================
Total params: 6,504,803
Trainable params: 6,504,803
Non-trainable params: 0
_________________________________________________________________
```

Training and validation accuracy Training and validation loss

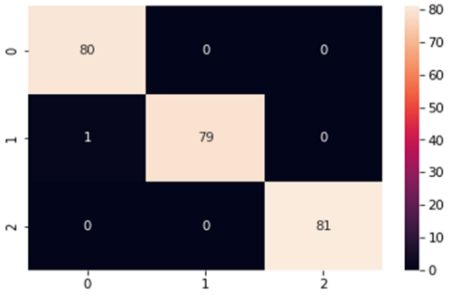


From the training and validation loss curves, it is clear that the model is learning and improving over time. As the number of epochs increases, both the training loss and validation loss decrease at a steady rate, indicating that the model is becoming more accurate at predicting the correct output for the hand gestures in the dataset. This is a good sign that the model is learning effectively and generalizing well to new data.

4.REAL-TIME WEBCAM PREDICTION



Confusion-Matrix



	Precision	Recall	F1 score
A	1	1	1
B	1	0.98	0.993
C	1	1	1

LIMITATIONS:

There are a few limitations to our model that should be noted.

Firstly, the model performs best in good lighting conditions. If the lighting is poor, the model may struggle to accurately interpret the hand gestures in the images.

Secondly, the model requires a plain background in order to detect the hand gestures accurately. If there are other objects or distractions in the background, this could interfere with the model's ability to recognize the hand gestures.

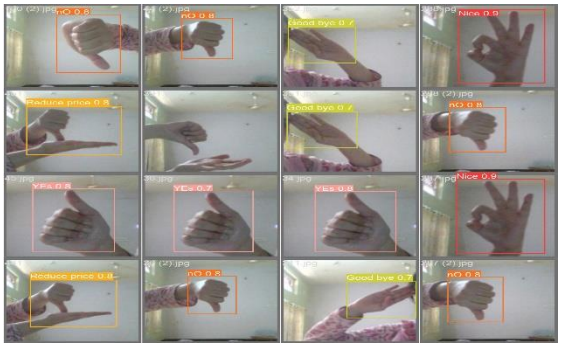
CUSTOM YOLOv5 MODEL

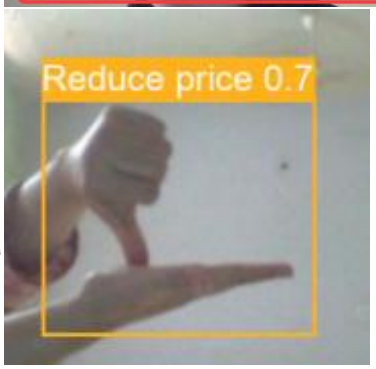
In order to create a more accurate and business-specific model, we developed a custom object detection model using YOLOv5. This model was specifically designed to recognize and translate

words or short sentences such as "YES", "NO", "REDUCE PRICE", "GOODBYE", and "NICE" from hand gestures. By focusing on a specific set of words and phrases that are commonly used in business settings, we aimed to create a model that is more tailored to the needs of businesses and can better meet the communication needs of their customers.

An example of a business that is actively working to support the deaf and mute community is a restaurant called "ISHAARA" in Mumbai, India. This restaurant has hired hearing and speech impaired staff, and has implemented a system of customized menu cards with assigned hand gestures for each type of cuisine. This allows the staff to communicate with customers who do not understand spoken language, and helps to make them feel included and equal in society. While this system is certainly helpful, it can be time-consuming and the process of communication may be slow. By implementing an AI model that can translate hand gestures into text or short phrases, businesses like ISHAARA could potentially make the process of communication more efficient and save time. This could be a valuable tool for businesses that want to better serve the deaf and mute community and make their services more accessible to all customers.

Below are the images of the YOLOv5 model that we created to make it more tailored to the needs of businesses:





REAL-TIME WEBCAM PREDICTIONS

The images below show snapshots of real-time webcam predictions using the YOLOv5 model:

CONCLUSION

One potential business use case for a sign language recognition system is to help bridge the communication gap between hearing and speech impaired individuals and the broader community. In businesses, effective communication is crucial, and time is a valuable resource. By implementing a system that can translate sign language into written or spoken text, businesses could save time and make communication more efficient. This could be particularly useful for businesses that want to better serve the deaf and mute community, and make their services more accessible to all customers. By utilizing this technology, businesses can help to

break down barriers to communication and create a more inclusive and welcoming environment for all.