# Research Proposal

ANKITA SINGH
(22MS0021)
Under the supervision of
Professor Subhashis Chatterjee
Department of Mathematics and Computing
IIT(ISM) Dhanbad

## Title

Decoding Silent Communication: Utilizing Advanced Deep Learning for Converting Sign Language to Text

## Introduction

Even in today's contemporary society, a lack of awareness about sign languages persists, making it challenging for many to communicate effectively with individuals who are deaf or hard of hearing.

A significant contributing factor to this challenge is the limited awareness surrounding sign languages, which are crucial mediums of communication for the Deaf community. Bridging this communication gap is not only a matter of inclusivity but also a step towards a more accessible and equitable world.

In our project I basically focus on producing a model which can recognise few conversational hand gestures. The gestures we aim to train are as given in the image below.


MACHINE LEARNING

Machine Learning (ML) is a branch of artificial intelligence (AI) that empowers computer systems to learn and improve from experience without explicit programming.

Supervised Learning: In supervised learning, algorithms are trained on labeled datasets, making predictions or decisions based on input data paired with corresponding output labels. This approach is prevalent in tasks such as image classification and speech recognition.

Unsupervised Learning: Unsupervised learning involves algorithms exploring unlabeled data to discover inherent patterns or structures. Clustering and dimensionality reduction are common applications, aiding in tasks like customer segmentation or anomaly detection.

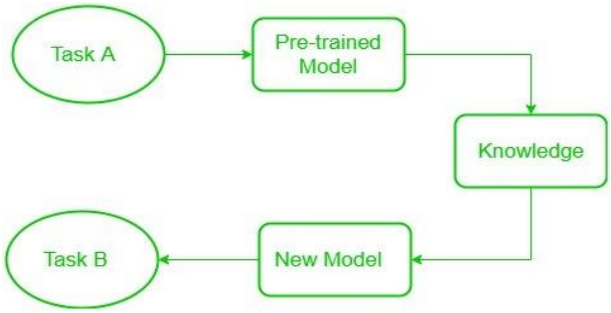
Reinforcement learning: Reinforcement learning, on the other hand, relies on agents learning from interactions with an environment to maximize cumulative rewards.

DEEP LEARNING

- Deep Learning is a branch of machine learning and is mainly used for image classification, object detection and natural language processing
- Deep Learning is an algorithm based on a neural network for automatic feature selection of data.
- It does not need a lot of artificial feature engineering.
- Its accuracy and generalization ability are improved compared to those of traditional methods in image recognition and target detection.

TRANSFER LEARNING

Transfer learning is like using knowledge gained from one task to help with another. Imagine you learned to ride a bicycle, and now you want to learn to ride a skateboard. The balance and coordination skills you gained from riding a bicycle can help you learn to ride a skateboard faster. Similarly, in transfer learning, a computer learns from one task and then uses what it learned to do better on a different task. This is different from some other algorithms that start learning each task from scratch. Transfer learning can be like building on what you already know, making it a smart and efficient way for computers to tackle new challenges.




## Literature Review

Image processing is one of the main task in ML algorithms for the correct classification of images into their respective classes based on common features. DL algorithms have been heavily used in research where proposed images are input into DL algorithms that extract features and classify images.Aggarwal, S.; Suchithra, M., et al conducts a thorough analysis of AI and ML techniques for rice disease diagnosis.

Focused specifically on image classification, this review outlines the various deep learning approaches employed in the field. It discusses the strengths and limitations of different

models, shedding light on their efficacy in diverse applications. Kumar, P.; Rajput, N.; Verma, R., et al. Deep Learning Approaches for Image Classification: A Review

## Problem Statement

Building a deep learning Model for conversion of sign to text and making it user friendly by building a web app using Streamlit.

## Research Objective

1. **Deep Learning Model Construction:**
    - Develop a deep learning model specifically tailored for the conversion of sign language gestures into textual representations.
2. **User-Friendly Web Application:**
    - Implement a user-friendly web application using Streamlit, ensuring a seamless and intuitive interface for users to interact with the sign-to-text conversion functionality.

## Methodology

**1. Data Collection:**
Utilize OpenCV to capture images, collecting a minimum of 5 images for each sign involved in the conversion process outlined in this thesis.

**2. Image Labeling:**
Establish the Label IMG framework for efficient image labeling.
Label collected images to facilitate subsequent image detection via Single Shot Multibox Detector (SSD).

**3. Data Preprocessing:**
Employ an appropriate sampling method to divide the collected dataset into training and testing datasets.

**4. Model Training:**
Apply transfer learning techniques to train the deep learning model.

**5. Update Checkpoints:**
Implement checkpoints to save the current state of the pre-trained model during the training process.

**6. Real-Time Detection:**
Develop an interactive web interface using Streamlit to enable real-time detection of sign language and its conversion to text.

## Timeline

| Month | January | February | March | April |
|---|---|---|---|---|
| Proposal Submission | | | | |
| Literature survey | | | | |
| Dataset Collection | | | | |
| Development of model | | | | |
| Thesis writing | | | | |

## Conclusion

This thesis explores using deep learning methodologies to turn sign language into written words. The goal is to create a new and helpful solution that makes it easier for deaf or hard-of-hearing people to connect with others. By using the latest technology, this research aims to add something important to the ongoing conversation about making sure everyone feels included and can communicate well in today's world.

## References


- Zaki, M.M., Shaheen, S.I.: Sign language recognition using a combination of new vision based features. Pattern Recognition Letters 32(4), 572–577 (2011)
- N. Mukai, N. Harada and Y. Chang, "Japanese Fingerspelling Recognition Based on Classification Tree and Machine Learning," 2017 Nicograph International (NicoInt), Kyoto, Japan, 2017, pp. 19-24. doi:10.1109/NICOInt.2017.9
- aeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks-Part-2/

- Pigou L., Dieleman S., Kindermans PJ., Schrauwen B. (2015) Sign Language Recognition Using Convolutional Neural Networks. In: Agapito L., Bronstein M., Rother C. (eds) Computer Vision - ECCV 2014 Workshops. ECCV 2014. Lecture Notes in Computer Science, vol 8925. Springer, Cham
- https://opencv.org/
- https://en.wikipedia.org/wiki/TensorFlow
- https://en.wikipedia.org/wiki/Convolutional_neural_network