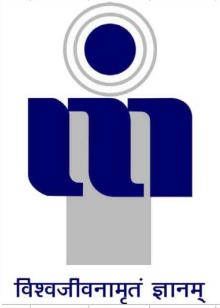
Sign Language to Text Translation

by
Aniket Sharma
2019BCS-008



Supervisor: Prof. Shashikala Tapaswi

Table of Contents

01

Introduction

- Sign languages are natural languages, with their grammar and lexicon.
- Out of more than 150 sign languages worldwide, ASL is the most widely studied.
- This work focuses on Fingerspelling component of ASL and aims to use Computer Vision for Sign Language Translation.

Objective

The main objective of the work is to develop an AI system capable of translating Sign Language without requiring any specialized hardware.

02
Literature Review

Three different strategies exists for automated Sign Language translation.

Using Specialized Hand Tracking Devices

- High accuracy
- Infeasible due to high cost requirements

Using Depth Images

- High accuracy
- Still need specialized hardware

Using Computer Vision

- Use daily-use cameras
- Difficult to achieve good accuracy

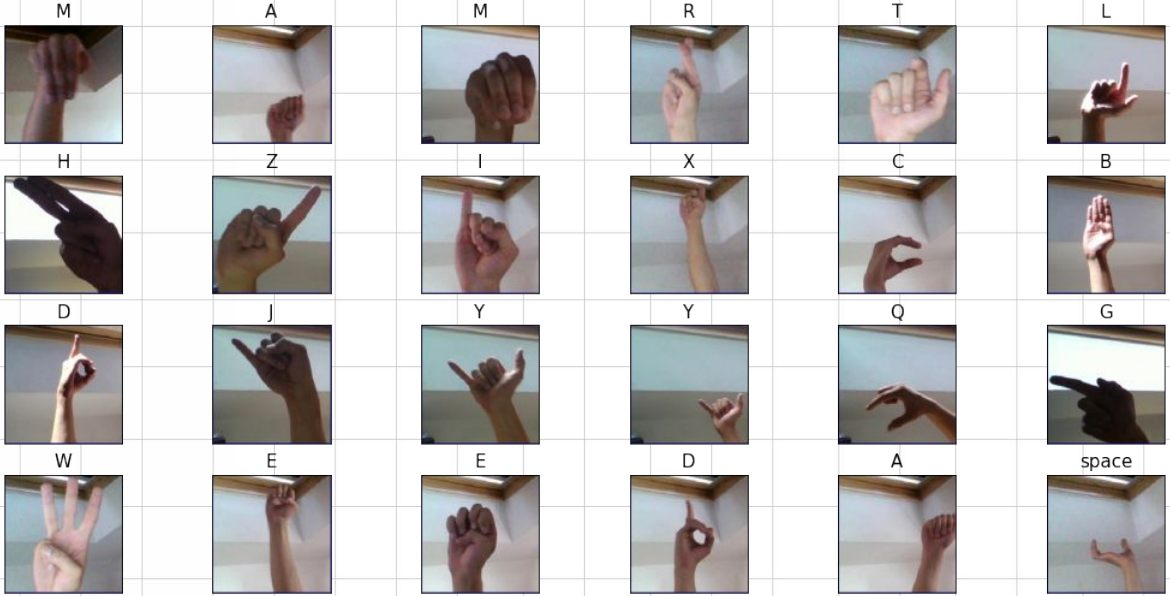
03

Methodology

Data used for Training

- consists of 84,000 images
- 28 classes
  - 26 for alphabets
  - 1 for ‘space’
  - 1 for ‘nothing’
- 200x200 pixels size
- colored

Data used for Training



Basic CNN Model

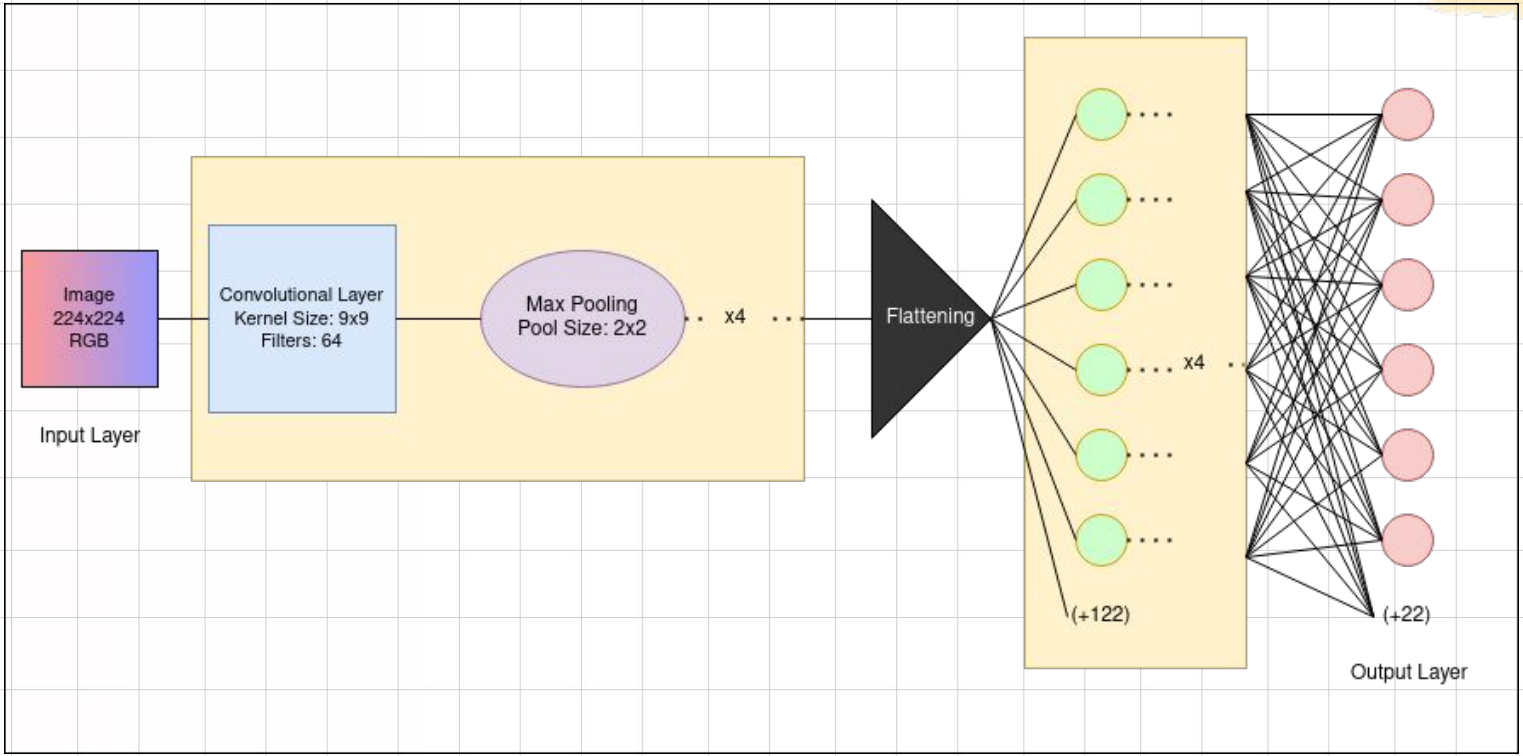


Image
224x224
RGB
Input Layer
Convolutional Layer
Kernel Size: 9x9
Filters: 64
Max Pooling
Pool Size: 2x2
x4
Flattening
x4
(+122)
(+22)
Output Layer

Transfer Learning

Pre-trained Model Performance

Model	Validation Accuracy	Training Time (sec.)
EfficientNetB0	89.00%	177
DenseNet201	88.73%	452
EfficientNetB5	88.62%	575
DenseNet169	88.45%	307
EfficientNetB7	88.23%	999
ResNet152	88.15%	688
EfficientNetB3	88.08%	331

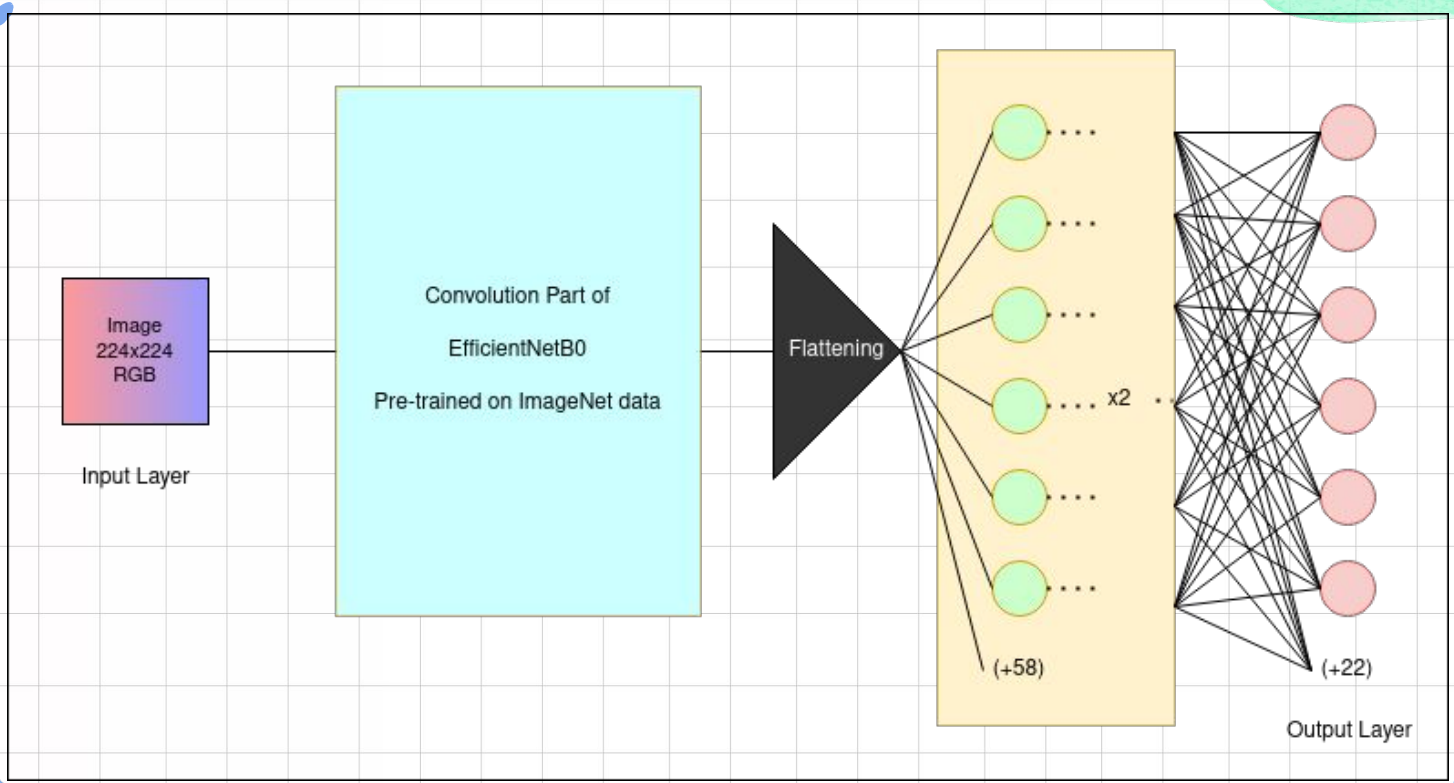


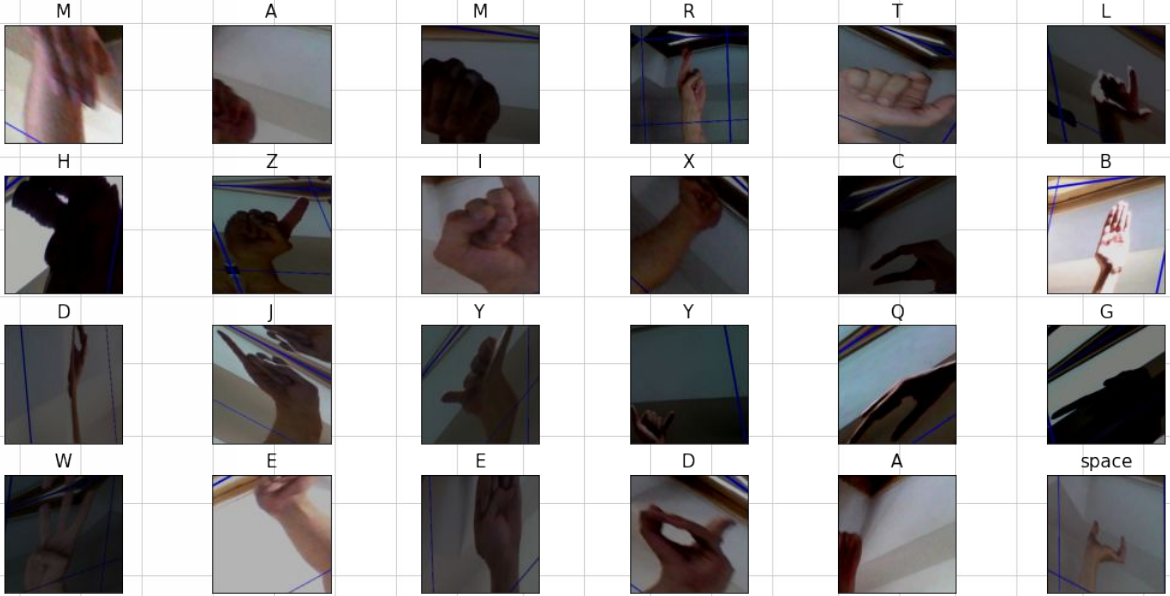
Image
224x224
RGB
Input Layer
Convolution Part of
EfficientNetB0
Pre-trained on ImageNet data
Flattening
x2
(+58)
(+22)
Output Layer

Data Augmentation

Data Augmentation

Augmentation Type	Value
Rotation	20°
Width Shift	10%
Height Shift	10%
Brightness	20%-100%
Shear	45°
Zoom	50%-150%
Channel Shift	100 px

Augmented Data



Pre-trained Model Performance

Model	Validation Accuracy	Training Time (sec.)
MobileNet	65.95%	933
ResNet152	62.62%	1177
EfficientNetB5	62.57%	1082
EfficientNetB7	61.69%	1251
DenseNet201	61.49%	1067
ResNet101	61.01%	1100
ResNet101V2	60.46%	1056

Ensemble Model

- Maximum Voting Ensemble
- Made of 5 Basic CNN models
- Each of the 5 models trained on a different set of Bagging data

Evaluation Metrics

Models have been evaluation on **four** classification metrics:

- Accuracy
- Precision Score
- Recall Score
- F-Score

04
Experiments & Results

ASL Dataset

- dataset used for training
- consists of 84,000 images
- 28 classes
  - 26 for alphabets
  - 1 for ‘space’
  - 1 for ‘nothing’
- 200x200 pixels size
- colored

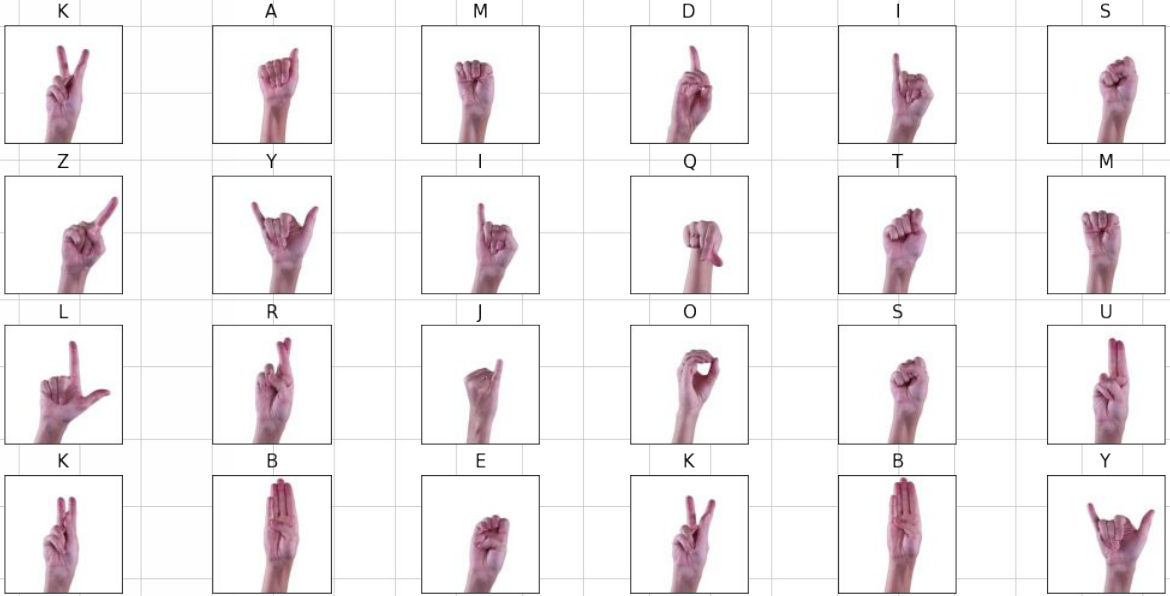
ASL Dataset

Model	Accuracy	Precision	Recall	F-Score
Basic CNN Model	95.71%	0.958	0.957	0.957
Transfer Learning CNN	98.12%	0.982	0.981	0.981
Basic CNN Model with Data Augmentation	95.72%	0.961	0.957	0.957
Transfer Learning CNN with Data Augmentation	94.95%	0.951	0.949	0.949
Ensemble Model	99.99%	0.999	0.999	0.999

ASL Alphabets

- consists of 832 images
- 26 classes
  - 26 for alphabets
  - 32 images per class
- 224x224 pixels size
- colored

ASL Alphabets



ASL Alphabets

Model	Accuracy	Precision	Recall	F-Score
Basic CNN Model	37.02%	0.230	0.344	0.257
Transfer Learning CNN	42.31%	0.323	0.407	0.315
Basic CNN Model with Data Augmentation	36.78%	0.259	0.368	0.269
Transfer Learning CNN with Data Augmentation	43.39%	0.422	0.434	0.380
Ensemble Model	44.11%	0.365	0.441	0.353

A

M

N

ASL Dataset (Training Data)

A

M

N

ASL Alphabets

Webcam

- video taken from laptop webcam
- colored

Webcam

05

Conclusion

Limitations

- must contain only the hand of the signer with an empty background
- dataset used for training was not diverse so the models are not able to accurately predict signers with different styles of representing the same symbol

Future Scope

- training the models on a diverse and large data containing different ways signers use to represent a character
- inclusion of object detection techniques to detect hands in a image
- adding more character like numbers and other commonly used characters.

Project Code

https://github.com/aniketsharma00411/sign-language-to-text-translator

Thank You!