DEPT OF COMPUTER SCIENCE AND ENGINEERING

RAJIV GANDHI INSTITUTE OF TECHNOLOGY,

Chola Nagar, Bengaluru, Karnataka - 560032

SUBJECT CODE: 17CSP85
PROJECT WORK - II

Sign Language Interpreter using Deep Learning

Guide : Asst. Prof. Pushplata Dubey (Dept of CSE)

- Bahaduri Prachiti Jagdish - 1RG17CS009
- Mohammed Azim - 1RG17CS030
- Mohammed Touseef - 1RG17CS031
- Prajwal B Mani - 1RG17CS037

CONTENTS

- ❑ ABSTRACT
- ❑ INTRODUCTION
- ❑ LITERATURE SURVEY
- ❑ PROBLEM DEFINITION
- ❑ PROPOSED WORK
- ❑ METHODOLOGY & IMPLEMENTATION
- ❑ RESULTS & DISCUSSIONS
- ❑ CONCLUSION
- ❑ REFERENCES

RAJIV GANDHI INSTITUTE OF TECHNOLOGY

ABSTRACT

Speech Impairment is a disability, which affects an individual's ability to communicate using speech and hearing. This brings about the difficulty for both the sign and non - sign language speakers to communicate with each other. With recent advances in deep learning and computer vision, the focus of our project is to create a vision of an end to end Convolutional Neural Network that will be trained on the ASL(American Sign Language) dataset then modeled on robust architectures like GoogLeNet/MobileNet architecture and deploy it on an android application so that it will have more accessibility and provides an ease of use, thus aiding communication between signers and non-signers. It is a challenging and interesting problem that if solved will bring a leap in social and technological aspects alike.

RAJIV GANDHI INSTITUTE OF TECHNOLOGY

INTRODUCTION

• A neural network is a series of algorithms that endeavors to recognize underlying relationships in a set of data.

• A neural net consists of thousands or even millions of simple processing nodes that are densely interconnected.

• <u>COMPONENTS OF NEURAL NETWORK:</u>
  • Input Layer
  • Hidden Layer
  • Output Layer

•<u>APPLICATIONS:</u>
  • Speech Recognition
  • Character Recognition
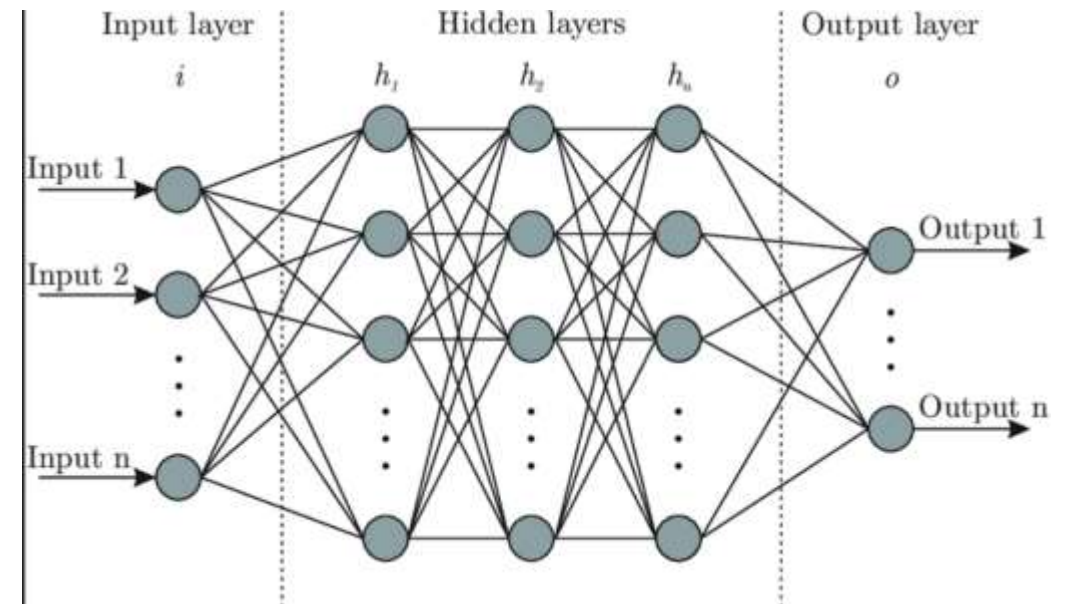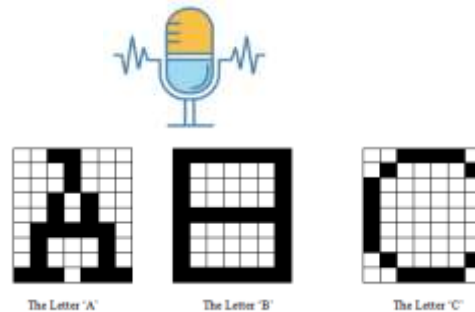
LITERATURE SURVEY

- We reviewed around 10 IEEE published paper and various online articles
- Techniques we adopt from the papers
  - Transfer learning
  - Optimizing the model
  - Image preprocessing methods
  - Model quantization

Real-time American Sign Language Recognition with Convolutional Neural Networks *Brandon Garcia & Sigberto Alarcon Viesca* *(Stanford University Stanford, CA)*	Hand Gesture Recognition Using Deep Learning *Soeb Hussain and Rupal Saxena*
Deaf Talk using 3D Animated Sign Language Mateen Ahmed, Mujtaba Idrees, Zain ul Abideen, Rafia Mumtaz, Sana Khalique	Deep Learning for American Sign Language Fingerspelling Recognition System *Huy B.D Nguyen, Hung Ngoc Do*
American Sign Language Recognition using Deep Learning and Computer Vision *Kshitij Bantupalli & Ying Xie*	American Sign Language Video Hand Gestures Recognition using Deep Neural Networks *Shivashankara S & Srinath S*
NumPy CNN Android : A Library for Straightforward Implementation of Convolutional Neural Networks for Android Devices *Ahmed Fawzy Gad*	American Sign Language Character Recognition using Convolutional Neural Network *Sarfaraz Masood, Manish Chandra Thuwal, Adhyan Srivastava*
A Deep Learning Based Video Classification System Using Multimodality Correlation Approach *Jungheon Lee, Youngsan Koh and Jihoon Yang*	Optimization of Transfer Learning for Sign Language Recognition Targeting Mobile Platform *Dhruv Rathi*

RAJIV GANDHI INSTITUTE OF TECHNOLOGY

American Sign Language Recognition using Deep Learning and Computer Vision

❏ Vision based Application

❏ Sign Language translation to text

❏ Inception Net

❏ CNN – spatial features

❏ RNN – temporal features

A Deep Learning Based Video Classification System Using Multimodality Correlation Approach

❑ Multimodality and correlation

❑ Better classification and performance

❑ Image feature vector extraction

❑ Audio feature vector extraction

❑ Normalization of feature vectors

RAJIV GANDHI INSTITUTE OF TECHNOLOGY

Optimization of Transfer Learning for Sign Language Recognition Targeting Mobile Platform

❑ Use of Transfer learning

❑ Real-Time recognizer of ASL

❑ Based on Mobile platform

❑ Monitoring continuous action movements

RAJIV GANDHI INSTITUTE OF TECHNOLOGY

PROBLEM DEFINITION

❑ Communication is one of the basic requirements for survival in society.

❑ Some of the existing methods

- Traditional based approach

- Glove based approach

❑ Lack of availability and expensive

PROPOSED WORK

We have introduced a refined approach using advanced deep learning techniques that is flexible to use by any android OS. In this presentation, we propose an architecture to detect signs. Image pre-processing makes the existing data and input data normalized. The pipeline allows us to automate machine learning workflow. Transfer learnings make it easy to make the models to learn even small details that are hard to capture in a small network. Android Studio gives us an advantages overview by giving us a docile structure to compile to any version of the android OS.

METHODOLOGY

**Development Life Cycle of our project**
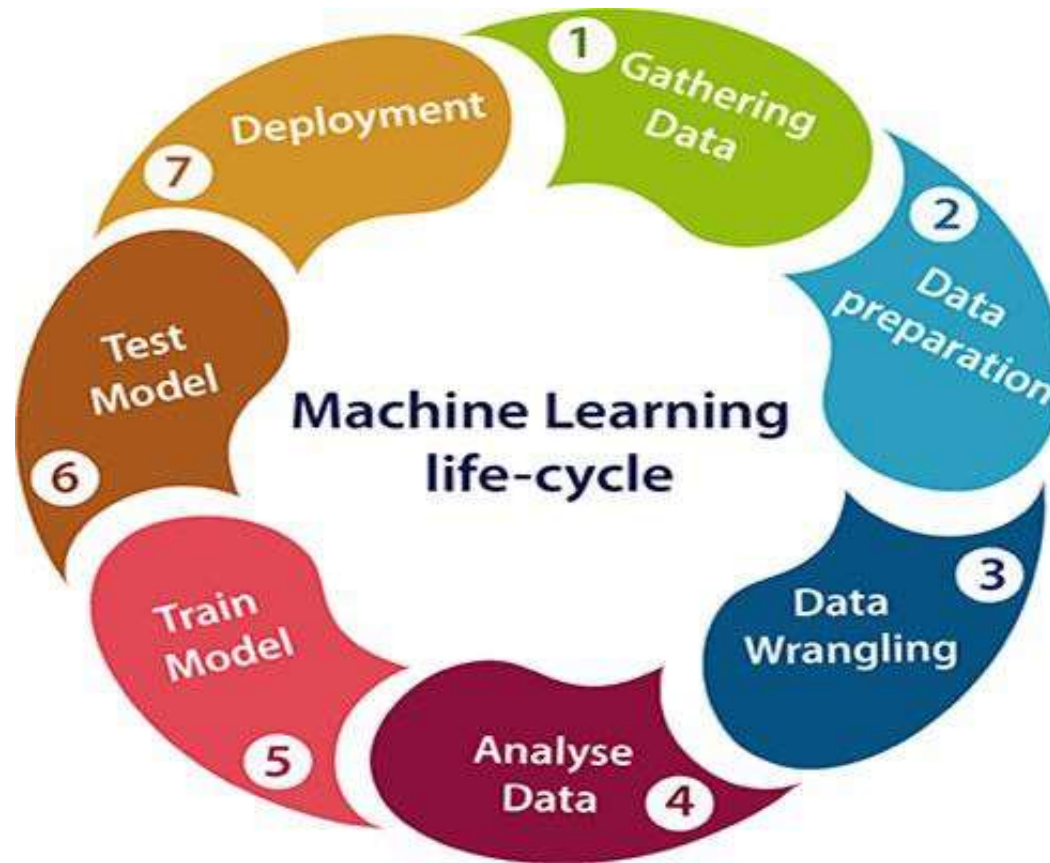
IMPLEMENTATION

- Image preprocessing : OpenCV and keras imagedatagenerator
- CNN model: VGG16 (ImageNet)
  - **<u>Image Classification</u>**
    - VGG16 (Online Dataset)
    - VGG16 (Our own custom Dataset)
    - VGG19
    - ResNet
  - **<u>Object Detection</u>**
    - SSD MobileNet FPN 320X320
    - SSD MobileNet FPN 640X640
    - ResNet FPN 1024X1024
- Language correction model
- Audio Processing
- Text to speech and speech to text
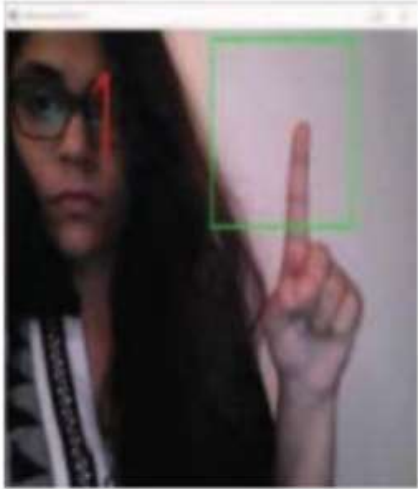- Giphy Api
- Deployment of the model

RESULTS & DISCUSSIONS

BASE PAPER ENHANCEMENTS
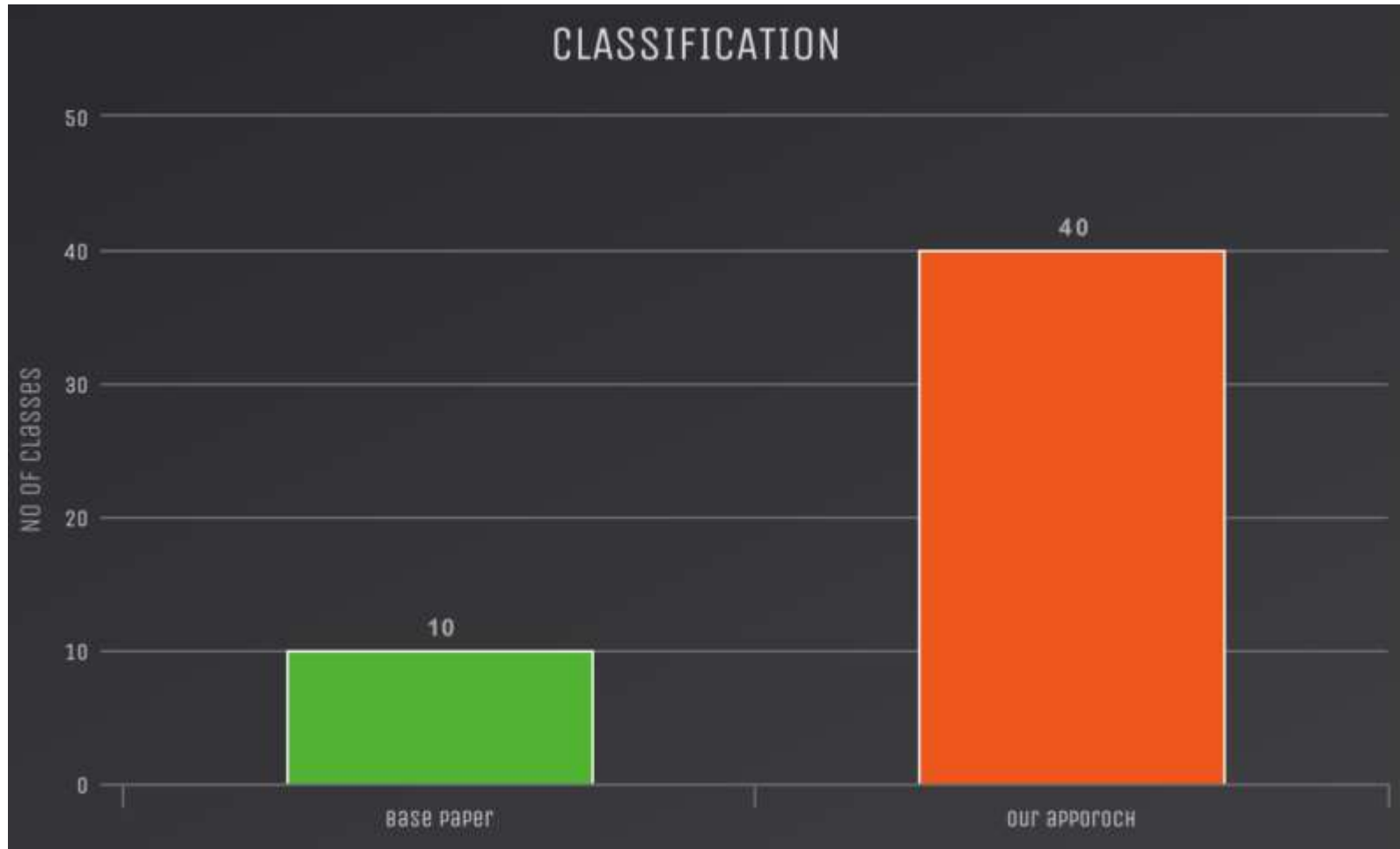
BASE PAPER IMAGES

(a)

(b)

CUSTOM IMAGES

And we deployed it on a android phone with less resources and without any cloud nor a server.

DEMO OF THE PROJECT

CONCLUSION

A sign language is a natural mode of communication used by the deaf community. India has large population of speech and hearing impaired but a very small number of certified sign language interpreters are available. Research in hand gesture recognition has gained attention with advancement in the field of computer vision. A Sign Language Interpreter (SLI) decodes and understands the information conveyed by signs. SLI can be a major breakthrough in helping a common people to communicate with the deaf and can help in bridging this communication gap. A SLI can be designed based on video/image processing and deep learning techniques which requires a standard dataset, determination of an optimal feature set, and an appropriate classification technique.

RAJIV GANDHI INSTITUTE OF TECHNOLOGY

REFERENCES

❑ **Reference Websites:**

[1] https://www.tensorflow.org/

[2] https://kivy.org/#home

[3] https://www.deeplearning.ai/

[4] https://ieeexplore.ieee.org/Xplore/home.jsp

[5] https://github.com/kivy/buildozer

❑ **Reference Text books:**

[1] "Machine Learning Yearning", by Andrew Ng

[2] "Machine Learning Techniques ", by Tom M Mitchell

[3] "AI and Machine Learning for Coders: A Programmer's Guide to Artificial Intelligence" by Laurence Moroney

[4] "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems", by Aurelien Geron

RAJIV GANDHI INSTITUTE OF TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SUBJECT CODE: 17CSP85
PROJECT WORK - II

Sign Language Interpreter using Deep Learning

THANK YOU