Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

# IDENTIFICATION OF INDIAN SIGN LANGUAGE

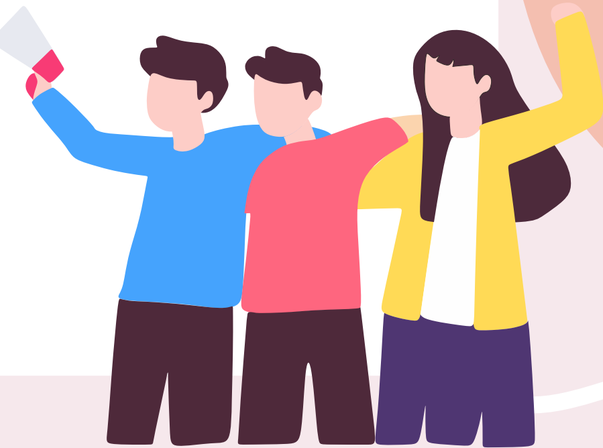
**Group 9**

Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

# Academic Year 2022-23

**Group Members**:
ODRIN RODRIGUES (201903039)
SHIVAM YADAV (201903054)
KAUSTUBH DESLE (20180307)
ANUSHKA SHINDE (2020032006)

**(Fourth Year Engineering)**



Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

# Contents:

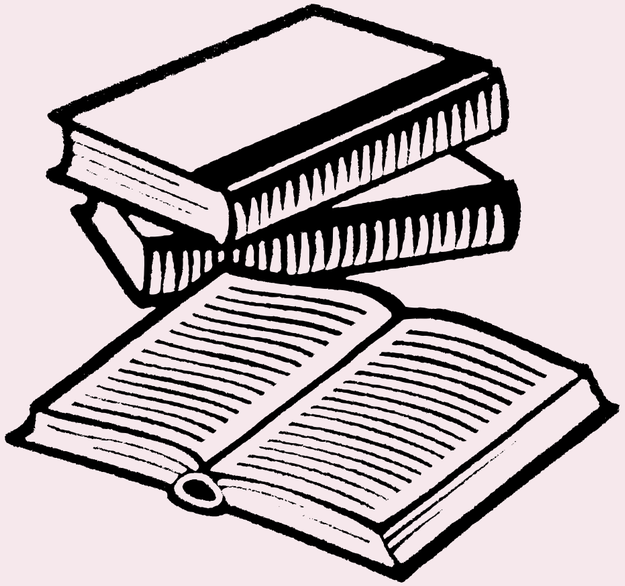
Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

# Introduction

- Deaf and hard-of-hearing persons use sign language, a visual-gestural language, to communicate.
- Hand gestures and other body movements are used to convey meaning. The syntax and vocabulary of this language are wholly different from those of spoken and written languages.
- Similar to how spoken language employs rules to generate comprehensive messages, sign language is structured by complex grammar.
- An AI converting sign language into text or speech providing environment for communication between normal and deaf community.

# LITERATURE SURVEY

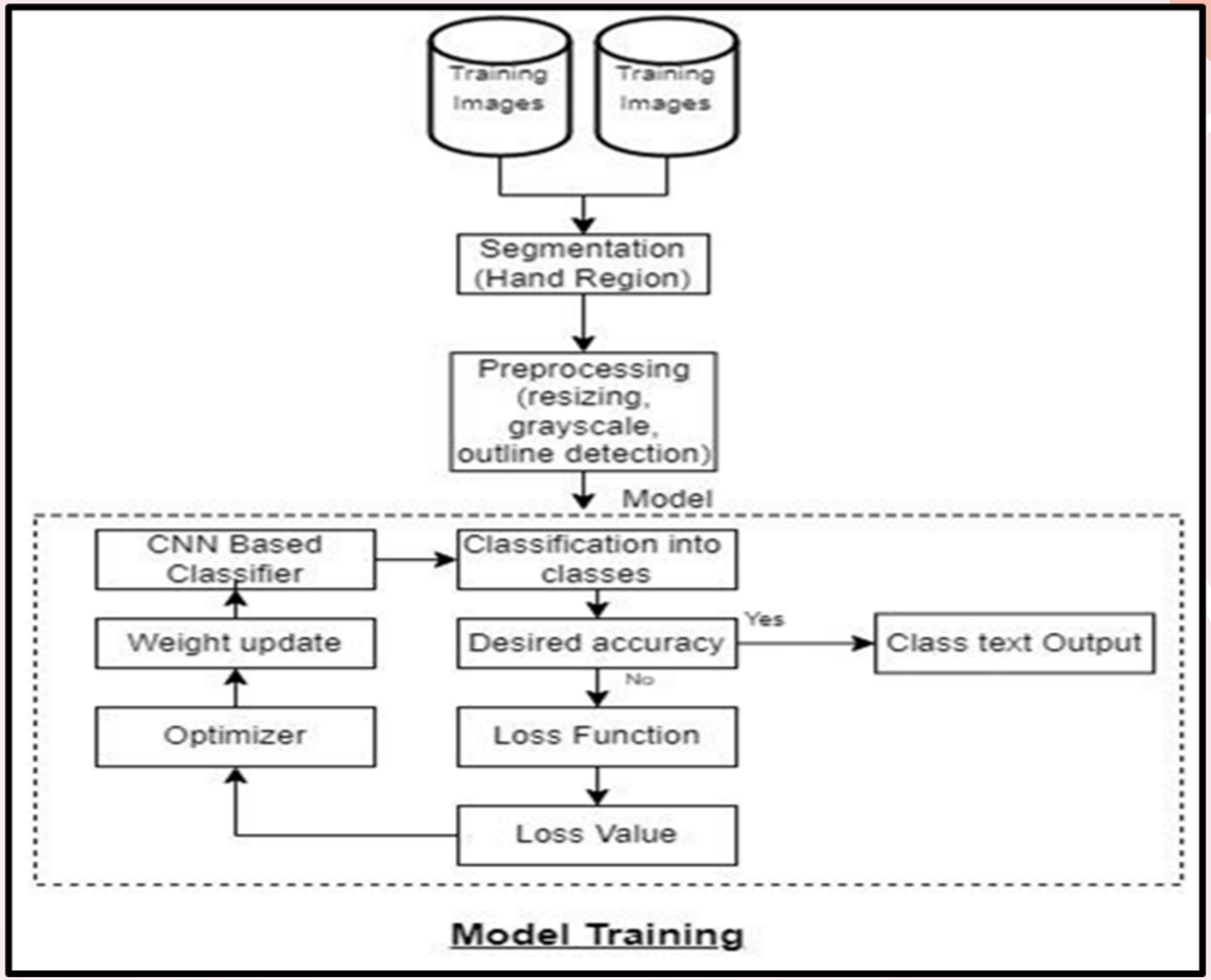
| Year | Authors | Title | Algorithm Architecture | Performance Results |
|---|---|---|---|---|
| 2015 | Lionel Pigou, Sander Dieleman, Pieter-Jan Kindermans, Benjamin Schrauwen [2] | Sign Language Recognition using Convolutional Neural Networks. | CNN, ANN | The accuracy on the test set is 95.68% and observe a 4.13% false positive rate, caused by the noise movements |
| 2018 | G.Anantha Rao K.Syamala, P.V.V.Kishore, A.S.C.S.Sastry [3]. | Deep Convolutional Neural Networks for Sign Language Recognition. | CNN | Deep Convolutional Neural Networks for Sign Language Recognition. CNN Comparing the proposed CNN model to existing state-of-the-art classifiers, its average recognition rate was higher at 92.88%. |
| Apr 2018 | Jie Huang, Wengang Zhou, Qilin Zhang, Houqiang Li, Weiping Li [4]. | Video-Based Sign Language Recognition without Temporal Segmentation. | Two-stream 3D CNN | Proposed method LS-HAN achieves 82.7 % accuracy which is more than LSTM-E 76.8%. |
| June 2018 | M.A Hossen, Arun Govindaiah, Sadia Sultana, Alauddin Bhuiyan [5]. | Bengali Sign Language Recognition Using Deep Convolutional Neural Network | Pre-trained VGG16, CNN | Proposed recognition system results - validation loss of 0.3523 (categorical cross-entropy) and validation accuracy of 84.68%. |
| 2018 | Kshitij Bantupalli, Ying Xie [6] | American Sign Language Recognition using Deep Learning and Computer Vision. | CNN, RNN, Machine learning, HMM. | Instead of using Pool Layer, the system was able to attain 93% accuracy with SoftMax Layer. |



| Year | Authors | Title | Algorithm Architecture | Performance Results |
|---|---|---|---|---|
| 2020 | Dongxu Li , Chenchen Xu, Xin Yu, Kaihao Zhang, Ben Swift , Hanna Suominen, Hongdong Li [11] | TSPNet: Hierarchical Feature Learning via Temporal Semantic Pyramid for Sign Language Translation | TSPNet, NMT | Conv2d-RNN and TSPNet-Single outperform the state-of-the-art SLT model, Conv2d-RNN, by a large margin. The relative improvement in BLEU-4 score is 39.80% (9.58 → 13.41), and the improvement in ROUGE-L is 9.94% (31.80 → 34.96). |
| 2020 | Necati Cihan Camg,Oscar Koller, SimonHadfifield and Richard Bowden [12]. | Sign Language Transformers: Joint End-to-end Sign Language Recognition and Translation. | RNN-based attention architectures, Connectionist Temporal Classification (CTC). | Compared to earlier approaches, the authors Language Transformers outperform both their recognition and their translation effectiveness with a 2% reduction in word error rate. |
| 2021 | Rachana Patil, Vivek Patil, Abhishek Bahuguna, and Mr. Gaurav Datkhile [13] | Indian Sign Language Recognition using Convolutional Neural Network. | CNN | The model achieved a validation accuracy of approximately 95%. |
| 13 Apr 2022 | Razieh Rastgoo, Kourosh Kiani, Sergio Escalera [14]. | Word separation in Continuous sign language using isolated signs and post-processing | CNN, LSTM | The proposed model obtains an average of recognized Softmax outputs equal to 0.98 and 0.59. |
| 30 Apr 2022 | Abdul Mannan, Ahmed Abbasi, Abdul Rehman Javed, Anam Ahsan, Thippa Reddy Gadekallu, Qin Xin [15]. | Hypertuned Deep Convolutional Neural Network for Sign Language Recognition | CNN | The proposed Deep CNN model can recognize the ASL alphabets with an accuracy rate of 99.67% on unseen test data. |

Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

# MODEL TRAINING



Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

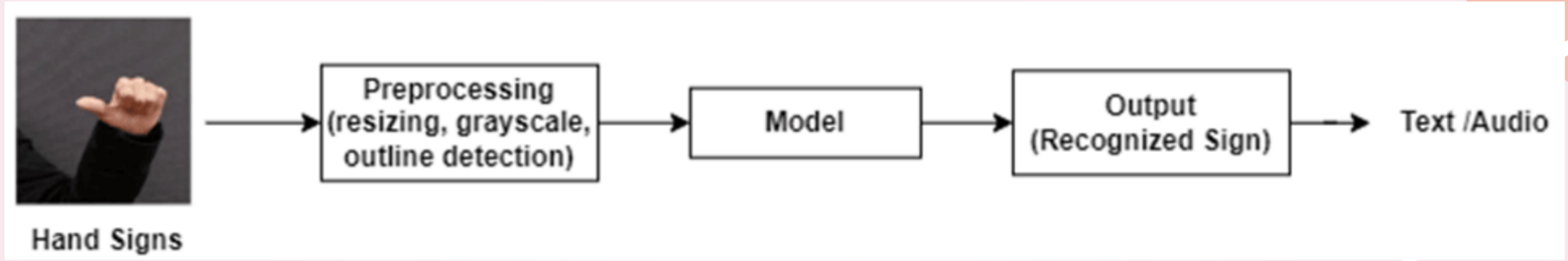



Fig. Block diagram of Hand Sign to Text/Audio conversion

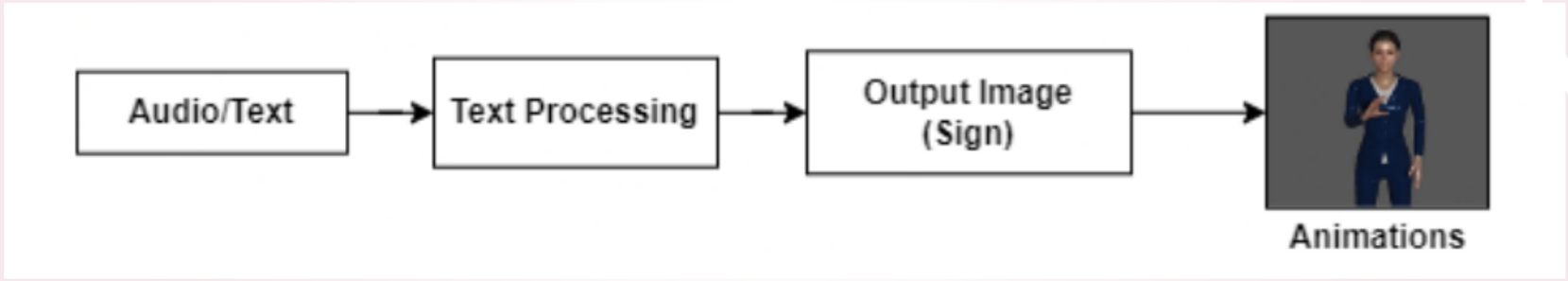



Fig. Block diagram of Audio/Text to Hand Sign conversion

Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

# REQUIREMENTS

## HARDWARE

COMPUTER/ LAPTOP

- MINIMUM 4GB RAM
- I3 PROCCESSOR 11TH GENERATION
- NVIDIA MX 330

## SOFTWARE

VS CODE

GOOGLE COLAB PLATFORM / GPU

GUI: HTML, CSS, JAVASCRIPT

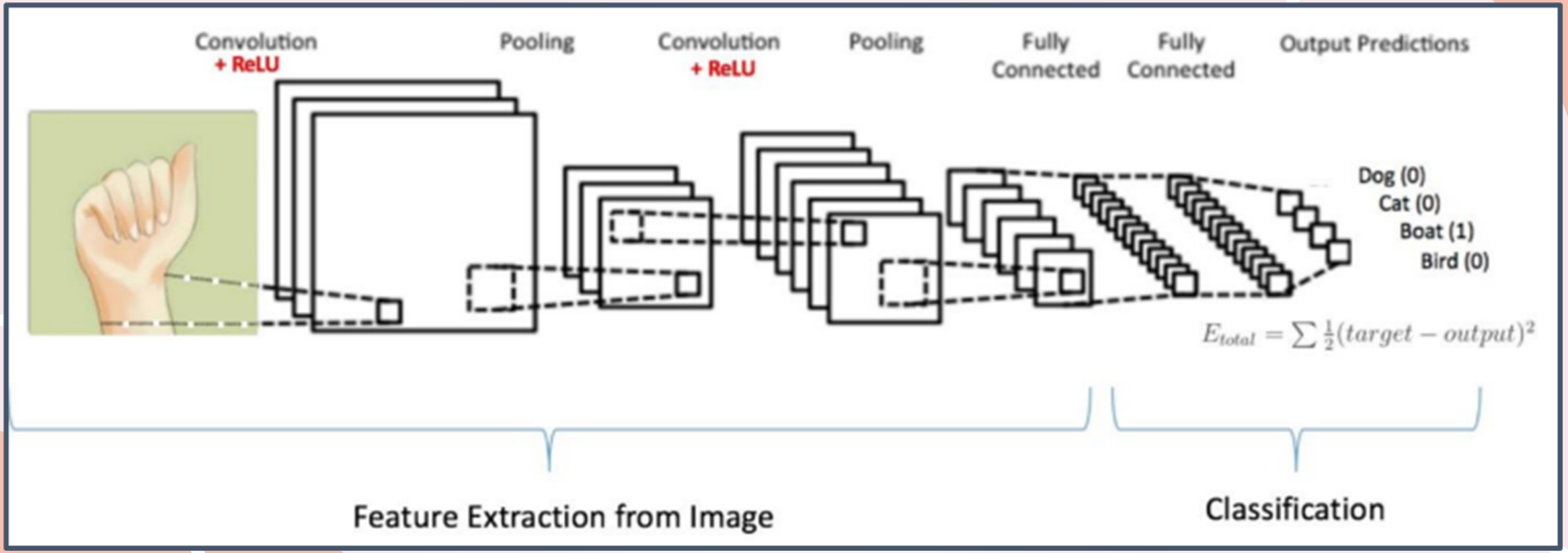
Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

# MODERN TOOLS

- Blender
- Tensorflow
- Google Mediapipe
- NLTK
- CNN
- KNN

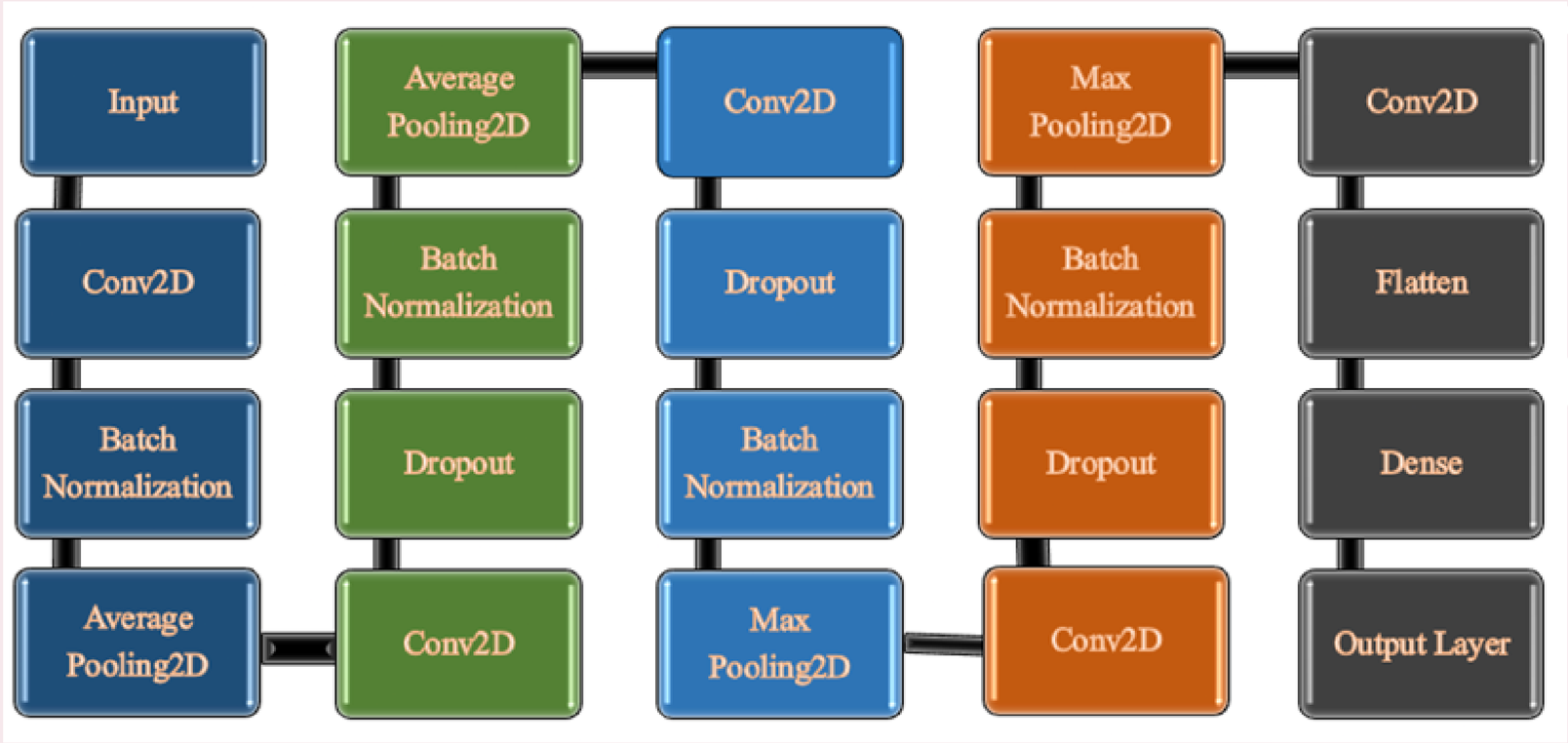
Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

# PROPOSED METHODOLOGY (ALGORITHM / FLOWCHART ETC)



Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

# PROPOSED CNN MODEL ARCHITETCURE



Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
|---|---|---|---|---|---|
| Pre-processing Approach | Gaussian Blur and Gray Scaling | Gaussian Blur, Gray scaling, Thresholding and Edge detection | Gaussian Blur, Gray scaling, Edge detection , Varying Augmentation parameters and 30% Dataset 2 mixture. | Gaussian Blur, Gray scaling, Edge detection and 70% of Dataset 1 + 70% of Dataset 2 | Mediapipe hand detection, Drawing finger joints, Gaussian Blur,Gray scaling and Edge detection |

Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

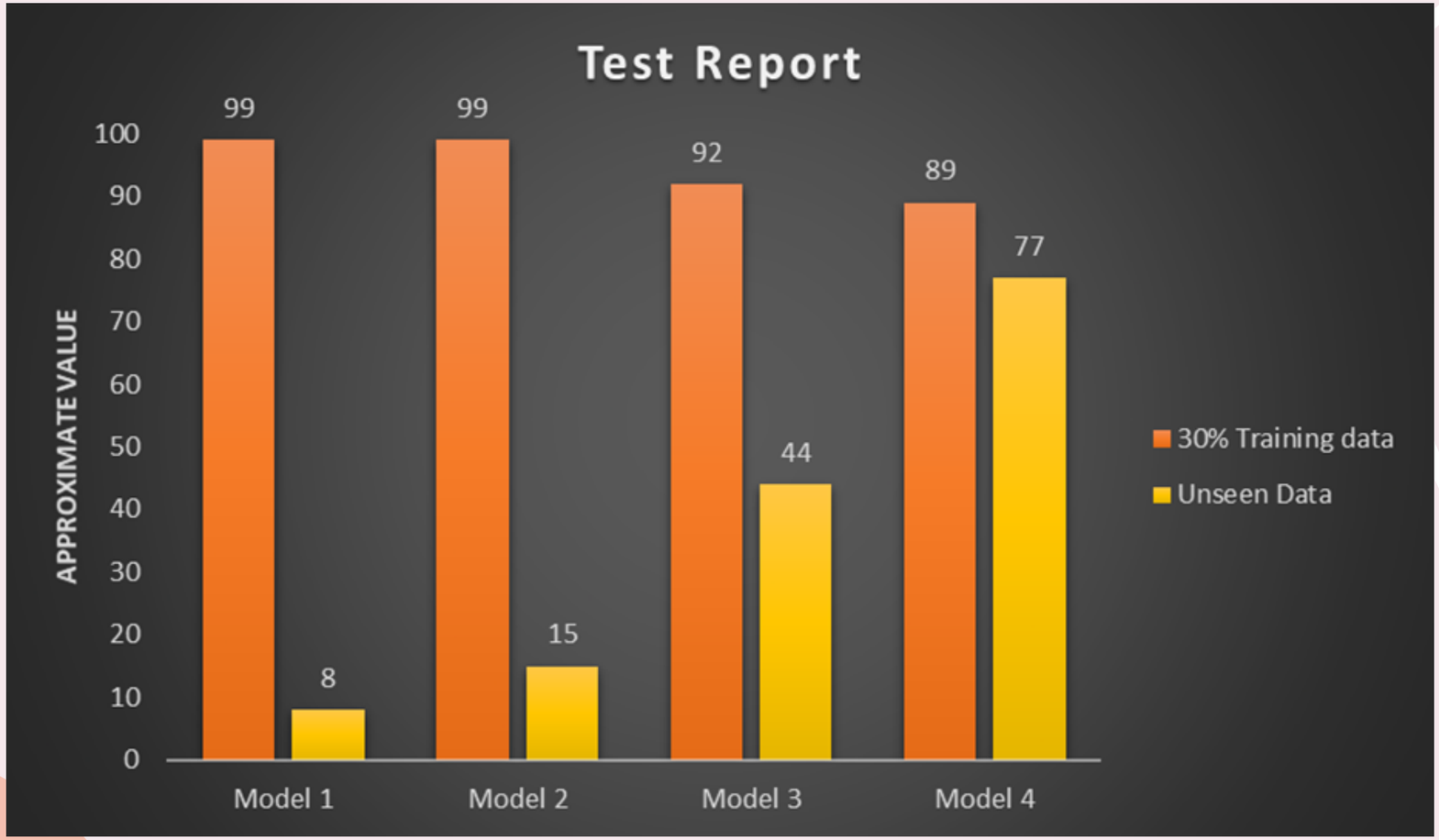
| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
|---|---|---|---|---|---|
| CNN Layers | 6 | 6 | 6 | 6 | 6 |
| Activation, Optimizer | ReLU,Adam | ReLU,Adam | ReLU,Adam | ReLU,Adam | ReLU,Adam |

Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

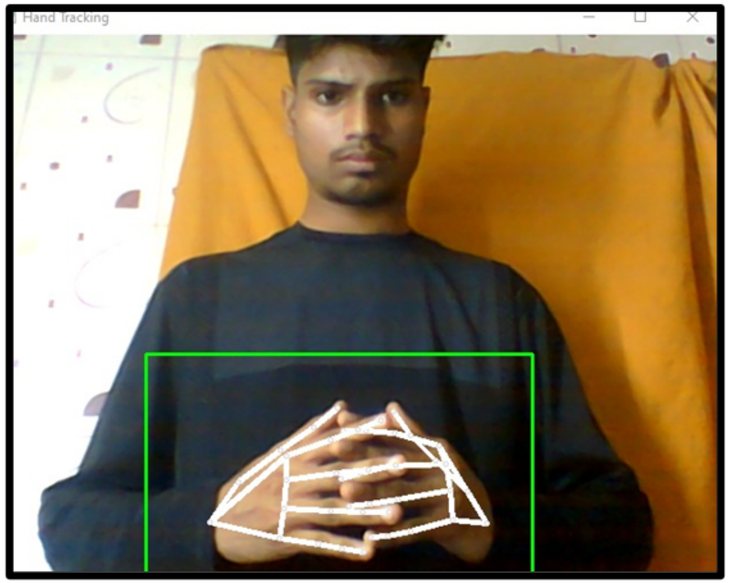
| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
|---|---|---|---|---|---|
| Result: | Accurate classification of test classes<br>Low recognition rate of 0.8% on unseen data<br>Inability to recognize in real-time video frames<br>Model needs improvement for practical use | Accurate classification of all test set classes<br>Improved recognition rate from 0.8% to 15% on unseen data<br>Indicates enhanced generalization capabilities<br>Still struggles to recognize hand signs in real-time video frames. | 30% of Dataset 2 included in training set<br>Recognition rate improved on unseen set<br>Improvement observed<br>Dataset 2 incorporation helped | Model achieved 77% accuracy on unseen test dataset<br>Poor performance in real-time classification<br>Unable to identify classes in real-time<br>Further improvements required for real-time usage. | Model achieved 99.50% accuracy.<br>Real-time sign language recognition.<br>K-Nearest Neighbor classifier used.<br>Effective communication for deaf individuals. |

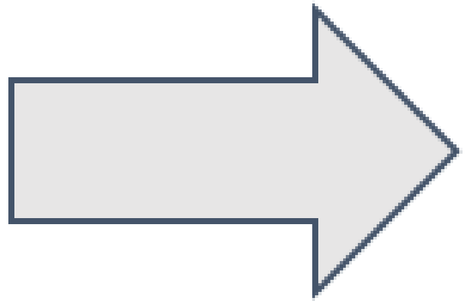
Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

# SUMMARIZED TEST RESULT OF MODELS 1–4



Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)



- Mediapipe's hand segmentation technique, which finds hands in the picture.
- Extracts the hand pose from given frame
- This extracted frame is the input to our Model

Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

Taking input from camera in grayscale

Applying preprocessing

Inverting pixel's

Transformed Image given to Model

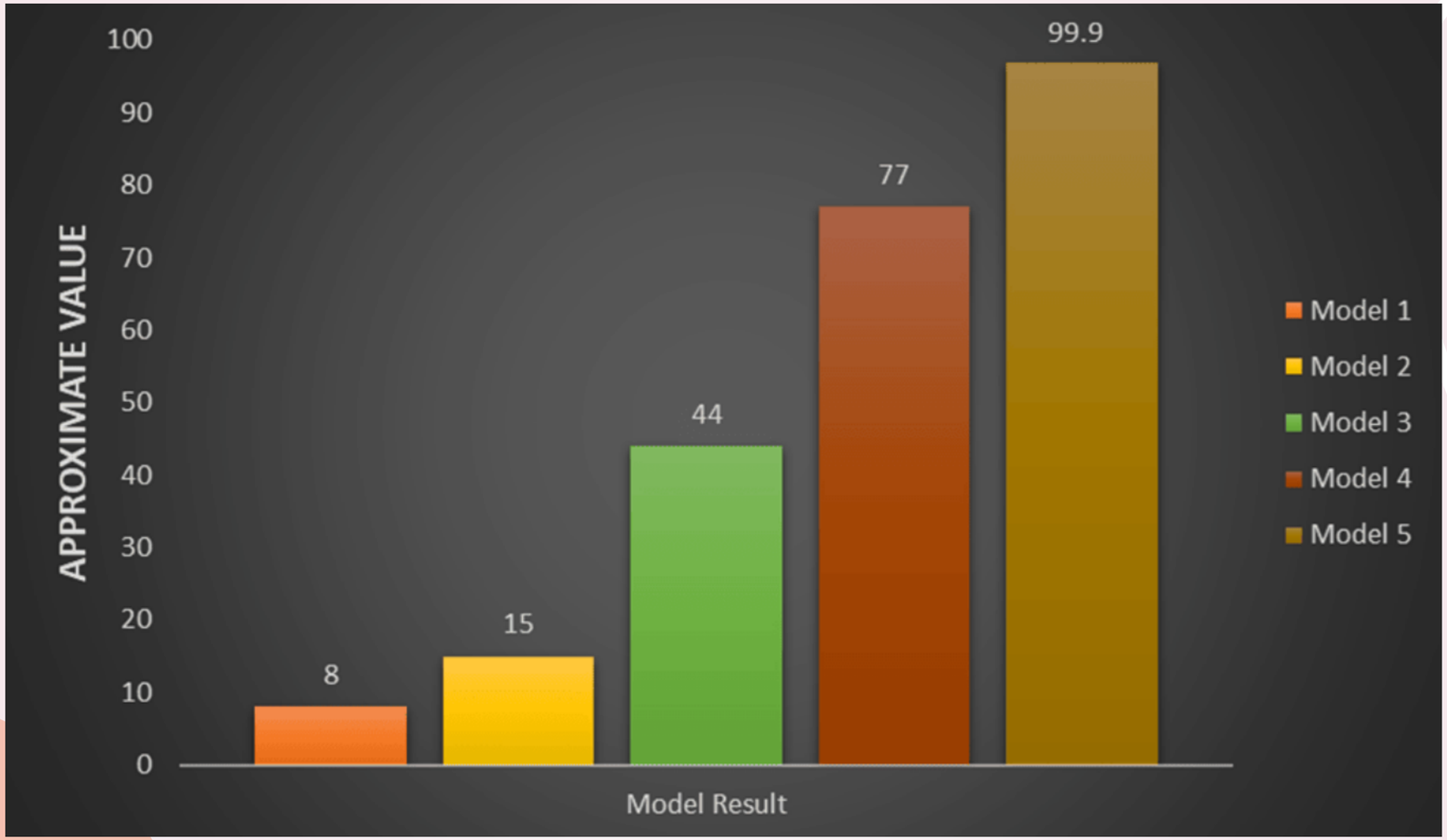
Model

Sign

Corresponding sign

Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

# SUMMARIZED TEST RESULT OF MODELS 1-5



Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

# TEXT TO SIGN LANGUAGE ANIMATION

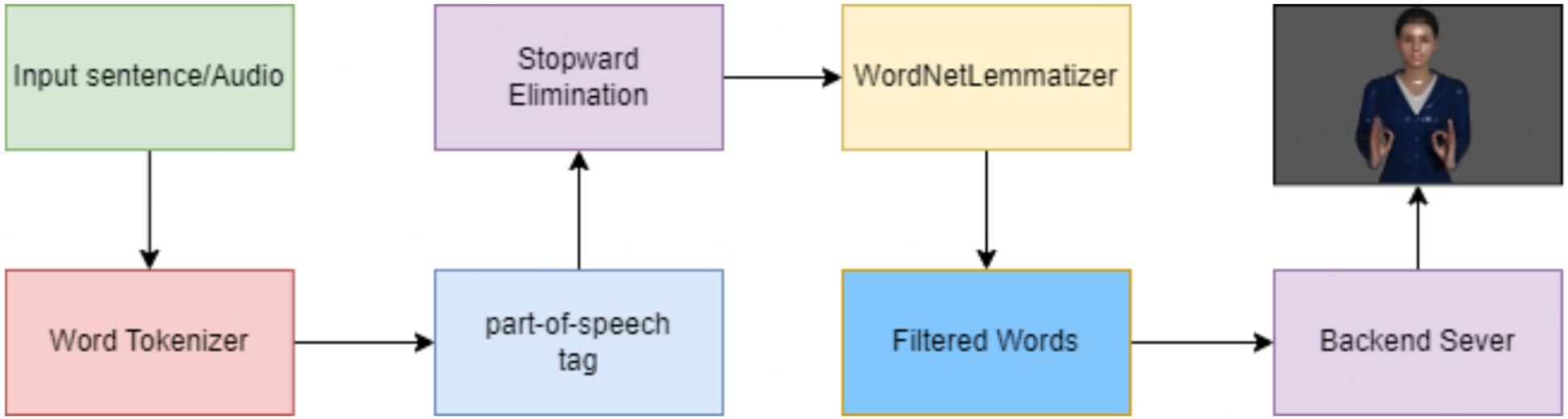



*Fig*. Text to ISL Animation

To create animations for Indian Sign Language hand motions, the team utilized Blender, a 3D modeling software.

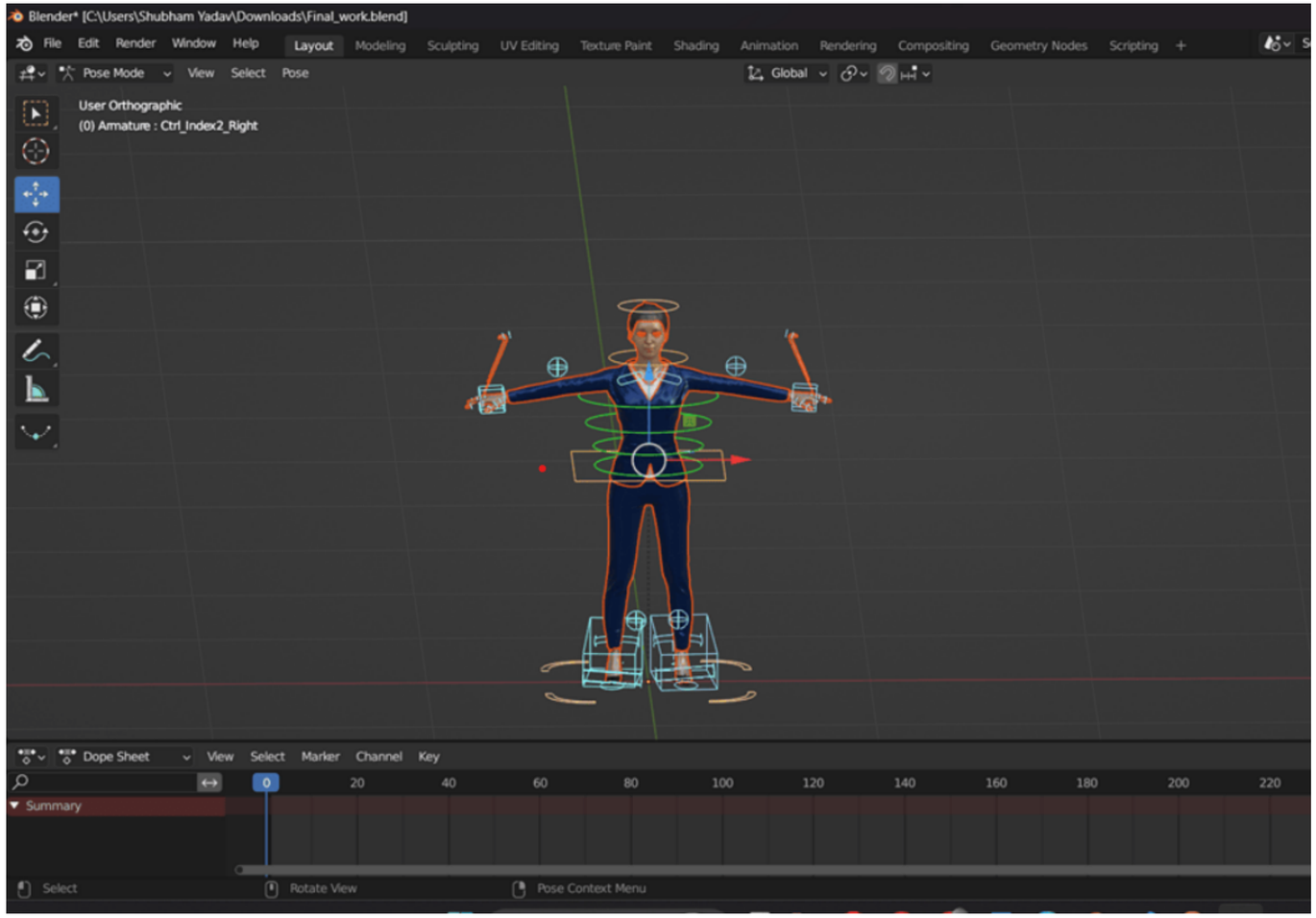



*Fig*. Blender 3D character used for animation

Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

# APPLICATION'S USER INTERFACE

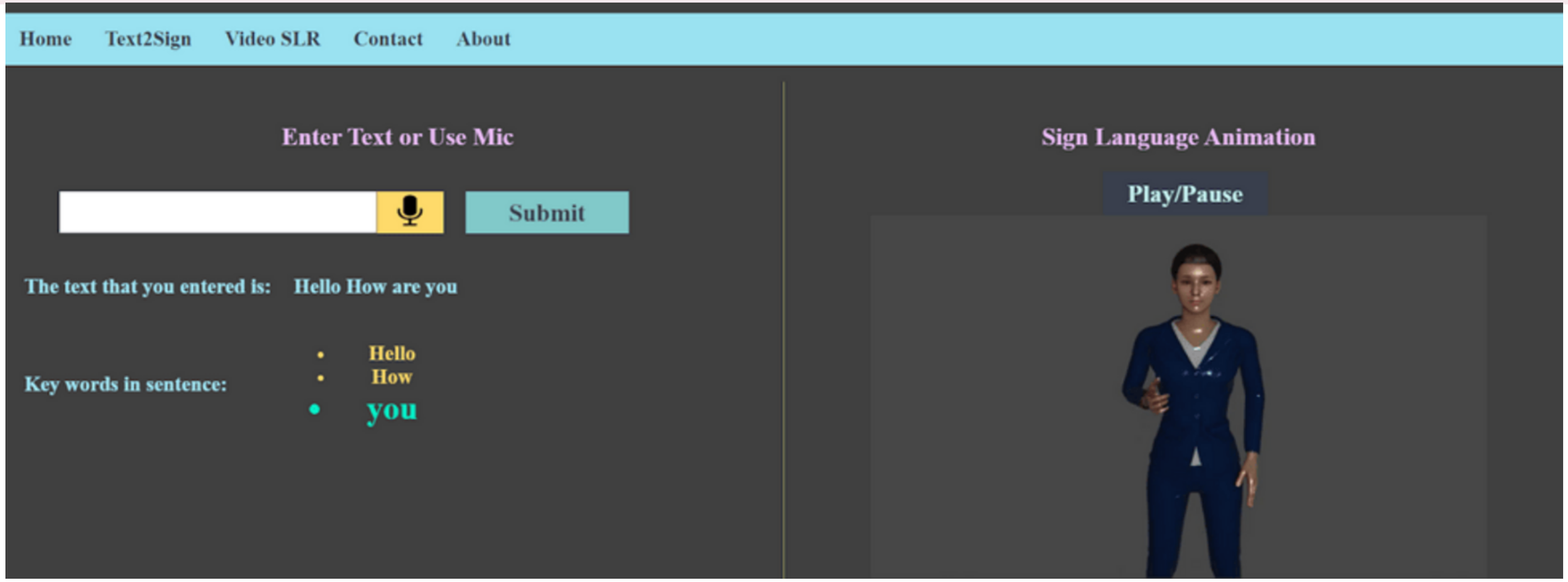



*Fig*. Animation playing on frontend

- The use of Blender and rigging systems allowed for the creation of realistic and accurate animations of Indian Sign Language hand motions, which are integrated into the application's user interface.
- This feature enables users to communicate with Sign Language in a more natural and intuitive way, without requiring prior knowledge of the language

Xavier Institute of Engineering
Mahim, Mumbai 400016
Department of Information Technology
(Affiliated to University of Mumbai)

# CONCLUSION

To build a model for real-time applications, careful attention should be given to preprocessing and feature extraction. Our experiments demonstrated that Model 5 outperformed previous models in real-time applications. It is observed that modifying the preprocessing steps improved the results on static data but resulted in poor performance in real-time applications. Thus, we recommend extracting features from video frames using model 5 approach or other techniques like optical flow and training the model using more complex neural network structures such as CNN-RNN to obtain better results.

# REFERENCES


**[1]** K.Shenoy,Tejas Dastane,Varun Rao,Devendra Vyavaharkar,“RealtimeIndianSignLanguage(ISL)Recognition” ,1 July 2021.
**[2]** Lionel Pigou, Sander Dieleman, Pieter-Jan Kindermans, Benjamin Schrauwen. ”Sign Language Recognition using Convolutional Neural Networks”, 2015.
**[3]** G.Anantha Rao, K.Syamala, P.V.V.Kishore, A.S.C.S.Sastry, “Deep Convolutional Neural Networks for Sign Language Recognition.” , 2018.
**[4]** Jie Huang, Wengang Zhou, Qilin Zhang, Houqiang Li, Weiping Li, “Video-Based Sign Language Recognition without Temporal Segmentation”, AAAI-18, April 2018.
**[5]** M.A Hossen, Arun Govindaiah, Sadia Sultana, Alauddin Bhuiyan “Bengali Sign Language Recognition UsingDeep Convolutional Neural Network” June 2018.
**[6]** Kshitij Bantupalli, Ying Xie, “American Sign language identification using Deep Learning and Computer Vision.”, IEEE,2018.
**[7]** Sruthi C. J and Lijiya A. Member, “Signet: A Deep Learning based Indian Sign Language Recognition System.”,IEEE, April 2019.
**[8]** Lean Karlo S. Tolentino, Ronnie O. Serfa Juan, August C. Thio-ac, Maria Abigail B. Pamahoy, Joni Rose R. Fortezaz and Xavier Jet O. Garcia. “Sign language identification using Deep Learning.”, IJMLC, December 2019.
**[9]** Ankita Wadhawan, Parteek Kumar, “Deep learning-based sign language recognition system for static signs” , Jan 2021.
**[10]** Prof. Radha S. Shirbhate, Mr. Vedant D. Shinde, Ms. Sanam A. Metkari, Ms. Pooja U. Borkar, Ms. Mayuri A. Khandge. ”Sign language Recognition Using Machine Learning Algorithm” March 2020.


# THANK YOU!