UML BASED DYNAMIC MODELING OF SIGN LANGUAGEDETECTION USING YOLOV5 FOR IDENTIFYING CONTINENTS

Mrs. Shailaja N Uke[1], Dr. Amol V Zade[2] , Dr. Nilesh J Uke[3]

1 Research Scholar, Vishwakarma Institute of Technology, Pune, Email: shailajauke274@gmail.com

2 Assistant Professor, G. H. Raisoni University, Amravati, Email: amol.zade@ghru.edu.in

3 Professor, Trinity Academy of Engineering, Pune Email: nilesh.uke@gmail.com

ABSTRACT–

Previously when Artificial Intelligence (AI) and Machine Learning (ML) was not so developed, it was very difficult to understand the sign language as it changes from person to person and country to country for depicting a particular word. Also, it was complicated for those having difficulties in speech and hearing to convey and understand the message to/from other. But nowadays, it has become very easy with the help of AI and ML by image processing and by creating the dataset of images and neural networks with hand contour. But still there are lots of problem with present system because we need faster and accurate recognition of the gestures to communicate in efficient way. Hence, we tried to model and create such system which has faster and accurate hand gesture processing with dataset of continents in ASL (American Sign Language) using Yolov5 and Makesense.ai. Hence, it will prove the quality of communication for betterment of current system.

Keywords- (Artificial Intelligence, Machine Learning, Hand Gesture, Yolov5, Makesense.AI, Unified Modeling Language)

1. INTRODUCTION

Over 5% of the world's population has hearing and speech problems. Hence, the only way to communicate for them is to communicate by gestures mainly hand gestures. But those who are facing the problems, people related to them only know the language of hand gesture or we called it as Sign Language. But, now as the Artificial Intelligence and Machine Learning has developed It has become possible to solve this problem by human-computer-human interaction which we called it as Hand Gesture Recognition in which we create a dataset for each and every hand gesture depicting a particular meaning to it and then we simply trained the system. After that when we put our camera towards person communicating with hand gesture it shows the meaning for it in real time.

Using above mentioned process we have also tried to create such type of system using YoloV5 in Python with better accuracy and precision and tried to make this whole process very effective.
It uses the dataset of depicting continents in American Sign Language (ASL). So, when we put our camera towards that person it will detect the message he is trying to convey at very faster rate with same accuracy and precision. This makes communication very effective.

We can also create our own gestures communication as many of the people do not follow particular language as they some sort variations according to the region they belong. So, from above mentioned process we can also create our own hand gestures our own language to communicate. Also, in future this system will able

eISSN 1303-5150

to customize your own hand gestures according to user need at a run time. Also, we can use this to give command to voice assistant such that when hand gesture are recognized it will go as command to voice assistants and then voice assistant will follow command or give output int the text form.

2. LITERATURE SURVEY

[1] In this paper authors reviewed the research on sign language which was carried from 2014 to 2020 in the vision-based hand gesture recognition system They reviewed more than 98 articles and said that the research in this field has lot to work on as they found that sign recognition by plotting points has average accuracy of 88.0% and by training the images found the accuracy of 78.05 and said there is lack of progress and one should focus in vision-based recognition.

[2] in this paper they have detected alphabets and numbers using hand gestures. They have used their bench mark with 36 distinct classes of dataset to train and evaluate it on YoloV5. As they have achieved great results with an average accuracy of more than 95%. Also, it is very accurate and fast as well as fps speed for YoloV5 is also quite high.

[3] In this paper they have discussed about the Yolo algorithm for image processing and image detection. Also, it talks about the algorithm and networks required for each version of the Yolo as the image processing and detection are done in single run only. After comparing the performance of YOLOv5, YOLOv3 and YOLOv4, they found that best suitable version for image processing is YOLOv5 as it has more accuracy and precision but training time required for YOLOv3 is faster than YOLOv4 and YOLOv5.

[4] this paper is focused on neutral environment as most of the times due to some background effects and low light sometimes it becomes difficult to recognize the hand gestures.
Hence, they have proposed a system where they are using camera with Kinect sensor to recognize the hand gestures without any environment restriction. Also, they have focused on reduced the hardware system and cost for the setup for image processing.

3. METHODOLOGY/EXPERIMENTAL

A. Problem Statement

To develop faster and effective hand gesture recognition system with higher accuracy and which can able to detect the hand gesture in a faster way with same accuracy.

B. Proposed Methodology

1. Collecting pictures and images for creating a dataset which we have to store in folder which consist of images and their labeled images.
2. For labelling the classes we have used Maksense.AI website which provides an interface where we can make classes and label our images.
3. After labelling our images we extract our labeled images in zip file which we further extract, it consists of the text file in which dimensions of label are defined. Which we further move to our label images folder for training it on the YoloV5 colab code.
4. After that we have to create a python file in which we defined our classes and file directory of our folder consist of images and label images. So, when we train our dataset, it gets the location of the folder.
5. Finally, we go on to YoloV5 colab code to train our data where we want to upload our image folder i.e., train_data and the python file i.e., custom_data and follow the installation process and training the data
6. After training we get the results as the images of our label which we created and the predicted images after training.
7. Finally, when we put our camera towards the person speaking sign language it will detect the gesture.

4. System architecture

- Program Structure

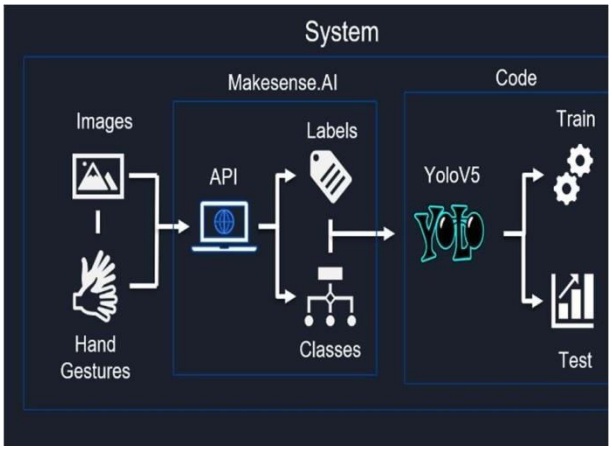

Fig.1 System Architecture

- Program Flow

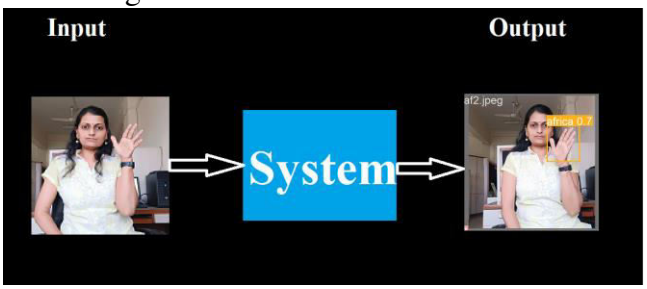

Fig 2 Block Diagram

5. SYSTEM DESIGN

It is the procedure of defining several segments of the system like the its architecture, different modules and components and its interfaces and the data which is exchanged or moved in system. Basically, its aim is to fulfill needs and wants of a user, business or organization through the engineering a efficient and smooth running system. This Unified Modeling approach is used by many researchers to model system from different views.[12]

USE CASE DIAGRAM

A use case diagram is used to capture a dynamic behavior of system. It basically models system functionality which comprises of different actors and set of use cases. Use cases refers to set of task, steps or functions that system is required to perform [11].

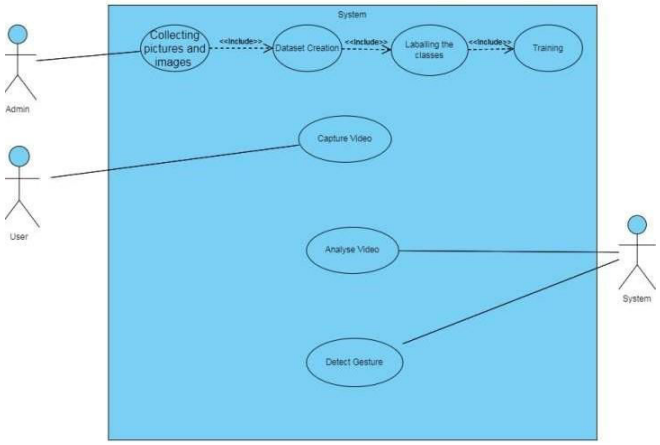

Fig 3 Use Case Diagram

ACTIVITY DIAGRAM

Dynamic aspect of system is captured by one of UML diagram which is nothing but Activity diagram. It is used to visualize the control flow in the system. It is used to model the concurrent and sequential activities. The control flow in system can be sequential, branched, or concurrent. It is basically kind of flowchart.

eISSN 1303-5150

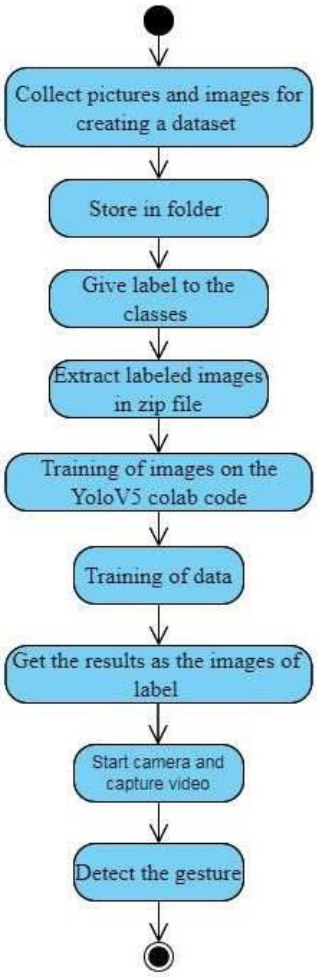

Fig 4 Activity Diagram

6. RESULTS AND DISCUSSIONS

As mentioned above we get all the dataset with precision more than 95% and recall equal to 100% for each class which more than enough to get faster and accurate hand gesture recognition.

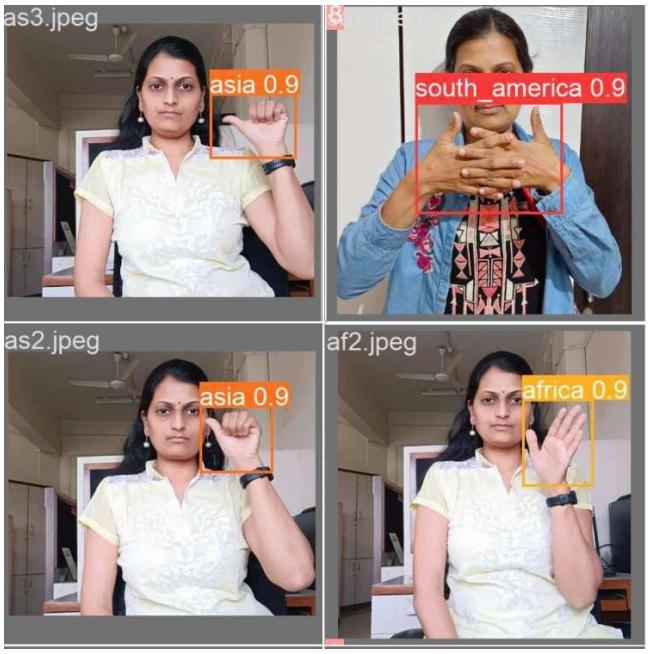

Fig 5 ASIA and AFRICA detected with precision 97% and 96% respectively

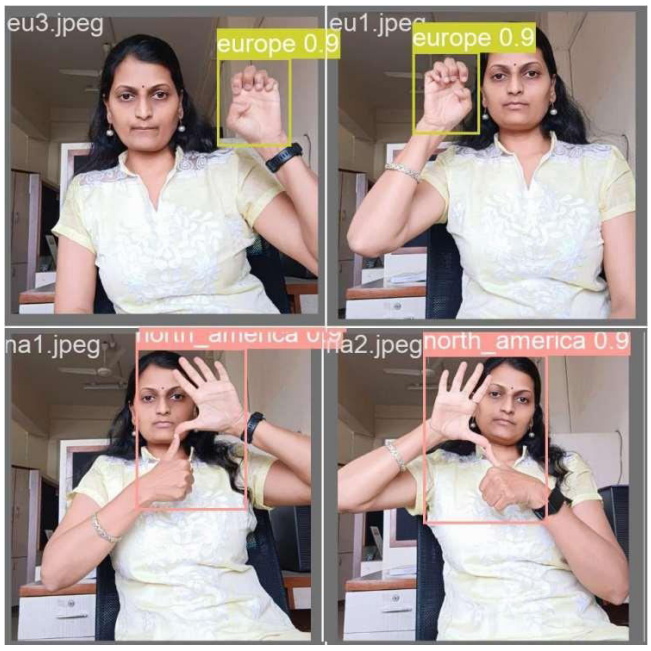

Fig 6 NORTH AMERICA and EUROPE detected with precision 95% and 96% respectively.

Fig 7 AUSTRALIA and SOUTH AMERICA detected with 96% and 98% respectively.

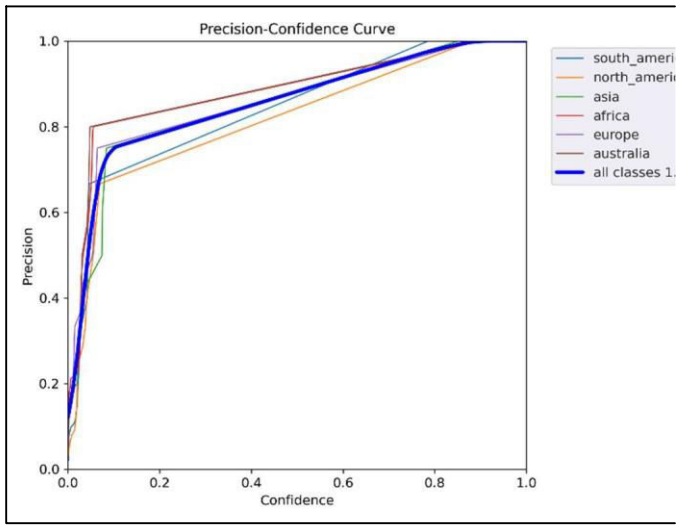

Fig 8 Precision Curve for trained dataset

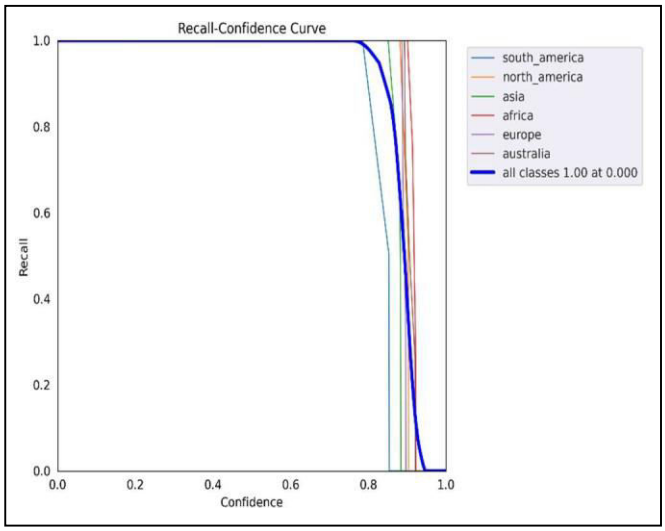

Fig 9 Recall Curve for trained dataset.

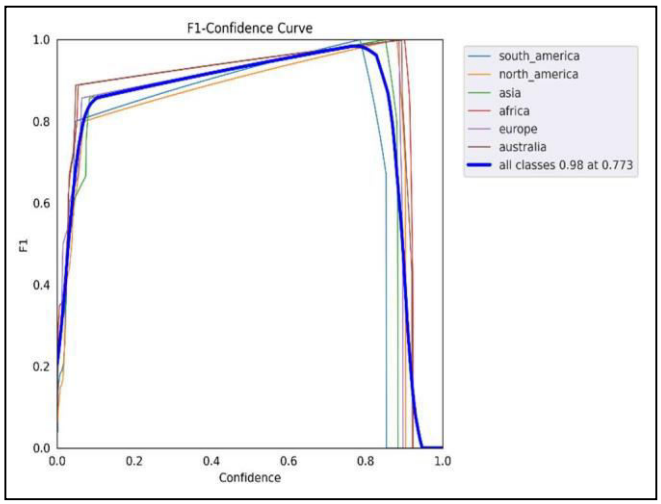

Fig 10 F1-Curve For trained dataset

CONCLUSION

After analyzing lots of research papers, we get to know different problems which are present in the current systems. After working on it we get the successful results as we are more focused on rate, accuracy and precision. As mentioned in the results we get image processing and detection at very faster rate. Due to which we have increase in the number of epochs and also we got an accuracy of more than 95% in average which greater than the present systems. Also, we observed that duringtraining Precision, Accuracy and F1-Curve have increased which is according to increase innumber of epochs

eISSN 1303-5150

and data loss was decreasing. So, these systems can be used for in real time. It can be used to interact between a person who is communicating via sign language and other recognizing it using our system.

FUTURE SCOPE

As Artificial Intelligence and Machine learning are increasing at CAGR of 40% day by day. So we can expect this world will move towards the automation. But still there are lots of problems in present system in hand gesture recognition such as user should able to customize his own sign language during the run time of algorithm which will be very beneficial as people do not have to rely on a particular language to convey their message, they can update their own hand gestures and assign them to depict particular word. Also, we can integrate these systems with chat bots and Voice assistant so that they can also have a benefit of using them by only communicating in Sign Language by converting text to voice input to give text output. Also, they are still questions on accuracy and rate for detection we are focusing on these problems in future.

REFERENCES

[1] N. Mohamed, M. B. Mustafa and N. Jomhari, "A Review of the Hand Gesture Recognition System: Current Progress and Future Directions," in *IEEE Access*, vol. 9, pp. 157422-157436, 2021, doi: 10.1109/ACCESS.2021.3129650.

[2] T. F. Dima and M. E. Ahmed, "Using YOLOv5 Algorithm to Detect and Recognize American Sign Language," *2021 International Conference on Information Technology (ICIT)*, 2021, pp. 603-607, doi: 10.1109/ICIT52682.2021.9491672. [3] Viraj Shinde, Tushar Bacchav, Jitendra Pawar, Mangesh Sanap "Hand Gesture Recognition System Using Camera" , International Journal of Engineering Research & Technology (IJERT) Vol. 3 Issue 1, January - 2014

[4] Sahla Muhammed Ali, "Comparative Analysis of YOLOv3, YOLOv4 and YOLOv5 for Sign Language Detection," International Journal Of Advance Research And Innovative Ideas In Education, vol. 7, no. 4, pp. 2393-2398, Jul-Aug 2021

[5]P. K. Pisharady and M. Saerbeck, ''Recent methods and databases in vision-based hand gesture recognition: A review,'' Comput. Vis. Image Understand., vol. 141, pp. 152–165, Dec. 2015, doi: 10.1016/j.cviu.2015.08.004.

[6] M. Yasen and S. Jusoh, ''A systematic review on hand gesture recognition techniques, challenges and applications,'' PeerJ Comput. Sci., vol. 5, p. e218, Sep. 2019.

[7] M. J. Cheok, Z. Omar, and M. H. Jaward, ''A review of hand gesture and sign language recognition techniques,'' Int. J. Mach. Learn. Cybern., vol. 10, no. 1, pp. 131–153, Jan. 2017, doi: 10.1007/s13042-017-0705-5.

[8] S. Kausar and M. Y. Javed, ''A survey on sign language recognition,'' in Proc. Frontiers Inf. Technol., 2011, pp. 95–98.

[9] H. Cooper, B. Holt, and R. Bowden, ''Sign language recognition,'' in Visual Analysis of Humans. London, U.K.: Springer, 2011, pp. 539–562.

[10] G. Fang, W. Gao, and D. Zhao, ''Large vocabulary sign language recognition based on fuzzy decision trees,'' IEEE Trans. Syst., Man, Cybern. A, Syst. Humans, vol. 34, no. 3, pp. 305–314, May 2004.

[11] Shailaja Uke, Ravindra Thool, "UML Based Modeling for Data Aggregation in Secured Wireless Sensor Network", Procedia Computer Science, Volume 78, 2016, Pages 706-713, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2016.02.120.

[12] Uke Nilesh & Thool Ravindra, "Objects tracking in video: A object-oriented approach using Unified Modeling Language", International Journal of Computational Vision and Robotics. 5. 202. 10.1504/IJCVR.2015.068800, Jan 2015.