Deep Learning based model for Sign Language Interpretation

Devaguptam Sreegeethi - 191IT110
Information Technology
National Institute of Technology Karnataka
Surathkal, India 575025
Email: sreegeethidev@gmail.com

Kogatam Thanmai - 191IT127
Information Technology
National Institute of Technology Karnataka
Surathkal, India 575025
Email: kthanmai.4@gmail.com

Nishka Kotian - 191IT136
Information Technology
National Institute of Technology Karnataka
Surathkal, India 575025
Email: nishka.191it136@nitk.edu.in

***Abstract*—With the huge leaps and bounds in technical advancements, human beings are able to overcome any form of disability. Hand gesture recognition is one of the most developing technical fields now. It has many applications and sign language interpretation is the most important among them. In sign interpretation, the main expressions considered are finger's and hand's orientations. This project provides a collaborative tool to learn the Sign Language to interact with the deemed community. Sign language interpreter acts as a communication tool among the communities. Recent developments have helped in interpreting without the physical presence of a in-person interpreter. It is used to learn the sign language by the deaf community to communicate with others. It is difficult for common people to understand their gestures so these interpreters will help us in understanding their intentions, ideas, thoughts and beliefs. In this project, an android application is developed, through which users can learn signs of different words and check their competence. It's also been installed with a live sign language interpreter which helps in interpreting the signs which are captured using the live camera.**

***Index Terms*—Hand gesture recognition, Sign Language Interpreter, C3D architecture, MLP, computer vision**

I. Introduction

To understand human body language, a computer must recognise the hand gestures. Hand gesture recognition has many applications in the stream of Human Computer Interaction (HCI) like hand gestures to control TV, Refrigerator, video games and virtual reality. We humans show our feelings and communication information in these subtle hand gestures. These primitives are very complex to recognise as the time dependence of these mounts makes it difficult in the Euclidean space. "American Sign Language", "Indian Sign Language" and "British Sign Language" are most frequently used sign languages [6]. The Indian Sign Language" dictionary is flourished by Indian Sign Research and Training Centre which contains around 3,000 Hindi and English words with their respective signs [8]. The Indian Sign Language dictionary was developed to remove the ambiguity in the Indian Sign Language across the nation. The dictionary is updated time to time based on the new addition of words. This will help us bring uniformity in Administration, policy implementation, teaching etc.,. , This is taxing because there is a taboo in the society against disabled persons, common people doesn't have prior knowledge and training regarding this and also cultural and language differences worsens the situation. During the course of the last 3 decades, many models have been developed to trace the hand movements with differing accuracy, speed and latency [7]. Many recognition models are present which recognises based on the hand's local configuration. These systems either take the segmented hand or pre-process it using hand color or black gloves. Nevertheless, these models work well for recognising alphabets and small sized words. Other approach of this problem is considering the global configuration and ignoring the local hand configuration. These models work for simple and accurate gestures but face difficulty in sign language interpreters.

We can classify the gestures using various artificial and machine learning algorithms and image processing [9]. We can divide these into two categories static and dynamic. However, most of the models are dependent on specific background or some devices like gloves, dresses.The proposed method will help the users in learning sign language efficiently. Here, in this project an android application is developed with vocabulary, quiz and live sign language interpreter fragments following the HCI guidelines. It's trained by EfficientNet-Lite model in the application.

Keys contribution of our work:

- An holistic android application which helps users in learning Sign Language.
- Live interpreter module which helps in live interpretation of sign using camera.
- User friendly interface and authentication, deployed in the application.
- Quiz module which helps users to assess their knowledge in sign language.

The report is divided into different sections as stated here. Section II contains description of studies related to Sign Language Interpretation and in Section III the problem statement and motivation behind selecting this topic has been discussed. Section IV comprises the methodology for implementing the android application and the model. Section V contains the Human Computer Interaction guidelines which were deployed in the implementation of the project . Section VI has the results of the project. Section VII concludes the paper with elaborate inferences.

II. Literature Survey

In [1] word and phrase level sign language videos from the KSU SSL dataset were used to train a deep learning model to recognise Saudi sign language. Their method consists of three stages - input preprocessing , feature learning and fusion and finally classification. Pretrained C3D architecture along with an optimization step is used for feature learning. MLP and autoencoder are used for feature fusion before passing the result to a classifier. [2] provides a framework for an interactive learning tool for hearing impaired children. Collaboration was considered to making the learning process effective. The 5 main components of the proposed architecture were - Pedalogy, Active space of the learners, Technology, Communication and Evaluation. The aspects of development that were taken into account are students literacy and communication ability, social skills, listening skill and concentration capacity.

III. Motivation and Problem Statement

Problem statement is to create an interactive tool to engage deaf community in learning and communication using Indian sign language. The motivation behind creating this interactive tool is to allow deaf community to have equal access to information and also provide opportunity for deaf children to educate themselves. Sign language can also be of great help to autistic children who have trouble expressing themselves. A large percentage of deaf children are born to parents with hearing ability and most of them would not have learnt sign language. So a tool that enables them to learn sign language would help them to converse with their children better.

A. Objectives

- Enable users to gain better communication skills by improving vocabulary in sign language
- Make the learning process easy along with assessment of learning
- Perform sign language interpretation in order to help deaf people to communicate with people that don't know sign language
- Create an effective UI that is simple and easy to use and is designed such that no additional training needs to be provided.

IV. Methodology

The app provides a real time sign language character recognition as well as a learning module where users are taught a set of words from the Indian Sign Language dictionary. A first time user is asked to create an account. Existing users have to login first, after which the home page is shown. As a part of continuous learning process each day the user is supposed to learn how different words are represented in sign language. This helps users to improve their vocabulary and enhance their communication skills. The students have an option to attempt a quiz where their learning is tested. A video of a signer performing a particular sign is displayed and the user has to select one of the four options and choose the correct word corresponding to it. An immediate toast message is shown which tells the user if the option they selected was correct or not. On finishing the quiz the final score of the user is displayed.

The real time sign language recognition module is developed by training deep learning models on Indian sign language dataset. Three different models that we trained are:
1) Traditional CNN model
2) MobileNetV2 model
3) EfficientNet-Lite model
Table II. contains details about the model sizes and training accuracy that was obtained. The size of traditional CNN model was high which is not suitable for running on mobile devices which have limited memory and processing power, so other models were trained. The model that was chosen to be deployed on the app is the EfficientNet-Lite model. The average inference time obtained was 81.573 ms which shows that the predictions are displayed quite spontaneously. It was calculated by loading the model and performing predictions over 100 times by showing different signs and finding the average of them. The generated model was converted to TFLite format. Camera input is captured and with the help of TFLite support library and the models generated real time classification is done.

Firebase database was added as database for the mobile application. Additionally firebase authentication is also included which is used for authentication of users using email and password. In the storage of firebase, 400 videos of sign language videos were added and stored in basic and intermediate folders according to the difficulty of the words. After adding the videos in firebase storage we the access token for the videos were obtained. The name and value were added in firebase realtime database with name as name of the sign and value as access token assigned for every video. This is used for the retrieval of the videos.

V. HCI Principles

Our mobile application follows Nielsen's Ten Heuristic guidelines. Guidelines followed are as follows :

1. Visibility of system status : Users need to be kept informed by the system about what is going on, through appropriate feedback within reasonable time. In our application, we added a bottom navigation bar which has three icons for

TABLE I
SUMMARY OF LITERATURE SURVEY

Authors	Methodology	Merits	Limitations	Additional Details
MUNEER et al.	C3D architecture with aggregate local(hand) and global(body) feature configurations	Recognition rate is high	Can identify only limited 40 signs	KSU SSL dataset
Asmaa Alsumait, Maha Fasial	Framework for collaborative learning tool	Support for collaborative learning	Validation of the system is required	-
S Mohammed Siddiq et al.	MobileNet and InceptionV3 architectures	Lightweight architecutre	Does not deal with complex background	Custom dataset
Advaith Sridhar et al.	Key point vectors, Pose and PAF video as features for encoder decoder model	Dataset videos resemble real world scenario	Output features may not be restricted to those relevant for sign recognition	INCLUDE dataset
Abhishek et al.	Hand glove with sensors	Wearable, designed such that even people with motor disability can use it	Not compatible with smartphone	-

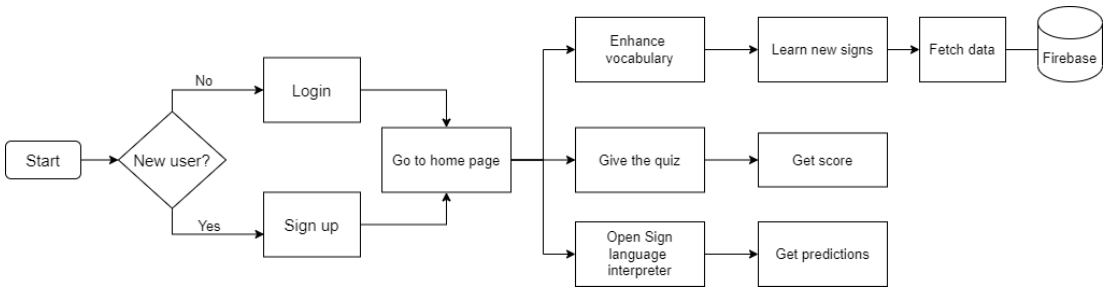

Fig. 1. Workflow diagram

TABLE II
SIZE AND ACCURACY OF THE TRAINED MODELS

Model	Size	Accuracy
CNN	8 MB	98.6
MobileNet v2	2.7MB	98.15
EfficientNet-Lite	3.9MB	97.22

three different fragments and the fragment which the user is currently in gets highlighted.

2. Consistency and Standards : Within an interface if multiple words or actions are used to mean the same thing, it can only lead to the confusion in the user due to perceived lack of consistency. Interaction pattern gets disrupted. When pattern becomes complex, user's cognitive load increases. Our application has similar visual elements like buttons of all same consistency in quiz fragment

3. Flexibility and Experience of Use : The system can cater to both inexperienced and experienced users. We have made sure to include basic words and text in our application so that inexperienced user can also use our application. We have also changed our user interface according to the theme (light mode or dark mode) of user's mobile.

4. Aesthetic and Minimalist Design : Relevancy, simplicity, minimum amount of labels, un-cluttered graphics result in efficient communication dialogue between the user and the interface. In our application we made sure to make the user interface as simple as possible. We added a single screen with two fragments for login and sign up in the home page. For the second page, we included 3 fragments namely Vocabulary, Quiz and Interpreter to make it is easy for user to choose.

5. Help Users Recognize, Diagnose Recover from Errors : Preventing a user who is about to make an error would be a good approach.Gentle wording of error messages, constructive suggestions, re-educating the user- can contribute towards a happy self-confident user who is not afraid of being caught unawares or penalized. In our application, we have added an error message in vocabulary section when the user is not connected to internet. This helps the user to recognize the absence of internet connection.

6.Match between system and real world : The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow the real-world conventions, making information appear in a natural and logical order. Our application uses language that is used in our day to day life so that user can find it easy to connect to the app. For example in our interpreter fragment we used the word "open" instead of initiate/load to make it user friendly.

7. User control and freedom : Users often choose system functions which they did not want. This calls for Support undo and redo. In our application when user wants to stop quiz before completing all the questions, he/she can click back and leave the quiz, allowing user control and freedom instead of

waiting till completing all the questions.

8. Recognition rather than recall : Loading the short term memory of the user beyond a limit has negative consequences. Given a navigation path, a user need not to remember or recall all the instructions. Users are better at recognizing things they have previously experienced. In our application , with usage of symbols (in bottom navigation bar) will ease recall which eliminates the need to think while interacting .Also no memory load on user is included.

VI. Empirical Analysis

At the end of interactive system design the user interface has to be evaluated with real users.Empirical analysis is the collection and analysis of end user data to determine the usability of the system.
As we were not able to find any deaf users for our study, we created a survey which was filled by 30 participants.The results of the study are shown below.

1) How much time did you spend on learning the vocabulary videos?

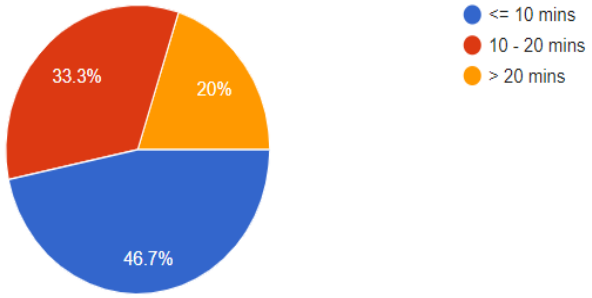

Fig. 2. 53.3% of the users spent more than 10 minutes on learning the vocabulary videos.

2) Did you have any difficulties running the app?

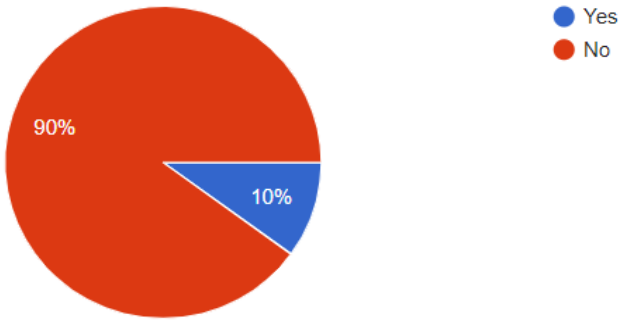

Fig. 3. 90% of the users didn't find any difficulties running the app.

3) How long did it take you to get familiar with the app?

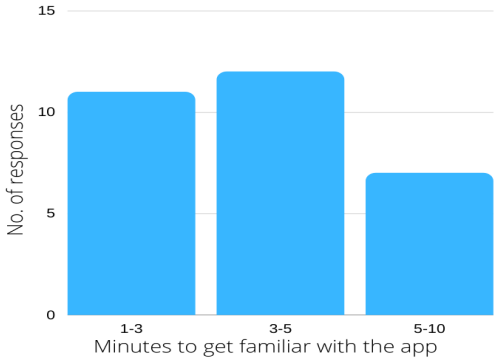

Fig. 4. 76.7% of the users got familiar with the app in less than 5 minutes.

4) How would you rate the effectiveness of the sign language interpreter?

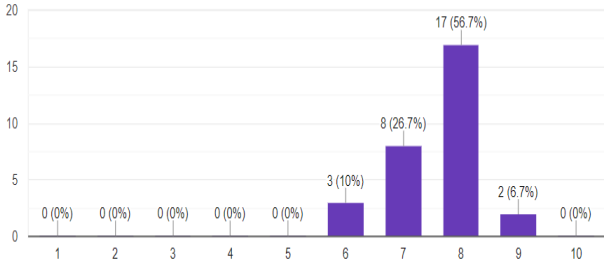

Fig. 5. Users on average gave 7.6 out of 10 for the effectiveness of the live interpreter.

5) Did the quiz module help you in evaluating your progress?

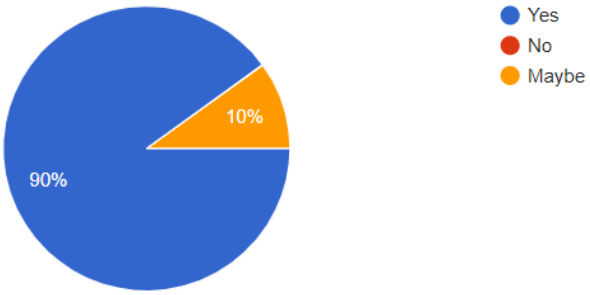

Fig. 6. 90% of the Users felt the quiz module helped them to evaluate their progress.

6) What was your score in the quiz?

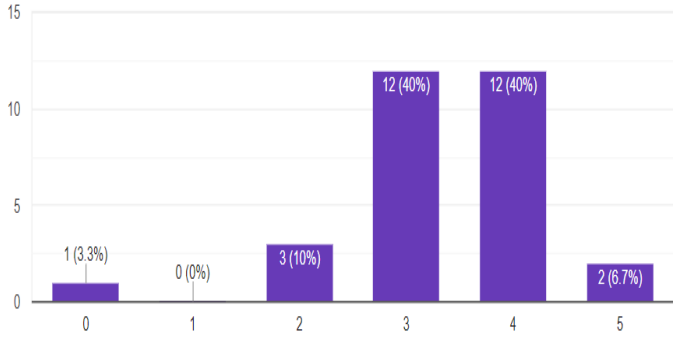

Fig. 7. Users on average scored 3.34 out of 5 in the quiz.

7) Did you notice any significant latency during loading of videos?

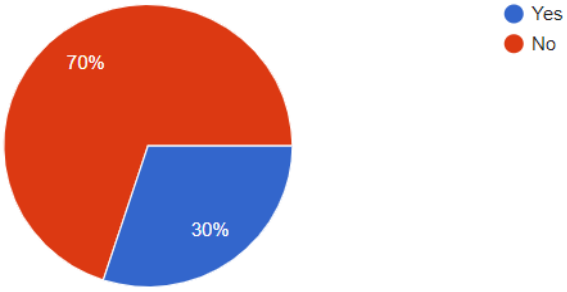

Fig. 8. 70% of the Users didn't face any latency during loading of videos.

VII. Results

The user interface of the developed application is as follows : Fig 9. shows the login page which is first screen displayed on opening the app. Fig 10 shows the sign up page which is it to be used by new users for registering. Fig 11. shows the vocabulary section, Fig 12 shows the quiz section, Fig 13 shows the score displayed when the quiz is completed and Fig 14 displays the Interpreter section. Fig 15 represents the working of the interpreter. Fig 16 and Fig 17 show the login and vocabulary page respectively in dark mode.

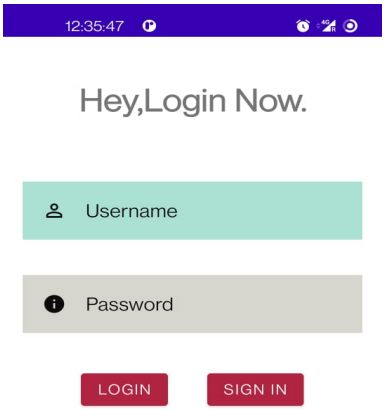

Fig. 9. Login page

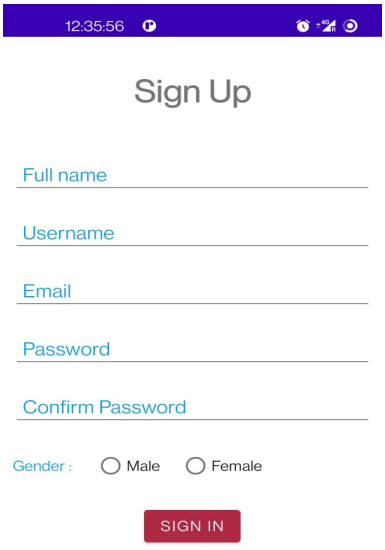

Fig. 10. SignUp page

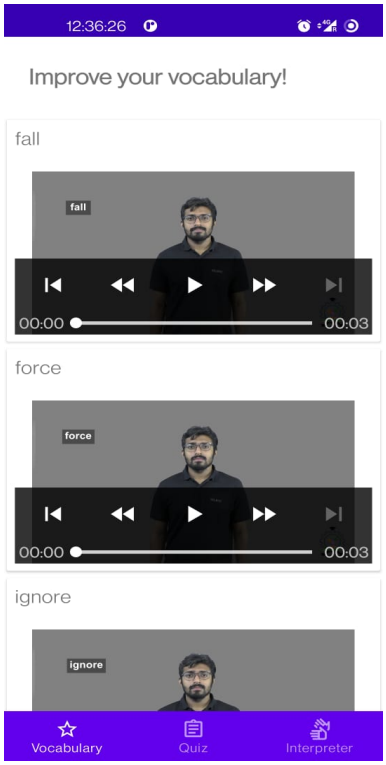

Fig. 11. Vocabulary fragment

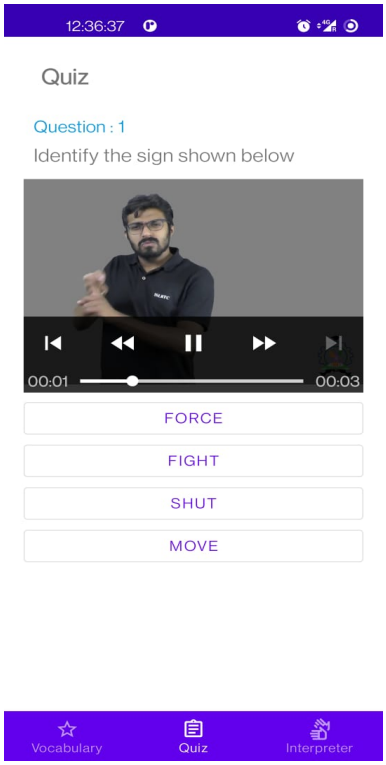

Fig. 12. Quiz fragment

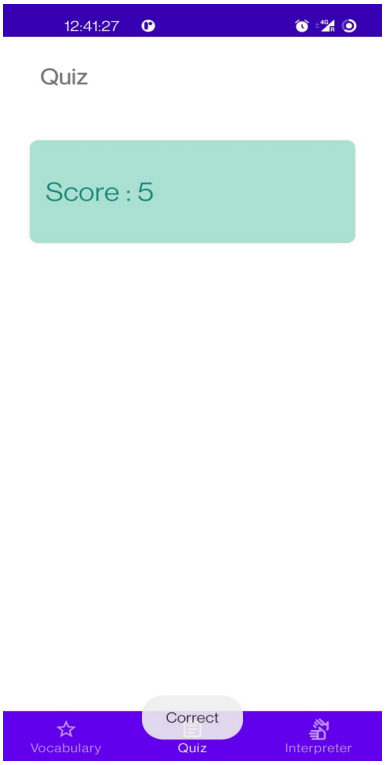

Fig. 13. Quiz score page

Fig. 14. Interpreter fragment

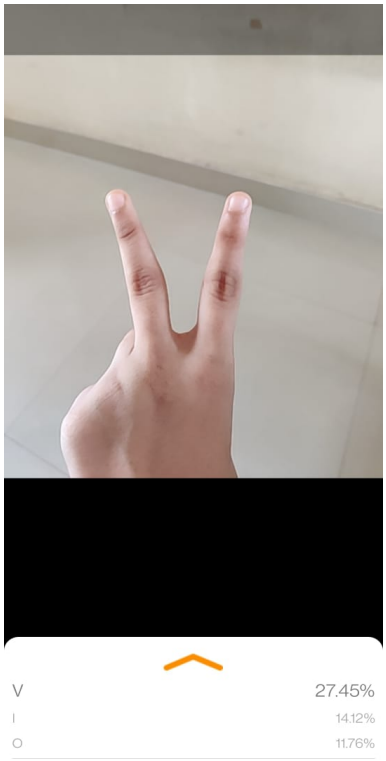

Fig. 15. Interpreter

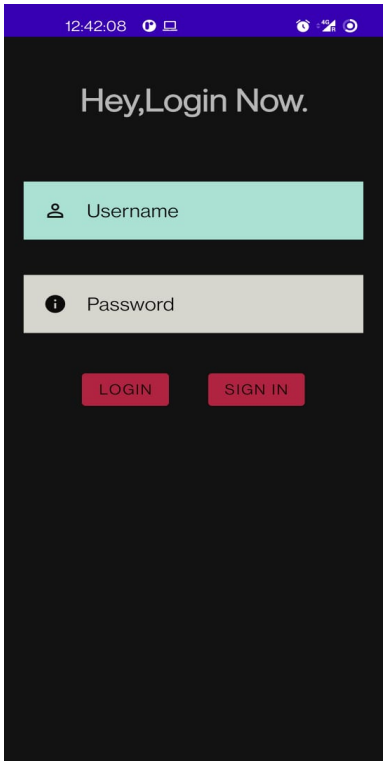

Fig. 16. Login page in Dark mode

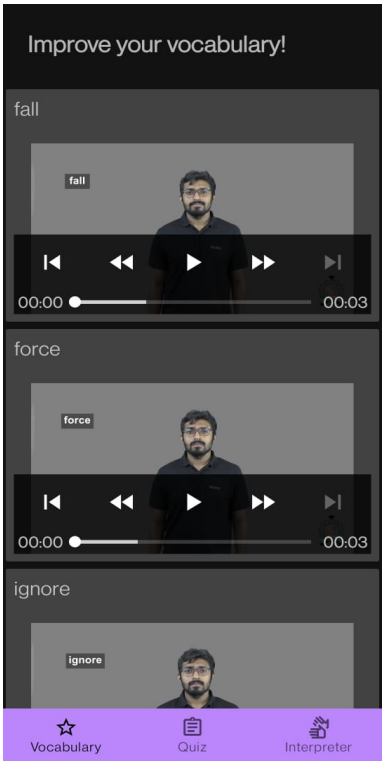

Fig. 17. Vocabulary fragment in Dark mode

VIII. Conclusion

- The developed application is easier to understand, more interesting and easy to use as the concept of learning Indian Sign Language.
- This application also helps in checking the progress of our learning by including quiz.
- Applications increase the interest of students to learn more about the sign language
- The developed application helps in building good communication between deaf community and people who don't know sign language.

IX. Individual Contribution

Devaguptam Sreegeethi - User authentication, database, Vocabulary section
Kogatam Thanmai - User Interface design for login and signup, database, Vocabulary section
Nishka Kotian - Interpreter, Vocabulary section, quiz section

Gantt Chart of the project -

X. Acknowledgement

The ISL Dictionary videos have been sourced from the Indian Sign Language Dictionary 3rd edition developed by the Indian Sign Language Research and Training Centre (ISLRTC), Department of Empowerment of Persons with Disabilities.

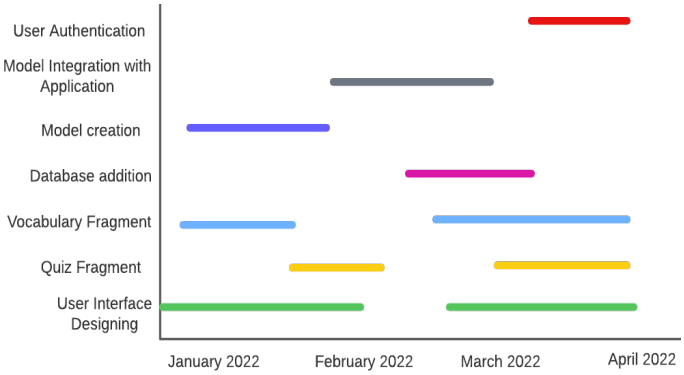

Fig. 18. Gantt chart

REFERENCES

[1] Muneer Al-hammadi,Ghulam Muhammad, Wadood Abdul,Mansour Alsulaiman, Mohammed A. Bencherif Tareq S. Alrayes, Hassan Mathkour, And Mohamed Amine Mekhtiche, Deep Learning-Based Approach for Sign Language Gesture Recognition With Efficient Hand Gesture Representation

[2] Asmaa Alsumait, Maha Faisal, CLTD, Collaborative Learning Tool for Deaf

[3] S Mohammed Siddiq, S Roopashree, Musfirah Suha, MRS Ruthvik, Kuruva Divyasree,SIGNify, A mobile solution for Indian sign language using MobileNet architecture

[4] Abhishek B. Jani, Nishith A. Kotak and Anil K. Roy, Sensor Based Hand Gesture Recognition System for English Alphabets used in Sign Language of DeafMute People

[5] Advaith Sridhar, Rohith Gandhi Ganesan,Pratyush Kumar,Mitesh Khapra, A Large Scale Dataset for Indian Sign Language Recognition

[6] Nair, A.V. and Bindu, V., A review on Indian sign language recognition

[7] Republic World, Press Trust of India, India's First Sign Language Dictionary Released, Mumbai, March 23, 2018.

[8] Lilha, H. and Shivmurthy, D., Evaluation of features for automated transcription of dual-handed sign language alphabets

[9] Mayyadah R. Mahmood1, Adnan Mohsin Abdulazeez and Zeynep Orman, A New Hand Gesture Recognition System Using Artificial Neural Network