AI-Enhanced E-Learning Platform for the Hearing impaired

TAF-24-25J-281

M.S.M Shazny

IT21173622

B.Sc. (Hons) Degree in Information Technology Specializing in Data Science.

Faculty of Computing

Sri Lanka Institute of Information Technology

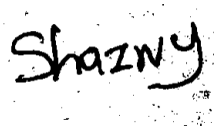
June 2024

DECLARATION

We declare that this is our own work, and this proposal does not incorporate without acknowledgement of any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Name: M.S.M Shazny

Student ID: IT21173622

Signature: 

The above candidate is carrying out research for the undergraduate Dissertation under my supervision.

Signature of the Supervisor: Ms. Thamali Kelegama

Signature of the Co-Supervisor: Ms. Vindhya Kalapuge

Date: 23/08/2024

ABSTRACT

In the current technological-based education systems, facilitating the education of learners with hearing impairment proves to be very fundamental challenge. ASL helps these students to be able to communicate with other people, though conventional forms of knowledge do not consider such children's requirements. This project fills this gap by designing a new real-time ASL translation system for educational videos for children between the ages of 8-10. The system employs masked and connected Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks for translating English to ASL with relevant non-manual signals including facial expressions, and body movements.

There are also features of captioning in real-time, as well as the summarization of videos, which comes in handy when trying to understand the content of the video clips. Coordinative techniques guarantee that ASL translations are well timed to match the video content and thus the learning process is smooth and uninterrupted. There is also a feedback option that enables users to input correcting feedback on the translation so that the model can be refined in real-time.

In the development process, the identification of ASL signs and expressions, testing, and revising the system to increase the level of translation accuracy and user satisfaction were conducted. The end product serves the purpose of making education more accessible by the deaf and hard of hearing students while at the same time allowing every learner in the classroom to participate freely. The system is also flexible and can be implemented at the classroom level and at a campus or even an online learning level. The goal of this project is to raise the bar of education to where with or without hearing impairment a student should be able to get whatever he/she wants.

TABLE OF CONTENTS

Introduction

As we enter the age of eLearning, ensuring that each and every student has equitable access to knowledge becomes all the more crucial. For students who are Deaf or hard-of-hearing, American Sign Language (ASL) is not just a mode of conversation; it is the most powerful means by which they could fully know more about as well as communicate with their world. Despite the advances made in educational technology, there are still major gaps: Much of our digital content is not designed with ASL users as a primary or even consequential audience.

Try to think of a universe in which every educational video, lecture and online resource is not only made available but also specifically catered on the individual language needs of ASL users. A world where students who use sign language can learn just as well from educational content in real-time, without being the last one there. This is not a dream, it's what we are using to drive innovative research.

This project aims to democratize access for ASL users into the digital education system by providing real-time translation between English and American Sign Languages. This system fundamentally relies on robust machine learning algorithms that easily convert spoken language into American Sign Language (ASL) visually enhanced sign language avatars in an equivalent process. But it doesn’t stop there. In understanding that all students have differing needs, the system also utilizes synchronized captions and summarizing bullets to provide not only sight of what was signed but an explanation on context and key points within each lesson.

Videos will therefore have their educational information converted into a fully interactive experience for ASL users which should make them just as accessible and natural to use. The captions will match the video exactly, and ASL avatars will provide an easy-to-view visual aid of the spoken words. In addition, the summary viewpoint will sum up the core points which provide students a short and effective method of reviewing what they have been taught.

This research is more than technology; it about tearing down walls and constructing an inclusive educational landscape where every student can achieve. Focusing on ASL, we open the door to a new community of learners who have otherwise been marginalized by traditional educational methods, thereby ensuring that they flourish in the marketplace.

If the world is constantly changing, our educational systems should do so as well— become more inclusive and accessible to all students. And this new real-time ASL translation system is not just a step in that direction; it's practically jumping into the promise of endless learning.

Background Literature

There is no doubt that as digital education becomes more widespread, creating learning spaces accessible to all — especially students who are deaf and hard of hearing — eclipses a necessity. Because ASL is the predominant form of communication for many in the Deaf community, it is important to adapt our educational content accordingly. However, because ASL has different grammar and syntax than spoken languages do, it can be even more difficult to match English words meaningfully into ASL. Many existing educational platforms were created for hearing students, so they are heavy in auditory and textual based content which makes it difficult to truly meet the needs ASL users have. While captioning is a step towards accessibility, it often fails to completely capture the rich linguistic and cultural nuances in ASL resulting into less comprehension or even disengagement from students needing this language [1].

Researchers have experimented with numerous technological solutions to counteract these issues, and the current evolution of machine learning techniques shows promise in overcoming some of them. Two of the existing end-to-end approaches for real-time ASL translation systems use Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. The hope is that these technologies will provide translations of spoken language into ASL with the same accuracy as they did for foreign languages, so as to ensure an equitable learning experience in ASL production for deaf children (Kambhampati et al., 2024) [2]. Reversible CNNs can be employed to translate spoken language and ASL, preserving the end-to-end flow of video contrasted with similar frame-by-frame approaches. It improves the precision of translations and make them reliable because forward and backward bookings both can be simultaneously trained on (Kambhampati et al., 2024) [2].

However, despite these improvements, multiple problems still hinder the development of real-time ASL translation systems. A huge problem is sync between the translated ASL content, and original video or audio material that accurately represents features of information being relayed. Dodandeniya et al. developed the "Visual Kids" platform, as shown in Fig. (2023) demonstrated up to 96% accuracy in sign language recognition which is an encouraging performance rate [2], but also pointed out that the alignment of ASL translations with educational content, a key component for effective communication and learning, was challenging. Not to mention the grammar and syntax of ASL is much more complex, as it does not map directly onto English. ASL is a context-based language where emotions, body motions and spatial orientation in addition to hand signs are used to express meaning. This complexity makes it challenging to create systems that can model the nuances of ASL adequately for real-time applications (Sindhu et al., 2024) [1].

One innovative model used by Hastha and collaborators is sign language avatars in online learning platforms to educate students with hearing impairments. It is a set of avatars that can serve as visual counterparts to ASL translations, which may be matched with educational elements for a more natural learning process. This, however, requires development of advanced algorithms that can comprehend and represent the complex details in grammar definition and syntax (Hastha et

al., 2024) [4]. Moreover, the findings of "Enhancing Remote Learning for Deaf and Hard-of-Hearing Students" suggest use of a multiple communication format integrating real-time captions with ASL to serve diverse needs among hearing-impaired students during distance learning. Doing so, especially in times of global crises where traditional one-on-one support might not be possible the need for more accessible educational platforms (Hastha et al., 2024) [5] is emphasized.

However, as it stands now, the advancement of more stable ASL translation systems is greatly associated with the readiness of rich and diverse datasets for training models. There is always the proviso that their usefulness will depend on the quality and variety of data that has been fed into the models. However, it is, however, the case that there is no large-scale, multi-regional, and multi-variety corpus of ASL available at the present time which decreases the applicability of these systems in the different educational institutions (Sindhu et al., 2024) [1]. The study specified as “Sign Language Recognition and Translation Systems for Enhanced Communication for the Hearing Impaired” also mention that the biggest and most varied datasets should be established to enhance the translation models of ASL, which is necessary for actual application (Sindhu et al. , 2024)[1] .

All in all, one can outline several key issues that remain critical despite the rather successful advancements in the area of translating ASL for education-related content. These include the requirement of coordinated time of translate and original texts, ASL grammar, lack of large amount of data set for the machine translation model training. Future studies and developments should be employed to combat these issues and make digital education effective and open to all deaf students. The current research activities, some of which are exemplified by Sindhu et al. (2024), Dodandeniya et al. (2023), Hastha et al. (2024) others include significant interventions to makes digital learning environment more accommodating to ASL users.

Research Gap

Strengthening research gaps is an informative stage in any research works as it makes clarifies areas or knowledge that would be useful to investigate to establish areas of knowledge that has not been explored adequately. New approaches also give the researchers an opportunity to take up new areas and make new discoveries thus contributing to the field. This study has pointed out the following rewards and difficulties in the specific domain of constructing real-time ASL translation systems for educational uses. These are the lack of more extensive correspondence between signed languages and instructional material, the challenge of translating sign language, such as ASL, and its structure and syntax, as well as the general absence of large data sets that can be used to train artificial intelligence systems. Based on an analysis of the selected research papers, this present study was able to identify several general research gaps as follows:

The first of the analyzed research papers [1] is devoted to the development of sign language recognition and translation systems for improving communication capabilities of hearing-impaired individuals. Although the study shows remarkable advances in the field, especially with the utilization of Convolutional Neural Networks (CNNs) for Sign Language Recognition, the authors do not contemplate on how to synchronize ASL translations in tandem to video content in classrooms. Consistent with learning contexts, another major failure is the absence of temporal correspondence between translated content and education material since failure in synchrony guarantees poor comprehension in translated environments. More research needs to be conducted on this problem with the aim of coming up with better techniques of enhancing the efficacy of ASL translation system for practical use.

The second research paper [2] focuses on an idea of improving the ASL translations' accuracy by applying Reversible CNNs. While the effectiveness of the study is observed to enhance the extent of translation accuracy, the study fails to capture the mechanistic and structural aspect of ASL that incorporates hand signs, facial expressions, and body gestures. The absence of such research is quite striking, especially when it comes to translating ASL for its subtle features are vital for the meaning. There appears to be a need for more research that will enable the modeling of the ASL expressions fully and map the new models obtained to the new context and culture appropriately.

The third research paper [3] presents on the development of "Visual Kids" an open educational platform designed for teaching hearing impaired Primary school students. When tested, using sign language recognition, the platform scored a high of the recognition accuracy, however, the researchers realized that there was problem with the temporal alignment of the ASL translations to the content of the educational videos. This may cause confusion as well as lowered understanding for students; hence the requirement for additional studies to establish the ways to achieve more accurate coupling of the translations with instructional materials.

The fourth research paper [4] deal with sign language avatars in technology ehanced learning system for hearing impaired students. The study provides a prescription on the possibility of solving the challenge on translating signs for ASL; but, the study does not proffer evidence on how to make the existing sets more expanded in order for these systems to learn from. Most of ASL translation systems suffer from the absence of large featuring diversified signs, gestures and comprehensible contexts in datasets. This gap implies the need for more research proposals that will involve the development and management of big data sets that can aid in easing training of more sophisticated models of ASL translation.

Last of all, the fifth research paper [5] captures the essence of the use of different modes of communication that can be useful in the current remote learning including real-time captioning for the Deaf, sign language. Although the study focuses on the aspects of accessibility during such global challenges as COVID—19 pandemic, the study lacks suggestion on how the quality of translation could be maintained when translated across platforms and contexts in such a process. Maintaining the ASL translations' quality in different learning contexts is relevant for their efficacy, therefore, the importance of further investigations with the aim of identifying the ways to sustain high translation quality is underlined.

Research Gap Feature	**Research 1 [1]**	**Research 2 [2]**	**Research 3 [3]**	**Research 4 [4]**	**Proposed Research Solution**
Synchronization of ASL translations with video content	✗	✗	✗	✗	✓
Comprehensive dataset inclusion for diverse ASL signs and expressions	✗	✗	✗	✗	✓
Accuracy of real-time ASL translation with full grammar and syntax	✗	✓	✗	✗	✓
Integration of facial expressions and body movements in ASL translation	✗	✗	✗	✓	✓
Real-time feedback for ASL users to correct translation errors	✗	✗	✗	✗	✓
Scalability of ASL translation systems for various educational settings	✗	✗	✗	✗	✓

Research Problem

There is no doubt that the general expansion of digital education poses various difficulties in the provision of equal opportunities to learners, including those who are deaf or hard of hearing. For most of the people in the Deaf community, ASL is the means of communication and avenue through which they can learn. Nonetheless, 95% of current digital educational content considered and/or developed needs the hearing students more rather than focusing on the ASL users since most of the information provided in this medium is based on audacious and textual information. This setup leads to a huge accessibility disconnect where minority students such as the deaf and the hard of hearing are disadvantaged in education. While the traditional way of captioning provides some assistance, it is not very effective when it comes to the grammar, syntax and cultural aspects of ASL due to which learners lose interest and fail to grasp the content fully.

Even this is made even more complex by the fact that ASL has a structure in terms of language that is very different to spoken languages since it is a visual language that uses hand shapes, movements, facial expressions and postures to communicate. While English has a written counterpart ASL does not and thus there is no direct translation from ASL to text. This poses major challenges when trying to create applications that provide close to real-time translation of spoken or written English to ASL. Current learning systems and applications, if they offer ASL translations at all, do it in a very limited manner or inaccurately and often do not support ASL at all, thus contributing to the existing education inequalities of the deaf and hard of hearing.

Furthermore, the concept of translation in ASL is even more than just a word-to-word translation process. Grammar in ASL is non-conventional in the sense that it does not relate in the conventional manner with the writing system of English as it can be understood both linear and non-linear in terms of message transmission which can involve the use of hands, face, head and other parts of the body simultaneously at once. This must involve sophisticated algorithms to interpret these multi-mode inputs and generate and output appropriately contextualized ASL. Recent developments in the machine learning, including CNNs and LSTMs, demonstrate some progress in this regard yet, numerous problems that must be solved regarding the high accuracy, in real-time systems.

Another important problem is the absence of large-scale data sets that contain the various ASL signs, the expressions and scenarios. When it comes to the training data, there is a kind of direct proportion between the quality and the quantity of the data and the performance of the machine learning models. However, there is not much availability of a big repository of annotated data in ASL that encompasses all the aspects of ASL such as regional dialects and further hand gesture with face and body actions. This limitation makes these models very limited in their applicability between different users or with various education systems and contexts.

In addition, another difficulty of conveying the implementation of ASL translations is that of achieving a consistency between sign language interpreting and educational material. The temporal synchronization of translations to the spoken or written is important in order to avoid gaps in information flow or, contrarily, information overload which might confuse the reader or listener. Current systems often fail on this account and as a result learner gets a very disconnected experience that hampers their learning.

Further, current approaches of ASL translation do not have the capacity to give feedback as they happen in real time hence proving useful for correcting mistakes. If the feedback is delayed improper interpretation of information can be over-emphasized and further worsening of the educational inequality tendencies will be inevitable. As there are no learning environments that can be adapted according to the students' performance, students who face difficulties in a particular type of a problem also will not be provided with an adequate support they require.

Last but not the least; the possibility of making the ASL translation systems portable across different educational environments and platforms and onto different devices is yet another problem area. These systems come in different platforms and can be accessed in different environments, and this may lead to variation in system performance rendering different performance to the users. Adapting the usability of those technologies consistently across different platforms is vital for practical implementation in schools.

These challenges are enough to warrant concrete overarching solutions that would be able to meet the complexity of ASL translation in digital education. This is through work with enhanced machine learning models that can translate ASL in real time, through elaboration of corresponding large and diverse datasets and through the proposals for systems that provide translations that not only are synchronized, but also contextually appropriate. Also, continuous evaluation of the methods, guaranteeing that they can be deployed in any learning settings and designing effective ASL learning environments are also critical features of improving ASL translation. Tackling these problems will be essential for narrowing the accessibility deficit in the experiences of deaf and hard of hearing students and for guaranteeing their equal learning opportunities to the provisions made for hearers in the digital environment

Research Objectives

MAIN OBJECTIVE

- The objective of this study is to design a real time translation system for educational content which will translate both spoken and written English into American sign language, caption, and summary. The system is designed to support the improvement of inclusive learning environment for deaf and hard of hearing students through real time ASL translations that reserve the value and importance of rich content being conveyed. The work propose to distil how advanced machine learning models can be used to translate ASL and include mechanisms for real-time feedbacks and learning environments. The goal is to attain equal learning quality for the deaf or the hard of hearing students to that of the normal hearing students to improve the embracing of the digital learning process. Thus, this goal of the study may contribute to the development of the field of educational accessibility, especially in any learning technologies using ASL translation.

SPECIFIC OBJECTIVES

- **Specific**
  - Develop a machine learning model capable of accurately translating spoken and written English into ASL in real time.
  - Use facial expressions and body movements in the interpretation to show all the ASL grammar and syntax.
  - Develop a coordination system that would enable the sign language translation to be in tangent with educational content.
  - Assemble a wide and extensive amount of data that encompasses many signs, expressions and contexts in ASL.
  - Propose and put in place an instant feedback mechanism to enable the ASL users to give feedback on translation errors instantly.

- **Measurable**
  - Test the quality of the ASL translation model by comparing the results of the model to native ASL interpreters for the same content.
  - Evaluate the performances of the synchronization mechanism by checking the degree of synchronization between ASL translations and educational content.
  - Assess the extent and the variety of the ASL materials by looking at the number of signs, facial gestures, and situations.
  - Quantify rotation speed_ Real-time feedback If the feedback channel is reliable, it means that users are correcting translation errors within the following time frame.

- **Achievable**
    - Ensure that the development of the ASL translation model is feasible using current machine learning techniques and available datasets.
    - Verify that the synchronization mechanism can be reliably implemented in real-world educational settings.
    - Confirm that the real-time feedback system is user-friendly and can be effectively integrated into the ASL translation process.

- **Relevant**
    - Mention the lack of accessibility for disability in the realm of technology incorporation in education for the deaf or hard of hearing students.
    - Meet the need for correct and culturally appropriate signs for the translations of educational materials in ASL.
    - Meet the demand for the more significant and diverse dataset to train ASL translation models and enhance their performance across different domains. Here is the need to have feedback systems that immediately enable the improvement of the learning environment for ASL users.

- **Time-bound**
    - Establish the ASL translation model in the next six months; conduct the ASL translation model testing and validation in the next two months.
    - Develop and trial the synchronization mechanism within a period of not more than three months.
    - Develop and populate the ASL dataset in a period of four months.
    - Devise and incorporate the real-time feedback system in three months and evaluate one more time and continuously do so for five months.

Methodology

4. 1 Data collection and preparation

The first phase of the methodology is to compile a large database of signs that pertain to several perspectives of ASL, frowns, motion of the body, and the verbal and written English translations. The dataset to be used will be collected from ASL databases, video databases and with the help of ASL teachers and members of the Deaf community. As will be seen, this dataset will contain a wide range of educational videos with the aim of testing the system with various content types. To this end, the data collected will be preprocessed to normalize frame of the videos; removal of noise; and segments the ASL signs into various parts such as hand shape, movements and the facial expressions.

4. 2 Development of the Translation Model

The essence of the approach is in creating an online machine learning model that will translate spoken and written English in videos into ASL. The model will be developed using some of the latest algorithms such as Convolutional Neural Networks (CNNs) for visual recognition problems, and Long Short-Term Memory (LSTM) for sequence data – signs. CNNs will perform feature extraction for the videos for shape, movement and facial expressions, while LSTMs will analyze the temporal aspect of the sign usually executed in sign language interpretation. This is a common practice in the model training, and it is expected that through several iterations of training and validation on the preprocessed dataset, the model will incorporate optimizations that enhance the accuracy of translation and eliminate errors.

4.3 Synchronization with Video Content

In order that ASL translations are in time with the educational video content, a synchronization system will be created. This mechanism will ensure that, the ASL translations done correspond with the audio or textual information in the video. The synchronization of signs in the translated videos will be done using DTW and temporal alignment algorithms, so as to match the signed and spoken/written languages precisely in terms of time.

4.4 Real–Time Captioning and Summary Generation

Other components included in the system will be real-time captioning and summarization capabilities. The captioning system will directly take text captions from the spoken content that appears in the videos through speech-to-text. At the same time, fine-grained summaries of the content of the video will be created with the help of a Natural Language Processing (NLP) model. These captions and summaries will be timed with the ASL translations in order to allow every student to watch the videos with ease.

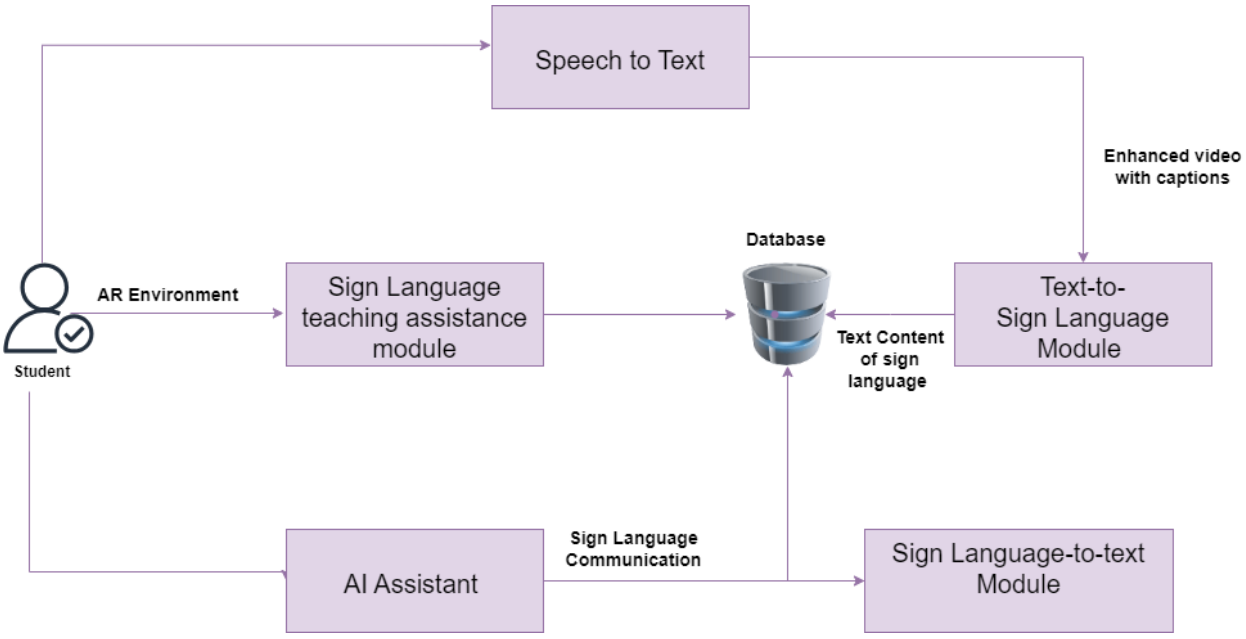
4.5 Testing and Validation

Afterward, this system will be tested and validated stringently to determine the efficiency and precision of the system. This will require a number of iterations of the tests with ASL users, the native signers and educators, the assessment of the quality of the translations, the synchronization of the materials, and the functions of the system. Therefore, objective time-related values (e. g. translation accuracy time, synchronization delay time) and subjective feedback (e. g. end user satisfaction, perceived ease of use) will be collected to determine areas that requires enhancement.

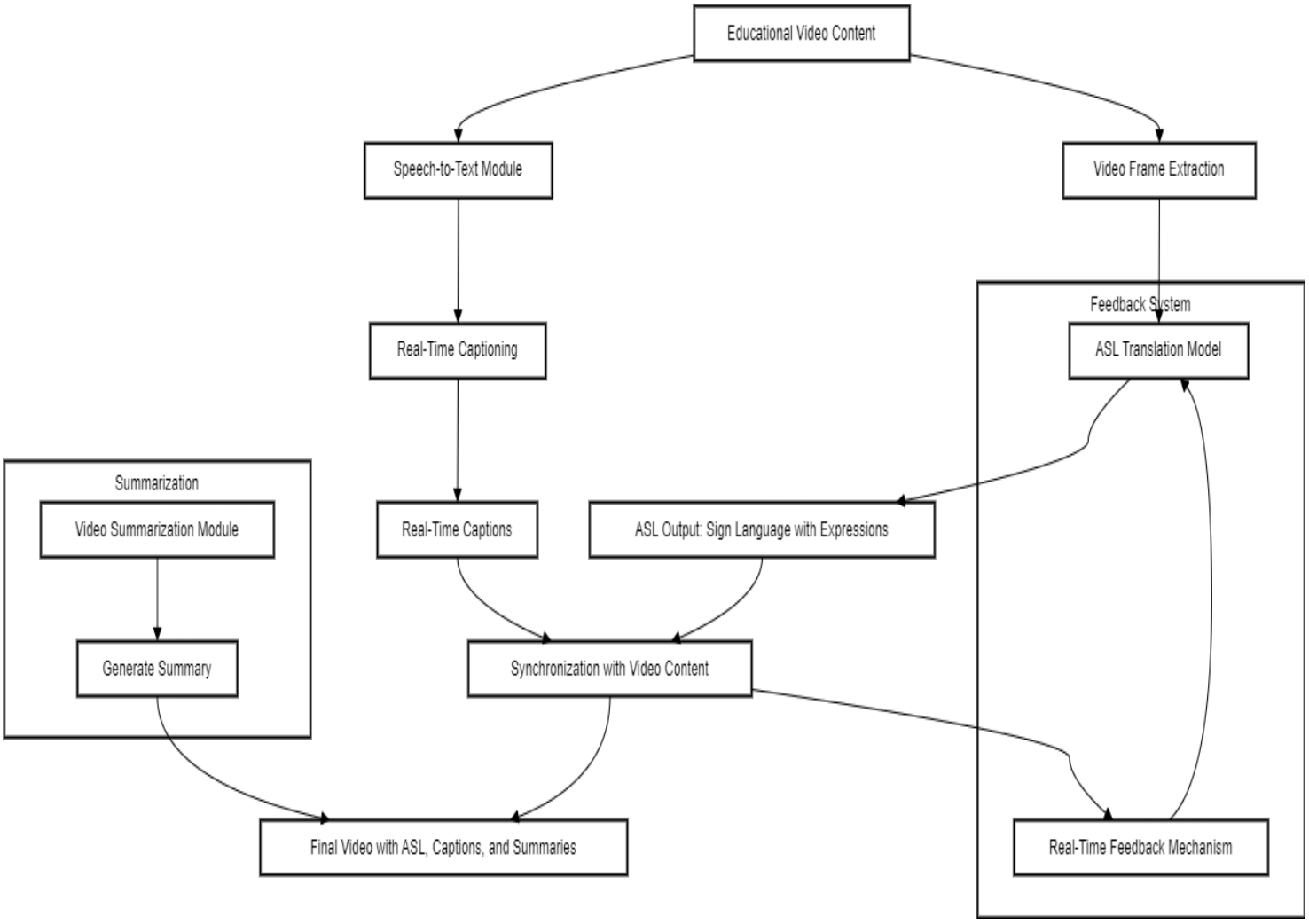
4.6 Deployment and Evaluation

The final step, the system will be incorporated into a live education environment like a school or an online learning environment to determine the effectiveness of the system on learning achievement for the deaf/Hoh students. Recommendations from the pilot study will be useful in any additional changes before a full-scale deployment of the method. The purpose is to incorporate the system into different platforms of education to make necessary learning for ASL available and convenient.

6.1 System Architecture



Component Diagram



Commercialization Of The Product

With focus on educational institutions, online learning platforms and the accessibility enthusiasts, the professional version of the real-time ASL translation system will be offered with subscription that allows access to all the features and additional functionality. All of the features of the product will be available to subscribers to provide them with accurate and in-sync ASL translations, real-time captions, and detailed video summaries. The proposed model of the service is based on subscription and will provide continuous revenues for further development and improvement of the product and constant individual assistance to the clients.

Apart from the professional edition of the product, there will be also the free edition of the product that will have some limitations but will be useful to the users who are not professional educators but teachers who are running their own classes, educational enthusiasts, students or people with disabilities or those users who want to contribute to the development of the product.

The commercialization strategy is to promote the research product as a market leader in accessibility to education and in particular ASL translations for digital content. Thus, the product can be both for paying customers and for the general population, which will make its use widespread, backed up by proper marketing and cooperation with educational organizations and initiatives for people with disabilities. This should not only help to establish a source of income but also serve a greater purpose which is to provide education to the students who have hearing impairment.

TARGET USERS

The ideal market for the real-time sign language interpreting system is schools for the deaf and dumb, elementary schools, and online classes for kids of 8-10 years. This system is especially appropriate for special education departments, accessibility coordinators, and educational managers who are concerned with the problem of making education accessible for young deaf and hard-of-hearing learners. Also, the product is aimed at non-profit organizations, advocacy groups by children's education and disability, as well as government organizations striving for inclusive education. Some of the secondary users are individual educators, ASL interpreters and parents who want to supplement learning activities of children within this age bracket.

MARKETING AND REVENUE

The commercialization strategy for the real-time ASL translation system includes a dual-tiered approach: Professional edition with subscription required, and the basic edition with limited features is available for download. The professional version meant for institutions and organizations working with young learners of 8 to 10 years will provide all the feature such as ASL translation, real time captioning and video summarization which will be in child friendly format. This version will have a recurring revenue stream through subscription, to fund the continuous improvement of the product, new releases, as well as providing dedicated specialized customer service to schools with a focus on this age group. The free edition, as it is less functional in comparison with the paid one, is going to attract small schools, individual teachers, and parents, thus driving the number of users and making the brand more recognizable.

MARKETING APPROACH

The marketing strategy for the piece of technology to translate ASL in real time will be the direct mail and partnership with and content marketing to educational institutions and other organizations that engage with children of 8-10 years of age. The strategy will include developing individual relationships with schools, educational technology firms, and organizations that support child education through a one-on-one engagement such as product presentation, virtual tours, and educational sessions. In order to address credibility and increase visibility, collaborations with relevant stakeholders in the educational field and organizations which promote accessibility will be sought. Some of the content marketing strategies to be implemented shall be blog articles, case studies, and video content illustrating the improvement of the system in terms of young learners' accessibility and engagement, for children in the target age group. Awareness activities such as advocacy on social media platforms and hosting of educational conferences will also be relevant in the promotion of the project and the establishment of a community of beneficiaries who are willing to ensure success of the targeted young deaf and hard of hearing students. The strategy is to popularize the ASL translation system as a crucial component of elementary education and ensure its demand by presenting data of its effectiveness for children of 8-10 years old.

Project Requirements

FUNCTIONAL REQUIREMENTS

- **Real-Time ASL Translation**
  - It requires the system to automatically transcribe English speech from educational videos and convert the texts to American Sign Language (ASL) for the group of children of 8 to 10 years old.
  - It should convey the meaning and context as well as cultural aspects of the translated content, as well as the features of ASL, such as facial expressions and body gestures.
- **Synchronization with Video Content**
  - The system must make certain that the ASL translations are perfectly in tandem with the audio as well as video segments of the videos.
  - This synchronization should help keep the flow of information so that when the signs are being made, the spoken words and visual cues within the context of the video are also being made.
- **Real-Time Captioning**
  - The text caption should be real-time and should be derived from the spoken content that is in the videos.
  - Captions have to be appropriate for children and given in simple words, the text has to be close to the level of children in the age of 8-10.
- **Video Summarization**
  - The system should also be able to provide brief overviews of the video material which can be comprehensible to children of a young age.
  - Such summaries should be shown at the end of the videos or as an option of the video, perhaps a ‘read more’ kind of a feature.
- **User Interface**
  - Usability should be considered, making user interface attractive to children of 8-10 years.
  - It should enable viewers to find their way through the videos easily, view ASL translations, captions and summaries The controls should be basic to enable the manipulation of the play back and links to other resources.
- **Real-Time Feedback System**
  - Real time feedback which would enable users such as teachers, ASL interpreters or parents to counteract any translation error on the system on the go should be incorporated in the system.
  - Such feedback should be used to enhance the quality of future translations and the algorithms of machine learning.

NON-FUNCTIONAL REQUIREMENTS

- Performance
  - To meet a user's needs, the system must offer ASL translations, captions, and summaries as soon as possible without having a considerable delay. It should be possible to get translated texts and captions in a short time, namely, within 1-2 seconds.
  - The system should be able to support large amounts of video content and at the same time handle multiple users with a possibility of users translating at the same instance.
- Scalability
  - The system has to be able to grow with the popularity and the amount of videos that are being offered and watched by its users. It should be equally suitable for point solutions (as in individual classes) and scaling up to entire school districts, or even online learning environments.
    The architecture of the system should be such that, it should support the addition of new features and interface with other educational apps and systems.
    Reliability and Availability
    It should be very reliable and should be at least 99% up time. 9%. It has to be always available and especially when the usage is high, for example during school times.
    Measures of data recovery and data backup should be implemented adequately in case of system failure is experienced.
- Security
  - The system must use an appropriate level of security to preserve the user data; for instance, all the information on the users, their feedbacks, and the videos, among others, should be encrypted.
  - Measures should be put in place to ensure that only genuine users (for example educators, students, administrators) are the only that should be allowed to both view and modify content and settings.
  - The system should adhere to the particularities of data protection legislation (e. g. , GDPR, COPPA) to protect the rights of students.

- Usability
  - It should be easy for the children of age 8-10 to be able to navigate in the user interface for the games that will be created.
  - The help should be easily available on the system in the form of tool tips, help, and even a help section where educators and students can find the help they need in using the system.
    The interface should have features for different users that include features such as text size, color and arrangement of items on the interface.

Anticipated Benefits

- Enhanced Accessibility for Deaf and Hard-of-Hearing Students
  - The real-time ASL translation system will significantly improve access to educational video content for deaf and hard-of-hearing students, particularly those aged 8-10 years old. By providing accurate and synchronized ASL translations, the system ensures these students receive the same quality of education as their hearing peers, promoting better engagement and understanding of educational material.
- Improved Educational Outcomes
  - By making educational content more accessible, the system is expected to enhance comprehension, retention, and overall academic performance among deaf and hard-of-hearing students. With better access to learning materials, these students are more likely to stay engaged, participate actively in the classroom, and achieve better educational outcomes.

Conclusion

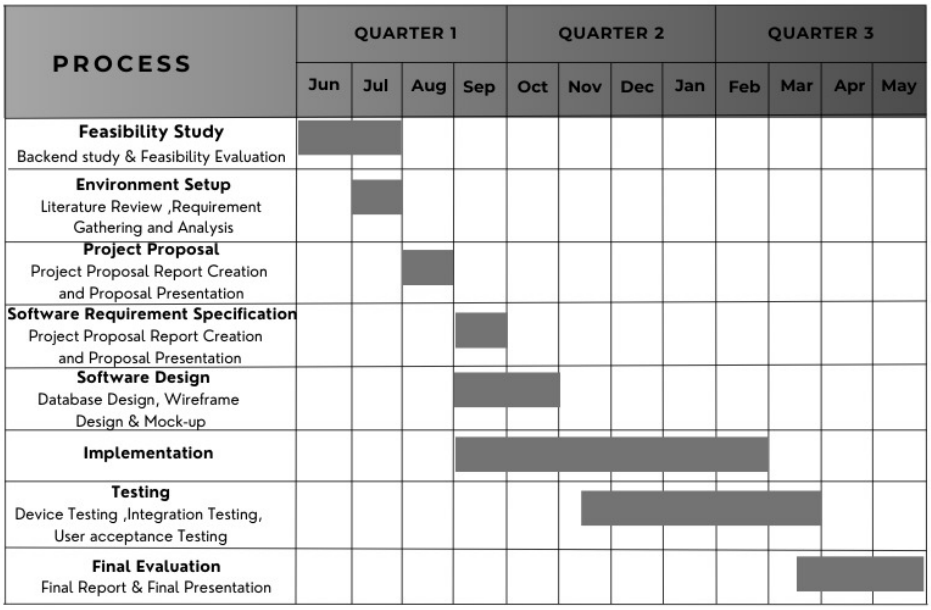
A true-time ASL translation of educational videos is a great enhancement to achieving education equity for the deaf and hard of hearing student. As the society goes more and more digital, it becomes even more important to provide equal opportunities for learning to all children, including those with hearing impairment. This project is doing more than filling the need for the lacking educational technology area but will pave the way for a better learning environment in which all students can succeed.

The concern with real-time translation into ASL and particularly for the children in the age bracket of eight to ten years is the kind of knowledge that only has a profound comprehension of the challenges that children in this particular age face. With the use of the state-of-the-art machine learning models, the system was able to incorporate all the ASL features such as the grammar, syntax and the important non-manual features such as the facial expressions and body movements. The integration of these translations with videos guarantee that students are provided an unbroken and consistent learning process that is necessary for their understanding and absorption of the content.

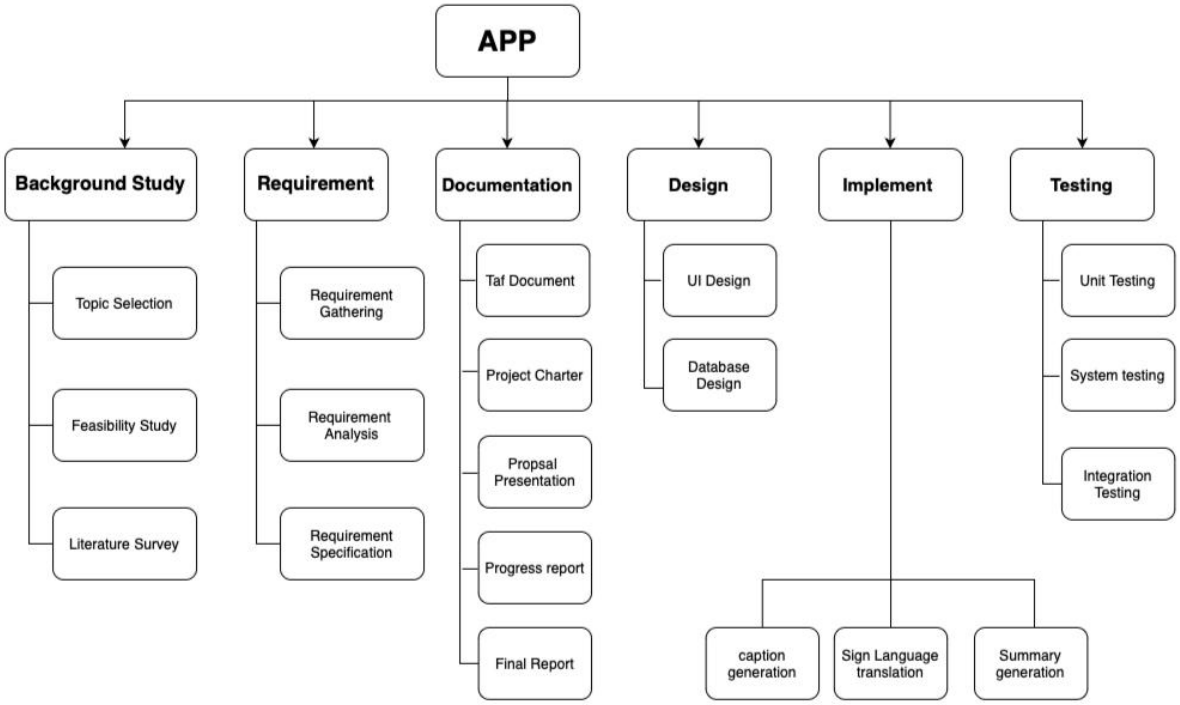
Furthermore, being applied with real-time caption and video summarization is another plus for the system since it provides learners with multiple ways of receiving information. This approach of instruction and communication is not only advantageous to the students who are deaf or hard of hearing but also advantageous to all students, as it increases its overall effectiveness and makes the classroom a more engaging place.

The following are the expected gains likely to accrue from this system though the gains are not limited to educational results only. Thus, supporting the ideas of integration of the members of the Deaf community into the mainstream schools and utilizing ASL as the communication tool, this project promotes the change of the culture in society. Since the system has a modular and flexible structure, it can be used in different educational settings from the single classroom to the large scale online environment, which makes it effective instrument in the struggle for equity in education.

11. Gant Chart



12. Work Breakdown Architecture



References:

1. Sindhu, K., et al. (2024). *Sign Language Recognition and Translation Systems for Enhanced Communication for the Hearing Impaired*. BVRIT HYDERABAD College of Engineering for Women.
2. Kambhampati, et al. (2024). *A Study on Reversible CNNs for Accurate Sign Language Translation*. IEEE Conference on Cognitive Computing.
3. Dodandeniya, J.M.D.G.K.M., et al. (2023). *Visual Kids: Interactive Learning Application for Hearing-Impaired Primary School Kids in Sri Lanka*. International Conference on Advancements in Computing.
4. Hastha, et al. (2024). *Online Learning Platform for Hearing Impaired*. IEEE Xplore.
5. Hastha, et al. (2024). *Improving Virtual Educational Platforms for the Deaf and Dumb under the COVID-19 Pandemic Circumstances*. Journal of Educational Technology & Society.