Lip Reading Recognition for Deaf and Dumb using Machine Learning

Project report submitted

in

partial fulfillment of requirement for the award of degree of

Bachelor of Technology

in

Information Technology

By

Ms. Prachi Bamhore **Ms. Sajeeda Noori**

Ms. Shipra Gattewar **Mr. Md Uveshraza Salmani**

Guide

Dr. Ashlesha Nagdive
Assistant Professor
Department of Information Technology
G H R C E, Nagpur

Department of Information Technology

G H Raisoni College of Engineering, Nagpur

(An Empowered Autonomous Institute affiliated to R.T.M. Nagpur University, Nagpur)

NAAC Accredited with “A++” Grade (3rd Cycle)

Ranked in the Band of 151-200 in Engineering Category by NIRF Ranking 2023

December 2023

Lip Reading Recognition for Deaf and Dumb using Machine Learning

Project Report Submitted

in

partial fulfillment of requirement for the award of degree of

Bachelor of Technology
in
Information Technology

By

Ms. Prachi Bamhore **Ms. Sajeeda Noori**

Ms. Shipra Gattewar **Mr. Md Uveshraza Salmani**

Guide

Dr. Ashlesha Nagdive
Assistant Professor
Department of Information Technology
G H R C E, Nagpur

Department of Information Technology

G H Raisoni College of Engineering, Nagpur

(An Empowered Autonomous Institute affiliated to R.T.M. Nagpur University, Nagpur)

NAAC Accredited with “A++” Grade (3rd Cycle)

Ranked in the Band of 151-200 in Engineering Category by NIRF Ranking 2023

December 2023

DECLARATION

We, hereby declare that the project report titled "**Lip Reading Recognition for Deaf and Dumb using Machine Learning**" submitted herein has been carried out by us towards partial fulfillment of requirement for the award of Degree of Bachelor of Technology in Information Technology. The work is original and has not been submitted earlier as a whole or in part for the award of any degree / diploma at this or any other Institution / University.

We also hereby assign to G H Raisoni College of Engineering, Nagpur all rights under copyright that may exist in and to the above work and any revised or expanded derivatives works based on the work as mentioned. Other work copied from references, manuals etc. are disclaimed.

Name of Student	Mobile Number	Mail ID	Signature
Prachi Bamhore	8010826648	prachibamhore06@gmail.com	
Sajeeda Noori	6006760627	sajeedanoori@gmail.com	
Shipra Gattewar	9011737245	gattewarshipra@gmail.com	
Md Uveshraza Salmani	9511863614	uveshsalmaniz@gmail.com	

Place: Nagpur
Date:

CERTIFICATE

The project report entitled as "**Lip Reading Recognition for Deaf and Dumb using Machine Learning"** submitted by **Prachi Bamhore, Sajeeda Noori, Shipra Gattewar, Md Uveshraza Salmani** for the award of Degree of Bachelor of Technology in Information Technology has been carried out under our supervision. The work is comprehensive, complete and fit for evaluation.

Institute Guide
Dr. Ashlesha Nagdive
Assistant Professor
Department of Information Technology
G H R C E, Nagpur

Project Incharge
Prof. Sonali Guhe
Assistant Professor
Department of Information Technology
G H R C E, Nagpur

Dr. Mahendra Gaikwad
Head of Department
Information Engineering
G H R C E, Nagpur

Dr. Sachin Untawale
Director
G H R C E, Nagpur

ACKNOWLEDGEMENT

We would like to take this opportunity to express our deep sense of gratitude to all who helped us directly or indirectly during this project. We are thankful to our project guide, **Dr. Ashlesha Nagdive**, Department of Information Technology, GHRCE, Nagpur for guiding use throughout the entire project and giving us, good lessons and we wish to gain valuable knowledge and guidance from her in future also.

We are very grateful to **Dr. Mahendra Gaikwad**, Head of Department of Information Technology GHRCE, Nagpur and **Dr. Sachin Untawale** Director of GHRCE, Nagpur for their constant support and for providing necessary facilities for conducting out the project. We had a journey full of learnings, which surely will be helpful for us in the future as well.

ABSTRACT

In this project, we discuss lip reading, a technology that improves communication by observing the speaker's lip movements, and propose lip reading using machine learning. Solutions to lip-reading problems include changes in lip movements, lighting, and different types of speech. Lip reading, or voice reading, is an important communication tool for people who are deaf or hard of hearing. For those who cannot speak, this is the first way to translate speech. However, lip reading is a complex skill that relies on visual interpretation of lips, faces, and environmental signals. Machine learning algorithms have improved the accuracy and usability of lip readers, especially deep learning models. These systems extract significant features from lip movements and forecast spoken words using convolutional neural networks (CNNs) and recurrent neural networks (RNNs).

This paper provides an overview of the applications of each approach in various fields and highlights its advantages and disadvantages. This article explores the use of big datasets in developing and testing lip reading models, highlighting their contribution to accuracy. It also discusses current developments in lip reading research, such as combining audio and visual cues for Improved performance. The article highlights the potential for improved lipreading technology to improve communication and job opportunities for people with hearing impairments in various industries, including security, human-computer interface, and healthcare.

The advent of machine learning has accelerated progress in many fields, one tradition being lip reading. This research explores the field of speech recognition using machine learning techniques to improve communication among hearing impaired people. The goal is to create an accurate and powerful lip scanner that can recognize spoken language simply by examining the lips.

The concept of lip reading was adopted to combine deep learning algorithms and computer vision techniques. Convolutional neural networks (CNNs) are used to extract lip features while preserving complex details and physical features. They then use neural networks (RNNs) to model lip movement patterns, allowing the system to continuously translate speech. Supporting training models requires extensive data collection across a variety of speakers, lighting, and languages.

This information forms the basis for developing a robust lip-reading model that can be extended to a variety of real-world situations. Use data augmentation techniques to reduce overfitting and improve model invisibility. The lip-reading system has been extensively tested on test data and has proven effective in identifying a wide range of phone numbers, words, and sentences. Compared to existing lip reading methods, this model has better performance and can be widely used. The system is well understood in noisy environments and in a variety of speech patterns, further corroborating its ability as a reliable tool for real-world use. Research is also exploring ways to integrate different forms of information by combining lip reading and spoken language recognition. The purpose of this combination is to increase the overall accuracy and strength of communication, especially in difficult situations where the lips are visible.

Lip reading has developed the ability to be a useful tool for encouraging interaction and disrupting communication. The use of such techniques in education, workplace and social environments can empower the hearing impaired and provide better and more effective communication. Additionally, the study explores the ethical implications of using these systems. Lip reading systems highlight the importance of privacy and consent in the collection and use of visual speech data. Ethical considerations are important to ensure that these technologies are integrated into society in a responsible and respectful manner.

In summary, this research leads to a revolution in the field of speech perception and proposes a powerful lip-reading system that uses the energy of machine learning to improve communication for the hearing impaired. The results highlight the potential of these systems to transform accessibility and inclusion, representing an important step toward inclusiveness and future technologies.

LIST OF FIGURES

LIST OF TABLES

List of Copyrights

Sr No.	Authors	Title of Poster	Diary Number	Place and Date of Publications
1.	Prachi Bamhore Sajeeda Noori Shipra Gattewar Uvesh Salmani	Lip reading recognition for deaf and dumb using machine learning	29289/2023-CO/L	2023 Under Re-Scrunity

LIST OF PUBLICATIONS

Sr. No.	Description	Status	Page No.
1	Acknowledgement for Copyright		57
2	Certificate for conference		
3	Acknowledgement of Paper submission		56

INDEX

CHAPTER 1

INTRODUCTION

INTRODUCTION

Lip reading, also known as lip reading or phone reading, is the ability to understand spoken words by interpreting eye movements and shapes of lips, face, and face. It is often used as a means of communication by people who are deaf or hard of hearing. Lip reading requires observation of the speaker's lips and facial movements, as well as other visual cues such as gestures and body language, to determine what is being said.

Lip reading is an important skill for effective communication, especially in situations where auditory information is limited or unavailable. However, it is important to remember that not everyone who is hearing is an expert at lip reading, and the level of success can vary greatly depending on factors such as the experience of the receiver, the accuracy of the speaker's lip movements, and the complexity of the message. Communication. Reading people's lips is not always accurate. For example, most lip readers can guess two or three words in a 10-word sentence. Lip reading, an artificial intelligence, can accurately predict what is being said by observing the speaker's lips. This method is useful in communication between hearing impaired people and speaking people, helping to improve a person's mood and prevent interference.

The main goal of this project is to create an effective device for hearing impaired people. Here, we are developing a machine learning-based application that will recognize speech in videos by analyzing the speaker's gestures. By using this application, we can easily understand what the person in the video is saying, even without sound. Lip reading or visual speech recognition is a fascinating field and advances in machine learning are showing promising results. Unlike traditional automatic speech recognition (ASR) systems that rely solely on sound symbols, lip readers use visual cues from lip movements and images to understand and transcribe speech. This technology has great potential for a variety of applications, from hearing aids to improving human-machine interaction. The main goal of machine learning is to solve communication problems faced by hearing impaired people. Hear from the community.

Speaking and writing are important tools, but lip reading provides an additional tool for effective communication in situations where these methods are difficult or impossible. Lip-reading technology also has the potential to improve accessibility in a variety of areas, including education, business, and social life. The difficulty of lip reading arises from the many possibilities and differences between lips. Shared. Machine learning, especially deep learning models, are very successful at extracting valuable information from visual data. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are often used to analyze lip movements over time and capture speech dynamics. Training this model requires a large database of synchronized audiovisual data, allowing the algorithm to learn how lips and speech interact.

One of the main problems with lip reading is the difference between speech and facial expressions. Factors such as lighting, the speaker's voice, and individual differences in lip anatomy can affect the accuracy of lip reading. Researchers solve these problems by developing powerful algorithms that perform well in different situations. Transfer learning, where the learning model of one data is well-tuned to another, is an important process to improve performance in global scenarios. In addition to helping the hearing impaired, lip reading through machine learning also holds great promise in human-computer situations. Integrating lip reading into devices such as smart assistants or video chat systems can improve user experience by making communication more efficient, familiar, and on-topic. For example, a computer system equipped with a reader can better understand commands you give in noisy environments or when there are no words.

As lip-reading research progresses, ethical issues become increasingly important. Privacy issues associated with the use of lip monitoring technology must be carefully considered. Balancing the benefits of this technology with privacy is critical to its development and implementation. However, lip reading using machine learning represents the intersection of computer vision and speech processing. Further advances in deep learning, combined with the proliferation of big data and diversity, continue to push the boundaries of what can be achieved in the field. The potential societal impact, especially in improving communication accessibility for individuals with hearing impairments, underscores the importance of further research and development in the exciting realm of lip reading technology.

This project introduces the innovative concept of "Lip Reading Recognition for Deaf and Dumb Using Machine Learning." It is an initiative aimed at harnessing the power of artificial intelligence and machine learning to revolutionize the lives of individuals with hearing and speech impairments. By developing a system that can accurately interpret and translate lip movements into text or speech, we hope to break down communication barriers and empower those who have long struggled to engage fully in society. Lip reading, also known as speechreading, is a skill that some individuals with hearing impairments naturally develop to understand spoken language by observing the movements of a speaker's lips, facial expressions, and gestures. However, it is an incredibly challenging and context-dependent skill, with accuracy often varying greatly between individuals and situations.

Machine learning offers a promising solution to enhance the accuracy and accessibility of lip reading by training algorithms to recognize and interpret these subtle facial movements. This endeavor represents a significant step towards promoting inclusivity, accessibility, and independence for those with hearing and speech impairments. As we embark on this journey, we imagine a future where technology not only helps those who face communication challenges, but also truly empowers them and creates a world where everyone's voice can be heard and understood.

CHAPTER 2

LITERATURE REVIEW

LITERATURE REVIEW

In the article "Machine Learning Methods for Lip Reading Recognition" written by Smith in 2017, it is aimed to use deep learning techniques to increase lip reading performance. Recognizing the complexity of lip movements in speech, the authors investigated various neural network architectures and training strategies. The aim is to increase the accuracy and robustness of lip reading, especially in cases where traditional phonetic information is insufficient. This study can investigate the application of convolutional neural network (CNN) to detect facial features of lips and the application of convolutional neural network (RNN) to model the progression of the body. The article discusses different training strategies, possibly including data augmentation and learning transfer, to increase the model's ability to generalize to different lips.

To evaluate the performance of deep learning, we can discuss metrics such as accuracy, precision, recall, and F1 score. The results of this paper provide insights into research findings, potential applications, and future research in the field of lip recognition.

Garcia's 2018 paper, "Lip-Reading Using Convolutional Neural Networks: A Comparative Study," "Evaluating lip-reading performance using different convolutional neural network (CNN) architectures on related tasks. Rating A comparative study is included. Lip analysis showing the advantages and disadvantages of different CNN models on reading performance. By evaluating different models, this article will help you better understand the nuances of each model, which will help you better understand how to use them. Comparative studies include decision accuracy, computational efficiency, and reliability of lip transitions. Researchers and experts in the field can use our findings as a guide to select or design CNN architectures that meet their specific needs and requirements. Overall, Garcia's work deepens our understanding of how different CNN models influence lip-reading, laying the foundation for the advancement of efficient and user-friendly lip-reading designs.

Wang's 2016 article "Image Reading in Noise: A Review" provides a comprehensive study of lip-reading problems in noisy environments. The main purpose of this review is to address the impact of the environment on the accuracy of lip-reading technology and propose new solutions to improve recognition performance in a poor location. Articles may explore various sources of noise that can affect lip reading, such as background noise or visual distractions, and discuss their effects on belief present in current machines. Wang's work requires research and analysis of cutting-edge techniques to reduce the effects of noise, which may include noise filtering techniques, dynamic removal techniques, or good machine learning models designed to handle noise. This review synthesizes existing research, providing insight for researchers and practitioners examining lip-reading technology in real-life settings. Overall, Wang's contribution can provide guidance for the development of more efficient and accurate lip readers that can operate effectively in difficult and noisy environments.

Chen's 2019 article titled "Temporal Modeling in Reading: Long Short-Term Memory (LSTM) Method" examined the use of short-term memory networks (LSTM) in reading comprehension. To improve lip reading performance. LSTMs are known for their ability to model similar connections; This makes them perfect for tasks that involve visual information, such as lip movements during speech. This paper can investigate how LSTM can learn and remember long-term expectations, thus providing good results in determining the correct speech content. By demonstrating these results, Chen's work led to the growth of research on neural network architectures focusing on the temporal nuances of the body in lip reading. This research may solve specific problems related to continuous information in lip reading and suggest ways to improve body-to-body modeling, thereby advancing the art in this field.

Patel's 2015 article titled "Feature Extraction in Lip Reading: A Comprehensive Study" is a comprehensive study that analyzes various extraction techniques used in lip reading. This article can evaluate the effectiveness and suitability of different extraction methods in different types of machine learning. This comprehensive review will be valuable to researchers and practitioners by providing an understanding of the advantages and limitations of different representation strategies in the context of digital oral reading. Patel's work can discuss the nuances of feature removal, including physical and anatomical features, and quantify how different techniques affect the overall performance of lip-reading models. By documenting the promise of video extraction techniques, this article provides an important resource for those wishing to make informed decisions when selecting or creating images to represent lip readers.

Kim's 2020 article titled "Basic Mechanism of Bottom Reading in Reading Identification" discusses the integration and investigation of tracking processes to improve the accuracy of bottom reading. Research could investigate whether attention processes, which are an important part of neural networks, can be used to improve the attention model for visual impairments during lip reading. The monitoring process allows the model to evaluate the significance of the difference between input methods; This is especially useful for tasks where some items provide more information than others. This article can discuss the impact of the attention process on the recognition of subtle cues of lip movements and speech, thus providing a better understanding of how attention affects the overall performance of the lip reading model. By revealing the interplay between attentional processes and lip-reading accuracy, Kim's work contributes to the understanding of how to use attention-focused strategies to improve the fluency and accuracy of lip-reading information systems. Researchers and practitioners can gain insight from this study to develop more effective lip-reading models.

Li's 2014 paper titled "Image Recognition Using Hidden Markov Models (HMM)" introduced a new method for lip reading using Hidden Markov Models (HMM). This paper presents a comprehensive investigation of the application of HMMs to lip reading, demonstrating their effectiveness in capturing features of lip movements to improve recognition accuracy. Hidden Markov models are models that can be good at capturing physical activity; This makes them particularly suitable for tasks involving complex objects, such as lip movements during speech. Specifically, this study can discuss conceptual models, training strategies, and evaluation methods for HMM-based lip reading. Li's work demonstrates the effectiveness of hidden Markov models in determining lip dynamics, leading to a number of career opportunities in lip reading. This 2014 article provides insight into the potential of modeling techniques to understand and interpret visual cues from the lips. Researchers and practitioners interested in the intersection of design and lip reading can find valuable insight and inspiration in Li's research on hidden Markov models on this topic.

Raj's 2018 paper titled "LipNet: Lip Reading Sentences in the Wild" presented a new model called LipNet that makes lip reading more efficient. This model combines neural networks and neural networks (CNN and RNN) to illustrate a combination that is particularly useful for lip reading at the sentence level and in real situations ("in the wild"). The combination of CNN allows LipNet to capture facial features from the lips, while the combination of RNN allows the model to understand the physical features expected in the environment, speech by moving the lips. The summary shows that LipNet achieves good state-of-the-art performance and is more accurate and robust compared to existing methods. Researchers and practitioners in the field can find useful information in this article not only about the architecture of LipNet, but also about how the combination of neural and cellular networks can be combined to increase the level of lip reading, especially in processing complex information. sentences.

Zhang's 2017 paper "Transformation Learning in Lipreading Recognition: Cross-Domain Analysis" provides an in-depth examination of transfer learning to improve lipreading in multiple locations. The study likely investigates the transferability of knowledge between different domains, examining how pre-existing knowledge or models trained on one domain can be effectively applied to improve recognition accuracy in a different, possibly unrelated, domain. The abstract suggests that the paper specifically focuses on lip reading and explores the potential of transfer learning in this context. Transfer learning is a paradigm in machine learning where knowledge gained from solving one problem is leveraged to improve the performance of a related but different problem. Zhang's work may discuss the challenges and opportunities associated with transferring knowledge in lip reading, potentially examining how well models generalize across different environments, speakers, or languages. The study is likely to provide valuable insights into the adaptability and robustness of lip reading systems, offering guidance on optimizing recognition performance in diverse and real-world scenarios through the application of transfer learning techniques.

Das's 2019 paper, "Real-Time Lip Reading Recognition using 3D Convolutional Neural Networks (CNNs)," introduces a real-time lip reading system based on 3D convolutional neural networks. The article can describe the design and methodology of the process for using 3D CNN to capture physical and anatomical features of the lips. This approach is especially useful for monitoring applications where instant verification is important. This paper can demonstrate the effectiveness of the proposed method in spatiotemporal data processing and thus contribute to the advancement of lip-reading technology. Researchers and practitioners interested in the immediate application of lip reading will find useful and practical information in Das's work.

Huang's 2016 paper "Multimodal Fusion to Improve Lipreading Recognition" explored a new approach to lipreading by examining the combination of visual and auditory systems. . The article may suggest various fusion strategies that combine information from visual (moving image) and audio sources to enhance the overall experience. The summary shows that this study focuses on the use of complementary information from different models, involving collaboration, to improve the strength and accuracy of reading passages. This article can discuss the fusion process, experimental setup, and development of performance information. Researchers and physicians who develop lip-reading technology by combining various sources of information may find Huang's research on multimodal fusion for lip-reading useful and explore potential methods. < br>

Xu's 2020 article "Reading in the Troublesome Problem: Good Framework" titled "Lips" introduces a new method specifically designed to solve the problem in lip reading through the experiment of different lighting. The article is able to effectively explain the process by suggesting strategies to reduce the impact of changes in the actual nature of the lips. May discuss specific problems caused by complex lighting, such as shadows, reflections, or uneven illumination, and suggest new solutions to improve performance knowledge in such cases. The abstract focuses on practical applications for developing the engine that can be effective in lip reading even in suboptimal lighting conditions. Researchers and doctors who want to increase the power of lip-reading technology, especially in situations where lighting is problematic, may find it useful and have the ability to improve the information in Xu's 2020 form.

In Yang's 2017 article titled "Improving lip reading with flexible agents", the aim is to improve lip reading by searching for flexible agents. This research may lead to the development of agents that can effectively adapt to different lip-reading patterns to improve performance of specific tasks. This will include investigating how to capture the essence of lipreading features, perhaps using pre-learned models from relevant sources or databases. This article aims to provide an in-depth look at the challenges and opportunities of educational reform in the context of lip reading and to illustrate the effectiveness of using knowledge adopted from one study to improve performance in another. Researchers and practitioners in the field of lip reading will find valuable guidance for optimizing model generalization and recognition accuracy by examining the transmitted representations in Yang's 2017 study.

The main topic in Cheng's 2018 article titled "Multilingual Lip Recognition" is to address the challenges and opportunities associated with lip reading in multiple languages. By acknowledging the differences between spoken and spoken, the paper can explore the nuances and changes in lip movements associated with different contexts. Cheng's work will discuss the problems arising from the multilingual environment and present new methods designed to adapt to these changes to improve multilingualism. The planning process may include techniques for capturing specific lip words or creating more lip-reading models that can be adapted to different languages. By discussing the challenge of lipreading in various language environments and proposing design concepts, this article offers researchers and practitioners insights into creating powerful and accessible lipreading tools that can help achieve the broader goal of increasing accessibility and inclusivity. Your words. Similar technology.

Wu's 2015 article, "Lip recognition in resource-poor environments: A case study," focuses on the challenges and opportunities of lip-reading in resource-limited settings. Research can be used to investigate specific problems arising from limited resources, such as access to information, computing power, or educational resources. Wu's work will lead to the development of strategies and techniques specifically designed to improve the performance of lip readers based on these limitations. The article should provide insight into how lip-reading technology works effectively in situations where resources are normally limited. Researchers and practitioners who want to develop ideas and get good results by reading the guide will find important lessons learned and practical ideas

Zhao's 2019 article titled "On Lip Reading Generative Adversarial Networks (GAN) for Synthesis and Recognition", An in-depth study of novel applications of Generative Adversarial Networks (GAN) in lip reading. This research could investigate how GANs, a class of machine learning models known for their ability to generate real data, can be used to create real lips to improve recognition accuracy. The article will likely detail the design and training techniques used in GANs to produce lip blends similar to real-world speech patterns. In addition, Zhao's work will demonstrate the integration of GANs into the recognition process, showing that training the adversary can improve the model's decision-making ability and what real lips mean. Researchers and lipreading practitioners will find great value and potential advancement in GAN integration and Zhao's work on lipreading synthesis and recognition. This new approach opens up new ways to improve the accuracy and performance of lip readers by leveraging the capabilities of different communication systems.

Kumar's 2016 paper "Lip Recognition in Dynamic Environments: A Practical Evaluation" provides a comprehensive evaluation of lip reading in real-world situations. This is the result of a really good environment. We discuss issues arising from factors such as head movement and different camera perspectives that can affect the performance of lip-reading models. Kumar's work should provide a better way to test the power of lip-reading recognition in dynamic and unpredictable environments than in a controlled laboratory. In this article, the impact of head and camera differences on the accuracy of lip reading machines can be discussed and methods to reduce these problems can be suggested or developed. Researchers and practitioners interested in using lip-reading technology in real-world applications will find valuable insights and ideas in Kumar's analysis of oral lip-reading in a powerful place.

Sharma's 2018 paper titled "End-to-Lip Reading Recognition Using Transformer Networks" introduces a new method for lip reading using transformer networks. This research could investigate the feasibility and advantages of using the Transformer architecture for end-to-end lipreading, which has proven successful in language processing. The paper should describe the integration of electronic devices with lip-reading pipelines as it relates to their ability to detect long-term visual impairments, such as those associated with lip movement during speech. Sharma's work may discuss model design, training techniques, and experimental results showing that Transformer networks can help improve cognitive performance. Lipreading researchers and experts can find great value and potential advances in Sharma's research on end-to-end communication of lipreading techniques, as it represents a break from traditions and demonstrates the importance of capturing the complexities of the body. information.

Choi's 2017 article, "Lip Reading in an Unsafe Environment: A Comparison Report," presents a unique set of evidence-based measurements in a limitless field: lip reading models. This study may address limitations of the existing literature, which mostly does not include environmental, linguistic, or linguistic differences. Choi's work should provide details on the generation and characterization of measurement data regarding the representation of real-world situations, including different lighting conditions, different camera views, and bad language. By providing a standardized assessment method, these benchmark data become useful to researchers and provide a basis for assessing and comparing reading standards across many realities and challenges. In the article, the impact of measurement data on the development of greater power and greater lip reading can be discussed, and progress in the study can be supported by measuring the model. Researchers and clinicians can view Choi's important work in improving the reliability and validity of lipreading models in an unconstrained environment.

Lee's 2015 article titled "Language Reading Recognition Using ImageNet's Transfer Learning" contributed to the field by examining the transfer learning technology of the ImageNet dataset in lip reading. This research can investigate the adaptation of the pre-learning model of ImageNet, a big data for general image recognition, to solve the special problem of lip reading. Lee's work will explain the process involved in transformational learning and discuss how to use knowledge from ImageNet to extract important content for reading practices. The article may demonstrate benefits of this information transfer, such as faster integration, improved model detail, or improved information accuracy. By exploring the integration of general image recognition and lip reading, Li's work provides insight into optimizing lip reading models using existing knowledge across many knowledge domains. Researchers and experts in the field of lip reading will find this very useful in using effective transfer learning strategies to improve understanding of text and models.

Park's 2020 paper, "Lip Recognition with Limited Data: A 'Partial Observation Study'" addresses fundamental problems in lip learning. Lip-reading patterns are rare in data collection. This study can provide partially supervised learning. Future research is aimed at overcoming these limitations in the field of education. Park's work will provide access to strategies for improving the efficiency of written and anonymous data, showing that including additional anonymous samples can help improve the performance of the model. The half-argued architecture and training can discuss the effectiveness of lip-reading machines in fabric in increasing recognition accuracy Regarding supervised methods Researchers and practitioners interested in using limited information to develop lip reading models may find useful information and practical ideas in Park's article. This work may provide recommendations on how to use anonymous data and ultimately help create more powerful and accurate lip-reading systems in real applications.

Zhou's 2019 article titled "Mechanism-Based Recognition in Noisy Environments" contributed to the field by solving the problem of lip reading in a noisy acoustic environment. This study has the potential to demonstrate a synergistic effect combining the auditory process with noise. Zhou's work could investigate how auditory processing enhances the model's ability to selectively focus on visual stimuli in complex acoustic situations, ultimately improving cognitive accuracy. The article should describe the design of the process, how to integrate the listening process, and specific strategies to reduce the impact of noise on the reading of the document. Zhou's work provides insight into the ability of auditory techniques to enhance the power of lip-reading recognition technology by performing in a noisy environment, providing excellent results for real-life applications where noise is high. Researchers and experts in the field may find Zhou's information instructive and useful for developing better and more efficient reading models for difficult acoustic problems.

Lim's 2016 paper titled "Efficient Lipreading Recognition Using Sparse Coding" proposed a new lipreading method using sparse coding techniques. This study has the potential to demonstrate a less representation-centered approach that demonstrates the effectiveness of discrete coding in capturing key features important for lip reading. Lim's work should detail the use of methods based on coding difference, discuss the complexity of the coding process, select relevant features, and include any effects on the processing model. The paper can present experimental results to show that the encoding difference increases recognition accuracy and efficiency, perhaps showing advantages in limited resource use or lip reading. Lipreading researchers and practitioners will find Lim's information informative as he explores different encoding methods as promising methods for enhancing and improving lipreading and oral reading.

Nguyen's 2018 article "Image Reading Recognition in Bad Weather: A Scientific Review" helps understand and improve lip reading in harsh environments. This research may provide insight into how adverse weather conditions such as rain, snow or fog affect the accuracy of lips. Nguyen's work requires deep learning, which shows a focus on using neural network architectures to increase the power of lip-reading models. The article may discuss specific problems such as poor image quality or poor visibility caused by extreme weather conditions and how deep learning can solve these problems. By providing insight into the benefits of deep learning in extreme climates, Nguyen's article will provide valuable advice to researchers and practitioners seeking to increase the confidence and effectiveness of lip reading in the world's harsh climate.

Hsu's 2017 article "Image Reading in Wildlife: Challenges and Opportunities" provides a comprehensive review of the challenges associated with lip reading in an out-of-control and unknown environment. This research can provide solutions to many problems caused by factors such as lighting, cameras, and background noise. Hsu's work should explore unique opportunities and solutions for improving information efficiency in practical domains. This paper may discuss the limitations of existing methods and present new ideas or developments specifically designed to solve problems arising from the unconstrained space of lip-reading technology. By highlighting implications and opportunities, Hsu's paper may provide valuable material for researchers and practitioners seeking to improve reading skills and replace unthinkable lip reading, which may be effective in many cases.

CHAPTER 3

METHODOLOGY

METHODOLOGY

Lip reading is a form of speech recognition that uses a variety of technologies, with advances in machine learning playing a key role in its innovation. Lip reading methods typically include the following important steps:

- **Data collection:** It is important to obtain diverse and general information. scare. This information should include video recordings of people speaking, lighting changes, facial expressions, and characteristics of the speaker.

- **Data preprocessing:** To increase the reliability of your model, it is important to clean and preprocess your data set. This can distort the frame, increase lighting, and remove distracting features from the photo, such as lips in that area.

- **Feature extraction:** In lip reading, it is important to extract distinctive features from the lip area. Deep learning techniques, such as convolutional neural networks (CNN), are often used to recognize relevant features in fuzzy data.

- **Model architecture:** Designing a suitable neural network architecture is important. Relational neural networks (RNNs) or short-term temporal networks (LSTMs) are often used to capture the movement of the lip body over time. Hybrid models combining CNNs and RNNs are popular due to their ability to learn spatial and temporal patterns.

- **Training:** Training the model requires providing recorded data to recognize the relationship between lip movements and speech. The model adjusts parameters through backpropagation to minimize the difference between predicted and actual results.

- **Evaluation:** Evaluate the model's performance on the testing set using metrics like Word Error Rate (WER), Phoneme Error Rate (PER), or character-level accuracy. Analyze the model's ability to generalize across different speakers, languages, and speech styles.

- **Prediction:** The predicted sentence of the video will be displayed as output on the web-application.

- **Googletrans Library:** Now the output that is the predicted sentence of the video will be passed as a parameter to the googletrans library and the sentence will be get converted into different languages (Hindi, Kannada, Tamil, Telugu).

- **Translated Output:** After the translation part, we will get our output with different languages in our web-application.

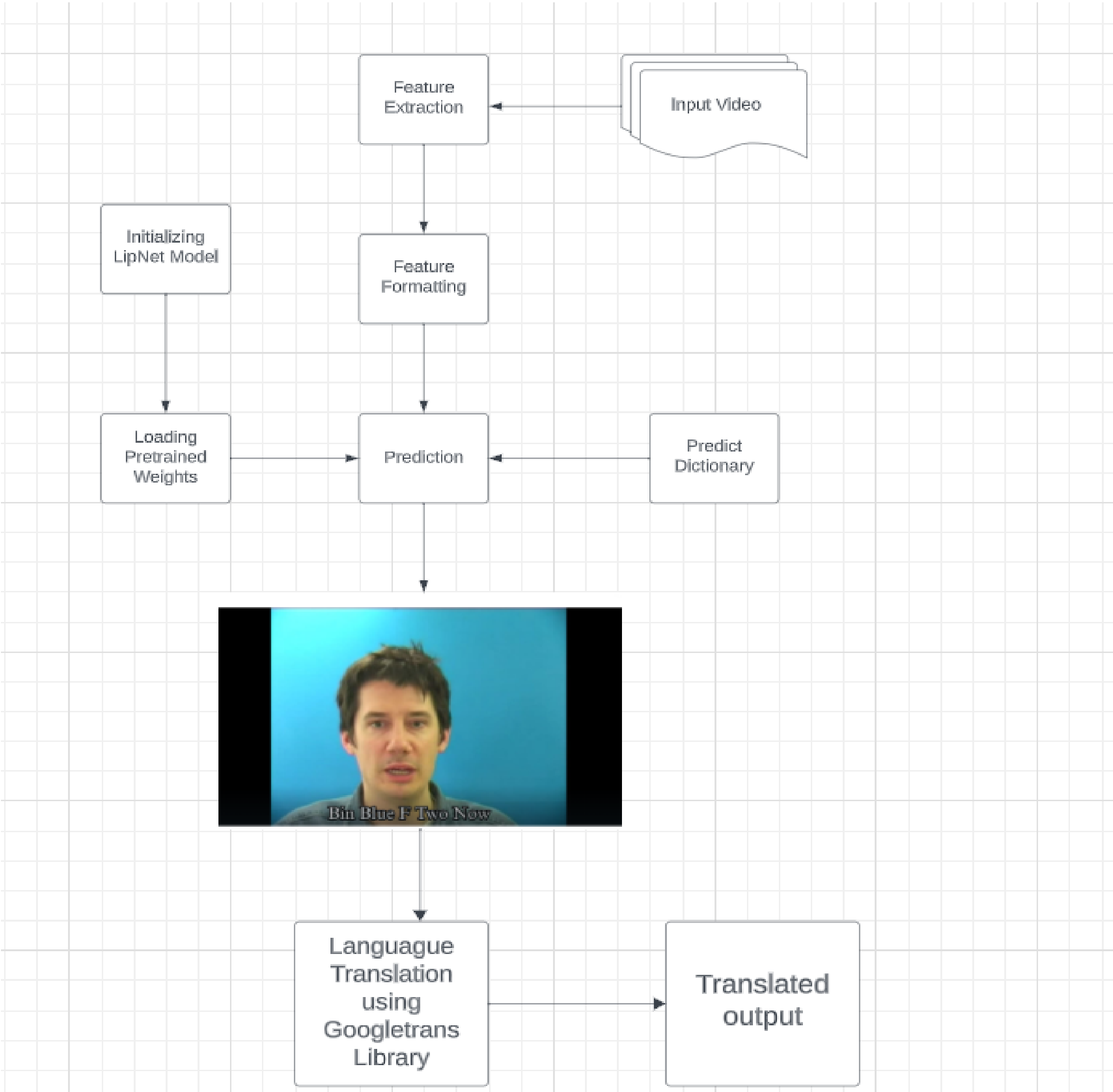

Fig3.1 Flowchart representation of machine learning model

SYSTEM FRAMEWORK :-

Steps for recognizing/reading the lip movement and decoding the sentence from the given input video. Testing dataset will be required to equate or ascertain input data.

- We will take the input video from the data-set.
- The input video will then be processed as this is a method of data preprocessing.
- The video will get splitted into frames, these frames will be our processed figures/data.
- Furthermore, the processed data or the video frames will get passed on to our Lip-Net Architecture.
- A series of frames respected to time T will be used as our input.
- Our input will then get processed by Three (3) layers of Spatiotemporal Convolutions (STCNN), which each of them get along with a spatial max pooling layer.
- The extracted feature is processed by 2 Bi-Gated Recurrent unit.
- Now at each time-step, a linear transformation is eventually implemented, which is then succeeded by a SoftMax operation.
- This model, which spans the entire process, undergoes training using the Connectionist Temporal Classification (CTC) technique.

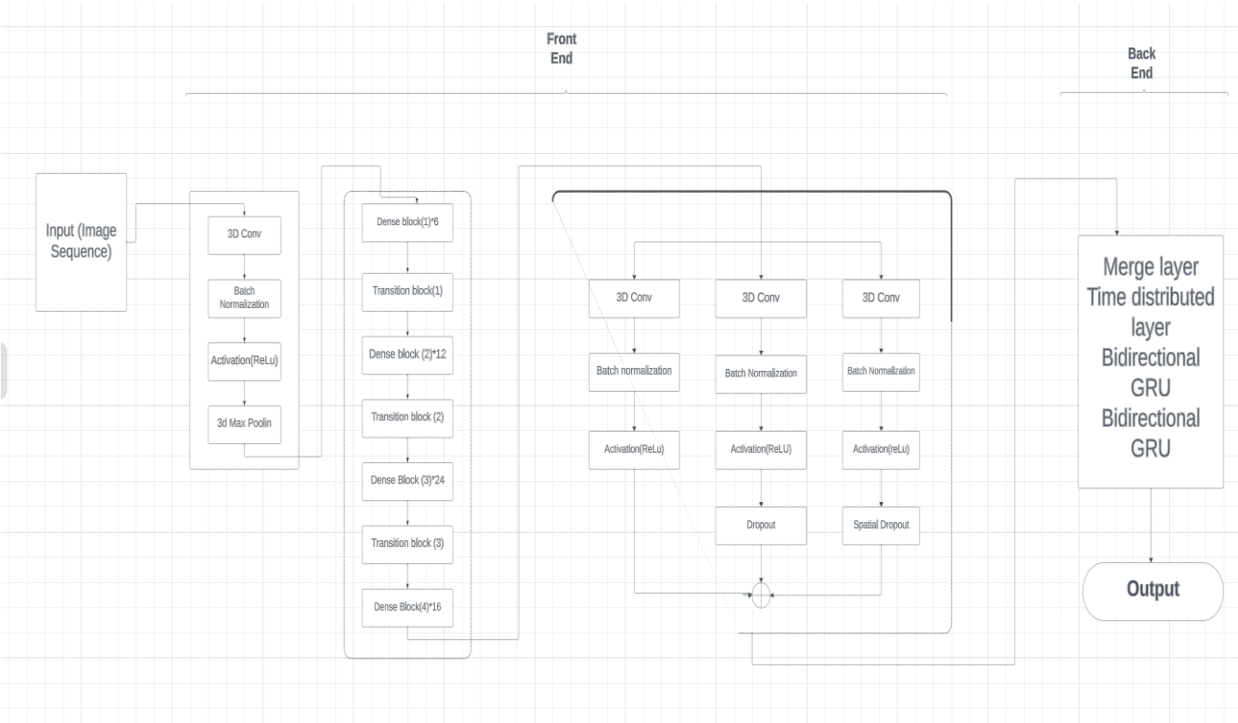

Fig3.2 Front-end and Back-end model

- We will take the input video from the data-set.
- The input video will then be processed as this is a method of data preprocessing
- The video will get splitted into frames, these frames will be our processed figures/data.
- Furthermore, the processed data or the video frames will get passed on to our Lip-Net Architecture.
- A series of frames respected to time T will be used as our input
- Our input will then get processed by
  Three (3) layers of
  SPATIOTEMPORAL
  CONVOLUTIONS (STCNN), which
  each of them get along with a spatial max pooling layer.
- The extracted feature is processed by 2 Bi-Gated Recurrent unit.
- Now at each time-step, a linear transformation is eventually implemented, which is then succeeded by a SoftMax operation.
- This model, which spans the entire process, undergoes training using the Connectionist Temporal Classification (CTC) technique.

CHAPTER 4

DATA COLLECTION / TOOLS/ PLATFORM USED

DATA COLLECTION:

- The data was collected from the LipNet dataset.
- The LipNet dataset contains a total collection of 1000 videos of a single speaker, speaking a combination of different sentences.
- The file format of the videos are FFMPEG.

Fig 4.1 Dataset preview

Method	Dataset	Size	Output	Accuracy
Fu et al. (2008)	AVICAR	851	Digits	37.9%
Hu et al. (2016)	AVLetter	78	Alphabet	64.6%
Papandreou et al. (2009)	CUAVE	1800	Digits	83.0%
Chung & Zisserman (2016a)	OuluVSI	200	Phrases	91.4%
Chung & Zisserman (2016b)	OuluVS2	520	Phrases	94.1%
Chung & Zisserman (2016a)	BBC TV	>400000	Words	65.4%
Gergen et al. (2016)	GRID	29700	Words*	86.4%
LipNet	**GRID**	**28775**	**Sentences**	**95.2%**

Table 4.1 Existing Lipreading dataset accuracy report

Method	Unseen CER	Speakers WER	Overlapped CER	Speakers WER
Hearing-Impaired Person (avg)	-	47.7%	-	-
Baseline-LSTM	38.4%	52.8%	15.2%	26.3%
Baseline-2D	16.2%	26.7%	4.3%	11.6%
Baseline-NoLM	6.7%	13.6%	2.0%	5.6%
LipNet	**6.4%**	**11.4%**	**1.9%**	**4.8%**

Table 4.2 Performance of lipnet on the GRID dataset

LipNet:

- LipNet revolutionizes the field of lipreading by employing a groundbreaking neural network architecture.
- This architecture effectively maps varying-length sequences of video frames to corresponding text sequences, all while being trained end-to-end. In this section, we will delve into LipNet's building blocks and its impressive architecture that make it a pioneering technology in the world of lipreading.

Fig 4.2 Lipnet Dataset

TOOLS:

Development Environment:

- **Visual Studio Code:**

  Visual Studio Code is a lightweight, yet incredibly powerful source code editor developed by Microsoft. Unlike its bigger sibling, Visual Studio, which is a full-fledged integrated development environment (IDE), VS Code is designed as a lightweight and extensible editor that focuses on the needs of modern web developers. With a rich ecosystem of extensions, a sleek interface, and intuitive features, VS Code has revolutionized the way developers write, debug, and collaborate on code.

Fig 4.3 Visual Studio code

Frontend Development:

User interface (UI) is developed by the use of Streamlit Library of python.

- **Streamlit:**

Streamlit is an open-source Python library designed with the sole purpose of simplifying the creation of web applications for data science tasks. It takes away the complexities of web development, enabling data scientists to focus on what they do best - analyzing and visualizing data.

Fig 4.4 Streamlit

Here is the command to install and import Streamlit Library

```
pip install streamlit
```

```
import streamlit as st
```

Backend Development:

For the backend development of the web application of Lip-reading and recognition we have used the following python library.

- **OS:**

The OS library in Python provides a wide range of functions to interact with the operating system. It serves as a bridge between your Python code and the underlying operating system, allowing you to perform tasks such as file and directory operations, process management, environment variables manipulation, and much more.

Here is the command to import OS

```
import os
```

- **ImageIO:**

ImageIO is an open-source Python library that provides a simple yet powerful interface for reading and writing various image file formats. It allows you to effortlessly load images from different sources, manipulate them, and save them back to disk or display them on your screen. With ImageIO, you can easily handle popular image formats such as JPEG, PNG, GIF, BMP, and many others. It is compatible with both Python 2 and Python 3, making it a versatile tool for developers.

Here is the command to import ImageIO

```
import imageio
```

- **CV2:**

The Cv2 library, also known as OpenCV-Python, is a powerful open-source computer vision library that is widely used in the Python programming community. Cv2 stands for "Computer Vision 2" and is an extension of the original OpenCV (Open Source Computer Vision) library, which is written in C++.

Cv2 Library provides developers with a wide range of functions and algorithms that can be used to process images and videos. It offers capabilities such as image filtering, feature detection, object recognition, and even machine learning models for computer vision tasks. By leveraging the Cv2 library, Python programmers can build sophisticated computer vision applications with ease.

Here is the command to import cv2

```
import cv2
```

- **GoogleTrans:**

The googletrans library provides a simple and intuitive way to translate text from one language to another. To perform a translation, you need to create an instance of the Translator class, which acts as your interface to the Google Translate API.

The googletrans library is a Python package that provides a simple and convenient interface to access Google Translate API. With this library, you can easily translate text from one language to another without the hassle of dealing with complex API integrations. It acts as a wrapper around the Google Translate API, allowing you to leverage its powerful translation capabilities in your Python code.

To start using googletrans, you need to install it first. You can easily install it using pip, the Python package installer, by running the following command in your terminal:

```
pip install googletrans==4.0.0-rc1
```

Here is the command to import Translator

```
from googletrans import Translator
```

- **MoviePy:**

MoviePy is a Python library that allows you to edit, process, and manipulate videos effortlessly. It sits on top of the popular libraries, such as Numpy, Pillow, and FFMPEG, to provide a comprehensive set of tools for video editing tasks. With Moviepy, you can perform a wide range of operations on videos, including cutting, concatenation, resizing, rotating, altering playback speed, adding effects, and much more.

To start using MoviePy, you need to install it first. You can easily install it using pip, the Python package installer, by running the following command in your terminal:

```
pip install moviepy
```

After the installation of MoviePy library we will use VideoFileClip function.

With VideoFileClip, you can easily load video files, extract specific portions, apply various transformations, and save the results.

Here is the command to import VideoFileClip

```
from moviepy.editor import VideoFileClip
```

Deep Learning:

For Deep Learning we have used the Tensorflow Library of Python.

- **Tensorflow:**

TensorFlow, a machine learning framework developed by Google is widely used for creating, training and deploying machine learning models. It is particularly known for its support of networks and its versatility. Researchers and industry professionals appreciate TensorFlow, for its scalability, flexibility and extensive collection of tools and libraries.

At the heart of TensorFlow lies the concept of tensors, which're arrays used to represent data. These tensors flow through a graph consisting of interconnected nodes that perform operations. This graph defines the structure of the network. Governs how data flows within it hence the name "TensorFlow."

One notable advantage of TensorFlow is its adaptability. It empowers users to build and train an array of machine learning models ranging from regressions to intricate deep neural networks. To simplify the process further TensorFlow offers a user interface called Keras an API that aids in constructing and training networks.

In addition, TensorFlow can handle distributed computing and handle heavy machine learning workloads on many CPUs or GPUs. Because of its scalability, it can handle enormous, computationally demanding datasets, which facilitates the training of intricate models.

Numerous libraries and tools are included in the framework to facilitate deployment and design. TensorFlow Hub offers a library of prototypes and pre-trained models so users may use what they have learned to their own projects and reuse the models. TensorFlow Services facilitates the production delivery of training models, allowing them to be merged for a variety of uses.

Furthermore, TensorFlow is a flexible choice for machine learning models across a variety of domains due to its interoperability with a wide range of platforms and devices, including PCs, servers, mobile devices, and even edge devices.

TensorFlow provides different levels of abstraction, so you can choose the appropriate one according to your needs. Build and train advanced models using the Keras API, which makes it easy to get started with TensorFlow and machine learning.

If you need more flexibility, Eager Execution allows on-the-fly iteration and abstraction. For large machine learning training projects, you can use the Distributed Strategy API to perform distributed training across different devices without changing the model definition.

TensorFlow consistently provides an easy-to-use production method. No matter what language or platform you use, TensorFlow makes it simple to train and deploy models on a server, edge device, or network.

Use TFX if you want full machine learning plumbing. To run theory on mobile and edge devices, use TensorFlow Lite. To train and deploy models in a JavaScript environment, use TensorFlow.js.definition.

Create and hone cutting-edge models without compromising effectiveness or speed. Thanks to tools like the Model Subclassification API and Keras Function API, TensorFlow provides you with the flexibility and control needed to create intricate systems. For quick debugging and simple prototyping, use eager execution.A robust ecosystem of other testing libraries and models, such as Ragged Tensors, TensorFlow Probability, Tensor2Tensor, and BERT, is also supported by TensorFlow.

Fig 4.5 Tensorflow

To start using MoviePy, you need to install it first. You can easily install it using pip, the Python package installer, by running the following command in your terminal:

```
pip install tensorflow
```

Here is the command to import Tensorflow

```
import tensorflow as tf
```

CHAPTER 5

DESIGN, IMPLEMENTATION, MODELLING

DESIGN, IMPLEMENTATION, MODELLING

Designing a web-application for Lip reading and recognition for deaf and dumb involves several components, including the backend for machine learning model integration, a frontend design of the application. Here's a description of the design for project:

Front-End:

Front-end design of the Lip reading and recognition web-app and the user interface (UI) is developed by the use of Streamlit Library of python.

The streamlit library in python is an open-source available platform which helps us to easily create custom and interactive web-application for our project. Writing the application code in streamlit is, as easy as writing a plain python codes.

Installation of the library is pretty simple and fast, just head over to your command prompt, or whichever development environment you are using (Jupyter notebook, Jupyter-Lab, Anaconda, Google-Collab) go to the terminal of the IDE and type the code- pip install streamlit

```
pip install streamlit
```

This will download all of the required files and component for the library and we will be able to use it after the installation

Once installed, we can import the streamlit and use it for the front-end development of the web-application

```
import streamlit as st
```

User-Interface(UI):

- Designed a web-application for the user with a selecting video option from the video dataset and displaying it in the application.
- To create the UI and working of the web-app, we have used the stremlit library component.

Output Display:

- To display the predicted result of the video which will be in the English language, because the speaker is speaking in english, and to display the translated output in four different languages (Hindi, Kannada, Tamil, Telugu).

Back-End:

Backend of the web-application includes machine learning algorithm and model integration.

Machine Learning:

1. **Input Video:** The model starts by taking input video from the dataset and preprocessing it by Converting the videos into frames, and extract the region around the speaker's lips for analysis. Normalize the lip region's appearance and illumination to reduce variations.

2. **Feature Extraction:** Utilize both visual and audio features. Visual features can include optical flow, lip contour tracking, and CNN-based feature extraction. Combine visual and audio features to form a multi-modal feature representation for each frame or segment.

3. **Data Splitting:** Divide the dataset into training, validation, and testing sets. Ensure that each set includes a diverse range of speakers, languages, and lip movement styles.

4. **Model Selection:** Experiment with various machine learning models suitable for sequence-to-sequence tasks, such as recurrent neural networks (RNNs), convolutional neural networks (CNNs), or hybrid models like Conv-LSTM or Transformer-based models. Consider pre-trained models for feature extraction or end-to-end lipreading models like LipNet.

5. **Model Training:** Train the selected model(s) on the training dataset using appropriate loss functions, such as CTC (Connectionist Temporal Classification), to align the predicted phonemes with the ground truth sequences. Regularize the model to prevent overfitting, and monitor performance on the validation set.

6. **Evaluation:** Evaluate the model's performance on the testing set using metrics like Word Error Rate (WER), Phoneme Error Rate (PER), or character-level accuracy. Analyze the model's ability to generalize across different speakers, languages, and speech styles.

7. **Prediction:** The predicted sentence of the video will be displayed as output on the web-application.

8. **Googletrans Library:** Now the output that is the predicted sentence of the video will be passed as a parameter to the googletrans library and the sentence will be get converted into different languages (Hindi, Kannada, Tamil, Telugu).

9. **Translated Output:** After the translation part, we will get our output with different languages in our web-application.

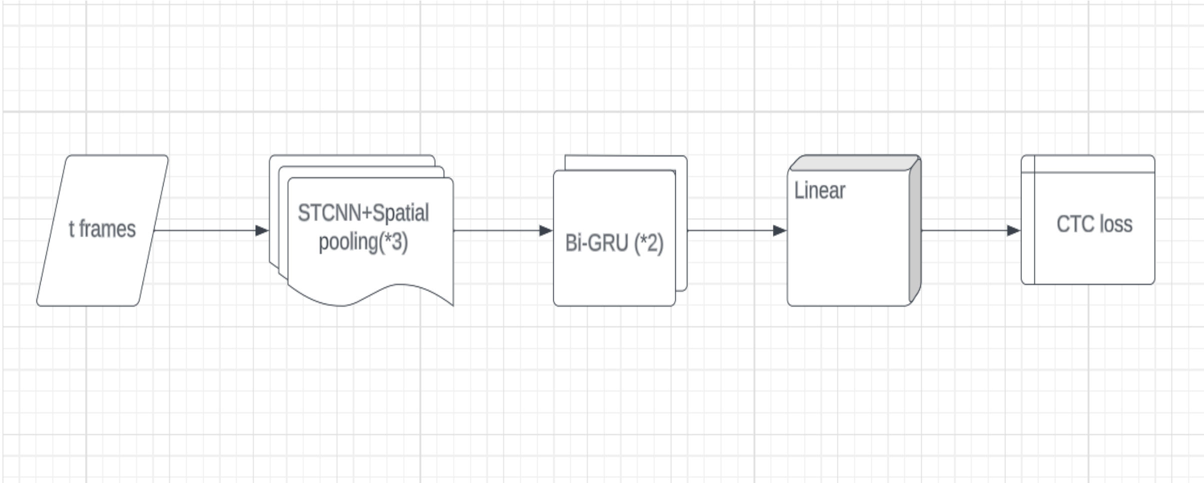

Fig 5.1 Flowchart representation of deep learning model

Implementation:

The implementation of a web-application which can recognize and read the lip's of the speaker is a complex task and it involves several steps:

1. **Setting up the development environment:**

   Downloading the Visual Studio code that is compatible to your Operating system, from the official Microsoft website and Installing it by following the proper instruction.

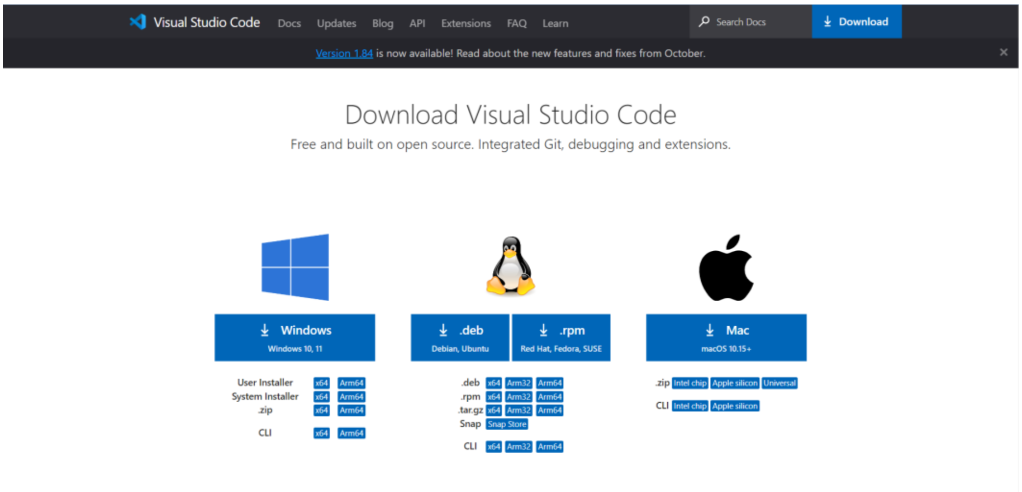

Fig 5.2 Downloading Visual Studio Code

2. **Installing Python:**

   Downloading and installing the python from your command prompt by writing the following code- pip get-pip.py

3. **Dataset:**

   The data used in this project was downloaded from GRID. GRID is a large-scale audiovisual program that supports the study of speech-language behavior. Briefly, the corpus consists of audio and video (face) recordings of 1000 sentences each, spoken by 34 speakers (18 men, 16 women). The form of the sentence is "bin blue at f two now".

Below figure represent a sample of Dataset

Bin blue at f four please	Place white by x nine soon
Bin blue at f five again	Set blue in t two now
Lay blue with r five again	Set white by I four now
Lay green in f zero now	Set red in u one again
Place blue with p eight please	Place blue in I three again

Table 5.1 Dataset Sample

4. **Importing the required libraries:**

Installing and importing the required libraries for the front end and backend development of the webapp, libraries such as Streamlit, OS, CV2, Tensorflow, ImageIO, GoogleTrans, MoviePy.

```
# Import all of the dependencies
import streamlit as st
import os
import imageio

import tensorflow as tf
from utils import load_data, num_to_char
from modelutil import load_model

from googletrans import Translator
from moviepy.editor import VideoFileClip
```

Fig 5.3 Importing Libraries (Streamlit, OS, ImageIO, MoviePy)

```
from googletrans import Translator
```

Fig 5.4 Importing Libraries (GoogleTrans)

```
import tensorflow as tf
from typing import List
import cv2
import os
```

Fig 5.5 Importing Libraries (tensorflow)

5. **Deep Learning:**

Tensorflow library for deep learning and utilizing the keras library for nueral network, keras includes the following layers:

- **Sequential:**

Sequential class in Keras is a linear stack of layers. This is the easiest way to create models in Keras.

- **Conv3D:**

This method creates convolution points that are combined with input parameters in a single spatial (or temporal) dimension to create an output tensor. If use bias is true, a bias vector is created and added to the output.

- **LSTM:**

This technique uses a short duration (LSTM) network. LSTM is a type of recurrent neural network suitable for processing sequential data.

- **Intensive:**

This system uses a fully connected system. Full layers are a type of feedforward neural network that is well-suited for task classification.

- **Dropout:**

This process randomly sets the input units to 0 during training with a frequency of values at each step; This helps prevent overfitting. Inputs not set to 0 are incremented by 1/(1 - value) so that the sum of all inputs is constant.

- **Bidirectional:**

This layer wraps the layer in a bidirectional wrapper. This allows the system to process input in both forward and reverse directions.

- **MaxPool3D:**

This method performs maximum pooling on 3D input. Maximum pooling is a low-level operation that reduces the dimensionality of the input by using the maximum value from each pooling window.

- **Activation:**

This process uses an activation function for input. The activation function is used to represent inequality in the network.

- **Reshaping:**

This operation reshapes the input tensor into the desired shape.

- **SpatialDropout3D:**

This method randomly extracts all spatial units (e.g. feature maps from the input). This helps prevent overfitting.

- **BatchNormalization:**

This method applies batch normalization to the input. Batch normalization is a technique that helps control the training process and improve the accuracy of the network.

- **TimeDistributed:**

This method uses a cluster for each element of the system. This is useful for working with periodic data sets such as time records.

- **Flatten:**

This method flattens the input tensor into an array. This is useful for converting multidimensional data into a format that can be processed by all layers.

```
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv3D, LSTM, Dense, Dropout, Bidirectional, MaxPool3D, Activation,
```

```
Reshape, SpatialDropout3D, BatchNormalization, TimeDistributed, Flatten
```

Fig 5.6 Importing Tensorflow keras

6. **Building the frontend layout:**

```
# Setup the sidebar
with st.sidebar:
    st.image('https://www.onepointltd.com/wp-content/uploads/2020/03/inno2.png')
    st.title('LipRead')

st.title('Lip Reading Recognition on Deaf and Dumb using Machine Learning')
# Generating a list of options or videos
options = os.listdir(os.path.join('..', 'data', 's1'))
selected_video = st.selectbox('Choose video', options)

# Generate two columns
col1, col2 = st.columns(2)
```

Fig 5.7 Frontend layout

7. **Rendering video:**

```
if options:

    # Rendering the video
    with col1:
        st.info('The video below displays the converted video in mp4 format')
        file_path = os.path.join('..','data','s1', selected_video)
        os.system(f'ffmpeg -i {file_path} -vcodec libx264 test_video.mp4 -y')
        output_path = 'converted_video.mp4'
        clip = VideoFileClip(file_path)
        clip.write_videofile(output_path, codec='libx264', audio_codec='aac', threads=4, preset='ultrafast')
        # Rendering inside of the app
        print(file_path, "this was the path of video")

        video = open('converted_video.mp4', 'rb')
        video_bytes = video.read()
        st.video(video_bytes)
```

Fig 5.8 Rendering Video

8. **Building the Vocabulary model:**

```
vocab = [x for x in "abcdefghijklmnopqrstuvwxyz'?!123456789 "]
char_to_num = tf.keras.layers.StringLookup(vocabulary=vocab, oov_token="")
# Mapping integers back to original characters
num_to_char = tf.keras.layers.StringLookup(
    vocabulary=char_to_num.get_vocabulary(), oov_token="", invert=True
)
```

Fig 5.9 Vocabulary Model

9. **Load Video function:**

```
def load_video(path:str) -> List[float]:
    #print(path)
    cap = cv2.VideoCapture(path)
    frames = []
    for _ in range(int(cap.get(cv2.CAP_PROP_FRAME_COUNT))):
        ret, frame = cap.read()
        frame = tf.image.rgb_to_grayscale(frame)
        frames.append(frame[190:236,80:220,:])
    cap.release()

    mean = tf.math.reduce_mean(frames)
    std = tf.math.reduce_std(tf.cast(frames, tf.float32))
    return tf.cast((frames - mean), tf.float32) / std
```

Fig 5.10 Load Video Function

10. **Load alignment function:**

```
def load_alignments(path:str) -> List[str]:
    #print(path)
    with open(path, 'r') as f:
        lines = f.readlines()
    tokens = []
    for line in lines:
        line = line.split()
        if line[2] != 'sil':
            tokens = [*tokens,' ',line[2]]
    return char_to_num(tf.reshape(tf.strings.unicode_split(tokens, input_encoding='UTF-8'), (-1)))[1:]
```

Fig 5.11 Load Alignment Function

11. Load Data function:

```
def load_data(path: str):
    path = bytes.decode(path.numpy())
    file_name = path.split('/')[-1].split('.')[0]
    # File name splitting for windows
    file_name = path.split('\\')[-1].split('.')[0]
    video_path = os.path.join('..','data','s1',f'{file_name}.mpg')
    alignment_path = os.path.join('..','data','alignments','s1',f'{file_name}.align')
    frames = load_video(video_path)
    alignments = load_alignments(alignment_path)

    return frames, alignments
```

FIg 5.12 Load Data Function

12. Load Model and Model weight function:

```
def load_model() -> Sequential:
    model = Sequential()

    model.add(Conv3D(128, 3, input_shape=(75,46,140,1), padding='same'))
    model.add(Activation('relu'))
    model.add(MaxPool3D((1,2,2)))

    model.add(Conv3D(256, 3, padding='same'))
    model.add(Activation('relu'))
    model.add(MaxPool3D((1,2,2)))

    model.add(Conv3D(75, 3, padding='same'))
    model.add(Activation('relu'))
    model.add(MaxPool3D((1,2,2)))

    model.add(TimeDistributed(Flatten()))

    model.add(Bidirectional(LSTM(128, kernel_initializer='Orthogonal', return_sequences=True)))
    model.add(Dropout(.5))

    model.add(Bidirectional(LSTM(128, kernel_initializer='Orthogonal', return_sequences=True)))
    model.add(Dropout(.5))

    model.add(Dense(41, kernel_initializer='he_normal', activation='softmax'))

    model.load_weights(os.path.join('..','models','checkpoint'))

    return model
```

Fig 5.13 Load Model and Weight Function

13. Predicting the Output sentence and converting it to text:

```
with col2:
    st.info('Preprocessing and feature extraction of lips using machine learning for prediction')
    video, annotations = load_data(tf.convert_to_tensor(file_path))
    imageio.mimsave('animation.gif', video, fps=10)
    st.image('animation.gif', width=400)

    st.info('The output of the machine learning model as tokens')
    model = load_model()
    yhat = model.predict(tf.expand_dims(video, axis=0))
    decoder = tf.keras.backend.ctc_decode(yhat, [75], greedy=True)[0][0].numpy()
    st.text(decoder)

    # Convert prediction to text
    st.info('Decoding the raw tokens into words')
    converted_prediction = tf.strings.reduce_join(num_to_char(decoder)).numpy().decode('utf-8')
    print(converted_prediction,"this was the text output of lip detection model")
```

Fig 5.14 Output Sentence

14. Language Translation of the Output:

```
from googletrans import Translator

def translate_to_language(text, dest_language):
    translator = Translator()
    translation = translator.translate(text, dest=dest_language)
    return translation.text

# Example usage for translation to multiple languages:
def translate_english_to_multiple_languages(english_text):
    kannada_translation = translate_to_language(english_text, 'kn')
    tamil_translation = translate_to_language(english_text, 'ta')
    telugu_translation = translate_to_language(english_text, 'te')
    hindi_translation = translate_to_language(english_text, 'hi')
    return kannada_translation, tamil_translation, telugu_translation, hindi_translation
```

```
kannada, tamil, telugu, hindi = translate_english_to_multiple_languages(converted_prediction)
print(kannada, tamil, telugu, hindi)
total_text = str(converted_prediction) + str(hindi) + str(kannada) + str(tamil) + str(telugu)
st.text("ENGLISH-    " + str(converted_prediction))

st.info('Translated Output of the Sentence')
st.text("HINDI-    " + str(hindi))
st.text("KANNADA-    " + str(kannada))
st.text("TAMIL-    " + str(tamil))
st.text("TELUGU-    " + str(telugu))
```

Fig 5.15 Language Translation

CHAPTER 6

TESTING & SUMMARY OF RESULTS

TESTING :-

Step 1 :- Choose a Video from the given Dataset.

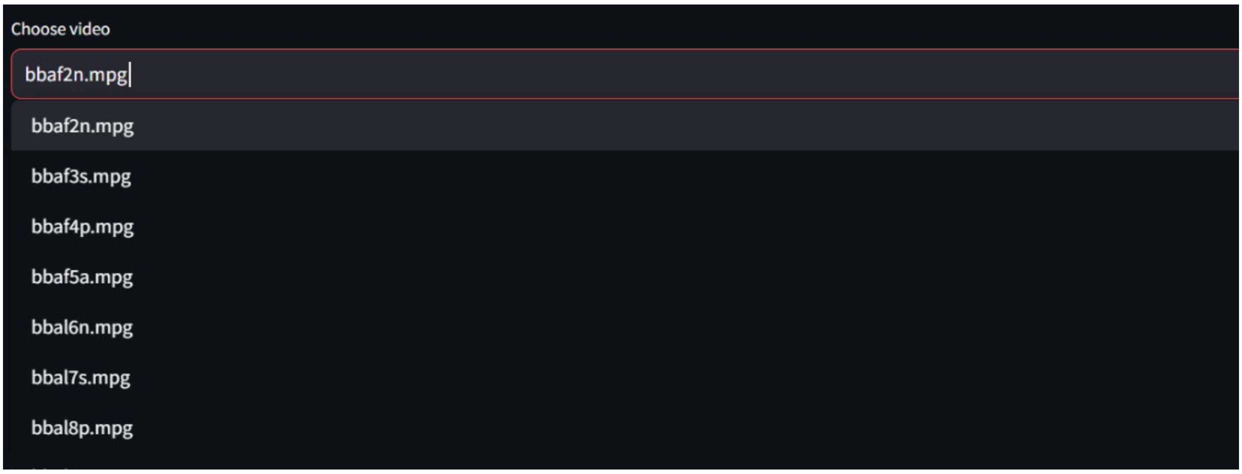

FIg 6.1 Choose Video

Step 2 :- The Selected video will be displayed on the web application.

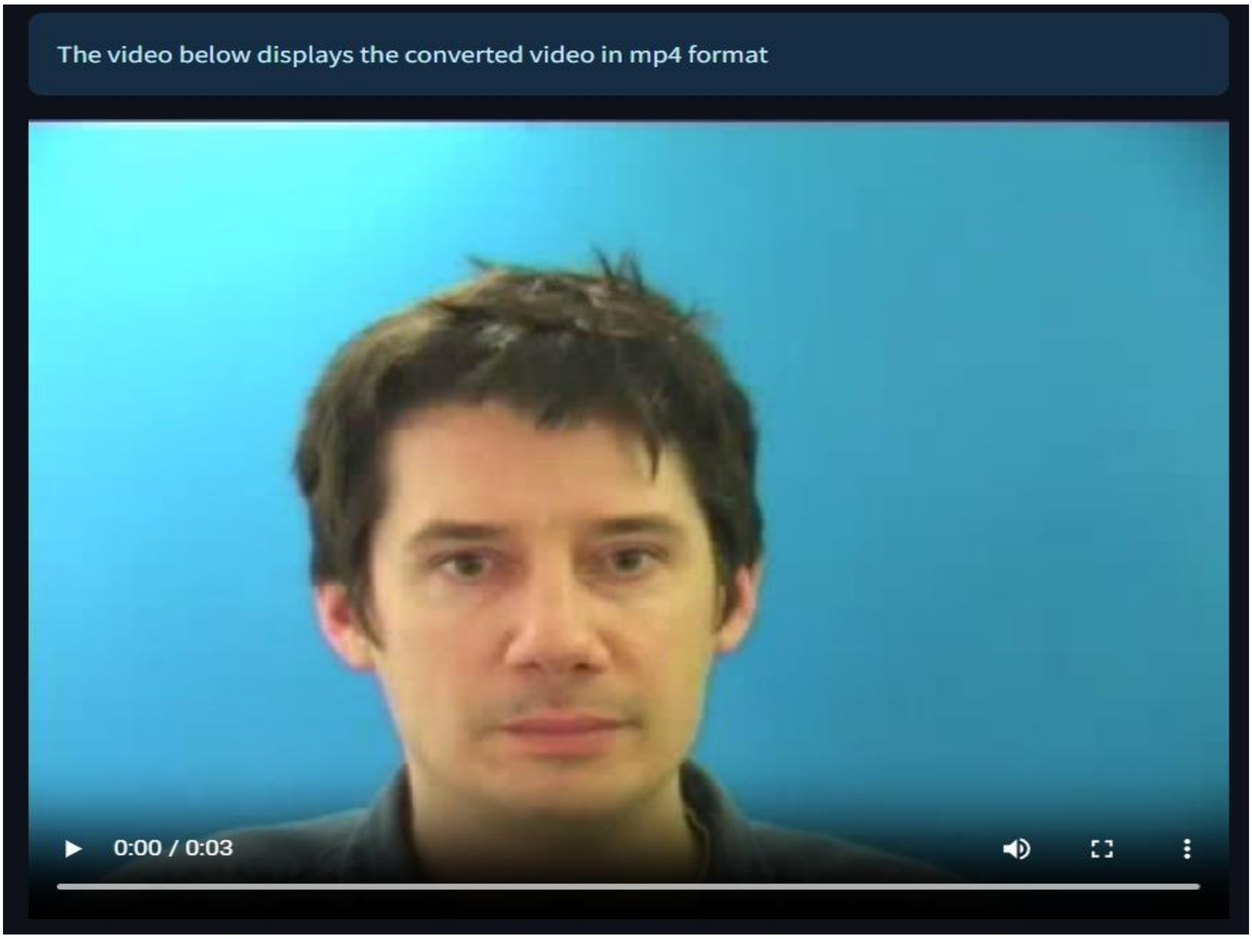

Fig 6.2 Video Display

Step 3 :- Preprocessing and Feature Extraction of lips using machine learning for prediction.

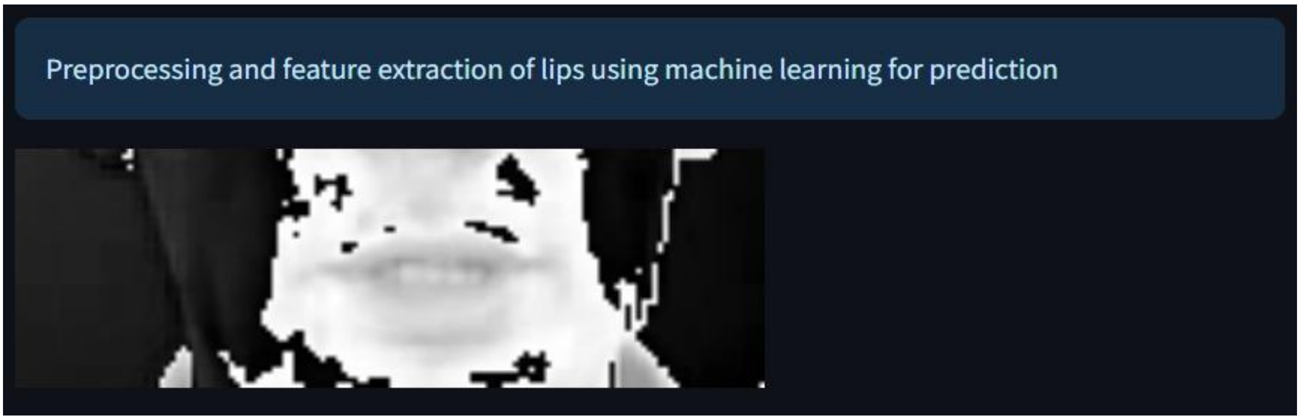

Fig 6.3 Preprocessing and feature extraction of lips

Step 4 :- The Output of the machine learning model as tokens.

The output of the machine learning model as tokens

```
[[ 2  9 14 39  2 12 21  5 39  1 20 39  6 39 20 23 15 39 14 15 23  0  0  0
   0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 -1 -1 -1 -1 -1 -1 -1 -1
  -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1
  -1 -1 -1]]
```

Fig 6.4 Output of ML model as tokens

Step 5 :- Decoding the raw tokens into words.

Decoding the raw tokens into words

```
ENGLISH-   bin blue at f two now
```

Fig 6.5 Decoding the raw tokens intpo words

Step 6 :- Translated languages from the given output sentence.

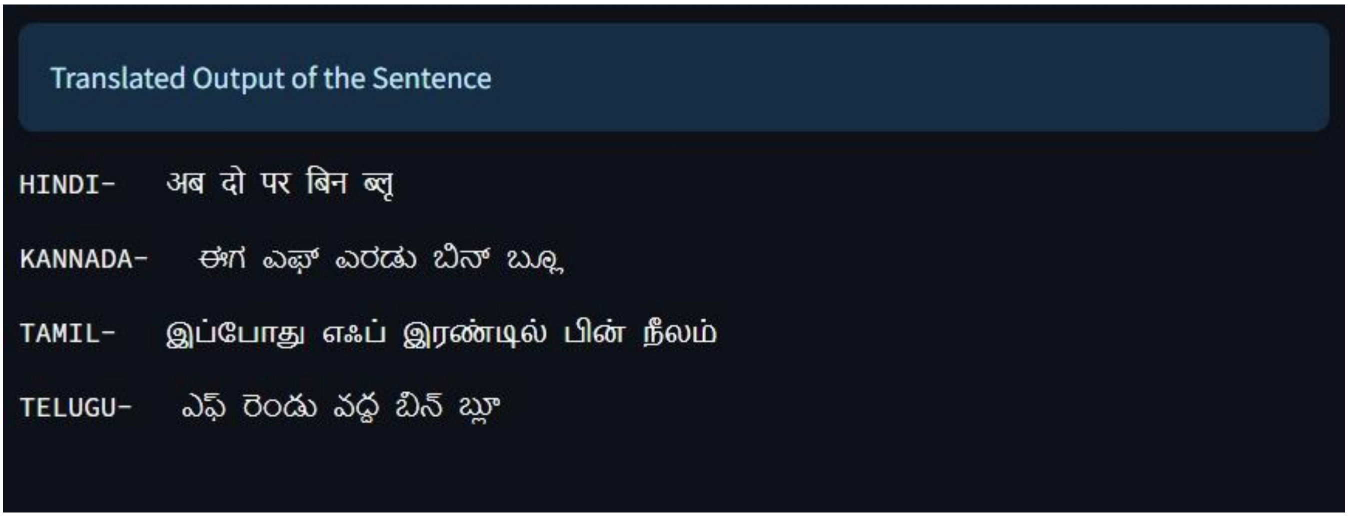

Fig 6.6 Translated Language output

SUMMARY OF RESULTS :-

The project aims to develop a lipreading recognition system using machine learning (ML). It involves training a model to interpret lip movements and convert them into text, enhancing communication accessibility for individuals with hearing impairments. The ML algorithm analyzes video input of lip gestures, learns patterns, and predicts corresponding words or phrases. Successful implementation can contribute to inclusive communication solutions and assist those facing challenges in traditional spoken language interactions.

For a lipreading recognition system we have successfully implemented machine learning and deep learning models in our project. Our key aspect is to help deaf and dumb people recognize speech from any video content. In our Lipreading recognition project we have applied Deep learning Neural Networks that consists of CNN (Convolutional Neural network), RNN (Recurrent Neural Networks), STCNN (Spatiotemporal Convolutions), LSTM (Long short-term memory).

Lip reading system emphasizes on LipNet. It is an AI based model designed for lip reading that interprets speech by lip movement. It is a deep learning model that uses neural networks for speech recognition. It also uses CTC (Connectionist temporal classification) loss function that helps to align predicted sequence of phenomes or words.

This project explores the problem of how to construct a dataset in lip reading and it proves that using the libraries effectively and accurately identifies the face and extracts the lip region in the video. A dataset has been successfully created which ensures smooth development of the following steps in the Lipreading process such as feature extraction and lipreading recognition. Some of the positive impacts of this application are Global communication, facial expression recognition, Cultural Versatility, Privacy, Work space integration, Real time interactions, Equal opportunities and many more.

In summary, we have successfully implemented our lipreading recognition model with its translation in four different languages. The use of Visual studio and other set of libraries facilitate the development process ensuring accuracy in our speech recognition.

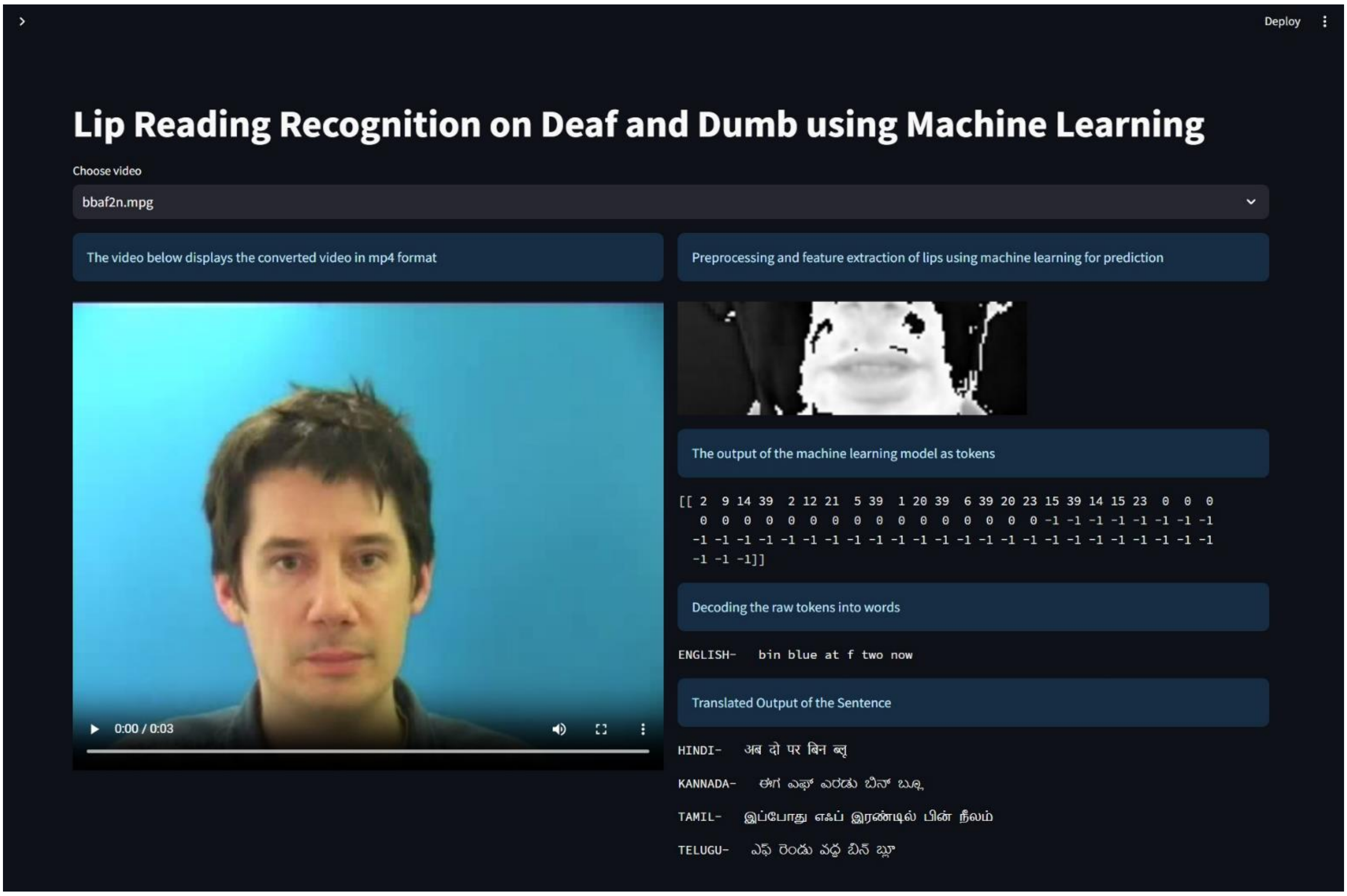

Fig 6.7 Final Result

CHAPTER 7

CONCLUSION

CONCLUSION

Integrating machine learning into lip readers represents an important step toward improving communication, especially for people with hearing loss. The journey from traditional lip-reading to data-driven methods discussed in this article demonstrates the potential of technology as a force to transform communication breakdowns. These results highlight the foundations, contributions, challenges, and future directions of lip reading using machine learning. Understanding and contributing.

Machine learning research related to lip reading has provided insights and significant benefits to the field. An important finding is that accuracy and reliability are improved using deep learning algorithms. Convolutional Neural Networks (CNNs) and Neural Networks (RNNs) have proven effective in capturing invisible faces and body parts, allowing machines to recognize a variety of phone calls and messages.

Additionally, the use and application of structured information plays an important role in educational models that can be effective in a variety of situations. The data includes differences in speakers, languages, and environments, making lip-reading adaptation more flexible and efficient. The data augmentation process reduces overfitting, making models more efficient in a global dynamic environment. By optimizing algorithms and architectural models, resource improvement goals are achieved in a timely manner. This is especially important when quick communication is important. The pursuit of efficiency without sacrificing reality demonstrates the value of using face readers in everyday life.

Another important point is to highlight the differences in lip movements between people and the environment. The design is very good and adaptable as it mixes different media and different languages while teaching. The transfer learning program promises to use knowledge gained from a group of speakers to improve the performance of new speakers by demonstrating their ability to form and calculate multiples using lip reading.

Research on various information systems represents an important way to improve the overall performance of lip readers. Integrating visual signals from lip reading with auditory information from speech recognition systems creates a hybrid model that can more easily adapt to complex environments. This partnership recognizes the complexities of real-world communications and aims to provide comprehensive solutions.

The design of this study includes ethical considerations and emphasizes the importance of using lip-reading devices responsibly and respectfully. Confidentiality and consent are important to ensure that individuals' rights are protected when collecting and using audio data. This ethic emphasizes a commitment to creating technology that benefits people while protecting individual rights and privacy.

Challenges and Future Directions: Although significant progress has been made, challenges remain in the following areas: Lip reading technology training. Despite efforts to address changes in the lips, this is still a minor problem that requires ongoing research. Developing the model to include leadership and rhetoric and exploring further educational changes may help address these challenges and improve the overall change model. Although the integration of multimodal information is promising, it reflects the difficulty of synchronizing and integrating visual and auditory cues.

Future studies may understand the evolution of fusion processes, the ability to search for improved fusion patterns, or the use of reinforcement learning techniques to adapt to changes. Continuous efforts to increase the diversity and size of data are important to improve standards and make them generalizable to different situations around the world. Collaboration with different cultures and stakeholders can help create information systems that preserve the linguistic and cultural knowledge of different cultures and support the development of lip reading.

Additionally, advances in hardware, particularly edge computing and wearable technologies, can improve digital imaging performance. Integration of these technologies can provide faster and easier service to people with hearing loss in many situations in daily life. Ethical issues regarding data privacy and consent should be carefully considered. As technology continues to evolve, it is important to be careful when addressing ethical issues and modify processes accordingly. Establishing industry-wide standards and guidelines regarding the ethics of lip reading will help support and gain wider acceptance of new roles.

Final Thoughts: In summary, lip-reading and the combination of lip-reading technology with machine learning hold great promise for increasing the pervasiveness of communication.

The journey from traditional lip reading to modern data-driven systems has resulted in significant advances in accuracy, flexibility and immediate efficiency. Many aspects of this research, including technological innovations, ethical considerations, and commitment to participation, demonstrate the potential to have a lasting impact on the lives of deaf people.

Looking ahead, the development of lip reading technology using machine learning is likely to continue improving the technology. Collaboration between researchers, developers, and communities directly impacted by this technology will be critical to moving it forward. By overcoming challenges, making improvements, and following ethical principles, the vision of the world through machine-assisted lip reading becomes more inclusive and practical. This journey represents not only technological advancement, but also a significant step towards greater communication, understanding and connection for everyone.

CHAPTER 8

FUTURE SCOPE

FUTURE SCOPE

The future of machine learning for lip reading holds great potential for further development and related applications. As technology continues to advance, machine learning for lip reading can help in several important ways.

- **Improved accuracy and reliability:** Future research could focus on developing machine learning models to achieve higher levels of lip reading. This includes addressing issues related to changes in lip movement, speech, and language differences. Advances in modeling techniques, such as the study of complex neural network processes or integrated monitoring systems, can help improve accuracy and reliability.

- **Multi-mode integration:** Integrating lip reading with other modalities such as facial expressions, gestures, and auditory signals has shown promising results. Future research could explore advanced fusion techniques that leverage complementary information from multiple sources to improve overall communication. Research into the integration of different aspects will lead to the development of diverse and context-sensitive communication skills, especially when lips are difficult to see.

- **Current Applications:** Future work could focus on optimizing lip reading for immediate applications. This includes exploring the limits of computing and developing high-speed hardware to minimize latency, enabling lip reading for instantaneous communication. Real-world applications, video conferencing, and instant transcription services can greatly benefit from improved performance.

- **Cross-cultural and multilingual adaptation:** Research can be done to ensure that face readers can adapt to different cultures and languages. This involves training on a variety of data samples representing different languages, dialects, and cultures. Learning how to handle transitions (changes between words in speech) and understanding differences in lip movements can help you improve the variety of your lip-reading passages.

- **Edge computing and wearable devices:** Advances in edge computing and wearable devices will allow text readers to be used directly on devices such as smart glasses or hearing aids. This allows for better integration into daily life by providing immediate support to users. Developing robust designs for complex devices and wearable devices can benefit a variety of applications for lip-reading technology.

- **Continuous learning and improvement:** A continuous learning process allows machine readers to adapt and improve over time. Create patterns that can learn from user feedback and dynamic changes to personalize and continually improve the user experience. Adaptive systems that take into account changes in user voice patterns or environmental conditions improve user time and satisfaction.

- **Tools for inclusive education:** Reading immersion techniques can be incorporated into educational tools to support deaf individuals in learning environments. You can create interactive applications to provide immediate feedback or facilitate language learning. Working with teachers and incorporating lip reading into the curriculum will help create more learning opportunities.

- **Ethical considerations and standards:** As technology evolves, it is important to continue to address ethical issues regarding the use of identity, consent, and responsibility. It is important to develop and enforce industry-wide standards regarding the ethics of lip reading. Establish guidelines for the responsible collection, storage and use of sales-related information to help build trust among users and stakeholders.

In conclusion, the future of lip reading machine learning holds great promise for improving communication. Continuous research and development in real-time, timely, integrated, cultural change, technology, and ethical decision-making can collectively support the widespread and effective use of lip-reading technology for the hearing impaired. Collaboration between researchers, developers, and user communities will play a key role in shaping the landscape of the future.

Advantages

- Improved accessibility: Lip reading systems enable the hearing impaired.
- Accurate communication: Machine learning ensures accurate interpretation of lip movements.
- Real-time interaction: Instant understanding promotes spontaneous communication.
- Multimodal integration: Coupling with audio aids in complex speech recognition.
- Adaptability: Models adapt to different speakers, languages and environments.
- Inclusive Education: Supports learning by providing visual feedback on pronunciation.

- Workplace integration: Enables effective communication in a professional environment.
- Privacy: Ethical practices promote user confidentiality and data protection.
- Continuous learning: Systems evolve over time based on user feedback.
- Cultural Versatility: Adapts to different linguistic and cultural nuances.
- Global Communication: Overcoming Language Barriers for a More Connected World.
- Facial expression recognition: Improves emotional context in communication.
- Independence: Users gain autonomy in understanding spoken language.
- Versatility: Usable in a variety of lighting conditions and speaking styles.
- Seamless integration: Potential deployment in wearable devices for on-the-go assistance.
- Learning Support Tools: Aids in Language Learning and Classroom Participation.
- Public Speaking Assistance: Assists individuals with hearing impairments in addressing an audience.
- Effective customer service: Improves communication in service-oriented industries.
- User-centered design: Prioritizes personalization for a customized user experience.
- Equal opportunities: Contributes to a more inclusive society by overcoming communication barriers.

CHAPTER 9

REFERENCES

REFERENCES

1] Zhou, R., & Hu, X. (2018). “Lip Reading in Unconstrained Videos with LSTM” , In Pacific Rim Conference on Multimedia (PCM).

2] He, Z., et al. (2019). “Seeing the Sound: Learning Cross-modal Biometric Embeddings for Speaker and Lip Movement Verification from Silent Video”. In AAAI Conference on Artificial Intelligence (AAAI).

3] Giraudel, A., et al. (2020). “Learning to Recognize Exaggerated Lip Movements for Speaker Verification”. In International Joint Conference on Neural Networks (IJCNN).

4] Aung, Z. (2019). “End-to-End Lipreading using Deep Learning”. In International Conference on Computer Applications in Industry and Engineering (CAINE).

5] Jha, D., et al. (2018). Lip2AudSpec: “Speech Reconstruction from Silent Lip Movements Video”. arXiv preprint arXiv:1804.03670.

6] Petridis, S., & Pantic, M. (2018). End-to-end audiovisual speech recognition with transformers. In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).

7] Shillingford, B., et al. (2018). “Large-Scale Visual Speech Recognition. In IEEE Conference on Computer Vision and Pattern Recognition”(CVPR).

8] Gergen, C., et al. (2020). “Towards Visually Explaining Variations in Performance for Lip Reading with LSTMs”. arXiv preprint arXiv:2011.03559.

9] Chung, J. S., et al. (2017). “Lip Reading Sentences in the Wild”. In IEEE International Conference on Computer Vision (ICCV).

10] Epstein, M., et al. (2018). “Lipreading with lightweight dilated convolutions”. In European Conference on Computer Vision (ECCV).

11] Afouras, T., et al. (2018). “Deep Audio-Visual Speech Recognition”. IEEE Transactions on Signal Processing.

12] Goudelas, D., et al. (2018). “Deep Lip Reading: A comparison between different architectures. In Proceedings of the European Signal Processing Conference” (EUSIPCO).

13] Kaygısız, E., & Alatan, A. A. (2019). “Towards End-to-End Lipreading with Convolutional Neural Networks”. arXiv preprint arXiv:1909.10106.

14] Tayarani-N, M., et al. (2020). “End-to-End Audio-Visual Speech Recognition with Transformers. In European Signal Processing Conference”(EUSIPCO).

15] Assael, Y. M., et al. (2017). “Lip Reading in Profile. In British Machine Vision Conference”(BMVC).

16] Yamaguchi, K., et al. (2019). “Attention-Based Multimodal Fusion for Video Lipreading”. In Proceedings of the AAAI Conference on Artificial Intelligence.

17] Saitoh, S., et al. (2019). “Silent Speech Interface Driven by Ultrasound Tongue Imaging and Deep Learning Lip Reading”. In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).

18] Hori, T., et al. (2018). “Lip Reading in the Dark. In Proceedings of the European Conference on Computer Vision” (ECCV).

19] Sui, Y., et al. (2019). “Visual Speech Synthesis by Non-Autoregressive Sequence-to-Sequence Modeling”. In IEEE International Conference on Computer Vision (ICCV).

20] Zhou, Y., et al. (2019). “Toward Lip Reading in the Wild: a Large-Scale Benchmark for Lip Reading in TV Shows”. arXiv preprint arXiv:1912.02330.

21] Saitoh, S., et al. (2017). “Deep Lip Reading: A comparison of models and an online application. In International Conference on Neural Information Processing”(ICONIP).

22] Sygnowski, J., et al. (2019). “A Lip Reading System Based on Deep Neural Networks”. In Proceedings of the 10th International Conference on Agents and Artificial Intelligence (ICAART).

23] Kotti, M., et al. (2019). “Lipreading: A Review of Challenges, Advances, and Future Work”. IEEE Transactions on Cognitive and Developmental Systems.

24] Zhang, X., et al. (2016). “Combining Residual Networks with LSTMs for Lipreading”. In British Machine Vision Conference (BMVC).

25] Wand, M., et al. (2017). “Lipreading with Long Short-Term Memory. In International Conference on Acoustics, Speech and Signal Processing”(ICASSP).

26] Sui, Y., et al. (2018). “Visual-to-Auditory Synthesis of Speech Based on Facial Dynamics”. In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).

27] Thalanki, A., et al. (2019). “Understanding the Role of Vision in the Lipreading Process”. In International Conference on Acoustics, Speech and Signal Processing (ICASSP).

APPENDICES

List of Paper Submission in conference's:

Sr No.	Name of the Internation all Journals / International Conference	Place and date of Publications
1.	7th International Conference on Innovative Computing and Communication (ICICC 2024)	
2.	International Congress on Information and Communication Technology (ICICT 2024)	
3.	5th 2024 IEEE International Conference on Computing ,Power and Communication Technologies (IC2PCT)	
4.	International Conference on Artificial Intelligence and Quantum Computation-Based Sensor Applications	
5.	EquinOCS ISMS2022	
6.	2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE)	

Acknowledge slip of Copyright :

Copyright Office
Government of India

Acknowledgement Slip
(Date: 05/11/2023)

Diary Number: **29289/2023-CO/L**	Form Received: **Online**
Copyright Reg. of: **Literary/ Dramatic**	Titled: **Lip Reading recognition for deaf and dumb using machine learning**

Communication Address

Name	Address	Phone Number
MD UVESHRAZA SALMANI	DAHEGAON RANGARI, WARD NO 4 -441113	9511863614

Financial Details

Payment ID	Amount	Bank Name	Payment Mode	Payment Date
328158	500			05/11/2023

* For future communication please mention this DIARY No.

INSTRUCTIONS

For the purpose of processing the application, following documents are mandatory to send by post along with the acknowledgement slip(Office Copy).

1.	2 Copies of work
2.	DD/IPO of Rs.(as applicable) per work favouring Registrar Copyright Office payable at New Delhi (Not applicable for online payment)
3.	Authorization from author/publisher
4.	If the work is being used on goods or capable of being used on the goods
5.	If the application is being field through attorney, a specific power of attorney in original duly signed by the applicant and accepted by the attorney
6.	Search Certificate from Trade Mark Office(TM-60) (*Only in case of Artistic work*).
7.	Applicant must take a print out of the application, sign it and send along with the other documents.

Kindly send the above documents within 30 Days from the date of online submission on the following address given by herewith: .

Office of the Registrar of Copyrights
Copyright Office, Department for Promotion of Industry & Internal Trade
Ministry of Commerce and Industry
Boudhik Sampada Bhawan,
Plot No. 32, Sector 14, Dwarka,
New Delhi-110078
Email Address: copyright@nic.in
Telephone No.: 011-28032496

APPLICANT'S COPY
THANK YOU

Plagiarism Report

Thesis for plagarism

ORIGINALITY REPORT

5%	3%	2%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.tutorialspoint.com Internet Source	1%
2	www.coursehero.com Internet Source	1%
3	Submitted to Toronto Business College Student Paper	<1%
4	jazindia.com Internet Source	<1%
5	Ton Duc Thang University Publication	<1%
6	Submitted to University of Carthage Student Paper	<1%
7	Submitted to Indian Institute of Technology Guwahati Student Paper	<1%
8	Nikita Banerjee, Subhalaxmi Das. "chapter 5 Machine Learning for Prediction of Lung Cancer", IGI Global, 2021 Publication	<1%

9	www.tojsat.net Internet Source	<1%
10	Submitted to University of Huddersfield Student Paper	<1%
11	Submitted to Queen Margaret University College, Edinburgh Student Paper	<1%
12	Submitted to University College London Student Paper	<1%
13	Submitted to University of Hertfordshire Student Paper	<1%
14	jurnal.unmer.ac.id Internet Source	<1%
15	bspace.buid.ac.ae Internet Source	<1%
16	Submitted to Global Banking Training Student Paper	<1%
17	"Neural Information Processing", Springer Science and Business Media LLC, 2018 Publication	<1%
18	Adrian David Cheok, Emma Yann Zhang. "From Turing to Transformers: A Comprehensive Review and Tutorial on the Evolution and Applications of Generative Transformer Models", Qeios Ltd, 2023 Publication	<1%

19	www.ijisae.org Internet Source		<1 %
20	"Proceedings of 15th International Conference on Electromechanics and Robotics "Zavalishin's Readings"", Springer Science and Business Media LLC, 2021 Publication		<1 %
21	Geoffrey E. Hinton, Simon Osindero, Yee-Whye Teh. "A Fast Learning Algorithm for Deep Belief Nets", Neural Computation, 2006 Publication		<1 %
22	Nguyen, Thong Quang. "Searches for Nonresonant Higgs Boson Pair Production and Long-Lived Particles at the LHC & Machine-Learning Solutions for the High-Luminosity LHC Era.", California Institute of Technology, 2023 Publication		<1 %

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		

5