



**Department of Electrical and Computer Engineering
North South University**

**Senior Design Project
Lip Reading App**

Jubair Mahmud Pulock ID# 1931111042

**Faculty Advisor:
Md. Shahriar Hussain
Senior Lecturer
ECE Department**

Spring, 2023

LETTER OF TRANSMITTAL

11 March, 2024

To,

Dr. Rajesh Palit

Chairman,

Department of Electrical and Computer Engineering

North South University, Dhaka

Subject: **Submission of Capstone Project Report on “Lip Reading App”**

Dear Sir,

With due respect, we would like to submit our **Capstone Project Report** on **“Lip Reading App”** as a part of our BSc program. This project was very much valuable to us as it helped us gain experience from practical field and apply in real life. We tried to the maximum competence to meet all the dimensions required from this report.

We will be highly obliged if you kindly receive this report and provide your valuable judgment. It would be our immense pleasure if you find this report useful and informative to have an apparent perspective on the issue.

Sincerely Yours,

.....

Jubair Mahmud Pulock

ECE Department

North South University, Bangladesh

APPROVAL

Jubair Mahmud Pulock (ID # 1931111042) from Electrical and Computer Engineering Department of North South University, have worked on the Senior Design Project titled “Lip Reading App” under the supervision of Md. Shahriar Hussain partial fulfillment of the requirement for the degree of Bachelors of Science in Engineering and has been accepted as satisfactory.

Supervisor’s Signature

.....

Md. Shahriar Hussain

Senior Lecturer

Department of Electrical and Computer Engineering
North South University
Dhaka, Bangladesh.

Chairman’s Signature

.....

Dr. Rajesh Palit

Professor

Department of Electrical and Computer Engineering
North South University
Dhaka, Bangladesh.

DECLARATION

This is to declare that this project is our original work. No part of this work has been submitted elsewhere partially or fully for the award of any other degree or diploma. All project related information will remain confidential and shall not be disclosed without the formal consent of the project supervisor. Relevant previous works presented in this report have been properly acknowledged and cited. The plagiarism policy, as stated by the supervisor, has been maintained.

Students' names & Signatures

1. Jubair Mahmud Pulock

ACKNOWLEDGEMENTS

The authors would like to express their heartfelt gratitude towards their project and research supervisor, Md. Shahriar Hussain, Senior Lecturer, Department of Electrical and Computer Engineering, North South University, Bangladesh, for his invaluable support, precise guidance and advice pertaining to the experiments, research and theoretical studies carried out during the course of the current project and also in the preparation of the current report. Furthermore, the authors would like to thank the Department of Electrical and Computer Engineering, North South University, Bangladesh for facilitating the research. The authors would also like to thank their loved ones for their countless sacrifices and continual support.

ABSTRACT

Lip Reading App

Effective communication is essential for promoting inclusivity and closing the gap between people who speak different languages in today's globally connected world. But hearing-impaired people frequently have trouble understanding spoken language and must rely significantly on visual signals to follow discussions. To solve this problem, we introduce a revolutionary Lip Reading App, a state-of-the-art mobile application that uses computer vision and artificial intelligence (AI) to enable real-time speech interpretation.

In order to effectively monitor and interpret lip movements and translate them into understandable text or audio output, the Lip Reading App makes use of the sophisticated capabilities of deep learning algorithms. The software recognizes and tracks facial landmarks using facial recognition technology, capturing the delicate movements and gestures involved in speech creation. The software can produce accurate transcriptions or audio renderings of the spoken content by processing these acquired visual cues and comparing them to a huge library of phonetic and linguistic patterns.

This project presents the development and training of a character-level sequence-to-sequence language model using recurrent neural networks (RNNs). The primary objective of the model is to generate text predictions given input sequences of characters. The model employs the 'binary_crossentropy' loss function, although it might be more suitable to use categorical cross-entropy for multi-class classification tasks. Training is performed over multiple epochs, with a learning rate of $1.0e-4$ (0.0001) and an average epoch time of 35026 seconds. The model's progress is evaluated based on the loss value, which steadily decreases with each epoch. By the end of 100 epochs, the average loss stands at 63.765. Despite the potential improvements in the loss function selection, the model exhibits promising capabilities in generating coherent and meaningful text predictions. Further fine-tuning and optimization could potentially enhance the performance and versatility of the language model for various text generation tasks.