



K. S. INSTITUTE OF TECHNOLOGY, BANGALORE-560109
DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
PROJECT PHASE 1 + SEMINAR (17CSP78)



PROJECT ABSTRACT SUBMISSION 7TH SEM A & B SEC 2020-21 (Odd Sem)

PROJECT TITLE: INTUITIVE PERCEPTION: SPEECH RECOGNITION USING MACHINE LEARNING

BatchNo.:2020_CSE_14 **Guide Name: Dr Swathi K**

Sl. No.	NAME	USN
1	Roopashree N	1KS17CS064
2	Sai Sneha S V	1KS17CS070
3	Spoorthi V	1KS17CS082

ABSTRACT

Lipreading is the task of understanding speech by analysing the movement of lips. Alternatively, it could be described as the process of decoding text from visual information generated by the speaker's mouth movement. The task of lipreading relies also on information provided by the context and knowledge of the language. It also known as visual speech recognition, and is a challenging task for humans, especially in the absence of context. Several seemingly identical lip movements can produce different words, therefore lipreading is an inherently ambiguous problem in the word level. Even professional lipreaders achieve low accuracy in word prediction for datasets with only a few words. Automated lipreading has been a topic of interest for many years.

A machine that can read lip movement has great practicality in numerous applications such as: automated lipreading of speakers with damaged vocal tracts, at crime scenes; it can be used to determine the conversation of the people at a crime scene during the time of crime caught on a surveillance camera ,biometric person identification, multi-talker simultaneous speech decoding, silent-movie processing and improvement of audio-visual speech recognition in general. The advancements in machine learning make automated lipreading possible.

Our proposed Automated Lip-Reading System includes the following:

- Surveys the state of the art in image and video recognition.
- Design and implement a real-time system capable of lip reading from video.
- Train the networks on a lipreading dataset, so that they can operate as lipreading systems.
- Evaluate the accuracy on simplified word classification task.
- Explore the possibility to recognize sentences in real-time.
- Build a prototype that present capabilities of deep learning algorithms.



K S I T
K. S. INSTITUTE OF TECHNOLOGY

K. S. INSTITUTE OF TECHNOLOGY, BANGALORE-560109
DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
PROJECT PHASE 1 + SEMINAR (17CSP78)

PROJECT ABSTRACT SUBMISSION 7TH SEM A & B SEC 2020-21 (Odd Sem)

System Requirements (H/W and S/W)

Minimum Hardware Requirements:

- 1.System: Intel Core i7 9750H
- 2.Speed: 4.5 GHz
- 3.Hard Disk: 20 GB
- 4.Monitor: LED/LCD Display
- 5.RAM: 8 GB
- 6.Keyboard: Standard Windows keyboard
- 7.Mouse: Optical mouse
- 8.GPU: NVIDIA GeForce GTX 1650

Software Requirements:

- 1.Operating System: Ubuntu/Windows 10 Home
- 2.Platform: Python
- 3.Frontend: Python interface
- 4.Toolkit: CUDA 10.0 and cuDNN
- 5.Packages: TensorFlow (1.0), Keras (2.0), OpenCV

Base Paper Submitted: (Yes/No): Yes

1. “**Large-Scale Visual Speech Recognition**”:Brendan Shillingford, Yannis Assael, Matthew W. Hoffman, Thomas Paine, Cían Hughes, Utsav Prabhu, Hank Liao, Hasim Sak, Kanishka Rao, Lorrayne Bennett, Marie Mulville, Ben Coppin, Ben Laurie, Andrew Senior, Nando de Freitas .
2. “**Lipreading by neural networks: Visual preprocessing, learning and sensory integration**.”Gregory J. Wolff,K. Venkatesh Prasad, David G. Stork,Marcus Hennecke

Note: Not for Student Use

ACCEPTED	REJECTED	RE SUBMIT
Reason for Rejection:		
Reason for Re Submit:		



K. S. INSTITUTE OF TECHNOLOGY, BANGALORE-560109
DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
PROJECT PHASE 1 + SEMINAR (17CSP78)



PROJECT ABSTRACT SUBMISSION 7TH SEM A & B SEC 2020-21 (Odd Sem)
Project Coordinator **HOD**