

SPEECH RECOGNITION USING DEEP LEARNING



Dr. Narendrababu Reddy G

Archers & Elevators Publishing House
ISBN:978-81-19385-08-9

Speech Recognition using Deep Learning

Dr. Narendrababu Reddy G,
Assistant. Professor, Dept of CSE,
G Narayanamma Institute of Technology and Science Shaikpet, Hyderabad.

PREFACE

Lip reading is a method of determining speech by looking at the movement of lips. Audio visual speech recognition (AVSR) is an approach that uses image processing abilities in lip reading to assist speech recognition systems. This system focuses on multimodal sensor architectures with deep learning for audio-visual speech recognition on the wild scenarios, particularly when the audio is corrupted by noise. It is a combination of both audio part and visual part, which implies integration of both lip reading and speech recognition processes working separately. In this research work, we propose new improvements for AVSR models by incorporating data augmentation techniques to generate more data samples for building the classification models. For the data augmentation techniques, we utilise a combination of conventional approaches (e.g., flips and rotations), as well as new approaches, such as generative adversarial networks (GANs). To validate the approaches, we used augmented data from well known datasets (LRS2-Lip Reading Sentences 2 and LRS3) in the training process and testing was performed using the original data. This framework enhances the performance of the AVSR framework in the wild for any datasets.

CONTENT

Sl.No.	Topic	Page No.
CHAPTER -1	Hand Gestures to Speech through CNN	
	1.1 Introduction	3-8
	1.2 Related Works on Multimodal-Sensor Architectures for Speech Recognition in Wild Environments	9-12
	1.3 Methodology	13-19
	1.4 Experiment and Results	20
	1.5 Conclusions and Future Enhancements	21-23
	1.6 References	
CHAPTER-2	Audio Visual Speech Recognition In Wild using Deep Learning	
	2.1 Introduction	26-27
	2.2 Background Work	28-33
	2.3 Proposed Method	34-39
	2.4 Datasets	40-44
	2.5 Results and Analysis	45-53
	2.6 Conclusion	54
	2.7 References	55-56
CHAPTER-3	Detection of stuttering using FluentNet model	
	3.1 Introduction	59-61
	3.2 Related Works	62-63
	3.3 Proposed Method	64-71
	3.4 Results Analysis	72-77
	3.5 Conclusion and Future Enhancements	78
	3.6 References	79