INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
SHORT ABSTRACT OF THESIS

Name of the Student	:	**Shikha Baghel**
Roll Number	:	**156102006**
Programme of Study	:	**Ph.D.**

Thesis Title: **Shouted, Overlapped and Competitive Speech Detection in Indian Television News Debates**

Name of Thesis Supervisor(s)	:	**Dr. Prithwijit Guha, Prof. S. R. Mahadeva Prasanna**
Thesis Submitted to the Department/ Center	:	**Electronics & Electrical Engineering**
Date of completion of Thesis Viva-Voce Exam	:	**03/09/2022**
Key words for description of Thesis Work	:	*Indian Broadcast News Debates, Shouted Speech, Overlapped Speech, Competitive Speech, Indian Broadcast News Debate Corpus, Excitation Source Features, Phase Features, Deep Learning*

SHORT ABSTRACT

Television (TV) news debates present expert opinions, analysis and discussions on contemporary events. These debates play a critical role in navigating public belief and understanding of socio-politically relevant topics. This encourages several agencies to analyze the TV news debate content for monitoring their influence. The availability of huge (and ever increasing) amount of news debate data calls for the necessity of automatic content analysis. TV news debates are generally argumentative in nature. Such arguments are often associated with the presence of *shouted*, *overlapped*, and *competitive* speech. In this context, the present thesis aims to detect these three speech categories in Indian TV news debates. The first contribution of this thesis is the development of an Indian Broadcast News Debate (IBND) corpus containing audio signals from 15 news debates (approximately 13 hours). A multi-level annotation procedure was followed to obtain the final annotations for the three targeted tasks of the thesis. The second contribution lies in the proposal of excitation source based Shouted Speech Detection (SSD). Both hand-crafted and learned features from excitation source-based representations are explored for SSD. An auto-encoder with Bi-GRU based architecture is used as classifier. The third aim of the thesis is to identify the overlapped speech in TV news debates. Phase-based representations of the speech signals are established as efficient features for Overlapped Speech Detection (OSD) using a CNN-LSTM based classifier. Finally, the shouted and overlapped speech classification network embeddings and their prediction scores are used as features to identify the competitive speech. It has been shown that the detection of competitive speech can be performed efficiently using high-level information of both shouted and overlapped speech.