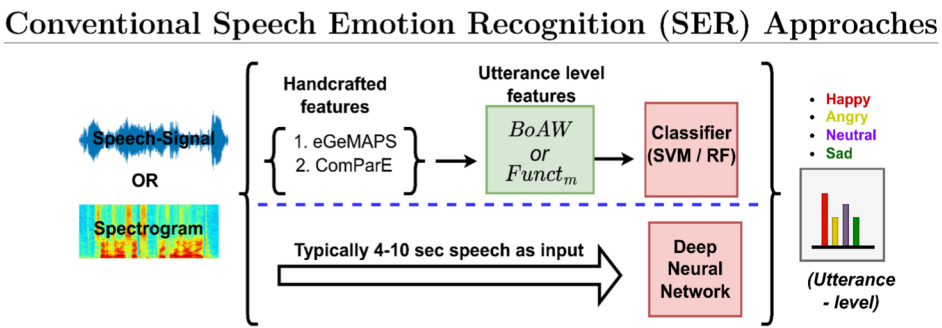
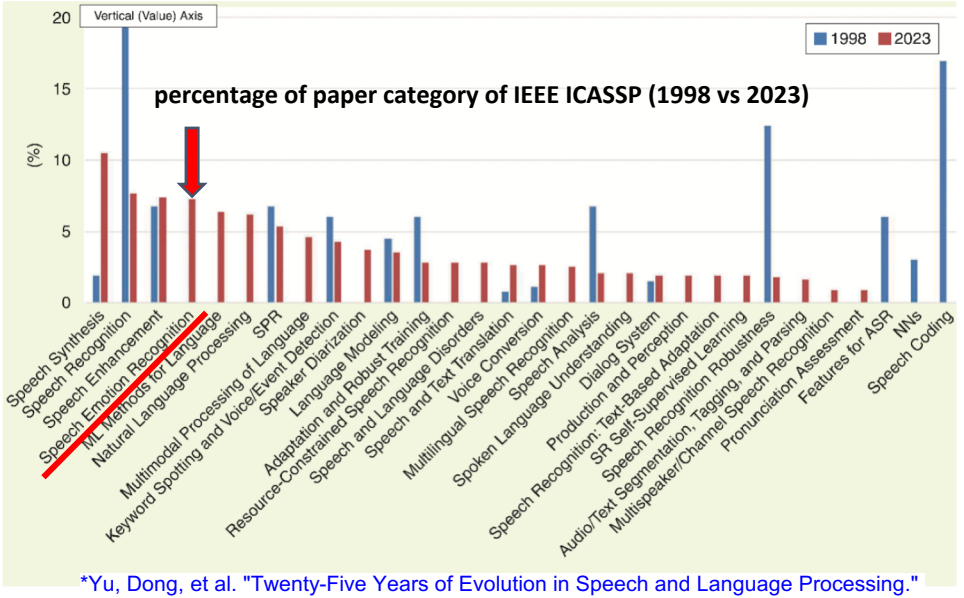
Towards Learning Emotion Information from Short Segments of Speech

Tilak Purohit

- For my PhD, working towards making Deep Brain Stimulation (DBS) for Parkinson's Disease patients Adaptive [DBS → aDBS]
  (Project: EMIL).
- The discipline of automatically recognizing human emotion and affective states from speech.

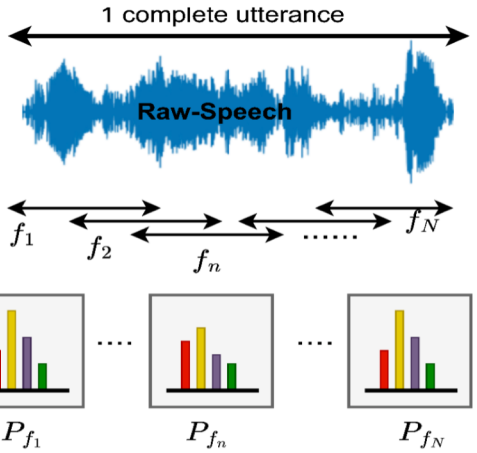


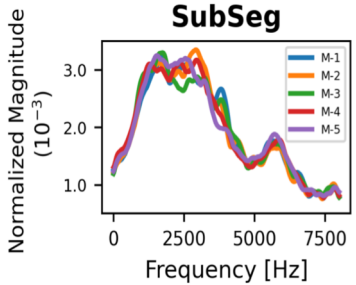


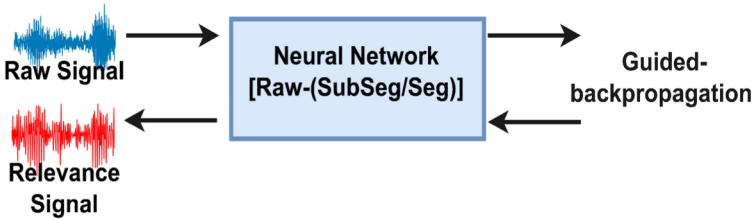
*Yu, Dong, et al. "Twenty-Five Years of Evolution in Speech and Language Processing." *IEEE Signal Processing Magazine* 40.5 (2023): 27-39.

Purohit, Tilak, et al. "Towards Learning Emotion Information from Short Segments of Speech." *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023.

Modelling Subword segment level information for SER







Utterance level modelling

ComParE$_{LLD\times F}$	RF	58.23
BoAW(ComParE$_{LLD}$)	RF	57.71
BoAW(eGeMAPS)	RF	55.90
BoAW(Wav2vec2)	RF	56.00

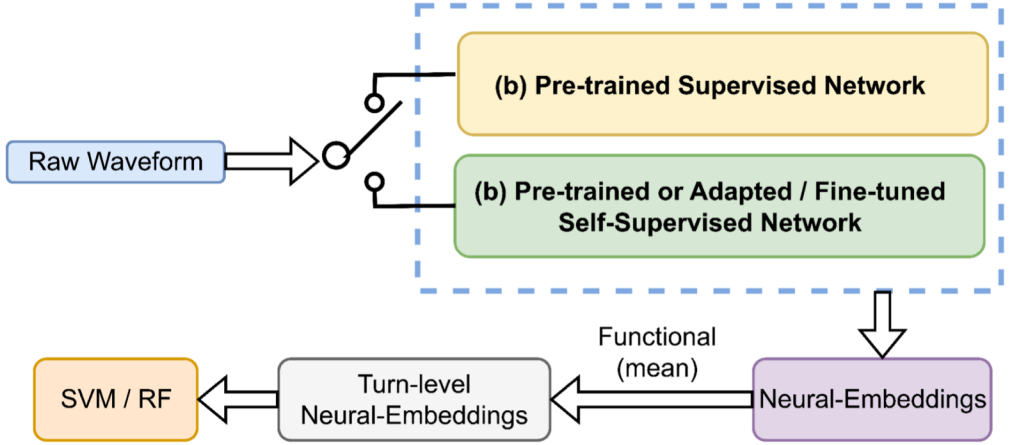
Short-segment level modelling

ComParE$_{LLD}$ [A]	MLP	45.88
eGeMAPS [A]	MLP	44.36
Raw SubSeg [B]	CNN-MLP	57.48

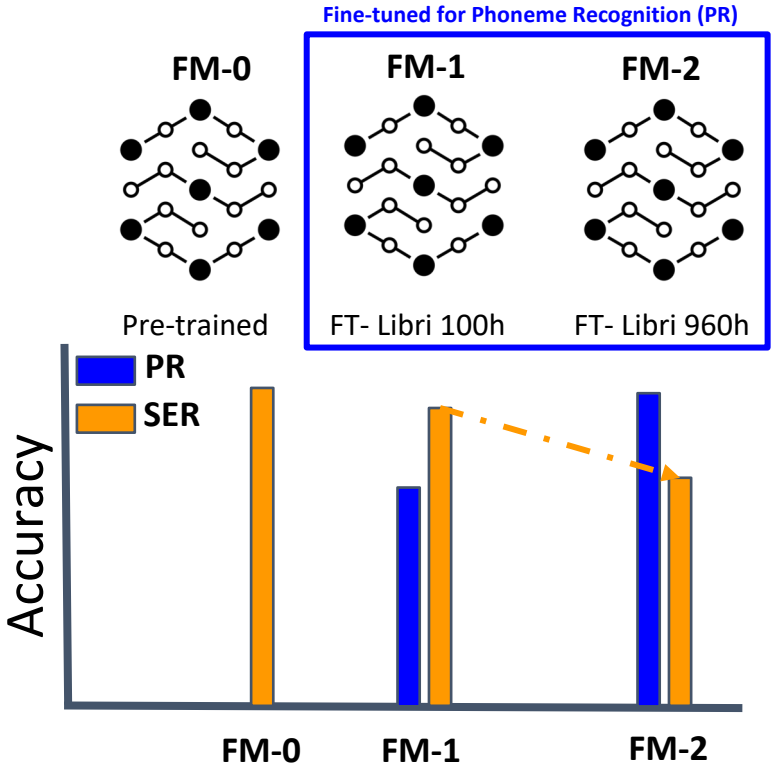
Raw-waveform	SubSeg-Rel
Spectral Flux	Cepstral
Spectral Harmonicity	Cepstral
Spectral (Auditory)	Cepstral
Spectral Harmonicity	Cepstral
Spectral (Auditory)	Cepstral
Spectral Flux	Cepstral
Spectral (Auditory)	Cepstral
Spectral Flux	Cepstral
Prosodic	Spectral Flux
Spectral Flux	Voice quality

Purohit, Tilak, et al. "Towards Learning Emotion Information from Short Segments of Speech." *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023.

Implicitly modeling Phonetic information via Foundation Models (FMs)



Purohit, Tilak, et al. "Implicit phonetic information modeling for speech emotion recognition" *INTERSPEECH* 2023.



- Phonetic embeddings yield improved SER performance compared to Handcrafted features.
- SER inverse relation with ASR.

Purohit, Tilak, et al. "Towards Learning Emotion Information from Short Segments of Speech." *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023.