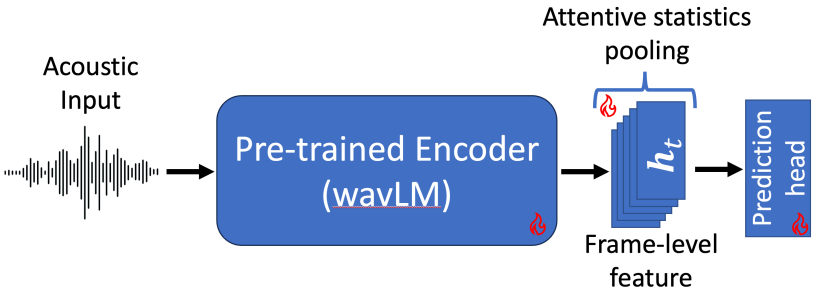
Model Overview



Training Settings:

1. Learning rate: 1e-5
2. Epochs: 20
3. Batch size: 32

Tasks Explored:

Task 1 - Categorical Emotions Model:

1. Weighted Loss Function: A weighted loss function is used to assign more weight to less frequent classes. For instance, using CrossEntropyLoss for a classification problem, we set the `weight` parameter to a tensor representing the inverse frequency of each class.
2. Loss: Categorical Cross Entropy loss with class weighting based on training set
3. Evaluation Metrics: F1-Micro, F1-Macro, Precision, Recall

Task 2 - Emotional Attributes Model:

1. Loss: Concordance Correlation Coefficient (CCC) loss
2. Evaluation Metrics: CCC

Results:

Task 1 – Categorical Emotions									
	Test 3					Development			
Model	F1-Mi	F1-Ma	Prec.	Rec.		F1-Mi	F1-Ma	Prec.	Rec.
WavLM	0.327	0.311	0.332	0.325		0.409	0.307	0.316	0.345

Task 2 - Emotional Attributes							
	Multi-Task Setup						
	Test 3				Development		
Model	Aro.	Val.	Dom.		Aro.	Val.	Dom.
WavLM	0.577	0.577	0.405		0.652	0.688	0.579
	Single-Task Setup						
	Test 3				Development		
	Aro.	Val.	Dom.		Aro.	Val.	Dom.
WavLM	0.566	0.607	0.424		0.651	0.709	0.584

References

-Chen, Sanyuan et al. “WavLM: Large-Scale Self-Supervised Pre-Training for Full Stack Speech Processing.” IEEE Journal of Selected Topics in Signal Processing 16 (2021): 1505-1518.

-Okabe, Koji et al. “Attentive Statistics Pooling for Deep Speaker Embedding.” ArXiv abs/1803.10963 (2018): n. pag.