Multimodal Emotion Recognition Results

Dataset: RAVDESS Audio-only (1440 samples)

Train: 979, Val: 173, Test: 288

Emotions: ['neutral', 'calm', 'happy', 'sad', 'angry', 'fearful', 'disgust', 'surprised']

Used Simulated Sentences: True

AudioCNN Test Accuracy: 0.6597

TextRNN Test Accuracy: 1.0000

EarlyFusion RNN Test Accuracy: 0.9549

LateFusion RNN Test Accuracy: 0.9479