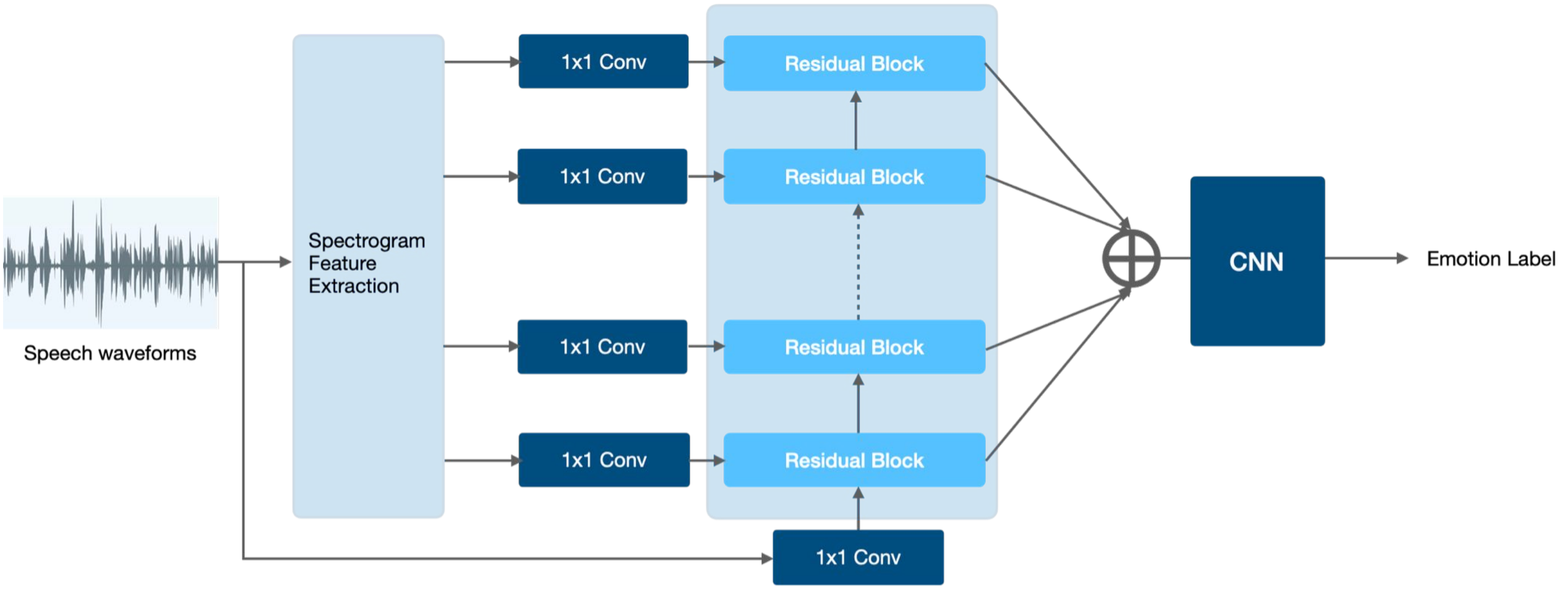
NANYANG TECHNOLOGICAL UNIVERSITY
SINGAPORE

School of Computer Science and Engineering
College of Engineering

Speech Emotion Recognition

with WaveNet

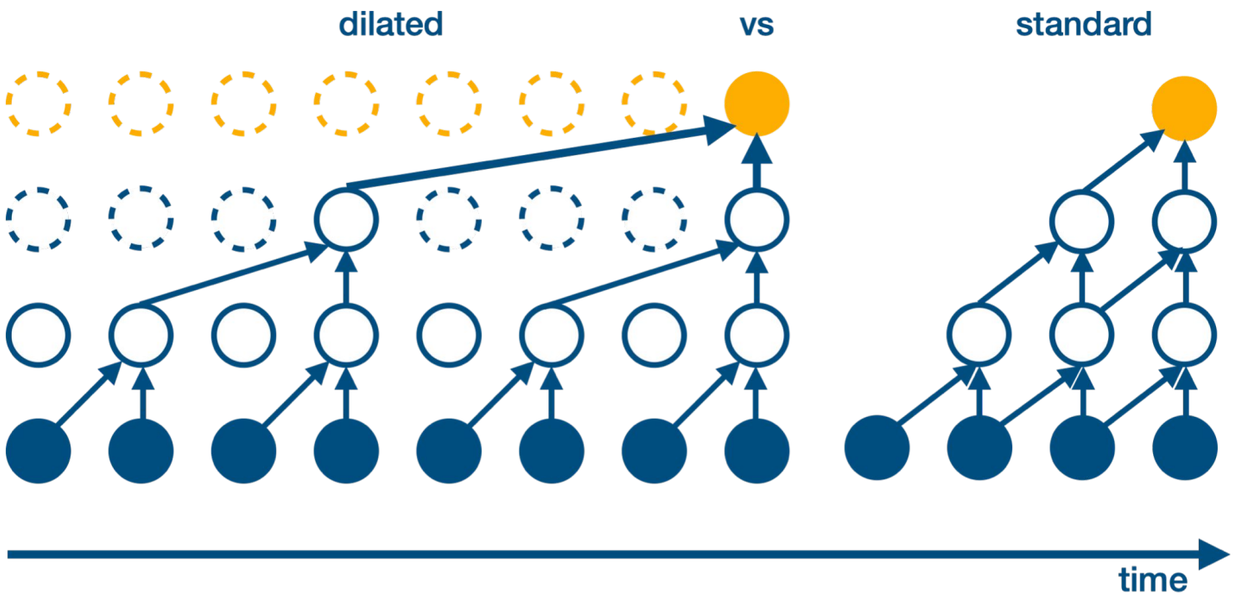
Student: Tan Zhi Yang Supervisor: Prof. Jagath C Rajapakse



Project Objectives:

The project proposes a speech emotion recognition (SER) architecture inspired by WaveNet that does not rely on tedious preprocessing or layers. The model depends on the use of both raw speech waveforms and audio features as inputs and causal dilated convolutions for processing. The model comprises of layers of residual blocks that utilise a gated activation function. This allows the model to outperform a traditional 1D CNN-LSTM model whilst maintaining a lower model complexity.

Dilated Causal Convolutions



Residual Block