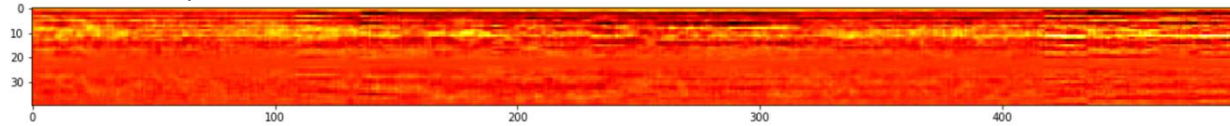


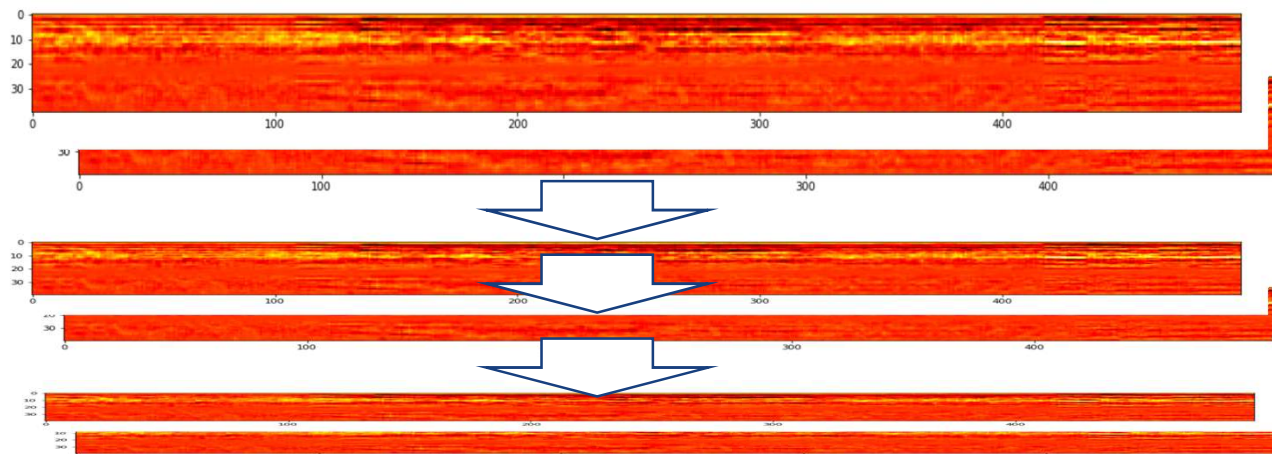
40 triangular band MFCC, 499 slots
(40ms window/20ms overlap,
10,000 audio clip

(n = # of samples)



nX40x500

nX40x500x96

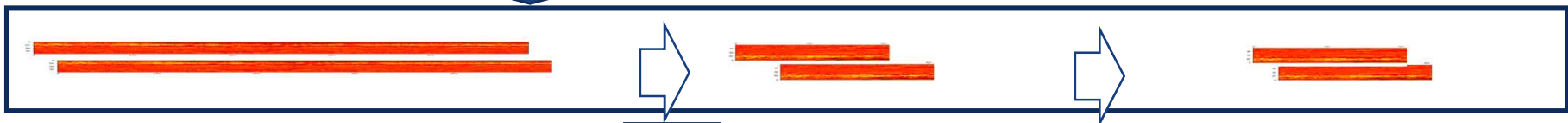


Frequency domain extraction

nX1x500x96

nX1x500x96 (n = # of samples)

"Temporal" domain extraction



Fully Connected Feed forward Neural network

Fully Connected Feed forward Neural network