Appendix II: The Architecture of CNN and RNN Models

1. Convolutional Neural Network

CNN Model	Training Parameter	Architecture
CNN Model1 1 (The best shallow CNN model we get)	Xavier Initializer Batch Size: 211 Learning rate: 0.005 AdamOptimizer Decay rate: 0.975 (after epoch > 40) Epoch: 100	Input: 2115 * 22 * 1000 * 1 **Conv (stride: (1, 2), kernel: (1, 40), filter size: 40)** Activation Map: 2115 * 22 * 481 * 40 **Batch Normalization (momentum: 0.1)** **Conv (stride: (1, 1), kernel: (22, 1), filter size: 40)** Activation Map: 2115 * 1 * 481 * 40 **Batch Normalization (momentum: 0.1)** **Square Activation** **Average Pool (stride: (1, 15), kernel: (1, 75))** Activation Map: 2115 * 1 * 28 * 40 **Dropout (p: 0.3)** **Fully Connect** Output: 2115 * 4 Softmax
CNN Model 2 (The original shallow CNN model proposed in [1])	Xavier Initializer Batch Size: 211 Learning rate: 0.005 AdamOptimizer No decay rate Epoch: 100	Input: 2115 * 22 * 1000 * 1 **Conv (stride: (1, 1), kernel: (1, 40), filter size: 40)** Activation Map: 2115 * 22 * 961 * 40 **Batch Normalization (momentum: 0.1)** **Conv (stride: (1, 1), kernel: (22, 1) filter size: 40)** Activation Map: 2115 * 1 * 961 * 40 **Batch Normalization (momentum: 0.1)** **Square Activation** **Average Pool (stride: (1, 15), kernel: (1, 75))** Activation Map: 2115 * 1 * 61 * 40 **Dropout (p: 0.3)** **Fully Connect** Output: 2115 * 4 Softmax
CNN Model1 3 (The shallow model with extra frequency domain transferred channel)	Xavier Initializer Batch Size: 105 Learning rate: 0.005 AdamOptimizer Decay rate: 0.975 (after epoch > 40) Epoch: 100	Input: 2115 * 44 * 1000 * 1 (for each channel, we used DFT to obtain the signal in frequency domain, and used it as extra features) **Conv (stride: (1, 1), kernel: (1, 40), filter size: 40)** **Batch Normalization (momentum: 0.1)** Activation Map: 2115 * 44 * 961 * 40 **Conv (stride: (1, 1), kernel: (44, 1) filter size: 40)** **Batch Normalization (momentum: 0.1)** Activation Map: 2115 * 1 * 961 * 40 **Square Activation** **Average Pool (stride: (1, 15), kernel: (1, 75))** **Dropout (p: 0.3)** Activation Map: 2115 * 1 * 61 * 40 **Fully Connect** Output: 2115 * 4 Softmax
CNN Model1 4 (The CNN with one more convolute layer comparing to Model 1)	Xavier Initializer Batch Size: 211 Learning rate: 0.005 AdamOptimizer Decay rate: 0.975 (after epoch > 40) Epoch: 200	Input: 2115 * 22 * 1000 * 1 **Conv (stride: (1, 1), kernel: (1, 25), filter size: 40)** Activation Map: 2115 * 22 * 976 * 40 **Batch Normalization (momentum: 0.1)** **Conv (stride: (1, 1), kernel: (22, 1) filter size: 40)** Activation Map: 2115 * 1 * 976 * 40 **Batch Normalization (momentum: 0.1)** **Max Pool (stride: (1, 6), kernel: (1, 6))** Activation Map: 2115 * 1 * 162 * 40 **Conv (stride: (1, 1), kernel: (1, 23) filter size: 40)** Activation Map: 2115 * 1 * 140 * 40 **Batch Normalization (momentum: 0.1)** **Square Activation** **Max Pool (stride: (1, 5), kernel: (1, 5))** Activation Map: 2115 * 1 * 28 * 40 **Dropout (p: 0.6)** **Fully Connect** Output: 2115 * 4 Softmax

* The architectures in the table represent the best architectures and the best training parameters among each category architectures (1. The shallow convolutional net with significantly less parameters in the last fully connected layer; 2. The original shallow net proposed in the previous study; 3. The shallow net with frequency domain channels; 4. A deep convolutional net.)

* For Model 3, we use DFT to extend from 22 channels to 44 channels, where 1 – 22 channels are original signals in the time domain, and 23 – 44 channels are the corresponding signals in the frequency domain.

2. Recurrent Neural Network

RNN Model	Hyper Parameter	Architecture
RNN Model1 1 (Simple RNN)	Batch Size: 50 Learning Rate: 0.001 adamOptimizer epoch: 30	**LSTM, units: 64** **Dropout (p: 0.5)** **Fully Connected, Output dimension: 4** Softmax
RNN Model 2 (Stacked RNN)	Batch Size: 50 Learning Rate: 0.001 adamOptimizer epoch: 30	**Bidirectional LSTM, units: 64** **Dropout (p: 0.5)** **Bidirectional LSTM, units: 64** **Flatten** **Dropout (p: 0.5)** **Fully Connected, Output dimension: 32** **Dropout (p: 0.5)** **Fully Connected, Output dimension: 4** Softmax