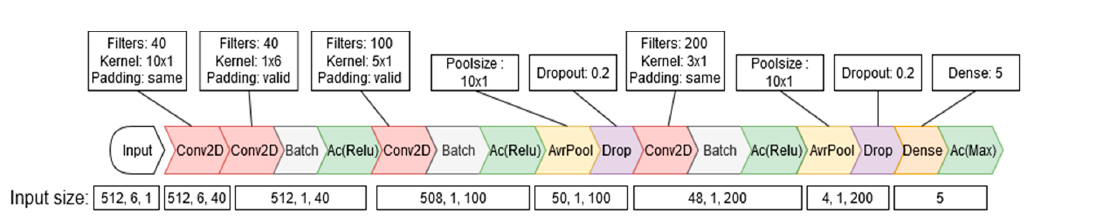
DAA GROUP-1

Classification of Vowels from Imagined Speech with

Convolutional Neural Networks

This Research uses imagined speech to recognize vowels while we are imagining pronouncing them. Imagined speech is very useful for communication. It will help many people who are physically unable to speak. In this research we used 15 persons thinking about pronouncing vowels (“a, e, I, o, u”) and six different words.

The proposed model uses Convolutional Neural Networks (CNN). The objective of this project is to develop a classifier that uses deep learning to classify EEG (Electroencephalogram) signals while imagining us pronouncing the vowels.



This model Consists of 4 convolutional 2d layers, followed by batch Neutralization and Relu (“Rectified layer Unit”) for every layer. And for 2 layers the conv2d with normalization and Relu is followed by Average pooling & Drop Out. The the 4th layer is Flatten which converts the multidimensional arrays to single dimensional array.

The next layer is Dense layer which takes input from previous layers.

Softmax is the last layer which converts the Raw data into probabilities.

The Inputs which we took were the same input which we are presented with in the proposed model. Such as the first convolutional 2d layer consists of input of 40 filters and (10X1) kernel size with Relu function as activation.

In the second convolutional layer “padding=’same’” It means that the spatial dimensions of the output feature map will be the same as the input feature map. Padding is added to the input to ensure this. partial dimensions of the output feature map will be the same as the input feature map. If padding=’valid’ = This means no padding is added to the input data, and the spatial dimensions of the output feature map will be reduced due to the convolution operation.

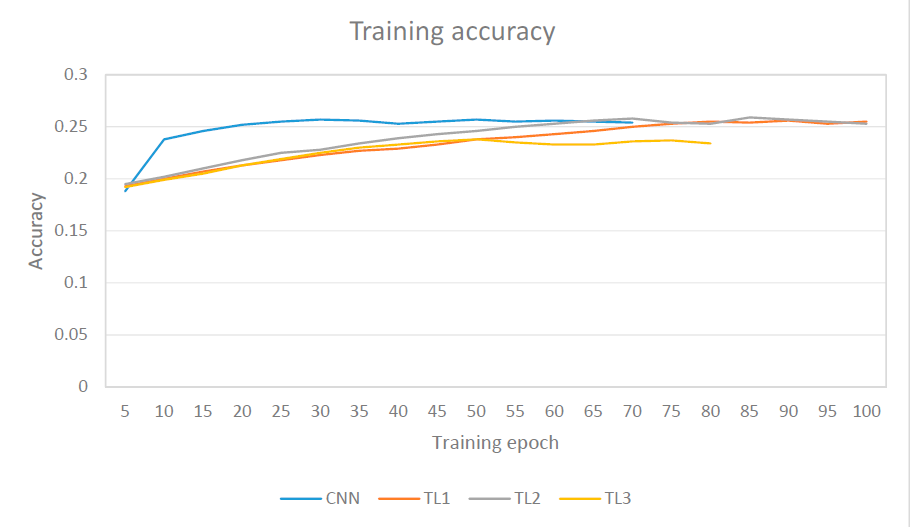
The DATA SET consists of EEG signals taken from 15 sources thinking of saying 5 vowels and six different spanish words such as (“up, down, left, right, back & front”). Then the signals are pre-processed using filtering, Down-sampling and normalization. The pre-processed data is split into Train and test data sets.

The EEG signals were recorded with an 18-channel analogue amplifier at a sampling rate of 1024 Hz, and the electrodes were positioned at six locations according to the 10–20 international system.

In this proposed model all CNN ’s were trained with the Adam optimizer and sparse categorical cross entropy loss function. And Accuracy as metric. Adam optimizer works well when compared to other adaptive optimizers.

**Transfer Learning**: 3 approaches were used with the proposed model to improve the accuracy of the model. The aim of transfer learning is to improve classification accuracy of the model. The TL method freezes the whole base model and then unfreezes the first two convolutional layers to be retrained on the new data. Frozen means keeping the weights of certain layers constant without updating them during training.

Two previous studies resulted in 2 Replicates sets-the 1st Replicate set accuracy was around-22% & The 2nd replicate set accuracy was around-32%. The model we used resulted in with the main model CNN accuracy was around-23% and 3-Transfer learning with accuracy of TL-1=21%, TL-2 = 20%, TL-3=21%.



The model achieves similar accuracy while using fewer layers.

So, This model provides a simple and more Efficient solution for Classification Of Vowels which is very helpful in the Real world. The given model for Classifying Imagined Speech of different vowels is a simpler and more efficient model than the previous models.

The use of Transfer Learning Methods were also very helpful because we were able to train the model along a subject. This model should be re-trainable on a single subject to achieve even higher individual classification performance. This research and help the field of communication a lot and makes the lives of physically impaired persons easier.

By:

1. Pranav Reddy,

N.Kaghuvahan Reddy,

Vemana Sumanth.