

Gradient-weighted Class Activation Mapping (Grad-CAM)[1] is a technique to increase the interpretability and transparency of convolutional neural networks. It is widely used in the field of computer vision to provide intuitive interpretability in a visual way. Porting it to our task, we do not display the heatmap superimposed on the input data because our input data is extracted EEG features. The y-axis of the graph is aesthetically meaningless, the x-axis represents a one-dimensional vector, and the thermal intensity is normalised to values from 0 to 1. X ranges from 270 to 1770 representing the feature of power spectrogram. In the figure, it can be seen that the heat values in this range are high in all the three classifications, indicating that the feature has a great influence on the model in making classification decisions for the final output.

1. Selvaraju R R, Cogswell M, Das A, et al. Grad-cam: Visual explanations from deep networks via gradient-based localization[C]//Proceedings of the IEEE international conference on computer vision. 2017: 618-626.