Task 1.2P

Project Description

Recently, Attention Technique (AT) has been recognized as a prominent solution to the existing Convolutional Neural Network (CNN). Often time, AT is applied in the area of natural language processing as well as image classification. In image classification, AT identifies the unique features of the image and thus improves the predictive capability of the machine learning model. The mechanism of AT can be infused into a wide range of CNN as an intermediate feature map which helps saturate the input map to obtain a more distinguishable feature of the image before reaching the final prediction at densely connected layer of CNN. By nature, the AT mechanism helps machine learning model to ignore the unnecessary background of images and focuses on the objects within them. However, it is doubtful whether this mechanism is effective in resolving image classification where image background is completely or nearly excluded. It is even more challenging when the model is applied in human emotion recognition using facial expressions since the background is minimized and the faces are subjected to small variations. How consistent it is for AT in resolving human face recognition in these scenarios.

To enlight this grey area, this study will carry out a benchmarking comparison between conventional CNN and AT-enabled CNN in predicting human emotions in a different scenario (with and without background) of facial-image datasets to find out whether AT can maintain its performance consistency. A well-known AT, called Convolutional Block Attention Module (CBAM), will be employed in the study to compare the accuracy and error rate of image classification. To the increase the reliability of the study, the experiments will be conducted under different sample settings, i.e. with and without facial-image pre-processing techniques such as illumination and augmentation.