

Proposed Novel model;

Past year Deepmind has released the paper on 'Mixing token in Fourier transform'. Or we can say fouryer –Net. Basically this model is in NLP for sentence-embedding. And is this exceptionally better result as good as Transformers I.e BeRT, RoBeRTA, ELMO.

My idea to use power of Fouryer transform in field of computer vision. Model has the learneable Foueryer paramters with three blocks of pre-Trained ViT. Wth this I got 81% train and 77% on val.

