

Improving Axial-Attention Network via Cross-Channel Weight Sharing

Authors : Nazmul Shahadat, Anthony S. Maida

Abstract : In recent years, hypercomplex inspired neural networks improved deep CNN architectures due to their ability to share weights across input channels and thus improve cohesiveness of representations within the layers. The work described herein studies the effect of replacing existing layers in an Axial Attention ResNet with their quaternion variants that use cross-channel weight sharing to assess the effect on image classification. We expect the quaternion enhancements to produce improved feature maps with more interlinked representations. We experiment with the stem of the network, the bottleneck layer, and the fully connected backend by replacing them with quaternion versions. These modifications lead to novel architectures which yield improved accuracy performance on the ImageNet300k classification dataset. Our baseline networks for comparison were the original real-valued ResNet, the original quaternion-valued ResNet, and the Axial Attention ResNet. Since improvement was observed regardless of which part of the network was modified, there is a promise that this technique may be generally useful in improving classification accuracy for a large class of networks.

Keywords : axial attention, representational networks, weight sharing, cross-channel correlations, quaternion-enhanced axial attention, deep networks

Conference Title : ICIAP 2023 : International Conference on Image Analysis and Processing

Conference Location : Vancouver, Canada

Conference Dates : September 25-26, 2023