

Attention as Activation

中国、デンマーク、ドイツ、アメリカの大学間研究チーム

目的と手段

活性化関数と注意メカニズムは別のものとして開発されているが、これらは非線形ゲート関数として定式化できる。

(注意は広範囲を考えた高度なものの反面、活性化はスカラーで単純)

そこで、活性化と注意メカニズムを統合したATAC(注意活性化ユニット)を提案する。

線形ユニットをATACに置き換えることで、大きく性能が向上した完全な注目ネットワークを構築できる。

主な貢献

- 活性化関数と注意メカニズムをまとめたアプローチを提供する
- 活性化関数の局所性と注意メカニズムの文脈的集約の両方を満たすローカルチャネルの注意モジュール(ATAC)を提案する
ポイントごとにチャネル間の特徴を集約する
- これにより、同程度のパラメータ数なら性能が向上することを示した
 - 画像分類タスク、ResNet50ベースのモデル間で最もよい精度
比較対象
ResNet50, SE-ResNet50, AA-ResNet50, FA-ResNet50, $GE^{\theta} + \sigma$ -ResNet50
 - CIFAR-10分類タスク、以下のモデル間で最もよい精度
比較対象
ResNet, SENet, GENet- σ^{θ}

詳細

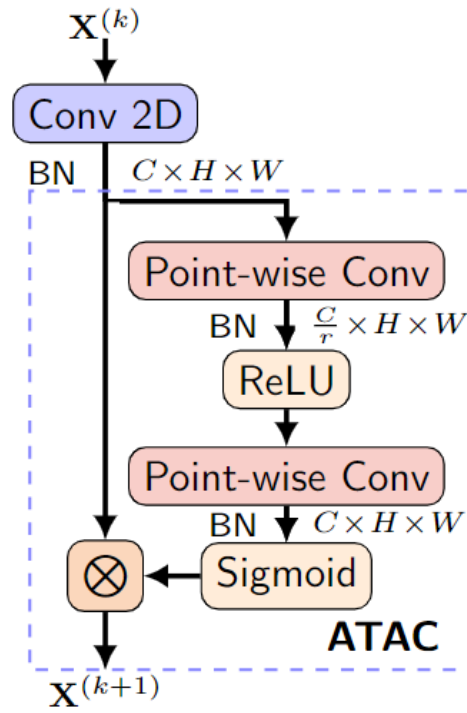


Fig. 1: The proposed attentional activation (ATAC) unit