

Peer Review

Review of: "Partial Convolution Meets Visual Attention"

Vinay Verma¹

1. Independent researcher

This paper talks about a new network called PATNet, which tries to balance both speed and accuracy for vision tasks. They've improved on FasterNet by fixing its main drawback — since FasterNet skips some channels during convolution, accuracy takes a hit. So these people added a smart attention mechanism called Partial Visual Attention (PAT) that fills this gap without making the model heavy. They've designed three types of blocks — channel, spatial, and self-attention — which help the model understand global and local features better. Their results on ImageNet and COCO show good improvement over FasterNet, both in terms of accuracy and speed. Overall, a simple idea but quite effective!

Innovation-wise, it's a nice balance between keeping the model light and still improving accuracy — especially when compared to heavy transformer-based models like ViTs. Practicality looks promising too, since attention-based models often struggle with speed on edge devices, but PATNet seems designed with that limitation in mind. Compared to other models like EfficientNet or ConvNeXt, this one looks easier to deploy on low-power devices while still staying competitive on accuracy. But the paper could've added more comparisons on real-world tasks to highlight this better.

Declarations

Potential competing interests: No potential competing interests to declare.