Combining Local and Global Features Using Transformer-Based Architecture for Efficient Low-Light Enhancement

Authors : Michał Wnuczyński, Maciej Nalewaj

Abstract : Transformer-based methods that aim to enhance dark, underexposed images, in most cases, require an excessive amount of available memory. The inference time of these models accelerates quickly with the increasing image resolution, making them unsuitable for mobile deployment due to their relatively low RAM memory. Proposed in this paper, architecture combines extraction of features across spatial and channel domains using two different self-attention mechanisms and U-Net structure. Using global and window-based self-attention mechanisms in channel and spatial domains effectively extracts important features and efficiently combines them, producing high-quality results. Comprehensive qualitative and quantitative evaluations show that the proposed method achieves state-of-the-art results while maintaining significantly lower memory needs and rapid inference speed.

Keywords : low-light enhancement, transformer, U-net, self-attention, image processing, computer vision
Conference Title : ICCV 2024 : International Conference on Computer Vision
Conference Location : Vancouver, Canada
Conference Dates : September 26-27, 2024