INFORMATICS INSTITUTE OF TECHNOLOGY

INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

Enhancer: An Effective Transformer Architecture for Single Image Super Resolution

A dissertation by

Mr. Chamath Peiris

Supervised by

Mr. John Sriskandarajah

Submitted in partial fulfilment of the requirements for the BEng (Hons) in Software Engineering degree at the University of Westminster.

July 2022

Abstract

A widely researched domain in the field of image processing in recent times has been single image super-resolution, which tries to restore a high-resolution image from a single low-resolution image. Many more single image super-resolution efforts have been completed utilizing equally traditional and deep learning methodologies, as well as a variety of other methodologies. Deep learning-based super-resolution methods, in particular, have received significant interest. As of now, the most advanced image restoration approaches are based on convolutional neural networks; nevertheless, only a few efforts have been performed using Transformers, which have demonstrated excellent performance on high-level vision tasks. The effectiveness of CNN-based algorithms in image super-resolution has been impressive. However, these methods cannot completely capture the non-local features of the data. Enhancer is a simple yet powerful Transformer-based approach for enhancing the resolution of images. A method for single image super-resolution was developed in this study, which utilized an efficient and effective transformer design. This proposed architecture makes use of a locally enhanced window transformer block to alleviate the enormous computational load associated with non-overlapping window-based self-attention. Additionally, it incorporates depth-wise convolution in the feed-forward network to enhance its ability to capture local context. This study is assessed by comparing the results obtained for popular datasets to those obtained by other techniques in the domain.

Keywords: Single Image Super Resolution; Computer Vision; Vision Transformers; Image Restoration; Encoder; Decoder; Self-Attention