

SwinIR: Image Restoration Using Swin Transformer

Rafael Monteiro Laranjeira
rml@ic.ufal.br



Introdução

Introdução

- Restauração de imagem engloba super resolução, remoção de ruído, redução de artefatos gerados pela compressão de JPEG e outras aplicações;
- Devido ao bom resultado dos transformers na área de PLN, começaram a investigar o desempenho dos transformers em outras áreas, como processamento de imagens;
- Com isso, o artigo propõe uma metodologia para a restauração de imagens utilizando o Swin Transformer.



Trabalhos relacionados

Convolutional Neural Networks

- A restauração de imagens geralmente é feita utilizando CNNs;
 - RNN;
 - ESRGAN;
 - RCAN;
- A performance é boa, especialmente com redes mais elaboradas;
- Entretanto, sofrem de dois problemas básicos:
 - As interações entre os kernels de convolução não dependem do conteúdo (posições);
 - Convolução não é efetiva para memórias longas.

Transformers

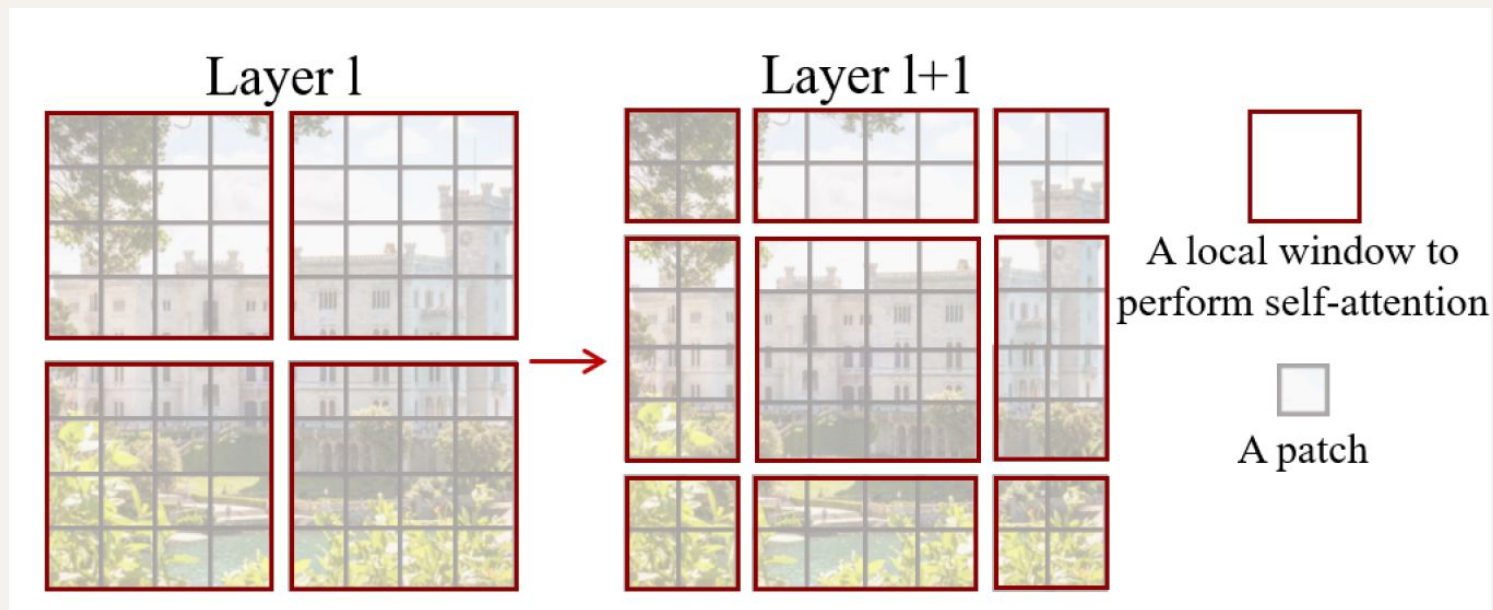
- Os transformers mostraram bom desempenho em diversos problemas de visão computacional;
 - DETR
 - DeiT
 - ViT;
- Os transformers utilizados para restauração de imagens geralmente trabalham com a divisão da imagem de entrada em pedaços de tamanho fixo;
- Devido a isso, possuem alguns pontos negativos:
 - Os pixels da borda não podem utilizar os vizinhos de maneira adequada;
 - A restauração da imagem pode introduzir artefatos(ruídos) nas bordas ao redor de cada pedaço;
 - Pode ser resolvido, mas tem um custo computacional elevado.



Swin Transformer

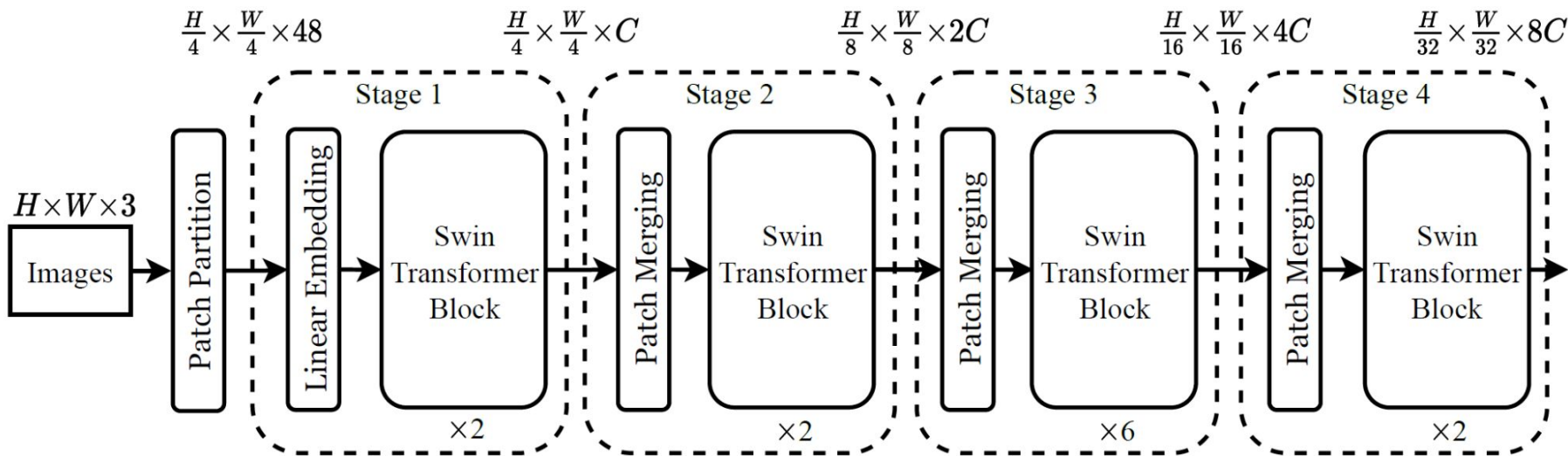
Swin Transformer

- Shifted **W**indows Transformer;



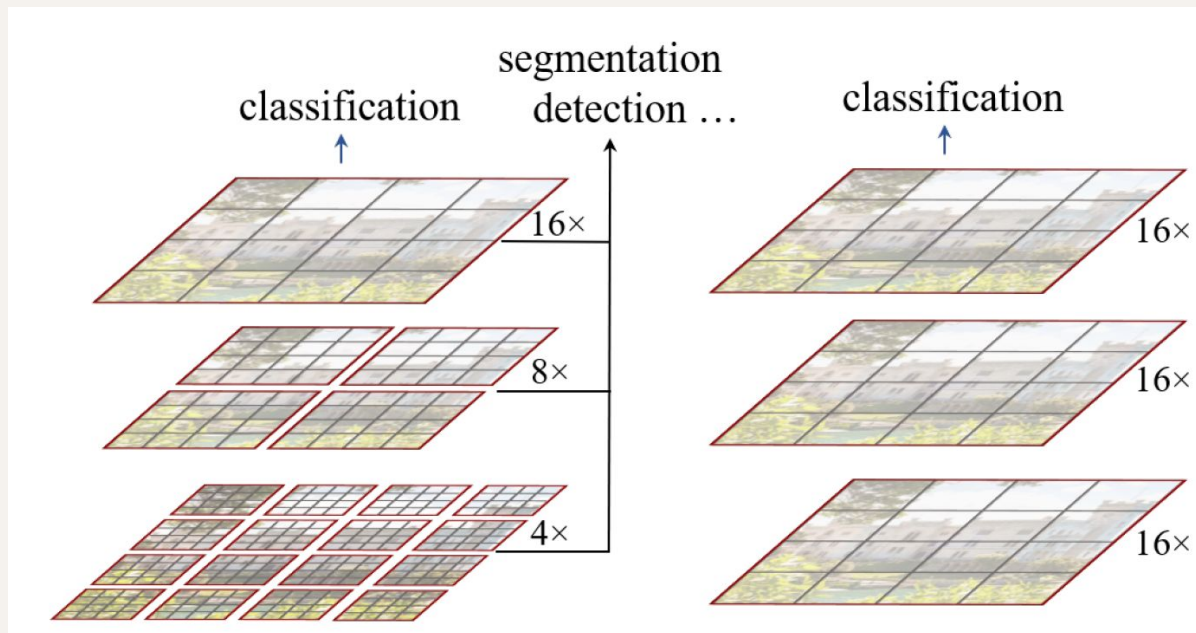
Swin Transformer

- Arquitetura do Swin Transformer.



Swin Transformer

- Features organizadas de maneira hierárquica.



Vantagens do Swin Transformer

- Integra as vantagens de uma CNN e de um Transformer;
- Constrói uma representação hierárquica das features;
- Complexidade computacional linear em relação à imagem;
- Performance superior a outros métodos;

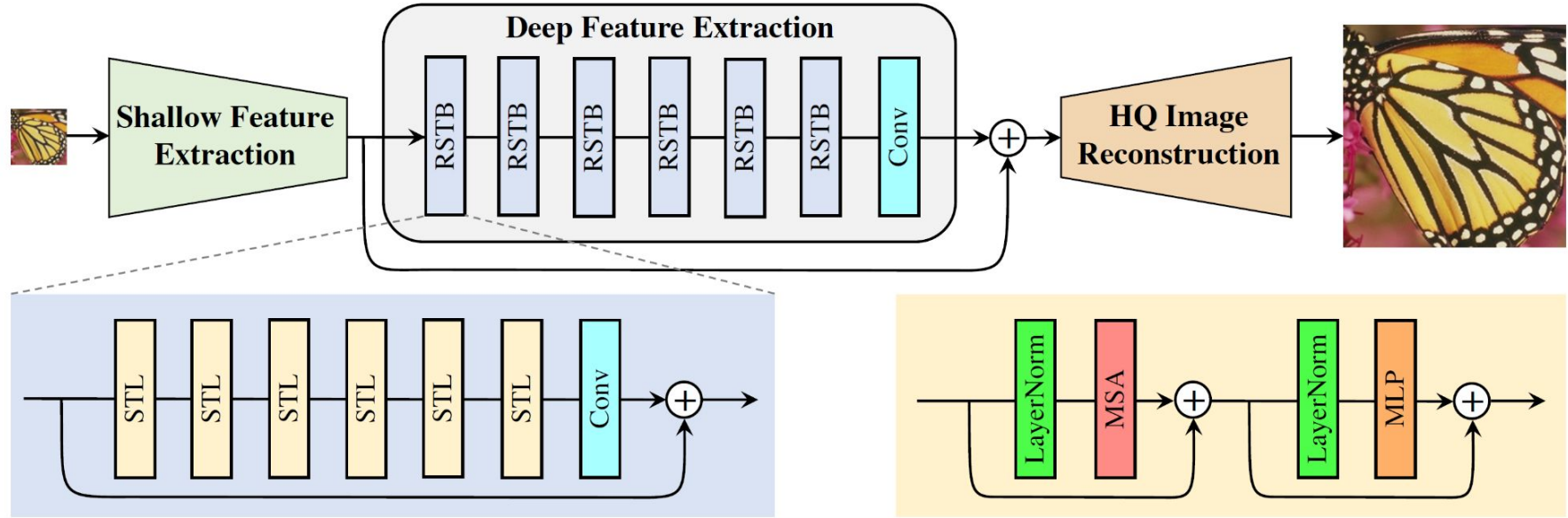
$$\Omega(\text{MSA}) = 4hwC^2 + 2(hw)^2C,$$

$$\Omega(\text{W-MSA}) = 4hwC^2 + 2M^2hwC,$$

The image features two thin, dark horizontal lines. The top line starts with a curve on the left side and then continues straight to the right edge. The bottom line starts straight from the left edge and ends with a curve on the right side.

SwinIR

Arquitetura do SwinIR

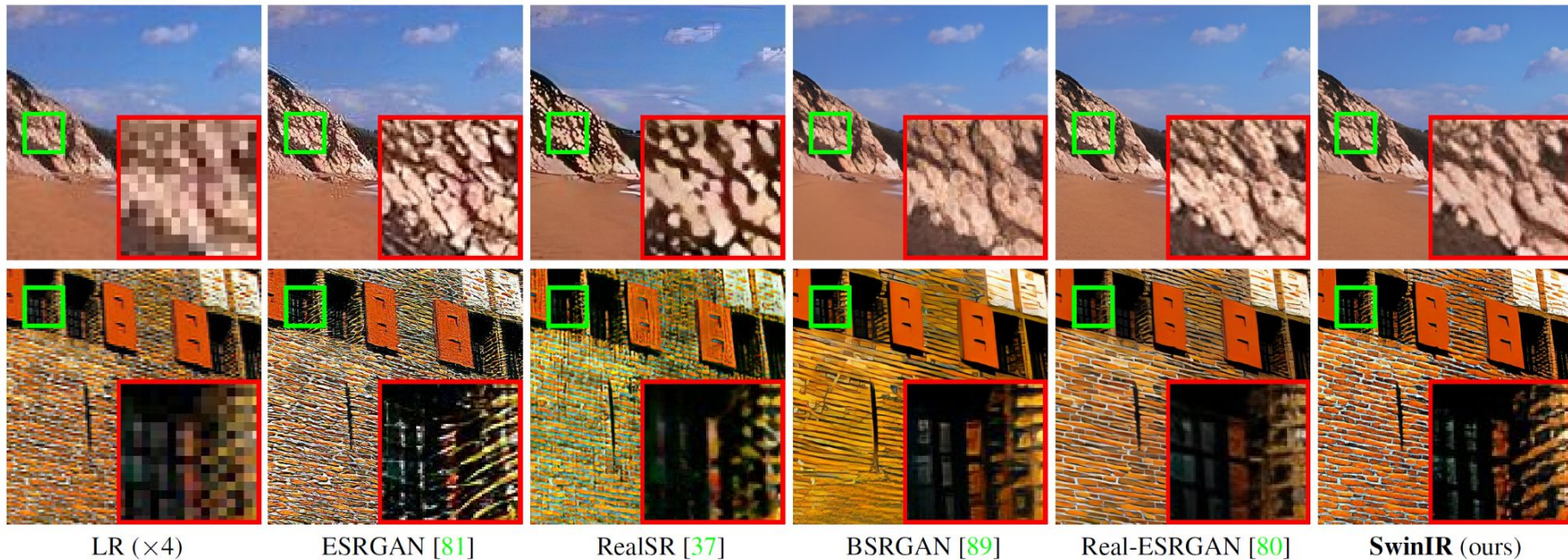




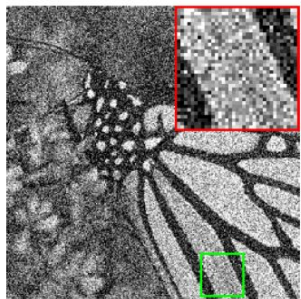
Resultados do SwinIR



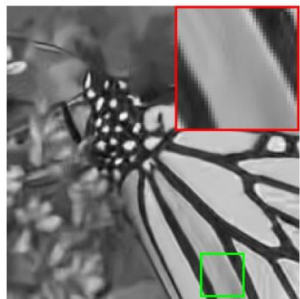
Super Resolução



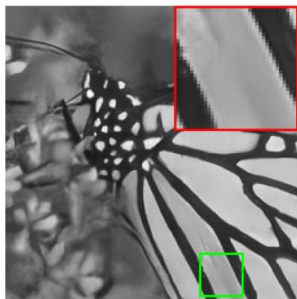
Remoção de ruído



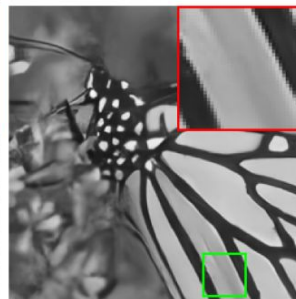
Noisy



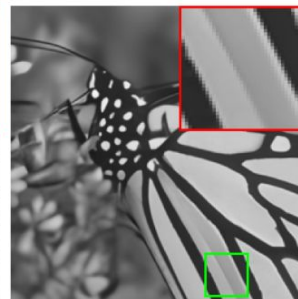
BM3D [14]



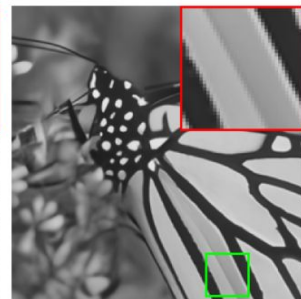
DnCNN [90]



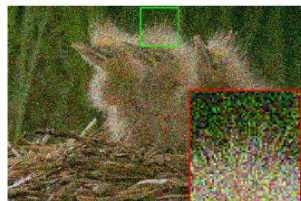
FFDNet [92]



DRUNet [88]



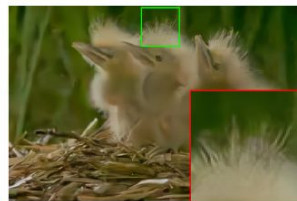
SwinIR (ours)



Noisy



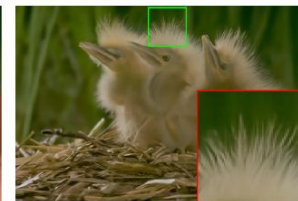
DnCNN [90]



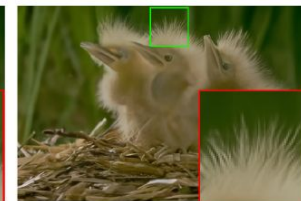
FFDNet [92]



IPT [9]



DRUNet [88]



SwinIR (ours)

Testes de super resolução



Original



realESRGAN



BRSGAN

Testes de super resolução



Original



SwinIR



SwinIR_large

Referências

Link para o paper: <https://paperswithcode.com/paper/swinir-image-restoration-using-swin>