MG-MLP: Multi-gated MLP for Restoring Images from Spatially Variant Degradations

Jaihyun Koh\textsuperscript{1}, Jaihyun Lew\textsuperscript{2}, Jangho Lee\textsuperscript{3} and Sungroh Yoon\textsuperscript{2}

Samsung Display\textsuperscript{1}, Data Science and AI Laboratory, Seoul National University\textsuperscript{2}, Incheon National University\textsuperscript{3}

Introduction and Background

1. Image Restoration Task
   - Low-level semantic, the relationship between pixels rather than abstract.
   - Sufficient receptive field is essential to search for similar patterns.
   - Architecture should be "fully convolutional": Not fixed resolution.

2. Restormer (CVPR 2022)
   - Channel-wise multi-head self-attention with quadratic complexities.
   - Leveraging 3rd order interaction of transformer architectures.
   - Invariant architecture to the resolution of an input image.

3. NAFnet (ECCV 2022)
   - MLP architecture using 1x1 conv. and 3x3 depth-wise conv.
   - Gating operation using element-wise multiplication.
   - Channel mixing (1x1) and token mixing (3x3) are critical components.

Proposed MG-MLP Block

1. Intra-token gating: controls the flow of information through the interaction of data in each token.
2. Cross-token gating: updates the resulting token from intra-token gating and simultaneously brings back the data discarded by intra-token gating by referring to the adjacent tokens.

Benefit / Experiments and Results

1. More flexible information flow control
   - Update gate, reuse gate inspired by LSTM, GRU.
   - Input and receptive field adaptive gating
   - 3rd order interaction like ViT
   - MG-MLP: Multiplication of three projection input

2. Multi-modal Gaussian Mixture
   - By two gating unit

Complexity and Performance

Generalization

Feature Visualization and Ablation

Macro Architecture

Feature Visualization

Ablation

GoPro Dataset

Real Blur-J and DVD Dataset

The two residuals refined by the two gating mechanisms contain complementary features. The two gating paths perform distinct feature refinements by dividing a region into more blurry and less blurry regions.