Visual Prompts for Guiding Image Restoration Tasks

Supervisors: Javier Vazquez-Corral, David Serrano-Lozano (CVC/UAB)

Image restoration represents a fundamental challenge in computer vision, representing an essential inverse image-to-image problem. It involves the task of recovering a high-quality image from its degraded version, including tasks ranging from enhancement, deblurring, and denoising to super-resolution and beyond. Over recent years, there has been a surge in methodologies addressing this challenge, with a notable trend toward using textual cues to guide image manipulation. These approaches leverage text prompts to describe and guide powerful models.

In this project, we plan to diverge from the text-guided-based methods and delve into visual prompts [1] to guide and refine the suite of image restoration tasks (see Figure 1). Rather than relying on textual input, our focus lies in generating a comprehensive set of features that encapsulate important image context. These features may include depth maps for spatial understanding, color naming as human perception features, optical flow for motion, and other tailored cues that can be learnt. By exploring the potential of guided image manipulation, we plan to leverage state-of-the-art models such as all-in-one methods (MAXIM [2], MIRNet [3]), GAN-based (VQGAN [4]) and diffusion-based models [5].
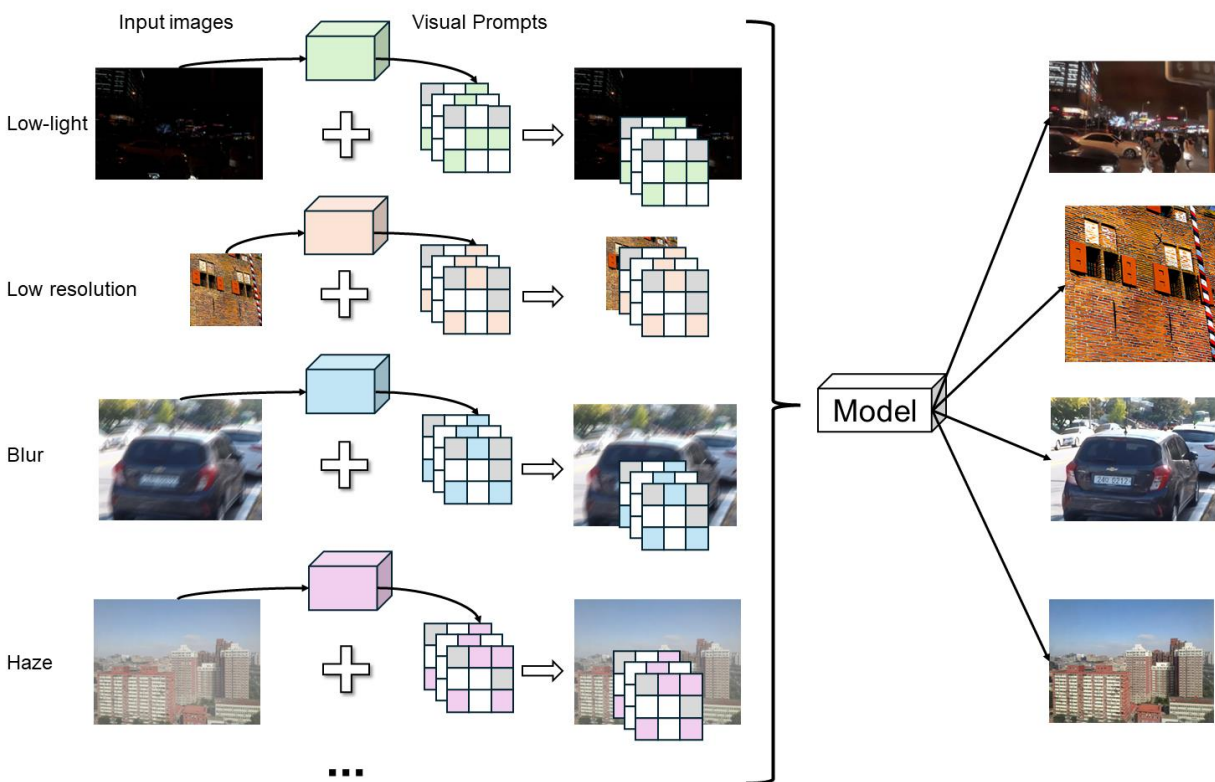

Figure 1: Example of visual prompt image restoration. The model receives as inputs a corrupted image and a prompt referred to that corruption –coming from an independent model- and outputs a restored image.

References:

[1] Chen, A., Yao, Y., Chen, P. Y., Zhang, Y., & Liu, S. Understanding and improving visual prompting: A label-mapping perspective. In CVPR 2023.

[2] Tu, Z., Talebi, H., Zhang, H., Yang, F., Milanfar, P., Bovik, A., & Li, Y. Maxim: Multi-axis mlp for image processing. In CVPR 2022

[3] Zamir, S. W., Arora, A., Khan, S., Hayat, M., Khan, F. S., Yang, M. H., & Shao, L. Learning enriched features for real image restoration and enhancement. In ECCV 2020

[4] Cao, S., Yin, Y., Huang, L., Liu, Y., Zhao, X., Zhao, D., & Huang, K. Efficient-VQGAN: Towards High-Resolution Image Generation with Efficient Vision Transformers. In CVPR 2023

[5] Ren, M., Delbracio, M., Talebi, H., Gerig, G., & Milanfar, P. Multiscale structure guided diffusion for image deblurring. In CVPR 2023