Project Topic:

"Quantifying the Impact of Transformer-based Image Restoration On Downstream Vision Tasks."

Problem Statement:

State-of-the-art object detection models perform very well on benchmark datasets. These datasets primarily consist of clean and high-quality images. But, real-world applications often have to deal with images that contain some level of degradation, such as sensor noise and motion blur. The difference between training data and actual prediction data reduces the model's performance.

This project aims to quantitatively analyze whether image restoration can help recover the performance lost due to degraded images in object detection tasks. To be precise, I aim to investigate how much the Restormer model can improve the performance of object detection on artificially degraded images.

Team Member:

1. Bitupan Arandhara

Learning Objectives:

1. Practical experience with modern architectures: Through this project, I aim to gain hands-on experience working with a transformer-based image restoration model and a state-of-the-art object detection model. I also want to learn how to utilize pre-trained models for performing specific tasks.
2. Building end-to-end pipelines: I want to learn how to build end-to-end computer vision pipelines that handle data preprocessing, model training, model inference, and quantitative evaluation.
3. Using performance evaluation metrics: I want to learn and apply standard evaluation protocols for object detection tasks.

Plan of actions:

1. Dataset Selection:

   I plan to use the MS-COCO 2017 validation set as the dataset for the project. It contains 5000 annotated images across 80 object categories.

2. Pipeline Setup and Implementation:

   I will implement a pipeline with three main stages: degradation, restoration, and detection. The pipeline will take a clean image, degrade it, and then restore it using Restormer.

   a. Degradation: We will start with a clean, annotated dataset and apply various types of degradations to the images. For example, add different levels of Gaussian noise or apply motion blur with varying kernel sizes.
   b. Restoration: Use the publicly available pre-trained Restormer models to restore the images degraded in the previous step.
   c. Detection: Use a pre-trained object detector, such as YOLOv5s or Faster R-CNN, and run inferences on the three sets of images: the original clean images, the degraded images, and the restored images.

3. Experimentation and Evaluation:

   I will measure the object detection performance using Mean Average Precision (mAP) on the original clean images, the degraded images, and the restored images separately. I will create detailed tables to compare the object detection performance on the three sets of images to visualize clearly:

   - The performance drop caused by each type and level of degradation.
   - The performance recovered by the Restormer's preprocessing step.
   - Compare the performance on the restored images to that on the clean images and on the degraded images.

   I will also visually analyze the detection results, showing them on the same image under three conditions: clean, degraded, and restored. It will provide intuitive, qualitative evidence to support the quantitative findings.