HArDFormer: Advancing Dehazing and Artefact Reduction through Transformer-based Image Restoration

Niccolò Camarlinghi[1], Antonio Di Tommaso[1], Benedetto Michelozzi[1], Giacomo Fontanelli[1], Andrea Masini[1]

[1]FlySight Srl, Via A. Lampredi 45, Livorno 57121, Italy

The visibility reduction caused by fog and other obscuring factors is known to have an impact on situational awareness. To address this issue, a number of algorithms have been developed to enhance the visibility of affected scenes. While artificial intelligence models like convolutional neural networks (CNNs), Generative Adversarial Networks (GANs) and Transformer-based models have shown impressive capabilities in generating realistic images, they are not immune to producing artefacts like blockiness, texture artefacts, distorted shapes and visual hallucination. Certain types of artefacts may not only degrade the quality of the scene but also have the potential to give rise to inaccurate predictions or interpretations of images, thereby compromising the system's reliability. Such compromise could translate into safety concerns, especially in critical applications like enhanced flight vision systems, autonomous vehicles as it may result in unsafe decisions or recommendations that could potentially lead to accidents or harm to individuals. For this reason, in this work we focus on formulating an efficient solution aimed at enhancing visibility in scenarios involving obscurants such as fog, smoke, and other atmospheric disturbances, while simultaneously minimizing the presence of artefacts. In order to give a quantitative estimate of the severity of the artefacts, the HArD (Hazy Artifact Detector) metric, which quantifies the presence of artefacts in defogged images, was used. In this study, we conducted a comparative study of eight commonly used single-image dehaze/enhancement algorithms, namely Dark Channel Prior, Non-Local Image Dehazing, CLAHE, AOD-Net, DehazeNet, MSCNN, Pix2Pix, and DehazeFormer. In order to conduct a comprehensive comparison between the different algorithms we created a dataset comprising haze-augmented images generated starting from the data available in three public datasets: LLVIP, M3FD and KAIST. Our investigation revealed that DehazeFormer - a vision Transformer for single image dehazing – outperforms the others across the majority of metrics relevant to this context. However, this model is among those that produces stronger artefacts. To limit the severity of the artefacts, we propose HArDFormer, an enhancement of original DehazeFormer, which is capable of removing the presence of obscurants similarly to the original model, while minimizing the presence of artefacts. To achieve this result, we introduced the HArD metric into the DehazeFormer‘s loss function.

By comparing DehazeFormer and HardFormer, we observed that our model outperforms DehazeFormer in terms of “No Reference” metrics, such as NIQE, PIQE and BRISQUE, with improvements of 9.84%, 27.68%, 25.54%, respectively. Additionally, it excels in custom metrics FADE and HArD, showing improvements of 1.13% and 67.20%, respectively. When considering the “Full Reference” metrics, we observed a 19.18% decrease in MSE and slight decreases for PSNR and SSIM, 4.33% and 1.19% respectively.

In conclusion, our study led us to choose a single-image dehazing model from among various models available in the literature and enhance it to efficiently eliminate obscurants in scenes while significantly reducing artefact presence in the output scenes.