# FLARE-FREE VISION: EMPOWERING UFORMER WITH DEPTH INSIGHTS

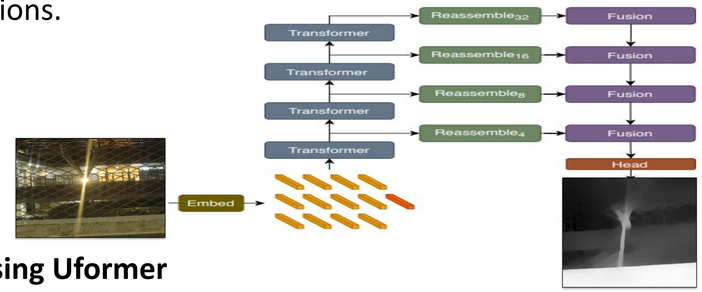
*Yousef Kotp, Marwan Torki*
*Alexandria University*
yousefkotp@outlook.com, mtorki@alexu.edu.eg

**Abstract -** Image flare is a common problem that occurs when a camera lens is pointed at a strong light source. It can manifest as ghosting, blooming, or other artifacts that can degrade the image quality. We propose a novel deep learning approach for flare removal that uses a combination of depth estimation and image restoration. We use a Dense Vision Transformer to estimate the depth of the scene. This depth map is then concatenated to the input image, which is then fed into a Uformer, a general U-shaped transformer for image restoration. Our proposed method demonstrates state-of-the-art performance on the Flare7K++ test dataset, demonstrating its effectiveness in removing flare artifacts from images. Our approach also demonstrates robustness and generalization to real-world images with various types of flare. We believe that our work opens up new possibilities for using depth information for image restoration.

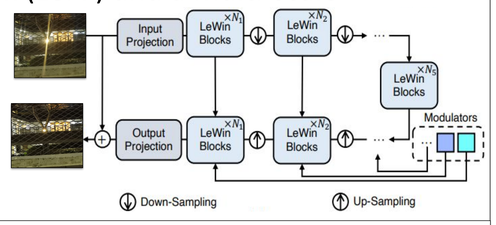
## Approach

The image passes through the following 2-stages:

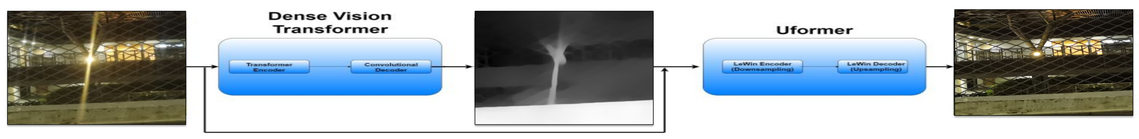
1. **Depth Estimation Network using Dense Vision Transformer**
   Consists of a vision transformer encoder for global image features across various resolutions and a convolutional decoder for integrating these features into full-resolution, detailed predictions.




2. **Image Restoration Network using Uformer**
   Consists of a hierarchical encoder-decoder structure using Transformer blocks for image restoration, featuring non-overlapping Window-based Multihead Self-Attention (W-MSA), Locally-enhanced Feed-Forward Network (LeFF) and a multi-scale restoration modulator for detail enhancement.




**Overall Pipeline:**




## References

1] Yuekun Dai, Chongyi Li, Shangchen Zhou, Ruicheng Feng, and Chen Change Loy, "Flare7k: A phenomenological night- time flare removal dataset," in Thirty-sixth Conference on Neural Information Processing Systems Datasets and Benchmarks Track, 2022
[2] Yuekun Dai, Chongyi Li, Shangchen Zhou, Ruicheng Feng, Yihang Luo, and Chen Change Loy, "Flare7k++: Mixing synthetic and real datasets for nighttime flare removal and beyond," 2023.
[3] Zhendong Wang, Xiaodong Cun, Jianmin Bao, Wengang Zhou, Jianzhuang Liu, and Houqiang Li, "Uformer: A general u-shaped transformer for image restoration," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2022, pp. 17683–17693.
[4] Rene Ranftl, Alexey Bochkovskiy, and Vladlen Koltun, "Vision transformers for dense prediction," in 2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021, pp. 12159–12168

## Results

1. **Visual results**


2. **Results on Flare7K++ real test dataset**

'*' denotes models with reduced parameters due to the limited GPU memory.

| Metric/Method | PSNR | SSIM | LPIPS | G-PSNR | S-PSNR |
|---|---|---|---|---|---|
| Flare7K [18] - NeurIPS 2022 | 26.978 | 0.890 | 0.0466 | 23.507 | 21.563 |
| Flare7K++ (U-Net) [19, 16] | 27.189 | 0.894 | 0.0452 | 23.527 | 22.647 |
| Flare7K++ (HINet) [19, 25] | 27.548 | 0.892 | 0.0464 | **24.081** | **22.907** |
| Flare7K++ (MPRNet*) [19, 26] | 27.036 | 0.893 | 0.0481 | 23.490 | 22.267 |
| Flare7K++ (Restormer*) [19, 27] | 27.597 | **0.897** | 0.0447 | 23.828 | 22.452 |
| Flare7K++ (Uformer) [19, 7] | 27.633 | 0.894 | 0.0428 | 23.949 | 22.603 |
| Zhou et al. [20] - ICCV 2023 | 25.184 | 0.872 | 0.0548 | 22.112 | 20.543 |
| Ours | **27.662** | **0.897** | **0.0422** | 23.987 | 22.847 |

## Conclusion

- We propose a novel 2-stage approach for a more accurate flare removal process by leveraging depth information in the process of image restoration.
- Our approach integrates a depth map estimated using a dense vision transformer with the RGB channel before employing the Uformer model.
- Our model eliminates flare while preserving the authenticity of the primary light source.
- The paper demonstrates the state-of-the-art performance with comprehensive qualitative and quantitative results on Flare7K++ dataset.