



Figure 1: Mitigation capability of our method without using base restoration methods for pre-processing.

## A Appendix / supplemental material

### A.1 MLP Network Details

Our network architecture consists of two MLPs: a content MLP and a deformation MLP. The content MLP has 4 fully-connected layers, with an input dimension of  $Q_1$ , a hidden dimension of 128, and an output dimension of 2, representing  $\Delta x$  and  $\Delta y$ . The deformation MLP comprises 6 fully-connected layers, with an input dimension of  $Q_2$ , a hidden dimension of 256, and an output of 3 channels representing RGB intensity.

### A.2 Position Encoding

Position encoding for spatial and temporal indices is embedded within the trainable feature map, as these indices are trainable. We directly use  $x$ ,  $y$ , and  $t$  to query the corresponding feature tensors from the feature maps. Notably, in the temporal feature map, neighboring features are shared across multiple frames, with each frame weighted differently due to explicit regularization.

### A.3 Additional Results on Mitigation Capability without Base Restoration Methods.

Additional results highlighting our method’s mitigation capability independently of base restoration techniques are presented in Figure 1.