Supplemental Material for ReCoNet: Recurrent Correction Network for Fast and Efficient Multi-modality Image Fusion

Anonymous ECCV submission

Paper ID 3864

Abstract. This document provides more information on three points: To begin, we display more realistic instances of our proposed ReCoNet, which completely demonstrates the improved visual effects as compared to the state-of-the-art. Then, we demonstrate how our network could be used to remove image noise and perform additional downstream tasks. Finally, we explore the model's limits and present several failure instances. The source code is included in this supplemental content.

1 More realistic instances of our ReCoNet

We provide more visual comparisons to verify the superiority of our proposed method against eight other state-of-the-art methods (DenseFuse[2], FusionGAN[6], RFN[3], GANMcC[7], DDcGAN[5], MFEIF[4], U2Fusion[9]) on *TNO* and *Road-Scene* Dataset, which are shown in Figure. 1 and Figure. 2, respectively.

2 Extension to high-level tasks

As our ReCoNet is designed for real scenarios, this section will exhibit the denoise effects (Figure 3), and show how our fusion method optimize the results of the following applications. The visual comparisions of realizing object detection with YoloV4[1], depth estimation with MiDaS[8] on fused images are present in Figure 4 and Figure 5.

3 Limitations and failure cases

By predicting deformation fields $\phi = R_\phi(I_{ir}, \tilde{I}_{vis})$, the micro registration module in our ReCoNet corrects visible images with slight misalignment. It will not perform very well if the objects in the visible images do not correspond to ones in the infrared images. As illustrated in Figure 6, the street light in the green box is successfully adjusted, however the pillar in the red box is unable to correct due to which in the infrared image being unmatched.

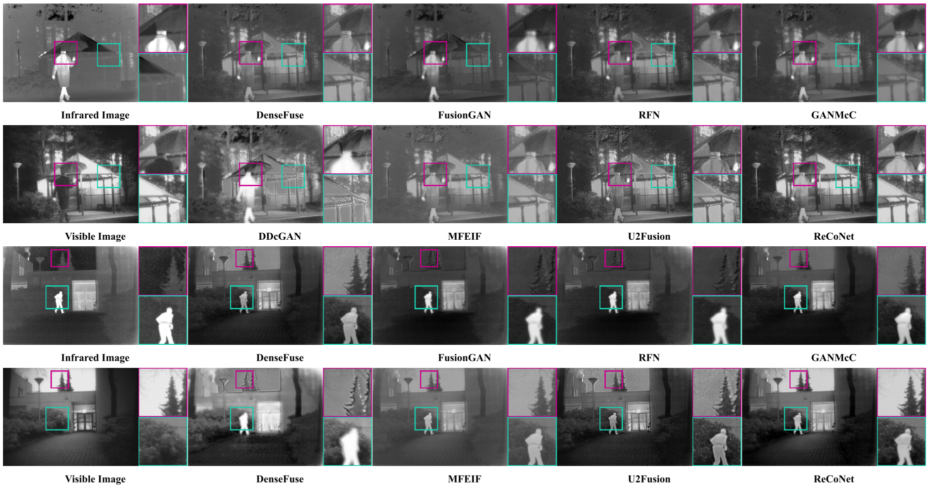

Fig. 1. Visual results comparison between different methods on *TNO* Dataset Best view on screen.

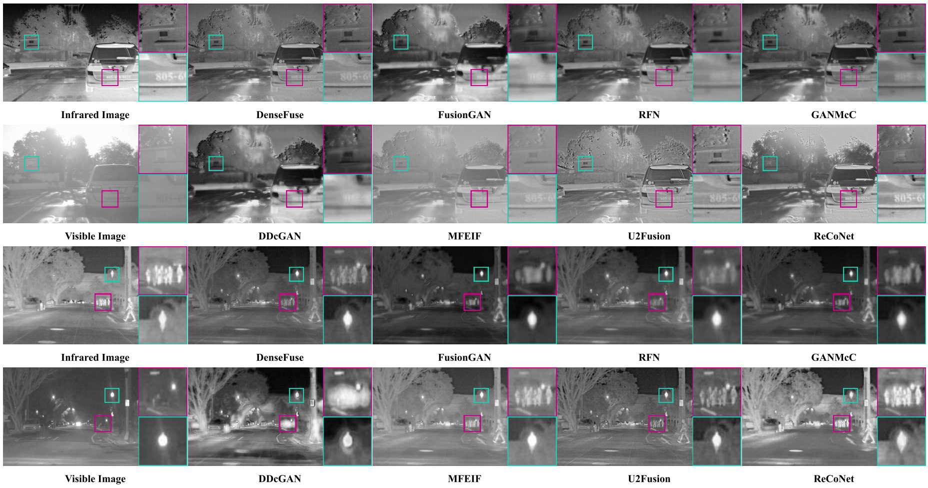

Fig. 2. Visual results comparison between different methods on *RoadScene* Dataset Best view on screen.

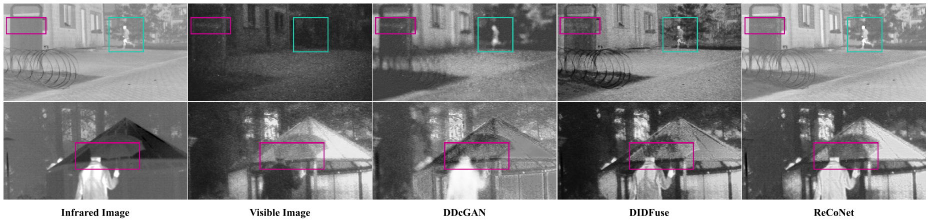

Fig. 3. Visual results comparison of de-noise between two representative methods and our ReCoNet.

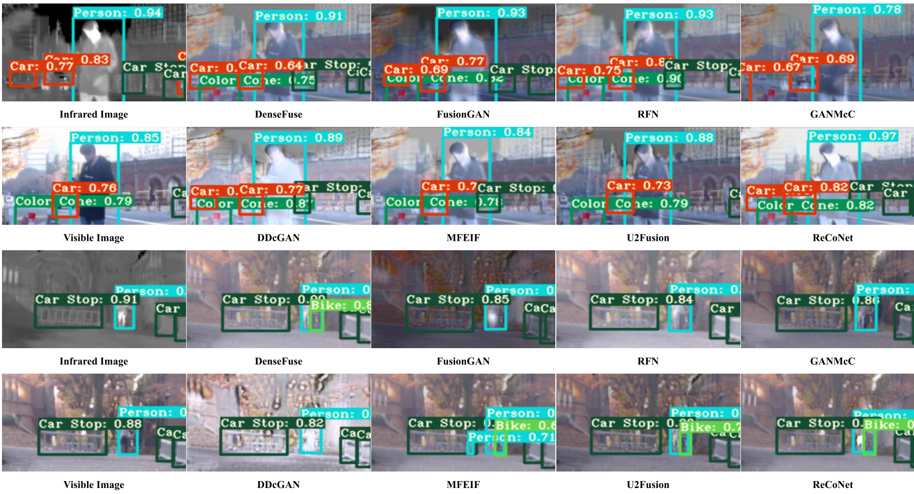

Fig. 4. Visual results comparison of object detection between different methods on *Multi Spectral* dataset, respectively.

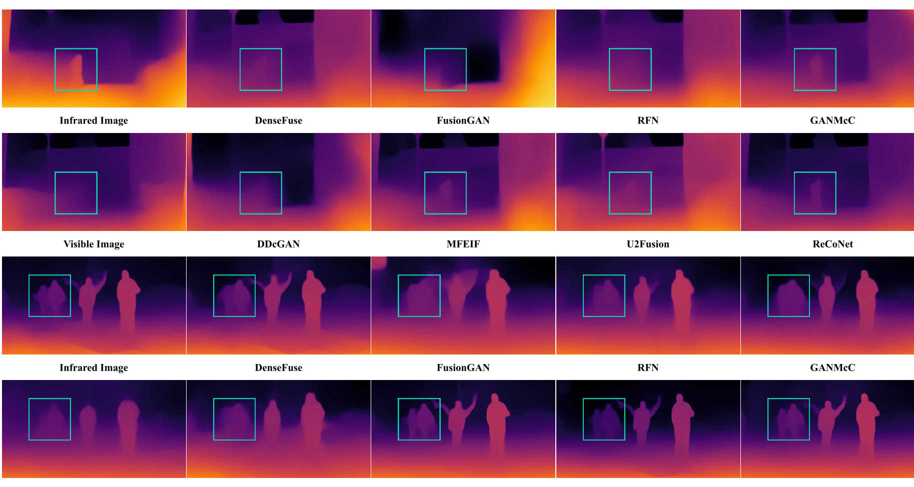

Fig. 5. Visual results comparison of depth estimation between different methods on *Multi Spectral* dataset, respectively.

Fig. 6. Visual results of our ReCoNet on asymmetric image pairs with misaligned. From left to right: infrared image, visible image w/o misaligned, fusion result on visible image w/o misaligned, visible image with misaligned, fusion result on visible image with misaligned.

References

1. Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M.: Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934 (2020)
2. Li, H., Wu, X.J.: Densefuse: A fusion approach to infrared and visible images. IEEE TIP **28**(5), 2614–2623 (2018)
3. Li, H., Wu, X.J., Kittler, J.: Rfn-nest: An end-to-end residual fusion network for infrared and visible images. Information Fusion **73**, 72–86 (2021)
4. Liu, J., Fan, X., Jiang, J., Liu, R., Luo, Z.: Learning a deep multi-scale feature ensemble and an edge-attention guidance for image fusion. IEEE TCSVT (2021)
5. Ma, J., Xu, H., Jiang, J., Mei, X., Zhang, X.P.: Ddcgan: A dual-discriminator conditional generative adversarial network for multi-resolution image fusion. IEEE TIP **29**, 4980–4995 (2020)
6. Ma, J., Yu, W., Liang, P., Li, C., Jiang, J.: Fusiongan: A generative adversarial network for infrared and visible image fusion. Information Fusion **48**, 11–26 (2019)
7. Ma, J., Zhang, H., Shao, Z., Liang, P., Xu, H.: Ganmcc: A generative adversarial network with multiclassification constraints for infrared and visible image fusion. IEEE TIM **70**, 1–14 (2020)
8. Ranftl, R., Lasinger, K., Hafner, D., Schindler, K., Koltun, V.: Towards robust monocular depth estimation: Mixing datasets for zero-shot cross-dataset transfer. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) (2020)
9. Xu, H., Ma, J., Jiang, J., Guo, X., Ling, H.: U2fusion: A unified unsupervised image fusion network. IEEE TPAMI (2020)