УДК 681.78, 004.932

COMPARISON OF TRADITIONAL AND DEEP LEARNING METHODS FOR IMAGE FUSION

Yuandong Shao (ITMO University)
Supervisor – Aleksandr S. Vasilev, Ph.D., Researcher, (ITMO University)

Introduction. A single shooting environment or an image obtained from a single sensor does not provide an effective and comprehensive description of the scene, so we desire to obtain different source images from different sensors. To make the acquired images more convenient for subsequent applications, image fusion technology comes into being. If we summarize image fusion technology in one sentence, it is to extract meaningful information from different source images and combine the extracted information. Currently, the mainstream image fusion techniques are mainly divided into traditional image fusion techniques and deep learning-based image fusion techniques. The main research method of traditional image fusion technology is the relevant mathematical transformation. When using traditional image fusion techniques, to adapt to different fused images and design specific fusion rules, researchers need to analyze the activity level of the image to be fused in the spatial or transformation domain. While in the existing deep learning-based image fusion techniques, researchers mainly focus on designing and improving the algorithmic architecture for the three aspects of feature extraction, feature fusion, and image reconstruction. These two existing mainstream methods face some problems, such as fusion of unaligned images, fusion of images with different resolutions, real-time image fusion, preservation of source image information, preservation of weights for different source image information, etc. So on the one hand, unaligned fusion algorithms can be developed, while specifically combining different resolution images and the characteristics of the image fusion task to design the fusion strategy, and on the other hand, according to the fusion of the results of the actual corresponding loss function, from the decision-making level to guide the fusion process.

Main part. Firstly, we will describe the scenarios where traditional image fusion methods [1] are preferred. For image fusion tasks where the amount of data is not huge, it is better to use traditional image fusion methods with designed targeting, because deep learning image fusion algorithms require a larger amount of data compared to traditional image fusion algorithms. Secondly, for simpler scenarios where the algorithm complexity required for the image fusion task is not high, traditional image fusion methods are also preferred. For some specific scenarios, traditional methods still have their advantages, such as for the need to capture the features of the image to be fused at different scales, Multi-scale Transform-based Methods are more suitable, and some image fusion based on spectral and spatial information is more suitable for the use of Subspace-based Methods. However, for some more complex image processing tasks, or image processing tasks with large amounts of data, the traditional methods are not suitable, so the next section will introduce deep learning-based image fusion methods.

In the field of deep learning-based image fusion, there are different kinds of image fusion methods, which can be mainly divided into autoencoder (AE)-based methods [2,3], conventional convolution neural network (CNN)-based methods [4], and generative adversarial network (GAN)-based methods [5-7], Task-driven methods [8]. deep learning-based image fusion methods mainly improve the corresponding functionality by introducing different architectures. For example, the Dense Fuse method in autoencoder (AE)-based methods, found that in previous image fusion, only a single feature map was used for fusion, and intermediate information was lost, so a Dense block structure is introduced so that the input of each layer is correlated with the output of the previous convolutional layer to reduce the loss of intermediate information. Meanwhile, when the fusion strategy is found to be bad and there are no multi-scale features, the Nest Fuse method can be used to improve the problem by introducing a mesh sampling structure. The method based on the RFN-Nest structure can design the fusion strategy as a learnable architecture. To avoid the manual design of fusion strategies Fusion GAN method can be used. Meanwhile, when using the Fusion GAN method,

it is found that the fused image is only similar to one source image, and the information of the other source image will be lost, we can use the DDcGAN method to input the source image and fused image to the discriminator, and let the discriminator judge the probability that the image is the source image. Further, thinking will find that in the previous fusion method, there is an imbalance in the place that a single omitted the infrared image in the details of the information or omitted the visible image in the contrast information at this point you can use the GANMcC method, the design of a gradient path and a contrast path for the gradient path, the introduction of a primary and secondary ideas to improve the fusion of the image effect.

Conclusions. In conclusion, we can state that there is no perfect image fusion method until now. By comparing traditional image fusion methods and deep learning-based image fusion methods, we can find the advantages and disadvantages of each method and understand the deep meaning of different structures for improving the fused image results. Then, when designing new image fusion methods and fusion strategies, we can carefully adjust the image fusion methods according to the image fusion results. For example, in security systems, we need as much image detail as possible, so we can design gradient paths and contrast paths separately to process the source image group that mainly focuses on visible images and the source image group that mainly focuses on infrared images, to avoid the loss of as little detail as possible in the infrared images and visible images. Finally, we can still pursue higher quality image fusion, and before further designing new image fusion strategies, we need to summarise the previous methods to lay the foundation for designing even better image fusion strategies in the future.

References:

1. Y. Mo, X. Kang, P. Duan, B. Sun, S. Li, Attribute filter based infrared and visible image fusion, Inf. Fusion 75(2021)41-54.
2. Li, Hui, Xiao-Jun Wu, and Tariq Durrani. "NestFuse: An infrared and visible image fusion architecture based on nest connection and spatial/channel attention models." IEEE Transactions on Instrumentation and Measurement 69.12 (2020): 9645-9656.
3. Li, Hui, Xiao-Jun Wu, and Josef Kittler. "RFN-Nest: An end-to-end residual fusion network for infrared and visible images." Information Fusion 73 (2021): 72-86.
4. Xu, Han, et al. "U2Fusion: A unified unsupervised image fusion network." IEEE Transactions on Pattern Analysis and Machine Intelligence 44.1 (2020): 502-518.
5. Ma, Jiayi, et al. "FusionGAN: A generative adversarial network for infrared and visible image fusion." Information fusion 48 (2019): 11-26.
6. Xu, Han, et al. "Learning a Generative Model for Fusing Infrared and Visible Images via Conditional Generative Adversarial Network with Dual Discriminators." IJCAI. 2019.
7. Ma, Jiayi, et al. "GANMcC: A generative adversarial network with multiclassification constraints for infrared and visible image fusion." IEEE Transactions on Instrumentation and Measurement 70 (2020): 1-14.
8. Liu, Jinyuan, et al. "Target-aware dual adversarial learning and a multi-scenario multi-modality benchmark to fuse infrared and visible for object detection." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022.