Rui GUO, Xuan-jing SHEN, Xiao-yu DONG, Xiao-li ZHANG, 2020. Multi-focus image fusion based on fully convolutional networks. *Frontiers of Information Technology & Electronic Engineering*, 21(7):1019-1033. https://doi.org/10.1631/FITEE.1900336

Multi-focus image fusion based on fully convolutional networks

Key words: Multi-focus image fusion; Fully convolutional networks; Skip layer; Performance evaluation

Corresponding author: Xiao-li ZHANG
E-mail: zhangxiaoli@jlu.edu.cn
ORCID: https://orcid.org/0000-0001-8412-4956

Main contributions

1. A new fully convolutional network (FCN) with skip layers, denoted as FD-FCN, is proposed for focus detection.

2. A new training dataset for FD-FCN is constructed based on dataset CIFAR-10.

3. The algorithm can be easily extended to fusing an image sequence.

Method

We detect focused regions using the theory of deep learning. Assume that two source images, denoted as ***A*** and ***B***, are well registered. A new FCN is designed to obtain a focus map. The overall flow diagram of the proposed method can be summarized as follows:

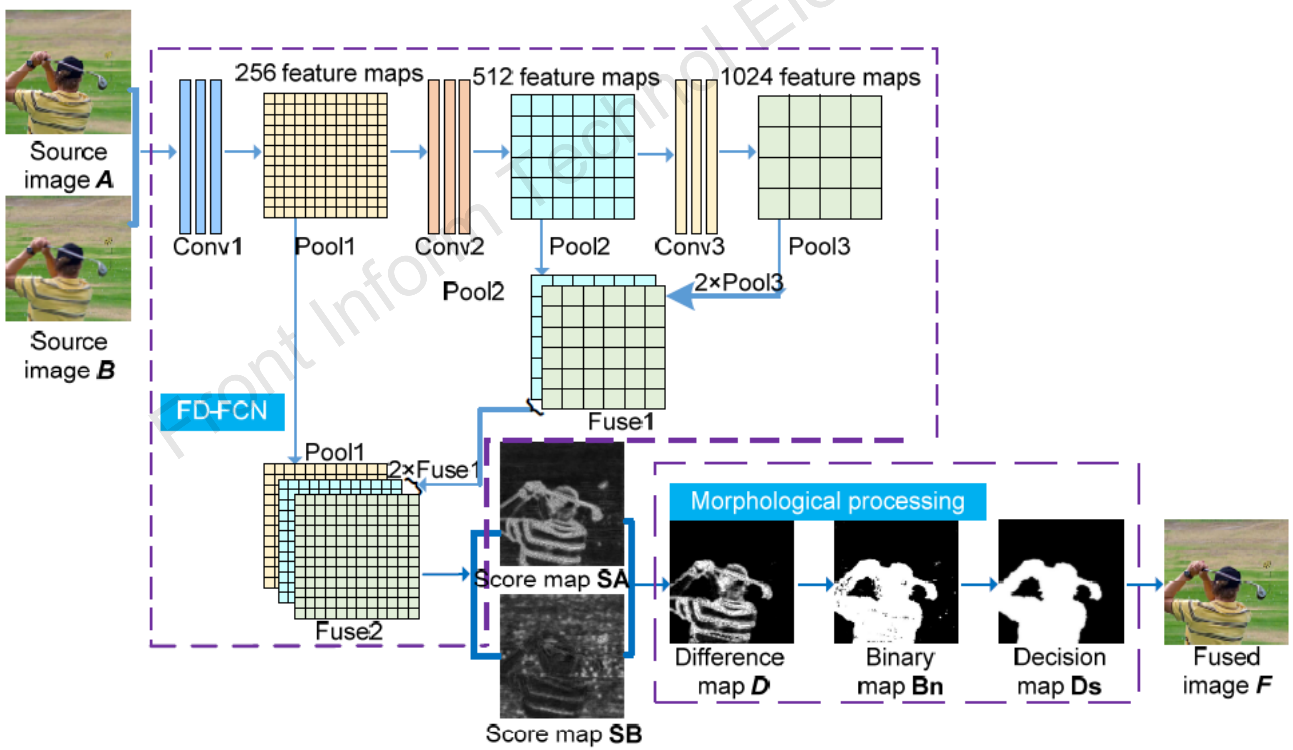

Fig. 1 Overall flow diagram of the proposed method

Method (Cont’d)

1. Input the source images ***A*** and ***B*** into our trained FD-FCN to obtain the focus maps (score maps) of ***A*** and ***B***, which are denoted as **SA** and **SB**, respectively.

2. Conduct a series of morphological processing steps on focus maps to obtain the final decision map **Ds**.

3. Based on **Ds**, the fused image can be obtained by

$\boldsymbol{F}=\mathbf{Ds}\times\boldsymbol{A}+(\mathbf{1}-\mathbf{Ds})\times\boldsymbol{B}$.

Training dataset

When generating a training dataset, we take the original image as the focused region. At the same time, Gaussian blur is applied to all images with a variance of 15, simulating the defocused region. After these processes, our training dataset eventually includes 60 000 clear images and 60 000 blurred images. Because the images in the CIFAR-10 dataset are from the actual image, they contain a lot of information.

Major results

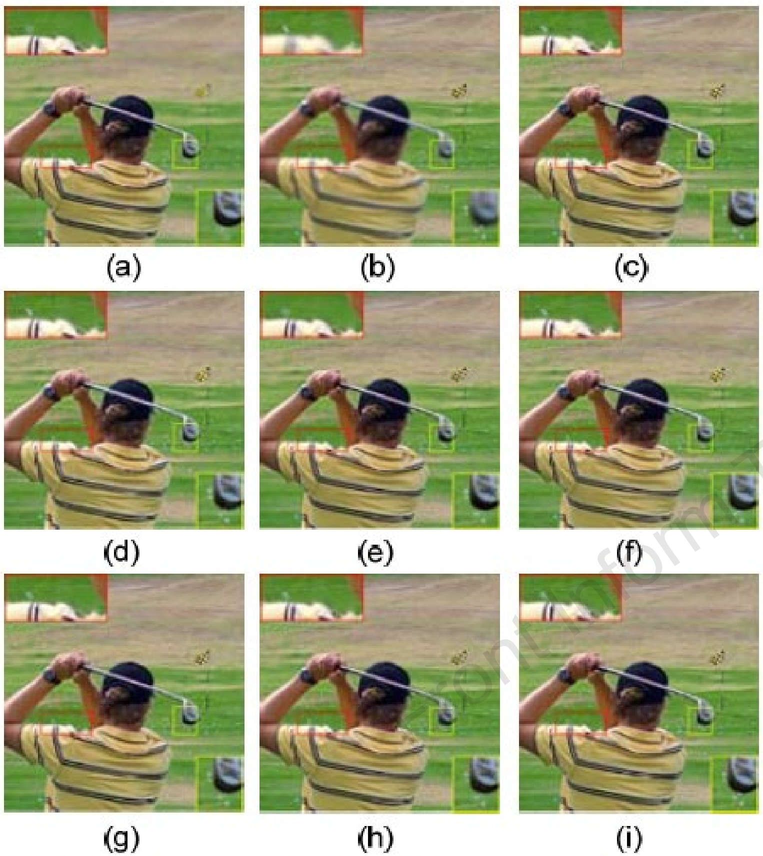

Fig. 5 The first set of colored images "Golf": (a) source image A; (b) source image B; (c) fused image by IMF; (d) fused image by DWT; (e) fused image by MWG; (f) fused image by NSCT; (g) fused image by GF; (h) fused image by CNN; (i) fused image by FD-FCN

Fig. 6 The second set of color images "Child": (a) source image A; (b) source image B; (c) fused image by IMF; (d) fused image by DWT; (e) fused image by MWG; (f) fused image by NSCT; (g) fused image by GF; (h) fused image by CNN; (i) fused image by FD-FCN

Major results (Cont'd)

Table 1 Quantitative evaluation of Fig. 5

Algorithm	MI	Q^F_{AB}	Q_e	Q_R	VIFF
IMF	7.199 840	0.741 351	0.789 912	1.742880	0.584 032
DWT	5.312 529	0.683 448	0.792 999	1.555027	0.490 717
MWG	7.080 758	0.740 648	0.800 766	1.691695	0.611 451
NSCT	6.435 876	0.745 110	0.802 505	1.699236	0.588 021
GF	7.180 822	0.754 383	0.801 949	1.734907	0.610 761
CNN	7.395 021	0.754 614	0.801 807	1.745298	0.613 490
FD-FCN	**7.556 523**	**0.755 076**	**0.802 554**	**1.753 202**	**0.616 000**

The best results are in bold

Table 2 Quantitative evaluation of Fig. 6

Algorithm	MI	Q^F_{AB}	Q_e	Q_R	VIFF
IMF	8.064 029	0.732 390	0.726 478	1.098 585	0.599 546
DWT	5.726 883	0.617 876	0.718 810	0.926 814	0.442 699
MWG	7.741 579	0.729 557	0.727 447	1.094 336	0.598 777
NSCT	6.398 577	0.708 367	0.734 095	1.062 551	0.559 535
GF	7.707 244	0.733 896	0.728 883	1.100 844	0.601 925
CNN	8.104 806	0.735 962	0.728 961	1.103 943	0.605 657
FD-FCN	**8.182 621**	**0.745 621**	**0.738 705**	**1.156 895**	**0.606 512**

The best results are in bold

Major results (Cont'd)

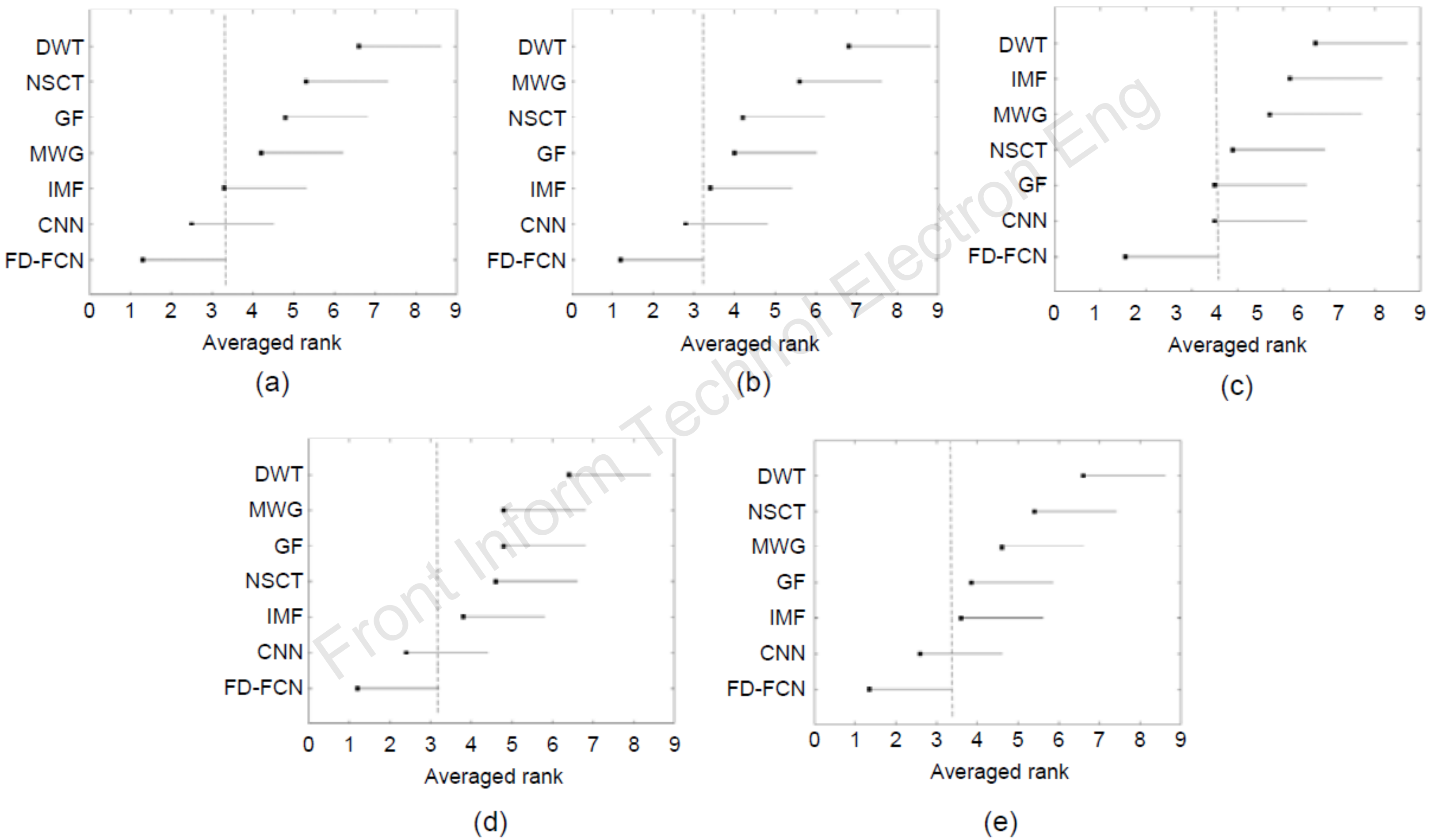

Fig. 9 Results of pairwise comparisons of all image fusion algorithms: (a) MI; (b) Q_{AB}^{F}; (c) Q_e; (d) Q_R; (e) VIFF

Conclusions

1. We have designed a fully convolutional network, FD-FCN, which can be used for multi-focus image fusion.
2. We have conducted both subjective and objective evaluations to compare the performance of the proposed algorithm with those of six state-of-the-art algorithms. The quantitative evaluations of 20 sets of test images showed that the proposed algorithm is the best on the measures MI, Q_e, Q_R, and VIFF.
3. In the future, we will design better-performing networks to obtain more accurate focus maps, or design a network that can directly generate higher level results, such as direct output of fused images.

Rui GUO received the B.E. and M.E. degrees from Jilin University, in 2010 and 2013, respectively. He is currently pursuing the Ph.D. degree at the College of Computer Science and Technology. His research interests include image fusion and machine learning.

Xiao-li ZHANG received the B.S. degree in Computer Science and Engineering from Dalian Nationalities University, in 2010, and the Ph.D. degree in Computer Science and Technology from Jilin University, in 2016. He works as a Research Assistant with the College of Computer Science and Technology, Jilin University. His research interests include machine learning and information fusion. He has authored more than 40 research articles.