













on Pattern Analysis & Machine Intelligence, 2015, 37(9):1904-1916.[DOI: 10.1109/TPAMI.2015.2389824]

[3] Girshick R. Fast R-CNN[J]. Computer Science, 2015: 1440-1448. [DOI: 10.1109/ICCV.2015.169]

[4] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149. [DOI: 10.1109/TPAMI.2016.2577031]

[5] Dai, Jifeng, Li, Yi, He, Kaiming, Sun, Jian. R-FCN: Object Detection via Region-based Fully Convolutional Networks[M]// D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, R. Garnett. Advances in Neural Information Processing Systems: 29. USA: Curran Associates, Inc. 2016:379--387.

[6] Redmon J, Divvala S, Girshick R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016:779-788.[DOI: 10.1109/CVPR.2016.91]

[7] Redmon J, Farhadi A. YOLO9000: Better, Faster, Stronger[C]// IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA:IEEE, 2017:6517-6525.[DOI: 10.1109/CVPR.2017.690]

[8] Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[C]// European Conference on Computer Vision. Cham: Springer International Publishing, 2016:21-37. [DOI: 10.1109/CVPR.2017.690]

[9] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation[M]// Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015. Springer International Publishing, 2015:234-241.[DOI: 10.1007/978-3-319-24574-4\_28]

[10] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society. Boston, MA, USA:IEEE, 2015:3431-3440.[DOI: 10.1109/CVPR.2015.7298965]

[11] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]// International Conference on Neural Information Processing Systems. USA: Curran Associates Inc, 2012, 60(2):1097-1105.

[12] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer Science, 2014.

[13] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]// Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016:770-778.[DOI: 10.1109/CVPR.2016.90]

[14] Badrinarayanan V, Kendall A, Cipolla R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, PP

(99):1-1.[DOI: 10.1109/TPAMI.2016.2644615]

[15] Zhao H, Shi J, Qi X, et al. Pyramid Scene Parsing Network[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society. Honolulu, HI, USA: IEEE, 2017:6230-6239.[DOI: 10.1109/CVPR.2017.660]

[16] Chen L C, Papandreou G, Kokkinos I, et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2018, 40(4):834-848.[DOI: 10.1109/TPAMI.2017.2699184]

[17] Lin G, Milan A, Shen C, et al. RefineNet: Multi-Path Refinement Networks for High-Resolution Semantic Segmentation[C]// IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017:5168-5177.[DOI: 10.1109/CVPR.2017.549]

[18] Bulò S R, Neuhold G, Kotschieder P. Loss. Max-Pooling for Semantic Image Segmentation[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017:7082-7091.[DOI: 10.1109/CVPR.2017.749]

[19] Lin T Y, Dollar P, Girshick R, et al. Feature Pyramid Networks for Object Detection[C]// IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017:936-944.[DOI: 10.1109/CVPR.2017.106]

[20] He K, Gkioxari G, Dollar P, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, PP(99):1-1. [DOI: 10.1109/TPAMI.2018.2844175]

[21] Ren S, He K, Girshick R, et al. Object Detection Networks on Convolutional Feature Maps[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence. 2017, 39(7):1476-1481.[DOI: 10.1109/TPAMI.2016.2601099]

[22] Lin T Y, Goyal P, Girshick R, et al. Focal Loss for Dense Object Detection[C]// IEEE International Conference on Computer Vision. IEEE Computer Society. Venice, Italy: IEEE, 2017:2999-3007.[DOI: 10.1109/ICCV.2017.324]

[23] Everingham M, Gool L V, Williams C K I, et al. The Pascal Visual Object Classes (VOC) Challenge[J]. International Journal of Computer Vision, 2010, 88(2):303-338. [DOI: 10.1007/s11263-009-0275-4]

[24] Shu X, Qi G J, Tang J, et al. Weakly-Shared Deep Transfer Networks for Heterogeneous-Domain Knowledge Propagation[C]// ACM International Conference on Multimedia. New York, NY, USA: ACM, 2015:35-44.[DOI: 10.1145/2733373.2806216]

[25] Lin T Y, Dollar P, Girshick R, et al. Feature Pyramid Networks for Object Detection[C]// IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, 2017:936-944.[DOI: 10.1109/CVPR.2017.106]



#### 第一作者简介:

肖锋(1993-), 男, 硕士研究生, 主要研究方向为深度学习、计算机视觉。  
E-mail:1193221332@qq.com。



#### 通信作者:

芮挺, 副教授, 主要研究方向为机器学习、计算机视觉。  
E-mail:rtinguu@sohu.com。

任桐炜, 男, 教授, 主要研究方向为图像分析与视觉显著性, E-mail:rentw@nju.edu.cn。

王东, 男, 讲师, 主要研究方向为智能信号处理, E-mail:dyhkxydfbb@163.com。

