

我们将图像缩放到 224×224 , 使用的是预训练的 VGG16 网络, 得到 conv5_3 卷积层输出的 14×14 特征图, 并分别使用 sum pooling、SPOC 和深度先验方法提取特征, 最终特征的维度为 512 维。可以看到, 实验结果都分别低于表 1 中的结果, 说明对于较高分辨率的图像而言, 在输入时进行缩放 (主要是缩小) 会使 CNN 图像特征的表达能力大幅下降, 一方面是缩小到固定尺寸破坏了原有图像的空间信息, 同时大尺度地缩小图像使得图像的细节信息被模糊。

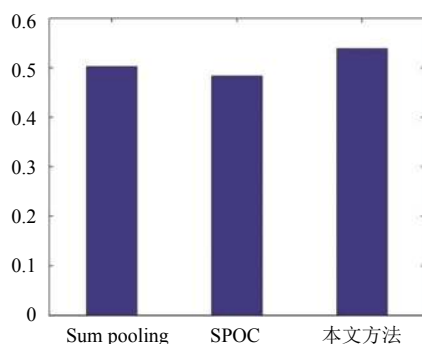


图6 对输入图像缩放的检索结果

5 结论

本文提出了基于深度先验的图像特征提取算法, 该算法可以有效聚焦近处景物特征。而图像检索实验验证了该特征能有效表征室外照片, 可以被应用于其他相关图像任务中。同时, 我们发现在图像特征提取过程中, 原始图像输入对特征的代表能力有极大地提高, 并且再一次验证了 CNN 模型的良好泛化能力和表征能力。

参考文献

- 1 宋维静, 刘鹏, 王力哲, 等. 遥感大数据的智能处理: 现状与挑战. 工程研究-跨学科视野中的工程, 2014, (3): 259-265.
- 2 Turner BL, Skole D, Sanderson S, *et al.* Land-use and land-cover change: Science/research plan. Global Change Report, 1995, 43(1995): 669-679.
- 3 Chi M, Sun Z, Qin Y, *et al.* A novel methodology to label urban remote sensing images based on location-based social media photos. Proceedings of the IEEE, 2017, 105(10): 1926-1936. [doi: 10.1109/JPROC.2017.2730585]
- 4 Lowe DG. Object recognition from local scale-invariant features. Proceedings of the Seventh IEEE International Conference on Computer Vision. IEEE. 1999. 1150-1157.

- [doi: 10.1109/ICCV.1999.790410]
- 5 Bay H, Tuytelaars T, Van Gool L. Surf: Speeded up robust features. Computer Vision-ECCV 2006. Gray, Austria. 2006. 404-417.
- 6 Dalal N, Triggs B. Histograms of oriented gradients for human detection. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). San Diego, CA, USA. 2005. 886-893. [doi: 10.1109/CVPR.2005.177]
- 7 Jegou H, Perronnin F, Douze M, *et al.* Aggregating local image descriptors into compact codes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(9): 1704-1716. [doi: 10.1109/TPAMI.2011.235]
- 8 Perronnin F, Dance C. Fisher kernels on visual vocabularies for image categorization. 2007 IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, MN, USA. 2007. 1-8. [doi: 10.1109/CVPR.2007.383266]
- 9 Arandjelovic R, Zisserman A. All about VLAD. Computer Vision and Pattern Recognition, 2013: 1578-1585.
- 10 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Communications of the ACM, 2013, 60(2): 2012.
- 11 Razavian AS, Azizpour H, Sullivan J, *et al.* CNN features off-the-shelf: An astounding baseline for recognition. Computer Vision and Pattern Recognition, 2014: 512-519.
- 12 Girshick R, Donahue J, Darrell T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA. 2014. 580-587. [doi: 10.1109/CVPR.2014.81]
- 13 Babenko A, Slesarev AV, Chigorin A, *et al.* Neural codes for image retrieval. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T, eds. Computer Vision - ECCV 2014. Lecture Notes in Computer Science, vol 8689. Springer. Cham. 2014. 584-599. [doi: 10.1007/978-3-319-10590-1_38]
- 14 Jegou H, Chum O. Negative evidences and co-occurrences in image retrieval: The benefit of PCA and whitening. European Conference on Computer Vision. Lecture Notes on Computer Science, vol 7573. Springer. Berlin, Heidelberg. 2012. 774-787. [doi: 10.1007/978-3-642-33709-3_55]
- 15 Gong Y, Wang L, Guo R, *et al.* Multi-scale orderless pooling of deep convolutional activation features. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T, eds. Computer Vision - ECCV 2014. Lecture Notes in Computer Science, vol 8695. Springer, Cham. 2014. 392-407. [doi: 10.1007/978-3-319-10584-0_26]

- 16 Ng JY, Yang F, Davis LS, *et al.* Exploiting local features from deep networks for image retrieval. 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Boston, MA, USA. 2015. 53–61. [doi: [10.1109/CVPRW.2015.7301272](https://doi.org/10.1109/CVPRW.2015.7301272)]
- 17 Tolias G, Sicre R, Jégou H. Particular object retrieval with integral max-pooling of CNN activations. arXiv preprint arXiv, 2015: 1511.05879.
- 18 Babenko A, Lempitsky V. Aggregating local deep features for image retrieval. Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile. 2015. 1269–1277. [doi: [10.1109/ICCV.2015.150](https://doi.org/10.1109/ICCV.2015.150)]
- 19 Saxena A, Chung SH, Ng AY, *et al.* Learning depth from single monocular images. Neural Information Processing Systems, 2006: 1161–1168.
- 20 Saxena A, Sun M, Ng A. Make3D: Learning 3D scene structure from a single still image. IEEE TPAMI, 2009, 31: 824–840. [doi: [10.1109/TPAMI.2008.132](https://doi.org/10.1109/TPAMI.2008.132)]
- 21 Liu B, Gould S, Koller D, *et al.* Single image depth estimation from predicted semantic labels. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, CA, USA. 2010. 1253–1260. [doi: [10.1109/CVPR.2010.5539823](https://doi.org/10.1109/CVPR.2010.5539823)]
- 22 Karsch K, Liu C, Kang SB, *et al.* Depth extraction from video using non-parametric sampling. In: Fitzgibbon A, Lazebnik S, Perona P, *et al.*, eds. Computer Vision – ECCV 2012. Lecture Notes in Computer Science, vol 7576. Springer. Berlin, Heidelberg. 2012. 775–788. [doi: [10.1007/978-3-642-33715-4_56](https://doi.org/10.1007/978-3-642-33715-4_56)]
- 23 Konrad J, Wang M, Ishwar P, *et al.* 2D-to-3D image conversion by learning depth from examples. 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. Providence, RI, USA. 2012. 16–22. [doi: [10.1109/CVPRW.2012.6238903](https://doi.org/10.1109/CVPRW.2012.6238903)]
- 24 Eigen D, Puhrsch C, Fergus R, *et al.* Depth map prediction from a single image using a multi-scale deep network. Neural Information Processing Systems, 2014: 2366–2374.
- 25 Liu F, Shen C, Lin G, *et al.* Deep convolutional neural fields for depth estimation from a single image. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA. 2015. 5162–5170. [doi: [10.1109/CVPR.2015.7299152](https://doi.org/10.1109/CVPR.2015.7299152)]
- 26 Philbin J, Chum O, Isard M, *et al.* Object retrieval with large vocabularies and fast spatial matching. 2007 IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, MN, USA. 2007. 1–8. [doi: [10.1109/CVPR.2007.383172](https://doi.org/10.1109/CVPR.2007.383172)]
- 27 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv, 2015: 1409.1556.
- 28 Deng J, Dong W, Socher R, *et al.* ImageNet: A large-scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, FL, USA. 2009. 248–255. [doi: [10.1109/CVPR.2009.5206848](https://doi.org/10.1109/CVPR.2009.5206848)]
- 29 He K, Zhang X, Ren S, *et al.* Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904–1916. [doi: [10.1109/TPAMI.2015.2389824](https://doi.org/10.1109/TPAMI.2015.2389824)]