SCIENCE CHINA
Information Sciences

• LETTER •

September 2019, Vol. 62 199101:1–199101:2
https://doi.org/10.1007/s11432-018-9639-7

Joint horizontal and vertical deep learning feature for vehicle re-identification

Jianqing ZHU[1], Huanqiang ZENG[2*], Xin JIN[3*], Yongzhao DU[1],
Lixin ZHENG[1] & Canhui CAI[1]

[1]*College of Engineering, Huaqiao University, Quanzhou 362021, China;*
[2]*College of Information Science and Engineering, Huaqiao University, Xiamen 361021, China;*
[3]*Department of Computer Science and Technology, Beijing Electronic Science and Technology Institute, Beijing 100070, China*

Received 27 March 2018/Revised 31 July 2018/Accepted 11 October 2018/Published online 12 June 2019

Citation Zhu J Q, Zeng H Q, Jin X, et al. Joint horizontal and vertical deep learning feature for vehicle re-identification. Sci China Inf Sci, 2019, 62(9): 199101, https://doi.org/10.1007/s11432-018-9639-7

Dear editor,
Vehicle re-identification is a procedure in which a query image of a vehicle can be matched to vehicle images captured by different cameras in a data gallery. This procedure is extremely important for the improvement of public safety; however, accurately re-identifying vehicles is a very challenging computer vision task because vehicle images are captured from different camera view-points and they usually contain large variations in appearances. Therefore, the development of an effective vehicle re-identification method is both challenging and socially meaningful.

Before recognition, images are usually represented by features (e.g., texture [1], sparse coding [2], deep learning [3]) to resist some adverse factors (e.g., illumination changes, deformations, noise). In re-identification field, the view-point variation is a key challenge. To address the view-point variation issue, a common strategy used for the person re-identification tasks is to model the horizontal symmetry of a pedestrian in either hand-crafted feature representation methods (e.g., ELF [4] and LOMO [5]) or deep feature learning methods (e.g., horizontal pooling module [3, 6]). These methods have been specifically developed for person re-identification, and they may not be suitable for direct application in vehicle re-identification tasks because the symmetry characteristics of vehicles are quite different from those of pedestrians. Furthermore, vehicle re-identification procedures should include the use of both horizontal and vertical pooling modules because vehicles contain not only strong left-right symmetry but also considerable up-down symmetry. Efficient use of these modules to describe vehicle images in both directions would improve the analysis of vehicle images in terms of view-point variations, thereby enhancing their performance.

Method. Based on these factors, an effective deep learning feature for vehicle re-identification, called the joint horizontal and vertical deep learning feature (JHV-DLF), is proposed herein to describe vehicle images in both horizontal and vertical directions, which makes re-identification robust toward view-point variations. As shown in Figure 1, the proposed method first exploits a shortly and densely connected convolutional neural network [7] as a basic deep feature learning module (B-DFLM) on the square input vehicle image to extract a $d \times d$ basic feature map. Then, a horizontal and vertical deep feature learning module (HV-DFLM) is used to jointly compress the basic $d \times d$ feature maps into $d \times 1$ and $1 \times d$ directional feature maps corresponding to the horizontal and vertical directions. Finally, these two directional feature maps are concatenated as the final feature for vehicle re-identification.

* Corresponding author (email: zeng0043@hqu.edu.cn, jinxinbesti@foxmail.com)

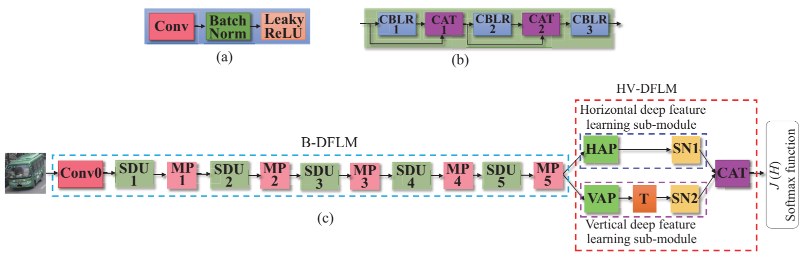

Figure 1 (Color online) Diagram for the proposed method. The acronyms MP, HAP, VAP, T, SN and CAT denote max pooling, horizontal average pooling, vertical average pooling, transposition, spatial normalization and concatenation layers, respectively. (a) The packaged block of convolutional layer, batch normalization and leaky ReLU layers (CBLR) block; (b) the short and dense unit (SDU); (c) the framework of joint horizontal and vertical deep feature learning.

Experiment. Our study presents performance comparisons between the proposed JHV-DLF and representative state-of-the-art methods (i.e., LOMO [5], BOW-SFIT [8], PROVID [9], NuFACT + Plate-SNN [9] and NuFACT + Plate-REC [9]) on the vehicle re-identification (VeRi) [9] database. Among all compared methods, the JHV-DLF method acquired the highest rank-1 identification rate, 84.74%.

By maintaining the B-DFLM as unchanged, we further comprehensively analyzed the contributions of the proposed HV-DFLM on performance improvement. The features learned in the configurations that use the HV-DFLM, only the horizontal deep feature learning sub-module, or only the vertical deep feature learning sub-module, are denoted as JHV-DLF, H-DLF, and V-DLF, respectively.

The performance of the proposed JHV-DLF on the VeRi database was superior to that of both H-DLF and V-DLF. In particular, the mean average precision (MAP) of JHV-DLF is 2.62% and 3.03% higher than that of H-DLF and V-DLF, respectively. Moreover, the rank-1 identification rate of JHV-DLF is 1.61% and 3.57% higher than that of H-DLF and V-DLF, respectively. This demonstrates that the use of JHV-DLF that comprehensively describes vehicles in both the horizontal and vertical directions is beneficial to the robustness of camera view-point variations, which allows for better performance.

Conclusion. This study presents an effective JHV-DLF for vehicle re-identification. Using the proposed HV-DFLM, vehicles can be comprehensively described in both the horizontal and vertical directions, which provides robustness against camera view-point variations. Our experiments show that the proposed JHV-DLF is superior to multiple state-of-the-art vehicle re-identification methods on the VeRi database.

Acknowledgements This work was supported in part by National Natural Science Foundation of China (Grant Nos. 61602191, 61871434, 61802136, 61672521), in part by Natural Science Foundation of Fujian Province (Grant Nos. 2018J01090, 2016J01308), in part by Promotion Program for Young and Middle-aged Teacher in Science and Technology Research of Huaqiao University (Grant Nos. ZQN-PY418, ZQN-YX403), and in part by Scientific Research Funds of Huaqiao University (Grant No. 16BS108).

Supporting information Detailed parameter configuration of the proposed method and performance comparison results. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Jin X, Zhu S Y, Xiao C E, et al. 3D textured model encryption via 3D Lu chaotic mapping. Sci China Inf Sci, 2017, 60: 122107
2 Guo L H, Guo C G, Li L, et al. Two-stage local constrained sparse coding for fine-grained visual categorization. Sci China Inf Sci, 2018, 61: 018104
3 Zhu J Q, Zeng H Q, Liao S C, et al. Deep hybrid similarity learning for person re-identification. IEEE Trans Circ Syst Video Technol, 2018, 28: 3183–3193
4 Chen Y C, Zheng W S, Lai J H, et al. An asymmetric distance model for cross-view feature mapping in person re-identification. IEEE Trans Circ Syst Video Technol, 2017, 27: 1661–1675
5 Liao S C, Hu Y, Zhu X Y, et al. Person re-identification by local maximal occurrence representation and metric learning. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Boston, 2015. 2197–2206
6 Zhu J Q, Zeng H Q, Du Y Z, et al. Person re-identification based on novel triplet convolutional neural network. J Electron Inf Technol, 2018, 40: 1012–1016
7 Zhu J Q, Zeng H Q, Lei Z, et al. A shortly and densely connected convolutional neural network for vehicle re-identification. In: Proceedings of International Conference on Pattern Recognition, Beijing, 2018
8 Zheng L, Wang S J, Zhou W G, et al. Bayes merging of multiple vocabularies for scalable image retrieval. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Columbus, 2014. 1963–1970
9 Liu X C, Liu W, Mei T, et al. PROVID: progressive and multimodal vehicle reidentification for large-scale urban surveillance. IEEE Trans Multimedia, 2018, 20: 645–658