Deep feature extraction based on an L2-constrained combination of center and softmax loss functions for ukiyo-e image recommendation

Lilang Xiong, Zhenao Wei, Wenwen Ouyang, Yulin Cai, Ruck Thawonmas,
Keiko Suzuki, and Masaaki Kidachi
Ritsumeikan University

VGG-19 is a high-performance CNN architecture [1] that has 19 layers, and in this work, we use the features extracted from its conv5_1 layer to represent the image style. According to Gatys et al. [2], the image style features extracted from the conv5_1 layer are better than others in image style transfer applications. They extracted an image style from an art painting and applied it to another image called content image so that a newly generated image has both content and style features. Based on our previous research [3], our results indicated that the way people judge whether two images are similar is based more on image styles than actual objects, or content, in ukiyo-e images. The dataset we used is a set of ten thousand ukiyo-e images from a database made publicly available by the Art Research Center (ARC), Ritsumeikan University. The 10,000 ukiyo-e images are separated into five categories which contain 2,000 images each.

We focus on ukiyo-e image recommendation. In our subsequent previous work [4], we examined different image-style feature-extraction methods for the ukiyo-e dataset using the nearest-neighbor algorithm (1-NN). 1-NN was used because the classification result of 1-NN indicates the features' ability in representing the image similarity used in a recommender. Wen et al. [5] proposed a new loss function called *center loss* that squeezes the distances of all data in the same class. More recently, L2-constrained softmax [6] was proposed that it is able to increase the classification performance by normalizing the features.

Our present work is about feature extraction of image styles from VGG-19 based on an L2-constrained combination of center and softmax loss functions for ukiyo-e image recommendation. 1-NN is used for evaluation of the features' quality. Experimental results on the aforementioned ukiyo-e data set show that the performance of the resulting features outperform features extracted by existing methods.

References

[1] Karen Simonyan and Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[2] Leon A Gatys, Alexander S Ecker, and Matthias Bethge. 2016. Image style transfer using convolutional neural networks. In Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on, 2414–2423.

[3] Zhenao Wei, Lilang Xiong, Kazuki Mori, Tung Duc Nguyen, Tomohiro Harada, Ruck Thawonmas, Keiko Suzuki, and Masaaki Kidachi. Deep Features for Image Classification and Image Similarity Perception. JADH 2017: 60.

[4] Lilang Xiong, Zhenao Wei, Wenwen Ouyang, Tomohiro Harada, Ruck Thawonmas,Keiko Suzuki, and Masaaki Kidachi. 2018. Validation of Deep Features Using the 1-NN Algorithm for Image Similarity Computation. In 2018 NICOGRAPH International, Tainan.

[5] Yandong Wen, Kaipeng Zhang, Zhifeng Li, and Yu Qiao. 2016. A discriminative feature learning approach for deep face recognition. In European Conference on Computer Vision, 499–515.

[6] Rajeev Ranjan, Carlos D Castillo, and Rama Chellappa. 2017. L2-constrained softmax loss for discriminative face verification. arXiv preprint arXiv:1703.09507.