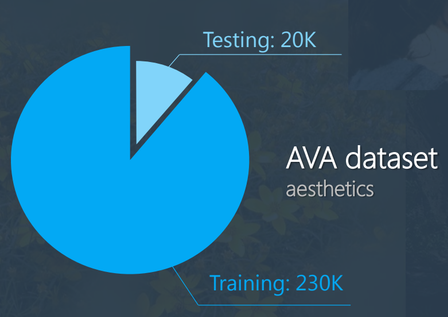
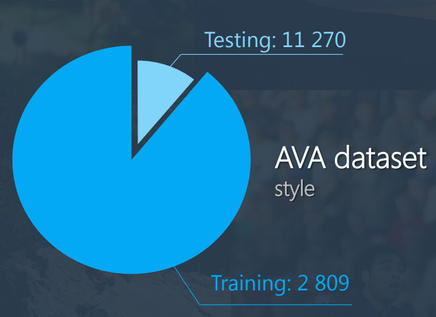
Filling the GAPs: Reducing the Complexity of Networks for Multi-attribute Image Aesthetic Prediction

Abstract. Computational aesthetics have seen much progress in recent years with the increasing popularity of deep learning methods. In this paper, we present two approaches that leverage on the benefits of using Global Average Pooling (GAP) to reduce the complexity of deep convolutional neural networks. The first model fine-tunes a standard CNN with a newly introduced GAP layer. The second approach extracts global and local CNN codes by reducing the dimensionality of convolution layers with individual GAP operations. We also extend these approaches to a multi-attribute network which uses a style network to regularize the aesthetic network. Experiments demonstrate the capability of attaining comparable accuracy results while reducing training complexity substantially.

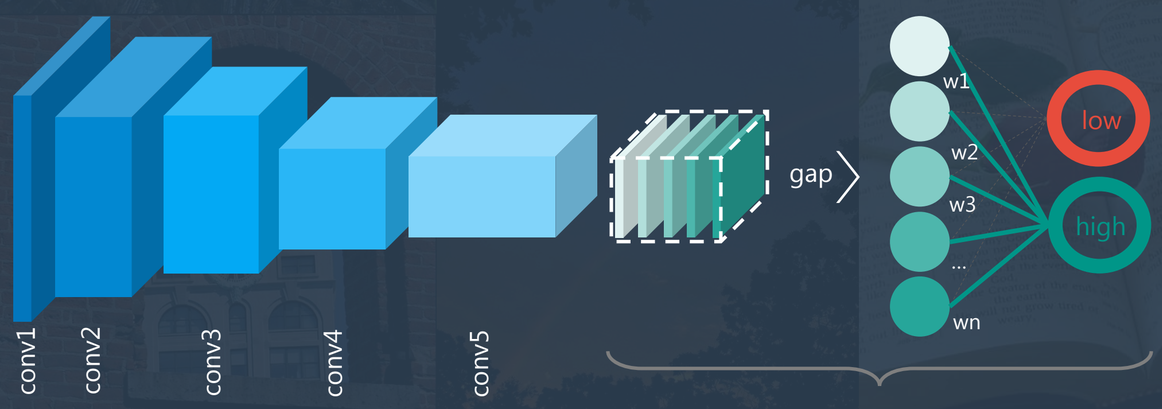
Dataset





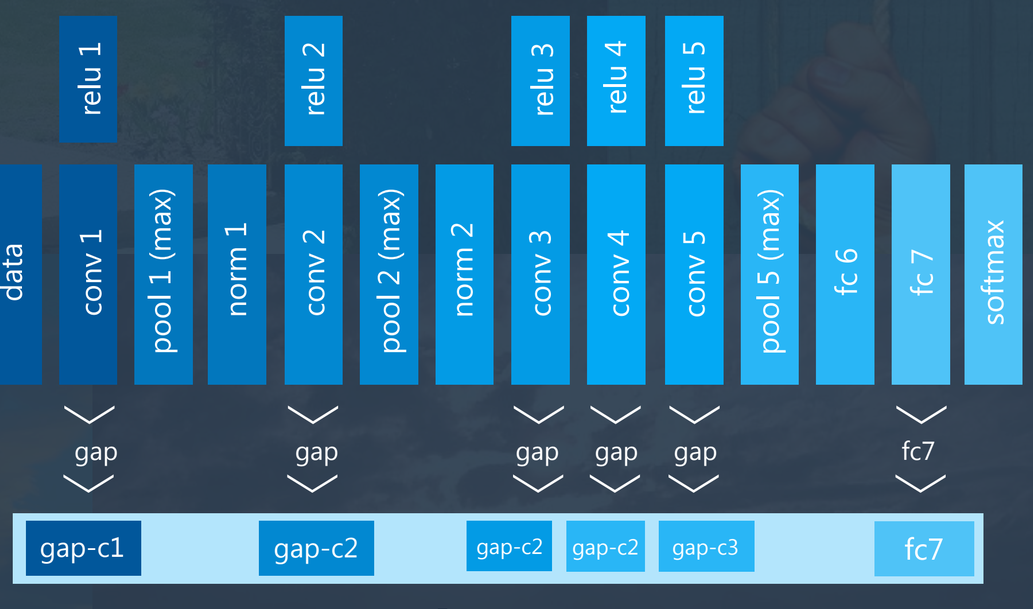
Ground truth data (human ratings ranging from **1** to **10**)

Single Column Architecture (AnGAP-Finetuned)

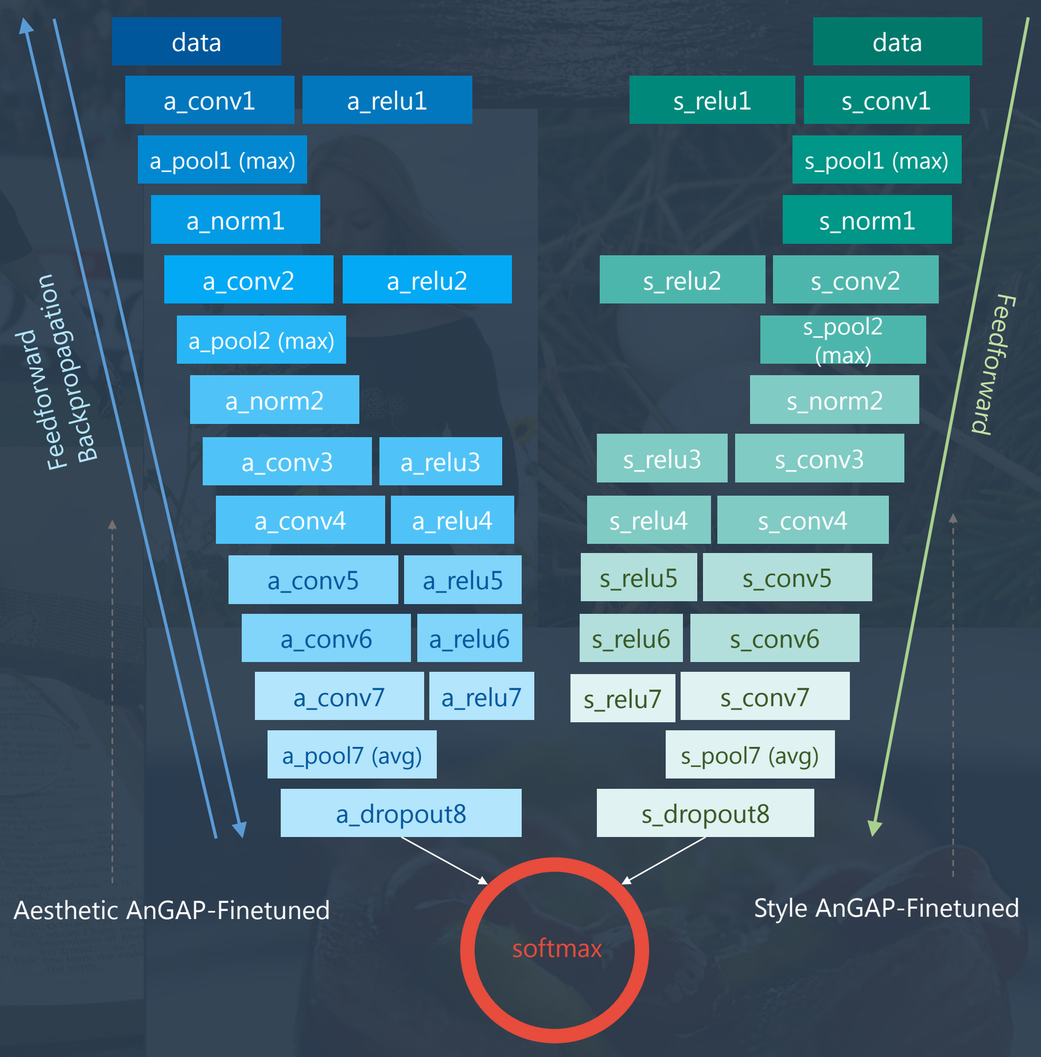


Class Activation Map (CAM)

Single Column Architecture (AnGAP-FeatEns)



Multi-attribute Net



Results

Method	Accuracy (%)	# params
AVA Baseline [1]	68.00	-
SPP [2]	72.85	-
DCNN [3]	73.25	-
RDCNN-style [3]	74.46	-
AnGAP-Finetuned	**74.84**	**~4K**
AlexNet-Finetuned	**75.13**	**~56K**
Multi Att. AnGAP-Finetuned	**75.16**	**~8K**
DMA-Net [2]	75.41	-
RDCNN semantic [4]	75.42	-
AnGAP-FeatEns.	**76.07**	**~56K**
Multi Att. AnGAP-FeatEns.	**76.32**	**~112K**

References

[1] Naila Murray, Luca Marchesotti, and Florent Perronnin, "Ava: A large-scale database for aesthetic visual analysis," in Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012, pp. 2408–2415.
[2] Xin Lu, Zhe Lin, Xiaohui Shen, Radomir Mech, and James Z Wang, "Deep multi-patch aggregation network for image style, aesthetics, and quality estimation," in Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 990–998.
[3] Xin Lu, Zhe Lin, Hailin Jin, Jianchao Yang, and James Z Wang, "Rapid: Rating pictorial aesthetics using deep learning," in Proceedings of the 22nd ACM international conference on Multimedia. ACM, 2014, pp. 457– 466.
[4] Xin Lu, Zhe Lin, Hailin Jin, Jianchao Yang, and James Z Wang, "Rating image aesthetics using deep learning," IEEE Transactions on Multimedia, vol. 17, no. 11, pp. 2021–2034, 2015.

Magzhan Kairanbay, John See, Lai-Kuan Wong, Yong-Lian Hii
Multimedia University, Center for Visual Computing, Faculty of Computing and Informatics, 63100 Cyberjaya, Malaysia, http://viprlab.github.io
magzhan.kairanbay@gmail.com, johnsee@mmu.edu.my,lkwong@mmu.edu.my, yonglian146@gmail.com

ICIP 2017
IEEE International Conference on Image Processing