Data Covariance Learning in Aesthetic Attributes Assessment

Zhihong Chen

School of Mathematics, South China University of Technology, Guangzhou, China
Email: chenzhihong12341234@gmail.com

Abstract

Image Aesthetic Assessment (IAA) is a widely considered problem given its usefulness in a wide range of applications such as the evaluation image capture pipelines, sharing and storage techniques media, but the intrinsic mechanism of aesthetic evaluation is seldom been explored due to its subjective nature and the lack of interpretability of deep neural networks. Noticing that the photographic style annotations of images (*i.e.* the score of aesthetic attributes) are more objective and interpretable compared with the Mean Opinion Scores (MOS) annotations for IAA, we evaluate the problem of Aesthetic Attributes Assessment (AAA) as to provide complementary information for IAA. We firstly introduce the learning of data covariance in the field of Aesthetic Attributes Assessment and propose a regression model that jointly learns from MOS as well as the score of all aesthetic attributes at the same time. We construct our method by extending the scheme of data uncertainty learning and propose data covariance learning. Our method achieves the state-of-the-art performance on AAA without architectural design that trains in a totally end-to-end manner and can be easily extended to existing IAA methods.

Keywords

Aesthetic Assessment, Aesthetic Attributes Assessment, Data Uncertainty, Data Covariance, Maximum Likelihood, Convolutional Neural Network

1. Introduction

Image Aesthetics Assessment (IAA) aims to predict the aesthetic quality score of a given image. Given the growing amount of digital photography IAA has become increasingly important in many downstream tasks such as image cropping, image aesthetic manipulation, and image search [1]-[7]. Due to the high degree

of subjectivity of the problem [2] [8] and low interpretability of deep neural networks, the intrinsic mechanism of IAA has seldom been explored [7], although many solutions have been proposed.

It has been shown that human aesthetic experience is connected with long-established photographic rules [9], thus it is beneficial to provide style-related information when evaluating the aesthetics of images, which is exactly what Aesthetic Attributes Assessment (AAA) is about. Compared to IAA, which predicts only the overall score of image aesthetics, AAA aims to provide not only the main score but also the score of photographic rules such as Rule of Thirds, Symmetry, Vivid Colors, etc. [10] [11]. Since the evaluation process of photographic rules is more objective than that of the overall aesthetics, the prediction result made by an AAA model can be viewed as a predictable IAA prediction as it contains tractable clues to interpret the result as shown in Figure 1.

Albeit the insightful information provided, AAA is overly underestimated [12]. Recent approaches that address IAA focus only on the overall score (MOS), but few care about what the underlying graphic styles (attributes) are and how they interact with each other (including MOS). Due to the ambiguous nature of aesthetics evaluation, interpretability brought about by these graphic styles (attributes) becomes substantially important. Unfortunately, the majority of proposed IAA approaches only predict one single score (MOS) at inference time, which is unreliable to use.

Since the IAA problem is ambiguous, the constructed dataset is annotated noisily [13] [14] [15] [16]. As AAA introduces more targets to predict (*i.e.* the photographic styles), it should suffer from noise more seriously. To make it even worse, these photographic styles might be intertwined with each other [2] [8] [13] and contribute differently to the overall aesthetic score. These challenges make AAA more complicated.
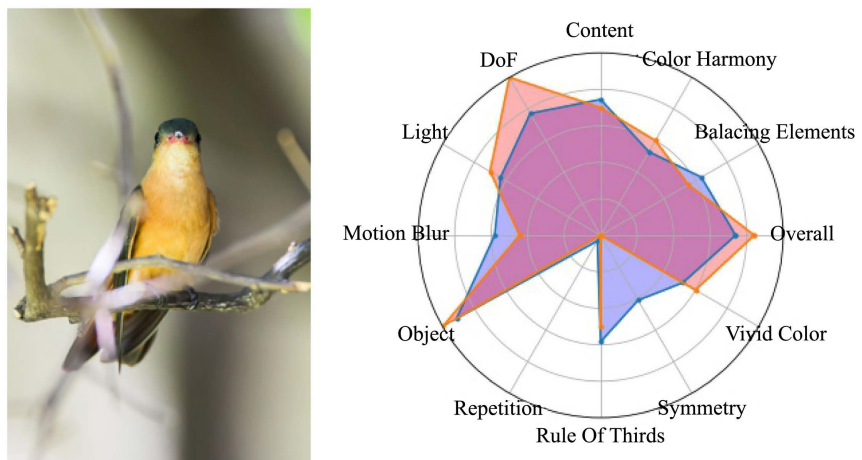
Figure 1. Image evaluation with photographic styles. Scores are normalized to 0 and 1 for illustration. Red represents the ground truth annotation in test set and blue represents prediction made by our model.

As pointed out in [17] modeling data uncertainty is beneficial in noisy data training [18] [19] [20] [21]. However, recent data uncertainty learning methods treat the modeled multivariate noise independent among components for simplicity, which is not an appropriate assumption for intertwined variables like photographic style. Hence we extend the framework of data uncertainty learning to data covariance learning to fit the AAA problem which is achieved by extending the set of independent variances into a covariance matrix. Then the model is trained with cross-entropy with the data covariance setting, which can be viewed as a natural extension of the commonly used mean squared error.

Our method, as shown in Figure 2, is universally extensible to all CNN-based architecture, and reaches the state-of-the-art performance with a standard ResNet-50 backbone without any handcrafted architectural design in the field of AAA. That means our method can inject the discrimination ability of aesthetic attributes for IAA methods that did not care about it before. The effectiveness of our method is verified by experiments. Our contributions lie in four folds:

1) Extend the commonly used cross-entropy loss for covariance learning;

2) Emphasize the importance of AAA as it provides compliment information to IAA without hurting the performance of predicting the overall score;

3) No handcrafted architecture design needed, thus our method can be transplanted to IAA method easily;

4) For the first time, we report that, without any handcrafted feature design, the CNN model is constantly doing better in one attribute called *Vivid Color* than the overall score, which supports our assumption that the photographic style is more objective to evaluate than the overall aesthetics score, which at the same time, unveil an underlying mechanism of CNN that it learns to evaluate photo aesthetics mainly through color.

2. Related Work

A. Aesthetic attributes assessment

There are many works that aim to tackle the IAA problem. Among them, early ones focus on designing handcrafted features [4] [16] [22] [23] [24] and recent works adopt deep neural networks [10] [13] [14] [15] [25]-[30]. However, very few works focus on the photographic styles. [10] is the first to study AAA, and propose a dataset called AADB that is annotated with the overall scores as well as the scores of 11 kinds of meaningful photographic styles. [11] is another one to study AAA, which learns 8 attributes in AADB dataset and develops a visualization method by back propagating the gradient.
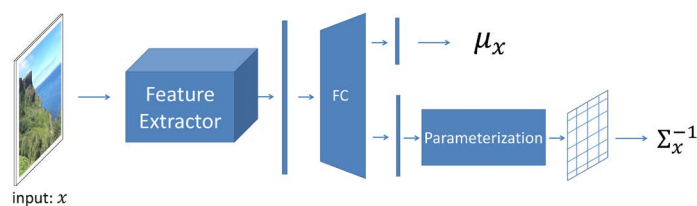
Figure 2. Pipeline of our proposed method.

B. Data uncertainty in deep learning

Modeling data uncertainty is beneficial to deep learning which describes data noise effectively by Gaussian distribution. [17] proposes to model random embeddings for facial recognition in the form of $p(z_i \mid x_i) = N(z_i; \mu_i, \sigma_i^2 I)$, which z_i is the embedding for the i-th sample x_i in the dataset, with a KL divergence term added as regularization to σ_i. The learned model can therefore predict the uncertainty σ_i of data for each cluster μ_i.

C. Data uncertainty in image aesthetics assessment

Modeling data uncertainty is beneficial to aesthetics assessment. [13] finds IAA problem benefit from data uncertainty modeling due to the ambiguity. They propose a unified probabilistic formulation to deal with IAA and introduce Gaussian uncertainty for score (MOS) distribution. However, the interaction among attributes is not concerned.

D. Cross-entropy (CE) serves as strong baseline in deep metric learning

[31] shows that when optimizing, CE can reflect Mutual Information (MI) which takes care of inter- and intra-class distance and the state-of-the-art performance is obtained in deep metric learning dataset with a standard Resnet-50 architecture. This method addresses deep metric learning by CE with label smoothing which can be viewed as indirect regularization to a scalar variance, albeit not explicitly modeled.

3. Algorithm

A. Training Objective

Since for each image in the dataset, the final result of the main score, as well as that of all other attributes, can be viewed as the accumulation from multiple independent human raters which roughly satisfies the prerequisite of big number theory, we can hypothesize that the multi-dimensional label y_i *i.e.* the main score along with the scores of all other attributes, conditioned on its input image x_i, satisfies the multi-variate normal distribution. In order to utilize the full power of data uncertainty modeling, we adopt a general multi-variate normal distribution with a full functioning covariance matrix Σ_i for each input image, *i.e.*

$$y_i \mid x_i \sim N\left(y \mid x, \mu_i, \Sigma_i^{-1}\right). \tag{1}$$

The model is learned by minimizing the conditional cross-entropy between the ground truth label and the model output, we have

$$loss = -\mathbf{E}_{(x_i, y_i) \sim \mathbf{F}(x,y)} \mathbf{I}(y_i \mid x_i) \ln \mathbf{P}(\hat{y}_i \mid x_i), \tag{2}$$

which F is the empirical distribution, $\mathbf{I}$ is the indicator function suggested by the label and $\hat{y}$ is the predicted result of the model to be fit. The above equation can be further simplified as,

$$loss = -\frac{1}{N} \sum_{(x_i, y_i) \sim D} \ln \mathbf{P}(\hat{y}_i \mid x_i), \tag{3}$$

which D stands for the dataset, and the equation is in line with the maximum li-

kelihood estimation of multivariate normal distribution.

In order to improve computation efficiency, we integrate the condition x_i into μ_i and Σ_i^{-1}, *i.e.*

$$y_i \mid x_i \sim \mathbf{N}\left(y \mid \mu_{\theta_1}(x_i), \Sigma_{\theta_2}^{-1}(x_i)\right). \tag{4}$$

Then the cross-entropy becomes:

$$loss = -\frac{1}{N} \sum_{(x_i, y_i) \sim D} \ln \mathbf{P}\left(\hat{y}_i \mid \mu(x_i), \Sigma^{-1}(x_i)\right). \tag{5}$$

Substitute (1) into (5), we have

$$loss = \frac{1}{N} \sum_{(x_i, y_i) \sim D} \left(-\frac{1}{2}\ln\left|\Sigma_i^{-1}\right| + \frac{1}{2}(y_i - \mu_i)^{\mathrm{T}} \Sigma_i^{-1}(y_i - \mu_i)\right), \tag{6}$$

Furthermore, if we set Σ_i to identity matrix, then this loss function degenerates to mean square error.

B. Positive Definiteness of Σ^{-1}

Since the covariance matrix $\Sigma^{-1}(x)$ should be kept symmetric and positive definite throughout the course of training, otherwise the $\ln\left|\Sigma^{-1}\right|$ will be undefined, we need to find a proper parameterization that complies with the constraint and is easy to optimize. Observing that the symmetry and positive definition can be satisfied simultaneously, we propose the following two-step parameterization.

Suppose $a(x)$ is an m^2-dimensional output generated by a neural network, m is the dimension of y_i. We reshape $a(x)$ into A such that $A \in \mathcal{R}^{m \times m}$. Firstly, a symmetric matrix is obtained by:

$$M \leftarrow A^{\mathrm{T}} + A. \tag{7}$$

Then a positive definite matrix is obtained by:

$$S \leftarrow \frac{1}{m}M + I. \tag{8}$$

Since a positive definite matrix is invertible, then we can obtain Σ^{-1} by:

$$\Sigma^{-1} \leftarrow S. \tag{9}$$

We can further obtain Σ simply by calculating the inverse of Σ^{-1} for more flexibility at inference time, but it is not necessary for training. Such arrangement enables computation efficiency without hurting flexibility of the model.

4. Experiments

A. Dataset

The dataset we concern is AADB [10], since it provides eleven meaningful aesthetic attributes in addition to the overall aesthetic score (MOS) which is collected from professional photographers.

B. Implementation details

Our method is trained jointly in an end-to-end manner on MOS and all eleven aesthetic attributes under Alexnet and ResNet-50 following the train-val-test

split in [10]. We use AdamW [32] with warmup under the following parameters: $\beta_1 = 0.9$, $\beta_2 = 0.99$, learning rate of 7e−4, weight decay of 1e−2 and epoch of 10. These parameters are picked as default except for the learning rate, which is picked by the range test [33].

C. Performance evaluation

We compare our method with all existing AAA methods [10] [11]. As is shown in Table 1, our method outperforms previous method on most attributes even with weak backbone like AlexNet. The metric we use to evaluate the performance of different methods is spearman rank order coefficient. Some data examples are shown in Figure 3.

Table 1. Spearman rank order coefficient ρ compared with other Aesthetics Attributes Assessment method under AADB.

Attribute	Kong et al. [10]	Malu et al. [11]	Proposed-AlexNet	Proposed-ResNet50
Balancing Element	0.220	0.186	**0.320**	**0.339**
Color Harmony	0.471	0.475	**0.494**	**0.539**
Interesting Content	0.508	0.584	0.537	0.580
Shadow Depth of Field	0.479	0.495	**0.511**	**0.553**
Good Lighting	0.443	0.399	**0.459**	**0.488**
Motion Blur	-	-	0.434	0.416
Object Emphasis	0.602	0.666	0.648	**0.673**
Repetition	-	-	0.508	0.530
Rule of Thirds	0.225	0.178	**0.288**	**0.279**
Symmetry	-	-	0.496	0.475
Vivid Color	0.648	0.681	**0.686**	**0.715**
Overall Aesthetic Score	0.654²/0.678	0.689	0.656	**0.708**
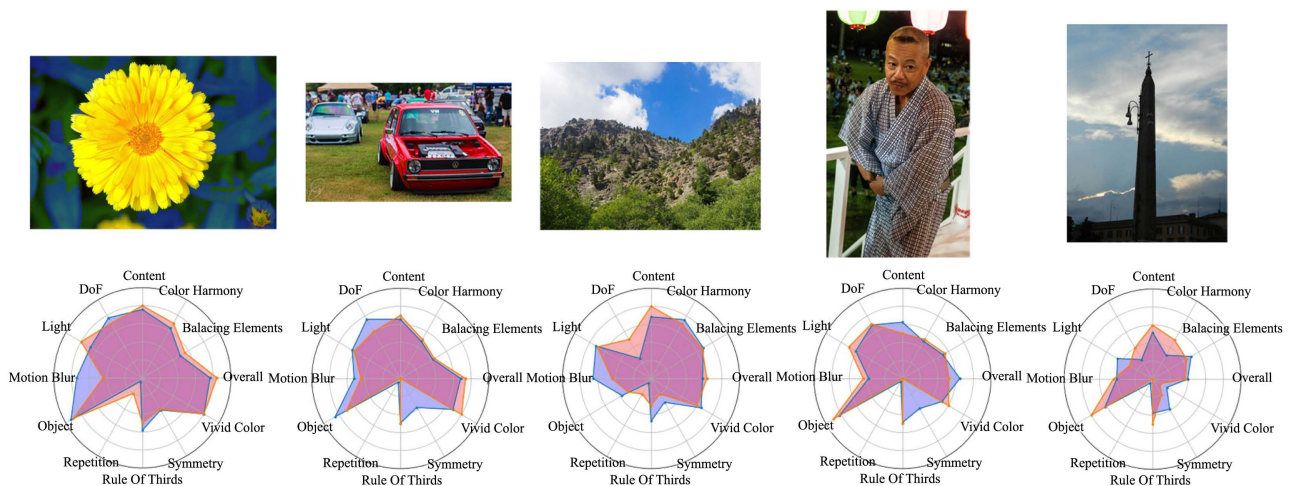
Figure 3. Test cases of our proposed method. Scores are normalized to 0 and 1 for illustration. Red represents the ground truth annotation in test set and blue represents prediction made by our model.

D. Ablation study

To verify the effectiveness of our method, we conduct experiments on two different backbones as AlexNet (Table 2) and ResNet-50 (Table 3). As our proposed data covariance learning scheme (COV) can be viewed as an extended form of mean square error (MSE), thus the ablation study is designed to use models trained with MSE to represent the absence of COV.

Table 2. Ablation study with AlexNet under AADB. COV/MSE means the model is trained with/without the proposed covariance learning branch.

Attribute	Alex-MSE	Alex-COV (Proposed)
Balancing Element	0.300	**0.320**
Color Harmony	0.469	**0.494**
Interesting Content	0.529	**0.537**
Shadow Depth of Field	0.504	**0.511**
Good Lighting	0.443	**0.459**
Motion Blur	0.407	**0.434**
Object Emphasis	0.633	**0.648**
Repetition	0.463	**0.508**
Rule of Thirds	0.231	**0.288**
Symmetry	0.452	**0.496**
Vivid Color	0.661	**0.686***
Overall Aesthetic Score	0.6566	**0.6569**
Average	0.479	**0.503**
Median	0.466	**0.502**

Table 3. Ablation study with Resnet50 under AADB. COV/MSE means the model is trained with/without the proposed covariance learning branch.

Attribute	Resnet50-MSE	Resnet50-COV (Proposed)
Balancing Element	0.338	**0.339**
Color Harmony	0.513	**0.539**
Interesting Content	**0.583**	0.580
Shadow Depth of Field	0.549	**0.553**
Good Lighting	0.482	**0.488**
Motion Blur	**0.432**	0.416
Object Emphasis	0.648	**0.673**
Repetition	0.511	**0.530**
Rule of Thirds	0.244	**0.279**
Symmetry	**0.497**	0.475
Vivid Color	0.693	**0.715***
Overall Aesthetic Score	0.706	**0.708**
Average	0.516	**0.525**
Median	0.512	**0.535**

5. Discussion

We model all of the 11 attribute in the original dataset in a TOTALLY end-to-end manner, without: 1) Sampling strategies of training pairs in [10]; 2) Handcrafted attributes selection and architectural design of CNN [11].

Our method is encapsulated in a loss-function style, which means high extensibility in a broad range of CNN-based deep learning method.

Full information of the target distribution can be extracted which can produce conditional distribution of any combinations of factors, which is useful in application, and further increase interpretability.

Interestingly, our method does not need to perform regularization on Σ as other data uncertainty methods, which may be because of the two-step parameterization that serves as a strong regularizer.

Our method can be further improved by: 1) using more complicated parameterization as Gaussian Process; 2) initialized with pretrained weight that better illustrates photographic styles, such as the pretrained weights derived from scene classification and texture recognition; 3) applying features that represent style; 4) train with conditional distribution that depends on one or more attributes.

6. Conclusion

We have developed a theoretically neat approach along with an efficient implementation to model the interaction of multiple photographic attributes which is extensible to a broad range of deep learning based methods. Our experiments show superior results especially on attributes that were previously thought difficult and provide better interpretability due to the meaningfulness of attributes while nearly no extra computing cost is introduced since it contains no architectural requirements. To sum up, our method is a good complement to image aesthetic assessment framework.

Source Code

Our experiments for Aesthetic Attributes Assessment are implemented with PyTorch [34] and fast.ai [35], which is available at
https://github.com/Hong123123/AAACov.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Deng, Y., Loy, C.C. and Tang, X. (2018) Aesthetic-Driven Image Enhancement by Adversarial Learning. *Proceedings of the* 26*th ACM International Conference on Multimedia*, Seoul, 22-26 October 2018, 870-878.
https://doi.org/10.1145/3240508.3240531

[2] Joshi, D., Datta, R., Fedorovskaya, E., Luong, Q.-T., Wang, J.Z., Li, J. and Luo, J.

(2011) Aesthetics and Emotions in Images. *IEEE Signal Processing Magazine*, **28**, 94-115. https://doi.org/10.1109/MSP.2011.941851

[3]   Liu, L., Chen, R., Wolf, L. and Cohen-Or, D. (2010) Optimizing Photo Composition. *Computer Graphics Forum*, **29**, 469-478. https://doi.org/10.1111/j.1467-8659.2009.01616.x

[4]   Luo, Y. and Tang, X. (2008) Photo and Video Quality Evaluation: Focusing on the Subject. In: *European Conference on Computer Vision*, Springer, Berlin, 386-399. https://doi.org/10.1007/978-3-540-88690-7_29

[5]   Yan, Z., Zhang, H., Wang, B., Paris, S. and Yu, Y. (2016) Automatic Photo Adjustment Using Deep Neural Networks. *ACM Transactions on Graphics* (*TOG*), **35**, 1-15. https://doi.org/10.1145/2790296

[6]   Talebi, H. and Milanfar, P. (2018) Nima: Neural Image Assessment. *IEEE Transactions on Image Processing*, **27**, 3998-4011. https://doi.org/10.1109/TIP.2018.2831899

[7]   Tu, Y., Niu, L., Zhao, W., Cheng, D. and Zhang, L. (2020) Image Cropping with Composition and Saliency Aware Aesthetic Score Map. *AAAI*, New York, 7-12 February 2020, 12104-12111. https://doi.org/10.1609/aaai.v34i07.6889

[8]   Deng, Y., Loy, C.C. and Tang, X. (2017) Image Aesthetic Assessment: An Experimental Survey. *IEEE Signal Processing Magazine*, **34**, 80-106. https://doi.org/10.1109/MSP.2017.2696576

[9]   Ang, T. (2012) Digital Photographer's Handbook. Penguin, London.

[10] Kong, S., Shen, X., Lin, Z., Mech, R. and Fowlkes, C. (2016) Photo Aesthetics Ranking Network with Attributes and Content Adaptation. In: *European Conference on Computer Vision*, Springer, Berlin, 662-679. https://doi.org/10.1007/978-3-319-46448-0_40

[11] Malu, G., Bapi, R.S. and Indurkhya, B. (2017) Learning Photography Aesthetics with Deep CNNS.

[12] Yang, H., Shi, P., He, S., Pan, D., Ying, Z. and Lei, L. (2019) A Comprehensive Survey on Image Aesthetic Quality Assessment. 2019 *IEEE*/*ACIS* 18*th International Conference on Computer and Information Science* (*ICIS*), Beijing, 17-19 June 2019, 294-299. https://doi.org/10.1109/ICIS46139.2019.8940355

[13] Zeng, H., Cao, Z., Zhang, L. and Bovik, A.C. (2019) A Unified Probabilistic Formulation of Image Aesthetic Assessment. *IEEE Transactions on Image Processing*, **29**, 1548-1561. https://doi.org/10.1109/TIP.2019.2941778

[14] Jin, X., Wu, L., Li, X., Chen, S., Peng, S., Chi, J., Ge, S., Song, C. and Zhao, G. (2017) Predicting Aesthetic Score Distribution through Cumulative Jensen-Shannon Divergence.

[15] Wang, Z., Liu, D., Chang, S., Dolcos, F., Beck, D. and Huang, T. (2017) Image Aesthetics Assessment Using Deep Chatterjee's Machine. 2017 *IEEE International Joint Conference on Neural Networks* (*IJCNN*), Anchorage, 14-19 May 2017, 941-948.

[16] Wu, O., Hu, W. and Gao, J. (2011) Learning to Predict the Perceived Visual Quality of Photos. 2011 *IEEE International Conference on Computer Vision*, Barcelona, 6-13 November 2011, 225-232. https://doi.org/10.1109/ICCV.2011.6126246

[17] Chang, J., Lan, Z., Cheng, C. and Wei, Y. (2020) Data Uncertainty Learning in Face Recognition. *IEEE*/*CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*), Seattle, 13-19 June 2020, 5709-5718. https://doi.org/10.1109/CVPR42600.2020.00575

[18] Hu, W., Huang, Y., Zhang, F. and Li, R. (2019) Noise-Tolerant Paradigm for Train-

ing Face Recognition CNNS. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Long Beach, 15-20 June 2019, 11887-11896.

[19] Wu, X., He, R., Sun, Z. and Tan, T. (2018) A Light CNN for Deep Face Representation with Noisy Labels. *IEEE Transactions on Information Forensics and Security*, **13**, 2884-2896. https://doi.org/10.1109/TIFS.2018.2833032

[20] Ng, H.-W. and Winkler, S. (2014) A Data-Driven Approach to Cleaning Large Face Datasets. 2014 *IEEE International Conference on Image Processing* (*ICIP*), Paris, 27-30 October 2014, 343-347.

[21] Yu, T., Li, D., Yang, Y., Hospedales, T.M. and Xiang, T. (2019) Robust Person Re-Identification by Modelling Feature Uncertainty. *Proceedings of the IEEE International Conference on Computer Vision*, Seoul, 27 October-2 November 2019, 552-561. https://doi.org/10.1109/ICCV.2019.00064

[22] Datta, R., Joshi, D., Li, J. and Wang, J.Z. (2006) Studying Aesthetics in Photographic Images Using a Computational Approach. In: *European Conference on Computer Vision*, Springer, Berlin, 288-301. https://doi.org/10.1007/11744078_23

[23] Ke, Y., Tang, X. and Jing, F. (2006) The Design of High-Level Features for Photo Quality Assessment. 2006 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1, 419-426.

[24] Marchesotti, L., Perronnin, F., Larlus, D. and Csurka, G. (2011) Assessing the Aesthetic Quality of Photographs Using Generic Image Descriptors. 2011 *IEEE International Conference on Computer Vision*, Barcelona, 6-13 November 2011, 1784-1791. https://doi.org/10.1109/ICCV.2011.6126444

[25] Jin, B., Segovia, M.V.O. and Süsstrunk, S. (2016) Image Aesthetic Predictors Based on Weighted CNNS. 2016 *IEEE International Conference on Image Processing* (*ICIP*), Phoenix, 25-28 September 2016, 2291-229. https://doi.org/10.1109/ICIP.2016.7532767

[26] Kao, Y., He, R. and Huang, K. (2017) Deep Aesthetic Quality Assessment with Semantic Information. *IEEE Transactions on Image Processing*, **26**, 1482-1495. https://doi.org/10.1109/TIP.2017.2651399

[27] Lu, X., Lin, Z., Shen, X., Mech, R. and Wang, J.Z. (2015) Deep Multi-Patch Aggregation Network for Image Style, Aesthetics, and Quality Estimation. *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, 7-13 December 2015, 990-998.

[28] Ma, S., Liu, J. and Chen, C.W. (2017) A-Lamp: Adaptive Layout-Aware Multi-Patch Deep Convolutional Neural Network for Photo Aesthetic Assessment. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 4535-4544. https://doi.org/10.1109/CVPR.2017.84

[29] Mai, L., Jin, H. and Liu, F. (2016) Composition-Preserving Deep Photo Aesthetics Assessment. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 497-506. https://doi.org/10.1109/CVPR.2016.60

[30] Murray, N. and Gordo, A. (2017) A Deep Architecture for Unified Aesthetic Prediction.

[31] Boudiaf, M., Rony, J., Ziko, I.M., Granger, E., Pedersoli, M., Piantanida, P. and Ayed, I.B. (2020) Metric Learning: Cross-Entropy vs. Pairwise Losses.

[32] Loshchilov, I. and Hutter, F. (2019) Decoupled Weight Decay Regularization. 7*th International Conference on Learning Representations*, New Orleans, 6-9 May 2019. https://dblp.org/rec/conf/iclr/LoshchilovH19.html

[33] Smith, L.N. (2017) Cyclical Learning Rates for Training Neural Networks. 2017 *IEEE Winter Conference on Applications of Computer Vision* (*WACV*), Santa Rosa, 24-31 March 2017, 464-472. https://doi.org/10.1109/WACV.2017.58

[34] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J. and Chintala, S. (2019) Pytorch: An Imperative Style, High-Performance Deep Learning Library. In: Wallach, H., Larochelle, H., Beygelzimer, A., Alché-Buc, F., Fox, E. and Garnett, R., Eds., *Advances in Neural Information Processing Systems*, Vol. 32, Curran Associates, Inc., Red Hook, 8024-8035.

[35] Howard, J., *et al.* (2018) Fastai. https://github.com/fastai/fastai