CNN Feature Similarity: Paintings Are More Self-Similar at All Levels

Seyed Ali Amirshahi
Norwegian Colour and Visual Computing Laboratory,
Norwegian University of Science and Technology (NTNU),
Gjøvik, Norway
L2TI, University Paris 13
France
Email: s.ali.amirshahi@ntnu.no

Stella X. Yu
UC Berkeley / ICSI
Berkeley, CA., USA
Email: stellayu@berkeley.edu

***Abstract*—Self-similarity is often an indicator of highly aesthetic paintings; a key aspect is what feature to use for evaluating self-similarity. Previous works use low-level features such as the gradient orientation to show that artworks and natural scenes share a similar degree of self-similarity. In this study, we take advantage of the AlexNet model to evaluate the changes of self-similarity at different convolutional layers in a CNN model. Compared to previous measures, our approach takes into account low, mid, and high level features. Different behaviors with regards to self-similarity at different layers is observed in paintings. The results confirm previous findings that artworks and natural scenes share similar degrees of self-similarity. For paintings and photographs with similar subject matters, while different degrees of self-similarity are observed at the first layer, other layers show closer values. Finally, the proposed measure of self-similarity is able to better differentiate between images which belong to a similar category but different datasets of images.**

I. Introduction

Self-similarity is one of the universal characteristics found in highly aesthetic paintings. That is, if we zoom in and out of a highly self-similar image we see similar patterns. In other words, a highly self-similar image has a fractal like property [1]. In this work, using a Convolutional Neural Network (CNN), we performed an in-depth analysis on how different low, mid, and high level features affect the degree of self-similarity seen in paintings. We wanted to find out which type of images closely resemble the degree of self-similarity seen in paintings at different convolutional layers. Previously, using a limited number of features related to the orientation seen in the image, artworks and natural scenes shared a similar degree of self-similarity. Our questions is, does such property hold when calculating self-similarity using CNN features? Simply, we are using a statistical approach to compare CNN features at different levels of the spatial resolution and extend that to see how self-similar an image is. Using the new approach, we showed that paintings and photographs with similar subject matters have close self-similarity values. Nevertheless, self-similarity values were different when judging self-similarity only based on features related to orientation.

The existence of universal characteristics in artworks has been a hot topic of research in the field of experimental aesthetics [1], [2], [3], [4]. The mentioned properties are seen independent of the subject matter it covers, the style or art period it is associate with, or the cultural background the creator has. Studies through different subjective and psychological experiments by philosophers, psychologists, artists, and vision experts has been the main approach taken in this field of studies. Works have related the mentioned characteristics to the human visual system and pointed out that they evoke aesthetic perception in all observers. Following this line of work, in the field of computational aesthetics, it has been shown that such universal characteristics are seen not only in paintings but other types of artworks such as handwritings, ornate prints, ornaments, and even print advertisements from different regions and cultural backgrounds [5], [6]. One such universal characteristic seen in paintings is self-similarity [5], [6], [7], [8], [9], [10], [11], [12], [13], [14]. In this study, we used the AlexNet [15] model which has been pre-trained on the ImageNet [16] dataset to study the degree of self-similarity in different convolutional layers in the network. We then took one step further and linked the calculated values to reach a single value which can represent the overall self-similarity in the image.

It has been pointed out in different studies that the high degree of self-similarity seen in artworks is close to those of natural scenes. In general, a number of research in different fields of study have pointed out the fact that artworks and complex natural scenes share common statistical properties [7], [17], [18], [19], [20]. Different studies suggest that similar coding mechanisms in the human visual system (efficient sensory code) might be used for both types of images [12], [13], [21]. Some works have went on to hypothesis that the shared property is intentionally mimicked by artists to resemble the properties seen in natural scenes [10], [12]. As an example, Amirshahi et al. [9] showed that although a dataset of portrait paintings is highly self-similar, the degree of self-similarity seen in a set of portrait photographs is quite low.

The rest of the paper is organized as follows: we will first give a short overview on previous works focused on evaluating self-similarity in paintings and other artworks in Section II. Section III is dedicated to introducing the proposed approach.

2018 Colour and Visual Computing Symposium (CVCS)
978-1-5386-5645-7/18/$31.0

A short description of the datasets used in our experiments is given in Section IV while the experimental results is presented in Section V. Finally, in Section VI, a conclusion of this work is discussed.

II. Previous works

Studies such as [10], [13], [14] were among the first works to take a computational approach for evaluting self-similarity in artworks. In the mentioend studies, the Fourier analysis was used to study statistical properties of Eastern and Western graphic artworks. The results showed that like complex natural scenes, graphic artworks have a scale invariant Fourier spectrum. This means that when we zoom in and out of an image the spatial frequency profile remains constant. From the results, they implied that images of both graphic artworks and complex natural scenes have a spatial frequency profile which is self-similar (fractal-like). Redies et al. [13] showed that unlike artworks and natural scenes, other types of images such as photographs of faces, plants, and simple objects do not follow the same spatial properties.

It can be implied from the mentioned works and other studies such as [22], [23] that graphic artworks show a high degree of self-similarity at different levels of the spatial resolution. Based on this finding and inspired by how PHOG (Pyramid of Histogram of Orientation Gradients) features [24] are calculated for an image, works such as [7], [9], [11] introduced a new measure of self-similarity. To calculate the measure of self-similarity, they first compute the HOG (Histogram of Orientation Gradients) feature vectors [25],

$$h(\mathcal{I}) = (h_1(\mathcal{I}), h_2(\mathcal{I}), \ldots, h_i(\mathcal{I}), \ldots, h_n(\mathcal{I})), \quad (1)$$

for a given image $\mathcal{I}$. In Eq 1, the HOG feature vector is presented by $h(\mathcal{I})$, the i^{th} bin in $h(\mathcal{I})$ is shown by $h_i(\mathcal{I})$, and n corresponds to the number of bins in the HOG feature vector. Similar to the PHOG approach, they then divide the image into four equal sub-regions and calculate the HOG feature vectors for each sub-region. In the proposed approach, the calculation is continued for level four of the spatial pyramid. The proposed approach is based on comparing different HOG vectors calculated for all the sub-regions in a specific level to their parent region. The next step in calculating the measure of self-similarity is to introduce

$$d_{HIK}(h(\mathcal{I}_1), h(\mathcal{I}_2)) = \sum_{i=1}^{n} \min(h_i(\mathcal{I}_1), h_i(\mathcal{I}_2)). \quad (2)$$

The Histogram Intersection Kernel (HIK) [26] is calculated between the HOG feature vectors in two different regions in an image $\mathcal{I}$ ($\mathcal{I}_1$ and $\mathcal{I}_2$). This allows to determine how similar two HOG feature vectors are. Finally,

$$m_{SS}(\mathcal{I}, L) = \text{median}\{d_{HIK}(h(\mathbf{S}), h(N(\mathbf{S})))\}, \quad (3)$$

is used to calculate self-similarity for image $\mathcal{I}$ at level L. The median value is calculated among all the HIK values calculated between all sub-regions of the image at level L of the spatial pyramid. In Eq. 3, $h(N(\mathbf{S}))$ corresponds to the parent region in the image to which sub-region $\mathbf{S}$ is compared to. The median value in Eq. 3 is used to avoid taking the overshoots into account.

Amirshahi et al. [7] pointed out to a number of drawbacks in the proposed measure of self-similarity. For instance, self-similarity is measured based on a single spatial level omitting the changes seen in other levels. For this reason, the measure of weighted self-similarity was proposed. In summary, using a simple weighting approach, the measure of weighted self-similarity

$$m_{WSS}(\mathcal{I}) = \frac{1 - \sigma(m_{SS}(\mathcal{I}))}{\sum_{l=1}^{L} \frac{1}{l}} \sum_{l=1}^{L} \frac{1}{l} \cdot m_{SS}(\mathcal{I}, l), \quad (4)$$

links self-similarity values calculated at different levels of the spatial pyramid. In Eq. 4,

$$\begin{aligned} m_{SS}(\mathcal{I}) = &(m_{SS}(\mathcal{I}, 1), m_{SS}(\mathcal{I}, 2), \cdots, \\ &m_{SS}(\mathcal{I}, z), \cdots, m_{SS}(\mathcal{I}, L)), \end{aligned} \quad (5)$$

corresponds to a self-similarity vector consisting of all self-similarity values calculated at different levels for a given image $\mathcal{I}$. In Eq. 4, σ corresponds to the standard deviation. It should be pointed out that [7] suggested that level L in $m_{SS}(\mathcal{I}, L)$ corresponds to the last level in the spatial pyramid which the smallest side of the smallest sub-region is greater than 64 pixels.

While the introduced approaches have led to acceptable results [7], [9], [11], self-similarity is only evaluated based on using local features in the image (different orientation angles). In this study, by using a Convolutional Neural Network (CNN) model, we not only focused on using local features but also took a step further and used mid and high level features in our calculation. In the next section the proposed approach is described in detail.

III. Proposed approach

In the last few years, a huge amount of attention has been paid to the use of different CNN models in various computer vision scenarios. The methods which take advantage of the use of CNNs have shown a better performance with regards to the accuracy they provide. This issue along with the high number of feature maps extracted from the image encouraged us to study the self-similarity seen in the image based on the extracted feature maps. This approach will allow us to not only provide a single value which represents the self-similarity in an image but we are also able to calculate self-similarity at different layers of the network based on different feature maps. Compared to previous proposed approaches in this field of study, taking into account low, mid, and high level features in our approach is another advantage of the introduced method.

The proposed approach was based on the measure of weighted self-similarity previously introduced by Amirshahi et al. [7]. In our approach, using the AlexNet [15] model which was pre-trained on the ImageNet [16] dataset we calculated the measure of self-similarity for each convolutional layer. The five self-similarity values were then linked with each other to

provide a single score for each image. We should point out that Amirshahi et al. [27] took a similar approach to propose a new image quality metric. The following steps were taken for calculating the self-similarity at each layer.

1) For each of the five convolutional layers n, we calculate histogram

$$h(n) = (\sum_{i=1}^{X}\sum_{j=Y}^{x}\mathcal{F}(n,1)(i,j), \sum_{i=1}^{X}\sum_{j=Y}^{x}\mathcal{F}(n,2)(i,j), \cdots, \sum_{i=1}^{X}\sum_{j=Y}^{x}\mathcal{F}(n,z)(i,j), \cdots, \sum_{i=1}^{X}\sum_{j=Y}^{x}\mathcal{F}(n,M)(i,j)). \quad (6)$$

In Eq. 6, $\mathcal{F}(n,z)$ corresponds to feature map z in the n^{th} convolutional layer and X and Y correspond to the height and width of the feature map. It is clear that each bin in the histogram *h* corresponds to the sum of the response of the feature maps at a given layer. It is interesting to point out that the sum over each feature map can be compared to the bin entry in the HOG vectors. Keep in mind that unlike the case of the measure of weighted self-similarity where only 16 bins were used, in the case of our method, we are dealing with a huge number of bin entries (each representing one of the feature maps). For example just in the first layer, 96 feature maps are extracted from the image.

2) Similar to the approach taken in [7], we then divided each feature map to four equal sub-regions and calculated histogram *h* for each sub-region.
3) The division and calculation was then continued as long as the smallest side of the smallest sub-region was equal or larger than 7 pixels. In other words, we did the calculations until the third level for the first convolutional layer, the second level for the second layer, and the first level for the third, fourth, and fifth layers.
4) Having the different histogram vectors at different levels and different layers, for each layer n, the measure of self-similarity was calculated by

$$m_{CLWSS}(\mathcal{I},n) = \frac{1-\sigma(m_{CSS}(\mathcal{I},n))}{\sum_{l=1}^{L}\frac{1}{l}}\sum_{l=1}^{L}\frac{1}{l}\cdot m_{CSS}(\mathcal{I},n,l). \quad (7)$$

In Eq. 7, $m_{CLSS}(\mathcal{I},n,l)$ corresponds to the self-similarity value calculated for image $\mathcal{I}$ at level l in the n^{th} layer. $m_{CSS}(\mathcal{I},n)$ is then the result of the concatenation of all m_{CSS} values for $\mathcal{I}$ at the n^{th} layer.
5) Since the self-similarity at each layer was calculated at different levels, to link all the self-similarity values, instead of using an arithmetic mean we used the geometric mean

$$m_{CWSS}(\mathcal{I}) = \prod_{n=1}^{5} m_{CLWSS}(\mathcal{I},n). \quad (8)$$

In Eq. 8, $m_{CWSS}(\mathcal{I})$ corresponds to the overall self-similarity of image $\mathcal{I}$.

IV. Image Databases

To be able to compare the self-similarity measure introduced in this work to previous results, we use the image datasets introduced in the work of Amirshahi et al. [7]. For an in-depth information about the datasets please refer to [7]. The image datasets used in our experiments can be categorized to four category of images.

A. Artworks

For the case of aesthetic paintings, we use the JenAesthetics Dataset [7], [28], [29] (Fig. 1(a)). The dataset consists of 1621 images from 11 different art periods and provides information about the three dominant subject matters observed in the paintings.

B. Natural Scenes

As mentioned earlier, different studies have pointed out that artworks and natural scenes share common statistical properties such as similar degrees of self-similarity. This category of images consists of three different datasets which are plant patterns (Fig. 1(b)), vegetation (Fig. 1(c)), and large vistas (Fig. 1(d)). It should be pointed out that the datasets have been captured in a way that they seem to be photographs captured from the same scene at different distances. While the plant pattern dataset are captured from a short distance the vegetation dataset are taken from a further distance and images in the large vista dataset are photographs taken from a far distance.

C. Highly Self-similar Natural Patterns

This category of images consists of four different datasets, lichen (Fig. 1(e)), clouds (Fig. 1(f)), turbulent water (Fig. 1(g)), and branches (Fig. 1(h)). The images can be used to evaluate the validity of the proposed measure.

D. Other Datasets

This category of images is used to compare artworks with datasets of photographs taken from other man-made structures and objects as well as portrait photographs. For example, in our experimental results, we compare paintings and photographs of urban scenes (Fig. 1(i)), buildings (Fig. 1(j)), facades (Fig. 1(k)), and simple objects (Fig. 1(l)). To compare portrait photographs and paintings, 500 images were randomly selected from the labeled faces in the wild dataset [30] (Fig. 1(m)). Similar to the case of the datasets in the natural scenes category it can be assumed that the images in the facades, buildings, and urban scenes datasets are taken from the same scene from different distances.

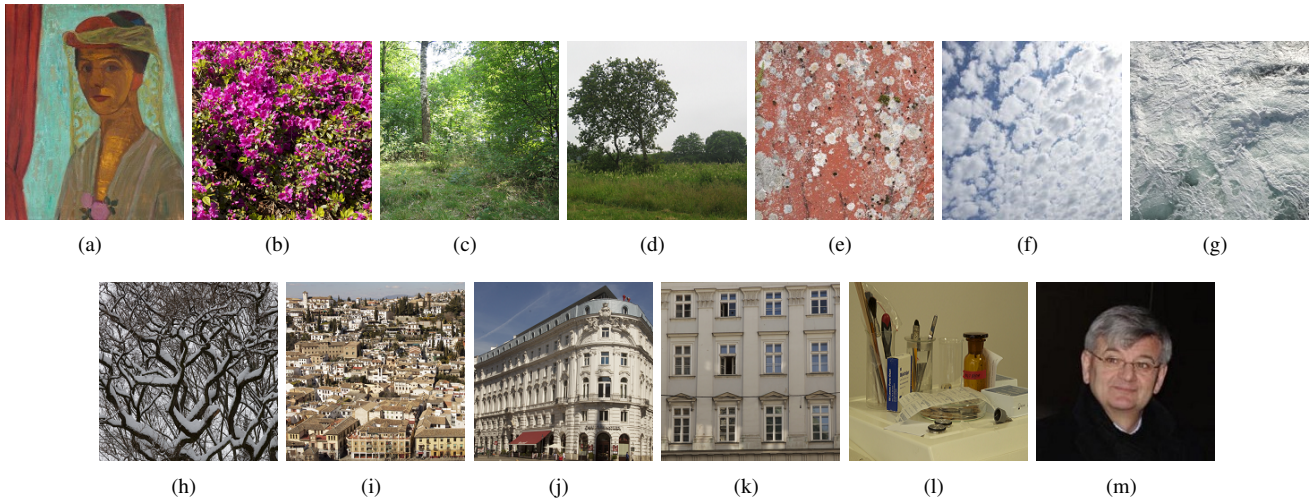


Fig. 1. Sample images from the (a) JenAesthetics with 1621, (b) plant patters with 331, (c) vegetation with 525, (d) large vistas with 584, (e) lichen with 280, (f) clouds with 248, (g) turbulent water with 425, (h) branches with 302, (i) Urban scenes with 225, (j) buildings with 528, (k) facades with 175, (l) simple objects with 207, and (m) portrait photographs with 500 images in our datasets.

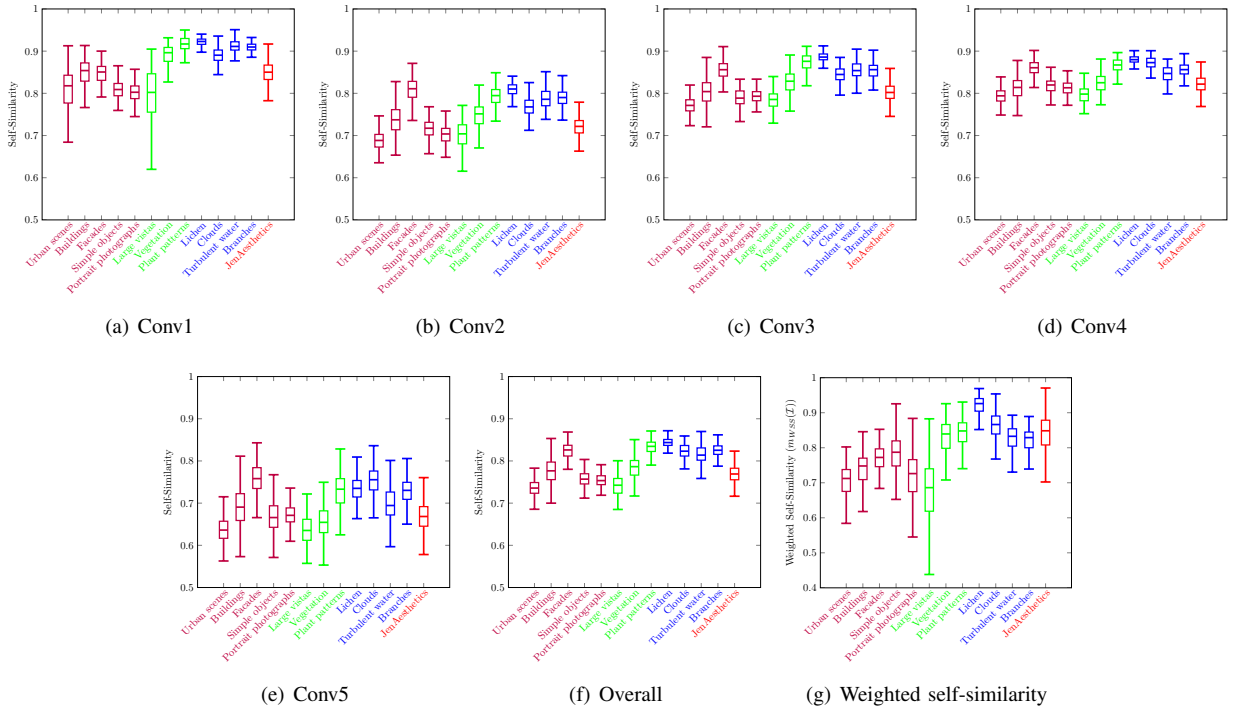


Fig. 2. Self-similarity values calculated for the images in different layers ((a)-(e)) as well as the overall self-similarity results for the images (f). Below each plot the layer the calculations is done for is presented. (g) The weighted self-similarity values [7] for different datasets.

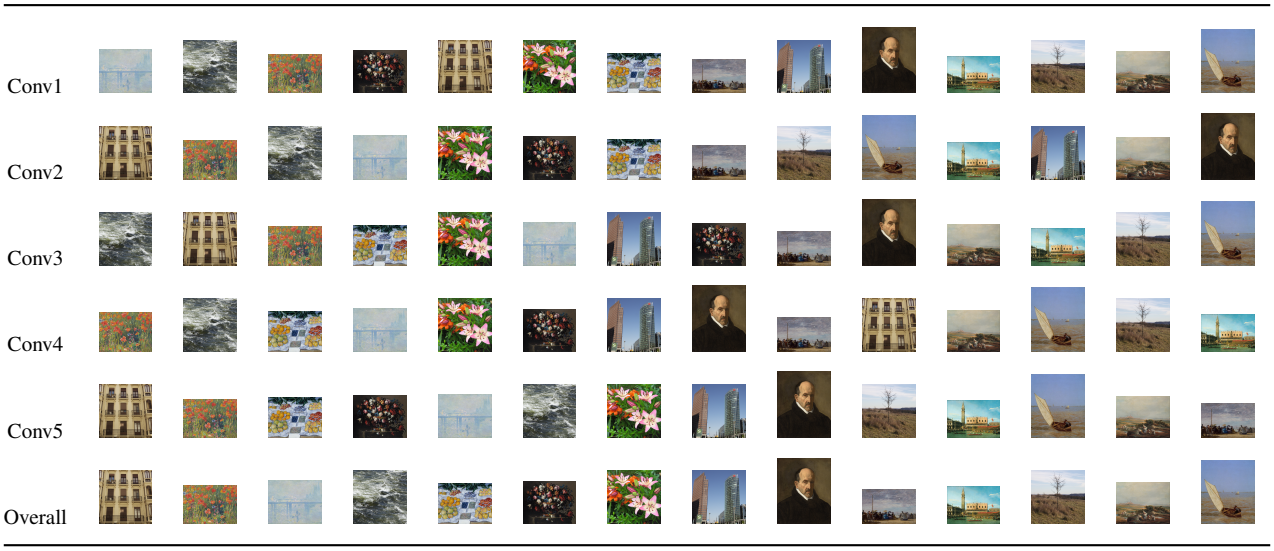
V. Experimental Results

Earlier, we introduced a new approach for calculating the self-similarity seen in an image which takes advantage of the use of CNNs. Based on this measure, we first calculated the self-similarity values in different convolutional layers for the images in our dataset. While Figs. 2(a)-(e) provides self-similarity values for different convolutional layers, Fig. 2(f) provides the results of the overall self-similarity in different datasets. Table I provides sample images ordered based on their self-similarity values at different convolutional layers along with the overall scores.

Self-similarity in the first convolutional layer: It is known that, the features in the first convolutional layer are mainly related to orientations seen in the image. This corresponds with our results (Fig. 2(a)) where datasets with a dominant orientation present in the image such as the urban scenes, large vistas, simple objects, and portrait photograph datasets showed lower self-similarity values. In the case of datasets where a dominant orientation was not observed, such as the datasets in

TABLE I

SAMPLE IMAGES ORDERED BASED ON SELF-SIMILARITY VALUES AT DIFFERENT LAYERS AND OVERALL SCORES. SELF-SIMILARITY VALUE DECREASE FROM LEFT TO RIGHT.



the self-similar category and the vegetation and plant patterns datasets, higher self-similarity values were observed. It was interesting to observe that paintings which lack a dominant orientation show higher self-similarity in the first convolutional layer compared to cases which do have dominant orientations. A similar observation was also seen in the case of photographs where a lack of dominant orientation results in high self-similarity values.

In the case of paintings in the first layer, a small increase is seen in paintings depicting interior scenes compared to the paintings with a subject matter of urban scenes and buildings. This finding could be related to the existence of more dominant orientations in the two later set of paintings. A drop of values can also be seen in paintings depicting seascape, port, and coast as well as sky which are similar to the images in the large vista dataset. Finally, the increase in the case of the paintings with a subject matter of flowers and vegetation show an increase in their self-similarity value similar to the case of the plant patterns and vegetation photographs.

Self-similarity in the middle layers: Results in the second, third, and fourth convolutional layers follow the expectations from the feature maps extracted at the mentioned layers. Generally, the feature maps in the mentioned layers are designed to detect textures in the image. This is in line with the results (Fig. 2(b)-(d)) where datasets containing images with less texture (urban scene and large vista) show lower self-similarity values compared to images containing higher texture (facades and plant pattern correspondingly). The drop in the self-similarity values for the case of the simple objects and portrait photograph datasets is also interesting. This could be due to the nature of the images in the two mentioned dataset (Fig. 1(l) and (m)) which mainly cover an object(s)/person with a uniform background. It is interesting to observe high self-similarity values in the three mentioned convolutional layers for the case of paintings and photographs with high texture. Lower values is observed in the case of paintings and photographs with less texture.

With regards to self-similarity seen in the mid layers of paintings, similar patterns like the case of photographs can be seen in the case of paintings. The higher the texture seen in the painting the higher the self-similarity values in the mid layers.

Self-similarity in the fifth layer: In the case of the fifth layer, the features are focused on detecting objects. This again justifies the difference between the calculated self-similarity for different datasets. While in the case of datasets which photographs have been taken from a closer distance (facades and plant patterns) a number of objects are detected in the photograph, in the case of photographs taken from a longer distance single or fewer objects can be detected. This justifies the higher self-similarity values taken from a closer distance.

Self-similarity values seen in paintings at the fifth layer show similar pattern to the observation we have in photographs. A change in self-similarity value can be observed depending on the number of objects depicted in the painting. A good example is the slight difference between the portrait paintings of a single person and the many persons where the latter show a higher self-similarity value.

Difference between the proposed measure and the measure of weighted self-similarity: Comparing the overall results (Fig. 2(f)) to the weighted self-similarity measure proposed by Amirshahi et al. [7] (Fig. 2(g)) shows interesting conclusions. On one hand, the weighted self-similarity measure shows a better accuracy when comparing all 13 datasets with each other. This is, higher self-similarity values for natural scenes datasets compared to portrait photographs, man-made objects, and structures (which at a first glance is what is expected). On the other hand, the proposed self-similarity

measure show more accurate results inside different image categories.

Subject matter and art periods in paintings: Finally, we focused our attention on observing the values calculated in the case of the subject matters and art periods covered in the JenAesthetics dataset. It is interesting to observe close self-similarity values in both cases compared to values calculated in the different photographic datasets in our experiments (Fig. 2). Again, this follows previous assumptions that artists intentionally change statistical properties in their creation.

Difference between paintings and photographs with same subject matters: Compared to photographs, paintings on average show higher self-similarity values when the same subject matter is compared. The existence of dominant orientations in photographs results in lower self-similarity values compared to paintings in the first convolutional layer. The lack of texture in photographs could be the reason behind higher self-similarity values in paintings. It is interesting to observe close values in the case of portraits which can be due to the similar nature of portrait photographs and paintings.

VI. Conclusion

In conclusion, a new measure of self-similarity based on using the AlexNet [15] model was introduced. We were not only able to measure the overall self-similarity of an image but also evaluated self-similarity at different convolutional layers. The proposed approach was inspired by a measure previously introduced by Amirshahi et al. [7]. Compared to the previous methods, our approach does not only use low level features in the image but also takes into account mid and high level features.

The proposed measure was tested on a large dataset of aesthetic paintings as well as a number of different datasets of photographs covering different subject matters. The results show that artworks and natural scenes share similar degrees of self-similarity. Compared to previous measures, the proposed approach is able to better differentiate between images in similar category of images.

References

[1] S. A. Amirshahi, *Aesthetic Quality Assessment of Paintings*. Verlag Dr. Hut, 2015.

[2] S. A. Amirshahi, S. H. Amirshahi, and J. Denzler, "Using color difference equations for calculating gradient images," in *13th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*. IEEE, 2017, pp. 275–282.

[3] S. A. Amirshahi and J. Denzler, "Judging aesthetic quality in paintings based on artistic inspired color features," in *International Conference on Digital Image Computing: Techniques and Applications (DICTA),*. IEEE, 2017, pp. 1–8.

[4] S. Zeki, "Art and the brain," *Journal of Consciousness Studies*, vol. 6, no. 6-7, pp. 76–96, 1999.

[5] T. Melmer, S. A. Amirshahi, M. Koch, J. Denzler, and C. Redies, "From regular text to artistic writing and artworks: Fourier statistics of images with low and high aesthetic appeal," *Frontiers in Human Neuroscience*, vol. 7, no. 106, 2013.

[6] J. Braun, S. A. Amirshahi, J. Denzler, and C. Redies, "Statistical image properties of print advertisements, visual artworks and images of architecture," *Frontiers in Psychology*, vol. 4, no. 808, 2013.

[7] S. A. Amirshahi, C. Redies, and J. Denzler, "How self-similar are artworks at different levels of spatial resolution?" in *Symposium on Computational Aesthetics*. ACM, 2013, pp. 93–100.

[8] C. Redies, A. Brachmann, and G. U. Hayn-Leichsenring, "Changes of statistical properties during the creation of graphic artworks," *Art & Perception*, vol. 3, no. 1, pp. 93–116, 2015.

[9] S. A. Amirshahi, M. Koch, J. Denzler, and C. Redies, "PHOG analysis of self-similarity in aesthetic images," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2012, pp. 82 911J–82 911J.

[10] D. J. Graham and C. Redies, "Statistical regularities in art: Relations with visual coding and perception," *Vision Research*, vol. 50, no. 16, pp. 1503–1509, 2010.

[11] C. Redies, S. A. Amirshahi, M. Koch, and J. Denzler, "PHOG-derived aesthetic measures applied to color photographs of artworks, natural scenes and objects," in *Computer Vision–ECCV 2012. Workshops and Demonstrations*. Springer, 2012, pp. 522–531.

[12] C. Redies, J. Hänisch, M. Blickhan, and J. Denzler, "Artists portray human faces with the fourier statistics of complex natural scenes," *Network: Computation in Neural Systems*, vol. 18, no. 3, pp. 235–248, 2007.

[13] C. Redies, J. Hasenstein, and J. Denzler, "Fractal-like image statistics in visual art: similarity to natural scenes," *Spatial Vision*, vol. 21, no. 1-2, pp. 137–148, 2007.

[14] D. J. Graham and D. J. Field, "Statistical regularities of art images and natural scenes: Spectra, sparseness and nonlinearities," *Spatial Vision*, vol. 21, no. 1-2, pp. 149–164, 2008.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[16] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009, pp. 248–255.

[17] G. J. Burton and I. R. Moorhead, "Color and spatial structure in natural scenes," *Applied Optics*, vol. 26, no. 1, pp. 157–170, 1987.

[18] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *JOSA A*, vol. 4, no. 12, pp. 2379–2394, 1987.

[19] W. S. Geisler, "Visual perception and the statistical properties of natural scenes," *Annu. Rev. Psychol.*, vol. 59, pp. 167–192, 2008.

[20] D. L. Ruderman and W. Bialek, "Statistics of natural images: Scaling in the woods," *Physical review letters*, vol. 73, no. 6, p. 814, 1994.

[21] C. Redies, "A universal model of esthetic perception based on the sensory coding of natural stimuli," *Spatial Vision*, vol. 21, no. 1, p. 97, 2008.

[22] R. Taylor, B. Spehar, C. Hagerhall, and P. Van Donkelaar, "Perceptual and physiological responses to Jackson Pollock's fractals," *Frontiers in Human Neuroscience*, vol. 5, p. 60, 2011.

[23] R. P. Taylor, "Order in Pollock's chaos," *Scientific American*, vol. 287, no. 6, pp. 116–121, 2002.

[24] A. Bosch, A. Zisserman, and X. Munoz, "Representing shape with a spatial pyramid kernel," in *ACM International Conference on Image and Video Retrieval*. ACM, 2007, pp. 401–408.

[25] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1. IEEE, 2005, pp. 886–893.

[26] A. Barla, E. Franceschi, F. Odone, and A. Verri, "Image kernels," in *Pattern Recognition with Support Vector Machines*. Springer Berlin Heidelberg, 2002, pp. 83–96.

[27] S. A. Amirshahi, M. Pedersen, and S. X. Yu, "Image quality assessment by comparing CNN features between images," *Journal of Imaging Science and Technology*, vol. 60, no. 6, pp. 60 410–1, 2016.

[28] S. A. Amirshahi, J. Denzler, and C. Redies, "JenAesthetics-a public dataset of paintings for aesthetic research," Computer Vision Group, University of Jena Germany, Tech. Rep., 2013.

[29] JenAesthetics Dataset. (2013) http://www.inf-cv.uni-jena.de/en/jenaesthetics. [Online]. Available: http://www.inf-cv.uni-jena.de/en/jenaesthetics

[30] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," Technical Report 07-49, University of Massachusetts, Amherst, Tech. Rep., 2007.