Pattern Recognition Letters 129 (2020) 56–62

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Learning the Principles of Art History with convolutional neural networks ☆

Eva Cetinic [a,*], Tomislav Lipic [a], Sonja Grgic [b]

[a] Rudjer Boskovic Institute, Bijenicka cesta 54, Zagreb 10000, Croatia
[b] University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, Zagreb 10000, Croatia

ARTICLE INFO

Article history:
Received 17 July 2019
Revised 29 October 2019
Accepted 7 November 2019
Available online 8 November 2019

Keywords:
Convolutional neural networks
Fine art
High-level image features
Wölfflin

ABSTRACT

Understanding the historical transformation of artistic styles implies the recognition of different stylistic properties. From a computer vision perspective, stylistic properties represent complex image features. In our work we explore the use of convolutional neural networks for learning features that are relevant for understanding properties of artistic styles. We focus on stylistic properties described by Heinrich Wölfflin in his book Principles of Art History (1915). Wölfflin identified five key visual principles, each defined by two contrasting concepts. We refer to each principle as one high-level image feature that measures how much each of the contrasting concepts is present in an image. We introduce convolutional neural network regression models trained to predict values of the five Wölfflin's features. We provide quantitative and qualitative evaluations of those predictions, as well as analyze how the predicted values relate to different styles and artists. The outcome of our analysis suggests that the models learn to discriminate meaningful features that correspond to the visual characteristics of concepts described by Wölfflin. This indicates that the presented approach can be used to enable new ways of exploring fine art collections based on image features relevant and well-known within art history.

1. Introduction

Analysing paintings is a complex task which includes understanding the subject matter, as well as properties of formal elements such as line, shape, colour, texture and composition. The majority of studies concerned with computational analysis of artworks focuses on automatic classification or recognizing objects in artworks, while not many attempts have been made to analyze artworks in terms of their specific stylistic properties. The successful performance of deep learning techniques for a wide variety of computer vision tasks, motivates us to explore their potential in enabling new ways of exploring digitized art collections. Particularly, regarding concepts which play an important role within art history such as the concepts defined by Heinrich Wölfflin in his book Principles of Art History (1915) [1]. Discussing the historical transformation of styles, particularly from Renaissance to Baroque, Wölfflin identifies five key visual principles. Each principle is defined by two contrasting visual schemes: (1) Linear and painterly, (2) closed and open form, (3) planar and recessional, (4) multiplicity and unity and (5) absolute and relative clarity. The transformation of styles from Renaissance to Baroque corresponds to the conversion of the arrangement of visual elements from the first to the second visual scheme in each principle. Although Wölfflin's comparative principles were developed based on the differences between characteristics of artworks of 16th and 17th centuries, they became a standard method of formal analysis of art and a conventional approach in understanding changes of artistic styles. The main goal of our work is to quantify and predict the level of presence of two contrasting visual schemes for each of the five principles in an image. Research outcomes presented in [2] suggest that convolutional neural networks trained to classify paintings according to different styles, implicitly learn features related to Wölfflin concepts. Based on those findings, we introduce convolutional neural networks models trained to predict values of each concept. We provide quantitative and qualitative evaluations of those predictions, as well as analyze how the predicted values relate to different styles and artists.

2. Related work

Most of the studies concerned with computational analysis of fine art images concentrate on the challenge of automatically classifying artworks based on categories such as artist, style or genre. Research progress in the domain of fine art automatic classifi-

☆ Handled by Associate Editor: Gabriella Sanniti di Baja.

* Corresponding author.
E-mail address: ecetinic@irb.hr (E. Cetinic).

https://doi.org/10.1016/j.patrec.2019.11.008
0167-8655/

Table 1
Feature IDs and descriptions for each of the five Wölfflin's principles.

Feature ID	Wölfflin's principle	Description
W1	linear vs. painterly	Linear - elements are clearly outlined and boundaries are clear Painterly - elements are fused together, contours and boundaries are blurred
W2	closed vs. open form	Closed form - elements are balanced with the frame, dominant horizontal and vertical composition Open form - impression of space beyond the edges of the picture, dominant diagonal composition
W3	planar vs. recessional	Planar - elements are arranged on successive planes parallel to the picture plane Recessional - elements are arranged on various planes, illusion of depth
W4	multiplicity vs. unity	Multiplicity - elements appear distinct and independent Unity - elements appear united and entangled, fused into a single whole
W5	absolute vs. relative clarity	Absolute clarity - explicit and articulated forms Relative clarity - less clearly articulated forms, intentionally avoiding objective clearness

cation was recently accelerated by the availability of large and well-annotated fine art datasets. The WikiArt collection[1] is the most commonly used dataset for fine art-related classifications tasks [3–9], followed by other available online sources such as the Web Gallery of Art (WGA) with more than 40k images [10], the Rijksmuseum challenge dataset [11,12], the OmniART dataset [13] and the newly introduced MultitaskPainting100k dataset [14]. The appearance of large annotated datasets enabled the use of convolutional neural networks for various large scale fine art classification tasks. Besides classification, the use of CNNs showed promising results in other areas of interest such as retrieving visual links in paintings collections [10] and recognizing objects in paintings [13,15]. CNN-based features were also used to address the difference between specific visual properties of artworks and natural images [16], as well as to explore quantitative approaches to highly subjective aspects of perceiving artworks [17]. To understand how internal representation of convolutional neural networks trained for style classification encode discriminative features, Elgammal et al. [2] performed a correlation analysis of learned features with art history related concepts. In particular, they introduced a dataset containing art historian's rating annotations (scale of 1 to 5) for each of the Wölfflin 's pairs. They showed that principal modes of variation obtained by applying Principal Component Analysis (PCA) on feature representations obtained from the last CNN layers correlate with the annotation values of certain Wölfflin concepts. Following the work of Elgammal et al. [2], we propose a methodology of quantifying and predicting Wölfflin features which could enable exploring digitized fine art collections based on features that are relevant in the context of art history.

3. Methodology

Wölfflin's principles are defined by two contrasting concepts. In this work, we refer to each principle as one high-level image feature that measures how much each of the contrasting concepts is present in an image. Table 1 lists feature IDs and brief conceptual descriptions of each of the five Wölfflin's principles.

Each feature can have a value in the range between 0 and 1. For example, the value of the first feature (W1) represents the relation between linear and painterly, with 0 representing a fully linear style and 1 representing a highly painterly style. Using the dataset containing art historian's rating annotations for each of the Wölfflin's pairs, presented by [2], we train a convolutional neural network regression model for each feature. The models are evaluated quantitatively by analyzing the predictions on the small annotated test set of paintings. Furthermore, the models are evaluated qualitatively by employing the models on a large unlabelled dataset of paintings and analyzing images with high and low prediction values of each feature, as well as by analyzing how the prediction

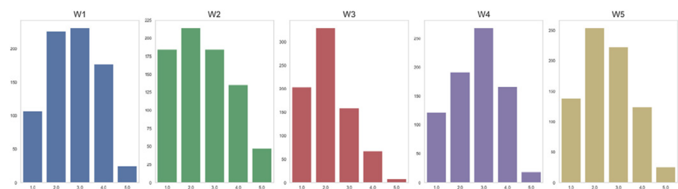

Fig. 1. Distribution of images per rating annotations in the Wölfflin's pairs' annotation dataset.

values of each feature change in regard to different art historical attributes. Detailed descriptions of the datasets and training settings are given in the following section.

4. Experimental setup

4.1. Datasets

For the purpose of training CNN models, we use the Wölfflin's pairs' annotation dataset presented by [2]. The dataset consists of 765 images. Each image in the dataset is annotated with values between 1 and 5 for each of the five Wölfflin's pairs. The annotation are provided by art historians. Fig. 1 shows the distribution of images in the dataset per value for each feature (W1-W5). For the purpose of training CNN models, we normalize the rating annotation values to be in the range between 0 and 1. For each task, the dataset is divided in order to keep 640 images for training, 65 for validation and 60 for testing. All images are resized to 256 x 256 pixels prior training.

After training, the models are employed on a new dataset - a large collection of paintings from the WikiArt dataset. WikiArt is a well organized collection of artworks which includes a broad set of metadata such as artist, style, genre, year of creation, technique, etc. At the time of our data collection process, the WikiArt collection included more than 130K images of digitized artworks (paintings, sculptures, illustrations, posters, etc.). For the purpose of our work, we used a subset of 105,121 color images of paintings.

4.2. CNN models

Because [2] showed that feature representations obtained from the last layers of CNN models trained for style classification correlate with the annotation values of certain Wölfflin concepts, in our work we decide to fine-tune CNN models pre-trained for style classification for the new task of predicting Wölfflin's pairs' feature values (W1-W5). Concretely, we used the best performing CNN model for style classification introduced in [9]. This model is trained to classify paintings based on 27 different style categories. The model architecture corresponds to the CaffeNet model [18], which is a slightly altered version of the AlexNet model [19]. For each of the five Wölfflin's concepts, we train one separate model

[1] http://www.wikiart.org.

Table 2
Summary of the model architecture (AlexNet). The type, name and size is given for each layer. The final column indicates if the layer was active during fine-tuning.

Type	Name	Size	Fine-tuning
Input	data	$227 \times 227 \times 3$	no
Convolutional	conv1	$55 \times 55 \times 96$	no
Max pooling	pool1	$27 \times 27 \times 96$	no
Convolutional	conv2	$27 \times 27 \times 256$	no
Max pooling	pool2	$13 \times 13 \times 256$	no
Convolutional	conv3	$13 \times 13 \times 384$	no
Convolutional	conv4	$13 \times 13 \times 384$	no
Convolutional	conv5	$13 \times 13 \times 256$	no
Max pooling	pool5	$6 \times 6 \times 256$	no
Fully connected	fc6	4096	yes
Fully connected	fc7	4096	yes
Fully connected	fc8	1	yes

Table 3
Mean suared error (MSE) and R^2 coefficient of determination of the best performing model for each feature value prediction task (W1-W5).

Task	MSE	R^2score
W1	0.0251	0.487
W2	0.0338	0.4179
W3	0.0185	0.475
W4	0.0258	0.399
W5	0.0369	0.3473

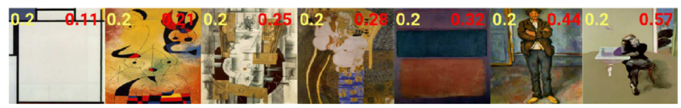

Fig. 3. Examples of artworks rated with the lowest W3 ground-truth score (yellow), ordered from left to right by their corresponding predicted W3 score (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

to predict the value of the corresponding feature (W1-W5). For all five models we use the style classification model as the base model when fine-tuning. Before fine-tuning, the last softmax classification layer is replaced with a regression layer minimizing the mean square error (MSE) loss. The output of the model is a score in the range from 0 to 1 that represents the predicted value of the particular Wölfflin's feature for which the model is being trained. Several different fine-tuning approaches were tested in order to identify the optimal setting for each task. We found that the best results for all five tasks are obtained when only the last three fully connected layers (fc6, fc7 and fc8) are being fine-tuned, while the weights of all other layers are kept frozen. The overview of the model architecture indicating which layers where kept frozen and which were modified while fine-tuning is summarized in Table 2. The models are fine-tuned using the Adam stochastic optimization method, with L2 regularization. For all tasks, the best results are achieved when training for 200 epochs with an initial learning rate of 10^{-4}.

5. Results and discussion

The experimental results are analysed and discussed from different perspectives. Firstly, we focus on the performance of regression models trained for each of the five different tasks on the test sets. Furthermore, we analyze the qualitative results by comparing images with the highest and lowest prediction values for each feature, obtained when employing the models on the large WikiArt dataset. Finally, we analyze how the predicted feature values change over time, as well as in relation to different stylistic categories.

5.1. Performance evaluation

To evaluate the predictive performance of different CNN regression models trained for predicting the values of Wölfflins features, we rely on the mean squared error and R^2 coefficient of determination. Table 3 lists the values of the mean squared errors and R^2 coefficients of determination of the best performing models on the tests set of each of the five prediction tasks. The results suggest a reasonable good performance of the models, particularly for the task of predicting the linear-painterly relation (W1 feature).

Besides analyzing the values of MSE and R2, we also evaluate the performance of the models based on regression plots. Fig. 2 shows the relationships between ground-truth and predicted values obtained with the best performing model for each of the five Wölfflin's features.

For some of the models, we can observe a rather high variance between values of the predicted feature scores in relation to the same ground-truth score. For example for the W3 model, predictions show high variability for images with the lowest ground-truth score, with predicted values ranging from 0.09 to 0.57. In order to understand this result, we visually inspect images manually annotated as having a highly planar style (having the lowest W3 ground-truth score). Fig. 3 shows several examples of artworks rated with the lowest W3 ground-truth score (GT = 0.2), ordered from left to right by their corresponding predicted W3 score.

Although examples in Fig. 3 are labeled as having a highly planar composition, we can observe a gradation of flatness. The model captures the difference between artworks in which the elements are arranged on only one plane or a few successive parallel planes. In the Supplemental material (Figure S1 - S5) we provide all test set examples for all five tasks, annotated with the values of ground-truth and predicted scores.

Because the tasks of predicting Wölfflin's features are introduced for the first time in this work, the evaluation of our approach cannot rely only on the performance of the models on the rather small test sets. Therefore, besides analysing the predictive strength of the models on the small-sized test sets, we employ the best performing models for each task on the large WikiArt dataset and analyze qualitative results and how the predicted values relate to well-known trends in art history.

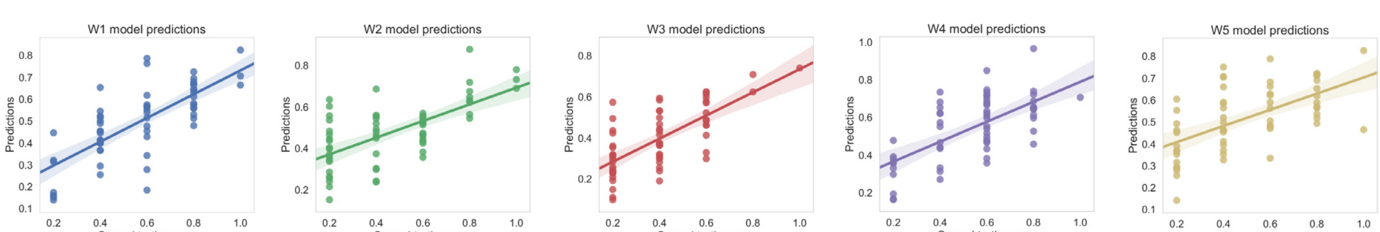

Fig. 2. Regression plots of the best performing model for each feature value prediction task (W1-W5).

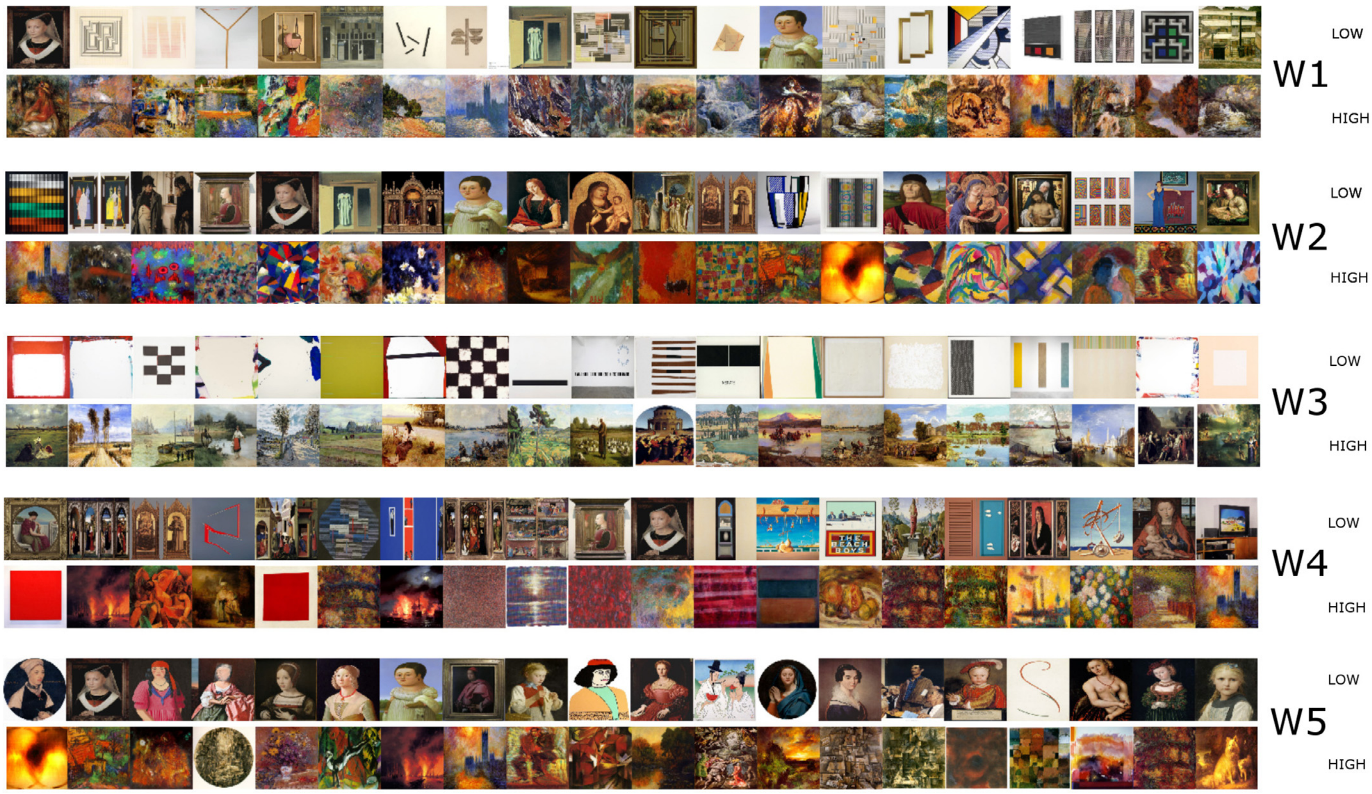

Fig. 4. Artworks from the WikiArt dataset with the top 20 lowest (upper row) and top 20 highest (bottom row) predicted values for each of the five Wölfflin's features (W1-W5).

5.2. Qualitative analysis of the predicted Wölfflin's features

To evaluate how the models generalize to a new dataset, we employ the fine-tuned models on the large WikiArt dataset. For each of the five Wölfflin's features, we analyze the images with the highest and lowest prediction scores in order to assess the performance of the models. Fig. 4 shows artworks with the top 20 lowest (upper row) and top 20 highest (bottom row) predicted scores for each feature (W1-W5).

When looking into the images with the highest and lowest scores for the W1 feature, we can observe that the images with low scores are indeed dominantly linear in style, as well as comprise both abstract and renaissance paintings. This diversity among artworks with low W1 scores indicates that the model learned to distinguish the particular feature of linearity and not some other content-related feature. Also, all artworks with high values of the W1 feature are exclusively painterly in style. In the case of the W3 feature, we can observe that all images with high values tend to show landscape paintings with elements that appear to diminish in size and create an illusion of depth. In contrast, artworks with low values of W3 are mostly abstract paintings characterized by simple forms arranged only on one plane. When looking at images with high scores of the W4 feature, we can observe two interesting patterns of unity in style - one is achieved by representing one dominant element, while the other is achieved by representing multiple strongly interwoven elements.

Besides exploring images with the lowest and highest predicted values of different features, we analyze images from the WikiArt dataset with different score values for each feature in order to understand the relation of visual style and the gradual increase of different feature scores. Fig. 5 show five examples with different values of predicted scores for each of the five Wölfflin's features.

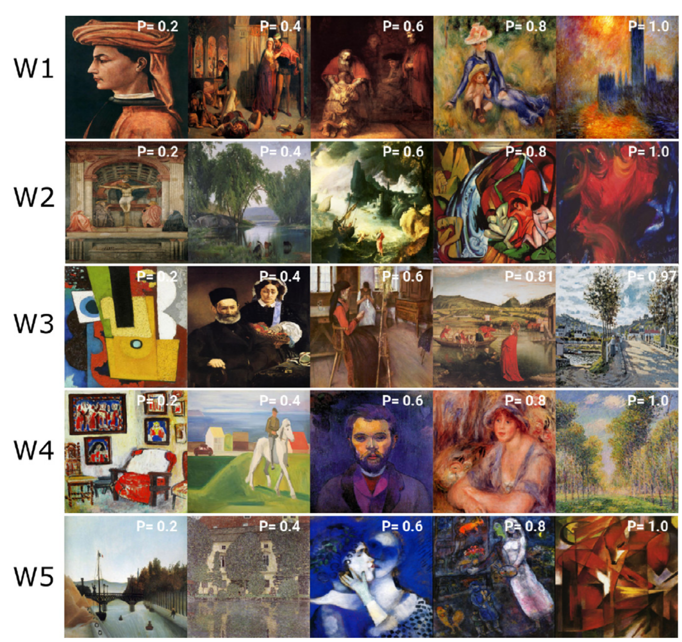

Fig. 5. Five examples with different (gradually increasing) values of predicted scores for each of the five Wölfflin features.

In the case of W1, we can see a clear transition from a purely linear to an intensely painterly style. Examples for different W2 score values show how images with lower values tend to have a more static composition, while images with higher values have a more dynamic composition with dominant diagonal lines. In the case of W3, we can observe a gradation in the compositional ar-

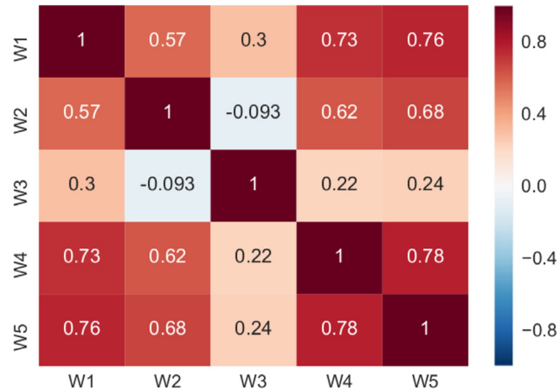

Fig. 6. Heatmap of Spearman's correlation coefficients between the predicted values of different features on the WikiArt dataset (p-values < 0.01).

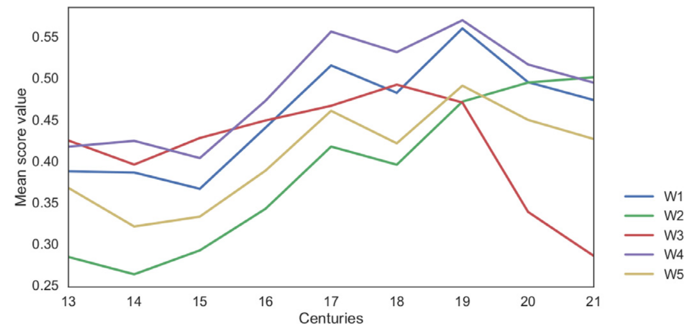

Fig. 7. Chronological plot of the per century mean values of the predicted Wölfflin's features' values on the WikiArt dataset.

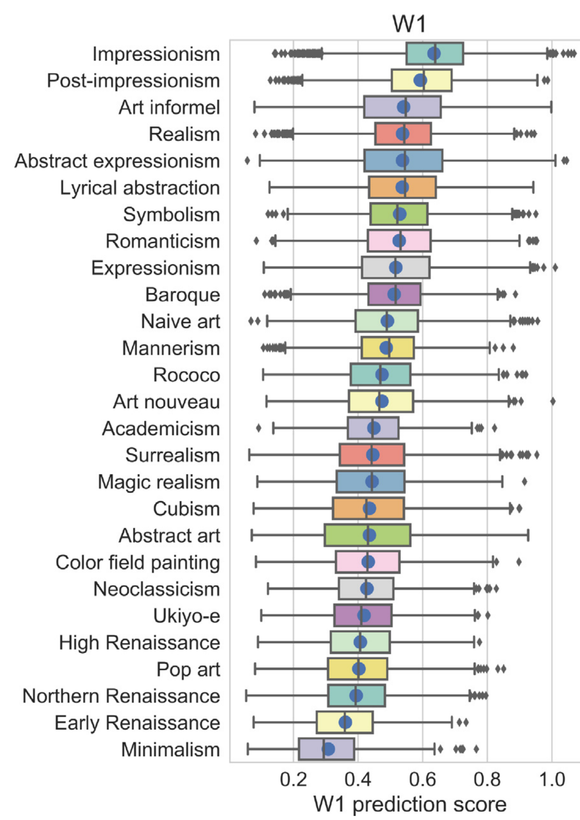

Fig. 8. Box plot distribution of the W1 (linear vs. painterly) feature prediction scores across artistic styles in the WikiArt dataset.

rangement of figures regarding the picture plane, ranging from artworks where all elements are placed only on one plane to artworks that achieve a full illusion of depth. Similarly for W4, the selected artworks indicate a transition from distinct to fused elements, while for W5 the images present a progression from representing elements clearly to intentionally avoiding objective clearness.

Analyzing the progression of style according to the change of predicted values also indicates similarity and overlapping for the different Wölfflin's features. It is known that some concepts are often mutually related. For instance, stylistic properties of unity and relative clarity usually imply a dominant painterly style. In order to better understand how different concepts relate to each other, we analyze the correlation between different feature scores. Fig. 6 shows the Spearman's correlation coefficients between the predicted scores of different Wölfflin's features on the WikiArt dataset.

Values of the correlation coefficients confirm that the features are mutually highly correlated. Particularly W1 feature (linear vs. painterly) is strongly correlated with W4 (multiplicity vs. unity) and W5 (absolute vs. relative clarity). On the other hand, feature W3 (planar vs. recessional) has the weakest correlation with other features. This is consistent with the findings presented by [2], where the correlation analysis between different modes of variations with Wölfflin's concepts often indicate different trends for the W3 feature in comparison to other features. Moreover, the obtained correlation results between the predicted values of different features on the WikiArt dataset are consistent with the correlation results of ground-truth scores.

5.3. Wölfflin's features in the context of art history

One necessary aspect that has to be verified in order to conclude that the predicted features are meaningful and truly correspond to Wölfflin's concepts, is the chronological behavior of the features values. To analyze how the feature values change over time we use a subset of 82 000 images from the WikiArt dataset that include the information about the year of creation. We group the images according to different centuries and calculate the mean of predicted W1-W5 values for each century. Fig. 7 shows how the per century mean values of the predicted Wölfflin's features change over time.

An important insight arising from the chronological ordering of mean values is that they ascent from the 15th to the 17th century for all the features. This corresponds to Wölfflin's theoretical analysis of the changes in style from Renaissance to Baroque. He defined the five key visual principles as contrasting concepts and the difference between Renaissance and Baroque as a shift of dominance of one concept to the other. Therefore, the rise of mean values indicates the predominance of the second concept and represent a confirmation of the accuracy of the learned features. Apart from the changes of values from Renaissance and Baroque, the high mean values of features in the 19th century is also in line with the commonly known stylistic properties of the fine art in that era. Particularly, having in mind the strong predominance of painterlines in impressionistic paintings.

In order to explore in more detail how the values of different Wölfflin's features change according to different styles, we use a subset of images from the WikiArt collections that belong to 27 different style categories. We calculate the mean of predicted W1-W5 values for each style. The per style distribution of the predicted W1 (linear vs. painterly) values is shown in Fig. 8. The boxes are ordered by the mean score marked with a blue dot. In the Supplemental materials we provide the per style distributions for other features (W2-W5).

As expected because of their strong painterly style, Impressionism and Post-impressionism have the highest average W1 score.

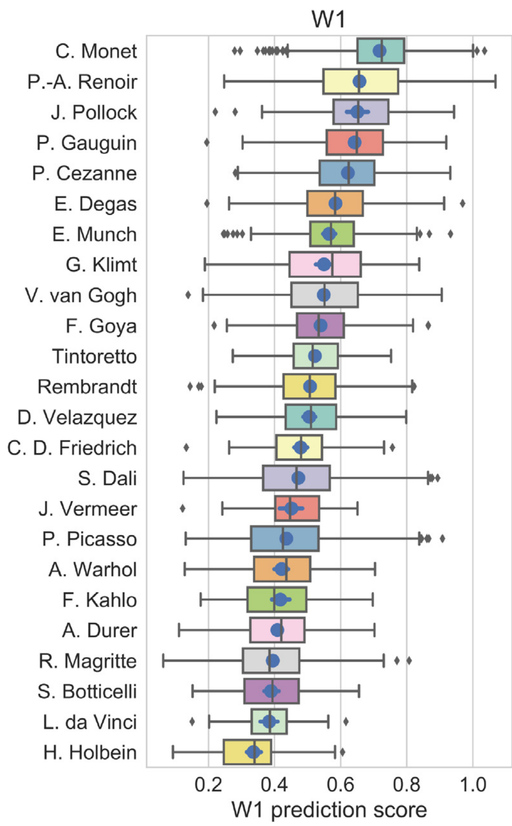

Fig. 9. Box plot distribution of the W1 (linear vs. painterly) feature prediction scores across different artists in the WikiArt dataset.

Renaissance, as well as Minimalism and Pop-art are among the styles with the lowest mean W1 values. Although Renaissance styles and abstract and contemporary styles such as Minimalism and Pop-art differ significantly regarding the subject matter and other visual properties, one common characteristic of these styles is indeed strong linearity expressed with low W1 values.

Apart from addressing the changes between different styles, we also analyze how the predicted Wölfflin's features values change in regard to various artists. The WikiArt collection includes artworks by more than 2000 different artist. For the purpose of our exploration, we choose a subset of 24 well known artists, belonging to different historical art movements. Box plots in Fig. 9 show the distribution of the predicted W1 scores for different artists. In the Supplemental material we provide the per artist distributions for other features (W2-W5).

The distribution of scores per artists reveals that impressionist painters Monet and Renoir have the highest mean W1 score values, which corresponds to their highly painterly style. On the contrary, Renaissance artists have the lowest mean W1 scores. Interestingly, Wölfflin states in his book that "Leonardo is more linear than Botticelli" ([1], page 30.) and the ordering of the predicted W1 mean values reflects this subtle difference.

6. Conclusion

In this article we introduce new high-level image features that quantify stylistic properties of paintings according to Wölfflin concepts. Deep learning based quantitative approaches are employed for the first time in order to predict the values of the newly introduced features. We use an existing small-sized annotated dataset to train five convolutional neural network regression models to predict the values of each of the five Wölfflin's features. Furthermore, we employ the trained models on a large collection of paintings and explore how the predicted values of the five Wölfflin's features correlate with each other, as well as how they relate to different styles and artists. We provide an analysis of qualitative results based on exploring images with high and low values of each feature. The outcome of our analysis suggests that the models learn to discriminate meaningful features that correspond to the visual characteristics described by Wölfflin. This indicates that our models can be used to enhance the search capabilities of online fine art collections. Particularly, by enabling users to search according to art historically well-known and relevant attributes such as Wölfflin's comparative principles. In our future work we plan to explore if further improvement can be achieved by using deep models of different architectures. We also aim to intensify our interdisciplinary collaboration and investigate the applicability of the presented models to concrete art history-related research challenges.

Declaration of Competing Interest

The authors declare no conflict of interest.

Supplementary material

Supplementary material consists of nine figures. Figures S1-S5 show examples from the test set of each task, marked with values of ground-truth and predicted scores. Figures S6-S9 represent box plot distribution of the W2-W5 features prediction scores across artistic styles and artists in the WikiArt dataset. Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.patrec.2019.11.008.

References

[1] H. Wölfflin, Principles of Art History, New York, NY: Dover Publications Inc., 1950.

[2] A. Elgammal, B. Liu, D. Kim, M. Elhoseiny, M. Mazzone, The shape of art history in the eyes of the machine, in: Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, New Orleans, Louisiana, USA, February 2–7, 2018, 2018.

[3] S. Karayev, M. Trentacoste, H. Han, A. Agarwala, T. Darrell, A. Hertzmann, H. Winnemoeller, Recognizing image style, in: British Machine Vision Conference, BMVC 2014, Nottingham, UK, September 1–5, 2014, BMVA Press, 2014.

[4] Y. Bar, N. Levy, L. Wolf, Classification of artistic styles using binarized features derived from a deep neural network, in: Computer Vision - ECCV 2014 Workshops - Zurich, Switzerland, September 6–7 and 12, 2014, Proceedings, Part I, 2014, pp. 71–84, doi:10.1007/978-3-319-16178-5_5.

[5] O.E. David, N.S. Netanyahu, Deeppainter: painter classification using deep convolutional autoencoders, in: Artificial Neural Networks and Machine Learning - ICANN 2016 - 25th International Conference on Artificial Neural Networks, Barcelona, Spain, September 6–9, 2016, Proceedings, Part II, Springer, 2016, pp. 20–28.

[6] R.B. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, in: 2014 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2014, Columbus, OH, USA, June 23–28, 2014, 2014, pp. 580–587, doi:10.1109/CVPR.2014.81.

[7] C. Hentschel, T.P. Wiradarma, H. Sack, Fine tuning CNNs with scarce training data adapting imagenet to art epoch classification, in: 2016 IEEE International Conference on Image Processing, ICIP 2016, Phoenix, AZ, USA, September 25–28, 2016, IEEE, 2016, pp. 3693–3697, doi:10.1109/ICIP.2016.7533049.

[8] B. Saleh, K. Abe, R.S. Arora, A. Elgammal, Toward automated discovery of artistic influence, Multimed. Tools Appl. 75 (7) (2016) 3565–3591.

[9] E. Cetinic, T. Lipic, S. Grgic, Fine-tuning convolutional neural networks for fine art classification, Expert Syst. Appl. 114 (2018) 107–118.

[10] B. Seguin, C. Striolo, I. diLenardo, F. Kaplan, Visual link retrieval in a database of paintings, in: Computer Vision - ECCV 2016 Workshops - Amsterdam, The Netherlands, October 8–10 and 15–16, 2016, Proceedings, Part I, Springer, 2016, pp. 753–767, doi:10.1007/978-3-319-46604-0_52.

[11] N. van Noord, E. Hendriks, E. Postma, Toward discovery of the artist's style: learning to recognize artists by their artworks, IEEE Signal Process. Mag. 32 (4) (2015) 46–54.

[12] T. Mensink, J. Van Gemert, The Rijksmuseum challenge: museum-centered visual recognition, in: Proceedings of International Conference on Multimedia Retrieval, ACM, 2014, p. 451.

[13] G. Strezoski, M. Worring, Omniart: a large-scale artistic benchmark, ACM Trans. Multimedia Comput.Commun. Appl. (TOMM) 14 (4) (2018) 88.

[14] S. Bianco, D. Mazzini, P. Napoletano, R. Schettini, Multitask painting categorization by deep multibranch neural network, Expert Syst. Appl. (2019).

[15] E.J. Crowley, A. Zisserman, In search of art, in: Computer Vision - ECCV 2014 Workshops - Zurich, Switzerland, September 6–7 and 12, 2014, Proceedings, Part I, 2014, pp. 54–70, doi:10.1007/978-3-319-16178-5_4.
[16] A. Brachmann, E. Barth, C. Redies, Using cnn features to better understand what makes visual artworks special, Front. Psychol. 8 (2017) 830, doi:10.3389/fpsyg.2017.00830.
[17] E. Cetinic, T. Lipic, S. Grgic, A deep learning perspective on beauty, sentiment, and remembrance of art, IEEE Access 7 (2019) 73694–73710.
[18] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, T. Darrell, Caffe: convolutional architecture for fast feature embedding, in: Proceedings of the 22nd ACM International Conference on Multimedia, ACM, 2014, pp. 675–678, doi:10.1145/2647868.2654889.
[19] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, 2012, pp. 1097–1105.