Simon Grützmacher*, Ralf Kemkemer, and Cristóbal Curio

Using Deep Correlation Features to define the Meta Style of Cell Images for Classification

https://doi.org/10.1515/cdbme-2019-0058

Abstract: Digital light microscopy techniques are among the most widely used methods in cell biology and medical research. Despite that, the automated classification of objects such as cells or specific parts of tissues in images is difficult. We present an approach to classify confluent cell layers in microscopy images by learned deep correlation features using deep neural networks. These deep correlation features are generated through the use of gram-based correlation features and are input to a neural network for learning the correlation between them. In this work we wanted to prove if a representation of cell data based on this is suitable for its classification as has been done for artworks with respect to their artistic period. The method generates images that contain recognizable characteristics of a specific cell type, for example, the average size and the ordered pattern.

Keywords: Machine Learning, Cell Imaging, Structural Images

1 Introduction

In the field of art, there are some well-known artists with their own artistic style. For humans it is relatively easy to distinguish between individual styles and identify the artist when looking at another painting if the same artistic style is used. This ability is based on the fact that the human visual system is highly adapted for recognizing structural information in images. As Chu et al. have shown in [1] it is possible to match paintings due to their structural information to their corresponding period of art. However, structural information is not only inherent to artistic work like paintings but can also be found in images of biological samples. Different cell types, growing in a cell culture dish, have different structural features, including for instance cell size and polarity as well as the nematic pattern of the cell monolayer [2]. These structural features may allow a skilled person to identify the type of cells visually and to assess if the culture is healthy or if there are already changes due to prolonged culture periods or contamination. However, a human expert is always vulnerable to bias due to various factors including lack of expertise, when working with a new cell type, and routine. Additionally, a manual visual inspection of cell cultures does not apply to high throughput approaches with large sample numbers. Ideally, an automated method enables inspecting cell cultures over time and can distinguish between different cell types in-vitro. To address this challenge, we employ structural images of phase contrast microscopy images of two different cell types to teach a classifier about the general structure of them. We will not only show that such a classifier is useful to distinguish the two cell types, a generally used model cell line for collective cell behavior (MDCK) and a breast cancer model cell line (MCF-7), but further that the use of structural images leads to better classifications.

***Corresponding author: Simon Grützmacher,** Dept. of Informatics, Reutlingen University, Reutlingen, Germany, Simon.Gruetzmacher@reutlingen-university.de
Ralf Kemkemer, Dept. of Applied Chemistry, Reutlingen University & Max Planck Institute for Medical Research, Heidelberg, Germany
Cristóbal Curio, Dept. of Informatics, Reutlingen University

2 Style Extraction

2.1 Architecture for Learning Style

The recognition of artistic style is a simple task for a human but is based on complex processes, and thus, it is challenging to teach a machine to extract, identify, and compare different styles. The approach by Gatys et al. [3] is addressing this problem by extracting the structure of an image and transferring it to an image with another content. For our purpose, we only use the extraction of structure from a given image in order to train our classifier. For extracting the structural information of an image we use a pre-trained VGG19 [4] neural network and the method used by Gatys et al. [3] outlined as follows. The structure of an image is determined with the help of the correlations between the different features in different layers of the CNN. The chosen depth dictates the level of detail in the structural representation as can be seen in Figure 1. The image structure for classification is extracted with the help of a white noise image and the distance between its gram-matrix and the gram-matrix of the respective CNN feature maps is iteratively minimized. The gram matrix $G^l \in \mathbb{R}^{N_l \times N_l}$ is defined by G^l_{ij} denoting the inner product between the vectorized feature maps i and j in layer l, with index k denoting the po-

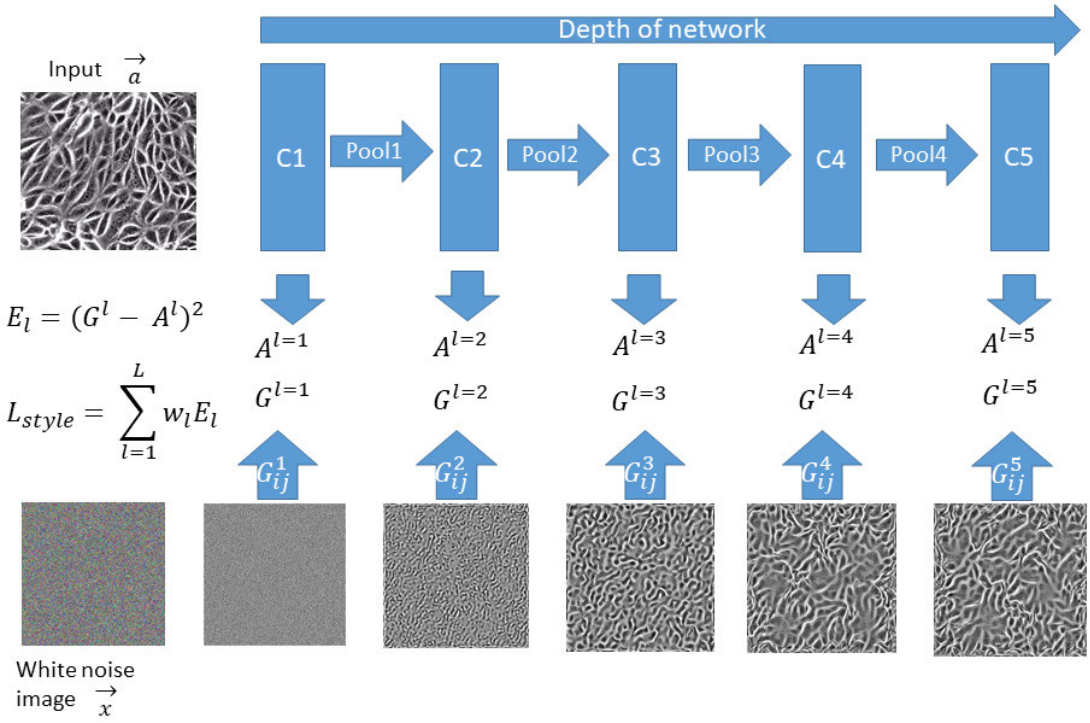

Fig. 1: Style extraction adapted from [3]. The detail of the generated structural images is increasing with the depth of the network.

sition in the vectorized feature, where F^l_{ij} is the activation on the i^{th} filter at position j in layer l, and reads as follows

$$G^l_{ij} = \sum_k F^l_{ik} F^l_{jk}. \quad (1)$$

The weight factors w_l determine the contribution of loss E_l for each layer to the overall loss L_{style} (Figure 1, left). The image $\vec{a}$ for which we are seeking the style for is passed through the network, and its structural representation A^l for all L layers are computed and stored, which can be seen on the upper row of Figure 1.

As such the white noise image is iteratively reconstructed to match the gram matrix of the deep correlation features extracted for each layer in the network. Further details are given in Gatys [3]. The structural images for our classifier were generated from three different videos, one video for the MDCK (Madin-Darby Canine Kidney) cells and two different videos for the MCF-7 (Michigan Cancer Foundation-7 - breast cancer) cells. The original images were taken from random frames and structural images of 786x786px resolution were computed. We chose this size to obtain a better visual representation of the structure not only for demonstration purposes, also we aim to investigate the similarity of the structural images depending on the augmentation of the original image in further work. A higher resolution increases the necessary time to compute the structural images, but also increases the level of detail in these images. Since it is possible to use a different depth in the network to generate the structural representation we decided to use a depth of four layers since the representation with a depth of five layers does not increase in visible structural information but in increased processing time. To ensure a relative similarity between the structural images we initialize the white noise image $\vec{x}$ and use it for all training, test and evaluation images. Our hardware setup consists of a TITAN Xp GPU, an Intel Core i7-8700K CPU, and 32 GB RAM. The GPU performs the training of the network as Keras offers support for CUDA enabled GPUs. The processing time for one structural image was roughly 1 hour.

2.2 Architecture of the Classifier

For classifying the structural cell images, we use the VGG16 architecture [4]. For classification, we chose two cell types with similar structural images, namely MDCK and MCF-7 cells. Since the structural images do not change a lot when the original image is augmented, we further rotated them by 180 degrees, flipped them vertically and horizontally to increase the amount of training and test data. This alteration has to be done carefully, since it is not guaranteed that the phase contrast images are invariant to rotation. The result of the augmented original data and the corresponding structural images can be seen in Figure 2. For classification the structural images were re-sized from 786x786px to 200x200px in order to

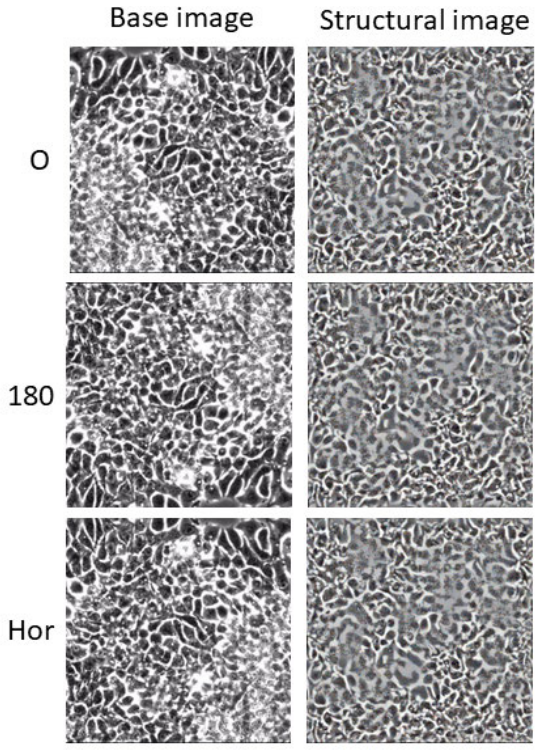

Fig. 2: Augmented cell images and their corresponding structural images. The applied augmentations are: O = Original with no augmentation, 180 = rotated by 180 degrees, Hor = horizontally mirrored

fit the VGG16 architecture. As an optimizer, we used stochastic gradient descent and as loss for the classifier the binary cross-entropy function.

3 Results

The data used for training, testing and evaluation are presented in Table 1. Note that the amount of data is limited because of the lack of more cell images from these cell types. To compensate this problem, we used image augmentation to create more data as can be seen in Figure 2. The output of our network is a classification of the used structural image of a learned cell type. Since cells do have a distinct structural pattern, the use of structural images to classify them seems adequate. Another advantage is the loss of artifacts within the original cell images since single artifacts like a bright spot or a blurred part do not affect the global structure of the cell image.

3.1 Evaluation

For the evaluation of our classifier, we used generated structural images from the two trained cell types, images of both cell types can be seen in Figure 3. One of the problems of cell

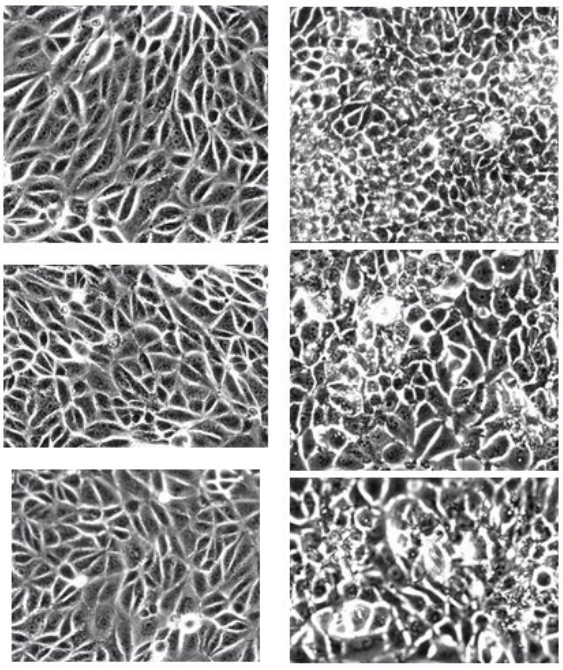

Fig. 3: Exemplary microscope images of cell types to classify. Left: MDCK cells, right: MCF-7 cells.

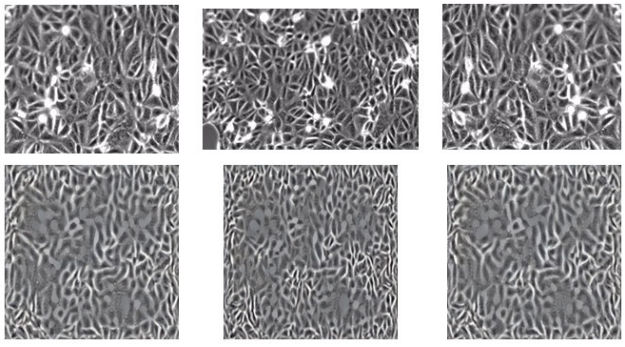

Fig. 4: Falsely classified MDCK images. Original images on top, respective structural images on bottom.

classification is that the alignment of cells. MDCK and MCF-7 cells can both appear in a round or elongated form. To distinguish between them, it is essential that the cell images have a fixed form when an expert is looking at them, since rotation or mirroring can affect the ability of an expert to classify them. Therefore, we use the structural images since it does not have a big impact how the original images have been altered. Also, cells from different videos may vary in brightness and quality of the recording. These factors can influence a human expert, but for the structural images, these differences result in almost no difference. All of the used evaluation images have not been used in the training or test data. To ensure an accurate evaluation, all original images of the cells have been augmented randomly, an example can be seen in Figure 2 . The evaluation results can be seen in Table 2.

Tab. 1: Used structural images for training, test and evaluation.

	MDCK	MCF-7
Training	**44**	**29**
Test	**27**	**20**
Evaluation	**16**	**11**

The result indicates that the classification of confluent grown cells by the learned meta-structure of cells is suitable. Since we used augmented images, it shows that structural images contain enough structural information of the cell types to classify them later correctly. Since there were only three falsely classified images, we investigated why there was a false classification with these three images. The three images classified as false can be seen in Figure 4, along with the original images. The top row shows the utilized original cell images, whereas the bottom shows corresponding structural images. Note that a high amount of fragments (observable as bright spots, which may be dead cells) compromise the global structure if they appear too often, the correctly classified MDCK cells do have less bright spots. If compared with the correctly classified images it can be seen that one difference is the lack of dead cells. To support this theory, we investigated three

Tab. 2: Result of the evaluation of the classifier trained with structural images.

	MDCK	MCF-7
MDCK	**13**	**3**
MCF-7	**0**	**11**

correct classified MCF-7 cell images, which can be seen in Figure 5. MCF-7 cells are cancer cells, during cell division cell contract and appear as white spots in the images. Whereas the bright spots in the MDCK cell line mostly indicate dead cells, the structural images cannot differentiate between those two cases. The bright spots lead to round grey spots in the structural images and can lead to false classification. For further evaluation, we used the original images, from which we generated the structural images, to train the classifier and the evaluation set. The result can be seen in Table 3. All of the falsely classified images were augmented with a 180 degrees rotation which can be a problem since not all cell types are invariant to rotation.

Tab. 3: Result of the evaluation of the classifier trained with original images.

	MDCK	MCF-7
MDCK	**14**	**2**
MCF-7	**6**	**5**

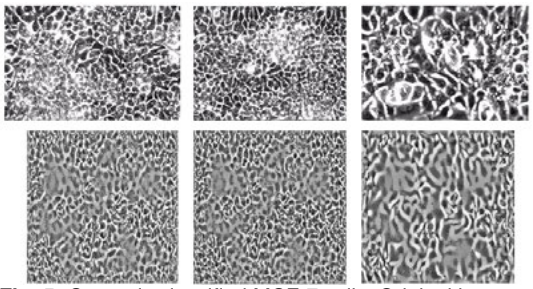

Fig. 5: Correctly classified MCF-7 cells. Original images on top, respective structural images on bottom.

4 Further Work

Additional different cell types will be used for classification in order to demonstrate the versatility of our tool in distinguishing between different cell types under in vitro conditions. In addition, further work will evaluate at which point artifacts in the original images are too numerous and affect the global structure of the cell image, and, therefore, the computed structural representation. Also, we aim to investigate the similarity of structural images, this is important since a human can see the similarity, yet it is challenging to quantify this similarity.

Acknowledgment: S.G. receives funding from "Kooperatives Promotionskolleg IPMB Reutlingen-Tübingen" (MWK, Baden-Württemberg). The project was partly funded by Reutlingen University.

Author Statement

Conflict of interest: Authors state no conflict of interest. Ethical approval: The conducted research is not related to either human or animals use.

References

[1] Wei Ta Chu and Yi Ling Wu. Image style classification based on learnt deep correlation features. *IEEE Transactions on Multimedia*, 20(9):2491–2502, 2018.

[2] R. Kemkemer, D. Kling, D. Kaufmann, and H. Gruler. Elastic properties of nematoid arrangements formed by amoeboid cells. *The European Physical Journal E*, 1(2):215, 2000.

[3] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. Image Style Transfer Using Convolutional Neural Networks. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016-Decem:2414–2423, 2016.

[4] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In Yoshua Bengio and Yann LeCun, editors, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015.