# TER2021 – 033 Deep Learning to paint like Van Gogh

Author: Stefano D'Angelo – Data Science (Master 2 EIT Digital)
Supervisor: Frédéric Precioso
Université Côte d 'Azur, Inria, Polytech Nice Sophia – Nice, France

## INTRODUCTION

The objective of this paper is to transfer the artistic style from artworks, such as paintings, to real images to transform them into new artworks. The principle is to start from real photos and apply the **style of Van Gogh** to transform the initial photo into a painting-like picture.

In this way, new data can be generated when real data is missing, and the dataset containing all of Van Gogh's paintings can be augmented.

There have been several techniques proposed to transfer artistic style, from **Convolutional Neural Networks** to **Generative Adversarial Networks**. However, most of them do not take into account class information related to the objects present in both images. Hence, this work tries to **improve** the state-of-the-art results, comparing different approaches involving semantic segmentation.

## PROPOSED APPROACHES

In this work, we explored three methods to transfer the style from a painting to a real photo, taking into account the semantics of both the images.

We first started from a paired dataset, containing couples of samples coming from two different domains, but that in some way share the same content. These domains are respectively the set of Van Gogh's paintings and some real pictures.

To extract the semantics of the images the approach followed was the one coined by Penhouët et al. [1], that, differently from other methods having the goal of segmenting everything [2], uses image segmentation and semantic grouping to merge minority classes in order for the masks of each pair of images to match.

Since all the pre-trained segmentation models have been trained on real pictures, it is hard to directly segment paintings. For this reason, we investigated an approach that converts the paintings to real images and then computes segmentation masks. Then, three different pipelines have been tested:

1. In the first one, all the masks and the corresponding images are cropped. Then, for each patch of the real photo, we retrieve the patch of the painting having the most similar semantic content. Based on this choice, we transfer the style from the corresponding patch of the painting to the patch of the photo. In the end, we recompose the whole image.

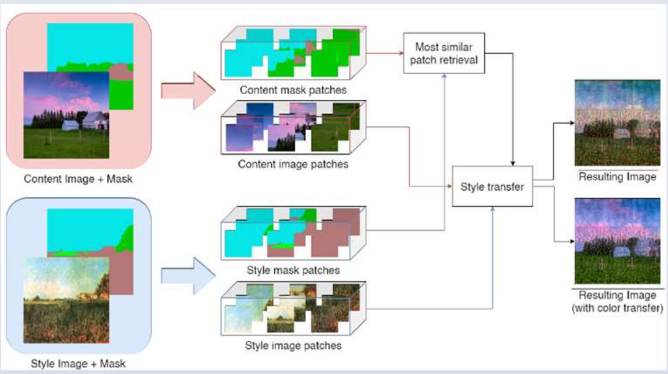


**Fig. 1**: Style is transferred patch-by-patch from the style image to the content image basing on the Euclidean Distance between the patches of both the content mask and the style mask.

2. The second pipeline is based on Champandard's model [3], where we transfer the style re-arranging the instances of the objects in the style image according to the semantic mask of the content image.

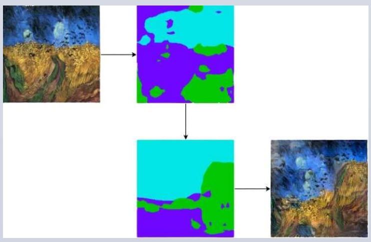

**Fig. 2**: The composition of each element in the image is re-shaped based on the semantics of the real photo.

3. In the third and last approach, the style is transferred in the real photos domain, and the resulting image is then converted back to the paintings domain.

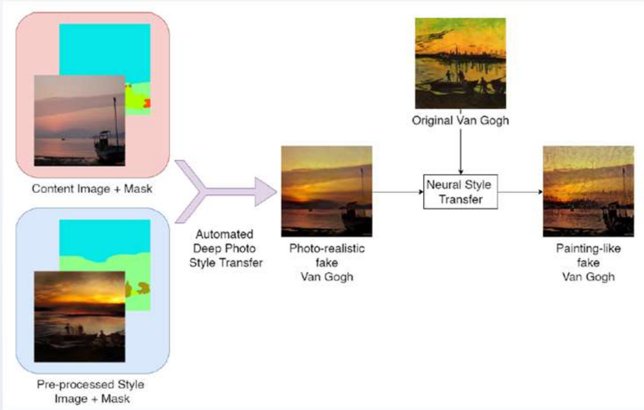


**Fig. 3**: In this workflow, the style is first transferred in the photo-realistic domain from the painting to the photo. The resulting image and the original painting are then fed into the Neural Style Transfer model [4] as content and style images, respectively.

All the segmentation masks of the painting have been generated from a pre-processing step, which is actually the main novelty added to the state-of-the-art by this work. In this phase:

1. Paintings have been converted into real photos through a CycleGAN [5]
2. Both paintings and photos have been segmented, and their masks have been *semantically* grouped [1].

## RESULTS

The last model has shown to be the best, improving even the state-of-the-art. In this work, three state-of-the-art models have been used as a source of comparison: Neural Style Transfer[4], CNNMRF[6], and CycleGAN [5]. A couple of results are shown in the following figure.

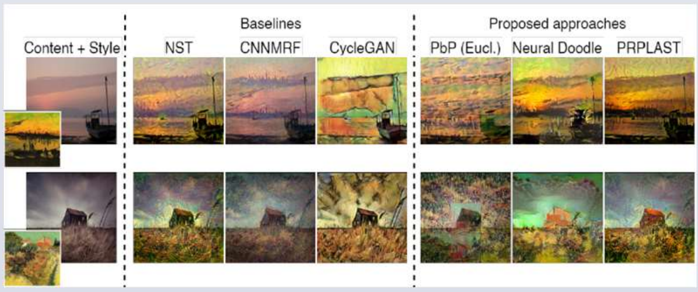


**Fig. 4**: Comparison between all the three models and the state-of-the-art.

## REFERENCES

[1] Sebastian Penhou¨et et al. Automated Deep Photo Style Transfer. 2019. arXiv: 1901.03915 [cs.CV].

[2] Ronghang Hu et al. Learning to Segment Every Thing. 2018. arXiv: 1711. 10370 [cs.CV].

[3] Alex J. Champandard. Semantic Style Transfer and Turning Two-Bit Doodles into Fine Artworks. 2016. arXiv: 1603.01768 [cs.CV].

[4] Leon A. Gatys et al. A Neural Algorithm of Artistic Style. 2015. arXiv: 1508.06576 [cs.CV].

[5] Jun-Yan Zhu et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. 2020. arXiv: 1703.10593 [cs.CV].

[6] Chuan Li and MichaelWand. Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis. 2016. arXiv: 1601.04589 [cs.CV].

## CONTACTS

Stefano D'Angelo – *stefano.dangelo.ct@gmail.com*
Frédéric Precioso – *frederic.precioso@univ-cotedazur.fr*

Université Côte d'Azur, Nice, France (https://univ-cotedazur.fr)
Inria, Sophia Antipolis (Nice), France (https://www.inria.fr/fr)