Project Proposal: 3MD3020- Deep Learning
Neural Style Transfer

Bergaoui, Khalil
khalil.bergaoui@student.ecp.fr

Hachicha, Mohamed Amine
mohamed-amine.hachicha@student.ecp.fr

1. Motivation and Problem Definition

Applying deep learning techniques to perform artistic style transfer, refered to as Neural Style Transfer (NST), has become a trending topic both in academic literature and industrial applications. It is receiving increasing attention and a variety of approaches are proposed to either improve or extend the original NST algorithm.[6].

The Style Transfer problem can be formulated as follows:

Given a set M of input images, and **Given** a style image S, **Generate**, for each input image I, an output image S(I) that adapts the style S to the image I.

Figure 1. Example of an application of NST with different styles

In practice, Neural Style Transfer enables us to guess how would Claude Monet or Vincent Van Gogh, for example, paint an image with their respective artistic styles.

It is already used to make interesting computer generated artwork by transforming an ordinary camera taken landscape into automatically generated paintings with different artistic styles (see https://deepart.io/ for example). The use of NST is also quite popular in photo editing software to create photo filters for social media like Instagram and Snapchat.

In our project, we suggest to compare different approaches to achieve NST: for now we have identified two main popular approaches: the first one is based on a Convolutional Neural Network (CNNs) architecture while the second one is mainly based on Generative Adversarial Networks (GANs).

We will base our first experiments on a collection of paintings of some of the most influential artists. We will then define a specific style according to a given artist. The full dataset can be found here: https://www.kaggle.com/ikarus777/best-artworks-of-all-time .

2. Methodology

Our first step is to read and understand recent scientific publications about implementations of NST. We have already identified two widely used approaches in the literature.

Once a thorough review is conducted, we propose to efficiently implement the two identified approaches and experimentally evaluate the results by comparing their performances in terms of visual output, memory and time efficiency, both during training and inference. Additionnally, if possible, we propose to explore a common quantitaive metric that allows to compare visual rendering of both approaches[3] .

2.1. CNN based approach

Deep CNN based architectures are commonly used to perform NST. This is the case for [1] and [4].

In [1], the authors assume that many artists have barely the same artistic style (impressionist artists for example) and therefore they use an N-style feedforward network where each style image is converted into a vector in an embedding space. The algorithm tries to produce an output that is similar to the input image in content (represented by high-level features) and similar to the style image in style (meaning in low-level features).

Therefore, in [1], the neural network has "style layers" and "content layers". The gradient descent tries to find an output image that minimizes a weighted sum of the style loss and the content loss (loss weights are themselves hyperparameters of the model).

In [4], the authors present a hierarchical multi-modal deep convolutional neural network architecture that is able to learn both large-scale texture distortion (input image content) and the fine brushwork of the style image. The network combines multiple models allowing the architecture to handle large size images.

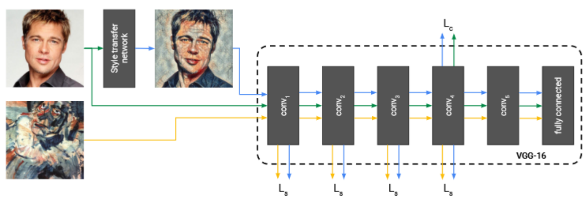

Figure 2. Style Transfer network training diagram

2.2. GAN based approach

Generative adversarial networks (GANs) have been widely studied for image generation and manipulation task. When applied to Style Transfer, these networks allow solving less constrained problems, compared to methods based on purely convolutional networks which only work with limited number of styles, and cannot apply to an unseen style images. Recent approaches, based on GANs, are therefore designed for arbitrary style transfer, where both the content and the style inputs can be unseen images [5].

In [5] , the authors explore adversarial training for arbitrary style transfer. In this case, an encoder-decoder architecture is used and the adversarial network learns the intrinsic property of image styles from large-scale multi-domain artistic image.

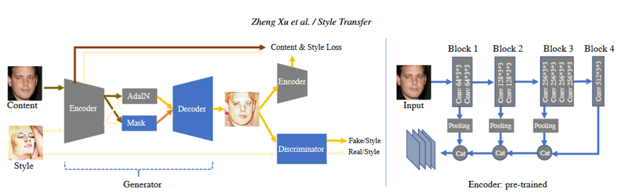

Figure 3. GAN architecture proposed in [5]

The adversarial training is challenging because both the input and output of the generator are diverse multi-domain image. The authors use both conditional generator and conditional discriminator to tackle multi-domain input and outpu

In [2], the authors investigate conditional adversarial networks as a general-purpose solution to image-to-image translation problems. In this perspective, Style Transfer can be considered as a specific case of a wider problem domain referred to as Image-To-Image Translation : The goal is to transform a given image from an input domain X to an output domain Y. For this project, since we are interested in NST, Y represents a particular artistic style. Using this problem formulation, we could investigate adversarial training applied to image to image translation and adapt it to the style transfer domain. We propose to compare the performance of conditional adversarial networks, mainly in terms of visual rendering on a style transfer task, when the adversarial training is being general-purpose (image to image translation) vs domain-specific (artistic style).

3. Evaluation

After implementing the two previously described approaches, we propose to evaluate the experimental results as follows:

On the one hand, we will compare their performances in terms of visual rendering. On the other hand, we propose to compare the runtime during training and inference modes. This last step can be conducted on different input/output image-style pairs.

Additionally, we could explore other quantitative metrics that address the visual aspects of the image obtained after style transfer. We might find inspiration from the literature to help us quantify this, rather qualitative, artistic task.

References

[1] Vincent Dumoulin, Jonathon Shlens, and Manjunath Kudlur. A learned representation for artistic style. *CoRR*, abs/1610.07629, 2016. 1, 2

[2] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks. *CoRR*, abs/1611.07004, 2016. 2

[3] Li Fei-Fei Justin Johnson, Alexandre Alahi. Perceptual losses for real-time style transfer and super-resolution. *CoRR*, abs/1611.07004, 2016. 1

[4] Xin Wang, Geoffrey Oxholm, Da Zhang, and Yuan-Fang Wang. Multimodal transfer: A hierarchical deep convolutional neural network for fast artistic style transfer. *CoRR*, abs/1612.01895, 2016. 1, 2

[5] Zheng Xu, Michael J. Wilber, Chen Fang, Aaron Hertzmann, and Hailin Jin. Beyond textures: Learning from multi-domain artistic images for arbitrary style transfer. *CoRR*, abs/1805.09987, 2018. 2

[6] Zunlei Feng-Jingwen Ye Yizhou Yu Mingli Song Yongcheng Jing, Yezhou Yang. Neural style transfer: A review. *CoRR*, abs/1610.07629, 2019. 1