"A Neural algorithm of Artistic style" a Revisit

Giovanni Regina, Riccardo Viviani

March 2021

1 Introduction

This work is a revisit of the Paper "A Neural algorithm of Artistic style" (citazione), made by Riccardo Viviani and Giovanni Regina. Convolutional networks have recently enjoyed a great success in large-scale image recognition, which has become possible due to the large public image repositories, such as ImageNet (citazione), high-performance computing systems, such as GPUs, and the cloud computing. The authors of (citazione di prima paper) exploit the structure of the CNN, in particolar the VGG19 (citazione), to mix the content of an image (like a picture of a building, etc..) with the style of another image (like a portrait or whatever) obtaining a result that makes the former looks like a portrait. We have implemented the network descripted in the paper, with the same architecture and the same settings, on the Google Colab platform (link Colab), that gives the fast (cloud) computing necessary for Deep Neural Network's calclus. Then we made some adjustments, like the introduction of the Total Variation Loss in the calculus of the loss (see Methods section), and we made some try varing the hyperparameters of our network to see which were the better ones in order to achieve better result (see Our work section), like decreasing the total loss or the right style-content mix in the result's appearance. The rest of the paper is organised as follows: In section 2 we explore the architecture of the VGG19 network, in order to understand why the authors have made their choices, and also to change the settings in order to achieve better results. In section 3 we show our results

2 Google's VGG Network

The Google's VGG19 (citazione) is a network that empower the Deep Convolutional Neural Networks's architecture to make high accurate predictions on the image recognition task, even using small convolutional filters (3x3). The achievement is reached by building a Deep Network with 19 hidden layers (originally the were 16, then improved to achieve better performances), trained on the Imagenet dataset , but scalable to other datasets.