HACETTEPE UNIVERSITY

Department of Computer Engineering

FALL 2018

BBM 413 IMAGE PROCESSING PROJECT PROGRESS REPORT

Supervisors: Dr. Erkut Erdem
TA Ozge Yalcinkaya

Authors: Ege Berke Balseven
Can Berk Sürmeli
Mehmet Harun Doğanay

IMAGE COLORIZATION

Ege Berke Balseven	21590776
Can Berk Surmeli	21427363
Mehmet Harun Doganay	21426878
21590776@cs.hacettepe.edu.tr	Hacettepe University
21427363@cs.hacettepe.edu.tr	Ankara, TURKEY
21426878@cs.hacettepe.edu.tr	

November 13th, 2018

Abstract

According to convolutional neural network approach mentioned in that paper(1), colorize grayscale images that combines both global priors and local image features. We have been inspired by this approach and try to apply this technique on gravures. Gravure is a printing method in which an image is applied to a printing substrate by use of a metal plate mounted on a cylinder. So, with this method the images have a unique pattern. When we look at how the project we took as reference, we reached the following informations:
They introduce a novel technique for automatically coloring grayscale images combining both global presets and local image functions. Based on Convolutional Neural Networks, their deep network has a fusion layer that allows us to elegantly link local information based on small image patches to global priorities that use the entire image. The entire framework, including global and local priorities and the colorization model, is rigorously trained. In addition, their architecture can handle images at any resolution, unlike most existing approaches based on CNN. They use an existing, extensive scene classification database to train their model, and use the record's class names to learn global priorities more efficiently and more discriminatively. They validate their approach with a user study and compare it with the state of the art, where they show significant improvements. In addition, they demonstrate their method extensively on many different types of images, including black-and-white photography more than a hundred years ago, and show realistic colorings.

1 Introduction

Conventional coloring requires significant user interaction, whether by placing many color patches, displaying related images, or performing segmentation. Instead, this article proposes a fully automated, data-driven approach to color grayscale images. Their approach uses a combination of global image priorities extracted from the entire image and local image functions computed from small image fields to color an image automatically. Global priorities provide information at an image level, e.g. Whether the image was taken indoors or outdoors, day or night, etc., while local features represent the local texture or object at a particular location. By combining both functions, they can use the semantic information to color, for example, the evening sky or the human skin - all without human interaction.
Their approach is based on Convolutional Neural Networks, which have strong learning capabilities. They propose a novel architecture that can collectively extract global and local features from an image and then combine them to achieve final coloration. They will show that this approach is significantly better only when using local functions. In addition, they can use semantic class names of existing records during training to learn more complex global functions. However, the semantic class designations are not needed when grayscale images

Figure 1: Ottoman Gravures

are colored. They use an existing large-scale scene database to train their model to predict the color saturation of a grayscale image using the CIE L * a * b * color space. Their method requires neither preprocessing nor post-processing: everything is consistently learned.
Their model consists of four main components: a low-level feature network, a mid-level feature network, a global feature network, and a colorization network. Conceptually, these networks work as follows: First, a generic set of low-level shared functions is extracted from the image. These functions compute a number of global image functions and mid-level image functions. Then, both the middle and the global structure are fused together by their proposed "fusion layer" and used as input to a colorization network that outputs the final chroma map. It is needless to say that this is not explicitly implemented as a sequential procedure; It is realized as a single network. Note that neither preprocessing nor postprocessing is performed. It calculates everything in a single step. As a byproduct of their approach they can also classify the scene. While the global features are computed with fixed-size images, their novel approach of merging global and local features enables their model for an arbitrary resolution image, unlike most convolutional neural networks.
Because of the separation between global and local features, it is possible to use global features that are calculated for one image in combination with local features calculated for another image to change the style of the resulting color. For example, if the global features at dusk are calculated on a picture and combined with a picture of a sunny beach, the color of a beach is displayed at dusk. A picture can also look like it was taken in another season. This underlines the flexibility of their model.

1.1 Problem to be addressed

Traditional coloring methods take the grayscale(black and white picture) and colorize picture in accordance with origin. In our approach, besides coloring the images we wanted to apply gravure technique to them. In this method instead of significant user interaction whether in the form of placing numerous color scribbles looking at related images, or performing segmentation. In the (1), they instead propose a fully automated data-driven approach for colorization of grayscale images. They have approached uses a combination of global image priors, which are extracted from the entire image, and local image features, which are computed from small image patches, to colorize an image automatically. Global priors provide information at an image level such as whether or not the image was taken indoors or outdoors, whether it is day or night, etc., while local features represent the local texture or object at a given location. By combining both features, we can leverage the semantic information to, for example, color the dusk sky or the human skin—all without requiring human interaction. This is what they said in (1).

1.2 Related work

The project we refer to in this study consists of a special CNN model. There are two different layers in parallel to this model. One of them is the global features network and the other is the mid level features network. This model uses both global and local information to color the image. First of all, there is also a low level features network. low level features are released to mid level and global features networks. After working these two branches have fusion layer. Combining inferences with the Fusion layer. Then there is the coloring layer. End-to-end adopted a training model. So no pre-processing and post-processing. In order to achieve such a complex model, the train was run with 2.5 million photos. The results are great.

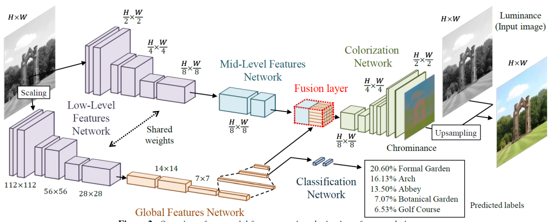


Figure 2: *Overview of our model for automatic colorization of grayscale images.*

1.3 Methodology to be employed

Our approach in this project is to colorize photos customized in a certain style rather than any photo. The example we work with is end-to-end learning. But since we work with a special photographic style, you may need to pre-process first. At this point, we plan to solve the project with two different approaches. The first is based on the method of inputting the learning model to any color image, not in particular colored engraving style. For this reason, the photos in the dataset should be given to the model by a special filtering process. In this way, the end-to-end learning method is intact. Then we will give the input as engraving gravure.

You need to design a special data set for this method. We are working on the coloration of human figures in our study. For this reason, our data set consists of human figures. At this stage, it is aimed to make colored human figures look like colored engravings by using various image processing methods. For this, various filters and filters should be applied.

The second method is based on the style transfer method. We transfer the style of the engraved images to the human images in our data color. In this way, our data consists of colored engraving photographs. Then, the model is trained with the end-to-end learning architecture. The advantages of this method are also there are disadvantages. We will explain them in experimental evaluations.

1.4 Experimental evaluation

Firstly we need to collect our dataset. Our dataset contains human figures and their dresses, scarfs, skirts and so on. We collect these images from google search. We have used a script to download related searchword's images. Our keywords like this:
kaftan
bindalli
abiye
yeniceri kiyafeti
salvar
cubbe modelle
kostum
pardesu
etek
kurk
red dress
green dress
blue dress
pink dress
black dress
yellow dresss

This dataset contains 1500 color images. Some of the images like that:

Behind of that there are many irrelevant data.

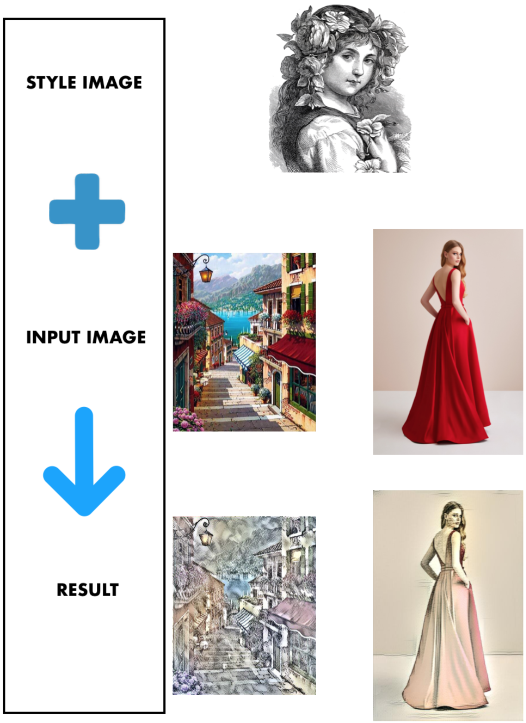
To get the idea behind the Iizuka's idea (1) firstly we tried original human figures. not gravures. We trained our model different configurations. Last work that we have done is 20 epoch with 1500 images in 2 batche size. We have use Google Colab to all our processes. So Nvidia K80 GPU takes nearly 3 hours to train. Let examine the result below:

This figure result of out test in sample from training set.

Then in the figure above, left side is original image, middle is L layer from LAB color space and right side is our result.
Iizuka says in the paper, there are some inherent ambiguity of the colorization problem and they gives that result from their work.

Input Ground truth Proposed

So we mean it is possible to fail specific color patches colorization.



The second approach is style transfer. We are going to do some preprocessing. First of all, we are planning to produce bunch of styled images. We use gravure images as a style image. Then our dataset that we have used in first approach again using. So each image in dataset replaces with style transfered images.

One of the other method for image style transferring.
Explanation of steps for image above:
Basically we are going to use a series of filters and image conversions.
First we downscale the image and then apply bilateral filter to get a cartoon flavour. Then again we upscale the image.

Next step is getting a blurred version of the original image. Now, we don't want the colours to interfere in this process. We only want the blurring of the boundaries. For this, we first convert the image to grayscale and then we apply the media blur filter.
Next step is to identify the edges in the image and then add this to the previously modified images to get a sketch pen effect. For this first we are using adaptive threshold. In last step, we compile the final images obtained from the previous steps.
We expect to get better results. Because we will get gravure textures in images. So learning process will became more accurate. We are going to improve this method upcoming days.

1.5 Reference paper result in gravures

Some of our test cases below. We are going to compare and contrast them in final report. Now you can see some examples here. (1)

References

[1] S. Iizuka, E. Simo-Serra, and H. Ishikawa. Let there be Color!: Joint End-to-end Learning of Global and Local Image Priors for Automatic Image Colorization with Simultaneous Classification. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2016)*, 35(4):110, 2016. 1, 2, 3, 4, 6