HACETTEPE UNIVERSITY

Department of Computer Engineering

FALL 2018

BBM 413 IMAGE PROCESSING PROJECT FINAL 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 Doğanay 21426878
b21590776@cs.hacettepe.edu.tr
b21427363@cs.hacettepe.edu.tr
b21426878@cs.hacettepe.edu.tr
Hacettepe University
Ankara, TURKEY

JANUARY 15th ,2019

Abstract

In this study, gravure images are discussed with CNN based models. Coloring gravure images is a problem that has not been studied before. For this reason, in this study, you will see common methods in unique domain. Then you are going to see the results are evaluated. We will represent our approach to solving this problem.

Image colorization is a very difficult problem because it requires different studies on different domains. It is very difficult to get successful results especially on the subjects that have not been studied. There are many parameters that cannot be taken into consideration in specific studies. In this study, we can not say that we have achieved successful results. But we will show how we get results in different approaches to gravure coloring problem.

1 Introduction

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.

We have adopted our domain into the (3) Iizuka's automatic image colorization technique. We have implemented gravure like images representation into the CNN model to train our model. We believe that this gravure colorization project is the one of the first attemps in the world. Our method colorizes gravures which contains human figures.

According to convolutional neural network approach mentioned in that paper(3), 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.
Some features of Lizuka(3):

- Two-stream architecture
- Regression loss
- Different database

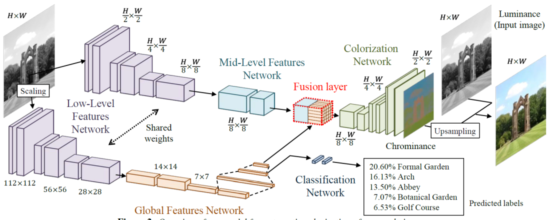
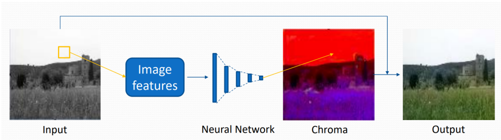


Figure 2: *Overview of our model for automatic colorization of grayscale images.*

2 Related work

- Automatic colorization with hand-crafted features(6). The work of Cheng combines three levels of features with increasing receptive field: the raw image patch, daisy features, and semantic features.



  - Uses existing multiple image features.

  - Computes chrominance via a shallow neural network.

  - Depends on the performance of semantic segmentation

  - Only handles simple outdoor scenes

- Learning Representations for Automatic Colorization from Larsson, Maire, Shakhnarovich: According to this article(4): Automatic colorizer takes each pixel, looks at a its surrounding and predicts a distribution over plausible colors.

- Zhang (7) similarly propose predicting color histograms to handle multi-modality. Some key differences include their usage of up-convolutional layers, deep supervision, and dense training. In comparison, we use a fully convolutional approach, with deep supervision implicit in the hypercolumn design, and memory-efficient training via spatially sparse samples.

- Tajbakhs (5) - Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?
In this study, it is investigated how pre-trained models and fine-tuned models give results in specific domains. According to the results of the study, fine-tuning process improves success. They use pre-trained models and also they train from scratch with medical images. They achieve more successful results than to train from scratch with medical images.

3 The Approach

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 the (3), they propose a fully automated data-driven approach for colorization of grayscale images. -as Zhang did-. 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.

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 gravure 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. Then we will give the input as gravure.

3.1 Dataset

At the first stage of the project, we decided to gather dataset from Google Images with specific keywords like related our domain.That dataset contains human figures and their dresses, scarfs, skirts and so on.Some of the keywords are like: ”kaftan, abiye, cubbe, etek, kurk, dress...” But we only get approximately 800 images. We have used that dataset to train our model. But it gives poor results and we see only reddish artifacts in the images. At behind of that we concurrently implemented our filters. So we throw through the rubbish that dataset then we have found a new dataset like our domain. Behind of that there are many irrelevant data.
But this data-set was very inadequate. As a result of long-term research and experiments we were able

to collect a better data-set. In order to use this(2) data-set, we signed the Dataset Release Agreement (1).The dataset was 28 GB however, this dataset was created for a different purpose. The part that can be used for our project is roughly 11 GB. We cleaned up unnecessary data for this dataset, but we do not able to use all this dataset, because although we used Google Colab servers to compile the project, a compilation process can take several hours even though we use a small number of data. When it takes few hours to see the results of the changes we made on the code. It takes a few days to the results if we use all the dataset and Google Colab does not allow for more than 12 hours of processes.
Our main idea is gravure colorization. So TC Kültür Bakanlığı has published several gravure achieves. We scanned one of them. So our testing dataset is hard copy scanned images. They are high resolution images with many details. We scanned 150 gravure examples that garments and portraits.

Explanation of converting steps for image above:

3.2 Filters

At this stage, it is aimed to make colored human figures look like colored gravure images by using various image processing methods. For this, various filters and filters should be applied.

First we downscaled the image using Gaussian pyramid and then repeatedly applied small bilateral filter instead of applying one large filter for edge preserving and to get a cartoon like image. Then again we upscaled the image to original size.

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 converted the image to grayscale and then we applied the median blur filter.

Next step is to identified the edges in the image and then added this to the previously modified images to get a sketch pen effect. For this first we are used adaptive threshold. Then we converted it back to rgb for merging with color image.

In last step, we get the final image combining color image and edge image with bitwise_and.

3.3 CNN

Knowing what kind of features will appear in each layer makes it easier to maintain the model. If we can analyze the results obtained after running the model, we can understand where the problem is and apply the necessary treatment. But if you want to try a model in a previously untried domain that you don't know how it works, you might need to do a lot of work. As Iizuka stated in his study, it was stated which features were extracted in different layers. We could analyze the problem in the gravure approach and maintain in the model.

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 proposed in work. So it actually works. That is drawback of the Zhang's work.

When we give the same gravure image to Iizuka and Zhang's model to comparison, while Iizuka classifies the sky and trees correctly with the human figure, and color the correct toning, Zhang fails. Because the sky in gravure images are also expressed by lines. This makes it difficult to detect the low-level features in the sky.

3.4 Style Transfer

The other method that we intend to apply is style transfer method. We transfer the style of the gravure images to the human images in our data colour. In this way, our data consists of colored engraving images. Then, the model is trained with the end-to-end learning architecture. For this purpose, we made some experiments shown in progress reports. Because of some limitations and drawbacks we couldn't keep it. But this will be a future work. Because it is an interesting topic.

4 Experimental Results

We have trained the model without applying filters. Then we test the model if is it worked. Then we tested the model on natural human images. Because of computational limitations, we trained our model with 2000-5000 images. and 3-10 epochs. As a result of our experiments, if we increase epoch, the color distributions becomes more certain.

As you see in the figure that below, it is more certain color edges than the gravures. Because gravure images don't have the continuous intensity values. Actually not, but they look like black and white pixels. Also background is not empty.

Until this stage we can easily see that with more training we get more accurate color edges and visible toning in human skin. Because our dataset has a lot girls images with short pants and strap tshirt. We can easily say better learning in more accurate patches.

The result of first stage shows us the human pictures became only one color likely red. More epochs and more train data makes more certain color patches. For example we have trained 5 epoch and 10 epochs training result with rest parameters same. Shown below figure left hand side is 5 epochs and background more red than the right hand side that 10 epoch training test result.

Because of that we have implemented gravure like filtering and color gravure filters. We have trained same model with same parameters. In addition we set the input and label of the training data with our filter outputs. Filter works in each iteration. In contrast to first method, it takes more and more time while training because of the more than 50 filters application each image. But as you can see the image above, there is no any color information. Directly natural image training more successful.

5 Conclusions

Our project was a project that has not been studied before.

The project was not as easy as we thought at the beginning. Problem domain was a unique domain. The resources and technical base were limited. We tried to produce solutions to the problem and took reasonable steps. We searched the literature and we chose to best project that can be made base to our project.

We got the appropriate dataset for our domain. We think we're using the right methods. The lack of technical competence limited our success rate. One of the most important points in the success of the project was fine-tuning. But we couldn't do fine tuning. The project, which we mainly refer to, did not share pre-trained model weights. That's why we had to do train from scratch. This did not yield very successful results. Zhang (7) share trained model weights. But since they use caffe to developed project, our technical infrastructure was not enough to deal with the complex structure of caffe.

Another point we are failed is that we cannot equalize the dimensions of the input and output images. We couldn't solve the problems like this because time limitations and hard semester. But as an idea, we think we are on the right way. When we solve the problems mentioned above, we think that the value of the gravure archives will increase greatly and will set light to historical research.

As a result, we have used the best deep learning algorithm that we can use. We have found a large dataset to suitable to our problem domain. We have integrated filters that we mentioned above, through the deep learning infrastructure. We tested our work with various parameters.

Although we didn't have a production level project, we tried to bring a solution to a different problem.

References

[1] Deepfashion dataset release agreement. http://mmlab.ie.cuhk.edu.hk/projects/DeepFashion/DeepFashionAgreement.pdf. 3

[2] Large-scale fashion (deepfashion) database. http://mmlab.ie.cuhk.edu.hk/projects/DeepFashion.html. 3

[3] 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):110:1–110:11, 2016. 1, 2

[4] G. Larsson, M. Maire, and G. Shakhnarovich. Learning representations for automatic colorization. In *European Conference on Computer Vision (ECCV)*, 2016. 2

[5] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang. Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE transactions on medical imaging*, 35(5):1299–1312, 2016. 2

[6] B. S. Zezhou Cheng, Qingxiong Yang. Deep colorization. https://arxiv.org/pdf/1605.00075.pdf. 1

[7] R. Zhang, P. Isola, and A. A. Efros. Colorful image colorization. In *ECCV*, 2016. 2, 7