Multiple and Diverse Image Colorization

Coloma Ballester, Lara Raad Cisa, Patricia Vitoria

Project description

Image colorization consists in recovering a color image from a grayscale one. This process attracts a lot of attention in the computer vision and image-editing communities with the goal of restoring old grayscale movies or pictures. While turning a color image into a grayscale one is only a matter of standard, the reverse operation is a strongly ill-posed problem as no information on which color has to be added is known. A recent strategy is to exploit semantic knowledge by using data-driven methods and priors learned by deep learning approaches.

FIGURE 1 – Left to right : grayscale input with three different colorization results [2].

Image colorization is in general dealt as a one-to-one mapping problem, where the purpose is to map a single color image to a given grayscale image. In fact, there is not just a unique solution but many plausible ones. Figure 1 displays an example. The main goal of this project is to tackle the real one-to-many mapping problem, i.e. mapping different plausible color images to a given grayscale image.

On the one hand, a few methods within the large literature of image colorization have been designed to generate diverse colorizations. The proposed approaches are based on different generative models such as variational auto-encoders [1], recursive neural networks [2, 3], vision transformers [4], among others. The results of these approaches show in general low diversity, semantic inconsistencies, or are “heavy” procedures as in the case of recursive networks.

On the other hand, the work of Zhang et al. [5] predicts per-pixel color distributions (marginals) conditioned to a grayscale input and, afterward, generates one deterministic color image by considering the annealed mean of these distributions. Besides, the network in [5] is simply trained as a classification problem (cross-entropy loss between the inferred distribution and a one modal ground truth distribution).

To meet the goal of this project, the idea is to leverage the simplicity of Zhang et al.'s approach in obtaining the marginal color distributions and develop a strategy to sample diverse plausible colorization results, being both spatially and semantically coherent, from these marginal distributions. On the one hand, we will focus on improving the lack of semantic description provided by the network in [5] as has already been explored in [7, 4] hoping to predict more realistic marginal color distributions. On the other hand, we will tackle the sampling problem which is not taken into account in [5] hence yielding a unique colorization for a given grayscale input image. In what follows, we list several questions that will be of interest for this project :

1. Are the marginal distributions predicted by Zhang et al. [5] improved when improving the semantic description of the network as in Pucci et al. [7] or Kumar el al. [4] ?
2. Can we generate multiple diverse colorizations in a 2-step approach where, first, the marginal distributions are predicted and, second, they are used in a clever method (e.g., variational method as in [6]) to sample a color value ? More precisely, in the spirit of [6], by optimizing a cost function composed of a reconstruction term on a color image sampled from the marginal distributions and a regularization term. Will a 2-step approach be enough to produce diverse color samples for a given grayscale image ?
3. Can we learn a sampling layer in the spirit of [8, 9] ? Following them, the diversity of colorization results could be obtained by learning stochastic filter generators. Moreover, the task of obtaining the stochastic filter generators is simplified in [8, 9] by replacing stochastic filter generation with basis generation.
4. Can transformers as in [4] really solve diversity and spatial consistency issues yielding plausible colorization results being both spatially and semantically coherent ?

Références

[1] A. Deshpande, J. Lu, M. -C. Yeh, M. Jin Chong, D. Forsyth. Learning diverse image colorization. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017

[2] S. Guadarrama, R. Dahl, D. Bieber, M. Norouzi, J. Shlens, K. Murphy. Pixcolor : Pixel recursive colorization. *arXiv preprint arXiv :1705.07208*, 2017

[3] A. Royer, A. Kolesnikov, C. H. Lampert. Probabilistic image colorization. *arXiv preprint arXiv :1705.04258*, 2017

[4] M. Kumar, D. Weissenborn, N. Kalchbrenner. Colorization transformer. *arXiv preprint arXiv :2102.04432*, 2021

[5] R. Zhang, P. Isola, A. A. Efros. Colorful image colorization. *European Conference on Computer Vision*, 2016

[6] T. Mouzon, F. Pierre, M.-O. Berger. Joint cnn and variational model for fully-automatic image colorization. *International Conference on Scale Space and Variational Methods in Computer Vision*, 2019

[7] R. Pucci, C. Micheloni, N. Martinel. Collaborative Image and Object Level Features for Image Colourisation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021

[8] Z. Wang, X. Cheng, G. Sapiro, Q. Qiu. Stochastic conditional generative networks with basis decomposition. *arXiv preprint arXiv :1909.11286*, 2019

[9] Q. Qiu, X. Cheng, G. Sapiro. DCFNet : Deep neural network with decomposed convolutional filters. *International Conference on Machine Learning*, 2018