Pattern Recognition Project Proposal

Jose Acitores Cortina 7007337
Martijn Dekker 6013368
Daniele Di Grandi 7035616
Alex Vermeer 5717329
Matthijs Wolters 5983592

December 2020

1 Goal

The steadily increasing performance of state of the art neural networks raises new questions regarding the limits of these neural networks. Although humans are far from perfect, in more abstract situations humans can still outperform advanced neural networks. One task in which humans used to outperform these intelligent systems, but no longer do, is the classification of images. A similar but more abstract task is the colourisation of images. In this study we will look into the development of a neural network that is able to colourise images in a manner that will result in images true to the original.

With this research we hope to find a way of correctly colourising images that were turned monochrome. If successful enough, this could be used to restore old photographs and turn old monochrome photographs and films into coloured ones. Currently this is often done by human specialists. This is, however, quite costly which means that it is not an available option for everyone. Therefore it would be desirable to develop a system that can perform on par, or even better, with the human specialists.

2 Dataset

The dataset to be used in the first part of this project will be the CIFAR-10 dataset, which contains 60000 images in 10 different classes [5]. The classes range from animals to machines and are generally used to classify objects, as this is a labeled dataset. However, we believe that this dataset can also be used as a starting point for training our colourization network. The images consist of 32 by 32 pixels in the RGB colourspace. These images will be transformed from RGB to gray scale using scripts from scikit-learn [9]. In this way we have the ground truth to compare the work of the neural network to how it is supposed to work. Once the network performs at a sufficient level with this dataset, we are planning on scaling up the size of the images. The larger images will probably come from the ImageNet dataset, as it is very large and has multiple different image size sets available [2].

3 Related Work

Early work on automatic image colourisation was focused on old physical materials that have degraded over time, such as old bleached film reels or oxidised paintings. These materials usually still contain colour, albeit not the original ones. These early works primarily used mathematical models for the colour restoration.

Pappas and Pitas [8] use one such mathematical model for digital painting restoration. Their method requires a human to restore sections of a painting manually, which the model uses as a reference for the colour palette of the entire painting. Their model would then use this reference using different distance and regression functions to restore the remaining sections of the painting.

Rizzi et al. [11] used a two phase algorithm for restoring bleached film reels, be it an originally coloured film, or a black-white film. Their algorithm works on the assumption that the film bleaches about equally for every pixel, but not necessarily for every colour channel. It thus readjusts the pixels related to the image as a whole and then uses the generated image to estimate what white and gray

are in the image and finally adjust the tone using this information.
More recently, exploration of complete automatic image colourisation utilizing neural network models has attracted the attention of many researchers. These works aim to create a model that can colour images and film correctly with as little manual work as possible.
Zhang et al. [12] attacked this problem proposing a feed-forward pass in a Convolutional Neural Network trained on over a million colour images and evaluating the performance by a "colourisation Turing test" while using, as a loss function, a classification loss with rebalanced rare classes. Larsson et al. [6] and Iizuka et al. [4] have developed similar systems, using large-scale data and Convolutional Neural Networks, though they use different loss functions: an un-rebalanced classification loss and a regression loss, respectively. Moreover, Larsson et al. used as CNN's architecture, an hypercolumns on a VGG network, Iizuka et al. used a two-stream architecture in which they fuse global and local features and Zhang et al. used a single-stream, VGG-styled network with added depth and dilated convolutions.
Nazeri et al. [7] tackled this problem using a different approach: fully generalize the colourisation procedure using a conditional Deep Convolutional Generative Adversarial Network (DCGAN) and they compare the results between their model and traditional deep neural networks models. It is also worth mentioning the implementation of Richart et al. [10], where they proposed a Multi Layer Perceptron Neural Network to classify the colour of a gray level pixel without the need of heavy image processing algorithms, achieving some good, although not optimal, results.

4 Model

As we are using images we will be using a Convolutional Neural Network. As the output of the network is also an image, we will be using an encoder-decoder network architecture. The implementation of our model will be dependent on the model design to be done in the first week of January.

5 Evaluation of the model

There is a high complexity when evaluating how well a model colourises an image due to the subjectivity of the issue. For a person, it can be really easy to evaluate whether the colours of an image fit or not, but for a model it all varies with the values of the image pixels. To overcome this issue in the best way possible, we use 2 different approaches for the evaluation of this model similar to the way done by Larsson et al.[6]. For that we create a dataset from colourised images turned into black and white in order to have a reference for the desired output of the model. After that we will evaluate the performance of the model with two markers, the RMSE of the difference in the colour space denoted by $\alpha\beta$ as explained by Deshpande et al. [3], and peak signal-to-noise ratio (PSNR) in RGB calculated per image using the arithmetic mean as used by Larsson et al. [6] in contrast to the geometric mean used by Cheng et al. [1] due to the high sensitivity of the latter.

6 Planning

Table 1 shows our preliminary planning for the project. The planning leaves some space for adjustment and team members will likely also work on other tasks than they are assigned to.

Period	Task	Assignee
Jan 4-10	Model design Data collection Data preprocessing	Martijn, Daniele, Alex Matthijs Jose, Alex
Jan 11-17	Model implementation Model tuning	Jose, Alex, Matthijs Daniele, Martijn
Jan 18-24	Model evaluation and refinement	Everyone
Jan 25-27	Document and presentation preparation	Everyone
Jan 27-29	Finish up document	Everyone

Table 1: Preliminary planning

References

[1] Z. Cheng, Q. Yangg, and B. Sheng. Deep colorization. *ICCV*, 2015.

[2] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*, 2009.

[3] A. Deshpande, J. Rock, and D. Forsyth. Learning large-scale automatic image colorization. *ICCV*, 2015.

[4] 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.

[5] A. Krizhevsky. *Learning multiple layers from tiny images.* 2009.

[6] G. Larsson, M. Maire, and G. Shakhnarovich. *Learning Representations for Automatic Colorization.* Springer International Publishing, Cham, 2016.

[7] K. Nazeri, E. Ng, and M. Ebrahimi. *Image Colorization Using Generative Adversarial Networks.* Springer International Publishing, Cham, 2018.

[8] M. Pappas and I. Pitas. Digital color restoration of old paintings. *IEEE Transactions on Image Processing*, 9(2):291–294, Feb 2000.

[9] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[10] M. Richart, J. Visca, and J. Baliosian. Image colorization with neural networks. pages 55–60, 2017.

[11] A. Rizzi, C. Gatta, C. Slanzi, G. Ciocca, and R. Schettini. Unsupervised color film restoration using adaptive color equalization. In S. Bres and R. Laurini, editors, *Visual Information and Information Systems*, pages 1–12, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[12] R. Zhang, P. Isola, and A. A. Efros. *Colorful Image Colorization.* Springer International Publishing, Cham, 2016.