IMAGE COLOURIZATION VIA CONVOLUTIONAL NEURAL NETWORKS AND DEEP LEARNING

Youssef Fikry
Student# 1005678901
youssef.fikry@mail.utoronto.ca

Harkirpa Kaur
Student# 1011242479
harkirpa.kaur@mail.utoronto.ca

Peter Leong
Student# 1005678901
peter.leong.utoronto.ca

Thulasi Thavarajah
Student# 10115358424
t.thavarajah@mail.utoronto.ca

ABSTRACT

This template should be used for all your project related reports in APS360 course. – Write an abstract for your project here. Please review the **First Course Tutorial** for a quick start —–Total Pages: 2

1 INTRODUCTION

While the colour photography processes first emerged in the 1890s, colour photography did not become widely accessible until the 1970s Science & Museum (2020). As a result, most historic photographs are black and white and lack the visual richness that modern viewers are accustomed to. In addition, individuals who have their cataracts removed as a part of vision restoration processes have shown to struggle with identifying objects in black and white images Vogelsang et al. (2024), making most historic photographs inaccessible to them. This project aims to use deep learning to automatically colorize black and white images, with the goal of restoring visual information and making historical imagery more accessible for all audiences. Traditional, non deep-learning based colorization methods often produce desaturated results and rely heavily on human guidance Cheng et al. (2016), making them non-scalable. Conversely, deep networks such as CNNs effectively capture spatial and semantic features and produce realistic colourized images without user interaction Zhang et al. (2016), making deep learning an ideal approach for image colorization.

1.1 BACKGROUND & RELATED WORK

The challenge of image colourization has been approached in a wide variety of ways. Even within solutions involving deep learning techniques, there numerous unique design choices. One grouping method Zěger et al. (2021) results in five categories: simple colourization neural networks, user-guided colourization neural networks, diverse colourization neural networks, multi-path colourization neural networks, and examplar-based neural networks.

Simple colourization neural networks use feedforward CNNs to map grayscale inputs to colour outputs. One of the foremost solutions proposed by Zhang et al. Zhang et al. (2016) used a fully convolutional network to predict the a and b channels of the CIELAB colour space from grayscale images. Their architecture is composed of several convulational layers, each followed by a ReLU activation function and a batch normailization layer.

User-guided colourization neural networks use user input to guide the colourization process. One such solution Zhang et al. (2017) uses a fully convolutional network to predict the a and b channels of the CIELAB colour space from grayscale images, but also takes user input in the form of user-provided colour scribbles. The network is trained to minimize the difference between the predicted and user-provided colours, allowing it to learn to colourize images in a way that is consistent with the user's input.

Diverse colourization networks produce multiple colourization outputs for a given grayscale input. One such solution Vitoria et al. (2020) uses a generative adversarial network (GAN) to produce multiple colourization outputs for a given grayscale input. The GAN is trained to minimize the difference between the predicted and ground truth colours, allowing it to learn to produce diverse colourization outputs.

Multi-path colourization neural networks differentiate features at different scales. Iizuka et al. Iizuka et al. (2016) proposed a multi-path colourization neural network that uses multiple convolutional layers to extract features at different scales. The network is trained to minimize the difference between the predicted and ground truth colours, allowing it to learn to produce colourization outputs that are consistent with the features at different scales.

Exemplar-based neural networks use a set of exemplar images to guide the colourization process. In Su et al Su et al. (2020), example images are used to trasnfer the colour to the target image. Each instance is ouput to two different colourization networks which fuse to yield the final result. This group of solutions is easier to implement, as learning to colourize instances is significantly easier than learning to colourize an entire image.

2 Methodology

2.1 Data Processing

2.2 Architecture

2.3 Baseline Model

3 Ethical Considerations

The dataset being used is public, so there are no copyright or consent issues. However, the dataset may contain racial or demographic imbalances, which could cause the model to generalize poorly or be biased towards specific skin tones. This may result in racially inaccurate or culturally insensitive outputs. A similar behaviour may be observed with animals, where a lack of diversity in breeds or fur colours in the dataset can result in misleading results. If the outputs produced by the model are used in educational contexts or in breed identification, they can contribute to misinformation. Furthermore, since the model results are plausible but cannot be verified, there is a risk that users may overtrust the outputs in sensitive contexts.

4 Project Plan

4.1 Team Communication & Coordination

5 Risk Register

References

Zezhou Cheng, Qingxiong Yang, and Bin Sheng. Deep colorization, 2016. URL `https://arxiv.org/abs/1605.00075`.

Satoshi Iizuka, Edgar Simo-Serra, and Hiroshi 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 (TOG)*, 35(4):110:1–110:11, July 2016. ISSN 0730-0301. doi: 10.1145/2897824.2925974. URL `https://doi.org/10.1145/2897824.2925974`.

National Science and Media Museum. A short history of colour photography, Jul 2020. URL `https://www.scienceandmediamuseum.org.uk/objects-and-stories/history-colour-photography`.

Jheng-Wei Su, Hung-Kuo Chu, and Jia-Bin Huang. Instance-aware image colorization, 2020. URL `https://arxiv.org/abs/2005.10825`.

Patricia Vitoria, Lara Raad, and Coloma Ballester. ChromaGAN: Adversarial Picture Colorization with Semantic Class Distribution. In *Proceedings of the 2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 2434–2443, Los Alamitos, CA, USA, March 2020. IEEE Computer Society. doi: 10.1109/WACV45572.2020.9093389. URL `https://doi.ieeecomputersociety.org/10.1109/WACV45572.2020.9093389`.

Marin Vogelsang, Lukas Vogelsang, Priti Gupta, Tapan K. Gandhi, Pragya Shah, Piyush Swami, Sharon Gilad-Gutnick, Shlomit Ben-Ami, Sidney Diamond, Suma Ganesh, and Pawan Sinha. Impact of early visual experience on later usage of color cues. *Science*, 384(6698):907–912, 2024. doi: 10.1126/science.adk9587. URL `https://www.science.org/doi/abs/10.1126/science.adk9587`.

Ivana Zěger, Sonja Grgic, Josip Vuković, and Gordan Šišul. Grayscale image colorization methods: Overview and evaluation. *IEEE Access*, 9:113326–113346, 2021. doi: 10.1109/ACCESS.2021.3104515.

Richard Zhang, Phillip Isola, and Alexei A Efros. Colorful image colorization. *European conference on computer vision*, 2016.

Richard Zhang, Jun-Yan Zhu, Phillip Isola, Xinyang Geng, Angela S. Lin, Tianhe Yu, and Alexei A. Efros. Real-time user-guided image colorization with learned deep priors, 2017. URL `https://arxiv.org/abs/1705.02999`.