

# 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 ?. 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 ?, 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 ?, making them non-scalable. Conversely, deep networks such as CNNs effectively capture spatial and semantic features and produce realistic colourized images without user interaction ?, 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 ? results in five categories: simple colourization neural networks, user-guided colourization neural networks, diverse colourization neural networks, multi-path colourization neural networks, and exemplar-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. ? 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 convolutional layers, each followed by a ReLU activation function and a batch normalization layer.

User-guided colourization neural networks use user input to guide the colourization process. One such solution ? 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 [?] 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. [?] 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 [?], example images are used to transfer the colour to the target image. Each instance is output 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