Image Colorization for Vintage Portraits

Project Proposal

Team Members:

Ahsan Jalil Mirza - 7028849

Kai Wittenmayer - 7005859

Patryk Dąbkowski - 7028711

Motivation & Tasks

In society photos play a vital role as a visual memory. They preserve a moment in time to be relived by generations to come, which gives them both historical and sentimental value. For this reason it is important to preserve and restore vintage (or printed) images. Traditionally colorization and restoration was done by hand, which is very time consuming and requires a special artistic skill-set. We propose to develop a convolutional neural network based model that performs two main tasks:

1. Image *colorization* of vintage portraits from grayscale to full color
2. Removal of salt and pepper noise and scratches as part of *restoration*

The primary benefit of this technique over the conventional methods is that restoration and colorization of images can be done at scale with a very consistent level of quality.

Goals & Challenges:

There are three primary challenges that need to be addressed in this project. Like any other machine learning based project the dataset that the model is trained on plays a vital role in the end product. In our case there is no openly accessible dataset of vintage portrait photographs and their corresponding full color images available to train this type of model. Therefore, we will need to work on a pre-processing mechanism that can simulate realistic vintage photography from available full-color portrait images and form a suitable dataset for supervised learning.

The next major challenge in this project is that in contrast to mainstream applications such as object detection, segmentation, the literature available on this type of application of deep learning is scarce.

Lastly, in the improvement phase of the project we plan on implementing a GAN based model to achieve a higher level of image quality. The challenge with this task is that generally GANs are notoriously unstable to train. We will need to perform a separate literature review to understand and implement techniques that help stabilize the training process of GANs.

Methods

Our project will involve the following five major phases:

Literature Review:

In this phase we will work on exploring both conventional image processing and deep learning based image colorization and restoration techniques present in the literature. This is important in order to understand the state-of-the-art for this and related applications.

Dataset Pre-processing:

For the dataset preprocessing we will use low-level image processing to implement a filter that simulates the attributes of actual vintage images (greyscale, salt & pepper noise, scratches, etc) and generates a supervised dataset for training. Python libraries like Scikit-Image, NumPy, SciPy and openCV will be used for this task.

Preliminary Model Implementation:

Based on the literature review done earlier, we will implement a DL-based model along with its training (using standard regression loss functions) and testing scripts. We will use PyTorch Deep Learning framework for this step.

Evaluation:

In the evaluation phase we will use both well-defined numerical metrics (like PSNR, SSIM etc) and visual fidelity (for Just-Noticeable Difference evaluation) in order to determine the model performance. We will also discuss any undesirable artifacts produced by the model due to color hallucination etc.

Improvement:

Lastly, we will work on implementation of a GAN configuration for model improvement. The exact architecture for the generator and discriminator is still to be decided.

Dataset

For training our model we have decided to use the UTKFace dataset which consists of over 20k portrait images captured both in the wild and controlled environments. The dataset is very evenly distributed in terms of age, gender and ethnicity. As mentioned earlier we will pre-process the dataset to produce vintage-like images for model training.

Evaluation

The performance evaluation of the model will be done in the following two parts:

Overall Performance Evaluation:

We will use the following metrics for evaluating the overall performance of the model:

- Peak-Signal to Noise Ratio (or MSE)
- Structural Similarity Index
- Perceptual Loss

Visual Fidelity Evaluation:

We will also perform a visual analysis to evaluate Just-Noticeable Difference (JND) in generated and ground-truth images. This is crucial to determine if any unwanted artifacts are produced by the model.

References

1. https://susanqq.github.io/UTKFace/
2. https://arxiv.org/pdf/2008.10774.pdf