FINAL PROJECT PROPOSAL

By
Group 1

Naiska Buyandalai
Sai Rachana Kandikattu
Snehitha Tadapaneni

Deep Learning: DATS 6303

Under the guidance of,

Prof. Amir Jafari

Problem Selection and Motivation

Problem Selected:

“ Colorization of grayscale images using Deep Neural Networks “

Motivation:

Restoring color to grayscale images is a long standing challenge in computer vision and graphics. It is an ill posed problem since many color combinations can correspond to the same grayscale intensity. Many Traditional methods required user input or produced desaturated colors. We aim to build a fully automatic and vibrant image colorization system using deep learning, trained on over a million color images, that can hallucinate plausible colors that are perceptually realistic to humans. Beyond aesthetics, this task also serves as a self supervised learning problem, helping models learn semantic and visual features useful for other downstream tasks like object recognition and segmentation.

Dataset Description

For our project, we will utilize the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) subset, which contains over 1.3 million high-resolution color images across 1,000 diverse object categories. ImageNet is one of the most comprehensive and widely used datasets in computer vision and has been extensively applied in colorization research, including the baseline paper Colorful Image Colorization (Zhang et al., 2016).

The dataset’s scale and diversity make it ideal for training deep learning models such as convolutional neural networks (CNNs) and generative adversarial networks (GANs), as it encompasses a broad range of semantics, textures, and color distributions that are essential for learning context-aware and globally consistent colorization.

Deep Network Architecture

We will use a customized network. The network will use standard blocks consisting of 2–3 repeated Conv + ReLU layers followed by Batch Normalization.
In the future, we may further change the architecture, have more customized networks, use pre-trained networks based on the training function, loss, or dataset characteristics.

Future Extensions:
As a potential extension, we may experiment with a conditional GAN (cGAN) architecture, where a generator (similar to the baseline CNN) predicts color images from grayscale inputs, and a discriminator learns to distinguish real vs. generated images. This could improve realism in colorization and allow the network to model more complex distributions.

Implementation Framework

We will use PyTorch because we find it is much more flexible compared to tensorflow. Its interface is closer to the mathematical interpretation of the network which makes building and testing layers easier.
It is also widely used in research, making it easier to follow existing papers to implement various methodologies.

Reference Materials and Background

Primary Reference:

- Zhang, R., Isola, P., & Efros, A. A. (2016, September). Colorful image colorization. In the European conference on computer vision (pp. 649-666). Cham: Springer International Publishing.

Supporting Literature:

- Nazeri, K., Ng, E., & Ebrahimi, M. (2018, June). Image colorization using generative adversarial networks. In International conference on articulated motion and deformable objects (pp. 85-94). Cham: Springer International Publishing.

These references collectively provide the theoretical foundation for CNN-based image colorization and adversarial image generation, the latter of which we aim to adopt as an extension to the former to improve the performance.

Evaluation Metrics

We will evaluate the model using metrics adopted in both Zhang et al. (2016) and Nazeri et al. (2018) to assess color accuracy and structural consistency.

- AUC: Measures the accuracy of the model's predicted color distributions, following Zhang et al. (2016).
- VGG Top-1 Classification Accuracy: Evaluates how well colorized images preserve semantic cues by testing a pretrained classifier on them, following Zhang et al. (2016).
- Color Accuracy: Reports the percentage of pixels whose color prediction error falls within a small tolerance range, as used by Nazeri et al. (2018).

Project Timeline

Week 1:
Conduct a literature review on the reference materials and set up an environment to preprocess the ImageNet dataset. Implement a baseline CNN generator.

Week 2:
Train baseline colorization model using weighted cross entropy loss. Reproduce the reference paper results.

Week 3:
Implement and integrate the discriminator. Begin adversarial training with a combined generator–discriminator framework and tune GAN hyperparameters.

Week 4:
Evaluate results and analyze GAN improvements over the baseline and prepare final report and presentation.