**Image Restoration with PaddlePaddle**

Table of Content

[1. Introduction 2](#_Toc102145809)

[2. Background 2](#_Toc102145810)

[3. Technical Difficulties Related 2](#_Toc102145811)

[4. Restoration Strategy 3](#_Toc102145812)

[4.1 Photo Restoration 3](#_Toc102145813)

[4.2 historical photo restoration 3](#_Toc102145814)

[5. Model Design 3](#_Toc102145815)

[5.1 Layer and Data Structure 3](#_Toc102145816)

[5.11 Lab data structure 3](#_Toc102145817)

[5.12 Input Layer 4](#_Toc102145818)

[5.13 Output Layer 4](#_Toc102145819)

[5.2 Loss Function 5](#_Toc102145820)

[5.3 Backbone Design 5](#_Toc102145821)

[5.4 Optimizer Design 6](#_Toc102145822)

[5.5 Dataset Summary - ImageNet 6](#_Toc102145823)

[5.6 Deep Learning Platform Summary - Paddle 6](#_Toc102145824)

[6. Code Link 7](#_Toc102145825)

# Introduction

This is a project created to have the function to colorize black and white picture’s color or restore faded-color pictures. Using Self-Attention GAN as the principal solution for the image restoration.

# Background

As a specific image record of a period, old photos and historical images carry valuable

memories, which have irreplaceable significance. However, in the past photography

technology wasn’t as advanced as in the present, and due to improper storage mechanisms, old photos faded, creased wearied, and torn. Hence, it is necessary to recover old photos. Photo restoration is meant to use edge detection and techniques to recover photos in which areas are broken or missing —— the traditional way of photo recovering, which does not perform well if the repair area contains complex, non-repetitive structures.

In recent years deep learning continues to make advances in computer vision and image processing, using large numbers of images to train networks, allowing trained models to have a large amount of prior knowledge of images, giving another way to restore photos.

# Technical Difficulties Related

In 2006, Deepak Pathak used a simple codec as the master network for image restoration. The encoding part yields a low-dimensional feature map, and the decoding part uses the extracted features to repair the defective region and restore the whole image and uses discriminator supervision to make the repaired region as close as possible to the original image, but due to the information bottleneck in the channel layer, the repaired areas of the output image often contain visual artifacts and exhibits blurred features with obvious repair traces.

Currently, the problem encountered with old image restoration are the follows:

1. There are no models that can realistically represent the degradation of old photographic artifacts.
2. Repairing complex degraded photos often requires different strategies, such as unstructured defects tend to be homogeneous, faded defects need to be repaired using information from adjacent pixel points, and structural defects like scratches need to be repaired using global image information.
3. For portrait photo restoration, viewers tend to pay more attention to the character’s face and consider portraits to account for a large portion of old photos, it is needed to design networks that are specific to face restoration.

# Restoration Strategy

## 4.1 Photo Restoration

Using Self-Attention GAN as the principal solution while also using some other traditional image processing techniques to cover the small holes and missing pieces of the image.

## historical photo restoration

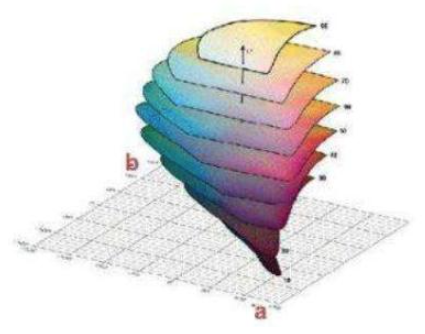
Based on *Colorful Image Colorization* and *Learning Representations for Automatic Colorization,* Low-level and semantic representations are utilized in the coloring process. Since many scene elements are currently generated naturally based on multi-peaked color distributions, our model is trained to predict pixel color histograms. This intermediate output can be used to automatically generate color images or to perform further manipulations before the images are formed.

# Model Design

## 5.1 Layer and Data Structure

### 5.11 Lab data structure

Lab mode is formed by three channels, but not R, G, and B channels. One of its channels is lightness as L, the other two are color channels A and B. A channel includes colors from dark green (low lightness level) to gray (mid lightness level) and then to light pinky red (high lightness level). B channel includes dark blue (low lightness level) to gray (mid lightness level) then to yellow (high lightness level). This featured color could mix and create brighter colors.

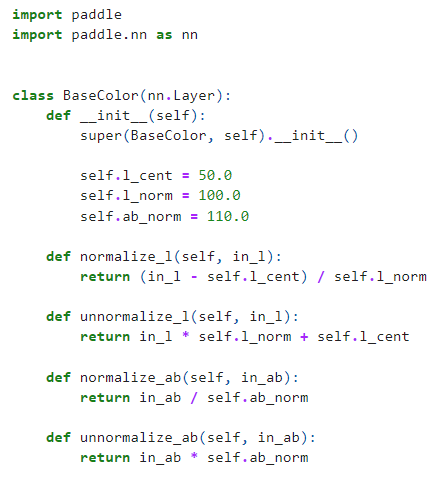


### 5.12 Input Layer

Since historic photos are mostly in black and white, the index converts the photo to greyscale upon input. This makes the photo a single channel instead of three. We are then able to use

[C, H, W] (here can be shown by [1, H, W]) to represent photos dimension, height, weight respectively.

To separate greyscale and color information from RGB photos, we can convert RGB to LAB as needed, thus the lightness in LAB represents outline and textures, A and B being the color channel. By inputting the L channel, we can make the model predict its A and B channels.



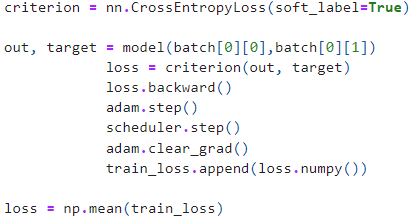
### 5.13 Output Layer

In case insufficient color is generated, A and B color channels (0~255) are compressed, so inputting 256\*256 pixels forms 313 groups of high-frequency color classification. Meanwhile

downgrading [B, H, W, C] format to [H, W, 313]. 313 represents 313 color combination probabilities.

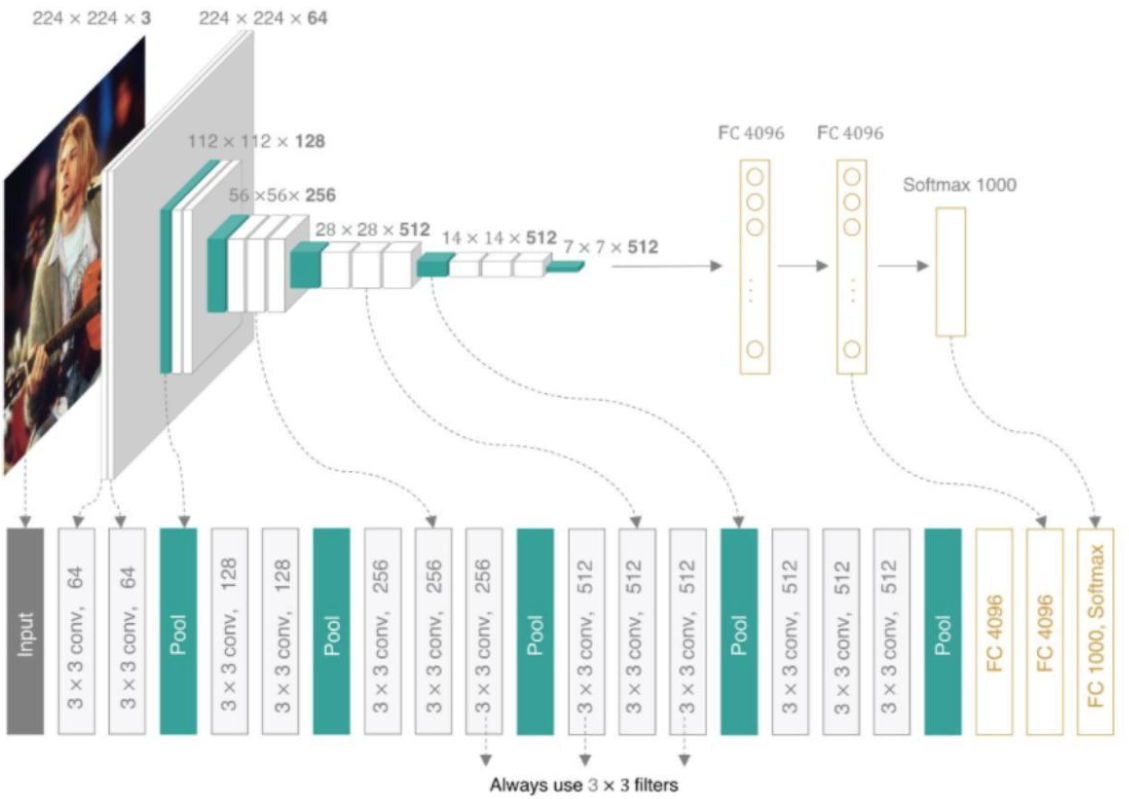
## 5.2 Loss Function

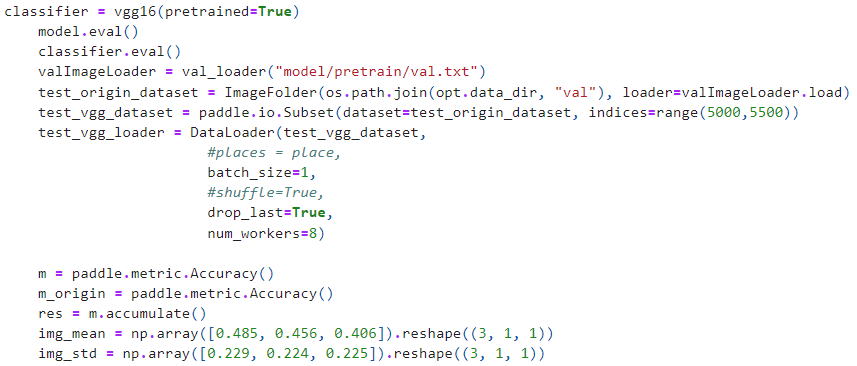
For calculating loss from A and B color channels



## 5.3 Backbone Design

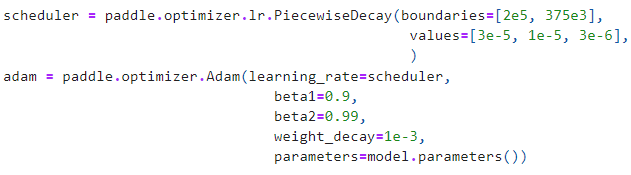
Use VGG as the classification backbone for the model, utilizing VGG 16 classifier to evaluate the semantic interpretability of coloring results.





## 5.4 Optimizer Design

Using Adam to optimize model parameters, and set the learning rate to decay down in steps of the default parameter.



## 5.5 Dataset Summary - ImageNet

ImageNet-mini is used in the project. ImageNet is an image database organized according to the WordNet hierarchy (currently only the nouns), in which each node of the hierarchy is depicted by hundreds and thousands of images.

## 5.6 Deep Learning Platform Summary - Paddle

PaddlePaddle, as the first independent R&D deep learning platform in China, has been officially open-sourced to professional communities since 2016. It is an industrial platform with advanced technologies and rich features that cover core deep learning frameworks, basic model libraries, end-to-end development kits, tools & components as well as service platforms. PaddlePaddle originated from industrial practices with dedication and commitment to industrialization. It has been widely adopted by a wide range of sectors including manufacturing, agriculture, enterprise service, and so on while serving more than 4 million developers, and 157,000 companies and generating 476,000 models.

# Code Link

https://github.com/harry334421/6364\_MechineLearningFinal