Find Color Project

Jaeyeon Kim, Changheon Kim, Saebyeol Shin, Mingeun Jang, Dongjin Jung

Capstone Design Project, Sungkyunkwan University

Abstract. Colorizing old grayscale images preserves the historical meaning and provides a reality to the photos. By using AI technology, colorization can be performed appropriate for the era and situation. However, in the process of performing colorization, there can be multiple correct answers and it shows poor performance with multiple objects. In this project, by using methodology of context-aware adaptive network and local-global hint network, we will develop user-interactive and adaptive colorization model and website.

Keywords: Instance-aware · User Interactive · Colorization

1 Introduction

Most of 1900's images are grayscale, and the need for colorization technology for the old records and contents is increasing these days. As the technology applying AI technology to media data such as images and videos develops, it is very meaningful in terms of social and economic value to restore data through colorization technology to grayscale images containing historical information. Fig. 1 is an example of colorization of old grayscale image.

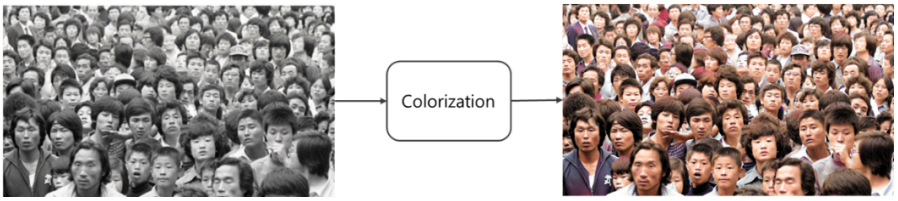

Fig. 1. Colorization of old grayscale image

Our team's goals are as follows:

(1) Development of AI restoration technology to color digital grayscale images containing Korean history for old records and contents
(2) Development of appropriate colorization models suitable for situations such as war, liberation, and demonstrations in images.
(3) Developing the website using above technology and model.

2 Motivation and Objective

Colorization of grayscale images containing historical information produces significant value in social and economic aspects. However, the current coloration technology for old records or content selects colors according to the judgment of the data learned by the model.

Suppose that a building contained in an old record remains to this day and can be observed. The column color of the actual building is red, and the image has been restored to dark brown by the AI model's own judgment. So, is this restoration socially and economically valuable?

If the perspective of the Colorization technology is a judgment criterion, its value is determined according to the quality of the restored image. However, from the perspective of restoring historical records, the value is low because the model made a wrong judgment in the restoration process.

To solve this problem, we aim to develop a model that uses user-interactive colorization methodology to provide a choice for color, and to set the color of the target in more detail. As a result of this project, we will produce services that anyone can use.

3 Background and Related Work

3.1 Background

Whether as a way of colorizing old and historical images which is socially and economically valuable, or expressing artistic creativity, people continue to be fascinated by colorization. So, many researchers have been creating colorization models and distribute them to free or paid websites. However, there still exists some problems that most of them have poor performance and prevent users from intervention. In addition, colorization is too slow in many websites. For this reason, we planned to develop a model that is faster and has better performance than other existing models, and distribute it to website for free.

3.2 Related Work

User Interactive Colorization. When colorizing grayscale images, there can be multiple correct answers not a single correct answer. User interactive colorization allows the user to progressively colorize the images by specifying colors at different locations in the grayscale image. Recently, Zhang *et al.* [1] proposed a learning-based method by extending an existing unconditional colorization model to produce color images given a grayscale image and user hints with local hints network and global hints network.

Instance-aware Image Colorization. Traditional colorization methods are dependent on user intervention to provide guidance such as color scribbles and reference images. However, there is a problem of poor performance when there are several objects because deep neural networks apply learning and colorization to the entire image. Instance-aware image colorization [2] is used which considers instance features and full-image features. Learning to color an instance is simpler than learning to color the entire image because it disregards complex background clutter. Every instance is forwarded to two different colorization neural networks which are instance colorization network and full-image colorization network. Fusion of the features is performed for the output. By using this method, better feature map is obtained leading to better colorization results.

Dynamic Region-Aware Convolution. Dynamic Region-Aware Convolution (DRConv) [3] introduces a filter generator module to learn specialized filters for region features at different spatial location. It can automatically assign filters to corresponding spatial-dimension region with learnable instructor. As a consequence, DRConv has powerful semantic representation ability and maintains translation-invariance property.

Image Colorization with Transformers. Unlike the widely-used convolution-based approach for image synthesis such as CNN, Colorization Transformer [4] proposes an autoregressive model for unconditional colorization which uses the transformer decoder architecture in order to generate diverse colorization results. However, the excessively slow inference speed of autogregressive models prevents user from application of intervention. Lee, S *et al.* [5] proposed a multi-head attention of the transfer encoder which reduces the inference time of model compared to autoregressive colorization.

4 Problem Statement Proposed Solution

4.1 Datasets for Effective Training

MHMD(Modern Historical Movies Dataset) MHMD is a dataset of 1,353,166 images which meets the requirements of different clothing types, eras, and nationalities preprocessed from 147 old movies and TV series. It can solve the problem that existing datasets' lack information about historical grayscale images or people's clothing colors. Thus, MHMD is a data set suitable for restoring old records without distortion.

YouTube-8M Dataset YouTube containes a lot of information on various situations, eras, nationalities, etc. It is suitable to effectively learn models according to the eras, situations, and nationality of people. It is also useful for the case of reinterpreting the images of the past in a modern way.

4.2 Our Method

In this project, we'll use the context-aware adaptive network as shown in Fig. 2. The instance detection information is not required, and the context confusion problem can be solved efficiently by performing instance adaptive convolution. Also, we'll use the local and global hint network [1] as shown in Fig. 3. By using this, the ill-posed problem can be solved by allowing the user to choose the desired color. After the context-aware adaptive network is trained, we will apply the methodology of local and global hint network to learn to change detailed colors and make user-interactive model.

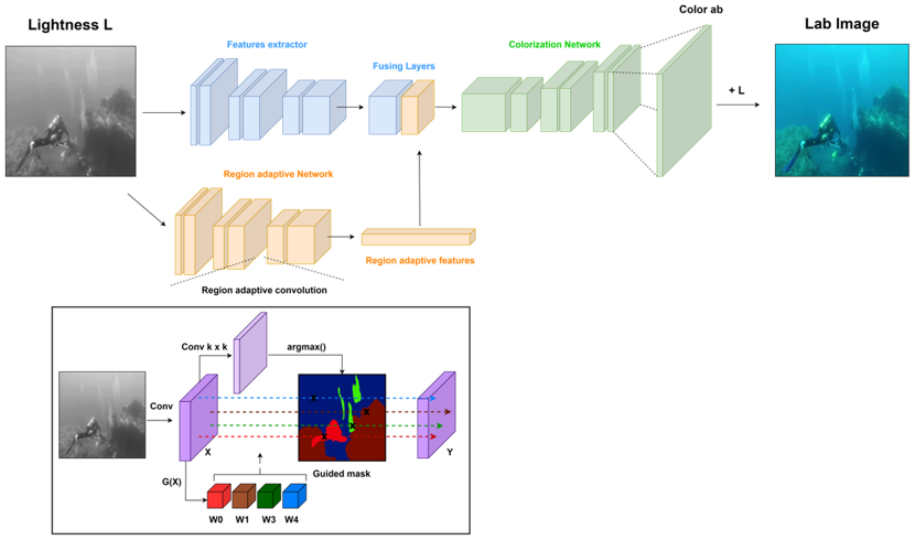

Fig. 2. Our Method: Context-aware Adaptive Network

5 Planning in Detail

5.1 Subtasks

Our project contains two parts largely. One is modeling part and the other is web UI/UX. And modeling part can be divided into three parts including data preprocessing, baseline model implementation and backbone model.

In data preprocessing part, we will collect datasets(MHMD, Youtube-8M, etc) and preprocess data. By gathering proper datasets, we can achieve valid datasets for the model. Also, to achieve better learning ability of model, we will add noise and augmentation with considering practical environments.

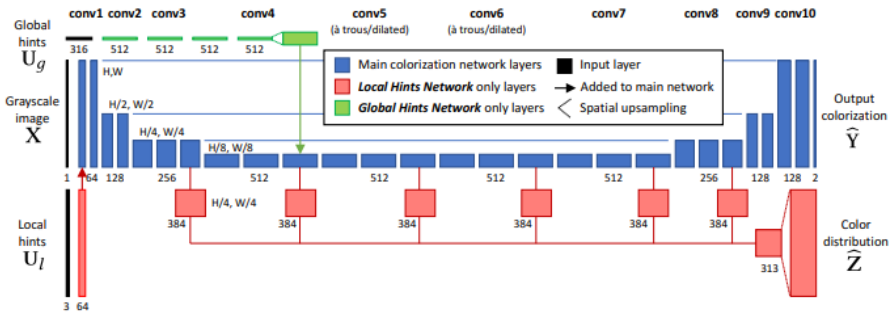

Fig. 3. Our Method: Local and Global Hint Network

After preprocessing part, we will conduct baseline model implementation. A baseline model is essentially a simple model that acts as a reference in a machine learning project. Its main function is to contextualize the results of trained models. We also do performance verification of our baseline model. By designing the baseline model properly, we can lay the groundwork firmly.

As the works for the baseline model are done, we'll start the jobs for backbone model. Backbone model translate the input image into a feature map. By adopting SOTA methodology including SCSNet and ColTran, we will conduct the test and evaluation for the model. Also, we'll implement adaptive inference model and do the test and evaluation for the model. By repeating trial and fixing, our model will be improved steadily.

During the model implementation part, the works for the web UI/UX parts will be progressed together. By using Adobe XD, we will design the web layout of the user interactive web page, and implement with PyQt5. Also, the output file of our model is pth file. So the job for linking pth file and web application will be need. For web page distribution, we will use Flask and Amazon Web Service(AWS). We will put the development files made by using PyQt5 into Flask container. Then, upload them on AWS EC2 server. During this process, we'll test the files in local environment with using Flask. If the test done, Uploading them on AWS EC2 server would be the next step. Specific roles for the project are in Fig. 4 and weekly schedules are in Fig. 5

After above jobs are done, we'll test our web app in web environment. Evaluation standards would be how the model colorize black-and-white photography properly and whether the application runs well. Our project requires quite a lot of tasks. As our team has five of teammates, and the project will be conducted over semester, jobs are neither too much nor less. We assigned proper part to each teammate.

Dongin Jung	Gathering Datasets(MHMD, Youtube-8M) & Data preprocessing
	Adding Noise & Augmentation to Data
Jaeyeon Kim	Baseline Model Implementation & Performance Verification
Sabyeol Shin	Model Testing & Evaluation With Adopting SOTA Methodology
	Adaptive Inference Model Implementation, Testing & Evaluation
Changheon Kim	Layout Designing with Adobe XD Implementation with PyQt5
Mingeun Jang	Designing UI/UX Implementation of Linking .pth and Web by using Flask & AWS

Fig. 4. Specific Roles for the Project

Categories		September				October				November				December			
Data	Works	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4
Data Preprocess	Gathering Datasets(MHMD, Youtube-8M) & Data preprocessing	■	■	■	■												
	Adding Noise & Augmentation to Data				■	■	■										
Baseline Model	Baseline Model Implementation Performance Verification						■	■	■	■	■						
Backbone Model	Model Testing & Evaluation With Adopting SOTA Methodology						■	■	■	■	■	■	■	■	■	■	
	Adaptive Inference Model Implementation, Testing & Evaluation													■	■	■	■
Web UI/UX	Layout Designing with Adobe XD Implementation with PyQt5	■	■	■	■	■	■	■	■	■							
	Designing UI/UX Implementation of Linking .pth and Web by using Flask & AWS								■	■	■	■	■	■	■	■	■

Fig. 5. Weekly Schedules

References

1. Zhang, Richard, et al. "Real-time user-guided image colorization with learned deep priors." arXiv preprint arXiv:1705.02999 (2017).
2. Su, Jheng-Wei, Hung-Kuo Chu, and Jia-Bin Huang. "Instance-aware image colorization." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020.
3. Chen, Jin, et al. "Dynamic region-aware convolution." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021.
4. Kumar, Manoj, Dirk Weissenborn, and Nal Kalchbrenner. "Colorization transformer." arXiv preprint arXiv:2102.04432 (2021).
5. Lee, Sanghyeon, Jooyeol Yun, and Minho Park. "iColoriT: Towards Propagating Local Hint to the Right Region in Interactive Colorization by Leveraging Vision Transformer." arXiv preprint arXiv:2207.06831 (2022).