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: Hint-Based · 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.

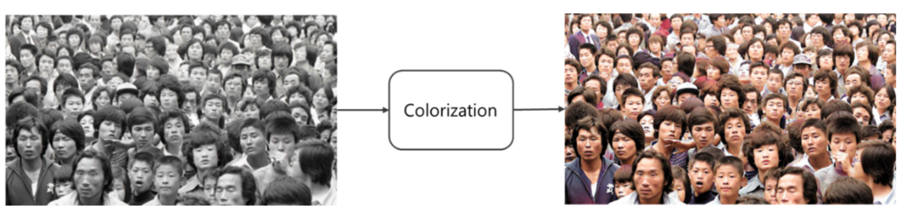

Fig. 1. Colorization of old grayscale image

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.

We developed a website that can check the colorized results through the developed model. Another page of the website allows users to download model and code files, so user can colorize in a local environment.
Evaluation of model was conducted through a comparison with existing models with PSNR, and a survey was conducted to evaluate how well people distinguish a real image from a restored image.

2 Background and Related Work

2.1 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.

2.2 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.

3 Design for the Proposed Service

3.1 Challenges

Context Confusion A context confusion in which various situations are contained in one image occurs. Object detection based coloration requires a lot of inference time compared to performance.

Fig. 2. Context Confusion

Many Objects, High Resolution AI model should efficiently processes photos with high resolution and many objects.

Ill-Posed Problem Overhead occurs when a person has to search for a reference image directly.

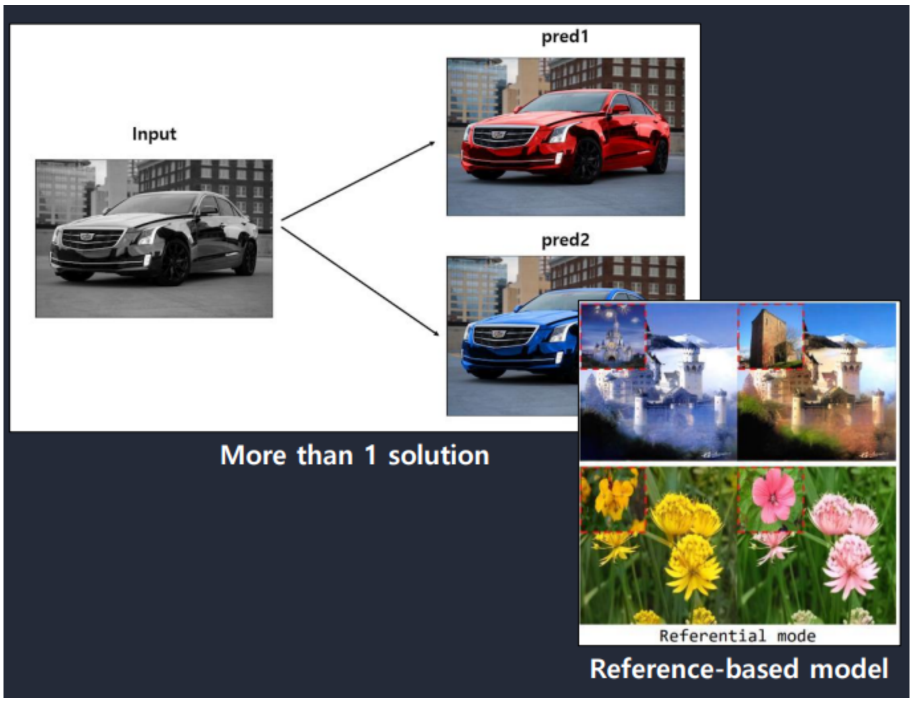

Fig. 3. Ill-Posed Problem

3.2 UI / UX

PyQt to Flask We changed PyQt code to flask code in order to distribute the colorization model and GUI to server. We embodied connection of uploaded image and model, function of colorization button, save button, restart button and save button. Besides, We specified the importance of colorization of old recorded image and some examples in main page. Unfortunately, we couldn't embody hint method function in PyQt GUI by using flask. So, we distribute GUI github link to the server too.

Main function of PyQt GUI in PyQt GUI, we embody the colorization button and hint based method. In original version, colorization is always applied when we use hint method. Instead, we added colorization button. by adding the button, the changed hint is applied when we click the button not applied immediately. GUI can be downloaded in github which is `https://github.com/SecAI-Lab/SWE3028-Fall-2022/tree/master/TeamB/Colorization_UI`.

3.3 Dataset

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.

ImageNet 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. The project has been instrumental in advancing computer vision and deep learning research. The data is available for free to researchers for non-commercial use.

ActivitiNet ActivityNet is a new large-scale video benchmark for human activity understanding. ActivityNet aims at covering a wide range of complex human activities that are of interest to people in their daily living. In its current version, ActivityNet provides samples from 203 activity classes with an average of 137 untrimmed videos per class and 1.41 activity instances per video, for a total of 849 video hours. Within this dateset, we extract the frames per 5 seconds.

3.4 Colorization Model

Context-adaptive Colorization We adopt Context-adaptive Colorization and Patch-wise inference. From the research, we found out that colorization based on object detection costs much and takes quite a long time. Also, if the image includes many of contexts, the result would be not good. So, we choose to inference the context patch by patch. By doing this, each of the patch will have less context than image-wise. The input RGB image would be transformed LAB image and only get L(Lightness) value. It means the image will be changed to gray scale image in the perspective of RGB image. During the procedure of encoding, feature extractor will extract the features from L image. After that, classification loss will be computed. Then, the decoding will follow with colorization network. The output would be the LAB image.

Hint-based Colorization To solve ill-posed problem, we choose to use hint-based colorization. After the colorization, the hint will be given to the model and do re-colorization. During the training, the model will learn the information of

Fig. 4. Patch & Image-wise inference

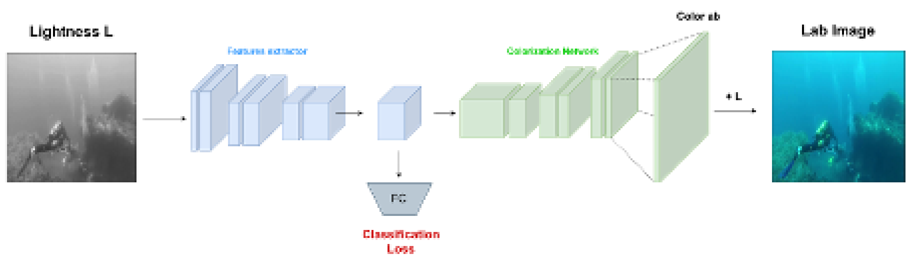

Fig. 5. Our method

the context and knows which colors would be appeared to the specific context. If the user choose the place to put the hint, the ab gamut will be appeared and the user can pick color they want to put. With this method, we can get the image that represents our past century manually.

Evaluation We conducted evaluation with PSNR and survey. The PSNR block computes the peak signal-to-noise ratio, in decibels, between two images. This ratio is used as a quality measurement between the original and a compressed image. The higher the PSNR, the better the quality of the compressed, or reconstructed image.

$$PSNR = 10log_{10}((MAX_I)^2/MSE) \tag{1}$$

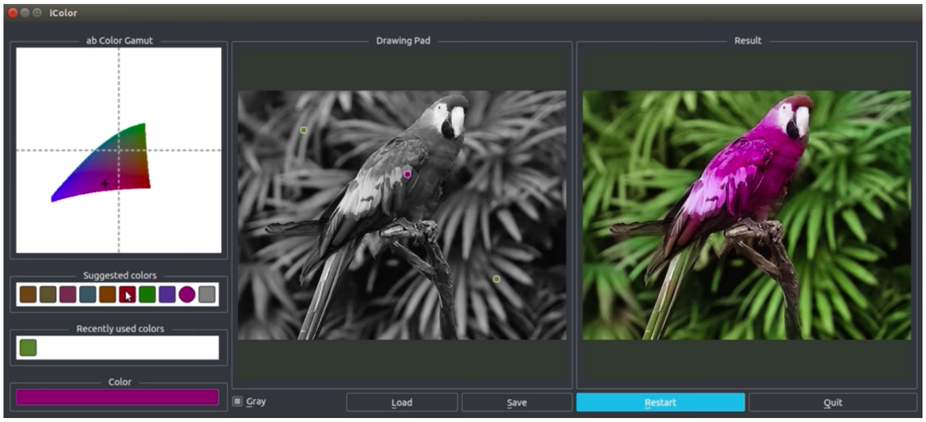

Fig. 6. Hint-based Colorization

Also, the survey items include 42 images and the answer would be one of "looks like real image", "Don't know", and "looks like creation of AI".

Fig. 7. Survey

Hyperparameters With the 0.3 million images, we set batch size to 256, epoch to 50 and initial learning rate to 1e-4. Within every epoch, learning rate will be updated with Adam optimizer. Adam is an alternative optimization algorithm that provides more efficient neural network weights by running repeated cycles of "adaptive moment estimation." Adam extends on stochastic gradient descent to solve non-convex problems faster while using fewer resources than many other optimization programs. With Adam, the model can converge to global minimum rather than local minimum.

4 Implementation

4.1 Demo

UI / UX In Fig. 8, there are Colorize, load, save, restart and quit button. images colorized by our models by clicking Colorize button. The colorized image can be downloaded by save button. In Fig. 9, hint based method is added. By selecting the color of image's location, colorization is influenced by its colors.

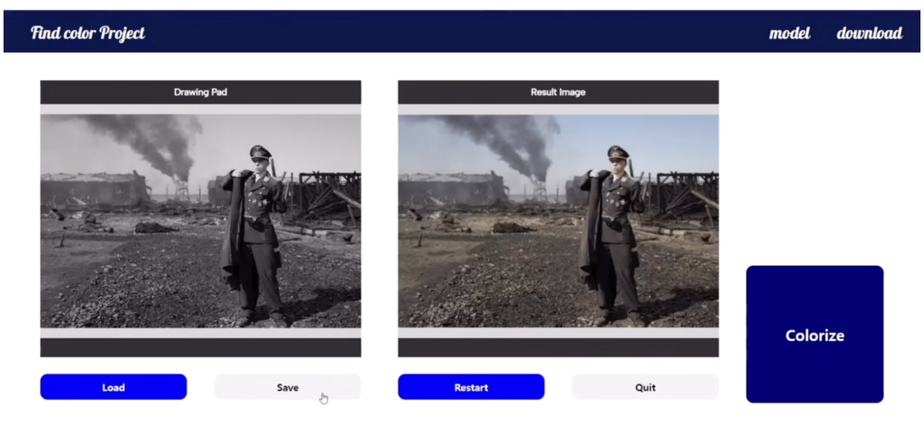

Fig. 8. Demo of UI / UX. By clicking colorize button, gray-scaled image is translated to colored image in 20 seconds.

Visualization We conducted visualization Fig. 10 with several other colorization method. One is instance-awareness colorization, and the other one is MyHeritage. Compare with Ground-Truth image, we can see how the images are well colorized.

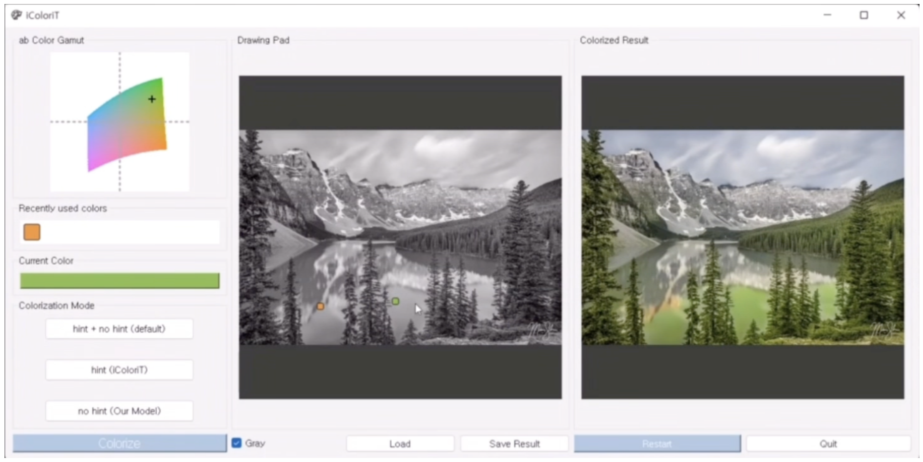

Fig. 9. Demo of GUI. It is similar with html UI. but hint method can be used in this GUI.

Fig. 10. Visualization

5 Conclusion

5.1 Evaluation

Quantitative Evaluation The quantitative evaluation was made with an index called PSNR, and it was confirmed that it had a better PSNR score than existing methods.

Category	InstanceAware	MyHeritage	Ours
Animal	31.70	28.70	31.65
Nature	29.61	28.57	29.67
Building	33.00	29.64	32.85
Soldiers	31.26	29.64	32.53
People	30.86	28.94	30.90
Total Average	30.88	28.93	31.08

PSNR

Fig. 11. Quantitative Evaluation by PSNR

Qualitative Evaluation The qualitative evaluation consisted of 17 surveys of 42 images, and people do not seem to be able to distinguish between real and restored images.

	Recall	Precision
Max	0.67	1.00
Min	0.00	0.00
Std	0.19	0.22
Mean	0.32	0.51

42장의 이미지에 대한 17명의 설문조사

Fig. 12. Quantitative Evaluation by PSNR

5.2 Limitations

As it produced its own UI/UX, there was a time limit and more survey samples were not secured. In addition, as it was conducted over a semester, various research and experiments could not be conducted, but a model suitable for the research neck was developed. If a little more time is given, more diverse research and experiments will be able to develop models that are more suitable for research purposes.

5.3 Evolution

Develop Our Own UI/UX With the distribution of web apps in mind, UI/UX was developed by having a development environment and framework suitable for it.

Survey Sample If a survey of more people targeting various age groups is conducted, meaningful results can be derived.

A Variety of Research and Experiments Based on existing research, various techniques were modified/combined and a baseline was developed solely for research purposes.

Collect meaningful datasets The dataset was constructed only with old photos of Korea, and the model was trained to meet the research purpose.

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).