International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 2 Feb. 2023, pp: 525-529 www.ijaem.net ISSN: 2395-5252

Old-school Image and Video colorization using Deep learning

Sonia Das [1], Abhinav Prakash[2], Ankit Aman[3], Bishwajit Paul[4], Charan Kumar K[5]

[1]*Asst. Professor, Department of Computer Science & Engineering, T John Institute of Technology, Bengaluru, India.*

[2,3,4,5]*Students, Department of Computer Science & Engineering, T John Institute of Technology, Bengaluru, India*

Date of Submission: 03-02-2023 Date of Acceptance: 17-02-2023

ABSTRACT - Colorization of old images and videos using deep learning involves using machine learning techniques to automatically add color to black and white or grayscale images and videos. There are several approaches to using deep learning for colorization, including using convolutional neural networks (CNNs) trained on large datasets of color images to predict the colors of specific pixels in the input image or video. One way to approach the problem is to use an autoencoder architecture, where the model is trained to reconstruct a colored version of the input image or video from its grayscale version. Another approach is to use a generative adversarial network (GAN), where the model is trained to generate realistic-looking colored versions of the input image or video. To train a deep learning model for colorization, a dataset of color images is typically used. The model is then trained to predict the colors of the pixels in the input image based on the colors of the pixels in the corresponding regions of the training Images. The model can then be used to colorize new images or videos by predicting the colors of the pixels in the input image or video based on the patterns and relationships learned from the training data.

I. INTRODUCTION

Old-school image and video colorization is the process of adding color to black-and-white images or videos using deep learning techniques. This is done by training neural networks on large datasets of colored images so that the networks can learn to predict the color of an image based on its grayscale input. This technology has improved significantly over recent years and has been used for a wide range of applications, from restoring old film footage to creating new and unique art styles. The use of deep learning has enabled high-quality colorization results to be achieved with much less manual effort than previously possible.

The process of colorization typically involves feeding the grayscale image through a deep convolutional neural network, which outputs a colored version of the image. The network is trained on large datasets of colored images, which allows it to learn the relationships between the grayscale input and the corresponding color output. The network can then be used to colorize new grayscale images, either by fine-tuning specific datasets or by applying the trained network directly to the input image. For video colorization, the process is similar, but the network must be able to handle sequential inputs and outputs, as the color of each frame is dependent on the previous frames. Recurrent neural networks (RNNs) or long short-term memory (LSTM) networks are often used to handle this. In conclusion, the use of deep learning has revolutionized the field of old-school image and video colorization. With its ability to learn complex relationships between grayscale inputs and color outputs, deep learning has enabled high-quality colorization results to be produced with much less manual effort than was previously possible. Additionally, deep learning-based colorization has opened up new possibilities for preserving and revitalizing historical images and videos. With its ability to automatically add color to old black-and-white media, this technology can bring new life to historical artifacts, making them more engaging and accessible to a wider audience. Furthermore, deep learning-based colorization has also found use in the entertainment industry, where it has been used to add color to classic movies and TV shows, allowing viewers to experience these works in a new way. Despite its advances, deep learning-based colorization is imperfect and still faces challenges, such as accurately handling complex scenes and colorizing certain textures, like hair and fur. However, as deep learning techniques continue to advance, these challenges will likely be overcome and colorization will become even more accurate and versatile.

Fig 1: Old School Image Colorization

1.1 OBJECTIVE

The objective of "old school image and video colorization" is to add color to black and white or grayscale images or videos. This can be done manually or with the use of software and algorithms that attempt to automatically determine the correct colors for the image or video. One reason for colorizing images and videos is to make them more visually appealing and engaging. Black and white images and videos can be perceived as dull and lacking in contrast, so adding color can help to bring them to life and make them more interesting to look at. Another reason for colorizing images and videos is to restore or preserve historical and cultural information. Many images and videos from the past were captured in black and white, and adding color can help to give a sense of what these scenes may have looked like at the time they were captured. This can be particularly important for historical documents and artifacts, as it can help to provide a more accurate representation of the period. Overall, the main objective of image and videocolorization is to enhance these media's visual appeal and cultural value.

1.2 PYTHON USED IN MACHINE LEARNING

Python is a simple and trustworthy programming language that enables programmers to create reliable, readable software.

Python was primarily employed as the programming language for projects involving many developments and cooperative implementation. Tests and complex machine learning tasks can be completed quickly, so you can build prototypes faster.

Python's amazing libraries and frameworks are another reason to master Python for machine learning.

Machine learning has had a major impact on the modern world.

In the world we live in, new applications are constantly being developed. Python is used by developers at all stages of problem-solving.

Python practitioners assert that they believe the language is well suited for AI and machine learning.

II. LITERATURE SURVEY

Old-school image and video colorization is an area of computer vision and multimedia processing that has received significant attention from researchers and practitioners in recent years. The goal of this task is to automatically add color to grayscale images or videos, typically based on historical or artistic considerations.

The earliest work in this area can be traced back to the 1990s when researchers first started exploring the use of machine learning algorithms for colorization. These early efforts were limited by the computational resources available at the time and were typically based on simple statistical models that were unable to capture the complex relationships between colors and patterns in real-world images.

In the early 2000s, several researchers proposed the use of more sophisticated machine learning algorithms, including neural networks, for image colorization. These methods typically focused on learning mappings from grayscale images to their corresponding color images and were able to achieve improved results compared to earlier methods.

In recent years, deep learning methods have become increasingly popular for image and video colorization, as these models can capture complex relationships between colors and patterns in large datasets. For example, convolutional neural networks (CNNs) and Generative Adversarial Networks (GANs) have been used to colorize images and videos with remarkable results.

One of the most notable examples of the use of deep learning for colorization is the "DeOldify" project, which uses a GAN architecture to colorize grayscale images and videos in real-time. This project has been widely adopted by researchers and practitioners and has set the benchmark for modern image colorization techniques.

In recent years, there has also been a growing interest in the use of deep learning for video colorization, which is a more challenging problem due to the temporal relationships between frames. Many of the same techniques used for image colorization can be adapted to video colorization, and researchers have proposed various methods for

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 2 Feb. 2023, pp: 525-529 www.ijaem.net ISSN: 2395-5252

combining frame-level colorizations into a coherent video.
Overall, the field of old-school image and video colorization has made significant progress in recent years, and there is ongoing research into developing more sophisticated methods for this task.

III. ALGORITHM USED

There are several algorithms and mathematical equations used in old-school image and video colorization, some of which include:
Histogram Matching: This method involves matching the color distribution of the grayscale image to a reference color image. The idea is to adjust the color channels of the grayscale image so that its color histogram matches that of the reference image.

Color Transfer: This method involves transferring the color distribution from a source color image to a grayscale image. The idea is to find a mapping between the color distributions of the source and target images and then apply this mapping to the grayscale image.

K-Nearest Neighbour (KNN) based colorization: This method involves finding the K-nearest neighbors in a reference color image for each pixel in the grayscale image and then using the average color of these neighbors as the color of the grayscale pixel.

Convolutional Neural Networks (CNNs): In recent years, CNNs have become the most popular deep-learning method for image colorization. A typical CNN-based colorization method involves training a network to predict the color channels of an image given its grayscale input. The network architecture is typically based on a combination of convolutional and up-sampling layers, and may also include skip connections to capture long-range dependencies between color and grayscale pixels.

Generative Adversarial Networks (GANs): GANs are another popular deep-learning method for image colorization. A GAN-based colorization method involves training two networks: a generator network that maps grayscale images to color images, and a discriminator network that evaluates the quality of the colorization. The goal is to find a mapping that generates realistic colorizations that are difficult for the discriminator to distinguish from real color images.

These are just a few examples of algorithms and equations used in classical image and video colorization. The specific method used depends on the requirements of the task at hand, and researchers are constantly exploring new methods and techniques for this field.

IV. RELATED WORKS

Our project is inspired by Richard Zhang, Philip Isola, and Alexey A. Efros - Colorization of colorful images. Zhang's system uses Alex Nets layers trained on the ImageNet dataset. Our method also learns to classify colors, but with some improvements to the loss function.

Another inspiration is ZHITONG HUANG, NANXUAN ZHAO, and JING LIAO. They propose the first unified framework Unicolor to support colorization in multiple modalities, including both unconditional and conditional ones, such as stroke, exemplar, text, and even a mix of them. Instead of learning a separate model for each type of condition, we used a two-stage colorization framework for combining various conditions into a single model.

Another inspiration is Brian Sam Thomas, Rajat Dogra, Bhaskar Dixit, and Aditi Raut They design and build a Convolutional Neural Network (CNN)- Long Short-Term Memory (LSTM) that accepts a black and white image as an input and convert it to a colorized version of the image as an output. The color information is stored in an LSTM and provided for coloring the next image. The system only colors images that have been studied in the past without direct human intervention.

V. SYSTEM ANALYSIS

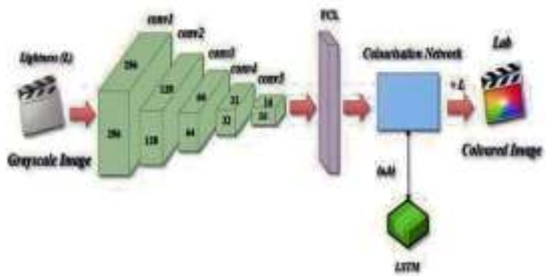

Fig 2: Network Architecture

As shown inNetwork Architecture. For video, we use LSTM. For single images, LSTM is omitted.

The following figure shows the overall pipeline or architecture of the network. as pictured. At 1, the input images (frames extracted from the video) are passed through a CNN layer. Each convolutional layer is a series of convolutions & Merge layers. Five such layers are used. The output of each layer goes through batch normalization before reaching the next layer.

The output of the fifth layer is passed directly to the fully connected layer for classification.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 2 Feb. 2023, pp: 525-529 www.ijaem.net ISSN: 2395-5252

At this level, various objects are identified/classified based on the SoftMax function. Lab color space data is generated. For video, this data is stored in an LSTM as each frame enters the colorization network. If the input is a single image, this data is used directly to colorize the image and the color image is used as output.

A. Abbreviations and Acronyms:

LSTM – Long Short-Term Memory is a module used with neural networks to store information after an output is generated. LSTM stores information as needed.

CNN – Convolutional Neural Network is a neural network that can process image data. It consists of 3 layers, 1 convolutional layer, 1 pooling layer, and 1 fully connected layer. We'll take a closer look at how each works in the next section.

B. Target function:

Luminance input channels $\mathbf{X} \in \mathbf{R}^{\mathbf{H} \times \mathbf{W} \times 1}$, Our goal is to study the mapping of $\widehat{\mathbf{Y}} = \mathbf{F}(\mathbf{x})$ to the two associated color channels $\mathbf{Y} \in \mathbf{R}^{\mathbf{H} \times \mathbf{W} \times 2}$, Where H and W are the sizes of the image. It does this in the CIE Lab color space. Since the distance in this spatial model is the perceptual distance, the natural objective function used for the Euclidean loss L2 between the predicted color and the actual color is:

$$L_2(\widehat{\mathbf{Y}}, \mathbf{Y}) = \frac{1}{2} \sum_{h,w} \|\mathbf{Y}_{h,w} - \widehat{\mathbf{Y}}_{h,w}\|_2^2$$

However, the loss is not stagnant to the parental ambiguity and the multimodal nature of the Coloring problems. If an object can have a unique set of ab values, the optimal Euclidean loss solution is the average of the set. In color prediction, this averaging effect produces a grayish, blurry result.Also, if the set of plausible colors is non-convex, the solution is outside the set, giving implausible results.

VI. CONCLUSION

Colorizing an image is the process of adding color to a grayscale image or video. There are several ways to colorize an image, including manual methods, semi-automatic methods, and fully automatic methods. In the process of coloring an image by hand, an artist or graphic designer uses image editing software such as Photoshop to manually add color to an image.This method can produce high-quality results, but it is time-consuming and requires a skilled artist. Semi-automatic image colorization methods involve using algorithms to generate a colorization that is then refined by a human artist. This can produce good results, but it requires human intervention and is somewhat time-consuming. Fully automatic image colorization methods use algorithms to generate a colorization without any human intervention. These methods have improved significantly in recent years, but they still have limitations and may not produce results that are as visually appealing as those produced by manual or semi-automatic methods. mage colorization in general is an active area of research, and there is still much to be improved in terms of the accuracy and speed of automated methods. However, it is an important task for many applications, including digitally restoring old photos, improving accessibility of grayscale documents and images for people with color vision impairments, and enhancing the visual appeal of videos and images.

REFERENCES

[1]. Richard Zhang, Phillip Isola, Alexei A. Efros. “Colourful Image Colourisation” In the University of CaliforniaBerkeley, October 2016.
[2]. http://richzhang.github.io/Colourisation/
[3]. ZHITONG HUANG, NANXUAN ZHAO, JING LIAO. “A Unified Framework for MultiModal Colorization with Transformer” In University of Hong Kong, China, September 2022.
[4]. https://luckyhzt.github.io/unicolor.
[5]. B. S. Thomas, R. Dogra, B. Dixit, and A. Raut, "Automatic Image and Video Colourisation using Deep Learning" 2018 International Conference on Smart City and Emerging Technology (ICSCET), 2018, pp. 1-4, doi: 10.1109/ICSCET.2018.8537308.
[6]. LvMin Zhang, Chengze Li, Tien-Tsin Wong, Yi Ji, and ChunPing Liu. 2018. Two-stage Sketch Colorization. ACM Trans.Graph. 37, 6, Article 261 (November 2018), 14 pages.
[7]. https://doi.org/10.1145/3272127.3275090.
[8]. Huiwen Chang, Ohad Fried, YimingLiu,StephenDiVerdi, Adam Finkelstein. “Palette-based Photo Recoloring” ACM Transactions on Graphics (Proc. SIGGRAPH), July 2015.
[9]. Jeff Hwang, You Zhou. “Image Colourisation with Deep Convolutional Neural Networks” In Stanford University Research Projects, 2016.
[10]. A. Krizhevsky, I. Sutskever, and G. E. Hinton. “Imagenet classification with deep convolutional neural networks.” In Advances in neural information processing systems, pages 1097–1105, 2012.
[11]. Raj Kumar Gupta, Alex Yong-Sang Chia, DeepuRajan, Ee Sin Ng, and Huang

Zhiyong "Image Colourisation using similar images." In Proceedings of the 20th ACM International Conference on Multimedia, ACM (2012) 369-378.

[12]. K. He, X. Zhang, S. Ren, and J. Sun. "Deep residual learning for image recognition." In arXiv preprint arXiv:1512.03385, 2015.

[13]. Larsson, G., Maire, M., Shakhnarovich, G.: Learning representations for automatic colorization. European Conference on Computer Vision (2016)