Vivid And Diverse Image Colorization Using Deep Learning Techniques

[1]K. Devendra Reddy, [2]G. Hitesh Reddy, [3] S.Pothumani

Department of computer science and Engineering, Sathyabama Institute of Science and Technology
Chennai 600119, India

Abstract: The task of assigning colors to grayscale images, referred to as image colorization, has been effectively tackled using deep neural networks. While various research and review papers have addressed this problem, they have been classified based on the criteria such that the number of the colored output images, colorization methods, technique or network used, and network paths, with a focus on commonly used datasets and comparison measure. To advance the field, it would be beneficial to unify methods and datasets to showcase the progress made by new models. An approach on the deep learning for automatic colorization can be proposed, whereby a Convolutional Neural Networks (CNN) is used to map grayscale images and user cues to output colorizations. This network integrates low level information from sources with high level info learned from large-scale data, resulting in real-time, desaturated outputs that users can edit. This differs from previous methods that depend on user input and produce non-real-time desaturated outputs. Neural networks have been trained on a large dataset to reduce dependency on specific approaches. Applications for image colorization systems include astronomical photography, CCTV photos, electron microscopy, and other domains.

Key Words: Image colorization, Generative Adversarial Network, Convolutional Neural Networks, Image Processing.

I. INTRODUCTION

There is currently two main approach to colorizing images in computer graphics: user driven edit propagation and automatic colorization. The former involves a user providing colored strokes to a grayscale image, with the enhancement strategy producing a colorized picture that adheres to both the user's strokes and defined picture priors. However, this approach requires significant user input and may produce incorrect colors in regions with few shades. The latter approach uses data-driven methods, such as parametric mappings or deep neural networks, to automatically colorize grayscale images. While this approach is fully automated, it may also produce inaccurate colors and artifacts. To overcome the limitations of both approaches, we propose a combined approach that leverages large-scale image data to improve color symbolism while also integrating user input. Specifically, we propose training convolutional neural network (CNN) on a large data to generate low-resolution images, which can then be colorized using user input. By using this approach, we aim to outperform both existing methods and improve the accuracy of colorized images. Furthermore, we plan to extend this approach to also detect deep fakes and colorize videos.

II. LITERATURE SURVEY

Colorization, which involves adding realistic colors to black and white images, has been extensively researched using convolutional neural networks specifically designed for image data. However, cartoon images have different color characteristics from natural images, and automatic coloring can be challenging due to the high uncertainty and subjective evaluation of cartoon colors. Different approaches have been proposed to address this challenge, including using an image classifier and fine-tuning models with specific datasets. Other related techniques aim to enhance existing color information or modify color palettes. These methods are often used to correct camera defects, and their simple transformations are used to describe the behaviours of more complex algorithms, even in the domain of CNNs. Overall, the idea of generating color information from other data is a popular image-to-image task in the computer vision field, and recent works have used various approaches to tackle this problem.

Shweta Salve utilized Google image classifier, Inception Resnet V2, in a similar approach for automatic colorization. Their system model includes an encoder, feature extractor, fusion layer, and decoder, and can produce satisfactory results with adequate resources and a large data set.

Yu Chen also proposed a method for colorizing Chinese films from the past by fine-tuning their own data set with existing data, using multi scale convolution kernel and low or middle features from VGG-16. V.K.

Putri method converts sketches and diagrams into colorful images using a sketch colour model and color prediction in CIE Lab color space, although its effectiveness is limited by the availability of data.

Richard Zhang's solution, which is based on a large data set and a single feed forward pass in CNN, was primarily focused on training and was able to deceive 32% of human subjects during testing.

While working with color channels in images is a well-researched area, generating color information from other data has received comparatively limited research and application. Color enhancement methods, which improve image quality by adjusting contrast or modifying color palettes, serve as a precursor to colorization techniques and are frequently used to describe the behaviours of more complex algorithms like CNNs. Despite its relatively limited application and research compared to other image-related problems, automatic colorization has come to be increasingly popular in the computer vision field. In recent years, various approaches have been used and several works have been published on this topic. In addition, there are numerous algorithms that aim to improve the color information in images or modify the color palette, which are considered color enhancements rather than colorization methods. These methods are often used to correct camera defects, such as overexposure or underexposure, by adjusting contrast through histogram equalization. Although non-parametric, these simple transformations are frequently used to describe the behaviour of more complex algorithms, even in the domain of CNNs. For example, it is possible to say that a CNN performs a transformation like histogram equalization.

III. EXISTING WORK

The previous research in the field of automatic colorization is reviewed with a focus on works that have influenced and shaped the current thesis. Though working with color channels in images is a well-researched area, the idea of generating color information given other data has not been extensively researched compared to similar problems. However, in recent years, several works have been published in this area, making it a popular image-to-image task in computer vision. Additionally, algorithms that improve poor color information or modify the color palette of an image are discussed as precursors to full colorization techniques. These methods often serve to remedy camera defects and their simple transformations are used to describe the behaviour of more complex algorithms like CNNs.

IV. PROPOSED WORK

Deep convolution neural network has proven to be effective on learning features from visual data and could be used in various computer vision tasks such as object discovery, segmentation, and image captioning. The success of these model depend on the amount of data used for training, and large-scale image datasets like ImageNet and COCO have been proposed for this purpose. However, collecting and annotating such datasets can be challenging, time-consuming, and prone to errors. Transfer learning is a common approach when a large, annotation dataset is not available, where pre-trained models with general knowledge are fine tuned for specific task. This approaches provide good starting point for building convolutional models with previously learned feature. Additionally, techniques such as transfer learning with progressive resizing can be used to improve the CNN's performance and achieve better results.

The proposed method involves training the Convolutional Neural Network (CNN) on randomized data to simulate a range of potential inputs, including those with common errors or mistakes. In addition, a separate neural network would be used to combine these randomized inputs with the trained CNN. To ensure the model is not overly reliant on the randomized inputs, a large-scale training dataset would also be used to help the model learn to make accurate predictions about natural color imagery. During training, randomized inputs would be introduced in a controlled manner to avoid the need for a large number of actual input examples and to reduce errors associated with these inputs. By relying more on the large-scale training dataset and less on the randomized inputs, the model can achieve greater accuracy and effectiveness.

A. Implementation Description:

- The proposed method involves training two models together, in which a generator model creates image examples, and a discriminator model evaluates those examples to determine if they are real or fake.
- This type of neural network architecture is commonly used in image recognition and processing. While there are other types of neural network, CNNs are favoured in identifying and recognizing objects.
- To train the models, the generator is first trained alone using feature loss, and then the critic is trained to be able to distinguish between the generator's outputs and real images. Finally, both models are trained together in a GAN environment. This approach is versatile and should perform well for various image modification tasks.
- The initial phase involves training the generator model using only feature loss. Then, the generated images can be used to train the critic model to distinguish between the generated and real images. Finally, both models are trained in a GAN environment. The advantage of this approach is that it can be used for various types of image modification and perform well.

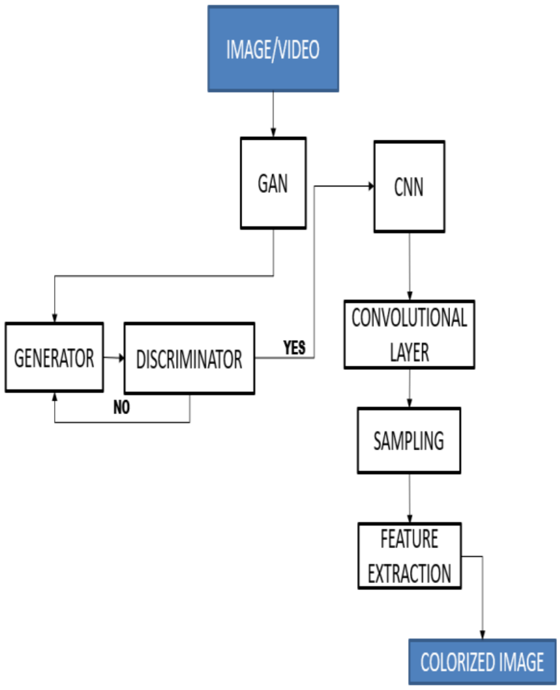


Fig 1. Workflow Diagram

B. Image Implementation

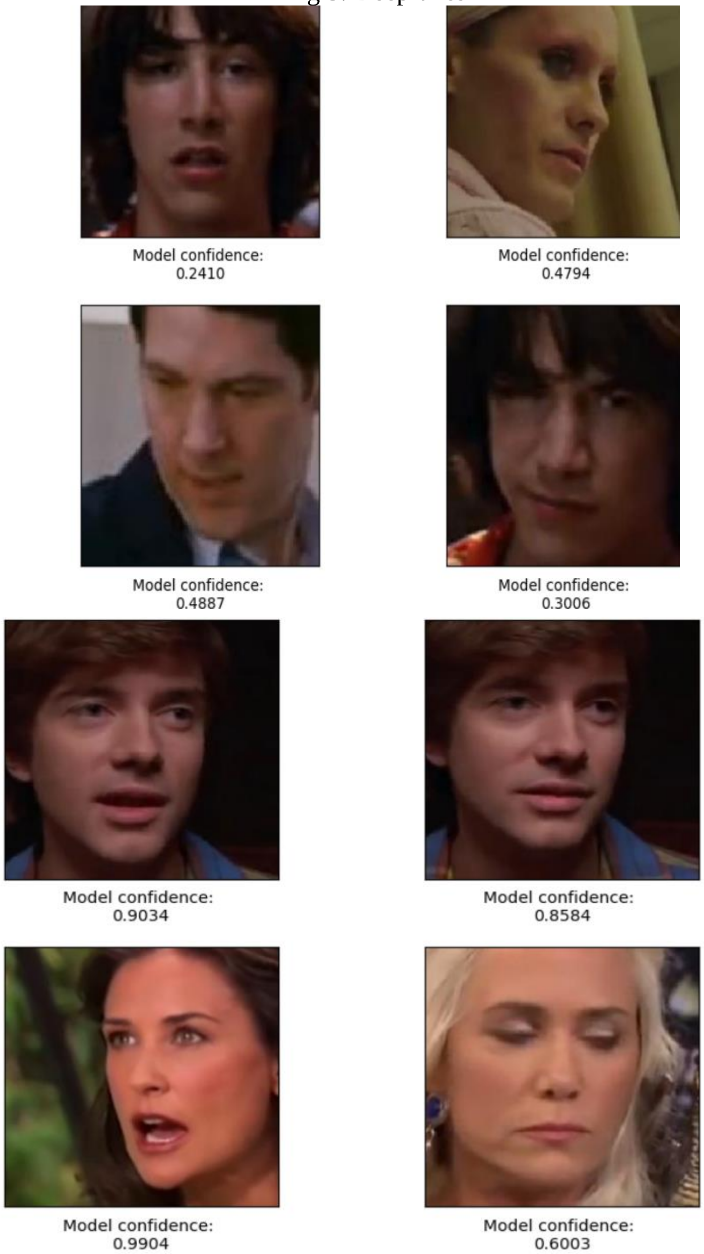
The process of adding realistic and plausible colors to grayscale images and videos is known as image and video colorization, respectively. The LAB color scheme is a commonly used method, where 'L' signifies brightness and 'a' and 'b' channels represent green, red and blue, yellow color component. Image colorization utilizes a Convolutional Neural Network (CNN) to process grayscale images, extract the luminance channel 'L,' and predict the corresponding 'ab' values. Video colorization, on the other hand, presents a challenge since each frame must be processed individually before merging them into the final colorized video, which can lead to stuttering, resulting in an unsmooth transition between frames. Nonetheless, efforts have been made to minimize stuttering while combining frames in video colorization.

Fig 2. Colorized grayscale images

C. Deep Fake Detection

The deployment of a deepfake detection system involves a series of significant steps. Initially, a varied collection of both genuine and counterfeit videos must be utilized to educate a deep learning model. The model should possess the capacity to precisely differentiate between legitimate and fake videos. After the completion of model training, it can be included in an existing platform, including web applications, mobile apps, or desktop software. The integration process should be designed to enable the system to process video inputs expeditiously and efficiently. The creation of a user-friendly interface that provides unambiguous feedback to users about video authenticity is also a crucial aspect. Additionally, regular monitoring and updating of the deepfake detection system are crucial to ensure its effectiveness against the ever evolving deepfake techniques. This can be achieved by routinely retraining the model and updating the deployment pipeline. In summary, deploying a deepfake detection system requires a systematic approach that involves meticulous planning, implementation, and monitoring to accurately identify counterfeit videos while ensuring user-friendliness. One way to distinguish deepfakes in images or videos is by presenting a probability-based confidence level.

Fig 3. Deepfakes



V. CONCLUSION

To utilize deep learning techniques to differentiate between genuine and counterfeit images or videos, as well as to colorize grayscale images or videos. This can be accomplished by implementing various Deep Learning algorithm such as Convolutional Neural Networks (CNN), Generative Adversarial Networks (GAN), and Deepfakes. Deep fakes are an algorithmic approach used for identifying fake images or videos. CNN can be utilized to extract background and object data from images or videos. Meanwhile, GAN can be utilized to add colors to grayscale images or videos based on input images or videos.

VI. FUTURE SCOPE

It is projected that the domain of deep learning for image and video processing will undergo advancements and refinements in the future. These advancements are expected to involve the creation of novel algorithms and architectures that enhance the precision and efficiency of deepfake detection and colorization operations. There is also mounting interest in utilizing deep learning techniques in other domain, such as the natural language process, speech recognition, and robotics. The availability of vast data repositories and cutting-edge computational resources is anticipated to expedite further advancements in this field. As the technology progresses, it is predicted that deep learning techniques will witness greater prevalence in diverse applications, including image and video processing, medical diagnosis, and autonomous vehicles.

VII. REFERENCES

1. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Deep lab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE Trans. Pattern Anal. Mach. Intell. 2017.
2. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017.
3. Pathak, D.; Krahenbuhl, P.; Donahue, J.; Darrell, T.; Efros, A.A. Context encoders: Feature learning by inpainting. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016.

4. Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z.; et al. Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
5. Zhang, R.; Isola, P.; Efros, A.A. Colorful image colorization. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016.
6. Thoma, M. A survey of semantic segmentation. arXiv 2016.
7. Noroozi, M.; Vinjimoor, A.; Favaro, P.; Pirsiavash, H. Boosting self-supervised learning via knowledge transfer. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 16–23 June 2018.
8. Beyer, L.; Hénaff, O.J.; Kolesnikov, A.; Zhai, X.; van den Oord, A. Are we done with the image net? arXiv 2020.
9. Jing, L.; Tian, Y. Self-supervised visual feature learning with deep neural networks: A survey. IEEE Trans. Pattern Anal. Mach. Intell. 2020.