ColorAI – Automatic Image Colorization using CycleGAN

Deven Bothra[1], Rishabh Shetty[2], Suraj Bhagat[3], Mahendra Patil[4]

[1]BE Computer Engineering, Atharva College of Engineering
[2]BE Computer Engineering, Atharva College of Engineering
[3]BE Computer Engineering, Atharva College of Engineering
4 Assistant Professor BE Computer Engineering, Atharva College of Engineering

---***---

Abstract -In this paper, we propose a fully automatic method for reproducing black and white images as color images. The methodology used was Cycle-consistent Generative Adversarial Networks (CycleGAN), which are very powerful for image-to-image translations. Image-to-image translation is a class of computer graphics and computer vision where we pass in an input image and get a reconstructed image as output. However, for this translation to happen, we need a paired image dataset which is not available for the problem of image colorization. This is where CycleGANs are powerful. CycleGANs are able to learn a mapping from an input domain A to another domain B without needing paired datasets. The model was trained on Intel Landscape Image dataset. The experiments show significant improvements on previously existing models, and gives a better, more general solution to the problem of Image Colorization

Key Words -Image colorization, Image processing, CycleGAN, Image translation, Unpaired training

1.INTRODUCTION

Automatic image colorization is the technology to automatically colour black-and-white images. It has been a popular area of research in the field of computer vision for several practical application areas including restoration and colorization of age-old images and videos, colorization of old black-and-white images, colorizing medical images such as chest X-rays, improving video quality of CCTV footages, etc. However, as most of the colors share similar gray values, the problem of effectively colorizing images is a very challenging. Traditional methods [1], [2], [3], [4], [5] are example-based colorization methods where example of colorized versions of grayscale images have to be provided while training. Methods [6], [16], [17], [18], are semi-automatic scribble-based approaches, where the user needs to provide some color hints to segment the images and also give information to color these segments. The user provides scribbles for the grayscale image which is then colorized through a colorization optimization alogirthm. However, example-based image colorization has limitations aspaired training data might not always be available for tasks such as restoration of old black-and-white photos and movies. Semi-automatic approaches require human intervention which needs a lot of time and effort.

In this paper, we use the CycleGAN architecture that can complete image translation from one domain to another without needing paired image datasets. Inspired by [12], [13], [14], our aim is to put out a more general model that can produce realistic colorization of grayscale images. The architecture comprises of two generators and two discriminators. The generator is responsible for transforming the image. The discriminators are convolutional neural networks that see the transformed image and judges them as real or fake. The model works on unpaired image datasets which makes the proposed method more generalized and easier than most of the existing methods.

2. Methodology

The CycleGAN architecture comprises of two generators – one for generating images from the first domain (Generator A -> Domain A) and the other for generating images from the second domain (Generator B -> Domain B), and two discriminators – one discriminator model that takes images from domain A and images generated by generator A and judges them as real or fake (Domain A -> Discriminator A -> Real/Fake) and the other discriminator that takes images from domain B and regenerated images generated by generator B and judges them as real or fake (Domain B -> Discriminator B -> Real/Fake). The generators and discriminators are trained just like normal GANs – in a zero-sum process. The generator learns to better decieve the discriminator and the discriminator learns to judge better.

Generator Network – The generator network consists of an encoder, a convolutional neural network and a decoder, a transpose convolutional neural network that converts a feature representation to a transformed image. The optimization of the generator network is done using four outputs from four loss functions – Adversarial Loss, Forward Cycle Loss, Backward Cycle Loss, and Identity Loss.

Discriminator Network – The two discriminators are CNNs that classify the generated image as real/fake. The model is optimized using mean squared error. The CycleGAN paper[10] suggested weighting the model to decelerate the changes to the discriminator. So, a weighting of 0.5 is used to have half the normal effect. The discriminator model is trained directly on the real and constructed images.

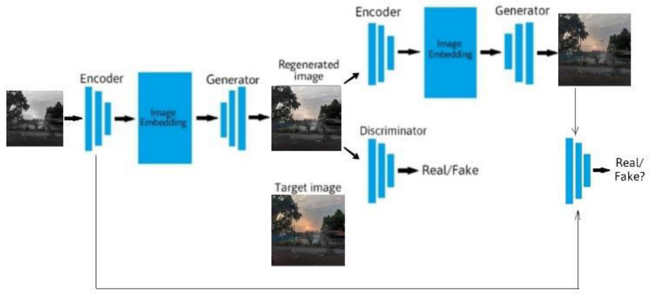

Figure 1: Data flow of CycleGAN Architecture

The output of the generator is fed to the discriminator as input. Assuming the source domain (Grayscale) is A and the target domain (Colored) is B, each domain consists of a pair of one generator and one discriminator namely Generator_AtoB, Discriminator A, Generator_BtoA, Discriminator B. The input image is fed into Generator_AtoB, giving an output of the colored image to discriminator B which then determines if the image is real or fake. A second input is fed into Generator AtoB which is expected to be output without any translation for identity mapping. Next, forward cycle loss is implemented by providing a grayscale image to Generator_AtoB which converts it to a colored output which is fed into Generator_BtoA in order to reconstruct it into grayscaled image again. Backward loss is implemented in the same process as above but in reverse, where the colored input is fed into Generator_BtoA to convert it to grayscale which is then reconstructed to a colored image by being passed through Generator_AtoB. These losses ensure transferring the underlying style in colored images and translate it in a realistic manner on grayscale images.
The generator and discriminator models do not converge. The model is saved periodically and is used to generate sample image colorizations during training. An equilibrium is found between the discriminator and generator after sufficient training of both the models.

3. Experiments and results

The model was trained of the Intel landscape Image Dataset. The model took around 19 hours to train on approximately 8k images of size 256X256. Once the model is trained, getting colorizations was pretty fast. The results achieved are pretty good and can identify most objects like trees, marshy lands, daylight, cars, humans, snow, stars, and the sky and appropriately color them without needing any human assistance. The colorizations are not true, and may differ with original scenes, but are reasonable and realisitic.

However, the model struggles to pick up different colors of human skin as the data it was trained on isn't sufficient for it to correctly discriminate skin colors. In order to overcome this drawback, the model can be trained and fine-tuned on a dataset that incorporates all possible scenarios in which the model can be used. This, however, would require huge computational power.

4. CONCLUSIONS

Automatic Image Colorization can be used in tasks like restoration of old black-and-white photos, colorization and enhancement of old movies, color amplification of medical images such as X-rays and MRIs for better diagnosis, colorization of low-quality CCTV footages, etc. A cycleGAN implementation makes the training and testing much more accessible with less constraints and provides a style transfer that is realistic in nature and appropriate in a real world setting which could be beneficial to enriching archival monochrome data be it image or video to appropriately mimic its long lost real life counterpart.

ACKNOWLEDGEMENT

We would like to express our special thanks and gratitude to Prof. Mahendra Patil. He was a constant force of help and guided us accurately and efficiently through all the problemsthat we faced during our research work. We would also like to extend our gratitude towardsour Principal, Dr. S. P. Kallurkar, and HOD of Computer Department, Prof. SuvarnaPansambal for their beneficial guidance and all the project coordinators, for their perseverancewithout which completing this research project would be very difficult.

REFERENCES

1. 1 G. Charpiat, M. Hofmann, and B. Scholkopf, "Automatic image colorization via multimodal predictions." In *European conference on computer vision*. Springer, 2008.
2. A. Y.-S. Chia, S. Zhuo, R. K. Gupta, Y.-W. Tai, S.-Y. Cho, P. Tan, and S. Lin, "Semantic colorization with internet images," in *ACM Transactions on Graphics (TOG)*, vol. 30, no. 6. ACM, 2011.
3. T. Welsh, M. Ashikhmin, and K. Mueller, "Transferring color to greyscale images," in *ACM transactions on graphics (TOG)*, vol. 21, no. 3. ACM, 2002, pp. 277–280.
4. X. Liu, L. Wan, Y. Qu, T.-T. Wong, S. Lin, C.-S. Leung, and P.-A. Heng, "Intrinsic colorization," in *ACM Transactions on Graphics (TOG)*, vol. 27, no. 5. ACM, 2008.
5. D. Patel and S. Parmar, "Image retrieval based automatic grayscale image colorization," *2013 Nirma University International Conference on Engineering (NUiCONE)*, Ahmedabad, India, 2013.
6. Jacob, Vivek & Gupta, Sumana. (2009). "Colorization of grayscale images and videos using a semiautomatic approach."
7. S. Wan, Y. Xia, L. Qi, Y. Yang and M. Atiquzzaman, "Automated Colorization of a Grayscale Image With Seed Points Propagation," in *IEEE Transactions on Multimedia*, vol. 22, no. 7, pp. 1756-1768, July 2020.
8. S. Titus and J. R. N.M., "Fast Colorization of Grayscale Images by Convolutional Neural Network," *2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET)*, Kottayam, India, 2018.
9. B. Rubben, B. Purnama and Adiwijaya, "The implementation of neighbor embedding on grayscale image colorization," *2016 4th International Conference on Information and Communication Technology (ICoICT)*, Bandung, Indonesia, 2016.
10. Jun-Yan Zhu, Taesung Park, Philip Isola, Alexei A. Efros, "Unpaired Image-to-image translation using Cycle-consistent Adversarial Networks," in *IEEE International Conference on Computer Vision*, 2017.
11. Y. Xiao, A. Jiang, C. Liu and M. Wang, "Single Image Colorization Via Modified Cyclegan," 2019 *IEEE International Conference on Image Processing (ICIP)*, Taipei, Taiwan, 2019.
12. Karen Simonyan and Andrew Zisserman, "Very deep convolutional networks for large-scale image recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2014.
13. Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, "Generative adversarial networks," in *Advances in Neural Information Processing Systems*, 2014.
14. Jason Antic, "A deep learning based project for colorizing and restoring old images," https://github.com/jantic/DeOldify/, 2018.

15. Jason Brownlee, "How to Implement CycleGAN model from scratch with Keras," https://machinelearningmastery.com/how-to-develop-cyclegan-models-from-scratch-with-keras/, 2019
16. Y.-C. Huang, Y.-S. Tung, J.-C. Chen, S.-W. Wang, and J.-L. Wu, "An adaptive edge detection based colorization algorithm and its applications," in *Proceedings of the 13th annual ACM international conference on Multimedia,* ACM, 2005.
17. Q. Luan, F. Wen, D. Cohen-Or, L. Liang, Y.-Q. Xu, and H.-Y. Shum, "Natural image colorization," in *Proceedings of the 18th Eurographics conference on Rendering Techniques,* Eurographics Association, 2007.
18. L. Yatziv and G. Sapiro, "Fast image and video colorization using chrominance blending," *IEEE transactions on image processing,* vol. 15, no. 5.