Dual Color-Image Discriminators Adversarial Networks for Generating Artificial-SAR Colorized Images from SENTINEL-1 Images

Youssef Mourchid, Marc Donias, Yannick Berthoumieu

Dual Color-Image Discriminators Adversarial Networks for Generating Artificial-SAR Colorized Images from SENTINEL-1 Images ⋆

Youssef Mourchid[1], Marc Donias[1], and Yannick Berthoumieu[1]

University of Bordeaux, CNRS, IMS, UMR 5218,
Signal and Image Group, F-33405 Talence, France
{youssef.mourchid, marc.donias, yannick.berthoumieu}@ims-bordeaux.fr

Abstract. In this paper, we introduce a new generative adversarial network (GAN) with dual image-color discriminators, to predict Artificial-SAR colorized images from SAR ones (Sentinel-1). Based on the conventional architecture of GANs, we employ an additional color discriminator that evaluates the differences in brightness, contrast, and major colors between images, while the image discriminator compares texture and content. To achieve the required level of colorization in the generation process, we employ non-adversarial color loss dedicated for color comparison, unlike conventional approaches that use only L_1 loss. Moreover, to overcome the vanishing gradient problem in deep architecture, and ensure the flow of low-level information inside network layers, we add residual connections to our generator that follows the general shape of U-Net. The performance of the proposed model was evaluated quantitatively as well as qualitatively with the SEN1-2 dataset. Results show that the proposed model generates realistic colorized images with fewer artifacts compared to the state-of-the-art approaches. This model helps to maintain color steadiness as well as visual recognizability at less textured large continuous regions, such as plantation and water areas, when it's difficult to be distinguished in SAR images.

Keywords: Colorization · GAN · SAR · Artificial.

1 Introduction

Synthetic Aperture Radar (SAR) images help to reveal important patterns on appointed regions and objects. So, it is useful to generate natural pseudo colors on high-resolution satellite images for understandability improvement. In the past few years, deep learning has had an enormous impact on the field of remote sensing [18, 21]. Among deep learning advantages, is to model highly non-linear relationships between remote sensing observations and the eventually desired geographical parameters that could not be represented by previous models. Data fusion is considered one of the most promising directions of deep learning in remote sensing [14]. A special case of this direction is a

⋆ This study has been carried out with financial support from the French Direction Générale de l'Armement (DGA) in the frame of the projet Man-Machine Teaming (MMT).

combined exploitation of SAR and Optical images, as these data modalities are completely different in terms of geometric and radiometric appearance. Optical images collect information about the chemical characteristics of the observed environment, while SAR images observe the physical properties of the target scene. On the other hand, SAR images are based on range measurements, while optical images are based on angular measurements. Most of the work for optical image colorization is majorly divided into three categories: scribble-based colorization [6, 9, 11], Example-based colorization [3, 10, 16] and deep learning-based colorization [2, 4, 8, 19, 20]. The scribble-based approaches require user intervention to scribble desired colors in certain regions. Then the colors are propagated to the whole image based on pixels adjacency criterion. The example-based approach, remove the burden of annotating the image with color scribbles by using a color reference image to transfer color. Whereas deep learning methods have considered the image colorization as a self-supervised problem, and they propose various convolutional network architectures with different loss functions. Moreover, a novel learning structure called generative adversarial network (GAN) [5] was popularly employed in this field for being able to instruct CNN to learn and generate realistic results. Isola *et al.* [7] (Pix2Pix) leveraged the power of conditional GANs in the context of image colorization, by coupling a DCGAN [7] based on the U-net architecture for the generator network and a Markovian definition for the discriminator, i.e. PatchGAN architecture. Nazeri *et al.* [12] proposed generative network-based on L1 loss in architecture with skip connections, within a "U-Net" shape. Xiao *et al.* [17] formulate the process of colorization as image-to-image translation based on a variation of GANs called CycleGAN to predict RGB color space, where high-level semantic identity loss and low-level color loss are employed. Caoe *et al.* [1] proposed a conditional generative adversarial network (cGAN) based on a fully convolutional generator with multi-layer. Suarez *et al.* [15] proposed an approach for colorizing the infrared image based on a triplet DCGAN architecture.

In remote sensing, generating artificial colorized images is an ill-posed problem with multiple solutions, since most of the natural objects can have several colored appearances. Therefore, holding a spatially stable color theme for these objects that belong to main land-cover categories is a higher priority. Inspired by GAN-based image colorization techniques, we propose a new Artificial-SAR colorization model based in GAN for Sentinel-1 [13] images. The contributions of this paper are as follows:

- Proposing a new model for Artificial-SAR colorized images that employs two discriminators to promote the restitution of fine-colors of the generated images. The first discriminator is dedicated to the pixel level and the second one for image color level.
- Employing an improved version of U-net architecture for the generator, using residual connections, to avoid the vanishing gradient problem, and ensure the flow of low-level information inside the network using skip connections.
- Training the generator with discriminators using non-adversarial color loss, in order to evaluate the difference in brightness, contrast and major colors between generated images and original ones.

2 Proposed Model

Generative Adversarial Networks (GANs) [5] was popularly employed for image generation and automatic colorization, and proved to be powerful in making up textural and color details. GANs consist of two neural networks competing with each other: a generator and a discriminator respectively. While the generator tries to fool the discriminator by generating realistic synthetic images, the discriminator tries to distinguish generated fake images from real ones. Both networks are trained simultaneously until the generator can consistently generate results that the discriminator cannot distinguish. To optimize the architecture of GAN so that it can effectively achieve grayscale remote sensing image colorization, the generator network is used to generate the colorized images from SAR images. The discriminator network discriminates whether the color image data is real or colorized by the generator network. This can be defined by the equation:

$$\begin{aligned} L_{GAN}(G,D) = & \mathop{\mathbb{E}}_{y\sim pdata(y)}[\log(D(y)] \\ & + \mathop{\mathbb{E}}_{x\sim pdata(x)}[\log(1-D(G(x))] \end{aligned} \tag{1}$$

The GAN model tries to solve the minimax problem which is defined as follows:

$$\min_{G}\max_{D}\left(\mathbb{E}_{y\sim p_{data}(y)}[\log(D(y))]+\mathbb{E}_{x\sim p_{x}(x)}[\log(1-D(G(x)))]\right) \tag{2}$$

where G(x) is the output of a generator network for a given x data, D denotes the discriminator network, y is a sample data from a real distribution and x is random noise. We further refers to data distributions as $y \sim pdata(y)$ and $x \sim pdata(x)$.

In the context of Artificial-SAR image colorization, we propose an extended model illustrated in fig 1. which employs two discriminators: an image discriminator Di and a color discriminator Dc. The first one discriminates real images (RGB) from colorized Artificial-SAR images by inspecting their pixel values, while the second discriminates real images from colorized ones by inspecting their color, brightness, and contrast difference. The combination of the two discriminators improves the accuracy of image discrimination. Adversarial training between the generator and multi-discriminators (Image and Color) forces Artificial-SAR images to converge with Optic images in terms of data and color distribution, to finally improve image visual perceptual quality. The architectures of our discriminators are shown in Table 1. Besides, our generator follows the U-Net shape and uses residual connections between deconvolutional layers, to avoid the vanishing gradient problem and preserve image information transfer across generator layers, which help to obtain generated image with fewer speckles. We choose to use the L*a*b color space for our Artificial-SAR colorization task. The reason why, is because L*a*b color space contains a dedicated channel (L) to depict the brightness of the image, and the color information is fully encoded in the remaining two channels (a*b). As a result, we avoid any unexpected variations in both color and brightness through small perturbations in intensity values that can be observed in using RGB color space.

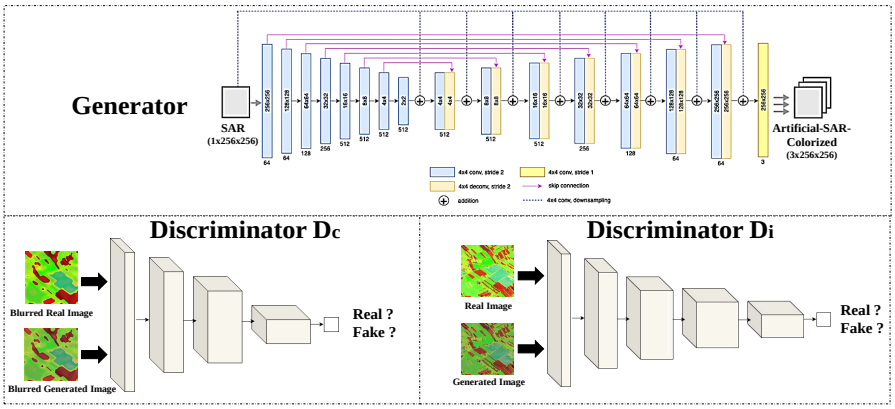

Fig. 1: Architecture of the proposed model, we employ two discriminators, an image discriminator in the pixel domain and a color discriminator. The generator follows the U-Net architecture with residual connections.

Image Discriminator	Color Discriminator
conv(k4,s2,n64),LeakyRelu	conv(k4,s2,n64),LeakyRelu
conv(k4,s2,n128),LeakyRelu,BN	conv(k4,s2,n128),LeakyRelu,BN
conv(k4,s2,n256),LeakyRelu,BN	conv(k4,s2,n256),LeakyRelu,BN
conv(k4,s2,n512),LeakyRelu,BN	conv(k4,s2,n1),LeakyRelu
conv(k4,s2,n1),LeakyRelu	

Table 1: The architecture of our dual image-color discriminators. n, k and s denote respectively the number of Conv filters, the size of the filter and the stride.

2.1 Conventional Loss function

In the context of automatic colorization, based on the previous colorization work [12], we consider a l_1-loss for pixel level. The l_1-loss is minimized by measuring the distance between the generated image $G(x)$ and the real image y. l_1-loss can be written as:

$$L_{l1} = \underset{x,y}{\mathbb{E}} ||y - G(x)||_1 \tag{3}$$

where $G(x)$ and y have the same size.

2.2 Proposed Loss functions

GAN loss $L_{M-Dis}(G, D_i, D_c)$: Based on the above loss functions (Eqs. 1,3), we trained the generator with the two discriminators by introducing a loss function that takes the form:

$$L_{M-Dis}(G, D_i, D_c) = \lambda_i L_{GAN}(G, D_i) + \lambda_c L_{GAN}(G, D_c) \tag{4}$$

where λ_i and λ_c denote a defined regularization factors, $L_{GAN}(G,D_i)$ refers to pixel GAN loss which represents high-frequency details in the pixel domain, $L_{GAN}(G,D_c)$ is the color GAN loss that characterizes color details.

The pixel GAN loss can be defined as follows:

$$L_{GAN}(G,D_i) = \mathop{\mathbb{E}}_{y\sim pdata(y)}[\log(D_i(y)] + \mathop{\mathbb{E}}_{x\sim pdata(x)}[\log(1-D_i(G(x))] \tag{5}$$

To compute the color GAN loss, for the discriminator D_c, we take a blurred image as input and discriminates images by color, brightness and contrast. The blurred image is obtained via blurred convolution B:

$$y_b = y * B,\ G(x)_b = G(x) * B$$

where y_b and $G(x)_b$ are respectively the blurred images of the real images and generated ones. * denotes the convolution operation, and B refers to the blur filter. The weights of the filter B are fitted to the Gaussian distribution, while the size of B is 21×21 with stride 1. B can be computed by the equation bellow:

$$B(x,y) = A\exp(-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{(y-\mu_y)^2}{2\sigma_y^2} \tag{6}$$

with $A = 0.053$, $\mu_{x,y} = 0$ and $\sigma_{x,y} = 3$

We fixed a constant $\sigma_{x,y}$ by visual inspection as the smallest value which ensures that texture and content are dropped as observed in figure 2.

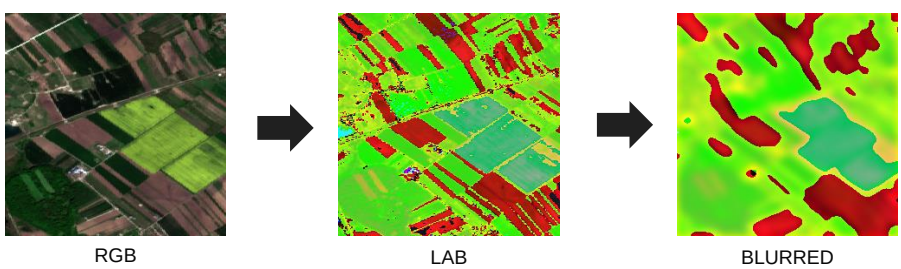

Fig. 2: Results of an RGB image after applying the Gaussian filter B, on its LAB transformation with $\sigma_{x,y} = 3$.

So, our color GAN loss can be defined by:

$$L_{GAN}(G,D_c) = \mathop{\mathbb{E}}_{y\sim pdata(y)}[\log(D_c(y_b] + \mathop{\mathbb{E}}_{x\sim pdata(x)}[\log(1-D_c(G(x)_b)] \tag{7}$$

Color loss L_{color}: Moreover, we suggest non-adversarial color loss to evaluate the difference in brightness, contrast and major colors between images, while eliminating texture and content comparison. It can be written as:

$$L_{color} = ||y_b - G(x)_b||_2^2 \tag{8}$$

where y_b and $G(x)_b$ are the blurred images of y_b and $G(x)_b$, resp.

Full Objective: Our full objective loss can be written by the equation:

$$L_{Multi-GAN}(G, D_i, D_c) = L_{M-Dis}(G, D_i, D_c) + \lambda_{l1} L_{l1} + \lambda_{color} L_{color} \tag{9}$$

where λ_{l1} and λ_{color} are parametric factors. Our goal is to solve the minimax game problem with the value function:

$$G, D_i, D_c = arg \min_{G} \max_{D_i, D_c} L_{Multi-GAN}(G, D_i, D_c) \tag{10}$$

3 Experimental Analysis

In this section, we discuss the performance of our proposed model. The quantitative and qualitative evaluations were performed on the SEN1-2 dataset [13]. SEN1-2 comprises 282 384 pairs of SAR and optical image patches extracted from versatile Sentinel-1 and Sentinel-2 scenes and collected from across the globe and throughout all meteorological seasons.

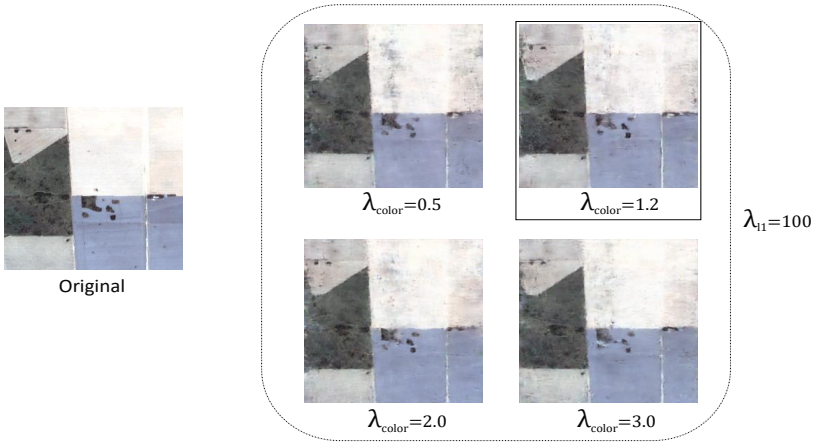

Fig. 3: Figure is better seen zoomed on the digital version of this document. Comparison results with different values of λ_{color}.

3.1 Training details

To train our networks, we first select two pairs of scenes (SAR & Optical) from SEN1-2 dataset to get a total of 2340 pairs images of size (256×256). Adam is selected as our optimizer for generator and discriminators, with an initial learning rate $\alpha_1 = 0.003$ for the generator and $\alpha_2 = 0.0003$ for discriminators. We set the regularization factors $\lambda_i = 1$ and $\lambda_f = 10^{-2}$ (Eq.4). As a result of extensive hyper-parameters optimization, $\lambda_{l1} = 100$ (Eq.9) and $\lambda_{color} = 1.2$ (Eq.8) because its produce stable Artificial-SAR colorization results with less artifacts as shown in figure 3. The number of training epochs is 200 where the batch size is 4. We have implemented our model in Tensorflow.

3.2 Result analysis

First, we have performed an ablation study. We have considered different cases by considering only the image discriminator (D_i), the image discriminator (D_i) + the color discriminator (D_c) and the image discriminator (D_i) + the color discriminator (D_c) + L_{color}. Some comparison results are presented in Fig.4. We can observe that our generator with the three components (D_i, D_c, L_{color}), produces better and stable Artificial-SAR colorization results compared to when we use only one of the component (D_c or L_{color}), which generate sometimes blurred images with artifacts.

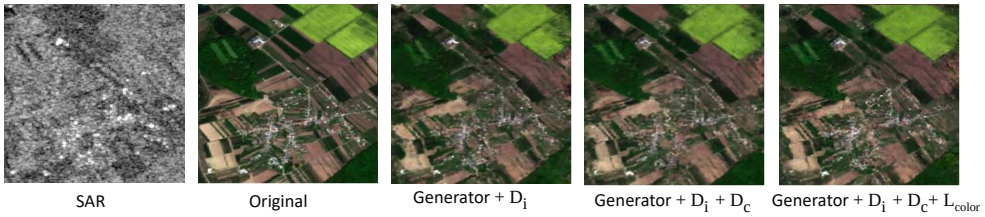

Fig. 4: Figure is better seen zoomed on the digital version of this document. Influence of each component in our proposed model.

We compare our proposed model with state-of-the-art methods Isola *et al.* [7] and Nazeri *et al.* [12]. The experiment results in fig.5 show that Pix2Pix is not suitable for generating Artificial-SAR colorized images. So, we limited our comparison with Nazeri *et al.*. Figure 6 shows that our proposed model outperforms Nazeri *et al.* by producing Artificial-SAR colorized images with less artifacts and realistic color. Moreover, we can observe that the proposed model generate images with well-sharper regions (black rectangles in fig.6), we can explain that by the residual connections that we insert between deconvolutional layers in our generator, which preserve the image structure.

For quantitative comparison by the PSNR, SSIM and color histogram difference by the correlation method (Histcomp), some models can not show significant improvement in PSNR and SSIM or they produce Artificial-color images which do not correspond to the real ones, so, the evaluation metrics could not be calculated. As shown in table 2,

Fig. 5: Colorization result by Pix2Pix approach [7]

the proposed model produces images with good PSNR, SSIM and Histcomp values and ensure a tradeoff with visual results. This is reflected by the generated images having sharper details and more fine-tuned structures.

Method	PSNR	SSIM	Histcomp
Nazeri *et al.* [12]	20.89	0.61	0.90
Proposed Model	**21.74**	**0.65**	**0.98**

Table 2: Quantitative comparison between the proposed model and Nazeri *et al.* [12].

4 Conclusion

In this paper, we proposed a novel Artificial-SAR image colorization GAN network for Sentinel-1 images. The proposed model employs two discriminators for image and color information. Moreover, a non-adversarial color was used to improve the generator output, by evaluating the difference in major colors between the images while eliminating texture and content comparison. Compared with state-of-the-art methods which use a single discriminator, a great reduction in overall color prediction error has been shown in experiments. Such model will serves to maintain visual recognizability at less textured large continuous regions, when it's difficult to be viewed in SAR images. Moreover the proposed model was able to consistently produce pleasing visual colorized images with less artifacts.

References

1. Cao, Y., Zhou, Z., Zhang, W., Yu, Y.: Unsupervised diverse colorization via generative adversarial networks. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases. pp. 151–166. Springer (2017)
2. Cheng, Z., Yang, Q., Sheng, B.: Deep colorization. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 415–423 (2015)

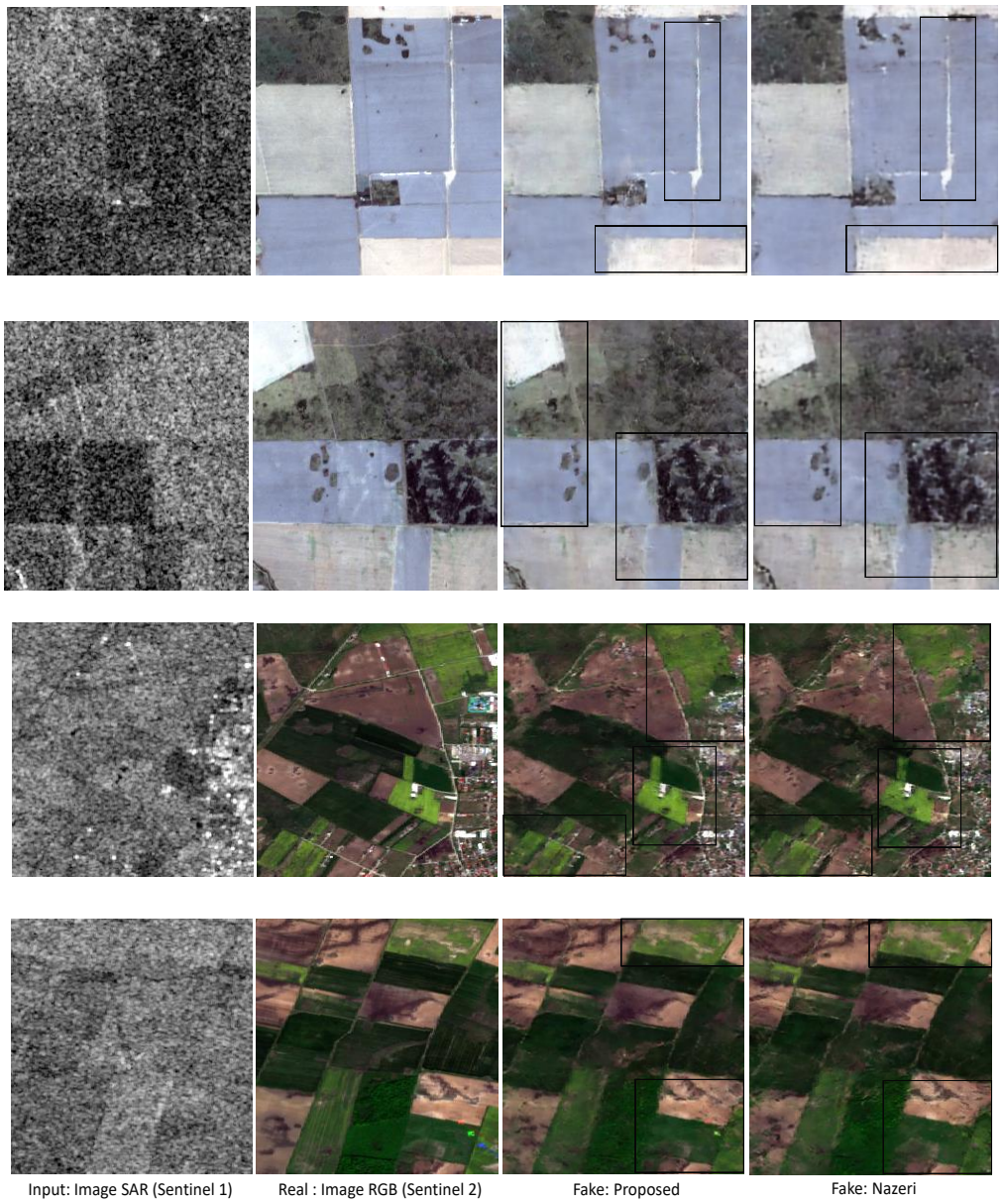

Fig. 6: Figure is better seen zoomed on the digital version of this document. Comparison results between our proposed model and Nazeri *et al.* [12]

3. Chia, A.Y.S., Zhuo, S., Gupta, R.K., Tai, Y.W., Cho, S.Y., Tan, P., Lin, S.: Semantic colorization with internet images. ACM Transactions on Graphics (TOG) **30**(6), 1–8 (2011)
4. Deshpande, A., Rock, J., Forsyth, D.: Learning large-scale automatic image colorization. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 567–575 (2015)
5. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: Advances in Neural Information Processing Systems. pp. 2672–2680 (2014)
6. Huang, Y.C., Tung, Y.S., Chen, J.C., Wang, S.W., Wu, J.L.: An adaptive edge detection based colorization algorithm and its applications. In: Proceedings of the 13th Annual ACM International Conference on Multimedia. pp. 351–354. ACM (2005)
7. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1125–1134 (2017)
8. Larsson, G., Maire, M., Shakhnarovich, G.: Learning representations for automatic colorization. In: European Conference on Computer Vision. pp. 577–593. Springer (2016)
9. Levin, A., Lischinski, D., Weiss, Y.: Colorization using optimization. In: ACM Transactions on Graphics (tog). vol. 23, pp. 689–694. ACM (2004)
10. Liu, X., Wan, L., Qu, Y., Wong, T.T., Lin, S., Leung, C.S., Heng, P.A.: Intrinsic colorization. In: ACM SIGGRAPH Asia 2008 papers, pp. 1–9 (2008)
11. Luan, Q., Wen, F., Cohen-Or, D., Liang, L., Xu, Y.Q., Shum, H.Y.: Natural image colorization. In: Proceedings of the 18th Eurographics Conference on Rendering Techniques. pp. 309–320 (2007)
12. Nazeri, K., Ng, E., Ebrahimi, M.: Image colorization using generative adversarial networks. In: International Conference on Articulated Motion and Deformable Objects. pp. 85–94. Springer (2018)
13. Schmitt, M., Hughes, L., Zhu, X.: The sen1-2 dataset for deep learning in sar-optical data fusion. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences pp. 141–146 (2018)
14. Schmitt, M., Zhu, X.X.: Data fusion and remote sensing: An ever-growing relationship. IEEE Geoscience and Remote Sensing Magazine **4**(4), 6–23 (2016)
15. Suárez, P.L., Sappa, A.D., Vintimilla, B.X.: Infrared image colorization based on a triplet dcgan architecture. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 18–23 (2017)
16. Welsh, T., Ashikhmin, M., Mueller, K.: Transferring color to greyscale images. In: Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques. pp. 277–280 (2002)
17. Xiao, Y., Jiang, A., Liu, C., Wang, M.: Single image colorization via modified cyclegan. In: 2019 IEEE International Conference on Image Processing (ICIP). pp. 3247–3251. IEEE (2019)
18. Zhang, L., Zhang, L., Du, B.: Deep learning for remote sensing data: A technical tutorial on the state of the art. IEEE Geoscience and Remote Sensing Magazine **4**(2), 22–40 (2016)
19. Zhang, R., Isola, P., Efros, A.A.: Colorful image colorization. In: European Conference on Computer Vision. pp. 649–666. Springer (2016)
20. Zhang, R.Y., Zhu, J.Y., Isola, P., Geng, X., Lin, A.S., Yu, T., Efros, A.A.: Real-time user-guided image colorization with learned deep priors. ACM Transactions on Graphics **36**(4), 119 (2017)
21. Zhu, X.X., Tuia, D., Mou, L., Xia, G.S., Zhang, L., Xu, F., Fraundorfer, F.: Deep learning in remote sensing: A comprehensive review and list of resources. IEEE Geoscience and Remote Sensing Magazine **5**(4), 8–36 (2017)