JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Effective Approaches for Auto-Colorization of Grayscale Images Using No GANs

Pavan Kumar Tadiparthi1, Divya Kanchi2, Divi Lakshmi Bhavani3, Thurumella Pratyusha4

Associate Professor1, Student2,3,4

Department of Information Technology

MVGR College of Engineering, India

Abstract: ***In centuries the images or pictures was drawn by manual process with human intervention. However, the technology evolved and photography was first invented and images were captured by low end resolution cameras which were black and white images. However, enhancing the black and white images into color images was an extension research in the field of computer vision and machine learning techniques. Auto-Colorization of a grayscale image has been an imminent topic in the image processing and also an interesting area to explore over the recent years with the importance given to achieve the artifact-free quality. It is being used to improve the visual appealing of images like the grayscale photos, historic illusions, degraded images or videos. In this article, we aim to perform auto-colorization by taking a grayscale image as an input and generate a realistic colorized output image using convolutional neural network (CNN) and generative adversarial network (GAN). Both these proposed models tries to map grayscale image with their respective RGB colour format. The proposed models are trained on the standard dataset Flickr by ourselves with 30k images having 200X200 resolutions. The obtained result is compared using the adam optimizer and root means squared error and is clearly stated that the performance of auto-colorization of the generative adversarial network is better when compared with convolutional neural network model.***

Index Terms: Auto-colorization, Generative adversarial network, generator, discriminator, Root mean squared error, Convolutional neural network

I. INTRODUCTION

In past the black and white images are captured due to the low end resolution cameras. Every image has some internal meaning which was depicted by its visual appearance. In order to enhance the visual appearance of grayscale image by using color photography is to interpret its hidden meaning.

Generally in centuries, paintings were done in the form of cloth painting, wall painting or manuscripts. For example, we see the wall paintings in historical place in all over the world. A person uses paints, photoshop softwares, editing software or other tools for image colorization which doesn't guarantee to give high quality results.

However, the quality of grayscale images can be enhanced by using techniques like auto-colorization which is a process of imposing colors to grayscale images or monochromic videos. We perform colorization operation on the input grayscale image without any human intervention and generate a realistic colorized output image. In other words, adding colors to a grayscale image and generating a plausible colored image.

Figure 1. A sample grayscale image showing a man sitting in the waiting hall.

Over the recent years, colorization has been one of the interesting areas of research and even color photography has now became a part of our daily lifestyle due to which people tend to prefer a realistic colored image over the grayscale image. Grayscale images are a monochromic image that doesn't have any information about the color, shades and is less attractive to the people, whereas colored image has more information stored and is more attractive to the eyes of the viewers.

Despite of manual paintings and photoshops in hand, it is also necessary to ensure that colorization safeguards the richness, meaning and conditions that are to be considered in the grayscale. Grayscale image being a one dimensional are to be mapped to their original color space. The color space here is RGB that represents the image into three-dimensional information named Red, Green and Blue. Initially data pre- processing is performed, where we have better results using LAB color space and for this conversion of images to LAB we use tensor flow and plot LAB color space. This process is useful to perform test on a single image but when we want to perform on a huge dataset, we can use various model based approaches that were time saving too on the other hand. In the literature, there are several model based approaches and among those, we will be outlining effective models like the convolutional neural network(CNN) using Keras and Generative Adversarial Network(GAN).

After the generation of colorized output image using CNN and GAN, the resultant performance of both the models are computed using performance metric like root mean squared error (rmse), which predicts the similarities with the expected output image. It is observed that GAN gives better results and quality than CNN. The concept of colorization has a vivid real time application like restoration of old images, enhancement of degraded images or videos, medical research, object detection, image classification.

The rest of the article is organized as follows: the literature review is presented in section-II.In section III discussed about our proposed methodology. Experimentation was described in section –IV .The results of the work presented in section V of this article. Conclusion work presented in section VI.

II. RELATED WORK

The process of colorization is a cryptic task since it doesn't have specific solution to rely on. Various approaches to the task of colorization can be divided into three groups: scribble-based, example-based, and learning-based.

In the scribble-based colorization approach, colors are interpolated in the grayscale image based on the scribbles given by an artist. Levin et al [8] proposed the principle that states that the neighbouring pixels with close color intensities must have near colors and applied this method for certain stationary images and videos. However this approach has a few problems on the image like the color blending effect and performance was based on the human skills.

In the example-based colorization approach, color translation is carried out based on the reference color image information by using the color correlation method proposed by Reinhard et al [11]. But this approach gives better results only when an input given to the algorithm is a proper color image.

The learning-based colorization approach in [14-17] applied to different machine learning models and algorithms so as to make the model learn complex patterns and extract features present in an image by using a large dataset. This approach is automatic and also overcomes the drawback of manual coloring methodologies. Over the past few years, CNN based approaches were used for the task of image colorization. In [16-18], authors detailed the concurrent extraction of both global and local features which are further fused together. CNNs became a vital player to handle tasks like the restoration, image-labelling apart from image colorization. But there were drawbacks in CNN like the image quality and object identification was not up to the mark for a few samples which the GANS were able to overcome. GANs were introduced in 2014 by Ian, et al and have been proven to yield the most clear-sighted outcomes for the image colorization problem. GANS has the capability of generating new data samples within itself and improve computer vision. GANS has the discriminator component that distinguish the images as fake or real.

III. PROPOSED METHODOLOGY

Intially we implement CNN, Autoencoder and CNN using Transfer Learning, alongside the trainable parameters. Followed by the proposed method implemented using Generative Adversarial Network (GAN). Both resultant images are compared with the ground truth image.

3.1 DATASET CONSIDERED

For our experimentation, we considered a Flickr 30k dataset for training and validation.

The Flickr dataset consists of colored images of various categories in order to train the model to predict the colors for different objects and shapes. It is a standard dataset that contains 30 thousand images in 200x200 resolution and perform training on those images. We also have a Flickr 8k dataset, subset of Flickr 30k dataset that contains 8 thousand images in 200x200 resolution. This was used for the initial training to understand the performance and accuracy.

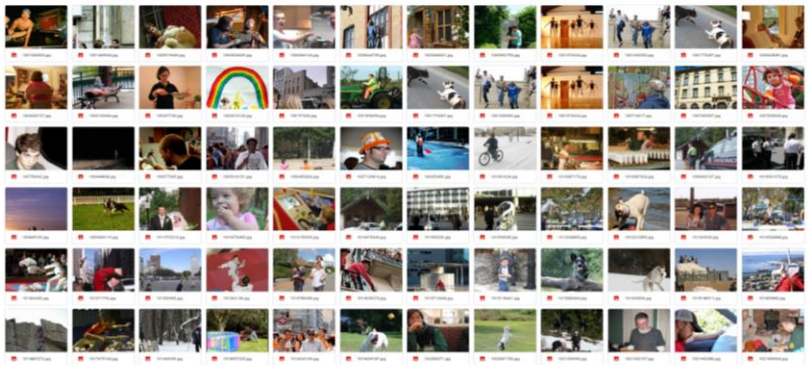

Figure 2. Some sample images taken from the flickr dataset

3.2 PRE-PROCESSING

The Flickr dataset are considered as an input values and are being pre-processed, including the conversion of CIELAB color space. For training and a validation set, we convert those rgb images to grayscale images.

After the pre-processing of data, implementation of the proposed model is preceded as follows:

3.3 CONVOLUTIONAL NEURAL NETWORK

The convolutional neural network (CNN) is a model that has three sections namely encoder, global feature extractor, decoder. It is built of 8 blocks of convolutional layers, without any pooling layers. At the end of each block, batch normalization is used as the only regularization. This CNN model does not have any activation layer towards the edge of the network. It uses Adam optimizer. CNN succeeded in their ability to learn and discern colors, patterns from the given image and associate them with the classes of objects. CNN is very useful as it combines the local features and global features together and takes less time duration, when compared to the manual paintings.

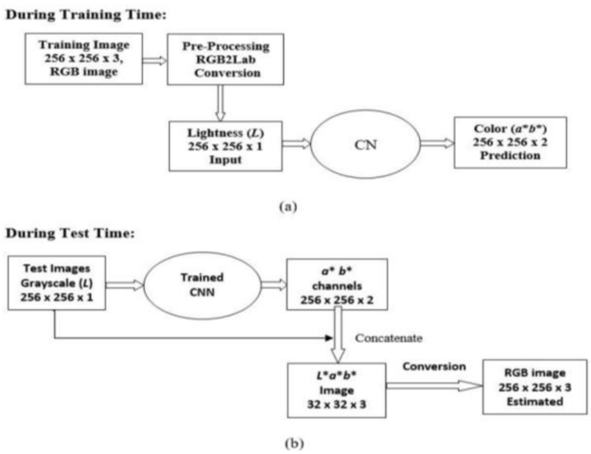

Figure 3.System Flow Diagram for Auto-colorization Model; (a) during training and (b) during testing.

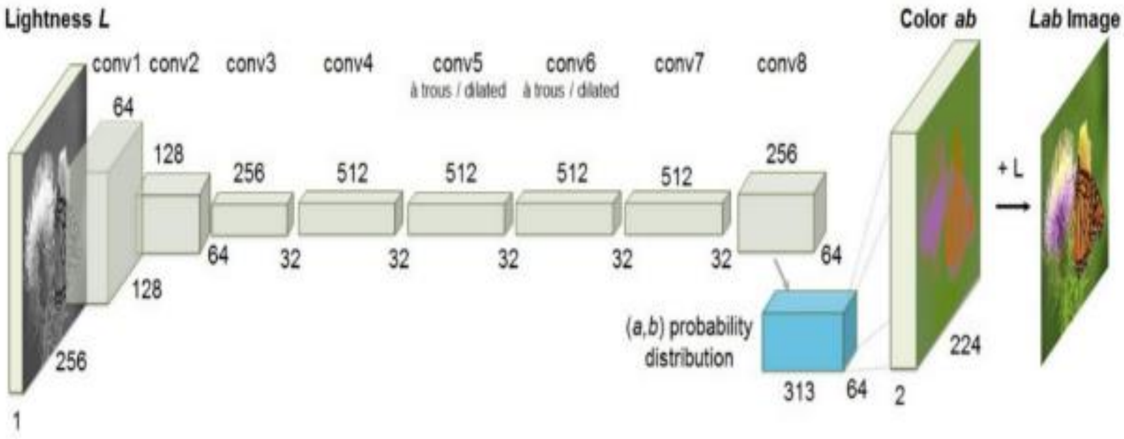

Figure 4.Overview of the CNN model architecture for auto-colorization of grayscale images.

3.4 AUTOENCODER MODEL

The Auto encoders is a type of neural network which reduces the dimensions of the input and the loss function. Auto encoders are built from two components named encoder and decoder. This autoencoder model consists of four convolutional layers namely the central block of layers, Flatten, two Dense and one Reshape layers that act as an encoder part. This model also have four convolutional transpose layers that act as decoder part. The encoder transforms the input data into a lower-dimensional representation while the decoder recovers the input data from low dimensional representation. CNN can be used in the encoding and decoding parts of the auto encoders for image colorization.

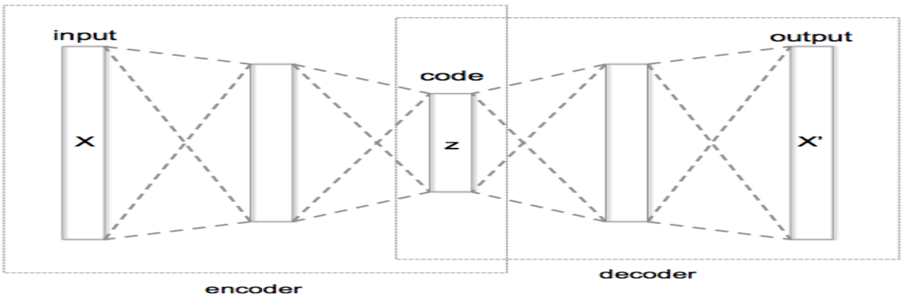

Figure 5.Overview of the Autoencoder model architecture for auto-colorization of grayscale images.

3.5 COMBINATION OF CNN AND XCEPTION PRETRAINED MODEL

The another approach is implementation of CNN using transfer learning which is referred as Xception pretrained model. In our auto-colorization we use Xception, which is an extreme inception. This model acts as encoder for the input and there are five convolutional layers that acts as decoder representing the expected output. In the training stage, Adam optimizer and root mean square error are used.

3.6 GENERATIVE ADVERSARIAL NETWORK

GAN are one of the recent innovation in the machine learning. GAN learns from a set of training data and further generates new data that is similar to the training data. NoGAN is subset of GAN training that is developed to solve some key problems. This provides the benefits of reducing time while spending direct GAN training and also eliminates glitches if any. In implementation of NoGAN, first generator is trained by itself in a conventional way with just the feature loss. After generation of images, discriminator differentiate the outputs and the real images as a binary classifier. Finally based on the render factor we sharpen the background.

IV. EXPERIMENTATION

The overall experimentation is carried out by the proposed model is outlined in section III of the article using the benchmark dataset outlined in the section 3.1 of the article is implemented in Google Collab with Python programming language and Tensor Flow as backend. The obtained results are evaluated with performance metric root mean squared error for similarity measures.

For suppose let us consider that $\hat{\theta}$ be an estimator with respect to a given estimated parameter θ, the Root Mean Square Error is formulated as

$$MSE = (1/MN) \sum_{n=1}^{M} \sum_{m=1}^{N} [\hat{g}(n,m) - g(n,m)]^2 \quad - (1)$$

$$RMSE(\hat{\theta}) = \sqrt{MSE(\hat{\theta})} \quad - (2)$$

V.RESULTS

5.1. Training

The models are trained approximately with 30 K images in the dataset, out of which 80% were used for training purpose and the remaining 20% were used for the testing purpose. This split is done based on the total number of samples in the dataset as well as actual model that is being trained. In addition, testing of colorization is carried out using root mean squared error. The best results were reflected with batch size 96 and split ratio 80:20, 250 epochs. Increasing the epochs from 250 to 500 has improvised the quality results and realistic outcome. Coding was carried out using Python 3.7.

5.2 Colorization Results on Test Images

After the training phase is completed, images from the test and validation dataset were loaded to the model. The obtained results were good for few objects like human being, animals, trees and weren't good for the light backgrounds with CNN model. However, with GAN the results were plausibly good and showed unbelievable outcomes with natural elements like trees, background.

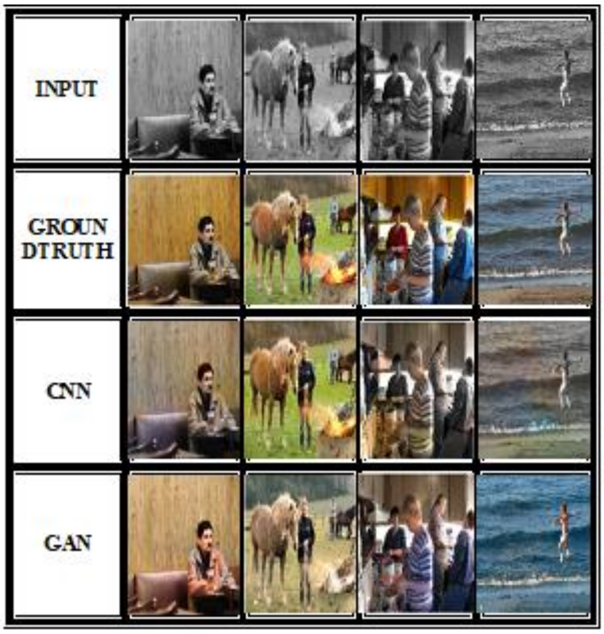

Figure 6. Results obtained from the model on a part of dataset as well as custom images. The 1st row represents the input grayscale images, the 2nd row represents the ground truth image that is the actual outcome, the 3rd row represents the predicted outcome with CNN model and the 4th row represents the GAN model.

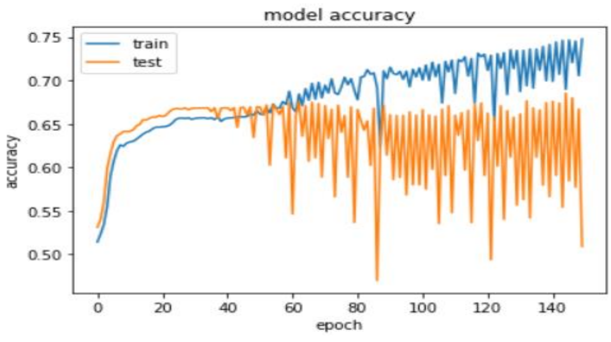

Figure 7. Graph showing the model accuracy vs. number of epochs (dataset: 30k, Number of Epochs= 500, Batch size: 96)

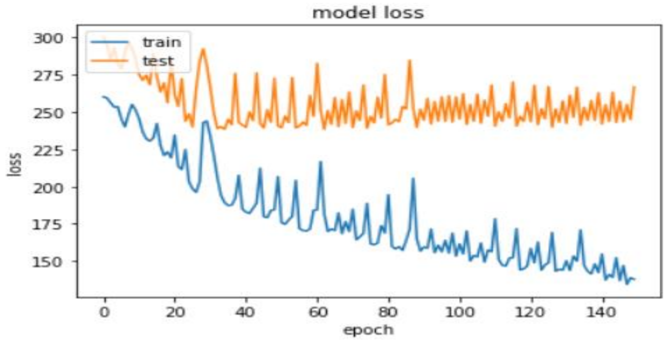

Figure 8. Graph showing the model loss vs. number of epochs (dataset: 30k, Number of Epochs= 500, Batch size: 96)

VI. CONCLUSION

In this article, we are focused on the colorization of grayscale image which generate a realistic color image as outcome. The model was trained on a standard flickr dataset of 30k images. We implemented this problem in two sections where the first section had CNN, Autoencoder and CNN using Transfer Learning model and the second section has GAN. The performance of the model from the two sections is compared for evaluating with the traditional methods such as root mean squared error (RMSE). The results are discussed in section VI. Hence, our proposed model for grayscale colorization gives better results with GAN compared with the first section CNN. The proposed model successfully color the image components such as sky, trees. The proposed model is a generalized approach can be experimented further with various datasets for various applications that involve image colorization.

REFERENCES

[1] Martínez, B.; Casas, S.; Vidal-González, M.; Vera, L.; García-Pereira, I. TinajAR (2018): An edutainment augmented reality mirror for the dissemination and reinterpretation of cultural heritage. Multimodal Technol. Interact. 2, 33. [CrossRef]

[2] Portalés, C.; Rodrigues, J.M.; Rodrigues Gonçalves, A.; Alba, E.; Sebastián, J. (2020)Digital Cultural Heritage. Available onlone: https://www.mdpi.com/2414-4088/2/3/58.

[3] Casas, S.; Gimeno, J.; Casanova-Salas, P.; Riera, J.V.; Portalés, C. (2020)Virtual and Augmented Reality for the Visualization of Summarized Information in Smart Cities: A Use Case for the City of Dubai. In Smart Systems Design, Applications, and Challenges; IGI Global: Pennsylvania, PA, USA; pp. 299–325.

[4] Guidi, G.; Beraldin, J.A.; Atzeni, C. (2004)High-accuracy 3D modeling of cultural heritage: The digitizing of Donatello's "Maddalena". IEEE Trans. Image Process. 13, 370–380. [CrossRef] [PubMed]

[5] Andreetto, M.; Brusco, N.; Cortelazzo, G.M. (2004) Automatic 3D modeling of textured cultural heritage objects. IEEE Trans. Image Process. 13, 354–369. [CrossRef] [PubMed]

[6] Elazab, N.; Soliman, H.; El-Sappagh, S.; Islam, S.; Elmogy, M. (2020) Objective Diagnosis for Histopathological Images Based on Machine Learning Techniques: Classical Approaches and New Trends. Mathematics, 8, 1863. [CrossRef]

[7] Kim, H.I.; Yoo, S.B. (2020) Trends in Super-High-Definition Imaging Techniques Based on Deep Neural Networks. Mathematics 8, 1907. [CrossRef]

[8] Levin, A.; Lischinski, D.; Weiss, Y. (2004)Colorization using optimization. In ACM SIGGRAPH 2004 Papers; ACM: Los Angeles, CA, USA, pp. 689–694.

[9] Huang, Y.C.; Tung, Y.S.; Chen, J.C.; Wang, S.W.; Wu, J.L. (2005) An adaptive edge detection based colorization algorithm and its applications. In Proceedings of the 13th Annual ACM International Conference on Multimedia, Hilton, Singapore, 6–11 November, pp. 351–354. Mathematics 2020, 8, 2258 16 of 17

[10] Yatziv, L.; Sapiro, G. Fast (2006) image and video colorization using chrominance blending. IEEE Trans. Image Process. 15, 1120–1129. [CrossRef] [PubMed]

[11] Reinhard, E.; Adhikhmin, M.; Gooch, B.; Shirley, P. (2001)Color transfer between images. IEEE Comput. Graph. Appl. 21, 34–41. [CrossRef]

[12] Welsh, T.; Ashikhmin, M.; Mueller, K. (2002) Transferring color to greyscale images. In Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques, San Antonio, TX, USA, 23–26 July, pp. 277–280.

[13] Ironi, R.; Cohen-Or, D.; Lischinski, D. (2020) Colorization by Example. Rendering Techniques. Available online: https://www.cs.tau.ac.il/~dcor/onlinepapers/papers/colorization05.pdf

[14] Bugeau, A.; Ta, V.T. (2012) Patch-based image colorization. In Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012), Tsukuba Science City, Japan, 11–15 November, pp. 3058–3061.

[15] Zeng, D.; Dai, Y.; Li, F.; Wang, J.; Sangaiah, A.K. (2019) Aspect based sentiment analysis by a linguistically regularized CNN with gated mechanism. J. Intell. Fuzzy Syst. 36, 3971–3980. [CrossRef]

[16] Zhang, R.; Isola, P.; Efros, A. (2016) Colorful image colorization. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 16–18 October; pp. 649–666.

[17] Luo, Y.; Qin, J.; Xiang, X.; Tan, Y.; Liu, Q.; Xiang, L. (2020)Coverless real-time image information hiding based on image block matching and dense convolutional network. J. Real-Time Image Process. 17, 125–135. [CrossRef]

[18] Iizuka, S.; Simo-Serra, E.; Ishikawa, H. (2016) Let there be color! Joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification. Acm Trans. Graph. (ToG) 35, 1–11. [CrossRef]

[19] Agrawal, M.; Sawhney, K. (2020) Exploring Convolutional Neural Networks for Automatic Image Colorization. Available online: http://cs231n.stanford.edu/reports/2017/pdfs/409.pdf

[20] Karpathy, A. (2020) Cs231n convolutional neural networks for visual recognition. Available online: http://cs231n.stanford.edu/2016/

[21] Vu, M.T.; Beurton-Aimar, M.; Le, V.L. (2018) Heritage Image Classification by Convolution Neural Networks. In Proceedings of the 2018 1st International Conference on Multimedia Analysis and Pattern Recognition (MAPR), Ho Chi Minh City, Vietnam, 9–10 May; pp. 1–6.

[22] Varga, D.; Szabo, C.A.; Sziranyi, T. (2017) Automatic cartoon colorization based on convolutional neural network. In Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing, Florence, Italy, June; pp. 1–6.

[23] Baldassarre, F.; Morín, D.G.; Rodés-Guirao, L. Deep koalarization: (2017) Image Colorization Using Cnns and Inception-Resnet-v2. arXiv, arXiv:1712.03400.

[24] Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. (2016) Inception-v4, Inception-Resnet and the Impact of Residual Connections on Learning. arXiv, arXiv:1602.07261.

[25] Kingma, D.P.; Ba, J. Adam: (2014) A method for stochastic optimization. arXiv, arXiv:1412.6980.

[26] Sara, U.; Akter, M.; Uddin, M.S. (2019) Image quality assessment through FSIM, SSIM, MSE and PSNR—A comparative study. J. Comput. Commun. 7, 8–18. [CrossRef]

[27] Liu, D.; Jiang, Y.; Pei, M.; Liu, S. (2018) Emotional image color transfer via deep learning. Pattern Recognit. Lett. 110, 16–22. [CrossRef]

[28] Gulli, A.; Pal, S. (2019)Deep Learning with Keras; Packt Publishing Ltd.: Birmingham, UK, 2019. 29. Isogawa, M.; Mikami, D.; Takahashi, K.; Kimata, H. Image quality assessment for inpainted images via learning to rank. Multimed. Tools Appl. 78, 1399–1418. [CrossRef]

[29] Nazeri, K.; Ng, E.; Ebrahimi, M. (2018) Image colorization using generative adversarial networks. In Proceedings of the International Conference on Articulated Motion and Deformable Objects, Palma de Mallorca, Spain, 12–13 July; pp. 85–94.

[30] Kiani, L.; Saeed, M.; Nezamabadi-pour, H. (2020) Image Colorization Using Generative Adversarial Networks and Transfer Learning. In Proceedings of the 2020 International Conference on Machine Vision and Image Processing (MVIP), Tehran, Iran, 18–20 February; pp. 1–6.