NTIRE 2019 Challenge on Image Colorization: Report

Shuhang Gu Radu Timofte Richard Zhang Maitreya Suin Kuldeep Purohit
A. N. Rajagopalan Athi Narayanan S. Jameer Babu Pinjari Zhiwei Xiong Zhan Shi
Chang Chen Dong Liu Manoj Sharma Megh Makwana Anuj Badhwar
Ajay Pratap Singh Avinash Upadhyay Akkshita Trivedi Anil Saini
Santanu Chaudhury Prasen Kumar Sharma Priyankar Jain Arijit Sur

Gökhan Özbulak

Abstract

This paper reviews the NTIRE challenge on image colorization (estimating color information from the corresponding gray image) with focus on proposed solutions and results. It is the first challenge of its kind. The challenge had 2 tracks. Track 1 takes a single gray image as input. In Track 2, in addition to the gray input image, some color seeds (randomly samples from the latent color image) are also provided for guiding the colorization process. The operators were learnable through provided pairs of gray and color training images. The tracks had 188 registered participants, and 8 teams competed in the final testing phase.

1. Introduction

Image colorization refers to the problem of estimating the color information from a gray scale image. It provides a practical solution to enhance old pictures as well as expressing artistic creativity.

Currently, two broad approaches to image colorization exist: user-guided colorization and data-driven automatic colorization image. The first category of approaches aims to propagate the color information (color strokes) provided by users and generate a visually plausible color image. Early colorization works follow the seminal work of Levin *et al*. [18] and formulate the guided colorization task as an optimization problem. Recently, deep learning based approach [33] has also been exploited for solving the guided colorization task. Another categories of works, which is more challenging, is the automatic colorization method. As there is not any color information of the scene, automatic colorization methods need to mine the semantics of the scene to help the colorization process. Different deep neural networks [17, 32, 5, 7] have been proposed to learn image priors from large scale training datasets.

The NTIRE 2019 colorization challenge is the first colorization challenge in the literature. The challenge consider both the two types of colorization settings, *i.e*. the guided colorization setting and the guidance free colorization setting. The NTIRE colorization challenge builds upon the previous NTIRE challenges, and we use the DIV2K [1] dataset as the official training and validation datasets. In the testing phase, another 20 images are utilized as testing set, the color versions of the testing images are not public available. We expect the challenge helps to benchmark the colorization research.

2. NTIRE 2019 Image Colorization Challenge

The objectives of the NTIRE 2019 challenge on image colorization are: (i) to gauge and push the state-of-the-art in image colorization; and (ii) to compare different solutions.

2.1. DIV2K Dataset

DIV2K Dataset [1] employed by NTIRE 2017 and NTIRE 2018 image super resolution challenges [27, 28] is also used as for training in our challenge. DIV2K has 1000 DIVerse 2K resolution RGB images with 800 for training, 100 for validation and 100 for testing purposes. The manually collected high quality images are diverse in contents. As the low-resolution color version of DIV2K testing set is available online, we take the training and validation set of DIV2K as the training and validation set used in the challenge, and prepare an additional 20 images as the testing set of the challenge.

S. Gu, R. Timofte and R. Zhang are the NTIRE 2019 challenge organizers, while the other authors participated in the challenge.
Appendix A contains the authors' teams and affiliations.
NTIRE webpage: http://www.vision.ee.ethz.ch/ntire19/
NTIRE workshop is in conjunction with CVPR 2019:
http://cvpr2019.thecvf.com/

Team	Track 1 No guidance PSNR	SSIM	Perceptual rank	Track 2 With guidance PSNR	SSIM	Perceptual rank
IPCV_IITM	22.1232	0.9406	1	22.8605	0.9443	1
Athi	20.8710	0.9229	2	20.6470	0.7997	2
VIDAR	22.1949	0.9419	4	23.2707	0.9461	3
Team_India	17.9643	0.8472	3			
pksvision_mm	21.2248	0.9279	5			
ITU-GO	21.0773	0.8526	6			

Table 1. NTIRE 2019 Colorization Challenge results and final rankings. IPCV_IITM team is the winner of the challenge followed by Athi and VIDAR teams.

Team	Toolbox	GPU Device	Running time [s]	Extra Data
IPCV_IITM	Pytorch	NVIDIA Titan X	3	-
Athi	Matlab	-	1.46	-
VIDAR	Pytorch	NVIDA 1080Ti	0.4	-
Team_India	Tensorflow	Nvidia DGX1	0.025	-
pksvision_mm	Pytorch	Nvidia DGX	0.72	-
ITU-GO	Keras	Nvidia Titan X	7.00	ILSVRC 2012

Table 2. NTIRE 2019 methods implementation details.

2.2. Tracks

The NTIRE 2019 challenge on image colorization had two competition tracks. Access to data and submission of colorization results required registration on Codalab competition track.

Track 1: Image Colorization without Guidance uses the gray version of color images as input, and aims to estimate the color information of the scene without any extra information. For the RGB to grayscale color space we use the standard commonly used linear transformation:

$$I = 0.2989 \cdot R + 0.5870 \cdot G + 0.1140 \cdot B, \quad (1)$$

where I are the grayscale values while R, G, B are the red, green, and blue channel values, respectively, of the original RGB color image.

Track 2: Image Colorization with Guidance provides, in addition to the input gray image from Track 1, some guiding color seeds as input. For each input image, we randomly sample between 5 and 15 pixels and provide the color information of these pixels along with their (x, y) image coordinates to help the colorization process.

2.3. Challenge phases

(1) Development phase: the participants got pairs of gray and color train images and the gray validation images of the DIV2K dataset; an online validation server with a leaderboard provided immediate feedback for the uploaded colorization results to the gray validation images.

(2) Testing phase: the participants got test gray images and were required to submit colorization results, code, and a factsheet for their method. After the end of the challenge the final results were released to the participants.

https://competitions.codalab.org

2.4. Evaluation protocol

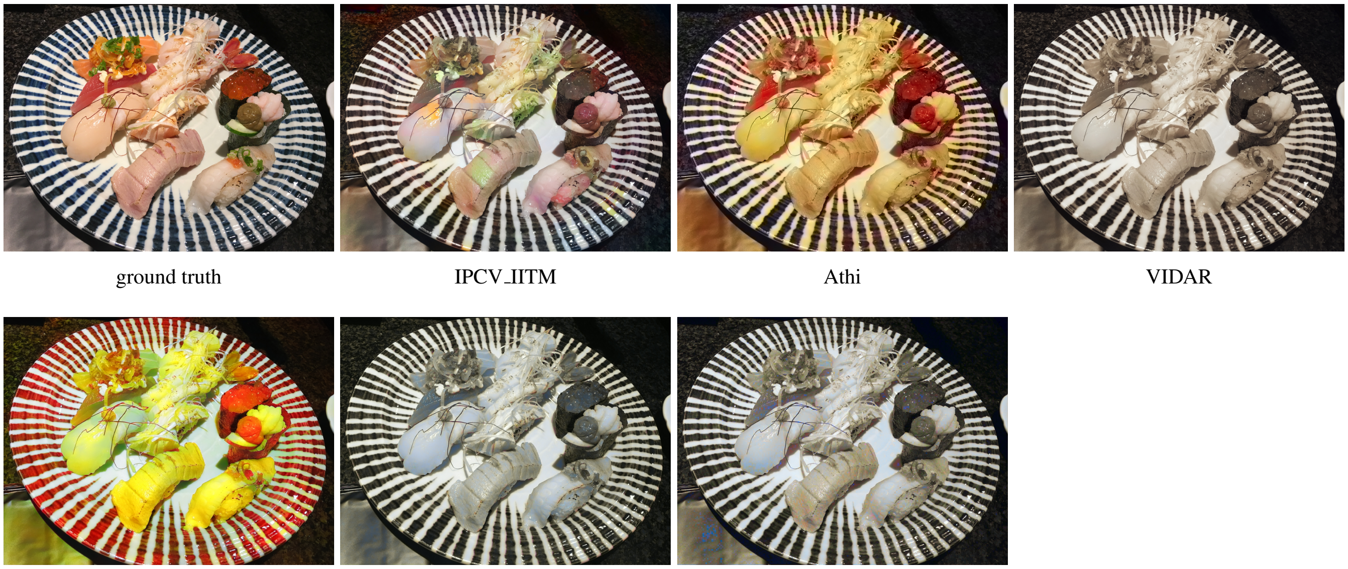
The quantitative measures are Peak Signal-to-Noise Ratio (PSNR) measured in deciBels [dB] and the Structural Similarity index (SSIM) [29], both full-reference measures computed between the colorization result and the ground truth (GT) color image. We report averages over sets of images. As we found both the PSNR and SSIM index can not reflect the visual quality of the colorization performance, we also report the rank of subjective visual quality in this report. The subjective comparison is conducted in a non-reference manner, three people who do not know the ground truth color image compare the colorization results, and we rank the colorization results based on the comparison results.

3. Challenge Results

Track 1 and Track 2 got 81 and 107 registered participants, respectively; and 6/3 teams entered in the final phase and submitted their results, codes/executables, and factsheets. Table 1 reports the final test results and rankings of the challenge, while in Table 2 2 the self-reported runtimes and major details are provided. The methods are briefly described in section 4 and the team members are listed in Appendix A.

3.1. Architectures and main ideas

All the proposed methods in the challenge are deep learning based methods, with the notable exception of the method adopted by Athi team in the guided colorization



ground truth IPCV_IITM Athi VIDAR

Team_India pksvision_mm ITU-GO

Figure 1. Visual examples of the colorization results (Track 1: without guidance) by different methods.

ground truth IPCV_IITM Athi VIDAR

Figure 2. Visual examples of the colorization results (Track 2: with guidance) by different methods.

track. To enlarge the receptive field, most of the methods have adopted an encoder-decoder or the structure based on U-net [21]. And the deep residual net (ResNet) architecture [9] and the dense net (DenseNet) architecture [10] are the basis for most of the proposed methods.

3.2. Restoration fidelity

Generating high quality colorization estimation is still a very challenging task. For both the non-guidance track and guided track, all the methods did not achieve estimations with high PSNR index. The best PSNR indexes on the testing dataset are 22.19dB and 23.27dB for the no guidance and guided tracks, respectively. The VIDAR team obtains the highest PSNR and SSIM scores on both tracks. In order to achieve high fidelity estimation they adopted only the RMSE loss in the training phase. We note that the guiding color pixels from Track 2 helped VIDAR to improve ~ 1.08dB PSNR.

3.3. Perceptual quality

Some visual examples of the colorization results provided by different teams can be found in Figure 1 for the Track 1 without guidance and in Figure 2 for Track 2 with guidance. As can be found in the figures, although team VIDAR has achieved best PSNR and SSIM indexes, they actually did not generate much color information and the estimation still looks greyish. In order to improve the perceptual quality, some teams have also adopted the perceptual loss [14, 11] or the Generative Adversarial Network (GAN) [6] loss in the training phase.

IPCV_IITM team provides the best perceptual quality of the results and is the winner of the NTIRE 2019 challenge on image colorization, while Athi and VIDAR teams comes on the next positions with consistent performance on both tracks.

3.4. Conclusions

The NTIRE 2019 image colorization challenge was the first of its kind. Despite the large number of registered participants only a small number of teams competed in the final

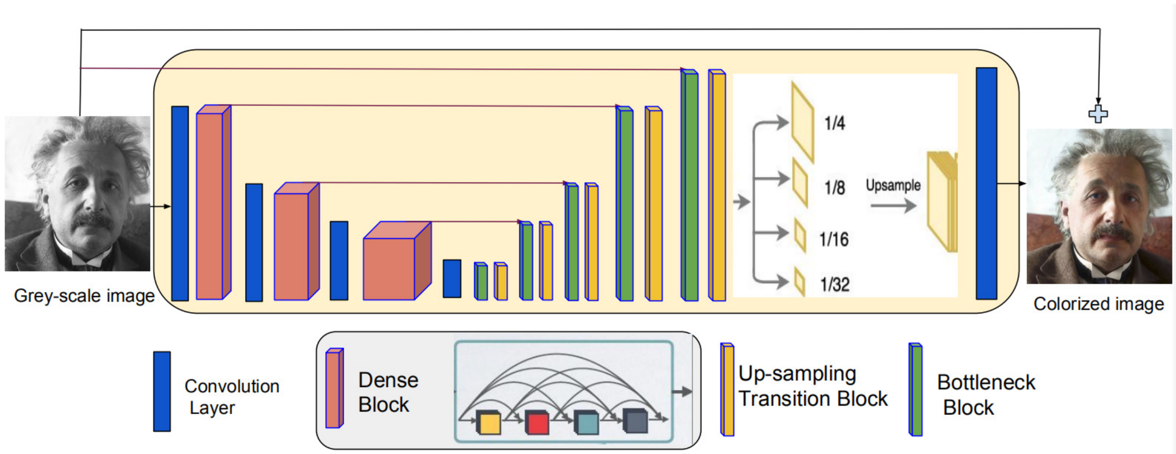


Figure 3. IPCV_IITM's network architecture.

test phase of the challenge. This and the perceptual quality achieved by the proposed solutions indicates the difficulty of the image colorization task and that there is plenty of room for further improvements. The proposed solutions in the challenge were rather diverse, including also a non-deep learning solution that came 2^{nd} in the final ranking (Athi team). As expected the use of extra information under the form of several guiding color pixels helped into stabilizing and improving the colorization result. The deployment of deep learning solutions validated on other image to image mapping including restoration and enhancement tasks [13, 35, 11, 28, 4, 2, 3] does not guarantee good perceptual results on the image colorization task. The image colorization challenge is a step forward into benchmarking and further research is necessary for defining better losses for this task and also for designing solutions capable to see large contexts and infer plausible colors.

4. Challenge Methods and Teams

4.1. IPCV_IITM

IPCV_IITM proposed a deep dense-residual encoder-decoder structure with multi-level pyramid pooling module for estimating the color image from the gray input (see Figure 3). The proposed encoder is made of densely connected modules. It helps to address the issue of vanishing gradients and feature propagation while substantially reducing the model complexity. The encoder has a similar structure to the network mentioned in [31]. The first, second, and the third dense-blocks contain 12, 16 and 24 dense units, respectively. Specifically, each dense unit consists of batch norm, ReLU, conv1×1 (with out_channels=$4\times$GR) followed by batch norm, ReLU, conv3×3 (with out_channels=GR). The growth rate (GR) is set to 32 in our network. The weights of these blocks are initialized using initial layers of pre-trained DenseNet-121 network [10]. Each layer in a block receives feature maps from all earlier layers, which strengthens the information flow during forward and backward pass making training deeper networks easier. The Decoder accepts the features estimated by the encoder at various levels and processes them using bottleneck blocks before increasing their spatial resolution through bilinear up-sampling and convolution. Bottleneck blocks have a similar structure to the dense unit in the encoder. The intermediate features with higher spatial resolution in the decoder are concatenated with the corresponding-sized encoder features. Finally, the decoder output is enhanced through multi-scale context aggregation through pooling and upsampling at 4 scales, before being fed to the final layer. The final output is the residual between the ground-truth colored image and the input image. The colorization network is first trained using the L1 loss, then fine-tuned with MSE loss. The authors have also tried to utilize custom losses defined over other color domains or adversarial loss, but these losses did not lead to stable training of the proposed model.

4.2. Athi

Athi team directly adopts the colorization network proposed in [33] to attend the no-guidance track.

While for the guided track, Athi team firstly creates a dummy image with colors from the guiding pixels and transfer the dummy image to the YCbCr space. For the input gray image, they assigned the gray values to R, G and B planes and converted from RGB to YCbCr color space. Then in the obtained Y channel of the input gray image, they check for the closest value in the Y channel of dummy image. And assign the corresponding Cb and Cr channel values to the output image, without changing the Y channel value. This approach does not achieve RGB reconstruction results with high PSNR. But, in our non-reference subjective evaluation test, the results generated by Athi outperform in perceptual quality the results generated by some other algorithms which have much higher PSNR index.

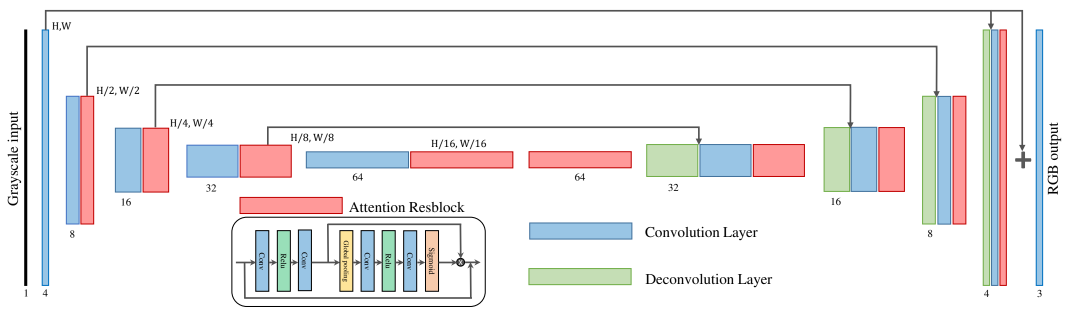


Figure 4. VIDAR's network architecture.

4.3. VIDAR

In order to achieve the best PSNR result, VIDAR team utilizes a regression method to colorize the grey images. VIDAR proposes a U-Net [21] like network using the residual channel attention block [34, 25] to reconstruct color images from grey images. An illustration of the network adopted by VIDAR is shown in Figure 4.2. For Track 2, the guided pixels are first reorganized to a grid of the original image, where the other pixels are set to zeros. VIDAR also utilizes a mask of which values are ones at the position of guided pixel pixels and the others are set to zeros. Then, the guided pixel image and mask are concatenated to form the input of the network besides the grey image. Before taking into the network, they are processed by a 9×9 Gaussian filter. In the network, the guided pixel image and mask first go through a convolution layer, and then concatenate the feature map of the grey image as a new input taken into the remaining network.

4.4. Team_India

Team_India proposes a framework for colorization of a gray scale image using a Self-Attention based progressive GAN [30]. The framework (see Figure 5) consists of state-of-the-art techniques for stabilization and better generalization for image colorization; (i) Self-Attention mechanism (ii) Progressive Growing of input channel (iii) Spectral Normalization for both generator and discriminator (iv) Higher critic to generator learning rate in our case 5:1 worked most efficiently (v) Hinge based adversarial loss. The generator is inspired by the U-Net [21] architecture with its encoder as a pretrained ResNet-34 [9] following a 34-layered decoder; the pre-trained encoder is used as a feature extractor. The output of the U-Net is then passsed to a pre-trained VGG-16 [26] to calculate perceptual/content loss which is then backpropagated to the generator. The weights are initialized by spectral normalization in both the generator and discriminator, self attention is used in between each convolution layer in both the generator and discriminator. Then discriminator uses the hinge version of the adversarial loss which is backpropagated to the generator as well as the discriminator. They have used batch normalization in between layers, and without any activation function in the generator; whereas in the discriminator they are using dropout with 0.2 value, as well as they are using LeakyReLU as the activation function with alpha = 0.2. The training procedure is inspired by Progressive Growing of GAN [15] although they do not progressively grow our model, they are just growing the input size in the training procedure by starting the training at a 64×64 down sampled image and growing it from 64, 96, 128, 160 progressively to 192×192 sample. The final model makes use of the $192\times$ training regime. They have used the ADAM optimizer [16] with hyperparameters of $\beta_1 = 0.0$ $\beta_2 = 0.9$, and a cycle learning rate finder for learning rate. The solution is described in detail in [24].

4.5. pksvision_mm

pksvision_mm team proposed a method, shown in Figure 6, which is based on the conditional Generative Adversarial Network (cGAN) framework [19, 13]. Specifically, the network takes a gray scale image as an input and predict the its corresponding chrominance components (Cb, Cr). The encoder sub-network comprises of six convolutional layers where each layer consists of 64 convolution kernels except the penultimate and last layer which consists of 32, 2 kernels respectively. The decoder sub-network consists of six transpose convolutional layers. Each filter in the generator model has a spatial dimension of 3×3 and stride of 1. Each layer in encoder and decoder sub-networks is followed by the Parametric ReLU (PReLU) [8] and Batch Normalization (BN) [12]. Residual skip connections are given between layer 2, 4 of encoder to 2, 4 of decoder, respectively, in-order to maintain the spatial information and also to reduce the effect of gray scale features if symmetric connections were used instead. The decoder layers consists of conventional transpose convolution operation. The spatial dimension of the input has been kept intact in order to preserve the features and avoid the blurred artifacts in the results. A representation network (shown in Figure 7) has also been proposed to extract the latent features of the input color image to reconstruct the given image in YCbCr color space during the cost computation of the main model.

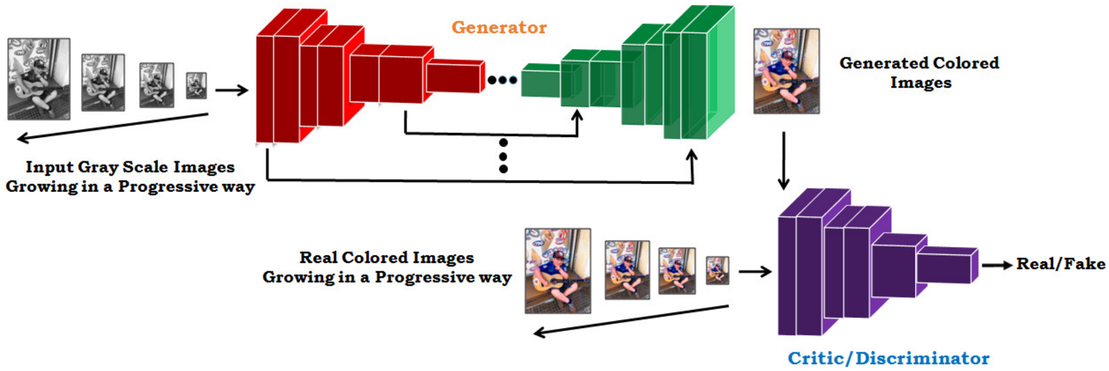


Figure 5. Team_India's proposed solution.

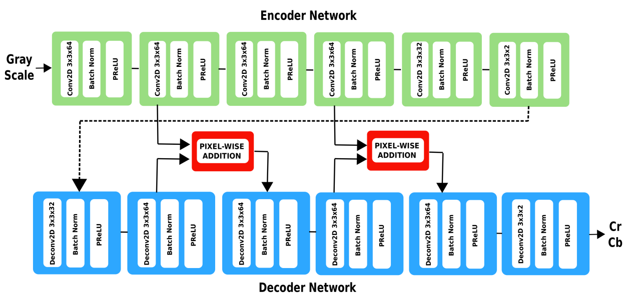


Figure 6. The pksvision_mm's network structure.

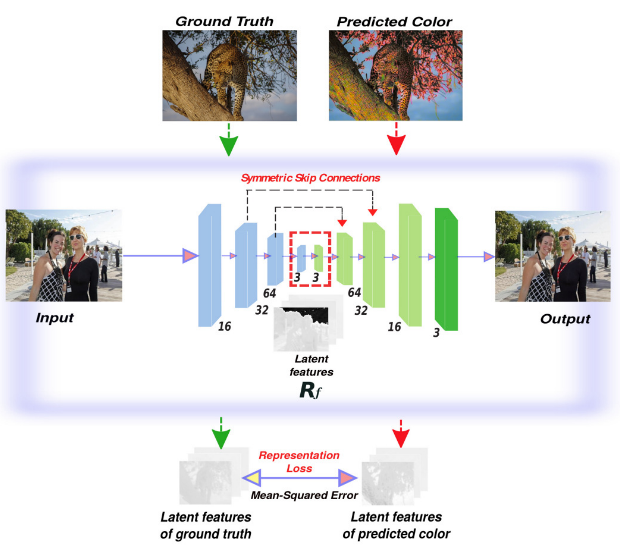


Figure 7. The pksvision_mm's network structure.

4.6. ITU-GO

ITU-GO team uses the Capsule Network (CapsNet) [23] for the problem of image colorization. The feature detector part of the original CapsNet model proposed in [23] is updated with first two layers of VGG-19 [26] and they are initialized from original VGG-19 model pre-trained on ImageNet [22]. The margin loss of original CapsNet model is modified with Mean Squared Error (MSE) for this image-to-image problem. Training and testing phases are performed as patch-based by using $9 \times 9 \times 3$ gray/color patch pairs. The Lab colormap is selected as colorspace for the color patches as in [32]. The network is trained 10 epoch first with ILSVRC 2012 dataset (ImageNet) by creating same kind of patch pairs ($9 \times 9 \times 3$). The resulting network is then re-trained with provided DIV2K dataset [1] (using both train and validation sets). The final model is used to predict the images from test set.

The network is trained via the RMSE, adversarial and feature losses in the CIE Lab space. The total number of parameters is 2,216,243 whose 768 are non-trainable and the total model size is about 8.9 MB. It takes the network about 7 seconds to process one input image.

The ITU-GO proposed solution is described in detail in [20].

Acknowledgements

We thank the NTIRE 2019 sponsors: Huawei Technologies Co. Ltd., NVIDIA Corp., Amazon.com, Inc., Samsung, MediaTek, Oppo and ETH Zurich (Computer Vision Lab).

A. Teams and affiliations

NTIRE2019 team

Title: NTIRE 2019 Challenge on Image Colorization
Members: *Shuhang Gu[1] (shuhang.gu@vision.ee.ethz.ch)*, Radu Timofte[1], Richard Zhang[2]
Affiliations:
[1] Computer Vision Lab, ETH Zurich, Switzerland
[2] Adobe, US

IPCV_IITM

Title: Efficient Image Colorization Using Dense Residual Network
Members: *Kuldeep Purohit (kuldeeppurohit3@gmail.com)*, Maitreya Suin, A. N. Rajagopalan
Affiliation:
Indian Instituite of Technology Madras

Athi

Title: Image Colorization using Color Mapping
Members: *Athi Narayanan S. (mail2athi@gmail.com)*, Jameer Babu Pinjari
Affiliation:
Amrita Vishwa Vidyapeetham, Amritapuri, Kerala, India

VIDAR

Title: Image Colorization Using Channel Attention U-Net
Members: *Zhiwei Xiong (zwxiong@ustc.edu.cn)*, Zhan Shi, Chang Chen, Dong Liu
Affiliation:
University of Science and Technology of China.

Team_India

Title: Improved Attention based Progressive GAN for Image Colorization
Members: *Manoj Sharma (mksnith@gmail.com)*, Megh Makwana, Anuj Badhwar, Ajay Pratap Singh, Avinash Upadhyay, Akkshita Trivedi, Anil Saini, Santanu Chaudhury
Affiliation:
CEERI-Delhi Center, IIT Delhi, CCS Computers Pvt Ltd, Delhi Technological University

pksvision_mm

Title: Representation Loss for Image Colorization using Generative Adversarial Network
Members: *Prasen Kumar Sharma (kumar176101005@iitg.ac.in)*, Priyankar Jain, Arijit Sur
Affiliation:
Indian Institute of Technology Guwahati, India

ITU-GO

Title: Image Colorization By Capsule Networks
Members: *Gökhan Özbulak (gokhan.ozbulak@itu.edu.tr)*
Affiliation:
Istanbul Technical University

References

[1] E. Agustsson and R. Timofte. NTIRE 2017 challenge on single image super-resolution: Dataset and study. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, July 2017. 1, 6

[2] C. Ancuti, C. O. Ancuti, and R. Timofte. Ntire 2018 challenge on image dehazing: Methods and results. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2018. 4

[3] B. Arad, O. Ben-Shahar, and R. Timofte. Ntire 2018 challenge on spectral reconstruction from rgb images. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2018. 4

[4] Y. Blau, R. Mechrez, R. Timofte, T. Michaeli, and L. Zelnik-Manor. The 2018 pirm challenge on perceptual image super-resolution. In *The European Conference on Computer Vision (ECCV) Workshops*, September 2018. 4

[5] A. Deshpande, J. Lu, M.-C. Yeh, M. Jin Chong, and D. Forsyth. Learning diverse image colorization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6837–6845, 2017. 1

[6] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 27*, pages 2672–2680. Curran Associates, Inc., 2014. 3

[7] S. Guadarrama, R. Dahl, D. Bieber, M. Norouzi, J. Shlens, and K. Murphy. Pixcolor: Pixel recursive colorization. *arXiv preprint arXiv:1705.07208*, 2017. 1

[8] K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, pages 1026–1034, 2015. 5

[9] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016. 3, 5

[10] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017. 3, 4

[11] A. Ignatov, N. Kobyshev, R. Timofte, K. Vanhoey, and L. Van Gool. Dslr-quality photos on mobile devices with deep convolutional networks. In *The IEEE International Conference on Computer Vision (ICCV)*, Oct 2017. 3, 4

[12] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015. 5

[13] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017. 4, 5

[14] J. Johnson, A. Alahi, and L. Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *European*

conference on computer vision, pages 694–711. Springer, 2016. 3

[15] T. Karras, T. Aila, S. Laine, and J. Lehtinen. Progressive growing of gans for improved quality, stability, and variation. *International Conference on Learning Representations (ICLR)*, 2018. 5

[16] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 5

[17] G. Larsson, M. Maire, and G. Shakhnarovich. Learning representations for automatic colorization. In *European Conference on Computer Vision*, pages 577–593. Springer, 2016. 1

[18] A. Levin, D. Lischinski, and Y. Weiss. Colorization using optimization. In *ACM transactions on graphics (tog)*, volume 23, pages 689–694. ACM, 2004. 1

[19] M. Mirza and S. Osindero. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*, 2014. 5

[20] G. Ozbulak. Image colorization by capsule networks. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2019. 6

[21] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015. 3, 5

[22] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al. Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115(3):211–252, 2015. 6

[23] S. Sabour, N. Frosst, and G. E. Hinton. Dynamic routing between capsules. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 3856–3866. Curran Associates, Inc., 2017. 6

[24] M. Sharma, M. Makwana, A. Upadhyay, A. P. Singh, A. Badhwar, A. Trivedi, A. Saini, and S. Chaudhury. Robust image colorization using self attention based progressive generative adversarial network. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2019. 5

[25] Z. Shi, C. Chen, Z. Xiong, D. Liu, Z.-J. Zha, and F. Wu. Deep residual attention network for spectral image super-resolution. In *ECCV Workshops*, 2018. 5

[26] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. 5, 6

[27] R. Timofte, E. Agustsson, L. Van Gool, M.-H. Yang, L. Zhang, et al. Ntire 2017 challenge on single image super-resolution: Methods and results. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, July 2017. 1

[28] R. Timofte, S. Gu, J. Wu, and L. Van Gool. Ntire 2018 challenge on single image super-resolution: Methods and results. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2018. 1, 4

[29] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, April 2004. 2

[30] H. Zhang, I. Goodfellow, D. Metaxas, and A. Odena. Self-attention generative adversarial networks. *arXiv preprint arXiv:1805.08318*, 2018. 5

[31] H. Zhang, V. Sindagi, and V. M. Patel. Multi-scale single image dehazing using perceptual pyramid deep network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 902–911, 2018. 4

[32] R. Zhang, P. Isola, and A. A. Efros. Colorful image colorization. In *European conference on computer vision*, pages 649–666. Springer, 2016. 1, 6

[33] R. Zhang, J.-Y. Zhu, P. Isola, X. Geng, A. S. Lin, T. Yu, and A. A. Efros. Real-time user-guided image colorization with learned deep priors. *arXiv preprint arXiv:1705.02999*, 2017. 1, 4

[34] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu. Image super-resolution using very deep residual channel attention networks. In *ECCV*, 2018. 5

[35] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *The IEEE International Conference on Computer Vision (ICCV)*, Oct 2017. 4