SEMI-DERAINGAN: A NEW SEMI-SUPERVISED SINGLE IMAGE DERAINING

Yanyan Wei¹,², Zhao Zhang*,¹,², Yang Wang¹,², Haijun Zhang³, Mingbo Zhao⁴, Mingliang Xu*,⁵, Meng Wang¹,²

¹ School of Computer Science and Information Engineering, Hefei University of Technology, China
² Key Laboratory of Knowledge Engineering with Big Data (Ministry of Education) & Intelligent Interconnected Systems Laboratory of Anhui Province, Hefei University of Technology, China
³ Department of Computer Science, Harbin Institute of Technology, Shenzhen, China
⁴ Department of Electronic Engineering, City University of Hong Kong, Hong Kong
⁵ School of Information Engineering, Zhengzhou University, China

ABSTRACT

Although supervised single image deraining (SID) have obtained impressive results, they still cannot obtain satisfactory results on real images for the weak generalization of rain removal capacity. In this paper, we mainly discuss the semi-supervised SID and propose a new GAN-based deraining network called Semi-DerainGAN, which can use both synthetic and real data in a uniform network based on two supervised and unsupervised processes. For this task, a semi-supervised rain streak learner termed SSRML sharing the same parameters of both processes is derived, which makes the real images contribute more rain streak information, so that the resulted model has a strong generalization power to the real SID task. We also contribute a new real-world rain image dataset called Real200 to alleviate the difference between both synthetic and real image domains. Extensive results on public datasets show that our model can obtain competitive results, especially on the real rain images.

Index Terms— Single image deraining, Semi-supervised learning, rain removal, dataset

1. INTRODUCTION

Single image deraining (SID), which is a classical image restoration task, has been a challenging and interesting topic in the areas of computer vision and artificial intelligence. The problem of SID can be formulated as follows:

\[ X = R + B, \]  

(1)

where \( X \) denotes a rain image that will be decomposed into a rain-streak component \( R \) and a clean background \( B \). Because Eqn. (1) is an ill-posed problem, some feasible approaches have been proposed to solve it, include both the traditional [1, 2, 3] and deep learning-based models [4, 5, 6, 7, 8, 9].

It is noteworthy that most existing deep deraining networks are supervised methods using paired information in synthetic datasets. The strong constraint can make the network convergence fast, however, the performance on the real rain images is still unsatisfactory due to lack of paired information. To address this issue, some researchers start to shift to study the semi-supervised deraining models, which can also use real images to enhance the model generalization ability. For example, a semi-supervised SID framework called Semi-Supervised Transfer Learning (or shortly SIRR) [4] was recently proposed to address the paired data restricted SID task. SIRR adds the real rain images without ground-truth for the network training, by taking the residual between the input rain image and expected network output (clear image without rain) as a specific parameterized rain fringe distribution. As such, SIRR adapts to the real unsupervised rain types by transferring the supervised synthetic rain, which can clearly alleviate the issue of insufficient training data and supervised sample bias. But the difference of the rain streaks in synthetic and real images is very large, so it may be inappropriate to train the model by adding a relatively hard constraint between both synthetic and real image domains in a single network, i.e., minimizing the KL divergence between a synthetic rain streak mask and the real one learned from real images via training. Note that the KL divergence is the most important part of SIRR, which prompts the semi-supervised model to be effective. As a result, SIRR may leave rain streak information based on the synthetic and real images (see Fig.1).
To solve this problem, we investigate how to improve the paired data restricted SID task and propose a new semi-supervised SID network that use two processes to train synthetic and real images respectively via a hybrid loss. The difference between our model and SIRR is illustrated in Fig.2, and the main contributions are summarized as:

- A new and effective semi-supervised SID network, termed Semi-DerainGAN, is technically proposed. It can use partial paired data in a uniform network by two separate processes. Semi-DerainGAN can clearly avoid the interaction to be degraded, which is caused by the large difference between the rain streak information of the synthetic and real image domains.

- A new semi-supervised rain mask learner (SSRML) is designed to learn rain streak information from both synthetic and real rain images, and two generators for two independent processes are used to generate the derained images by training.

- A new real rain image dataset Real200 is also created, which can be applied by other models to alleviate the difference between synthetic and real image domains. To the best of our knowledge, this is the first real rain image dataset with more complex carefully-designed rain streak directions and shapes.

2. RELATED WORK

To solve the SID problem, many traditional algorithms have been proposed, such as low-rank representation methods [1], Gaussian mixture models [2], and sparse coding-based models [3]. In recent years, some deep learning-based deep network models have been proposed. For example, a contextualized dilated network [5] was recently proposed to jointly detect and remove the rain streaks from single image. [10] use the residual block to reduce the mapping range from the input to output directly, which makes the learning process easier. A novel density-aware multi-stream dense convolutional network-based framework [6] was proposed to jointly estimate the rain density and deraining. More recently, a hybrid block [7] has been proposed to extract the rain streak more precisely, especially in heavy rain condition. Similarly, a better and simpler baseline deraining network is proposed in [8] by repeatedly unfolding a shallow ResNet to take advantage of the recursive computation.

Because of the powerful capacity of generating realistic images, Generative Adversarial Networks (GAN) [11] have achieved superior performance in many vision tasks, including the SID. [9] proposed a conditional GAN-based network model which considers quantitative, visual and discriminative performance into the objective function. [12] proposed a raindrop removal method which uses GAN to produce the attention map and uses it to generate a raindrop-free image through a contextual auto-encoder. Note that these existing

\[ m_s = \mathcal{S}(x_s), m_r = \mathcal{S}(x_r), \]

where \( \mathcal{S}(\cdot) \) denotes the function of SSRML module, \( m_s \) and \( m_r \) are the rain masks extracted from synthetic rain image \( x_s \) and real rain image \( x_r \), respectively.

3. PROPOSED METHOD

3.1. Network Architecture

We show the framework of Semi-DerainGAN in Fig.3, which has two supervised and unsupervised processes. More specifically, the framework has a semi-supervised rain mask learner (SSRML), three generators \((G_s, G_r, G_r^{'})\) that dispose on synthetic rain images and real rain images respectively, and three discriminators \((D_s, D_r, D_p)\) corresponding to the generators. The synthetic and real rain images are denoted as \( \{x_s^i, y_s^i\}_{i=1}^{M_s} \in S \) and \( \{x_r^i, \hat{y}_r^i\}_{i=1}^{N_r} \in R \), respectively, where \( x_s \) and \( y_s \) are rain image and its corresponding rain-free image (i.e., label) in the synthetic dataset \( S \), while \( x_r \) and \( \hat{y}_r \) are rain image and corresponding rain-free image in real dataset \( R \). Since \( x_r \) has no corresponding ground truth, we randomly choose \( \hat{y}_r \) from \( \{\hat{y}_r^i\}_{i=1}^{M_r} \in S \) as the ground truth (i.e., fake label). In what follows, we will introduce detailed information of each part respectively.

3.1.1. Semi-Supervised Rain Mask Learner (SSRML)

The original rain mask learner (RML), as an attentional rain drops extractor by [12], performs in supervised mode, which includes a LSTM unit and five Conv_ReLU_Conv_ReLU units. We design a structure to use RML as an attentional rain streak extractor in semi-supervised mode, called semi-supervised rain mask learner (SSRML), to learn rain streak information (i.e., shapes and directions) from both the synthetic and real-world rain image domains. Under the circumstances, SSRML can obtain better deraining result in the real-world SID task. Overall, this process can be formulated as:

\[ m_s = \mathcal{S}(x_s), m_r = \mathcal{S}(x_r), \]
3.1.2. Generators in Synthetic and Real Image Domains

We describe the three generators $G_s, G_r, G_r'$. $G_s$ and $G_r$ can generate the derained images from synthetic data $S$ and real data $R$ in training process respectively, and we utilize the U-net. Then, $G_r'$ reconstructs the real rain image for a consist purpose which is proposed in CycleGAN [13], and is consisted of a SSRML and a U-net. The formulation for the generators can be represented as:

$$\hat{y}_s = G_s(m_s, x_s), \quad \hat{y}_r = G_r(m_r, x_r), \quad \tilde{x}_r = G_r'(\hat{y}_r),$$

where $\hat{y}_s$ and $\hat{y}_r$ are the derained results of images $x_s$ and $x_r$, and $\tilde{x}_r$ is reconstructed rain image of the derained image $\hat{y}_r$.

3.1.3. Discriminators in Synthetic and Real Domains

There are two types of discriminators in our model. The first type contains $D_s$ and $D_r$ that apply multi-scale structures, where the feature maps at each scale go through five convolution layers and then are fed into sigmoid outputs [14]. We use 3 different scales for $D_s$ and $D_r$. The second type of discriminator $D_p$ is a paired discriminator [15]. In our network, we simplify the input of $D_p$ and use paired images, i.e., rain and derained images, to make the network generate more realistic derained images. The efficiency of the paired discriminator is verified by simulations. The adversarial losses for used generators and discriminators can be defined as follows:

$$\mathcal{L}_{adv, super} = \mathbb{E}_{x_s \sim S}[\log D_s(y_s)] + \mathbb{E}_{x_r \sim S}[\log(1 - D_s(G_s(S(x_s), x_s)))] + \mathcal{L}_{adv, pair},$$

$$\mathcal{L}_{adv, pair} = \mathbb{E}_{x_r \sim S}[\log D_p(x_r, y_s)] + \mathbb{E}_{x_r \sim S}[\log(1 - D_p(x_r, G_s(S(x_s), x_s)))] + \mathcal{L}_{adv, unsup} = \mathbb{E}_{x_r \sim R}[\log D_r(\hat{y}_r)] + \mathbb{E}_{x_r \sim R}[\log(1 - D_r(G_r(S(x_r), x_r)))] + \mathcal{L}_{adv, unsup} = \lambda_1 \mathcal{L}_{adv, super} + \lambda_2 \mathcal{L}_{adv, super} + \lambda_3 \mathcal{L}_{adv, unsup},$$

where $\mathcal{L}_{adv, super}$ is the adversarial loss which contains a loss of pair discriminator $\mathcal{L}_{adv, pair}$ for supervised process and $\mathcal{L}_{adv, unsup}$ is an adversarial loss for unsupervised process.

3.2. Supervised Process

In the supervised process, we use the synthetic data $\{x_s^i, y_s^i\}_{i=1}^M$ to learn the parameters of network, i.e., SSRML, $G_s$, $D_s$ and $D_p$, and the loss function denotes as:

$$\mathcal{L}_{super} = \lambda_1 \mathcal{L}_{adv, super} + \lambda_2 \mathcal{L}_{adv, super} + \lambda_3 \mathcal{L}_{ssim},$$

where $\mathcal{L}_{adv, super}$ is an adversarial loss for the process in Eqn. (4), $\mathcal{L}_{adv, super}$ denotes the perceptual loss [16] that can encode the difference between the derained image $\hat{y}_s$ and the corresponding label image $y_s$, $\mathcal{L}_{ssim}$ is the SSIM loss [17] that can keep structural similarity between the two images, $\lambda_n$ is tradeoff parameter to control the contributions of each loss. The two losses $\mathcal{L}_{per, super}$ and $\mathcal{L}_{ssim}$ can be defined as:

$$\mathcal{L}_{per, super} = \|\partial_c(y_s) - \partial_c(\hat{y}_s)\|^2_2,$$

$$\mathcal{L}_{ssim} = -SSIM(y_s, \hat{y}_s),$$

where $\partial_c(\cdot)$ denotes the feature extractor of the conv2_3 layer of the VGG-16 network [18] that are pre-trained on the ImageNet [19]. $SSIM(\cdot)$ is the SSIM function to calculate the similarity between two images $y_s$ and $\hat{y}_s$.

3.3. Unsupervised Process

In the unsupervised process, we use the real-world rain image data $\{x_r^i, \tilde{y}_r^i\}_{i=1}^N \in R$ without paired information to learn the parameters of the network, i.e., SSRML, $G_r$, $G_r'$ and $D_r$. Specifically, we minimize the unsupervised loss function:

$$\mathcal{L}_{unsup} = \lambda_4 \mathcal{L}_{adv, unsup} + \lambda_5 \mathcal{L}_{per, unsup} + \lambda_6 \mathcal{L}_{cc} + \lambda_7 \mathcal{L}_{tv},$$

Fig. 3: The framework of our Semi-DerainGAN that includes two processes, i.e., supervised and unsupervised processes.
Table 1: Comparison with each model in terms of PSNR. w/o R means without real data.

| Dataset     | Input | Rain1400 |                   | Rain1400&SIRR-Data |                   |
|-------------|-------|----------|-------------------|--------------------|-------------------|
|             |       | DSC      | GMM               | DDN                | JORDER            | DID-MDN          | SIRR | Ours w/o R | Ours |
| Sparse      |       | 24.14    | 25.05             | 25.67              | 26.88             | 24.22            | 25.66 | 26.98       | 26.15 | **28.21** |
| Dense       |       | 17.95    | 19.00             | 19.27              | 19.90             | 18.75            | 18.60 | 21.60       | 20.32 | **23.15** |

Table 2: Quantitative comparison on the synthetic datasets. * denotes the model is trained on synthetic and real dataset.

| Datasets   | Rain100H | Rain100L | Rain12 |
|------------|----------|----------|--------|
| Metrics    | PSNR/SSIM| PSNR/SSIM| PSNR/SSIM|
| DSC        | 13.77/0.320 | 27.34/0.849 | 30.07/0.866 |
| GMM        | 15.23/0.450 | 29.05/0.872 | 32.14/0.916 |
| DDN        | 22.85/0.725 | 32.38/0.926 | 34.04/0.933 |
| SIRR*      | 22.47/0.716 | 32.37/0.925 | 34.02/0.935 |
| Ours*      | 22.89/0.801 | **34.12/0.958** | **35.86/0.960** |

Fig. 4: Real images in two datasets (SIRR-Data and Real200).

\[ \mathcal{L}_{adv,unsup} \] is the adversarial loss for unsupervised process, \( \mathcal{L}_{cc} \) ensures that the derained image \( \tilde{y}_r \) can be translated back to the original rain image \( x_r \) for preserving the contents of images, \( \mathcal{L}_{per,unsup} \) similarly defined as \( \mathcal{L}_{per,super} \), \( \mathcal{L}_{tv} \) is the TV function to make generated real derained image more realistic. \( \lambda_n \) is the trade-off parameter to balance the losses. Technically, \( \mathcal{L}_{cc}, \mathcal{L}_{per,unsup} \) and \( \mathcal{L}_{tv} \) denote as:

\[
\mathcal{L}_{cc} = \mathbb{E}_{x_r \sim R} \| x_r - \bar{x}_r \|_1, \tag{10}
\]

\[
\mathcal{L}_{per,unsup} = \| \partial_v(x_r) - \partial_v(\tilde{y}_r) \|_2^2, \tag{11}
\]

\[
\mathcal{L}_{tv} = TV(\tilde{y}_r). \tag{12}
\]

### 3.4. Objective Function

The overall loss function of Semi-DerainGAN used for training the proposed network is defined as follows:

\[
\mathcal{L}_{total} = \mathcal{L}_{super} + \lambda_{unsup} \mathcal{L}_{unsup}, \tag{13}
\]

where \( \lambda_{unsup} \) is a tradeoff parameter to balance the supervised process \( \mathcal{L}_{super} \) and unsupervised process \( \mathcal{L}_{unsup} \).

### 3.5. Created Real rain Image Dataset Real200

Recently, [4] proposed a real rain image dataset termed SIRR-Data, which has 147 real rain images with more shape and density information than synthetic ones. However, SIRR-Data still have several drawbacks: (1) it contain many unsuitable weather conditions, e.g., snow and mist, which can hardly prompt the convergence in the semi-supervised training process; (2) it contain some low-resolution images whose rain streak information are difficult to be detected; (3) it is raw and not carefully prepared, as they may contain rain-free regions, or too light to see the rain streaks in sky or rainwater ground, or contain the rain drops or rain spray. Maybe some researchers think that these situations can enhance the generalization ability of the deep network, but in fact it may be negative for training due to the interference. Although these images are realistic, they are unsuitable for semi-supervised mode. As such, we create a new real rain image dataset Real200. The images are carefully selected from Internet and manually cropped to highlight the rain regions (see Fig.4).

The advantages of Real200 are twofold: (1) all images are high-resolution and only contain rain conditions; (2) it is cropped manually to highlight the rain streak regions, while preserving the diversity of the rain streaks. Note that the rain streaks of Real200 are not significantly different from those in synthetic images, so it will be beneficial for training a semi-supervised model, as can be seen in Table 4.

### 4. EXPERIMENTS AND ANALYSIS

#### 4.1. Training Details

Semi-DerainGAN is trained by Pytorch in Python environment on a NVIDIA GeForce GTX 1080ti GPU with 12GB memory. We crop each rain image to many patches of 100×100 by stride of 80. Adam is used as the optimization algorithm with a batch size of 4. The model is trained for total 200 epochs. The learning rates of supervised and unsupervised processes are set to 1e-4 and 1e-3, respectively. The learning rate is decayed with a policy of Pytorch after 100 epochs. The trade-off parameters in Eqn. (6) and Eqn. (13) are set to 1, and the trade-off parameters in Eqn. (9) are set to 1.5e-5, 10, 1 and 100, respectively. All the tradeoff parameters are chosen empirically.
4.2. Datasets and Evaluation Metrics

4.2.1. Evaluated Synthetic and Real rain image Datasets

For synthetic images, we use five datasets: Rain100H [5], Rain100L [5], Rain1400 [10], Rain12 [2], Rain20 [4]. For real images, we use two datasets: SIRR-Data [4] that contains 147 real rain images; Real200 that is created in this paper, which contains 200 real-world rain images. Since our method and SIRR are all semi-supervised methods, we use the synthetic dataset plus real-world dataset for training, which is denoted by &, such as Rain1400&SIRR-Data.

4.2.2. Evaluation Metrics and Compared Methods

The Peak Signal-to-Noise Ratio (PSNR) [20] and Structural Similarity Index (SSIM) [17] are used for evaluating the rain images with ground-truth. For the real rain images that have no ground-truth, we only provide the visual deraining results.

In this experiment, seven popular SID methods are added for comparing with, including two model-driven methods (i.e., DSC [3] and GMM [2]), four supervised deep learning methods (i.e., DDN [10], JORDER [5], DID-MDN [6] and PReNet [8]), and one most related semi-supervised deep learning method (i.e., SIRR [4]).

4.3. Results on Synthetic rain images

We first evaluate each method on synthetic datasets. The supervised models are directly trained on Rain1400 and tested on Rain20, the semi-supervised deep models are trained on Rain1400&SIRR-Data. From Table 1, we see that our model can obtain the best performance on both sparse and dense test data. To show the effect of unsupervised process in our model, we also train it using a supervised mode totally (i.e., without real data). The result shows that the performance of adding real images into the training process can be improved. The illustration of the deraining results can also be seen in Fig. 5.

We also evaluate several SID methods on synthetic datasets (i.e., Rain100H, Rain100L and Rain12) in Table 2 and visualize some SID results in Fig.6. The image deraining results also demonstrate the superior performance of our method.

4.4. Results on Real rain Images

We compare the result of our network with others on Rain1400&SIRR-Data. We visualize the SID results in Fig.7, from which we see that our network performs better than the other methods, especially for the most related SIRR, since SIRR leaves more rain streak information, which keeps consistent with the above numerical evaluation results.

4.5. Ablation Study

We discuss the selections of loss functions and module in our network, and demonstrate the availability of our created real rain image dataset Real200. For fair comparison, we use Rain100L&SIRR-Data as the training data.

| Setting | w/o. $L_{per}$ | w/o. $L_{tv}$ | w/o. $D_p$ | All |
|---------|----------------|----------------|-------------|-----|
| PSNR    | 32.25          | 33.82          | 29.94       | 34.12 |
| SSIM    | 0.939          | 0.944          | 0.949       | 0.958 |

4.5.1. Loss Functions and Modules

We explore the deraining results of our method without $L_{per}$ or $L_{tv}$, in Table 3, from which we see that the deraining performance goes down without partial loss, especially on the metric of PSNR. Besides, we also evaluate the importance of $D_p$, and we can see that without $D_p$, the resulted performance drops fast. The results demonstrate that: (1) the hybrid loss in our supervised and semi-supervised processes can optimize our network effectively. Since the perceptual loss constrains the feature maps between both paired and unpaired images, and the TV loss can constrain the model to generate more realistic images; (2) the discriminator $D_p$ can distinguish image pairs and make the model generate realistic derained images.

4.5.2. Evaluation on SIRR-Data and Our Real200

We investigate the effectiveness of created real dataset Real200. As can be seen in Table 4, the deraining results of
Table 4: Ablation study of real datasets.

| Datasets          | Rain100L& | Rain100L& |
|-------------------|-----------|-----------|
|                   | SIRR-Data | Real200   |
| Metrics           | PSNR/SSIM | PSNR/SSIM |
| Semi-DerainGAN    | 34.12/0.958 | 35.27/0.988 |

our model on Rain100L& Real200 is clearly better than that on Rain100L& SIRR-Data. That is, Real200 can clearly enhance the deraining performance for SID by providing more rich, useful and accurate rain streak information to improve the model generalization power during the training process.

5. CONCLUSION

We propose a new semi-supervised GAN-based deraining network that can use both the synthetic and real rain images to a uniform network with two supervised and unsupervised processes. Based on the semi-supervised rain mask learner that makes the real images contribute more rain streak information to improve the results and paired discriminator, our network can clearly perform better than the recently proposed method SIRR. We also contribute a new real image dataset Real200, and based on Real200 we can obtain the enhanced SID results. In the future, we will also study better training architectures and skills for cross-domain semi-supervised SID, and discover more connections between different domains.

Acknowledgment

This work is partially supported by the National Natural Science Foundation of China (62072151, 61672365), Anhui Provincial Natural Science Fund for Distinguished Young Scholars (2008085330), and the Fundamental Research Funds for the Central Universities of China (JZ2019HGPA0102). Zhao Zhang is the corresponding author of this paper.

6. REFERENCES

[1] He Zhang and Vishal M Patel, “Density-aware single image de-raining using a multi-stream dense network,” in CVPR, 2018, pp. 695–704.
[2] Yu Li, Robby T Tan, Xiaojie Guo, Jiangbo Lu, and Michael S Brown, “Rain streak removal using layer priors,” in CVPR, 2016, pp. 2736–2744.
[3] Yu Luo, Yong Xu, and Hui Ji, “Removing rain from a single image via discriminative sparse coding,” in ICCV, 2015, pp. 3397–3405.
[4] Wei Wei, Deyu Meng, Qian Zhao, Zongben Xu, and Ying Wu, “Semi-supervised transfer learning for image rain removal,” in CVPR, 2019, pp. 3877–3886.
[5] Wenhan Yang, Robby T Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, and Shuicheng Yan, “Deep joint rain detection and removal from a single image,” in CVPR, 2017, pp. 1357–1366.
[6] He Zhang and Vishal M Patel, “Density-aware single image de-raining using a multi-stream dense network,” in CVPR, 2018, pp. 695–704.
[7] Yanyan Wei, Zhao Zhang, Haijun Zhang, Richang Hong, and Meng Wang, “A coarse-to-fine multi-stream hybrid deraining network for single image deraining,” in ICDM, 2019, pp. 628–637.
[8] Dongwei Ren, Wangmeng Zuo, Qinghua Hu, Pengfei Zhu, and Deyu Meng, “Progressive image deraining networks: a better and simpler baseline,” in CVPR, 2019, pp. 3937–3946.
[9] He Zhang, Vishwanath Sindagi, and Vishal M Patel, “Image de-raining using a conditional generative adversarial network,” TCSVT, 2019.
[10] Xueyang Fu, Jiabin Huang, Delu Zeng, Yue Huang, Xinghao Ding, and John Paisley, “Removing rain from single images via a deep detail network,” in CVPR, 2017, pp. 3855–3863.
[11] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, “Generative adversarial nets,” in NIPS, 2014, pp. 2672–2680.
[12] Rui Qian, Robby T Tan, Wenhan Yang, Jiajun Su, and Jiaying Liu, “Attentive generative adversarial network for raindrop removal from a single image,” in CVPR, 2018, pp. 2482–2491.
[13] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in ICCV, 2017, pp. 2223–2232.
[14] Hsin-Ying Lee, Hung-Yu Tseng, Jia-Bin Huang, Maneesh Singh, and Ming-Hsuan Yang, “Diverse image-to-image translation via disentangled representations,” in ECCV, 2018, pp. 35–51.
[15] Qiao Gu, Guanzhi Wang, Mang Tik Chiu, Yu-Wing Tai, and Chi-Keung Tang, “Ladn: Local adversarial disentangling network for facial makeup and de-makeup,” in ICCV, 2019, pp. 10481–10490.
[16] Justin Johnson, Alexandre Alahi, and Li Fei-Fei, “Perceptual losses for real-time style transfer and super-resolution,” in ECCV, 2016, pp. 694–711.
[17] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli, “Image quality assessment: from error visibility to structural similarity,” TIP, vol. 13, no. 4, pp. 600–612, 2004.
[18] Karen Simonyan and Andrew Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv preprint arXiv:1409.1556, 2014.
[19] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in CVPR, 2009, pp. 248–255.
[20] Quan Huynh-Thu and Mohammed Ghanbari, “Scope of validity of psnr in image/video quality assessment,” Electronics letters, vol. 44, no. 13, pp. 800–801, 2008.