FACE HALLUCINATION WITH FINISHING TOUCHES

Mr. SANDESH R1, RAKSHITH S V2, SUMANTH B M2, SUSHMITHA C R2, VARSHA C2
1Assistant Professor, Department of CSE, ATMECE, Mysuru
2 8th Semester students, Department of CSE, ATMECE, Mysuru

Abstract— Recovering high-resolution (HR) face images from their low-resolution (LR) counterparts, known as face hallucination, obtaining a high-quality frontal face image from a low-resolution (LR) non-frontal face image is primarily important for many facial analysis applications. Main-streams either focus on super-resolving near-frontal LR faces or frontalizing non-frontal high-resolution (HR) faces. We present a novel Vivid Face Hallucination Generative Adversarial Network (VividGAN) for simultaneously super-resolving and frontalizing tiny non-frontal face images. VividGAN consists of coarse-level and fine-level Face Hallucination Networks (FHnet) and two discriminators, i.e., Coarse-Dand Fine-D. The coarse-level FHnet generates a frontal coarse HR face and then the fine-level FHnet makes use of the facial component appearance prior. Face recognition and expression classification, compared with other state-of-the-art methods. Face super-resolution (FSR), also known as face hallucination. First, we summarize the problem formulation of FSR and introduce popular assessment metrics and loss functions. Second, we elaborate on the facial characteristics and popular datasets used in FSR. Third, we roughly categorize existing methods according to the utilization of facial characteristics. In each category, we start with a general description of design principles, then present an overview of representative approaches, and then discuss the pros and cons among them. Fourth, we evaluate the performance of some state-of-the-art methods. Fifth, joint FSR and other tasks, and FSR-related applications are roughly introduced. Finally, we envision the prospects of further technological advancement in this field. Hallucination, or hallucination followby frontalization produces Conventional and emerging face frontalization methods. Often rely on facial landmarks warping 2D face images onto 3D models, and thus require the input images to have a sufficient resolution where such landmarks are detectable. This renders them ineffective for tiny face images. Without a proper frontalization, directly applying face hallucination methods may cause severe artifacts due to large pose variations and misalignments, for very low-resolution non-frontal face images, applying either face frontalization followed by degraded results.

Keywords— face super-resolution, Face hallucination, super-resolution, face frontalization, generative adversarial network.

I. INTRODUCTION

Face recognition, a highly researched subject in computer vision, has gained significant attention due to its broad range of applications in law enforcement, biometrics, marketing, and more. In recent times, deep learning-based techniques have led to substantial advancements in face recognition. In recent years, the process of recovering high-resolution (HR) face images from their low-resolution (LR) counterparts, commonly known as face hallucination, has garnered significant attention. Current face hallucination methods primarily focus on super-resolving nearly frontal faces, which provide crucial perceptual information for the human visual system. However, in many cases, LR faces may not be in a frontal pose. Super-resolving non-frontal LR faces either requires frontalizing them first and then applying existing face hallucination techniques, or super-solving first (which relies heavily on an available pose-specific exemplar dataset) and then frontalizing. Nevertheless, both of these options pose significant challenges

Traditional and modern techniques for face frontalization often depend on detecting facial landmarks to warp 2D face images onto 3D models. As a result, these methods require input images with high enough resolution for accurate landmark detection, making them unsuitable for tiny face images. When dealing with low-resolution non-frontal face images, using face frontalization followed by face hallucination, or vice versa, can lead to poor results due to significant pose variations and misalignments causing artifacts.

Our goal is to simultaneously perform face frontalization and face hallucination on a given input image to mitigate the artifacts that may arise when these tasks are performed separately. The Transformative Adversarial Neural Network (TANN) is an advanced technique in computer vision that focuses on face frontalization and face hallucination. The Transformative Adversarial Neural Network (TANN) is a method that seamlessly combines face frontalization and face hallucination by automatically transforming low-resolution (LR) faces into frontalized LR feature maps, while also upsampling the images by a factor of 8x in an end-to-end manner. TANN aims to generate high-
quality frontalized face images, even from low-resolution inputs, making it valuable for applications such as face recognition, facial expression analysis, and virtual reality. TANN represents a significant advancement in facial image synthesis and holds great potential for various real-world applications in the field of computer vision and image processing.

To begin with, a subnetwork is created with the aim of transforming a non-frontal LR face into a latent representation. This transformation ensures that the latent representation of the non-frontal face closely resembles the latent representation of its frontal counterpart in the latent subspace. Next, the resulting frontalized LR feature maps are fed into another subnetwork that comprises deconvolutional and spatial transformer layers. The ultimate objective of this subnetwork is to produce HR outputs.

To train our network effectively, we incorporate not only the conventional image appearance and class-wise similarity constraints that were used in our earlier works, but also introduce a triplet loss. This loss helps us ensure that the latent representations of the input non-frontal faces and their corresponding ground-truth frontal LR ones are similar, while being distant from other LR frontal faces in the latent subspace. This allows us to generate upscaled frontalized faces that not only resemble their HR frontal counterparts, but are also distinct from other artificially generated faces, since the same upsampling subnetwork is used for super-resolution. As a result of using the proposed triplet loss, the upscaled frontalized faces retain the identity information implicitly and share similar facial characteristics with their corresponding ground-truth ones after super-resolution. It is important to note that unlike the traditional triplet loss, where both positive and negative examples are used to compute the neural network gradients and updated simultaneously, we update only the latent representations of LR side-view faces. We ensure that they are like the representations of their ground-truth frontal faces without affecting the positive and negative LR frontal faces. Additionally, we use a feature-wise similarity constraint, referred to as perceptual loss, to enhance the visual quality by making the artificially generated facial characteristics like the ground-truths.

II. LITERATURE SURVEY

Ping Liu, Yuewei Lin, Zibo Meng, Weihong Deng, Joey Tianyi Zhou, and Yi Yang (Point Adversarial Self-Mining: A Simple Method for Facial Expression Recognition in the Wild). In this article, we propose a simple yet effective approach, called point adversarial self mining (PASM), to improve the recognition accuracy in facial expression recognition (FER). Unlike previous works focusing on designing specific architectures or loss functions to solve this problem, PASM boosts the network capability by simulating human learning processes: providing updated learning materials and guidance from more capable teachers. Specifically, to generate new learning materials, PASM leverages a point adversarial attack method and a trained teacher network to locate the most informative position related to the target task, generating harder learning samples to refine the network. The searched position is highly adaptive since it considers both the statistical information of each sample and the teacher network capability. Other than being provided new learning materials, the student network also receives guidance from the teacher network. After the student network finishes training, the student network changes its role and acts as a teacher, generating new learning materials and providing stronger guidance to train a better student network. The adaptive learning materials generation and teacher/student update can be conducted more than one time, improving the network capability iteratively. Extensive experimental results validate the efficacy of our method over the existing state of the arts for FER (2020). Xin Yu, Basura Fernando, Bernard Ghanem, Fatih Porikli and Richard Hartley (Face Super-resolution Guided by Facial Component Heatmaps). State-of-the-art face super-resolution methods leverage deep convolutional neural networks to learn a mapping between low-resolution (LR) facial patterns and their corresponding high-resolution (HR) counterpart parts by exploring local appearance information. However, most of these methods do not account for facial structure and suffer from degradations due to large pose variations and misalignments. In this paper, we propose a method that explicitly incorporates structural information of faces into the face super-resolution process by using a multi-task convolutional neural network (CNN). Our CNN has two branches: one for super-resolving face images and the other branch for predicting salient regions of a face coincidental facial component heatmaps. These heatmaps encourage the upsampling stream to generate super-resolved faces with higher-quality details. Our method not only uses low-level information (i.e., intensity similarity), but also middle-level information (i.e., face structure) to further explore spatial constraints of facial components from LR inputs images. Therefore, we are able to super-resolve very small unaligned face images (16x16 pixels) with a large upsampling factor of 8x, while preserving face structure. Extensive experiments demonstrate that our network achieves superior face hallucination results and outperforms the state-of-the-art (2015). Z. Wang, X. Yu, M. Lu, Q. Wang, C. Qian, and F. Xu. (Single image portrait relighting via facial component heatmaps). Portrait relighting aims to render a face image under different lighting conditions. Existing methods do not explicitly consider some challenging lighting effects such as specular and shadow, and thus may fail in handling extreme lighting conditions. In this paper, we propose a novel framework that explicitly models multiple reflectance channels for single image portrait relighting, including the facial albedo, geometry as well as two lighting effects, i.e., specular and shadow. These channels are finally composed to generate the relit results via deep neural networks. Current datasets do not support learning such multiple reflectance channel modeling. Therefore, we present a large-scale dataset with the ground-truths of the channels, enabling us to train the deep neural networks in a supervised manner. Furthermore, we develop a novel module named Lighting guided Feature Modulation (LFM). In contrast to existing methods which simply incorporate the given lighting in the bottleneck of a network, LFM fuses the lighting by layer-wise feature modulation to deliver more convincing results. Extensive experiments demonstrate that our proposed method achieves better results and is able to generate challenging lighting effects (2018). L. Tran, X. Yin, and X. Liu. (Representation learning by Rotating your faces). The large pose discrepancy between two face images is one of the fundamental challenges in automatic face recognition. Conventional approaches to pose-invariant face recognition either perform face frontalization on, or learn a pose-invariant representation from a non-frontal face image. We argue that this is more desirable to perform both tasks jointly to allow them to leverage each other. To this end, this paper proposes a Disentangled Representation learning-Generative Adversarial Network (DR-GAN) with three distinct novelties. First, the encoder-decoder structure of the generator enables DR-GAN to learn a representation that is both generative and discriminative, which can be used for
face image synthesis and pose-invariant face recognition. Second, this representation is explicitly disentangled from other face variations such as pose, through the pose code provided to the decoder and pose estimation in the discriminator. Third, DR-GAN can take one or multiple images as the input, and generate one unified identity representation along with an arbitrary number of synthetic face images. Extensive quantitative and qualitative evaluation on a number of controlled and in-the-wild databases demonstrate the superiority of DR-GAN over the state of the art in both learning representations and rotating large-pose face images (2019).

III. PRELIMINARIES

Face Super-resolution (SR) aims at establishing the intensity relationships between input LR and output HR face images. The previous works are generally categorized into three mainstreams: holistic-based, part-based, and deep learning based methods. The basic principle of holistic-based techniques is to upsample a whole LR face by a global face model. Wang et al. [22] formulate a linear mapping between LR and HR images to achieve face SR based on an Eigen-transformation of LR faces. Liu et al. [23] incorporate a bilateral filtering to mitigate the ghosting artifacts. Kolouri and Rohde [24] morph HR faces from aligned LR ones based on optimal transport and subspace learning. However, they require LR inputs to be precisely aligned and reference HR faces to exhibit similar canonical poses and natural expressions. To address pose and expression variations, part-based methods are proposed to make use of exemplar facial patches to upsample local facial regions instead of imposing global constraints. The approaches [25]–[27] super-resolve local LR patches based on a weighted sum of exemplar facial patches in reference HR database. Liu et al. [28] develops a locality-constrained bi-layer network to jointly super-resolve LR faces as well as eliminate noise and outliers. Moreover, SIFT flow [29] and facial landmarks [30] are introduced to locate facial components for further super-resolution.

3.1 CHARACTERISTICS OF FACE IMAGES

Human face is a highly structured object with its own unique characteristics, which can be explored and utilized in FSR task. In this section, we simply introduce these facial characteristics. 3.1 Prior Information As shown in Fig. 1, structural priors can be found in face images, such as facial characteristics. 3.1.1 Prior Information As shown in Fig. 1, facial landmarks: These locate the key points of facial components. The number of landmarks varies in different datasets, such as CelebA [55], which provides five landmarks while Helen [56] offers 194 landmarks.

- Facial landmarks: These locate the key points of facial components. The number of landmarks varies in different datasets, such as CelebA [55], which provides five landmarks while Helen [56] offers 194 landmarks.

- Facial heatmaps: These are generated from facial landmarks. Facial landmarks give accurate points of the facial components, while heatmaps give the probability of the pointbeing a facial landmark. To generate the heatmaps, every landmark is represented by a Gaussian kernel centered on the location of the landmark. • Facial parsing maps: These are semantic segmentation maps of face images separating the facial components from face images, including eyes, nose, mouth, skin, ears, hair, and others. These face structure prior information can provide the location of facial components and facial structure information. We can expect to recover more reasonable target face images if we incorporate these prior knowledge to regularize or guide the FSR models.

IV. TANN

Our network has two components: (i) a transformative upsampling network, which transforms different poses to the frontal one and also super-resolves the frontalized LR feature maps; and (ii) a discriminative network, which forces the generated HR frontal faces to lie on the manifold of authentic HR face images. Figure 2 illustrates the overall architecture of TANN. In the training phase, the entire network is trained in an end-to-end fashion to compensate for possible artifacts induced by any of the frontalization and hallucination tasks. As shown in Fig. 3(k), when we train the upsampling network separately, i.e., generating frontalized LR faces as intermediate results, the transformer subnetwork may suffer from the loss of information contained in its feature maps because it is forced to output 3 channel LR faces as its objective function rather than 32 channel feature maps. This may lead to accumulated errors and obvious deviations in the output of the upsampling subnetwork due to the incorrect input images for upsampling. Thus, feeding 32 feature maps directly to the upsampling network is a better choice.

4.1 Transformative Upsampling Network (TUN)

In Fig. 2, our transformative upsampling network is shown (red box). TUN is composed of two parts: a transformer subnetwork and an upsampling subnetwork. The transformer part (purple box) aims at encoding non-frontal LR faces into latent representations which are close to the latent representations of their corresponding frontal LR ones. By doing so, we can achieve the latent codes of frontalized LR faces. Our transformer subnetwork is constructed by convolutional layers, a fully-connected layer, deconvolutional layers and spatial transformer layers. Since the input LR faces undergo in-plane rotations, translations and scale changes, multiple spatial transformer networks (STN) [3] are embedded as intermediate layers to compensate for such affine transformations. Moreover, because STNs learn 2D affine warps rather than out-of-plane rotations, they cannot recover self-occluded parts of faces. To solve this problem, our intuition is that we can project different views of a face into a subspace, where their encoded representations are enforced to lie close to the representations of their corresponding frontal one. Therefore, we incorporate a fully-connected layer to encode the feature maps of LR profile faces as well as design a triplet loss to force the similarity between the representations of LR profile and frontal ones. To illustrate the effectiveness of the transformer subnetwork, we change the channel number of its output layer to 3, and use LR frontal faces as ground-truth images to train this subnetwork. As shown in Fig. 3(j) and Fig. 4(d), it can successfully generate an LR frontal face image. Note that, when training our TANN, we do not employ LR frontal faces as supervision to prevent the aforementioned drift issue. After obtaining the feature maps of LR frontal faces generated by the transformer subnetwork, we apply an upsampling subnetwork (green box in Fig. 2) to hallucinate the high-frequency facial details of frontal faces. Because the resolution of LR input images is very low, STNs in our transformer subnetwork may not align LR faces accurately. The LR feature maps generated
by the transformer network may still contain misalignments. We employ the upsampling structure used in our previous works\cite{2, 26} for further alignment and super-resolution. As shown in Fig. 3(h), simply applying the method of \cite{2} to LR profile faces cannot provide high-quality HR frontal face images. This manifests that upsampling LR non-frontal faces with large pose variations is more difficult compared to LR frontal faces and also indicates the necessity of our transformer subnetwork. Since the mapping between common LR patterns and HR facial details can be easily learned from frontal faces, we frontalize LR inputs first and then hallucinate them.

4.2 Discriminative Network

As demonstrated in our previous works\cite{2, 22, 26}, only using Euclidean distance (pixel-wise $L_2$ loss) between the upsampled faces and the ground-truth HR faces tends to generate oversmoothed results. Therefore, a class-specific discriminative objective is also incorporated into our TUN, aiming to force the hallucinated HR face images to lie on the same manifold of real frontal face images. As shown in Fig. 2 (blue box), the discriminative network consists of convolutional layers, max-pooling layers, dropout layers, and fully-connected layers. It is designed to determine whether an image is sampled from real face images or the hallucinated ones. The discriminative loss, also known as adversarial loss, will be back-propagated to update the parameters of TUN as well. With the help of the adversarial loss, we can generate more realistic HR frontal faces. Figure 4 illustrates the impact of the adversarial loss on the final results.

4.3 Training Details of TANN

We construct LR profile and HR frontal ground-truth face image pairs $\{l_i, h_i\}$ for our training purpose, where $h_i$ represents the aligned frontal HR face images (only eyes are aligned), and $l_i$ is the synthesized LR side-view face images from $h_i$. For each HR frontal face image $h_i$, we generate five different views, i.e. $\{0^\circ, \pm 40^\circ, \pm 75^\circ\}$, to construct LR/HR training pairs. Using these five distinct poses is a trade-off between a sufficient coverage of pose variations and the reasonable size of the training dataset and also suggested in \cite{31}. More details are provided in Sec. 4. In training our TANN, we not only enforce the conventional pixel-wise intensity similarity, known as pixel-wise $L_2$ loss, but also the feature-wise similarity, known as perceptual loss \cite{28}, to obtain high-quality results. Similar to the works\cite{22, 26}, the adversarial loss is also employed to attain visually appealing frontalized HR face images. As mentioned in Sec. 3.1, we also develop a triplet loss to force the representations of LR profile faces to be similar to the representations of their frontal faces. In this manner, we can frontalize LR profile faces without degrading super-resolution of frontal ones. Pixel-wise intensity similarity loss: We constrain the generated HR frontalized face $h_i$ to be similar to its ground-truth frontal counterpart $h_i$ in terms of image intensities. Thus we employ a pixel-wise $L_2$ regression loss $\mathcal{L}_{\text{pix}}$ to impose the appearance similarity constraint, expressed as:

$$
\mathcal{L}_{\text{pix}} = \mathbb{E}_{(h_i, l_i) \sim p(h, l)} [ | h_i - l_i |^2 ],
$$

where $t$ and $T$ are the parameters and the output of TUN, $p(h, h')$ represents the joint distribution of the frontalized HR faces and their corresponding frontal HR ground-truths, and $p(l, h)$ indicates the joint distribution of the LR and HR face images in the training dataset. Feature-wise similarity loss: As mentioned in \cite{22}, pixel-wise $L_2$ loss leads to over-smoothed super-resolved results. Here, we employ a feature-wise similarity loss, known as perceptual loss \cite{28}, to constrain the super-resolved HR faces to share the same facial details as their ground-truth counterparts, thus attaining high-quality results with rich facial details. The perceptual loss $\mathcal{L}_{\text{perm}}$ measures Euclidean distance between the feature maps of HR frontalized and ground-truth faces extracted by a deep neural network, written as:

$$
\mathcal{L}_{\text{perm}} = \mathbb{E}_{(h_i, l_i) \sim p(h, l)} [ \Phi(h_i) - \Phi(l_i) ]^2,
$$

where $\Phi(\cdot)$ denotes feature maps extracted by the ReLU32 layer in VGG-19 \cite{54}, which gives good empirical performance in our experiments.

4.4 Hallucinating Frontal HR from Non-frontal LR

The discriminative network is only employed in the training phase. In the testing phase, we feed an unaligned LR profile face image into the transformative upsampling network to obtain its upright and frontal HR version. Note that, only in the training stage, we need to feed the network with triplet samples due to employing the triplet loss. In the testing stage, our network is able to super-resolve and frontalize a single image. Since aligned HR frontal face images are employed as ground-truths, TUN will output aligned and frontalized HR faces directly. As a result, our method does not need to estimate the face orientations or align very low-resolution images beforehand, and provides an end-to-end and highly nonlinear mapping from an unaligned LR profile face image to its frontal HR version.

Fig. 2. TANN consists of two parts: a transformative upsampling network (red box) and a discriminative network (blue box). In our transformer subnetwork, we also employ skip connections between our encoding layers and decoding layers, indicated by the purple line. For simplicity, we only draw the first skip connections.

Fig. 3. Illustrations of influence of different losses. (a) The input 16x16 LR images. (b) The original 128x128 HR frontal images. (c) The downsampled version of (b). (d) The frontalized LR faces by our transformer subnetwork. (e) The upsampling results only using pixel-wise loss. (f) The upsampling results using the pixel-wise and perceptual losses. (g) The upsampling results without using the triplet loss. (h) Our final results.

4.5 Super-resolving LR faces without Frontalization
Since our method is an extension of our previous face superresolution methods our network can be also applied to super-resolve LR face images without frontalizing them. To this end, we use the ground-truth HR faces that have the same poses as the input LR ones as supervision and remove the triplet loss in training. As seen in Fig. 14, after retraining our network, we can effectively super-resolve LR faces in different poses, similar to our previous methods.

4.6 Limitations
Since our method uses a generic face model to generate faces in different poses, we do not contain different expressions in the training dataset. Therefore, our network does not account for different facial expressions. Furthermore, limited by the generic 3D face model, we do not model eyeglasses or sun-glasses in the training dataset either. When frontalizing occluded regions, general occluded regions and sun-glasses should yield different frontalization results due to the symmetry of sun-glasses and asymmetry of general occlusions. This may introduce further ambiguity in the frontalization process without exploiting any high-level semantic information. Our training dataset is generated from face images captured in normal illumination conditions where facial landmarks can be detected for generating different poses. Since facial landmark detectors may fail to localize landmarks accurately under extreme illumination conditions and the generated faces by the 3D model may suffer severe artifacts, we do not contain those faces for training.

V. EXPECTED OUTCOME
In this concept we are going to make use of deep learning and GAN architecture to develop the model where the input image is given and the face image is hallucinated using the train model. The image with missing pixels is filled using algorithms. Obtaining a high-quality frontal face image from a low-resolution (LR) non-frontal face image is primarily important for many facial analysis applications. Face recognition and expression classification, compared with other state-of-the-art methods.

VI. CONCLUSION
We have developed a transformative adversarial network that is capable of simultaneously upsampling and frontalizing low-resolution, unaligned face images in an end-to-end manner. Our network has the ability to learn the frontalization and alignment of LR faces while upsampling by a factor of 8x. By leveraging our proposed triplet loss, we are able to ensure that LR profile faces are closely aligned with their frontal counterparts in the latent subspace, resulting in improved frontalization performance. Additionally, our network benefits from intra-class discriminative information and feature constraints, which help generate realistic facial details in the generated images.

REFERENCES
[1] Y. Liang, J. H. Lai, W. S. Zheng, and Z. Cai. A survey on face hallucination. In CCBR, pages 83–93, 2012.
[2] N. Wang, D. Tao, X. Gao, X. Li, and J. Li. A comprehensive survey to face hallucination. International Journal of Computer Vision, 106(1):9–30, 2014.
[3] M. P. Autee, M. S. Mehta, M. S. Desai, V. Sawant, and A. Nagare. A review of various approaches to face hallucination. Procedia Computer Science, 45:361–369, 2015.
[4] S. Kanakaraj, V. K. Govindan, and S. Kalady. Face superresolution: A survey. International Journal of Image, Graphics and Signal Processing, 9:54–67, 05 2017.
[5] K. Nguyen, C. Fookes, S. Sridharan, M. Tistarelli, and M. Nixon. Super-resolution for biometrics: A comprehensive survey. Pattern Recognition, 78:23–42, 2018.
[6] J. Li, S. Liu, J. Yang, and M.-H. Yang. “Generative face completion,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Honolulu, HI, USA, Jul. 2017, pp. 3911–3919.
[7] C. Dong, C. C. Loy, K. He, and X. Tang. “Learning a deep convolutional network for image super-resolution,” in Proc. Eur. Conf. Comput. Vis. Zurich, Switzerland: Springer,2014, pp. 184–199.
[8] J. Kim, J. K. Lee, and K. M. Lee. “Accurate image super-resolution using very deep convolutional networks,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, Jun. 2016, pp. 1646–1654.
[9] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi. “Photo-realistic single image super-resolution using a generative adversarial network,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Honolulu, HI, USA, Jul. 2017, pp. 4681–4690.
[10] M. S. Rad, B. Bozorgtabar, U.-V. Marti, M. Basler, H. K.Ekenel, and J.-P. Thiran. “SROBB: Targeted perceptual loss for single image super-resolution,” in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), Seoul, South Korea, Oct. 2019, pp. 2710–2719.
[11] J. Kim, J. K. Lee, and K. M. Lee. “Deeply-recursive convolutional network for image super-resolution,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, Jun. 2016, pp. 1637–1645.
[12] Y. Chen, Y. Tai, X. Liu, C. Shen, and J. Yang. “FSRNNet: End-to-end learning face super-resolution with facial priors,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Salt Lake City, UT, USA, Jun. 2018, pp. 2492–2501.