A Prior-Guided Face Image Super-Resolution Network Based on Attention Mechanism

Yun Liu, Remina Yano, Hiroshi Watanabe
Waseda University
Takuya Suzuki, Zheming Fan, Takeshi Chujoh, Tomohiro Ikai
Sharp Corporation

***Abstract*—In this paper, we propose a face image super-resolution method, aiming to reconstruct high quality face images from a low resolution input. The proposed method introduces an attentional multi-scale feature fusion block. In addition, the facial prior information is utilized by adding a separate prior branch. Experiments show that the face images reconstructed by the proposed method exhibit noticeable quality improvement compared to the low-resolution images.**

Keywords—face super-resolution; channel attention; face prior

I. Introduction

Due to the limitation of camera sensor hardware or the distance between the camera and the target, most of the captured face images are of poor quality. By using face super-resolution, we can obtain high-resolution face images, thus providing help for subsequent tasks like face detection, face recognition, etc. Relative to general images, face images have distinctive features (e.g., the positions of facial organs are approximately the same), which can be applied in super-resolution tasks.

II. Proposed Method And Experiments

In this paper, we adopt the structure of SRResNet [1] and modify its res-block by introducing the channel attention mechanism. The whole structure of the model is shown in Figure 1.

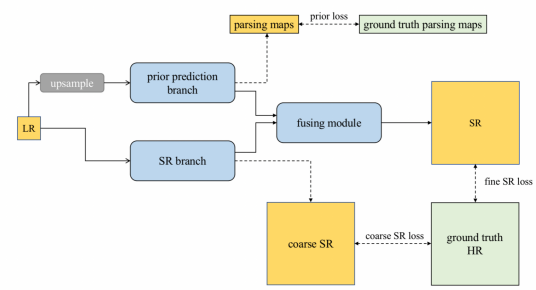

Fig. 1. The structure of proposed method.

The basic block of SR branch is shown in Figure 2, we use convolution kernels of different sizes in the feature extraction part and fuse the features extracted from each convolution layer to improve the feature extraction capability of the model. Softmax operator is used to calculate the weights of each channel and these weights are multiplied with the feature map, so as to emphasize the useful channel features and suppress the useless ones. In the prior branch, the hourglass structure is used to predict the parsing map of the face image. Low resolution face images provide limited information, it will be helpful to reconstruct the facial components if the accurate parsing map can be predicted and used as the prior information.

The Fusing module is used to fuse the information from the two branches for the final information integration and reconstruction. It is worth mentioning that the super-resolution branch and the prior prediction branch are trained simultaneously and the whole network is end-to-end.

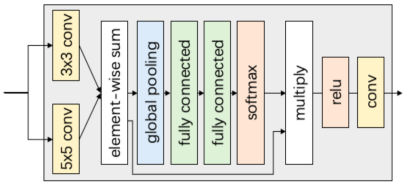

Fig. 2. Basic block of SR branch

We use 28000 images from CelebAMask-HQ Dataset [2] to train the model and 1000 images are used for testing. The evaluation metrics are PSNR and SSIM, results (8x) are shown in TABLE I.

TABLE I. Experimental Results

Method	PSNR	SSIM
BICUBIC	*20.72*	*0.5200*
VDSR[3]	*20.08*	*0.4972*
SRResNet[1]	*24.47*	*0.7201*
PFSR[4]	*22.73*	*0.6459*
EIPNET[5]	*24.47*	***0.7284***
Proposed	***24.66***	*0.7230*

III. Conclusion

We proposed a prior-guided face super-resolution method, by introducing channel attention and prior information, high quality face images can be reconstructed.

References

[1] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, "Photo-realistic single image super-resolution using a generative adversarial network," in CVPR, pp. 105-114, Jul. 2017.

[2] C. H. Lee, Z. Liu, L. Wu, P. Luo, Maskgan, "Towards diverse and interactive facial image manipulation," in IEEE/CVF CVPR, pp. 5549-5558, Jun. 2020.

[3] J. Kim, J. K. Lee, K. M. Lee, "Accurate image super-resolution using very deep convolutional network.," in CVPR, pp. 1646-1654, Jun. 2016.

[4] D. Kim, M. Kim, G.Kwon, D. S. Kim, Progressive face super-resolution via attention to facial landmark. arXiv preprint arXiv:1908.08239, 2019.

[5] J. Kim, G. Li, I. Yun, C. Jung, J.Kim, "Edge and identity preserving network for face super-resolution," in Neurocomputing, 2021.