2019 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)

Single Image Super-resolution Reconstruction of Enhanced Loss Function with Multi-GPU Training

Jianqiang Huang[a,b,*], Kai Li[a,b], Xiaoying Wang[a]
[a]*Department of Computer Technology and Application*
Qinghai University, China
[b]*Department of Computer Science and Technology*
Tsinghua University, China
hjq16@mails.tsinghua.edu.cn, qhu.kaili@gmail.com, wangxiaofu163@163.com
[*]*Corresponding Author: Jianqiang Huang*
Jianqiang Huang and Kai Li contributed equally to the writing of this article.

***Abstract*—According to research on super-resolution (SR), SR image reconstruction using generated anti-networks can produce images that are more realistic than using convolutional neural networks. At present, SR technology based on convolutional neural networks ignores the impact of loss function on image reconstruction; the results lack detail and accuracy. In this paper, we use SR method and combine Generative Adversarial Networks to design a super-resolution (Lapras-GAN) model of the enhanced loss function. The proposed enhancement loss function is a Mix loss function that combines the multiscale SSIM and L_1 loss functions to obtain realistic images. We performed qualitative and quantitative analysis of the performance of different loss functions and demonstrated the advantages of the Mix loss function. In addition, the neural network is accelerated by multiple GPUs of multiple nodes, which can be 3-4 times faster than a single node single GPU. Experimental results show that the proposed Lapras-GAN method can generate images consistent with images produced by human perception. Further comparisons show that our Lapras-GAN has excellent performance and test time in the PIRM2018 experimental test data set. Finally, we obtained a perception index of 1.83 and a test time of 0.031s in the PIRM2018 competition test set.**

***Keywords*-Deep Learning, Generative Adversarial Networks, Loss Function, GPU, Single Image Super Resolution**

I. Introduction

Super resolution (SR) refers to the technique of recovering high-resolution (HR) images from low-resolution (LR) or sequential images. Single-image SR (SISR) reconstruction technology is widely used in hyperspectral imaging, medical imaging, satellite remote sensing, and other fields.

Traditional SISR methods include interpolation-based, reco- nstruction-based, and learning methods that learn either the potential internal similarities of the same image or the mapping functions of external LR and HR sample pairs. Due to the recent breakthroughs in deep learning in other computer vision fields, several researchers have attempted to construct deep networks to conduct end-to-end training and introduce deep neural networks to solve the problem of SR image reconstruction.

Commonly used deep learning models can be divided into models based on interpolation preprocessing, original image processing, hierarchical features, and high-frequency details of image input information. However, advanced methods are based on a single-loss function, and SR image reconstruction methods employing the Mix loss function are rarely studied. We propose an enhanced loss function network for the proposed SR (enhanced loss function network for SR; Lapras-GAN) method that combines the L_1 and multi-scale structural similarity index (MS-SSIM) loss functions.

Most traditional SR methods use the L_2 loss function [1]–[3] for loss calculation because different depth learning frameworks provide the L_2 loss function method, and the theoretical results obtained using this function are generally good. When SSIM evaluates the image quality, a certain correlation between the human perception system and the local structure of the image is indicated. Some studies have shown that the local quality of an image has a certain relation with the distance between the image and the human body. Therefore, MS-SSIM [4] has been proposed to address these problems, and SSIM is calculated and weighted based on perceptions of the image by the human eye at different scales. In this work, we combine the MS-SSIM and L_1 loss functions as a Mix loss function to design an SR model and enhance the final loss function. We conduct generative adversarial net (GAN) optimization and post-process the generated images. Our contributions can be summarized as follows:

(1) We apply the Mix loss function [5] to the GAN model for image super-resolution reconstruction and discuss its feasibility. To the best of our knowledge, this is the first attempt in the single-image super-resolution field.
(2) The enhanced SR GAN (SRGAN) is modified and a larger HR patch is used to supplement the details of the Lapras-GAN input image block. Experiments have shown that larger HR patches can be used to extract larger image features.

978-1-7281-4328-6/19/$31.00 ©2019 IEEE

DOI 10.1109/ISPA-BDCloud-SustainCom-SocialCom48970.2019.00085

Authorized licensed use limited to: Tsinghua University. Downloaded on June 04,2022 at 17:14:00 UTC from IEEE Xplore. Restrictions apply.

(3) A multi-node parallel processing method for SR images generated by Lapras-GAN is implemented. Accelerate image super-resolution training time by using multiple GPU computing resources for multiple nodes.

II. Related model

Deep learning has achieved good results in solving computer vision problems. This article focuses on solving image SR problems through deep learning. In this section, we briefly review several current key points on SR.

Secondly, in the process of using the generated anti-network training, a large number of GPUs are also needed for acceleration. Among them, Meng et al. [6] proposed how to train a deeper model through GPU memory optimization on Tensorflow. Since the training of deep learning is performed at the nodes of the cluster, Zhang et al. [7] propose an efficient communication architecture for distributed deep learning of multi-GPU clusters. Specific distributed deep learning training will be introduced in model optimization.

The mainstream methods used for SR image reconstruction based on convolutional neural networks [8], include SR convolutional neural network (SRCNN) [9], sparse coding network (SCN) [10], very deep SR (VDSR) [11], deeply recursive convolutional network [12] and fast SRCNN [13]. Although all of these methods produce good image SR effects, the CNN and RNN based methods learn the mapping features of LR images to SR images from LR image interpolation, resulting in smoother images and lack of details and texture of the original image; unpleasant artifacts may also occur in deeper neural networks. The perceptual and content losses proposed by GANs can solve the problems generated by CNN and RNN so that the obtained SR image conforms better to that of the human perception system in terms of detail and texture. The SRGAN [14] proposed by Ledig et al. solves the shortcomings of the results of conventional and deep learning methods, which generally lack high-frequency information and details. In Section III, we explain how we use GANs in detail.

III. Method

SRGAN [14] is the first model proposed to improve the task of using the GAN network to solve image super-resolution. As mentioned earlier, the magnified image of SRGAN is often accompanied by unpleasant artifacts. Therefore, our proposed Lapras-GAN model draws on the SRGAN model to improve the quality of the generated image.

A. Innovation of Method

First, we use a GAN [8] as the basic framework for Lapras-GAN. The use of GAN for SR addresses the shortcoming of a lack of high-frequency information and detail in the results of traditional methods, including deep learning methods. Textures may be relatively simple, but they always have great detail, the visual feel is better, and the details are rich.

Compared with SRGAN [14], Lapras-GAN presents some innovations in the GAN model. The original GAN framework is changed to RaGAN (relativistic average GAN) [15] because the discriminator should also reduce the probability that "the actual data are true" and increase the probability of "the pseudo-data are true" to determine whether the image is more realistic than others rather than whether the image is true or false. RaGAN also helps learn sharper edges and finer textures.

Lapras-GAN removes the batch normalization [16] layer from SRGAN to achieve consistent performance without artifacts. It does not degrade performance but saves computing resources and memory usage. Models with batch normalization layers are likely to introduce unpleasant artifacts.

In the Lapras-GAN generation network section, SRResNet [14] is used as the basic network architecture. By calculating most of the calculations to obtain the eigenvalues of the LR image, the amount of calculation required to produce image feature values is reduced, thereby accelerating the training time of the SR model. As shown in Fig. 1, each residual block contains two 3x3 convolutional layers. After the convolutional layer, PReLU acts as an activation function and uses two 2x sub-pixel convolution layers [17] to increase the feature size.

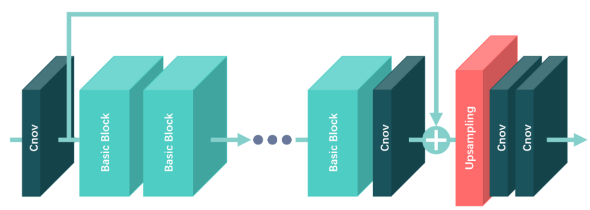

Figure 1. SRResNet uses Parametric ReLU to help it adaptively learn some of the negative coefficients; SRResNet uses a sub-pixel convolutional layer for image upsampling.

Sub-pixel convolution layers are also used for upsampling in the generation network. As shown in Fig. 2, the types of layers included in each module of the discrimination network include a convolution layer, a Leaky ReLU layer, and a Batch Normalization layer.

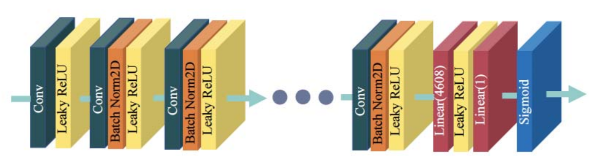

Figure 2. We use the VGG19 network to link two fully connected layers, the Leaky ReLU function and the Sigmoid function. The network is finally used as the discriminant network of the Lapras-GAN.

Authorized licensed use limited to: Tsinghua University. Downloaded on June 04,2022 at 17:14:00 UTC from IEEE Xplore. Restrictions apply.

IV. Optimizing direction

As research on computer version is basically related to images, the optimization direction should be adjusted around the relevant parameters of the image. Optimization of the structure of the neural network is also necessary. In this section, we introduce the optimization direction for Lapras-GAN.

A. HR patch-size

We have found through extensive experimentation that using larger patch sizes can achieve better image quality when training deep neural networks, because the expanded image sensing range can help deep neural networks capture more useful image information. We changed the HR patch-size to 128x128 and 192x192 on the basis of 10 and 23 RRDB models, respectively, and used the same training set (Urban100) for training. The results are shown in Figs. 3 and 4. Two neural networks of different depths were observed to achieve good image quality with large patch sizes.

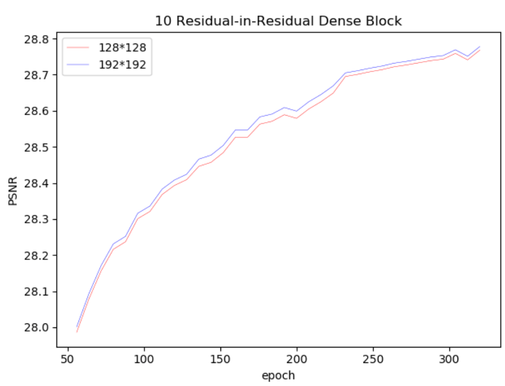

Figure 3. The number of RRDBs is 10 different patch-size renderings.

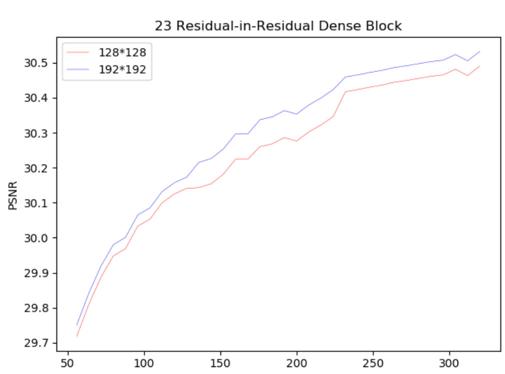

Figure 4. The number of RRDBs is 23 different patch-size renderings.

B. Loss function

The loss function is the basis for judging the predictions of the neural network. The following mainly explains the L_1, L_2, and Mix loss functions used in Lapras-GAN. The L_1 loss function considers the sum of the absolute values of the difference between the target variable and the predicted value. Therefore, to achieve image quality that is better than that of human perception, we combine L_1 and SSIM to obtain the Mix loss function. Equation (1) shows the SSIM loss function.

$$L^{SSIM}(P) = 1 - SSIM(\widetilde{p}) \tag{1}$$

The MS-SSIM method [4] is proposed to combine image details under different resolutions and observation conditions into a quality evaluation algorithm. Equation (2) reveals the MS-SSIM loss function.

$$L^{MS-SSIM}(P) = 1 - MS - SSIM(\widetilde{p}) \tag{2}$$

MS-SSIM and SSIM are not particularly sensitive to uniform deviations, which can result in brightness changes or color shifts. L_1 preserves the weighting error of color and brightness regardless of the local structure, but it does not produce exactly the same contrast as MS-SSIM. The Mix loss function [5] is obtained by combining MS-SSIM and L_1, as proposed by Hang Zhao et al. Equation (3) describes the Mix loss function. We will test it in the subsequent experiments section.

$$L^{Mix} = \alpha L^{MS-SSIM} + (1-\alpha) G_{\sigma_G^M} L^{l_1} \tag{3}$$

C. Distributed computing

We use Pytorch's distributed computing approach to divide small batches into smaller samples and run each smaller batch of samples in parallel. We run the adjusted parameters and structure of Lapras-GAN on a card and do not join the Pytorch parallelization module. Then, we use torch.nn.DataParallel in Pytorch for data parallelism, which can be parallelized on multiple GPUs in a batch dimension. Pytorch's current recommendation is a multi-process single GPU implemented using torch.nn.DistributedDataParallel. In this work, we consider a single-process single GPU, a multi-process single GPU, and a multi-process multi-GPU (notably, the cluster node currently only has two graphics cards). Table 2 shows the times we calculated using Lapras-GAN for the same data size. The training time obtained using multiple GPUs and multiple processes is much smaller than that obtained using a single GPU single process.

Table I
Lapras-GAN training time results of different training methods.

Result	Single-process single-GPU	Multi-process single GPU	Multi-process multi-GPU
Time(h)	300.96	195.6	125.04

Authorized licensed use limited to: Tsinghua University. Downloaded on June 04,2022 at 17:14:00 UTC from IEEE Xplore. Restrictions apply.

The training time results demonstrate that better performance is obtained under a multi-process multi-GPU setup by using distributed computing because, in the case of multi-processing, full use of the computing performance of each GPU can be made to achieve the fastest training time. We also consider whether the training time can be further shortened in the case of multi-process multi-GPU between multiple nodes. However, because our cluster does not include a large number GPUs, it cannot verify the proposed multi-level multi-GPU.

V. Experiment

In this section, we illustrate the dataset and hyperparameters of our model training. Then, we compare our optimized Lapras-GAN with several other models using multiple benchmark datasets.

A. Data set

First, we used the DIV2K and Flickr datasets to obtain a total in 3,450 2K images for training dataset selection. We enriched the training set by merging these datasets. We used Set5, Set14, Urban100 [18], and the dataset provided in PIRM 2018 [19] for verification.

B. Training details

The SR image is calculated under the scale factor of x4. Thus, we used Matlab to downsample the HR image and obtain the LR image. The generation network of the Lapras-GAN model is SRResNet, the structure is Conv-Activate, and the RRDB is set to 23. Through experiments, we found that more RRDBs yield clearer textures; unfortunately, more time is needed to complete the training. Thus, we selected 23 RRDB modules to achieve the fastest training time without affecting the results. A large HR patch-size can force the model to learn richer image feature information, but it will also cause the training speed to become too slow. Thus, we chose a relatively large HR patch-size of 192x192 and a multi-process multi-GPU to speed up the training time.

We used the Pytorch framework to implement our Lapras-GAN method in this SR experiment and six NVIDIA Tesla V100 Tensor Cores for training.

C. Experimental results

We performed SR reconstruction of a baseline dataset using the EDSR model optimized for PSNR and WDSR [20] and the SRGAN, ESRGAN, and Lapras-GAN models for vision system optimization. In addition, because no standard for evaluating image quality currently exists, we chose the PIRM official calculation indicators, PI, and RMSE to determine quality. Fig. 7 shows the results of different models on the same picture.

As can be seen in Fig. 5, in contrast to these techniques, our method produces a visually pleasing SR image that is similar to a real image without much noise or smooth image details and textures.

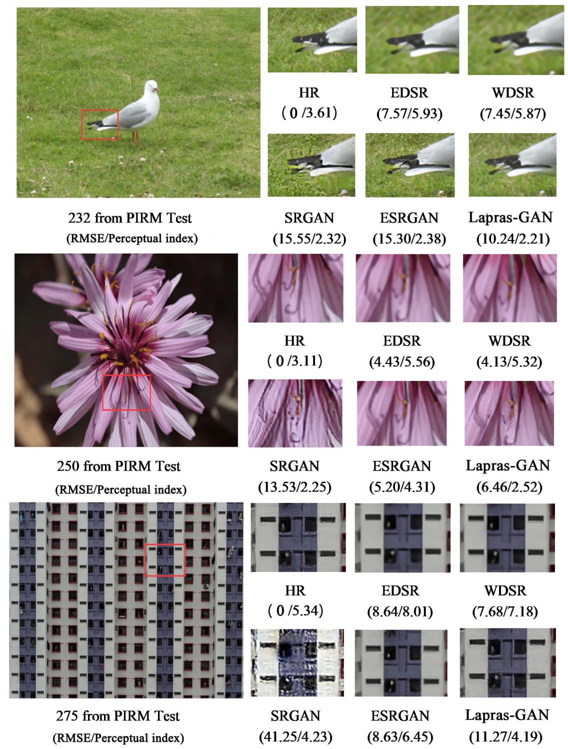

Figure 5. Qualitative results of Lapras-GAN. Lapras-GAN produces realistic textures and details. Although SRGAN can achieve very low perceptual indices, it also produces many unrealistic artifacts.

Fig. 5 reveals that Lapras-GAN optimized from SRGAN will be better in texture and detail than SRGAN and other models. From the different results obtained from image 232, we can see that the results processed in SRGAN and Lapras-GAN have clearer textures and details than the results of the other models. The GAN-based SR reconstruction model generates x4 SR images but often produces unnatural textures and noise that does not conform to the vision system. CNN-based models (EDSR, WDSR) yield image details that are often too smooth and lack realism of detail because they focus on PSNR optimization. To verify the effect of Lapras-GAN, we applied it to different test sets and show the results in Table 3.

We calculated the test times for different models in the same hardware, as shown in Table 4. We calculated the test times for different models and observed that the Lapras-GAN method we designed presents a certain advantage in terms of computational performance and speed under the same hardware configuration, mainly because we use a large HR patch and a skip connection to deepen our network and stabilize training.Since we used skip connections to get a larger receptive field, we reduced the number of discriminant network layers of Lapras-GAN and accelerated the test speed without changing the image quality. After removal of the batch normalization layer, training can save about 40% of

Authorized licensed use limited to: Tsinghua University. Downloaded on June 04,2022 at 17:14:00 UTC from IEEE Xplore. Restrictions apply.

Table II
RESULTS OF DIFFERENT SR MODELS.

Model	Training Set	Valid Set	Test Set	Perceptual Index	RMSE
ESRGAN	DIV2K	DIV2K	PIRM	2.1383	14.2212
Lapras-GAN	DF2K	DIV2K	PIRM	1.8671	16.884
WDSR_A	DIV2K	Urban100+Set5	PIRM	3.8012	14.4879
RDN	DIV2K	Set5	Urban100	4.3211	14.974
EDSR	NONE	NONE	PIRM	4.9035	10.7299

the space to speed up the test time.

Table III
THE TIME IT TAKES FOR DIFFERENT MODELS TO PERFORM SUPER-RESOLUTION TESTING ON A SINGLE IMAGE AND 100 IMAGES.

Model	Single Image	100 Images
EDSR	0.144s	13.123s
WDSR	0.102s	9.832s
SRGAN	0.152s	14.853s
ESRGAN	0.098s	8.365s
Lapras-GAN	0.031s	4.421s

By analyzing the experimental results, it can be found that the GAN model produces over-interpreted image textures and details when generating SR images, which leads that the perception index of the images is increased (the lower the index value, the better the image quality). We use the back projection method to post-process the SR images. The back projection method records the pixel fit of given images as a distribution of pixels in the histogram model. By combining it with the SR images, we have improved the back projection method. First, the result of the deep learning model is downsampled by x4 to obtain an image with the same resolution as the LR image. The pixel-by-pixel comparison with the LR image is then used to eliminate the details and textures produced by over-learning of the model. Finally, by up-sampling the processed result, the image generated by the model is reconstructed twice to generate a back-projected SR image. Figure 6 shows the results obtained by back projection.

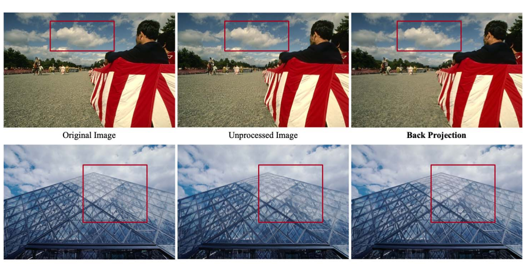

Figure 6. The back projection method can eliminate the texture and details generated by GAN over-improvement to improve image perception quality.

We conducted several experiments and concluded that the above results are consistent. The use of GAN provides richer details and textures for the generated SR images, but it also introduces so much noise that the RMSE results are relatively large. Using the improved back projection method to post-process the results, we obtained more natural details without introducing too much noise. Therefore, in this paper, we use back projection as an image post-processing operation.

D. Loss Function

In the SISR method, the mainstream method for evaluating image quality is still PSNR and SSIM. As mentioned above in the third account, these methods do not conform well to human perception systems. Here, we used the Mix loss function instead of the L_1 loss function only because MS-SSIM shows the contrast of images well while L_1 addresses issues related to color and brightness. The Fig. 7 below shows the experimental result we obtained for different loss functions.

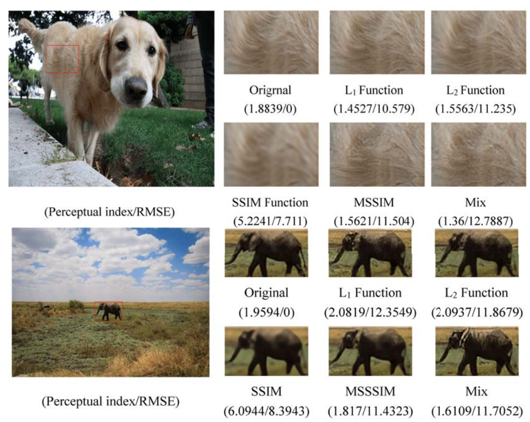

Figure 7. The experimental results of different loss functions, through the perceptual index, can prove that the use of the Mix loss function can restore more natural image details and textures.

The experimental results show that the PI obtained using the L_1 loss function is smaller than those obtained from L_2 and SSIM; in addition, the brightness of the picture tested by MS-SSIM is better than that produced by other methods. Experimental results further show that the PI obtained by using the Mix loss function is better than that provided by other functions.

To prove that our Lapras-GAN method is also suitable for LR images taken in natural conditions, we used the CameraSR [21] dataset and an LR image taken by an iPhone

Authorized licensed use limited to: Tsinghua University. Downloaded on June 04,2022 at 17:14:00 UTC from IEEE Xplore. Restrictions apply.

Figure 8. Low-resolution image processing results for natural shots using different loss functions. The results of L_2 and SSIM processing are ambiguous and lack detail, while those of L_1 and MS-SSIM recover some of the image details.

X as the test set. Fig. 8 shows the results obtained using Lapras-GAN with different loss functions.

The experimental results confirm that the loss function has a positive effect on the LR image reconstruction of the focal length imaging of an optical lens. According to the experimental results, the L_1, MS-SSIM, and Mix loss functions can better restore image details and texture. The Mix loss function combines the advantages of the L_1 and MS-SSIM loss functions, and the restored SR images show brightness, contrast, and structures consistent with those of LR images; greater image details are also captured.

In addition, we conducted some researches on texture loss and content loss, and optimized the feature loss and content loss of the SRGAN model generation network. According to the experimental result, it is found that the application of the Mix loss function to calculate the feature loss and content loss is more accurate, and the generated image has richer texture and detail. In this case, we compared the performance of each network model and introduced the combinations of different loss functions concretely in Table 5. Where L_1 represents the L_1 loss function, L_P is the perceptual loss, Mix is the Mixed loss function, and L_A is the counter loss. The Fig. 9 below shows the results of a comparison of different tests on loss function.

The calculation method of perceptual loss and adversarial loss is Equation 4 and Equation 5. Equation 4 is the perceived loss of the Lapras-GAN model. Where Φ is the calculation function of the feature map. Equation 5 is the adversarial loss calculation method for the Lapras-GAN model. Where D is the discriminating network and G is the generating network.

$$\mathcal{L}_P = \|\Phi(I_{GT}) - \Phi(I_{HR})\|_2^2 \quad (4)$$

$$\mathcal{L}_A = -\log_{10}(D(G(x))) \quad (5)$$

Table IV
TRAIN THE SAME NETWORK WITH DIFFERENT LOSS FUNCTIONS.

Networks	Loss Function	Description
Lapras-GAN-$\mathcal{L}_1$	$\mathcal{L}_1$	Only $\mathcal{L}_1$ Function
Lapras-GAN-P	Mix	Mix Loss Function
Lapras-GAN-LP	$\mathcal{L}_1$ + $\mathcal{L}_P$	$\mathcal{L}_1$ + Perceptual
Lapras-GAN-MP	Mix + $\mathcal{L}_P$	Mix + Perceptual
Lapras-GAN-LPA	$\mathcal{L}_1$ + $\mathcal{L}_P$ + $\mathcal{L}_A$	Lapras-GAN-LP + Adversarial
Lapras-GAN-MPA	Mix + $\mathcal{L}_P$ + $\mathcal{L}_A$	Lapras-GAN-MP + Adversarial

This experiment performed different super-resolution reconstructions on images from Urban100 at 4x magnification. The results of the Lapras-GAN-P appear to be clearer than the Lapras-GAN-L_1 texture, but they all appear blurry. The image produced by Lapras-GAN-MP is very clear but partially blurred. It is found that Lapras-GAN-MPA can generate richer textures, making the resulting image more accordant with the original HR image.

VI. Summary

We studied several excellent SISR models proposed in recent years and compared their respective characteristics. To the best of our knowledge, this is the first application of the single-image super-resolution of the Mix loss function in the GAN model. Through extensive research, we found that using the hybrid loss function has a positive impact on image recovery. We designed an Lapras-GAN method that produces images with more natural detail and texture than those obtained from previous SR models. We developed a basic structure of Lapras-GAN based on SRGAN and introduced different optimization directions for the Lapras-GAN model. We focused on the effect of the loss function on the SR image and are more realistic for the generator to use the blended-loss function than the image generated by the single-loss function. Finally, experiments were conducted on simulated and real LR datasets. The experimental results show that Lapras-GAN can produce better SR images for natural focal length and analog images and exceeds the results of SR models with simulated LR images.

Acknowledgement

The authors are grateful to the reviewers for valuable comments that have greatly improved the paper. This paper is partially supported by the national Natural Science Foundation of China (No.61762074,No.61962051), National Natural Science Foundation of Qinghai Province (No.2019-ZJ-7034)."Qinghai Province High-end Innovative Thousand Talents Program - Leading Talents" Project Support. The Open Project of State Key Laboratory of Plateau Ecology and Agriculture, Qinghai University (2020-ZZ-03). National

Authorized licensed use limited to: Tsinghua University. Downloaded on June 04,2022 at 17:14:00 UTC from IEEE Xplore. Restrictions apply.

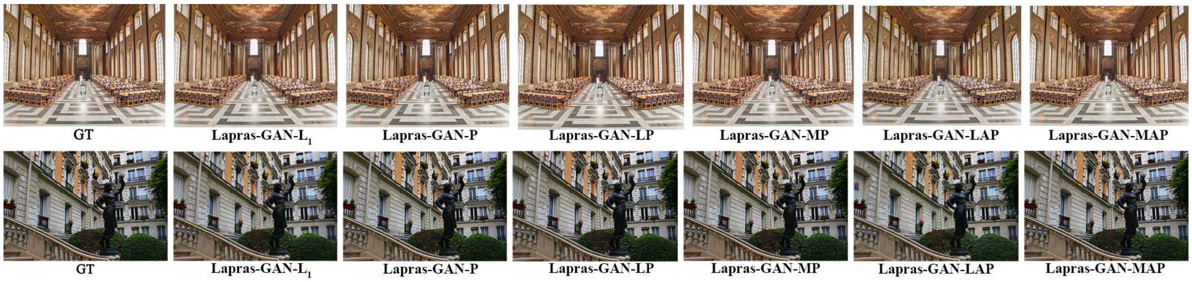

Figure 9. Compare the results of the different loss combinations of the Lapras-GAN model.

College Students Innovation and Entrepreneurship Training Program (No.201910743002)

REFERENCES

[1] Xudong Mao, Qing Li, Haoran Xie, Raymond YK Lau, and Zhen Wang. Multi-class generative adversarial networks with the l2 loss function. *arXiv preprint arXiv:1611.04076*, 5, 2016.

[2] Shuang Zhang, Ada Zhen, and Robert L Stevenson. Gan based image deblurring using dark channel prior. *arXiv preprint arXiv:1903.00107*, 2019.

[3] Yangchun Qing, Yafei Zhao, Yongjie Shi, Dan Chen, Yining Lin, and Yao Peng. Improve cross-domain face recognition with ibn-block. In *2018 IEEE International Conference on Big Data (Big Data)*, pages 4613–4618. IEEE, 2018.

[4] David M Rouse and Sheila S Hemami. Analyzing the role of visual structure in the recognition of natural image content with multi-scale ssim. In *Human Vision and Electronic Imaging XIII*, volume 6806, page 680615. International Society for Optics and Photonics, 2008.

[5] Hang Zhao, Orazio Gallo, Iuri Frosio, and Jan Kautz. Loss functions for neural networks for image processing. *arXiv preprint arXiv:1511.08861*, 2015.

[6] Chen Meng, Minmin Sun, Jun Yang, Minghui Qiu, and Yang Gu. Training deeper models by gpu memory optimization on tensorflow. In *Proc. of ML Systems Workshop in NIPS*, 2017.

[7] Hao Zhang, Zeyu Zheng, Shizhen Xu, Wei Dai, Qirong Ho, Xiaodan Liang, Zhiting Hu, Jinliang Wei, Pengtao Xie, and Eric P Xing. Poseidon: An efficient communication architecture for distributed deep learning on {GPU} clusters. In *2017 {USENIX} Annual Technical Conference ({USENIX} {ATC} 17)*, pages 181–193, 2017.

[8] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.

[9] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2):295–307, 2016.

[10] Zhaowen Wang, Ding Liu, Jianchao Yang, Wei Han, and Thomas Huang. Deep networks for image super-resolution with sparse prior. In *Proceedings of the IEEE international conference on computer vision*, pages 370–378, 2015.

[11] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. Accurate image super-resolution using very deep convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1646–1654, 2016.

[12] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. Deeply-recursive convolutional network for image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1637–1645, 2016.

[13] Chao Dong, Chen Change Loy, and Xiaoou Tang. Accelerating the super-resolution convolutional neural network. In *European conference on computer vision*, pages 391–407. Springer, 2016.

[14] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690, 2017.

[15] Alexia Jolicoeur-Martineau. The relativistic discriminator: a key element missing from standard gan. *arXiv preprint arXiv:1807.00734*, 2018.

[16] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015.

[17] Wenzhe Shi, Jose Caballero, Ferenc Huszár, Johannes Totz, Andrew P Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1874–1883, 2016.

[18] Jia-Bin Huang, Abhishek Singh, and Narendra Ahuja. Single image super-resolution from transformed self-exemplars. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5197–5206, 2015.

[19] Yochai Blau, Roey Mechrez, Radu Timofte, Tomer Michaeli, and Lihi Zelnik-Manor. The 2018 pirm challenge on perceptual image super-resolution. In *European Conference on Computer Vision*, pages 334–355. Springer, 2018.

[20] Jiahui Yu, Yuchen Fan, Jianchao Yang, Ning Xu, Xinchao Wang, and Thomas S Huang. Wide activation for efficient and accurate image super-resolution. *arXiv preprint arXiv:1808.08718*, 2018.

[21] Xinmei Tian Zheng-Jun Zh Feng Wu Chang Chen, Zhiwei Xiong. Camera lens super-resolution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019.

Authorized licensed use limited to: Tsinghua University. Downloaded on June 04,2022 at 17:14:00 UTC from IEEE Xplore. Restrictions apply.