Australian National University

Single Image Super-Resolution Using a Channel Attention Generative Adversarial Network

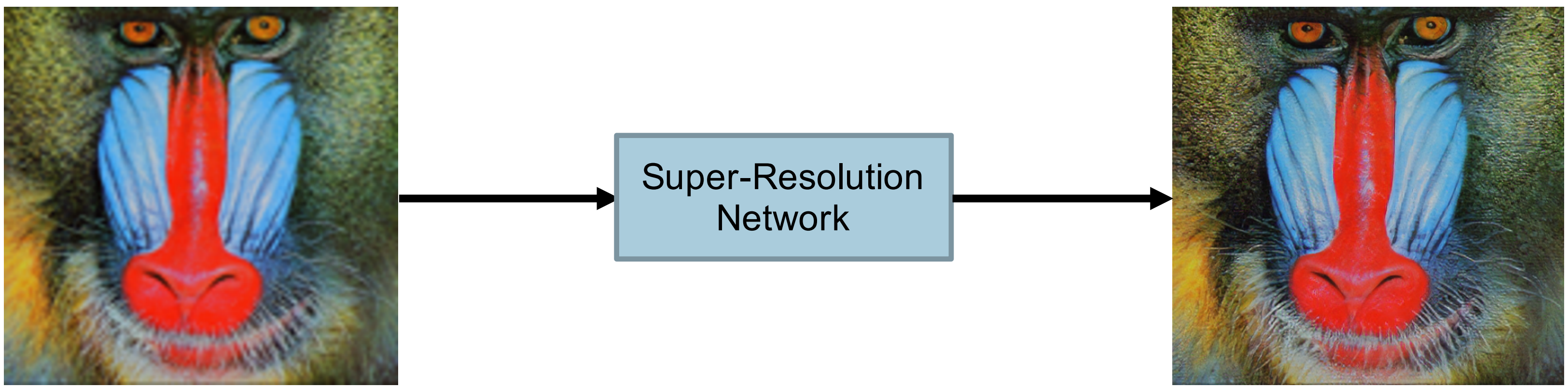
Siyuan Yan
U7050317

Gefei Zhang
U7138112

Hao Wang
U7195922

ENGN8536 Project Presentation
Group 8

Super-resolution



Low-resolution

Super-resolution

Key idea: reconstruct required details!

Introduction

- Motivations
  - Require texture information to reconstruct details
  - Features between channels and pixels
  - Using Generative Adversarial Network
- Our Method
  - Proposed a Channel Attention based GAN

Attention + GAN

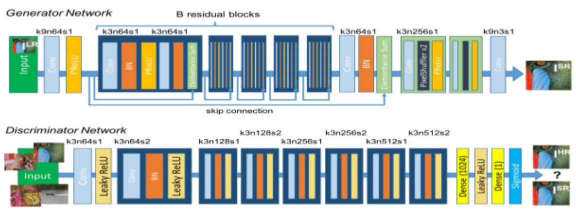
texture reconstructed

LR image

Super-resolution using Attention & GAN

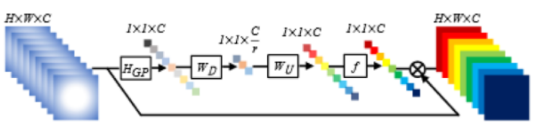
SR image

Related work

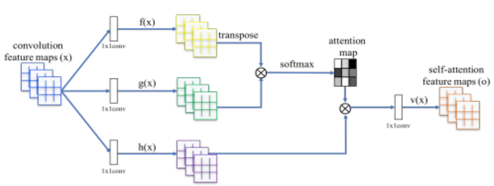


SRGAN

Perceptual Loss: visually satisfying results



Channel Attention

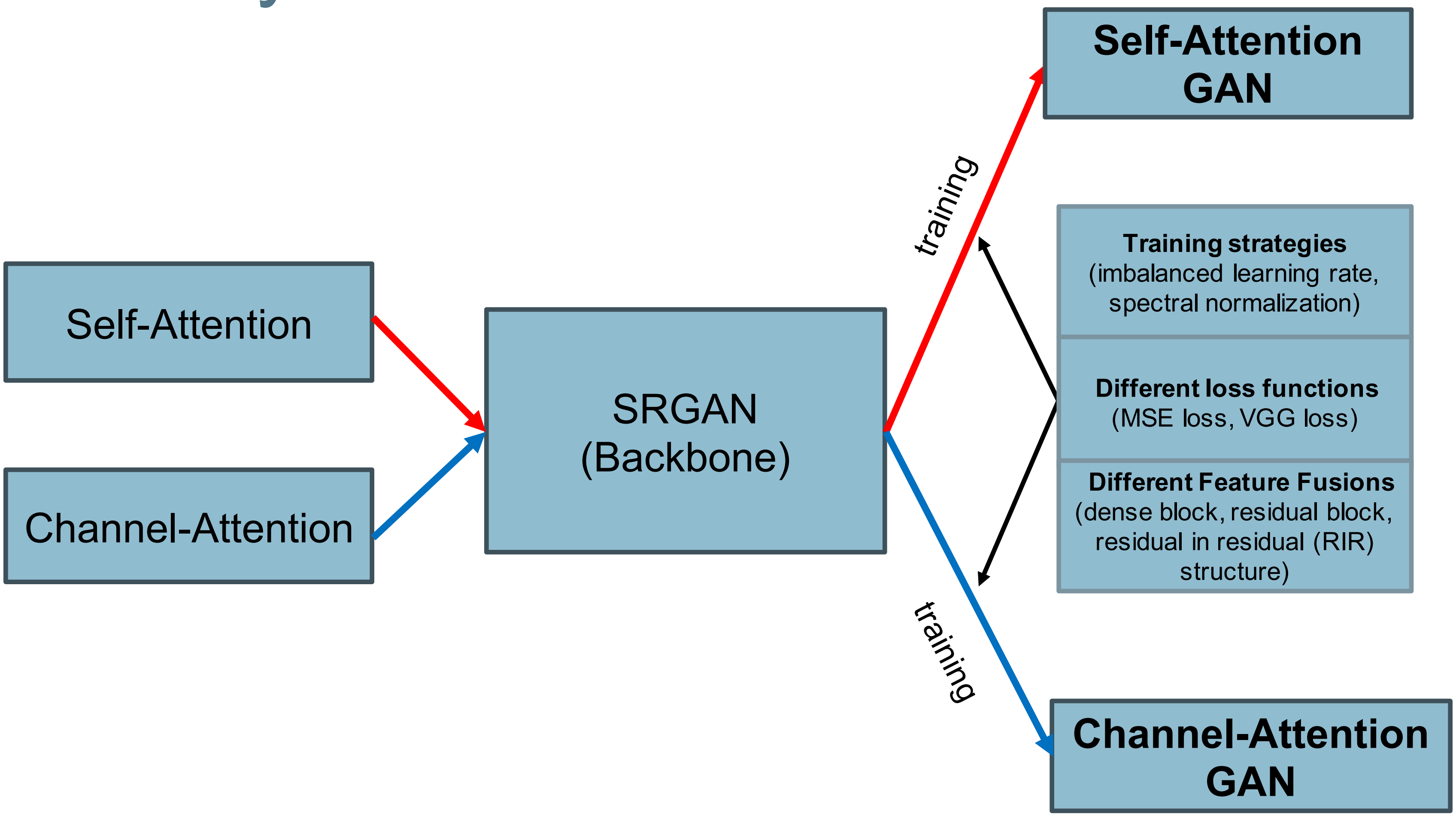


Self Attention

Attention strategies

Australian National University

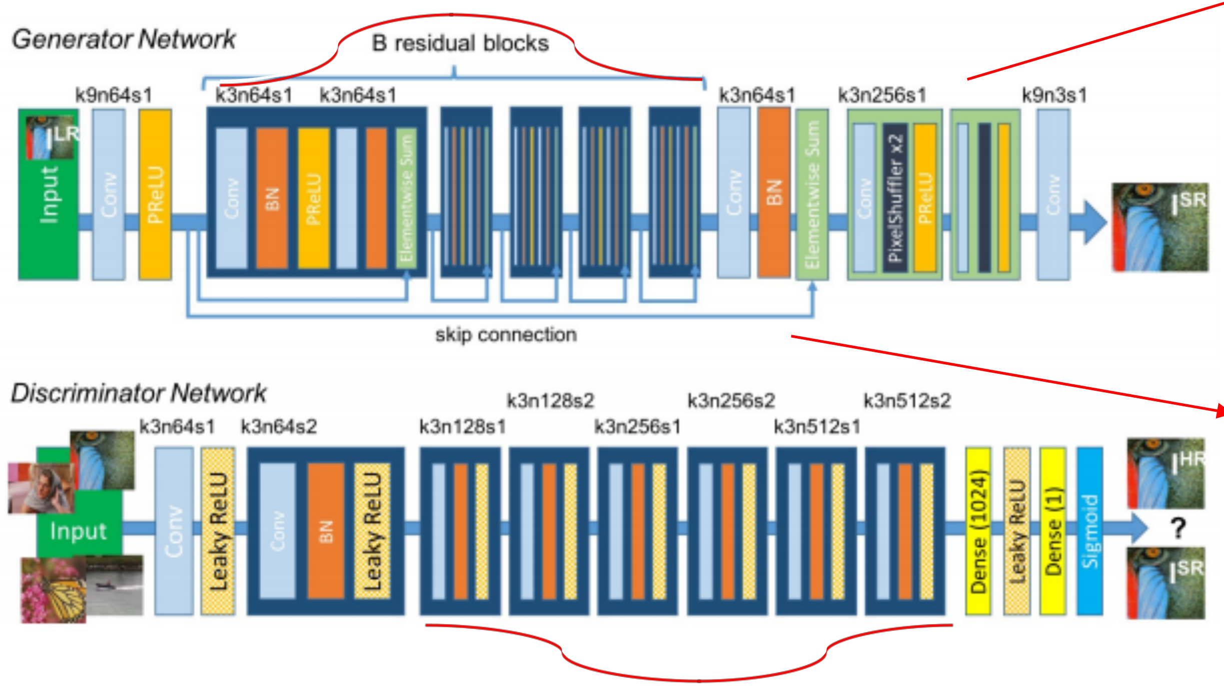
Novelty



Channel Attention GAN Architecture

Channel Attention Blocks

Subpixel convolutional block



Feature fusion of low-level features and high-level features

Channel Attention Blocks

Loss of SR-CAGAN

$$l^{SR} = \underbrace{l_X^{SR}}_{\text{content loss}} + \underbrace{10^{-3} l_{Gen}^{SR}}_{\text{adversarial loss}}$$

Content Loss:

MSE loss *vs* VGG loss

$$l_{MSE}^{SR} = \frac{1}{r^2WH} \sum_{x=1}^{rW} \sum_{y=1}^{rH} (I_{x,y}^{HR} - G_{\theta_G}(I^{LR})_{x,y})^2$$

$$l_{VGG/i.j}^{SR} = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^{HR})_{x,y} - \phi_{i,j}(G_{\theta_G}(I^{LR}))_{x,y})^2$$

Loss of SR-CAGAN

$$l^{SR} = \underbrace{l^{SR}_{X}}_{\text{content loss}} + \underbrace{10^{-3} l^{SR}_{Gen}}_{\text{adversarial loss}}$$

Adversarial Loss:

$$l^{SR}_{Gen} = \sum_{n=1}^{N} -\log D_{\theta_D}(G_{\theta_G}(I^{LR}))$$

Channel Attention vs Self Attention

Channel Attention mechanism:

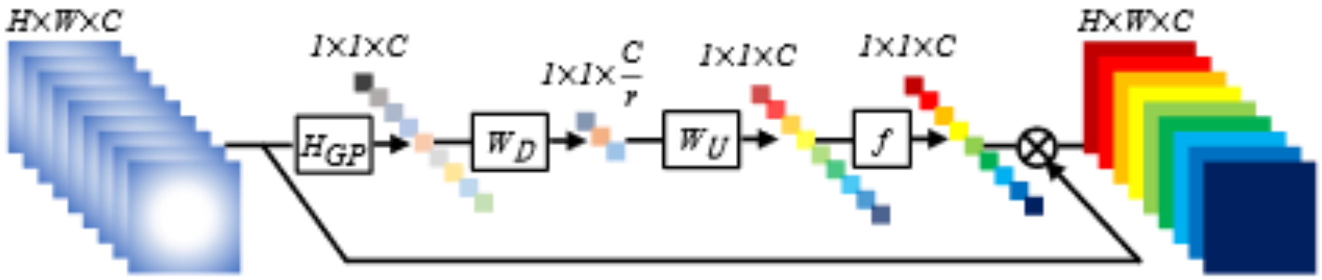


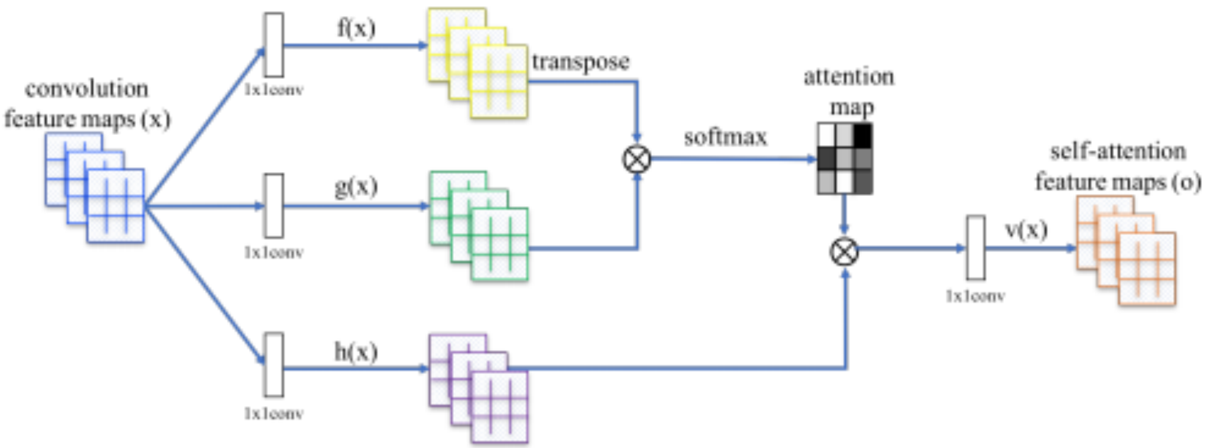
Fig. 3. Channel attention (CA). ⊗ denotes element-wise product

Attention through **channels**!

Capture channel relationships

Channel Attention vs Self Attention

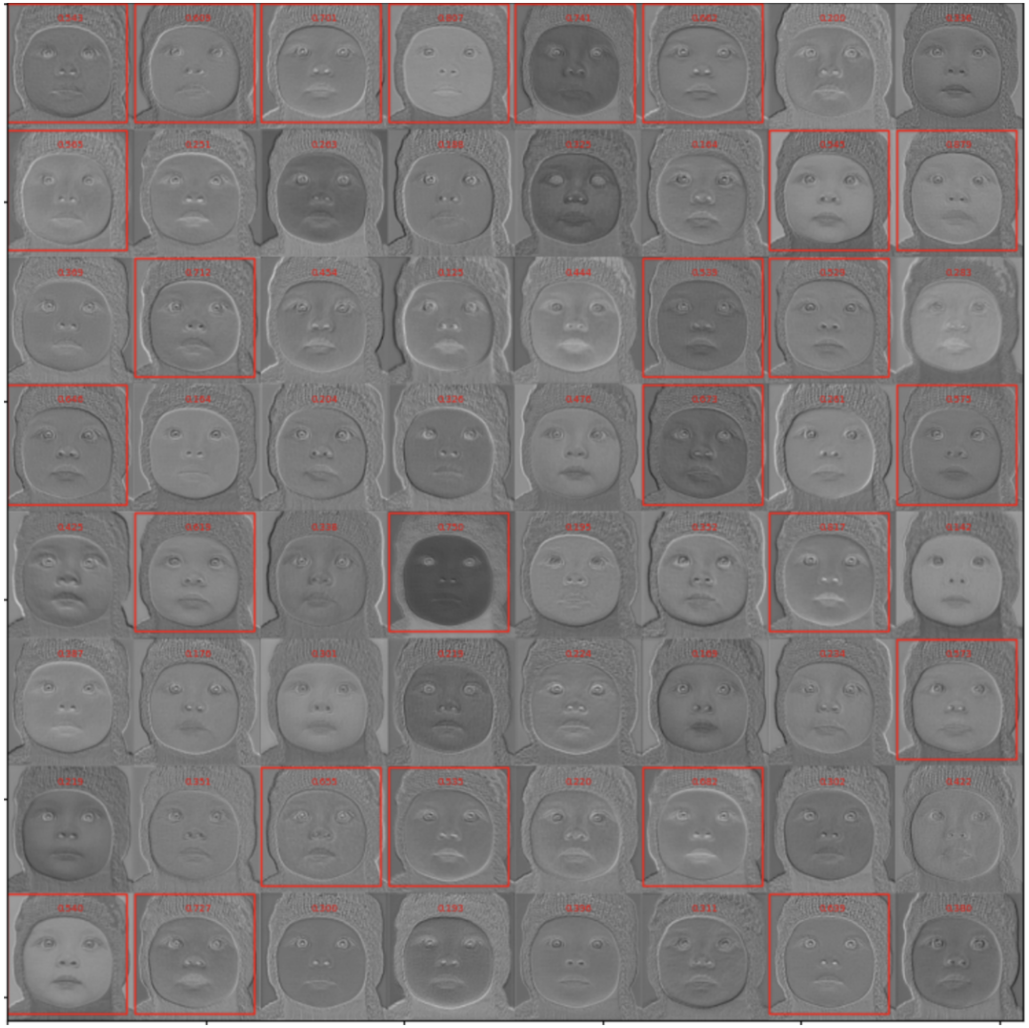
Self Attention mechanism:



Attention through **pixels**!

Capture spatial relationships

Visualization of Channel Attention



Red-boxed channels have larger channel weights(>0.5).

Emphasizing informative channels

Suppresses less useful channels.

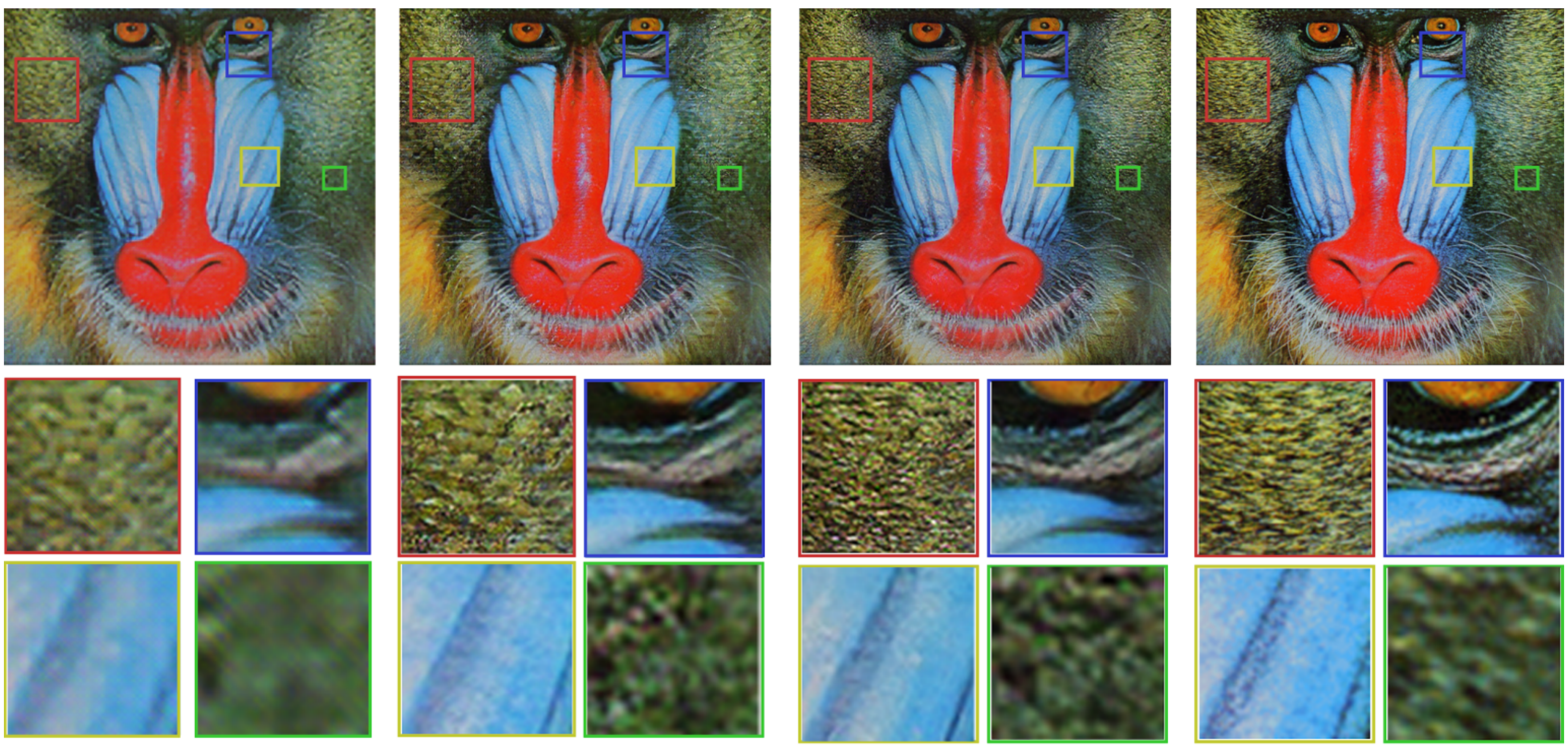
Comparison between different models

dataset / models	Set5			Set14			BSD100		
	PSNR	SSIM	MOS	PSNR	SSIM	MOS	PSNR	SSIM	MOS
SRGAN	26.304	0.825	3.46	24.269	0.690	3.56	23.126	0.634	3.53
Self Attention	27.135	0.828	3.62	25.474	0.727	3.57	23.599	0.646	3.61
Channel Attention	**30.757**	**0.869**	**3.87**	**26.455**	**0.743**	**3.76**	**25.572**	**0.686**	**3.79**

training data: 82k coco images (train2014)
validation data: Set5,Set14 and BSD100
batch size:16
epochs:100

SRGAN vs our Models (4x scaling)

Self Attention vs Channel Attention



bicubic | self attention | **channel attention** | original

Ablation study 1:Channel Attention

Experiments: Different numbers of channel attention
Conclusion: More channel attentions have better performance!

dataset / number of CA	Set5			Set14			BSD100		
	PSNR	SSIM	MOS	PSNR	SSIM	MOS	PSNR	SSIM	MOS
0	27.504	0.828	3.67	25.269	0.660	3.62	23.416	0.674	3.63
4	30.742	0.854	3.77	26.384	0.726	3.67	25.441	0.670	**3.78**
8	**30.811**	0.856	3.78	26.333	0.739	3.63	25.519	0.672	3.70
16	30.757	**0.869**	**3.91**	**26.455**	**0.743**	**3.76**	**25.572**	**0.686**	**3.78**

Ablation study 2:Perceptual loss

Experiments: VGG loss vs MSE loss

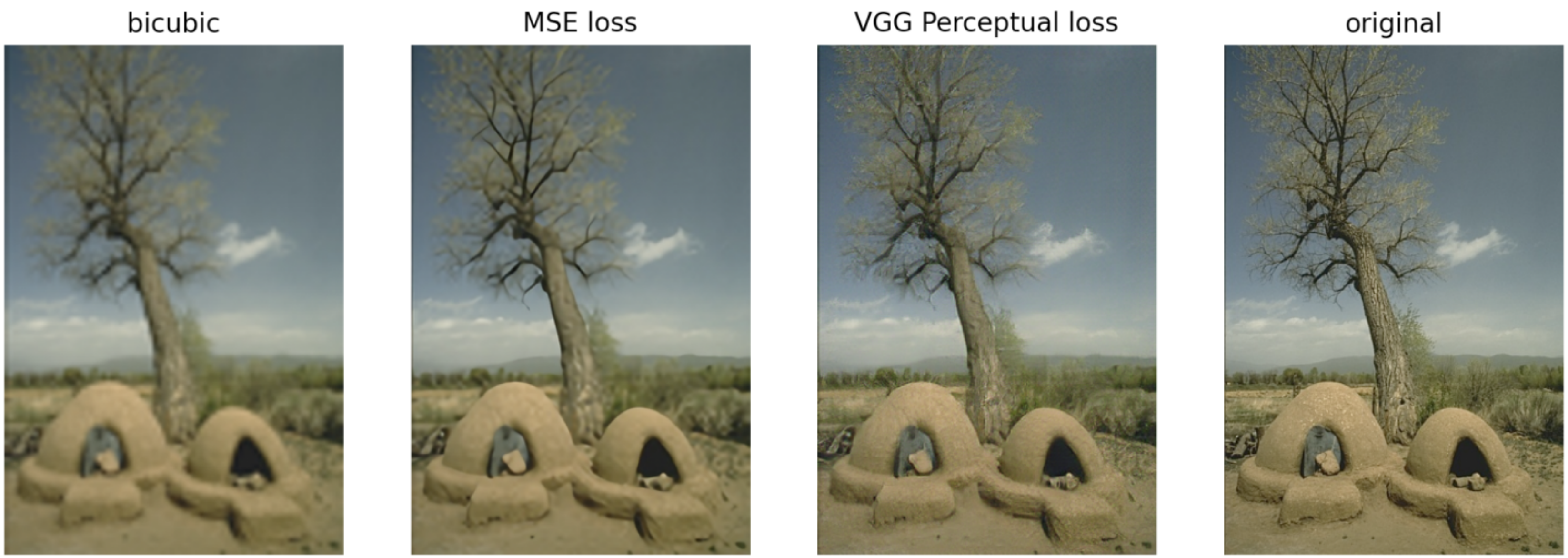
Conclusion:

- MSE loss can get higher PSNR and SSIM score
- But VGG loss has better visual results

number of CA	Set5			Set14			BSD100		
	PSNR	SSIM	MOS	PSNR	SSIM	MOS	PSNR	SSIM	MOS
MSE loss	**33.335**	**0.923**	3.33	**28.588**	**0.800**	3.33	**27.187**	**0.722**	3.37
VGG loss	30.942	0.854	**3.91**	26.384	0.726	**3.88**	25.441	0.670	**3.79**

VGG loss vs MSE loss

PSNR does not necessarily reflect the perceptually better SR results!



PSNR: 28.367 26.476

Ablation study 3:Stabilization strategy

Experiments: Fixed vs Imbalanced learning rate

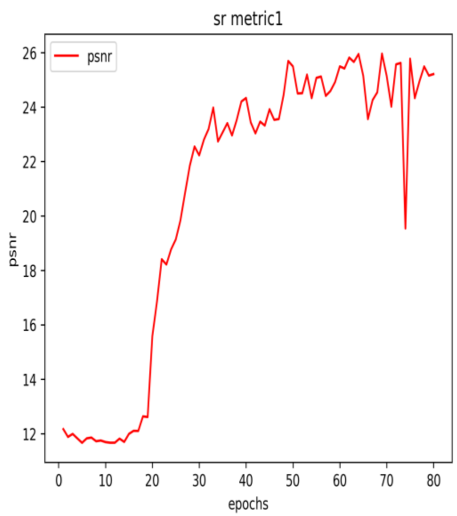
Model is trained on a smaller COCO dataset (6k images)

Conclusion: model with imbalanced learning rate gets better results.

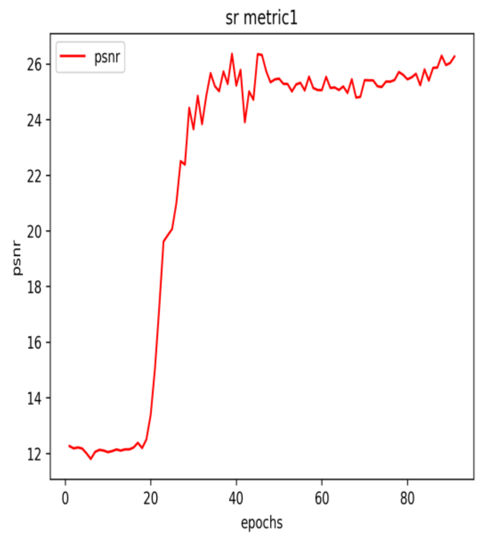
	Set5			Set14			BSD100		
	PSNR	SSIM	MOS	PSNR	SSIM	MOS	PSNR	SSIM	MOS
fixed learning rate	26.039	0.808	3.33	24.848	0.721	3.41	**24.517**	**0.667**	3.**66**
imbalanced learning rate	**26.146**	**0.812**	**3.60**	**25.077**	**0.727**	**3.54**	24.093	0.656	3.54

Stabilization strategies

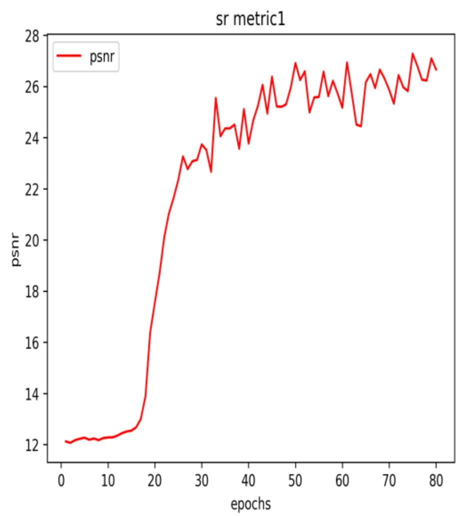
Imbalanced learning rate and **spectral norm** speed up and stabilize the training process



psnr for base model



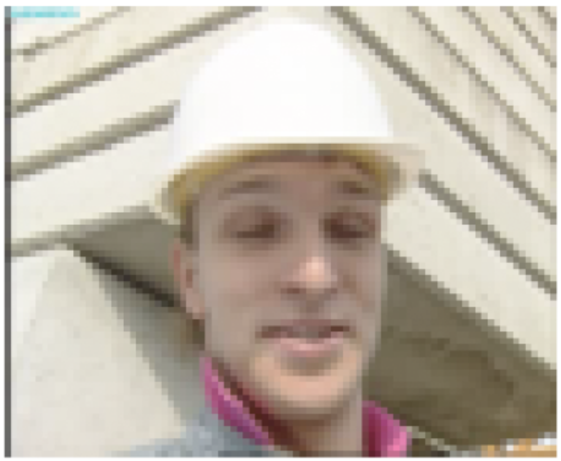
psnr for imbalanced lr



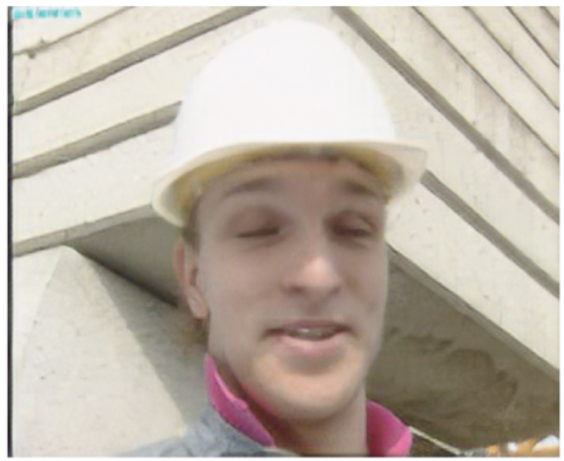
psnr for spectral norm

More results(4x scaling)

4x downsampled lr image



channel attention gan



original

lr: low resolution image (4x downsampled result of the original image)

More results(4x scaling)

4x downsampled lr image

channel attention gan

original

More results(4x scaling)

4x downsampled lr image

channel attention gan

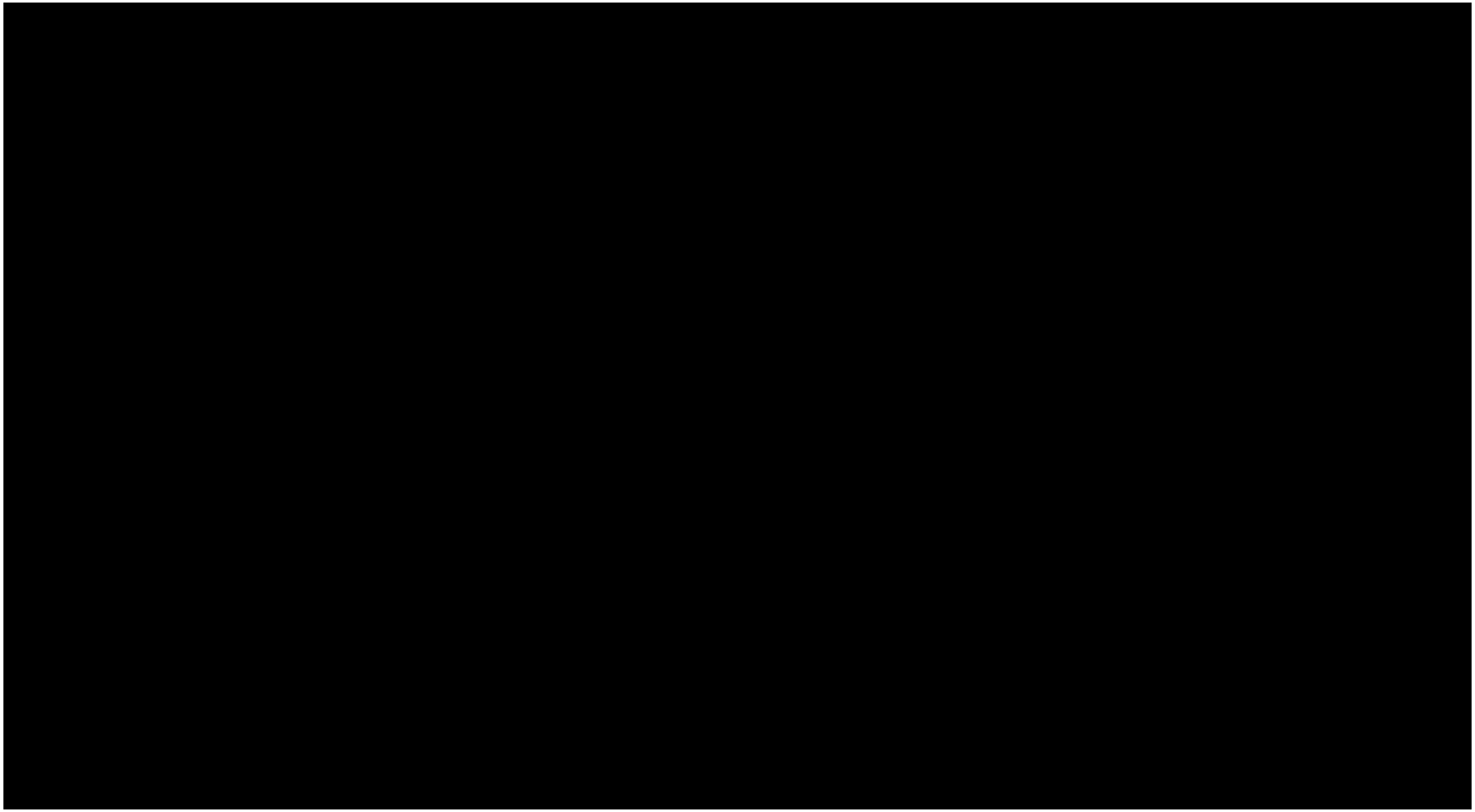
original

More results(4x scaling)

4x downsampled lr image

channel attention gan

Video demo(4x scaling)



Conclusion

- We proposed the Channel-attention GAN and showed channel information is more useful than spatial information on SR tasks.

- Our models outperform SRGAN on SR tasks.

Future work

- To research a better feature-fusion method(e.g. RIR structure)

- To improve the efficiency of our model

References

1 Saeed Anwar, Salman Khan, and Nick Barnes. A Deep Journey into Super-resolution: A Survey. In arXiv:1904.17523, 2020.

2 Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Learning a deep convolutional network for image super-resolution. 01 2014.

3 C. Dong, C. C. Loy, K. He, and X. Tang. Image super-resolution using deep convolutional networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(2):295–307, 2016.

4 M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The pascal visual object classes (voc) challenge. International Journal of Computer Vision, 88(2):303338, June 2010.

5 Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, editors, Advances in Neural Information Processing Systems 27, pages 2672–2680. Curran Associates, Inc., 2014.

6 P. Gupta, P. Srivastava, S. Bhardwaj, and V. Bhateja. A modified psnr metric based on hvs for quality assessment of color images. In 2011 International Conference on Communication and Industrial Application, pages 1–4, 2011.

7 Z. Wang, E. P. Simoncelli, and A. C. Bovik. Multiscale structural similarity for image quality assessment. In The Thrity Seventh Asilomar Conference on Signals, Systems Computers, 2003, volume 2, pages 1398–1402 Vol.2, 2003.

8 Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. 2015. cite arxiv:1512.03385Comment: Tech report.

9 J. Hu, L. Shen, and G. Sun. Squeeze-and-excitation networks.In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 7132–7141, 2018.

10 Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi. Photo-realistic single image super-resolution using a generative adversarial network. pages 105–114, 07 2017.

References

11 Han Zhang, Ian Goodfellow, Dimitris Metaxas, and Augustus Odena. Self-attention generative adversarial networks. In arXiv preprint arXiv:1805.08318, 2018.

12 Yulun Zhang, Kunpeng Li, Kai Li, Lichen Wang, Bineng Zhong and Yun Fu, Image Super-Resolution Using Very Deep Residual Channel Attention Networks. In arXiv:1807.02758, 2018.

13 Wenming Yang, Xuechen Zhang, Yapeng Tian, Wei Wang, Jing-Hao Xue and Qingmin Liao. Deep Learning for Single Image Super-Resolution: A Brief Review. 2019. cite arXiv:1808.03344.

14 Tong T , Li G , Liu X , et al. Image Super-Resolution Using Dense Skip Connections[C]// IEEE International Conference on Computer Vision. IEEE Computer Society, 2017.

15 Yan Y , Liu C , Chen C , et al. Fine-grained Attention and Feature-sharing Generative Adversarial Networks for Single Image Super-Resolution[J]. 2019.


Q&A