Resolution Enhancement – white paper

Aug 2024

saiwa

simple artificial intelligence web application

Resolution Enhancement or Super-Resolution (SR) is the application of increasing spatial resolution of input images which are provided in limited resolution. One of the main categories of SR is **single-image super-resolution** that recovers a high-resolution (HR) image from a single low-resolution (LR) one[1]. SR has been around as a challenging and inherently ill-posed image processing problem for decades. There are a variety of classical methods for scale-up images, including: prediction-based methods, edge-based methods, statistical methods, patch-based methods and sparse representation methods.

[1] The other variant uses multiple LR images.

As a recent approach several deep learning techniques for super-resolution and several deep learning methods have been proposed, including: Convolutional Neural Networks (CNN) based methods, Generative Adversarial Nets (GAN) based methods and more. These methods differ from each other in aspects such as: network architectures, loss functions, learning algorithms, etc. For further details on deep learning-based SR techniques please refer to [1].

In saiwa Resolution Enhancement service, three single image resolution enhancement methods using deep learning are provided: Residual Dense Network (**RDN**) [2], Residual in Residual Dense Network (**RRDN**) [3] and **Real-ESRGAN** (Enhanced Super-Resolution Generative Adversarial Networks) [3]. Both RDN algorithms use residual learning proposed in ResNet for learning residuals instead of a thorough mapping. Residual learning has also been widely adopted to ease the training process, either in image-level or feature-level. Real-ESRGAN uses synthetic data to improve image details and offers customizable magnification ratios.

Resolution Enhansement

Residual Dense Network (RDN) consists of four parts: shallow feature extraction net (SFENet), residual dense blocks (RDBs), dense feature fusion (DFF), and the up-sampling net (UPNet). Figure 1 shows an overview of a RDN and its RDBs. RDN creates a super-scaled version of an input image by the construction of a very large amount of full-sized (same size of the input image) intermediate representations of the input LR image (the outputs of the RBDs) which are then recombined to form a larger one. Using local and global features, RDN leads to a dense feature fusion and deep supervision. The local feature fusion (LFF) stabilizes the training wider network and also adaptively controls the preservation of information from current and preceding RDBs. Moreover, a global feature fusion (GFF) is proposed to extract hierarchical features in the LR space.

Finally, an up-sampling net (UPNet) in the HR space is employed. For more details of residual connections and how RDN works please consult [2]. RDN model, trained with Adversarial and VGG features losses, also provides compression artifact cancelling and retrieves improved super-scaled images starting from a poor ground truth [4].

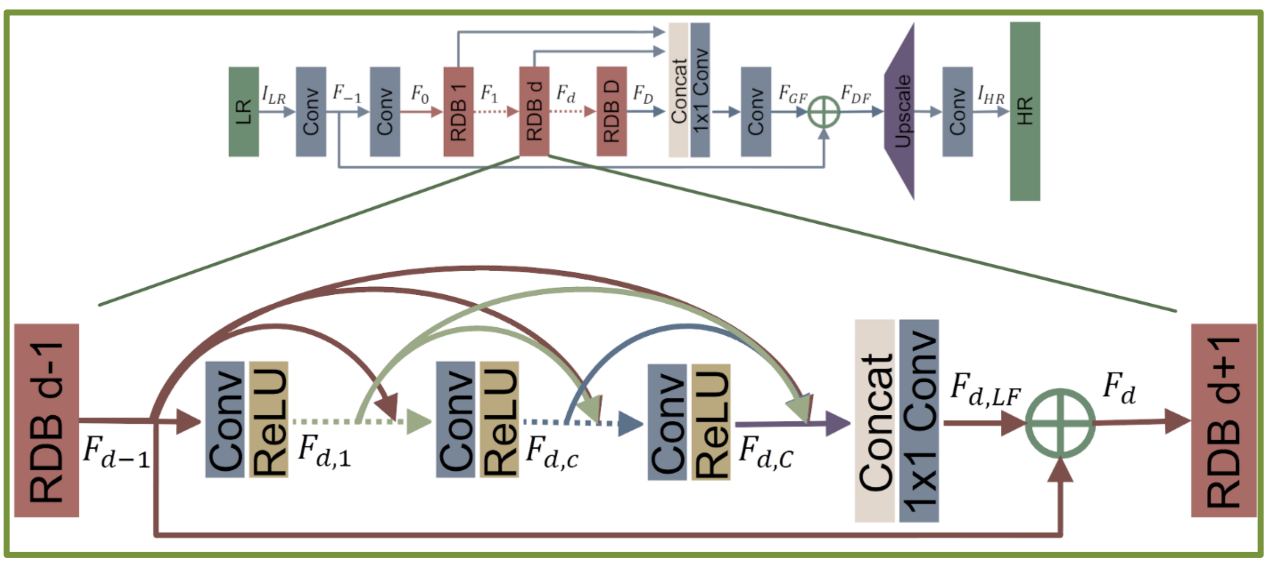

Figure 1. An overview of a Residual Dense Network and its Residual Dense Blocks (printed from [4])

Figure 2 shows a few LR and HR pair images (scaled up by a factor of 2) that were reconstructed using saiwa interface for RDN SR method.

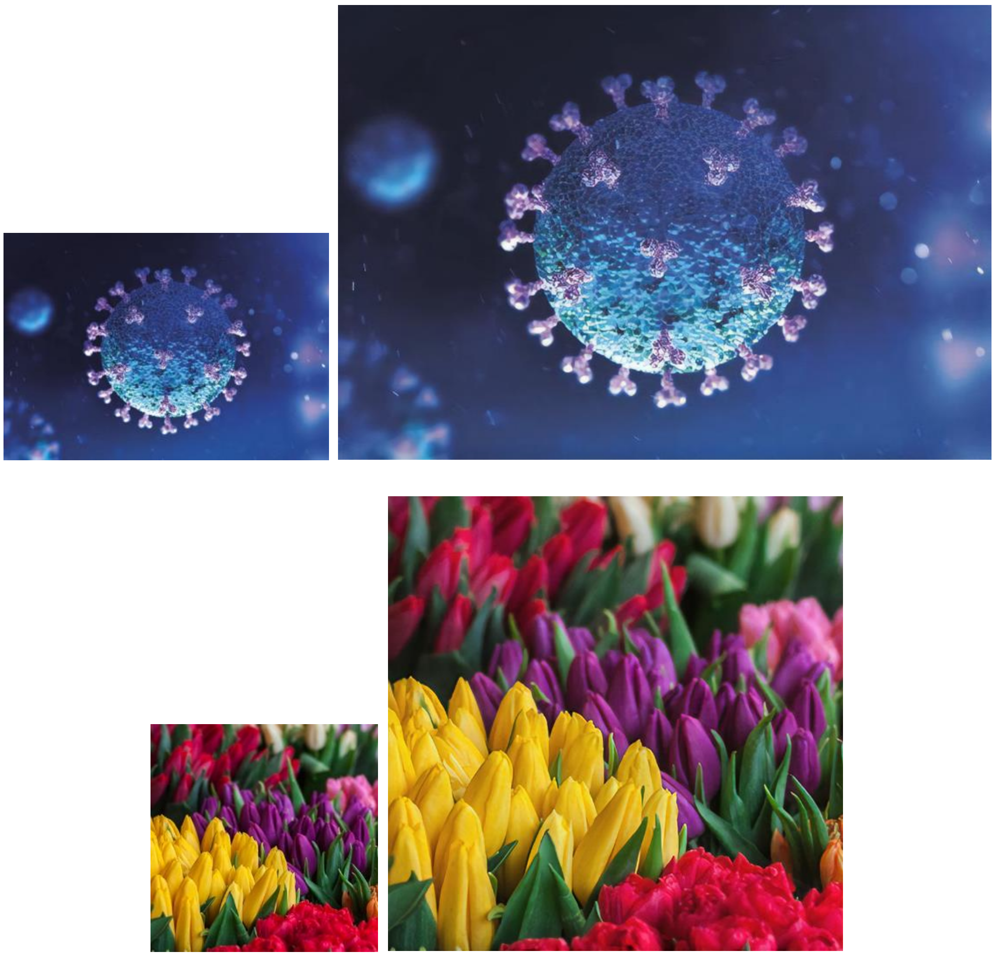

Figure 2. Visual results of RDN plus artifact cancelling method using saiwa Resolution Enhancement service.

Residual in Residual Dense Network (RRDN) uses effectively residual dense blocks for SR as we have seen in RDN. RRDN utilizes residual-in-residual structure, where residual learning is used in multi-level and, similar to the RDN, in the main network path where the network capacity becomes higher, benefiting from the dense connections. Figure 3 depicts the overall structure of Residual in Residual Dense Block.

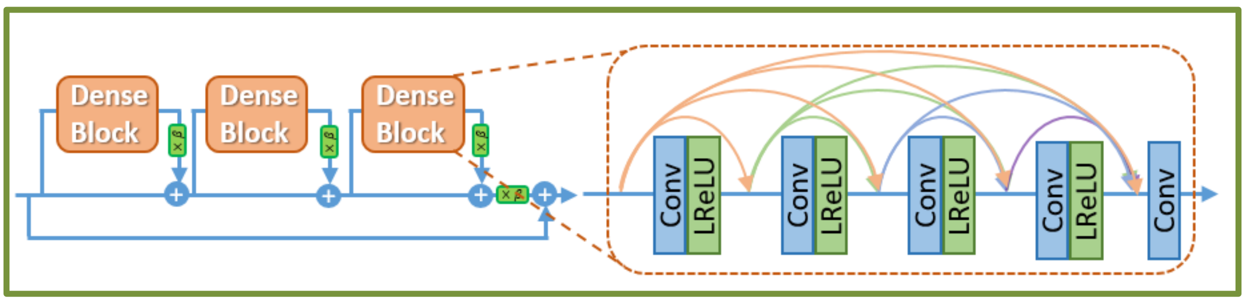

Figure 3. Residual in Residual Dense Block (printed from [3])

In addition to the improved architecture, RRDN has other main features including: 1) residual scaling to prevent instability, and 2) smaller initialization to ease the training. More discussion can be found in [3]. Figure 4 shows a few LR and HR pair images (scaled up by a factor of 4) that are reconstructed using saiwa interface for RRDN SR method.

Figure 4. Visual results of RRDN SR method using saiwa Resolution Enhancement service.

Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN) further enhance the visual quality, by utilizing three key components of SRGAN network architecture[5] : adversarial loss and perceptual loss, and an enhanced version of each, to derive an enhanced SRGAN. ESRGAN employs the basic architecture of SRResNet, where most computations are done in the LR feature space. we can select or design components, such as residual blocks, dense blocks, or RRDB. Figure 5 depicts the basic architecture of SRResNet.

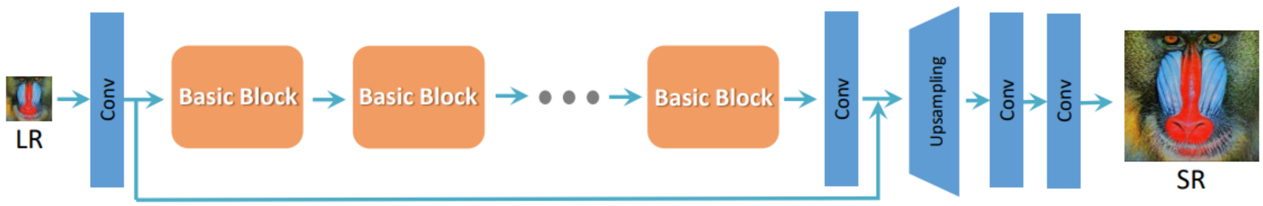

Figure 5. The basic architecture of SRResNet (printed from [3])

In particular, ESRGAN utilizes the Residual-in-Residual Dense Block (RRDB) without batch normalization (BN) as the basic network building unit, as depicted in Figure 6.

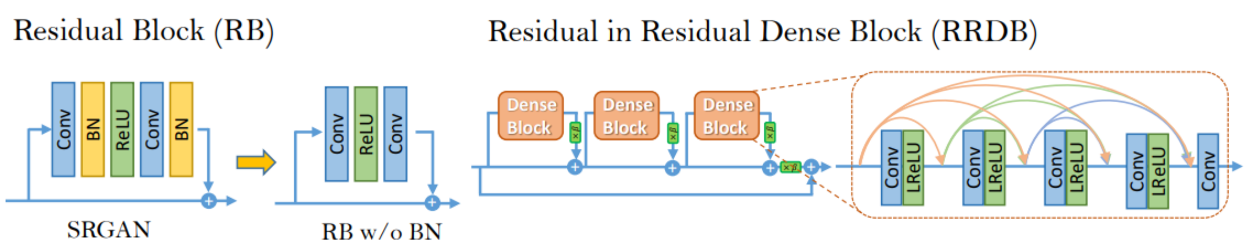

Figure 6. **Left:** Remove the BN layers in residual block in SRGAN. **Right:** Residual in Residual Dense Block (printed from [3])

Besides the improved structure of generator, ESRGAN borrows the concept from relativistic GANs, enabling the discriminator to predict relative realness instead of the absolute value. Finally, ESRGAN improves perceptual loss by using the features before activation, providing stronger supervision for brightness consistency and texture recovery. Figure 7 shows a few LR and HR pair images (scaled up by a factor of 4) that are reconstructed using saiwa interface for Real-ESRGAN method.

Figure 7. Visual results of Real-ESRGAN method using saiwa Resolution Enhancement service.

References:

[1] Wang, Zhihao, Jian Chen, and Steven CH Hoi. "Deep learning for image super-resolution: A survey." IEEE transactions on pattern analysis and machine intelligence 43.10 (2020): 3365-3387.

[2] Zhang, Yulun, et al. "Residual dense network for image super-resolution." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

[3] Wang, Xintao, et al. "Esrgan: Enhanced super-resolution generative adversarial networks." Proceedings of the European conference on computer vision (ECCV) workshops. 2018.

[4] Francesco Cardinale. "Zoom in... enhance: a Deep Learning based magnifying glass - part 2" tutorial, 2019.

[5] Ledig, C., Theis, L., Husz´ar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al.: Photo-realistic single image super-resolution using a generative adversarial network. In: CVPR. (2017).

info@ saiwa.ai

720 Guelph Line Burlington, ON L7R 4E2 Building
+15148131809
www.linkedin.com/company/saiwa
www.instagram.com/saiwa.ai