Exploring Image Super Resolution Techniques

Varsha R[1], Vibush Shanmugam[2], Neha Ghaty[3]

[1,2,3]*Student, Dept of CSE, PES University, Bangalore, India*

Abstract - *Super resolution is a technique that reconstructs a high resolution image from the observed low resolution images. Most Super Resolution techniques aim to improve the spatial resolution of an image. But as two-dimensional signal records, digital images with a higher resolution are always desirable in most applications. Imaging techniques have been rapidly developed in the last decades, and the resolution has reached a new level. The question we are trying to address is whether image resolution enhancement techniques are still required and if they are, what are the techniques that are state of the art, what type of super resolution enhancements are known to produce the best results, what are the drawbacks with these architectures and to come up with a work around to handle these drawbacks and improve the resolution of images*

***Key Words*: Computer vision, high resolution, super resolution, spatial resolution, image resolution**

1. INTRODUCTION

The main objective of super-resolution is to estimate the high-resolution visual output of a corresponding low resolution visual input, which can either be a low-resolution image (single-image) or a set of images (multi-image), for example, corresponding to frames in a video sequence. The goals range from providing better content visualization for traditional image processing application to achieving better visual recognition, including computer vision tasks. Image super-resolution is important in many applications of multimedia, such as playing a video on a higher-resolution screen. Due to some technical limitations in imaging devices and systems, like, the presence of optical distortions and lens blur, insufficient sensor sampling density and aliasing, motion blur due to low shutter speed, the presence of noise due to sensor limitations and lossy coding, super-resolution technique is actually needed. The high-resolution visual output can be obtained either by providing devices with excellent spatial resolution, at the cost of a very high market price of the imaging device or with the use of software-related tools. The former is achieved by some hardware-related tools which includes – reducing the pixel size (which unfortunately leads to an increasing appearance of shot noise as the amount of light captured by the device decreases), increasing the chip size to accommodate a larger number of pixel sensors (which unfortunately results in an increased capacitance), reducing the shutter speed (which leads to an increasing noise level), adoption of high-precision optics and sensors (which invariably results in an increase in the price of the device). The advantage of post-processing the captured visual data is that it allows us to balance computational and hardware costs. Thus, on one hand we may have a lower market price and, on the other we can work with contemporary imaging devices and systems.

Super-resolution allows a high-resolution image to be generated from a lower resolution image, with the trained model inferring photo-realistic details while up-sampling. In this work, we will explore super resolution GANs and their applications in detail.

Super-resolution GAN applies a deep network in combination with an adversary network to produce higher resolution images. SRGAN is more appealing to a human with more details. During the training, a high-resolution image is downsampled to a low-resolution image. A GAN generator upsamples the low resolution images to super-resolution images. We use a discriminator to distinguish between the original high resolution images and the super resolution image generated by SRGAN. The GAN loss is then backpropagated to train the discriminator and the generator. The SRGAN model [9] adds an adversarial loss component which constrains images to look like natural images, producing convincing solutions.

2. LITERATURE SURVEY

2.1 Image super-resolution

The task of estimating a high-resolution image from its low-resolution counterpart is referred to as super-resolution (SR). The optimization target of super-resolution algorithms is usually the minimization of the Mean Square Error (MSE) between the generated image and the ground truth image. Minimizing MSE also maximizes Peak Signal to Noise Ratio (PSNR) which is a common measure that is used to evaluate super resolution algorithms.

2.2 History of super-resolution techniques

In 1964, Harris established the foundation for super resolution as a technique by solving the diffraction problem [19]. The milestones of spectroscopy have been achieved almost entirely by using readily available detection technology while minimizing background levels. The lesson from the progress in both fields is basically that anything can be detected if the background is low enough.

In 1984, Tsai and Huang first addressed the idea of super resolution to improve the spatial resolution of a dataset containing the Landsat images. After analysing the results

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 06 | June 2019 www.irjet.net p-ISSN: 2395-0072

from these experiments the super resolution techniques were categorized into Interpolation based methods, Reconstruction based methods and Experiment based methods.

In the period of 1984 to 2000, most methods concentrated on frequency domain based super resolution technique. This technique comes under reconstruction methods to obtain high resolution images that obtains high computational efficiency. But it was observed that these models were sensitive to errors and could not handle complicated inputs.

2000 to 2010 was a decade of spatial domain methods. Most of these methods produced state of the art results in that era. But however now, these methods are obsolete because of the advent of experiment based techniques. Interpolation was one of the easiest and common methods. Iterative back projection, regularization, etc were a few other methods that were designed for super resolution.

From 2010 until the present days, machine learning and deep learning methods have changed the way we used to solve problems. Computer vision has revolutionized the image processing domain. Example based methods were widely popularized and regression based methods, SR-CNN, SR-GANs were the commonly used methods.

2.3 Comparison

1. **Interpolation based methods:** Interpolation is the technique of using points with known values or sample points to estimate values at other unknown points.

Advantages: Needs lesser computational complexity and hence these methods are better suited for real-time applications.

Disadvantages: It's not possible to obtain high-frequency information and it is not possible to find missing spectral contents. The results obtained are over-smooth and have jagged artifacts at edges.

2. **Reconstruction based methods:** Data collected in two-dimensional projections give planar images of object at each projection angle. To obtain information along the depth of the object, tomographic images are reconstructed using these projections.

Advantages: The high resolution images are very close to the original low resolution images with respect to features. There is also an added smoothness because of down-sampling.

Disadvantages: Not suited for arbitrary images as ringing artifacts may appear.

3. **Example based methods:** The algorithms of example-based super-resolution problems are based on machine learning models exploiting available examples.

Advantages: These are the most successful in producing state of the art, best quality images because it is based on example learning neural networks.

Disadvantages: Due to insufficient training examples available for the model to learn from, high frequency artifacts may appear in the output high resolution images. Learning time also significantly increases which forces the need for hardware resources like high memory or a GPU.

3. RESEARCH RESULTS

3.1 Interpolation based SR techniques

1. Nearest Neighbour Interpolation: This method manipulates pixel values of the nearest pixels which have the same value as the neighbour pixel. This method is one of the simplest and easiest but does not produce hig sub-pixel accuracy.

2. Bilinear Interpolation: This method passes a straight line between two consecutive pixel locations. This method is known to be better than Nearest Neighbour Interpolation but still does create artifacts and poor preservation image details.

3. Quadratic Interpolation: This method uses three points for interpolation and results in one point at the centre and another two points on each side. This has shown one of the best performances.

4. Bicubic Interpolation: This method extends into four number of pixel neighbours where the function is defined with two pixels on each side. This performs better than quadratic interpolation too.

3.2 Reconstruction based SR techniques

1. Non-uniform interpolation: This method allows for the reconstruction of samples from other samples taken at non-uniformly distributed locations. It is a basic and intuitive method of super resolution and is known to have relatively low computational complexity. But it assumes that the blur and noise characteristics are identical across all low resolution images.

2. Frequency domain: This method is basically reconstructing a high-resolution image from multiple low resolution images based on the aliasing images present in the low resolution images. This method is simple to implement and produces high quality output but it is only efficiently applied provided the noise has zero mean the blurring is either absent or identical across the low resolution images.

3. Regularization: In this method, we assume that the registration parameters are estimated and deterministic regularization is done by taking proper prior information about the solution. There is no need of larger training datasets as the image details preservations is high. But the performance degrades with higher magnification factor. It also takes more time for computation.

4. Projection onto Convex Sets: In this method, an estimate of the high resolution version of the reference image is determined iteratively starting from some arbitrary initialization. It solves the effect of under sampling but suffers from image blurring.

5. Iterative back projection: In this image, a high resolution image is estimated by back projecting the difference between the simulated low resolution image and captured low resolution image on the interpolated image. This method removes noise and blurry effects from the image but there is no unique solution as it is difficult to choose the ideal back projecting operator.

3.3 Example based SR techniques

1. Neighbour Embedding: In this method, each input data vector can be described as a linear combination of its nearest neighbours on the natural image manifold of low resolution patches. This is an unsupervised learning method with both external and internal learning and where external performs better.

2. Sparse Coding: This follows the previous method with the additional constraint of a compact and optimized dictionary that is obtained through the training process. This provides robust nearest neighbour decomposition. This is also unsupervised on both the low and high resolution pairs. This allows only external learning. Also ensures no overlapping and requires low computation.

3. Anchored Regression: In this method, an external database composed of a low resolution dictionary and a set of linear regression matrices map the low resolution examples to their high resolution counterparts. The running cost and computational cost is greatly reduced by removing the sparsity constraint from the inference stage in sparse coding technique. This allows external learning methods.

4. Regression Trees and Forests: In this method, the input image is divided into patches. Each patch traverses the tree from root node to the most suitable leaf node, and the corresponding regression model is used to generate the high-resolution patch. This method is computationally faster than other example based techniques. But there is a limitation of high memory requirements for storing the regression parameters because they are expanded in the whole set of training data points.

5. Deep Learning: In this method, we use the power of back-propagation algorithms in order to learn the hierarchical representations that allow for minimizing the error at the end of the network. Deep learning is one of the current alternatives with supervised learning approach based on deep convolutional neural networks. These algorithms have the power to determine the hierarchical descriptions of the visual data and this is learned directly from the data. But the fine tuning of all the parameters in the network takes a considerably large amount of time than the classical machine learning approaches.

3.4 Challenges faced today

1. Image registration: Bayesian approach can be used but computation cost can be very high.

2. Computational efficiency: Interpolation restoration algorithms work but computation goes up with non-translation models.

3. Robustness aspects: Median estimation to combine the upsampled images to cope with outliers from noise works but showed improvements for outliers assumed on the validation data and not much on real data.

4. Performance limits: motion estimation, decimation factor, number of frames and prior information work but they only suggest ways and are far from enough.

3.5 GAN based SR techniques

Generative Adversarial Networks also commonly known as GANs are deep neural-network architecture comprising of the generator and discriminator, that are pitted against each other . GANs have large potential since they can learn to generate by itself any type of data. [1]

The SRGAN model proposed by Ledig et al. [9] adds an adversarial loss component to the GAN loss function which constrains images to appear close to natural images. The SRGAN generator is conditioned on low-resolution input and infers photo-realistic natural images with 4x upsampling. Along with adversarial loss there is a perceptual loss (which is a weighted sum of content loss and adversarial loss) from a pre trained classifier and regularization loss that encourages spatially coherent images. SRGAN set a new state-of-the-art for image super-resolution with high upscaling factors (4x) as measured by both PSNR and Structural Similarity (SSIM). They confirm with an extensive mean opinion score (MOS) test on images from three public benchmark datasets (Set5, Set14 and BSD100) that SRGAN is the new state of the art, by a large margin, for the estimation of photo-realistic SR images with high upscaling factors (4×) .It also shows that perceptual loss is more invariant to changes in pixel space and hence performs better than MSE based content loss. Despite this SRGAN is not optimized for video super-resolution in real time and perceptually convincing reconstruction of text or structures scenes is still beyond the scope of the model.

Liu et al. [10] attempted to further improve on the SRGAN model proposed by Ledig et al. by making changes to the model network. They trained models for both 2x and 4x upsampling of images. They tried three different loss function/optimizer combinations, namely, softmax cross entropy loss with an Adam optimizer, which did not train beyond a few hours until either the generator or discriminator gained an advantage over the other. Wasserstein GAN with gradient penalties did not produce useful results either. Finally, using the loss function from Least-Squares GAN: least square loss and Adam optimizer resulted in the most stable training and best-looking images of the three approaches. They received results that were better than bicubic interpolation methods, sometimes at the expense of added background colour noise and artifacts and some preliminary results on video super-resolution. The model however had a hard time generating details across all classes when the input dataset had many class labels. They trained their 4x upsampling GAN only on anime images and hence this model may not work well on real images. Their preliminary work on video super resolution led to a generative network that could learn color and spatial structure well, however there was still a little bit of blur.

Xiangyu Xu et al. [11] created an algorithm to directly restore a clear high-resolution image from a blurry low-resolution input. They focus on text and face images and learn a category specific prior to solve this problem. They designed two models called MCGAN (Multi-class GAN) and SCGAN (Single-class GAN) in order to generate high resolution images. SCGAN has a single generator and discriminator whereas MCGAN has a single generator and K discriminators which are trained to classify real and generated images for each of the K classes. After training the learned generator can be used to generate images from any of the K classes. They added two additional components to the basic GAN loss function, namely pixel-wise loss and a feature matching loss term. The pixel-wise loss penalizes the difference between generated images and the ground truth image. The feature matching loss function forces restored and real images to have similar feature responses at the intermediate layers of the discriminator network. This creates more realistic features and structural information in generated images. Results showed that SCGAN performed better than MCGAN and state-of-the-art super resolution methods on both text and face images. On the down side some of the reconstructed faces contained checkerboard artifacts.

Karras et al. [12] design a model that progressively trains the generator and the discriminator. Generation of high resolution images is difficult due to the gradient problem and memory constraints. Progressively growing the generator and the discriminator starting from low resolution images and progressively adding higher resolution images not only improves the stability in the high resolution images generated, but also significantly reduces the training time. GANs have a tendency to capture only a part of the variation in the training data which hampers high resolution image generation. They proposed a new way to increase variation in the generated data which uses minibatch (Salimans et al. [13]) standard deviation to produce a new feature map. This layer is then inserted towards the end of the discriminator to produce optimal results. To disallow the scenario where the magnitudes in the generator and discriminator spiral out of control as a result of competition, they normalize the feature vector in each pixel to unit length in the generator after each convolutional layer. In addition, they also introduce a new metric for evaluating quality and variation of generated images, which uses sliced Wasserstein distance (SWD). This is because current techniques such as MS-SSIM (Odena et al., 2017) are good at finding large-scale mode collapses easily but fail to react to smaller effects such as loss of variation in colors or textures. The experimental results achieved state of the art inception scores of 8.80 for the CIFAR10 dataset and also created a high-quality version of the existing CELEBA dataset consisting of 30000 of the images at 1024 × 1024 resolution. However, they maintain that there is still a long way to true photorealism and room for improvement in the micro-structure of the images.

Andrew Beers et al. [14] took this concept of PG-GANs a little further and used them specifically for medical image synthesis and the resolution of medical image data. The same training scheme of progressively growing of GANs has been used to create photorealistic and phenotypically diverse fine-grained images at high resolution. The GAN is trained in phases with each phase adding an upsampling layer and a pair of convolutional layers to both the discriminator and generator using an Adam optimizer and the loss calculated as Wasserstein Loss. In this experiment, High quality and variation images were produced including some very unrealistic images. They did suffer from some overly distinct edges mostly caused by the pressure on the generator to create segmentation maps but importance was given to certain pathological features. PGGAN had the ability to produce a great variety of images and the ability to generate vessel trees from outside its original training set. It was also observed that the latent space of GANs often encodes semantic information about the images produced, and that latent vectors similar to each other in latent space produce qualitatively similar output images. This method can produce images of unprecedented size and its latent space can be used to learn imaging features in an unsupervised manner for high resolution imaging.

Huikai et al. [15] developed a high resolution conditional image framework called GP-GAN that uses both GANs and gradient based image blending methods. They build a network called Blending GAN for generating low-resolution realistic images by proposing the Gaussian-Poisson equation to combine gradient information and colour information. Blending GAN leverages Wasserstein GAN for supervised learning tasks using the encoder-decoder architecture. Gaussian-Poisson equation fashioned by the well known Laplacian Pyramid is proposed to make use of the natural images produced by Blending GAN to generate high resolution realistic image by approximating the color. The conditional GAN is good at generating natural images from a particular distribution but weak in capturing high frequency image details like textures and edges. Gradient based methods on the other hand perform well at generating high-resolution images with local consistency but generated images tend to be unnatural and have many artifacts. Combined together, these methods could result in a conditional image generation system and also for image-to-image translation tasks. The drawback of the algorithm is that it fails to generate realistic images when the composited images are far away from the distribution of the training dataset.

In order to boost network convergence and achieve good looking high resolution results, Curt´o et al. [16] proposed a model called HDCGAN, with the goal being to generate indistinguishable samples to push GANs to scale well and maintain context information of high resolution images. They create an image generation tool that samples from a very precise distribution whose instances resemble or highly correlate with real sample images of the underlying true distribution. These generated image points fit well into the originals and add additional information such as redundancy, poses or generate highly-probable scenarios. Self-normalizing Neural Networks (SNNs) keep the activations normalized when propagating through the layers and Scaled Exponential Linear Units (SELU) is used as the activation function for feed forward neural networks to construct a mapping leading to SNNs. SELU + Batchnorm is used when numerical points move away from the usual point. BatchNorm ensures it is close to a desired value thus maintaining convergence. This technique stabilizes training, allows fewer GPU resources with steady diminishing errors in the generator and discriminator thus accelerating convergence speed. High-Resolution Deep Convolutional Generative Adversarial Networks (HDCGAN) by stacking SELU + BatchNorm (BS) layers generates high-resolution images in circumstances where all other former methods fail. It exhibits a steady and smooth training mechanism.

Bousmalis et al. [6] use unsupervised learning to learn a transformation in the pixel space from one domain to another. The model makes source-domain images appear as if they are drawn from the target domain. They train a model to change images from the source to appear as if they are from the target domain, while maintaining the original content and this is known as pixel level domain adaptation method or PixelDA. In their model they make use of content similarity loss that penalizes the difference between source and generated image for foreground pixels. They also use pairwise mean squared error (PMSE) which penalizes the difference between pairs of pixels rather than absolute difference between input and output. The different domain adaptation scenarios that were considered are MNIST to USPS, MNIST to MNIST-M and Synthetic Cropped LineMod to Cropped LineMod. They perform a quantitative and qualitative evaluation on the different domain adaptation scenarios. qualitative evaluation involves the examination of the ability of the method to learn the underlying pixel adaptation process from the source to the target domain by visually inspecting the generated images. Qualitative evaluation involves comparison of PixelDA with Source only(train on source training data and evaluate on target test data) and Target only(train on target training data and evaluate on target test data) baselines. The PixelDA models outperforms previous work on a set of unsupervised domain adaptation scenarios, and in the case of the challenging "Synthetic Cropped Linemod to Cropped Linemod" scenario, the model more than halves the error for pose estimation

Phillip et al. [17] investigate conditional GANs as a general purpose solution for image to image translation problems because they learn a loss that adapts to the data. In other words, they learn a structured loss which penalizes the entire output instead of treating each output pixel independently. Conditional GANs learn a mapping from the combination of observed image and random noise vector to output image.The noise is provided in the form of dropout to the generator during both training and test time. In image translation, a lot of low level information is common between the input and output images. In order to move this directly across the network, skip connections are added to the generator which follow the shape of a U-Net [18]. The discriminator is a PatchGAN which penalizes the image structure at the patch level. The discriminator predicts real or fake images by averaging all the responses. The model is evaluated on toe metrics, AMT perceptual studies which involves testing plausibility to a human observer and FCN-score which uses an off the shelf classifier to see how well the synthesized images can be classified. The results showed that using GANs along with L1 loss function fooled participants 18.9% which is significantly higher than previous methods. However designing conditional GANs that produce highly stochastic output which capture the full entropy of the conditional distributions they model is an important question left open by the present work.

3.6 Analysis

Based on our analysis of various super resolution techniques, we find that a lot of the challenges in super resolution have been addressed and rectified. However there are still some challenges or limitations that need to be addressed.

1. Noise and motion blur
2. Image quality preservation
3. Oversmoothing in images

Interpolation, although simple to implement leaves much to be desired in terms of visual quality, as the details (eg:- sharp edges) are often not preserved. Sparse coding makes use of low resolution and high resolution images, however the pipeline involves multiple steps all of which can not be optimized.

4. EXPERIMENTS

In an attempt to understand the performance of the various super resolution techniques, we chose different architectures to measure the performance of the three techniques, namely, reconstruction based, interpolation based and example based super resolution.

We used the flickr8k dataset across all three methods. All images were cropped to dimensions of 300x300 and we perform 4x upscaling of images. We measure the Peak Signal to Noise Ratio (PSNR) in order to compare the methods.

For interpolation techniques, we perform bilinear, bicubic, area, nearest neighbour and Lanczos interpolation.

For a reconstruction based method using CNN to produce low resolution images, we perform reconstruction of the low resolution features to higher resolution images.

For example based methods we implemented a deep convolutional network which uses skip connections to preserve image quality. The PSNR as well as SSIM values for these turned out to be better overall for this method than for the previous methods. In terms of subjective quality, the images generated by the deep network are more visually pleasing than those generated by the other methods.

5. RESULTS

Our experiments show that example based techniques produce better images than interpolation and reconstruction based methods. Example based techniques also produce higher PSNR and SSIM scores than interpolation and reconstruction based methods. Among the different interpolation techniques, bicubic interpolation consistently outperforms the other interpolation techniques, producing a higher PSNR and SSIM score.

	Interpolation	Reconstructio n	Example -based
PSNR	33.642	32.90	32.067

Table1. PSNR and SSIM values comparison

Table 1. shows the PSNR and SSIM values for a particular image of our professor that we obtained from the internet. It can be observed that the PSNR values are higher for Interpolation and Reconstruction based methods while the SSIM values are better(higher the better) in Example-based method.

6. CONCLUSIONS

Through our experiments we see that example based techniques produce images that are of better quality than images produced by interpolation and reconstruction based techniques.

But even though our experimental results do produce good quality images for experiment based techniques and especially GANs, the PSNR values don't seem to be the right measure for proving the experimental results. While the results obtained by deep learning are more pleasing to the eye compared to other methods, they suffer from PSNR statistics. We would like to analyse why deep learning methods suffer from this and explore different loss functions to make PSNR a suitable evaluation metric for deep learning methods.

7. NEXT STEPS

The aim of our research project is to be able to generate high resolution images using example based methods. By implementing ideal GAN or CNN architectures using the right activation and loss functions, our model should be able to yield images of high quality and variety. We plan on using an image Dataset like ImageNet and then train a GAN model known to produce the best results. From the literature survey, we can observe that PG-GAN is observed to generate images of high quality and also follow a continuous learning process and thus helps the discriminator perform a better job at identifying new samples generated by the generator. We plan on also tweaking and adding changes to the basic GAN Loss function by merging it with the Wasserstein Loss and the Adam Optimizer. It has been established that PG-GAN does not yet produce photorealistic images or generate micro-structure details of images. Thus using PG-GAN along with a strong loss function, our goal of research in improving resolution of images could be realized hoping our model succeeds in producing great results. The main goal is to use a different accuracy measure other than the PSNR value to determine the performance of example based techniques and realize its power and conclude our model as state of the art.

REFERENCES

[1] Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in Advances in Neural Information Processing Systems, 2014, pp.2672–2680.

[2] E. L. Denton, S. Chintala, R. Fergus et al., "Deep generative image models using a laplacian pyramid of adversarial networks,"in Advances in Neural Information Processing Systems, 2015, pp.1486–1494.

[3] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," in Proceedings of the 5th International Conference on Learning Representations (ICLR) - workshop track, 2016.

[4] V. Dumoulin, I. Belghazi, B. Poole, O. Mastropietro, A. Lamb, M. Arjovsky, and A. Courville, "Adversarially learned inference," in (accepted, to appear) Proceedings of the International Conference on Learning Representations, 2017.

[5] J. Donahue, P. Krähenbühl, and T. Darrell, "Adversarial feature learning," in (accepted, to appear) Proceedings of the International Conference on Learning Representations, 2017.

[6] K. Bousmalis, N. Silberman, D. Dohan, D. Erhan, and D. Krishnan, "Unsupervised pixel-level domain adaptation with generative adversarial networks," in IEEE Conference on Computer Vision and Pattern Recognition, 2016.

[7] S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee, "Generative adversarial text to image synthesis," in International Conference on Machine Learning, 2016. [Online]. Available: https://arxiv.org/abs/1605.05396

[8] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to image translation using cycle-consistent adversarial networks," in Proceedings of the International Conference on Computer Vision, 2017. [Online]. Available: https://arxiv.org/abs/1703.10593

[9] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in IEEE Conference on Computer Vision and Pattern Recognition, 2017.

[10] J. Liu, M. Spero, A. Raventos, "Super-Resolution on Image and Video,"

[11] X. Xu, D. Sun, J. Pan, Y. Zhang, H. Pfister and M. Yang, "Learning to Super-Resolve Blurry Face and Text Images," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, 2017, pp. 251-260.

[12] T. Karras, T. Aila, S. Laine, J .Lehtinen (NVIDIA and Aalto University), "Progressive Growing of GANs for Improved Quality, Stability, and Variation," in ICLR Conference on Computer Graphics, Machine Learning and Artificial Intelligence, 2018.

[13] Diederik P Kingma, Tim Salimans, Rafal Jozefowicz, Xi Chen, Ilya Sutskever, and Max Welling. Improved variational inference with inverse autoregressive flow. In NIPS, volume 29, pp. 4743–4751. 2016.

[14] A. Beers, J. Brown, K. Chang, J. P. Campbell, S. Ostmo, M. F. Chiang, and J. Kalpathy-Cramer. High resolution medical image synthesis using progressively grown generative adversarial networks. arXiv preprint arXiv:1805.03144, 2018.

[15] Huikai Wu, Shuai Zheng, Junge Zhang, and Kaiqi Huang. Gp-gan: Towards realistic high resolution image blending. arXiv preprint arXiv:1703.07195, 2017.

[16] Curtó, J.D., Zarza, I.C., Torre, F.D.L., King, I. and Lyu, M.R. [2018] High-resolution deep convolutional generative adversarial networks. arXiv:1711.06491v9 [cs.CV].

[17] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In CVPR, 2017

[18] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In MICCAI, pages 234–241. Springer, 2015

[19] T.D Harris, R.D Grober, J.K Trautman, Super Resolution Image Spectroscopy, 1964

[20] Image Restoration: Fundamentals and Advance - edited by Bahadir Kursat Gunturk, Xin Li

[21] L. Yue, H. Shen, J. Li, Q. Yuan, H. Zhang, and L. Zhang, "Image super-resolution: The techniques, applications, and future"

[22] Sabyasachi Moitra, "Single-Image Super-Resolution Techniques: A Review"

[23] Aparna K V1 , Lisha P P2, "A Survey on Super-Resolution Techniques"

[24] Kathiravan Srinivasan, "A Study on Super-Resolution Image Reconstruction Techniques"

[25] Sina F, Dirk R, Michael E, Peyman M, "Advances and Challenges in Super-Resolution"

[26] Lisha PP, Jayasree VK, "Single image super-resolution - A quantitative comparison"