







Real-World Thermal Image Super-Resolution

Moaaz Allahham¹(✉) , Andreas Aakerberg¹ , Kamal Nasrollahi^{1,2} ,
and Thomas B. Moeslund¹ 

¹ Visual Analysis and Perception Laboratory, Aalborg University, Aalborg, Denmark

² Research Department, Milestone Systems A/S, Milestone Systems,
Brøndby, Denmark

{anaa, kn, tbm}@create.aau.dk

Abstract. Thermal cameras are used in various domains where the vision of RGB cameras is limited. Thermographic imaging enables the visualizations of objects beyond the visible range, which enables its use in many applications like autonomous cars, nightly footage, military, or surveillance. However, the high cost of manufacturing this type of camera limits the spatial resolution that it can provide. Real-World Super-Resolution (RWSR) is a topic that can be used to solve this problem by using image processing techniques that enhance the quality of a real-world image by reconstructing lost high-frequency information. This work adapts an existing RWSR framework that is designed to super-resolve real-world RGB images. This framework estimates the degradation parameters needed to generate realistic Low-resolution (LR) and High-resolution (HR) image pairs, then the SR model learns the mapping between the LR and HR domains using the constructed image pairs and applies this mapping to new LR thermal images. The experiments results show a clear improvement in the perceptual quality in terms of clarity and sharpness, which surpasses the performance of the current SotA method for thermal image SR.

Keywords: Thermal imaging · Super-resolution · Image registration

1 Introduction

In recent years, thermal imaging has grown considerably and is being used in various domains where a typical RGB camera can not get the job done, like nightly footage, surveillance, or in autonomous cars. However, thermal images generally have some shortcomings like insufficient details and blurred edges, and most importantly considerably low-resolution. This makes it too hard to observe the structure and recognize objects in an image. However, having a thermal camera that is capable of capturing high-resolution images is not as affordable as using RGB cameras. Even the most expensive thermal cameras, which can vary from US\$200 to more than US\$20,000 [20], still can not deliver sufficient resolutions. To the best of our knowledge, the highest resolution that a thermal camera can

provide as for today is 1920×1200 pixels for the Vayu HD [24], thus enhancing real images captured by thermal cameras is therefore important. However, although increasing the resolution of a thermal image with an image processing algorithm would not compensate for the true information that is not captured by the camera’s sensor, having an enhanced and higher resolution image makes it easier to recognize objects and structure in an image. The efficiency of this process can be improved by taking advantage of computer vision techniques that can assist in enhancing these images. Many methods were developed to perform image super-resolution, however, most of these methods perform poorly when used on real LR images. This is because they follow the approach of downsampling HR images to construct LR and HR pairs and then they super-resolve the LR image back to match the HR image quality. Such methods fail when given a real-world image as the degradation process is not entirely known. Therefore recent studies have been working on developing methods that would be more robust to previously unseen real-world images that are acquired directly from cameras with unknown degradation parameters. This RWSR issue also applies to the thermal imaging domain, making it an interesting area to investigate since it has not been widely explored. Hence, the goal of this project was to explore the state-of-the-Art (SotA) SR algorithms that deal with RGB images and investigate its usability in the thermal imaging domain, and explore the possibility of tuning these methods to fit the thermal domain. The main contributions of this work are:

- A comparison of the performance of existing RGB-based RWSR solutions in the thermal imaging domain.
- SotA results within the real-world thermal SR domain are achieved.

2 Related Work

2.1 RGB Image Super-Resolution

Zero-Shot Methods. In 2017, ZSSR [23] was introduced as the first blind SR algorithm (self-learning-based) that performed SR on LR real-world images without relying on any prior image examples or prior training. Instead, ZSSR trains an image-specific CNN using the recurrence of small patches across different scales within the same image at test time. This was done by downscaling the test image to smaller versions of itself, then applying data-augmentation to the smaller versions to fulfill the need of having multiple examples as a training dataset. The image-specific CNN learns to reconstruct the original LR image using the downscaled examples, then they finally apply the trained CNN to the original test image to construct the desired HR output. ZSSR outperformed external-based SotA methods in some regions when tested on images with salient recurrence of information. A drawback of ZSSR is the fact that the learning process fully depends on the internal information in the test image, which makes it require thousands of back-propagation gradient updates. This yields slow testing time as well as poor results in some regions compared to other external-based

methods [26]. Inspired from ZSSR, Meta-Transfer Learning for Zero-Shot Super-Resolution (MZSR) [26] was introduced, where the authors of MZSR utilize the powerful parts of ZSSR and improve upon it by introducing the concept of Meta-Transfer learning. The idea behind how meta-learning works is to make the model adapt fast to new blur kernel scenarios by adding a meta-training step, then utilize transfer-learning by pre-training the SR network using a large-scale dataset DIV2K [1]. The combination of Meta-transfer learning and ZSSR exploits both the internal (the test image) and external (the DIV2k) information. The main advantage that was introduced in the MZSR work, was the flexibility and fast running time compared to the ZSSR method, as well as outperforming other supervised SotA algorithms such as CARN [2] and RCAN [28]. Different zero-shot methods were designed following the ZSSR principle, however the most recent study that was able to achieve competitive SotA results was Dual Super-resolution (DualSR) [8]. DualSR addresses the RWSR problem in a similar way to the way it was addressed in the ZSSR work, where they learn the image-specific LR-HR relations by training their proposed network at the test time using patches extracted from the test image. Their proposed network is split into mainly two parts, the downsampler which learns the degradation process using a generative adversarial network (GAN), and an upsampler that learns to super-resolve the LR image. Both the up-sampler and down-sampler are trained simultaneously by improving each other using the cycle-consistency loss, the masked interpolation loss, and the adversarial loss.

Learned Degradation Based Super-Resolution. Many supervised SR approaches make the assumption that LR images are a bicubically downsampled version of their HR counterpart, and that Gaussian noise is usually used to simulate the sensor noise. However, these approaches fail when tested on real images because those images were not degraded using ideal degradation operation (bicubic kernel + Gaussian noise). For this reason, Fritsche et al. [9] introduced DSGAN (the winner of AIM2019 RWSR challenge [17]), which is a GAN network that learns to generate the appropriate LR images, which have the same corruptions as the original HR images. Bell-Kligler et al. [3] introduced another realistic degradation method KernelGAN, an image-specific InternalGAN, which trains solely on the LR test image at test time and learns its internal distribution of patches. The generator of the network is trained to produce a lower resolution image such that the network’s discriminator can not distinguish between the patch distribution of the generated image and the patch distribution of the original LR image. Ji et al. [11] proposed their method RealSR, which is divided into two stages. They first use KernelGAN to estimate the degradation from the real data and use it to construct the LR images, and then they train an SR model based on the constructed data. RealSR method was the winner of the NTIRE 2020 challenge [17], and by the time of doing this work, RealSR is considered to be the SotA in the real-world super-resolution field for RGB images.

2.2 Thermal Image Super-Resolution

All the methods mentioned in 2.1 are examples of super-resolution methods that deal with images in the RGB spectrum. However, there are only a few studies that developed methods for super-resolving LR thermal images. Cho et al. [4] conducted a study where they tried to enhance thermal images by training a CNN using different image spectrums aiming to find the best representation that would fit the thermal domain. They found that a grayscale trained network provided the best enhancement. Lee et al. [14] proposed a similar CNN-based on enhancement for thermal images, where they evaluated four RGB-based domains with a residual-learning technique. That improved the enhancement in comparison to the previous work by [4]. Rivadeneira et al. [5] was motivated by the two previously proposed methods, so he proposed the Thermal Enhancement Network (TEN), which was the first CNN-based method to be trained specifically using thermal dataset unlike the two previous proposals by [4,14]. TEN was based on the SRCNN model [7], which utilizes the residual net and dense connections technique. TEN was able to outperform the previously proposed methods, which was due to training the network using thermal images instead of RGB-based domains. Recently, Rivadeneira et al. [20] proposed another thermal SR method that is based on the well-known CycleGAN [29] architecture. Two-way Generative-Adversarial-network (CycleGAN) is a technique that is used to map information from one domain to another. So the authors of [20] used the CycleGAN network to map information from the LR domain to the HR domain. They trained their proposed network to perform x2 scale SR following two scenarios, LR to medium-resolution (MR) and MR to HR. Chudasama et al. [6] proposed TherISuRNet, which is another method to super-resolve thermal images by progressively upscaling the LR test image to obtain the final SR image. They achieve different upscaling factors (x2, x3, and x4) by applying residual learning. The TherISuRNet network consists of four main modules: low-frequency feature extraction modules, high-frequency feature extraction modules, second high-frequency feature extraction modules, and finally an image reconstruction module that is responsible for reconstructing the final SR image. They measured the performance of their proposed method by comparing its performance to the most common SotA methods [5,15,16,18,28] and bicubic interpolation, and they were able to surpass all the other methods when testing on thermal images. TherISuRNet was the winning method of the Thermal Image Super-Resolution Challenge PBVS 2020 [19], which makes the TherISuRNet the SotA method for the thermal image SR domain.

Constraints Noted from Related Works. Having reviewed the relevant literature on super-resolution applied to both RGB and thermal images, we witness that, to the best of our knowledge:

- None of the studies try to investigate the performance of RGB-based SR methods in the thermal domain.
- All the existing thermal SR methods were trained using synthetically constructed image pairs.

3 Dataset

One of the challenges when working with RWSR methods is the lack of ground truth data that could be used for supervised learning and to evaluate the performance of the SR methods leading to unreliable performance when testing on single real world images. For this work, the PBVS dataset [19,20] was used as it offers three subsets called *Domo*, *Axis* and *GT* with different native resolutions (160×120 , 320×240 , 640×512 , respectively), which were acquired using three different cameras. For this work, the Domo and GT subsets are used as the source and target domains respectively. Each of these subsets includes a total of 951 training images and 50 images for validation. The Axis subset was discarded since the goal of this work was to super-resolve a given resolution with an upscaling factor of $s = 4$ and later evaluate the performance by comparing it to the ground-truth, which has a native resolution that matches the SR output images. Therefore, it was decided to super-resolve the input images (Domo validation subset) and compare the output with the ground truth (GT validation subset). However, one of the problems with the PBVS dataset is the limited number of images in each subset, which is considered too little to be used for training a neural network. Therefore, we used the augmented version of the PBVS dataset, which was provided by the authors of the TherISuRNet [6]. The augmentation operations they apply on the original dataset are horizontal flipping, 180° rotation, and two affine operations, resulting in a total of 4755 training images for each subset.

4 Thermal RealSR

This section describes the two-step pipeline that T-RealSR uses to achieve the final SR results. The first step aims to realistically degrade the HR from the target domain Y , such that the degraded images have the same image characteristics as the LR images in the source domain X . The second step is to use the LR-HR image pairs to train a SR model that can be used to super-resolve real-world thermal images.

4.1 Realistic Degradation Using KernelGAN and Noise Injection

To understand how we can construct a realistic LR image that does not have ideal blurring and noise characteristics, let's assume an LR image is obtained following the degradation operation [11]:

$$I_{LR} = (I_{HR} * k) \downarrow_s + n \quad (1)$$

Where k denotes the kernel used to blur the image, n denotes the noise added to the image, and s denotes the downscaling factor. Instead of using ideal kernels (e.g. Bicubic downscaling), T-RealSR explicitly utilizes KernelGAN to create a pool of kernels, and it extract noise patches from a real LR images to create a noise patches pool. Then both these pools are used to construct the realistic LR-HR image pairs.

Kernel Degradation. In general, KernelGAN is an image-specific InternalGAN [22] that trains solely on a given LR image at test time and learns its internal distribution of patches. Its generator (G) is trained to generate a down-scaled version of the given image, such that its discriminator (D) can not distinguish between the patch-distribution of the generated image and the patch distribution of the original image. D is trained to output a heat map, referred to as D -map, indicating for each pixel how likely is its surrounding patch to be drawn from the original patch-distribution. The loss is the pixel-wise MSE difference between the output D -map and the label map. Where the label map is all the ones in the crops extracted from the original image, and all the zeros in the crops extracted from the downscaled image [3].

Noise Extraction. In addition to creating the kernel pool, T-RealSR introduces a simple filtering rule for extracting noise patches from source images. The idea behind extracting these noise patches is to inject them into the degraded images, so LR images from the two different domains (source LR and generated LR images) will have similar noise distribution. The filtering rule used to choose the relevant noise patch is as follows:

$$\sigma(n_i) < v \quad (2)$$

Where $\sigma(\cdot)$ denotes the function used to calculate the noise variance, and v is the max value of variance.

Having created a series of kernels $\{k_1, k_2, \dots, k_l\}$ and a series of noise patches $\{n_1, n_2 \dots n_m\}$, the degradation process is performed as follows:

$$I_{LR} = (I_{HR} * k_i) \downarrow_s + n_j, i \in 1, 2, \dots, l, j \in 1, 2, \dots, m \quad (3)$$

Where s denotes the sampling stride.

4.2 Super-Resolution Model

As mentioned in Sect. 2.1, T-RealSR consists of two phases, the first is constructing the realistic image pairs using KernelGAN and the second phase is training the SR model, which is based on ESRGAN with some modification. To understand the T-RealSR SR backbone, we need to first understand how ESRGAN works and then understand how T-RealSR adjust the ESRGAN architecture to make it more flexible to different image sizes. ESRGAN [27] stands for Enhanced Super-Resolution Generative Adversarial Networks, which is a generative adversarial network that is based on SRGAN [13]. SRGAN is a GAN network that is capable of generating realistic textures during single-image SR, whose discriminator aims to base its prediction on perceptual quality. However, ESRGAN improves SRGAN by adjusting the SRGAN architecture where they introduce their Residual-in-Residual Dense Block (RRDB) without batch normalization, as well as improving the SRGAN discriminator by making it judge whether an image is more realistic than another rather than judging whether an image is

real or fake. ESRGAN improvement over SRGAN resulted in sharper and more visually pleasing results [27].

From the name Enhanced Super-Resolution GAN, we can tell that the architecture should contain the two main modules, discriminator D and generator G networks. The G network takes a low-resolution image (LR) as input, and it passes it through a 2D convolutional layer (Conv1) with small 3×3 kernels and 64 feature maps. It is then passed through 23 Residual in Residual Dense Blocks (RRDB). The image is then passed through another convolutional layer (Conv2) in which its output is summed with the output of the first (Conv1). At this stage, the image gets upscaled with a factor of 4 by passing it through an upsampling block that consists of two convolutional layers for reconstruction, with LeakyReLU (LReLU) activation ($\alpha = 0.2$) on each layer. After upsampling, the image is passed through another convolutional layer (Conv3) with LReLU activation ($\alpha = 0.2$). Finally, the image is passed through the final convolutional layer (Conv4) that final super-resolved image. The other part of the network is the discriminator D , and to be more specific it is called the Relativistic Discriminator [12]. Following [27] this specific discriminator was used rather than using the standard discriminator used in SRGAN [13]. This is because the relativistic discriminator estimates the probability that a real image x_r is relatively more realistic than a fake one x_f . Where a standard discriminator estimates only whether an image x is natural enough to be real.

We adapted the ESRGAN structure and trained it using the constructed paired data $\{I_{LR}, I_{HR}\}$. Several losses were used during the training including:

- **Pixel loss** L_1 : or so called Mean Absolute Error (MAE), which measures the mean absolute pixel difference of all pixels in two given images.
- **Perceptual loss** L_{per} : proposed to enhance the visual quality by minimizing the error in feature space instead of pixel space. It uses the inactive features of VGG-19 [25] and aims to enhance the visual quality of low-frequency information like edges.
- **Adversarial loss** L_{adv} This loss is used to enhance the texture details to make the image look more realistic.

The final loss function was the weighted sum of all the above losses as follows:

$$L_{total} = \lambda_1 L_1 + \lambda_{per} L_{per} + \lambda_{adv} L_{adv} \quad (4)$$

where λ_1 , λ_{per} , and λ_{adv} are constants used to specify the weight of each of the losses on the total loss.

PatchGAN Discriminator. The discriminator (VGG-128) used in the ESRGAN may introduce many artefacts, so PatchGAN [10] was used instead for two reasons: First is that VGG-128 used by ESRGAN limits the size of the generated image to 128, making multi-scaling training not as simple, Second is that the VGG-128 fixed fully connected layer makes the discriminator pays more attention to the global features and ignore the local ones. Where the PatchGAN has a fully convolutional structure that maintains a fixed receptive field that restricts

the discriminator’s attention to the local image patches. The structure of PatchGAN only penalizes structure at the scale of patches, meaning that it tries to classify if each $N \times N$ patch in an image is real or fake. The responses of all patches get averaged afterward forming the final D output to guarantee global consistency, then gets fed back to the generator.

5 Experiments and Results

5.1 Evaluation Metrics

Usually, the most challenging part when dealing with RWSR images is the lack of GT reference images. However, despite having the GT images, which the PBVS dataset provides, the SR and GT images are not perfectly aligned together. Making it difficult to use reference-based IQA methods such as SSIM, PSNR, or LPIPS, however we still use them for reference purposes. Additionally, it was decided to take another evaluation approach by following the IQA evaluation protocol from the NTIRE2020 challenge, where they used non-reference-based IQA methods including PIQE, NIQE, and BRISQUE. In addition to that, the Mean Opinion Score (MOS) method was used to support the previously mentioned non-reference-based methods, which correlate poorly with human opinion. For the MOS, a total of 20 participants were given a set of 13 SR images that were generated using different methods. Then the participants were asked to give unique scores that range between 1 and 6 (best to worse respectively) to each individual images based on the perceived clarity and sharpness of the images. The results of 6 different SR methods were used, where the methods were shuffled randomly when presented to the participants to avoid bias. The scores were then averaged for the individual images for each method, and were then used to calculate the final MOS scores.

5.2 Comparison with the State of the Art

To the best of our knowledge, an evaluation of the adapted T-RealSR method as well as the other mentioned SotA SR rgb-methods within the thermal domain, in comparison to the SotA thermal SR method has not be done before. Therefore, we compare the adapted method to bicubic upscaling, as well as with a number of RWSR methods including two zero-shot SR methods (DualSR [8], KernelGAN+ZSSR [3, 23]) and the ESRGAN [27] RWSR method, and for the thermal SotA SR method TherISuRNet [6]. To ensure a fair comparison, ESRGAN [27] was retrained using the same dataset used to train the adapted T-RealSR, and employing the settings suggested by the authors of the ESRGAN. For DualSR [8] and KernelGAN+ZSSR [3, 23], a training is not needed, as it is a part of the inference phase; the settings suggested by the authors were used. For the TherISuRNet [6], the retraining was needed as pretrained weights were not provided by the authors, and the same settings were adapted because the method was designed specifically for the utilized PBVS dataset.

Image Registration. We explained in Sect. 3 how the PBVS subsets (Domo and GT) were acquired using different cameras. Despite the effort by the authors to acquire two identical pictures of the same scene using different cameras, the process was physically impossible. That introduced some challenges when having to evaluate the performance of the different SR methods. Besides the different light conditions and different sensors’ noise that resulted in brightness and contrast differences, the images were not perfectly aligned together. The imperfect alignment of the images meant that reference-based IQA methods in general and PSNR in specific, will be inaccurate to be used on their own. Therefore, we decided to apply image registration between the SR images and the GT reference images prior to evaluating the images using the non-reference-based methods. To do so, the ORB detector [21] with a target number of features $N = 5000$ was used to align the images together as illustrated in Fig. 1. The central crop (50%) of both the SR and GT images was used for evaluation. This was done to discard the black areas around the registered images and to make the comparison as fair as possible, since lens distortion is at its minimum in the central part of the image.

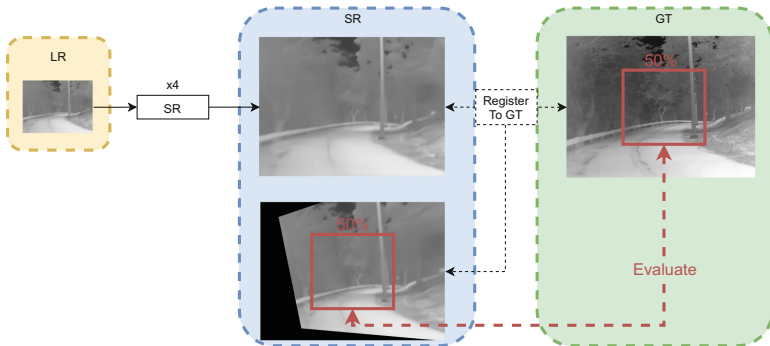


Fig. 1. The evaluation pipeline used to evaluate the super-resolved LR image in comparison to the GT.

Quantitative and Qualitative Evaluation. We evaluate the performance of the methods on the PBVS test dataset, where we show the quantitative results in Table 1. For the qualitative results a number of patches taken from some test images are shown in Fig. 2. The adapted T-RealSR method outperforms the other thermal and rgb-based SR methods by a large margin. Where it is possible to see that the traditional non-reference-based IQA methods (PIQE, NIQE, BRISQUE) correlate well with the human-opinion based MOS method. However, the reference-based IQA methods (SSIM and PSNR) correlate poorly with the other IQA methods. This is due to brightness and contrast differences. A method such as PSNR, will penalize the performance in case the registered image is shifted one pixel in any direction, and we know for sure that this is most likely the case with our test data.

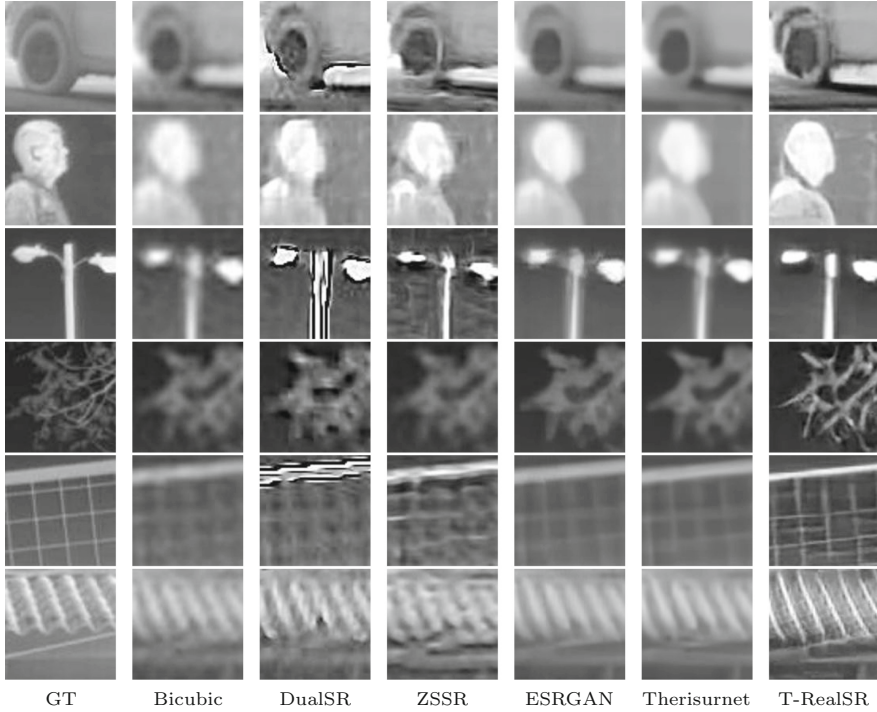


Fig. 2. Qualitative comparison of SotA methods for x4 SR of real LR images from the Domo validation subset.

Table 1. Comparison between the SotA methods that have been tested. The best values are in bold text.

Method	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	PIQE \downarrow	NIQE \downarrow	BRISQUE \downarrow	MOS \downarrow
Bicubic	20.11	0.70	0.46	67.39	5.55	57.20	4.10
DualSR [8]	18.77	0.59	0.43	56.48	4.18	43.03	4.74
ZSSR+KernelGAN [3]	19.01	0.57	0.44	60.79	5.71	46.14	4.15
ESRGAN [27]	18.37	0.65	0.43	76.77	5.72	53.74	2.98
TherISuRNet [6]	20.10	0.71	0.42	88.69	5.20	55.34	3.20
T-RealSR [11]	18.78	0.52	0.37	36.33	3.31	34.31	1.45

6 Conclusion

In this work we investigate the possibility of using rgb-based RWSR methods to super-resolve real-world thermal images. The images used for evaluation were upscaled with a factor of 4, and we found that tuning the T-RealSR by training it using thermal images is able to achieve SotA performance that surpasses the current SotA thermal-based SR method by a large margin in terms of perceived quality. This was proven by the different IQA methods, which showed results that

correlate with the human-based MOS evaluation method. This work is, up to our knowledge, the first work that train on thermal images using realistically degraded image pairs, making it robust to real images that contain some of the most common degradation types (blurring and sensor noise).

References

1. Agustsson, E., Timofte, R.: NTIRE 2017 challenge on single image super-resolution: dataset and study. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, July 2017
2. Ahn, N., Kang, B., Sohn, K.A.: Fast, accurate, and lightweight super-resolution with cascading residual network. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 252–268 (2018)
3. Bell-Kligler, S., Shocher, A., Irani, M.: Blind super-resolution kernel estimation using an internal-gan (2020)
4. Cho, Y., Bianchi-Berthouze, N., Marquardt, N., Julier, S.J.: Deep thermal imaging. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, April 2018. <https://doi.org/10.1145/3173574.3173576>
5. Choi, Y., Kim, N., Hwang, S., Kweon, I.S.: Thermal image enhancement using convolutional neural network. In: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 223–230. IEEE (2016)
6. Chudasama, V., et al.: Therisurnet - a computationally efficient thermal image super-resolution network. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, June 2020
7. Dong, C., Loy, C.C., He, K., Tang, X.: Image super-resolution using deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **38**(2), 295–307 (2015)
8. Emad, M., Peemen, M., Corporaal, H.: Dualsr: Zero-shot dual learning for real-world super-resolution. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), pp. 1630–1639, January 2021
9. Fritsche, M., Gu, S., Timofte, R.: Frequency separation for real-world super-resolution. In: 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), pp. 3599–3608 (2019)
10. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1125–1134 (2018)
11. Ji, X., Cao, Y., Tai, Y., Wang, C., Li, J., Huang, F.: Real-world super-resolution via kernel estimation and noise injection. In: The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, June 2020
12. Jolicœur-Martineau, A.: The relativistic discriminator: a key element missing from standard GAN (2018)
13. Ledig, C., 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, pp. 4681–4690 (2017)
14. Lee, K., Lee, J., Lee, J., Hwang, S., Lee, S.: Brightness-based convolutional neural network for thermal image enhancement. *IEEE Access* **5**, 26867–26879 (2017). <https://doi.org/10.1109/ACCESS.2017.2769687>
15. Li, J., Fang, F., Mei, K., Zhang, G.: Multi-scale residual network for image super-resolution. In: The European Conference on Computer Vision (ECCV), September 2018

16. Lim, B., Son, S., Kim, H., Nah, S., Lee, K.M.: Enhanced deep residual networks for single image super-resolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 136–144 (2017)
17. Lugmayr, A., et al.: NTIRE 2020 challenge on real-world image super-resolution: Methods and results. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 494–495 (2020)
18. Park, S.J., Son, H., Cho, S., Hong, K.S., Lee, S.: Srfeat: single image super-resolution with feature discrimination. In: Proceedings of the European Conference on Computer Vision (ECCV), September 2018
19. Rivadeneira, R.E.: Thermal image super-resolution challenge - PBVS 2020. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 432–439 (2020). <https://doi.org/10.1109/CVPRW50498.2020.00056>
20. Rivadeneira, R.E., Sappa, A.D., Vintimilla, B.X.: Thermal image super-resolution: a novel architecture and dataset. In: International Conference on Computer Vision Theory and Applications, pp. 1–2 (2020)
21. Rublee, E., Rabaud, V., Konolige, K., Bradski, G.: Orb: An efficient alternative to sift or surf. In: 2011 International Conference on Computer Vision, pp. 2564–2571 (2011). <https://doi.org/10.1109/ICCV.2011.6126544>
22. Shocher, A., Bagon, S., Isola, P., Irani, M.: Ingan: capturing and remapping the DNA of a natural image (2019)
23. Shocher, A., Cohen, N., Irani, M.: Zero-Shot super-resolution using deep internal learning (2017)
24. Sierra-Olympic: Vayu HD feature specification. https://sierraolympic.com/wp-content/uploads/2020/06/2020_VayuHD_Sell-Sheet_FINAL.pdf. Accessed 02 June 2021
25. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: International Conference on Learning Representations (2015)
26. Soh, J.W., Cho, S., Cho, N.I.: Meta-transfer learning for zero-shot super-resolution. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3516–3525 (2020)
27. Wang, X., et al.: Esrgan: enhanced super-resolution generative adversarial networks. In: Proceedings of the European Conference on Computer Vision (ECCV) Workshops (2018)
28. Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., Fu, Y.: Image super-resolution using very deep residual channel attention networks. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 286–301 (2018)
29. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2223–2232 (2017)