Thermal Image Super-Resolution Challenge - PBVS 2020

Rafael E. Rivadeneira Angel D. Sappa Boris X. Vintimilla
Lin Guo Jiankun Hou Armin Mehri Parichehr Behjati
Ardakani Heena Patel Vishal Chudasama Kalpesh Prajapati
Kishor P. Upla Raghavendra Ramachandra Kiran Raja Christoph Busch
Feras Almasri Olivier Debeir Sabari Nathan Priya Kansal
Nolan Gutierrez Bardia Mojra William J. Beksi

Abstract

This paper summarizes the top contributions to the first challenge on thermal image super-resolution (TISR) which was organized as part of the Perception Beyond the Visible Spectrum (PBVS) 2020 workshop. In this challenge, a novel thermal image dataset is considered together with state-of-the-art approaches evaluated under a common framework. The dataset used in the challenge consists of 1021 thermal images, obtained from three distinct thermal cameras at different resolutions (low-resolution, mid-resolution, and high-resolution), resulting in a total of 3063 thermal images. From each resolution, 951 images are used for training and 50 for testing while the 20 remaining images are used for two proposed evaluations. The first evaluation consists of downsampling the low-resolution, mid-resolution, and high-resolution thermal images by $\times 2$, $\times 3$ and $\times 4$ respectively, and comparing their super-resolution results with the corresponding ground truth images. The second evaluation is comprised of obtaining the $\times 2$ super-resolution from a given mid-resolution thermal image and comparing it with the corresponding semi-registered high-resolution thermal image. Out of 51 registered participants, 6 teams reached the final validation phase.

1. Introduction

Single image super-resolution (SR) is a challenging, ill-posed problem, that is still solved using conventional methods. In recent years, deep learning techniques have shown better results. Most of these methods have been largely used in the visible spectral domain. In contrast to visible spectrum images, thermal images tend to have poor resolution which could be improved by using learning-based traditional SR methods. These methods work by down-sampling and adding noise and blur to the given image. The poor quality noisy and blurred images, together with the given ground truth images, are used in the learning process.

The approach mentioned above has been frequently used to tackle the SR problem, however there are few contributions where the learning process is based on the usage of a pair of images (low and high-resolution images) obtained from different cameras. A novel thermal image dataset has been created containing images with three different resolutions (low-resolution (LR), mid-resolution (MR), high-resolution (HR)) obtained with three distinct thermal cameras.

The TISR Challenge[1] consists of creating a solution capable of generating a SR thermal image in $\times 2$, $\times 3$, and $\times 4$ scales from cameras with different resolutions, in the conventional way by downsampling, and adding noise to the given ground truth image. Additionally, a $\times 2$ SR image must be generated from the image obtained with a MR camera. This $\times 2$ SR image is evaluated with respect to the corresponding image obtained from a HR camera.

The remainder of this paper is organized as follows. Section 2 introduces the objectives of the challenge, and presents the dataset and evaluation methodology. Section 3 summarizes the results obtained by the different teams. In Section 4, a short description of each teams' approach is provided. Finally, the paper is concluded in Section 5

2. TISR Challenge

The objectives of the TISR challenge are the following: (i) promote state-of-the-art approaches for the SR problem

Rafael E. Rivadeneira* (rrivaden@espol.edu.ec), Angel D. Sappa*+ and Boris X. Vintimilla* are the TISR Challenge - PBVS 2020 organizers, while the other authors participated in the challenge.
*Escuela Superior Politécnica del Litoral, ESPOL, Guayaquil, Ecuador.
+Computer Vision Center, Campus UAB, 08193 Bellaterra, Barcelona, Spain.
Appendix A contains the authors' teams and affiliations.

[1]http://vcipl-okstate.org/pbvs/20/challenge.html

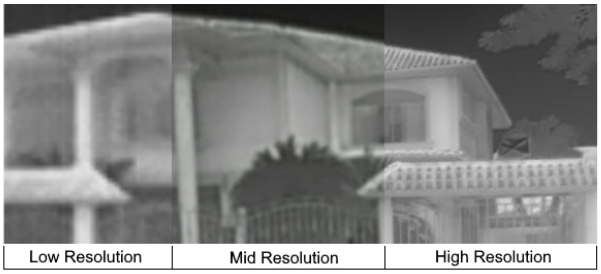

Figure 1. A mosaic with three different resolution thermal images from each camera for visual comparison: $(left)$ crop from a LR image; $(middle)$ crop from a MR image; $(right)$ crop from a HR image [10].

in the thermal image domain; (ii) evaluate and compare the different solutions; and (iii) promote a novel thermal image dataset to be used as a benchmark by the community working on the thermal image SR problem.

2.1. Thermal Image Dataset

The dataset used in this challenge was recently presented in [10]. It consists of a set of 1021 thermal images acquired by using three thermal cameras with different resolutions. The dataset contains images from indoor and outdoor scenarios under various lighting conditions (e.g., morning, afternoon, and night) and objects (e.g., buildings, cars, people, vegetation). The cameras were mounted in a rig that minimizes the baseline distance between the optical axis such that the acquired images are almost registered. Figure 1 presents a mosaic obtained with images from each camera (i.e., LR, MR, and HR). The camera parameters are given in Table 1 and illustrations from each camera depicted in Figure 2.

2.2. Evaluation Methodology

Peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) measures are computed over a small region of the images in order to evaluate the performance of the proposed solution. In this challenge two kinds of evaluations are performed. For the first evaluation, a set of 10 down-sampled and noisy images from each resolution (LR, MR, and HR) are considered. Downsampling scale factors of $\times 2$, $\times 3$, and $\times 4$ are performed and Gaussian noise of 10% is added. Figure 3 presents an illustration of this first evaluation process.

The second evaluation consists of computing the PSNR and SSIM of the obtained SR images with respect to the corresponding ground truth images. The ground truth images are of the same resolution as the computed SR, but are acquired with a different higher resolution camera. For the second evaluation, a set of 10 MR images is considered. The obtained SR images, from the MR set, are compared with the corresponding HR images which have been acquired with another camera. SIFT, SURF, and ORB descriptors are used to acquire characteristic keypoints between the MR and HR thermal images. Obtaining the mapping parameters allows for overlapping the two images (for more details see [10]). The evaluation measures (PSNR and SSIM) are performed over a central cropped region of the image. Figure 4 illustrates the second evaluation process.

3. Challenge Results

From 51 participants registered in the challenge, 6 teams made it to the final phase and submitted results together with their corresponding extended abstracts. Table 2 shows the average results (PSNR and SSIM) for each team in the two evaluations. A brief description of the thermal SR approach proposed by each team is presented in Section 4. Information about the team members and their affiliations is provided in Appendix A. According to the figures presented in Table 2, the winner of the TISR Challenge - PBVS 2020 is the MLCV-Lab_SVNIT_NTNU team, who achieved the top results in most of the evaluation tasks. The COUGER AI team achieved the best results for the second evaluation. Teams that did not reach the baseline results (i.e., bicubic interpolation) were not considered in this report.

4. Proposed Approaches and Teams

This section briefly presents the approaches proposed by the different teams.

4.1. HPZ-OSU

HPZ-OSU team follows the s-LWSR super-resolution framework [6, 5, 13], where the images are processed in small patches through residual networks. To deal with the added noise and the noise from thermal images' nature, a noise reduction process, referred to as PZ, is proposed. This approach has been originally designed for color noise, but it also works well with other kind of noise, such as white noise or Gaussian noise. The proposed network has been trained for the three super-resolution tasks of Evaluation 1.

The proposed architecture is presented in Figure 5; it first converts the image into the YUV space, then applies two-time-filtering along the horizontal and vertical direction. For each target pixel, the filter kernel is defined by the YUV values of the local pixels together with the relative distances to the target pixel.

All experiments were performed on a workstation with an 8GB NVIDIA 1070 GPU, using the Python programming language with PyTorch as a platform. The given dataset has been split up into 800 images for training and 151 images for testing. For the first evaluation, per each epoch, different noise values have been added, the model with best PSNR score has been selected. For the second

Figure 2. An example of thermal images acquired by each camera. $(left)$ LR image with 160×120 native resolution from an Axis Domo P1290. $(middle)$ MR image with 320×240 native resolution from an Axis Q2901-E. $(right)$ HR image with 640×480 resolution from an FC-6320 FLIR (native resolution is 640×512) [10].

Table 1. Thermal Camera Specifications (Note: HR images have been cropped to 640×480) [10].

Image Description	Camera Brand	FOV	Focal Length	Native Resolution	Total # of Images
Low (LR)	Axis Domo P1290	35.4	4mm	160×120	1021
Mid (MR)	Axis Q2901-E	35	9mm	320×240	1021
High (HR)	FC-632O FLIR	32	19mm	640×512*	1021

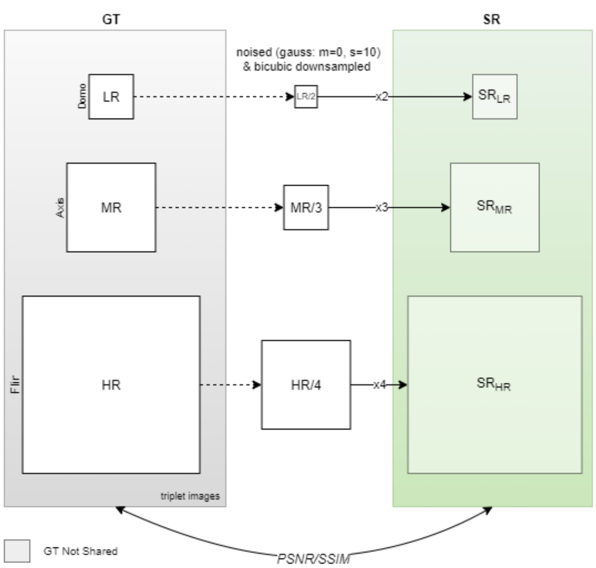

Figure 3. An illustration of the first evaluation process (×2 for low, ×3 for mid, ×4 for high).

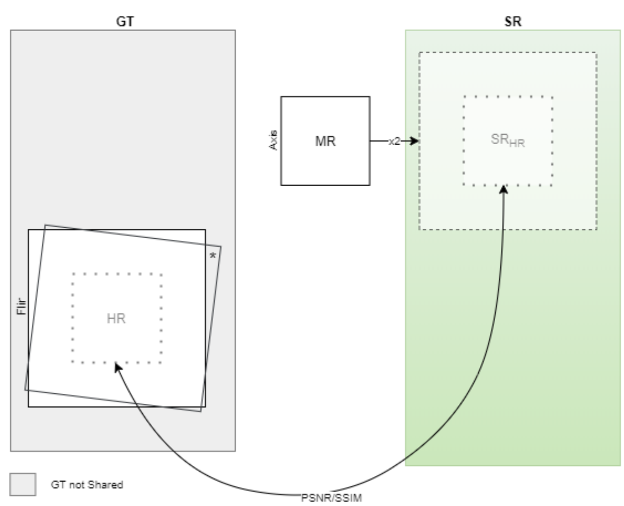

Figure 4. A illustration of the second evaluation process. Note that this evaluation is applied over a set of 10 MR images.

evaluation, the same network is trained using HR (640x480) images downsampled by 2.

4.2. CVC-UAB

CVC-UAB team proposes a Lightweight Multi-Path Residual Network (LMPRNet) intended for thermal image super-resolution. This architecture makes the network pay attention to learning more abstract features by letting abundant low-frequency features to be avoided via multiple connections. Additionally, to seek a better trade-off between performance and applicability, a novel module is introduced, referred to as Residual Module (RM), which contains Residual Concatenation Blocks that connected to each other with global skip-connection; build with a set of Adap-

Team	Evaluation 1						Evaluation 2	
	×2		×3		×4		×2 (MR to HR)	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
HPZ-OSU	**26,06**	0,8686	26,11	0,8373	27,32	0,8589	19,98	0,7416
CVC-UAB	26,04	0,8651	25,97	0,8326	27,12	0,8555	19,93	0,7419
MLCV-Lab_SVNIT_NTNU	25,81	**0,8858**	**26,35**	**0,8531**	**27,72**	**0,8758**	20,02	0,7452
LISA-ULB	25,57	0,8401	25,17	0,7583	26,31	0,7824	20,09	0,7385
COUGER AI	25,45	0,8529	25,96	0,8271	27,31	0,8498	**20,36**	**0,7595**
RVL-UTA	24,72	0,8325	25,37	0,8211	26,11	0,8415	19,90	0,7391
Bicubic	24,47	0,8511	25,37	0,8172	26,74	0,8421	20,24	0,7515

Table 2. TISR Challenge: the average results from the evaluations detailed in Section 2.2. The bold and underline values correspond to the first and second best results, respectively.

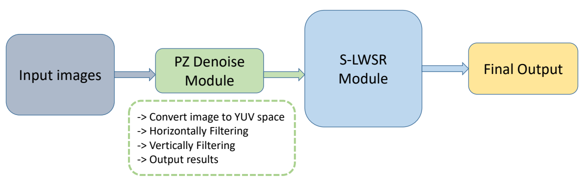

Figure 5. HPZ-OSU proposed architecture with color noise reduction process, where the $s - LWSR$ module uses the framework in [5].

tive Residual Blocks (ARB) with local skip-connection, see Figure 6. Each ARB is defined as a divers residual pathways learning to make use of all kinds of information form LR space. The LMPRNet design has the benefits of a multi-level learning connection and also takes advantage of propagating information throughout the network. As a result, each block has access to information of the preceding block via local and global skip-connections and passes on information that needs to be preserved. By concatenating different blocks followed by 1×1 convolutional layer the network can reach both intermediate and high-frequency information, resulting in better image reconstruction. Finally, a new practical attention mechanism (TFAM) is proposed by focusing on both channel and spatial information. The main objective of TFAM is to enhance the representation power of the model by emphasizing informative features and reduce worthless ones. In contrast to [12] that applies sequentially two modules by two joint sigmoid operations, which is not practical for lightweight models and edge devices. The proposed TFAM applies channel and positional units simultaneously with a different set of operations and a single hard sigmoid function.

In the training stage, input patches with a size of 60×60 from each of the randomly selected 64 training images were used. The number of patches is augmented by random horizontally flips and 90-degree rotation. The Adam optimizer with default setting has been employed. The initial learning rate set to 10^{-3} and its halved every 4×10^5 steps. $L1$ is used as a loss function to optimize the model. The proposed LMPRNet model is implemented in the PyTorch framework. The proposed network has been trained using a 3.2 eight-core processor with 32GB of memory with a NVIDIA GeForce GTX TITAN X GPU;

4.3. MLCV-Lab_SVNIT_NTNU

Figure 7 depicts the framework proposed by MLCV-Lab_SVNIT_NTNU team for thermal image super-resolution. A new ResBlock module, inspired from Inception network [1], is designed (see Figure 8). A channel attention (CA) module [13] is also adopted to adaptive re-scale the channel-wise features by considering inter-dependencies between channels. Furthermore, the local skip connection is utilized in each ResBlock in which the higher layer gradients are bypassed to the lower layer. In addition, long skip connections after six number of ResBlocks, which bypasses the higher layer gradients directly to the first convolution layer, are used. These skip connections help to solve the problem of exploding or vanishing gradient. In the proposed method, the parametric exponential linear unit (i.e., PeLU) activation function [11] is utilized. The feature maps are up-scaled to the desired resolution level by using the sub-pixel convolution layer.

All experiments have been performed on a workstation with the following specifications: Intel Core $i7 - 7700K$ processor, 32 GB RAM with NVIDIA GeForce GTX 1070 8GB GPU. The code is implemented using Tensorflow library. The proposed network is trained using l_1 loss function with a learning rate of 10^{-4} and the same is optimized using Adam optimizer with $\beta = 0.5$. The proposed model was trained up to 50,000 iterations with a batch size of 4. The given training images are augmented using flipping and rotating operations. In the up-sample block, the value of f is set to 63 for $\times 3$ and 64 for $\times 2$ and $\times 4$.

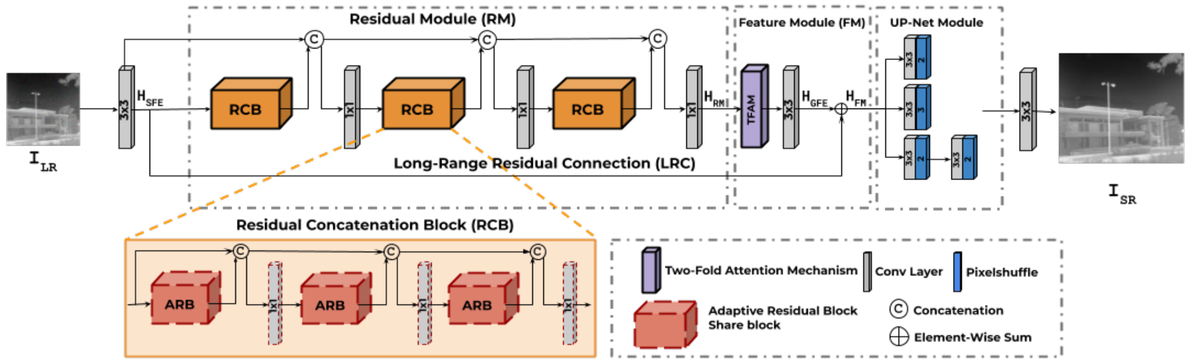

Figure 6. CVC-UAB architecture of the proposed Lightweight Multi-Path Residual Network (LMPRNet).

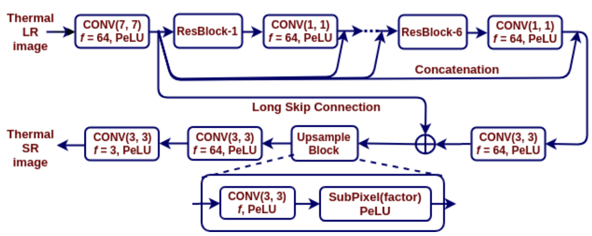

Figure 7. MLCV-Lab_SVNIT_NTNU proposed architecture.

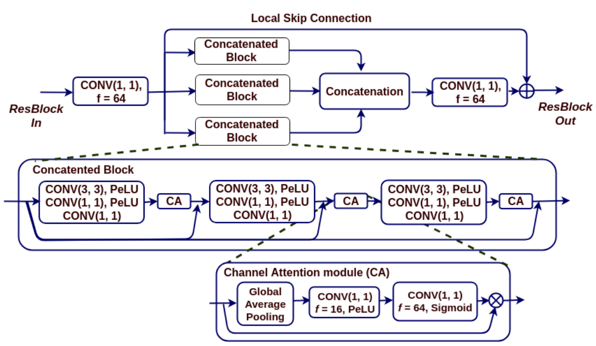

Figure 8. ResBlock design for the MLCV-Lab_SVNIT_NTNU architecture.

4.4. LISA-ULB

The LISA-ULB team introduces a model referred to as VCycles BackProjection (VCBP); it is designed to be scalable to meet the requirements of upscaling the image by ($\times 2$, $\times 3$, $\times 4$) factors while maintaining the performance with a small number of parameters. The main contributions are: (i) an iterative module of shared parameters and Backprojection procedures between cycles; (ii) a new training strategy by constructing the model backwardly.

The model shown in Figure 9 consists of four modules: Encoder (E), Decoder (D), Upsampler and Downsampler. D and E are one convolutional layer map of the image to and from its multidimensional features. The downsampler is a one convolutional layer with stride=2 to downscale the features from the high dimension to the low dimension space. The upsampler is one dense network with 4 layers and one deconvolution layer at the end when the module is last in the sequence. All upsampler modules at the same level (L1, L2, L3, L4) share the same parameters.

The model first upsamples the input image to the target size using bicubic interpolation and uses this image with the low-resolution (LR) image as inputs for the model. The model is responsible to generate the residual high-frequency (HF) information and add them to the encoded features before decoding them back into the image space. In each VCycle the model downscales the accumulated features from the previous and the current cycle to be the input for the next cycle. This procedure enforces the model to generate features in the high-resolution (HR) space while maintaining the similarity in its LR space. All downsampler modules share the same parameters to ensure that all earlier down-sampled features are similar to the encoded features of the LR input image. Only the last down-sampled features are used for the backprojection loss.

$\mathcal{L}^{L1}(x, y) = \mathbb{E}[|x - y|]$ is used for the content loss and the Backprojection loss. The total loss function is:

$$\mathcal{L}^{L1}(SR, HR) + 0.01 * \mathcal{L}^{L1}(BP, a),$$

(BP) and (a) are the encoded features in the LR space of the down-sampled super-resolved features and the input image features respectively.

The VCBP network has been implemented in Pytorch and performed on a NVIDIA TITAN XP. The proposed model was trained using the AdamW optimizer followed by SGDM. Instead of building the model forwardly, a new building and training procedure is proposed, which add models backwardly for each new upscaling factor as shown in Figure 10. The last model in the sequence responsible for generating all the natural super-resolved images, while added earlier stages responsible for producing intermediate super-resolved images. This allows to train the model on larger image sizes. The model has 883K parameters and each upscaling factor module was trained on the three training sets.

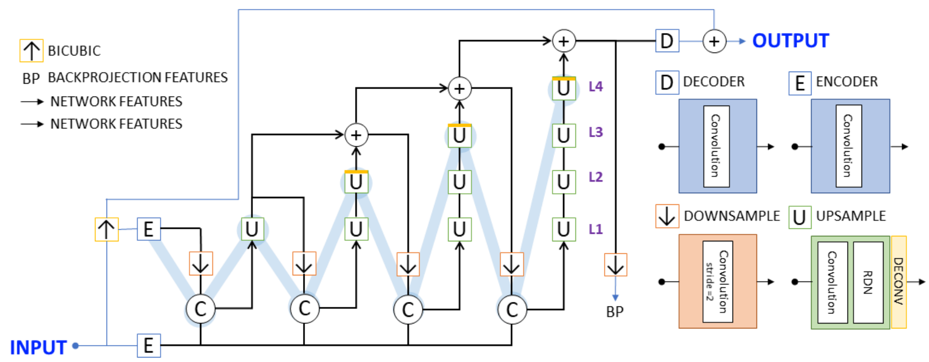

Figure 9. LISA-ULB proposed architecture.

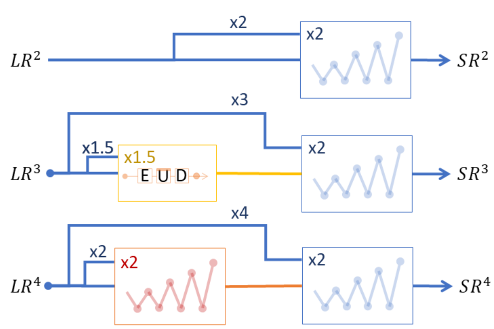

Figure 10. LISA-ULB proposed building model.

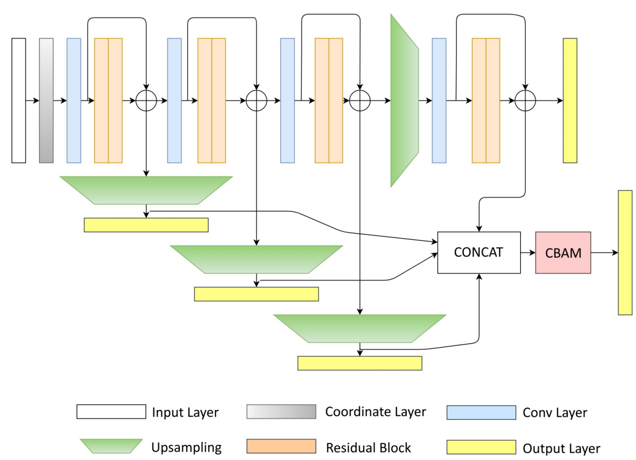

Figure 11. COUGER AI proposed architecture: A Multi-Level Supervision Model

4.5. COUGER AI

The COUGER AI team proposes an architecture for the task of generating $\times 2$, $\times 3$ and $\times 4$ resolution images, which are acquired at three different resolutions. The proposed approach is based on a neural network that utilizes the coordinate convolutional layer [8] and residual units [2], along with the multi-level supervision and attention unit to map the information between LR images to MR and HR images.

As shown in Figure 11, firstly, the bicubic interpolated input image is mapped in Cartesian space using the coordinate convolutional layer [8]. In each base block, two residual units along with the one convolutional layer are used. The output of each base block is up-sampled according to the output resolution requirement. All the up-sampled outputs are then fused to the Convolutional Block Attention Module (CBAM) [12]. Also, to improve the pixel-wise resolution, a multi-level supervision is applied on each up-sampled layer, inspired on [4] [9].

To supervise the model output, a combination of three losses are used: mean squared error (MSE), SSIM, and Sobel, i.e.,

$$Total_{Loss} = MSE + SSIM_{Loss} + SOBEL_{Loss}.$$

In total, three ($\times 2$,$\times 3$,$\times 4$) networks are trained for generating the high-resolution images from the low-resolution input images in Keras 2.2.4. Input images are normalized between 0 to 1 and introduced with Gaussian noise (mean = 0 and sigma = 10). The dataset was trained using a NVIDIA 1080 GTX GPU.

4.6. RVL-UTA

The RVL-UTA team presents a novel network for thermal image super-resolution (SR) called the Multiscale Residual Channel Attention Network (MSRCAN). The architecture is inspired by state-of-the-art methods to recover details from low-resolution (LR) RGB images such as: very deep residual channel attention networks (RCAN) [13], learning a mixture of deep networks for single image SR (MSCN) [7], and multiscale convolutional neural networks (CNNs) (MSSR) [3]. RCAN allows deeper CNN models, which result in more feature representation. MSCN uses multiple parallel inference modules with sequentially increased dilation factors and an adaptive weight (AW) module allowing for multiscale SR outputs and pixel-wise AW summation. MSSR provides a model for parallel CNN paths with different depths corresponding to multiscale SR image outputs. The proposed MSRCAN implements a combination of all these networks to produce higher PSNR and SSIM scores, as well as sharper SR output images.

The inputs to MSRCAN are bicubicly up-sampled LR images. These LR images pass through parallel RCAN SR inference modules, which produce high-resolution (HR) estimates that are aggregated using AW modules at the pixel level. The receptive field of the convolutions within each of the SR modules linearly increases against the number of modules. This is done by increasing the dilation factor by two for each of the parallel modules. Different receptive fields of each SR inference module allow the network to produce an HR estimate for varying scales as was done with MSSR. Each SR inference module is pixel-wise multiplied with its corresponding AW module according to the architecture of MSCN. The sum of these pixel-wise products produces the HR estimate. An overview of the network architecture is shown in Figure 12.

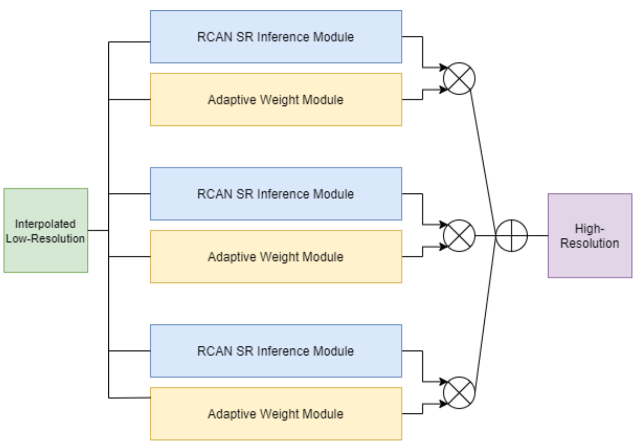

Figure 12. The RVL-UTA proposed MSRCAN architecture.

MSRCAN was trained on a workstation with a NVIDIA Quadro P4000 GPU, Intel Core i7-8700 CPU, and 32GB of RAM. It was written with the Python 3.7.6 programming language and the Tensorflow v1 library. In addition, the following modules were used: PIL, Pyelastix, OpenCV, ImageIO, TQDM.

5. Conclusions

This paper summarizes the best contributions to the Thermal Image Super-Resolution Challenge - PBVS 2020, where 51 teams from 17 different countries have participated and 6 teams reached the final validation phase. This was the first time this challenge has been proposed and a wide interest from the research community has been observed. Undoubtedly, the results from this year will be used as the benchmark for next year's challenge. This challenge has also been an opportunity to promote the evaluation dataset used by the participating teams.

Acknowledgements

This work has been partially supported by the ESPOL project PRAIM (FIEC-09-2015); the Spanish Government under Project TIN2017-89723-P; and the "CERCA Programme / Generalitat de Catalunya". The first author has been supported by government of Ecuador under a SENESCYT scholarship contract.

A. Teams and Affiliations

TISR Challenge 2020 Team

Title: Thermal Image Super-Resolution Challenge - PBVS 2020

Members: Rafael E. Rivadeneira[1] (rrivaden@espol.edu.ec), Angel D. Sappa[1,2] and Boris X. Vintimilla[1]

Affiliations:
[1]Escuela Superior Politécnica del Litoral, ESPOL, Facultad de Ingeniería en Electricidad y Computación, CIDIS, Campus Gustavo Galindo Km. 30.5 Vía Perimetral, P.O. Box 09-01-5863, Guayaquil, Ecuador
[2]Computer Vision Center, Campus UAB, 08193 Bellaterra, Barcelona, Spain

A.1. HPZ-OSU

Members: Lin Guo (guli@ostatemail.okstate.edu) and Jiankun Hou

Affiliation:
Oklahoma State University, Stillwater, OK, USA

A.2. CVC-UAB

Members: Armin Mehri (amehri@cvc.uab.es) and Parichehr Behjati Ardakani

Affiliation:
Computer Vision Center, Campus UAB, 08193 Bellaterra, Barcelona, Spain

A.3. MLCV-Lab_SVNIT_NTNU

Members: Heena Patel[1] (hpatel1323@gmail.com), Vishal Chudasama[1], Kalpesh Prajapati[1], Kishor P. Upla[1,2], Raghavendra Ramachandra[2], Kiran Raja[2] and Christoph Busch[2]

Affiliations:
[1]SVNIT, Surat, India
[2]NTNU, Gjøvik, Norway

A.4. LISA-ULB

Members: Feras Almasri (falmasri@ulb.ac.be) and Olivier Debeir
Affiliation:
Universite Libre de Bruxelles, Belgium

A.5. COUGER AI

Members: Sabari Nathan (sabari@couger.co.jp) and Priya Kansal
Affiliation:
Couger Inc, Japan

A.6. RVL-UTA

Members: Nolan Gutierrez (nolan.gutierrez@mavs.uta.edu), Bardia Mojra and William J. Beksi
Affiliation:
Robotic Vision Laboratory
Department of Computer Science and Engineering
University of Texas at Arlington, Arlington, TX, USA

References

[1] François Chollet. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1251–1258, 2017. 4

[2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016. 6

[3] Xiaoyi Jia, Xiangmin Xu, Bolun Cai, and Kailing Guo. Single image super-resolution using multi-scale convolutional neural network. In *Pacific Rim Conference on Multimedia*, pages 149–157. Springer, 2017. 6

[4] Priya Kansal and Sabari Nathan. Eyenet: Attention based convolutional encoder-decoder network for eye region segmentation, 2019. 6

[5] Biao Li, Jiabin Liu, Bo Wang, Zhiquan Qi, and Yong Shi. s-lwsr: Super lightweight super-resolution network. *arXiv preprint arXiv:1909.10774*, 2019. 2, 4

[6] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRw)*, July 2017. 2

[7] Ding Liu, Zhaowen Wang, Nasser Nasrabadi, and Thomas Huang. Learning a mixture of deep networks for single image super-resolution. In *Asian Conference on Computer Vision*, pages 145–156. Springer, 2016. 6

[8] Rosanne Liu, Joel Lehman, Piero Molino, Felipe Petroski Such, Eric Frank, Alex Sergeev, and Jason Yosinski. An intriguing failing of convolutional neural networks and the coordconv solution, 2018. 6

[9] Sabari Nathan and Priya Kansal. Skeletonnet: Shape pixel to skeleton pixel. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRw)*, 2019. 6

[10] Rafael E. Rivadeneira, Angel D. Sappa, and Boris X. Vintimilla. Thermal image super-resolution: a novel architecture and dataset. In *Proceedings of the 15th International Conference on Computer Vision Theory and Application (VISAPP), Valletta, Malta*, pages 111–119, February 2020. 2, 3

[11] Ludovic Trottier, Philippe Gigu, Brahim Chaib-draa, et al. Parametric exponential linear unit for deep convolutional neural networks. In *16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 207–214. IEEE, 2017. 4

[12] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. Cbam: Convolutional block attention module. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 3–19, 2018. 4, 6

[13] Yulun Zhang, Kunpeng Li, Kai Li, Lichen Wang, Bineng Zhong, and Yun Fu. Image super-resolution using very deep residual channel attention networks. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 286–301, 2018. 2, 4, 6