A new deep-learning neural network for super-resolution up-scaling of thermal images

by P. Więcek, D. Sankowski

Institute of Applied Computer Science, Lodz University of Technology, *pwiecek@gmail.com*

Abstract

This paper presents a novel architecture of deep learning convolutional neural network for up-sampling of thermal images. The proposed solution is based on Kernel-Sharing Atrous Convolution (KSAC) filtering block. The developed system ensures the high accuracy of the up-sampling with scales up to 6 with much lower algorithm complexity compared to the reference methods widely used for visual image processing. The learning process uses the high resolution RGB visual images available in the DIV2K database. The examples of up-sampling of thermal images generated by IR cameras with160x120 and 32x24 sensors are presented.

1. Introduction

Deep learning networks have recently become very useful in different applications, in particular in image processing. One of possible applications is up-scaling already successfully implemented for visual images. Thermal images can also be processed using such deep-learning networks to increase the pixel resolution.

Interpolation and up-scaling of images, more than other algorithms, faces with a problem of vanishing gradients during learning [1,2]. In the critical conditions, learning process can get stuck. It is because the input and output data are typically very similar to each other in contrast e.g. to the clustering or classification of images. One of the possible solutions of this problem is the application of residual networks that are using short-cut lines (skip connections) allowing the original data passing over many feature blocks as shown in Figure 1a,b. In this way, the residuals blocks are learning with the significantly different input-output data. The typical residual network consists of many residual blocks (R) with additional local skip connections, as shown in Figure 1a.

a)   b)
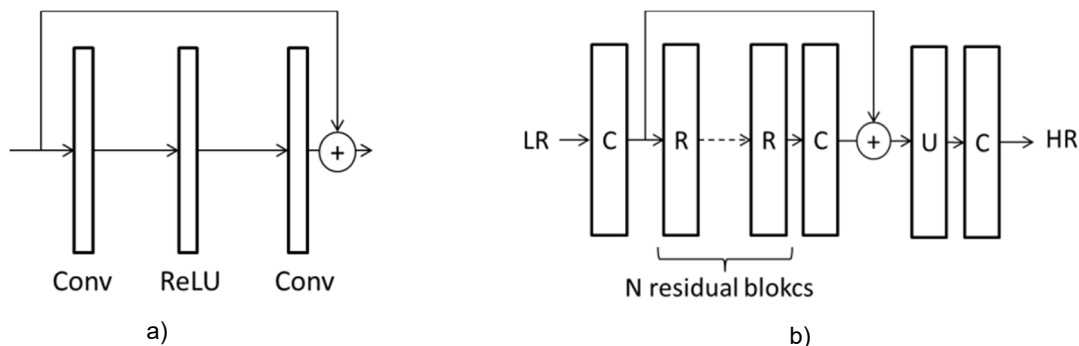
Fig. 1. *The concept of residual deep-learning network, a) residual block, b) residual network, C-convolutional, R-residual, U-up-scaling blocks*

There are already a few existing architectures available for up-scaling of visual images [3-11]. Some of them have been already tested for thermal images as well [14]. Another problem in processing of thermal images is that there are no large databases with thermal images available for learning. In the case of lack of training data, the so-called transfer learning is only possible. In this research, the learning and validation of the proposed neural networks was made using the DIV2K database [13]. It consists of high-resolution and corresponding low-resolution versions of RGB colour visual images.

The important problem of deep learning systems is their computational complexity and in consequence the relatively long execution time of implemented algorithms. Recently, the efforts have been undertaken to diminish the algorithm complexity with simultaneous retaining the effectiveness and accuracy of image processing. This paper presents the results of super-resolution up-scaling using a new deep-learning convolutional neural network offering the reasonable compromise between the accuracy and complexity of image processing.

2. Architecture of a new deep-learning neural network for thermal image up-sampling

The novel algorithm of deep-learning for super-resolution up-scaling proposed in this paper uses the Kernel-Sharing Atrous Convolution (KSAC) block [15]. The concept of KSAC is based on Atrous Spatial Pyramid Polling (ASPP). Originally, this idea of feature filtering has been successfully used for image segmentation [15]. Term "Atrous" comes from

French and stands for the algorithm with holes (gaps) in data sets [15]. In other applications, the Atrous convolution is called dilated convolution. The concept of the feature filtering using KSAC is presented in Figure 2. It operates as parallel 2D convolutional filtering with different input data for each filter using 2D sub-sampling as shown in Figure 2. For example, if the rate of sub-sampling is equal to 6, every sixth pixel in each direction is taken for processing. After parallel filtering, the output data is concatenated forming the next layer of the feature map.

In the concept of the presented research, the KSAC block is inserted into the Residual Convolutional Network which uses the global skip connections with reduced number of maps' layers as it is shown in Figure 3. As a result, the new network is named KSAC-ResNet. Typically, the residual networks consist of many residual blocks (R) with local skip connections. In the network presented in this paper, a few residual blocks incorporate with classical convolutional blocks (C) and KSAC module as shown in Figure 3. The low-resolution image (LR) enters the KSAC-ResNet neural network and is processed first by the input convolutional block. In addition, for super-resolution applications, the up-scaling block (U) is necessary to generate high-resolution images (HR). Similarly to the input data, the output HR image is generated by the convolutional deep-learning block.
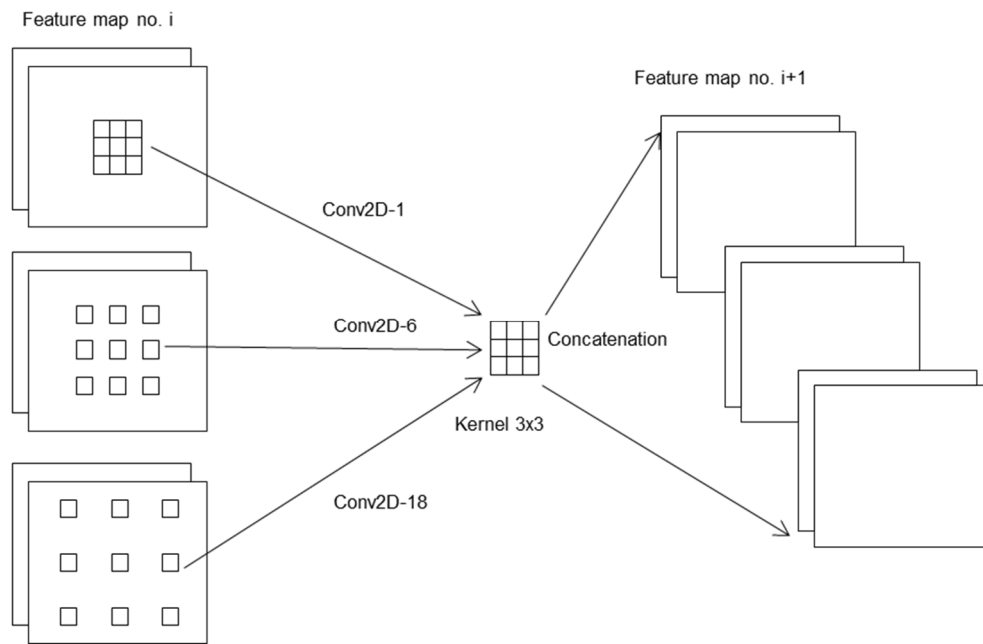
**Fig. 2.** *The concept of Kernel-Sharing Atrous Convolution (KSAC) feature filtering*
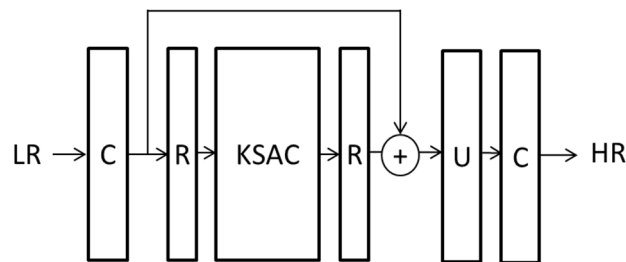
**Fig. 3.** *The proposed architecture of the Residual Network with KSAC block inserted*

The detailed structure of the new network for up-scaling of thermal images is presented in Figure 4. The network was implemented in PyTorch framework using Python environment [12]. The network was implemented on a computer with 4x Xeon 2.2 GHz CPU with 20 GB memory. The computer was equipped with the Tesla 4 16 GB GPU graphical card.

3. Learning process

Learning process was performed using the DIV2K database of visual images of different resolutions [13]. This database consists of the images of high and low resolutions. Typical resolution of an image from the DIV2K database is 1228x1125 RGB pixels. Learning is implemented with sub-images containing 192x192 pixels only. Each epoch was using 16000 such images for processing. The holdout validation was applied using full-size images. The set of 800 full-size images was used for learning and 10 images randomly selected were taken for validation. The process of learning was monitored using PSNR and L1loss measures. The examples of PSNR and L1loss curves during learning for 650 epochs are presented in Figure 5.
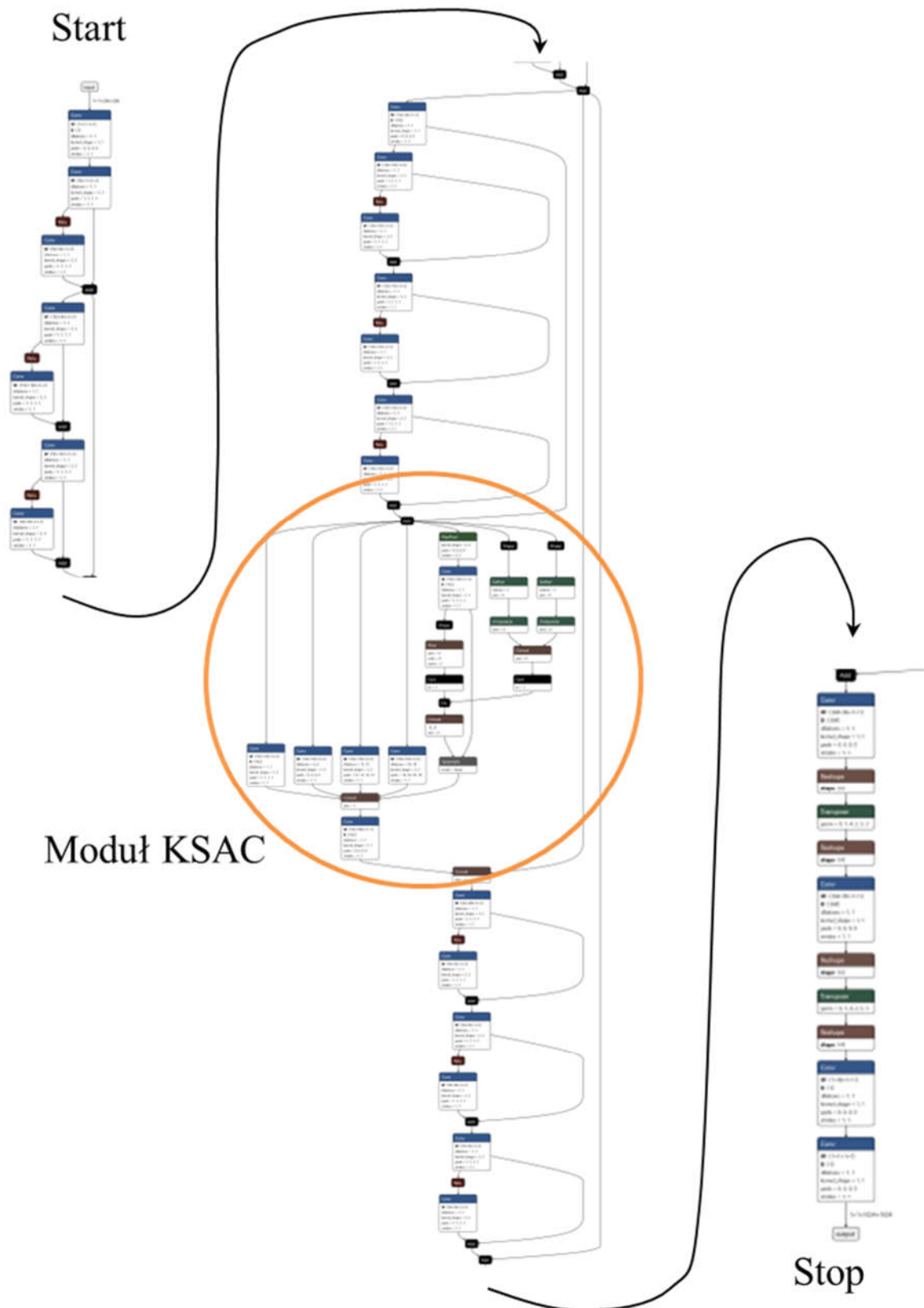
Fig. 4. *The structure of the propsed KSAC-ResNet neural network for super-resolution up-sampling of thermal images implemented in PyTorch environment*
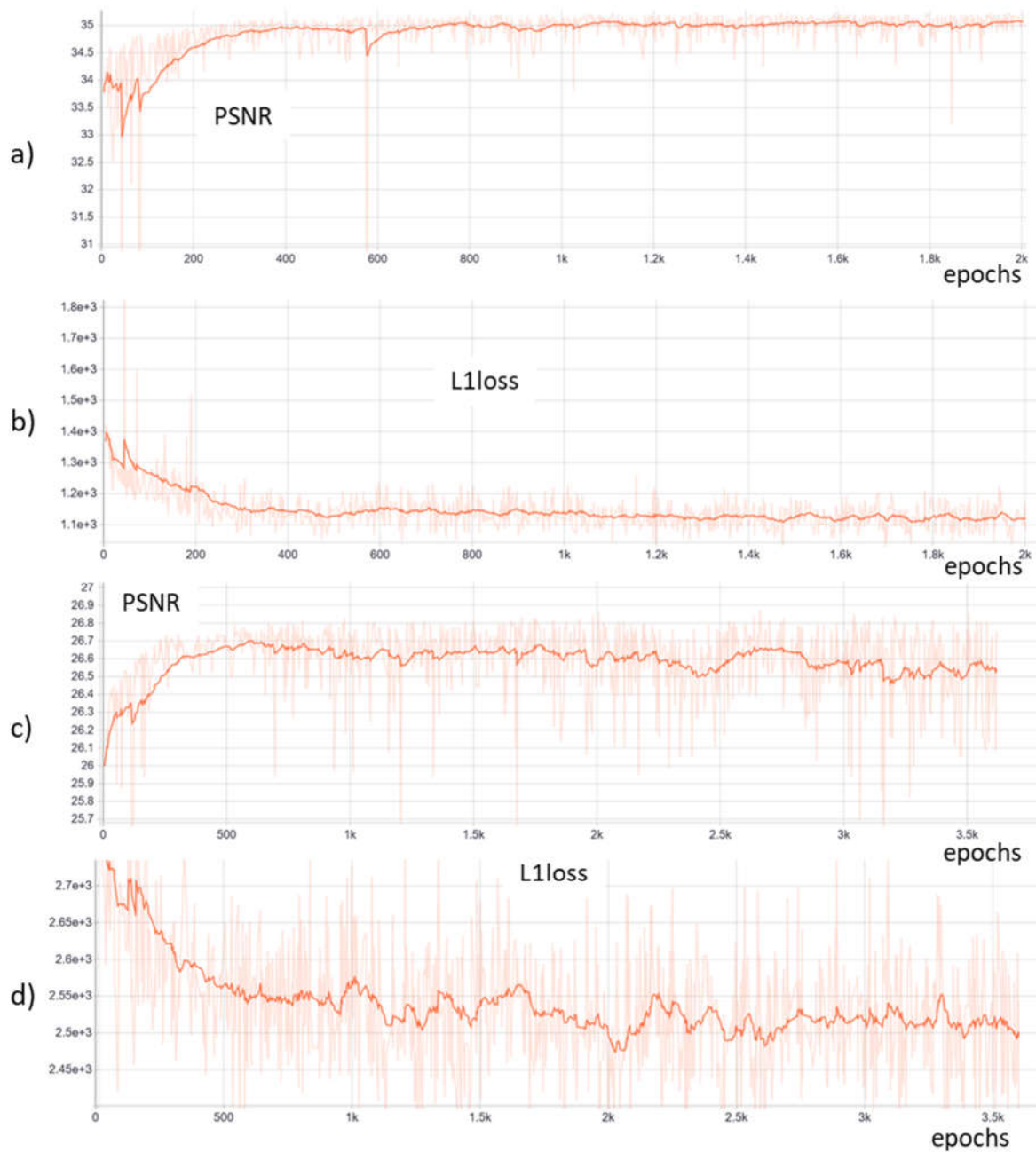
***Fig. 5.** Examples PSNR and L1Loss curves during learning for KSAC-ResNet using monochromatic high resolution images, a), b) for scale s = 2, and c), d) for scale s = 6*

4. Results

The proposed deep learning system was developed for monitoring hardening of metal samples by inductive heating. It is the industrial, high-temperature thermal process. The intention was to develop low-cost monitoring system with low-resolution IR camera and up-scaling algorithms implemented to improve the resolution. Different IR cameras were used in the research with the sensor matrix of 640x480, 160x120 and 32x24 pixels. The setup with 32x24 miniature IR LWIR sensor is presented in Figure 6.
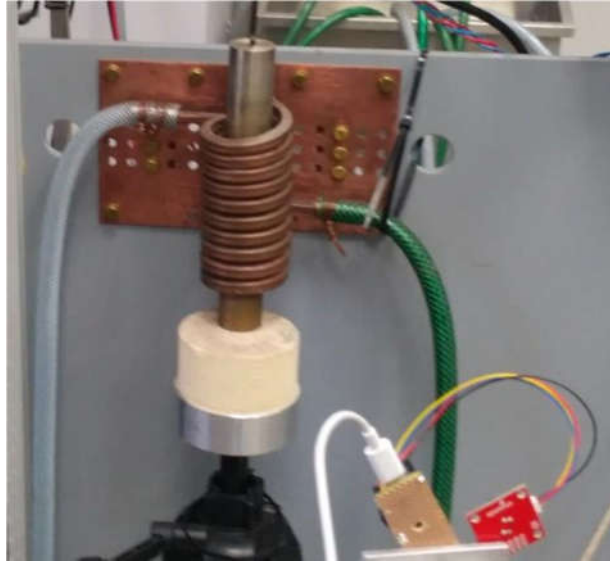
Fig. 6. *The coil of power driver for industrial inductive heating*

The first example presents a bras sample heated by inductive Eddy current system. The original IR image was of 160x120 pixel size. The results of up-sampling with scale s=6 up to 960x720 resolution by different methods are presented in Figure 7. The bicubic interpolation (Figure 7a) definitely shows the worst result. The up-sampling results achieved by EDSR32 [4] (Figure 7c) and KSAC-ResNet (Figure 7d) deep learning solutions are comparable, but KSAC-ResNet requires much lower computational power due the lower algorithm complexity.
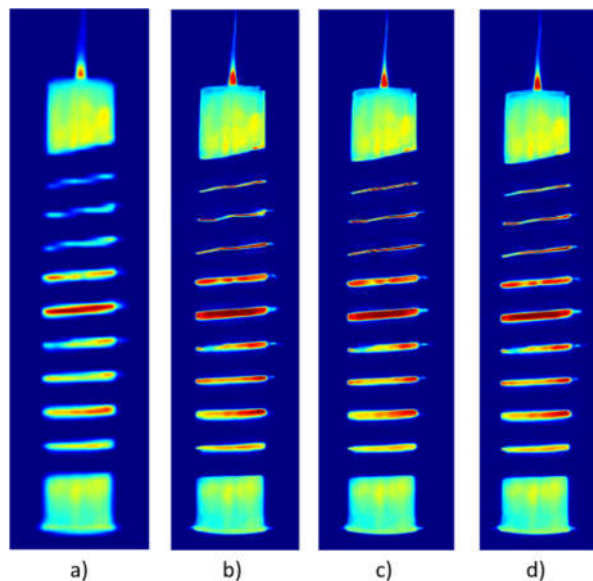
Fig. 7. *Up-sampled thermal images (960x720 pixel resolution) of the coil of power driver and the bras sample during hardening performed by industrial inductive heating, original image resolution 160x120, scale s =6, a) bicubic interpolation, b) EDSR16 [4], c) EDSR32 [4], d) KSAC-ResNet*

The developed system was also applied for low-cost low-resolution IR camera with 32x24 sensors. The qualitative results are presented in Figure 8. The deep learning increasing of image pixel resolution (Figure 8b,d) is better compared to classical, widely used in practice bicubic interpolation (Figure 8a,c).
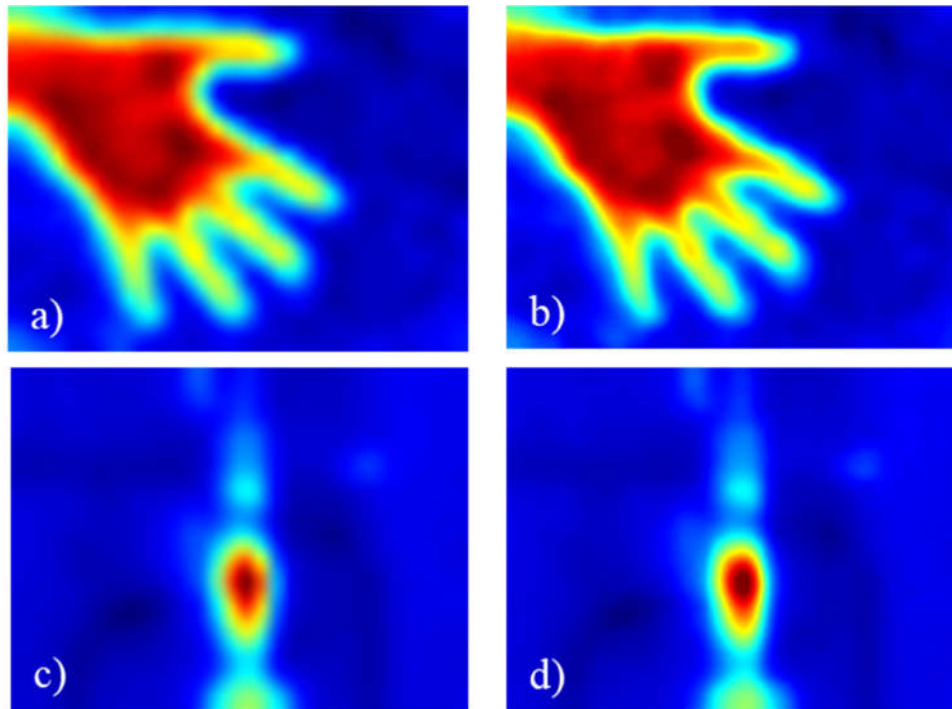
Fig. 8. *Up scaling of images with 32x24 IR sensors with scale s=6, a), c) – bicubic interpolation, b), d) – deep learning algorithm*

The algorithm complexity can be quantitatively determined by counting the number of Multiply–Accumulate (MAC) operations. MAC measure is very suitable for convolutional networks as it sums many 2-operand products. Table 1 presents clearly the advantage of KSAC-ResNet over the reference algorithm of the super-resolution EDSR32 [4]. Both deep learning neural networks were implemented in PyTorch framework and then run on the same computer for complexity comparison. It was a computer with 4x Xeon 2.2 GHz CPU and 20 GB memory. The computer was equipped with the Tesla 4 16 GB GPU graphical card. KSAC-ResNet and EDSR32 deep learning networks can perform up-sampling tasks for thermal images with the very similar accuracy.

Table 1. Comparison of KSAC-ResNet and EDSR32 complexity indexes for images with original pixel resolution 160x120 and 32x24 for scale s=6

Deep learning algorithm	Algorithm complexity (GMac)	No. of parameters (10^6)	Mean time of 1000 runs (s)
160x120			
KSAC – ResNet x6	212.29	7.75	28.814
EDSR32 x6	1146.08	59.6	153.896
32x24			
KSAC – ResNet x6	8.5	7.75	2.048
EDSR32 x6	45.84	59.6	8.903

5.   Conclusions

The deep learning convolutional neural network was successfully used in up-scaling tasks of low-resolution thermal images. The new architecture of the deep learning network with Kernel-Sharing Atrous Convolution (KSAC) block based on the Atrous Spatial Pyramid Polling ensures the acceptable accuracy of resolution increase with relatively lower algorithm complexity compared to the reference method based on deep learning residual networks.

REFERENCES

[1] Y. Bengio, P. Simard, and P. Frasconi. Learning long-term dependencies with gradient descent is difficult. IEEE Transactions on Neural Networks, 5(2):157–166, 1994.

[2] X. Glorot and Y. Bengio. Understanding the difficulty of training deep feedforward neural networks. In AISTATS, 2010.

[3] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In ICML, 2015.

[4] B. Lim Sanghyun, S. Heewon Kim, S, Nah K. Mu Lee, Enhanced Deep Residual Networks for Single Image Super-Resolution, The IEEE Conference on Computer Vision and Pattern Recognition Workshops, July 21- 26, Honolulu, 2017.

[5] J. Li, F. Fang, K. Mei, G. Zhang, Multi-scale Residual Network for Image Super-Resolution, 15th European Conference on Computer Vision, Munich, September 8 – 14, 2018.

[6] J-H. Kim, J-S Lee, Deep Residual Network with Enhanced Upscaling Module for Super-Resolution, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2018, Salt-Lake City, June 18-22, 2018, pp. 800-808.

[7] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, Y. Fu, Image Super-Resolution Using Very Deep Residual Channel Attention Networks, 15th European Conference on Computer Vision, Munich, September 8 – 14, 2018.

[8] J. Kim, J, J. K. Lee, K. M. Lee. Accurate Image Super-Resolution Using Very Deep Convolutional Networks, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016, pp. 1646-1654.

[9] X. Zhang, C. Li, Q. Meng,S. Liu,Y. Zhang, J. Wang, Infrared Image Super Resolution by Combining Compressive Sensing and Deep Learning, Sensors (Basel). 2018 Aug; 18(8): 2587.

[10] Z. Lu, Y. Chen, Single Image Super Resolution based on a Modified, arXiv:1911.09428 [eess.IV], 21 Nov. 2019.

[11] X. Hu, M. A. Naiel, A. Wong, M. Lamm, P. Fieguth, RUNet: A Robust UNet Architecture for Image Super-Resolution, Computer Vision and Pattern Recognition Workshop, Long Beach, California, June 16th - June 20th, 2019.

[12] https://pytorch.org

[13] https://data.vision.ee.ethz.ch/cvl/DIV2K/

[14] P. Więcek, D. Sankowski, Low-cost, low-resolution IR system with super-resolution interpolation of thermal images for industrial applications, Measurement Automation Monitoring, MAM 2018, no. 04, pp. 108-111

[15] Y. Huang, Q. Wenjing, W. Jia, L. Yue, X. He, See More Than Once -- Kernel-Sharing Atrous Convolution for Semantic Segmentation, arXiv:1908.09443 [cs.CV] 16 Nov. 2019.