Super-Resolution of Synthetic Aperture Radar Complex Data by Deep-Learning

1th Pia Addabbo
Departimento di Ingegneria
Università degli studi del Sannio
Benevento, Italy
paddabbo@unisannio.it

2st Mario Luca Bernardi
Dipartimento di Ingegneria
Università degli studi del Sannio
Benevento, Italy
bernardi@unisannio.it

3nd Filippo Biondi
Dipartimento di Ingegneria
Università degli studi dell'Aquila
L'Aquila, Italy
biopippop@gmail.com

4th Marta Cimitile
Department of Law and Economics
Univeristà Unitelma Sapienza
Roma, Italy
marta.cimitile@unitelmasapienza.it

5rd Carmine Clemente
Department of
Electronic and Electrical Engineering
University of Strathclyde
Glasgow, UK
carmine.clemente@strath.ac.uk

6th Nicomino Fiscante
Dipartimento di Ingegneria
Università degli studi Roma TRE
Rome, Italy
n.fiscante@uniromatre.it

7th Gaetano Giunta
Dipartimento di Ingegneria
Università degli studi Roma TRE
Rome, Italy
gaetano.giunta@uniromatre.it

8th Danilo Orlando
Università degli Studi
"Niccolò Cusano"
Rome, Italy
danilo.orlando@unicusano.it

***Abstract*—One of the greatest limitations of Synthetic Aperture Radar imagery is the capability to obtain an arbitrarily high spatial resolution. Indeed, despite optical sensors, this capability is not just limited by the sensor technology. Instead, improving the SAR spatial resolution requires large transmitted bandwidth and relatively long synthetic apertures that for regulatory and practical reasons are impossible to be met. This issue gets particularly relevant when dealing with Stripmap mode acquisitions and with relatively low carrier frequency sensors (where relatively large bandwidth signals are more difficult to be transmitted). To overcome this limitation, in this paper a deep learning based framework is proposed to enhance the SAR image spatial resolution while retaining the complex image accuracy. Results on simuated and real SAR data demonstrate the effectiveness of the proposed framework.**

***Index Terms*—Expectation Maximization, Polarimetric Radar, Radar, Synthetic Aperture Radar**

I. Introduction

Synthetic Aperture Radar (SAR) imagery has become an important Earth observation technique for obtaining physical and shape information about targets. As coherent radar technology improves, there is an increasing demand for image quality with the ability to discriminate more and more detail. Image resolution is a key factor in discriminating two separately located targets at ever smaller distances. In order to meet these demands, it is necessary to process conspicuous amounts of both electromagnetic (EM) and Doppler bands, which may not always be feasible. In order to solve this problem, we have developed a super-resolution (SR) system based on Deep-Learning (DL).

Author of [1] proposes a recovery solution for SAR Single Look Complex (SLC) images that are corrupted by noncoherent EM noise covering only the higher frequency spectrum. The solution consists of, first, exporting the spectrum damages that occur in the native data and, second, focusing only the survived spectrum information at lower resolution. The recovery of the original image is done by SR signal processing based on spectrum extrapolation and implemented by convex programming. In [2] a new method of single image SR based on DL features and dictionary model is proposed. The experimental results indicate that the proposed algorithm can produce good SR visual results with respect to state-of-the-art algorithms.

In [3] authors propose a deep network architecture for a SR-aided hyperspectral image classification with classwise loss (SRCL). First, a three-layer SR convolutional neural network (SRCNN) is employed to reconstruct a high-resolution image from a low-resolution image. Second, an unsupervised triplet-pipeline CNN (TCNN) with an improved classwise loss is built to encourage intraclass similarity and interclass dissimilarity. Finally, SRCNN, TCNN, and a classification module are integrated to define the SRCL, which can be fine-tuned in an end-to-end manner with a small amount of training data. Experimental results on real hyperspectral images demonstrate that the proposed SRCL approach outperforms other state-of-the-art classification methods, especially for the task in which

only a small amount of training data are available. In [4] a new approach that combines the advantages of multiple-image fusion with learning the low-to-high resolution mapping using deep networks, is defined. In [5], authors proposes a full-polarimetric SAR image SR reconstruction method based upon deep convolutional neural network for nonlinear model fitting and, then, apply residual compensation to network reconstruction results using low-resolution image information. So far, DL has been successfully applied to single image SR, which aims at reconstructing a high-resolution (HR) image from its low-resolution (LR) counterpart. In [6], authors use a scheme based in the frequency domain to reconstruct the HR image at various frequency bands. Further, authors propose a method that incorporates the wavelet transform (WT) and the recursive Res-Net. The WT is applied to the LR image to divide it into various frequency components. To validate the effectiveness of the proposed method, extensive experiments are performed using the NWPU-RESISC45 data set, and the results demonstrate that the proposed method outperforms several state-of-the-art methods in terms of both objective evaluation and subjective perspective. Recently in [7], a novel CNN-based technique that exploits both spatial and temporal correlations to combine multiple images, is devised. This novel framework integrates the spatial registration task directly inside the CNN, relying on a single CNN with three main stages. Finally, another tool that comes in handy to obtain SR images is the generative adversarial network (GAN) as corroborated by [8].

II. The Proposed Methodology

This study proposes the two channel Deep CNN with Residual Net, Skip Connection and Network in Network (DC2SCN) as an extension of the DCSCN model described in [9], [10]. The overall network consists of a combination of a feature extraction and reconstruction tasks. The feature extraction part of the network learns, from the complex SAR image, a wide range of features used by the subsequent reconstruction part to rebuild the input SAR image at an higher resolution. The proposed DC2SCN variant exploits two main channels, instead of a single one, for both feature extraction and reconstruction of each SAR image component: one for real part and the other for imaginary part of the signal (recall that processed SAR data are complex).

The feature extraction part is comprised, for each channel, of a cascade of N_F sets of 3×3 CNN, bias and Parametric Rectified Linear Units (ReLU). Skip connections are used to send extracted features at each level to the reconstruction subsequent part of the network.

The reconstruction part of the network is comprised for each channel of a 3×3 CNN layer in parallel with two 3×3 CNN blocks (each one followed by a bias and a parametric ReLu units) that are shared through the real and imaginary channels as shown at the center of the figure. This shared portion of the network allows for: (*i*) it improves the computation performance by reducing the dimensions of the previous layers with a low information loss, and (*ii*) it allows the entire network to learn relationships across real and imaginary parts of the signal to preserve phase information during the high resolution image reconstruction.

For each channel an estimation of the up-sampled original image is obtained including these outputs to the image constructed using bicubic interpolation. Notice that, the input layer of the reconstruction network is characterized by large dimensions since all the features are included. For this reason a parallelized 1×1 CNN [11] is used to reduce the dimension before generating the final images ($\mathrm{X}^{'}$ and $\mathrm{Y}^{'}$). It also enhances the final representation by including additional nonlinearity.

III. Experimental results

The data set used in the experiments are SAR SLC images sensed by COSMO-SkyMed (CSK) in the StripMap (SM) acquisition mode [12]. The SM mode implements aproximately a spatial resolution of 3×3 m in ground coordinates. The set is made by 10 scenes of different size acquired over nothern Italy. Since each whole scene consists in wide geographical areas containing milions of pixels, the images are divided into blocks of size 512×512 pixels, hereafter referred to as tiles. Precisely, the whole dataset is divided into a total of 20239 tiles. Figure 1 shows the entire first scene that covers the river Po valley in northen Italy. This scene covers an area of about 50×65 km and is made by 1344 tiles. We trained the network employing degraded resolution SAR images. Spatial resolution degradation was done in both the range and azimuth directions. In this context, we focused the images by employing adapted filters that consider half-band, both range and azimuth. The same images that we focused at half spatial resolution were also focused at maximum resolution, then using adapted filters set to consider all available chirp and Doppler bandwidth. The maximum resolution images were used as ground truth during the training phase of the neural network.

Precisely, we selected 16191 tiles for training and validation and 4048 tiles for test. This experimental setup allowed to an extensive training of the network as well as a computational efficiency.

The 25 tiles highlighted in the boxes in Figure 1 are here considered for assessing the performance of the proposed architecture. To quantitatively evaluate the performance in terms of the quality of the reconstructed super-resoluted image, hereafter indicated as $\mathbf{I_r}$, with respect to the original full-resoluted one, namely $\mathbf{I_f}$, we use the following metrics:

- the Mean Absolute Error (MAE) expressed as

$$\mathrm{MAE} = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} [|\mathbf{I_f}(m,n)| - |\mathbf{I_r}(m,n)|];$$

- the Root Mean Square Error (RMSE) expressed as

$$\mathrm{RMSE} = \sqrt{\frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} [|\mathbf{I_f}(m,n)| - |\mathbf{I_r}(m,n)|]^2};$$

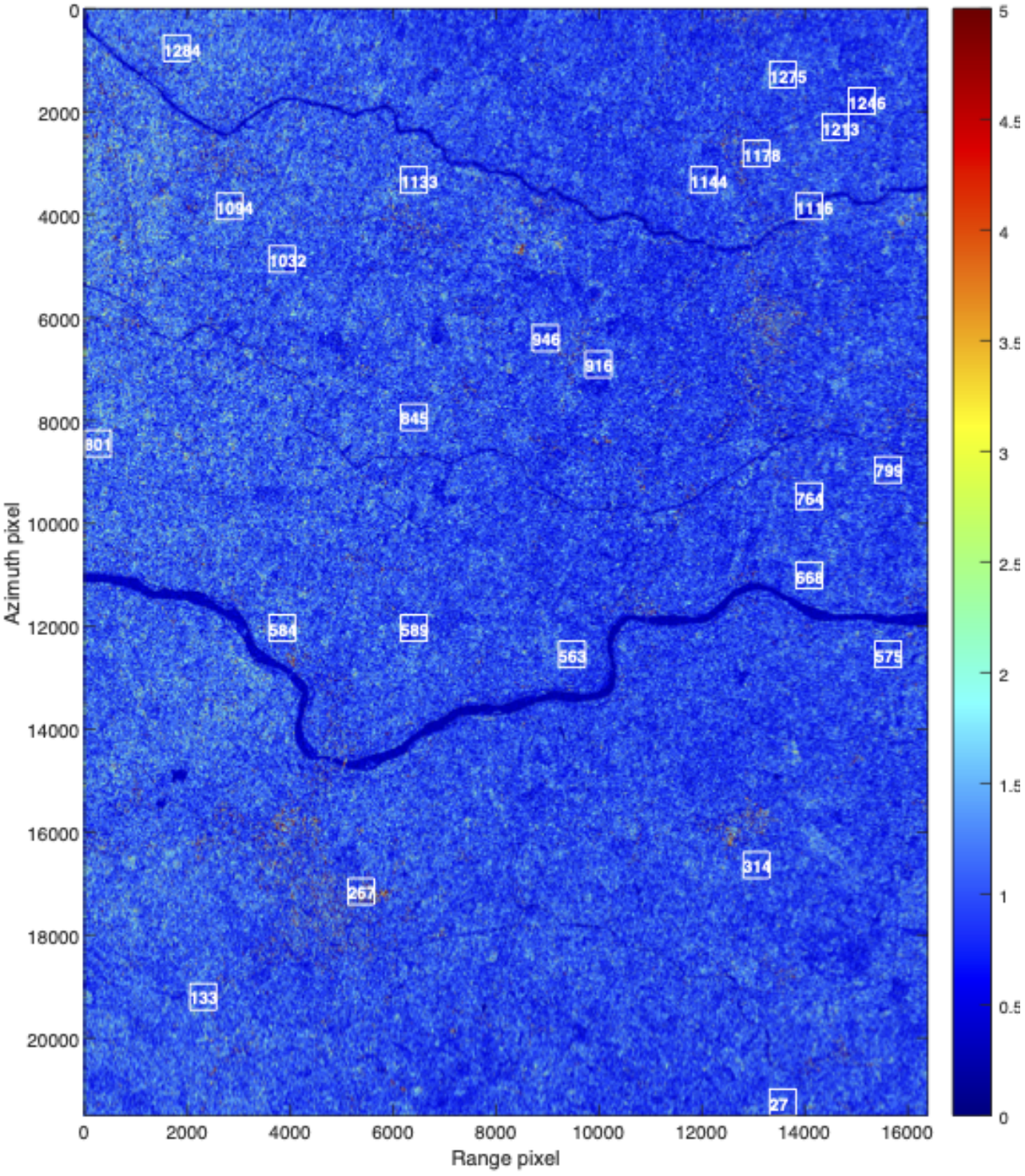

Figure 1: The intensity image is represented with a zoomed color scale from 0 to 5 for image enhancement.

- the Peak Signal-to-Noise Ratio (PSNR), expressed in d defined as the ratio of the maximum pixel intensity to t power of the distorsion,

$$\text{PSNR} = 10\log_{10}\frac{\max(|\mathbf{I_r}|^2)}{\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}[|\mathbf{I_f}(m,n)| - |\mathbf{I_r}(m,n)}$$

- the SSIM, a widely used perceptual image quality metr that can be expressed as

$$\text{SSIM} = l \cdot c \cdot s,$$

with l, c, s the luminance, the contrast and the structu changes as defined in [13].

Particularly, the PSNR, the RMSE, the MAE and the SSI are evaluated both two network configuration of respective 20 and 40 layers (respectively called L20 and L40) which a compared with the corresponding degraded images used input. This is done to evaluate the gain obtained through the reconstructed resolution of the two networks.

The experimental results show an increase in the values of PSNR and SSMI for all the tiles and for both L20 and L40 configurations. The RMS of the PSNR gains of 13.97 dB and 12.54 dB with respect to degraded image are obtained by L40 and L20, respectively. As for the SSIM, the RMS value for the degraded image is 0.17 whereas 0.60 and 0.67 are obtained with L40 and L20, respectively. Both the RMSE and MAE metrics also confirm the superiority of L40 with respect to L20 with slightly lower values.

In Figure 2 (top), the tile 267 of the full scene in Figure 1 is shown representing a mixed urban and vegetated area. Moreover, a zoom plot on a bright target of the degraded image is shown at the top in Figure 2 (bottom). In Figure 3, the improvement with respect to the degraded target of Figure 2 (bottom) is shown in term in the amplitude and phase domains. In this case, the PSLR values are 12.9809, 5.9038, 13.0155 and 13.0461 dB for the original full-resoluted image, the degraded image and the reconstructed images with L20 and L40, respectively. In terms of 3 dB main lobe amplitude, 40 interpolated azimuth pixels (corresponding to 4 original azimuth pixels) for the degraded image are found whereas 20 pixels are obtained for both L20 and L40 which coincide with the full resolution of the original image. At the bottom of Figure 3, the phase values are reported. Interestingly, it can be noticed that the reconstructed phases show a very good agreement with respect to the original one. On the contrary, the degraded phase exhibits a phase discrepancy of π in the main lobe position (azimuth pixel 234).

These results confirm that the proposed architecture is able to superesolve as well as to reconstruct both the module and the phase of the SAR complex data.

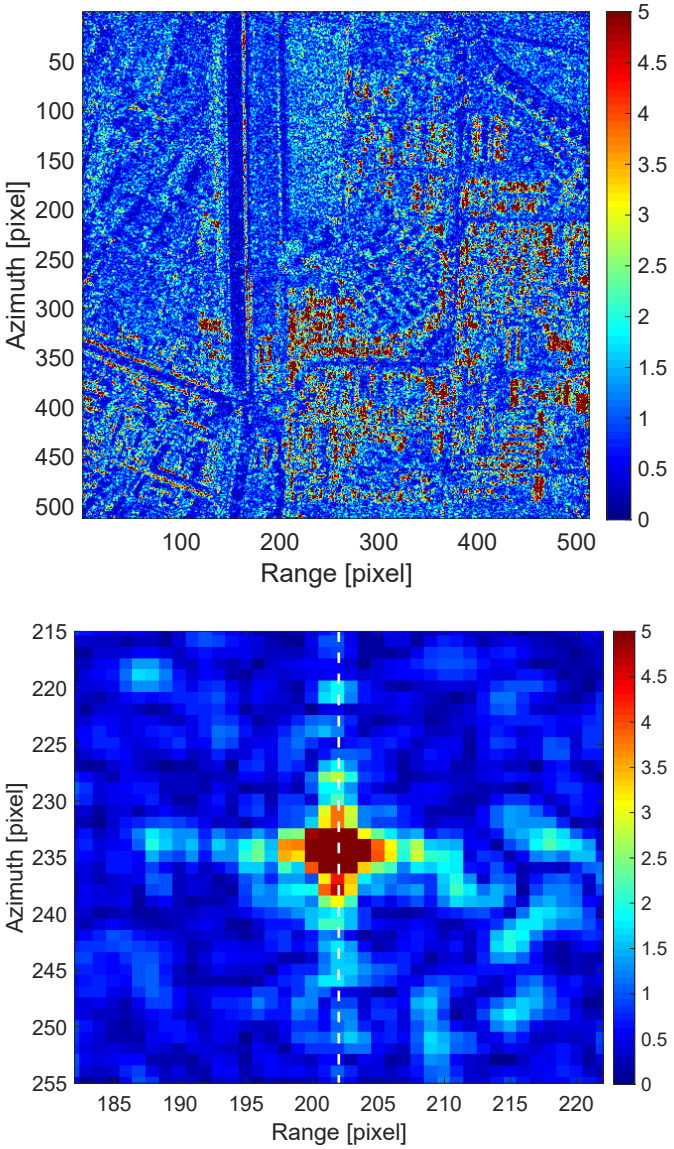

Figure 2: Tile 267 of the full scene in Figure 1. Top original image and bottom a zoom over a point target of the degraded image.

IV. Conclusions

In this paper a novel deep learning approach based on a DC2SCN fully convolutional neural network has been introduced to reconstruct super resolved Single Look Complex SAR images. The proposed method employs two channels to deal with both real and imaginary parts of the SAR image and combines a feature extraction stage with a reconstruction stage in order to reconstruct a super-resolved SAR image. Real COSMO-skymed SAR data have been used to quantitatively assess the effectiveness of the proposed framework. Different figures of merit have been used to asses the performance over rural and anthropized areas, all confirming the capability of the proposed framework to reliably create super-resolution images. Finally, the capability to preserve phase information has been also demonstrated, thus enabling the application of this framework to advanced SAR processing techniques such as interferometry.

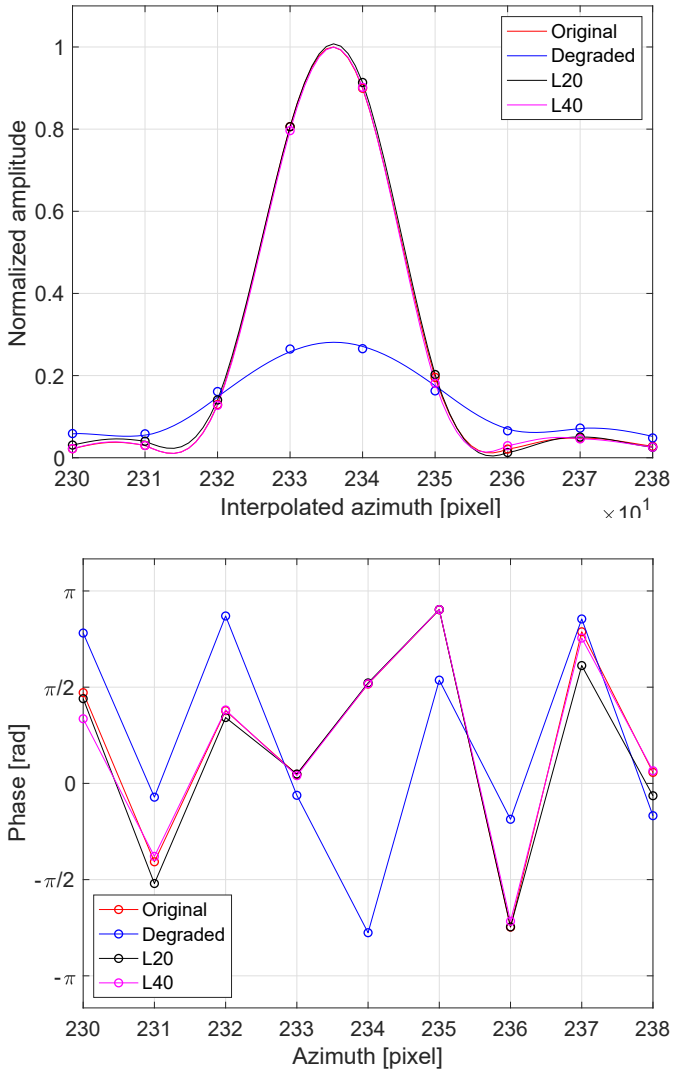

Figure 3: Azimuth section module (top) and phase (bottom) of the Tile 267 of the bright target shown in Figure 2 (bottom).

[6] W. Ma, Z. Pan, J. Guo, and B. Lei, "Achieving super-resolution remote sensing images via the wavelet transform combined with the recursive res-net," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 6, pp. 3512–3527, 2019.

REFERENCES

[1] F. Biondi, "Recovery of partially corrupted sar images by super-resolution based on spectrum extrapolation," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 2, pp. 139–143, 2016.

[2] L. Zhao, Q. Sun, and Z. Zhang, "Single image super-resolution based on deep learning features and dictionary model," *IEEE Access*, vol. 5, pp. 17 126–17 135, 2017.

[3] S. Hao, W. Wang, Y. Ye, E. Li, and L. Bruzzone, "A deep network architecture for super-resolution-aided hyperspectral image classification with classwise loss," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 8, pp. 4650–4663, 2018.

[4] M. Kawulok, P. Benecki, S. Piechaczek, K. Hrynczenko, D. Kostrzewa, and J. Nalepa, "Deep learning for multiple-image super-resolution," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 6, pp. 1062–1066, 2019.

[5] L. Lin, J. Li, Q. Yuan, and H. Shen, "Polarimetric sar image super-resolution via deep convolutional neural network," in *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2019, pp. 3205–3208.

[7] A. B. Molini, D. Valsesia, G. Fracastoro, and E. Magli, "Deepsum: Deep neural network for super-resolution of unregistered multitemporal images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 5, pp. 3644–3656, 2019.

[8] S. Lei, Z. Shi, and Z. Zou, "Coupled adversarial training for remote sensing image super-resolution," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 5, pp. 3633–3643, 2019.

[9] K. Peng and D. Ma, "Tree-structure CNN for automated theorem proving," in *Neural Information Processing - 24th International Conference, ICONIP 2017, Guangzhou, China, November 14-18, 2017, Proceedings, Part II*, ser. Lecture Notes in Computer Science, D. Liu, S. Xie, Y. Li, D. Zhao, and E. M. El-Alfy, Eds., vol. 10635. Springer, 2017, pp. 3–12.

[10] J. Yamanaka, S. Kuwashima, and T. Kurita, "Fast and accurate image super resolution by deep cnn with skip connection and network in network," in *ICONIP*, 2017.

[11] M. Lin, Q. Chen, and S. Yan, "Network in network," *CoRR*, vol. abs/1312.4400, 2014.

[12] ASI, "Italian space agency (asi) cosmo skymed mission, cosmo skymed sar products handbook." [Online]. Available: https://www.asi.it/wp-content/uploads/2019/08/COSMO.pdf

[13] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.