SUPER RESOLUTION APPROACH FOR THE SATELLITE DATA BASED ON THE GENERATIVE ADVERSARIAL NETWORKS

Mykola Lavreniuk[1], *Nataliia Kussul*[1,2], *Andrii Shelestov*[1,2], *Alla Lavrenyuk*[2], *Leonid Shumilo*[3,2]

[1] Space Research Institute NASU-SSAU, Kyiv, Ukraine
[2] National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine
[3] Department of Geographical Sciences, University of Maryland, College Park, MD, USA

ABSTRACT

In the past few years, medium and high-resolution data became freely available for downloading. It provides great opportunity for researchers not to select between solving the task with high-resolution data on small territory or on global scale, but with low-resolution satellite images. Due to high spectral and spatial resolution of the data, Sentinel-1 and Sentinel-2 are very popular sources of information. Nevertheless, in practice if we would like to receive final product in 10 m resolution we should use bands with 10 m resolution. Sentinel-2 has four such bands, but also has other bands, especially red-edge 20 m resolution bands that are useful for vegetation analysis and often are omitted due to lower resolution. Thus, in this study we propose methodology for enhancing resolution (super-resolution) of the existing low-resolution images to higher resolution images. The main idea is to use advanced methods of deep learning – Generative Adversarial Networks (GAN) and train it to increase the resolution for the satellite images. Experimental results for the Sentinel-2 data showed that this approach is efficient and could be used for creating high resolution products.

Index Terms— deep learning, Generative Adversarial Networks, GAN, super-resolution, Sentinel-2.

1. INTRODUCTION

Large variety of available satellite data allows us to solve many important tasks such as land cover classification, yield forecasting, logging detection, wild fire detection, and so on. The choice of a suitable satellite depends mainly on the tasks as well as on the investigated territory. For the small target territory very high resolution images are more suitable [1]. However, for the national and global cases medium resolution mostly utilized [2] - [4]. Sentinel-1 and Sentinel-2 data have the highest spatial and spectral resolutions among the data that are in open access. Thus, researchers in many studies used Sentinel data [5], [6]. Nevertheless, in practice if we would like to receive crop classification map in 10 m resolution we should use bands with 10 m resolution [7], [8]. Sentinel-2 has four such bands, but also has other bands, especially red-edge 20 m resolution bands that are useful for vegetation analysis and not utilized for crop mapping in 10 m resolution. Thus, and important tasks arises of enhancing resolution (super-resolution) of the existing low-resolution images to higher resolution images.

In general, super resolution task arises in many applied tasks for image and video processing [9]. Spatial image resolution enhancement methods can use multiple low spatial resolution images to create a single high-resolution image, or work with only one low spatial resolution input image. Solving the problem using several input images is described in detail in [9]. Let us consider the solution of the problem if there is only one input image.

The simplest method to increase the spatial resolution of the image is to use interpolation (nearest neighbor method, bilinear interpolation, bicubic interpolation, splines, etc.). The main disadvantage of such approaches is that it does not take into account the semantics of the image and the contours of the objects, and as a result, the resulting image contains a lot of blur.

With the development of deep learning methods and models for solving various applied problems, the problem of improving the spatial resolution of the image started to be solved using deep learning methods. In particular, one of the first applications of deep learning methods, namely Super-Resolution Convolutional Neural Network (SRCNN) to this task was proposed in 2015 [10]. However, the authors of this paper do not propose an end-to-end solution, but first increase the image to the desired size using the standard method of bicubic interpolation, and only then apply the proposed neural network to improve the quality of the image.

Improvements to the SRCNN method were proposed by the authors of [11] using the Very Deep Super Resolution (VDSR) model. The proposed VDSR model consisted of 20 layers, in contrast to the three layers that were in SRCNN model. In addition, the VDSR tries to learn the residual value between the target image and the interpolated one, rather than fully learning the translation from one image to another, as was proposed in case of SRCNN.

The problem of the previous two methods is the usage of interpolation at the initial stage, which leads to the use of a large number of parameters in the models, as well as the inability to learn the model to increase the image resolution. To address these issues, the Fast Super-Resolution Convolutional Neural Networks (FSRCNN) model has been proposed [12]. Low-resolution images used as an input of the model without any prior transformations. Methods that work on a similar principle were proposed in [13] – [15].

An alternative approach to increase image resolution is to use Generative Adversarial Networks (GAN) [16], which consists of two neural networks – a generator and a discriminator. Such approach was first proposed in [17] based on the Super Resolution Generative Adversarial Networks (SRGAN). The idea of the model is as follows. One neural network (generator) generates a higher resolution image, and another neural network (discriminator) tries to guess whether this image is real or generated by the network. Training of the GAN model stopped when the first network learns to generate "real" high-resolution images that the other network will not be able to distinguish from real ones.

Several other papers [18] – [20] proposed to use GAN models also for image super resolution tasks and shows that such models provide the most accurate results and are the most suitable for solving super resolution problem. However, these methods have been applied to ordinary images, not to satellite multispectral data. In [21] and [22] papers researchers proposed to use GAN models to remote sensing tasks such as increasing the resolution of Digital Elevation Models (DEMs) and increasing the resolution of weather radar data, which allows produce more accurate weather forecasts. Nevertheless, there are no researches about increasing the resolution of Sentinel-2 bands. Thus, we would like to close this gap and propose the methodology for increasing the resolution of all other available 20m resolution Sentinel-2 bands to 10m.

2. STUDY AREA AND MATERIALS DESCRIPTION

To conduct experiment, we downloaded Sentinel-2 tile 35UQR for 1 June 2021, located in Kyiv region of Ukraine and include all main land cover types – cropland, water bodies, artificial objects, forest, grassland and include 11% of clouds. Preprocessing of images was conducted with use of Sen2Core software and include atmosphere and radiometric correction. We used green, blue and NIR bands for training, while red band was used for testing. For each 10 m resolution band we created band with 20 m resolution by warping it using the cubic interpolation. After that training data were cut into small patches with the size of 256x256 and 128x128 pixels for 10 m and 20 m bands respectively.

3. METHODOLOGY

In this paper, we propose to use GAN model to solve the image super resolution task, in particular for converting Sentinel-2 20m bands into 10m resolution images. The workflow of the proposed methodology is shown in the figure 1. Our GAN consists of two neural networks – a generator and a discriminator. As a baseline model, we utilized an idea of the SRGAN [17] that was used for solving super resolution tasks of regular image and adopt it to satellite data.

At the first stage, GAN model trained to convert back to 10m the warped red, green, blue and NIR bands at 20m resolution. After that, trained GAN model could be used to convert other bands from 20m to 10m resolution. In contrarily to traditional images satellite image should be divided into patches. The input image size is equal to 128x128x1 pixels, the output image size is equal to 256x256x1 pixels. Taking into account that the edges of the image divided objects into parts, thus we propose to generate patches with overlap in 1\4 of image size from all sides. As a result, we cut off central part of the obtained image (256x256x1 pixels) with the size of 128x128x1 pixels, except the patches that were taken from the edges of satellite image.

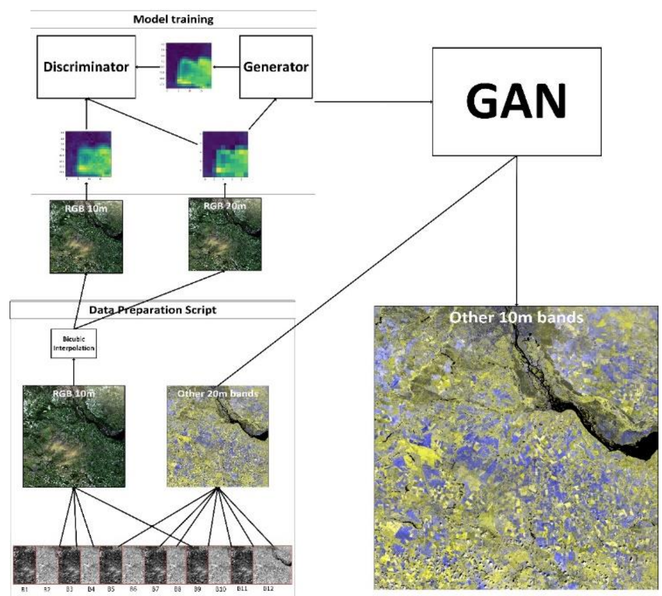

Figure 1. The workflow for solving the image super resolution task.

The discriminator network is commonly used neural network for binary classification (real 10m image or generated from 20m) with sigmoid layer in the end. Thus, we chose ResNet as a baseline discriminator, however, any backbone could be used for the purpose. ResNet can be easily changed to state-of-the-art backbones, such as Efficientnet, Normalizer Free Network (NFnet) or transformers.

The generator network in our case consists of 16 identical blocks connected with skip connections with element wise sum operation on it. Each block has two convolutional layers with 3×3 kernels and 64 feature maps followed by batch-normalization layers and activation function. For both our models we utilized Leaky ReLU function as an activation function. After that, we have specific block for increasing the spatial resolution of the feature map. Taking into account that we would like to obtain accurate values of the pixels without blur, thus interpolation or other approaches like upsample layer with transpose convolution are not appropriate. For this purpose, we chose efficient sub-pixel convolutional layer that is specifically designed for super pixel tasks [13].

For joint training the discriminator and generator networks total loss function consist of two parts. First one is the standard binary cross-entropy for discriminator. The second one is the loss for the generator. In the basic GAN the generator, loss is calculated based on the discriminator outputs for the generated data. For super resolution task, the main goal of generator is not only learned how to generate images that looks like real one, but also to preserve pixel values and object shapes. Thus, we have added content loss that is calculated on feature maps of the Resnet model like in [17], which is more invariant to changes in pixel space on contrarily to commonly used MSE loss [13]. So if X is input image, Y_{pred} is generated image, $\widehat{Y_{pred}}$ is discriminator output for generated image, Y is original high resolution image, $loss$ of GAN can be estimated by formula (5) based on the discriminator loss ($loss1$) and generator loss ($loss2$), where α and β are weight coefficients.

$$Y_{pred} = generator(X). \quad (1)$$

$$\widehat{Y_{pred}} = discriminator\left(Y_{pred}\right) \quad (2)$$

$$loss1 = \frac{1}{n}\sum_{i=1}^{n} Y_{pred_i} * log\widehat{Y_{pred_i}} + \left(1 - Y_{pred_i}\right) * \log(1 - \widehat{Y_{pred_i}}) \quad (3)$$

$$loss2 = MSE\left(\mathrm{Y} - Y_{pred}\right) \quad (4)$$

$$loss = \alpha * loss1 + \beta * loss2 \quad (5)$$

4. RESULTS

The model was initialized with random weight coefficient for all layers. The input layer contains 64 convolutional layers with 9 kernel size and PReLU activation function. Then the signal is going through 16 consecutive residual blocks with 64 convolutional layers with 9 kernel size and PReLU activation and 64 convolutional layers with the same kernel size and batch normalization. After this we have one more block with 64 convolutional layers with 3 kernel size and batch normalization and summing with output from input layer skip connection. The last step is propagation of information through transpose filter for up-sampling and output convolutional filter that generate resulting image. For generator loss estimation we are using weight coefficients: 1 for MSE of generator and 0.1 for binary cross-entropy of discriminator.

After 10 epochs of training with 10 iteration in each and 32 batch size, resulting the model reached resulting MSE equal to 2075, for Sentinel-2 not normalized pixels with values between 0 to 10,000. Testing was conducted by calculation of MSE between generated image by GAN from 20 m warped red channel and original 10 m red channel. Figure 2 demonstrate output of GAN model for input sample. Usually in use of transpose filters, result on the center of input sample (when transpose values calculation perfumed maximum times) rather than on corners, where filtration was conducted only 1 or few times. So, in result image we take only the central part of output sample with 128x128 size and sliding window for modeling with 64 pixels step size. Figure 3 demonstrate result on the full image prediction.

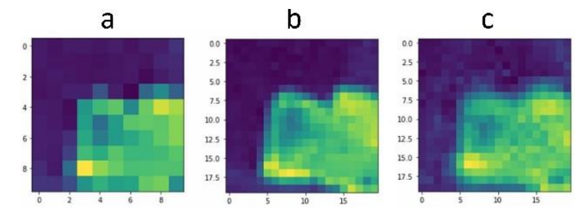

Figure 2. The super resolution on one sample: a – 20 m warped image, b – 10 m original image, c – GAN output

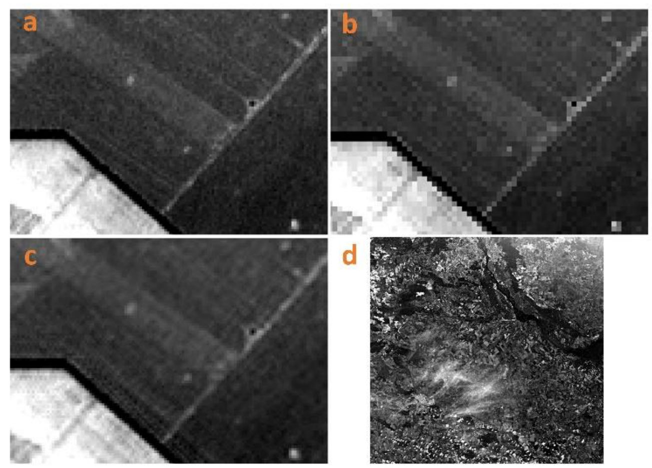

Figure 3. Result of GAN super resolution approach on the image: a – original 10 m red channel image, b – warped to 20 m red channel, c – result of super resolution model, d – result on the full image.

5. DISCUSSION AND CONCLUSIONS

In this work, the methodology for increasing resolution of satellite images is proposed based on the GAN model. It was successfully trained on red, green, blue and NIR channels of Sentinel-2 satellite images and allows us to increase the resolution of all other available 20m resolution bands to 10m. Such bands with increased resolution can be used for solving different applied tasks, especially red edge bands for more accurate vegetation analysis, for example for crop yield forecasting [23]. The limitation of proposed methodology based on the GAN model is the possibility of appropriate increasing of the resolution with the scale factor of two or four. If we would like to increase the resolution with scale factor of larger than four we should use additional information of high resolution, for example NDVI, land cover map, etc.

6. ACKNOWLEDGMENT

The authors acknowledge the funding received by the National Research Foundation of Ukraine from the state budget 2020.02/0292 "Deep learning methods and models for applied problems of satellite monitoring" (NRFU Competition "Support for research of leading and young scientists").

7. REFERENCES

[1] Kislov, Dmitry E., et al. "Extending deep learning approaches for forest disturbance segmentation on very high-resolution satellite images." *Remote Sensing in Ecology and Conservation,* vol. 7, no. 3, pp. 355-368, (2021).

[2] Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., ... & Townshend, J. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, 342(6160), 850-853.

[3] Waldner F., Schucknecht A., Lesiv M. et. al., "Conflation of expert and crowd reference data to validate global binary thematic maps," *Remote sensing of environment*, vol. 221, pp. 235-246, 2019.

[4] Shelestov A., Lavreniuk M., Kussul N. et. al., "Exploring Google Earth Engine Platform for Big Data Processing: Classification of Multi-Temporal Satellite Imagery for Crop Mapping," *Frontiers in Earth Science*, vol. 5, 2017.

[5] d'Andrimont, Raphaël, et al. "From parcel to continental scale--A first European crop type map based on Sentinel-1 and LUCAS Copernicus in-situ observations," *arXiv preprint arXiv:2105.09261*, 2021.

[6] Yi Zhiwei, Li Jia, and Qiting Chen, "Crop classification using multi-temporal Sentinel-2 data in the Shiyang River Basin of China," *Remote Sensing,* 12.24, pp 1-21, 2020.

[7] Shelestov A., Lavreniuk M., Vasiliev V., et. al., "Cloud Approach to Automated Crop Classification Using Sentinel-1 Imagery," *IEEE Transactions on Big Data*, vol. 6, no. 3, pp. 572-582, 2019.

[8] Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. (2017). Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, 14(5), 778-782.

[9] Yang, Jianchao, and Thomas Huang. "Image super-resolution: Historical overview and future challenges." *Super-resolution imaging*, pp. 20-34, 2010.

[10] Dong, Chao, et al. "Image super-resolution using deep convolutional networks." *IEEE transactions on pattern analysis and machine intelligence* 38.2 (2015): 295-307.

[11] Kim, Jiwon, Jung Kwon Lee, and Kyoung Mu Lee. "Accurate image super-resolution using very deep convolutional networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

[12] Dong, Chao, Chen Change Loy, and Xiaoou Tang. "Accelerating the super-resolution convolutional neural network." *European conference on computer vision*. Springer, Cham, 2016.

[13] Shi, Wenzhe, et al. "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

[14] Lim, Bee, et al. "Enhanced deep residual networks for single image super-resolution." *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2017.

[15] Tai, Ying, Jian Yang, and Xiaoming Liu. "Image super-resolution via deep recursive residual network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

[16] Ian, Goodfellow, et al. "Generative adversarial nets." *In Advances in neural information processing systems*. (2014): 2672-2680.

[17] Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

[18] Yuan, Yuan, et al. "Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2018.

[19] Wang, Xintao, et al. "Esrgan: Enhanced super-resolution generative adversarial networks." *Proceedings of the European conference on computer vision (ECCV) workshops*. 2018.

[20] Johnson, Justin, Alexandre Alahi, and Li Fei-Fei. "Perceptual losses for real-time style transfer and super-resolution." *European conference on computer vision*. Springer, Cham, 2016.

[21] Demiray, Bekir Z., Muhammed Sit, and Ibrahim Demir. "D-SRGAN: DEM super-resolution with generative adversarial networks." *SN Computer Science* 2.1 (2021): 1-11.

[22] Chen, Hongguang, et al. "Generative adversarial networks capabilities for super-resolution reconstruction of weather radar echo images." *Atmosphere* 10.9 (2019): 555.

[23] Kolotii, A., Kussul, N., Shelestov, A., Skakun, S., Yailymov, B., Basarab, R., ..., Ostapenko, V. "Comparison of biophysical and satellite predictors for wheat yield forecasting in Ukraine." *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 40(7W3), 2015, pp. 39-44.