INFORMATION TECHNOLOGY

UDC 004.021
DOI: 10.15587/1729-4061.2024.313690

REMOVING CLOUDINESS ON OPTICAL SPACE IMAGES BY A GENERATIVE ADVERSARIAL NETWORK MODEL USING SAR IMAGES

Mykola Romanchuk
Corresponding author
PhD, Senior Researcher,
Deputy Head of Scientific Center*
E-mail: romannik@ukr.net
Andrii Zavada
PhD, Senior Researcher, Deputy Head
of Scientific-Research Department*
Olena Naumchak
Adjunct*
Leonid Naumchak
Senior Researcher*
Iryna Kosheva
Lecturer
Department of Computer Integrated Technologies
and Cybersecurity**
*Scientific Center**
**Korolov Zhytomyr Military Institute
Myru ave., 22, Zhytomyr, Ukraine, 10004

The object of this study is the process of removing cloudiness on optical space images. Solving the cloudiness removal task is an important stage in processing data from the Earth remote probing (ERP) aimed at reconstructing the information hidden by these atmospheric disturbances. The analyzed shortcomings in the fusion of purely optical data led to the conclusion that the best solution to the cloudiness removal problem is a combination of optical and radar data. Compared to conventional methods of image processing, neural networks could provide more efficient and better performance indicators due to the ability to adapt to different conditions and types of images. As a result, a generative adversarial network (GAN) model with cyclic-sequential 7-ResNeXt block architecture was constructed for cloud removal in optical space imagery using synthetic aperture radar (SAR) imagery. The model built generates fewer artifacts when transforming the image compared to other models that process multi-temporal images.

The experimental results on the SEN12MS-CR data set demonstrate the ability of the constructed model to remove dense clouds from simultaneous Sentinel-2 space images. This is confirmed by the pixel reconstruction of all multispectral channels with an average RMSE value of 2.4 %. To increase the informativeness of the neural network during model training, a SAR image with a C-band signal is used, which has a longer wavelength and thereby provides medium-resolution data about the geometric structure of the Earth's surface. Applying this model could make it possible to improve the situational awareness at all levels of control over the Armed Forces (AF) of Ukraine through the use of current space observations of the Earth from various ERP systems

Keywords: remote probing, image reconstruction, cloud removal, generative adversarial network

Received date 06.08.2024
Accepted date 15.10.2024
Published date 30.10.2024

***How to Cite**: Romanchuk, M., Zavada, A., Naumchak, O., Naumchak, L., Kosheva, I. (2024). Removing cloudiness on optical space images by a generative adversarial network model using SAR images. Eastern-European Journal of Enterprise Technologies, 5 (2 (131)), 6–12. https://doi.org/10.15587/1729-4061.2024.313690*

1. Introduction

Given current global and regional challenges, developed countries are revising their approaches to information technology support to national security. One of the possible ways to achieve this is the active and systematic use of space images, namely multi-spectral data from ERP.

The change in emphasis in the methods and forms of conducting modern warfare has led to a significant expansion of tasks that are solved with the help of ERP data from various space vehicles (SVs) of medium and high resolution and when they are integrated. A promising option in this case is the use of open multispectral data from SVs.

Under the conditions of rapid development of technologies, including in the field of ERP, there are many opportunities to use open ERP data obtained from various sensors in various formats. Such information is often complementary, and its integration could make it possible to obtain more detailed and accurate observational data. This would provide a more reliable interpretation of data from space images. But the main disadvantages are the complexity of integrating different time information from different sensors and the insufficient number of available multimodal data sets.

Obtaining a clear optical image is affected by such factors as clouds, fog, and smoke. The experience of a full-scale war in Ukraine highlighted this problem. The first cloud-free multispectral image of Mariupol from the Sentinel-2 spacecraft was obtained only on March 14, 2022, after the Russian invasion on February 24, 2022. Since the combat zone is the object of special attention and the priority direction of space surveillance, the influence of smoke, cloudiness, etc. greatly complicates the possibility of using satellite images as a source of intelligence.

These factors have a minor effect on the performance of SAR sensors. This make it possible to use them under any weather conditions and time of day. However, the geometric distortions and characteristics inherent in SAR images make the study of discriminative features a more difficult task.

Thus, for most common ERP datasets, SAR images generally have lower classification accuracy than multispectral ones. Considering their complementary nature, an important task is the integration of multispectral and SAR data.

Privately-held companies launch their own ERP SVs, promoting the development of competition and innovation, offering open access to their own space data, which creates conditions for their use for military purposes. For example, the Copernicus mission provides access to space images from the Sentinel-2 spacecraft with a spatial resolution of about 10 m. Such images can be used to solve various tasks. The application of integrated multi-time and multi-parameter data from ERP systems, which are publicly available for use, is one of the ways to increase the level of information technology support, namely situational awareness at all levels of management of the Armed Forces of Ukraine. But considering that atmospheric obstacles, such as clouds, are regular weather phenomena, their presence is a problem for space observation of the Earth, which is attempted to be solved by automated reconstruction of the noisy or cloud-covered underlying surface [1]. Therefore, the improvement of space image processing algorithms is an urgent task.

2. Literature review and problem statement

Identification of cloudiness on space images obtained by ERP passive sensors in the visible and infrared parts of the electromagnetic spectrum is an important stage of pre-processing for the construction of high-quality geoinformation products. Cloudiness is the main limiting factor for the use of time series data obtained by ERP optical sensors.

Removal of cloudiness becomes an indispensable stage of pre-processing of ERP data and aims to reconstruct the missing information caused by these natural phenomena. According to the results of a comprehensive review [2], conventional methods can be divided into three groups: multispectral, multitemporal, and "painting" methods, and the rest are their hybrid combination. However, the disadvantage is that the issue of automation and increasing the accuracy of information recovery for complex weather conditions remains insufficiently resolved. The reason for this may be that most existing methods are focused on processing images with fixed time intervals between images, which does not make it possible to adequately process images with dynamic changes in cloudiness.

Multispectral approaches are used in the case of haze and thin cirrus clouds, when optical signals are not completely blocked but partially depend on the wavelength of absorption and reflection. In such cases, information about the surface can be restored using mathematical or physical models [3]. The advantage of multispectral methods is the possibility of using information from original space images, which does not require additional data. However, they can be effective only for translucent clouds since such models use optical information from the same image to restore, which in the case of dense covers does not eliminate the cloudiness problem or distorts the resulting information.

Multitemporal approaches reconstruct cloudy space images by integrating information from reference images obtained under clear sky conditions [4]. Multitemporal methods of catalog learning are also sometimes used for this purpose. Multi-time data can also come from different sensors at different spacecraft [5]. These methods are popular due to their simplicity and practicality. Their drawback is the need to select the freshest, least cloudy observations, as they replace corrupted cloudy pixels with real cloud-free observations. Therefore, the use of such approaches has certain limitations: if there are no completely cloud-free observations at a certain point in time, then the resulting forecast will also contain cloudy pixels. The reason for this is the time dependence between the acquisition of the reference and reconstructed image, especially under conditions of rapid changes in the underlying surface. The only way to overcome this limitation is to increase the sampling rate or time range of the observation, but cloudiness can persist for a long time depending on terrain and seasonality, so the closest cloud-free observation may be irrelevant.

The first attempts to solve the problem of cloud removal in space images were based on the assumption that cloudy areas and the rest of the image have the same statistical and geometric structure. "Correction" approaches fill in damaged fragments using surface information from clean regions of the same space image. For example, the interpolation method using a sparse dictionary for adaptive restoration of images with dense clouds [6]. Such methods do not require additional data but achieve good results only in the case of slight cloudiness. To overcome this problem, the process of selecting the most similar pixel for cloning is often based on the use of auxiliary data such as time-lapse [7] or SAR data [8]. Although such methods give good results, their disadvantage is increased complexity due to the need to use a large amount of multi-temporal or multi-sensor auxiliary data.

Close to them are interpolation methods that extract information hidden by clouds due to spatial interference from neighboring or sufficiently close pixels without clouds. Such methods are based on the method of nearest neighbors or kriging. Their disadvantage is that clouds may cover large contiguous areas continuously, in which case the proximity assumption is not effective.

There are two aspects to using SAR data for cloud removal from multispectral imagery. The first is global fusion, which controls the relationship between all local optical windows to preserve the structure of the reconstructed region in accordance with the rest of the cloud-free regions. The second is local fusion, which transfers the additional information embedded in the SAR image corresponding to the cloud areas to create reliable textural details of the missing parts and uses dynamic filtering to reduce the performance degradation caused by speckle noise. The approach make it possible to create high-quality cloud-free images.

Prepared models based on neural networks are able to handle different types of clouds and other atmospheric disturbances. The basis for approaches based on convolutional neural networks (CNN) was the essence of the cloud removal process, which involves restoring an image from a noisy analog [9]. Their disadvantage is the need for a large amount of training data to achieve greater accuracy. The reason for this is the complexity of the variety of cloud types, which can affect the effectiveness of model training.

A more modern architecture is the conditional generative adversarial network (cGAN). In [10], a cGAN model is trained to remove detected clouds from visible (RGB) images from Worldview-2 using images in the near-infrared region as auxiliary data. The McGAN model [5] is a GNN built on the basis of the pix2pix model, which maps cloud images in the visible and near-infrared ranges of the spectrum. The next model, Cycle-GAN, does not require paired cloudy and cloudless space images for training [11]. The disadvantage of

cGAN is the high computational cost and the instability of training, and therefore prediction, in the case of low-quality input data (for example, satellite images with heavy clouds).

The models above give good results, but the data sets used are very limited. The disadvantages of the considered approaches are the need for large sets of real data; focus on narrow areas of interest; dependence on synthesized cloud samples. All this makes it impossible to generalize the data to more extensive areas and real conditions. Another significant disadvantage is the impossibility of using a set of weights from one area to process another, which differs under climatic conditions.

Therefore, approaches based on combining optical and SAR data deserve special attention. In [12], SAR data is combined with information from auxiliary optical imaging using the closest spectral matching method [13]. Although optical and radar sensors measure different quantities, which makes their comparison difficult, this became the basis for the development of deep learning approaches [14].

Since SAR is invariant to daylight conditions and robust to atmospheric noise, but differs in measured values from optical sensors, the problem of bridging the gap between their modalities arises. Although the approach of combining optical and radar data demonstrates the possibility of removing cloudiness from images, not all spectral characteristics of the earth's surface can be obtained from the corresponding radar measurements, which fundamentally limits the sar2opt method [15].

Hence, the ability of GAN models to learn to transform images from one modality to another can help fill gaps caused by cloudiness; as well as generalize and adapt to new regions and conditions. Therefore, in the future, this approach will be used as a basis for building a model for cloudiness removal on optical space images.

3. The aim and objectives of the study

The aim of our study is to improve the process of converting optical space images by removing cloudiness by reconstructing them with a generative adversarial network using additional SAR data. This will enable the process of cloudiness removal on optical space images with the necessary quality indicators.

To achieve the set goal, the following tasks must be solved:

– to build a model for cloudiness removal on optical space images;

– to conduct an experimental study of the model for removing cloudiness on optical space images.

4. The study materials and methods

The object of research is the process of removing cloudiness on optical space images. As the main hypothesis, it was suggested that the use of generative adversarial networks to remove cloudiness in optical space images could provide the necessary image quality for further processing.

The following limitations and assumptions were adopted during the research:

– input images are georeferenced cloud-free and cloud-covered Sentinel-2 multispectral space images in 12 bands (spatial resolutions of RGB bands: 2, 3, 4 and 8–10 m band, 5, 6, 7, 8A, 11 and 12–20 m, and bands 1, 9–60 m), band 10 is used for atmospheric correction, so it is excluded;

– input SAR images are Sentinel-1 radar measurements of the same region in 2 polarization ranges VV and VH;

– the size of the input image is 256×256 px^2;

– digital image processing procedures are applied;

– indicators that comprehensively ensure high accuracy of visual analysis of images and their correspondence to the original were chosen to evaluate the results, namely: Precision, Recall, F1, MAE, RMSE, PSNR, SSIM and SAM;

– the following hardware was used during the experimental study of the model: 12th Gen Intel(R) Core (TM) i5-12400F 2.50 GHz processor, 64 Gb RAM, NVIDIA GeForce RTX 3050 Gaming OC 8G video adapter;

– the following software was used for the experiment: Ubuntu 20 LTS OS with CUDA 12.1 and GPU support, PyTorch 2.3.0, torchvision 0.18.0.

5. Results of investigating the model for cloudiness removal on optical space images

5. 1. Basic elements of the model for cloudiness removal in optical space images

Taking into account the approaches described in [11, 15], a model for the reconstruction of cloud-obscured information from optical satellite images is proposed, which explicitly processes a continuous mask of cloudiness values to preserve cloud-free pixels while correcting cloudy areas based on SAR data.

This model is trained without strict pixel-by-pixel alignment of cloudy and cloudless observations using fine-grained masks with continuous values. This is a major difference from previously discussed models that use binary cloudiness information or are more cloudiness agnostic.

The construction of a neural network for cloud removal in optical space images corresponds to the architecture of a cyclic-sequential 7-ResNeXt (in order to use a large receptive field due to aggressive convolution) block GAN with two output discriminators. The model consists of three parts: encoder, decoder, and generator (Fig. 1) with residual blocks (Fig. 2) [12].

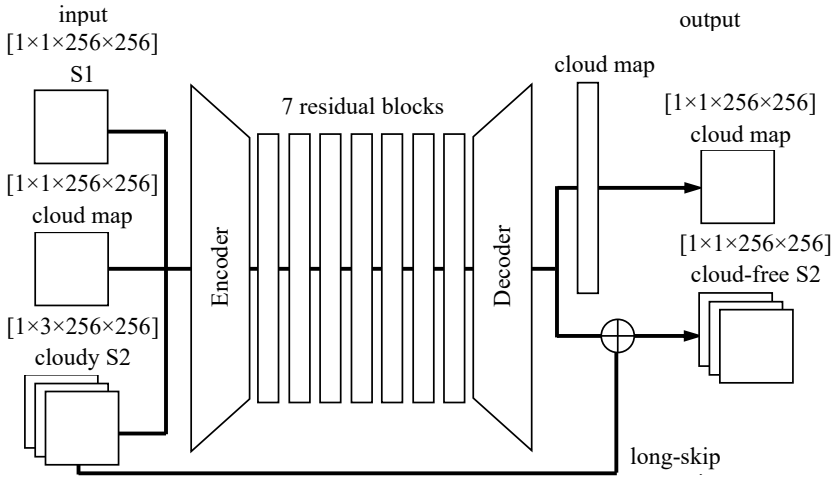

Fig. 1. Architecture of a cloud removal model in optical satellite imagery using generative adversarial networks

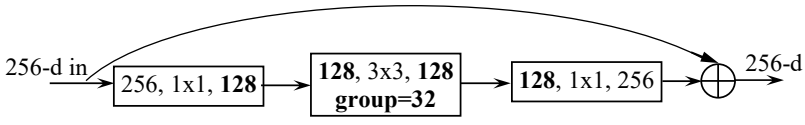

Fig. 2. Residual block of the cloud removal model

During training, pairs of SAR and optical space images and corresponding cloud masks are fed to the encoder input. The output of the decoder is a map of predicted cloudiness masks. After aggregating the feature maps with the overcast images, the inverse hyperbolic tangent of the slope angle is applied to demodulate the effect of the output nonlinearity on the long-passed pixels of the overcast input image. As a result, when used, the trained model reconstructs a cloudless optical image.

For the loss function, the cycleGAN approach [16] was used for both the generator network and the discriminator, which can be expressed as:

$$L_{cGAN} = E_{x,y\sim P_{data}(x,y)}\left[\log\left(D(x,y)\right)\right] + \\ + E_{x,z\sim P_{data}(x,z)}\left[\log\left(1-D\left(x,G(x,z)\right)\right)\right], \quad (1)$$

where L_{cGAN} is the loss function, E is the encoder, D is the decoder, G is the generator, and D tries to maximize while G tries to minimize the objective.

To reduce blurring and bring the original image closer to the target, the L loss function was used as follows:

$$L_{adv} = \left(D_{S1}\left(\acute{S}1-1\right)\right)^2 + \left(D_{S2}\left(\acute{S}2-1\right)\right)^2, \quad (2)$$

$$L_{cyc} = \left\|m\cdot\left(S1-\breve{S}1\right)\right\|_1 + \left\|(1-m)\cdot\left(S2-\breve{S}2\right)\right\|_1, \quad (3)$$

$$L_{idt} = \left\|m\cdot\left(S1-\dot{S}1\right)\right\|_1 + \left\|(1-m)\cdot\left(S2-\dot{S}2\right)\right\|_1, \quad (4)$$

$$L_{aux} = \left\|(1-m)\cdot\left(m-\acute{m}\right)\right\|_1, \quad (5)$$

$$L_{all} = \lambda_{adv}L_{adv} + \lambda_{cyc}L_{cyc} + \lambda_{idt}L_{idt} + \lambda_{aux}L_{aux}, \quad (6)$$

where $S1$, $S2$ are input SAR and optical data, respectively: $\acute{S}1$, $\acute{S}2$ – synthesized, $\breve{S}1$, $\breve{S}2$ – regularized, $\dot{S}1$, $\dot{S}2$ – intra-domain transformations, D_{S1}, D_{S2} – discriminant networks of these data; m is cloudiness mask, λ_{adv}=5.0, λ_{cyc}=5.0, λ_{idt}=1.0, λ_{aux}=10.0 – hyperparameters for linear fusion of individual losses for L_{all}.

Loss weights are set similarly to [16], with small manual adjustments. L_{adv} is the adversary loss originally proposed in LSGAN [17], which implements the least square error function for discriminator classifications. L_{cyc} and L_{idt} are the same as in [16], but weighted pixel by pixel with the cloud map. L_{aux} is the cloud map regression loss introduced to ensure that the learned residual feature maps of the optical image data are sparse so that non-cloud pixels are almost uncorrectable.

The neural network model was evaluated using Precision and Recall, $F1$ [18], Mean Absolute Error (*MAE*), Root Mean Square Error (*RMSE*), Peak Signal-to-Noise Ratio (*PSNR*), Structural Similarity (*SSIM*), and Spectral Reflection Angle (SAM) [14].

MAE and *RMSE* are pixel-level metrics that measure the average deviation between the target and predicted images in absolute terms and in the units of interest:

$$MAE(x,y) = \frac{1}{C\cdot H\cdot W}\sum_{c=h=w=1}^{C,H,W}\left|x_{c,h,w}-y_{c,h,w}\right|, \quad (7)$$

$$RMSE(x,y) = \\ = \sqrt{\frac{1}{C\cdot H\cdot W}\sum_{c=h=w=1}^{C,H,W}\left(x_{c,h,w}-y_{c,h,w}\right)^2}, \quad (8)$$

PSNR is an image-level metric that measures how well the reconstructed image matches the original target image in signal-to-noise ratio:

$$PSNR(x,y) = 20\cdot\log_{10}\left(\frac{1}{RMSE(x,y)}\right), \quad (9)$$

SSIM is a metric that quantifies the structural differences between the target and predicted images, designed to capture the perceived change in structural information between two given images, as well as differences in brightness and contrast:

$$SSIM(x,y) = \frac{\left(2\mu_x\mu_y+\epsilon_2\right)}{\left(\mu_x+\mu_y+\epsilon_1\right)\left(\sigma_x+\sigma_y+\epsilon_2\right)}, \quad (10)$$

SAM is a measure that quantifies the spectral angle between the target and predicted images by measuring their similarity in terms of rotations in spectral range space:

$$SAM(x,y) = \cos^{-1}\left(\frac{\sum_{c=h=w=1}^{C,H,W}x_{c,h,w}\cdot y_{c,h,w}}{\sqrt{\sum_{c=h=w=1}^{C,H,W}x_{c,h,w}^2\cdot\sum_{c=h=w=1}^{C,H,W}y_{c,h,w}^2}}\right), \quad (11)$$

where x, y are images to be compared with pixel values $x_{c,h,w}$, $y_{c,h,w}\in[0,1]$, dimensionality C=3, H=W=256, mean values μ_x, μ_y, standard deviations σ_x, σ_y, covariance σ_{xy} and small numbers ϵ_1, ϵ_2 to stabilize calculations.

5. 2. Experimental study of the model for cloudiness removal on optical space images

The cloud removal data set SEN12MS-CR was used for the experiment. It consists of selected orthocorrected, terrain-referenced cloud and cloud-free 13-band multispectral Sentinel-2 images, as well as a corresponding Sentinel-1 image from globally distributed regions of interest, evenly distributed over all continents and in all meteorological seasons, having an average coverage size of approximately 52×40 km^2, corresponding to a full-scale image size of 5200×4000 px^2. Each scene in the dataset is converted to the Mercator coordinate system and then partitioned into 256×256 px^2 plots with 50 % spatial overlap between adjacent plots. The average value of cloudiness is approximately 47.93±36.08 %, i.e., about half of all optical image information is cloudy. This value is close to the real value of 55 % cloud cover over land. To increase the number of training samples in the data set, synthetic data were generated using the "copy" method, similar to the method used in [19]. It is based on increasing the cloud cover in the image using cloud masks from other real cloud images.

During the training of the neural network model, subsets of 10,000 satellite images from several regions were used as input data. The hyperparameters of training the neural network model are as follows: the weights of the network w are initialized with a sample from the Gaussian distribution $w\sim N(\mu=0, \sigma_2=0{,}02)$; ADAM optimizer with initial learning rate lr=0.0002 and impulse parameters β=(0.5, 0.999). The

model was trained over 50 epochs with initial learning rate lr, then for n_{decay}=25 epochs and further with a multiplicative decay of the learning rate to ensure well-controlled optimization process during training [16]. The results of model training were compared with the results of model training described in [19] and are given in Table 1.

Table 1

Comparing the results of models for cloud removal in optical space images

Model	Precision	Recall	*F*1	*MAE*	*RMSE*	*PSNR*	*SSIM*	*SAM*
Known model [20]	0.458	0.586	0.514	0.017	0.023	35.802	0.904	9.936
Constructed model	0.461	0.589	0.517	0.018	0.024	36.028	0.909	9.866

The results show that the built model is capable of reconstructing cloud-covered areas while preserving information from non-cloudy areas of the image. The proposed neural network model improves image reconstruction by predicting uncertainty and outperforms the approach reported in [19]. Metrics such as *SSIM* and *F*1 show improvements in image restoration quality. In particular, the *SSIM* value increased by 0.005, which indicates an improvement in the visual similarity of the restored image with the original. This indicates an improvement in the quality of reproduction of textures and image structures. However, forecasting uncertainty marginally degrades *RMSE* by 0.001. This is consistent with the fact that the proposed loss function implies a trade-off between mean optimization and variance estimation.

6. Discussion of results based on the model for cloud removal in optical space images

A neural network model was built for cloud removal on optical space images based on the block generative competitive network architecture (Fig. 1). The choice of GAN is due to its high quality of generating realistic images, which does not require modeling of the probability distribution to obtain new data, and the possibility of applying transfer learning. GAN blocks consist of cyclic-sequential 7-ResNeXt, which provides detection of more complex and multi-level features of input data without significantly increasing the number of block parameters. To train the model built, a loss function based on the cycleGAN approach was used, which ensures the preservation of key characteristics with minimal distortions during image generation (1) to (6).

To evaluate the results of our experiment, the following indicators were chosen: Precision, Recall, F1, MAE, RMSE, PSNR, SSIM, and SAM, calculated from expressions (7) to (11), as they comprehensively provide high accuracy of visual analysis of images, evaluating their quality and correspondence to the original.

The results of our experiment (Table 1) demonstrate the ability of the proposed model to successfully restore cloud-covered areas from simultaneous Sentinel-2 space images, while preserving information from non-cloudy parts of the image. This is confirmed not only by visually evaluating the RGB channels of the input image but also by the pixel reconstruction of all multispectral image channels with an average RMSE value of 2.4 % (8). To improve the informativeness of the neural network during model training, a SAR image with a C-band signal is used, which has a longer wavelength and thereby provides medium-resolution data about the geometric structure of the Earth's surface. Accordingly, the use of auxiliary SAR images at the stage of pre-processing of images is effective, which is confirmed by the results given in Table 1. This enables fulfillment of most tasks in the field of national security that use data from ERP.

Unlike the method proposed in [5], in which matrix factorization and error correction methods were used to remove clouds, the proposed model provides a more accurate restoration of image textures and structure. For heterogeneous clouds and multispectral images, the matrix factorization method is less effective. The model built achieves a greater degree of consistency with the original images due to a more complex architecture and adaptation to different types of data.

The approach from [6] is effective for high spatial resolution images. The proposed model provides more versatile results, especially for medium resolution images. Sparse dictionaries have limited efficiency when working with multispectral channels, and generative networks can model more complex spectral features.

A method from [7] is less adaptable to complex textures because interpolation does not always make it possible to preserve the geometric structures of objects, which are important for the analysis of the earth's surface. The model using ResNeXt (Fig. 1) make it possible to more accurately model high-level features in large spaces, unlike [7], in which a model of Markov random fields is used to replace overcast pixels. This is confirmed by the better results of SSIM (10) and RMSE (8) metrics, especially in cases with dense cloud cover.

The proposed model produces fewer artifacts during image transformation compared to other models that process multi-temporal images. As a result of the combination of ERP data, the model ensures the removal of dense clouds in the images. Comparison of the results with existing models demonstrates the possibility of its application for image reconstruction during preprocessing of multispectral space images.

When applying in practice, as well as in further theoretical studies, the following main limitations must be taken into account:

1. In complex cases, when there is a dense cloud cover of a complex shape, the model does not provide accurate reconstruction of the image due to the sampling of the SEN12MS-CR data set for all latitudes of the Earth. The model may perform better in regions similar to those included in the training set, and worse under new, previously unknown conditions, such as other climate zones or specific ecosystems. To eliminate this shortcoming, it is necessary to retrain the constructed model on a data set corresponding to the space image processing area, or on data from a wider geographic and climatic spectrum to improve the universality of resulting models.

2. Images in the training data set were divided into smaller areas (256×256 pixels) with a spatial overlap of 50 %. This provides a larger training sample but may result in loss of context for large objects or phenomena in the image. In real-world scenarios where the context of large areas is important, the model may be less effective at removing cloud cover for large areas. In the future, we need to build models for removing clouds in optical space images that can work with larger image fragments or take into account the full context of the scene.

3. Sentinel-1 and Sentinel-2 data are used as a basis for model training, but these spacecraft have certain limitations in image quality and frequency. Further research into methods for integrating additional data sources or expanding the dataset could improve the quality of forecasting and cloud removal.

4. The model may show different results depending on which metric is prioritized. For example, in some tasks where the visual aspect is important, SSIM (10) may be key, while in tasks where precise numerical values are important, RMSE (8) may be more critical. What is needed is a method that would make it possible to balance the metrics or optimize the model depending on the priorities of the task, for example, adapting the loss functions to specific practical applications. In real-world applications, it is important to properly tune the model for specific scenarios – whether it is a clearer image restoration or a priority on minimizing uncertainty.

Our experiment has confirmed the effectiveness of the proposed neural network model for cloud removal in optical space images, showing improvement in image restoration quality. However, to achieve optimal results under real conditions, it is necessary to consider the limitations, adapt the model, and conduct further research to improve its performance.

Successful cloud removal is important for numerous applications such as environmental monitoring, agriculture, resource management, and natural disasters. The results of the research could prove useful for the implementation of more accurate satellite observations.

Prospects for further research are related to the improvement of the methodical apparatus for cloudiness removal from multispectral space images based on medium spatial resolution ERP data.

7. Conclusions

1. A neural network model for removing cloudiness in optical space images has been built, based on the architecture of a cyclic-sequential 7-ResNeXt block GAN with two output discriminators. Such an architecture makes it possible to effectively detect complex multi-level features of input data without significantly increasing the number of parameters. The use of GANs is due to its ability to generate realistic images without the need to model probability distributions to generate new data, as well as the ability to apply transfer learning. A CycleGAN-based loss function was used to train the model, which minimizes distortion while preserving the key features of the images. The model combines SAR and optical data, using their synergy to improve the reconstruction. Additional information from SAR images helps increase the reliability of reconstructed images.

2. The experimental results show that the proposed model effectively restores cloudy areas in Sentinel-2 satellite images, preserving information from cloud-free parts. This is confirmed not only by the visual evaluation of the RGB channels but also by the accurate pixel recovery of all multispectral channels with an average RMSE of 2.4 %. Metrics such as SSIM and *F*1 indicate an improvement in the reconstruction quality, in particular, the SSIM value increased by 0.005, indicating a higher visual similarity of the reconstructed images to the originals, especially in terms of textures and structures. To increase the efficiency of the model during training, SAR images from the C-band were used, which, owing to a longer wavelength, provide additional information about the geometry of the Earth's surface. The model generates fewer artifacts compared to other multi-temporal image processing approaches and effectively removes dense clouds, making it suitable for most national security tasks based on remote probing data.

Conflicts of interest

The authors declare that they have no conflicts of interest in relation to the current study, including financial, personal, authorship, or any other, that could affect the study, as well as the results reported in this paper.

Funding

The study was conducted without financial support.

Data availability

All data are available, either in numerical or graphical form, in the main text of the manuscript.

Use of artificial intelligence

The authors used artificial intelligence technologies within acceptable limits to ensure reliable verification of the SEN12MS-CR dataset, built by the model, which is described in the chapter on the research results related to the model for cloud removal in optical space images.

References

1. Rees, W. G. (2012). Physical Principles of Remote Sensing. Cambridge University Press https://doi.org/10.1017/cbo9781139017411
2. Shen, H., Li, X., Cheng, Q., Zeng, C., Yang, G., Li, H., Zhang, L. (2015). Missing Information Reconstruction of Remote Sensing Data: A Technical Review. IEEE Geoscience and Remote Sensing Magazine, 3 (3), 61–85. https://doi.org/10.1109/mgrs.2015.2441912
3. Xu, M., Jia, X., Pickering, M., Jia, S. (2019). Thin cloud removal from optical remote sensing images using the noise-adjusted principal components transform. ISPRS Journal of Photogrammetry and Remote Sensing, 149, 215–225. https://doi.org/10.1016/j.isprsjprs.2019.01.025
4. Ji, T.-Y., Yokoya, N., Zhu, X. X., Huang, T.-Z. (2018). Nonlocal Tensor Completion for Multitemporal Remotely Sensed Images' Inpainting. IEEE Transactions on Geoscience and Remote Sensing, 56 (6), 3047–3061. https://doi.org/10.1109/tgrs.2018.2790262
5. Li, X., Wang, L., Cheng, Q., Wu, P., Gan, W., Fang, L. (2019). Cloud removal in remote sensing images using nonnegative matrix factorization and error correction. ISPRS Journal of Photogrammetry and Remote Sensing, 148, 103–113. https://doi.org/10.1016/j.isprsjprs.2018.12.013

6. Meng, F., Yang, X., Zhou, C., Li, Z. (2017). A Sparse Dictionary Learning-Based Adaptive Patch Inpainting Method for Thick Clouds Removal from High-Spatial Resolution Remote Sensing Imagery. Sensors, 17 (9), 2130. https://doi.org/10.3390/s17092130
7. Cheng, Q., Shen, H., Zhang, L., Yuan, Q., Zeng, C. (2014). Cloud removal for remotely sensed images by similar pixel replacement guided with a spatio-temporal MRF model. ISPRS Journal of Photogrammetry and Remote Sensing, 92, 54–68. https://doi.org/10.1016/j.isprsjprs.2014.02.015
8. Eckardt, R., Berger, C., Thiel, C., Schmullius, C. (2013). Removal of Optically Thick Clouds from Multi-Spectral Satellite Images Using Multi-Frequency SAR Data. Remote Sensing, 5 (6), 2973–3006. https://doi.org/10.3390/rs5062973
9. Zhang, Q., Yuan, Q., Zeng, C., Li, X., Wei, Y. (2018). Missing Data Reconstruction in Remote Sensing Image With a Unified Spatial–Temporal–Spectral Deep Convolutional Neural Network. IEEE Transactions on Geoscience and Remote Sensing, 56 (8), 4274–4288. https://doi.org/10.1109/tgrs.2018.2810208
10. Isola, P., Zhu, J.-Y., Zhou, T., Efros, A. A. (2017). Image-to-Image Translation with Conditional Adversarial Networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). https://doi.org/10.1109/cvpr.2017.632
11. Zhang, X., Zhang, T., Wang, G., Zhu, P., Tang, X., Jia, X., Jiao, L. (2023). Remote Sensing Object Detection Meets Deep Learning: A metareview of challenges and advances. IEEE Geoscience and Remote Sensing Magazine, 11 (4), 8–44. https://doi.org/10.1109/mgrs.2023.3312347
12. He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). https://doi.org/10.1109/cvpr.2016.90
13. Meng, Q., Borders, B. E., Cieszewski, C. J., Madden, M. (2009). Closest Spectral Fit for Removing Clouds and Cloud Shadows. Photogrammetric Engineering & Remote Sensing, 75 (5), 569–576. https://doi.org/10.14358/pers.75.5.569
14. Schmitt, M., Zhu, X. X. (2016). Data Fusion and Remote Sensing: An ever-growing relationship. IEEE Geoscience and Remote Sensing Magazine, 4 (4), 6–23. https://doi.org/10.1109/mgrs.2016.2561021
15. Wang, L., Xu, X., Yu, Y., Yang, R., Gui, R., Xu, Z., Pu, F. (2019). SAR-to-Optical Image Translation Using Supervised Cycle-Consistent Adversarial Networks. IEEE Access, 7, 129136–129149. https://doi.org/10.1109/access.2019.2939649
16. Zhu, J.-Y., Park, T., Isola, P., Efros, A. A. (2017). Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. 2017 IEEE International Conference on Computer Vision (ICCV). https://doi.org/10.1109/iccv.2017.244
17. Mao, X., Li, Q., Xie, H., Lau, R. Y. K., Wang, Z., Smolley, S. P. (2017). Least Squares Generative Adversarial Networks. 2017 IEEE International Conference on Computer Vision (ICCV). https://doi.org/10.1109/iccv.2017.304
18. Goodfellow, I., Bengio, Y., Courville, A. (2016). Deep learning. MIT Press. Available at: https://www.deeplearningbook.org/
19. Ebel, P., Meraner, A., Schmitt, M., Zhu, X. X. (2021). Multisensor Data Fusion for Cloud Removal in Global and All-Season Sentinel-2 Imagery. IEEE Transactions on Geoscience and Remote Sensing, 59 (7), 5866–5878. https://doi.org/10.1109/tgrs.2020.3024744