DOI 10.1007/s10559-023-00535-9
Cybernetics and Systems Analysis, Vol. 58, No. 6, November, 2022

REVIEWING DEEP LEARNING METHODS IN THE APPLIED PROBLEMS OF ECONOMIC MONITORING BASED ON GEOSPATIAL DATA*

M. Lavreniuk,[1†] L. Shumilo,[2] B. Yailymov,[1‡] and N. Kussul[3]

UDC 004.8, 004.62

Abstract. *Development of modern observation technologies, increase of the amount of open data, and development of new approaches to their processing open new opportunities in carrying out applied research in the economic activity of people. The central approach in this field is the use of the deep learning methods in the data processing and analysis of their time series. In this paper, we review the basic (in terms of geospatial analysis) sections of deep learning, namely, increasing the resolution of graphical data, using transfer learning for optimization of learning processes, scaling deep neural network models, and analyzing time series using recurrent neural networks.*

Keywords: *deep learning, transfer learning, satellite data, geospatial data, recurrent neural networks.*

INTRODUCTION

The last years are characterized by the emergence of geospatial data in the open access, in particular, satellite data largely differing in time and space. Petabytes of satellite images became publicly available, and most algorithms of extracting information out of these images are widely used today, while the modern computing devices based on cloud technologies allow for these algorithms to be used on a global scale. As a result, today geospatial data is put to efficient use when solving a wide range of applied problems.

An example of a problem that can be solved based on the geospatial data is the economic monitoring of activities in regions by determining such indicators as economic activity and population poverty [1]. Most modern studies of this direction consider the agricultural economics or the economic assessment of the consequences of natural disasters or emergencies. For example, [2] carried out an economic assessment of the damages caused by a fire in Haifa (Israel) in 2016. The use of the satellite data provided by the company "Planet" made it possible to estimate the total area of damaged trees and their number (according to the estimate, the amount of damages in the urban forest was 41±10 mill. USD). An approach to assessing drought damage based on the technology of merging satellite data of different nature and different types of sensors is presented in [3]. The estimated economic damages in the Odessa oblast (region of Ukraine) from the death of grain crops due to adverse conditions in 2020 were about 338 thousand ha. of crops. Another problem to be solved based on the satellite data is the assessment of economic losses from floods. This information is very useful in the allocation of

* The research is conducted within the project №2020.02/0292 "The Methods and Models of Deep Learning for Applied Problems of Satellite Monitoring" funded by the National Research Foundation of Ukraine within the contest "Support research of leading and young scientists."

[1]Space Research Institute, National Academy of Sciences of Ukraine and State Space Agency of Ukraine, Kyiv, Ukraine, [†]*nick_93@ukr.net*; [‡]*yailymov@gmail.com*. [2]The University of Maryland, College Park, USA, *shumilo.leonid@gmail.com*. [3]National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine, *nataliia.kussul@gmail.com.*
Translated from Kibernetyka ta Systemnyi Analiz, No. 6, November–December, 2022, pp. 177–192. Original article submitted June 27, 2022.

1060-0396/22/5806-1008 ©2022 Springer Science+Business Media, LLC

resources for the recovery and reconstruction or in the preliminary economic assessment of the flood damages, which contributes to the crisis settlement after the disaster [4]. One of the most important aspects of applying these approaches to the analysis of satellite data is the possibility to conduct retrospective analysis of natural disasters for further risk assessment. The analysis of natural disasters in Hungary in 1998–2001 is presented in [5].

Today, geospatial data is an integral tool that makes it possible to predict and visualize emergencies for decision support systems used by the international disaster management community [6] and to carry out economic evaluation in this manner. Modern methods of the satellite data processing and analysis make it possible to estimate economic indicators of the gross domestic product [7], the impact of agricultural practices on the soil productivity [8] and land value [9] or the impact of a military conflict on land cover and land use [10]. The problem of determining and monitoring the dynamics of changes of indicators over time is relevant, and the solution to this problem is given in [11], but only in the case of a very specific problem of determining soil moisture. Therefore, it is necessary to adapt this approach to the classification of geospatial data and to ensure the possibility of applying deep learning methods to solve the problems of economic monitoring based on geospatial data.

This paper analyzes the existing methods of processing and analyzing satellite data that can be used to solve relevant applied problems of economic monitoring based on deep learning. The methods of working with satellite data described below allow us to solve a number of problems associated with the limitations of the classical approaches to the geospatial analysis and the Big Data problems, as well as making it possible to improve the accuracy and quality of the research results in the field of remote sensing of the Earth. The key technologies in the development of these approaches are the image resolution enhancement, which allows us to improve the quality of data at the stage of their pre-processing, the transfer learning, which allows us to optimize the process of training and scaling the deep learning models, and the classification of time series of the multivariate data.

ANALYZING THE EXISTING METHODS AND MODELS OF DEEP LEARNING TO SOLVE APPLIED PROBLEMS

In recent years, the intelligent methods of information processing and deep learning methods are actively developed to solve the applied problems based on large amounts of data. These methods are developed most intensely in the fields of computer vision and text information processing [12]. The modern methods of intellectual analysis of satellite data are based on the deep learning methods using convolutional or recurrent neural networks, allowing us to solve the problems of classification and regression of satellite data, as well as to carry out their semantic markup. An important challenge in this area is the need to develop methods for harmonizing satellite data of different nature (optical and radar data) with different spatial and temporal resolution, as well as the formation of high-resolution analysis ready data series. These mathematical methods are developed for other kinds of image data, namely, for ordinary photos [13, 14], texts [15], and sounds [16].

In recent years, a new type of neural networks was developed that has great prospects in solving these problems, namely, the generative adversarial networks (GANs) [17]. The problem of developing and adapting the GANs for working with satellite data while taking into account the specifics of satellite imagery is relevant. Here, recurrent neural networks with the long short-term memory (LSTM) show high efficiency, and they are already used to solve problems of satellite data classification [18].

When solving applied problems, the hybrid neural network models are often used. For example, the deep learning methods are applied in [19] to estimate the hourly global solar radiation (GSR) from a geostationary satellite where a hybrid deep network consisting of a convolutional neural network (CNN). a multilayer perceptron (MLP) for model connectivity, and additional time/location information are proposed to estimate the GSR.

One of the problems of solving applied problems based on satellite data is that most of it has low spatial resolution. For example, in [20], satellite products with a spatial resolution from 250 m to 1 km are used in order to detect smoke in a forest fire using a backpropagation neural network. The use of the low spatial resolution data is stipulated by the fact that it makes it possible to estimate how the smoke spreads during a fire due to its frequent updating and the large coverage of the territory that is photographed at a time. However, due to the low spatial resolution, it is difficult to detect active fires and their spread.

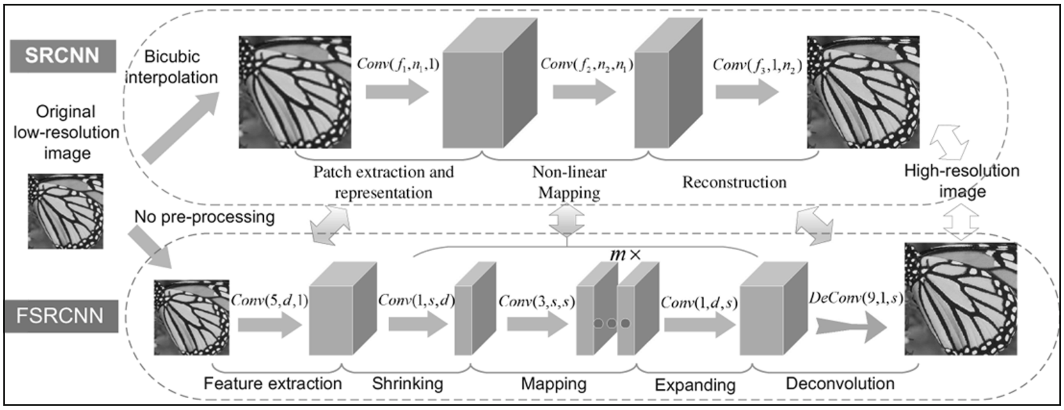

Fig. 1 Comparing the SRCNN and FSRCNN model architectures [28].

ANALYZING THE AVAILABLE METHODS TO IMPROVE THE RESOLUTION OF PICTURES AND IMAGES

The problem of increasing the spatial resolution of an image (of the super-resolution) arises in many application areas when processing both images and videos [21]. The spatial resolution is characterized by the density of pixels in an image and is measured in pixels per unit area [21]. The methods for the spatial resolution enhancement in images can either use multiple low spatial resolution images to create a single high-resolution image or work exclusively on a single input image with a low spatial resolution. The solution to the problem using multiple input images is described in detail in [21]. Let us consider the solution to the problem with only one input image.

The simplest method to increase the spatial resolution of an image is to apply the interpolation (the nearest neighbor method, bilinear interpolation, bicubic interpolation, splines, etc.). The disadvantage of this approach is that it does not consider the semantics of the image and the object contours. This can result in the image being very blurry. Another approach used to increase the spatial resolution of an image is based on the sparse coding [22]. It is assumed in most of the sparse methods that each pair of patches in low- and high-resolution images has the same coding coefficients in the patch space [23]. As a result, high resolution image patches can be represented by sparsely coded image patches with a low resolution. Thus, pairs of dictionaries of low- and high-resolution image patches are taught first. After that, a high-resolution image can be obtained using the taught dictionary of high-resolution image patches and the coding obtained using the taught dictionary of low-resolution image patches [23].

With the advent of the era of developing deep learning methods and models for solving various applied problems, the problem of improving image spatial resolution began to be solved using deep learning methods as well. The method of using convolutional neural networks to solve this problem was proposed in 2015 bearing the name of the Super-Resolution Convolutional Neural Network (SRCNN) [24]. However, this paper does not propose an “end-to-end” solution, but recommends first to enlarge the image to the required size using the standard bicubic interpolation method, and only then to apply the proposed neural network to improve the quality of the resulting image. The standard Peak Signal to Noise Ratio (PSNR) [25] was used to assess the performance of the model for improving image resolution.

An improvement of the SRCNN method was proposed in [26] by using the Very Deep Super Resolution (VDSR) model. The SRCNN model is used as a basis, but the convolutions with a large size (5×5 and 9×9) are replaced by a sequence of convolutions of the size 3×3 as it was proposed in the VGG model for image classification [27]. By contrast to the SRCNN model that had three layers, the VDSR model consisted of 20 layers. Moreover, the VDSR model tried to learn the residue between the target image and the interpolated image rather than learning the entire transitioning from one image to the other as in the case with the SRCNN model.

The problem of the previous two methods is the use of the interpolation at the initial stage as it leads to the use of a large number of parameters in the models, as well as making it impossible for the models to learn the parameters for the procedure of increasing image dimensionality. To solve these problems, the Fast Super-Resolution Convolutional Neural Networks (FSRCNN) model was proposed [28]. A low-resolution image was fed to the model input without any preliminary transformations (Fig. 1). Moreover, here, the 1×1 convolution is used in addition, and the activation function from the Rectified

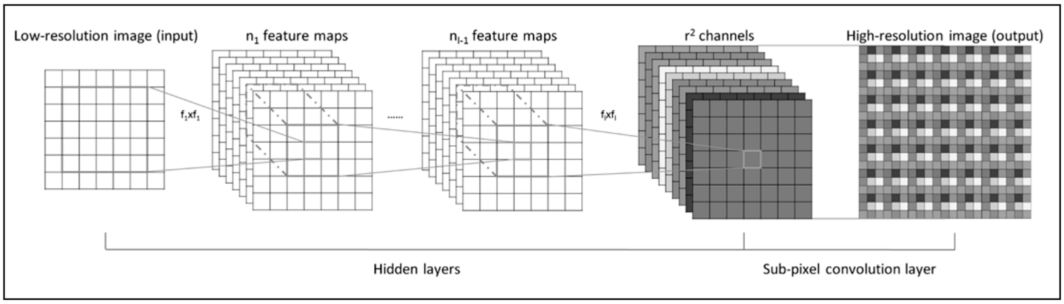

Fig. 2. An example of using subpixel convolutional layers in the ESPCN model [29].

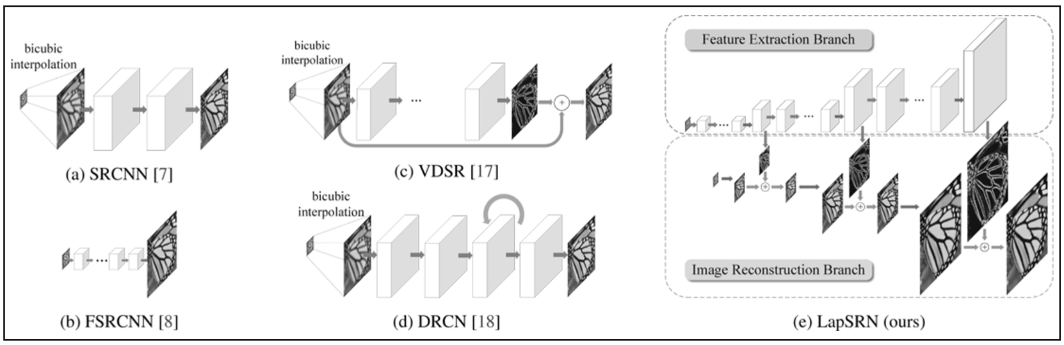

Fig. 3. The schematic SRCNN, FSRCNN, VDSR, DRCN, and LapSRN models [34].

Linear Unit (ReLU) is changed to the Parametric Rectified Linear Unit (PReLU). To train the model of enhancing image resolution, the deconvolution with layers was used, which allowed us to train the model using the "end-to-end" method.

A method that works on a similar principle is proposed in [29] and it is called the Efficient Sub-Pixel Convolutional Neural Network (ESPCN). The architecture of the network is based on the use of convolutional layers on a similar principle, but the deconvolution with layers is not used at the end. Instead, the subpixel convolutional layers (Fig. 2) are used. This idea is similar to the use of the position-sensitive score maps when detecting objects [30].

The Enhanced Deep Super-Resolution (EDSR) network and SRResNet models [17] use the idea of the ResNet to improve the image resolution. However, in the SRResNet, the nonlinear layer with the ReLU activation is removed, while the batch normalization layers are removed in the EDSR. It is experimentally shown in [17] that these modifications allow us to obtain better images at the model output and help us to reduce the amount of memory used for model learning.

The Cascading Residual Network (CARN) model based on the residual blocks is proposed in [31]. In [32], the Deep Recursive Residual Network (DRRN) model is proposed, which is a combination of two ideas where the first idea is to use the residual blocks, while the other is to use the Deep-Recursive Convolutional Network (DRCN) [33], which is based on applying one and the same convolution in an iterative manner to improve the result. As a result, the DRRN consists of 52 convolutional layers that are taught in the "end-to-end" manner.

The problem of the above methods is that they can improve the image resolution by 2–4 times. When it comes to improving it by eight times and more, the considered models produce results of not such a high quality. To solve this problem, the Laplacian Pyramid Super-Resolution Network (LapSRN) is proposed [34]. The idea of the model is to gradually increase the image using a pyramid (Fig. 3). The loss function is the sum of the loss functions at each level of the pyramid. To calculate the loss function at each level of the pyramid, the bicubic interpolation is applied here to reduce the resolution of the image to be reproduced. The Charbonnier Loss is taken as a more resistant to emissions loss function according to the authors [34].

All the above methods work the same in any part of the image. However, in general, a model should focus on some parts of the image more than on the others. For this purpose, the mechanism of attention is used in deep learning [25]. To implement this idea, the Squeeze-and-Excitation (SE) block is proposed in [35] that allows the neural network to pay more attention to important features and less attention to less important ones.

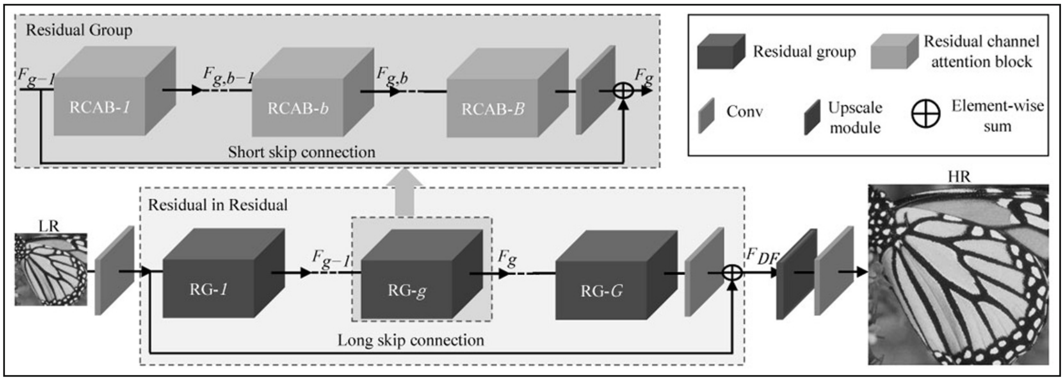

Fig. 4. The architecture of the Residual Channel Attention Networks model [36].

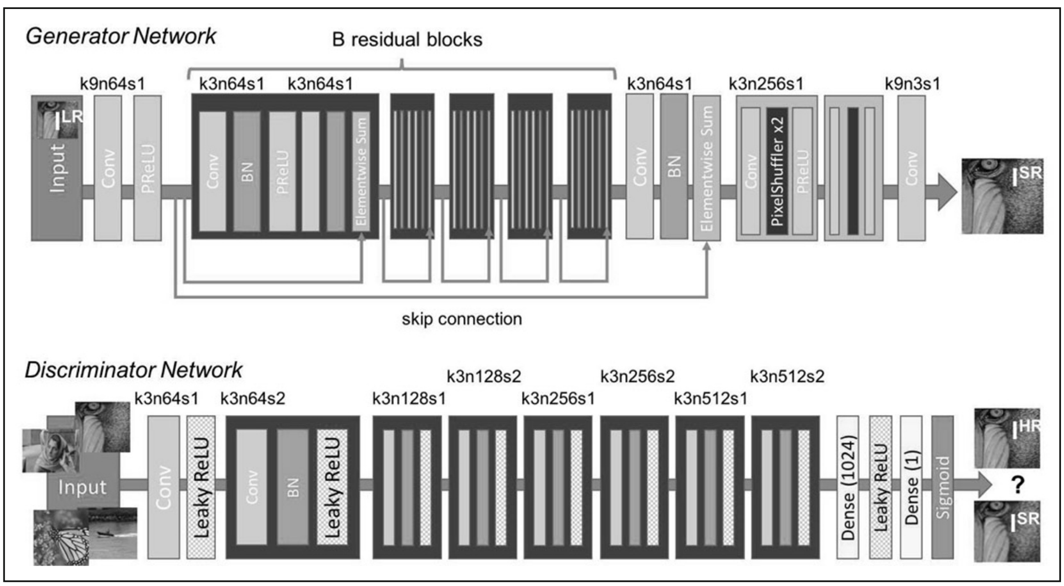

Fig. 5. The generator and discriminator architectures in the SRGAN model [38].

The Residual Channel Attention Networks (RCAN) model, which uses blocks with residual long and short links, as well as the Residual Channel Attention Blocks (RCABs), which are analogous to the SE blocks, are proposed in [36]. The general architecture of this network is presented in Fig. 4. In the model, the L1 loss function is used to compare it to the above models during learning.

An alternative approach to improving image resolution is to use Generative Adversarial Networks (GANs) [37], which consist of two neural networks, namely, a generator and a discriminator, and are learning by the rule of learning without a teacher. This approach was first proposed in [38] under the name Super Resolution Generative Adversarial Networks (SRGANs). The idea of the model is as follows: one neural network (a generator) generates a higher-resolution image, while another neural network (a discriminator) tries to determine whether this image is real or generated by the network. The learning of the models is completed when the first network learns to generate "real" higher-resolution images that the other network cannot distinguish from the real ones. The discriminator and generator architectures of the SRGAN model are based on the ResNet network and are shown in Fig. 5. Here, to improve the results, it is also proposed to replace the standard MSE loss function with a "perceptual loss" function [39], which is computed based on the features obtained by a pre-taught VGG network. To train the GAN as an adversarial loss function, the standard loss function, i.e., the binary cross-entropy [38], is used.

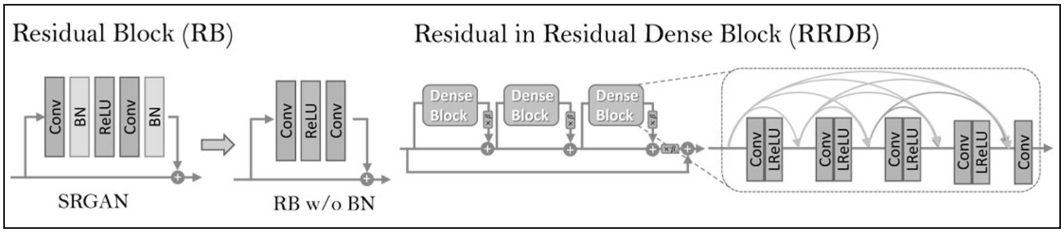

Fig. 6. An example of basic blocks, namely, the residual block (on the left-hand side) and the RRDB block (on the right-hand side) [43].

In [40], the idea of using the GAN to improve the resolution is modified by using the CycleGAN [41], which is based on training a model to transform a low-resolution image into a high-resolution image, but without the noise and blur, and vice versa. The proposed method is called the Cycle-in-Cycle Generative Adversarial Networks (CinCGANs). Moreover, the standard adversarial loss function (the binary cross entropy) is replaced by the MSE [42], which allows us to obtain clearer results. The penalty functions of cycle consistency loss and of identity loss are added in the form of L2 and L1 norms, respectively. During model learning, the total loss function is calculated as the sum of all proposed loss functions with certain coefficients chosen empirically.

In [43], the Enhanced Super-Resolution Generative Adversarial Networks (ESRGANs) model based on the SRResNet model is proposed, but the basic residual blocks are replaced by the Residual in Residual Dense Block (RRDB) (Fig. 6). The adversarial loss function is also modified to not only predict whether the input images are real or fake, but also to predict the probability of the real images being relatively more realistic than the fake ones.

Taking into account all the above methods, it can be concluded that the most feasible methods now are those that are based on the idea of using the GAN to improve image resolution. In particular, this idea is not only used in the case of ordinary photos, but also when solving problems with other data. For example, in [44], a model based on the SRGAN is proposed to improve the resolution of the Digital Elevation Models (DEMs) by a factor of four. In [45], the GAN is used to improve the resolution of weather radar data, which allows us to obtain a more accurate weather forecasting. Several scientific papers are devoted to enhancing the multispectral data, satellite data in particular. In [46], the ESRGAN is used to increase the spatial resolution of the Sentinel-2 images to the spatial resolution of the WorldView images, which allows us to increase the spatial resolution by a factor of 5. The ESRGAN is used to improve the resolution of the images from unmanned aerial vehicles [47]. The Edge-Enhancement Generative Adversarial Networks (EEGANs) method is proposed to improve the resolution of the Jilin-1 and Digitalglobe images [48]. The modified SRGAN method with replacing the 2D convolutional layers by the 3D convolutional layers and adding the mechanism of attention is proposed to improve the resolution of the multispectral data [49]. Additional examples of applying the GAN models and deep learning models to the data obtained by the remote Earth sensing methods are given in [50].

ANALYZING THE AVAILABLE METHODS FOR TRANSFER LEARNING IN THE REMOTE SENSING OF THE EARTH

The transfer learning technology is a section in machine learning where the main problem is to save and transfer the received information and artificial intelligence models to another or related problem. When applying the transfer learning approach to the remote Earth sensing data, the main problem is to transfer the artificial intelligence models taught on a dataset for a given territory and a given time period to other territories or time periods. In this problem statement, the above problems are not trivial as even the same territory present in the satellite data is significantly different each year in terms of multispectral indicators. This data is usually obtained at different dates of the year and under different weather conditions, which affect both the data availability and the illumination of the ground surface during recording.

Formally, the transfer learning problem can be described as follows [51] using mathematics: the domain $D = \{X, P(X)\}$ consists of the space of features X and the marginal probability distribution $P(X)$ where the domain $T = \{Y, f(\cdot)\}$ consists of the space of labels Y and the predictive function $f(\cdot)$ that models $P(y|x)$ where $y \in Y$ and

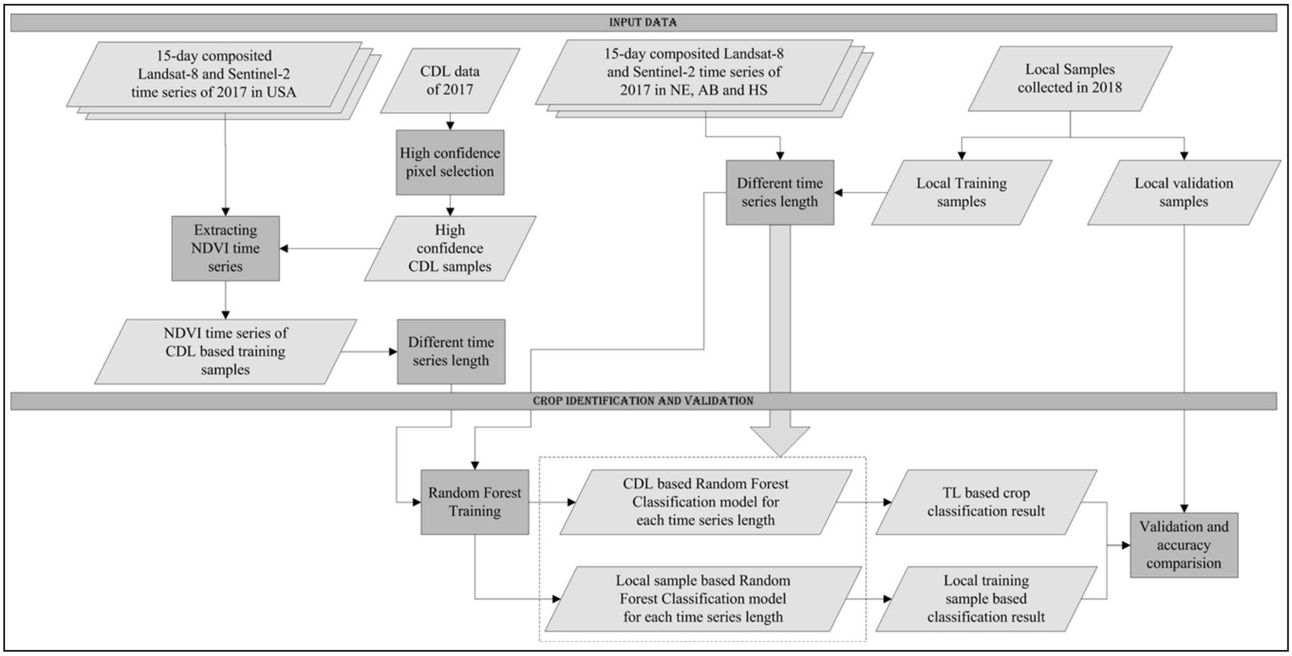

Fig 7. The scheme training the random forest model on the Crop Layer Data for transferring it onto other countries [54].

$x \in X$. With the initial domain D_s and the learning problem T_s together with the target domain D_t and the target learning problem T_t, let us form an approach focused on improving the learning process of the target predictive function $ft(\cdot)$ that belongs to T_t by using the knowledge and the information from D_s and T_t where $D_s \neq D_t$ and/or $T_s \neq T_t$. This problem is especially relevant when $|Dt| \leq |Ds|$.

An interesting example of the use of the transfer learning approach on the remote Earth sensing data is the transfer of information from the ImageNet object recognition problem by using the high-resolution spatial satellite images acquired during the day to the problem of predicting the intensity of night illumination for mapping poverty indicators and the compliance with the sustainable development goals. This study was carried out by a team from the Stanford University [52]. When carrying out the study, a convolutional neural network model with the VGG architecture was used. To map poverty indicators, the transfer learning graph in the form of a linear chain with three transfer learning problems was constructed where P_1 is the problem of object recognition on the ImageNet data, P_2 is the problem of predicting the night illumination being present on the satellite data, and P_3 is the problem of predicting the poverty indicators based on the obtained satellite data.

In [53], a very relevant example of using the deep learning and transfer learning methods for solving the problem of crop yield forecasting using the remote Earth sensing data is given. The problem is that there are very few crop yield statistics available for the developing countries. At the same time, the crop yield forecasting is one of the keys to the sustainable agricultural development and food security for these countries. During the experiment, the MODIS satellite data and the soybean yield statistics were used for Argentina and Brazil at the levels of country and province. To train the deep learning model, the MODIS images containing multispectral and temperature channels are presented as separate histograms. Then, all histograms are grouped in the form of a matrix that is transformed into a tensor, which is used at the input of the LSTM model.

During the experiment, the authors of [53] taught the crop yield forecasting model first for Argentina. Next, the dense layer of the taught model was replaced with a new untaught layer and the model was retaught using a small amount of data available for Brazil. Thus, the transfer learning approach was applied.

A classic experiment showing how the transfer learning works is described in [54]. Here, the Crop Data Layer [55] that is produced annually by the United States Department of Agriculture was used as the training ground data for the crop classification that uses 15-day composites of the harmonized Sentinel-2 and Landsat-8 data collections. The scheme of this experiment is presented in Fig. 7. The authors trained the random forest model classifier on the territory of the United States and transferred it onto the territories of China and Canada, obtaining a high classification accuracy of 97.79% and

94.86% for the territories of America and Canada, respectively, and 86.45% for the territory of China. It is obvious that the greater the differences in climate, agricultural practices, landscape, and other natural and anthropomorphic characteristics of the area, the lower the classification accuracy of the resulting model. However, this problem can be reduced by increasing the amount of useful information, for example, by providing information for several years that covers the full growing season. When transferring the model, the number of classes should be considered as well. In this experiment, the so-called corn belt of America was chosen for training, i.e., the area where corn, winter wheat, and cotton are the majority crops. Thus, the model recognized only three crops and caould not work correctly in areas with a greater crop variety.

The approach is efficient for a more complex visual data as well. For example, this is shown in [56] where an experiment of the road quality assessment in Nigeria based on very high spatial resolution of the satellite data is described. However, the main problem of this study is the lack of the sufficient amount of marked data for the area of interest. Therefore, the authors of the paper applied the transfer learning approach by training a convolutional neural network model on 53686 images covering 2400 kilometers of roads in the United States and adjusted and tested the parameters on 1000 marked images in Nigeria.

ANALYZING THE EXISTING APPROACHES TO THE USE OF THE RECURRENT NEURAL NETWORKS WORKING WITH TEMPORAL DATA

In the case of problems with satellite data, it is often necessary to obtain models that provide a vector of outputs for individual areas or time series. This is especially common in the case of problems that classify agricultural fields in countries with warm climates or developed agricultural practices where there is double cropping for one season. The simplest approach to solving this classification problem with the classical machine or deep learning methods is coding the double crops by introducing new classes for the classification map and by applying the usual stages of the classification map construction. However, these methods are not very efficient as these classification maps have a rather low accuracy in practice since the number of classes is very high and the amount of ground data is limited. The second and a more common way to solve this problem, which is now being actively developed and improved by modern scientists, is the use of the modern neural network architectures that are optimized for working with time series where each time interval is given a corresponding class label. As a result of this approach, the number of classes does not increase and it is possible to obtain more than one classification map for an agricultural season by using and training the model only once. The second important problem is to create deep learning models that allow us to classify the long-term satellite data and to provide information on the type of land cover or crops for each year. The most common type of the deep learning models solving these problems are the feedback neural networks.

A classic example of the use of the recurrent neural network architecture to solve the problem of the long-term land cover classification based on the satellite data is given in [57] where the classification of the land cover from 1982 to 2015 was performed by using the satellite data provided by the CNLUCC land cover map set for China as the training data. The model had 10 land cover classes, such as farmland, forest, shrubland, grassland, water, glaciers, marshes, cities, desert, and tundra for seven years, namely, 1980, 1990, 1995, 2000, 2005, 2010, and 2015. The architecture based on the two-way neural network with a short-term and long-term memory shown in Fig. 8 is used as the classification model. The neural network consists of the two conventional LSTM neural networks with their time vectors directed in the opposite directions and being called the forward and the backward LSTMs. In this study, the ability of the LSTM models to generate a vector of outputs is used to create a set of temporal features that are pooled and passed through the Mean Pooling and normalized exponential function (Softmax) layers to produce a probabilistic output from which the most probable class is selected.

A detailed description on how to use the recurrent neural networks and the long- and short-term memory (LSTM) neural network to classify the time series of the satellite data to obtain crop maps is presented in [18]. Here, an important feature of the recurrent neural networks is that each neuron situated at the output that receives its time interval and the information from past neuron outputs as the input, i.e., the information from the past time intervals, produces an output based on the normalized exponential function (Softmax) and is transformed into class probabilities for the classification map. Thus, a classification map is generated that corresponds to the corresponding input time interval. Figure 9 shows a diagram of the LSTM neural network model in action. Using this approach, it is possible to obtain several crop classification maps for one season.

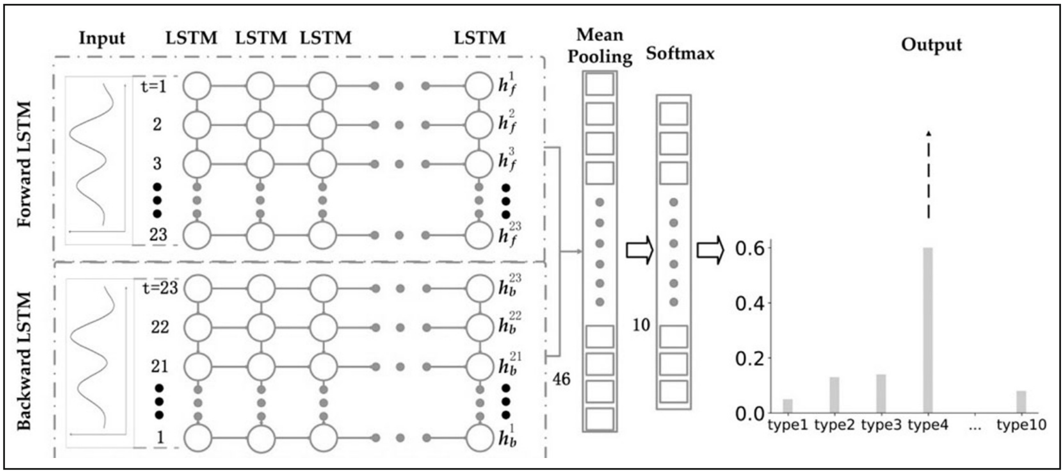

Fig. 8. The architecture of the two-way neural network with a long-term and short-term memory for the land cover classification [57].

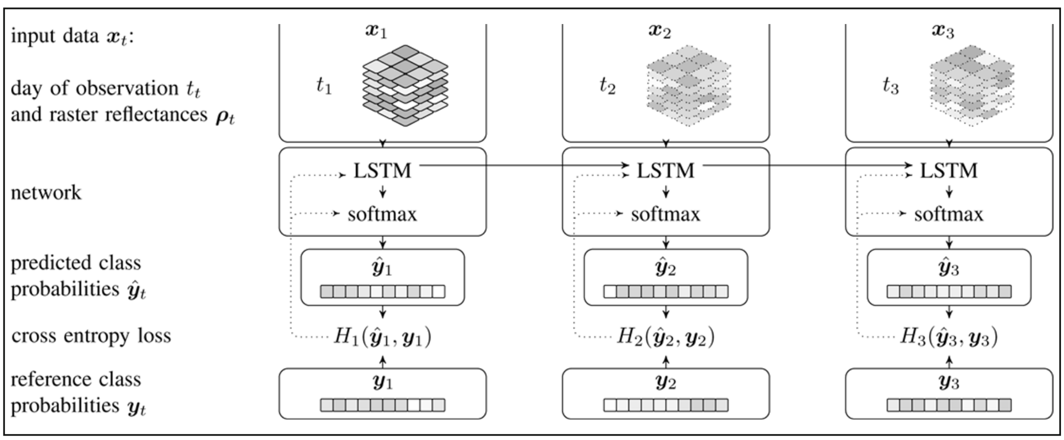

Fig. 9. The scheme of the LSTM neural network with outputs at each time interval in action [18].

A similar example of applying the LSTM model on the time series of the satellite images, but for the problem of obtaining land cover classification maps for several years using a single neural network, was published in [58]. Here, the Landsat and CDL mission data for 2013 to 2016 is used to train the neural network with the multi-year satellite data input to produce land cover classification maps for the corresponding years. Figure 10 shows the scheme of the neural network architecture in action. It is evident that this principle is similar to the principle of the neural network from [18], but it differs in the length of its time series, which is four years.

In addition to the advantages in the construction of the highly accurate classification maps for the time series of the satellite data, the LSTM architectures have a great advantage over the classical methods when it comes to the prediction problems. The recurrent neural networks working in this direction have already demonstrated the highest quality of the produced results and make it possible to efficiently predict crop yields on a large time series of the heterogeneous data at the level of districts, countries, and even continents [59]. Since this approach has already demonstrated its efficiency and prospects, scientists are researching and developing more complex neural network architectures. In [60], the use of the attention layers in the LSTM architecture is proposed to predict the yield of tomatoes. Thus, this architecture makes it possible to improve the properties of the classical recurrent neural networks by storing the information with a high level of importance in the space of exogenous variables that affect the value of the dependent variables in the long-term memory.

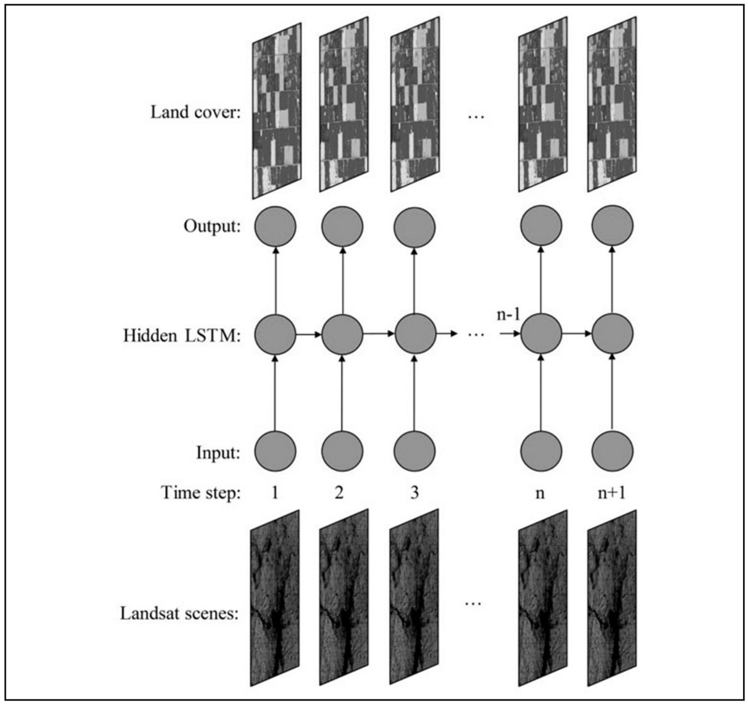

Fig. 10. The scheme of the neural network for obtaining land cover classification maps for several years in action [58].

When recurrent neural networks are working with the data in the form of multidimensional matrices or images, they can be combined with the convolutional neural networks, i.e., hybrid architectures can be created. In [61], it is proposed to use the DeepYield architecture to predict the crop yields by the satellite data at the country level, which is based on the use of the so-called ConvLSTM layers. The neural network layers that combine the convolutional filters and recurrent neurons make it possible to produce more complex features reflecting not only the change of the multispectral characteristics of pixels over time, but also the dependence between these changes in the neighboring pixels, as well as taking into account the texture of crop yields and the terrain.

CONCLUSIONS

The paper provides an overview of the modern technologies for Earth observation based on the geospatial data. Here, the main directions and fields of the applied problems of data processing and analysis of its time series are considered in detail, namely, the increase of the graphical data resolution, the use of the transfer learning approach to optimize the learning processes and scaling of the deep neural network models, and the analysis of the time series using the recurrent neural networks. The vast majority of the considered methods are based on the mathematical methods of machine learning, the deep learning in particular. Thus, this paper describes the integration of the modern deep learning methods for solving applied problems using the available geospatial data in detail, the high-resolution satellite data in particular.

REFERENCES

1. T. P. Sohnesen, P. Fisker, and D. Malmgren-Hansen, “Using satellite data to guide urban poverty reduction,” Review of Income and Wealth, Vol. 68, No. S2, S282–S294 (2021). https://doi.org/10.1111/roiw.12552.
2. Y. Michael, I. M. Lensky, S. Brenner, A. Tchetchik, N. Tessler, and D. Helman, “Economic assessment of fire damage to urban forest in the wildland-urban interface using planet satellites constellation images,” Remote Sens., Vol. 10, No. 9, 1479 (2018). https://doi.org/10.3390/rs10091479.

3. Y. Bilokonska, H. Yailymova, B. Yailymov, A. Shelestov, L. Shumilo, and M. Lavreniuk, “Losses assessment for winter crops based on satellite data and fuzzy logic,” in: 2020 IEEE 5th Intern. Symp. on Smart and Wireless Systems within the Conferences on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS) (Dortmund, Germany, September 17–18, 2020), IEEE (2020), pp. 1–5. https://doi.org/10.1109/IDAACS-SWS50031.2020.9297059.
4. B. Merz, H. Kreibich, R. Schwarze, and A. Thieken, “ Review article “Assessment of economic flood damage,” Nat. Hazards Earth Syst. Sci., Vol. 10, Iss. 8, 1697–1724 (2010). https://doi.org/10.5194/nhess-10-1697-2010.
5. M. Lelkes, G. Csornai, and C. Wirnhardt, “Natural disaster monitoring by remote sensing in Hungary: Waterlogging and floods in the 1998–2001 period,” in: G. Begni (ed.), Observing Our Environment from Space — New Solutions for a New Millennium, CRC Press, London (2002), pp. 259–264.
6. D. M. Tralli, R. G. Blom, V. Zlotnicki, A. Donnellan, and D. L. Evans, “Satellite remote sensing of earthquake, volcano, flood, landslide and coastal inundation hazards,” ISPRS J. Photogramm. Remote Sens., Vol. 59, Iss. 4, 185–198 (2005). https://doi.org/10.1016/j.isprsjprs.2005.02.002.
7. L. R. Martinez, “How much should we trust the dictator’s GDP growth estimates?,” SSRN eJournal (2019). http://doi.org/10.2139/ssrn.3093296.
8. N. Kussul, K. Deininger, L. Shumilo, M. Lavreniuk, D. A. Ali, and O. Nivievskyi, “Biophysical impact of sunflower crop rotation on agricultural fields,” Sustainability, Vol. 14, No. 7, 3965 (2022). https://doi.org/10.3390/su14073965.
9. L. Shumilo, M. Lavreniuk, S. Skakun, and N. Kussul, “Is soil bonitet an adequate indicator for agricultural land appraisal in Ukraine?,” Sustainability, Vol. 13, No. 21, 12096 (2021). https://doi.org/10.3390/su132112096.
10. S. Skakun, C. J. Justice, N. Kussul, A. Shelestov, and M. Lavreniuk, “Satellite data reveal cropland losses in South-Eastern Ukraine under military conflict,” Front. Earth Sci., Vol. 7, 305 (2019). https://doi.org/10.3389/feart.2019.00305.
11. N. Efremova, D. Zausaev, and G. Antipov, “Prediction of soil moisture content based on satellite data and sequence-to-sequence networks,” arXiv:1907.03697v1 [eess.IV] 5 Jun (2019). https://doi.org/10.48550/arXiv.1907.03697.
12. M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and Prabhat, “Deep learning and process understanding for data-driven Earth system science,” Nature, Vol. 566 (7743), 195–204 (2019). https://doi.org/10.1038/s41586-019-0912-1.
13. H. Chen, X. He, L. Qing, Y. Wu, C. Ren, R. E. Sheriff, and C. Zhu, “Real-world single image super-resolution: A brief review,” Inf. Fusion, Vol. 79, 124–145 (2022). https://doi.org/10.1016/j.inffus.2021.09.005.
14. F. Bellavia, M. Fanfani, C. Colombo, and A. Piva, “Experiencing with electronic image stabilization and PRNU through scene content image registration,” Pattern Recognit. Lett., Vol. 145, 8–15 (2021). https://doi.org/10.1016/j.patrec.2021.01.014.
15. M. Nadif and F. Role, “Unsupervised and self-supervised deep learning approaches for biomedical text mining,” Briefings in Bioinformatics, Vol. 22, Iss. 2, 1592–1603 (2021). https://doi.org/10.1093/bib/bbab016.
16. M. Chaiani, S. A. Selouani, M. Boudraa, and M. S. Yakoub, “Voice disorder classification using speech enhancement and deep learning models,” Biocybern. Biomed. Eng., Vol. 42, Iss. 2, 463–480 (2022). https://doi.org/10.1016/j.bbe.2022.03.002.
17. B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, “Enhanced deep residual networks for single image super-resolution,” in: Proc. 2017 IEEE Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW) (Honolulu, HI, USA, July 21–26, 2017), IEEE (2017), pp. 1132–1140. https://doi.org/10.1109/CVPRW.2017.151.
18. M. Ruβwurm and M. Körner, “ Temporal vegetation modelling using long short-term memory networks for crop identification from medium-resolution multi-spectral satellite images,” in: Proc. 2017 IEEE Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW) (Honolulu, HI, USA, July 21–26, 2017), IEEE (2017), pp. 1496–1504. https://doi.org/10.1109/CVPRW.2017.193.
19. H. Jiang, N. Lu, J. Qin, W. Tang, and L. Yao, “A deep learning algorithm to estimate hourly global solar radiation from geostationary satellite data,” Renew. Sust. Energ. Rev., Vol. 114, 109327 (2019). https://doi.org/10.1016/j.rser.2019.109327.
20. X. Li, W. Song, L. Lian, and X. Wei, “Forest fire smoke detection using back-propagation neural network based on MODIS data,” Remote Sens., Vol. 7, No. 4, 4473–4498 (2015). https://doi.org/10.3390/rs70404473.
21. J. Yang and T. Huang, “Image super-resolution: Historical overview and future challenges,” in: P. Milanfar (ed.), Super-Resolution Imaging, CRC Press, Boca Raton (2011), pp. 3–34. https://doi.org/10.1201/9781439819319.
22. J. Zhang, Z. Wang, Y. Zheng, and G. Zhang, “Cascade convolutional neural network for image super-resolution,” arXiv:2008.10329 [cs.CV] 24 Aug (2020). https://doi.org/10.48550/arXiv.2008.10329.
23. J. Yang, J. Wright, T. S. Huang, and Y. Ma, “Image super-resolution via sparse representation,” IEEE Trans. Image Process., Vol. 19, Iss. 11, 2861–2873 (2010). https://doi.org/10.1109/TIP.2010.2050625.

24. C. Dong, C. C. Loy, K. He, and X. Tang, “Image super-resolution using deep convolutional networks,” IEEE Trans. Pattern Anal. Mach. Intell., Vol. 38, Iss. 2, 295–307 (2016). https://doi.org/10.1109/TPAMI.2015.2439281.
25. W. Zhihao, J. Chen, and S. C. H. Hoi, “Deep learning for image super-resolution: A survey,” IEEE Trans. Pattern Anal. Mach. Intell., Vol. 43, Iss. 10, 3365–3387 (2021). https://doi.org/10.1109/TPAMI.2020.2982166.
26. J. Kim, J. K. Lee, and K. M. Lee, “Accurate image super-resolution using very deep convolutional networks,” in: Proc. 2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) (Las Vegas, NV, USA, June 27–30, 2016), IEEE (2016), pp. 1646–1654. https://doi.org/10.1109/CVPR.2016.182.
27. K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition arXiv:1409.1556 [cs.CV] 4 Sep (2014). https://doi.org/10.48550/arXiv.1409.1556.
28. C. Dong, C. C. Loy, X. Tang, “Accelerating the super-resolution convolutional neural network,” in: B. Leibe, J. Matas, N. Sebe, and M. Welling (eds.), Computer Vision — ECCV 2016, ECCV 2016; Lecture Notes in Computer Science, Vol. 9906, Springer, Cham (2016), pp. 391–407. https://doi.org/10.1007/978-3-319-46475-6_25.
29. W. Shi, J. Caballero, F. Huszár, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, “Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network,” in: Proc. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (Las Vegas, NV, USA, June 27–30, 2016), IEEE (2016), pp. 1874–1883. https://doi.org/10.1109/CVPR.2016.207.
30. J. Dai, Y. Li, K. He, and J. Sun, “R-FCN: Object detection via region-based fully convolutional networks,” in: NIPS’16: Proc. of the 30th Intern. Conf. on Neural Information Processing Systems (Barcelona, Spain, December 5–10, 2016), Vol. 1 (2016), pp. 379–387.
31. N. Ahn, B. Kang, and K. A. Sohn, “Fast, accurate, and lightweight super-resolution with cascading residual network,” in: V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss (eds.), Computer Vision — ECCV 2018, ECCV 2018; Lecture Notes in Computer Science, Vol. 11214, Springer, Cham (2018), pp. 256–272. https://doi.org/10.1007/978-3-030-01249-6_16.
32. Y. Tai, J. Yang, and X. Liu, “Image super-resolution via deep recursive residual network,” in: Proc. 2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) (Honolulu, HI, USA, July 21–26, 2017), IEEE (2017), pp. 2790–2798. https://doi.org/10.1109/CVPR.2017.298.
33. J. Kim, J. Kwon Lee, K. Mu Lee, “Deeply-recursive convolutional network for image super-resolution,” in: Proc. 2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) (Las Vegas, NV, USA, June 27–30, 2016), IEEE (2016), pp. 1637–1645.
34. W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, “Deep Laplacian pyramid networks for fast and accurate super-resolution,” in: Proc. 2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) (Honolulu, HI, USA, July 21–26, 2017), IEEE (2017), pp. 624–632.
35. J. Hu, L. Shen, and G. Sun, “Squeeze-and-excitation networks,” in: Proc. 2018 IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR) (Salt Lake City, UT, USA, June 18–23, 2018), IEEE (2018), pp. 7132–7141.
36. Y. Zhang, Kunpeng Li, Kai Li, L. Wang, B. Zhong, and Y. Fu, “Image super-resolution using very deep residual channel attention networks,” in: V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss (eds.), Computer Vision — ECCV 2018, ECCV 2018; Lecture Notes in Computer Science, Vol. 11211, Springer, Cham (2018), pp. 294–310. https://doi.org/10.1007/978-3-030-01234-2_18.
37. I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in: Proc. of the 27th Intern. Conf. on Neural Information Processing Systems (Montreal, Canada, December 8–13, 2014), MIT Press (2014). pp. 2672–2680.
38. C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, “Photo-realistic single image super-resolution using a generative adversarial network,” in: Proc. 2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) (Honolulu, HI, USA, July 21–26, 2017), IEEE (2017), pp. 105–114. https://doi.org/10.1109/CVPR.2017.19..
39. J. Johnson, A. Alahi, and L. Fei-Fei, “Perceptual losses for real-time style transfer and super-resolution,” in: B. Leibe, J. Matas, N. Sebe, and M. Welling (eds.), Computer Vision — ECCV 2016, ECCV 2016; Lecture Notes in Computer Science, Vol. 9906, Springer, Cham (2016). pp. 694–711. https://doi.org/10.1007/978-3-319-46475-6_43.
40. Y. Yuan, S. Liu, J. Zhang, Y. Zhang, C. Dong, and L. Lin, “Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks,” in: Proc. 2018 IEEE/CVF Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW) (Salt Lake City, UT, USA, June 18–22, 2018), Computer Vision Foundation / IEEE Computer Society (2018), pp. 814–823. https://doi.org/10.1109/CVPRW.2018.00113.

41. J.-Y. Zhu, T. Park, P. Isola and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in: Proc. 2017 IEEE Intern. Conf. on Computer Vision (ICCV) (Venice, Italy, October 22–29, 2017), IEEE (2017), pp. 2242–2251. https://doi.org/10.1109/ICCV.2017.244.
42. X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang and S. P. Smolley, "Least squares generative adversarial networks," in: Proc. 2017 IEEE Intern. Conf. on Computer Vision (ICCV) (Venice, Italy, October 22–29, 2017), IEEE (2017), pp. 2813–2831. https://doi.org/10.1109/ICCV.2017.304.
43. X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. C. Loy, "ESRGAN: Enhanced super-resolution generative adversarial networks," in: L. Leal-Taixé, S. Roth (eds.), Computer Vision — ECCV 2018 Workshops, ECCV 2018; Lecture Notes in Computer Science, Vol. 11133, Springer, Cham (2019), pp. 63–79. https://doi.org/10.1007/978-3-030-11021-5_5.
44. B. Z. Demiray, M. Sit, and I. Demir, "D-SRGAN: DEM super-resolution with generative adversarial network," arXiv:2004.04788 [cs.CV] 9 Apr (2020). https://doi.org/10.48550/arXiv.2004.04788.
45. H. Chen, X. Zhang, Y. Liu, and Q. Zeng, "Generative adversarial networks capabilities for super-resolution reconstruction of weather radar echo images," Atmosphere, Vol. 10, Iss. 9, 555 (2019). https://doi.org/10.3390/atmos10090555.
46. L. Salgueiro Romero, J. Marcello, and V. Vilaplana, "Super-resolution of Sentinel-2 imagery using generative adversarial networks," Remote Sens., Vol. 12, No. 15, 2424 (2020). https://doi.org/10.3390/rs12152424.
47. M. Pashaei, M. J. Starek, H. Kamangir, and J. Berryhill, "Deep learning-based single image super-resolution: An investigation for dense scene reconstruction with UAS photogrammetry," Remote Sens., Vol. 12, No. 11, 1757 (2020). https://doi.org/10.3390/rs12111757.
48. K. Jiang, Z. Wang, P. Yi, G. Wang, T. Lu, and J. Jiang, "Edge-enhanced GAN for remote sensing image superresolution," IEEE Trans. Geosci. Remote Sens., Vol. 57, No. 8, 5799–5812 (2019). https://doi.org/10.1109/TGRS.2019.2902431.
49. X. Dou, C. Li, Q. Shi, and M. Liu, "Super-resolution for hyperspectral remote sensing images based on the 3D attention-SRGAN network," Remote Sens., Vol. 12, No. 7, 1204 (2020). https://doi.org/10.3390/rs12071204.
50. G. Tsagkatakis, A. Aidini, K. Fotiadou, M. Giannopoulos, A. Pentari, and P. Tsakalides, "Survey of deep-learning approaches for remote sensing observation enhancement," Sensors, Vol. 19, No. 18, 3929 (2019). https://doi.org/10.3390/s19183929.
51. S. J. Pan and Q. Yang, "A survey on transfer learning," IEEE Trans. Knowl. Data Eng., Vol. 22, No. 10, 1345–1359 (2009). https://doi.org/10.1109/TKDE.2009.191.
52. M. Xie, N. Jean, M. Burke, D. Lobell, and S. Ermon, "Transfer learning from deep features for remote sensing and poverty mapping," arXiv:1510.00098 [cs.CV] 1 Oct (2015). https://doi.org/10.48550/arXiv.1510.00098.
53. A. X. Wang, C. Tran, N. Desai, D. Lobell, and S. Ermon, "Deep transfer learning for crop yield prediction with remote sensing data," in: COMPASS'18: Proc. of the 1st ACM SIGCAS Conf on Computing and Sustainable Societies (Menlo Park and San Jose, CA, USA, June 20–22, 2018), Article No. 50, Association for Computing Machinery, New York (2018), pp. 1–5. https://doi.org/10.1145/3209811.3212707.
54. P. Hao, L. Di, C. Zhang, and L. Guo, "Transfer Learning for Crop classification with Cropland Data Layer data (CDL) as training samples," Sci. Total Environ., Vol. 733, 138869 (2020). https://doi.org/10.1016/j.scitotenv.2020.138869.
55. CropScape — Cropland Data Layer. URL: https://nassgeodata.gmu.edu/CropScape/.
56. E. Brewer, J. Lin, P. Kemper, J. Hennin, and D. Runfola, "Predicting road quality using high resolution satellite imagery: A transfer learning approach," PLoS ONE, Vol. 16, No. 7, e0253370 (2021). https://doi.org/10.1371/journal.pone.0253370.
57. H. Wang, X. Zhao, X. Zhang, D. Wu, and X. Du, "Long time series land cover classification in China from 1982 to 2015 based on Bi-LSTM deep learning," Remote Sens., Vol. 11, No. 14, 1639 (2019). https://doi.org/10.3390/rs11141639.
58. Z. Sun, L. Di, and H. Fang, "Using long short-term memory recurrent neural network in land cover classification on Landsat and Cropland data layer time series," Int. J. Remote Sens., Vol. 40, Iss. 2, 593–614 (2019).
59. J. Sun, Z. Lai, L. Di, Z. Sun, J. Tao, and Y. Shen, "Multilevel deep learning network for county-level corn yield estimation in the U.S. corn belt," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., Vol. 13, 5048–5060 (2020). https://doi.org/10.1109/JSTARS.2020.3019046.
60. W. Cho, S. Kim, M. Na, and I. Na, "Forecasting of tomato yields using attention-based LSTM network and ARMA model," Electronics, Vol. 10, No. 13, 1576 (2021). https://doi.org/10.3390/electronics10131576.
61. K. Gavahi, P. Abbaszadeh, and H. Moradkhani, "DeepYield: A combined convolutional neural network with long short-term memory for crop yield forecasting," Expert Syst. Appl., Vol. 184, 115511 (2021). https://doi.org/10.1016/j.eswa.2021.115511.