The 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications
21-23 September, 2017, Bucharest, Romania

Convolutional Neural Network Based Solar Photovoltaic Panel Detection in Satellite Photos

Vladimir Golovko [1], Sergei Bezobrazov [1], Alexander Kroshchanka [1], Anatoliy Sachenko [2,3], Myroslav Komar [3], Andriy Karachka [3]
[1] Brest State Technical University, Belarus, bescase@gmail.com
[2] Silesian University of Technology, Poland, sachenkoa@yahoo.com
[3] Ternopil National Economic University, Ukraine, mko@tneu.edu.ua

Abstract—**The aim of this work is the detection of solar photovoltaic panels in low-quality satellite photos. It is important to receive the geospatial data (such as country, zip code, street and home number) of installed solar panels, because they are connected directly to the local power. It will be helpful to estimate a power capacity and an energy production using the satellite photos. For this purpose, a Convolutional Neural Network was used. For training and testing dataset consists of 3347 low-quality Google satellite images was used. The experimental results show a high rate accuracy of detection with low rate incorrect classifications of the proposed approach. The proposed approach has enormous implementation and can be improved in future.**

Keywords—convolutional neural network; solar panels detection; satellite photos; geospatial data; power capacity; energy production

I. INTRODUCTION

The popularity of using the photovoltaic cells all over the world is growing from year to year. Solar's increasing competitiveness against other technologies has enabled it to increase quickly its share of the total electrical generation. With over 100% year-on-year growth in photovoltaic (PV) system installation PV module makers increased dramatically their shipments of solar modules in 2010. Thus, the U.S. solar industry has installed 742 MW of PV in the second quarter of 2012, up to 45% sequentially and 116% annually now. The industry has installed 2.85 GW of solar [1]. Nearly 260,000 Americans work in solar - more than double the number in 2012 - at more than 9,000 companies in every U.S. state. After reaching one million in 2016, two million should be hit in In connection with the growing popularity of PV the problem of their servicing is important. Many service companies are interested in the information about the potential clients. As a result the automatic detection of solar panels using the satellite photos is very actual.

The discovering the theory of deep learning by Hinton [3] and using the parallel computing based on GPU gave a huge impulse for the developing the different variants of Convolutional Neural Networks (CNN) for recognition and classification tasks which can be trained in the acceptable period of time.

The paper is organized as follows: a Section 2 contains relative works in the research area. A Section 3 describes the proposed approach and an architecture of CNN. A Section 4 provides some experimental results. The final Section 5 concludes this paper.

II. RELATIVE WORKS

CNN are developing the ideas of both Neocognitron introduced by Fukushima [4] and weight-sharing networks. In the first time a CNN, in its classical point of view (LeNet-5), was introduced by LeCun in [5]. The convolutional neural network integrates three approaches, namely local receptive field, shared weights and spatial subsampling [5-7]. Using local receptive areas the neural units of the first convolutional layer can extract primitive features such as edges, corners etc. The convolutional neural network represents a combination of convolutional and pooling layers (Fig. 1), which perform the nonlinear hierarchical transformation of the input data space. The last block of the convolutional neural network is a multilayer perceptron, Support Vector Machine (SVM) or other classifier.

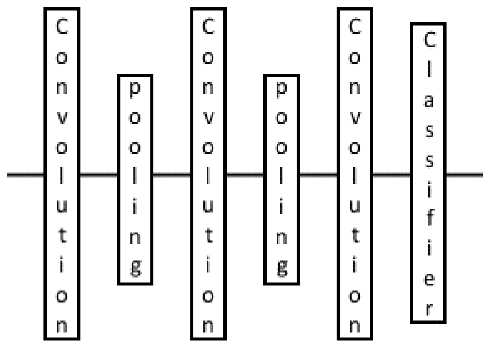

Figure 1. General representation of convolutional neural network

J. Malof et al. [8] developed the approach based on a Support Vector Machine for automatically identification rooftop PV installations using the high resolution color satellite orthoimagery. The approach first applies a pre-screening operation that identifies regions which are then considered for the more sophisticated feature processing. Features are extracted from each region and then a trained support vector machine is used to classify the features. The output of this approach is a list of regions and their respective confidence values, indicating how likely they are corresponding to a rooftop PV installation. The approach shows the detection rate of 94%.

III. PROPPOSED APPROACH

Our approach differs from the described in [8] one by using a CNN that enables the accuracy of detection and improve the flexibility. Another important feature of the proposed approach is using the low-quality satellite imagery (for instance, Google Maps photos), instead of the high resolution color satellite orthoimagery that enables decreasing the requirements for the approach.

We used the three satellite channels (RGB) images with the size 200x200 pixels. The images are put into the first convolution layer followed by a pooling layer (we used max-pooling) and propagates forward until last layer. Last fully connected layer of the network has one neuron only, which produces the probability of presence solar panels on image.

Fig.2 represents an architecture of the proposed CNN for the detection of solar panels.

The presented CNN consists of six layers. The first three layers are convolutional layers and form high level features of the images.

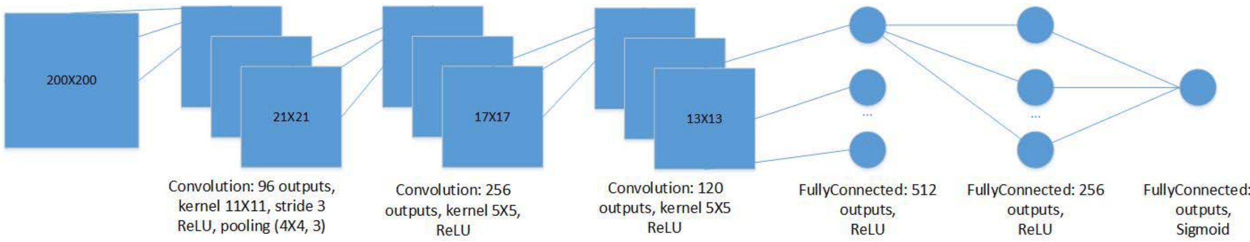

Figure 2. Architecture of the proposed model for solar panel detection

The first layer is a convolutional layer. A convolutional layer consists of neurons that connect to subregions of the input images (in our case there are satellite images with the size 200x200) or the outputs of the layer before it. A convolutional layer learns the features localized by these regions while scanning through an image. The first convolutional layer consists of 96 future maps and each future map contains 64x64 neurons.

A convolutional layer is usually followed by a nonlinear activation function. It is usually a rectified linear unit (ReLU) function, specified by a ReLU layer. It performs a threshold operation to each element, where any input value less than zero is set to zero:

$$f(x) = \max(0, x) . \tag{1}$$

Max-pooling layers follow the convolutional layers for down-sampling, and reduce the number of connections and parameters to be learned in the following layers. They also help to reduce overfitting. A max-pooling layer returns the maximum values of inputs rectangular regions [9]. We use the polling operation with the size 4x4. In both the convolutional layer and the pooling layer we use the stride equals 3. This argument determines the step size of the training function while it scans along the image. It is the additional way to decrease the computational load. Finally, after the first convolutional and pooling layers we have 96 feature maps and each map contains 21x21 neurons.

The second convolutional layer contains the 256 feature maps with 17x17 neurons in each map.

The third convolutional layer consists of 120 feature maps with the 13x13 neurons in each map.

The last three layers are fully connected layers that perform the classification task and have 512, 256 and one neurons correspondingly. The convolutional (and down-sampling) layers are followed by one or more fully connected layers. All neurons in a fully connected layer connect to the neurons in the previous layer. This layer combines all of the features learned by the previous layers across the image to identify the larger patterns. The last fully connected layer combines them to classify the images. That is why, the `number of neurons in` the last fully connected layer is equal to the number of classes in the target data [9]. Since we have two classes of images (1 – images that contains the solar panel, 2 – images without the solar panels) the last layer consists of 1 neuron. Thus the value '1' of the output neuron corresponds to the images with the solar panel. In opposite, the images without the solar panels will generate '0' on the output.

The last fully connected layer of the network has the sigmoid activation function

$$\sigma(x) = 1/(1 + e^{-x}) , \tag{2}$$

while the two first fully connection layers have the ReLU activation function (equation 1).

The goal of learning of the Convolutional Neural Network is to minimize the total mean square error (MSE), which characterizes the difference between real and desires outputs of the neural network. In order to minimize a MSE the gradient descent technique was used [10, 11]. The mean square error for L samples is defined as:

$$E_s = \frac{1}{2} \sum_{k=1}^{L} \sum_{j=1}^{m} (y_j^k - e_j^k)^2 , \tag{3}$$

where y^k_j and e^k_j–real and desired output of *j*-th unit for *k*-th sample respectively.

Then using gradient descent approach we can write in case of mini-batch learning, that

$$w_{cij}(t+1) = w_{cij}(t) - \alpha \frac{\partial E(r)}{\partial w_{cij}(t)} \quad (4)$$

where α is a learning rate, E(r) is a mean square error for r samples (size of minibatch).

Since the units of each feature map in a convolutional layer contain the same set of weights then the partial derivative is equal to the sum of partial derivative $\frac{\partial E(r)}{\partial w_{cij}(t)}$ for all neurons of the feature map:

$$\frac{\partial E(r)}{\partial w_{cij}(t)} = \sum_{i,j} \frac{\partial E(r)}{\partial w_{cij}(t)} \quad (5)$$

As a result in case of batch learning we can obtain the following delta rule to update synaptic weights:

$$w_{cij}(t+1) = w_{cij}(t) - \alpha(t) \sum_{i,j} \sum_{k} \gamma_{ij}^{k} F'(s_{ij}^{k}) x_{c}^{k} \quad (6)$$

where $c=1,\ \rho^2,\ F'(S_{ij}^k) = \frac{\partial y_{ij}^k}{\partial S_{ij}^k}$ – the derivative of activation function for k-th sample, S_{ij}^k – the weighted sum, γ_{ij}^k the error of ij-th unit in a feature map for k-th sample, x_c^k – the c-th input.

Taking into account that the proposed network is a deep one, we recommend to use the pre-trained network for achieving the better results.

IV. EXPERIMENTAL RESULTS

During the experiment we used the three channels satellite images with the size 200x200 pixels.

For training process and testing the dataset of 3347 satellite images (where 1643 images contain rooftop solar panels and 1704 without PV installations) from Google Maps was employed. The whole dataset was devised as 80% for training and 20% for testing (it is a standard ratio for neural network approach). Fig. 3 shows the examples of the images with PV installations and without them.

Figure 3. Examples of the images with solar panels and without it

As can be seen from Fig. 3, the images 1, 2 and 5 contain the solar photovoltaic panels on the roof while the images 3, 4 and 6 are without them.

Besides, the Convolutional Neural Network with the following parameters (a learning rate is equals to 0.001; a momentum parameter is equals to 0.9; a weight-decay is equals to 0.0005; a mini-batch size is equals to 20) was trained during the 70 epochs. It took about 2 hours with the NVidia GeForce 750TI video adapter (2GB memory on the board) and 8GB RAM.

Additionally, a dropout with 0.5 probability for fully connected layers of the proposed Convolutional Neural

Network was used. A dropout layer randomly sets the layer's input elements to zero with a given probability. Although the output of a dropout layer is equal to its input, this operation corresponds to temporarily dropping a randomly chosen unit and all of its connections from the network during training. So, for each new input element, a `training algorithm` randomly selects a subset of neurons, forming a different layer architecture [9].

In the test mode a simple transformation output data by applying threshold function was used. In this way

$$b_s = y_s > 0.5 \tag{7}$$

where y_s is a real output a CNN; b_s is a binarized form; $s = 1, 2, 3, ..., L$, L – a number of images in a dataset.

The accuracy was calculated using the next equation:

$$A = \frac{S}{L} * 100,$$
$$S = \sum_{S=1}^{L} 1(b_s = e_s) \tag{8}$$

where $1()$ is indicator function, e_s – target output of CNN ("Yes" or "No").

To verify the accuracy of the proposed approach we used 670 satellite photos where 313 images contain PV modules and 357 images without rooftop solar panels. The experiment results showed the overall accuracy of detection equals to 87% (a table 1 contains the confusion matrix of the results and the detection accuracy per each class of data).

TABLE I. CONFUSION MATRIX FOR PROPOSED MODEL

	Predicted "No"	Predicted "Yes"	Accuracy, %
Expected "No"	325	32	91
Expected "Yes"	52	261	83
Summary	377	293	87

Fig. 4 represents ROC-characteristics of the proposed CNN classificator.

Fig. 5 contains the visualisation of outputs for the first convolutional layer. Using local receptive areas the first convolutional layer of a well-trained Convolutional Neural Network can extract different primitive features as edges, corners etc. As can be seen on the Fig. 5, the features on this layer are well organized.

It was significant to note that not all satellite images are with the satisfied quality (Fig. 6).

It means that it is possible to increase the detection rate significantly by implementing the additional filter for 'bad' images.

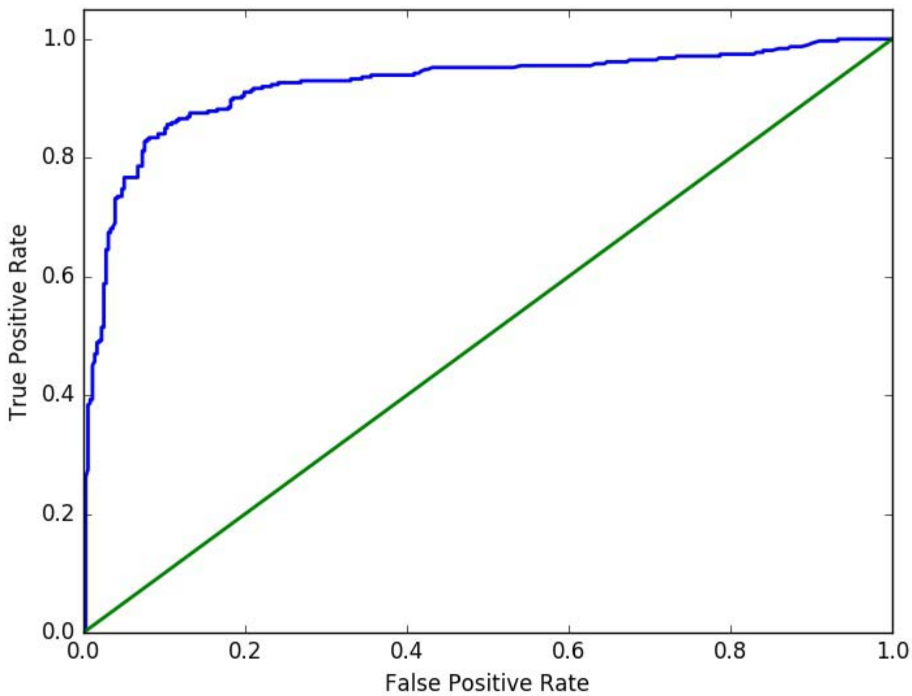

Figure 4. ROC-curve: AUC = 0.92, Recall = 0.8339, Specificity = 0.9104, Precision = 0.8907, F-measure = 0.8614.

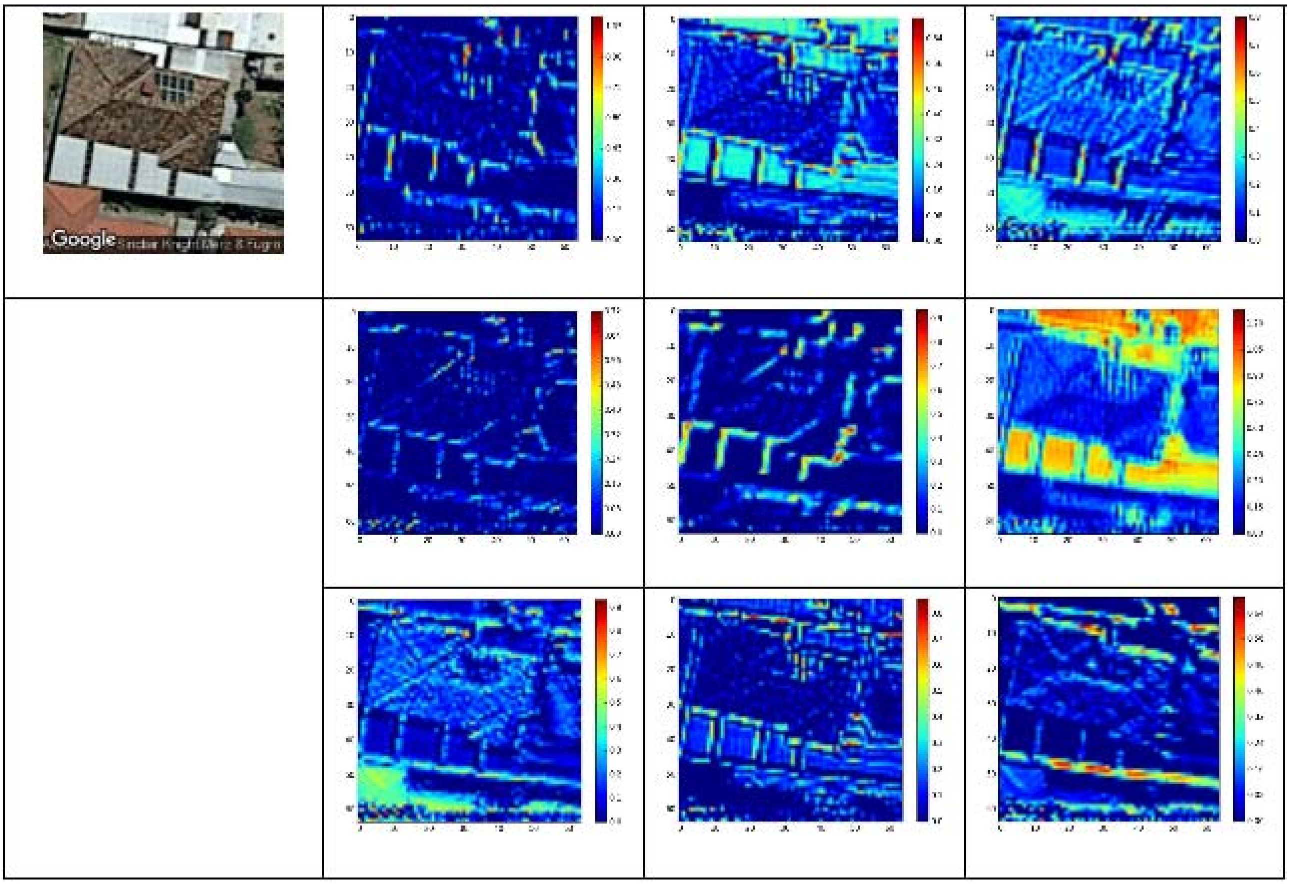

Figure 5. Visualization of outputs maps from the first layer of CNN

Figure 6. The example of unsatisfactory satellite images

V. CONCLUSION

Authors proposed the approach based on a Convolutional Neural Network with the six layers for rooftop solar panels detection on satellite photos. The first three layers are convolutional layers and form high level features of images. The last three layers are fully connected layers that perform classification task. That will enable to estimate a power capacity and an energy production of rooftop photovoltaic installations using the satellite photos.

The experimental results confirmed a high rate accuracy of the detection with the low rate of incorrect classifications. In comparison with the existing approaches that operate with the high resolution color satellite orthoimages we use the the low-quality satellite imagery (for instance, Google Maps photos), that enables decreasing the requirements for the approach.

In future authors plan to explore the quality of the proposed approach using the additional filter for satellite images with the "bad" quality and inviting the experts for organizing the quality train set because some fragments of the roofs can looks like solar panels but not they are.

REFERENCES

[1] Solar stats that will blow your mind, [Online]. Available: https://www.aol.com/article/2012/09/12/solar-stats-that-will-blow-your-mind/20322104

[2] Solar industry data, [Online]. Available: http://www.seia.org/research-resources/solar-industry-data

[3] G. E. Hinton, S. Osindero, and Y. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, pp. 1527-1554, 2006.

[4] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, vol. 36, issue 4, pp. 193-202, 1980.

[5] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, issue 11, pp. 2278-2324, November 1998.

[6] Y. LeCun, B. Boser, J. Denker, R. Henderson, R. Howard, W. Hubbard and L. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, issue 4, pp. 541-551, 1989.

[7] Y. LeCun, P. Haffner, L. Bottou and Y. Bengio, "Object recognition with gradient based learning," in Proceedings of the Shape, Contour and Grouping in Computer Vision, Part of the Lecture Notes in Computer Science book series (LNCS), Springer, vol. 1681, pp. 319-345, 1999.

[8] J. Malof, R. Hou, L. Collins, K. Bradbury, and R. Newell, "Automatic solar photovoltaic panel detection in satellite imagery," in *Proceedings of the International Conference on Renewable Energy Research and Applications ICRERA'2015*, 2015.

[9] Specify Layers of Convolutional Neural Network, [Online]. Available: https://www.mathworks.com/help/nnet/ug/layers-of-a-convolutional-neural-network.html

[10] V. Golovko, A. Kroshchanka, D. Treadwell, "The nature of unsupervised learning in deep neural networks: A new understanding and novel approach," *Optical Memory and Neural Networks*, vol. 25, issue 3, pp. 127-141, 2016.

[11] V. Golovko, E. Mikhno, A. Brich, A. Sachenko, "A shallow convolutional neural network for accurate handwritten digits classification," Springer International Publishing AG 2017, V. V. Krasnoproshin and S. V. Ablameyko (Eds.): *Communications in Computer and Information Science*, pp. 77–85, 2017.