Application of AI tools to the inventory of technical and transportation infrastructure based on UAV data

Paulina ZACHAR, Mateusz BUDA, Maksymilian FOLTYN, Wojciech OSTROWSKI, Radosław PALAK, Konrad SOSNOWICZ, Krzysztof BAKUŁA, Poland

Key words: Artificial Intelligence, deep learning, CNNs, UAV, object detection, inventory, technical and transport infrastructure, managing an investment

SUMMARY

The dynamic development of artificial intelligence (AI) technologies, including deep learning methods, has meant that these solutions can be widely applied in photogrammetry and remote sensing. Recent years have also been marked by the development of techniques for acquiring spatial data from unmanned aerial vehicles (UAVs). The photogrammetric data obtained in this way can be processed and applied in the inventory and modelling of technical and transport infrastructure objects using deep learning algorithms.

The development and research work in this project, which is carried out in cooperation with SkySnap company, aims to elaborate and deploy innovative services in the production process that allow managing an investment at every stage of its functioning. The project develops a methodology based on machine learning (precisely convolutional neural networks) using photogrammetric products to recognise corridor objects in transport and energy (e.g. traction masts, streetlamps, sleepers).

In this study, the methodology was developed to acquire photogrammetric data from UAVs closely (including images and point clouds) related to the input data requirements for neural networks. The process of preparing datasets for training neural network models was automated and tools were created to use data from completed projects.

Preliminary experiments were related to the detection of rail sleepers on orthophotomap, for which the accuracy of the trained network was over 90%. This research is still ongoing, with more objects of interest and other approaches.

Application of AI tools to the inventory of technical and transportation infrastructure based on UAV data

Paulina ZACHAR, Mateusz BUDA, Maksymilian FOLTYN, Wojciech OSTROWSKI, Radosław PALAK, Konrad SOSNOWICZ, Krzysztof BAKUŁA, Poland

1. INTRODUCTION

Deep learning (DL) can be described as a set of state-of-the-art methods within the artificial intelligence (AI) domain and is a technology that is rapidly evolving. The main promise of supervised DL methods is the ability to learn from large volumes of labelled data to solve tasks and problems on their own when deployed in a new test environment. Using these technologies, algorithms can be trained to perform specific tasks by processing large amounts of data and recognising patterns in datasets.

Despite many tasks that the technology of broad artificial intelligence and deep learning methods can handle, many challenges remain. One of which is for AI to be able to equal a human perception of stimuli from the real world. Additionally, some tasks and problems require intuition, informal knowledge, or even subjective opinion in addition to acquired skills to solve. Currently, the aim is, for systems that rely on AI to be able to absorb the knowledge and infer patterns found in data.

The methods of artificial intelligence (AI), and more specifically deep learning (DL), have been widely used in the fields of photogrammetry and remote sensing, in areas related to Earth observation (Chew et al., 2020; Lyu et al., 2021). These methods have been particularly widespread in the area related to 2D image processing. Thanks to the application of these algorithms, automation of time-consuming processes has become possible, including, among others, the detection of objects in aerial images or the classification of satellite scenes. The acceleration of image data processing is significant because many new datasets are acquired every day. For the acquisition of information, and at a later phase, knowledge from the acquired data to be possible with the continuous expansion of databases, the aspect of automation is precious. Currently, the scale of automation still appears too small in relation to the amount of acquired data. However, the project described in this paper, is an example of automating these processes, not only in typical research solutions but also in commercial applications.

The advances in deep learning methods and neural networks have resulted in many successful solutions for 2D image processing, yet many difficulties remain within processing 3D data (Kowalczuk and Szymański, 2019). For example, despite emerging algorithms and large volumes of data, challenges remain with point cloud classification extraction from dense image matching (DIM) or Light Detection and Ranging (LiDAR) point cloud data obtained from Airborne Laser Scanning (ALS). The first research studies that attempted to develop point clouds using deep learning methods originated around a decade ago. However, more attention was focused on this topic a few years ago. Also, many of theses have been presented so far, proving the great potential of using 3D data in neural network training processes.

In addition to airborne and satellite-derived data, recent years have seen a significant development of sensors designed to acquire data from unmanned aerial vehicles (UAVs). Photogrammetric data and products acquired from drones can also be processed and form the basis of the inventory process and modelling of technical and transportation infrastructure objects using AI tools and deep learning algorithms.

Moreover, there is a perceived need for an efficient way to capture and process construction data, as currently it often comes down to lengthy fieldwork and inventory taking. The approach developed in the project will enable the automation of data processing, acceleration of advanced analysis, and increased control over the construction process in a construction project lifecycle.

The project's main challenge will be developing deep machine learning models, more precisely convolutional neural networks, to detect various objects of technical and transport infrastructure in BIM technology. The project goal is to create an IT environment and develop a new service to enable fast, accurate mapping of the situation of an investment. The support of the service by AI solutions aims to develop algorithms capable of transforming drone-acquired and BIM/GIS data into a form suitable for further analysis.

The project focuses on corridor transportation and energy investments. Objects of interest include linear facilities and point and surface infrastructure elements.

The paper is organised as follows. Section 2 describes related works. Section 3 develops the project description and assumptions. The research methodology is presented in Section 4. Then it is further divided into subsections about corridor objects recognition, used data and preparation of training datasets, and details of an implementation of the neural network models. The results of the research work carried out so far are presented in Section 5. Finally, section 6 describes future work and provides the final conclusions.

2. RELATED WORKS

UAVs provide an efficient approach to acquiring high-resolution airborne data on technical and transportation infrastructure objects. The number of research articles dealing with UAV-based technical and transportation infrastructure object detection has increased during the last few years. Furthermore, detection and segmentation algorithms based on the deep learning framework have grown rapidly to process images acquired from UAVs (Singh et al., 2021). For object detection and effective infrastructure monitoring with computer vision techniques and deep learning methods, relevant information can be delivered by photogrammetric products such as 3D point clouds (LiDAR and DIM-based), aerial images, and videos.

Recently, conducting inspection and monitoring of infrastructure using UAVs has gained interest in industries such as railroads, power generation, and road construction. Advantages of using UAVs for this purpose include, for example, elimination of disturbance of train traffic due to required earthworks, the possibility of inspecting places that are difficult to access, reduction in manpower, fast damage detection (Bojarczak and Lesiak, 2021), cost and time savings of inspections.

One challenging problem in this area is the lack of publicly available datasets that can be used for training. These limitations are related to the high costs of data acquisition and annotation. While publicly available training datasets for detecting objects, such as cars in UAV imagery exist (Heo et al., 2020; Yang et al., 2019; Kyrkou et al., 2018; Li et al., 2020), in the case of engineering infrastructure, these are only just being made available. An example training dataset, with annotations provided by the authors, is for power line detection (Zhang et al., 2019).

Details of the project are described in the following section, but it is worth mentioning that our project focuses on railroad, highway, and energy objects. Still, this work will present the initial results of using deep learning methods to detect railroad sleepers and railroad rails. Similar solutions for power line corridor surveys are present in the literature (Mammeri et al., 2021), where based on UAV images, they segment the railway tracks. In their paper, Singh et al. (2019) explore the possibilities of computer vision-based monitoring through drone imagery. The paper by Mao et al. (2021) presents a pipeline for embedding road sign models based on CNNs.

There are also works where deep learning models are used for 3D point cloud segmentation obtained from UAV (Hu et al., 2020), or for detection and 3D modelling of rail tracks using dense point clouds obtained from UAV images (Sahebdivani et al., 2020). The possibilities provided by modern remote sensing sensors in power line corridor surveys (also drones) have been described in detail in the literature review article (Matikainen et al., 2016).

3. PROJECT DETAILS

This project's main objective is to develop a service for inventory and modelling key technical and transportation infrastructure objects in BIM technology, using AI tools. As mentioned in the introductory chapter, the demand for construction data continues to grow. Currently, available technologies are time-consuming and costly, especially when it comes to site work. At the forefront of technology, addressing these issues, is the use of drone data. To date, there are no developed standards detailing the feasibility of using drones or artificial intelligence tools in the inventory of technical and transportation infrastructure objects. Research in this area is innovative, and its results will create knowledge surrounding the optimal parameters of the flight mission and selecting appropriate AI tools.

The work carried out under the project develops a methodology for correct drone data acquisition and the possibility of using photogrammetric products to recognise transport and energy infrastructure objects using AI tools. The expected primary outcomes of the conducted research include:

- development of a methodology for data acquisition from drones that matches input requirements for convolutional neural networks,
- development and training of neural network models which can automatically, quickly, and accurately support the process of acquiring the necessary information in 3D,
- methods of using BIM in the process of machine learning and integration of artificial intelligence results with BIM and GIS environments,
- the possibility of automatic terrain modelling based on objects detected by the algorithms.

The research results will have practical applications in pre-project inventories, analysing the progress of work for the investor and updating investment information for managers at the stage of investment operation and before reconstruction.

The project includes specific research problems that were defined at the beginning and should be solved after the work has been completed:

- research on how to transfer results from DL models from 2.5D images to 3D,
- research on finding objects on a point cloud, for which the point location accuracy is approximately 5-15 cm,
- research on integrating data and different photogrammetric products to create solutions where, for example, in addition to the RGB image, the input to learn the network is a Digital Surface Model (DSM).

On the other hand, the defined technical problems are:

- problems related to data resolution,
- problems related to the planning of automatic photogrammetric missions to acquire adequate data consistent with the input requirements for DL models,
- problems associated with the optimisation of work on large datasets.

4. METHODOLOGY

Research on data processing methodologies using AI techniques will enable a methodology to be developed, optimised for solving measurement and inventory problems encountered during surveying linear, technical, and transport infrastructure objects. The project's main objective is to create a solution that will enable automation of the inventory process and investment monitoring during construction and after completion.

The works related to the development and implementation of deep learning algorithms consist of three main stages:

- preparation of data for the model training process
- design of neural network architectures

– testing and optimising algorithms, examining the performance of models considering the iterative process of improving the results.

This chapter provides an overview of the methodology and approaches used in the later stages of the project. The method for railroad sleepers, traction poles, and railway tracks will be described in more detail, as a workflow was created and validated for these objects first.

4.1 CORRIDOR OBJECT RECOGNITION

Based on consultations with technology partners regarding the demand for the functionality of the target solution, a list of objects of interest was defined, which should be subject to detection and segmentation by the artificial intelligence system being under construction in this project.

As shown in Table 1Table 1Error! Reference source not found., the infrastructure elements on which the research work related to model training is focused, will be objects from the road, rail, and energy projects. As far as road objects are concerned, the following have been selected: streetlamps, elements of groundworks, and elements of construction layers. Selected railroad objects include rail sleepers, rail tracks, and traction networks. The executed groundworks and heaps, power poles, and powerlines are energy infrastructure objects for which detection and segmentation models will be designed.

No.	Object	Reference	Assumed data inputs to network	Assumed classes being the result of model	Network operation	Development networks/ pipelines for dimensions
1	streetlamps	reference of streetlamps locations from completed projects in the form of points	orthophotomaps and images for 2D/2.5D; point clouds for 3D	bounding box in 2D/2.5D; point cloud segment in 3D	detection	2D and 3D
2	executed ground works and heaps	polygons manually labelled from orthophotos	orthophotomaps and images for 2D/2.5D; point clouds for 3D	-excavation (polygon in 2D/2.5D, point cloud segment in 3D) -embankment (polygon in 2D/2.5D, point cloud segment in 3D) -heaps (polygon in 2D/2.5D, point cloud segment in 3D)	segmentation	2D and 3D
3	construction layers	polygons manually	orthophotomaps and images for	-road chippings (polygon in	segmentation	2D and 3D

	(aggregate and asphalt/concrete layers)	labelled from orthophotos	2D/2.5D; point clouds for 3D	2D/2.5D, point cloud segment in 3D) -asphalt/ concrete (polygon in 2D/2.5D, point cloud segment in 3D)		
4	rail sleepers	point layer with marking the centre of the sleeper railroad	orthophotomaps and images for 2D/2.5D; point clouds for 3D	bounding box in 2D/2.5D; point cloud segment in 3D	detection	2D and 3D
5	railway track	railway track axis as a line layer	orthophotomaps and images for 2D/2.5D	railway lines - polygon in 2D/2.5D	segmentation	2D
6	traction network	line in the form of point cloud	point clouds for 3D	point cloud segment in 3D	segmentation	3D
7	executed ground works and heaps	polygons manually labelled from orthophotos	orthophotomaps and images for 2D/2.5D; point clouds for 3D	-excavation (polygon in 2D/2.5D, point cloud segment in 3D) -embankment (polygon in 2D/2.5D, point cloud segment in 3D) -heaps (polygon in 2D/2.5D, point cloud segment in 3D)	segmentation	2D and 3D
8	power poles	reference of power poles locations in the form of points layer	orthophotomaps and images for 2D/2.5D; point clouds for 3D	bounding box in 2D/2.5D; point cloud segment in 3D	detection	2D and 3D
9	powerlines	line in the form of point cloud	point clouds for 3D	point cloud segment in 3D	segmentation	3D

Table 1 Objects of interest for the detection and segmentation tasks of artificial intelligence system under construction.

The input data for the neural network training process is: drone images, point clouds (both LiDAR and from DIM), orthophotos, and vector BIM spatial layers. The final products (as a result of the neural network model) will be classified point clouds with retrieved objects, vector (2D and 3D) forms of representation of retrieved elements.

4.2 PREPARATION OF TRAINING DATASETS

Nine objects from three different industries are selected for the work. Archival investment data and data acquired during the project will be used as input data for network training. To support the automation of the process of preparing the training datasets, a methodology (Figure 1) has been adopted.

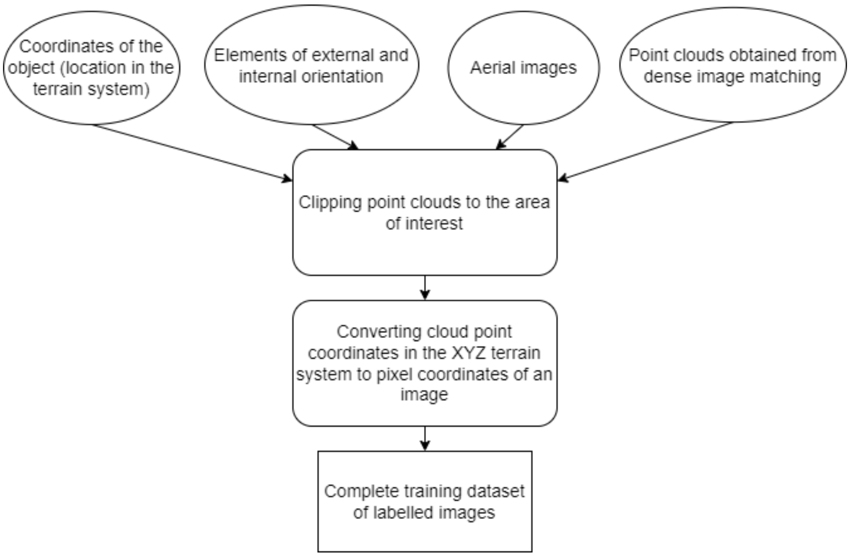


Figure 1 A block diagram showing the process of automating the creation of training datasets for neural networks.

Examples from the training collections are shown below (Figure 2).

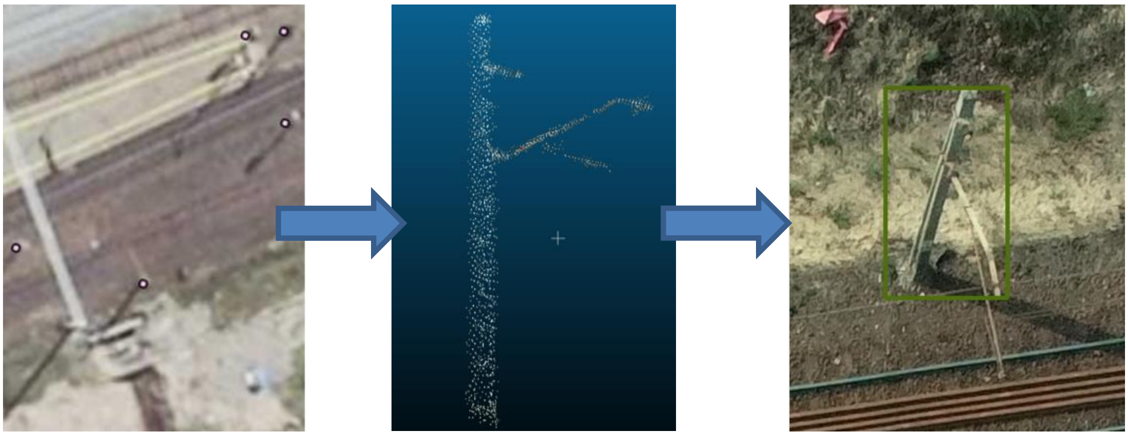


Figure 2 Examples from a training dataset on a traction pole example using a methodology to automate training data preparation.

4.3 DESIGN AND TRAINING OF NEURAL NETWORKS

In the project, we develop a reliable method to solve each task defined by model input (orthophotomaps, images, or 3D point clouds) and output (segmentation or detection). For each project task, we reviewed different approaches to the use of neural networks in task-specific context, the impact of transfer learning methods, and designed evaluation approaches created as part of the project methodology; the datasets are divided into learning, testing and validation sets. At the initial stage of creating subsequent models and network architectures, the initialisation strategy for weights, loss activation functions, optimisation and regularisation algorithms, and metrics are established. Network layers are defined for each model, and model tuning is automated.

We decided to use the same model validation strategy, that prevents data leakage and overoptimistic test performance, for all project tasks. We divided the input dataset into training, validation, and test subsets to ensure there were no locations between them. An additional criterion for the split was to sufficiently represent data collection settings in each subset. Then, we developed network type and architecture for each task, an initialisation strategy for model weights, loss functions, optimisation, and regularisation algorithms. All hyper-parameters related to networks and training were selected based on performance on the validation set. In addition, the optimisation of some hyper-parameters were automated to achieve the best performance using the Tree-structured Parzen Estimator algorithm for sampling with 20 trials.

The initial part of the project focused on the first set of tasks to build neural network models for detection and segmentation in 2D input data. First, orthophotomaps were taken as research subjects, as products of photogrammetric studies. For 2D detection task, we used the RetinaNet model (Lin et al., 2017) with a feature-pyramid network and ResNet-50 backbone (He et al., 2016). For 2D segmentation tasks, we used a fully convolutional neural network, U-Net model (Ronneberger et al., 2015) with 5 encoding and decoding layers, having 64 filters in the initial layer. The models were developed in PyTorch (a high-performance deep learning library (Paszke et al., 2019)), where solutions were adopted as the working environment, and the MLFlow library (Zaharia et al., 2018) was used to control the training process.

The first objects of railroad infrastructure - railroad sleepers, traction poles, and railway track - were taken into consideration. In this approach, based on training sets, a network for detection and recognition of railroad infrastructure facilities in images was trained - the first methodology and knowledge related to the use of photogrammetric data with modern IT solutions were created.

5. EXPERIMENTS AND RESULTS

The foundation for research in the project is the process of building a complete solution, based on machine learning including pre- and post-processing activities. The first approaches to developing methodologies using AI tools involved experiments and initial test runs to create neural network architectures for rail infrastructure facilities.

Many techniques are used in object detection and classification in images. While an approach using typical detection, i.e., detecting pixels that are corners of bounding boxes locating the object of interest, can be applied to single objects such as sleepers or traction poles, for continuous objects (such as railway tracks), it is more reasonable to use segmentation, which involves finding the probability distribution of class membership for each pixel of the processed image.

Thus, as a result, the model that performs the detection task was used to detect traction poles of railroad networks and railroad sleepers, and the identification of railway track fragments from images was performed by the model for segmentation. As beforementioned, the initial work started with 2D data adaptation. First of all, orthophotomaps were taken as a subject of research.

This chapter presents further approaches and iterations of model creation and training. The first experimental results are described, along with an accuracy assessment.

5.1 DETECTION OF TRACTION POLES AND SLEEPERS ON ORTHOPHOTOMAPS

Developing neural networks requires tuning multiple hyper-parameters, notably affecting the final performance. Throughout the development of the deep learning model, the use of single and dual-phase detectors was considered. The characteristics of the data-heavy, 3-channel true-orthophoto images and rasters- representing a digital surface model- indicated the advantage of using a single-phase detector. This type of detector is faster and has a significant advantage in the inference process. However, older versions of single-phase detectors (e.g., YOLO or SSD) were not up to par with two-phase detectors (e.g., Faster R-CNN), mainly due to the class imbalance problem. This is because most bounding boxes (especially at the beginning of the optimisation) come from the background class. In contrast, the more effective two-phase detectors (in cases where real-time detection is not required) are much slower and more complicated. Furthermore, their models took up more disk space when written to a disk file. Considering the above advantages and disadvantages of the different architectures, it was decided to conduct a series of more detailed experiments with the architecture based on RetinaNet, which proved to be a good compromise between the efficiency of two-phase detectors and the speed of single-phase detectors.

Because research has shown the best results are obtained when using RetinaNet architecture with the ResNet-50 backbone network, it was decided to use this approach for the first attempts to detect traction sleepers and poles on orthophotos acquired from drone imagery.

Of course, training data had to be prepared before training the model on data with labels of these two classes. The orthophotos were cut into 1024x1024 pixel images, all of which contained some target objects. The best model achieved a detection recall of 93.3% on the validation set for the railroad sleeper class. However, only 50.0% recall was obtained for the traction pole class. For all model evaluations, we assumed the predicted bounding box to be true positive if the intersection over union with a ground truth box was at least 0.5. The first results and the result from the model are shown in the figure below (Figure 3).

Figure 3 Examples of prediction on a validation dataset. The green bounding boxes are references, the red boxes are sleeper predictions, and the blue bounding boxes are traction pole predictions.

Since the results for traction poles were poor, the training dataset was verified, and inaccuracies in the annotations for this class were identified, which were caused by distortions and deformations resulting from the characteristics of the product, which is an orthophotomap. Based on the obtained results, it was decided that detecting high elements, such as traction masts or in the other part of the project, streetlamps will be done based on georeferenced images. The best model achieved a 94.8% recall on the validation set for the railroad sleeper class. The validation results for the sleeper class were satisfactory, as shown in the figure below (Figure 4).

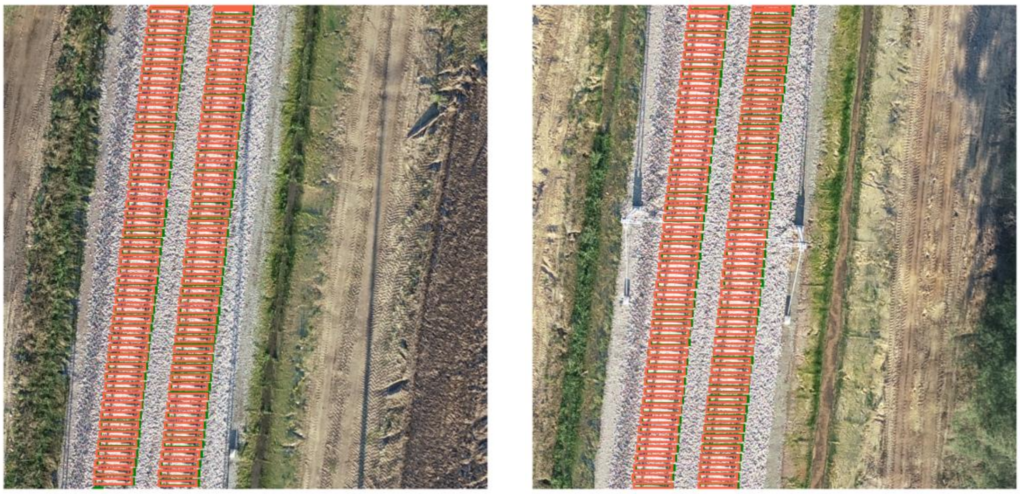


Figure 4 Examples of prediction on a validation dataset for the second approach. The green bounding boxes are references; the red boxes are sleeper predictions.

Qualitative analysis (visual assessment) and accuracy analysis were performed for the test set. For the test design, the results are shown in the following table (Table 2) and the following figure (Figure 5).

Parameter	Value
Recall	90.12%
Precision	91.12%
F1 score	90.62%

Table 2 Results for the sleeper detection model for the test subset.

Figure 5 Examples of prediction on a test dataset for the second approach. The green bounding boxes are references; the red boxes are sleeper predictions.

Independent evaluation and model tests were performed on data that was not part of the training process. The test data showed different characteristics and different resolutions. A significant effect of orthophotomaps resolution was noticed, and it was found that the training set size was still too small. Considering the conclusions after testing, another iteration of training was conducted using a more significant number of more diverse data, and data augmentation (scale, rotation, and colour augmentation) was used. The results of the second training of the model are shown in the table below (Table 3) and Figure 6 and Figure 7 are visual examples of the results.

Parameter	Value
Recall	92.10%
Precision	82.80%
F1 score	86.60%

Table 3 Results after added augmentation methods and more training examples for the rail sleepers detection model for the test dataset.

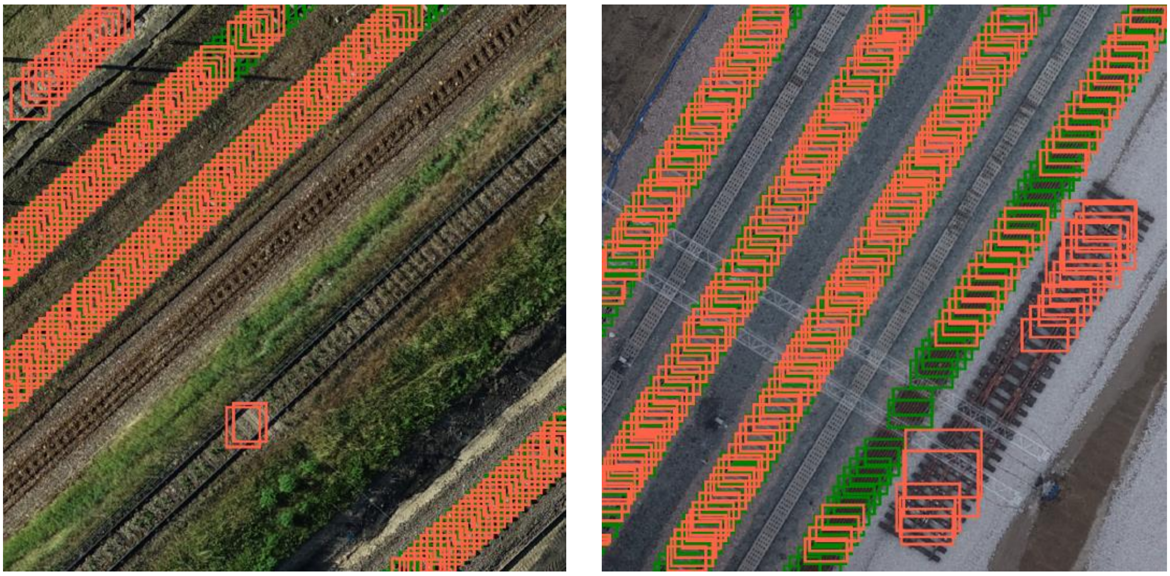


Figure 6 Examples of prediction on the validation dataset for the next iteration after added augmentation methods and more training examples for the rail sleepers.

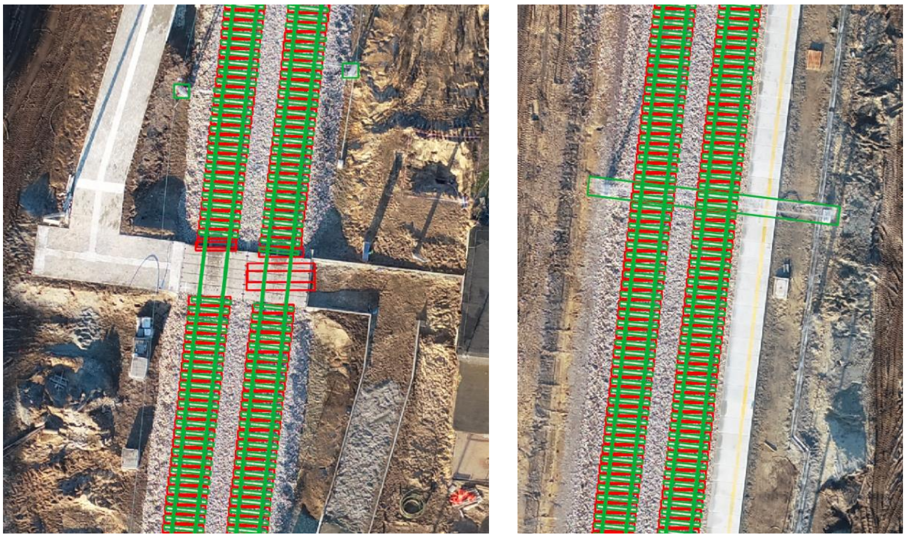


Figure 7 Examples of prediction on the test dataset for the next iteration after added augmentation methods and more training examples for the rail sleepers.

Both quantitative and qualitative evaluations revealed that adding data augmentation methods and expanding the training dataset to include additional examples improved the results. A comparison with the first trained model is shown in the figure below (Figure 8).

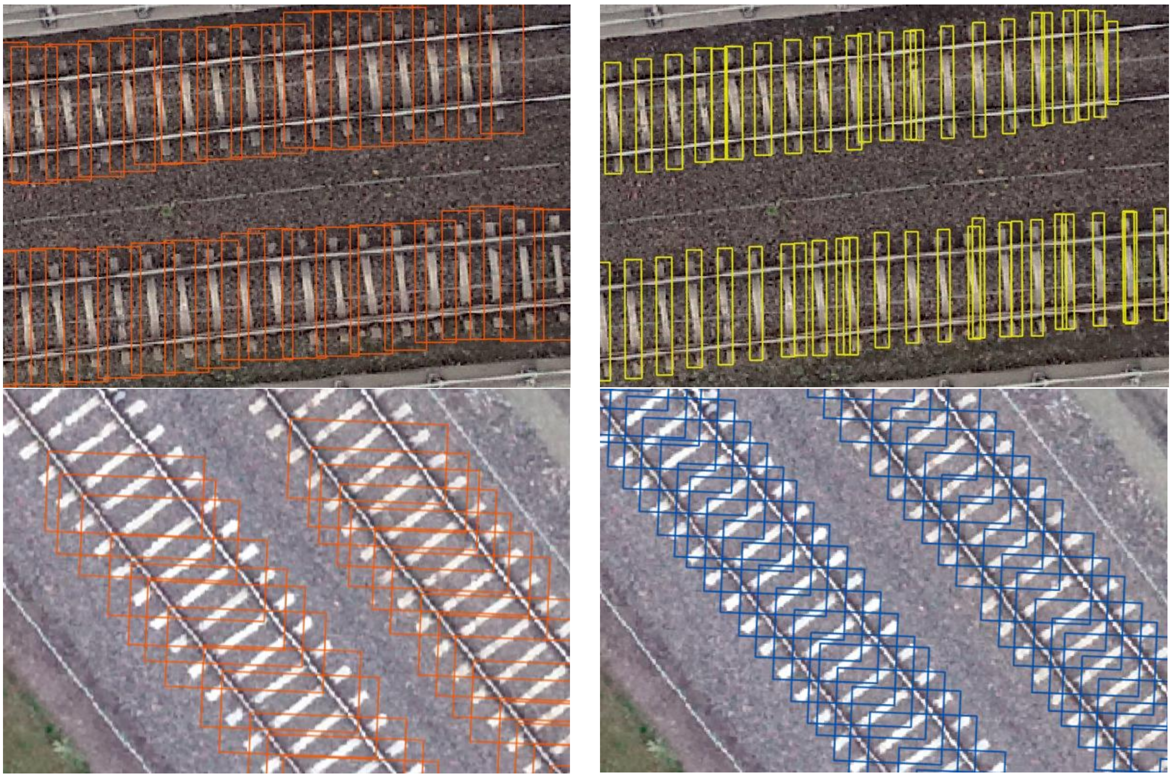


Figure 8 Comparison of results on an independent data set. Respectively, on the left, the results from the first training of the sleeper detection model; on the right, the model prediction results with added data augmentation and more training examples for the rail sleepers.

5.2 SEGMENTATION OF RAILWAY TRACKS ON ORTHOPHOTOMAPS

Another problem solved in this project was image segmentation, which was used to recognise railway tracks. In this case, the input data for the training process are orthophotos with shapefiles as labels. The expected result of the training process is a model that can effectively classify individual pixels of the input images. The model used for the segmentation task was a fully convolutional U-Net (Ronneberger et al., 2015) neural network, which is based on the encoding-decoding of the input.

The mIoU (mean Intersection over Union) metric was used to evaluate the model. For this particular task, the only annotated class was railway track. Relative to the total area of the input orthophotos, a very small fraction of them were areas containing rails. This is a typical example of unbalanced classes (in this case: the ratio of pixels containing rails to pixels being background). Running the training process on such data gave poor results (mIoU ~= 0.23). The literature provides a variety of possible solutions to this problem (Buda et al., 2018). In this case, undersampling was applied by selecting only those areas that contained at least one pixel of a non-background class. This improved the result to the 0.87 mIoU achieved with a training process almost twice as short. Examples of results of the segmentation of railway tracks on orthophotomaps are shown in the image below (Figure 9).

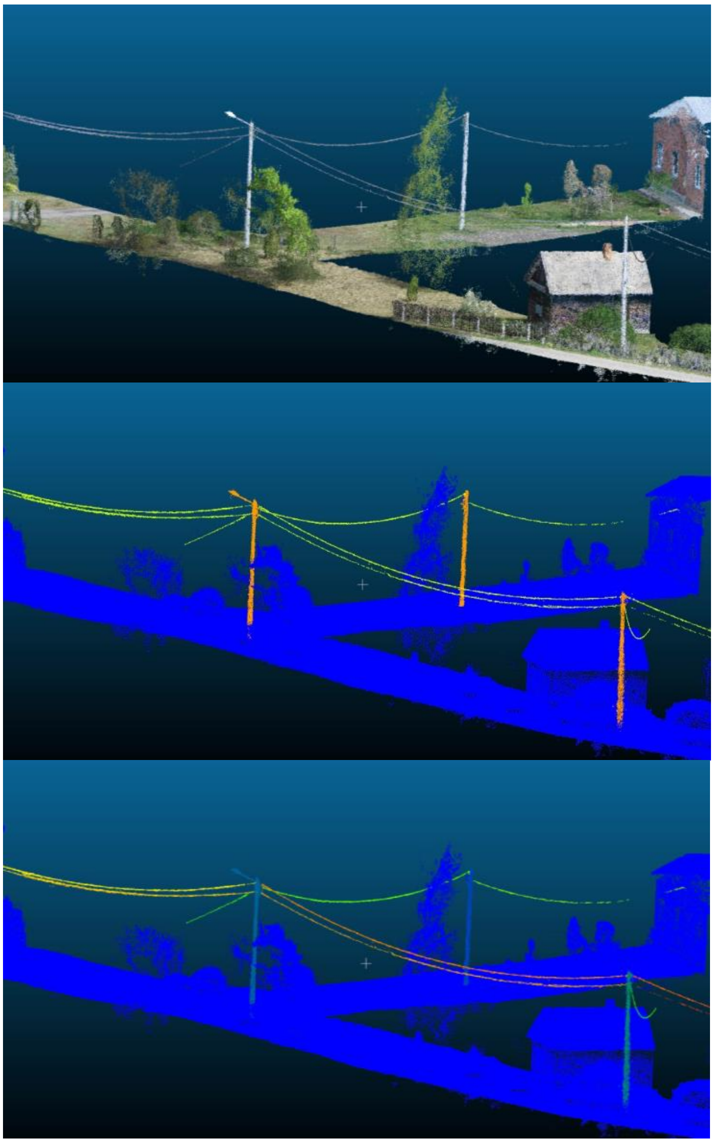
Figure 9 Segmentation results on different data from test sets.

Analysis of the results showed that the results were satisfactory, but the model in future iterations will be retrained using newly acquired data to improve accuracy and generalisation.

6. FUTURE WORKS AND CONCLUSIONS

This paper outlines the methodology adopted to recognize corridor objects using deep convolutional neural networks from UAV imagery. Delivering orthophotomap to the customer is no longer a required . Machine learning brings new opportunities for surveyors to automatically process data using a neural network to classify the objects of customers' interest. The project research work presented in this article focused on the first approaches related to creating the design and development of machine learning models. The first attempts consisted of 2D data adaptation, and more precisely, orthophotos, i.e., products of photogrammetric processing, were taken as the subject of research in the first place.

Due to unsuccessful attempts and problems connected with deformations of high objects (in this case, traction poles) on orthophotomaps, the second approach for objects extending above the ground level- using georeferenced photos directly - was started. Work is also in progress on tools that make full use of 3D point cloud data and training networks for subsequent objects of interest with the preparation of training sets. An example of the training set data for the neural network approach using 3D point clouds is shown below (Figure 10).



a)

b)

c)

Figure 10 An example from training dataset as input to neural network model in 3D point cloud approach. From the top: a) RGB point cloud from dense image matching, b) point cloud by class, c) point cloud by object ID number.

Summarising, the project demonstrates the potential of using artificial intelligence methods in the inventory and modelling of technical and transport infrastructure objects. The results so far are the basis for the conclusion that using proposed solutions with modern technologies accelerates the processing of data and reduces the workload.

ACKNOWLEDGMENT

The project is carried out by an interdisciplinary research team by SkySnap Sp. z o.o. company and Faculty of Geodesy and Cartography of Warsaw University of Technology within the project entitled, “R&D works in inventory and modelling of key technical and transport infrastructure objects in BIM technology using AI tools in the process of drone data processing", financed by the National Centre for Research and Development, grant no POIR.01.01.01-00-0980/20

REFERENCES

Bojarczak, P., & Lesiak, P. (2021). UAVs in rail damage image diagnostics supported by deep-learning networks. *Open Engineering*, *11*, 339–348. https://doi.org/10.1515/eng-2021-0033

Buda, M., Atsuto M., and Mazurowski, M.A., A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks* 106 (2018): 249-259.

Chew, R., Rineer, J., Beach, R., O’Neil, M., Ujeneza, N., Lapidus, D., Miano, T., Hegarty-Craver, M., Polly, J., & Temple, D. S. (2020). Deep Neural Networks and Transfer Learning for Food Crop Identification in UAV Images. *Drones*, *4*(1). https://doi.org/10.3390/drones4010007

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)*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Heo, W. Y., Kim, S., Yoon, D., Jeong, J., & Sung, H. (2020). *Deep learning based moving object detection for oblique images without future frames*. 2. https://doi.org/10.1117/12.2552347

Hu, Q., Yang, B., Khalid, S., Xiao, W., Trigoni, N., & Markham, A. (2020). *Towards Semantic Segmentation of Urban-Scale 3D Point Clouds: A Dataset, Benchmarks and Challenges (CVPR’2021)*.

Kowalczuk, Z., & Szymański, K. (2019). Classification of objects in the LIDAR point clouds using Deep Neural Networks based on the PointNet model. *IFAC-PapersOnLine*, *52*(8), 81–6. https://doi.org/10.1016/j.ifacol.2019.08.099

Kyrkou, C., Plastiras, G., Theocharides, T., Venieris, S. I., & Bouganis, C.-S. (2018). DroNet: Efficient convolutional neural network detector for real-time UAV applications. *2018 Design, Automation Test in Europe Conference Exhibition (DATE)*, 967–972. https://doi.org/10.23919/DATE.2018.8342149

Li, C., Yang, T., Zhu, S., Chen, C., & Guan, S. (2020). Density Map Guided Object Detection in Aerial Images. *CoRR*, *abs/2004.05520*. https://arxiv.org/abs/2004.05520

Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal Loss for Dense Object Detection. *2017 IEEE International Conference on Computer Vision (ICCV)*, 2999–3007. https://doi.org/10.1109/ICCV.2017.324

Lyu, Y., Vosselman, G., Xia, G., & Yang, M. Y. (2021). Bidirectional Multi-scale Attention Networks for Semantic Segmentation of Oblique UAV Imagery. *CoRR*, *abs/2102.03099*. https://arxiv.org/abs/2102.03099

Mammeri, A., Jabbar Siddiqui, A., & Zhao, Y. (2021). UAV-assisted Railway Track Segmentation based on Convolutional Neural Networks. *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, 1–7. https://doi.org/10.1109/VTC2021-Spring51267.2021.9448887

Mao, Z., Zhang, F., Huang, X., Jia, X., Gong, Y., & Zou, Q. (2021). Deep Neural Networks for Road Sign Detection and Embedded Modeling Using Oblique Aerial Images. *Remote Sensing*, *13*(5). https://doi.org/10.3390/rs13050879

Matikainen, L., Lehtomäki, M., Ahokas, E., Hyyppä, J., Karjalainen, M., Jaakkola, A., Kukko, A., & Heinonen, T. (2016). Remote sensing methods for power line corridor surveys. *ISPRS Journal of Photogrammetry and Remote Sensing*, *119*, 10–31.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., … Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d Alché-Buc, E. Fox, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems* (Vol. 32). Curran Associates, Inc. https://proceedings.neurips.cc/paper/2019/file/bdbca288fee7f92f2bfa9f7012727740-Paper.pdf

Ronneberger, O., Fischer, P., Brox, T. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*. Springer, Cham, 2015.

Sahebdivani, S., Arefi, H., & Maboudi, M. (2020). Rail Track Detection and Projection-Based 3D Modeling from UAV Point Cloud. *Sensors*, *20*(18). https://doi.org/10.3390/s20185220

Singh, A. K., Dwivedi, A. K., Nahar, N., & Singh, D. (2021). Railway Track Sleeper Detection in Low Altitude UAV Imagery Using Deep Convolutional Neural Network. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 355–358. https://doi.org/10.1109/IGARSS47720.2021.9554542

Singh, A. K., Swarup, A., Agarwal, A., & Singh, D. (2019). Vision based rail track extraction and monitoring through drone imagery. *ICT Express*, *5*(4), 250–255. https://doi.org/https://doi.org/10.1016/j.icte.2017.11.010

Yang, F., Fan, H., Chu, P., Blasch, E., & Ling, H. (2019). Clustered Object Detection in Aerial Images. *CoRR*, *abs/1904.08008*. http://arxiv.org/abs/1904.08008

Zaharia, M. A., Chen, A., Davidson, A., Ghodsi, A., Hong, S. A., Konwinski, A., Murching, S., Nykodym, T., Ogilvie, P., Parkhe, M., Xie, F., & Zumar, C. (2018). Accelerating the Machine Learning Lifecycle with MLflow. *IEEE Data Eng. Bull.*, *41*, 39–45.

Zhang, H., Yang, W., Yu, H., Zhang, H., & Xia, G.-S. (2019). Detecting Power Lines in UAV Images with Convolutional Features and Structured Constraints. *Remote Sensing*, *11*(11). https://doi.org/10.3390/rs11111342

BIOGRAPHICAL NOTES

Paulina Zachar

M.Sc. Eng., Ph.D. Student at Warsaw University of Technology. She graduated with her bachelor and master degrees from the faculty of Geodesy and Cartography (Department of Photogrammetry, Remote Sensing, and GIS). Her research spans the problem of object detection in photogrammetric data (mainly in oblique aerial images) using AI technologies, specifically convolutional neural networks. Her other current research interests are related to archaeology and urban planning. These include spatial analyses of ancient cities, as well as implementation of computer vision and image recognition methods for numismatic studies. She is involved in several research projects financed by Foundation for Polish Science National Centre for Research and Development and The National Science Centre. Paulina performed an implementation project with OPEGIEKA Sp. z o.o., “Methodology for automating the creation of object databases from synchronously acquired hybrid aerial photogrammetric data”. She won 1st place in the competition for the best thesis defended in the field of geodesy and cartography, edition 2019/20, organised by the Association of Polish Geodesists and the Chief Surveyor of the country under the patronage of the Geodesy Committee of the Polish Academy of Sciences.

Mateusz Buda

M.Sc. Eng. Mateusz Buda is a PhD candidate at Warsaw University of Technology. He received his undergraduate degree from the faculty of Mathematics and Information Sciences at Warsaw

University of Technology and master degree in Machine Learning from the faculty of Computer Science at KTH Royal Institute of Technology in Stockholm, Sweden. He worked as an Associate in Research at Carl E. Ravin Advanced Imaging Laboratories, affiliated with the Department of Radiology, at Duke University, North Carolina, USA. He was involved in applied machine learning project in various domains, including photogrammetry, medical imaging, and marketing. His research focuses on class imbalance problem, transfer learning, and multitask learning in convolution neural networks. Currently, he is responsible for the development of deep learning solutions and experimental environment for training object detection and segmentation models based on imaging and cloud point data at SkySnap.

Maksymilian Foltyn
Graduate from Warsaw University of Technology's Faculty of Geodesy and Cartography as well as Computer Science Faculty of Polish-Japanese Academy of Information Technology. Connecting a background in photogrammetry and remote sensing, along with an interest in Artificial Intelligence techniques, Maksymilian utilized deep learning for processing aerial and satellite image data within the Drone Powered Solutions Team of PwC's consulting department. His work concerned aerial data coming from several industries, from mining, through capital projects and infrastructure up to the energy industry. He gained experience during work on projects conducted globally for major clients in Northern California, central Africa, Europe, and Asia. Currently, he is responsible for the development of deep learning models and experiments environment within SkySnap. The main area of focus includes object detection and image segmentation, especially concerning capital projects (road, railway, energy) construction monitoring purposes. Utilised datasets consist of both 2- and 3-dimensional photogrammetric data.

Wojciech Ostrowski
He received his M.Sc. degree in geodesy and cartography (specialization: photogrammetry and remote sensing) in 2013 from the Warsaw University of Technology, Warsaw, Poland, where he has been working as a research-teaching assistant in the Department of Photogrammetry, Remote Sensing, and Spatial Information Systems, Faculty of Geodesy and Cartography, since 2013. His research interests include aerial and UAV photogrammetry and UAV lidar. The main topic of his research is the photogrammetric processing of oblique aerial images and the application of Machine Learning for urban semantic 3D meshes. His other research interest involves GIS and photogrammetry applications in archaeology. He is a member of the Polish Society for Photogrammetry and Remote Sensing, Computer Applications and Quantitative Methods in Archaeology (CAA), and Aerial Archaeology Research Group (AARG).

Radosław Palak
Eng. Radoslaw Palak graduated from the faculty of Geodesy and Cartography at Warsaw University of Technology in 2019 and since that time he has been working at SkySnap where he is the head of the photogrammetry department. His main responsibilities are: processing automatisation, management and consultation. In the project, he is responsible for combining photogrammetric know-how with artificial intelligence standards, for example, by seamlessly

converting datasets between different types of photogrammetric data. His work focuses on improving and scaling photogrammetric processes for commercial use.

Konrad Sosnowicz
Co-owner of SkySnap. Graduate from Warsaw University of Technology's Faculty of Geodesy and Cartography, specialisation in satellite navigation as well as Faculty of Spatial Planning. Since 2015, responsible for team building and technology development related to GIS, CAD and BIM in SkySnap. Currently Machine Learning Development Director of in R&D Department. As the main researcher, he implements, and coordinates R&D projects related to the development of artificial intelligence based on spatial vector and image data.

Krzysztof Bakuła
He received a Doctoral degree in 2015 from the Faculty of Geodesy and Cartography, Warsaw University of Technology. He is employed as an assistant professor in the Department of Photogrammetry, Remote Sensing and Spatial Information Systems. He is the author of over sixty scientific publications related to the analysis and the application of multi-sourced photogrammetric data, especially in issues associated with airborne laser scanning and aerial images. Over the past few years, he was involved in several research projects financed by i.e. Foundation for Polish Science National Centre for Research and Development. He is a contractor of forensic studies in photogrammetry and photo interpretation and several implementations works. He is a member of board of the Polish Society for Photogrammetry and Remote Sensing, vice-president of the Association of Polish Surveyors. He is involved in works of international organisations such as ISPRS, EuroSDR and FIG.

CONTACTS

Paulina Zachar
Warsaw University of Technology, Faculty of Geodesy and Cartography,
Department of Photogrammetry, Remote Sensing and Spatial Information Systems
Pl. Politechniki 1
00-661 Warsaw
POLAND
Email: paulina.konarzewska.dokt@pw.edu.pl

Wojciech Ostrowski
Warsaw University of Technology, Faculty of Geodesy and Cartography,
Department of Photogrammetry, Remote Sensing and Spatial Information Systems
Pl. Politechniki 1
00-661 Warsaw
POLAND
Tel. + 48 22 234 76 94
Email: wojciech.ostrowski@pw.edu.pl

Radosław Palak
SkySnap sp. z o.o.
Aleje Jerozolimskie 81
02-001Warsaw
POLAND
Email: radoslaw.palak@skysnap.pl

Mateusz Buda
SkySnap sp. z o.o.
Aleje Jerozolimskie 81
02-001Warsaw
POLAND
Email: mateusz.buda@skysnap.pl

Maksymilian Foltyn
SkySnap sp. z o.o.
Aleje Jerozolimskie 81
02-001Warsaw
POLAND
Email: maksymilian.foltyn@skysnap.pl

Konrad Sosnowicz
SkySnap sp. z o.o.
Aleje Jerozolimskie 81
02-001Warsaw
POLAND
Email: konrad.sosnowicz@skysnap.pl

Krzysztof Bakuła
Warsaw University of Technology, Faculty of Geodesy and Cartography,
Department of Photogrammetry, Remote Sensing and Spatial Information Systems
Pl. Politechniki 1
00-661 Warsaw
POLAND
Tel. + 48 222347694
Email: krzysztof.bakula@pw.edu.pl