Technology for Automated Monitoring of Construction Object Using Aerial Photography and Neural Networks

Dmitrii Zenkov[1*], Aleksandr Chibunichev[2]

[1] Moscow State University of Geodesy and Cartography, Moscow, Russia – zenkov858@gmail.com
[2] Moscow State University of Geodesy and Cartography, Moscow, Russia – agchib@miigaik.ru

KEY WORDS: Technical supervision, Aerial photography, Unmanned aerial vehicles (UAVs), Neural networks, Building inspections, Machine learning

ABSTRACT:

The quality of finished construction products is often compromised by various negative factors, including flawed engineering decisions, unqualified labor, and budget constraints. To mitigate these issues, technical supervision is essential for ensuring adherence to project and regulatory standards regarding timelines, costs, and quality. Traditional methods of technical supervision rely heavily on manual measurements and site inspections, necessitating the use of high-precision geodetic equipment operated by specialists. This research introduces an automated technology for monitoring the actual positions of capital construction objects through aerial photography captured by unmanned aerial vehicles (UAVs) and enhanced by neural networks. It aims to monitor structural geometry by comparing measured coordinates with design specifications. The proposed methodology is exemplified through the monitoring of pile positions in foundation construction, emphasizing the economic advantages of UAV-based monitoring over traditional geodetic methods, particularly in complex soil conditions. By utilizing calibrated UAV-mounted cameras and ensuring appropriate image overlap, this approach enhances the accuracy and reliability of coordinate measurements, ultimately contributing to improved construction quality and compliance with regulatory standards.

1. INTRODUCTION

The quality of finished construction products is often subject to many negative factors. These include erroneous engineering and design decisions, unqualified labor, and budget constraints. In such situations, technical supervision over the construction process becomes a necessary condition for delivering high-quality construction products.

During technical supervision, the timelines, costs, and quality of work are assessed in accordance with the project and regulatory requirements.

The utilization of classical methods in construction monitoring is characterized by the reliance on manual measurements, which are conducted through a series of sequential site inspections coupled with in-depth expert analysis of the data collected during these inspections. This traditional approach necessitates the employment of high-precision geodetic equipment, such as total stations and GPS devices, and requires the involvement of highly trained specialists who possess the expertise to operate this sophisticated technology effectively. The accuracy and reliability of the monitoring process hinge on the proficiency of these specialists and the quality of the equipment used.

The subject of this research is the development of an automated technology for monitoring the actual position of capital construction objects using aerial photography from unmanned aerial vehicles (UAVs) and neural networks

In recent years, rapid advancements in remote sensing technologies, along with the emergence of powerful neural networks, have significantly transformed the landscape of construction monitoring practices. This literature review aims to provide a comprehensive summary of recent studies that meticulously examine the application of these cutting-edge technologies in the automatic detection, segmentation, and monitoring of various construction projects. These advancements have paved the way for more efficient and accurate monitoring processes that can adapt to the dynamic nature of construction sites.

Currently, aerial photography utilizing UAVs is increasingly employed on construction sites for a variety of purposes. These include monitoring earthworks, conducting detailed inspections of buildings, and ensuring strict compliance with established construction schedules (Avetisyan R.T. et al., 2020). Furthermore, high-precision digital models of capital construction objects are meticulously created from the data obtained through aerial photography. These digital models serve as a crucial foundation for analyzing the condition of construction objects at all stages of their life cycle, which encompasses design, construction, operation, and eventual demolition (Adamtsevich L.A. et al., 2021; Vogel and Chakhkiev, 2021). This comprehensive approach allows stakeholders to make informed decisions based on accurate and up-to-date information regarding the status of construction projects.

The application of neural networks within the construction industry has a rich history that dates back to as early as the late

* Corresponding author

Publisher's note: Copernicus Publications has not received any payments from Russian or Belarusian institutions for this paper.

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-2-W9-2025-289-2025 | © Author(s) 2025. CC BY 4.0 License.

1980s (Mishchenko, 2024; Achanccaray et al., 2023). Over this substantial period, neural networks have been employed to address a wide array of tasks, including forecasting future trends, optimizing processes, modeling complex workflows, architectural design, tracking work schedules and sequences, among many others. To date, the range of areas where this machine learning approach is applied has expanded significantly and now encompasses various aspects such as construction object design, the creation of architectural projects, the pursuit of optimal engineering solutions, calculations of building characteristics, electrical circuit design, development of autonomous systems, and smart home technologies, among other applications.

Additionally, Li and Yilei (2020) propose the innovative use of a convolutional neural network specifically for image classification and the detection of undocumented structures. This represents just one example of how neural networks can be leveraged to enhance efficiency in construction practices. Furthermore, there are numerous other applications of neural networks that have emerged in the construction field (Zhou et al., 2019), showcasing the versatility and potential impact of these technologies on improving accuracy and effectiveness in various construction-related tasks. As these technologies continue to evolve, they promise to further revolutionize the construction industry by enabling more streamlined processes and better resource management.

A study of the publications of the above-mentioned authors allows us to conclude that the issue of using UAS technology and neural networks for the purpose of monitoring the representations of structural elements in the process of construction of capital construction projects has been little studied, which indicates a low level of coverage of this problem.

This article proposes the use of neural networks for the automatic detection and measurement of the coordinates of construction elements based on aerial photography data obtained from UAVs, with the aim of monitoring the geometry of the structure by comparing the obtained coordinates with the design ones.

2. METHODOLOGY

Figure 1 shows the proposed technological scheme for the automatic monitoring of construction objects using aerial photography and neural networks, illustrated by the example of monitoring the position of piles. During the construction of large-area nonlinear and non-point structures on loose and subsiding soils, pile fields are used for laying foundations. These piles are typically arranged in rows, and the position of each pile must comply with regulatory requirements. According to SP 45.13330.2017 "Earthworks, Foundations, and Pile Foundations," the tolerance for piles in a pile field under the entire building or structure should not exceed 40% of the pile diameter. The number of such piles can amount to hundreds. Therefore, the use of traditional geodetic methods to monitor the position of each pile is not always economically justified. The use of unmanned aerial vehicles for this purpose is preferable.

Aerial photography should be performed using a calibrated camera mounted on a UAV at a height that ensures the required accuracy for determining the coordinates of the piles. The overlap of images should be set at around 70-80% to enable the measurement of the coordinates of each pile across multiple images, thereby increasing the accuracy and reliability of automatic coordinate measurement using a neural network. The georeferencing of aerial photographs should be performed in accordance with the requirements of the national standard [9].

Photogrammetric processing of aerial photography data can be performed in any photogrammetric software that allows for phototriangulation and orthophoto transformation of each image individually. Since the original images are obtained in a central projection, the same pile appears differently in each image (Figure 2). If each image is transformed using a digital elevation model, the base of the pile will be visible in the orthophotos at different angles. This allows for multiple measurements of the coordinates of the pile base in the object coordinate system (based on the number of orthophotos).

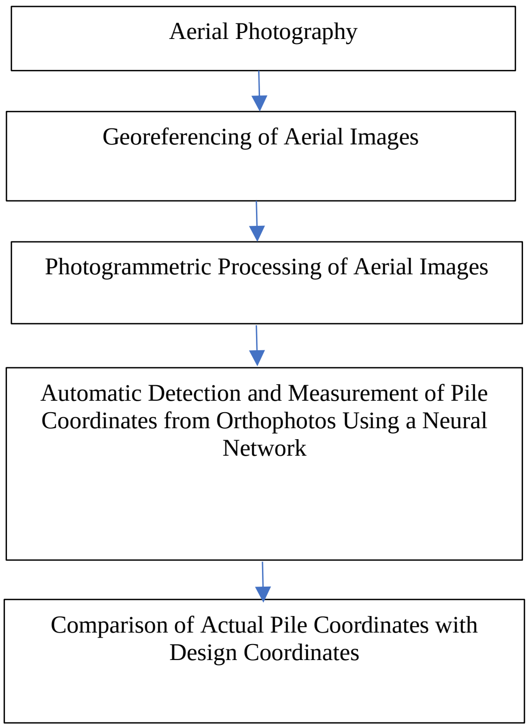

Figure 1. Technological scheme for automated monitoring of pile position

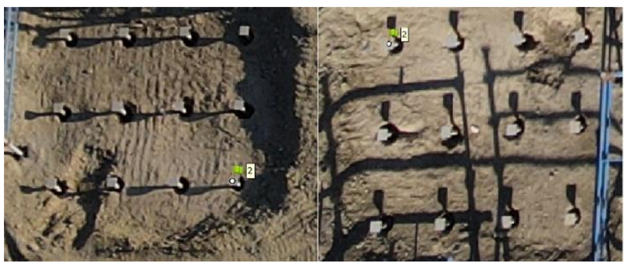

Figure 2. Image of the same pile in different orthophotos

3. RESULTS

To conduct experimental research on the proposed technology, it was first necessary to train a neural network to recognize piles and measure the coordinates of the pile bases.

The YOLOv8 model was chosen as the neural network model. According to Wu and Dong (2023), this model has proven to be the most effective in the field of object detection using remote sensing data, demonstrating high accuracy and speed.

https://doi.org/10.5194/isprs-archives-XLVIII-2-W9-2025-289-2025 |

Training was conducted using images of 5000 piles. The piles were labeled using rectangular bounding boxes with the CVAT software. During labeling, it was necessary to ensure that the base of the pile was located at the center of the bounding box. An example of pile labeling is shown in Figure 3.

Figure 3. Labeled piles

The neural network was trained with predefined hyperparameters. The quality of training was monitored using the mAP (mean Average Precision) metric. Satisfactory performance of the neural network is described by average values of $0.5 \leq mAP < 0.75$. If the metric falls below these values, it may be necessary to increase the volume of training data, improve labeling quality, or adjust hyperparameters. The medium-sized neural network model underwent several training cycles until the mAP reached 65%. The results of training the neural network are presented in Figure 4, where the x axis is epoch and the ordinate axis are the metric values according to the name of the graph.

train/box_loss and val/box_loss plots: loss metrics on the training and validation sets. Both curves show that the loss decreases as training progresses, indicating that the model is getting better at finding bounding boxes. The validation error (val/box_loss) fluctuates in the middle of training but eventually stabilizes. train/obj_loss and val/obj_loss: object function errors, which represent the model's confidence in the presence of objects in the frame. These errors also decrease as training progresses, indicating that the model is improving in confidence. train/cls_loss and val/cls_loss: classification errors (cls_loss) are irrelevant in this case since only one class of objects is recognized. metrics/precision and metrics/recall: precision and recall metrics increase and reach high values near the end of training, indicating that the model is working. metrics/mAP_0.5 and metrics/mAP_0.5:0.95: Mean Average Precision metrics at different IoU (intersection over union) levels. It is clear that mAP at 0.5 is quite high (around 0.7), and at the stricter mAP 0.5:0.95 it also increases to ~0.3.

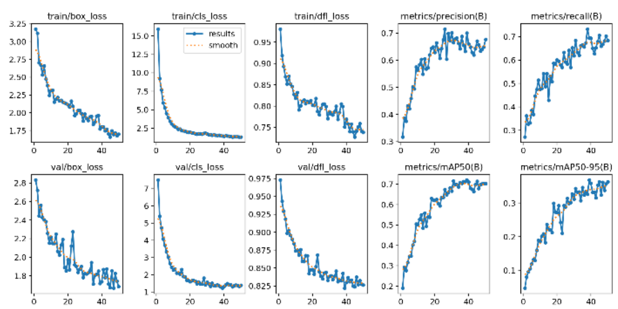

Figure 4. Layout of images and control points

For the experimental verification of the proposed pile position monitoring technology, aerial photography data of a pile field consisting of 40 piles was used. The aerial photography was performed using a DJI Mavic Air 2S UAV at flight height of 105 meters with a ground resolution of 3.5 cm/pixel. The longitudinal and transverse overlap was set at 80%. The aerial images were georeferenced using 12 control points, the coordinates of which were determined using GNSS equipment in RTK mode with an accuracy of 0.2 cm. The layout of the images and control points is shown in Figure 5.

Figure 5. Layout of images and control points

Phototriangulation and orthophoto transformation of all images were performed using the Agisoft Metashape software. The accuracy of orthophoto transformation was assessed using 4 control points. The root mean square errors (RMSE) are presented in Table 1.

Number of Points	RMSE X (cm)	RMSE Y (cm)	RMSE Z (cm)	RMSE XY (cm)	Total RMSE (cm)
4	1.1	0.9	0.7	1.5	1.6

Table 1. The root mean square errors (RMSE)

Figure 6 shows an example of an orthophoto of the pile field.

Figure 6. Orthophoto of the pile field

The orthophotographs obtained as a result of photogrammetric processing are divided into fragments of the same size in accordance with the training data set. Fig. 7-11 shows the results of pile recognition based on the orthophotographs obtained, displaying the recognition accuracy assessment from 0 to 1.

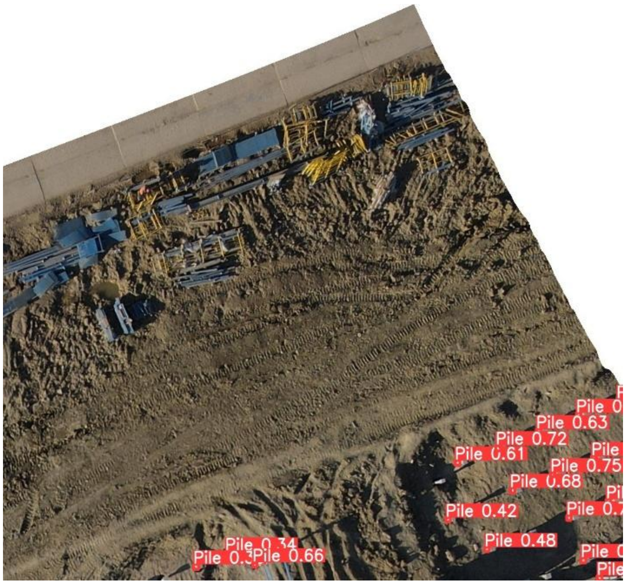

Figure 7. Example of pile recognition results on an orthophoto

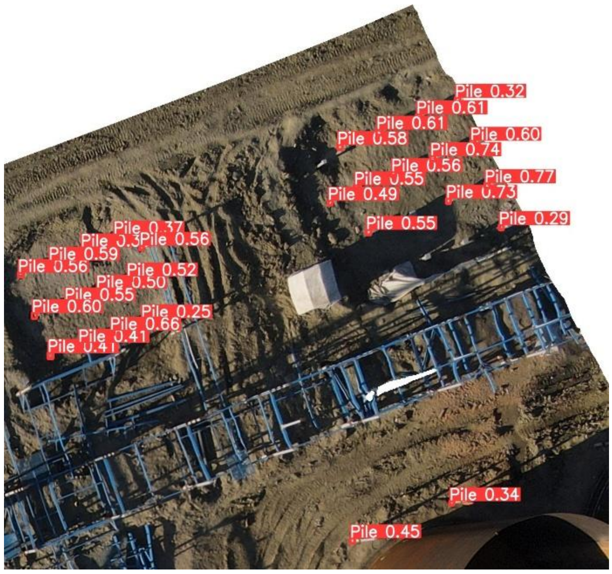

Figure 9. Example of pile recognition results on an orthophoto

Figure 8. Example of pile recognition results on an orthophoto

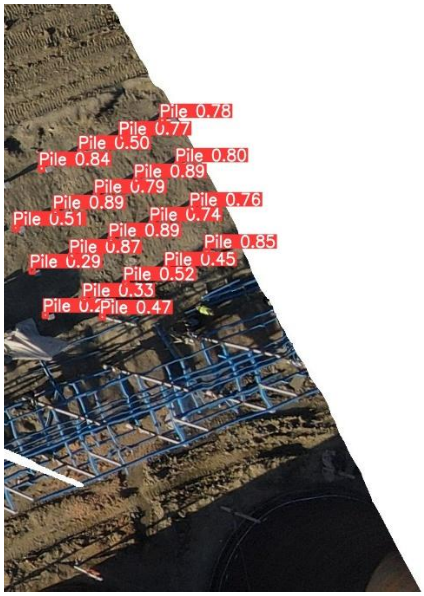

Figure 10. Example of pile recognition results on an orthophoto

Figure 11. Example of pile recognition results on an orthophoto

Figure 12 shows the results of automatic pile detection using the neural network. The coordinates of the piles measured automatically were compared with the design coordinates (a fragment of the design layout of the piles and pile heads is shown in Figure 13). Piles whose coordinates differ from the design by less than 8.8 cm (40% of the pile diameter) are shown in green, while those exceeding this tolerance are shown in red.

Figure 12. Results of pile detection

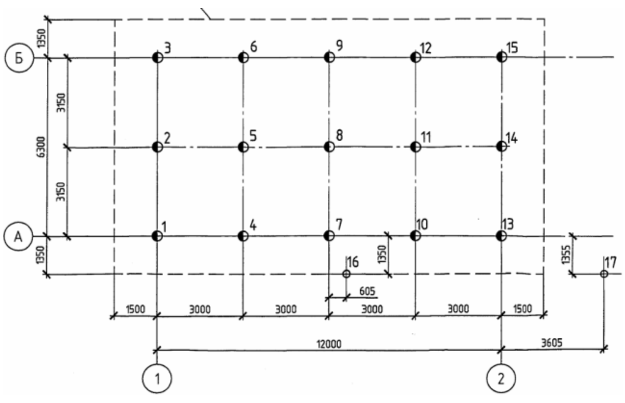

Figure 13. Fragment of the design layout of piles and pile heads

To verify the accuracy of measuring the coordinates of the pile bases using the neural network, these coordinates were manually measured from the orthophotos. The coordinates of each pile in the object coordinate system were measured across all orthophotos (7-10 images) in which it appeared, and the root mean square errors (RMSE) were calculated for both manual and automatic measurements, as shown in Table 2. Points exceeding the allowable coordinate deviations are highlighted in red.

Manual Measurements (m)			Automatic Measurements Using Neural Network (m)		
d_x RMSE	d_y RMSE	d_{xy}	dx RMSE	dy RMSE	d_{xy}
0.08	0.02	0.09	0.08	0.03	0.08
0.10	0.04	0.11	0.11	0.02	0.11
0.07	0.03	0.08	0.07	0.04	0.08
0.10	0.04	0.11	0.08	0.03	0.09
0.08	0.03	0.09	0.07	0.03	0.08
0.07	0.03	0.08	0.05	0.03	0.05
0.08	0.02	0.08	0.07	0.02	0.08
0.07	0.05	0.08	0.07	0.06	0.10
0.04	0.07	0.08	0.07	0.06	0.09
0.03	0.04	0.05	0.04	0.05	0.06
0.08	0.02	0.08	0.06	0.02	0.06
0.08	0.03	0.08	0.07	0.04	0.08
0.06	0.03	0.07	0.05	0.04	0.07
0.06	0.02	0.06	0.04	0.02	0.04
0.05	0.04	0.06	0.06	0.03	0.06
0.08	0.02	0.08	0.07	0.03	0.07
0.05	0.04	0.07	0.06	0.05	0.08
0.06	0.07	0.09	0.06	0.09	0.11
0.04	0.05	0.06	0.04	0.05	0.06
0.00	0.04	0.04	0.01	0.04	0.04
0.04	0.03	0.05	0.03	0.07	0.08
0.02	0.05	0.06	0.06	0.08	0.10
0.09	0.02	0.09	0.07	0.03	0.07
0.04	0.03	0.06	0.06	0.01	0.06
0.05	0.03	0.06	0.07	0.02	0.07
0.05	0.03	0.06	0.06	0.05	0.08
0.05	0.04	0.07	0.07	0.02	0.08
0.05	0.04	0.07	0.04	0.08	0.09
0.01	0.01	0.01	0.03	0.01	0.03
0.06	0.02	0.06	0.06	0.02	0.06
0.05	0.03	0.06	0.05	0.05	0.07
0.06	0.02	0.07	0.05	0.04	0.06
0.05	0.03	0.06	0.04	0.04	0.06
0.06	0.04	0.07	0.04	0.11	0.12
0.06	0.04	0.07	0.05	0.06	0.08
0.05	0.01	0.05	0.04	0.03	0.05
0.07	0.02	0.07	0.04	0.04	0.06
0.03	0.04	0.05	0.06	0.07	0.09
0.04	0.02	0.05	0.10	0.05	0.11
0.03	0.04	0.05	0.05	0.07	0.09

Table 2. Allowable coordinate deviations

As can be seen from this table, the average deviation of the coordinates of the pile bases measured using the neural network from manual measurements was less than 2 cm, i.e., approximately 0.5 of a pixel size. This indicates a sufficiently effective method for automatic detection and measurement of pile coordinates using a neural network. However, to improve the reliability of detecting piles that exceed the tolerance for deviation from design values, aerial photography with better ground resolution should be performed.

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-2-W9-2025-289-2025

4. CONCLUSION

An automated technology for monitoring the position of piles in construction structures has been developed, based on the use of aerial photography from unmanned aerial vehicles and neural networks.

The comprehensive research conducted in this area has yielded compelling evidence demonstrating that this technology is not only sufficiently effective but also versatile in its applications. While its primary focus is on the monitoring of pile positions, the technology can be seamlessly adapted for the monitoring of a wide range of other critical elements found within construction structures. By employing sophisticated algorithms and analytical techniques, the system is capable of comparing the actual coordinates of various construction elements with their corresponding design coordinates. This comparison process is essential for ensuring that construction projects adhere to their intended specifications and standards.

Moreover, the implications of this technology extend beyond just pile monitoring. It opens up new avenues for enhancing overall construction site management and quality assurance. By facilitating real-time assessments and providing detailed insights into the positional accuracy of various structural components, this automated monitoring system contributes to improved decision-making processes among construction professionals. As such, it represents a significant advancement in the field of construction monitoring, offering a reliable solution that can be leveraged for various applications within the industry. This technology not only streamlines monitoring procedures but also enhances the ability to maintain compliance with design specifications, ultimately leading to more successful construction outcomes.

REFERENCES

Achanccaray, P., Gerke, M., Wesche, L., Hoyer, S., Thiele, K., Knufinke, U., and Krafczyk, C.: ON THE ASSESSMENT OF INSTANCE SEGMENTATION FOR THE AUTOMATIC DETECTION OF SPECIFIC CONSTRUCTIONS

Adamtsevich L.A., Vorobyov P.Yu., Zheleznov E.M. Technology for monitoring capital construction objects at life cycle stages using remote sensing methods with unmanned aerial vehicles (drones) based on high-precision digital models of objects // Construction and Architecture. 2021. Vol. 9 No. 3(32). P. 51--55.

Avetisyan R.T., Bilonda T.E., Kazaryan R.R. Prospects for the development of automated construction site monitoring based on drone data // Construction Production. 2019. No. 3. P. 29--32.

Li, Qingyu & Shi, Yilei & Auer, Stefan & Glock, Clemens & Schmitt, Michael & Zhu, Xiao. (2020). Detection of Undocumented Building Constructions from Official Geodata Using a Convolutional Neural Network. Remote Sensing. 12. 10.3390/rs12213537.

Mishchenko A.S. THE USE OF NEURAL NETWORK TECHNOLOGY IN CONSTRUCTION ACTIVITIES // Bulletin of Polotsk State University. Series D. Economic and Legal Sciences. 2024. No. 2. URL: https://cyberleninka.ru/article/n/ispolzovanie-tehnologii-neyronnyh-setey-v-stroitelnoy-deyatelnosti (accessed: 11.08.2024).

National standard of the Russian Federation. Technology of aerial phototopographic survey performed for the purpose of creating topographic maps and plans and ensuring cadastral works. Moscow: JSC "Kodeks", 2021. 54 p.

Vogel A.A., Chakhkiev I.M. Application of an automated monitoring system using unmanned aerial vehicles // World Scientific Research and Development in the Era of Digitalization. Proceedings of the XV International Scientific and Practical Conference. 2021. P. 42--51.

Wu, Tianyong & Dong, Youkou. (2023). YOLO-SE: Improved YOLOv8 for Remote Sensing Object Detection and Recognition. Applied Sciences. 13. 12977. 10.3390/app132412977.

Zhou, K., Chen, Y., Smal, I., Lindenbergh, R., 2019. Building segmentation from airborne VHR images using Mask R-CNN. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42(2/W13)