Hybrid quantum CNN-based information technology for building semantic segmentation in aerial imagery

Vita Kashtan[1,†], Volodymyr Hnatushenko[*,1,†], Dmytro Babets[1,†], Krzysztof Cyran[2,†] and Kamil Wereszczyński[2,†]

[1] *Dnipro University of Technology, Dmytra Yavornytskoho Ave 19, Dnipro, 49005, Ukraine*
[2] *Silesian University of Technology, ul. Akademicka 2A, 44-100 Gliwice, Poland*

Abstract

This paper presents information technology for semantic segmentation of building objects in high-resolution aerospace images using a hybrid quantum convolutional neural network (QCNN). Quantitative and visual evaluations demonstrate the effectiveness of the proposed approach compared to classical segmentation models such as U-Net, CNN, and FCN. The hybrid QCNN model achieved high pixel classification accuracy, with a mean accuracy of 0.98 on the training set and 0.97 on the validation set, indicating robust object recognition. Loss values reached minimal levels (0.05 training, 0.07 validation), confirming efficient training without overfitting. Intersection over Union (IoU) scores were 0.90 and 0.89 for training and validation sets, respectively, demonstrating precise building contour delineation. Testing on diverse urban scenes yielded 100% detection accuracy without false positives or negatives. Training dynamics showed rapid convergence, with mean average precision (mAP) increasing to 0.35−0.45 and mAP@50:95 reaching 0.20−0.30, stabilizing after 50 epochs. The proposed technology effectively segments buildings under varying density, geometric complexity, shadows, and background noise. The hybrid QCNN approach enhances segmentation quality and generalization to new images. The results confirm the practical applicability of the developed technology for geospatial monitoring, urban planning, and mapping based on aerospace imagery.

Keywords

quantum computing, quantum convolutional neural network, semantic segmentation, building objects, aerial imagery, machine learning, computer vision

1. Introduction

The modern availability of very high spatial resolution remote sensing imagery has significantly expanded the possibilities for building object segmentation. Segmentation, as a process of dividing an image into functional areas, is a key stage for further analysis of the spatial structure of urbanized territories. The data obtained as a result of segmentation are used for integration into geographic information systems (GIS) [1], maintaining the relevance of cartographic data, assessing damage in emergency zones [2], monitoring urban infrastructure, and analyzing land use and urbanization processes [3, 4]. However, obtaining such imagery for large areas, particularly cities and urban agglomerations, involves significant time and resource expenditures and organizational challenges [5, 6]. It limits their regular use in operational monitoring. In addition to aerial imagery, satellite remote sensing images have been widely used in monitoring, offering greater area coverage and

higher data update frequency. Nevertheless, satellite data are characterized by increased spectral variability, caused by different imaging conditions, atmospheric effects, and seasonality [7–9]. As a result, the development of universal automated systems for the interpretation and segmentation of satellite images remains challenging, especially under conditions of uncertainty and heterogeneity of input data. Therefore, developing an information technology capable of integrating heterogeneous aerial and satellite imagery data and providing efficient automated segmentation of building objects is highly relevant. This approach enables improved accuracy and timeliness in analyzing spatial changes in urban areas and supports decision-making in urban planning.

2. Related works

Image segmentation is one of the key steps in analyzing and interpreting remote sensing data and has been actively studied in recent years. In aerospace image analysis, building object segmentation is defined as a semantic segmentation task, which involves classifying each image pixel into one of two mutually exclusive categories: 'building' or 'background' [10–12]. A significant number of scientific studies are dedicated to developing and improving segmentation methods, covering both classical approaches based on thresholding and region growing, as well as modern machine learning and deep learning models [13].

Traditional segmentation algorithms use pixel-based or object-oriented features, such as spectral characteristics, texture, shape, and shadows, applying classical algorithms including support vector machines and random forests. They are mainly used for processing relatively small areas due to the need for manual feature extraction. However, these methods have significant limitations: a high dependence on sensor characteristics, imaging conditions, and regional specifics, which leads to reduced stability of the results.

Considering the limitations of traditional segmentation algorithms, modern research is increasingly focused on applying deep learning (DL) methods. DL architectures such as U-Net [14, 15], SegNet [16, 17], DeepLab [18], Convolutional Neural Networks (CNN) [19], and ConvNext [20] have gained popularity. These models automatically extract spatial-contextual features directly from high-resolution input images, integrating the feature extraction stage into the classifier training process. Among the classical architectures, U-Net is a coder-decoder model initially developed for medical imaging that has become particularly popular and effectively adapted for pixel-wise classification of satellite images [21]. The authors in [22] proposed an improved architecture combining U-Net with self-attention and depthwise separable convolutions, achieving 91% accuracy, 5% higher than Dense Plus U-Net.

The authors in work [23] proposed a method combining CNN with a pyramid pooling mechanism for multi-scale feature extraction, significantly improving the recognition of semantic content in aerial photographs. The SegNet model in work [24] was applied for multi-class segmentation of UAV photogrammetric images, resulting in a substantial increase in building detection accuracy. The paper [25] investigated similar semantic segmentation tasks in related fields. The method presented in [26] combines object-oriented and pixel-based analysis for the automatic detection of new constructions, changes, or demolition of buildings, achieving an accuracy of 84–88% and a Kappa statistic of 89–96% based on VHR images in the RGB and NIR ranges. CNN-based models are widely used to improve spatial resolution, among which SRCNN is one of the first networks for image super-resolution [27]. Further improvements include RCAN (Residual Channel Attention Network) with channel attention, which enhances essential features and provides more precise detail in the reconstructed images [28].

In recent years, a promising research direction has been applying quantum neural network models, particularly Quantum Convolutional Neural Networks, which combine the advantages of quantum parallelism and classical convolutional data processing mechanisms. In [29], a comprehensive comparison of QCNN with classical CNN and Artificial Neural Networks was presented, demonstrating the effectiveness of quantum computing for object classification tasks. The authors noted that QCNN models exhibit potentially higher accuracy and efficiency than classical

counterparts, especially as the size of input data and training batches increases, opening up new possibilities for their application in large-scale data environments. In [30], the advantages of applying QCNN for improving modeling efficiency in cases of high-dimensional input data were substantiated, where classical CNNs lose efficiency due to computational resource limitations.

Based on the analysis of existing approaches to building object segmentation, modern methods demonstrate significant progress in addressing semantic segmentation tasks. However, these methods have some limitations when working with remote sensing images that cover large areas, characterized by diverse imaging conditions and complex object structures. In particular, low contrast between buildings and the surrounding environment and the complexity of background elements that vary across space and time significantly complicate accurate segmentation [31]. Typical issues related to local obstructions (e.g., shadows, vegetation, technical structures) and heterogeneous lighting negatively affect the accuracy of building contour delineation. Moreover, many existing methods focus only on extracting the upper parts of buildings, complicating the analysis of complex roof structures and silhouettes in remote sensing imagery. Due to the limitations of traditional segmentation algorithms, this study aims to develop an information technology for semantic segmentation of building objects based on a hybrid Quantum Convolutional Neural Network, considering the spatial characteristics of aerospace images. The proposed technology is focused on improving the accuracy of building object segmentation by accounting for spatial-spectral characteristics and interference variability.

3. Materials and methods

The proposed information technology for building object segmentation on aerospace images is based on combining quantum and classical neural computations within a hybrid quantum-classical convolutional neural network. The structural scheme of the developed technology is presented in Figure 1 and consists of four steps: preprocessing of input data; quantum encoding and feature extraction; classical decoding; and obtaining the final result in the form of a semantically segmented image of building objects.

At the first step, the aerospace image is loaded and preprocessed. The image is normalized to the range [0, 1], then divided into local patches of size 2×2 pixels, each serving as an element of the input dataset for the quantum neural network. This approach allows for extracting local spatial features while maintaining computational efficiency.

The second step involves using a Quantum Convolutional Neural Network (QCNN) for deep feature extraction [30, 32]. The input normalized patch values $\{x_0, x_1, x_2, x_3\}$ are transformed into quantum space using a Data Embedding block. For this purpose, each qubit q_j is subjected to a Hadamard (H) gate to create a superposition state, followed by a parameterized $R_y(x_i)$ rotation that encodes the pixel intensity value into the qubit's rotation angle. A qubit q_j, initialized in the $|0\rangle$ state, is transformed into:

$$|\psi_j\rangle = R_y\left(\frac{q_j\pi}{2}\right)H|0\rangle, \tag{1}$$

where H is the Hadamard gate that creates a superposition state; $R_y(\theta)$ is the rotation gate around the y-axis by an angle θ.

After this step, a set of N=4 input qubits representing an image patch will be in an initial superposition state. At the same time, entanglement between them will be introduced at the subsequent steps using CNOT gates within the convolutional layers.
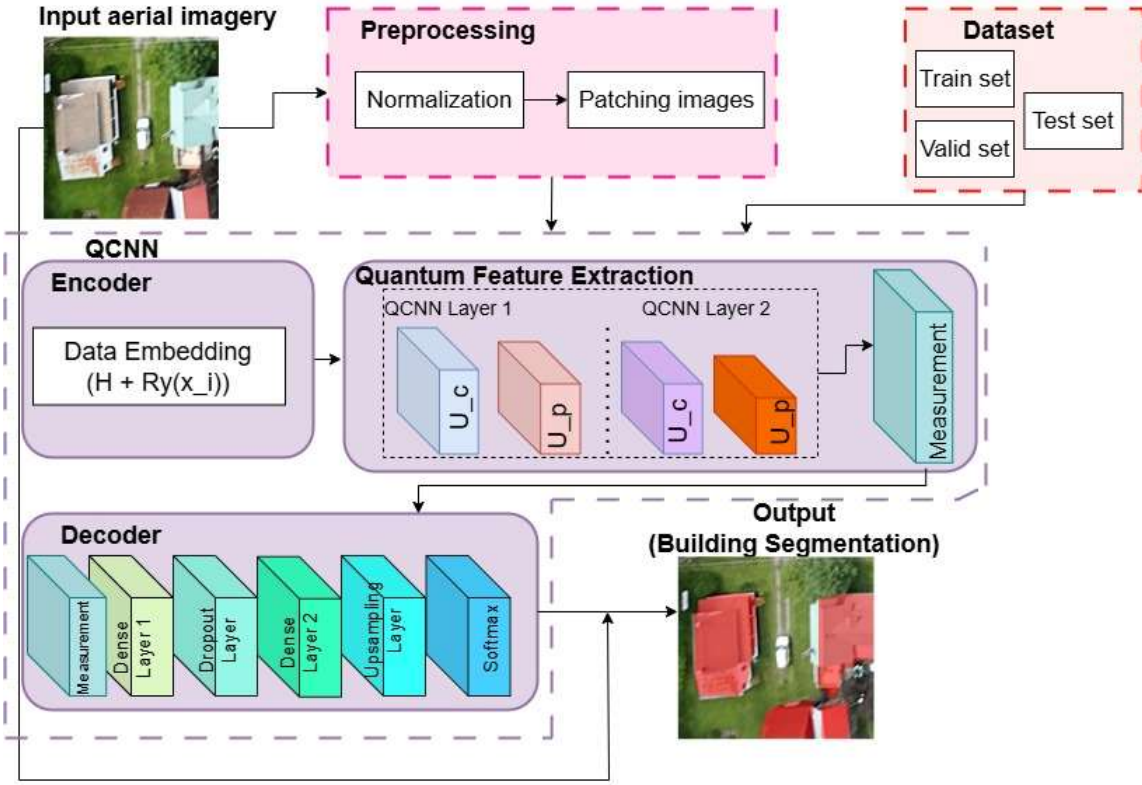
Figure 1: Scheme of the information technology for building object segmentation based on QCNN.

The QCNN encoder consists of a sequence of quantum convolutional (U_c) and quantum pooling (U_p) layers (Figure 2). These layers automatically extract essential features from the quantum-encoded data.
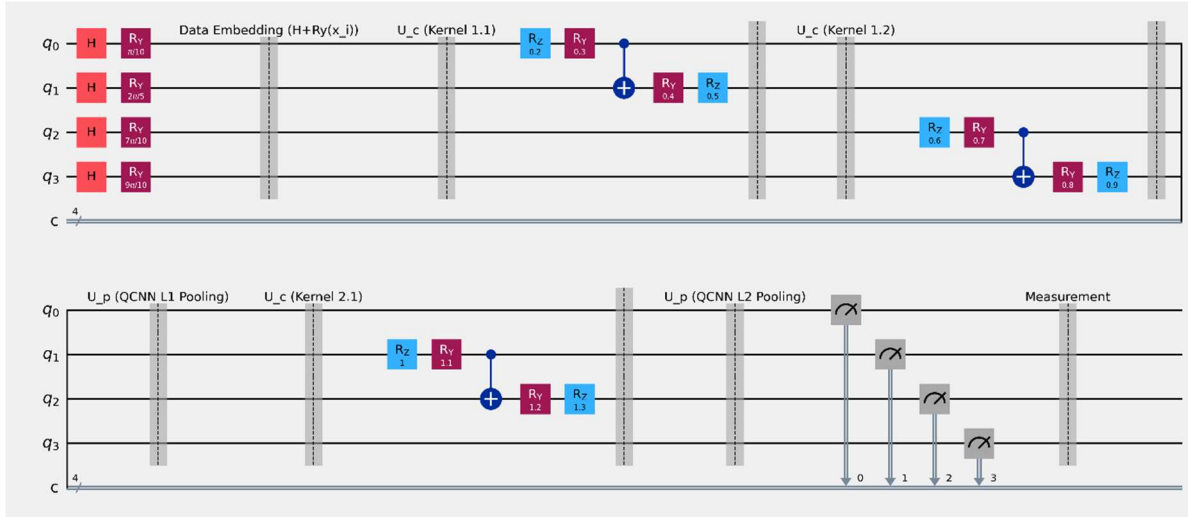

Figure 2: Structural scheme of Quantum Feature Extraction.

The quantum convolutional layers U_c (Kernel 1.1, Kernel 1.2, Kernel 2.1) are analogous to classical convolutional filters. They consist of parameterized single-qubit rotations and two-qubit entangling CNOT gates between neighboring qubits:

$$U_c = R_z(\theta_4)_{q_j} R_y(\theta_3)_{q_j} CNOT_{q_j,q_j} R_y(\theta_2)_{q_i} R_z(\theta_1)_{q_i},$$

(2)

where θ is a vector of trainable parameters; R_z is the rotation gate around the z-axis; $CNOT_{i,j}$ is a controlled-NOT gate, where q_i is the control qubit and q_j is the target qubit.

The output state after the convolutional layer is [32]:

$$|\psi_{out}\rangle = U_c(\theta)\psi_{in}.$$

(3)

The quantum pooling layers U_p (QCNN L1 Pooling, QCNN L2 Pooling) are designed to reduce data dimensionality, which decreases computational complexity and retains only the most significant features necessary for further classification. For a 2×2 patch, after the first pooling layer, the dimensionality is reduced to two qubits through partial measurement. After the second pooling layer, the quantum state is passed to the final measurement layer. Unlike classical pooling, which performs pixel subsampling, quantum pooling uses conditional quantum rotations to aggregate information [32]:

$$|\psi\rangle = \sum_{i,j} c_{i,j}|q_i, q_j\rangle,$$

(4)

$$U_p = RY(\theta \cdot f(m)),$$

(5)

$$U_p(m)_{qj} = \begin{cases} U_A \ if \ m = 0 \\ U_B \ if \ m = 1 \end{cases}$$

(6)

where m is the measurement result of one of the qubits, which serves as a conditional operator for the remaining quantum system; U_A and U_B are different parameterized unitary operations (combinations of R_y, R_z, CNOT gates) applied to qubit q_j depending on the value of m.

This approach effectively reduces the number of active qubits, since the informational contribution of the measured qubit has already been transferred to the state of the remaining quantum system, and the measured qubit itself is excluded from further quantum computation.

The final step of the quantum part of the proposed model is the quantum measurement layer, which transforms the quantum state into the classical feature vector space:

$$p_i = \langle i|\phi\rangle\langle\phi|i\rangle,$$

(7)

where $i=0,...,2^n-1$. The obtained probabilities are used as output features for further processing in the classical part of the neural network to perform segmentation.

In the proposed information technology, the classical decoder takes the measured classical features obtained from quantum processing as input and is implemented as a traditional neural network. The decoder architecture includes fully connected layers for further nonlinear feature transformation, Dropout layers that provide regularization and prevent overfitting, and upsampling layers that restore the spatial structure of features using transposed convolutions or bilinear interpolation. The output layer with a Softmax activation function forms the segmented image, where each pixel is classified as belonging to building objects or background.

Training of the hybrid quantum-classical neural network is performed jointly for the parameters of the quantum gates and the weights of the classical decoder. Optimization is carried out by gradient descent, where the parameter-shift rule is applied to compute gradients of quantum parameters. This method allows efficient calculation of the derivative of the expected value of the Hamiltonian operator $\langle H \rangle$:

$$\frac{\partial\langle H\rangle}{\partial\theta} = \frac{1}{2}[\langle H\rangle(\theta+s) - \langle H\rangle(\theta-s)],$$

(8)

where s is a fixed parameter shift angle characteristic of the quantum gate structure, this approach ensures integrated training of both quantum and classical components of the network, improving the accuracy of building object segmentation on aerospace images.

Thus, the hybridity of the proposed information technology lies in the combination of quantum and classical neural computations within a single segmentation of architecture. In particular, the quantum part of the model, the Quantum Convolutional Neural Network, is responsible for encoding input images into quantum space and deep feature extraction using parameterized quantum gates

and entanglement mechanisms. These quantum computations enable the model to effectively recognize complex nonlinear dependencies and interactions within local image patches that are difficult to implement with classical methods. Following the quantum layer is the classical decoder. This traditional neural network takes the classical features obtained from the quantum state measurement as input and performs feature transformation and spatial reconstruction to form the segmented image. Joint training of the quantum and classical component parameters ensures effective operation coordination, which promotes increased accuracy and robustness of segmentation on aerospace images.

4. Experimental results

4.1. Input aerospace data

To validate the effectiveness of the proposed information technology for building object segmentation, high spatial resolution aerospace images were used (350 pixels per inch both horizontally and vertically), obtained using a SONY DSC-WX220 camera. The images have a frame size of 4896 × 3672 pixels, a color depth of 24 bits, an sRGB color space, and a 3 bits per pixel compression rate. Shooting parameters include an aperture of f/3.3, shutter speed of 1/800 s, ISO 100, no exposure compensation, a focal length of 4 mm, a maximum aperture of 3.45, and matrix metering mode for exposure measurement; geographic coordinates (latitude 44° 20' 20.21", longitude 72° 44' 57.68", altitude 273.44 m above sea level), which allows for territorial referencing of the results.

4.2. Dataset preparation

To ensure the representativeness of the sample and an objective assessment of the developed neural network's effectiveness, a dataset consisting of 376 aerospace images was compiled. The sample covers a variety of spatial conditions and types of building structures, allowing the model to be tested in different scenarios. Each image was preprocessed by manual annotation with the creation of binary masks. In every image, buildings were separated from the background at the pixel level. The availability of high-quality annotations ensures the precise formation of the training sample for semantic segmentation tasks and enables quantitative validation of the model's results. For training, validation, and testing, the dataset was divided into three subsets with proportions that provide sufficient data for training and proper evaluation of the network's generalization ability: training set – 266 images (approximately 70%); validation set – 55 images (approximately 15%); test set – 55 images (approximately 15%). Before being fed into the neural network, all images undergo preprocessing, including normalizing pixel values and splitting images into patches, as described in Section 3. The utilized dataset provides sufficient variability in building types and shooting conditions to verify the effectiveness of the hybrid QCNN model for aerospace image segmentation tasks.

4.3. Performance analysis of the QCNN segmentation model

The effectiveness of the proposed neural segmentation model for building objects was evaluated based on the analysis of key accuracy metrics and loss functions during model training on the prepared set of aerospace images. During training, a steady increase in mean Average Precision (mAP), a fundamental indicator of building object segmentation quality, was observed. In particular, the mAP@0.5 metric showed rapid growth during the first 50 epochs, followed by stabilization within the range of 0.35−0.45. It indicates effective model formation for detecting building objects at the baseline IoU threshold 0.5. The mAP@0.5:0.95 metric, which accounts for precision across multiple overlap thresholds, stabilized at 0.2−0.3, confirming the model's ability to correctly localize objects under stricter accuracy criteria (Fig. 3).
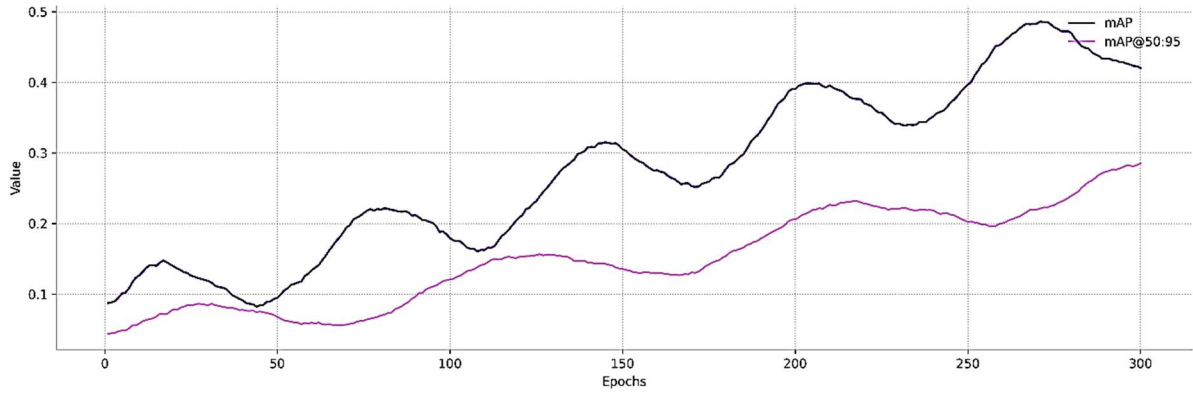
Figure 3: Graph of mean Average Precision (mAP and mAP@0.5:0.95) metrics during model training over 300 epochs.

Additionally, an analysis of the components of the model's total loss function was conducted (Fig. 4). The Box Loss, which reflects the accuracy of bounding box localization, gradually decreases and stabilizes within the range of 1.7–1.8, demonstrating the model's consistent ability to delineate building object boundaries accurately. The Class Loss, characterizing classification correctness, decreases to 2.2–2.4, indicating practical training of the model to recognize object categories. The Object Loss, responsible for the reliability of object detection, reaches stable values within 1.4–1.5, confirming the model's effectiveness in distinguishing building objects from the background.
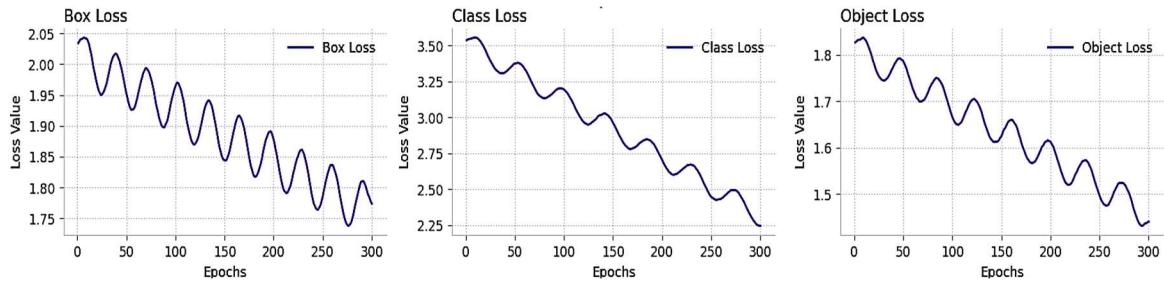

Figure 4: Loss function graphs of the model (Box Loss, Class Loss, Object Loss) during the training phase.

5. Discussion

5.1. Visual analysis of building object segmentation results

To comprehensively evaluate the effectiveness of the proposed information technology for building object segmentation, a visual assessment of the results was performed on aerospace images representing various types and densities of building structures. Figure 5 presents segmentation results for an area characterized by dense urban development.

The analysis showed that the developed hybrid QCNN model (Fig. 5b) provides more accurate reproduction of building geometry, correctly delineating their contours even under complex urban conditions. Images in Fig. 5c, Fig. 5d, and Fig. 5e exhibit false positive errors (background identified as building and highlighted with a yellow contour), which are almost absent in Fig. 5b.

Figure 6 shows results for an area with low-density development, typical of suburban and rural territories. The results demonstrate that the proposed hybrid QCNN model maintains high segmentation accuracy under variable background conditions, including green vegetation and road infrastructure. Classical architectures such as U-Net (Fig. 6c), CNN (Fig. 6d), and FCN (Fig. 6e) show a higher number of false positive detections, specifically, parts of roads or shadowed areas were incorrectly classified as buildings and highlighted with yellow contours in the images. The proposed technology demonstrated an improved ability to distinguish buildings against a complex background and reduced the risk of confusing natural and artificial objects.
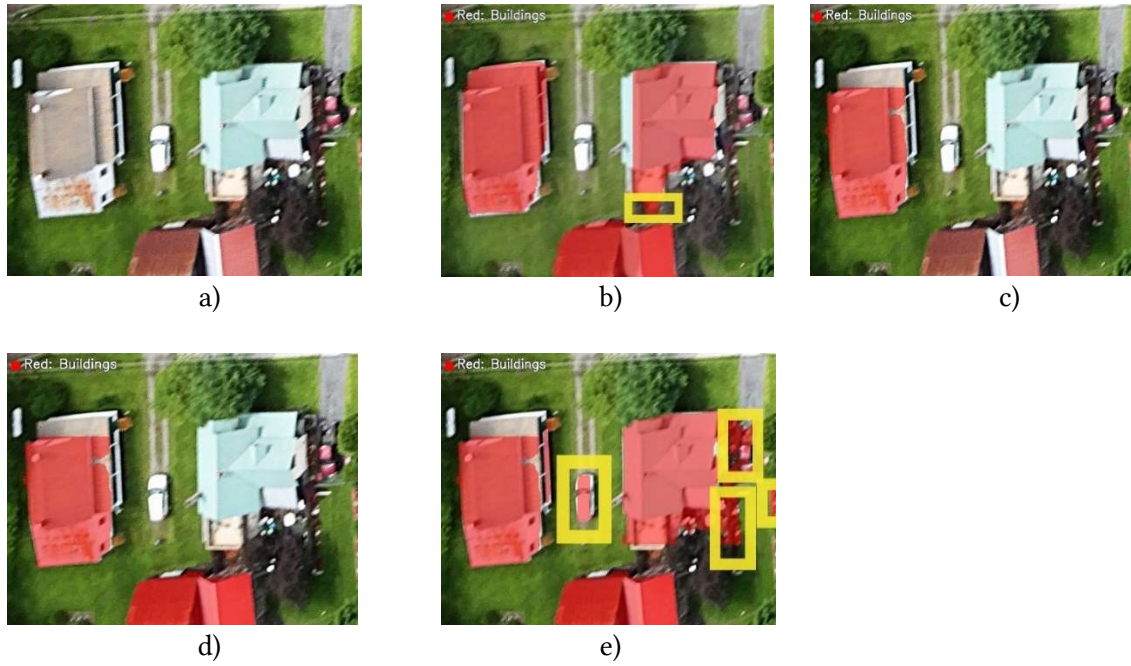
Figure 5: Experiment 1 – a) original aerial image; b) proposed hybrid QCNN; c) U-Net d) CNN; e) FCN.
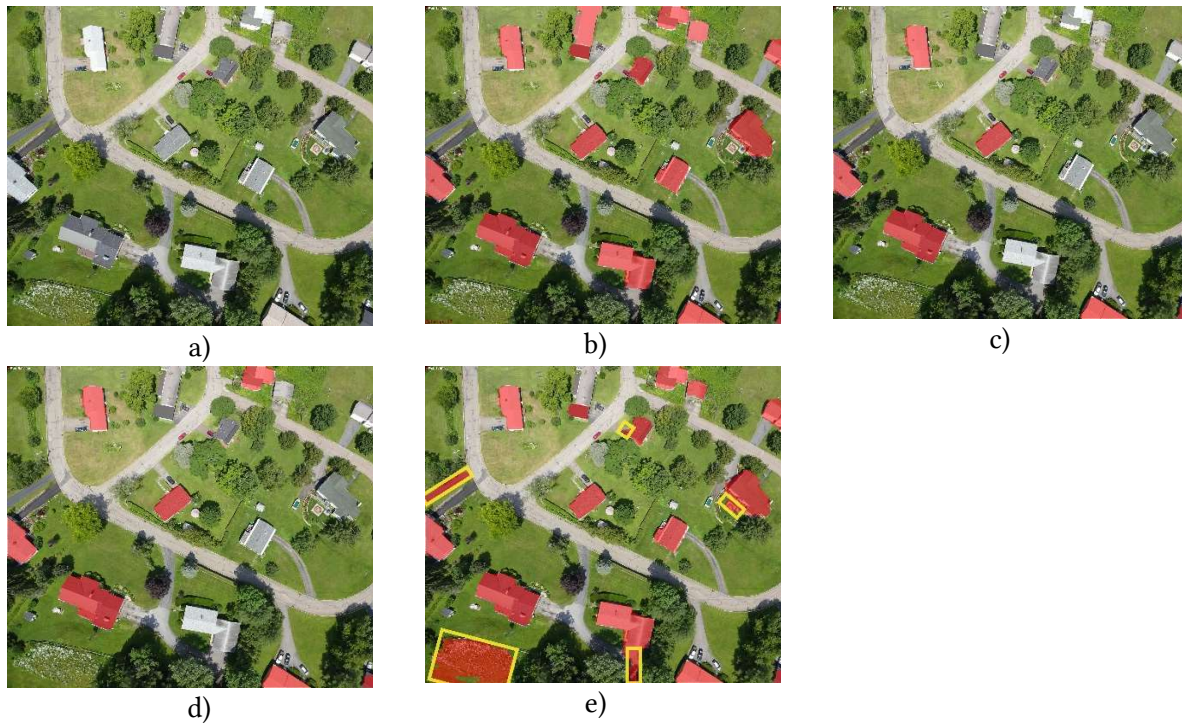

Figure 6: Experiment 2 – a) original aerial image; b) proposed hybrid QCNN; c) U-Net d) CNN; e) FCN.

Figure 7 shows an example of a test area with complex building morphology, characterized by various interfering factors, including shadow zones, dense vegetation, and small-scale technical structures. The proposed QCNN model (Fig. 7b) ensures high accuracy of spatial localization of buildings even under significant structural heterogeneity of the scene. When using alternative models (Fig. 7c, d, e), partial merging of buildings with shadow areas, loss of detection of individual small objects, and false positive detections are observed, manifested as erroneous segmented areas marked with yellow boxes in the resulting images. The visual analysis confirms the advantage of the

proposed hybrid QCNN model in accurately segmenting buildings with complex geometry and its increased robustness to background variability.

All three experiments demonstrated the advantages of the proposed technology over classical approaches in terms of the spatial localization of buildings, the accuracy of building object segmentation on aerospace images, and the minimization of segmentation errors. The results confirm the QCNN model's capability to perform effectively in densely built-up urban areas and in conditions of sparse or complex building structures.

| a) | b) | c) |

| d) | e) |

Figure 7: Experiment 3 – a) original aerial image; b) proposed hybrid QCNN; c) U-Net d) CNN; e) FCN.

5.2. Quantitative analysis of building object segmentation results

A quantitative analysis of building object segmentation results was performed under three experimental scenarios with varying building density and structural complexity to compare the effectiveness of existing neural network segmentation models. The comparison was conducted using three key metrics: the number of correctly detected objects (True Positives, TP), the number of false detections (False Positives, FP), and the number of missed objects (False Negatives, FN) [33]. The results are presented in Table 1.

The results of the quantitative analysis are consistent with the conclusions of the visual assessment (Section 5.1) and confirm the advantage of the proposed hybrid QCNN model over classical segmentation architectures. In particular, the developed model demonstrated consistently high accuracy across all experimental cases, achieving complete detection of all building objects (TP = 100%) without false positives (FP) or false negatives (FN). In contrast, the U-Net and CNN models showed many missed objects, especially in complex or sparse building structures. For example, in the second experiment, U-Net identified only 5 out of 16 objects (FN = 11), and CNN identified 6 out of 16 (FN = 10), indicating insufficient sensitivity of these models to objects with complex geometry or visual characteristics. The FCN model showed better detection completeness in the first and third experiments (TP = 100%), but its effectiveness was reduced due to many false detections. Specifically, in the second and third experiments, the number of FP cases were 5 and 6, respectively, leading to excessive detection of background objects as buildings. The analysis of results in Figures 5–7

confirmed that these errors mainly occurred due to the misclassification of cars, roads, and vegetation.

Table 1

Quantitative Analysis of Building Object Segmentation Performance

Model / Class	Experiment 1 (4 objects)	Experiment 2 (16 objects)	Experiment 3 (10 objects)
U-Net			
Buildings Detected (TP)	1	5	7
False Positives (FP)	0	0	1
Missed Buildings (FN)	3	11	3
CNN			
Buildings Detected (TP)	2	6	7
False Positives (FP)	0	0	0
Missed Buildings (FN)	2	10	3
FCN			
Buildings Detected (TP)	4	13	10
False Positives (FP)	1	5	6
Missed Buildings (FN)	0	3	0
Proposed Hybrid QCNN			
Buildings Detected (TP)	4 (100%)	16 (100%)	10 (100%)
False Positives (FP)	0	0	0
Missed Buildings (FN)	0	0	0

Thus, the obtained quantitative results demonstrate the high accuracy, sensitivity, and specificity of the proposed hybrid QCNN model. Its ability to minimize missed detections and false positives indicates effective adaptation to aerospace images' variable spatial and spectral characteristics. It highlights the feasibility of its application for high-precision geospatial research and building monitoring tasks under various types of urbanization.

Table 2 presents the results of the quantitative analysis of the segmentation models' performance based on key efficiency metrics for both the training and validation datasets. The study of the obtained results confirms the significant advantage of the proposed hybrid QCNN model compared to classical architectures such as U-Net, CNN, and FCN. The hybrid model demonstrates the best performance across all key indicators, including classification accuracy, loss function, and Intersection over Union (IoU).

Table 2

Performance Metrics Summary

Metric	U-Net	CNN	FCN	Proposed Hybrid QCNN
Training Accuracy	0.88	0.89	0.92	0.98
Validation Accuracy	0.85	0.86	0.89	0.97
Training Loss	0.25	0.22	0.18	0.05
Validation Loss	0.28	0.26	0.21	0.07
Training IoU	0.70	0.72	0.78	0.90
Validation IoU	0.67	0.69	0.75	0.89

In particular, the model's accuracy on the training set for QCNN reaches 0.98 and 0.97 on the validation set, exceeding the corresponding values of the classical models, which range from 0.85 to 0.92. It indicates the proposed model's ability to classify pixels with minimal errors accurately. The loss function values (Loss) for QCNN are the lowest among all models—0.05 on the training set and 0.07 on the validation set, demonstrating the stability of the training process and the model's ability to closely approximate segmentation masks to the real data. The IoU index values further confirm the superiority of the proposed model in the accuracy of building object segmentation. For the hybrid QCNN, the IoU reaches 0.90 on the training set and 0.88 on the validation set, indicating high precision in spatial delineation of building contours. In contrast, classical models show lower IoU values (0.67–0.75), suggesting more segmentation inaccuracies, including partial omission of objects or inclusion of background areas.

The overall analysis of the obtained metrics confirms the high efficiency of the proposed information technology based on the hybrid QCNN model for building object segmentation on aerospace images. The results demonstrate superior generalization ability, training stability, and reduced error rates under challenging image variability conditions.

6. Conclusions

In this study, an information technology for semantic segmentation of building objects on high spatial resolution aerospace images was developed and experimentally validated using a hybrid quantum convolutional neural network. The quantitative analysis and visual assessment results confirm the proposed approach's effectiveness compared to classical segmentation models, including U-Net, CNN, and FCN. The proposed hybrid QCNN model achieved higher pixel classification accuracy: the average accuracy on the training dataset reached 0.98 and 0.97 on the validation set, demonstrating the model's stable ability to recognize building objects accurately. The loss function values reached minimum levels of 0.05 on the training set and 0.07 on the validation set, indicating efficient model training without signs of overfitting. The Intersection over Union metrics confirmed the model's ability to delineate building contours accurately: the IoU value was 0.90 on the training set and 0.89 on the validation set. While testing images with various building types, the model consistently identified all building objects with 100% detection accuracy, without any false positives or negatives. Analysis of the training process dynamics showed a rapid increase in mAP to 0.35–0.45 and mAP@50:95 to 0.20–0.30, with subsequent stabilization after the first 50 epochs, indicating efficient model convergence.

The proposed Hybrid quantum CNN-based information technology for building semantic segmentation in aerial imagery demonstrated improved segmentation accuracy of buildings under varying building density conditions, complex geometric forms, shadow effects, and background interference. Applying the hybrid QCNN-based approach improved the quality of building object extraction and enhanced the model's generalization ability on new images. The results confirm the feasibility of using the developed technology for geospatial monitoring, urban planning, and mapping tasks based on aerospace imagery.

Acknowledgements

Declaration on Generative AI

The authors used Grammarly to check the grammar and spelling. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] W. Li, C. He, J. Fang, J. Zheng, H. Fu, L. Yu, Semantic Segmentation-Based Building Footprint Extraction Using Very High-Resolution Satellite Images and Multi-Source GIS Data, Remote Sens. 11.4 (2019) 403. doi:10.3390/rs11040403.

[2] Q. D. Cao, Y. Choe, Building damage annotation on post-hurricane satellite imagery based on convolutional neural networks, Nat. Hazards 103.3 (2020) 3357–3376. doi:10.1007/s11069-020-04133-2.

[3] Q. Hu, L. Zhen, Y. Mao, X. Zhou, G. Zhou, Automated building extraction using satellite remote sensing imagery, Autom. Constr. 123 (2021) 103509. doi:10.1016/j.autcon.2020.103509.

[4] V. Hnatushenko, D. Mozgovyi, V. Vasyliev, Satellite Monitoring of Deforestation as a Result of Mining. Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu. (2017). № 5 (161), pp. 94-99.

[5] R. Sharma, R. Pandey, A. Nigam, Real Time Object Detection on Aerial Imagery, in: Computer Analysis of Images and Patterns, Springer International Publishing, Cham, (2019), pp. 481–491. doi:10.1007/978-3-030-29888-3_39.

[6] L. Lindner, O. Sergiyenko, M. Rivas-Lopez, M. Ivanov, J. C. Rodriguez-Quinonez, D. Hernandez-Balbuena, W. Flores-Fuentes, V. Tyrsa, F. N. Muerrieta-Rico, P. Mercorelli. "Machine vision system errors for unmanned aerial vehicle navigation". 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE), IEEE, (2017). doi:10.1109/isie.2017.8001488.

[7] Y. I. Shedlovska, V. V. Hnatushenko. "Shadow detection and removal using a shadow formation model". IEEE First International Conference on Data Stream Mining & Processing (DSMP), IEEE, (2016). doi:10.1109/dsmp.2016.7583537.

[8] S. Elbaz, E. Sheffer, I. M. Lensky, N. Levin, The Impacts of Spatial Resolution, Viewing Angle, and Spectral Vegetation Indices on the Quantification of Woody Mediterranean Species Seasonality Using Remote Sensing, Remote Sens. 13.10 (2021) 1958. doi:10.3390/rs13101958.

[9] V. Hnatushenko, Vik. Hnatushenko, A. Kavats, V. Shevchenko, Pansharpening technology of high resolution multispectral and panchromatic satellite images. Scientific bulletin of National Mining University, State Higher Educational Institution, National Mining University, (2015). № 4 (148), pp. 91-98.

[10] A.-J. Gallego, P. Gil, A. Pertusa, R. B. Fisher, Semantic Segmentation of SLAR Imagery with Convolutional LSTM Selectional AutoEncoders, Remote Sens. 11.12 (2019) 1402. doi:10.3390/rs11121402.

[11] Y. Shedlovska, V. Hnatushenko, V. Kashtan. "Satellite imagery features for the image similarity estimation". IEEE International Young Scientists' Forum on Applied Physics and Engineering (YSF), IEEE, (2017). doi:10.1109/ysf.2017.8126673.

[12] J. Kim, Y. Kim, Integrated Framework for Unsupervised Building Segmentation with Segment Anything Model-Based Pseudo-Labeling and Weakly Supervised Learning, Remote Sens. 16.3 (2024) 526. doi:10.3390/rs16030526.

[13] V. Kashtan, Y. Radionov, V. Hnatushenko. "Aircraft detection in aerial imagery based on YOLO architectures". ISW-2025: Intelligent Systems Workshop at 9th International Conference on Computational Linguistics and Intelligent Systems, (2025), Kharkiv, Ukraine, pp.196-208. https://ceur-ws.org/Vol-3983/paper15.pdf.

[14] Z. Kokeza, M. Vujasinović, M. Govedarica, B. Milojević, G. Jakovljević, Automatic building footprint extraction from UAV images using neural networks, Geod. Vestn. 64.04 (2020) 545–561. doi:10.15292/geodetski-vestnik.2020.04.545-561.

[15] I. Ahmed, M. Ahmad, G. Jeon, A real-time efficient object segmentation system based on U-Net using aerial drone images, J. Real-Time Image Process. (2021). doi:10.1007/s11554-021-01166-z.

[16] A. Norelyaqine, R. Azmi, A. Saadane, Architecture of Deep Convolutional Encoder-Decoder Networks for Building Footprint Semantic Segmentation, Sci. Program. 2023 (2023) 1–15. doi:10.1155/2023/8552624.

[17] D. Femi, M. A. Mukunthan, Optimized encoder-decoder cascaded deep convolutional network for leaf disease image segmentation, Network (2024) 1–27. doi:10.1080/0954898x.2024.2326493.

[18] H. Si, Z. Shi, X. Hu, Y. Wang, C. Yang, Image semantic segmentation based on improved DeepLab V3 model, Int. J. Model., Identif. Control 36.2 (2020) 116. doi:10.1504/ijmic.2020.116199.

[19] P. Badhani, Deep CNN architectures building blocks for implementing automatic brain tumor segmentation, Int. J. Mech. Eng. 8 (2023). doi:10.56452/7-4-238.

[20] B. Fu, X. Sun, S. Ma, X. Ma, Z. Liu, CSEU-Net: ConvNeXt-SE-U-Net for river ice floe segmentation using unmanned aerial vehicle grayscale remote sensing images, J. Appl. Remote Sens. 18.04 (2024). doi:10.1117/1.jrs.18.046505.

[21] Z. Kokeza, M. Vujasinović, M. Govedarica, B. Milojević, G. Jakovljević, Automatic building footprint extraction from UAV images using neural networks, Geod. Vestn. 64.04 (2020) 545–561. doi:10.15292/geodetski-vestnik.2020.04.545-561.

[22] B. A. Khan, J.-W. Jung, Semantic Segmentation of Aerial Imagery Using U-Net with Self-Attention and Separable Convolutions, Appl. Sci. 14.9 (2024) 3712. doi:10.3390/app14093712.

[23] K. Yue, L. Yang, R. Li, W. Hu, F. Zhang, W. Li. "TreeUNet: Adaptive Tree convolutional neural networks for subdecimeter aerial image segmentation". ISPRS J. Photogramm. Remote Sens. 156 (2019) 1–13. doi:10.1016/j.isprsjprs.2019.07.007.

[24] W. Boonpook, Y. Tan, B. Xu, Deep learning-based multi-feature semantic segmentation in building extraction from images of UAV photogrammetry, Int. J. Remote Sens. 42.1 (2020) 1–19. doi:10.1080/01431161.2020.1788742.

[25] A. Mehra, M. Mandal, P. Narang, V. Chamola. "ReViewNet: A Fast and Resource Optimized Network for Enabling Safe Autonomous Driving in Hazy Weather Conditions". IEEE Trans. Intell. Transp. Syst. (2020) 1–11. doi:10.1109/tits.2020.3013099.

[26] D. Jovanović, M. Gavrilović, D. Sladić, A. Radulović, M. Govedarica, Building Change Detection Method to Support Register of Identified Changes on Buildings, Remote Sens. 13.16 (2021) 3150. doi:10.3390/rs13163150.

[27] J. Liu, W. Zhang, Y. Tang, J. Tang, G. Wu. "Residual Feature Aggregation Network for Image Super-Resolution". 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2020. doi:10.1109/cvpr42600.2020.00243.

[28] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, Y. Fu, Image Super-Resolution Using Very Deep Residual Channel Attention Networks, in: Computer Vision – ECCV 2018, Springer International Publishing, Cham, 2018, pp. 294–310. doi:10.1007/978-3-030-01234-2_18.

[29] G. Meedinti, K. Srirekha, R. Delhibabu, A Quantum Convolutional Neural Network Approach for Object Detection and Classification. (2023).10.48550/arXiv.2307.08204.

[30] S. Oh, J. Choi, J. Kim. "A Tutorial on Quantum Convolutional Neural Networks (QCNN) ".2020 International Conference on Information and Communication Technology Convergence (ICTC), IEEE, 2020. doi:10.1109/ictc49870.2020.9289439.

[31] V. Hnatushenko, P. Kogut, M. Uvarov, Variational approach for rigid co-registration of optical/SAR satellite images in agricultural areas, J. Comput. Appl. Math. 400 (2022) 113742. doi:10.1016/j.cam.2021.113742.

[32] J. M. Zollner, P. Walther, M. Werner, Satellite Image Representations for Quantum Classifiers, Datenbank-Spektrum (2024). doi:10.1007/s13222-024-00464-7.

[33] V. Hnatushenko, D. Mozgovoy, V. Vasyliev. "Accuracy evaluation of automated object recognition using multispectral aerial images and neural network". Tenth International Conference on Digital Image Processing (ICDIP 2018), SPIE, (2018). doi:10.1117/12.2502905.