https://doi.org/10.33271/nvngu/2025-5/112

V. Yu. Kashtan[1],
orcid.org/0000-0002-0395-5895,
V. V. Hnatushenko[*1],
orcid.org/0000-0003-3140-3788,
K. Wereszczyński[2],
orcid.org/0000-0003-1686-472X,
K. Cyran[2],
orcid.org/0000-0003-1789-4939

1 – Dnipro University of Technology, Dnipro, Ukraine
2 – Silesian University of Technology, Gliwice, Republic of Poland
[*] Corresponding author e-mail: vvgnat@ukr.net

QUANTUM MACHINE LEARNING FOR FUSION OF MULTICHANNEL OPTICAL SATELLITE IMAGES

Purpose. To develop a novel approach for fusion of optical satellite images based on machine learning and quantum optimization for integrating spatial-spectral information from RGB and IR channels.

Methodology. The proposed approach involves sequential processing of input data, including geometric, radiometric, and atmospheric corrections. Each channel is decomposed into low-frequency and high-frequency components using a Gaussian filter. The Independent Component Analysis (ICA) method is applied to reduce the dimensionality of input data. A quantum optimizer approximation algorithm is applied to analyze the infrared channel. A deep convolutional neural network with residual dense blocks is used to extract spatial structural features from RGB channels. After integrating features through fully connected layers, the quantum block optimizes the weight coefficients for the final channel fusion.

Findings. Quantitative evaluation demonstrates that the proposed approach outperforms classical fusion methods, including Brovey, Gram-Schmidt, IHS, HCS, HPFC, ATWT, PCA, and CNN, in spectral and spatial information integration accuracy. The method achieves the lowest mean squared error (MSE = 191.8), high structural similarity index (SSIM = 0.43), entropy (Entropy = 7.54), and Sobel filter range (Sobel Sharp = 19.19–21.67 across R, G, B channels). Visual analysis also confirms qualitative advantages: images exhibit clear structure without artifacts and balanced color reproduction consistent with the spectral characteristics of the original RGB data.

Originality. A novel approach to utilizing information of the IR channel is proposed, which integrates a quantum-classical algorithm within a deep convolutional neural network architecture for synergistic processing of multichannel optical images using multilevel frequency decomposition and a weighted feature fusion mechanism.

Practical value. The proposed approach can be implemented in Earth remote sensing systems to enhance the quality of satellite image processing, particularly for mapping, land resource monitoring, agricultural control, and environmental analysis tasks. Applying quantum algorithms opens new opportunities for improving efficiency and accuracy in processing multidimensional geoinformation data containing IR channel information.

Keywords: *quantum machine learning, data fusion, neural network, satellite imagery*

Introduction. Image fusion is one of the approaches in remote sensing data processing and computer vision that involves integrating multiple images of the same scene, acquired from different sensors, at other times, or under varying imaging conditions, to create a single, more informative image for further analysis [1]. The image fusion process aims to preserve meaningful information from all input sources while simultaneously reducing redundancy and eliminating non-informative data. As a result, an image is produced with enhanced spatial, spectral, and temporal characteristics, which improves accuracy in scene content interpretation, object classification, and decision-making based on the obtained data [2, 3]. In satellite imaging, no single optical sensor can simultaneously provide high spatial, spectral, and temporal resolution due to technical limitations of the sensor hardware. For instance, sensors with high spatial resolution typically offer a limited number of spectral bands, whereas multispectral remote sensing systems provide broader spectral coverage but with lower spatial detail. Therefore, image fusion becomes particularly relevant, combining the strengths of different data types while minimizing their limitations. It is essential for land cover mapping [4], drought monitoring, flood extent assessment, and others. Fusion algorithms must meet two key requirements: on the one hand, they should preserve essential image features (such as textures and contours), and on the other hand, they should minimize distortions in the form of artifacts, such as noise or loss of detail.

An important direction in the development of satellite image fusion methods is the integration of the infrared (IR) channel [5] together with the visible spectrum (RGB), which enhances the spectral and spatial informativeness of the image. Infrared data contain features related to the thermal characteristics of the surface that are not available in the visible spectrum, especially under low lighting conditions, fog, or smoke. It opens up additional opportunities for analyzing and monitoring complex environments, particularly in applications such as land surface change detection, wildfire monitoring, soil moisture assessment, vegetation condition analysis, and others.

The fusion of RGB and IR data enables a synergistic effect: on the one hand, it preserves the high level of texture and contour detail characteristic of the visible spectrum; on the other hand, the IR channel provides additional spectral features that enhance object recognition, improve classification accuracy, and ensure better informativeness under adverse imaging conditions. Therefore, the effective integration of the infrared channel is a technical challenge and a crucial requirement for developing more informative remote sensing systems. However, effective integration of RGB and IR data remains challenging due to differences in channel characteristics, especially when processing large volumes of multispectral images. In this context, quantum computing is gaining particular relevance, as it opens new possibilities for accelerated and efficient image fusion by leveraging the principles of superposition and parallelism. Quantum algorithms enable the processing of large volumes of visual information with high accuracy and efficiency, which is especially important for remote sensing data processing.

Literature review. Image fusion is an essential field within image processing that holds significant potential for application in remote sensing, medical diagnostics, military intelligence, and more, due to its ability to enhance the informativeness of images. Of particular scientific interest are multispectral fusion methods, which allow data integration from different channels to improve the reliability of detecting structural, spectral, and spatial features of objects in satellite imagery. There are two main classes of image fusion methods: spatial methods, which process images directly at the pixel level, and transform-based methods, which operate in alternative representation domains (such as frequency, wavelet, discrete cosine, or curvelet transforms).

Spatial domain methods focus on directly combining pixel-level information and derived features within images. These methods are classified based on the level of data representation into three subclasses: pixel-level fusion, feature-level fusion, and decision-level fusion. Pixel-level fusion methods [6–8] integrate information directly at the level of the intensity values of corresponding pixels. Although this approach is relatively simple to implement, its effectiveness significantly decreases in the presence of geometric distortions or noise in the input images, which can degrade the quality of the result. Feature-level fusion methods [9] are based on the preliminary extraction of relevant image characteristics, such as texture, morphological, or geometric features, and their subsequent integration, which improves algorithm robustness against inhomogeneities and artifacts in the source data. Decision-level fusion methods [10] involve the independent processing of each data channel and the integration of results at the decision-making level. This approach efficiently classifies objects, phenomena, or land cover types, where fusion is performed based on probabilistic estimates.

The methods that involve transforming images into alternative representation domains include wavelet transform-based fusion methods, curvelet transform-based methods, and discrete cosine transform-based methods. Wavelet transform-based fusion methods [11–13], which provide multilevel image decomposition according to spatial frequencies, allow for adaptive integration of details at different scales. Methods that utilize curvelet transforms [14] more effectively capture directional and curvilinear structures, which is particularly important when processing images containing linear or arc-shaped objects, such as roads or field boundaries. Discrete cosine transform-based methods [15] are characterized by high efficiency in detecting periodic structures and reducing redundant information, making them well-suited for image compression and signal filtering tasks.

With modern systems' increasing computational power and large-scale datasets' growing availability, deep learning-based image fusion methods are gaining widespread adoption. These approaches are generally divided into supervised [16] and unsupervised [17] methods. Among the supervised methods is the approach proposed in [18], which implements a deep convolutional neural network for multi-focus image fusion. The model architecture includes a multi-scale feature extraction block and a visual attention mechanism, enabling the adaptive selection of the most relevant regions for integration. In [19], a model for satellite image fusion is presented that utilizes a dual-branch deep neural network architecture to separately extract spectral and spatial features, enhanced by a residual learning mechanism. This approach enables the creation of links between data with different resolutions, contributing to improved fusion results. Another example is the method proposed in [20], which introduces the Laplacian Sharpening Network model that includes a high-frequency modification block for extracting detailed spatial information from panchromatic images. To improve the visual quality of the synthesized result, a perceptual loss function is applied along with model optimization that accounts for high-level infrared features. Generative Adversarial Networks (GANs), as introduced in [21], are also actively explored in the context of satellite image fusion. While they demonstrate impressive performance, they require large volumes of annotated data and substantial computational resources and present significant challenges during training. In [22], an innovative unsupervised self-supervised learning approach to multi-focus image fusion is presented. Due to insufficient annotated datasets, the authors use a pretext task to restore high-resolution images from low-resolution inputs to pretrain the network. The model includes a residual feature extraction network and a fusion module that uses activity level estimation and boundary refinement to merge focused regions accurately. Similarly, in

[23], the authors propose an unsupervised GAN-based approach with gradient and adaptive constraints, where the decision-making block identifies the degree of pixel focus by modeling re-blurring.

A promising area attracting particular attention is quantum image processing, which lies at the intersection of quantum computing and digital visual data processing. The development of this field is driven by the potential of quantum algorithms to accelerate computationally intensive tasks that are traditionally resource-demanding for classical systems. In [24], a surface segmentation method based on hybrid quantum-classical optimization is proposed, combining a graph-theoretic approach to graph partitioning with smoothness constraints to achieve realistic segmentation results. Meanwhile, the Quantum Approximate Optimization Algorithm (QAOA), discussed in [25], is a promising variational quantum algorithm designed to solve combinatorial optimization problems that are computationally challenging in the classical context. The review in this work covers an analysis of QAOA's performance under various conditions, its sensitivity to hardware limitations (such as noise and errors), and its potential applications in optimization tasks, particularly in image processing. Comparative studies of modifications and extensions of the algorithm help outline promising directions for its development and adaptation to practical problems.

Despite significant achievements, existing image fusion methods have several limitations. Most classical approaches focus on low-level image features (such as brightness, contrast, and textures) without considering the semantic structure of the scene. It reduces their effectiveness in real-world conditions, particularly when working with high spatial resolution satellite images or under complex weather and lighting situations. Deep learning methods demonstrate higher performance but remain resource-intensive, do not always guarantee global optimality of fusion weights, and can be prone to overfitting. In this context, a promising direction is the application of quantum machine learning, specifically the Quantum Approximate Optimization Algorithm (QAOA), which can efficiently solve discrete optimization problems. QAOA allows adaptive determination of weighting coefficients for channel fusion, which is especially important for complex spectral ranges, including infrared, where classical methods show reduced effectiveness. Therefore, this work proposes a novel hybrid approach for the fusion of optical satellite images that combines classical preprocessing, supervised learning methods, and a quantum computational block based on QAOA. This approach aims to improve the fusion results' informativeness, accuracy, and consistency.

Purpose. This research aims to develop a novel approach for the fusion of optical satellite images based on machine learning and quantum optimization, enabling the efficient integration of spatial-spectral information from different channels (RGB and IR) to produce a multispectral image with enhanced spatial detail and preserved spectral characteristics. To achieve the stated goal, the following tasks were formulated and addressed in this work:

- to perform preprocessing of optical satellite images, including geometric, radiometric, and atmospheric corrections;
- to develop a spatial-frequency image analysis method that accounts for low-frequency (pixel intensity) and high-frequency (details, contours) components;
- to integrate a quantum computational block aimed at optimizing weighting coefficients for the IR channel using QAOA (Quantum Approximate Optimization Algorithm) after dimensionality reduction;
- to develop a machine learning algorithm based on convolutional neural networks for feature representation and preparation of input data for fusion;
- to implement an image fusion procedure that combines weighting coefficients and edge information to form a normalized four-channel (RGB+IR) image with enhanced informativeness;
- to evaluate the effectiveness of the proposed approach on real datasets.

Methods. The proposed approach, which aims to enhance spatial resolution while preserving the spectral informativeness of multispectral images, is shown in Fig. 1.

At the first stage, an optical satellite image is loaded. Next, the input data preparation takes place, which involves the preprocessing of individual channels: geometric alignment (coregistration) of the red, green, blue, and near-infrared bands. Then, the following corrections are performed: geometric correction to eliminate spatial distortions; radiometric normalization to bring pixel values to a standard display scale [26]; and atmospheric correction using the Dark Object Subtraction method to remove atmospheric scattering effects. As a result, new channels $Band\ R^*(x, y)$, $Band\ G^*(x, y)$, $Band\ B^*(x, y)$, $Band\ IR^*(x, y)$ are obtained. Each channel is decomposed into low-frequency and high-frequency components to enhance the extraction of structural and spatial features of the image. The decomposition is performed using convolution with a Gaussian filter (x), with variance σ

$$Band\ R^*_{LF}(x,y) = Band\ R^*(x,y) \cdot G_\sigma(x,y);$$
$$Band\ R^*_{HF}(x,y) = Band\ R^*(x,y) - Band\ R^*_{LF}(x,y);$$
$$Band\ G^*_{LF}(x,y) = Band\ G^*(x,y) \cdot G_\sigma(x,y);$$
$$Band\ G^*_{HF}(x,y) = Band\ G^*(x,y) - Band\ G^*_{LF}(x,y);$$
$$Band\ B^*_{LF}(x,y) = Band\ B^*(x,y) \cdot G_\sigma(x,y);$$
$$Band\ B^*_{HF}(x,y) = Band\ B^*(x,y) - Band\ B^*_{LF}(x,y);$$
$$Band\ IR^*_{LF}(x,y) = Band\ IR^*(x,y) \cdot G_\sigma(x,y);$$
$$Band\ IR^*_{HF}(x,y) = Band\ IR^*(x,y) - Band\ IR^*_{LF}(x,y),$$

where $Band\ R^*(x, y)$, $Band\ G^*(x, y)$, $Band\ B^*(x, y)$, $Band\ IR^*(x, y)$ are images of the respective satellite channels; LF is a low-frequency component that preserves the brightness and overall structure of the image; HF is a high-frequency component that highlights local spatial features containing information about edges and textures.

Pixel values within each channel are normalized to improve the quality of the input channels before feeding them into the neural network. It helps reduce brightness variability and provides a unified data representation, a crucial condition for practical model training. Additionally, the Independent Component Analysis (ICA) method is employed to reduce dimensionality, allowing

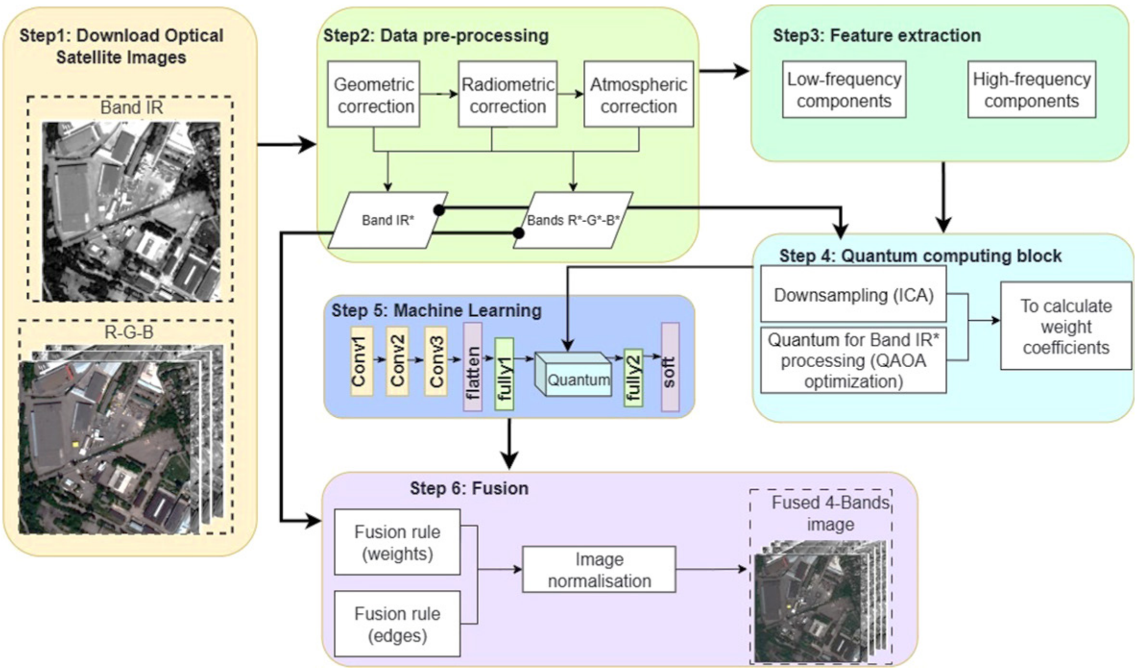

Fig. 1. Structural diagram of the proposed approach for fusion of optical satellite images

computational load reduction while preserving the most informative image features.

For further processing of the infrared channel, the proposed architecture utilizes the Quantum approximate optimization algorithm, a hybrid quantum-classical algorithm designed to solve discrete combinatorial optimization problems. In this context, QAOA is applied to determine the optimal parameters of the infrared spectrum that influence the fusion of IR data with RGB components, considering spatial-frequency features (Fig. 2). The QAOA algorithm is based on a variational approach. It consists of a sequence of p layers, each comprising a cost and mixer layer. Its goal is to construct a quantum state that minimizes the expected value of the cost function [27].

All qubits are initialized into a uniform superposition at the first step by applying Hadamard gates

$$|\psi_0\rangle = H^{\otimes n}|0\rangle,$$

where H is the Hadamard operator; n is the number of qubits corresponding to the number of principal components in the IR feature space.

The cost layer implements the problem Hamiltonian H_C, which encodes the optimization objective, such as selecting the most informative frequency components of the IR channel. The evolution operator describes the action of this layer

$$U_C(\gamma_j) = e^{-i\gamma_j H_C},$$

where γ_j is a variational parameter to be optimized.

In typical problems, such as Ising optimization, the Hamiltonian H_C takes the form

$$H_c = \sum_{k<l} J_{kl}, Z_k, Z_l + \sum_k h_k Z_k,$$

where Z_k is the Pauli-Z operator acting on the k^{th} qubit; J_{kl}, h_k are the interaction weight coefficients defining the problem structure.

The mixer layer facilitates quantum tunneling between bit configurations to explore the solution space

$$U_M(\beta_j) = e^{-i\beta_j H_M}, \quad H_M = \sum_k X_k,$$

where X_k is the Pauli-X operator; β_j is a variational parameter. This Hamiltonian corresponds to applying RX-type rotation gates to each qubit in the circuit.

The full quantum circuit for p layers is expressed as

$$|\psi(\gamma, \beta)\rangle = U_M(\beta_p) U_C(\gamma_p) \ldots U_M(\beta_1) U_C(\gamma_1) |\psi_0\rangle,$$

where $\gamma = (\gamma_1, \ldots, \gamma_p)$, $\beta = (\beta_1, \ldots, \beta_p)$ are vectors of parameters. The objective function is the expectation value of the Hamiltonian

$$C(\gamma, \beta) = \langle \psi(\gamma, \beta) | H_C | \psi(\gamma, \beta) \rangle. \quad (1)$$

The value in equation (1) is minimized by a classical SPSA optimizer, which guides the tuning of parameters γ and β.

After executing the quantum circuit, the qubits are measured in the Z-basis. The resulting bit distributions are decoded into a vector of weighting coefficients

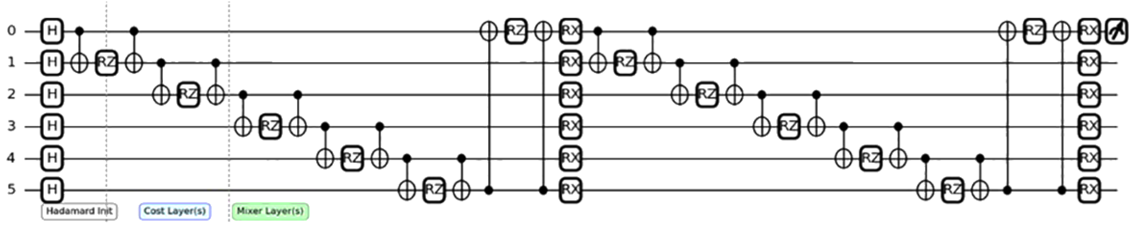

Fig. 2. Quantum circuit for implementing the QAOA algorithm

$$W_{IR} = Measure(|\psi(\gamma, \beta)) \in \mathbb{R}^n.$$

These coefficients characterize the importance of corresponding frequency features of the infrared channel and are used for weighted image fusion in multispectral analysis.

In parallel with quantum computations, a machine learning block processes the features of the RGB channels. This part of the architecture employs a convolutional neural network (CNN) that provides automatic hierarchical feature extraction from input images. At each convolutional layer (Conv1, Conv2, Conv3), filters are applied to detect local features such as edges, textures, and color gradients

$$(K \cdot F)(x,y) = \sum_u \sum_v F(x-u, y-v) \cdot K(u,v),$$

where $K(u, v)$ is the convolution kernel; $F(x, y)$ is the convolution kernel; (R, G, B or IR); (x, y) are the pixel coordinates in the output image.

After the final convolutional layer, the multidimensional feature tensors are transformed into a one-dimensional vector through a flattening operation. It is necessary for further processing by fully connected layers.

The fully connected layers (Fully1, Fully2) perform linear and nonlinear transformations of the features. At each layer, the following computation is performed

$$h^{(l)} = \sigma(W^{(l)} h^{(l-1)} + b^{(l)}),$$

where $W^{(l)}$ is the weight matrix; $b^{(l)}$ is the bias vector; σ is the ReLU activation function: $f(x) = \max(0, x)$. This transformation enables the network to learn complex dependencies between the input features and the output labels.

The hybrid quantum block is integrated into the neural network structure after one of the fully connected layers and optimizes the feature fusion weight coefficients based on quantum computation. A quantum approximate optimization algorithm is used to achieve this, which models the quantum system as a Hamiltonian H_C.

After quantum processing, the resulting feature vector is passed to the classification layer, where the Softmax function is applied to transform the vector of normalized weight values z_i into a probability distribution

$$y_i = \frac{e^{z_i}}{\sum_j z_j},$$

where z_i is the vector of normalized weight values for class *i*. In the context of information fusion from multiple channels, the Softmax function can be used to determine relative weights or to select the most relevant feature integration rule learned from the multidimensional feature space.

The final stage is the fusion step, in which data from the infrared and visible spectral bands are integrated to produce an enhanced four-channel image with higher quality, improved contrast, and better preservation of scene details. The fusion of information from different channels is performed based on the weight coefficients obtained in the previous processing steps. Specifically, the weights for the IR channel W_{IR} are generated by the quantum computational block. In contrast, the weights for the visible spectrum channels $W_{RGB} = \{\beta_R, \beta_G, \beta_B\}$ are determined based on the output of a convolutional neural network. These coefficients are then used to define the overall fusion rule

$$I_{Fused}(x,y) = \alpha \cdot Band^*_{IR}(x,y) + \sum_{k \in \{R,G,B\}} \beta_k \cdot Band^*_k(x,y),$$

where α is the weight coefficient for the IR channel; β_k is the weight coefficient for the corresponding channel $k \in \{R, G, B\}$; $Band^*_k(x,y)$ is he processed (frequency-filtered) component of the respective channel.

Additional integration of high-frequency information is performed to preserve localized spatial details and enhance image sharpness, especially in areas with pronounced object boundaries. The high-frequency component $F_{HF}(x, y)$, obtained as a result of spatial-frequency decomposition, is added to the previously fused low-frequency level image

$$I_{Fused}(x,y) = I_{LowFreq}(x,y) + F_{HF}(x,y),$$

where $I_{LowFreq}(x, y)$ represents the weighted fusion of the low-frequency components of all channels. This approach ensures a balance between spectral consistency and preservation of spatial details.

At the final step, linear normalization of the pixel values of the fused image is performed to bring them into the standard range [0,255] or visualization in the RGB format. This method provides data consistency for subsequent thematic analysis and contributes to the improvement of the visual quality of the results

$$I_{Normalized}(x,y) = \frac{I_{Fused(x,y)} - I_{\min}}{I_{\max} - I_{\min}} \cdot 255,$$

where $I_{fused}(x, y)$ s the pixel value in the fused image; $I_{\min}$, $I_{\max}$ are the minimum and maximum pixel values in the fused image, respectively.

Findings. The study used WorldView-3 images containing multichannel data, including visible spectrum and infrared bands. The dataset covers heterogeneous scenes, such as urban areas, infrastructure, forests, water bodies, and agricultural lands. This diversity ensures the model's better generalization capability and applicability to various landscape types. To form a sample set oriented for neural network training, the scenes were divided into non-overlapping patches of size 400 × 400 pixels, which enabled efficient processing of large images and provided a sufficient number of training examples. The dataset was split into three subsets: training (80 %) for model parameter optimization; validation (10 %) for monitoring generalization ability and preventing overfitting; and testing (10 %) for final independent performance evaluation. For training the deep convolutional neural network with Residual-in-Residual Dense Blocks, the Adam optimization algorithm was used, providing stable and fast convergence. The initial learning rate was set to 1e-4, with stepwise decay every 50 epochs, with a factor $\gamma = 0.5$. The model was trained for 200 epochs with a batch size of 16. An early stopping strategy based on the validation loss metric was applied to reduce the risk of overfitting.

The loss function was a combination of Mean Squared Error (MSE) and perceptual loss, which enabled a better balance between numerical errors and the visual quality of the reconstructed images. Fig. 3 presents the results of the WorldView-3 satellite image fusion using various methods. Visual comparison allows assessment of the quality of spatial detail restoration, color consistency, and presence of artifacts. The image in Fig. 3, *a* demonstrates relatively high spatial detail, especially in the contours of buildings and the road network. However, significant spectral distortions are observed, particularly with green areas appearing unnaturally saturated, and the overall image tone tends toward bluish-green. The image in Fig. 3, *b* visually resembles the Brovey method. It generally preserves structure but shows atypical grayish areas, indicating an impact on spectral information. The image in Fig. 3, *c* has high sharpness and clarity of details. Object edges look very pronounced. However, this sharpness is achieved at the expense of significant color distortion and artifacts around sharp transitions. The colors appear washed out, indicating a loss of spectral information. The image in Fig. 3, *d* demonstrates a good balance between spatial sharpness and preservation of natural colors. Details are precise, and colors look more natural and saturated compared to the Brovey, Gram-Schmidt, and IHS methods. The image in Fig. 3, *e* is visually similar to the Brovey and Gram-Schmidt methods in detail, but it shows moderate color changes (the spectral component contains artifacts or distortions). The image in Fig. 3, *f* combines high spatial resolution with preservation of spectral information. Colors look natural, details are sharp, and the overall visual impression is balanced. The image in Fig. 3, *g* results from applying a simple high-frequency filter. It has high sharpness and emphasized details, but, like HCS, tends to color distortions and looks less natural due to excessive enhancement of high-frequency components. The image in Fig. 3, *h* demonstrates preservation of color information. Detailing is also improved, and colors look natural, although contrast is somewhat lower than in other methods. The image in Fig. 3, *i*, obtained using the proposed approach, visually stands out with an optimal combination of spatial detail and preservation of the spectral component. Objects in the satellite image (buildings, roads, shadows) have sharp edges without visible artifacts. Colors maintain a natural appearance, closely matching the original visible spectrum image. At the same time, detail is preserved and enhanced thanks to the infrared channel, which provides increased sensitivity to differences in material reflectance, particularly improving classification of urban elements such as roof types and road surfaces. Such an image is visually the most informative, confirming the effectiveness of the proposed algorithm in preserving both spatial and spectral characteristics of the scene.

For a quantitative assessment of the effectiveness of multispectral image fusion, a comparative analysis was conducted with existing methods: Brovey, Gram-Schmidt, IHS, HCS, HPFC, ATWT, as well as a deep convolutional neural network and the proposed approach. For objective evaluation, both the True Color Image and individual spectral bands were used along with the following metrics (Table 1): Mean Squared Error, Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), entropy, Average Gradient, Spatial Frequency, and Sobel Sharpness. The quantitative analysis of the obtained results (Table 1) demonstrates that incorporating the infrared channel into the multispectral image (R-G-B-IR) significantly enhances the spectral expressiveness of the data. The proposed approach shows the lowest Mean Squared Error value (MSE = 191.8), indicating minimal spectral distortion relative to the original visual image. For comparison, traditional methods such as Brovey (MSE = 6,081.97) and Gram-Schmidt (MSE = 5,901.88) have substantially higher error values, indicating considerable loss of spectral components. However, it should be noted that some modern transform-based methods, like HPFC (MSE = 194.70) and ATWT (MSE = 195.46), show an error level similar to the proposed method. PSNR also confirms the advantage of the proposed method over classical approaches. However, the absolute value is low (PSNR = 9.94), it is compensated by balanced structural similarity and absence of color distortions, which is especially critical when integrating infrared information.

The SSIM index for the proposed method (SSIM = = 0.43) is higher than most traditional methods (Brovey, IHS, Gram-Schmidt) but lower than the transformative approaches ATWT (SSIM = 0.8348) and HPFC (SSIM = 0.8539). It is explained by a less aggressive transfer of the luminance component to preserve balanced spectral accuracy. The average gradient and Sobel Sharpness for the proposed method (Sobel Sharpness = = 142.19) indicate sufficient detail presence, although they are lower than those of HCS (Sobel Sharpness = = 220.39) or HPFC (Sobel Sharpness = 149.19). However, this spatially moderate detail is balanced and does not lead to artifacts or oversaturation with high-frequency information. The entropy of the fused image in the proposed method (Entropy = 7.54) exceeds the values typical for most methods, indicating high spectral informativeness. An entropy increase of 2.88 % relative to the input image suggests a moderate spectral range expansion without significant structural consistency loss.

To evaluate the effectiveness of channel fusion, a detailed analysis was conducted of intensity metrics, textural features, and spatial homogeneity of the image in each spectral channel (Table 2).

The selected metrics enable a multidimensional evaluation of image fusion quality, covering brightness and contrast preservation, textural detail retention, and spatial homogeneity across spectral channels. Intensity-based indicators reveal how well the fusion process maintains the original radiometric properties, while texture-related measures quantify the integrity of fine structural details essential for accurate interpretation. Spatial homogeneity metrics assess the uniformity and consistency of fused regions, allowing detection of potential distortions or artifacts. Such a comprehensive analysis facilitates the identification of fusion methods that achieve an optimal balance between enhancing local features and maintaining overall spectral stability, providing a clear basis for the comparative discussion presented below.

Methods that provide high contrast, particularly HCS in the R channel and PCA in the R channel, demonstrate increased sensitivity to local brightness variations. The high coefficient of variation in the HCS method indicates instability in brightness distribution, leading to a loss of spectral consistency.

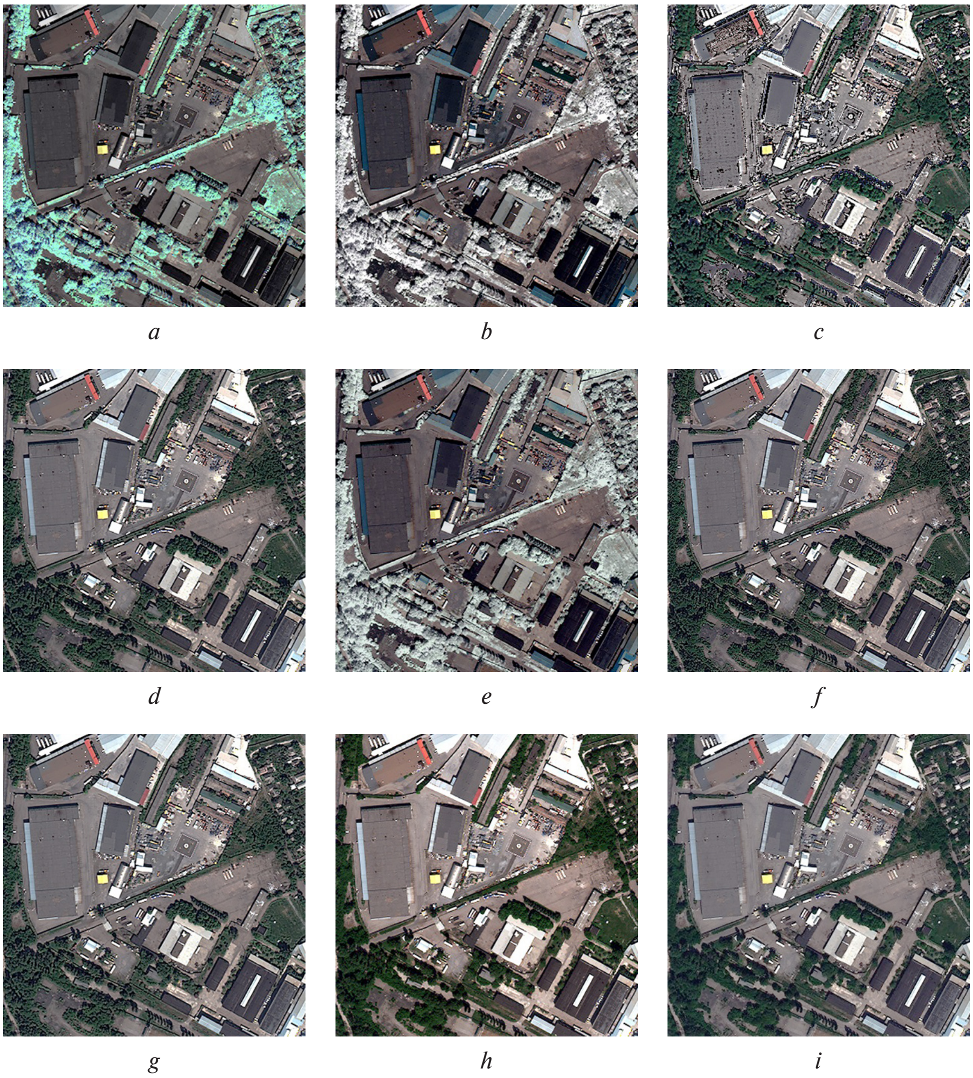

Fig. 3. Fusion results for the WorldView-3 satellite image:
a — Brovey; b — Gram-Schmidt; c — HCS; d — HPFC; e — IHS; f — ATWT; g — HPF; h — PCA; i — proposed approach

Table 1

Quantitative metrics for assessing the fusion quality of the WorldView-3 optical image

Method	MSE	PSNR	SSIM	Avg Grad	SF	Rel Avg Grad Inc, %	Entropy	Entropy Inc, %	Avg Sobel Sharp	Sobel Sharp Inc, %
Brovey	6,081.97	10.29	0.35	159.08	48.16	39.27	7.49	0.91	159.08	39.26
Gram-Schmidt	5,901.88	10.42	0.34	155.58	46.54	36.19	7.52	1.22	155.57	36.19
HCS	2,050.23	15.01	0.49	220.39	83.54	92.93	6.98	–6.02	220.39	92.93
HPFC	194.70	25.26	0.85	149.19	47.86	30.59	7.51	1.15	149.18	30.59
IHS	5,899.76	10.42	0.34	155.59	46.55	36.21	7.52	1.19	155.59	36.21
ATWT	195.46	25.22	0.83	146.39	51.49	28.15	7.46	0.38	146.38	28.15
CNN	192.51	10.12	1.20	114.23	38.95	30.00	7.54	1.20	114.23	30.00
PCA	1,236.72	17.21	0.87	149.28	47.98	114.23	7.74	2.22	149.28	30.67
Proposed approach	191.80	9.94	0.43	142.19	47.86	24.47	7.54	2.88	142.18	53.47

The proposed approach shows balanced contrast values in the R, G, and B channels with reduced variation coefficients, indicating the transmitted channels' spectral stability without oversaturation.

A key indicator is the Quantum_Band channel, which demonstrates the highest contrast among all channels of the proposed approach, while having the lowest coefficient of variation, indicating stable spatial generalization of high-frequency components while maintaining spectral consistency across channels. The entropy metric in most classical methods ranges between 7.10 and 7.50, with somewhat lower values in

Table 2

Analysis of metrics for evaluating image fusion quality

Method	Bands	Standard deviation	Variation coefficient	Entropy
Brovey	R	45.54	0.63	7.29
Brovey	G	67.56	0.79	7.44
Brovey	B	57.38	0.69	7.50
Gram-Schmidt	R	57.11	0.70	7.48
Gram-Schmidt	G	58.19	0.73	7.49
Gram-Schmidt	B	52.90	0.65	7.46
HCS	R	74.84	0.96	6.95
HCS	G	70.69	0.89	6.91
HCS	B	70.54	0.89	6.90
HPFC	R	61.60	0.71	7.48
HPFC	G	57.87	0.67	7.40
HPFC	B	57.18	0.66	7.43
IHS	R	54.63	0.67	7.46
IHS	G	59.20	0.74	7.50
IHS	B	54.53	0.67	7.47
ATWT	R	61.58	0.76	7.40
ATWT	G	58.07	0.73	7.35
ATWT	B	57.51	0.71	7.40
CNN	R	62.46	0.78	7.44
CNN	G	59.04	0.76	7.37
CNN	B	58.61	0.73	7.44
R-G-B-IR	R	59.99	0.75	7.43
R-G-B-IR	G	55.58	0.69	7.37
R-G-B-IR	B	55.19	0.68	7.18
R-G-B-IR	IR	60.26	0.58	7.63
PCA	R	73.64	0.65	7.17
PCA	G	65.66	0.60	7.11
PCA	B	69.06	0.68	6.93
Proposed approach	R	59.99	0.74	7.44
Proposed approach	G	55.58	0.69	7.37
Proposed approach	B	55.19	0.68	7.19
Proposed approach	Quantum_Band	65.79	0.52	7.54

HCS, indicating some loss of spectral content due to excessive amplification of high-frequency components (over-enhancement of spatial details). The highest entropy values are observed on the IR channel and Quantum_Band, indicating successful preservation and integration of additional spectral information.

Results for spatial detail assessment obtained using the Sobel operator are shown in Fig. 4. According to the data presented in Fig. 4, the HCS method provides the highest level of spatial structure detailing (Sobel operator = up to 39.59), but at the expense of spectral consistency loss. The proposed approach achieves Sobel Sharp values ranging from R-channel (Sobel operator = 21.67) to B-channel (Sobel operator = 19.19), which is moderate and balanced. Meanwhile, Quantum_Band demonstrates a high level of spatial structure detailing (Sobel operator = 31.27), confirming the effectiveness of quantum processing in preserving edge structures without sacrificing spectral balance.

Conclusions. This work proposes a novel approach to the fusion of optical satellite images based on quantum machine learning, combining methods of spatial-frequency analysis, convolutional neural networks, and QAOA for integrating R-G-B and infrared channel data. Experiments conducted on real WorldView-3 satellite images confirmed that the developed approach effectively preserves the spectral information of the visible range while improving spatial resolution through high-frequency components of the IR channel.

According to quantitative evaluation results, the proposed approach demonstrates higher accuracy in integrating spectral and spatial information than classical fusion methods. In particular, among the considered methods (Brovey, Gram-Schmidt, IHS, HCS, HPFC, ATWT, PCA, CNN), the proposed method achieved the minimum MSE value of 191.8, indicating the least deviation from the original visible image; a high structural similarity index SSIM = 0.43; high entropy values (Entropy = 7.54), indicating preservation of informational richness; balanced spatial detail metrics: Sobel Sharpness (ranging from 19.19 to 21.67 across R, G, B channels); as well as a significant increase in sharpness in the quantum channel (Sobel Sharpness = 31.27), con-

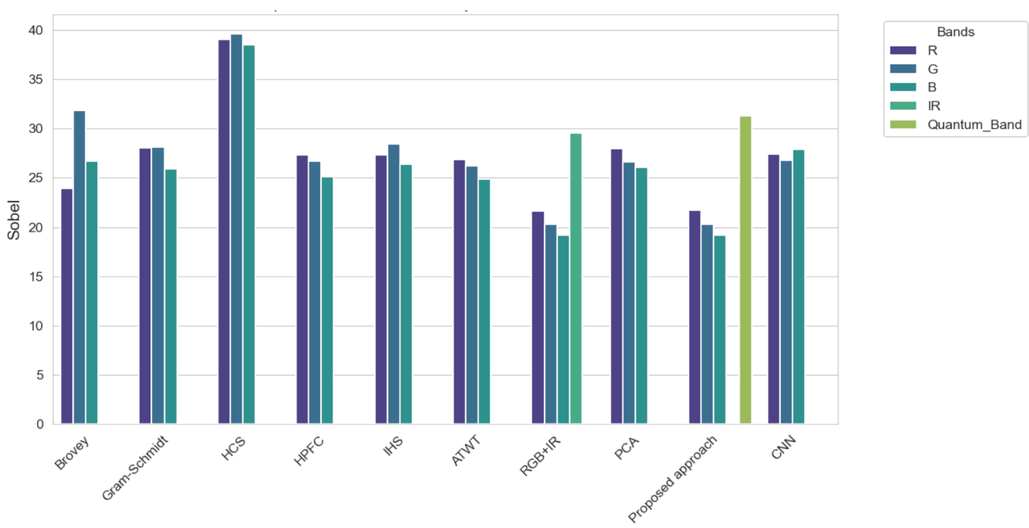

Fig. 4. Results of spatial detailing using the Sobel operator

firming the effectiveness of quantum optimization in preserving spatial structures.

Visual analysis confirms that images obtained by the proposed approach exhibit clear spatial structure without artifacts, and color reproduction corresponds to the spectral characteristics of the original RGB data. Notably, the algorithm can distinguish fine textural features of objects on the terrain, including buildings, roads, green vegetation, and soil cover.

Thus, the proposed quantum machine learning-based approach is an effective tool for multichannel satellite image fusion, ensuring a high level of spatial-spectral information integration, and can be applied in remote sensing tasks, landscape monitoring, and mapping.

Acknowledgements. *The article was prepared within the framework of the OptiQ project. This project has received funding from the European Union's Horizon Europe programme under grant agreement No. 101080374 – OptiQ. Additionally, the project is co-financed by the Polish Ministry of Science and Higher Education under the International Cofinanced Projects programme. Disclaimer: Funded by the European Union. The views and opinions expressed in this article are those of the authors only and do not necessarily reflect the views of the European Union or the European Research Executive Agency (REA – granting authority). Neither the European Union nor the granting authority can be held responsible for them.*

References.

1. Meher, B., Agrawal, S., Panda, R., & Abraham, A. (2019). A survey on region based image fusion methods. *Information Fusion*, *48*, 119-132. https://doi.org/10.1016/j.inffus.2018.07.010

2. Hnatushenko, V., Hnatushenko, Vik., Kavats, A., & Shevchenko, V. (2015). Pansharpening technology of high resolution multispectral and panchromatic satellite images. *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*, (4), 91-98.

3. Liu, Y., Wang, L., Cheng, J., Li, C., & Chen, X. (2020). Multi-focus image fusion: A Survey of the state of the art. *Information Fusion*, *64*, 71-91. https://doi.org/10.1016/j.inffus.2020.06.013

4. Kashtan, V., & Hnatushenko, V. (2023). Deep Learning Technology for Automatic Burned Area Extraction Using Satellite High Spatial Resolution Images. *Lecture Notes in Computational Intelligence and Decision Making. ISDMCI 2022. Advances in Intelligent Systems and Computing*, *1246*, 664-685. Springer, Cham. https://doi.org/10.1007/978-3-031-16203-9_37

5. Panda, M. K., Parida, P., & Rout, D. K. (2024). A weight induced contrast map for infrared and visible image fusion. *Computers and Electrical Engineering*, *117*, 109256. https://doi.org/10.1016/j.compeleceng.2024.109256

6. Qiu, X., Li, M., Zhang, L., & Yuan, X. (2019). Guided filter-based multi-focus image fusion through focus region detection. *Signal Processing: Image Communication*, *72*, 35-46. https://doi.org/10.1016/j.image.2018.12.004

7. Zhang, H., Xu, H., Tian, X., Jiang, J., & Ma, J. (2021). Image fusion meets deep learning: A survey and perspective. *Information Fusion*, *76*, 323-336. https://doi.org/10.1016/j.inffus.2021.06.008

8. Farid, M. S., Mahmood, A., & Al-Maadeed, S. A. (2019). Multi-focus image fusion using Content Adaptive Blurring. *Information Fusion*, *45*, 96-112. https://doi.org/10.1016/j.inffus.2018.01.009

9. Liu, Y., Liu, S., & Wang, Z. (2015). Multi-focus image fusion with dense SIFT. *Information Fusion*, *23*, 139-155. https://doi.org/10.1016/j.inffus.2014.05.004

10. Abhyankar, M., Khaparde, A., & Deshmukh, V. (2016). Spatial domain decision based image fusion using superimposition. *2016 IEEE/ACIS 15th International Conference on Computer and Information Science*. IEEE. https://doi.org/10.1109/icis.2016.7550766

11. Kashtan, V., & Hnatushenko, V. (2019). Computer technology of high resolution satellite image processing based on packet wavelet transform. *International Workshop on Conflict Management in Global Information Networks CMiGIN 2019*, Lviv, Ukraine, (pp. 370-380). Retrieved from https://ceur-ws.org/Vol-2588/paper31.pdf

12. Liu, Y., Liu, S., & Wang, Z. (2015). A general framework for image fusion based on multi-scale transform and sparse representation. *Information Fusion*, *24*, 147-164. https://doi.org/10.1016/j.inffus.2014.09.004

13. Kashtan, V., & Hnatushenko, V. (2020). A Wavelet and HSV Pansharpening Technology of High Resolution Satellite Images. *Intelligent Information Technologies & Systems of Information Security Intel, 2020*, 67-76. Retrieved from https://ceur-ws.org/Vol-2623/paper7.pdf

14. Merianos, I., & Mitianoudis, N. (2019). Multiple-Exposure Image Fusion for HDR Image Synthesis Using Learned Analysis Transformations. *Journal of Imaging*, *5*(3), 32. https://doi.org/10.3390/jimaging5030032

15. Martorell, O., Sbert, C., & Buades, A. (2019). Ghosting-free DCT based multi-exposure image fusion. *Signal Processing: Image Communication*, *78*, 409-425. https://doi.org/10.1016/j.image.2019.07.020

16. Lin, Chenyoukang, Liu, Tao, Wang, Zixi & Wang, Bo (2025). CTIUFuse: A CNN-Transformer-based Iterative Feature Universal Fusion Algorithm for Multimodal Images. *Knowledge-Based Systems*, *329*, 114313. https://doi.org/10.1016/j.knosys.2025.114313

17. Liu, Z., Zheng, Y., & Han, X.-H. (2021). Deep Unsupervised Fusion Learning for Hyperspectral Image Super Resolution. *Sensors*, *21*(7), 2348. https://doi.org/10.3390/s21072348

18. Lai, R., Li, Y., Guan, J., & Xiong, A. (2019). Multi-Scale Visual Attention Deep Convolutional Neural Network for Multi-Focus Image Fusion. *IEEE Access*, *7*, 114385-114399. https://doi.org/10.1109/access.2019.2935006

19. Shao, Z., & Cai, J. (2018). Remote Sensing Image Fusion With Deep Convolutional Neural Network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(5), 1656-1669. https://doi.org/10.1109/jstars.2018.2805923

20. Wang, J., Shao, Z., Huang, X., Lu, T., & Zhang, R. (2021). A Dual-Path Fusion Network for Pan-Sharpening. *IEEE Transactions on Geoscience and Remote Sensing*, 1-14. https://doi.org/10.1109/tgrs.2021.3090585

21. Benzenati, T., Kessentini, Y., & Kallel, A. (2022). Pansharpening approach via two-stream detail injection based on relativistic generative adversarial networks. *Expert Systems with Applications*, *188*, 115996. https://doi.org/10.1016/j.eswa.2021.115996

22. Wang, Z., Li, X., Duan, H., & Zhang, X. (2022). A Self-supervised Residual Feature Learning Model for Multi-focus Image Fusion. *IEEE Transactions on Image Processing*, *1*. https://doi.org/10.1109/tip.2022.3184250

23. Liu, S., & Yang, L. (2022). BPDGAN: A GAN-Based Unsupervised Back Project Dense Network for Multi-Modal Medical Image Fusion. *Entropy*, *24*(12), 1823. https://doi.org/10.3390/e24121823

24. Nam, H. Le, Sonka, M., & Toor, F. (2023). A Quantum Optimization Method for Geometric Constrained Image Segmentation. *Quantum Physics*. https://doi.org/10.48550/arXiv.2310.20154

25. Blekos, K., Brand, D., Ceschini, A., Chou, C.-H., Li, R.-H., Pandya, K., & Summer, A. (2024). A review on Quantum Approximate Optimization Algorithm and its variants. *Physics Reports*, *1068*, 1-66. https://doi.org/10.1016/j.physrep.2024.03.002

26. Kashtan, V. J., Hnatushenko, V. V., & Shedlovska, Y. I. (2017). Processing technology of multispectral remote sensing images. *2017 IEEE International Young Scientists' Forum on Applied Physics and Engineering*, Lviv, (pp. 17-20), October 2017. https://doi.org/10.1109/ysf.2017.8126647

27. Zou, P. (2025). Multiscale quantum approximate optimization algorithm. *Physical Review A*, *111*(1). https://doi.org/10.1103/physreva.111.012427

Квантове машинне навчання для злиття багатоканальних оптичних супутникових знімків

В. Ю. Каштан[1], *В. В. Гнатушенко*[*1], *К. Верещинський*[2], *К. Циран*[2]

1 – Національний технічний університет «Дніпровська політехніка», м. Дніпро, Україна

2 – Сілезький технологічний університет, м. Глівіце, Республіка Польща

[*] Автор-кореспондент e-mail: vvgnat@ukr.net

Мета. Розробка нового підходу до злиття оптичних супутникових зображень на основі машинного навчання й квантової оптимізації для інтеграції просторово-спектральної інформації RGB- і IR-каналів.

Методика. Запропонований підхід передбачає послідовну обробку вхідних даних, що включає геометричну, радіометричну й атмосферну корекції. Кожен канал розкладається на низькочастотну та високочастотну компоненти за допомогою Гаусового фільтра. Для зменшення розмірності вхідних даних застосовується метод незалежних компонент (ICA). Для аналізу інфрачервоного каналу застосовано алгоритм апроксимації квантового оптимізатора. Глибока згорткова нейронна мережа із залишковими щільними блоками застосовується для вилучення просторових структурних ознак із RGB-каналів. Після інтеграції ознак через повнозв'язні шари квантовий блок оптимізує вагові коефіцієнти для остаточного злиття каналів.

Результати. За результатами кількісної оцінки запропонований підхід продемонстрував перевагу у точності інтеграції спектральної та просторової інформації порівняно із класичними методами злиття, зокрема Brovey, Gram-Schmidt, IHS, HCS, HPFC, ATWT, PCA та CNN. Підхід досягає найнижчого значення середньоквадратичної помилки (MSE = 191,8), високих значень структурної подібності (SSIM = 0,43), ентропії (Entropy = 7,54), фільтра Собеля (діапазон Sobel Sharp = 19,19–21,67 по каналах R, G, B). Візуальний аналіз також підтверджує якісні переваги запропонованого підходу: зображення характеризуються чіткою структурою без артефактів, збалансованою кольоровою передачею, що відповідає спектральним характеристикам оригінальних RGB-даних.

Наукова новизна. Запропоновано підхід використання інформації IR-каналу, що інтегрує квантово-класичний алгоритм у структуру глибокої згорткової нейронної мережі, для синергетичного оброблення багатоканальних оптичних зображень із використанням багаторівневої частотної декомпозиції та механізму вагового злиття ознак.

Практична значимість. Запропонований підхід може бути впроваджений у системах дистанційного зондування Землі задля покращення якості оброблення супутникових знімків, зокрема у завданнях картографування, моніторингу земельних ресурсів, сільськогосподарського контролю й екологічного аналізу. Застосування квантових алгоритмів відкриває нові можливості для підвищення ефективності й точності оброблення багатовимірних геоінформаційних даних, що містять інформацію IR-каналу.

Ключові слова: *квантове машинне навчання, злиття даних, нейронна мережа, супутникове зображення*

The manuscript was submitted 03.05.25.