UDC 004.93

HYBRID SATELIT IMAGE RECOGNITION SYSTEM COMBINING NEURAL NETWORK FEATURE EXTRACTION AND AN INFORMATION-EXTREMAL CLASSIFIER

Dovbysh A. S. – Dr. Sc., Professor, Professor of the Department of Computer Science, Sumy State University, Sumy, Ukraine. ROR: https://ror.org/01w60n236. ORCID:https://orcid.org/0000-0003-1829-3318.

Piatachenko V. Y. – PhD, Assistant of the Department of Computer Science, Sumy State University, Sumy, Ukraine. ROR: https://ror.org/01w60n236. ORCID: https://orcid.org/0000-0002-7464-3119.

Serhieiev V. M. – Post-graduate student of the Department of Computer Science, Sumy State University, Sumy, Ukraine. ROR: https://ror.org/01w60n236. ORCID:https://orcid.org/0009-0009-3838-0153.

Hrytsenko O. M. – Post-graduate student of the Department of Computer Science, Sumy State University, Sumy, Ukraine. ROR: https://ror.org/01w60n236. ORCID: https://orcid.org/0009-0004-1382-708X.

ABSTRACT

Context. The study solves the relevant task of developing an interpretable and adaptive recognition system for semantic segmentation of satellite imagery by integrating neural network feature extractors with an information-extreme classifier.

Objective. To improve the accuracy of satellite land cover classification by developing a hybrid machine learning model that combines a deep convolutional neural network for extracting informative features with an information-extreme classifier, enabling the construction of highly reliable decision rules even in the presence of overlapping recognition classes in the feature space.

Method. A hybrid model is proposed that combines efficient spatial feature extraction using a convolutional neural network (CNN) with an information-extreme intelligent (IEI) technology for data analysis, based on maximizing the information capacity of the recognition system during machine learning. For feature aggregation, GlobalAveragePooling is applied instead of the classical Flatten operation. Additionally, regularization techniques such as weight decay and cyclical learning rate scheduling are implemented. The optimization of IEI model parameters is carried out using a modified Kullback information criterion, interpreted as a measure of recognition class diversity.

Results. The developed model achieves high classification accuracy (95%) on the test set and demonstrates stable performance, with improved efficiency of the neural feature extractor due to a reduced number of training epochs enabled by regularization techniques. As a result of the information-extreme machine learning process, the optimal geometric parameters of the hyperspherical recognition class containers were determined, allowing the construction of highly reliable decision rules even under conditions of recognition class overlap in the feature space.

Conclusions. The proposed hybrid model enables the construction of highly reliable decision rules through information-extreme machine learning, even in cases of a priori fuzzy partitioning of recognition classes in the feature space, based on the input training matrix formed during feature extraction.

KEYWORDS: information-extreme machine learning, convolutional neural network, information criterion, optimization, hybrid model, recognition feature extraction, land cover images.

ABBREVIATIONS

CNN is a Convolutional Neural Network;

IEI is an Information-Extreme Intellectual (Machine) Learning;

ReLU is a Rectified Linear Unit;

Swish is a Sigmoid-Weighted Linear Unit;

AdamW is an Adaptive Moment Estimation with Weight Decay;

SVM is a Support Vector Machine;

k-NN is a k-Nearest Neighbors;

GAP is a Global Average Pooling.

NOMENCLATURE

M is a number of recognition classes;

N is a number of recognition features;

n is a number of feature vectors of recognition classes in the training matrix;

x_m is an averaged binary feature vector of the recognition class X_m^0;

Z is a recognition feature space;

X is a working binary training matrix;

f_1 is an operator for forming matrix Y;

f_2 is an operator for forming matrix X;

G_E is an admissible domain of the information criterion function used for machine learning parameter optimization;

$K_{1,m}(d)$ is a number of events where realizations of class X_m^0 are mistakenly not assigned to their own class;

$K_{2,m}(d)$ is a number of events where alien realizations are mistakenly assigned to class X_m^0;

10^{-r} is a sufficiently small number introduced to avoid division by zero;

$\overline{D_1}$ is an average first-order confidence across the recognition class alphabet;

$\overline{\beta}$ is an average second-kind error across the recognition class alphabet;

$\overline{x_i}^B$ is an averaged structured binary feature vector defining the geometric center of the container for class X_i;

DOI 10.15588/1607-3274-2026-1-6

R_i is a radius of the hyperspherical container for class X_i;

δ_i is a parameter equal to half the tolerance field width for recognition features;

$\hat{\delta}$ is a normalized tolerance field over the recognition features;

D_{ij} is an inter-center Hamming code-based distance between class X_i and its nearest neighbor X_j in the binary feature space;

w is an image width;

h is an image height;

c is an image channel;

W is a convolution kernel;

k is a filter size;

y_i is a ground truth label;

$\hat{y}_i$ is a model prediction;

λ is a regularization coefficient.

INTRODUCTION

An important research direction in the field of intelligent image analysis is the investigation of effective approaches to feature extraction, which enhances the informativeness of the input mathematical representation and enables high-confidence classification decisions in visual recognition tasks. These tasks have broad applications in areas such as remote sensing, autonomous navigation, environmental monitoring, agroanalytics, military intelligence, and others. In particular, satellite imagery plays a key role in identifying land cover types, detecting changes in terrain structure, and classifying infrastructure-related objects.

Despite the widespread use of convolutional neural networks (CNNs) in image processing tasks, they have a number of significant limitations. For instance, their decisions are often difficult to interpret, which complicates their use in sensitive domains where transparency is required. In addition, neural networks demand large amounts of training data and are inflexible during retraining, especially when the number of machine learning objects increases under conditions of substantial overlap in the recognition feature space.

To overcome these limitations, hybrid models that combine the representational power of deep neural networks with the strengths of classical decision-making frameworks offer a promising solution. These strengths include interpretability, robustness, and statistical consistency. One such framework is the information-extreme intelligent technology (IEIT), which is based on maximizing the information capacity of the recognition system during training. The integration of CNNs for feature extraction with IEIT as a transparent classification mechanism provides a foundation for the development of flexible, adaptive, and controllable next-generation artificial intelligence systems.

The object of study is the process of land cover image classification based on satellite data.

Unlike traditional neural networks, the information-extreme intelligent (IEI) technology implements a geometric representation of classes in the form of containers within the feature space and enables the construction of decision rules that are invariant to the dimensionality and distribution of features. When combined with a convolutional neural network (CNN) that performs preliminary processing of input images and generates a descriptive feature representation, this approach offers flexible control over the trade-off between model accuracy, generalization, and interpretability.

The subject of study is a hybrid machine learning model that integrates a convolutional neural network with an information-extreme approach to recognition.

The work investigates the architecture of the hybrid model, develops algorithms for constructing recognition class containers, and conducts an experimental evaluation of model performance using the EuroSAT image dataset. Special attention is given to analyzing the role of the CNN as a feature extraction mechanism and examining the influence of geometric parameters of the categorical model on decision-making within the IEI framework.

The purpose of the work is to increase the accuracy and interpretability of satellite image classification by developing a hybrid machine learning model.

1 PROBLEM STATEMENT

Let us consider a formalized formulation of the information-extreme machine learning problem for image classification using a hybrid model. Let X be the alphabet of recognition classes corresponding to image categories (in our case, types of terrain). Each image X is input into a convolutional neural network (CNN), which performs feature extraction and transforms the image into a feature vector $z \in \Re^n$ in the feature space. The collection of such vectors forms a three-dimensional brightness matrix Z of dimensions $w \times h \times c$, which serves as the input to the information-extreme classifier. For each recognition class $X_i \in X$ the machine learning parameters are defined in the radial basis of the Hamming feature space:

$$\Omega_i = (\overline{x}_i^B, R_i, \delta_i). \qquad (1)$$

At the same time, the machine learning parameters are subject to the following constraints:

1) the value of the radius R_i must satisfy the inequality

$$R_i \le D_{ij},$$

2) the parameter δ_i is constrained to lie within the domain

$$\delta_i \le \hat{\delta}.$$

DOI 10.15588/1607-3274-2026-1-6

During the machine learning process, the following steps must be performed:

1) In accordance with the concept of the information-extreme intelligent (IEI) technology, the feature matrix Z must be transformed into a binary working matrix X^B by quantizing the features according to control tolerance levels. This transformation allows the model to adapt to the condition of maximizing the complete probability of correct classification decisions.

2) Based on the optimized parameters of the containers, decision rules must be constructed in the form of categorical mappings. These rules enable the determination of the class membership of an input feature vector z using its binary representation in the Hamming feature space (1).

3) When the recognition system operates in examination mode, the functional performance of the hybrid model must be validated by evaluating the classification accuracy and the statistical consistency of the decision rules generated during the training phase.

2 REVIEW OF THE LITERATURE

In recent years, the increasing complexity of recognition tasks has fueled growing interest in hybrid approaches within neural network models, particularly those that combine convolutional neural networks (CNNs) with classical machine learning methods. Such integration contributes to improved classification accuracy, especially under conditions of limited computational resources.

In computer vision tasks, studies on the effectiveness of hybrid models [1–3] have shown that employing CNNs for feature extraction followed by classification using linear algorithms can significantly reduce computational complexity without a substantial loss in accuracy. This highlights the potential of hybrid methods for deployment in resource-constrained environments.

The work presented in [2] proposes hybrid CNN-SVM and CNN-KNN architectures for image classification. These approaches treat the CNN as a feature extractor, combined with either a support vector machine (SVM) or k-nearest neighbors (k-NN) classifier, and demonstrate higher accuracy compared to standard CNNs. This approach has proven particularly effective when working with small datasets, as linear classification algorithms can generalize extracted features more efficiently in such scenarios.

The study presented in [3] demonstrated that, following deep feature extraction via a convolutional neural network (CNN), the use of k-nearest neighbors (k-NN) can enhance recognition accuracy by reducing the number of misclassifications.

In parallel with hybrid approaches, recent literature has devoted considerable attention to increasing model robustness against adversarial attacks, concept drift, and data imperfections. In [4, 5], the authors propose resilient classifier architectures that implement specialized training mechanisms to adapt to variations in the structure of input data. Specifically, [4] describes a method for constructing a resilient classifier capable of handling concept shift and fault injection, while [5] introduces a training architecture and algorithm that explicitly accounts for these factors.

Another research direction that has garnered interest in the context of interpretable models is information-extreme machine learning. This approach is based on maximizing the information capacity of the recognition system during training. Studies [6, 7] successfully apply the information-extreme intelligent (IEI) technology to photographic and video analytics tasks. Both examples demonstrate the advantages of geometric class representation and the use of an information criterion as a training objective function.

The direct foundation for this study lies in [8, 9], where information-extreme algorithms are proposed for onboard recognition systems targeting ground objects. In particular, [8] introduces a mechanism for constructing recognition class containers in the feature space, followed by optimization of their geometric parameters. Study [9] presents an extension of this model that enables selection of a base recognition class, improving the functional performance of multiclass machine learning.

The hybrid model proposed in this paper may be applied to solve the problem of deep neural feature extraction for recognition tasks.

3 MATERIALS AND METHODS

The concept of hybrid models is well established and widely supported in the literature [10, 11], as it leverages the ability of neural networks to effectively extract informative features from data while preserving the robustness and interpretability of classical recognition algorithms.

Accordingly, this study employs a deep convolutional neural network (CNN) to perform efficient extraction of informative recognition features from images for their subsequent use in an information-extreme classifier (Figure 1).

The model consists of three convolutional blocks, each comprising a convolutional layer, normalization (to stabilize the distribution of activations), and a subsampling operation that reduces dimensionality and enhances shift invariance.

To improve parametric efficiency, depthwise separable convolutions are employed, which factorize the standard $k \times k$ convolutional kernel into separate depthwise and pointwise operations. This reduces the number of parameters from $O(k^2 C_{in} C_{out})$ to $O(k^2 C_{in} + C_{in} C_{out})$. Formally, the spatial convolution operation for each channel is defined by Equation (2):

$$Y_{i,j,c} = \sum_{m=0}^{k-1} \sum_{n=0}^{k-1} X_{i+m,j+n,c} W_{m,n,c} + b_c , \qquad (2)$$

DOI 10.15588/1607-3274-2026-1-6

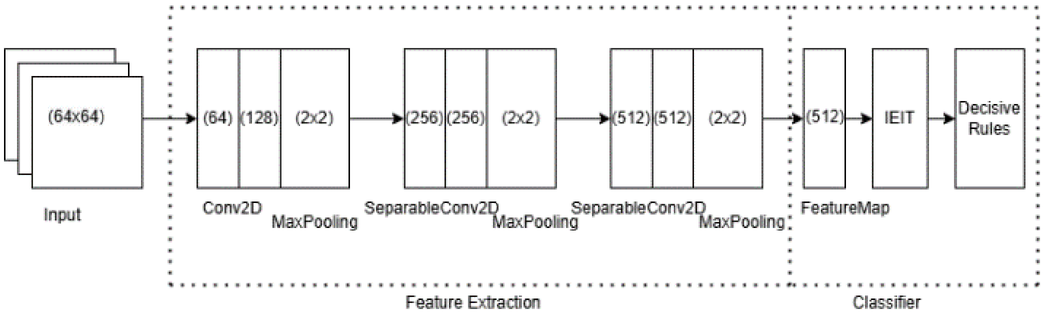


Figure 1 – Hybrid machine learning model architecture for the recognition system

The use of a global average pooling layer (GlobalAveragePooling2D) allows the aggregation of information across the entire activation map, significantly reducing the number of trainable parameters and lowering the risk of overfitting. Equation (3) expresses the computation of the average feature value for each channel c:

$$y_c = \frac{1}{H \times W} \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} X_{i,j,c} \ . \qquad (3)$$

The activation function used is Swish [12], which demonstrates improved properties compared to ReLU due to the absence of "dead neurons" and its smooth transition between activation regions:

$$\phi(x) = x \cdot \sigma(\beta x), \quad \sigma(x) = \frac{1}{1+e^{-x}}. \qquad (4)$$

Model weights are optimized using the AdamW algorithm. The total loss function, incorporating L2 regularization, is defined as:

$$L = -\sum_{i=1}^{N} y_i \log(\hat{y}_i) + \lambda \sum_j \left\| w_j \right\|^2 \ . \qquad (5)$$

The overall structure of the model is presented as a sequential architecture that combines a convolutional feature extractor with an information-extreme classifier (Figure 1).

At the feature extraction stage, three convolutional blocks are employed: the first uses a standard Conv2D operation, while the subsequent two utilize depthwise separable convolutions (SeparableConv2D), which reduce the number of parameters without compromising feature informativeness. Each convolutional block is followed by feature downsampling using MaxPooling.

In the final stage, features are aggregated into a fixed-length vector using GlobalAveragePooling2D, resulting in a 512-dimensional feature map. This vector is passed to the input of the information-extreme classifier (IEC), which constructs a relevant input training matrix, transforms it into a working binary matrix, and – by optimizing the machine learning parameters using an information criterion – reconstructs recognition class containers in the Hamming feature space.

At the output, decision rules are formed based on the optimized geometric parameters of the recognition class containers obtained during the training process.

To formalize the process of information-extreme machine learning, we present it as a functional categorical model that reflects the relationships among the sets involved in constructing optimal recognition class containers. This approach provides a clear representation of how the transition from input features to classification decisions is performed, taking into account both geometric and informational criteria.

Figure 2 illustrates the functional categorical model of information-extreme machine learning in the form of a directed graph, where the terminal set E of information criterion values for optimizing machine learning parameters is shared across all optimization loops.

At each step of the learning process, the operator ξ reconstructs recognition class containers in the radial basis of the feature space, forming, in the general case, a fuzzy partition $\tilde{\Re}^{|M|}$.

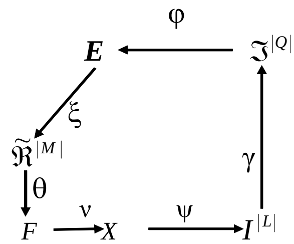


Figure 2 – Categorical model of information-extreme machine learning

The operator θ projects the constructed fuzzy partition $\tilde{\Re}^{|M|}$ onto the distribution of binary feature vectors of the binary training matrix X, while the operator ψ tests the main statistical hypothesis regarding the membership of feature vectors to the corresponding recognition class. Based on the results of statistical hypothesis testing, the set of statistical hypotheses $I^{|L|}$ is formed, and the operator γ generates the set of accuracy metrics $\Im^{|Q|}$, where $Q = L^2$. The set F corresponds to the feature vectors extracted by the convolutional neural network from raw input data. The operator ν performs the

DOI 10.15588/1607-3274-2026-1-6

transformation of features into a binary form X, which is subsequently used to construct fuzzy clusters $\tilde{\Re}^{|M|}$. The operator φ computes the set E of information criterion values used for optimizing the machine learning parameters.

The mappings between the sets in the graph correspond to the individual stages of the algorithm: starting with feature extraction, followed by binary vector construction, computation of informational measures, optimization of container parameters (radii, centers, tolerances), and culminating in the decision-making process regarding the class membership of a given object. This form of representation clearly delineates the logical structure of the system and the interrelationships between its components.

According to the categorical model (Figure 2), the information-extreme machine learning algorithm can be formalized as the following procedure:

$$P^* = \arg\max_{G_P} \{ \max_{G_R \cap \{S\}} \overline{E}_S \} . \tag{6}$$

Within the framework of the information-extreme intelligent (IEI) technology, machine learning under binary decision-making and equally probable a priori hypotheses is implemented through a targeted search for the global maximum of the information criterion, averaged over the alphabet of recognition classes. As the optimization criterion, we consider a modified version of the Kullback information measure, proposed by the authors, in the following form:

$$E_M^{(k)} = \frac{1}{n_m}[K_{1,m}^{(k)} - K_{2,m}^{(k)}] \times \times \log_2 [\frac{10^{-\omega} + n_m + [K_{1,m}^{(k)} - K_{2,m}^{(k)}]}{10^{-\omega} + n_m - [K_{1,m}^{(k)} - K_{2,m}^{(k)}]}]. \tag{7}$$

Thus, information-extreme machine learning consists in optimizing the geometric parameters of recognition class containers, reconstructed in the Hamming feature space, according to the information criterion.

4 EXPERIMENTS

To solve the problem of land cover image classification, data from the open EuroSAT dataset [13] were used. This dataset represents a collection of images corresponding to different types of terrain and land cover, based on Sentinel-2 satellite imagery obtained under the Earth observation program Copernicus.

The images used in this study were selected based on a proximity criterion, which a priori guaranteed their overlap in the recognition feature space, thereby determining the need for deep machine learning.

An alphabet consisting of six recognition classes was constructed, characterizing the following types of digital terrain image frames: “Annual Crop”, “Forest”, “Highway”, “Residential”, “Pasture” and “Industrial” (Figure 3). Each recognition class was characterized by a set of 2,000 images, each with a resolution of 64 by 64 pixels.

Figure 3 – Matrix of representative images for the recognition classes

All images were pre-sorted by class to ensure uniform representation of each category within the resulting subsets. The dataset was then stratified and divided into three parts: a training set (70% of the data), a validation set (15%), and a test set (15%), while preserving the proportional distribution of classes. This approach not only ensured balanced subsets but also minimized the risk of model overfitting on dominant classes.

Image augmentation (e.g., geometric transformations, brightness adjustments, flipping, etc.) was not applied, as the goal of the study was not to increase model robustness to variations in input data, but rather to evaluate the effectiveness of the hybrid feature extraction and classification approach under controlled conditions.

During the training stage of the feature extractor, a convolutional neural network consisting of three convolutional blocks with normalization and subsampling operations was used. After the final convolutional layer, GlobalAveragePooling2D was applied to aggregate spatial information across the entire activation map and to generate a fixed-size feature vector.

The model was trained for 50 epochs using the AdamW optimizer, which combines adaptive weight

DOI 10.15588/1607-3274-2026-1-6

updates with weight decay (L2 regularization), with a batch size of 32. The Swish activation function was used, as it provides smooth transitions across activation regions and eliminates the "dead neuron" problem.

After completing the training stage of the CNN feature extractor, the features obtained for each image were quantized into binary form based on a system of control tolerances. In this way, a binary training matrix of features was formed, which served as input data for the information-extreme classifier.

Within this approach, optimization of the geometric parameters of the class containers (radius, center, and tolerance field) was carried out based on the authors' modified version of the Kullback information criterion, taking into account statistical hypotheses regarding the belonging of realizations to their corresponding recognition classes.

To ensure reproducibility of the experiment, the random number generator was fixed, and identical initial conditions were applied at each model run.

To evaluate the effectiveness of the hybrid model, training was conducted for a convolutional neural network (CNN), which was used to extract features from the input images. Based on the resulting feature vectors, two types of classifiers were built: SVM and k-NN, operating on the feature representations formed by the CNN.

In order to enable a correct comparison of results, additional SVM and k-NN classifiers were implemented, which operated directly on the raw pixel values (flattened RGB images of size 64×64) without prior CNN-based feature extraction.

During the training stage of the CNN, a model was used consisting of three convolutional blocks with normalization and subsampling, followed by a GlobalAveragePooling2D layer to convert the spatial representation into a fixed-length feature vector.

The resulting feature vector was passed through a fully connected layer of size 512 with a Swish activation function and Dropout regularization (rate = 0.5).

The network was trained for 50 epochs using a batch size of 32, the AdamW optimizer (learning rate = 0.001, weight decay = 1e-4), and the sparse categorical crossetropy loss function.

For the purpose of comparative analysis of different classification algorithms, two experimental groups were formed, differing in the way the input data were preprocessed:

1. Baseline models: The SVM and k-NN classifiers were applied directly to the input images represented as unstructured vectors of pixel values (flattened RGB), without any prior feature extraction.

2. Hybrid models: The same classification algorithms (SVM, k-NN), as well as the information-extreme intelligent technology classifier (IEIT), were applied to feature vectors that had been previously extracted by the convolutional neural network (CNN).

This design enabled the evaluation of how prior deep image processing affects the accuracy and stability of classification.

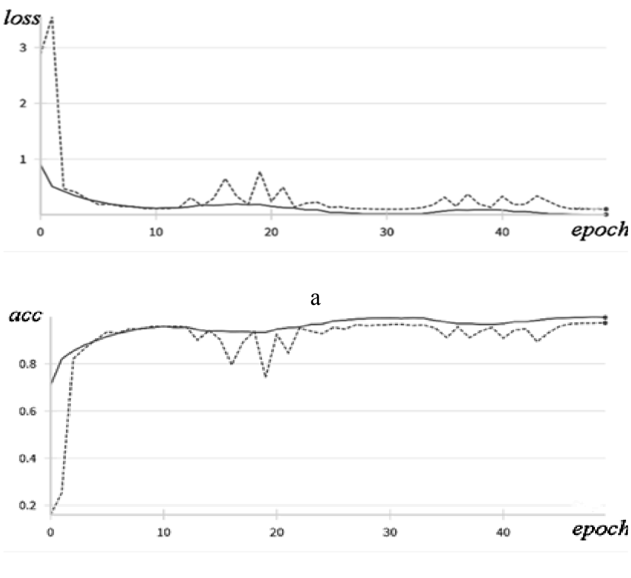
To ensure the reproducibility of the experiments, the random number generator was initialized with a fixed seed (random_state = 42).

This guarantees identical outcomes across repeated runs, as all procedures involving randomness – such as the stratified splitting of the dataset into training, validation, and test sets, model parameter initialization, and stochastic elements of training – were conducted under consistent conditions.

5 RESULTS

The software implementation of the hybrid algorithm was designed as a sequential procedure that includes training an artificial neural network to extract significant features from the training matrix and constructing decision rules based on the optimal geometric parameters of the recognition class containers, which were obtained during the process of information-extreme machine learning.

Figure 4 shows the plots illustrating the model's error reduction and accuracy as functions of training epochs. The analysis of Figure 4 reveals that at the early stages of machine learning, a significant gap is observed between the training accuracy (60.82%) and the validation accuracy (15.99%), which indicates the initial adaptation phase of the model before generalization is formed.



b

Figure 4 – Results of machine learning: a – model error reduction; b – model accuracy.

As the training progresses and the number of epochs increases, a trend toward stabilization is observed, and the model reaches a performance plateau at a level of 96–97% correct classification decisions.

At the same time, after the 35th epoch on the training set, the validation accuracy decreases, while the validation loss increases. Nevertheless, the regularization

DOI 10.15588/1607-3274-2026-1-6

mechanisms and adaptive machine learning strategies stabilize the accuracy at the level of 97%, which is likely a consequence of using cyclic learning rate adjustment to escape local minima. In the course of information-extreme machine learning, the recognition class containers are reconstructed through the optimization of their geometric parameters.

Figure 5 presents the plots showing the dependency of the information criterion (7) on the radii of the recognition class containers obtained during the process of information-extreme machine learning.

The analysis of the machine learning plots shown in Figure 5 demonstrates that all recognition classes from the defined alphabet formed valid working domains; that is, they are separable, since the first and second validity measures within the working domains exceed, respectively, the Type I and Type II error rates.

At the same time, the optimal radii of the recognition class containers (given in code units) are as follows: 53 for class X_1^0 – "Annual Crop", 230 for class X_2^0 – "Forest", 56 for class X_3^0 – "Highway", 55 for class X_4^0 – "Residential", 56 for class X_5^0 – "Pasture", and 55 for class X_6^0 – "Industrial".

The quality of machine learning was evaluated using a test set of 400 images per recognition class. The overall recognition accuracy during testing reached 95%.

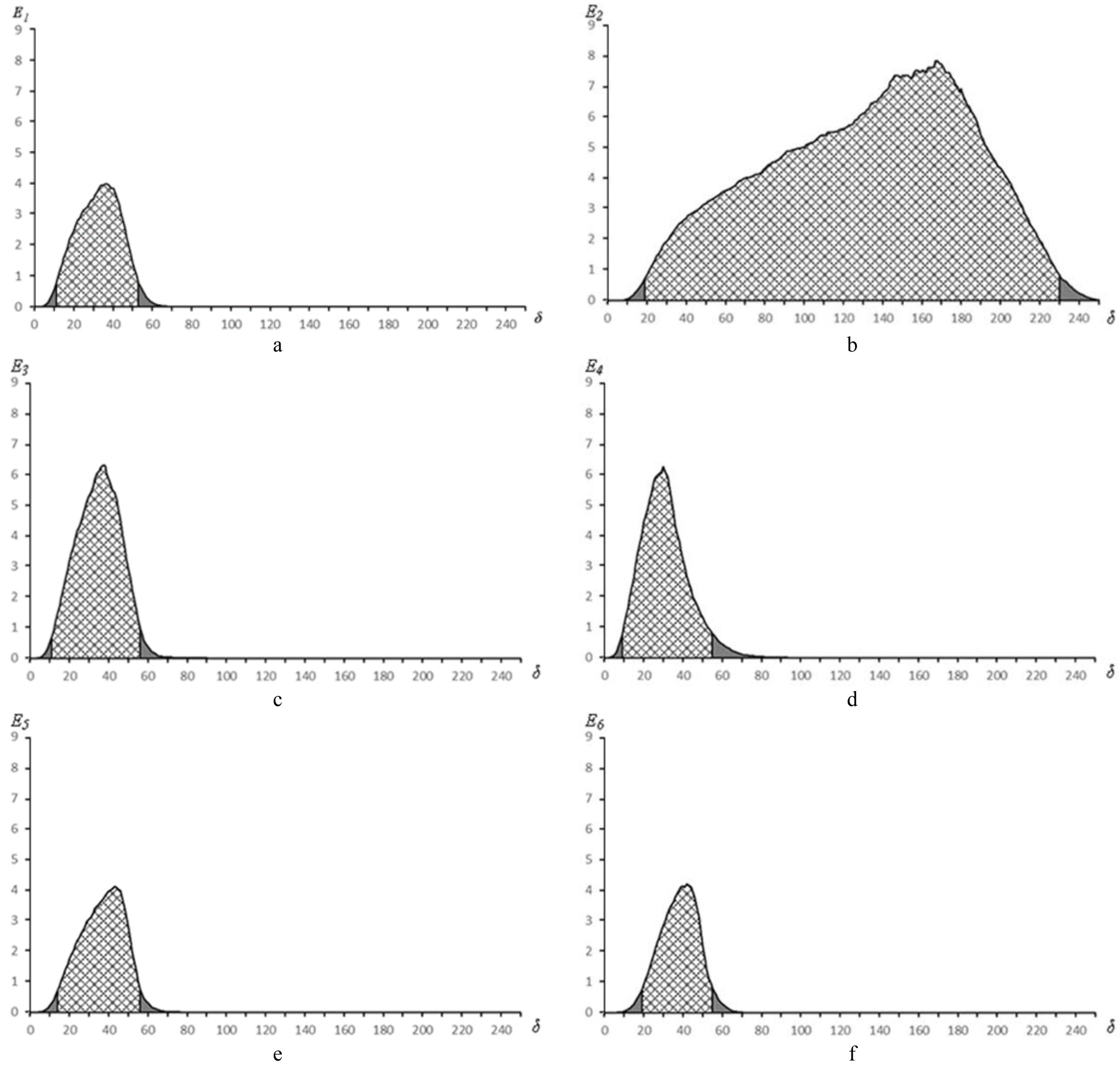


Figure 5 – Plots of the dependency between the information criterion and the radii of the recognition class containers: a – class X_1^0 "Annual Crop", b – class X_2^0 "Forest", c – class X_3^0 "Highway", d – class X_4^0 "Residential", e – class X_5^0 "Pasture", f – class X_6^0 "Industrial"

DOI 10.15588/1607-3274-2026-1-6

Special attention should be paid to the interpretability of the models. While deep neural networks with softmax output provide high classification accuracy, their internal mechanisms remain largely opaque to the user due to the large number of parameters and complex layer interactions.

The k-NN and SVM algorithms partially improve interpretability by allowing the analysis of proximity to training examples or positioning relative to the decision hyperplane.

At the same time, the proposed IEIT model ensures significantly greater transparency of decision-making due to its geometric interpretation in the form of recognition class containers in the space of binarized features. Unlike deep learning "black-box" models, this approach defines clear boundaries between classes in the form of hyperspherical regions, whose parameters – centers, radii, and tolerances – have explicit mathematical meaning.

This structure not only determines the class membership of an object but also allows assessing the degree of proximity or remoteness of the feature vector from the center of the corresponding container. In turn, this enables the identification of borderline or atypical instances.

Moreover, the geometric interpretation provides a foundation for building trust mechanisms into the model, which can be used to develop additional decision criteria or to detect anomalous cases that do not correspond to any of the training classes.

To summarize the obtained results visually, Table 1 presents a comparative analysis of the efficiency of different image classification approaches. The comparison includes both classical algorithms applied directly to raw pixel data and modern neural and hybrid architectures with prior feature extraction. In addition to classification accuracy, the models are also evaluated in terms of their generalization ability via macro F1-score, class-wise recall range, and interpretability of the decision process. This allows for a comprehensive assessment not only of classification performance but also of the model's practical suitability for analysis and real-world deployment.

Table 1 – The fragment of experimental results on model building by the formed samples

№	Architecture	Features	Classifier	Accuracy	Macro F1	Recall (min/max)	Interpretability
1	Raw pixels + SVM	64×64 grayscale	SVM	47.3%	0.44	0.10 / 0.92	low
2	Raw pixels + k-NN	64×64 grayscale	k-NN	38.6%	0.27	0.00 / 0.99	low
3	CNN + Softmax	CNN-features	Softmax	93.8%	0.94	0.80 / 0.99	limited
4	CNN + k-NN	CNN-features	k-NN	96.75%	0.97	0.94 / 0.99	limited
5	CNN + SVM	CNN-features	SVM	96.4%	0.96	0.93 / 0.99	limited
6	CNN + IEIT	CNN binary features	IEIT	96.5%	0.95	0.93 / 0.99	high

6 DISCUSSION

The analysis of the plots presented in Figure 4 demonstrates a typical dynamic of adaptation during the early stages of machine learning: the noticeable difference between accuracy on the training and validation sets gradually decreases as the model begins to form generalizations. After epochs 30–35, the accuracy reaches a plateau, which indicates the stabilization of the learning process. An increase in validation loss in the later stages signals the beginning of overfitting. Nevertheless, the application of regularization techniques – particularly weight decay implemented in the AdamW optimizer – and the use of the GlobalAveragePooling layer made it possible to maintain consistently high validation accuracy at the level of 96–97%.

The information-extreme component of the model shows a clearly expressed dependence of the information measure on the geometric parameters of the recognition class containers (Figure 5), which indicates the effectiveness of using the information criterion for optimizing machine learning parameters. In particular, for each recognition class, a specific optimal range of radius values was determined, within which the maximum of the information measure (7) is achieved. This measure takes into account the differences between intra-class and inter-class distributions. Such optimization allows the formation of compact yet sufficiently capacious hyperspherical containers that accurately reflect the statistical regularities of the input data.

The obtained optimal parameters confirm the model's ability to form well-separated regions in the multidimensional feature space, with features preliminarily extracted by the convolutional neural network. As a result, each recognition class is provided with an individually adapted geometric representation, which reduces the risk of inter-class overlap and improves the accuracy of classification decisions. Moreover, the presence of clearly pronounced maxima of the functional indicates the model's stability to small variations in the container parameters – an important property for practical use in conditions of limited or partially noisy input data.

Compared to the classical CNN architecture that uses a softmax classifier, the proposed hybrid approach has several advantages. First, the model provides higher interpretability due to the geometric representation of recognition classes in the form of hyperspherical containers. Second, the separation of feature extraction and classification processes allows for more flexible control of the system's behavior, in particular, the ability to adapt the classifier to changes in the feature space without the need to retrain the entire network. This is especially important in cases with small datasets or a high level of inter-class overlap.

At the same time, the construction of class containers in the space of binarized features has certain limitations. This approach assumes the isotropy of the feature space, which is not always satisfied in practice. Additionally, the effectiveness of decision-making is highly dependent on

DOI 10.15588/1607-3274-2026-1-6

the proper tuning of parameters – particularly the control tolerance δ and the boundary values of the recognition class container radii require prior optimization or the implementation of adaptive mechanisms.

During the experimental study, a comparative evaluation of various image classification approaches was carried out. The use of classical linear classifiers (SVM and k-NN) directly on pixel-level image representations proved to be ineffective: SVM achieved an accuracy of only 47.3%, while k-NN reached 38.6%. These results confirm the limitations of raw pixel features and emphasize the necessity of deep preprocessing of input data.

Applying a softmax classifier within the convolutional neural network framework significantly improved classification accuracy – up to 93.8%. However, a more detailed analysis revealed considerable discrepancies in the recognition performance of individual classes, particularly in cases with high visual similarity between them.

The highest performance was achieved using hybrid models that combine CNN-based feature extraction with classification via separate algorithms. The CNN + SVM model achieved an accuracy of 96.4%, while CNN + k-NN reached 96.75%. These results indicate that such hybrid approaches not only ensure high recognition accuracy but also preserve the structural coherence and stability of the model's behavior.

Thus, the hybrid architecture demonstrates superiority over both classical and standalone neural approaches, especially in tasks where a combination of high accuracy, adaptability, and interpretability of results is required.

Based on the confusion matrices (Figure 6), a more detailed analysis of model classification quality can be

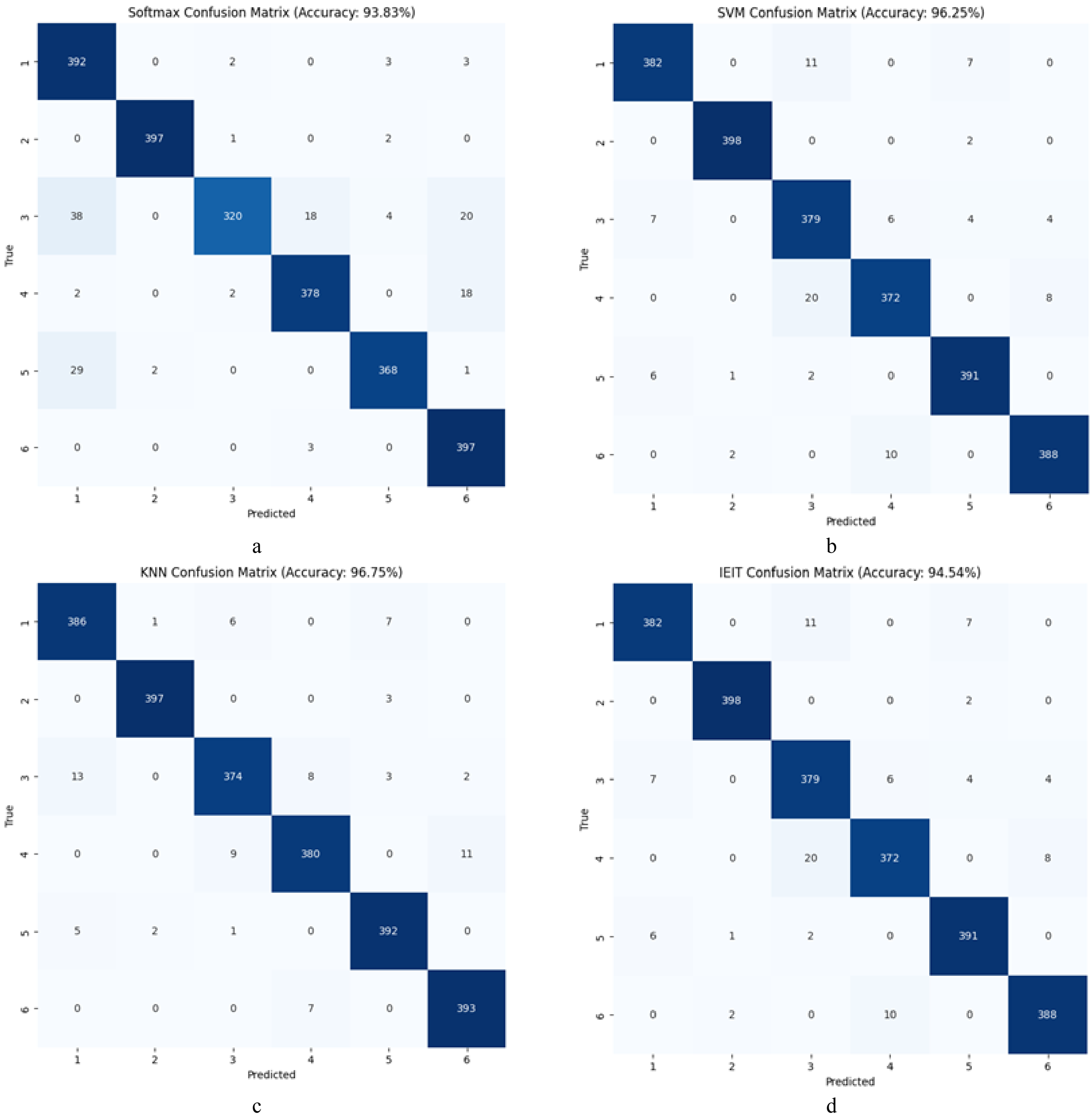


Figure 6 – Confusion matrixes for classification neural based models:
a – CNN + Softmax, b – CNN + k-NN, c – CNN + SVM, d – CNN + IEIT

DOI 10.15588/1607-3274-2026-1-6

performed. Despite an overall accuracy of 93.83%, the softmax classifier exhibits a pronounced bias in recognizing class X_3^0 (“Highway”) – with 38 misclassifications toward the “Annual Crop” class and an additional 42 errors scattered across other classes. This behavior indicates a high degree of inter-class overlap for this category and an insufficient capacity of softmax-based decisions to separate closely situated representations in the feature space.

The CNN + k-NN (96.75%) and CNN + SVM (96.25%) models demonstrate a significantly more stable error distribution. In particular, the number of misclassifications in the “Highway” and “Residential Area” classes has been reduced by at least half. This indicates that geometric methods relying on local or global decision boundaries in the feature space perform better in cases of partial overlap between classes.

The proposed CNN + IEIT model achieves an overall accuracy of 94.54%, with its confusion matrix nearly replicating the pattern of errors observed with the SVM classifier. However, unlike SVM, IEIT operates with clearly defined geometric containers featuring adaptive radii and enables the identification of uncertainty zones. For the “Highway” and “Residential” recognition classes, the number of errors is identical to that of SVM, though IEIT exhibits better clarity in other classes (e.g., “Annual Crop”), avoiding mixing with more distant categories.

CONCLUSIONS

The relevant task of constructing an interpretable and statistically grounded image classification system based on deep feature extraction has been addressed through the development of a hybrid model that combines a convolutional neural network (CNN) with information-extreme machine learning (IEIML).

The scientific novelty of the obtained results lies in the proposed approach that enables the formation of hyperspherical recognition class containers in the feature space generated by the CNN, followed by optimization of their parameters using an information-based criterion. This approach ensures a balance between classification accuracy and interpretability of the decision-making process.

The practical significance of the proposed solution is defined by the possibility of effectively applying the model for satellite image classification within a modular architecture, where the feature extractor and the classifier can be independently adapted or enhanced for various tasks and datasets.

Prospects for further research include the adaptation of the proposed hybrid architecture to object detection tasks, particularly the localization of multiple objects within a single image and the extension of the categorical model to operate with spatially oriented representations.

ACKNOWLEDGEMENTS

The research was concluded in the Computer Science Department at Sumy State University with the financial support of the Ministry of Education and Science of Ukraine in the framework of state budget scientific and research work of DR No. 0121U109466 “Intelligent Technologies in Cybersecurity”.

DECLARATIONS

Conflict of interest: The authors declare that they have no conflict of interest in relation to this research, whether financial, personal, authorship, or otherwise, that could affect the research and its results presented in this paper.

Authors’ contributions: Anatolii Dovbysh: research concept and methodology, scientific supervision, manuscript review;

Vladyslav Piatachenko: coordination of the research activities, hybrid system architecture, experimental design and analysis;

Viktor Serhieiev: development and investigation of hybrid machine learning models, parameter tuning, validation of information-extreme learning algorithms;

Oleksii Hrytsenko: image preparation and processing, training of neural network feature extractors, evaluation and visualization of results.

Data availability: The manuscript has no associated data.

Software availability: The software will be made available on reasonable request.

Use of artificial intelligence tools: The authors confirm that they did not use artificial intelligence technologies in creating the submitted work.

REFERENCES

1. Ab Wahab M. N., Nazir A., Ren A. T. Z., Noor M. H. M., Akbar M. F., Mohamed A. S. A. Efficientnet-Lite and Hybrid CNN-KNN Implementation for Facial Expression Recognition on Raspberry Pi, *IEEE Access*, 2021, Vol. 9, pp. 134065–134080. DOI:10.1109/ACCESS.2021.3113337
2. Ghosh S., Singh A., Kavita, Jhanjhi N. Z., Masud M., Aljahdali S. SVM and KNN Based CNN Architectures for Plant Classification, *Computers, Materials and Continua,* 2022, Vol. 71, № 3, P. 4257. DOI:10.32604/CMC.2022.023414
3. Lanjewar M. G., Parab J. S., Shaikh A. Y. Development of framework by combining CNN with KNN to detect Alzheimer’s disease using MRI images, *Multimedia Tools and Applications,* 2023, Vol. 82, № 8, pp. 12699–12717. DOI:10.1007/S11042-022-13935-4/METRICS
4. Moskalenko V., Kharchenko V., Moskalenko A., Petrov S. Model and Training Method of the Resilient Image Classifier Considering Faults, Concept Drift, and Adversarial Attacks, *Algorithms,* 2022, Vol. 15, № 384, pp. 1–24. DOI:10.3390/a15100384
5. Moskalenko V. V., Moskalenko A. S., Korobov A. G., Zaretsky M. O. Image Classifier Resilient To Adversarial Attacks, Fault Injections And Concept Drift – Model Architecture And Training Algorithm, *Radio Electronics, Computer Science, Control,* 2022, Vol. 3, № 86, pp. 1–16. DOI:10.15588/1607-3274-2022-3-9
6. Shelehov I., Prylepa D., Khibovska Yu. Information-extreme machine learning of an ophthalmic diagnostic system with a hierarchical class structure, *Artificial Intelligence*, 2024, Vol. 29, №3, pp. 114–125. DOI:10.15407/JAI2024.03.114

DOI 10.15588/1607-3274-2026-1-6

7. Dovbysh A. S. Shelekhov I. V., Prylepa D. V., Khibovska Yu. O., Nikitenko K. O. Information-Extreme Method For Ball Detection In Intelligent Video Analysis Systems Of Volleyball Matches, *Visnyk Kremenchutskoho natsionalnoho universytetu imeni Mykhaila Ostrohradskoho*, 2024, Vol. 5, pp. 41–51. DOI:10.32782/1995-0519.2024.5.6
8. Dovbysh A. S., Budnyk M. M., Piatachenko V. Y., Myronenko M. I. Information-extreme machine learning of on-board vehicle recognition system, *Cybernetics and Systems Analysis*, 2020, Vol. 56, pp. 534–543. DOI:10.1007/s10559-020-00269-y
9. Naumenko I., Piatachenko V., Myronenko M., Savchenko T. Information-Extreme Machine Learning of an On-board Ground Object Recognition System with a Choice of a Base Recognition Class, *6th International Conference on Computational Linguistics and Intelligent Systems, Gliwice, 12–13 May 2022: proceedings*. Gliwice, CEUR, 2022, pp. 1139–1148.
10. Tan M., Le Q. V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, *36th International Conference on Machine Learning, 9th June 2019: proceedings*. Long Beach:arXiv, 2019, pp. 10691–10700. DOI: 10.48550/arXiv.1905.11946
11. Sandler M., Howard A., Zhu M., Zhmoginov A., Chen L. C. MobileNetV2: Inverted Residuals and Linear Bottlenecks, *Conference on Computer Vision and Pattern Recognition, Salt Lake City, 2018: proceedings*. Salt Lake City, IEEE, 2018, pp. 4510–4520. DOI:10.1109/CVPR.2018.00474
12. Dasgupta R., Chowdhury Y. S., Nanda S. Performance Comparison of Benchmark Activation Function ReLU, Swish and Mish for Facial Mask Detection Using Convolutional Neural Network, *Algorithms for Intelligent Systems, Singapore, 2021: proceedings*. Singapore, Springer, 2021, pp. 355–367. DOI:10.1007/978-981-16-2248-9_34
13. Helber P., Bischke B., Dengel A., Borth D. Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2019, pp. 1–9. DOI: 10.48550/arXiv.1709.00029

Received 13.08.2025.
Accepted 17.11.2025.
Published 27.03.2026.

УДК 004.93

ГІБРИДНА СИСТЕМА РОЗПІЗНАВАННЯ СУПУТНИКОВИХ ЗОБРАЖЕНЬ З НЕЙРОМЕРЕЖЕВИХ ЕКСТРАКТОРОМ ТА ІНФОРМАЦІЙНО ЕКСТРЕМАЛЬНИМ КЛАСИФІКАТОРОМ

Довбиш А. С. – д-р техн. наук, професор, професор кафедри комп’ютерних наук Сумського державного університету, Суми, Україна. ROR: https://ror.org/01w60n236. ORCID:https://orcid.org/0000-0003-1829-3318.

П’ятаченко В. Ю. – канд. техн. наук, асистент кафедри комп’ютерних наук Сумського державного університету, Суми, Україна. ROR: https://ror.org/01w60n236. ORCID: https://orcid.org/0000-0002-7464-3119.

Сергєєв В. М. – аспірант кафедри комп’ютерних наук Сумського державного університету, Суми, Україна. ROR: https://ror.org/01w60n236. ORCID:https://orcid.org/0009-0009-3838-0153.

Гриценко О. М. – аспірант кафедри комп’ютерних наук Сумського державного університету, Суми, Україна. ROR: https://ror.org/01w60n236. ORCID: https://orcid.org/0009-0004-1382-708X.

АНОТАЦІЯ

Актуальність. Розв’язано актуальну задачу побудови інтерпретованої та адаптивної системи розпізнавання для семантичної сегментації супутникових знімків місцевості шляхом поєднання нейромережевих екстракторів з інформаційно-екстремальним класифікатором .

Мета роботи. Підвищення точності класифікації супутникових знімків місцевості шляхом розробки гібридної моделі машинного навчання, яка об’єднує глибоку згорткову нейронну мережу для відбору інформативних ознак та інформаційно-екстремальний класифікатор, що дозволяє будувати високодостовірні вирішальні правила за умови перетину класів розпізнавання в просторі ознак.

Метод. Запропоновано гібридну модель, яка поєднує ефективну екстракцію просторових ознак за допомогою згорткової нейронної мережі та інформаційно-екстремальну інтелектуальну технологію аналізу даних, яка базується на максимізації інформаційної спроможності системи розпізнавання в процесі машинного навчання. Водночас при екстракції ознак розпізхнавання замість класичного Flatten використано GlobalAveragePooling для узагальнення ознак, а також впроваджено регуляризаційні механізми, зокрема вагове затухання та циклічне навчання. Оптимізація параметрів інформаційно-екстремального машинного навчання виконується за модифікованим авторами інформаційним критерієм Кульбака, який розглядається як міра різноманітності класів розпізнавання.

Результати. Побудована модель забезпечує високу точність класифікації (95%) при тестуванні, а також демонструє стабільність та підвищення оперативності нейромережевого екстрактора шляхом зменшення кількості епох його навчання завдяки застосованню регуляризації. За результатами інформаційно-екстремального машинного навчання визначено оптимальні геометричні параметри гіперсферичних контейнерів класів розпізнавання, що дозволяє побудувати високодостовірні вирішальні правила за умови перетину класів розпізнаванняс в просторі ознак.

Висновки. Запропонована гібридна модель дозволяє для апріорно нечіткого розбиття в просторі ознак класів розпізнавання за сформованою в результаті екстракції ознак вхідною навчальною матрицею побудувати в процесі інформаційно-екстремального машинного навчання високодостовірні вирішальні правила.

КЛЮЧОВІ СЛОВА: інформаційно-екстремальне машинне навчання, згорткова нейронна мережа, інформаційний критерій, оптимізація, гібридна модель, екстракція ознак розпізнавання, зображення місцевості.

DOI 10.15588/1607-3274-2026-1-6

ЛІТЕРАТУРА

1. Efficientnet-Lite and Hybrid CNN-KNN Implementation for Facial Expression Recognition on Raspberry Pi / [M. N. Ab Wahab, A. Nazir, A. T. Z. Ren et al.] // IEEE Access. – 2021. – Vol. 9. – P. 134065–134080. DOI:10.1109/ACCESS.2021.3113337
2. SVM and KNN Based CNN Architectures for Plant Classification / [S. Ghosh, A. Singh, Kavita, N. Z. Jhanjhi et al.] // Computers, Materials and Continua. – 2022. – Vol. 71, № 3. – P. 4257. DOI:10.32604/CMC.2022.023414
3. Lanjewar M. G. Development of framework by combining CNN with KNN to detect Alzheimer's disease using MRI images / M. G. Lanjewar, J. S. Parab, A. Y. Shaikh // Multimedia Tools and Applications. – 2023. – Vol. 82, № 8. – P. 12699–12717. DOI:10.1007/S11042-022-13935-4/METRICS
4. Model and Training Method of the Resilient Image Classifier Considering Faults, Concept Drift, and Adversarial Attacks / [V. Moskalenko, V. Kharchenko, A. Moskalenko, S. Petrov] // Algorithms. – 2022. – Vol. 15, № 384. – P. 1–24. DOI:10.3390/a15100384
5. Image Classifier Resilient To Adversarial Attacks, Fault Injections And Concept Drift – Model Architecture And Training Algorithm / [V. V. Moskalenko, A. S. Moskalenko, A. G. Korobov, M. O. Zaretsky] // Radio Electronics, Computer Science, Control. – 2022. – Vol. 3, № 86. – P. 1–16. DOI:10.15588/1607-3274-2022-3-9
6. Shelehov I. Information-extreme machine learning of an ophthalmic diagnostic system with a hierarchical class structure / I. Shelehov, D. Prylepa, Yu. Khibovska // Artificial Intelligence. – 2024. – Vol. 29, №3. – P. 114–125. DOI:10.15407/JAI2024.03.114
7. Інформаційно-Екстремальний Метод Детектування М'яча В Системах Інтелектуального Відеоаналізу Волейбольного Матчу / [А. С. Довбиш, І. В. Шелехов, Д. В. Прилепа та ін.] // Вісник КрНУ імені Михайла Остроградського – 2024. – Вип. 5. – С. 41–51. DOI:10.32782/1995-0519.2024.5.6
8. Information-extreme machine learning of on-board vehicle recognition system / [A. S. Dovbysh, M. M. Budnyk, V. Y. Piatachenko, M. I. Myronenko] // Cybernetics and Systems Analysis. –2020. – Vol. 56. – P. 534–543. DOI:10.1007/s10559-020-00269-y
9. Information-Extreme Machine Learning of an On-board Ground Object Recognition System with a Choice of a Base Recognition Class / [I. Naumenko, V. Piatachenko, M. Myronenko, T. Savchenko] // 6th International Conference on Computational Linguistics and Intelligent Systems, Gliwice, 12–13 May 2022: proceedings. – Gliwice : CEUR, 2022. – P. 1139–1148.
10. Tan M. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks / M. Tan, Q. V. Le // 36th International Conference on Machine Learning, 9th June 2019: proceedings. – Long Beach:arXiv,2019. – P. 10691–10700. DOI: 10.48550/arXiv.1905.11946
11. MobileNetV2: Inverted Residuals and Linear Bottlenecks / [M. Sandler, A. Howard, M. Zhu et al.] //Conference on Computer Vision and Pattern Recognition, Salt Lake City, 2018: proceedings. – Salt Lake City: IEEE, 2018. – P. 4510–4520. DOI:10.1109/CVPR.2018.00474
12. Dasgupta R. Performance Comparison of Benchmark Activation Function ReLU, Swish and Mish for Facial Mask Detection Using Convolutional Neural Network / R. Dasgupta, Y. S. Chowdhury, S. Nanda // Algorithms for Intelligent Systems, Singapore, 2021: proceedings – Singapore: Springer, 2021. – P. 355–367. DOI:10.1007/978-981-16-2248-9_34
13. Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification / [P. Helber, B. Bischke, A. Dengel, D. Borth] // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. – 2019. – P. 1–9. DOI: 10.48550/arXiv.1709.00029

DOI 10.15588/1607-3274-2026-1-6