UDC: 004.891.032.26:629.7.01.066

A MODEL AND TRAINING ALGORITHM OF SMALL-SIZED OBJECT DETECTION SYSTEM FOR A COMPACT AERIAL DRONE

Moskalenko V. V. – PhD, Associate professor of Computer Science department, Sumy State University, Sumy, Ukraine.

Moskalenko A. S. – PhD, Teaching assistant of Computer Science department, Sumy State University, Sumy, Ukraine.

Korobov A. G. – Postgraduate student of Computer Science department, Sumy State University, Sumy, Ukraine.

Zaretsky M. O. – Postgraduate Student of Computer Science department, Sumy State University, Sumy, Ukraine.

ABSTRACT

Context. Lightweight model and effective training algorithm of on-board object detection system for a compact drone are developed. The object of research is the process of small object detection on aerial images under computational resource constraint and uncertainty caused by small amount of labeled training data. The subject of the research are the model and training algorithm for detecting small objects on aerial imagery.

Objective. Goal of the research is developing efficient model and training algorithm of object detection system for a compact aerial drone under conditions of restricted computing resources and the limited volume of the labeled training set.

Methods. The four-stage training algorithm of the object detector is proposed. At the first stage, selecting the type of deep convolutional neural network and the number of low-level layers that is pre-trained on the ImageNet dataset for reusing takes place. The second stage involves unsupervised training of high-level convolutional sparse coding layers using modification of growing neural gas to automatically determine the required number of neurons and provide optimal distributions of the neurons over the data. Its application makes it possible to utilize the unlabeled training datasets for the adaptation of the high-level feature description to the domain application area. At the third stage, there are reduction of output feture map using principan component analysis and building of decision rules. In this case, output feature map is formed by concatenation of feature maps from different level of deep network using upscaling upper maps to uniform shape for each channel. This approach provides more contextual information for efficient recognition of small objects on aerial images. In order to build classifier of output feature map pixels is proposed to use boosted information-extreme learning. Besides that, the extreme learning machine is used to build of regression model for predict bounding box of detected object. Last stage involves fine-tuning of high-level layers of deep network using simulated annealing metaheuristic algorithm in order to approximating the global optimum of complex criterion of training efficiency of detection model.

Results. The results obtained on open datasets testify to the suitability of the model and training algorithm for practical usage. The proposed training algorithm utilize 500 unlabeled and 200 labeled training samples to provide 96% correctly detection of objects on the images of the test dataset.

Conclusions. Scientific novelty of the paper is a new model and training algorithm for object detection, which enable to achieve high confidence of recognition of small objects on aerial images under computational resource constraint and limited volume of the labeled training set. Practical significance of the paper results consists in the developed model and training algorithm made it possible to reduce requirements for size of labeled training set and computation resources of on-board detection system of aerial drone in training and inference modes.

KEYWORDS: growing neural gas, convolutional neural network, boosting, objects detector, information criterion, simulated annealing algorithm, extreme learning.

ABBREVIATIONS

CNN is a convolutional neural network;

ELM is a extreme learning machine;

IoU SHFN is a Intersection over Union;

is a single-hidden-layer feedforward network

PCA is a Principal Component Analysis;

SSD is a Single Shot Detector;

VGG is a Visual Geometry Group;

XOR is a exclusive OR logical operation;

YOLO is a You Only Look Once.

NOMENCLATURE

B_k is a set of ground-truth bounding boxes which correspond to objects of interest on k-th image;

b_r is a the bias of the r-th hidden node;

b_z is a support vector of data distribution in class X_z^o;

$\{d\}$ is a set of concentric radii of hyperspherical container in binary Hamming space;

E_z is a training efficiency criterion of decision rule for X_z^o class;

e_{ξ_1} is a ξ_1-th parameter which impacts on feature representation, $\xi_1 = \overline{1, \Xi_1}$;

f_{ξ_2} is a ξ_2-th parameter which impacts on efficiency of decision rules, $\xi_2 = \overline{1, \Xi_2}$;

H is a the hidden layer output matrix of the SHFN;

I_k is a k-th RGB image

IoU_i is a IoU between ground truth box and appropriate i-th predicted box;

J_{Cls} is a information effectiveness criterion of classification analysis;

J_{Loc} is a effectiveness criterion of bounding box prediction;

K_1 is a the size of training sets;

K_2 is a the size of test sets;

M is a number of regression model outputs;

N is a dimension of instance;

N_2 is a number of induced binary features;

n is a size of dataset;

n_z is a the volume of the training set of X_z^o class;

o_j is a the output of the network with respect to the j-th input vector x_j;

uniform_random random is the function of generation of a random number from the uniformed distribution from the assigned range;

step_size is the size of a range of the search for new solutions, neighboring to $s_{current}$;

$s_{current}$ is a current solution of simulated annealing algorithm

w_r is a the weight vector linking the input layer with the r-th hidden node, $w_r = (w_{r1}, w_{r2}, ..., w_{rN})^T$;

Δw_{s0} is the correction vector of weight coefficients of the neuron-winner;

Δw_{sn} is the correction vector of weight coefficients of the topological neighbors of neuron-winner;

x_j is a j-th input vector, $x_j = (x_{j1}, x_{j2}, ..., x_{jN})^T$;

y is a the target output vectors, $y = (y_1, y_2, ..., y_n)^T$;

y_j is a label of j-th instance;

α_z is a false-positive rates of classification decisions regarding belonging of input vectors to the X_z^o class;

β_r is a the weight vector linking the output layer with the r-th hidden node, $\beta_r = (\beta_{r1}, \beta_{r2}, ..., \beta_{rM})$;

β_z is a false-negative rates of classification decisions regarding belonging of input vectors to the X_z^o class;

ε_b is the constants of the update force of weight coefficients of the neuron-winner;

ε_n is the constants of the update force of weight coefficients of topological neighbors of the neuron-winner;

ς is an any small non-negative number;

η_0 is the initial value of learning rate;

η_{final} is the final value of learning rate;

η_t is the current value of learning rate;

$\varphi(x)$ is a activation function that can be any bounded non-constant piecewise continuous functions;

$\hat{n}$ is a the number of matched predicted box.

INTRODUCTION

Aerial drones are widely used in the tasks of search, rescue, remote inspection and robotic aerial security services. One of the ways to increase the functional efficiency of the aerial drone is to integrate artificial intelligence technologies for on-board sensors data

analysis. The surveillance cameras is the most informativeness sensor and object detection function is in demand. Development of accurate detectors of visual objects of interest is a promising direction, but the limited computing resources and the weight of the drone complicate the task. Resource restrictions do not make it possible to implement in compact drone the models of visual data analysis, adapted to a full range of possible observation conditions and a variety of modifications to objects of interest. This causes the need for the development of computationally effective models and algorithms of adaptation to new condition of operation, inherent in the specific field of application area.

In terms of computational efficiency and generalizing power, the leader among the models for visual information analysis is convolutional neural networks. However, training and retraining of convolutional networks requires significant computing resources and large amount of labeled training samples. It is possible to reduce computational load of retraining for adaptation of model to the new conditions of functioning due to the use of the transfer learning techniques. It based on copying the first layers of the network, trained on the ImageNet dataset or another large dataset [1, 2]. However, the layers of high-level feature representation needs learning from scratch. In this case, it is difficult to estimate in advance the required number of neurons in each convolutional layer. Therefore, to use the principles of growing neural gas is promising approach to unsupervised learning of high-level layers and to determine automatically the necessary number of neurons. In this case, the output layers of detector model require fine-tuning, which is typically implemented as one of modifications of the error backpropagation algorithm [2, 3]. However, this algorithm is characterized by a low convergence rate and getting stuck in local minima of loss function. There are alternative metaheuristic search optimization algorithms, however, effectiveness of using these algorithms in problems of networks fine tuning is scantily explored [3].

That is why the research, aimed at development of model and effective training algorithm of the objects detector under conditions of limited computing resources and learning data, is relevant.

1 PROBLEM STATEMENT

Let $D_{train} = \{ I_k, B_k \mid k = \overline{1, K_1} \}$ and $D_{test} = \{ I_k, B_k \mid k = \overline{1, K_2} \}$ are collected training and test datasets, respectively. Let the set $\{ X_z^o \mid z = \overline{1, Z} \}$ is characterized by small-sized objects on aerial photos, be given. In this case, total number of ground-truth bounding boxes per class does not exceed 200 samples. Moreover, the structure of the vector of model parameters of object detector is known

$$g = < e_1, ..., e_{\xi_1}, ..., e_{\Xi_1}, f_1, ..., f_{\xi_2}, ..., f_{\Xi_2} >,$$
$$\Xi_1 + \Xi_2 = \Xi.$$

In this case, the constraints $R_{\xi_1}(e_1,...,e_{\xi_1},...,e_{\Xi_1}) \le 0$, $R_{\xi_2}(f_1,...,f_{\xi_2},...,f_{\Xi_2}) \le 0$ are impose on parameters.

It is necessary to find optimal values of parameters g (1) which provide to achieve the maximum value of complex criterion J

$$J = J_{Cls} \cdot J_{Loc}. \qquad (2)$$

$$g^* = \arg\max_{G} \{J(g)\}. \qquad (3)$$

When the object detector functions in its inference mode, it is necessary to provide high confidence of localization and classification of objects of interest on test images.

2 REVIEW OF THE LITERATURE

The works in [4, 5] proposed the models of the visual feature extraction based on Haar-like filters, histograms of oriented gradients, local binary patterns, histograms of visual words and other models of local information descriptors. In this case, high-level contextual information is ignored that leads to decrease effectiveness of smile-sized objects detection under conditions of limited volume of training sets. In addition, non-hierarchical feature representation models are characterized by high labor consuming computing under conditions of a large variation of observations [5].

In the image analysis problems, numerous models of hierarchical feature representation model based on CNN are widely used [6]. VGG-16, VGG-19, ResNet-50, GoogleNet, MobileNet, SqueezeNet are the most popular of them [5, 7]. These networks differ in the number of layers, existence of residual connections and multiscale filters in each of the layers. In this case, it known that the models trained on dataset ImageNet accumulate in themselves important information regarding the analysis of visual images [5, 8]. Despite many target domains are far from the ImageNet context, the few layers of trained networks are possible to be reused. In addition, narrowing the scope of application, for example, by reducing the number of classes, makes it possible to reduce the resource requirements of objects detector.

It was proposed in works [8, 9] to perform fine-tuning of object detector based on convolutional network VGG-16 using mini-batch stochastic gradient descent. However, this would require a significant volume of training set and a few days of working on the graphics processing unit for successful training. In work [10], it was proposed to scan of the normalized high-level feature map by a sliding window, in each position of which classification analysis was carried out. In research work [11], it was proposed to carry out classification analysis of a high-level feature representation using information-extreme decision rules. The main idea of this approach is to transform input space of primary features into binary Hamming space where building radial-basis decision rules. This approach provides high computation efficiency because simple operation of comparison with thresholds and Hamming distance calculation based on logical XOR and counting are used. However, solutions of issue of adaptation of high-level layers of feature extractor to domain area of use and fast optimization algorithm of thresholds for features was not proposed. Random forest based feature induction and boosted similarity sensitive coding are two promising approaches to speedup thresholds optimization for binary feature encoding, but integration aspects are not investigated [12].

The works in [7, 13] proposed the unsupervised learning of convolutional layers based on autoencoder or the restricted Boltzmann machine that require a large amount of resources for obtaining an acceptable result. In articles [14, 15], it is proposed to combine the principles of neural gas and sparse coding for learning convolutional filters for unlabeled datasets. This approach has a soft competitive learning scheme, which facilitates robust convergence to close to optimal distributions of the neurons over the data. In this case, embedding the sparse coding algorithms makes it possible to improve the noise immunity and generalization ability of feature representation. However, the number of neurons is not known in advance and is assigned at the discretion of a developer.

Usage of growing neural gas principles to automatically determine the required number of neurons is a promising approach to training of high-level convolutional layers of convolutional network [15]. However, the mechanism of the insertion of new neurons based on setting the insertion period leads to distortion of the learned structures and instability of the learning process. However, it was shown in article [16] that it is possible to ensure stability of learning by replacement of the neurons insertion period with the threshold of a maximum distance of a neuron from every datapoint of the training set, referred to it. However, the mechanisms of neurons updating for adaptation of the learning process to the sparse representation of observations have not been revised yet.

The problem of objects detection based on feature maps of convolutional network is solved by using detection layers YOLO, Faster R-CNN, and SSD [17, 18]. An important part of these layers is bounding box regression model which provides precise object localization on image. However, training such layers under conditions of a limited size of training dataset and computing resources based on stochastic gradient descent is ineffective. One of the promising ways to implement of bounding box regression model is using the ELM which characterized by rapid training to obtain the least squares solution of regression problem [19]. In order to eliminate overfitting which occurs when the number of hidden layer nodes is large, the incremental learning via successively adds hidden nodes is actual to investigate. In this case, the task of fine-tuning of feature extractor can be solve by application of metaheuristic algorithms as alternative of gradient descent approach. Among them, it is worth

highlighting the simulated annealing algorithm which characterized by better convergence and less probability of getting stuck in the "bad" local optimum [20]. However, its use in the problems of fine tuning of convolutional filters remains insufficiently studied.

3 MATERIALS AND ALGORITHMS

To solve the problem of the development of the model for data analysis under conditions of limited volume of training set and computing resources of a compact aerial drone, it is essential to maximize the use of all available a priori information. Transfer learning technique is one of the examples of using a priori information, accumulated in the trained network for the future reuse [1, 2]. This technique allows the lower layers of the model to be borrowed from a pre-trained deep learning network and the top layers to be adapted to the particular domain requirements. However, when objects of small sizes are detected, the zone of interest contains little information and the context, which allows to eliminate uncertainty, becomes increasingly more important. Fig. 1 depicts a proposed architecture for the detection of objects of small sizes based on a combination of transfer learning and contextual information obtained by concatenating the feature maps of different artificial neural network layers. Upscaling is used to provide uniform shape of each channel of feature map. Concatenation with upscaling are considered as single upscale-concatenation layer.
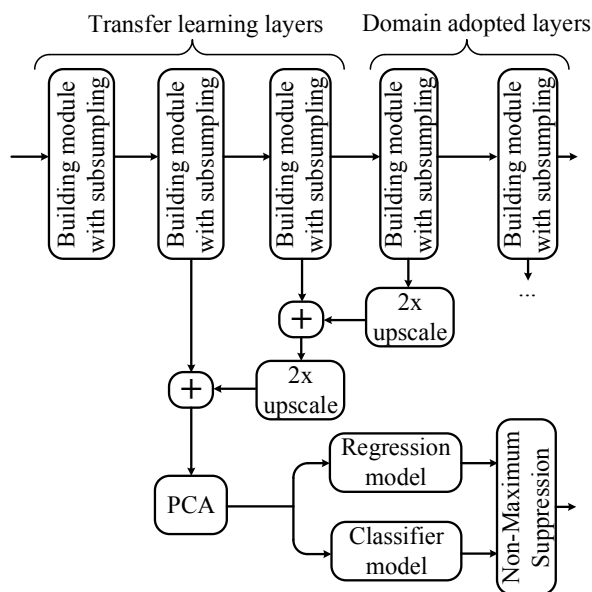

Figure 1 – Generalized architecture of the detector

Inception, Xception, VGG and Fire, among others, are some of the popular choices of modules to be used when constructing deep convolutional networks. These modules have different microarchitecture, which, in turn, implies different computational complexity and learning efficiency. We propose to adopt the lower layers from a pre-trained Squeezenet network, which consists of Fire modules and is characterized by the high computational efficiency. Upper layers of the network in this case can be built with the simple VGG modules which afford significant flexibility where different learning techniques are concerned,

At the first stage of training algorithm, we propose to use unsupervised pre-training of the high-level layers of the network to maximize utilization of the unlabeled domain training samples. In this case, to ensure the noise immunity and informativeness of feature representation, it is proposed to calculate the activation of each feature map pixel based on orthogonal matching pursuit algorithm [12].

It is proposed to carry out unsupervised training of the high-level layers of a neural network using a growing sparse coding neural gas, based on the principles of growing neural gas and sparse coding. In this case, the dataset for training high-level filters of the convolutional network is formed by partitioning input images or feature maps into the patches. These patches reshape to 1D vectors, arriving at the input of the algorithm of growing sparse coding neural gas, the basic stages of which are given below:

1) initialization of the counter of learning vectors $t := 0$;

2) two initial nodes (neurons) w_a and w_b are assigned by random choice out of the training dataset. Nodes w_a and w_b are connected by the edge, the age of which is zero. These nodes are considered non-fixed;

3) the following vector x, which is set to the unit length (L2– normalization), is selected;

4) set each basis vector $w, k = \overline{1, M}$ to unit length (L2-normalization);

5) calculation of the measure of similarity of input vector x to basis vectors $w_{s_k} \in W$ for sorting

$$-(w_{s_0}^T x)^2 \le ... \le -(w_{s_k}^T x)^2 \le ... \le -(w_{s_{M-1}}^T x)^2;$$

6) the nearest node w_{s0} and the second by proximity node w_{s1};

7) to increase by unity the age of all incidents w_{s0};

8) if w_{s0} is fixed, we proceed to step 9, otherwise – to step 10;

9) if $(w_{s0}^T x)^2 \ge v$, we proceed to step 12. Otherwise, we add new non-fixed neuron w_r to the point that coincides with input vector $w_r = x$, besides, a new edge that connects w_r and w_{s0} is added, then we proceed to step 13;

10) node w_{s0} and its topological neighbors (the nodes, connected with it by the edge) are displaced in the direction to input vector x according to the Oja's rule [14] by formulas

$$\Delta w_{s0} = \varepsilon_b \eta_t y_0 (x - y_0 w_{s0}), \; y_0 := w_{s0} x,$$

$$\Delta w_{s0} = \varepsilon_n \eta_t y_n (x - y_n w_{sn}), \; y_n := w_{sn}^T x,$$

$$0 < \varepsilon_n = 1, \quad 0 < \varepsilon_n = \varepsilon_b,$$

$$\eta_t := \eta_0 (\eta_{\text{final}} / \eta_0)^{t/t_{\max}},$$

11) if $(w_{s_0}^T x)^2 \geq v$, we label neuron w_{s0} as fixed;

12) if w_{s0} and w_{s1} are connected by the edge, its age is nulled, otherwise, a new edge with the zero age is formed between w_{s0} and w_{s1};

13) all edges in the graph with the age of more than $a_{\max}$ are removed. In the case when some nodes do not have any incident edges (become isolated), these nodes are also removed;

14) if $t < t_{\max}$, proceed to step 15, otherwise – increment of the step counter $t:=t+1$ and proceed to step 3;

15) if all neurons are fixed, the algorithm implementation stops, otherwise, proceed to step 3 and a new learning epoch begins (repetition of training dataset).

Concatenations of the feature maps of different layers of the artificial neural network leads to a high dimensionality problem of feature representation. To counter that, we propose to use one of the simplest techniques – Principal Component Analysis. This would allow removal of the features from the low levels of the network which are insensitive to the specific domain context.

In addition, it is proposed to carry out such classification analysis of the feature map in the framework of boosting and the so-called information-extreme technology. This makes it possible to synthesize a classifier with low computational complexity and relatively high accuracy under of limited training sets size constraint [11].

Boosted information-extreme classifier that evaluates belonging of j-th datapoint x_j (pixel of feature map) with N_1 features to one of the Z classes performs feature encoding using boosted trees and decision rules constructed in radial basis of binary Hamming space. In this case, there is the training set $D = \{x_j, y_j \mid j = \overline{1, n}\}$. The training of boosted information-extreme classifier is performed according to the following steps.

1. Initialize weight $w_j = 1/n$.

2. For $k = 1 \dots K$ do

3. Bootstrap D_k from D using probability distribution $P(X = x_j) = w_j$

4. Train decision tree T_k on D_k using entropy criterion to measure the quality of split.

5. Binary encoding of x_j datapoint from D using concatenation of decision paths from $T_1, \dots, T_k$. A datapoint x_j is classified in the leaf node by boosted trees. Each decision node receives a unique identifier. If a test is satisfied in a node, then the corresponding bit is asserted. Finally, the encodings for each tree are combined by concatenation (or more generally by hashing

the features ID onto a smaller dimensional space) [12]. Binary matrix $\{b_{z,s,i} \mid i = \overline{1, N_2}; s = \overline{1, n_z}; z = \overline{1, Z}\}$ is output of the step. Hence the equality $n = \sum\limits_z n_z$ is met.

6. Build information-extreme decision rules in radial basis of binary Hamming space and compute optimal information criterion :

$$E_z^* = \max_{\{d\}} E_z(d),$$

where $\{d\} = \{0, 1, \dots, \left(\sum\limits_i b_{z,i} \oplus b_{c,i} - 1\right)\}$ with center b_z of data distribution in class X_z^o, which computed using rule

$$b_{z,i} = \begin{cases} 1, & \text{if} \quad \dfrac{1}{n_z} \sum\limits_{s=1}^{n_z} b_{z,s,i} > \dfrac{1}{Z} \sum\limits_{c=1}^{Z} \dfrac{1}{n_c} \sum\limits_{s=1}^{n_c} b_{c,s,i}; \\ 0, & \text{otherwise}. \end{cases}$$

where E_z is computed as the normed modification of the S. Kullback's information measure [11]:

$$E_z = \frac{1 - (\alpha_z + \beta_z)}{\log_2(2 + \varsigma) - \log_2 \varsigma} \cdot \log_2 \left[\frac{2 - (\alpha_z + \beta_z) + \varsigma}{(\alpha_z + \beta_z) + \varsigma} \right]. \quad (4)$$

In order to increase learning efficiency, it is common to reduce the problem of the multi-class classification to the series of the two-class one by the principle "one-against-all". In this case, to avoid the problems of class imbalance, due to the majority of negative datapoints in datasets, a synthetic class is an alternative to the X_z^o class. The synthetic class is represented by n_z datapoints of the remaining classes, which are the closest to support vector b_z.

7. Test obtained information-extreme rules on dataset D and compute error rate for each sample from D. Under the inference mode, decision on belonging of datapoint b to one class from set $\{X_z^o \mid z = \overline{1, Z}\}$ is made according to maximum value of membership function $\mu_z(b)$ through the expression $\arg \max\limits_z \{\mu_z(b)\}$. In this case membership function $\mu_z(b)$ of binary representation b of input datapoint x to X_z^o class, the optimal container of which has support vector b_z^* and radius d_z^*, is derived from formula

$$\mu_z(b) = \exp\left(-\sum\limits_{i=1}^{N_2} b_i \oplus b_{z,i}^* / d_z^* \right).$$

8. Update $\{w_j\}$ proportional to errors on datapoint x_j.

9. If $|E_k^* - E_{k-1}^*| < \varepsilon$ abort loop.

Another important task in detecting precise object boundaries hampered by subsampling is bounding box prediction. We propose to implement bounding box prediction on the basis of a regression model based on a SHFN.

Consider n arbitrary distinct datapoints $\{(x_j, y_j) \mid x_j \in R^N, y_j \in R^M, 1 \le j \le n\}$. The SHFN with R additive hidden nodes and activation function $\varphi(x)$ can be represented by

$$\sum_{r=1}^{R} \beta_j \varphi(w_r^T x_j + b_r) = o_j, 1 \le j \le n.$$

The network with R hidden nodes can approximate these N samples with zero error when all the parameters are allowed to be adjusted freely, i.e., there exist β_r, w_r and b_r. The above n equations can be compactly rewritten as the matrix equation

$$H\beta = Y,$$

where

$$H = \begin{bmatrix} \varphi(w_1^T x_1 + b_1) & ... & \varphi(w_R^T x_1 + b_R) \\ ... & ... & ... \\ \varphi(w_1^T x_n + b_1) & ... & \varphi(w_R^T x_n + b_R) \end{bmatrix}_{n \times R},$$

$$\beta = \begin{bmatrix} \beta_1^M \\ ... \\ \beta_R^M \end{bmatrix}_{R \times M}, \quad Y = \begin{bmatrix} y_1^M \\ ... \\ y_n^M \end{bmatrix}_{n \times M}.$$

In order to solve such issues of network training as redundant hidden nodes and slow convergence rate, the orthogonal incremental ELM is proposed to use as regression model. It avoids redundant nodes and obtains the least squares solution of equation $H\beta = Y$ through incorporating the Gram-Schmidt orthogonalization method into well-known incremental extreme learning machine. Rigorous proofs in theory of convergence of orthogonal incremental extreme learning are given by Li Ying [19]. The training of orthogonal incremental ELM is performed according to the following steps.

1. Set maximum number of iterations L_{max} and expected learning accuracy E_0.

2. For $L = 1 \dots$ do

3. Increase by one the number of hidden nodes: $r = r + 1$.

4. Randomly generate one hidden node and calculate its output vector h_r.

5. If $r = 1$ then $v_r = h_r$ else

$$v_r = h_r - \frac{\langle v_1, h_r \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle v_2, h_r \rangle}{\langle v_2, v_2 \rangle} v_2 - ... - \frac{\langle v_{r-1}, h_r \rangle}{\langle v_{r-1}, v_{r-1} \rangle} v_{r-1}.$$

6. If $\| v_r \| \ge \varepsilon$ calculate the output weight for the new hidden node $\beta_r = v_r^T E / (v_r^T v_r)$ and calculate the new residual error $E = E - v_r \beta_r$ else $r = r - 1$.

7. If $\| E \| \ge E_0$ abort loop.

For training classifier and regression model or fine-tuning of feature extractor we collect training set using matching strategy to determine which default bounding boxes correspond to ground truth bounding boxes. Default bounding boxes are defined as feature map pixels reprojected on input image. We consider condition of aerial survey with down oriented camera and high altitude (higher than 100m) therefore multiple anchor boxes associated with feature map pixel are not used. In this case each default box is associated with feature map pixel reprojected on input image. Each ground truth box is matched to the default box with the best Jaccard overlap and default boxes is matched to any ground truth box with IoU higher than 0.4. Regression model is trained only on positive samples (matched default boxes).

The complex criterion J (2) of learning efficiency of object detector should be takes into account both effectiveness of classification analysis J_{Cls} and bounding box prediction J_{Loc}. It is proposed to calculate the criterion of object classification effectiveness from formula

$$J_{Cls} = \frac{1}{Z} \sum_{z=1}^{Z} E_z.$$

Effectiveness of bounding box prediction is proposed to calculate from formula

$$J_{Loc} = \frac{1}{\hat{n}} \sum_{i=1}^{\hat{n}} IoU_i.$$

At last stage of training algorithm, it is necessary to fine tune of high-level layers of feature extractor after unsupervised learning in order to take into account significant imbalance between objects of interest and background patches. It is proposed to use simulated annealing as metaheuristic search optimization algorithms. The efficiency of the simulated annealing algorithm depends on the implementation of the *create_neighbor_solution* procedure, forming a new solution s_i on the i-th iteration of the algorithm. Fig. 2 shows a pseudocode of the simulated annealing algorithm, which is implemented by the *epochs_max* iterations, on each of which function f() is calculated by passing a labeled training dataset through the model of the system of detection and calculation of a complex criterion (2) [3, 20].

$s_{current} \leftarrow create\_initial\_solution()$
$s_{best} \leftarrow s_{current}$
$T \leftarrow T_0$
$c \leftarrow \varepsilon, \ 0 < \varepsilon < 1$
$for(i = 1 \ to \ epochs\_max)$
$\quad s_i \leftarrow create\_neighbor\_solution(s_{current})$
$\quad if \ f(s_i) \geq f(s_{current})$
$\qquad s_{current} \leftarrow s_i$
$\qquad if \ f(s_i) \geq f(s_{best})$
$\qquad\quad s_{best} \leftarrow s_i$
$\qquad end \ if$
$\quad elseif \ \exp\left(\dfrac{f(s_{current}) - f(s_i)}{T}\right) > uniform\_random(0,1)$
$\qquad s_{current} \leftarrow s_i$
$\quad end \ if$
$\quad T \leftarrow c \times T$
$end \ for$
$return(s_{best})$

Figure 2 – Pseudocode of simulated
annealing algorithm

An analysis of the pseudocode in Fig. 2 shows that current solution $s_{current}$, in relation to which new best solutions s_{best} are sought for, is updated in case of providing a new solution of the criterion increase (2), or randomly from the Gibbs distribution. In this case, an initial search point that is formed by the *create_initial_solution* procedure can be either randomly generated or a result of the preliminarily training by another algorithm. To generate new solutions in the *create_neighbor_solution* procedure, it is proposed to use the simplest non-adaptive algorithm, which can be represented as formula [3]:

$$s_{current} = s_{current} + uniform\_random(-1,1) \cdot step\_size.$$

The non-maximum suppression algorithm is used for filtering unnecessary actions of the detector to one and the same image object [17, 18].

Thus, we propose a model and object detector training algorithm based on the fusion of different techniques with the aim of maximizing obtained domain context information under the small labelled training set and limited computational resources constraints.

4 EXPERIMENTS

To train the object detector, 200 images from dataset of Inria Aerial Image Labeling Dataset were used [21]. Each image has the resolution of 5,000×5,000 pixels. 5,00 unlabeled 224×224 pixel images were generated through random crop with rotation for unsupervised learning. Also 200 labeled images 224×224 pixel images were generated for supervised learning. Labeled training set was augmented to create 1000 instances by adding noise, a contrast change, rotation and cropping.

A large number of vehicles in the urban area were presented in the Inria Aerial Image Labeling Dataset. Vehicles were selected as the objects of interest and the urban area was considered a usage domain. In this case, the set of classes contained Z=3 classes, where the first class corresponded to cars, the second class corresponded

to trucks, and the third one – to the background. The size of objects in pixels in random images varied in the range of [7×7, …, 10×10].

In accordance with transfer learning technique, the first 7 Fire modules of pretrained convolutional neural network Squeezenet were adopted. As a result, each input image was encoded into the feature map with 13×13×384 pixel dimensions.

The subsequent layer is trained unsupervised on the unlabeled datasets from the usage domain area consisting of filters with kernels of 3×3 with stride=1. Output feature map is formed by concatenation of feature maps from Fire 6, Fire 7 modules and last convolutional layer.

It is proposed first to train the detector using the unsupervised pretrained last layer of feature extractor via growing sparse coding neural gas without fine tuning. In this case, a fixed value of the training dataset reconstruction v=0.8 is used during training. In addition, in the information-extreme classifier of the feature map pixels, the number nodes in decision trees is constrained to 16. The depth of each tree is set to 6.

In order to improve the results of machine learning of the detector, informativeness of feature description is increased by fine-tuning of the unsupervised trained convolutional layers. In this case, the following parameters of simulated annealing algorithm were used: c=0.98, T_0 =10, *epochs*_max=6000, *step_size*=0.001. Each fine-tuning step is involved a re-training of a regressor and classifier. To maximize the model's generalizing ability and minimize computational complexity, we implemented sequential tuning of the hyperparameter v responsible for the density of neuron distribution with a step of 0.1.

Prior unsupervised learning of the upper convolutional layers on unlabelled samples from the intended usage domain is aimed at increasing the subsequent supervised machine learning efficiency. It is worthwhile considering the influence the parameters of the growing sparse coding neural gas algorithm used in unsupervised learning have on the results of supervised learning. Table 1 presents the machine learning results and quantity N_c of generated convolutional filters (neurons) as a function of the parameter v, which characterises the accuracy of coverage of the training set by the convolutional filters.

Table 1 – Detector learning results at various values of the unsupervised learning training set coverage parameter

v	N_c	J_{Cls}	J_{Loc}	J	Percentage of discovered objects in the test set
0.4	196	0.0912	0.450	0.04104	0.67
0.5	337	0.1502	0.511	0.07675	0.73
0.6	589	0.2508	0.621	0.15574	0.87
0.7	889	0.4203	0.700	0.29421	0.93
0.8	1519	1.0000	0.921	0.92100	0.96
0.9	3058	1.0000	0.923	0.92300	0.95

Analysis of the Table 1 demonstrates that quantity of neurons and values of partial and general optimisation criteria increases with the growth in hyperparameter v (2). However, at $v \leq 0.8$ model accuracy as per the test set increases with the increase in parameter v, however a further increase in this parameter leads to decrease in results quality due to overfitting. At the same time, selection of the quantity of the principal components for each value of v was made in accordance with the Kaiser criterion : selecting only the principal components with eigenvalues exceeding 1.

Fig. 3 provides the graphical representations of the dependency of learning efficiency information criterion (4) on container radii of each class. These can be used to evaluate the accuracy and noise immunity of the synthesized classification decision rules.

Analysis of Fig. 3. shows that all classes X_1^o have error-free radial-basis decision rules by the training set. Optimal container radii in this case are equal to $d_1^* = 30$, $d_2^* = 50$. Corresponding distances between the container centres are equal to $d_{12}^* = 95$.

We have thus been able to formulate highly accurate decision rules on a training set of a limited size, ensuring that datapoint distribution of each class are relatively compact in the feature space with substantial clearance between classes.
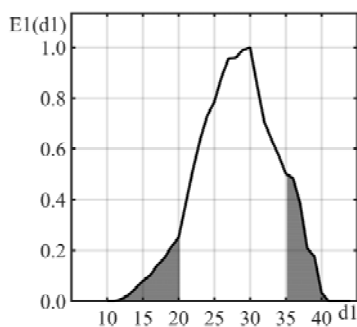
a
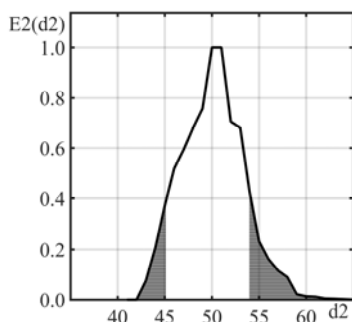
b

Figure 3 – Dependency of classifier learning efficiency information criterion on class container radius:

a – class X_1^o ; b – class X_2^o

6 DISCUSSION

To understand the advantages of prior unsupervised learning, let's consider the results of simulated annealing machine learning algorithm before and after unsupervised training with growing sparse coding neural gas algorithm.

Fig 4. shows the dependency of the simulated annealing algorithm learning information criteria (2) on the number of training epochs with parameters c=0.998, T_0=10, $epochs$_max=5000, $step\_size$=0.001.

Analysis of Fig. 4 shows that prior unsupervised learning on the basis of sparsely coded growing neural gas algorithm allows to improve the final outcome of the supervised learning with simulated annealing algorithm. The use of prior unsupervised learning allows reaching the global maximum of criterion (2) more than 10 times faster. Apart from that, the resulting model test on the training set indicates that the use of prior unsupervised learning allows to reduce the overfitting effects when the labelled training set is limited.

The information criterion (2) calculated for the training set was equal to J_{train}=0.921 which provides of 96% correctly detection on test dataset where prior unsupervised learning was applied. At the same time, there was a substantial difference between the criteria J_{train}=0.3011 and 85% correctly detection on test dataset where unsupervised learning was not applied.

Thus the proposed algorithm of the prior unsupervised training of the upper layers allows to increase the learning efficiency criteria and percentage of the objects detected in the test images. In addition, the use of such algorithm allows to reduce the overfitting effect and increase the speed of convergence on the global maximum in the subsequent supervised learning process when the labelled training set size is limited. The relationship between the learning efficiency criterion and the parameters of the simulated annealing algorithm was not considered in the present study, however. Thus, further research will be focused on improving the detector model and development of the algorithms parameter tuning on the process of machine learning.
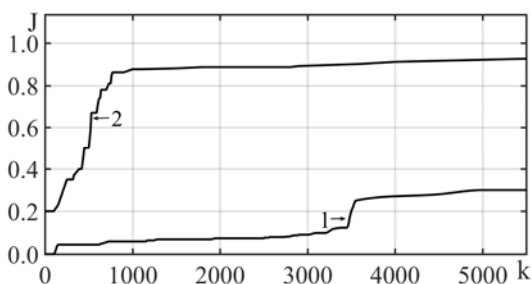

Figure 4 – Dependency of the optimization criteria (2) on the number of learning epochs:
1 – before prior unsupervised learning; 2 – after prior unupervised learning

CONCLUSIONS

The scientific novelty of the work lies in the new model and training algorithm of small-sized object detection under computational resource constraint and limited volume of the labeled training set for a compact drone are developed.

The model contains seven first modules of the computationally effective convolutional Squeezenet network, one convolutional sparse coding layer, upscaling-concatenation layer, PCA-transformer, the information-extreme classifier and extreme learning regression model. This approach provides more contextual information for efficient recognition of small objects on aerial images.

The training algorithm consist of multiple stages : transfer learning, unsupervised learning of additional high-level convolutional sparse coding layers, PCA-transformation of output feature map, information-extreme learning of pixel classifier for object detection, orthogonal incremental extreme learning of regression model for bounding box prediction and fine tuning of the upper layers by using the metaheuristic simulated annealing algorithm. It was shown that the use of prior unsupervised learning makes it possible to decrease the overfitting effect and to increase by more than 10 times the rate of finding the global maximum at supervised training on dataset of limited size.

The practical significance of the achieved outcomes is that the requirements for size of labeled training set and computation resources consumption are reduced. The proposed model provides acceptable for practical use reliability of detection of the objects, the size of which in pixels is by 7...14 times smaller than the size of the smallest side of the aerial image. The proposed training algorithm utilize 500 unlabeled and 200 labeled training samples to provide 96% correctly detection of objects on the images of the test Inria Aerial Image Labeling dataset.

ACKNOWLEDGMENTS

REFERENCES

1. Patricia N., Caputo B. Learning to Learn, from Transfer Learning to Domain Adaptation: A Unifying Perspective, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, Ohio, 23–28 June, 2014 : proceedings. Conference Publishing Services,* 2014, P. 1442–1449. DOI: 10.1109/CVPR.2014.187.
2. Nguyen A., Yosinski J., Clune J. Deep neural networks are easily fooled: High confidence predictions for unrecognizable images, *IEEE Conference on Computer Vision and Pattern Recognition, (CVPR), Boston, MA, 7–12 June, 2015 : proceedings, Conference Publishing Services,* 2015, P. 427–436. DOI: 10.1109/CVPR.2015.7298640.
3. Ayumi V., Rere L. M. R., Fanany M. I., Arymurthy A. M. Optimization of convolutional neural network using microcanonical annealing algorithm / V. Ayumi, // *International Conference on Advanced Computer Science and Information Systems (ICACSIS), Malang, Indonesia, 15–16 Oct., 2016: proceedings. Institute of Electrical and Electronics Engineers,* 2016, pp. 506–511. DOI: 10.1109/ICACSIS.2016.7872787.
4. Antipov G. Berrani S., Ruchaud N., Dugelay J. Learned vs. Hand-Crafted Features for Pedestrian Gender Recognition, *23rd ACM International Conference on Multimedia, Brisbane, Australia, 26–30 Oct., 2015: proceedings*, ACM. New York, NY, USA, 2015, pp. 1263–1266. DOI: 10.1145/2733373.2806332.
5. Carrio A. Sampedro C., Rodriguez-Ramos A., Campoy P. A Review of Deep Learning Methods and Applications for Unmanned Aerial Vehicles, *Journal of Sensors,* 2017, Vol. 2017. – P. 1–13. DOI:10.1155/2017/3296874.
6. Subbotin S. The special deep neural network for stationary signal spectra classification, *14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET).* Lviv-Slavske, Ukraine, 20–24 Feb, 2018, proceedings, IEEE, 2018, pp.123–128.
7. Xu X., Ding Y., Hu S. X. Scaling for edge inference of deep neural networks, *Nature Electronics,* 2018, Vol. 1(4), pp. 216–222. DOI:10.1038/s41928-018-0059-3.
8. Loquercio A. Maqueda A. I., del-Blanco C. R., Scaramuzza D. DroNet: Learning to Fly by Driving, *IEEE Robotics and Automation Letters*, 2018, Vol. 3, No. 2, pp. 1088–1095. DOI: 10.1145/2733373.2806332.
9. Mathew A., Mathew J., Govind M., Mooppan A. An Improved Transfer learning Approach for Intrusion Detection, *Procedia Computer Science,* 2017, Vol. 115, pp. 251–257. DOI: 10.1016/j.procs.2017.09.132.
10. Radenović F., Tolias G., Chum O. Fine-tuning CNN Image Retrieval with No Human Annotation, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 2018, Mode of access: https://ieeexplore.ieee.org/document/8382272. DOI: 10.1109/TPAMI.2018.2846566.
11. Moskalenko V., Dovbysh S. , Naumenko I., Moskalenko A., Korobov A. Improving the effectiveness of training the on-board object detection system for a compact unmanned aerial vehicle, *Eastern-European Journal of Enterprise Technologies,* 2018, No. 4/9 (94), pp. 19–26. DOI: 10.15587/1729-4061.2018.139923
12. Vens C., Costa F. Random Forest Based Feature Induction, *IEEE 11th International Conference on Data Mining, Vancouver, Canada, 11–14 Dec, 2011: proceedings*, pp. 744–753. DOI: 10.1109/ICDM.2011.121
13. Feng Q., Chen C. L. P., Chen L. Compressed auto-encoder building block for deep learning network, *3rd International Conference on Informative and Cybernetics for Computational Social Systems (ICCSS), Jinzhou, Liaoning, China, 26–29 Aug, 2016: proceedings, IEEE,* 2016, pp. 131–136. DOI: 10.1109/ICCSS.2016.7586437
14. Labusch K., Barth E., Martinetz T. Sparse coding neural gas: learning of overcomplete data representations, *Neurocomputing,* 2009, Vol. 72, I. 7–9, pp. 1547–1555. DOI:10.1016/j.neucom.2008.11.027.
15. Mrazova I., Kukacka M. Image Classification with Growing Neural Networks, *International Journal of Computer Theory and Engineering,* 2013, Vol. 5, No. 3, pp. 422–427. DOI:10.7763/IJCTE.2013.V5.722.

16. Palomo J., López-Rubio E. The Growing Hierarchical Neural Gas Self-Organizing Neural Network, *IEEE Transactions on Neural Networks and Learning Systems,* 2017, Vol. 28, No. 9, pp. 2000–2009. DOI:10.1109/TNNLS.2016.2570124.

17. Nakahara H., Yonekawa H., Sato S. An object detector based on multiscale sliding window search using a fully pipelined binarized CNN on an FPGA, *International Conference on Field Programmable Technology (ICFPT), Melbourne, VIC, 11–13 Dec, 2018: proceedings, IEEE,* 2017, pp. 168–175. DOI: 10.1109/FPT.2017.8280135

18. Chen X., Xiang S., Liu C.-L., Pan C.-H. Aircraft Detection by Deep Convolutional Neural Networks, *IPSJ Transactions on Computer Vision and Applications,* 2015, Vol. 7, pp. 10–17. DOI: 10.2197/ipsjtcva.7.10.

19. Zou W., Xia Y., Li H. Fault Diagnosis of Tennessee-Eastman Process Using Orthogonal Incremental Extreme Learning Machine Based on Driving Amount / W. Zou, // *IEEE Transactions on Cybernetics,* 2018, pp. 1–8. DOI: 10.1109/TCYB.2018.2830338.

20. Rere R. L. M., Fanany M. I., Arymurthy A. M. Metaheuristic Algorithms for Convolution Neural Network, *Computational Intelligence and Neuroscience,* 2017, Vol. 2016, pp. 1–13. DOI:10.1155/2016/1537325.

21. Maggiori E., Tarabalka Y., Charpiat G., Alliez P. High-Resolution Aerial Image Labeling With Convolutional Neural Networks, *IEEE Transactions on Geoscience and Remote Sensing,* 2017, Vol. 55, No. 12, P. 7092–7103. DOI:10.1109/TGRS.2017.2740362.

УДК 004.891.032.26:629.7.01.066

МОДЕЛЬ І АЛГОРИТМ НАВЧАННЯ СИСТЕМИ ДЕТЕКТУВАННЯ МАЛОРОЗМІРНИХ ОБ'ЄКТІВ ДЛЯ МАЛОГАБАРИТНОГО ЛІТАЮЧОГО АПАРАТУ

Москаленко В. В. – канд. техн. наук, доцент кафедри комп'ютерних наук, Сумський державний університет, Суми, Україна.

Москаленко А. С. – канд. техн. наук, асистент кафедри комп'ютерних наук, Сумський державний університет, Суми, Україна.

Коробов А. Г. – аспірант кафедри комп'ютерних наук, Сумський державний університет, Суми, Україна.

Зарецький М. О. – аспірант кафедри комп'ютерних наук, Сумський державний університет, Суми, Україна.

АНОТАЦІЯ

Актуальність. Розроблено обчислювально просту модель і ефективний алгоритм навчання бортової системи детектування об'єктів на місцевості. Об'єкт дослідження – процес детектування малорозмірних об'єктів на аерофотознімках в умовах обмежених обчислювальних ресурсів і невизначеності, зумовленої малим об'ємом розміченої навчальної вибірки. Предмет дослідження – модель і метод навчання моделі для детектування малорозмірних об'єктів на аерофотознімках.

Мета дослідження – розробка ефективних моделі і алгоритм навчання системи детектування об'єктів малогабаритним літаючим апаратом в умовах обмежених обчислювальних ресурсів і обмеженого обсягу розміченої вибірки.

Методи дослідження. В роботі запропоновано чотирьох етапний алгоритм навчання детектора об'єктів. На першому етапі здійснюється вибір згорткової нейромережі, попередньо навченої на наборі ImageNet, а також вибір кількості шарів, які будуть запозичені. Другий етап передбачає додавання згорткового розріджено-кодуючого шару, що навчається без вчителя, для формування карти ознак. Для реалізації другого етапу розроблено модифікацію алгоритму зростаючого нейронного газу, яка забезпечує автоматичне визначення необхідної кількості нейронів і їх оптимальний розподіл за даними. Це дозволяє ефективно використовувати нерозмічені набори даних для адаптації високорівневого ознакового опису до доменної області використання. На третьому етапі проводиться редукція вихідної карти ознак шляхом застосування методу головних компонент й навчання вирішальних правил. При цьому вихідна карта ознак будується шляхом конкатенації карт ознак нижніх рівнів з масштабованими картами ознак верхнього рівня. Даний підхід до побудови карти ознак покликаний забезпечити підвищення ефективності розпізнавання малорозмірних об'єктів за рахунок збільшення кількості контекстної інформації в кожному пікселі. Побудова моделі класифікаційного аналізу пікселів вихідної карти ознак пропонується здійснювати на основі принципів бустінгу і інформаційно-екстремального машинного навчання, а прогнозування меж детектованих об'єктів пропонується реалізувати на основі машини екстремального навчання. Останній етап алгоритму навчання полягає в тонкій настройці високорівневих шарів екстрактора ознак на основі метаевристичного алгоритму симуляції відпалу з метою максимального наближення до глобального оптимуму значення комплексного критерію ефективності навчання детектора.

Результати. Результати, отримані на відкритому наборі даних, підтверджують придатність запропонованих моделі та алгоритму навчання до практичного застосування. Запропонований метод навчання забезпечує 96% точності детектування малорозмірних об'єктів на тестових зображення при використанні 500 нерозмічених і 200 розмічених навчальних зображень.

Висновки. Наукова новизна статті полягає в розробці нових моделі і методу навчання детектора об'єктів, що забезпечують високодостовірне розпізнавання малорозмірних об'єктів на аерофотознімках в умовах обмежених обчислювальних ресурсів і малого обсягу розмічених навчальних вибірок. Практичне значення одержаних результатів роботи полягає в розробці моделі та алгоритму навчання детектора об'єктів, що дозволяють знизити вимоги до обсягу навчальної вибірки і обчислювальних ресурсів бортовий системи літаючого апарата в режимах навчання та екзамену.

КЛЮЧОВІ СЛОВА: зростаючий нейронний газ, згорткова мережа, бустінг, детектор об'єктів, інформаційний критерій, симуляція відпалу, екстремальне навчання.

УДК 004.891.032.26:629.7.01.066

МОДЕЛЬ И АЛГОРИТМ ОБУЧЕНИЯ СИСТЕМЫ ДЕТЕКТИРОВАНИЯ МАЛОРАЗМЕРНЫХ ОБЬЕКТОВ ДЛЯ МАЛОГАБАРИТНОГО ЛЕТАЮЩЕГО АППАРАТА

Москаленко В. В. – канд. техн. наук, доцент кафедры компьютерных наук, Сумской государственный университет, Сумы, Украина.

Москаленко А. С. – канд. техн. наук, ассистент кафедры компьютерных наук, Сумской государственный университет, Сумы, Украина.

Коробов А. Г. – аспирант кафедры компьютерных наук, Сумской государственный университет, Сумы, Украина.

Зарецкий Н. А. – аспирант кафедры компьютерных наук, Сумской государственный университет, Сумы, Украина.

АННОТАЦИЯ

Актуальность. Разработано вычислительно простую модель и эффективный алгоритм обучения бортовой системы детектирования объектов на местности. Объект исследования – процесс детектирования малоразмерных объектов на аэроснимках в условиях ограниченных вычислительных ресурсов и неопределенности, обусловленной малым объемом размеченной обучающей выборки. Предмет исследования – модель и алгоритм обучения модели для детектирования малоразмерных объектов на аэроснимках.

Цель исследования – разработка эффективных модели и метода обучения системы детектирования объектов малогабаритным летающим апаратом в условиях ограниченных вычислительных ресурсов и ограниченного объема размеченной выборки.

Методы исследования. В работе предложен четырех этапный алгоритм обучения детектора объектов. На первом этапе производится выбор сверточной нейросети, предобученной на наборе ImageNet, а также выбор количества слоев, которые будут позаимствованы. Второй этап предусматривает добавление сверточного разреженно-кодирующего слоя, обучаемого без учителя, для формирования карты признаков. Для реализации второго этапа разработано модификацию алгоритма растущего нейронного газа, которая обеспечивает автоматическое определение необходимого количества нейронов и их оптимальное распределение по данным. Это позволяет эффективно использовать неразмеченные наборы данных для адаптации высокоуровневого признакового описания к доменной области использования. На третьем этапе производится редукция выходной карты признаков путем применения метода главных компоненты и обучение решающих правил. При этом выходная карта признаков строится путем конкатенации карт признаков нижних уровней с масштабированными картами признаков верхнего уровня. Данный подход к построению карты признаков призван обеспечить повышение эффективности распознавания малоразмерных объектов за счет увеличения количества контекстной информации в каждом пикселе. Построение модели классификацинного анализа пикселей выходной карты признаков предлагается производить на основе принципов бустинга и информационно-экстремального машинного обучения, а прогнозирование границ детектированных объектов предлагается реализовать на основе машины экстремального обучения. Последний этап алгоритма обучения состоит в тонкой настройке высокоуровневых слоев экстрактора признаков на основе метаэвристического алгоритма симуляции отжига с целью максимального приближения к глобальному оптимуму значение комплексного критерия эффективности обучения детектора.

Результаты. Результаты, полученные на открытом наборе данных, подтверждают пригодность предложенных модели и алгоритма обучения к практическому применению. Предложенный алгоритм обучения обеспечивает 96% точности детектирования малоразмерных объектов на тестовых изображениях при использовании 500 неразмеченных и 200 размеченных обучающих изображений.

Выводы. Научная новизна статьи состоит в разработке новых модели и алгоритма обучения детектора обьектов, обеспечивающих высокодостоверное распознавание малоразмерных объектов на аэроснимках в условиях ограниченных вычислительных ресурсов и малого объема размеченных обучающих выборок. Практическая значимость результатов работы состоит в разработке модели и метода обучения детектора обьектов, позволяющих снизить требования к объему обучающей выборки и вычислительным ресурсам бортовой системы летающего аппарата в режимах обучения и экзамена.

КЛЮЧЕВЫЕ СЛОВА: растущий нейронный газ, сверточная сеть, бустинг, детектор объектов, информационный критерий, симуляция отжига, экстремальное обучение.

ЛІТЕРАТУРА / ЛИТЕРАТУРА

1. Patricia N. Learning to Learn, from Transfer Learning to Domain Adaptation: A Unifying Perspective / N. Patricia, B. Caputo // IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, Ohio, 23–28 June, 2014 : proceedings. – Conference Publishing Services, 2014. – P. 1442–1449. DOI: 10.1109/CVPR.2014.187.
2. Nguyen A. Deep neural networks are easily fooled: High confidence predictions for unrecognizable images / A. Nguyen, J. Yosinski, J. Clune // IEEE Conference on Computer Vision and Pattern Recognition, (CVPR), Boston, MA, 7–12 June, 2015 : proceedings. – Conference Publishing Services, 2015 – P. 427–436. DOI: 10.1109/CVPR.2015.7298640.
3. Optimization of convolutional neural network using microcanonical annealing algorithm / [V. Ayumi, L. M. R. Rere, M. I. Fanany, A. M. Arymurthy] // International Conference on Advanced Computer Science and Information Systems (ICACSIS), Malang, Indonesia, 15–16 Oct., 2016: proceedings. – Institute of Electrical and Electronics Engineers, 2016. – P. 506–511. DOI: 10.1109/ICACSIS.2016.7872787.
4. Antipov G. Learned vs. Hand-Crafted Features for Pedestrian Gender Recognition / G. Antipov, S. Berrani, N. Ruchaud, J. Dugelay // 23rd ACM International Conference on Multimedia, Brisbane, Australia, 26–30 Oct., 2015: proceedings. – ACM : New York, NY, USA, 2015. – P. 1263–1266. DOI: 10.1145/2733373.2806332.
5. A Review of Deep Learning Methods and Applications for Unmanned Aerial Vehicles / [A. Carrio, C. Sampedro, A. Rodriguez-Ramos, P. Campoy] // Journal of Sensors. – 2017. – Vol. 2017. – P. 1–13. DOI:10.1155/2017/3296874.

6. Subbotin S. The special deep neural network for stationary signal spectra classification / S. Subbotin // 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET), Lviv-Slavske, Ukraine, 20–24 Feb, 2018: proceedings.– IEEE, 2018. – P.123–128.

7. Xu X. Scaling for edge inference of deep neural networks / X. Xu, Y. Ding, S. X. Hu // Nature Electronics. – 2018. – Vol. 1(4). – P. 216–222. DOI:10.1038/s41928-018-0059-3.

8. Loquercio A. DroNet: Learning to Fly by Driving / A. Loquercio, A. I. Maqueda, C. R. del-Blanco, D. Scaramuzza // IEEE Robotics and Automation Letters, 2018. – Vol. 3, № 2. – P. 1088–1095. DOI: 10.1145/2733373.2806332.

9. An Improved Transfer learning Approach for Intrusion Detection / [A. Mathew, J. Mathew, M. Govind, A. Mooppan] // Procedia Computer Science. – 2017. – Vol. 115. – P. 251–257. DOI: 10.1016/j.procs.2017.09.132.

10. Radenović F. Fine-tuning CNN Image Retrieval with No Human Annotation / F. Radenović, G. Tolias, O. Chum // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2018. – Mode of access: https://ieeexplore.ieee.org/document/8382272. DOI: 10.1109/TPAMI.2018.2846566.

11. Improving the effectiveness of training the on-board object detection system for a compact unmanned aerial vehicle / [V. Moskalenko, S. Dovbysh, I. Naumenko, et al.] // Eastern-European Journal of Enterprise Technologies. – 2018. – № 4/9 (94). – P. 19–26. DOI: 10.15587/1729-4061.2018.139923

12. Vens C. Random Forest Based Feature Induction / C. Vens, F. Costa // IEEE 11th International Conference on Data Mining, Vancouver, Canada, 11–14 Dec, 2011: proceedings. – P. 744–753. DOI: 10.1109/ICDM.2011.121

13. Feng Q. Compressed auto-encoder building block for deep learning network / Q. Feng, C. L. P. Chen, L. Chen // 3rd International Conference on Informative and Cybernetics for Computational Social Systems (ICCSS), Jinzhou, Liaoning, China, 26–29 Aug, 2016: proceedings. – IEEE, 2016. – P. 131–136. DOI: 10.1109/ICCSS.2016.7586437

14. Labusch K. Sparse coding neural gas: learning of overcomplete data representations / K. Labusch, E. Barth, T. Martinetz // Neurocomputing. – 2009/ – Vol. 72, I. 7–9. – P. 1547–1555. DOI:10.1016/j.neucom.2008.11.027.

15. Mrazova I. Image Classification with Growing Neural Networks / Mrazova I., Kukacka M. // International Journal of Computer Theory and Engineering. – 2013. – Vol. 5, №. 3. – P. 422–427. DOI:10.7763/IJCTE.2013.V5.722.

16. Palomo J. The Growing Hierarchical Neural Gas Self-Organizing Neural Network / J. Palomo, E. López-Rubio // IEEE Transactions on Neural Networks and Learning Systems. – 2017. – Vol. 28, №. 9. – P. 2000–2009. DOI:10.1109/TNNLS.2016.2570124.

17. Nakahara H. An object detector based on multiscale sliding window search using a fully pipelined binarized CNN on an FPGA / H. Nakahara, H. Yonekawa, S. Sato // International Conference on Field Programmable Technology (ICFPT), Melbourne, VIC, 11–13 Dec, 2018: proceedings. – IEEE, 2017. – P. 168–175. DOI: 10.1109/FPT.2017.8280135

18. Aircraft Detection by Deep Convolutional Neural Networks / [X. Chen, S. Xiang, C.-L. Liu, C.-H. Pan] // IPSJ Transactions on Computer Vision and Applications. – 2015. – Vol. 7. – P. 10–17. DOI: 10.2197/ipsjtcva.7.10.

19. Zou W. Fault Diagnosis of Tennessee-Eastman Process Using Orthogonal Incremental Extreme Learning Machine Based on Driving Amount / W. Zou, Y. Xia, H. Li // IEEE Transactions on Cybernetics. – 2018. – P. 1–8. DOI: 10.1109/TCYB.2018.2830338.

20. Rere R. L. M. Metaheuristic Algorithms for Convolution Neural Network / R. L. M. Rere, M. I. Fanany, A. M. Arymurthy // Computational Intelligence and Neuroscience. – 2017. – Vol. 2016. – P. 1–13. DOI:10.1155/2016/1537325.

21. High-Resolution Aerial Image Labeling With Convolutional Neural Networks / [E. Maggiori, Y. Tarabalka, G. Charpiat, P. Alliez] // IEEE Transactions on Geoscience and Remote Sensing. – 2017. – Vol. 55, № 12. – P. 7092–7103. DOI:10.1109/TGRS.2017.2740362.