International Journal on
"Technical and Physical Problems of Engineering"
(IJTPE)
Published by International Organization of IOTPE

ISSN 2077-3528
IJTPE Journal
www.iotpe.com
ijtpe@iotpe.com

December 2023	Issue 57	Volume 15	Number 4	Pages 434-442

CREATING A METHOD FOR IDENTIFYING OBJECTS IN IMAGE WITH THEIR DIFFERENT ORIENTATIONS

V.Y. Terekhov [1] **E.Y. Linskaya** [1] **A.K. Lampezhev** [2]

1. Science and Technology Park of Biomedicine, I.M. Sechenov First Moscow State Medical University, Ministry of Health, Moscow, Russia, terekhovaleryy@bk.ru, linskaya.yelena@bk.ru
2. Institute of Design and Technology Informatics, Russian Academy of Sciences, Moscow, Russia abas.lampezhev@mail.ru

Abstract- This article describes a set of research studies conducted for evaluating new opportunities aimed at improving methods for identifying images of any three-dimensional object which will ensure its guaranteed recognition, considering various conditions of observations against the background of noise blur and occlusion. The research presents a number of significant solutions based on a completely new mechanism of hierarchical simulation modeling, which includes information about some spatial indicators inherent in the depicted three-dimensional object. This simulation makes it possible to form representations of the object being assessed and a natural image. It enables to achieve an increase in efficiency in terms of accuracy, recall, and the proportion of correctly recognized images, by providing stable recognition under conditions of various viewing angles, and various types of visual noise. Experimental verification showed its advantage over alternative solutions. Thus, for example, the recognition accuracy when using the proposed method (enabling to achieve the required resistance to occlusions) increases by 7-10%, and the recognition accuracy when changing the orientation of the object in the image experiences deviations of no more than 4-5%. The peculiarities of the presented approach can be used for ensuring the solution of highly specific problems and identifying various kinds of patterns in video surveillance and self-learning of the detected structures.

Keywords: Computer Vision, Efficiency of Computer Vision Systems, Image Recognition, Simulation Modeling, Video Surveillance.

1. INTRODUCTION

Visual image recognition is the most important component of the complex that performs the information data control and processing by an automated system, and a system that makes a decision [1]. Numerous tasks concerning the classification and identification of an object or a pulse characterized by a complete list of qualities arise in the field of searching and evaluating visual information, and studying issues related to artificial intelligence [2]. Algorithmic identification and classification of each image, which is used in most systems, ensure security and access control, and a search system in video surveillance [3]. Currently, production is actively using complexes that recognize text, or faces [4]. This is used in the interface of almost every software product, a system that provides security, identification, and in a combination of other practical areas.

Some practical studies in this area are characterized by a long history [5, 6]. Thus, over the past decade, there have been tangible shifts in the field of visual image recognition due to the development of methods that reduce the dimensionality [7], application of convolutional neural networks [8], and creation of a constellation model [9]. At the same time, even considering the successes achieved, the results of each advanced experiment allow us to conclude that the methods for recognizing an object in an image do not yet have full-fledged detection tools, for example, the probability of functioning on wide sets of recognition groups does not have the proper stability in the field of non-linear transformations.

The topical problems also include the category of issues of recognizing each depicted object based on affine transformation, which can radically change the quality of images, practically without affecting the belonging of objects to the recognition categories [10]. The attempts to find a solution to these issues related to the theoretical foundations of image recognition that are called the problem of inversions are reflected in some techniques, and in a multilayer convolutional network [11]. However, to date, most of these techniques can offer only a number of partial solutions that can guarantee only some degree of resistance to limited subsets of transformations. The severity of this range of problems mainly concerns those industries in which image recognition is performed in natural environments (video surveillance, evaluation of visual information coming from a monitoring camera, etc.), when visual sensors, as a rule, have arbitrary and limited viewing angles relative to the observed object [12].

This research aims to implement a comprehensive study to streamline each methodological approach improving methods that ensure the recognition of an image of a three-dimensional object, performing their stable identification at different viewing angles against the background of various visual noises (blurs, occlusions, etc.). The main tasks, the solution of which should ensure the achievement of the set goals, include the creation of methodological approaches that can guarantee the qualitative recognition of each object in the image, regarding its probable non-linear transformations

2. PECULIARITIES OF ADVANCED APPROACHES TO IMAGE RECOGNITION

The specifics of image recognition using a large range of various paradigms (Table 1) is determined by the fact that information expressed in visual signals shows relatively low information capacity. In other words, the most points on raster images (for example, corresponding to the area of monochromatic or equally distributed backgrounds), as a rule, do not have useful information that could affect recognition [13]. At the same time, the dimensionalities of the image used in the complexes that perform information processing are primarily quite large [14]. Thus, most cutting-edge multimedia tools, graphic displays and sensors can provide mass transmission of images (photos, video frames, etc.), in high resolution, with a dimensionality of millions of pixels [15]. This confirms the need to evaluate the potential of each alternative technique capable of recognizing any image. A number of general techniques are often used to obtain the numerical value of such an estimation; these techniques rely mainly on mathematical statistics, and special characteristics enabling to evaluate algorithms that perform machine learning.

Thus, a simple metric capable of evaluating the performance of an algorithm that recognizes a standard image is expressed as the percentage of each qualitatively defined image when the researcher receives a class corresponding to the previously established one on the output channel of an ordered chain of operations [16]. In this case, the estimation is based on a sample built in the same way as the training samples; however, including a number of images that the algorithm did not consider during training. For this purpose, the initial samples are divided into several parts (the sizes of some samples may have significant differences, sometimes reaching 25-35% of the sizes of the pooled samples [17]). The use of this generalized indicator is applicable in the course of recognition with a limited number of classes.

However, solutions in which the number of classes is not limited, and recognition is reduced to the problem of identifying the subject of specific parameters in the aggregate of other, almost limitless, categories, can use a pair (F-degree, the relationship indicator proposed by Matthews, a series of regression indicators, etc.) of qualitative characteristics, accuracy (P), and completeness (R) of estimation. Their practical application distinguishes between false positives (FP): the sorter made a positive decision regarding the image that does not contain the desired object; and false negatives (FN): the sorter failed to identify the object in the image volume, due to recognition errors, or Type 1 and Type 2 errors [18]. Therefore, estimation accuracy indicators will be expressed as the percentage of each specific image (TP) that actually corresponds to this group in relation to all images assigned by the system to this group. It can be expressed like this:

$$P = TP / (TP + FP) \tag{1}$$

The system completeness indicators are determined by the percentage of each image identified by the sorter that belongs to this group in relation to all images of this group in specific test samples:

$$R = TP / (TP + FN) \tag{2}$$

In addition, some indicators that evaluate most metrics of the functioning of ordered chains of operations that are not directly related to the process of making a particular classifying decision are of the essence when evaluating the productivity of recognition methods. These include the dimensions of confidence intervals, estimation of the duration of the algorithm functioning during training, which is necessary to ensure convergence, etc. [19].

The following information supplements the content of Table 1, regarding the effectiveness of evaluating each existing approach to image recognition. Approaches to object recognition, based on techniques that use standard simulations of supervised learning sorters, are characterized by a direct dependence on the dimensions of information from training samples, the duration of study, and a number of indicators that describe convergence when improving the simulation. The existence of a set of characteristics will negatively affect the degree of performance of the models, and often leads to problems associated with the need for retraining if the recognition functions approximated by the models will correctly classify the current samples. At the same time, it is rather difficult to attribute such a picture to the category of key characteristics; therefore, these models will demonstrate relatively low accuracy rates in test samples. The solution to this set of problems is to use an approach that can search for optimal image visualizations, that is, to distinguish a limited number of key features that carry useful information needed during recognition (Table 1).

In the optimal image display approach, only a small percentage of the constituent bitmaps will be critical for recognition. This circumstance explains the low sensitivity of the image to random noise and a number of global distortions. These features are actively used in every algorithm that provides compression with little loss. For example, the JPEG algorithm allows for image compression up to 9-12%, but, as a rule, the level of changes is unnoticed by the eye. Optimal representation opens up the prospect of solving dimensional problems by converting complex images into an optimal format. At the same time, this group of approaches demonstrates acceptable performance in relation to objects that have similar formats (for example, a human face), but it cannot effectively process objects that are characterized by significant visual differences.

Table 1. Units for magnetic properties [20-27]

Approaches based on common image recognition algorithms	Approach principle
Approaches based on standard simulations of supervised learning sorters [20, 21]	This group of approaches is used to recognize such images as a digit, a face, information from a visual sensor belonging to a highly technical system. They are expressed by regression simulations by artificial neural networks, techniques based on the support vector. This group also includes a model-ensemble, which is a combination of most of the simulations mentioned. To train such simulations, a labeled sample of information is usually used, including an array of images, and an array of signs that identifies the group to which certain images belong. The learning of information arrays is divided into several parts: training samples, and test ones. Further, because of the learning rule, which is rather peculiar for certain algorithms, the simulation parameters are adjusted, where the training sample is used. At the same time, the received incoming images are necessarily marked by the model; therefore, they receive a special sign on its output channel, demonstrating the correspondence of the image to a certain class.
Approaches based on models that select optimal image descriptions (distinguish a limited number of key characteristics that carry useful information needed for recognition) [22, 23]	This group of approaches is based on the use of the mentioned technique in images, which reduces the dimensionality. It includes classical methods of basic components, their derivatives, simulations of sparse encodings (there are ordered chains of operations that search for sparse ciphers for lists covering some images, for example, orthogonal ordered operations, and use a special neural network – a sparse autoencoder). Another group of methods, capable of creating a complete vision of objects, is expressed by the restricted machines proposed by Boltzmann. They are represented by a stochastic neural network trained to form a series of special distributions on its own inputs. Usually, it is a derivative of the standard Boltzmann machine.
Approaches that use groups of techniques based on models for identifying each local characteristic. They are the components of the depicted objects [24, 25]	This group of approaches is based on the application in the image of several techniques based on a convolutional neural network, which are among the best simulations in identifying any image. A convolutional neural network is a derivative of multilayer perceptron's. 1) It covers every convolutional and subsample layer, and every fully connected layer. They take advantage of the two-dimensional structures of incoming information through local connectivity techniques, by limiting the number of interconnections among neurons in hidden convolutional layers and incoming information. 2) These approaches are based on techniques that use electronic vision without involving training simulations, and techniques that select key points using a limited identifier.
A group of approaches that apply techniques based on models considering a number of spatial relationships in an environment of limited characteristics during recognition [26, 27]	This group of approaches is based on the use of constellation models in the image. They should be referred to the probabilistic simulations that classify objects with electronic vision. Constellation models are capable of forming a display of objects as a set of N-components of an image selected according to certain geometric patterns. Constellation models consider the relative positions of the components, ignoring data on the spatial location of the key characteristics of objects in the image. The result of creating constellation models when recognizing each item is a guaranteed display of image segments that belong to certain objects. The possibility of identifying characteristics that localize all the key features is the principle of these simulations. This is expressed by the product of individual and oriented values, forms, scales, and other characteristics that affect the image of objects.

In addition, some three-dimensional objects can change their own forms due to geometric transformations. Therefore, to ensure reliable recognition while maintaining integrity, a number of representations obtained using this group of approaches will be vulnerable to problems caused by incomplete information, when some fragments of objects are obstructed or difficult to distinguish against the background of noise. To obtain compact holistic representations, one should consider a matched sample, covering items aligned according to a common principle. The creation of such samples involves some degree of the researcher's participation and the subsequent recognition of the original image by a specialist. This specificity and a number of limitations when using a model that forms an integral representation ensured the development of a new class of techniques based on a set of local image characteristics. They are expressed as stable components of the depicted objects (Table 1). The need for such an approach is expressed by the constancy of each image, if the object can move within the outlined boundaries, and the desired result of identifying algorithms simply correlates the totality of each non-linear representation of objects into one class.

The specificity of the approach based on a convolutional neural network is that such models simply do not have data on the localization of the depicted desired object, and its spatial location. For this reason, when solving practical problems related to the control and processing of information data, the available information about localizations is a key requirement. Thus, having an idea about the location of objects, systems that process data can classify images according to the functions assigned to them. In addition, without knowing the mentioned characteristics, convolutional neural networks will be vulnerable to a general group of questions regarding integrity. That is, models trained on a number of local characteristics of depicted objects will perform positive classification of artificial images when certain components are arranged chaotically, which is a category of errors that people usually do not make. However, this approach has disadvantages associated with convolutional neural network models. Experts also note a number of difficulties regarding the processing of any small object, and difficulties with the localization of noise. Moreover, convolutional neural networks can solve problems due to qualitative identification, and the identification of any highly specific characteristic that is difficult for humans to detect [28].

At the same time, operations for convolution of convolutional neural networks can provide the proper degree of nonlinearity only with respect to the translations of each local feature. Their structure does not provide resistance to other transformations, for example, rotations, or mirror reflections. For this purpose, several heuristic techniques are used (aligning the image along the horizon, attracting spatial pyramids and copies of the originals). Thus, objects having different orientations are modeled by teaching various lists of characteristics. During recognition, it is required to train a convolutional neural network and several other approaches based on local non-linear characteristics lists of features that optimize the level of non-linearity for current samples, however, this does not provide a discriminatory display of objects.

Several alternative approaches capable of extracting local characteristics are covered by standard computer observation, where training operations are not required. Such techniques are engaged in the calculation of only individual image fragments that correspond to a number of predefined conditions. These include the probability of detecting an edge/border, the probability of detecting a corner, or a key point, and the probability of detecting unequal fragments. The above list of local features is actively used for visual tracking, and for keeping track of any objects. However, in the initial version, they cannot perform qualitative identification, since they have a non-discriminatory nature. But there are a number of effective approaches that satisfy most requirements mentioned. Most of them are based on the idea of searching for a "point of interest". In this case, local identifiers are used, which are expressed by compositions of image segments that are limited in a certain way. Here, each fragment needs to calculate exact characteristics, scales, etc., which makes it possible to identify the location of characteristic that will correspond to these identifiers [29].

Characteristics based on multiple local identifiers can be applied to image recognition by transforming it in 3D spaces. The locality of characteristics ensures the exclusion of occlusions, which makes it possible to compare parts of objects. The key problems of such methods include low resistance to nonlinearity. Ultimately, this does not allow such algorithms to establish the belonging of objects that differ in texture to the same group. However, techniques that are able to identify local characteristics make it possible to eliminate the complexity of recognizing each image, thereby guaranteeing increased resistance to occlusions, and reducing computational costs in the image recognition process with a large dimensionality, which opens up the possibility of creating non-linear characteristics that can detect objects using a non-linear transformation. At the same time, with all the variety of models and paradigms that provide training and identification of each classifying feature, key techniques recognize images by non-linear 3D transformations. The existence of a complex of problems associated with the localization of an object in images and the vulnerability to the disordered placement of each local feature suggests the need to find a solution that provides identification methods with an effective set of tools (Table 1).

Constellation models have significant differences from the models that have been considered before. Thus, from integral, and local identifications, they differ in a list of pre-established hypotheses regarding how exactly the structure of an object should be organized. It is also assumed that any objects that are difficult to evaluate should be made up of spatial components. Any structured object represented by local characteristics can be described by key features, but it should be considered that not every one of them will be useful.

Trained constellation simulations can identify the constituent elements of objects and make predictions about whether images fit into a group. Training has its own specifics. In this case, the system understands that the supervisor is either absent, or he has little effect on the creation of the simulation characteristics. Therefore, prior to training, the algorithm is not familiar with the rules that allow one to correlate the constituent parts of the simulation with local characteristics. They also do not know the data format of the components. The incoming information is expressed in the form of samples covering images marked with the sign "objects" and "noises". Upon completion of training, those characteristics are selected that increase degree of probability that images belong to a particular group, guided by spatial parameters.

Due to evaluation of the R and P indicators, recognition by constellation models demonstrates that these models can process large samples (up to 100 classes), covering various arrays. At the same time, it should be considered that standard constellation simulations, as a rule, do not guarantee the proper degree of sensitivity to orientations when the corresponding image belongs to the sample. In addition, with an increase in the number of links that form the skeleton of constellation models, the computation costs increase significantly. For this reason, the number of studies on the modernization of classical constellation models has been increased. Thus, today the following models are proposed: derivatives of constellation models that consider a number of established requirements for the use of characteristics that are not linear to rotations, and improved constellation models, in which each limited segment is determined by a priori assumptions about the nature of such characteristics.

Any trainable simulation is expressed by a relatively small range of recognition techniques, moreover, theoretical ones based on samples covering exclusively aligned images, in which each object is spatially oriented in the same way. A completely new solution for evaluating the recognition technology of each 3D image of an object, aimed at qualitative recognition, made it possible to develop a fundamentally new simulation based on a hierarchical correlation of each local feature containing useful information about the correspondence of images to any group. The simulation is characterized by a limited size, a limited image and can determine the characteristics of the presence of objects, and their location in the 3D space. As a novel approach that teaches such a simulation, a learning method based on information flows is proposed. It enables to obtain information about the spatial relationship of any local attribute of objects, and to find solutions for image recognition containing randomly oriented 3D objects. It is suggested to do the following:

1) Identify each local characteristic with the calculation of indicators that determine the visibility of any image sectors;
2) Provide optical tracking. Its potential is capable of processing video fragments and effectively tracking the detected fragments in the volume of a complete and continuous stream of video files;
3) Implement the simulation by the modernized optimizer using the gradient descent technique and the best aspects of trained transforming autoencoders.

3. OPTIMAL APPROACH TO ENSURE THE QUALITATIVE RECOGNITION OF EACH IMAGE

A completely novel hierarchical mechanism that includes information about each volumetric characteristic of the depicted object is the basis of such approaches, capable of solving a set of problems to ensure the qualitative recognition of any image. Here, the hierarchical correlation of each local feature is a logical basis [30]. This simulation makes it possible to form representations of the object being assessed and a natural image. Such a somewhat formalized display of this simulation and its components serves as the basis for displaying a two-level abstraction: a bounded detector and a representative layer. This imitation is expressed by an ordered hierarchical chain, including levels $L_0, L_1, ..., L_C$. Any of them covers the bounded detectors of useful characteristics D_i and processes the incoming information obtained during the operation earlier. Then the processing result is calculated, after which the information is fed to the input channels of the higher stages. As a result, we have the value of the sorter dependency and the characteristics of localized objects evaluated using a bounded detector.

The incoming information to the L_0 levels is a list of chains, including local image segments and transformations obtained from video fragments offered by the electronic sensor. Based on the frame-by-frame chain, the equivariant detector is trained. In this case, the model is trained at each level. Upon completion of the 1st level training, the initial combinations of video fragments are transmitted as constellation graphs. By any i-levels, where $i \neq 0$, information is clustered by combining disparate nodes of incoming graphs. At the higher levels, the training is carried out in the same way. Further operations will be repeated. At the same time, trained simulations are able to offer high-quality images, but as limited numbers (1-3), including high-order detectors.

As a result, models can be described by:
- The number of representations $L_0, L_1, ..., L_C, L$;
- The number and location of each detector of characteristic D_i;
- The number of values K_C, which corresponds to the level L_C;
- A number of internal characteristics inherent in standard detectors.

Thus, bounded equivariant feature detectors form the basis of the proposed simulation capable of finding a stable image recognition. They use a set of formats of 3D objects, and can set the initial formats of objects and the transformations of objects included in the current projection for these cases. The mission performed by bounded detectors is similar to the mission of each neuron in artificial neural networks: bounded detectors are also included in a hierarchical chain.

The operations studied in [30] related to optical tracking, and training of the structural components of simulations - bounded detectors of characteristics act as an integral part of this approach, in addition to identifying each local feature. A pair of extreme operations opens up the possibility of extracting information necessary for training and obtained from incoming video streams. In this case, labeled samples and specialist assistance are not required.

An algorithm that performs the classification of each object using pre-trained models is the main, final, component of this technique, which provides high-quality recognition of any image. The principle of operation of this ordered chain is shown in Figure 1. Identification is made up of chains of inclusion of model detectors. In the case when we have a trained model M_i that operates with some images c_i, or with several models, to solve the problem of multiclass identifications, and to determine images I, the recognition operation requires:

1. Forming cumulative memory (CM). It can be represented by several levels (stage 1).
2. The identifying dependence of the detectors should be applied to any l-th simulation level L_l, and any detector $b_j^{\langle l \rangle}$, with regard to each image sector, (stage 2).
3. When the series of levels L_l activating all detectors is negative, the images will correspond to the class c_i.
4. When the complex of recognition tasks belongs to multiclass ones, all operations of stage 1 should be conducted, in relation to simulating the subsequent categories M_{i+1}. If not, everything goes back to stage 3.
5. When the L_l layer detector is activated, and the detectors of the remaining layers are activated, the algorithm of stage 2 is repeated.
6. When the activation of the extreme levels is positive, the images are included in the c_i stage (stage 4).

Multiclass recognitions will be repeated in relation to the sets of trained chains $M_0, M_1, ..., M_n$, but the degree of image correspondence to any group will be determined relative to the highest active level (stage 5). Guided by the fact that some image sectors are expressed by edges, boundaries and corners, it is important to indicate the increase in the diversity of each detector at other stages of the simulation.

Experiments demonstrate that the number of first-level detectors during training does not exceed 100, with models trained on objects of different categories being able to share some of the first-level detectors between them, demonstrating an effect reminiscent of transfer learning or pre-training without a teacher. Recognition algorithm using a two-level model can be effectively applied to solve problems of recognizing such objects as human faces, in some cases (certain scale) - silhouettes, cars. In general, the heuristic rule for selecting the number of levels for the model depends on the number of degrees of freedom of the object of interest, and can't be unambiguously established from the image due to the effects of noise and limited observation space.

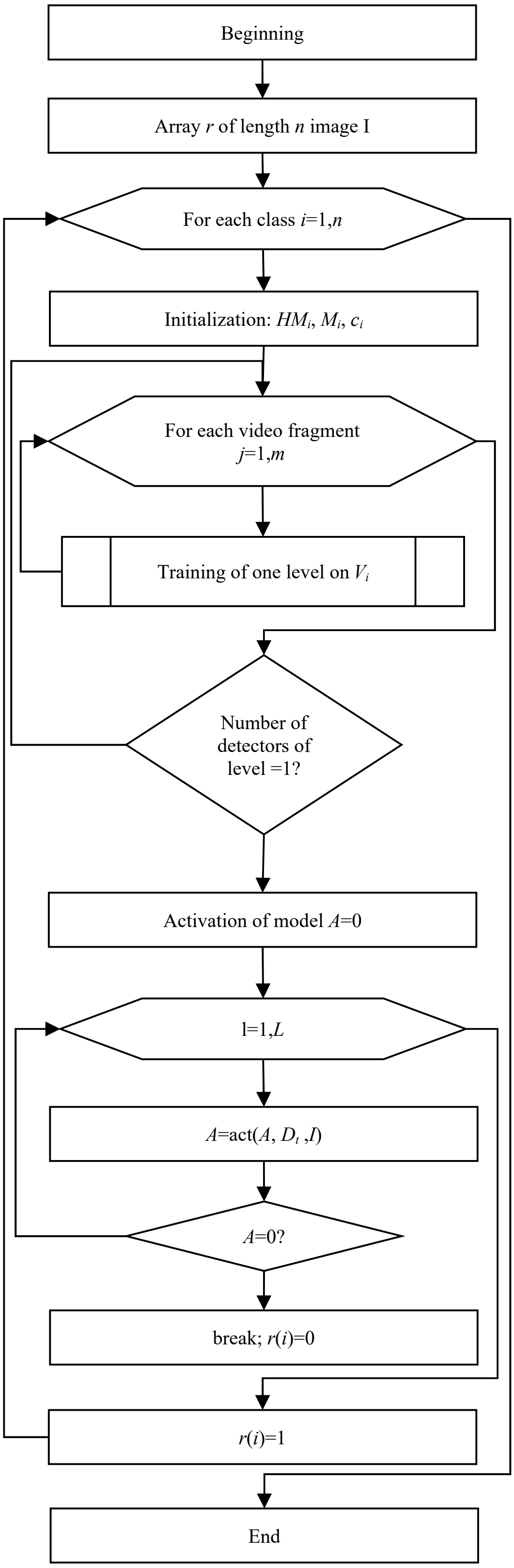

Figure 1. The block diagram of image recognition operation

4. RESULTS

An effective, in terms of accuracy, completeness and the proportion of correctly recognized images, methodical approach was developed that enables to find a solution to ensure high-quality image recognition, regarding probable non-linear transformations. In this case, a completely novel hierarchical modeling is proposed, covering information about a number of spatial indicators inherent in the depicted object. The operation of hierarchical matching of key features is its logical basis. The recognition performance was evaluated on the basis of open databases, and the HPID database containing labeled images of human faces in different orientations relative to the camera was used to test the performance of the transforming autoencoder. The confidence level value was chosen to be 0.95. Accuracy, completeness, and proportion of correctly recognized images were used to evaluate the performance of the methods involved in the experiment. The method of confidence intervals calculation was used to estimate the sample.

For images that semantically correspond to the video fragment used for training, the detector demonstrates flexibility in conditions of limited sampling - for example, 2-3 video fragments obtained using computer graphics were sufficient to recognize a sample of HPID with the presence of 15 people differing in gender, race, and facial expressions. Since equivariance - the ability to identify objects in an image in different orientations - is positioned as one of the main advantages of the model under consideration, the experimental samples are grouped as follows: some initial object position with Euler angle rotation coordinates is considered, while the group includes images of objects that have undergone rotation transformations.

Face recognition was tested by comparing Viola-Jones methods, SVM classifier combined with histogram computation of oriented gradients, random forest and convolutional network trained on ImageNet sample. Experimental verification, using open databases, showed its advantage, even in comparison with its main competitor, a convolutional neural network. Thus, the recognition accuracy when using the proposed method increases by 7-10%, and the recognition accuracy when changing the orientation of the object in the image experiences deviations of no more than 4-5%. Table 2 and Figure 2 show face recognition accuracy between existing methods and the proposed method.

Table 2. Face recognition accuracy rates for groups of different orientations, %

Orientation angle	Viola-Jones method	SVM+HOG	CaffeNet	Proposed author's method
(0°,15°)	85 ± 4	84 ± 5	87 ± 4	92 ± 3
(15°,30°)	68 ± 4	75 ± 3	85 ± 4	91 ± 4
(30°, 45°)	61 ± 2	72 ± 3	86 ± 2	90 ± 2
(45°, 60°)	66 ± 3	78 ± 5	83 ± 3	89 ± 4
(60°, 90°)	86 ± 3	77 ± 2	82 ± 3	92 ± 3
(90°, 120°)	64 ± 3	80 ± 2	83 ± 3	90 ± 2

The recognition according to the evaluated approach opens up the possibility of identifying the depicted objects with the calculation of indicators of their limitation in the 3D space. At the same time, the implementation of the mentioned algorithm for creating a simulation is capable of quickly making recognition without requiring significant resources.

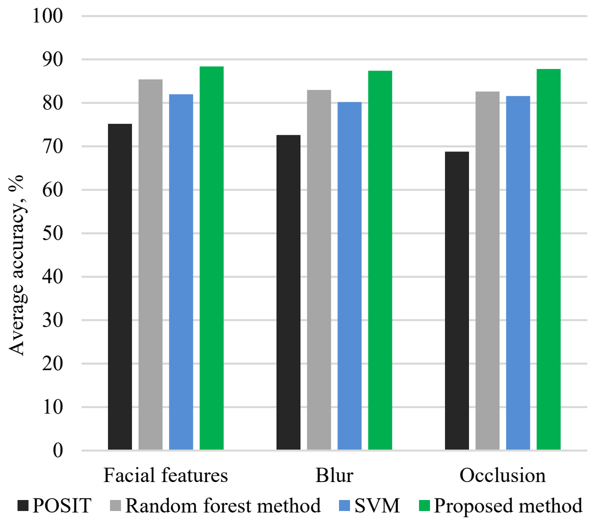

Figure 2. The graph is showing the difference in performance between the methods

The most problematic cases for the work of the equivariant detector are anomalous situations not encountered in the video fragment - for example, occlusion of the searched area of the image by third-party objects (for example, the detector demonstrates an error growth of up to 10% when processing images of faces with glasses on them). During the experiments, a slight decrease of up to 5% of recognition accuracy was found in the situation when objects in the image were occluded. Such errors are consistent with the principle of the model, which requires the presence of activated detectors of the lower level for positive activation of higher-level detector.

Since the simulation detectors are optimized discriminative neural networks, each bounded detector is resistant to a variety of interferences in the local sphere, opening up the possibility for identification algorithms to solve basic object recognition problems. The use of the mentioned approach in ordered operations makes it possible to achieve the required resistance to occlusions: the decrease in the recognition accuracy of the developed method is 3-5% against 5-10% for alternative methods.

5. CONCLUSIONS

Within the framework of this research, an approach to ensuring high-quality image recognition has been developed, which has novelty, significance, theoretical and practical contributions, regarding probable non-linear transformations, which is based on simulation, including information about the spatial indicators of the depicted objects. The use of this approach makes it possible, by providing stable recognition under conditions of different viewing angles, and various types of visual noise (blurring, occlusion), to achieve an increase in the efficiency of computer vision and decision-making systems through the use of compact hierarchical representations that require a much lower computational load compared to alternative methods. The peculiarities of the presented approach make it possible, in future theoretical and practical studies, to use it for ensuring the solution of highly specific problems and for identifying various kinds of patterns in video surveillance and self-learning of the detected structures.

The program-target aspects concerning the management of the further development of this research should primarily be focused on planning and programming, providing and conducting empirical studies, using quantitative and qualitative methods when designing information control and processing systems in computer vision and image recognition. The results given in the paper are of practical interest in the design of control and information processing systems in the field of computer vision and image recognition, for those tasks where there is a need to determine the spatial parameters of the depicted objects.

ACKNOWLEDGEMENTS

The presented results were obtained under the Subsidy Agreement No. 075-11-2021-050 of June 24, 2021, "Development of a multi-format information system for automated distributed monitoring and on-line quality control of information and medical services using artificial intelligence technologies" with the Ministry of Science and Higher Education of Russia.

REFERENCES

[1] M. Vasanthi, K. Seetharaman, "Facial Image Recognition for Biometric Authentication Systems Using a Combination of Geometrical Feature Points and Low-Level Visual Features", Journal of King Saud University - Computer and Information Sciences, Vol. 34, No. 7, pp. 4109-4121, 2022.
[2] G.S. Lebedev, E.Yu. Linskaya, V.Yu. Terekhov, A.A. Tatarkanov, "Monitoring and Quality Control of Telemedical Services via the Identification of Artifacts in Video Footage", International Journal of Intelligent Systems and Applications in Engineering, Vol. 11, No. 2, pp. 82-92, 2023.
[3] U. Kosarkar, G. Sarkarkar, S. Gedam, "Revealing and Classification of Deepfakes Video's Images Using a Customize Convolution Neural Network Model", Procedia Computer Science, Vol. 218, pp. 2636-2652, 2023.
[4] Y. Guan, et al., "Face Recognition of a Lorisidae Species Based on Computer Vision", Global Ecology and Conservation, Vol. 45, p. e02511, 2023.
[5] K. Sharadhi, V. Gururaj, S.P. Shankar, M.S. Supriya, N.S. Chogule, "Face Mask Recognizer Using Image Processing and Computer Vision Approach", Global Transitions Proceedings, Vol. 3, No. 1, pp. 67-73, 2022.
[6] H. Hofbauer, A. Uhl, "Identifying Deficits of Visual Security Metrics for Images", Signal Processing: Image Communication, Vol. 46, pp. 60-75, 2016.

[7] W. Wang et al., "TNNL: A Novel Image Dimensionality Reduction Method for Face Image Recognition", Digital Signal Processing, Vol. 115, p. 103082, 2021.
[8] X. Wang, W. Zhang, "Anti-Occlusion Face Recognition Algorithm Based on a Deep Convolutional Neural Network", Computers and Electrical Engineering, Vol. 96, p. 107461, 2021.
[9] S. Gerke, A. Linnemann, K. Müller, "Soccer Player Recognition Using Spatial Constellation Features and Jersey Number Recognition", Computer Vision and Image Understanding, Vol. 159, pp. 105-115, 2017.
[10] Y. Zhao, L. Li, Z. Liu, "A Novel Algorithm Using Affine-Invariant Features for Pose-Variant Face Recognition", Computers and Electrical Engineering, Vol. 46, pp. 217-230, 2015.
[11] L. Jiao, S. Zhang, L. Li, F. Liu, W. Ma, "A Modified Convolutional Neural Network for Face Sketch Synthesis", Pattern Recognition, Vol. 76, pp. 125-136, 2018.
[12] B. Gaikwad A. Karmakar, "End-to-end Person Re-Identification: Real-Time Video Surveillance Over Edge-Cloud Environment", Computers and Electrical Engineering, Vol. 99, p. 107824, 2022.
[13] R. Chellappa, J. Ni, V.M. Patel, "Remote Identification of Faces: Problems, Prospects, and Progress", Pattern Recognition Letters, Vol. 33, No. 14, pp. 1849-1859, 2012.
[14] A. Tatarkanov, A. Lampezhev, D. Polezhaev, R. Tekeev, "Suboptimal Biomedical Diagnostics in the Presence of Random Perturbations in the Data", International Journal of Engineering Trends and Technology, Vol. 70, No. 11, pp. 129-137, 2022.
[15] P. Hu, H. Ning, T. Qiu, Y. Xu, X. Luo, A.K. Sangaiah, "A Unified Face Identification and Resolution Scheme Using Cloud Computing in Internet of Things", Future Generation Computer Systems, Vol. 81, pp. 582-592, 2018.
[16] S.A. Chowdhury, M.M.S. Kowsar, K. Deb, "Human Detection Utilizing Adaptive Background Mixture Models and Improved Histogram of Oriented Gradients", ICT Express, Vol. 4, No. 4, pp. 216-220, 2018.
[17] A. Salim, Juliandry, L. Raymond, J.V. Moniaga, "General Pattern Recognition Using Machine Learning in the Cloud", Procedia Computer Science, Vol. 216, pp. 565-570, 2023.
[18] H. Sufikarimi, K. Mohammadi, "Role of the Secondary Visual Cortex in HMAX Model for Object Recognition", Cognitive Systems Research, Vol. 64, pp. 15-28, 2020.
[19] Z. Chen, Z. Sun, Y. Pei, L. Yin, "Generalized Sparse Codes for Non-Gaussian Channels: Code Design, Algorithms, and Applications", Fundamental Research, Vol. 2, No. 2, pp. 284-295, 2022.
[20] S. Zhao, W. Liu, S. Liu, J. Ge, X. Liang, "A Hybrid-Supervision Learning Algorithm for Real-Time un-Completed Face Recognition", Computers and Electrical Engineering, Vol. 101, p. 108090, 2022.
[21] M. Badea, C. Florea, A. Racoviţeanu, L. Florea, C. Vertan, "Timid Semi-Supervised Learning for Face Expression Analysis", Pattern Recognition, Vol. 138, p. 109417, 2023.
[22] R.D. Parizs, D. Torok, T. Ageyeva, J.G. Kovacs, "Machine Learning in Injection Molding: An Industry 4.0 Method of Quality Prediction", Sensors, Vol. 22, No. 7, p. 2704, 2022.
[23] R.T. Tomihama, S. Dass, S. Chen, S.C. Kiang, "Machine Learning and Image Analysis in Vascular Surgery", Seminars in Vascular Surgery, Vol. 36, No. 3, pp. 413-418, 2023.
[24] R. Thanki, "A deep Neural Network and Machine Learning Approach for Retinal Fundus Image Classification", Healthcare Analytics, Vol. 3, p. 100140, 2023.
[25] O. Topal, R. Chitic, F. Leprevost, "One Evolutionary Algorithm Deceives Humans and Ten Convolutional Neural Networks Trained on ImageNet at Image Recognition", Applied Soft Computing, Vol. 143, p. 110397, 2023.
[26] T. Czerniawski, F. Leite, "Automated Digital Modeling of Existing Buildings: A Review of Visual Object Recognition Methods", Automation in Construction, Vol. 113, p. 103131, 2020.
[27] P. Karthick, S.A. Mohiuddine, K. Tamilvanan, S. Narayanamoorthy, S. Maheswari, "Investigations of Color Image Segmentation Based on Connectivity Measure, Shape Priority and Normalized Fuzzy Graph Cut", Applied Soft Computing, Vol. 139, p. 110239, 2023.
[28] N. Remzan, K. Tahiry, A. Farchi, "Automatic Classification of Preprocessed MRI Brain Tumors Images Using Deep Convolutional Neural Network", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 54, Vol. 15, No. 1, pp. 68-73, March 2023.
[29] A. Dhandayuthapani, J. Lawrence, "Plant Disease Recognition Using Optimized Image Segmentation Technique", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 50, Vol. 14, No. 1, pp. 211-218, March 2022.
[30] J. Shi, Z. Li, H. Zhao, "Feature Selection via Maximizing Inter-Class Independence and Minimizing Intra-Class Redundancy for Hierarchical Classification", Information Sciences, Vol. 626, pp. 1-18, 2023.

BIOGRAPHIES

Name: **Valery**
Middle Name: **Yurevich**
Surname: **Terekhov**
Birthday: 02.02.1961
Birthplace: Minsk, Belarus
Bachelor: Military Radio Engineer, Military Unit 63002 Istrinsky District, Moscow Region, Minsk Higher Engineering School of Antiaircraft Missile Defense Forces, Minsk, Belarus, 1981
Master: Military Radio Engineer, Military Unit 63002 Istrinsky District, Moscow Region, Minsk Higher Engineering School of Antiaircraft Missile Defense Forces, Minsk, Belarus, 1981

The Last Scientific Position: Leading Engineer, Science and Technology Park of Biomedicine, I.M. Sechenov First Moscow State Medical University, Ministry of Health, Moscow, Russia, Since 2019
Research Interests: Artificial Intelligence, Telemedical Services, Quality Control
Scientific Publications: 4 Papers, 1 Project, 13 Theses

Name: **Elena**
Middle Name: **Yurievna**
Surname: **Linskaya**
Birthday: 07.07.1988
Birthplace: Moscow, Russia
Bachelor: Applied Mathematics, Department of Applied and Computational Mathematics, Faculty of Informatics, Moscow State University of Instrument Engineering and Informatics, Moscow, Russia, 2016
Master: Mathematical Engineer, Applied Mathematics, Department of Applied and Computational Mathematics, Faculty of Informatics, Moscow State University of Instrument Engineering and Informatics, Moscow, Russia, 2020.
The Last Scientific Position: Leading Engineer, Science and Technology Park of Biomedicine, I.M. Sechenov First Moscow State Medical University, Ministry of Health, Moscow, Russia, Since 2021
Research Interests: Automated Control Systems, Informatics and Information Systems
Scientific Publications: 14 Papers, 6 Patents, 4 Theses

Name: **Abas**
Middle Name: **Khasanovich**
Surname: **Lampezhev**
Birthday: 07.07.1997
Birthplace: Murmansk, Russia
Bachelor: Navigation and Ballistic Support for Use of Space Technology, Department of Dynamics and Flight Control of Rockets and Spacecraft, Faculty of Special Mechanical Engineering, Bauman Moscow State Technical University, Moscow, Russia, 2016
Master: Navigation and Ballistic Support for Use of Space Technology, Department of Dynamics and Flight Control of Rockets and Spacecraft, Faculty of Special Mechanical Engineering, Bauman Moscow State Technical University, Moscow, Russia, 2020
The Last Scientific Position: Researcher, Institute of Design and Technology Informatics, Russian Academy of Sciences, Moscow, Russia, Since 2020
Research Interests: Multi-Parameter Optimization and Mathematical Modeling, Multi-Parameter Optimization and Mathematical Modeling
Scientific Publications: 12 Papers, 3 Theses