UDC 681.5.01

Y.Timoshin, Y.Yuzhda

ANALYSIS OF FEATURES OF APPLICATION OF NEURAL NETWORKS FOR INTELLECTUAL PROCESSING OF VIDEO FLOWS OF TECHNICAL VISION SYSTEMS

Abstract: The article considers topical issues of application of modern technologies and methods of object detection and recognition. The article is devoted to the analysis of the peculiarities of the application of different types of neural networks in the process of step - by - step processing of video data obtained from the systems of technical vision of robots, video monitoring systems, intelligent security systems. A review of modern literature, which describes the method of forming a space of features of the description of objects and methods of their recognition. The review shows that the process of intelligent processing of video data consists of many stages of image processing, one of which is processing using neural networks as intelligent components. The validity of the stages of processing in real time requires justification for the use of different types of neural networks in different processing processes in order to improve the quality and optimize the processing time of such data. The structure of the video image processing model is given. The article also defines the types of neural network at different stages of data processing (such as identification of parameters and characteristics of the group, finding group objects, sector-by-sector image processing, object classification, object recognition, creating a contour model of the object, detection object in the sector, evaluation of sector parameters, definition of information sectors, division of frames into sectors, processing of information frames) according to the proposed hierarchical model, with subsequent use of the obtained results for multiagent system of distributed intelligent processing of video monitoring objects and examples of further application the results obtained.

Keywords: neural networks; intellectual processing; real-time processing; video data; video stream; computer vision systems; processing model; multiagent system.

Introduction

A review of modern literature, which describes the method of forming the space of features of the description of objects and methods of their recognition, showed that the process of intelligent processing of video data from various systems of technical vision consists of many stages of image processing, one of which is processing using neural networks as intelligent components [1]. Considers a hierarchy of stages, which includes image recovery, object selection, segment parameter estimation, trajectory tracking, object parameter estimation, object detection, recognition and maintenance, which requires processing and information at several hierarchical levels. and enables the creation of real-time processing systems [1]. Considers the application of the idea of Huff transformation to build a neural network that normalizes and filters the video signal, which then goes through the stage of highlighting the basic properties of images, as well as compressing information for easy recognition [2]. The last layer is the classification layer. The accuracy with which the neural network can recognize images directly depends on its structure and settings. The disadvantage of the object recognition approach according to the proposed standard as a whole is the large amount of primary data.

The work [3] is devoted to solving the problem of detecting objects in the image from the standpoint of the theory of active perception (TAV), the use of which allows to recognize objects of arbitrary classes with acceptable complexity and high accuracy, due to a new approach to forming appropriate features using Q- a transformation corresponding to the image pre-processing step and a set of 4x4 filters that are similar to Walsh filters. The characteristic description for the images which are in 2 databases - "positive" and "negative" samples on the basis of which the spatial description of characteristic images is formed is calculated.

In [4], the application of the Fuzzy Evolutionary Classifier Takagi-Sugeno for the tasks of detecting and tracking objects in the video stream is analyzed, for which a fuzzy evolution classifier is proposed, which separates pixel blocks into background and objects of interest (foreground objects) and predicts position of moving objects. They are then compared with the input data of neighboring blocks in order to cluster parameters into two classes: belonging to significant foreground objects or image background, with high efficiency of the algorithm for selecting foreground objects by taking into account the movement of the objects themselves.

In [5] the algorithm of object detection based on training of the Convolutional neural network (ZNM) is presented. The peculiarity of the algorithm is that in a single network two tasks are solved simultaneously: the selection of rectangular blocks containing objects, and determining the affiliation of an object of a certain class. To recognize the original image for each cell is formed a label that characterizes its belonging to the object of a class. The advantage of the proposed algorithm is the high speed of image processing. Disadvantages include the cost of network training, reduced detection accuracy compared to similar algorithms, limitations associated with the size of the grid cells (there are difficulties in detecting small objects and objects located close to each other).

In [6] shows an effective approach for processing low quality images. When forming the description of the object, an autocoder is used, which is a neural network with a symmetrical architecture, which was first described in [7]. The autocoder consists of input and output layers of the same dimension and hidden layers with a smaller dimension than the input and output layers. The autocoder has a hidden layer called the bottleneck layer, which has the smallest dimension of all the hidden layers. The dimensionality of the layers of the autocoder when searching from the input layer to the bottleneck layer is successively reduced, and when searching from the bottleneck layer to the output layer is sequentially increased. A set of autocoder layers from the input to the bottleneck layer is called an encoder, which converts some input signal into an output signal. The output signal from the bottleneck layer of the autocoder is then fed to the input of the polynomial regression module RVM (relevance vector machine) [8], which calculates the probability of an image of an object of a certain class on an elongated area of the photograph.

In [9], Recurrent Neural Networks (RNS) are analyzed - feedback neural networks (NN), which include, for example, Hopfield neural networks.

This analysis shows the complexity and ambiguity of video data processing processes and the variety of methods to achieve positive results, which requires further analysis and research.

Shows the main stages of video data processing on board the UAV [10], which include traditional operations of primary processing (image preparation and restoration) and semantic processing operations (segmentation, feature selection, classification, recognition, interpolation, feature selection, identification and classification) with using procedures - the operation of convolution in the spatial area; filtering in the space-frequency domain; shading correction, rank filtering; local averaging procedures; analysis of logical connections in the image, including those formed on board the UAV. In [11] it is shown that the increase of efficiency of NM application is possible due to the use of several technologies within one information security (IS) system, for example, a team (ensemble) of neural networks. When constructing an ensemble of neural networks, a finite set of pre-trained neural networks is used at the same time, the output signals of which are combined into a combined assessment, which exceeds the quality of the results obtained using local networks included in the ensemble. An important area is the development of classification models based on the application of a multi-agent approach that allows you to generate neural networks - classifiers of the ensemble, taking into account the characteristics of subject areas and use them as intelligent agents. The paper [12] is devoted to the problem of using neural networks in the process of data mining as a tool for their clustering based on the detection of hidden patterns. The analysis of the Kohonen neural network shows that it is mainly designed for unmanaged learning, which means that the network learns to understand the very structure of the data. It is argued that one of the most significant shortcomings of the Kohonen neural network is that the model does not provide for the determination of the number of clusters, but it is able to operate in interference conditions due to the fact that the number of clusters is fixed in advance. Analyzes and compares different algorithms related to the Harris and Stevens Detector, RANSAC (RANdom Samp le Consensus), Moravets Detector, and others for the use of detectors and descriptors in computer vision problems [13]. The issues of designing a system based on a neural network of fuzzy logic are considered. It is shown that it can turn into an extremely difficult task, because with the expansion of the task domain, the size of the rule base and computational needs will increase. Due to the resource and complexity of designing classification systems, fuzzy logic theory is mostly used in earlier stages of computer vision, such as segmentation or filtering.

Problem statement

The problem of processing video streams using computer vision systems using neural network (NM) technology is uncertainty - which models of neural networks and at what stages of processing is better to use to obtain the best intermediate and final results for distributed processing after technical vision (with prior data processing) and the following stages of analysis: definition, identification and classification of monitoring objects. For greater identification accuracy, it is important to detail the components of the object, and I often do it against the background of land or man-made structures.

The aim of the article is to determine the types of neural network at different stages of data processing according to the proposed hierarchical model, followed by the results for a multi-agent system of distributed intelligent processing of video monitoring objects and examples of further application of the results.

Video stream processing model using computer vision systems and neural network technologies

The model proposed for analysis has a hierarchical structure with the appropriate stages of primary and secondary processing of video images of monitoring objects, the images of which are transmitted from different technical systems - from ground-based surveillance systems to UAV optical systems.

Components of data processing stages

Components of data processing stages for identification of monitoring objects:

1) Monitoring objects that are on some natural background (land / water / air) or on the background of artificial structures.

2) On-board video cameras that form images of monitored objects.

3) The stage of improving the image quality, which has several sub-stages

* Primary processing

- image filtering using Fuzzy Neural Networks (Fuzzy Networks) - frame recovery (interpolation of points in image fragments)

* Final finishing

- normalization of the frame image (possibly based on the OpenCV library)

- frame segmentation (dividing it into sectors) using FNN

4) Analysis of images in sectors (leads to parallelization of calculations)

a) selection of metrics for each type of objects (based on metric database)

b) identification of key features for specified objects

c) search for objects by features

d) selection of sign points of objects in the i-th sector of the k-th frame of the image

e) identification of "empty" sectors in the frame and their removal from the processing

1. clustering of the image background;

2. clustering of objects using a database of key features and "fuzzy evolutionary classifier Takagi-Sugeno";

3. normalization of the image of the object {0; 1} on the plane on X * Y;
4. normalization of the image of the object {0; 1} on the height H;
5. formation of a set of main components of the image of the object;
6. use of the wrapped neural network (WNM);

- activation of ZNM neurons (transmission to the input of the set network data);
- convolution of functions;
- ZNM training;
- obtaining the structure of NM;
- determining the number of channels;
- determining the dimension of the image (grid of pixels);

7. identification of the object in the sector;
8. classification of the object in the sector;

- use of image library;
- use of ZNM;

9. object recognition;

- select the image of the object from the background of the frame;
- clarification of object parameters;
- calculation of location parameters in the frame in X * Y coordinates (based on the "fuzzy Takagi-Sugeno model");
- search for structural elements in the image of the object of analysis using the "Huff Transformation";
- analysis and identification of structural elements of the object and their parameters;

10.sector-by-sector processing of the frame;

5) Creating a new or refining a digital model of the optical image of the object that is in the image database.

6) Go to processing in other image frames, search for new objects (go to step (4) / s) ".

7) Forming a message with the results of the analysis and the model of the object.

8) Transfer of results to the customer.

9) Deciding on the location of the object in a particular sector with the appropriate coordinates.

10) Replenishment of the library of optical images of monitoring objects.

When clustering an object, the following processing steps are performed:

- several clusters are defined for each object;
- for each of the latter, the number of key points of the image assigned to it is calculated;
- a description of each image is built from clusters.

For example, clusters of the nose, eyes, chin, etc. are defined for the face. On each part of the face there is a model of critical points which are investigated at the analysis.

Summarizing the above, we can determine in Table 1 - for which stages of video stream processing can be used appropriate models of neural networks and their components to optimize time and computing resources, improve the quality of the estimates. The initial stages are located at the bottom of the table. The processing sequence is determined according to the specified steps from bottom to top.

Table 1

Levels of application of neural networks at different stages of processing

Stages of processing	Neural network application levels
Multi-agent processing of group objects	
Identification of parameters and characteristics of the group	
Finding group objects	Fuzzy evolutionary classifier Takagi-Sugeno
Sector-by-sector image processing	Based on the "fuzzy Takagi-Sugeno model"
Identification of object parameters: • Coordinates of the center • Geometric dimensions • Brightness / contrast level • Object shape • Special external signs	Using the "Huff Transformation"
Object classification	
Object recognition	Convolutional neural network (CNN)
Formation of a generalized search image (with min digital code)	Using an autocoder
Creating a contour model of the object	
Outlining an object	
Create a simplified object model (based on sharp contrast lines)	
Detect an object in a sector	
Estimation of sector parameters: • Coordinates of the center • Geometric dimensions • Brightness level and contrast • The shape of the boundaries of the Sector	
Definition of information sectors	Fuzzy evolutionary classifier Takagi-Sugeno
Division of the frame into sectors	
Information frame processing	Fuzzy Neural Network (FNN)
Definition of information images and their normalization along the axes	Fuzzy Neural Network (FNN)
Primary image processing • Noise filtration • Smoothing frame sequences • Interpolation of missing points	
Formation of digital images	

Secondary data processing tasks that use the results of the proposed model

I. Creating an electronic model of the object of analysis or monitoring

An electronic model of an object may have a different combination from other models and consist of their compilation for use in recognition and identification tasks for different purposes, which include, for example, the following:

a) Digital model of the optical image of the object.

b) Digital model of the infrared image of the object.

c) Digital model of ultrasound image of the object.

d) Digital model of laser object range monitoring object detection.

Taken together, the data in these models should create a multidimensional object model that is more informative than any of these.

II. Correction of the trajectory of the monitored object on the route map

A system for comparing digital optical image objects of a terrain map with an electronic route map while driving.

In order to clarify the coordinates of the location of the object on the route of movement often use at the final stage of movement embedded in the memory of the object electronic route map for comparison with digital optical images of the map, which are compared by a set of parameters and their characteristics, and on the basis of which the trajectory of the object can be adjusted.

To do this, "passports" are developed for different types of objects, which are displayed and captured in the image in the frame of the video stream, as well as for individual "beacons" of the background and the ground. The background can be a combination of natural and artificial objects.

It is then necessary to correlate or relate the characteristics of the "passports" of the objects belonging to the image with the characteristics of these "beacons" and define metrics for their comparison, including fuzzy metrics for the use of NNM technologies. Lighthouses must have GPS coordinates and the corresponding associated geographical coordinates as characteristics.

Summary

The analysis demonstrates the complexity of the process of processing video streams of technical systems, which is due to many features for different stages of data processing depending on their requirements to the accuracy of the estimates and their subsequent use. It is shown that to increase the efficiency of video data processing of technical devices it is necessary to use different models and types of neural networks, as well as the distribution of processing at different stages.

 ISSN 1560-8956

REFERENCES

1. B.A. Alpatov, P.V. Babayan. Methods of image processing and analysis in on-board object detection and tracking systems. - j. "Digital signal processing", №2, 2006, p.45-51.

2. D.S. Azarenko. Detecting an object in an image and determining its displacement in two different images. - j. "Artificial Intelligence", № 3, 2013, P 90-97.

3. V.E. Guy, V.A. Utrobin, E.N. Vikulova, N.A. Nikiforov, N.N. Makarov, D.A. Lyakhmanov. Method of detection of objects on images from the positions of the theory of active perception - "Informatics and management in technical and social systems", Proceedings of NGTU. RE. Alekseeva – 2018, – № 1 (120), P 9-15.

4. V.V. Buryachenko, M.N. Favorskaya, A.I. Tomilina. Application of the fuzzy evolution classifier takagi - sugeno for problems of detection and tracking of objects on video sequence. - in "Information management systems". – 2016, №5, P 15-23; doi: 10.15217 / issn1684-8853.2016.5.15.

5. J. Redmon. The IEEE / J. Redmon [et al.]// Conference on Computer Vision and Pattern Recognition (CVPR), – 2016. – P. 779–788.

6. A.S. Victorov. Algorithm for detecting objects in photographs with low image quality. - j. Software products and systems, 2017. T. 30. № 1. S. 130–137; DOI: 10.15827 / 0236-235X.030.1.130-1377. A. Krizhevsky, G.E. Hinton. Using very deep autoencoders for content-based image retrieval. Conf. ESANN 2011, 19th Europ. Sympos. on Artificial Neural Networks, Bruges, Belgium, – 2011, P 44–51.

7. M.E. Tipping. Sparse bayesian learning and the relevance vector machine. The Jour. of Machine Learning Research. – 2001, vol. 1, P 211–244.

8. M. Shevchenko, Y. Yuzhda. Recurrent neural networks.- K., NTUU "KPI", конференція «Ukrainian and foreign science: yesterday, today, tomorrow», December 2, – 2020, P 40-42

9. M.M. Protsenko. Analysis of methods of digital video image processing by unmanned aerial vehicle equipment - Zhytomyr, VISNYK ZhDTU. 2012. № 3 (62), P 67-72.

10. A.A. Mykrukov, A.B. Babash, V.A. Sizov. Classification of events in information security systems based on neural network technologies. Open education V. 23. № 1. 2019, P 57-63; DOI: http://dx.doi.org/10.21686/1818-4243-2019-1-4-13.

11. O.L. Leshchinsky, A.O. Ishchenko. The use of neural networks in the process of intellectual (cluster) data analysis - Mukachevo: Sat. "ECONOMY AND SOCIETY", issue 11, 2017, P 578-581.

12. R.M. Timchyshyn, O.E. Volkov, O.Y. Gospodarchuk, Y.P. Bogachuk. Modern approaches to solve computer vision problems. - K. USiM, 2018, № 6, P 46-73.