DESIGNING A NEURAL NETWORK IDENTIFICATION SUBSYSTEM IN THE HARDWARE-SOFTWARE COMPLEX OF FACE RECOGNITION

Vyacheslav I. Voronov,
Moscow Technical University of Communications and Informatics, Moscow, Russia, **Vorvi@mail.ru**

Ivan A. Zharov,
Moscow Technical University of Communications and Informatics, Moscow, Russia, **303.08@mail.ru**

Aleksej D. Bykov,
Moscow Technical University of Communications and Informatics, Moscow, Russia, **79778742616@yandex.ru**

Artem S. Trunov,
Moscow Technical University of Communications and Informatics, Moscow, Russia, **greek17@yandex.ru**

Lilia I. Voronova,
Moscow Technical University of Communications and Informatics, Moscow, Russia, **Voronova.lilia@ya.ru**

With the development of information technologies the popularity of access control systems and personal identification systems is growing. One of the most common methods of access control is biometric identification. Biometric identification is more reliable than traditional identification methods, such as login/password, card, PIN-code, etc. In recent years, special attention has been paid to biometric identification based on facial recognition in access control systems, due to sufficient accuracy, scalability and a wide range of applications: face recognition of intruders in public places, providing access control, etc. The purpose of this article is to design the architecture of the subsystem "Identifier" in the hardware-software complex face recognition. In article methods and models for recognition of the face on the image and in video stream are considered. As a base method the deep neural network is chosen, the basic advantages and lacks of the chosen approach are considered. Special attention is paid to the description of architecture and scenario of work of subsystem "Identifier" of neural network identification software and hardware complex, which implements face recognition in real time from the incoming video stream of IP and USB cameras. Improvements of the traditional algorithm of face recognition using the k-neighbor method are described in detail. The results of the conducted experiments including the influence of head rotation angle on the accuracy of identification are given, and conclusions about the applicability of this method in security systems are made. On the basis of carried out researches the software and hardware complex of biometric identification on the basis of neural network recognition of faces, for the subsequent integration into the security system of the Moscow Technical University of Communications and Informatics (MTUCI) is created.

Information about authors:

Vyacheslav I. Voronov, Moscow Technical University of Communication and Informatics, Associate Professor of the department "Intelligent systems in control and automation", PhD in engineering, Moscow, Russia

BMoscow Technical University of Communication and Informatics, undergraduate, Moscow, Russia

Aleksej D. Bykov, Moscow Technical University of Communication and Informatics, undergraduate, Moscow, Russia

Artem S. Trunov, Moscow Technical University of Communication and Informatics, senior teacher, Moscow, Russia

Lilia I. Voronova, Moscow Technical University of Communication and Informatics, head of the department "Intelligent systems in control and automation", D.Sc. in Physical and Mathematical Sciences, Moscow, Russia

I. INTRODUCTION

Modern biometric identification systems are designed to work under rather difficult conditions. The number of objects per unit of time to be identified is measured in tens and hundreds, and the cost of computing resources and information storage facilities, depending on the type of biometric data used, can be very significant. That's why the choice of identification methods and the task of designing an optimal system architecture is an extremely important and urgent task.

The task of biometric identification of a person can be solved in several ways. For identification can be used a number of static and dynamic characteristics such as: papillary finger pattern, hand geometry, iris, face geometry (2D and 3D), vein pattern, handwriting, silhouette, gait, voice.

The number of video surveillance systems in public places is constantly increasing. Maintenance of video surveillance systems is expensive and requires significant labor costs, due to the need to analyze the video by people, with human capacity to quickly process large amounts of information is limited. The shift to video analytics with neural network face recognition reduces the cost of maintaining these systems, and allows for operations such as detection, tracking and identification in real time. All this has led to the development of methods for effective biometric identification of people using video and photo images, with the vast majority of methods focused on facial recognition.

Biometric identification is the presentation by a user of his or her unique biometric parameter and the process of comparing it with a database.

Video analytics is a technology that uses computer vision methods to automatically obtain various data based on the analysis of the sequence of images coming from video cameras in real time or from archive records. Video analytics allows to help a person or replace him, in cases of multilateral analysis of the situation and significantly reduces the negative impact of human factor.

The world's largest companies are developing their own software solutions that implement face recognition functionality in photos and video streaming. Such giants as Amazon (USA), Facebook (USA), Apple (USA), as well as start-ups NTechLab (Russia), Macroscope (Russia) have implemented intelligent analysis of video stream and offer the market ready-made competitive solutions.

Face recognition system operation includes several stages: face detection, alignment, localization, normalization, facial features selection and matching.

For several years, the MTUCI Department «Intelligent systems in control and automation» has been working on the use of artificial intelligence in various fields. In particular, neural networks and associated methods were used in the development of Smart City and Industry 4.0 tools [1], forecasting the state of hydraulic systems [2], analysis of environmental pollution [3], imitation and recognition of sign language [4, 5, 6], prediction of the likelihood of bronchial asthma [7].

In Moscow Technical University of Communications and Informatics (MTUCI) in the framework of the grant "Development of a software complex of biometric identification based on face recognition for the security system of the university with the use of neural network methods and modern software solutions" was developed a prototype of software-hardware complex of biometric identification (SHC BI) using modern methods of computer vision and neural network methods of face recognition with the possibility of subsequent integration into the security system of the university. The created prototype of SHC BI allows to provide on the territory of the university (for any set of premises) access control and monitoring of movement of people.

II. METHODS AND MODELS OF NEURAL NETWORKS USED FOR FACE IDENTIFICATION

There are various methods for face recognition: the method of deep comparison on graphs, deep neural networks, hidden Markov models, methods of the main components, etc. Nowadays deep neural networks are the most widespread for face recognition because of less computational complexity of recognition procedure, higher accuracy of algorithms and absence of necessity to select parameters for each data set [8]. Neural network models are constantly being developed and upgraded, and the probability of error in these models is significantly reduced when recognizing the frontal face on a still image.

The largest companies in the world are developing their own neural network architectures. Companies such as Facebook (USA), Microsoft (USA) have implemented neural networks that identify people on test data sets with great accuracy. The main neural network architectures used for face recognition in images are discussed below. Deep neural networks have a drawback; the degree of model accuracy depends on the number of images-views in the database used for training. Training sets are publicly available, have poor markup and are poorly structured. These networks have been selected because of the open access architectures and pre-trained models.

A convolution neural network is a type of neural network architecture in which at least one layer is a convolution layer. Usually a convolution neural network consists of combinations of convolution, fully connected and pooling layers [9, 10]. The convolution layer at the input takes a brightness matrix of the image characteristics, and with the help of the convolution core (a matrix smaller than the input matrix and having weight coefficients that are set during training) it passes along the input matrix summing up the elements of two matrices. The result of the convolution is transferred to the activation function, which generates and transmits the output nonlinear value. The pooling layer implements the compaction of the output matrix of the convolution layer by applying the maximum or average value over the combined area. A fully connected layer is a layer where each node connects to each node in a subsequent layer.

DeepFace [11] is a deep neural network with over 120 million parameters. This algorithm uses several locally linked layers without weight distribution, rather than standard convolution layers. DeepFace uses the ReLU activation function expressed by a formula:

$$f(x) = \begin{cases} x, x > 0 \\ 0.01x, 0 \le x \end{cases} \qquad (1)$$

where x is the input value.

DeepFace uses image pre-processing: changing the angle of the face so that the face is facing the camera. The DeepFace network uses 2D and 3D normalization. Normalized feature vector, which is similar to Local Binary Patterns (LBP). LBP is an efficient operator that represents each pixel of an image as a binary

number depending on the intensity of neighboring image pixels. Linear SVM applied to element vectors is used to select weighting parameters.

The method of reference vectors (SVM) refers to linear qualifiers. The main idea of the algorithm is to build a decisive boundary by drawing a straight line through the middle of a segment connecting the mass centers of positive and negative examples.

At the input, the neural network receives a face image with the size of 152 by 152 pixels. The network has two convolution layers with 32 and 16 filters of $11 \times 11 \times 3$ and $9 \times 9 \times 16$, a pulling layer that takes a maximum of more than 3×3 spatial neighborhoods in steps 2, separately for each channel. The pulling layer is a non-linear compaction of the feature matrix. Subsequent layers are linked locally, like a convolution layer, they apply a filter set, but each area in the feature map uses a different filter set. The output of the first fully linked layer is used as a vector feature of the facial representation.

DeepID [12] refers to a type of convolution neural networks, which by means of a set of high-level representations of objects, called hidden identification functions, is used to identify a person. DeepID contains four convolution layers with a pooling, for hierarchical extraction of objects, followed by a fully linked DeepID layer and an output layer using softmax activation function, indicating the identity class is mathematically described by (2).

$$\sigma(x_j) = \frac{e^{x_j}}{\sum_i e^{x_i}} \tag{2}$$

where $\sigma(x_j)$ – output value of the layer, e – neuron of a softmax group, i – class number.

The convolution operation is expressed by as:

$$y^{j(r)} = \max\left(0, b^{j(r)} + \sum_i k^{ij(r)} * x^{i(r)}\right) \tag{3}$$

where x^i – i-th input Signature Map, y^j – i-th exit sign map.

Input data are $39 \times 31 \times k$ for rectangular images and $31 \times 31 \times k$ for square images, where k = 3 for color images and k = 1 for gray images.

FaceNet [13] does not use 2D and 3D alignment as opposed to DeepFace. FaceNet uses a special loss function called TripletLoss.

$$Loss = \sum_{i=1}^{N} \left[\left\| f_i^a - f_i^p \right\|_2^2 - \left\| f_i^a - f_i^n \right\|_2^2 + a \right]_+ \tag{4}$$

where f_i^a – anchor encoding; f_i^p – encoding of similar faces (positive); f_i^n – encoding of dissimilar faces (negative); a – constant.

It minimizes the distance between a set of features and images that contain similar looks, and maximizes the distance between images that contain different looks.

The FaceNet architecture is based on the Zeiler&Fergus model [14], this model is a convolution neural network with alternating convolution layers, layers of nonlinear activation of ReLU (1), local normalization and pooling layers. FaceNet uses additional layers of 1×1×d convolution. This network uses mixed layers that combine several different convolution and pooling layers in parallel and combine their responses.

FaceNet directly explores the comparison of facial images with compact Euclidean space, where distances correspond to the similarity of the face.
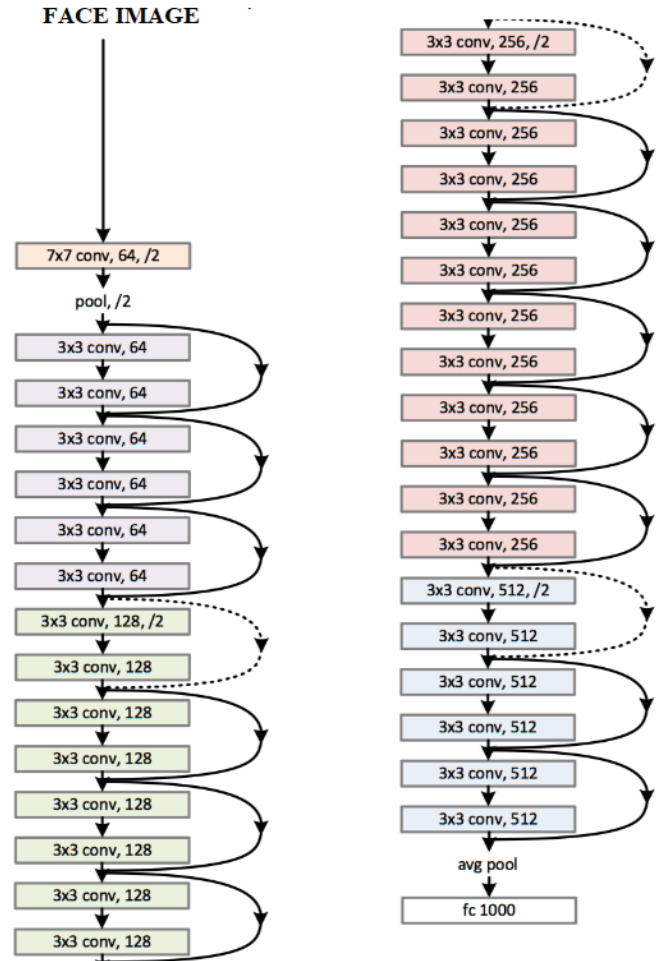

Figure 1. An example network architecture for Microsoft ResNet with 34 layers (3.6 billion FLOP)

Microsoft ResNet [15] is a neural network implementing the residual learning method. This neural network is built on the basis of VGG Net neural network.

In VGG Net network refused to use filters larger than 3×3, a layer with a filter 7×7 is equivalent to three layers with filters 3×3, and in the latter case is used by 55% less parameters with ReLU activation function (1). This network does not have the best accuracy, but due to simplicity it is used in more complex neural networks. Increasing the number of layers in a convolution neural network does not allow increasing accuracy. As the number of layers increases, a convolution neural network may begin to degrade – Its accuracy decreases on the validation set. Since the accuracy also decreases on the validation set, we can conclude that the problem is not in retraining the network.

ResNet has suggested that if the convolution neural network has reached its accuracy limit on some layer, all subsequent layers will have to degenerate into an identical transformation, but because of the difficulty of learning deep networks, this is not the case. To improve network learning, it has been suggested that we introduce Shortcut Connections. The Shortcut Connections layer allows you to skip the signal further without changes.

The Shortcut Connections layer is mathematically described in (5).

$$y = F(x, \{W_i\}) + x \tag{5}$$

where x, y – input and output vectors of the layer, $F(x, \{W_i\}) -$ function of residual display.

This architecture allows you to make a deeper neural network that will not degrade. Figure 1 shows the Microsoft ResNet architecture.

In the works of the authors [11][12][13][15] testing was carried out on the publicly available test data sets for facial identification (Label Faces in the Wild [16], YouTube Faces Database [17]) presented in Table 1. LFW is a data set that contains 13000 facial images. Each face is marked with the name of the person depicted. LFM contains data from 1680 classes. YTF is a data set that contains 3425 videos of 1595 different people. On average, there are 2.15 videos for each class.

To evaluate the quality of ID training, the metric accuracy was used - the proportion of sampling for which the ID made the right decision. Accuracy is considered by the formula:

$$Accuracy_c = \frac{TP + TN}{TP + FP + FN + TN} \tag{6}$$

where $Accuracy_c$ – class c-validity; TN-true-negative solution; TP-true-positive solution; FP-false positive solution; FN-false-negative solution.

Table 1

Comparison of facial identification algorithms

Algorithm name	LFW (Label Faces in the Wild) [9]	YTF (YouTube Faces Database) [10]
DeepFace		91,4%
DeepID		-
DeepID2		93,2%
FaceNet	99,63%	95,1%
Microsoft ResNet	99,88%	-

The U.S. National Institute of Standards and Technology (NIST) publishes monthly reports on neural network tests for face recognition [18]. NIST measures the accuracy, speed, memory capacity of automatic face recognition technologies used in a wide range of civil, law enforcement, and domestic applications. At present, the NEC-2 algorithm [19] shows the smallest error in facial identification. In Japan, NEC created the NEC-2 algorithm in 2018. NEC-2 uses technology to highlight the most characteristic facial points, which provides a higher level of human identification and learning with memory organization. With the help of memory organization, the algorithm allows to accelerate the updating of scales of stochastic gradient descent to the whole network. This algorithm is closed. The NEC-2 algorithm has a false identification probability of 0.26% when searching in a database of 640,000 patterns and only 0.31% when searching in a database of 12 million patterns.

The algorithms considered fall behind the described algorithm NEC-2, but this algorithm is closed, which does not allow us to use it when designing the identification subsystem. After studying the available algorithms and test results, to solve the task, we chose a neural network with Microsoft resNet architecture pre-trained on a large data set.

III. HARDWARE AND SOFTWARE COMPLEX FOR BIOMETRIC IDENTIFICATION

SHC of biometric identification [23] using neural network face recognition is the server software presented in Figure 2, which consists of a subsystem of face detection, identification subsystem, database (database) and web interface. The system software is implemented on the Python 3 platform. For realization libraries dlib [20], scikit-learn [21], opencv [22] were used.
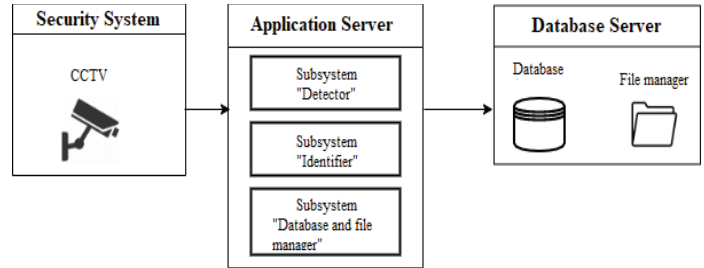

Figure 2. Biometric identification hardware-software complex architecture [23]

The system consists of the following subsystems:
- video surveillance system;
- the "Detector" subsystem;
- subsystem "Identifier";
- subsystem "Database and file manager";
- database;
- file storage.

SHC biometric identification was designed for integration into the university security system, which requires significant computing power on computing machines (CM). The architecture of the complex biometric identification allows you to run subsystems on different computers, distributing the load from the VM, which reduces the functional requirements for the performance of the VM.

IV. «IDENTIFIER» SUBSYSTEM

The subsystem "Identifier" realizes reception of the image with the detectable person, repeated detection of faces of "small" images for neutralization of error of detection and neural network identification of the person on a face.

The diagram of classes of the subsystem "Identifier" is shown in figure 3:
- main – the main class in which there is an initiation of initial parameters and classes, as well as the launch of the subsystem itself;
- TrainModel – the class conducting training of the classifier;
- FaceRecognition – the class, which implements the repeated detection and identification by means of the previously trained model;
- ReTrain – the class, implements the retraining of the classifier, if there are new data.

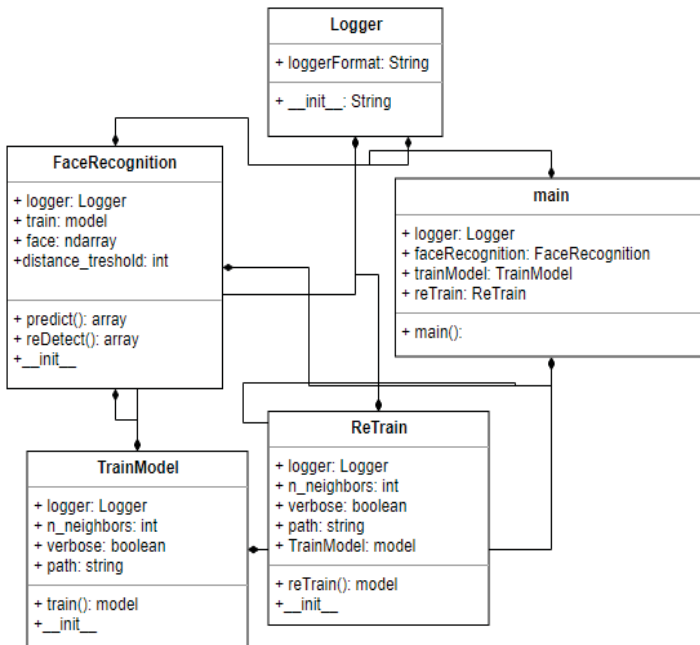- Logger – the class, records all the data generated by the subsystem.

Figure 3. The diagram of classes of the subsystem "Identifier"

[-0.12616129   0.00791075   0.02922434   -0.06563014   -0.12112911
-0.03997219   -0.02698989   -0.07678688   0.1261778   -0.07740152
0.26403227   0.02185132   -0.23205237   -0.04792541   0.03580608
0.15406483   -0.08407264   -0.13449208   -0.0907556   -0.09695972
0.10336196   0.01695136   -0.04402567   0.07327586   -0.21195126
-0.2377865 -0.12497076 -0.05062533 0.03289974 -0.02835255
0.01159448   0.01666744   -0.133228   -0.04311804   0.07258895
0.07394607   -0.06391457   -0.07646051   0.16769679   -0.00151986
-0.18195102   0.0316122   0.05988811   0.23168679   0.17555423
0.06508835   0.0909338   -0.06405875   0.02826674   -0.29183176
0.10869311   **0.15089822**
-0.19860986 -0.00483399 0.26444244 -0.2909506 0.15632941
0.05403928   -0.11835575   -0.04728038   -0.06832737   0.1933302
0.18850563   -0.21718842   -0.11698081   0.23235707   -0.18203288
-0.08930888   -0.01170634   -0.15678045   -0.26796865   -0.36144692
0.07562776   0.41509724   0.17437397   -0.22592078   0.00610069
0.00445438   0.05961983   0.0363998   0.09337185   -0.06003059
-0.14371039 -0.08897015 -0.02648188 0.19062081 0.01273332
-0.00051378   0.25421643   0.04770367   -0.03287114   0.02315203
0.13193734   -0.08855681   -0.08092593   -0.06442997   0.02985705
-0.08262224   -0.13092032   0.0301197   0.15032028   -0.14215843
0.2182014   -0.03664138   -0.03492382   -0.08867846   0.04167698
-0.07619464   -0.06421942   0.12646358   -0.21660684   0.18788326
0.25281885 0.08210047 0.18551196 0.07995273 0.11073787
0.04247311   -0.00351085   -0.19471884   -0.08157935   -0.03292821
-0.1122886 0.04450573 0.11372029]

Figure 4. Composition of facial features vector

SHC BI has a microservice architecture, the subsystem "Identifier" is a microservice. The use of microservice architecture simplifies project scalability and development. Each subsystem is considered a separate unrelated program and can be modified depending on the goals and objectives.

- Scenario of work of the subsystem "Identifier":
- The Identifier waits for the message from the subsystem "Detector". The message contains: bit array, which represents the image, time of detecting, unique identifier of image, unique identifier of camera.
- After receiving the message, in order to neutralize the detection error, the face on the image is re-detected using the directional gradient histogram (HOG) [24].
- The detected face is transferred to the input of the identification function. This function searches for 128 key points on the face.

Biometric identification consists in finding of euclidean distance between vectors of key points. With the help of the Euclidean metric of distance between vectors and on its basis refers the face to one of the specified classes. The Euclidean distance is calculated by the formula:

$$d(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \qquad (7)$$

where x, y – face descriptors (key points), n – dimension of space.

The model with resNet architecture is implemented in dlib libraries and used to calculate facial features vectors. In the facial feature vector 128 key points are used. Key points represent normalized numbers from -1 to 1 which characterize features of a face. Figure 4 shows the content of one of the composed vectors of facial features.

The use of SHC in the university's security system provides for a large database of students, and the university has dozens of entry points, so the system must be able to handle a large data stream from all cameras. Because of the described requirements to HAC in the university security system classical methods are not applicable, the speed and accuracy of identification is important. To increase the speed and accuracy of identification in the implemented system the classifier of k-neighbors, implemented in Sklearn library, is used. K-neighbourhood is a metric algorithm used for classification and regression. The object is assigned a class that is most common among its k neighbors.

On the basis of testing, a neural network with Microsoft resNet architecture, pre-trained on a large data set, was chosen to solve this problem.

When teaching the classifiers, the method of teaching "with the teacher" is used. When teaching "with teacher" there are objects (data) and true answer (data type). During classification it is necessary to restore general dependence on object-response pairs, build algorithm, which will predict answers by objects. To determine the responses by objects in training, the characteristic description of objects (attributes) is set.

The application of the classifier allows comparing the key points of the test image with several images from the database, which increases the time required for database iteration during identification and improves the accuracy of identification by using weighted voting. At the weighted voting the neighbors who have the minimum euclidean distance are defined and by means of quantity and distance the relation to a class is defined. The votes are determined by the following (8):

$$votes(class) = \sum_{i=1}^{n} \frac{1}{d^2(XY_i)} \qquad (8)$$

where $d^2(XY_i)$ – quare of the distance from the known record Y_i to new X, n – number of known class records, class – name of the class.

The knn classifier first learns on a set of marked (known) faces, and then predicts a person on an unknown image. Having found k most similar faces (focusing on euclidean distance of key points of a face) in a training set and having made comparison on the basis of "voices" with faces by means of function kneighbors received from a video stream, the given algorithm predicts the person. This algorithm allows using several photos for training.

Before the system starts working, the knn classifier on the training dataset is trained, which is stored locally on the server.

To determine the optimal identification threshold (distance euclidean) was tested at different values of distance euclidean. Figure 5 shows a graph showing the dependence of identification accuracy on the distance euclidean.
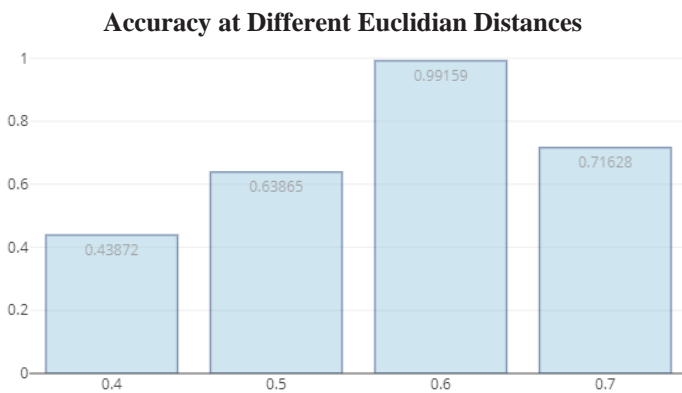
Accuracy at Different Euclidian Distances

Figure 5. Diagram of the dependence of identification accuracy on the Euclidean distance

As a result of testing, it was found that the euclidean distance 0.6 is the optimal value for identification.

Table 2 shows the response format of the "Identifier" subsystem that contains the identification results.

Table 2

Subsystem output format identifier

Field name	Type	Example
id	Stri	c073832e08214867af912bc9
datetime		2019-11-30 12-30-04
camId		1
prediction	String	Zharov

Conclusion

In the article the analysis of methods and models of application of neural networks for face recognition on the image is conducted. The architecture of the neural network identification subsystem "Identifier" in the hardware-software complex on the basis of computer vision is designed, which provides accelerated recognition of faces from a video stream due to the integration of methods of re-detection of "small" images, neural network modeling and kNN.

The UML class diagram and the script of the "Identifier" subsystem operation are given. According to the test results, the accuracy of recognition subsystem operation on the test data set for the metric "accuracy" is 0.99159.

References

1. Voronova L.I., Bezumnov D.N., Voronov V.I. (2019). Development of the Research Stand «Smart City Systems» INDUSTRY 4.0. *Proceedings of the 2019 IEEE International Conference&Quality Management, Transport and Information Security, Information Technologies; IT and QM and IS.*
2. Bykov A.D., Voronov V.I., Voronova L.I. (2019). Machine Learning Methods Applying for Hydraulic System States Classification. *2019 Systems of Signals Generating and Processing in the Field of on Board Communications, SOSG 2019.*
3. Usachev V.A., Voronova L.I., Voronov, V.I., Zharov I.A., Strelnikov V.G. (2019). Neural Network Using to Analyze the Results of Environmental Monitoring of Water. *2019 Systems of Signals Generating and Processing in the Field of on Board Communications, SOSG 2019.*
4. Goncharenko A., Voronova L., Artemov M., Voronov V., Bezumnov D. (2019). Sign language recognition information system development using wireless technologies for people with hearing impairments. *Conference of Open Innovation Association, FRUCT 2019.*
5. Ezhov A.A., Voronova L.I., Voronov V.I., Goncharenko A.A., Artemov M.D. (2018). Program to support non-verbal communication, using machine learning based on a convolution neural network. Certificate of registration for a computer program RU 2019610179, 09.01.2019. Request № 2018664377 from 13.12.2018. (*in Russian*)
6. Goncharenko A.A., Voronova L.I., Voronov V.I., Ezhov A.A., Artemov M.D. (2019). Software package for data management in the information and communication system of social accessibility for people with hearing disabilities. Certificate of registration for a computer program RU 2019610962, 18.01.2019. Request № 2018665275 from 26.12.2018. (*in Russian*)
7. Bashirov A.N., Voronov V.I. (2019). Prediction of the probability of bronchial asthma in children using a random forest algorithm. *Modern high technologies*. No. 12-2. P. 249-255. (*in Russian*)
8. Mishhenkova E.S. (2013). Comparative analysis of facial recognition algorithms. *Science Jornal of Volgograd State University: Young Researchers' Work*. No. 11. (*in Russian*)
9. Ciresan, Dan, Ueli Meier, Jonathan Masci, Luca M. Gambardella, Jurgen Schmidhuber. (2013). Flexible, High Performance Convolutional Neural Networks for Image Classification. *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*. Vol. 2. P. 1237-1242.
10. Voronova L.I. (2013). *Intellectual databases: a training manual*. MTUCI. 35 p. (*in Russian*)
11. Kaiming He, Xiangyu Zhang, Shaoqing Ren Jian Sun (2015). Deep Residual Learning for Image Recognition. 2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*
12. Y. Taigman, M. Yang, M. Ranzato and L. Wolf. (2014). DeepFace: Closing the Gap to Human-Level Performance in Face Verification. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH. P. 1701-1708.
13. Sun Y., Wang X., & Tang X. (2014). Deep Learning Face Representation from Predicting 10,000 Classes. *2014 IEEE Conference on Computer Vision and Pattern Recognition*. P. 1891-1898.

14. F. Schroff, D. Kalenichenko and J. Philbin (2015). FaceNet: A unified embedding for face recognition and clustering. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* P. 815-823.

15. M.D. Zeiler and R. Fergus. (2013). Visualizing and understanding convolutional networks. CoRR, abs/1311.2901.

16. http://vis-www.cs.umass.edu/lfw.

17. https://www.cs.tau.ac.il/~wolf/ytfaces.

18. Patrick Grother, Mei Ngan, Kayee Hanaoka (2019). Ongoing Face Recognition Vendor Test (FRVT) Part 1: Verification. *Information Access Division Information Technology Laboratory.*

19. Alexander Pritzel, Benigno Uria, Sriram Srinivasan, Adrià Puigdomènech, Oriol Vinyals, Demis Hassabis, Daan Wierstra, Charles Blundell (2017). *Neural Episodic Control.*

20. Dlib d entation // dli et UR http://dlib.net/python/index.html.

21. Scikit-learn documentation // scikit-learn.org URL: http://scikit-learn.org/stable/user_guide.html.

22. Opencv documentation // opencv.org URL: https://docs.opencv.org.

23. Voronov V.I., Voronova L.I., Bykov A.A., Zharov I.A. (2019). Software Complex of Biometric Identification Based on Neural Network Face Recognition. *Proceedings of the 2019 IEEE International Conference & Quality Management, Transport and Information Security, Information Technologies; IT and QM and IS 2019* 8928297. P. 442-446.

24. Dalal N., Triggs B. (2005). Histograms of oriented gradients for human detection. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.* Vol. 1. P. 886-893.

ПРОЕКТИРОВАНИЕ ПОДСИСТЕМЫ НЕЙРОСЕТЕВОЙ ИДЕНТИФИКАЦИИ В ПРОГРАММНО-АППАРАТНОМ КОМПЛЕКСЕ РАСПОЗНАВАНИЯ ЛИЦ

Воронов Вячеслав Игоревич, *Московский Технический Университет Связи и Информатики, Россия, Москва, Vorvi@mail.ru*
Жаров Иван Александрович, *Московский Технический Университет Связи и Информатики, Россия, Москва, 303.08@mail.ru*
Быков Алексей Денисович, *Московский Технический Университет Связи и Информатики, Россия, Москва, 79778742616@yandex.ru*
Трунов Артем Сергеевич, *Московский Технический Университет Связи и Информатики, Россия, Москва, greek17@yandex.ru*
Воронова Лилия Ивановна, *Московский Технический Университет Связи и Информатики, Россия, Москва, Voronova.lilia@yandex.ru*

Аннотация

С развитием информационных технологий растет популярность систем контроля доступа и систем идентификации личности. Одним из распространенных способов контроля доступа является биометрическая идентификация. Биометрическая идентификация является более надежной, чем традиционные способы идентификации, такие как логин/пароль, карта, ПИН-код и др. В последние годы особое внимание уделяется биометрической идентификации на основе распознавания лиц в системах контроля доступа из-за высокой точности, возможностей масштабирования и широкой области применения: распознавание лиц злоумышленников в общественных местах, обеспечение контроля управления доступа и др. Целью данной статьи является проектирование архитектуры подсистемы "Идентификатор" в программно-аппаратном комплексе распознавания лиц. Рассматриваются методы и модели для распознавания лица на изображении и в видеопотоке. В качестве базового метода выбрана нейронная сеть глубокого обучения, рассмотрены основные достоинства и недостатки выбранного подхода. Особое внимание уделено описанию архитектуры и сценария работы подсистемы "Идентификатор" программно-аппаратного комплекса нейросетевой идентификации, реализующей распознавание лиц в реальном времени с входящего видеопотока IP и USB камер. Подробно описаны улучшения традиционного алгоритма распознавания лиц с помощью метода k-ближайших соседей. Приведены результаты проведенных экспериментов, включающие влияние угла поворота головы на точность идентификации, сделаны выводы о применимости данного метода в системах безопасности. На основе проведенных исследований создается программно-аппаратный комплекс биометрической идентификации на основе нейросетевого распознавания лиц, для возможной последующей интеграции в систему безопасности Московского Технического Университета Связи и Информатики (МТУСИ).

Ключевые слова: идентификация по лицу, биометрическая идентификация, нейросетевое моделирование, нейронная сеть, Microsoft ResNet, метод kNN, система биометрической идентификации, системы контроля доступа.

Литература

1. *Voronova L.I., Bezumnov, D.N., Voronov, V.I.* Development of the Research Stand "Smart City Systems" INDUSTRY 4.0 // Proceedings of the 2019 IEEE International Conference&Quality Management, Transport and Information Security, Information Technologies; IT and QM and IS 2019.

2. *Bykov, A.D., Voronov, V.I., Voronova, L.I.* Machine Learning Methods Applying for Hydraulic System States Classification // 2019 Systems of Signals Generating and Processing in the Field of on Board Communications, SOSG 2019.

3. *Usachev, V.A., Voronova, L.I., Voronov, V.I., Zharov, I.A., Strelnikov, V.G.* Neural Network Using to Analyze the Results of Environmental Monitoring of Water // 2019 Systems of Signals Generating and Processing in the Field of on Board Communications, SOSG 2019.

4. *Goncharenko, A., Voronova, L., Artemov, M., Voronov, V., Bezumnov, D.* Sign language recognition information system development using wireless technologies for people with hearing impairments // Conference of Open Innovation Association, FRUCT 2019.

5. *Ezhov A.A., Voronova L.I., Voronov V.I., Goncharenko A.A., Artemov M.D.* Program to support non-verbal communication, using machine learning based on a convolution neural network. Certificate of registration for a computer program RU 2019610179, 09.01.2019. Request № 2018664377 from 13.12.2018.

6. *Goncharenko A.A., Voronova L.I., Voronov V.I., Ezhov A.A., Artemov M.D.* Software package for data management in the information and communication system of social accessibility for people with hearing disabilities. Certificate of registration for a computer program RU 2019610962, 18.01.2019. Request № 2018665275 from 26.12.2018.

7. *Bashirov A.N., Voronov V.I.* Prediction of the probability of bronchial asthma in children using a random forest algorithm. Modern high technologies. 2019. № 12-2. С. 249-255.

8. *Мищенкова Е.С.* Сравнительный анализ алгоритмов распознавания лиц // Вестник ВолГУ. Серия 9: Исследования молодых ученых. 2013. №11.

9. *Ciresan, Dan, Ueli Meier, Jonathan Masci, Luca M.* Gambardella, Jurgen Schmidhuber. "Flexible, High Performance Convolutional Neural Networks for Image Classification". *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence. 2013.* Vol. 2, pp. 1237-1242.

10. *Воронова Л.И.* Интеллектуальные базы данных: учебное пособие. М.: МТУСИ, 2013. 35 с.

11. *Kaiming He, Xiangyu Zhang, Shaoqing Ren Jian Sun.* Deep Residual Learning for Image Recognition // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015.

12. *Y. Taigman, M. Yang, M. Ranzato and L. Wolf.* DeepFace: Closing the Gap to Human-Level Performance in Face Verification // 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, 2014, pp. 1701-1708.

13. *Sun, Y., Wang, X., & Tang, X.* Deep Learning Face Representation from Predicting 10,000 Classes // 2014 IEEE Conference on Computer Vision and Pattern Recognition. 2014, pp. 1891-1898.

14. *F. Schroff, D. Kalenichenko and J. Philbin.* FaceNet: A unified embedding for face recognition and clustering // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015. pp. 815-823.

15. *M.D. Zeiler and R. Fergus.* Visualizing and understanding convolutional networks. CoRR, abs/1311.2901, 2013.

16. http://vis-www.cs.umass.edu/lfw.

17. https://www.cs.tau.ac.il/~wolf/ytfaces.

18. *Patrick Grother, Mei Ngan, Kayee Hanaoka.* Ongoing Face Recognition Vendor Test (FRVT) Part 1: Verification // Information Access Division Information Technology Laboratory, 2019.

19. *Alexander Pritzel, Benigno Uria, Sriram Srinivasan, Adri Puigdomnech, Oriol Vinyals, Demis Hassabis, Daan Wierstra, Charles Blundell.* Neural Episodic Control. 2017.

20. Dlib documentation // dlib.net URL: http://dlib.net/python/index.html.

21. Scikit-learn documentation // scikit-learn.org URL: http://scikit-learn.org/stable/user_guide.html.

22. Opencv documentation // opencv.org URL: https://docs.opencv.org.

23. *Voronov V.I., Voronova L.I., Bykov A.A., Zharov I.A.* Software Complex of Biometric Identification Based on Neural Network Face Recognition // Proceedings of the 2019 IEEE International Conference & Quality Management, Transport and Information Security, Information Technologies; IT and QM and IS 2019 8928297. C. 442-446.

24. *Dalal N., Triggs B.* Histograms of oriented gradients for human detection // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2005. Vol. 1, pp. 886-893.

Информация об авторах:

Воронов Вячеслав Игоревич, *Московский Технический Университет Связи и Информатики, доцент каф. ИСУиА, к.т.н., Москва, Россия*

Жаров Иван Александрович, *Московский Технический Университет Связи и Информатики, магистрант, Москва, Россия*

Быков Алексей Денисович, *Московский Технический Университет Связи и Информатики, магистрант, Москва, Россия*

Трунов Артем Сергеевич, *Московский Технический Университет Связи и Информатики, старший преподаватель, Москва, Россия*

Воронова Лилия Ивановна, *Московский Технический Университет Связи и Информатики, зав. каф. ИСУиА, профессор, д.ф.-м.н., Москва, Россия*