International Journal of Civil Engineering and Technology (IJCIET)
Volume 10, Issue 03, March 2019, pp. 1424-1434, Article ID: IJCIET_10_03_139
Available online at http://www.iaeme.com/ijciet/issues.asp?JType=IJCIET&VType=10&IType=03
ISSN Print: 0976-6308 and ISSN Online: 0976-6316

RECOGNITION OF EMOTIONS BY FACIAL GEOMETRY USING A CAPSULE NEURAL NETWORK

Liudmyla Tereikovska
Kyiv National University of Construction and Architecture, Kyiv, Ukraine

Ihor Tereikovskyi
Faculty of Applied Mathematics, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

Shynar Mussiraliyeva, Gulmaral Akhmed, Aiman Beketova and Aizhan Sambetbayeva
Department of Information Systems, Al-Farabi Kazakh National University, Almaty, Kazakhstan

ABSTRACT

The article is devoted to the problem of improving the efficiency of neural network means of emotion recognition by the geometry of the human face. It is shown that one of the most significant drawbacks of modern neural network means of emotion recognition, which are used in General-purpose information systems, is the lack of recognition accuracy under the influence of characteristic interference. It is proposed to improve the accuracy of recognition through the use of capsule neural network model, which has increased adaptability to the analysis of noisy images. As a result of the research, a neural network model of the CapsNet type was developed, designed to recognize basic emotions taking into account such interference as face rotation. It is shown experimentally that in the analysis of undistorted images CapsNet slightly exceeds the accuracy of the classical convolutional neural network type LaNet, which is approximately equal to its resource intensity. The accuracy of CapsNet recognition of undistorted images is somewhat inferior to modern types of convolution networks, which have a much higher resource consumption compared to it. When detecting emotions on rotated images, the accuracy of CapsNet is comparable with the accuracy of modern types of convolution networks and significantly exceeds the accuracy of LaNet. Prospects for further research in the field of neural network recognition of emotions on the geometry of the face can be associated with the improvement of architectural solutions of the capsule neural network in the direction of reducing the number of training iterations while ensuring acceptable recognition accuracy.

Keywords: neural network model, recognition of emotions, a capsule of neural network, face geometry, convolutional neural network.

Cite this Article: Liudmyla Tereikovska, Ihor Tereikovskyi, Shynar Mussiraliyeva, Gulmaral Akhmed, Aiman Beketova and Aizhan Sambetbayeva, Recognition of Emotions By Facial Geometry Using a Capsule Neural Network, International Journal of Civil Engineering and Technology, 10(3), 2019,pp. 1424-1434

http://www.iaeme.com/IJCIET/issues.asp?JType=IJCIET&VType=10&IType=03

1. INTRODUCTION

Currently, the means of automatic recognition of human emotions are widely used in various fields of human activity: medicine, entertainment industry, smart home control systems, cybersecurity, and distance learning systems. The basis of recognition is the analysis of biometric characteristics, which clearly reflects the emotional state of a person. The analysis of the market of modern means of recognition of an emotional state allows to claim that in information systems of the General purpose the means which are based on the analysis of the parameters describing geometry of the person's face have the greatest distribution. It should be noted that the face is the front part of the human head, from above limited by the border of the scalp, below - the corners and the lower edge of the lower jaw, with the sides - the edges of the branches of the lower jaw and the base of the auricles [4, 5].

The advantages of facial geometry recognition tools include ease of use, high classification accuracy, good testing, low cost and prevalence of reading devices (video cameras). The last advantage predetermines high universality of use of such means in information systems of different function. At the same time, practical experience, as well as the results of scientific and applied works [1, 4, 5, 8-10] indicate the need for significant modernization of modern recognition tools in the direction of reducing resource intensity, increasing recognition accuracy, reducing the development time and increasing adaptation to many features of modern information systems. Along with the use of more efficient hardware, one of the main directions of modernization is to improve the mathematical support of the recognition process [11, 12], which determines the relevance of research in this direction.

2. ANALISYS OF EXISTING PUBLICATIONS

The basis for the analysis of the current state of developments in the field of recognition were works [2-5, 8-12, 19-21] which describe both tested solutions and modern approaches in this direction. As a result of the first stage of the analysis, a general characteristic of the known recognition methods is obtained. It is determined that from the position of distinguishing the features of the analyzed image, the known methods of emotion recognition can be divided into holistic and local. Holistic-emotions are defined based on the whole image. Local-emotions are determined by a set of individual control points or individual parts of the face. The methods based on the holistic approach include the eigenvector method, the optical flow analysis method, and graph matching methods. The local approach is based on the methods of analysis of the set of anthropometric points of the face, methods of probabilistic evaluation, methods of comparison with the standard, which can use elements of fuzzy logic. In this case, neural network solutions are used in both approaches. From the standpoint of taking into account the dynamics of the image of the face in time, these approaches are divided into static and temporal. Temporal-emotions are determined based on the dynamics of the image of a person's face. The static approach does not take into account dynamics. It should be noted that the use of a temporal approach requires continuous video monitoring of the individual. Therefore, a static approach is now more common in general-purpose information systems. At the same time, many studies indicate that the effectiveness of recognition depends largely on the compliance of the approach on which the method is based and the conditions of the task of recognition. It should also be noted that for the period 2016-2018 published several very complete databases, which contain images of people's faces with different emotions. The most famous are the bases: OMG-Emotion challenge, EmotiW challenge, AffectNet, AFEW-VA, EmotioNet challenge, EmoReact, Cohn-Kanade. In these databases, the number of emotions represented varies from 6 to 17. The emergence of these databases, on the one hand, signals great progress in the field of emotion recognition tools, and on the other hand, predetermines the possibility of improving such tools.

At the second stage of the analysis the features of modern solutions are considered. So in [4] the variant of realization of system of recognition of an emotional state for support of communication of the person with service anthropomorphic robots is offered. It is indicated that a significant obstacle in the development of emotion recognition systems is the limited availability of databases, as well as a high proportion of individual characteristics in the manifestation of an emotion in different people. This significantly increases the requirements for the generalizing capabilities of the applied machine learning methods. In addition, the accuracy of the recognition is significantly affected by the change in the position of the face in the image, the presence of glasses, makeup or a covering eyebrow hairstyle. To compensate for these difficulties it is proposed to use various algorithms local filtering of the image when determining informative features of the face, and the assessment of the severity of the emotions to count with the help of multiclassloader. The system developed in [7] allows to recognize 7 basic emotions on the basis of filtered local features of the degree of expression of 20 motor units of the face (Action Units, AU) included in the system of coding facial movements (Facial Action Coding System, FACS), developed by P. Ekman. Basic emotions for classification are also chosen according to FACS: joy, anger, sadness, disgust, fear, surprise. The obtained values were normalized with respect to the neutral facial expression of the same subject, and then the degree of expression of each of the basic emotions was calculated by the classifiers of three types: a probabilistic neural network of the multilayer perceptron type and a system of logical rules. The final degree of emotion is calculated as the sum of the responses of the probability classifier and neural network. Logical rules are used only to resolve disputes where several emotions receive a similar high degree of severity. It is experimentally proved that in itself it does not have sufficient accuracy, so its independent use is impractical. The declared recognition accuracy is 85%. In [2, 3] several types of features are combined (optical flow, SIFT, hierarchical Gaussian) followed by support vector Machine (SVM) classification. The authors [18] declare the possibility of a significant increase in recognition accuracy through the use of spatial-temporal modification of local binary patterns (LBP-TOP) as features. In [3, 15] demonstrated the algorithm for calculating the intensity of the AU and comparing the effectiveness of different groups of traits and their associations. An interesting approach to the classification is proposed in [5], where AU degrees of expression are transformed into markers of the presence of emotions using logical decision trees specific to different ethnic groups.

The use of fuzzy set theory for emotion recognition is described in [14, 17]. A typical set of fuzzy logic rules is a number of different variants of emotion realizations, based on the recommendations of P. Ekman. For each feature, the minimum threshold of positive or negative expression is experimentally selected, and then a logical rule is drawn up. An example of a logical rule for the emotion "joy" is the expression:

$$v=x_{5+}\&!(x_{2+}|x_{3+})\&x_{14+}\&x_{15+}\&x_{18+}\&x_{4+}$$

where & — logical operator «and»; | — logical operator «or»; ! — logical operator «not», v - recognized emotion «joy», x_{i+} - i-th AU, positive expression of which exceeds the minimum threshold.

It is indicated that for classification using methods of fuzzy logic theory, the priority is to configure the parameters of membership functions, as well as the completeness of expert rules. Thus, in [17] it is proposed to use the improved algorithm of Principal Component Analysis to configure the parameters of membership functions. The results of experimental studies according to which the average accuracy of recognition of seven emotions is 98.32%. At the same time, it is pointed out that a significant obstacle in the development of such tools is the need to create representative databases of fuzzy expert rules that take into account the diversity of recognition conditions.

Also, the analysis allows us to formulate a conclusion that today the most effective are neural network means of emotion recognition. An important prerequisite for their effectiveness is the availability of open and representative databases that can be used for network training and testing. In this case, neural networks can be used both for emotion recognition based on the analysis of characteristic points of the face, and on the basis of a holistic comparison of the image of the face with some standards.

As a rule, the basis of modern solutions are convolutional neural networks (CNN) of different architecture. Other types of tested types of neural network models are less effective both in terms of recognition accuracy and resource consumption. In addition to CNN, recurrent neural networks of the LSTM type are sometimes used in recognition tools to account for the time component [6]. At the same time, practical experience and data [15, 19-23] indicate the need to increase the level of adaptation of modern neural network models of emotion recognition to the typical noise that occurs when fixing the image of the face in General-purpose information systems. Thus, the aim of this study is to develop a neural network model that allows to realize the recognition of emotions based on the geometry of the human face under the influence of noise characteristic of general-purpose information systems.

3. DEVELOPING THE ARCHITECTURE OF THE NEURAL NETWORK MODEL

Let us clarify the problem of emotion recognition. Suppose the need to recognize only the basic emotions on the basis of static images of the faces of the flow of people recorded with video cameras with average characteristics. At the same time, the issues of personality recognition, pre-filtering of the image, the influence of illumination, isolation of individuals on the image and leveling the conscious distortion of the person's face in order to hide their emotional state are not considered. Results [2, 5, 13, 18, 19] point out that in this case in information systems of General purpose the main hindrances arise in a consequence of turn of the image of the person. In accordance with [11], it is possible to eliminate these shortcomings by using a neural network model based on a capsule neural network, which is a modification of CNN adapted to the analysis of rotated and noisy images.

The developed model is based on a shallow capsule neural network of the CapsNet type proposed in [16]. Note that CapsNet allows not only to recognize the analyzed image, but to decode it, that is, to restore the image standard. Thus, the network can be divided into a recognition unit and a decoding unit. The structure of the CapsNet network adapted to the problem of emotion recognition is shown in Fig. 1.

The main structural units of the CapsNet recognition unit are: the input layer of neurons that corresponds to the analyzed image, the convolutional layer (Conv), the layer of primary capsules (PrimaryCaps) and the layer of convolutional capsules (DigitCaps), each of which corresponds to one of the recognized states.

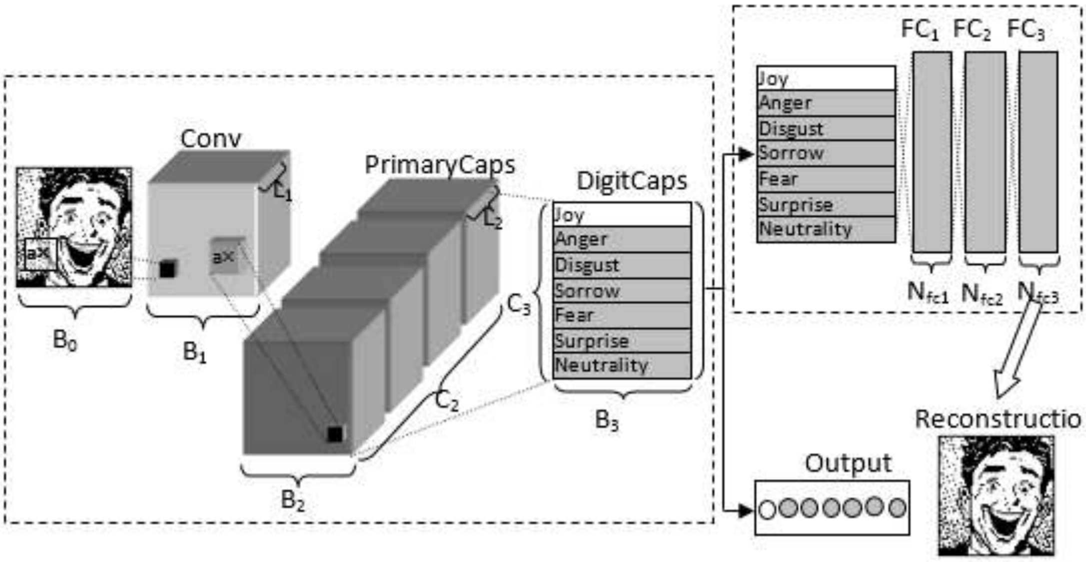

Figure 1. Structure of CapsNet

The decoding unit additionally includes three fully connected layers of neurons (FC1, FC2, FC3). For example, in Fig. 1 shows that the output layer DigitCaps evidence of recognition joyful state. Also shown is a reconstructed image of the face, which corresponds to the activation of neurons of the FC3 layer.

For figure 1 the following designations are used, which correspond to the structural parameters of the network:

- B_0 – vertical and horizontal size of the analyzed image.
- B_1 – size of feature maps for Conv layer.
- B_2 – the mesh size in the layer PrimaryCaps.
- B_3 - number of convolutional units in each of the DigitCaps end layer capsules.
- C_2 – the number of channels in the PrimaryCaps layer.
- C_3 – the number of capsules in the layer DigitCaps.
- L_1 – number of feature maps in the Conv layer.
- L_2 – the number of convolutional units in each channel of the PrimaryCaps layer.
- a – the size of the convolution kernel.
- N_{fc1}, N_{fc2}, N_{fc3} – the number of neurons in the first, second and third (output) fully connected layer of the decoder.
- B_3 – number of convolutional units in the DigitCaps end layer capsule.

By analogy with the classic capsule network the size of the convolution kernel a=9, the step of the convolution kernel Conv to d_1=1, the step of the convolution kernel for PrimaryCaps d_2=2, the number of feature maps L_1=256, the number of channels in the layer

PrimaryCaps C_2=32, the number of convolutional units in the layer PrimaryCaps L_2=8, the number of convolutional units in each of the capsules layer DigitCaps B_3=8, the value of padding the boundaries of the input image r_1=1, the value of padding the border feature maps for layer Conv r_2=0. Numerical values of other architectural parameters of the neural network model are determined on the basis of data [6, 7, 11] taking into account the number of recognizable emotions and the size of the analyzed images. The size of the input image B_0=48, the size of the feature maps of a convolutional layer B_1=41, the grid size in the layer PrimaryCaps B_2 =17, number of capsules in the layer DigitCaps equal to the number of recognizable emotions C_3 = 7, the number of neurons in the first layer fully connected N_{fc1} =512, the number of neurons in the second fully connected layer N_{fc2} =1024, the number of neurons in the third fully connected layer equal to the number of pixels into a recognizable image N_{fc3} =2304.

Calculation of input and output signals for neurons in Conv, FC1, FC2, FC3 layers corresponds to calculations in convolutional neural networks [1, 13, 19]. In the layers Conv, FC1, FC2 uses ReLU activation function:

$$y = \max(0, x) \tag{1}$$

where x – the total input of the neuron, y – output

In the FC3 layer uses a sigmoidal activation function of the form:

$$y = 1/(1 + \exp(-x)) \tag{2}$$

The input of some neuron in the Conv layer is calculated as:

$$x_k^{(i,j)} = x_{0,k} + \sum_{s=1}^{a}\sum_{t=1}^{a} w_{k,s,t} x^{(i+s,j+t)} \tag{3}$$

Where $x_k^{(i,j)}$ - input (*i,j*) -th neuron *k*-th feature map, $x_{0,k}$ - the displacement of neurons *k*-th feature map, *a* - the size of the convolution kernel, $w_{k,s,t}$ - weight coefficient (*s,t*)-th synaptic connections of the neuron the *k*-th feature map, $x_0^{(i,j)}$ – the value of the input signal (*i,j*)-th neuron of the input layer.

In this case, the calculation of the input and output signals of the capsules is realized using the expressions (4-9).

$$v_j = \frac{\|s_j\|^2}{1+\|s_j\|^2}\frac{s_j}{\|s_j\|} \tag{4}$$

$$s_j = \sum_{i=1}^{C_3}\left(c_{i,j}\hat{u}_{j|i}\right) \tag{5}$$

$$\hat{u}_{j|i} = W_{i,j}u_i \tag{6}$$

$$c_{i,j} = \exp\left(b_{i,j}\right)\Big/\sum_{k=1}^{C_2}\exp\left(b_{i,j}\right) \tag{7}$$

$$\Delta b_{i,j} = v_j\hat{u}_{j|i} \tag{8}$$

$$b_{i,j} = b_{i,j} + \Delta b_{i,j} \tag{9}$$

where v_j- output vector of the *j*-th capsule in the layer DigitCaps, s_j - component of the *j*-th DigitCaps layer capsule in the network output signal, $c_{i,j}$- the weighting factor of the degree of coherence between the i-th capsule in the layer PrimaryCaps и j-th capsule in

DigitCaps layer, $\hat{u}_{j|i}$- the predicted value of the output signal of the i-th capsule in the layer PrimaryCaps, $W_{i,j}$ - the matrix of weight coefficients of the connections between the i-th capsule in the layer PrimaryCaps and j-th capsule in the layer DigitCaps, $b_{i,j}$- the logarithm of the probability of connection between i-th capsule in the layer PrimaryCaps and j-th capsule in the layer DigitCaps, u_i - output signal (a vector) of the i-th capsule in the layer PrimaryCaps, $\Delta b_{i,j}$ - corrective coefficient in iterative calculation $b_{i,j}$.
Note that an expression of the form (4) is a so-called squash function, and an expression of the form (7) is a softmax function.

The process of learning a neural network model is that the k-th capsule in the PrimaryCaps layer has a long implementation vector only for images of persons with the corresponding emotion. To do this, the functionality of the form is minimized:

$$\sum_{k=1}^{K} E_k \to \min \tag{10}$$

$$E_k = T_k \max\left(0, m^+ - \|v_k\|\right)^2 + \lambda(1 - T_k)\max\left(0, \|v_k\| - m^-\right)^2 \tag{11}$$

where $K = C_3$ - number of capsules per layer PrimaryCaps.

In this case, for the capsule, which corresponds to the k-th emotion $T_k = 1$ only if this emotion is displayed on the face image. Otherwise $T_k = 0$. Values of other coefficients: $m^+ = 0.9$, $m^- = 0.1$, λ=0,5.

For training and recognition of neural network model it is proposed to use algorithms described in detail in [7, 11].

4. EXPERIMENTS

For training and testing of the constructed neural network model, data sets were used, which were formed on the basis of the Fer2013-images database available on the website www.kaggle.com. The database contains 35494 jpg-files with photos of people's faces, expressing 6 basic emotions and neutral state. There are photos in which the face is fixed in the frontal projection, as well as photos in which the face is fixed in the rotated state. Angle of rotation -45^0 to +45^0. The list of emotions presented in Fer2013-images corresponds to Fig. 1. Each photo is a grayscale image with a resolution of 48x48 pixels. An example of images corresponding to the Joy emotion is shown in Fig. 2.

Figure 2. Examples of photos with Joy emotion.

At the first stage of research experiments on recognition of emotions on the basis of the analysis of geometry of the frontal, well-lit image of the face of the person are carried out. The second stage of research is related to the recognition of emotions based on the analysis of images of the turned face of a person. Experiments are realized by means of specially developed computer program which was based on the mathematical apparatus given by expressions (1-11). The obtained histograms of the accuracy of basic emotions recognition on the frontal and rotated face images are shown in Fig.3 and Fig.4. Also, to test the hypothesis of sufficient efficiency of the capsule neural network model in the conditions typical for General-purpose information systems, a comparison of the recognition accuracy of the constructed CapsNet with the classical convolutional neural network LaNet and one of the most modern modifications of convolutional neural networks VGG is carried out.

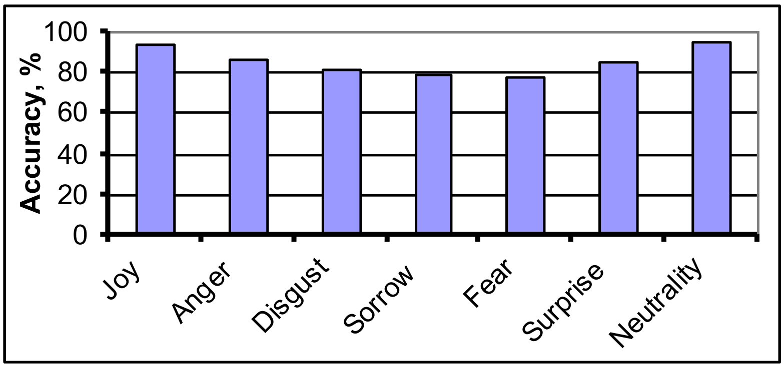

Figure 3. Capsnet recognition accuracy of different emotions on the front face image.

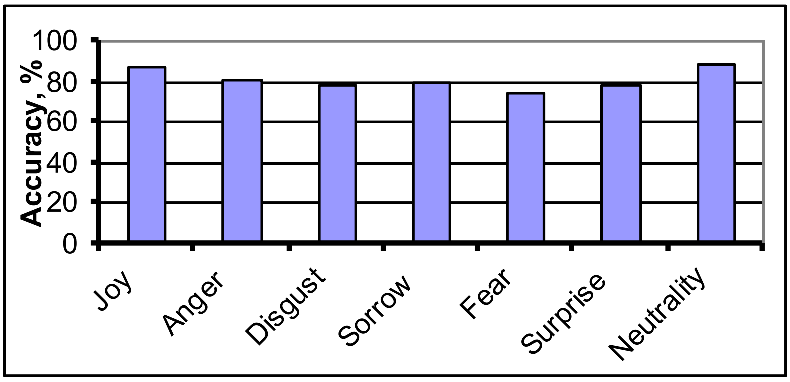

Figure 4. CapsNet recognition accuracy of different emotions on a rotated face image.

Data are used for comparison [3, 4, 9, 12]. The results of the comparison are summarized in table 1

TABLE 1. The average accuracy of recognition of basic emotions

Type of neural network model	Average accuracy	
	For front images	For rotated images
CapsNet	85,3	81,2
Lanet	78,3	71,4
VGG	92,7	87,4

Analysis of graphs shown in Fig. 2, 3 indicates that the frontal image is the least accurately recognized emotion of Sorrow and Fear. In this case, the rotated images worst recognized emotions Fear and Surprise. This can be explained by the non-proportional representation of the corresponding photos in the Fer2013-images database. At the same time, the data table 1 indicate that the recognition accuracy of the front image CapsNet higher recognition accuracy LaNet and slightly lower recognition accuracy VGG. At the same time, the accuracy of emotion recognition on rotated images using CapsNet is commensurate with the accuracy of modern types of convolution networks and significantly exceeds the accuracy of the LaNet. It can also be argued that the overall accuracy of recognition CapsNet rotated images about 5% lower than the front. When using LaNet and VGG accuracy deteriorates by about 9%. In addition, experiments have shown that the training time of the CapsNet network is much longer than the training time of the LaNet network. At the same time, the cost of implementation of the CapsNet far below the carrying capacity of the network VGG.

5. CONCLUSIONS

As a result of the research, a neural network model of the CapsNet type is developed, designed to recognize basic emotions, taking into account such a characteristic noise for general-purpose information systems as face rotation.

It is shown experimentally that in the analysis of undistorted images CapsNet slightly exceeds the accuracy of the classical convolutional neural network type LaNet, which is approximately equal to its resource intensity. However, LaNet is superior to CapsNet in terms of the number of training iterations needed to achieve an acceptable learning error. The accuracy of CapsNet recognition of undistorted images is somewhat inferior to modern types of convolution networks, which have a much higher resource consumption compared to it. When detecting emotions on rotated images, CapsNet is comparable with the accuracy of modern types of convolutional networks and significantly exceeds the accuracy of LaNet. Prospects for further research in the field of neural network recognition of emotions on the geometry of the face can be associated with the improvement of architectural solutions of the capsule neural network in the direction of reducing the number of training iterations while ensuring acceptable recognition accuracy. In addition, improving the efficiency of neural network recognition of emotions is associated with the classification of blurred and partially hidden images of the face.

REFERENCES

[1] Berik Akhmetov, Igor Tereykovsky, Aliya Doszhanova, Lyudmila Tereykovskaya (2018) Determination of input parameters of the neural network model, intended for phoneme recognition of a voice signal in the systems of distance learning. International Journal of Electronics and Telecommunications. Vol **64**, No 4 (2018), pp. 425-432. DOI: 10.24425/123541.

[2] Anderson K., McOwan Peter W. A realtime automated system for the recognition of human facial expressions // Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on , vol.**36**, no.1, pp.96-105, 2006.

[3] J. C. Batista, V. Albiero, O. R. Bellon, and L. Silva. Aumpnet: simultaneous action units detection and intensity estimation on multipose facial images using a single convolutional neural network. In Automatic Face & Gesture Recognition (FG 2017), 2017 12th IEEE International Conference on, pages 866–871. IEEE, 2017.

[4] Bobe A.S., Konyshev D.V., Vorotnikov S.A. Sistema raspoznavanija bazovyh jemocij na osnove analiza dvigatel'nyh edinic lica. Inzhenernyj zhurnal: nauka i innovacii, 2016, T. 9. pp. 1-16, http://dx.doi.org/10.18698/2308-6033-2016-09-15.

[5] Chandrani S., Washef A., Soma M., Debasis M. Facial Expressions: A Cross-Cultural Study. Emotion Recognition: A Pattern Analysis Approach. Wiley Publ., 2015, pp. 69–86.

[6] I. Dychka, I. Tereikovskyi, Sh. Mussiraliyeva, L. Tereikovska, V. Pogorelov Deobfuscation of Computer Virus Malware Code with Value State Dependence Graph.

Advances in Intelligent Systems and Computing, volume **754**, Springer, pp 370-379, DOI:10.1007/978-3-319-91008-6_37

[7] O. Ertugrul, L. A. Jeni and J. F. Cohn, "FACSCaps: Pose-Independent Facial Action Coding with Capsules," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, 2018, pp. 2211-221109. doi: 10.1109/CVPRW.2018.00287

[8] S. Ghosh, E. Laksana, S. Scherer, and L.-P. Morency. A multi-label convolutional neural network approach to crossdomain action unit detection. In Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on, pages 609–615. IEEE, 2015.

[9] Goodfellow, J., Erhan, D., Carrier, L. (2015), Challenges in representation learning: A report on three machine learning contests, Neural Networks 2015, vol. 64, pp. 59–63.

[10] A. Gudi, H. E. Tasli, T. M. Den Uyl, and A. Maroulis. Deep learning based facs action unit occurrence and intensity estimation. In Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on, volume 6, pages 1–5. IEEE, 2015.

[11] G. Hinton, S. Sabour and N. Frosst, ”Matrix capsules with EM routing”, ICLR 2018 Conference, Vancouver Convention Center, Vancouver, BC, Canada April 30 - May 3, 2018

[12] J. He, D. Li, B. Yang, S. Cao, B. Sun, and L. Yu. Multi view facial action unit detection based on cnn and blstmrnn. In Automatic Face & Gesture Recognition (FG 2017), 2017 12th IEEE International Conference on, pages 848–853. IEEE, 2017.

[13] Hu, Z., Tereykovskiy, I., Zorin, Y., Tereykovska, L., Zhibek, A. Optimization of convolutional neural network structure for biometric authentication by face geometry // Advances in Intelligent Systems and Computing. 2018. Volume **754**, pp 567-577.

[14] Ilbeygi, M., Shah-Hosseini, H., “A novel fuzzy facial expression recognition system based on facial feature extraction from color face images”. In: Engineering Applications of Artificial Intelligence 25, 2012, pp. 130–146.

[15] Erik Learned-Miller, Gary B. Huang, Aruni RoyChowdhury, Haoxiang Li, and Gang Hua. Labeled Faces in the Wild: A Survey. In Advances in Face Detection and Facial Image Analysis, Springer, pages 189-248, 2016.

[16] Sabour S., Frosst N. and Hinton G., ”Dynamic Routing Between Capsules”, 2017 Advances in Neural Information Processing Systems 2017-December, pp. 3857-3867

[17] Seyed M., Marjan A., Face emotion recognition system based on fuzzy logic using algorithm improved Particle Swarm, International Journal of Computer Science and Network Security, VOL.16 No.7, July 2016, pp 157-166. (20160718)

[18] Tariq U., Lin K., Li Z., Zhou Z., Wang Z., Le V., Huang T.S., Lv X., Han T.X. Emotion Recognition from an Ensemble of Features. Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions, 2012, vol. 42 (4), pp. 1017–1026

[19] Tereykovska, L., Tereykovskiy, I., Aytkhozhaeva, E., Tynymbayev, S., Imanbayev, A. (2017), “Encoding of neural network model exit signal, that is devoted for distinction of graphical images in biometric authenticate systems”, News of the National Academy of Sciences of the Republic of Kazakhstan, Series of Geology and Technical Sciences, Volume **6**, Number 426, pp. 217 – 224, 2017.

[20] Z. Toser, L. A. Jeni, A. L˝orincz, and J. F. Cohn. Deep learning for facial action unit detection under large head poses. In European Conference on Computer Vision, pages 359–371. Springer, 2016.

[21] Tereikovskyi I., Mussiraliyeva S., Kosyuk, Y., Bolatbek M.,Tereikovska L. An experimental investigation of infrasound influence hard drives of a computer system. International Journal of Civil Engineering and Technology (IJCIET) Volume 9, Issue 6, June 2018, pp. 1558–1566.

[22] Z.Wang, Y. Li, S.Wang, and Q. Ji. Capturing global semantic relationships for facial action unit recognition. In Computer Vision (ICCV), 2013 IEEE International Conference on, pages 3304–3311. IEEE, 2013.

[23] K. Zhao, W.-S. Chu, and H. Zhang. Deep region and multilabel learning for facial action unit detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 3391–3399, 2016.