UDK 004: 811.512.11
https://doi.org/10.47533/2023.1606-146X.2

Y. N. AMIRGALIYEV[1,2], S. B. MUKHANOV[1]*, D. B. ZHEXENOV[1], N. K. KALZHIGITOV[1], A. S. LI[1], D. D. YEVDOKIMOV[1], C. KENSHIMOV[2]

[1]International University of Information Technologies, Almaty, Kazakhstan
[2]Institute of Information and Computer Technologies, Almaty, Kazakhstan

A COMPARATIVE ANALYSIS OF NEURAL NETWORK MODELS FOR HAND GESTURE RECOGNITION METHODS

In recent years, gesture recognition methods have undergone major changes. Because the very demand for it has reached a different level. Humans have increasingly begun using various areas of human activity. The purpose of gesture recognition is to record gestures formed in a certain way and then tracked by a device such as a camera. Hand gestures can be used as a form of communication within many different applications. It can be used by people with various disabilities, including those with hearing and speech impairments, to communicate and interact socially with others. Our research demonstrates various methods for implementing hand gesture recognition based on Hidden Markov Model (HMM), Convolutional Neural Network (CNN), Diffractive Deep Neural Network (D^2NN) and other neural networks. This research reviews previous approaches and results of hand gesture recognition methods, hypotheses, diagrams, as well as a comparative review between various gesture recognition methods are given in this paper.

***Key words**: Hand gesture recognition, gesture recognition video-chat, neural networks, dynamic gestures.*

Introduction. Nowadays in Kazakhstan there are more than 200 thousand mute people, 80 thousand deaf people, and in the world more than 430 million people suffer from hearing loss, and the number of mute people exceeds more than 70 million [1]. For many of them, this is a problem they face every day. Many of these people suffer in terms of communication and are often forced to use the services of an interpreter for normal communication. And the solution to this problem is too expensive for many. The average price of a hearing aid is 300,000 tenge and more, not to mention how often they must be changed. This problem covers a huge contingent of people, so we decided to take on the solution to this problem. We thought of different ways to help the deaf and mute people. The key problem for mute and deaf people was communicating with ordinary people and being unable to support or understand them.

We have considered different solutions to this problem:

1. Online site for learning sign language
2. Gesture recognition and translation
3. Platform for finding people as sign language interpreters
4. Video chat with gesture recognition

After much deliberation, our choice fell on the fourth option. Since quite a few people today use chat roulette for fun, we decided to use a fun learning method. We plan to embed gesture recognition into online chat, where the user will be able to communicate with mute or deaf people. Thus, users can learn sign language in a playful and conversational way. In

* E-mail корреспондирующего автора: kvant.sam@gmail.com

addition, it is also useful for - most mute or deaf people, they will be able to learn various sign languages such as Kazakh, world and Russian and even those who do not know sign language themselves.

Goal: Development of a convenient platform for real-time hand gesture recognition.

Hypothesis:

1. The topic of research is the recognition of sign language or dactyl alphabet based on the training of neural networks.
2. Analysis of analogues and recognition systems in different sign languages.
3. Selection and training of various neural network architecture models, as well as a comparative analysis of gesture recognition methods.
4. Designing an application for sign language recognition.

An Overview of Gesture Capture and Recognition Techniques And Methods. In our research, the Python pipeline library was used directly to capture hand gestures. The remaining methods, and the methodology of further work itself, were used based on the works of other authors. Because a lot of work has been done on gesture recognition. Maria Abastillas in Real Time Hand Gesture Recognition and Classification Technique (June 17, 2011) this work is discussed, a way to recognize and classify hand movements performed by healthy subjects in real time. In addition, the analysis of the choice of the simplest features and the best classifier for the indicated hand movements is discussed. 2 hand movement modes: relaxed hand and closing hand are classified on the Lab View platform. The classifier model was trained using 75% of the data as the training set and the remaining 25% as model test/validation data. The results show that the proposed model is an efficient and accurate method for classifying hand movements with approximately 96.58% accuracy for offline classification. The classified model was also implemented in the LabView software by pairing it with Python. Thomas Bravenek, Thomas Friza in "Multiplatform Hand Gesture Recognition System" (2019). This article collects available gestures and finger detection in still images and video snapshots. The document also contains a quick test of various approaches to convenient gesture detection, including because the platform implementation is an application written in Python using the OpenCV and PyTorch libraries that will display a specific image or play a video sequence with highlighted recognized gestures.

Google company's project GoogleAI shows an approach to accurate and effective hand and finger tracking using machine learning on the MediaPipe cross-platform framework, which in turn builds data processing pipelines (video, audio, and time series) [2]. The authors use models such as a palm detector (blazePalm), a model for determining key points on the hand, and a gesture recognition algorithm. These models form a single basis of the above structure. Every model is unique and defines the key points for identifying special elements in gesture recognition.

The authors of the Russian Sign Language Dactyl Recognition article compared several real-time sign language recognition systems and presented a new model based on deep convolutional neural networks. These systems are capable of recognizing letters of the Russian alphabet represented as static characters in Russian Sign Language used by people from the deaf community. In this approach, they recognize Russian natural language words represented by successive gestures of each letter. The authors evaluate their approach to Russian Sign Language (RSL), for which they collected their own dataset and evaluated dactyl recognition [3].

A. Dhawan et al [5] focus on only one camera as an input source, while H.S. Yeo et al [4] compare image recognition results with a single camera and a depth camera. In H.S. Yeo et al [4] have proven that a user can interact with a computer without using any physical controllers. Another discovery of H.S. Yeo et al [4], the single camera performs poorly in low light conditions. To solve this problem, there is a wide range of - YCbCr thresholds. be used. However, this limitation does not occur when using the Kinect Depth Camera due to its robustness to preserve the contours of the hand by applying a fixed depth threshold. Table 1 summarizes of some hand gesture recognition systems: a comparison of methods in which hand gesture recognition methods are used.

Table 1 – Results of Hand Gesture Recognition Systems Using Different Types of Cameras.

Research	Method	Detection	Feature Extraction	Recognition (%)
M.M. Hassan, P.K. Mishra	Vision-based using single camera	HSV color Threshold method	Classical normalization: Divide the hand image into blocks of intensity features, which is then being extract by using edge information technique	Classical normalization 91%
M.M. Hassan & P.K. Mishra	Vision-based using single camera	HSV color Threshold method	Comparison between classical normalization technique and block scaling using center of mass normalization technique	Classical normalization 83.3%
				Block scaling normalization using center of mass 96.6%
Ginus Thomas	Vision-based using single camera	RGB color Otsu threshold method	Image captured are being tested in few methods: a) Pixel by pixel comparison b) Edges method	Pixel by pixel method 86%,
				Edges method 92%
				Orientation

Based on the table above, we can conclude that applying different image processing techniques, the results of recognition success will be different, according to this the best option can be chosen for your system.

Table illustrates number of researchers, recognition methods of a system that used other tracking devices such as gloves, detection, feature extraction and percentage of recognition.

Table 2 – Summary of Hand Gesture Recognition Systems Using Tracking Devices.

Research	Method	Detection	Feature Extraction	Recognition (%)
Y. Yao, Y. Fu	14-patches color glove	RGB-D using Kinect sensor	Database indexing technique is used by collecting training samples. Similarity between training samples and gesture are tested	51.87%
J. Nagi, A. Giusti, L. Gambardella, G.A.Caro	Colored Glove	HSV color space	Support Vector Machine (SVM) is trained in a cascaded multi-binary-class configuration, where the spatial gestures are effectively learned and recognized by a swarm of UAVs.	Spatial gesture recognition is robust and scales well with swarm sizes of up to 20 robots
Kılıboz, Nurettin Çağrı Güdükbay, Uğur	Six-degree-of-freedom (DOF) Magnetic motion tracking device	Collected motion data is converted into relative position data in gradient form (x-y axis)	Finite State Machine (FSM) is used as a gesture recognizer method. Needleman-Wunsch sequence matching algorithm is applied in order to produce similarity scores between two sequences.	Recognition rate: 73%. The outcome can be improved by using a more powerful tracking approach which has better sensor capability

Convolutional neural network, hidden markov model and diffractive deep neural network architectures. Based on the data of these tables and to obtain our own results, we have studied some models and types of Neural Network (NN). So, Convolutional Neural Network (CNN) is a neural network consisting of several layers that perform convolutional calculations based on mathematical algorithms. The CNN consists of convolutional layers, pooling layers, and fully connected layers. These layers create feature maps, after which they use special operations and rules to reduce the sample size of these features step by step. Further, information about the objects remaining after the previous operations is collected in the layers of the union. The CNN has good learning abilities, qualitatively extracting convolutional features and classifying related layers [6].

Benefits of Using CNN for Deep Learning:

• For image recognition applications, CNNs are especially useful as they provide highly accurate results, especially when a large amount of input data is used. Such training removes the need for self-extraction of features.

• CNN can be retrained for completely different recognition goals and built based on existing networks. Thanks to this, you can use this type of network, and it will not be very resource intensive.

• CNNs are more computationally efficient than regular neural networks because they share parameters. These models are easy to deploy and able to run on any device.

The Hidden Markov Model (HMM) is a model in which the system being modeled is treated as a Markov process with hidden states. The HMM gives us the ability to consider observable and hidden events that are contained in our model. One of the main applications of this model is gesture recognition.

Benefits of HMM model:

• HMM has a strong statistical foundation and efficient learning algorithms where learning can be done from the raw sequence data.

• It provides the ability to sequentially handle insertion and deletion penalties as locally trained methods and can handle variable length inputs.

• It can also perform a large number of operations, including multiple alignment, data analysis and classification, structural analysis, and pattern detection.

• It is also easy to combine this model into libraries.

An Optical Neural Network (ONN) is a physical implementation of an artificial neural network with optical components.

In the ONN architecture, the neural network is physically formed by multiple diffractive surfaces working together to perform an arbitrary function. The ONN takes characteristics of input information with amplitude, phase and light intensity based on optical modulation. Although the inference and prediction of a physical network is entirely optical, the part of the training that determines its design can be performed by computers.

Meanwhile, different The ONN models have emerged, among them, Diffractive Deep Neural Network (D^2NN), based on the all-optical structure of machine learning, provides an extremely high degree of freedom for model training, and facilitates important applications in a wide range of fields.

The Diffractive Deep Neural Network (D^2NN) has demonstrated its importance in performing various all-optical machine learning tasks, such as classification, segmentation, etc, the operation algorithm of this network is shown in Figure 1.

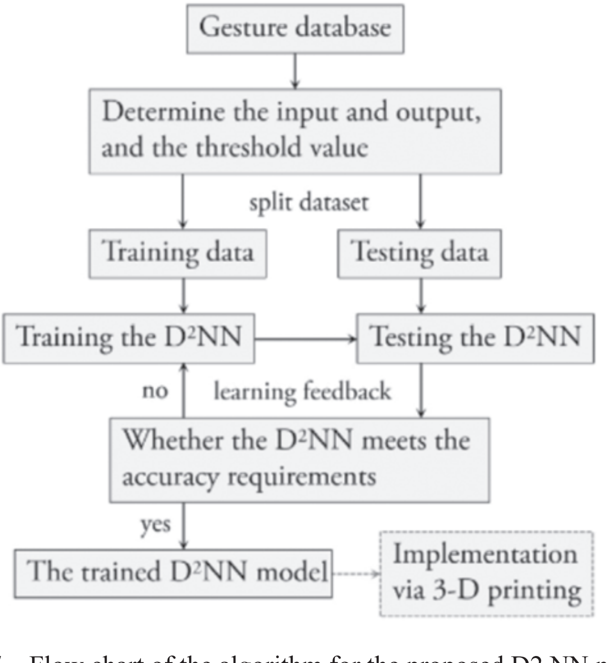

Figure 1 – Flow chart of the algorithm for the proposed D2 NN model

In D^2NN framework, each neuron has a complex transmission coefficient, i.e.,

$$t_i^l\ (x_i, y_i, z_i) = a_i^l\ (x_i, y_i, z_i) \exp\ (j\phi_i^l\ (x_i, y_i, z_i)), \tag{1}$$

where *i* and *l* denote the neuron and diffractive layer number, respectively. a_i^l and ϕ_i^l are represented during the network training as functions of two latent variables, α and β, defined in the following form:

$$a_i^l = sigmoid\ (\alpha_i^l), \tag{2}$$

$$\phi_i^l = 2p \times sigmoid\ (\beta_i^l), \tag{3}$$

where, $sigmoid(x) = \frac{e^x}{e^x + 1}$, is a non-linear, differentiable function. In fact, the trainable parameters of a D^2NN are these latent variables, α_i^l and β_i^l, and eq. (1) defines how they are related to the physical parameters (a_i^l and ϕ_i^l) of a diffractive optical network. When we combine D^2NN training related changes reported in the earlier sub-section on the parametrization of neuron modulation (eq. (2)), with the cross-entropy loss outlined above, a significant improvement in the classification performance of an all-optical diffractive neural network is achieved [7, 8].

Authors in the research [9] used real-time tracking and hidden Markov models to track the moving hand and extract the region of the hand. They then used the Fourier descriptor to characterize the spatial features.

A scheme for extracting and classifying 2D motion in an image sequence based on motion trajectories was presented by the authors in the paper [10]. They used multi-scale segmentation, affine transformations, and a time-delayed neural network to show that hand gesture motion patterns can be extracted and accurately recognized using motion trajectories.

The authors [11] presented an algorithm for extracting and classifying 2D motion in an image sequence based on a motion trajectory. First, multi-scale segmentation is performed to create homogeneous regions in each frame. Regions between, then consecutive frames are matched to obtain two kinds of matches. Affine transforms are computed for each pair of matching regions to determine pixel matches. The pixel hits in successive pairs of images are combined to obtain the pixel level of the motion path through the sequence of images. Motion patterns are learned from the extracted trajectories using a neural network with a time delay. They applied the proposed method to recognize 40 American Sign Language hand gestures. Experimental results show that hand gesture motion patterns can be extracted and accurately recognized using motion trajectories.

The authors used the glove in [12] to create a hand-machine interface. The glove is supposed to be equipped with sophisticated sensors capable of detecting bending movements, measuring finger flexion, etc. The position and orientation of the hand are measured either by ultrasound, which provides five degrees of freedom, or by magnetic flux sensors, which provide six degrees of freedom. Thus, the glove owner will have an interface to the visual programming language.

In research [13] the authors proposed a two-level approach to solving the problem of classifying hand gestures based on real-time vision. The bottom layer of the approach implements pose recognition with Haar-like features and the AdaBoost learning algorithm. The top layer implements linguistic hand gesture recognition using grammar-based context-free parsing. Given an input gesture based on the extracted poses, composite gestures can be parsed and recognized using a set of primitives and production rules.

Paul S. Heckbert has presented a seed-fill algorithm in the paper [14]. Authors in the research [15] have presented a modified new seed-fill algorithm. Authors of this research proposes a two-level approach to solving Hand gesture classification problems based on real-time vision The lower level of the approach implements posture recognition with Haar-like features and the AdaBoost learning algorithm. With this algorithm, real-time performance, and high recognition accuracy it might work. The top level implements linguistic hand gesture recognition using a context-free syntax based on grammar parsing. Given an input gesture based on the extracted poses, composite gestures can be parsed and recognized using a set of primitives and production rules [14].

Authors proposed various methodologies for detecting and finding out whether pixels are aligned in a straight line or not [16, 17, 18]. Our method uses a very cheap device that includes LEDs and a glove to implement a hand gesture recognition library. These gestures can be used to simulate various mouse operations. The methods used are simple, efficient and easy to implement as they do not require too many complicated operations.

Results. Based on the research of other authors and the literature review, as well as through our experiments, we have determined the approximate architecture of our application. For clarity, various types of UML diagrams were created, they are shown below. These schemes do not describe the methods and ways of recognition of gestures, they describe the main functions, components, and principles of interaction of these components with each other. Those, in the future, regardless of the chosen method of working with gestures, according to these schemes, it will be possible to create a fully functioning application.

Figure 2 shows the class diagram of our video-chat application with gesture recognition. It describes the main classes and the types of relationships between them, as well as the methods that these classes implement.

So, we have defined our main class - User. The user can have different roles, be it a regular user, an admin, or a moderator. He directly sets all the chat settings himself, selects the type of chat, and other functions appear depending on the type he has chosen, such as random or closed chat.

Figure 3 is a deployment diagram that describes the system execution architecture, including nodes and the interfaces that link them.

In our case, there are 4 blocks:

1. User interface

– Renders to the user all information, and interaction tools.

2. Main server

– Processes the data sent by the user and sends the results of queries back to the user and receives data from the database and ML servers for further processing.

3. Database server

– Stores all the necessary information, such as user data and gesture datasets.

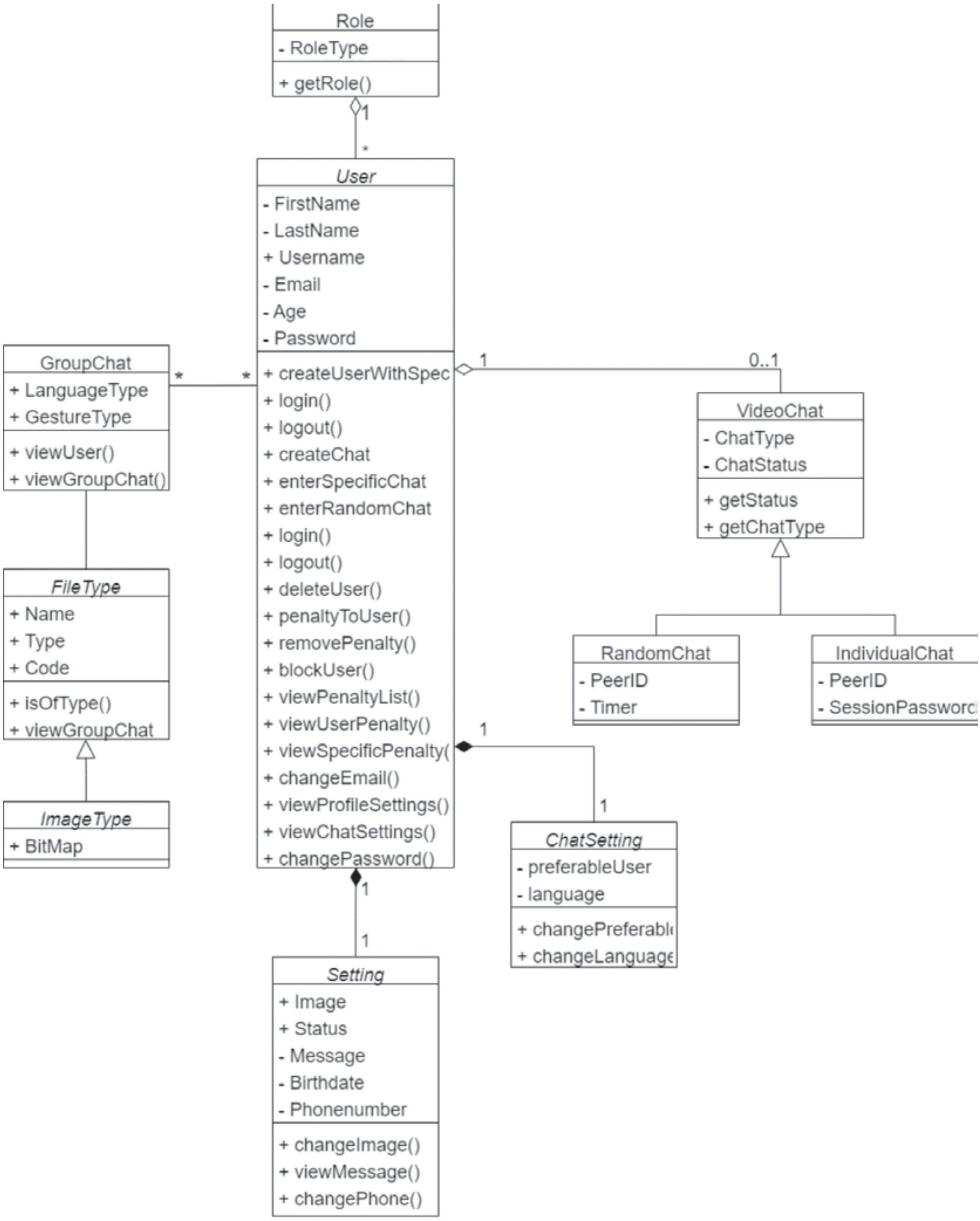

***Figure 2** – UML class diagram*

4. ML server

– Dynamically processes data from the webcam, translating gestures from images into the selected language

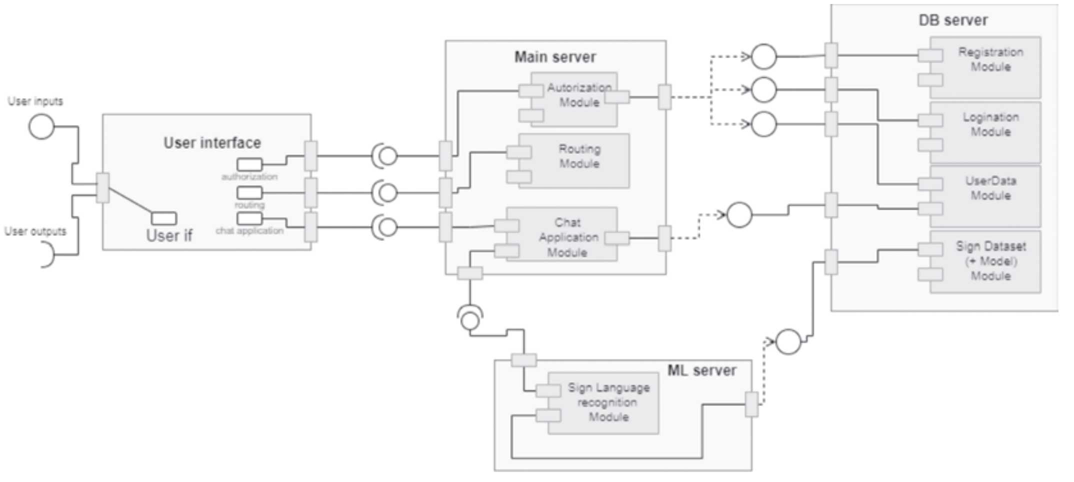

Figure 3 *– Deployment architecture*

This diagram illustrates software development application for gesture recognition of sign language. This picture describes solution architecture and logical structure of frontend and backend side interaction.

In addition to describing the architecture, we also conducted test experiments on the recognition of the Kazakh sign language, using the dataset collected for the paper [19]. To implement this experiment, we used the YOLOV5 model.

The YOLO (You Only Look Once) model works based on KNN, used for high performance object detection. The model divides the target image into a system of grids, and these, in turn, detect objects within themselves. They are used for real-time object detection based on so-called data streams. They do not require a lot of computing resources.

After that, the trained model was tested on the same images. Figure 5 shows examples of gesture recognition results from an image from the dataset.

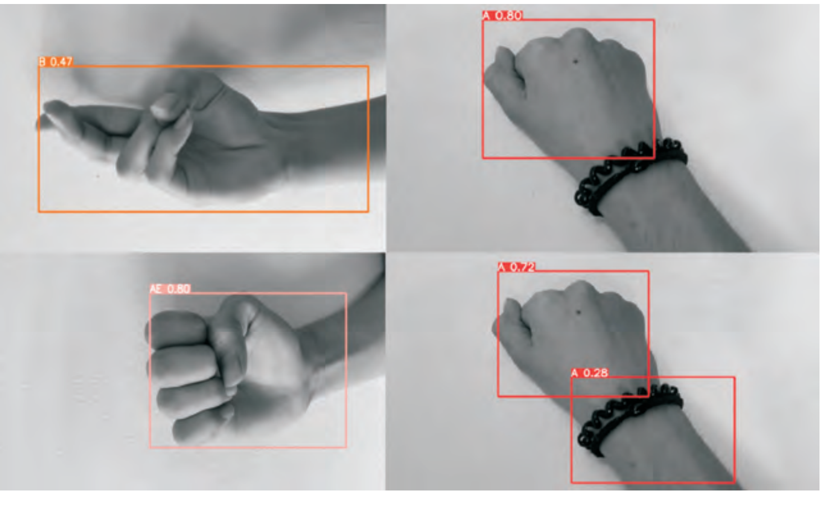

Figure 4 *– The result of gesture recognition by the YOLO model*

The numbers above the border of the object's rectangle indicate the probability of the correctness of the prediction made by the gesture, and the letters next to it are the symbols themselves that these gestures represent.

As you can see from Figure 5, in general the model did a good job but in the fourth image, an erroneous choice was made with a low probability, most likely this was due to the small amount of data on which the model was trained. The solution to the problem is simple - we need to expand the database of gesture images to train the model.

After that, tests of the model with a webcam were also carried out, which showed a certain result, but due to the inadequate quality of the camera, it was decided to repeat this experiment on an already better trained model, and with better equipment.

Conclusion. This overview paper discusses various methods of gesture recognition, including neural network, HMM, CNN, etc. For dynamic gestures, HMM tools are excellent and have proven their effectiveness, especially for robot control. CNNs are used as a classifier and to capture the shape of the hand. Some techniques and algorithms are required for feature extraction, even for capturing the shape of a hand, as in a 2D Gaussian function to match the segmented hand, which was used to minimize the effect of rotation. The choice of a specific recognition algorithm depends on the required application. This article presents the areas of application of the gesture system. An explanation of the problems of gesture recognition is also given, a detailed discussion of the latest recognition systems. Some selected systems are also listed.

We noticed that Convolutional neural network will be effective in recognizing the Kazakh sign language, an example of the work of CNN and a learning model based on the Kazakh language dataset. HMM also having its advantages in the form of a high speed of the algorithm, as well as the presence of many studies by other authors, is a good option to consider. And D^2NN, due to its novelty, has great advantages in the form of the ability to give a good percentage of accuracy with relatively small volumes of the dataset used, as well as the ability to recognize even very low-quality images. In connection with the above, the choice of our main model for the project fell on D^2NN, so its advantages seemed to us more significant.

Also, based on the experiment with the YOLO model, we concluded that the Kazakh language dataset we have can be used as a training for our models, with the prospect of expanding it in the future.

In future work, we plan to create our own workable application based on the research and materials studied, capable of recognizing several different sign languages, including Kazakh, as well as able to recognize gestures like swiping, pointing, and clicking and testing it on real people.

Acknowledgment. This research has been funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan, Grant AP14871625.

REFERENCES

1 Anna Vidyanova, "In the USA, they are interested in the development of Kazakhs for the deaf", Capital, 2022 | https://kapital.kz/business/105455/v-ssha-zainteresovalis-razrabotkoy-kazakhstantsev-dlya-glukhikh.html.

2 Bazarevsky, V., Fan, Zh. (2019). On-device, real-time hand tracking with mediapipe. Google AI Blog. Available at: https://ai.googleblog.com/2019/08/on-device-real-time-hand-tracking-with.html.

3 Russian Sign Language Dactyl Recognition, Ilya Makarov, Nikolay Veldyaykin, Maxim Chertkov, and Aleksei PokoevNational Research University Higher School of Economics, Moscow, Russia.

4 H.-S. Yeo, B.-G. Lee, and H. Lim, "Hand tracking and gesture recognition system for humancomputer interaction using low-cost hardware," Multimed. Tools Appl., May 2013.

5 A. Dhawan and V. Honrao, "Implementation of Hand Detection based Techniques for Human Computer Interaction," Int. J. Comput. Appl., vol. 72, no. 17, pp. 6–13, 2013.

6 J.O. Pinzón-arenas, R. Jiménez-moreno, J.E. Herrera-benavides, Convolutional neural network for hand gesture recognition using 8 different emg signals[C]//2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA) IEEE, Bucaramanga, Colombia (2019), pp. 1-5.

7 Deniz Mengu, Yi Luo, Yair Rivenson, and Aydogan Ozcan "Analysis of Diffractive Optical Neural Networks and Their Integration with Electronic Neural Networks", University of California.

8 Buribayev, Z., Merembayev, T., Amirgaliyev, Y., & Miyachi, T. (2021, April). The Optimized Distance Calculation Method with Stereo Camera for an Autonomous Tomato Harvesting. In 2021 IEEE International Conference on Smart Information Systems and Technologies (SIST) (pp. 1-5). IEEE.

9 Bretzner. L, Laptev. I, Lindeberg. T, Hand gesture recognition using multi-scale colour features, hierarchical models and particle filtering, Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition, Pages 423-428, 2002.

10 Feng-Sheng Chen, Chih-Ming Fu, Chung-Lin Huang, Hand gesture recognition using a realtime tracking method and hidden Markov models, Image, and Vision Computing, Volume 21, Issue 8, Pages 745–758, 1 August 2003.

11 Ming-Hsuan Yang, Ahuja, N. Tabb, Extraction of 2D motion trajectories and its application to hand gesture recognition, Pattern Analysis and Machine Intelligence, IEEE Transactions on (Volume:24 , Issue: 8), Page(s):1061 - 1074, 2002.

12 Arnaud Lemoine, Sean Mcgrath, Anthony Vernon Walker Smith, Alistair Ian Sutherland, Hand gesture recognition system and method, Patent US 6128003 A, 2000.

13 Thomas G. Zimmerman, Jaron Lanier, Chuck Blanchard, Steve Bryson, Young Harvill, A hand gesture interface device, CHI '87 Proceedings of the SIGCHI/GI Conference on Human Factors in Computing Systems and Graphics Interface, ACM, Pages 189-192, 2000.

14 Qing Chen, Georganas, Nicolas D., Petriu E.M., Real-time Vision-based Hand Gesture Recognition Using Haar-like Features, Instrumentation and Measurement Technology, Conference Proceedings, IMTC 2007, IEEE, Pages 1-6, 2007.

15 Paul S Heckbert, A seed fill algorithm, Graphics gems, Pages 275-277, Academic Press Professional, Inc. San Diego, CA, USA ©1990, ISBN:0-12-286169-5.

16 Malik Sikandar Hayat Khayal, Aihab Khan, Saba Bashir, Farhan Hassan Khan, Salma Aslam, Modified new Algorithm for Seed Filling, Publication of Little Lion Scientific, Journal of Theoretical and Applied Information Technology, Vol. 26 No. 1, 2011, ISSN 1992-8645.

17 Rafael Grompone von Gioi· Jérémie Jakubowicz, Jean-Michel Morel·Gregory Randall, On Straight Line Segment Detection, Journal of Mathematical Imaging and Vision, DOI 10.1007/s10851-008-0102-5, Springer Science and Business Media, LLC 2008.

18 Agnes Desolneux, Lionel Moisan, Jean-Michel Morel, Meaningful Alignments, International Journal of Computer Vision 40(1), 7–23, 2000.

19 Kenshimov, C., Mukhanov, S., Merembayev, T., Yedilkhan, D. (2021). A comparison of convolutional neural networks for Kazakh sign language recognition. Eastern-European Journal of Enterprise Technologies, 5 (2 (113)), 44–54.

Е. Н. ӘМІРҒАЛИЕВ*[1,2]*, С. Б. МУХАНОВ*[1]*, Д. Б. ЖЕКСЕНОВ*[1]*,
***Н. К. ҚАЛЖІГІТОВ*[1]*, А. С ЛИ.*[1]*, Д. Д. ЕВДОКИМОВ*[1]* , Ш. КЕНШІМОВ*[2]**

[1]*Халықаралық Ақпараттық Технологиялар Университеті, Алматы, Қазақстан*
[2]*Ақпараттық және Есептеуіш Технологиялар Институты, Алматы, Қазақстан*

ҚОЛ ЖЕЛІЛЕРІН ТАНУ ӘДІСТЕРІНІҢ НЕЙРЛІК ЖЕЛІЛЕРДІҢ САЛЫСТЫРМАЛЫ ТАЛДАУЫ

Соңғы жылдары ым-ишараны тану әдістері үлкен өзгерістерге ұшырады. Өйткені оған деген сұраныстың өзі басқа деңгейге жетті. Адамдар адам қызметінің әртүрлі салаларын көбірек пайдалана бастады. Қимылдарды тану мақсаты белгілі бір жолмен қалыптасқан, содан кейін камера сияқты құрылғы арқылы бақыланатын қимылдарды жазу болып табылады. Қол қимылдарын көптеген әртүрлі қолданбаларда байланыс нысаны ретінде пайдалануға болады. Оны әртүрлі мүгедектер, соның ішінде есту және сөйлеу қабілеті бұзылған адамдар басқалармен қарым-қатынас жасау және әлеуметтік қарым-қатынас жасау үшін пайдалана алады. Біздің зерттеуіміз жасырын Марков моделі (HMM), конволюционды нейрондық желі (CNN), дифракциялық терең нейрондық желі (D2NN) және басқа нейрондық желілер негізінде қол қимылын тануды жүзеге асырудың әртүрлі әдістерін көрсетеді. Бұл зерттеуде қол қимылын тану әдістерінің алдыңғы тәсілдері мен нәтижелері, гипотезалар, диаграммалар қарастырылады, сондай-ақ ым-ишараны танудың әртүрлі әдістері арасындағы салыстырмалы шолу осы мақалада келтірілген.

***Түйін сөздер**: қол қимылын тану, ым-ишараны тану видео-чаты, нейрондық желілер, динамикалық қимылдар.*

Е. Н. АМИРГАЛИЕВ*[1,2]*, С. Б. МУХАНОВ*[1]*, Д. Б. ЖЕКСЕНОВ*[1]*,
***Н. К. КАЛЖИГИТОВ*[1]*, А. С. ЛИ*[1]*, Д. Д. ЕВДОКИМОВ*[1]* , Ш. КЕНШИМОВ*[2]**

[1]*Международный Университет Информационных Технологий, Алматы, Казахстан*
[2]*Институт информационных и вычислительных технологий, Алматы, Казахстан*

СРАВНИТЕЛЬНЫЙ АНАЛИЗ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ ДЛЯ МЕТОДОВ РАСПОЗНАВАНИЯ ЖЕСТОВ РУК

В последние годы методы распознавания жестов претерпели серьезные изменения, потому что сам спрос на нее вышел на другой уровень. Человек все шире стал использовать различные сферы человеческой деятельности. Целью распознавания жестов является запись жестов, сформированных определенным образом, а затем отслеживаемых таким устройством, как камера. Жесты рук могут использоваться как форма общения во многих различных приложениях. Его могут использовать люди с различными ограниченными возможностями, в том числе с нарушениями слуха и речи, для общения и социального взаимодействия с другими людьми. Наше исследование демонстрирует различные методы реализации распознавания жестов рук на основе скрытой марковской модели (HMM), сверточной нейронной сети (CNN), дифракционной глубокой нейронной сети (D2NN) и других нейронных сетей. В этом исследовании рассматриваются предыдущие подходы и результаты методов распознавания жестов рук, гипотезы, диаграммы, а также сравнительный обзор различных методов распознавания жестов.

***Ключевые слова**: распознавание жестов рук, видеочат с распознаванием жестов, нейронные сети, динамические жесты.*