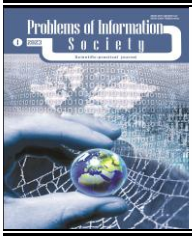


www.jpis.az

14 (1)
2023

A Comparative Analysis of Gesture Recording Technologies and Recognition Methods

Kamala Sh. Gurbanova[1], Fargana J. Abdullayeva[2]

[1,2]Ministry of Science and Education Republic of Azerbaijan Institute of Information Technology, B.Vahabzade str., 9A, AZ1141, Baku, Azerbaijan

[1]kemalewamil@gmail.com, [2]a_fargana@mail.ru

orcid.org/0000-0002-9234-0066[1], orcid.org/0000-0003-2288-6255[2]

ARTICLE INFO

http://doi.org/10.25045/jpis.v14.i1.05

Article history:
Received 30 June 2022
Received in revised form
15 September 2022
Accepted 23 November 2022

Keywords:

Sign language
Dynamic gestures
Gesture recording technologies
Hidden Markov model
Artificial neural network method
Color segmentation algorithm
Viola-Jones algorithm

ABSTRACT

The dynamic development of computing techniques and communication tools and the improvement of network technology have increased the role of information as the main resource in society. The application of information communication technologies has stimulated the development of intellectual and scientific potential all over the world and has been successfully applied to all fields. Gestures are the only means of communication for hearing and speech impaired people. Automatic recognition of gestures in order to facilitate communication through gestures is an urgent issue from both scientific and practical point of view. This article highlights the static and dynamic gestures. The process of gesture recognition collects data by means of various sensor technologies. The article analyzes the image-based and non-image-based technologies and presents their advantages and disadvantages. It also comparatively analyzes the working principle of existing methods proposed for the gestures identification, explores their advantages and disadvantages and interprets the performance of the software which localizes the hand showing the gesture in the video frame. As a result, it develops a machine learning method based on neural networks for high accuracy identification of gestures. The developed method testing on database open for research shows high performance.

1. Introduction

Scientific and technical progress has a positive effect on the development of all areas of social life and causes a fundamental qualitative change in the development of technology with a great leap. Along with other fields, information communication technology (ICT) is developing rapidly. The demand for using ICT is increasing day by day. One of the important issues facing the society is providing hearing and speech impaired people to benefit from the achievements of ICT as all other people.

World Health Organization statistics for 2021 reports about 466 million people to have hearing and speech impairment problems [1].

Gestures are the only means of communication for hearing and speech impaired people. Hardware and software that recognize gestures and translate them into text, voice or vice veCSA simplify the socialization complications of people with speech and hearing impairments. The automatic recognition system of elements of the sign language and the tactile alphabet will encourage children with hearing and speech disabilities to complete secondary education, have higher education, acquire any profession, integrate into society, and create new opportunities for the development of their lifestyle. There is a need to create a national sign language corpus in our country and develop effective methods to recognize national gestures and convert them into text and voice and vice veCSA. Consequently, the automatic recognition of gestures is a relevant issue from both scientific and practical point of view [2].

This article analyzes the working principles, effectiveness, detection accuracy and stability of existing methods for gesture recognition, highlights their advantages and shortcomings, and conducts their comparative analysis.

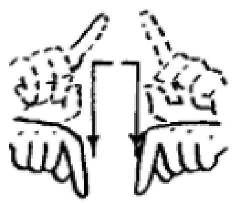
2. Communication with gestures and their automatic recognition stages

Communicating with gestures. People start to get used to a social lifestyle from the birth, and as a result, they cannot live outside of society. In general, society refers to a community of people. Everyone is a member of society, regardless of physical limitations and financial status. The integration into society bases on the communication factor. Communication problems hinder the adaptation of people with speech and hearing impairments to society, and complicate receiving information or sharing thoughts with other people.

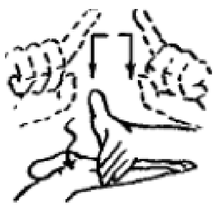
Sign language, which is the only means of communication for hearing impaired people, develops independently and shapes the unique sign language of each nation. Communication with sign language refers to the communication with facial expressions, hand movements and body language [3].

According to the expression forms, gestures are divided into two groups: static and dynamic gestures. Static gestures are motionless, immobile, stable position of the hand in space. Dynamic gestures refer to the sequential movement of the hand in space in a specific time period [4].

Unlike static gestures, dynamic gestures include a sequence of video frames, and therefore the working principle of dynamic gesture recognition methods is more complex than those for static ones. If the gesture is static, its recognition process is simple. During a certain fixed time-interval, the hand remains motionless, during which an image of the hand is recorded and the image is transmitted to the system for processing. In this case, the main task of the system is to analyze the palm configuration and compare it with the corresponding template. Dynamic gesture recognition process depends on hand configuration and movement trajectory. Dynamic gestures can be simple or complex. Figure below illustrates the word "road map" as an example of a simple dynamic gesture, and the word "map" as an example of a complex dynamic gesture (Fig. 1) [5]. When performing simple dynamic gestures, the configuration of the hand remains fixed, while the hand performing the gesture is in motion. In this case, the task of the recognition system is to record the hand configuration and the trajectory of movement. Compared to the simple dynamic gestures, the complex dynamic gestures are more difficult to recognize. In complex dynamic gestures, the hand configuration changes as the movement trajectory changes. In this case, several hand configurations and movement trajectories are recorded. The abundance of information complicates the computation task [6].



a) Road map



b) Map

Fig. 1. Dynamic gestures: a) simple dynamic gesture, b) complex dynamic gesture

3. Stages of automatic recognition of gestures

Advances in programming and the emergence of more sophisticated technologies have created the need to revise real-time automatic gesture recognition systems.

Automatic gesture recognition refers to hardware and software that recognizes gestures and converts them into text, voice, or vice veCSA. The process of real-time recognition of hand gestures is performed in sequence in 4 stages: collecting information with special technologies; finding and segmenting a hand pattern from an image; processing segmented part and detecting key feature; obtaining the result [7].

Collecting information. The initial stage in the process of gesture recognition is the recording and collection of hand patterns or video frame showing the gesture. Data is recorded and collected with special technologies (devices). Since these technologies differ for the operation principle, they can be categorized as: image-based and non-image-based technologies. Image-based technologies collect gestures as a sequence of images or video frames.

While the working principle of non-image-based technologies is to continuously record the position and movement trajectory of fingers, hand, palm through devices equipped with sensors (gloves, bracelets, rings, etc.) on the human body. In both cases, the data set obtained for gesture recognition is analyzed with the help of certain recognition methods. Obviously, all technologies have advantages and disadvantages, and there is no one technology applicable for all recognition methods. The most promising of image-based technologies are ToF (time-of-flight) cameras. The advantage of these cameras is that they can process data quickly. The most promising of non-image based technologies is sensor technology that does not come into contact with the human body, that is, it is not worn on body. These types of technologies do not require direct contact with the user. Not requiring mechanical contact with the object and the ability to work with different environments are the advantages of non-contact sensors. Non-contact sensors instantly identify the characteristic points of a hand image showing a gesture. Table 1 shows a comparative analysis of gesture recording technologies [8, 9]:

Table 1. Advantages and disadvantages of gesture recording technologies

№	Technologies	Advantages	Disadvantages
1.	Marker	Simplicity of computation	Coloring the user's body with markers
2.	Camera	Simplicity of device installation	Low reliability
3.	Stereo camera	Reliability	Complexity of computation
4.	ToF camera	High frame rate	Dependence on illumination and light reflectance
5.	Microsoft Kinect	Software support for body gesture recognition	Limitation of the distance between the device and the person (non-applicable for gesture recognition at a distance of more than 2 meters)
6.	Glove	Fast response	Discomfort of the device for the user
7.	Bracelet	Response speed, the width of the effect area	The presence of the device on the human body
8.	Wireless connection	Not requiring body contact	Low resolution accuracy

Segmentation of hand image detection from image. In the gesture recognition process, the next step following the data collection stage is the detection and framing of the hand showing the gesture from the collected raw data. The segmented part in the image is the framed part of the hand object.

The dynamic movement of the hand is recorded in the form of video frames and consists of a sequence of images. The figure below shows the detection of the hand from the images and its segmentation in a rectangle according to its position (Fig. 2) [10, 11].

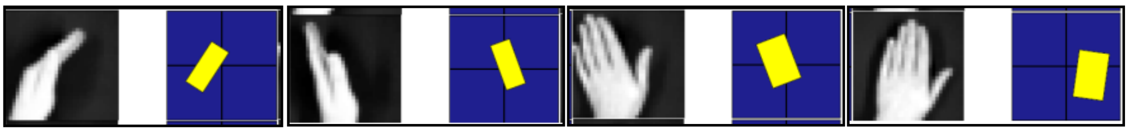


Fig. 2. Detecting and segmenting a hand pattern from an image

The segmentation process will be easier if the hand showing the gesture in the image is on a single background. But it is very unlikely that the hand will be on a single background. In most cases, the exposed skin part of the hand, face or body showing the gesture in the image can intersect, which causes errors in the segmentation and drastically reduces the recognition reliability. Furthermore, if the gesturing hand is perpendicular to the sensor plane, it will be easier to detect the palm and fingers from the recorded image. Correspondingly, gloves, rings, and other similar accessories have a negative effect on the recognition process and reduce the reliability of the algorithm [13].

Since the gestures are shown sequentially during communication, it is difficult to determine the beginning and end of gestures separately in segmentation work, that is, to accurately determine the boundaries of the trajectory of each gesture in the continuous movement of the hand. Therefore, the recognition of complex movements often results in a number of errors and inaccuracies. Another problem is that the gesture shown by the same person at different times can differ in terms of form and duration.

Feature detection. The next step in the gesture recognition process is feature detection. Accurate analysis of the segmented part, classification of the results and detection of key features must be performed in order to identify the gestures without error. The methods detecting the main features of the hand image showing the gesture can be categorized as: the methods based on the analysis of the external features of the hand; the methods based on the analysis of the three-dimensional model of the hand. Methods based on the analysis of the external features of the hand in the image analyze only the appearance (shape, position, etc.) of the object, rather than its physical properties. Methods based on the analysis of a three-dimensional model of the hand, on the other hand, analyze the detailed 3D configuration of the hand describing the gesture. The 3D configuration of the hand includes the position and direction of the key points of the palm and fingers in three-dimensional space [10].

Solution of the problem of feature detection are possible with Hidden Markov Model, Artificial Neural Networks, Viola-Jones algorithm, marker-based gesture recognition technique, Random forest method, K-nearest neighbors algorithm, Support vector method, etc.

Obtaining the result. The parameters of the gesture image are compared to the parameters of the images in the database to recognize the gesture and obtain the result. The "American Sign Language letter database" is used as this type of database [9]. The American Sign Language letter database consists of letters and 24 classes. Each training and test data contains the letters of the alphabet from A to Z marked with numbers in the range 0-25. Here, the letters 9=J and 25=Z are not included in the database because these gestures require movement. The training data contains 27,455 samples and the test data contains 7172 samples. The columns of the database are 784 (pixel1, pixel2….pixel784), and each row is a 28x28 pixel image with gray scale values in the range of 0-255. The figures in the database column representing the classes indicate the corresponding letter of the alphabet. Images of hand gestures are collected from several users.

4. Gesture recognition methods and their comparative analysis

In order for the gesture recognition system to be effective, high-quality light-sensitive cameras are required. The function of recognition systems includes the detection, processing and optimization of the hand image in the video frames recorded by the camera.

Hand gesture recognition is built using various recognition methods. The basis of the system is the extraction of key indicators of the image of the gesturing hand from raw materials obtained by web cameras or other devices. The raw frames are transferred via USB to the video frame processing software module. In the processing process, the hand image is selected as a segment, in other words, an object is assigned to continue the recognition process. After the object is detected, the features of the segmented part are searched to classify the gesture. A set of features must be sufficient to identify an object. Hand object data is processed and feature vectors are extracted with the methods. All feature vectors are used for classifiers' training and as a result the gesture's parameters are determined [13].

Hidden Markov Model (HMM). One of the most common approaches for gesture recognition is HMM. HMM is an efficient model for data analysis. HMM is a univeCSAl tool successfully used in gesture recognition. HMM-based gesture recognition methods represent each gesture as a set of states associated with three probabilities (initial, transition, observation). In general, the mathematical apparatus of HMM is a stochastic process controlled by probabilities. If the random process varies over time, then it is called a stochastic process [13]. A stochastic process is a homogeneous Markov chain with a finite number of states. It is hidden since the sequence of states is not observed. Specifically, the outside observer is aware of only the sequence of output symbols, not the state of the model. In a Markov chain, one process affects another random process. In HMM, the probability of transition from one state to another depends only on the state one step before, not on all previous states. An active state can produce a sequence of feature vectors at a moment. However, the exact sequence of states that the system goes through to create the set of generated feature vectors is impossible to know, because each state is determined by the probability density of the feature vector along with the transition probabilities. HMM selects the model with the best probability and classifies the gesture according to the appropriate category. Although HMM-based recognition systems select the model with the best probability, the uncertainty of the boundaries of the gesturing hand complicates the real-time gesture recognition. Thus, gesture recognition is based on the probability associated with it [15, 16, 17].

Selection of trainable features is important for successful gesture recognition with HMM training. The direction of the segment and the angle between the adjacent segments are taken as the key feature. The segmenting the hand image sequentially in the images in the video frame increases the accuracy of the recognition result. Proper segmentation partially solves the problem of computational complexity.

Figure 3 illustrates the general structure of HMM in the process of gesture recognition. HMM is based on a finite transition of N number of states {X1,X2,….,XN} called hidden. At each discrete time t, transitions between states occur not deterministically, but according to a probability law, and the transition probability is described by the ANN matrix. The matrix A of dimention N*N determines the probability of transition from one state to another. At a random time t, the next transition to the new state is performed and the output vector yt is generated. As a result, a sequence of feature vectors {y1,y2,...,yt} of length T, which determine the parameters of the gesture, is generated (Fig. 3) [18].

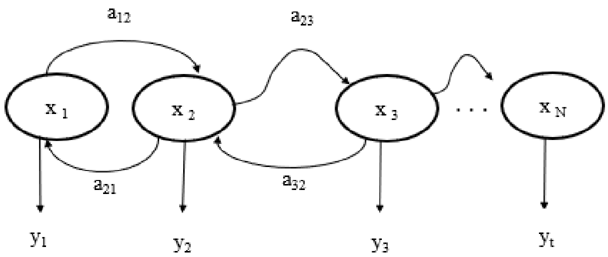


Fig. 3. General structure of HMM

HMM has several advantages: HMM's mathematical structure is simple; HMM structure allows modeling a complex chain of states; capacity of HMM to process sequences and images of different lengths increases the quality of modeling and has a significant positive effect on the recognition speed; HMM automatically resolves transitions with equal probability; HMM can work as hybrid with several algorithms. Along with the advantages of HMM, it also has several disadvantages: HMM requires a sufficient number of samples to calibrate the models; new parameters are required to be evaluated for each situation separately, which leads to an increase in calculations and an extension of the training period; HMM requires a large storage [19].

Artificial Neural Networks (ANN). To solve the problem of detection and classification of features of gestures, ANN is used along with other methods. ANN is a mathematical model built on the principles of organization and activity of biological neural networks. In the training process, ANN selects the features necessary for classification from the data set. In the recognition task, the characteristics of the gesture are normalized and transferred to the input of the ANN, and the gesture is obtained at the output. In order to increase the recognition quality, layers are added to the ANN basis (Fig. 4). ANNs differ from each other according to the training method and structure: Recurrent neural network, self-organizing Kohonen maps, Elman artificial neural network, Hebbian competitive learning, etc. [9, 20, 21].

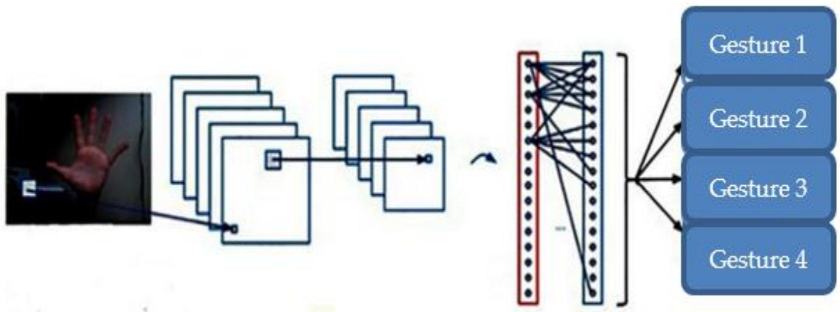


Fig. 4. The general structure of ANN

ANN includes two neural networks with different training processes. The first one is designed to localize the hand gesture in the original image. The training of the second model starts after the training of the first model is finished. As a result of the training of the second model, the two-dimensional coordinates of the main points of the hand are generated as an output signal, which is sufficient to recognize static gestures. If the gesture is dynamic, three-dimensional coordinates of the main points of the hand are generated as an output signal as a result of its training in the second model and applied to recognize complex gestures (Fig. 5) [20].

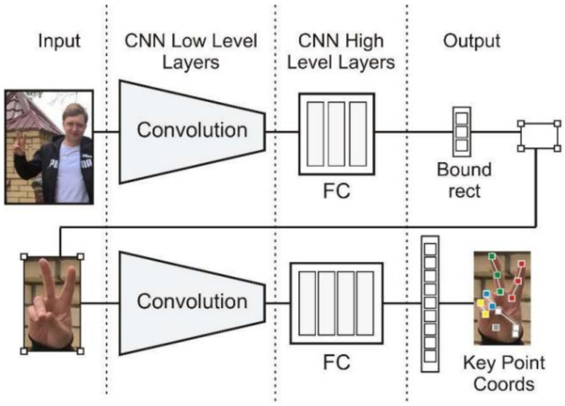


Fig. 5. The training process of ANN

ANN has a number of advantages: the use of ANN significantly improves the accuracy of gesture recognition; ANN is capable to detect static and dynamic gestures; ANN is resistant to changes in scale, displacement, rotation at a certain angle and other deformations of the hand showing the gesture; ANN is capable to detect complex non-linear dependencies between variables. Along with the advantages, ANN also has some shortcomings: the calculation process is complicated; not applicable if data is insufficient for training; not resistant to a weak lighting effect; multicolor background of the image negatively effects the method performance [21].

Viola-Jones algorithm (VCA). VCA is one of the effective methods for real-time object detection out of digital images and video sequences. The method builds a detector that can run in real time. Instead of image pixels, image features are used to build detectors. At the beginning, the input image of the hand showing the gesture is entered into the algorithm as a sequence of video frames. An image of the hand is detected and segmented in each frame. A separate descriptor is calculated for each frame to partially eliminate the difficulty in segmentation work. The hand image passes to the next stage as a sequence of Oriented Gradients Histogram (HOG) descriptors. HOG descriptors consist of a sequence of points indicating the key features of an image recorded by a computer vision system. At the stage of gesture recognition, the input sequence of descriptors is converted into an identifier command, and identification is performed through various methods [22].

For real-time hand gesture detection and recognition in video frames, VCA reduces the probability of error detection and expands system performance. This algorithm is based on the use of Haar-like features and consists of the following steps [23, 24]:

- The image is morphologically processed (image resizing, grayscale and image binarization). This speeds up calculations and reduces the probability of detection error;
- The required object is found out of the image using Haar features. The training process for Haar features selection is performed based on the AdaBoost algorithm. If a hand image is found in the conditional block of the algorithm, the gesture recognition process continues as follows:
  - Those that match the target object are classified as “true” and those that do not match the target object are classified as “false”;
  - Feature cascade is used to quickly remove areas in the image where the target object is not found.
- The specified Haar classifiers run, the parameters for that gesture are found and finally the gesture is identified.

VCA has a number of advantages: It is faster than the systems working with pixels since it is based on features in the process of object recognition; this algorithm differs from other algorithms in recognition efficiency and high speed; the probability of error in the operation of the VCA is low; if the object searched for in the image rotates at a small angle (up to 30 degrees), the algorithm recognizes it; it has a high processing speed due to the simplicity of the calculation; it is resistant to changes in hand orientation and scale; it is resistant to changes in lighting. Along with the advantages, VCA also has a number of limitations: the possibility of false detection if the gesturing hand in the image partially coincides with a part of the body or is on a colored background close to the skin color; requires the use of sufficient classifier cascades; not resistant to color change [17, 25].

Color segmentation algorithm (CSA). CSA is one of the widely used methods to detect a gesturing hand from an image or video frame. The algorithm performs the gesturing hand detection from the image in two stages: searching and selecting the areas corresponding to the human skin color in each pixel of the image (Fig. 6); morphological processing of selected areas in skin color. At the output of CSA, a morphological filter mask of skin-colored pixels in the image is generated. Morphological filter mask sequentially expands and compresses the image, evaluates the probability that each color refers to the skin color in the discretized color space, and removes parts that do not refer to the skin color. The detected skin-colored areas are segmented. The basic idea of CSA is to map probable skin-colored areas. However, not all segmented areas are needed in the gesture recognition process. Because, when performing a gesture, the hand can sometimes intersect with the open part of the human body. It is necessary to find and exclude segmented "wrong" areas in the image. Bayes classification is used for this. CSA uses found image segments and calculates their geometric characteristics [26, 27].

Fig. 6. Skin-color segmentation result

CSA has several advantages: simple calculation process; high processing speed; resistance to changes in hand orientation and scale; resistance to changes in lighting; capabilities to summarize and interpolate missing information. Along with advantages, CSA also has a number of disadvantages: potential for false detection of objects of color close to skin color; unapplicable in case of little data for training [26, 28].

As a result of the analysis carried out in this research, the advantages and disadvantages of the available methods in the field of gesture recognition are presented in Table 2.

Table 2. Comparative analysis of HMM, ANN, VCA and CSA methods for hand gesture recognition

Indicators	HMM	ANN	VCA	CSA
Mathematical structure	Simple	Complex	Simple	Simple
Processing speed	Fast	Fast	Fast	Fast
Resistance to lighting changes	Sustainable	Not sustainable	Sustainable	Sustainable
Detection rate of "wrong" areas	High	High	High	Low
Resistance to hand orientation	Sustainable	Sustainable	Sustainable	Sustainable
Resistance to changes in hand scale	Sustainable	Sustainable	Sustainable	Sustainable
Resistance to color change	Sustainable	Not sustainable	Not sustainable	Sustainable
Recognition accuracy	High	High	High	High

The low performance of object detection algorithms is caused by being strongly influenced by external factors such as poor lighting effect, lack of one-colored background of the image, difference in the object size, etc. These algorithms can be combined to increase the efficiency and accuracy of localization. Combining detection algorithms to track the image of the hand in the video frames can provide better quality, accurate and faster results.

[25] considers the issue of gesture recognition with the help of CSA and VCA. The results of both algorithms are compared separately and with a hybrid method. CSA's advantage is its speed, and its disadvantage is inaccuracy. Because the algorithm cannot unambiguously determine the probability that the segmented part of human skin color belong to the hand showing the gesture. VCA is used to overcome this problem. The idea of VCA is to process all segmented parts as a whole while passing through a cascade of classifiers. In the end, the algorithm determines the probability that the segmented object belongs to the hand showing the gesture, and deletes the segmented part that does not belong to the hand showing the gesture (Fig. 7) [25].

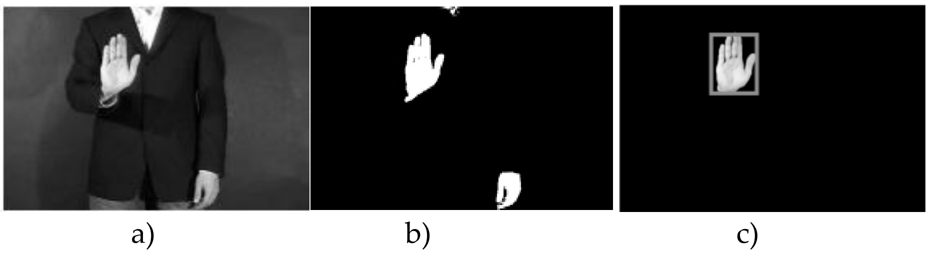


Fig. 7. Segmentation result: a) original image; b) segmented image; c) hybrid segmented image

[25] uses the dynamic gesture database of the ChaLearn Gesture Challenge for testing. The ChaLearn Gesture Challenge database contains a collection of independent packages. For training, each package contains 10 gesture images and 30-40 videos with a size of 230×240 pixels (Table 3.).

Table 3. Results of segmentation algorithms [25]

Algorithms	Working time, ms	positive result, %	error, %
Color segmentation algorithm	12,75	93,5	35,7
Viola-Jones algorithm	55,18	85,4	8,3
Hybrid method	20,71	93,5	4,1

As the table shows, the percentage of error is drastically reduced by the hybrid method.

As a result of the conducted research, in order to improve the effectiveness of the existing methods, a new approach for the recognition of sign language hand gestures is proposed in the presented article. In the presented work, the proposed system for real-time hand gesture recognition consists of two phases: the training phase and the test phase (Fig. 8).

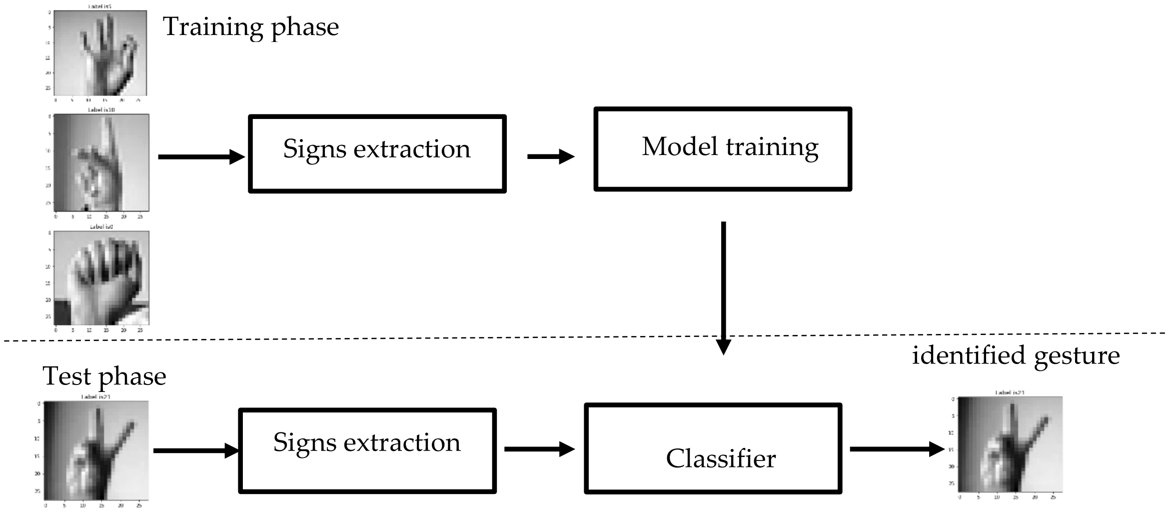


Fig. 8. The structure of the gesture recognition system

The training phase processes the training data of hand gesture images and extracts a feature vector. The feature vectors of all classes are then used to train the classifiers. In the test phase, the feature vectors of the images are re-extracted and identified by comparing them with the training data.

In the proposed architecture for gesture recognition, various machine learning techniques are developed to perform high-accuracy gesture recognition. The structure of the classification method based on neural networks is described in Table 4.

Total params: 91,024

Trainable params: 91,024

The parameters of the proposed model given in Table 4 are created as a result of experiments conducted on the "American Sign Language letter database" [29]. 100 neurons are used in the hidden layers of the neural network, and 24 neurons in the last output layer according to the classes in the database.

Table 4. Neural network model

Layer (type)	Output Shape	Param #
dense_11 (Dense)	(None, 100)	78500
dense_12 (Dense)	(None, 100)	10100
dense_13 (Dense)	(None, 24)	2424

The optimization function of the model is rmsprop, the loss function is categorical_crossentropy functions. Batch_size=50, epochs=30 are taken during the model training. The experiments are performed in the Python programming environment.

Table 5 presents the results of the experiments conducted on the American Sign Language letter database.

Table 5. Evaluation of the model effectiveness

Classes	Accuracy	Precision	Recall	F1-score
0	1.0	0.82	1.00	0.90
1	0.95	0.92	0.95	0.94
2	0.93	0.87	0.93	0.90
3	0.92	0.90	0.92	0.91
4	0.95	0.94	0.95	0.94
5	0.98	0.75	0.98	0.85
6	0.84	0.82	0.84	0.83
7	0.80	0.99	0.80	0.89
8	**0.64**	0.90	0.64	0.74
9	0.79	0.88	0.79	0.83
10	0.90	0.82	0.90	0.86
11	0.89	0.80	0.89	0.84
12	**0.62**	0.61	0.62	0.61
13	0.83	0.95	0.83	0.89
14	0.98	1.00	0.98	0.99
15	0.88	0.72	0.88	0.79
16	0.72	0.66	0.72	0.69
17	**0.48**	0.71	0.48	0.57
18	0.88	0.70	0.88	0.78
19	**0.54**	0.64	0.54	0.58
20	0.71	0.84	0.71	0.77
21	0.75	0.50	0.75	0.60
22	0.84	0.78	0.84	0.81
23	**0.66**	0.90	0.66	0.76

As Table 5 shows, the proposed model is capable to recognize most letters of the alphabet with high accuracy. Thus, the first letter of the alphabet is recognized with high accuracy and received a value of 1.0 according to the accuracy metric. The model could not accurately recognize letters 8, 12, 17, 54 and 23 and received values of 0.64, 0.62, 0.48, 0.54, 0.66 for each class according to the accuracy metric.

5. Conclusion

The study explained the working principle of image-based and non-based technologies that collect by tracking hand gestures, showed their comparative analysis and highlighted their pros and cons. Based on the collected information, the working principles of the Hidden Markov Model, the Artificial Neural Network method, the Viola-Jones algorithm and the color segmentation algorithm, which are methods of solving the problem of gesture recognition, were interpreted and comparatively analyzed. Experimental results of the hybridization of color segmentation and Viola-Jones algorithms to obtain a fast and partially accurate result with less error were shown as an example. The article developed a system based on artificial neural networks for automatic recognition of hand gestures. The effectiveness of the developed system was tested on an open research database and high performance was obtained.

References

1. World Hearing Day on March 3, 2022 will be held under the slogan "To hear all your life, take care of your hearing!". 25 February 2022, https://kurgan-med.ru/news/vsemirnyy-den-slukha-3-marta-2022-goda-proydet-pod-lozungom-chtoby-slyshat-vsyu-zhizn-beregite-svoy-/
2. Yeo, H.S., Lee, B.G., Lim, H. (2015). Hand tracking and gesture recognition system for human-computer interaction using low-cost hardware. Multimedia Tools and Applications, 74(8), 2687-2715.
3. Prozorova, E.V. (2007). Russian sign language as a subject of linguistic research. Questions of Linguistics, 1, 44–61. (Russian)
4. Kataev, M.Yu., Shirokov, L.V. (2013). Method for determining hand gestures observed with a video camera. Reports of TUSUR, 1(27), 45-49. (Russian)
5. Sign Language Lessons. 23 November 2022, https://www.pinterest.com/pin/742460688548321728/?mt=login
6. Muratov, E.R., Nikiforov, M.B., Skachkov, A.M. (2020). Artificial intelligence in the tasks of recognizing dynamic gestures. Modern information technologies and IT education, 16(4), 883-892. (Russian)
7. Bauer, A.V. (2015) Real-time gesture recognition implementation. Microsoft Technologies in the Theory and Practice of Programming: Proceedings of the 12th All-Russian Scientific and Practical Conference of Students, Postgraduates and Young Scientist, Tomsk, March 25-26, (pp. 56-58) (Russian)
8. Gesture recognition for interaction with AI: from theory to the latest achievements. 23 November 2022, integral-russia.ru
9. Abdullayeva, F.J., Gurbanova, K.S. (2022). Sign Language Hand Gesture Recognition Method based on Machine Learning, The 16th IEEE International Conference on Application of Information and Communication Technologies, Washington DC, 12-14 october 2022, 1-5.
10. Nguyen, T.T., Bolotova, Yu.A., Spitsyn, V.G. (2012). Real-time processing of video sequence data based on a hierarchical time network. Scientific Bulletin of NSTU, 2(47), 33-43. (Russian)
11. Chudnovsky, M.M. (2014). An algorithm for recognizing human hand gestures on a video sequence in real time for the implementation of human-machine interaction interfaces. Siberian Journal of Science and Technology, 3(55), 162-167. (Russian)
12. Muratov, E.R., Nikiforov, M.B., Skachkov, A.M. (2020). Artificial intelligence in the tasks of recognizing dynamic gestures. Modern information technologies and IT education, 16(4), 883-892. (Russian)
13. Chuikov, A.V., Vulfin, A.M. (2017). Gesture recognition system based on neural network Technologies. Bulletin of Ugatu, 3 (77), 113–122. (Russian)
14. Rustamov, G.A. (2012). Automatic regulation theory. Textbook for higher technical schools, 729. (Azerbaijan)
15. Potopakhin, A.A. (2017). Analysis of software tools and methods for dynamic gestures recognition. Computer science, control systems, mathematical and computer modeling: Proceedings of the 8th International Scientific and Technical Conference within the framework of the Third International Scientific Forum of the Donetsk People's Republic, Donetsk, May 25, 2, (pp.416-419) (Russian)
16. Yu, S. (2010). Hidden semi-Markov models. Artificial Intelligence, 174, 215-243.
17. Bauer, A.V. (2014). Dynamic Hand Gesture Recognition Based on Hidden Markov Models. Modern Techniques and Technologies: Proceedings of the 10th International Scientific and Practical Conference of Students, Postgraduate Students and Young Scientists, Tomsk, April 14-18, 2, (pp. 145-146) (Russian)
18. Gefke, D.A., Zatsepin, P.M. (2012). Application of hidden Markov models for recognition of sound sequences. Scientific journal "Izvestiya ASU", 2(73), 72-76. (Russian)
19. Nikolaev, E.I., Zakharov, V.V., Zakharova, N.I. (2017). Technology for training deep models on a synthesized data set for recognition and detection of gestures. Vestnik SevKavGTI, 4, 237-240. (Russian)
20. Dolgov, G.A. (2021). Choosing the optimal neural network architecture for gesture recognition. MNSK-2021: Proceedings of the 59th International Scientific Student Conference, Novosibirsk, April 12–23, (pp. 87) (Russian)
21. Matveev, A.S., Shestakova, V.V., Prokhorov, A.V. (2018). Tasks of professional orientation in the context of changing the educational system of undergraduate studies on the example of Tomsk Polytechnic University. Power industry through the eyes of youth: Proceedings of the 9th International Youth Scientific and Technical Conference, Kazan, 1, (pp. 69–72) (Russian)
22. Arindam, M., Abe, T., Takayuki, O., Koichiro, D. (2011). Hand Gesture Recognition using Histogram of Oriented Gradients and Partial Least Squares Regression. MVA 2011 IAPR Conference on Machine Vision Applications, Nara, Japan, June 13-15, (pp. 479-482)
23. Fan, N.Kh., Buy, T.T., Spitsyn, V.G. (2013). Real-time gesture recognition on a video sequence based on the application of the Viola-Jones method, the Camshift algorithm, the Wavelet transform and the principal component method. Tomsk State University Bulletin, Management, Computer Engineering and Informatics, 2(23), 102-111. (Russian)
24. Yang, J., Xu, Y. (1994). Hidden Markov Model for Gesture. Recognition: Technical Report CMURI-TR-94-10. – Pittsburgh, PA: Carnegie Mellon University, Robotics Institute, 27.
25. Nosov, A.V. (2019). Algorithm for localizing an image of a hand from a video sequence.REDS: Telecommunication Devices and Systems, 9(1), 22-26. (Russian)
26. Favorskaya, M.N., Zotin, A.G. (2012). Methods for motion estimation in video sequences and their software implementation. Textbook allowance; Sib. state aerospace un-t. Krasnoyarsk, 132. (Russian)
27. Bolotova, Yu.A., Fedotova, L.S., Spitsyn, V.G. (2014). Algorithm for detecting areas of faces and hands in an image based on the Viola-Jones method and the color segmentation algorithm. Fundamental Research, 11(10), 2130-2134. (Russian)
28. Kozyr, M.V. (2017). Algorithm for detecting areas of faces and hands in an image based on the Viola-Jones method. Proceedings of the Second Regional Scientific and Technical Conference, Omsk, June 06–07, (pp. 26-30) (Russian)
29. ASL Alphabet: Image data set for alphabets in the American Sign Language. 23 November 2022, https://www.kaggle.com/datasets/grassknoted/asl-alphabet