

# THE SIGN LANGUAGE RECOGNITION PROJECT FOR DISABLED PEOPLE WITH HEARING VIOLATION ABOUT

**K.V. Genchel, V.I. Voronov, L.I. Voronova**

*Moscow Technical University of Communications and Informatics, Moscow, Russia (111024, Moscow, Aviamotornaya Str. 8a.), e-mail: [voronova.lilia@ya.ru](mailto:voronova.lilia@ya.ru)*

---

**Abstract.** The article describes the program complex "Surdotelephone" for recognizing the sign language of disabled people with hearing impairment based on machine learning algorithms.

---

**Keywords:** Innovation, science, sign language, neural networks, machine learning.

Adaptation of disabled people in modern society is an important social task. Various events are held on this topic in Russia at the government level. In particular, the Social project of "United Russia" party, dedicated to IT-socialization of disabled people, held in Moscow International Conference and Conference on Electronic Technologies [1]. Understanding the importance of these tasks, at Moscow University of Communications and Informatics there are also developments of this kind [2]. The group of undergraduates and teachers of the Department of Intelligent Systems in Control and Automation is developing a large project related to the implementation of the so-called "Surdotelephone" which allows using modern information technologies and methods of Data Mining (DM) to convert a sign language into text or speech.

The goal of the project is to develop a software package that allows people with hearing impairment to communicate with people who do not know the Russian Sign Language (RSL), as well as make emergency calls in critical situations, such as firefighters, ambulance, etc.

**The essence of the program's work.** The user "calls" to the appropriate service, not directly, but through the application "Surdoreceiver" (for example, on the smartphone) and with a set of gestures informs about the critical situation before the video camera. The project provides recognition and transmission of information, with the addition of individual information about the caller and GPS data.

**Structure of the software complex "Surdotelephone".** The software complex consists of three large physical blocks: a smartphone (another device that has a video camera), a cloud platform, and a recipient device. In Fig. 1 shows the architecture of the complex.

**The "Surdoreceiver" application** (located on a smartphone, tablet, computer). It provides reception of the sign information, its transfer to the cloud platform for recognition.

**Cloud platform** - provides reception, storage and processing of user data (signs) received, as well as transmission of recognition results to their intended use.

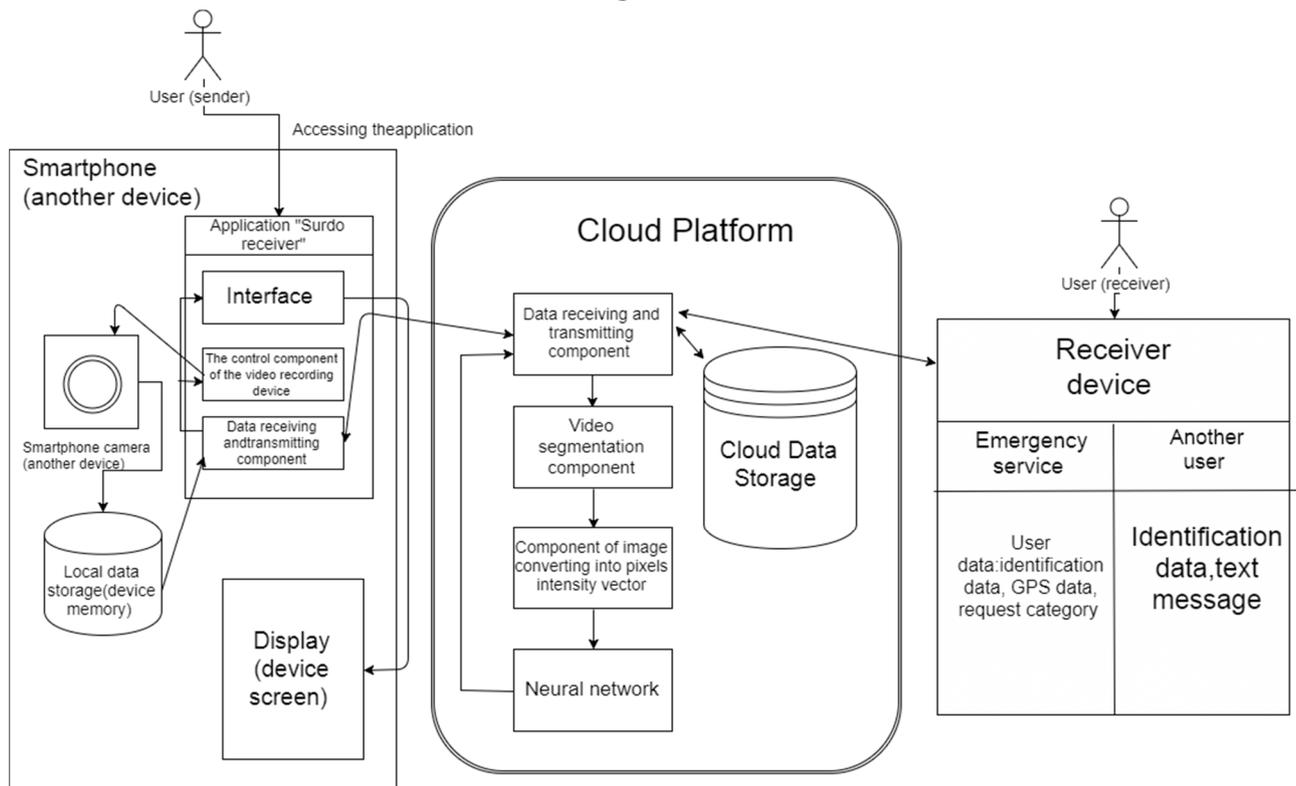


Figure 1. Architecture of the "Surdotelephone" software complex.

**The recipient's device** provides reception of the recognized information (or user data). This can person with normal hearing or emergency services. In an emergency, an additional message includes the geolocation and user identification data, which is also delivers to emergency services.

Through the **User Interface**, the user on his device (smartphone, PC) starts the application. The user interface works in two modes:

- Automatic recording of video information at the push of a button,
- Pre-identification - filling in a form with personal information.

**The control component of the video recording device** (web camera, smartphone camera, etc.) accesses the video camera of the smartphone, includes a recording, a continuous stream of video is recorded in the memory of the local data store with video blocks for 5 seconds in mp4 format.

**Data receiving and transmitting component.** Provides block-by-block video translation to the cloud (RTCP protocol). Gets data from the cloud and displays information on the device's screen for the user.

The outgoing data stream is compressed to save memory, increase data transfer speed, and simplify segmentation.

The incoming data is written to the cloud storage, to the directory, according to the user ID (Fig. 2).

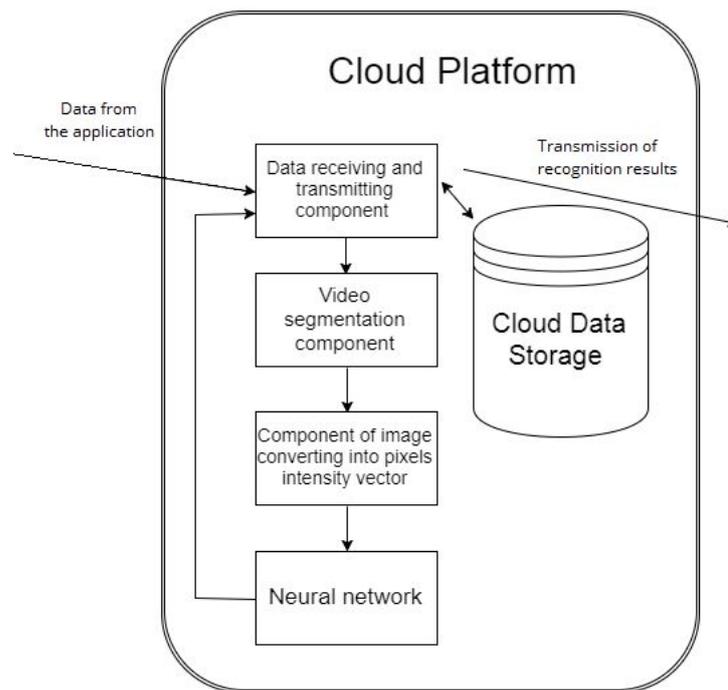


Figure 2. Cloud platform.

#### Components of the cloud platform:

- **Receiving and transmitting.** Accepts incoming data according to the identifier, passes them to the segmentation, receives the result of recognition of the neural network and sends it to the user or to emergency services [3].
- **Video segmentation.** Produces segmentation of the received video into JPEG frames, according to a given algorithm. It transmits frames to the transformation component in the intensity vector.
- **Transformations into an intensity vector.** Produces transformations of received frames, into pixels intensity vector and their recording in the storage. Sends data to the input layer of the neural network.
- **Recognition (Neural network).** Produces recognition of incoming information, transmits the result of components and obtains user data [4].
- **Cloud storage.** It stores initial, intermediate and service information. The information is sorted and placed according to the user identifier.

To create a training set, you need to shoot a video and apply the program to segment it into frames, in which to set the necessary parameters, and place the received frames in a directory whose name should coincide with the recognized letter.

The next step, a script, which labels the images at the name of the folders to assign a gesture image to the specific letter is run.

Then a script that brings the photos to the correct size and translates the image into a vector is run. And using the next script, the model is trained on this data.

Figure 3 shows the result of the recognition program. The program sequentially displays the original image, and then the image with the result of the work (the class of the recognized letter).

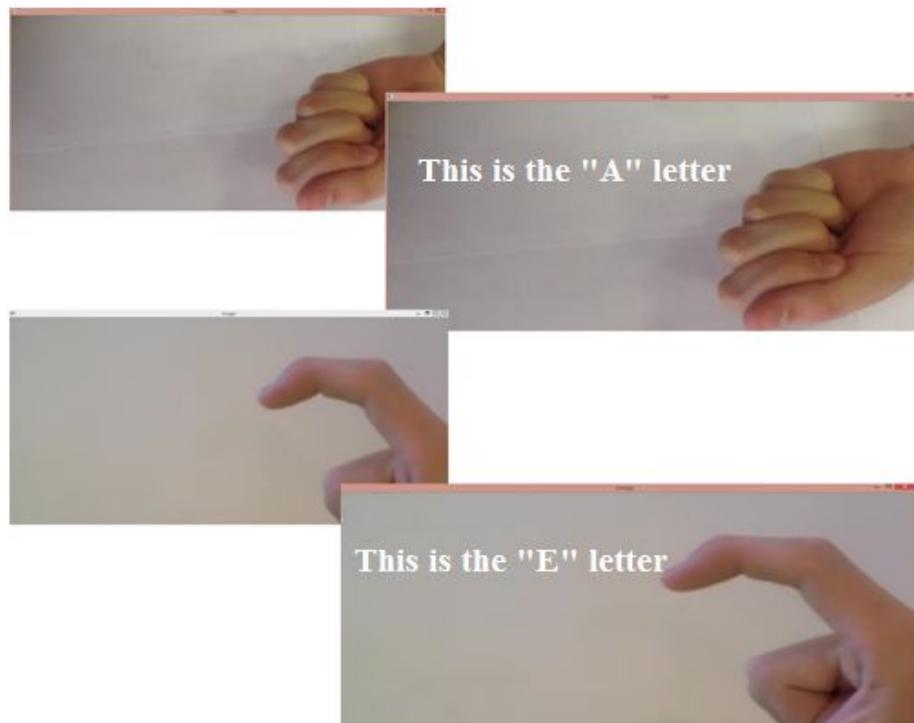


Figure 3. Result of program execution.

Currently, work continues on increasing the functionality of the software package and automating the main and supporting processes, such as segmenting video and storing frames in the correct directory. To recognize the dynamic sequences of gestures that are used by hard-of-hearing people in reality, both convolutional and recurrent neural networks are being developed [4].

## References

1. <http://pnu.edu.ru/ru/news/2015-09-22-IT1/> - Interregional scientific-practical conference
2. Voronov V.I., Voronova L.I. O povyshenii rezul'tativnosti masterskikh programmah v usloviyah innovacionnoj ehkonomiki / Voronov V.I., Voronova L.I., V knige: Innovacionnye podhody v nauke i obrazovanii: teoriya, metodologiya, praktika Monografiya. Pod obshchej redakciej G.YU. Gulyaeva. Penza, 2017. C. 35-44.
3. Tolmachev R.V., Voronova L.I. Razrabotka prilozheniya dlya kontent-analiza internet-publikacij /Telekommunikacii i informacionnye tekhnologii. 2016. T. 3. № 1. C. 104-107.
4. L.I.Voronova, V.I.Voronov Machine Learning: regressionnye metody intellektual'nogo analiza dannyh: uchebnoe posobie / MTUCI-M., 2017.-92c