IOSR Journal of Dental and Medical Sciences (IOSR-JDMS)
e-ISSN: 2279-0853, p-ISSN: 2279-0861.Volume 16, Issue 4 Ver. III (April. 2017), PP 107-113
www.iosrjournals.org

Using the Artificial Neural Networks for Identification Unknown Person

Kozan N.M. [1], Kotsyubynska Yu.Z. [1], Zelenchuk G.M. [1]
[1](*Department Of Pathology And Forensic Medicine*
SHEE "Ivano-Frankivsk National Medical University" Ukraine)

Abstract: *This article contains the information about the use of the original software Dermatoglyphics For Prediction (DFP) to identify the total phenotypic traits of a person (ethno-territorial belonging, gender, antroposcopic and anthropometric parameters based on the dermatoglyphics parameters.*
The object of the study were anthropometric, antroposcopic and dermatoglific parameters obtained through the description and measurement of persons male and female, aged 18-60 years who belonged to Boiko (85), Lemko (74) and Hutsul (126) ethnic groups, and 90 to the control group. Prints of comb picture of the toes obtained by scanning the scanner Futronic's FS80 USB2.0 Fingerprint Scanner using the program ftrScanApiEx.exe. with following transferring of data to the personal computer. For statistical processing of the obtained data we use STATISTICA 12 from the company StatSoft. Construction of the neural networks was carried out using the Neural Networks. Also as a result of realizing work we received the core functionality of the program and is designed the GUI (Graphical User Interface) to alleviate data entrying and the possibility of using designed by us complex by other researchers. The work resulted in the obtained data which allowing with 89-93% probability of to predict the total phenotypic features of person based on the dermatoglyphics parameters of fingers.

Keywords: *Forensic medicine, dermatoglyphics, identification of the person, artificial neural networks*

I. Introduction

One of the open questions in forensic medicine, remains a question posed by the problem of identifying unknown persons and carrying identification fragmented and impersonal corpses. Today, the most popular method is the method of DNA identification, which, despite the accuracy of the initial results is expensive and not available diagnostic method for all. Therefore, in recent years become more popular method of dermatoglific, as a simpler and cheaper method to use. The study of inheritance patterns skin of the fingers and feet is to establish disputed paternity and family relationship with the forensic examination of unknown persons. Dermatoglyphics method can have a significant advantage in determining relationship to other genetic methods. In particular, with help of the computer program, it can become widespread routine method of diagnosis in forensic medicine. Unfortunately, for today such a program is not exist. The closest analogues of it can be dactyloscopic computer programs which used in criminalistics for the purpose of automated identification of the fingerprint. Such kind of programs we can use today. With appropriate software revision, these programs probably could be useful in medical genetics and anthropology. In particular, Fokyn V.A. and co-author (2002), was created the program "KYBERINFORM" of medical genetic destination for analysis of main dermatoglyphics parameters which can serve as a complement to the programs which implemented their automatic identification, such as criminalistic dactyloscopic program "Songdo" (version 4.2 of "Pathfinder"), which is widely used for forensic and criminalistic purposes since the early 90's and is well prove itself. This program "Songdo" capable, in particular, automatically identify and classify basic types of papillary patterns of the terminal phalanges of both hands. This program distinguishes 19 types of such pictures and classifies them into arches, loops and curls on System Galton-Henry. Results of identifying each individual person can be seen on the computer screen. Based on these images, we can quantify the degree of asymmetry of the right and left hands, to calculate basic dermatoglific indicators for the control and experimental groups, which may be useful for the mass population anthropometric and genetic research [1]. Recent years become increasingly popular the artificial neural networks on which created identification program, including and forensic [2].

An artificial neural network (ANN) - a software - hardware complex built on a mathematical model that resembles the principle of operation of biological neural networks. This concept formed the study of processes occurring in the brain [3,4]. An artificial neural network (the neural network) is a system of connected and interacting processors (neurons). Neuron (basic neural network element) is a simple computer processor which can handle perceive and transmit simple signals (information). When you combine a large number of neurons in a network, the system can solve non-trivial task. Neural network equally well appropriate for solving both linear and nonlinear problems. Neural network, depending on your architecture [5,6] can be divided by level of difficulty. The basic type of neural network is a network of direct distribution - that is network to which

signals propagating only in one direction (for complex problems using recurrent neural network model). Neural networks have been widely applied in various types of complex problems needing analytical calculations of similar to those that the human`s brain does. Among the major classes of problems for neural networks can be distinguished classification and prediction. When we are talking about classification we talk about the searching and partitioning of the database for some parameters, in anticipation - the opportunity to provide some event or step for a given array input.

Aims and Objectives

The aim of our study was to investigate the current state of computer identification applications, such as artificial neural networks.

II. Material And Method

The object of the study were anthropometric, anthoposcopic and dermatoglific parameters obtained through the description and measurement of those men and women aged 18-60 years who belonged to Boyky (85), Lemky (74) and Hutsuly (126) ethnic groups, and 90 control individuals. Prints of comb pattern of the toes obtained by scanning the scanner Futronic's FS80 USB2.0 Fingerprint Scanner using the program ftrScanApiEx.exe. followed by the transfer of data to a personal computer. For statistical processing of the obtained data we use STATISTICA 12 from the company StatSoft [7]. Construction of neural networks was carried out using Neural Networks [8].

III. Results and discussion

While creating real system for processing and forecasting, the main attention should be paid to the choice of network architecture, method of training and systematization of the input data. Also important is the choice of software package which will set up and operate neural network. For statistical analysis of their data, we use STATISTICA 12 from the company StatSoft. This package allows to obtain a wide range of statistical data and build their own neural network. In our studies, we often turned to such STATISTICA capabilities as finding the correlations between variables, tables of probability distributions, etc. [9].

Also very helpful is to visualize the data, for example the construction of three-dimensional graphs of signs distribution (Fig.1) and the graph of the dependency between different values (Fig. 2).

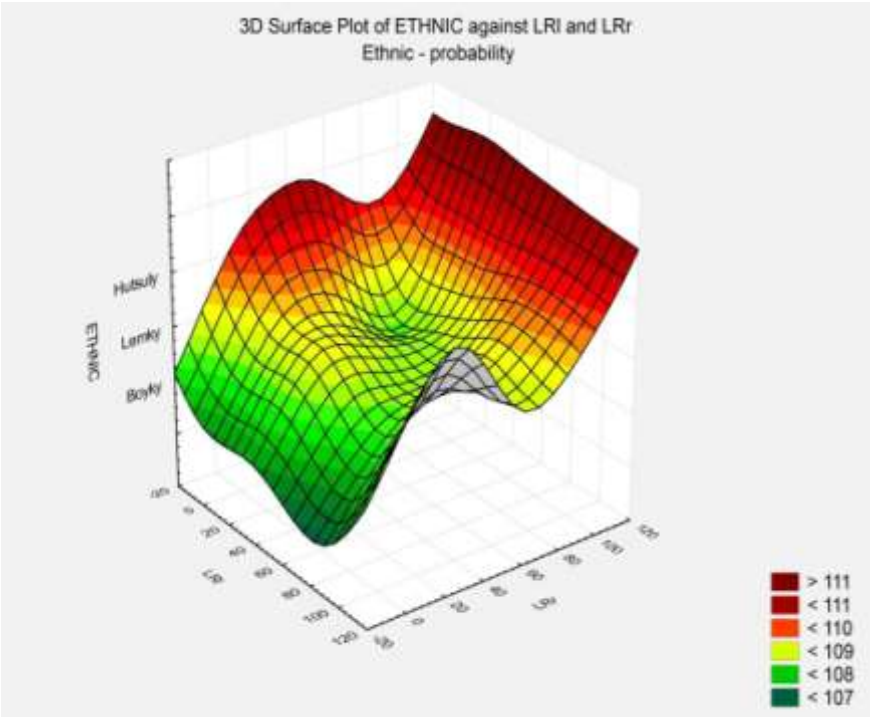

Fig.1. Distribution of probabilistic dependence (surface probabilities between ethno-racial implementing) and the frequency of manifestation LR on the fingers of the right (LRr) and left (LRl) arms, respectively.

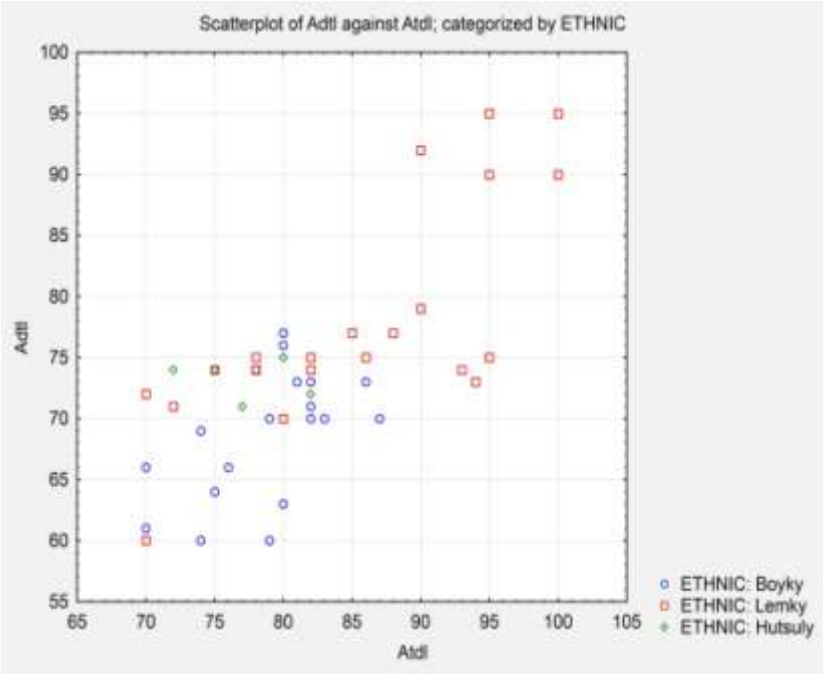

Fig.2. Communication between angles Adt and Atd on the left hand for different ethnic groups.

The basic instrument for forecasting the dermatoglific dependencies we chose the neural network [10]. To construct a new neural network in the program STATISTICA you just need to enter the relevant input data in the table and choose an item from the menu Neural Networks [8].

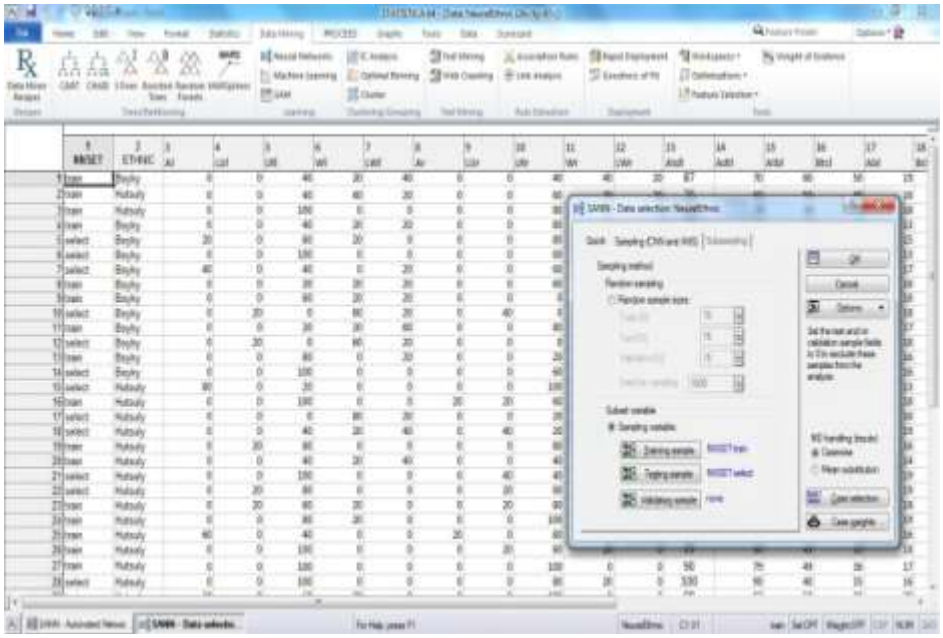

Fig.3. The input data (left) and setting (right) for instruction neural network.

As shown in Figure 3 except for inputs data (Al, LUl, Wr.) we added one more category codes, namely - "train" for training and "select" to control subsample. After we have selected the target category (in our case, Ethnic) and continuous input variables (Al, LUl, Atdl.) we have a choice between three strategies for building models (ANS - automated neural network, CNN user-defined neural network and Subsampling - method of repeated subsample). For simplicity of implementation we choose automated neural network. By using the variable NNSET we ask training (train) and control (select) set. Then we choose the strategy of forming the sub-samples with input parameters: 5 random sub-samples with the relative percentage parity of 70-15-15% - study-control-test subsample, respectively.

Function of activation we choose Logistic, Tanh (logistic, hyperbolic,) to output neurons and Logistic, Tanh, Exponential (logistic, hyperbolic, exponent) for latent neurons. Also as output parameters we specify the number of latent neurons - from 30 to 50, the attenuated regularization of neurons weights (weight decay) - from 0.001 to 0.01 (for latent layers), the number of networks for education - 20. All these data are experimentally selected and the tasks can vary within wide limits depending on the type and complexity [11]. After providing the training, we have 10 models of neural network with different indicators of performance (percentage of correct classification - the closer the number is to 100 the better model classifies the data) learning and test performance (Fig. 4).

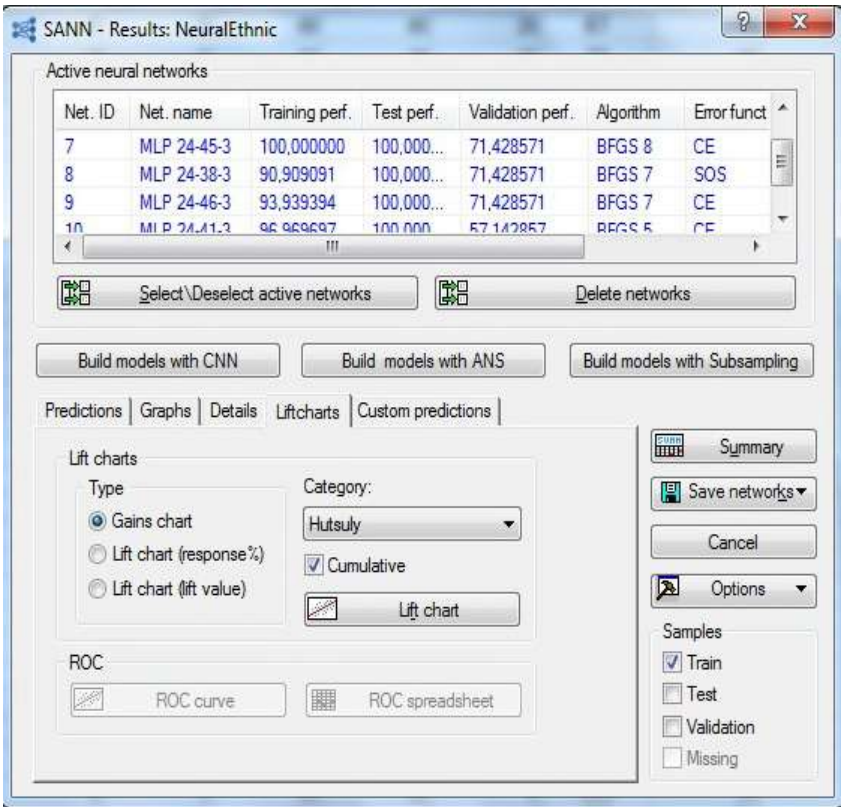

Fig.4. A set of neural networks with different architectures and data of performance training.

According to the received data we choose the network which has the maximum performance of studies. In this case this is the network № 10 with architecture MLP 24-42-3 (24 entrance, 42 latent, 3 output neurons). If the result does not suit us (such as 85% for network №1 actually quite a few), we turn to models CNN, where we expose specifying input data - namely, the type of network RBF (radial basis function), 20 networks for training, amount of neurons 50, the amount of periods 10,000 is turned on the interactive learning (for visual analysis of errors). Then perform similar operations to the moment when the test performance and productivity training will not establish us. At the end we got neural network №7 with architecture MLP 24-45-3, which with the data of performance satisfies our requirements figure 5.

Summary of active networks (NeuralEthnic)

Index	Net. name	Training perf.	Test perf.	Validation perf.	Training algorithm	Error function	Hidden activation	Output activation
1	MLP 24-43-3	90,9091	100,0000	42,85714	BFGS 7	SOS	Tanh	Identity
2	MLP 24-46-3	90,9091	100,0000	57,14286	BFGS 8	SOS	Logistic	Tanh
3	MLP 24-44-3	90,9091	85,7143	57,14286	BFGS 8	Entropy	Exponential	Softmax
4	MLP 24-46-3	93,9394	100,0000	57,14286	BFGS 6	Entropy	Exponential	Softmax
5	MLP 24-48-3	90,9091	100,0000	71,42857	BFGS 7	SOS	Tanh	Tanh
6	MLP 24-33-3	93,9394	100,0000	42,85714	BFGS 7	SOS	Tanh	Tanh
7	MLP 24-45-3	100,0000	100,0000	71,42857	BFGS 8	Entropy	Logistic	Softmax
8	MLP 24-38-3	90,9091	100,0000	71,42857	BFGS 7	SOS	Tanh	Tanh
9	MLP 24-46-3	93,9394	100,0000	71,42857	BFGS 7	Entropy	Logistic	Softmax
10	MLP 24-41-3	96,9697	100,0000	57,14286	BFGS 5	Entropy	Identity	Softmax

Fig.5. The neural networks which are based on the model of CNN with clarifying data.

For correct selection of the resulting network we should also pay attention not only to test performance, but also on the matrix of errors figure 6:

	ETHNIC (Classification summary) (NeuralEthnic) Samples: Train				
		ETHNIC-Boyky	ETHNIC-Hutsuly	ETHNIC-Lemky	ETHNIC-All
7.MLP 24-45-3	Total	14,0000	15,0000	4,0000	33,0000
	Correct	14,0000	15,0000	4,0000	33,0000
	Incorrect	0,0000	0,0000	0,0000	0,0000
	Correct (%)	100,0000	100,0000	100,0000	100,0000
	Incorrect (%)	0,0000	0,0000	0,0000	0,0000

Fig.6. Matrix of errors for network №7

Analysis of the neural network includes building trust levels (confidence levels) on all samples. As we can see from Fig.7, shown in red color is incorrect prediction belonging to a particular category.

Fig.7. Confidence level of neural networks №5 and №7.

Also pay attention to the importance of analysis (sensitivity analysis) variables that are part of our neural network (Fig. 8). The data analysis shows that feature LWl (W on the left hand) is more important than in forecasting LU, Ar. That is, in the further analysis when we will reconstruct models, we should pay more attention to sign LWl.

	Sensitivity analysis (NeuralEthnic) Samples: Train								
Networks	LWl	Cdl	Wl	LRl	Adtl	Atbr	Abr	Cdr	Bcr
7.MLP 24-45-3	3,194162	2,735918	1,859091	1,640607	1,508309	1,486017	1,410322	1,406043	1,375629

Fig.8. Analysis of importance (sensitivity analysis) of the input variables.

A lot of data about quality of classification the network we can get from the elevator cards (Liftcharts) and various graphs which makes it possible to build a software package STATISTICA (Figure 1-2).

After that when we establish the working version of the of neural network the program STATISTICA can save it as a code of basic programming languages: C and Java. In order for practically realize the resulting neural network, we used the language Java. After opening STATISTICA generated code in the integrated development environment (in this case the popular Java IDE - IntelliJ IDEA [12]), is easy to see that the code is well structured and easy to understand.

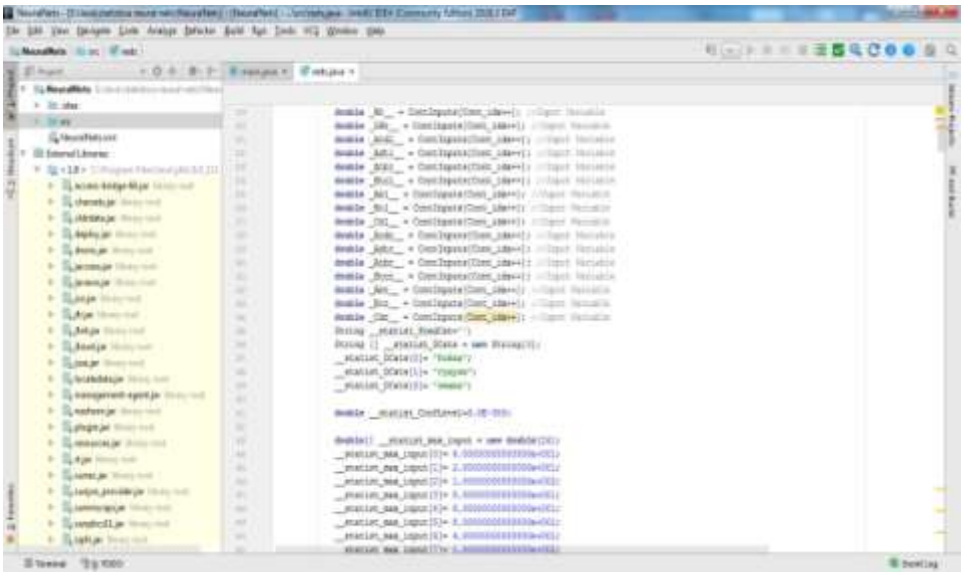

Fig.9. The initial Java code of neural network, created in STATISTICA and edited in IntelliJ IDEA

As shown in Figure 9, in order to get the desired response, you must specify neural network input fractional values (Al, LUl, Atdl.) then the network will conduct their analysis. At the output we get the text string (String) from the array of possible answers (in this case "Hutsuly", "Boyky".).

To expand capacity analysis and forecasting need to increase the number of variables that will handle the neural network. Above, we briefly describe how to create a network for data papillary drawings and signs palms ethno-racial groups (Boyky, Hutsuly, Lemky). If we add to the above table (Fig. 3) another variable, then the neural network will be able to more accurately predict the belonging to this or that category (ethnic group Of course, with this will increase complexity of construction and process the results of calculations neural network.

For practical realization the neural network with the sufficiently large amount of input variables we must or have to edit the resulting code (generated in STATISTICA), or enter into a network all the relevant data of the person ethos of which we define. (But how lead itself such neural network when we introduce only the data of angles on the right hand? This neural network give an error.)

In this formulation of the task, there is a simple and effective way of realization a multi-function: divide it into small pieces. In other words to allocate to the categories the input data (eg. category 1 - all data of the right hand, category 2 - all data of both hands, category 3 - data of both legs ... etc) create a separate neural network of each planned categories, and then unite them in a program that will determine the category (or categories) of input data and accordingly make informed prediction on this base (Fig. 10).

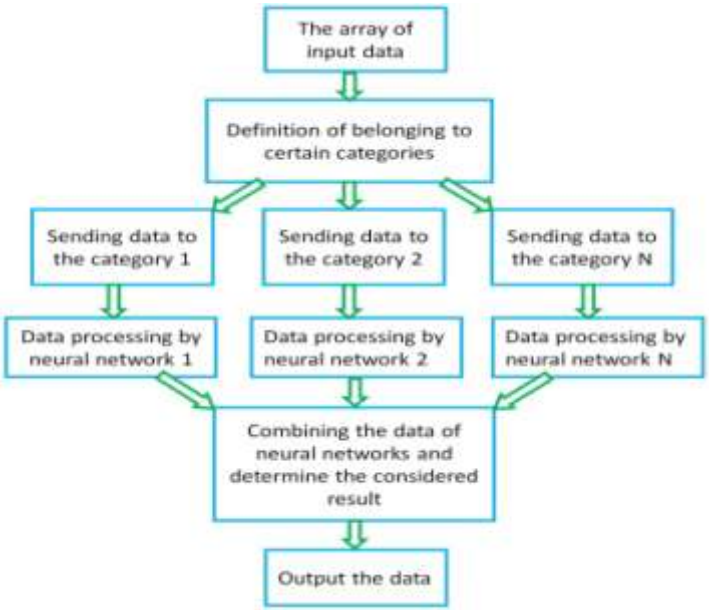

Fig.10. Block diagram of the program for the prediction of ethno-racial identity.

IV. Conclusions

Thus, we have the core of functionality of the program. In order to be able to fully use this core we developed a GUI (Graphical User Interface) to facilitate data entry and the possibility of using the complex developed by us other researchers. In general, we have developed a program (Dermatoglyphics for Prediction (DFP)), consists of a core based on neural networks (Fig.8, Fig.9) and shell created mainly through classes javafx.application. *, Javafx.stage. *, Java.awt. * and the large amount of classes Java Development Kit 1.8. [13].Graphical user interface created using Java FX 8. This framework allows you to quickly and efficiently realize the graphical shell of basic code, and also it is a multi-platform (programs which are written in JavaFX are launched on different operating systems) [14]. Also, to improve the design and following generally accepted standards Material Design we used the library JFoenix [15]. Graphic objects created using this library meet all current trends in programming for Desktop (in this case Windows).

For start our program (DFP) you just need to run the executable file DFP.jar (eg. the program no need to separately install). In the main window you need to choose the appropriate drop-down menu, and introduce at them the input data of unknown person. In case of insufficient of incoming data the program pointing to the error of input. After entering the data necessary to press "Submit". The program will process the data in a separate window give the category (ethno-racial belonging, etc.) to which the unknown person is refer, the probability of correct classification, etc. As we can see from the description, to work with the program (DFP) you do not need special skills, and the work of the program is very fast (due to the fact that the entire classification was made while studying neural networks). The program takes very little space (<5 Mb) and can be written in any modern carrier (or even forwarded in the email).

V. Prospects for further research in this direction

Development of neural networks is a promising direction in forensic medicine, because it will simplify, automate the method of identifying an unknown person and minimize routine errors.

References

[1]. V.A. Fokin, N.N. Ilyinskikh, A.V. Chantsev, Yu.A. Kozlov, The experience of creating a computer dactyloscopic program of medicogenetic destination, *Bulletin of the Siberian medicine, 1, 2002, 117-120.*

[2]. A. Azazy, *The image processing system in the diagnosis of hereditary diseases using method dermatoglyphics* (M.: Nauka, 2011)

[3]. S. Haikeyin, *Neural Networks : A Comprehensive Foundation* (Bei jing: Qing hua da xue chu ban she, 2001).

[4]. S. Haykin, *Neural networks and learning machines.* (Upper Saddle River: Pearson Education, 2009).

[5]. Donald F. Specht, Probabilistic neural networks, *Neural networks, 3, 1990, 109-118.*

[6]. P. Wasserman, *Advanced methods in neural computing* (New York: Van Nostrand Reinhold, 1993).

[7]. Statsoft.ru. STATISTICA Base. [online] Available at: http://statsoft.ru/products/STATISTICA_Base [Accessed 2 Apr. 2017].

[8]. Statsoft.ru. STATISTICA_Neural_Networks. [online] Available at: http://statsoft.ru/products/ STATISTICA_Neural_Networks [Accessed 5 Apr. 2017].

[9]. A.A. Borovkov, *Theory of probability(*M.: Nauka, 1986).

[10]. En.wikipedia.org. Artificial neural network. [online] Available at: https://en.wikipedia.org/wiki/Artificial_neural_network [Accessed 5 Apr. 2017].

[11]. Nielsen, M. Neural Networks and Deep Learning. [online] Neuralnetworksanddeeplearning.com. Available at: http://neuralnetworksanddeeplearning.com/chap2.html [Accessed 5 Apr. 2017].

[12]. JetBrains. Getting Started. [online] Available at: https://www.jetbrains.com/idea/documentation [Accessed 4 Apr. 2017].

[13]. Oracle.com. Oracle Software Downloads | Oracle Technology Network | Oracle. [online] Available at: http://www.oracle.com/technetwork/indexes/downloads/index.html#java [Accessed 5 Apr. 2017].

[14]. Docs.oracle.com. JavaFX: Getting Started with JavaFX Release 8 - Contents. [online] Available at: http://docs.oracle.com/javase/8/javafx/get-started-tutorial/index.html [Accessed 5 Apr. 2017].

[15]. Jfoenix.com. JFoenix. [online] Available at: http://www.jfoenix.com/documentation.html [Accessed 5 Apr. 2017].