Ukrainian finger language recognition using genetic algorithms

Nataliia Stulova

June 7, 2012

Abstract

This paper contains an overviev of existing approaches to pattern recognition problem and of existing systems for ukrainian finger language recognition. An approach to recognising elements of ukrainian finger language with cluster analysis, fuzzy C-means and k Nearest Neighbours is being proposed. The approach is demonstrated on graphical files in JPG format and the results are compared width the brightness gradient method, which is used on the clustering stage.

Keywords: pattern recognition, cluster analysis, classification, fuzzy C-means, FCM, k Nearest Neighbours, k-NN, human-machine interfaces.

1 Introduction

Nowadays about 2% of people suffer from various problems with hearing. For succesful communication and interaction they use finger language, and there are many software solutions which make human-machine interaction via different signs available.

At this time in Ukraine there are several implementations of such sign recognition systems. All of them use techniques of unsupervised sign extraction, but have different implementations, based on using neural networks, graphical filters and 2D or 3D modelling.

A proposed approach is based on using genetic algorithms and fuzzy clustering for extracting both hand and background cluster regions from a picture and further classifying clustered image by k-NN algorithm.

2 Pattern recognition problem

Pattern recognition problem in general can be divided in two stages: feature extraction stage and classification stage. In [1] there is proposed next formal model: let $O = \{o_1, o_2, ..., o_m\}$ be a set of patterns, $\Omega = \{c_1, c_2, ..., c_n\}$ – a set of classes, $V = (v_1, v_2, ..., v_n)$ – a set of feature vectors, R – a set of relationships among vectors and $X = \{x \mid x = \langle V, R\rangle\}$ be a feature representation.

Pattern classification problem then can be formally defined as a composite function

$$P = C(F(O)), \tag{1}$$

where $F : O \to X$ is a feature selection function and $C : X \to \Omega$ is a classification function.

This way, both of the stages can be viewed as optimisation problems. On the first stage of process (1) there emerges a problem to find the "best" feature set for classification, and on the second – to minimize a number of misclassified patterns and increase accuracy of classifier.

Nowadays there are next pattern recognition approaches: partial enumeration, statistical, syntactic and neural-net approaches.

First of them, partial enumeration, can be implemented in two ways – either using branch-and-bound search of using dynamic programming. Both implementations look for exact solution for discrete optimisation problem and adopt a divide-and-conquer strategy.

Second of them, statistical approach, represents each pattern in terms of D attributes(tests) and considers patterns as points in D-dimensional space [2]. The goal is to select those features that allow vectors of patterns belonging to different categories occupy compact and disjoint regions of D-dimensional feature space.

Third of them, syntactic approach, is used when it is acceptable to choose a hierarchical representation of the pattern, where it is regarded as a composition of simpler subpatterns, which in turn consist of even more simple ones, so-called primitives. Each pattern is represented as a set of relationships between primitives. The formal analogy, which gave the name of this approach to the problem of pattern recognition is the syntax of language.

Fourth of them, neural networks, can be regarded as a massive parallel computing system consisting of a large number of simple processors (neurons) with multiple relationships. The network is a weighted directed graph in which vertices are artificial neurons and directed edges – relationships between inputs and outputs of neurons. Most problems in recognition use feedforwards nets and Kohonen maps samoorhanizatsiyni (mainly in the problems of clustering).

3 Genetic algorithms

Unlike the other methods mentioned above, genetic algorithms are rather a set of stochastic algorithms, that represent a broad and uncomplicated tool to study complex optimization problems.

Genetic algorithm is evolutionary search algorithm used to solve optimization problems by successive selection, combination and variation of the desired parameters using biological evolution-like mechanisms. The peculiarity of the genetic algorithm is focused on the use of "crossover" operator, which performs the recombination operation, making the candidate. It's role is similar to the role of crossing in nature.

The most general implementation of GA (which was used in proposed approach) has next stages:

- create initial population of chromosomes
- repeat next steps until current optimal criteria is satisfied:
  1. select parent chromosomes;
  2. perform crossover to make child population;
  3. perform mutation to child population;

The most fundamental description of genetic algorithms is provided in [3]. The possibilities of genetic algorithms as studied in [4–6]. The efficiency of GA was shown in schema theorem and proven in 1975 by John Holland in [3].

4 Cluster analysis and classification in pattern recognition

Cluster analysis is associated with the first optimisation problem, caused by recognition problem (1). The paper [7] proposed to use genetic algorithms to select the optimal cluster centers from disordered and unlabeled original data set. This paper describes a genetic algorithm optimized to improve the classification of data by NPC (classification by the nearest prototype).

Also, detailed analysis of clustering techniques using GA held in the works [8–15]. In this paper, clustering is considered as a method of obtaining the key features of the image. The output from cluster analysis, viewed as coloured areas from converted original image, can be considered as input to the classifier.

Classification is associated with the second problem,caused by recognition problem (1). Among the many approaches to the implementation of this process, there are three main [16]:

- based on the symbolic approach;
- based on neural networks;
- based on social and emergency principles.

Also hybrid implementations are possible. The trend of recent years is the addition of the first two categories of algorithms in the third.

The methods that combine the first and third approaches, often include a combination of decision trees and genetic algorithms. In papers [17, 18] there are described algorithms, that use genetic algorithm for finding concepts for inductive algorithm, which constructs decision trees. A modified implementation of this approachis presented [19].

The second class of methods generally include those, that use neural networks. By [20], there are three basic approaches to combining neural networks and genetic algorithms: configuration weights in neural networks using GA [21, 22], construction of network topology [23], selection of training data to the network.

The methods belonging to the third approach, rank those that use multiple adaptation of classical genetic algorithms. The paper [24] presented modification of the classifier k-NN (k nearest neighbors) using genetic algorithms. The author proposes to use GA for accuracy checking of distribution of test cases from a given training set of classes. A similar approach is described in [25]. General research methods in this category are held in [26]. In papers [27–30] describes learning classifiers using genetic algorithms with variable-length chromosomes to determine the hyperplane or piecewise linear boundaries that outline the boundaries of classes.

5 Existing systems

There are several systems for ukrainian finger language recognition [31, 32]. There are three main stages of sign language recognition in these systems:

1. gesture data receiving;
2. obtaining additional information about the gesture;
3. processing information about the gesture.

On the first and second stages there is often used specific hardware. On the third stage of processing the received information there are next possible methods: palm shape recognition, gesture trajectory recognition, processing the whole image without extracting some of its parts.

In this paper static image recognition methods that perform palm shape recognition are considered.

The paper [31] proposes two algorithms that work with data obtained from a frontal camera. The main idea of the proposed methods is to obtain palm contours and further their comparison with the standard. The first way is to get the contours using values of the gradient of the brightness. Another way to get contours is to use the Sobel operator.

The paper [33] presents a series of algorithms which identify the elements of gesture pixel by pixel. In the first method there are used one webcam and gloves, the color of which is very different from colors of objects that fall within the field of view of the webcam. Based on the input color image a square is built, and it's size corresponds to a hand with outstretched fingers. Then a square divided into 16 equal cells and a primitive model of the gesture is built. This way a set of primitives is obtained and further recognition process is divided into two stages, at first of which a gesture is associated with one of the primitives and at second – it is compared with the standard gestures, which correspond to the same primitive.

The second method has different format of the input data used. It uses a glove with different color of each finger. It allows precisely determine the position of each single finger and approximately determine the condition of finger – folded or unfolded.

The paper [34] suggests using analysis of connected regions in the input data to recognize elements of finger language. Authors propose to use segmentation methods on the incoming image to define gesture area Reg. To select the characteristic features of the palm it is proposed to choose the following parameters: area of the palm, perimeter of it, compactness, main axis orientation and elongation.

In the paper [32] it is proposed to use the neural network to distinguish regions of the ends of fingers, which are used as characteristic features of the problem of recognizing gesture. As realization of the network a multilayer perceptron was chosen. To distinguish ends of fingers from the background and determine the number of pixels on the image, that belong to fingers, a black and white standard method was chosen.

Experimentally it was determined that none of components of colour format RGB, which is used for input image encoding, does not allow to distinguish the hand region from background details. Therefore, the colors were presented in the format of YCbCr.

6 Proposed approach

A proposed method is based on clustering algorithm for satellite images recognition, called VGA-FCM, as described in [35], and the classification algorithm is k-NN (k-Nearest Neighbours).

The first stage of pattern recognition problem, feature selection stage, is implemented as a modified VGA-FCM algorithm. This modification performs a two-stage clustering of the original image into two regions. In [35] VGA-FCM is used to select regions corresponding to roads from the satellite image area. Modified VGA-FCM helps to distinguish regions that correspond to palm and to background.

Modified VGA-FCM has the following stages:

1. Creeate the initial population and calculate its value for fitness function J_{old}
2. Repeat the steps 1-7 until $|J_{old} - J_{new}| > 0,01$:
   (a) Initialize the best fitness value for parent (old) population J_{old} with the best fitness value of descendants (new) population J_{new};
   (b) Apply selection operator;
   (c) Apply crossover operator;
   (d) Apply mutation operator;
   (e) Replace th old population with the new population;
   (f) Clear new population;
   (g) Initialize best fitness value of new population J_{new};
3. Return the vector of obtained clusters.

The second stage of pattern recognition problem, classification, is implemented with the k-NN algorithm. After receiving clustered image as input algorithm compares it with every standard from standards set pixel by pixel in a following way: it color of the pixel with (i, j) coordinates on the sample is compared with the color that is obtained by majority voting among the 5 standard pixels that make the "cross" (Fig. 1) with the center at (i, j) pixel on the standard image.

Pseudocode of the majority voting algorithm has the following stages:

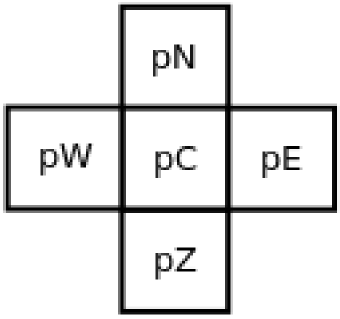

Figure 1: “cross” of the five pixels standard

1. Get the position of the center “cross” pC at the coordinate (i, j);
2. Initialize the color values “cross” ends:
   (a) pN with color value of the pixel at coordinate $(i, j-1)$;
   (b) pZ with color value of the pixel at coordinate $(i, j+1)$;
   (c) pE with color value of the pixel at coordinate $(i+1, j)$;
   (d) pW with color value of the pixel at coordinate $(i-1, j)$;
3. Calculate number of pixels in the “cross” with color that matches the background color, N_{bg};
4. Calculate number of pixels in the “cross” with color that matches the hand color, N_{hd};
5. Choose a color from standard as hand color, if $N_{hd} > Nbg$, otherwise as background color;
6. Return 1 if the pixel color on sample matches the color on standard, 0 otherwise.

Pseudocode of the classification algorithm has the following stages:

1. Determine the color of background and hand on the clustered image;
2. Reconcile sizes of sample and standards;
3. Paint sample clustered areas according to colors of standards;
4. For each i-th standard build number S_i, corresponding to the number of pixels of the sample color that correspond to hand on sample and on standard (in this case the boundary pixels of images are not considered);
5. Return the index of the sample, for which S_i is maximal.

Overall block-scheme is shown in Figure 2.

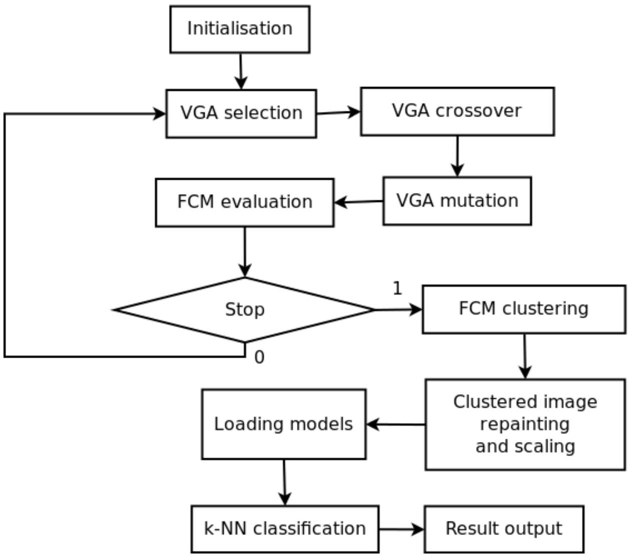

Figure 2: The general block-scheme

7 Input data

On the clustering stage a color image of the right palm which shows a gesture on a homogeneous background is chosen as input data. Input image is encoded in JPG format, and it’s size is 120‘ $\times$ ‘120 pixels. To be able to process images by color, a RGB color model was chosen. In this model white is $(255,\ 255,\ 255)$ – i.e. 100% of each of the three basic colors, and black – $(0,\ 0,\ 0)$, their complete absence.

According to this model, the original image can be represented as N-element array of points in three-dimensional color space RGB. In the array $X = \{x_1,\ x_2,\ \dots\ ,\ x_n\}$, where N is the total number of pixels of a color image, and every $x_j = \{x_{j1},\ x_{j2},\ x_{j3}\} \in X$ is a three-dimensional point.

Examples of original image and it‘s clustered result are shown in figures 3 and 4:

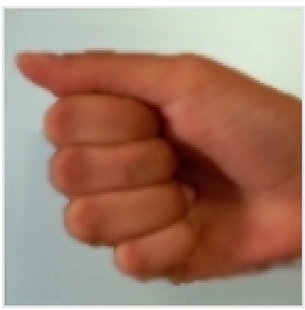

Figure 3: Image of letter A

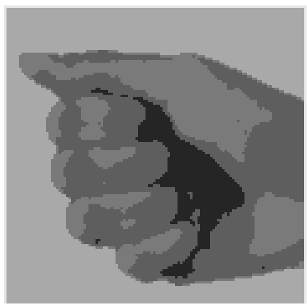

Figure 4: Clustered image of letter A

On the classification stage input consists of one bicolor image with total 1024 (32×32) pixels and an array of same sized bicolor images-models. Examples of one image-model and one unrecognized pattern are shown in figures 5 and 6:

Figure 5: Image-model of letter T

Figure 6: Clustered image of letter T

Before unrecognised pattern is compared with the first model it is repainted in the same colours, as model is.

8 GA implementation

Chromosome representation, fitness function and crossover operator are used the same as in [35]. Only selection and mutation operators were modified.

Selection is still performed with the use of roulette, but after each two "turns" into new population 4 chromosomes are added: two parents, that were chosen randomly with roulette, and their two childs. Number of chromosomes in population(N) is a fixed number, so selection operator is called $\frac{N}{4}$ times.

Mutation operator changes one component of random gene of randomly chosen chromosomes with the integer in range $[0; 255]$.

9 Results

Results of the proposed approach (GA) were compared with ones, that were received in [31] with the brightness gradient method (BGM). Two groups of people were taken into experiment: men aged from 21 to 23 years (M2123) and women aged from 16 to 19 (W1619). Results are shown in table 1:

Table 1: Comparison of results

Person	BGM		GA	
	Correct of 50	%	Correct of 50	%
M2123	30	60	42	84
W1619	33	66	44	88

As seen from the table, proposed approach shows higher results than the gradient of brightness method. Distributions of clusters and number of correct gestures recognized using genetic algorithms for both cases are presented in figures 7 and 8:

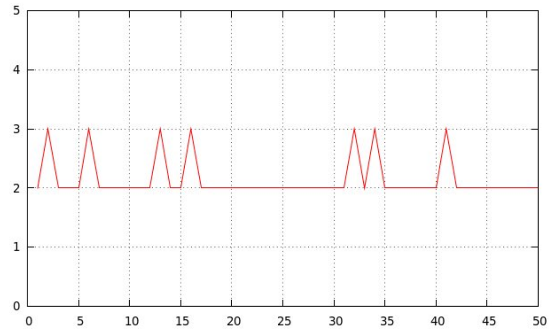

Figure 7: The distribution of clusters in recognizing the hand of men 21-23 years

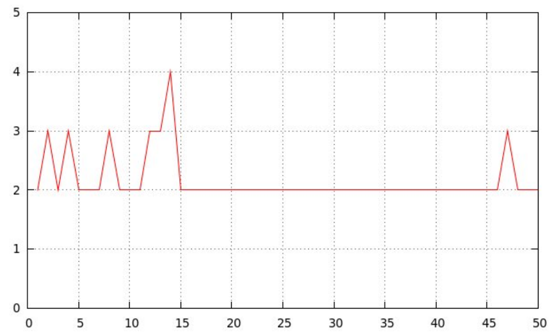

Figure 8: The distribution of clusters in recognizing the hand of women 17-19 years

10 Analysis of the modified GA

The results show that the highest percentage of correctly recognized gestures corespond to $P_M = 5\%$ and $P_M = 8\%$ values. Values of other parameters are following: $C_{min} = 2$, $C_{max} = 5$, $m = 3$, $N = 100$. This dependence is shown in figure 9 (x-axis shows P_M variation, y-axis shows recognition score):

The effect of changing the number of individuals in the population on the total number of generations was investigated. Figure 10 presents the results of 10 successfully recognized gesturesfor the following values for the number of individuals: $N = \{10, 20, 30, 40, 50, 80, 150\}$. The values of other parameters are following: $C_{min} = 2$, $C_{max} = 5$, $m = 3$, $P_m = 8\%$.

Almost all numbers of generations do not exceed limit in 12, so the conclusion is that for $N \in [20; 150]$

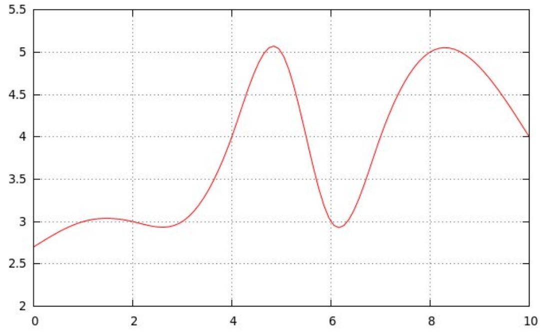

Figure 9: Recognition score with different P_M

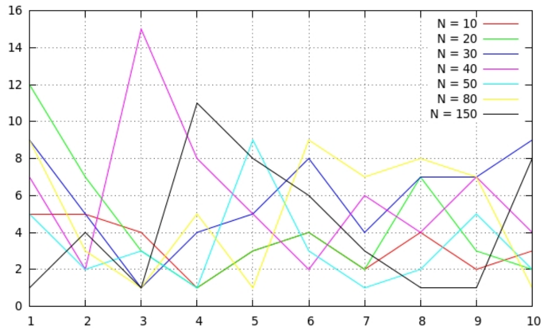

Figure 10: Number of generations with different population size

there is no significal win in increasing the number of individuals.

The effect of changing the maximal number of clusters on the total number of generations was investigated. Figure 11 presents the results of 10 successfully recognized gesturesfor the following values for the number of clusters: $C_{max} = \{2, 3, 4, 5, 8, 12, 100\}$. Values of other parameters are following: C_{min}‘ = ‘2, N‘ = ‘20, m‘ = ‘3, P_m‘ = ‘8%.

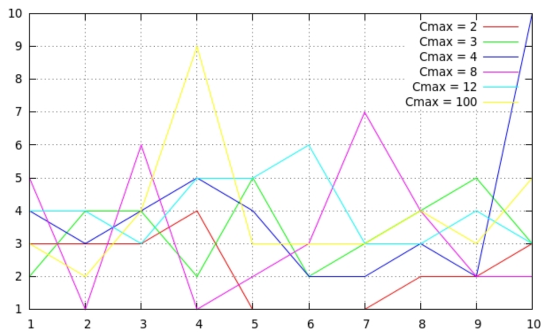

Figure 11: Number of generations with different C_{max} value

Figure 11 shows that with decreasing the maximal number of clusters, number of generations, and, consequently, the time needed for algorithm to finish, reduces.

11 Further investigations

- faster clustering algorithm;
- reading input data from the video stream;
- extension the database of gesture images;
- combining genetic algorithms with neural networks in order to increase recognition speed;
- using special input devices.

References

[1] Yuan-Kai Wang and Kuo-Chin Fan. Applying genetic algorithms on pattern recognition: an analysis and survey. *Pattern Recognition*, 2:740–744, 1996.

[2] Anil K. Jain, Robert P.W. Duin, and Jianchang Mao. Statistical pattern recognition: A review. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 22(1):4 – 37, Jan 2000.

[3] John H. Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*. MIT Press, Cambridge, MA, USA, 1992.

[4] Kenneth De Jong. Learning with genetic algorithms: An overview. *Machine Learning*, 3:121–138, 1988.

[5] David E. Goldberg and John H. Holland. Genetic algorithms and machine learning. *Machine Learning*, 3:95–99, 1988.

[6] O. Cordon, F. Gomide, F. Herrera, F. Hoffmann, and L. Magdalena. Ten years of genetic fuzzy systems: current framework and new trends. *Fuzzy Sets and Systems*, 141(1):5 – 31, 2004. <ce:title>Genetic Fuzzy Systems: New Developments</ce:title>.

[7] Ludmila I. Kuncheva and James C. Bezdek. Nearest prototype classification: Clustering, genetic algorithms, or random search? *IEEE Transactions on Systems, Man, and Cybernetics*, 28(1):160–164, February 1998.

[8] Rowena Marie Cole. Clustering with genetic algorithms. Master's thesis, University of Western Australia, 1998.

[9] Ayhan Demiriz, Kristin Bennett, and Mark J. Embrechts. Semi-supervised clustering using genetic algorithms. In *In Artificial Neural Networks in Engineering (ANNIE-99)*, pages 809–814. ASME Press, 1999.

[10] Sanghamitra Bandyopadhyay. Genetic algorithms for clustering and fuzzy clustering. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(6):524–531, 2011.

[11] Saeed Parsa and Omid Bushehrian. Genetic clustering with constraints. *Journal of Research and Practice in Information Technology*, 39(1):47 – 60, February 2007.

[12] Dharmendra K Roy and Lokesh K Sharma. Genetic k-means clustering algorithm for mixed numeric and categorical data sets. *International Journal of Artificial Intelligence & Applications*, 1(2):23 – 28, April 2010.

[13] Hwei-Jen Lin, Fu-Wen Yang, and Yang-Ta Kao. An efficient ga-based clustering technique. *Tamkang Journal of Science and Engineering*, 8(2):113 – 122, 2005.

[14] Виталий Снитюк. ЭВОЛЮЦИОННАЯ КЛАСТЕРИЗАЦИЯ СЛОЖНЫХ ОБЪЕКТОВ И ПРОЦЕССОВ. In Gladun V.P., Kr.K. Markov, A.F. Voloshin, and Kr.M. Ivanova, editors, *Proceedings of the XI-th International Conference "Knowledge-Dialogue-Solution"*, volume 1, pages 232– –237. FOI-COMMERCE, June 2005.

[15] Ujjwal Maulik and Sanghamitra Bandyopadhyay. Genetic algorithm-based clustering technique. *Pattern Recognition*, 33(9):1455 – 1465, 2000.

[16] Джорож Ф. Люгер. *Искуственный интеллект: стратегии и методы решения сложных проблем*. М.: Издательский дом "Вильямс", 2003.

[17] Peter D. Turney. Cost-sensitive classification: Empirical evaluation of a hybrid genetic decision tree induction algorithm. *Journal of Artificial Intelligence Research*, 2:368–409, 1995.

[18] J. Bala, J. Huang, and H. Vafaie. Hybrid learning using genetic algorithms and decision trees for pattern classification. In *In 14th Int. Joint Conf. on Artificial Intelligence (IJCAI)*, pages 719 –724. Morgan Kaufmann, 1995.

[19] Kenneth A. De Jong and William M. Spears. Learning concept classification rules using genetic algorithms. In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence*, pages 651–656. Morgan Kaufmann, 1991.

[20] D. Whitley. Genetic algorithms and neural networks. In *Genetic Algorithms in Engineering and Computer Science*, pages 191–201. John Wiley, 1995.

[21] Bernd Dachwald Lampros Tsinas. A combined neural and genetic learning algorithm. In *Proceedings of the First IEEE Conference on Evolutionary Computation (ICEC '94)*, June 1994.

[22] Ricardo B. C. Prudencio and Teresa B. Ludermir. Neural network hybrid learning: Genetic algorithms & levenberg-marquardt. In Martin Schader, Wolfgang Gaul, and Maurizio Vichi, editors, *Between Data Science and Applied Data Analysis*, Studies in Classification, Data Analysis, and Knowledge Organization, pages 464–472. Springer Berlin Heidelberg, 2003.

[23] Daniel Polani and Thomas Uthmann. Training kohonen feature maps in different topologies: An analysis using genetic algorithms. In Stephanie Forrest, editor, *Proceedings of the Fifth International Conference on Genetic Algorithms*, pages 326–333. San Francisco, CA: Morgan Kaufmann, 1993.

[24] Ludmila I. Kuncheva. Editing for the k-nearest neighbors rule by a genetic algorithm. *Pattern Recognition Letters*, 16(8):809–814, 1995.

[25] M. Pei, E. D. Goodman, and W. F. Punch. Feature extraction using genetic algorithms. In *Proceeding of International Symposium on Intelligent Data Engineering and Learning'98 (IDEAL'98), Hong Kong*, page 98, 1997.

[26] Faten Hussein, Nawwaf Kharma, and Rabab Ward. Genetic algorithms for feature selection and weighting, a review and study. In *Proceedings of the Sixth International Conference on Document Analysis and Recognition*, ICDAR '01, pages 1240–1244, Washington, DC, USA, 2001. IEEE Computer Society.

[27] S. Bandyopadhyay, C.A. Murthy, and S.K. Pal. Pattern classification using genetic algorithms: Determination of h. pages 1–22, May 1998.

[28] Sanghamitra Bandyopadhyay and Sankar K. Pal. Pattern classification with genetic algorithms: Incorporation of chromosome differentiation. *Pattern Recognition Letters*, 18(2):119–131, 1997.

[29] S. Bandyopadhyay, C.A. Murthy, and S.K. Pal. Pattern classification with genetic algorithms. *Pattern Recognition Letters*, 16:801–808, 1995.

[30] Sanghamitra Bandyopadhyay and Sankar K. Pal. Pixel classification using variable string genetic algorithms with chromosome differentiation. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, 39:303–308, February 2001.

[31] Ю.В. Крак and Д.В. Шкільнюк. Аналіз елементів дактильної жестової мови. *Искусственный интеллект*, 3:322 – 328, 2010.

[32] М.В. Давидов, Ю.В. Нікольський, and О.В. Пасічник. Дослідження ефективності методів розпізнавання у моделях жестової мови. *Інформаційні системи та мережі*, (621):117 – 123, 2008.

[33] Ю.В. Крак and Д.В. Шкільнюк. Технологія розпізнавання елементів дактильно-жестової мови. *Искусственный интеллект*, 3:564 – 572, 2009.

[34] Юрій Васильович Крак and Дмитро Валерійович Шкільнюк. ЗНАХОДЖЕННЯ ХАРАКТЕРИСТИЧНИХ ОЗНАК НА ЗОБРАЖЕННЯХ РУКИ ДЛЯ ЗАДАЧ РОЗПІЗНАВАННЯ ЕЛЕМЕНТІВ ДАКТИЛЬНОЇ МОВИ.

[35] H. Liu, J. Li, and M. A. Chapman. Automated road extraction from satellite imagery using hybrid genetic algorithms and cluster analysis. *Journal of Environmental Informatics*, 1 (2):40–47, 2003.