image recognition, data classification, neural network, decision support systems

Piotr PORWIK*

THE DEDICATED DECISION SUPPORT SYSTEM IN RECOGNITION OF SOME UNCERTAIN DISEASE ENTITIES

This work presents the principles of image recognition, where quality-based methods are applied. The neural networks and additional software have been proposed. This goal was achieved by using non-parametric recognition algorithms. In this paper the two-state hybrid classification method has been proposed, where artificial intelligence algorithm is included. In recognition process, the learning method, selection and optimization of diagnostic parameters have been introduced. The integrated part of the classifier structure is voting mechanism, which indicates incorrect states of the system – for example the unrecognized images. Effectiveness of the system has been shown by means of examples, where ambiguous data have been incorporated – it is very often a practice of medical diagnostics.

1. INTRODUCTION

It is well known that decision support systems are widely used in medicine. It results from the fact that medical diagnostics, even in a simple case, include "fuzzy" features: subjectivity of a diagnostician and medical diagnostics of a patient. It is obvious that the large role in taking correct decision plays a doctor experience. Frequently, for a large number of practiced doctors, the diagnostic rules are often very complicated and occasionally even risky. In medical practice, assistance of decision support systems should minimize the risk of incorrect decision.

One of prevalent decision tasks is repeated recognition of the class of objects. It is a typical task in computer support medical diagnostic, where patient status is recognized on the basis of the patient findings. In widely used diagnostics, recognition task boils down to estimate a given object to one of defined class. Presented description in medical diagnostics reduces the interpreting formula: (object -> patient), (features -> clinical symptoms), (class -> patient status or disease entity).

Recognition process is introduced in a case, when object classification rules are unknown, although classes (disease entity) are defined. Generally, recognition methods base on preliminary measurement of features, which describe recognised objects and in the next step on the analysis and selection such features.

The measured values of the selected symptoms (features) make a pattern of the vector features:

$$\underline{x} = [x_1, x_2, ..., x_n]^T \qquad (1)$$

the vector (1) is determined for the known *a priori M* classes of objects (disease entities):

$$\omega_1, \omega_2, ..., \omega_M \qquad (2)$$

On the basis of a set of classes (2), classification can be conducted:

$$\Psi(\underline{x}) = \omega_k, \quad k = 1, 2, ..., M \qquad (3)$$

In the proposed method learning algorithms are used – it means that embedded intelligent software incorporates teaching algorithm with learning sequences (coefficients or weight matrices), and the set of medical knowledge gathered during clinical practice. The last knowledge dataset is regarded as experiential material.

The teaching sequence can be formally described as:

$$TS = \left\{ (\underline{x}_1, \omega_1), (\underline{x}_2, \omega_2), ..., (\underline{x}_m, \omega_m) \right\} \qquad (4)$$

where $\underline{x}_1, \underline{x}_2, ..., \underline{x}_n$ are the vectors of successive features and $\omega_1, \omega_2, ..., \omega_M$ describe the results of correct disease recognition.

Recognition algorithm allows to classify new unknown features vector $\underline{x}$ to appropriate class ω - at the moment there is realized classification process (3).

* Institute of Informatics, University of Silesia, Będzińska 39, 41-200 Sosnowiec, Poland, e-mail: piotr.porwik@us.edu.pl

2. SELECTION CRITERIA OF RECOGNITION ALGORITHM

Today, the learning systems are applied in major recognition procedures. It follows from the fact that in modern medical diagnostics, probability of classes and density of probability are mostly unknown.

The statistical and parametrical recognition algorithms count the Bayes method among methods without learning [1].

The cases, where full probabilistic information is accessible can be considered as a theoretical accident only. In practice only expert rules are formed as mapping of an empirical practice.

Due to approximate estimation of *a posteriori* distribution probability, complicated calculations and requirement of learning sequences in practice the minimal-distance, non-parametric algorithms are used very often [1,2,6]. In such an approach, simple decision rules can be introduced and so-called Nearst-Mean (NM) algorithms are preferred.

The basic criterion of parameters selection in observation space is increasing of the diagnosis error probability. In many cases the set of data (features) has inconsistent character. For this reason, decision area partitioning can be difficult. It should be also emphasized, that in this time there is no other data sets, which can improve classification results.

It is obvious that signal parameters should concentrate images on one class and all areas should be separated for images which belong to different classes.

Unfortunately not always it is secure – it can be often observed in human physiological or anthropological data acquisition.

Fig. 1 presents the classified dataset, for M=4 different classes: healthy patient (image ω_1) and manifested disease entities (images ω_2, ω_3,...,ω_4). The decision areas were determined by means of the largest credibility method, where decision Bayesian function was introduced [1]. In Fig. 1 can be clearly observed ambiguous classification areas.

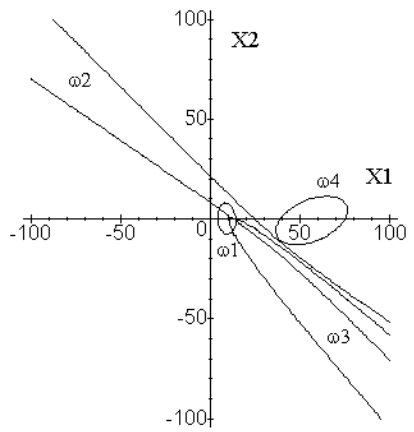

Fig. 1. Classification areas obtained by means of the Bayesian decision function

Fig. 2 depicts the same dataset divided into two classes (M=2) according to principle: healthy patient (image ω_1) or unhealthy patient (image ω_2), where image ω_2 is a sum of symptoms of the same disease entities form Fig. 1.

Decision areas have been determined with the aid of modification of the classical NM algorithm. The modification includes additional algorithm with linear separating functions.

Implementation of such a method in detail has been presented in work [5]. Similarly like on Fig. 1 some features of the vectors are also wrongly classified.

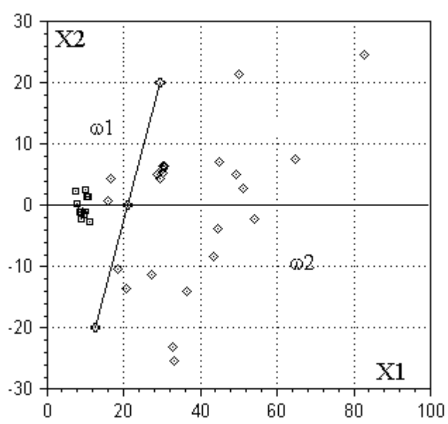

Fig. 2. Two-state classification by means of linear separating function

Classification difficulties can be overcame by appropriate transformation of the primary parameter space. The new space should have better discrimination properties. There are such sets of data (Figs. 1-2) that discover the transformation formulas can be difficult.

3. THE HYBRID RECOGNITION SYSTEM

From the text presented above follows, that in the exceptional cases can appear the risk of the inaccurate classification. For these cases the special type of classifiers can be build – then the risk of wrong classification can be minimized, because improper classifier states will be additionally indicated.

The scheme of such a hybrid classifier was depicted in Fig. 3. The concluding block was realized by means of artificial neural network. The *NM* classifier by means of minimal-distance algorithm was implemented.

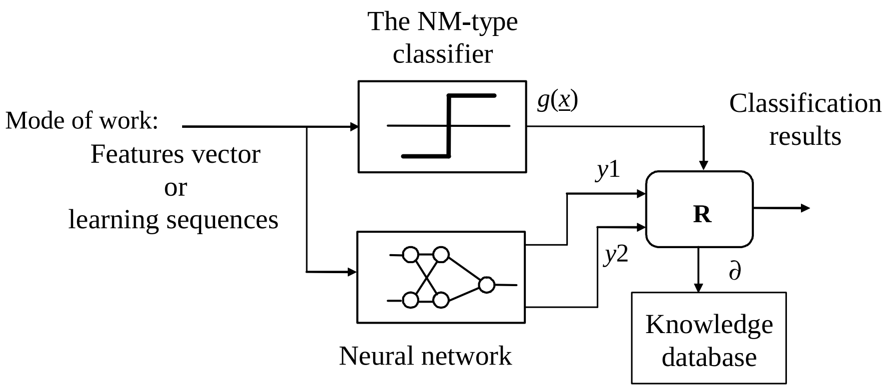

Fig. 3. Scheme of the hybrid recognition system. (Description of elements is presented in the text below)

Presented on Fig. 3 connections of the modules can work as a two-phase circuit:

- The first phase (off-line mode) create a learning cycle. In this stage, by means of the learning sequence *LS*, the neural knowledge base is formed. In the next stage the characteristic of the classifier *NM* is determined.
- In the second phase (on-line mode) data stream is recognized and appropriate classified.

The *NM*-type classifier works on the basis of linear decision function (it is so-called discriminant function). Discrimination separates the classes along the straight-line:

$$g(\underline{x}) = ax_1 + bx_2 - \Delta \qquad (5)$$

where: a, b stand weight coefficients, and the value Δ stands threshold of classification. The classification is performed according to the rules:

$$\begin{aligned} &\text{if } g(\underline{x}) > 0 \quad \text{then} \quad \underline{x} \in \omega_1 \\ &\text{if } g(\underline{x}) \le 0 \quad \text{then} \quad \underline{x} \in \omega_2 \end{aligned} \qquad (6)$$

Equation (6) can be easily transformed to the polar function [1,8].

Used in the classifier the perceptron neural network has a multi-layer structure. The perceptron is the simplest and well known kind of feed forward neural network with a linear classifier. The network input $\underline{x}$ and output $\underline{y}$ signals can be described as:

$$\underline{y} = \varphi\{W_n\varphi ... \{W_2\varphi\{W_1\underline{x}\}\}\} \qquad (7)$$

where: $W_1, W_2, ..., W_n$ – weights matrices of the connections between network layers, n – number of network layers, $\varphi\{\cdot\}$ – neuron activating function.

In the proposed approach, neural network realises recognition according to dependence:

$$\begin{aligned} &\text{if } y_1 \approx 1 \quad \text{then} \quad \underline{x} \in \omega_1 \\ &\text{if } y_2 \approx 1 \quad \text{then} \quad \underline{x} \in \omega_2 \end{aligned} \qquad (8)$$

During learning session, the appropriate weights are initialized and some parameters are established: speed of learning (η) and momentum (α). It is well known, that the appropriate selection of the η coefficient is a hard task. If the value of η is too small, then neural network learn very slow and number of iterations will be large. If mentioned value is too large, then learning process has divergent character and is not stable.

The learning acceleration of the algorithm convergence can be increased if momentum parameter α will be introduced [8]. If this parameter has large values, then learning process is more secure and stable. However, if the parameter α will be inappropriately selected, then the weights coefficients matching are very hard.

It should be emphasized that primary selection of the mentioned above parameters depends mainly on programmer's experience and can by distinctly different as theoretical considers [1,8].

In proposed solution the best results of neural network learning were achieved for neural layer type 2-3-3-2. The parameters of the network were established as η=4, α=0.1, weights distribution=0.5. The arbiter *R* works according to the rules:

$$
\begin{array}{llll}
y_1 > y_2 & \text{and} & g(\underline{x}) > 0 & \text{then } \underline{x} \in \omega_1 \\
y_1 > y_2 & \text{and} & g(\underline{x}) \le 0 & \text{then error } \partial \\
y_1 \le y_2 & \text{and} & g(\underline{x}) > 0 & \text{then error } \partial \\
y_1 \le y_2 & \text{and} & g(\underline{x}) \le 0 & \text{then } \underline{x} \in \omega_2
\end{array}
$$

In addition, different, more restricted rules between the values y_1 i y_2 can be also introduced. It allows to adopt the classifier to another input dataset or change the recognition parameters for current data.

4. CONCLUDING REMARKS

Presented in the paper recognition and classification method can give practical effects – the method can be used in supplementing of the knowledge bases for the cases when disease entities were classified wrongly. Because rules and criteria of the arbiter module *R* can be freely changed, the proposed method can has widely application. Proposed approach can be fully realized as specialized programme.

The error back-propagation method has been greatly used for the supervised training of feed-forward neural networks. However, as it is well-known, this method has a slow convergence. However in our approach the parameter η was accurately selected.

The formulated algorithms were tested for selected group of real data. For such dataset, correctness of recognition achieves 88%. During conducted tests, the knowledge database capacity was increased about 5% - it means that the new decision rules were included into knowledge database.

REFERENCES

[1] BASZTURA C., Komputerowe systemy diagnostyki akustycznej. PWN, Warszawa 1996.

[2] BASZTURA C., Źródła, sygnały i obrazy akustyczne. WKŁ, Warszawa 1988.

[3] KORBICZ J., i inni, Sztuczne sieci neuronowe. Podstawy i zastosowania. Akademicka Oficyna Wydawnicza, Warszawa 1994.

[4] NIENIEWSKI, M., Probabilistic methods in pattern recognition and computer vision. Lecture Notes and Computer Vision and Artificial Intelligence, Ossolineum 1990, pp. 123-161.

[5] PORWIK P.,ADAMCZEWSKI J., The object states classification method in diagnostics experiment. XXIV Ogólnopolskie Sympozjum DIAGNOSTYKA MASZYN. Politechnika Śląska, Zeszyt 1/ 1997, pp. 163-168.

[6] SOBCZAK W., MALINA W., Metody selekcji i redukcji informacji. WNT, Warszawa 1985.

[7] RUTKOWSKA D. i inni, Sieci neuronowe, algorytmy genetyczne i systemy rozmyte. PWN, Warszawa 1997.

[8] TADEUSIEWICZ R., Sieci neuronowe. Akademicka Oficyna Wydawnicza, Warszawa 1993.