UDC 519.65+616.4
DOI: 10.20535/1810-0546.2017.1.93128

G.V. Kriukova[1*], S.P. Radchenko[2], O.O. Sudakov[2]
[1]National University of "Kyiv-Mohyla Academy", Kyiv, Ukraine
[2]Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

ARTIFICIAL NEURAL NETWORK FOR MULTICLASS RECOGNITION AND ITS APPLICATION TO THE THYROID FUNCTIONAL STATE[**]

Background. Development of automated diagnostic requires selection and improvement of appropriate machine learning methods, in particular multiclass recognition. Artificial Neural Networks (ANN) of various architecture are considered as an approach to the problem.
Objective. The goal is to analyze and compare performance of ANN-based classifiers on various datasets for further improvement of model selection strategy.
Methods. ANN-based models of the distribution of class labels in terms of predictor features are constructed, trained and validated for datasets of clinical records. Varying training algorithms for multi-layer perceptrons, Kohonen neural network, linear functional strategy with multi-parameters regularization are considered.
Results. Performance of the classifiers is compared in terms of accuracy, sensitivity, and specificity. Linear functional strategy classifier outperforms the other with more complex ANN-architecture and exhibits relative steadiness to overfitting. Performance of Kohonen neural network on large dataset exceeds 90 % in terms of specificity for each class, withal sensitivity for distinct classes is more than 95 %.
Conclusions. The understanding of the strengths and limitations of each method is crucial for careful choice of ANN-based classifier, particularly its architecture, regularization and training algorithm.
Keywords: artificial neural network; classification; ultrasound image processing; regularization; inverse problem.

Introduction

Ultrasound imaging (ultrasonic imaging, ultrasound introscopy, USI) is a common medical diagnostic tool. USI is widely used for medical screening and automated USI images procession techniques are of great interest. There were many automated diagnostic techniques based on procession of USI images suggested [1–4]. These techniques involve image segmentation, features extraction and data classification. Many methods of classification are based on artificial neural networks (ANN), particularly multi-layer perceptrons (MLP). For example, various convolutional neural network (CNN) architectures are widely used for image classification. In present work we are focused on thyroid gland diseases recognition based on screening data that involve USI images features, hormonal analysis results, other patients' data.

It was shown in [5] that Kohonen neural network (NN) based classifier is efficient for ultrasound image processing and classification with 9 classes corresponding to diagnoses (normal, hypothyroidism, cancer, nodes, thyroxicosis, goiter, diffuse goiter, and samples with unknown and multiple diagnoses simultaneously). Despite so called deep CNNs have demonstrated impressive results on a number of computer vision problems, their computational costs and the need for huge amount of data for training make CNNs not always applicable for considered problem.

In this paper we demonstrate that application of Linear Functional Strategy (LFS) for diagnosis classification based on USI images can provide the same or even better efficiency compared to neural networks. In addition LFS requires significantly less data and resources.

Problem statement

Let $x = (x^1, x^2, x^3, ..., x^n) \in \mathbf{R}^n$ be a vector of medical measurements (or features extracted from image). We can access clinical records of patients, which contain the historical data, such as measurements or extracted features $x_j = (x_j^1, x_j^2, x_j^3, ..., x_j^n) \in \mathbf{R}^n$, $j = 1, 2, ..., m$, and retrospectively estimated class label y_j for the corresponding diagnosis. The goal is to construct a classifier assigning class labels y to the incoming instances with known predictor features x.

ANN-based classifier model

Trying to achieve better results in performance measurements we suggest to train the system to

*corresponding author: galyna.kriukova@gmail.com
** The authors gratefully acknowledge the support of the project AMMODIT funded within EU H2020-MSCA-RICE

recognize the risk of each diagnosis separately y_j^i, where i is class number, i.e. we teach the network to separate just one cluster versus all of other data, and then to recognize each diagnosis separately by selecting the optimum weight coefficients for risk combination of y_j^i. It corresponds to additional layer in neural network, aggregating the outputs for each class.

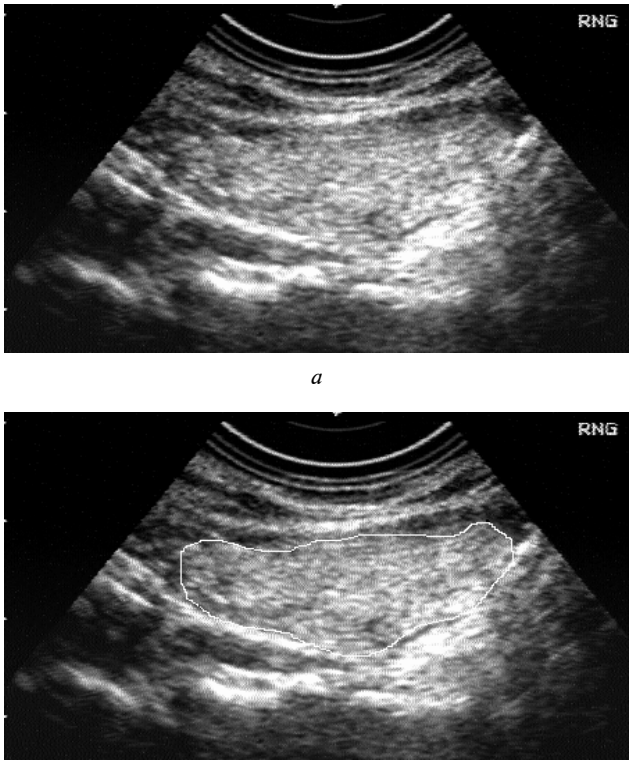
In this study we consider two main approaches: Kohonen neural networks for dataset with large number of features (see [5]), and linear functional strategy (LFS) for dataset with relatively small number of features (see [6]), where use of large neural networks is inappropriate because of propensity to overfitting. LFS can be interpreted as a neural network with one hidden layer, where the basic constructed rankers play the role of activation function. LFS shows the promising results for medical application, it is reliable and interpreted algorithm for predicting nocturnal hypoglycemia in common in patients with insulin-treated diabetes [7].

Below we recall the standard metrics for measuring the performance of classifiers: true positive rate (TPR), false negative rate (FNR), and the classification error percentage (CEP), i.e. the percenttage of incorrectly classified patterns, finally, accuracy = 100 % − CEP.

Kohonen NN for ultrasound image processing

For application of Kohonen neural network to diagnosis classification the data set for more than 600 patients was applied [5]. The data set includes introscopic images of thyroid gland, general formal information (name, surname, age, sex, pilot diagnoses etc.) and results of biochemical assay to investigate hormones concentration (protein-bound triiodothyronine T3 and thyroxine T4, free FT3 and FT4, additional thyroid stimulating hormone TSH and the thyroglobulin TG), description and complaints for each patient. USI system had a sector transducer ASU-32 WL-7.5. All ultrasonic introscopic images were obtained by ultrasound B-mode scan at working frequency of 7.5 MHz.

Thyroid gland images of different patients (Fig. 1) were processed for classification features extraction. At first images were preprocessed by speckle noise filter [4]. Thyroid gland ROI boundaries were selected by the physician. Quantitative characteristics of image texture were calculated using Haralick's texture features [8]. These characteristics of image texture are extracted as intensity-based statistical features. The basis for these features is the gray level co-occurrence matrix. This matrix is square with $N \times N$ dimension, where N is 256 that corresponds to the number of gray levels in the image. Matrix element ij is generated by counting the number of pixel's occurrences with value i adjacent to a pixel with value j and then dividing the entire matrix by the total number of such comparisons made. After that the mean intensity, maximum intensity, minimum intensity, central pixel's intensity, variance, standard variance, median intensity, skewness, kurtosis, correlation, covariance, inertia, entropy, energy, inverse difference moment of the ROI were computed. Total number of features was 800.



a

b

Fig. 1. Ultra sound introscopy image for features extraction: *a* – initial USI-image of thyroid gland tissue; *b* – region of interest selection made by physician

Patients were subdivided according to nosological state of thyroid gland (without pathology, hypothyrosis, diffusive goiter, thyrotoxicosis, chronic thyroiditis, nodular goiter, combined goiter) and patients' age (younger than 30 years, from 30 to 40 years, from 40 to 50 years, from 50 to 60 years, older than 60 years).

Kohonen neural network is a nonlinear model for data clustering and classification [9]. Such a network is a rectangular grid of nodes called neurons. Each neuron has vector of coordinates w_i in fea-

tures space. Each object corresponds to features vector x_i. For each object the distance $d_i = |w_i - x|$ to each neuron is calculated and the winner neuron with the less distance is selected. Then coordinates of the winner neuron are modified

$$w_i(t+1) = w_i(t) + r(x_i - w_i(t)),$$

where r is small parameter.

The Kohonen map of size 3×3 was chosen. It allows classifying data into 9 clusters corresponding to diagnosis. After convergence of iterations the probability to obtain appropriate diagnose while hitting certain neuron was computed. It was shown [5] that features of the same diagnosis hit different neurons with compared probability. It means that classical Kohonen algorithm requires modifications. The improvement involves power coefficients p_i that defines the influence of different features on the distance

$$d^2 = \sum_{i=1}^{N} p_i (x_i - w_i)^2.$$

System was trained to recognize each diagnosis separately by selecting the optimum power coefficients combination.

This approach is proved to provide significant improvements in true positive (sensitivity) and false negative (specificity) rates (Fig. 2). For instance for thyrotoxicosis and mixed diagnoses the true positive rate of prediction is more than 90% and false negative rate is more than 95 %.

Thus selection of appropriate features provides significant improvements to classification accuracy. Selection of such features evidently is not related to neural networks and may be applied to much simpler techniques as linear functional strategy (LFS) and one can expect improvements in this case also. On the other hand, LFS requires much less data and computing resources for training and classification.

Linear Functional Strategy for thyroid dataset

The thyroid dataset was created from real medical tests screening for hypothyroid problems [10]. Based on the patient query data and patient examination data, the task is to determine whether a patient thyroid has overfunction, normal function, or underfunction. Therefore three classes are built: normal (not hypothyroid), hyperfunction and subnormal functioning. Since most people were healthy 92.58 % of cases belong to the normal group. 21 medical tests were made in most cases, with 6 continuous and 15 binary values, and about 10 % of values missing. A total of 7200 cases are given: 3772 results from one year, and 3428 from the next one. Thus from the classification point of view this is a 3 class problem with 22 attributes.

One of the first attempts in multi-class clustering by means ANNs was done in [11]. The training set here consists of 3772 measurement vectors from the first year, and 3428 measurements of the second year are available for testing. It was shown that it's hard to train Backpropagation ANNs with the dataset: for fixed fully interconnected 3-layer network the best achieved CEP is at most 2.42 %, for batch mode even worse (7.15 %). It was shown, that with various learning rate adaptation techniques the best achieved CEP is 1.55 %. The cascade correlation algorithms clearly outperforms all other algorithms (1.52 % of CEP), but it is not directly comparable with them, as it differs in many ways, as the architecture of the network is not fixed, new hidden units are trained and added one by one.

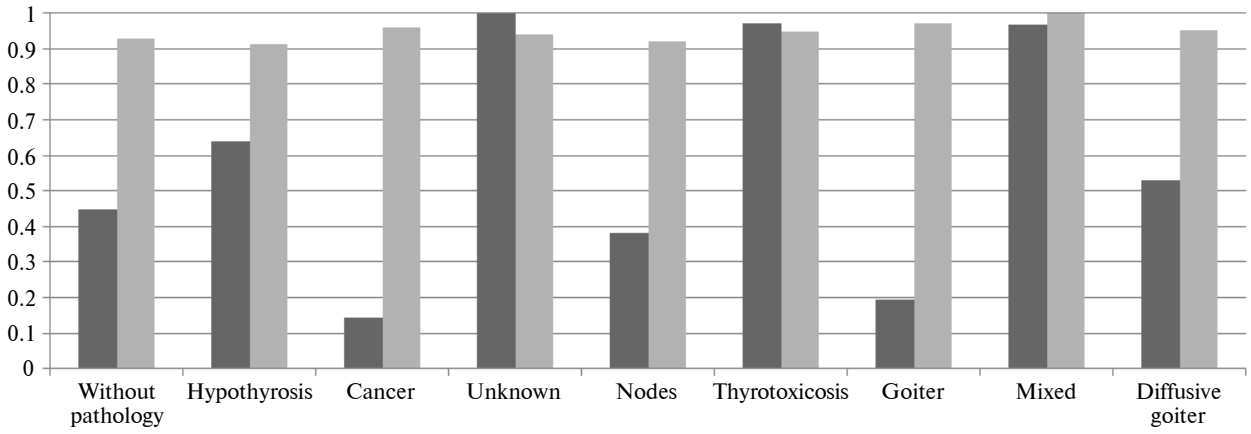


Fig. 2. Performance of modified Kohonen neural network-based classifier in terms of sensitivity and specificity for each class correspondingly: ■ — sensitivity; ■ — specificity

Several algorithms were used to train ANNs in [12]: the Backpropagation algorithm, the Levenberg–Marquardt algorithm, a Generic Algorithm, a hybrid between Generic Algorithm and Backpropagation, and a hybrid between Genetic Algorithm and Levenberg-Marquardt. The thyroid dataset was split into two sub-groups: one for training the ANN (5400 of all 7200 items in the dataset, i.e. 75 %), while the testing set is the remaining 1800, or 25 % of the items. The architecture of the network (multilayer perceptron) used for the task is composed of three layers: the input with 21 neurons, output with 3 neurons and additional hidden layer with 6 neurons. The activation function used in every artificial neuron of the network is the sigmoid function. It was shown that all algorithms except the Levenberg–Marquardt algorithm always obtain 7.28 % of CEP, while the latter one gives 2.14 %.

The conventional neural networks with various training algorithms seem to be too complicated and easy to overfit, therefore LFS seems to be more appropriate.

We split the training set into two subsets: training set per se, and cross-validation set. This cross-validation set is used for selection of LFS coefficients. For each class i we consider ranking function $y_j^i(\gamma)$ for various values of parameter γ of Gaussian kernel

$$K(x, \bar{x}) = \exp\left(-\frac{\|x - \bar{x}\|^2}{\gamma}\right), \quad \gamma = 0.1, 0.5, 1, 5, 10.$$

As the problem is ill-posed, some regularization is required. We use multi-parameter regularization, which allows for simultaneous feature selection. Then for aggregated risks $y_j^i(\gamma)$ we use all training data to construct y_j^i. The proposed technique allows selecting 5 most important features that gives accuracy 99.5 % on training set and 99.07 % on test set, that obviously outperforms previously described conventional ANNs. Note, that multi-parameter regularization and corresponding feature selection increase performance in term of accuracy from approximately 95 % to nearly 99 %.

Conclusions

We studied ANN-based approach to classification problem on medical datasets. It was shown, that the Kohonen neural network algorithm is efficient for datasets with large amount of features (800), whereas for datasets with relatively small amount of features (21) linear functional strategy is more reliable approach comparing to classical ANNs.

Unfortunately the most popular training algorithms for ANNs can be very slow for practical applications. Moreover, in case of large data dimension or large number of classes (number of output units) computational complexity of algorithm increases. Considering the actual level of data distribution (data stored in geographically distributed clinics and organizations), classification becomes a challenging problem. Fortunately, large-scale computational and data intensive ANN based classification can be efficiently applied on Grid system (see [5]).

Another issue of ANN-based classification is that its functioning is a “black box”, it does not give any insights on the structure of function being approximated. Therefore logical rules should be preferred over ANN-based classifications and other methods, provided that the complexity of the set of rules will not be too large and their accuracy will be sufficiently high. This is especially true in medical applications, where understanding which features contribute to classification and which are irrelevant is of extreme importance. Moreover, feature selection allow for detailed control over complexity of the classification model. Proper features extraction and selection are shown to be significant also for such important applied problems as epilepsy detection and prediction [13, 14]. In this way, development of regularization methods for causality detection and feature selection becomes a crucial task for further research. This research is supposed to be efficient in combination with simulations of underlying physical, physiological and other processes [15, 16] for extraction of features relevant for classification.

References

[1] J.A. Noble, “Ultrasound image segmentation and tissue characterization”, *Proc. In. Mech. Eng. H, J. Eng. Medicine*, vol. 223, pp. 1–10, 2009. doi: 10.1243/09544119JEIM604

[2] E.C. Kyriacou *et al.*, “A review of noninvasive ultrasound image processing methods in the analysis of carotid plaque morphology for the assessment of stroke risk”, *IEEE Trans. Inform. Technol. Biomed.*, vol. 14, iss. 4, pp. 1027–1038, 2010. doi: 10.1109/TITB.2010.2047649

[3] G. Slabaugh *et al.*, “Statistical region-based segmentation of ultrasound images”, *Ultrasound Med. Biol.* vol. 35, no. 5, pp. 781–795, 2009. doi: 10.1016/j.ultrasmedbio.2008.10.014

[4] M.K. Novoselets *et al.*, "Ultrasound introscopic image quantitative characteristics for medical diagnostics and refinements of physical noise rise reasons", *Proc. SPIE, Medical Imaging*: *Physics of Medical Imaging*, vol. 2163, 1994, pp. 442–446. doi: 10.1117/12.174281

[5] R.V. Byvalkevich *et al.*, "Tools for ultrasonic diagnostic image classification", in *Proc. 8th IEEE Int. Conf. Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, Warsaw, Poland, 2015, pp. 977–981. doi: 10.1109/IDAACS.2015.7341450

[6] G. Kriukova *et al.*, "A linear functional strategy for regularized ranking", *Neural Networks*, vol. 73, pp. 26–35, 2016. doi: 10.1016/j.neunet.2015.08.012

[7] P. Tkachenko *et al.*, "Prediction of nocturnal hypoglycemia by an aggregation of previously known prediction approaches: proof of concept for clinical application", *Comput. Methods Programs Biomed*, vol. 134, pp. 179–186, 2016. doi: 10.1016/j.cmpb.2016.07.003

[8] R.M. Haralick *et al.*, "Textural features for image classification", *IEEE Trans. Syst. Man Cybern.*, vol. SMC-3, no. 6, pp. 610–621, 1973. doi: 10.1109/TSMC.1973.4309314

[9] S. Aksenov and V. Novoseltsev, *Neural Networks Organization and Using (Methods and Techniques)*. Tomsk, Russia, 2006 (in Russian).

[10] M. Lichman, *UCI Machine Learning Repository*. Irvine, CA: University of California, School of Information and Computer Science, 2013. Available: http://archive.ics.uci.edu/ml/datasets/Thyroid+Disease

[11] W. Schiffmann *et al.*, "Comparison of optimized backpropagation algorithms", in *Proc. Europ. Symp. Artificial Neural Networks (ESANN'93)*, M. Verleysen Ed., Brussels, Belgium: de Facto Press, 1993, pp. 97–104.

[12] E. Alba and F. Chicano, *Generic Algorithms*, in *Metaheuristic Procedures for Training Neural Networks (Operations Research/Computer Science Interfaces Series 35)*, E. Alba and R. Marti, Eds. Springer Science & Business Media, 2006, pp. 120–138.

[13] I. Osorio *et al.*, "Performance reassessment of a real-time seizure-detection algorithm on long ECoG series", *Epilepsia*, vol. 43, no. 12, 2002, pp. 1522–1535. doi: 10.1046/j.1528-1157.2002.11102.x

[14] W. Christian, *Automatic Seizure-Detection in Rats using ECoG: A Comparison of Methods*. Lübeck, Germany, 2013.

[15] A. Salnikov *et al.*, "Integrated grid environment for massive distributed computing in neuroscience", in *Proc. 6th IEEE Int. Conf. Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, Prague, Czech Republic, 2011, pp. 198–202. doi: 10.1109/IDAACS.2011.6072740

[16] V.O. Gaidar and O.O. Sudakov, "Archiving and analysis of electroencephalograms in Ukrainian grid: the first application", in *Proc. 8th IEEE Int. Conf. IDAACS*, Warsaw, Poland, 2015, pp. 961–965. doi: 10.1109/IDAACS.2015.7341447

Г.В. Крюкова, С.П. Радченко, О.О. Судаков

ШТУЧНА НЕЙРОННА МЕРЕЖА ДЛЯ МУЛЬТИКЛАСОВОЇ КЛАСИФІКАЦІЇ ТА ЇЇ ЗАСТОСУВАННЯ ДО ВИЗНАЧЕННЯ ФУНКЦІОНАЛЬНОГО СТАНУ ЩИТОВИДНОЇ ЗАЛОЗИ

Проблематика. Розробка засобів автоматичної діагностики вимагає вибору й удосконалення відповідних методів машинного навчання, зокрема мультикласової класифікації. Щоб адресувати цю задачу, розглядаються штучні нейронні мережі (ШНМ) різних конструкцій.

Мета дослідження. Метою роботи є аналіз та порівняння ефективності класифікаторів, що базуються на ШНМ, на різних даних для подальшого вдосконалення стратегії вибору моделі.

Методика реалізації. Побудовано моделі на основі ШНМ розподілу класових міток ознак предиктора, треновані та оцінені на клінічних даних. Досліджено різні алгоритми навчання багатошарового персептрона, нейронну мережу Кохонена та лінійну функціональну стратегію з багатопараметричною регуляризацією.

Результати досліджень. Ефективність класифікаторів порівнюється з точки зору точності, чутливості та специфічності. Класифікатор, побудований за допомогою лінійної функціональної стратегії, випереджає інші, що базуються на ШНМ зі складною архітектурою, а також демонструє відносну стійкість до перенавчання. Продуктивність нейронної мережі Кохонена на великих наборах даних по специфічності перевищує 90 % для кожного класу, одночасно з тим чутливість для окремих класів перевищує 95 %.

Висновки. Розуміння сильних і слабих сторін кожного методу є надзвичайно важливим для ретельного добору класифікатора на основі ШНМ, зокрема його архітектури, алгоритмів регуляризації та навчання.

Ключові слова: штучна нейронна мережа; класифікація; обробка ультразвукових зображень; регуляризація; обернена задача.

Г.В. Крюкова, С.П. Радченко, А.А. Судаков

ИСКУССТВЕННАЯ НЕЙРОННАЯ СЕТЬ ДЛЯ МУЛЬТИКЛАССОВОЙ КЛАССИФИКАЦИИ И ЕЕ ПРИЛОЖЕНИЕ К ОПРЕДЕЛЕНИЮ ФУНКЦИОНАЛЬНОГО СОСТОЯНИЯ ЩИТОВИДНОЙ ЖЕЛЕЗЫ

Проблематика. Развитие методов автоматической диагностики требует выбора и совершенствования подходящих методов машинного обучения, в частности мультиклассовой классификации. В качестве подхода к этой задаче рассматривается искусственная нейронная сеть (ИНС) разного строения.

Цель исследования. Целью работы является анализ и сравнение производительности классификаторов, базирующихся на ИНС, на различных данных для дальнейшего совершенствования стратегии выбора модели.

Методика реализации. Построены модели на основе ИНС распределения классовых меток в зависимости от признаков предиктора, обучены и проверны на клинических данных. Рассмотрены различные алгоритмы обучения мультислойного перцептрона, нейронная сеть Кохонена и линейная функциональная стратегия с многопараметрической регуляризацией.

Результаты исследований. Производительность классификаторов сопоставляется относительно точности, чувствительности и специфичности. Классификатор на основе линейной функциональной стратегии опережает другие на основе более сложных ИНС, а также демонстрирует сравнительную устойчивость к переобучению. Продуктивность нейронной сети Кохонена на больших наборах данных по специфичности превышает 90 % для каждого класса, при этом чувствительность для отдельных классов превышает 95 %.

Выводы. Понимание сильных и слабых сторон каждого метода имеет особое значение для точного выбора классификатора на основе ИСН, в частности его архитектуры, алгоритмов регуляризации и обучения.

Ключевые слова: искусственная нейронная сеть; классификация; обработка ультразвуковых изображений; регуляризация; обратная задача.

Рекомендована Радою
факультету прикладної математики
КПІ ім. Ігоря Сікорського

Надійшла до редакції
27 грудня 2016 року