Liliana BYCZKOWSKA-LIPIŃSKA, Agnieszka WOSIAK

Institute of Information Technology, Technical University of Lodz

Image signal metadata analysis in diagnosis of Alzheimer's disease

Abstract. *The paper presents the possibility of using the methods of multimedia databases and imaging data analysis in the diagnosis of Alzheimer's disease. The aim of this study was to investigate for the statistical analysis of the available data, especially signal data. In the study, the relationships between test results obtained in different forms have been examined. The correlations between the external metadata for DICOM objects and signal selected metadata extracted from images have been explored.*

Streszczenie*. W pracy przedstawione zostały możliwości zastosowania metod multimedialnych baz danych oraz analizy danych obrazowych w diagnostyce choroby Alzheimera. Celem przeprowadzonych badań było zbadanie pod kątem analizy statystycznej dostępnych danych, przede wszystkim danych sygnałowych. W ramach badań sprawdzone zostały zależności pomiędzy wynikami badań pozyskiwanych w różnych formach. Zbadane zostały zależności korelacyjne pomiędzy metadanymi zewnętrznymi dla obiektów DICOM a wyłuskanymi z obrazów wybranymi metadanymi sygnałowymi. (**Analiza danych obrazowych w diagnozowaniu medycznym choroby Alzheimera**).*

Słowa kluczowe: systemy informatyczne w medycynie, obrazowanie medyczne, systemy diagnostyczne, analiza danych
Keywords: computer medical systems, medical imaging, diagnostic systems, data analysis

Introduction

Technological capabilities in the field of medical imaging contribute to the increasing use of image analysis in diagnostic medical systems. The image data can come from a number of tomographic studies, including ultrasonography (USG), computed tomography (CT) and magnetic resonance imaging (MRI). Medical imaging is one of the key sources of information for health professionals, which is largely due to the fact that the accuracy of the conclusions drawn by the physicians from this form of presentation of the data is very high in comparison with other forms (verbal description and numerical data) [1]. Therefore, the implementation of functions for the medical imaging in medical diagnostics computer systems is desirable, or even necessary. The process of collecting and storage of medical imaging results does not present a problem using almost any database system, however, the analysis of this type of information is a complex issue that requires much more advanced information technology [2].

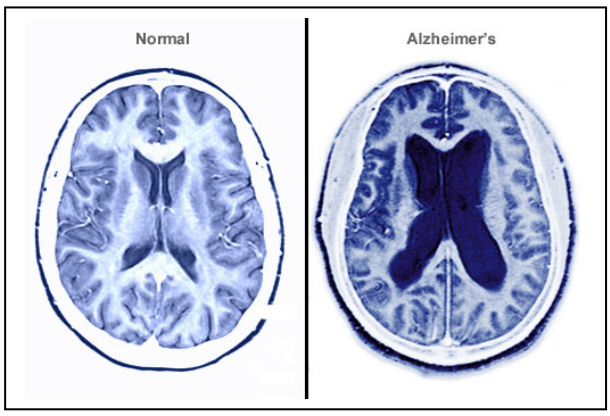

Fig. 1 MRI image of the brain of a healthy patient (left) and a patient with Alzheimer's disease (right) [3].

Currently, imaging studies are used in many applications. Medical imaging is a key element in the diagnosis of atrophic and degenerative changes of the bone - joint and gives an accurate picture of the position changes of malignancy [4, 5]. The results of imaging studies, in particular the magnetic resonance imaging, are particularly useful in detection of multiple sclerosis, vascular disorders and Alzheimer's disease (Fig. 1). This last case of medical imaging application is the subject of this paper.

Medical diagnostics

Special, therapeutic devices are sources of numerous data in medical, diagnostic systems, and this is the reason for development of computer, medical software. This is also the feature that differs medical systems from other diagnostic systems. A huge number of data collected makes human consideration of all factors impossible. Therefore, there is a strong need for use of advanced analytical tools, which have been primarily related to business or industrial processes [6,7].

The data collected in the system can be of different forms: numerical, text or multimedia. Information about the visual characteristics plays an extremely important role in the diagnosis, treatment planning and monitoring. This is due to the fact that this type of data is easily legible and particularly effectively interpreted by physicians and other medical staff.

Database, that meet the expectations of medical diagnostics including imaging, require specific objectives of the project and more resources than databases built for other purposes. Therefore, multimedia databases can be widely used in medicine, because they offer additional software solutions for complex multimedia data.

In medical, diagnostic systems, important functions belong to modules that are responsible for data mining. With the methods of data analysis, in particular statistical analysis it is possible to detect patterns and dependencies that have not been reported, and thus to improve the quality of diagnosis. In carrying out the medical data, first it is necessary to identify the most important attributes of the assessment of the patient, and then to apply the selected method to check the relationship between the values of various attributes and between attribute values and the classification of patients.

Medical imaging

Important issues in the technical development of tools for collecting diagnostic information are medical imaging techniques.

The main applications of medical imaging include:
- visualization,
- quantitative analysis or quantitation,
- localization,
- screening.

The basic method for the use of medical images is to collect them and share. The medical computer systems called RIS (Radiological Information System) and PACS

(Picture Archiving and Communication System) are dedicated for this applications. The results of the analysis of the image (the original, or subjected to a predefined computer processing operations) can be made available to physicians for their evaluation and interpretation, as well as stored in the hospital database (HIS), but may be also subject to further, computer assisted interpretation, and therefore assist medical diagnosis.

Searching and comparing multimedia objects, in particular medical imaging, requires obtaining information about the content stored in a graphical object (metadata). Metadata extracted are the source of information for the user of the system, but also for the system itself, which on this basis can automatically analyze the data. There are three basic types of metadata: external, signal, and semantic. External metadata format for medical imaging has been standardized by the DICOM standard (Digital Imaging and Communications in Medicine) awarded by the National Electrical Manufacturers Association. This standard defines a way to store bits of data that make up the image, as well as about 2,000 attributes defining the metadata associated with the image and additional information (eg. patient data, case study) [8].

Obtaining the signal metadata is associated mainly with extracting color information and its location, shape and texture from the content of the image. The extraction of this information is often complex and time consuming hence the permanent need to find new methods for extracting metadata [9,10].

Multimedia databases make it possible to compare the internal metadata of images, which is extremely important in the medical analysis of individual cases. It is possible to compare the average color (SI.AverageColor), distribution of color in the image (SI_PositionalColor), color histogram (SI_ColorHistogram), texture (SI_Texture) and all the features with specific weights (SI_FeatureList).

The value of the adjustment of the image properties to the standard list of properties is defined by the weighted average for each property. The specificity of medical images requires a modification of the weighting factors. Defining the value of the weighting factors can be realized by the system user during the analysis of imaging data. The significance of individual features in such cases is determined arbitrarily and adjusted to the search results. These actions may require multiple iterations, and consequently they are time consuming. Use the relevance feedback in the process of comparing of medical images enables determining the significance of individual features automatically, and the same releases the user from the decision-making process.

The idea of relevance feedback in information retrieval (IR) systems is to use a pre-search results returned on the basis of the available features and to perform re-query dependent on the similarity of results obtained in the first step. The process can be repeated many times until a satisfactory result.

During the implementation of this algorithm, it is necessary to consider the method of determining k most similar results (ranking function) and to determine the weight factors (significance) descriptors for each of the features. Determination of weighting factors for descriptors of similar features in a multimedia medical database can be carried out considering the variance of the features of the entire set of images, and in the set specified by the top k [11].

Experimental studies

For the purpose of this study, a diagnostic medical system module was implemented. Two groups of patients were evaluated: a control group of healthy cases and a group of test patients with known Alzheimer's disease. The basic element of analysis in both groups was the result of magnetic resonance imaging. According to the National Institute of Neurologic, Communicative Disorders and Stroke / Alzheimer 's Disease and Related Disorders Association criteria for scientific recognition of Alzheimer's disease include:

- The presence of early and significant episodic memory impairment, which is the primary criterion,
- Noticing the failure of the medial temporal lobe documented by MRI (Fig. 2) - reducing the volume of the hippocampus, entorhinal cortex, amygdala nuclei, the changes are a supporting criterion.

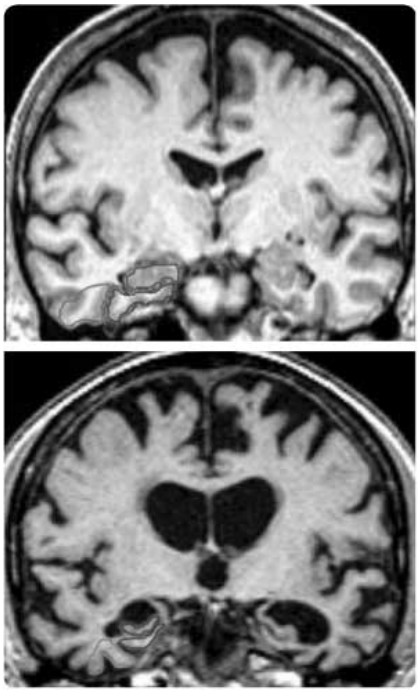

Fig. 2. MRI image of the healthy patient's brain (top) and patient with Alzheimer's disease (bottom photo) with marked atrophic areas [12].

The aim of experimental studies has been examination of the relationship between examination results obtained in different forms: numbers, text and graphics. The correlations between the external metadata for DICOM objects and selected, signal metadata extracted from images selected have been examined. The significance of the selected, imaging attributes was assessed on the basis of relevance feedback techniques and determined by the standard deviation for the attributes.

The analysis of the relationship between the parameters studied, both in the control group and in the group studied, has been performed by calculating the Pearson correlation coefficient or Spearman's rank correlation coefficient. The selected results are shown in Table 1 and on the graphs (Fig. 3 and Fig. 4). The threshold for statistical significance was set at $p < 0.05$.

The significant correlation was found between the selected visual attributes and known Alzheimer's disease.

Tabel 1. Selected correlations

Correlation type	r(X,Y) value	Significance p
SI_AverageColor - AD	0,33	0,03
SI_PositionalColor loc. 1 - AD	0,50	0,00
SI_PositionalColor loc. 2 - AD	0,38	0,02
SI_PositionalColor loc. 3 - AD	0,37	0,01
SI_PositionalColor loc. 6 - AD	0,36	0,02
SI_PositionalColor loc. 8 - AD	0,49	0,00

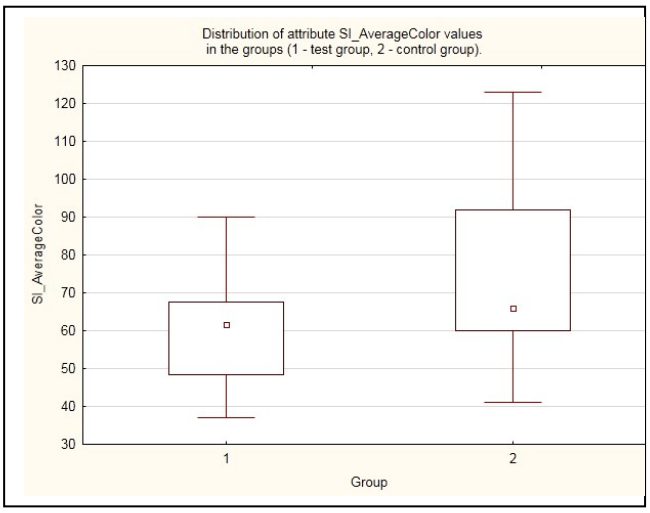

Fig. 3. Distribution of attribute SI_AverageColor values in the groups (1 - test group, 2 - control group).

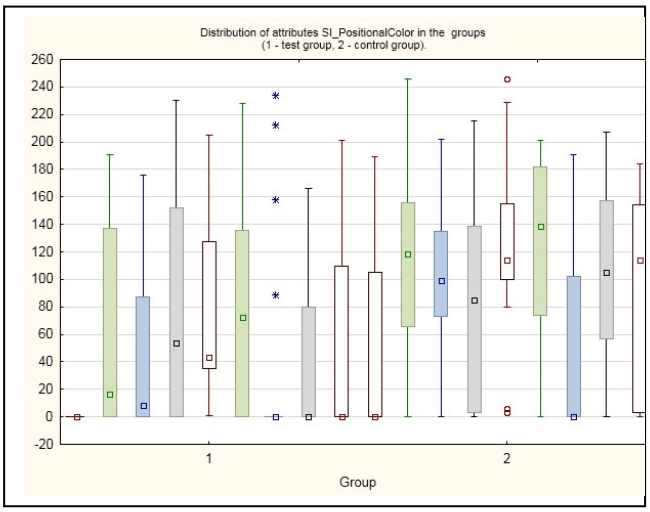

Fig. 4. Distribution of attributes SI_PositionalColor in the groups (1 - test group, 2 - control group).

Experimental studies were carried out using the Oracle Database tools [13]. Data for the analysis were obtained from the research unit servers, providing their resources for scientific purposes (among others. MedPix - Medical Image Database and Radiology Portal, Science Photo Gallery).

The results of the analysis have shown that the signal inclusion image metadata in the analysis process is reasonable, and the method itself is so fast and effective that it can become the initial diagnostic phase of the screening

Conclusions

The paper presents the possibility of using the methods of multimedia databases and imaging data analysis in the diagnosis of Alzheimer's disease. The aim of this study was not to establish the existence of the disease, but tested for statistical analysis of the available data, in particular signal data.

Most of the existing diagnostic systems allows to perform statistical analysis on the basis of numerical data, without examining their relationship with the image data. In this study, the relationship between test results obtained in different forms has examined. The correlations between the external metadata for DICOM objects and selected, signal metadata extracted from images have been calculated. The significance of the selected image attributes was assessed on the basis of relevance feedback techniques and determined by the standard deviation of the visual attributes. The relationship between the signal coming from the metadata of MRI and diagnosed with Alzheimer's disease has been stated.

These conclusions may lead to the statement that that the evaluation of image data in terms of signal metadata may be the initial stage of automatic medical diagnosis.

REFERENCES

[1] Tadeusiewicz R.: Informatyka medyczna, Uniwersytet Marii Curie-Skłodowskiej w Lublinie, Instytut Informatyki, Lublin 2011

[2] Stanchev P. L., Fotouhi F.: MEDIMAGE – A Multimedia Database Management System for Alzheimer's Disease Patients, Lecture Notes in CS 2314, Recent Advances in Visual Information Systems, 2002, pp. 187 - 193

[3] http://www.doctortipster.com/7491-alzheimers-disease-research.html

[4] Daoqiang Zhang, Yaping Wang, Luping Zhou, Hong Yuan, Dingngang Shen: Multimodal classification of Alzheimer's disease and mild cognitive impairment, NeuroImage, Vol. 55, Elsevier Ireland Ltd 2011, pp. 856–867

[5] Matoug S., Abdel-Dayem A., Passi K., Gross W., Alqarnil M.: Predicting Alzheimer's disease by classifying 3D-Brain MRI images using SVM and other well-defined classifiers, Journal of Physics: Conference Series 341 (2012) 012019

[6] Wiak S., Drzymała P., Welfle H.: Using ORACLE tools to generate Multidimensional Model in Warehouse, Przegląd Elektrotechniczny, ISSN 0033-2097, R. 88 NR 1a/2012, pp. 257-262

[7] Rymaszewski J., Lebioda M., Korzeniewska E.: Propagation of normal zone in superconducting tapes due to heating in near-electrode area, Materials Science and Engineering: B, Volume 176, Issue 4, 15 March 2011, pp. 334-339

[8] National Electrical Manufacturers Association: Digital Imaging and Communications in Medicine (DICOM), 2009

[9] Stasiak B., Yatsimirsky M.: Frequency Domain Methods for Content-Based Image Retrieval in Multimedia Databases, Methods and Supporting Technologies for Data Analysis, D. Zakrzewska,E. Menasalvas, L. Byczkowska-Lipinska (Eds), Springer 2009, pp. 137 – 166

[10] Pryczek M., Tomczyk A., Szczepaniak P. S.: Active Partition Based Medical Image Understanding with Self-Organised Competitive Spatch Eduction, Journal of Applied Computer Science, Vol. 18, No. 2, 2010, pp. 67–78

[11] Aksoy S., Haralick R.M., Cheikh F.A. Gabbouj M.: A weighted distance approach to relevance feedback, Proceedings of the 15th International Conference on Pattern Recognition, 2000, Vol. 4, pp. 812-815

[12] http://www.elements4health.com/mri-scans-accurately-diagnose-alzheimers-disease.html

[13] Oracle® Multimedia DICOM Developer's Guide 11g Release 2 (11.2), Oracle Database Documentation Library

***Authors**: prof. dr hab. inż. Liliana Byczkowska-Lipińska, Politechnika Łódzka, Instytut Informatyki, ul. Wólczańska 215, 90-924 Łódź, e-mail: liliana.byczkowska-lipinska@p.lodz.pl, dr inż. Agnieszka Wosiak, Politechnika Łódzka, Instytut Informatyki, ul. Wólczańska 215, 90-924 Łódź, e-mail: agnieszka.wosiak@p.lodz.pl*