Features of Implementation of Prototype Integrated Expert System for Comprehensive Diagnosis of Breast Diseases: Integration with Image Processing

Galina V. Rybina [1] and Andrey A. Grigoryev [2]

[1] *National Research Nuclear University "MEPhI", Kashirskoe highway 31, Moscow, 115409, Russian Federation*
[2] *National Research Nuclear University "MEPhI", Kashirskoe highway 31, Moscow, 115409, Russian Federation*

Abstract

The paper is devoted to the further development of research and development in the field of integrated expert systems (IES), the construction of which is carried out on the basis of a problem-oriented methodology.The focus is on the integration of the basic components of the IES with new functional components that implement some methods of image processing obtained as a result of X-ray and ultrasound studies. Complex diagnostics of breast diseases was considered as an experimental problem area for the implementation of the IES prototyping processes.

Keywords

Integrated expert systems, problem-oriented methodology, AT-TECHNOLOGY workbench, integration, image processing, medical diagnostics, mammary gland

1. Introduction

Interest in the creation of applied intelligent systems intended for such a socially significant problem area (PA) as medicine is constantly increasing and is constantly being updated in accordance with the emergence of new medical technologies and the improvement of the used firmware and platforms, which significantly expands the range of used models, methods and means of artificial intelligence, for the tasks of medical diagnostics in particular.

It should be noted that the traditional view of the possibilities of effective use of AI methods and technologies is usually related to the description of the goals and objectives of various knowledge-based systems, expert systems (ES) in particular. The main emphasis here is on supporting decision-making in order to improve the quality and accuracy of doctors' work in identifying the most complex and difficult to diagnose diseases, detecting signs of diseases in the early stages and prescribing medical measures, for training medical personnel, etc. [1-6]. An important requirement in medical expert systems is always the presence of an explicit description of the knowledge used, as well as the implementation of convenient consultation regimes for doctors with expert systems explaining the results of diagnostics.

In recent years, the ontological approach has begun to actively develop, within the framework of which various applied ontologies are created, for example, the ontology "Medical Diagnostics" [7] and others are widely known. The issues of building some universal ontology of knowledge about diagnostic processes in various PAs are also being discussed, which in general can be useful for developers of intelligent systems [8].

However, a significant "breakthrough" in the field of medical technologies focused on the use of a wide range of new types of equipment for disease diagnostics, showed the need to integrate the methods

Russian Advances in Fuzzy Systems and Soft Computing: Selected Contributions to the 8th International Conference on "Integrated Models and Soft Computing in Artificial Intelligence (IMSC-2021)", May 17–20, 2021, Kolomna, Russian Federation
EMAIL: GVRybina@mephi.ru (A. 1); grigandal625@gmail.com (A. 2);
ORCID: 0000-0002-4077-3660 (A. 1)

and tools of traditional ES with image processing systems (ultrasound, X-ray, etc.), radiation and computed tomography systems, big data processing (Data Mining, Text Mining), etc.

Research in the field of integration processes emerging from the joint use of semantically heterogeneous objects, components, tools, etc., as well as the presence of a holistic methodology and technology for the design and development of integrated intelligent systems, are especially important for overcoming the autonomy of software and hardware technologies and their lack of influence on decision-making processes. Of course, at this stage we are not talking about overcoming hardware dependence on types of electronics and equipment, but we are only talking about developing complex approaches to disease diagnosis processes, both through knowledge engineering and by reducing the influence of the human factor in the perception and interpretation of visual information in images.

Currently, the most developed in the context of the integration problem and well tested in practice for solving problems of real complexity and significance are the architectures of IES with flexible extensible functionality and a scalable architecture, developed on the basis of the problem-oriented methodology [9] and the intelligent software environment of the AT- TECHNOLOGY workbench [9-11].

The purpose of this work is to expand the ontology of typical architectures of IES by developing models and methods for integrating the basic components of integrated expert systems with new functional components that implement some methods for processing the results of ultrasonic and X-ray examinations. As an experimental problem area for the implementation of prototyping processes of IES, a comprehensive diagnosis of breast diseases was considered, including clinical, radiological (mammographic) and ultrasound diagnostics.

2. Features of implementation of technology application of integrated expert systems on the basic version of the AT-TECHNOLOGY workbench

Methods and techniques of implementing integration processes in an IES are determined by the basic principles of problem-oriented methodology [9], which is based on a multi-level integration model, modeling of specific types of tasks, relevant technologies of traditional expert systems, methods and techniques of constructing a program architecture of IES and its components at each level of integration with complex consideration of the specifics of solved problems for a given PA.

Modern experience of application of problem-oriented methodology and various versions of constantly developing tools for automated support of this methodology of intelligent software environment of AT- TECHNOLOGY workbench [9-11] already has dozens of applied integrated expert systems of different architectural typology, developed for tasks of diagnostics, design, planning, management and training, which are described in monographs [9,10] and various publications of recent years.

Accordingly, the technology of prototyping applied IES using the basic functional means of the AT-TECHNOLOGY complex has developed and is constantly being improved, for which refactoring is carried out in some cases, and new and modifiable components of the intelligent software environment [10.11], such as an intelligent scheduler, ontology of typical IES architectures, reusable components, typical design procedures, etc. are actively being developed and experimentally investigated.

Given the context of this work, let us consider some features of the process of prototyping applied IES, which is traditionally carried out in two stages: the development of a set of models provided for by a problem-oriented methodology (for example, for static PA, this is the IES architecture model, PA model, dialogue model, explanation model); design and software implementation of the IES prototype based on the tools of the basic version of the AT-TECHNOLOGY complex.

At the stage of system requirements analysis, a model of the architecture of the designed prototype of the IES is being built, based on the ideas of deep integration of components [9], which is associated with the possible expansion of the functionality of traditional ES by including functions implemented by other components. The average level of integration in the architecture model of each particular IES is determined by the specifications of function sets of both a simple ES (non-formalized operation) and functions that are not characteristic of the ES (formalized operation), which reflects the composition and structure of all components of the IES and their information and control connections. Based on this, the architecture model is a hierarchy of extended data flow diagrams (EDFDs).

The next important task is the construction of the PA model. This process involves automated knowledge acquisition using basic tools to support the combined knowledge acquisition method (CKAM) from different knowledge sources [9,10], resulting in the phased construction of an intermediate view (knowledge fields) and then after verification and consolidation of knowledge field fragments (if there are several sources of knowledge), they are converted to the knowledge base (KB) in the knowledge representation language provided for by the problem-oriented methodology.

From the point of view of analysis of peculiarities of stages of design and program implementation of individual components of IES prototype, two main aspects of technology of prototyping of applied IES based on AT- TECHNOLOGY workbench are to be highlighted. On the one hand, depending on the model of the IES architecture, the IES software is configured on the basis of selection for the use of some basic components (AT-SOLVER, explanatory component, etc.), and on the other hand, the IES prototype is expanded by developed specialized components, for example, an integration component with external software tools reflecting the specifics of a particular PA.

In this case, the functioning of the communication subsystem in the developed IES will result from the interpretation of the constructed model of dialogue with the user (consultation mode), presented in a special language of description of dialogue scenarios [9,10], which provides for the description and sending of messages to other components of the IES (internal and external).

Now we will directly consider the results of experimental software modeling in the context of implementing integration processes with the joint use of IES construction technology and methods for processing and interpreting images obtained from radiographic (mammographic) and ultrasound medical examinations (ultrasound) of breast diseases (BD).

3. General characteristics of problem area "Comprehensive diagnosis of breast diseases" and tasks related to image processing

When creating the IES demonstration prototype, the system analysis (examination) of this PA was carried out with the participation of experts - Kovynev A.V. (expert doctor in ultrasound equipment, researcher at the Department of Ultrasound Diagnostics and Surgery of the Peoples' Friendship University of Russia) and Mazo M.L. (Candidate of Medical Science, Senior Researcher, National Medical Radiological Research Center).

In agreement with the experts and taking into account [12,13], for the first stage of prototyping the IES, a range of informal tasks was identified in which the logic of decision making based on KB (PA model) includes the use of the results obtained during the processing of images on various shots. Below you may find short setting of those tasks.

Task of clinical radiographic diagnostics of lacteal gland (mammography)

Given: Set of lacteal gland mammograms in different projections. Clinical symptoms of female breast.

Required: to assess the general state of lacteal glands; to detect and, if possible, to diagnose primary formations.

Task of ultrasonic diagnostics of lacteal glands in B-Mode (sonography)

Given: ultrasound image of lacteal glands (sonogram). Mammography results.

Required: to re-evaluate the general state of lacteal glands; to detect and, if present (dependent and independent of mammography), to diagnose the formation.

Task of ultrasonic diagnostics of lacteal glands in Doppler mode (Colour Doppler Imaging (CDI)

Given: an ultrasound image of the lacteal glands with a color display of the places of blood flow and its level. Results of mammography and sonography.

Required: in the presence of education, the risk of malignancy shall additionally be assessed.

In addition, the formalized task of interpreting the initial features of lacteal glands by image was highlighted:

Given: ultrasound or RG-image of lacteal glands obtained as a result of patient examination.

Required: If there is education, to identify its signs for diagnosis, as well as identify signs for diagnosing of the general state of lacteal glands and diffuse changes.

The general diagram of the process of solving the problem of complex diagnostics of lacteal glands is given in Figure 1.

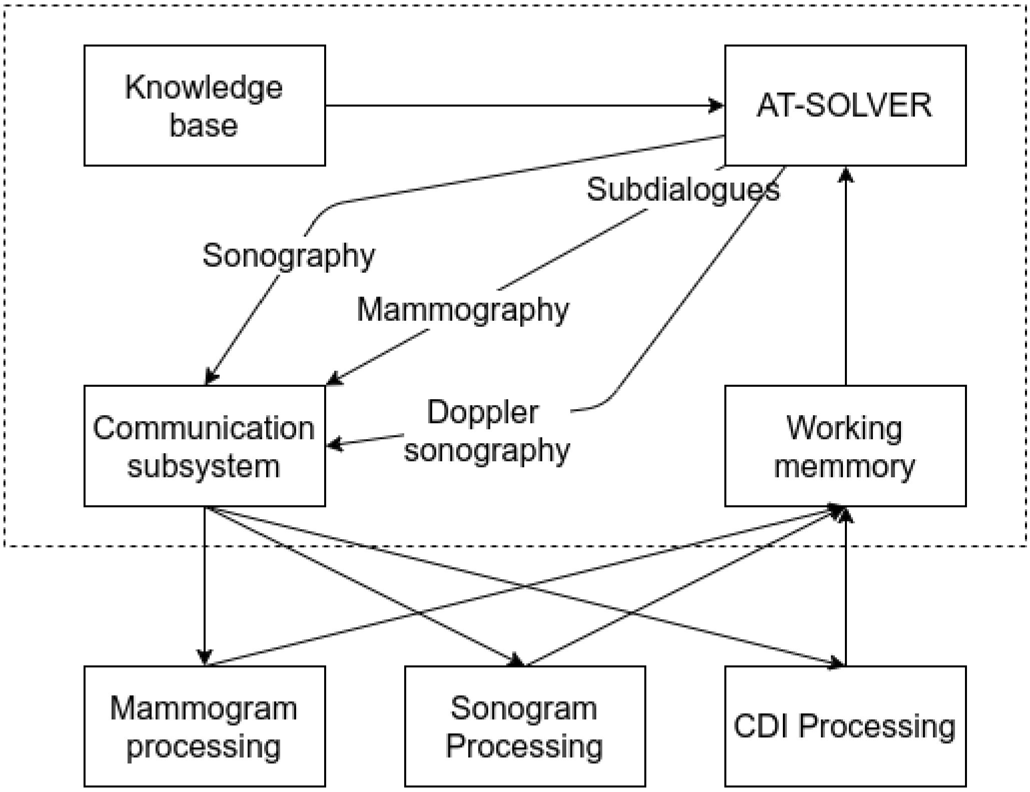

Figure 1: Diagnostic diagram

Let us briefly consider the features of selected and implemented methods of processing, interpreting and describing images in order to identify diagnostic features.

Currently, there are a fairly wide variety of image processing methods. However, it is important to notice that medical images of lacteal glands - mammograms and sonograms, are monochrome images, so their processing is reduced to processing the brightness level. The main goals solved by the doctor during the analysis of mammograms and echograms are: description of the general state of glandular and adipose tissues, search and determination of the nature of calcareous inclusions - diffuse analysis of mammograms and sonograms in general; localization of formations and identification of their diagnostic features - local analysis of sections of mammograms and sonograms.

In order to describe the state of glandular and adipose tissues, the tasks of determining the degree of severity and ratio of these tissues on a mammogram and an echogram are set. From the point of view of digital image processing, in order to solve these problems, it is necessary to first reduce the visible effects of various noise and image defects, and then classify the state of the tissue. Here, as shown in [15], the following brightness correction methods can be used to improve image rendering: logarithmic transformation; gamma correction (power transformation); piecemeal linear transformations; histogram image transformation; different methods of spatial filtering of images.

The classification of the state of lacteal glands tissues is carried out according to the indicators of various textural characteristics of the image, among which the most versatile, in accordance with [15], are: an adjacent matrix of gray tone levels; energy (second angular momentum); contrast (inertia); maximum probability; local uniformity (homogeneity); entropy (texture heterogeneity); trace of the matrices of the grayscale levels; mathematical expectation of gray tone; standard deviation of gray tone; histogram of image brightness.

Further, it is necessary to solve problems related to localization of formation and detection of its diagnostic, mainly, geometric and textural features [12-14]. The localization process at this stage is most conveniently assigned to the user-doctor, providing the necessary means for this by image cropping, since the doctor has more clear information about the location of the formation (if available) through a clinical examination of lacteal glands. Then, in order to interpret the diagnostic features of formation, it is necessary to first isolate the contours of malignant growth, and then describe and classify the selected contour, for which various gradient methods [15] are used (for example, the Marr-Hildreth algorithm, the Canny algorithm, the Prewitt operator). When solving the problem of describing and interpreting found contours, characteristics such as curvature of the boundary, as well as its statistical characteristics, are most often used.

It should be noted that a series of Python experiments using the Numpy, Python Image Library (PIL), Open-CV libraries in the Jupyter Notebook specialized environment were conducted to select the most suitable of all described image processing methods, based on the following conclusions:

- Gamma correction is the most suitable method for correction of image brightness;
- All described texture and contour characteristics are informative because they have a sufficiently tangible response;
- The path extraction method with the Pruitt operator is selected as the most efficient for this task;
- In order to eliminate noise during contouring, it is necessary to pre-localize the node formation (if any);
- in order to ensure greater accuracy of the contour extraction, it is necessary to perform threshold binarization of the image, wherein the binarization threshold is selected individually for each image.

The specialized software package "See 5" was used as a tool to build an image classifier. The selected methods and approaches to image processing served as the basis for the development and programmatic implementation of algorithms for detecting diagnostic signs of lacteal gland lesions by mammograms and echograms, namely: an algorithm for forming image classifiers, an algorithm for interpreting lacteal gland medical images using classifiers, etc.

4. Example of functioning of integrated expert system demonstration prototype for comprehensive diagnosis of breast diseases

Currently, the current KB volume for a given PA is about 120 rules, among which there are rules containing various types of NON-factors of knowledge, such as uncertainty, inaccuracy and nebulosity [9,10].

The functionality of the AT-SOLVER refinement subdialogs was significantly expanded in order to support calls not only to the user, but also to organize the interaction of IES components with external programs using the Scripter integration component, which provides calling functions and procedures in JavaScript and VETAcript languages. In order to do this, specialized software (SF) for processing medical images of the breast was developed and incorporated into the architecture, the interaction with which is provided by the integration component.

Let's take a brief look at how the prototype works in consultation mode.

After launching the consultation mode with the IES, a dialog box is displayed in which you need to register the patient and indicate such characteristics as: hereditary diseases; presence/absence of cancer, trauma, allergic diseases; the number of pregnancies, births and abortions; presence/absence of non-periodic lactations.

Then there is a transition to a clinical examination of the patient, where symptoms such as soreness, the presence of palpable formations, etc. are clarified. The clinical examination of the patient takes place in several stages depending on the symptoms introduced at each step, while the doctor has the opportunity to administer the symptoms taking into account his or her degree of confidence.

If the need for deeper diagnostics based on clinical examinations is identified, the consultation stage related to the processing of medical images of lacteal glands (mammograms and sonograms (echograms)) begins. Let us consider this step using the example of mammogram processing.

In this case, the doctor is given the opportunity to choose one of the mammographic projections (Figure 2.), on the basis of which the image is analyzed and the required diffuse signs of lacteal gland are distinguished: such as the state of tissues, the presence and type of calcareous inclusions, etc.

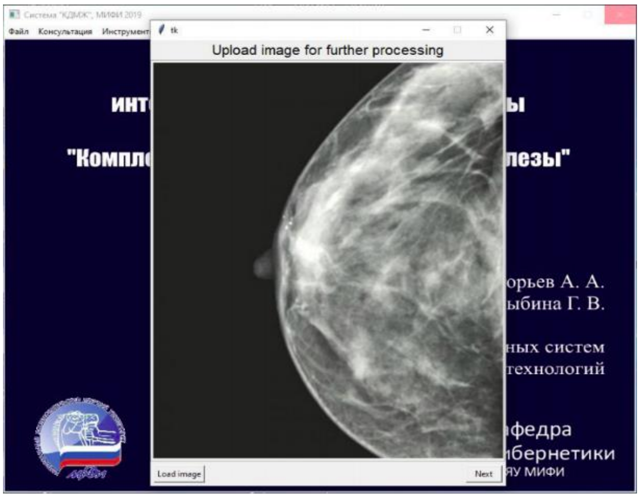

Figure 2: Loading a mammogram

Next, the doctor is asked to localize the formation by highlighting the rectangular area of the image containing the malignant growth (Figure 3). If there is no malignant growth, then this stage is skipped, and if the formation is localized, the selected part of the image will be analyzed and all necessary characteristics of the malignant growth (shape, appearance and clarity of the contours) will be determined.

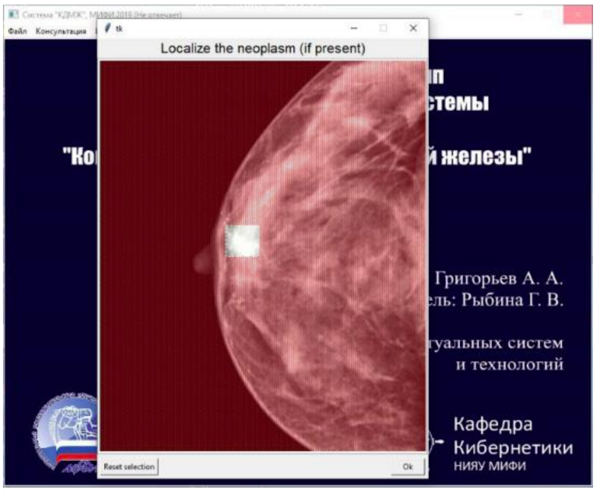

Figure 3: Malignant growth focalization

Based on the results of the consultation, the dialog box will display the results of diagnostics of the general condition of the breast and the found malignant growths, as well as possible recommendations.

5. Conclusion

The conducted research and study have shown a fairly high efficiency of the technology of prototyping applied IES using the basic functional means of the AT-TECHNOLOGY complex and the developed specialized software for image processing. It is planned to develop the demonstration prototype of the IES to the level of research and its experimental testing for this PA. Moreover, it is important to consider the possibility of creating a version in the context of the architecture of the training IES for educational purposes.

6. Acknowledgements

This work was carried out with the financial support of the Russian Foundation for Basic Research, project No. 18-01-00457

7. References

[1] Kobrinsky B.A. "Retrospective analysis of medical expert systems". Artificial Intelligence News No. 2, 2005, p. 6-17.

[2] Kobrinsky B.A. "The triunicity of confidence factors in the tasks of medical diagnostics". Artificial intelligence and decision-making. 2018. №2. p. 62-72.

[3] Singla J., Grover D., Bhandari A. "Medical expert systems for diagnosis of various diseases". International journal of computer applications. 2014. Vol. 93 - No.7, p. 36-43.

[4] Mohammed Abbas Kadhim, M. Asfar Alam, Harleen Kaur. "Design and implementation of Fuzzy expert system of back pain diagnosis". International journal of innovative technology & creative engineering. 2011 Vol. 1, No. 9, p. 16-22.

[5] Singla J. "The diagnoses of some lung diseases in a PROLOG expert system". International journal of computer applications. 2013. Vol. 78 - No.15, p. 37¬40.

[6] Gribova V.V., Ostrovsky G.E. "Intellectual learning environment for the diagnosis of acute and chronic diseases". Fifteenth national conference on artificial intelligence with international participation of CAI-2016 (October 3-7, 2016, Smolensk). The proceedings of the conference. T.Z. - Smolensk: University. 2016. p. 171-179.

[7] Kleshchev A.S., Moskalenko F.M., Chernyakhovskaya M.Yu. "Ontology model of the subject area "Medical diagnosis." Part 1. Informal description and definition of basic terms". STI. Series 2. 2005. №2. p. 1-7.

[8] V.V. Gribova, E.A. Shalfeeva. "Ontology of process diagnostics". Ontology of design. 2019. №4 (34). p. 449-461.

[9] Rybina G.V. Theory and Technology for the Construction of Integrated Expert Systems. - M.: Nauchtechlitizdat, 2008. - 482 p.

[10] Rybina G.V. Intellectual systems: from A to Z. A series of monographs in three books. Book 1. Knowledge-based systems. Integrated expert systems. M.: Nauchtekhlitizdat, 2014. - 224 p.

[11] Rybina G.V., Blokhin Y.M. "Methods and software tools of intelligent planning for building integrated expert systems". Artificial intelligence and decision making. 2018. №1. p. 12-28.

[12] Dominique A. Lobar Approach to Breast Ultrasound. - M.: Springer International Publishing AG. Part of Springer Nature 2018. - 346 p.

[13] Jacobs O.E., Rozhkova N.I., Mazo M.L., Mikushin S.Y. "Experience using virtual breast sonography". Bulletin of Radiology and X-ray radiography. Radiation diagnostics. 2014. No 1. p. 23-33.

[14] Rozhkova N.I., Gorshkov V.A., Mesky E.V., Mazo M.L., Prokopenko S.P., Kireeva M.N. and others. Digital Mammology Clinic: Imaging Technologies. M.: SIMK; 2012.

[15] Rafael S. Gonzalez, Richard E. Woods. Digital image processing. M:" TECHNOSPHERE". 2012. - 1105 p.