COMPUTER-AIDED INTERPRETATION OF MEDICAL IMAGES: MAMMOGRAPHY CASE STUDY

*Institute of Radioelectronics, Nowowiejska 15/19, 00-665 Warszawa, Poland
**Institute of Biocybernetics and Biomedical Engineering PAS, Trojdena 4, 02-109 Warszawa, Poland

Abstract. This paper presents the current limitations and challenges of computer-aided interpretation of radiological examinations. The analysis and the proposed improvements of interpretation arose from our experience, knowledge and observations with the collected suggestions and conclusions. The emphasized topics are as follows: computer understanding of human determinants of diagnosis, characteristics and enhancement of observer performance, diagnostic accuracy measures of image examinations, computer-aided diagnosis (CAD) systems, and numerical description of medical image-based content. All of these diagnosis support concepts can be integrated into an intelligent diagnosis interface and enhanced, basing on a formal description of semantic image content, i.e. ontology implied as a reliable, dynamic platform of medical knowledge, useful for diagnosis. CAD for mammography and content-based image indexing supported by the ontology were integrated for the needs of enhanced diagnostic workstation applied in tele-information medical systems. A design of an effective human–machine interface has arisen as the leading problem of the current challenges.

Key words: radiological interpretation, ontology, diagnostic accuracy, computer-aided diagnosis, content-based indexing

1. Introduction

Undoubtedly, radiological examinations have become an essential source of diagnostic information [8, 28]. Clinicians routinely employ a variety of imaging modalities in the diagnostic practice. Therefore, a fundamental and vital aspect is the correct interpretation of the collected image data, referring to the physician’s knowledge of typical, healthy and pathological anatomy and physiology of the examined organs and structures, completed by experience and cognitive intuition [19]. The full process of radiological interpretation, i.e. understanding and assessment of medical image content, involves image-based detection of disease, defining disease extent, determining etiology of the disease process, assisting in the designing of clinical management plans for the patient based on imaging findings, and following the response to therapy. Yet the initial and a key constituent are true detection and the defining of the disease.

However, permanent development of imaging capabilities, almost unlimited number of imaged object variations and subjective human factors cause the lack of stable standards in diagnosis. Information overload caused by the increasing volume of diversified
medical image information, the difficulty in interpretation in the context of known limitations of the human observer, make this area the one in which computer–based image understanding or interpretation can make significant improvements in the detection and treatment of illness. Hence, computer–based aid is sought to be an essential remedy for challenges of revolutionary changes in the medical imaging world. Supported radiological interpretation is expected to be more reliable and objective, repetitive, time- and even cost effective.

Computer-aided radiology is useful and becomes real help because of the development of semantic image understanding concept, which means additional perspectives on the same image, taking into account the meaning and significance of its content. It is originated in the conceptions of semantic information theory\(^1\). Semantic content is the essence of image information. Thus, semantic content measure, assessment, description, extraction and enhancement are fundamental for the automatic interpretation of diagnostic image information. Computer–aided interpretation tools are “semantic eyes” approximating human possibilities of image understanding.

Generally, the purpose of computer–aided radiological interpretation is the improvement of the efficiency of interpretation process of large image data sets, decreasing diagnosis (medical) errors, improving the integration and time–effectiveness of radiological (medical) information exchange (communication), and the development of standardization and medical (diagnostic) interoperability. The ultimate goal is the enhanced quality and safety of patient care \([40]\).

A design of human–machine interface is the most important aspect of computer–aided interpretation of medical image exams. Assists include decision support, reminder and navigation techniques to help avoid diagnosis errors, content-based data mining capabilities, and access to reference libraries. Human–machine systems should take advantage of computer capabilities to increase physicians’ interpretation capabilities.

The original contribution of this paper includes indication and characteristics of the most important problems in computer–aided interpretations:

- reliable observer performance evaluation and exam accuracy measures,
- numerical description of diagnostic image content,
- improved semantic perception and reliability of CADs,
- ontology as formal description of medical semantic technologies,
- methodology of interpretation: rules and protocols of effective diagnosis,
- integration of digital medical systems and aiding tools.

Besides synthetic overview of the problem, the most important challenges, hopes and selected methods originated in our experiences, knowledge and observations with collected suggestions and conclusions were described. In details, we presented hypothesis of descriptor-based measure of diagnostic accuracy to make accuracy assessment more

\(^1\)http://plato.stanford.edu/entries/information-semantic/
reliable and put it closer to the algorithm optimization procedures. Moreover, the perception improvement of mammograms by wavelet–based processing, the methods of automatic detection of microcalcification clusters, the mammographic ontology interfaced to CAD and retrieval system in integrated interface of human–machine feedback for diagnosis were proposed.

Taking into account human conditioning of diagnosis, reliable modeling of observer performance is the key issue in the radiological performance improvement, where the first challenge is enhanced semantic perception of image content (section 2). Computer-based answer to human requirements is an automatic understanding of diagnostic image content, based on formal medical knowledge description (ontology) interfaced to numerical features, descriptors and other signal (texture and edge) characteristics (section 3). Ontology clarifies pathology descriptors, diagnostic protocols and interpretation consequences including important hierarchical content dependencies. Computer-aided diagnosis (CAD) systems use numerical descriptors of image content (lesions, pathologies, normal structures, general image features) to extract, detect and classify semantic content (confirmed subjectively with semantic (cognitive) understanding of visualized data). Set of semantic descriptors for image content characteristics were used to select similar pathologies for content-based data retrieval. Section 4 suggests using all of these methods as integrated interpretation support, i.e. realized interpretation interface for diagnostic workstation in tele–information environment of diagnosis, as a remedy to improve interpretation effectiveness with conclusions in section 5.

2. Understanding of human limitations

Medical image perception is a very complex activity involving interplay between vision and cognition, and requiring detection, classification and actionable decision tasks performed on highly variable objects (i.e. lesions). The remark that an experienced physician usually sees a visible lesion clearly but there are times when he does not, is still up-to-date. This is a baffling problem, apparently partly visual and partly psychological. It constitutes the still unexplained human equation in diagnostic procedures. Observer variation is due to many factors – the degree of individual training, experience, and interest, imaging equipment and the quality of technical work performed. It is also partly chargeable to some unknown factor that results in an occasional “blind spot” on the part of the best trained workers. This human equation can be mitigated in part by dual observation [1].

Errors (false positives, false negatives and misclassifications) are correlated with training, experience and environmental factors, as well as technical limitations, lack of cognitive information (i.e. clinical history, prior images, clearly formalized knowledge), and cognitive overload (i.e. lack of selective procedures and rules of diagnosis). Inter-and intra-observer variation is large and non-uniform.
Observer performance improvement, i.e. increased stability, perfection, repeatability and reliability, is the main purpose of computer assistance. It starts with clarification (standardization, normalization) in more objective description of normal and abnormal findings including variability of symptoms regardless of acquisition parameters and technical conditions. Better understanding of working issues such as presentation of information, impact of the environment and reader fatigue, and better understanding and control of image quality issues including the physics, psychophysics and diagnostic (semantic) measures lead to "intelligent", assistant interfaces, clearer rules of diagnosis, and generally improved methodology of diagnosis. Extraction of underlying image-based diagnostic information (CAD applications), integration and navigation (clear synthesis) of all accessible patient data and appreciation (modeling, pointing out) of the phenomenon of observer variation all should lead to improvement in the diagnosis of disease.

2.1. Improved semantic perception

Research in human perception of lesion symptoms requires objective methodologies for optimal image presentation, i.e. image diagnostic quality enhancement, semantic information measure, description and extraction, fitting of display parameters. Psycho-physical models for detection of abnormalities based on the understanding of what is desired by an image observer, what properties of radiological images are the most useful in their interpretation, and how these properties can be enhanced to improve the accuracy of interpretation, are sought.

However, visualization enhancing semantic (diagnostic) features is very challenging. The term visualization was derived from "visual data analysis" to emphasize analysis and interpretation. "Intelligent" visualization of data possibly communicates selected, extracted and empowered diagnostic image features (i.e. lesion symptoms) to the human visual system instead of large amounts of disordered information.

Exemplary application of improved semantic image perception is MammoViewer\(^2\) developed in our lab. Wavelet-based multiresolution decomposition of images into a set of subbands of different scales was used for effective mammogram data denoising and enhancement. We tested two decomposition methods (dyadic and undecimated) and many kernels (of over 40 orthogonal and biorthogonal filter banks) and thresholding or empowering procedures (e.g. fixed elimination of low frequency subbands, scale-dependent adaptive threshold estimation). The application modifies image coefficients in the multiresolution domain, which causes (after reconstruction) the conspicuous improvement of the usable signal perception (perception of microcalcifications and masses, especially spiculated ones).

Wavelet kernel that was used is Taswell 10/10 biorthogonal symmetric most-regular: low pass for analysis = \( [0.02691342,-0.03230335,-0.24110982,0.05410042,0.89950611, \ldots] \),

\(^2\)http://www.ire.pw.edu.pl/MammoViewer/
low pass for synthesis = [0.01984354,0.0238176,-0.02325784,0.14557075,0.54113273,0.54113273,0.14557075,-0.02325784,0.0238176,0.01984354].

The denoising step of the method is achieved by cutting high-pass subbands of the first transform level and removing the smallest transform coefficients in subbands 2 to 6 (according to Fig. 1a). Modified wavelet data are reconstructed in a process of image synthesis. Then in the processed image local contrast is enhanced by applying the adjusted curves in high-pass subbands of level 3 to 6 (according to Fig. 1b).

Fig. 1. Transformation grids and curves used for (a) denoising and (b) contrast enhancement of mammograms. The grids are visualization of kinds of processing in dyadic wavelet transform subbands: white boxes means removing of all the subbands – coefficients are set to zero, black subbands are not changed and gray ones means that the shown transformation curves are applied to them. Below the curves scaling parameters are shown, $D_n$ means coefficients in a processed subband.

Verification tests were performed on a set of 16 mammograms from DDSM\(^3\) (digitized at a pixel size of 43.5 microns and a 12-bit grayscale) containing pathologies - spiculated and circumscribed masses and with accompanying microcalcifications in two cases. Radiologist, expert in mammogram diagnosis, compared processed images with original ones and gave their opinion which was measured with subjective, comparative (relative) measures of quality by diagnostic symptoms analysis (tab. 1). The radiologist

\(^3\)http://marathon.csee.usf.edu/Mammography/Database.html
confirmed the effectiveness of the proposed methods. The results are shown in tab. 2. Exemplary pathologies – original and processed images are shown in Fig. 2.

<table>
<thead>
<tr>
<th>Mark scale</th>
<th>Wordy description of diagnostic image quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>+3</td>
<td>definitely (arbitrarily) better</td>
</tr>
<tr>
<td>+2</td>
<td>better</td>
</tr>
<tr>
<td>+1</td>
<td>slightly better</td>
</tr>
<tr>
<td>0</td>
<td>comparable with the original</td>
</tr>
<tr>
<td>-1</td>
<td>slightly worse</td>
</tr>
<tr>
<td>-2</td>
<td>worse</td>
</tr>
<tr>
<td>-3</td>
<td>definitely (arbitrarily) worse</td>
</tr>
</tbody>
</table>

Tab. 1. Subjective, comparative mark scale that was used in perception improvement tests. All important image features were scored consequently and conditions of true diagnosis were assessed.

<table>
<thead>
<tr>
<th>Overall mark</th>
<th>Spiculated masses</th>
<th>Circumscribed masses</th>
</tr>
</thead>
<tbody>
<tr>
<td>+2.5</td>
<td>+2.58</td>
<td>+2.25</td>
</tr>
</tbody>
</table>

Tab. 2. Mean mark of radiologist taking part in the tests.

2.2. Observer performance modeling and measure

Reliable measures of diagnostic accuracy of radiological examinations are still sought. Easy to use objective or numerical diagnostic quality factors designed to help define the level of ”safety” or ”enhancement” have the limited correlation with subjectively perceived diagnostic accuracy of images. Observer performance characteristics should be included in the exam accuracy measurement. Diagnostic image content and the process of image interpretation is a complex phenomenon where many technical, medical, human, dynamic, and even random factors contribute to the resulting decision. For example, in case of mammography, distinct conclusions are difficult to draw. One radiologist noted no masses on any of the films read, while another reported that 7.9% of the films contained a mass. Similarly, the proportion of calcifications and fibrocystic changes detected on the films ranged from 0% to 21.3% and from 1.6% to 27.8%, respectively. Moreover, radiologists in their 40s who completed medical school within the last 5 to 15 years were nearly four times more likely to have a higher false-positive rate than those in their 60s or 70s with more than 20 years since medical school graduation [21].

Furthermore, subjective methods of accuracy assessment do not provide any constructive methodology for performance improvement of the computer-based aid. It is hard to include such methods in the optimization process of observer performance and
Fig. 2. Examples of perception improvement of spiculated masses and microcalcifications. On the left – original images and on the right – processed ones.
image processing algorithms basing on more formal data description and modeling. Neverthe-
theless, quasi-objective measures based on statistical analysis of large enough samples of individual interpretation in terms of pathology detection, collected according to clinical practice, are preferred in many applications. Widely accepted methods of measuring image diagnostic accuracy used ROC (receiver operating characteristic) analysis, which has its origins in theory of signal detection, or its modifications LROC and FROC [2, 4]. However, ROC analysis has no natural extension to the evaluation of measurement accuracy in processed medical images [5]. It artificially simplifies ordinary practice (artificial scales, binary decisions, simplified lesion descriptions etc.), relies on suspect Gaussian assumptions and is often too complex and costly to be effectively fitted into the dynamic factors contributing to the diagnostic decision. Statistical nature of ROC–based results refers to a large set of diagnostic decisions grounded in perception (so subjective). But the nature of an individual decision is not statistical, only temporary, dynamic and case-dependent. An important conclusion of the largest data–gathering experiment of diagnostic accuracy of mammograms estimation (still only a pilot study in statistical meaning) is that great care must be devoted to any statistical analysis of diversified radiologists opinions [11].

Standard meaning of diagnostic accuracy is that it refers to the decision made by a radiologist in terms of pathology detection and classification. Single image information is interpreted in appropriate terms of diagnosis on the grounds of a radiologist’s knowledge and experience, and combined with the other available information (case context). Thus, assessment of a single, concrete image in diagnostic and possibly objective terms common and important for every abnormality case is particularly important because it reflects radiological practice and is useful for compression effects acceptance.

Hypothesis of descriptor-based measure of diagnostic accuracy. Our suggestion is that pathology detection test is a somewhat coarse method. Another, even more sensitive and rigorous way is the observation and assessment of selected image features, which are important for diagnosis of a given case. Such assessment can be useful for design of reliable numerical measures of diagnostic accuracy [32]. It was confirmed experimentally that subjective rating of diagnostic symptoms occurred more useful, reliable and consequently objective method of diagnostic quality measurement for acceptable image compression range estimation [26].

Reliable evaluation of a single image in diagnostic terms is really important because it reflects radiological practice. Lexicon (first of all BI-RADS) or ontology-based textual descriptors are useful in subjective rating of image accuracy. However, lower-level local features which influence detection decision (i.e. image interpretation for the purpose of diagnosis) are sought. Such features are as follows: edge gradient, smoothness, blurring, contrast, shape, outline and sharpness of certain particular details, texture clearness, relation of inside and outside textures of selected structures etc. They play an essential role in detection and classification of any lesion. Thus, images with lesions could be
evaluated in diagnostic accuracy terms by rating quality (state of visibility, perception) of these “diagnostic features”.

Accuracy measure should signal a reduction of processed image accuracy. The procedure of evaluation concerns, referring to difficulties with observing of the lesion symptoms, all initially pointed out abnormalities, and rating specified image local features (diagnostically important) on a multilevel scale. Perception of those elements is conclusive in a final decision of radiologists relating to the lesion detection and classification. The character of evaluation is more qualitative - the improved quality of selected local image features affecting diagnostic accuracy is decisive. Analysis of lesion features is done on a lower level, categorization and description of symptoms are more formal, detailed, and therefore, constrains more objective scores.

Moreover, diagnostic image features are useful in design of numerical lesion descriptors for automatic detection, following subjectively verified accuracy of the images. To make some kind of interface, observers are asked to describe breast lesions or abnormalities in terms acceptable to both, medical and technical (image processing) communities, for example: contrast differentiation (related to tissue density), interpretation clarity (visibility of the disorder, noticeability of indicated abnormalities, mostly related to detection ability), shape and margin (outline, contour distinction) of chosen structures, including lesions and other abnormalities. Observer rates perception of these image diagnostic features in indicated, automatically or manually, ROIs on a scale of weak, indistinct, scarcely perceptible, distorted features to distinct, clearly perceptible, regular, beyond a doubt signs.

“Cognitive resonance” between textual descriptors of medical knowledge and numerical descriptors are fundamentals of human–machine interface.

3. Automatic understanding of diagnostic content

The important aspect of computer–based aid development and technological progress is the understanding of the relevant semantic contents of the radiological examinations on the basis of numerical features extracted from the image data [30]. The subject of interests is the application of the cognitive-based approach for intelligent semantic analysis, allowing the description of diagnostic image content automatically. The most important part of this analysis depends on the “cognitive resonance” process, in which the features of real images are compared with some kind of expectation taken from the knowledge base, containing the characteristics of the pathological cases originated from medical practice. The importance of cognitive resonance in medical image understanding was noticed and confirmed by the research of R. Tadeusiewicz and M. Ogiela [31].

High-level analysis, description and recognition includes the use of artificially intelligent techniques, functional analysis of data, approximation theory methods and human visual models for image interpretation and human-following understanding. The rela-
tionship between image components, objects and patterns, its context - that play an 
important role in diagnosis is related to and completed with a priori knowledge gained 
from a range of sources. Some important approaches were based on a special kind of 
image description language and grammar formalism. During the linguistic analysis of 
medical patterns, one can solve the problem of generalization of features of a selected 
image and obtaining semantic content description of the image [38].

A different, signal processing based methods of computational understanding of med-
ical images was proposed in CAD systems.

3.1. Computer-Aided Detection and Diagnosis

Leveraging CAD is critical to utilizing all the information available, and while validation 
is slowly occurring for more diseases, integration of CAD into the practice is essential. 
Application of CADs is intended to maximize the amount of useful diagnostic information 
being reliably extracted from radiological examinations and moreover, to select and 
interpret image data by true positive or negative indications. The combination of a 
radiologist’s interpretation and target–based CAD was superior to either radiologist 
alone and can be treated as alternative to second human opinion.

CAD methods are typically of two types: target–based such as object detection, char-
acterization and recognition (automatic understanding and diagnosis), and non-target-
based including change detection methods and atlas–based morphometric approaches 
focused on screening studies and quantitative morphometric methods for new types of 
diagnoses). But despite significant demands existing in many clinical areas, world–wide 
research and long–term clinical trials, only two clinical problems have FDA–approved 
CAD systems – mammography and lung lesion detection. Impediments to more rapid 
advancement include limited, which means expensive, access to image databases with 
large numbers of verified cases for training and testing of algorithms, the lack of a gold 
standard to measure performance objectively, compare algorithms with constructive con-
cclusions, the poor technological integration of CAD tools into the clinical workflow, i.e. 
with picture archiving and communication systems (PACS), tele–radiology systems and 
radiology information systems (RIS).

Even though typically CAD systems do only a small part of the work required when 
interpreting medical imaging examination, they play an important and irreplaceable role 
in increasing of the efficiency and completeness of radiological interpretation process. 
Development of computer–based systems for automated detection and diagnosis is an 
important goal of researches in digital mammography. One of the important aspects 
in this technological progress is the understanding of the relevant, semantic contents of 
medical images on the basis of technical features.

CADs for mammography are precursors of successfully applied computer aid. Com-
mercial CAD systems in the area of microcalcification detection achieve sensitivity up 
to 98% (R2 ImageChecker, Second Look) with a mean number of false positives per
Microcalcifications can be defined as follows: individual microcalcifications appear as tiny spots slightly brighter (with low local contrast) than a variable background of their surrounding tissue. Calcifications associated with a malignancy are usually from 0.05 to 1 mm in size and variable in shape—curvilinear or branched. Benign calcifications are larger and smoother. Only clusters of microcalcifications including at least 3 to 5 particles within an area of 1 cm$^2$ are regarded as clinically suspicious [16]. The probability of cancer increases as the number of calcifications in a cluster grows. There are many approaches for localization and segmentation of microcalcifications, i.e.:

- localization of regions with microcalcifications using statistical texture feature space (e.g., surrounding region-dependent or gray level run length), fractal dimensions for estimating region roughness etc;
- localization of individual microcalcification objects, e.g., Laplacian of Gaussian filtering, wavelet-based decompositions;
- segmentation based on texture statistics, mathematical morphology, local thresholding and region growing etc.

Clustering algorithms are commonly used in the simplest way (a window 1 cm$^2$ passing the entire image with criterion 5 or 3 particles within the window). Our study aims at making progress in the development of a computer-based system for the automatic detection of microcalcification clusters and segmentation of individual objects on mammograms [37]. Optimized detection and classification methods were validated in clinical tests using diagnosed databases (Digital Database for Screening Mammography from the University of South Florida and exams collected from 2 hospitals in Warsaw, Poland). Microcalcification detection technique which is developed by us is a combination of a) wavelet-based methods and convolutions with Laplacian filters of different scales for localization of bright spots [17] (potential microcalcifications), b) optimized DBSCAN (Density Based Spatial Clustering of Applications with Noise) [7] for grouping signals and thus verifying detection, and c) finally morphological and region growing methods for shape reconstruction (see Fig. 3). Sensitivity of approx. 90% was initially achieved with a mean number of 1.6 false positive clusters per image (within tests localization threshold was selected manually).

3.2. Content-based indexing of medical images

Databases can be useful in finding relevant prior examinations quickly, and in finding the key images within that exam set as precisely as possible. Content-based retrieval will likely become more commonly used for medical information retrieval. Content-based visual information retrieval (CBVIR) or content-based image retrieval (CBIR) has been one of the most vivid research areas in the field of computer vision over the last years. The availability of large and continuously growing amounts of visual and multimedia data, and the development of the Internet point the need to create content-based access.
methods that offer more than simple text-based queries.

It needs to be stated that the purely visual image queries as they are executed in the computer vision domain will not most likely be able to ever replace text-based methods as there will always be queries for all images of a certain patient, but they have the potential to be a very good complement to text-based search based on their characteristics. Still, the problems and advantages of the technology have to be stressed to obtain acceptance and the use of visual and text-based access methods up to their full potential. A scenario for hybrid, textual and visual queries is proposed in the CBIR2 system [34, 23].

The benefits from CBIR engine for PACS user seem to be incontrovertible. Even if CBIR is relatively simple and able to distinct only modalities and some body parts - it might be a good supplement to classic text query engine, giving an opportunity to verify misclassified DICOM-based information, which happens in clinical practice. Nowadays, the CBIR improvement is based on semantic understanding of technologies in the context of semantic-based knowledge management as promising development factor of IT world [39]. In a heterogeneous world of ubiquitous information flow, they allow a flexible and seamless integration of applications and data sources (i.e. radiological chain

of patient, imaging system, interpretation, and therapy). They provide an intelligent access, understand context and content, give answers and generate knowledge including as objective as possible object description (e.g. lesion, structure relation, pathology features).

The inclusion of visual and semantic features into medical studies is an interesting point for several medical research domains. Content–based features do not only allow the retrieval of cases with patients having similar diagnoses, but also cases with visual and semantic similarity but with different diagnoses. In teaching, it can help lecturers as well as students to browse educational image repositories and visually inspect the results found. Image interpretation is expected to be improved through the use of content-based access methods to the existing large repositories of diagnoses cases.

Reference database of image exams indexed by content is an exciting challenge for semantic analysis of image data. It provides the examples of clinically verified cases of pathologies. The retrieval of diagnostically similar exams over as wide as possible distributed databases of reliably labeled cases was used to support diagnosis of cases difficult to assess. Image retrieval techniques are based on similarity matching of image features describing characteristic symptoms of pathologies. To search the necessary image exams, we have to index the images according to their diagnostic content and find best matches for a query image according to a given similarity measure. Between interesting algorithms, those based on multiresolution histogram matching, multiscale data distributions and local structure correlations, subband texture analysis are worth mentioning [25].

Our research was directed to optimization of multiscale extraction of semantic image features basing on quality and similarity measures and ontology–based particular description on lesions and context information [42]. Directional subbands, curvelets (contourlets, beamlets, wedgelets, 2D kernels of wavelets, etc.) are tested to increase packing of diagnostically important information and selectivity of image characteristics. The patterns of abnormality signatures were sought in hierarchical redundant tree of semantic points. Mean value of retrieval precision of modality differentiation approx. 80% was achieved in initial experiments with CT, MR, radiograms and mammography exams (especially 75% for mammograms).

Moreover, the proposed CBIR system used ontology-based semantic descriptors of abnormalities and CAD-based, i.e. MammoViewer-based, numerical descriptors correlated to the pathology symptoms. The precision of image retrieval depends on CAD sensitivity (approx. 90% of cluster detection), lesion characteristics, precision of breast tissue differentiation, the estimates of local and global features of the image, and textual and numerical descriptors based on ontology (e.g. texture, shape, outline, localization).

Exemplary interfaces of indexed database are presented in Fig. 4. The content-based retrieval engine is part of larger telemedicine system (Fig. 5), ImageShark (IShark), which offers many interesting features like:
IShark consists of several modules that provide far more efficient way of interaction with image database than classic text-based queries from PACS system. The important thing is that the system does not take out any search functionality known from classic systems. The new features concentrate mainly on distributed search with the possibility of sharing interesting cases (anonymized and encrypted) between medical centers, content-based search and effective, interactive JPEG2000-based image communication.

3.3.Ontology as reliable medical knowledge platform

Nowadays, roles played by ontologies are multiple. Primary goal of ontology is to effectively represent domain knowledge, adequately and exhaustively define relevant concepts, object characteristics and relationships between them, to provide a common, standardized vocabulary comprehensible by humans and machines by which users and computer systems can communicate. Thus, ontology means systematization, objectification and verification, knowledge base of the model populated with concept’s instances constitutes standard diagnostic knowledge database. Ontologies are the foundation of the Semantic Web, where integration and interoperability of heterogeneous sources of information is needed [33], [29]. Ontologies also form the basis foundation of evidence-based-medicine and standardization efforts [27].

Standardized terminological system in mammography already exists, it is BI-RADS [9] (Breast Imaging Reporting and Database System). But BI-RADS has not been univocally accepted, and what is more important, there is no clear evidence that its use improves the completeness or accuracy in mammogram descriptions [6, 13] Moreover, there are indications of its limited expressiveness for encoding mammography findings in databases [10]. Other shortcomings of BI-RADS are: lack of hierarchical terms organization, simplified, lesions descriptions and inconsistent diagnostic criteria. More exhaustive, consistent and formal system for lesion description classification and interpretation is necessary. Ontology is the best candidate to face these problems, and help build useful and necessary computer-based aid in mammography.

For that purpose, mammographic ontology has been constructed [36]. It has three main goals. The first: to provide standard vocabulary and formal, exhaustive definitions of concepts for description and interpretation of mammograms. The second: modeling of mammography report. The third: to use ontology as specification for designing the database of mammography reports and graphical editor for pathologies description. Other potential uses of mammographic ontology are: educational tasks, as an assistant tool for diagnosis in mammography, and content–based indexing of mammograms database. Knowledge of ontology construction has been extracted from three sources: corpus of routine, free–text mammography reports (close to 400), long–term interviews...
Fig. 4. Interfaces for indexing of medical image database.

and consultations with radiologists at local hospital and careful wide range analysis of medical literature [14, 15, 3, 24]. During the phase of knowledge acquisition manual methods have been used.

Despite of its known draw–backs BI–RADS is a good starting point in construction of mammographic vocabulary. It contains rudimentary lexicon for allowing the description of basic visual features of masses and calcifications. It has to be stated, that there is a discrepancy between terminology found in Polish radiology reports and BI–RADS. There are terms in BI–RADS that were never used by Polish radiologists, and terms often mentioned in the reports but not included in BI–RADS. Mammographic ontology incorporates more detailed and more diagnosis–oriented, definitions of such concepts as: breast composition, calcifications, architectural distortion and axillary lymph nodes. In authors opinion, BI–RADS system has to be extended to a national vocabulary, although we are fully aware of the fact that reaching a consensus within mammographers community represents a great challenge.

After gathering a relatively complete vocabulary of mammographic terms, an initial
set of concepts, their properties and relationships between them have been identified. Concepts have been structured into subsumption hierarchy, properties of concepts have been modeled and described in natural language. This informal model has been formalized using ontology editor Protégé-2000 version 3.1, allowing export of the ontology to machine readable formats (RDF, XML). Ontology has been divided into small, internally coherent components-modules. A goal of modular design is to achieve explicitness of the ontology and to support knowledge reuse and maintainability.

3.3.1. Overview of the model

Basic general concepts that allow description and interpretation of mammograms are: Breast Composition, Breast Lesions and Lymph Nodes. Each of those concepts is a root-concept in corresponding module in the model. The module Features of Mammographic Lesions and Observations contains features definitions of all concepts in mammographic ontology and form its crucial part. It is divided into four sub-modules grouping separate levels of breast lesions and their features. The first sub-module contains visual features of mammographic objects as seen on X-ray film, second and third levels- correlations between lesions feature and facts from patients clinical data. The fourth sub-module captures abstracts, non-visual attributes of mammographic pathologies like lesion diagnosis and interpretation. Definitions of all concepts in the model and their instances form Knowledge Base of the ontology. Part of the Knowledge Base containing instances of classes from module Breast Lesions deserves special attention, because it is a described set of images of different mammographic pathologies. Basic ideas are presented in Fig. 6.

At the moment, mammographic ontology contains 141 concepts-classes, arranged on 7 hierarchy levels (BREAST LESIONS is the most nested module) and is permanently developed. To enable feature description of ontology concepts, 145 slots were defined and total count of objects in the model was 687: classes, slots and instances of classes. Assumed and target advantages of our system are: reliable and sound knowledge representation based on diversified acquisition sources, use of frame-based knowledge representation which is cognitively simple, intuitive and easy to understand for radiologists, domain knowledge sharing between applications such as: database and editor for mammography reports, CAD and system for indexing mammogram reference database.

To sum up, the ontology exists in a tree format, structuring key mammographic features of breast disease, linking directly both with the methods a radiologist would use to identify the abnormality as well as how a CAD method might also identify, measure, classify and understand the abnormality. The ontology further enables the inference of breast disease according to the visible features of the abnormality as it appears on a mammogram.
Ontology for mammography, in essence, is a computer representation of how breast diseases may be described.
3.3.2. Support of diagnosis

The ontology, initially, may be used as a common platform for the comparison of a radiologist’s and a CAD’s effectiveness at identifying breast pathologies on a mammogram. In addition, the ontology may be used as a foundation of content–based data exchange between humans and computers to make human knowledge more objective and computers more human–like (i.e. understanding the semantic information). Other potential uses of mammographic ontology are content-based description and indexing of mammogram reference database (e.g. IShark), educational tasks, modeling of exam reports and GUI for description of pathologies.

Mammographic ontology implied as reliable, dynamic platform of medical knowledge useful for the observer performance improvement is helpful for the diagnosis. Reliable numerical description of medical image features constitutes the base of an integrated environment for diagnosis realized as integrated intelligent diagnosis interface for diagnostic workstation in tele–information medical systems (Fig.5). Particular results are a set of common numerical diagnostic features for indexing, ontology-CAD feedback of automatic pathology indications and descriptions, and integrated, diagnosis supporting interface for mammogram description [43]. Ontology was used for more objective modeling of lesion signatures in multiscale domain, the extraction of semantic image features basing on diagnostic quality and similarity measures.

The idea was to design a visual ontology, whose concepts are described by numerical descriptors and methods to compute their values and associate them with concepts in mammographic ontology. This is the way to have two complementary descriptions of objects in an image, the first one corresponding to numerical description, and the second one to a semantic, human-like vision of the image content. However, the idea is difficult because of lack of reliable numerical measures related to diagnostic content. Moreover, several numerical representations might correspond to one semantic description (e.g. "round shape" can have more than one numerical representation). The visual ontology under development was based on the extension of MPEG-7 visual descriptors [20]. Combining numerical and semantic description of image content is a promising perspective to make up for the lack of semantic information in CAD systems (Fig. 7).

Formal characteristics of abnormalities were used as the expression of standardized medical knowledge and observer interpretation preferences, which is potentially the best model of human beings (radiologists) which interfaced to machine agents (CAD systems with better description and modeling of image semantics) creates computational intelligence able to enhance diagnosis.

4. Integrated interpretation support

Radiology is extremely susceptible to computer-based integrated support in decision–making. Integration of databases, decision–aiding tools, systems, networks can help to
Fig. 7. The idea of associating two descriptions of image content. Semantic, human description represented by domain mammographic ontology and numerical descriptors of the content.

determine what information is needed that a user does not have. Information such as the reason for the imaging examination being ordered by the referring clinician, laboratory data, and patient clinical history, in addition to cross-specialty and cross-modality imaging should be integrated and easily accessed. Decision support tools will provide more information to end-users, but need to be more fully integrated into the PACS database.

The purpose of designed computer-assisted medical systems is the completeness of supplied aid to improve the diagnosis. However, visualization and navigation of all accessible medical information requires efficient selection of all the information necessary to make effective clinical decisions without distracting the user, followed by informa-
tion synthesis. Automation by integration can improve database information quality, as well as facilitate improved user interfaces, computer-aided tools, and preemptive detection of errors before they propagate. The importance of provided information quality, CAD-based support and tele-radiology, flexible adjustment to various needs, and easily accessible examples of pathology cases, quick and precise retrieval, are crucial.

Advances in integration may require new open standards, continued evolution of the National Electrical Manufacturers Association (NEMA), Digital Imaging and Communications in Medicine (DICOM) standards, and increased adoption of the framework Integrating the Healthcare Enterprise (IHE) in imaging systems. Greater acceptance of IHE concepts means the integration of HIS and RIS with PACS. Real-time CADs at the PACS display and tele-consultation desktop with Web-based reference systems (e.g. content-based indexed image databases) will be useful through better integration of information richness.

End-users want more functionality but simpler user interfaces. Intelligent user interfaces (IUI) should be designed with the following features and demands in mind: efficient and reliable visualization of all useful information, adoption to human performance and limitations, intuitiveness and consistency, flexibility and easy configurability (to accommodate different types of users and different types of imaging examinations), appeal and idiot-proof. IUI is based on effective models and metrics for function task lists to read different case types, workflow guidelines, common language and best practice for hanging protocols. Diagnostic IUI including important information sources and interpretation support was presented in Fig. 8.

For any medical condition, there would be huge gains if one had a pan-national database – so long as that (federated) database appears to the user as if it were installed in a single site. Such a geographically distributed (pan-European) database can be implemented using the so-called Grid technology. Interesting initiative of integration was done for mammography. MammoGrid\(^4\) project was developed towards providing a collaborative Grid database analysis platform, in which statistically significant sets of mammograms can be shared between clinicians across Europe. Mammogram standardization system was applied to enable the comparison of mammograms in terms of tissue properties independently of scanner settings, and to explore its place in the context of medical image formats. MammoGrid will provide a channel to make software solutions available to a large number of European screening centers to perform quality control and set a common ground for epidemiological studies. Access to image databases with large numbers of verified cases for training and testing of CAD algorithms is a priceless benefit.

To sum up, the integration is essential to improved diagnostic data evaluation, decision making and finally high-quality patient care. However, data gathering, flow and exchange standardization, IUI design and communication of the results is a great chal-

\(^4\)http://mammogrid.vitamib.com/
Fig. 8. Diagnostic user interfaces designed for enhancement of mammogram interpretation.
lenge for life science community. An integrated tele–information system for clinical and research purposes was developed in collaboration between Warsaw University of Technology, Wolski Hospital and software houses\textsuperscript{5}. The basic structure of the system is presented in Fig. 9.

![Diagram of the tele–information system](image)

**Fig. 9.** Integration of computer–aided possibilities: Tele–information Radiology System as networked environment to support the interpretation of medical images.

5. Conclusions

Improvement of radiological interpretation process includes image diagnostic accuracy verification and optimization, image data interpretation (ways of assessment), communication of imaging results (reporting), a reduction in medical errors (observer performance improvement), integration of all services and information sources, workflow and efficiency within the health care enterprise, and ultimately, the overall quality of care.

Generally, computer-assisted interpretation of image exams is a complex problem of physician performance support, integrating different methods and forms of data processing, visualization, analysis, classification, understanding and navigation. Summarized concept of proposed diagnostic image interpretation was presented in Fig. 10.

Computer-based interpretation of medical images proposed in our research was based on the following concepts:

\textsuperscript{5}http://telemedycyna.evernet.com.pl/system/

Fig. 10. Complex computer-based aiding of radiological interpretation process. Acronyms are as follows: RIS – radiology information system, HIS – hospital information system, PACS – picture archiving and communication system, HVS – human visual system.
• extraction and numerical description of image diagnostic information in terms of semantic local and global features, descriptors and diagnostic quality indicators; e.g. wavelet-based and other multi-scale sensing technologies may be used for representation, separation and recognition of the signal useful-in-diagnosis;
• constituting objective, as formal as possible medical knowledge platform which is gold standard reference of meaning and significance;
• realization of dynamic, easy-to-access source of up-to-date information based on knowledge-by-example concepts; content-based image indexing and retrieval is really useful;
• multiple stages of the "cognitive resonance" process which means optimization of recognition, classification and the similarity identification according to the semantic criteria; the essence is the modeling of numerical image description to be fitted as reliably as possible to the standardized, objectivated, formalized and up-to-date diagnostic knowledge.

Advanced techniques of artificial intelligence were applied in order to recognize and understand semantic content of the images to help its interpretation. Computational intelligence (NN, fuzzy sets used as classifiers and decision support for CAD and ontology) can improve the intellectual behavior of machines by better description, extraction and cognitive-based identification of semantic information. Making semantic data interpretation more amenable to computational methods is the main subject of the presented research. Formal description of abnormalities, interpretation criteria and protocols, the relationships (properties) between these concepts, reasoned to computational methods of determining of diagnostic accuracy indicators (recognition and understanding in diagnostically important terms), as selective, representative and compact as possible, are well readable by both human beings (radiologists) and machine agents (CAD systems). Computational intelligence can be improved by more effective mathematical, numeric modeling of semantic information according to the following formulas: -selection of computational information (compaction with redundancy reduction), -reliable determining of semantic contexts, -adaptive, cognitive recognition of medical knowledge (patterns of abnormalities), -verification by image example (indexing, retrieval) of interpreted information and concluded computational intelligence to be able to support diagnosis.

Even though a design of human–machine interface is the most challenging aspect of computer–aided interpretation of medical image exams, both the technology improvements and the increased sensitivity and specificity of human performance in content–based information use are the fundamental conditions for expected success.

References
1961
1975

1986


1990


1994


1996


1996


1997


1997


1997


1997


1997


1998


1998


1998


1999


1999


2000


2000


2001


2002


2002


2002


2004


Pisanelli D.M., Gangemi A., If Ontology is the Solution, what is the Problem, Ontologies in Medicine vol. 102 in Studies in Health Technology and Informatics, IOS Press, 2004


2005


2006


2007
