1 Administrative data

Project Title

Bayesian Decision Support in Medical Screening

Project Acronym

B-SCREEN

Principle Investigator

Applicant:	Dr. Peter Lucas, MD
Institute:	Institute for Computing and Information Sciences University of Nijmegen
Address:	Toernooiveld 1, 6525 ED Nijmegen, The Netherlands
Telephone:	+ 31 24 3652611
Telefax:	+ 31 24 3653366
Email:	peterl@cs.ru.nl

Applicant:	Dr. Ir. Nico Karssemeijer
Institute:	Department of Radiology, UMC St. Radboud
Address:	Postbus 9101, 6500 HB Nijmegen, The Netherlands
Telephone:	+ 31 24 3614548
Telefax:	+ 31 24 3540866
Email:	n.karssemeijer@rad.umcn.nl

2 Summaries

English Summary

It is expected that the availability of large data sets will lead to important changes in health care as these data can be exploited for the construction of decision support systems which may change the quality of patient care. Probabilistic graphical models, in particular Bayesian networks, are considered appropriate tools for mining medical data. However, learning Bayesian networks (automatically) from data is time consuming and the quality of resulting structures is debatable. Research has shown that simple naive Bayesian classifiers often outperform sophisticated Bayesian networks for classification purposes. The latter have been shown to improve, though, by adding additional dependences between variables, if medical domain knowledge can be modelled properly. This project intends to develop new improved classifiers using Bayesian networks based on advanced image analysis and domain knowledge from the breast cancer screening domain to be used in decision support systems for radiologists. Recent breast cancer research has revealed that the reading of mammograms by radiologists is the weakest link in breast cancer screening. From audits it is known that in the Netherlands more than 25% of all cancers detected in the screened population show relatively clear signs of abnormality in previous screening mammograms. Methods for computer-aided detection (CAD) have been developed to support radiologists, but are usually intended to be used to avoid perception errors instead of addressing the problem of interpretation failure by radiologists. There is strong evidence that interpretation failure is a more common cause of missing cancers in screening than perceptual

oversight. This project has as goal to develop new CAD methods to support the interpretation process of mammograms by radiologists. Soon the digitization will start of the Dutch breast cancer screening program, which will make use of broadband technology to archive images in a national archive. This will lead to a very large database of known abnormalities. With several thousands of new breast cancer cases each year this provides a unique opportunity to develop decision support systems for radiologists. Due to screening breast cancer mortality in the Netherlands has decreased by 800 cases per year. By developing more effective screening methods cancer mortality can be reduced even further.

Dutch Summary

Verwacht wordt dat de opkomst van enorme datasets zal leiden tot grote veranderingen in de gezondheidszorg aangezien deze gebruikt kunnen worden voor het ontwikkelen van betere beslissingsondersteunende technieken. Probabilistische grafische modellen, in het bijzonder Bayesiaanse netwerken, worden door menigeen gezien als een geschikte techniek om onzekere medische kennis te representeren. Daar tegenover staat dat het (automatisch) leren van Bayesiaanse netwerken tijdrovend is en dat de kwaliteit van de gevonden netwerkstructuren betwistbaar is. Onderzoek heeft aangetoond dat zeer eenvoudige netwerken vaak beter zijn dan complexere netwerken, hoewel ze nog wel verbeterd kunnen worden d.m.v. het toepassen van domeinkennis. Dit project heeft als doel om classificatiemethoden te ontwikkelen die leiden tot beslissingsondersteunende systemen voor radiologen in de context van borstkanker screening. In het bevolkingsonderzoek naar borstkanker blijkt de beoordeling van mammogrammen door radiologen de zwakste schakel te zijn. Uit gegevens van visitaties is bekend dat van de carcinomen die gevonden worden in de gescreende bevolkingsgroep in ruim 25% van de gevallen al een duidelijke afwijking zichtbaar was in voorgaande screeniningsmammogrammen. Methoden voor ‘computer-aided detection’ (CAD) zijn ontwikkeld om radiologen te ondersteunen, maar richten zich voornamelijk op het voorkomen van perceptiefouten, i.p.v. de interpretatie van gedetecteerde afwijkingen. Uit onderzoek is gebleken dat aan interpretatiefouten het grootste deel van de gemiste carcinomen kan worden toegeschreven. Dit project heeft als doel om nieuwe CAD methoden te ontwikkelen om de interpretatie van mammogrammen te ondersteunen. Op korte termijn zal begonnen worden met de digitalisatie van de Nederlandse borst kanker screening, waarbij gebruik zal worden gemaakt van breedbandtechnologie voor het opzetten van een landelijk beeldarchief. Dit zal resulteren in een enorme database waarin jaarlijks enkele duizenden nieuwe borstkanker gevallen worden opgenomen. Hierdoor onstaan interessante mogelijkheden voor het ontwikkelen van beslissingsmethoden voor radiologen in het screenings proces. Op basis van recente gegevens is geschat dat in Nederland de jaarlijkse sterfte aan borstkanker door screening is gedaald met 800 gevallen. Met behulp van effectievere screenings methoden is het mogelijk de sterfte door borstkanker nog verder te reduceren.

3 Classification

Computer Science, subdiscipline: Interaction - image processing,
NOAG-i: Intelligent Systems (5.6)
Bayesian network modelling, Radiology

4 Composition of the Research Team

Institute for Computer and Information Science, Radboud University Nijmegen (RU) **Toernooiveld 1, 6525 ED Nijmegen, The Netherlands**		
Name	**Discipline**	**Paid by**
Dr. P.J.F. Lucas, MD	Computing Science (AI, Medical Informatics), Medicine	RU
Dr. P.C. Groot	Computing Science (AI, Mathematics)	NWO Hefboom
Prof.dr ir Th.P. van der Weide	Computing Science (IR, Information Systems)	RU
postdoc – requested	Computing Science (AI)	NWO

Department of Radiology, UMC St. Radboud **Postbus 9101, 6500 HB Nijmegen, The Netherlands**		
Name	**Discipline**	**Paid by**
Dr. Ir. N. Karssemeijer	Medical Physics	UMC
Dr. C. Boetes	Radiology	UMC
postdoc – requested	Computing Science	NWO
Scientific programmer – requested	Computing Science	NWO

National Expert and Training Center for Breast Cancer Screening (LRCB) **UMC St. Radboud, Postbus 9101, 6500 HB Nijmegen, The Netherlands**		
Name	**Discipline**	**Paid by**
Prof.Dr A. den Heeten	Radiology	LRCB

5 Research School

Research School for Information and Knowledge Systems (SIKS).

6 Description of the Proposed Research

6.1 Problem Description

It is expected that one of the most important changes of health care by the introduction of information technology is the future availability of huge sets of data. This will allow exploiting these data in the construction of decision-support systems, which may play a role in improving the quality of patient care. Many consider probabilistic graphical models, in particular Bayesian Networks (BNs), appropriate and also intuitive tools for incorporating the qualitative (cause-effect) and quantitative (uncertainty) information amply available in medicine as many of the crucial elements of data mining are embedded in important and subtle statistical concepts [34].

Much emphasis has already been put on building BNs based on (medical) expert knowledge [3, 37, 36, 61, 63], however building such networks is very time consuming. With the availability of huge data sets automatic learning of BNs becomes interesting [20, 55, 82, 17]. One particular area which has already received a lot of interest in the research community is the use of BNs for classification (see for example a recent special issue on probabilistic graphical models for classification [57]).

One of the simplest and yet most consistently well-performing set of classifiers are the naive Bayesian classifiers [24]. In the naive Bayesian classifier all attributes are assumed to be conditionally independent given the class variable, an assumption which is clearly violated in many real world domains. This has inspired many researchers to investigate extensions of the model. However, although crude as BNs, naive Bayesian classifiers have been shown to often outperform more sophisticated network structures as well as other types of classifiers [22, 78, 56, 29]. One extension, however, tree-augmented Bayesian Networks (TANs) [29], which incorporate extra dependences among features in the form of a tree structure, have been shown to outperform naive Bayesian classifiers [29]. It is also confirmed in [59] that using expert causal domain knowledge to add dependences to the model may improve the Bayesian classifier performance. Additional benefits of using (medical) knowledge in learning BNs are that restrictions can be put on the topology of the network such that it can be learned in less time. Without restrictions, the BN topology space is super-exponentially large. Finally, [59] shows that adding dependence information may make the classifier more faithful w.r.t. the data, which is still a weak aspect of current classifiers. Faithfulness w.r.t. the data is often not considered as one usually only focusses on the classification performance which does not necessarily improve with a more faithful representation. However, if a BN is faithful w.r.t. the data it can also be used to compute a marginal probability distribution, which may be useful in supporting the medical decision making process. In practice, a trade-off has to be made between expressiveness and ease of learning of the BN structure and parameters.

In 2006 digitization of the Dutch breast cancer screening will start. It has been decided that all screening mammograms will be stored in one national archive, which will be facilitated by the use of broadband technology. As a consequence, an enormous database of normal and cancer cases will become accessible in only a few years (over 3000 cancer are detected yearly in screening). This provides a unique opportunity for the development of decision-support in this domain. In breast cancer screening, radiologists select a small percentage of women for referral, based on suspicious abnormalities in their mammogram. To maintain high specificity, radiologists do not refer all abnormalities they see. Consequently, not all cancers initially detected are acted upon because of *interpretation failure*. There is strong evidence that interpretation failure is a more common cause of missing cancers in screening than perceptual oversights [90, 33, 45, 7, 94, 10]. From audits it is known that in the Netherlands more than 25% of all cancers detected in the screened population show relatively clear signs of abnormality on previous screening mammograms, while another 25% show minimal signs.

Computer aided detection (CAD) methods have been introduced as a way to avoid perception errors and has already gained wide acceptance in screening. However, current CAD technology does not address the problem of interpretation failure. *The* **aim** *of this project is therefore to use BNs or classifiers to further develop CAD technology to address the problem of interpretation failures by radiologists.* More specifically, the **objectives** are:

- Development of novel classification methods using BNs or Bayesian classifiers such that:
  - Medical background knowledge of the breast cancer domain is incorporated,
  - The resulting classifiers are faithful w.r.t. to the data. i.e., the dependences and independences of the data are represented correctly.
- Develop new image representation techniques using quantitative image analysis and automated detection of potential abnormalities.
- Extend mammographic breast cancer data sets by annotation and processing of representative samples from the screening archive.
- Determine experimentally how CAD can be used as a decision aid in screening practice through data acquired in observer studies.

6.2 Background

Bayesian Networks and Bayesian Classifiers

Bayesian Network (BN) approaches offer substantial advantages in terms of flexibility and potential for knowledge reuse in the development of medical decision support systems due to the declarative, i.e., task-non-specific nature of the knowledge embodied in such systems [62]. BNs are examples of so-called graphical statistical models, and offer an encoding of a joint probability distribution on the variables in a problem domain in terms of local (conditional) probability distributions and a graph which represents the statistical dependences and independences [21, 40, 75, 76]. The BN formalism can easily be augmented with utility information and by taking some variables as decision variables can be used for the selection of optimal (sequences) of decisions. The resulting formalism is known as the *influence diagram* or *decision network* formalism [81, 40]. Experience, both by the research team and others, suggest that clinicians find this computational formalism intuitive and appealing [36, 35, 61, 63].

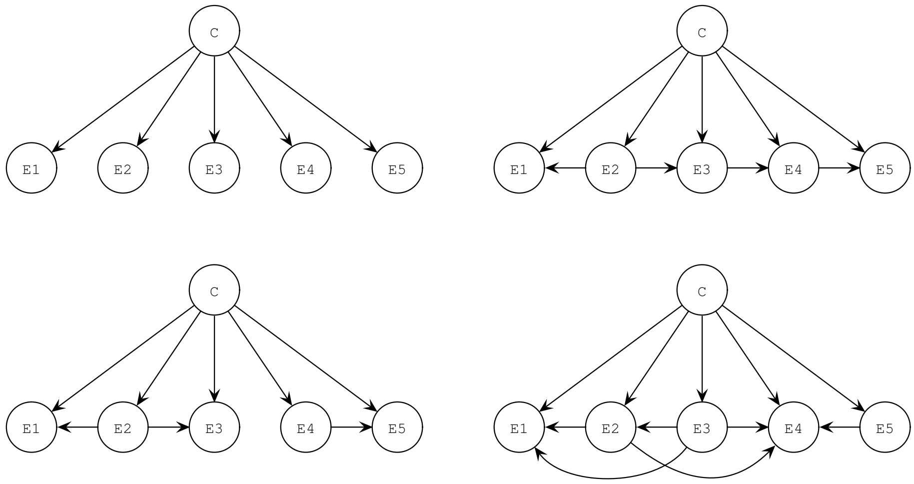


Figure 1: Bayesian Network classifiers based on different models for representing dependencies between attributes. (top left) The Naive Bayesian Network classifier (NB), (top right) Tree-Augmented Naive Bayesian Network classifier (TAN), (bottom left) Forrest-Augmented Naive Bayesian Network classifier (FAN), and (bottom right) Bayesian Network Augmented Naive Bayesian classifier (BAN).

BNs conforming to the topology of Figure 1 correspond to the situation where a distinction is made between *evidence variables* E_i and a class variable C, with the evidence variables assumed to be conditionally independent given the class variable. Such networks are normally used for classification purposes by determining the class value with maximum a posteriori probability, i.e., $max_C\{Pr(C \mid E_1, \ldots, E_m)\}$. Even with such strong simplifying assumptions, [22] convincingly shows that such naive Bayesian classifiers yield surprisingly powerful classifiers. Subsequently, [29] showed that by linking evidence variables into a tree, i.e., a tree-augmented Bayesian network (TAN), this network again outperforms the naive Bayesian network. Further extensions are possible, i.e., forest-augmented Bayesian networks (FANs) [59], which allows clusters of linked evidence variables, or at the extreme end of representations fully unrestricted Bayesian classifiers.

A BN for a real-life problem is usually constructed with the help of one or more expert clinicians [89]. Although feasible, this can be very time consuming. To make the construction of BNs more feasible it is inevitable to base its representation on observational clinical data as much as possible. For the current research proposal, the access to a unique national database once digital screening is in place, and a concentration of expertise in the area of breast cancer and mammography in participating institutes, provides a solid research environment. Very few researchers have actually tried to learn BNs for real-life problems. In many of the papers, experimental results are based on data sets generated by Monte-Carlo simulation given a BN. This makes comparing results straightforward, but may be substantially different from results achieved with real-life data sets. In this project, the learning of BNs will be studied in a real-life context with input from expert radiologists, thereby addressing a gap in current research on BNs.

Breast Cancer Screening and Computer Aided Detection

Breast cancer screening programs have been established as an effective way to reduce mortality from breast cancer [74, 25]. It is well known that the effectiveness of these programs strongly depends on high quality of the screening procedure. In the past, a lot of attention has been focused on technical quality assurance to guarantee optimal mammographic image quality. Currently, however, experts agree that the quality of mammographic interpretation is the weakest link in the process. Review studies [90, 33, 45, 7, 94, 10] have revealed that observer errors are frequent in breast cancer screening. On the basis of these studies, it is estimated that 20%-30% of cancers could be detected in an earlier screening without an increase in the recall rate to an unacceptable level.

The Radboud University Nijmegen Medical Center is playing an active role in the breast cancer screening program in the Netherlands. It houses the 'National Expert and Training Center for Breast Cancer Screening' (LRCB) which is responsible for quality assurance and training of screening radiologists. Research is conducted in cooperation with the department of radiology, where a successful program on image processing of mammograms and CAD is carried out. In this program large annotated databases of digitized film mammograms and expertise on all aspects of mammogram reading are available. These databases can expand rapidly when direct digital mammography is going to be employed in screening. New tools will need to be developed though, to handle digital data from different manufacturers, and to annotate and describe abnormalities that are visible in the images.

Computer-aided detection (CAD) in mammography is an active area of research and commercial systems are widely used, nowadays, in particular in the United States. It is the first successful pattern recognition application in medical imaging. CAD has primarily been designed to generate prompts on suspicious areas on mammograms and has been shown to be effective in detecting cancers at an early stage, in which radiologists often fail to detect them [10, 83, 6]. Typically, these prompting systems operate at a high sensitivity, but their specificity is only modest. In a prospective study it was found that the sensitivity of screening increased with 19.5% by using CAD in combination with single reading [26]. Although positive results have been confirmed by a number of other reports, the benefit of CAD is still debated, in particular within a European context where screening practice is rather different than screening in the US. It is felt that CAD should also help with interpretation of abnormalities, rather than just prompting them, in order to help decide which cases need to be referred to a hospital. A number of studies have shown that radiologists may improve interpretation of lesions using CAD programs for classification of masses [39, 16, 32] or for classification of microcalcifications [42, 41, 43, 92, 93].

6.3 Related Research and Collaborations

Dr. Lucas, who has been trained as a medical doctor and computer scientist, has extensive experience in building BN based systems for clinical problems in collaboration with clinicians, e.g., [30, 47, 46, 60, 31, 54, 58, 61, 63]. Dr. Lucas also has 20 years experience in dealing with the methodological questions which arise in research in medical decision support, and has close contact with other groups in Utrecht (Computing Science) and Nijmegen (Medical Physics) with complementary expertise in the area of BNs.

Dr. Ir. Karssemeijer received his MS degree in Physics from Delft University of Technology. He became involved in medical imaging when he joined the laboratory of Medical Physics and Biophysics of the Radboud University in Nijmegen, where he received a PhD for his research on development of statistical methods for automated medical image analysis. Since then he is employed by the Department of Radiology of the Radboud University Nijmegen Medical Center, where he holds a position as associate professor in image analysis. His research is aimed at development of computer aided diagnosis and understanding perception in medical imaging, with a focus on breast cancer screening, e.g., [91, 88, 86, 73, 66, 87, 84, 93, 50, 48, 49, 83, 92].

There is a close cooperation with regional screening organizations in the area of Nijmegen and Utrecht. The combined volume of these organizations is more than 100,000 screenings each year. In Nijmegen an experimental breast cancer screening program was already setup in 1975, which makes it the second oldest breast cancer screening program world-wide. The expertise built over the years on radiological and histological aspects breast cancer led to the foundation of the 'National Expert and Training Center for Breast Cancer Screening' (LRCB) in Nijmegen, which trains radiologist, pathologists, and radiographers involved in screening. The LRCB and the department of Radiology work in close cooperation. Furthermore, there are contacts with a number of research groups that work in the area of digital mammography and CAD.

6.4 Research Methodology

Image Analysis and Data Representation

Detection of breast cancer in mammograms can be modelled as a multi-stage process. In a first step a search is carried out to identify locations of interest in the images. Sensitive methods for automating this step have been developed in the past and will be used in this project. These methods comprise detection of masses, microcalcifications, architectural distortion, and asymmetry (e.g., [86, 66, 84, 87, 86, 50]), which are all signs of breast cancer. Alternatively, radiologists may identify these locations. In a second stage, locations of interest are inspected further, to identify women who are at high risk for having breast cancer. These are referred to hospitals for further diagnosis. This second stage will be the focus of this project. It is a complex classification problem in which in a large sample of initial image locations only few abnormalities need to be found. Complexity arises because of subtlety of the image features and of the heterogeneous nature of the classification problem. Sometimes an abnormality is only visible in one view, in other cases it can be seen in two projections of the breast (typically two views of each breast are acquired at screening). In other cases a previous screening mammogram is available in which the abnormality may be visible as well. In that case growth needs to be taken into account. Lesions may be obscured, however, by normal parenchymal patterns, so that importance of visibility of features in additional views varies. In addition, various signs like masses and microcalcifications may occur alone or in combination. Simple classification methods used thus far in CAD mostly treat images and signs separately but cannot deal well with combinations of views and signs. BNs are expected to be more suited for this task. In the project we will focus on two aspects. The first will be further development of image feature extraction and standardized data representation based on classification of local image features in single views. The second aspect will be the combination of information extracted from different views and global pattern analysis in a BN. Global pattern analysis will include breast density measurement, which is one of the most important risk factors for breast cancer.

Learning Bayesian Networks from Data

Learning a BN consists of two tasks, namely learning the *structure*, i.e., identifying the topology of the network, and learning the *parameters*, i.e., determining the associated conditional probability distributions for a given network topology. As the number of possible BN structures for a given set of variables is super exponentially large, it is necessary to use heuristic methods to construct a BN automatically using medical data [20, 55, 82, 17].

The variables in the BN will represent image features computed from the mammograms. Usually these are computed after the images are pre-processed, i.e., a segmentation of breast tissue, background tissue, and pectoral muscle [48], peripheral enhancement to correct for differences in tissue thickness, and removal of the sharp transition in grey level from the breast area to the pectoral region [88]. Image features will include stellate pattern features [50] (malignancies more often are surrounded by a radiating pattern of spicules), mass features (malignancies tend to have a greater density), location features (most malignancies (45%) develop in the upper outer quadrant [12]), morphological features (malignant masses are on average larger than benign ones [86]), linear texture features [50] (linear structures often indicate the presence of normal breast tissue), and border features [91] (benign lesions tend to have a smooth circle like shape). Currently up to 80 image features are computed to represent masses in a mammogram.

Additional features representing microcalcifications and features extracted from other views of the breast will lead to more features than a BN can handle. Therefore, we will first investigate reduction of feature sets by feature selection and principal component analysis. Most effort will be put though in using expert knowledge to train a series of simple classifiers on separate subsets (representing different signs, views, or distinct aspects like lesion boundary and effect of a lesion on its surroundings). The output of these simple classifiers will then be used as input for the BN. Initial experience has shown already that with a subset of 11 features a Naive Bayesian classifier performs as well as a Support Vector Machine [85] in classification of benign and malignant masses.

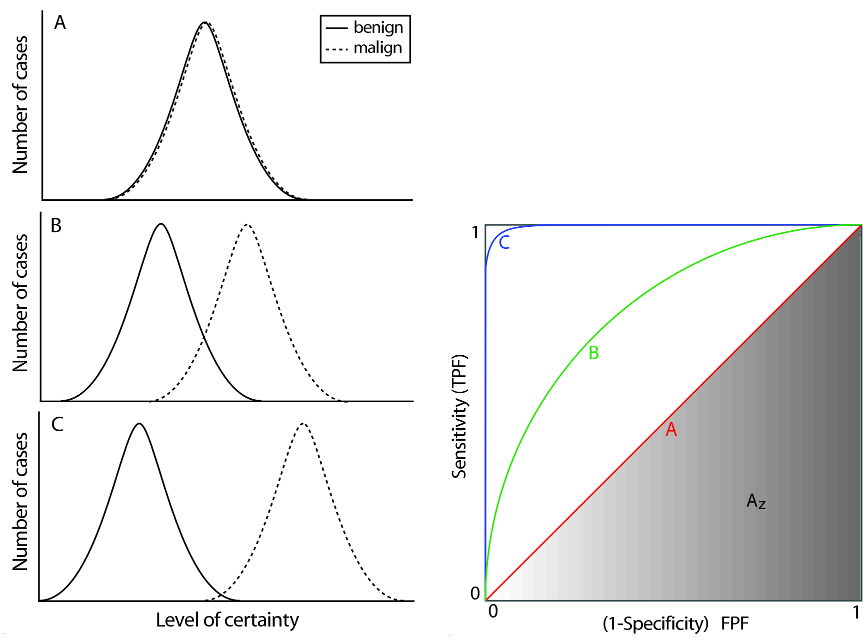
As most image features are continuous by nature and discretizing continuous attributes may risk losing information that is relevant for the classification task [28], image features will be represented using continuous variables. Although discrete variables are more often used for successfully building BNs, plenty of experience has been acquired in the last decade for continuous variables [27, 71, 77, 79]. Most software packages now offer support for continuous variables, although possibly using different modeling assumptions [70, 53]. A preliminary investigation [80] has shown us continuous variables to be possible when using Matlab [69], which supports Gaussian nodes analytically, after some data normalizing transformations [44, 65].

Classifier Training and Evaluation

We will start BN learning using existing algorithms [1, 18, 11, 20, 55, 82, 17] and existing software packages [38, 2, 70, 69, 19]. A preliminary investigation has been performed with Matlab, but also commercially available software for BN construction offering analogous support for continuous variables is available within the ICIS department [38, 2, 64]. New algorithms will be developed during the project as demanded by the problem domain. For the scoring algorithm, we will use N-fold cross-validation. This method divides the dataset in N subsets of (approximately) equal size. The network is trained N times, each time leaving out one of the subsets from training. Only the omitted set is used to evaluate the classifier performance. Good results have been obtained for 10-fold cross-validation [52, 8] and averaging the performance error of all N BNs will give a good estimate of the true generalization error [95, 52].

As already stated an extension we intend to investigate of the simple classifier model is the combination of simple classifiers into a more complex classifier [51, 9]. Furthermore, we will explore the introduction of background knowledge (e.g., age, breast density) into less restricted BNs as we expect this to improve the performance and especially the faithfulness of the representation. In addition the use of hidden nodes [4, 5] to create dependences between similar image features taken from a different view or taken at a different time will be studied.

The evaluation of the classifiers will be based on Receiver Operator Characteristic (ROC) [68, 67] or Free-response Receiver Operator Characteristic (FROC) methodology [15, 14, 13], which are widely accepted evaluation methodology standards in Radiology. Classifying abnormalities as normal (including benign) or malignant will be evaluated with ROC methodology. ROC curves usually plot sensitivity, also called true positive fraction (TPF), as a function of 1-specificity, also called false positive fraction (FPF) (Figure 2). The performance of the system will be based on both image based and lesion based. The evaluation of one case, which may consist of multiple images, is often obtained by taking the minimum, maximum, or average of the image- or lesion based evaluations. To evaluate breast cancer detection we use FROC methodology, which plots the average number of false positive detections per images against the fraction of correctly detected lesions (sensitivity).



(a) (b)

Figure 2: In (a) the certainty of a tumor being benign relative to malign is given and (b) shows their corresponding ROC curves. The area under the curve, the value A_z is used as a performance measure. The better we are to distinguish malign from benign cases, the higher A_z will be for the corresponding ROC curve.

Datasets

In breast cancer screening at the first screening usually two mammographic views are taken, the medio lateral oblique (MLO) and cranio caudal (CC) views. At subsequent screenings usually only MLO views are taken unless additional views are necessary. This makes breast cancer datasets somewhat complex, e.g., some cases have multiple views while others do not, some cases have prior mammograms while others do not, or some cases have lesions in two views while others have lesions only in one view. Datasets that are currently available are mostly digitized from film. These include normal and abnormal screening exams. All abnormal cases were referrals from screenings or interval cancers of which 50% are cancer cases and the rest is benign or normal, verified by biopsy or follow-up. As in the Netherlands the positive predictive value of screening is roughly 40%, this explains the high value of cancer cases.

When collecting datasets it is standard practice in our group to obtain the last screening mammogram and up to two previous screening examinations of a patient, if available. In previous projects many mammographic datasets have been digitized. Currently our database contains more than 8000 digitized films and 4000 digitally acquired mammograms. Abnormalities in these mammograms have been annotated by drawing the outline and describing pathology based on BI-RADS [23, 72]. Collection of new data will be an ongoing task in this project, where we focus on centers that have converted to digital acquisition. In the coming years, it may become feasible to use ultrasound in screening when certain kind of abnormalities are seen by the radiographer. These images will be included when available.

New Methods for Feature Extraction

The success of CAD is strongly dependent on adequate representation of the image data. In previous projects we have developed methods for detection and classification of masses and microcalcifications, and to some extent to architectural distortion of parenchymal patterns [50]. In this project we will develop some of these methods further, to address known weaknesses. In particular we will focus on representation of linear patterns and architectural distortion. Linear patterns (e.g., vessels) often cause false positive detections. To reduce these, methods for detection of vessels and other normal linear structures should be investigated. For architectural distortion, our current method is adequate only as long as they can be modelled by a pattern of radiating straight lines. Our method is based on multi-scale statistical analysis of vector fields. In this project we will generalize this method to make it sensitive for curved diverging linear patterns. The idea is to quantify local divergence and to use non-linear diffusion to accumulate evidence for curved diverging patterns of lines to a center of radiation.

Although digitization of the screening program has great advantages for application of CAD, the used equipment will need to rely on digitized film for a substantial period. In this project we will address this practical issue by implementing image normalization techniques to convert mammograms to normalized space independent of vendor and equipment used.

Observer Studies and estimation of the value of CAD

We will use data from observer studies to determine the effect of computer assisted detection using the methods developed in the project. The potential effect of CAD on interpretation of abnormalities can be determined by treating the radiologist and CAD as independent observers. By computing performance of a weighted combination of CAD and the observer scores an estimate of combined performance can be obtained. We have previously conducted a large observer study in which 15 experienced screening radiologists read 500 mammograms (including prior views). In each mammogram these readers marked locations of interest and rated the likelihood that cancer was present. In total around 6000 observations were recorded. We have shown that by combining observer scores with CAD performance in breast cancer screening can be improved, and that better CAD systems lead to better combined performance. Current performance of CAD, however, is still worse though than any of the observers in the study. Similar observer studies will be carried out in parallel to the proposed project, which will increase the amount of available observer data.

We intend to use the methodology outlined above to determine the potential effect of the use of new methods developed in this project. It allows great flexibility to experiment with different decision strategies. In fact, the observer can also be included as a node in the BN to explore ways to understand and maximize the performance of a combined observer. In addition to these experiments, it will be also be investigated how radiologists may interact with a CAD system in practice. A prototype viewing system will be developed to conduct observer studies with computer assisted reading using methods developed in the project. One of the participating institutes is the 'National Expert and Training Center for Breast Cancer Screening' (LRCB) in Nijmegen, which trains radiologists, pathologist, and radiographers involved in screening. The radiologists trained at the LRCB may be involved in the observer studies.

6.5 Relevance

Although doubt still remains about the effectiveness of CAD in mammography, in particular when used for the detection of masses, it is our belief that this is partly due to a misunderstanding of the problem. Current CAD systems are intended to be used to avoid oversights, i.e., prompting systems, instead of being used as an aid in interpreting findings. Problems in reading mammograms are therefore only partly addressed by current technology. This project intends to close this gap. Contrary to previous research, which was mostly targeted at highly suspect regions with low levels of suspiciousness, this project will focus on mostly normal regions with low levels of suspiciousness, i.e., those that raise suspicion in screening but for which radiologists hesitate to either recall or not. Preliminary work [49] shows that improving decisions in this category has the potential to lead to a very significant improvement of detection, without increasing referral rates.

Using recent data it is estimated that due to screening breast cancer mortality in the Netherlands has decreased by 800 cases per year. New computer-aided detection and diagnosis methods developed in this project can lead to higher sensitivity (without loss of specificity) and less late stage detection, which is essential for increasing the effect of screening and consequently a further reduction of breast cancer mortality.

6.6 Embedding of Research

The proposed research is in line with research carried out by Dr. Lucas and Dr. Ir. Karssemeijer. Currently, the Nijmegen Institute for Computing and Information Sciences is involved in two projects on the learning of BNs from data. The first project, TimeBayes, focusses on learning dynamic BNs, whereas the second project, ProBayes, focusses on BNs for the purpose of prognosis. Both projects are clearly related to the current proposal and some results may even be exchanged. However, all projects also complement each other and address the application of BNs to particular domains, which is mostly lacking in current literature.

The department of Radiology of the Radboud University Nijmegen Medical Centre is involved in research projects aimed at development and evaluation of new diagnostic techniques. The current application is in line with the 'Computer-aided diagnosis and digital mammography' research line, which is part of the Oncology research program (KUN MED 1) of the medical faculty of the University of Nijmegen. In the past decade the Dutch Cancer Society has funded several research projects on CAD of breast cancer and a new project in this area has recently been accepted. The department also plays an active role in advising on the conversion of the Dutch breast cancer screening program from film to digital. Furthermore, two EU 5th framework IST projects (SCREEN and SCREEN-TRIAL) were initiated together with the German research institute MeVis in Bremen in which a dedicated workstation for breast cancer screening was developed. As part of the SCREEN-TRIAL project the first digital screening unit in the Netherlands, which uses Computer-aided detection, has been set up at the Preventicon in Utrecht.

7 Description of the Proposed Plan of Work

This research proposal consists of the following tasks:

- Setting up collaboration with expert radiologists.
- Analyze the decision making process of radiologists.
- Collecting and annotating digital mammograms.
- Study the literature (i.e., machine learning, Bayesian Networks, image processing).
- Study techniques for image analysis and data representation.
- Study techniques for multi-stage image classification.
- Develop new image feature extraction methods and image normalization techniques.
- Identification of suitable background knowledge in breast cancer screening.
- Development of Bayesian classifiers.
- Software development to support observer studies with radiologists, e.g., a user interface connected to a viewer.

The required tasks will be divided into two closely linked subprojects:

Postdoc ICIS

- Startup Phase (months 0-6)
  - Collecting and studying literature
  - Analysing breast cancer data sets
  - Study decision making process
- Development Phase (months 7-24)
  - Exploiting background knowledge
  - Using multi stage classifiers
  - Using multiple image views
- Evaluation Phase (months 25-30)
  - Experimentation
  - Fine tuning and refinement theory
- Dissemination Phase (months 31-36)
  - Writing scientific journal and conference papers

Postdoc UMC St. Radboud

- Startup Phase (months 0-6)
  - Collecting and studying literature
  - Training mammogram reading
- Development Phase (months 7-24)
  - Identify image features
  - Develop new image feature extraction techniques
  - Develop Image normalization techniques
- Evaluation Phase (months 25-30)
  - Optimizing decision strategies using observer data
  - Pilot studies
  - Observer Studies
- Dissemination Phase (months 31-36)
  - Writing scientific journal and conference papers

Figure 3: Work programme

The first project will be supervised by Dr. Lucas and can be done by a postdoc. The focus will be on the development of classifiers for CAD. Starting using existing data sets and background knowledge provided by the Radiology department. Finishing with experiments to evaluate and fine tune the classifiers. The second project will be supervised by Dr. Ir. Karssemeijer and done by a postdoc. The focus will be on obtaining the data sets, extending the data base annotations using image normalization and feature extraction techniques, and finally doing observer studies to investigate whether a CAD system can interactively be used in the interpretation of mammogram screening. The time schedule of both projects is given in Figure 3. As can be seen, both projects run in parallel and provide feedback to each other, e.g., data sets with new image features for classifier development, or classifiers for doing the observer studies with the digitized Dutch breast cancer data base, which will be made available in 2006 and 2007. For the digitization and database maintenance a scientific programmer will be needed. In addition a software environment needs to be created to support observer studies. Some software is already being developed to make the transition smoother.

8 Expected Use of Instrumentation

For each new staff member a PC is required, including basic software.

9 Publications

Selected Publications by Applicants

1. P.J.F. Lucas, H. Boot, and B.G. Taal. Computer-based decision-support in the management of primary gastric non-Hodgkin lymphoma. *Methods of Information in Medicine*, 37:206–219, 1998.
2. P.J.F. Lucas, N.G. de Bruijn, K. Schurink, and I.M. Hoepelman. A Probabilistic and decision-theoretic approach to the management of infectious disease at the ICU, *Artificial Intelligence in Medicine*, 19(3):251–279, 2000.
3. N. Karssemeijer, J.D.M. Otten, A.L.M. Verbeek, J.H. Groenewoud, H.J. de Koning, J.H.C.L. Hendriks, and R. Holland. Computer-aided detection versus independent double reading of masses on mammograms. *Radiology*, 227:192–200, 2003.
4. P.J.F. Lucas. Bayesian network modelling through qualitative patterns. Artificial Intelligence, 163(2):233–263, 2005.
5. K.J. McLoughlin, P.J. Bones, and N. Karssemeijer. Noise equalization for detection of microcalcification clusters in direct digital mammogram images. *IEEE Transactions on Medical Imaging*, 23(3):313–320, 2004.

References

[1] S. Acid, L.M. de Campos, and J.G. Castellano. Learning bayesian network classifiers: Searching in a space of partially directed acyclic graphs. *Machine Learning, Special Issue on Probabilistic Graphical Models for Classification*, 59(3):213–235, 2005.

[2] S.K. Andersen, K.G. Olesen, F.V. Jensen, and J. Jensen. HUGIN – a Shell for Building Bayesian Belief Universes for Expert Systems. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence*, 1989.

[3] S. Andreassen, M. Woldbye, B. Falck, and S.K. Andersen. MUNIN – a causal probabilistic network for interpretation of electromyographic findings. In *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, pages 366–372, Milan, Italy, 1987.

[4] J.W. Bang and D. Gillies. Using Bayesian Networks to Model the Prognosis of Hepatitis C. In *7th Workshop on Intelligent Data Analysis in Medicine and Pharmacology*, pages 7–15, Lyon, France, 2002.

[5] J.W. Bang and D. Gillies. Using Bayesian Networks with Hidden Nodes to Recognise Neural Cell Morphology. In M. Ishizuka and A. Satter, editors, *7th Pacific Rim International Conference on Artificial Intelligence*, pages 385–394, Tokyo, New York, 2002. Springer.

[6] R.L. Birdwell, D.M. Ikeda, K.F. O'Shaughnessy, and E.A. Sickles. Mammographic characteristics of 115 missed cancers later detected with screening mammography and the potential utility of computer-aided detection. *Radiology*, 219:192–202, 2001.

[7] R.G. Blanks, M.G. Wallis, and S.M. Moss. A comparison of cancer detection rates achieved by breast cancer screening programmes by number of readers, for one and two view mammography: results from the UK National Health Service breast screening programme. *Journal of Medical Screening*, 5(4):195–201, 1998.

[8] L. Breiman. Heuristics of instability and stabilization in model selection. *Annals of Statistics*, 24:2350=–2383, 1996.

[9] H. Bunke and A. Kandel, editors. *Hybrid Methods in Pattern Recognition*, volume 47 of *Machine Perception and Artificial Intelligence*, 2002.

[10] L.J.W. Burhenne, S.A. Wood, C.J.D. D'Orsi, S.A. Feig, D.B. Kopans, K.F. O'Shaugnessy, E.A. Sickles, L. Tabar, J. Vyborny, and R.A. Castelino. The potential contribution of computer aided detection to the sensitivity of screening mammography. *Radiology*, 215:554–562, 2000.

[11] R. Castelo and T. Kocka. On inclusion-driven learning of Bayesian networks. *Journal of Machine Learning Research*, 4:527–574, 2003.

[12] S. Caulkin, S. Astley, J. Asquith, and C. Boggis. Sites of occurrence of malignancies in mammograms. In N. Karssemeijer, M. Thijssen, J. Hendriks, and L. Erning, editors, *Digital Mammography Nijmegen*, pages 279–282. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

[13] D. Chakraborty. Statistical power in observer-performance studies. Comparison of the Receiver-Operating characteristic and free-response methods in tasks involving localization. *Acad. Radiol*, 9:147–156, 2002.

[14] D. Chakraborty and L. Winter. Free-response methodology: alternate analysis and a new observer-performance experiment. *Radiology*, 174:873–881, 1990.

[15] D.P. Chakraborty. Maximum likelihood analysis of free-response receiver operating characteristic (FROC) data. *Medical Physics*, 16:561–568, 1989.

[16] H.P. Chan, B. Sahiner, M.A. Helvie, N. Petrick, M.A. Roubidoux, T.E. Wilson, D.D. Adler, C. Paramagul, J.S. Newman, and S. Sanjay-Gopal. Improvement of radiologists' characterization of mammographic masses by using computer-aided diagnosis: an roc study. *Radiology*, 212(3):817–827, 1999.

[17] J. Cheng and D. Bell. Learning Bayesian networks from data: an efficient approach based on information theory. In *Proceedings of the 6th ACM International Conference on Information and Knowledge Management*, 1997.

[18] D.M. Chickering. Learning equivalence classes of Bayesian network structures. *Journal of Machine Learning Research*, 2:445–498, 2002.

[19] Elvira Consortium. Elvira: An environment for creating and using probabilistic graphical models. In *Probabilistic Graphical Models*, 2002.

[20] G.F. Cooper and E.A. Herskovits. A Bayesian method for the induction of probabilistic networks from data. *Machine Learning*, 9:309–347, 1992.

[21] R.G. Cowell, A.P. Dawid, and D.J. Spiegelhalter. *Probabilistic Networks and Expert Systems*. Springer, 1999.

[22] P. Domingos and M. Pazzani. On the optimality of the simple Bayesian classifier under zero-one loss. *Machine Learning*, 29:103–130, 1997.

[23] C. D'Orsi and D. Kopaas. Mammography interpretation: the BI-RADS method. *American Family Physician*, 55(5):1548–1552, 1997.

[24] R.O. Duda and P.E. Hart. *Pattern classification and scene analysis*. John Wiley & Sons, New York, 1973.

[25] J. Fracheboud, S.J. Otto, J.A. van Dijck, M.J. Broeders, A. L. Verbeek, and H. J. de Koning. Decreased rates of advanced breast cancer due to mammography screening in the Netherlands. *British Journal of Cancer*, 91:861–867, 2004.

[26] T.W. Freer and M.J. Ulissey. Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. *Radiology*, 220(3):781–786, 2001.

[27] N. Friedman, M. Goldszmidt, and T.J. Lee. Bayesian network classification with continuous attributes: getting the best of both discretization and parametric fitting. In *Proceedings of the 15th International Conference on Machine Learning*, pages 179–187, San Fransisco, CA, 1998. Morgan Kaufmann Publishers.

[28] N. Friedman, M. Linial, and I. Nachman. Using Bayesian Networks to analyze expression data. *Journal of Computational Biology*, 7(3-4):601–620, 2000.

[29] N.I.R. Friedman, D. Geiger, and M. Goldszmidt. Bayesian network classifiers. *Machine Learning*, 29:131–163, 1997.

[30] M. Gerven, P.J.F. Lucas, and T. van der Weide. A Qualitative Characterisation of Causal Independence Models using Boolean Polynomials. In L. Godo, editor, *Symbolic and Quantitative Approaches to Reasoning with Uncertainty: 8th European Conference (ECSQARU 2005)*, volume 3571 of *Lecture Notes in Computer Science*, pages 244–256, Barcelona, Spain, 2005. Springer-Verlag GmbH.

[31] M. van Gerven and P.J.F. Lucas. Employing Maximum Mutual Information for Bayesian Classification. In *Biological and Medical Data Analysis: 5th International Symposium (ISBMDA 2004)*, volume 3337 of *Lecture Notes in Computer Science*, pages 188–199, Barcelona, Spain, 2004. Springer-Verlag GmbH.

[32] L.M. Hadjiiski, H. Chan, B. Sahiner, M.A. Helvie, M.A. Roubidoux, C.E. Blane, C. Paramagul, N. Petrick, J.E. Bailey, K. Klein, M. Foster, S. Patterson, D.D. Adler, A. Nees, and J. Shen. ROC study: effects of computer-aided diagnosis on radiologists' characterization of malignant and benign breast masses in temporal pairs of mammograms. *SPIE Medical Imaging, Image Processing*, 5032:94–101, 2003.

[33] J.E. Harvey, L.L. Fajardo, and C.A. Inis. Previous mammograms in patients with impalpable breast carcinoma: retrospective vs blinded interpretation. *American Journal of Roentgenology*, 161:1167–1172, 1993.

[34] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York, 2001.

[35] D. Heckerman. A tutorial on learning with Bayesian networks. In M. Jordan, editor, *Learning in Graphical Models*. MIT Press, 1999.

[36] D.E. Heckerman, E.J. Horvitz, and B.N. Nathwani. Toward normative expert systems: part II – probability-based representations for efficient knowledge acquisition and inference. *Methods of Information in Medicine*, 31:106–116, 1992.

[37] D.E. Heckerman and B.N. Nathwani. Toward normative expert systems: part I – the Pathfinder project. *Methods of Information in Medicine*, 31:90–105, 1992.

[38] HUGIN Expert [online, cited 1 Februari 2006]. Available from: `http://www.hugin.com`.

[39] Z. Huo, M.L. Giger, C.J. Vyborny, D.E. Wolverton, R.A. Schmidt, and K. Doi. Automated computerized classification of malignant and benign masses on digitized mammograms. *Academic Radiology*, 5(3):155–168, 1998.

[40] F.V. Jensen. *An Introduction to Bayesian Networks*. UCL Press, London, 1996.

[41] Y. Jiang, R.M. Nishikawa, and J. Papaioannou. Dependence of computer classification of clustered microcalcifications on the correct detection of microcalcifications. *Medical Physics*, 28(9):1949–1957, 2001.

[42] Y. Jiang, R.M. Nishikawa, R.A. Schmidt, C.E. Metz, M.L. Giger, and K. Doi. Improving breast cancer diagnosis with computer-aided diagnosis. *Academic Radiology*, 6(1):22–33, 1999.

[43] Y. Jiang, R.M. Nishikawa, R.A. Schmidt, A.Y. Toledano, and K. Doi. Potential of computer-aided diagnosis to reduce variability in radiologists' interpretations of mammograms depicting microcalcifications. *Radiology*, 220(3):787–794, 2001.

[44] J.A. John and N.R. Draper. An alternative family of transformations. *Applied Statistics*, 29(2):190–197, 1980.

[45] R. Jones, L. McLean, and J. Young. Proportion of cancers detected at first incident screen which were false negative at the prevalent screen. *The Breast Journal*, 5:339–343, 1996.

[46] R. Jurgelenaite and P.J.F. Lucas. Exploiting Causal Independence in Large Bayesian Networks. *Knowledge-Based Systems*, 18:153–162, 2005.

[47] R. Jurgelenaite, P.J.F. Lucas, and T. Heskes. Exploring the Noisy Threshold Function in Designing Bayesian Networks. In M. Bremer, F. Coenen, and T. Allen, editors, *AI–2005 the Twenty–fifth SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence*, pages 133–146, Berlin, Germany, 2005. Springer–Verlag.

[48] N. Karssemeijer. Automated classification of parenchymal patterns in mammograms. *Physics in medicine & biology*, 43:365–378, 1998.

[49] N. Karssemeijer, J.D.M. Otten, A.L.M. Verbeek, J.H. Groenewoud, H.J. de Koning, J.H.C.L. Hendriks, and R. Holland. Computer-aided detection versus independent double reading of masses on mammograms. *Radiology*, 227:192–200, 2003.

[50] N. Karssemeijer and G.M. te Brake. Detection of stellate distortions in mammograms. *IEEE Transactions on Medical Imaging*, 15:611–619, 1996.

[51] J. Kittler and R. Fabio, editors. *Multiple Classifier Systems*, volume 2096 of *Lecture Notes in Computer Science*, Cambridge, UK, 2001.

[52] R. Kohavi. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. In *International Joint Conference on Artificial Intelligence*, pages 1137–1145, 1995.

[53] K.B. Korb and A.E. Nicholson. *Bayesian Artificial Intelligence*. Chapman & Hall / CRC Press, 2003.

[54] M. Korver and P.J.F. Lucas. Converting a rule-based expert system into a belief network. *Medical Informatics*, 18(3):219–241, 1993.

[55] W. Lam and F. Bacchus. Learning Bayesian belief networks: an approach based on the MDL principle. *Computational Intelligence*, 10:269–293, 1994.

[56] P. Langley, W. Iba, and K. Thompson. An analysis of Bayesian classifiers. In *Proceedings of the National Conference on Artificial Intelligence*, pages 223–228, 1992.

[57] P. Larrañaga, J.A. Lozano, J.M. Peña, and I. Inza, editors. *Special issue on Probabilistic Grahpical Models for Classification*, volume 59 of *Machine Learning*, 2005.

[58] P.J.F. Lucas. Knowledge acquisition for decision-theoretic expert systems. *AISB Quarterly*, 94:23–33, 1996.

[59] P.J.F. Lucas. Restricted Bayesian Network Structure Learning. In G. A. Gomez and S. Moral, editors, *Advances in Bayesian Networks, Studies in Fuzziness and Soft Computing*, pages 217–232. Springer-Verlag, Berlin, 2004.

[60] P.J.F. Lucas. Bayesian network modelling through qualitative patterns. *Artificial Intelligence*, 163(2):233–263, April 2005.

[61] P.J.F. Lucas, H. Boot, and B.G. Taal. Computer-based decision-support in the management of primary gastric non-Hodgkin lymphoma. *Methods of Information in Medicine*, 37:206–219, 1998.

[62] P.J.F. Lucas, H. Boot, and B.G. Taal. Decision-theoretic network approach to treatment management and prognosis. *Knowledge-based Systems*, 11:321–330, 1998.

[63] P.J.F. Lucas, N.G. de Bruijn, K. Schurink, and I. M. Hoepelman. A Probabilistic and decision-theoretic approach to the management of infectious disease at the ICU. *Artificial Intelligence in Medicine*, 19(3):251–279, 2000.

[64] A.L. Madsen, F. Jensen, U.B. Kjrulff, and M. Lang. HUGIN - The Tool for Bayesian Networks and Influence Diagrams. *International Journal of Artificial Intelligence Tools*, 14(3):507–543, 2005.

[65] B. Manly. Exponential data transformations. *The Statistician*, 25:37–42, 1976.

[66] K.J. McLoughlin, P.J. Bones, and N. Karssemeijer. Noise equalization for detection of microcalcification clusters in direct digital mammogram images. *IEEE Transactions on Medical Imaging*, 23(3):313–320, 2004.

[67] C. Metz, B. Herman, and C. Roe. Statistical comparison of two ROC-curve extimates obtained from partially-paired datasets. *Medical Decision Making*, 18(1):110–121, 1998.

[68] C. Metz, B. Herman, and J. Shen. Maximum likelihood estimation of receiver operating characteristic (ROC) curves from contiuously-distributed data. *Statistics in Medicine*, 17(9):1033–1053, 1998.

[69] K. Murphy. The Bayes Net Toolbox for Matlab [online]. 2001 [cited 1 Februari 2006]. Available from: `http://bnt.sourceforge.net/`.

[70] K. Murphy. Software Packages for Graphical Models / Bayesian Networks [online]. 2005 [cited 1 Februari 2006]. Available from: `http://www.cs.ubc.ca/~murphyk/Software/bnsoft.html`.

[71] C. Needham, J. Bradford, A. Bulpitt, and D. Westhead. Application of Bayesian Networks to two classification problems in bioinformatics. In *Proceedings of the 24th Leeds Annual Statistical Research Workshop*, pages 87–90, 2005.

[72] S. Orel, N. Kay, C. Reynolds, and D. Sullivan. BI-RADS categorization as a predictor of malignancy. *Radiology*, 211:845–850, 1999.

[73] J.D.M. Otten, N. Karssemeijer, J.H. Hendriks, J.H. Groenewoud, J. Fracheboud, A.L.M. Verbeek, H.J. de Koning, and R. Holland. The effect of recall rate on earlier detection of both interval and screen-detected cancers in breast screening programs based on the dutch performance indicators. *JNCI*, 2005.

[74] S. J. Otto, J. Fracheboud, C.W. Looman, M.J. Broeders, R. Boer, J.H. Hendriks, A.L. Verbeek, and H.J. de Koning. Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: as systematic review. *The Lancet*, 361(9367):1411–1417, 2003.

[75] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Francisco, 1998.

[76] M. Ramoni and P. Sebastiani. Bayesian methods for intelligent data analysis. Technical report, Knowledge Media Institute, The Open University, Milton Keynes, UK, 1998.

[77] J.R. Rice, Y. Tu, and G. Stolovitzky. Reconstructing biological networks using conditional correlation analysis. *Bioinformatics*, 2005.

[78] I. Rish. An empirical study of the naive Bayes classifier. In *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*, 2001.

[79] S. Rogers and M. Girolami. A Bayesian regression approach to the inference of regulatory networks from gene expression data. *Bioinformatics*, 2005.

[80] S. Samulski. Computer-aided detection of lesions in digital mammograms using Bayesian classifiers. Master Thesis, Unpublished, 2006. Available from: `http://www.cs.ru.nl/mtl/`.

[81] R.D. Shachter. Evaluating influence diagrams. *Operations Research*, 34(6):871–882, 1986.

[82] M. Sigh and M. Voltorta. Construction of Bayesian network structures from data: a brief survey and an efficient algorithm. In *Uncertainty in Artificial Intelligence*, pages 259–265. Morgan Kaufmann, 1993.

[83] G.M. te Brake and N. Karssemeijer. Automated detection of breast carcinomas that were not detected in a screening program. *Radiology*, 207:465–471, 1998.

[84] G.M. te Brake and N. Karssemeijer. Segmentation of suspicious densities in digital mammograms. *Medical Physics*, 28:259–266, 2001.

[85] S. Timp. Ph.d. thesis. Unpublished, 2006.

[86] S. Timp, S. van Engeland, and N. Karssemeijer. A regional registration method to find corresponding mass lesions in temporal mammograms. *Medical Physics*, 32(8):2629–2638, 2005.

[87] S. Timp and N. Karssemeijer. A new 2D segmentation method based on dynamic programming applied to computer aided detection in mammography. *Medical Physics*, 31(5):958–971, 2004.

[88] S. Timp and N. Karssemeijer. Interval change analysis to improve computer aided detection in mammography. *Medical Image Analysis*, 10(1):82–95, 2005.

[89] L.C. van der Gaag, S. Renooij, C.L.M. Witteman, B. Aleman, and B.G. Taal. How to elicit many probabilities. In K. Laskey and H. Prade, editors, *Uncertainty in Artificial Intelligence*, pages 647–654, San Francisco, 1999. Morgan Kaufmann.

[90] J.A.A.M. van Dijck, A.L.M Verbeek, J.H.C.L. Hendriks, and R. Holland. The current detectability of breast cancer in a mammographic screening program. *Cancer*, 72:1933–1938, 1993.

[91] C. Varela, S. Timp, and N. Karssemeijer. Use of border information in the classification of mammographic masses. *Physics in medicine & biology*, 51:425–441, 2006.

[92] W.J.H. Veldkamp and N. Karssemeijer. An improved method for detection of microcalcification clusters in digital mammograms. *SPIE Medical Imaging*, 3661:512–522, 1999.

[93] W.J.H. Veldkamp, N. Karssemeijer, J.D. Otten, and J.H. Hendriks. Automated classification of clustered microcalcifications into malignant and benign types. *Medical Physics*, 27:2600–2608, 2000.

[94] B Vitak. Invasive interval cancers in the Ostergotland mammographic screening programme: radiological analysis. *European Radiology*, 8(4):639–646, 1998.

[95] S.M. Weiss and C.A. Kulikowski. *Computer Systems That Learn*. Morgan Kaufmann, 1991. ISBN 1-55860-065-5.

10 Requested Budget

aanstelling postdoc ICIS (3 years)	169.299
aanstelling postdoc UMCN (3 years)	169.299
aanstelling scientific programmer (2 years, 0.5fte)	60.006
persoonsgebonden benchfee postdoc	10.000
Total	408.604