Improved Mammographic CAD Performance Using Multi-View Information: a Bayesian Network Framework[1]

Marina Velikova [a] Maurice Samulski [a] Peter J.F. Lucas [b]
Nico Karssemeijer [a]

[a] *Radboud University Nijmegen Medical Centre, The Netherlands*
[b] *Radboud University Nijmegen, The Netherlands*

1 Introduction

In the coming years, the practical application of Computer-Aided Detection (CAD) systems in the Dutch breast cancer screening program is likely to expand considerably. This trend is greatly aided by the recent introduction of digitisation of breast imaging in the entire program. Screening radiologist deal with a challenging task due to the tremendous workload with more than 800,000 mammographic exams per year in The Netherlands, the low breast cancer incidence rates, and the low subtlety of the lesions, which makes them hard to distinguish from normal breast tissue. Therefore by providing a "second opinion" in image analysis, CAD can help increasing accurate early diagnosis of breast cancer. The research presented here demonstrates a CAD system that integrates the knowledge and the working principles of radiologists in the computer programs can improve the analysis of breast images in the screening programs.

One important principle in radiologists' practice is reading and comparing two different views of the same breast: mediolateral oblique (MLO) view, taken under 45° angle, and craniocaudal (CC) view, taken top-down; see Fig. 1. The general rule is that a lesion is to be observed in both views. Most computerized systems, however, are based on a single-view principle where each view and the regions within a view are analyzed independently. Hence, the multi-view dependencies in the breast are ignored and the breast cancer detection can be obscured. To tackle these problems, we propose a *Bayesian network framework for exploiting multi-view dependencies for the analysis of screening mammograms*. The main idea of our methodology lies in combining the information available as detected regions from a single-view CAD system in MLO and CC to obtain a single likelihood measure for a patient being cancerous. In comparison to previous methods, we can outline a number of advantages of our probabilistic framework: (i) *Handling noise and missing information* - specifying and learning the network parameters in a probabilistic manner allows uncertain information to be incorporated based on the values of all the non-missing variables; (ii) *Incorporating domain knowledge* - unlike black-box approaches such as neural networks, our framework captures explicitly view dependencies through the Bayesian network structure and the definition of the conditional probability tables and (iii) *Using context information over the whole breast* - breast classification is done on the basis of simultaneous consideration of the regions automatically detected in each breast view and their links to the other view of the same breast.

2 Bayesian Network Multi-View Detection

The objective of the multi-view detection is to determine whether or not a breast (and thus the patient) is suspicious for cancer, by establishing correspondences between the 2D image region features of the breast. Fig. 1 depicts the multi-view detection scheme, where a cancerous lesion, represented by the circle, is

[1]The full version of this paper appeared in *Physics in Medicine and Biology*, Vol. 54, pp. 1131-1147, 2009.
This research is funded by the Netherlands Organisation for Scientific Research under BRICKS/FOCUS grant number 642.066.605.

projected in both views of the same breast; hence, the whole breast is cancerous. In both views a single-view CAD system detects potential cancerous regions described by a number of real-valued extracted features. In Fig. 1 regions A_1 and B_1 are true positive (TP) regions whereas A_2 and B_2 are false positive (FP) regions. Since we deal with projections of the same breast we introduce *links* (L_{ij}) between the detected regions in both views, A_i and B_j. Every link has a class with values of $true$ if A_i *OR* B_j are TP and $false$ otherwise.

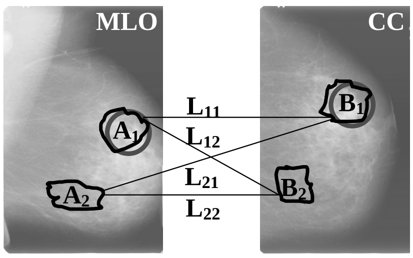

Figure 1: Multi-view analysis problem

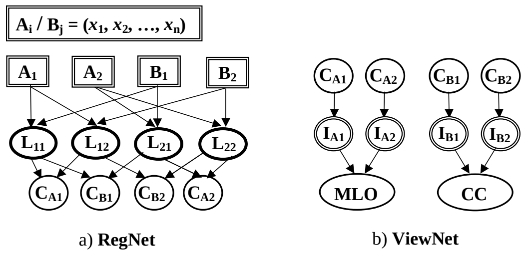

Figure 2: Bayesian network multi-view analysis

Our modelling scheme is based on two Bayesian networks with a hand-constructed (fixed) structure to explicitly represent the multi-view dependencies in the detection problem. In the context of Bayesian networks, the region dependence can be modelled by the so-called *v-structure* where directed arcs are drawn from the region nodes to the link node: $A_i \longrightarrow L_{ij} \longleftarrow B_j$. Furthermore, by definition the link variable is discrete and the regions are represented by a vector of real-valued features $(x_1, x_2, \ldots, x_n)$ extracted from a single-view CAD system. Therefore we apply logistic regression to compute $P(L_{ij} = \ell_{ij}|A_i, B_j)$. It is straightforward to construct a causal structure where all the links are modelled in parallel as shown in the first top layer in the network depicted in Fig. 2(a). Thus, using the context modelling capabilities of Bayesian networks we consider at once all the information available about the breast.

Next we estimate the probabilities $P(C_{A_i} = \textit{true}|\{L_{ij} = \ell_{ij}\}_{j=1}^{N_B})$ and $P(C_{B_j} = \textit{true}|\{L_{ij} = \ell_{ij}\}_{i=1}^{N_A})$ where C_{A_i} (C_{B_j}) is the class of region $A_i(B_j)$, $N_A(N_B)$ is the total number of regions in MLO (CC) and $\{L_{ij} = \ell_{ij}\}_{j(i=1)}^{N_B(N_A)}$ denotes the set of all links containing $A_i(B_j)$. Given our link class definition, we can easily model these conditional dependencies through a causal model using the logical OR ([1]). Next we construct a second Bayesian network (Fig. 2(b)) that combines the computed region probabilities from `RegNet` to obtain the probability of a view being cancerous using again a causal-independence model with the logical OR. Finally, we combine the view probabilities obtained from `ViewNet` into a single probabilistic measure for the breast by using two approaches: (i) averaging both view probabilities and (ii) a logistic regression model with the estimated view probabilities as input variables. The probability that a patient has cancer is the maximum out of the probabilities for the left and right breasts.

3 Experiments

We applied the proposed Bayesian network multi-view system to a dataset of 1063 screening exams from which 385 were cancerous and compared its performance to the single-view CAD system in [2]. The results showed that the incorporation of expert knowledge in a probabilistic manner led to a significantly higher breast cancer detection rate compared to the single-view CAD system. This improvement was achieved at a view, breast and patient level. In this study we also demonstrated the potential of the multi-view CAD system for prescreening purposes, where the most suspicious cases are selected as a set of difficult cases. In this way, the selected cases would get more attention from radiologists by providing additional reading and help increase the cancer detection rate. Finally, we note that the straightforward nature of the proposed Bayesian network framework allows its relatively easy application to any type of multi-view detection problems.

References

[1] D. Heckerman and J. S. Breese. Causal independence for probability assessment and inference using Bayesian networks. *IEEE Transactions on Systems, Man and Cybernetics, Part A*, 26(6):826–831, 1996.

[2] S. van Engeland, S. Timp, and N. Karssemeijer. Finding corresponding regions of interest in mediolateral oblique and craniocaudal mammographic views. *Medical Physics*, 33(9):3203–3212, 2006.