Learning Contextual Relationships in Mammograms Using a Hierarchical Pyramid Neural Network

Paul Sajda*, *Member, IEEE*, Clay Spence, *Member, IEEE*, and John Pearson

***Abstract*—This paper describes a pattern recognition architecture, which we term *hierarchical pyramid/neural network (HPNN)*, that learns to exploit image structure at multiple resolutions for detecting clinically significant features in digital/digitized mammograms. The HPNN architecture consists of a hierarchy of neural networks, each network receiving feature inputs at a given scale as well as features constructed by networks lower in the hierarchy. Networks are trained using a novel error function for the supervised learning of image search/detection tasks when the position of the objects to be found is uncertain or ill defined. We have evaluated the HPNN's ability to eliminate false positive (FP) regions of interest generated by the University of Chicago's (UofC) Computer-aided diagnosis (CAD) systems for microcalcification and mass detection. Results show that the HPNN architecture, trained using the uncertain object position (UOP) error function, reduces the FP rate of a mammographic CAD system by approximately 50% without significant loss in sensitivity. Investigation into the types of FPs that the HPNN eliminates suggests that the pattern recognizer is automatically learning and exploiting contextual information. Clinical utility is demonstrated through the evaluation of an integrated system in a clinical reader study. We conclude that the HPNN architecture learns contextual relationships between features at multiple scales and integrates these features for detecting microcalcifications and breast masses.**

***Index Terms*—Computer-aided diagnosis, context, hierarchical pyramid neural network, mammography.**

I. Introduction

COMPUTER-AIDED diagnosis (CAD) can be defined as a diagnosis made by a radiologist who incorporates the results of computer analyses of the radiographs [1]. The goal of CAD is to improve radiologists' performance by indicating the sites of potential abnormalities, to reduce the number of missed lesions, and/or by providing quantitative analysis of specific regions in an image to improve diagnosis. CAD systems typically operate as automated "second-opinion" or "double-reading" systems that indicate lesion location and/or type. Since individual human observers overlook different findings, it has been shown that "double reading" (the review of a study by more than one observer) increases the detection rate of breast cancers by 5%–15% [2]–[4]. Double reading, if not done efficiently, can significantly increase the cost of screening. Methods to provide improved detection with little increase in costs will have significant impact on the benefits of screening. Automated CAD systems are a promising approach for low-cost double-reading.

Several CAD systems have been in development and the first have been approved by the U.S. Food and Draug Administration [5]. Complete systems have been rigorously characterized, both in retrospective and prospective trials [6]. Though many have demonstrated clinical utility, there is still a need to reduce false positive (FP) rates generated by CAD systems. For example, prospective clinical studies have shown lower sensitivities and specificities than originally found in retrospective studies—80% cancers detected with 2.4 FPs per case in prospective studies versus 85%–90% sensitivity at 1–2 FPs per image in retrospective studies [7].

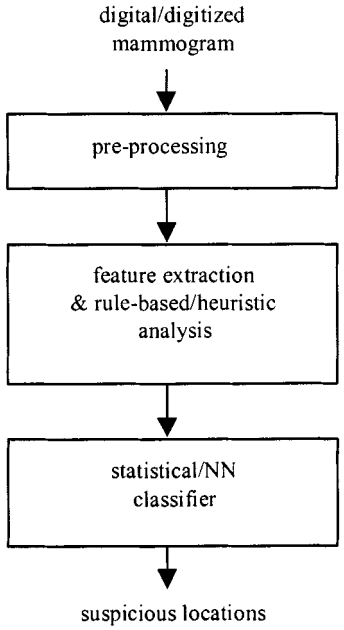

Fig. 1. Processing in a CAD system.

A. The Role of Neural Networks in CAD

CAD systems usually consist of two distinct subsystems, one designed to detect microcalcifications and one to directly detect masses [8]. A common element in both subsystems is a neural network, used to improve detection and reduce FP rates. Fig. 1

Manuscript received December 22, 2000; revised January 24, 2002. This work was supported in part by the Office of Naval Research (ONR) under Contract N00014-93-C-0202, in part by U.S. Department of the Army under Contract DAMD17-98-1-8061, in part by the Office of Women's Health under DHHS Contract 282-96-0026, in part by the Murray Foundation and in part by the National Information Display Laboratory. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was Z.-P. Liang. *Asterisk indicates corresponding author.*

*P. Sajda was with the Adaptive Image and Signal Processing Group, Sarnoff Corporation, Princeton, NJ 08540 USA. He is now with the Department of Biomedical Engineering, Columbia University, New York, NY 10027 USA (e-mail: ps629@columbia.edu).

C. Spence and J. Pearson are with the Sarnoff Research Center, Princeton, NJ 08540 USA.

Publisher Item Identifier S 0278-0062(02)04087-9.

shows a typical CAD system processing flowchart, generalized for either microcalcification or mass detection. The first two stages of the CAD system increase the overall signal-to-noise levels in the image and apply rules/heuristics to define a set of candidate regions-of-interest (ROIs). These stages have adjustable parameters that typically are set to produce a very high sensitivity, usually at a cost of low specificity. The final stage is a statistical model or neural network, whose parameters are found using error-based optimization given a set of training data. The function of this last stage is to reduce FPs (i.e., increase specificity) without significant loss in sensitivity. Neural networks are a particularly important class of statistical model in CAD because they are able to capture complicated, often nonlinear relationships in high dimensional feature spaces not easily captured by heuristic or rule based algorithms.

Several groups have developed neural networks architectures for CAD. Some of these architectures exploit well-known features that might also be used by radiologists [9]–[11], while others utilize more generic feature sets [12]–[15]. The general performance of these two approaches has been compared, particularly in the case of microcalcification detection [16]. There also have been efforts to combine the two approaches, for example using a mixed feature neural network [17]. Ultimately, performance largely depends upon the choice of features—i.e., the discriminative information in the features. Much emphasis has been placed on methods for choosing sets of discriminative features, whereas less emphasis has been placed on the specific nature of how features should be integrated. In fact, this has been a major reason for using neural networks, given they are statistical classifiers that presumably learn an optimal integration of features to maximize discrimination. However, prior information, incorporated in the architecture and training of the network, can be critical for learning an optimal integration strategy. Therefore, to maximize CAD system performance there is a need to co-optimize the feature set together with the training and architecture of the neural network.

B. Multiscale Approaches to Mass and Microcalcification Detection

Several groups have investigated using multiscale approaches for feature extraction in mammographic mass and microcalcification detection. Brzakovic and Neskovic [18] use a fuzzy pyramid approach for detecting masses while Ng and Bischof [19] apply a template at a several scales. One advantage of these approaches is that performance is independent of mass size, as further demonstrated by Miller and Ramsey [20]. Li *et al.* [21] use an oriented wavelet transform to construct features selective to spicules, thereby creating multiscale signatures for mass detection. Similarly, Netsch and Peitgen [22] use a Laplacian pyramid to develop scale-space signatures for detecting individual microcalcifications. Though these authors show that the signatures are useful for detecting individual calcifications, the signatures do not fully exploit the coarse-scale information that may be indicative of contextual relationships, critical for determining clinical significance (e.g., calcification clustering).

A few groups have investigated combining multiscale approaches and neural networks. te Brake and Karssemeijer [23] compare several multiscale feature extraction methods against a single-scale method for detecting mammographic masses. They use a single neural network (five hidden units, trained to minimize the root-mean-square error using backpropagation) to simultaneously integrate the multiscale features. They note only minor improvements for the multiscale approach over the single-scale method, given that an optimal scale is chosen *a priori*. Aghdasi [24] uses a neural network (i.e., single neuron) to learn the optimal set of weights for integrating wavelet coefficients as features for microcalcification detection, an approach similar to that proposed by Yoshida *et al.* [25]. In all cases, integration is done using a single network trained with a squared error cost function.

C. Exploiting Context in Mammographic Image Analysis

Context can be defined as nearby or surrounding structure that establishes the meaning or identity of an object. In image analysis, contextual information is often used to detect and classify visual objects. For example, detecting a small building in an aerial image can be facilitated by searching along roads, since buildings tend to lie in close proximity to roads. Both human observers and computer vision systems (e.g., [26]) have been developed to exploit contextual relationships in imagery. The exploitation of context is consistent with the global-focal models of visual search proposed by Kundel [27]. Using eye tracking to monitor gaze, Kundel and colleagues have shown that human observers search radiographs by alternating between a global and local feature analysis, integrating both sets of features in their decision process.

In mammographic image analysis, context is exploited by radiologists and mammographers for detecting and identifying breast abnormalities. The clustering of calcifications, their proximity to ductal tissue, the architectural distortion surrounding potential lesions, are all contextual cues used by radiologists and mammographers [28]. The predictive value, as determined by radiologists, of both local and contextual (global) features for calcification and mass detection has been reported elsewhere [29].

Contextual relationships can be integrated into mammographic CAD systems, being made explicit, given known pathology, through incorporation of preset rules and/or feature detectors tuned to capture the context. Alternatively, contextual relationships can be learned from the data, allowing for more complicated and less obvious contextual cues to be uncovered by the pattern recognition system.

D. Overview of Hierarchical Pyramid/Neural Network Architecture

We have developed a pattern recognition architecture that learns contextual relationships between structure in images for detection and classification of objects. Fundamental to the architecture is the multiscale decomposition of an image, via pyramid transforms [33] and the subsequent integration of multiscale image features by a hierarchy of neural networks. These fundamental aspects of the architecture lead to the name *hierarchical pyramid neural networks (HPNN)*. Several variants of the HPNN can be defined, dependent upon the direction of processing in the hierarchy. Fig. 2 illustrates the general coarse-to-fine and fine-to-coarse architectures. These two

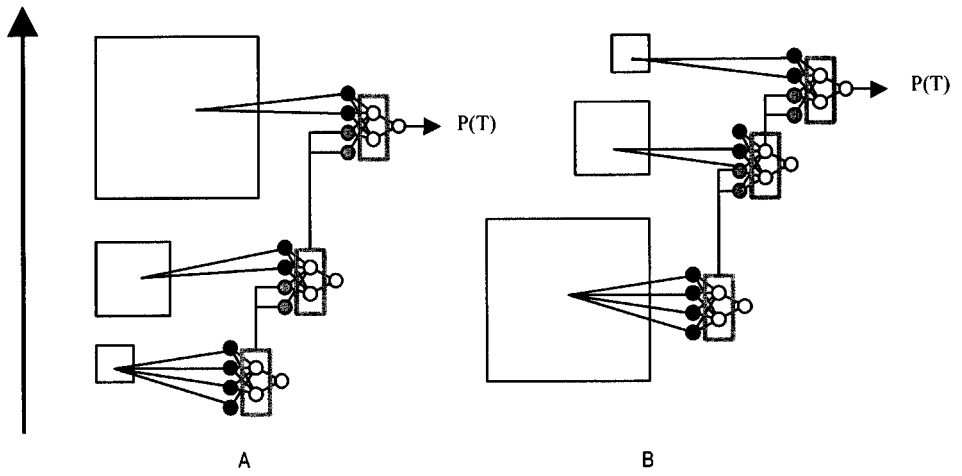

Fig. 2. Hierarchical pyramid/neural network architectures. (A) Coarse-to-fine and (B) fine-to-coarse. In (A), context is propagated from low to high resolution via the hidden units of low-resolution networks. In (B), small scale detail information is propagated from high- to low-resolution. In both cases, the output of the last integration network is an estimate of the probability that a target is present. Arrow shows direction of information flow.

architectures detect small or large target objects by exploiting coarse-scale (low resolution) or fine-scale (high-resolution) information associated with the target. For example, in the coarse-to-fine HPNN networks, operating at low resolution learn contextual features that are passed to networks operating at high resolution and integrated to detect the object of interest (i.e., the contextual inputs condition the probability of target present). For the fine-to-coarse HPNN architecture, networks extract detail structure at fine resolutions of the image and then pass this detail information to networks operating at coarser scales [see Fig. 2(B)]. For many types of objects, information about the fine detail structure is important for discrimination between different classes, i.e., fine resolution structure occurring within the context of the coarse resolution structure is indicative of an object class.

We have previously reported on how the HPNN architectures and learning algorithms can improve detection for a general class of image search/detection problems [30]–[32]. For example, we have shown that for the problem of detecting small buildings in aerial imagery, the coarse-to-fine HPNN architecture has higher accuracy than both conventional neural network architectures and standard statistical classification techniques [30]. In this paper, we present our results of applying the HPNN framework to two problems in mammographic CAD; detecting microcalcifications and masses in digital/digitized mammograms. The coarse-to-fine HPNN architecture is well suited for the microcalcification problem, while the fine-to-coarse HPNN is suited for mass detection. We evaluate the performance and utility of the HPNN framework by considering its effects on reducing FP rates in a well-characterized CAD system developed by The University of Chicago (UofC). In both cases (microcalcification and mass detection), the HPNN acts as a postprocessor of the UofC CAD system.

II. Methods

In this section, we describe four critical elements of the HPNN: 1) integrated feature pyramid representation; 2) neural network hierarchy; 3) the learning algorithm; and 4) training procedure.

A. Integrated Feature Pyramids

Image features are extracted and represented as integrated features pyramids (IFPs) [33]. Multiscale pyramid transforms are used to construct the IFP, which is the representation that serves as input into the neural network hierarchy. The pyramid transformation for the current set of experiments is based on a general class of filters that measure orientation energy and image intensity gradients.

For the coarse-to-fine IFP, steerable filters [34] are used to compute local oriented gradient information across scale. The steering properties of these filters enable the direct computation of the orientation having maximum energy. Features are constructed which represent, at each pixel location, the maximum energy (energy at orientation $\theta_{\max}$), the energy at the orientation perpendicular to $\theta_{\max}$ ($\theta_{\max} - 90°$), and the energy at the diagonal (energy at $\theta_{\max} - 45$). A pyramid decomposition using steerable oriented filters provides a rather generalized basis for capturing both coarse and fine scale structure predictive for clinically significant microcalcifications (e.g., calcification shape, branching, clustering, cluster shape [28], [29]). Fig. 3(A) illustrates the form of the IFP input into the coarse-to-fine network hierarchy.

The IFP for mass detection is slightly different from the coarse-to-fine IFP for microcalcification detection [Fig. 3(B)]. For mass detection, input to the fine-to-coarse neural network hierarchy is an IFP having radial and tangential gradient components at each resolution, relative to the mass center. The significance of radial and tangential features for mammographic mass detection has been reviewed by Giger *et al.* [8]. The features are filtered versions of the image, with filter kernels given by

$$\psi_{q,p}(r,\theta) = \left(\frac{q!}{\pi(q+|p|)!}\right)^{1/2} r^{|p|} e^{-r^2/2} L_q^{|p|}(r^2) e^{ip\phi} \qquad (1)$$

in polar coordinates, with $(q, p) \in \{(0, 1), (1, 0), (0, 2)\}$, and $L_q^{|p|}$ being the Laguerre polynomial [35]. The function $\psi_{q,p}$ is the normalized solution of Schroedinger's equation for the

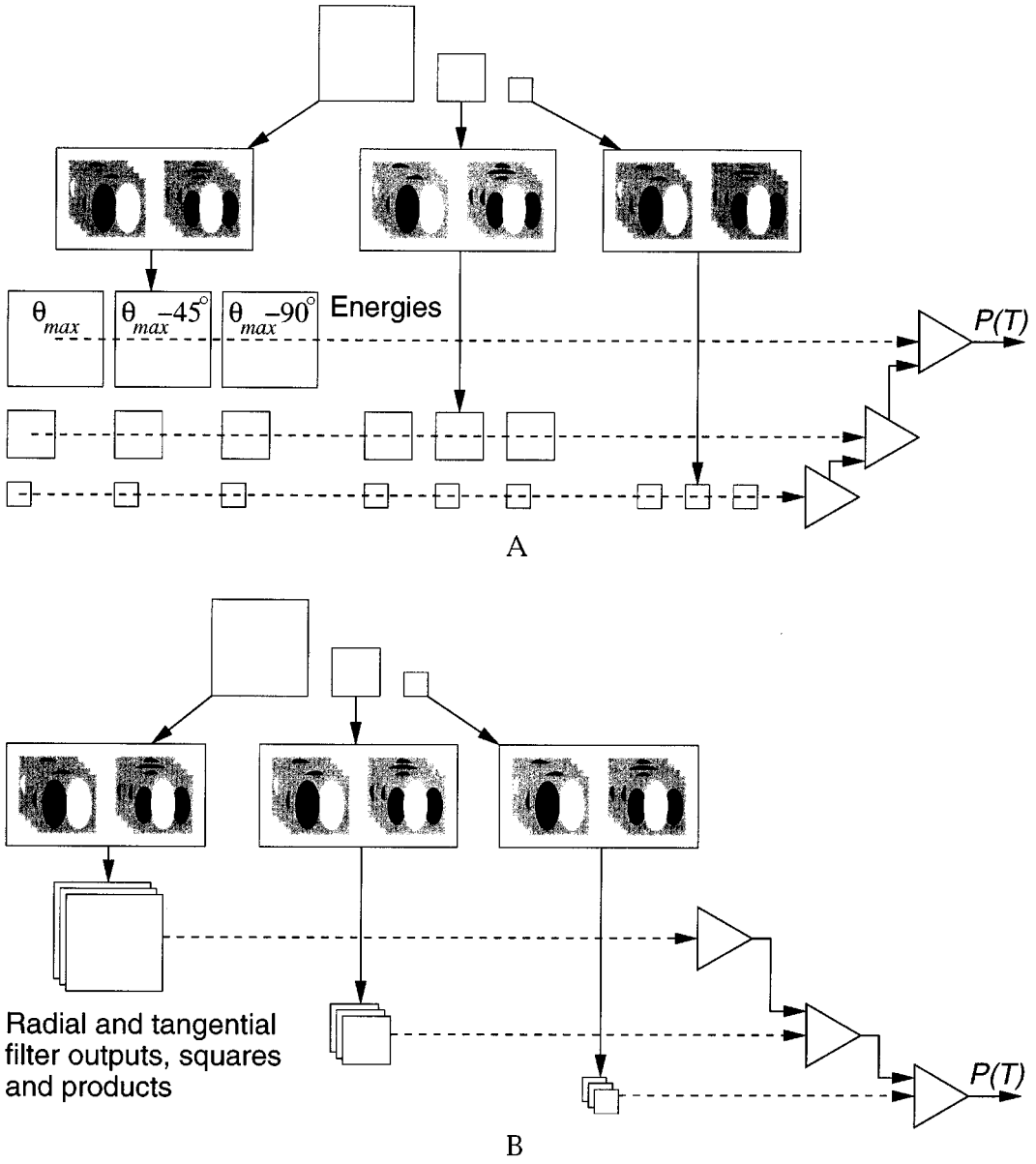

Fig. 3. IFPs for (A) coarse-to-fine and (B) fine-to-coarse HPNN.

two-dimensional (2-D) harmonic oscillator in polar coordinates. There are several motivations for this choice. First, the filtering function is steerable, since steering is simply a multiplication by a phase factor. Second, the set of functions for all q and p is complete, in that we can construct arbitrarily complex features, subject to the finite resolution of the image. Finally, because the harmonic oscillator also has solutions in Cartesian coordinates, we can express the steerable solutions in polar coordinates as sums of a small number of solutions that are separable in Cartesian coordinates, thereby ensuring computational efficiency. These filters are steered in the radial and tangential directions relative to the mass centers, using the real and imaginary components and their squares and products, as features.[1] Features were extracted at each level of the Gaussian pyramid representation of the mass ROI, and used as inputs to networks at the same level.

[1]The center coordinates of the masses are generated by earlier stages of the CAD system.

B. Neural Network Hierarchy

The neural networks in the HPNN are multilayer perceptrons, having one hidden layer with between four and eight hidden units. The number of hidden units is chosen via cross-validation [36]. All units in a network perform a weighted (w_i) sum of their inputs (x_i), subtracting an offset or threshold (θ) from that sum to get the activation (a)

$$a = \sum_i w_i x_i - \theta. \tag{2}$$

The activation is transformed into a unit's output, y, by passing it through a sigmoid function

$$y = \sigma(a) = \frac{1}{1+e^{-a}}. \tag{3}$$

Each network in the HPNN hierarchy receives input from the integrated feature pyramid and hidden unit input from networks lower in the hierarchy. Networks are trained either coarse-to-fine or fine-to-coarse, depending on the architecture. In the coarse-to-fine HPNN, the network lowest in the hierarchy is first trained until convergence and then all parameters in this network are held fixed while the next network on the hierarchy is trained. Coarse-to-fine training is possible because the positions of the small objects are well defined when the resolution is decreased. For the fine-to-coarse HPNN, extended objects do not have a definite location at high resolution. The entire hierarchy of networks is, therefore, trained as a single N-layered network (N being a function of the number of layers per network and the number of networks in the hierarchy).[2] Input for both training and testing is raster scanned into each network so that the output of a network at any level is an image. For both HPNN architectures, the output of the network is an image representing the probability that an object is present at each x, y position. For the coarse-to-fine architecture, each output pixel represents the probability of a point-like object (e.g., a microcalcification), while for the fine-to-coarse architecture each output pixel represents the probability that a large extended object (e.g., a mass) is within that low-resolution pixel.

C. Learning Algorithm

The conventional error function for training a neural network on a binary detection problem is the cross-entropy error function, which is the negative logarithm of the probability that the network produces detection decisions that agree with the targets in the training data. It is given by

$$E = -\sum_i [d_i \log y_i - (1 - d_i) \log(1 - y_i)] \tag{4}$$

where $d_i \in \{0, 1\}$ is the desired output and y_i is the actual output of the neuron, given by (3). For image-based detection, since networks are typically applied across a set of pixels, both y_i and d_i are a function of position; $y_i(x, y)$, $d_i(x, y)$. Thus, every position in an image is either associated with the presence, $d_i(x, y) = 1$, or absence, $d_i(x, y) = 0$, of a target.

In examining the truth data for the mammographic ROI datasets, we found that radiologists often make small errors in localizing individual microcalcifications and masses. For microcalcifications, these errors appear to be within ± 2 pixels of the correct position. For masses, the positional error also includes the extent of the mass—masses have ill-defined borders that are not easily ground-truthed, even by an expert. If the exact positions of the objects are unknown then the probability of detecting the objects at the correct positions cannot be evaluated and using (4) will result in poor performance, as will be illustrated below.

Consider instead the probability of detecting an object of interest when detection is defined as at least one pixel detected within a certain region known to contain the object. For a dataset with a coordinate vector for each object, let $\vec{x}_i$ represent the coordinates of the ith object.[3] Define a region P_i as set of pixel locations for the ith object that incorporate the known magnitude of the uncertainty or positional error in the truth data. A single detection within P_i will represent the detection of the ith object. Denote the output of the network when applied to the input vector derived from the neighborhood of $\vec{x}_i$ to be $y(\vec{x}_i)$. The probability of the network producing at least one detection in P_i is one minus the probability of producing no detection in P_i, or $1 - \prod_{\vec{x} \in P_i} (1 - y(\vec{x}))$. As with cross-entropy, the probability of not detecting an object at a negative position $\vec{x}_i$ is $1 - y(\vec{x}_i)$. If we define N as the set of all know negative locations then the new error function becomes

$$E_{\mathrm{UOP}} = -\sum_i \log \left(1 - \prod_{\vec{x} \in P_i} (1 - y(\vec{x})) \right) - \sum_{\vec{x}_i \in N} \log(1 - y(\vec{x})). \tag{5}$$

We call this the *Uncertain Object Position (UOP)* error function. The first term of (5) is the probability of detecting at least one pixel in a positive region while the second term is the probability of no detection in a negative region. The gradient of E_{UOP} with respect to the network weights is

$$\frac{\partial E_{\mathrm{UOP}}}{\partial w} = \sum_i \left\{ \frac{\prod_{\vec{x} \in P_i} (1 - y(\vec{x}))}{\prod_{\vec{x} \in P_i} (1 - y(\vec{x})) - 1} \sum_{\vec{x} \in P_i} \frac{\partial y(\vec{x}) / \partial w}{(1 - y(\vec{x}))} \right\} + \sum_{\vec{x} \in N} \frac{1}{1 - y(\vec{x})} \frac{\partial y(\vec{x})}{\partial w} \tag{6}$$

which is used in an optimization loop for training.

As an illustration of the utility of the UOP error function, we compare the detection performance, with a network trained using cross-entropy, for a "toy problem" as shown in Fig. 4. A 10×10 grid of single-pixel objects was embedded in a noisy background. Single-pixel objects were assigned a pixel value of one, while background pixels had a value of one-half or zero randomly assigned with equal probability. Errors were introduced into the truth data by randomly shifting the truth data within a 3×3 pixel neighborhood centered around the object's true position [see Fig. 5(B)]. A "network" consisting of a single sigmoidal neuron, with activation and transfer functions as in (2) and (3), was used to search the image for the objects. At a given location $\vec{x} = \{x, y\}$the inputs to the network are nine pixel values from a 3×3 window in the input image, centered on $\vec{x}$.

In Fig. 4, the truth image shows both the single-point truth data and the square 3×3 region around these pixels. The images in Fig. 4(D) and (E) are the outputs of the network trained using the cross-entropy error function. The cross-entropy-trained network with the output in Fig. 4(D) was trained using single-point truth data while the network with the output shown in Fig. 4(E) was trained using the 3×3 region truth data. Fig. 4(F) is the

[2]Error backpropagation through the pyramid reduction operations is straightforward, since this operation is linear.

[3]Note that for analysis of 2-D imagery, such as mammograms, $\vec{x}_i = \{x, y\}$. However, the formulation can be extended across an arbitrary coordinate space, so we use $\vec{x}_i$ for generality.

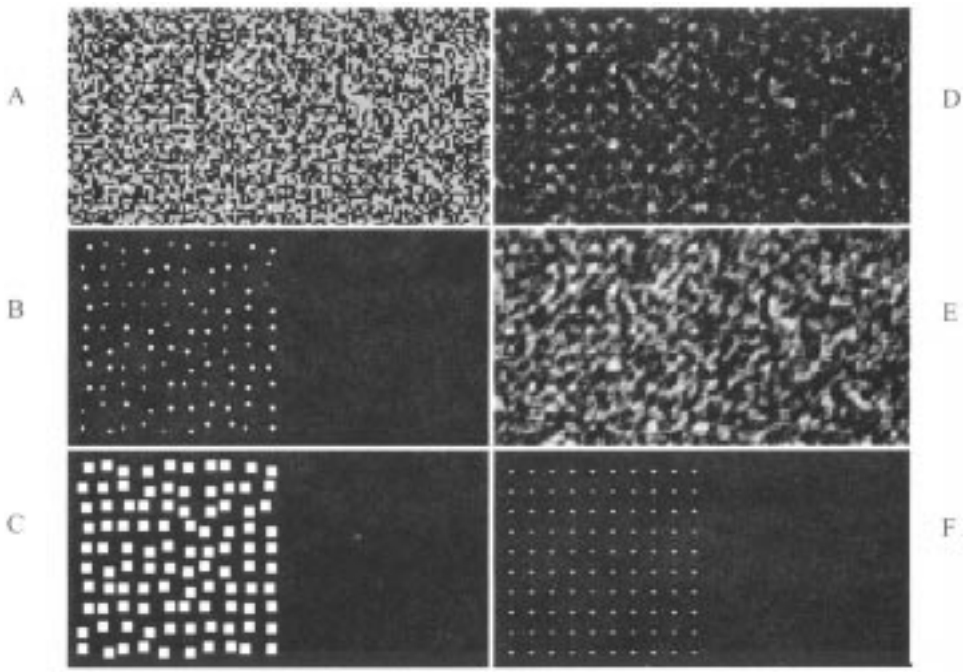

Fig. 4. "Toy problem" illustrating performance of UOP error function versus cross-entropy error. (A) Image consisting of 10 × 10 grid of white dots in a background of random binary noise. (B) Single-point truth data with positional error. (C) Truth data created by considering the magnitude of the positional error (±1 pixel results in 3 × 3 regions). (D) Output for network trained using cross-entropy error and truth data in B. (E) Output of network trained using cross-entropy error and truth-data in C. (F) Output of network trained using UOP error and truth data in C.

output of the network trained using the UOP error function with positive regions P_i as shown in Fig. 4(C). As is evident from the figure, the UOP trained network produces qualitatively superior results.

We measured, quantitatively, the detection performance of the networks by computing the sensitivity and FP rates on the data. For the cross-entropy-trained networks, sensitivity was 90% with a 7.5% pixel FP rate. For the UOP trained network, sensitivity was 100% with a 0% FP rate.

D. Training Procedure

All training was performed with the sequential quadratic programming (SQP) routine E04UCF of the commercial numerical Fortran subroutine library available from the Numerical Algorithms Group [37]. This routine optimizes a function of several variables using a quasi-Newton method, iteratively refining an approximation to the Hessian of the objective function as the search proceeds, and using the approximation to guide the search. For the HPNN, this results in a batch algorithm. The objective function is the summed error on the training set plus the regularization term. The SQP routine terminates when it determines that an acceptable approximation to a minimum has been reached. The HPNN parameters are initialized to uniformly distributed pseudorandom values between ±1.

To avoid "over-training," we used a "weight decay" regularization term

$$r = \frac{\lambda}{2} \sum_i w_i^2.$$

λ was adjusted to minimize the cross-validation error, computed by dividing the training data into disjoint subsets whose union is the entire set. The network was first trained on all of the training data, and then, starting from this set of weights, the network was retrained on the data with one of the subsets left out. The resulting network was tested on the "holdout" subset. This retraining and testing with a holdout set was repeated for each of the subsets, and the average of the errors on the subsets is the cross-validation error, an unbiased estimate of the average error on new data.

Using 68 ROIs and a network hierarchy having four networks, each with a single hidden layer with four hidden units, the time required for training was typically three to four days (50-MHz Sun Sparc 10). Note that this time also included training cross-validation sets to search for an optimal λ. With tenfold cross-validation, four pyramid levels in the coarse-to-fine HPNN, and approximately four values of λ searched for each network, this amounts to roughly 160 networks trained or retrained during the training period.

We also use cross-validation to optimize λ for the fine-to-coarse HPNN. This procedure requires one to three days (200 MHz Sun Ultra Sparc 2), depending on the complexity of the HPNN. Note that this is a much faster workstation than that used for the coarse-to-fine HPNN, so the fine-to-coarse HPNN is more expensive to train, even though the number of parameters is comparable. We speculate that there are two main reasons for this. First, the fine-to-coarse HPNN has to be trained as a single network, so we are searching a higher dimensional space. Second, the fine-to-coarse HPNN is effectively a feed-forward network with many layers. It is well known that the error signal attenuates as one backpropagates from the output layer through hidden layers, and the training problem is ill conditioned. Training is still effective, but it is somewhat slower.

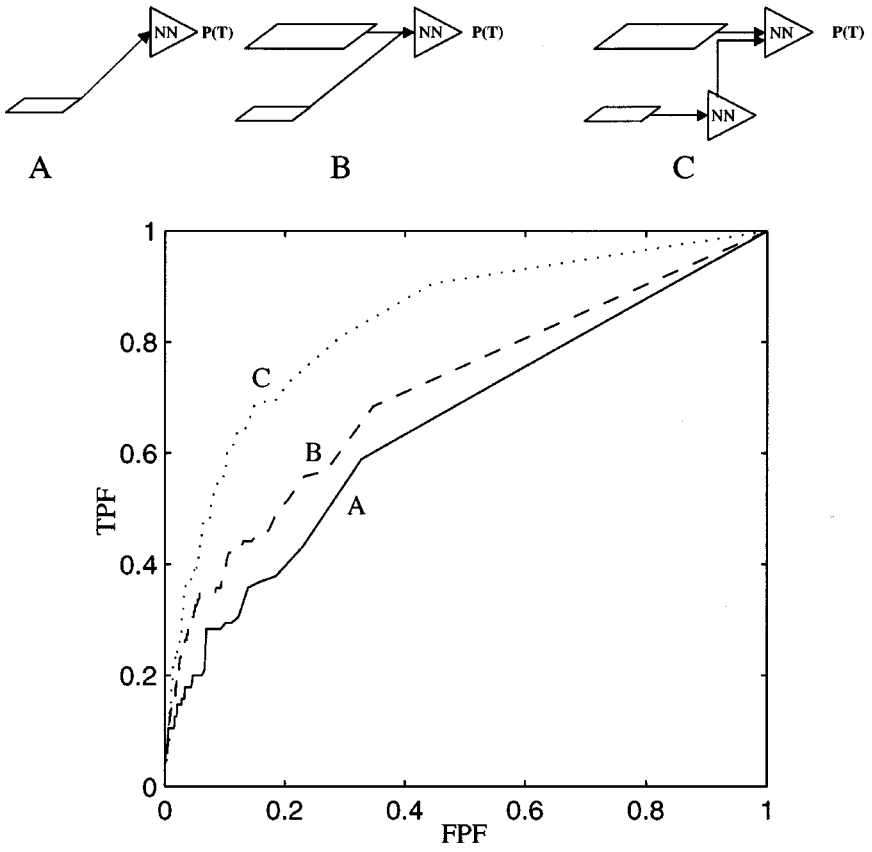

Fig. 5. Raw ROC curves for the three networks A, B, and C (HPNN).

Once trained, the HPNN can process images very quickly (roughly a second for a 100×100 pixel ROI on a 200-MHz Sun Ultra Sparc 2). This can be made even faster (factor of four) if the features for each ROI are precomputed by earlier stages of the CAD system.

III. Results

A. The Experimental Paradigm

We conducted a series of experiments to determine the utility of the HPNN architecture for mammographic CAD. The goal of the first set of experiments was to validate our hierarchical network architecture and learning algorithms for capturing contextual information and to demonstrate improved detection performance, relative to traditional neural network architectures. The second set of experiments focused on a quantitative and rigorous evaluation of the HPNN, in particular evaluation of two architectures for reducing the FP rate of the state-of-the-art CAD systems developed by UofC. Finally, as a demonstration of clinical utility, we integrated the HPNN with a UofC CAD system and evaluated its performance in a Reader Study.

B. Validation of the Network Hierarchy Architecture

Three neural network architectures were evaluated, each having one hidden layer with four to eight hidden units.[4] A two level coarse-to-fine IFP was constructed and used as input to the different network architectures. As shown in Fig. 5, network A consists of a single network processing data from the coarsest resolution of the IFP, network B is a single network receiving input from all levels of the IFP and network C is a two-level coarse-to-fine HPNN. The networks had activation and transfer functions described previously [(2) and (3)] and were trained using cross-entropy error (4).

We trained the networks on five mammograms. Each mammogram had one or two clusters with approximately 20 microcalcifications per mammogram, for a total of 97. The results given below were measured on five test mammograms with one cluster each, for a total of 95 microcalcifications.

Results for the three networks are shown as receiver operating characteristic (ROC) curves [38] in Fig. 5. Note the improvement as finer resolution information is added to the network (networks A versus B) and especially the very large improvement when using the hierarchical network architecture (networks A and B versus C). These results should be considered within the context of those reported by te Brake and Karssemeijer [23], who also noted only small improvements over single-scale features when integrated by a single neural network. Our results show that the hierarchical architecture of the HPNN is crucial for the integration of the multiscale features and, therefore, optimizing a CAD system's detection performance.

We considered whether network C was in fact taking advantage of context information by examining the representations developed by various hidden units in the network. Fig. 6 shows outputs of two classes of hidden units. The first class [Fig. 6(B)]

[4]Model complexity was controlled for by adding/subtracting hidden units using a cross-validation error.

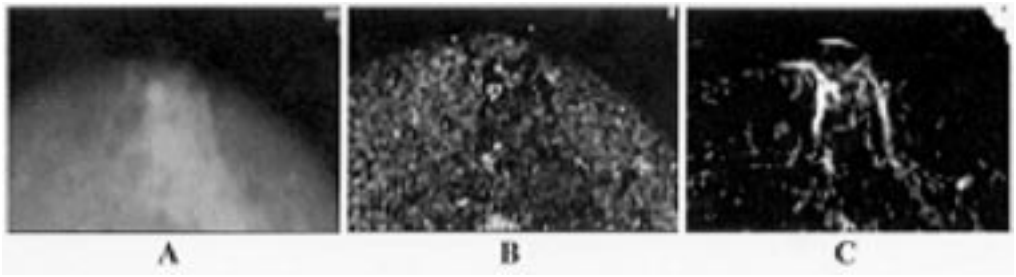

Fig. 6. (A) Original mammogram, (B) hidden unit representations for networks operating at high resolutions, (C) hidden unit representations for networks operating at low resolutions. Radiologists have noted that some of the structure in C appears to correlate with specific anatomy in the breast (e.g., ducts and/or blood vessels), indicating that these hidden units may represent contextual information.

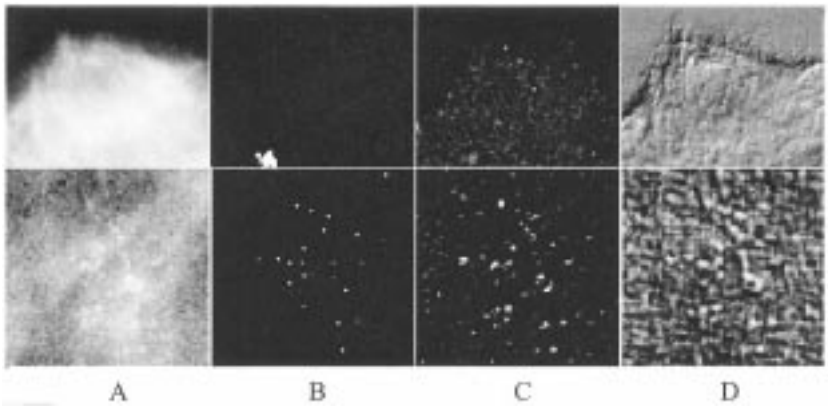

Fig. 7. Detecting microcalcifications using UOP error function. The upper row contains reduced resolution images from one full size test mammogram. The lower row shows a region of interest at full resolution. (A) Image, (B) truth data, (C) output of UOP trained network, (D) output of cross-entropy-trained network.

appears to represent point-like structure, similar to the structure of an individual microcalcification. The second class of hidden unit [Fig. 6(C)] has a different representation. In this case, the unit is selective for long, extended, and oriented structure. When shown to radiologists, they noted that this hidden unit structure appeared correlated with the ductal and vascular anatomy. As mentioned previously, the development of breast cancer is often correlated with these anatomical structures. Results for this experiment suggest that the coarse-to-fine hierarchical neural network is able to automatically extract information that is consistent with known contextual relationships and that this may result in the observed improvement in detection performance. These results are also consistent with our previous work showing that hidden units in the HPNN learn representations of image context for object detection (e.g., learning representations of roadways for building detection in aerial imagery [30]).

C. Validation of UOP for Microcalcification Detection

To validate the utility of our UOP error function (5) for mammographic CAD we conducted experiments comparing detection performance with the cross-entropy error function (4). We trained and tested a single neuron network to detect microcalcifications, using the dataset described in the previous experiment. Expert radiologists constructed the truth data, however, inspection of the data indicated positional errors of up to two pixels. At a given location $\vec{x}$, the inputs to the network were the 25 pixel values in a 5×5 window in the input, centered on $\vec{x}$. We expect that the average local brightness is not related to the detection problem. Therefore, to enforce invariance to average local brightness we constrained the weights of the single unit network to sum to one.

Fig. 7 shows results for a test mammogram. Note that the network trained using UOP generates fewer FPs than the conventional cross-entropy error function. If thresholds are applied to the networks so that 50% of the true positives (TPs) are detected, the UOP trained network has 50% fewer FPs that the cross-entropy network.

D. Results on Research Database: Microcalcification Detection

Given results for the previous two experiments we next evaluated the performance of an HPNN architecture trained using the UOP error. In the remaining experiments described in this paper, we evaluated the performance of the HPNN as a post processor or adjunct for the UofC CAD system.

UofC provided data used for the microcalcification experiments. The first set of data consists of 50 TP and 86 FP ROIs. These ROIs are 99×99 pixels and digitized at 100-μm resolution. A second set of data from the UofC clinical testing database included 47 TPs and 103 FPs, also 99×99 and sampled at 100-μm resolution.

We trained a coarse-to-fine HPNN [Fig. 2(A)], using UOP error function, to detect individual microcalcifications. Training and testing were done using a jackknife protocol [39], whereby one half of the data (25 TPs and 43 FPs)] was used for training and the other half for testing. Results were compiled for five different random splits of the data. For a given ROI, the probability map produced by the network was thresholded at a given value

TABLE I
COMPARISON OF HPNN AND SIANN NETWORKS

cc	HPNN				SIANN			
	A_z	σ_{Az}	FPF TPF=1.0	σ_{FPF}	A_z	σ_{Az}	FPF TPF=1.0	σ_{FPF}
1	.93	.03	.24	.11	.88	.04	.50	.11
2	.94	.02	.21	.11	.91	.02	.43	.10
3	.94	.03	.39	.19	.91	.03	.48	.19
4	.93	.03	.48	.15	.90	.05	.56	.21
5	.93	.03	.51	.06	.88	.05	.68	.21

to produce a binary detection map. Region growing was used to count the number of distinct detected regions. The ROI was classified as a positive if the number of regions was greater than or equal to a given cluster criterion.

Table I compares ROC results for the HPNN and the shift-invariant artificial neural network (SIANN) network that had been used in the UofC CAD system [15]. Reported are the area under the ROC curve (A_z), the standard deviation of A_z across the subsets of the jackknife (σ_{Az}), the FP fraction at a TP fraction of 1.0 ($\mathrm{FPF@TPF} = 1.0$) and the standard deviation of the FPF across the subsets of the jackknife (σ_{FPF}). A_z and $\mathrm{FPF@TPF} = 1.0$ represent the averages of the subsets of the jackknife. Note that both networks operate best when the cluster criterion (cc) is set to two. For this case, the HPNN has a higher A_z than the SIANN network while also halving the FP rate. This difference, between the two networks' A_z and FPF values, is statistically significant (z test: $p_{Az} = 0.0018$ and $p_{\mathrm{FPF}} = 0.000\,01$).

The second set of data was tested using a coarse-to-fine HPNN trained on the first dataset. 150 ROIs taken from a clinical study and classified as positive by the full UofC CAD system for microcalcification detection (including the SIANN neural network) were used to test the HPNN. Though the UofC CAD system classified all 150 ROIs as positive, only 47 were in fact positive while 103 were negatives—this dataset was overpopulated with FPs. We applied the HPNN trained on the entire previous data set to this new set of ROIs. The HPNN was able to reclassify 47/103 negatives as negative, without loss in sensitivity, i.e., no false negatives were introduced.

On examining the negative examples rejected by the coarse-to-fine HPNN, we found that many of these ROIs contained linear, high-contrast structure that would otherwise be FPs for the SIANN network (see Fig. 8). One possible reason for this is that the coarse-to-fine HPNN also learns context for the FPs. SIANN presumably interprets the "peaks" on the linear structure as calcifications. However, because the coarse-to-fine HPNN also integrates information from low resolution it can associate these "peaks" with linear structure at low-resolution and, thus, determine that these peaks are not microcalcifications. This is an interesting difference from our earlier results, in which the networks appeared to learn contextual relationships associated with positive examples—ductal and vascular anatomy. Thus, it appears that the HPNN can exploit contextual relationships to both detect TPs and eliminate FPs.

E. Results on a Research Data Base: Mass Detection

The next set of experiments applied a fine-to-coarse HPNN architecture to detect masses in digitized mammograms. Radiologists often distinguish malignant from benign masses based on the detailed shape of the mass border and the presence of spicules along the border [28]. We evaluate the fine-to-coarse HPNN, Fig. 2(B), for its ability to integrate high-resolution information within the context of coarse-scale mass structure.

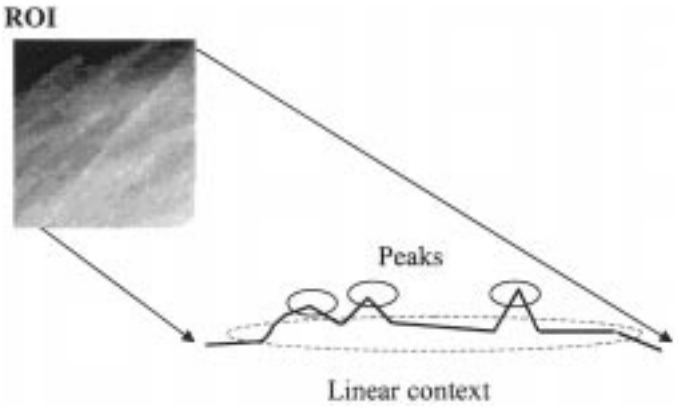

Fig. 8. Typical negative ROI that was eliminated by the coarse-to-fine HPNN for microcalcification detection. The HPNN is able to associate the intensity peaks, which in isolation may be interpreted as microcalcfications, with the coarse-scale linear structure in order to classify the ROI as a negative.

TABLE II
SENSITIVITY AND SPECIFICITY FOR FINE-TO-COARSE HPNN FOR MASS DETECTION

Sensitivity	Specificity
100%	51%
95%	57%
90%	67%
80%	79%

The experimental paradigm is similar to the microcalcification experiments in that we apply the HPNN as a postprocessor to the UofC CAD system for mass detection. The data in our study consists of 72 positive and 100 negative ROIs. The negative ROIs are FPs of the earlier stages of the CAD system. These are 256×256 pixels and are sampled at 200-μm resolution.

Results for the fine-to-coarse HPNN system are shown in Table II. The A_z value on the test set was 0.85. These results show a 51% reduction in FP rate of the UofC mass-detection system without loss in sensitivity.

F. Results in Clinical Evaluation

As a final test of the utility of the HPNN architecture a clinical reader study was conducted to evaluate the performance of the combined HPNN/UofC system for microcalcification

TABLE III
SUMMARY OF READER STUDY PROTOCOL

899 cases (4 standard views, original mammograms)
• *501 normals (including 10 atypia)*
• *199 benign*
• *199 malignant (58 DCIS+141 invasive) (22%)*
two reading conditions:
• *film only*
• *film + computer results*
• **films were mounted on alternators**
• **computer results were shown on CRT monitors**
standard observer study protocol
• *training session randomized reading order, etc.*
10 readers:
• *5 specialists (>50% breast imaging)*
• *5 general radiologists (MQSA certified)*

TABLE IV
FP RATES OF CAD SYSTEM

CAD Program	Number false positives per image (at fixed sensitivity)
Mass detection	1.6
Microcalc detection (no HPNN)	1.04
Microcalc detection (with HPNN)	0.88

detection.[5] A coarse-to-fine HPNN was integrated as the last stage of processing in the UofC CAD system for microcalcification. The HPNN was trained on a set of TPs and FPs generated by the UofC system. After training the parameters of the HPNN were fixed. Integration of the HPNN with the UofC system was done via a simple UNIX script for exchanging files containing the ROI data. Additional details of the reader study have been described previously [40]. In this paper, we summarize the results.

Table III outlines the protocol. Approximately 900 retrospective mammographic cases were collected and read by ten readers. Five readers were considered experts in mammography (spent over 50% of their time reading mammograms) and the other five were general radiologists who were MQSA certified [41]. Films were read in two conditions; film only (unaided) or film + computer results (aided).

Results of the computer output alone are shown in Table IV. Note that on this new dataset the HPNN continues to reduce the FP rate of the microcalcification CAD system.

The clinical utility of the complete system, which includes the CAD systems for mass detection and the HPNN-enhanced system for microcalcification detection, is shown in Table V, comparing reader performance with and without the computer aid. Expert readers showed a statistically significant improvement when using the CAD system, however, the improvement was not statistically significant for the general radiologists. One possible reason is that FPs continue to be an issue, since experts are better than general radiologists at negating or ignoring these FPs. Additional analysis is required to understand the difference between the two groups. However, the overall results show that the CAD system, which included the HPNN, can potentially improve performance of mammographic screening, in this case for more experienced radiologists.

[5]In this clinical evaluation, only the coarse-to-fine HPNN for microcalcification was integrated with the UofC CAD and evaluated.

IV. DISCUSSION

In this paper, we have demonstrated coarse-to-fine and fine-to-coarse HPNN architectures that learn contextual relationships for detecting microcalcifications and masses in digital/digitized mammograms. Though the architectures are novel, they bear some resemblance to previous network architectures. For example, the fine-to-coarse HPNN is similar to the convolution network proposed by Le Cun, [42] (which has been applied to mammographic image analysis by [14] and [43]), however, with a few notable differences. The fine-to-coarse HPNN receives as inputs preset features extracted from the image (in this case radial and tangential gradients) at each resolution, compared to the convolution network, whose inputs are the original pixel values at the highest resolution. Secondly, in the fine-to-coarse HPNN, the inputs to a hidden unit at a particular position are the pixel values at that position in each of the feature images, one pixel value per feature image. Thus, the HPNN's hidden units do not learn linear filters, except as linear combinations of the filters used to form the features. Finally, the fine-to-coarse HPNN is also trained using the UOP error function, which is not used in the convolution network.

The two architectures we have described can be combined into a more general architecture that integrates information both coarse-to-fine and fine-to-coarse. This bi-directional integration, shown in the architecture of Fig. 9, is attractive in that most objects can be considered to have a "natural scale"—typically some measure of their size. Classification of the object might be improved through integration of finer and coarser resolution information, relative to this natural scale. Since size can vary within a class of objects, it may be worthwhile to include outputs at more than one level of the HPNN. In this case, the UOP error (5) needs to be modified to include uncertainty over scale, but this is easily accomplished by changing the product to range over positions at all output levels. We can further generalize the architecture by adding connections between the fine-to-coarse and coarse-to-fine paths, but one must be careful to avoid loops when deciding where these connections should be added. We are currently investigating the application of this generalized HPNN architecture to mass detection.

Most of our results were reported relative to the UofC CAD mammographic systems, since they are considered to be well characterized and state-of-the-art. UofC is continuing to improve upon their systems and our current results are only meant as a comparison to a given standard at a given point in time. An issue in CAD research is the need for the development of appropriate benchmarks for comparing different algorithms. Several datasets are being developed

TABLE V
READER STUDY RESULTS USING CAD SYSTEM

Reader	Specialists		General Radiologists	
	Unaided	Aided	Unaided	Aided
1	0.851	0.878	0.813	0.824
2	0.891	0.911	0.862	0.876
3	0.878	0.898	0.881	0.888
4	0.911	0.914	0.876	0.863
5	0.884	0.903	0.899	0.892
avg	0.883	0.901	0.866	0.869
p value	0.01		0.19	

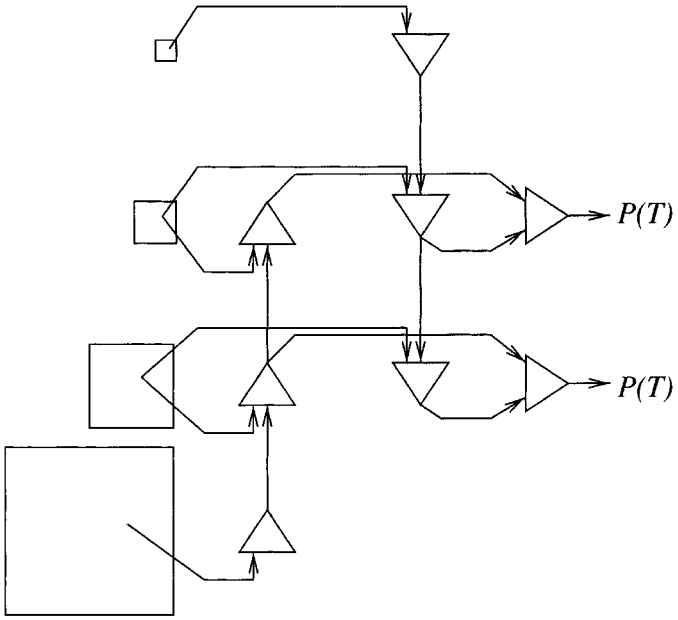

Fig. 9. Generalized HPNN architecture. Integration is bi-directional with output networks at the natural scale of the object. The natural scale may be known *a priori* or it can be searched for by optimizing over several output networks (e.g., search for the best one over the two output networks shown above).

which might eventually support such comparisons though they have yet to be widely accepted.[6]

V. CONCLUSION

We have presented the application of hierarchical pyramid neural network architectures to two problems in CAD; the detection of microcalcifications in mammograms and the direct detection of masses in mammograms. In the case of microcalcifications, the coarse-to-fine HPNN architecture successfully discovered large-scale context information that improves the system's performance in detecting small objects. We have demonstrated the performance of the HPNN framework by considering its utility as a postprocessor for a state-of-the-art CAD system. In addition, clinical utility has been demonstrated for a coarse-to-fine HPNN, which has been directly integrated with the UofC CAD system for microcalcification detection, with the complete system having been tested in a reader study. In the case of mass detection, a fine-to-coarse HPNN architecture was used to exploit information from fine resolution detail in order to eliminate FPs. One of the unique aspects of our approach which differentiates it from those that were previously discussed, is that it provides a single, unified framework for detection of both mammographic masses and microcalcifications. In general, we have found that the HPNN is a useful class of network architecture for exploiting context and integrating information at multiple scales for a variety of image analysis problems.

[6]Databases include the Digital Database for Screening Mammography (DDSM), Mammographic Image Analysis Society (MIAS) database, and Lawrence Livermore National Laboratories (LLNL)/University of California at San Francisco (UCSF) database. Information on these and other databases can be obtained from The Digital Mammography Home Page http://www.rose.brandeis.edu/users/mammo/digital.html.

ACKNOWLEDGMENT

The authors would like to thank Dr. R. Nishikawa and Dr. M. Giger of The University of Chicago for fruitful collaborations and discussions as well as providing the mammographic data.

REFERENCES

[1] K. Doi, M. L. Giger, R. M. Nishikawa, K. Hoffmann, H. MacMahon, R. A. Schmidt, and K. G. Chua, "Digital radiography: A useful clinical tool for computer-aided diagnosis by quantitative analysis of radiographic images," *Acta Radiologica*, vol. 34, pp. 426–439, 1993.
[2] R. E. Bird, "Professional quality assurance for mammography screening programs," *Radiology*, vol. 177, pp. 8–10, 1990.
[3] C. E. Metz and J. H. Shen, "Gains in accuracy from replicated readings of diagnostic images: Prediction and assessment in terms of ROC analysis," *Med. Decision Making*, vol. 12, pp. 60–75, 1992.
[4] E. L. Thurfjell, K. A. Lernevall, and A. S. Taube, "Benefit of independent double reading in a population-based mammography screening program," *Radiology*, vol. 191, pp. 241–244, 1994.
[5] "R2 Technology Pre-market approval (PMA) of the M1000 Image Checker,", U.S. Food and Drug Administration (FDA) application #P970058, approved, June 26, 1998.
[6] L. J. Burhenne, S. A. Wood, C. J. D'Orsi, S. A. Feig, D. B. Kopans, K. F. O'Shaughnessy, E. A. Sickles, L. Tabar, C. J. Vyborny, and R. A. Castellino, "Potential contribution of computer-aided detection to the sensitivity of screening mammography," *Radiology*, vol. 215, pp. 554–562, 2000.
[7] R. M. Nishikawa, R. A. Schmidt, R. B. Osnis, M. L. Giger, K. Doi, and D. E. Wolverton, "Two-year evaluation of a prototype clinical mammographic workstation for computer-aided diagnosis," *Radiology*, vol. 201, no. P, p. 256, 1996.
[8] M. L. Giger, Z. Huo, M. A. Kupinski, and C. J. Vyborny, "Computer-aided diagnosis in mammography," in *Handbook of Medical Imaging; Volume 2. Medical Image Processing and Analysis*, M. Sonka and J. M. Fitzpatrick, Eds. Bellingham, WA: SPIE, 2000, pp. 917–986.
[9] Z. Huo, M. L. Giger, C. J. Vyborny, D. E. Wolverton, R. A. Schmidt, and K. Doi, "Automated computerized classification of malignant and benign mass lesions on digital mammograms," *Acad. Radiol.*, vol. 5, pp. 155–168, 1998.
[10] C. E. Floyd, J. Y. Lo, A. J. Yun, D. C. Sullivan, and P. J. Kornguth, "Prediction of breast cancer malignancy using an artificial neural network," *Cancer*, vol. 74, pp. 2944–2948, 1994.

[11] Y. Jiang, R. M. Nishikawa, D. E. Wolverton, C. E. Metz, M. L. Giger, R. A. Schmidt, and K. Doi, "Automated feature analysis and classification of malignant and benign microcalcifications," *Radiology*, vol. 198, pp. 671–678, 1996.

[12] H. P Chan, B. Sahiner, K. L. Lam, N. Petrick, M. A. Helvie, M. M. Goodsitt, and D. D. Adler, "Computerized analysis of mammographic microcalcifications in morphological and feature spaces," *Med. Phys.*, vol. 25, pp. 2007–2019, 1998.

[13] J. Y. Lo, J. Kim, J. A. Baker, and C. E. Floyd, "Computer-aided diagnosis of mammography using an artificial neural network: Predicting the invasiveness of breast cancers from image features," in *Proc. SPIE Medical Imaging 1996: Image Processing*, vol. 2710, M. H. Loew, Ed., 1996, pp. 725–732.

[14] S. C. Lo, H. P. Chan, J. S. Lin, H. Li, M. T. Freedman, and S. K. Mun, "Artificial convolution neural network for medical image pattern recognition," *Neural Networks*, vol. 8, no. 7/8, pp. 1201–1214, 1995.

[15] W. Zhang, K. Doi, M. L. Giger, Y. Wu, R. M. Nishikawa, and R. Schmidt, "Computerized detection of clustered microcalcifications in digital mammograms using a shift-invariant artificial neural network," *Med. Phys.*, vol. 21, no. 4, pp. 517–524, 1994.

[16] K. S. Woods, J. L. Solka, and C. E. Priebe *et al.*, "Comparative evaluation of pattern recognition techniques for detection of microcalcifications," presented at the IS&T/SPIE, 1993.

[17] B. Zheng, W. Qian, and L. P. Clarke, "Digital mammography: Mixed feature neural network with spectral-entropy decision for detection of microcalcifications," *IEEE Trans. Med. Imag.*, vol. 15, pp. 589–597, Oct. 1996.

[18] D. Brzakovic and M. Neskovic, "Mammogram screening using multiresolution-based image segmentation," in *State of the Art in Digital Mammographic Image Analysis*. ser. Machine Perception and Artificial Intelligence, K. W. Bowyer and S. M. Astley, Eds. New York: World Scientific, 1994, vol. 9, pp. 103–127.

[19] S. L. Ng and W. F. Bischof, "Automated detection and classification of breast tumors," *Comput. Biomed. Res.*, vol. 25, pp. 218–237, 1992.

[20] L. Miller and N. Ramsey, "The detection of malignant mass by nonlinear multiscale analysis," in *Digital Mammography*, K. Doi, M. L. Giger, R. M. Nishikawa, and R. A. Schmidt, Eds. Amsterdam, The Netherlands: Elsevier, 1996, pp. 335–340.

[21] L. Li, W. Qian, and L. P. Clarke, "Digital mammography: Computer-assisted diagnosis method for mass detection with multiorientation and multiresolution wavelet transforms," *Acad. Rad.*, vol. 11, no. 4, pp. 724–731, 1997.

[22] T. Netsch and H. O. Peitgen, "Scale-space signatures for the detection of clustered microcalcifications in digital mammograms," *IEEE Trans. Med. Imag.*, vol. 18, pp. 774–786, Sept. 1999.

[23] G. M. te Brake and N. Karssemeiijer, "Single and multiscale detection of masses in digital mammograms," *IEEE Trans. Med. Imag.*, vol. 18, pp. 628–639, July 1999.

[24] F. Aghdasi, "Application of neural network adaptive wavelets for signal representation and classification in digital mammography," in *Digital Mammography*, K. Doi, M. L. Giger, R. M. Nishikawa, and R. A. Schmidt, Eds. Amsterdam, The Netherlands: Elsevier, 1996, pp. 307–310.

[25] H. Yoshida, W. Zhang, W. Cai, K. Doi, R. M. Nishikawa, and M. L. Giger, "Optimizing wavelet transform based on supervised learning for detection of microcalcifications in digital mammograms," in *Proc. Int. Conf. Image Processing*, vol. 3, 1995, pp. 152–155.

[26] "Proc. IEEE Workshop Context-based Vision," J. Mundy and T. Strat, organizers, Cambridge, MA, 1995..

[27] H. L. Kundel, "Visual search in medical images," in *Handbook of Medical Imaging; Volume 1. Physics and Psychophysics*, J. Beutel, H. L. Kundel, and R. L. Van Metter, Eds. Bellingham, WA: SPIE, 2000, pp. 837–858.

[28] D. B. Kopans, *Breast Imaging*. Baltimore, MD: Lippincott Williams & WIlkins, 1989.

[29] P. Whatmough, A. G. Gale, and A. R. M. Wilson, "Do radiologists agree on the importance of mammographic features," in *Digital Mammography*, K. Doi, M. L. Giger, R. M. Nishikawa, and R. A. Schmidt, Eds. Amsterdam, The Netherlands: Elsevier, 1996, pp. 111–116.

[30] P. Sajda, C. D. Spence, S. Hsu, and J. C. Pearson, "Integrating neural networks with image pyramids to learn target context," *Neural Networks*, vol. 8, no. 7/8, pp. 1143–1152, 1995.

[31] C. Spence, P. Sajda, S. Hsu, and J. Pearson, "Neural network/pyramid architectures that learn target context," in *Proc. DARPA Image Understanding Workshop*, 1994, pp. 853–862.

[32] C. Spence, "Supervised learning of detection and classification tasks with uncertain training data," in *Proc. ARPA Image Understanding Workshop*, 1996, pp. 1395–1402.

[33] P. J. Burt, "Smart sensing within a pyramid vision machine," *Proc. IEEE*, vol. 76, pp. 1006–1015, Aug. 1988.

[34] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 891–906, Sept. 1991.

[35] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, ser. National Bureau of Standards Applied Mathematics, 55. New York: Dover, 1972.

[36] C. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Oxford Univ., 1995.

[37] NAG Inc., 1400 Opus Place, Suite200, Downers Grove, IL 60515-5702. (1993) NAG Fortran Library Manual, Mark 16. Routine E04UCF. [Online]. Available: http://www.nag.com.

[38] C. Metz, "Current problems in ROC analysis," in *Proc. Chest Imaging Conf.*, Madison, WI, Nov. 1988, pp. 315–333.

[39] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. New York: Academic, 1990.

[40] R. M. Nishikawa, C. Gatsonis, M. D. Schnall, M. L. Giger, P. Sajda, and M. Chen, "Large scale observer study to measure the benefits of computer-aided detection to screening mammography," *Radiology*, vol. 213, no. P, p. 150, 1999.

[41] *Mammography Quality Standards Act (MQSA) of 1992*.

[42] Y. Le Cun, B. Boser, J. S. Denker, and D. Henderson, "Handwritten digit recognition with a back-propagation network," in *Advances in Neural Information Processing Systems 2*, D. S. Touretzky, Ed. San Mateo, CA: Morgan Kaufmann, 1990, pp. 396–404.

[43] A. Strauss, C. Monrocq, T. P. Baum, and M. Le Gal, "Selection of cluster candidates before clustered microcalcification detection: A modular and hierarchical approach based on neural networks," in *Digital Mammography*, K. Doi, M. L. Giger, R. M. Nishikawa, and R. A. Schmidt, Eds. Amsterdam, The Netherlands: Elsevier, 1996, pp. 297–300.