Physiologically Motivated Image Fusion for Object Detection using a Pulse Coupled Neural Network

Randy P. Broussard, Steven K. Rogers, Fellow, IEEE, Mark E. Oxley, and Gregory L. Tarr

Abstract—This paper presents the first physiologically motivated pulse coupled neural network (PCNN)-based image fusion network for object detection. Primate vision processing principles, such as expectation driven filtering, state dependent modulation, temporal synchronization, and multiple processing paths are applied to create a physiologically motivated image fusion network. PCNN’s are used to fuse the results of several object detection techniques to improve object detection accuracy. Image processing techniques (wavelets, morphological, etc.) are used to extract target features and PCNN’s are used to focus attention by segmenting and fusing the information. The object detection property of the resulting image fusion network is demonstrated on mammograms and Forward Looking Infrared Radar (FLIR) images. The network removed 94% of the false detections without removing any true detections in the FLIR images and removed 46% of the false detections while removing only 7% of the true detections in the mammograms. The model exceeded the accuracy obtained by any individual filtering methods or by logical ANDing the individual object detection technique results.

Index Terms—Automatics target recognition, breast cancer, CAD, CADx, computer-aided diagnosis, image fusion, neural networks, object detection, pulse coupled network, segmentation, wavelets.

I. INTRODUCTION

DIGITAL image processing is being investigated for object detection in applications such as breast cancer detection, and automatic target recognition [1]–[6]. Image processing is used to reduce unwanted information from an image with the hope that the improved signal-to-noise ratio will allow a pattern recognition process to detect and possibly identify the desired object. In general, no single image processing technique can be selective to all patterns for a given object, and still perform well at removing the many possible variations of unwanted information. Often, several techniques are used and the results are combined.

To perform this information fusion, primate vision processing principles are used to design a pulse coupled neural network (PCNN) based image fusion network for the purpose of improved object detection. Observed biological phenomenon such as temporal synchronization and state-dependent modulation are theorized as methods the vision system uses to combine the information and focus attention on an object [7]–[9]. The role these biological phenomena perform in information fusion and in the image fusion network is discussed. The PCNN is chosen as an architecture for the fusion network because it performs information linking at the neuronal pulse level. Through a combination of image segmentation, information fusion, and attention focus an object detection property emerges from the PCNN fusion network. Last, actual infrared and mammographic images are used to demonstrate the object detection accuracy of the network.

II. A BIOLOGICAL FOUNDATION FOR A FUSION NETWORK

Despite the enormous complexity of the primate cortical visual system, studies suggest it can be modeled by two basic hierarchical pathways, the parvocellular pathway and the magnocellular pathway [10]. The former pathway predominantly processes color information, and the later processes form and motion.

A. A Simplified Model of the Primate Vision System

Fig. 1 shows a model of these two pathways. The boxes refer to areas of the visual cortex which are believed to process distinct types of visual information [11]. The ovals denote the specific type of information processed within each area. The visual areas are almost fully connected in each direction of flow, but for clarity, only the stronger forward connections are shown. The entry point of an image into the model is the retina. The area marked LGN models the biological lateral geniculate nucleus. The areas of the model labeled with names starting with the letter V model specific areas in the human visual cortex. Each of these areas is believed to maintain one or more processed, topographically correct images of the light pattern that falls upon the retina. Areas V3, V4, and V5 are called specialty areas and it is believed that they process only selective information such as form, color, and motion, respectively.

As can be seen in Fig. 1, the information contained within a visual image is separated into various visual features. The specialty areas are of particular interest because they represent the final stages of the visual cortex and the visual features they produce are still separate. There is no known single place in the brain where these features (color, form, and motion) are brought back together and combined. Many current theories propose that the neuronal pulses that transport these features synchronize in a way which combines the information to
Fig. 1. Forward information flow of the visual system model.

represent the original object [7], [8], [12], [13]. These theories are used to design an image fusion network that segments an object, combines features, and isolates the object from the rest of the image.

B. Temporal Synchronization Provides Object Segmentation and Fusion

In 1987, stimulus-related neural oscillations were discovered in the primary visual cortex of monkeys and cats [7], [8]. These findings together with theoretical proposals (e.g., Grossberg 1983; Reitboek 1983, 1989; von der Malsburg 1985; Damasio 1989) support the hypothesis that neuronal pulse synchronization might be a mechanism that links local visual features into coherent global percepts. Synchronization is believed to be produced via a self-organizing process among local neural oscillations that are mutually connected. It is believed that these synchronizations mainly support the formation of more complex, “attentive percepts” that require iterative interactions among different processing levels and memory [8]. Visual segments that are related in some fashion will synchronize and pulse in unison. These synchronized segments represent objects, or segments of objects within a visual scene. This segmentation provides objects through which dissimilar features can be associated. Gray and Singer theorize that this association is performed by temporal synchronization [7]. Through this synchronization, the visual image is represented as an ensemble of synchronously pulsing objects. This biological principle is used by the PCNN fusion network to segment an image and fuse features into those segments. The result is a single segmented image with associated features superimposed. For object detection and recognition to be performed, a method is needed to select and extract a particular segment from the resulting image. This means the PCNN fusion network needs a method of focusing attention on individual groups of synchronously pulsing neurons.

C. State Dependent Modulation Provides Focus of Attention

Biological studies show the vision system performs substantial editing to focus attention and deemphasizes irrelevant information [9]. Even at early stages of processing, preference is given to elements to which the observer is paying attention. The response of many neurons double when the stimulus is a target of attention. State dependent signals are believed to be the stimulus that causes this preferential treatment. These are signals that originate from visual areas other than the retina, and are believed to modulate a neuron’s response to any object upon which attention is focused. The signals may originate from areas in the visual cortex, or from the higher processing areas in the parietal and temporal lobes. The signals modulate a neuron’s response to a stimulus within its receptive field causing a state of focused attention on the object causing the stimulus. This phenomenon is called state dependent modulation and is a method for one area of processing to superimpose its findings, or expectations.
Fig. 2. PCNN fusion architecture used to fuse breast cancer and FLIR images.

Fig. 3. Feeding and linking connections of a single PCNN neuron.

on another area [9]. The modulatory effect of state dependent modulations are believed to focus attention by elevating the perception of objects of interest effectively suppressing unnecesary information in a visual scene. The PCNN fusion network uses this biological principle to focus attention on objects that best fit the criteria of a desired object. By using the relative presence of a desired feature as a state dependent modulation signal, the networks response to the desired object is elevated. This elevated response facilitates detecting and isolating of a particular object in a visual scene.

III. A PCNN IMAGE FUSION NETWORK FOR OBJECT DETECTION

A. The PCNN Fusion Network

To perform object detection, the PCNN fusion network takes an original and several filtered versions of a gray-scaled image and outputs a single image in which the desired objects are the brightest and thus easily detected. Fig. 2 shows the PCNN fusion network which is the connection architecture used to fuse the original and filtered images. Each PCNN has one neuron per input image pixel. The pulse rate of each neuron in the center PCNN is used as a brightness value for the pixels in the output image. The neurons within the PCNN are arranged as a single two dimensional layer network with lateral linking. Fig. 3 shows the feeding and linking connections of a single neuron within the PCNN. Every neuron receives linking inputs from all neighboring neurons within a radius of three (radius of one is shown in Fig. 3 for clarity). Each neuron receives feeding inputs which are the intensity of the corresponding pixels in the input image. The pulse based linking mechanisms of the PCNN use temporal synchronization to segment the original image. The outer PCNN’s provide state dependent modulation signals used to focus attention on segments of interest.

Fig. 4 shows the inputs and output of the fusion process when used on a small portion of a mammogram which contains microcalcifications. The pulse rate of each output neuron is used as a brightness value for the pixels in the output image. Fig. 4(a)–4(c) is the images used as input to the fusion network. The fusion results are shown in Fig. 4(d). A threshold has been applied to remove the background and lower intensity segments. The segments that remain are the desired objects.

B. Using Image Processing Filters to Extract Features

To simulate biological information fusion, we need to simulate the output information of the specialty areas (V3, V4, and V5) and then fuse the information. To simulate the performed feature extraction, we will observe the hypothesis that neuronal processing units are best described as filters that are selective along multiple stimulus directions [14]. Information extracted by filters is used as features in the PCNN fusion network. Table I gives a list of possible filters that can be used to approximate each visual area of the static form pathway.

Since all images used in this paper (mammogram and FLIR) are static and gray scaled, only the static form pathway is included in this table. Wavelet, morphological (hit-and-miss), and difference-of-Gaussian (DoG) filters are chosen to provide features because they have been used successfully to detect breast cancer in mammograms [19]–[21] and in FLIR images [6]. These filters are used to separate visual information into various size and wavelength components. The filtered components serve as individual features which will be fused by the PCNN fusion network into a single image which combines yet exploits the selectivity of each individual filter.

When applied to the mammograms, the filters are tuned to be selective to microcalcifications which can be an early indication of cancerous growth [2]. For the FLIR images, the filters are tuned for selectivity to features of a mobile SCUD missile launcher. Since these filters are selective to a particular object, the outputs can be used as state dependent modulation

| Vision Model Area | Possible Filter Models |
|-------------------|-----------------------|
| V1 wavelength selective | Gabor filters, Gaussian or Wavelet filters |
| V1 orientation selective | Gabor filters, Gaussian or Wavelet filters |
| V1 layer 4B orientation selective (V1,4B0) | Gaussian or Wavelet filters, orientation selective filters |
| V1 layer 4B orientation + direction selective | Gaussian or Wavelet filters, orientation selective filters |
| V2 wavelength selective | Gabor filters, Gaussian or Wavelet filters |
| V2 orientation selective | Gabor filters, Gaussian or Wavelet filters |
| Static Form | Gabor filters, Gaussian or Wavelet filters |
Fig. 4. 128-by-128 pixel region containing microcalcifications (segmented from a 1000-by-2000 pixel mammogram). (a) Original image. (b) hit-and-miss filtered image. (c) wavelet filtered image. (d) PCNN fused image after a threshold has been applied.

Fig. 5. The Eckhorn model neuron in a PCNN.
Fig. 6. PCNN segmentation (output pulse periods produced by a PCNN with a linear zero to one input).

signals where the current state of attention is focused on detecting objects that resemble the target object.

C. The Pulse Coupled Neural Network

The heart of the fusion network is the PCNN. The PCNN is a physiologically motivated artificial neural network that is composed of artificial spiking neurons which are interconnected via multiplicative links. The PCNN uses the Eckhorn model spiking neuron [8] which is shown in block diagram form in Fig. 5. The neuron models the pulse height, duration, repetition rate, and modulatory interneural linking observed in biological dendrites. This neuron model is chosen because it contains the modulatory pulse-based linking necessary to simulate the temporal synchronization and state-dependent modulation observed in the primate visual cortex. The most notable aspects of the neuron are the dendritic branch and the pulse generator sections. The dendritic branch contains feeding inputs which are modulated by linking inputs. Each input contains a leaky integrator which models a dendritic synapse. The leaky integrator converts incoming pulses into a persistent signal. The time constant ($\tau_p$ or $\tau_L$) of the leaky integrator models the decay rate of neurotransmitters within the synapse. The pulse generator section is an oscillator that produces an output pulse train of very short duration pulses whose frequency is based on input magnitude. The pulse generator time constant $\tau_S$ models the refractory period that occurs after a biological dendrite fires. For detailed discussion of the inner workings of the PCNN see Eckhorn [8] and Johnson [22]–[24].

As used in this model, a PCNN neuron receives feeding inputs ($X_j$) from a gray-scaled image and receives linking inputs ($Y_j$) from neighboring neurons, and neurons in other PCNN’s (Fig. 3).

1) Pulse Coupling Performs Temporal Synchronization: Pulse based synchronization is the key characteristic that distinguishes the PCNN from other types of neural networks. The image segmentation property of the PCNN comes from this synchronization. Neurons with similar inputs pulse in synchrony to represent a segment of the input image. Neurons with related feeding input characteristics (color, intensity, etc.) have similar pulsing rates. The linking connections cause neurons, in close proximity and with related characteristics, to pulse in unison (synchronization). The PCNN links pixels based on similarity. This similarity is defined by an image pixel’s intensity value relative to the intensity values of the neighboring pixels within its linking radius. A pixel is similar to any pixel that is within its linking radius and has an intensity value within $F/\beta L$ greater than its own, where $F$ is the total feeding input value to the neuron, $L$ is the total linking input value, and $\beta$ is the value of the linking strength between neurons. Shown in equation form, a pixel with an intensity of $I_1$ is similar to a pixel with intensity of $I_2$ if

$$0 \leq F_{I_2} - F_{I_1} \leq F_{I_1} \beta L_{I_1}.$$  

Because of the multiplicative linking connections, this relation is not as simple and straightforward as it first appears. The following discussion will make some simplifying assumptions to demonstrate the complexity of determining which neurons are similar.

The pulse period ($T$) of a digitally simulated neuron with constant linking inputs is defined by the equation

$$T = -\tau_S \ln(1 + U/V^S)$$

where the neuron’s internal activity $U$ is defined as $U = F(1 + \beta L)$ and $[\cdot]$ denotes the ceiling function. Without any linking inputs ($L = 0$), bandwidth limitations of the neuron
would cause input values between zero and one to fire in nonoverlapping logarithmic sized groups as shown in Fig. 6 (much higher values of \( \tau_s \) are typically used). Notice that if \( L = 0 \), \( U \) is equal to the total feeding inputs \( F \). The scale of the output pulse period axis is time units where one unit is the maximum pulse firing rate the neuron bandwidth will support. For a digital implementation, each unit would be one time-step on the simulation clock. The values of \( U \) that pulse each time slice (without linking present) are shown by the bold lines. The set \( P(t) \) is defined to be the values of \( U \) that pulse at time \( t \) when no linking is present. Adding a constant linking input to a neuron extends the lower limit of \( P(t) \) by \( F/\beta L \) (shown as the thin line in Fig. 6). The set \( S \) of real numbers that are added to \( P(t) \) due to linking is defined to be the synchronization range of a PCNN neuron

\[
S = [F, F + F/\beta L].
\]

This synchronization range defines the similarity in pixel intensity which will cause neurons to synchronize and form a segment. A neuron that would not normally fire at time \( t \) will fire in synchrony with other neurons that fire at time \( t \) if

\[
S \cap P(t) \neq \emptyset.
\]

This criteria must be met for a neuron to synchronize with other neurons pulsing at a particular pulse frequency \( f = 1/T \).

Notice in Fig. 6 the total pulse range \( (P(t) \cup S) \) for each time \( t \) overlaps the total pulse range for time \( t+1 \). This means a neuron with internal activity value \( U \) in the overlapping region can fire at either time \( t \) or \( t+1 \) depending on linking inputs. So will a particular neuron \( N_i \) fire at time \( t \) or \( t+1 \) ?

Expanding the earlier constant linking input signal assumption to state linking inputs originate as the constant outputs of neighboring neurons as shown in Fig. 3, makes \( L \) a function of the feeding and linking inputs of neighboring neuron’s. Since the value of \( L \) originates as the output of neighboring neurons and the synchronization range \( S \) is a function of \( L \), (3) implies segmentation is image content dependent. For two adjacent neurons that are linked, the output of each neuron is dependent upon the output of the other. Since linked neurons are dependent upon one another, finding the output pulse period of a particular neuron requires solving simultaneous equations. For example, the output period of neurons connected in a \( 3 \times 3 \) array is described by the following matrix (assuming \( V^S = 1 \) ) shown at the bottom of the page. Since the value of each \( L \) is dependent on the output period of neighboring neurons, finding the output of any single neuron requires solving the nine equalities simultaneously. In essence, this is what the PCNN does. The assumption of a constant linking input simplifies the problem significantly. Since the PCNN is based on a spiking neuron, all linking signals are pulses which means linking inputs are not constant.

The actual operation of the PCNN is more complex than this simplified example, but the functional concept is the same. The actual PCNN solves the interneuron dependencies in a unique way. No linking signals are present until the first neuron fires. The brightest points within an image cause their corresponding neurons to fire first. This firing initiates a linking signal (linking wave) which travels through the multiplicative linking interconnects causing other neurons with similar inputs to fire [22]. Since linking fields overlap, pixel grouping occurs beyond the limits of a neuron’s linking radius. A single neuron can fire and cause a domino effect that continues until all neurons with similar inputs fire in phase synchrony with the first neuron. This group of synchronously firing neurons represents a distinct segment within the image. The segmentation process repeats each time step, on neurons that have not fired, until all neurons within the PCNN have fired and the image is completely segmented.

2) Pulse-Based Multiplicative Linking Performs State Dependent Modulation: As shown in the Fig. 5, the total input to a PCNN neuron \( (U) \) can be describe by the multiplicative equation

\[
U = F(1 + \beta L)(1 + \beta L E_x)
\]

where \( LE_x \) is the value of total linking inputs from sources external to the PCNN (possibly other PCNN’s). The signal \( U \) feeds directly into the pulse generator section of the PCNN which produces the output pulse train. The output frequency of pulses produced by the pulse generator is

\[
f = \frac{1}{-\tau_s \ln(1 + U/V^S)}
\]

which is the reciprocal of the output period shown in (2). From (5) it can be seen that the linking inputs of the PCNN modulate the feeding inputs. This modulatory property of the PCNN can be used to simulate the state dependent modulation observed in the visual system. Without linking inputs, \( U \) would equal \( F \) and the feeding input would drive the pulse generator section. A positive linking input \( (L > 0) \) would increase the value of \( U \) which would increase the frequency of the output pulse train (2). If filter outputs that are selective to features of a desired target are used as linking inputs, then neurons connected to image areas that resemble the desired target would have greater linking inputs than those that do not. The filter outputs represent the state dependent modulation signals when the current state of the PCNN is a focus of attention on the desired target. The neurons whose inputs most match the desired target would have the greatest modulatory input, thus having the highest frequency output. This increased output effectively separates the neurons from the rest of the image.

For the PCNN fusion network, this modulatory mechanism provides a method of associating filtered features with segments in the original image. It also provides a focus of

\[
[T] = \begin{bmatrix}
-\tau_s \ln(F_{11}(1 + \beta L_{11})) & -\tau_s \ln(F_{12}(1 + \beta L_{12})) & -\tau_s \ln(F_{13}(1 + \beta L_{13})) \\
-\tau_s \ln(F_{21}(1 + \beta L_{21})) & -\tau_s \ln(F_{22}(1 + \beta L_{22})) & -\tau_s \ln(F_{23}(1 + \beta L_{23})) \\
-\tau_s \ln(F_{31}(1 + \beta L_{31})) & -\tau_s \ln(F_{32}(1 + \beta L_{32})) & -\tau_s \ln(F_{33}(1 + \beta L_{33}))
\end{bmatrix}.
\]
attention to isolate the segment. Segments with a greater number of desired features present will be more active than other segments, therefore the most active segments are those that fulfill more of the target criteria. These segments are easily separable from the rest of the image.

D. How Information Is Fused

The cornerstone of the PCNN fusion network is the segmentation performed by the pulse synchronizations. This temporal synchronization groups the image pixels into individual, disjoint segmented regions (objects) that pulse in different time steps. These individual regions represent either the desired target(s) (cancer or SCUD), or other objects such as the background. The fusion process exploits the fact that an object of interest is represented as a group of pixels. This allows dissimilar and possibly spatially disjoint features such as brightness, edges, and gradients to be mapped into a single region of space that represents the object. This allows fusion of several dissimilar features of an object into a single representation of an object that contains more information than any individual feature, or subgroup of features. In a search for a specific object, size alone may remove many segments from consideration.

In the PCNN fusion network, the original image is used as a basis for object segmentation, and the filtered versions of the original image are used as the dissimilar features. The filters are selective to particular features of the desired object, thus the filtered images represent selective image information with a focus of attention on the desired object. The outer PCNN’s (Fig. 2) convert the filtered images into pulsed signals for use as state dependent modulation signals. These pulsed signals are linked to the original image using the center PCNN’s linking inputs. These signals both fuse features into its associated segment, and modulate the center PCNN’s neuronal response to the object of interest. The modulation signal increases the pulsing frequency of the synchronous group of neurons effectively separating the desired object from the background. The segments of the original image that best fulfill the selective criteria of the filters will be represented by the fastest pulsing groups of output neurons. An output image is created by using the output neuron pulse frequencies as pixel intensity values. Only simple thresholding of the fused output image is needed to extract the group of pixels that contain the desired objects.

E. The PCNN Produces a Time Signal as an Added Benefit

With many image fusion/object detection techniques available, why use a method based on a PCNN? The PCNN provides an additional benefit which is unique. The pulsed output of the PCNN forms a time signal that can be used as a translation, rotation, scale, distortion, and intensity invariant signature for each object in the image [23], [24]. This signal can be used for identification purposes on any object detected by the PCNN fusion network. The time signal is inherently produced during PCNN operation and thus requires no additional processing. The time signal was not utilized in this project, but could have been. The next stage of this research will use the signature property of the time signal to increase object detection accuracy.

IV. DETECTION RESULTS ON CANCER AND RADAR IMAGES

One hundred FLIR images, from aircraft training runs, were used to calibrate and test the object detection capability of the PCNN fusion network. Fifty images were used to calibrate the PCNN weights, linking radius, β and threshold parameters. Each image contained a single SCUD mobile missile launcher, a truck, a van, and four surrounding flash pods to mark the target location. The flash pods function as a guide for the photographer and are not used by the detection algorithms. After PCNN calibration, the object detection capability of the fusion network was tested on the remaining 50 FLIR images. The goal of the test was to detect the SCUD launcher while minimizing the number of false alarms. Since only one target is present, and the output of the PCNN fusion network can be processed by a pattern recognition engine for additional identification, a large false alarm rate is preferable to a missed target. For this reason, all filters were tuned conservatively to ensure SCUD detection. Any detected object other than the SCUD, truck, van, pods, and image edge effects were considered false targets. Detection of the truck, van, and flash pods was considered optional.

Table II presents the detection accuracy achieved by each method (only 12 representative images shown). Fig. 7 shows an example binary (thresholded) image used as input and the resulting output of the PCNN fusion network. For every image inside the desired target detection range (ranges of interest from a munitions release perspective), the selective filters and the PCNN fusion network detected the SCUD mobile missile launcher. As can be seen in Fig. 7(a) and 7(b), conservative tuning can cause the selective filter routines to produce a large number of false alarms. The Hit/Miss filter algorithm averaged 8.2 false targets per image, the DoG filter algorithm averaged 20.3 false targets per image, and PCNN network averaged 0.6 false alarms per image. When compared to the best filter accuracy, the PCNN network removed 94% of the false alarms without removing any true detections. The accuracy produced by the PCNN network also exceeds the accuracy produced by ANDing the binary filter outputs.

In the second test, the algorithms were used to detect microcalcifications in mammograms. Microcalcification density
is often used by computer aided diagnosis (CADx) systems for early detection of cancerous breast regions [1]. Microcalcifications are present in healthy tissue, but a high density (5+ per square centimeter) can be an early indication of cancer. In this test, the selective filters were tuned to detect radiologist identified microcalcifications. The goal of the test was to maximize the detection of identified microcalcifications while minimizing the number of other detections. All detected objects that did not represent an identified microcalcification were considered false targets. The identified microcalcifications were visually detectable, but others may exist. Since this test does not attempt to detect all microcalcifications, but only ones identified by radiologist, the resulting accuracy should not be directly compared to other cancer detection algorithms. The purpose of the test is to demonstrate information fusion by a PCNN.

Thirty 256 × 256 pixel regions segmented from full breast mammograms were used to test microcalcification detection. Eighteen of the regions were used to calibrate the PCNN network and the filter algorithms, and the remaining 12 were used to test the detection accuracy. Table III presents the detection accuracy achieved by each algorithm. Since the PCNN network fuses the results of the selective filters, no additional

| Image Name | Number of Cals Found/Number of False Alarms |
|------------|---------------------------------------------|
| a002a01m   | 16/15 21/27 18/15                        |
| a002b01m   | 31/15 41/26 38/12                        |
| a002b02m   | 20/17 24/19 21/8                         |
| a002c01m   | 32/7  49/18 15/0                         |
| a002d01m   | 25/29 32/27 26/20                       |
| a003a01m   | 5/41  7/22  3/34                        |
| a004b01m   | 3/24  3/24  3/24                        |
| a004b02m   | 4/26  5/22  4/11                        |
| a005a01m   | 6/16  7/29  6/21                        |
| a005b01m   | 2/9   4/23  0/0                         |
| a006a01m   | 9/14  9/10  8/2                         |
| a006b01m   | 10/11 10/7  10/1                        |
| Average Ratio | 0.76  0.60  1.24                      |

Fig. 7. Input and output images of a mobile SCUD launcher and flash pods. (a) Original FLIR Image. (b) DoG filtered image. (c) Morphological (Hit/miss) filtered image. (d) PCNN fusion network output image.

| TABLE III |
| DETECTION RESULTS OF MICROCALCIFICATIONS IN MAMMOGRAMS USING THE PCNN FUSION NETWORK |
|----------------------------------|----------------|----------------|----------------|----------------|
| Image Name | Number of Cals | Found/Number of False Alarms |
|------------|----------------|----------------|----------------|----------------|
| a002a01m   | 16/15          | 21/27          | 18/15          |                |
| a002b01m   | 31/15          | 41/26          | 38/12          |                |
| a002b02m   | 20/17          | 24/19          | 21/8           |                |
| a002c01m   | 32/7           | 49/18          | 15/0           |                |
| a002d01m   | 25/29          | 32/27          | 26/20          |                |
| a003a01m   | 5/41           | 7/22           | 3/34           |                |
| a004b01m   | 3/24           | 3/24           | 3/24           |                |
| a004b02m   | 4/26           | 5/22           | 4/11           |                |
| a005a01m   | 6/16           | 7/29           | 6/21           |                |
| a005b01m   | 2/9            | 4/23           | 0/0            |                |
| a006a01m   | 9/14           | 9/10           | 8/2            |                |
| a006b01m   | 10/11          | 10/7           | 10/1           |                |
| Average Ratio | 0.76          | 0.60           | 1.24           |                |
true detections were expected or achieved. The results do show that the number of false detections were significantly reduced with only a small reduction in true detections. The Hit/Miss algorithm averaged 1.3 false detections for each true detection and the DoG algorithm averaged 1.7 false detections per true one. The PCNN network reduced these ratios to 0.8 false detections per true detection. When compared to the best filter result, the PCNN network removed 46 percent of the false detections while removing only 7% of the true detections.

V. CONCLUSION

The fusion network provided a greater accuracy increase on the FLIR images than on the mammogram images. The network reduced the false alarm rate from 8.2 to 0.6 false alarms per image in the FLIR images and from 1.7 to 0.8 false detections per true detections in the mammograms. In the fusion process, the PCNN network does not add true detections to the output, but instead removes false detections. Since the FLIR images contained many objects such as trees and roads that were larger than the target, the PCNN could easily segment and remove the large objects. Because the mammograms contained few large objects with consistent brightness and boundaries, the PCNN segmented the image into many small objects which prevented any significant object removal based on size. The majority of the information removal was performed by the state dependent modulation. These results imply the PCNN fusion network is better suited for processing images which contain structures that differ in size from the targets. The PCNN was able to map many false detections, such as road and forest edges, into the larger original object and subsequently remove the false detections. The tests have shown the network is suitable for removing false detections from conservatively tuned filter outputs while preserving a majority of the true detections. The network removed 94% of the false detections without removing any true detections in the FLIR images and removed 46% of the false detections while removing only 7% of the true detections in the mammograms.

In general, the PCNN fusion network provides a method of improving object detection accuracy by fusing the outputs of multiple object detection algorithms. The accuracy of the fusion network surpassed the accuracy provided by the results of any single filtered output, or the logical AND of all filter results. The network inputs pixel-based information and produces an object-based output with a time signal that can be used as an identification signature for each object. The brightness values of the objects in the output image represent the degree to which each object matches the characteristics of the desired object. The PCNN also provides a good computer architecture for implementing physiologically-based fusion and other pulse-based physiologically observed phenomenon.

REFERENCES

[1] M. L. Giger, “Computer-aided diagnosis,” *RSNA Categorical Course Physics*, pp. 283–298, 1993.
[2] R. M. Nishikawa, M. L. Giger, K. Doi, C. Mete, F. Yin, C. Vyborny, and R. Schmidt, “Computer-aided detection and diagnosis of masses and clustered microcalcification from digital mammograms,” *SPIE*, vol. 1905, pp. 422–432, 1994.
[3] R. N. Strickland and H. I. Hahn, “Wavelet transforms for detecting microcalcifications in mammograms,” *IEEE Trans. Med. Imag.*, vol. 15, pp. 218–229, 1995.
[4] A. Laine and S. Schulter, “Hexagonal wavelet processing of digital mammography,” *SPIE Image Processing*, pp. 1011–1018, 1993.
[5] A. Laine, S. Schuler, J. Fan, and W. Huda, “Mammographic feature enhancement by multiscalar analysis,” *IEEE Trans. Med. Imag.*, vol. 13, Dec. 1994.
[6] S. K. Rogers, J. M. Colombi, C. E. Martin, J. C. Gainey, K. H. Fielding, T. J. Burns, D. W. Ruck, M. Kabrisky, and M. Oxley, “Neural networks for automatic target recognition,” *IEEE Trans. Neural Networks*, vol. 8, pp. 1153–1184, 1995.
[7] C. M. Gray, P. Konig, A. K. Engel, and W. Singer, “Oscillatory responses in cat visual cortex exhibit intercolumnar synchronization which reflects global stimulus properties,” *Nature*, vol. 358, pp. 334–337, Mar. 1999.
[8] R. Eckhorn, H. J. Reitboeck, M. Arndt, and P. Dicke, “Feature linking via synchronization among distributed assemblies: Simulations of results from cat visual cortex,” *Neural Comput.*, vol. 2, pp. 293–307, 1990.
[9] J. H. R. Maunsell, “The brain’s visual world: Representation of visual targets in cerebral cortex,” *Science*, vol. 270, pp. 764–768, 1995.
[10] M. P. Young, “The large scale organization of the primate cortical visual system,” *SPIE*, vol. 2054, pp. 185–193, 1999.
[11] S. Zeki, *A Vision of the Brain*. Oxford, U.K.: Blackwell, 1993.
[12] H. Neven and A. Aertsen, “Rate coherence and event coherence in the visual cortex: A neuronal model of object recognition,” *Biol. Cybern.*, vol. 67, pp. 309–322, 1992.
[13] M. T. Wallace, M. A. Meredith, and B. E. Stein, “Integration of multiple sensory modalities in cat cortex,” *Experimental Brain Res.*, vol. 91, no. 3, pp. 484–488, 1992.
[14] C. Van Essen, H. C. Anderson, and D. J. Felleman, “Information processing in the primate visual system: An integrated systems perspective,” *Science*, vol. 255, pp. 419–422, 1992.
[15] S. Grossberg, *Neural Networks and Natural Intelligence*, Massachusetts: MIT Press, Cambridge, 1988.
[16] S. Grossberg and E. Mingolla, “Neural dynamics of perceptual grouping: Textures, boundaries and emergent segmentation,” in *Perception Psychophys.*, vol. 38, no. 2, pp. 141–171, 1985.
[17] G. A. Carpenter, S. Grossberg, and C. Mehanian, “Invariant recognition of cluttered scenes by a self-organizing art architecture: Cort-x boundary segmentation,” *Neural Networks*, vol. 2, pp. 169–181, 1989.
[18] M. R. Turner, “Texture discrimination by gabor functions,” *Biol. Cybern.*, vol. 55, pp. 71–82, 1986.
[19] R. Dak, S. Rogers, J. Hoffmeister, D. Ruck, and R. Raines, “A comparative study of texture measures for the classification of breast tissue,” *AFIT Tech. Rep.*, Jan. 1996.
[20] D. McCandless, S. Rogers, J. Hoffmeister, D. Ruck, and R. Raines, “Model based detection of clustered microcalcifications using wavelets and phantom derived images,” *IEEE Trans. Med. Imag.*, 1996.
[21] W. Polakowski, S. Rogers, D. Ruck, J. Hoffmeister, and R. Raines, “Computer aided breast cancer detection and diagnosis of masses using difference of gaussian and derivative-based feature saliency,” *IEEE Trans. Med. Imag.*, vol. 16, pp. 811–819, 1997.
[22] J. L. Johnson, “Pulse coupled neural networks,” in *Proc SPIE Int. Symp. Opt. Eng. Photonics Aerospace Sensing, Conf. Adaptive Comput.: Math., Electron. Opt. (Critical Reviews)*, Orlando, FL, 1994.
[23] J. L. Johnson, “Pulse coupled neural nets: Translation, rotation, scale, distortion, and intensity signal invariance for images,” *Appl. Opt.*, vol. 33, no. 26, pp. 6239–6253, 1994.
[24] J. L. Johnson, “Time signatures of images,” in *Proc. IEEE Int. Conf. Neural Networks*, Orlando, FL., 1994.

Randy P. Broussard. Major, USAF, received the B.S. degree in electrical engineering from Tulane University, New Orleans, LA, in 1986, and the M.S. degree in computer engineering from Florida Institute of Technology, Melbourne, in 1991, and the Ph.D. degree in electrical engineering from the Air Force Institute of Technology, Wright-Patterson Air Force Base, OH, in 1997. He presently works in the Air Force Research Laboratory. Sensors directorate, Reconnaissance, Strike, and Combat ID branch. His research interests are neural networks, computer vision, and pattern recognition.
Steven K. Rogers (S’76–M’78–SM’95’F’98) received the B.S. degree in electrical engineering/computer science from the University of Colorado, Boulder, the M.S. degree in electrical engineering from the Air Force Institute of Technology, Wright-Patterson Air Force Base, OH, and the Ph.D. degree in electrical engineering from the University of Colorado.

He is a Professor of Electrical Engineering in the Department of Electrical and Computer Engineering, Air Force Institute of Technology. He is presently conducting an extensive research program on neural networks and pattern recognition. The research program addresses the problems inherent in making smart weapons. He has published more than 200 papers in the areas of neural networks, pattern recognition, and optical information processing and several textbooks including Introduction to Biological and Artificial Neural Networks for Pattern Recognition.

Dr. Rogers is an SPIE Fellow.

Mark E. Oxley received the B.S. degree in mathematics in 1978 from Cumberland College, Williamsburg, KY, the M.S. degree in applied mathematics in 1980 from Purdue University, West Lafayette, IN, and the Ph.D. degree in mathematics in 1987 from North Carolina State University, Raleigh.

He is an Associate Professor of Mathematics in the Department of Mathematics and Statistics at the Air Force Institute of Technology, Wright-Patterson Air Force Base, OH. His current research interests include neural networks, wavelet analysis, functional analysis, and nonlinear partial differential equations.

Dr. Oxley is a member of Pi Mu Epsilon, the American Mathematical Society (AMS), the Society for Industrial and Applied Mathematics (SIAM), the American Geophysical Union (AGU), and SPIE the International Society for Optical Engineering.

Gregory L. Tarr, photograph and biography not available at the time of publication.