ELSEVIER

A novel cognitive interpretation of breast cancer thermography with complementary learning fuzzy neural memory structure

T.Z. Tan [a], C. Quek [a,*], G.S. Ng [a], E.Y.K. Ng [b]

[a] *Centre for Computational Intelligence, School of Computer Engineering, Nanyang Technological University, Blk N4, #B1a-02, Nanyang Avenue, Singapore 639798, Singapore*
[b] *School of Mechanical and Aerospace Engineering, Nanyang Technological University, Blk N2, #01a-29, Nanyang Avenue, Singapore 639798, Singapore*

Abstract

Early detection of breast cancer is the key to improve survival rate. Thermogram is a promising front-line screening tool as it is able to warn women of breast cancer up to 10 years in advance. However, analysis and interpretation of thermogram are heavily dependent on the analysts, which may be inconsistent and error-prone. In order to boost the accuracy of preliminary screening using thermogram without incurring additional financial burden, *Complementary Learning Fuzzy Neural Network* (CLFNN), FALCON-AART is proposed as the *Computer-Assisted Intervention* (CAI) tool for thermogram analysis. CLFNN is a neuroscience-inspired technique that provides intuitive fuzzy rules, human-like reasoning, and good classification performance. Confluence of thermogram and CLFNN offers a promising tool for fighting breast cancer.
© 2006 Elsevier Ltd. All rights reserved.

Keywords: FALCON-AART; Fuzzy adaptive learning control network fuzzy neural network; Thermogram; Breast cancer diagnosis; Complementary learning

1. Introduction

Breast cancer is the second most deadly cancer among women. Each year, 211,240 women are diagnosed with breast cancer and 40,870 of them will die in 2005 (American Cancer Society, 2005). In United States alone, it is estimated that there are 1 million women with undetected breast cancer; to date, the figure of women affected has surged to 1.8 million and 45, 000 women die per year (Diakides & Diakides, 2003). This high death rate has stimulated extensive researches in breast cancer detection and treatment. Recent studies have determined that the key to breast cancer survival rests upon its earliest detection possible. If discovered in its earliest stage, 95% cure rates are possible (Gautherie, 1999; Pacific Chiropractic and Research Center). On the other side, it is reported that 70 to 90% of the *excisional*

biopsies performed are found to be benign (Lay, Crump, Frykberg, Goedde, & Copeland, 1990). Owing to this high false positive rate, many endeavors have been putted into ameliorate the breast cancer early detection.

Breast imaging is a noninvasive and inexpensive cancer detection technology. Amongst, mammography is accepted as the most reliable and cost-effective imaging modality. However, its false-negative rates is high (up to 30%) (Elmore, Wells, & Carol, 1994; Rajentheran, Rao, Lim, & Lennard, 2001). In addition, the danger of ionizing radiation and tissue density, which has been associated with increased cancer risk (Boyd, Byng, & Jong, 1995), is linked with patient who underwent mammography screening. It is also uncomfortable, because the breast has to be compressed between flat surfaces to improve image quality. Furthermore, obtaining adequate images from radiologically dense breasts (with little fat) or in women with breast implants are difficult (Foster, 1998), and it is difficult to detect breast cancer in young women (Gohagan, Rodes,

---

* Corresponding author. Tel.: +65 6790 4926; fax: +65 6792 6559.
E-mail address: ashcquek@ntu.edu.sg (C. Quek).

Blackwell, & Darby, 2004). Despite of these limitations, mammogram remains the gold standard for screenings (Gohagan et al., 2004; Moore, 2001). Since early detection is important, new technologies such as *Magnetic Resonance Imaging* (MRI), *Positron Emission Tomography* (PET), *Computed Tomography-Single Photon Emission Computed Tomography* (CT-SPECT) (Del Guerra, Di Domenico, Fantini, & Gambaccini, 2003), and ultrasound have been applied as complement to mammogram (Ng & Fok, 2003). Fig. 1 and Table 1 show the available modalities for breast cancer detection at present, and the reported accuracy, respectively. Note that the reported accuracy is
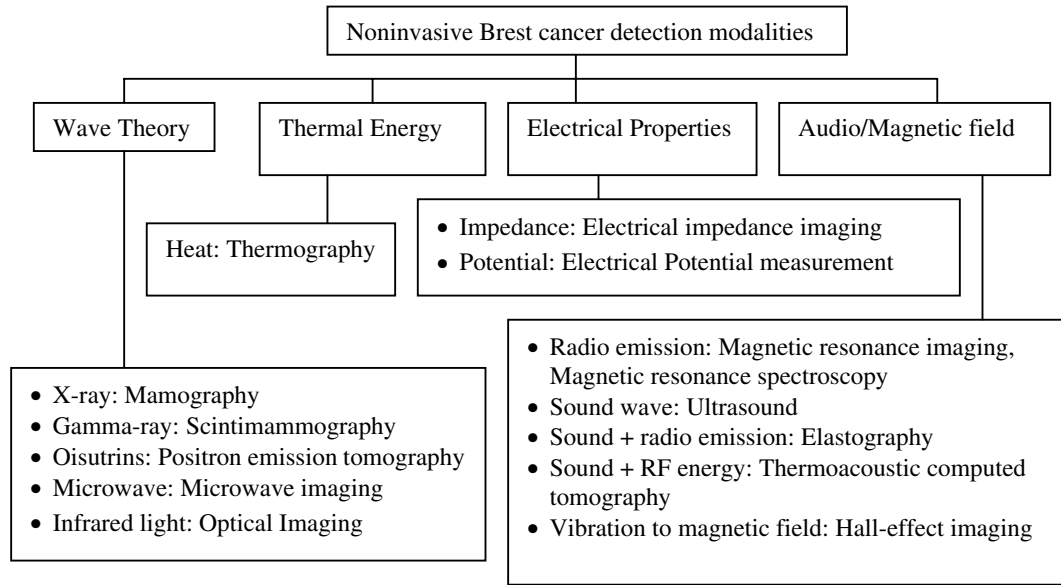

Fig. 1. Noninvasive breast cancer detection modalities (adapted from Fok et al., 2002).

Table 1
Accuracy of breast cancer diagnosis modalities

Technique	Sensitivity (%)	Specificity (%)	References
Clinical examination	48.3–59.8	90.2–96.9	Barton (2002), McDonald et al. (2004)
Biopsy			
Surgical/Open biopsy	≈100	≈100	Imaginis, Breast Cancer Diagnosis (2004)
Vacuum-assisted biopsy (Mammotome)	95	98	Simmon et al. (2000)
Large core biopsy	74–97	91–100	Delle Chiaie and Terinde (2004), Meyer et al. (1999), Puglisi et al. (2003)
FNA (biopsy)	85–88	55.6–90.5	Pisano et al. (2001)
Core needle biopsy	91–99	73–100	Brenner et al. (2001), Pisano et al. (2001)
Breast cyst aspiration	79	94	Lucas and Cone (2003)
Imaging			
FNA (cytology)	65–99	64–100	Fajardo et al. (1990), Reinikainen (2003)
Mammography	13–95	14–90	Fajardo et al. (1990), Fletcher et al. (1993), Singhal and Thomson (2004)
Full-Field Digital Mammography (FFDM)	64.3	88.2	Irwig et al. (2004), Lewin et al. (2002)
Thermography	90	90	Amalu (2003)
Ultrasound/Sonography	13–98.4	67.8–94	Houssami et al. (2003), Singhal and Thomson (2004), Stavros et al. (1995)
MRI	86–100	21–97	Cecil et al. (2001), Orel (2000), Singhal and Thomson (2004), Yeung et al. (2001)
Proton Magnetic Resonance Spectroscopy (MRS)	83–100	73–87	Cecil et al. (2001), Reinikainen (2003), Yeung et al. (2001)
Scintigraphy (CT)	55–95	62–94	Brem et al. (2003), Singhal and Thomson (2004)
PET	96	100	Singhal and Thomson (2004)
Positron Emission Mammography (PEM)	80–86	91–100	Levine et al. (2003), Murthy et al. (2000)
Electrical Impedance Scanning (EIS)	62–93	52–69	Glickman et al. (2002), Malich et al. (2003)
Gene screening			
Serum protein expression profiling	90	93	Vlahou et al. (2003)
Gene Profiling	83–91	72.7–81.8	van't Veer et al. (2002)
Gene Testing	63–85	Not mentioned	Berry et al. (2002)

only an estimate because these modalities perform differently on different types of breast cancer, on different age group, apart from the fact that most of the tests are done on small populations.

As shown in Table 1, none of the methods possesses high sensitivity (correctly identify women with breast cancer), and high specificity (correctly weed out women without breast cancer), albeit a lot of endeavors have been put in Foster (1998), Moore (2001). Each has its limitations. For example, clinical examination is insensitive, examiner-dependent (McDonald, Saslow, & Alciati, 2004); biopsy is invasive, causes complications, leaves scars, and requires long recovery time (Imaginis, Breast Cancer Diagnosis, 2004; Simmon, Kalbhen, Cooper, & Flisak, 2000); MRI is inconsistent, costly and low-resolution (Cardillo, Starita, Caramella, & Cilotti, 2001); PET, CT-SPECT, is expensive and scarce; ultrasound images are of poor resolution (Kotre, 1993; Moore, 2001), and operator-dependent (Chen, Chang, & Huang, 2000); microwave imaging requires accurate modeling of the relation between various tissues' frequency dependency, and its sensitivity is affected by many factors (Bond, Li, Hagness, & Van Venn, 2003; Fear, Hagness, Meaney, Okoniewski, & Stuchly, 2002; Kosmas, Rappaport, & Bishop, 2004); PEM (Thompson, Murthy, Picard, Weinberg, & Mako, 1995) is expensive and insensitive (Moses, 2004); FNA is operator-dependent (Pisano, Fajardo, Caudry, & Sneige, 2001), and incurs complications (Lucas & Cone, 2003); Gene expression analysis on genes BRCA1 and BRCA2, whose mutations are associated with breast cancer, is difficult as the genes are highly complex. The costly blood storage worsens the matter (Spengler, 2003); MRS is technically demanding, and only of confimatory value to MRI (Cecil, Siegelman, & Lenkinski, 2001; He & Shkarin, 1999); EIS requires localization of lesion before hand (Glickman et al., 2002), insensitive, and observer-dependent (Malich et al., 2003). These methods are often too cumbersome, costly inaccessible or invasive to be used as first-line detection modalities alongside clinical examination and mammography (Keyserlingk, Ahlgren, Yu, Belliveau, & Yassa, 2000; Qi & Diakides, 2003).

Thus, thermogram appears as one of the most promising and suitable alternatives for preliminary screening (Amalu, 2003). Thermogram monitors the breast health based on the heat pattern variation that correlates with the patients' medical condition (Gautherie, 1999; Head, Wang, Lipari, & Elliott, 2000). It is cheap, noninvasive, simple, painless, low cost, and highly accurate if done right, safe (no side effect known), practical, and it requires no contact nor compression, no radiation or venous access (Aksenov et al., 2003; Bamberg, 2002; Gautherie, 1989; Head et al., 2000; Keyserlingk et al., 2000). Infrared breast thermography can increase sensitivity at the critical early detection phase by providing an early warning of an abnormality that is not evident by other approaches (Keyserlingk et al., 2000). It is able to warn women up to 10 years before a cancer is found (Amalu, 2003; Pacific Chiropractic and Research Center). Furthermore, thermography is the only

physical method that mediates significant information on breast physiology (Gautherie, 1989). In contrast to other techniques, its result is independent of nodal status, and unrelated to age, tumor location (right or left breast), and estrogen, progesterone receptor status (Head et al., 2000). Hence, thermogram plays a pivotal role in breast cancer, be it risk assessment (Amalu, 2003), detection, diagnosis, or prognosis (Gautherie, 1989; Head et al., 2000).

Unfortunately, despite of the strengths reported, thermogram is associated with some of the limitations such as environment-dependent, operator-dependent (Fok, Ng, & Tai, 2002; Ng & Fok, 2003), not descriptive (Aksenov et al., 2003; Bamberg, 2002), difficult to interpret (Amalu, 2003), nonspecific (Jones, 1998), inconsistent (Frize, Herry, & Roberge, 2002; Head, Hoekstra, Keyserlingk, Elliott, & Diakides, 2003), and no standard analysis procedure (Ohashi & Uchida, 2000; Kaczmarek & Nowakowski, 2003), as pointed out in *Breast Cancer Detection Demonstration Projects* (BCDDP). As a result, breast thermography is yet to be widely used and is not recommended by National Breast Cancer Centre (National Breast Cancer Centre Position Statement, 2004). Apparently, thermogram performs no better than other modalities. All in all, if the thermography is done right, it offers a very powerful tool for fighting breast cancer. Thus, by providing decision aids using intelligent system (Ng & Fok, 2003; Ng, Fok, Peh, Ng, & Sim, 2002), good and consistent diagnosis performance can be maintained using breast thermography. At the same time, these intelligent tools can lighten the pressures upon the physicians, and ease the burden of examining large number of images (e.g., 1 million pairs of X-ray images per year is needed to be reviewed Kotre, 1993). A summary of the use of complementing breast cancer detection modalities with intelligent tools is given in Table 2.

As shown in Table 2, intelligent tools contribute significantly in improving the breast cancer detection and prognosis. This is consistent with a recent review that computer-aided diagnosis shows incremental improvement in sensitivity (Irwig, Houssami, & van Vliet, 2004). MLP or BP is the favorite algorithm to complement various modalities, in spite of its limitations such as slow learning, likely to be trapped in local minima, etc. SOM is another common adjunct for imaging modalities, albeit its poor classification performance, and high memory requirement. Statistical methods like LDA, Bayesian network, and logistic regression are often applied in assisting diagnosis and prognosis. However, statistical methods are difficult to develop, and oftentimes they work under the assumption that the underlying data is normally distributed. Whereas RBF has heavy computation and memory requirements, decision tree is limited in its representation power due to the use of crisp rule. On the other hand, evolving ANN, although it is able to achieve optimal performance, is time-consuming to develop since it may take a few hundreds to thousands runs before it can find the appropriate parameters. Furthermore, due to the stochastic nature of the algorithm, it may generate

Table 2
Reported accuracy on computer-aided diagnosis

Methods & applications	Accuracy (%)	Training/Sample size
Prognostic factors & BP for breast cancer prognosis (Burke et al., 1994)	85	NM
Prognostic factors & SOM for nodal metastasis detection (Naguib et al., 1996)	55–84	50/81
Prognostic factors & (a) Logistic regression (b) MLP (c) Decision tree for breast cancer survivability prediction (Delen et al., 2005)	(a) 89.2 (b) 91.2 (c) 93.6	10-fold cross validation, 202,932
Patient physiological and history data & (a) ANN (b) Data Employment Analysis (c) LDA for breast cancer diagnosis (Pendharkar et al., 1999)	(a) 81.5 (b) 66.5 (c) 66.1	227/227
Clinical pathological data & MLP & single threshold system for breast cancer prognosis (Gómez-Ruiz et al., 2004)	96	828/1035
FNA & (a) Fuzzy k-NN (b) logic regression (c) MLP for breast cancer prognosis (Seker et al., 2001)	(a) 88 (b) 82 (c) 87	Leave-one-out, 100
FNA & (a) Rank NN (Bagui et al., 2003) (b) Evolving ANN (Land & Albertelli, 1998) (c) Memetic pareto ANN and (d) BP (Abass, 2002) (e) CLFNN (Tan, 2005) (f) DA and (g) MARS and (h) BP and (i) MARS and BP (Chou et al., 2004) (j) Evolving ANN and (k) ANN ensembles (four MLP) (Yao & Liu, 1999) (l) Hybrid fuzzy genetic (Andrés et al., 1999) (m) SVM (Liu et al., 2003) for breast cancer diagnosis	(a) 97 (b) 97.1 (c) 98.1 (d) 97.5 (e) 97.81 (f) 95.91 (g) 97.66 (h) 98.25 (i) 98.25 (j) 96–99.8 (k) 99.99 (l) 97.5 (m) 97.07	(a), (b), (k): 524/699 (d): 400/683 (e), (j), (k), (l): 349/699 (f), (g), (h), (i): 398/569 (m) 547/699
Wavelet features of mammogram & MLP for breast cancer diagnosis (Kocur et al., 1996)	88	NM
Breast cancer tissue image & fuzzy co-occurrence matrix & MLP for breast cancer diagnosis (Cheng et al., 1995)	100	60/90
Biopsy image & (a) RBF (Schnorrenberg et al., 1997) (b) receptive field function and (c) ANN (Schnorrenberg et al., 2000) (d) singular value decomposition & MLP with Levenberg–Marquardt algorithm (Tsapatsoulis et al., 1997) for breast cancer nuclei detection	(a) 83.7–84.6 (b) 76.4–78.1 (c) 79.3–80.7 (d) 76.8	Sensitivity
Mammogram & (a) Evolving ANN (Fogel et al., 1998) (b) DA (Leichter et al., 1996) (c) SOM & MLP (Santos-André & da Silva, 1999) (d) Bayesian belief network (Wang et al., 1999) for breast cancer diagnosis	(a) 84.64–91.96 (b) 93.8 (c) 60 (d) 89	(a) Leave-one-out, 216 (b) NM (c) 247/272 (d) 335/419
(a) Mammogram & LDA for parenchymal patterns identification. (b) Mammogram & one-step rule-based & ANN for breast cancer diagnosis (Huo et al., 1998)	(a) 91 (b) 94	NM
(a) Mammogram & patient history data & ANN for breast cancer diagnosis, (b) and for mammographic invasion prediction (Lo & Floyd, 1999)	(a) 82–86 (b) 77.96	Leave-one-out
Mammogram & patient history data & (a) evolutionary programming & Adaboosting (Land et al., 2000) (b) Constraint satisfaction ANN [Tourassi01] for breast cancer diagnosis	(a) 86.1–87.6 (b) 84	(a) 400/500 (b) 250/500
Mammogram & RBF for (a) abnormalities detection (b) breast cancer diagnosis (Christoyianni et al., 2002)	(a) 88.23 (b) 79.31	(a) 119/238 (b) 119/119
Ipsilateral mammogram & ANN (BP and Kalman filter) for breast cancer diagnosis (Sun et al., 2004)	About 65	60/100
MRI & BP for breast cancer diagnosis (Cardillo et al., 2001)	Improved accuracy	NM
(a) Spectrum of radio frequency echo signals in ultrasound (b) B-mode ultrasound & DA for axillary lymph node classification (Tateishi et al., 1998)	(a) 92.5 (b) 80	NM
Sonography & (a) SOM (Chen et al., 2000) (b) ANN (Lo & Floyd, 1999) for breast cancer diagnosis (Chen et al., 2000)	(a) 85.6 (b) 96	(a) 10-fold cross validation (b) NM
Thermogram & (a) Image histogram & Co-occurrence matrix (Jakubowska et al., 2003) (b) Microwave radiation & Karhunen–Loeve transformation (Varga & De Muynck, 1992) (c) CLFNN (Tan et al., 2004) (d) BP for breast cancer diagnosis	(a) Almost 100 (b) Compared well to physician (c) 74–94 (d) 53–64	(a), (b): NM (c) 39/78 (d) 65/78
Gene expression & k-means clustering & principal component analysis & Bayesian classification tree for (a) lymph-node metastasis and (b) relapse (Huang et al., 2003)	(a) 90 (b) 90	Leave-one-out, (a) 37, (b) 52

Abbreviations: BP: Backpropagation, ANN: Artificial Neural Network, SVM: Support Vector Machine, DA: Discriminant Analysis, LDA: Linear DA, MLP: Multilayer Perceptron, SOM: Self-Organizing Map, NN: Nearest Neighbor, MARS: Multivariate Adaptive Regression Splines, RBF: Radial Basis Function, NM: Not Mentioned.

inconsistent knowledge base. Most of all, these methods (except decision tree) do not provide any explanations for their computations and reasoning. As a result, the physicians have no way to validate the system operation, and hence, they find it difficult to trust the system.

Complementary Learning Fuzzy Neural Network (CLFNN) is therefore proposed to be *Computer-Assisted Intervention* (CAI) for breast thermography. CLFNN is a neuroscience-inspired, evolving, and autonomous fuzzy neural network that based on positive and negative learning. CLFNN not only provides good performance in classification, but also fast in learning. Most importantly, CLFNN offers human-like reasoning as well as intuitive fuzzy rules to explain its computations. Since human observer's image interpretation is often lack of thoroughness and lack of consistency (Bick, 2000), the capacity of CLFNN in providing cognitive interpretation on given thermogram is of great importance for aiding image analysis. Psychophysical evidence demonstrates that even imperfect prompts can enhance human ability in pattern detection (Kotre, 1993). Therefore, CLFNN is believed to enhance the overall accuracy of breast thermography. On the other hand, most of the disease detection works in CAI adopted physic or physically inspired models (Ellis & Peters, 2004), statistical methods such as *Bayesian theory* and *nearest neighbor* (Sajda, Spence, & Parra, 2003), or *Artificial Neural Network* (ANN) (Frigyesi, 2003; Joo, Yang, Moon, & Kim, 2004). These methods however possess some shortcomings: statistical methods and ANN do not justify, and provide no explanation for their computation. As a result, the output is difficult to trust because it comes without reason. As for model-based system, it is difficult to develop, and many a times requires assumption to be made. This applies to statistical methods as well, as many statistical methods assume that the data is normally distributed. Conversely, other than superior accuracy, CLFNN provides positive and negative fuzzy rules to reason its decisions, and this reasoning is closely akin to diagnostician's decision-making process. These rules not only can be used to countercheck physician's diagnosis, they could potentially guide junior physician. Besides, CLFNN can also be adopted to confirm or investigate hypothesis associated with breast cancer such as women having family history of breast cancer belong to high risk group (Cancer Research UK, 2002), temperature difference between left and right breast suggests possible case of cancer (Gautherie, 1989), and so on.

2. FALCON-AART

FALCON-AART is a CLFNN that forms its fuzzy partitions based on visual cortical plasticity, and adjusts its parameters based on psychological theory of learning (Tan, Quek, & Ng, 2004) (for details, see Tan et al., 2004). It generates fuzzy rules autonomously in the form described by Eq. (1).

IF x_1 *is* **A** *and* x_2 *is* **B**, **THEN** y_1 *is* **C** *and* y_2 *is* **D** (1)

The fuzzy rule in Eq. (1) is an example of a system with two inputs and two outputs. It consists of five elements:

1. Input linguistic variables (x_1, x_2).
2. Input linguistic terms (**A**, **B**). This represents fuzzy entities such as tall, short, thin, fat, and so on. FALCON-AART represents input linguistic terms by using trapezoidal membership function.
3. If–Then rule: links the antecedent part (i.e., input linguistic variables and terms) with the consequent part (i.e., output linguistic variables and terms).
4. Output linguistic variables (y_1, y_2).
5. Output linguistic terms (**C**, **D**).

FALCON-ART has five layers and each layer is mapped onto the elements of the fuzzy rule (Fig. 2). Before training commences, FALCON-AART consists of input and output layers only. As training progresses, FALCON-AART evolves and automatically constructs its hidden layer by modified Fuzzy ART algorithm (Tan et al., 2004). This algorithm is based on complementary learning paradigm that comprises positive (learn from positive patterns) and negative learning (learn from negative patterns). The modified fuzzy ART algorithm (known as Another ART) improves Fuzzy ART (Baraldi & Bonda, 1999) by functionally models and incorporates the human visual cortical plasticity. With this, FALCON-AART structural learning becomes a function of time (age), which enables FALCON-AART to alleviate the stability–plasticity dilemma as well as to avoid the problem of generating bad clusters as suffered by most competitive learning algorithm. It dynamically partitions the input and output spaces into trapezoidal fuzzy clusters, and subsequently these clusters are finetuned using modified adaptive back-propagation algorithm (Tan et al., 2004). The tuning is done simultaneously to the slope and the location of fuzzy sets. When new training patterns are presented, the stored cluster will resonate if the new training patterns are sufficiently similar to them. The resonant cluster will then expand to incorporate these patterns using the Another ART algorithm. Training terminates when the mean square errors between two consecutive epochs are sufficiently equal.

The neural memory structure between Layers 2 and 3 is the construct of the complementary learning. Complementary learning refers to positive and negative learning, which is believed to be a mechanism underlies human recognition. When a positive pattern is presented, positive rules will be excited, and negative rules will be inhibited simultaneously, and vice versa. The complementary learning is often practiced in daily life: a child will learn how to recognize an apple more efficiently, if he/she were presented an apple (positive pattern) and other fruits (negative patterns).

Likewise, a radiologist will have to have seen/learned, both abnormal medical image (positive learning) and normal medical images (negative learning), in order for him/her to recognize or analyse the images effectively. Evidences for this complementary learning can be drawn from vari-
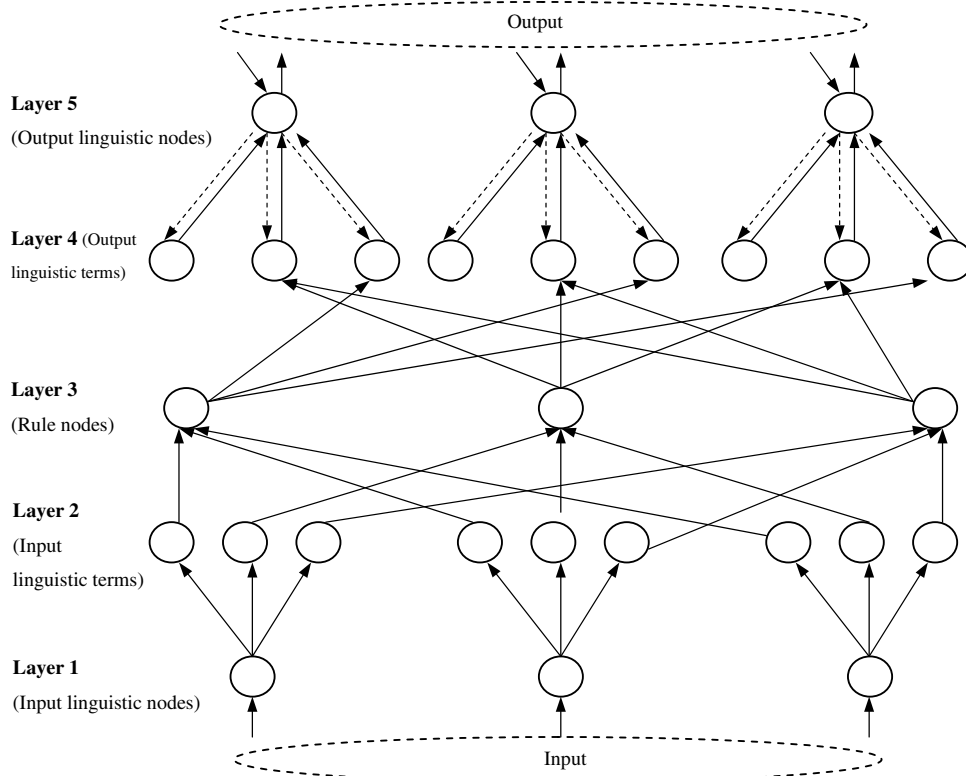
Fig. 2. Architecture of FALCON-AART.

ous neuroscience studies. For instance, *hippocampus* possesses both positive and negative reinforcement signals; the existence of excitatory (positive) and inhibitory (negative) neurotransmitter systems inside human brain, etc. As shown in Fig. 3, different objects are registered into different brain areas, lending further support to the complementary learning conjecture. Hence, whenever a car is presented (positive), only areas registered for car (positive

rules) will be activated, while the areas registered for other objects (negative rules) will be inhibited simultaneously.

Thus, FALCON-AART functionally models the biological complementary learning, and is formalized as Eqs. (2) and (3). Given a positive sample, $\{\mathbf{x}^+ = (x_1, x_2, \ldots, x_I), d = 1\}$, $x \in U$, $d \in V$, and $\mu_{R^+}(x) =$ membership function of positive rule, $\mu_{R^-}(x) =$ membership function of negative rule, then:
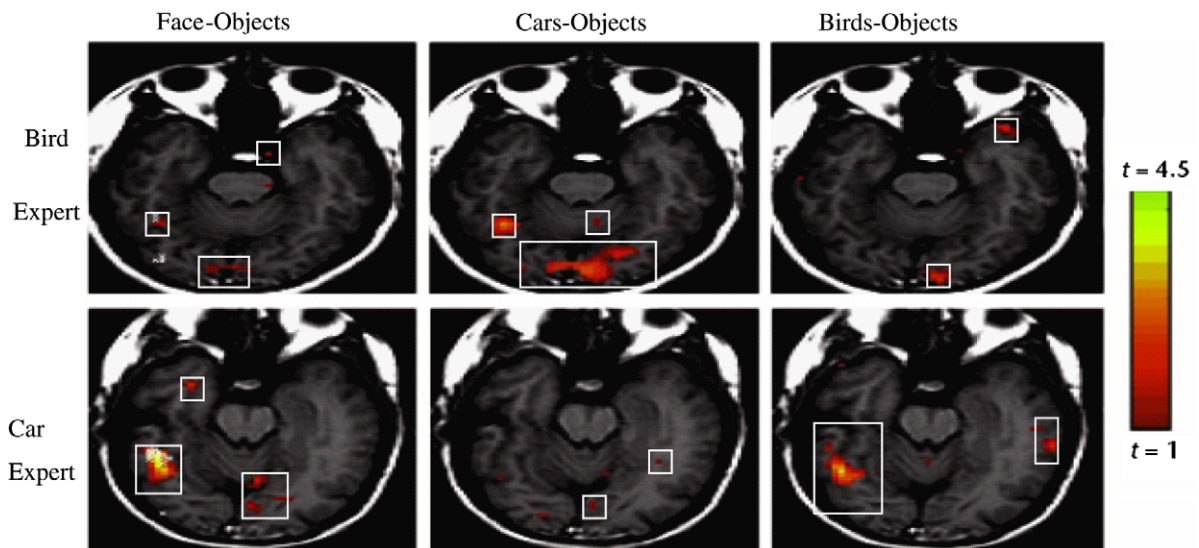

Fig. 3. Slices of fusiform gyrus of car and bird expert in face, car, and bird recognition. The rectangular boxes show the activated areas of brain for different recognition task (Adapted from Gauthier et al., 2000).

$$\mu_{R^+}(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} = \mathbf{x}^+ \\ 0, & \text{otherwise} \end{cases} \qquad (2)$$

$$\mu_{R^-}(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} = \mathbf{x}^- \\ 0, & \text{otherwise} \end{cases} \qquad (3)$$

Hence, whenever a positive sample is presented to the system, $\mu_{R^+}(\mathbf{x}^+) > \mu_{R^-}(\mathbf{x}^-)$, which leads to a correct decision, i.e., $d = 1$.

3. Experiment and results

3.1. Dataset

The thermograms are obtained from voluntary patients at the Singapore General Hospital (SGH) (Ng, Fok, Ng, & Sim, 2001; Ng, Peh, Fok, Ng, & Sim, 2002). The thermograms are captured using the AVIO thermal camera TVS-2000 MkIIST system. The thermography process is shown in Fig. 4.

The patient's thermal image is captured using thermal camera. The imager component of thermal system converts infrared emitted by the object under observation into electrical signals. Subsequently, the processor components collects these signals, store them in frame memory, then displays them on a LCD display, either as real-time sixteen bit color or monochrome thermographic images. The thermogram is stored, and feature extraction is done to compute the temperatures of the left and right breasts using the AVIO software. Example of thermogram is given in Fig. 5.
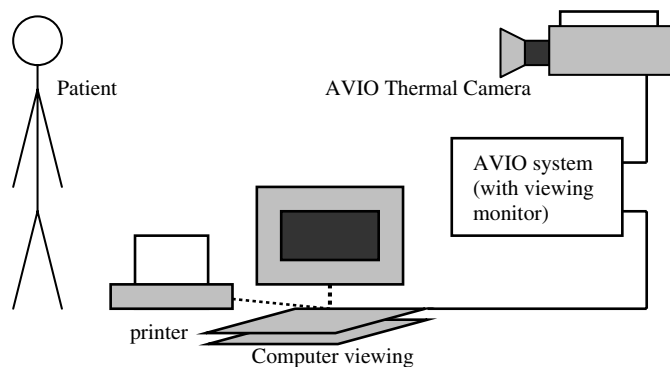
The volunteers are between the ages of 27 and 90. Screening was carried out from 9.00 am to 11.30 pm of the day as this is the most stable period (Gautherie, 1989). All volunteers were briefed the methodology and process of thermography in advance in order to relief them from any possible emotional stress as well as to obtain their consent. They were advised not to put on any powder, ointments, perfume, or any other wipes that will affect the conduction through the skin, around regions to be examined. Before the examination was carried out, volunteers were required to rest for 15–20 min for acclimatization to room temperature upon arrival at the examination room. This is important to keep patients in basal metabolic rate which will result minimal surface temperature changes for satisfactory thermograms. Since standardized ambient conditions are necessary to minimize variations in thermography, the ambient temperature was carefully observed for the examination. The examination environment was a controlled, air-conditioned room maintained at an ambient temperature of 20–22 °C (maximum variation is ±0.1 °C), with humidity between 55% and 65%. Direct draughts are avoided in the areas where the patient is positioned. Volunteers wore loose gowns that do not restrict airflow for equilibration and do not constrict the skin surface during this equilibration period. It was ensured that patients were within the period of the 5th–12th and 21st day after the onset of menstrual cycle as this is the most suitable period for imaging. This is because women body temperature is known to be stable in this period (Gautherie, 1989), and the vascularisation is at basal level with least engorgement of blood vessels (Ng et al., 2001).

Three thermograms were taken for each patient: one front view and two lateral views. There are total of 78 patients with 28 healthy patients, 43 benign tumor patients, and 7 cancer patients. Mean, median, mode, standard deviation and skewness of each breast temperature are extracted from front-view thermograms using histograms of the temperature distribution, and calculated using the *Statistical Package for the Social Sciences* (SPSS). Population of patients is shown in Table 3.

Table 3 shows that carcinoma patients generally have higher breast temperature compared to healthy patients.
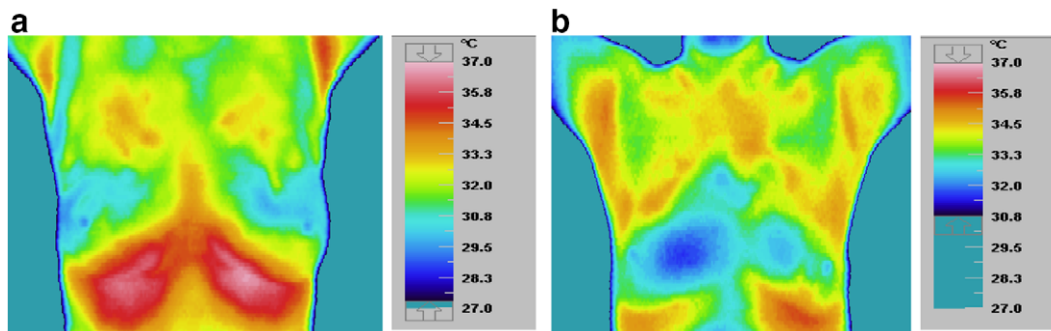

Fig. 4. Thermography process.

Fig. 5. Thermogram of (a) healthy patient-symmetrical temperature (b) unhealthy patient-unsymmetrical temperature.

Table 3
Average mean and modal temperatures of healthy and unhealthy breasts

	Healthy patients	Benign patients	Carcinoma patients
Average mean temperature of normal breast (left and/or right) (°C)	32.66	32.81	33.43
Average mean temperature of abnormal breast (left or right) (°C)	Not available	33.00	33.51
Average modal temperature of normal breast (left and/or right) (°C)	32.67	33.05	33.40
Average modal temperature of abnormal breast (left or right) (°C)	Not available	33.00	33.51

This temperature difference arises because the cancerous breast has higher metabolism. The blood vessels in the vicinity of the tumor are engorged with blood and therefore, cancerous breast emits more heat (Ng et al., 2002).

3.2. Experiment

The experiment is to diagnose whether a patient belongs to normal, benign, or malignant based on breast temperatures extracted from thermogram. Five types of file are created, and three training/testing sets are created for each type of file for cross-validation purpose. Each of the stratified training sets contains randomly selected 50% samples from the dataset, and the remaining unseen samples made up the testing sets. The sets are presented below:

- File FH: contains patient age, family history, hormone replacement therapy, age of menarche, presence of palpable lump, previous surgery/biopsy, presence of nipple discharge, breast pain, menopause at age above 50 years, and first child at age above 30 years.
- File T: contains mean, median, modal, standard deviation and skewness of temperature for left and right breasts.
- File TH: combination of FH and T.
- File TD: contains temperature difference of mean, median, modal, standard deviation and skewness for left and right breasts.
- File TDH: combination of TD and FH.

The averaged performance of FALCON-AART is benchmarked against *Linear discriminant analysis* (LDA) (Hanm & Kamber, 2001), *k-Nearest neighbor* (kNN) (Hanm & Kamber, 2001), *Naïve Bayesian* (Hanm & Kamber, 2001), *logistic regression* (LR) (Hanm & Kamber, 2001), *Self-Organizing Map* (Chen et al., 2000), *Radial Basis Function* (RBF) (Hanm & Kamber, 2001), *Support Vector Machine* (SVM) (Hanm & Kamber, 2001), *C4.5* (Hanm & Kamber, 2001), *Multilayer Perceptrons* (MLP) (Hanm & Kamber, 2001). Apart from that, comparison is made with FALCON-AART ancestors: FALCON-ART (Lin & Lin, 1997) and FALCON-MART (Tung & Quek, 2001). The result is listed in Table 4. Recall refers to the classification accuracy on the training set, whereas predict refers to the classification accuracy on testing set.

It is shown that FALCON-AART outperforms the common methods in medical image analysis and its ancestors in all the training/testing sets. While having good recall and relatively superior generalization capability, the aver-

age training time of FALCON-AART is significantly shorter than MLP, SOM, and LR. Though statistical algorithms require only one pass of training dataset, it does not necessarily means they are faster than FALCON-AART as this depends on the computational complexity of the algorithm. In this particular case, FALCON-AART is as fast as kNN, LDA, SVM, and Naïve Bayesian classifier in learning ($\approx$245 ms). In contrast to statistical methods, FALCON-AART did not make assumption on the data distribution, and this may give superior classification performance even for non-normally distributed data. Note that this result is not comparable to the one in Tables 1 and 2 as this is a different classification task. This classification task involves normal, benign, and malignant whereas the task in Tables 1 and 2 involves only benign and malignant. In other words, from the experimental result shown in Table 4, complementary learning displays superior capacity in multi-class classification than conventional methods.

One significant advantage FALCON-AART offers is the ability to explain its computed output. In contrast to conventional methods, FALCON-AART constructs intuitive positive and negative fuzzy rules dynamically to depict its reasoning process; these rules can be scrutinized by the physicians and decide upon whether to adopt the system suggestion. In addition, accurate rules identified may be used as a guideline for inexperience physicians in diagnosis. As shown in Table 4, rule generation capability of FALCON-AART is better than its ancestor, in which lesser rules are generated but greater accuracy are attained. Some authors have proposed a few criteria for measuring system interpretability: compactness (lesser number of rule in rule base), coverage (every value in universe of discourse should belong to one of the rule), normality (every rule has at least one pattern exhibit full-matching), and so on (Casillas, Cordón, Herrera, & Magdalena, 2003). FALCON-AART learning is a data-centered learning and therefore, it fulfills the coverage and normality criteria. From this experiment, it can be seen that FALCON-AART generates a smaller rule base then its ancestors. Thus, from this aspect, FALCON-AART offers a more interpretable system than its ancestors. Examples of the rules generated are given in Table 5.

As shown in Table 5, fuzzy rules generated by FALCON-AART are highly similar to the diagnostic rules practiced by diagnosticians. Aside from the capacity for uncertainty handling (allowing vagueness in linguistic terms), FALCON-AART rule is relatively more expressive compared to decision-tree rule. FALCON-AART rule

Table 4
Breast cancer diagnosis result (desired values are in bold)

Method		FH	T	TH	TD	TDF
Linear discriminant analysis	Recall (%)	65.79	62.86	88.57	37.14	71.43
	Predict (%)	34.21	47.37	28.95	28.95	36.84
	No. of epoch	1	1	1	1	1
	No. of rules	Not applicable				
Multilayer perceptron	Recall (%)	**82.86**	77.14	97.14	65.71	88.57
	Predict (%)	42.11	55.26	57.89	47.37	42.11
	No. of epoch	100	100	100	100	100
	No. of rules	Not applicable				
Naïve Beyesian classifier	Recall (%)	57.14	57.14	55.88	57.14	54.29
	Predict (%)	54.29	54.29	57.14	54.29	22.86
	No. of epoch	1	1	1	1	1
	No. of rules	Not applicable				
k-Nearest neighbor	Recall (%)	25.71	88.57	97.06	74.29	91.43
	Predict (%)	40	45.71	48.57	42.86	45.71
	No. of epoch	1	1	1	1	1
	No. of rules	Not applicable				
Support vector machine	Recall (%)	62.86	62.86	77.14	54.29	68.57
	Predict (%)	42.11	52.63	57.89	52.63	42.11
	No. of epoch	1	1	1	1	1
	No. of rules	Not applicable				
C4.5	Recall (%)	74.29	74.29	91.43	68.57	88.57
	Predict (%)	42.11	57.89	55.26	44.74	50
	No. of epoch	1	1	1	1	1
	No. of rules	7	**4**	**9**	**4**	**9**
Logistic regression	Recall (%)	80	77.14	100.0	60.0	82.86
	Predict (%)	36.84	57.89	50	44.74	42.11
	No. of epoch	1	1	1	1	1
	No. of rules	Not applicable				
Self organizing map	Recall (%)	77.14	82.86	80	77.14	77.14
	Predict (%)	34.21	50.0	21.05	36.84	34.21
	No. of epoch	2000	2000	2000	2000	2000
	No. of rules	Not applicable				
Radial basis function	Recall (%)	60.0	62.86	57.14	57.14	62.86
	Predict (%)	50.0	55.26	52.63	44.74	50.0
	No. of epoch	1	1	1	1	1
	No. of rules	Not applicable				
FALCON-ART	Recall (%)	85.71	71.43	65.71	70.37	88.57
	Predict (%)	50.0	52.63	52.63	51.28	55.26
	No. of epoch	50	50	50	50	50
	No. of rules	137	189	148	134	180
FALCON-MART	Recall (%)	37.14	**97.14**	**100.0**	70.37	97.14
	Predict (%)	34.21	65.79	52.63	51.28	52.63
	No. of epoch	6	17	10	6	8
	No. of rules	34	23	78	54	31
FALCON-AART	Recall (%)	77.14	94.29	97.14	**100.0**	**100.0**
	Predict (%)	**65.79**	**65.79**	**68.42**	**55.26**	**63.16**
	No. of epoch	4	4	4	4	10
	No. of rules	**4**	22	30	38	31

encapsulates unnecessary details using linguistic term, and allows the use of linguistic hedges such as "very", "rather", etc. Moreover, rules generated by FALCON-AART do not have the confusing repeated antecedent term as in decision tree. Furthermore, because FALCON-AART adopts complementary learning, positive and negative rules are generated. This, aside from better classification performance,

models the problem space closer than positive or negative learning systems (system with only positive or negative rule base) because no assumptions are made for the uncovered space by the rule base. Fig. 6 depicts the FALCON-AART reasoning process.

As shown in Fig. 6, the reasoning process of FALCON-AART is closely akin to how a diagnosis is made: A diag-

Table 5
Fuzzy rules generated by FALCON-AART

Fuzzy rules (FALCON-AART)	Crisp rules (C4.5)	Diagnostic rule Varga and De Muynck (1992)
IF *mean difference* is small, **AND** *median difference* is rather big, **AND** *modal difference* is medium, **AND** *standard deviation* is big, **AND** *skewness difference* is very small, **THEN** Normal	**IF** *modal difference* < 0.1 **AND** *modal difference* < 0.03 **AND** *mean difference* < 0.25 **AND** *median difference* < 0.24 **AND** *skewness difference* < 0.05 **THEN** Normal	**IF** temperature is generally 0.3–1.5 °C higher than the surrounding normal tissues, **THEN** Tumor
IF *mean difference* is medium, **AND** *median difference* is very small, **AND** *modal difference* is very big, **AND** *standard deviation* is medium, **AND** *skewness difference* is small, **THEN** Benign	**IF** *modal difference* $\geqslant 0.1$ **AND** *median difference* < 0.07 **THEN** Benign	**IF** temperature drops linearly with time **THEN** normal tissue
IF *mean difference* is very small, **AND** *median difference* is small, **AND** *modal difference* is very big, **AND** *standard deviation* is big, **AND** *skewness difference* is medium, **THEN** Malignant	**IF** *modal difference* < 0.1 **AND** *modal difference* < 0.03 **AND** *mean difference* $\geqslant 0.25$ **AND** *modal difference* < 0.02 **AND THEN** Malignant	**IF** temperature remains high but sometimes drops a little **THEN** Tumor

1) Presenting a Cancer sample.

$\mathbf{x} = [0.070594412, 0.153538261, 0.190580085, 0.197984934, 0.120725836]$

2) Compute matching degree of antece -dents of the 3 rules.

The fuzzy sets describing the antecedents:

$\mathbf{u}_1 = [0.2549 \quad 0.121549 \quad 0.139751 \quad 0.195671 \quad 0.0520233]$

$\mathbf{v}_1 = [0.2555 \quad 0.121577 \quad 0.139798 \quad 0.195848 \quad 0.171412]$

Therefore, the matching degree: $\mathbf{z}_1 = [0 \quad 0.3 \quad 0.2 \quad 0.7 \quad 1]$

3) Compute the overall rule matching degree for every rule.

$z_{rule\,1} = 0.44$

$z_{rule\,2} = 0.22$

$z_{rule\,3} = 1$

The fuzzy sets describing the antecedents:

$\mathbf{u}_2 = [0.384539 \quad 0.00451737 \quad 0.0746796 \quad 0.125762 \quad 0.111001]$

$\mathbf{v}_2 = [0.38454 \quad 0.00451737 \quad 0.0746825 \quad 0.125764 \quad 0.111003]$

Therefore, the matching degree: $\mathbf{z}_2 = [0 \quad 0 \quad 0 \quad 0.3 \quad 0.8]$

4) Derive the consequent:

$(u_1, v_1) = (1,1);$
$(u_2, v_2) = (1,1);$
$(u_3, v_3) = (1,1)$

The fuzzy sets describing the antecedents:

$\mathbf{u}_3 = [0.0705747 \quad 0.153534 \quad 0.190578 \quad 0.197985 \quad 0.120725]$

$\mathbf{v}_3 = [0.0706131 \quad 0.153546 \quad 0.190581 \quad 0.197985 \quad 0.120725]$

Therefore, the matching degree: $\mathbf{z}_3 = [1 \quad 1 \quad 1 \quad 1 \quad 1]$

5) Conclusion: $y_1 = \dfrac{1+1}{2} \times 0.44 = 0.44 \qquad y_2 = \dfrac{1+1}{2} \times 0.22 = 0.22 \qquad y_3 = \dfrac{1+1}{2} \times 1 = 1$
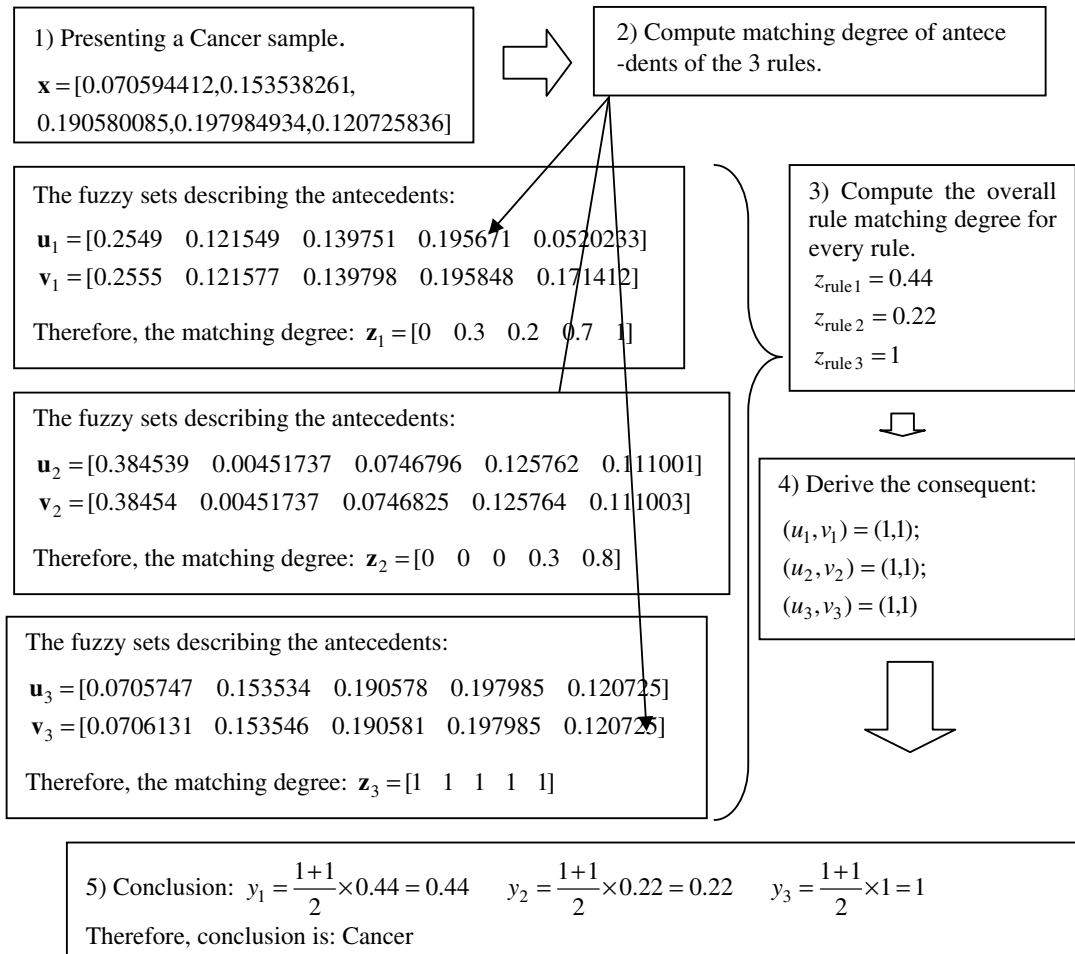
Therefore, conclusion is: Cancer

Fig. 6. Reasoning process of FALCON-AART.

nostician will first observe (presents sample), generates a set of hypotheses (a set of rules), evaluates each hypotheses (compute matching degree of rules), and subsequently derives the conclusion. This human-like reasoning, together with the fuzzy rules generated, which provide insights and interpretations to the thermograms, are useful to aid diagnosti-

cian. Table 6 shows the similarity between FALCON-AART and thermogram analyst's reasoning process. As shown, there is one-to-one mapping of the reasoning process, suggesting the closeness between the two reasoning processes. This is paramount as it facilitates the physicians in analyzing or validating a system, in that he/she can do so

Table 6
FALCON-AART and analyst reasoning

Steps	FALCON-AART	Analyst
1	Take in the extracted features from thermogram	Examine the thermogram. Looks for abnormal heat patterns, temperature variations, etc.
2	Compare the feature values with own positive and negative diagnostic rules (knowledge/ experiences). Computes their matching degree (firing strength/ similarity)	Compare the examined thermogram with previous benign (negative) and malignant (positive) thermograms. Judges and determines their similarity based on own diagnostic rules and experiences
3	Select the rule with maximum matching degree, and inhibits others	Select the knowledge that best describes the current situation. Eliminates those hypotheses that are not relevant
4	Determine the consequent linked by the winning rule	Determine the conclusion derived from the knowledge applied
5	Perform defuzzification and outputs the conclusion	Give the diagnostic conclusion and decision

in his/her familiar terms, as well as in his/her familiar thought process.

FALCON-AART can be used to assess/affirm certain medical hypothesis as well. For example, from Table 4, one can see that the classification accuracy of training/testing set using only breast temperatures alone is lower than that of using breast temperatures and family history. This confirms that family history is important risk factor for breast cancer, lending support to the hypothesis that women who have family history of breast cancer belong to the high-risk group. Another example: the performance of FALCON-AART trained on files TH (FA_{TH}) and TDF (FA_{TDF}) seemed to be inconsistent with the belief that temperature asymmetry between left and right breast suggests possible case of cancer. This happens because the classification task is to classify three classes, instead of classifying out the cancerous case. In fact, the temperature asymmetry between left and right breast may be more useful in determining the stages of cancer instead of cancer detection (Usuki, Maeta, Maeba, & Wakabayashi, 2000). Nevertheless, the result of detecting cancerous case is illustrated in Task 1 of Table 7.

Though both FA_{TH} and FA_{TDF} attain same accuracy, FA_{TDF} is better when assessed using *Receiver Operating Curve* (ROC) plot, which suggests that it is relative easier to classify using temperature asymmetry of left and right breasts. The 45° line signifies the random guessing. As shown in Fig. 7, FALCON-AART trained on either files deviates far away from the 45° line, achieving good performance for breast cancer detection. The *Area Under the Curve* index (A_Z) is often used in ROC analysis. $A_Z = 0.5$ symbolizes random guessing, and the closer A_Z is to 1.0, the better the classifier is. A_Z for FA_{TH} and FA_{TDF} are 0.867 and 0.93, respectively, hence, confirming that asym-
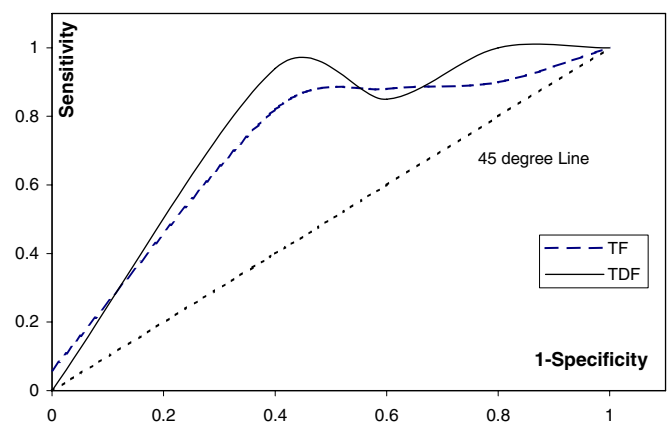

Fig. 7. ROC plot of FALCON-AART trained on files TF and TDF.

metry temperature between the left and right breast is an alarm for breast cancer, and the fact that FALCON-AART is a competent classifier.

Thermogram is often employed to detect the presence of breast tumor. Hence, experiment to classify patient with breast tumor is conducted using FALCON-AART. The result is summarized in Task 2 of Table 7. The experimental result reveals that FALCON-AART can detect patient with breast tumor accurately. Therefore, FALCON-AART could assist the physicians in identifying suspected cases where follow-ups are needed. With overall performance close to 90%, good recall and generalization capability is exhibited by FALCON-AART.

Sometimes, it is desired to classify benign and malignant breast cancer. Misdiagnose benign breast tumor as malignant causes unnecessary physical and emotional agony, because the only way to remove breast tumor is surgical

Table 7
Performance of FALCON-AART on breast thermography

Tasks		Recall (%)	Predict (%)	Sensitivity (%)	Specificity (%)	No. of epoch	No. of rules
1. Cancer detection	TH	100.0	94.74	100.0	60.0	4	14
	TDF	100.0	94.74	100.0	60.0	7	21
2. Breast tumor detection	TH	100.0	84.0	33.33	90.91	11	10
	TDF	100.0	71.05	76.0	61.54	11	17
3. Breast tumor classification	TH	100.0	88.0	33.33	95.45	4	8
	TDF	100.0	84.0	33.33	90.91	11	10

biopsy. On the other hand, misdiagnose malignant breast tumor as benign brings fatal consequences. Thus, it is required to diagnose breast tumor as accurate as possible. The experimental result listed in Task 3 of Table 7 demonstrates that FALCON-AART is able to assist in this diagnostic task as well. Giving an overall accuracy about 93%, FALCON-AART demonstrates its competency in tumor classification task. This shows that complementary learning paradigm is a promising recognition approach.

From the results presented in Tables 1, 2 and 7, complementary learning exhibits itself as a promising tool for aiding breast cancer diagnosis. Applying FALCON-AART with thermogram shows an improved performance in cancer detection as well as breast tumor classification. This confluence of thermography and CLFNN subsides the problem of high variability in accuracy of breast thermogram analysis. Besides, sensitivity and specificity are offered as high/higher than the reported accuracy on breast thermography alone, as well as other modalities. However, CLFNN is not to replace, rather, is to complement the breast thermography and to assist the physicians in breast cancer diagnosis. The contribution of CLFNN-breast thermography in enhancing the consistency of breast cancer diagnosis accuracy is believed to bring forth better patient outcome. Comparing the results of Tables 2 and 7, confluence of CLFNN and breast thermography shows a superior performance in breast cancer detection over different conventional methods in medical diagnosis and medical imaging analysis. Medley of CLFNN and breast thermography gives as accurate result, if not better, compared to other combinations of ANN and breast imaging modalities in tumor classification and detection. In general, CLFNN has relatively good generalization capability, in that it can classify well using only a small fraction of the data. Together, this supports the application of CLFNN and breast thermography. This also suggests that the confluence of breast thermography and CLFNN is a promising system for fighting breast cancer.

4. Discussion and conclusion

In this study, it is shown that CLFNN complements breast thermography in various ways. The combination of breast thermography and CLFNN gives better or more consistent result than using breast thermography alone. Whether it is cancer detection, tumor classification or breast-cancer diagnosis (multi-class problem), CLFNN outperforms conventional methods, showing the strength of complementary learning in recognition task. FALCON-AART assists the physicians in different diagnostic tasks by providing relative accurate decision support, and hence could potentially enhance patient outcome. FALCON-AART not only gives superior result than conventional methods, but it also offers intuitive positive and negative fuzzy rules to explain its reasoning process. FALCON-AART satisfies the criteria of an interpretable system: normality, compactness, coverage, and therefore is a more interpretable system. The rules generated are useful because it gives insight to the problem space, provides simple cognitive interpretation of medical image, and could potentially serve as guidelines or arguments for its decision to the physicians. Apart from assisting physician in diagnosis, FALCON-AART can also be used to investigate or to support hypothesis associated with the problem domain, i.e., *concept validation* (Qi & Diakides, 2003). In this study, only two hypotheses were analyzed. In future, more hypotheses can be assessed using CLFNN by proper experiment setup. Examples are *thermal challenge test* (Eccles, 2003), *cold stress* (Usuki et al., 2000) or *cooling-rewarming tests* (Gautherie, 1999) (outside cooling of the breast will increase the temperature contrast if the breast is cancerous), *injection of vasoactive substances* (Gautherie, 1999), *microwave or ultrasonic irradiation* (Gautherie, 1999) and so on. Likewise, FALCON-AART can be applied with advanced technologies, which provides more information in thermography: *dynamic thermography* (Ohashi & Uchida, 2000), *3-dimensional thermography* (Aksenov et al., 2003), or *thermal texture map* (Hassan, Hattery, & Gandjbakhche, 2003), or *Dynamic Area Telethermometry* (DAT) (Anbar et al., 2000). Conversely, FALCON-AART can complement thermogram in other application areas such as injuries monitoring (Bamberg, 2002), neurology, vascular disorders (e.g., diabetes), rheumatic diseases, tissue viability, oncology (especially breast cancer), dermatological disorders, neonatal, ophthalmology, surgery (Jones, 1998), as well as *Severe Acute Respiratory Syndrome* (SARS) (Diakides & Diakides, 2003). Alternatively, CLFNN can be used to complement other medical imaging modalities such as MRI, MRS, PET, etc., as well as to serve as a concept validation tool for techniques such as *nipple fluid bFGF* (Liu, Wang, Chang, Barsky, & Nguyen, 2000), *Electrical Impedance Tomography* (EIT) (Cherepenin et al., 2001), etc. In current study, FALCON-AART does not perform feature analysis, which is an important area that may improve the system performance and deserved to be studied, as recognition requires one to make decision based on some ''important features''. Moreover, performing feature analysis can reduce the number of antecedents of the rule, and hence improve the interpretability of the system. This will be investigated in future.

References

Abass, H. A. (2002). An evolutionary artificial neural networks approach for breast cancer diagnosis. *Artificial Intelligence in Medicine, 25*, 265–281.

Aksenov, P., Clark, I., Grant, D., Inman, A., Vartikovski, L., & Nebel, J. C. (2003). 3D thermography for quantification of heat generation resulting from inflammation. *8th 3D modelling symposium*. Paris, France.

Amalu, W. C. (2003). A review of breast thermography. International Academy of Clinical Thermology. Available from http://www.iact-org.org/articles/articles-review-btherm.html.

American Cancer Society (2005). Cancer facts and figures 2005. Atlanta: American Cancer Society.

Anbar, M., Milescu, L., Brown, C., Carty, C., AlDulaimy, K., Geronimo, C., et al. (2000). In *Proceedings of 22nd annual international conference*

of the IEEE Engineering in Medicine and Biology (Vol. 2, pp. 1215–1218).

Andrés C., Reyes, P., & Sipper M. (1999). Designing breast cancer diagnostic systems via a hybrid fuzzy-genetic methodology. In *Proceedings of the 8th IEEE international conference fuzzy system* (Vol. 1. pp. 135–139).

Bagui, S. C., Bagui, S., Pal, K., & Pal, N. R. (2003). Breast cancer detection using rank nearest neighbor. *Pattern Recognition, 36*, 25–34.

Bamberg, P. (2002). *Infrared thermography*. Biomedical engineering seminar. University of Rhode Island.

Baraldi, A., & Bonda, P. (1999). A survey of fuzzy clustering algorithms for pattern classification-part B. *IEEE Transactions on System, Man and Cybernetic, 29*(6), 786–801.

Barton, M. B. (2002). *The clinical breast examination*. Harvard Pilgrim Health Care, Harvard Medical School.

Berry, D. A., Iversen, E. S., Jr., Gudbjartsson, D. F., Hiller, E. H., et al. (2002). BRCAPRO validation, sensitivity of genetic testing of BRCA1/BRCA2, and prevalence of other breast cancer susceptibility genes. *Journal of Clinical Oncology, 20*(11), 2701–2712.

Bick, U. (2000). Computer-assisted data analysis in breast imaging. *European Radiology, 10*(9), F1.

Bond, E. J., Li, X., Hagness, S. C., & Van Venn, B. D. (2003). Microwave imaging via space-time beamforming for early detection of breast cancer. *IEEE Transactions on Antennas and Propagation, 51*(8), 1690–1705.

Boyd, N. F., Byng, J. W., Jong, R. A., et al. (1995). Quantitative classification of mammographic densities and breast cancer risk. *Journal of the National Cancer Institute, 87*, 670–675.

Brem, R. F., Kieper, D. A., Rapelyea, J. A., & Majewski, S. (2003). Evaluation of a high-resolution, breast-specific, small-field-of-view gamma camera for the detection of breast cancer. *Nuclear Instruments and Methods in Physics Research A, 497*, 39–45.

Brenner, R. J., Bassett, L. W., Fajardo, L. L., Dershaw, D. D., et al. (2001). Stereotactic core-needle breast biopsy: A multi-institutional prospective trial. *Radiology, 218*, 866–872.

Burke, H. B., Philip, H., Goodman, H., & Rosen, D. B. (1994). Neural networks for measuring cancer outcomes. In *Proceedings of IEEE 10th advanced technologies in instrumentation and measurement technology conference IMTC/94* (Vol. 1(1), pp. 157–159).

Cancer Research UK. (2002). About breast cancer: Risks and causes. Available at http://www.cancerhelp.org.uk/help/default.asp?page=3293.

Cardillo, F. A., Starita, A., Caramella, D., & Cilotti, A. (2001). A neural tool for breast cancer detection and classification in MRI. In *Proceedings of the 23rd annual EMBS international conference* (pp. 2733–2736).

Casillas, J., Cordón, O., Herrera, F., & Magdalena, L. (Eds.). (2003). *Interpretability issues in fuzzy modeling. Studies in fuzziness and soft computing* (Vol. 128). Heidelberg: Springer-Verlag.

Cecil, K. M., Siegelman, E. S., & Lenkinski, R. E. (2001). The evaluation of human breast lesions with magnetic resonance imaging and proton magnetic resonance spectroscopy. *Breast cancer Research and Treatment, 68*(1), 45–54.

Chen, D., Chang, R., & Huang, Y. (2000). Breast cancer diagnosis using self-organizing map for sonography. *Ultrasound in Medicine & Biology, 26*(3), 405–411.

Cheng, H. D., Chen, C. H., & Freimanis, R. I. (1995). A Neural Network for breast cancer detection using fuzzy entropy approach. *IEEE International Conference on Image Processing, 3*, 141–144.

Cherepenin, V., Karpov, A., Korjenevsky, A., Kornienko, V., Mazaletskaya, A., Mazourov, D., et al. (2001). *Physiological Measurement, 22*, 9–18.

Chou, S., Lee, T., Shao, Y. E., & Chen, I. (2004). Mining the breast cancer pattern using artificial neural networks and multivariate adaptive regression splines. *Expert Systems with Applications, 27*, 133–142.

Christoyianni, I., Koutras, A., Dermatas, E., & Kokkinakis, G. (2002). Computer aided diagnosis of breast cancer in digitized mammograms. *Computerized Medical Imaging and Graphics, 26*, 309–319.

Del Guerra, A., Di Domenico, G., Fantini, A., Gambaccini, M., et al. (2003). A dedicated system for breast cancer study with combined SPECT-CT modalities. *Nuclear Instruments and Methods in Physics Research, 497*, 129–343.

Delen, D., Walker, G., & Kadam, A. (2005). Predicting breast cancer survivability: A comparison of three data mining methods. *Artificial Intelligence in Medicine, 34*(2), 113–127.

Delle Chiaie, L., & Terinde, R. (2004). Three-dimensional ultrasound-validated large-core needle biopsy: Is it a realiable method for the histological assessment of breast lesions? *Ultrasound in Obstetrics and Gynecology, 23*(4), 393–397.

Diakides, N. A., & Diakides, M. (2003). Infrared imaging in medicine worldwide. *1st international conference of Thermal Texture Maps (TTM) technology in medicine and engineering*. September 23–25, Houston, TX, USA.

Eccles, N. (2003). Thermography – Its role in early breast cancer detection and pain monitoring. Institute of Complementary Medicine Journal. Available at http://www.icmedicine.co.uk/journal/2004/feb/002.htm.

Ellis, R. E., & Peters, T. (2004). Editorial. *Medical Image Analysis, 8*(4), 167.

Elmore, J. G., Wells, C. F., Carol, M. P. H., et al. (1994). Variability in radiologists interpretation of mammograms. *New England Journal of Medicine, 331*(22), 1493.

Fajardo, L. L., Davis, J. R., Wiens, J. L., & Trego, D. C. (1990). Mammography-guided stereotactic fine-needle aspiration cytology of nonpalpable breast lesions: Prospective comparison with surgical biopsy results. *American Journal of Roentgenology, 155*, 977–981.

Fear, E. C., Hagness, S. C., Meaney, P. M., Okoniewski, M., & Stuchly, M. A. (2002). Near-field imaging for breast tumor detection. *IEEE Microwave Magazine, 3*, 48–55.

Fletcher, S. W., Black, W., Harris, R., Rimer, B. K., & Shapiro, S. (1993). Report of the international workshop on screening for breast cancer. *Journal of the National Cancer Institute, 85*, 1644–1656.

Fogel, D. B., Wasson, E. C., III, Boughton, E. M., & Porto, V. W. (1998). Evolving artificial neural network for screening features from mammograms. *Artificial Intelligence in Medicine*, 317–326.

Fok, S. C., Ng, E. Y. K., & Tai, K. (2002). Early detection and visualization of breast tumor with thermogram and neural network. *Journal of Mechanics in Medicine and Biology, 2*(2), 185–195.

Foster, K. R. (1998). Thermographic detection of breast cancer. *IEEE Engineering in Medicine and Biology Magazine, 17*, 10–14.

Frigyesi, A. (2003). An automated method for the detection of pulmonary embolism in V/Q-scans. *Medical Image Analysis, 7*(2), 341–349.

Frize, M., Herry, C., & Roberge, R. (2002). Processing of thermal images to detect breast cancer: Comparison with previous work. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society] EMBS/BMES Conference. In *Proceedings of the Second Joint Engineering in Medicine and Biology* (Vol. 2, pp. 1159–1160).

Gautherie, M. (1989). *Atlas of breast thermography*. San Doneto Milanses, Italy: PADUSA s.r.l.

Gautherie, M. (1999). Thermopathology of breast cancer: Measurement and analysis of in vivo temperature and blood flow. *Annal New York Academy of Science*, 383–415.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3*(2), 191–197.

Glickman, T. A., Filo, O., Nachaliel, U., Lenington, S., Amin-Spector, S., & Ginor, R. (2002). Novel EIS postprocessing algorithm for breast cancer diagnosis. *IEEE Transactions on Medical Imaging, 21*(6), 710–712.

Gohagan, J. K., Rodes, N. D., Blackwell, C. W., Darby, W. P., et al. (2004). Individual and combined effectiveness of palpation, thermography and mammography in breast cancer screening. *Preventive Medicine, 9*(6), 713–721.

Gómez-Ruiz, J. A., Jerez-Aragonés, J. M., Muñoz-Pérez, J., & Alba-Conejo, E. (2004). A neural network based model for prognosis of early breast cancer. *Applied Intelligence, 20*, 231–238.

Hanm, J., & Kamber, N. (2001). *Data mining: Concepts and techniques*. San Francisco: Morgan Kanfmann Publishers.

Hassan, M., Hattery, D., & Gandjbakhche, A. H. (2003). Thermal texture map—A new technique for disease assessment and treatment monitoring. *1st International conference of Thermal Texture Maps (TTM) technology in medicine and engineering*. September 23–25, Houston, TX, USA.

He, Q., & Shkarin, P. (1999). Proton magnetic resonance spectroscopy and imaging of human breast cancer by selective multiple quantum coherence transfer. In *Bioengineering, proceedings of the Northeast conference* (pp. 102–103).

Head, J. F., Hoekstra, P., Keyserlingk, J., Elliott, R. L., & Diakides, N. A. (2003). Comparison of breast infrared imaging results by three independent investigators. In *Proceedings of the 25th annual international conference of the IEEE EMBS* (pp. 1125–1128).

Head, J. F., Wang, F., Lipari, C. A., & Elliott, R. L. (2000). The important role of infrared imaging in breast cancer. *IEEE Engineering in Medicine and Biology Magazine, 19*(3), 52–57.

Houssami, N., Irwig, L., Simpson, J. M., McKessar, M., Blome, S., & Noakes, J. (2003). Sydney breast imaging accuracy study: Comparative sensitivity and specificity of mammography and sonography in young women with symptoms. *American Journal of Roentgenology, 180*, 935–940.

Huang, E., Cheng, S. H., Dressman, H., Pittman, J., et al. (2003). Gene expression predictors of breast cancer outcomes. *The Lancet, 361*, 1590–1596.

Huo, Z., Giger, M. L., Vyborny, C. J., Olopade, F. I., & Wolverton, D. E. (1998). Computer-aided diagnosis: Analysis of mammographic parenchymal patterns and classification of masses on digitized mammograms. In *Proceedings of the 20th annual international conference of the IEEE Engineering in Medicine and Biology Society* (Vol. 20(2), pp. 1017–1020).

Imaginis, Breast Cancer Diagnosis. (2004). Available from http://imaginis.com/breasthealth/biopsy/biopsy_methods.asp.

Irwig, L., Houssami, N., & van Vliet, C. (2004). New technologies in screening for breast cancer: A systematic review of their accuracy. *British Journal of Cancer, 90*, 2118–2122.

Jakubowska, T., Wiecek, B., Wysocki, M., & Drews-Peszynski, C. (2003). Thermal signatures for breast cancer screening comparative study. In *Proceedings of the 25th annual international conference of the IEEE EMBS* (pp. 1117–1120).

Jones, B. F. (1998). A reappraisal of the use of infrared thermal image analysis in medicine. *IEEE Transactions on Medical Imaging, 17*(6), 1019–1027.

Joo, S., Yang, Y. S., Moon, W. K., & Kim, H. C. (2004). Computer-aided diagnosis of solid breast nodules: Use of an artificial neural network based on multiple sonographic features. *IEEE Transactions on Medical Imaging, 23*(10), 1292–1300.

Kaczmarek, M., & Nowakowski, A. (2003). Analysis of transient thermal processes for improved visualization of breast cancer using IR imaging. In *Proceedings of the 25th annual international conference of the IEEE EMBS* (Vol. 2(2), pp. 1143–1146).

Keyserlingk, J. R., Ahlgren, P. D., Yu, E., Belliveau, N., & Yassa, M. (2000). Functional infrared imaging of the breast. *Journal of IEEE Engineering in Medicine and Biology, 19*(3), 30–31.

Kocur, C. M., Rogers, S. K., Myers, L. R., Burns, T., et al. (1996). Using neural networks to select wavelet features for breast cancer diagnosis. *IEEE Engineering in Medicine and Biology, 15*(3), 95–102.

Kosmas, P., Rappaport, C. M., & Bishop, E. (2004). Modeling with the FDTD method for microwave breast cancer detection. *IEEE Transactions on Microwave Theory and Techniques, 52*(8), 1890–1897.

Kotre, J. (1993). Image processing in the fight against breast cancer. *Engineering Science and Education Journal, 2*(1), 41–46.

Land W., Jr., Masters, T., & Lo, J. (2000). Application of a new evolutionary programming/adaptive boosting hybrid to breast cancer diagnosis. In *Proceedings of the 2000 congress on evolutionary computation* (Vol. 2, pp. 1436–1442).

Land, W. H., & Albertelli, L. E. (1998). Breast cancer screening using evolved neural networks. *Proceedings of the IEEE International Conference on Systems, Man & Cybernetics, 2*, 1619–1624.

Lay, S. F., Crump, J. M., Frykberg, E. R., Goedde, T. A., & Copeland, E. M. (1990). Breast biopsy – Changing patterns during a five-year period. *The American Surgeon, 56*(2), 79–85.

Leichter, I., Bamberger, P., Novak, B., Fields, S., Lederman, R., & Behar, D. (1996). Computerized diagnostics in digital mammography. *19th Convention of electrical and electronics engineers in Israel. Proceedings* (pp. 406–408).

Levine, E. A., Freimanis, R. I., Perrier, N. D., Morton, K., et al. (2003). Positron emission mammography: Initial clinical results. *Annals of Surgical Oncology, 10*, 86–91.

Lewin, J. M., D'Orsi, C. J., Hendrick, R. E., Moss, L. J., Isaacs, P. K., Karellas, A., et al. (2002). Clinical comparison of full-field digital mammography and screen-film mammography for detection of breast cancer. *American Journal of Roentgenology, 179*, 671–677.

Lin, C. J., & Lin, C. T. (1997). An ART-based fuzzy adaptive learning control network. *IEEE Transaction on Fuzzy Systems, 5*, 477–496.

Liu, H. X., Zhang, R. S., Luan, F., Yao, X. J., Liu, M. C., Hu, Z. D., et al. (2003). Diagnosing breast cancer based on support vector machines. *Journal of Chemical Information and Computer Sciences, 43*, 900–907.

Liu, Y., Wang, J. L., Chang, H., Barsky, S. H., & Nguyen, M. (2000). Breast-cancer diagnosis with nipple fluid bFGF. *The Lancet, 356*, 567.

Lo, J. Y., & Floyd, C. E., Jr. (1999). Application of artificial neural networks for diagnosis of breast cancer. In *Proceedings of the 1999 congress on evolutionary computation* (Vol. 3, pp. 1755–1759).

Lucas, J. H., & Cone, D. L. (2003). *Breast cyst aspiration*. American Family Physician, November 15.

Malich, A., Böhm, T., Facius, M., Kleinteich, I., Fleck, M., Sauner, D., et al. (2003). Electrical impedance scanning as a new imaging modality in breast cancer detection—a short review of clinical value on breast application, limitations and perspectives. *Nuclear Instruments and Methods in Physics Research A, 497*, 75–81.

McDonald, S., Saslow, D., & Alciati, M. H. (2004). Performance and reporting of clinical breast examination: A review of the literature. *CA–A Cancer Journal for Clinicians, 54*, 345–361.

Meyer, J. E., Smith, D. N., Lester, S. C., Kaelin, C., Di Piro, P. J., Denison, C. M., et al. (1999). Large-core needle biopsy of nonpalpable breast lesions. *J. Am. Med. Assoc., 281*, 1638–1641.

Moore, S. K. (2001). Better breast cancer detection. *IEEE Spectrum, 38*(5), 50–54.

Moses, W. W. (2004). Positron emission mammography imaging. *Nuclear Instruments and Methods in Physics Research: Section A, 525*(1–2), 249–252.

Murthy, K., Aznar, M., Thompson, C. J., Loutfi, A., Lisbona, R., & Gagnon, J. H. (2000). Results of preliminary clinical trails of the positron emission mammography system PEM-I: A dedicated breast imaging system producing glucose metabolic images using FDG. *Journal of Nuclear Medicine, 41*(11), 1851–1858.

Naguib, R. N. G., Adams, A. E., Horne, C. H. W., Angus, B., Sherbet, G. V., & Lennard, T. W. J. (1996). The detection of nodal metastasis in breast cancer using neural network techniques. *Physiological Measurement, 17*, 297–303.

National Breast Cancer Centre Position Statement. (2004). Available from http://www.nbcc.org.au/media/nbccposition.html.

Ng, E. Y. K., & Fok, S. C. (2003). A framework for early discovery of breast tumor using thermography with artificial neural network. *The Breast Journal, 9*(4), 341–343.

Ng, E. Y. K., Fok, S. C., Ng, F. C., & Sim, L. S. J. (2001). *Breast tumor detection by thermography*. Applied research fund (RG69/98). School of Mechanical and Production Engineering, Nanyang Technological University.

Ng, E. Y. K., Fok, S. C., Peh, Y. C., Ng, F. C., & Sim, L. S. J. (2002). Computerized detection of breast cancer with artificial intelligence and thermograms. *Journal of Medical Engineering & Technology, 26*(4), 152–157.

Ng, Y. K., Peh, Y. C., Fok, S. C., Ng, F. C., & Sim, L. S. J. (2002). Noninvasive diagnosis of breast cancer using thermography with artificial neural network. *Journal of medical Engineering and Technology, 26*(4), 152–157.

Ohashi, Y., & Uchida, I. (2000). Applying dynamic thermography in the diagnosis of breast cancer. *IEEE Engineering in Medicine and Biology Magazine, 19*(3), 42–51.

Orel, S. G. (2000). MR imaging of the breast. *Radiologic Clinics of North America, 38*, 899–913.

Pacific Chiropractic and Research Center. *Breast cancer and early detection*. Available at http://www.breastthermography.com.

Pendharkar, P. C., Rodger, J. A., Yaverbaum, G. J., Herman, N., & Benner, M. (1999). Association, statistical, mathematical and neural approaches for mining breast cancer patterns. *Expert Systems with Applications, 17*, 223–232.

Pisano, E. D., Fajardo, L. L., Caudry, D. J., Sneige, N., et al. (2001). Fine-needle aspiration biopsy of nonpalpable breast lesions in a multicenter clinical trial: Results from the radiologic diagnostic oncology group V. *Radiology, 219*, 785–792.

Puglisi, F., Zuiani, C., Bazzocchi, M., Valent, F., et al. (2003). Role of mammography, ultrasound and large core biopsy in the diagnostic evaluation of papillary breast lesions. *Oncology, 65*, 311–315.

Qi, H., & Diakides, N. A. (2003). Thermal infrared imaging in early breast cancer detection – A survey of recent research. In *Proceedings of the 25th annual international conference of the IEEE EMBS* (pp. 1109–1112).

Rajentheran, R., Rao, C. M., Lim, E., & Lennard, T. W. J. (2001). Palpable breast cancer which is mammograpically invisible. *The Breast, 10*(5), 416–420.

Reinikainen, H. (2003). Complementary imaging of solid breast lesions: Contribution of ultrasonography, fine-needle aspiration biopsy, and high-field and low-field MR imaging. Academic dissertation. Faculty of Medicine, University of Oulu, Finland.

Sajda, P., Spence, C., & Parra, L. (2003). A multi-scale probabilistic network model for detection, synthesis, and compression in mammographic image analysis. *Medical Image Analysis, 7*(2), 187–204.

Santos-André, T. C. S., & da Silva, A. C. R. (1999). A neural network made of a Kohonen's SOM coupled to a MLP trained via backpropagation for the diagnosis of malignant breast cancer from digital mammograms. *International Joint Conference on Neural Networks, 5*, 3647–3650.

Schnorrenberg, F., Pattichis, C. S., Kyriacou, K. C., & Schizas, C. N. (1997). Computer-aided detection of breast cancer nuclei. *IEEE Transactions on Infomration Technology in Biomedicine, 1*(2), 128–140.

Schnorrenberg, F., Tsapatsoulis, N., Pattichis, C. S., Schnizus, C. N., et al. (2000). Improved detection of breast cancer nuclei using modular neural networks. *IEEE Engineering in Medicine and Biology Magazine, 19*(1), 48–63.

Seker, H., Odetayo, M., Petrovic, D., Naguib, R. N. G., Bartoli, C., Alasio, L., et al. (2001). Prognostic comparison of statistical neural and fuzzy methods of analysis of breast cancer image cytometric data. In *Proceedings of 23rd annual EMBS international conference* (pp. 3811–3814).

Simmon, J. R., Kalbhen, C. L., Cooper, R. A., & Flisak, M. E. (2000). Accuracy and complication rates of US-guided vacuum-assisted core breast biopsy: Initial results. *Radiology, 215*, 694–697.

Singhal, H., & Thomson, S. (2004). Breast cancer evaluation. eMedicine.com, Inc. Available from http://www.emedicine.com/med/topic3287.htm.

Spengler, R. (2003). *Breast cancer (BRCA) gene testing, healthwise*. Available from http://my.webmd.com/hw/breast_cancer/tu6462.asp.

Stavros, A. T., Thickman, D., Rapp, C. L., Dennis, M. A., Parker, S. H., & Sisney, G. A. (1995). Solid breast nodules: Use of sonography to distinguish between benign and malignant lesions. *Radiology, 196*, 123–134.

Sun, X., Qian, W., & Song, D. (2004). Ipsilateral-mammogram computer-aided detection of breast cancer. *Computerized Medical Imaging and Graphics, 28*, 151–158.

Tan, T. Z., Quek, C., & Ng, G. S. (2004). FALCON-AART: An improved version of FALCON-ART. Technical report. C2i-TR-04/002. School of Computer Engineering, Nanyang Technological University, Singapore.

Tan, T. Z. (2005). *Complementary learning fuzzy neural network for medical domain and bioinformatics domain applications*. First year report, Department of Computer Science, School of Coumpter Engineering, Nanyang Technological University.

Tan, T. Z., Quek, C., Ng, G. S., & Ng, E. Y. K. (2004). Breast cancer diagnosis using thermography and complementary learning fuzzy neural network. In *Proceedings of the first international bioengineering conference (IBEC)* (pp. 65–68).

Tateishi, T., Machi, J., Feleppa, E. J., Oishi, R., et al. (1998). In vitro diagnosis of axillary lymph node metastases in breast cancer by spectrum analysis of radio frequency echo signals. *Ultrasound in Medicine & Biology, 24*(8), 1151–1159.

Thompson, C. J., Murthy, K., Picard, Y., Weinberg, I. N., & Mako, R. (1995). Positron emission mammography (PEM): A promising technique for detecting breast cancer. *IEEE Transaction on Nuclear Science, 42*(1), 1012–1017.

Tsapatsoulis, N., Schnorrenberg, F., Pattichis, C., & Kollias, S. (1997). An image analysis system for automated detection of breast cancer nuclei. *IEEE International Conference on Image Processing, 3*, 512–515.

Tung, W. L., & Quek, C. (2001). A novel approach to the derivation of fuzzy membership functions using Falcon-MART architecture. *Pattern Recognition Letters, 22*, 941–958.

Usuki, H., Maeta, H., Maeba, T., Wakabayashi, H., et al. (2000). Standardization of thermographic breast cancer detection-role of qualitative findings and quantitative findings. In *Proceedings of the 22nd annual international conference of the IEEE Engineering in Medicine and Biology Society* (Vol. 2(2), pp. 219–1222).

van't Veer, L. J., Dai, H., van de Vijver, M. J., He, Y. D., et al. (2002). Gene expression profiling predicts clinical outcome of breast cancer. *Nature, 415*, 530–536.

Varga, M. J., & De Muynck, P. (1992). Thermal analysis of infra-red mammography. In *Proceedings of 11th IAPR international conference on pattern recognition. Conference C: Image, speech and signal analysis* (Vol III, pp. 360–364).

Vlahou, A., Laronga, C., Wilson, L., Gregory, B., et al. (2003). A novel approach toward development of a rapid blood test for breast cancer. *Clinical Breast Cancer, 4*(3), 203–209.

Wang, X., Zheng, B., Good, W. F., King, J. L., & Chang, Y. (1999). Computer-assisted diagnosis of breast cancer using a data-driven Bayesian belief network. *International Journal of Medical Informatics, 54*, 115–126.

Yao, X., & Liu, Y. (1999). Neural networks for breast cancer diagnosis. In *Proceedings of the 1999 congress on evolutionary computation* (Vol. 3, pp. 1760–1767).

Yeung, D. K. W., Cheung, H. S., & Tse, G. M. K. (2001). Human breast lesions: Characterization with contrast-enhanced in vivo proton MR spectroscopy-initial results. *Radiology, 220*, 40–46.