

Automatic Determination of Hormone Receptor Status in Breast Cancer Using Thermography

Siva Teja Kakileti, Krithika Venkataramani^(✉), and Himanshu J. Madhu

Xerox Research Centre India, Bangalore, India

{SivaTeja.Kakileti,Krithika.Venkataramani,Himanshu.Madhu2}@xerox.com

Abstract. Estrogen and progesterone hormone receptor status play a role in the treatment planning and prognosis of breast cancer. These are typically found after Immuno-Histo-Chemistry (IHC) analysis of the tumor tissues after surgery. Since breast cancer and hormone receptor status affect thermographic images, we attempt to estimate the hormone receptor status before surgery through non-invasive thermographic imaging. We automatically extract novel features from the thermographic images that would differentiate hormone receptor positive tumors from hormone receptor negative tumors, and classify them through machine learning. We obtained a good accuracy of 82 % and 79 % in classification of HR+ and HR- tumors, respectively, on a dataset consisting of 56 subjects with breast cancer. This shows a novel application of automatic thermographic classification in breast cancer prognosis.

Keywords: Thermography · Breast cancer prognosis · Hormone receptor status

1 Introduction

Breast cancer has the highest incidence among cancers in women [1]. Breast cancer also has wide variations in the clinical and pathological features [2], which are taken into account for treatment planning [3], and to predict survival rates or treatment outcomes [2, 4]. Thermography offers a radiation free and non-contact approach to breast imaging and is being re-investigated in recent times [5–8] with the availability of high resolution thermal cameras. Thermography detects the temperature increase in malignancy due to the increased metabolism of cancer [9] and due to the additional blood flow generated for feeding the malignant tumors [6]. Thermography may also be sensitive to hormone receptor status as these hormones release Nitric Oxide, which causes vasodilation and temperature increase [6, 10]. Both these effects could potentially lead to evaluation of hormone receptor status of malignant tumors using thermography. If this is possible, it provides a non-invasive way of predicting the hormone receptor status of malignancies through imaging, before going through Immuno-Histo-Chemistry (IHC) analysis on the tumor samples after surgery. This paper investigates this possibility and the prediction accuracy. Most other breast imaging techniques including

mammography are not able to detect hormone receptor status changes. Though the paper by Chaudhuri et al. [11] claims that Dynamic Contrast Enhanced (DCE) MRI can be used for prediction of Estrogen status, it is invasive, and has been tested only on a small dataset of 20 subjects with leave-one-out cross-validation.

There has been a study to analyze the effect of hormone receptor status of malignant tumors on thermography [12] though quantitative analysis of average or maximum temperatures of the tumor, the mirror tumor site and the breasts. [12] reports a significant difference in these temperature measurements for hormone receptor positive and negative status using thermography. In this paper, we automatically extract features from the thermographic images in the region of interest (ROI), i.e. the breast tissue, using image processing and attempt to classify the hormone receptor status of malignant tumors using machine learning techniques. The determination of whether or not a subject has breast cancer using thermography, i.e. screening for cancer, is out of scope for this paper. There are other algorithms for breast cancer screening using thermography [8, 13], which the reader may refer to based on interest.

The paper is organized as follows. Section 2 provides details on the effect of hormone receptor positive and negative breast cancers on thermography from the existing literature. Section 3 describes our approach to automatic feature extraction from the ROI for HR+ and HR- malignant tumor classification. Section 4 describes the dataset used for our experiments and our classification results are provided in Sect. 5. Conclusions and future work are given in Sect. 6.

2 Effect of Hormone Receptor Status on Thermography

There is usage of readily available tumor markers such as Estrogen Receptor (ER), Progesterone Receptor (PR), Human Epidermal growth factor Receptor 2 (HER2) and tumor cell growth protein marker Ki67, for treatment planning [3, 14], and survival rate prediction [2, 4], especially in resource constrained developing countries like India. [2] uses ER, PR and HER2 for estimating breast cancer mortality risk from a large dataset of more than 100,000 patients with invasive breast cancer. They find that there is variability in the 8 different ER/PR/HER2 subtypes, and the ER status has the largest importance. ER+ tumors have a lower risk than ER- tumors. PR status has a lesser importance than ER status and PR+ tumors have lower risk than PR- tumors. HER2 status has variations in risk across the different hormone receptor subtypes, depending on the stage of the cancer, with the lowest risk for the ER+/PR+/HER2- tumors, and the highest risk for ER-/PR-/HER2- tumors. The effect of the Ki-67 marker indicates the rate of tumor cell growth [14]. More aggressive tumors may have higher temperatures due to their increased metabolism [9] and so the Ki-67 marker status may play a role in thermography, but it has not been formally investigated in any study yet.

Estrogen leads to increase in vasodilation due to the production of Nitric Oxide with a resultant temperature increase [6, 15]. Progesterone is also

associated with locally high concentrations of Nitric Oxide generation [10] for prolonged periods of time. [12] find there is a significant difference in average and maximum temperature of the tumor site between PR+ and PR- tumors, with the PR- tumors being hotter. The same pattern holds for ER status although in a non-significant manner. Their study showed that the more aggressive ER-/PR- tumors were hotter than the less aggressive ER+/PR+ tumors. Their study also indicates that the difference in average temperatures of the tumor and its mirror sites in contra-lateral breasts is higher in ER- tumors than in ER+ tumors, although in a non-significant manner. The same pattern holds for the PR status too. Since the hormone sensitivity of both breast tissues are similar, it is probable that there is a thermal increase on both breasts for estrogen or progesterone positive cases. [12] don't specifically analyze the four different subtypes of ER/PR status, probably because the difference in temperatures are small for just one hormone receptor status. Using these medical reasons and empirical observations, in the next section, we design a set of novel features along with a few existing features that would either extract these observations automatically or would correlate with these findings for classifying hormone receptor positive and negative tumors.

3 Automatic Feature Extraction for Hormone Receptor Status

We attempt to classify all combinations of Hormone Receptor (HR) positive (ER+/PR+, ER+/PR-, ER-/PR+) tumors from the HR negative (ER-/PR-) tumors. We extracted features from elevated temperature regions in the ROI, and the overall ROI. The elevated temperature regions, i.e., the hot-spots are extracted as below.

3.1 Abnormal Region Extraction

The entire ROI is divided into abnormal regions and normal regions based on their regional temperatures. The malignant tumor region is typically an abnormal region with an elevated temperature. The abnormal regions have the highest regional temperature in the Region of Interest (ROI). To segment an abnormal region, we used an algorithm proposed in [16], where segmentation areas are combined from multiple features defined by Eqs. 1 and 2 using a decision rule.

$$T_1 = Mode(ROI) + \rho * (T_{max} - Mode(ROI)) \quad (1)$$

$$T_2 = T_{max} - \tau \quad (2)$$

In the above equations, T_{max} represents the overall maximum temperature in all views and $Mode(ROI)$ represents the mode of the temperature histogram obtained using temperature values of pixels from the ROIs of all views. The parameters ρ , τ and the decision fusion rule are selected based on the accuracy

of classification on a training/cross-validation subset and diversity in the segmentation decisions. Decision fusion results in better hot-spot detection than simple thresholding techniques [16]. Heat transmission from deep tumors results in diffused lower temperatures on the surface and these parameters play a large role in the deep tumor detection. Research on determining the combined depth and size of tumors that can be detected needs to be done.

As discussed in [12], HR– tumors are hotter compared to HR+ tumors while temperature increase on both sides is observed for HR+ tumors due to the presence of similar hormone sensitive tissues. To capture these properties, we extract the following features from these detected abnormal regions.

Distance Between Regions. The malignant tumor region is hotter than the surrounding region, but the relative difference is higher for HR– tumors. In case of HR+ tumors, the entire breast region is warmed up, and so this difference is lesser. We use the normalized histogram of temperatures, or probability mass function (PMF), to represent each region, and find the distance between regions using a distance measure between PMFs. Here, the Jensen-Shannon Divergence (JSD) is used as a measure, as it is a symmetric measure. The JSD is defined as

$$JSD(P||Q) = \frac{1}{2} * \sum_i (\log \frac{P(i)}{M(i)})P(i) + \frac{1}{2} * \sum_i (\log \frac{Q(i)}{M(i)})Q(i), \quad (3)$$

where $M = \frac{1}{2}(P + Q)$. The value of $JSD(P||Q)$ tends to zero when P and Q have identical distributions and has a very high value when the distributions are very different. To include a measure of distance between multiple regions, one or more of the PMFs of one region is modified by the mean temperature of another region. The JSD between $P - \mu_2$ and $Q - \mu_1$, where P is the PMF of the abnormal region on the malignant side, Q is the PMF of the normal region on the malignant side, μ_1 is the mean of the contra-lateral side abnormal region and μ_2 is the mean of the contra-lateral side normal region, is taken as a feature. In case of absence of an abnormal region on the contralateral side, μ_1 is taken to be equal to μ_2 . A subtraction of the contralateral region means corresponds to a relative increase in the heat with respect to the contralateral regions. For HR– tumors, there may be no abnormal regions on the contra-lateral side, due to which this JSD will be higher.

Relative Hotness to the Mirror Site. HR+ tumors have a lower temperature difference between the tumor site and the mirror tumor site on the contra-lateral side. To capture this, we use the mean squared distance between the temperature of the malignant side abnormal region pixels and the mean temperature of the contra-lateral side abnormal region, as defined in Eq. 4.

$$RH = \frac{1}{|A|} \sum_{x \in A} \sum_{y \in A} ||T(x, y) - \mu||^2 \quad (4)$$

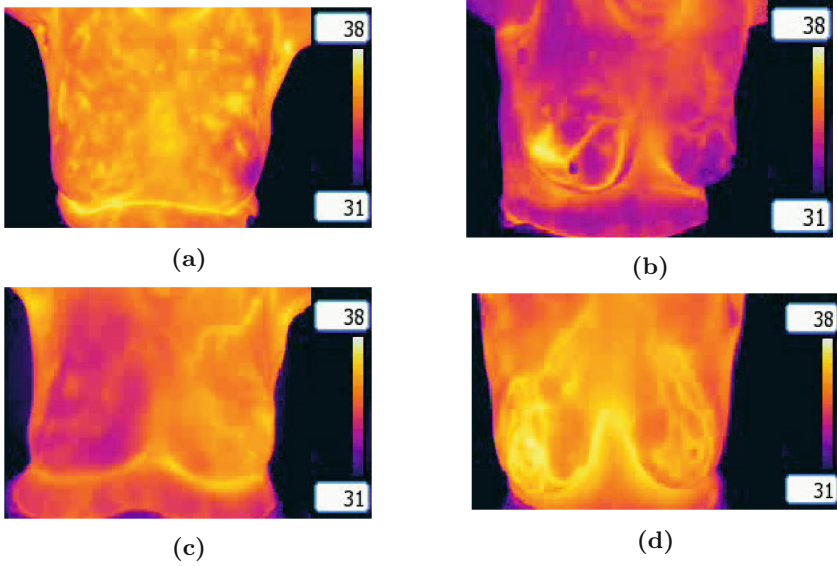


Fig. 1. Shows subjects with malignant tumors having a. ER+/PR+ status b. ER-/PR- status c. ER+/PR+ status with asymmetrical thermal response d. ER-/PR- status with some symmetrical thermal response

where $T(x, y)$ represents temperature of the malignant side abnormal region pixels at location (x, y) in the image, μ represents mean temperature of the contralateral side abnormal region and $|A|$ represents the cardinality of abnormal region A on the malignant side. This value is lower for HR+ tumors compared to HR- tumors, as hormone sensitive tissues will be present on both sides. As shown in Fig. 1a and b, we see thermal responses on both sides for HR+ tumors and no thermal response on the normal breast for HR- tumors. However, there might be outliers like Fig. 1c and d.

Thermal Distribution Ratio. In addition to the temperature change, the areas of the abnormal regions on both sides are also considered as features. We used the ratio of areas of abnormal regions on the contralateral side to the malignant side. This value tends to be zero for HR- tumors, as there may be no abnormal region on the contralateral side, and is higher for HR+ tumors.

3.2 Entire ROI Features

Textural features are used here to extract the features from the entire ROI. However, instead of using the original temperature map of the ROI, a modified temperature map is used. The thermal map formed by subtracting the malignant side ROI with the contra-lateral side mean temperature, i.e. the relative temperature from the contralateral side, is used to determine the textural features. The

Run Length Matrix (RLM) is computed from the thermal map, after quantizing the temperature into l bins. Gray level non-uniformity and Energy features from the RLM are computed, as mentioned in [7]. The non-uniformity feature would be higher for HR– tumors as their tumors have more focal temperatures.

4 Dataset Description

We obtained an anonymized dataset of 56 subjects with biopsy confirmed breast cancer with age varying from 27 to 76 years through our collaboration with Manipal University. The FLIR E60 camera with a spatial resolution of 320×240 pixels is used to capture the initial 20 subjects and a high-resolution FLIR T650Sc camera with an image resolution of 640×480 pixels is used for the remaining subjects. A video is captured for each subject, and the acquisition protocol involved asking the subject to rotate from right lateral to left lateral views. The data for each subject included the mammography, sono-mammography, biopsy reports, the ER/PR status values, with surgery reports and HER2 Neu status values, where available of the tumors. From this data, there are 32 subjects with HR+ malignant tumors and rest of them have HR– tumors.

5 Classification Results

From the obtained videos, we manually selected five frames that correspond to frontal, right & left oblique and lateral views, and manually cropped the ROIs in these. Consideration of multiple views helps in better tumor detection since it might not be seen in a fixed view. From these multiple views, the view corresponding to maximum abnormal region area with respect to the ROI area is considered as the best view. This best view along with its contra-lateral side view is used to calculate the features from the abnormal regions and the entire ROI as mentioned in Sect. 3. The training set and testing set comprise of a randomly chosen subset of 26 and 30 subjects, respectively, with an internal division of 14 HR+ & 12 HR– and 18 HR+ & 12 HR– tumors, respectively. The abnormal region is located using $\rho = 0.2$, $\tau = 3^\circ C$ using the *AND* decision rule, to optimize for the accuracy in classification. All 11 deep tumors of size 0.9 cm and above have been detected in this dataset. The bin width of the PMFs used is $0.5^\circ C$. The step size of the temperature bins in the RLM computation is $0.25^\circ C$.

A two-class Random Forest ensemble classifier is trained using the features obtained. The Random Forest (RF) randomly chooses a training sub-set & a feature sub-set for training a decision tree, and combines the decisions from multiple such trees to get more accuracy in classification. The mode of all trees is taken as the final classification decision. RFs with increasing number of trees have a lower standard deviation in the accuracies over multiple iterations. The standard deviation in (HR–, HR+) accuracies of the RFs using all features with 5, 25 and 100 trees over 20 iterations is (9.1 %, 11.1 %), (6.4 %, 4.8 %), (2.5 %, 2.0 %), respectively, and hence a large number of 100 trees is chosen. Table 1 shows the

max. accuracies over 20 iterations of RFs with 100 trees using individual and combined features proposed in our approach. We tested with different textural features obtained from both RLM and Gray Level Co-occurrence Matrix, but we found out that gray-level non-uniformity from the RLM is having better accuracy than others. Using an optimal combined set of region based features and textural features, we obtained an accuracy of 82 % and 79 % in classification of HR+ and HR– tumors respectively.

Table 1. Accuracies with different features obtained using our approach

Feature set	Features	HR–Accuracy	HR+Accuracy
Abnormal Region Features	Distance between regions	74 %	56 %
	Relative Hotness	79 %	73 %
	Thermal Distribution Ratio	63 %	27 %
	Combination of above three features	84 %	73 %
Entire ROI Features	Gray-level non-uniformity	68 %	64 %
Overall features	Combination of features from abnormal and entire ROI regions	79 %	82 %

From Table 1, it is clear that Abnormal Region features plays an important role compared to textural features. Among these abnormal region features, features corresponding to relative temperatures, i.e., Relative Hotness and Distance Between Regions, have an important role in the classification of HR+ and HR– tumors, thus validating the findings of [12].

6 Conclusions and Future Work

We have come up with a novel application to automatically classify breast cancer tumors into HR+ tumors from HR– tumors using thermography with a reasonably good accuracy of around 80 %. This is a first approach through image processing features and machine learning algorithms for such automatic classification. This also presents an advantage to thermography over other imaging modalities in estimating prognosis and treatment planning of breast cancer without invasive surgery. In future work, we will test our algorithm on larger datasets with more variation in data and modify the algorithm to detect sub classes within HR+ tumors. Additionally, we will try to determine the role of Ki-67 status in thermography to refine the automatic classification.

Acknowledgement. We thank Manipal University and Dr. L. Ramachandra, Dr. S. S. Prasad and Dr. Vijayakumar for sharing the data and assisting us in thermographic image interpretation.

References

1. Fitzmaurice, C., et al.: The global burden of cancer 2013. *JAMA Oncol.* **1**(4), 505–527 (2015)
2. Parise, C.A., Caggiano, V.: Breast cancer survival defined by the er/pr/her2 subtypes and a surrogate classification according to tumor grade and immunohistochemical biomarkers. *J. Cancer Epidemiol.* **2014**, 11 p. (2014). Article ID 469251
3. Alba, E., et al.: Chemotherapy (CT) and hormonotherapy (HT) as neoadjuvant treatment in luminal breast cancer patients: results from the GEICAM/2006-03, a multicenter, randomized, phase-ii study. *Ann. Oncol.* **23**(12), 3069–3074 (2012)
4. Cheang, M., Chia, S.K., Voduc, D., et al.: Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J. Nat. Cancer Inst.* **101**(10), 736–750 (2009)
5. Keyserlingk, J., Ahlgren, P., Yu, E., Belliveau, N., Yassa, M.: Functional infrared imaging of the breast. *Eng. Med. Biol. Mag.* **19**(3), 30–41 (2000)
6. Kennedy, D.A., Lee, T., Seely, D.: A comparative review of thermography as a breast cancer screening technique. *Integr. Cancer Ther.* **8**(1), 9–16 (2009)
7. Acharya, U.R., Ng, E., Tan, J.H., Sree, S.V.: Thermography based breast cancer detection using texture features and support vector machine. *J. Med. Syst.* **36**(3), 1503–1510 (2012)
8. Borchardt, T.B., Conci, A., Lima, R.C., Resmini, R., Sanchez, A.: Breast thermography from an image processing viewpoint: a survey. *Signal Process.* **93**(10), 2785–2803 (2013)
9. Gautherie, M.: Thermobiological assessment of benign and malignant breast diseases. *Am. J. Obstet. Gynecol.* **147**(8), 861–869 (1983)
10. Vakkala, M., Kahlos, K., Lakari, E., Paakko, P., Kinnula, V., Soini, Y.: Inducible nitric oxide synthase expression, apoptosis, and angiogenesis in in-situ and invasive breast carcinomas. *Clin. Cancer Res.* **6**(6), 2408–2416 (2000)
11. Chaudhury, B., et al.: New method for predicting estrogen receptor status utilizing breast mri texture kinetic analysis. In: *Proceedings of the SPIE Medical Imaging* (2014)
12. Zore, Z., Boras, I., Stanec, M., Oresic, T., Zore, I.F.: Influence of hormonal status on thermography findings in breast cancer. *Acta Clin. Croat.* **52**, 35–42 (2013)
13. Madhu, H., Kakileti, S.T., Venkataramani, K., Jabbireddy, S.: Extraction of medically interpretable features for classification of malignancy in breast thermography. In: *38th Annual IEEE International Conference on Engineering in Medicine and Biology Society (EMBC)* (2016)
14. Urruticoechea, A.: Proliferation marker ki-67 in early breast cancer. *J. Clin. Oncol.* **23**(28), 7212–7220 (2005)
15. Ganong, W.F.: *Review of Medical Physiology*. McGraw-Hill Medical, New York (2005)
16. Venkataramani, K., Mestha, L.K., Ramachandra, L., Prasad, S., Kumar, V., Raja, P.J.: Semi-automated breast cancer tumor detection with thermographic video imaging. In: *37th Annual International Conference on Engineering in Medicine and Biology Society, pp. 2022–2025* (2015)