Detection of Breast Cancer Using Infrared Thermal Images for Improved Accuracy by Using Random Forest and Multilayer Perceptron

Thejeshwar M[1] , Stella Jenifer Isbella S[2,*]

[1]Research Scholar, [2]Project guide, Department of Medical Instrumentation, Saveetha School of Engineering, saveetha Institute of Medical and Technical Sciences, Saveetha University,Chennai, Tamilnadu, India,

Abstract: At the present time, breast cancer is one of the most often diagnosed forms of cancer in females. Mammography is the most common form of screening imaging used to identify breast cancer in its earlier stages. Nevertheless, thermal infrared pictures (thermography) can be utilized to detect lesions in dense breasts. In this study, the typical areas reflect warmer temperatures than malignant areas. In this study, we offer a unique approach for modeling the temperature variations in normal and abnormal breasts by combining the Random forest and Multilayer perceptron techniques. The project aims to study the accuracy, sensitivity, and specificity of the infrared breast cancer images using infrared thermal images using random forest and multilayer perceptron algorithms and comparing the accuracy, specificity, and sensitivity. **Materials and Methods**: The information for this study was s gained from thermal images from Visual labs DMR-IR. The samples were considered as (N=60) for Random Forest and (N= 60) for MultiLayer Perceptron. Novel Matlab software is used to calculate accuracy, specificity, and sensitivity. **Results:** The result demonstrates the accuracy of the thermal breast images using SPSS software. A statistically insignificant difference exists, with Random Forest accuracy (92.5%) with specificity (90%) and with sensitivity (95%) and demonstrated a better outcome in comparison with Multilayer Perceptron accuracy (90%), specificity (91.6%) and sensitivity (88.3%). **Conclusion:** Random Forest gives better accuracy, specificity, and sensitivity than Multilayer Perceptron to detect breast cancer.

KEYWORDS: Breast cancer disease; novel thermal IR images; Random Forest ,Multilayer Perceptron, MatLAB;

INTRODUCTION

Breast cancer affects the bodies of thousands of people all over the world . (Malvezzi et al. 2018)Mammographic scans are frequently able to identify breast cancer in its early stages (Abdel-Nasser, Moreno, and Puig 2016). [Summary] [Citation needed] On the other hand, a number of studies have demonstrated that thermal infrared pictures, which

* Corresponding author : bitmist2011@gmail.com

are often referred to as thermographies, can produce better results in the case of thick breasts (breasts of young girls) (Chiarelli et al. 2015).Mammographic images are commonly used to detect breast cancer in its early stages. The cost of the dynamic thermography procedure is significantly lower than that of the mammography and magnetic resonance imaging (MRI) procedures. In addition, it is a diagnostic process that does not cause discomfort to patients, is non-invasive, does not emit any ionizing radiation, and is safe. Thermography operates based on two facts: first, the temperatures of breast cancer regions are higher than the temperatures of the tissues that surround them, and second, the rates at which metabolic heat and blood perfusion are generated in tumors are significantly higher than the rates at which they are generated in normal regions. Thermal infrared cameras have the capability of capturing the difference in temperature that exists between normal and diseased breast tissue (Faust et al. 2014).

By using an external stimulus to boost the thermal contrast, dynamic infrared thermography may improve the detection results of static infrared imaging (Zhou and Herman 2018). This is accomplished by applying an infrared thermography technique known as dynamic infrared thermography.

Procedures involving either cooling or heating the patient's breasts can be used as a stimulus to stimulate the patient's breast tissue thermally. It is important to point out that the method of cooling is significantly more secure than the procedure of heating since the temperature range of a woman's body is between 36.5 and 37.5 degrees Celsius, and thus higher heating may cause damage to the living tissues in the breasts. When the breasts are subjected to the chilling method, the temperature of healthy tissues drops together with an attenuation of the vascular diameter. On the other hand, the temperature of diseased tissues does not change (or it increases along with a vascular dilatation), as was described in (Kennedy, Lee, and Seely 2009). In this method, the study's findings of the similarity (or the dissimilarity) between infrared pictures taken before and after the cooling technique may be used to detect breast cancer. This can be done by comparing the infrared images to each other.

Multiple methods, including mammography and magnetic resonance imaging (MRI), have been introduced over the past two decades for the early diagnosis of breast cancer (Litjens et al. 2017). In their review of the many methods available for detecting breast cancer, Sebastien et al. (Mambou et al. 2018) highlighted the key drawbacks of each. In a recent review article, Hamidinekoo et al. (Hamidinekoo et al. 2018) summarized the current state-of-the-art deep learning-based computer-aided diagnosis (CAD) systems created for mammography and breast histopathology pictures. A method was devised to link mammographic abnormalities with their histological counterparts. Breast density categorization in mammograms proposed by Rampun et al. (Rampun et al. 2018), who suggested using local quinary patterns (LQP) on different neighborhood topologies. Mammography is the gold standard for detecting breast cancer early, yet, it has been shown to have a high rate of false positives (Mambou et al. 2018). Consequently, women with genetic mutations are encouraged to get MRIs in addition to regular mammograms (Cho et al. 2017). However, MRI's biggest drawback is its poor spatial resolution, which results in a lack of sensitivity for lesions smaller than a centimeter in size (Boogerd et al. 2017).

Since infrared thermal imaging may specifically maximize the contrast in regions of dense tissues (young women), it can help overcome the limitations of the mammography approach (Kosus et al. 2010). Several approaches have been presented in the literature (Díaz-Cortés et al. 2018),,] for analyzing breast cancer using dynamic thermograms. Using thermography as an example, the authors of (Sathish et al. 2016) surveyed the state of the art in computer-aided cancer detection techniques across a variety of medical imaging modalities. In (Selle et al. 2015), the authors introduced a novel approach to ROI extraction from breast thermograms, one that considers both the breasts' lateral and frontal perspectives. The discovered ROIs assist doctors in distinguishing between normal and

aberrant biomarkers. Discrete wavelet transform, texture descriptors, fuzzy and decision tree classifiers were all tested and assessed by Mookiah et al. (Mookiah, Rajendra Acharya, and Ng 2012). From a sample size of 50 thermograms, they could determine an average sensitivity of 86.70 percent, a specificity of 100 percent, and an accuracy of 93.30 percent. (Saniei et al. 2015)devised a five-step method of assessing thermal pictures to detect cancer. The breast region is extracted from the infrared images with the connected component labeling method, the infrared images are aligned with a registration method, the blood vessels are segmented with morphological operators, the branching points of each vascular network are exploited as thermal minutiae points. The branching points are fed into a matching algorithm to classify breast regions as normal or abnormal. By providing each infrared image's extracted co-occurrence matrix and run length matrix texture features into a support vector machine algorithm, (Acharya et al. 2012). were able to distinguish between normal and malignant cases with a mean sensitivity of 85.71%, specificity of 90.48%, and accuracy of 88.10%. In addition, showed that the Lazy snapping method (an interactive picture cutoff procedure) is useful for rapidly and easily adjusting the hottest/coldest parts of thermographic images.(Etehadtavakol et al. 2019)

Most of the methods proposed in the literature focus on the extraction of features from the images and do not consider the temporal evolution of temperature during the dynamic procedure (i.e., they ignore the temporal information in the infrared image sequences). In this paper, we present an effective and efficient method to improve the accuracy of the thermographic breast images and identify breast cancer for classification as normal or abnormal (without cancer or with cancer. we proposed the use of a random tree and multilayer perceptron to compare the accuracy of the thermal breast images, we also present the comparison of results with random tree forest (TRF), and multilayer perceptron (MLP).

MATERIALS AND METHODS

The research study was carried out using Matlab programming, at Saveetha School of Engineering. The data sets were collected from the Visual labs DMR-IR website. It is worth mentioning that the DMR-IR dataset includes segmented images containing only the temperatures of breasts (excluding the temperatures of other body parts). After training we used Gray level co-occurrence matrix (GLCM) for feature extraction with Matlab software. We trained the data using Random Forest and Multilayer perceptron.. Once the data was trained, we generated the confusion matrix. By using the confusion matrix, accuracy (%), sensitivity (%) and specificity (%) were determined.

statistical analysis: To Evaluate the proposed method we use one of their metrics : the accuracy . To determine the accuracy TP(true positive), TN (true negative),FP(False Positive),FN (False Negative) are to be determined . Table 1 represents the statistical analysis of RF and MLP. Here in the table 1 we can determine that RF show improved accuracy of 2% when compared to the MLP techniques. From table 2 : Anova test was done to determine the accuracy of the malignant and adjacent healthy samples using breast thermal images.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Group Statistics

groups	N	Mean	Std. Deviation	Std. Error Mean

accuracy	RF	60	92.7052	.13574	.01752
	MLP	60	90.3993	.18061	.02332

Table 1. Statistics analysis for Random Tree and Multiplayer Perceptron

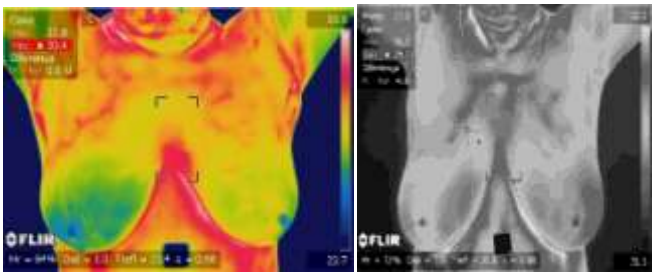
Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
accuracy	Equal variances assumed	6.366	.013	79.054	118	<.001	<.001	2.30583	.02917	2.24807	2.36359
	Equal variances not assumed			79.054	109.530	<.001	<.001	2.30583	.02917	2.24803	2.36364

Table 2. Independent sample test for RF and MLP

TP refers to the true positive cases
TN refers to true negative cases
FP refers to False Positive cases
FN refers to False negative cases.

RESULTS



Breast IR images **Grey Scale Images**

Figure .1 Breast Infrared Thermal images

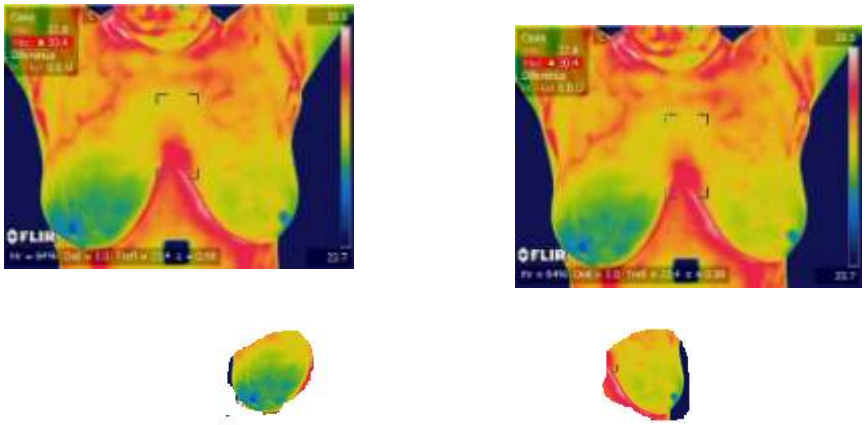


Figure 2. GLCM segmentation method

Figure 1. show the thermal images of the breast. The temperature contrast of cancer and the other nearby regions are also displayed in the Fig.1. Figure 2 show the images are segmented and extracted using GLCM techniques. From the extracted feature, cancer and the normal region is classified and their accuracy is determined. In breast cancer detection study, both techniques appeared to provide different results with accuracy ranging from (90%-92%), specificity (91-90%) and sensitivity (88%-95%). Table 1 shows the mean accuracy (60.42%), of the breast using Random Forest, and Table 2 representsFigure 1. show the thermal images of the breast. Te other nearby regions are also displayed in the Fig.1. breast cancer detection study, both techniques appeared to provide different results with accuracy ranging from (90%-92%), specificity (91-90%) and sensitivity (88%-95%). Table 1 shows the mean accuracy (60.42%), of the breast using Random Forest, and Table 2 represents the mean accuracy (87-90%) of breast cancer using Multiple Layer Perceptron. Table 2. explains the comparison of mean, accuracy, and Figure 2 show the images are segmented and extracted using GLCM techniques. From the extracted feature, cancer and the normal region is classified and their accuracy is determined. In using Random Forest and Multilayer Perceptron.

To determine the accuracy, Figure 3 and 4a & 4b explains the comparison of accuracy between both the classifiers. Figure 3 shows the accuracy percentage. From the graph, it is understood that the Random forest shows 2 times higher accuracy than the multilayer perceptron. Figure 4a shows the Confusion Matrix of Random Forest, no of Normal and Abnormal outputs, overall accuracy, and it is declaredhe temperature contrast of cancer and th to be (92.5%). Figure 4b shows the Confusion Matrix of Multilayer Perceptron and overall accuracy and its declared to be (90%).
the mean accuracy (87-90%) of breast cancer using Multiple Layer Perceptron. Table 2. explains the comparison of mean, accuracy, and using Random Forest and Multilayer Perceptron.

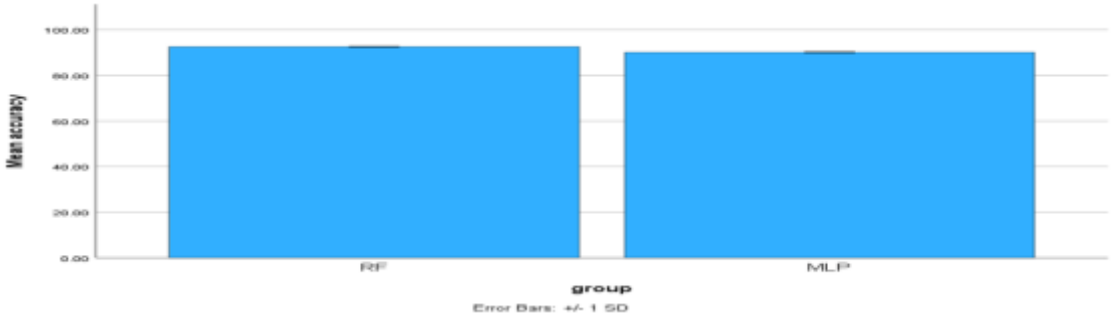


Fig. 3. Simple bar mean of accuracy, using Random Tree and Multilayer perceptron. Both techniques appear to produce the same variable results with the accuracy ranging from (92.5%-90%), . X asis: Random Tree algorithm vs Multilayer Perceptron, algorithm vs Y axis: Mean accuracy of detection +/- 1 SD.

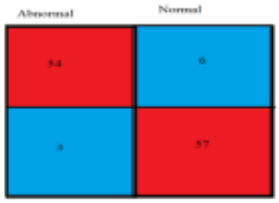


Fig. 4a. Confusion matrix of Cosine KNN classifier. True positive accounts for 54, false positive accounts for 6, false negative accounts for 3 and true negative accounts for . The total accuracy 54 of 60 was found to be (92.5%)

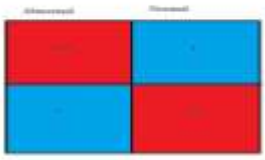


Figure .4b Confusion matrix of Multilayer Perceptron True positive accounts for 53, false positive accounts for 5, false negative accounts for 7 and true negative accounts for 53. The total accuracy was found to be (90%%).

To determine the accuracy, Figure 3 and 4a & 4b explains the comparison of accuracy between both the classifiers. Figure 3 shows the accuracy percentage. From the graph, it is understood that the Random forest shows 2 times higher accuracy than the multilayer perceptron. Figure 4a shows the Confusion Matrix of Random Forest, no of Normal and

Abnormal outputs, overall accuracy, and it is declared to be (92.5%). Figure 4b shows the Confusion Matrix of Multilayer Perceptron and overall accuracy and its declared to be (90%).

DISCUSSION

In this paper, we discuss a novel technique for using dynamic thermograms to diagnose breast cancer. Using RF and Multilayer Perceptron methods, we simulate the rise and fall in breast temperature experienced during dynamic thermography operations. With our technique, the whole series of thermograms for each example is shown in a way that is both concise and illustrative. As far as classification accuracy goes, the suggested technique achieves remarkable outcomes. Compared to similar approaches, its performance is superior as well. Future work will center on leveraging sparse dictionary learning to produce a more robust description of infrared pictures, which will, in turn enhance classification results.

CONCLUSION

In terms of breast cancer prediction, the Random Forest with the accuracy (92.5%), which uses Matlab programming, Random Forest appeared to produce better outcomes when compared to Multilayer Perceptron with accuracy (90%). In addition, unlike other approaches, the algorithm's performance improved as the amount of deep learning. This unique model shows great potential for improving breast cancer accuracy, making it suitable for usage in endocrine centers and hospitals.

REFERENCES

[1] Abdel-Nasser, Mohamed, Antonio Moreno, and Domenec Puig. 2016. "Temporal Mammogram Image Registration Using Optimized Curvilinear Coordinates." *Computer Methods and Programs in Biomedicine*. https://doi.org/10.1016/j.cmpb.2016.01.019.

[2] Acharya, U. Rajendra, U. Rajendra Acharya, E. Y. K. Ng, Jen-Hong Tan, and S. Vinitha Sree. 2012. "Thermography Based Breast Cancer Detection Using Texture Features and Support Vector Machine." *Journal of Medical Systems*. https://doi.org/10.1007/s10916-010-9611-z.

[3] Boogerd, Leonora S. F., Henricus J. M. Handgraaf, Hwai-Ding Lam, Volkert A. L. Huurman, Arantza Farina-Sarasqueta, John V. Frangioni, Cornelis J. H. van de Velde, Andries E. Braat, and Alexander L. Vahrmeijer. 2017. "Laparoscopic Detection and Resection of Occult Liver Tumors of Multiple Cancer Types Using Real-Time near-Infrared Fluorescence Guidance." *Surgical Endoscopy*. https://doi.org/10.1007/s00464-016-5007-6.

[4] Chiarelli, Anna M., Maegan V. Prummel, Derek Muradali, Rene S. Shumak, Vicky Majpruz, Patrick Brown, Hedy Jiang, Susan J. Done, and Martin J. Yaffe. 2015. "Digital versus Screen-Film Mammography: Impact of Mammographic

Density and Hormone Therapy on Breast Cancer Detection." *Breast Cancer Research and Treatment*. https://doi.org/10.1007/s10549-015-3622-x.

[5] Cho, Nariya, Wonshik Han, Boo-Kyung Han, Min Sun Bae, Eun Sook Ko, Seok Jin Nam, Eun Young Chae, et al. 2017. "Breast Cancer Screening With Mammography Plus Ultrasonography or Magnetic Resonance Imaging in Women 50 Years or Younger at Diagnosis and Treated With Breast Conservation Therapy." *JAMA Oncology*. https://doi.org/10.1001/jamaoncol.2017.1256.

[6] Díaz-Cortés, Margarita-Arimatea, Noé Ortega-Sánchez, Salvador Hinojosa, Diego Oliva, Erik Cuevas, Raúl Rojas, and Anton Demin. 2018. "A Multi-Level Thresholding Method for Breast Thermograms Analysis Using Dragonfly Algorithm." *Infrared Physics & Technology*. https://doi.org/10.1016/j.infrared.2018.08.007.

[7] Faust, Oliver, U. Rajendra Acharya, E. Y. K. Ng, Tan Jen Hong, and Wenwei Yu. 2014. "Application of Infrared Thermography in Computer Aided Diagnosis." *Infrared Physics & Technology*. https://doi.org/10.1016/j.infrared.2014.06.001.

[8] Hamidinekoo, Azam, Erika Denton, Andrik Rampun, Kate Honnor, and Reyer Zwiggelaar. 2018. "Deep Learning in Mammography and Breast Histology, an Overview and Future Trends." *Medical Image Analysis*. https://doi.org/10.1016/j.media.2018.03.006.

[9] Kennedy, Deborah A., Tanya Lee, and Dugald Seely. 2009. "A Comparative Review of Thermography as a Breast Cancer Screening Technique." *Integrative Cancer Therapies*. https://doi.org/10.1177/1534735408326171.

[10] Kosus, Nermin, Aydin Kosus, Muzeyyen Duran, Serap Simavli, and Nilgun Turhan. 2010. "Comparison of Standard Mammography with Digital Mammography and Digital Infrared Thermal Imaging for Breast Cancer Screening." *Journal of the Turkish German Gynecological Association*. https://doi.org/10.5152/jtgga.2010.24.

[11] Litjens, Geert, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A. W. M. van der Laak, Bram van Ginneken, and Clara I. Sánchez. 2017. "A Survey on Deep Learning in Medical Image Analysis." *Medical Image Analysis*. https://doi.org/10.1016/j.media.2017.07.005.

[12] Malvezzi, M., G. Carioli, P. Bertuccio, P. Boffetta, F. Levi, C. La Vecchia, and E. Negri. 2018. "European Cancer Mortality Predictions for the Year 2018 with Focus on Colorectal Cancer." *Annals of Oncology*.

https://doi.org/10.1093/annonc/mdy033.

[13] Mambou, Sebastien, Petra Maresova, Ondrej Krejcar, Ali Selamat, and Kamil Kuca. 2018. "Breast Cancer Detection Using Infrared Thermal Imaging and a Deep Learning Model." *Sensors*. https://doi.org/10.3390/s18092799.

[14] Mookiah, Muthu Rama Krishnan, U. Rajendra Acharya, and E. Y. K. Ng. 2012. "Data Mining Technique for Breast Cancer Detection in Thermograms Using Hybrid Feature Extraction Strategy." *Quantitative InfraRed Thermography Journal*. https://doi.org/10.1080/17686733.2012.738788.

[15] Rampun, Andrik, Bryan Scotney, Philip Morrow, Hui Wang, and John Winder. 2018. "Breast Density Classification Using Local Quinary Patterns with Various Neighbourhood Topologies." *Journal of Imaging*. https://doi.org/10.3390/jimaging4010014.

[16] Saniei, Elham, Saeed Setayeshi, Mohammad Esmaeil Akbari, and Mitra Navid. 2015. "A Vascular Network Matching in Dynamic Thermography for Breast Cancer Detection." *Quantitative InfraRed Thermography Journal*. https://doi.org/10.1080/17686733.2015.1005398.

[17] Sathish, Dayakshini, Surekha Kamath, K. V. Rajagopal, and Keerthana Prasad. 2016. "Medical Imaging Techniques and Computer Aided Diagnostic Approaches for the Detection of Breast Cancer with an Emphasis on Thermography - a Review." *International Journal of Medical Engineering and Informatics*. https://doi.org/10.1504/ijmei.2016.077446.

[18] Selle, J. Josephine, J. Josephine Selle, A. Shenbagavalli, B. Venkatraman, M. Menaka, and M. Jayashree. 2015. "Automated Segmentation for Quantitative Analysis of Breast Thermograms." *Proceedings of the 2015 Asia International Conference on Quantitative InfraRed Thermography*. https://doi.org/10.21611/qirt.2015.0035.

[19] Zhou, Yan, and Cila Herman. 2018. "Optimization of Skin Cooling by Computational Modeling for Early Thermographic Detection of Breast Cancer." *International Journal of Heat and Mass Transfer*. https://doi.org/10.1016/j.ijheatmasstransfer.2018.05.129.