Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 20 No. 4 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v20i04.46387

PAPER

CNN-Based Approach for Non-Invasive Estimation of Breast Tumor Size and Location Using Thermographic Images

Zakaryae Khomsi(✉), Mohamed El Fezazi, Larbi Bellarbi

ENSAM-ENSIAS, ST2I Laboratory, E2SN Team, Mohammed V University in Rabat, Rabat, Morocco

zakaryae_khomsi@um5.ac.ma

ABSTRACT

The characterization of tumors is crucial for guiding appropriate treatment strategies and enhancing patient survival rates. Surface thermography shows promise in the non-invasive detection of thermal patterns associated with the existence of breast tumors. Nevertheless, the precise prediction of both tumor size and location using temperature characteristics presents a critical challenge. This is due to the limited availability of thermal images labeled with the corresponding tumor size and location. This work proposes a deep learning approach based on convolutional neural networks (CNN) in combination with thermographic images for estimating breast tumor size and location. Successive COMSOL-based simulations are conducted, including a 3D breast model with various tumor scenarios. Thus, different noise levels were included in the development of the thermographic image dataset. Every image was accordingly labeled with the corresponding tumor location and size to train the CNN model. Mean absolute error (MAE) and the coefficient of determination (R^2) were considered as evaluation metrics. The results show that the proposed CNN model achieved a reasonable prediction performance with MAE–R^2 values of 0.872–98.6% for tumor size, 1.161–96.8% for x location, 1.086–97.1% for y location, and 0.954–96.7% for z location. This study indicates that the combination of surface thermography and deep learning is a convenient tool for predicting breast tumor parameters.

KEYWORDS

3D tumor localization, finite element method (FEM), surface thermography, deep learning, breast cancer

1 INTRODUCTION

Breast cancer is a significant global health concern, affecting numerous women annually. The World Health Organization (WHO) estimated 2.3 million new cases of breast cancer and 685,000 deaths worldwide in 2020 [1]. Various factors can lead to

Khomsi, Z., Fezazi, M.E., Bellarbi, L. (2024). CNN-Based Approach for Non-Invasive Estimation of Breast Tumor Size and Location Using Thermographic Images. *International Journal of Online and Biomedical Engineering (iJOE)*, 20(4), pp. 160–175. https://doi.org/10.3991/ijoe.v20i04.46387

Article submitted 2023-10-31. Revision uploaded 2023-12-08. Final acceptance 2023-12-08.

breast cancer, including family history, radiation exposure, obesity, age, radiation, and several other factors [2–4]. The presence of a breast tumor can progress from early-stage to more advanced stages [5], [6]. Precise knowledge of tumor parameters informs the selection of appropriate therapeutic interventions, with smaller tumors often warranting less invasive treatments compared to their larger counterparts [7]. Furthermore, understanding the spatial location of tumors within the breast plays a fundamental role in planning surgical procedures, optimizing radiation therapy, and ultimately improving patient outcomes [8]. Therefore, the reliable estimation of tumor size and location not only enhances the quality of patient care but also contributes to the advancement of breast cancer management and prognosis [9].

Mammography is a commonly used screening tool to characterize breast tumors [10]. It is an imaging modality that uses an important dose of X-rays to create detailed images of the breast tissue. It works on the principle of differential attenuation of X-rays by the tissue. The breast is compressed between two plates to achieve a uniform thickness, which reduces scatter and ensures maximum contact between the breast and the image receptor. The X-rays are then passed through the breast, and the resulting image is captured on a digital detector [11]. A radiologist analyzes the captured image to detect breast tumors [12]. However, there are several limitations associated with mammography. Firstly, mammography has lower sensitivity for women with dense breast tissue, which can increase the chances of false negatives. Secondly, mammography involves exposure to ionizing radiation, which can increase the risk of radiation-induced cancer in the long term [13], [14]. Additionally, mammography can be uncomfortable for some women due to the compression of the breast during the imaging process [15]. Therefore, there is a need for alternative, non-invasive screening methods for breast tumor characterization.

Breast thermography is an emerging modality for early detection of breast cancer due to its non-invasiveness, no radiation exposure, and low cost [16], [17]. The principle of breast thermography is based on the fact that cancer cells generate more heat than normal cells due to their increased metabolic activity [18], [19]. By analyzing the heat difference captured by an infrared camera, image processing techniques are used to create a thermographic image of the breast. The resulting image shows temperature variations in the breast tissue, indicating any potential abnormalities that could be indicative of cancer. In their study, Kakileti et al. [20] explored the recent advancements in thermal cameras with computer-aided diagnosis systems (CAD) to enhance the accuracy of breast cancer detection. Singh et al. [21] conducted a comprehensive survey outlining the key steps involved in CAD for breast thermography. They recommended the exploitation of advanced simulation methods for better prediction. Kandlikar et al. [22] examined the advancements in employing surface thermography for breast cancer detection spanning the past three decades. The authors identified distinct areas needing refinement to enhance thermography as a dependable diagnostic tool for breast tumor prediction. They recommended the incorporation of inverse modeling using advanced simulation methods to formulate artificial neural networks dedicated to early breast tumor detection.

As revealed in the prior studies, surface thermography is a secure method with significant capacity for identifying the existence of breast tumors. Nevertheless, the precise determination of tumor location and size using thermal properties presents a notable challenge, primarily due to the scarcity of thermographic images containing labeled tumor sizes and locations. In this context, advanced simulation methods emerge as a critical solution, offering the ability to generate controlled and diverse datasets that closely mimic the thermal patterns of the breast [23], [24]. Through the finite element method (FEM) and other advanced techniques, these simulations

provide representative data, facilitating rigorous AI model training and testing. For example, Majdoubi et al. [25] introduced a non-intrusive method for the estimation of tumor location and size based on a feed-forward neural network (FFNN). By using FEM, they generated 122 cases of thermal distribution of the breast by varying tumor size and location. The researchers demonstrated the FFNN model's capability in predicting tumor parameters. Venkatapathy et al. [26] generated 599 thermal images of breast models using FEM. Subsequently, they conducted training utilizing a range of machine learning (ML) models. The authors employed scatter plots for result analysis and employed the MAE as an evaluative metric. Their findings indicated the superior predictive capabilities of the decision tree regression (DTR) model in comparison to other ML models. The main limitations associated with these approaches can be outlined as follows: (1) These studies did not incorporate realistic breast modeling into their data-generation processes. It is essential for breast modeling to account for biological components such as skin, fat, glands, and muscle to achieve a thermal approximation of real breast tissues [27]. (2) These approaches were evaluated only for noise-free thermal data and not investigated in noisy conditions. (3) The selected tumor cases are confined to specific scenarios, resulting in a limited amount of training data.

In this study, we proposed a comprehensive framework for an accurate estimation of tumor size and location using breast thermographic images, which contributes to addressing the above critical challenges. As previously mentioned, numerous studies have been conducted on this subject; nevertheless, our approach differs in multiple aspects. First, we improved the FEM-based breast modeling through a realistic 3D breast model. Second, we selected multiple tumor scenarios, considering noisy thermal images, to enhance the amount of training data. Third, we exploited the potential of CNN model for automatically extracting features from thermal images, thus characterizing tumor parameters. Our contributions could be summarized as follows:

- We proposed a CNN model for estimating the breast tumor location and size based on thermographic images. This novel approach enhances the use of thermography for non-invasive breast tumor characterization.
- We generated a diverse and well-labeled dataset that incorporates noise-free and noisy thermal images for different sizes and locations of breast tumors. The application of different noise intensities allows the approximation to realistic conditions.
- We investigated the impact of various noise levels on thermographic images in the estimation of tumor parameters. That provides relevant insights into selecting an acceptable range of noise for better prediction performance.

The paper is organized as follows: In Section 2, we describe our method. Section 3 presents the results, while Section 4 discusses the findings. Finally, in Section 5, we conclude the paper.

2 METHODOLOGY

The methodology comprises the following phases: FEM-based data generation (including noise-free and noisy thermal images), data preparation (which includes data labeling and splitting), and finally the formulation of the CNN model. Figure 1 illustrates the conceptual diagram of the proposed framework.

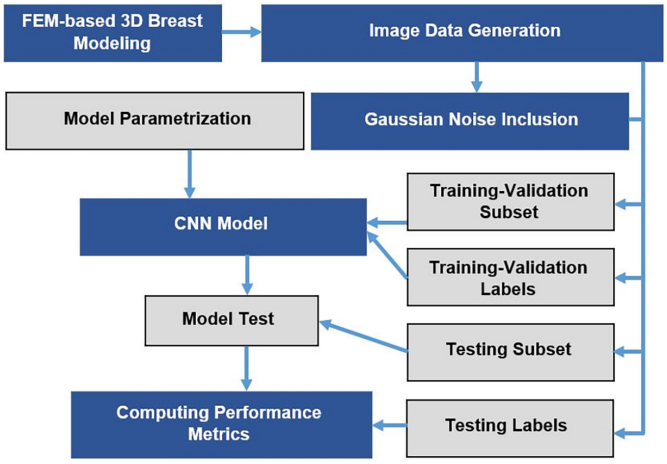

Fig. 1. Conceptual diagram of the proposed approach

2.1 FEM-based breast tissue modeling

The mammary tissue is a complex anatomical structure that consists of multiple layers, including the skin, adipose tissue (fat), glandular tissue, and muscular tissue. Cancer is a malignant tumor that can occur in glandular tissue in different scenarios. Figure 2 illustrates the representation of the mammary tissue structure utilized in the breast cancer modeling process based on COMSOL software.

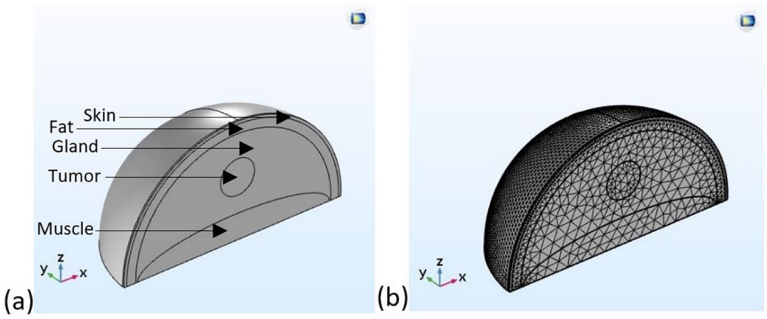

Fig. 2. (a) Geometry of breast layers; and (b) Mesh generated

We applied the Pennes equation (1) to describe the thermal dynamics within the tissues [28]. The equation accounts for heat generation due to metabolism, heat conduction through tissues, and convective heat transfer through blood perfusion. By incorporating these factors, the Pennes equation provides a comprehensive framework for simulating and analyzing the temperature distribution in tissues. This modeling allows us to gain insights into the surface temperature variations associated with different tumor sizes and locations.

$$\nabla(k_l \cdot \nabla T_l) + c_b \cdot \rho_b \cdot \omega_{b,l}(T_b - T_l) + q_{m,l} = 0 \quad (1)$$

Where l represents the breast layer, including muscle, gland, fat, skin, and tumor. T_b is the arterial blood temperature (37°C), K_l is the thermal conductivity for each tissue, $\omega_{b,l}$ is the blood perfusion rate for each tissue. T_l, and $q_{m,l}$ is respectively the temperature and the metabolic heat generation rate of the tissues. The specific heat c_b and the density ρ_b of the blood are set at 3770 J/(kg.K) and 1060 kg/m^3 respectively.

The temperature of the room is considered to be 25°C, representing a typical low medical room temperature. The specific thermophysical properties used in the model are provided in Table 1 [29–31].

Table 1. Thermophysical properties of breast tissue

	Thickness (mm)	k (W/m.K)	q_m (W/m^3)	ω_b (ml.s^{-1}.ml^{-1})
Skin	1.6	0.45	368.1	0.00018
Fat	5.0	0.21	400	0.00022
Gland	43.4	0.48	700	0.00054
Muscle	15	0.48	700	0.00270
Tumor	–	0.62	70,000	0.01600

The first simulation was performed in a steady-state condition, revealing that the presence of the tumor induces an observable thermal gradient on the breast's surface, as depicted in Figure 3.

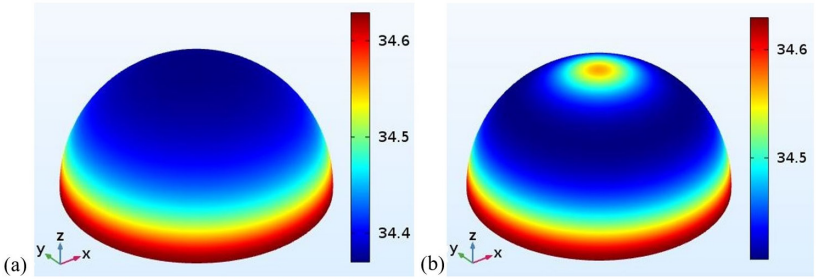

Fig. 3. Breast surface temperature: (a) normal; and (b) with tumor

2.2 Range selection of tumor parameters

In the field of cancer staging, researchers have delineated multiple categories, differentiating tumors based on their size and progression from early to advanced stages. The initial category pertains to tumors with sizes less than 20 mm, followed by a second category encompassing tumors falling within the range of 20 mm to 50 mm [32]. The subsequent categories encapsulate tumors exceeding 50 mm in size, indicative of advanced-stage malignancies, which are out of the scope of this study. In our investigation, we focused on early-stage tumors ranging from 2 mm to 40 mm, with increments of 1 mm, all incorporated into the modeling. As indicated in Section 2.1, the modeling framework was adopted to export thermographic images through numerical simulations. We selected different tumor location coordinates (x, y, and z) ranging from (–52, –51, and 15) to (53, 51, and 57) while respecting the boundaries delineated by the gland of breast tissue. The various permutations of tumor sizes and locations yielded a total of 1406 different cases within our simulated models.

2.3 Data collection

The data collection process for this study started with the extraction of 1406 thermal images, each representing a unique case generated through

COMSOL software. These images serve as the foundational dataset, offering a diverse range of thermal patterns for our CNN model to analyze and predict tumor size and location.

To further enrich our dataset and evaluate the robustness of the CNN model, we introduced controlled variations in the form of noise. We applied different noise levels to the imported thermal images. Specifically, three distinct Gaussian noise levels, 0.01, 0.05, and 0.1, were introduced to the noise-free images, replicating the realistic thermographic scenarios [33]. The mathematical expression of the Gaussian noise applied to thermal images is formulated in equation (2) [34]. This inclusion of noisy data allows for assessing the model's generalizability under varying conditions. Figure 4 shows an example of the thermographic images with and without noise.

$$N(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2}$$

Where,
x: represents the gray value
μ: the mean value of x
σ: the standard deviation of x (noise level)

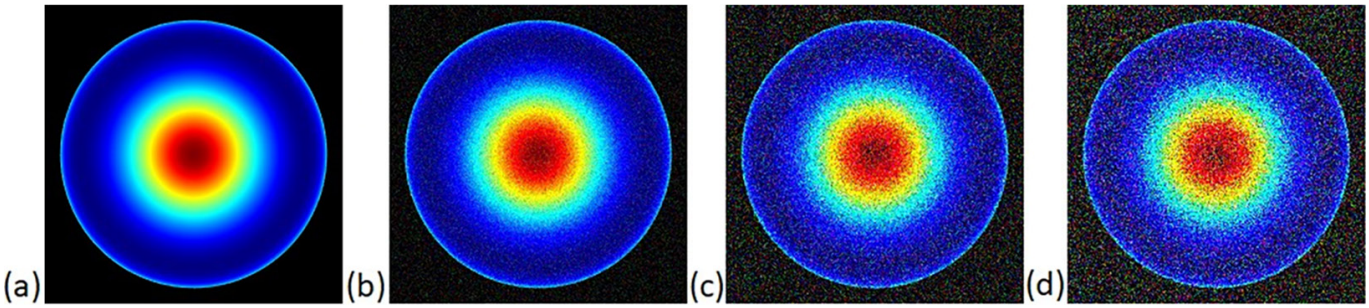

Fig. 4. Example of thermal images generated with included tumor for different noise levels: (a) Noise-Free; (b) 0.01 noise; (c) 0.05 noise; and (d) 0.1 noise

2.4 Architecture of the CNN model

The architecture of our proposed CNN model is designed to make accurate predictions of tumor parameters from thermal images. The CNN comprises a series of interconnected layers, including the convolutional layer, which uses filters to identify local features and spatial relationships, as shown in Figure 5, the pooling layers to reduce data dimensionality, and the fully connected layers to make predictions from extracted features. The mathematical expression of standard convolution is formulated in equation (3) [35].

$$G_{k,l,n} = \sum_{i,j,m} K_{i,j,m,n} . I_{k+i-1,l+j-1,m} \tag{3}$$

Where,
$G_{k,l,n}$: output of standard convolution
I: input image
K: convolution kernel

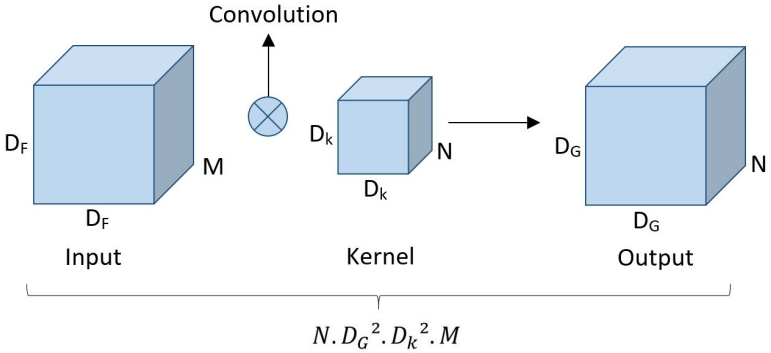

Fig. 5. Convolution principle

Figure 6 represents the global architecture adopted for predicting tumor size and its locations *x*, *y*, and *z*. The model's architecture employed begins with an input layer having dimensions of 224 × 224 × 3, corresponding to an RGB image. The initial convolutional layer is equipped with 32 filters of size 3 × 3, allowing the extraction of relevant features from input data. To introduce non-linearity and capture complex patterns in the data, the rectified linear unit (ReLU) activation function (4) [36] is applied to this convolutional layer. Subsequently, a max-pooling layer with a 2 × 2 pool size is employed to efficiently decrease spatial dimensions and computational complexity by downsampling the feature maps. The extracted features are subsequently flattened and pass through the fully connected layer, consisting of two hidden layers with 64 units each. Ultimately, the output layer employs a linear activation function to produce the final prediction of tumor parameters, including size, *x*, *y*, and *z*.

$$ReLU(z) = \begin{cases} 0 if\ z < 0 \\ z if\ z \geq 0 \end{cases} \tag{4}$$

Where, *z* is the input.

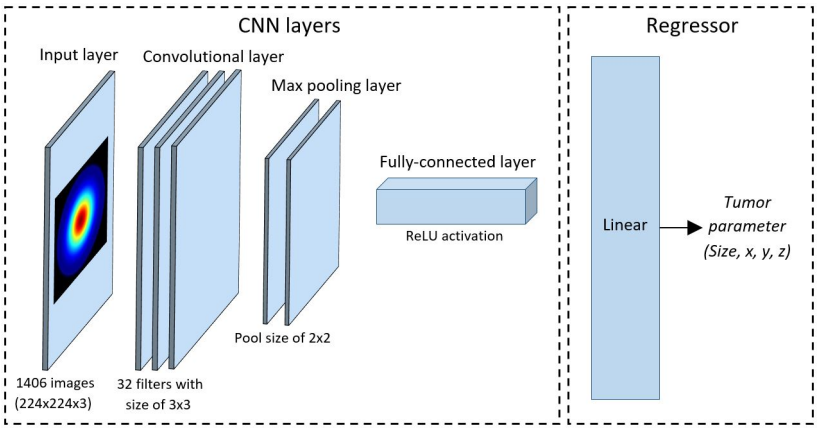

Fig. 6. CNN architecture adopted for prediction of tumor parameters

3 RESULTS

In this section, we provide a comprehensive evaluation of the proposed model for predicting breast tumor parameters (size and location) from thermographic images under various conditions. By fine-tuning the hyperparameters of the CNN in

the training process, we examined the model's response to varying Gaussian noise levels, from noise-free conditions to heightened noise. We considered two primary metrics for the model's evaluation, especially the mean absolute error (*MAE*) and the coefficient of determination (R^2) computed by the equations (5) and (6), respectively [37].

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right| \tag{5}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{6}$$

Where y_i is the actual value, $\hat{y}_i$ is predicted value, N is number of examples, $\bar{y}$ is the mean of y_i.

We split our dataset into three parts with the following ratios: 80% for the training set, 10% for the validation set, and 10% for the test set. The training set, constituting the largest portion, serves as the foundation for training the CNN model, enabling it to learn and adapt to the patterns present in the data. The validation set plays a crucial role in fine-tuning the model during training, allowing for the adjustment of hyperparameters and the prevention of overfitting. Finally, the test set allows for the assessment of the model's generalization capabilities, providing an efficient evaluation of its predictive performance.

3.1 Tumor size prediction

For the noise-free scenario, the model achieved higher accuracy compared to other cases (MAE = 0.872) and showed good correlation (R^2 = 98.6%) between predicted and actual tumor sizes, as shown in Figure 7a. When the noise level rose to 0.01 and 0.05, the model still performed well, with only slight increases in MAE (1.041 for 0.01 noise and 1.419 for 0.05 noise), while maintaining good predictive capabilities (Figure 7b and c). At a noise level of 0.1, the model encountered more dispersed scatter plots (R^2 = 95.1%) as shown in Figure 7d, causing a small rise in MAE (1.804).

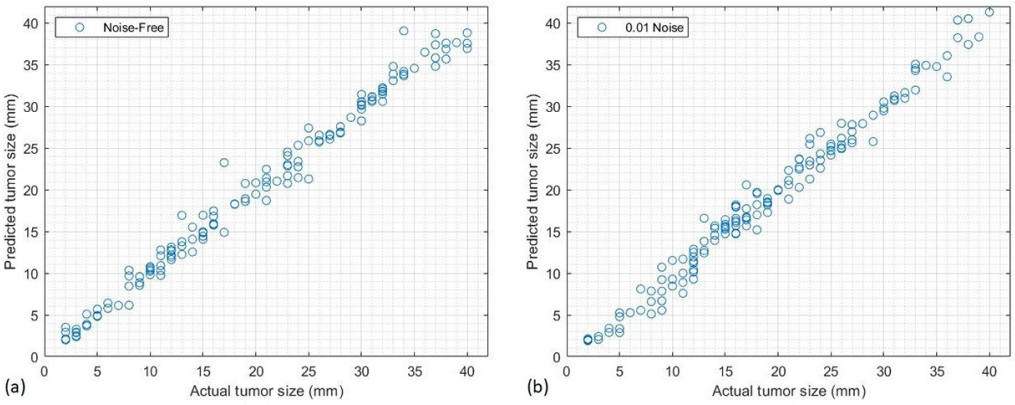

Fig. 7. *(Continued)*

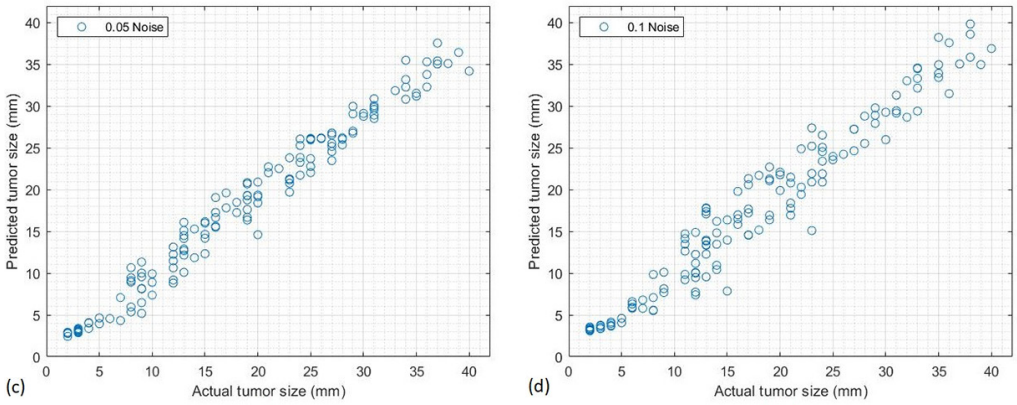

Fig. 7. Scatter plots of actual vs. predicted tumor size values at different noise levels: (a) noise-free; (b) $\sigma = 0.01$; (c) $\sigma = 0.05$; and (d) $\sigma = 0.1$

3.2 Location coordinates (x, y, and z) prediction

We evaluated the model's efficacy in predicting spatial coordinates (x, y, and z) under these varying noise conditions. The following scatter plots depict the relationship between the actual and predicted x, y, and z locations, as shown in Figures 8, 9, and 10, respectively.

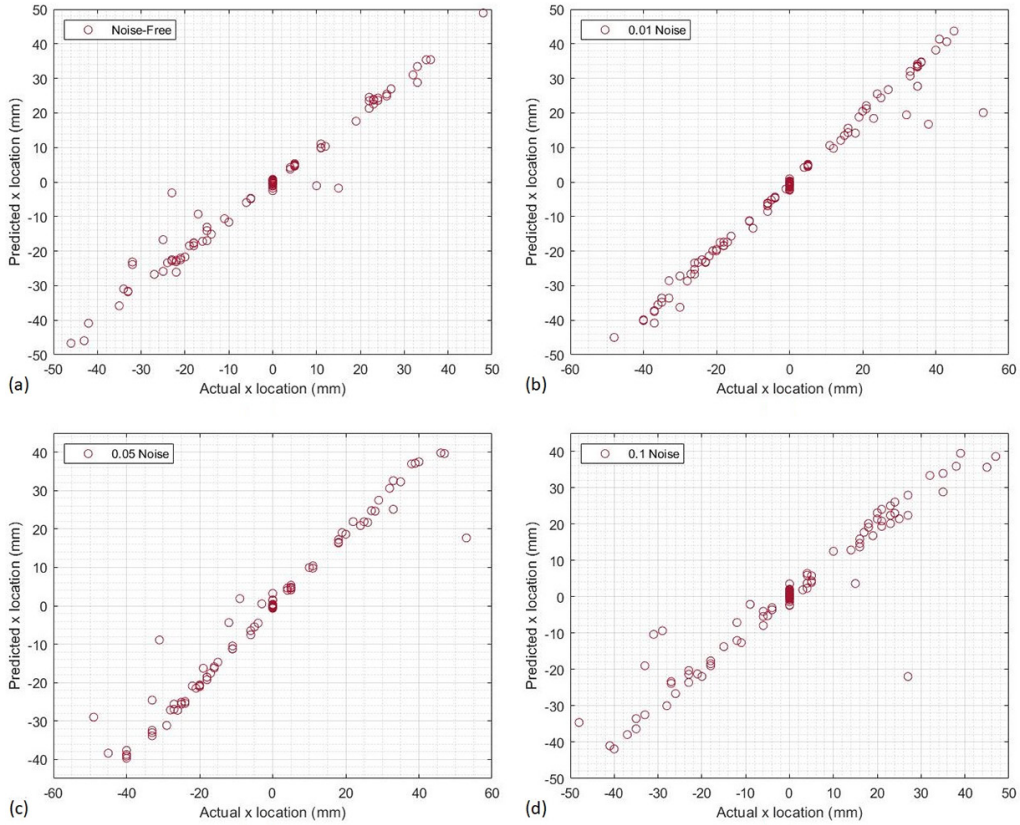

Fig. 8. Scatter plots of actual vs. predicted x location values at different noise levels: (a) noise-free; (b) $\sigma = 0.01$; (c) $\sigma = 0.05$; and (d) $\sigma = 0.1$

In all the noise-free scenarios, the model achieved a strong correlation, as indicated by the R^2 values (96.8% for x, 93.7% for y, and 96.7% for z) and acceptable MAE values (1.161 for x, 1.615 for y, and 0.954 for z). In the presence of heightened noise at 0.01, 0.05, and 0.1, the model is affected relatively by the noise perturbations, especially in the case of the z location, as shown in Figure 10c and d. Besides, it adeptly retained its predictive accuracy for estimating the location coordinates in challenging conditions.

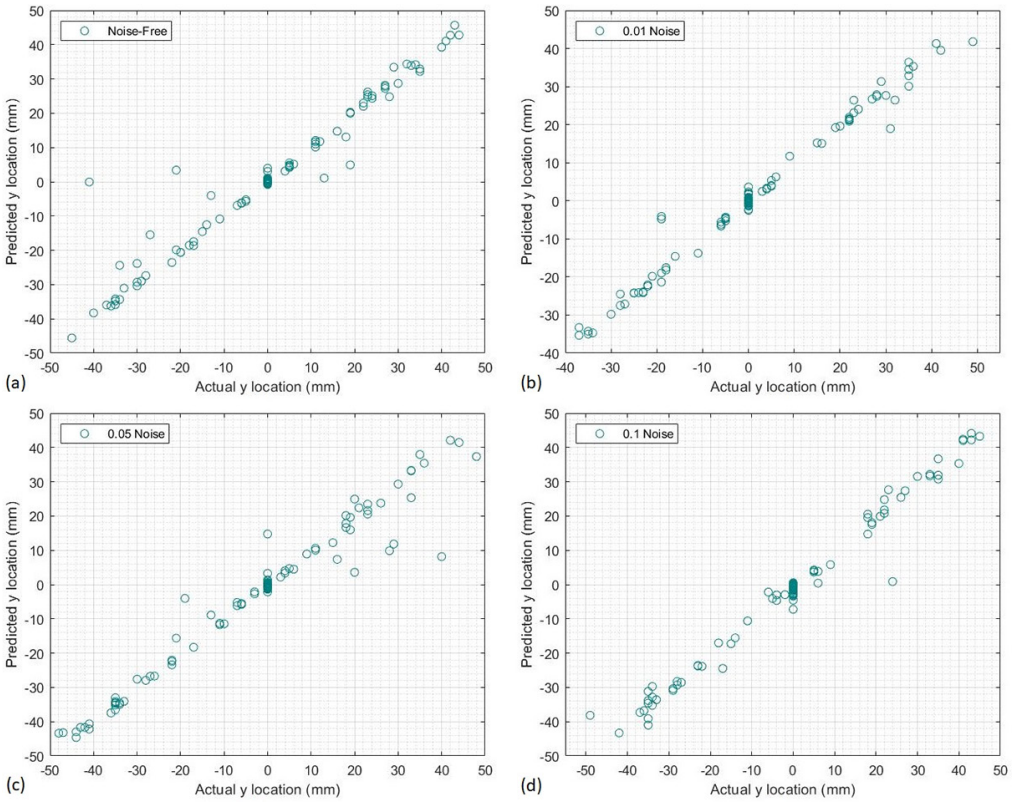

Fig. 9. Scatter plots of actual vs. predicted y location values at different noise levels: (a) noise-free; (b) $\sigma = 0.01$; (c) $\sigma = 0.05$; and (d) $\sigma = 0.1$

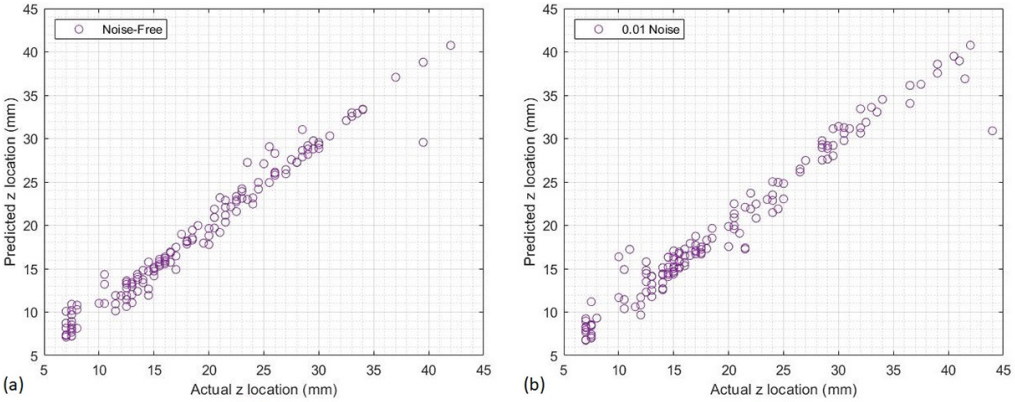

Fig. 10. *(Continued)*

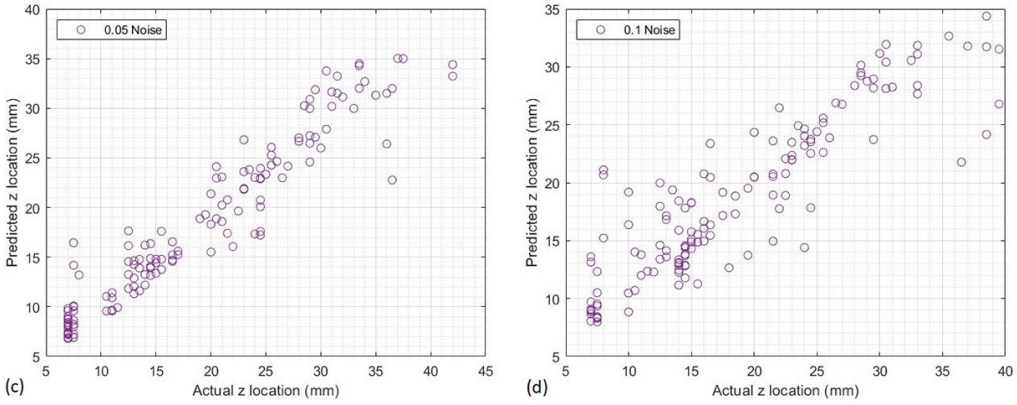

Fig. 10. Scatter plots of actual vs. predicted z location values at different noise levels: (a) noise-free; (b) $\sigma = 0.01$; (c) $\sigma = 0.05$; and (d) $\sigma = 0.1$

Table 2 and Figure 11 summarize the obtained results for the prediction performance of the proposed model. Generally, we note that the CNN model presents less prediction performance in noisy images compared to noise-free images. Nevertheless, in some cases; adding a controlled amount of Gaussian noise to images can lead to improved accuracy. This phenomenon is observed in the prediction of y location cases, as we add noise at 0.01, the MAE becomes smaller and R^2 becomes larger compared to noise-free cases. That means that the Gaussian noise can sometimes act as a form of regularization during the training of deep neural networks.

Table 2. Comparison of the prediction performance results

	Size		x		y		z	
	MAE	R^2	MAE	R^2	MAE	R^2	MAE	R^2
Noise-Free	0.872	98.6%	1.161	96.8%	1.615	93.7%	0.954	96.7%
0.01 Noise	1.041	97.9%	1.386	96.5%	1.086	97.1%	1.172	95.7%
0.05 Noise	1.419	97%	1.58	94.7%	1.835	94.9%	2.068	90.2%
0.1 Noise	1.804	95.1%	2.163	89.8%	1.651	97.6%	2.681	78%

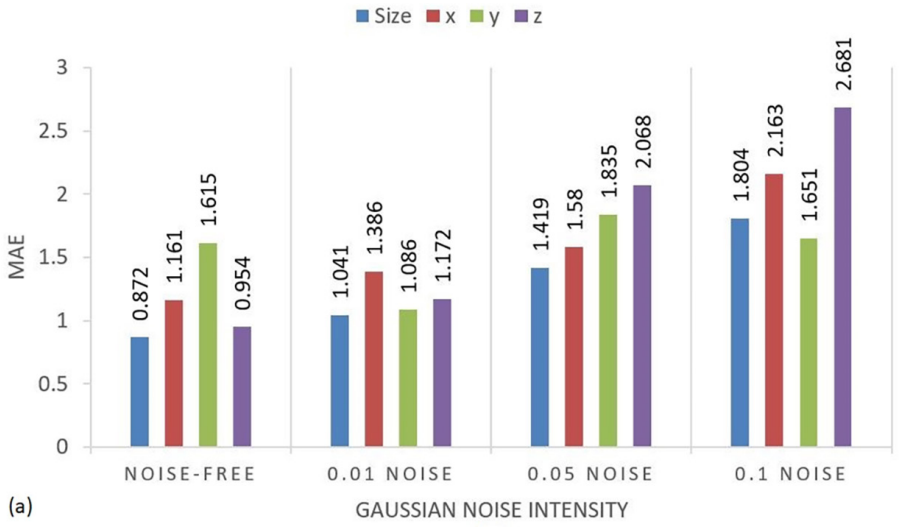

Fig. 11. *(Continued)*

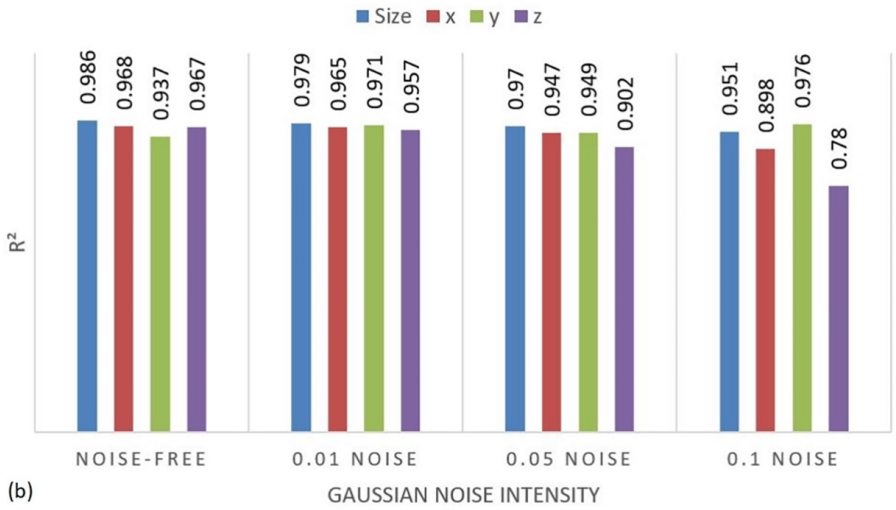

Fig. 11. Performance histogram for different tumor parameters: (a) MAE; and (b) R^2

4 DISCUSSION

This paper introduces a CNN-based approach for the estimation of breast tumor size and location using thermographic images. Using FEM, we first generated 1406 thermographic images for different breast tumor scenarios exported from a realistic breast model (muscle, fat, gland, skin, and tumor). Second, we applied various Gaussian noise intensities (0.01, 0.05, and 0.1), which enriched the capabilities of the predictive model across different scenarios. We prepared the dataset by labeling each image with the corresponding tumor size and location. Thus, we split our dataset into training, validation, and testing. Furthermore, through the computation of the performance metrics MAE and R^2, we presented the overall assessment of the CNN model. Specifically, we reported an MAE and R^2 ranging from 0.872 to 1.804 and 95.1% to 98.6%, respectively, for size, 1.161 to 2.163 and 89.8% to 96.8% for x location, 1.086 to 1.835 and 93.7% to 97.6% for y location, 0.954 to 2.681 and 78% to 96.7% for z location.

These results are compared with other studies in the literature. For example, Majdoubi et al. [25] created a dataset comprising 122 instances of temperature vectors for defined tumor scenarios exported from a simplified breast model. Based on a simple FFNN (18 units, 1 hidden layer), they conducted tumor size estimation using raw temperature data from the surface. The model's evaluation was carried out only on seven testing samples, which did not reflect a credible assessment of the predictive model's performance. Their results indicated R^2 and MAE values of 93.3% and 1.471, respectively, for tumor size, 86.43% an 5.42 for x location, 98.3% and 2.14 for y location, and 85.94% and 2 for z location. Venkatapathy et al. [26] utilized a dataset of thermal images containing 599 tumor cases, simulated on a simplified breast model. Various machine learning models (including SVR, LR, DTR, and KNNR regression models) were employed to estimate tumor size and location from thermal images. They considered the area of hot spots, the number of hot spots, and the maximum surface temperature as the input features for the prediction of tumor parameters. Their findings revealed that the DTR model exhibited superior performance compared to the other machine learning models, achieving an MAE value of 0.025 for size, 0.0048 for x, and 0.037 for z, respectively. These differences within our study can be explained by several

reasons. The authors did not consider the combination of realistic breast components in the breast modeling. They approximated the breast only as fat or skin tissue. This affects the reliability of the generated thermal data, which is investigated by studies [38] and [39]. In addition, the thermographic data are generated under noise-free conditions and do not investigate the influence of noisy scenarios.

Nonetheless, our study introduces certain limitations. Only standard breast size is considered, underscoring the importance of investigating different breast sizes. The estimation of tumor parameters is performed only for the size and location of the tumor, whereas other tumor characteristics can be explored, such as metabolism and heat generation in the tumor.

Finally, the utilization of the CNN model in this study has achieved acceptable predictive accuracy, affirming the potential of this approach for non-invasive estimation of breast tumor size and location from thermal images.

5 CONCLUSION

In this paper, we present a convenient CNN-based framework for non-invasive estimation of breast tumor size and location from thermal images. Firstly, we performed the FEM-based 3D breast modeling, incorporating key tissue components, namely skin, fat tissue, mammary gland, and muscle, using COMSOL software. A large number of thermographic images were generated, including a wide range of tumor scenarios. Then, we introduced three Gaussian noise levels, 0.01, 0.05, and 0.1, to the noise-free images. That enhanced the predictive model's behavior in various conditions. Next, the final dataset was prepared by labeling the thermographic images with the corresponding tumor size and location values to train the deep-learning model. Based on the results, our approach indicates the potential of the proposed CNN to estimate tumor location and size with an acceptable prediction performance. Thus, the progressive inclusion of noise led to a slight elevation in the prediction error. It is important to note that in certain specific cases, noise contributes to the regularization, enhancing the model's ability to generate more precise spatial estimations. In conclusion, this study's results offer a detailed method that combines a realistic breast model, noisy conditions, and a deep-learning model to improve the characterization of breast tumors using a non-invasive tool. In future work, we aim to investigate multiple breast sizes and the prediction of additional tumor parameters.

6 ACKNOWLEDGMENT

We thank Al-Khawarizmi program, Rabat, Morocco for encouraging research in artificial intelligence and its applications.

7 REFERENCES

[1] World Health Organization, "Breast cancer." https://www.who.int/news-room/fact-sheets/detail/breast-cancer. [Accessed: Feb. 01, 2023].

[2] M. E. Barnard, C. E. Boeke, and R. M. Tamimi, "Established breast cancer risk factors and risk of intrinsic tumor subtypes," *Biochim. Biophys. Acta – Rev. Cancer*, vol. 1856, no. 1, pp. 73–85, 2015. https://doi.org/10.1016/j.bbcan.2015.06.002

[3] Y. S. Sun *et al.*, "Risk factors and preventions of breast cancer," *Int. J. Biol. Sci.*, vol. 13, no. 11, pp. 1387–1397, 2017. https://doi.org/10.7150/ijbs.21635

[4] M. Kamińska, T. Ciszewski, K. Łopacka-Szatan, P. Miotła, and E. Starosławska, "Breast cancer risk factors," *Prz. Menopauzalny*, vol. 14, no. 3, pp. 196–202, 2015. https://doi.org/10.5114/pm.2015.54346

[5] J. S. Michaelson *et al.*, "Predicting the survival of patients with breast carcinoma using tumor size," *Cancer*, vol. 95, no. 4, pp. 713–723, 2002. https://doi.org/10.1002/cncr.10742

[6] S. C. Tate, V. Andre, N. Enas, B. Ribba, and I. Gueorguieva, "Early change in tumour size predicts overall survival in patients with first-line metastatic breast cancer," *Eur. J. Cancer*, vol. 66, pp. 95–103, 2016. https://doi.org/10.1016/j.ejca.2016.07.009

[7] P. Shahmirzalou, M. J. Khaledi, M. Khayamzadeh, and A. Rasekhi, "Survival analysis of recurrent breast cancer patients using mix Bayesian network," *Heliyon*, vol. 9, no. 10, p. e20360, 2023. https://doi.org/10.1016/j.heliyon.2023.e20360

[8] R. Ranjbarzadeh *et al.*, "Breast tumor localization and segmentation using machine learning techniques: Overview of datasets, findings, and methods," *Comput. Biol. Med.*, vol. 152, p. 106443, 2023. https://doi.org/10.1016/j.compbiomed.2022.106443

[9] A. A. Alsalihi, H. K. Aljobouri, and E. A. K. Altameemi, "GLCM and CNN deep learning model for improved MRI breast tumors detection," *Int. J. online Biomed. Eng.*, vol. 18, no. 12, pp. 123–137, 2022. https://doi.org/10.3991/ijoe.v18i12.31897

[10] K. Loizidou, R. Elia, and C. Pitris, "Computer-aided breast cancer detection and classification in mammography: A comprehensive review," *Comput. Biol. Med.*, vol. 153, p. 106554, 2023. https://doi.org/10.1016/j.compbiomed.2023.106554

[11] F. Sardanelli and T. H. Helbich, "Mammography: EUSOBI recommendations for women's information," *Insights Imaging*, vol. 3, no. 1, pp. 7–10, 2012. https://doi.org/10.1007/s13244-011-0127-y

[12] R. Murakami, N. Uchiyama, H. Tani, T. Yoshida, and S. Kumita, "Comparative analysis between synthetic mammography reconstructed from digital breast tomosynthesis and full-field digital mammography for breast cancer detection and visibility," *Eur. J. Radiol. Open*, vol. 7, p. 100207, 2020. https://doi.org/10.1016/j.ejro.2019.12.001

[13] A. B. De González and G. Reeves, "Mammographic screening before age 50 years in the UK: Comparison of the radiation risks with the mortality benefits," *Br. J. Cancer*, vol. 93, no. 5, pp. 590–596, 2005. https://doi.org/10.1038/sj.bjc.6602683

[14] A. B. de González and S. Darby, "Risk of cancer from diagnostic X-rays: Estimates for the UK and 14 other countries," *Lancet*, vol. 363, no. 9406, pp. 345–351, 2004. https://doi.org/10.1016/S0140-6736(04)15433-0

[15] V. A. Loving, S. Aminololama-Shakeri, and J. W. T. Leung, "Anxiety and its association with screening mammography," *J. Breast Imaging*, vol. 3, no. 3, pp. 266–272, 2021. https://doi.org/10.1093/jbi/wbab024

[16] A. Lozano and F. Hassanipour, "Infrared imaging for breast cancer detection: An objective review of foundational studies and its proper role in breast cancer screening," *Infrared Phys. Technol.*, vol. 97, pp. 244–257, 2019. https://doi.org/10.1016/j.infrared.2018.12.017

[17] A. Mashekova, Y. Zhao, E. Y. K. Ng, V. Zarikas, S. C. Fok, and O. Mukhmetov, "Early detection of the breast cancer using infrared technology – A comprehensive review," *Therm. Sci. Eng. Prog.*, vol. 27, p. 101142, 2022. https://doi.org/10.1016/j.tsep.2021.101142

[18] H. A. Coller, "Is cancer a metabolic disease?" *Am. J. Pathol.*, vol. 184, no. 1, pp. 4–17, 2014. https://doi.org/10.1016/j.ajpath.2013.07.035

[19] G. Kroemer and J. Pouyssegur, "Tumor cell metabolism: Cancer's achilles' heel," *Cancer Cell*, vol. 13, no. 6, pp. 472–482, 2008. https://doi.org/10.1016/j.ccr.2008.05.005

[20] S. T. Kakileti, G. Manjunath, H. Madhu, and H. V. Ramprakash, "Advances in breast thermography," *New Perspect. Breast Imaging*, 2017. https://doi.org/10.5772/intechopen.69198

[21] D. Singh and A. K. Singh, "Role of image thermography in early breast cancer detection-past, present and future," *Comput. Methods Programs Biomed.*, vol. 183, p. 105074, 2020. https://doi.org/10.1016/j.cmpb.2019.105074

[22] S. G. Kandlikar *et al.*, "Infrared imaging technology for breast cancer detection – Current status, protocols and new directions," *Int. J. Heat Mass Transf.*, vol. 108, pp. 2303–2320, 2017. https://doi.org/10.1016/j.ijheatmasstransfer.2017.01.086

[23] H. Murtaza, M. Ahmed, N. F. Khan, G. Murtaza, S. Zafar, and A. Bano, "Synthetic data generation: State of the art in health care domain," *Comput. Sci. Rev.*, vol. 48, p. 100546, 2023. https://doi.org/10.1016/j.cosrev.2023.100546

[24] Z. Khomsi, A. Elouerghi, and L. Bellarbi, "Towards development of synthetic data in surface thermography to enable deep learning models for early breast tumor prediction," *Artificial Intelligence and Industrial Applications*, pp. 356–365, 2023. https://doi.org/10.1007/978-3-031-43520-1_30

[25] J. Majdoubi *et al.*, "Estimation of tumor parameters using neural networks for inverse bioheat problem," *Comput. Methods Programs Biomed.*, vol. 205, p. 106092, 2021. https://doi.org/10.1016/j.cmpb.2021.106092

[26] G. Venkatapathy, A. Mittal, N. Gnanasekaran, and V. H. Desai, "Inverse estimation of breast tumor size and location with numerical thermal images of breast model using machine learning models," *Heat Transf. Eng.*, vol. 44, no. 15, pp. 1433–1451, 2022. https://doi.org/10.1080/01457632.2022.2134081

[27] Z. Khomsi, M. Elfezazi, and L. Bellarbi, "Deep learning-based approach in surface thermography for inverse estimation of breast tumor size," *Sci. African*, vol. 23, p. e01987, 2024. https://doi.org/10.1016/j.sciaf.2023.e01987

[28] H. H. Pennes, "Analysis of tissue and arterial blood temperatures in the resting human forearm," *J. Appl. Physiol.*, vol. 1, no. 2, pp. 93–122, 1948. https://doi.org/10.1152/jappl.1948.1.2.93

[29] S. Hossain and F. A. Mohammadi, "Tumor parameter estimation considering the body geometry by thermography," *Comput. Biol. Med.*, vol. 76, pp. 80–93, 2016. https://doi.org/10.1016/j.compbiomed.2016.06.023

[30] Y. Zhou and C. Herman, "Optimization of skin cooling by computational modeling for early thermographic detection of breast cancer," *Int. J. Heat Mass Transf.*, vol. 126, pp. 864–876, 2018. https://doi.org/10.1016/j.ijheatmasstransfer.2018.05.129

[31] M. Gautherie, "Thermopathology of breast cancer: Measurement and analysis of in vivo temperature and blood flow," *Ann. N. Y. Acad. Sci.*, vol. 335, no. 1, pp. 383–415, 1980. https://doi.org/10.1111/j.1749-6632.1980.tb50764.x

[32] A. E. Giuliano *et al.*, "Breast cancer-major changes in the American joint committee on cancer eighth edition cancer staging manual," *CA. Cancer J. Clin.*, vol. 67, no. 4, pp. 290–303, 2017. https://doi.org/10.3322/caac.21393

[33] Y. Ye, Y. Li, R. Ouyang, Z. Zhang, Y. Tang, and S. Bai, "Improving machine learning based phase and hardness prediction of high-entropy alloys by using Gaussian noise augmented data," *Comput. Mater. Sci.*, vol. 223, p. 112140, 2023. https://doi.org/10.1016/j.commatsci.2023.112140

[34] J. Liu, L. Zhang, Y. Li, and H. Liu, "Deep residual convolutional neural network based on hybrid attention mechanism for ecological monitoring of marine fishery," *Ecol. Inform.*, vol. 77, p. 102204, 2023. https://doi.org/10.1016/j.ecoinf.2023.102204

[35] H. Wang, F. Lu, X. Tong, X. Gao, L. Wang, and Z. Liao, "A model for detecting safety hazards in key electrical sites based on hybrid attention mechanisms and lightweight Mobilenet," *Energy Reports*, vol. 7, pp. 716–724, 2021. https://doi.org/10.1016/j.egyr.2021.09.200

[36] M. H. Aabidi, A. E. Makrani, B. Jabir, and I. Zaimi, "A model proposal for enhancing leaf disease detection using Convolutional Neural Networks (CNN): Case study," *Int. J. Online Biomed. Eng.*, vol. 19, no. 12, pp. 127–143, 2023. https://doi.org/10.3991/ijoe.v19i12.40329

[37] A. A. A. A. Alrashed, M. S. Gharibdousti, M. Goodarzi, L. R. de Oliveira, M. R. Safaei, and E. P. Bandarra Filho, "Effects on thermophysical properties of carbon based nanofluids: Experimental data, modelling using regression, ANFIS and ANN," *Int. J. Heat Mass Transf.*, vol. 125, pp. 920–932, 2018. https://doi.org/10.1016/j.ijheatmasstransfer.2018.04.142

[38] A. A. A. Figueiredo, H. C. Fernandes, F. C. Malheiros, and G. Guimaraes, "Influence analysis of thermophysical properties on temperature profiles on the breast skin surface," *Int. Commun. Heat Mass Transf.*, vol. 111, p. 104453, 2020. https://doi.org/10.1016/j.icheatmasstransfer.2019.104453

[39] M. A. S. Al Husaini, M. H. Habaebi, F. M. Suliman, M. R. Islam, E. A. A. Elsheikh, and N. A. Muhaisen, "Influence of tissue thermophysical characteristics and situ-cooling on the detection of breast cancer," *Appl. Sci.*, vol. 13, no. 15, p. 8752, 2023. https://doi.org/10.3390/app13158752

8 AUTHORS

Zakaryae Khomsi has a Master degree in Electrical Engineering from ENSAM, Mohammed V University, Rabat, Morocco, in 2018 and is currently a Ph.D. student at ENSAM-ENSIAS with ST2I laboratory, and member of E2SN team, Mohammed V University in Rabat, Morocco (E-mail: zakaryae_khomsi@um5.ac.ma; ORCID: http://orcid.org/0000-0003-2321-9622).

Mohamed El Fezazi has a Master degree in Electrical Engineering from ENSAM, Mohammed V University, Rabat, Morocco, in 2018 and is currently a Ph.D. student at ENSAM-ENSIAS with ST2I laboratory, and member of E2SN team, Mohammed V University in Rabat, Morocco (E-mail: elfezazi.med@gmail.com; ORCID: https://orcid.org/0000-0001-6072-325X).

Larbi Bellarbi is a Professor at the Electrical Engineering department of ENSAM Rabat, and research director of the Electronic Systems, Sensors, and Nanobiotechnology (E2SN) group, Mohamed V University in Rabat, Morocco (E-mail: l.bellarbi@um5r.ac.ma; ORCID: https://orcid.org/0009-0004-5074-3117).