Breast Cancer Detection using Thermal Infrared Image Analysis based on Dempster-Shafer Decision Fusion of CNN Classifiers

Iulia-Ramona Macașoi
Faculty Electronics, Telecommunications and Information Technology
"Politehnica" University of Bucharest
Bucharest, Romania
iuliamacasoi@gmail.com

Victor-Emil Neagoe
Life Senior Member, IEEE
Faculty Electronics, Telecommunications and Information Technology
"Politehnica" University of Bucharest
Bucharest, Romania
victoremil@gmail.com

2023 15th International Conference on Electronics, Computers and Artificial Intelligence (ECAI) | 979-8-3503-2138-8/23/$31.00 ©2023 IEEE | DOI: 10.1109/ECAI58194.2023.10194213

***Abstract*—Thermography is a promising technology for breast cancer detection. We propose a new model to detect breast cancer based on thermography using an ensemble composed by two Convolutional Neural Networks (CNNs). The considered classifier applies Dempster-Shafer decision fusion. The two CNN modules have an identical architecture, but they use an asymmetric training procedure. The ratio between the number of cancer training thermograms and the normal training thermograms corresponding to first CNN module is denoted by β. The corresponding ratio for the second CNN module is chosen to be (1/β). The influence of the asymmetry training parameter β over the decision fusion classifier performances is evaluated. We have obtained the best result concerning overall accuracy (OA) of 98.02%, by choosing the parameter β of 1.2.**

Keywords— breast cancer, thermal imaging, thermogram, convolutional neural networks (CNNs), decision fusion, Dempster-Shafer theory

I. INTRODUCTION

Breast cancer is one of the most challenging diseases in the world, as it is often difficult to be detected early enough for effective treatment. For most cancers, the major cause of the disease has been identified, but for breast cancer it is still unknown. Therefore, early diagnosis is more than necessary to support early treatment [1].

Nowadays, medical imaging is an important tool in medicine. Technological advances have enabled doctors to get a clearer look inside the body and to identify any issues faster and more accurately. This information helps in diagnosing the disease, planning treatment and monitoring treatment [2].

The amount of data from medical imaging is increasing rapidly due to advancements in technology. It can be difficult for doctors to analyze a large amount of data. It is evident that new tools are needed to help in diagnosis and decision-making [2].

There are various types of medical imaging used in breast cancer detection. Thermography is one of the most common methods for breast cancer diagnosis [1]. The ability to detect breast cancer relies on the local temperature differences of the breast areas. Thermal infrared imaging is one of the most promising imaging techniques. It is a key factor in recognizing abnormal temperature variations of the human body. The temperature of the human body is thought to be an important factor in monitoring human health and detecting various diseases [3].

Considering the limitations of human vision, such temperature distribution can be complicated to analyse. As a result, a system that can help to discover changes in body temperature distribution is highly necessary. Deep learning models can be used to interpret thermography. Convolutional neural networks (CNNs) can successfully classify and analyze thermograms [4].

A reliable breast cancer diagnosis requires professional experience and expertise in the field. Without these, misdiagnosis can occur. In fact, many computer-aided diagnostic (CAD) systems are often used to assist doctors in the early detection of breast tumors. The CAD systems have shown advancement in breast cancer diagnosis. These help doctors to recognize abnormalities by using medical imaging techniques and to reduce the mortality rate from 70% to 30%. Deep learning models can efficiently handle the complexity and difficulties of automatic detection of breast cancer [5].

Deep learning techniques are able to detect breast cancer earlier than conventional methods [6]. The combination of artificial intelligence and thermal imaging is a promising solution for the early detection of breast cancer [7].

In our paper, we present the following contributions:

• We develop a deep learning model based on Dempster-Shafer decision fusion [8], which has been used to compute the belief functions and to deduce the final decision. The above theory is applied here for an ensemble of two CNN classifiers. The combination Dempster-Shafer-CNN Ensemble is inspired by the model proposed by Neagoe et al. [9] for COVID-19 detection, but the proposed classifier is used here for breast cancer detection based on thermograms.

• We have chosen the same architecture for the two Convolutional Neural Networks (CNNs) and used an asymmetric training. Thus, we have obtained two CNNs trained for different tasks. The training process is controlled by using the asymmetry parameter β.

• The parameter β has been defined as the ratio between the number of training thermograms of the two classes for a considered CNN. We have assessed performances of our

979-8-3503-2138-8/23/$31.00 ©2023 IEEE

Authorized licensed use limited to: NATIONAL INSTITUTE OF TECHNOLOGY CALICUT. Downloaded on October 11,2023 at 05:38:48 UTC from IEEE Xplore. Restrictions apply.

proposed breast cancer detection classifier across a range of values for the parameter β, specifically in the interval [1, 20].

In this paper we present the application of Dempster-Shafer theory of evidence for breast cancer detection, the architecture of the proposed model, and the influence of asymmetric training. We also present the experimental results and the appropriate conclusions.

II. BACKGROUND RESEARCH

There are several papers focusing on breast cancer diagnosis. Different studies show that early diagnosis of breast cancer is important. In this paper, we will refer to a few of them.

One of the papers that caught our attention was the study of Ucuzal et al. [1]. They have investigated different pre-trained networks, including the VGG-19 model that we used in our work. For all the networks, an accuracy value of over 90% was obtained, while for the VGG-19 model the accuracy was equal to 93,1%. A useful study has been developed by N. Aidossov et al. [10]. They have implemented multiple AI techniques for breast cancer diagnosis based on thermograms. The best result has been obtained with MobileNet, which provided 93.8% accuracy. Farooq et al. [3] have conducted research on thermal breast images, analyzing a classification system for breast tumor detection. Their system obtained an overall accuracy of 80%.

N. Aidossov et al. [11] explored the performance of binary classification of thermograms from a multicenter database without any preprocessing. The results highlight the usefulness of deep learning models for the analysis of thermograms, leading to an accuracy of 80.77%. H. G. Zadeh et al. [12] have analyzed breast thermography images using an autoencoder neural network to classify images as healthy and unhealthy. The accuracy of their proposed system was 94.87%. An amazing study has been developed by P. Kanimozhi et al. [13], where a convolutional neural network model has been analyzed for breast cancer detection. They combined the Spatial pyramid pooling method with a modified U-Net architecture to achieve accurate results. Using this method, an accuracy of 96.13% was achieved.

There are also many other valuable research papers that have focused on the diagnosis of breast cancer using thermogram classification [14-19]. All of these have contributed to the possibility of an early diagnosis and efficient treatment.

III. DEMPSTER-SHAFER DECISION FUSION OF CNN CLASSIFIERS

In this paper, a classification system for breast cancer has been proposed. The purpose of this research is to improve the medical imaging diagnosis. Thermograms obtained from the platform Visual Lab were used to implement this technique. The classification is binary and consists of predicting the class of patients (healthy/sick) based on thermograms. It is necessary to build the model that facilitates this classification. In that case, a model based on Dempster-Shafer decision fusion of CNN classifiers can improve the diagnosis.

A. *Dataset*

The Database for Mastology Research (DMR) - Visual Lab [20] has been used for our research. This is an online platform that stocks and manages mastological images to detect breast cancer at an early stage.

Data from 109 patients have been selected, 51 of them belonging to the sick (cancer) class and 58 belonging to the healthy class. The data set contains 2700 images with the same number of images in either class.

For the CNN models, an image size of 224 × 224 × 3 is required. Therefore, the input thermal images have been resized to the required values.

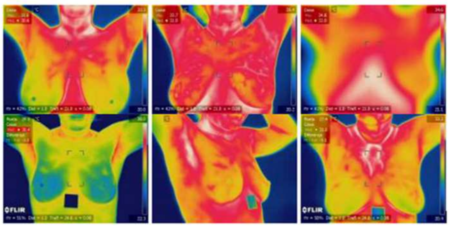

Fig. 1. Examples of images corresponding to the healthy class

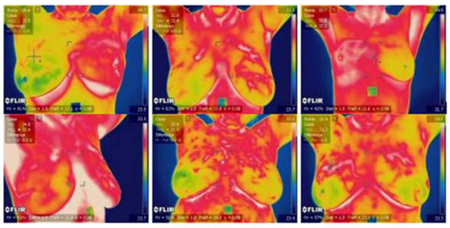

Fig. 2. Examples of images corresponding to the cancer class

B. *CNN architecture*

In this paper, two CNNs of identical architecture have been used to classify breast tumors. More specifically, the architecture used was the pre-trained VGG-19 model.

The dataset contains the same number of images from either class. We have divided the dataset and used a test size equal to 0.3.

By using the parameter β we can control the training process of the two CNNs. The first CNN model has been trained with 945 images corresponding to the cancer class and 945/β images corresponding to the healthy class, and the second CNN model was trained with 945 images corresponding to the sick class and 945/β images corresponding to the healthy class.

By this method, we obtained two CNNs trained for contrary tasks. The first CNN has a better performance in identifying cancer class images, while the second CNN has a better performance in identifying healthy class images. Each model has been trained for 20 epochs.

Authorized licensed use limited to: NATIONAL INSTITUTE OF TECHNOLOGY CALICUT. Downloaded on October 11,2023 at 05:38:48 UTC from IEEE Xplore. Restrictions apply.

C. Dempster-Shafer Decision Fusion

The architecture of the system consists of two CNNs which produce decision probabilities from contrary perspectives. The aim is to integrate the probabilities from the CNN classifiers to obtain a final decision. To achieve this, we have selected the Dempster-Shafer theory of evidence [8], in the variant inspired by the model of Neagoe et al. for COVID-19 detection [9]. The system architecture is shown in Fig. 3.

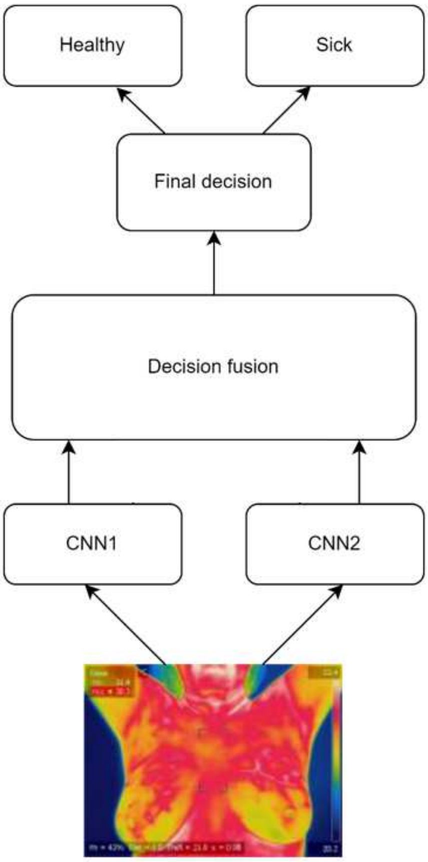

Fig. 3. System architecture

The two classifiers consist of the modules CNN1 and CNN2, with the output NET function of either unit being denoted by $net_p^{(n)}(\beta, z)$, where: p=the class index; n=the CNN index; β=the asymmetry control parameter; z=the test image.

The evidence of belief in each classifier's decision according to [9] is given by relation

$$m_p^{(n)}(\beta, z) = \frac{1}{1+e^{-net_p^{(n)}(\beta,z)}} \quad (1)$$

We deduce $m_p^{(n)}(\beta, z)$ for p=1 and 2, and n=1 and 2. After that, we calculate the belief functions, and make a final decision. The belief functions are given by the following relations deduced in [9]

$$m_1(\beta, z) = \frac{m_1^{(1)}(\beta,z)\cdot m_1^{(2)}(\beta,z)}{m_1^{(1)}(\beta,z)\cdot m_1^{(2)}(\beta,z)+m_2^{(1)}(\beta,z)\cdot m_2^{(2)}(\beta,z)} \quad (2)$$

$$m_2(\beta, z) = \frac{m_2^{(1)}(\beta,z)\cdot m_2^{(2)}(\beta,z)}{m_1^{(1)}(\beta,z)\cdot m_1^{(2)}(\beta,z)+m_2^{(1)}(\beta,z)\cdot m_2^{(2)}(\beta,z)} \quad (3)$$

According to the final decision, the input test image z is attributed to class j, where

$$m_j(\beta, z) = MAX\{m_1(\beta, z), m_2(\beta, z)\}. \quad (4)$$

IV. Experimental Results

The performance indices considered for evaluation of results are the following: Overall Accuracy (OA), Miss Alarm Rate (MAR) and False Alarm Rate (FAR). They can be computed according to relations given in [9] and [14].

A large set of β values have been considered and the corresponding classification performances have been evaluated.

We have further present both the influence of the asymmetry training parameter k on each of the CNN modules but also on the decision fusion classifier.

The experimental results regarding the impact of the parameter β on first CNN classifier performances are shown in Table I.

TABLE I. Performances of the First CNN

β	OA_1	MAR_1	FAR_1
1	96.41	1.97	5.18
1.1	97.28	2.22	3.20
1.2	97.53	0.98	3.95
1.3	97.16	2.22	3.45
1.5	96.79	2.96	3.45
2	96.79	1.48	4.93
2.5	96.04	1.23	6.66
3	93.45	0.74	12.34
5	91.97	0.74	15.30
10	88.51	0.49	22.46
20	75.18	0.24	49.38

The best result for Overall Accuracy (OA) is 97.53%, achieved for a value of the β parameter equal to 1.2. It can be seen that, for a higher value of the asymmetry parameter, the MAR rate decreases, for β=20 having a value of 0.24.

Table II contains the experimental results for the second CNN.

TABLE II. Performances of the Second CNN

β	OA_2	MAR_2	FAR_2
1	96.91	4.93	1.23
1.1	97.28	3.20	2.22
1.2	97.77	2.46	1.97
1.3	96.17	6.41	1.23
1.5	96.66	1.97	4.69
2	96.54	4.44	2.46
2.5	96.04	5.43	2.46
3	96.17	5.92	1.72
5	93.20	12.59	0.98
10	87.03	25.43	0.49
20	80.37	39.01	0.24

Authorized licensed use limited to: NATIONAL INSTITUTE OF TECHNOLOGY CALICUT. Downloaded on October 11,2023 at 05:38:48 UTC from IEEE Xplore. Restrictions apply.

The best result for OA is 97.77% and it was obtained for a value of the asymmetry parameter equal to 1.2. For higher values of the β parameter, the MAR rate increases, while the FAR rate decreases. The best (minimum) value for MAR has been achieved for a value of the β parameter equal to 1.5, while for FAR, minimum value has been obtained for β=20.

Table III shows the results of decision fusion.

TABLE III. PERFORMANCES OF DECISION FUSION

β	OA_DS	MAR_DS	FAR_DS
1	97.65	1.5	3.15
1.1	97.77	1.74	2.68
1.2	98.02	2.67	1.25
1.3	97.40	1.02	4.06
1.5	97.40	3.39	1.75
2	97.16	2.01	3.64
2.5	96.29	2.53	4.81
3	97.28	2.24	3.17
5	95.18	7.44	1.84
10	92.83	3.48	10.29
20	85.92	9.24	17.88

The optimal result for OA has been achieved for a value of the β parameter equal to 1.2. The best (minimum) MAR has been achieved for β=1.3, while the minimum FAR has been reported for β=1.2.

TABLE IV. PERFORMANCES OF OVERALL ACCURACY

β	OA_1	OA_2	OA_DS
1	96.41	96.91	97.65
1.1	97.28	97.28	97.77
1.2	97.53	97.77	98.02
1.3	97.16	96.17	97.40
1.5	96.79	96.66	97.40
2	96.79	96.54	97.16
2.5	96.04	96.04	96.29
3	93.45	96.17	97.28
5	91.97	93.20	95.18
10	88.51	87.03	92.83
20	75.18	80.37	85.92

The Table IV shows the results for Overall Accuracy (OA) for the two CNNs as well as for the decision fusion. For an asymmetry parameter β=1.2, the optimal results have been obtained, with OA_fusion having a value of 98.02%.

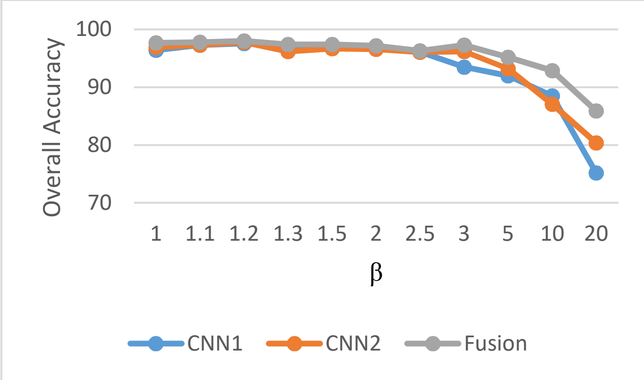

Fig. 4. Accuracy index (OA) as a function of parameter β

In Table IV and Fig. 4 we can remark that the Overall Accuracy of this system has been improved by using the Dempster-Shafer fusion model.

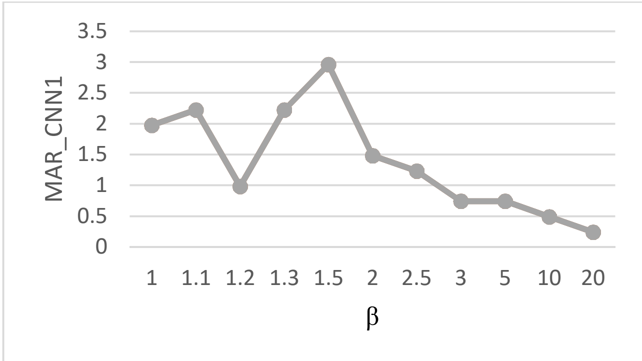

Fig. 5. Miss Alarm Rate of the first CNN

Fig. 5 illustrates that the MAR for CNN1 decreases, when β is increased.

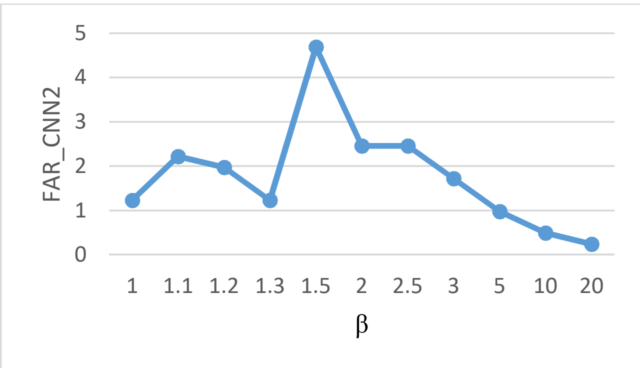

Fig. 6. False Alarm Rate of the second CNN

We can remark in Fig. 6 that the FAR for CNN2 is reduced as β is increased.

Authorized licensed use limited to: NATIONAL INSTITUTE OF TECHNOLOGY CALICUT. Downloaded on October 11,2023 at 05:38:48 UTC from IEEE Xplore. Restrictions apply.

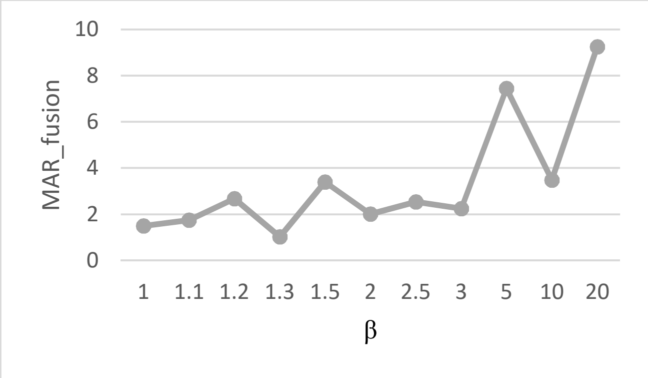

Fig. 7. Miss Alarm Rate after the decision fusion

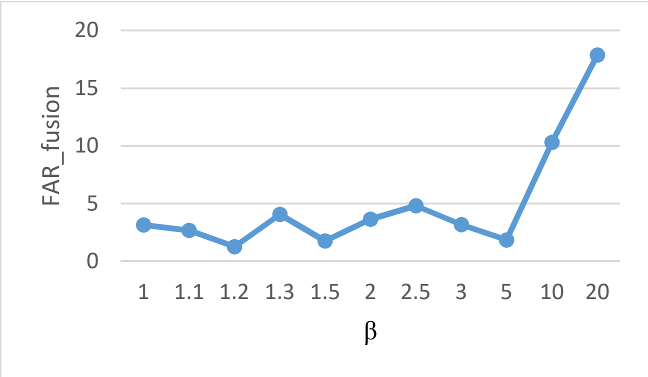

Fig. 8. False Alarm Rate after the decision fusion

One can deduce from Figs. 7 and 8 how the MAR/FAR rates are modified depending on the β parameter.

To evaluate the computational complexity, an analysis of training time has been also performed. In Table V the results obtained for the training time of each of the two CNNs are shown, as well as the value of decision fusion classifier total training time.

TABLE V. THE EVALUATION OF TRAINING TIME

β	Training time_1 [s]	Training time_2 [s]	Total training time [s]
1	353.95	343.01	696.96
1.1	336.78	334.74	671.52
1.2	317.99	317.67	635.66
1.3	307.56	303.54	611.1
1.5	291.69	284.32	576.01
2	269.43	265.99	535.42
2.5	250.61	250.93	501.54
3	234.09	229.36	463.45
5	215.48	212.87	428.35
10	201.87	198.31	400.18
20	189.93	186.61	376.54

The training time decreases by increasing the value of the asymmetry parameter β. This is explained by the fact that the number of images is reduced when the dataset for training of either class is divided by the factor β.

CONCLUDING REMARKS

In the research reported in this article, we have designed, implemented and assessed an interesting method of thermogram classification to detect breast cancer. We have mixed various techniques to obtain a model with high accuracy classification performances: asymmetric controlled training of CNN modules and Dempster-Shafer decision fusion.

For the parameter β equal to 1.2, the Dempster-Shafer decision fusion leads to an Overall Accuracy (OA) of 98.02%. This performance is higher than the accuracy for the same β obtained by any of the two CNN modules.

By choosing an optimum asymmetry parameter β, the performance of a single CNN can also be improved, even in the absence of decision fusion. For example, for β=1 the first CNN obtains an OA of 96.41% and a MAR rate of 1.97%; at the same time, for β=1.2 one obtains an OA of 97.53% and a MAR of 0.98%.

References

[1] H. Ucuzal, M. Baykara, Z. Kucukakcali, "Breast Cancer Diagnosis Based On Thermography Images Using Pre-Trained Networks," The Journal of Cognitive Systems (JCS), vol. 6, no. 2, 2021.

[2] Z. Salahuddin, H. C. Woodruff, A. Chatterjee, P. Lambin, "Transparency of deep neural networks for medical image analysis: A review of interpretability methods," Computers in Biology and Medicine, vol. 140, 2022.

[3] M. A. Farooq, P. Corcoran, "Infrared Imaging for Human Thermography and Breast Tumor Classification using Thermal Images," 2020 31st Irish Signals and Systems Conference (ISSC), Letterkenny, Ireland, 2020, pp. 1-6.

[4] R. Roslidar et al., "A Review on Recent Progress in Thermal Imaging and Deep Learning Approaches for Breast Cancer Detection," IEEE Access, vol. 8, pp. 116176-116194, 2020.

[5] M. F. Mridha et al., "A Comprehensive Survey on Deep-Learning-Based Breast Cancer Diagnosis," , 2021.

[6] M. Nasser, U. K. Yusof, "Deep Learning Based Methods for Breast Cancer Diagnosis: A Systematic Review and Future Direction,", 2023.

[7] M. A. S. Al Husaini, M. H. Habaebi, S. A. Hameed, MD. R. Islam, T. S. Gunawan, "A Systematic Review of Breast Cancer Detection Using Thermography and Neural Networks," IEEE Access, vol. 8, pp. 208922-208937, 2020.

[8] G. Shafer, A Mathematical Theory of Evidence, vol. 42. Princeton, NJ, USA : Princeton Univ. Press, 1976.

[9] V. E. Neagoe, G. L. Ghenea, "An approach of Dempster-Shafer decisionfusion to diagnose COVID-19 in Chest X-ray Imagery by Using Controlled Asymmetric Training of the two CNNs Ensemble," Proc. 14th International Conference on Electronics, Computers, and Artificial Intelligence (ECAI), June 30-July 1, 2022, Ploiești, Romania pp. 1-6.

[10] N. Aidossov et al., "Evaluation of Integrated CNN, Transfer Learning, and BN with Thermography for Breast Cancer Detection," Appl. Sci. 2023.

[11] N. Aidossov et al., "Intelligent Diagnosis of Breast Cancer with Thermograms using Convolutional Neural Networks," In Proceedings of the 14th International Conference on Agents and Artificial Intelligence (ICAART 2022), vol. 2, p. 598-604.

[12] H. G. Zadeh, A. Fayazi, B. Binazir, M. Yargholi, "Breast Cancer Diagnosis Based on Feature Extraction Using Dynamic Models of Thermal Imaging and Deep Autoencoder Neural Networks," Journal of Testing and Evaluation, 2020.

[13] P. Kanimozhi, S. Sathiya, M. Balasubramanian, P. Sivaraj, "Novel Segmentation Method to Diagnose Breast Cancer in Thermography Using Deep Convolutional Neural Network," Annals of R.S.C.B. , vol. 25, 2021.

Authorized licensed use limited to: NATIONAL INSTITUTE OF TECHNOLOGY CALICUT. Downloaded on October 11,2023 at 05:38:48 UTC from IEEE Xplore. Restrictions apply.

[14] M. A. Grigore, V. E. Neagoe, "A Deep CNN approach using thermal imagery for breast cancer diagnosis," Proc. 13th International Conference on Electronics, Computers, and Artificial Intelligence (ECAI), July 1-3, 2021, Romania, pp. 1-4.

[15] R. Roslidar, K. Saddami, F. Arnia, M. Syukri, K. Munadi, "A Study of Fine-Tuning CNN Models Based on Thermal Imaging for Breast Cancer Classification," IEEE International Conference on Cybernetics and Computational Intelligence (IEEE CYBERNETICSCOM), August 22-24, 2019.

[16] K. Rautela, D. Kumar, V. Kumar, "An Interpretable Network to Thermal Images for Breast Cancer Detection," Proc. Of the International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), 16-17 November 2022, Maldives.

[17] L. S. Garia, M. Hariharan, "Ensemble Transfer Learning Approach for Breast Cancer Classification using Thermograms," IEEE, 2021.

[18] A. K. Aggarwal, Alpana, M. Pandey, "Deep Learning based Breast Cancer Classification on Thermogram," International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), 2022.

[19] S. Mishra, A. Prakash, S. K. Roy, "Breast Cancer Detection using Thermal Images and Deep Learning,", IEEE 2020.

[20] "Visual Lab," Banco de Imagens mastológicas, [Online]. Available: http://visual.ic.uff.br/dmi/

Authorized licensed use limited to: NATIONAL INSTITUTE OF TECHNOLOGY CALICUT. Downloaded on October 11,2023 at 05:38:48 UTC from IEEE Xplore. Restrictions apply.