RESEARCH PAPER
Contrib. Sci. & Tech Eng, 2025,2(3)
Homepage: https://cste.journals.umz.ac.ir/
DOI: 10.22080/cste.2025.28717.1012

Breast Mass Detection with a Nonlinear Model of Thermography Images

Faezeh Roshanravan Yazdi [1], Mohammad Mahdi Khalilzadeh [1]*, Faramarz Firouzi [1], Mahdi Azarnoosh [1]

[1] *Department of Biomedical Engineering, Ma.C., Islamic Azad University, Mashhad, Iran*

Article Info
Received 26 February 2025
Accepted 20 May 2025
Available online 20 July 2025

Keywords:
Nonlinear Transform;
CNN;
Temperature Pattern.

Abstract:
Detecting a breast mass is a common and stressful event for women. Although most breast masses are benign, the risk of malignancy highlights the importance of appropriate screening. Different imaging methods have different precisions and accuracies, so choosing an appropriate imaging method, especially for women with dense breast tissue, is very important. Since vascular structure regional temperatures differ between normal and abnormal tissues, thermography can detect masses earlier than conventional imaging methods. 237 cases, including 152 healthy individuals and 85 cases with breast masses examined in this study. The raw recorded images of these cases are gray-level, which are given to a nonlinear transform to become colorful and increase the thermal contrast. Then, these color-scaled images are given to convolutional neural networks. The used networks in this research are AlexNet and GoogLeNet. The extracted features are given to different classifiers as input. The classifiers used in this study are KNN, SVM, and NB. The best result was achieved when GoogLeNet and SVM were used together. The results of this study have a remarkable accuracy and sensitivity, which are 95.8% and 100%, respectively. The developed system, combining nonlinear color scaling and deep learning, shows potential as an effective tool for early breast screening.

***Corresponding Author:** mmkhalilzadeh@iau.ac.ir

Supplementary information: Supplementary information for this article is available at https://cste.journals.umz.ac.ir/

Please cite this paper as: Roshanravan Yazdi, F. , Khalilzadeh, M. M. , Firouzi, F. , & Azarnoosh, M. (2025). Breast mass detection with nonlinear model of thermography images. Contributions of Science and Technology for Engineering, 2(3), 26-36. doi:10.22080/cste.2025.28717.1012.

1. Introduction

Breast cancer is the most well-known type of disease among women all throughout the world [1]. Based on World Health Organization (WHO) reports, by the end of 2020, breast cancer became the most common cancer in the world, with over 7.8 million diagnosed cases in the past 5 years. Also, 685 million global deaths were reported due to breast cancer. Breast cancer might occur in women at any age after puberty, but the probability increases in older age [2] Early diagnosis of breast cancer plays a significant role in curing breast cancer, because it has a tendency to metastasize in the middle and late stages. Therefore, it is vital to diagnose breast cancer in the early stages, which can improve the survival rate of patients [3] and provides fewer loss surgeries (lumpectomy vs. mastectomy). A variety of techniques are utilized for the early diagnosis of breast cancer.

Mammograms have been the gold standard for breast cancer screening since 1960. However, it is affected by parameters like age, breast tissue density, BMI, and family history [4]. This method is based on X-rays that can stimulate some cells. Although mammography is the most commonly used imaging method for breast cancer screening, the false negative rates of this method can reach up to 30% [5]. Additionally, it usually can't detect cancer in younger women due to their breast density [6]. The masking effect on mammography that occurs in dense breast tissue reduces sensitivity. Moreover, mammography is potentially an independent risk factor that can develop breast cancer itself [7].

Sonography or ultrasound imaging is usually used to further investigate suspicious breast regions found during breast examination or mammography. Moreover, this method is used for complementary screening in patients with dense breast tissue. Sonography is a complementary method to mammography since it is widely accessible, moderately inexpensive, and does not discomfort patients. The disadvantage of this technique is its inability to detect breast cancer at an early stage and its higher rate of false-positive results [8].

Magnetic resonance imaging (MRI) of the breast is a test used to detect breast cancer and other abnormalities in the breast. In MRI, a strong magnetic field is used to acquire high-resolution images of the breasts in different locations. To increase image quality, a contrast agent is utilized. MRI has some advantages, like its high sensitivity, high

*ISSN **3060-6578***

specificity, and being non-invasive. The disadvantages of this method are that it is not suitable for patients with Claustrophobia and it high price [3]. Moreover, women are not permitted to breastfeed their children for several hours [9].

However, despite the wide use of mammography, ultrasound, and MRI in breast cancer diagnosis, each of these imaging techniques suffers from specific limitations. Mammography shows reduced sensitivity in women with dense breast tissue and can yield false-negative results. Ultrasound, although non-invasive, is highly dependent on operator experience, and MRI, while providing detailed anatomical information, is both expensive and unsuitable for all patients due to specific conditions such as claustrophobia or breastfeeding limitations. These shortcomings have driven researchers to explore non-invasive, safe, and more accessible complementary imaging techniques, one of which is thermal imaging.

Another method that can be used for breast cancer detection is infrared imaging, which is also known as thermography. Thermography captures thermal radiation emitted from tissue in the range of 2-20 micrometers.

Metabolic activities, blood perfusion, and the temperature of the environment are parameters that affect body temperature. If the metabolic activity of a tissue changes for any reason, like a tumor or any abnormality, the temperature of the skin changes too. This variation in temperature is because of having different biology and a higher metabolic rate [10].

In spite of the fact that the utilization of breast cancer detection via thermography started in the 1960s, it was criticized after the Breast Cancer Detection Demonstration Project (BCDDP) for its low precision. Poor imaging protocol and the absence of experienced thermographers to interpret the collected thermal images lead to low precision [11]. However, due to the progress in resolution of thermal imaging cameras and computer-aided diagnosis (CAD) systems, thermography can now be considered as an alternate imaging method for the early detection of breast cancer [1].

To help the radiologists with breast mass differentiation, various machine learning methods have been proposed. For example, Zavvar et al. performed the breast cancer diagnosis of thermography images using Support Vector Machine (SVM). The inputs of their classifier were pixel-based features, which were extracted from the images. They achieved high accuracy [12]. Abdel Nasser et al. [13] have used dynamic thermograms with fixed time steps to detect breast cancer. They used the LTR and texture analysis methods to model the changes in the breast's temperature. Then they gave this model to the MLP classifier. They differentiate between normal and cancerous cases with AUC=0.989. In [14] The proposed method was used for the classification of normal, benign, and malignant cases according to their thermographic images of the breast. For feature selection, ant colony optimization (ACO) and particle swarm optimization (PSO) were used. Then, the outputs are given to a multi-class SVM with a polynomial kernel function. The results showed that their method had a good performance. In [15], to detect the tumor region, the authors used the active contours technique by employing different energy functions. The accuracy of their method was 91.98%. Recent studies are more interested in using deep networks. For instance, in one study, sparse principal component analysis was used to extract features from a ResNet-50 pre-trained model. At the end, in order to classify the participants, a random forest model was used, which led to the classification of normal and abnormal subjects with an accuracy of 78.16% [16]. In a newer research, a machine learning model based on convolutional neural networks (CNN) was used. The performance of this model was observed in two modes: of simple CNN and adding clinical data decisions to the CNN. The final result reported an accuracy of 93.8% for the model that used clinical data versus 85.4% for the one that didn't [17].

Several recent studies have explored the application of deep learning techniques in breast thermography for cancer detection. For instance, Tamrin et al. [18] employed a hybrid deep learning approach combining CNN and SVM classifiers to achieve high accuracy in classifying breast thermograms. Similarly, Lak and Najafi [19] utilized machine learning and computer vision techniques to analyze thermographic images for breast cancer diagnosis.

The main objectives of this study can be grouped as i) Imaging protocol, ii) Assigning a specific nonlinear transform on raw images, iii) using a pre-trained CNN network as feature extraction, and iv) comparing different methods of classification.

While prior studies have demonstrated the potential of thermography combined with machine learning for breast cancer diagnosis, most rely on raw grayscale images or basic feature extraction methods, which limit the model's sensitivity and specificity. Furthermore, data imbalance and the reliance on manual feature selection often reduce the generalizability of the results. Our approach tackles these issues by enhancing the contrast using a tailored nonlinear transform and using transfer learning from pre-trained CNNs to extract robust features. These steps improve both the interpretability and classification performance, offering a more reliable framework for breast mass detection.

2. Materials and Methods

In this section, the methodology of the breast thermography procedure is explained. The whole methodology is made up of four stages: Breast image data acquisition, proposing a nonlinear transform for color scaling, and feature extraction and classification. The framework is shown in Figure 1. The database, which was used for system development, consists of 152 healthy and 85 cases with mass in their breasts (total = 237). The average age of patients is 39.63. Written informed consent was obtained from the patients. The dataset used in this study was self-collected over a six-month period from patients who visited Chamran Clinic in Mashhad for ultrasound screening. Only those participants who strictly followed the pre-thermography instructions were included in the imaging phase. All procedures were carried out under controlled

conditions based on a standardized thermal imaging protocol. Ethical approval and patient consent details are provided in the "Ethics approval" section at the end of the paper. The authors affirm that human research participants provided informed consent for publication of the images.

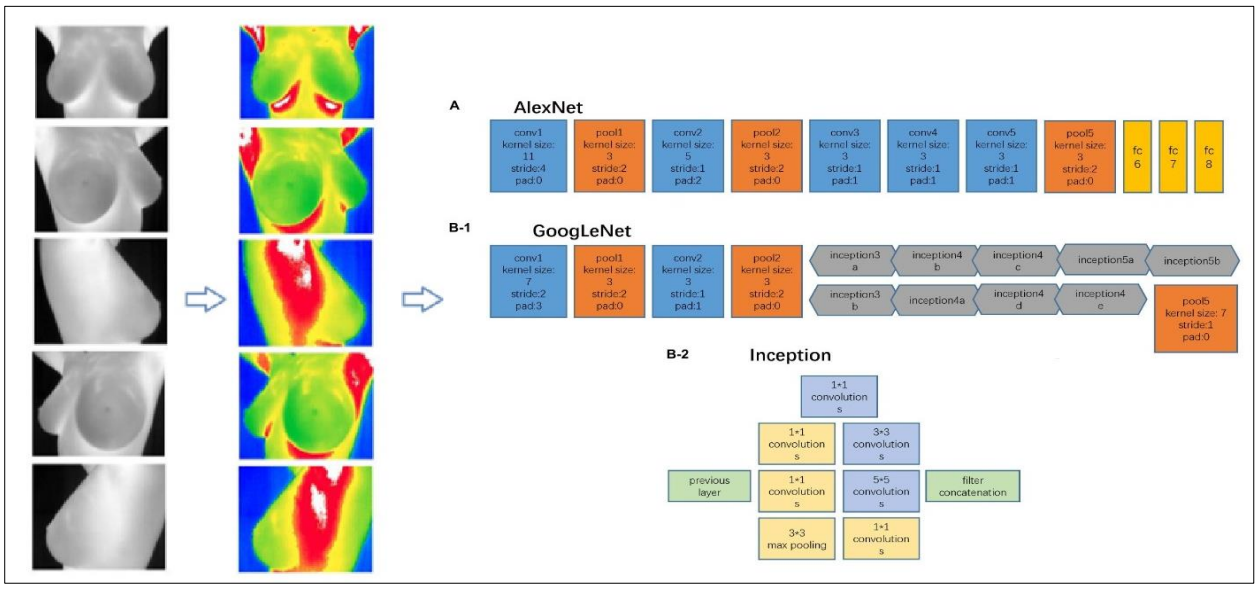


Figure 1. Breast thermography procedure

The original thermography images were captured using a FLIR Lepton 3 infrared camera at a resolution of 120×160 pixels. Since AlexNet and GoogLeNet architectures require input images of 224×224×3, all images were resized to this input dimension using bicubic interpolation before being fed into the networks.

To improve the model's generalization, data augmentation techniques, including horizontal flipping and random translation within a pixel range of [-30, 30] were applied dynamically during training. This approach increased the diversity of the data input while the actual number of samples remained 237.

2.1. Imaging Protocol

Because of human physiology, a controlled environment is a necessity for thermal imaging. Any changes in temperature, clothing, air conditioning and etc., during image recording can cause thermal artifacts. There are some tips that all patients should follow in order to be prepared for imaging. All patients are commanded to desist from sun exposure, stimulation or treatment of the breasts, lotions, antiperspirants, cosmetics, deodorants, bathing, and exercising before the test [20]. It should be noted that the metabolic rate of the body can change depending on the food that the patient consumes before the imaging. This can lead to variations in the observed thermograms. Large meals, cold or hot beverages such as tea, coffee, or iced water, as well as alcohol consumption and smoking, should be avoided a few hours before the examination.

The examination room's temperature and humidity must be controlled (18 and 25 ℃ with a relative humidity between 40% and 75%). The room should also be free from drafts and infrared sources of heat. Shielded windows can prevent external radiation from entering the room [21]. Lastly, the patients must undergo 15 minutes of waist-up nude while their arms have no contact with the body to acclimate to the environment. It should be noted that the wall behind the patient must be covered with a cloth so there will be no reflection, and the examination must be performed in a dark room to prevent any heat reflection from light sources. Images were taken in 5 directions: front, 45°, and 90° rotations to both left and right. For thermal contrast enhancement of areas with abnormal vascularity and metabolic activities, cold stress is suggested. One common cold stress, called the cold challenge, consists of plunging the hands in cold water, which reduces the overall temperature of the breasts [22].

The captured breast thermal images were obtained using an infrared camera FLIR Lepton3 with thermal sensitivity of <50mK (0.05 ℃). The unprocessed images of this camera are in gray-level. Each pixel value of this image corresponds to an amount of temperature. In these images, the lowest value corresponds to black, which indicates the coldest point, and the highest value corresponds to white, which indicates the hottest point. Figure 2 shows an example of a raw image.

Figure 2. **Raw thermal images of the breast in all directions**

2.2. Proposed Nonlinear Transform for Color Scaling

Thermograms are normally in a gray level form of images, but discoveries show that human beings can't interpret gray images well. Human beings can distinguish a number of colors, while they can only recognize a few dozen gray level values. To analyze the images better, a nonlinear transform for color scaling is offered to thermograms. This color scale starts and ends with white, which blue and white represent cold and hot areas, respectively. The transformation function, which is used in this study, is according to Equation 1. In these equations, x is the Normalized gray level, which is between 0 and 1. Then these three channels of R, G, and B are concatenated. Results of applying this nonlinear transform on images are shown in Figures 3, and 4. To further enhance the interpretability of the color-scaled images, the R channel is defined by Equation 1, and the G and B channels are defined similarly. The breakpoints in these equations were determined through an empirical process of iterative experimentation and comparative analysis of the resulting color spectrums. This process aimed to optimize the visual representation of temperature variations in thermography images, ensuring that subtle temperature differences are clearly discernible in the color-scaled thermograms

.

$$
R(x)=\begin{cases}0 & 0\le x\le \frac{1}{2}\\ 4x-2 & \frac{1}{2}<x<\frac{4}{5}\\ 1 & \frac{4}{5}\le x\le 1\end{cases}
$$

$$
G(x)=\begin{cases}5x & 0\le x<\frac{1}{5}\\ -\frac{20}{9}x+\frac{13}{9} & \frac{1}{5}\le x<\frac{7}{20}\\ \frac{20}{27}x+\frac{2}{5} & \frac{7}{20}\le x<\frac{3}{4}\\ -\frac{15}{2}x+7 & \frac{3}{4}\le x<\frac{13}{14}\\ 60x-56 & \frac{13}{14}\le x<\frac{18}{19}\\ 1 & \frac{18}{19}\le x\le 1\end{cases}
$$

$$
B(x)=\begin{cases}1 & 0\le x\le \frac{1}{5}\\ -4x+\frac{9}{5} & \frac{1}{5}<x<\frac{2}{5}\\ 1/5 & \frac{2}{5}\le x<\frac{13}{14}\\ 60x-56 & \frac{13}{14}\le x<\frac{18}{19}\\ 1 & \frac{18}{19}\le x\le 1\end{cases}
$$

(1)

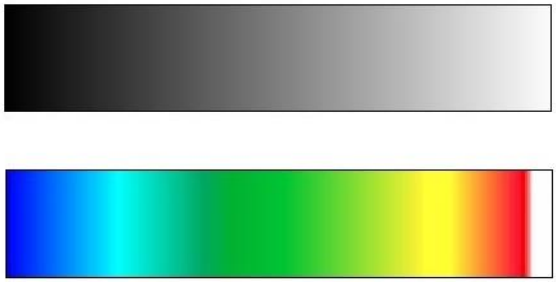


Figure 3. **Spectrum in raw thermal images vs color scaled images. White color shows the maximum temperature**

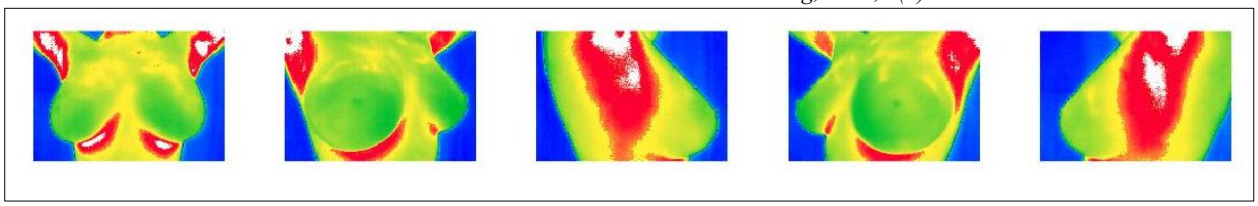


Figure 4. Proposing a nonlinear transform for color scaling in Figure 2

Furthermore, it is worth noting that the proposed nonlinear transform, while effective for enhancing thermal contrast, may introduce mathematical discontinuities. However, these discontinuities are not visually perceptible in the final color-scaled thermograms due to factors such as clipping of negative values, the limited input range (0-1), and the visual insensitivity to small color variations. Moreover, the spatial averaging performed by the convolutional neural networks and the emphasis on overall thermal patterns contribute to mitigating the impact of these discontinuities, ensuring the generation of visually coherent and informative images for breast mass detection.

2.3. Feature Extraction

Convolutional Neural Network (CNN) is a specialized neural network for processing data, which is mostly used in image processing and recognition. A CNN acts very similar to a multilayer perceptron that has been designed to reduce processing requirements. The layers of a CNN consist of an input layer, hidden layers, and an output layer. Hidden layers include multiple convolutional layers, pooling layers, normalization layers, and fully connected layers. Because of its effective models, deep learning extracts features better than conventional machine learning. To have an excellent performance, CNNs need a large amount of data for training. To resolve this problem and to achieve efficient models of classification, employing pre-trained deep learning models is suggested. These models have been trained with a large amount of data and have shown noticeable performance in image classification [23]. Moreover, the CNN network has simpler and quicker fine-tuning than a network with randomly initialized weights from scratch [24-26]. Rather than training and testing the CNN with medical images, transfer learning is suggested. In transfer learning, the pre-trained CNN that is trained with millions of non-medical images is used for medical images [27].

In the proposed method, AlexNet [28] and GoogLeNet [29] models were used to extract features. The comparison of these networks is shown in Table 1.

Table 1. Comparison of AlexNet and GoogLeNet

	Number of layers	Number of convolutional layers	Number of full connection layers	Inception model	Local response normalization	Top 5 error rate	Year of publication
AlexNet	8	5	3	×	✓	15.3%	2012
GoogLeNet	22	21	1	✓	✓	5.5%	2014

AlexNet: Over the other non-deep learning methods, AlexNet achieved significantly improved performance for the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2012 [30]. ImageNet is a set of data consisting of 1.2 million images with 1000 different categories. AlexNet has eight layers: the first five are convolutional layers and the last three are fully connected layers.

GoogLeNet: The GoogLeNet model is remarkably more complicated and deeper than AlexNet. Significantly, it also includes “Inception” as a new module. In this module, filters of different sizes and dimensions are connected into a single new filter. The network with the inception architecture is faster than the network with a non-inception architecture [31]. This model includes two convolutional layers, two pooling layers, and nine “Inception” modules, which are comprised of six convolution layers and one pooling layer.

To lessen the reliance on training data and to have a more accurate model, the data was augmented by rotating and mirroring images. The output of these two networks is separately applied to the classifiers. For model training and evaluation, the dataset was split into 70% for training, 15% for validation, and 15% for testing. The partitioning was performed randomly, ensuring a balanced distribution between healthy and mass cases.

2.4. Classification

In the classification stage, the performance of the following classifiers has been studied: K Nearest Neighbor (KNN) [32], Support Vector Machine (SVM) [33], and Naive Bayesian (NB) [34]. To ensure the statistical reliability of the classification results and minimize the risk of overfitting, we performed 4-fold cross-validation on the extracted features during the classifier training process. In each fold, the data was split into training and validation subsets, ensuring balanced representation from both classes. Additionally, it should be noted that the CNN models used for feature extraction (AlexNet and GoogLeNet) were pre-trained on ImageNet and were not fine-tuned on our thermography dataset, so they did not see any of the training or testing samples during this process.

KNN is one of the easiest and applicable algorithms in machine learning. Its basic principle is based on geometric measurement, and for measuring the distances between the data points, it often uses the Euclidean distance metric. In

our implementation, we used k = 5 for kNN with correlation as the distance metric instead of the default Euclidean distance. Goals like regression and classification are achievable by measuring the distance between different feature values.

$$D(i.j) = \sqrt{\sum_{t=1}^{m}(x_{it} - x_{jt})^2} \quad (2)$$

SVM is another algorithm of supervised learning with a special position in engineering practice [35]. In this study, we used the radial basis function (RBF) as the kernel function for SVM, with a BoxConstraint value of 10 and KernelScale set to 'auto'. For a two-class problem, the fundamental idea of this algorithm is to find the best hyperplane Equation 3. where; w is a vector normal to the hyperplane, x is training data, b is bias, and f(x) represents the discriminant function, where the sign of *f(x)* determines the class assignment, respectively.

$$f(x) = (w \times x) + b(x \in R^m) \quad (3)$$

In commonly used forms of SVM, the variable *m* represents the number of support vectors. This method is also able to solve multi-class problems. If the data isn't linearly separable, a mapping function, called a kernel, is used to determine the hyperplane. So, the hyperplane can be formulated according to Equation 4. In Equation 5, f (x_d), x_j, and x_d represent the class of new input data, a support vector, and input data, respectively.

$$min\frac{1}{2}\|w\|^2 = min\frac{1}{2}\|w\|^T w \quad (4)$$

$$f(x_d) = \sum_{j=1}^{Ns} a_j y_j x_j x_d + b \quad (5)$$

NB is also easy and fast to predict the classes of data sets. This algorithm is based on Bayes Theorem which is shown in Equation 6. In this equation; $P(X|Y)$, $P(Y)$, $P(X)$, and $P(Y|X)$ represents the conditional probability of the class for the given attribute, probability of the class, probability of the attribute, and the conditional probability that attribute belongs to the class, respectively.

$$P(Y|X) = \frac{P(Y)P(X|Y)}{P(X)} \quad (6)$$

3. Results

For better analysis of the images, a nonlinear transform is proposed to color scale images. The effect of this transform can be seen in Figure 6. where a, b, c, and d are gray-level image of a healthy case, gray-level image of a case with mass in breast, color scaled image of a healthy case, and color scaled image of a case with mass in breast, respectively. What seems clear is that not all the information can be extracted from gray-level images. In both gray-level images, there are warmer areas than the rest of the tissue, but the details are more obvious in colorful images. Heat distribution of hot spots can be a symptom of a disorder. In addition to high temperature, asymmetry is an important parameter in diagnosis because breast structure, temperature, and density are usually symmetrical. In this study, temperature and symmetry are considered as two important detection criteria alongside each other. It is because of the fact that when images are approximately symmetrical, small asymmetries can show a suspicious region.

It should be noted that the higher temperature of the areas under the breast and underarms is due to the higher probability of sweating in those areas and has no diagnostic value. As also shown in Figure 5, the hot spots in (a) are approximately symmetrical, whereas in (b) there is temperature asymmetry in the left and right breast.

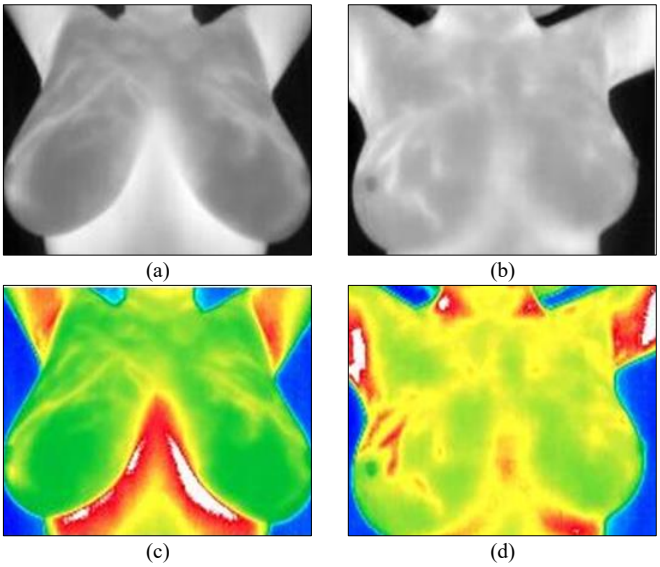


Figure 5. Comparing front images of: (a) gray-level image of a healthy case. (b) gray-level image of a case with a mass in the breast. (c) Color-scaled image of a healthy case. (d) color-scaled image of a case with mass in the breast

In this study, 237 participants, including 152 healthy individuals and 85 cases with a mass in their breasts, took part. The input size for CNN networks for feature extraction is 120*160. Comparing AlexNet and GoogLeNet, as expected, GoogLeNet, with a more complex architecture, is

a better method for feature extraction using the same classifier.

In the following, the simulation results have been briefly described. The classifiers that are used in the comparison of results are K-nearest neighbor (KNN), Support Vector Machine (SVM), Naive Bayesian (NB). The classification performance is evaluated using random train/test splits. This can get persuasive results by eliminating the effect of manually dividing of train and test sets. It should be mentioned that the kernel function, which is used in the SVM, is a radial basis function (rbf).

To assess the classification performance of the model, the accuracy, sensitivity, specificity, positive predictive value(PPV), and negative predictive value (NPV) are calculated.

The train/ test accuracy obtained for each transfer learning technique presented is shown in Table 2. As seen, among all classifiers, SVM has the best accuracy in both the training and testing phases of AlexNet and GoogLeNet. The performance of each classifier is illustrated in Table 3. It should be noted that the performance values shown in Table 3 were averaged over repeated runs and include both training and testing results. In contrast, the confusion matrix presented in Figure 6 corresponds to a single run on the test dataset using GoogLeNet and SVM. Slight differences may occur between the two due to randomness in data splits and classifier initialization.

Table 2. Comparison of training and testing accuracies for different classifiers

	Feature Extraction Method: AlexNet		Feature Extraction Method: GoogLeNet	
	train data	test data	train data	test data
KNN	92.1%	87.5%	93.9%	91.7%
SVM	99.4%	90.3%	98.8%	95.8%
NB	93.9%	90.3%	95.2%	91.7%

Table 3. The performance of the applied classifiers

	Accuracy	Sensitivity	Specificity	PPV	NPV
AlexNet + KNN	87.5%	97.8%	69.2%	84.9%	94.7%
AlexNet + SVM	90.3%	97.8%	76.9%	88.2%	95.2%
AlexNet + NB	90.3%	97.8%	76.9%	88.2%	95.2%
GoogLeNet + KNN	91.7%	95.7%	84.6%	91.7%	91.7%
GoogLeNet + SVM	95.8%	100%	88.5%	93.9%	100%
GoogLeNet + NB	91.7%	95.7%	84.6%	91.7%	91.7%

The best result in this section is reached when the classifier is SVM and the feature extraction method is GoogLeNet (train error=1.2% and test error=4.2%). It is obvious that, due to the fact that the initial weights of the network are chosen randomly, the declared results are not definite. The reported result is the average of repeated network executions.

To assess the details of classification of best performance (GoogLeNet+SVM), the confusion matrix of the corresponding classification result is demonstrated in Figure 6. where class1 are healthy cases and class2 are cases with mass in their breasts. In this case, the accuracy of classification of train data is 98.8% and test data is 95.8%. The receiver operating characteristic curve (ROC) is calculated and shown in Figure 7. as well. An ROC curve is a graph that plots True Positive Rate and False Positive Rate which shows the performance of the classification model. Any ROC curve is actually a step function generated from a limited set of samples, that approaches a true curve as the number of samples increases and comes close to infinity. In this study, the ROC curves were computed using the perfcurve function in MATLAB, based on the class scores output by the SVM classifier. Although our task is binary, we plotted a curve for each class using a one-vs-rest strategy. This allows each class to be considered as the "positive class" separately. We report both curves for completeness, but the AUC of the "mass" class (Class 2) is used for primary performance evaluation.

4. Discussion

In the presented study, a fully automatic computer-aided breast screening and classification system based on a nonlinear transform is developed. The developed system provides precious information to the specialists. It should be mentioned that our fully automatic system is based on the analysis of thermographic pictures. It is necessary to know that temperature is not the only important parameter for analyzing images, but the asymmetrical hot patterns also play a role. If there is a higher temperature on a breast forming a pattern, it doesn't necessarily confirm the presence of a mass or tumor. The symmetrical points in the other breast should also be checked. If the hot spots were symmetrical in both breasts, there would not be a problem. If there were hot spots in a breast with no symmetrical sign on the other breast, that can confirm the presence of a mass or tumor. To have a better vision of the temperature pattern, this nonlinear transform was performed on the images.

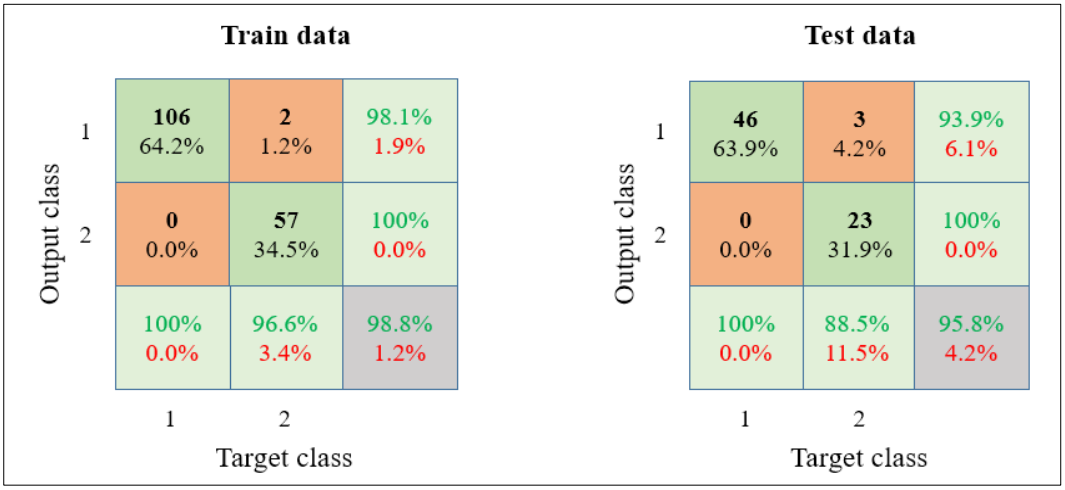


Figure 6. **Confusion matrix of the best result where the feature extraction method is GoogLeNet and the classifier is SVM**

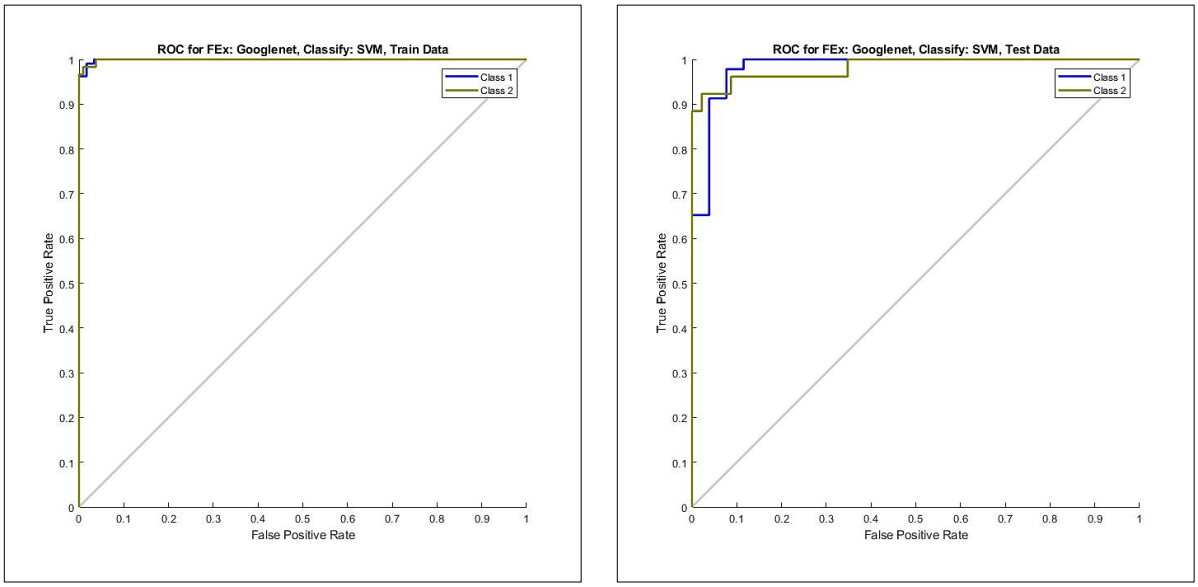


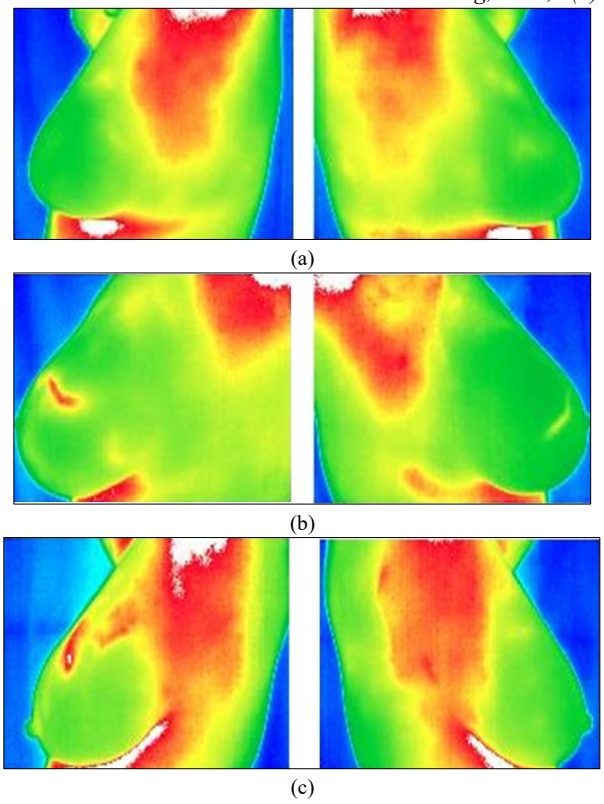
Figure 7. **ROC curves of the best result, where the feature extraction method is GoogLeNet and the classifier is SVM**

As seen in Figure 8-a, the whole breast is completely green with no hot spots. Therefore, this case is considered a healthy person. Needless to say, hot spots (red and white colors) in locations such as the neck, armpits, and under breasts are due to perspiration and less susceptibility to environmental temperature, which is considered to be normal. This usually happens in most subjects. The hot points in Figure 8-b attract attention. As mentioned before, although heat is a necessary sign to detect suspicious regions, it is not enough, and point symmetry must also be considered. Regarding the image, it is seen that the hot region observed in the upper-outer quarter of the left breast is symmetrical to the upper-outer quarter position of the right breast. Therefore, since the hot points are symmetrical, the case is also assumed to be a normal case with no problem. The importance of symmetry can be seen in the decision-making of these kinds of images. Moreover, hot spots in the breasts can sometimes be associated with age. Hot points with symmetry mostly happen in younger women. One important factor to remember is that the lesion distance to the skin strongly affects temperature. Finally, according to the images of Figure 8-c, in the upper-outer quadrant of the left breast, close to the auxiliary 2 o'clock, a hot region is seen which doesn't have a similar correspondence in the right breast. Since no symmetry was found for it in the right breast, this case is assumed to have a mass or tumor in her breast. It should be emphasized that all of the above observations are approved by the ultrasound reports.

Ultrasound and thermography reports of Figure 8-c case were compared, and the suspicious lesion observed in the left breast of thermography was confirmed by the radiologist with BIRADS 4a.

Compared to previous works in the literature [14, 17, 18], the proposed method shows competitive or superior performance in terms of classification accuracy and AUC.

Table 4 presents a comparative summary of classification performance metrics from several recent studies alongside the results of our proposed method. As shown, our model achieves higher or competitive performance across key metrics such as accuracy and AUC.



(a)

(b)

(c)

Figure 8. (a) a healthy person without any hot spot and no mass or tumor. (b) a healthy person with symmetrical hot spots in the breasts and no mass or tumor. (c) Abnormal case with asymmetrical hot spots with mass or tumor

Table 4. Comparison of classification performance between the proposed method and recent studies on breast thermography

Study	Imaging Modality	Classifier	Accuracy (%)	AUC
Ahmed et al. (2019) [14]	Thermography	SVM + ACO-PSO	93.2	0.90
Mammoottil et al. (2022) [17]	Thermography	CNN	93.8	0.94
Rezazadeh et al. (2023) [18]	Thermography	CNN-SVM hybrid	94.6	0.95
Our study	Thermography	Pre-trained CNN + SVM	**95.8**	**0.996**

Besides the better vision of temperature pattern, which was obtained from the nonlinear transform, it can be seen that AlexNet and GoogLeNet do well even in the classification of the medical domain, which has a limited amount of data. What makes a difference in the performance of these networks is the architecture. The Inception Modules in GoogLeNet filter different sizes of convolutions and concatenate the filters for the next layer. Instead, layers input in AlexNet, processed by the former layer rather than a filter concatenation. As the depth of AlexNet is less when it attempts to learn features from image sets, it requires more time to attain better results compared to GoogLeNet.

Analyzing the above results shows that GoogLeNet with SVM as a classifier has higher accuracy compared to other methods. It is shown that the specificity and sensitivity of this method are the highest values as well. As expected, the ROC curve also reached infinity with fewer steps, which means a lower error rate in identifying the classes.

5. Conclusion

In this work, we demonstrated the use of a deep learning framework in the classification of healthy cases and cases with mass in their breasts from thermal images. Thermal data can be used in two forms of gray-level images and in color-scaled scaled. The color-scaled images are obtained from proposing a nonlinear transform on raw images, which are the input of the CNN network. The aim of using a CNN network is transfer learning, which is used for feature extraction. Then, the extracted features are applied to different classifiers. Then, the output becomes the input for classifiers, which are SVM, KNN, and NB. The best accuracy was assessed using GoogLeNet for feature extraction and an SVM classifier with 95.8%. It should be mentioned that these results are for color-scaled images. The accuracy of raw images with the same method was 83%, and that's why the nonlinear transform of color scaling is performed. Thus, the proposed system could be used as an effective screening system for early breast screening.

For future work, as the feature extraction and classification are two separate tasks, further study should focus on linking these two tasks in order to reduce computational complexity. Another upgrade for future work can concentrate on improving the CNN model, which can lead to a reduction in the time spent on computation and training.

6. Statements & Declarations

6.1. Ethics statement

The protocol adopted in this research was matched with the Declaration of Helsinki, approved by the committee of Islamic Azad University of Mashhad's Department of Biomedical Engineering. In addition, the research purpose was clearly explained to all participants, and they were asked to provide their written informed consent.

6.2. Funding statement

No funding was received for conducting this study.

6.3. Declaration of interests

The authors have no relevant financial or non-financial interests to disclose.

7. References

[1] Kakileti, S. T., Dalmia, A., & Manjunath, G. (2020). Exploring deep learning networks for tumour segmentation in infrared images. Quantitative InfraRed Thermography Journal, 17(3), 153–168. doi:10.1080/17686733.2019.1619355.

[2] World Health Organization (WHO). (2025). Cancer. World Health Organization(WHO), Geneva, Switzerland. Available online: https://www.who.int/news-room/fact-sheets/detail/cancer (accessed on July 2025).

[3] He, Z., Chen, Z., Tan, M., Elingarami, S., Liu, Y., Li, T., Deng, Y., He, N., Li, S., Fu, J., & Li, W. (2020). A review on methods for diagnosis of breast cancer cells and tissues. Cell Proliferation, 53(7). doi:10.1111/cpr.12822.

[4] Liu, L., Hao, X., Song, Z., Zhi, X., Zhang, S., & Zhang, J. (2021). Correlation between family history and characteristics of breast cancer. Scientific reports, 11(1), 6360. doi:10.1038/s41598-021-85899-8.

[5] Berg, W. A., Blume, J. D., Cormack, J. B., Mendelson, E. B., Lehrer, D., Böhm-Vélez, M., Pisano, E. D., Jong, R. A., Evans, W. P., Morton, M. J., Mahoney, M. C., Larsen, L. H., Barr, R. G., Farria, D. M., Marques, H. S., & Boparai, K. (2008). Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. Jama, 299(18), 2151–2163. doi:10.1001/jama.299.18.2151.

[6] Yao, X., Wei, W., Li, J., Wang, L., Xu, Z. L., Wan, Y., Li, K., & Sun, S. (2014). A comparison of mammography, ultrasonography, and far-infrared thermography with pathological results in screening and early diagnosis of breast cancer. Asian Biomedicine, 8(1), 11–19. doi:10.5372/1905-7415.0801.257.

[7] Jarvis, H., Mi, L., Patel, B., Cube, R. P., Pruthi, S., & Vegunta, S. (2024). Factors Influencing Patient Confidence in Screening Mammography. Journal of the American Board of Family Medicine, 36(6), 942–951. doi:10.3122/jabfm.2023.230055R1.

[8] Prasad, S. N., & Houserkova, D. (2007). The role of various modalities in breast imaging. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub, 151(2), 209–218. doi:10.5507/bp.2007.036.

[9] Husaini, M. A. S. Al, Habaebi, M. H., Hameed, S. A., Islam, M. R., & Gunawan, T. S. (2020). A Systematic Review of Breast Cancer Detection Using Thermography and Neural Networks. IEEE Access, 8, 208922–208937. doi:10.1109/ACCESS.2020.3038817.

[10] Navid, M., Hamidpour, S. S. F., Khajeh-Khalili, F., & Alidoosti, M. (2020). A novel method to infrared thermal images vessel extraction based on fractal dimension. Infrared Physics & Technology, 107. doi:10.1016/j.infrared.2020.103297.

[11] Kakileti, S. T., Manjunath, G., Madhu, H., & Ramprakash, H. V. (2017). Advances in Breast Thermography. New Perspectives in Breast Imaging, InTechOpen, London, United Kingdom. doi:10.5772/intechopen.69198.

[12] Zavvar, T., Ridhawi, I. Al, & Abbas, A. (2023). Thermography-Based Early-Stage Breast Cancer Detection Using SVM. 2023 International Conference on Intelligent Data Science Technologies and Applications, IDSTA 2023, 27–34. doi:10.1109/IDSTA58916.2023.10317862.

[13] Abdel-Nasser, M., Moreno, A., & Puig, D. (2019). Breast cancer detection in thermal infrared images using representation learning and texture analysis methods. Electronics (Switzerland), 8(1). doi:10.3390/electronics8010100.

[14] Ahmed, A. A., Ali, M. A. S., & Selim, M. (2019). Bio-inspired based techniques for thermogram breast cancer classification. International Journal of Intelligent Engineering and Systems, 12(2), 114–124. doi:10.22266/IJIES2019.0430.12.

[15] Zadeh, H. G., Haddadnia, J., Seryasat, O. R., & Isfahani, S. M. M. (2016). Segmenting breast cancerous regions in thermal images using fuzzy active contours. EXCLI Journal, 15(532), 532–550. doi:10.17179/excli2016-273.

[16] Yousefi, B., Akbari, H., & Maldague, X. P. V. (2020). Detecting Vasodilation as Potential Diagnostic Biomarker in Breast Cancer Using Deep Learning-Driven Thermomics. Biosensors, 10(11). doi:10.3390/BIOS10110164.

[17] Mammoottil, M. J., Kulangara, L. J., Cherian, A. S., Mohandas, P., Hasikin, K., & Mahmud, M. (2022). Detection of Breast Cancer from Five-View Thermal Images Using Convolutional Neural Networks. Journal of Healthcare Engineering, 2022, 15. doi:10.1155/2022/4295221.

[18] Tamrin, H. R., Saniei, E., & Barough, M. S. (2023). Intelligent diagnosis of breast cancer through hybrid deep networks using thermography images. Journal of Health Informatics and Biomedical Engineering, 10(3), 260-268. doi:10.34172/jhbmi.2023.30

[19] Lak, B., & Najafi, P. (2021). Breast cancer diagnosis using a combination of machine learning and machine vision methods in thermography images. Advanced Signal Processing, 5(1), 41-52. doi:20.1001.1.26763397.1400.5.1.5.0

[20] Mohamed, N. A. E.-R. (2015). Breast Cancer Risk Detection Using Digital Infrared Thermal Images. International Journal of Bioinformatics and Biomedical Engineering, 1(2), 185–194.

[21] Kandlikar, S. G., Perez-Raya, I., Raghupathi, P. A., Gonzalez-Hernandez, J. L., Dabydeen, D., Medeiros, L., & Phatak, P. (2017). Infrared imaging technology for breast cancer detection – Current status, protocols and new directions. International Journal of Heat and Mass Transfer, 108, 2303–2320. doi:10.1016/j.ijheatmasstransfer.2017.01.086.

[22] Gonzalez-Hernandez, J. L., Recinella, A. N., Kandlikar, S. G., Dabydeen, D., Medeiros, L., & Phatak, P. (2019). Technology, application and potential of dynamic breast thermography for the detection of breast cancer. International Journal of Heat and Mass Transfer, 131, 558–573. doi:10.1016/j.ijheatmasstransfer.2018.11.089.

[23] Al-Sabaawi, A., Hasan, R. I., Fadhel, M. A., Al-Shamma, O., & Alzubaidi, L. (2021). Employment of Pre-trained Deep Learning Models for Date Classification: A Comparative Study. International Conference on Intelligent Systems Design and Applications, 1351, 181–189. doi:10.1007/978-3-030-71187-0_17.

[24] Pratap, T., & Kokil, P. (2019). Computer-aided diagnosis of cataract using deep transfer learning. Biomedical Signal Processing and Control, 53(101533). doi:10.1016/j.bspc.2019.04.010.

[25] Huynh, B. Q., Li, H., & Giger, M. L. (2016). Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. Journal of Medical Imaging, 3(3), 034501. doi:10.1117/1.jmi.3.3.034501.

[26] Elaraby, N., Barakat, S., & Rezk, A. (2023). A Novel Siamese Network for Few/Zero-Shot Handwritten Character Recognition Tasks. Computers, Materials and Continua, 74(1), 1837–1854. doi:10.32604/cmc.2023.032288.

[27] Angriawan, M. (2023). Transfer Learning Strategies for Fine-Tuning Pretrained Convolutional Neural Networks in Medical Imaging. Research Journal of Computer Systems and Engineering, 4(2), 73–88. doi:10.52710/rjcse.79.

[28] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60(6), 84–90. doi:10.1145/3065386.

[29] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition. 7-12 June, 2015, Boston, United States.

[30] Yang, Y., Yan, L. F., Zhang, X., Han, Y., Nan, H. Y., Hu, Y. C., Hu, B., Yan, S. L., Zhang, J., Cheng, D. L., Ge, X. W., Cui, G. Bin, Zhao, D., & Wang, W. (2018). Glioma grading on conventional MR images: A deep learning study with transfer learning. Frontiers in Neuroscience, 12, 804. doi:10.3389/fnins.2018.00804.

[31] Chelghoum, R., Ikhlef, A., Hameurlaine, A., Jacquir, S. (2020). Transfer Learning Using Convolutional Neural Network Architectures for Brain Tumor Classification from MRI Images. Artificial Intelligence Applications and Innovations. AIAI 2020. IFIP Advances in Information and Communication Technology, vol 583. Springer, Cham, Switzerland. doi:10.1007/978-3-030-49161-1_17..

[32] Cover, T. M., & Hart, P. E. (1967). Nearest Neighbor Pattern Classification. IEEE Transactions on Information Theory, 13(1), 21–27. doi:10.1109/TIT.1967.1053964.

[33] Cortes, C. and Vapnik, V. (1995) Support-Vector Networks. Machine Learning, 20, 273-297. doi:10.1007/BF00994018.

[34] Russell, S., & Norvig, P. (2020). Artificial Intelligence: A Modern Approach: A Modern Approach. Pearson, London, United Kingdom.

[35] Liu, W., Wei, J., & Meng, Q. (2020). Comparisions on KNN, SVM, BP and the CNN for Handwritten Digit Recognition. 2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications(AEECA), 587–590. doi:10.1109/aeeca49918.2020.9213482.