Journal of Information Hiding and Multimedia Signal Processing ©2025 ISSN 2073-4212
Ubiquitous International Volume **16**, Number **1**, March **2025**

Image Classification using Decision Tree Classifier and Features Extraction using Hough transform and Genetic Algorithm

Hend A. Elsayed *
Electrical Engineering Department,
Faculty of Engineering, Damanhour
University, Damanhour, Egypt
hend.alielsayed@yahoo.com; Hend.Elsayed@eng.dmu.edu.eg

Mustafa F. Hameed
Department of Information Technology,
Institute of Graduate Studies and Researches,
Alexandria University, Alexandria, Egypt
mustafa.fehmy@gmail.com

Mohammed M. El Sherbiny
Department of Information Technology,
Institute of Graduate Studies and Researches,
Alexandria University, Alexandria, Egypt
mmmsherbiny@gmail.com

*Corresponding author: Hend A. Elsayed
Received August 11, 2024, revised October 17, 2024, accepted November 22, 2024.

Abstract. *Image classification for machines has proven to be a difficult task in the past even with the rapid development in computer and information technology, also has proven to be a needed task with the rapid increase of the amounts of data which has led the task of image classification to be a needed field of development. Digital images are used in nearly every field. With the increased use of medical digital images, it has become more and more necessary to develop more rapid ways for analysis and classification, with increase of technology and modern computing power it is now possible to improve the quality of output results so it can effectively be used to lift the burden that faces every person in the field of computer technology. This work presents three classification techniques for digital mammography medical images. The three techniques were Classification via Hough transform, feature peaks and decision tree combination, Classification via Hough transform, genetic algorithm and decision tree combination. and Classification via Hough transform, Feature peaks, genetic algorithm and Decision tree combination. The final technique has proven to be the best in terms with the least number of errors, the highest accuracy and the least processing time.*
Keywords: Image classification; features extraction; hough transform; genetic algorithm; decision tree classifier

1. Introduction.

An approach to classification that is based on the contextual details of the images is called contextual image classification. The idea of this classification focuses on the neighbouring pixels. Classification comes with a degree of certainty as is the case with most data mining solutions. The classification of objects is often based on previous examples of other similar object with a reference to the original model of the data inquired.

Classification is a classical data mining technique based-on machine learning. It is used to classify items into predefined sets or groups. There are several methods for classification used such as decision tree and linear programming. There are many feature extraction techniques used in digital image classification such as Hough transform that can be used to find objects “like circles” in an image [1]. Another technique is feature extraction, which is a way to take small amounts of semantically significant information out of images. [2]. Another excellent technique for reducing the dimensions of the data used in data mining and machine learning is non-negative matrix factorization, or NMF [3]. Divide-and-Conquer (DC) is another approach based on feature space decomposition for classification. When large datasets are available, this approach manages to separate between the classes at will [4]. In our approach, we use Hough Transform (HT) Decision Trees, Genetic Algorithm (GA) and Feature Extraction. Decision trees are tree-structured classifiers with the ability to predict. They create criterion of decisions where the intermediate prediction is available in samples [5]. The implementation of Hough Transform and Decision Trees will allow us to bring out the method proposed. Several methods can be used to classify images for different things. For instance, abnormal thermograms can be detected by the implementation of the CLAHE method. The filtering of the enhanced images in the method uses K-Means and Fuzzy C-Means and a comparison is set using the SVM and Bayesian classifiers [6]. Breast cancer classification is made using decision tree algorithms without using feature extraction and selection in [7] and using deep learning xception Algorithm in [8]. The rest of this paper is organized as follows. Section 2 describes the related work, section 3 describes the materials and methods and section 4 shows the experimental results and analyses. Finally, section 5 concludes the paper.

2. **Related Work.** Batik image classification using treeval and treefit as decision tree is made in [9] and automated classification of bitmap images using decision trees is made in [10]. Face recognition using hough transform based feature extraction is shown in [11,12]. A genetic algorithm-based feature selection is used in [13] and Fault Classification of a Centrifugal Pump in Normal and Noisy Environment with Artificial Neural Network and Support Vector Machine Enhanced by a Genetic Algorithm is made in [14]. A fruit classification system using Tensorflow mode is shown in [15]. Multispectral recognition using genetic and evolutionary feature extraction is used in [16]. The Operator-N Layer Construction is used for Optimizing Capsule Network Methods in the image classification problems is shown in [17]. [18] Shows feature selection based on hybridization of genetic algorithm and particle swarm optimization. Optimal Parameter-Feature Selection Using Binary PSO for Enhanced Classification Performance is shown in [19]. The applying of image processing technology for the face recognition is made in [20] and spam classification is made in [21].

3. **Materials and methods.** The suggested procedure is presented in this part, and Figure 1 illustrates the key components of the suggested system. Before the image can be successfully classified, it will need to undergo some preparation first, including the following steps.

3.1. **Image preprocessing.** Due to the specific purpose of this system, some image data can safely be discarded to simplify processing and reduce processing time and processing power required to perform the task. This is the first step of the proposed system. When input images to the system, preprocessing will be applied to images, which image resize of the image in accordance with the measured rate of (256 * 256) and then converted from (RGB) to (gray scale) two-dimensional matrix (Dual Matrix dimension by the rate of (256 * 256).

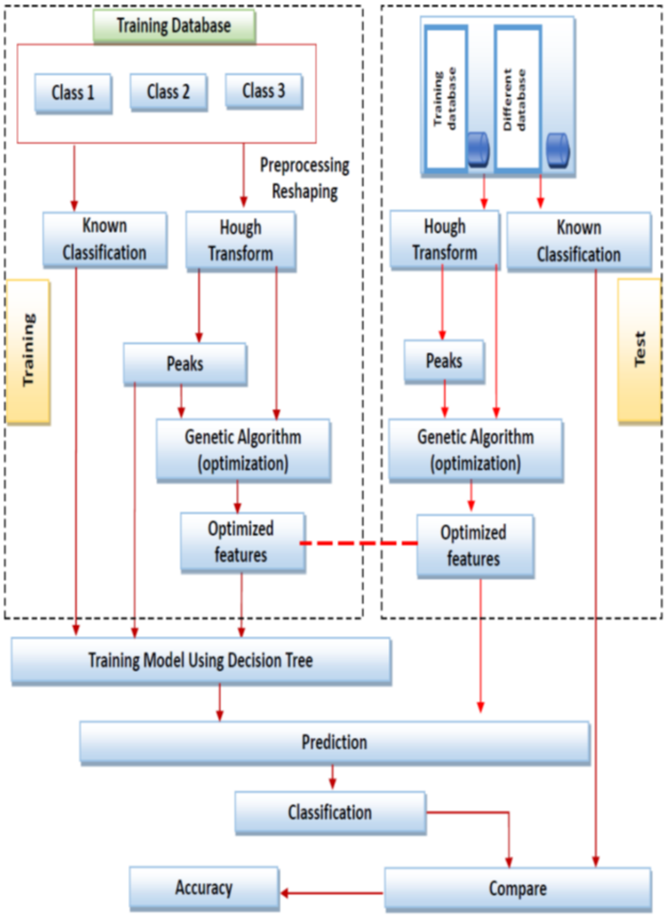

FIGURE 1. Main steps for the proposed system

3.2. **Feature extraction.** This part of study aims to determine the most advantageous combination of these techniques:

(1) Hough transform.

a. Edges Detection. In this part, the training image set (Normal + Abnormal) is initially processed by the system into black and white images. It then uses a cunning technique to detect edges, as shown below:

To eliminate noise from the image, smooth it with a Gaussian filter. A Gaussian filter kernel of dimension $(2k+1)*(2k+1)$ can be found using the following equation:

$$H_{ij} = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{(i-(k+1))^2+(j-(k+1))^2}{2\sigma^2}\right)} ; 1 \leq i, j \leq (2k+1) \tag{1}$$

Find the image's intensity gradients.

Use non-maximum suppression to eliminate erroneous edge detection responses.

To identify possible edges, use a double threshold.

Track edges using hysteresis: To complete edge detection, suppress all weak edges that are unconnected to strong edges.

b. Calculate Hough transform. Determine the standard Hough transform applied to the final edges image [22]. The parametric representation of a line is used by the Standard Hough Transform (SHT):

$$\rho = x cos(\theta) + y sin(\theta) \tag{2}$$

where the distance along a vector perpendicular to the line from the origin to the line is known as the variable ρ and θ is pronounced θ that is the angle expressed in degrees clockwise from the positive x-axis of the perpendicular projection from the origin to the line.

where θ has a range of -90°≤ θ< 90°measuring clockwise with regard to the positive x-axis, the lines angle is θ+90°. As seen in Fig. 2, the result of this conversion is a matrix (H) of θ (T) and matching ρ (R). The SHT is a matrix in parameter space whose rows and columns represent the values of θ and ρ, respectively. Accumulator cells are represented by the elements in the SHT. Every cell ρ each non-background point in the image. In SHT, ρ is rounded to the closest permitted row. A cell in the accumulator is increased. A value of Q in SHT(r,c) at the conclusion of this process indicates that Q points in the xy-plane lie on the line indicated by θ(c) and ρ(r). Potential lines in the input image are represented by peak values in the SHT. A matrix (H) containing the corresponding ρ (R) and θ (T) is the result of this transformation as shown in Figure 2 [22].

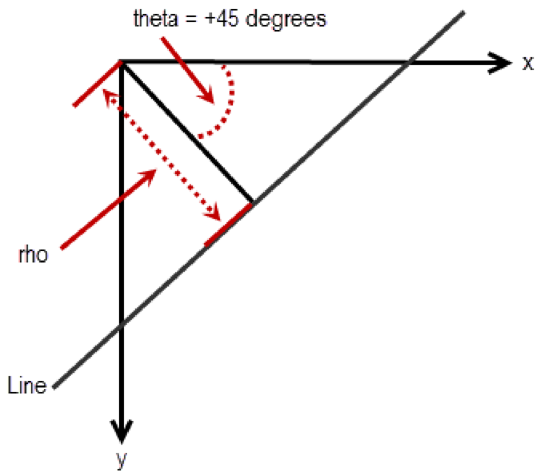

FIGURE 2. Hough transform [22]

c. Finding the peaks. To further reduce the results to most significant points, since the purpose of this test is the classification of the image, certain number of peak values of angles (θ) and distances (ρ) are selected, the optimum number of peaks for this training set is to be found by several trials.

(2) Genetic algorithm. This is among the most used techniques for feature extraction. Genetic Algorithm (GA) for feature selection is shown in one paper by Oluleye et al. The performance of the relevant classifiers was improved by dimensionality reduction using a binary genetic algorithm. The researchers used a series of photos from the Flavia dataset to extract one hundred (100) characteristics. The characteristics that were retrieved were Zernike Moments (ZM), Fourier Descriptors (FD), Legendre Moments (LM), Hu 7 Moments (Hu7M), Texture Properties (TP), and Geometrical Properties (GP) [13].

In a different study, Dharani et al. suggest using an I-Genetic method to classify objects based on their intensity ratings. To increase the speed of execution and object identification, all pixel concerns are removed. The object interest's intensity values are taken into account by the I-Genetic algorithm during cross-over operations. This algorithm effectively reduced the execution time [17]. An approach to feature selection based on the combination of particle swarm optimization and evolutionary algorithms is put out in another paper [16] and Feature Selection Using Binary particle swarm optimization for enhanced Classification Performance in [19].

A fitness value was determined by looking at the support vector machine classifier's accuracy on validation samples. The hyperspectral data set from Indian Pines was used to test that methodology. Results from this work demonstrate that the new approach does not require users to pre-specify the amount of required features; instead, it is able to automatically select the most informative characteristics in terms of classification accuracy within an acceptable CPU processing time frame. The outcomes demonstrate the ability to distinguish between background and road pixels, and outperform the other methods under evaluation in terms of performance parameters. Setting the number of desired features a priori is not necessary for the new feature selection technique. The findings demonstrate that the approach can find informative bands with respect to classification accuracies in a reasonable amount of CPU time; it can be applied to road detection; it is independent of data set distribution and does not require parameter initialization; it is also based on evolutionary techniques. The approach can automatically select the most valuable features with respect to classification accuracies, so there is no need to set the number of output features [18].

3.3. **Classification.** Decision trees were employed in this study as the classification technique. A classic and natural model of learning is the decision tree by definition. It has a lot in common with the basic computer science concept of "divide and conquer." Decision trees use a sorting technique to categorize instances by moving them from the root of the tree to a leaf node, which indicates the instance's categorization. Every node in the tree indicates a test of a particular instance attribute, and every branch that descends from that node represents one of the attribute's potential values [14,15]. Classifying an instance begins at the tree's root node. After testing the attribute that this node specifies, descend the tree branch that corresponds to the attribute's value.

3.4. **Testing system.** In this study, 120 gray scale indexed GIF images for digital mammography medical images have been used in the training set, split into 3 categories: normal, benign, and diseased, with 40 images in each category, also 60 images were used to test the proposed system, 30 of them were taken from the training set, and the other 30 were taken from another images from the same database and different from in the training set. All Images are courtesy of University of South Florida, Digital Mammography, database name is DDSM: Digital Database for Screening Mammography [23].

These images will be split into training sets and testing sets and will be used to train and test the system through the following steps:

1. Importing the images into the system, ensuring that the imported images are in the correct color depth (grayscale) and correct size (256 x 256) pixels.
2. Feature extraction using Hough transform.
3. Features extracted by Hough transform undergo further optimization using genetic algorithm.
4. Classification of the image using decision tree.
5. Compare the results with Hough transform results.
6. Along with the main comparison, the system is designed to validate its results by comparing it to these combinations.

1. Hough transform, feature peaks and decision tree.
2. Hough transform, genetic algorithm and decision tree.
3. Hough transform, Feature peaks, genetic algorithm and Decision tree.

For validation purposes training and test sets can be mixed, exchanged or combined to see the effects on the results.

These comparisons are being done with 2 datasets:

1. Same dataset used for training (training set).
2. Different dataset (testing set).

3.5. **Measurements.** For each classification methods, these measurements were taken:

1. Accuracy.
2. Error.
3. Processing time.

3.6. **Test Rig Specification (Hardware and Software).** Application has been built using "Mathworks Matlab" software version, 2013, it was running on Microsoft windows 7 ultimate SP1 machine, with these hardware specifications:

1. Model: Hewlett – Packard HP pavilion g6 notebook PC.
2. Processor: 2 core, 4 logical threads, Intel core i3 running at 2.2 GHz.
3. RAM: 6 GB of DDR 3 physical memory.
4. Hard drive: Hitachi 500 GB notebook HDD.

4. **Experimental results and analyses.**

4.1. **Classification via Hough transform, feature peaks and decision tree combination (HT_Peaks_DT).** (1) Using same training set shown in Table 1. As show in the Table 1, the accuracy varied according to the number of peaks selected since the number of peaks is chosen arbitrarily from 4 to 13 peaks to conclude the optimal number of feature peaks. According the obtained result the lowest number of peaks to give adequate results was 7 feature peaks, the process took 51.7 ms to complete.

TABLE 1. Hough transform, feature peaks and decision tree combination (training set)

Feature peaks count	**Error**	**Accuracy %**	**Time (ms)**
4	2	93.3333	24.5
6	1	96.6667	21.6
7	0	100	51.7
8	1	96.6667	23.7
9	1	96.6667	27.1
10	1	96.6667	23.1
11	1	96.6667	23.6
12	0	100	22.7
13	0	100	25.1

(2) Using different dataset (testing set) shown in Table 2. As shown in the Table 2, the accuracy didn't change much with the feature peaks count, yet the overall accuracy was poor (60% - 66%) which is not satisfactory in practical applications where higher accuracy is needed.

TABLE 2. Hough transform, feature peaks and decision tree combination (testing set)

Feature peaks count	**Error**	**Accuracy %**	**Time (ms)**
4	4	66.66	22.4
6	11	63.33	23.3
7	12	60	24.9
8	11	63.33	22.4
9	11	63.33	28.2

Feature peaks count	Error	Accuracy %	Time (ms)
10	11	63.33	28.3
11	11	63.3333	23.96
12	11	63.3333	22.8
13	11	63.3333	32.3

4.2. **Classification via Hough transform, genetic algorithm and decision tree combination (HT_GA_DT).** In this experiment, error, accuracy and processing time were measured for each combination of selected number of generations (from 5 to 11) and population count from (6 to 12).

(1) Using same training set shown in Table 3. As per the results in the Table 3, and since the test set and the training set are identical, it showed no change in accuracy by changing the process variables (Generation number and population count), however the processing time is much higher than the Hough transform, feature peaks and Decision tree combination.

TABLE 3. Hough transform, genetic algorithm and decision tree combination using (training set)

Gen. no.	Population count	Error	Accuracy (%)	Time (sec.)
5	6	0	100	3.70
6	7	0	100	4.13
7	6	0	100	4.13
7	7	0	100	3.52
8	7	0	100	4.03
10	10	0	100	3.42
10	28	0	100	1.81
11	12	0	100	3.67
15	15	0	100	3.25
15	30	0	100	3.60
22	10	0	100	3.45
30	30	0	100	3.55

(2) Using different dataset (testing set) shown in Table 4. The Table 4 shows that the accuracy increases in general with the increase of number of generations and population count, while processing time was almost constant except in rare cases, while accuracy has been noticeably increased in general.

increased in general.

TABLE 4. Hough transform, genetic algorithm and decision tree combination using (testing set)

Gen. no.	Population count	Err.	Accuracy %	Time (Sec)
5	6	9	70	3.56
6	7	8	73.3333	3.46
7	6	9	70	3.57
7	7	8	73.3333	3.38
8	7	8	73.3333	3.47
10	10	6	80	3.49

Gen. no.	Population count	Err.	Accuracy %	Time (Sec)
10	28	6	80	1.80
11	12	8	73.3333	3.83
15	15	8	73.3333	3.23
15	30	6	80	1.87
22	10	6	80	3.39
30	30	6	80	3.62

4.3. **Classification via Hough transform, Feature peaks, genetic algorithm and Decision tree combination (HT_Peaks_GA_DT).** In this proposed system, all previous methods were combined to perform the feature extraction and classification process. Aiming to have better classification performance with higher accuracy and lower processing time.

(1) Using same training set shown in Table 5. As per the Table 5, this experiment has 3 variables to control and test, the number of feature peaks, the initial population count, and number of generations for the genetic algorithm.

The results show that using the same training database, changing these variables has little or no effect on the measurements. But the processing time does show as low as if genetic algorithm were not used.

TABLE 5. Hough transform, Feature peaks, genetic algorithm and Decision tree combination (training set)

Feature	Gen. No.	Pop Count	Error	Acc.	Time (ms)
4	5	5	2	93.33	27.9
6	22	10	0	100	21.6
7	8	8	0	100	27
8	8	8	0	100	26.1
9	8	8	0	100	24.3
10	30	30	0	100	21.5
11	8	8	0	100	21.1
12	12	12	0	100	22.7
13	8	8	0	100	21.5
14	30	30	0	100	22.3
15	5	5	0	100	22.1
16	22	10	0	100	28.8

(2) Using different dataset (testing set) shown in Table 6. The results in the Table 6 show significant increase in accuracy and even shorter processing time comparing to the other tested methods, scoring up to 86.66 % Accuracy and 83.33% in the normal trend. It is also shown that changing the variables didn't have significant effect on processing time using a modern day computers.

For further verification, the number and mix of training and test images have been changed, and used to test the proposed method.

TABLE 6. Hough transform, Feature peaks, genetic algorithm and Decision tree combination (testing set)

Feature	Gen. No.	Population count	Error	Accuracy %	Time (ms)
4	5	5	4	86.67	20.8
6	22	10	9	70	21.6
7	8	8	6	80	22.1
8	8	8	9	70	21.9
9	8	8	6	80	22.2
10	30	30	9	70	22.8
11	8	8	9	70	25.9
12	12	12	5	83.33	22.4
13	8	8	8	73.33	21.3
14	30	30	5	83.33	27.2
15	5	5	6	80	23.8
16	22	10	5	83.33	20.6

4.4. **Additional verification.** Using 3 sets of images which were the total of 112 medical images: 46 of them were normal images, 38 carried benign diseases and the final 28 images were of patients who had cancer. In this step, the training and test image set will include:

(1) Using all 112 images for both training and testing shown in Table 7.

TABLE 7. Hough transform, Feature peaks, genetic algorithm and Decision tree combination using all 112 images for both training and testing

Feature	Error	Accuracy	Time
7	0	100	0.0271
10	0	100	0.0215
12	0	100	0.0227
13	0	100	0.0251
15	0	100	0.0221
16	0	100	0.0288
20	0	100	0.0268

(2) Using 46 images from the “normal” class shown in Table 8.

TABLE 8. Hough transform, Feature peaks, genetic algorithm and Decision tree combination using 46 images from the “normal” class using 46 images from the “normal” class

Feature	Error	Accuracy	Time
7	0	100	0.0271
12	0	100	0.0227
13	0	100	0.0251
15	0	100	0.0221
10	0	100	0.0215
16	0	100	0.0288

(3) Using 28 images from “benign” class shown in Table 9.

TABLE 9. Hough transform, Feature peaks, genetic algorithm and Decision tree combination using 28 images from “benign” class

Feature	Error	Accuracy	Time
7	0	100	0.0271
10	0	100	0.0215
12	0	100	0.0227
13	0	100	0.0251
15	0	100	0.0221
16	0	100	0.0288

(4) Using 38 images from the “Diseased” class shown in Table 10.

TABLE 10. Hough transform, Feature peaks, genetic algorithm and Decision tree combination

Feature	Error	Accuracy	Time
7	0	100	0.0271
12	0	100	0.0227
13	0	100	0.0251
15	0	100	0.0221
10	0	100	0.0215
16	0	100	0.0288

(5) Using 30 images with 10 images from each class shown in Table 11.

TABLE 11. Hough transform, Feature peaks, genetic algorithm and Decision tree combination using 30 images with 10 images from each class

Feature	Error	Accuracy	Time
7	0	100	0.0271
10	0	100	0.0215
12	0	100	0.0227
13	0	100	0.0251
15	0	100	0.0221
16	0	100	0.0288

(6) Using random images from all 3 classes shown in Table 12.

TABLE 12. Hough transform, Feature peaks, genetic algorithm and Decision tree combination Using random images from all 3 classes

Feature	Error	Accuracy	Time
7	0	100	0.0271
10	0	100	0.0215
12	0	100	0.0227
13	0	100	0.0251
15	0	100	0.0221
16	0	100	0.0288

The above tables show that the changes in the input images didn't have any effect on accuracy of the test and very little effect on the processing time, proving that the proposed system is not limited to the image set used beforehand.

The following graphs compare different methods performance in comparison to each other in Figures from Figure 3 to Figure 8. Figure 3, Figure 4, and Figure 5 for Error, accuracy, and processing time using training set and Figure 6, Figure 7, and Figure 8 for Error, accuracy, and processing time using testing set.

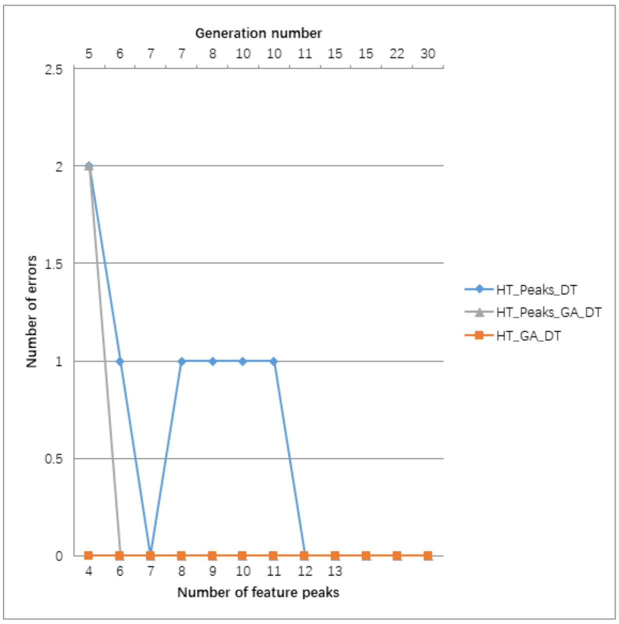

FIGURE 3. Error using training set

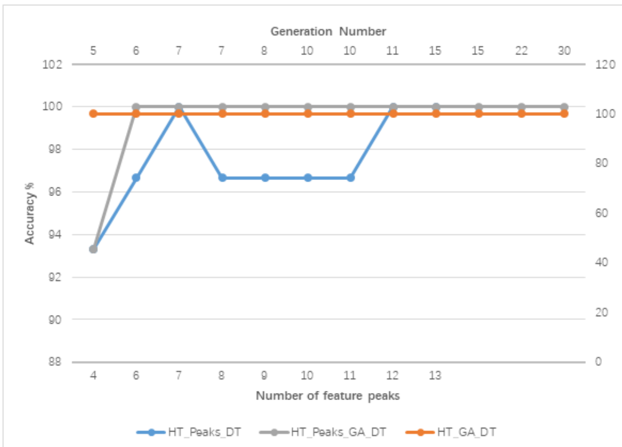

FIGURE 4. Accuracy using training set

The previous tables and figures show that there is a significant effect on choosing a different database other than the original database as the training set accuracy results ranged between 93.33% and 100 % while the highest accuracy obtained for a different dataset (testing set) was 86.67%, which is even more than what has been proposed earlier for Alzheimer disease diagnostics system [24] where accuracy didn't exceed 85% where however every method responded to this change differently as shown in Table 13.

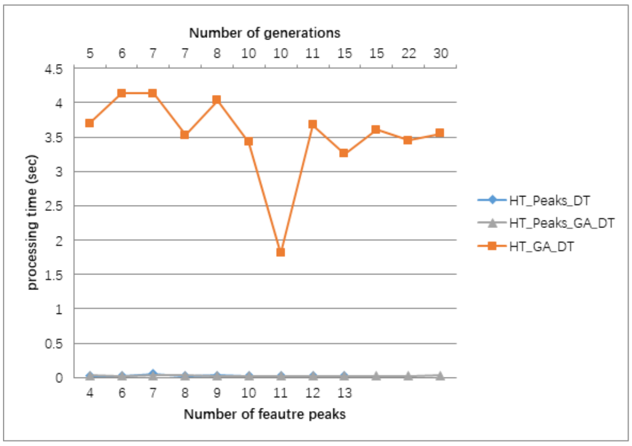

FIGURE 5. Processing time using training set

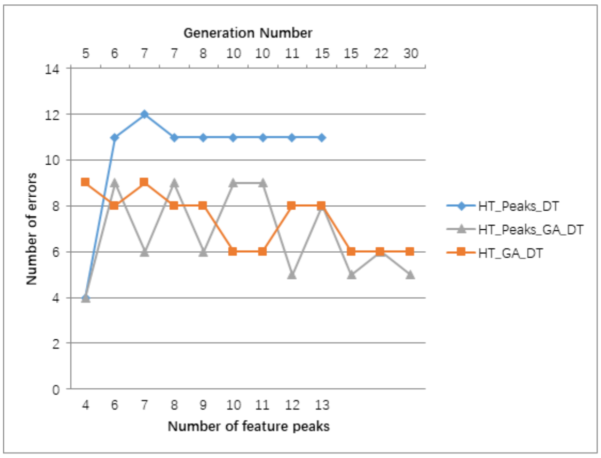

FIGURE 6. Error using testing set

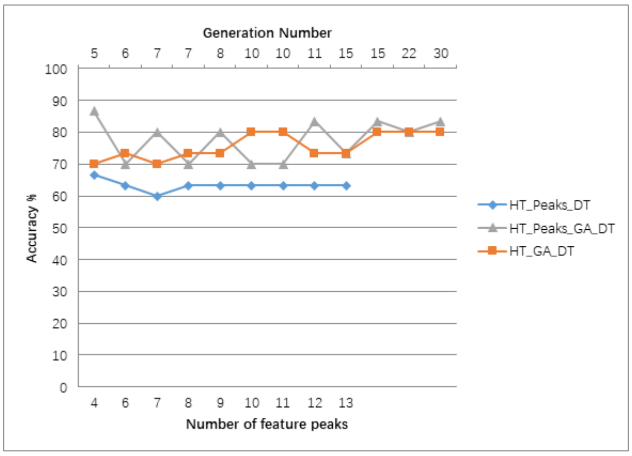

FIGURE 7. Accuracy using testing set

5. **Conclusion.** In this work, we have presented a novel classification technique for medical images which are digital mammography medical images. Three techniques were developed in order to come up with the best technique possible when it came to the errors that can happen, the accuracy of the classification, and finally the time needed to run the classification. Through the three techniques that were developed, one has proven to be

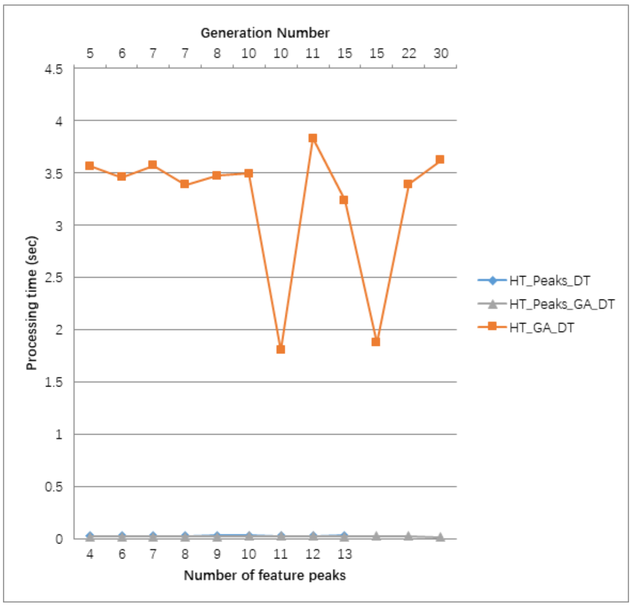

FIGURE 8. Processing time using testing set

TABLE 13. Results of the Alzheimer disease detection system

Method	**Accuracy**
Healthy subjects vs AD patients	
PCA	82.86%
NNMF	84.29%
Haralick	80.00%
Proposed method (multikernel)	85.00%
Proposed method (3 classifiers)	85.00%
MCI vs AD patients	
PCA	75.71%
NNMF	76.43%
Haralick	72.14%
Proposed method (multikernel)	79.29%
Proposed method (3 classifiers)	78.57%

the best out of them in terms of error, accuracy, and speed. The three techniques used two tests were undertaken, one had images that the machine was familiar with (training set) and the other test underwent images that the machine was unfamiliar with (testing set). The three techniques were Classification via Hough transform, feature peaks and decision tree combination, Classification via Hough transform, genetic algorithm and decision tree combination, and Classification via Hough transform, Feature peaks, genetic algorithm and Decision tree combination. The final technique that was used has proven to be the best in terms with the least number of errors, the highest accuracy and the least time undertaken to run the experiments. The results show that the training set accuracy results ranged between 93.33% and 100 % while the highest accuracy obtained for a different dataset (testing set) was 86.67%, which is even more than what has been proposed earlier for Alzheimer disease diagnostics system [22] where accuracy didn't exceed 85%. From these results we conclude that the performance of the classifier depends on the technique used, the type of the dataset used for training and the type of the dataset used to test the classifier.

Acknowledgment. We would like to thank the Institute of Graduate Studies and Researches in Alexandria University for supporting and sponsoring our studies and work, especially the Department of Information Technology and also Faculty of Engineering in Damanhour University.

References

[1] E. R.DAVIES. Vision, the challenge, *Machine Vision*, 1-15, 1990.

[2] Y. cai. Supervised non-negative matrix factorization for analysis of microbiome data,[Master Thesis], Dalhousie University; 2014.

[3] H. Lee, S. Choi. Group, Nonnegative Matrix Factorization for EEG Classification, *Lecture Notes in Computer Science* [Internet], 2009.

[4] Q. Guo, B-W. Chen, S. Rho, W. Ji, F. Jiang, X. Ji, et al., Divid-and-Conquer Classification Based on Feature-Space Decomposition, *Harbin, China. : Tsinghua University, Beijing, China; Princeton University, Princeton, NJ, USA; Harbin Institute of Technology*, 2015.

[5] P. Bulo SR Kontschieder, A. Criminisi, P. Kohli, M. Pelillo, H. Bischof, Context-Sensitive Decision Forests for Object Detection, *In: Proceedings of NIPS*, 440-8, 2012.

[6] D. Dinsha, N. Manikandapradu, Breast tumor segmentation and classification using SVM and byesian from thermogram images, *Unique Journal of Engineering and Advanced Sciences,* Vol. 2, No. 2, 147-51, 2014.

[7] O. Tarawneh, M. Otair, M. Husni, H. Y.Abuaddous, M. Tarawneh, M. Almomani, Breast Cancer Classification using Decision Tree Algorithms, *(IJACSA) International Journal of Advanced Computer Science and Applications,* Vol. 13, No. 4, 2022.

[8] B. S. Abunasser, M. R. AL-Hiealy, S. I. Zaqout, S. S. Abu-Naser, Breast Cancer Detection and Classification using Deep Learning Xception Algorithm, *(IJACSA) International Journal of Advanced Computer Science and Applications*, Vol. 13, No. 7, 2022.

[9] A. H. Rangkuti, Z. E. Rasjid, D. J. Santoso, Batik image classification using treeval and treefit as decision tree function in optimizing content based batik image retrieval,*Procedia Comput Sci,* Vol.59, 577–83, 2015.

[10] P. Surynek, I. LukÅ¡ovÃ¡, Automated classification of bitmap images using decision trees,*Polibits*, Vol. 44, 11-8, 2011.

[11] R. Varun, Y. V. Kini, K. Manikantan, S. Ramachandran, Face recognition using hough transform based feature extraction,*Procedia Comput Sci,* 46: 1491-500, 2015.

[12] R. Zou, H. Sun, Building extraction in 2d imagery using Hough transform. Bachelor,*Thesis. Faculty of Engineering and Sustainable Development, University of GÃ¤vle, GÃ¤vle, Sweden*, 2014.

[13] B. Oluleye, A. Leisa, J. Leng, and D. Dean, A genetic algorithm-based feature selection,*International Journal of Electronics Communication and Computer Engineering*, Vol. 5, No. 4, 899-905, 2014.

[14] A. Nourmohammadzadeh, S. Hartmann, Fault Classification of a Centrifugal Pump in Normal and Noisy Environment with Artificial Neural Network and Support Vector Machine Enhanced by a Genetic Algorithm,*1st ed. Springer International Publishing*, 2015.

[15] P. N. Huu, T. Nguyen, and T. Mai, Proposing a fruit classification system using Tensorflow model, *Journal of Information Hiding and Multimedia Signal Processing, Ubiquitous International*,vol. 14, no. 1, 2023.

[16] P. A. Arias, J. Shelton, K. Roy, G. V. Dozier, F. Ahmad, Multispectral recognition using genetic and evolutionary feature extraction, 2015.

[17] R. N. R. Wijaya, B. Setiyono, M. Yunus and D. R. Sulistyaningrum, Operator-N Layer Construction for Optimizing Capsule Network Methods in Image Classification Problems, *Journal of Information Hiding and Multimedia Signal Processing*, Vol. 14, No. 3, pp. 90-101, September 2023.

[18] P. Ghamisi, Feature selection based on hybridization of genetic algorithm and particle swarm optimization, *IEEE Geoscience and Remote Sensing Letters*, Vol. 12, No. 2, 309-13, 2015.

[19] T. Ngo, H. Nguyen-Cong, T. Nguyen, and T. Dao, Optimal Parameter-Feature Selection Using Binary PSO for Enhanced Classification Performance, *Journal of Information Hiding and Multimedia Signal Processing, Ubiquitous International*, vol. 14, no. 4, December, 2023.

[20] T. Liang, M. Chiu, Y. Wu, M. Yeh, Chao-Hsing Hsu and Yi-Nung Chung, Applying Image Processing Technology to Face Recognition, *Journal of Information Hiding and Multimedia Signal Processing*, Vol. 13, No. 2, pp. 106-112, June 2022.

[21] T. Tu, N. Loan and V. Toan, Assessing some formula of spam probability and applying in spam classification, *Journal of Information Hiding and Multimedia Signal Processing*, Vol. 13, No. 4, pp. 204-213, December 2022.

[22] The Mathworks, Inc. Hough: Hough Transform. Mathworks 2016.

[23] University of South Florida. Digital Mammography: Digital database for screening mammography (DDSM). Marathon 2006.

[24] F. Segovia, J. M. Gorriz, J. Ramirez, C. Phillips, Combining feature extraction methods to assist the diagnosis of Alzheimer's, Disease, *Curr Alzheimer Res*, Vol. 13, No. 7, 831-7, 2016.