

Received June 14, 2020, accepted June 29, 2020, date of publication July 8, 2020, date of current version July 24, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3008038

Fusion of Orthogonal Moment Features for Mammographic Mass Detection and Diagnosis

MOHAMED W. ABO EL-SOUD^{1,2}, (Member, IEEE), IMAD ZYOUT³, (Member, IEEE),
KHALID M. HOSNY⁴, (Senior Member, IEEE),
AND MOHAMED MESELHY ELTOUKHY^{5,6}, (Member, IEEE)

¹Department of Computer Science and Information, College of Science, Majmaah University, Al Majma'ah 11952, Saudi Arabia

²Department of Information System, Faculty of Computers and Informatics, Suez Canal University, Ismailia 41522, Egypt

³Communications, Electronics Engineering, and Computer Engineering Department, Tafila Technical University, Tafila 66110, Jordan

⁴Department of Information Technology, Faculty of Computers and Informatics, Zagazig University, Zagazig 44519, Egypt

⁵Computer Science Department, Faculty of Computers and Informatics, Suez Canal University, Ismailia 41522, Egypt

⁶College of Computing and Information Technology, Khulais, University of Jeddah, Jeddah 21921, Saudi Arabia

Corresponding author: Mohamed Meselhy Eltoukhy (mmeltoukhy@uj.edu.sa)

This work was supported by the Deanship of Scientific Research at Majmaah University under Project R-1441-141.

ABSTRACT Masses are mammographic nonpalpable signs of breast cancer. These masses could be detected using screening mammography. This paper proposed a system utilizing orthogonal moment invariants (OMIs) features for mammographic masses detection and diagnosis. In this work, three sets of OMIs features were extracted. These OMIs features are Gaussian-Hermite moments (GHMs), Gegenbauer moments (GeMs), and Legendre moments (LMs). The extracted features are fused and presented to the particle swarm optimization (PSO) algorithm for feature selection. The classification step is achieved using the support vector machine (SVM). The proposed system is evaluated using 400 regions, extracted from the DDSM dataset. The obtained results reveal the promising application of OMIs features for masses detection and identification. It shows that fusing the OMIs features produces an acceptable detection performance where the area under the receiver operating characteristics (ROC) curve is $A_z = 0.969 \pm 0.01$ and the best performance of OMIs features is $A_z = 0.856 \pm 0.053$ for characterizing the malignancy of masses.

INDEX TERMS Feature extraction, mammographic mass detection, orthogonal moment invariants, particle swarm optimization (PSO), support vector machine (SVM), PSO-SVM.

I. INTRODUCTION

Breast cancer is the most frequent cancer among women worldwide. It is impacting 2.1 million women each year. In 2018, it is estimated that 627,000 women died from breast cancer. It is approximately 15% of all cancer deaths among women [1].

Early detection of breast cancer is the key to reduce mortality rates. The treatment is easy to handle when breast cancer detected at an early stage, whereas late detection decreases treatment options.

Mammography is an effective tool that helps in the early detection of breast cancer. It has the ability to detect the abnormality before physical symptoms appear. Regular screening may help to detect breast cancer in its early stage, before it is developed into a systemic disease and may even

invade other body parts. The most common symptom of breast cancer is a lump or mass in the breast. Mammogram interpretation is a difficult task even for skilled radiologists due to the overlapping of subtle signs of breast abnormalities and tissue. Computer Aided Detection/Diagnosis (CAD) systems were developed to assist the radiologist in pinning out the suspicious regions in mammograms. Therefore, there is a pushing need for developing CAD systems to identify and classify mammograms in order to reduce the high miss-detection rate [2].

The CAD system is consists of four steps, namely, image segmentation, feature extraction, features selection, and classification. The image segmentation step is used to identify the region of interest (ROI). Feature extraction aims to calculate the features that able to determine whether the ROI is normal or abnormal. Feature selection is the step that identifies the high significant features that able to distinguish between different classes. The classification step is to determine the

The associate editor coordinating the review of this manuscript and approving it for publication was Yakoub Bazi.

ROI class according to the presented feature vector. The overall performance of any CAD system is affected by the achievement obtained from each different step. The feature extraction step is a significant step in developing CAD systems; however, the complexity of breast tissues makes it difficult to find the prominent features that able to distinguish between different classes [3], [4].

Several attempts have been done to overcome the challenging issue of extracting an efficient set of features for mammogram classification [5]–[8]. The studies indicated that the feature space of ROIs is vast and complicated because of the wide diversity of healthy tissues and the abnormalities. The use of excessively many features may degrade the classification methods' performance and increase the classifier's complexity. Therefore, the extraction of features is an essential task for classifying breast tumors. Moreover, with the advances in information technology, the problem size and search space increase exponentially for any image processing problem.

Dahabi *et al.* [8] applied the *t*-test to rank the curvelet and moment features according to their capabilities [4]. Their results indicated that curvelet moments are efficient and useful for breast cancer diagnosis.

Gardezi *et al.* [9] presented a feature extraction method based on decomposing the mammogram image using a curvelet. They calculated the grey level co-occurrence matrix (GLCM) of the curvelet. They reported an accuracy of 88.6%, the sensitivity of 76.53%, and specificity of 91.3%.

Jiang *et al.* [10] proposed a CAD system to define the masses in digital mammography using scale-invariant shift transform (SIFT) features. They reported an accuracy of 86.9%. Eltoukhy and Faye [6] applied a statistical method to maximize the distance between two different mammographic classes. Their proposed method was applied over wavelet and curvelet coefficients. The reported results showed an accuracy of 91.19% classification rate.

Metaheuristic algorithms have been growing for optimization problems. These algorithms are driven by simulating the behaviors of natural phenomena. They have the capabilities of extracting information from a set of features and often generate the best features in practice. Therefore, several algorithms have been developed to address optimization problems. Among them, some metaheuristic search algorithms with population-based frameworks such as GAs, simulated annealing (SA), ant colony optimization (ACO), and PSO [11], [12] have proved adequate capabilities for managing high-dimensional optimization problems.

Ramos *et al.* [13] proposed a system to classify mammographic images as normal or mass. They compared the performance of Haralick, wavelet, and ridgelet features. A genetic algorithm (GA) was employed to select the prominent features, and the random forest was used for the classification task. Their results demonstrated that wavelet-based features using GA achieved an area under curve (AUC) = 0.90.

Rouhi *et al.* [14] proposed to use Zernike moment features to identify the benign mass from malignant. GA is used

as a feature reduction method to improve the accuracy and decrease the computational cost. They concluded that multi-layer perceptron (MLP) classifier is promising compared with the existing methods.

Zyout *et al.* [7] proposed a false positive reduction algorithm, including the PSO model selection, to define an appropriate set of features from wavelet and GLCM features. SVM based classification method is used to determine the class of the suspicious region. Their results recommended using PSO to reduce the false-positive rate.

Chen *et al.* [15] proposed to develop a CAD scheme building an initial feature pool containing four different groups of features. Next, a particle swarm optimization (PSO) algorithm was applied to select optimal features so that redundant features can be removed from the feature pool. Finally, an SVM was used to determine either the case is benign or malignant.

Other meta-heuristic algorithms have been used to solve different problems, including mammographic mass segmentation [16]. Khehra and Pharwaha [17] developed a comparison study between three different metaheuristic methods, namely GA, PSO, and biogeography based optimization (BBO), to select an optimal set of features from 50 features. In their work, they reported a 91% classification rate is using PSO-SVM. The classification accuracy rate has been achieved with half of the total features.

From the literature, it is concluded that metaheuristic algorithms have great potential for selecting an optimal subset of features from a set of features extracted from mammograms. It has been observed that these biologically motivated metaheuristic algorithms perform better than the classical optimization approaches. The main advantage of these algorithms is domain-independent nature. Secondly, these algorithms could find optimal or near-optimal solutions in an ample search space.

On the other hand, a group of studies focused on the fusion of texture features of mediolateral oblique (MLO) and cranial-caudal (CC) views of mammograms to improve the performance of mammography CADx. Several features extraction techniques were used including local binary patterns (LBP) [18] and [19], binary Gabor patterns (BGP) [20], K-Gabor filters [21], gray-level co-occurrence matrix (GLCM), Law's texture, GRFCM, and gray level difference method (GLDM) [22]. The feature vector obtained using either serial or parallel approach were further engineered and improved by applying principle component analysis (PCA), Genetic Algorithm (GA), Particle swarm optimization (PSO), binary bat algorithm (BBA), binary Firefly algorithm (BFA), and canonical correlation analysis (CCA). The performance of LBP fused features was further processed, and the performance improved after applying feature selection based on Firefly and optimum-path Forest classifier [19]. In [20], BGP texture features were found more efficient than LBP features. For classification, classical machine learning algorithms based on K-nearest neighbor, support vector machines and optimum-path forest classifiers were applied. Reported

results, from the DDSM (screen-film) and INbreast (digital full-field) mammography, showed that the fusion of textures of both mammographic views had improved the performance compared to pertinent studies.

Wang *et al.* [23] fused 20 deep features from CNN with five texture features, five morphological features, and seven density features of the 400 mammographic images. Patil and Biradar [24] applied a hybrid deep classifier based on a combination of convolutional neural network (CNN) and recurrent neural network (RNN) namely (CRNN). They extracted the Grey level co-occurrence matrix (GLCM) and gray level run-length matrix (GLRLM) features. Their system produced an accuracy of 90.59%. The deep convolutional neural networks (DCNN), indeed, surprisingly solve many image classification problems. The fantastic trait of DCNN could be the seamless approach for feature engineering and pattern classification. However, the screen-film mammography data, available for the research community, is limited. We believe that the use of a classical machine learning approach by focusing on the feature extraction through cascade fusion and embedded feature selection remains an adequate approach for successfully solve the mammography CADE and CADx tasks.

Eltoukhy *et al.* [25] utilized exact Gaussian-Hermite moments with K-NN, random forests and AdaBoost classifiers. Their system is evaluated using mammographic image analysis society (MIAS) [26] and image retrieval in medical applications (IRMA) [27] datasets. The obtained accuracies are 90.56% for MIAS and 93.27% for IRMA. The successful utilization of orthogonal Gaussian-Hermite moments [25] motivated the authors to fuse three kinds of orthogonal moments, GHMs, LMs, & GeMs, for extracting the fine features from the input images. The orthogonal moments, GHMs, LMs, & GeMs have the following attractive characteristics:

- 1) These moments are orthogonal, which means their ability to represent digital images with minimum information redundancy.
- 2) These moments are invariant with respect to rotation, scaling and translation, which enables the computer-based systems to discriminate between similar images, whatever their orientation, location and distance to the camera.
- 3) These moments are computed with highly accurate methods that reflect the accuracy of extracted features from the input images.
- 4) These moments are robust against the well-known kinds of noise.

Thus, in this study, we hypothesized that it is possible to identify and fuse the global and local features computed from the ROIs of mammographic masses. These moment-based features could produce a high-performance CADE/CADx. Besides, applying the PSO could reduce the requirement of an extensive training dataset as the conventional deep learning approach. Thus the objective of this study is to analyze the

combination of three moment-based features to find out the best set of features that have the capability to distinguish between different mammographic masses, Either normal or abnormal, and the abnormal class is distinguished into benign or malignant.

Feature Fusion is applied to combine the advantage of different moments features and enable the extraction of coarse and fine features from the input images. The idea is to combine the substantial information of several moment features to ensure that no details are lost. Hence, extracting different texture descriptors from each image will produce a good representation of the processing image. We proposed combining three-moment features to investigate the performance of their combination because of the increasing need to integrate different moments features. In the following we summarize the contributions of this work:

- 1) We exploit the behavior of three distinct moments features for breast mass detection CADE and breast mass diagnosis CADx, where each feature type has its characteristics. A further goal of an advanced fusion method is to increase the classification accuracy rate by combining different features sets.
- 2) A more comprehensive system based on diverse characteristics of various moment features produced. The obtained results encourage the claim that combining distinct features will gain the advantage of each type and remove their drawbacks.
- 3) Extracting and identifying an adequate set of features that have high capability to distinguish between the different mammographic images' types, either mass or normal, and then benign or malignant.
- 4) We are applying embedded feature selection using PSO-SVM to accomplish: hyper-parameter selection, feature dimensionality reduction, and SVM-classifier parameter-tuning and performance optimization.

Orthogonal moments in cartesian coordinates such as LMs, GeMs, and GHMs are defined by multiplying two basis functions in the x- and y-directions. The x-direction basis functions of LMs, GeMs and GHMs are computed for different orders, 0, 1, 2, 3, 4, & 5. For simplicity, the computed basis functions for the order 5 are plotted and displayed in the Figure (1.a), (1.b) and (1.c), respectively, while the fused basis functions are plotted and displayed in figure (1.d). The plotted curves clearly show that both LMs and GeMs are very close and oscillates at a lower frequency and uniformly distributed over the interval $(-1 \leq x \leq 1)$ while the GHMs are oscillates at higher frequencies and non-uniformly distributed over the same interval.

This paper proposes a system to identify the mammographic mass from the normal regions, and then distinguish the benign from malignant regions. Particularly, the proposed approach takes advantage of combining the theory of orthogonal moments with the power of the parameter selection based on PSO-SVM algorithms. Three sets of OMIs features, GHMs, GeMs, and LMs, are extracted where the extracted

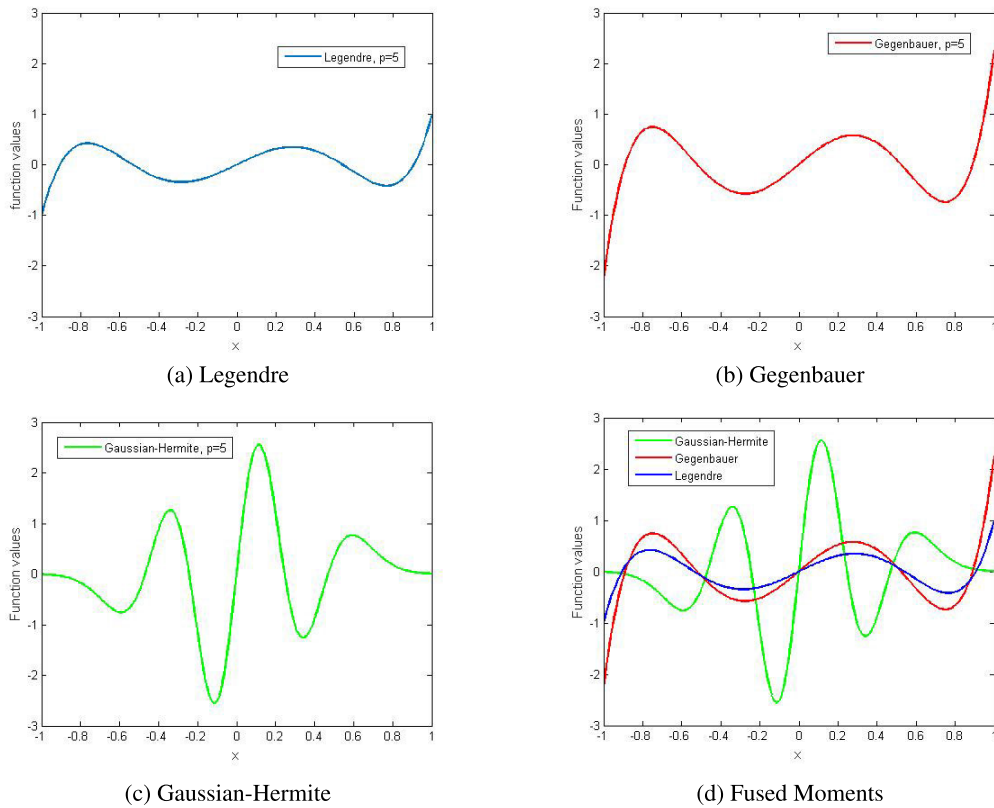


FIGURE 1. The basis functions of the orthogonal moments in cartesian coordinates.

features are presented to an SVM classifier. The proposed system is evaluated using 400 regions obtained from the DDSM dataset [28]. To the best of the authors’ knowledge, there is no previous study has applied orthogonal moments feature and PSO-SVM for implementing the detection and diagnosis tasks of mammographic masses.

The rest of this paper is organized as follows. Section II briefly presents the preliminaries of OMIs for feature extraction. Section III discusses the PSO-SVM algorithm. Section IV explains the proposed CAD system. The results are presented and discussed in Section V. Finally, the presented work is concluded in Section VI.

II. ORTHOGONAL MOMENT INVARIANTS

This section presents the preliminaries of proposed orthogonal moment invariants (OMIs) features. Since the images of the DDSM dataset are captured and rasterized using Cartesian pixels, orthogonal moments in Cartesian coordinates are preferable where no need for coordinate conversion nor image mapping. Orthogonal moments in Cartesian coordinates of order $(n + m)$ are:

$$A_{n,m} = \frac{1}{C_n C_m} \int_{-1}^1 \int_{-1}^1 f(x, y) P_n(x) P_m(y) w(x) w(y) dx dy \quad (1)$$

where $n \geq 0, m \geq 0$; C_n and C_m are the normalization factors; $P_n(x)$ and $P_m(y)$ are the orthogonal polynomial functions; $w(x)$ and $w(y)$ are the weight function in x - and y - directions, respectively. Table 1 gives a concise and

explicit mathematical formula for the utilized three orthogonal moments.

Where $\Gamma(\cdot)$ and σ are gamma function and the standard deviation and $\lfloor \frac{n}{2} \rfloor$ is:

$$\lfloor \frac{n}{2} \rfloor = \begin{cases} \frac{(n + 1)}{2} & \text{for odd } n \\ \frac{n}{2}, & \text{for even } n \end{cases} \quad (2)$$

For moment order, Max, the total number of independent moments is:

$$Total = \frac{(Max + 1)(Max + 2)}{2} \quad (3)$$

The Gegenbauer polynomials are generic polynomials with the scaling parameter $\alpha > -0.5$. Chebyshev polynomials of the second kind and Legendre polynomials are special cases of Gegenbauer polynomials when $\alpha = 0.5$ and $\alpha = 1$, respectively. Pawlak [29] shows that the scaling parameter enabled Gegenbauer moments to extract both local and global features of an image.

A. ACCURATE COMPUTATION OF ORTHOGONAL MOMENTS

Orthogonal moments as defined in Table 1 are computed accurately by using the following form:

$$A_{nm} = \frac{1}{C_n C_m} \sum_{i=1}^M \sum_{j=1}^N T_{nm}(x_i, y_j) f(x_i, y_j) \quad (4)$$

TABLE 1. Mathematical formulae for the utilized three orthogonal moments.

Moment	Normalization Factor	Polynomial Function $P_n(x)$	Weight Function
Gaussian-Hermite	$C_n(\sigma) = \sqrt{2^n n! \sigma \sqrt{\pi}}$	$H_n\left(\frac{x}{\sigma}\right) = \sum_{p=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^p \frac{n!}{p!(n-2p)!} (2x)^{(n-2p)}$	$w(x) = e^{-x^2}$
Gegenbauer	$C_n(\alpha) = \frac{2\pi \Gamma(n+2\alpha)}{2^{2\alpha} n! (n+\alpha) [\Gamma(\alpha)]^2}$	$G_n^\alpha(x) = \sum_{p=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^p \frac{\Gamma(n-k+\alpha) 2^{\binom{n-2p}{2}}}{(p!(n-2p)! \Gamma(\alpha))} (x)^{(n-2p)}$	$w^\alpha(x) = (1-x^2)^{(\alpha-0.5)}$
Legendre	$\frac{2}{(2n+1)}$	$L_n(x) = \sum_{p=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^p \frac{(2n-2p)!}{2^n p!(n-p)!(n-2p)!} (x)^{n-2p}$	$w(x) = 1$

where

$$T_{nm}(x_i, y_j) = \int_{x_i - \frac{\Delta x_i}{2}}^{x_i + \frac{\Delta x_i}{2}} \int_{y_j - \frac{\Delta y_j}{2}}^{y_j + \frac{\Delta y_j}{2}} P_n(x_i) P_m(y_j) w(x_i) w(y_j) dx dy \tag{5}$$

Double integration represented by equation (5) could be evaluated separately. For a digital image of size $M \times N$, the points $(x_i, y_j) \in [-1, 1] \times [-1, 1]$ are:

$$x_i = \frac{-2 + (2i - 1)}{2} \Delta x; \quad y_j = \frac{-2 + (2j - 1)}{2} \Delta y; \tag{6}$$

The constant steps are $\Delta x = \Delta x_i = \frac{2}{M}$ and $\Delta y = \Delta y_j = \frac{2}{N}$, where $i = 1, 2, 3, \dots, M, j = 1, 2, 3, \dots, N$.

Equation (4) could be rewritten as follows:

$$A_{nm} = \frac{1}{C_n C_m} \sum_{i=1}^M \sum_{j=1}^N IX(x_i) IY(y_j) f(x_i, y_j) \tag{7}$$

where the kernels are defined as follows:

$$IX(x_i) = \int_{x_i - \frac{\Delta x_i}{2}}^{x_i + \frac{\Delta x_i}{2}} P_n(x_i) w(x_i) dx \tag{8}$$

$$IY(y_j) = \int_{y_j - \frac{\Delta y_j}{2}}^{y_j + \frac{\Delta y_j}{2}} P_m(y_j) w(y_j) dy \tag{9}$$

Representing the orthogonal polynomials and the weight functions, $P_n(y_j), P_m(x_i), w(x_i)$ and $w(y_j)$. Using Table 1 yields six kernels in x - and y - directions for computing GHMs, GeMs and LMs, respectively. The kernels, $IX_n(x_i)$ and $IY_m(y_j)$, are exactly computed by using the principle of Calculus. Details for this computational approach for the three sets of orthogonal moments are presented in LMs [30], GeMs [31] and GHMs [32].

Based on the extreme importance of the invariance to rotation, scaling and translation (RST) in pattern recognition applications. Highly accurate methods were proposed for Legendre moment invariants (LMIs) [33], [34], Gegenbauer moment invariants (GeMIs) [35] and Gaussian-Hermite moment invariants (GHMIs) [36]. Easily use of the OMI, required converting the 2D matrices of the three kinds of moment invariants into 1D vectors which achieved by the pseudo-code [37].

Each one of the orthogonal moments, LMs, GeMs & GHMs, are used individually to extract the features from the input images. Based on equation (3), the total number of the extracted features is the same where the length of each feature vector is *Total*. The fusion of the features is carried by combining the feature vectors of three moments for each ROI. The extracted features are fused in one feature vector of the length, $3 \times Total$.

III. PSO-SVM ALGORITHM

Feature engineering aiming at extracting a few but discriminative features, which is the essential components of any machine learning algorithms, including SVMs. To improve the generalization capacity of the non-linear SVM classifier, hyper-parameters need to be optimized. The SVM hyper-parameters include the input features, the kernel type and its control parameters. The straightforward and optimal solution is through the exhaustive or even the grid search to find the optimal parameters. However, such search methods are impractical and the computational complexity is extreme. A suboptimal but efficient alternative of the grid search and exhaustive search is the metaheuristic algorithms such as PSO and GA. Among these algorithms, PSO has been shown to be very efficient with many features; simple structure, easiest to implement, and its ability to avoid local minima. Since PSO algorithm was introduced by Eberhart and Kennedy in 1995, many studies have reported various PSO modifications and successfully used it [7], [38].

Particle Swarm Optimization (PSO) is a biology-inspired metaheuristic search approach. The critical concept of PSO is the swarm intelligence in which members of the swarm collaborate to solve the optimization problem. Each member (candidate solution or particle) of the population belongs to the multidimensional parameter space of the objective function of the problem. PSO algorithm, in particular, accomplish the global optimization task by utilizing both local fitness of individual particles and the experience and fitness so far achieved by the entire population. The fitness of each particle is continuously (each step of the search process) evaluated and its characteristics are updated using the best experience achieved so far by other members in the swarm. Unlike genetic algorithms, PSO uses a velocity operator rather than the cross-over and mutation to control and update the location

and search direction of each particle. Also, use a large population helps PSO to escape from local minima. For a better generalization capacity and higher classification accuracy, both classifier's parameters and hyper-parameters (kernel function and feature spaces) need to be optimized. For this purpose, we adopted the PSO-SVM algorithm, an embedded feature selection and model selection approach from [7]. The classifier type and the feature extraction methods determine the dimensionality of the PSO search space to be used for optimizing the performance of PSO-SVM algorithm. The procedure describing the proposed PSO-SVM is presented in Algorithm 1.

IV. THE PROPOSED CAD SYSTEM

This section presented a description of the proposed CAD system. This work is primarily applies the OMIs features, first, for classifying a mammographic region into normal and abnormal classes by characterizing the presence of masses in the ROI, and second, for classifying an abnormal region into benign or malignant. The CAD system consists of four steps, segmentation, feature extraction, features selection, and classification.

In this work, the image segmentation is achieved manually using the given center of the suspicious region as given in the dataset by the radiologists. Extracting and identifying adequate features is a vital step to achieve a high classification performance. For each ROI, three sets of OMIs features were extracted. Each one of the orthogonal moments, LMs, GeMs & GHMs, are used individually to extract the features from the input ROI. The total number of the extracted features for each ROI is 231 features. The fusion of the features is carried by combining the feature vectors of three moments for each ROI. The extracted features are fused in one feature vector of length 693 features.

The extracted features are input to an SVM classifier. Common data partitioning and cross-validation procedures are applied to the validation datasets to obtain training and testing sets that will be used to evaluate the discriminative power of the moment based features.

To further improve the overall performance of the proposed CAD system, we used the PSO-SVM embedded feature selection approach, for both determining the most important features and also for tuning the parameters of the SVM classifier. As for the settings of the PSO algorithm, we followed the work in [7]. The proposed PSO-SVM parameter selection, OMIs features extraction, and SVM based classification methods were all compiled in MATLAB. Figure 2, illustrates the steps of the proposed CADe/CADx system.

A. PSO-SVM SETUP

In this work, the PSO-SVM settings include the fitness criterion based on the Az-value of the ROC curve, the swarm size is set to 100 particles structured as described in Section III, and the maximum number of iteration of 50. Further, the population of the swarm is initialized assuming that each parameter belongs to a random variable that is uniformly distributed

Algorithm 1 PSO-SVM Algorithm

- 1: Initialize population of dimension $L = N + 2$ is represented as $P = [P_1, P_2, \dots, P_M]^T$ with size M . The dimension $N + 2$ consists of the N features to be optimized and the two dimensions are used for tuning the classifier parameters γ and C . Also T denotes the transpose operator;
- 2: Set two random numbers g_1, g_2 generated among $[0, 1]$, an inertia weight ω , and parameters c_1, c_2 ;
- 3: Initialize positions $P_x = [P_{x,1}, P_{x,2}, \dots, P_{x,L}]^T$ of each particle P_x where $x = 1, 2, \dots, M$ of population;
- 4: Initialize velocity $V_x = [V_{x,1}, V_{x,2}, \dots, V_{x,L}]$ of each particle V_x where $x = 1, 2, \dots, M$ of population;
- 5: Evaluating the fitness of each particle $F_p^i = f(P_x^i), \forall x$, and get the best particle of population up to iteration i ;
- 6: Set iteration count $i = 1$;
- 7: Comparing the fitness values and select the local best particle $Lbest_p^i = P_x^i$, and global best particle $Gbest^i = P^i$;
- 8: Compute an inertia weight ω ;
- 9: Update of each particle the position and velocity

$$V_{x,y}^{i+1} = \omega \times V_{x,y}^i + c_1 g_1 (Lbest_{x,y}^i - P_{x,y}^i) + c_2 g_2 (Gbest_{x,y}^i - P_{x,y}^i) \quad (10)$$

$$P_{x,y}^{i+1} = P_{x,y}^i + V_{x,y}^{i+1} \quad (11)$$

where the index $x = 1, 2, \dots, M$ and the index $y = 1, 2, \dots, L$.

- 10: Evaluate updated fitness $F_x^{i+1} = f(P_x^{i+1}), \forall x$ (of each particle) and get the best particle;
- 11: Update $Lbest \forall x$ (of each particle)

$$\begin{aligned} &\text{If } F_x^{i+1} < F_x^i \text{ then } Lbest_x^{i+1} = P_x^{i+1}; \\ &\text{else } Lbest_x^{i+1} = Lbest_x^i; \end{aligned}$$

- 12: Update $Gbest$

$$\begin{aligned} &\text{If } F_{b1}^{i+1} < F_b^i \text{ then set } b \\ &\quad = b1 \text{ and } Gbest^{i+1} = Lbest_{b1}^{i+1}; \\ &\text{else } Gbest^{i+1} = Gbest^i; \end{aligned}$$

- 13: If $i < \max$ of iteration then $i = i + 1$ and go to step 7; else go to step 14;
- 14: Get optimum solution: print the outputs of generation as $Gbest^i$;
- 15: Retrain SVM with optimum parameters and features; then identify unknown samples on testing dataset.

in the corresponding search space. As for the control parameter PSO parameters controlling the search process: c_1 and c_2 were both set to 2, inertia ω monotonically decreased from 1.2 to 0.4 as the number of iterations increased. The PSO search process was terminated either if the maximum number of iterations of 50 was reached or perfect Az-value achieved.

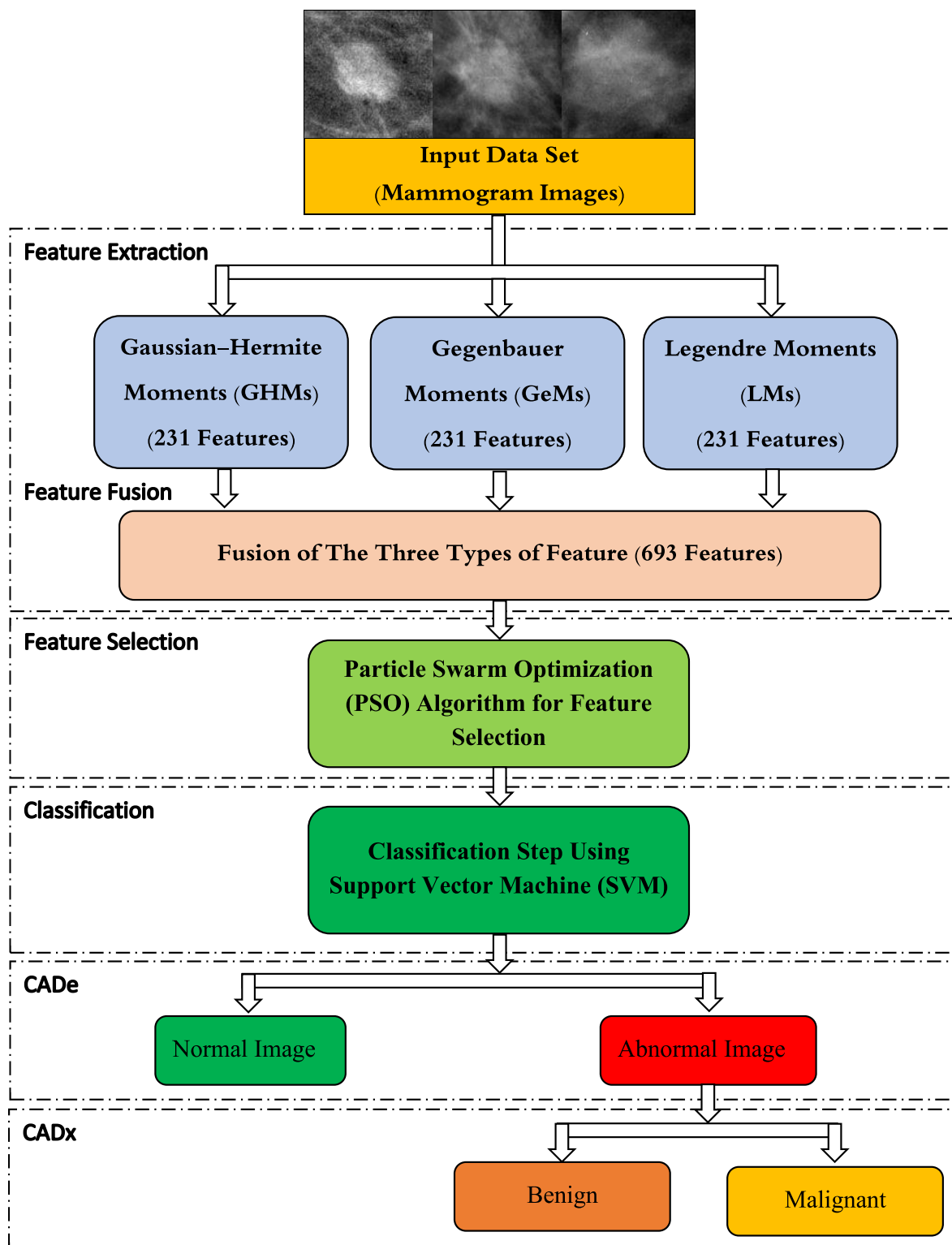


FIGURE 2. The proposed CADe/CADx system steps.

B. PERFORMANCE EVALUATION

For evaluating the discriminating ability of the OMI features, each dataset was randomly partitioned into training and testing sets. With the two-third of the dataset used for training

and the remaining one-third was held for testing. To achieve a better generalization capacity of the C-SVM classifier, 10-fold cross-validation was further applied to the training set to accomplish SVM learning and parameters optimization.

Namely, the cross-validation and heuristic search were used to tune the classifier regularization C constant and the radial basis function control parameter γ .

With the use of the SVM classifier for solving the two-class classification problem, the classifier's decision value (or the class membership) was used to produce the receiver operating characteristics curve (ROC). The area under the ROC curve or Az-value was estimated. Additionally, we included, for each ROC curve, the corresponding accuracy, false-positive fraction (FPF), and false-negative fraction (FNF). For the final evaluation, we used the average of all trials of the cross-validation and testing stages.

V. EXPERIMENTAL RESULTS AND DISCUSSION

For evaluating and obtaining the results of the applying the proposed OMIs features to distinguish between abnormal and normal classes (CADE or detection) and to differentiate malignant from benign classes (CADx or diagnosis), we used a set of mammographic regions of interest (ROIs) extracted from the digital dataset for screening mammography (DDSM) [28]. The following subsection is briefly describe the used validation datasets.

A. MAMMOGRAPHY DATASETS (DDSM)

The DDSM dataset is the largest, public and free mammography dataset, which is commonly used by the mammography image analysis research community. Two ROI datasets, as described in Table 2, named detection and diagnosis sets were formed. The dataset used for validating the detection task contained 200 ROIs depicting annotated masses and 200 regions representing normal breast parenchyma. Each case, in the DDSM dataset, includes radiologists annotation (a chain code representing the radiologist delineation of the abnormality), which we have used to extract a rectangular patch containing the mass in the center.

On the other hand, normal suspicious regions, each of size 512×512 related to normal breast parenchyma, were manually extracted, from normal mammograms, such that a selected ROI does not overlap with pectoral muscle or radiographic background. Moreover, for normal regions were obtained from Normal 11 and Normal 09 volumes of the DDSM dataset. The sample of the regions used are shown in Figure 3. We excluded normal ROIs and formed another dataset, CADx dataset, from the 200 abnormal (100 benign and 100 malignant) ROIs to apply the proposed features to classify benign and malignant masses. The summary of the regions in the detection and diagnosis datasets are described in Table 2.

B. RESULTS OF THE PROPOSED CADE SYSTEM

Three different OMIs features were first used for implementing the CADE system (i.e. classification of abnormal and normal regions). In other words, we have applied the different feature sets, namely, GeMs-features, GHMs-features, and LMs-feature for distinguishing between normal breast parenchyma related ROIs and mass depicting

TABLE 2. Description of the ROI datasets.

Dataset	Abnormal ROI		Normal ROI
	Benign	Malignant	
CADE dataset (Detection)	100	100	200
CADx dataset (Diagnosis)	100	100	0

(or abnormal) ROIs. Table 3 presents average cross-validation and test results of applying different OMIs features to the CADE dataset (i.e. the dataset with 200 abnormal and 200 normal ROIs). Obtained results, namely, ROC analysis (Az-value) of SVM classification results show an excellent performance of using various moment features for characterizing the presence of masses. GeMs-features produced average test classification results of 0.929 ± 0.021 (0.935 ± 0.014) which are a little higher than 0.926 ± 0.021 (0.935 ± 0.014) and 0.918 ± 0.021 (0.926 ± 0.011) respectively produced by GHMs-features and LMs-features.

To further examine the adequacy of applying moment features for implementing both CADE and CADx systems, we have examined whether the early fusion (combining) of different feature subsets can boost the classification performance. The results of different feature combinations are also presented in Table 3. The obtained results, from Table 3, show that combining the three sets of features, (GeMs + GHMs + LMs) produced an average classification performance of 0.935 ± 0.022 (0.948 ± 0.013). These results are higher than the best classification performance produced by individual feature sets, namely, GeMs-features. However, the cost of such a small performance improvement was obviously higher dimensionality of the features space of ($231 \times 3 = 693$) compared to 231 features in case of using an individual features subset for classification. We have also combined features pair-wise that led to features space with the dimensionality of 462 in each combination. The combination of GeMs-features and LMs-features, called GeMs/LMs-features, produced the highest classification results 0.943 ± 0.027 (0.95 ± 0.013). This is slightly higher than the performance achieved using GHMs/LMs-features. Consequently, the use of either individual feature subset or combined feature subsets is promising for the detection of abnormalities.

C. RESULTS OF THE PROPOSED CADx SYSTEM

We also applied OMIs features for characterizing the malignancy of masses in mammographic regions. Table 4 presents the results of classifying ROIs in the CADx dataset. This dataset consists of regions depicting 200 masses of which 100 are benign and 100 are malignant. The overall average Az-value obtained using various feature sets show that the performance of the proposed features, for implementing CADx systems (i.e. distinguishing between ROI depicting malignant masses and those regions depicting benign masses), is not as efficient as for the CADE systems

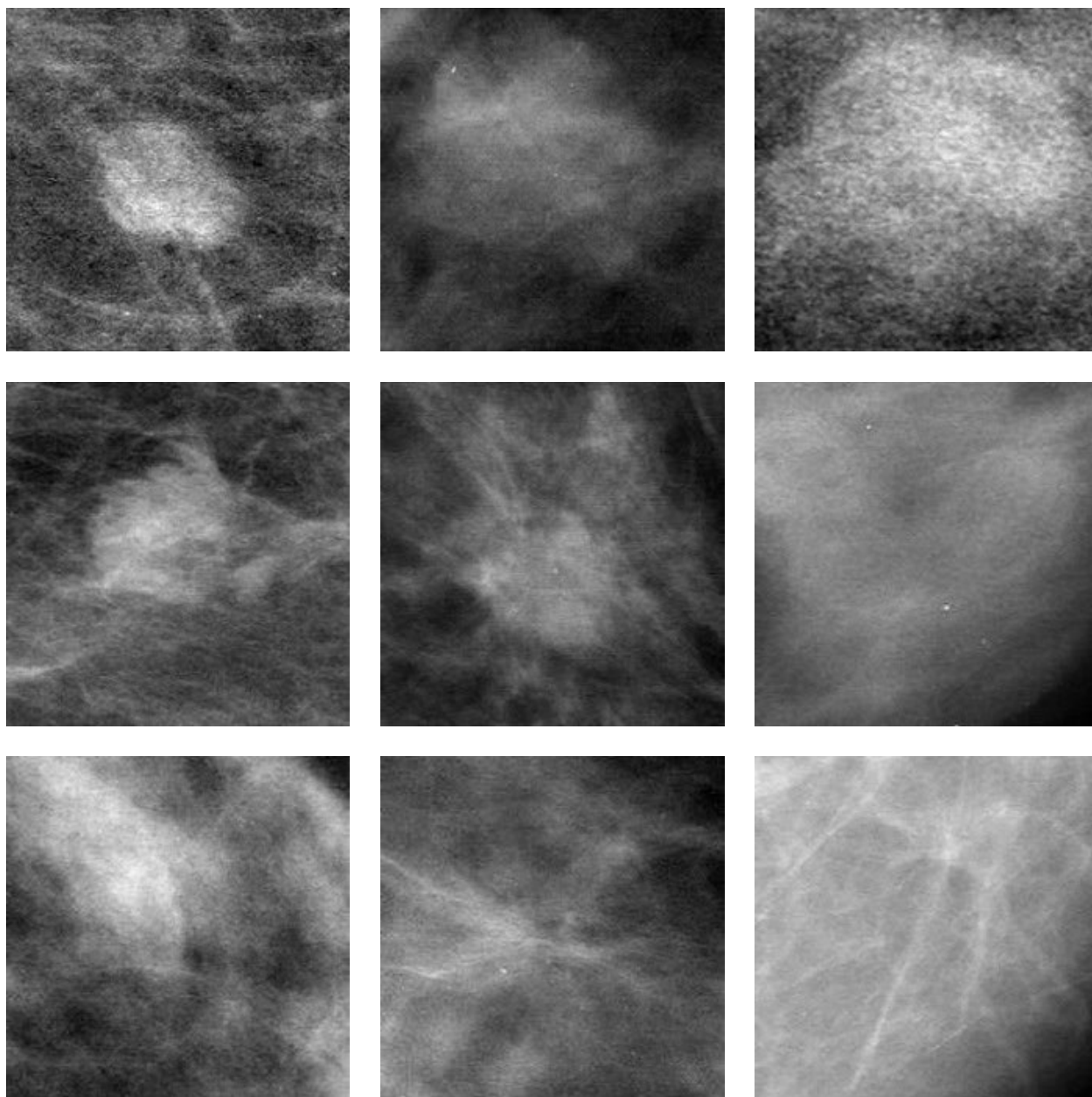


FIGURE 3. Samples regions of the detection and diagnosis datasets; regions with malignant masses (first row), regions with benign masses (second row), and regions related to the normal breast parenchyma (third row).

(i.e. characterizing the presences of masses). Our justification for such a performance of CADx algorithms is that the task of characterizing the malignancy of abnormality, even when done by expert radiologists is much more difficult than accomplishing the detection task. This is mainly due to the obscured and vast variation of the appearance of malignant masses on mammograms.

From the results presented in Table 4, the highest classification performance, in terms of average Az-value, of CADx algorithms using different OMIs features is 0.765 ± 0.065 (0.798 ± 0.052) which was produced by GHMs-features, which slightly outperformed GeMe-features and

LMs-features as well. Combining different features subsets, including pair-wise combinations, was not useful and even provided inferior classification performance. For instance, results of combined features, also included in Table 4, show that the ABC-features with 693 features provided an average classification performance of 0.764 ± 0.054 (0.803 ± 0.040). The obtained classification performance, on the test dataset, was almost equal to that produced using B-features (GHMs) and A-features (GeMe) individually applied.

Moreover, obtained classification results demonstrated that different features sets (including combined features) produced FPF results that is significantly higher than FNF

TABLE 3. Classification results of various feature sets applied to distinguish between abnormal and normal classes(A = GeMs; B = GHMs; C = LMs).

Feature Set	Cross-Validation				Test			
	FNF	FPF	Accuracy	Az	FNF	FPF	Accuracy	Az
A	0.116 ± 0.022	0.107 ± 0.017	0.888 ± 0.012	0.935 ± 0.014	0.149 ± 0.043	0.152 ± 0.034	0.849 ± 0.024	0.929 ± 0.021
B	0.119 ± 0.024	0.109 ± 0.024	0.886 ± 0.018	0.932 ± 0.015	0.153 ± 0.051	0.143 ± 0.050	0.852 ± 0.037	0.926 ± 0.025
C	0.116 ± 0.022	0.106 ± 0.02	0.889 ± 0.014	0.926 ± 0.011	0.145 ± 0.051	0.150 ± 0.046	0.853 ± 0.036	0.918 ± 0.021
AB	0.100 ± 0.027	0.103 ± 0.021	0.899 ± 0.0158	0.948 ± 0.013	0.139 ± 0.043	0.149 ± 0.041	0.856 ± 0.029	0.935 ± 0.022
BC	0.122 ± 0.016	0.105 ± 0.031	0.886 ± 0.017	0.936 ± 0.014	0.161 ± 0.049	0.149 ± 0.03	0.845 ± 0.022	0.927 ± 0.023
AC	0.097 ± 0.020	0.10 ± 0.017	0.902 ± 0.013	0.95 ± 0.013	0.133 ± 0.039	0.140 ± 0.059	0.863 ± 0.038	0.943 ± 0.027
ABC	0.096 ± 0.020	0.099 ± 0.016	0.90 ± 0.0125	0.95 ± 0.0138	0.131 ± 0.048	0.138 ± 0.054	0.865 ± 0.03	0.941 ± 0.018

TABLE 4. Classification results of OMIs feature sets applied to distinguish between malignant and benign classes (A = GeMs; B = GHMs; C = LMs).

Feature Set	Cross-Validation				Test			
	FNF	FPF	Accuracy	Az	FNF	FPF	Accuracy	Az
A	0.158 ± 0.116	0.323 ± 0.123	0.759 ± 0.029	0.784 ± 0.044	0.217 ± 0.184	0.406 ± 0.152	0.688 ± 0.05	0.761 ± 0.055
B	0.173 ± 0.062	0.292 ± 0.097	0.767 ± 0.031	0.798 ± 0.052	0.285 ± 0.104	0.353 ± 0.128	0.681 ± 0.056	0.765 ± 0.065
C	0.105 ± 0.106	0.395 ± 0.134	0.75 ± 0.021	0.763 ± 0.041	0.145 ± 0.146	0.520 ± 0.135	0.668 ± 0.038	0.736 ± 0.032
AB	0.141 ± 0.113	0.324 ± 0.127	0.767 ± 0.023	0.792 ± 0.037	0.239 ± 0.190	0.412 ± 0.167	0.675 ± 0.05	0.758 ± 0.053
BC	0.132 ± 0.089	0.303 ± 0.144	0.783 ± 0.037	0.815 ± 0.05	0.249 ± 0.170	0.439 ± 0.141	0.656 ± 0.044	0.738 ± 0.044
AC	0.154 ± 0.091	0.31636 ± 0.14	0.765 ± 0.036	0.800 ± 0.053	0.250 ± 0.161	0.382 ± 0.118	0.684 ± 0.063	0.758 ± 0.067
ABC	0.144 ± 0.087	0.312 ± 0.104	0.772 ± 0.030	0.803 ± 0.040	0.194 ± 0.143	0.467 ± 0.177	0.67 ± 0.059	0.764 ± 0.054

TABLE 5. Classification results of selected OMIs features applied to the CADe(A = GeMs; B = GHMs; C = LMs).

Feature Set	Number of Features	Cross-validation				Test			
		FNF	FPF	Accuracy	Az	FNF	FPF	Accuracy	Az
A	88.35 ± 10.69	0.070 ± 0.0193	0.070 ± 0.016	0.930 ± 0.011	0.963 ± 0.009	0.126 ± 0.05	0.142 ± 0.055	0.866 ± 0.029	0.933 ± 0.022
B	89.05 ± 16.31	0.073 ± 0.0193	0.067 ± 0.021	0.930 ± 0.014	0.961 ± 0.0131	0.134 ± 0.047	0.154 ± 0.060	0.856 ± 0.027	0.926 ± 0.021
C	86.7 ± 11.842	0.069 ± 0.0167	0.069 ± 0.019	0.931 ± 0.015	0.9561 ± 0.016	0.129 ± 0.042	0.154 ± 0.046	0.859 ± 0.027	0.923 ± 0.022
AB	90.52 ± 9.60	0.070 ± 0.019	0.062 ± 0.02	0.934 ± 0.013	0.965 ± 0.008	0.148 ± 0.053	0.135 ± 0.058	0.858 ± 0.034	0.926 ± 0.027
BC	174.64 ± 23.91	0.079 ± 0.022	0.077 ± 0.016	0.922 ± 0.013	0.952 ± 0.015	0.132 ± 0.035	0.137 ± 0.043	0.866 ± 0.029	0.929 ± 0.023
AC	175 ± 32.40	0.058 ± 0.019	0.071 ± 0.015	0.935 ± 0.013	0.968 ± 0.01	0.121 ± 0.038	0.128 ± 0.058	0.876 ± 0.036	0.945 ± 0.027
ABC	177.72 ± 30.55	0.057 ± 0.017	0.072 ± 0.020	0.935 ± 0.011	0.969 ± 0.01	0.118 ± 0.041	0.127 ± 0.043	0.880 ± 0.030	0.952 ± 0.017

results, which implies that using various CADx system benign masses were misclassified at higher rate than malignant masses.

D. RESULTS OF APPLYING PSO-SVM FEATURE SELECTION

To examine whether applying the feature selection can improve the classification performance of the proposed features, we have adopted the embedded feature selection and classification system based on PSO-SVM algorithm from [7]. Such an approach provided a simultaneous framework for accomplishing the parameters and feature selection tasks

such that both dimensionality reduction of the feature space and higher classification performance can be attained. The classification results, after applying PSO-SVM for parameters and feature selection, produced from individual and combined OMIs feature sets are presented in Tables 5 and 6.

For distinguishing between abnormal and normal regions, as shown by results in Table 5, using the feature selection based on PSO-SVM, we were not only able to significantly reduce the dimensionality of the feature space but also we have improved the classification performance. For instance, using individual feature sets, the highest

TABLE 6. Classification results of selected OMI features applied to the CADx or detection dataset (A = GeMs; B = GHMs; C = LMs).

Feature Set	Number of Features	Cross-Validation				Test			
		FNF	FPF	Accuracy	Az	FNF	FPF	Accuracy	Az
A	90.2 ± 13.34	0.143 ± 0.063	0.120 ± 0.108	0.829 ± 0.042	0.856 ± 0.053	0.278 ± 0.124	0.350 ± 0.123	0.686 ± 0.053	0.736 ± 0.066
B	87.22 ± 16.011	0.150 ± 0.06	0.221 ± 0.108	0.815 ± 0.044	0.843 ± 0.05	0.314 ± 0.130	0.348 ± 0.130	0.669 ± 0.055	0.733 ± 0.069
C	90.82 ± 16.634	0.168 ± 0.082	0.242 ± 0.128	0.795 ± 0.039	0.811 ± 0.055	0.279 ± 0.150	0.412 ± 0.169	0.65 ± 0.061	0.70 ± 0.074
AB	169.96 ± 30.3925	0.157 ± 0.063	0.207 ± 0.104	0.818 ± 0.037	0.841 ± 0.048	0.319 ± 0.133	0.345 ± 0.124	0.669 ± 0.049	0.735 ± 0.057
BC	177.4 ± 32.405	0.142 ± 0.073	0.242 ± 0.1261	0.808 ± 0.046	0.838 ± 0.056	0.299 ± 0.157	0.379 ± 0.138	0.661 ± 0.059	0.738 ± 0.073
AC	176.48 ± 29.893	0.150 ± 0.052	0.223 ± 0.104	0.813 ± 0.042	0.848 ± 0.0442	0.282 ± 0.144	0.355 ± 0.137	0.681 ± 0.054	0.736 ± 0.058
ABC	275.2 ± 45.473	0.138 ± 0.073	0.290 ± 0.101	0.786 ± 0.0347	0.817 ± 0.046	0.229 ± 0.133	0.417 ± 0.131	0.677 ± 0.054	0.777 ± 0.050

TABLE 7. The comparison between the proposed system and the other related work.

Reference	Dataset	No. of Images	No. of Features	Feature Selection Method	Classifier	Accuracy		Az	
						CADe	CADx	CADe	CADx
Eltoukhy et al. [6]	MIAS	322	N/A	Diff. of Means	SVM	91.19	91.18	N/A	N/A
Dhabhi et al. [8]	MIAS	252	10	PCA	KNN	91.27	N/A	97.3	84.0
Ramos et al. [13]	DDSM	120	670	Genetic Algorithm	Random Forest	90.0	N/A	N/A	N/A
Rouhi et al. [14]	DDSM	170	51	Genetic Algorithm	MLP	N/A	86.66	N/A	88.15
Chen et al. [15]	private	275	59	PSO	SVM	N/A	N/A	N/A	0.81±0.03
Sasikala et al. [19]	Inbreast	N/A	N/A	N/A	SVM	N/A	86.8	N/A	N/A
Wang et al. [23]	Private	400	37	N/A	SVM	N/A	83.75	N/A	N/A
Patil & Birada [24]	Private	N/A	N/A	N/A	CRNN	90.59	N/A	N/A	N/A
Eltoukhy et al. [25]	MIAS	265	231	N/A	AdaBoost	90.56	N/A	88.5	N/A
Rabidas & Arif [40]	DDSM	500	63	N/A	FLDA	N/A	80.72±0.01	N/A	89±0.01
Proposed Method	DDSM	400/200*	693	PSO	SVM	93.5±0.011	82.9±0.042	96.9±0.01	85.6±0.053

classification results of 0.933 ± 0.022 was produced from GeMe-features with the average size of the optimized feature space of 88.35 ± 10.69 features (original feature space is 231). When combined feature sets were used for classification, the fused set of features provided the best classification performance of 0.969 ± 0.01 with the feature space reduced from 693 to an average of 177.72 ± 30.55 features, which is almost 25% of the original features set.

For classifying regions with benign or malignant masses with the PSO-SVM approach used for selecting the most relevant OMI features and for performance optimization, results from Table 6 show that applying the PSO-SVM optimization, indeed, reduced the dimensionality but, unexpectedly, degraded the classification performance provided by most features sets. The highest classification performance of 0.817 ± 0.04 , however, produced by ABC-features that slightly outperformed previous performance attained without applying the feature selection step. The dimensionality of the feature space, in case of using all features, was reduced from 693 to an average of 275.2 ± 45.473 features, which is almost 39.6% of the total features set.

The high dimensionality of the feature space, namely, the fused space, presented a practical and main computational challenge as it is expected to require larger population swarm size or higher number of iterations to enable the PSO-SVM escape from the local or suboptimal solution. The next subsection presents a comparison study of the proposed system for solving the detection (CADe) and diagnosis (CADx) tasks of masses in mammographic images against state of the art.

E. COMPARISON STUDY

This work proposed to combine the advantages of orthogonal moments features with the PSO-SVM to achieve high performance CAD system. The results of the proposed detection and diagnosis system, as shown in Table 7, are promising and competitive when compared with the state of the art. Moreover, the PSO based SVM classifier was not only able to select the appropriate input features but also optimized the SVM performance. However, the straight comparison of the proposed system results with those from existing systems cannot be direct because of differences, among systems, in terms of datasets used for the training and validation of the

algorithm. The community of mammography CAD research mostly used two common and public datasets of screen mammography: the DDSM [28] and MIAS [26] datasets. Another difference is the performance metrics used for the evaluation. For this, we reported our results using both A_z -values of the ROC curve and accuracies. In addition to the mammography data. Other differences as shown in Table 7, however, including the machine learning algorithms used for classification and feature selection approach, including the PCA and genetic algorithms.

Considering the key difference in the literature of mammographic mass CAD systems, namely, the validation datasets and for providing meaningful and fair comparison, the focus will be on those studies [13], [14], [40] used DDSM dataset for the evaluation of the CAD systems. For accomplished mammography detection task, the proposed work, by combining PSO-SVM model selection and orthogonal moments had outperformed, in terms of the classification accuracy or the A_z -value, existing systems [13], [40]. For instance, the proposed CADe system provided an average accuracy of 93.5 ± 0.01 that higher than 90 achieved from combining GA and random forest classifier [13]. As for accomplishing the CADx task, the proposed system is competing very well with FLDA approach from [40] but inferior to the performance obtained from the GA and MLP system [14].

VI. CONCLUSION

This paper proposed a CAD system combining orthogonal moments for feature extraction and the PSO-SVM for hyperparameter and parameter optimization to accomplish the two CAD tasks. The first task classified regions of interest (ROIs) of digitized mammograms into mass and normal breast tissue regions. The second task focused on characterized the malignancy of mammographic masses. This paper focused on the examining whether the use of OMIs features is adequate or not. Three sets of OMIs features were extracted from mammographic ROI dataset with 200 masses. On the CADe or detection's dataset, the combined OMIs features achieved an average A_z value of 0.969 ± 0.01 . As for the application of OMIs features for solving the CADx task, the classification performance was good but not as high as for the detection task. However, obtained results show that the potential application of OMIs features for solving the feature extraction task in CADe and CADx systems.

REFERENCES

- [1] (Nov. 15 2019). WHO. [Online]. Available: <http://www.who.int/cancer/prevention/diagnosis-screening/breast-cancer/en/>
- [2] M. Helvie, "Improving mammographic interpretation: Double reading and computer-aided diagnosis," *Radiologic Clinics North Amer.*, vol. 45, no. 5, pp. 801–811, Sep. 2007.
- [3] M. M. Eltoukhy, I. Faye, and B. B. Samir, "A comparison of wavelet and curvelet for breast cancer diagnosis in digital mammogram," *Comput. Biol. Med.*, vol. 40, no. 4, pp. 384–391, Apr. 2010.
- [4] M. M. Eltoukhy, I. Faye, and B. B. Samir, "A statistical based feature extraction method for breast cancer diagnosis in digital mammogram using multiresolution representation," *Comput. Biol. Med.*, vol. 42, no. 1, pp. 123–128, Jan. 2012.
- [5] M. M. Eltoukhy and I. Faye, "An adaptive threshold method for mass detection in mammographic images," in *Proc. IEEE Int. Conf. Signal Image Process. Appl.*, Melaka, Oct. 2013, pp. 374–378, doi: 10.1109/ICSIPA.2013.6708036.
- [6] M. M. Eltoukhy and I. Faye, "An optimized feature selection method for breast cancer diagnosis in digital mammogram using multiresolution representation," *Appl. Math. Inf. Sci.*, vol. 8, no. 6, pp. 2921–2928, Nov. 2014.
- [7] I. Zyouf, J. Czajkowska, and M. Grzegorzek, "Multi-scale textural feature extraction and particle swarm optimization based model selection for false positive reduction in mammography," *Computerized Med. Imag. Graph.*, vol. 46, pp. 95–107, Dec. 2015.
- [8] S. Dhabbi, W. Barhoumi, and E. Zagrouba, "Breast cancer diagnosis in digitized mammograms using curvelet moments," *Comput. Biol. Med.*, vol. 64, pp. 79–90, Sep. 2015.
- [9] S. J. S. Gardezi, I. Faye, and M. M. Eltoukhy, "Analysis of mammogram images based on texture features of curvelet sub-bands," in *Proc. 5th Int. Conf. Graphic Image Process. (ICGIP)*, Hong Kong, vol. 9069, Jan. 2014, p. 906924, doi: 10.1117/12.2054183.
- [10] M. Jiang, S. Zhang, H. Li, and D. N. Metaxas, "Computer-aided diagnosis of mammographic masses using scalable image retrieval," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 2, pp. 783–792, Feb. 2015, doi: 10.1109/TBME.2014.2365494.
- [11] Y. Bazi and F. Melgani, "Semisupervised PSO-SVM regression for biophysical parameter estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1887–1895, Jun. 2007, doi: 10.1109/TGRS.2007.895845.
- [12] J.-Y. Yeh and S. Chan, "Population-based metaheuristic approaches for feature selection on mammograms," in *Proc. IEEE Int. Conf. Agents (ICA)*, Beijing, China, Jul. 2017, pp. 140–144, doi: 10.1109/AGENTS.2017.8015321.
- [13] R. P. Ramos, M. Z. D. Nascimento, and D. C. Pereira, "Texture extraction: An evaluation of ridgelet, wavelet and co-occurrence based methods applied to mammograms," *Expert Syst. Appl.*, vol. 39, no. 12, pp. 11036–11047, Sep. 2012.
- [14] R. Rouhi, M. Jafari, S. Kasaei, and P. Keshavarzian, "Benign and malignant breast tumors classification based on region growing and CNN segmentation," *Expert Syst. Appl.*, vol. 42, no. 3, pp. 990–1002, Feb. 2015.
- [15] X. Chen, A. Zargari, A. B. Hollingsworth, H. Liu, B. Zheng, and Y. Qiu, "Applying a new quantitative image analysis scheme based on global mammographic features to assist diagnosis of breast cancer," *Comput. Methods Programs Biomed.*, vol. 179, Oct. 2019, Art. no. 104995, doi: 10.1016/j.cmpb.2019.104995.
- [16] K. B. Soulam, N. Kaabouch, M. N. Saidi, and A. Tamtaoui, "An evaluation and ranking of evolutionary algorithms in segmenting abnormal masses in digital mammograms," *Multimedia Tools Appl.*, pp. 1–39, Mar. 2020, doi: 10.1007/s11042-019-08449-5.
- [17] B. S. Khehra and A. P. S. Pharaoh, "Comparison of genetic algorithm, particle swarm optimization and biogeography-based optimization for feature selection to classify clusters of microcalcifications," *J. Inst. Eng. (India) B*, vol. 98, no. 2, pp. 189–202, Apr. 2017, doi: 10.1007/s40031-016-0226-8.
- [18] S. Sasikala and M. Ezhilarasi, "Fusion of two view binary patterns to improve the performance of breast cancer diagnosis," in *Proc. Int. Conf. Commun. Signal Process. (ICCCSP)*, Chennai, India, Apr. 2017, pp. 792–796, doi: 10.1109/ICCCSP.2017.8286473.
- [19] S. Sasikala, M. Ezhilarasi, and S. A. Kumar, "Detection of breast cancer using fusion of MLO and CC view features through a hybrid technique based on binary firefly algorithm and optimum-path forest classifier," in *Applied Nature-Inspired Computing: Algorithms and Case Studies* (Springer Tracts in Nature-Inspired Computing), N. Dey, A. Ashour, and S. Bhattacharyya, Eds. Singapore: Springer, 2020.
- [20] S. Sasikala, M. Bharathi, M. Ezhilarasi, M. R. Reddy, and S. Arunkumar, "Fusion of MLO and CC view binary patterns to improve the performance of breast cancer diagnosis," *Current Med. Imag. Rev.*, vol. 14, no. 4, pp. 651–658, Jul. 2018, doi: 10.2174/1573405614666180104162408.
- [21] S. Sasikala and M. Ezhilarasi, "Fusion of k-Gabor features from medio-lateral-oblique and cranio-caudal view mammograms for improved breast cancer diagnosis," *J. Cancer Res. Therapeutics*, vol. 14, no. 5, pp. 1036–1041, 2018, doi: 10.4103/jcrt.JCRT_1352_16.
- [22] S. Sasikala and M. Ezhilarasi, "Comparative analysis of serial and parallel fusion on texture features for improved breast cancer diagnosis," *Current Med. Imag. Rev.*, vol. 14, no. 6, pp. 957–968, Oct. 2018, doi: 10.2174/1573405613666170926164625.

- [23] Z. Wang, M. Li, H. Wang, H. Jiang, Y. Yao, H. Zhang, and J. Xin, "Breast cancer detection using extreme learning machine based on feature fusion with CNN deep features," *IEEE Access*, vol. 7, pp. 105146–105158, 2019, doi: [10.1109/ACCESS.2019.2892795](https://doi.org/10.1109/ACCESS.2019.2892795).
- [24] R. S. Patil and N. Biradar, "Automated mammogram breast cancer detection using the optimized combination of convolutional and recurrent neural network," *Evol. Intell.*, pp. 1–16, Apr. 2020, doi: [10.1007/s12065-020-00403-x](https://doi.org/10.1007/s12065-020-00403-x).
- [25] M. M. Eltoukhy, M. Elhoseny, K. M. Hosny, and A. K. Singh, "Computer aided detection of mammographic mass using exact Gaussian–Hermite moments," *J. Ambient Intell. Humanized Comput.*, vol. 2018, pp. 1–9, Jun. 2018, doi: [10.1007/s12652-018-0905-1](https://doi.org/10.1007/s12652-018-0905-1).
- [26] (Sep. 2019). *MIAS*. [Online]. Available: <http://peipa.essex.ac.uk/info/mias.html>
- [27] J. E. E. Oliveira, M. O. Gueld, A. de A. Araújo, B. Ott, and T. M. Deserno, "Toward a standard reference database for computer-aided mammography," in *Proc. SPIE*, vol. 6915. Bellingham, WA, USA: SPIE, 2008, doi: [10.1117/12.770325](https://doi.org/10.1117/12.770325).
- [28] M. Heath, K. Bowyer, D. Kopans, W. P. Kegelmeyer, R. Moore, K. Chang, and S. MunishKumaran, "Current status of the digital database for screening mammography," in *Digital Mammography*. Norwell, MA, USA: Kluwer, 1998, pp. 457–460.
- [29] M. Pawlak, *Image Analysis by Moments: Reconstruction and Computational Aspects*. Wrocław, Poland: Oficyna Wydawnicza Politechniki Wrocławskiej, 2006.
- [30] K. M. Hosny, "Exact legendre moment computation for gray level images," *Pattern Recognit.*, vol. 40, no. 12, pp. 3597–3605, Dec. 2007, doi: [10.1016/j.patcog.2007.04.014](https://doi.org/10.1016/j.patcog.2007.04.014).
- [31] K. M. Hosny, "Image representation using accurate orthogonal gegenbauer moments," *Pattern Recognit. Lett.*, vol. 32, no. 6, pp. 795–804, Apr. 2011, doi: [10.1016/j.patrec.2011.01.006](https://doi.org/10.1016/j.patrec.2011.01.006).
- [32] K. M. Hosny, "Fast computation of accurate Gaussian–Hermite moments for image processing applications," *Digit. Signal Process.*, vol. 22, no. 3, pp. 476–485, May 2012, doi: [10.1016/j.dsp.2012.01.002](https://doi.org/10.1016/j.dsp.2012.01.002).
- [33] K. M. Hosny, "Refined translation and scale legendre moment invariants," *Pattern Recognit. Lett.*, vol. 31, no. 7, pp. 533–538, May 2010, doi: [10.1016/j.patrec.2009.12.008](https://doi.org/10.1016/j.patrec.2009.12.008).
- [34] M. K. Hosny, "New set of rotationally legendre moment invariants," *Int. J. Electr. Comput. Syst. Eng.*, vol. 4, no. 3, pp. 176–180, 2010.
- [35] K. M. Hosny, "New set of gegenbauer moment invariants for pattern recognition applications," *Arabian J. Sci. Eng.*, vol. 39, no. 10, pp. 7097–7107, Oct. 2014, doi: [10.1007/s13369-014-1336-8](https://doi.org/10.1007/s13369-014-1336-8).
- [36] B. Yang, J. Flusser, and T. Suk, "Design of high-order rotation invariants from Gaussian–Hermite moments," *Signal Process.*, vol. 113, pp. 61–67, Aug. 2015, doi: [10.1016/j.sigpro.2015.01.005](https://doi.org/10.1016/j.sigpro.2015.01.005).
- [37] K. M. Hosny, "Exact and fast computation of geometric moments for gray level images," *Appl. Math. Comput.*, vol. 189, no. 2, pp. 1214–1222, Jun. 2007, doi: [10.1016/j.amc.2006.12.025](https://doi.org/10.1016/j.amc.2006.12.025).
- [38] M.-Y. Cho and T. T. Hoang, "Feature selection and parameters optimization of SVM using particle swarm optimization for fault classification in power distribution systems," *Comput. Intell. Neurosci.*, vol. 2017, pp. 1–9, 2017, doi: [10.1155/2017/4135465](https://doi.org/10.1155/2017/4135465).
- [39] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw. (ICNN)*, Perth, WA, Australia, vol. 4, Nov./Dec. 1995, pp. 1942–1948.
- [40] R. Rabidas and W. Arif, "Characterization of mammographic masses based on local photometric attributes," *Multimedia Tools Appl.*, pp. 1–19, May 2020, doi: [10.1007/s11042-020-08959-7](https://doi.org/10.1007/s11042-020-08959-7).



MOHAMED W. ABO EL-SOUD (Member, IEEE) received the Ph.D. degree in network security from TUHH University, Hamburg, Germany, and in channel system from Suez Canal University, Ismailia, Egypt, in 2010. He is currently an Assistant Professor with the Information System Department, Faculty of Computers and Informatics, Suez Canal University, and the Department of Computer Science and Information, College of Science, Majmaah University, Zulfi, Saudi Arabia.

He has published in several peer-reviewed reputed journals and conference proceedings. His research interests include network security, medical image analysis, and artificial intelligence. He is a Reviewer for several international journals.



IMAD ZYOUT (Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from the Jordan University of Science and Technology, Jordan, in 1999 and 2005, respectively, and the Ph.D. degree in electrical and computer engineering from Western Michigan University, USA, in 2010. He was a Visiting Scholar at the Pattern Recognition Research Group, Siegen University, Germany. He is currently an Associate Professor in electrical engineering at Tafila Technical University, Jordan.

His research experience and interests include image compression and processing, feature engineering, heuristic feature search and selection, and classical and modern machine learning algorithms focused on the development of efficient algorithms for the pattern recognition and analysis of breast cancer in mammography. He is a member of the Jordan Engineers Association (JEA). He was a recipient of the Australian Endeavour Post-Doctorate Fellowship, in 2013, which allowed him to join Signal and Information Processing Research Laboratory, University of Western Australia (UWA), in 2014.



KHALID M. HOSNY (Senior Member, IEEE) was born in Zagazig, Egypt, in 1966. He received the B.Sc., M.Sc., and Ph.D. degrees from Zagazig University, Egypt, in 1988, 1994, and 2000, respectively. From 1997 to 1999, he was a Visiting Scholar with the University of Michigan, Ann Arbor, and with the University of Cincinnati, Cincinnati, USA. He is currently a Professor of information technology with the Faculty of Computers and Informatics, Zagazig University. He has

published three edited books and more than 70 articles in international journals. His research interests include image processing, pattern recognition, multimedia, and computer vision. He is a Senior Member of ACM. He is an Editor and a Scientific Reviewer for more than 35 international journals.



MOHAMED MESELHY ELTOUKHY (Member, IEEE) received the Ph.D. degree in electrical and electronic engineering from Universiti Teknologi PETRONAS, Malaysia, in 2011. He is currently an Associate Professor with the Computer Science Department, Faculty of Computers and Informatics, Suez Canal University, Egypt. He is also an Associate Professor with the College of Computing and Information Technology, Khulais, University of Jeddah, Saudi Arabia. He has published

in several peer-reviewed reputed journals and conference proceedings. His research interests include artificial intelligence, medical image analysis, cancer detection, feature extraction and selection, and multiresolution representation. In 2012, he received a Postdoctoral Fellowship from Universiti Teknologi PETRONAS. He is a Reviewer for several international journals.

...