

# Multi-Images Recognition of Breast Cancer Histopathological via Probabilistic Neural Network Approach

Heba M. Afify<sup>1</sup>, Kamel. K. Mohammed<sup>2</sup>, Aboul Ella Hassanien<sup>3,4</sup>

<sup>1</sup> Electrical Engineering Department, Egyptian Academy for Engineering and  
Advanced Technology, Cairo, Egypt

<sup>2</sup> Center for Studies and Research Viral, Al Azhar University, Cairo, Egypt

<sup>3</sup> Faculty of Computers and Information, Cairo University, Giza, Egypt

<sup>4</sup> Scientific Research Group in Egypt (SRGE), Cairo, Egypt

*hebaaffify@yahoo.com; tawfickamel@gmail.com; aboitcairo@gmail.com*

**Abstract.** This paper suggests an approach for automated diagnosis of different types of breast carcinoma histopathological images via machine learning algorithms. The approach starts by extracting some features then use the nondominated genetic algorithm II (NSGA-II) algorithm and Ant Colony Optimization (ACO) to select the best features. Additionally, two classifiers including probabilistic neural network (PNN) and multi-class support vector machine (MSVM) is proposed to determine the type of breast cancer. These histopathological images database consider as a multi-classification task that used 848 BreakHis images with 400× magnification factors. The data divided into 160 samples for the testing phase and 688 samples for the training phase. The experiment results evaluated by four parameters including accuracy, sensitivity, specificity, and precision. For the NSGA-II algorithm, these outcomes refer to the maximum accuracy obtained through the combined features set with the accuracy being 100% by the training dataset and 82.5% by the testing dataset for the PNN classifier.

**Keywords:** BreakHis images, NSGA-II, ACO, PNN, MSVM.

## 1. Introduction

The hazard of breast cancer is associated with age, diagnosis stage, race, gene mutations, and cancer type (Howlader *et al.*, 2019). According to the statistics by ductal carcinoma in situ (DCIS), it is reported that 41,760 women will die from this disease (Howlader *et al.*, 2019). The awareness of breast cancer has spurred attention in advanced diagnosis programs and medical imaging problems to support radiologists. Although the development of diagnostic imaging devices, but the radiology technologies of images are insufficient for early diagnosis and useful treatment (Aswathy and Jagannath, 2017). The sensitivity and specificity rate of digital mammography in the U.S. is 86.9% and 88.9%<sup>3</sup> respectively (Lehman *et al.*, 2016). Therefore, the microscopic images of breast tissue are presented with more accurate information for the discovery of this disease by using the biopsy techniques and pathologists' experience (Burt *et al.*, 2018). Generally, the issues specific to histopathological image analysis are huge image size, unlabeled images, color variation, and magnification level. To address these problems, the sorting of histopathology images of breast cancer are supported by the Computer-Assisted Detection/Diagnosis (CAD/CADx) program (Filipczuk *et al.*, 2018). The high computational complexity to recognize the different types of breast images is solved by advanced image analysis and the right selection of feature extraction via machine learning algorithms (Wu *et al.*, 2015). Kowel *et al.* applied nuclei related features to classify breast cancer images with an accuracy between 84 to 93%. While Belsare *et al.* applied statistical texture features on breast histology images with an accuracy between 70% to 100%. Cruz-Roa *et al.* applied the nuclei related features and overall tissue-related features for breast cancer recognition with an accuracy of 78%. George *et al.* proposed the support vector machine (SVM) and artificial neural network (ANN) on cytological images for breast cancer detection. Another research is focused on the SVM algorithm to analyze histopathological images and determine the position of the ROI in images (BenTaieb *et al.*, 2017). The random framework classifier is used for distinguishing microscopic images to obtain a high accuracy (Zhang *et al.*, 2013). The experimental results informed that the SVM algorithm produced the highest accuracy of 97.13% (Asri *et al.*, 2016).

Many previous studies are suffered from Whole-Slide Imaging and small datasets that lead to poor results especially in multi-class classification (Han *et al.*, 2017). Consequently, Spanhol *et al.* used the BreakHis database on K-Nearest Neighbor, Quadratic Discriminant Analysis, Random Forest, and SVM to achieve the accuracy of 85%. Nawaz *et al.* suggested a pattern for recognizing the multi-class breast cancer images with an accuracy of 95.4%. Recently, deep learning algorithms are implemented on BreakHis images to enhance the classification accuracy based on the Inception\_V3 and Inception\_ResNet\_V2 network (Xie *et al.*, 2019). In this paper, the feature extraction performed on BreakHis images by using forty-six textural features. After the feature extraction step, we utilized the feature selection to obtain the optimal features by using non-dominated sorting genetic

algorithm II (NSGA-II) as well as the classification algorithms by probabilistic neural network (PNN) and multi-SVM classifiers. The content of the paper divided as follows: materials and methods, proposed methodology, and comparative results.

## 2. Materials and Methods

### 2.1. Data Description

Breast cancer histopathological Database (BreakHis) (Spanhol *et al.*, 2016) consists of 7909 microscopic images divided into eight classes including four classes for benign and four classes for malignant. In this work, we selected 848 images with 400× magnification factors. The selected data consists of 106 digital images from each class. The four kinds of benign cancer are Adenosis, Fibroadenoma, Tubular Adenoma, and Phyllodes Tumor while the four kinds of malignant cancer are Ductal Carcinoma, Lobular Carcinoma, Mucinous Carcinoma, and Papillary Carcinoma (Spanhol *et al.*, 2017). This data gathered from 82 patients with 400× amplification variables. Some samples of BreakHis database with 400× magnification factors are presented in Fig.1. Each pathological image represented in RGB with a size of 700 × 460.

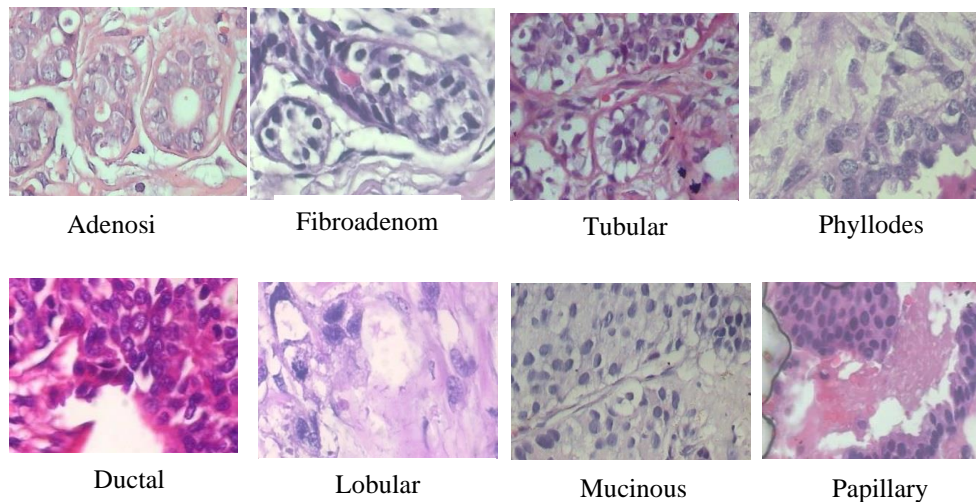


Fig.1: Samples of BreakHis database when the magnification factor of these images is 400×

### 2.2. NSGA-II algorithm

Generally, the NSGA-II algorithm (Deb *et al.*, 2002) is used for handling the optimization problem that depended on the selection of the optimal features for the identification phase. This algorithm is based on a fast ranking process by the computation of the crowding distance and crowding selection operator. This algorithm produces sets of features by utilizing genetic factors, crossover and mutation phases.

### 2.3. Ant Colony Optimization (ACO) algorithm

Ant colony optimization (ACO) is introduced as a novel nature-inspired metaheuristic for solving the hard combinatorial optimization (CO) problems (Dorigo *et al.*, 2005). This algorithm is based on the creation of several ants as shown in Fig.2. Each ant begins by path construction at random feature via graph to reach to the best feature.

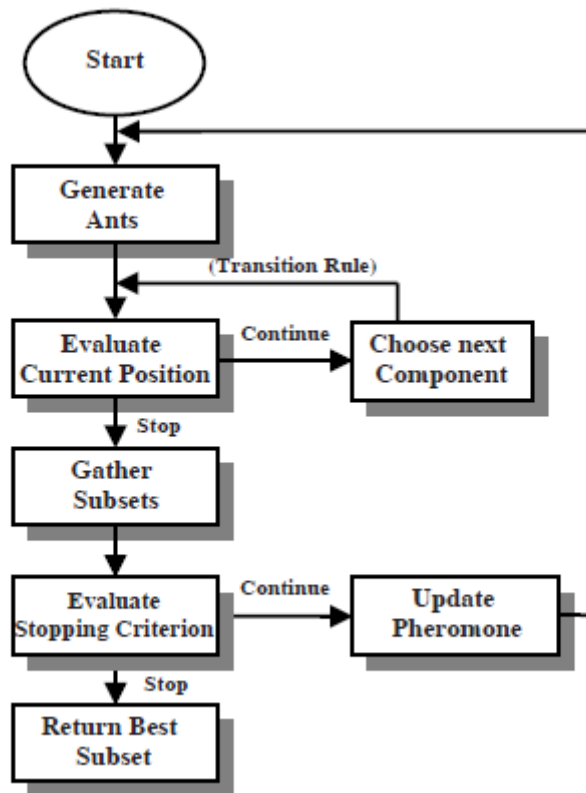


Fig.2: Ant Colony Optimization (ACO)

### 2.4. Probabilistic neural network (PNN)

This algorithm (Azar and S. El-Said, 2012) is a feed-forward neural network based on the probability density function and bayesian classifier technique. The advantages of PNN are faster and more accurate classifier than back-propagation. This algorithm consists of four layers including input, pattern, classes, and output as in Fig. 3.

### 2.5. Multi-class support vector machine (MSVM)

This algorithm is used to avoid errors when handling the classification problem of imbalanced training samples. It takes more computational time than the traditional SVM algorithm because of large samples of classes. While traditional SVM

algorithm is used for classification between two classes only. This algorithm is based on kernel functions for a nonlinear classifier. This algorithm generates sets of traditional SVM algorithms that composed to form the MSVM as in Fig.4.

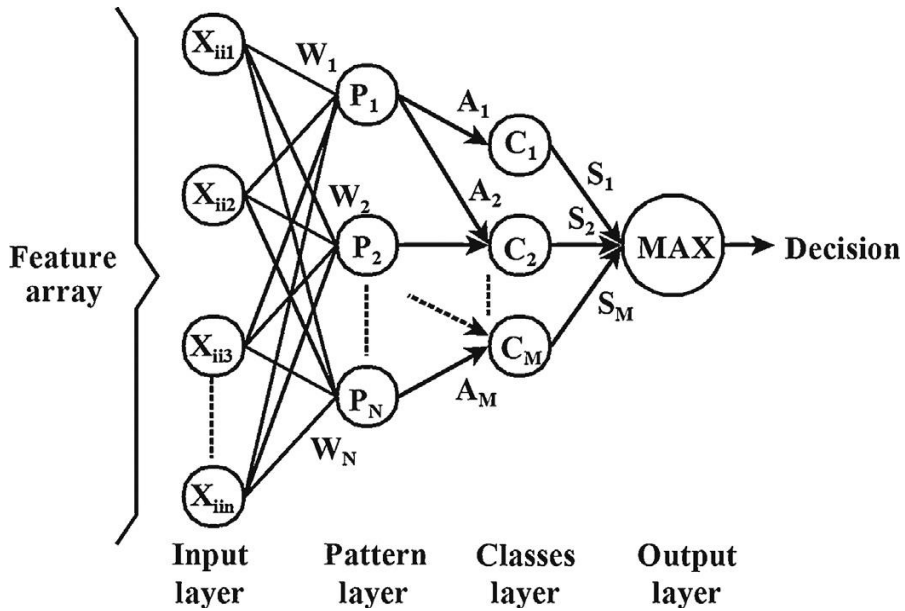


Fig.3: Probabilistic neural network

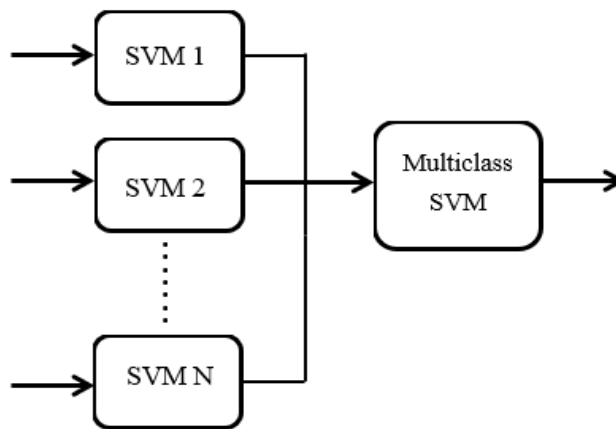


Fig.4: Multi-class SVM

### 3. Proposed Methodology

The steps of the suggested method are summarized in Fig.5. The histopathological image classification is based on adequate feature extraction. The recommended algorithm divided into four phases. The first phase is statistical features extraction that divided into six features for intensity histogram (Suematsu *et al.*, 2002), twenty

two features for gray-level co-occurrence matrix (GLCM) (Vasantha *et al.*, 2010), eleven features for gray-level run-length matrix (GLRLM) (Selvarajah and Kodituwakku, 2011), and seven features for invariant moments (Hu, 1962). The second phase is feature selection that used to reduce the classification errors. Then, the selection features are passed through the classification phase and classification outputs are assessed by evaluation parameters like accuracy, sensitivity, specificity, and precision.

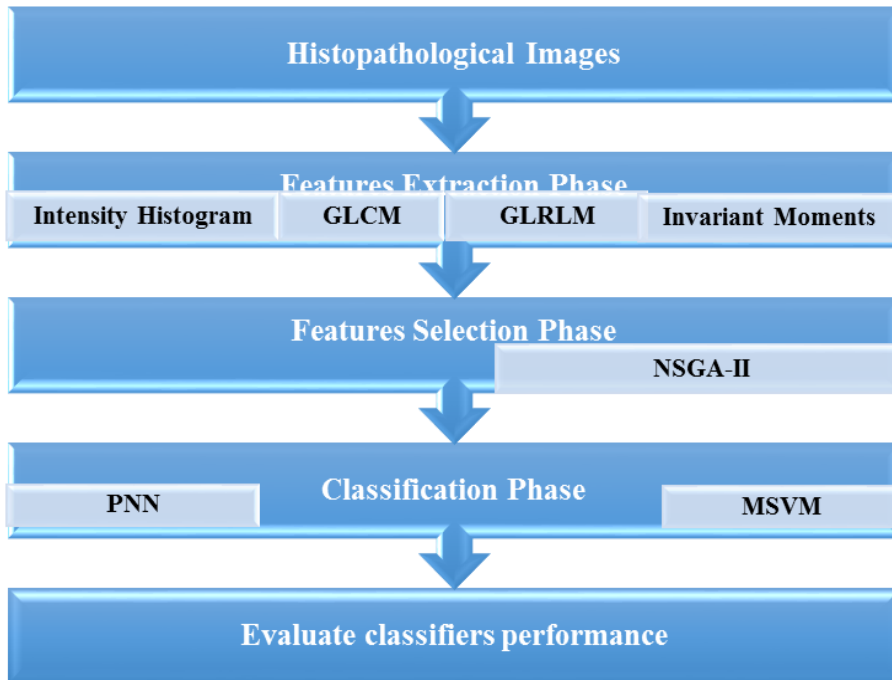


Fig.5: The proposed method for classification of BreakHis images

### 3.1. Feature Extraction Phase

For data implementation, we generated four types of feature extraction including intensity histogram, GLCM, GLRLM, and invariant moments. Therefore, the total number of feature extraction was forty-six textural features. The different types of features that conducted on original BreakHis images are combined for shape recognition and thus for learning the characteristic of pathological images.

The issue of feature selection for multi-objectives and the larger set of features extraction is solved by applying the NSGA-II algorithm (Tarek *et al.*, 2007). The NSGA-II algorithm is critical for improving the recognition rate and minimize computational time (Khan and Baig, 2015). This algorithm is based on a random sorting method to gain the best ranking of features used for multi-class breast cancer recognition problems. In this paper, the 12 selected features using NSGA-II algorithm are skewness, gray-level nonuniformity, low gray-level run emphasis, run

percentage, short-run high gray-level emphasis, cluster shade, cluster prominence, sum average, information measure I of correlation, Information measure II of correlation, and two invariant moments. By using the ACO algorithm, the four selected features are used in the classification phase.

### **3.2. Classification Phase**

In this paper, the features selection act as an input to classification algorithms by using PNN (Azar and El-Said, 2012) and multi-class SVM (Weston and Watkins, 1999) classifiers. The classification step is used to differentiate the eight types of breast cancer images. We assumed 80% for the training samples while 20% for the testing samples. The testing samples are applied to supply an independent test of the system efficiency during and after training.

## **4. Results and Discussion**

In this paper, the suggested approach is developed by MATLAB 2017b environment in a standalone personal computer using Intel i7 3770 processor @3.40 GHz with 8 GB RAM and 64-bit Windows 10 operating system. The proposed algorithm of breast cancer histopathological detection was trained and tested based on 848 microscopic images having four kinds of benign cancer and four kinds of malignant cancer and they were collected from the BreakHis database. The presented CAD method contains two major phases are feature extraction from breast cancer image and classification of breast cancer images and the performance of the classifier is measured by accuracy rate. In the first phase, the breast cancer image is further subjected to 46 features. Four different texture sets (a total of 46 features) were produced from the automatic breast cancer images. In the second phase, the traditional classifiers (i.e., MSVM and PNN) based on the selection of optimal features. The major target of utilizing the feature selection technique is to minimize the large set of features for identifying the different classes of breast cancer histopathological images. Twelve features were selected from multi-class feature selection by NSGA II while four features were selected from multi-class feature selection by ACO. Finally, classification results. The classification of breast cancer images was done by MSVM and PNN classifiers. The confusion matrix is applied for evaluating the performance of the MSVM and PNN classifier by using NSGA II and ACO. For the NSGA II algorithm, Figs.6-7 explain the performance of the PNN classifier and Figs.8-9 for the MSVM classifier. For the ACO algorithm, Figs.10-11 explain the performance of the PPN classifier and Figs.12-13 for the MSVM classifier. Total 848 image samples of breast cancer histopathological images having Adenosis, Fibroadenoma, Tubular Adenoma, Phyllodes Tumor, Ductal Carcinoma, Lobular Carcinoma, Mucinous Carcinoma, and Papillary Carcinoma. Among 848 samples, 688 samples are utilized for training phase which composed of 86 samples of Adenosis, 86 samples of Fibroadenoma, 86

samples of Tubular Adenoma, 86 samples of Phyllodes Tumor, 86 samples of Ductal Carcinoma 86, samples of Lobular Carcinoma, 86 samples of Mucinous Carcinoma and 86 samples Papillary Carcinoma. Outcomes of the training data of the MSVM represent that selected features with multi-objective feature selection with NSGA II yield an overall accuracy of 88.1%, sensitivity of 88.74%, specificity of 98.32%. The number of the input images filled in the MSVM was 688 samples. For example, 81 samples of 86 samples of Adenosis are correctly classified as Adenosis breast cancer histopathological. This corresponds to 95% of all 86 samples of Adenosis breast cancer histopathological. Similarly, 63 samples of 86 samples of Fibroadenoma are correctly classified as Fibroadenoma breast cancer histopathological. This corresponds to 73% of all 86 samples of Fibroadenoma breast cancer histopathological and so on. Findings of the testing data of the MSVM represent that selected features with multi-objective feature selection with NSGA II produce an overall accuracy of 71.9%, sensitivity of 74.51%, and specificity of 96.02%. The number of the input images filled in the MSVM was 160 samples that are utilized for testing phase which composed of 20 samples of Adenosis, 20 samples of Fibroadenoma, 20 samples of Tubular Adenoma, 20 samples of Phyllodes Tumor, 20 samples of Ductal Carcinoma, 20 samples of Lobular Carcinoma, 20 samples of Mucinous Carcinoma and 20 samples Papillary Carcinoma. For example, 17 samples of 20 samples of Adenosis is correctly classified as Adenosis breast cancer histopathological. This corresponds to 85% of all 20 samples of Adenosis breast cancer histopathological. Similarly, 11 samples of 20 samples of Fibroadenoma is correctly classified as Fibroadenoma breast cancer histopathological. This corresponds to 55% of all 20 samples of Fibroadenoma breast cancer histopathological and so on. The efficiency of the PNN was estimated by the analysis of a confusion matrix. Results of the training data of the PNN represent that selected features with Multi-objective Feature Selection with NSGA II yield an overall accuracy of 100%, sensitivity of 100%, and specificity of 100%. The number of the input images filled in the PNN was 688 samples (i.e. 86 samples multiply 8 classes). For example, 86 samples of Adenosis are correctly classified as Adenosis breast cancer histopathological. This corresponds to 100% of all 86 samples of Adenosis breast cancer histopathological. Similarly, 86 samples of Fibroadenoma are correctly classified as Fibroadenoma breast cancer histopathological. This corresponds to 100% of all 86 samples of Fibroadenoma breast cancer histopathological and so on. Findings of the testing data of the PNN represent that selected features with multi-objective feature selection with NSGA II produce an overall accuracy of 82.5%, sensitivity of 92.71%, and specificity of 97.57%. The number of the input images filled in the PNN was 160 samples (i.e. 20 samples multiply 8 classes). For example, 20 samples of Adenosis are correctly classified as Adenosis breast cancer histopathological. This corresponds to 100% of all 20 samples of Adenosis breast cancer histopathological. Similarly, 16 samples of



20 samples of Fibroadenoma is correctly classified as Fibroadenoma breast cancer histopathological. This corresponds to 80% of all 20 samples of Fibroadenoma breast cancer histopathological and so on. These outcomes refer to the best accuracy obtained by the combined features set with the accuracy being 100% by the training dataset with PNN and 82.5% with PNN by the testing dataset as in Table.1. In Table 2, accuracy, sensitivity, specificity, and precision are calculated using ACO for PNN and MSVM classifiers.

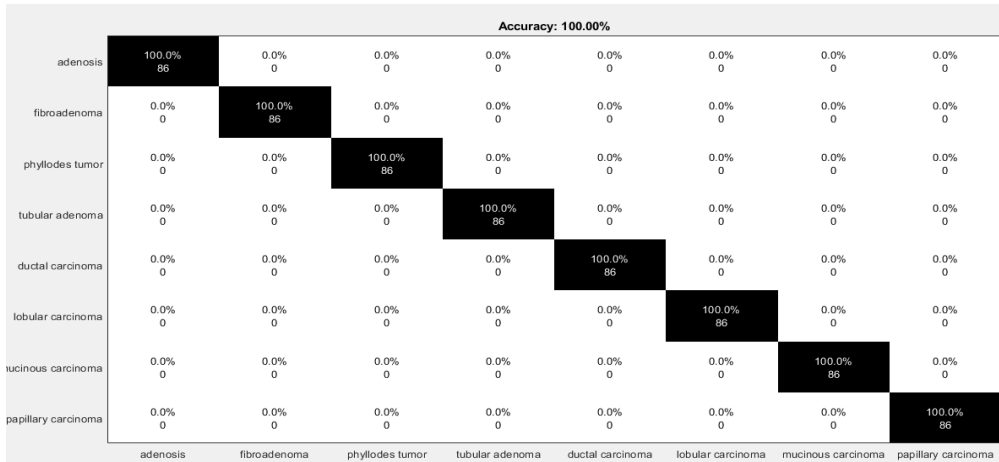


Fig.6: Confusion matrix of mixed features in training data for PNN classifier by using NSGA II

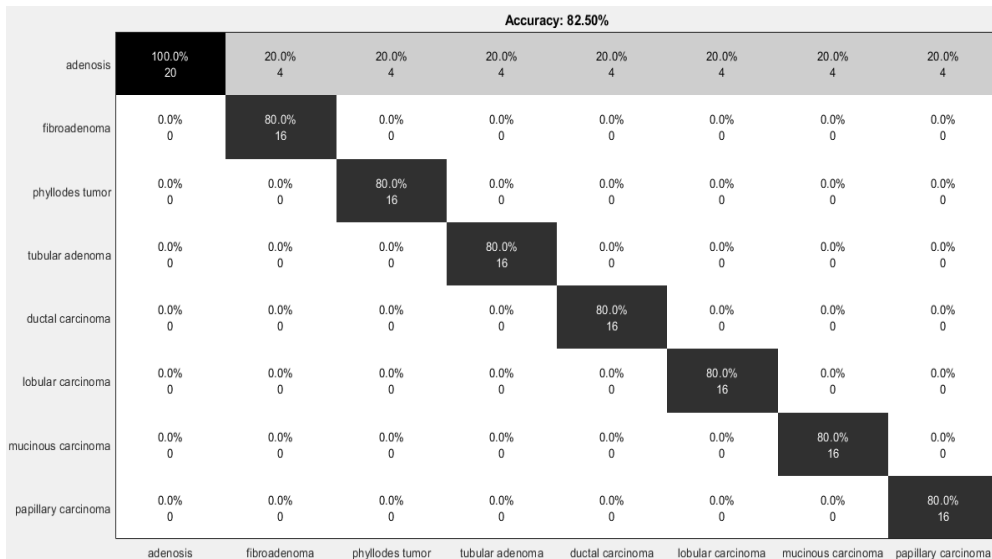


Fig.7: Confusion matrix of mixed features in testing data for PPN classifier by using NSGA II

**Accuracy: 88.08%**

adenosis	95.3% 82	3.5% 3	8.1% 7	0.0% 0	0.0% 0	0.0% 0	12.8% 11	0.0% 0
fibroadenoma	0.0% 0	73.3% 63	1.2% 1	3.5% 3	0.0% 0	3.5% 3	0.0% 0	0.0% 0
phyllodes tumor	1.2% 1	9.3% 8	77.9% 67	0.0% 0	0.0% 0	0.0% 0	1.2% 1	0.0% 0
tubular adenoma	0.0% 0	2.3% 2	0.0% 0	91.9% 79	0.0% 0	0.0% 0	0.0% 0	0.0% 0
ductal carcinoma	0.0% 0	0.0% 0	2.3% 2	0.0% 0	100.0% 86	0.0% 0	2.3% 2	8.1% 7
lobular carcinoma	3.5% 3	9.3% 8	10.5% 9	0.0% 0	0.0% 0	96.5% 83	1.2% 1	0.0% 0
mucinous carcinoma	0.0% 0	2.3% 2	0.0% 0	2.3% 2	0.0% 0	0.0% 0	77.9% 67	0.0% 0
papillary carcinoma	0.0% 0	0.0% 0	0.0% 0	2.3% 2	0.0% 0	0.0% 0	4.7% 4	91.9% 79
	adenosis	fibroadenoma	phyllodes tumor	tubular adenoma	ductal carcinoma	lobular carcinoma	mucinous carcinoma	papillary carcinoma

Fig.8: Confusion matrix of mixed features in training data for MSVM classifier by using NSGA II

**Accuracy: 71.88%**

adenosis	85.0% 17	15.0% 3	10.0% 2	0.0% 0	0.0% 0	15.0% 3	25.0% 5	10.0% 2
fibroadenoma	0.0% 0	55.0% 11	0.0% 0	5.0% 1	10.0% 2	0.0% 0	0.0% 0	0.0% 0
phyllodes tumor	0.0% 0	20.0% 4	65.0% 13	0.0% 0	0.0% 0	5.0% 1	0.0% 0	5.0% 1
tubular adenoma	5.0% 1	0.0% 0	5.0% 1	70.0% 14	0.0% 0	0.0% 0	5.0% 1	0.0% 0
ductal carcinoma	0.0% 0	0.0% 0	5.0% 1	0.0% 0	85.0% 17	0.0% 0	5.0% 1	10.0% 2
lobular carcinoma	0.0% 0	10.0% 2	10.0% 2	0.0% 0	0.0% 0	80.0% 16	0.0% 0	5.0% 1
mucinous carcinoma	0.0% 0	0.0% 0	0.0% 0	0.0% 0	5.0% 1	0.0% 0	65.0% 13	0.0% 0
papillary carcinoma	10.0% 2	0.0% 0	5.0% 1	25.0% 5	0.0% 0	0.0% 0	0.0% 0	70.0% 14
	adenosis	fibroadenoma	phyllodes tumor	tubular adenoma	ductal carcinoma	lobular carcinoma	mucinous carcinoma	papillary carcinoma

Fig.9: Confusion matrix of mixed features in testing data for MSVM classifier by using NSGA II

**Accuracy: 100.00%**

adenosis	100.0% 86	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0
fibroadenoma	0.0% 0	100.0% 86	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0
phyllodes tumor	0.0% 0	0.0% 0	100.0% 86	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0
tubular adenoma	0.0% 0	0.0% 0	0.0% 0	100.0% 86	0.0% 0	0.0% 0	0.0% 0	0.0% 0
ductal carcinoma	0.0% 0	0.0% 0	0.0% 0	0.0% 0	100.0% 86	0.0% 0	0.0% 0	0.0% 0
lobular carcinoma	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0	100.0% 86	0.0% 0	0.0% 0
mucinous carcinoma	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0	100.0% 86	0.0% 0
papillary carcinoma	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0	100.0% 86
	adenosis	fibroadenoma	phyllodes tumor	tubular adenoma	ductal carcinoma	lobular carcinoma	mucinous carcinoma	papillary carcinoma

Fig.10: Confusion matrix of mixed features in training data for PNN classifier by using ACO.

**Accuracy: 81.25%**

adenosis	90.0% 18	10.0% 2	20.0% 4	15.0% 3	10.0% 2	10.0% 2	5.0% 1	0.0% 0
fibroadenoma	5.0% 1	80.0% 16	0.0% 0	0.0% 0	10.0% 2	0.0% 0	0.0% 0	10.0% 2
phyllodes tumor	0.0% 0	0.0% 0	80.0% 16	0.0% 0	0.0% 0	5.0% 1	5.0% 1	10.0% 2
tubular adenoma	0.0% 0	5.0% 1	0.0% 0	80.0% 16	0.0% 0	0.0% 0	5.0% 1	0.0% 0
ductal carcinoma	0.0% 0	0.0% 0	0.0% 0	5.0% 1	80.0% 16	5.0% 1	5.0% 1	0.0% 0
lobular carcinoma	5.0% 1	0.0% 0	0.0% 0	0.0% 0	0.0% 0	80.0% 16	0.0% 0	0.0% 0
mucinous carcinoma	0.0% 0	5.0% 1	0.0% 0	0.0% 0	0.0% 0	0.0% 0	80.0% 16	0.0% 0
papillary carcinoma	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0	80.0% 16
	adenosis	fibroadenoma	phyllodes tumor	tubular adenoma	ductal carcinoma	lobular carcinoma	mucinous carcinoma	papillary carcinoma

Fig.11: Confusion matrix of mixed features in testing data for PPN classifier by using ACO.

**Accuracy: 52.76%**

adenosis	96.5% 83	12.8% 11	44.2% 38	2.3% 2	1.2% 1	24.4% 21	24.4% 21	0.0% 0
fibroadenoma	0.0% 0	38.4% 33	2.3% 2	14.0% 12	0.0% 0	1.2% 1	8.1% 7	100.0% 86
phyllodes tumor	0.0% 0	4.7% 4	18.6% 16	0.0% 0	1.2% 1	1.2% 1	0.0% 0	0.0% 0
tubular adenoma	0.0% 0	1.2% 1	0.0% 0	51.2% 44	1.2% 1	0.0% 0	0.0% 0	0.0% 0
ductal carcinoma	1.2% 1	3.5% 3	5.8% 5	14.0% 12	95.3% 82	0.0% 0	8.1% 7	0.0% 0
lobular carcinoma	2.3% 2	26.7% 23	25.6% 22	7.0% 6	0.0% 0	69.8% 60	5.8% 5	0.0% 0
mucinous carcinoma	0.0% 0	10.5% 9	0.0% 0	5.8% 5	0.0% 0	3.5% 3	52.3% 45	0.0% 0
papillary carcinoma	0.0% 0	2.3% 2	3.5% 3	5.8% 5	1.2% 1	0.0% 0	1.2% 1	0.0% 0
	adenosis	fibroadenoma	phyllodes tumor	tubular adenoma	ductal carcinoma	lobular carcinoma	mucinous carcinoma	papillary carcinoma

Fig.12: Confusion matrix of mixed features in training data for MSVM classifier by using ACO.

**Accuracy: 48.75%**

adenosis	80.0% 16	45.0% 9	75.0% 15	0.0% 0	0.0% 0	15.0% 3	20.0% 4	45.0% 9
fibroadenoma	0.0% 0	25.0% 5	0.0% 0	10.0% 2	0.0% 0	0.0% 0	10.0% 2	10.0% 2
phyllodes tumor	0.0% 0	0.0% 0	0.0% 0	5.0% 1	0.0% 0	0.0% 0	0.0% 0	5.0% 1
tubular adenoma	5.0% 1	0.0% 0	5.0% 1	65.0% 13	0.0% 0	0.0% 0	0.0% 0	0.0% 0
ductal carcinoma	5.0% 1	5.0% 1	5.0% 1	10.0% 2	90.0% 18	5.0% 1	15.0% 3	25.0% 5
lobular carcinoma	0.0% 0	15.0% 3	0.0% 0	10.0% 2	10.0% 2	70.0% 14	10.0% 2	0.0% 0
mucinous carcinoma	5.0% 1	10.0% 2	0.0% 0	0.0% 0	0.0% 0	0.0% 0	45.0% 9	0.0% 0
papillary carcinoma	5.0% 1	0.0% 0	15.0% 3	0.0% 0	0.0% 0	10.0% 2	0.0% 0	15.0% 3
	adenosis	fibroadenoma	phyllodes tumor	tubular adenoma	ductal carcinoma	lobular carcinoma	mucinous carcinoma	papillary carcinoma

Fig.13: Confusion matrix of mixed features in testing data for MSVM classifier by using ACO.

Table 1: Accuracy, sensitivity, specificity, and precision by using NSGA II

<b>Classifier</b>	<b>Accuracy %</b>	<b>Sensitivity%</b>	<b>Specificity%</b>	<b>Precision%</b>
MSVM of training database	88.1	88.74	98.32	88.1
MSVM of testing database	71.88	74.51	96.02	71.88
PNN of training database	100	100	100	100
PNN of testing database	82.5	92.71	97.57	82.5

Table 2: Accuracy, sensitivity, specificity, and precision by using ACO.

<b>Classifier</b>	<b>Accuracy %</b>	<b>Sensitivity %</b>	<b>Specificity %</b>	<b>Precision %</b>
MSVM of training database	52.76	54.58	93.52	52.76
MSVM of testing database	48.75	48.27	93	48.75
PNN of training database	100	100	100	100
PNN of testing database	81.25	84.22	97.33	81.25

## 5. Conclusion

In this paper, we have proposed a classification algorithm for multi-class breast cancer images that created the predictive model of eight types of breast cancer. The outcomes of the suggested method on BreakHis images showed the PNN classifier is more accurate than the MSVM classifier for the training phase. The NSGA II algorithm supported the proposed classifier rather than ACO algorithm by using the testing dataset. It means that the NSGA II algorithm acts as a good choice for reducing the error rate created from using a large set of features.

## References

- A. Azar and S. El-Said. (2012). Probabilistic neural network for breast cancer classification. *Neural Computing and Applications*, 23(6), 1737–1751.
- A. Cruz-Roa *et al.* (2014). Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. in *Medical Imaging 2014: Digital Pathology*, 9041- 9043.
- A. D. Belsare, M. M. Mushrif, M. A. Pangarkar, and N. Meshram. (2015). Classification of breast cancer histopathology images using texture feature analysis. in *TENCON 2015-2015 IEEE Region 10 Conference*, 1–5.
- A. Khan, A. R. Baig (2015). Multi-Objective Feature Subset Selection using Non-dominated Sorting Genetic Algorithm. *Journal of Applied Research and Technology*, 13 (1),145-159.
- Asri, H., Mousannif, H., Al Moatassime, H., and Noel, T. (2016). Using machine learning algorithms for breast cancer risk prediction and diagnosis. *Procedia Computer Science*, 83, 1064–1069.
- Aswathy, M., and Jagannath, M. (2017). Detection of breast cancer on digital histopathology images: present status and future possibilities. *Informatics in Medicine Unlocked*, 8, 74–79.
- Burt, J. R. *et al.* (2018). Deep learning beyond cats and dogs: Recent advances in diagnosing breast cancer with deep neural networks. *The British Journal of Radiology*. 20170545.
- BenTaieb, A., Li-Chang, H., Huntsman, D., Hamarneh, G. (2017). A structured latent model for ovarian carcinoma subtyping from histopathology slides. *Medical Image Analysis*,39,194–205.
- Deb, K., Samir, A., Amrit, P., and Meyarivan, T. (2002). A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: Nsga-ii. *IEEE Transactions on Evolutionary Computation*, 6(2), 103–112.
- Dorigo, M., Blum, C. (2005). Ant colony optimization theory: A survey. *Theoretical Computer Science*, 344, 243–278.
- F. A. Spanhol, L. S. Oliveira, P. R. Cavalin, C. Petitjean, and L. Heutte. (2017). Deep features for breast cancer histopathological image classification. In

*Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff, Canada, October, 1868–1873.*

Han, Z., Wei, B., Zheng, Y., Yin, Y., Li, K., and Li, S. (2017). Breast cancer multiclassification from histopathological images with structured deep learning model. *Scientific Report*, 7(1), 4172.

J. Weston and C. Watkins. (1999). Multi-class support vector machines. In M. Verleysen, editor, *Proceedings of ESANN99, Brussels, D. Facto Press*, 219–224.

Juanying Xie, Ran Liu, Joseph Luttrell IV, Chaoyang Zhang. (2019). Deep Learning Based Analysis of Histopathological Images of Breast Cancer. *Frontiers in Genetics*, 10-80.

Lehman, C. D. et al. (2016). National Performance Benchmarks for Modern Screening Digital Mammography: Update from the Breast Cancer Surveillance Consortium. *Radiology*, 283, 49–58.

M. Kowal and P. Filipczuk. (2014). Nuclei segmentation for computer-aided diagnosis of breast cancer. *International Journal of Applied Mathematics and Computer Science*, 24 (1), 19–31.

M. Vasantha, V. Bharathi, and R. Dhamodharan. (2010). Medical image feature, extraction, selection and classification. *International Journal of Engineering Science and Technology*, 2 (6), 2071–2076.

M. Hu. (1962). Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, 8, 179–187.

N. Suematsu, Y. Ishida, A. Hayashi, and T. Kanbara. (2002). Regionbased image retrieval using wavelet transform. in *Proceedings of the 15th International Conference on Vision Interface, Calgary, Canada, May, 9–16.*

Nawaz, M., Sewissy, A. A., and Soliman, T. H. A. (2018). Multi-class breast cancer classification using deep learning convolutional neural network. *International Journal of Advanced Computer Science and Applications*, 9, 316–322.

N Howlader, AM Noone, M Krapcho, et al., eds. (2019). SEER Cancer Statistics Review, 1975 - 2016. Based on the November 2018 SEER data submission, posted to the SEER website April 2019. *Surveillance, Epidemiology, and End Results Program, National Cancer Institute, National Institutes of Health; 2019.*

P. Filipczuk, T. Fevens, A. Krzyżak, and R. Monczak. (2013). Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies. *IEEE Transactions on Medical Imaging*, 32 (12), 2169–2178.

Spanhol, F., Oliveira, L. S., Petitjean, C., Heutte, L. (2016). A Dataset for Breast Cancer Histopathological Image Classification. *IEEE Transactions on Biomedical Engineering (TBME)*, 63(7), 1455-1462.

S. Selvarajah and S. Kodituwakku, 2011. Analysis and comparison of texture features for content based image retrieval. *International Journal of Latest Trends in Computing*, 2 (1).

Tarek M. Hamdani, Jin-Myung Won, Adel M. Alimi, and Fakhri Karray, (2007). Multi-objective Feature Selection with NSGA II. *International Conference on Adaptive and Natural Computing Algorithms*. 240–247.

Wu H, Phan JH, Bhatia AK, Shehata B, Wang MD. (2015). Detection of Blur Artifacts in Histopathological Whole-Slide Images of Endomyocardial Biopsies. *Conf Proc Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Conf*, 727–30.

Y. M. George, H. L. Zayed, M. I. Roushdy, and B. M. Elbagoury. (2014). Remote computer-aided breast cancer detection and diagnosis system based on cytological images. *IEEE Systems Journal*, 8 (3), 949–964.

Zhang, Y., Zhang, B., Coenen, F., and Lu, W. (2013). Breast cancer diagnosis from biopsy images with highly reliable random subspace classifier ensembles. *Machine Vision Applications*, 24, 1405–1420.