IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 17 February 2024, accepted 28 February 2024, date of publication 12 March 2024, date of current version 18 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3374890

APPLIED RESEARCH

NSGA-II-DL: Metaheuristic Optimal Feature Selection With Deep Learning Framework for HER2 Classification in Breast Cancer

TARIK AHMED RASHID[1], (Member, IEEE), JAFAR MAJIDPOUR[2,3], RAJERMANI THINAKARAN[4], MALATHY BATUMALAY[4], DESHINTA ARROVA DEWI[4], BRYAR A. HASSAN[5], HABIBOLLAH DADGAR[6], AND HOSSEIN ARABI[7]

[1]Department of Computer Science and Engineering, University of Kurdistan Hewlêr, Erbil 44001, Iraq
[2]Department of Software and Informatics Engineering, University of Raparin, Rania, Kurdistan 46012, Iraq
[3]Department of Computer Science, Faculty of Science, Soran University, Soran, Kurdistan 44008, Iraq
[4]Faculty of Data Science and Information Technology, INTI International University, Nilai 71800, Malaysia
[5]Department of Computer Science, College of Science, Charmo University, Chamchamal, Sulaimani, Kurdistan 46023, Iraq
[6]Imam Reza Cancer Research, Nuclear Medicine and Molecular Imaging Department, Razavi Hospital, Mashhad 1696700, Iran
[7]Division of Nuclear Medicine and Molecular Imaging, Department of Medical Imaging, Geneva University Hospital, 1211 Geneva, Switzerland

Corresponding author: Tarik Ahmed Rashid (tarik.ahmed@ukh.edu.krd)

This work was supported by the INTI International University-Malaysia.

ABSTRACT Immunohistochemistry (IHC) slides are graded for breast cancer based on visual markers and morphological characteristics of stained membrane regions. The usage of whole slide images (WSIs) from histology in digital pathology algorithms for computer-assisted evaluations has increased recently. Human epidermal growth factor receptor 2 (HER2)-stained microscopic images are challenging, time-consuming, and error-prone to evaluate manually. This is due to different staining, overlapped regions, and huge, non-homogeneous slides. Additionally, the classification of HER2 images by the selection of fundamental features must be used to capture the difficult elements of the images, such as the irregular cell structure and the coloring of the tissue of the cells. To solve the above problems, a transfer learning model-based, trainable metaheuristic method for choosing the best features is suggested in this paper. Moreover, the suggested model is efficient in reducing model complexity and computational costs as well as avoiding overfitting. The four main components of the proposed cascaded design are: (1) converting WSIs to tiled images and enhancing contrast with fast local Laplacian filtering (FlLpF); (2) extracting features with a ResNet50 CNN technique based on transfer learning; (3) selecting the most informative features with the help of a non-dominated sorting genetic algorithm (NSGA-II) optimizer; and (4) using a support vector machine (SVM) to classify HER2 scores. Results from the HER2SC and HER2GAN datasets show that the suggested model is superior to other methods already in use, with 94.4% accuracy, 93.71% precision, 98.07% specificity, 93.83% sensitivity, and a 93.71% F1-score for the HER2SC dataset being achieved.

INDEX TERMS HER2, CNN, transfer learning, NSGA-II optimizer, FlLpF.

I. INTRODUCTION

HER2 stands for human epidermal growth factor receptor 2, and it is a protein that stimulates the development of cancerous epithelial cells. In instances of breast cancer (BC) with invasive tumors, the HER2 gene is amplified. When we look at HER2+ (positive) and HER2 (negative) cases of BC, it can be seen that neoplastic changes happen more often in HER2+ cases, which means that tumors can grow without being stopped. Overexpression of the HER2 protein is linked with poor prognosis, poorer survival, and increased recurrence [1], [2], and is thus suggested for all patients with invasive breast cancer [3]. Recent research [4] has described HER2 status as a prognostic factor that correlates invasive tumors with mortality and recurrence-free survival. This emphasizes the need for an accurate assessment of HER2 overexpression

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

in HER2+ patients so that they may get the most effective anti-HER2 medication possible. Immunohistochemistry (IHC) and fluorescence in situ hybridization (FISH) were two common ways to measure HER2 expression [5]. Samples are usually put into one of four HER2 classes based on how many and how strongly the membranes of invasive cancer cells are stained. Based on the IHC examination hypothesis for figuring out the BC [6], here's how to figure out the HER2 score:

- 3+ (More than 10% of invasive cancer cells had strong, complete, and consistent staining of the cell membrane);
- 2+ (More than 10% of tumor cells exhibited weak to moderate full membrane staining);
- 1+ (More than 10% of tumor cells had faint, almost invisible incomplete membrane staining);
- 0 (Less than 10% of invasive cancer cells exhibited no staining or patchy, weak staining of the cell membrane).

Pathologists visually examine the biopsy tissue slides under the microscope as part of a standard clinical procedure. This kind of visual evaluation is often open to mistakes. HER2 testing becomes hard to do in places with limited resources and where qualified pathologists aren't always available right away [7]. IHC slide digitization and quantitative image analysis have become crucial for image preservation and repeatable diagnosis to overcome both of these problems. Thus, pathologists primarily use computer-assisted digital approaches for quantitative image analysis. The repeatability and ease of use of digital pathology are its key benefits. Digital imaging technology is also a pixel-based technology, which lowers false positives and inter-observer variability by increasing detection and segmentation accuracy, among other things [8].

Assessing the HER2 score manually is challenging, time-intensive, and prone to errors. The reason for this is the presence of overlapping regions, extensive and heterogeneous slides, and diverse stains. To accurately depict the intricate features of the images, such as the irregular cell structure and the pigmentation of the cell tissue, the HER2 images must be classified based on a predefined set of criteria. The study proposes a hybrid paradigm that effectively tackles these challenges. The model comprises a transfer learning model, a metaheuristic optimizer, and a machine learning algorithm. The features were extracted from the transfer learning model, specifically Resnet50. Additionally, we developed a trainable multi-objective metaheuristic optimizer called NSGA-II. An SVM machine learning algorithm is used to classify the selected features and provide feedback to the NSGA-II algorithm for the purpose of reselecting the minimal optimal features. The results showcase the exceptional, adaptable, and optimal performance of the proposed model.

Contributions made by this paper include the following:

1. Optimal feature selection from the irregular structure of the cells and the coloring of the tissue of the cells.
2. Utilizing the NSGA-II to select minimum-perfect features.
3. High performance accuracy to classify the HER2 score.
4. Calculation of the error value and sending it to the NSGA-II optimizer as feedback
5. Reduce time consumption for HER2 score decisions using DL and ML networks.

An application of our proposed method is its potential installation on imaging devices for real-time, clinical examination, rapid, and accurate image classification. Additionally, it can assist specialists in identifying the type of HER2 score while reviewing the images.

A review of earlier research on the same problem that we are addressing in this work was done in Part 2, and a detailed technical background of the major themes is given in Section III. The findings from the experiments that were applied to the data using the suggested model are shown in Section IV. Lastly, the paper's general findings are presented in Section V.

II. RELATED WORK

A lot of research was published on various machine learning (ML) and deep learning (DL), including segmentation, texture features, classification, etc., to solve the difficulties in HER2 score diagnosis [9], [10], [11], [12], [13], [14], [15], [16], [17], [18]. Numerous studies have employed the HER2 classification using DL technology, which eliminates the need to manually set up the feature extractor [9]. Instead, a deep neural network with multi-layer nonlinear transformation is used to automatically learn the internal structure feature representation of the data from the input data. HER2GAN presents a novel supervised deep learning-based method for overcoming the scarcity of HER2 datasets [10]. To achieve optimal results via transfer learning and generate HER2 images of superior quality, a GAN-based model is suggested. The InceptionResNetV2 model achieved 94.2% accuracy by using a mix of generated and original images for training and testing. To automatically classify breast imaging cancers, [11] presented a complete connectivity layer structure utilizing the maxout activation function and incorporated it into the CNN model. The model's AUC increased from 0.787 to 0.822. To choose more discriminative patches, [12] suggested a patch screening approach based on the clustering algorithm and CNN. This method achieved 88.89% accuracy on the whole test set. Convolutional, pooling, and fully connected layers were used to create the model [13]. The probability of passing the exam is 99.7 percent. Reference [14] presented a deep reinforcement learning method for automatically scoring IHC-stained HER2 breast cancer slides. The approach treats the IHC score as a series of choices, making it simple to zero in on regions with high diagnostic potential. We just employed ResNet50 DL for feature extraction in our proposed model.

Several academics have done their studies based on the texture of the images. The authors of [19] used two ML techniques, logistic regression, and SVM, as well as uniform local binary patterns as a texture descriptor for feature

extraction, to determine the HER2 score. The accuracy rates of the two models were 91% and 93%, respectively. Unlike the semi-supervised model used in [19] and [20] uses a supervised classification framework. Reference [21] proposed a semi-supervised method to discover lesions in colorectal biopsies that was simple to grasp and had a classification accuracy of 90.19% based on ML and feature aggregation approaches. Performance would be improved by training tiles for HER2 and then combining them all with CNN in a single optimization procedure. Using HER2-ResNet, a DL and pretrained model, [9] completed the identical test with 91% accuracy. The work detailed in [22] and [23] included the extraction of characteristics from the HER2 2+ tissue and the application of ML algorithms for classification, producing mostly acceptable findings.

In addition, DL-based HER2 image segmentation is a growing area of study for determining HER2 scores in diagnostic settings. The authors of [8] suggested a unique methodology for automatically segmenting, categorizing, and quantifying IHC breast cancer images inside an ML framework. They started by using an SVM feature learning classifier to divide the WSI into epithelial and stromal components. To automatically segregate or categorize epithelial (EP) and stromal (ST) regions from digital tumor tissue microarrays, [24] introduced a Deep Convolutional Neural Networks (DCNN)-based feature learning technique. To achieve a 98% success rate in HER2 image segmentation, [25] employed a U-Net network with a trapezoidal long short-term memory (TLSTM) in its latent layers.

The selection of acceptable features for the HER2 image was the subject of fewer investigations since it is difficult due to the complex structure of the HER2 image. Using a binary pixel classification method, [26] showed a new way to pick out representative features in order to separate images of breast cancer that show high levels of HER2. They were able to preserve good classification performance by reducing the original collection of 210 color and texture characteristics to 65 features. In cases when the medical images have unique morphology and complex structure, such as HER2 images, metaheuristic approaches may play a crucial role in selecting the optimal feature. A multi-objective metaheuristic optimizer was used in this investigation. In multi-objective optimization problems [27], [28], [29], [30], there are two or more optimization goals that compete with each other. This means that reaching one goal will make it harder to reach another.

Pretrained deep learning models can be used to detect the HER2 score and analyze breast ultrasound images by extracting relevant features. References [31] and [32] employed pretrained models for feature extraction and applied NCA (neighborhood component analysis) to identify the most optimal features. The Kaplan et al., [31] developed a Breast Imaging Reporting and Data System (BI-RADS) and achieved accuracy rates of 79.29%, 80.42%, and 88.67% for Case 1, Case 2, and Case 3, respectively. The image classification approach provided by [32], which utilizes grid-based deep feature creation, achieved a classification accuracy of 97.18% on ultrasonic pictures for three distinct classes: malignant, benign, and normal.

TABLE 1. Summary of HER2 scoring approaches.

Study	Year	Method	Objective	Evaluation (Accuracy)
Saha et al. [8]	2018	SVM	Segmentation and classification HER2 score	98.33%
Wang et al. [9]	2022	Resnet	HER2 classification	92%
Mirimoghaddam et al. [10]	2023	GAN and pretrained models	Scarcity of HER2 image	94.2%
Arevalo et al. [11]	2015	CNN	HER2 classification	82.2%
Li et al. [12]	2019	CNN	Breast cancer classification	88.9
Qaiser et al. [14]	2019	CNN & LSTM	HER2 score	79.4%
Mukundan et al. [19]	2019	Logistic regression, SVM	Automated scoring of HER2	93%
Xu et al. [24]	2016	DCNN	Automatically segment and classify	98%
Kaplan et al. [31]	2022	CNN, NCA	Classification of breast ultrasound lesions	88.67%
Liu et al. [32]	2022	CNN, NCA	Breast ultrasound lesions diagnosis	97.18%

In cases where the medical images have unique morphology and complex structure, such as HER2 images, metaheuristic approaches may play a crucial role in selecting the optimal feature. NSGA-II [33] is a very effective multi-objective optimization technique. Several metaheuristic algorithms make use of the NSGA-II idea, such as ant colony optimization (ACO) [34], bee colony optimization (BCO) [35], the Differential Evolution (DE) [36] method, and others; however, the Pareto front estimate technique is specific to the NSGA-II. In contrast to other feature selection approaches such as Chi2 [37], neighborhood component analysis (NCA) [38], and ReliefF [39], multi-objective optimization techniques like NSGA-II are trainable and may find optimum and minimum features in problems. Each iteration of multi-objective optimization provides a set of minimum and optimal features. Table 1 presents a concise overview of the several methods used for scoring HER2.

III. MATERIAL AND METHODS

A method that combines DL and a metaheuristic optimizer is proposed, as shown in Figure 1, to diagnose the HER2

score with the best accuracy, the fewest errors, and the least amount of time. Algorithm 1 offers a thorough explanation of the whole procedure. The structure of our proposed model includes four steps: (i) a preprocessing step containing cropping the WSI image and then enhancing the cropped image based on the FlLpF method; (ii) a TL-based Resnet50 model is used for feature extraction; (iii) a fine-tuned NSGA-II optimizer is used for optimal feature selection; (iv) SVM is used for the HER2 score diagnostic. Following are detailed explanations of each stage.

A. DATASET

The model's capability to generate and classify convincing images was evaluated using two distinct sets of data.

For the first dataset, the Department of Computer Science at the University of Warwick in the United Kingdom donated the publicly available HER2 image dataset (HER2SC), which was used to evaluate the suggested strategy [40], [41]. The dataset was gathered using a Hamamatsu NanoZoomer C9600 scanner and mostly consisted of 158 WSIs. A monoclonal antibody against the HER2 gene was used to stain 79 of the 158 WSIs, while the other 79 were stained with hematoxylin and eosin (H&E). Each WSI was 100,000 by 80,000 pixels in size (width by height). The WSIs could be seen at magnifications ranging from 4 to 40. Figure 2 includes several instances of HER2SC images.

For the second dataset, the Department of Pathology at Mashhad University of Medical Sciences' Ghaem Educational Research and Treatment Center collected IHC-stained tissue (HER2GAN) samples from 126 individuals. After diagnosis of Invasive Ductal Carcinoma of the breast, all female patients had core needle or excisional biopsies stained for the HER2 marker. 32 patients scored zero, 40 scored one, 30 scored two, and 24 scored three, which is an imbalanced dataset. To protect patient privacy, we deidentified our data during screening. Pathology slides were digitized using a Jenus microscope, and manual digital photography was done using a Tucsen Photonics TrueChrome II digital camera with a CMOS sensor with magnifications of 10–20 and 40 in diagnostic areas suitable for HER2 scoring [10].

B. TRANSFER LEARNING MODEL

In some situations, the challenges of putting up a large dataset may reduce the model's performance accuracy, even if a large dataset is necessary for CNN training to reach the necessary accuracy. Training and testing data pairs are notoriously difficult to get in the real world [42]. To address this issue, ''transfer learning'' was put forth as a solution. To rephrase, transfer learning is an ML strategy in which we use a previously trained model as the basis for a model applied to a different problem. A model that has been optimized for one task and then applied to another enables quick progress to be made while modeling the second task. Compared to training with a minimal quantity of data, the results of applying transfer learning to a new task are far more impressive [43]. For

Algorithm 1 Pseudocode of the proposed method

Inputs:
***D1**: HER2SC Dataset*
***D2**: HER2GAN*
***img1**: a HER2's image which is selected from D1*
***img2**: a HER2's image which is selected from D2*
*For **img1**, **img2** in (**D1** and **D2**)*
  *[$\mathbf{Img_{D1}}$, $\mathbf{Img_{D2}}$] = Select 2048 features from [**img1**, **img2**]*
End for
For 1: iterations
  *[$\mathbf{M_{D1}}$, $\mathbf{M_{D2}}$] = Build a model **M** based on **NSGA-II** model which select optimal minimum features from [$\mathbf{Img_{D1}}$, $\mathbf{Img_{D2}}$]*
  *[$\mathbf{C_{D1}}$, $\mathbf{C_{D2}}$] = Build a model **C** based on **SVM** classifier which classify [$\mathbf{M_{D1}}$, $\mathbf{M_{D2}}$]*
  ***[MSE, accuracy, specificity, sensitivity, precision, F1-score]** =*
  *Calculate metrics **[MSE, accuracy, specificity, sensitivity, precision, F1-score]** from [$\mathbf{C_{D1}}$, $\mathbf{C_{D2}}$]*
End for
End

feature extraction, our recommended method makes use of the Resnet50 transfer learning model, which is detailed in the section below.

1) RESNET50

When researchers attempted to add layers to the CNN design, they ran into issues that led to the development of ResNet. The performance of a CNN design improves as more layers are added, but then it begins to degrade. The ResNet model built to address this issue does not carry out gradient computation. To avoid this, a fast cut is given by simply adding x to the f(x) function (See Figure 3).

ResNet has a flat 34-layer network architecture with fewer filters and a simpler design. Learn More By including jump connections or residual blocks into this flat network, the design is then transformed into a residual network. In a 34-layer network, Resnet50 is created by swapping out every 2-layer block with a 3-layer bottleneck block [44], [45]. The research conducted for feature extraction employed ResNet50. A 50-layer convolutional neural network, ResNet-50, consists of 48 convolutional layers, 1 MaxPool layer, and 1 average pool layer.

C. NSGA-II OPTIMIZER

A lot of problems have been pointed out with the first non-dominated sorting genetic algorithm (NSGA) [46], such as the fact that it doesn't have any elitism, you have to set a sharing parameter to keep the diversity, and it takes a lot of time to run. NSGA-II architecture, on the other hand, exhibits elitism and does not need a sharing parameter. Its fast processing allows it to be called the ''Fast Elitist NSGA-II.'' The maximum complexity of NSGA-II is O(MN2), where M is the total number of objective functions and N is the population size. In addition, it uses the crowding distance operator as part of its variety preservation strategy. The elite preserving operator, crowding distance, selection operator,

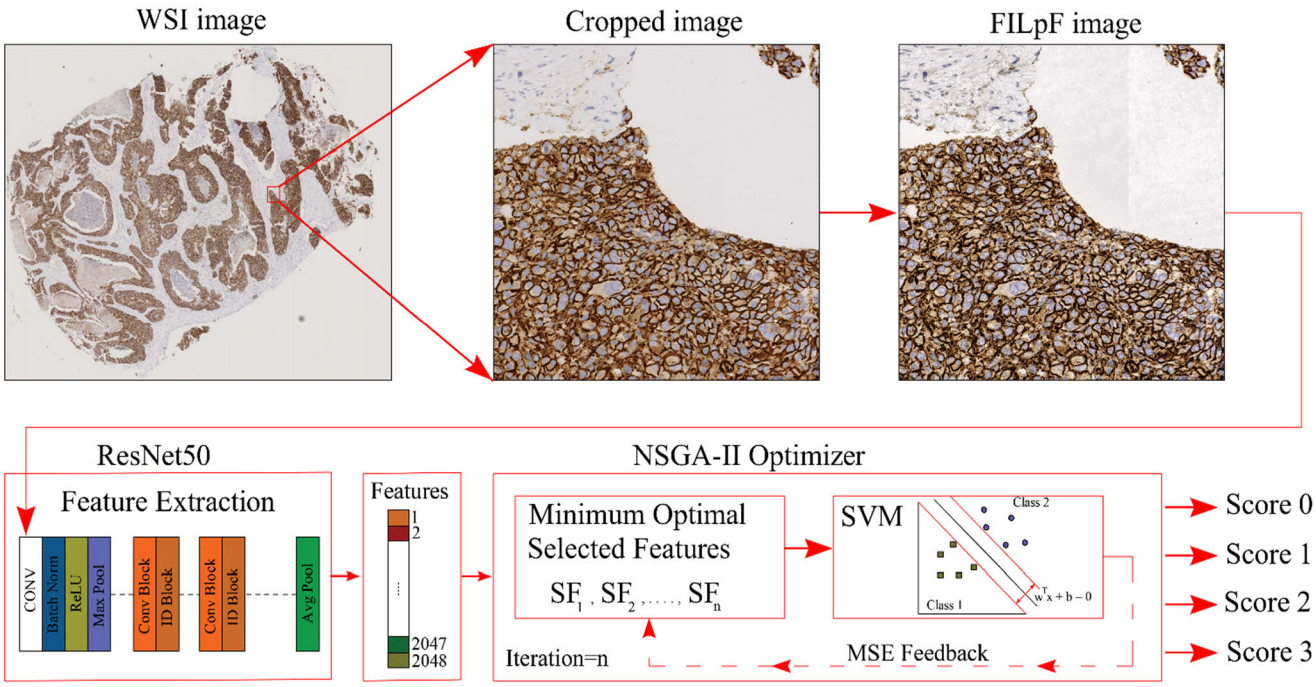


FIGURE 1. Proposed structure.

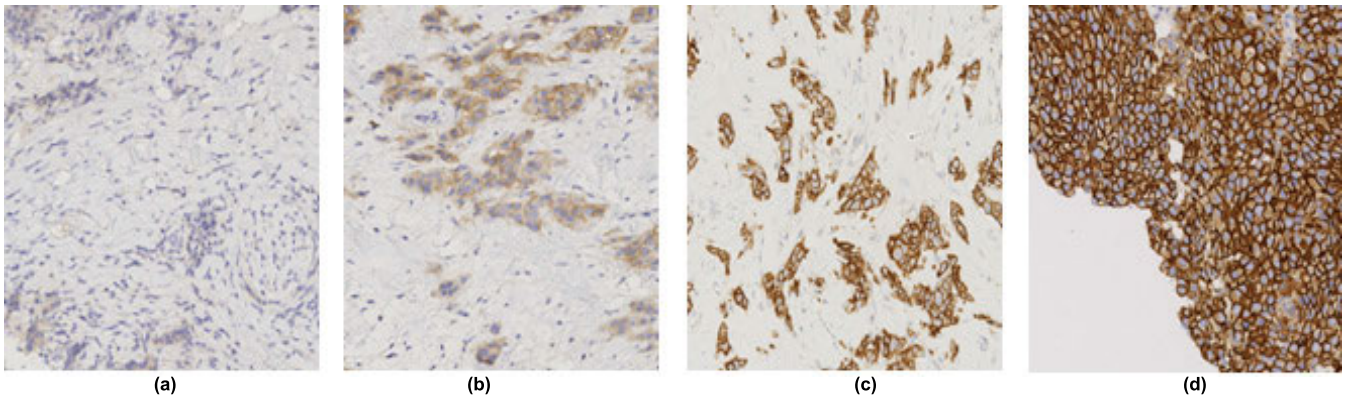


FIGURE 2. HER2-stained IHC image of breast cancer. Images from the HER2SC dataset a) score 0 b) score 1 c) score 2 d) score 3.

and non-dominated sorting are the four fundamental ideas that form the basis of the NSGA-II-based foundation. Below is a quick discussion of each of them [33].

1) NON-DOMINATED SORTING

Using the Pareto dominance principle, the whole population is sorted using this strategy. The non-dominated individuals of the starting population are given the top rank at the beginning of the procedure by non-dominated sorting. The first front is filled with first-rank members, who are then excluded from the original population. The remaining population members are then sorted using a non-dominant method. Non-dominant members of the remaining population are given the second rank and are seated in the second front [27]. As shown in Figure 4(a), the procedure is repeated until every member of the remaining population is given a rank.

2) ELITE PRESERVING OPERATOR

A population's elite solutions are preserved by the direct transmission of those solutions to the next generation. In other words, unless certain solutions take control, the non-dominant solutions of each generation are passed on to the next [47].

3) CROWDING DISTANCE

The crowding distance calculates how close an individual is to their neighbors. Better population variety will be the outcome of a larger average crowding distance. By adopting binary

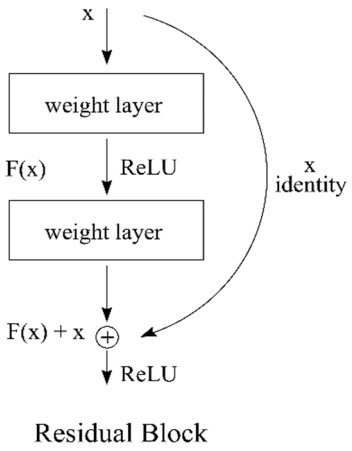
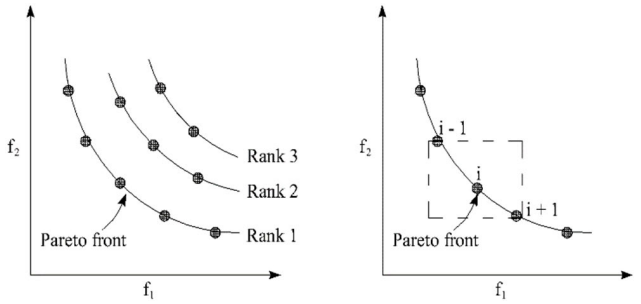


FIGURE 3. ResNet architecture.



a) Non-dominated sorting b) Crowding distance calculation

FIGURE 4. Non-dominated sorting procedure and Crowding distance calculation.

tournament selection based on rank and crowding distance as shown in Figure 4(b), parents are chosen from the population. Equations 1 and 2 compute the crowding distance [27]. The maximum and minimum values of the j^{th} objective function across all individuals is f_j^{max} and f_j^{min}, respectively, if f_j^i is the j^{th} value of an objective function for the i^{th} individual. Then, the crowding distance of an i^{th} individual by averaging the distances of the two closest solutions on either side of it is calculated.

$$d_i^j = \frac{\left| f_j^{previous} - f_j^{next} \right|}{f_j^{max} - f_j^{min}} \tag{1}$$

$$d_i = d_i^1 + \ldots + d_i^m = \sum_{j=1}^{m} d_i^j \tag{2}$$

4) SELECTION OPERATOR

A crowded tournament selection operator is used to determine the next generation's population based on the crowding distances and ranks of the current population. If two individuals in a population have the same rank, then the one with the greater crowding distance will be chosen to go on to the next generation. If their ranks are different, then the individual with the higher rank will be chosen [27], [48].

5) NSGA-II PROCEDURE

To begin, NSGA-II creates a population of starting candidates, denoted by P_t, with a size of N. Next, the population P_t experiences crossover and mutation to produce the new

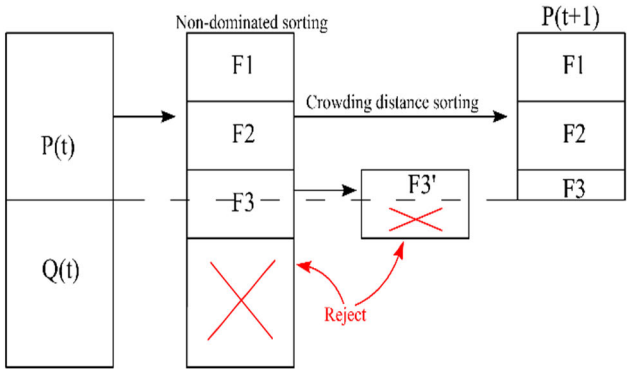


FIGURE 5. Procedure of NSGA-II.

population Q_t. Next, we merge P_t and PQ_t into a new population, R_t, and apply the non-dominated sorting method to R_t. In the next step, the R_t the population is stratified into distinct "fronts" based on their non-domination scores. After that, choose N people from Rt to form the next population, P_{t+1}. If there are fewer than N individuals in the first front, those individuals are promoted to the next generation immediately, and the remaining individuals are selected from the second front's least congested zone and added to P_{t+1}. If the size of the first front is larger than or equal to N, however, then only N members will be chosen from the least congested section of the first front to create P_{t+1}. If P_{t+1} is still less than N in size, the process is repeated for each succeeding front. With the help of Figure 5, we can see how NSGA-II operates. The rationale for adopting a multi-objective optimizer algorithm to determine the most effective features for breast cancer, particularly HER2, is that it provides measurements for both quality and regularity. Both measures are part of NSGA-II, and in this research, quality is more important than regularity.

The primary benefits of using NSGA-II for feature selection over other methods like Chi2, NCA, ReliefF, etc. are its multi-objectiveness and trainability. In order to provide NSGA-II feedback on MSE in each iteration and improve performance in later iterations, we set up NSGA-II with an SVM classifier. Another component of NSGA-II that may choose the minimum and greatest features in each iteration is trainability.

D. SVM

The SVM method generates a hyperplane decision boundary, which is situated at the dataset's extremes. Where to place a decision border and which features to use to get there [49]. There is some flexibility in terms of where and how the line is drawn. Because there is no unique decision boundary, incorrect classifications may be assigned. Thus, as illustrated in Figure 6(a), the edge of a data point exerts upward pressure on all points in the vicinity of the opposite class, which are known as support vectors [50]. The classifiers must find out where the outliers are and use those points to set their boundaries. Margin refers to the separation hyperplane, which is the D+ plus the D-. The shortest distance to the nearest positive

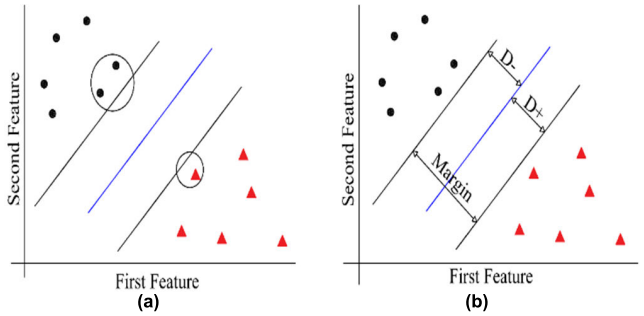


FIGURE 6. Support vector points a) samples from two classes left image, b) The Margin and D-, D+ with samples.

point (denoted D+) and the shortest distance to the nearest negative point (denoted D-) are shown in Figure 6(b) [51]. The SVM is used as a classifier to determine the HER2 score in this research.

Python was used to implement the ResNet50 model, the NSGA-II model, and the SVM classifier. For all training methods, the NVIDIA GeForce GTX 1080 Ti graphics processing unit was employed.

E. EVALUATION CRITERIA

A range of observational error measures is employed to assess each classification's effectiveness. One metric for determining a classification model's effectiveness is how well it can classify data. Here, ''accuracy'' refers to the frequency with which our model makes correct predictions. To be exact, in technical terms, means:

$$Accuracy = \frac{TP + TN}{TN + TP + FN + FP} \tag{3}$$

where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.

Precision, also known as a positive predictive value, is the number of relevant examples found among those found. It measures how many predicted positive classes include members from that class. Precision is defined by the following equation (4).

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

Specificity is sometimes used to show or assess how well a test can accurately rule out the existence of an ailment or disease state. This indicator of a test's classification accuracy is crucial in circumstances where a false positive might be exceedingly expensive. The equation defines specificity (5).

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{5}$$

In statistics, sensitivity is found by dividing the number of accurate positive predictions by the total number of positives. It's also known as the true positive rate or recall. Sensitivity is defined in Equation (6).

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{6}$$

The F-score, also known as the F1-score, measures how well a model fits the given data. The harmonic mean of the model's accuracy and recall is called the F-score, and it is calculated. Use Equation (7) to get an F1-score:

$$F1 - Score = \frac{2TP}{2TP + FP + FN} \tag{7}$$

The MSE calculates the degree of accuracy in statistical models. Between the observed and predicted values, it evaluates the average squared difference. The formula for the MSE is Equation (8). We used the MSE as a loss function to evaluate the NSGA-II in our proposed model.

$$MSE = \frac{\sum (y_i - \hat{y}_i)^2}{n} \tag{8}$$

where y_i is the i^{th} observed value, $\hat{y}_i$ is the corresponding predicted value, and n is the number of observations.

The Matthews Correlation Coefficient (MCC) is a metric used to assess the accuracy of binary classifications, especially when dealing with imbalanced classes. The analysis considers the number of correctly identified positive cases, correctly identified negative cases, incorrectly identified positive cases, and incorrectly identified negative cases. The MCC is a numerical measure that ranges from -1 to $+1$. A value of $+1$ represents perfect prediction, 0 shows no better than random prediction, and -1 indicates complete disagreement between the prediction and the observation (Equation 9).

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{9}$$

Cohen's Kappa score is a metric used to evaluate the effectiveness of machine learning classification models by quantifying the level of agreement between two raters, taking into account both the perfect agreement and the agreement that could occur by chance (Equation 10).

$$kappa = \frac{(p_o - p_e)}{(1 - p_e)} \tag{10}$$

where p_o is relative observed agreement among raters and p_e is hypothetical probability of chance agreement.

IV. EXPERIMENTAL

The goal of this work is to detect the HER2 score in breast cancer tissue images using an ideal feature selection approach by selecting the pertinent features. The main method has two steps: the first is to choose the most useful features based on the primary structure and morphological texture of HER2 images, and the second is to reduce error and improve the accuracy of HER2 score diagnosis. Although there are many feature selection approaches, choosing the appropriate one for a given job is always a challenge. For this research, a combination of several preprocessing, feature extraction, feature selection, and classifier algorithms has been set up, which are assessed using various metrics. In this part, the

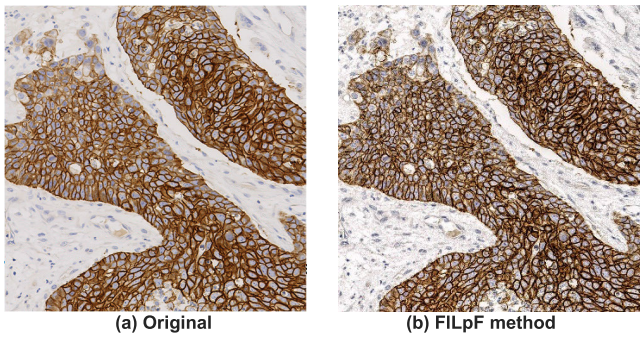


FIGURE 7. HER2 image example used by the FILpF technique.

preprocessing, training, and testing details will precede the actual implementation of the system. Experiments that use the proposed approaches are then shown and analyzed.

A. PREPROCESSING

We used the HER2SC and HER2GAN datasets to evaluate the suggested method, which included classifying HER2 images. HER2SC is a dataset that was pre-processed in several ways to diagnose the HER2 score and evaluate the proposed model. The dataset consists of WSI images. All WSI images are more than 50,000 pixels wide and 100,000 pixels tall and are available in a wide range of resolutions. For practical purposes, these images will need to be cropped. In this study, the WSI images were loaded into memory, and their dimensions were calculated using the OpenSlide package [52]. All of the WSI images were scrolled through in 2048 by the 2048-pixel window, and that window was used to create a series of stored images. Information on HER2 scores is complete in the archived images. This allows for very precise identification in addition to producing high-quality images.

Following the preparation of the WSI images, 1600 images were chosen, with 400 images chosen for each score, which is a balanced dataset. The FILpF [53] is used to improve the colors. In order to determine the rate of change of the first derivative, a Laplacian filter computes the second derivative of the picture. This indicates if the values of nearby pixels have changed because of an edge or a smooth transition. Laplacian filter kernels, located in the middle of the array, often make use of negative values arranged in a cross pattern. Values for corners might be either 0 or 1. The central value might be negative or positive. All cropped photographs are then processed using the FILpF technique to enhance their quality. Figure 7 displays an example built on the FILpF platform.

B. TRAINING AND TESTING PHASES

Our major focus is the development and training of a multi-objective optimizer model with the following objectives: i) feature reduction and optimum feature selection and ii) Improve the model's performance accuracy and minimize its error rate. Despite these challenges, it is possible to get a high degree of diagnostic accuracy for the HER2 score

TABLE 2. Set of the NSGA-II parameters.

Parameters	Value
Maximum Number of Iterations	100
Population Size	600
Crossover Percentage	70%
Mutation Percentage	40%
Mutation Rate	0.1

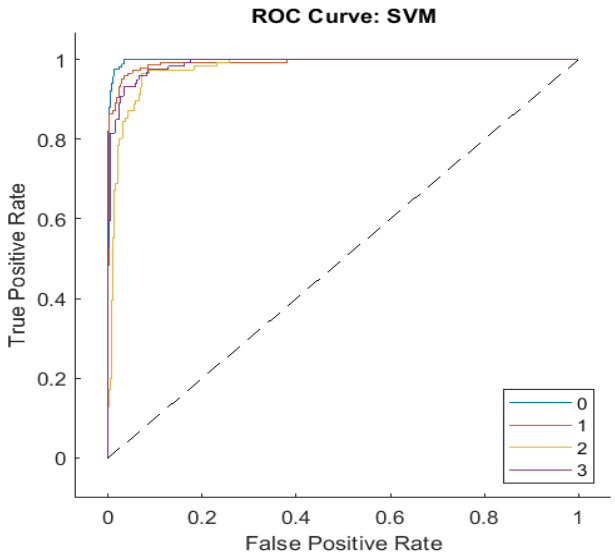


FIGURE 8. ROC curve for the SVM classifier on different scores.

using just a small variety of features, as shown in Tables 3 and 4. There are two parts to the proposed procedure at this point. In the first step, all HER2 preprocessed image features were extracted using a Resnet50 fine-tuning transfer learning model from the global average pooling layer. Each image has 2048 features that were extracted from it. The NSGA-II is developed in the second stage using two objective functions: MSE and optimal and minimal feature selection.

To determine the performance model accuracy and MSE objective function, the minimal and ideal features are chosen for each iteration and fed to the SVM classifier. The NSGA-II receives the calculated MSE back. The proposed model is processed through 100 iterations to get the lowest MSE with the best possible feature set. Setting the key parameters of NSGA-II as indicated in Table 2. Table 4 and Figure 9(a) both illustrate the impressive outcomes of the NSGA-II for the Pareto front. Following the calculation of the initial values and setting of the parameters in the third phase, all subsequent stages are used in each iteration to obtain the optimal and minimized features, as well as the reduction of the model MSE.

1. Non-Dominated Sorting of the population.
2. Calculate Crowding Distance on the population
3. Sort the population
4. Truncate the extra population

TABLE 3. The efficiency of using optimal feature selection on the proposed model's performance on the both HER2SC and HER2GAN datasets (%).

Model	NF	Rf	Accuracy	Specificity	Sensitivity	Precision	F1-score	MCC	Kappa
ResNet50+SVM (HER2GAN)	2048	100	88.86	87.35	86.34	87.42	87.79	0.821	86.12
Resnet50 + NSGA-II + SVM (HER2GAN)	894	43.65	90.8	91.18	90.19	90.31	90.03	0.852	87.91
ResNet50+SVM (HER2SC)	2048	100	90.89	97.05	89.74	90.47	89.76	0.853	87.95
Resnet50 + NSGA-II + SVM (HER2SC)	633	30.91	94.4	98.07	93.73	93.81	93.71	0.90	92.3
Resnet50 + NSGA-II + SVM (HER2SC)	549	26.81	90.75	96.98	89.98	89.96	89.93	0.851	87.88

NF= Number of Features RF= Ratio of selected features

TABLE 4. The results of the pareto front from the last iteration on the HER2SC dataset (%).

F	NF	Rf	MSE	Accuracy	Specificity	Sensitivity	Precision	F1-score
1	695	33.94	0.0616	93.84	97.96	93.42	93.55	93.45
2	549	26.81	0.098	90.75	96.98	89.98	89.96	89.93
3	633	30.91	0.0645	94.4	98.07	93.73	93.81	93.71
4	551	26.9	0.0856	91.66	97.26	91	91.03	91
5	623	30.42	0.0692	93.19	97.78	92.44	92.49	92.43
6	562	27.44	0.0794	92.39	97.52	91.51	91.5	91.47
7	566	27.64	0.0758	92.75	97.62	92.1	92.14	92.09
8	604	29.49	0.0692	93.3	97.81	92.75	92.7	92.7
9	583	28.47	0.0754	92.57	97.57	91.93	92.01	91.92
10	597	29.15	0.0699	93.11	97.75	92.37	92.49	92.41
11	594	29	0.0732	93.11	97.74	92.61	92.54	92.53
12	561	27.39	0.0842	91.69	97.26	91.07	91.11	91.05
13	592	28.91	0.075	92.71	97.61	92.18	92.2	92.14
14	560	27.34	0.0852	91.69	92.27	91.12	91.08	91.07

F= Number of front NF= Number of features RF= Ratio of selected features

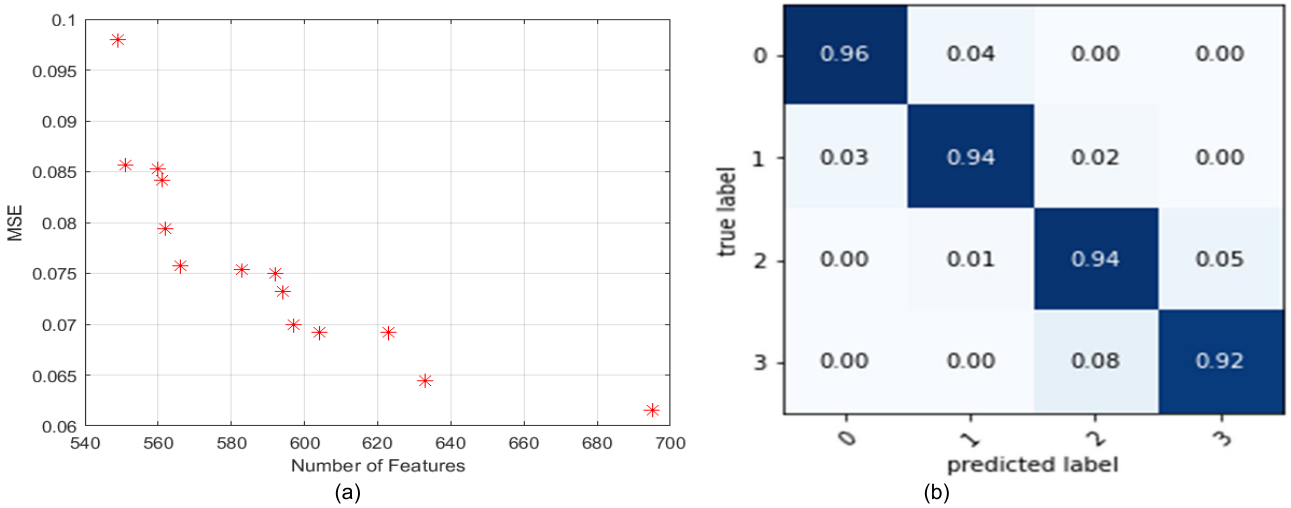


FIGURE 9. a) Pareto front of the proposed model, b) Confusion matrix of the first value of the Pareto front based on (a).

5. Non-Dominated Sorting of the new population
6. Calculate Crowding Distance on the new population
7. Sort the new population
8. Store first Front (F1)

In addition, when the selected features are fed to the SVM classifier, The system performance accuracy and MSE objective function based on a cross-validation (CV) approach [54] are evaluated. We trained and tested the SVM classifier five

times with different data using the fivefold cross-validation (CV) approach, assigning 80% of the data to train and 20% to testing each time. We averaged across all CVs in a batch of five to get the mean CV. The averaged result is sent back to NSGA-II. The summaries of all of the tests are included in Tables 3 and 4.

C. DISCUSSION

The main objective of this work is to develop a composite model that accurately classifies the HER2 score, aiming for high model accuracy. The Resnet50 model is utilized for extracting features via TL. The NSAGA-II algorithm is employed to identify the optimal and minimal set of features from a collection of retrieved features, characterized by their distinct and intricate structure. The NSAGA-II algorithm differs from traditional feature selection models in that it does not aim to pick a predetermined number of features. Instead, it is a trainable model that iteratively selects the optimal minimal number of features. This process ensures that the best features are selected gradually and systematically. Furthermore, NSGA-II has the capability to utilize feedback from a classifier in order to choose the most effective set of minimum features. The SVM classifiers are utilized in our suggested model.

Several limitations must be acknowledged in this investigation. The fundamental constraint that specialist pathologists must ascertain is the diagnosis of the HER2 score. Furthermore, in the WSI image, due to its extensive size and the presence of many HER2 types, it is imperative that an expert pathologist individually diagnose and separate each distinct HER2 score inside the image. The third constraint pertains to the limited availability of the HER2 dataset. Acquiring an adequate dataset with comprehensive labeling is exceedingly challenging.

As can be seen in Tables 3 and 4, one of the primaries aims of this study is to develop an automated trainable model that selects the best characteristics while keeping them to a minimum, thereby improving the system's performance accuracy. Table 2 shows that the model's accuracy was 88.86 and 90.89% before we used the NAGA-II optimizer for both the HER2SC and HER2GAN datasets, respectively. Using the NSGA-II optimizer for least-cost feature selection, we were able to boost the model's performance accuracy to 94.4% while utilizing only 633 features (or 30.91% of all features) for the HER2SC dataset. In addition, the model's accuracy was 90.75% while utilizing the minimal optimum feature set of 549 features, which is equal to 26.81% of all features. We improved the HER2GAN dataset's accuracy to 90.8% with 43.65% of all features. This increase in the model's ability to find HER2 scores suggests that NSGA-II is one of the best multi-objective optimizers for picking the best features from images with certain structures and morphologies.

Table 4 displays the Pareto front results of the HER2SC dataset for our proposed models. The model could provide the best 14 responses based on minimum features and minimum MSE rate after 100 iterations, as shown by the details of all metrics in Table 4 and Figure 9(a). We proved that a model's performance accuracy may reach higher than 90% by choosing less than 34% of all features. Moreover, the MSE for every response was lower than 1%. Figure 9(b) shows the confusion matrix, and Figure 8 shows the ROC curve which shows the sample response with 94.4% model performance accuracy of the Pareto front.

TABLE 5. Performance accuracy of our proposed model compared to existing deep frameworks (%).

Methods	Ratio of used features	Classification of accuracy
[9] HER2-ResNet	100	93
[19] Logistic regression SVM	100	91 averages
[14] $L_\theta, L_{sc}, L_{IoR}$	100	79.4
[25]	100	87
[20] CNN+MLP	100	83.3
[10] HER2GAN (original images)	100	90.5
[10] HER2GAN (combined generated and original images)	100	94.2
Resnet50 + NSGA-II + SVM (HER2GAN)	**43.65**	**90.8**
Resnet50 + NSGA-II + SVM (HER2SC)	**30.91**	**94.4**

TABLE 6. CNN classifier's average time test (ATT) (sec).

Model	ATT
ResNet50	0.203
ResNet50+SVM	0.0958
Resnet50 + NSGA-II + SVM	0.563

ATT = Average Time Test

Table 5 presents a comparison between the proposed model and the most advanced research. Researchers employed various ML and DL algorithms to assess the HER2 image scores. By employing a composite model comprising the transfer learning model, the multi-objective NSGA-II optimizer, and the SVM classifier, we achieved a 4% improvement in the accuracy of the system's performance for the HER2SC dataset, despite utilizing less than 34% of the total features. The HER2GAN has shown a substantial enhancement, increasing by 2%. The proposed model outperforms [10] in terms of both the number of features and model performance. Reference [10] integrated created and original images using all available features. Our combination model, in contrast to the models shown in references [14], [19], [20], and [25], which utilized ML and DL algorithms, demonstrated the ability to simultaneously decrease the number of features and substantially enhance the accuracy of the performance model.

Our proposed method only needs 100 iterations of the NSGA-II model to pick optimum minimal features, even when systems take a long time to train the model. Table 6

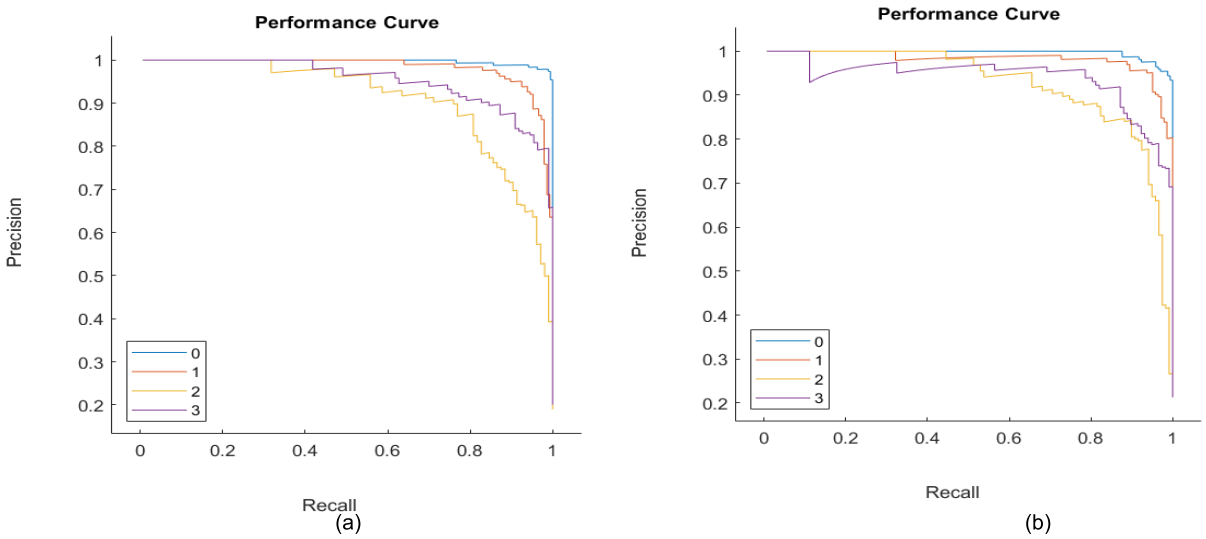


FIGURE 10. a) precision-recall curves of SVM classifier for HER2SC dataset, b) precision-recall curves of SVM classifier for HER2GAN dataset.

shows that it takes less than 0.563 seconds on average to evaluate every image.

FLOPs, short for Floating Point Operations, serve as a widely employed metric for quantifying the computational complexity of deep learning models. FLOPs provide a rapid and straightforward means of quantifying the number of arithmetic operations necessary to execute a specific computation in deep learning models. The Resnet50 model, when trained with an input size of 224*224, has a time complexity of 497 giga FLOPs. The NSGA-II algorithm has a temporal complexity of $O(MN^2)$ for creating non-dominated fronts in a single iteration, given a population size of N and M objective functions. The time complexity for training the SVM classifier is $O(Numbe\ Of\ Samples^2 * Number\ Of\ Features)$. Due to the varying number of features in each iteration of our proposed model, it is not possible to precisely calculate the time complexity of the SVM classifier.

It is evident from the results of Figure 9(a)'s Pareto front design from the last iteration that the features that were selected ranged from 540 to 700 (corresponding to 26% to 34% of all features). The Pareto front in Figure 9(a) is a complete Pareto Front, demonstrating that the NSGA-II is an appropriate approach for selecting the best features. Figure 10 depicts SVM classifier precision-recall curves of a classification model's performance at different classification thresholds for both the HER2SC and HER2GAN datasets.

V. CONCLUSION

In conclusion, it is presented in this paper how to choose the best features to improve performance in the HER2 score diagnostic problem. Based on breast cancer images, specifically HER2 images, which have a unique structure and morphological, we showed that selecting the best features may improve the accuracy of model performance. In addition, reducing the number of features reduced the time complexity of the image assessment. According to the findings, metaheuristic optimizers may be effective in determining the minimal optimal feature set. Moreover, instead of repeatedly utilizing a single-objective optimizer, we used an NSGA-II multi-objective optimizer for optimum feature selection, which boosted the model's accuracy and saved time.

In the future, it is necessary to research and compare the performance of different metaheuristic optimization algorithms, specifically when fine-tuning parameters for optimal minimum feature selection in complex structure images. Furthermore, it is necessary to optimize the parameters of the chosen metaheuristic optimizers in order to enhance their effectiveness and efficiency in picking the smallest optimal feature set.

REFERENCES

[1] M. Chivukula, R. Bhargava, A. Brufsky, U. Surti, and D. J. Dabbs, "Clinical importance of HER2 immunohistologic heterogeneous expression in core-needle biopsies vs resection specimens for equivocal (immunohistochemical score 2+) cases," *Modern Pathol.*, vol. 21, no. 4, pp. 363–368, Apr. 2008, doi: 10.1038/modpathol.3801021.

[2] C. A. Hudis, "Trastuzumab—Mechanism of action and use in clinical practice," *New England J. Med.*, vol. 357, no. 1, pp. 39–51, Jul. 2007, doi: 10.1056/nejmra043186.

[3] G. H. Vance, T. S. Barry, K. J. Bloom, P. L. Fitzgibbons, D. G. Hicks, R. B. Jenkins, D. L. Persons, R. R. Tubbs, and M. E. H. Hammond, "Genetic heterogeneity in HER2 testing in breast cancer: Panel summary and guidelines," *Arch. Pathol. Lab. Med.*, vol. 133, no. 4, pp. 611–612, Apr. 2009, doi: 10.5858/133.4.611.

[4] J. M. Chen, A. P. Qu, L. W. Wang, J. P. Yuan, F. Yang, Q. M. Xiang, N. Maskey, G. F. Yang, J. Liu, and Y. Li, "New breast cancer prognostic factors identified by computer-aided image analysis of HE stained histopathology images," *Sci. Rep.*, vol. 5, pp. 1–13, May 2015, Art. no. 10690, doi: 10.1038/srep10690.

[5] E. S. Yi, J. M. Boland, J. J. Maleszewski, A. C. Roden, A. M. Oliveira, M.-C. Aubry, M. R. Erickson-Johnson, B. L. Caron, Y. Li, H. Tang, S. Stoddard, J. Wampfler, K. Kulig, and P. Yang, "Correlation of IHC and FISH for ALK gene rearrangement in non-small cell lung carcinoma: IHC score algorithm for FISH," *J. Thoracic Oncol.*, vol. 6, no. 3, pp. 459–465, Mar. 2011, doi: 10.1097/jto.0b013e318209edb9.

[6] A. C. Wolff, M. E. H. Hammond, K. H. Allison, B. E. Harvey, P. B. Mangu, J. M. S. Bartlett, M. Bilous, I. O. Ellis, P. Fitzgibbons, W. Hanna, R. B. Jenkins, M. F. Press, P. A. Spears, G. H. Vance, G. Viale, L. M. McShane, and M. Dowsett, "Human epidermal growth factor receptor 2 testing in breast cancer: American society of clinical oncology/college of American pathologists clinical practice guideline focused update," *Arch. Pathol. Lab. Med.*, vol. 142, no. 11, pp. 1364–1382, Nov. 2018, doi: 10.5858/arpa.2018-0902-sa.

[7] M. Saha, R. Mukherjee, and C. Chakraborty, "Computer-aided diagnosis of breast cancer using cytological images: A systematic review," *Tissue Cell*, vol. 48, no. 5, pp. 461–474, Oct. 2016, doi: 10.1016/j.tice.2016.07.006.

[8] M. Saha and C. Chakraborty, "Her2Net: A deep framework for semantic segmentation and classification of cell membranes and nuclei in breast cancer evaluation," *IEEE Trans. Image Process.*, vol. 27, no. 5, pp. 2189–2200, May 2018, doi: 10.1109/TIP.2018.2795742.

[9] X. Wang, C. Shao, W. Liu, H. Liang, and N. Li, "HER2-ResNet: A HER2 classification method based on deep residual network," *Technol. Health Care*, vol. 30, pp. 215–224, Feb. 2022, doi: 10.3233/thc-228020.

[10] M. M. Mirimoghaddam, J. Majidpour, F. Pashaei, H. Arabalibeik, E. Samizadeh, N. M. Roshan, and T. A. Rashid, "HER2GAN: Overcome the scarcity of HER2 breast cancer dataset based on transfer learning and GAN model," *Clin. Breast Cancer*, vol. 24, no. 1, pp. 53–64, Jan. 2024, doi: 10.1016/j.clbc.2023.09.014.

[11] J. Arevalo, F. A. González, R. Ramos-Pollán, J. L. Oliveira, and M. A. Guevara Lopez, "Representation learning for mammography mass lesion classification with convolutional neural networks," *Comput. Methods Programs Biomed.*, vol. 127, pp. 248–257, Apr. 2016, doi: 10.1016/j.cmpb.2015.12.014.

[12] Y. Li, J. Wu, and Q. Wu, "Classification of breast cancer histology images using multi-size and discriminative patches based on deep learning," *IEEE Access*, vol. 7, pp. 21400–21408, 2019, doi: 10.1109/ACCESS.2019.2898044.

[13] T. Pitkäaho, T. M. Lehtimäki, J. McDonald, and T. J. Naughton, "Classifying HER2 breast cancer cell samples using deep learning," in *Proc. Irish Mach. Vis. Image Process. Conf.*, 2016, pp. 1–104.

[14] T. Qaiser and N. M. Rajpoot, "Learning where to see: A novel attention model for automated immunohistochemical scoring," *IEEE Trans. Med. Imag.*, vol. 38, no. 11, pp. 2620–2631, Nov. 2019, doi: 10.1109/TMI.2019.2907049.

[15] F. Özyurt, J. Majidpour, T. A. Rashid, A. Majidpour, and C. Koç, "Multi-transfer learning techniques for detecting auditory brainstem response," *Appl. Acoust.*, vol. 212, Sep. 2023, Art. no. 109604, doi: 10.1016/j.apacoust.2023.109604.

[16] J. Majidpour, A. M. Ahmed, B. A. Hassan, M. H. Abdalla, S. M. Qader, N. B. Tayfor, and T. A. Rashid, "Thermal face image reidentification based on variational autoencoder, cascade object detector using lightweight architectures," in *Practical Artificial Intelligence for Internet of Medical Things*. Boca Raton, FL, USA: CRC Press, 2023, pp. 291–303.

[17] S. K. Jameel, S. Aydin, N. H. Ghaeb, J. Majidpour, T. A. Rashid, S. Q. Salih, and P. S. JosephNg, "Exploiting the generative adversarial network approach to create a synthetic topography corneal image," *Biomolecules*, vol. 12, no. 12, p. 1888, Dec. 2022, doi: 10.3390/biom12121888.

[18] H. A. Ahmed, J. Majidpour, M. H. Ahmed, S. K. Jameel, and A. Majidpour, "Enhancing auditory brainstem response classification based on vision transformer," *Comput. J.*, vol. 2023, Nov. 2023, Art. no. bxad107, doi: 10.1093/comjnl/bxad107.

[19] R. Mukundan, "Analysis of image feature characteristics for automated scoring of HER2 in histology slides," *J. Imag.*, vol. 5, no. 3, p. 35, Mar. 2019, doi: 10.3390/jimaging5030035.

[20] S. P. Oliveira, J. Ribeiro Pinto, T. Gonçalves, R. Canas-Marques, M.-J. Cardoso, H. P. Oliveira, and J. S. Cardoso, "Weakly-supervised classification of HER2 expression in breast cancer haematoxylin and eosin stained slides," *Appl. Sci.*, vol. 10, no. 14, p. 4728, Jul. 2020, doi: 10.3390/app10144728.

[21] P. C. Neto, S. P. Oliveira, D. Montezuma, J. Fraga, A. Monteiro, L. Ribeiro, S. Gonçalves, I. M. Pinto, and J. S. Cardoso, "IMIL4PATH: A semi-supervised interpretable approach for colorectal whole-slide images," *Cancers*, vol. 14, no. 10, p. 2489, May 2022, doi: 10.3390/cancers14102489.

[22] Z. Jiang, L. Song, H. Lu, and J. Yin, "The potential use of DCE-MRI texture analysis to predict HER2 2+ status," *Frontiers Oncol.*, vol. 9, Apr. 2019, doi: 10.3389/fonc.2019.00242.

[23] H. Lu and J. Yin, "Texture analysis of breast DCE-MRI based on intratumoral subregions for predicting HER2 2+ status," *Frontiers Oncol.*, vol. 10, p. 543, Apr. 2020, doi: 10.3389/fonc.2020.00543.

[24] J. Xu, X. Luo, G. Wang, H. Gilmore, and A. Madabhushi, "A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images," *Neurocomputing*, vol. 191, pp. 214–223, May 2016, doi: 10.1016/j.neucom.2016.01.034.

[25] F. D. Khameneh, S. Razavi, and M. Kamasak, "Automated segmentation of cell membranes to evaluate HER2 status in whole slide images using a modified deep learning network," *Comput. Biol. Med.*, vol. 110, pp. 164–174, Jul. 2019, doi: 10.1016/j.compbiomed.2019.05.020.

[26] A. Aguilera, R. Pezoa, and A. Rodríguez-Delherbe, "A novel ensemble feature selection method for pixel-level segmentation of HER2 overexpression," *Complex Intell. Syst.*, vol. 8, no. 6, pp. 5489–5510, Dec. 2022, doi: 10.1007/s40747-022-00774-x.

[27] S. Verma, M. Pant, and V. Snasel, "A comprehensive review on NSGA-II for multi-objective combinatorial optimization problems," *IEEE Access*, vol. 9, pp. 57757–57791, 2021, doi: 10.1109/ACCESS.2021.3070634.

[28] J. M. Abdullah, T. A. Rashid, B. B. Maaroof, and S. Mirjalili, "Multi-objective fitness-dependent optimizer algorithm," *Neural Comput. Appl.*, vol. 35, no. 16, pp. 11969–11987, Jun. 2023, doi: 10.1007/s00521-023-08332-3.

[29] C. M. Rahman, T. A. Rashid, A. M. Ahmed, and S. Mirjalili, "Multi-objective learner performance-based behavior algorithm with five multi-objective real-world engineering problems," *Neural Comput. Appl.*, vol. 34, no. 8, pp. 6307–6329, Apr. 2022, doi: 10.1007/s00521-021-06811-z.

[30] Y. Cao, H. A. Dhahad, H. Togun, H. M. Hussen, T. A. Rashid, A. E. Anqi, N. Farouk, and A. Issakhov, "Exergetic and financial parametric analyses and multi-objective optimization of a novel geothermal-driven cogeneration plant; adopting a modified dual binary technique," *Sustain. Energy Technol. Assessments*, vol. 48, Dec. 2021, Art. no. 101442, doi: 10.1016/j.seta.2021.101442.

[31] E. Kaplan, W. Y. Chan, S. Dogan, P. D. Barua, H. T. Bulut, T. Tuncer, M. Cizik, R.-S. Tan, and U. R. Acharya, "Automated BI-RADS classification of lesions using pyramid triple deep feature generator technique on breast ultrasound images," *Med. Eng. Phys.*, vol. 108, Oct. 2022, Art. no. 103895, doi: 10.1016/j.medengphy.2022.103895.

[32] H. Liu, G. Cui, Y. Luo, Y. Guo, L. Zhao, Y. Wang, A. Subasi, S. Dogan, and T. Tuncer, "Artificial intelligence-based breast cancer diagnosis using ultrasound images and grid-based deep feature generator," *Int. J. Gen. Med.*, vol. 15, pp. 2271–2282, Mar. 2022, doi: 10.2147/ijgm.s347491.

[33] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002, doi: 10.1109/4235.996017.

[34] M. Dorigo, M. Birattari, and T. Stutzle, "Ant colony optimization," *IEEE Comput. Intell. Mag.*, vol. 1, no. 4, pp. 28–39, Nov. 2006, doi: 10.1109/mci.2006.329691.

[35] P. Luč ić and D. Teodorović, "Computing with bees: Attacking complex transportation engineering problems," *Int. J. Artif. Intell. Tools*, vol. 12, no. 3, pp. 375–394, Sep. 2003, doi: 10.1142/s0218213003001289.

[36] K. Price, R. M. Storn, and J. A. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*. Cham, Switzerland: Springer, 2006.

[37] H. Liu and R. Setiono, "Chi2: Feature selection and discretization of numeric attributes," in *Proc. 7th IEEE Int. Conf. Tools Artif. Intell.*, May 1995, pp. 388–391, doi: 10.1109/TAI.1995.479783.

[38] J. Goldberger, G. E. Hinton, S. Roweis, and R. R. Salakhutdinov, "Neighbourhood components analysis," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 17, 2004, pp. 513–520.

[39] R. J. Urbanowicz, M. Meeker, W. La Cava, R. S. Olson, and J. H. Moore, "Relief-based feature selection: Introduction and review," *J. Biomed. Informat.*, vol. 85, pp. 189–203, Sep. 2018, doi: 10.1016/j.jbi.2018.07.014.

[40] (Nov. 30, 2017). *HER2 Scoring Contest*. [Online]. Available: http://www2.warwick.ac.uk/fac/sci/dcs/research/tia/her2contest/

[41] N. Rajpoot and T. Qaiser, "HER2 challenge contest: A detailed assessment of automated HER2 scoring algorithms in whole slide images of breast cancer tissues," *Histopathology*, vol. 72, no. 2, pp. 227–238, Jan. 2018, doi: 10.1111/his.13333.
[42] L. Yang, S. Hanneke, and J. Carbonell, "A theory of transfer learning with applications to active learning," *Mach. Learn.*, vol. 90, no. 2, pp. 161–189, Feb. 2013, doi: 10.1007/s10994-012-5310-y.
[43] S. Khan, N. Islam, Z. Jan, I. Ud Din, and J. J. P. C. Rodrigues, "A novel deep learning based framework for the detection and classification of breast cancer using transfer learning," *Pattern Recognit. Lett.*, vol. 125, pp. 1–6, Jul. 2019, doi: 10.1016/j.patrec.2019.03.022.
[44] Y. Liu, G.-R. She, and S.-X. Chen, "Magnetic resonance image diagnosis of femoral head necrosis based on ResNet18 network," *Comput. Methods Programs Biomed.*, vol. 208, Sep. 2021, Art. no. 106254, doi: 10.1016/j.cmpb.2021.106254.
[45] A. V. Ikechukwu, S. Murali, R. Deepu, and R. C. Shivamurthy, "ResNet-50 vs VGG-19 vs training from scratch: A comparative analysis of the segmentation and classification of pneumonia from chest X-ray images," *Global Transitions Proc.*, vol. 2, no. 2, pp. 375–381, Nov. 2021, doi: 10.1016/j.gltp.2021.08.027.
[46] N. Srinivas and K. Deb, "Muiltiobjective optimization using nondominated sorting in genetic algorithms," *Evol. Comput.*, vol. 2, no. 3, pp. 221–248, Sep. 1994, doi: 10.1162/evco.1994.2.3.221.
[47] L. Wang, Y. Gao, J. Li, and X. Wang, "A feature selection method by using chaotic cuckoo search optimization algorithm with elitist preservation and uniform mutation for data classification," *Discrete Dyn. Nature Soc.*, vol. 2021, pp. 1–19, Jun. 2021, doi: 10.1155/2021/7796696.
[48] P. B. Myszkowski, M. Laszczyk, and J. Lichodij, "Efficient selection operators in NSGA-II for solving bi-objective multi-skill resource-constrained project scheduling problem," in *Proc. Federated Conf. Comput. Sci. Inf. Syst. (FedCSIS)*, Sep. 2017, pp. 83–86.
[49] J. Suykens and T. Van Gestel, *Least Squares Support Vector Machines*. Singapore: World Scientific, 2002.
[50] H. S. Dadi and G. K. Mohan Pillutla, "Improved face recognition rate using HOG features and SVM classifier," *IOSR J. Electron. Commun. Eng.*, vol. 11, no. 4, pp. 34–44, Apr. 2016, doi: 10.9790/2834-1104013444.
[51] S. K. Al-Salihi, S. Aydin, and N. H. Ghaeb, "SWFT: Subbands wavelet for local features transform descriptor for corneal diseases diagnosis," *TURKISH J. Electr. Eng. Comput. Sci.*, vol. 29, no. 2, pp. 875–896, Mar. 2021, doi: 10.3906/elk-2004-114.
[52] A. Goode, B. Gilbert, J. Harkes, D. Jukic, and M. Satyanarayanan, "OpenSlide: A vendor-neutral software foundation for digital pathology," *J. Pathol. Informat.*, vol. 4, no. 1, p. 27, Jan. 2013, doi: 10.4103/2153-3539.119005.
[53] S. Paris, S. W. Hasinoff, and J. Kautz, "Local Laplacian filters: Edge-aware image processing with a Laplacian pyramid," *ACM Trans. Graph.*, vol. 30, no. 4, 2011, Art. no. 68.
[54] D. M. Allen, "The relationship between variable selection and data agumentation and a method for prediction," *Technometrics*, vol. 16, no. 1, pp. 125–127, Feb. 1974, doi: 10.1080/00401706.1974.10489157.

TARIK AHMED RASHID (Member, IEEE) received the Ph.D. degree in computer science and informatics from the College of Engineering, Mathematical and Physical Sciences, University College Dublin (UCD), in 2006. From 2006 to 2007, he was a Postdoctoral Fellow with the School of Computer Science and Informatics, UCD. In 2017, he joined the University of Kurdistan Hewlêr. His research interests include three fields. The first field is the expansion of machine learning and data mining to deal with time series applications. The second field is the development of DNA computing, optimization, swarm intelligence, and nature-inspired algorithms and their applications. The third field is networking, telecommunication, and telemedicine applications. He is a member of ACM.

JAFAR MAJIDPOUR received the master's degree in software engineering from East Azarbaijan Science and Research Branch, Tabriz, Iran, in 2013. Since 2013, he has been with the Department of Software and Informatics Engineering, University of Raparin, Iraq, as a Lecturer. His primary research interests include computer vision, machine learning, deep learning, and image processing.

RAJERMANI THINAKARAN received the Bachelor of Science degree in computer science from Universiti Teknologi Malaysia (UTM), Malaysia, in 1995, the master's degree in IT from Universiti Kebangsaan Malaysia (UKM), in 2012, and the Ph.D. degree from UTM, in 2019. She is currently an Associate Professor with the Faculty of Data Science and Information Technology, INTI International University, Negeri Sembilan, Malaysia. She supervises both bachelor's and master's students (master's and Ph.D. levels). She has more than 30 papers in international and local journals, international and national conference proceedings, as well as book chapters and lecture notes. Her research interests include artificial intelligence, assistive technology in empowering disabled students, elearning, and gaming ranging from theory to design to implementation. She also serves as a member of the Editorial Board and a Technical Reviewer for local and international journals and conferences.

MALATHY BATUMALAY is currently focuses on the development of D-shaped fiber, microbottle resonators (MBRs) and also kretschmann surface plasmon resonance for food sensing, relative humidity, and gas sensing. The development of sensors also emphasizes on the type of nanostructure coating materials to enhance the sensing performance. She focused on the research of photonics engineering, fiber optics, and lasers. In her previous research work, she developed fiber optics into sensors to monitor the relative humidity and chemical concentration.

DESHINTA ARROVA DEWI received the Ph.D. degree in computer science from the National University of Malaysia (UKM), Malaysia, in 2019. She joined INTI International University in 2010, where she is actively conducting teaching, learning, and research. She was appointed as an Associate Professor in 2023, and the Managing Editor for the *Journal of Data Science*. Her research interests are in artificial intelligence, software engineering, data science, and education.

BRYAR A. HASSAN pursued the bachelor's and master's degrees in software engineering from the University of Southampton, and the joint Ph.D. degree in information technology. He is a renowned Assistant Professor in computer science with Charmo University. He is in the list of the World's Top 2% Scientists Rankings 2023 (by Stanford University and Elsevier). He has worked as a Software Developer and Entrepreneur for several years. He is currently a renowned Assistant Professor, he inspires students through innovative teaching and cutting-edge research in artificial intelligence and optimization algorithms. His research interests include artificial intelligence, optimization algorithms, medical computing, semantic web, and NLP.

HABIBOLLAH DADGAR is currently a Nuclear Medicine Physicist with the Nuclear Medicine and Molecular Imaging Research Center, PET-CT Section, Razavi Hospital, with more than ten years of experience in internal dosimetry and clinical assessments of theranostics. His research interests include the application of deep learning and AI in medical imaging research.

HOSSEIN ARABI received the Ph.D. degree in medical physics from the University of Geneva, in 2016. From 2016 to 2021, he was a Postdoctoral Fellow in medical image processing and artificial intelligence with the University of Geneva, where he was a Senior Research Assistant, in 2021. He is currently a Senior Research Assistant with the Department of Nuclear Medicine, Geneva University Hospital. His research interests include artificial intelligence, medical image processing, PET & SPECT instrumentation, image reconstruction, and Monte Carlo simulation.