ISTP

RESEARCH ARTICLE

A Bio-Inspired Method for Breast Histopathology Image Classification Using Transfer Learning

R. K. Chandana Mani,[1] **J. Kamalakannan,**[1] **Y. Pandu Rangaiah,**[2] **and Sesham Anand**[3]

[1]Vellore Institute of Technology, Vellore, TN, India

[2]G. Pullaiah College of Engineering and Technology, Kurnool, Andhra Pradesh, India

[3]Maturi Venkata Subba Rao Engineering College, Hyderabad, India

Abstract: Breast cancer is one of the deadly cancer among the female population and still a developing area of research in the field of medical imaging. The fatality rate is more in patients who are not early diagnosed and are given delayed treatment. Hence, researchers are keeping their lot of efforts in developing breast cancer detection systems that could provide accurate diagnosis in the initial stages which are relied on medical imaging. Deep learning is offering key solutions to overcome many image classification tasks. Although deep learning techniques have extended their root to many medical fields even it suffers from the problem of lack of sufficient data. Convolutional neural networks are more preferred for medical image classification tasks. In this paper, we propose a transfer learning method that overcomes the issue of insufficient data. We use a familiar pre-trained network VGG-16 (Visual Geometric Group) + with logistic regression as a binary classifier. Since hyper-parameters of every CNN have a closer impact on the performance of the entire deep learning model, our method focuses on optimizing hyper-parameters using particle swarm optimization which is a bio-inspired algorithm. The proposed model performs the classification of breast histopathology images into benign and malignant images and produce better results. We use BreakHis Dataset to test our method and achieve an accuracy of around 96.9%. The experimental results show that this study has improved performance metrics when compared to other transfer learning methods.

Keywords: bio-inspired algorithms; breast cancer; convolutional neural networks; data augmentation; swarm optimization

I. INTRODUCTION

Recent studies summarized that breast cancer is one of the most prevalent forms of cancer, among women worldwide. From the knowledge of the cancer mortality survey, this disease is the leading cause of death with 5,24,000 deaths and 2.4 million new cases per year [1]. Early diagnosis and accurate medical treatment play an important role to minimize the mortality rate. Generally, inevitable mutations of breast cells result in the development of a mass called tumor or cyst. These tumors are usually categorized into two classes; carcinoma namely malignant and non-carcinoma namely benign [2]. Few blossoming methods for purpose of screening and diagnosis of tumors are magnetic resonance imaging [MRI], mammography, and histopathology [3]. Despite all other approaches, cancer detection from histopathology is the gold standard. Generally, pathologists examine the cell shapes, tissue structure, and density to define malignancy degree and cancerous locality, by analyzing a narrow slice of biopsy sample with a microscope. This manual analysis needs enormous domain knowledge; also, it is an error-prone and time-taking process [4–6]. By using whole-slide digital scanners, histological slides of the suspected tissues can be computerized and preserved in form of a digital image. Digitalization figures out other advantages such as better visualization, quick storage, and analysis of images that in turn lead to automatic image analysis methods [7]. In

recent years, the use of deep learning in conjunction with image processing techniques has grown quickly. To address challenges like identifying specific areas of infected organs, defending the classification of tumors, corresponding treatment, and disease stage, deep learning-based algorithms are particularly adaptive. Particularly, deep learning techniques based on convolutional neural networks (CNNs) have made room for integrating CAD methodologies with image processing jobs [8]. In the case of medical data analytics, AI and ML techniques are frequently for speech-based COVID-19 detection from speech signals [9,10].

The use of deep learning in conjunction with image processing techniques has grown quickly. To address challenges like identifying specific areas of infected organs, defending the classification of tumors, corresponding treatment, and disease stage, deep learning-based algorithms are particularly adaptive. Particularly, deep learning techniques based on CNNs have made room for integrating CAD methodologies with image processing jobs [11–14]. The origin of the aforementioned methods directed the path to researchers in developing CAD models assisted by CNN for several medical applications such as breast cancer, skin cancer, lung cancer, brain cancer, and liver cancer [15]. In contradiction with hand-crafted methods, CNNs have the ability to learn features from the given medical images. Ting *et al.* [16] developed a CAD system by cascading pre-trained networks and obtained an accuracy of around 90.5%. Xu *et al.* [17] performed feature extraction by using fine-tuned CNN for image classification regarding colon cancer and brain tumor. However, there is a lot of space provided for the development of pre-trained networks in the context of breast cancer

Corresponding author: Y. Pandu Rangaiah (e-mail: y.pandurangaiah@gmail.com).

histopathological images. As the task of breast histological images is much complicated, recent studies started using pre-trained networks such as AlexNet, VGG-16, and VGG-19. For the problem of classification, these pre-trained networks are accompanied by other classifiers like logistic regression, KNN (K-Nearest Neighbor), SVM (Support Vector Machine), and decision trees.

Although the fast interventions of CNN models are increasing day by day, in point of the accuracy of the proposed CNN model, the hyper-parameters have a great impact on the performance. Selecting feasible hyper-parameters can be considered as an important step during the training of the CNN model. Hence, it is required to undergo optimization of CNN's hyper-parameters to achieve good performance and success rate. For the process of optimizing the CNN model, bio-inspired algorithms can be used to produce desired optimal hyper-parameters. Particle Swarm Optimization (PSO) [18], Artificial Bee Colony, Gray wolf optimization, and Ant Colony Optimization are some of the familiar algorithms used for optimizing deep learning models [18–21].

In this study, we introduce a novel method of CNN for breast image classification. Since training the CNN model from scratch is a time-taking process, we use the concept of transfer learning. Transfer learning by using VGG-16 with logistic regression as a binary classifier is the major part of the proposed classification method. In this way, this study decreases time complexity for training the entire network and makes it more reliable. This developed model is further optimized by using the PSO algorithm. The desired classification model produces better results and also avoids the over-fitting problem. The major contributions of this study are listed below

- Optimizing hyper-parameters of CNN using PSO algorithm
- An optimized CNN architecture to classify breast histopathological images
- Comparison of the proposed model with existing CNN models used for breast image classification.

The remainder of this paper is organized as follows. Related works are discussed in section II. Section III denotes the preliminary concepts, and the proposed methodology is explained in section IV. Section V illustrates the dataset. Performance metrics and results are described in section VI. In last, we conclude our study in section VII.

II. RELATED WORKS

For the last four decades, CAD systems for cancer detection have been the developing field of research with its intrinsic complexities of histopathological images, but it is yet a tedious task. Considering the classification of histopathological images into malignant and benign subtypes, many techniques have been developed in the leaflet of medical image classification. The classification of histological images has been reviewed precisely with their relevant works of research in this section.

In recent years, deep learning methodologies have been widely used in the classification of images. Even they are applied in various fields that include image recognition, drug discovery, automatic speech recognition, and bioinformatics [22,23]. The profound AlexNet CNN [24] was developed in the year 2012, which has shown admirable results in an application with a dataset named ImageNet. This success invited deep learning-relied approaches that used CNN to develop a path of attractive performance. Thus it extended its root, even to the area of medical tasks [25].

Spanhol *et al.* [26] investigated the pre-trained LeNet model that relies on a deep learning framework for the classification of breast cancer. The classification accuracy was 72%, in this method. In addition, they used a modified AlexNet model that includes two fully connected and three convolutional layers to train the network. This lead to the improved classification accuracy of 80.8%.

Wei *et al.* [27] developed an enhanced CNN architecture that inspects individual image labels as former knowledge as well as information of subclasses of the images. The authors reported that resultant CNN can provide improved binary classification in learning features. Rao *et al.* [28] have proposed a modified CNN that includes a fully connected layer, convolutional layer, and pooling layer. For the task of BreakHis Image classification, several hyper-parameters were examined to figure out the most appropriate framework. This study reported that binary image classification accuracy of 90%.

Simonyan *et al.* [29] proposed the architecture of VGG which lead to a new era of CNN methods. In a recent study, several architectures based on VGG lead to excellent results in the analysis of histopathological images [30,31]. Various cancer classification methods such as skin cancer [32], colorectal cancer [33,34], and breast cancer [29] have made use of VGG-19. In addition, fundus image classification has taken the opportunity to use the singular value decomposition along with principal component analysis complemented VGG-19 [35].

A magnification independent method was developed by Bayramoglu *et al.* [36] that has dual CNN architectures namely multi-task CNN and single-task CNN. Malignancy of the image and magnification factor was evaluated by multi-task CNN, whereas malignancy was evaluated by a single-task CNN model. The range of accuracies that used multi-task CNN methodology was between 80.6% and 83.3%; further, the range of accuracies that used single-task CNN methodology was between 82.1% and 83.0%.

Hamindineko *et al.* [37] developed a Nottingham grading system that classifies the image samples into three subclasses. Zhang *et al.* [38] investigated a class Kernel principal component analysis (KPCA) for feature extraction. Further, a KCPA method is trained on the individual feature that is extracted. The same procedure is repeated for every image in the provided dataset; thus eventually, the pre-trained KPCA architecture has the ability to make the decisions. This study reported the accuracy of binary classification as 92%.

Sudarshan *et al.* [39] have proposed a novel approach for the analysis of histopathological images known as multiple instance learning (MIL). This study uses two different types of scenarios such as Patient as a bag and Image as a bag. The authors in this study examined several methods of MIL namely algorithms based on Diversity density, Seminal axis Parallel rectangular method, SVM, K-Nearest Neighbor, CNN, and non-parametric algorithm. Their study claimed the accuracies of binary image classification ranging from 83.4% to 92.1%.

Like other neural networks, even DL-based models suffer from the issues such as multiple local optima, lack of hyper-parameter tuning, and increased time complexity. To overcome the above problems, optimizing hyper-parameters and the preferred NN architecture has become a challenging task. The evolutionary algorithms are widely used in regular optimization problems in the area of medical diagnosis that cover the selection of cancer pathway genes [40], estimating ultrasonic echo [41], DNA micro-array's classification [42], reducing bioinformatics data dimensionality [43], localization of retinal blood vessels [44], and detection of

diabetes disease [45]. The aforementioned algorithms have been not frequently utilized in optimizing outputs of feature extraction, segmentation, and classification of CNN. Henceforth, our work interrogates to enforce hybridization of the proposed CNN architecture that classifies breast histopathological images with help of evolutionary algorithms.

To optimize the hyper-parameters and kernel, Rubio *et al.* [46] proposed a novel method that uses an evolutionary approach. In addition, FABOLAS [47], which was Bayesian optimization-based method, was designed for hyper-parameter selection in architectures of SVM and CNN. Stoean *et al.* [48] proposed a newly designed method that fine tunes the convolutional layers by taking the advantage of evolutionary algorithms. To optimize the hyper-parameters of the LSTM model, Riegler *et al.* [49] proposed a structure that makes use of RS, GS, and BAO approaches.

Senthilpandi & Mahadevan [50] used CNN for brain tumor detection and diagnosis. Renith & Senthilselvi [51] improved the performance of diabetic retinopathy detection using deep learning network. Mohammed Thaha *et al.* [52] proposed extended CNN for brain tumor detection and segmentation. Deep learning network plays vital role in all area of research such as food adulteration [53], noise removal [54], mobile phone recycling process [55], and sign language understanding [56]. The optimization algorithm provides an optimized solution in all area of researches such as noise removal [57,58], micro-grid protection [59], food adulteration [60], etc.

From the above discussion, it is clear that the average accuracy of the machine learning-based models is not up to the mark in all the datasets. Additionally, the hyper-parameters are selected manually, and no optimization techniques are applied. Bio-inspired techniques have been used in the applications of signal processing more specifically speech processing for selecting the best parameters. This problem has been taken in this paper to perform hyper-parameter optimization regarding CNN that may produce excellent results in the field of medical diagnosis. As the throughput of CNN is directly related to efficient disease classification, in this study we carry our work to improvise the performance of CNN by optimizing it with help of bio-inspired algorithm; thus, our study will be used to expertise performance of CNN used for breast histopathological images.

III. PRELIMINARIES

In this section, we explore the background concepts used for our proposed model. We clearly explain the basic CNN model, basic idea behind transfer learning, and standard form of PSO algorithm.

A. CNN

In a recent study, CNNs achieved enormous adoration in the leaflet of deep learning [61,62]. CNNs are a special variant of ANNs that are designed to handle a large amount of multi-dimensional data. The primary elements of CNN are as follows: convolutional layer, fully connected, and pooling layer [63].

The images are considered as input to the convolutional layer that further produces a corresponding feature map. The major parameters of this layer are the number of filters, filter width, and stride [64]. The * symbol defines the convolutional operation. The feature map x*(t)* is obtained with given input *I(t)* when applied with the filter *K(a)*. The following eq. 1 represents the resultant
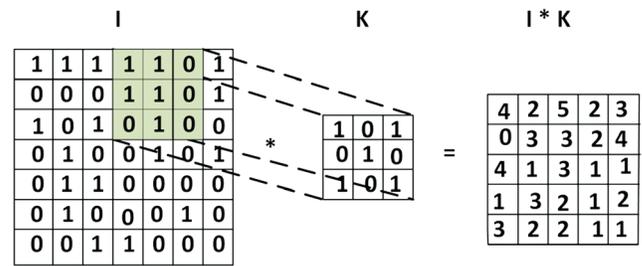
Fig. 1. Convolution operation in CNN.

matrix after the convolution operator is applied on the input feature map. The below Fig. 1 clearly exhibits the convolutional operation more precisely.

$$x(t) = (I * K)(t) \qquad (1)$$

The most common activation function that is used for CNNs is rectified linear unit (RELU) that turns the input values, which are in negative form to zero. The below equation represents the RELU activation function in mathematical form, where x is the input to the neuron. Even other activation functions, like sigmoid, leaky RELU, tanh, etc.., can be used depending on the designed architecture of CNN.

$$f(x) = \max(0,x) \qquad (2)$$

The primary motive of the pooling layer is dimensionality reduction of the given input image. This layer is infused in between RELU and convolutional layers and has complete authority on spatial invariance. Based on the requirement of the CNN model, max-pooling or average pooling operation is performed at this layer. Figure 2 clearly illustrates the max-pooling. As learnable parameters are not present in the pooling layer, it is often referred to as a component of the convolutional layer.

The fully connected layer is located beside the pooling layer and designed in such a way that all neurons in the previous layer are connected to all neurons in the FC layer. The final output is determined by the Loss function associated with the FC layer, which helps in identifying the actual and the predicted labels. Softmax is a highly recommended loss function for most CNN models.

The problem of over-fitting is resolved by regularization methods. Two prominent regularization methods are Drop-connect and Drop-out [65]. Figure 3 demonstrates a CNN model that holds the fundamental components [66]. VGg-16 [29], ResNet [67], AlexNet [13], and GoogleNet [68] are most considerable variants of CNN models when compared to other architectures. The basic hyper-parameters of CNN are learning rate, mini-batch size,
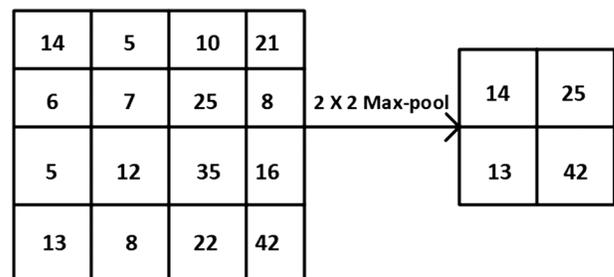
Fig. 2. Max pooling operation in CNN.
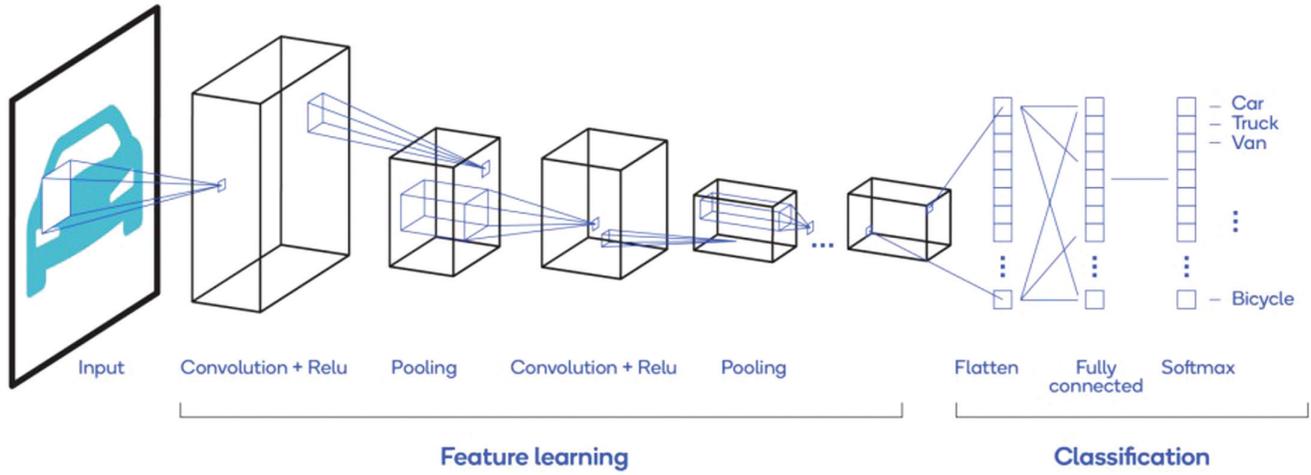
Fig. 3. CNN block diagram.
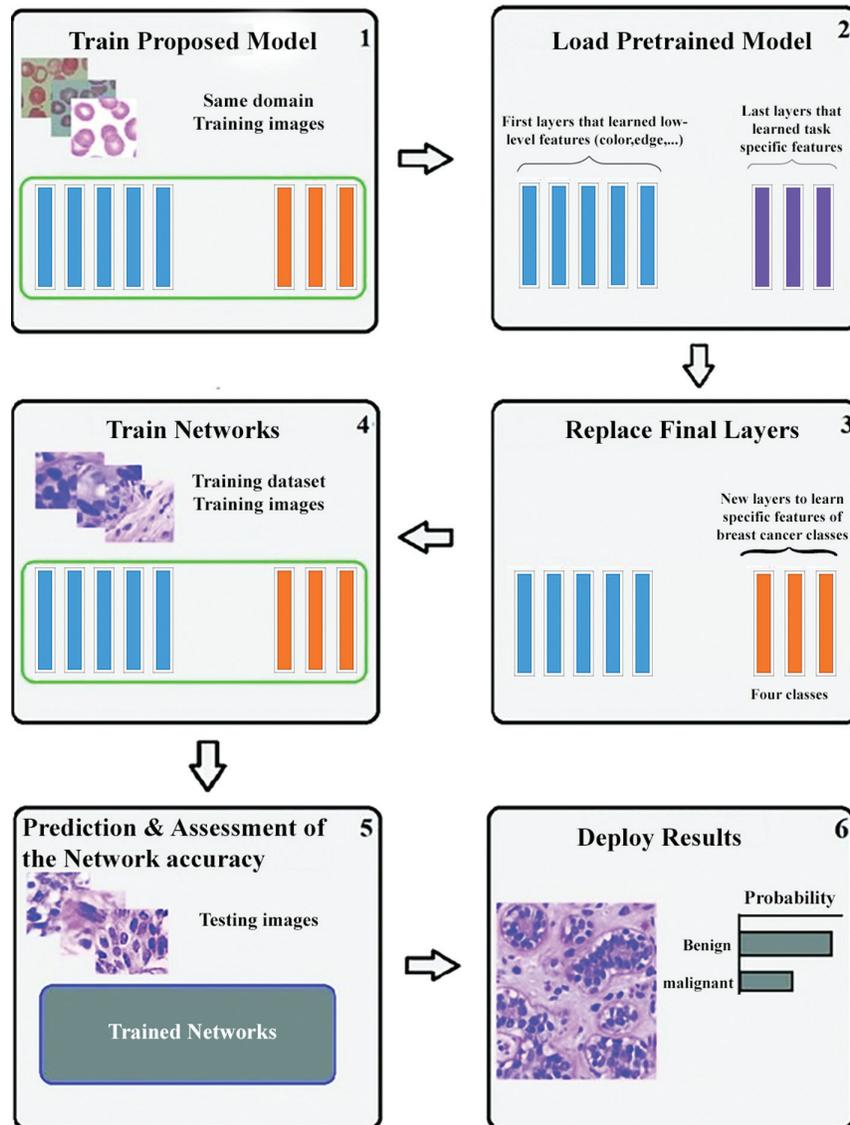
Fig. 4. Transfer learning pipeline.

regularization parameter, maximum epochs, momentum, and shuffle type.

B. TRANSFER LEARNING USING VGG-16

The provided datasets are small to make the network learnable and produce more efficient results of accuracy. The solution to overcome this issue is achieved with the help of transfer learning. The process of dumping the information from the precedent source to the target is known as transfer learning [69–71].

The steps involved in the process of transfer learning are very simple and easy to implement them for any application. Initially, the pre-trained model of ConvNet is loaded. This model is further classified into the following categories. The first class incorporates layers for minimum standards of learning such as edge, shape, and color, whereas the high-level standards of learning are included in the second class. Considering class 1, some particular layers are merged with it and applied for general problems like cancer classification. The resultant model is used to train the available training dataset. In the next step, the pre-trained network model is applied to test data. Finally, accuracy is calculated by the task of transfer learning. The architectural flow diagram of transfer learning is denoted in below Fig. 4 [72].

VGG-16 is a pre-trained network model proposed by simonyan *et al.* [29]. This model achieved 92.7% classification accuracy for the ImageNet dataset. VGG-16 can be used for the tasks of pattern detection and image recognition. The terminology 16 indicates that it holds 16 learnable layers. In our study, the BreakHis dataset with a 7909 image sample is the target that is lesser when compared to the Source dataset ImageNet.

C. PSO

In 1995, Kennedy and Beernaert proposed a novel evolutionary algorithm, namely PSO [73]. The cognitive and social behavior of the swarms is the motivation of this algorithm. This method is influenced by the flying of the herd of birds. According to the terminologies of PSO, a particle is defined as an individual solution, and a swarm is a collection of total solutions. Generally, PSO is associated with a particle that includes velocity and position [74], as explained in Fig. 5. Considering the present position and velocity of the particle, the next position of the particle is updated.

In this optimization algorithm, all particles recollect and utilize their past best position ($p_{best i}$) and the global best ($g_{best i}$) in the swarm. Using the below equations, the position and velocity of the particles are updated.

$$\text{Vel}_i^{(t+1)} = w.\,\text{Vel}_i^t + c_1 r_1 (p_{best}^t - x_i) + c_2 r_2 (g_{best}^t - x_i) \qquad (3)$$

$$x_i^{(t+1)} = x_i + \text{Vel}_i^{(t+1)} \qquad (4)$$

In the above equations, $x_{i,d}$, and $\text{Vel}_{i,d}$ represent the position and velocity of the particles, i_{th} particle's best position is given by $p_{best i}$, whereas $g_{best i}$ represents the current best position of the particle, where d denotes the D dimensional space. The terms c_1 and c_2 are considered as learning factors that represent the social behavior of the swarms. The parameters r_1 and r_2 are the randomly generated integers that range between 0 and 1. The major advantage of PSO is that its convergence rate is higher than Genetic Algorithms [75]. This characteristic of PSO helps in finding optimal CNN models.
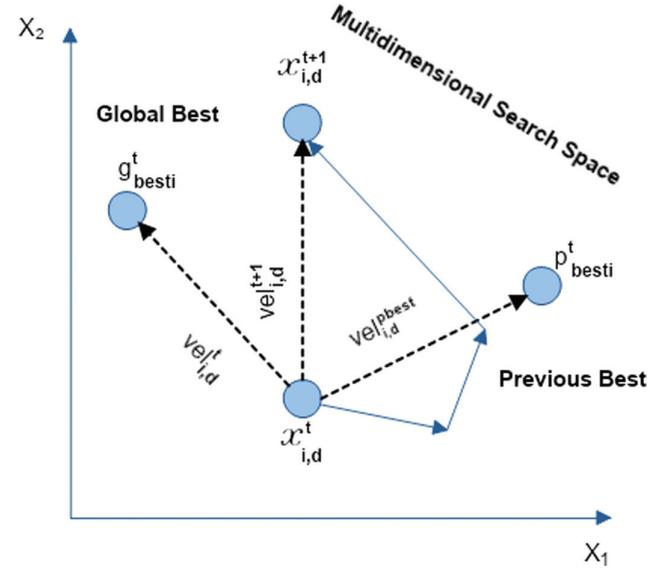
Fig. 5. Parameter updation in PSO.

IV. THE PROPOSED METHOD

In this section, we discuss the proposed working procedure of the developed model and also clearly demonstrate the correlation between PSO & CNN.

A. WORKFLOW

The recent pieces of literature have determined that the CNN model shows promising results in the problems of medical image classification. We employ a novel approach for building the CNN model by the task of transfer learning. Using pre-trained networks for transfer learning tasks is a simple methodology. Adjusting the accurate set of hyper-parameters is a challenging task since dynamic relation holds among the parameters. The process of selecting hyper-parameters could be solved by considering it as an optimization problem. For this, robust optimization algorithms that find the global optimum solutions are necessary for selecting hyper-parameters, as they consequently affect the entire performance of the CNN model.

The important steps of the proposed method are listed in algorithm 2. The familiar BreakHis dataset of breast histopathological images is loaded as the primary step of this algorithm. Further, the data are split into training set and test set. The training dataset is then divided into trainopt and testopt datasets. In the next step, the pre-trained network VGG-16 model is loaded for the task of transfer learning. These new datasets along with the basic model are considered as input parameters for the PSO algorithm. In the training process of the proposed CNN model, the optimized hyper-parameter values obtained from the algorithm are used as parameters of the training. The loss function can be used to evaluate the fitness of the given algorithm. Equation 3 represents the mathematical expression of the loss function.

$$\text{Loss} = \sum_{j=1}^{N} y_{ij} \log(p_{ij}) \qquad (5)$$

In the above expression, *N* is the number of classifications categorized, p_{ij} is the prediction probability that sample *i* belongs to
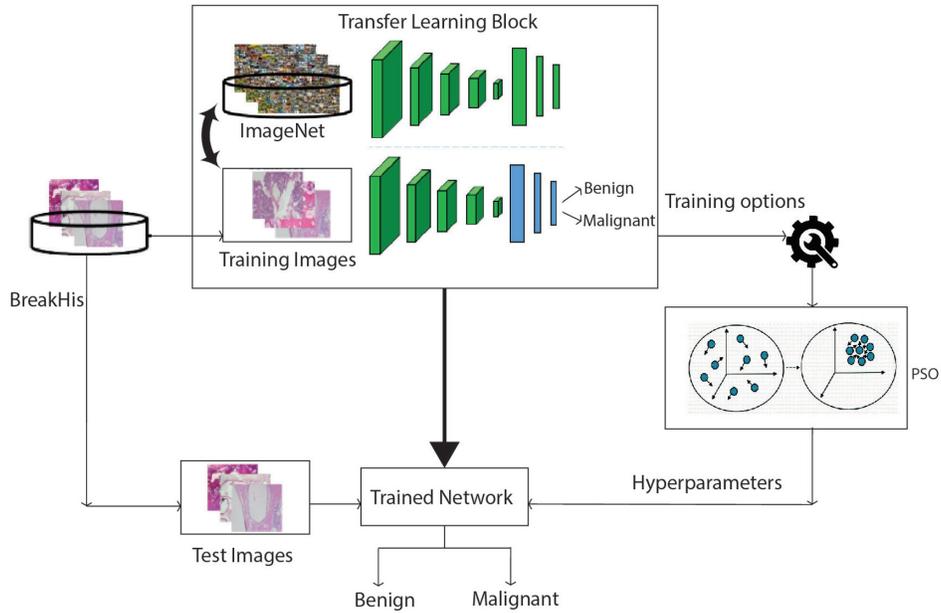
Fig. 6. The workflow diagram of the proposed model.

category j and y_{ij} is the integer 0 or 1 (0 if i and j belongs to the same class, else 1).

The workflow of the proposed methodology is described in the following Fig. 6. Once the termination criteria are satisfied, the algorithm returns the best solution that is optimal values for hyper-parameters. Then, new training is performed with the resultant optimal hyper-parameter values with the initial training dataset loaded on the VGG-16 model. Further, the test dataset is used to calculate the overall performance of the proposed method.

B. MAPPING PSO & CNN

In this study, we apply PSO for parameters of CNN to derive optimal hyper-parameters, so it is necessary to understand how we correlate our neural architecture with the evolutionary algorithm.

Algorithm 1:   The proposed optimized CNN algorithm for breast histopathology image classification

---

Input : Breast$_{\text{training}}$: training set of BreakHis dataset

        Breast$_{\text{test}}$: testing set of BreakHis dataset

Output: Classification of BC images

Step 1: Load Breast$_{\text{trainig}}$ and further split it into Train$_{\text{BC}}$ and Validation$_{\text{BC}}$

Step 2: Load the VGG-16 model

Step 3: Send Train$_{\text{BC}}$ and Validation$_{\text{BC}}$ and pre-trained VGG-16 model to the algorithm 2 as functional parameters.

Step 4: Initialize vectors of hyper-parameters.

Step 5: Get optimized hyper-parameter values of VGG-16 obtained by algorithm 2.

Step 6: Embed the resultant optimized parameters as Hyper$_{\text{opt}}$ into the pre-trained VGG-16 model.

Step 7: Train the network with Breast$_{\text{training}}$

Step 8: Evaluate the performance of the network using Breast$_{\text{test}}$.

Every vector X of the given particle P denotes a CNN, and dimensions residing in vector X denote hyper-parameters of the proposed CNN model. Variables of vector X, x_1, $x_2 \ldots x_{15}$,

---

Algorithm 2:   PSO algorithm for selection of optimal hyper-parameters of proposed CNN model

---

Input: t_m : maximum number of iterations

        Train$_{\text{BC}}$: training set

        Validation$_{\text{BC}}$: validation set

        N_s: swarm size

Output: optimal hyper-parameters

Step 1: Initialize the parameter set of PSO algorithm.

Step 2: Randomly generate an initial population x_i using Table I.

Step 3: for i=0 to N_s

        calculate fitness(x_i) using Equation 2

        $p_{\text{besti}} = x_i$ // Particle best position

        $g_{\text{besti}} = p_{\text{besti}}$ // swarm best position

    end for

Step 4: for t = 1 to t_m

        for i = 1 to N_s

            update velocity of the particle Vel$_i$ using eq. (1)

            update position of the particle using eq. (2)

            if fitness(x_i) < fitness (p_{besti}) then

            update best position of particle : $p_{\text{besti}} = x_i$

            endif

      if fitness(p_{besti}) < fitness(g_{besti})

        update best position of swarm : $g_{\text{besti}} = p_{\text{besti}}$

      end if

    end for

end for

Table I. Particles in PSO mapped to CNN

Variable	Hyper-parameter mapped in CNN
x_1	Number of filters in the first layer of convolution
x_2	Filter size of convolutional layer-1
x_3	Activation function used in convolutional layer-1
x_4	Type of pooling layer-1
x_5	Number of filters in the second layer of convolution
x_6	Filter size of convolutional layer-2
x_7	Activation function used in convolutional layer-2
x_8	Type of pooling layer-2
x_9	Number of neurons in FC layer-1
x_{10}	Activation function used in FC layer-1
x_{11}	Drop-out of FC layer-1
x_{12}	Number of neurons in FC layer-2
x_{13}	Activation function used in FC layer-2
x_{14}	Drop-out of FC layer-2
x_{15}	learning rate

Fig. 7. Breast histopathological images from BreaKHis dataset with magnification factors 40×, 100×, 200×, and 400×.

Table II. BreakHis dataset image distribution

	Class	
Magnification factor	**Benign**	**malignant**
40×	625	1370
100×	644	1437
200×	223	1390
400×	588	1232

correspond to hyper-parameters of CNN. Table I lists out the relation between variables and hyper-parameters, with their distinct purpose.

As we know that vector X is the representation of CNN's hyper-parameters, hence every dimension of it has a unique meaning and also varying range. Initially, a few of them are denoted by integers, like the number of convolutional layers, number of FC layers, and number of neurons. Further, Kernel size, pooling type, and type of activation function are presented by discrete values. We may encode these discrete values with any integer, for instance in the context of activation functions, Tanh is denoted by 1, Sigmoid with 2, and so on. The remaining hyper-parameter such as learning rate and drop-out is considerably real numbers. From Table I, we clearly understand the particle encoding mechanism for different dimensions. To update the velocity and position, we use Eq. (3) and Eq. (3). Additionally, we decode them with hyper-parameters of CNN to get the required results.

V. DATASETS

A. BreakHis DATASET

The proposed method is employed on the publicly available dataset, namely BreakHis to evaluate the performance of the model. This dataset comprises 7909 breast histopathological images which are collected from 82 patients. In this, 2480 images are from the class of benign, whereas 5429 images are from the class of malignant. These acquisition of breast images are with four types of magnification factors such as 40X, 100X, and 200X. These samples are collected from breast biopsies and then performed H&E staining [76]. The original images of this dataset are the 3-channel RGB images of the pixel size 700×480. The samples of breast histopathology images with different magnification factors are visualized in Fig. 7. Statistical data of the dataset are briefly listed in Table II, which includes the number of images classified according to the cancer class and also with different factors of magnification. Out of these samples, benign data were collected from 24 patients and malignant samples were taken from 58 patients. The web link for accessing the dataset used in this study

is as follows which is used in various works: https://web.inf.ufpr.br/vri/databases/breast-cancer-histopathological-database-breakhis/.

B. DATA AUGMENTATION

The throughput of the CNN models is highly reliable on the size of the dataset to be trained. Hence, the augmentation of the provided data could be considered as a major step in improving the cardinality of the dataset [77,78]. Each image of the training dataset is initially resized to images of size 224×224. The Augmentor python library [79] performs the additional data augmentation methods that include rotating, flipping, and cropping methods. The following Fig. 8 demonstrates the resultant images of data augmentation. In the next step, all the images are transformed into matrix form and undergo normalization. On the other hand, the test set is loaded to the model, such that images are normalized with specific input size without any further change; that is, raw image will be input for the test dataset.

VI. EXPERIMENTS AND RESULTS

Experiments of the proposed model have been implemented on a system that has built-in NVIDIA GEFORCE GTX. The VGG-16 model pre-trained with ImageNet is performed by the Keras Deep Learning framework. The VGG-16 model with the last layers replaced and logistic regression is used to classify the histopathological images from the publicly available BreakHis dataset, and the proposed CNN model is optimized in the selection of hyper-parameters. Hyper-parameter optimization uses the particle swarm
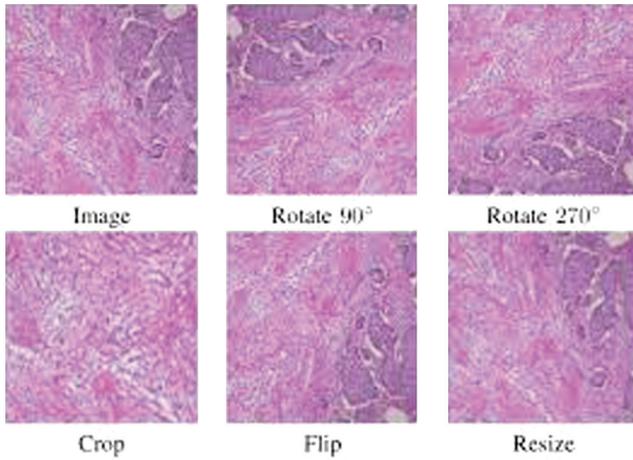
Fig. 8. BreakHis images after applying data augmentation techniques.

optimization algorithm. We split the training and test data in a ratio of 80:20. The parameters of PSO algorithm used in this study are as follows: inertia weight w = 0.7298, population size and number of iteration are taken as 20 and acceleration coefficients c1 and c2 are considered as 1.496.

A. PERFORMANCE METRICS

For evaluating the performance of the proposed system, there are certain aspects that help in a good analysis of the proposed model in comparison with other existing methods. Accuracy is determined by the number of images that are correctly classified, divided by the total number of images. Equations 1, 2, 3, and 4 represent the mathematical expression of Accuracy, Precision, F1-Score, and Sensitivity.

$$\text{Accuracy} = \frac{T_P + T_N}{T_P + F_N + T_N + F_P} \tag{6}$$

$$\text{Precision} = \frac{T_P}{T_P + F_P} \tag{7}$$

$$F1 - \text{score} = \frac{2 * \text{precision} * \text{recall}}{\text{Precision} + \text{recall}} \tag{8}$$

$$\text{sensitivity} = \frac{T_P}{T_P + F_N} \tag{9}$$

In the above equations, true positive presents the number of images that are correctly classified as malignant, false-positive denotes the number of images that are miss-classified as malignant, true
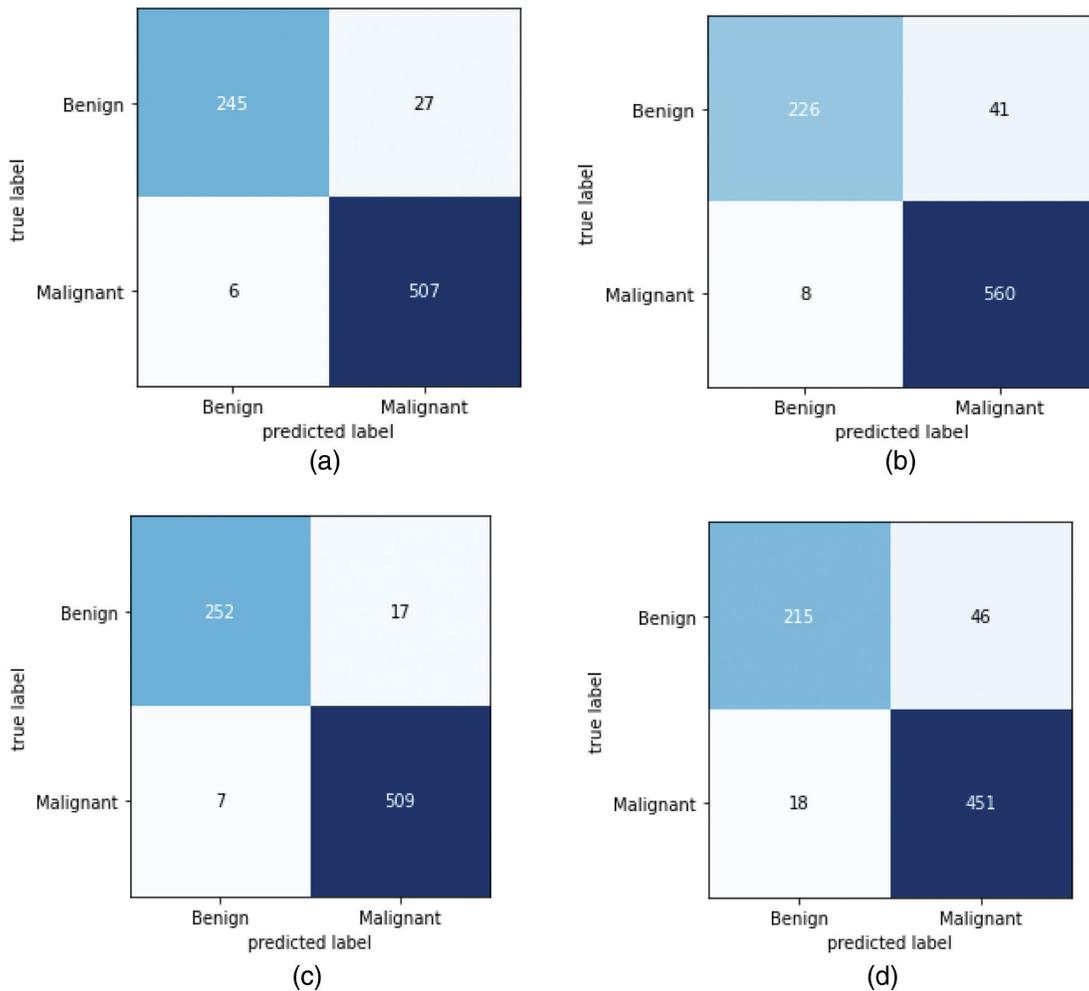
Fig. 9. Confusion matrix of the proposed model. (a) 40× images; (b) 100× images; (c) 200× images; (d) 400× images.
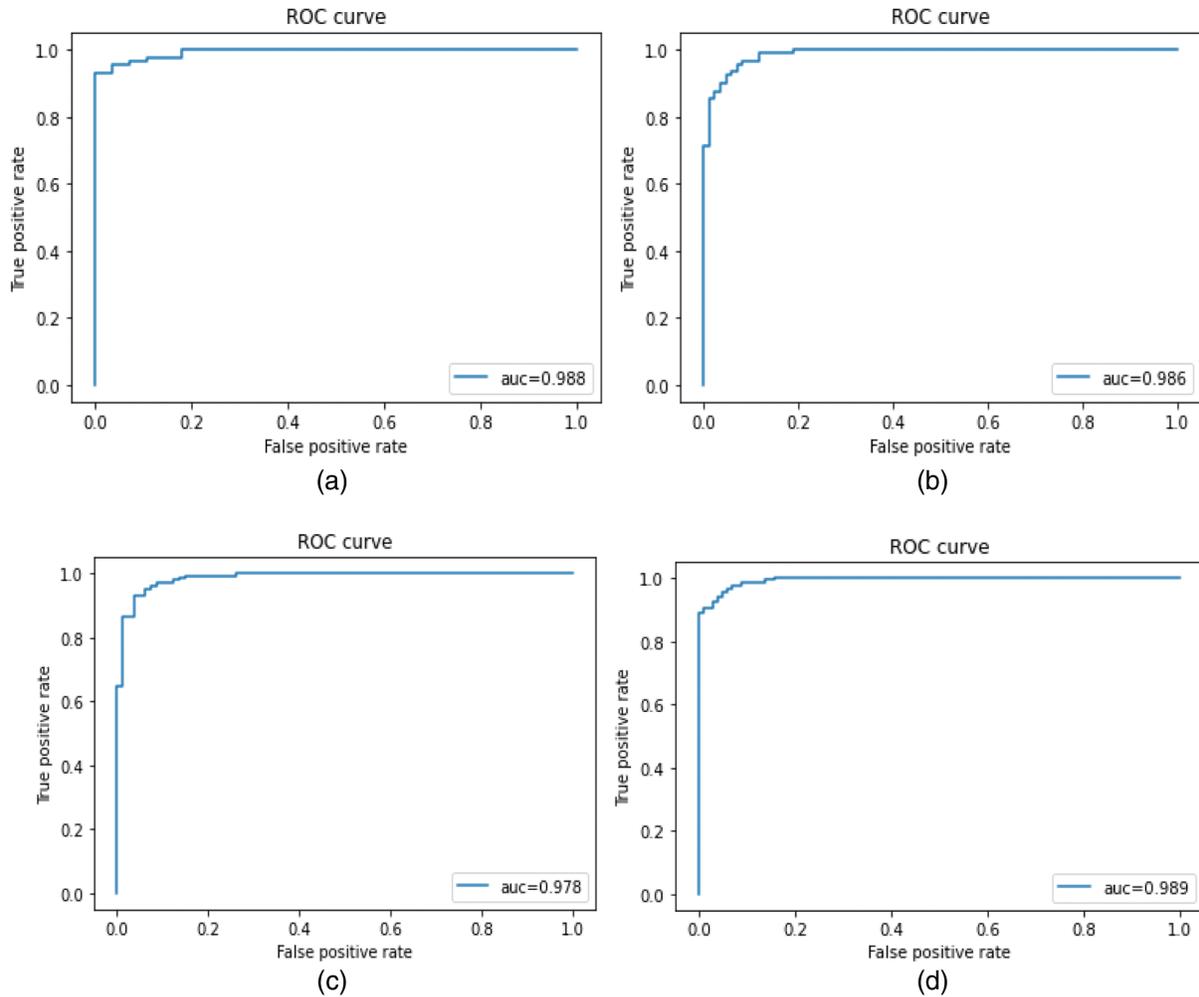
Fig. 10. ROC curves of the proposed model. (a) 40× images; (b) 100× images; (c) 200× images; (d) 400× images.

negative shows the number of images correctly classified as benign, and false-negative shows the number of images miss-classified as benign.

B. RESULTS

In this study, we used a pre-trained CNN model along with optimized hyper-parameters to achieve the best performance of the model. We have experimented with the model with optimized hyper-parameters and classified whether the collected biopsy sample belongs to benign or malignant. We conducted our study in a magnification-dependent manner with 40×, 100×, 200×, and 400× as distinct magnification factors.VGG-16 model is used to classify the images, and the hyper-parameters of this model are obtained with the help of a bio-inspired algorithm namely PSO, during the training phase of the model. Further, to the resultant trained VGG-16 model, test images are provided as input, and also, the parameters of the model are previously optimized. Further, the proposed model performs feature extraction and undergoes image classification tasks, thus classifying them into respective cancer classes using fully connected layers and LR classifier.

The performance of the proposed optimized CNN model is illustrated with the confusion matrix as shown in Fig. 9. From Fig. 9, it is clear that there is several image classification that is correctly classified and miss-classified with different magnification factors. The confusion matrix presents 752 correctly classified images among 785 test images with 40× magnification factor, whereas images with 100× magnification factor 786 images are correctly classified out of 835 images. Also, 761 images were correctly classified among 790 test images with 200× magnification factor and 666 images were correctly classified from 730 test images with 400× magnification factor. Also, we retrieve the ROC curves as mentioned in Fig. 10 for the four types of magnification factors.

The accuracy obtained by the proposed model for distinct magnification factors is mentioned in Table III. For the images with magnification factors with 40×, 100×, 200×, and 400×, the proposed model achieved accuracies of 95.7%, 94.1%, 96.9%, and 91.2%, respectively. From the results, it is shown that the proposed CNN model obtained maximum accuracy with a magnification factor of 200×. Table IV represents the sensitivity, precision, and F1-score achieved by the proposed approach. The experiments show that our model achieved the highest precision of 97.2%, maximum F1-score of 95.2%, and recall of 93.6%.

In Table V, we compare our model with the existing methods that are defined by un-optimized CNN models to predict breast cancer. From Table V, it is clear that the obtained accuracies are less when compared to the proposed CNN model. By comparing

Table III.   Accuracy of the proposed CNN model

Magnification factor	Accuracy in %
40×	95.7
100×	94.1
200×	96.9
400×	91.2

Table IV.   Performance metrics of the proposed CNN model with different magnification factors

Performance metric	Magnification factor			
	40×	100×	200×	400×
Sensitivity (%)	97.6	96.5	97.3	92.2
Precision (%)	90.07	84.6	93.6	82.3
F1-score (%)	93.6	90.2	95.4	87.0

the results of the proposed model with standard deep learning models, it can be observed that the accuracy with magnification factor 40×, the proposed model is at the first rank, and the Deniz *et al.* [80] model and K. Das *et al.* [81] models are at the second and third position. Similarly, for the magnification factor 100×, the first three performing models are proposed model, Saxena *et al.* [82], and Deniz *et al.* [80]. For the magnification factor of 200× and 400×, the best performing models are proposed model, Saxena *et al.* [82], and Deniz *et al.* [80], and the proposed model. It can be noted at lower magnification factor, and the proposed model is performing the best which justifies the use of bio-inspired techniques. Figure 11 clearly illustrates the summary of existing methods used for breast cancer classification and the proposed method. Our model achieved the best accuracy of 96.9% with a magnification factor of 200X. Also, Table VI depicts the clear-cut comparison of other VGG models. From the table, it is clear that fully trained accuracy (63.2%) was improvised to 84.28% for the pre-trained VGG-16 model. Also, pre-trained VGG-16+SVM provides an accuracy of 86.36% and VGG-16 +LR retrieves an accuracy of 92.2%. Hence, the proposed model obtains better accuracy when compared to other VGG models.

Table V.   Comparison of the proposed model with existing deep learning methods

Year	Author	Method	Accuracy with magnification factor			
			40×	100×	200×	400×
2018	K. Das *et al.* [81]	MIL +VGG-16	89.52	89.06	88.84	87.67
2018	S Cascinalle *et al.* [83]	VGG_VD-16	87.0	85.2	85.0	81.3
2020	Saxena *et al.* [82]	ResNet-50 + SVM	89.46	92.6	93.92	89.7
2017	Han *et al.* [84]	AlexNet	85.6	83.5	83.1	80.8
2019	Mahesh *et al.* [85]	RestHit	87.4	87.2	91.2	86.2
2017	Spanhol *et al.* [86]	AlexNet	89.6	85.0	84.2	81.6
2019	Sudharshan *et al.* [39]	AlexNet	87.8	85.6	80.8	82.9
2018	Deniz *et al.* [80]	Fine-tuned AlexNet	90.96	90.58	91.37	91.3
		Our Proposed Model	95.7	94.1	96.9	91.2
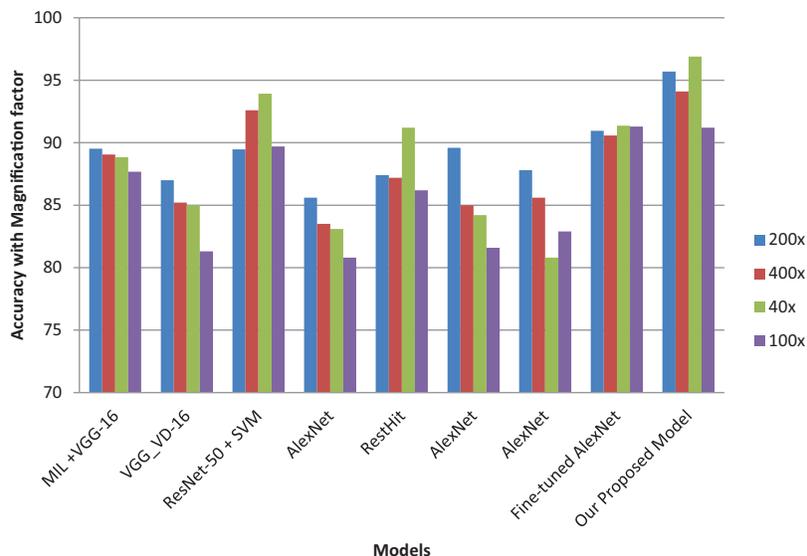
Fig. 11. Comparison summary of existing and proposed models.

Table VI.   Summary of various VGG-16 models

Model	Accuracy in %
Fully trained VGG-16	63.2
Pre-trained VGG-16	84.28
VGG-16 + SVM	86.36
VGG-16 + LR	92.2
Proposed optimized VGG-16	96.9

VII. CONCLUSION

In this paper, we presented a method for hyper-parameter selection of the desired CNN architecture that classifies breast histopathological images. This study uses a pre-trained network of VGG-16 model accompanied by an LR classifier to perform binary classification of breast histopathological images. The transfer learning approach makes the model simpler and easier to implement. LR classifier is placed at the end of the model which outperforms binary classification more efficiently, and the proposed model is trained separately for different magnification factors. Also, it has been observed at the lower magnification factor, the proposed model provides better results as compared to the state-of-the-art techniques. Since network architecture involves the selection of hyper-parameters, we use an evolutionary algorithm namely PSO which optimizes the hyper-parameters of the given CNN model. It often helps to decrease the time complexity when compared to other un-optimized CNN models. We use a publicly available dataset and obtain a classification accuracy of around 96.9 %. In the end, we compare our proposed optimized CNN model with various existing deep learning methods. In the future, we further study other bio-inspired algorithms that help in optimizing the CNN model. Also, future work may be designed in such a way that classifies the subtypes of cancer. Additionally, the optimized model should provide superior performance on all datasets irrespective of the selected magnification factor.

AUTHOR CONTRIBUTION

Chandana mani R K Interpretation of results and manuscript preparation.
Kamalakannan J: Study conception, design, data collection, and analysis.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

CONFLICT OF INTEREST STATEMENT

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

REFERENCES

[1] C. Fitzmaurice et al., "Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study," *JAMA Oncol.*, vol. 3, no. 4, pp. 524–548, 2017.

[2] G. M. Cooper and R. E. Hausman, "The development and causes of cancer," *Cell: Mol. Approach*, vol. 2, pp. 719–728, 2000.

[3] A. Jalalian, S. Mashohor, R. Mahmud, B. Karasfi, M. I. B. Saripan, and A. R. B. Ramli, "Foundation and methodologies in computer-aided diagnosis systems for breast cancer detection," *EXCLI J.*, vol. 16, p. 113, 2017.

[4] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological image analysis: a review," *IEEE Rev. Biomed. Eng.*, vol. 2, pp. 147–171, 2009.

[5] R. K. Chandana Mani and J. Kamalakannan, "Computer-aided diagnosis using white shark optimizer with attention-based deep learning for breast cancer classification," 1 Jan. 2023, pp. 2641–2655.

[6] L. He, L. R. Long, S. Antani, and G. R. Thoma, "Histology image analysis for carcinoma detection and grading," *Comput. Methods Programs Biomed.*, vol. 107, no. 3, pp. 538–556, 2012.

[7] N. Brancati, G. De Pietro, M. Frucci, and D. Riccio, "A deep learning approach for breast invasive ductal carcinoma detection and lymphoma multi-classification in histological images," *IEEE Access*, vol. 7, pp. 44709–44720, 2019.

[8] M. Veta, J. P. Pluim, P. J. Van Diest, and M. A. Viergever, "Breast cancer histopathology image analysis: a review," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1400–1411, 2014.

[9] T. K. Dash, C. Chakraborty, S. Mahapatra, and G. Panda, "Gradient boosting machine and efficient combination of features for speech-based detection of COVID-19," *IEEE J. Biomed. Health. Inf.*, vol. 26, no. 11, pp. 5364–5371, 2022.

[10] T. K. Dash, C. Chakraborty, S. Mahapatra, and G. Panda, "Mitigating information interruptions by COVID-19 face masks: a three-stage speech enhancement scheme," *IEEE Trans. Comput. Soc. Syst.*, 2022.

[11] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 580–587.

[12] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, 3431–3440.

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Adv. N.A. Inf. Process. Syst.*, vol. 25, pp. 1097–1105, 2012.

[14] O. Russakovsky et al., "Imagenet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, 2015.

[15] Z. Hu, J. Tang, Z. Wang, K. Zhang, L. Zhang, and Q. Sun, "Deep learning for image-based cancer detection and diagnosis−a survey," *Pattern Recognit.*, vol. 83, pp. 134–149, 2018.

[16] F. F. Ting, Y. J. Tan, and K. S. Sim, "Convolutional neural network improvement for breast cancer classification," *Expert Syst. Appl.*, vol. 120, pp. 103–115, 2019.

[17] Y. Xu et al., "Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features." *BMC Bioinf.*, vol. 18, no. 1, pp. 1–17, 2017.

[18] R. M. Ghoniem and K. Shaalan, "FCSR-fuzzy continuous speech recognition approach for identifying laryngeal pathologies using new weighted spectrum features," in *International Conference on Advanced Intelligent Systems and Informatics*. Springer, Cham, Sept. 2017, pp. 384–395.

[19] J. C. Bansal, A. Gopal, and A. K. Nagar, "Stability analysis of artificial bee colony optimization algorithm," *Swarm Evol. Comput.*, vol. 41, pp. 9–19, 2018.

[20] J. Decerle, O. Grunder, A. H. El Hassani, and O. Barakat, "A hybrid memetic-ant colony optimization algorithm for the home health care problem with time window, synchronization and

working time balancing," *Swarm Evol. Comput.*, vol. 46, pp. 171–183, 2019.

[21] G. I. Sayed, A. E. Hassanien, and G. Schaefer, "An automated computer-aided diagnosis system for abdominal CT liver images," *Procedia Comput. Sci.*, vol. 90, pp. 68–73, 2016.

[22] K. J. Cios, H. Mamitsuka, T. Nagashima, and R. Tadeusiewicz, "Computational intelligence in solving bioinformatics problems," *Artif. Intell. Med.*, vol. 35, no. 1–2, pp. 1–8, 2005. https://doi.org/10.1016/j.artmed.2005.07.001

[23] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights Imaging*, vol. 9, no. 4, pp. 611–629, 2018.

[24] D. Shen, G. Wu, and H. I. Suk, "Deep learning in medical image analysis," *Annu Rev. Biomed. Eng.*, vol. 19, pp. 221–248, 2017.

[25] G. Litjens et al., "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, 2017.

[26] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "Breast cancer histopathological image classification using convolutional neural networks," in *2016 Int. Joint Conf. Neural Netw. (IJCNN)*, IEEE, July 2016, pp. 2560–2567.

[27] B. Wei, Z. Han, X. He, and Y. Yin, "Deep learning model based breast cancer histopathological image classification," in *2017 IEEE 2nd international conference on cloud computing and big data analysis (ICCCBDA)*: IEEE, Apr. 2017, pp. 348–353.

[28] K. S. Kumar, B. V. Rao, and G. S. Kumar, "Fractional order PLL based sensorless control of PMSM with sliding mode observer," in *2018 Int. Conf. Power, Instrum., Control Comput. (PICC)*: IEEE, Jan. 2018, pp. 1–6.

[29] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[30] A. Kumar et al., "Deep feature learning for histopathological image classification of canine mammary tumors and human breast cancer," *Inf. Sci.*, vol. 508, pp. 405–421, 2020.

[31] R. Mehra, "Breast cancer histology images classification: Training from scratch or transfer learning?," *ICT Exp.*, vol. 4, no. 4, pp. 247–254, 2018.

[32] P. Xie, K. Zuo, Y. Zhang, F. Li, M. Yin, and K. Lu, "Interpretable classification from skin cancer histology slides using deep learning: a retrospective multicenter study," *arXiv preprint arXiv:1904.06156*, 2019.

[33] K. Faust, Q. Xie, D. Han, K. Goyle, Z. Volynskaya, U. Djuric, and P. Diamandis, "Visualizing histopathologic deep learning classification and anomaly detection using nonlinear feature space dimensionality reduction," *BMC Bioinf.*, vol. 19, no. 1, pp. 1–15, 2018.

[34] J. N. Kather et al., "Predicting survival from colorectal cancer histology slides using deep learning: a retrospective multicenter study," *PLoS Med.*, vol. 16, no. 1, p. e1002730, 2019.

[35] M. Mateen, J. Wen, S. Song, and Z. Huang, "Fundus image classification using VGG-19 architecture with PCA and SVD," *Symmetry*, vol. 11, no. 1, p. 1, 2019.

[36] N. Bayramoglu, J. Kannala, and J. Heikkilä, "Deep learning for magnification independent breast cancer histopathology image classification," in *2016 23rd Int. Conf. Pattern Recognit. (ICPR)*: IEEE, Dec. 2016, pp. 2440–2445.

[37] A. Hamidinekoo, E. Denton, A. Rampun, K. Honnor, and R. Zwiggelaar, "Deep learning in mammography and breast histology, an overview and future trends," *Med. Image Anal.*, vol. 47, pp. 45–67, 2018.

[38] Y. Zhang, B. Zhang, F. Coenen, J. Xiao, and W. Lu, "One-class kernel subspace ensemble for medical image classification," *EURASIP J. Adv. Signal Process.*, vol. 2014, no. 1, pp. 1–13, 2014.

[39] P. J. Sudharshan, C. Petitjean, F. Spanhol, L. E. Oliveira, L. Heutte, and P. Honeine, "Multiple instance learning for histopathological breast cancer image classification," *Exp. Syst. Appl.*, vol. 117, pp. 103–111, 2019.

[40] P. Agarwalla and S. Mukhopadhyay, "Bi-stage hierarchical selection of pathway genes for cancer progression using a swarm based computational approach," *Appl. Soft Comput.*, vol. 62, pp. 230–250, 2018.

[41] A. L. Qi, G. M. Zhang, M. Dong, H. W. Ma, and D. M. Harvey, "An artificial bee colony optimization based matching pursuit approach for ultrasonic echo estimation," *Ultrasonics*, vol. 88, pp. 1–8, 2018.

[42] B. A. Garro, K. Rodríguez, and R. A. Vázquez, "Classification of DNA microarrays using artificial neural networks and ABC algorithm," *Appl. Soft Comput.*, vol. 38, pp. 548–560, 2016.

[43] T. Prasartvit, A. Banharnsakun, B. Kaewkamnerdpong, and T. Achalakul, "Reducing bioinformatics data dimension with ABC-kNN," *Neurocomputing*, vol. 116, pp. 367–381, 2013.

[44] A. E. Hassanien, E. Emary, and H. M. Zawbaa, "Retinal blood vessel localization approach based on bee colony swarm optimization, fuzzy c-means and pattern search," *J. Vis. Commun. Image Represent.*, vol. 31, pp. 186–196, 2015.

[45] F. Beloufa and M. A. Chikh, "Design of fuzzy classifier for diabetes disease using modified artificial bee colony algorithm," *Comput. Methods Programs Biomed.*, vol. 112, no. 1, pp. 92–103, 2013.

[46] G. Rubio, H. Pomares, I. Rojas, and L. J. Herrera, "A heuristic method for parameter selection in LS-SVM: application to time series prediction," *Int. J. Forecast.*, vol. 27, no. 3, pp. 725–739, 2011.

[47] A. Klein, S. Falkner, S. Bartels, P. Hennig, and F. Hutter, "Fast bayesian optimization of machine learning hyperparameters on large datasets," in *Artificial Intelligence and Statistics*: PMLR, Apr. 2017, pp. 528–536.

[48] R. Stoean, "Analysis on the potential of an EA–surrogate modelling tandem for deep learning parametrization: an example for cancer classification from medical images," *Neural Comput. Appl.*, vol. 32, no. 2, pp. 313–322, 2020.

[49] K. Kousias, M. Riegler, Ö. Alay, and A. Argyriou, "HINDSIGHT: an R-based framework towards long short term memory (LSTM) optimization," in *Proc. 9th ACM Multimed. Syst. Conf.*, June 2018, pp. 381–386.

[50] S. P. Sankareswaran and M. Krishnan, "Unsupervised end-to-end brain tumor magnetic resonance image registration using RBCNN: rigid transformation, B-spline transformation and convolutional neural network," *Curr. Med. Imaging*, vol. 18, no. 4, 2022. DOI: 10.2174/1573405617666210806125526.

[51] L. Ali, C. Chakraborty, Z. He, W. Cao, Y. Imrana, and J. J. Rodrigues, "A novel sample and feature dependent ensemble approach for Parkinson's disease detection," *Neural Comput. Appl.*, vol. 35, no. 22 pp. 15997–16010, 2023.

[52] M. Mohammed Thaha, M. Pradeep, K. Murugan, B. S. Dhanasekar, S. Vijay Karthick, and P. A. Senthilselvi, "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images," *J. Med. Syst.*, vol. 43, no. 9, July 2019.

[53] V. Surya and A. Senthilselvi, "A qualitative analysis of the machine learning methods in food adultery: a focus on milk adulteration detection," *J. Adv Res Dyn. Control Syst. Titled*, vol. 12, no. 4, May 2020.

[54] H. K. Bhuyan, C. Chakraborty, Y. Shelke, and S. K. Pani, "COVID-19 diagnosis system by deep learning approaches," *Expert Syst.*, vol. 39, no. 3, p. e12776, 2022.

[55] L. Garg, C. Chakraborty, S. Mahmoudi, and V. S. Sohmen (eds.), *Healthcare Informatics for Fighting COVID-19 and Future Epidemics*: Springer, 2022.

[56] A. Senthilselvi, S. Pasayat, K. Sivakumar, and M. Reene Stephanie, "An adaptive, dynamic and semantic approach for understanding of sign language based on convolution neural network," *Turk. Online J. Qual. Inquiry (TOJQI)*, vol. 12, no. 5, pp. 1331–1344, June 2021.

[57] A. Senthilselvi, M. K. Pradeep, S. Dhanasekar, P. Uma Maheswari, S. Ramesh, and S. Senthil Pandi, "Denoising of images from salt and pepper noise using hybrid filter, fuzzy logic noise detector and genetic optimization algorithm (HFGOA)," *Multimed. Tools Appl.*, vol. 78, no 14, July 2019.

[58] A. Senthilselvi, R. Sukumar, and S. Senthil Pandi, "Hybrid fuzzy logic and gravitational search algorithm based multiple filters for image restoration," *Int. J. Data Anal. Tech. Strategies*, vol. 12, no. 1, pp. 76–97, Feb. 2020,

[59] R. Hannah Lalitha, G. Emily Manoranjitham, D. Weslin, and A. Senthilselvi, "Multi–objective approach for protection of microgrids using surrogate assisted particle swarm optimization (SAPSO)," *Appl. Nanosci.*, pp. 1–14, 2021. DOI: 10.1007/s13204-021-02044-7.

[60] V. Surya, and A. Senthilselvi, "Identification of oil authenticity and adulteration using deep long short-term memory-based neural network with seagull optimization algorithm," *Neural Comput. Appl.*, 2022. DOI: 10.1007/s00521-021-06829-3

[61] X. Li, T. Pang, B. Xiong, W. Liu, P. Liang, and T. Wang, "Convolutional neural networks based transfer learning for diabetic retinopathy fundus image classification," in *2017 10th Int. Congr. Image Signal Process. Biomed. Eng. Inf. (CISP-BMEI)*: IEEE, Oct. 2017, pp. 1–11.

[62] M. Gao et al., "Holistic classification of CT attenuation patterns for interstitial lung diseases via deep convolutional neural networks," *Comput. Methods Biomech. Biomed. Eng.: Imaging Visual.*, vol. 6, no. 1, pp. 1–6, 2018.

[63] T. Ozcan and A. Basturk, "Lip reading using convolutional neural networks with and without pre-trained models," *Balkan J. Electr. Comput. Eng.*, vol. 7, no. 2, pp. 195–201, 2019.

[64] T. Özcan and A. Baştürk, "Static image-based emotion recognition using convolutional neural network," in *2019 27th Signal Process. Commun. Appl. Conf. (SIU)*: IEEE, Apr. 2019, pp. 1–4.

[65] E. Kilic, Classification of mitotic figures with convolutional neural networks (Doctoral dissertation, M. Sc. thesis, Erciyes University Graduate School of Natural and Applied Sciences), 2016.

[66] C. Murray, Deep learning CNN's in tensorflow with GPUs, 2018. hackernoon.com/deep-learning-cnns-in-tensorflow-with-gpus-cba6efe0acc2.

[67] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recogn.*, 2016, pp. 770–778.

[68] C. Szegedy et al., "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recogn.*, 2015, pp. 1–9.

[69] H. Daume III and D. Marcu, "Domain adaptation for statistical classifiers," *J. Artif. Intell. Res.*, vol. 26, pp. 101–126, 2006.

[70] B. Zadrozny, "Learning and evaluating classifiers under sample selection bias," in *Proc. Twenty-First Int. Conf. Mach. Learn.*, 2004, July, p. 114.

[71] L. Gui, R. Xu, Q. Lu, J. Du, and Y. Zhou, "Negative transfer detection in transductive transfer learning," *Int. J. Mach. Learn. Cybern.*, vol. 9, no. 2, pp. 185–197, 2018.

[72] T. Ozcan and A. Basturk, "Transfer learning-based convolutional neural networks with heuristic optimization for hand gesture recognition," *Neural Comput. Appl.*, vol. 31, no. 12, pp. 8955–8970, 2019.

[73] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. ICNN'95-Int. Conf. Neural Netw.*: IEEE, Nov. 1995, vol. 4, pp. 1942–1948.

[74] O. U. Rehman, S. Tu, S. U. Rehman, S. Khan, and S. Yang, "Design optimization of electromagnetic devices using an improved quantum inspired particle swarm optimizer," *Appl. Comput. Electromagn. Soc. J.*, vol. 33, no. 9, 2018.

[75] W. Hu and G. G. Yen, "Adaptive multiobjective particle swarm optimization based on parallel cell coordinate system," *IEEE Trans. Evol. Comput.*, vol. 19, no. 1, pp. 1–18, 2013.

[76] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A dataset for breast cancer histopathological image classification," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1455–1462, 2015.

[77] A. A. A. Setio et al., "Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1160–1169, 2016.

[78] S. C. Wong, A. Gatt, V. Stamatescu, and M. D. McDonnell, "Understanding data augmentation for classification: when to warp?," in *2016 Int. Conf. Digital Image Comput.: Tech. Appl. (DICTA)*: IEEE, Nov. 2016, pp. 1–6.

[79] M. D. Bloic, "Augmentor-image augmentation library in python for machine learning," *Retrieved June*, vol. 9, p. 2020, 2017.

[80] E. Deniz, A. Şengür, Z. Kadiroğlu, Y. Guo, V. Bajaj, and Ü. Budak, "Transfer learning based histopathologic image classification for breast cancer detection," *Health Inf. Sci. Syst.*, vol. 6, no. 1, pp. 1–7, 2018.

[81] K. Das, S. Conjeti, A. G. Roy, J. Chatterjee, and D. Sheet, "Multiple instance learning of deep convolutional neural networks for breast histopathology whole slide classification," in *2018 IEEE 15th Int. Symp. Biomed. Imaging (ISBI 2018)*: IEEE, Apr. 2018, pp. 578–581.

[82] S. Saxena, S. Shukla, and M. Gyanchandani, "Pre-trained convolutional neural networks as feature extractors for diagnosis of breast cancer using histopathology," *Int. J. Imaging Syst. Technol.*, vol. 30, no. 3, pp. 577–591, 2020.

[83] S. Cascianelli, R. Bello-Cerezo, F. Bianconi, M. L. Fravolini, M. Belal, B. Palumbo, and J. N. Kather, "Dimensionality reduction strategies for cnn-based classification of histopathological images," in *International Conference on Intelligent Interactive Multimedia Systems and Services*. Cham: Springer, May 2018, pp. 21–30.

[84] Z. Han, B. Wei, Y. Zheng, Y. Yin, K. Li, and S. Li, "Breast cancer multi-classification from histopathological images with structured deep learning model," *Sci Rep*, vol. 7, no. 1, pp. 1–10, 2017.

[85] M. Gour, S. Jain, and T. Sunil Kumar, "Residual learning based CNN for breast cancer histopathological image classification," *Int. J. Imaging Syst. Technol.*, vol. 30, no. 3, pp. 621–635, 2020.

[86] F. A. Spanhol, L. S. Oliveira, P. R. Cavalin, C. Petitjean, and L. Heutte, "Deep features for breast cancer histopathological image classification," in *2017 IEEE Int. Conf. Syst., Man, Cybern. (SMC)*: IEEE, Oct. 2017, pp. 1868–1873.