Computers and Electrical Engineering 103 (2022) 108382

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

Classification of Breast Tumors Based on Histopathology Images Using Deep Features and Ensemble of Gradient Boosting Methods ☆

Mohammad Reza Abbasniya [a,1], Sayed Ali Sheikholeslamzadeh [a,1], Hamid Nasiri [b,*], Samaneh Emami [a]

[a] Department of Electrical and Computer Engineering, Semnan University, Semnan, Iran
[b] Department of Computer Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran

ARTICLE INFO

Keywords:
BreakHis
Breast cancer
Ensemble classification
Grad-CAM
Inception-ResNet-v2
Transfer learning

ABSTRACT

Breast cancer is the most common cancer among women worldwide. Early-stage diagnosis of this disease can significantly improve the efficiency of treatment. Computer-Aided Diagnosis (CAD) Systems are adopted widely in this regard due to their reliability, accuracy and affordability. There are different imaging techniques for a breast cancer diagnosis; one of the most accurate ones is histopathology which is used in this paper. Deep feature transfer learning is used as the main idea of the proposed CAD system's feature extractor. As such, the present paper works on sixteen different pre-trained networks with a focus on their classification phase, something that has not been studied enough. The Inception-ResNet-v2, which has both residual and inception networks profits together, has shown the best feature extraction capability in the case of breast cancer histopathology images among all tested Convolutional neural networks (CNNs). In the classification phase, the ensemble of Categorical Boosting (CatBoost), Extreme Gradient Boosting (XGBoost) and Light Gradient boosting Machine (LightGBM) has given the best average accuracy. The Breast Cancer Histopathological Image Classification (BreakHis) dataset helps evaluating the proposed method, i.e., IRv2-CXL, with the experimental results indicating that IRv2-CXL outperforms other state-of-the-art methods.

1. Introduction

The term cancer is a generic word for an extensive group of diseases that may affect different body parts. To understand cancer, we must first know about Tumors. Tumors can be either benign or malignant, also known as cancerous. Tumors that are deemed to be a benign (noncancerous) increase in size slowly and do not spread, while cancerous tumors grow swiftly, overrun, and dismantle nearby healthy tissues throughout the body. Breast cancer arises in lining cells of ducts or lobules in the breast's glandular tissue. According to the World Health Organization (WHO), breast cancer is the most prevalent cancer worldwide and the fifth most common cause of cancer death in 2020.[2] If breast cancer is detected at its early stages, it can be treated quite effectively before it spreading to other body parts [1].

☆ This paper is for regular issues of CAEE. Reviews processed and recommended for publication to the Editor-in-Chief by Guest Editor Mohamed-Rafik Bouguelia.

* Corresponding author.
E-mail address: h.nasiri@aut.ac.ir (H. Nasiri).

[1] These authors contributed equally to this work and should be regarded as co-first authors.

[2] https://www.who.int/news-room/fact-sheets/detail/cancer, Retrieved on 2022-2-20.

https://doi.org/10.1016/j.compeleceng.2022.108382
Received 9 June 2022; Received in revised form 20 August 2022; Accepted 2 September 2022

Breast cancer is diagnosed by means of various imaging techniques such as mammography, ultrasound, thermography, and pathological tests. Among these, for those patients who have undergone other forms of scannings, histopathology images are considered to be the gold standard for improving the results' accuracy as well as providing reliable information for the assessment of the disease's impacts on surrounding tissues. Histopathological images of tissues from breast tumor patients are obtained by applying certain chemicals to dye the nucleus of the cells and then dyeing the other components with another chemical in different shades so as to accentuate various parts of tissue structures and cellular characteristics [2].

After completing the biopsy, diagnosis will be made by a pathologist who will examine the stained tissues using a microscope. Although these images are very comprehensive, the occurrence of misdiagnoses is fairly possible due to various difficulties such as scant contrast in images, noise, and lack of appreciation by the human eye. Add to this a lack of well-trained specialists and the extremely time-consuming analysis of Hematoxylin and Eosin-stained (H&E-stained) images using microscopes. It becomes understandable why an alternative method to using human specialists is needed [3]. With the onset of pattern recognition, Machine Learning (ML), and Convolutional Neural Networks (CNNs), to overcome existing problems, researchers' focus has shifted to the use of Computer-Aided Diagnosis (CAD) in order to improve tests' accuracy [2]. CAD systems are classified into two different categories. One category sorts the systems via staining method of images used as input, which are H&E-stained and Immunohistochemistry (IHC) images. Images obtained from the use of haematoxylin and eosin are more suited to determine whether a neoplasm is cancerous or not; however, analysis of these images calls for more sophisticated image processing and machine learning methods. Immunohistochemistry images marked by certain biomarkers to illustrate certain cells or regions are especially handy when cancer has already been diagnosed by means of the H&E method [4].

The other category of CAD systems is their grouping based on the images they use, i.e. whether they are Whole Slide Images (WSI) or Region of Interest (ROI). Both H&E-stained and IHC images are used as input in these systems. For diagnosis, using WSI as input, the entirety of images is used without tampering. However, in systems that use ROI, key parts are concentrated on segmentation or image detection [4]. When analyzing WSI's, no segmentation or path determination are done on images with the goal of the system recognizing cancerous and noncancerous slides. With this fact in mind, experts are only required to label the images in the training phase as either benign or malignant [4]. CAD systems employ different algorithms and structures for tumor classification. This study proposes a new method for binary classification of tumors based on histopathology whole slide images through deep feature transfer learning.

Transfer learning is a deep learning method where the knowledge from one model is passed on to another model. Using transfer learning algorithms reduces the training time of models drastically, allowing for different solutions to be built straight away. Furthermore, it can save us the difficulty of establishing an expensive and complicated cloud Graphics Processing Unit (GPU)/Tensor Processing Unit (TPU). Another noteworthy benefit of using transfer learning is tackling a lack of data. For deep learning models to solve a task successfully, a good deal of data is needed. However, cases, where data is abundantly available are few and far between, leaving us to deal with a lack of data most of the time, as evident in the case of breast tumor classification [2].

2. Related works

An early-stage breast cancer diagnosis can substantially decrease its mortality rate; CAD systems using ML approaches could be widely effective here. As with any other ML problem, researchers and data scientists need enough data to be involved in this context. Therefore, many different datasets have been provided with various imaging modalities such as Mammograms (MGs), Ultrasound (US), Magnetic Resonance Imaging (MRI), and Histopathology (HP).

For example, the Bioimaging dataset containing 156 annotated H&E-stained images, sized 2048 × 1536, divided into four classes (normal, benign, in situ, and invasive) was used in a study by Fondon et al. (2018) [5]. In their work, contrast level adaptive histogram equalization is used for preprocessing images, and in the next step, 250 numerical features based on nuclei, color region, and texture characteristics are extracted. Finally, a Support Vector Machine (SVM) classifier is adopted to classify images in accordance with the extracted features.

In another study by George et al. (2020) [6], 200× magnification images of the BreakHis [7] dataset were used to evaluate their method. Their suggested method, NucTraL+BCF, consists of four sections; first, the Macenko [8] strategy is used for stain normalization as the preprocessing phase, then nucleus paths are extracted from images and enter into three individual transfer learning-based processes. Each process utilizes an SVM classifier, and in the final section, belief theoretical classifier fusion is used to determine the final result according to the results provided by three SVM classifiers.

In addition, invasive ductal carcinoma (IDC) is an imbalanced dataset that contains 277,524 color images divided into two classes (negative and positive). This dataset was used by Choudhary et al. (2021) [9]. Their study made use of several CNNs as well as the pruned version of them were employed to train on 70% of the dataset's images by the following two strategies: training the whole CNN or only the last layer. The pruned version of ResNet50, in which all layers were trained, showed the best performance in that case. Another study by Barsha et al. (2021) [10] tried to use an ensemble of two pre-trained CNNs on the same dataset (IDC), but the training size was 80% of the dataset. In that paper, a dense layer with softmax activation was added to the end of the DenseNet-121 and DenseNet-169 networks for binary classification. Test time augmentation was employed to increase model accuracy and, the final decision was reached by using the mean of prediction probabilities.

Thermography infrared is another technique for medical imaging that can be used in the breast cancer diagnosis process. This technique is used in collecting the DMR-IR dataset, which is adopted widely to evaluate state-of-the-art breast cancer diagnosis methods such as the proposed methods by Gonçalves et al. (2022) [11] and Chatterjee et al. (2022) [12]. Similar to other mentioned

Table 1
Distribution of BreakHis samples by different magnifying factors.

BreakHis magnification factor	Benign	Malignant	Total
40×	652	1370	1995
100×	644	1437	2081
200×	623	1390	2013
400×	588	5429	1820
Total	2480	5429	7909

works, Gonçalves et al. also used pre-trained CNNs. They utilized original weighted networks of VGG-16, ResNet-50 and, DenseNet-201. They then searched for the best architecture of the fully connected layer to be adopted as a binary classifier at the end of each CNN. In that paper, two bio-inspired algorithms (i.e., Genetic Algorithm and Particle swarm optimization) were suggested as searching methods. In the study by Chatterjee et al. [12] pre-trained version of VGG-16 was also used to extract image features. The distinctive point of this study was to use a memory-optimized version of the Dragonfly Algorithm to reduce the dimension of the features vector by about 40% before adopting an SVM classifier.

3. Materials and methodology

3.1. Dataset

In this article, the Breast Cancer Histopathological Image Classification (BreakHis) dataset is used to examine our method's viability [7]. The BreakHis dataset was collected in a clinical study by Pathological Anatomy and Cytopathology Laboratory, Parana state, Brazil, from January to December of 2014. BreakHis is a public and well-known dataset that contains 7909 high-quality images taken from the microscopic biopsy of benign and malignant (2480 and 5429 images, respectively) breast tumors. BreakHis dataset contains two procedures of biopsy: Surgical Open Biopsy (SOB) and Core Needle Biopsy (CNB) section of ~ 3 μm thickness. In addition, H&E stained breast tissue biopsy slides are used in the process of collecting images. Image acquisition is made by Olympus BX-50 microscope with a 3.3× magnification relay lens coupled with SCC-131AN, a Samsung digital color camera with 6.5 μm pixel size. These images were collected from 82 patients at four magnifying factors: 40×, 100×, 200×, and 400×. In fact, these magnifications are produced by a 4×, 10×, 20×, and 40× objective lens, respectively, and a 10× ocular lens. No normalization nor color standardization was undertaken on the raw images. Each image is in 700 × 460 pixels dimension with a 3-byte color depth (3-channels: Red, Green, Blue) and saved in Portable Network Graphics (PNG) format. The BreakHis dataset determines the class of tumors (Benign or Malignant) as well as their type. There are four types of each tumor class. For benign tumor, they include Adenosis, Fibroadenoma, Tubular Adenoma, and Phyllodes Tumor; and for malignant ones, they are Ductal Carcinoma, Lobular Carcinoma, Mucinous Carcinoma (Colloid), and Papillary Carcinoma [7]. Some samples of these tumour types are shown in Fig. 1 with the magnification-specific class distribution in Table 1. For this research, we used 70% of the images in the dataset as a training set and the remaining 30% as the test set.

Since BreakHis uses the Histopathology (HP) method, which provides colored images and the possibility of in-depth tissue analysis, there is a higher chance of early-stage cancer diagnosis along with greater confidence in the results, compared to other methods. Known as an invasive method, this model requires much more care in the process of biopsy.

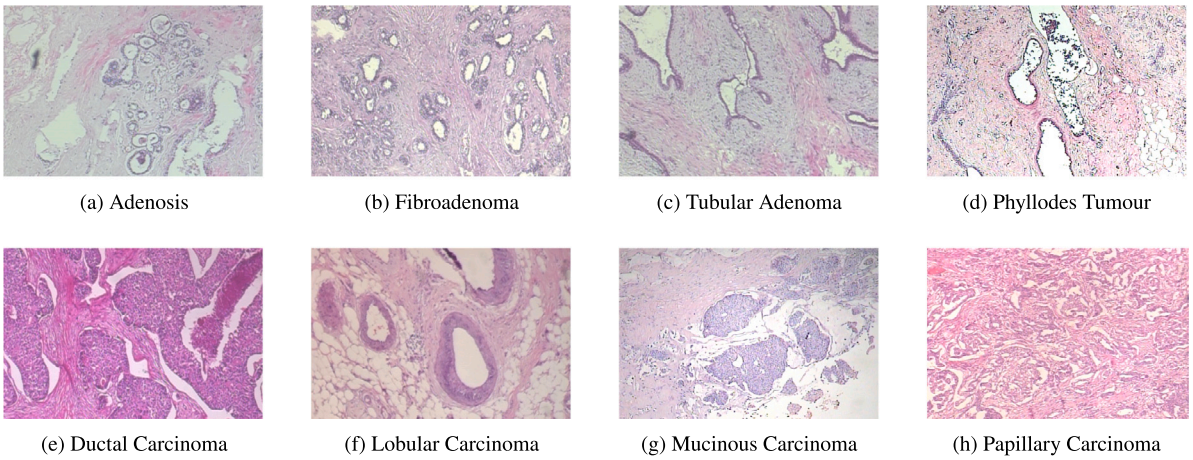
3.2. Methodology

Transfer learning is the main idea for the proposed model's binary classification on the BreakHis dataset. Stevo Bozinovski and Ante Fulgosi were the first to suggest using transfer learning in a neural network [13]. Their paper was published in 1976. Five years later, they reported the transfer learning application in training a neural network on a dataset of computer terminal letter images.

Deep feature transfer learning is the idea of using a neural network that is trained on a massive dataset such as ImageNet with hundreds of classes to predict labels on another dataset. This idea could be achieved by removing fully connected layers, employed to classify images in accordance with the features extracted by convolutional layers and adding another classifier to the network; therefore, allowing us to use the features, extracted by a network with a massive dataset with another classification method in different models. The classifier we add to the model can either be a fully-connected layer or based on other classification algorithms.

3.2.1. Feature extractor

After testing 16 pre-trained neural networks, Inception-ResNet-v2, which is trained on ImageNet, showed the most promising results in our case. Accuracies of extracting features by different neural networks are available in Section 4. The Inception-ResNet-v2 was first introduced by Szegedy et al. in 2016 [14]. The idea of this neural network involved residual connections and architecture of Inception. To achieve the advantages of residual connection and its efficiency in computation at the same time, the filter connection stage of Inception architecture was replaced with residual connections. In [14], the authors argued that using residual connections causes great improvement in training speed. The architecture of Inception-ResNet-v2 is illustrated in Fig. 4. Every convolution marked by "V" is a valid-padded layer (filter window stays inside the input) and those that are not marked by "V" are same-padded (output window is the same size as the input). There are three inception blocks in the architecture of Inception-ResNet-v2 and in each block, just before matching the depth of input, a 1 × 1 convolutional layer is placed to increase the filter bank dimensionality [14].



(a) Adenosis (b) Fibroadenoma (c) Tubular Adenoma (d) Phyllodes Tumour

(e) Ductal Carcinoma (f) Lobular Carcinoma (g) Mucinous Carcinoma (h) Papillary Carcinoma

Fig. 1. Samples of benign (a–d) and malignant (e–h) tumor types in 40× magnification.

3.2.2. Gradient Class Activation Map (Grad-CAM)

Spatial information is naturally preserved in convolutional layers in general. As a result, it is anticipated that the last convolutional layer of the feature extractor has the best balance between high-level semantics and detailed spatial information. Neurons in the last convolutional layer search for semantic class-specific data in the given image. Grad-CAM uses gradient information streaming into these layers to appoint significance values to each neuron for a specific decision of interest, thus emphasizing the identification area for each class in the input image. In other words, Grad-CAM determines the gradient of a particular class grade with regard to the pixel of the feature map and then averages gradients of each channel to come by channel-wise weight. With this information, it is possible for us to disclose the internal logic of CNNs while providing clinical references about patients' data and health status to a greater extent [15,16]. We can also regard these visual insights to determine where the model is failing and why and then address these problems by changing the model's architecture. Examples of applying Grad-CAM on features extracted by Inception-ResNet-v2 are shown in Table 2.

3.2.3. Classifier

Three different classifiers were used in a variety of different model structures in this study to determine whether a tumor is benign or malignant based on the deep features extracted by the deep neural network; however, here only discuss the model that lent to the best results and only mention the others. Extreme Gradient Boosting (XGBoost), Categorical Boosting (CatBoost), and Light Gradient Boosting Machine (LightGBM) were used alone in three of the seven models, while the other four used a soft voting ensemble method of two or all three of the mentioned classifiers instead. The best results were achieved by LightGBM, XGBoost, and CatBoost classifier's soft voting described in 3.2.4.

LightGBM:. Gradient boosting decision tree (GBDT) is a time-honored model, utilizing a weak classifier, namely a decision tree, with iterative training to achieve a top-notch model that enjoys advantages such as competent training effect and resistance to overfitting. LightGBM is a speedy, open-source lifting framework based on GBDT, which has a faster training speed, lower memory consumption, and better accuracy [17]. LightGBM can quickly process massive data by instigating leaf-wise tree growth (see Fig. 2) with depth limitation, gradient-based one-side sampling, and exclusive feature bundling approaches [18]. LightGBM hyperparameters were tuned by trial and error. The output received from LightGBM is acquired using the formula provided below:

$$F_M(x) = \sum_{m=1}^{M} h_m(x) \tag{1}$$

In which $h_m(x)$ represents the base decision tree and M, the highest number of iterations.

CatBoost:. This algorithm integrates GBDT and categorical features, hence the name CatBoost. CatBoost is an oblivious tree-based with fewer parameters and categorical variable support. The mentioned implementation also handles gradient bias and prediction shift. Doing so improves the generalization and robustness of the algorithm. CatBoost uses permutation, target-based statistics, and one-hot-max-size (OHMS) to emphasize categorical columns while utilizing the greedy algorithm at each new tree split to resolve the exponential feature combination increase. CatBoost boasts the characteristics of high classification accuracy and fast training speed. In fact, the highest single magnification classification accuracy result (i.e., 97.49%) was achieved by using CatBoost as the model's classifier; however, the ensemble method worked better across all magnifications on average. In this study, CatBoost hyperparameters were tuned by trial and error.

Table 2
Inception-ResNet-v2 Grad-CAM visualization samples.

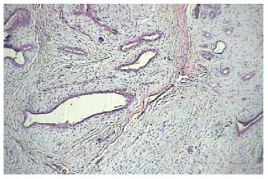
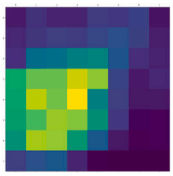
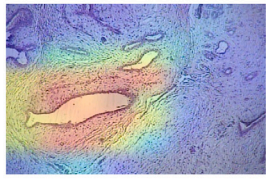
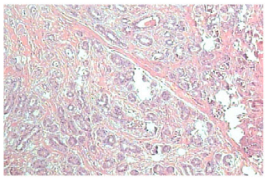
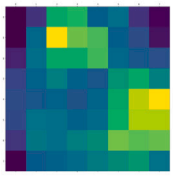
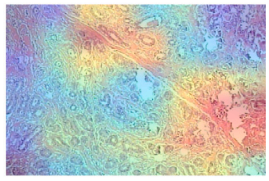
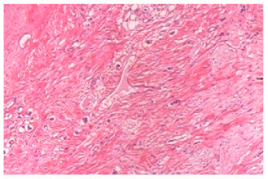
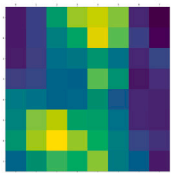
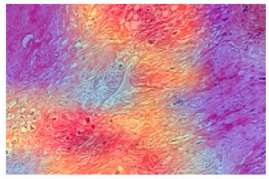
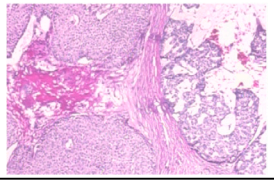
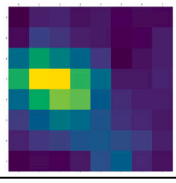
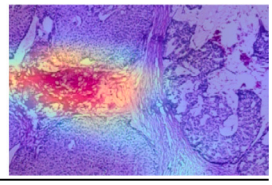
Tumour class	Original biopsy image	Heatmap	Grad-CAM
Benign			
			
Malignant			
			

Fig. 2. LightGBM leaf-wise tree growth. This method can quickly process massive data by instigating leaf-wise tree growth with depth limitations.

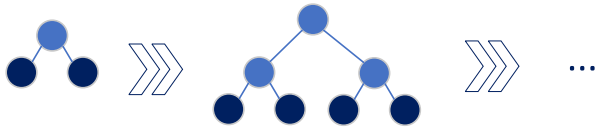


Fig. 3. XGBoost level-wise tree growth. This method expands the tree level-wise by constantly dividing features.

XGBoost:. XGBoost, like the other two classifiers used in the proposed majority voting ensemble method, is a technique that employs the results of several decision trees to generate a prediction score [19]. This method expands the tree level-wise (see Fig. 3) by constantly dividing features. Every tree then computes the feature and threshold with the best branch effect and proceeds to complete the split. Simply put, XGBoost yields better results than conventional decision trees by utilizing classification and regression trees as base learners to consecutively blend various tree predictions with the use of gradient boosting as an error minimizer [15]. XGBoost also lowers the likelihood of over-fitting through regularization by employing a second-order Taylor series to approximate the value of the loss function [19–21]. This makes XGBoost a gradient boosting library optimized for efficiency, versatility, and portability [22]. XGBoost hyperparameters were tuned by trial and error. Let $\hat{y}_i$ denote the XGBoost's predicted output, results from the summation of all K scores provided by the trees, as shown in the equation below:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \in \Gamma \tag{2}$$

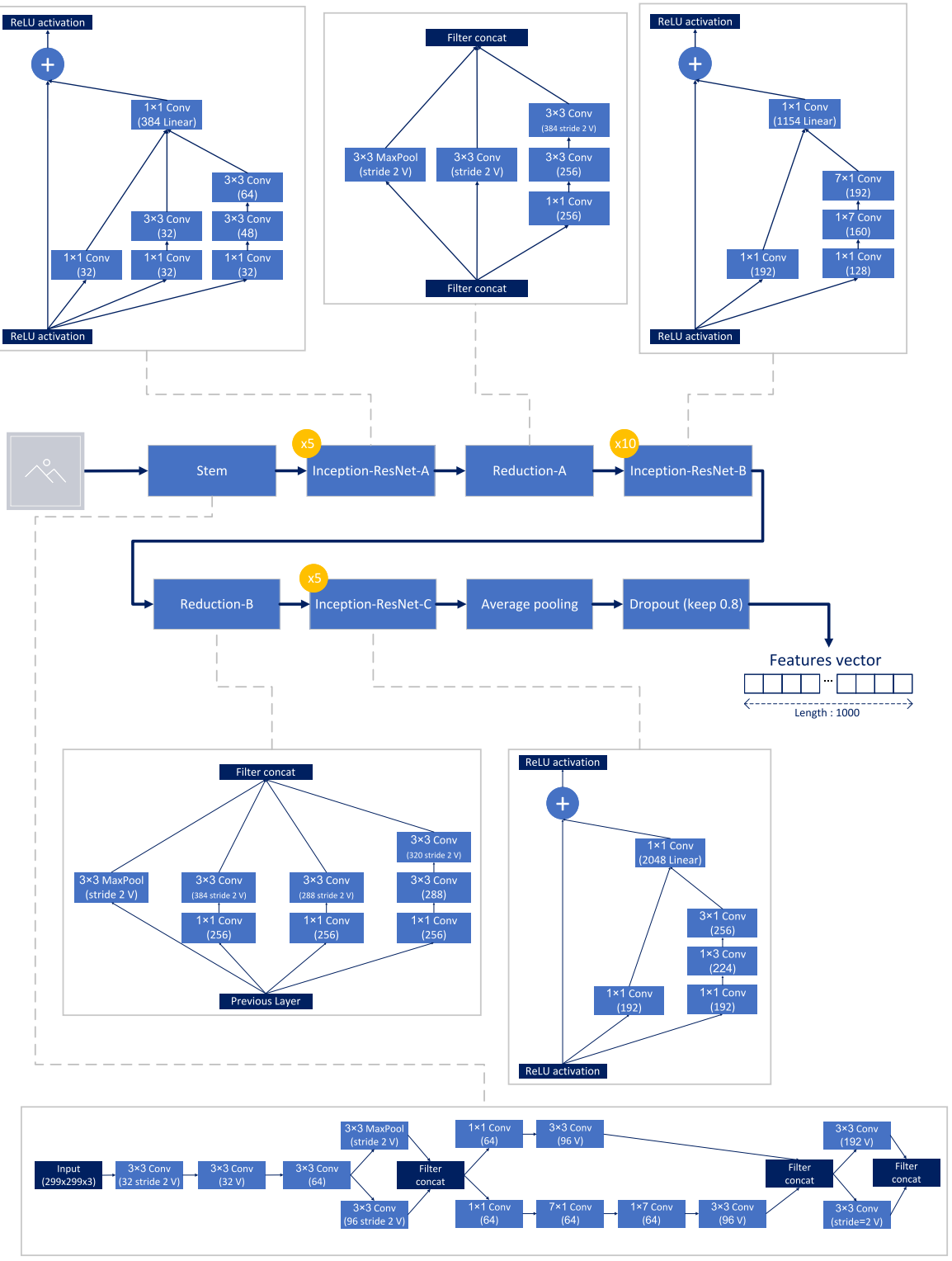


Fig. 4. Architecture of Inception-ResNet-v2 as a feature extractor. Histopathology images are fed to this network as input; output is a vector of 1000 features extracted from the images.

where K, f_k, x_i and Γ stand for the number of regression trees, marks predicted by the kth tree, the input feature vector, and the function space which holds all viable regression trees, respectively.

3.2.4. Ensemble classification

This study made use of three fine-tuned classifiers to predict the probability of malignancy, based on features vector obtained from pre-trained CNNs. We decided to implement different combinations of these classifiers, which use a form of ensemble learning themselves, by utilizing another ensemble method. There are multiple ways to adopt ensemble learning in classification, but here we employed one of the most common ones, called Soft-Voting. In this approach, every classifier produces a probability for each class, the mean of which is considered the final probability of that class [23]. Finally, the class with the highest probability will be the agreement of ensembled classifiers. This approach has no limitation on the number of classifiers, while another method like Majority-Voting can make only an ensemble an even number of classifiers.

4. Results and discussion

4.1. Evaluation metrics

In our binary classification of the BreakHis dataset, predicted labels can be split into four different categories by comparing them to actual labels of the test set: TP (number of correct malignant predictions), TN (number of correct benign predictions), FP (number of incorrect malignant predictions), and FN (number of incorrect benign predictions). The following are different evaluation metrics used in this paper:

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

$$Balanced\ accuracy = \frac{1}{2}\left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP}\right) \tag{4}$$

$$F_1 score = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \tag{5}$$

4.2. Results and proposed method

In this paper, 16 pre-trained networks were used to extract features from input images. Features were then supplied to seven different classifiers (three single and four soft voting ensembles) to predict our final labels. Results of the performance of individual and ensemble models on the test set are available in Table 3; X, L, and C stand for XGBoost, LightGBM, and CatBoost, respectively. The Keras implementation of the pre-trained networks (CNN) is used in this research. For each network, a simple preprocessing method recommended by Keras is employed before feeding images to the CNNs.[3] The Google-Colaboratory free environment (12 GB RAM, 78 GB disk, and cloud GPU) has been used for model implementation.

As evident by the results presented in Table 3, the combination of Inception-ResNet-v2 as feature extractor and an ensemble of all three classifiers produced the best accuracy among all 112 tested models (results with the f1-score as an evaluation metric are also available in Table 4), so it was selected as the proposed method in this study. It consists of Inception-ResNet-v2 and an ensemble of CatBoost, XGBoost, and LightGBM, which we named IRv2-CXL. The full architecture of IRv2-CXL is illustrated in Fig. 5. By reviewing Benhammou [7] and the other studies mentioned in related works, we concluded that ResNet50, DensNet201, and Inception-v3 networks are three of the more commonly used pre-trained networks in breast tumors' classification. Therefore, we decided to compare the Receiver Operating Characteristic (ROC) curve of these networks with Inception-ResNet-v2, using the chosen classifier (i.e., CXL). Results are shown in Fig. 6. When it comes to evaluating a model's performance, ROC is one of the most notable metrics available. ROC measures performance for classification problems at several threshold settings. In short, ROC demonstrates the model's ability to differentiate between different classes. The area under the curve (AUC), a measure of this ability, is also presented in the magnification-specific format in Fig. 6.

Other evaluation metrics applied to the proposed method are provided in Table 5, with Fig. 7 demonstrating the magnification specific confusion matrix.

4.3. Discussion

Analysis of breast cancer histopathological images by utilizing image processing methods developed for the medical field has recently received much attention. Non-automatic examination of breast cancer histopathological images is susceptible to various problems, including observer variability, human error, and an extremely tedious procedure. A solution that alleviates these obstacles is the use of CAD systems. In this study, we proposed a new model for breast cancer histopathology images binary classification based on deep feature transfer learning. In comparing the IRv2-CXL results to other existing binary breast tumor classification methods, the proposed deep feature method attains better or comparable performance without using any image augmentation (see Table 6), signifying transfer learning and GBDTs potential in this task while demonstrating ensemble methods' power in predictive accuracy and bias/variance trade-off. Regarding improved results, we believe a better CAD system can be developed using IRv2-CXL as its foundation.

[3] https://github.com/keras-team/keras/tree/master/keras/applications, Retrieved on 2021-8-5.

Table 3
Inceptio-ResNet-v2 with X&L&C classifiers achieved the highest average **accuracy** among all the models and was therefore selected as the proposed model, using an ensemble of XGBoost, LightGBM, and CatBoost as its classifier.

Feature extractor	Classifier(s)	Result (accuracy %) 40×	100×	200×	400×	Avg.	Feature extractor	Classifier(s)	Result (accuracy %) 40×	100×	200×	400×	Avg.
VGG 16	X	85.64	81.44	83.44	78.20	82.18	Inception v3	X	94.49	93.60	94.70	93.58	94.09
	L	85.97	82.56	85.26	78.38	83.04		L	95.82	94.56	94.53	94.68	94.90
	C	86.47	83.68	85.59	78.93	83.67		C	95.15	93.44	93.70	94.13	94.11
	X & L	87.31	82.88	84.76	79.48	83.61		X & L	94.99	93.92	94.86	95.23	94.75
	X & C	86.64	82.24	84.76	79.30	83.24		X & C	95.15	93.76	94.70	94.87	94.62
	L & C	86.31	84.00	85.09	79.48	83.72		L & C	95.32	94.08	94.70	94.87	94.74
	X & L & C	86.97	83.04	84.76	78.93	83.43		X & L & C	95.49	94.24	94.53	95.23	94.87
VGG 19	X	85.64	80.32	83.27	78.38	81.90	Inception ResNet v2	X	96.32	95.36	95.36	95.97	95.75
	L	84.47	83.04	83.60	78.20	82.33		L	96.82	95.52	96.85	96.33	96.38
	C	85.97	82.72	85.43	79.67	83.45		C	**97.49**	95.84	96.19	95.97	96.37
	X & L	85.64	81.60	85.43	80.58	83.31		X & L	96.49	95.84	96.68	**96.52**	96.38
	X & C	86.81	81.92	85.59	80.40	83.68		X & C	96.82	96.16	96.35	95.78	96.28
	L & C	85.14	82.56	84.60	79.67	82.99		L & C	96.99	96.16	96.68	95.42	96.31
	X & L & C	85.80	82.40	85.92	80.03	83.54		X & L & C	96.82	95.84	**97.01**	96.15	**96.46**
ResNet 50	X	87.81	89.60	90.39	90.65	89.61	DenseNet 121	X	93.99	95.20	93.54	95.23	94.49
	L	89.81	89.92	90.72	90.10	90.14		L	94.15	95.84	94.53	95.23	94.94
	C	89.14	90.08	91.05	89.92	90.05		C	94.15	94.72	93.87	94.50	94.31
	X & L	89.81	90.56	91.05	90.29	90.43		X & L	94.65	96.16	94.37	94.87	95.01
	X & C	89.31	90.56	91.39	90.29	90.39		X & C	93.99	95.20	94.03	94.87	94.52
	L & C	90.15	90.08	90.89	90.84	90.49		L & C	93.99	95.36	94.37	95.42	94.79
	X & L & C	89.48	90.72	91.22	90.84	90.57		X & L & C	94.65	95.20	95.03	95.23	95.03
ResNet 101	X	86.14	86.40	87.58	87.36	86.87	DenseNet 169	X	94.65	96.00	94.20	94.68	94.88
	L	86.47	88.16	89.40	88.09	88.03		L	94.65	**96.64**	94.20	93.77	94.82
	C	87.14	88.00	90.06	89.56	88.69		C	93.15	96.16	94.86	94.50	94.67
	X & L	86.14	87.68	88.90	87.72	87.61		X & L	94.65	95.84	94.20	94.13	94.71
	X & C	86.97	88.00	89.56	88.46	88.25		X & C	93.99	95.68	94.20	94.68	94.64
	L & C	87.31	88.64	89.73	88.82	88.63		L & C	94.15	96.48	94.70	93.95	94.82
	X & L & C	86.64	88.32	89.56	88.46	88.25		X & L & C	94.32	96.00	94.70	93.95	94.74
ResNet 152	X	86.14	84.96	88.57	86.26	86.48	DenseNet 201	X	92.98	95.20	93.37	95.23	94.20
	L	86.14	85.92	88.57	86.63	86.82		L	93.99	95.36	95.03	94.87	94.81
	C	86.14	87.04	88.74	85.34	86.82		C	93.82	95.36	94.53	93.58	94.32
	X & L	86.47	85.76	89.56	86.26	87.01		X & L	94.49	95.04	94.70	95.05	94.82
	X & C	86.97	86.24	88.90	86.81	87.23		X & C	93.65	94.88	94.20	94.68	94.35
	L & C	86.31	85.92	89.40	86.44	87.02		L & C	94.15	95.68	94.86	94.32	94.75
	X & L & C	86.47	85.92	89.07	87.17	87.16		X & L & C	94.32	95.52	95.03	94.50	94.84
ResNet 50-v2	X	95.49	94.24	93.54	92.67	93.99	Xception	X	92.15	91.04	92.88	92.85	92.23
	L	94.65	95.52	94.20	94.13	94.63		L	93.65	92.16	92.05	94.13	93.00
	C	94.65	95.52	93.70	93.04	94.23		C	90.98	92.00	91.72	90.47	91.29
	X & L	95.82	95.84	94.37	93.77	94.95		X & L	93.65	92.32	93.04	93.40	93.10
	X & C	95.82	95.04	94.03	92.67	94.39		X & C	92.48	92.32	91.72	92.12	92.16
	L & C	94.99	95.68	94.53	93.40	94.65		L & C	92.98	92.64	92.38	93.40	92.85
	X & L & C	95.49	95.68	94.37	93.22	94.69		X & L & C	93.15	92.00	92.38	93.04	92.64
ResNet 101-v2	X	94.15	95.04	93.37	93.04	93.90	NASNet Large	X	94.49	94.08	93.21	92.12	93.48
	L	94.82	96.00	93.87	95.42	95.03		L	94.82	94.56	94.20	93.58	94.29
	C	93.15	94.88	94.37	93.58	94.00		C	94.49	93.44	94.37	92.49	93.70
	X & L	94.32	95.20	93.54	94.68	94.44		X & L	94.65	95.04	93.54	93.40	94.16
	X & C	94.49	95.36	93.54	93.04	94.11		X & C	93.65	94.40	94.20	92.30	93.64
	L & C	94.32	95.36	93.70	95.05	94.61		L & C	94.65	94.56	94.20	93.40	94.20
	X & L & C	94.15	95.20	93.87	94.13	94.34		X & L & C	94.49	94.56	94.03	93.22	94.08
ResNet 152-v2	X	93.99	92.48	94.37	93.04	93.47	EfficientNet B6	X	84.97	85.92	88.24	88.46	86.90
	L	94.65	94.08	94.03	94.13	94.22		L	86.47	85.60	88.74	88.64	87.36
	C	94.49	93.28	93.21	93.22	93.55		C	86.47	85.44	87.91	89.56	87.35
	X & L	94.49	94.08	94.03	93.40	94.00		X & L	85.14	86.08	88.07	89.01	87.08
	X & C	94.49	92.80	93.70	93.40	93.60		X & C	85.47	85.44	88.24	89.19	87.09
	L & C	94.65	93.76	94.03	94.13	94.14		L & C	86.47	86.08	88.90	89.74	87.80
	X & L & C	94.82	94.08	94.37	93.95	94.31		X & L & C	85.80	85.76	88.41	89.19	87.29

As illustrated in Table 6, IRv2-CXL achieved better results than the 7 models included in the table in three of the four magnifications factors. In the 40× magnification factor, the proposed model outperformed almost all the other models by a considerable margin, the only close ones being Sharma and Kumar [24] with 96.25%, and Kumar et al. [25] with 94.11%. IRv2-CXL attained a 95.84% accuracy in the 100× magnification factor, only outdone by the 96.25% achieved by Sharma and Kumar [24] with a 0.41% margin. The 200× magnification factor led to the best individual magnification result of the proposed model. In this

Table 4
The proposed model acquired the best **F1-score**. This model consists of a classifier which is an ensemble of CatBoost, XGBoost, and LightGBM.

Feature extractor	Classifier(s)	Result (F1-Score × 100) 40×	100×	200×	400×	Avg.	Feature extractor	Classifier(s)	Result (F1-Score × 100) 40×	100×	200×	400×	Avg.
VGG 16	X	90.23	87.22	88.45	84.36	87.57	Inception v3	X	96.13	95.51	96.21	95.21	95.77
	L	90.67	88.17	89.87	85.18	88.47		L	97.06	96.15	96.10	96.01	96.33
	C	90.97	88.89	90.06	85.50	88.86		C	96.60	95.41	95.48	95.62	95.78
	X & L	91.48	88.38	89.47	85.68	88.75		X & L	96.47	95.72	96.34	96.42	96.24
	X & C	90.99	87.90	89.50	85.38	88.44		X & C	96.61	95.60	96.21	96.16	96.15
	L & C	90.91	89.22	89.75	85.93	88.95		L & C	96.71	95.83	96.21	96.14	96.22
	X & L & C	91.28	88.45	89.52	85.42	88.67		X & L & C	96.82	95.94	96.09	96.42	96.32
VGG 19	X	90.23	86.50	88.27	84.47	87.37	Inception ResNet v2	X	97.42	96.69	96.69	96.98	96.95
	L	89.70	88.38	88.76	85.07	87.98		L	97.75	96.80	97.75	97.24	97.39
	C	90.71	88.29	89.93	86.04	88.74		C	**98.23**	97.04	97.28	96.98	97.38
	X & L	90.44	87.46	89.86	86.41	88.54		X & L	97.52	97.03	97.64	**97.38**	97.39
	X & C	91.11	87.62	89.94	86.23	88.73		X & C	97.76	97.27	97.41	96.84	97.32
	L & C	90.14	88.17	89.35	86.11	88.44		L & C	97.87	97.25	97.64	96.56	97.33
	X & L & C	90.52	87.99	90.24	86.19	88.74		X & L & C	97.76	97.03	**97.87**	97.11	**97.44**
ResNet 50	X	91.70	92.55	93.16	93.10	92.63	DenseNet 121	X	95.80	96.58	95.37	96.42	96.04
	L	92.98	92.77	93.35	92.78	92.97		L	95.92	97.05	96.09	96.41	96.37
	C	92.54	92.94	93.62	92.62	92.93		C	95.91	96.23	95.60	95.88	95.91
	X & L	92.96	93.24	93.60	92.87	93.17		X & L	96.26	97.27	95.96	96.11	96.40
	X & C	92.68	93.27	93.85	92.89	93.17		X & C	95.79	96.57	95.70	96.40	96.12
	L & C	93.21	92.94	93.49	93.28	93.23		L & C	95.79	96.70	95.98	96.55	96.26
	X & L & C	92.77	93.39	93.73	93.28	93.29		X & L & C	96.25	96.58	96.45	96.39	96.42
ResNet 101	X	90.49	90.44	91.04	90.69	90.67	DenseNet 169	X	96.25	97.16	95.85	95.97	96.31
	L	90.64	91.61	92.34	91.25	91.46		L	96.24	**97.61**	95.85	95.28	96.25
	C	91.14	91.47	92.92	92.35	91.97		C	95.20	97.27	96.32	95.86	96.16
	X & L	90.47	91.30	92.03	90.98	91.20		X & L	96.25	97.05	95.85	95.56	96.18
	X & C	91.03	91.56	92.56	91.50	91.66		X & C	95.79	96.94	95.84	95.97	96.14
	L & C	91.26	91.94	92.65	91.79	91.91		L & C	95.90	97.50	96.19	95.44	96.26
	X & L & C	90.80	91.77	92.54	91.52	91.66		X & L & C	96.03	97.17	96.20	95.45	96.21
ResNet 152	X	90.54	89.41	91.83	89.82	90.40	DenseNet 201	X	95.14	96.58	95.29	96.42	95.86
	L	90.43	90.00	91.93	90.07	90.61		L	95.79	96.71	96.44	96.14	96.27
	C	90.62	90.93	92.02	89.22	90.70		C	95.66	96.69	96.09	95.17	95.90
	X & L	90.70	89.97	92.65	89.82	90.79		X & L	96.16	96.48	96.21	96.28	96.28
	X & C	91.08	90.40	92.15	90.19	90.96		X & C	95.58	96.35	95.86	96.02	95.95
	L & C	90.66	90.11	92.52	90.00	90.82		L & C	95.91	96.94	96.32	95.75	96.23
	X & L & C	90.74	90.11	92.25	90.51	90.90		X & L & C	96.04	96.82	96.45	95.88	96.30
ResNet 50-v2	X	96.83	95.90	95.34	94.54	95.65	Xception ‘	X	94.46	93.55	94.84	94.66	94.38
	L	96.24	96.80	95.84	95.60	96.12		L	95.55	94.31	94.34	95.60	94.95
	C	96.27	96.82	95.50	94.82	95.85		C	93.72	94.16	94.06	92.99	93.73
	X & L	97.07	97.03	95.95	95.33	96.35		X & L	95.53	94.43	95.04	95.07	95.02
	X & C	97.06	96.46	95.71	94.57	95.95		X & C	94.74	94.43	94.03	94.15	94.34
	L & C	96.48	96.91	96.08	95.08	96.14		L & C	95.07	94.66	94.59	95.08	94.85
	X & L & C	96.84	96.91	95.95	94.94	96.16		X & L & C	95.18	94.20	94.56	94.81	94.69
ResNet 101-v2	X	95.89	96.46	95.25	94.79	95.60	NASNet Large	X	96.15	95.72	95.11	94.05	95.26
	L	96.37	97.15	95.60	96.56	96.42		L	96.37	96.09	95.81	95.19	95.87
	C	95.20	96.35	95.99	95.20	95.69		C	96.14	95.32	95.97	94.36	95.45
	X & L	96.02	96.57	95.34	96.03	95.99		X & L	96.26	96.44	95.34	95.03	95.77
	X & C	96.12	96.69	95.40	94.79	95.75		X & C	95.58	95.99	95.84	94.21	95.41
	L & C	96.02	96.69	95.49	96.29	96.12		L & C	96.26	96.10	95.83	95.03	95.81
	X & L & C	95.90	96.58	95.60	95.62	95.93		X & L & C	96.15	96.11	95.71	94.91	95.72
ResNet 152-v2	X	95.77	94.67	95.96	94.77	95.29	EfficientNet B6	X	89.73	90.20	91.70	91.33	90.74
	L	96.24	95.81	95.77	95.60	95.86		L	90.70	90.07	92.07	91.55	91.10
	C	96.13	95.23	95.12	94.92	95.35		C	90.83	90.05	91.50	92.29	91.17
	X & L	96.12	95.81	95.75	95.04	95.68		X & L	89.85	90.41	91.59	91.87	90.93
	X & C	96.12	94.90	95.49	95.04	95.39		X & C	90.10	90.03	91.73	92.02	90.97
	L & C	96.24	95.58	95.75	95.60	95.79		L & C	90.74	90.47	92.24	92.39	91.46
	X & L & C	96.35	95.80	95.99	95.47	95.90		X & L & C	90.31	90.25	91.88	91.99	91.11

magnification, IRv2-CXL and Kumar et al.'s [25] model surpassed the other models with an accuracy gap of 1.27% at best, and 9.26% at worst, by achieving the same 97.01% accuracy. In the 400× magnification factor, the proposed model secured the best result by a 2.04% margin with an accuracy of 96.15%, the second-best model being Sharma and Kumar's [24]. In average IRv2-CXL model attained the best accuracy (96.46%), followed by Sharma and Kumar's [24] model with an average accuracy of 95.58%.

We decided to use GBDT-based methods as a classifier thanks to their high prediction accuracy, adaptability to non-linear characteristics, and good training effect. XGBoost, CatBoost, and LightGBM were chosen as classifier candidates. XGBoost got selected

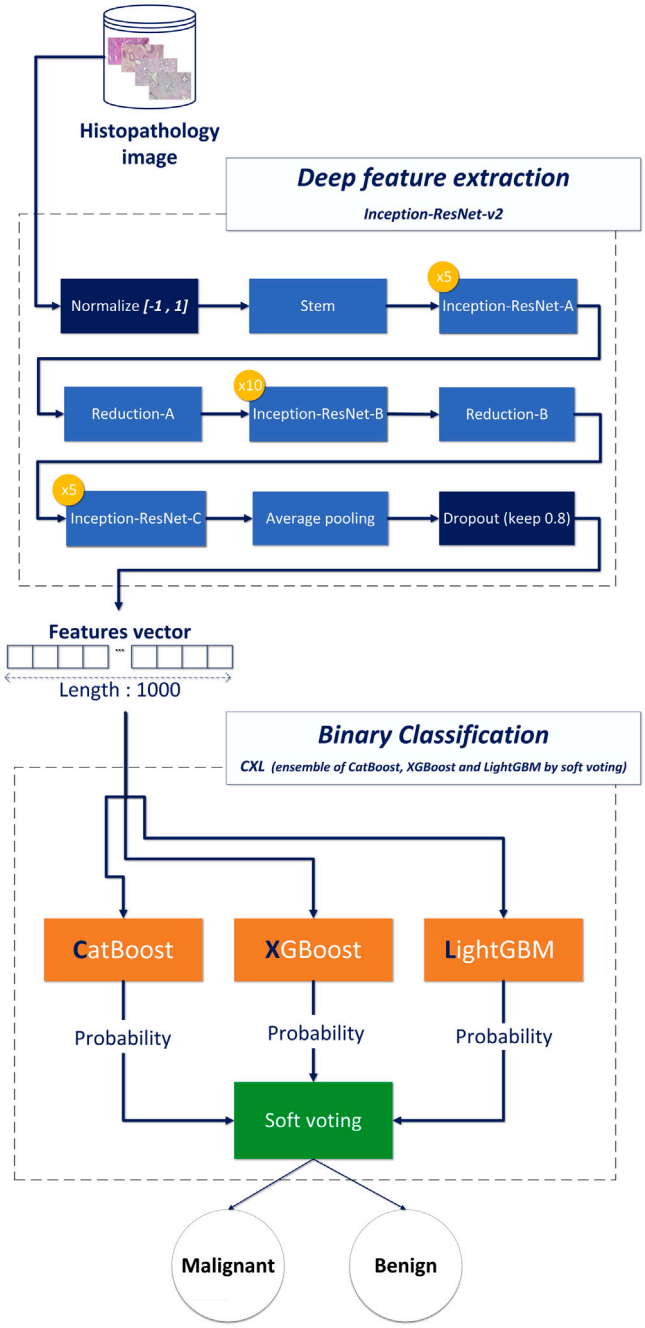


Fig. 5. Proposed method (IRv2-CXL) architecture.

Table 5
Proposed method (IRv2-CXL) evaluation.

Metric	Results				
	40×	100×	200×	400×	**Average**
Accuracy	96.83%	95.84%	97.01%	96.15%	**96.46%**
Balanced accuracy	95.64%	94.90%	95.89%	95.15%	**95.40%**
F_1 *score*	97.76%	97.03%	97.87%	97.11%	**97.44%**

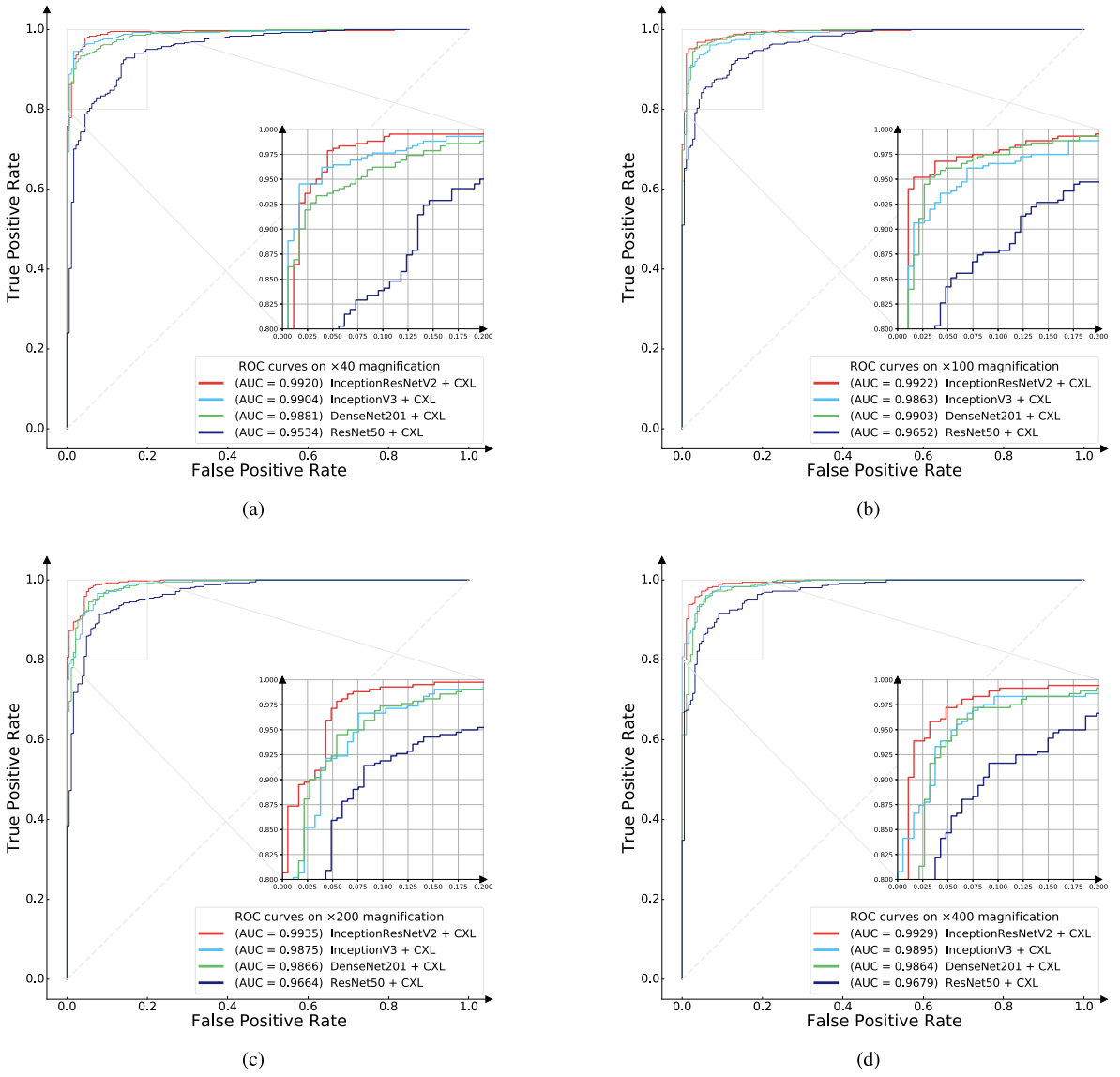


Fig. 6. ROC curves of different feature extractors classified by CXL. (a) 40×, (b) 100×, (c) 200×, (d) 400×.

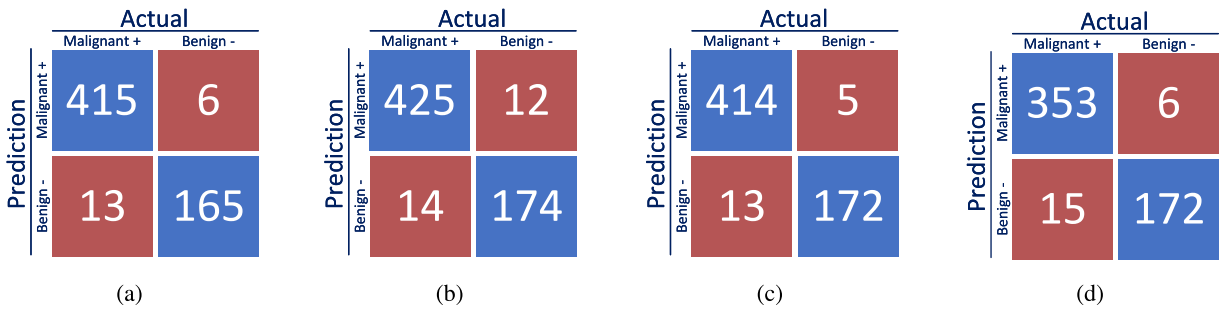


Fig. 7. IRv2-CXL confusion matrix on different test set magnifications. (a) 40×, (b) 100×, (c) 200×, (d) 400×.

because of its new regularization method that resists overfitting [26], thus making the model more robust. CatBoosts improved generalization, and its ability to capture high-order dependencies made it a viable candidate. Lastly, LightGBM is an attractive

Table 6
Comparison of IRv2-CXL with state-of-the-art models in all magnification factors of the BreakHis dataset.

Study (Year)	Results (Accuracy %)				
	40×	100×	200×	400×	Average
Alirezazadeh et al. [4] (2018)	89.10	87.30	91.00	86.60	88.50
Du et al. [27] (2018)	90.69	90.46	90.64	90.96	90.68
Kumar et al. [25] (2020)	94.11	95.12	**97.01**	93.40	94.91
Li et al. [28] (2021)	87.85	86.68	87.75	85.30	86.89
Sharma and Kumar [24] (2022)	96.25	**96.25**	95.74	94.11	95.58
Joseph et al. [29] (2022)	90.87	89.57	91.58	88.67	90.17
Zerouaoui and Idri [30] (2022)	92.61	92.00	93.93	91.73	92.56
IRv2-CXL (2022)	**96.82**	95.84	**97.01**	**96.15**	**96.46**

method to use as a classifier due to its capability to handle large-scale data. In addition, CatBoost's leaf-wise growth approach made it more accurate than other level-wise GBDT methods. As all three of our classifier candidates performed relatively well on this task, we decided to incorporate a soft majority voting method in the proposed method to enhance its predictive performance and improve the model's robustness by lowering its variance. We found that IRv2-CXL, a model with Inception-ResNet-v2 as its feature extractor and soft majority voting ensemble of XGBoost, CatBoost, and LightGBM achieves better magnification-wise average accuracy than other models we experimented.

IRv2-CXL has several advantages in comparison to other existing models. Using a pre-trained model as a feature extractor, its training time is much less than those models attempting to extract features from images. The use of transfer learning also reduced the proposed model's need for huge amounts of training data, which is almost always in short supply in the medical field. Incorporating an ensemble of GBDT-based algorithms in the IRv2-CXL classifier not only improved its accuracy compared to other models, but also drew attention to the fact that there had not been enough scrutiny in choosing the classifiers in previous studies. Most of the researchers tackling this task have focused on developing better feature extractors. The use of soft majority voting provides IRv2-CXL with a lower variance relative to other models. Since the proposed model performed well with all magnifications, present in BreakHis, having a low accuracy variance across all of them, a CAD system based on it would have an extremely low dependency on the magnification factor. While the proposed model had many advantages, there are some weaknesses associated with is, as well. The first weakness came from a lack of feature extractor fine-tuning. Inception-ResNet-v2's large image data belongs mostly to the general domain, such as ear, corn, and vase; nevertheless, the images in this task involved breast cancer histopathology images with non-identical visual appearances. That may make IRv2-CXL's feature extractor biased with the source data and less generalized on the target data. The second weakness arose from the dataset we used to develop this model and the blatant imbalance in its class distributions. While our model performed well on the F1 score (Table 5) and ROC curve (Fig. 6) as well as accuracy, and its high accuracy did not rely on this imbalance, we believe a more balanced dataset would produce even better results.

In the future, we aim to resolve the proposed model's lack of generalization on the target dataset through fine-tuning and addressing the imbalanced dataset issue with a more comprehensive preprocessing, including an image augmentation that makes the model more balanced. Further improvements can be achieved by incorporating pathologist expertise and the Grad-CAM images we have already produced. These two can help us determine the nonconformities between the areas that help the experts and IRv2-CXL decide on tumor classification. Then, we can construct an attention-based model targeting those areas most helpful to correct classification.

5. Conclusion

This paper, tackled the issue of breast tumor type (benign or malignant) estimation by processing histopathological images obtained from the lab. This tumor-type assessment remains an unsolved problem in early cancer detection. By resolving it with the help of an automatic classification system, we can contribute greatly to the medical community and ordinary people, saving them money, time, and above all, lives.

IRv2-CXL, a novel binary classification method for CAD systems is introduced in this work. Most CAD systems consist of three phases. The focus here was on the classification phase. This approach implemented deep features extracted from a pre-trained network and a classifier to distinguish between benign and malignant tumors. Sixteen different pre-trained networks were tested with seven classifiers (three single and four ensembles). The best model was an Inception-ResNet-v2 network, and a soft majority voting ensemble of gradients boosted decision trees (i.e., CatBoost, ...) named IRv2-CXL. IRv2-CXL was trained on and then validated, using the BreakHis dataset. IRv2-CXL achieved an accuracy of 96.83%, 95.84%, 97.02%, and 96.16% on the 40×, 100×, 200×, and 400× magnification factors, respectively. Furthermore, a class activation mapping was also implemented to identify important regions in the image that helped IRv2-CXL classify tumors to ensure the proposed model did not concentrate on any unrelated parts.

We believe that further improvements to this model or the ideas it is based on are entirely possible and within reach. We hope this study will attract others to explore the presented ideas and build upon our and other existing research findings to introduce even better methods with more outstanding results.

6. Data and code availability

The BreakHis dataset (version 1) is publicly available at the Laboratory of Vision, Robotics and Imaging of the Federal University of Parana website.[4] In addition, the source code of the proposed model required to reproduce the predictions with the results is available at the public Github repository.[5]

CRediT authorship contribution statement

Mohammad Reza Abbasniya: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Sayed Ali Sheikholeslamzadeh:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Hamid Nasiri:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Writing – review & editing, Supervision. **Samaneh Emami:** Writing – review & editing, Supervision, Project administration.

Declaration of competing interest

No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to https://doi.org/10.1016/j.compeleceng.2022.108382.

Acknowledgments

The authors would like to thank Saeed Chehreh Chelgani (from the Lulea University of Technology) and Mohammad Ezzoddin (from the University of Tehran), who voluntarily helped us with this paper.

References

[1] Hajiabadi Hamideh, Babaiyan Vahide, Zabihzadeh Davood, Hajiabadi Moein. Combination of loss functions for robust breast cancer prediction. Comput Electr Eng 2020;84:106624.
[2] Mohammed Mazin Abed, Al-Khateeb Belal, Rashid Ahmed Noori, Ibrahim Dheyaa Ahmed, Abd Ghani Mohd Khanapi, Mostafa Salama A. Neural network and multi-fractal dimension features for breast cancer classification from ultrasound images. Comput Electr Eng 2018;70:871–82.
[3] Punitha S, Al-Turjman Fadi, Stephan Thompson. An automated breast cancer diagnosis using feature selection and parameter optimization in ANN. Comput Electr Eng 2021;90:106958.
[4] Alirezazadeh Pendar, Hejrati Behzad, Monsef-Esfahani Alireza, Fathi Abdolhossein. Representation learning-based unsupervised domain adaptation for classification of breast cancer histopathology images. Biocybern Biomed Eng 2018;38(3):671–83.
[5] Fondón Irene, Sarmiento Auxiliadora, García Ana Isabel, Silvestre María, Eloy Catarina, Polónia António, Aguiar Paulo. Automatic classification of tissue malignancy for breast carcinoma diagnosis. Comput Biol Med 2018;96:41–51.
[6] George Kalpana, Faziludeen Shameer, Sankaran Praveen, et al. Breast cancer detection from biopsy images using nucleus guided transfer learning and belief based fusion. Comput Biol Med 2020;124:103954.
[7] Benhammou Yassir, Achchab Boujemaa, Herrera Francisco, Tabik Siham. BreakHis based breast cancer automatic diagnosis using deep learning: Taxonomy, survey and insights. Neurocomputing 2020;375:9–24.
[8] Macenko Marc, Niethammer Marc, Marron James S, Borland David, Woosley John T, Guan Xiaojun, Schmitt Charles, Thomas Nancy E. A method for normalizing histology slides for quantitative analysis. In: 2009 IEEE international symposium on biomedical imaging: From nano to macro. IEEE; 2009, p. 1107–10.
[9] Choudhary Tejalal, Mishra Vipul, Goswami Anurag, Sarangapani Jagannathan. A transfer learning with structured filter pruning approach for improved breast cancer classification on point-of-care devices. Comput Biol Med 2021;134:104432.
[10] Barsha Nusrat Ameen, Rahman Aimon, Mahdy MRC. Automated detection and grading of Invasive Ductal Carcinoma breast cancer using ensemble of deep learning models. Comput Biol Med 2021;139:104931.
[11] Gonçalves Caroline Barcelos, de Souza Jefferson R, Fernandes Henrique. CNN architecture optimization using bio-inspired algorithms for breast cancer detection in infrared images. Comput Biol Med 2022;105205.
[12] Chatterjee Somnath, Biswas Shreya, Majee Arindam, Sen Shibaprasad, Oliva Diego, Sarkar Ram. Breast cancer detection from thermal images using a Grunwald-Letnikov-aided Dragonfly algorithm-based deep feature selection method. Comput Biol Med 2022;141:105027.
[13] Bozinovski Stevo. Reminder of the first paper on transfer learning in neural networks, 1976. Informatica 2020;44(3).
[14] Szegedy Christian, Ioffe Sergey, Vanhoucke Vincent, Alemi Alexander A. Inception-v4, inception-resnet and the impact of residual connections on learning. In: Thirty-first AAAI conference on artificial intelligence. 2017.
[15] Nasiri Hamid, Alavi Seyyed Ali. A novel framework based on deep learning and ANOVA feature selection method for diagnosis of COVID-19 cases from chest X-Ray images. Comput Intell Neurosci 2022;2022:4694567.
[16] Nasiri Hamid, Hasani Sharif. Automated detection of COVID-19 cases from chest X-ray images using deep neural network and XGBoost. Radiography 2022;28(3):732–8.
[17] Nasiri Hamid, Kheyroddin Ghazal, Dorrigiv Morteza, Esmaeili Mona, Nafchi Amir Raeisi, Ghorbani Mohsen Haji, Zarkesh-Ha Payman. Classification of COVID-19 in chest X-ray images using fusion of deep features and LightGBM. In: 2022 IEEE World AI IoT Congress. AIIoT, IEEE; 2022, p. 201–6.
[18] Ezzoddin Mobina, Nasiri Hamid, Dorrigiv Morteza. Diagnosis of COVID-19 cases from chest X-ray images using deep neural network and LightGBM. In: 2022 international conference on machine vision and image processing. MVIP, IEEE; 2022, p. 1–7.
[19] Chelgani Saeed Chehreh, Nasiri Hamid, Alidokht Mehdi. Interpretable modeling of metallurgical responses for an industrial coal column flotation circuit by XGBoost and SHAP-A "conscious-lab" development. Int J Mining Sci Technol 2021;31(6):1135–44.

[4] https://web.inf.ufpr.br/vri/databases/breast-cancer-histopathological-database-breakhis
[5] https://github.com/mohammadAbbasniya/BreastCancer-Detection

[20] Chelgani Saeed Chehreh, Nasiri Hamid, Tohry Arash. Modeling of particle sizes for industrial HPGR products by a unique explainable AI tool-A "Conscious Lab" development. Adv Powder Technol 2021;32(11):4141–8.
[21] Akhavan Zeinab, Esmaeili Mona, Sikeridis Dimitrios, Devetsikiotis Michael. Internet of Things-enabled passive contact tracing in smart cities. Internet of Things 2021;100397.
[22] Nasiri Hamid, Homafar Arman, Chelgani Saeed Chehreh. Prediction of uniaxial compressive strength and modulus of elasticity for Travertine samples using an explainable artificial intelligence. Results Geophys Sci 2021;8:100034.
[23] Gupta Deepak, Rani Rinkle. Improving malware detection using big data and ensemble learning. Comput Electr Eng 2020;86:106729.
[24] Sharma Shallu, Kumar Sumit. The Xception model: A potential feature extractor in breast cancer histology images classification. ICT Express 2022;8(1):101–8.
[25] Kumar Abhinav, Singh Sanjay Kumar, Saxena Sonal, Lakshmanan K, Sangaiah Arun Kumar, Chauhan Himanshu, Shrivastava Sameer, Singh Raj Kumar. Deep feature learning for histopathological image classification of canine mammary tumors and human breast cancer. Inform Sci 2020;508:405–21.
[26] Fatahi Rasoul, Nasiri Hamid, Dadfar Ehsan, Chehreh Chelgani Saeed. Modeling of energy consumption factors for an industrial cement vertical roller mill by SHAP-XGBoost: a" conscious lab" approach. Sci Rep 2022;12(1):1–13.
[27] Du Baolin, Qi Qi, Zheng Han, Huang Yue, Ding Xinghao. Breast cancer histopathological image classification via deep active learning and confidence boosting. In: International conference on artificial neural networks. Springer; 2018, p. 109–16.
[28] Li Xin, Li HongBo, Cui WenSheng, Cai ZhaoHui, Jia MeiJuan. Classification on digital pathological images of breast cancer based on deep features of different levels. Math Probl Eng 2021;2021.
[29] Joseph Agaba Ameh, Abdullahi Mohammed, Junaidu Sahalu Balarabe, Ibrahim Hayatu Hassan, Chiroma Haruna. Improved multi-classification of breast cancer histopathological images using handcrafted features and deep neural network (dense layer). Intell Syst Appl 2022;14:200066.
[30] Zerouaoui Hasnae, Idri Ali. Deep hybrid architectures for binary classification of medical breast cancer images. Biomed Signal Process Control 2022;71:103226.

Mohammad Reza Abbasniya started his B.Sc. at Semnan University, Iran, in 2018 and took his Computer Engineering degree in 2022. His research interest consists of artificial intelligence applications in healthcare and medicine. Also, he has experienced natural language processing and speech recognition.

Sayed Ali Sheikholeslamzadeh graduated from Semnan University, Semnan, Iran, with a bachelor of computer engineering degree in 2022. His research interests include natural language processing, computer vision and deep learning.

Hamid Nasiri received the M.Sc. degree in Computer Engineering from the Amirkabir University of Technology, Tehran, Iran, in 2016. He is currently a Ph.D. candidate at the Department of Computer Engineering, Amirkabir University of Technology, under the supervision of Dr. Mohammad Mehdi Ebadzadeh. His research interests include machine learning, deep learning, medical image analysis, and fuzzy systems.

Samaneh Emami is an Assistant Professor in the Department of Electrical and Computer Engineering at Semnan University. She finished her Ph.D. studies in the Department of Computer Engineering at the Amirkabir University of Technology in 2017 and her B.Sc. and M.Sc. degrees (with great distinction) from Shahid Beheshti University.