Bulletin of Environment, Pharmacology and Life Sciences
Bull. Env. Pharmacol. Life Sci., Special Issue [1]2022 : 1580-1590

Online ISSN 2277-1808
Journal's URL:http://www.bepls.com
CODEN: BEPLAD

ORIGINAL ARTICLE **OPEN ACCESS**

An Automated Histopathology Image Classification using a Hybrid Model

Nishant Behar*and Manish Shrivastava
[1]Guru Ghasidas University Bilaspur, India
*Corresponding author: nishant.itggv@gmail.com

ABSTRACT

Breast cancer is a disease where cells grow abnormally inside the breast. Malignant tumors are a threat for women worldwide. Early diagnosis is necessary to reduce the mortality rate due to breast cancer. Recently, many attempts have been made for its detection using computer-based systems. This paper also focuses on the automatic classification of breast cancer images into two classes: ductal carcinoma and benign. The sole purpose of this study is to propose an accurate model to correctly determine the classes for a set of unknown (Test) images. In this experimental work, a total of four machine learning techniques have been developed and their results are compared. Two models present a hybrid approach based on the pre-trained Inception-V3 with Random Forest (RF) and Support Vector Machine (SVM) models. These hybrid models have performed excellently, and when compared together, the Inception based SVM model has delivered better performance in all aspects with 0.33% higher accuracy, 0.4% improved ROC-AUC value, and 0.69% higher Cohen's kappa coefficient value. Additionally, other parameters' precision, recall and F1 values were also significantly enhanced. The proposed Inception based SVM model is a valid and effective classifier for the employed dataset based on the experimental outcomes.

Keywords: *Benign; Ductal Carcinoma; Breast cancer; Classification; Image processing; CNN; SVM; Random Forest*

Received 10.03.2022 Revised 16.03.2022 Accepted 19.04.2022

INTRODUCTION

According to a cancer statistics report,[1] 2,81,550 new breast cancer cases is expected in 2021 in the United States only, out of which43,600 deaths may occur due to it. A review study[2] revealed that breast cancer is a significant cause of mortality worldwide and highly impacts the quality of life of women in all respects. Breast cancer can be cured or controlled when detected in the initial stage, unlike other cancers. Accurate detection of breast cancer is a crucial step for deciding the line of treatment. There are invasive methods like Fine needle aspiration(FNA), open surgery for tumor detection.

In thenon-invasive radiology process, unique images are analyzed for finding cancer regions. For example, in the mammogram examination, the patient is exposed to low-powered X-rays to detect micro-calcification or mass in the breast to detect probable tumor areas. One another technique involves analyzing histopathology images, which provides more details about cancer. Usually, histopathology images have large size [3] and may have different color intensity levels and complex patterns. These image variations make the manual diagnostic method difficult. A whole image analysis is very tedious, laborious, and time-expensive task and mainly influenced by the expertise level of the specialist. Computer-aided diagnosis (CAD) systems, especially deep learning methods, can investigate and automatically classify images and put into the appropriate class. This classification process involves training a model to learn a particular disease pattern. Each pixel of a picture is read in this process, which results according to the learning algorithm used.

The advantages of using automated systems are that they provide fast, reliable, and consistent results and increase the ability to detect minor abnormalities, which is generally not possible using human faculties alone.

Our study proposes a computer-based, effective technique for automatically identifying Ductal Carcinoma (DC) images over the Benign classes. The novelty lies in the two fusion models, combining Convolutional Neural Network (CNN) with the traditional methods. In the first stage, the fully connected (FC) layer of the Inception model was redesigned, and feature reduction was carried out. In the other stage, the hyper parameters of the traditional methods were optimized before applying them for the classification tasks. Further, instead of applying a pre-trained weight of the model, entire layers were retrained for the more

suitable weights and model learning, specifically for the employed dataset. Section 3 provides the complete details about the methodology used. All DC and Benign images are collected from an open benchmark dataset [4].

RELATED WORK

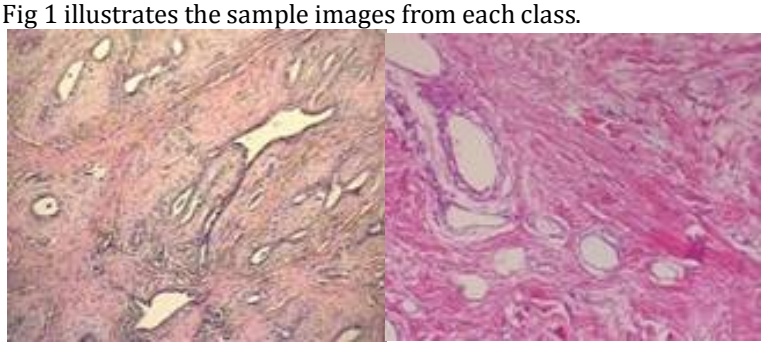
Literature survey on computer-aided diagnostic (CAD) systems using medical images

Machine learning methods and their advancements attract numerous researchers,who offer much novel work in this area, putting continuous effort into building more consistent and reliable CAD systems. One research [4] has developed a computer-aided system for finding the biomarker classifying breast cancer patients through Convolutional Neural Network (CNN). Stroma properties from biopsy tissue images have been assessed with these networks' potentially discriminatory power. The authors found that stroma properties could be used as a biomarker for cancer detection. This system obtained 92% as the area under the curve for breast cancer diagnosis. A novel experiment for lung cancer classification extracted and used deep features. [5].Another study[6] developed an edge detector system for diabetic foot ulcer images. Various edge detectors were employed in this work, finding the highest mean structural similarity value of 99.92%. Autoencoder was used for the feature extraction, and binary decision tree was used for classification purposes. Classification accuracy was 75.01% for benign and malignant lung nodules. Research[7] has demonstrated that residual networks have a two-fold advantage: (a)easy to optimize (b)higher accuracy on increasing network depth. The analysis reported that increasing residual nets with 152 layers on ImageNet test data showed lower complexity. Experiments were extended for benchmark datasets CIFAR-10, CIFAR-100, and 1000. Some authors[8] analyzed the impact of manual features and the patient's information for improved performance of the CNN network. A dataset of 45000 was used in the research. The researchers concluded that location and context features are useful for network learning and yield a better result. In another study [9], the authors used different magnification levels of breast cancer histopathology images. The first model for single-task-CNN was trained to predict the malignant class. In contrast, in multitasks, CNN was applied in the second model to detect malignancy and image magnification level on the Break His dataset. The authors in [4] have contributed 7909 breast cancer histopathology images of binary classes Benign and Malignant. These images were acquired from 82 patients. The dataset details are given in the experimental section of this paper. The diagnosis and classification task was the primary objective of this dataset. With the standardized process, accuracy ranges between 80–85% were obtained. In an extended experimental work, authors [10] used CNN for feature extraction (DeCAF) and applied it for classification. This paper shows that features extracted from CNN outperform traditional handcrafted features. In one important research work [11], the authors proposed a solution for difficulties associated with the traditional diagnostic procedure when different specialists offered ambiguous opinions on the same case. As a solution, a Deep Neural Network was used for feature extraction. The classification was done on Support Vector Machine (SVM). Accuracy was 77.8% for multiple classes and 83.3% for binary classification. In one interesting research work [12], the authors successfully applied a VGG-16 pre-trained model and SVM with early and late fusion models to classify the Rose leaf disease detection.They obtained good accuracy of 90.26%. In another study [13],the researchers proposed an automatic technique and evaluated Google and AlexNet for the detection of mass lesions in mammography images. They also concluded that CNN is a suitable approach for the classification of mass lesions and performs exceptionally well on a small number of image samples. Yet another study [14] presented a novel framework SSDLite for the purpose of object segmentation. The authors also developed a reduced DeepLAbv3 as Mobile DeepLAbv3, which was based on the inverted residual structure. Imagenet dataset was applied to evaluate the performance of the proposed model. The tradeoff between accuracy—the number of operations—and actual latency—the number of parameters—was also measured. One research[15] observed that convolution network (CNN) showed better performance when connection densities were shorter at the input and output ends. The experiment was carried out on different datasets CIFAR-10, CIFAR-100, SVHN, and ImageNet, where DenseNet showed significant improvement. The study by[16] focused on CNN architectural issues for the same classification accuracy of a dataset. They proposed a squeeze net, which achieved the same accuracy with only half of the parameters applied in the Alexnet on the image dataset. The authors in [17] proposed a machine learning-based segmentation model for retinal blood vessels.The color features had been extracted from eye fundus images and fed into the back propagation neural network for the classification. The model can segment blood vessels precisely in less time compared to other similar work.

The research detailed above highlights the importance of computer-based CNN systems and their adoption for analyzing medical images precisely, pixel by pixel, to identify underlying patterns from images. These are also suitable for distinguishing Normal, Benign, or Malignant test samples by using a

large number of high-resolution images and machine learning algorithms. Furthermore, CNN also assistsin segmentation and feature reduction and classification [18]–[24].

Therefore, this research paper offers a practical model with a high accuracy rate to classify histopathology images into Ductal Carcinoma and Benign classes. The proposed solution is based on a pretrained inception-v3 model, and performance was tested on test datasets which were unseen for the developed models. Fig 1 illustrates the sample images from each class.



Benign tumor Malignant (DC) tumor

Figure 1: Histopathology images of (a) Benign and (b) Malignant cells [4]

MATERIAL AND METHODS

THE INCEPTION MODEL

Inception-v3 has been used in this present work, which is already trained on the ImageNet dataset. The complete architecture is shown in Fig2 [25]. The main impact of this Inception model lies in their inception layers, which are the combinations of filters 1*1, 3*3, and 5*5 to construct convolution layers. In this research, the convolution and other layers, except FC layers, have been kept trainable to exploit feature extraction ability, especially for local and general features.

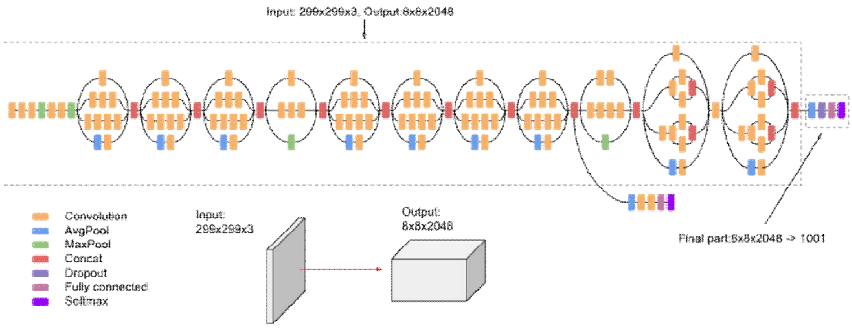


Figure 2 : Illustrating Inception-v3 architecture [25]

Feature extraction by convolution layers and pooling layers

Convolution and pooling layers play a vital role in the feature extraction process. The convolution process uses a mathematical operation, as shown in Fig 3. Several convolved layers can be applied in a single CNN architecture for processing one-, two-, or three-dimensional images. Filter or feature extractors are provided as n*n matrix, where 'n' is the number of rowand column dimensions and n = 2,3,4,5… All pixels that come under the filter area (n*n) over a 2D are read. All the corresponding pixel values of image pixels are multiplied and summed up to generate a feature map value for that specified region. Each filter slides from left to right and top to bottom to read a different set of pixels at different stages. Filter size must be smaller than the input image size.

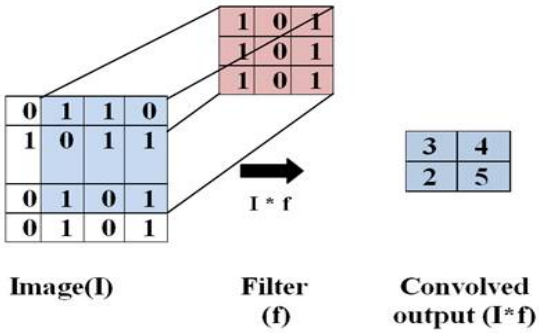


Figure 3: Illustration of Convolution layer process

The stride value determines the pixel shift over the input image to read the next block of pixels. This process is repeated until all possible pixels have been read. A mathematical operation (dot product) of the convolved image (input matrix) patch and various filter matrices gives the layer's output. Different filters are used to detect different shapes and textures, edges, and gradients. Generally, this matrix is followed by the pooling layer, which significantly reduces the spatial matrix size to almost half the size (on stride=1), refer Fig 4. Reducing the size of the matrix helps to reduce overall parameter size and the mathematical operation, which ultimately offers faster operation in the network. The pooling layer stores the most important values received from the convolution layer. These values or weights can determine specific patterns and distinguish images as per their possible classes. The convolution layer may use padding for extracting more useful values at the border of the input image. These values are applied as layer weights for extracting relevant features from the unknown image.

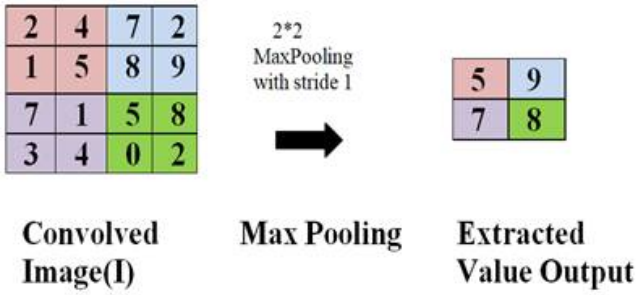


Figure 4: Illustration of application of filter and Max Pooling functions

Support Vector Machine (SVM)

SVM algorithm uses statistical theory, which can separate different classes using small datasets. A few specific data points, i.e., support vectors, are used to determine the hyperplane. This hyperplane splits data into two or more classes. The line on both sides of the hyperplane defines the decision boundary [26]. The main intention of SVM is to maximize the distance between boundaries that presents both sides of the hyperplane. Different kernel tricks are available for non-linear data samples [27]–[29]. A higher hyperplane margin indicates a higher classification rate.

Random Forest (RF)

The ensemble classifier Random Forest (RF) technique [30]–[34] is applied to improve the classification accuracy. Random Forest has a high classification rate and a minimal overfitting issue compared to other well-known classification algorithms. Several decision trees are generated to form a random forest. The subsamples and the nodes are selected randomly. Each tree predicts a label when a sample is supplied to it. The final label is determined based on the maximum vote having a label from all the trees.

The Proposed Model

Four models, RF, SVM, IRF, and ISVM were evaluated in the experimental work. The supervised models RF and SVM were straightforward in following the steps (a) preprocessing of all input images, (b) training data used for model learning, (c) test subset being used to evaluate the models. The two hybrid models (IRF and ISVM) were based on the transfer learning concept and employed the Inception pretrain fine-tuned model. To obtain the optimum benefit from the Inception-v3 model, all FC layers were redesigned and retrained according to the configurations given in Table 1. The purpose of using transfer learning was to extract prominent features only and reduce the feature set.

Table1: Hyper-parameters details and parameter configuration for fine-tuned inception model

Attributes	FC-1	FC-2	FC-3	Output layer
Neurons	1024	512	128	1
Activation	Relu	Relu	Relu	Sigmoid

In the first model, i.e. IRF, an Inception model was hybrid with Random forest (RF) classifier developed to perform binary classification; refer to Fig 5(top).
The fusion technique was applied between inception-v3 and SVM in the second model, as depicted in Fig 5 (bottom). A benchmark image dataset (BreakHis) was used to collect all Benign and Ductal carcinoma (DC) images. The database contains several Ductal Carcinoma images, almost equal to the number of all Benign images.

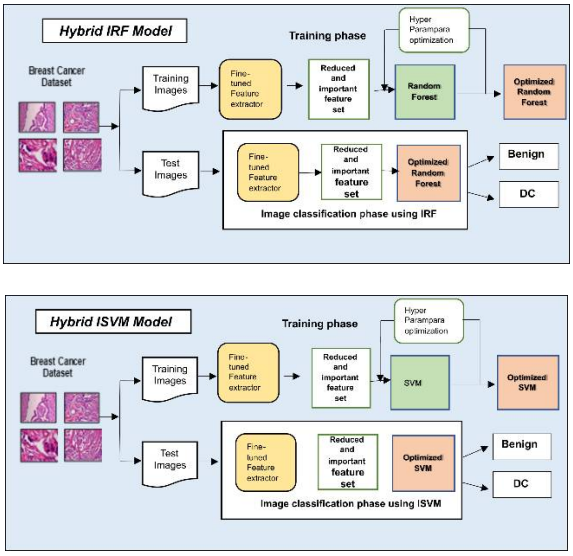


Figure 5: Figure showing the hybrid models (a) IRF model (top) (b) ISVM model(bottom)

Data prepossessing

Well-processed images are easy to visualize, and help to distinguish different elements present in the image. Preprocessing is the prime task in improving the image quality and composing images ready for feeding to the CNN model. If an image is improper or of substandard quality, it is hard to understand important patterns using the CNN model or it may mislead; therefore, the expected outcome is not achieved. The deep neural network requires a large number of samples to learn themodel effectively. The number of sample images available for this experiment was low and insufficient to train the proposed models properly; therefore, an augmentation technique was implemented with two objectives: (a) to increase the sample size with new images and (b) to handle the over-fitting issue effectively. Input datasets are augmented with various features like flip horizontal, vertical, and rotate.
The models' performance was robustly evaluated on a set of 1470 images for validation, and a separate set of 297 images were used as test images. The image batch size was set to 32, and all three channels for each input image were adopted. Image sizes were also downsized to 256*256 pixels. The validation and test dataset were only resized to 256*256 and normalized pixel-wise, as given in Equation (1), to maintain original image values within the range [0, 1].

Normalized pixel (P_n) =

$$\sum \frac{(\text{pixel actual value } (P_i) - \text{mean of all pixel } (P_m))}{255} \quad (1)$$

Here, Pn is the processed pixel and Pi, Pm are the original and mean pixel values.
At the model evaluation stage, actual class labels of test datasets were eliminated before allocating them to the built models (RF, SVM, IRF, and ISVM). Then the model's predictions are compared with the actual label of the test subset images. Benign and Ductal Carcinoma are labelled as 0 and 1, respectively, as per their alphabetical order. The main aim of this work was to classify histopathology images effectively into binary classes as Benign or Ductal Carcinoma (i.e., Malignant tumors).

Feature Extraction

Feature extraction provides strength to the image classification and recognition of local or global features. This convolution layer scans the complete image in a certain order and extracts only valuable information [35]–[41], i.e., feature maps. In the hybrid models,a fine-tuned inception model was executed, and once the specified epochs were completed the feature extraction took place. The same features were fed into both hybrid models to assess model performance fairly; refer to Fig 6. The ImageNet weights were initially supplied to begin model training. All layers were kept re-trainable, which took computational time but provided better features.

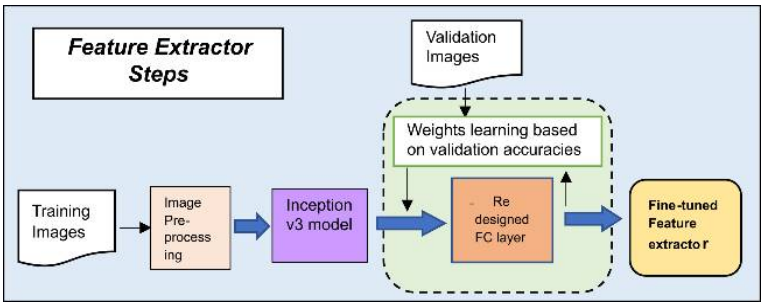


Figure 6 : Feature Extraction and reduction process

Evaluation parameters and metrics

Model performance was evaluated based on the correct classification rate and confusion matrix generated on test samples. The confusion matrix for binary classification is a 2*2 matrix; Table 2 shows the classification characteristics between the model prediction labels and actual class labels of test image samples. Because of inherent biases caused by unbalanced class samples, classification accuracy alone cannot reveal actual model performance. Several metrics such as recall, F1, and ROC-AUC score have been computed for robust model evaluation.

		Predicted classes	
		No disease	disease
Actual classes	No disease	TP	FN
	Disease	FP	TN

Table 2: Representing 2*2 Confusion matrix for the binary classification

Mathematical expressions for Accuracy, Recall, Precision, and the F1 score are given in Equations (2) to (5).TP, TN, FN, FP represent True Positive, True Negative, False Negative, and False Positive, respectively. Other important parameters are defined as follows :

$$\text{accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (2)$$

$$\text{Precision} = \frac{TP}{(TP+FP)} \quad (3)$$

$$\text{Recall} = \frac{TP}{(TP+FN)} \quad (4)$$

$$\text{F1 score} = \frac{2*(\text{Precesion}*\text{Recall})}{(\text{Precesion}+\text{Recall})} \quad (5)$$

RESULTS AND DISCUSSION

The network was trained for 45 epochs, and the batch size was set to 32. The dataset augmented for training, validation, and test samples had 4119, 1470, and 297 images, respectively. The supervised SVM algorithm was optimized with the C=9 and kernel='Linear', whereas Random Forest was tuned with 3615 estimators. Hyperparameters' values were chosen randomly instead of applying every option. The ISVM model outperformed the test dataset among all four models, where fewer features were extracted using the Inception-v3 model. These features have high discrimination values, which were fed into the optimized SVM classifier. The same feature set was applied to the IRF model. The same hyperparameter values were applied in the traditional SVM as well as Random Forest classification. All the practical work was carried out using scikit-learn [42] on python 4.1 employing theColab cloud platform.

Test Accuracies

Among all the classifiers, IRF and ISVM architectures illustrated promising results, where IRF provided a test accuracy of 99.33%, and hybrid ISVM with Linear kernel provided the highest test accuracy of 99.66 %.

Confusion Matrix

Table 2 : Comparison of RF, SVM, IRF, and ISVM using a confusion matrix

Model	Class	Precision	Recall	F1-score
RF	0	0.9	0.84	0.87
	1	0.89	0.94	0.91
SVM-Linear	0	0.78	0.66	0.71
	1	0.78	0.87	0.82
IRF	0	**0.99**	**0.99**	**0.99**
	1	**0.99**	**0.99**	**0.99**
ISVM (proposed)	0	**0.99**	**1**	**1**
	1	**1**	**0.99**	**1**

Class wise evaluation of all models

The prime objective of this research work was to achieve a high classification accuracy rate for identifying unlabeled binary class DC and Benign images collected from theBreakHis dataset.Other essential parameters, such as precision, recall value, and F1 score were also observed for theactual model performance; refer to Table3.

Table3: Class-wise individual model performance on precision, recall, and F1 accuracy

F		SVM	
103	20	81	42
11	163	23	151
IRF		ISVM	
122	1	123	0
1	173	1	173

Here,'0' and '1' represent the Benign and DC ***classes***.The class support is 123 for '0' (Benign) and 174 for the '1' (DC) classes.

Area under the curve (AUC) comparison

The AUC is a graphical representation of model performance for classification problems at different thresholds. It indicates how well the model can differentiate across classes (see Figure 8). A higher accuracy value of AUC indicates better ability of the model to distinguish samples into their exact class. The AUC value for the IRF was 0.9930 and for the ISVM it was 0.4 % higher at 0.9971.

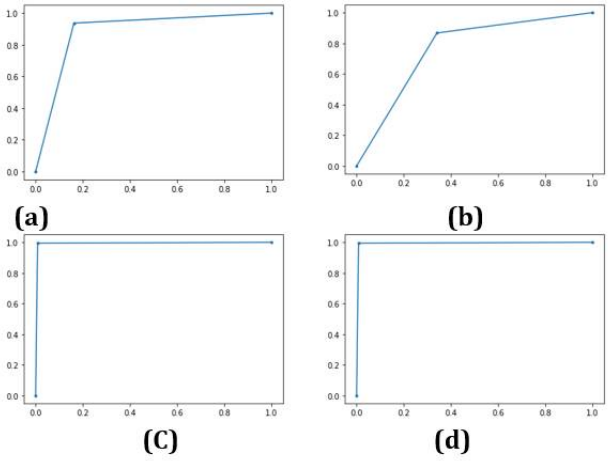


Figure 8: ROC visualization for models (a) RF (top left) (b) SVM (top right) (c) IRF (bottom left) and (d) ISVM based on ROC

Performance measures on different evaluation parameters

Table 4: Weighted average accuracies of each classifier (test samples= 297)

Model	Accuracies	Precision	Recall	F1-score
RF	0.89	0.9	0.9	0.89
SVM-Linear	0.78	0.78	0.78	0.78
IRF (proposed)	**0.9933**	**0.99**	**0.99**	**0.99**
ISVM (proposed)	**0.9966**	**1**	**1**	**1**

The top two performers, IRF and ISVM, were further compared on all the evaluated parameters, as shown in Figure9.

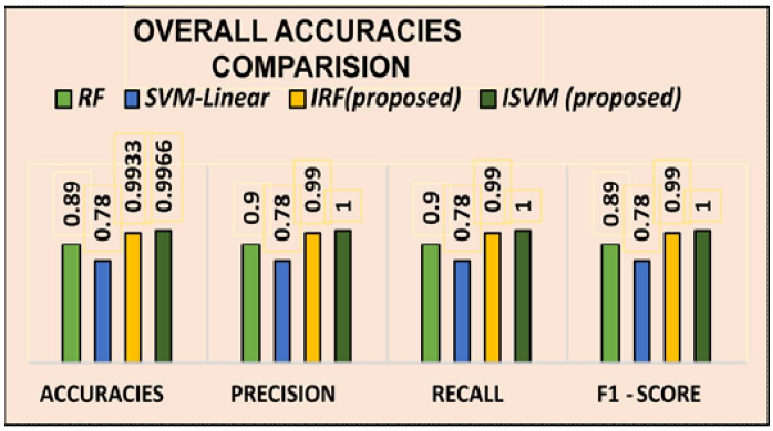


Figure 9 : Performance comparison of IRF and ISVM on precision, recall, F1 score

Cohen's kappa is a measure to evaluate the validity of the classification model based on the confusion matrix, which determines a coefficient that represents the agreement (correct prediction) and disagreement (incorrect prediction) on the classification of the data samples. A coefficient value near 1 represents the model is suitable, whereas a value closer to 0 demonstrates that the model is inadequate. While comparing the top two performers, it is found that ISVM has 0.69% higher coefficient value than the IRF model. Furthermore, ROC-AUC is also 0.4% better than the IRF model. These two cases qualify the ISVM model as the best classifier (refer Figure 10).

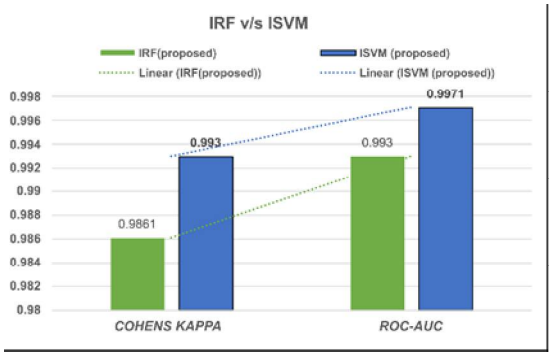


Figure 10: Illustration of performance of IRF and ISVM based on Cohen's kappa and ROC-AUC

Comparison with the similar work

The main objective of this research was to effectively distinguish DC images and benign images available in Breakhis dataset. For this purpose, all DC images have been collected in a group, while all benign images have been gathered in another class. Any similar study, using the above classes, could not be

found. Therefore, we have compared results with the four algorithms, RF, SVM, IRF and ISVM, as shown in section 4. However, in several studies, it has been found that the experimental setup used all malignant and benign classes. Therefore, comparison with these similar works has been provided in Table 5.

Table 5: Comparative analysis of the IRF and ISVM with similar work.

Paper	Year	Classification Technique	Highest Accuracy for Breakhis Dataset	Classification type	Magnification type and factor	Classes used for binary classification
[43]	2018	CNN + SVM	98.33	Binary	Magnification dependent (40X)	All Benign and all Malignant
[44]	2020	Resnet50 with 40layers	99.26	Binary	magnification independent	All Benign and all Malignant
[45]	2021	DenseNet201 and XGBoost	97%	Binary	Whole Slide Image	All Benign and all Malignant
[46]	2021	DCNN (Binary class)	98.36	Binary	Magnification dependent 200X	All Benign and all Malignant
[47]	2021	DenseNet-201 + NasNetMobile + VGG16 + Fine-tuned	99%	Binary	patient wise	All Benign and all Malignant
Proposed Models					WSI and Magnification independent	DC and all benign classes
	IRF	**Inception+ RF**	**99.33%**	Binary		
	ISVM	**Inception+ SVM**	**99.66%**			

CONCLUSION AND FUTURE SCOPE

The primary goal of this study was to propose an accurate model for identifying ductal carcinoma and benign tumor classes from given histopathology images. Our proposed Inception-SVM emerged on top in every criterion when compared with other models such as Random Forest, Support vector machine and inception-Random Forest classifier. The suggested model is rapid, accurate and cost-effective with the highest accuracy rate of 99.66%. The valuable and reduced features were extracted for the hybrid models. When comparing both hybrid models together then Inception-SVM model has performed better than the Inception-RF model with 0.33% higher accuracy, 0.4% increased ROC-AUC value and 0.69% better Cohen's kappa score. Precision, recall, and F1 scores were also found better than the Inception- RF model. The model reliability was determined by the accuracy, precision, recall, and F1 scores, whereas Cohen's kappa value defines the model's legitimacy.

This study can be further extended for including all malignant class images. The FC layer can be analyzed further for its hyperparameter tuning, such as dropouts, number of neurons, hidden layers, etc.

Although numerous improvements have been made to CAD systems, healthcare specialists' opinion is necessary to make final decisions due to the intricacy of the work. This study offers an automated tool to classify DC images from benign images as a primary screening tool.

ACKNOWLEDGEMENT

The authors would like to acknowledge VRI lab b (Spanhol et al.) for providing the Breast Cancer Histopathological Image Dataset on request for this academic and research purpose at https://web.inf.ufpr.br /vri/databases/breast-cancer-histopathological-database-breakhis/.

The authors also acknowledge google.com for making available the Inception-v3 high-level diagram model at https://cloud.google.com /tpu/docs /inceptionv3-advanced

FUNDING STATEMENT

The author(s) received no specific funding for this study.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest to report regarding the present study

REFERENCES

1. R. L. Siegel, K. D. Miller, H. E. Fuchs, and A. Jemal, (2021).“Cancer Statistics, 2021,” *CA: A Cancer Journal for Clinicians*, vol. 71, no. 1, pp. 7–33. doi: 10.3322/caac.21654.
2. M. M. Sunilkumar, C. G. Finni, A. S. Lijimol, and M. R. Rajagopal, (2021).“Health-Related Suffering and Palliative Care in Breast Cancer,” *Current Breast Cancer Reports*, vol. 13, no. 4, pp. 241–243. doi: 10.1007/s12609-021-00431-1.
3. Y. Liu *et al.*,(2021).“Detecting Cancer Metastases on Gigapixel Pathology Images,” *arXiv:1703.02442 [cs]*, Mar. 2017, Accessed: [Online]. Available: http://arxiv.org/abs/1703.02442.
4. F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, (2016).“A Dataset for Breast Cancer Histopathological Image Classification,” *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 7, pp. 1455–1462. doi: 10.1109/TBME.2015.2496264.
5. D. Kumar, A. Wong, and D. A. Clausi, (2015). “Lung Nodule Classification Using Deep Features in CT Images,” in *2015 12th Conference on Computer and Robot Vision*, Halifax, NS, Canada, pp. 133–138, doi: 10.1109/CRV.2015.25.
6. R. Mwawado, B. Maiseli, and M. Dida, (2020).“Robust Edge Detection Method for the Segmentation of Diabetic Foot Ulcer Images,” vol. 10, no. 4, p. 7. doi: 10.48084/etasr.3495.
7. K. He, X. Zhang, S. Ren, and J. Sun, (2016).“Deep Residual Learning for Image Recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 770–778, doi: 10.1109/CVPR.2016.90.
8. T. Kooi *et al.*, (2017).“Large scale deep learning for computer aided detection of mammographic lesions,” *Medical Image Analysis*, vol. 35, pp. 303–312, doi: 10.1016/j.media.2016.07.007.
9. N. Bayramoglu, J. Kannala, and J. Heikkila, (2016). “Deep learning for magnification independent breast cancer histopathology image classification,” in *2016 23rd International Conference on Pattern Recognition (ICPR)*, Cancun, pp. 2440–2445, doi: 10.1109/ICPR.2016.7900002.
10. F. A. Spanhol, L. S. Oliveira, P. R. Cavalin, C. Petitjean, and L. Heutte,(2017). “Deep features for breast cancer histopathological image classification,” in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Banff, AB, pp. 1868–1873, doi: 10.1109/SMC.2017.8122889.
11. T. Araújo *et al.*, (2017).“Classification of breast cancer histology images using Convolutional Neural Networks,” *PLOS ONE*, vol. 12, no. 6, p. e0177544. doi: 10.1371/journal.pone.0177544.
12. S. Nuanmeesri, (2021).“A Hybrid Deep Learning and Optimized Machine Learning Approach for Rose Leaf Disease Classification,” *Engineering, Technology & Applied Science Research*, vol. 11, no. 5, pp. 7678–7683 doi: 10.48084/etasr.4455.
13. F. Jiang, H. Liu, S. Yu, and Y. Xie, (2017). “Breast mass lesion classification in mammograms by transfer learning,” in *Proceedings of the 5th International Conference on Bioinformatics and Computational Biology*, Hong Kong Hong Kong, pp. 59–62, doi: 10.1145/3035012.3035022.
14. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “MobileNetV2: Inverted Residuals and Linear Bottlenecks,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, Jun. 2018, pp. 4510–4520, doi: 10.1109/CVPR.2018.00474.
15. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger,(2017 “Densely Connected Convolutional Networks,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, Jul. 2017, pp. 2261–2269, doi: 10.1109/CVPR.2017.243.
16. F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, (2016). “SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size,” *arXiv:1602.07360 [cs]*, [Online]. Available: http://arxiv.org/abs/1602.07360.
17. A. N. Saeed, (2020). “A Machine Learning based Approach for Segmenting Retinal Nerve Images using Artificial Neural Networks,” *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5986–5991, doi: 10.48084/etasr.3666.
18. J.-H. Kim and J.-W. Jeong, (2022). “Prediction of Cardiovascular Disease Using Machine Learning Technique—A Modern Approach,” *Computers, Materials & Continua*, vol. 71, no. 1, pp. 855–869, doi: 10.32604/cmc.2022.021582.
19. P. Gao, M. Perkowski, Y. Li, and X. Song, (2022).“An Automated Deep Learning Based Muscular Dystrophy Detection and Classification Model,” *Computers, Materials & Continua*, vol. 71, no. 1, pp. 305–320. doi: 10.32604/cmc.2022.020914.
20. V. Rajinikanth, S. Kadry, and Y. Nam, (2021). “Convolutional-Neural-Network Assisted Segmentation and SVM Classification of Brain Tumor in Clinical MRI Slices,” *Information Technology and Control*, vol. 50, no. 2, pp. 342–356, doi: 10.5755/j01.itc.50.2.28087.
21. V. Jusas and S. G. Samuvel, (2019). “Classification of Motor Imagery Using Combination of Feature Extraction and Reduction Methods for Brain-Computer Interface,” *Information Technology and Control*, vol. 48, no. 2, pp. 225–234. doi: 10.5755/j01.itc.48.2.23091.
22. A. Devipriya, P. Prabhu, K. Venkatachalam, and A. Z. Ibrahim, (2022). “Kernel Granulometric Texture Analysis and Light RES-ASPP-UNET Classification for Covid-19 Detection,” *Computers, Materials & Continua*, vol. 71, no. 1, pp. 651–666. doi: 10.32604/cmc.2022.020820.
23. T. Uktveris and V. Jusas, (2017).“Application of Convolutional Neural Networks to Four-Class Motor Imagery Classification Problem,” *Information Technology And Control*, vol. 46, no. 2, pp. 94–107. doi: 10.5755/j01.itc.46.2.17528.

24. A. M. Alajlan, (2022). "Automatic Heart Disease Detection by Classification of Ventricular Arrhythmias on ECG Using Machine Learning," *Computers, Materials & Continua*, vol. 71, no. 1, pp. 17–33. doi: 10.32604/cmc.2022.018613.
25. Advanced Guide to Inception v3 on Cloud TPU," (2020). *Google Cloud*. https://cloud.google.com /tpu/docs/inception-v3-advanced.
26. P. Tao, Z. Sun, and Z. Sun, (2018). "An Improved Intrusion Detection Algorithm Based on GA and SVM," *IEEE Access*, vol. 6, pp. 13624–13631. doi: 10.1109/ACCESS.2018.2810198.
27. A. Astorino, E. Gorgone, M. Gaudioso, and D. Pallaschke, (2011). "Data preprocessing in semi-supervised SVM classification," *Optimization*, vol. 60, no. 1–2, pp. 143–151. doi: 10.1080/02331931003692557.
28. E. Carrizosa and D. Romero Morales, (2013). "Supervised classification and mathematical optimization," *Computers & Operations Research*, vol. 40, no. 1, pp. 150–165. doi: 10.1016/j.cor.2012.05.015.
29. F. Nie, W. Zhu, and X. Li, (2020). "Decision Tree SVM: An extension of linear SVM for non-linear classification," *Neurocomputing*, vol. 401, pp. 153–159. doi: 10.1016/j.neucom.2019.10.051.
30. T. Sathiya, R. Reenadevi, and B. Sathiyabhama, (2021). "Random Forest Classifier based detection of Parkinson's disease," vol. 25, no. 5, p. 8.
31. S. Thaseen and C. A. Kumar,(2013). "An analysis of supervised tree based classifiers for intrusion detection system," in *2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering*, Salem, 294–299, doi: 10.1109/ICPRIME.2013.6496489.
32. N. Farnaaz and M. A. Jabbar, "Random Forest Modeling for Network Intrusion Detection System," *Procedia Computer Science*, vol. 89, pp. 213–217, 2016, doi: 10.1016/j.procs.2016.06.047.
33. A. D. Kulkarni and B. Lowe,(2020). "Random Forest Algorithm for Land Cover Classification," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 4, no. 3, p. 7.
34. B. J. Samajpati and S. D. Degadwala,(2016). "Hybrid approach for apple fruit diseases detection and classification using random forest classifier," in *2016 International Conference on Communication and Signal Processing (ICCSP)*, Melmaruvathur, Tamilnadu, India, pp. 1015–1019, doi: 10.1109/ICCSP.2016.7754302.
35. Y. Benhammou, S. Tabik, B. Achchab, and F. Herrera, (2018). "A first study exploring the performance of the state-of-the art CNN model in the problem of breast cancer," in *Proceedings of the International Conference on Learning and Optimization Algorithms: Theory and Applications - LOPAL '18*, Rabat, Morocco, pp. 1–6, doi: 10.1145/3230905.3230940.
36. R. Rouhi, M. Jafari, S. Kasaei, and P. Keshavarzian, (2015). "Benign and malignant breast tumors classification based on region growing and CNN segmentation," *Expert Systems with Applications*, vol. 42, no. 3, pp. 990–1002, doi: 10.1016/j.eswa.2014.09.020.
37. S. Charan, M. J. Khan, and K. Khurshid, (2018). "Breast Cancer Detection in Mammograms using Convolutional Neural Network," p. 5.
38. R. Rasti, M. Teshnehlab, and S. L. Phung, (2017). "Breast cancer diagnosis in DCE-MRI using mixture ensemble of convolutional neural networks," *Pattern Recognition*, vol. 72, pp. 381–390. doi: 10.1016/j.patcog.2017.08.004.
39. A. Azamimi Abdullah, A. F. Dickson Giong, and N. A. Hanin Zahri,(2019). "Cervical cancer detection method using an improved cellular neural network (CNN) algorithm," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 14, no. 1, p. 210. doi: 10.11591/ijeecs.v14.i1.pp210-218.
40. S. A. Agnes, J. Anitha, S. I. A. Pandian, and J. D. Peter, (2020). "Classification of Mammogram Images Using Multiscale all Convolutional Neural Network (MA-CNN)," *Journal of Medical Systems*, vol. 44, no. 1, p. 30. doi: 10.1007/s10916-019-1494-z.
41. P. Matlani and M. Shrivastava, (2019). "Hybrid Deep VGG-NET Convolutional Classifier for Video Smoke Detection," *Computer Modeling in Engineering & Sciences*, vol. 119, no. 3, pp. 427–458. doi: 10.32604/cmes.2019.04985.
42. F. Pedregosa *et al.*, (2011)."Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830.
43. D. Bardou, K. Zhang, and S. M. Ahmad, (2018). "Classification of Breast Cancer Based on Histology Images Using Convolutional Neural Networks," *IEEE Access*, vol. 6, pp. 24680–24693. doi: 10.1109/ACCESS.2018.2831280.
44. Y. Yari, T. V. Nguyen, and H. T. Nguyen, (2020)."Deep Learning Applied for Histological Diagnosis of Breast Cancer," *IEEE Access*, vol. 8, pp. 162432–162448, doi: 10.1109/ACCESS.2020.3021557.
45. X. Y. Liew, N. Hameed, and J. Clos, (2021). "An investigation of XGBoost-based algorithm for breast cancer classification," *Machine Learning with Applications*, vol. 6, p. 100154. doi: 10.1016/j.mlwa.2021.100154.
46. R. Karthiga and K. Narashimhan,(2021). "Deep Convolutional Neural Network for Computer-Aided Detection of Breast Cancer Using Histopathology Images," *Journal of Physics: Conference Series*, vol. 1767, no. 1, p. 01204, doi: 10.1088/1742-6596/1767/1/012042.
47. S. I. Khan, A. Shahrior, R. Karim, M. Hasan, and A. Rahman, (2021). "MultiNet: A deep neural network approach for detecting breast cancer through multi-scale feature fusion," *Journal of King Saud University - Computer and Information Sciences*, p. S1319157821002068. doi: 10.1016/j.jksuci.2021.08.004.

CITATION OF THIS ARTICLE

N Behar and M Shrivastava. An Automated Histopathology Image Classification using a Hybrid Model. Bull. Env.Pharmacol. Life Sci., Spl Issue [1] 2022 : 1580-1590