Breast Cancer Image Multi-Classification Using Random Patch Aggregation and Depth-Wise Convolution based Deep-Net Model

Vandana Kate ^(✉),Pragya Shukla
IET DAVV, Indore, India
vandana.kate@gmail.com

Abstract—Adapting the profound, deep convolutional neural network models for large image classification can result in the layout of network architectures with a large number of learnable parameters and tuning of those varied parameters can considerably grow the complexity of the model. To address this problem a convolutional Deep-Net Model based on the extraction of random patches and enforcing depth-wise convolutions is proposed for training and classification of widely known benchmark Breast Cancer histopathology images. The classification result of these randomly extracted patches (size 50X50X3) is aggregated using majority vote casting in deciding the final image classification type. It has been observed that the proposed Deep-Net model implementation results when compared with classification results of the VGG Net (16 layers) learned features, outclasses achieving accuracy up to 89.6% on multi-class classification for 40X magnified images. The results further indicate model trained for images of one optical magnification factor (eg. 40X) might not classify images captured on different magnification (like 100X, 200X, and 400X) with similar accuracy. Thus, different classifiers are required at different magnifications.

Keywords—Image-Multi-Classification, Histopathology-Images, Breast-Cancer (BC), Feature-Extraction, Deep-Neural-Networks

1   Introduction

A rapid increase has been observed in the occurrences of breast cancer, especially in Asian nations like China, India, and Malaysia [1][2]. Recognizing breast cancer is yet AI's most philanthropic and intricate challenge. The vital symptom of breast cancer is normally a lump or a tumor that feels peculiar from the rest of the breast tissue. However, it is not always easy to distinguish a malignant tumor from a benign one because of their structural similarities. To effectively understand the structural versions, physicians should carefully look at an individual's medical records and make various clinical examinations such as mammography or ultrasound. However, a precise and accurate analysis of a breast tumor can be acquired through few forms of biopsy wherein a small pattern of cells or tissue is removed and stained (using Hema-

toxylin and Eosin stains) for examination [3]. The process of monitoring these images is a task of fatigue and requires expertise in the field [4]. There may be the possibility of missing important indicators when perceptibly searching for signs of cancer and may return a false negative.

Light microscopes are used to examine fine details and enlarge images of small matter. Enlarging an image is known as magnification and the amount of refined specification that can be visible is called resolution. H & E-stained tissue images are often captured at different optical magnification degrees, where each magnification can constitute specific facts. The lowest magnification captures the larger vicinity of tissue, while other larger magnification captures the zoomed-in view of the tissue. This explains the use of different magnifications which can doubtlessly yield varying discriminated statistics.

Deep learning methods have been exceedingly used to extract the applicable facts from the raw images and use them for classification tasks [5][6][7]. There are lots of contributions that had been proposed to enhance generalization capacity for various heavily used benchmark Cancer datasets by making use of deep convolutional networks. Most of the images concentrate on using complex model architecture consisting of Alex Net [8] to VGG-16 [9], ResNet [10], Inception-V3 [11], and DenseNet [12] with functional manipulations such as dropout regularization, batch normalization, transfer learning [13], and zero-shot training [14] have been evolved to try and expand Deep Learning on image data set.

The Table 1 summarizes the work done in literature on Breast Cancer images. It compiles research accomplished with Unsupervised Methods, Supervised Methods, Ensemble techniques, and Deep Methods. Table 2 summarizes the methods employed for Binary and Multi-class classification using Histopathology Breast Cancer Image analysis.

Though Deep convolutional neural network has confirmed its effectiveness for numerous image classification responsibilities, gaining comprehensive details present in biomedical images is difficult and the task adds on another difficulty when we stumble upon images of various magnification levels. Learning Deep-Net works with huge-size input dimensions requires a longer training time and a notably large network shape with more no. of hidden layers and hardware memory. To ease the task of training Deep-Net work with huge size images as input literature [15][16] represented each image with one randomly cropped patch, and labeled the patch with the same label as of original image. But this approach leads to ambiguity in training examples as one patch may not be the good representative of the entire image. To deal with this problem, we propose and evaluate a deep-NET structure using SeparableConv2D to enforce depth-wise convolutions where we represent an input image with a small set of patches or tiles cropped from it and associating each tile with the corresponding image label. We also compared the proposed system with a widely used Pre-trained CNN model used for global feature extraction and classification using hard and soft voting techniques.

Table 1. Summary of Machine Learning Methods employed in Histopathology Breast Cancer Image analysis

Year	Technique	Classifier	Research By
2009	Supervised Classification	Bayes classifier	Marugame et al. [17]
2013		Random Forests	Basavanhally et al. [18]
2014		MLP, SVM, Tree	Irshad et al. [19]
2015		LSSVM	Korkmaz and Poyraz [20]
2015		SVM and RF	Tashk et al. [21]
2016		Random Kitchen Sink	Beevi et al. [22]
2016		Random Forests	Balazsi et al. [23]
2017		Random Forests	Valkonen et al. [24]
2017		Cascade SVM	Peikari et al. [25]
2010	Un-Supervised Classification	K-means	Fatakdawala et al. [26]
2011		K-means	Roullier et al. [27]
2015		Thresholding	Sirinukunwattana et al. [28]
2015		Mean-shift clustering	Huang [29]
2013	Using Ensembles	SVM + Random Subspaces using Perceptron	Filipczuk et al. [30]
2013		Adaboost	Vink et al. [31]
2016		Bagging and Tree using sum and variance	Fern´andez-Carrobles et al. [32]
2019	Deep Learning	Stacked Sparse Autoencoder (SSAE) based DNN	Sawon et al. [33]
2017		CNN (Convolutional Neural Network)	Spanhol, Fabio et al. [34]
2013		Deep max-pooling CNN	Cireşan et al.[35]
2017		CNN	Araújo et al.[36]
2015		SSAE (Sparse Stacked AutoEncoders)	Xu et al.[37]

Table 2. Summary of Methods employed for Binary and Multiclass Classification using Histopathology Breast Cancer Image analysis

| Year	Technique	Classifier	Research By
2016	Binary Classification	Global feature extraction Methods (LBP, GLCM etc.)	Spanhol et al.[38]
2017		shallow features	Gupta et al. [39]
2017		Multiple feature vector (MFV) & transfer learning	Fabio et al. [40]
2018		Graph-manifold & BI-LSTM models	Pratiher et al.[41]
2016		convolution neural network (CNN) model with fusion rule (FR)	F. A. Spanhol et al. [42]
2017		ConvNet based fisher vector (CFV) & Gaussian mixture model (GMM)	Song et al., [43]
2017		Deep CNN	Wei et al.[44]
2019		Incremental boosting convolution networks	Vo, Duc et al. [45]
2017	Multi Class Classification	Class structured deep CNN (CSDCNN) model	B. Wei et al.[46]
2018		CNN based approach	D. Bardou et al.[47]
2019		DNN	Sawon et al. [48]

2 Dataset Used

Spanhol et al. [38] offered a public dataset (BreakHis) of 82 breast cancer sufferers with diagnoses of eight forms of breast tumors. A framework for the evaluation of the classification strategies for 8 forms of tumor was also likewise proposed. Baseline accuracy was obtained using six feature extractors and four classifiers. The dataset includes 7,909 BC images of four magnification factors. Pictures are 460 pixels in pinnacle and 700d pixels huge. Table 3. illustrates distribution of classes and sub-classes in benign, malignant breast cancer dataset at different magnification factor (40X, 100X, 200X, 400X).

Table 3. Distribution of Benign and Malignant classes, sub-classes in Breast Cancer (BreaKHis) Dataset @ different magnification factors

Classes	Sub-Classes	40X	100X	200X	400X	Total Images	Patient count
Benign	Adenosis	114	113	111	106	444	4
	Fibroadenoma	253	260	264	237	1014	10
	Tubular adenoma	109	121	108	115	453	3
	Phyllodes tumor	149	150	140	130	569	7
Malignant	Ductal carcinoma	864	903	896	788	3451	38
	Lobular carcinoma	156	170	163	137	626	5
	Mucinous carcinoma	205	222	196	169	792	9
	Papillary carcinoma	145	142	135	138	560	6
	Total	**1995**	**2081**	**2013**	**1820**	**7909**	**82**

3 Convolutional Neural Network (CNN)

CNN's are a significant tool for image classification, retrieval and detection tasks these days. There are four predominant operations in CNNs, such as Convolution, Non-Linearity, Pooling/Sub Sampling, FC layer and Classification as shown in Fig. 1. CNN generally require plenty of training information, and has the capability to address big, high-resolution images and remodeling them without losing significant characteristics.
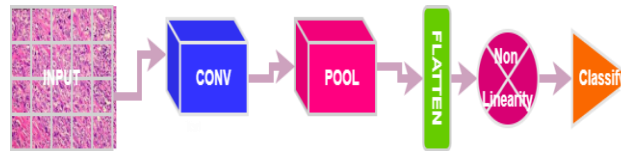

Fig. 1. Convolutional Neural Net

4 Experiment 1- Global Feature Extraction using Transfer Learning

Here an effort is made to construct accurate models for BC image classification problems through transfer learning in a quicker or better way. Transfer learning is usually used for preventing over-fitting problem considering the fact that ConvNet contains more generic features (e.g. horizontal /vertical edge detectors or color detectors) at initial layers, and later layers comprise more problem precise information of the multiple classes contained in the original data set. Typically, just the weights of trained convolutional layers are copied, rather than the entire network structure. This could be very effective as almost all data sets share low-level spatial traits that are better learned with big data sets.

A pre-trained CNN based model VGG-16 trained on an Imagenet benchmark dataset was imported deliberating the high computational cost of training such models.

VGG16 is chosen for its gratifying performances in image recognition tasks [9], usefulness toward real-time applications, and feasibility of transfer learning for constrained datasets. A typical CNN has two essential components: first is the Convolutional basis, which is composed of an assembly of convolutional and pooling layers. The main goal of the first part is to generate low-level features from the image. The second part consists of using the traditional machine learning approach, to categorize images from extracted features. We can compare and improve classification accuracy by training various classification models inclusive of linear Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and Multilayer Perceptron Model (MLP) classifiers on the features extracted by the Convolutional base part and using k-fold cross-validation to estimate the error of the classifier. For data augmentation usually, images are transformed using affine transforms (such as horizontal and vertical flipping, contrast enhancement, rotating, zooming, mirroring, fill mode= nearest, etc.) to avoid class imbalance. Fig. 2 describes the Global feature extraction method for BC image classification.

In addition, voting classifier strategies popularly called the weight-adjusted vote casting method is also used and compared. The idea of Ensemble Classifiers [49] is to combine the predictions made with the aid of more than one classifier, which can be more expressive than a single classifier prediction. Also, the outcomes obtained are much less dependent on the peculiarities of the training set. The varieties of voting used are hard and soft voting. The various classifiers used for voting were -Decision Tree Classifier, K Nearest Neighbor, Naive Bayes, Random Forest, Quadratic Discriminant Analysis, and AdaBoost.
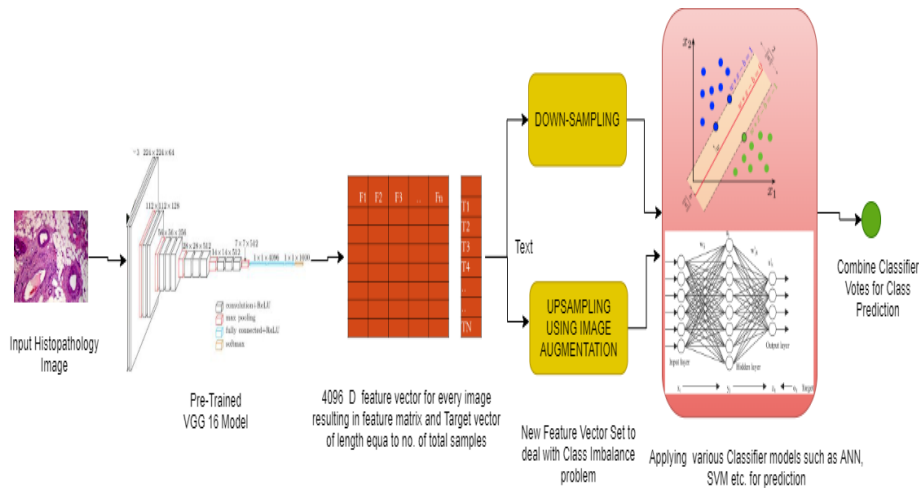
Fig. 2. Global Feature Extraction Model

4.1 Hard voting

In this very last prediction depends on the prediction that appears most even when making classification with the usage of specific models. It is likewise known as majority vote casting. Prediction $\hat{y}$ depends on prediction made thru unbiased classifiers $C_1(x), C_1(x) \ldots \ldots \ldots \ldots C_m(x)$.

$$\hat{y} = \text{mode}\{C_1(x), C_1(x) \ldots \ldots \ldots \ldots C_m(x)\} \tag{1}$$

The equation shows final output prediction is obtained by taking mode of different classifiers used

4.2 Soft voting

In this final prediction $\hat{y}$ relies upon on the predicted probability (p) of individual classifier.

$$\hat{y} = \underset{i}{\text{argmax}}\left\{\sum_{j=1}^{m} w_j\, p_{ij}\right\} \tag{2}$$

Where, p_{ij} is the ith probability of jth classifier, w_j is the weight assigned to j^{th} classifier and m is the total number of classifiers used.

The experimental results demonstrates that Average Accuracy using Soft Voting classifier was better than other conventional classifiers and Ensemble voting. The order of classifier accuracy was as follows Soft Voting > Hard Voting > MLP > SVC > KNN. Average Accuracy order for different magnification levels was as 40X (A) with 81.2% > 40X with 78.6% > 100X with 78% > 100X(A) with 75.6% > 200X with 75.6% > 200X(A) with 75% > 400X (A) with 72.2% > 400X with 69.8% , here A stands for Augmentation.

5 Proposed Deep-NET Model

Adapting the prevailing deep neural models for large images can result in extra complicated architectures, with massive units of parameters. Exceptional tuning these large ranges of parameters can appreciably grow the complexity of the model. We thus propose a Deep-Net Model to resolve these issues. We also use various parameter tuning and regularization methods adopted in literature for speeding Deep-Net works outputs such as-

5.1 Batch size and epochs

Batch size is the number of samples used from the learning set to estimate of the error gradient and is used to update the model weights. The more training examples we use the more accurately the weights will be adjusted towards improving the accuracy of the model. A batch size sample will be used to estimate the error gradient before the model weights are updated after completion of one epoch. Commonly used

learning algorithms are Batch Gradient Descent where batch size equals total training samples, Stochastic Gradient Descent where batch $size \, = \, = 1$, Mini-batch Gradient Descent. Batch < 1 and $<$ no. of samples in the data set.

5.2 Dropout

Dropout is a way where randomly selected neurons are overlooked for the duration of training by means of zeroing their activation values. Dropout is implemented to hidden layer neurons of the Deep-Net model. Dropout is used for regularization by using adding noise to the output function maps of every layer, adding robustness to the model in managing variations of the test image.

5.3 Batch normalization

Batch normalization is again a regularization technique that provides some noise to each hidden layer by normalizing the set of activations in the layer. Normalization works by subtracting the batch average value from each of the calculated activations and dividing by the batch standard deviation. This pre-processing technique allows each layer of a network to analyze/learn by way of itself and more independently of other layers.

To further speed up the task of training Deep-Net work with huge size images which is the major problem considered in this paper, the proposed Deep-Net Model is based on the extraction of random patches and enforcing separable convolutions for training. Further aggregate result of these patches is used for final image class recognition. Separable convolutions first perform a depth wise spatial convolution on each input channel separately followed by a point wise convolution that mixes all the channels as shown in Fig 3. Various steps implemented in sequence to build Deep-Net Model for 40X magnified images are explained as follows:

Step 1: Patch extraction: Here we extract multiple patches of $50X50$ from images of both the positive and negative classes. This dataset holds 16,176 patches of size 50X50 extracted from 2022 histopathology images of breast cancer specimens scanned at $40X$. Of these, $10,960$ patches test malignant and $5,216$ patches test benign. The dataset is available in public domain with the name Breakhis for breast cancer Dataset. In addition, a text file is created containing list of paths to image patches.

Step 2: Database creation: Using the patches we create a mean file, which is the mean value of a pixel as seen through all the patches found in learning database. This mean value is subtracted from the pixel to roughly "zero" the data, improving the efficiency of the Deep Learning algorithm. Then, using text file created in the previous step, we calculate an index by multiplying the length of this list in text file by 0.8 so we can slice this list to get sub-lists for the training and testing datasets. Next, 20% of remaining list is kept for testing. Now, datasets are a list with tuples (holding paths and the base path (training, validation or testing) with class label for each image) for information about the training, validation, and testing sets. We now actually build

training, validation and testing image set using above tuples in separately named folders. If the base path does not exist, we'll create the new directory.

Step 3: Build model: As a next step we suggest a deep-NET structure using Separable-Conv2D to enforce depth-wise convolutions. The purpose of doing convolution is to extract useful capabilities from the input. In Convolutional Neural Network, numerous features are extracted via convolution using filters whose weights are automatically learned throughout training. All those extracted features are then combined to make decisions. Convolution also takes spatial relationships of pixels into concern. Deep Learning have come up with different styles of convolutions (e.g. $2D, 3D, (1x1)$, Transposed, Dilated, Spatially Separable, Depth-wise Separable, Flattened, Grouped, Shuffled etc.). Here, in this work depthwise separable convolutions is used, commonly used in deep architectures (e.g. Mobile Net and Xception) which is applied to single channel at a time rather than applying to all channels at the same time, this requires lesser number of learnable parameters thus computations are reduced and resolves overfitting issues. The depth wise separable convolutions encompass steps involving depth-wise convolutions and $1x1$ convolutions.

For a histopathology input image of length $50\ X\ 50\ X\ 3$ (color images) we need to do 2D convolution (with stride=1, padding=0) with N kernels of size $h\ X\ h\ X\ 3$ in which h is even, which maps input layer $(50X50XD)$ into the output layer $(50 - h + 1)X(50 - h + 1)\ X\ N)$. The normal multiplications required are $NX\ (h\ X\ h\ X\ 3)\ X\ (50 - h + 1)\ X(50 - h + 1)$. On the contrary, for the identical transformation, the multiplication required for depth wise separable convolution is:

$$3\ X\ h\ X\ h\ X\ 1\ X\ (50 - h + 1)^2 + N\ X\ 1\ X\ 1\ X\ 3\ X\ (50 - h + 1)^2 = (h^2 + N)\ X\ 3\ X\ (50 - h + 1)^2 \tag{3}$$

The ratio of multiplications among depth wise separable convolution and 2D convolution is now

$$\frac{1}{N} + \frac{1}{h^2} \tag{4}$$

Assuming that output layer has many channels i.e $(N_C \gg h)$, the above expression reduces all the way down to

$$1/h/h \tag{5}$$

Asymptotically, if 3X3 filters are used, 2D convolutions spend nine times more in multiplications than depth wise separable convolutions. For 5 X 5 filters, 2D convolutions spend 25 times more time in doing multiplications.

We will build a classifier to train on 80% of a original breast cancer histopathology image dataset. Of this, we'll keep 21% of the data for validation/dev set, thus splitting our dataset into training, validation, and testing sets in the ratio - 59% (for training): 21% (for validation): 20% (for testing). Using Keras, we'll define a CNN (Convolutional Neural Network), calling it as BCNet, and train it on our images. We'll then derive metrics such as confusion matrix, accuracy, f score, recall and precision to analyze the performance of the model. We use ImageDataGenerator from Keras for

image data augmentation and extract batches of images to avoid overloading the entire dataset in memory at once.

The Deep-NET model performs the following operations:

Use $3X3$ convolutional _lters

Stack these filters on top of each other

Perform max-pooling

Use SeparableConv2D for depth-wise separable convolution

The class deep-NET has a static method build () that takes four parameters- width and height of the image, its depth (the number of color channels in each image), and the number of classes the network will predict between, which, for us, is 8 (0 and 7). In this method, we initialize model and shape. Here we used channels first, to update the shape and the channel dimension. Now, we'll define three $DEPTHWISE\ CONV =$ $> RELU => POOL$ layers; each with a higher stacking and a greater number of filters. The softmax classifier outputs prediction percentages for each class. In the end, we return the model.

Step 4: Train model: Here we train and evaluate our model. Here, we'll import from keras, sklearn, Deep-Net , config, imutils, matplotlib, numpy, and os. In this script, first, we set initial values for the number of epochs to 200-400, the learning rate $= 1e - 2$, and the batch size of 64. We'll get the number of paths in the three directories for training, validation, and testing. Then, we adjust class weights to deal with class imbalance problem. We will initialize the model using the Adagrad optimizer and compile it with a categorical cross-entropy loss function. The model is compiled and trained for the given dataset.

Step 5: Image labeling: After patch classification result is computed image-wise classification is obtained using majority voting, where the most happening patch label is selected as image label.

Step 6: Evaluation: In this section, we present an extensive experimental evaluation shown in Table 5 for the proposed model with architecture shown in Table 4 on the BreakHis dataset, in order to show case its better performance than Global feature extraction method using metrics such as Accuracy, F-Score, Precision and Recall. The steps were defined for $40X$ i.e we trained the model at $40X$ magnification and as it has the tendency of being maximum informative magnification factor, we test it with images of other stage magnifications i.e.,$100X,\ 200X, 400X$.

However, it would seem that one magnification model may not be able to handle images with other magnifications, and different classifiers are required at different magnifications. Moreover, decision in such cases where large variation in patient score exists, may not be reliable by just considering one magnification level. However, it would seem that one magnification model may not be able to handle images with other magnifications, and different classifiers are required at different magnifications. Moreover, a decision in such cases where large variation in patient score exists, may not be reliable by just considering one magnification level. However, it appears that one magnification trained model might not classify images with similar accuracy for different magnification images.
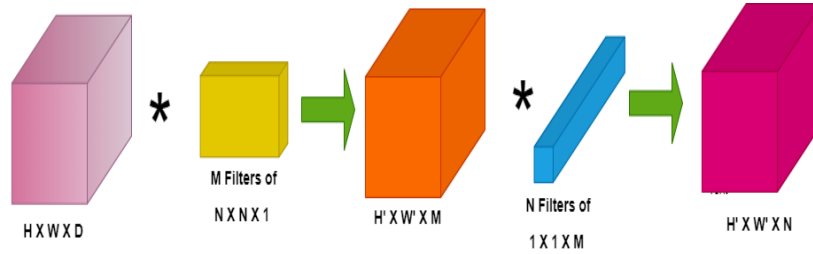
Fig. 3. Depth wise Separable Convolution

Table 4. Architecture of Sequential Deep-Net Model

Architecture of Sequential DeepNet Model		
{Layer Type}	*{Output Shape}*	*{Learnable Parameters}*
InputLayer	(None, 50, 50, 3)	0
Separable_Conv2D (padding='same')	(None, 48, 48, 32)	155
Activation('relu'), Batch_Normalization, MaxPooling2D((2,2), Dropout(0.25)	(None, 24, 24, 32)	128
Separable_Conv2D (padding='same')	(None, 24, 24, 64)	2400
Activation('relu'), Batch_Normalization,	(None, 24, 24, 64)	256
Separable_Conv2D (padding='same')	(None, 24, 24, 64)	4736
Activation('relu'), Batch_Normalization, MaxPooling2D((2,2), Dropout(0.25)	(None, 24, 24, 64)	256
Separable_Conv2D (padding='same')	(None, 12, 12, 128)	8896
Activation('relu'), Batch_Normalization,	(None, 12, 12, 128)	152
Separable_Conv2D (padding='same')	(None, 12, 12, 128)	17664
Activation('relu'), Batch_Normalization,	(None, 12, 12, 128)	152
Separable_Conv2D (padding='same')	(None, 12, 12, 128)	17664
Activation('relu'), Batch_Normalization, MaxPooling2D((2,2), Dropout(0.25)	(None, 12, 12, 128)	512
Flatten	(None, 4608)	0
Dense Layer	(None, 256)	
Activation('relu'), Batch_Normalization, Dropout(0.25)	(None, 256)	1179904
Dense Layer	(None, 8)	1024
Activation('Softmax')	(None, 8)	
Total Trainable Parameters		**{1,233,533}**

Table 5. Results of Proposed DeepNet Model using metrics such as F-score (F-S), Recall (R) and Precision (P) for all the 8 sub-classes (A, F, TA, PT, DC, LC, MC, PC) of all magnification levels (40X, 100X, 200X, 400X)

Metrics used ->	F-S	R	P	Acc	F-S	R	P	Acc
Sub-Classes\Magnification	*40X*				*100X*			
Adenosis (A)	0.8	0.84	0.86		0.76	0.68	0.85	
Fibroadenoma (F)	0.9	0.88	0.92		0.9	0.95	0.87	
Phyllodes Tumor (PT)	0.79	0.76	0.93		0.77	0.68	0.88	
Tubular Adnoma	0.91	0.93	0.9	**86.89**	0.91	0.96	0.87	**88.69**
Ductal Carcinoma (DC)	0.76	0.85	0.78		0.7	0.68	0.73	
Lobular Carcinoma (LC)	0.87	0.82	0.93		0.88	0.89	0.86	
Mucinous Carcino-ma(MC)	0.77	0.72	0.93		0.72	0.65	0.81	
Papillary Carcinoma (PC)	0.9	0.93	0.97		0.89	0.93	0.85	
Metrics used ->	F-S	R	P	Acc	F-S	R	P	Acc
Sub-Classes\Magnification	200X				400X			
Adenosis (A)	0.66	0.65	0.88		0.63	0.55	0.73	
Fibroadenoma (F)	0.85	0.86	0.84		0.86	0.91	0.81	
Phyllodes Tumor (PT)	0.72	0.64	0.84		0.7	0.67	0.73	
Tubular Adnoma	0.9	0.94	0.85	86.53	0.87	0.89	0.86	86.37
Ductal Carcinoma (DC)	0.7	0.62	0.79		0.61	0.58	0.66	
Lobular Carcinoma (LC)	0.88	0.93	0.85		0.84	0.87	0.82	
Mucinous Carcinoma(MC)	0.66	0.59	0.75		0.53	0.44	0.66	
Papillary Carcinoma (PC)	0.86	0.91	0.82		0,83	0.89	0.77	

Table 6. Comparison of the proposed work with literature (results in bold demonstrates the best result achieved).The metric chosen for comparison is Accuracy

	Magnification Factor			
Methods	*40X*	*100X*	*200X*	*400X*
[38]	83.8 ± 4.1	82.1 ± 4.9	85.1 ± 3.1	82.3 ± 3.8
[42]	88.6±5.6	84.5 ± 2.4	85.3 ± 3.8	81.7 ± 4.9
[50]	83.0 ± 3.0	83.1 ± 3.5	84.6 ± 2.7	82.1 ± 4.4
This work	86.89 ± 4.0	88.69±3.9	86.53±4.2	86.37±4.3

6   Comparison with State-of-The-Art

To endorse the viability of proposed deep architecture for BC sub-classification we balance our results with some state of art work. In Table-7 we assemble the best outcomes got in this work along with other CNN-based approach presented in, [38], [42] and [50] applied on the same Breast Cancer dataset as used in this paper. Results depicts that the proposed strategy beats many of contemporary strategies except in some cases. The objective of this work is to examine and comprehensively analyze the sub-class classification performance of the proposed model across all optical magnification frontiers.

7   Conclusion

The work presented here proposed a Deep-Net Model based on the extraction of random patches and enforcing depth-wise convolutions which is an enhancement over traditional way of using Deep-Net models for image classification. An Exhaustive comparison of proposed model for 8 sub-class classification for Histopathology Breast Cancer images is presented over transfer learning based Global Feature Extraction method using Ensemble voting classifiers. Results show case better performance of proposed Deep-Net model over Global feature extraction method. Above result examination also reveals that one magnification model may not be able to handle images with other magnifications thus it raises the need for a magnification independent model utilizing deep learning to classify the benign and malignant cases. The study opens some questions about scale in variance properties of feature-classifier combination, role of ensemble classification, considering that the magnification specific model requires relatively less training than the magnification independent model. This research is a foundation for our future publication in the integration of deep learning and block-chain technology.

8   Declaration

We as author declare that there is no conflict of interest.

9   References

[1] Agarwal, Gaurav, P. V. Pradeep, Vivek Aggarwal, Cheng-Har Yip, and Polly SY Cheung." Spectrum of breast cancer in Asian women." World journal of surgery 31, no. 5 (2007): 1031-1040. https://doi.org/10.1007/s00268-005-0585-9

[2] Ameer SF, Nayyef ZT, Fahad ZH, ALRubee IR. "Using Morphological Operation and Watershed Techniques for Breast Cancer Detection". International Journal of Online and Biomedical Engineering (iJOE). 2020 May 14;16(05):140-9. https://doi.org/10.3991/ijoe.v16i05.12999

[3] De Matos, Jonathan, Alceu de Souza Britto Jr, Luiz ES Oliveira, and Alessan- dro L. Koerich." Histopathologic image processing: A review." arXiv preprint arXiv:1904.07900 (2019)

[4] Muljo HH, Perbangsa AS, Lie Y, Pardamean B. Improving early cancer detection knowledge through mobile learning application. International Journal of Online and Bio- medical Engineering (iJOE). 2019 Jan 31;15(02):60-70. https://doi.org/10.3991/ijoe.v15i02.9678

[5] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton." Deep learning." Nature 521.7553 (2015): 436-444. https://doi.org/10.1038/nature14539

[6] Bengio, Yoshua, Aaron Courville, and Pascal Vincent." Representation learning: A review and new perspectives." IEEE transactions on pattern analysis and machine intelligence 35, no. 8 (2013): 1798-1828. https://doi.org/10.1109/tpami.2013.50

[7] Muniasamy A, Alasiry A. Deep Learning: The Impact on Future eLearning. Interna- tional Journal of Emerging Technologies in Learning (iJET). 2020 Jan 15;15(01):188-99. https://doi.org/10.3991/ijet.v15i01.11435

[8] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton." Imagenet classification with deep convolutional neural networks." In Advances in neural information pro- cessing systems, pp. 1097-1105. 2012. https://doi.org/10.1145/3065386

[9] Simonyan, Karen, and Andrew Zisserman." Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).

[10] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun." Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. 2016. https://doi.org/10.1109/cvpr.2016.90

[11] Szegedy, Christian, Vincent Vanhoucke, Serge Ioffe, Jon Shlens, and Zbigniew Wojna." Rethinking the inception architecture for computer vision." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2818-2826. 2016. https://doi.org/10.1109/cvpr.2016.308

[12] Huang, Gao, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q. Weinberger." Densely connected convolutional networks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4700-4708. 2017. https://doi.org/10.1109/cvpr.2017.243

[13] Shao, Ling, Fan Zhu, and Xuelong Li." Transfer learning for visual categorization: A sur- vey." IEEE transactions on neural networks and learning systems 26, no. 5 (2014): 1019-1034. https://doi.org/10.1109/tnnls.2014.2330900

[14] Palatucci, Mark, Dean Pomerleau, Geoffrey E. Hinton, and Tom M. Mitchell." Zero-shot learning with semantic output codes." In Advances in neural information processing sys- tems, pp. 1410-1418. 2009.

[15] L. Kang, P. Ye, Y. Li, and D. Doermann. Convolutional neural networks for no- reference image quality assessment. In CVPR, pages 1733–1740, 2014 https://doi.org/10.1109/cvpr.2014.224

[16] X. Lu, Z. Lin, H. Jin, J. Yang, and J. Wang. RAPID: Rating pictorial aesthetics using deep learning. In ACM MM, pages 457–466, 2014 https://doi.org/10.1145/2647868.2654927

[17] A. Marugame, T. Kiyuna, M. Ogura, A. Saito, Categorization of HE stained breast tissue samples at low magnification by nuclear aggregations, in: IFMBE Proceedings, volume 25, pp. 173–176. https://doi.org/10.1007/978-3-642-03882-2_45

[18] A. Basavanhally, S. Ganesan, M. Feldman, N. Shih, C. Mies, J. Tomaszewski, A. Madabhushi, Multi-Field-of-View Framework for Distinguishing Tumor Grade in ER+

Breast Cancer From Entire Histopathology Slides, IEEE Transactions on Biomedical Engineering 60 (2013) 2089–2099.

[19] Irshad, Humayun, Alexandre Gouaillard, Ludovic Roux, and Daniel Racoceanu." Multispectral band selection and spatial characterization: Application to mito- sis detection in breast cancer histopathology." Computerized Medical Imaging and Graphics 38, no. 5 (2014): 390-402. https://doi.org/10.1016/j.compmedimag.2014.04.003

[20] Korkmaz, Sevcan Aytac, and Mustafa Poyraz." Least square support vector machine and minumum redundacy maximum relavance for diagnosis of breast cancer from breast microscopic images." Procedia-social and behavioral sciences 174 (2015):4026-4031. https://doi.org/10.1016/j.sbspro.2015.01.1150

[21] Tashk, Ashkan, Mohammad Sadegh Helfroush, Habibollah Danyali, and Mojgan Akbarzadeh-Jahromi." Automatic detection of breast cancer mitotic cells based on the combination of textural, statistical and innovative mathematical features." Applied Mathematical Modelling 39, no. 20 (2015): 6165-6182. https://doi.org/10.1016/j.apm.2015.01.051

[22] Sabeena, Madhu S. Nair, and G. R. Bindu." Detection of mitotic nuclei in breast histopathology images using localized ACM and Random Kitchen Sink based classifier." In 2016 38th annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 2435-2439. IEEE, 2016. https://doi.org/10.1109/embc.2016.7591222

[23] Balazsi, Matthew, Paula Blanco, Pablo Zoroquiain, Martin D. Levine, and Miguel N. Burnier." Invasive ductal breast carcinoma detector that is robust to image magnification in whole digital slides." Journal of Medical Imaging 3, no. 2 (2016): 027501. https://doi.org/10.1117/1.jmi.3.2.027501

[24] Valkonen, Mira, Kimmo Kartasalo, Kaisa Liimatainen, Matti Nykter, Leena La-tonen, and Pekka Ruusuvuori." Metastasis detection from whole slide images using local features and random forests." Cytometry Part A 91, no. 6 (2017): 555-565. https://doi.org/10.1002/cyto.a.23089

[25] Peikari, Mohammad, Sherine Salama, Sharon Nofech-Mozes, and Anne L. Martel." Automatic cellularity assessment from post-treated breast surgical specimens." Cy- tometry Part A 91, no. 11 (2017): 1078-1087. https://doi.org/10.1002/cyto.a.23244

[26] Fatakdawala, Hussain, Jun Xu, Ajay Basavanhally, Gyan Bhanot, Shridar Ganesan, Michael Feldman, John E. Tomaszewski, and Anant Madabhushi." Expectation–maximization-driven geodesic active contour with overlap resolution (emagacor): Application to lymphocyte segmentation on breast cancer histopathology." IEEE Transactions on Biomedical Engineering 57, no. 7 (2010): 1676-1689. https://doi.org/10.1109/tbme.2010.2041232

[27] Roullier, Vincent, Olivier L´ezoray, Vinh-Thong Ta, and Abderrahim Elmoataz." Multiresolution graph-based analysis of histopathological whole slide images: Ap- plication to mitotic cell extraction and visualization." Computerized Medical Imag- ing and Graphics 35, no. 7-8 (2011): 603-615. https://doi.org/10.1016/j.compmedimag.2011.02.005

[28] Sirinukunwattana, Korsuk, Adnan M. Khan, and Nasir M. Rajpoot." Cell words: Modelling the visual appearance of cells in histopathology images." Computerized Medical Imaging and Graphics 42 (2015): 16-24. https://doi.org/10.1016/j.compmedimag.2014.11.008

[29] Huang, Chao-Hui." Semi-supervised color decomposition for histopathological images using exclusive component analysis." In 2015 IEEE 25th International Work- shop on Ma-

chine Learning for Signal Processing (MLSP), pp.1-6. IEEE, 2015. https://doi.org/10.1109/mlsp.2015.7324359

[30] Filipczuk, Pawe l, Bartosz Krawczyk, and Micha Wo´zniak." Classifier ensemble for an effective cytological image analysis." Pattern Recognition Letters 34, no. 14 (2013): 1748-1757. https://doi.org/10.1016/j.patrec.2013.04.030

[31] Vink, Jelte Peter, M. B. Van Leeuwen, C. H. M. Van Deurzen, and Gerald de Haan." Efficient nucleus detector in histopathology images." Journal of microscopy 249, no. 2 (2013): 124-135. https://doi.org/10.1111/jmi.12001

[32] Fernandez-Carrobles, M. Milagro, Ismael Serrano, Gloria Bueno, and Oscar D´eniz." Bagging tree classifier and texture features for tumor identification in histological images." Procedia Computer Science 90 (2016): 99-106. https://doi.org/10.1016/j.procs.2016.07.030

[33] Pratiher, Sawon, and Subhankar Chattoraj." Diving Deep onto Discriminative Ensemble of Histological Hashing Class-Specific Manifold Learning for Multi-class Breast Carcinoma Taxonomy". In ICASSP 2019-2019 IEEE International Confer- ence on Acoustics, Speech and Signal Processing (ICASSP), pp. 1025-1029.IEEE,2019.https://doi.org/10.1109/icassp.2019.8683856

[34] Spanhol, Fabio A., Luiz S. Oliveira, Paulo R. Cavalin, Caroline Petitjean, and Laurent Heutte." Deep features for breast cancer histopathological image classification." In 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 1868-1873. IEEE, 2017. https://doi.org/10.1109/smc.2017.8122889

[35] Cire¸san, Dan C., Alessandro Giusti, Luca M. Gambardella, and Ju¨rgen Schmid- huber." Mitosis detection in breast cancer histology images with deep neural networks." In International conference on medical image computing and computer- assisted intervention, pp. 411-418. Springer, Berlin, Heidelberg, 2013. https://doi.org/10.1007/978-3-642-40763-5_51

[36] Arau´jo, Teresa, Guilherme Aresta, Eduardo Castro, Jos´e Rouco, Paulo Aguiar, Catarina Eloy, Antonio Polonia, and Aur´elio Campilho." Classification of breast cancer histology images using convolutional neural networks." PloS one 12, no. 6 (2017). https://doi.org/10.1371/journal.pone.0177544

[37] Xu, Jun, Lei Xiang, Qingshan Liu, Hannah Gilmore, Jianzhong Wu, Jinghai Tang, and Anant Madabhushi." Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images." IEEE transactions on medical imaging 35, no. 1 (2015): 119-130. https://doi.org/10.1109/tmi.2015.2458702

[38] Spanhol, Fabio A., Luiz S. Oliveira, Caroline Petitjean, and Laurent Heutte." A dataset for breast cancer histopathological image classification." IEEE Transactions on Biomedical Engineering 63, no. 7 (2015): 1455-1462. https://doi.org/10.1109/tbme.2015.2496264

[39] Gupta, Vibha, and Arnav Bhavsar." Breast cancer histopathological image classification: is magnification important?" In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 17-24. 2017https://doi.org/10.1109/cvprw.2017.107

[40] Spanhol, Fabio A., Luiz S. Oliveira, Paulo R. Cavalin, Caroline Petitjean, and Laurent Heutte." Deep features for breast cancer histopathological image classification." In 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 1868-1873. IEEE, 2017. https://doi.org/10.1109/smc.2017.8122889

[41] S. Pratiher, S. Chattoraj, S. Agarwal and S. Bhattacharya," Grading Tumor Malig- nancy via Deep Bidirectional LSTM on Graph Manifold Encoded Histopathological Image,"

2018 IEEE International Conference on Data Mining Workshops (ICDMW), Singapore, Singapore, 2018, pp. 674-681. https://doi.org/10.1109/icdmw.2018.00104

[42] F. A. Spanhol, L. S. Oliveira, C. Petitjean and L. Heutte," Breast cancer histopathological image classification using Convolutional Neural Networks," 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, 2016, pp. 2560-2567. https://doi.org/10.1109/ijcnn.2016.7727519

[43] Song, Yang, Ju Jia Zou, Hang Chang, and Weidong Cai." Adapting fisher vectors for histopathology image classification." In 2017 IEEE 14th International Sympo- sium on Biomedical Imaging (ISBI 2017), pp. 600-603. IEEE, 2017. https://doi.org/10.1109/isbi.2017.7950592

[44] Han, Zhongyi, Benzheng Wei, Yuanjie Zheng, Yilong Yin, Kejian Li, and Shuo Li." Breast cancer multi-classification from histopathological images with structured deep learning model." Scientific reports 7, no. 1 (2017): 1-10. https://doi.org/10.1038/s41598-017-04075-z

[45] Vo, Duc My, Ngoc-Quang Nguyen, and Sang-Woong Lee." Classification of breast cancer histolog images using incremental boosting convolution networks." Infor- mation Sciences 482 (2019): 123-138. https://doi.org/10.1016/j.ins.2018.12.089

[46] Han, Zhongyi Benzheng Wei, Yuanjie Zheng, Yilong Yin, Kejian Li, and Shuo Li." Breast cancer multi-classification from histopathological images with structured deep learning model." Scientific reports 7, no. 1 (2017): 1-10. https://doi.org/10.1038/s41598-017-04075-z

[47] Bardou, Dalal, Kun Zhang, and Sayed Mohammad Ahmad." Classification of breast cancer based on histology images using convolutional neural networks." IEEE Access 6 (2018): 24680-24693https://doi.org/10.1109/access.2018.2831280

[48] Pratiher, Sawon, and Subhankar Chattoraj." Diving Deep onto Discriminative Ensemble of Histological Hashing Class-Specific Manifold Learning for Multi-class Breast Carcinoma Taxonomy." In ICASSP 2019-2019 IEEE International Confer- ence on Acoustics, Speech and Signal Processing (ICASSP), pp. 1025-1029. IEEE, 2019https://doi.org/10.1109/icassp.2019.8683856

[49] Opitz, David, and Richard Maclin." Popular ensemble methods: An empirical study." Journal of artificial intelligence research 11 (1999): 169-198. https://doi.org/10.1613/jair.614

[50] Bayramoglu N, Kannala J, Heikkil J. Deep Learning for magnification independent breast cancer histopathology image classification. In2016 23rd International conference on pattern recognition (ICPR) 2016 Dec 4 (pp. 2440-2445). IEEE. https://doi.org/10.1109/icpr.2016.7900002

10   Authors

Prof. Vandana Kate received the BE degree in Computer Engineering from Jabalpur Engineering College in 2002, ME degree in Computer Engineering from Institute of Engineering & Technology affiliated to DAVV University, Indore in 2010 and Currently pursuing PhD in Computer Engineering from the same university. She is working as Associate Professor in Acropolis Institute of Technology and Research and has a vast experience in teaching of more than 15 years. Her research interests include pattern recognition, machine learning, and medical image analysis and computer vision. She has many publications in leading international journals and peer

reviewed conferences, in various fields of Computer Science. Email: vandana.kate@gmail.com

Dr. Prof. Pragya Shukla received ME, PhD degree in Computer Engineering from Institute of Engineering and Technology affiliated to DAVV University, Indore. She has guided many BE, ME student projects and managing many PhD research scholars working in the field of Artificial Intelligence, Image Processing and AI. She is at present holding position of Professor in IET, DAVV, with huge teaching experience and research understanding. She has more than 25 publications in prestigious scientific international journals and peer-reviewed conferences covering various fields of Computer Science. She is also the PhD guide of Prof. Vandana Kate who is working in the field Breast Cancer cell processing using soft computing techniques.