B.Tech. Project

End-semester Report

on

Classification of breast cancer histology images

Submitted by

Aniket Raj
(201452027)

&

Rajeev Kumar Singh
(201452062)

under the supervision of

Dr. Ashish Phophalia

Assistant Professor

IIIT-Vadodara

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY VADODARA

2018

Declaration

We, Aniket Raj and Rajeev Singh, declare that this written submission represents our ideas in our own words and where others ideas or words have been included, We have adequately cited and referred the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated any idea /data/fact/ source in our submission. We fully understand that any violation of the above will cause for disciplinary action by the institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained.

Aniket Raj

201452027

Rajeev Kumar Singh

201452062

Date:27-04-2018

Abstract

This report describes the attempt made to classify the hematoxylin and eosin stained breast images into different classes. This work is inspired by ICIAR 2018 grand challenge on breast cancer histology images which is held as part of the ICIAR 2018 conference [6]. Breast cancer is becoming one of common cancer in women [2]. A method of feature extraction using pretrained model combined with some hand-picked features is used to classify the images into four classes. Further, an attempt to classify images into three classes has been made where the noncancerous classes are merged together as the cancerous images are of more significance to the pathologists. Three class classification is not part of ICIAR 2018 challenge. Machine-aided classification would reduce the cost significantly as manual inspection requires well-trained personnel. The number of features is kept small to reduce the computation time. The features extracted are used to train an SVM classifier with the linear kernel. Accuracies of 65% for four class and 74% for three class are achieved.

Contents

Chapter 1

Introduction

Breast cancer is one of top cancer in women worldwide [2]. In most of the cases, it causes death due to delay in diagnosis. So, it is necessary to find out the problem in the early stage in order to save lives. The real cause of breast cancer is still unknown but various risk factors are identified [3]. There have been cases of breast cancer in male too but it is comparatively lower [3]. Breast cancer in men is fatal because they do not generally notice it which results in a delay in diagnosis and by that time, cancer has already spread. Detecting breast cancer manually by visualizing image is difficult and the average diagnostic concordance or the agreement among experts is approx 75% [13]. Clinically, breast cancer can be detected by regular checkups using ultrasonic imaging followed by breast tissue biopsy and if there is any indication of the irregular growth of tissue then these tissues are accessed at the microscopic level. These images can be broadly classified into carcinoma and non-carcinoma categories. These categories can further be divided into subcategories where non-carcinoma is divided into normal and benign and carcinoma is divided into in-situ carcinoma and invasive carcinoma. The histology allows distinguishing between normal tissue, non-malignant (benign) and malignant lesions. Benign lesions represent changes in normal structures of the breast that are not directly related to progression to malignancy. In in-situ carcinoma, the cells are restrained inside the mammary ductal-lobular system, whereas in invasive carcinoma the cells spread beyond that structure. The work discussed in this report is inspired by ICIAR 2018 challenge which consists in automatically classifying HE stained breast histology microscopy images in four classes [6]. So, an attempt to classify the images into four classes is discussed.

In addition to this, the images are also categorized into three classes by combining the Normal and Benign class as one class because they are non-cancerous. The samples of in situ carcinoma and invasive carcinoma is of prime importance for the pathologists as they are cancerous. The classification in three categories is not part of the task in ICIAR 2018 grand challenge.

In the approach discussed in the report, a pretrained model is used to extract feature which is combined with some hand-picked features to train a support vector machine classifier. The number of features is kept small to avoid overfitting and reduce the computation time.

Chapter 2

Literature Survey

Computer aided diagnosis systems increases the efficiency and reduces workload of experts. Various attempts have been made in order to build such systems. Teresa Araujo et al.[1] used convolutional neural network (CNN) to build such classifier. In their work, 32 x 32 and 64 x 64 pixels patches were extracted from the initial images and used for training the CNN. The final classification was obtained by combining the patch probabilities with sum, product or maximum rules. The extraction of patches allowed to reduce the complexity of the model by decreasing the size of the input in subsequent layers. Image-wise classification achieve an overall accuracy of 77.8% for four classes. M.A.Aswathy et al. [7] has reviewed and summarized different digital image processing techniques on histopathological images for the detection of breast cancer in their article. In their article, there is critical analysis of different algorithms like Active Contour Model with General Classifier Neural Network (GCNN) by A.Jain and others [4]. They used active contours which was based on both region based and boundary based functions. This algorithm has an accuracy of 83.47% but at the same time it fails for the images containing high noise. It is also unable to detect overlapping cells but at the same time it is efficient in detecting shapes occuring at irregular intervals. For this, they used some predefined contours which were defined inside as well as outside the contour region using some suitable function. Some authors has worked on 3-classes classification of breast cancer histology images. Brook et al. [9] and Zhang et al. [10] classified breast cancer images in normal, in situ carcinoma and invasive carcinoma. Using multiple image threshold Brook et al. [9] binarized each image and used statistics of connected component to train

support vector machine (SVM). The classification into three classes instead of four classes would increase the overall accuracy of the model without hurting the detection of cancerous images. So, in addition to classifying image in four class, this report also discusses the accuracy score when they are classified into three class.

Chapter 3

Materials and methods

3.1 Dataset

The dataset has been taken from ICIAR 2018 Grand challenge on Breast Cancer Histology images [6]. It is organised as a part of the ICIAR 2018 conference. The dataset is composed of Hematoxylin and eosin stained breast histology microscopy images. The annotation was performed by two medical experts and images where there was disagreement were discarded. Figure 3.1 to 3.4 are sample of images from the dataset. The dataset contains a total of 400 microscopy images, distributed as follows:

- Normal: 100
- Benign: 100
- In situ carcinoma: 100
- Invasive carcinoma: 100

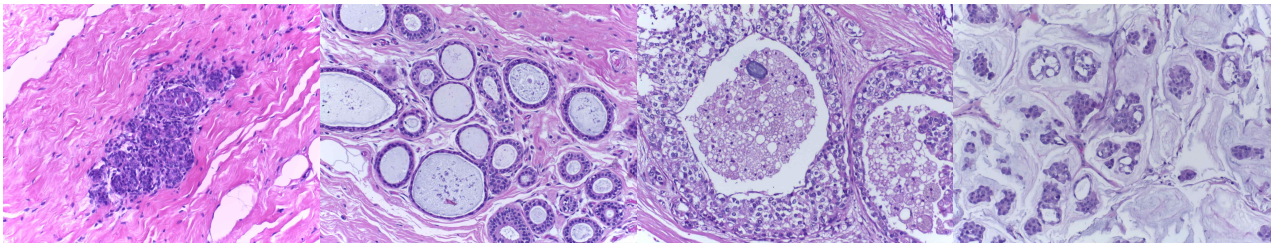

Figure 3.1: Normal Figure 3.2: Benign Figure 3.3: In situ Figure 3.4: Invasive

Microscopy images are on .tif format and have the following specifications:

- Color model: RGB
- Size: 2048 x 1536 pixels
- Pixel scale: 0.42 m x 0.42 μm
- Memory space: 10-20 MB (approx.) per image
- Type of label: image-wise

Different approaches have been tried to extract features from the data. First of all, different hand engineered featured which can be seen on manual inspection has been tried to be extracted. In addition to hand engineered features, pretrained model is used to extract features and the accuracy score is calculated. Extraction of features is followed by the selection of features where redundant or less significant features are removed. The selected features are used to train a support vector classifier.

3.2 Feature extraction

3.2.1 Hand engineered methods

Active Contour Count

Active contour model is used for finding the boundaries of an object [5]. Active contour model is also known as snakes works on the principle of minimizing energy of snakes. Energy function of snake [5]:

$$E^*_{snake} = \int_0^1 E_{snake}(\mathbf{v}(s))\, ds = \int_0^1 (E_{internal}(\mathbf{v}(s)) + E_{image}(\mathbf{v}(s)) + E_{con}(\mathbf{v}(s)))\, ds \qquad (3.2.1)$$

This model is generally used in application like shape recognition, segmentation, object tracking, edge detection etc. Figure 3.6 shows an image with contours. There are 1600 contours in this image. The number of contours in each image has been counted in the attempt to classify images. The accuracy score of this feature is only 20% and so has less significance.

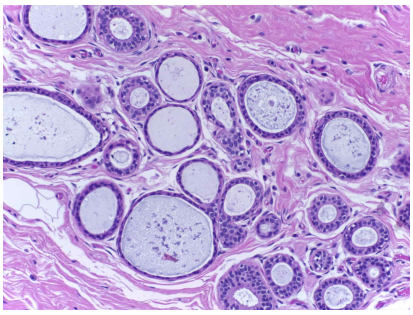

Figure 3.5: Original image

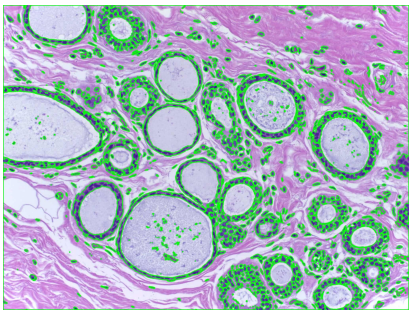

Figure 3.6: active contours

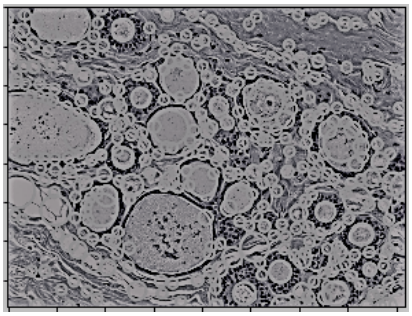

Figure 3.7: Blobs(LoG)

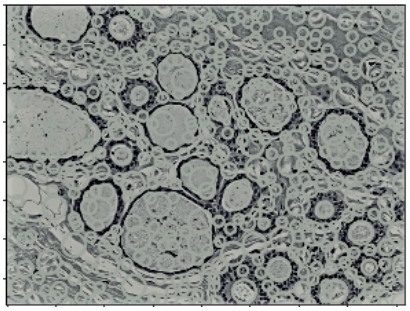

Figure 3.8: Blobs(DoG)

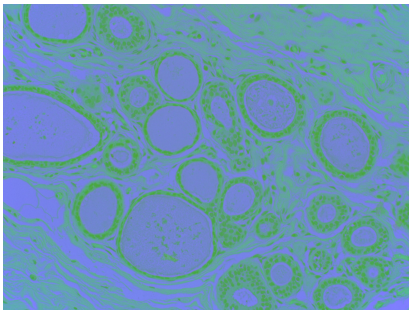

Figure 3.9: LAB colour space

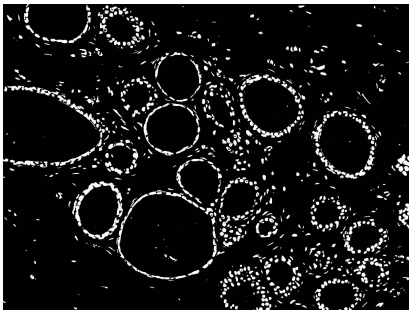

Figure 3.10: thresholding

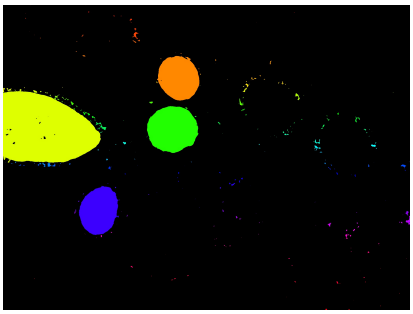

Figure 3.11: connected components

Suppose the test data is evenly distributed into four classes then even if the model predicts the same label every time and doesn't learn anything then the accuracy would be 25% and so the threshold for selection of any hand-picked feature in this report would be 25%.

Connected Components

Connected Components algorithm scans the image and groups its pixel into components based on its pixel connectivity. All those pixels who share same pixel density value are some way connected in same groups. After grouping these pixel, each group get coloured with different colour. Figure 3.11 shows an image with labelling of connected components. It is visible from

manual inspection that it is not possible to extract feature in order to classify image so, as of now, it has not been used to train the model.

Blob detection

Blob detection is a method that detects regions in digital image based on the difference in properties like colour or brightness. Blobs appears dark on bright region and bright on dark region of an image [8]. Blob detection can be done using different approaches like LoG , DoG etc.

LoG(Laplacian of Guassian): In this approach laplacian of gaussian images is calculated with increasing standard deviation in each step. LoG as an operator or convolution kernel defined as [15] :

$$LoG \triangleq \triangle G_\sigma(x,y) = \frac{\partial^2}{\partial^2 x} G_\sigma(x,y) + \frac{\partial^2}{\partial^2 y} G_\sigma(x,y) = \frac{x^2+y^2-2\sigma^2}{\sigma^4} e^{-(x^2+y^2)/2\sigma^2} \tag{3.2.2}$$

Figure 3.7 shows an image with blobs drawn.The image is in gray scale and the small circles represent the blobs detected. In Figure 4.3 the number of blobs are 351. The number of blobs in each image has been counted to train the model.

DoG(Difference of Gaussian): This is a faster approximation of LoG approach. In this case the image is blurred with increasing standard deviations and the difference between two successively blurred images are stacked up in a cube. DoG as an operator or convolution kernel defined as [15] :

$$DoG \triangleq G_{\sigma_1} - G_{\sigma_2} = \frac{1}{\sqrt{2\pi}} \left(\frac{1}{\sigma_1} e^{-(x^2+y^2)/2\sigma_1^2} - \frac{1}{\sigma_2} e^{-(x^2+y^2)/2\sigma_2^2} \right) \tag{3.2.3}$$

This method suffers from the disadvantage in detecting larger blobs. Blobs are again assumed to be bright on dark. Figure 3.8 shows an image with blobs drawn.The image is in gray scale and the small circles represent the blobs detected.The number of blobs in the image shown (Figure 3.8) are 446.

Image thresholding

Image thresholding method is used for image segmentation. In this method each pixel in image is replaced with black pixel if it's intensity is less than certain constant and with white pixel if the intensity of image is greater than constant. The number of components and area(in pixel) of each component in each image has been calculated. The number of components in Figure 3.10 are 1303.
In this approach, the image has been converted into Lab color space (Figure 3.9) and then thresholding (Figure 3.10) is done to better visualize the components. This thresholding operation can be expressed as:

$$\texttt{dst}(x,y) = \begin{cases} 0 & \texttt{if src(x,y)} > \texttt{thresh} \\ \texttt{maxval} & \texttt{otherwise} \end{cases}$$

The accuracy of this method is 0.21 for four class classification problem.

Area of contours

In this approach, the area of each contour in each image is calculated and given the label same as the label of the image. For example, the value of the area of all contours from class 1 i.e normal is labeled as normal. So, there would be thousands of data points of the normal category. Similarly for the class 2, 3 and class 4 images. The main idea behind trying this approach is to check if the images of the same category have similar contour area. The data points are used to train an SVM. Whenever a new test image comes, the area of all the contours of it is calculated and passed to the classifier which in turns return label for each

area. After that, majority voting for the returned label is done to assign the label to the image.

The idea doesn't work and the model gets biased towards the class having most data points i.e the label having the maximum number of contours. It can be concluded that the area of contours is not similar in images of the same class. Further, it is also evident from the manual inspection that contours of the similar area are present in every class and so it is difficult to classify on the basis of this feature.

Average area of contours

In this approach, the sum of the area of closed contours is calculated and then divided by the number of contours to get the average area. This process is repeated for every image which is then used as a feature for training the SVM classifier.

3.2.2 Transfer learning

Transfer learning can be defined as learning and gain knowledge from one task and apply that knowledge to some other task [14]. Suppose a neural network has learned to classify animals and then that knowledge or part of the knowledge can be used to classify X-ray images. This is termed as transfer learning. The weights of the network can be retrained using the new dataset where initial weights act as starting position. This method is very helpful when there is some model which has been trained on a lot of data and there are comparatively very few data for the new problem. It would not make any sense if the model being used has been trained on lesser data than what is available for the new problem. In the problem discussed in this report, since there is only hundred image per class which would be further split into training and test set, feature extraction has been done and not retraining and reassignment of weights has been done. In the course of extracting features using transfer learning for this problem, VGG16 and DesnseNet have been tried to verify if they both extract similar features or if the features can be combined to get better accuracy.

- **VGG16**

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Figure 3.12: ConvNet configuration from source paper [11]

The VGG [11] network was introduced by Simonyan and Zisserman in their paper, Very Deep Convolutional Networks for Large Scale Image Recognition. This team has also secured the second position in the classification track of ImageNet challenge 2014. This network follows a simple architecture of 3x3 convolutional layers on top of each other where volume size is managed by max pooling. It is then followed by two fully connected layers and a softmax classifier. In order to classify images of the given problem, a SVM with linear kernal has been trained on top of features extracted through this model. Figure 3.12 shows the configuration from source paper [11]. The block with blue boundary in the figure is the 16 weight layer configuration which is used in this report.

- **Densenet**

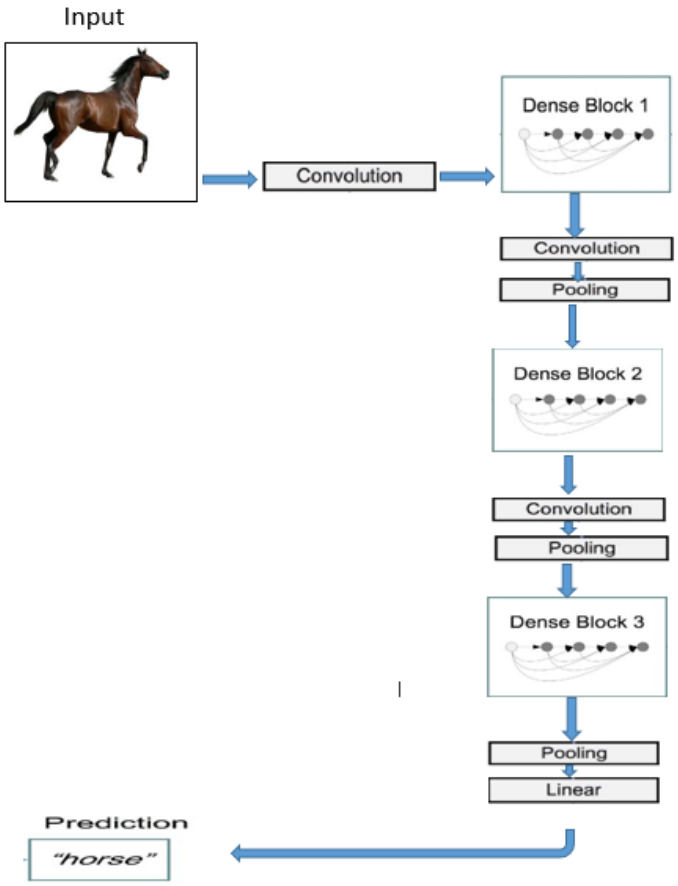

Figure 3.13: Full densenet example with 3 blocks from source paper [12]

In the DenseNet [12] architecture, each layer is connected to every other layer in a feed-forward fashion. DenseNet is very powerful and has achieved state of the art performance for many datasets. This model has been used to extract features which in turn has been used to train a svm.

3.3 Feature selection

In the previous section 3.2, different features have been extracted through different methods but not all of them are going to be helpful. This section of feature selection attempts to select the features which would be used to build the model for the given problem. Taking a lot of features would take a lot of training time and hence only appropriate features need to

be taken carefully into consideration. It would also help the users to understand the model in a simpler way and avoids overfitting. In the section 3.2.1, different handpicked features have been discussed along with the discussion whether we would be using it in our model for classifying images. Now, this section would attempt to narrow down the features extracted by pre-trained model. Table 3.1 lists the mean accuracy score and standard deviation using pretrained models. These values are calculated by repeating the experiment five times on different split of training and test data. The third row of the table lists the accuracy score when features extracted by both pretrained models are combined together to verify if the combination of two gives a better result.

Table 3.1: Pre-trained model accuracy score

Methods	Mean	SD
VGG16	0.473	0.025768
Densenet	0.473	0.025768
VGG16+Densenet	0.473	0.025768

It is evident from Table 3.1 that both VGG16 and Densenet are giving same accuracy score for our dataset. The idea behind trying different model was to combine the principal features extracted from both the models in order to get a feature set having much higher accuracy but the third row of Table 3.1 clearly illustrates that combining features would be of no help. It may be because these models have been trained on a very different dataset from the problem addressed in this report. So,the features extracted from Densenet would be discarded as it would be redundant. In the next step, the total features extracted by VGG16 are in thousands which must be narrowed down. In order to accomplish this task, the accuracy score after taking a particular number of principal features from VGG16 are combined with hand-picked features and corresponding accuracy score has been monitored. The accuracy score when the number of features from VGG16 is 100, 10, 5, 3, 2 is 0.65, 0.64, 0.64, 0.65 and 0.67 which indicates that the model would work better if only two principal features are taken into consideration. So, only two features extracted by the VGG16 model is used in chapter 4 for calculation of accuracy score.

Chapter 4

Results

This section attempts to find accuracy score of different features extracted and their various combinations. The dataset available has been divided into the training set (70%) and test set (30%) for the experiment. In section 4.1 an attempt has been made to classify images in the dataset into four classes which can be merged to form three classes as discussed in chapter 1. In the section 4.2, an attempt has been made to classify images in three classes where normal and benign has been merged into one class. Support vector machines with linear kernal have been used for the problem. Given a set of training examples, each marked as belonging to one or the other class, it builds a model that assigns new examples to one category or the other. Different combination of training and test data has been used and the corresponding mean and standard deviation has been calculated and are mentioned in tables. The mean and standard deviation has been calculated by repeating the experiment five times with different training and test split.

4.1 Classification into four classes

In this section, the images are classified into four classes i.e normal, benign, in situ carcinoma and invasive carcinoma based on the features extracted and the corresponding mean of accuracy score and standard deviation has been noted down in tables.

4.1.1 Hand engineered Features

Nomenclature : For simplicity, features has been renamed in the graph, count of blobs(LoG) as h1, count of blobs(DoG) as h2 and average area of contours as h3.

Table 4.1 lists down the accuracy score of individual features on different combination of training and test data. The Figure 4.1 corresponds to the data in table 4.1 and gives a better visualization for comparison between different features. The values are calculated by repeating the experiment five times on different training and test split.

Table 4.1: Accuracy score of individual features for four class classification

Methods	Mean	SD
LOG	0.3152	0.034098
DOG	0.374	0.057706
Average Area	0.336	0.029665

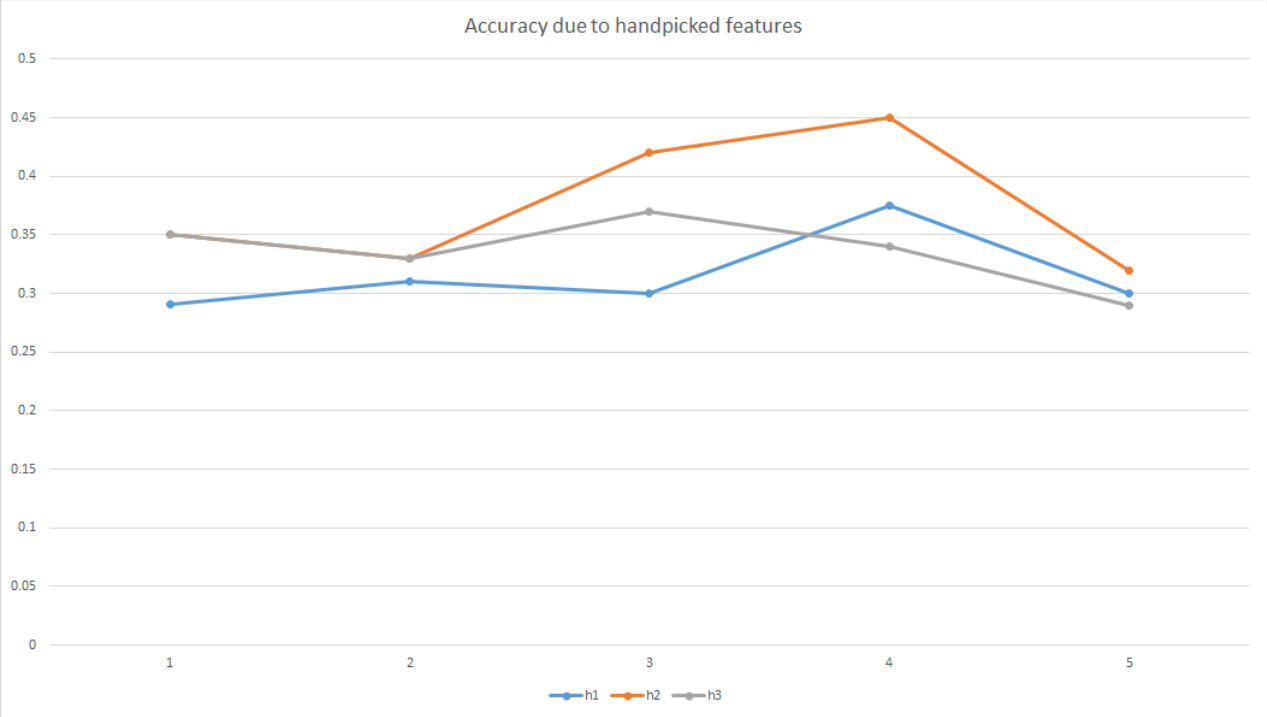

Figure 4.1

The table 4.2 note down the mean and standard deviation when SVM is trained on different

combinations of the hand-picked features. It helps in understanding whether the features are redundant or not. Figure 4.2 corresponds to the table 4.2 and shows the change in accuracy of the combination of features over different training and test split.

Table 4.2: Accuracy score of combination of features for four class classification

methods	Mean	SD
LOG + DOG	0.382	0.049193
LOG + Average Area	0.404	0.02881
DOG + Average Area	0.4152	0.054085

The values are calculated by repeating the experiment five times on different training and test split.

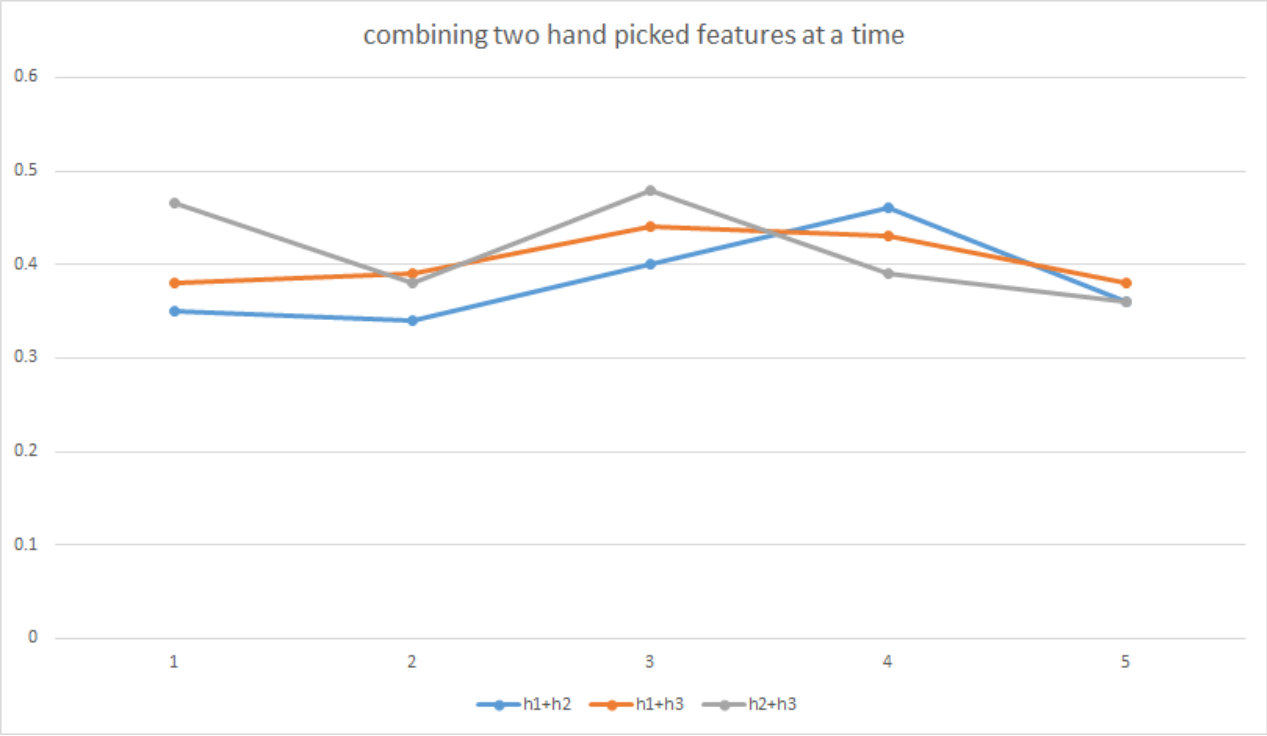

Figure 4.2

Further, in Table 4.3, all the three selected hand engineered features are merged and figure 4.3 gives an insight that all three features combined give a higher accuracy than pairwise features for most of the training and test set combination.

Table 4.3: Accuracy score for all three features

methods	Mean	SD
LOG + DOG + Average Area	0.443	0.033466

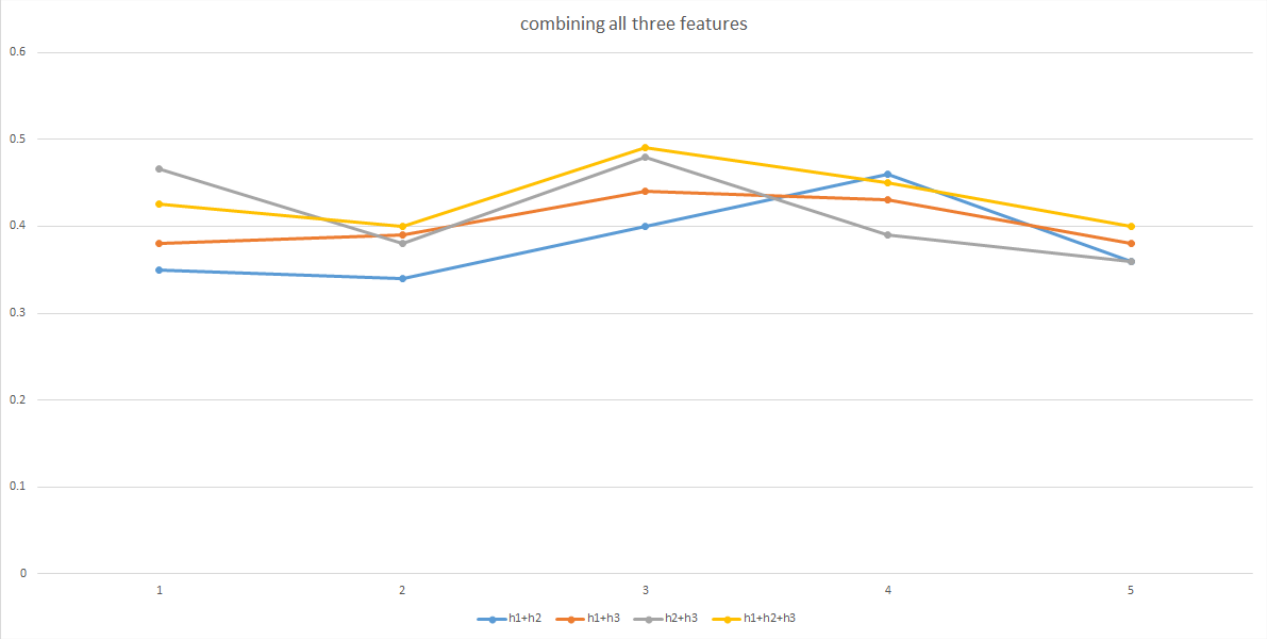

Figure 4.3

4.1.2 Pretrained Models

In this section, accuracy score of features extracted through pre-trained models is shown. The models are trained on the imagenet dataset and the same weights and biases are used to extract features from the new dataset. Since the available dataset is quite small so the model is left as it is and not even a single weight is reassigned as there is a chance of overfitting. In section 3.3, it has been tried to extract features from two different models and combine them so that the overall accuracy increases but combining features do not lead to an increase in accuracy. It may be because of the fact that both are trained on imagenet dataset which

is quite different the breast cancer image dataset. It is also noteworthy that both give the same accuracy score when an SVM is trained on top of the features extracted by them. It is also noteworthy that only two features extracted from VGG16 have been selected for further analysis.

4.1.3 Combining features

In this section, the features extracted through hand engineered methods are combined with the features extracted through VGG16. It can be seen in Fig 4.3 that features LoG,DoG and average area combined outperform other combinations most of the times but at some points, DoG and average area combined and some point LoG and DoG combined outperforms it. So, table 4.5 lists the performance of three different combinations.

Table 4.4: Experimental data

Methods	Mean	SD
DOG + Average Area + VGG16	0.652857	0.034983
LOG + DOG + VGG16	0.628571	0.045981
LOG + DOG + Average Area + VGG16	0.651429	0.037273

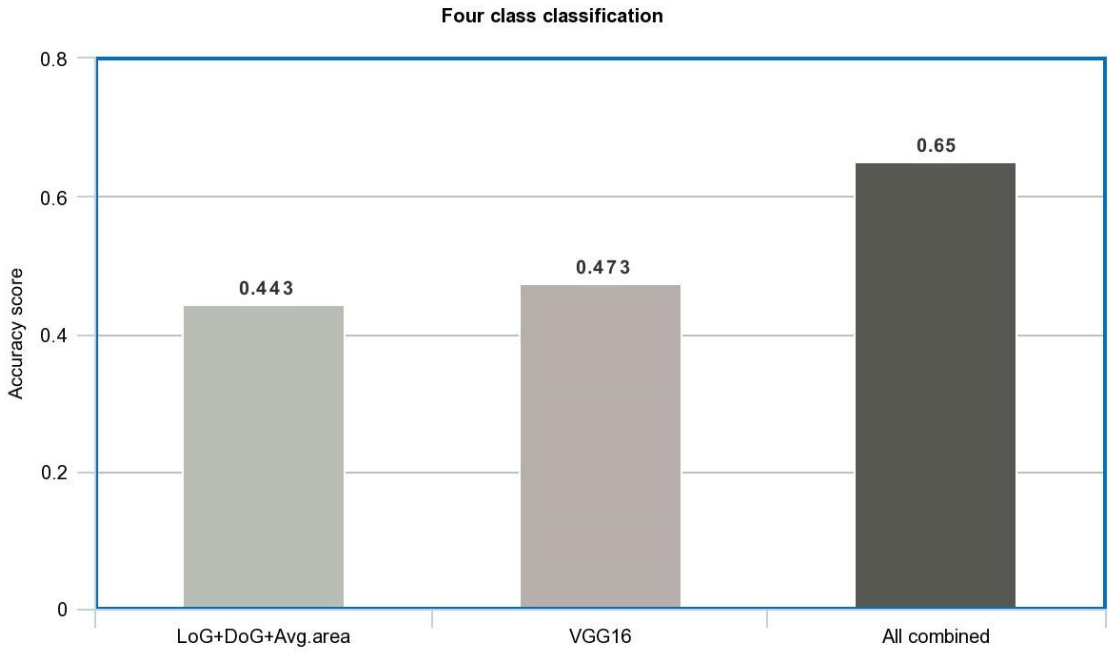

Figure 4.4

The figure 4.4 illustrates different features and the corresponding accuracy score.

4.2 Classification into three classes

This part is inspired by the fact that detecting cancerous images is of prime importance for pathologists. This task is not part of ICIAR 2018 challenge. Moreover, the work done by Brook et al. [9] also classifies images into three classes. In this section, the images are classified into three classes based on the features extracted and the corresponding accuracy score has been noted down in tables. In this section, Normal and Benign has been combined to form one class as they are non-cancerous and the same procedure of classification as in section 4.1 is followed for classification.

4.2.1 Hand engineered Features

Table 4.5 lists down the accuracy score of individual features on different combination of training and test data. These values has been calculated by repeating the experiment five times with different training and test set splitting.

Table 4.5: Accuracy score of individual features for three class classification

Methods	Mean	SD
LOG	0.428	0.043243
DOG	0.526	0.06269
Average Area	0.514	0.04669

4.2.2 Combination of features

In this section, hand-engineered features have been combined with the features extracted from VGG16. Table 4.6 lists the accuracy score of various combination of features.

Table 4.6: Accuracy score of combination of features for three class classification

methods	Mean	SD
VGG	0.628	0.038987
LOG + DOG + Average Area	0.628	0.022804
VGG + LOG + DOG + Average Area	0.748	0.038987

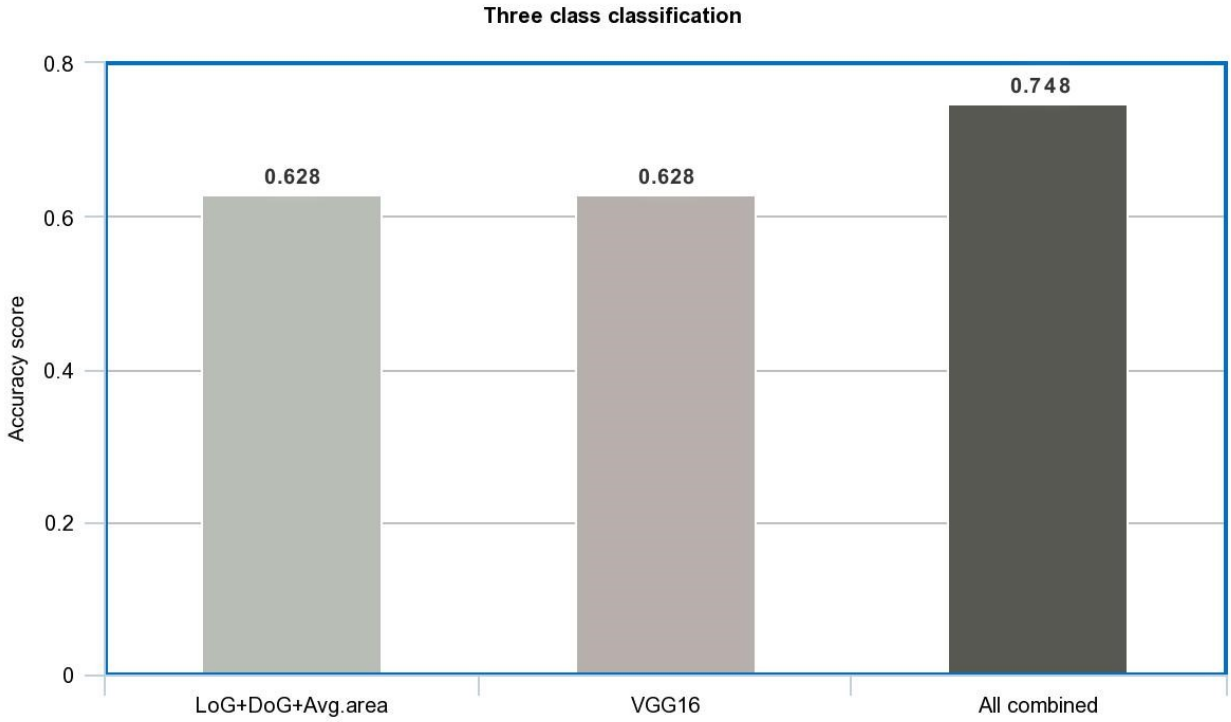

Figure 4.5

The figure 4.5 illustrates different features and the corresponding accuracy score for three class classification.

Chapter 5

Conclusion

In this report, an approach of classifying the dataset into four classes is discussed. In addition to this, the same approach is applied to classify the dataset into three categories by treating normal and benign classes as one class. The three category classification groups the normal and benign in the same group as they are non-cancerous while in-situ and invasive carcinoma is kept separate as these are of interest for pathologists. The approach has been able to obtain a mean accuracy of 0.65 ± 0.03 for the four class classification problem and 0.74 ± 0.03 for the three class classification problem. The features have been extracted from the dataset through different methods and combined together to train a support vector machine with the linear kernel. The accuracy obtained is good given the fact that the training set is quite small. The pretrained model is not fine-tuned due to the smaller size of available training data. It is also noteworthy that only five features are being used to train and so the training time is reduced drastically. It reduces the complexity of the model and hence easier for the user to understand.

In practical scenarios, The classification in the wrong category can be disastrous as early detection of cancerous tissue is required for proper treatment. This is prevented by giving different images of the tissue of the same patient to the classifier.

Bibliography

[1] Araujo T, Aresta G, Castro E, Rouco J, Aguiar P, Eloy C, et al. (2017) Classification of breast cancer histology images using Convolutional Neural Networks. PLoS ONE 12(6): e0177544. retrieved from https://doi.org/10.1371/journal.Pone.0177544

[2] Breast Cancer: Prevention and control. Retrieved feb 28 2018 from http://www.who.int/cancer/detection/breastcancer/en/index1.html

[3] Jerry R. Balentine ,DO, FACEP, Editor:Melissa Conrad Stppler,MD. Article on Breast Cancer,retrieved 26-04-2018 from https://www.medicinenet.com/breast_cancer_facts_stages/article.htm

[4] A. Jain, S. Atey, S. Vinayak, V. Srivastava:Cancerous cell detection using histopathological image analysis Int J Innov Res Comput Commun Engng, 2 (12) (2014), pp. 7419-7426

[5] Inderpreet Kaur, Dr. Amandeep Kaur https:Modified Active Contour Snake Model for Image Segmentation using Anisotropic Filtering,retrieved from //www.irjet.net/archives/V3/i5/IRJET-V3I5438.pdf

[6] Dateset from ICIAR 2018 Grand challenge on Breast Cancer Histology images. https://iciar2018-challenge.grand-challenge.org/home/

[7] M.A.Aswathy, M.Jagannath :Detection of breast cancer on digital histopathology images: Present status and future possibilities. https://doi.org/10.1016/j.imu.2016.11.001

[8] Documentation-scikitimage:retrieved on 01-03-2018 from http://scikit-image.org/docs/dev/`auto_examples/features_detection/plot_blob`.html

[9] Brook A, El-Yaniv R, Issler E, Kimmel R, Meir R, Peleg D. Breast Cancer Diagnosis From Biopsy Images Using Generic Features and SVMs. 2007; p. 116.

[10] Zhang B. Breast cancer diagnosis from biopsy images by serial fusion of Random Subspace ensembles. In: 2011 4th International Conference on Biomedical Engineering and Informatics (BMEI). vol. 1. Shanghai: IEEE; 2011. p. 180186.

[11] Karen Simonyan Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. CoRRabs/1409.1556 2014 http://arxiv.org/abs/1409.1556

[12] Gao Huang,Zhuang Liu,Laurens van der Maaten. Densely Connected Convolutional Networks. CoRRabs/1608.06993 2016 http://arxiv.org/abs/1608.06993

[13] Elmore JG, Longton G2, Carney PA, Geller BM, Onega T, Tosteson ANA,et al. Diagnostic concordance among pathologists interpreting breast biopsy specimens. JAMA.2015 Mar 17;313(11):1122-32. https://doi.org/10.1001/jama.2015.1405

[14] Jason Brownlee. A gentle introduction of transfer learning for deep learning retrived 26-04-2018 from https://machinelearningmastery.com/transfer-learning-for-deep-learning/

[15] Ruye Wang. 2016-10-19 Retrieved April 27 2018 from http://fourier.eng.hmc.edu/e161/lectures/gradient/node8.html