



An empirical study of handcrafted and dense feature extraction techniques for lung and colon cancer classification from histopathological images

Naresh Kumar^a, Manoj Sharma^{b,*}, Vijay Pal Singh^c, Charanjeet Madan^d, Seema Mehandia^e

^a Department of ECE, UIET, Panjab University, Chandigarh, India

^b Department of ECE, Giani Zail Singh Campus College of Engineering & Technology, MRSPTU, Bathinda, India

^c Department of Electronics & Communication Engineering, GJU S & T, Hisar, Haryana, India

^d Department of Electrical Engineering, GJU S & T, Hisar, Haryana, India

^e Department of Biotechnology, UIET, Panjab University, Chandigarh, India



ARTICLE INFO

Keywords:

Lung cancer
Colon cancer
Deep learning
Feature extraction
Machine learning

ABSTRACT

According to a 2020 WHO report, cancer is one of the main causes of deaths worldwide. Among these deaths, lung and colon cancer collectively responsible for nearly 2.735 million deaths. So, detection and classification of lung and colon cancer is one of the utmost priority research areas in the field of biomedical health informatics. In this article, comparative analysis of two feature extraction methodologies has been presented for lung and colon cancer classification. In one approach, six handcrafted features extraction techniques based on colour, texture, shape and structure are presented. Gradient Boosting (GB), SVM-RBF, Multilayer Perceptron (MLP) and Random Forest (RF) classifiers with handcrafted features are trained and tested for lung and colon cancer classification. In another approach, using the notion of transfer learning, seven deep learning frameworks for deep feature extraction from lung and colon cancer histopathological images are presented. The extracted deep features (as input attributes) are applied into conventional GB, SVM-RBF, MLP and RF classifiers for lung and colon cancer classification. However, in contrast to handcrafted features a significant improvement in classifiers performance is observed with features extracted by deep CNN networks. It has been found that the proposed technique obtained excellent results in terms of accuracy, precision, recall, F1 score and ROC-AUC. The RF classifier with DenseNet-121 extracted deep features can identify the lung and colon cancer tissue with an accuracy and recall of 98.60%, precision of 98.63%, F1 score of 0.985 and ROC-AUC of 01.

1. Introduction

Lung and colon cancer is one of the deadliest diseases in the world and it is reported that both may develop synchronously. As reported by Koich Kurishima et al. in 2018 [1] 17 patients of lung cancer out of 3102 developed colon cancer within one month. Lung cancer is primarily known to be associated with upper aero digestive tract cancer, but its association with gastrointestinal cancer should not be overlooked. Hence, it is important to study detection of lung and colon cancer together from the medical images for assisting early diagnosis and treatment.

The World Health Organisation (WHO) has reported that cancer has caused nearly 10 million deaths across the globe in the year 2020 [2,3]. Among various types of cancers, Lung cancer is reported with 2.21 million new patients and Colon cancer with 1.93 million cancer patients,

a combination of both types of cancers amounting to almost 21.4 % of all types of cancers as shown in Fig. 1 [2,3]. Mortality is also highest in case of Lung cancer & Colon cancer with 18 % and 9.4 % respectively as shown in Fig. 1. Lung Cancer is one of the most sophisticated diseases and if not treated in the early stages, causes even death. It majorly falls into two categories namely Non-Small-Cell-Lung-Cancer (NSCLC) and Small-Cell-Lung-Cancer (SCLC) [4,5]. NSCLC is the major cause of lung cancer occurrence with almost 80 to 85 percent of cases being estimated. Thirty percent of these tumours originated in cells that comprise the body's cavities inner layer along with surfaces. This category generally persists in the external part of the lungs (Adenocarcinomas) [5].

The detection process of Lung Cancer consists of three steps, namely, Segmentation, Feature Extraction, and Classification. Lung and colon cancer detection using Artificial Intelligence technique is becoming one of the most famous research areas nowadays. This is due to an increase

* Corresponding author.

E-mail addresses: naresh_uiet@pu.ac.in (N. Kumar), neelmanoj@gmail.com (M. Sharma).

in the death rate due to lung and colon cancer [2]. The incidence and humanity rate of lung cancer rank first among various cancers across the world [3].

In medical image classifications there are four main processes which includes (i) selection of dataset for training & testing, (ii) pre-processing of images from the selected dataset, (iii) feature extraction based on various parameters, (iv) classification of the images using different classifiers. Conventional feature extraction techniques include feature extraction based on texture, color structure, and shape known as handcrafted features and explained in detail in section 3. These feature extraction techniques have major concerns such as sensitivity to light intensity, quantization errors, high dimensionality on representation etc. Extracting spatial information, image indexing and retrieval are other challenges. These problems result in poor accuracy and precision, hence given scope for implementing classifications using machine learning. Machine learning and artificial Intelligence can employ a variety of feature extraction, feature optimization and classification techniques on the biomedical images so as to cure the disease at an early stage. Transfer learning can be employed in two ways: (i) feature extraction from a pre-trained model and then training a classifier on top of it (ii) Fine tuning the pre-trained model keeping learnt weights as initial parameters. It is common to pre-train a convolutional neural network (CNN) on a very large data-set and then use the pre-trained model as a fixed feature extractor for the task of interest. Primary advantages of transfer learning are to save time and resources from having to train multiple machine learning models from scratch to complete similar tasks.

Wang et al. in 2020 [6] have used a transfer learning-based approach in classification of multiple sclerosis images and recently Zhang et al. 2021 [7] have shown COVID-19 diagnosis using combined DenseNet and optimization of transfer learning setting (OTLS) on chest computed tomography (CT) images. Methodologies based on deep learning are getting attention from researchers in multiple areas like cybersecurity, natural language processing, bioinformatics, robotics and control, and especially in medical imaging, or medical image investigation [8].

In the current scenario Transfer Learning and Artificial Intelligence has been observed playing an important role, especially in the diversified domain of the medical sector. These computational intelligence techniques are non-invasive techniques to diagnose the disease without harming the patient. Recent statistics observe that the detection and classification of lung cancer disease are one of the most tiresome tasks in the field of medical care. It is difficult to create a database of Machine Learning ready medical images for the implementation of algorithms for

classification as well as early diagnosis of cancers. Recently, an image dataset (LC25000) with 25,000 color images in 5 classes wherein every class contains 5,000 images of colon adenocarcinoma, benign colonic tissue, lung adenocarcinoma, lung squamous cell carcinoma, and benign lung tissue [9].

We have implemented classification algorithms using this dataset for training, testing of lung and colon cancer together. The rest of the manuscript is organised as follows: Literature survey is discussed in Section 2. Section 3 describes the proposed methodology. Section 4 presents the performance evaluation of models with handcrafted and dense CNN extracted features. Section 5 presents the discussion and conclusion is presented in section 6.

2. Literature survey

A good amount of work has been reported for classification of histopathological images of various types of cancer using deep learning algorithms for different datasets. LC 25000 image dataset was used by Garg et al. in 2020 [10] for a pre-trained CNN-based model for identification of lung and colon cancer. Eight CNN based models were reported for classifying malignant and benign images with good accuracy but authors considered colon and lung cancer separately. No result with respect to accuracy, F1 score & other parameters are presented for complete lung and colon dataset. Teramoto et al. in 2017 [11] used Deep Convolutional Neural Networks (DCNN) for classification of lung cancer types from cytological images and showed around 71% classification efficiency. Classification is based on DCNN which uses three convolutional layers, three pooling layers, and two fully connected layers giving accuracy of 71% only. The image dataset was considered for lung cancer only. Images were resized to 256X256 by cropping & resampling which have resulted in quantization error. Wang et al. in 2019 [12] also observed through their study that combination of histopathological imaging with deep learning methods can greatly help patients with lung cancer. Authors used CNN model to classify malignant and non-malignant images after resizing 300x300 pixels but could achieve only 89.8% accuracy. Shapcott et al. in 2019 [13] implemented a cell identification algorithm using deep learning using systematic sampling of tissue regions for performance enhancement without compromising accuracy for colon cancer. Authors considered classification based on the cell density and morphological features were extracted using deep learning for detection and prediction of colon cancer. Nur Ibrahim et al. in 2021 [14] proposed a non-complex Convolutional Neural Network (CNN) model for the image dataset 2500 which consists of tumor,

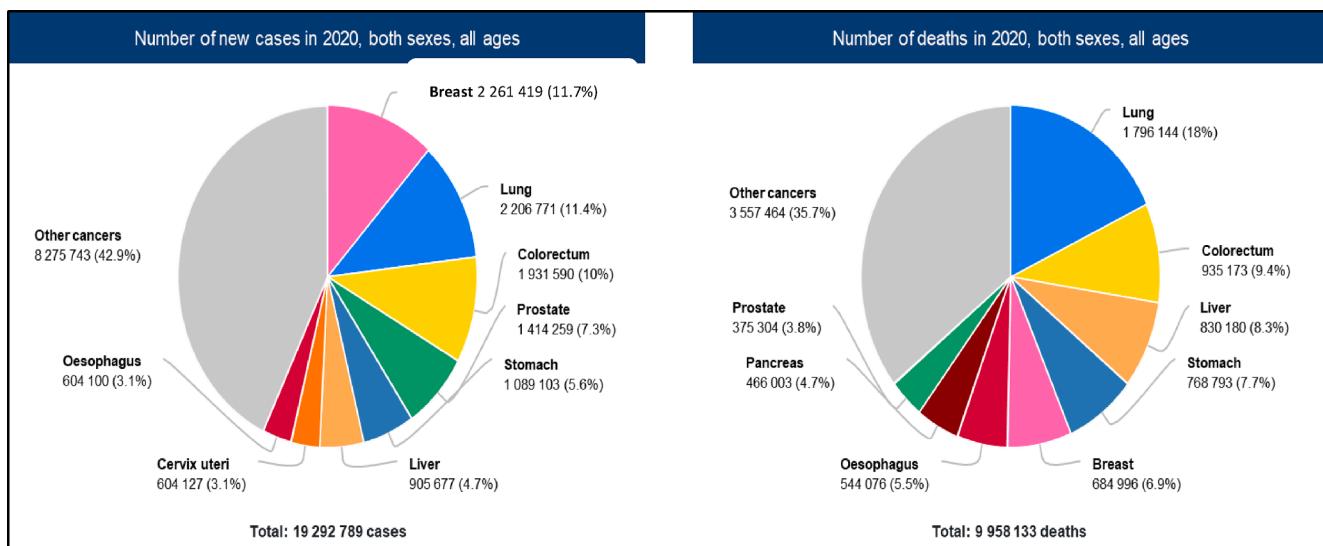


Fig. 1. Cancer incidence and mortality statistics worldwide in the year 2020 [2,3].

complex, lymphoma, and stroma and reported 83% accuracy for detecting four classes of colon cancer. Authors have extracted features from 150×50 -pixel textures using histological pictures and provided results for four types of colon cancer. They have used smaller datasets and accuracy is also less. No information is provided regarding sensitivity and selectivity of the classifiers. Wang et al. in 2021 [15] proposed Machine Learning based algorithms for gender specific classification of Lung cancer and achieved prediction accuracy of 82.9% and 73.2 % for females and males respectively. More investigation is required for identification of gender specific features which may or may not exist for improving accuracy. We still need to have a better system which can provide higher accuracy irrespective of gender for simplicity of implementation. Bhatia et al. (2019) [16] showed a strategy to detect lung cancer, from CT scan using deep residual learning. The researchers have described a pipeline of pre-processing methods to underline lung cancerous regions and take the features using UNet and ResNet models. The feature set is fetched to many classifiers viz. XGBoost and Random Forest, and the individual predictions are ensembles for predicting the possibility of a CT scan to be cancerous. 84% of accuracy is achieved in LIDC-IRDI than conventional methods [16]. FCH histogram feature extraction technique was used by [17] in 2002. It was also used to reduce the impact of noise because of illumination changes and quantization errors. Authors have presented feature extraction techniques for the application of image indexing and retrieval and not for classification.

From the literature survey, we examine that a limited number of studies are performed on the LC25000 dataset for colon cancer and lung cancer. Earlier studies are carried out on a smaller sample size for classification of lung and colon cancers. A very limited number of models have been studied for classification of lung and colon cancer together. Most of the models presented have used pre-trained CNN models and transfer learning-based feature extraction for lung or colon cancer whereas in our approach, we have used transfer learning-based feature extraction for lung and colon cancer together. The objective of this research work is to develop an efficient cancer (lung and colon together) detection technique that can have a high level of accuracy and performance.

The main highlights of manuscript are:

i) To the best of knowledge, it is for the first-time that two approaches of feature extraction: (a) conventional handcrafted features extraction techniques and; (b) transfer learning approach using pre-trained CNN networks as feature extractor are presented for histopathological images of Lung and Colon cancer.

ii) Transfer learning approach using pre-trained Deep CNN networks with conventional classifiers is used for lung and colon cancer histopathological images classification. In the proposed study, the transfer learning approach is used as a feature extractor that differentiates our study from previous studies [18,19].

iii) A detailed comparative performance analysis of transfer learning approach as feature extractor and handcrafted features is presented. The extracted features are utilized with conventional supervised machine learning algorithms for classification and detection of lung and colon cancer. The performances of models are compared in terms of accuracy, sensitivity, specificity, F-1 score, MCC, RoC-AUC curve and precision recall curve.

3. Proposed methodology

The proposed methodology is shown in Fig. 2. The proposed methodology consists of a) Lung and Colon cancer histopathological image dataset analysis b) image resize and pre-processing c) feature extraction from image dataset using handcrafted and deep learning technique d) comparative analysis of features extracted as input attribute to supervised learning classifier for detection and classification of lung and colon cancer tissue.

3.1. Dataset description and pre-processing:

The proposed research worked with image dataset known as the LC25000 dataset for classification of lung and colon cancer tissues into five classes. The LC25000 dataset images were collected at James A. Haley Veterans' Hospital situated in Tampa, Florida. The dataset was assembled by Andrew A. Borkowski et al. [9]. The dataset consists of 25,000 (5000 images for each class) coloured histopathological images of Lung Adenocarcinoma, Lung Squamous Cell Carcinoma, Benign Lung Tissues, Colon Adenocarcinoma and Benign Colonic Tissues. Originally 1250 images (250 images for each class) of cancer tissues from pathology glass slides were collected. The authors [9] use image augmentation techniques: left and right rotations (upto 25 degrees, 1.0 probability) and horizontal and vertical flips (0.5 probability) to expand the dataset to 25,000 images. The original 1024×768 pixels' images were resized to 768×768 pixels to make them square. All images are hematoxylin-eosin (H&E) stained, de-identified, Health Insurance Portability and Accountability Act compliant, and are freely available for download [20]. Colon Adenocarcinoma is accountable for more than 95% of colon cancer and primarily found in the large intestine due to development of a particular type of polyp (growth of tissue) known as Adenoma, which at later stage transformed into cancer. Lung Adenocarcinoma is the most common type of lung cancer and accounted for more than 40% of all lung cancer. It mainly grows in glandular cells and then extends towards alveoli in the lungs. Squamous Cell Carcinoma is the second most common type of lung cancer which accounts for nearly 30% of all lung cancer. This cancer appears in the lung's airways or bronchi. The benign lung and colon tumours are normally not life threatening, still they need to be surgically removed.

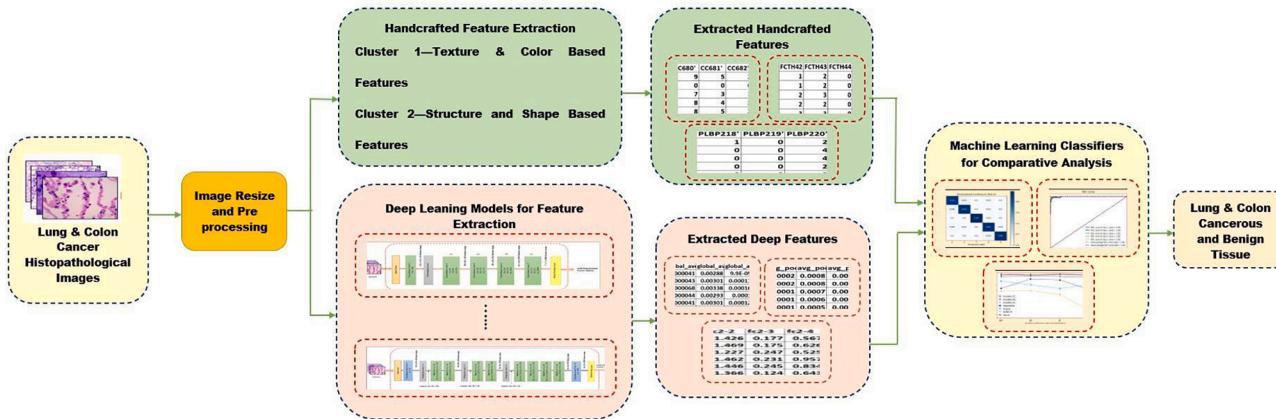


Fig. 2. Proposed methodology.

Table 1 presents the lung and colon cancer tissue classes, their class ID used in further sections and number of samples. **Fig. 3** shows the lung and colon tissues from LC 25000 dataset.

3.2. Handcrafted feature extraction

In this approach, for an unbiased comparison, the 768×768 sized labelled histopathological images are resized to dimensions of 224×224 . The images are then randomly divided training (80%) images and testing (20%) images. In the proposed manuscript, handcrafted features are broadly divided into two clusters. Cluster 1 extracts texture and color-based features and cluster 2 extracts structure and shape-based features. Cluster 1 feature extraction includes: Fuzzy Color and Texture Histogram Features (FCTH), Color Correlogram (CC) and Color Layout (CL) feature extraction algorithms. Cluster 2 feature extraction includes: Edge Histogram (EH), Pyramid Histogram of Oriented Gradients (PHOG) and Pyramid based Local Binary Patterns (PLBP) approaches to extract feature vectors. The extracted features from cluster 1 and cluster 2 are utilized as input attributes to conventional classifiers for classification and detection of lung and colon cancer.

3.2.1. Cluster 1: Texture and color based handcrafted features

In this cluster the features of histopathological images are extracted on the basis of texture and color of the image. FCTH, CC and CL-based feature extraction approaches are employed for feature extraction.

FCTH feature extraction technique: FCTH feature extraction was proposed by Savvas et al. in 2008 [21]. FCTH is a low-level feature which is a combination of histogram, colour and texture. There are 3 Fuzzy systems which are combined to extract FCTH features. First fuzzy system is applied on the colour and its related features like Hue, Saturation and Value; their membership values are redefined and new limits are set to calculate fuzzy color segmentation. Second fuzzy system is the Fuzzy texture in which energy from high frequency bands of wavelets are used to extract information about luminosity from the YIQ colour space. In the third fuzzy system, each image block is transformed with Haar Wavelet transform and a set of texture elements are exported which are applied at the input of the fuzzy system which converts a 24-bins histogram into a 192-bins histogram. This feature is also implemented for image retrieval and various other applications. In the present paper, 192 such features are extracted and used in implementation of Lungs and colon cancer classifications.

CC feature extraction technique: Color Correlogram [22] extracts the color distribution of pixels in images like color histogram and it also extracts the spatial information of pixels in the images. The color correlogram uses a single-color space. It may not have robust discriminative features that this feature is clubbed with other features for classification whereas it is widely used in content-based image retrieval (CBIR) and Image indexing etc. In the present paper, 1024 such features are extracted & used in implementation. It is observed that cancer cells are mostly heterogeneous in shape and non-uniform in color.

CL based features: Color Layout extracts information about spatial distribution of color in an image [23]. We have used CL based features and extracted 33 such features. As it gives information about distribution of color in a grid like structure and hence it is a useful feature for image classification.

Table 1
LC 25000 Dataset Cancer Category and Class ID.

Cancer Category	Class ID	Number of Samples
Lung Adenocarcinoma	Class 1	5000
Lung Squamous Cell Carcinoma	Class 2	5000
Benign Lung Tissue	Class 3	5000
Colon Adenocarcinoma	Class 4	5000
Benign ColonicTissue	Class 5	5000

3.2.2. Cluster 2: Structure and shape based handcrafted features

In this cluster, structure and shape based handcrafted features of histopathological images are extracted for image classification. In this, Edge Histogram, PHOG and PLBP approaches are employed for feature extraction.

Edge Histogram feature extraction approach: The edge-detection in images is an efficient representation of the original image in which one can easily make out about the image after discarding information from the majority of the pixels that do not belong to any edges [24,25]. The differentiating information in medical images lies in the relatively small number of edges in that image and edge detection helps in detecting those small areas. Convolution is generally used for detecting edges in images. An edge histogram contains the frequency and the directionality of the brightness changes which is a unique feature which can't be replaced by color histogram or the homogeneous texture features. In the present work 80 edge histogram features were calculated and implemented for classification.

PHOG feature extraction approach: Histogram of oriented gradients (HOG), and pyramid histogram of oriented gradients (PHOG) features extracted from medical images are reported to be useful for classification of X-ray images [26,27]. PHOG is a flexible and compact vector descriptor to be used in standard learning algorithms with kernels. PHOG is based on spatial pyramid matching, and can be changed with varying degrees of spatial correspondence. Features related to appearance and shape representations can be extracted using PHOG. In the present work, 630 PHOG features are extracted from the medical images for the classification.

PLBP feature extraction approach: Locally Binary Patterns (LBP) is an efficient method used for texture feature extraction for images which thresholds the neighbouring pixels based on the value of the current pixel [28–29]. This method is very popular for face detection and pattern recognition approaches also. The LBP operator transforms an image into an array or image of integer labels describing small-scale appearance of the image. In PLBP histogram, the histogram is composed of a separate bin for every uniform pattern, and all non-uniform patterns are assigned to a single bin & helps in extracting significant texture feature vectors having texture resolution variations. We have extracted 756 such features using the PLBP approach for classification of medical images.

3.3. Convolutional neural networks and transfer learning approach for feature extraction

With the development in the field of Machine Learning (ML), nowadays ML algorithms have been used in the domain of medical and biomedical for classification of various types of disease and biomedical signals. Progression in the ML field and with development of deep learning algorithms, machines are enabled to process high dimensional data such as images and videos. Deep learning can be thought of as a subset of machine learning where the structure and function of human brains gives rise to development of learning algorithms [30]. The Convolutional Neural Networks (CNN) are sub-classification of deep neural networks and constructed by neurons with their respective weights, biases and activation functions. CNNs are structured with sequences of convolutional, sub-sampling/pooling and activation layers for feature extraction. Finally, at the end a multi-layered neural network SoftMax function/layer is added to generate probability distribution at output [31–32]. After proper supervised learning/training, the CNN can be used as classifier or feature extractor in case of transfer learning. Transfer learning involves utilization of pre-existing networks for obtaining parameters to train deep learning models. Transfer learning techniques are used for improving machine learning performance by taking the advantage of knowledge obtained from pre-existing networks/tasks [33]. Pang and Yang [34] defined transfer learning as: Assume feature space $F = \{f_1, f_2, f_3, \dots, f_n\}$ with probability distribution $P(F)$ in domain S defined by $S = \{F, P(F)\}$. For a given task T, with two

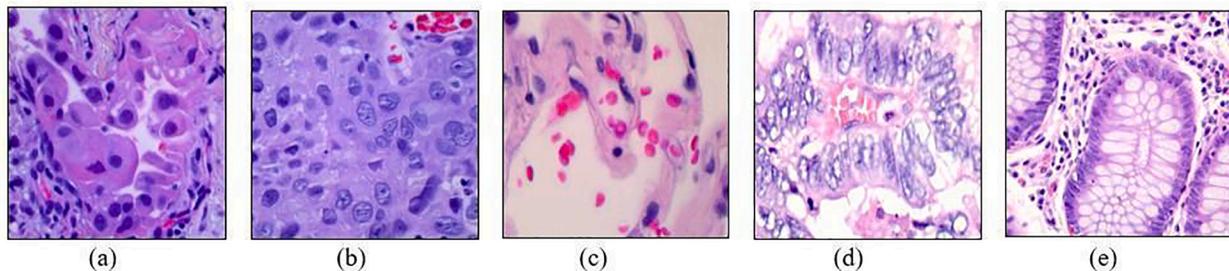


Fig. 3. Samples of tissues from LC 25000 dataset [9] (a) lung adenocarcinoma tissue; (b) lung squamous cell carcinoma tissue; (c) lung benign tissue; (d) colon adenocarcinoma tissue and; (e) colon benign tissue.

elements: $Y = \{y_1, y_2, y_3, \dots, y_n\}$ as the actual truth and another element as objective function $T = \{Y, g(\hat{A})\}$ learned through a training database and $g(\hat{A})$ can be used to predict class $g(f)$ for a new unknown instance f . In terms of probability, $g(f)$ can be given as $P(y|f)$. A training database F with actual output $Y = \{f_i, y_i\}$ is employed to train and learn the function $g(\hat{A})$ until it attains an acceptable error rate between $g(f)$ and actual value Y . For a given source domain space $Q_s = \{(f_{s1}, y_{s1}), (f_{s2}, y_{s2}), (f_{s3}, y_{s3}), \dots, (f_{sn}, y_{sn})\}$ and the learning task T_s and the target domain $Q_T = \{(f_{T1}, y_{T1}), (f_{T2}, y_{T2}), (f_{T3}, y_{T3}), \dots, (f_{Tn}, y_{Tn})\}$ and learning task T_T , the process of transfer learning seeks to improve the target predictive function $g_T(\hat{A})$ from the knowledge extracted from Q_s and T_s , assuming $Q_T \neq Q_s$ and $T_T \neq T_s$. In the transfer learning approach, the pre-trained networks can be employed for fine-tuning of pre-trained networks, feature extractor and baseline model. In transfer learning as fine-tuning of the network, the weights of the pre-trained model are transferred to the target model. In the baseline transfer learning approach, only pre-trained model's architecture is utilized and the targeted model is trained from the scratch with a given new dataset by initializing the weights randomly. In transfer learning as a feature extractor the conventional base of a pre-existing model is employed in its primary form, without altering the weights of the pre-trained model. Additionally, the classifier is trained on top of it by replacing the dense layers of the pre-existing model [35].

For the proposed methodology, we are using seven pre-trained CNN models as feature extractors. The performance of pre-trained models DenseNet (121, 169 and 201) [36], SqueezeNet [37], VGGNet [38], Xception [39] and ResNet [40,41] as feature extractor with conventional classifiers are analyzed and compared with handcrafted feature extractors.

3.3.1. Lung and colon feature extraction using DenseNet

Dense Convolutional Network (DenseNet) architecture was proposed by G. Huang et al. [36] in 2017. Fig. 4 shows the architecture of DeneseNet with layers and dense blocks. In this architecture, the layers

are connected in a feed-forward method i.e., every layer receives added inputs (feature maps) from previous layers and proceeds its own features to successive layers. In DenseNet architecture the features are combined by concatenating them. Depending upon the number of deep layers, we implement DenseNet-121, DenseNet-169 and DenseNet-201 for extracting deep features from lung and colon cancer histopathological images. The main advantage of Densenet is that it requires less parameters and computation to achieve better performance [36].

The numbers n_1 , n_2 and n_3 shown above deep blocks in Fig. 4 represent the number of repetitions of deep blocks for DenseNet-121, DenseNet-169 and DenseNet-201 respectively. A feature size of 1024, 1664 and 1920 are extracted from DenseNet-121, DenseNet-169 and DenseNet-201 respectively.

3.3.2. Lung and colon feature extraction using SqueezeNet

To achieve desired accuracy level, a smaller CNN network known as SqueezeNet was proposed by landola et al. [37]. Using fewer parameters SqueezeNet CNN network achieves comparable level of accuracy with AlexNet CNN network on ImageNet dataset. Fig. 5 shows SqueezeNet architecture as a feature extractor that comprises convolutional layer (Conv 1), max pooling layers (3 in numbers), 8 fire modules (fire 2–fire 9), convolutional layer (conv 10) and global pooling layer. The “fire module” incorporates a 1×1 filter squeeze convolution layer and two parallel 1×1 and 3×3 convolution filters in the expanded layer. The network follows three approaches during designing of architecture: (a) 3×3 filters are replaced with 1×1 filters to reduce the parameters as 1×1 filters has 9X less parameters in contrast to 3×3 filters, (b) “Squeeze layers” are used to reduce the number of input channels to 3×3 filters and, (c) Delayed down sample to have large activation maps. Approaches (a) and (b) are used to decrease the number of parameters and (c) to maximize the accuracy. Total 1000 deep features are extracted by the SqueezeNet network.

3.3.3. Lung and colon feature extraction using VGGNet

VGGNet (Visual Geometry Group) is similar to the AlexNet deep

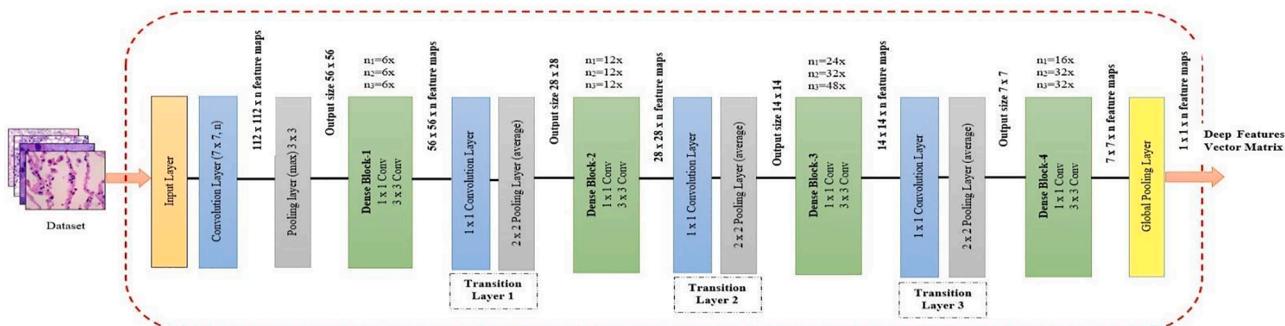


Fig. 4. The DenseNet architecture as feature selector with convolutional layers, pooling layers(max), dense blocks and transition layer. The number inside the block ($k \times k, n$) represents filter size ($k \times k$) and number of channels(n).

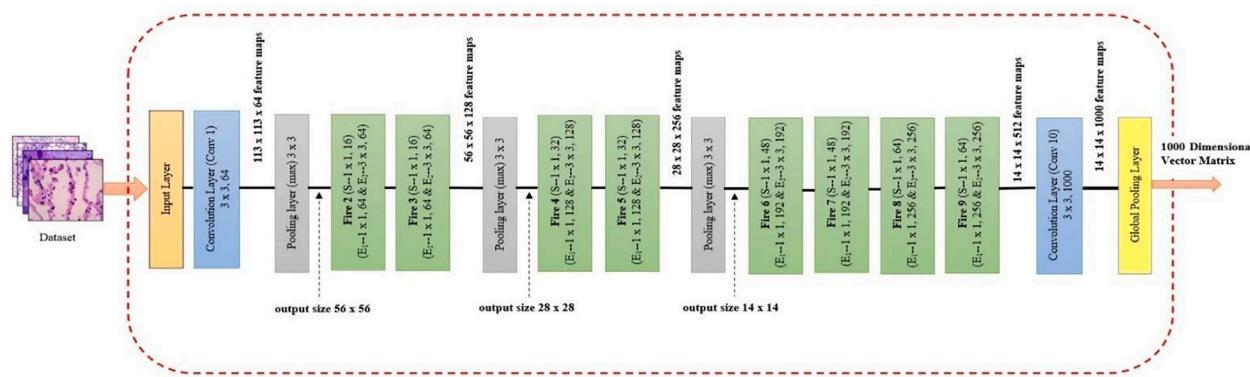


Fig. 5. The SqueezeNet architecture as feature selector with convolutional layers, pooling layers(max) and fire modules. The number inside the block ($k \times k$, n) represents filter size ($k \times k$) and number of channels(n).

network proposed by Simonyan, K et al. [38]. The VGGNet observes better performance in contrast to the AlexNet network [36]. Fig. 6 shows VGGNet architecture as a feature extractor with a number of extracted features. The VGG deep network extracts 4096 features from lung and colon histopathological images. In the VGG network additional convolutional layers with small 3x3 convolution filters are added to increase the depth of the network. The deep model consists of 13 convolutions layers, rectification and pooling layers with three fully connected layers. The VGG convolutional network was implemented with the aid of 3x3 filter and 2x2 pooling network. The proposed VGGNet-16 deep network extracts 4096 features from lung and colon histopathological images.

3.3.4. Lung and colon feature extraction using Xception

Xception deep CNN model was proposed by F. Chollet [39]. The model was initially inspired by Inception CNN architecture. Fig. 7 shows the architecture of the Xception network as a feature extractor. In the Xception network, initial inception modules are substituted by depth-wise separable convolutions with residual connections and are implemented without non-linearities [39]. For feature extraction the Xception deep model uses 36 convolution layers structured into 14 modules/blocks. Except uppermost and lowermost modules/blocks, all modules have linear residual connections around them. As shown in Fig. 7, the image information initially goes to the entry flow, then it goes to the middle flow where it repeats itself 8 times, and finally the data goes to the exit flow. Total 2048 features are extracted with the Xception model.

3.3.5. Lung and colon feature extraction using ResNet-50

A deeper residual neural network learning framework known as ResNet was proposed by K. He et al. [40,41]. Fig. 8 shows the ResNet-50 network. Instead of using unreference functions, the ResNet network uses learning residual functions at input layers. The ResNet achieves

good classification results on the ImageNet dataset. The ResNet has an advantage of reduced training time because of combined multiple sized convolution filters to manage the degradation problem [40]. The ResNet-50 extracts 2048 deep features from lung and colon histopathological images.

In Fig. 8, the numbers on top of convolutional blocks indicate recurrences of convolutional blocks and the output dimensional array of extracted feature maps by convolutional layers are mentioned at output of each convolutional layer block.

In our proposed manuscript, these seven pre-trained CNN architectures are employed to share their characteristics through transfer learning approach. The features extracted by respective CNN architectures are fed into traditional classifiers for classification of lung and colon cancer tissues.

4. Performance evaluation of models with handcrafted and dense CNN extracted features

Several experiments have been conducted using RF, MLP, GB and SVM-RBF classifiers with handcrafted and dense CNN extracted features for classifying the lung and colon cancer histopathological images. To evaluate the performance, the data set is randomly divided between training and test datasets. The training dataset (80%) is used to build and train the classifier and test dataset (20%) is used to test the model. The experiments were simulated using Python 3.8 on an IBM PC with Intel Core i7-6700 CPU @ 3.40 GHz processor, 8 GB RAM and NVIDIA GeForce GPU. The classification performance is measured in terms of accuracy, precision, recall, F1 score, Matthew's Correlation Coefficient (MCC) and Root Mean Square Error (RMSE). The term accuracy is used to define the correctly predicted number of events from the total number of test events. The parameter Precision/ Positive Predicted Value (PPV)

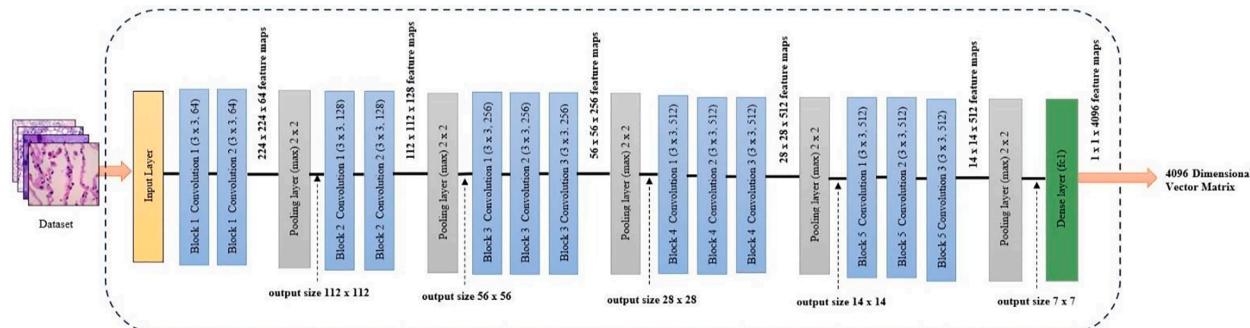


Fig. 6. VGGNet-16 architecture as feature selector with convolutional layers and pooling layers(max). The number inside the block ($k \times k$, n) represents filter size ($k \times k$) and number of channels(n).

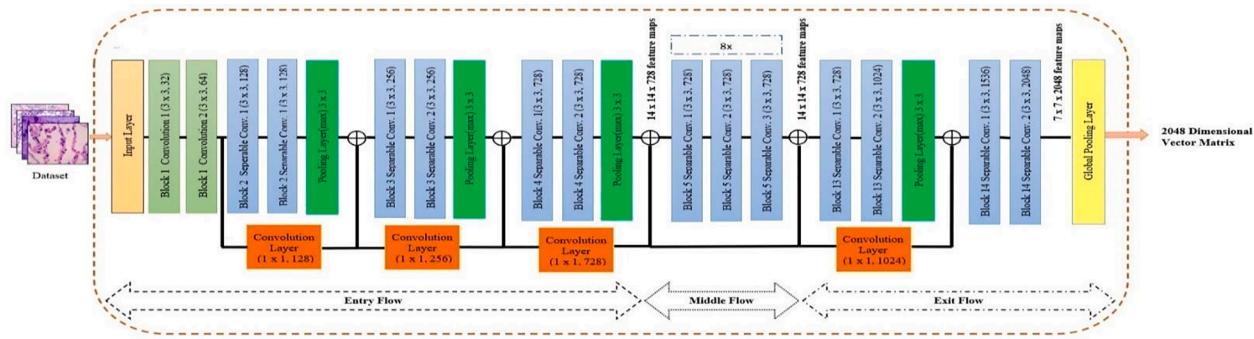


Fig. 7. Xception Network as feature selector with separable convolutional and pooling layers. The number inside the block ($k \times k, n$) represents filter size ($k \times k$) and number of channels(n). The combining point represents filter concatenation.

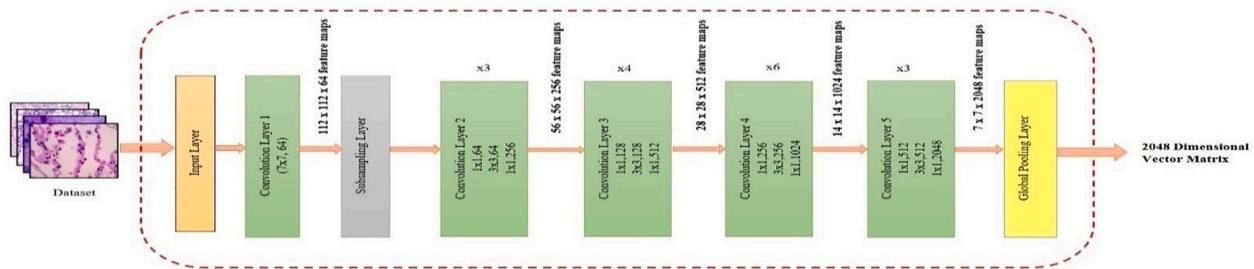


Fig. 8. ResNet-50 Network as feature selector with number of repeating blocks mentioned on top of each block. The number inside the block ($k \times k, n$) represents filter size ($k \times k$) and number of channels(n).

defines what fraction of predicted positives that are actually positives. The metric Recall/sensitivity defines the percentage of correct positive prediction from the total number of positives. High value of recall signifies percentage of correctly predicted cancerous tissues amongst actual cancerous tissues. The term F1 Score provides a weighted average value of the precision and recall. A model with a value of 1 as F1 score is assumed to be its best value and value of 0 to be its worst value. The metric MCC is used to find the Pearson product-moment correlation coefficient between true value and predicted value. The parameter RMSE measures the dissimilarity between predicted values and true values. It is given as the square root of the difference between squares of true values and predicted values for all events. The value of RMSE is proposed to be kept as low as possible. The detailed simulation parameters and performance measurement matrices are provided in supplementary material.

4.1. Handcrafted features-based lung and colon cancer classification

Cluster 1 and cluster 2 handcrafted features extracted from lung and colon cancer histopathological images are used as input to MLP, GB, RF and SVM-RBF classifiers for classification and detection of lung and colon cancer. **Table 2** presents classifiers performance for lung and colon cancer classification with cluster 1 handcrafted features. From **Table 2**, it has been found that MLP, GB, RF and SVM-RBF classifiers achieve more than 90% accuracy, precision and recall with color correlogram based extracted features as an input attribute. With color correlogram features, the classifiers attain more than 0.900 F1 score. The classifiers with CC features classify the lung and colon cancer histopathological images with an accuracy and recall in the range of 90.40% to 93.20%, precision in the range of 90.40% to 93.28% and F1 score in the range of 0.900 to 0.932.

In contrast to classifiers with CC features, the classifiers with FCTH features observe decline in performance. Classifiers with FCTH features evaluate accuracy and recall in range of 86.60% to 89.20%, precision in the range of 86.72% to 89.41% and F1 score in the range of 0.865 to

Table 2

Classification using Cluster 1(Texture and colour) feature extraction approaches.

Features	Classifier	Accuracy (%)	Precision (%)	Recall (%)	F1 Score
FTCH Features (192 features)	MLP	89.20	89.41	89.20	0.891
	GB	88.00	88.04	88.00	0.879
	RF	86.60	86.72	86.60	0.865
	SVM-RBF	86.80	86.99	86.80	0.867
CC Features (1024 features)	MLP	93.20	93.20	93.20	0.932
	GB	93.20	93.24	93.20	0.930
	RF	93.20	93.28	93.20	0.931
	SVM-RBF	90.40	90.40	90.40	0.900
CL Features (33 features)	MLP	74.80	75.19	74.80	0.748
	GB	87.00	87.04	87.00	0.869
	RF	86.20	86.31	86.20	0.861
	SVM-RBF	77.20	77.45	77.20	0.771

0.891. Declination of approximately 3% to 4% in accuracy, recall, precision and 0.035 to 0.04 in F1 score was observed. In cluster 1, an abrupt decline in performance of MLP and SVM-RBF classifiers with CL features was detected in classifying the lung and colon cancerous histopathological images. The MLP classifier with CL features shows unsatisfactory performance with an accuracy and recall of 74.80%, precision of 75.19% and F1 score of 0.748. In contrast to RF classifiers with CC features, MLP with CL features marks deterioration of approximately 18% in accuracy, precision, recall and 0.183 in F1 score. However, the GB and RF classifiers with CL features mark similar performance as that of classifiers with FCTH features.

Table 3 presents the classifiers performance analysis with cluster 2 feature extraction techniques. From **Table 3** it was found that classifiers with edge and shape-based features showed unsatisfactory performance in classifying the lung and colon cancerous and non-cancerous histopathological images.

Table 3

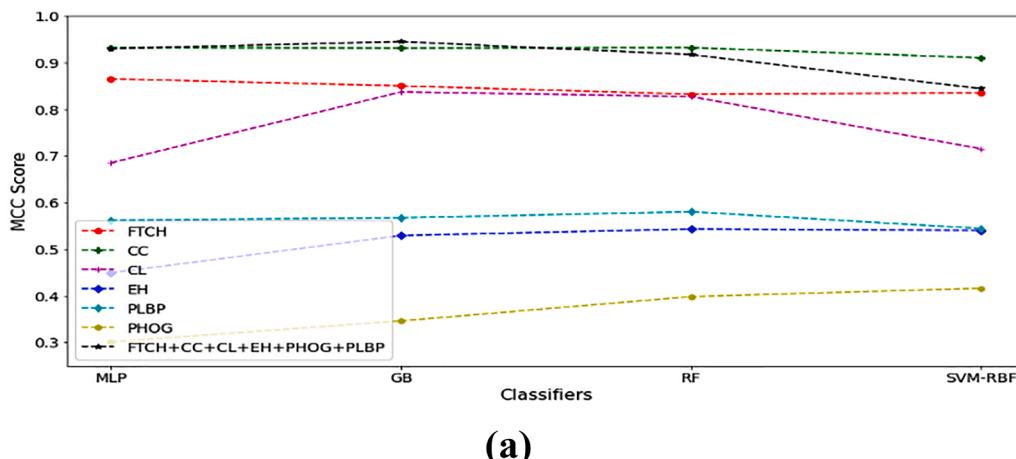
Classification using Cluster 2 (structure and shape) and combined (Cluster 1 + 2) feature extraction approaches.

Features	Classifier	Accuracy (%)	Precision (%)	Recall (%)	F1 Score
Edge Histogram (EH) features (80 features)	MLP	55.40	56.75	55.40	0.548
	GB	62.40	62.03	62.40	0.621
	RF	63.40	62.54	63.40	0.627
	SVM-RBF	63.20	63.00	63.20	0.628
PLBP (756 features)	MLP	64.80	65.81	64.80	0.642
	GB	65.20	65.95	65.20	0.654
	RF	66.40	67.01	66.40	0.663
	SVM-RBF	63.20	65.09	63.20	0.627
PHOG (630 features)	MLP	43.60	42.73	43.60	0.417
	GB	47.59	47.49	47.59	0.473
	RF	51.60	51.70	51.60	0.510
	SVM-RBF	53.20	53.02	53.20	0.528
Combined (Cluster 1 + 2) feature extraction approach					
FTCH + CC + CL + PHOG + EH+PLBP	MLP	94.39	94.55	94.39	0.944
	GB	95.60	95.62	95.60	0.955
	RF	93.40	93.56	93.40	0.934
	SVM-RBF	87.60	87.72	87.60	0.876

Classifiers with edge histogram features, PLBP features and PHOG features achieve classification accuracy in the range of 43.60% to 66.40% only. Classifiers with PLBP features as input attributes, demonstrates lung and colon cancer classification with an accuracy and recall ranging from 63.20% to 66.40%, precision ranging from 65.09% to 67.01% and F1 score in the range of 0.627 to 0.663. It has been noticed that classifiers with PHOG features as input attributes demonstrate unfavourable lung and colon cancer classification results. The classifiers obtained accuracy and recall in the range of 43.60% to 53.20%, precision in the range of 42.73% to 53.02% and F1 score in the range of 0.417 to 0.528. In contrast to classifiers with PLBP features, the classifiers with PHOG features perceive decline in classification performance. The classifiers marked decline of 13% to 19% in accuracy and recall, decline of 14% to 23% in precision and decline of 0.13 to 0.21 in F1 score. The classifiers with edge histogram features demonstrate accuracy, precision and recall in the range of 55.40% to 63.40%. The classifiers with edge histogram features marked improved performance in contrast to classifiers with PHOG features.

An improvement in classifiers performance has been observed by Combining color, texture, structure, and shape-based features into one feature vector. In contrast to the GB classifier with CC feature vectors, the GB classifier with combined features into one feature vector marked an improvement of approximately 2.40% in accuracy, precision and recall. The model showed an improvement of 0.02 F1 score.

Fig. 9 shows classifiers performance with handcrafted features in terms of MCC score and RMSE. From Fig. 9 (a) it has been found that a



(a)

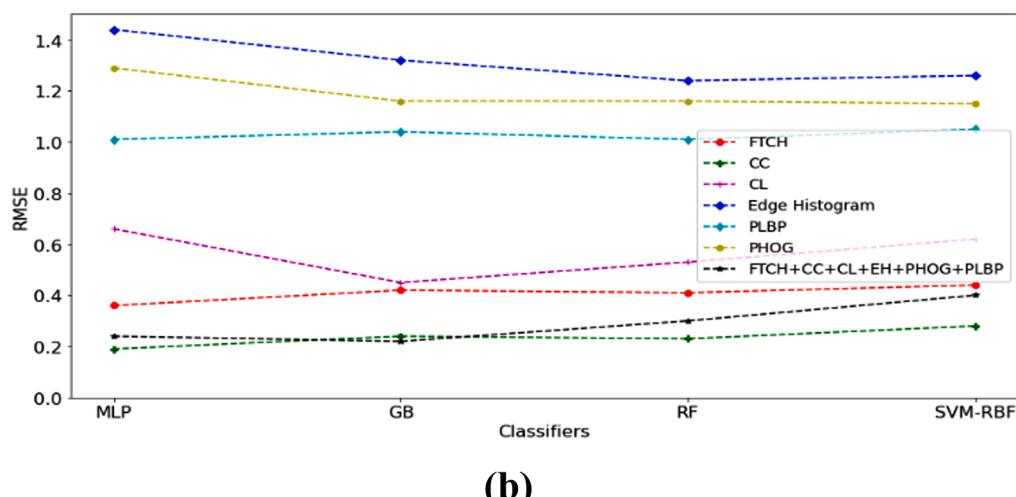


Fig. 9. Classifiers performance with extracted features in terms of (a) MCC; (b) RMSE.

GB classifier with combined features into one feature vector attains a maximum value of MCC. Classifiers with CC features as input attribute obtained MCC value of greater than 0.9. The higher value of MCC signifies high correlation between the actual and predicted values. With higher value of MCC, it is observed that GB classifier with combined features into one feature vector is effectively discriminating cancerous and non-cancerous histopathological images. Fig. 9 (b) shows classifiers performance in terms of RMSE. The classifier with CC features obtained least values of RMSE in the range of 0.19 to 0.28. GB classifier with combined features into one feature vector observes minimum value of RMSE. The model shows an error of 0.19 while predicting the results. From Table 2 and 3 it has been found that the texture and color based are more efficient in classifying the lung and colon cancerous and non-cancerous histopathological images in contrast to edges, shape and oriented gradient features. However, the classification accuracy is improved by combining all the color, texture, structure and shape features into one feature vector. The spatial shapes and edges-based features comprises low weight and less significant features in classifying the lung and colon cancer images.

4.2. Pre-trained deep CNN based lung and colon cancer detection and classification

For comprehensive analysis, the extracted features using pre-trained deep CNN networks are tabulated into two groups as shown in Table 4 and Table 5. Table 4 presents classifiers performance using features extracted by SqueezeNet, Xception, ResNet-50 and VGGNet-16 pre-trained networks. Table 5 presents the classifiers performance with deeply extracted features using the DenseNet CNN network with variation in the number of layers within the network. From Table 4 and Table 5, it has been observed that the classifiers with features extracted by SqueezeNet, Xception, ResNet and VGGNet pre-trained networks demonstrate deteriorated performance in contrast to classifiers with features extracted by DenseNet pre-trained networks.

From Table 4, it has been found that classifiers with ResNet-50 extracted deep features show the dissatisfied results in discriminating the lung and colon cancer histopathological images. The classifiers demonstrate accuracy and recall in the range of 48.19% to 76.04%,

Table 4
Classifiers performance with SqueezeNet, Xception, ResNet-50 and VGGNet-16 deep features.

Feature Extractor	Classifier	Accuracy (%)	Precision (%)	Recall (%)	F1 Score
SqueezeNet (1000 features)	MLP	79.20	79.55	79.20	0.790
	GB	91.80	91.95	91.80	0.918
	RF	90.60	90.79	90.60	0.906
	SVM-RBF	77.60	77.66	77.60	0.775
	Xception (2048 features)	89.00	89.08	89.00	0.890
ResNet-50 (2048 Features)	GB	84.00	84.28	84.00	0.840
	RF	79.20	79.26	79.20	0.791
	SVM-RBF	84.39	84.51	84.39	0.844
	MLP	75.40	76.49	75.40	0.740
	GB	76.04	76.16	76.04	0.760
ResNet-101 (2048 Features)	RF	71.20	71.06	71.20	0.705
	SVM-RBF	48.19	53.93	48.19	0.436
	MLP	81.20	82.08	81.20	0.806
	GB	76.80	77.10	76.80	0.768
	RF	73.20	73.02	73.20	0.725
VGGNet-16 (4096 features)	SVM-RBF	55.80	56.14	55.80	0.551
	MLP	95.00	95.49	95.00	0.949
	GB	95.60	95.65	95.60	0.956
	RF	95.19	95.34	95.19	0.951
	SVM-RBF	94.00	94.18	94.00	0.940

Table 5

Classifiers performance with DenseNet-201, DenseNet-169 and DenseNet-121 deep features.

Feature Extractor	Classifier	Accuracy (%)	Precision (%)	Recall (%)	F1 Score
DenseNet-201 (1920 features)	MLP	98.20	98.26	98.20	0.981
	GB	97.00	97.06	97.00	0.970
	RF	97.80	97.82	97.80	0.978
	SVM-RBF	97.00	97.18	97.00	0.970
	Xception	97.21	97.21	97.20	0.972
DenseNet-169 (1664 features)	MLP	97.21	97.21	97.20	0.972
	GB	98.20	98.22	98.20	0.982
	RF	98.00	98.02	98.00	0.979
	SVM-RBF	97.39	97.44	97.39	0.974
	DenseNet-121	98.40	98.39	98.40	0.983
DenseNet-121 (1024 Features)	MLP	98.40	98.39	98.40	0.983
	GB	98.00	98.02	98.00	0.980
	RF	98.60	98.63	98.60	0.985
	SVM-RBF	97.60	97.65	97.60	0.975

precision in the range of 53.93% to 76.16% and F1 score in range of 0.436 to 0.760. However, the classifiers marks improved in performance with features extracted by ResNet-101 CNN network and the classifiers attains accuracy and recall ranging from 55.80% to 81.20%, precision in the range of 56.14% to 82.08% and F1 score in range of 0.551 to 0.806. The classifiers with the same number of deep features extracted by Xception network shows significant improvement in performance in contrast to classifiers with ResNet-101 extracted features. The classifiers demonstrate accuracy and precision in the range of 79.20% to 89.00%, precision in the range of 79.26% to 89.08% and F1 score in range of 0.791 to 0.890. It has been observed that RF and GB classifiers with 1000 deep significant features extracted by SqueezeNet pre-trained network marks classification accuracy, precision and recall of more than 90% and more than 0.90 F1 score. However, MLP and SVM-RBF classifiers with SqueezeNet extracted features demonstrate degraded performance with approximately 11% to 13% decrease in accuracy, precision and recall. It was also found that except for SVM-RBF classifier with ResNet-50 and ResNet-101 extracted deep features, all the classifiers with input attributes as deep features extracted by SqueezeNet, Xception, ResNet-50 and ResNet-101 mark their classification accuracy ranging from 71.20% to 91.80%. The classifiers with VGGNet extracted deep features show significant improvement in classifying the lung and colon cancer histopathological images. Classifiers with VGGNet extracted features attained more than 94% accuracy, precision and recall. From Table 4 it was found that the GB Classifier with VGGNet extracted features shows impressive performance with more than 95% accuracy, recall and precision.

From Table 5 it has been found that classifiers with DenseNet extracted deep features demonstrate excellent classification results in terms of accuracy, precision, recall and F1 score. Classifiers with input attributes as deep features extracted by DenseNet-201, DenseNet-169 and DenseNet-121 attain more than 97% accuracy, precision, recall in classifying the lung and colon cancerous and non-cancerous histopathological images.

It has been found that with the DenseNet network as a feature extractor as the number of dense layers decreases, the classifiers mark an improvement in performance. Classifiers with DenseNet-121 extracted features, having the least number of dense layers and extracted features, demonstrates excellent classification results in terms of accuracy, precision, recall and F1 score. MLP, GB and RF classifiers with deep 1024 features extracted by DenseNet-121 network, predicts the classification and detection of lung and colon cancer with more than 98% accuracy, recall and precision. RF classifier with features extracted by DenseNet-121 feature extractor appraises the classification with maximum accuracy and recall of 98.60%, precision of 98.63 and F1 score of 0.985.

Fig. 10 (a) and (b) shows the MCC score and RMSE attained by the

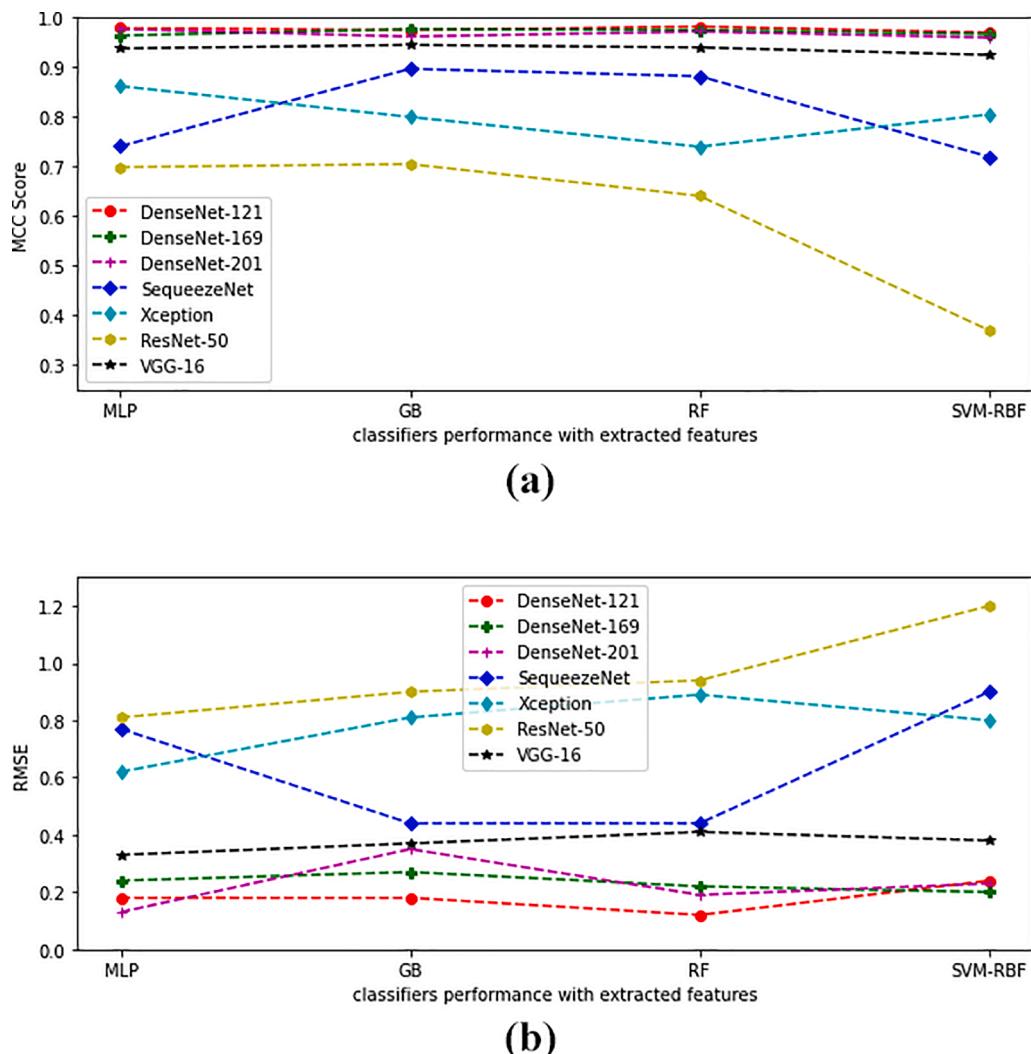


Fig. 10. (a) Classifiers performance with deeply extracted features in terms of MCC (b) classifiers performance with deeply extracted features in terms of RMSE.

classifiers using deeply extracted features. It has been found that classifiers with features extracted by all variants of DenseNet evaluate approximately equal and high MCC scores. The high value signifies that all the classifiers are capable of classifying the cancerous as well as non-cancerous tissues with high accuracy. However, DenseNet-121 attains least value of RMSE amongst all the classifiers. The Low RMSE indicates closeness of predicted value with true value. The low value of RMSE for all the classifiers with Densenet 121extracted deep features indicates

high significance of extracted features. From Table 4 and 5 it has been observed that classifiers outperform in lung and colon classification with features extracted by DenseNet pre-trained networks in contrast to features extracted by other proposed pretrained networks. DenseNet with 121 layers' extracts highly significant deep features in contrast to other extracted deep features.

Fig. 11 shows the confusion matrix of classifiers with features extracted by pre-trained deep networks. From Fig. 11 (a) it has been

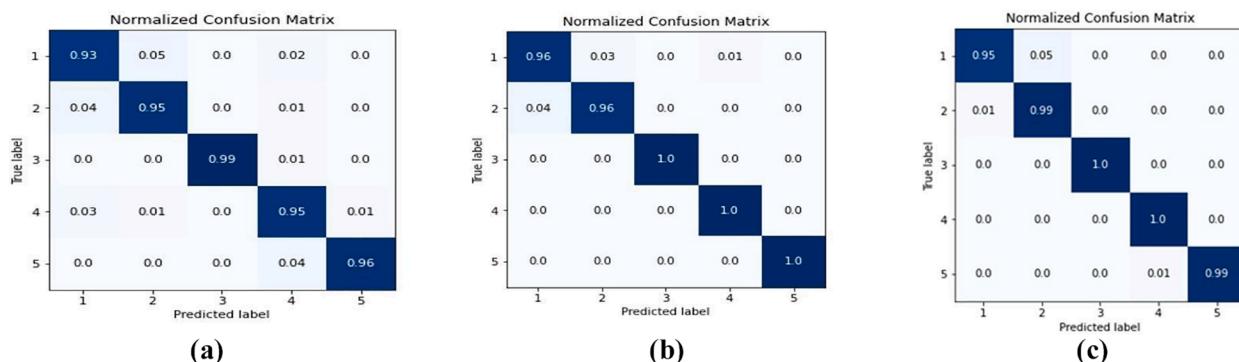


Fig. 11. Confusion matrices for: (a) GB classifier with VGGNet-16 extracted features, (b) MLP classifier with DenseNet-121 extracted features and (c) RF classifier with DenseNet-121 extracted features.

found that GB classifier with VGGNet-16 extracted deep features correctly classifies and detects Colon Adenocarcinoma tissue and Benign Colonic Tissue with an accuracy of 95% and 96% respectively. An improvement in classification accuracy was observed for Benign lung tissue classification and accuracy of 99% was observed. Though the classifier with VGGNet-16 extracted deep features shows significant performance nevertheless the classifier misclassified the lung cancer tissue as colon cancer tissue with an error of 1% to 2%.

From Fig. 11 (b) and (c), an improvement in classification accuracy for lung and colon cancer tissue has been observed for RF and MLP classifiers with DenseNet-121 extracted deep features as input attributes. The MLP and RF classifiers classify benign lung tissues and Colon Adenocarcinoma tissues with 100% accuracy. The MLP classifier classifies the Benign Colonic Tissues with 100% accuracy. Classification accuracy of 96% was observed for Lung Adenocarcinoma, Benign Lung Tissue and Lung Squamous Cell Carcinoma tissue by MLP classifiers. The RF classifier classifies the Lung Squamous Cell Carcinoma tissues and benign colonic tissues with an accuracy of 99%. From Fig. 11 (b) and (c) it was found that RF and MLP classifiers reveals approximately same performance with deep features extracted by DenseNet-121 however, the MLP classifiers shows downside results for Lung Adenocarcinoma tissues and misclassified these tissues as Colon Adenocarcinoma with an error of 1%.

Fig. 12 shows the ROC-AUC curves for classifiers with features extracted by pre-trained deep networks. ROC-AUC curve represents the classifier performance in terms of True Positive Rate (TPR) and False Positive Rate (FPR) with varying threshold values [42]. The TPR gives measure (in percentage) of true cancerous tissues that are correctly identified and FPR gives the measure about predicting results as cancerous tissue for a benign tissue. From Fig. 12 (a), (b) and (c) it has been observed that although the classifiers reveal excellent prediction performance and attain micro-average and macro-average value of ROC-AUC as 1, nevertheless the classifiers attain dissimilar TPR and FPR values. For a perfect classification, the ROC-AUC curve should be flat at the top left corner i.e., value at x-axis shall be equals to 0 and y-axis shall be equals to 1. A flatness in curves at top left has been observed from Fig. 12 (a) to 12 (c) indicating an improvement in classifiers performance in terms of TPR and FPR. Differential variation in TPR and FPR values has been observed with increase in TPR and decrease in FPR values. In Fig. 12 (a) the GB classifier with features extracted by VGGNet-16 indicates least flatness in the ROC-AUC curve in contrast to Fig. 12(b) and 12(c).

However, a meaningful improvement in flatness in curves at top left is noted in Fig. 12(b) and 12(c) representing an improvement in performance in terms of predicting the true benign and true cancerous tissues. Maximal flatness in the ROC-AUC curve at top left corner has

been observed in Fig. 12(c) for RF classifiers with DenseNet-121 extracted deep features. Here the classifier gains maximum value of TPR and minimum value of FPR amongst all the classifiers. The high value of TPR and low value of FPR acquired by RF classifier indicates its capability to predict the cancerous tissues as cancerous tissues with very high rate and predicted the benign tissue falsely as cancerous tissue with minimal rate. Hence, the flatness on the left top in Fig. 12(c) indicates excellent prediction performance of RF classifiers not in terms of accuracy, precision and recall but also in terms of TPR and FPR.

4.3. Interpretability of model with deep CNN and handcrafted extracted features

From the previous section, it was found that classifiers with deeply extracted features showed preferable prediction results. However, the machine learning models and artificial neural networks demonstrate excellent performance in the medical field, still the models face the issue of interpretability and explainability [43,44]. The machine learning models accordingly detect, learn and extract the hierarchical data representations needed for classification and detection. The increasing complexity and huge amount of data boosts the system prediction and classification ability at the cost of reducing the ability to explain the inner workings and mechanism of the models [45]. As a consequence, the results predicted by these "black boxes" are hard to understand and interpret [45]. Many researchers proposed interpretability methods to explain deep learning models [46–48] and the concept of eXplainable AI (XAI) has emerged and provides a trade-off between explainability and predictive utility to enable the models to be more understandable to humans [49]. Herein, we have used two frameworks SHapley Additive explanation (SHAP) proposed by Lundberg and Lee [50] for global interpretation i.e., to explain the complete behaviour of the model and Locally Interpretable Model Agnostic Explanations (LIME) proposed by Ribeiro et al. [51] for local interpretation i.e., to explain the prediction for a single test instance. The SHAP framework defines the prediction in terms of a linear combination of binary variables that are used to describe whether an input feature is present in the model or not [49]. The framework defines results in terms of Shapley values. Shapley (SHAP) values define the feature importance and impact of features on the prediction model by considering three required properties: (a) local accuracy; (b) missingness and; (c) consistency [50]. Fig. 13 shows the impact of feature variables on the model and impact of feature variables on lung and colon classification. Fig. 13(a) presents the impact of feature variables on the model output. The x-axis represents mean SHAP value and Y- axis refers to feature variables. For instance, feature variable avg_pool262 attains a higher SHAP value for Lung Adenocarcinoma tissue as compared to other classes of lung and colon cancer tissues. The

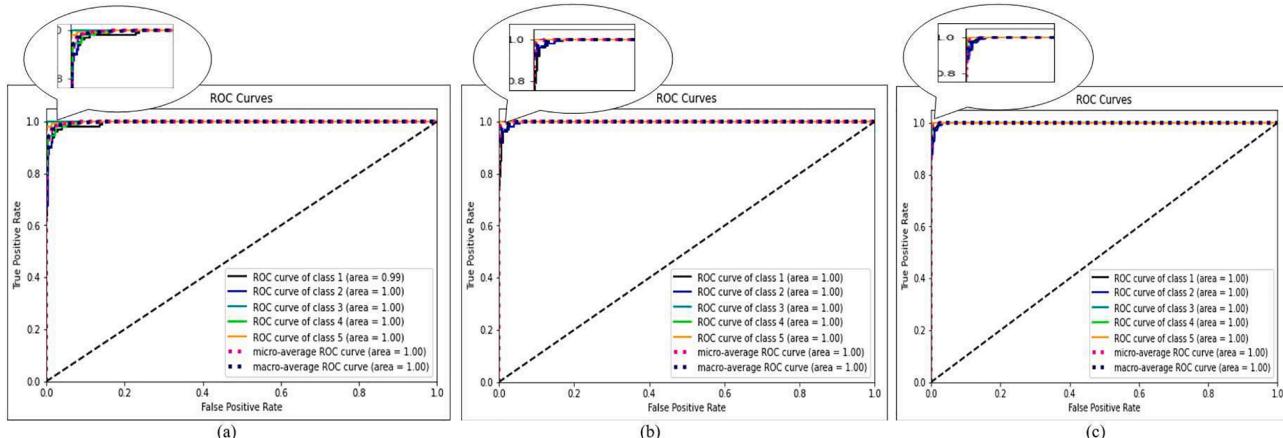
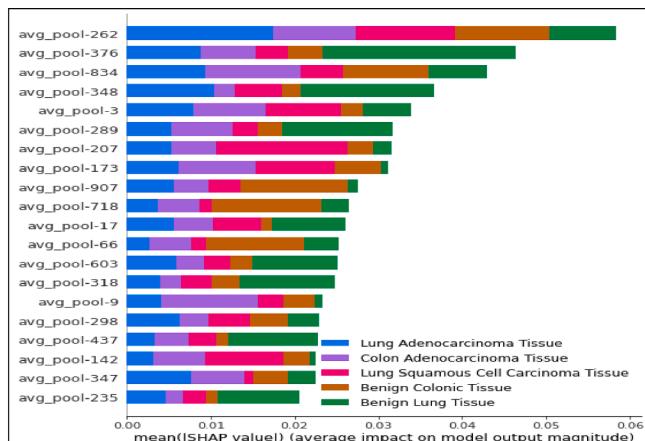
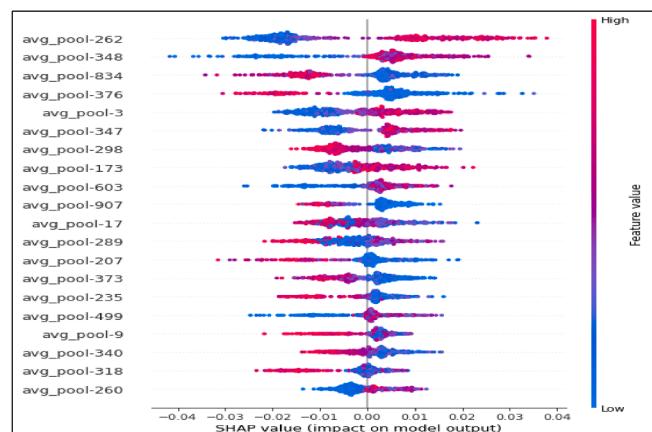


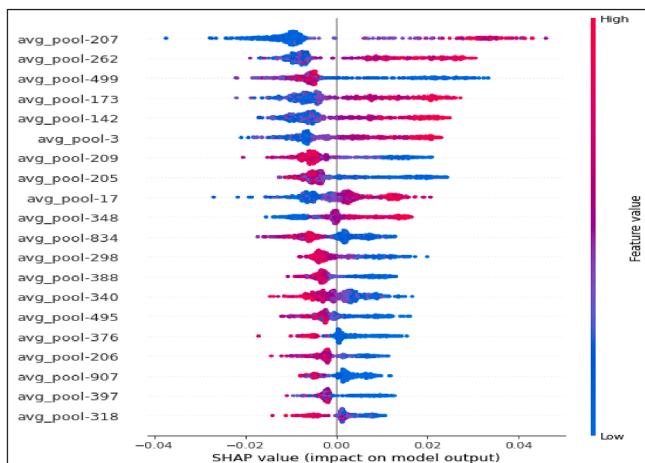
Fig. 12. ROC-AUC curves: (a) GB classifier with VGGNet-16 extracted features, (b) MLP classifier with DenseNet-121 extracted features and (c) RF classifier with DenseNet-121 extracted features.



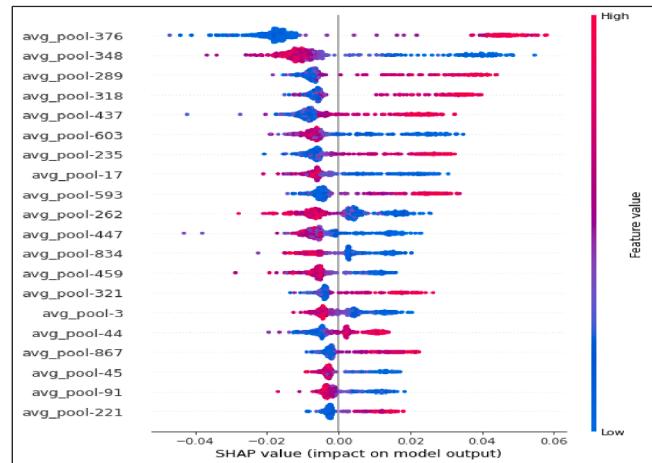
(a) Impact of feature variables on the model output



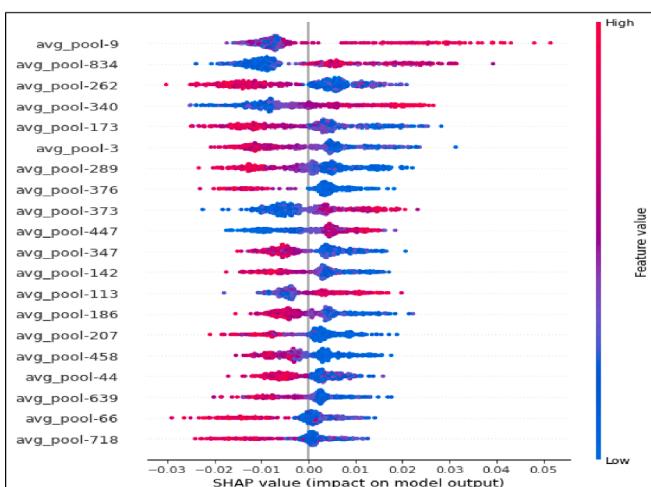
(b) Impact of features variables on Lung Adenocarcinoma Tissue classification



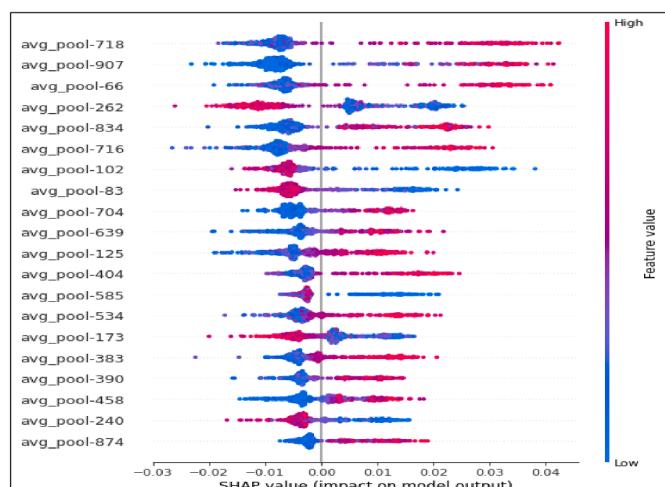
(c) Impact of features on Lung Squamous Cell Carcinoma Tissue classification



(d) Impact of features variables on Benign Lung Tissue classification



(e) Impact of features variables on Colon Adenocarcinoma Tissue classification



(f) Impact of features variables on Benign Colonic Tissue classification

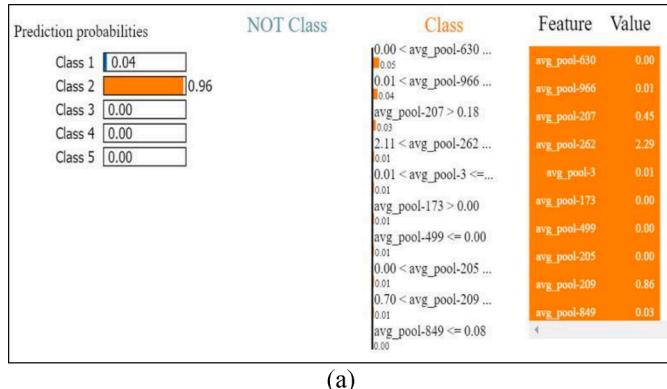
Fig. 13. Impact of features on model and Lung and Colon Classification Using SHAP Values.

higher SHAP value signifies higher impact of feature variables in classifying Lung Adenocarcinoma tissue. The feature variable avg_pool262 marks approximately identical impact on classification of Lung Squamous Cell Carcinoma Tissue and Benign Colonic Tissue as the feature attains approximately identical SHAP value. A lower SHAP value for Benign Lung tissue is attained by feature variable avg_pool262 indicating trivial significance in classifying Benign Lung tissue as compared to other classes. Similarly, the feature variable avg_pool376 achieved higher SHAP for Benign Lung tissue indicating the higher impact of the feature variable in categorizing Benign Lung tissue in contrast to other lung and colon cancer classes. The feature has the least impact in classifying Lung Squamous Cell Carcinoma Tissue as the feature attains least SHAP value for this class of tissue. Fig. 13 (b)-(f) shows the impact of feature variables on the classification of lung and colon cancer tissues. In Fig. 13 (b)-(f), the y-axis points to the input variables indicating their impact on the model. The input variables on the y-axis are arranged according to their importance. The values on the x-axis indicate SHAP values and points on the plot indicate Shapley values of input variables for the instances. The positive and negative SHAP values on x-axis indicate positive or negative impact of the feature variable on model output. The colour gradient (blue to red) indicates variable importance from low to high. The higher the SHAP value, the higher is the variable impact on the model. As shown in Fig. 13 (b), the feature variables avg_pool262, avg_pool348, avg_pool3, avg_pool347, avg_pool173, avg_pool603, avg_pool499 and avg_pool260 influence the classification of Lung Adenocarcinoma Tissue with positive correlation. i.e., the predicted values increase with increase in feature values. Changing the feature values from low (blue hue) to high (red hue) pushes the prediction higher. On contrary, the feature variables avg_pool834, avg_pool376, avg_pool298, avg_pool907, avg_pool17, avg_pool289, avg_pool207, avg_pool373, avg_pool235, avg_pool9, avg_pool340 and avg_pool318 influence the prediction of Lung Adenocarcinoma Tissue with negative correlation i.e., incrementing the feature values from low (blue hue) to high (red hue) pushes the prediction lower. As observed, the features avg_pool262 and avg_pool260 are highest and lowest impact features respectively for Lung Adenocarcinoma Tissue classification. Similarly, the impact of feature variables on Lung Squamous Cell Carcinoma Tissue, Benign Lung Tissue, Colon Adenocarcinoma Tissue and Benign Colonic Tissue are presented in Fig. 13(c)-(f).

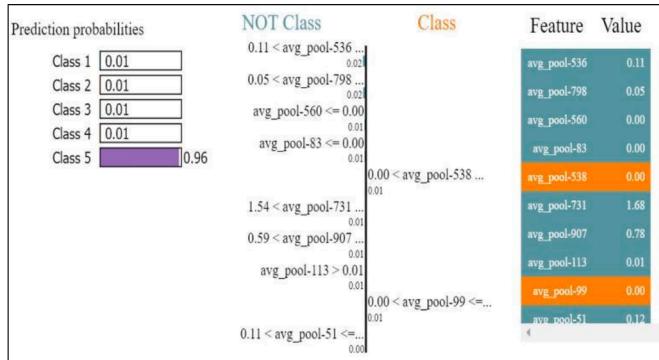
Figure 14 (a)-(b) shows the prediction probability and feature importance for two test instances using LIME. LIME can be used for explicit individual prediction scores by the classifiers. It has an advantage that LIME can be used explain the behaviour of any classifier regardless of algorithm used for forecasting the classification results [51]. LIME works locally i.e.; the LIME framework provides prediction explanations relative to individual instances or observation. LIME framework estimates relationship between feature variables and prediction by training explainable models like linear model, decision tree etc. for particular test instances [45]. Fig. 14 (a) shows the prediction

probability for class 2 (Lung Squamous Cell Carcinoma Tissue) for particular test instances. As observed from the figure, in the first column the model is 96% sure that the current test instance belongs to Lung Squamous Cell Carcinoma Tissue. However, the model also predicts the test instance as Lung Adenocarcinoma Tissue but with a very low probability of only 4%. Column 2 in Fig. 14 (a) shows the features with their relative importance to predict the results. Here we have shown 10 features for ease of explainability. Features in "Class" of column 2 shows positive correlation with prediction probability i.e., an increase in feature value increases the prediction probability. As shown, the feature avg_pool630 has highest relative importance and avg_pool849 has lowest feature importance for the specific test instance. Third column of Fig. 14 (a) shows feature variables with their actual values. Similarly, Fig. 14 (b) shows the prediction probability and feature importance for another test instance. In this, the model is 96% confident that the test instance belongs to Benign Colonic Tissue. The features in "NOT Class" of column 2 shows negative correlation with the predicted class i.e., an increase in the feature value decreases the prediction probability. The features in "Class" show positive correlation with the predicted class.

Similarly, Fig. 15 presents the prediction probability of the GB model for two test instances with a vector of concatenated handcrafted features. Fig. 15(a) presents prediction probability for class 3 (Benign Lung Tissue) test instance. For this instance, the model has 100% assured that the test instance belongs to benign lung tissue. Column 2 in Fig. 15 (a) reveals that handcrafted features CC388 (color correlated feature) and PLBP 294 (PLBP feature) in "NOT Class" show negative correlation with prediction probability of target class. An increase in feature values will lead to decrease in prediction probability value. For this particular test instance, color correlation handcrafted feature 'CC388' has highest feature importance with feature weight of 0.38. The features CC640, CC867, PLBP 741, PLBP 651, CC33, PLBP 453 and CC16 marks positive correlation with benign lung tissue. Increase in their feature values will increase the prediction probability value. Third column of Fig. 15 (a) shows feature variables with their actual values. The prediction probability of another test instance is shown in Fig. 15 (b). For this instance, the model marks 99% accuracy and classifies it as Lung Squamous Cell Carcinoma Tissue (Class 2). In this case, the handcrafted feature FCTH0 (Fuzzy Texture and color histogram feature) has highest feature importance with an assigned weight of 0.45 and the handcrafted feature PLBP486 has low significance with a weight of 0.05. The features FCTH0, CC640, CC420 and PLBP mark a positive correlation with prediction probability and show increase in prediction probability results with increase in feature values. However, the features CC867, CCT04, CC54, CC497, CC833 and PLBP show negative correlation with prediction probability and mark decrease in prediction probability with increase in feature values. So, by looking into the Figs. 13–15, the model's behaviour can be easily explainable and understandable i.e., prediction probability, feature importance and their impact on models can be better interpretable.



(a)



(b)

Fig. 14. Prediction probability and feature importance using LIME for two test instances.

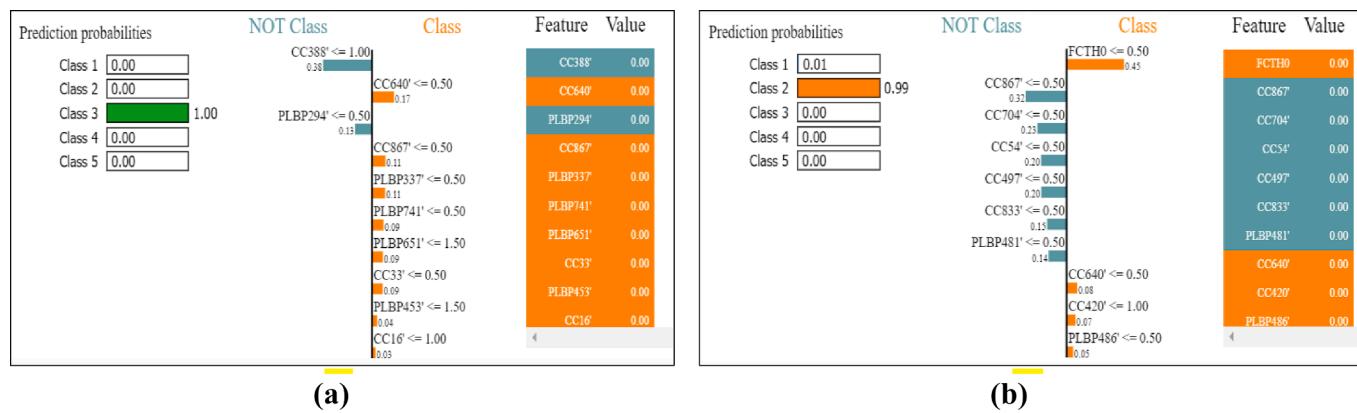


Fig. 15. GB Classifier Prediction probability and feature importance (handcrafted features) using LIME for two test instances.

5. Discussion

As observed from previous sections, the features extracted by DenseNet pre-trained network bears more significant information in comparison to handcrafted features and features extracted by other deep CNN networks. Proposed RF classifiers with DenseNet-121 extracted features show better performance in terms of accuracy, precision and recall and predict comparable results with existing state of art models. Table 6 presents the comparison of proposed models with existing state of art models. As seen from Table 6, proposed RF classifier DenseNet-121 extracted features obtained accuracy and recall of 98.60% and precision of 98.63%. Masud M et al. [52] presented CNN model with features extracted by two-dimensional Discrete Fourier transform (2D-DFT) and Single-level discrete two-dimensional wavelet transform (2D-DWT) for classifying lung and colon cancer tissue. Their proposed model obtained an accuracy of 96.33%, precision of 96.39% and recall of 96.37%. Their model obtained an MCC score of 0.954 in contrast to the MCC score of 0.985 attained by our proposed model.

Mangal S. et al. [53] presented two distinct models for lung and colon cancer tissue classification. The model's performance was evaluated in terms of accuracy only and obtained classification accuracy of 97.89% and 96.61% for lung and colon tissues respectively. However, no information about classification results in terms of precision and recall was provided. A CNN based deep learning model was presented by B. K. Hatuwal et al. [54]. Their presented model samples only lung tissues from the dataset. The proposed model was able to classify two cancerous and one benign lung tissue only and no information regarding classification of colon cancer was provided. An accuracy of 97.20%, precision and recall of 97.33% was achieved by their proposed model for lung tissue classification. KNN classifier with deep features extracted by DenseNet-121 pretrained network was proposed by D. Sarwinda et al. [55] for colon tissues. Their model samples colon tissues from the

dataset and identifies cancerous and benign colon tissues from the sampled tissues. Their model achieved an accuracy of 98.53% and recall of 98.63% for colon classification. However, their model was not able to sample lung tissues and delivers no information with reference to lung classification. Mesut Toğacar [56] proposed lung and colon cancer classification using DarkNet-19 model and Support Vector Machine (SVM) method. Features of lung and colon cancers from histopathological images were extracted from the DarkNet-19 model. Equilibrium and Manta Ray Foraging optimization algorithms were employed for selection of inefficient features. Using the complement rule of set theory, efficient features were obtained by eliminating the inefficient features. SVM classifier was employed for classification of lung and colon cancer. Their prediction model achieved excellent classification results for efficient and inefficient features. Our proposed model identifies five types of lung and colon cancer tissues (three cancerous and two benign) and marks an improvement of 2.27% in accuracy, 2.24% in precision and 2.23% in recall in comparison to the five-class classification model proposed by Masud M. et al [52].

It was also observed that features extracted by DenseNet-121 are more significant as compared to other features extracted by other CNN pre-trained networks. This is due to small linking implemented between the initial and final layer so as to improve accuracy and efficiency of the network. Herein, the feature map from each layer is concatenated instead of adding with previous ones as the input progresses through subsequent layers and hence the network provides the access of feature maps of any layer by the subsequent layers. Due to dense connection, the gradient problem is significantly reduced and enables feature reusing. The DenseNet-121 also marked an advantage of extracting the least number of significant features as compared to other CNN based deep feature extractor. The overfitting problem in DenseNet-121 is significantly reduced regularization effect of the network. However, besides the advantages, our model suffers with more training time as the RF

Table 6
Comparison of the proposed work with existing state of art methods.

Authors	Methodology	Accuracy (%)	Precision (%)	Recall (%)	Remark
Masud M. et al. [52]	Lung and Colon cancer using CNN classifier with features extracted by 2D-DFT and 2D-DWT methods.	96.33	96.39	96.37	Identify five types of lung and colon cancer tissues (three cancerous and two benign)
Mangal S. et al. [53]	Two distinct shallow CNN architectures for classification of lung cancer and colon cancer histopathological images.	Lung: 97.89 Colon: 96.61	—	—	Unable to discriminate between lung and colon tissues.
B.K.Hatuwal et al. [54]	CNN for classifying histopathological images of lung cancer tissues.	97.20	97.33	97.33	Can identify only three types of lung tissues (two cancerous and one benign tissue)
D. Sarwinda et al. [55]	DenseNet-121 with KNN classifier	98.53	—	98.63	Can identify only two types of colon tissues (one cancerous and one benign tissue)
Proposed	Lung and Colon cancer tissue classification using RF classifier with features extracted by DenseNet-121 pre-trained network	98.60	98.63	98.60	Identify five types of lung and colon tissues (three cancerous and two benign tissues)

classifier combines many decision trees to find out the target class and requires more computational power as the classifier has to build several trees to amalgamate their outputs.

6. Conclusion

In this paper, handcrafted features and deep CNN extracted features were analyzed and compared for LC 25000 lung and colon cancer histopathological image dataset. From experimental results, it was found that texture and color based handcrafted features had significant weight in contrast to edges and shape-based features to classify and detect lung and colon cancer using histopathological images. As the color and spreading of color in cancer cells are not homogenous, hence, features based on these two parameters are more significant. For handcrafted features, Color Correlogram feature extraction techniques extract more significant features in contrast to other proposed handcrafted feature extraction techniques. However, the CC feature extraction technique extracts more features in comparison to other handcrafted feature extraction techniques.

The experimental results reveal that classification performance in terms of accuracy, precision and recall can be improved with deep features extracted by CNN pre-trained networks. From experimental results, it was also found that with an approximately similar number of features, DenseNet-121 pre-trained network demonstrates an improvement of approximately 5% in accuracy, precision and recall in contrast to CC extracted features. This improvement is due to extraction of more significant deep features. In contrast to existing similar state of the art lung and colon diagnosis methods, the proposed RF classifier with DenseNet-121 extracted features show superior performance results in terms of accuracy, precision and recall. So, using transfer learning-based methods in medical imaging analysis will allow pathologists to diagnose the disease with less expertise, effort and cost.

The absence of stain normalization as a pre-processing step is the limitation of present work and we propose comparative analysis of handcrafted features extracted from stain normalization images and CNN features as future work. It would also be interesting to see the effect of dimensionality reduction and feature ranking techniques on classification of lung and colon histopathological images. More studies need to be conducted on histopathological images of this new LC 25000 dataset.

Funding

There is no funding source.

Availability of data

Dataset is available at Kaggle [20]. Available online: <https://www.kaggle.com/andrewmvd/lung-and-colon-cancer-histopathological-images>

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

CRediT authorship contribution statement

Naresh Kumar: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Manoj Sharma:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Vijay Pal Singh:** Validation, Visualization, Investigation, Methodology, Writing – review & editing. **Charanjeet Madan:** Formal analysis, Investigation, Methodology, Writing – review & editing. **Seema Mehandia:** Conceptualization, Investigation, Methodology,

Visualization, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bspc.2022.103596>.

References

- [1] K. Kurishima, K. Miyazaki, H. Watanabe, T. Shiozawa, H. Ishikawa, H. Satoh, N. Hizawa, Lung cancer patients with synchronous colon cancer, *Mol. Clin. Oncol.* 8 (1) (2018) 137–140, <https://doi.org/10.3892/mco.2017.1471>.
- [2] J. Ferlay, M. Ervik, F. Lam, M. Colombet, L. Mery, M. Piñeros, et al. Global Cancer Observatory: Cancer Today. Lyon: International Agency for Research on Cancer; 2020 (<https://gco.iarc.fr/today>, accessed June 2021).
- [3] WHO Cancer <https://www.who.int/news-room/fact-sheets/detail/cancer> (Last accessed June 2021).
- [4] J.R. Molina, P. Yang, S.D. Cassivi, S.E. Schild, A.A. Adjei, Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship, *Mayo Clin. Proc.* 83 (5) (2008) 584–594, <https://doi.org/10.4065/83.5.584>.
- [5] S.A. El-Regaily, M.A. Salem, M.H. Abdel Aziz, M.I. Roushdy, Survey of computer aided detection systems for lung cancer in computed tomography, *Curr. Med. Imaging Rev.* 14 (1) (2017) 3–18.
- [6] S.-H. Wang, Y.-D. Zhang, DenseNet-201-based deep neural network with composite learning factor and precomputation for multiple sclerosis classification, *ACM Trans. Multimedia Comput. Commun. Appl.* 16 (2s) (2020) 1–19.
- [7] Y.D. Zhang, S.C. Satapathy, X. Zhang, et al., COVID-19 diagnosis via DenseNet and optimization of transfer learning setting, *Cogn. Comput.* (2021), <https://doi.org/10.1007/s12559-020-09776-8>.
- [8] L. Alzubaidi, J. Zhang, A.J. Humaidi, A. Al-Dujaili, Y.e. Duan, O. Al-Shamma, J. Santamaría, M.A. Fadhel, M. Al-Amidie, L. Farhan, Review of deep learning: concepts, CNN architectures, challenges, applications, future directions, *J. Big Data* 8 (1) (2021), <https://doi.org/10.1186/s40537-021-00444-8>.
- [9] A.A. Borkowski, M.M. Bui, L.B. Thomas, C.P. Wilson, L.A. DeLand, S.M. Mastorides, Lung and colon cancer histopathological image dataset (LC25000), ArXiv Prepr. ArXiv1912.12142 (2019) 1–2, <https://arxiv.org/abs/1912.12142v1>.
- [10] Satvik Garg and Somya Garg. 2020. Prediction of lung and colon cancer through analysis of histopathological images by utilizing Pre-trained CNN models with visualization of class activation and saliency maps. 2020 3rd Artificial Intelligence and Cloud Computing Conference. Association for Computing Machinery, New York, NY, USA, 38–45. doi: 10.1145/3442536.3442543.
- [11] A. Teramoto, T. Tsukamoto, Y. Kiriyama, H. Fujita, Automated classification of lung cancer types from cytological images using deep convolutional neural networks, *Biomed Res. Int.* 2017 (2017) 1–6, <https://doi.org/10.1155/2017/4067832>.
- [12] S. Wang, D.M. Yang, R. Rong, X. Zhan, J. Fujimoto, H. Liu, J. Minna, I.I. Wistuba, Y. Xie, G. Xiao, Artificial intelligence in lung cancer pathology image analysis, *Cancers* 11 (11) (2019) 1673, <https://doi.org/10.3390/cancers1111673>.
- [13] S. Mary, K.J. Hewitt, R. Nasir, Deep learning with sampling in colon cancer histology, *Front. Bioeng. Biotechnol.* 7 (2019) 52, <https://doi.org/10.3389/fbioe.2019.00052>.
- [14] N. Ibrahim, N.K.C. Pratiwi, M.A. Pramudito, F.F. Taliningsih, (2021) Non-Complex CNN Models for Colorectal Cancer (CRC) Classification Based on Histological Images. In: Triwiyanti, Nugroho H.A., Rizal A., Caesarendra W. (eds) Proceedings of the 1st International Conference on Electronics, Biomedical Engineering, and Health Informatics. Lecture Notes in Electrical Engineering, vol 746. Springer, Singapore. .
- [15] Y. Wang, S. Liu, Z. Wang, Y. Fan, J. Huang, L. Huang, Z. Li, X. Li, M. Jin, Q. Yu, F. Zhou, A machine learning-based investigation of gender-specific prognosis of lung cancers, *Medicina* 57 (2) (2021) 99, <https://doi.org/10.3390/medicina57020099>.
- [16] S. Bhatia, Y. Sinha, L. Goel, Lung cancer detection: a deep learning approach, *Soft Computing for Problem Solving, Advances in Intelligent Systems and Computing* 817 (2019), https://doi.org/10.1007/978-981-13-1595-4_55.
- [17] J.u. Han, K.-K. Ma, Fuzzy color histogram and its use in color image retrieval, *IEEE Trans. Image Process.* 11 (8) (Aug. 2002) 944–952, <https://doi.org/10.1109/TIP.2002.801585>.
- [18] P. Nanglia, S. Kumar, A.N. Mahajan, P. Singh, D. Rathee, A hybrid algorithm for lung cancer classification using SVM and neural networks, *ICT Express.* 7 (3) (2021) 335–341.
- [19] A. Bhattacharjee, S. Majumder (2019), Automated computer-aided lung cancer detection system, in: *Advances in Communication, Devices and Networking*, Springer, Singapore, 2019, pp. 425–433.
- [20] Kaggle. <https://www.kaggle.com/andrewmvd/lung-and-colon-cancer-histopathological-images>.

- [21] S.A. Chatzichristofis, Y.S. Boutalis, "FCTH: Fuzzy Color and Texture Histogram - A Low Level Feature for Accurate Image Retrieval," 2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services, 2008, pp. 191-196, doi: 10.1109/WIAMIIS.2008.24.
- [22] Y.-E. An, J.-A. Park, Color correlogram using combined RGB and HSV color spaces for image retrieval, *J. Korean Inst. Commun. Information Sci.* 32 (5C) (2007).
- [23] H.A. Jalab, "Image retrieval system based on color layout descriptor and Gabor filters," 2011 IEEE Conference on Open Systems, 2011, pp. 32-36, doi: 10.1109/ICOS.2011.6079266.
- [24] Chee Sun Won, Dong Kwon Park, and Soo-Jun Park, "Efficient Use of MPEG-7 Edge Histogram Descriptor", ETRI Journal, Volume 24, Number 1, February 2002, pp. 23-30.
- [25] Won C.S. (2004) Feature Extraction and Evaluation Using Edge Histogram Descriptor in MPEG-7. In: Aizawa K., Nakamura Y., Satoh S. (eds) Advances in Multimedia Information Processing - PCM 2004. PCM 2004. Lecture Notes in Computer Science, vol 3333. Springer, Berlin, Heidelberg. doi: 10.1007/978-3-540-30543-9_73.
- [26] A. Chauhan, D. Chauhan, C. Rout, H.A. Kestler, Role of gist and PHOG features in computer-aided diagnosis of tuberculosis without segmentation, *PLoS ONE* 9 (11) (2014) e112980, <https://doi.org/10.1371/journal.pone.0112980>.
- [27] A. Bosch, A. Zisserman, X. Munoz, Representing shape with a spatial pyramid kernel, *Image Process* 5 (2007) 401–408.
- [28] N.J. Sairamya, L. Susmitha, S. Thomas George, M.S.P. Subathra, Chapter 12 - Hybrid Approach for Classification of Electroencephalographic Signals Using Time-Frequency Images with Wavelets and Texture Features, Editor(s): D. Jude Hemanth, Deepak Gupta, Valentina Emilia Balas, In Intelligent Data-Centric Systems, Intelligent Data Analysis for Biomedical Applications, Academic Press, 2019, Pages 253-273, ISBN 9780128155530, .
- [29] X. Qian, X.-S. Hua, P. Chen, L. Ke, PLBP: An effective local binary patterns texture descriptor with pyramid representation, *Pattern Recogn.* 44 (10-11) (2011) 2502–2515.
- [30] J. Schmidhuber, Deep Learning in neural networks: an overview, *Neural Netw.* 61 (2015) 85–117.
- [31] Z.C. Horn, L. Auret, J.T. McCoy, C. Aldrich, B.M. Herbst, Performance of convolutional neural networks for feature extraction in froth flotation sensing, *IFAC-PapersOnLine* 50 (2) (2017) 13–18, <https://doi.org/10.1016/j.ifacol.2017.12.003>.
- [32] S. Sharma, R. Mehra, Conventional machine learning and deep learning approach for multi-classification of breast cancer histopathology images—a comparative insight, *J. Digit. Imaging* 33 (3) (2020) 632–654, <https://doi.org/10.1007/s10278-019-00307-y>.
- [33] E. Ribeiro, A. Uhl, G. Wimmer, M. Häfner, Exploring deep learning and transfer learning for colonic polyp classification, *Comput. Math. Methods Med.* 2016 (2016) 1–16, <https://doi.org/10.1155/2016/6584725>.
- [34] S.J. Pan, Q. Yang, A survey on transfer learning, *IEEE Trans. Knowl. Data Eng.* 22 (10) (2010) 1345–1359.
- [35] K. Weiss, T.M. Khoshgoftaar, D. Wang, A survey of transfer learning, *J. Big Data* 3 (9) (2016), <https://doi.org/10.1186/s40537-016-0043-6>.
- [36] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger (2017), Densely Connected Convolutional Networks, in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, pp. 2261-2269. doi: 10.1109/CVPR.2017.243.
- [37] F.N. Iandola, S. Han, M.W. Moskewicz, K. Ashraf, W.J. Dally, K. Keutzer, 2016. SqueezeNet: Alexnet-level accuracy with 50 × fewer parameters and < 0.5MB model size. arXiv: 1602.07360.
- [38] K. Simonyan, A. Zisserman, (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
- [39] François Chollet (2017), Xception: Deep Learning with Depthwise Separable Convolutions, arXiv:1610.02357.
- [40] K. He, X. Zhang, S. Ren, J. Sun (2016), "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), pp. 770–778.
- [41] K. He, X. Zhang, S. Ren, J. Sun (2016), "Identity mappings in deep residual networks," in European conference on computer vision (ECCV). Springer, 2016, pp. 630–645.
- [42] D.M.W. Powers, Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation, *J. Mach. Learn. Technol.* 2 (1) (2011) 37–63.
- [43] C. Rudin, Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead, *Nature Machine Intelligence* 1 (5) (2019) 206–215, <https://doi.org/10.1038/s42256-019-0048-x>.
- [44] E. Tjoa, C. Guan, A survey on explainable artificial intelligence (XAI): toward medical XAI, *IEEE Trans. Neural. Networks and Learning Syst.* 32 (11) (2021) 4793–4813.
- [45] P. Linardatos, V. Papastefanopoulos, S. Kotsiantis, Explainable AI: A Review of Machine Learning Interpretability Methods. *Entropy* 2021, 23, 18. <https://dx.doi.org/10.3390/e23010018>.
- [46] V. Petsiuk, A. Das, K. Saenko, RISE: Randomized Input Sampling for Explanation of Black-box Models. In Proceedings of the British Machine Vision Conference 2018, BMVC 2018, Newcastle, UK, 3–6 September 2018; BMVA Press: Durham, UK, 2018; p. 151.
- [47] G. Montavon, S. Lapuschkin, A. Binder, W. Samek, K.R. Müller, Explaining nonlinear classification decisions with deep taylor decomposition, *Pattern Recognit.* 65 (2017) 211–222.
- [48] W.J. Murdoch, C. Singh, K. Kumbier, R. Abbasi-Asl, B. Yu, Definitions, methods, and applications in interpretable machine learning, *Proc. Natl. Acad. Sci. USA* 116 (2019) 22071–22080.
- [49] A. Gramegna, P. Giudici, SHAP and LIME: an evaluation of discriminative power in credit risk, *Front. Artif. Intell.* 4 (2021), 752558, <https://doi.org/10.3389/frai.2021.752558>.
- [50] S.M. Lundberg, S.I. Lee, (2017). A Unified Approach to Interpreting Model Predictions. *Adv. Neural Inf. Process. Syst.* 30, 4765–4774. Available at: <https://arxiv.org/abs/1705.07874>.
- [51] M.T. Ribeiro, S. Singh, C. Guestrin, Why should i trust you? *Knowledge Discov. Databases* 16 (2016) 1135–1144, <https://doi.org/10.1145/2939672.2939778>.
- [52] M. Masud, N. Sikder, A.-A. Nahid, A.K. Bairagi, M.A. AlZain, A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework, *Sensors* 21 (3) (2021) 748, <https://doi.org/10.3390/s21030748>.
- [53] S. Mangal, A. Chaurasia, A. Khajanchi, Convolution neural networks for diagnosing colon and lung cancer histopathological images. arXiv 2020, arXiv:2009.03878.
- [54] B.K. Hatuwal, H.C. Thapa, Lung cancer detection using convolutional neural network on histopathological images, *Int. J. Comput. Trends Technol.* 68 (10) (2020) 21–24.
- [55] D. Sarwinda, A. Bustamam, R.H. Paradisa, T. Argyadiva, W. Mangunwardoyo, Analysis of Deep Feature Extraction for Colorectal Cancer Detection, 2020 4th International Conference on Informatics and Computational Sciences (ICICoS), 2020, pp. 1-5, doi: 10.1109/ICICoSS1170.2020.9298990.
- [56] M. Toğacıçar, Disease type detection in lung and colon cancer images using the complement approach of inefficient sets, *Comput. Biol. Med.* 137 (2021) (2021), 104827, <https://doi.org/10.1016/j.combiomed.2021.104827>.