Detection of Lung and Colon Cancer Using Transfer Learning

Deepali Sharma[1], Dilip Kumar Choubey[1,*], Neeraj Kumar Yadav[1], and Aditya Kumar Pandey[1]

ARTICLE INFO

Article history:
Received: 16 June 2024
Revised: 17 September 2024
Accepted: 3 October 2024
Online: 15 May 2026

Keywords:
Transfer Learning
VGG-16
ResNet-50
MobileNet_V2
Lung and Colon Cancer
HSCT

ABSTRACT

Lung and colon cancers rank among the most prevalent cancers worldwide and are responsible for a significant number of deaths. Individuals must be treated at an early stage and this helps recovery. To diagnose the patients, the medical professionals have been utilizing microscopic pathology images of the lung and colon tissue biopsies of the potentially infected areas of the body. Majority of these cancers are slow to detect and usually necessitate a lot of testing procedures where traditional techniques are used. The study primarily addresses colon adenocarcinoma and colon benign tissue, along with three lung tissue types—squamous cell carcinoma, adenocarcinoma, and benign tissue. This manuscript employs transfer learning, specifically VGG-16, MobileNet_V2, and ResNet-50 classification methods to identify lung and colon cancer. Out of all the implemented techniques, ResNet-50 demonstrated superior performance, achieving 99% and 98.8% accuracy in both training and testing subset of dataset. The future work will focus on identifying the biological components that affect human lung and colon cancers and use these factors as indicators for cancer detection.

1. INTRODUCTION

Cancer is one of the most life-threatening diseases which can be defined as uncontrolled proliferation of a typical cells capable of in filtrating nearby tissues and spreading to distant organs. Among its various forms, lung cancer is consistently ranked among the deadliest forms of cancer which contributes almost 70-71% of global cancer related fatalities [1]. It ranks as the second leading malignancy in females after breast cancer and correlates with less survival rate of approximately 19%. Another type of lethal cancer is Colorectal cancer (CRC), it also ranks third among the deadliest varieties of cancer. Its severity is largely dependent on tumor invasion of bowel wall, regional nodal spread, and distant metastasis.

Diagnosis of these cancers typically involves a range of imaging and clinical techniques. Some of the medical diagnosis methods which are widely used in traditional methods are Computed Tomography (CT) scan, X-Rays, bronchoscopy, PET-CT and biopsy [2]. To find the severity and type of lung cancer doctors often examine tissue samples using a common test known as Hematoxylin and Eosin (H&E) staining [3]. The detection of colorectal cancer relies on imaging and laboratory techniques such as colonoscopy, CT-colonography, and blood analysis [4][5]. Since treatment strategies are strongly influenced by histological type, molecular profile, and disease stage, swift and precise classification of histopathology images is essential to enhancing clinical prognosis [6].

Over the past decade, sophisticated AI methods like deep learning (DL) have appeared as a transformative tool in digital pathology, lessening the diagnostic workload for pathologists and enhancing classification precision [7]. Deep learning models called convolutional neural networks are very good at recognizing cancer from tissue images across multiple types, like skin, brain, lung, and colorectal cancers [8]. This work utilizes three commonly applied architectures, namely VGG16, ResNet50, and MobileNetV2. These models are utilized for extraction of features and analysis of lung and colon cancer images because they can provide balance between depth, efficiency and computational cost which is suitable for this experiment.

Generative Artificial Intelligence (GenAI) can be defined as an advanced computational model which is capable of creating new content or insights by learning from vast number of datasets. It enables systems to generate human-like text, summarize complex information, and provide decision-support using Natural Language Processing (NLP). Within this field, GenAI plays a transformative role by assisting with a range of tasks such as clinical documentation, offering patient-specific education, enhancing diagnostic interpretation, and supporting treatment planning. Its ability to synthesize vast amounts of biomedical data into accessible knowledge has the capacity

[1]*Department of Computer Science and Engineering, Indian Institute of Information Technology Bhagalpur, Bihar, India.*
**Corresponding author:* Dilip Kumar Choubey; Email: dilipchoubey_1988@yahoo.in, dkchoubey.cse@iiitbh.ac.in.

to lessen the workload of healthcare professionals while improving the accuracy and efficacy of patient care. Generative AI can be an exceptionally useful tool in combating lung cancer through the provision of tailored education and the strengthening of patient-provider trust.

This work seeks to improve the precision of diagnosis as well as to support doctors and nursing staffs in determining cancer subtypes more reliably. Such automated approaches can help to accelerate clinical workflows and ensure timely treatment decisions so that it contribute to enhancing survival outcomes in lung and colorectal cancers.

Doctors use multiple approaches to diagnose lung and colon cancer, including medical imaging, tissue sampling, and specialized scans like PET-CT and bronchoscopic procedures. In biopsy, a small portion of tissue from the lung or colon is collected and studied under a microscope to check for abnormalities. It is important to examine the tissue closely to understand the cancer's characteristics. However, analyzing this manually can take a long time and may be influenced by the person doing it, which may result in varying results. To accelerate diagnosis and also ease the burden on pathologists, we use a technique called transfer learning. Transfer learning has become a key method in digital pathology because it helps integrate and improve workflows. It allows for better and faster analysis by adapting existing methods to new tasks. These tools also help doctors decide on the best treatment by identifying the type and stage of the cancer. This study examines three deep learning models that use previously trained convolutional networks, including ResNet-50, VGG16, and MobileNet_V2, to analyze medical images

This paper proceeds as follows to present the research findings. Section 2 provides a review of current literature in the field. The research design, outlining the methodologies and approaches, is presented in Section 3. Section 4 delivers the study's results alongside a detailed analysis. Section 5 concludes by discussing the implications of these outcomes and identifying directions for future research.

2. LITERATURE REVIEW

Deep learning is a widely used method for detecting and categorizing things. Many researchers apply various deep learning architectures on lung data sets to train and test their models. Previous research has explored various approaches to lung and colon cancer classification. One such study by Masud et al. (2021) [9] introduced a model using the LC25000 dataset. They employed the 2D Fourier transformation and wavelet decomposition to analyze the images in a manner that separates the channels. The model based on CNN obtained its highest accuracy at 96.33%. Several feature extraction methods like VGG16, ResNet50, and InceptionV3 were used by Garg & Garg (2021) [10] to build eight CNN models to recognize colon and lung cancers from the LC25000 histopathological data. The models demonstrated exceptional performance, yielding accuracy levels between 96% and 100%. In 2020, Abdul [11] introduced a CNN-based approach using CT scans from the LIDC-IDRI dataset to distinguish between cancerous and non-cancerous lung tumors. They used expert marking to separate nodules. An accuracy of 97.2% was attained by the model. Subramanian et al. (2020) [12] made use of AlexNet and SoftMax activation functions to detect lung cancer. The model was effective in analyzing CT lung images, attaining an accuracy of 99.52%. Sajja et al. (2019) [13] used AlexNet, Google Net, and ResNet50, pre-trained models to conduct a performance comparison on CT images of the LIDC dataset. They developed a CNN model featuring a sparsified dense layer and 60% dropout to lower computational costs and prevent overfitting. This approach outperformed other approaches, achieving 99.03% testing accuracy and 100% validation accuracy.

A hybrid model integrating CNN and Gradient Boosting via the LightGBM algorithm was presented by Elraouf et al. (2023) [14] for classifying lung cancer. The approach, evaluated on the LC25000 dataset, yielded outstanding accuracy and sensitivity values of 99.6%. Anjum et al. (2023) [15] tested different variants of Efficient Net model for classifying lung colon malignancies based on histopathological scans from the LC25,000 datasets. The model aimed to adjust the CNN in terms of depth, resolution, and width based on available resources which resulted in accuracy of 97.24%. In their 2023 study, Wang et al. [16] explored if a pre-trained ResNet performs better when all layers are fine-tuned or when only the last layer is trained while keeping the rest fixed. They used a histopathological dataset of lymphatic tissue and tested various ResNet architectures (e.g., ResNet-34, -18, -152, and -50) with the key finding that ResNet-152 architecture had the highest AUC equalling 0.989. They proved that fine-tuning is better than training only the last layer. Pang et al. (2020) [17] carried out experiments using CT scans from Shandong Provincial Hospital to classify different categories of lung cancer. They used the VGG16-T model with boosting and got better results than VGG16-T without boosting, ResNet-34, AlexNet, and DenseNet. Their model had an accuracy of 86.58%. Hutuwal and Thapa (2020) [18] proposed a convolutional neural network utilizing cross-entropy as the objective function. They used histopathological scans from the LC25000 biopsied tissue dataset. The model's performance was robust across both datasets, recording an accuracy of 96.11% on the training set and 97.2% on the validation set. In 2023, Tummala et al. [19] presented a model leveraging progressive learning and integrated compound scaling known as the EfficientNetV2 family, which features mid-sized, small, and large iterations. They also highlighted key areas in GradCAM's target class prediction. They used the EfficientNetV2-L, -M, and -S models to sort the LC25000 histopathology dataset into five groups of colon as well as lung cancer. On the test dataset, they achieved over 99.97% accuracy.

Geng et al. (2019) [20] combined lung segmentation techniques using the VGG-16 network and dilated convolution on CT images. They worked in three steps: first, they used the first three components of VGG-16 for image pooling and convolution. Next, they used multi-scale convolutions, combined features from the multi-scale convolutional layers, and used MLP for pixel prediction during segmenting the parenchymal region. The experiment achieved a 0.9867 Dice similarity coefficient. Saric et al. (2019) [21] performed classification of lung cancer and normal by training two CNN models, VGG and ResNet, to categorize lung cancer and normal. They analyzed a full slide containing histopathological images stained using hematoxylin and eosin (H&E), and they used a microscope to automatically magnify them 20 times. They checked the results based on how well each patch was classified. The ResNet-50 model had an accuracy of 0.7205, while VGG-16 had 0.7541. (Zhou et al., 2020) [22] introduced DenseNet-NSCR, which is a method that uses non-negative sparse representation for classification. The model achieved 99.10% accuracy in separating lung tumors into cancerous and non-cancerous categories. A comparison of this model was made against AlexNet, GoogleNet, and DenseNet using SVM, SRC, or NSCR. (Naseer et al. (2023) [23] introduced a seven-layer LungNet-SVM architecture comprising convolutional, pooling, and fully connected layers for effective feature extraction. They utilized a SVM classifier to differentiate malignant from benign lung cancer cases, which resulted in a notable accuracy of 97.64%. Learning of the model was performed on CT scans sourced from the dataset. Fang (2018) [24] applied median intensity projection (MIP) on fine-tuned GoogleNet models using CT scans from the LIDC-IDRI dataset. This model's performance was benchmarked against a fine-tuned AlexNet with MIP. They trained their model over 300 epochs and achieved accuracy rate of 81%. (Shen et al., 2016) [25] created a domain-adaptation framework leveraging transfer learning for deep features to predict cancerous cells at the patient level. The model reached 70.69% accuracy and recorded an ROC-AUC score of 0.66. The study by Ausawalaithong et al. (2018) [26] employed a DenseNet-121 model enhanced with transfer learning for lung cancer detection on JSRT chest X-ray images. Evaluation results indicated that the model reached an average accuracy of 74.43 ± 6.01%, alongside specificity of 74.96 ± 9.85% and sensitivity of 74.68 ± 15.33%. (Al-Yasriy et al., 2020) [27] used the AlexNet architecture to classify lung cases as healthy, non-cancerous, or cancerous. They aimed to improve the model's performance in detecting and classifying lung cancer cases. Using CT images from the IQ-OTH/NCCD dataset, the model reached an impressive accuracy rate of 93.55%. Teramoto et al. (2017) [28] utilized a DCNN to classify cytological images into multiple cancer subtypes, achieving a classification accuracy of 71%. The technique was useful for classifying microscopic images. (Shimazaki et al., 2022) [29] developed a deep neural network-based algorithm to identify cancer visible on chest radiographs, demonstrating a low low mFPI and a 0.52 average Dice coefficient. They validated the model using 5-fold cross entropy. Fan et al. (2017) [30] developed a 3D CNN combined with traditional image processing techniques to identify lung nodules in CT scans. They converted grayscale images to RGB. The model demonstrated an accuracy of 67.7%. Similarly, (Jangir et al., 2021) [35] used a functional link CNN to categorize real, localized diabetes datasets. They assessed their method using a non-parametric Friedman statistical test. Neural networks have been applied by researchers, fuzzy decision trees, and SVMs based on kernel functions to classify diabetes. For selecting features, they employed GA [36, 38, 39] and PSO-SVM [37], integrating these with earlier classification methods for better results. (Choubey et al., 2016) [40] introduced a hybrid AI-based model for identifying diabetes. (Choubey et al., 2020) [41] explored diabetes classification algorithms. (Choubey et al., 2020) [42] compared classification methods with and without attribute selection techniques. (Choubey et al., 2020) [43] evaluated classification method performance using PCA and PSO for diabetes. Inspired by these works, the idea emerged to study HSCT-related mortality prediction. (Sharma et al., 2020) [44] discussed soft computing techniques for breast cancer diagnosis. Researchers [45] compared leukemia classification using ML and data mining. Additionally, researchers [46] used soft computing to predict liver outcomes. (Choubey and Paul, 2016) [47], (Srivastava et al., 2021) [48] conducted a thorough comparison of ML, DL, soft computing, and knowledge discovery methods for diabetes and heart disease. Their discussion covered performance, benefits, and challenges of these techniques. (Verma et al., 2023) [50] proposed CNN+LSTM for classifying hate speech and offensive language, outperforming existing methods. (Bhardwaj et al., 2024) [51] developed a computational classifier on protein data, where Random Forest outperformed previous approaches. In their 2023 and 2024 studies, Choubey et al. [49, 33] investigated the application of computational intelligence, knowledge discovery, and automated learning methods in the context of identifying cases of ovarian cancer and dengue fever.

The Literature Review is also presented in a detailed table, including the techniques used, their advantages, purposes, performance, and future work.

Table 1. Discussion of existing works for lung and colon cancer

Ref. No	Dataset	Techniques	Purpose	Advantages	Performance	Future Work
[9]	LC25000	CNN model using 2D Fourier transformation and wavelet decomposition.	Identification of the various types of cancer will enable the pathologists in the medical centre detect more incidences of colon as well as lung cancer with less efforts, finances, and time.	When making decisions, the model is less inclined towards bias in favour of one class over another.	96.33% peak classification accuracy	Put some architectural work and Engineering of new sets of features from image dataset to increase the model performance.
[10]	LC25000	VGG16, NASNetMobile, InceptionResNetV2, InceptionV3, Xception, ResNet50, DenseNet169, and MobileNet are eight distinct pre-trained CNN models, Image Augmentation, GradCam and SmoothGrad.	To apply improved augmentation techniques to histopathological images in order to detect lung and colon cancer with CNN-based pre-trained model.	Attention images from pre-trained CNN models are pictured using GradCAM and SmoothGrad in order to examine saliency maps and class activation, respectively.	Accuracy achieved between 96% and 100%.	Classify subtypes of cancer with intersection of data and apply multi-class classification of various sub-types of cancer.
[11]	LIDC and IDRI-CT images,	CNN	The model is designed to identify and categorise lung tumours as benign or malignant.	The system performed better in terms of accuracy	96.1% specificity, 95.6% sensitivity, and 97.2% accuracy.	…….
[12]	The dataset was collected from a number of sources, including hospitals and open-source software.	AlexNet + SVM, AlexNet + DeepkNN, AlexNet Model + SoftMax, Image augmentation	A python-tkinter user interface has been developed for easy use by the general public.	Image segmentation is used to avoid overfitting problem	98.62%, 97.75%, 99.52% accuracy respectively.	Identify more influential factors for lung cancer detection. Design a IoT application to extract physiological parameters. Using fog-assisted cloud computing, classification and analysis can be completed effectively.
[13]	LIDC and IDRI-CT images, Initiated by NCI.	CNN model, sparse network, SoftMax activation function, AlexNet, GoogleNet, ResNet50	To categorise the lung CT scans in the malignant and benign tissue.	Network is sparsified and Use of dropout layer reduces the overfitting and reduce computing cost.	100% validation accuracy, 99.03% training accuracy, 100% sensitivity.	Test the model with and without dropout at various dropout ratios. Find the importance of inception layer and number of inception layers adequate to attain improved performance.
[14]	LC25000	CNN-Light gradient boosting (LightGBM) model, feature	To build computer based	Takes only three seconds to categories photos and extract	99.6% accuracy.	Boost performance with novel techniques like

		extraction	Identification method	characteristics. Can also be utilised for breast, prostate, and throat cancer detection		Natural Gradient Boosting and Categorical Boosting,
[15]	LC25000	SoftMax's parameter tuning techniques, Transfer Learning, and EfficientNet variations EfficientNetB0-B7	To identify five distinct types of lung, and colon cancer.	Seeks to methodically upgrade a CNN model in each of the three dimensions—depth, resolution, and width—based on easily accessible resources.	Accuracy of proposed model (EfcientNetB2) - 97.24%	To experiment with adding more photos of different sizes in order to widen the piece. To increase the total number of categories that are compatible to the data availability, potentially enhancing the model of the accuracy.
[16]	Whole slide images on Kaggle Histopathological Cancer Detection HCD public dataset	Resnet18, 34,50, 152, Fine tuning	The richness of the dataset's semantic features should determine performance rather than the depth of the layers, and choosing a model should also take the time and resources needed for training into account.	Outperforms traditional machine learning models or models trained from scratch in terms of overall accuracy.	AUC value of ResNet-18, 34, 50, 152 is 0.9914, 0.9922, 0.9921, 0.9888.	To utilize multiple model configurations and a range of fine-tuning methods, while also conducting further data augmentation to produce more training samples.
[17]	CT images of Shandong Provincial Hospital	VGG16-T with boosting	To utilize CT images to determine the type of lung cancer.	Training multiple weak classifiers significantly improves the performance of classification	Accuracy 86.58%, Sensitivity 86%, Specificity 72.63%.	**…….**
[18]	LC25000	CNN model, SoftMax function	To detect the kinds of lung cancer	……	97.11%	……..
[19]	LC25000	EfficientNetV2 -Small, -medium, -Large, Compound Scaling, gradCAM,	The approach facilitates the complete automation of interpreting and subtyping colon and lung cancer, based on histopathological images.	GradCAM was used to highlight the regions where more attention should be given during model prediction.	EfficientNetV2 -L: Accuracy 99.97%, AUC: 99.99%	To suggest lightweight designs that could be implemented on edge devices for the same task.
[20]	……	VGG16, Dilated Convolution, Feature Fusion, Lung Segmentation method	To demonstrate the lung parenchymal area's effective segmentation using this method.	Feature fuse technique increased the robustness of the algorithm and get accurate outcomes.	99.23%	……
[21]	ACDC@LUNGHP	ResNet-50, VGG16, SGD optimizer, Patch Level Image Classification	The pathologist would benefit and the entire process would be greatly accelerated by automated cancer region detection.	Assist specialists in the lung cancer identification.	Resnet50 accuracy: 0.7205, VGG16 accuracy: 0.7541	Use stain normalisation and picture augmentation to expand the size of the training set. Create a fresh CNN model from

						the ground up.
[22]	CT images	NCR-DenseNet, AlexNet, GoogleNet	AlexNet, GoogleNet, and DenseNet-201 models have been used in comparison experiments to show that the proposed method has improved robustness and generalisation ability.	Solve problems of data redundancy, increase in dimensionality on feature extraction	Accuracy 99.10%	…….
[23]	CT images of LUNA16 dataset	Modified AlexNet architecture, SVM, LungNet-SVM	LungNet-SVM for segmenting and classifying lung nodules from CT scans as benign or malignant	The experimental findings indicate that the proposed LungNet-SVM model was strikingly good with regard to precision on a LUNA16 data.	Accuracy: 97.64%	To improve lung cancer detection performance, the suggested method can be utilized to additional publicly accessible datasets using different CNN architectures and optimizers.
[24]	LIDC and IDRI-CT images, Initiated by NCI.	Transfer Learning, GoogleNet, Median Intensity Projections, Fine Tuning	To categorize the images of lung cancer as malignant or normal.	The accuracy, sensitivity, and specificity metrics for the proposed system was greater than most lung cancer detection that are available today.	Accuracy: 81%	This channel can be used to gather testing data for system adjustment.
[25]	LIDC and IDRI-CT images, Initiated by NCI.	CNN model, CNN-MIL	To decrease the need for pathologically validated data, potentially enabling the use of CT imaging datasets from multiple sources to improve cancer diagnosis	Created the domain transfer model based on the observation that, with widespread CT screening, the Discovery Set is reasonably accessible at an early stage of diagnosis.	Accuracy: 70.69%	Plans to increase the size of its multimodal image sets in order to enhance its predictive diagnostic capabilities.
[26]	Chest X-ray14 dataset, JSRT (Japanese Society of Radiological Technology) dataset	Deep Learning, DenseNet-121	This work demonstrates the potential to enhance the application of deep learning for detecting lung carcinoma using chest CT scans	Solves the small dataset problem.	Accuracy: 74.43 ± 6.01%.	An additional method of data augmentation could be to crop the image and add Gaussian Noise at random.

3. PROPOSED METHODOLOGY

The proposed approach involves around **f**ive key stages, described in detail below:

3.1 Dataset overview

This experiment employed the LC25000 dataset. It has been sourced from GitHub [31]. It contains 15,000 lung and 10,000 colon photos, evenly split into five groups: lung squamous cell cancer, lung adenocarcinoma, lung benign, colon adenocarcinoma and colon benign tissue. The exact data set has been implemented in [32]. The flowchart representation is provided in Fig. 1

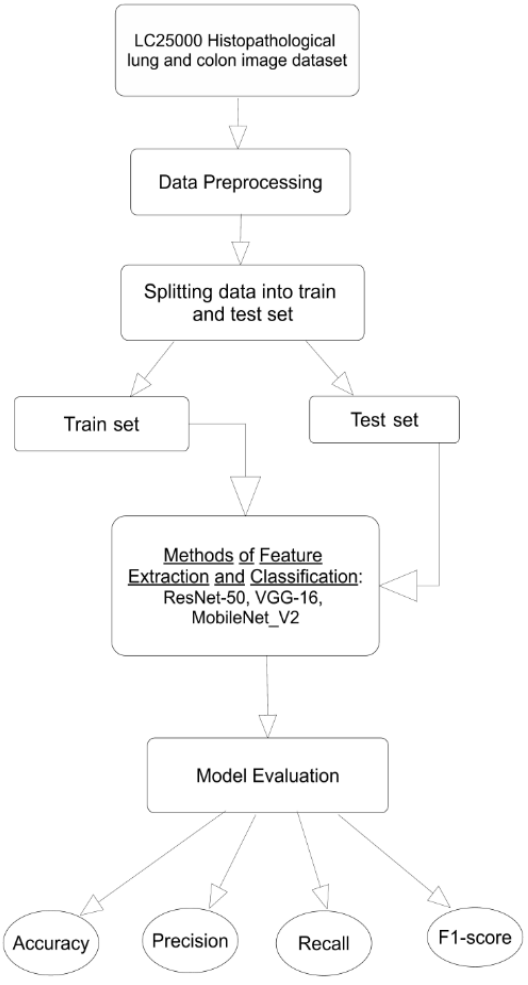

Fig. 1. Flowchart depicting the proposed methodology

3.2 Data Preprocessing

The images produced by LC25000 [31] are 768 by 768 pixels in size. The implemented models crop every image from its original 768 x 768-pixel dimensions to a square of 224 x 224 pixels. The LC25000 image collection is subjected to data augmentation, image flipping horizontally and verticacally (with a 0.5 probability) and a rotation of up to 25 degrees in either direction. [31].

- The steps involved in data pre-processing stage for MobileNet_V2 model is given below:
- To implement the model, we cropped each image from its original 768 x 768-pixel dimensions to a square of 224 x 224 pixels. The batch size used is 128, and the class mode is sparse.

3.3 Splitting Data

- Dividing dataset into two sets namely train and test set. Each class has 5000 photos, and there are five total classes which makes total number of images to 25000. The dataset is split with an 80:20 ratio, resulting in 20,000 images allocated for training and 5,000 images reserved for testing.

3.4 Feature Extraction with Classification

From study in [32], It has already been stated that performing feature extraction followed by classification helps reduce computational costs and speeds up the overall system process.

3.4.1 VGG16

The VGG16 approach is explained in the section below:

- Its structure is made up of 16 layers, comprising 13 convolutional layers that handle feature extraction and 3 completely connected layers dedicated to classification. The model is structured as a chain of convolutional layers, each accompanied by progressively deeper max-pooling operations.
- The resulting feature map (7x7x512) is flattened into a one-dimensional vector of length 25,088.
- The system includes three completely connected layers, each using ReLU activation. The first tier takes an input of size 25,088 and produces an output of 256. The second tier receives 256 inputs and outputs 128 features, while the third tier reduces the 128 features to 5 outputs, corresponding to the classification categories. For final classification, Softmax activation is applied in the last completely connected layer.

3.4.2 ResNet-50

The ResNet-50 technique is discussed below:

- The ResNet-50 architecture, structured into five stages with multiple residuals. During the initial stage, the input dimensions are reduced through a convolutional layer followed by max-pooling. The second stage made up of three residual blocks; each built with two convolutional layers and a shortcut connection. The subsequent third, fourth, and fifth stages include four, six, and three residual blocks, respectively. Across these stages, each block integrates multiple convolutional layers together with a shortcut connection.

- Create a vector of size 10035 by flattening the resulting feature map (7x7x2048).
- ReLU activates the system's three fully connected layers. The five classes are represented by the first layer, which is then activated by ReLU. First layer has an input size of 10035 which means the features coming from convolutional layer flattened into a vector of length 10035, it is connected to 256 nodes in output layer. The second layer has 256 nodes as input and 128 nodes as output and the third layer has 128 nodes as input and 5 nodes as output. Then comes the Softmax activation function which takes the value from the third layer and perform classification.

3.4.3 Mobile_V2

The MobileNet_V2 technique is described below:

- The base model MobileNet_v2 is utilized, followed by supplementary layers. To begin with, a Global Average Pooling 2D layer is incorporated, after which a hidden layer containing 1024 neurons with ReLU activation is added. Subsequently, another dense layer with five units serving as the output layer is introduced. Finally, a softmax activation function is applied to show the classification task.
- The Adam optimizer was employed to train the models. The loss function is sparse categorical cross-entropy, along with the metric used being accuracy. We have also utilized optimizers and callbacks for early stopping and implemented checkpoints for each iteration of the epoch to prevent overfitting of the model. This methodology ensures that we only save the optimal model while patiently monitoring the validation loss with a tolerance of 5 and setting the mode to minimum to find the optimal value of loss function. The model was optimized for 5 epochs while processing data in batches of 128, using accuracy as the primary evaluation metric. In addition, classification metrics like precision, recall, and F1-score were reported to ensure a thorough assessment of performance.
- Compilation and Training of the CNN Model: The CNN structure, incorporating kernel filtering, max-pooling, and fully connected layers, was utilized for both training and testing images. The model was trained for five epochs using the Adam optimizer, employing a batch size of 128, with categorical cross-entropy selected as the loss function.

3.5 Model Evaluation

The model's performance is assessed through multiple metrics such as accuracy, precision, recall, and F1-score, which are briefly outlined in [32].

4. EXPERIMENTAL RESULTS AND EVALUATION

The VGG16, MobileNet_V2, and ResNet-50 classified images of lung and colon cancer. The colon cancer dataset had two classes: benign and adenocarcinoma, while the lung cancer dataset included three classes: benign, adenocarcinoma, and squamous cell carcinoma. Implementation used the LC25000 lung histology image dataset. Software and hardware requirements are listed in Table 2.

Table 2. Software and Hardware Details

Software Requirements	Hardware Requirements
• Windows 10 • Web Browser- Google Chrome, Brave • Python Package Manager • Jupyter Notebook IDE • PyCharm	• Processor: Ryzen 5 4500U • AMD Radeon Graphics • RAM: 8.00 GB • 64-bit operating system • 512 GB SSD • Monitor

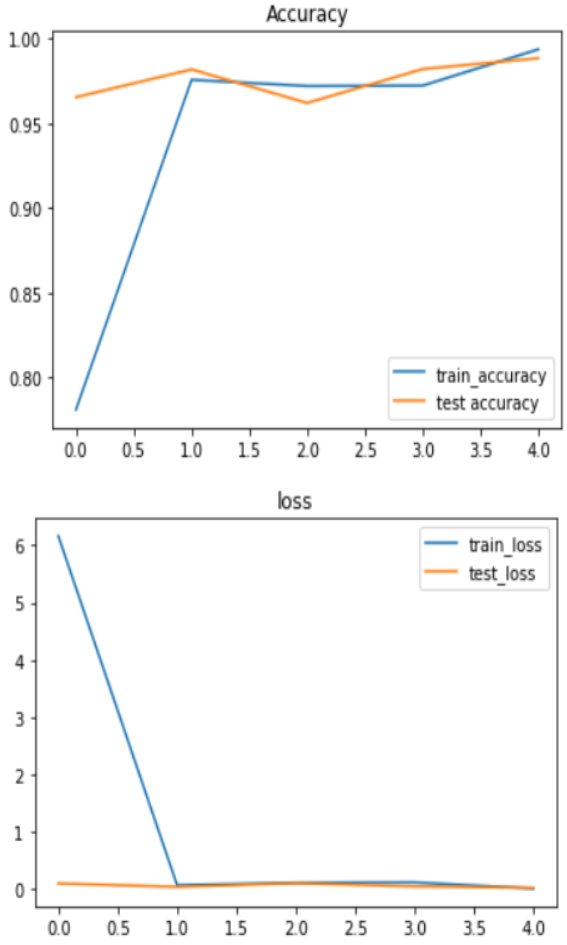

Fig. 2. Loss and Accuracy between training set and test set of ResNet-50 model.

Fig. 2 depicts the value of loss and accuracy between the training set as well as test set of ResNet-50.

In Fig. 2, we may observe that training loss as well as test loss decrease with an increase in epochs. Furthermore, test and train accuracy rise.

Table 3 illustrates the accuracy of each of the categories with respect to ResNet-50 model.

Table 3. Performance of the ResNet-50 for various categories

Model (ResNet-50)	Accuracy (%)	Precision	Recall	F1-score
Lung adenocarcinoma	99	0.99	0.99	0.99
Lung benign tissue	99	0.98	0.97	0.97
Lung squamous cell carcinoma	99	1.00	0.99	1.00
Colon adenocarcinoma	99	1.00	1.00	1.00
Colon benign tissue	99	0.97	0.99	0.98

Table 3 reveals that colon adenocarcinoma achieved the highest performance, whereas lung benign tissue achieved the lowest. The category performance was ranked from highest to lowest as follows: colon adenocarcinoma, lung squamous cell carcinoma, lung adenocarcinoma, colon benign tissue, and lung benign tissue.

Fig. 3 shows the confusion matrix for ResNet-50.

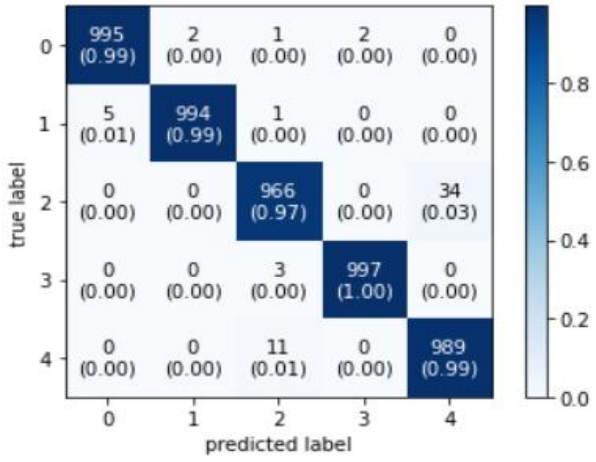

Fig. 3. Confusion matrix of ResNet-50.

Figure 3 shows the relationship between true class labels and predicted labels for the test dataset. The main misclassifications involved 34 lung adenocarcinoma images labeled as lung squamous cell carcinoma and 11 squamous cell carcinoma images identified as adenocarcinoma.

Fig. 4 depicts the value of loss and accuracy between the training and test sets of VGG16.

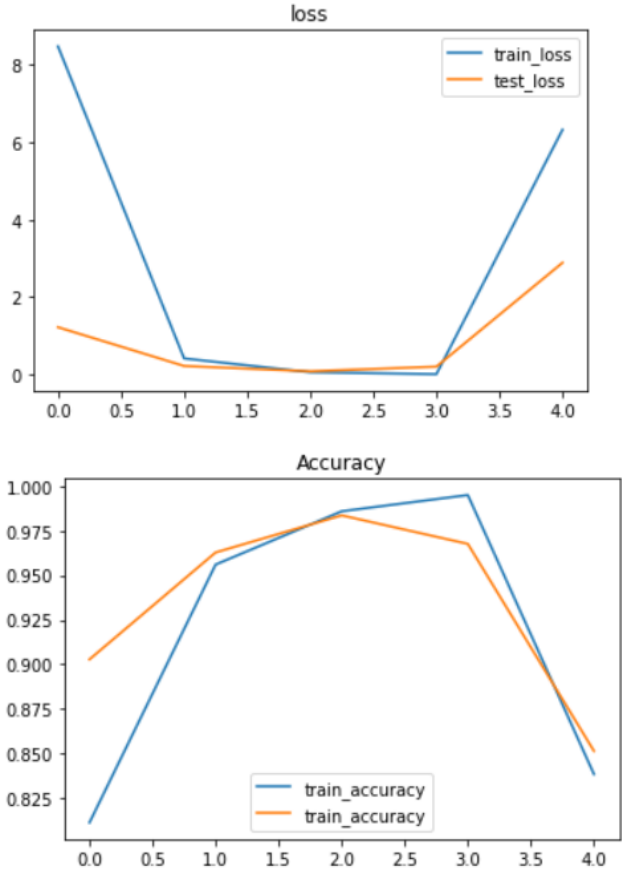

Fig. 4. Loss and Accuracy between training set and test set of VGG16 model

In Fig. 4, we may observe that training loss as well as test loss decrease with an increase in epochs, and accuracy on tests and in training rises. Fig. 5 shows the confusion matrix of VGG16.

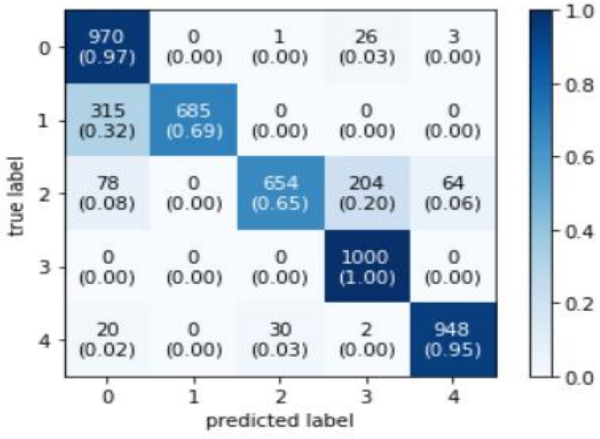

Fig. 5. Confusion matrix of VGG16.

Figure 5 shows the comparison between the true image labels and the predicted labels across various categories. Specifically, 315 colon benign images were misclassified as

colon adenocarcinoma, 346 as lung adenocarcinoma, 52 as lung squamous cell carcinoma, while the lung_n category had no misclassifications. These are the main misclassifications. Table 4 highlights the class-wise performance of the VGG16 classifier.

Table 4. VGG16 performance for various categories

Model (VGG16)	Accuracy (%)	Precision	Recall	F1-score
Lung adenocarcinoma	98	0.70	0.97	0.81
Lung benign tissue	97	0.95	0.65	0.78
Lung squamous cell carcinoma	99	1.00	0.69	0.81
Colon adenocarcinoma	99	0.81	1.00	0.90
Colon benign tissue	99	0.93	0.95	0.94

In Table 4, the highest performance is achieved in colon_benign tissue, whereas the worst performance is achieved in lung _benign tissue. The performance ranking from lowest to highest was as follows: colon benign tissue, colon adenocarcinoma, lung squamous cell carcinoma, lung adenocarcinoma, and lung benign tissue.

Fig. 6. shows the loss value and accuracy between the training and test sets of MobileNet_V2.

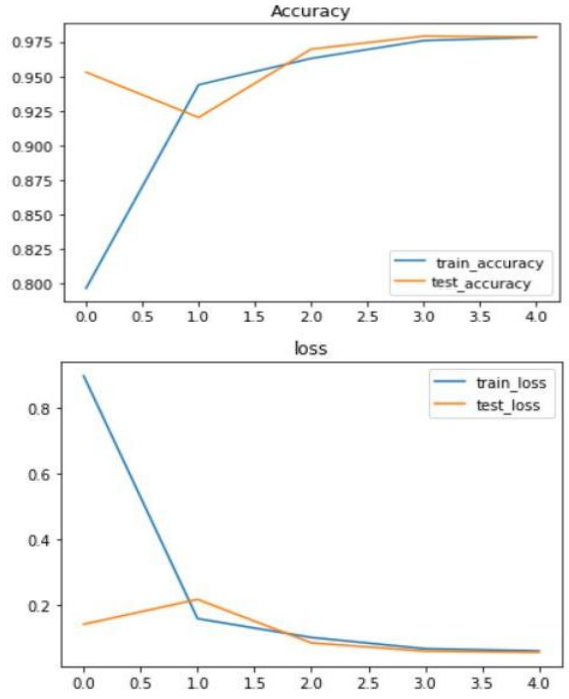

Fig. 6. Value of loss and Accuracy of training set as well as test set of MobileNet_V2.

In Fig. 6, The MobileNet_V2 model underwent training for five epochs, and the outcome was plotted between the test and train accuracy and test and train loss. It achieved a test accuracy of 0.9602 and a test loss of 0.1118.

Fig. 7 shows the confusion matrix for MobileNet_V2.

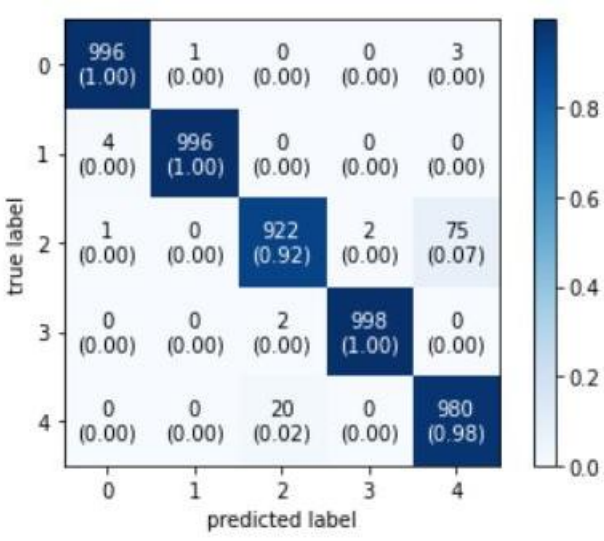

Fig. 7. Confusion matrix of MobileNet_V2.

In Fig. 7, the major misclassifications can be seen as 75 images of the lung_aca class classified as lung_scc, 20 images of lung_scc classified as lung_aca, Colon_aca has only four misclassifications, colon_n has four misclassifications, and lung_n has only two misclassifications.

Table 5 highlights the staging of each category for the MobileNet_V2 classifier.

Table 5. Performance of the MobileNet_V2 for various categories

Model (MobileNet_V2)	Accuracy (%)	Precision	Recall	F1-score
Colon adenocarcinoma	99	1.00	1.00	1.00
Colon benign tissue	99	1.00	1.00	1.00
Lung adenocarcinoma	98	0.98	0.92	0.95
Lung benign tissue	99	1.00	1.00	1.00
Lung squamous cell carcinoma	98	0.93	0.98	0.95

In Table 5, the highest performance is achieved in colon adenocarcinoma, colon benign, and lung benign tissues, whereas the worst performance is achieved in lung adenocarcinoma. The classes of colon adenocarcinoma, colon benign tissue, and lung benign tissue showed similar performance, while the performance of lung squamous cell carcinoma and lung adenocarcinoma followed in a decreasing order. Table 6 presents the training, test accuracy, testing time duration, number of epochs, and

number of misclassified images for ResNet-50, VGG16, and MobileNet_V2.

Table 6. Performance comparison of ResNet-50, VGG-16, and MobileNet_V2

Model	Train accuracy	Test accuracy	Testing time duration	Number of epochs	Number of misclassified images (from 5000)
ResNet-50	0.99	0.988	4 hrs.	5	58
VGG16	0.85	0.851	4.5 hrs.	5	743
MobileNet_V2	0.97	0.97	15mins	5	106

Table 6 highlights the performance comparison of these three models. ResNet-50, VGG16, and MobileNet were improved in terms of accuracy. Among them, ResNet-50 acheived the highest accuracy of 98.8%, followed by MobileNet_V2 with 97% and VGG16 with 85.1%.

Figure 8 shows the training and testing accuracies of the implemented models.

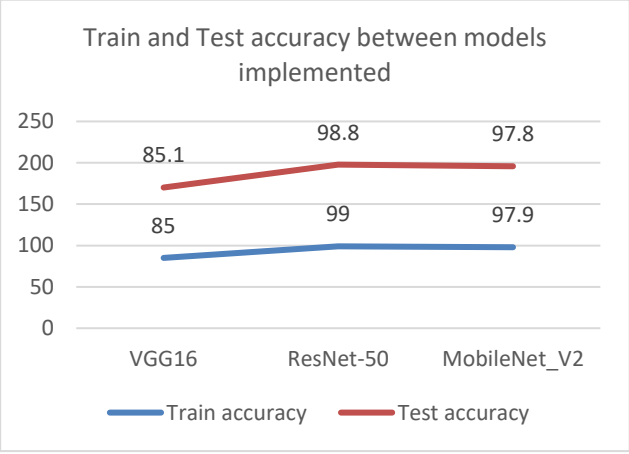

Fig. 8. Training and testing accuracy of the proposed models.

In the above Fig. 8, it may be observed that the train and test accuracy is higher in ResNet-50 than the other two models implemented.

Fig. 9 depicts the misclassified images for the proposed model.

In the above Fig. 9, it is illustrated that the number of misclassified test images in ResNet-50 is 58, which is much less than the other two models implemented, and the total misclassified images from testing set in Vgg16 is high—that is, 743.

Table 7 presents a well analysis of comparison between the proposed models and existing approaches.

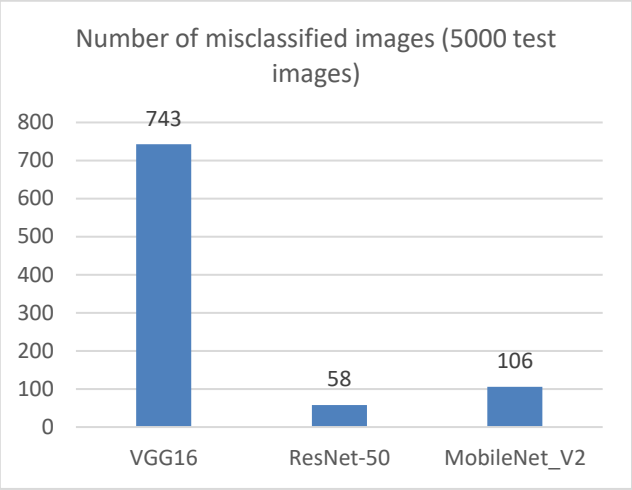

Fig. 9. Number of misclassified images proposed models.

Table 7. Performance evaluation of proposed and existing models using the LC25000 dataset

Reference	Dataset	Models	Accuracy
[9]	LC25000 (lung and colon)	Custom CNN	96.33 %
[10]	LC25000 (lung and colon)	VGG16, InceptionResNetV2, InceptionV3, Xception, ResNet50, DenseNet169 and MobileNet.	96% to 100%
[14]	LC25000 (lung)	CNN-Light gradient boosting	99.6%
[15]	LC25000 (lung and colon)	EfficientNetB0-B7	97.24%
[18]	LC25000 (lung)	CNN	99.97%,
[34]	LC25000 (lung and colon)	AlexNet	98.40%
[32]	LC25000 (lung and colon)	VGG16 with Random Forest	95.1%
		VGG16 with Xgboost	97.46%
		GLCM with LightGBM	93.52%
Proposed Work	LC25000 (lung and colon)	ResNet-50	98.8%
		VGG16	85.1%
		MobileNet_V2	97.8%

Table 7 above demonstrates how well the proposed models outperform with existing models.

5. DISCUSSION AND FUTURE DIRECTIONS

To improve accuracy and determine which design is most effective with this dataset, the three transfer learning architectures are evaluated in this manuscript. All three transfer learning models—VGG16, MobileNet_V2, and ResNet-50—demonstrated strong accuracy. In this study, ResNet-50 reached the highest accuracy of 98.8%, followed by MobileNet_V2 with 97.8%, whereas VGG16 showed the lowest accuracy of 85.1%. Looking forward, our plans include improving the overall model and developing novel feature groups derived from additional histopathological image data to boost classification effectiveness. Future work will aim to identify critical physiological factors that influence lung cancer, which can then be incorporated as predictive markers in upcoming cancer detection models. Physiological suggestions, including body temperature, oxygen saturation, and pressure, are used to make prognostications. We will be looking into more sophisticated data processing techniques, like conventional normalisation and data augmentation strategies, to enhance results. Further optimisation will involve fine-tuning variables such as learning rate, network depth, and training iterations. We also study CNN models that are not the conventional choices for deep learning, like Inception Net, Dense Net, and UNet.

References

[1] Youlden DR, Cramb SM, Baade PD. The International Epidemiology of Lung Cancer: geographical distribution and secular trends. J Thorac Oncol. 2008 Aug;3(8):819-31. doi: 10.1097/JTO.0b013e31818020eb. PMID: 18670299.

[2] Purandare NC, Rangarajan V. Imaging of lung cancer: Implications on staging and management. Indian J Radiol Imaging. 2015 Apr-Jun;25(2):109-20. doi: 10.4103/0971-3026.155831. PMID: 25969634; PMCID: PMC4419420.

[3] Kanavati F, Toyokawa G, Momosaki S, Takeoka H, Okamoto M, Yamazaki K, Takeo S, Iizuka O, Tsuneki M. A deep learning model for the classification of indeterminate lung carcinoma in biopsy whole slide images. Sci Rep. 2021 Apr 14;11(1):8110. doi: 10.1038/s41598-021-87644-7. PMID: 33854137; PMCID: PMC8046816.

[4] Medical Advisory Secretariat. Screening methods for early detection of colorectal cancers and polyps: summary of evidence-based analyses. Ont Health Technol Assess Ser. 2009;9(6):1-65. Epub 2009 Sep 1. PMID: 23074536; PMCID: PMC3377498.

[5] Shaukat, A., Levin, T.R. Current and future colorectal cancer screening strategies. *Nat Rev Gastroenterol Hepatol* 19, 521–531 (2022). https://doi.org/10.1038/s41575-022-00612-y

[6] Inamura K. Lung Cancer: Understanding Its Molecular Pathology and the 2015 WHO Classification. Front Oncol. 2017 Aug 28;7:193. doi: 10.3389/fonc.2017.00193. PMID: 28894699; PMCID: PMC5581350.

[7] Wang L. Deep Learning Techniques to Diagnose Lung Cancer. Cancers (Basel). 2022 Nov 13;14(22):5569. doi: 10.3390/cancers14225569. PMID: 36428662; PMCID: PMC9688236.

[8] Anjum S, Ahmed I, Asif M, Aljuaid H, Alturise F, Ghadi YY, Elhabob R. Lung Cancer Classification in Histopathology Images Using Multiresolution Efficient Nets. Comput Intell Neurosci. 2023 Oct 16;2023: 7282944. doi: 10.1155/2023/7282944. PMID: 37876944; PMCID: PMC10593544.

[9] Masud M, Sikder N, Nahid AA, Bairagi AK, AlZain MA. A Machine Learning Approach to Diagnosing Lung and Colon Cancer Using a Deep Learning-Based Classification Framework. Sensors (Basel). 2021 Jan 22;21(3):748. doi: 10.3390/s21030748. PMID: 33499364; PMCID: PMC7865416.

[10] Garg, Satvik & Garg, Somya. (2021). 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.

[11] W. Abdul, "An Automatic Lung Cancer Detection and Classification (ALCDC) System Using Convolutional Neural Network," 2020 13th International Conference on Developments in eSystems Engineering (DeSE), Liverpool, United Kingdom, 2020, pp. 443-446, doi: 10.1109/DeSE51703.2020.9450778.

[12] R. Raja Subramanian, R. Nikhil Mourya, V. Prudhvi Teja Reddy, B. Narendra Reddy, Srikar Amara. "Lung Cancer Prediction Using Deep Learning Framework" International Journal of Control and Automation Vol. 13, No. 3, (2020), pp. 154-160

[13] Sajja, Tulasi & Devarapalli, Retz Mahima & Kalluri, Hemantha kumar. (2019). Lung Cancer Detection Based on CT Scan Images by Using Deep Transfer Learning. Traitement du Signal. 36. 339-344.

[14] Abd Elraouf, Esraa & Salem, Mohammed Abdel-Megeed Mohammed & Badr, Nagwa & Tolba, Mohamed. (2023). An Efficient Combination of Convolutional Neural Network and LightGBM Algorithm for Lung Cancer Histopathology Classification.

[15] Anjum, Sunila & Ahmed, Imran & Asif, Muhammad & Aljuaid, Hanan & Alturise, Fahad & Ghadi, Yazeed & Elhabob, Rashad. (2023). Lung Cancer Classification in Histopathology Images Using Multiresolution Efficient Nets. Computational Intelligence and Neuroscience. 2023. 1-12. 10.1155/2023/7282944.

[16] Wang, Z., Gao, J., Kan, H., Huang, Y., Tang, F., Li, W., & Yang, F. (2023). ResNet for Histopathologic Cancer Detection, the Deeper, the Better? Journal of Data Science and Intelligent Systems. https://doi.org/10.47852/bonview JDSIS3202744

[17] Pang, Shanchen & Meng, Fan & Wang, Xun & Wang, Jianmin & Song, Tao & Wang, Xingguang & Cheng, Xiaochun. (2020). VGG16-T: A Novel Deep Convolutional Neural Network with Boosting to Identify Pathological Type of Lung Cancer in Early Stage by CT Images. International Journal of Computational Intelligence Systems. 13. 10.2991/ijcis.d.200608.001.

[18] Hatuwal, Bijaya & Thapa, Himal. (2020). Lung Cancer Detection Using Convolutional Neural Network on Histopathological Images. International Journal of Computer

Trends and Technology. 68. 21-24. 10.14445/22312803/ IJCTT-V68I10P104.

[19] Tummala S, Kadry S, Nadeem A, Rauf HT, Gul N. An Explainable Classification Method Based on Complex Scaling in Histopathology Images for Lung and Colon Cancer. Diagnostics (Basel). 2023 Apr 29;13(9):1594. doi: 10.3390/diagnostics13091594. PMID: 37174985; PMCID: PMC10178684.

[20] Geng L, Zhang S, Tong J, Xiao Z. Lung segmentation method with dilated convolution based on VGG-16 network. Comput Assist Surg (Abingdon). 2019 Oct;24(sup2):27-33. doi: 10.1080/24699322.2019.1649071. Epub 2019 Aug 12. PMID: 31402721.

[21] Šarić, Matko & Russo, Mladen & Stella, Maja & Sikora, Marjan. (2019). CNN-based Method for Lung Cancer Detection in Whole Slide Histopathology Images. 1-4. 10.23919/SpliTech.2019.8783041.

[22] Zhou, Tao & Bingqiang, Huo & Huiling, Lu & Yang, Zaoli & Hongbin, Shi. (2020). NSCR-Based DenseNet for Lung Tumor Recognition Using Chest CT Image. BioMed Research International. 2020. 1-9. 10.1155/2020/6636321.

[23] Naseer, T. Masood, S. Akram, A. Jaffar, M. Rashid et al., "Lung cancer detection using modified alexnet architecture and support vector machine," *Computers, Materials & Continua*, vol. 74, no.1, pp. 2039–2054, 2023.

[24] T. Fang, "A Novel Computer-Aided Lung Cancer Detection Method Based on Transfer Learning from GoogLeNet and Median Intensity Projections," 2018 IEEE International Conference on Computer and Communication Engineering Technology (CCET), Beijing, China, 2018, pp. 286-290, doi: 10.1109/CCET.2018.8542189.

[25] Shen, Wei & Zhou, Mu & Yang, Feng & Dong, Di & Yang, Caiyun & Zang, Yali & Tian, Jie. (2016). Learning from Experts: Developing Transferable Deep Features for Patient-Level Lung Cancer Prediction. 124-131. 10.1007/978-3-319-46723-8_15.

[26] Ausawalaithong, Worawate & Thirach, Arjaree & Marukatat, Sanparith & Wilaiprasitporn, Theerawit. (2018). Automatic Lung Cancer Prediction from Chest X-ray Images Using the Deep Learning Approach. 1-5. 10.1109/ BMEiCON.2018.8609997.

[27] F. Al-Yasriy, Hamdalla & Al-Huseiny, Muayed & Mohsen, Furat & Khalil, Enam & Hassan, Zainab. (2020). Diagnosis of Lung Cancer Based on CT Scans Using CNN. IOP Conference Series: Materials Science and Engineering. 928. 022035. 10.1088/1757-899X/928/2/022035.

[28] Teramoto A, Tsukamoto T, Kiriyama Y, Fujita H. Automated Classification of Lung Cancer Types from Cytological Images Using Deep Convolutional Neural Networks. Biomed Res Int. 2017; 2017:4067832. doi: 10.1155/2017/4067832. Epub 2017 Aug 13. PMID: 28884120; PMCID: PMC5572620.

[29] Shimazaki, Akitoshi & Ueda, Daiju & Choppin, Antoine & Yamamoto, Akira & Honjo, Takashi & Shimahara, Yuki & Miki, Yukio. (2022). Deep learning-based algorithm for lung cancer detection on chest radiographs using the segmentation method. Scientific Reports. 12. 10.1038/s41598-021-04667-w.

[30] L. Fan, Z. Xia, X. Zhang and X. Feng, "Lung nodule detection based on 3D convolutional neural networks," 2017 International Conference on the Frontiers and Advances in Data Science (FADS), Xi'an, China, 2017, pp. 7-10, doi: 10.1109/FADS.2017.8253184.

[31] "LC25000 Lung and Colon Cancer Histopathological Image Dataset." [Online]. Available: https://github.com/tampapath/ lung_colon_image_set/blob/ master/README.md.

[32] Sharma, D., Choubey, D. K., & Thakur, K. (2025). Lung and Colon Cancer Detection using Deep Learning Techniques. Procedia Computer Science, 258, 4136-4146.

[33] Choubey, D. K., Newton, R., Ojha, M. K., & Kumar, S. (2023, February). Soft Computing and Data Mining Techniques for Dengue Detection: A Review. In International Conference on Innovative Computing and Communication (pp. 295-309). Singapore: Springer Nature Singapore.

[34] S. Mehmood et al., "Malignancy Detection in Lung and Colon Histopathology Images Using Transfer Learning with Class Selective Image Processing," in IEEE Access, vol. 10, pp. 25657-25668, 2022, doi: 10.1109/ACCESS.2022. 3150924.

[35] Jangir, S. K., Joshi, N., Kumar, M., Choubey, D. K., Singh, S., & Verma, M. (2021). Functional link convolutional neural network for the classification of diabetes mellitus. International Journal for Numerical Methods in Biomedical Engineering, 37(8), e3496.

[36] Choubey, D. K., Paul, S., & Dhandhania, V. K. (2019). GA_NN: An intelligent classification system for diabetes. In Soft Computing for Problem Solving: SocProS 2017, Volume 2 (pp. 11-23). Springer Singapore.

[37] Choubey, D. K., Paul, S., Bala, K., Kumar, M., & Singh, U. P. (2019). Implementation of a hybrid classification method for diabetes. In Intelligent Innovations in Multimedia Data Engineering and Management (pp. 201-240). IGI Global.

[38] Choubey, D. K., & Paul, S. (2017). GA_SVM: A classification system for diagnosis of diabetes. In Handbook of research on soft computing and nature-inspired algorithms (pp. 359-397). IGI Global.

[39] Choubey, D. K., Tripathi, S., Kumar, P., Shukla, V., & Dhandhania, V. K. (2021). Classification of Diabetes by Kernel based SVM with PSO. Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science), 14(4), 1242-1255.

[40] Choubey, D. K., & Paul, S. (2016). GA_MLP NN: a hybrid intelligent system for diabetes disease diagnosis. International Journal of Intelligent Systems and Applications, 8(1), 49-59.

[41] Choubey, D. K., Paul, S., Shandilya, S., & Dhandhania, V. K. (2020). Implementation and analysis of classification algorithms for diabetes. Current Medical Imaging, 16(4), 340-354.

[42] Choubey, D. K., Kumar, M., Shukla, V., Tripathi, S., & Dhandhania, V. K. (2020). Comparative analysis of classification methods with PCA and LDA for diabetes. Current diabetes reviews, 16(8), 833-850.

[43] Choubey, D. K., Kumar, P., Tripathi, S., & Kumar, S. (2020). Performance evaluation of classification methods with PCA and PSO for diabetes. Network Modeling Analysis in Health Informatics and Bioinformatics, 9(1), 1-30.

[44] D. Sharma, P. Jain, and D. K. Choubey, "A Comparative Study of Computational Intelligence for Identification of Breast Cancer," in International Conference on Machine Learning, Image Processing, Network Security and Data

Sciences, 2020, pp. 209–216.
[45] A. Parthvi, K. Rawal, and D. K. Choubey, “A Comparative study using Machine Learn-ing and Data Mining Approach for Leukemia,” Proc. 2020 IEEE Int. Conf. Commun. Signal Process. ICCSP 2020, pp. 672–677, 2020, doi: 10.1109/ICCSP48568.2020.9182142.
[46] S. Pahari and D. K. Choubey, “Analysis of Liver Disorder Using Classification Techniques: A Survey,” Int. Conf. Emerg. Trends Inf. Technol. Eng. ic-ETITE 2020, pp. 1–4, 2020, doi: 10.1109/ic-ETITE47903.2020.300.
[47] Choubey, D. K., & Paul, S. (2016). Classification techniques for diagnosis of diabetes: a review. International Journal of Biomedical Engineering and Technology, 21(1), 15-39.
[48] Srivastava, K., & Choubey, D. K. (2021). Soft Computing, Data Mining, and Machine Learning Approaches in Detection of Heart Disease: A Review. In Hybrid Intelligent Systems: 19th International Conference on Hybrid Intelligent Systems (HIS 2019) held in Bhopal, India, December 10-12, 2019 19 (pp. 165-175). Springer International Publishing.
[49] Choubey, D. K., Choubey, A., Mahto, S., & Soni, V. (2024). Soft Computing Approaches for Ovarian Cancer: A Review. GMSARN International Journal, 18, 223-239.
[50] Verma, A., Singh, A., Bihari, A., Tripathi, S., Agrawal, S., Pandey, S. K., & Verma, S. (2023). Identification of Hate Speech on Social Media using LSTM. GMSARN International Journal, 17, 468-474.
[51] Bhardwaj, S. K., Vishwakarma, S., Bihari, A., Tripathi, S., Agrawal, S., & Joshi, P. (2024). Protein Enzyme Sequence Class Prediction using Computational Model. GMSARN International Journal, 18, 62-70.