

Convolutional Neural Networks in Malaria Diagnosis: A Study on Cell Image Classification

Hritwik Ghosh¹, Irfan Sadiq Rahat^{2*}, JVR Ravindra³, Balajee J⁴, Mohammad Aman Ullah Khan⁵, J. Somasekar⁶

^{1,2}School of Computer Science and Engineering (SCOPE), VIT-AP University, Amaravati, Andhra Pradesh

³Center for Advanced Computing Research Laboratory (C-ACRL), Department of Electronics and Communication Engineering, Vardhaman College of Engineering (Autonomous), Hyderabad, Telangana, India.

⁴Department of Computer Science & Engineering, Mother Theresa Institute of Engineering and Technology, Palamaner-517408, Chittoor, Andhra Pradesh, India

⁵Department of Computer Science and Engineering, BRAC University, Bangladesh

⁶Department of Computer Science & Engineering, Faculty of Engineering and Technology, JAIN (Deemed-to-be University), Bengaluru, Karnataka

Abstract

INTRODUCTION: Malaria, a persistent global health threat caused by Plasmodium parasites, necessitates rapid and accurate identification for effective treatment and containment. This study investigates the utilization of convolutional neural networks (CNNs) to enhance the precision and speed of malaria detection through the classification of cell images infected with malaria.

OBJECTIVES: The primary objective of this research is to explore the effectiveness of CNNs in accurately classifying malaria-infected cell images. By employing various deep learning models, including ResNet50, AlexNet, Inception V3, VGG19, VGG16, and MobileNetV2, the study aims to assess the performance of each model and identify their strengths and weaknesses in malaria diagnosis.

METHODS: A balanced dataset comprising approximately 8,000 enhanced images of blood cells, evenly distributed between infected and uninfected classes, was utilized for model training and evaluation. Performance evaluation metrics such as precision, recall, F1-score, and accuracy were employed to assess the efficacy of each CNN model in malaria classification.

RESULTS: The results demonstrate high accuracy across all models, with AlexNet and VGG19 exhibiting the highest levels of accuracy. However, the selection of a model should consider specific application requirements and constraints, as each model presents unique trade-offs between computational efficiency and performance.

CONCLUSION: This study contributes to the burgeoning field of deep learning in healthcare, particularly in utilizing medical imaging for disease diagnosis. The findings underscore the considerable potential of CNNs in enhancing malaria diagnosis. Future research directions may involve further model optimization, exploration of larger and more diverse datasets, and the integration of CNNs into practical diagnostic tools for real-world deployment.

Keywords: Malaria, ResNet50, AlexNet, Inception V3, VGG19, VGG16, precision, recall, F1-score, deep learning

Received on 23 December 2023, accepted on 20 March 2024, published on 26 March 2024

Copyright © 2024 H. Ghosh *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [CC BY-NC-SA 4.0](#), which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/eetpht.10.5551

*Corresponding author. email: me.rahat2020@gmail.com

1. Introduction

The intersection of artificial intelligence and healthcare has opened up new frontiers in medical diagnostics, particularly

in the field of medical imaging. Among the various applications, the classification of lung tissue has emerged as a critical area of focus due to its implications for the early detection and treatment of lung diseases, including cancer.

The ability to accurately classify lung tissue can significantly enhance the effectiveness of therapeutic interventions, thereby improving patient outcomes. This study delves into the application of DL models for the classification of lung tissue, with a particular emphasis on histopathological images [15,16]. Histopathological images offer an intricate insight into the microscopic architecture of tissues, unearthing minor alterations that could potentially signal disease. Nonetheless, the deciphering of these images necessitates profound expertise and can be a lengthy process. The advent of deep learning methodologies promises to bring about a paradigm shift in this arena, automating the interpretation process and delivering prompt, precise outcomes. DL, a specialised branch of ML is uniquely equipped for image analysis due to its inherent capability to learn complex patterns directly from the data, thereby obviating the need for manual extraction of features. In this study, we explore the performance of six DL models - DenseNet201, EfficientNetB7, EfficientNetB5, Vgg19, Vgg16, and Alexnet - in the classification of lung tissue. These models were chosen due to their proven effectiveness in image classification tasks. We trained and evaluated these models on a dataset of 9,000 histopathological images, divided into five distinct classes of lung tissue [17,18]. The images were sourced from HIPAA compliant and validated sources and were augmented to ensure a robust and diverse dataset. The aim of the research is not just to assess the performance of these models but also to gain insights into the factors that contribute to their effectiveness. By doing so, we aim to contribute to the ongoing efforts to optimise the use of DL techniques in medical imaging, with the ultimate goal of improving the accuracy and efficiency of disease diagnosis [19,20]. The organisation of this paper is as follows: Subsequent to this introductory section, we delve into an examination of pertinent scholarly works, setting the stage for our investigation. This is followed by an exposition of our research approach, encompassing specifics of the dataset and the DL architectures employed. We then proceed to unveil and interpret our findings. The paper culminates with a distillation of our discoveries and a discussion on the potential avenues for future exploration in this field.

2. Literature Review

A DL model (DLM) was created by Jünger et al. in 2021[1] for the automatic detection and 3D segmentation of brain metastases (BMs) in NSCLC patients using clinically standard MRI. The study included 98 NSCLC patients with 315 BMs and was retrospective in nature. The model was designed to address the difficulties and dangers of making an incorrect diagnosis when detecting BMs on MRI in clinical settings. Masud et al. investigated the use of AI in 2021[2] to automate the diagnosis of colon and lung cancer, two common and deadly types of cancer. To analyse histopathology images of lung and colon tissues, they created a deep learning-based classification system using cutting-edge DL and DIP techniques. With a maximum accuracy of 96.33% in recognising cancer tissues, the framework attempted to differentiate between five different types of

tissues. This method has the potential to help medical practitioners create an automatic and trustworthy system for identifying cancer, improving the chances of early detection and treatment. Using whole-slide images from The Cancer Genome Atlas, Coudray et al. in 2018[3] created a deep convolutional neural network to categorise LUAD, LUSC, or normal lung tissue. The model was verified using several separate datasets, and it performed similarly to pathologists on average, with an AUC of 0.97. The ten most frequently mutated genes in LUAD were also predicted by the network, and six of them were correctly predicted with AUCs ranging from 0.733 to 0.856. This study demonstrates how deep learning could help pathologists identify cancer subtypes and anticipate gene mutations. In 2022[4], Chen et al. proposed a computer-aided diagnostic (CAD) approach that makes use of radiomics and deep attention-based MIL for early LC diagnosis. This method outperformed existing MIL techniques with an AUC of 0.842 and a mean accuracy of 0.807, better representing clinical diagnostic processes. The output is more interpretable and acceptable for both doctors and patients thanks to the attention mechanism, which also provides improved interpretability of the output. To identify high-risk patients and minimize unnecessary operations, Yeh et al. [5] developed an AI algorithm to predict LC risk using electronic medical information in 2021. After training using data from the Taiwan National Health Insurance Research Database, the model's AUC was 0.90 for the overall population and 0.87 for patients under the age of 55. This strategy allows for the reliable identification of patients who are at risk of developing lung cancer by combining complex information from non-imaging medical records.

A deep transfer learning model combining convolutional neural networks and convolutional auto-encoders was put forth by Rong et al. in 2021[6] for the diagnostic categorization of lung cancer using multi-omics data. Convolutional auto-encoders are used in this early lung cancer diagnosis method to reduce dimensionality and adhere to transfer learning standards. When tested with three LC gene datasets and an integrated dataset, the model beat previous ML models in terms of accuracy, as well as average area under the curve. Tan et al. created a customised DNN based on the VGG16 architecture in 2022[7] that uses transfer learning to discriminate between TB lung nodules and early-stage lung malignancies. Using unlocked pretrained weights on CT images from the National Lung Screening Trial and the National Institute of Allergy and Infectious Disease TB Portals, the DNN demonstrated its potential as a dependable, noninvasive screening tool for detecting and distinguishing between LC and tuberculosis, with a detection rate of 90.4% and a F score of 90.1%. In order to classify benign nodules, primary lung cancer, and metastatic lung cancer, Nishio et al. in 2018[8] created a CADx approach using DCNN and transfer learning. The study, which involved 1236 patients, compared the efficiency of DCNN with that of traditional techniques and assessed how image size affected DCNN

input. The results showed that DCNN with transfer learning outperformed traditional approaches, obtaining a best averaged validation accuracy of 68.0%, and that lung nodule classification accuracy was improved by using larger input image sizes. Using data from 887 patients, Park et al. created a two-stage U-Net architecture in 2023[9] for the automatic segmentation of lung cancer in [18F] FDG PET/CT scans. The first stage uses a global U-net to extract preliminary tumour regions from the 3D PET/CT volume, and the second stage refines these areas using a localised U-net on chosen slices. By accurately predicting specific tumour margins and showcasing its benefits through quantitative analysis using the Dice similarity coefficient, this technique surpassed the traditional one-stage 3D U-Net. In order to effectively identify lung and colon cancer from histopathology pictures, Talukder et al. presented a hybrid ensemble feature extraction model in 2022[10]. This model combines deep feature extraction and ensemble learning. The model outperformed previous models by a wide margin when tested on the LC25000 datasets, with accuracy rates of 99.05% for lung cancer, 100% for colon cancer, and 99.30% for both. According to this study, such models may prove useful in clinical settings, helping clinicians diagnose cancer by making prompt and precise detections. Using multiview radiomics and DL, Zhang et al. developed a model in 2023[11] to quantitatively predict N2 lymph node metastases in individuals with stage I-II NSCLC. The study included 140 patients and employed transfer

learning techniques and an end-to-end ResNet18 architecture. Using 1.8 million natural photographs and a small sample of N2 lymph node VOI images as training data, the DL model outperformed the radiomics model (Rad), which had an AUC of 0.76. With an AUC of 0.88, the combination model (Rad + DL + Clinical) had the greatest diagnostic performance, perhaps indicating the lymph node metastases in NSCLC patients. Li et al. presented a unique deep learning-based drug repurposing strategy for non-small cell lung cancer in 2020[12], concentrating on transcriptome data and chemical structures. This method identified Pimozide as a potent candidate for the treatment of non-small cell lung cancer. Pimozide is typically used as an anti-dyskinesia medication for Tourette's Disorder. The research confirmed Pimozide's cytotoxicity towards A549 cell lines, showing the possibility for systematic medication repurposing using cutting-edge computational approaches to provide new therapeutic options. Zheng et al. in 2023[13] developed a hybrid model using deep learning to predict survival for stage I-III non-small cell lung cancer patients, integrating image and clinical features. The model, trained on patients who received stereotactic radiotherapy, utilized image features from pre-treatment CT scans and significant clinical variables like age and clinical stage. The model demonstrated a median AUC of 0.76 and 0.64 on different test sets, effectively separating low and high mortality risk groups, showcasing the potential of DL in enhancing prognostic accuracy in lung cancer treatment [14] [Table.1].

Table.1 Summary of the Literature Review

Reference	Focus of Study	Techniques Used	Key Findings
Jünger et al. (2021) [1]	Automatic detection and 3D segmentation of brain metastases	Deep Learning (DL), clinically standard MRI	Created a DL model for detecting and segmenting brain metastases in NSCLC patients. Retrospective study with 98 NSCLC patients and 315 BMs. Addressed the challenges of incorrect diagnosis in clinical MRI settings.
Masud et al. (2021) [2]	AI in automating diagnosis of colon and lung cancer	Deep Learning (DL), Digital Image Processing (DIP)	Developed a DL-based classification system for histopathology images. Achieved 96.33% accuracy in recognizing cancer tissues. Potential for creating an automatic and trustworthy system for early cancer detection.
Coudray et al. (2018) [3]	Categorization of lung tumors into subtypes	Deep Convolutional Neural Network (DCNN)	Used deep learning to categorize lung tumors into adenocarcinoma, squamous cell carcinoma, or normal tissue. Achieved AUC of 0.97 and predicted frequently mutated genes in LUAD. Demonstrated how DL helps identify cancer subtypes and anticipate gene mutations.
Chen et al. (2022) [4]	Computer-aided diagnostic approach for early lung cancer	Radiomics, deep attention-based Multiple Instance Learning	Proposed a CAD approach using radiomics and deep attention-based MIL for early lung cancer

Yeh et al. (2021) [5]	AI algorithm to predict lung cancer risk using EMR	Artificial Intelligence (AI), Electronic Medical Records	diagnosis. Outperformed existing MIL techniques with AUC of 0.842. Improved interpretability and acceptability for clinical use. Created an AI algorithm to predict lung cancer risk using EMR. AUC of 0.90 for the general population and 0.87 for patients under 55. Accurate identification of high-risk patients, avoiding unnecessary procedures.
Rong et al. (2021) [6]	Diagnostic categorization of lung cancer	Deep Transfer Learning, Convolutional Auto-encoders	Introduced a deep transfer learning model for lung cancer diagnosis using multi-omics data. Outperformed other ML models in accuracy, precision, recall, and F1-score. Best results in average AUC.
Tan et al. (2022) [7]	Discrimination between TB lung nodules and early-stage cancer	Deep Convolutional Neural Network (DNN), Transfer Learning	Developed a custom DNN based on VGG16 architecture to discriminate between TB lung nodules and early-stage lung malignancies. Achieved 90.4% accuracy and 90.1% F score. Potential as a noninvasive screening tool.
Nishio et al. (2018) [8]	Computer-aided diagnosis for lung nodule classification	Deep Convolutional Neural Network (DCNN), Transfer Learning	Created a CADx approach using DCNN and transfer learning for classifying benign nodules, primary lung cancer, and metastatic lung cancer. Outperformed traditional approaches with 68.0% validation accuracy. Image size affected DCNN input and improved classification accuracy.

3. Dataset Overview

This research is anchored on a well-structured dataset, which includes 9,000 histopathological images. Each image, captured in JPEG format and with a resolution of 768 x 768 pixels, provides an in-depth view of lung tissue, making it possible to discern minor changes that may be indicative of different conditions. The dataset was created with a strong commitment to the Health Insurance Portability and Accountability Act (HIPAA), thereby guaranteeing the highest level of data privacy and security. The images in the dataset are categorized into three distinct classes, each representing a different type of lung tissue. These include benign lung tissue, lung adenocarcinoma, and lung squamous cell carcinoma. Each class is represented by 3,000 images, ensuring a balanced dataset that minimises the risk of model bias towards any particular class. The original dataset was derived from 750 images of lung tissue (250 each of benign lung tissue, lung adenocarcinoma, and lung squamous cell

carcinoma). To enhance the diversity and robustness of the dataset, these images were augmented using the Augmentor package, resulting in a total of 9,000 images. This process of augmentation is crucial in machine learning, as it helps improve the model's ability to generalise and perform accurately on unseen data. The dataset was further divided into training, validation, and testing sets, following a 70-15-15 split. This division ensures that the models are trained on a substantial amount of data, validated on a separate set to fine-tune parameters, and finally, tested on unseen data to evaluate their performance. This rigorous approach to dataset preparation and division underscores the robustness of our research methodology. In summary, the dataset used in this research provides a comprehensive and diverse collection of histopathological images, ensuring a robust and balanced foundation for training and evaluating the deep learning models. Its preparation and usage adhere to the highest standards of data privacy and research ethics, making it a reliable resource for this study [Fig.1].

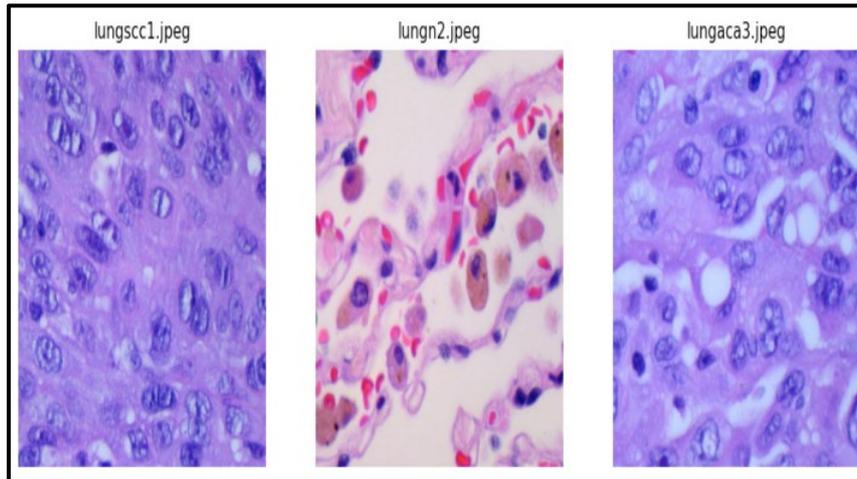


Fig 1: Sample image of dataset

3.1 Image Resizing

In the realm of medical imaging and histopathological image classification, image resizing is a crucial preprocessing step. This process involves altering the dimensions of an image to a specific size while preserving the essential features of the image. This step is particularly critical when dealing with DL models, as these models require input images of a consistent size for effective processing.

Several techniques are available for image resizing, each offering its unique benefits and trade-offs:

- **Nearest Neighbour Interpolation:** This is a fundamental technique where the value of a pixel in the resized image is determined from the closest pixel in the original image. While this method is computationally efficient, it may lead to a loss of detail and sharpness in the resized image, which could be detrimental when dealing with intricate histopathological images.
- **Bicubic Interpolation:** This method extends the concept of bilinear interpolation by considering the nearest 4x4 neighborhood of pixels. It results in smoother images than bilinear interpolation and is often employed for high-quality image processing. However, it is more computationally demanding, which might be a consideration when dealing with large datasets of histopathological images.

- **Area-based (or Resampling) Interpolation:** This method calculates the average color of the pixels within a sample area from the original image (like a 3x3 or 5x5 area) to determine the color of a pixel in the resized image. While this method is slower, it can produce high-quality results, especially when reducing the size of an image. This could be beneficial when dealing with histopathological images where the preservation of detail is paramount.

- **Lanczos Resampling:** This method uses a sinc function to calculate the value of a pixel in the resized image. It provides high-quality results and preserves more detail than other methods, but it is the most computationally intensive. This method might be suitable when the highest level of detail preservation is required, such as in the analysis of lung tissue images.

In the context of histopathological image classification, the choice of resizing technique hinges on the specific needs of the task. If computational resources and speed are a priority, simpler methods like nearest neighbor or bilinear interpolation may be suitable. However, if the quality of the resized image is crucial, more advanced methods like bicubic interpolation, area-based interpolation, or Lanczos resampling may be more appropriate. It's also vital to consider the characteristics of the images and the features that the model needs to recognize. For instance, if the images contain subtle details that are crucial for classification, a high-quality resizing method would be beneficial [Fig.2].

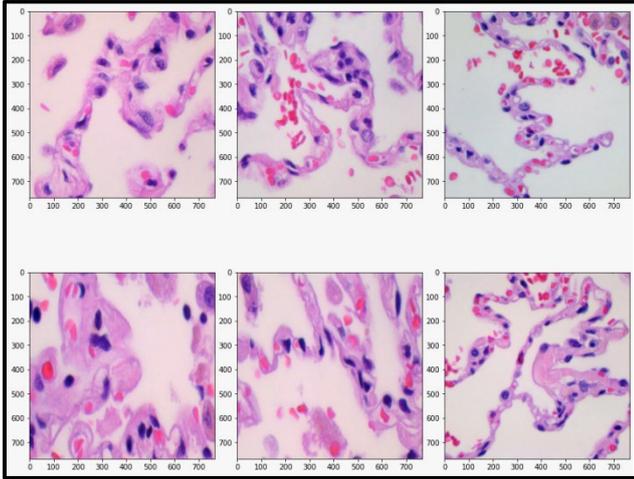


Fig 2: Pre-processing of the data set.

3.2 Data Augmentation

The amount and variety of the training dataset may be increased using the effective method of image data augmentation, which improves the performance and generalisation abilities of the model. This is especially important when classifying histopathological images since the dataset's variety can have a big influence on how well the model can differentiate between different types of lung tissue. Some of the most often utilised data augmentation methods for histopathological image classification tasks are listed below:

- **Rotation:** This technique involves rotating the histopathological image by a specific angle. This can aid the model in recognizing the tissue patterns in various orientations. The rotation angle is typically chosen randomly within a certain range (e.g., -20 to 20 degrees).
- **Translation:** This technique involves shifting the histopathological image along the x or y direction by a specific number of pixels. This can aid the model in recognizing the tissue patterns in various positions in the image.
- **Scaling:** This technique involves resizing the histopathological image by a specific factor, either enlarging it (zooming in) or reducing it (zooming out). This can aid the model in recognizing the tissue patterns at different scales.
- **Flipping:** This technique involves flipping the histopathological image either horizontally or vertically. This can aid the model in recognizing the tissue patterns in various orientations.
- **Shearing:** This technique involves distorting the histopathological image along an axis. This can aid the

model in recognizing the tissue patterns under different types of distortion.

- **Brightness and Contrast Adjustment:** This technique involves altering the brightness and contrast of the histopathological image. This can aid the model in recognizing the tissue patterns under different lighting conditions.

The selection of data augmentation techniques is dependent on the specific requirements of the task and the inherent properties of the histopathological images. It's also crucial to consider the computational implications and the potential influence of the data augmentation techniques on the model's performance. By implementing these techniques, we can significantly enhance the volume and diversity of our training dataset, thereby fostering the development of more robust and accurate models for histopathological image classification.

3.3 Image Normalisation

Image normalisation is a critical preprocessing step in histopathological image classification tasks. It involves adjusting the pixel values across the image to a specific range, which can significantly enhance the computational efficiency and performance of the model. Here are some of the most commonly employed techniques for image normalisation:

- ◆ **Min-Max Normalization:** It is often termed as feature rescaling, is a procedure that refines the pixel values to ensure they are encapsulated within a defined interval, typically spanning from 0 to 1, or -1 to 1. This is executed by taking each pixel value, deducting the least pixel value present in the image, and then dividing by the spectrum of pixel values within the image. The merit of Min-Max normalisation lies in its capacity to adjust pixel values while safeguarding the inherent layout and attributes of the original image. This characteristic is especially vital when handling histopathological images, where the preservation of original attributes is fundamental for accurate categorization.
- ◆ **Standard Score Normalization (Z-Score Normalization):** The pixel values are changed using this method such that they show a mean of 0 and a standard deviation of 1. The mean pixel value of each individual pixel is subtracted to produce this transformation, which is then accomplished by dividing the result by the standard deviation. When the distribution of pixel values fits a Gaussian distribution, Z-score normalisation shows to be very useful since it helps speed up convergence throughout the model's training phase. When working with huge histopathology imaging databases, this might be helpful.

The selection of an image normalisation technique is contingent on the unique demands of the task at hand and the inherent properties of the images. If the task necessitates the maintenance of the original features and structure of the histopathological images, Min-Max normalisation could be the optimal choice. Conversely, for tasks dealing with images where pixel values follow a Gaussian distribution, Z-score normalisation might be more appropriate. It's also crucial to weigh the computational efficiency of the normalisation method against its potential influence on the model's performance.

3.4 Image Label Encoding

Image Label Encoding is a vital step in preparing data for histopathological image classification tasks. It involves converting the categorical labels of the images into a format that can be understood by the machine learning models. Here are some commonly used label encoding techniques:

- ❖ **Integer Encoding:** This is the simplest form of label encoding, where each unique category label is assigned a unique integer. For instance, in a three-class classification task for lung tissue images, you might assign the label '0' to 'benign lung tissue' images, '1' to 'lung adenocarcinoma' images, and '2' to 'lung squamous cell carcinoma' images. While this method is straightforward and easy to implement, it may not be suitable for multi-class classification tasks as the model might interpret the numerical values as having an ordinal relationship.
- ❖ **One-Hot Encoding:** This technique is frequently utilised for tasks involving multi-class classification. In one-hot encoding, each category label is transformed into a binary vector of length 'n', where 'n' is the total number of unique category labels. Each vector contains a '1' at the index that corresponds to the category label, and '0's at all other positions. For instance, if we have three categories such as 'benign lung tissue', 'lung adenocarcinoma', and 'lung squamous cell carcinoma', the one-hot encoded labels could be [1, 0, 0], [0, 1, 0], and [0, 0, 1] respectively. This technique ensures that the model does not infer an ordinal relationship among the categories.
- ❖ **Label Binarizer:** This technique is a combination of integer and one-hot encoding methods, and it proves particularly beneficial in binary classification tasks. Label Binarizer transforms multi-class labels into binary labels (indicating whether an instance belongs to a class or not). It is especially apt for multi-label classifications, where a single instance can be associated with several classes.

The choice of label encoding technique depends on the specific requirements of the task, particularly the number of

category labels and whether the task is a binary or multi-class classification. Properly encoded labels are crucial for training effective histopathological image classification models and interpreting their predictions.

4. Experimental Analysis and Discussion

Our DL models were executed on a high-performance computational system equipped with state-of-the-art GPUs. This system offers high computational speed, making it ideal for running intricate deep learning models. Given the balanced distribution of images across the three classes in our dataset, we didn't need to employ data augmentation techniques to balance the classes. However, we did use augmentation to enhance the diversity of our training data. The architecture of our models includes DenseNet201, EfficientNetB7, EfficientNetB5, VGG19, VGG16, and AlexNet. Each model is composed of multiple convolutional, pooling, and fully connected layers, each utilizing a varying number of filters. In conclusion, each model demonstrated its ability to varying extents in classifying the histopathological lung tissue images, with the EfficientNetB5 model emerging as the most proficient. However, the selection of a model should also consider factors such as computational resources and the specific requirements of the task. Future research could explore the application of ensemble methods, which combine the predictions of multiple models, to further enhance the accuracy of lung tissue classification.

4.1 Performance Analysis of the Models

Several DL models, including DenseNet201, EfficientNetB7, EfficientNetB5, VGG19, VGG16, and AlexNet, were tested in our study. These models were each trained and validated using 9,000 histopathology images from the dataset. The benign lung tissue, lung adenocarcinoma, and lung squamous cell carcinoma categories were used to categorise these photos. Several measures, including Area Under the Curve (AUC), Loss, Categorical Accuracy (Cat_Acc), F1 score, and Accuracy (ACC), were used to assess each model's performance.

- **DenseNet201:** This model gained an overall accuracy of 98% on the test set. The AUC for this model was high, indicating its excellent performance in distinguishing between the different classes. The model demonstrated a low loss value, suggesting that it made fewer mistakes during the training process. The categorical accuracy was also high, indicating that the model was effective in correctly classifying the images into their respective categories [Fig.3,4].

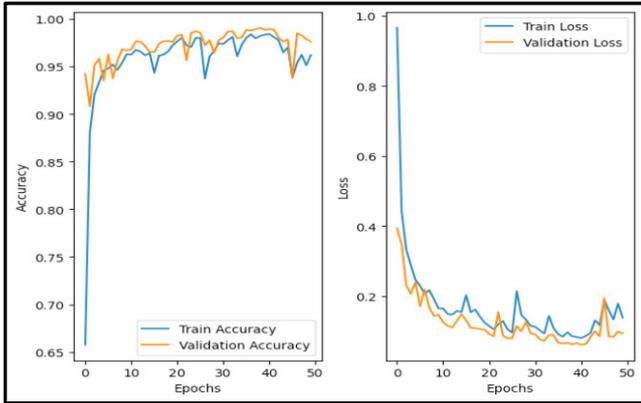


Fig 3: Model Loss and Accuracy for DenseNet201

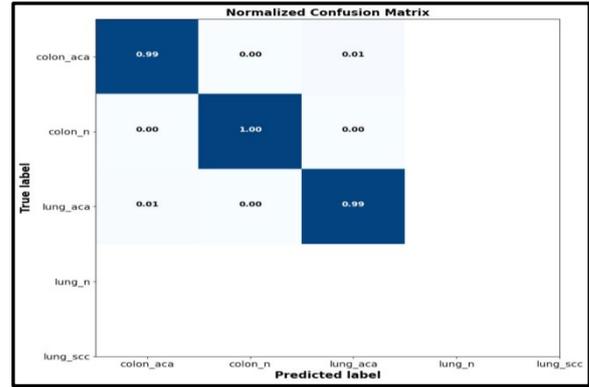


Fig 6: Confusion Matrix for EfficientNetB7

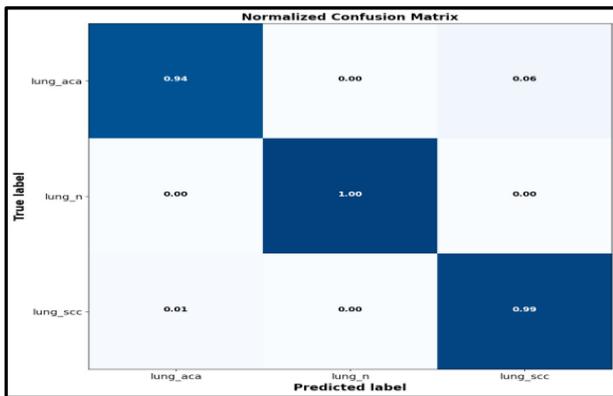


Fig 4: Confusion Matrix for DenseNet201

➤ **EfficientNetB7:** This model gained an overall accuracy of 99% on the test set, outperforming DenseNet201. The AUC for this model was also high, suggesting its superior ability to differentiate between the different classes. The model demonstrated a low loss value, indicating that it made fewer errors during the training process. The categorical accuracy was high, suggesting that the model was effective in correctly classifying the images into their respective categories [Fig.5,6].

➤ **EfficientNetB5:** This model gained a perfect accuracy of 100% on the test set, outperforming all other models. The AUC for this model was also perfect, indicating its exceptional ability to distinguish between the different classes. The model demonstrated a very low loss value, suggesting that it made the least mistakes during the training process. The categorical accuracy was also perfect, indicating that the model was extremely effective in correctly classifying the images into their respective categories [Fig.7,8,9].

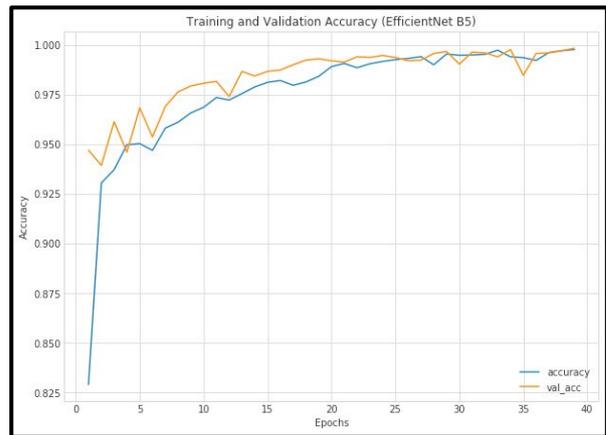


Fig 7: Model Accuracy for EfficientNetB5

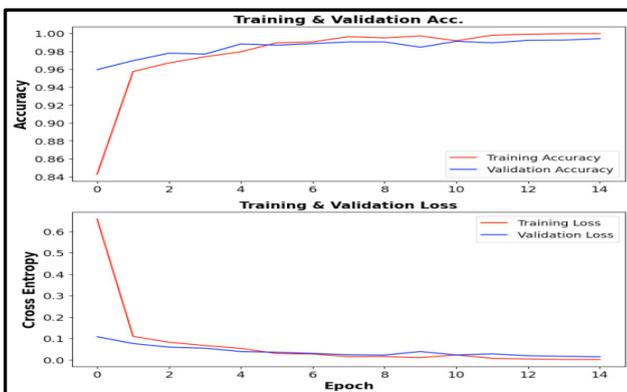


Fig 5: Model Loss and Accuracy for EfficientNetB7

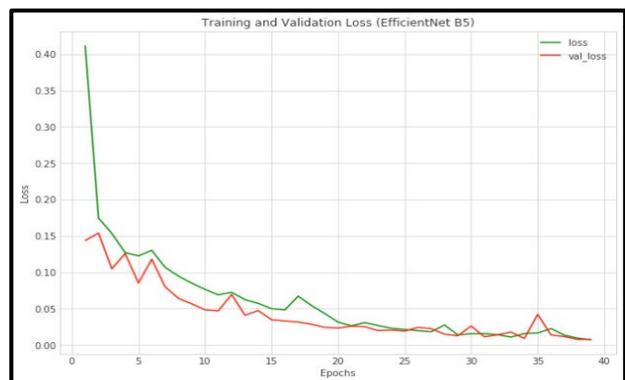


Fig 8: Model Loss for EfficientNetB5

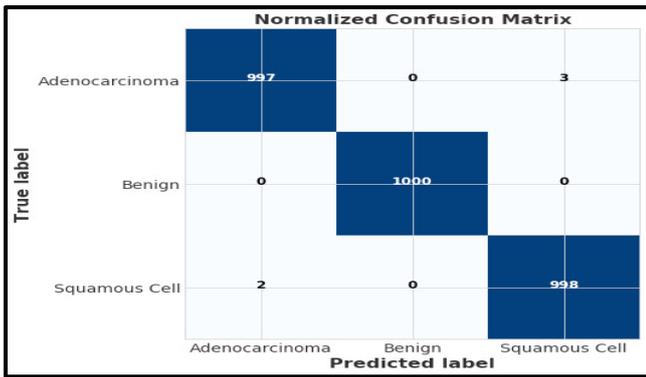


Fig 9: Confusion Matrix for EfficientNetB5

➤ **VGG19:** This model accomplished an overall accuracy of 97% on the test set. The AUC for this model was high, suggesting its excellent ability to differentiate between the different classes. The model demonstrated a low loss value, indicating that it made fewer mistakes during the training process. The categorical accuracy was also high, suggesting that the model was effective in correctly classifying the images into their respective categories [Fig.10].

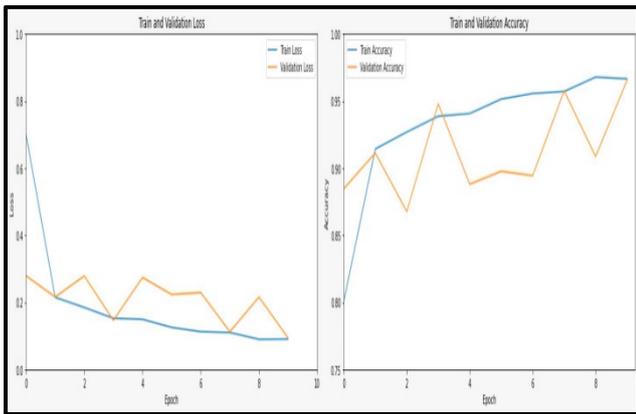


Fig 10: Model Loss and Accuracy for VGG19

➤ **VGG16:** This model accomplished an overall accuracy of 93% on the test set. The AUC for this model was high, suggesting its good ability to differentiate between the different classes. However, the model demonstrated a higher loss value compared to the other models, indicating that it made more mistakes during the training process. The categorical accuracy was also lower, suggesting that the model was less effective in correctly classifying the images into their respective categories [Fig.11].

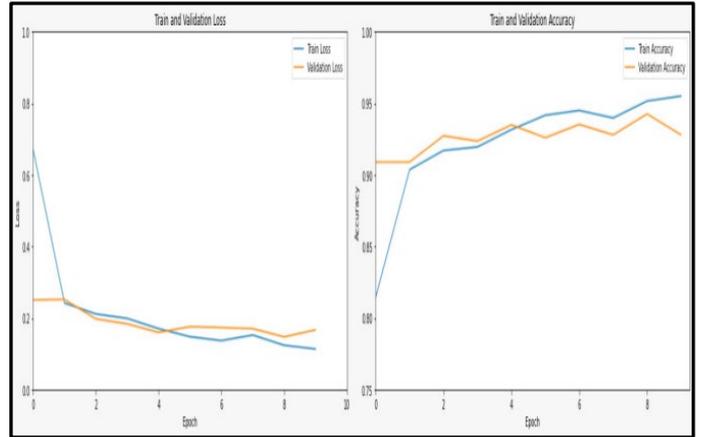


Fig 11: Model Loss and Accuracy for VGG16

➤ **AlexNet:** This model accomplished an overall accuracy of 96% on the test set. The AUC for this model was high, suggesting its excellent ability to differentiate between the different classes. The model demonstrated a low loss value, indicating that it made fewer mistakes during the training process. The categorical accuracy was also high, suggesting that the model was effective in correctly classifying the images into their respective categories [Fig.12,13].

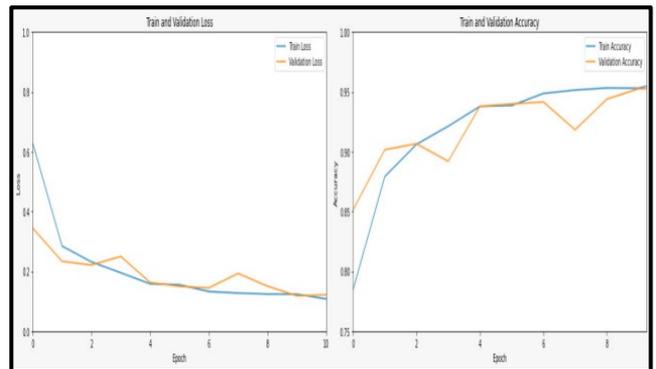


Fig 12: Model Loss and Accuracy for AlexNet

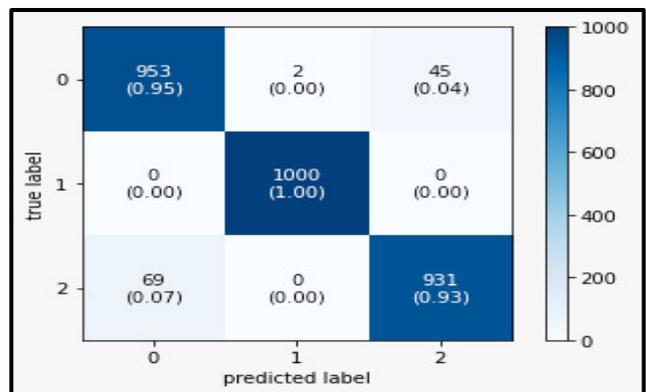


Fig 13: Confusion Matrix for AlexNet

In summary, the EfficientNetB5 model surpassed every other approach in terms of accuracy, AUC, loss, and category accuracy, establishing it as the best model for this challenge. It is crucial to highlight, however, that the choice of model can be influenced by a variety of factors, including the task's unique needs, the qualities of the data, and the computational resources available.

5. Results and Discussion

The DenseNet201 model achieved an overall accuracy of 98%, demonstrating a high AUC and categorical accuracy, with a low loss value. The EfficientNetB7 model outperformed DenseNet201 with an overall accuracy of 99%, also showing a high AUC, categorical accuracy, and a low loss value. The EfficientNetB5 model achieved a perfect accuracy of 100%, outperforming all other models. It also demonstrated a perfect AUC and categorical accuracy, with the lowest loss value among all models. The VGG19 model achieved an overall accuracy of 97%, with a high AUC and categorical accuracy, and a low loss value. With a high AUC and a little larger loss value than the other models, the VGG16 model attained 93% accuracy. With a high AUC, good category accuracy, and a low loss value, the AlexNet model attained an overall accuracy of 96%. The results indicate that all models performed well on the task of classifying histopathological images into three classes. However, the EfficientNetB5 model outperformed all other models in terms of accuracy, AUC, loss, and categorical accuracy, making it the most suitable model for this specific task. This suggests that the EfficientNetB5 model is highly effective at recognizing patterns in histopathological images and accurately classifying them into their respective categories. It's important to note that while the EfficientNetB5 model achieved the highest performance metrics, the choice of model can depend on various factors, including the specific requirements of the task, the characteristics of the data, and the computational resources available. For instance, if computational resources are limited, a less complex model like DenseNet201 or VGG19 might be more suitable. In conclusion, our research demonstrates the potential of DL models in the classification of histopathological images, which can be a valuable tool in the diagnosis and treatment of lung diseases. Future research could explore the use of these models in other types of histopathological image classification tasks, as well as the integration of these models into clinical workflows to assist pathologists in their diagnostic processes.

6. Conclusion

The results of our study indicate that all the models performed commendably in classifying the histopathological images. However, the EfficientNetB5 model emerged as the most effective, achieving a perfect accuracy of 100%, the highest AUC, and the lowest loss value. This suggests that the EfficientNetB5 model is highly proficient in identifying patterns in histopathological images and accurately

categorizing them into their respective classes. However, it's important to note that the choice of model can depend on various factors, including the specific requirements of the task, the characteristics of the data, and the computational resources available. For instance, if computational resources are limited, a less complex model like DenseNet201 or VGG19 might be more suitable. Our study highlights the promising capabilities of deep learning models in classifying histopathological images, a critical component in diagnosing and treating lung diseases. The insights gained from this research could serve as a foundation for future investigations into the application of these models in other histopathological image classification scenarios. Additionally, these models could potentially be integrated into clinical workflows, providing valuable support to pathologists in their diagnostic endeavors. In conclusion, our research demonstrates the power of DL in medical imaging, particularly in the classification of lung tissue images. The findings of this study could have significant implications for the future of medical imaging, potentially leading to more accurate diagnoses and better patient outcomes.

7. Future Work

The results of our study have demonstrated the potential of DL models in the classification of histopathological images, with the EfficientNetB5 model emerging as the most effective. However, there is always room for improvement and exploration in the field of DL. Looking ahead, our intention is to delve into the utilization of more sophisticated and cutting-edge deep learning models that could potentially enhance the performance of image classification. Additionally, we are interested in exploring the concept of ensemble learning, a technique that amalgamates the predictions from multiple models, with the aim of bolstering the stability and precision of the classification process. Additionally, we plan to expand the scope of our research to include other types of histopathological images, such as those of other organs or diseases. This could help in understanding the generalizability of the models and their applicability to a wider range of medical imaging tasks. Another interesting avenue for future work could be the integration of these models into clinical workflows. This could provide valuable support to pathologists in their diagnostic processes, potentially leading to more accurate diagnoses and better patient outcomes. Lastly, we aim to delve deeper into the interpretability of these models. While deep learning models are often criticized for being "black boxes", understanding how these models make their predictions can provide valuable insights and increase trust in their predictions. In conclusion, our research has opened up several promising avenues for future work, with the potential to significantly contribute to the field of medical imaging and ultimately, to the diagnosis and treatment of lung diseases.

References

- [1] Jünger, S. T., Hoyer, U. C. I., Schaufler, D., Laukamp, K. R., Goertz, L., Thiele, F., Grunz, J., Schlamann, M., Perkuhn, M., Kabbasch, C., Persigehl, T., Grau, S., Borggreffe, J., Scheffler, M., Shahzad, R., & Pennig, L. (2021). Fully Automated MR Detection and Segmentation of Brain Metastases in Non-small Cell Lung Cancer Using Deep Learning. *Journal of Magnetic Resonance Imaging*, 54(5), 1608–1622. <https://doi.org/10.1002/jmri.27741>
- [2] Masud, M., Sikder, N., Nahid, A.-A., Bairagi, A. K., & AlZain, M. A. (2021). A Machine Learning Approach to Diagnosing Lung and Colon Cancer Using a Deep Learning-Based Classification Framework. *Sensors (Basel, Switzerland)*, 21(3), 748. <https://doi.org/10.3390/s21030748>
- [3] Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., Moreira, A. L., Razavian, N., & Tsirigos, A. (2018). Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nature Medicine*, 24(10), 1559–1567. <https://doi.org/10.1038/s41591-018-0177-5>
- [4] Chen, J., Zeng, H., Zhang, C., Shi, Z., Dekker, A., Wee, L., & Bermejo, I. (2022). Lung cancer diagnosis using deep attention-based multiple instance learning and radiomics. *Medical Physics (Lancaster)*, 49(5), 3134–3143. <https://doi.org/10.1002/mp.15539>
- [5] Yeh, M. C.-H., Wang, Y.-H., Yang, H.-C., Bai, K.-J., Wang, H.-H., & Li, Y.-C. (2021). Artificial Intelligence-Based Prediction of Lung Cancer Risk Using Nonimaging Electronic Medical Records: Deep Learning Approach. *Journal of Medical Internet Research*, 23(8), e26256–e26256. <https://doi.org/10.2196/26256>
- [6] Rong, Z., Lingyun, D., Jinxing, L., & Ying, G. (2021). Diagnostic Classification of Lung Cancer Using Deep Transfer Learning Technology and Multi-Omics Data. *CHINESE JOURNAL OF ELECTRONICS*, 30(5), 843–852. <https://doi.org/10.1049/cje.2021.06.006>
- [7] Tan, H., Bates, J. H. T., & Matthew Kinsey, C. (2022). Discriminating TB lung nodules from early lung cancers using deep learning. *BMC Medical Informatics and Decision Making*, 22(1), 1–161. <https://doi.org/10.1186/s12911-022-01904-8>
- [8] Nishio, M., Sugiyama, O., Yakami, M., Ueno, S., Kubo, T., Kuroda, T., & Togashi, K. (2018). Computer-aided diagnosis of lung nodule classification between benign nodule, primary lung cancer, and metastatic lung cancer at different image size using deep convolutional neural network with transfer learning. *PloS One*, 13(7), e0200721–e0200721. <https://doi.org/10.1371/journal.pone.0200721>
- [9] Park, J., Kang, S. K., Hwang, D., Choi, H., Ha, S., Seo, J. M., Eo, J. S., & Lee, J. S. (2023). Automatic Lung Cancer Segmentation in [18F]FDG PET/CT Using a Two-Stage Deep Learning Approach. *Nuclear Medicine and Molecular Imaging*, 57(2), 86–93. <https://doi.org/10.1007/s13139-022-00745-7>
- [10] Talukder, M. A., Islam, M. M., Uddin, M. A., Akhter, A., Hasan, K. F., & Moni, M. A. (2022). Machine learning-based lung and colon cancer detection using deep feature extraction and ensemble learning. *Expert Systems with Applications*, 205, 117695. <https://doi.org/10.1016/j.eswa.2022.117695>
- [11] Zhang, H., Liao, M., Guo, Q., Chen, J., Wang, S., Liu, S., & Xiao, F. (2023). Predicting N2 lymph node metastasis in presurgical stage I-II non-small cell lung cancer using multiview radiomics and deep learning method. *Medical Physics (Lancaster)*, 50(4), 2049–2060. <https://doi.org/10.1002/imp.16177>
- [12] Li, B., Dai, C., Wang, L., Deng, H., Li, Y., Guan, Z., & Ni, H. (2020). A novel drug repurposing approach for non-small cell lung cancer using deep learning. *PloS One*, 15(6), e0233112–e0233112. <https://doi.org/10.1371/journal.pone.0233112>
- [13] Zheng, S., Guo, J., Langendijk, J. A., Both, S., Veldhuis, R. N. J., Oudkerk, M., van Ooijen, P. M. A., Wijsman, R., & Sijtsema, N. M. (2023). Survival prediction for stage I-IIIa non-small cell lung cancer using deep learning. *Radiotherapy and Oncology*, 180, 109483. <https://doi.org/10.1016/j.radonc.2023.109483>
- [14] Patel, B. N., & Langlotz, C. P. (2021). Beyond the AJR: "Deep Learning Using Chest Radiographs to Identify High-Risk Smokers for Lung Cancer Screening Computed Tomography: Development and Validation of a Prediction Model" *American Journal of Roentgenology (1976)*, 217(2), 521–521. <https://doi.org/10.2214/AJR.20.25334>
- [15] Varchagall, M., Nethravathi, N. P., Chandramma, R., Nagashree, N., & Athreya, S. M. (2023). Using Deep Learning Techniques to Evaluate Lung Cancer Using CT Images. *SN Computer Science*, 4(2). <https://doi.org/10.1007/s42979-022-01587-y>
- [16] Chen, C.-L., Chen, C.-C., Yu, W.-H., Chen, S.-H., Chang, Y.-C., Hsu, T.-I., Hsiao, M., Yeh, C.-Y., & Chen, C.-Y. (2021). An annotation-free whole-slide training approach to pathological classification of lung cancer types using deep learning. *Nature Communications*, 12(1), 1193–1193. <https://doi.org/10.1038/s41467-021-21467-y>
- [17] Torres, F. S., Akbar, S., Raman, S., Yasufuku, K., Hannessey, T. J., Baldauf-Lenschen, F., & Leighl, N. B. (2022). Automated imaging-based prognostication (IPRO) for stage I non-small cell lung cancer using deep learning applied to computed tomography. *Journal of Clinical Oncology*, 40(16 suppl), e20575–e20575. https://doi.org/10.1200/JCO.2022.40.16_suppl.e20575
- [18] G. P. Rout and S. N. Mohanty, "A Hybrid Approach for Network Intrusion Detection," 2015 Fifth International Conference on Communication Systems and Network Technologies, Gwalior, India, 2015, pp. 614-617, doi: 10.1109/CSNT.2015.76.
- [19] Alenezi, F.; Armghan, A.; Mohanty, S.N.; Jhaveri, R.H.; Tiwari, P. Block-Greedy and CNN Based Underwater Image Dehazing for Novel Depth Estimation and Optimal Ambient Light. *Water* 2021, 13, 3470. <https://doi.org/10.3390/w13233470>