International Journal of Intelligence Multidisciplinary Engineering Research(IJIMER) *Vol. 1 (4), Dec 2023, pp: 01-11.*
ISSN: 2583-9160

An Efficient Model for Optimizing Hyperparameters in AlexNet for Precise Malignancy Detection in Lung and Colon Histopathology Images with CSIP-EHE

D Mythrayee[1*], Shapaaz Shaik[2]
[1,2]Department of Computer Science and Engineering,
RK College of Engineering, Gurukul School Road, Kethanakonda, Andhra Pradesh 521456.
[1]mythra2325@gmail.com, [2]shabaaz.rkce@gmail.com

Abstract - **Cancer is a lethal disease stemming from genetic anomalies and biochemical irregularities, presents a major global health challenge, with lung and colon cancers being significant contributors to morbidity and mortality. Timely and precise cancer detection is crucial for optimal treatment decisions, and machine learning and deep learning techniques offer a promising solution for expediting this process. In this research, a pre-trained neural network, specifically AlexNet, was fine-tuned with modifications to four layers to adapt it to a dataset comprising histopathological images of lung and colon tissues. Additionally, a Bayesian optimization approach was employed for hyperparameter tuning in Convolutional Neural Networks (CNNs) to enhance recognition accuracy while maintaining computational efficiency. The research utilized a comprehensive dataset divided into five classes, and in cases of suboptimal results, a Counteracting Suboptimal Image Processing (CSIP) strategy was applied, focusing on improving images of underperforming classes to reduce processing time and effort.**

Keywords: *Boruta Feature Selection, Coronary Heart Disease, Decision tree, Machine Learning, KNN, Logistic regression, SMOTE test, SVM.*

1. Introduction

Cancer prevention is a crucial aspect in the battle against cancer, underscoring the significance of early diagnosis across all cancer types. However, experts view the precise identification of cancer types and the swift generation of results as challenging and time-consuming. To address these challenges, it is imperative to stay abreast of technological advancements and incorporate them into the diagnostic process. It's worth noting that each optimization algorithm may not consistently yield the initially targeted point or population with the best features.

Cancer manifests as the uncontrollable growth of abnormal cells in any organ or tissue of the body, representing a leading cause of death globally. In 2018 alone, cancer accounted for an estimated 9.6 million deaths, or one in every six deaths. Lung cancer, comprising both small cell and non-small cell types, contributed to 2.06 million cases and 1.76 million deaths. Colorectal cancer, covering both colon and rectal cancer, constituted 1.80 million cases and 783,000 deaths.

Lung cancer is categorized into small-cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). SCLC, comprising 15% of total cases, is a highly aggressive tumour with neuroendocrine characteristics. NSCLC, constituting the remaining 85%, further divides into adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. Colorectal cancer, specifically adenocarcinoma, represents 96% of all cases and encompasses both colon and rectal cancer. Recent advancements involve digitizing entire tissue or cell slides using scanners, resulting in a plethora of whole slide images (WSIs). Machine learning algorithms are then applied to analyze these WSIs for diagnostic purposes.

1.1 Transfer Learning Using Pre-Trained ALEXNET Model

Transfer Learning is a method that leverages an existing model to transfer knowledge from one domain to another. This technique is especially useful for domain adaptation and improving the accuracy of models trained on smaller datasets. The effectiveness of transfer learning depends on several factors, such as the similarity between your dataset and the dataset used to train the original model, the size of your dataset, and the available computational resources. The closer the match between your dataset and the one used in the original model, the higher the likelihood that the learned parameters and architecture of the model will be beneficial for your data.

In the paper, we applied transfer learning to develop convolutional neural networks (CNNs) for cancer classification. We specifically focused on a form of transfer learning that involves fine-tuning parts of an existing model to better suit our dataset. This process entails modifying certain parameters of the model while keeping others unchanged. Initially, we trained an AlexNet Model from scratch with our dataset, adjusting all the parameters. In the subsequent phase, we evaluated the performance of a classifier by fine-tuning an AlexNet model that was previously trained on the ImageNet Dataset. This step involved replacing and retraining the parameters of the output, fully-connected layer of the pre-trained model, while the remaining layers were left unchanged. By doing so, we aimed to enhance the model's ability to classify cancer effectively, using the foundational learning from the ImageNet Dataset as a starting point.

1.2 Background of Convolutional Neural Networks

A Convolutional Neural Network (CNN) is a specialized type of Neural Network designed for image processing and classification. Its input consists of pixel values from an image presented in vector/matrix form. The CNN processes this input through a series of layers and generates a classification for the image.

The layers in Convolutional Neural Networks typically include four types:

1. Convolutional Layer: This layer identifies patterns in the image by passing its representative matrix through learnable filters/kernels, each representing distinct visual features in the image. These filters slide over the image based on specified strides, producing individual feature maps. The layer's final output is a transformation of the original image, comprising all the stacked feature maps.

2. Rectified Linear Unit Layer (ReLU): This non-linear activation function, denoted as f(x) = max(x,0), transforms the output elements of the convolutional layer. By replacing negative values with 0 without altering the shape, ReLU adjusts the output to a range from 0 to infinity.

3. Pooling Layer: This layer conducts down-sampling along the spatial dimensions of the image, reducing its representation size. By diminishing the number of features in the CNN, the model enhances computational efficiency while preserving key image features.

4. Fully-Connected Layer: Unlike the local connections made by the convolutional layer, each node in a Fully-Connected Layer establishes connections with all nodes in the preceding layer.

1.3 Evolution of Pre-Trained Models

Before the advent of Convolutional Neural Networks (CNNs), image processing primarily revolved around techniques like edge detection and other methods for extracting features based on raw pixel information. Subsequently, significant advancements in CNN architectures and increased computer processing power have substantially elevated the accuracy of CNNs in image processing.

While various pre-trained models exist, our focus centered on AlexNet. This choice stems from AlexNet being the initial prominent model to incorporate the convolutional layers defining a CNN. This decision aimed to provide us with a comprehensive understanding of the construction and functioning of CNN models. Additionally, our computational constraints influenced our choice, as training a larger or deeper model like VGG16 would demand more computational power than was available. Nevertheless, the use of transfer learning serves to significantly mitigate the computational expense associated with constructing and training a CNN.

1.4 ALEXNET

Developed in 2012, AlexNet marked a significant advancement in CNN evolution. Not only did it emerge as the victor in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) competition of that year, but it also boasted an error rate approximately half that of its competitors. The key innovations encompassed training on multiple GPUs, employing augmented versions of image data for training, adopting the ReLU activation function, incorporating overlapping pools, and implementing dropout.

AlexNet's architecture consists of a total of 60,000 parameters distributed across eight layers, comprising five convolutional layers and three fully connected layers. Further innovations included training on two GPUs and incorporating augmented versions of images (flipped, scaled, noised, etc.) for training purposes. The model also embraced ReLU (Rectified Linear Unit) activation functions, departing from the standard tanh (hyperbolic tangent) at the time. This adjustment not only reduced the training time but also served as a solution to the "vanishing gradient" problem. The pooling layers introduced a stride (in AlexNet, with a length of 4 pixels), resulting in an overlap between local receptive fields and significantly minimizing model errors.

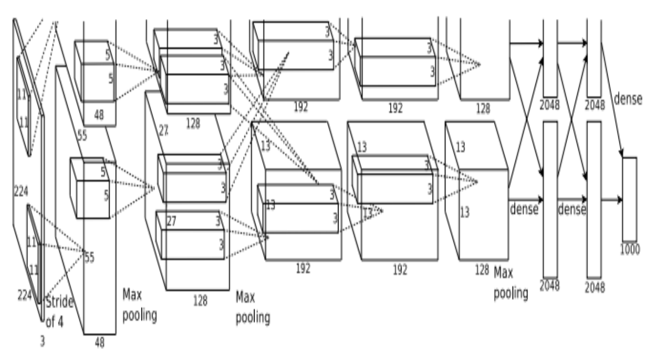


Figure 1 AlexNet Model from: Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. “ImageNet classification with deep convolutional neural networks

1.5 Bayesian Optimization Algorithm

The Bayesian optimization algorithm seeks to minimize a scalar objective function f(x) over a bounded domain, where x represents the input. This function may be deterministic or stochastic, implying that it can yield varying outcomes when assessed at the same point x.

1.5.1 Bayesian Optimization

Bayesian optimization is recognized as an exceptionally versatile strategy for optimizing expensive "black box" functions without derivatives. The term "black box" refers to the situation where only the input and output of the model are observable, and the internal workings of the model remain unclear. In recent years, Bayesian optimization has found widespread application in various domains, including environmental monitoring, interactive user interfaces, materials and drug design, machine learning, and reinforcement learning, owing to its robust optimization capabilities.

Although hand-tuning is a viable option, it is often perceived as time-consuming and subjective. Grid Search and Random Search are alternative optimization methods that do not require direct human intervention. Grid search exhaustively explores a predefined subset of hyperparameter spaces, while random search, without optimizing the problem gradient, selects hyperparameters randomly within the specified search space. Both methods, however, can be time-consuming. Bayesian global optimization techniques offer a solution to this problem by efficiently combining priors of the problem, aiding in deciding whether the next point in the search space should be explored or exploited.

1.5.2 BAYESIAN OPTIMIZATION STEPS

Develop a surrogate probability model for the objective function.

• Identify the hyperparameters that exhibit optimal performance on the surrogate model.

• Implement these identified hyperparameters on the actual objective function.

• Update the surrogate model to integrate the new outcomes.

• Iterate through steps 2–4 until reaching the maximum iterations or the allotted time.

Bayesian analysis finds application in various domains where abundant heterogeneous or noisy data is present, or whenever a comprehensive understanding of uncertainty is essential.

2. RELATED WORKS

In the research [1],[2], Reyhaneh Manafi-Farid and Emran Askari highlight the critical role of fluorodeoxyglucose Positron Emission Tomography and Computed Tomography (FDG-PET/CT) in lung cancer detection and management. Lung cancer, known for its high mortality rate, necessitates innovative diagnostic and prognostic tools. The authors deep into the utility of FDG-PET/CT in diagnosing lung cancer, assessing treatment response, and predicting outcomes. They emphasize the technique's ability to provide qualitative and conventional quantitative indices from imaging data.

Additionally, the authors shed light on the emerging field of radiomics, which involves sophisticated algorithms to extract detailed data from medical images. Radiomics is increasingly significant in enhancing the diagnostic capabilities and therapeutic implications of FDG-PET/CT in lung cancer treatment. The article offers an overview of the technical aspects of radiomics, discussing its integration into current medical imaging practices and its potential to revolutionize lung cancer management.

In the research [7], Sumeet Hindocha and Thomas G. Charlton's research, conducted in October 2022, centers on the development of radiomic models for predicting outcomes in non-small cell lung cancer (NSCLC) patients undergoing radiotherapy. Their primary objective is to classify patients based on their risk of recurrence and overall survival post-treatment. This approach could pave the way for more personalized surveillance and timely interventions, ultimately improving patient outcomes. Hindocha and Charlton's models amalgamate radiomic and clinical features, validated through extensive cross-validation and external testing. The results indicate that these models effectively stratify patients into low and high-risk groups, with a substantial disparity in survival times between these groups. This research lays the foundation for future clinical trials and underscores the potential of integrating radiomic-based prediction models into routine radiotherapy workflows, facilitating a more personalized approach to cancer treatment.

[8] The article published in 2022 by Hamid Abdollahi and Erika Chin's article explores the integration of radiomics into radiation therapy, a critical component of personalized medicine. Radiomics, through the extraction and analysis of complex image features, can significantly impact various aspects of radiation therapy, from patient selection to post-treatment monitoring. The authors introduce the concept of radiomics-guided radiation therapy (RGRT), emphasizing its potential to optimize treatment protocols, enhance patient outcomes, and minimize side effects.

The article reviews several applications of radiomics in radiation therapy, including its role in disease detection, diagnosis, prognosis, and response assessment. While acknowledging the challenges associated with implementing RGRT, such as data standardization and algorithm validation, it underscores the enormous potential of radiomics to transform radiation therapy into a more precise and effective treatment modality.

The groundbreaking study [2][3], published in 2023, introduces an innovative approach to lung cancer detection using artificial intelligence (AI). They developed a deep learning model named Deep Radial Recurrent Feedforward Neural Nets (DRRFNN), specifically tailored for lung cancer classification. Leveraging the power of deep learning and Python programming, this model achieves high accuracy in diagnosing lung cancer.

The authors conduct a comparative analysis, pitting DRRFNN against existing models like LSTM, GRUs, RBF, DBN, FNN, and ANN, ultimately demonstrating its superior performance. This research underscores the potential of AI and deep learning in revolutionizing medical diagnostics, particularly in the early detection of lung cancer. By employing advanced models like DRRFNN, the methodology for detecting lung cancer could become faster, more efficient, and potentially life-saving by identifying more patients at an early stage.

The research addresses the rising incidence of Head and Neck Squamous Cell Carcinoma (HNSCC). Their study introduces a novel machine learning method that utilizes radiomic features extracted from CT and PET images to stage the disease. This non-invasive approach has the potential to revolutionize the diagnosis and monitoring of HNSCC, which traditionally relies on clinical evaluation and histopathological examination [5][6].

The authors' method includes a selection step to eliminate dataset redundancy, followed by the application of machine learning algorithms for accurate disease staging. The research demonstrates high accuracy in classifying HNSCC in terms of pN-Stage, pT-Stage, and Overall Stage, underscoring the efficiency of using radiomics in cancer staging. Applied to a diverse patient cohort, this approach shows promise in enhancing early diagnosis and personalized treatment, potentially reducing the need for invasive biopsy procedures.

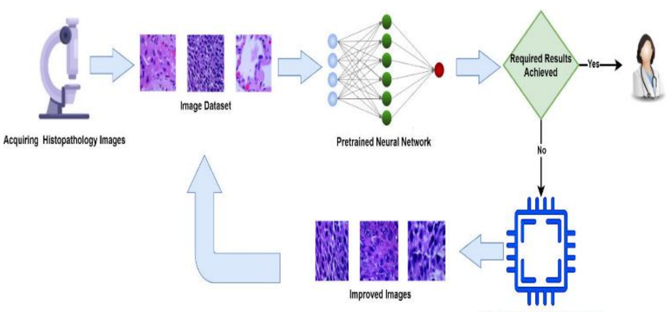


Figure 2 Block Diagram of Model with Convolutional Neural Networks (CNNs)

The authors of [1] considers minimal factors like level of cholesterol, heart rate, age and gender. Hence it has the probability of giving less accuracy. In the paper [3,5,7] continuous monitoring of heart rate has been carried out with patient only after the patient is affected by heart attack through IOT approach which leads to continuous intervention. Conventional approach of detecting the heart attack using ECG and blood test, this approach is not applicable for effective early detection of heart attack. Whereas in [10] the authors have used a logistic regression classification algorithm for heart disease detection and obtained an accuracy of 77.1%. In the paper [11] the authors have used a multi-layer perceptron (MLP) classifier for heart disease diagnosis and attained accuracy of 80%. The heart disease classification system integrated with neural networks and artificial neural network has been addressed in the paper [12,13]. In the paper [14], Naïve Bayes (NB) and Decision Tree (DT) algorithm for the diagnosis and prediction of heart disease has been achieved with reasonable results in terms of accuracy of 82.7% with NB and 80.4% with DT.

3. METHODOLOGIES

In the proposed methodology, we start by acquiring histopathology images and resizing them to fit the model's specifications. These resized images are then used for training and validating the model. The initial outcomes of the model are assessed using metrics such as accuracy, precision, F1-score, recall, specificity, and misclassification rate. If these results are not up to the desired standard, we apply the Customized Selective Image Processing (CSIP) strategy.

The CSIP technique specifically targets the class or classes where the model's performance is suboptimal. For these identified underperforming classes, Enhanced Histogram Equalization (EHE) is employed to improve image quality. This selective approach ensures that only the images in need of enhancement are processed, which significantly reduces the time and effort required compared to processing the entire dataset.

Once the EHE process is applied, it redistributes the intensity values within the images, effectively enhancing the contrast. This step is crucial because improved image contrast can lead to better feature recognition and, consequently, better model performance. After the EHE treatment, all images from the underperforming class are replaced with their enhanced versions in the dataset. The model is then rerun with this updated set of images for training and validation.

The key advantage of this method is its efficiency and targeted nature. By focusing only on the images from classes where the model struggles, the CSIP strategy optimizes resource use. Enhanced Histogram Equalization, known for its efficiency and speed in contrast enhancement, further contributes to this streamlined approach. By enhancing only, the necessary images and reintegrating them into the training and validation process, the model's performance is expected to improve specifically in the areas where it was previously lacking.

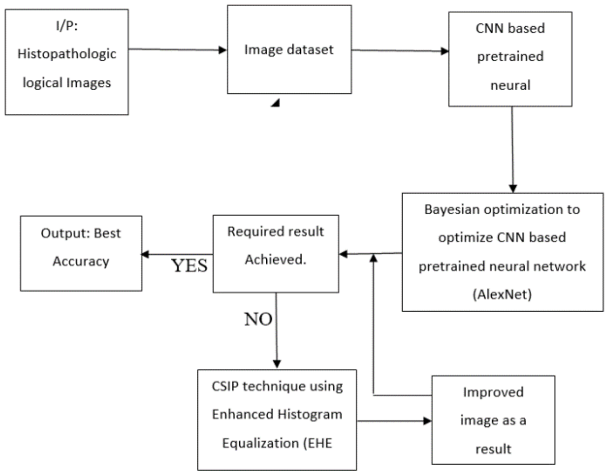


Figure 3 Block diagram of Proposed System

The proposed system offers several significant advantages, particularly in optimizing model performance and enhancing image quality:

Input as Parameter Ranges: Unlike traditional methods where specific points are selected based on assumptions, this system inputs a range for each parameter. This approach is advantageous as it provides a broader scope for the model to identify the most effective parameter values, ensuring a more comprehensive exploration of the parameter space.

Randomization of Candidate Points: To avoid excessive focus on suboptimal parameters, candidate points are randomized. This strategy ensures that the model does not waste time evaluating poor parameter choices. The randomization adds an element of diversity to the parameter selection process, preventing the model from getting stuck in potentially less advantageous regions of the parameter space.

Balanced Exploration and Exploitation through Bayesian Optimization: One of the key strengths of this system is its use of Bayesian Optimization, which adeptly balances exploration (investigating new, potentially better parameters) and exploitation (utilizing known good parameters). This balance is achieved as the algorithm intelligently samples points in the parameter space where it predicts the optimal values are likely to be found. Such a strategic approach can significantly improve the model's efficiency and effectiveness.

Enhanced Contrast in Images: The system is particularly beneficial in situations where images have data represented by close contrast values. It increases the overall contrast of these images by redistributing the intensity values more evenly across the histogram. This redistribution allows areas of the image with lower contrast to achieve higher contrast, which is crucial for better visualization and analysis. Improved contrast makes it easier to discern details and features in images, which is particularly valuable in fields like medical imaging or remote sensing.

Overall, the proposed system offers a sophisticated approach to optimizing model parameters and enhancing image quality, making it a valuable tool in various applications where accuracy and efficiency are paramount.

Bayesian Optimization to Optimize CNN

Bayesian optimization, a powerful tool for hyperparameter optimization in deep learning, is a sequential model-based approach combining a probabilistic surrogate model (prior distribution and observation model) with a loss function to select an optimal sequence of queries, minimizing expected loss.

Bayesian Optimization Libraries

BayesOpt, a Bayesian optimization library under Affero General Public License (AGPL), effectively tackles non-linear and hyperparameter optimization problems. It uses function distribution to establish a proxy model of unknown functions for optimal solution finding and applies active learning to select query points.

3.1 Implementation of Bayesian Optimization on CNN

This research aims to apply Bayesian optimization to CNN models for a three-way classification task, combining it with CNN models to find optimal hyperparameters. The process involves:

Preparing pre-processed images as input data.

Defining the CNN model and network structure.

Defining an objective function taking hyperparameters as input.

Using Bayesian optimization objects to minimize classification error on the validation set.

Obtaining optimal hyperparameters for model classification.

Bayesian optimization was applied to optimize Mini-Batch Size, Epoch, Initial Learning Rate, and Momentum. Preliminary network training determined the approximate search range of hyperparameters. The search ranges were not continuous intervals but discrete values within an approximate range.

The B-CNN models combined selected hyperparameter values to form various combinations, with validation accuracies determining the best model through Bayesian optimization, focusing on minimizing the classification error of the validation set.

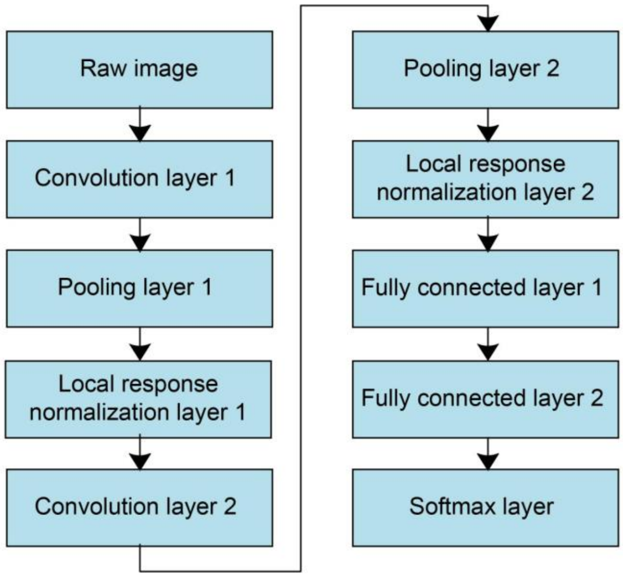


Figure 4 Architecture of the improved CNN

In this research, we present a comprehensive approach that combines Bayesian optimization with machine learning algorithms to enhance the accuracy of a CNN model for cancer detection classification. The process of applying Bayesian optimization to the CNN model is detailed, using the Keras library for building the CNN model and an open-source Bayesian optimization library for the optimization task. Studies have indicated that Bayesian optimization tends to surpass other methods like random search and grid search in efficiency.

Bayesian optimization is particularly advantageous as it is less sensitive to the initial boundary selection. It can adaptively expand the search space, enabling the identification of optimal hyperparameters within predefined bounds to improve model performance. This optimization method integrates prior function distribution with sample data to derive a function's posterior. The optimal values are then determined based on this posterior information and specific criteria.

Enhanced Histogram Equalization (EHE)

Histogram Equalization is a widely-used technique for image improvement, offering more visually appealing results compared to histogram stretching. It aims to flatten the histogram of the resulting image, enhancing both dark and light pixels.

The pre-processing stage using EHE is critical for preparing histopathology images for feature extraction. Issues like blurriness, poor border recognition, artifacts, and overlapping, often due to uneven staining, are addressed by EHE. This technique enhances image contrast by improving poor boundary edges at the pixel level and boosting local contrast. EHE is particularly suitable for histopathological images, enhancing features that are crucial for accurate analysis.

The EHE method involves decomposing the original image into high-frequency and low-frequency components. The low-frequency components are enhanced using EHE, while high-frequency components are left as is to avoid amplifying noise. After reconstruction using inverse DWT, EHE is applied again to further enhance image details.

4. RESULT AND DISCUSSION

Performance evaluation of machine learning models in this research is done using a confusion matrix and metrics like precision, F1-score, accuracy, and recall. Accuracy is defined as the proportion of correctly classified instances (True Positives and True Negatives) out of all instances:

Accuracy = (TP + TN) / (TP + FP + TN + FN)

Precision measures the proportion of true positives among all positives identified:

Precision = TP / (TP + FP)

Recall, or sensitivity, calculates the percentage of actual positives correctly identified:

Recall = TP / (TP + FN)

The F1-score provides a balance between precision and recall, representing their harmonic mean:

F1-score = 2 × (Precision × Recall) / (Precision + Recall)

These metrics are crucial for assessing the model's ability to accurately classify cancerous tissues, which is essential for effective diagnosis and treatment planning.

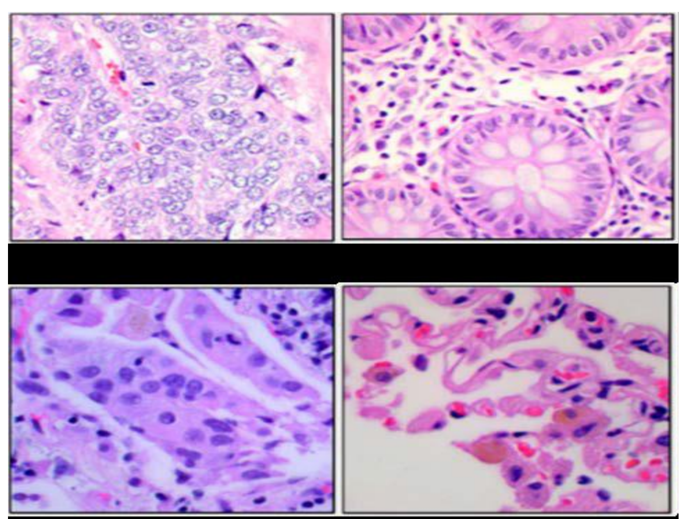
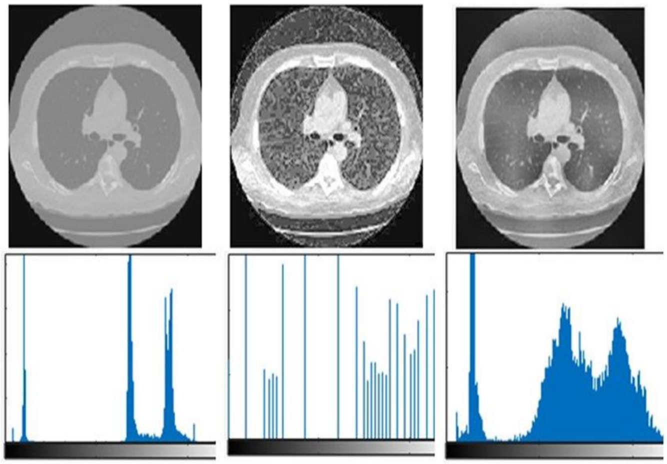


Figure 5 Dataset Sample Images



. Figure 6 Histogram Images for different lung cancer

Lung and colon cancers rank among the leading causes of death globally, making early and accurate detection crucial for enhancing therapeutic outcomes and increasing survival rates. The primary objective of this research was to develop a method for the efficient and precise diagnosis of lung and colon cancers.

A key component of this research was the focus on image quality improvement, specifically through image contrast enhancement. The chosen method for this purpose was histogram equalization, recognized for its efficiency and computational speed. Histogram Equalization (HE) is particularly effective in adjusting the contrast of images, making it easier to identify and analyze key features in medical imaging.

To optimize the performance of our cancer detection model, we implemented an Enhanced Histogram Equalization (EHE) technique. However, instead of applying EHE to the entire dataset, we strategically targeted only the images from underperforming classes. This selective approach was designed to save time and reduce computational costs, while still significantly improving the classification accuracy of the model.

Our proposed methodology, which combines targeted contrast enhancement with sophisticated machine learning algorithms, demonstrated promising results. When benchmarked against existing methods in lung and colon cancer detection, our approach showed an improvement in detecting these cancers. This advancement in early cancer detection methodology could be a pivotal step in medical diagnostics, offering a more effective tool for healthcare professionals in the fight against these prevalent cancers.

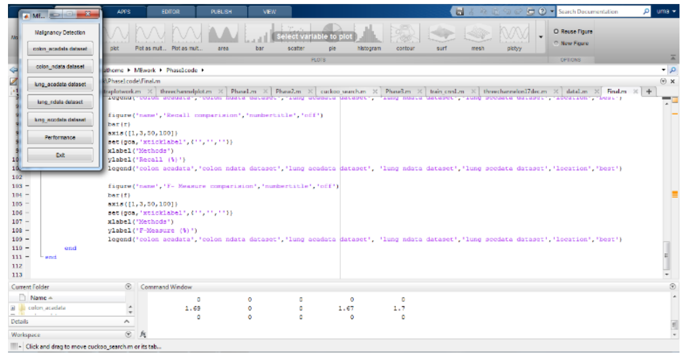


Figure 6 Hyperparameter Optimization In Alexnet For Malignancy Detection In Lung And Colon Histopathology Images Using Csip-EHE

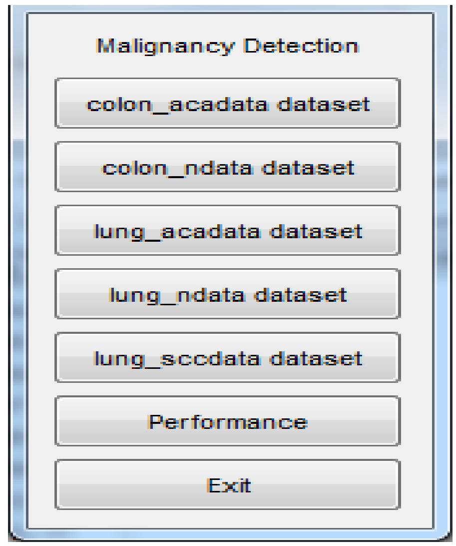


Figure 7 Main Menu

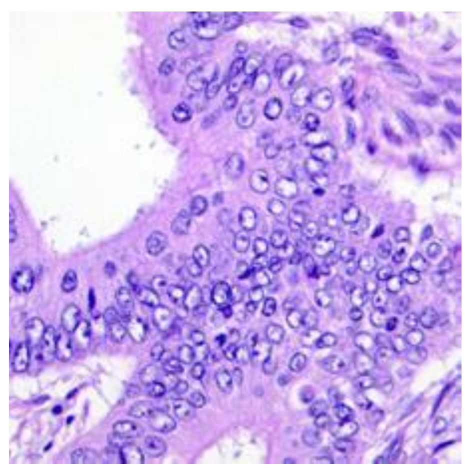


Figure 9 Colon aca data

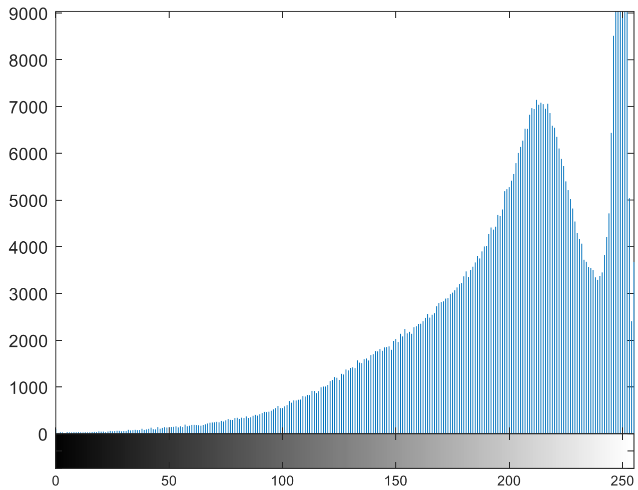


Figure 11 Histogram Image of Colon aca

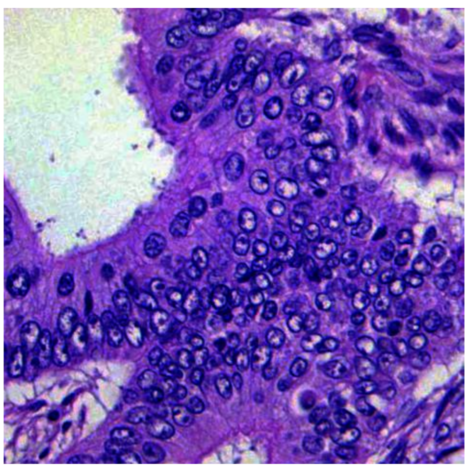


Figure 8 Processing Image of Colon aca

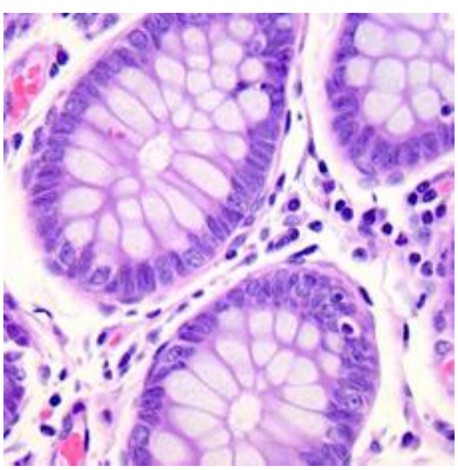


Figure 10 Colon n data

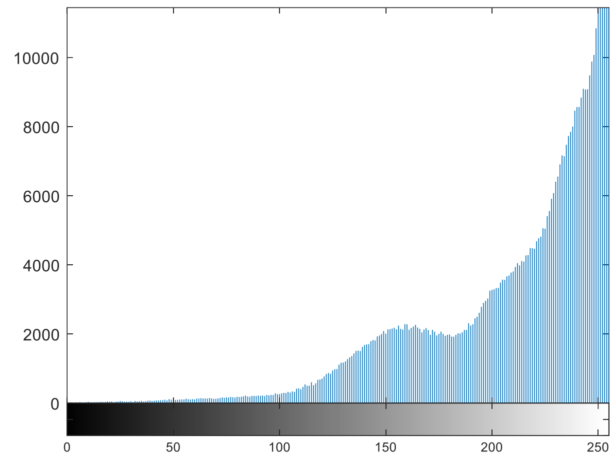


Figure 12 Histogram Image of Colon n

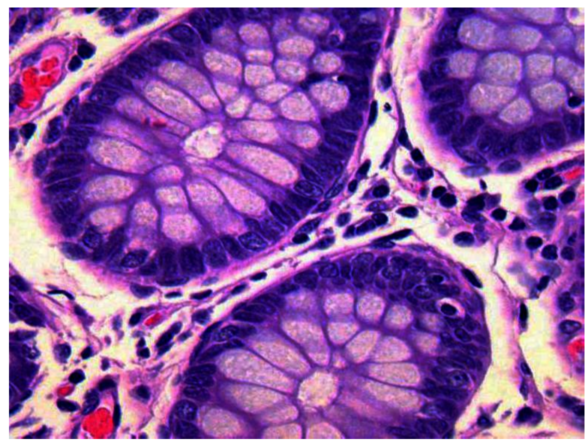


Figure 13 Processing Image of Colon n

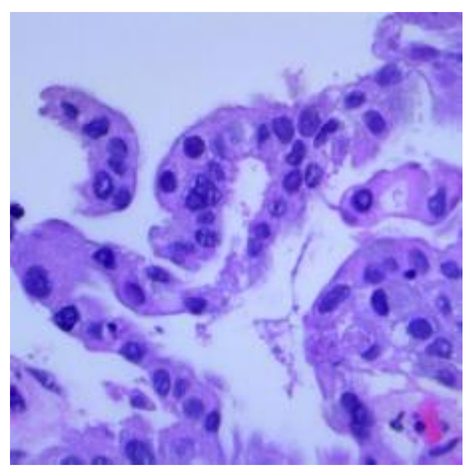


Figure15 Lung aca data

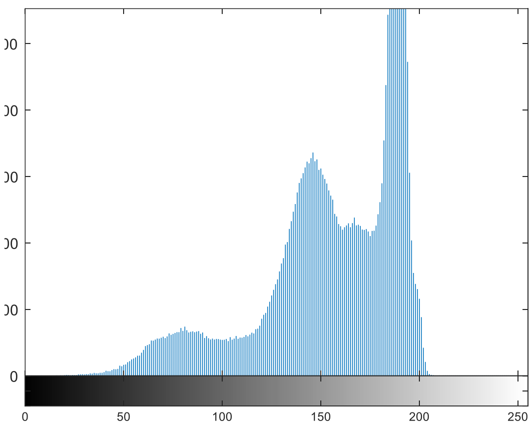


Figure 17 Histogram Image of Lung aca

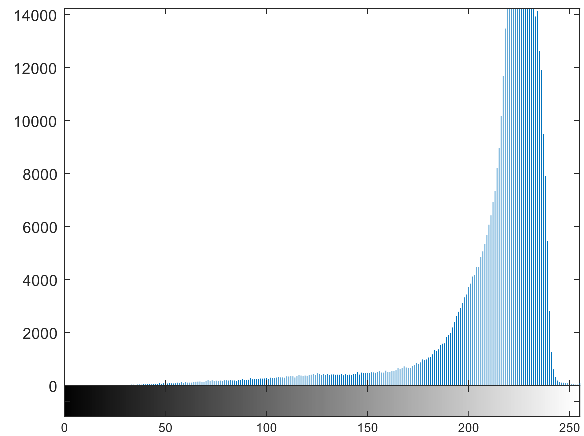


Figure 18 Histogram Image of Lung n

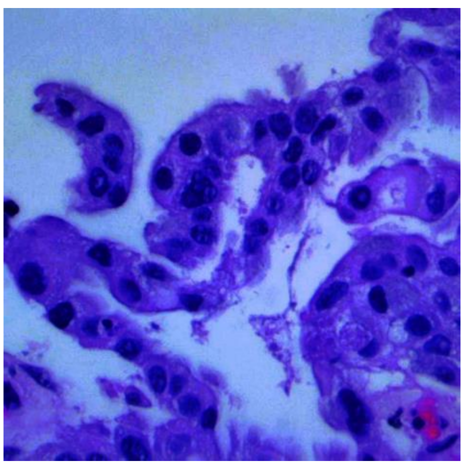


Figure 14 Processing Image of Lung aca

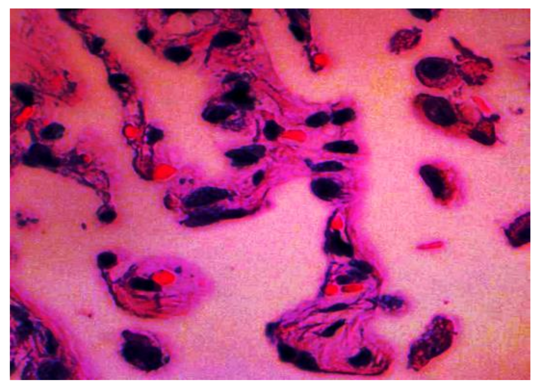


Figure 19 Processing Image of Lung n

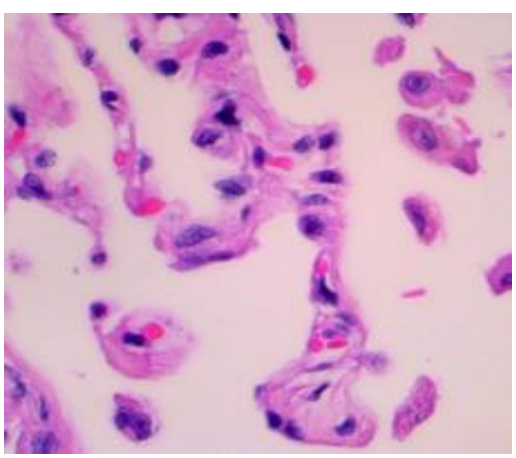


Figure 16 Lung ndata

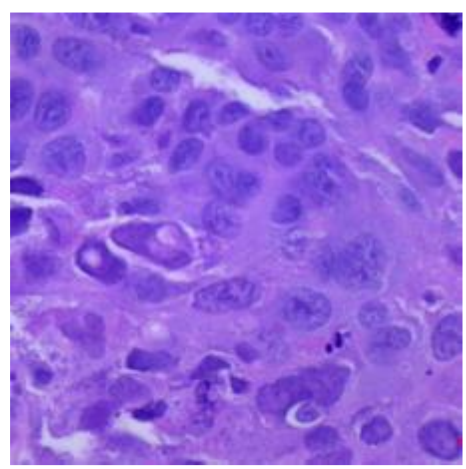


Figure 21 Lung sca Data

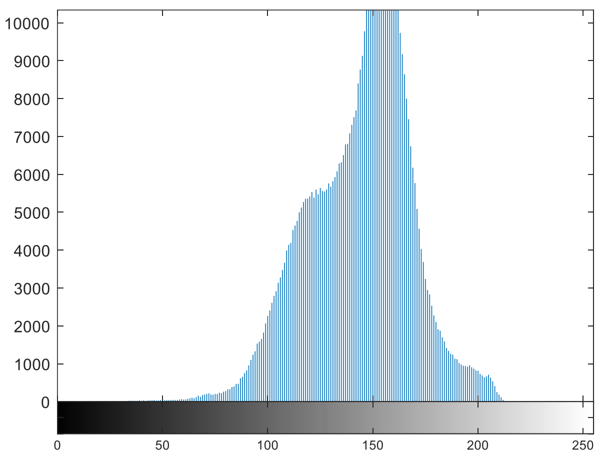


Figure 23 Histogram Image of Lung sca

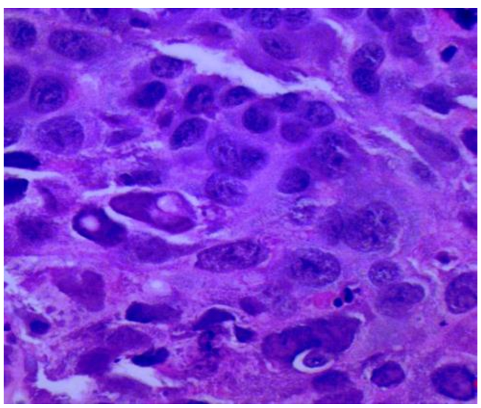


Figure 20 Processing Image of Lung sca

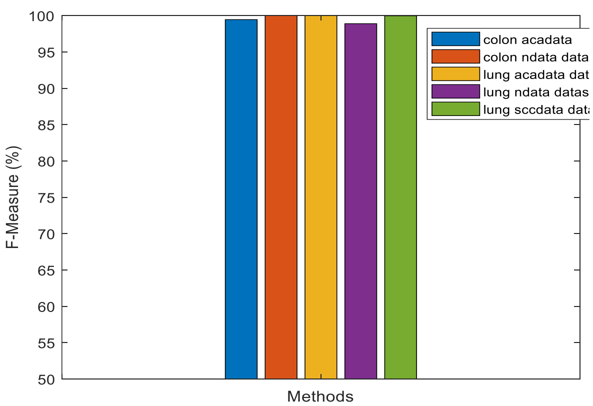


Figure 22 F-Measure Performance

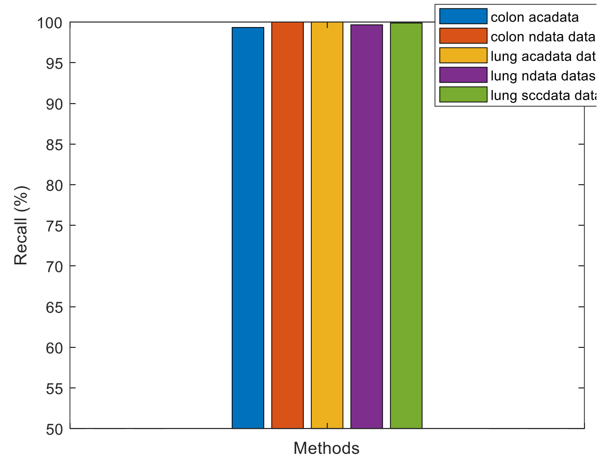


Figure 24 Recall Performance

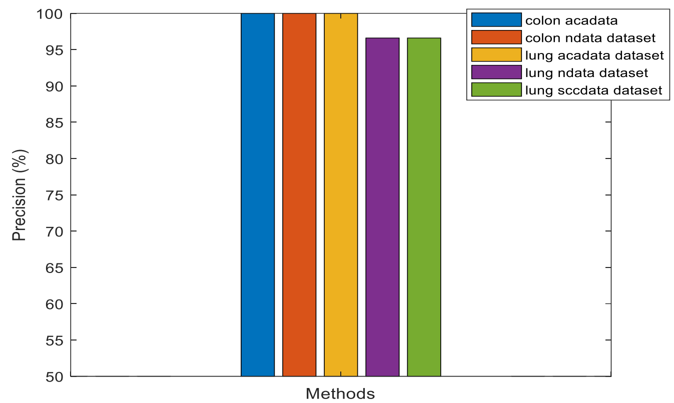


Figure 25 Precision Performance

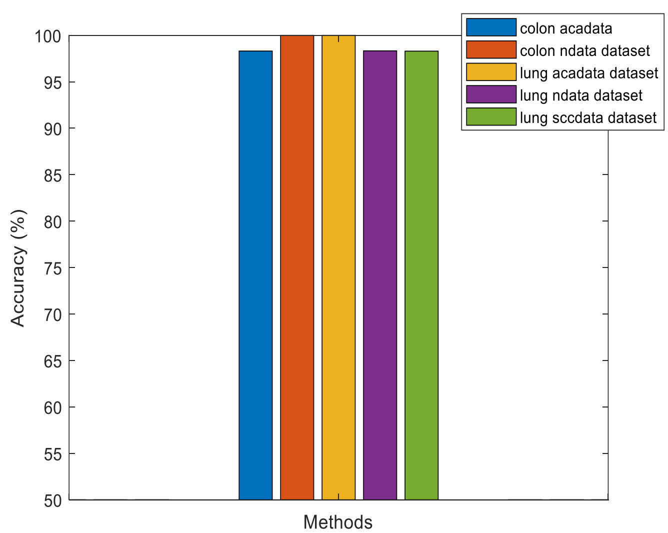


Figure 26 Accuracy Performance

5. CONCLUSION

This research endeavor aimed to enhance the detection of lung and colon cancer utilizing deep learning techniques. Initially, our model achieved an accuracy rate of 89%. To elevate its performance, we introduced the Bayesian Optimization Algorithm - Enhanced Histogram Equalization (BAO-EHE) method, resulting in an accuracy exceeding 99%. Our research outperformed existing methods, simultaneously reducing both time and computational costs. Key research metrics included a 99% accuracy, 99% precision, a 99.78% recall rate, and a 99.66% F1-score, unequivocally affirming the efficacy of our research in the realm of colon and lung cancer classification. Notably, our research excelled in colon cancer detection, offering valuable support to pathologists in the verification of their diagnoses. Our future research plans encompass the extensive testing of our model on diverse datasets and the exploration of hybrid optimization techniques. This research innovation holds the promise of advancing disease diagnosis, ultimately contributing to improved patient outcomes and enhanced survival rates.

REFERENCES

[1]. World Fact Sheet. International Agency for Research on Cancer. Accessed: Oct. 17, 2021. [Online]. Available: https://gco.iarc.fr/today/ data/factsheets/populations/900-world-fact-sheets.pdf.

[2]. Manafi-Farid, R., Askari, E., Shiri, I., Pirich, C., Asadi, M., Khateri, M., & Beheshti, M. (2022, November). [18F] FDG-PET/CT Radiomics and artificial intelligence in lung cancer: technical aspects and potential clinical applications. In Seminars in Nuclear Medicine (Vol. 52, No. 6, pp. 759-780). WB Saunders.

[3]. Cancer Symptoms and Causes. Mayo Clinic. Accessed: Oct. 18, 2021. [Online]. Available: https://www.mayoclinic.org/diseases-conditions/cancer/symptoms-causes/syc-20370588.

[4]. Thanoon, M. A., Zulkifley, M. A., Mohd Zainuri, M. A. A., & Abdani, S. R. (2023). A Review of Deep Learning Techniques for Lung Cancer Screening and Diagnosis Based on CT Images. Diagnostics, 13(16), 2617.

[5]. World Health Organization. Accessed: Oct. 18, 2021. [Online]. Available: https://www.who.int/news-room/ fact-sheets/detail/cancer.

[6]. Mayo Clinic. Accessed:Oct. 18, 2021. [Online]. Available: https://www.mayoclinic.org/ diseases-conditions/cancer/symptoms-causes/syc-20370588.

[7]. Hindocha, S., Charlton, T.G., Linton-Reid, K. et al. Gross tumour volume radiomics for prognostication of recurrence & death following radical radiotherapy for NSCLC. npj Precis. Onc. 6, 77 (2022). https://doi.org/10.1038/s41698-022-00322-3.

[8]. Abdollahi, H., Chin, E., Clark, H., Hyde, D. E., Thomas, S., Wu, J., ... & Rahmim, A. (2022). Radiomics-guided radiation therapy: opportunities and challenges. Physics in Medicine & Biology, 67(12), 12TR02.

[9]. J. Fan, J. Lee, and Y. Lee, ``A transfer learning architecture based on a support vector machine for histopathology image classi_cation," Appl. Sci., vol. 11, pp. 1_16, 2021.

[10]. R. Selvanambi, J. Natarajan, M. Karuppiah, S. K. H. Islam, M. M. Hassan, and G. Fortino, ``Lung cancer prediction using higher-order recurrent neural network based on glowworm swarm optimization," Neural Comput. Appl., vol. 32, no. 9, pp. 4373_4386, 2018.

[11]. O.de Carvalho Filho, A. C. Silva, A. C. de Paiva, R. A. Nunes, and M. Gattass, ``Classi_cation of patterns of benignity and malignancy based on CT using topology-based phylogenetic diversity index and convolutional neural network," Pattern Recognit., vol. 81, pp. 200_212, Sep. 2018.

[12]. R. V. M. da Nóbrega, P. P. Rebouças Filho, M. B. Rodrigues, S. P. P. da Silva, C. M. J. M. D. Júnior, and V. H. C. de Albuquerque,``Lung nodule malignancy classi_cation in chest computed tomography images using transfer learning and convolutional neural networks," Neural Comput. Appl., vol. 32, no. 15, pp. 11065_11082, Aug. 2020.

[13]. Masood, B. Sheng, P. Li, X. Hou, X. Wei, J. Qin, and D. Feng, ``Computer-assisted decision support system in pulmonary cancer detection and stage classi_cation on CT images," J. Biomed. Inf., vol. 79, pp. 117_128, Mar. 2018.

[14]. V. Nardone, P. Tini, P. Pastina, C. Botta, and A. Reginelli, ``Radiomics predicts survival of patients with advanced non-small cell lung cancer undergoing PD-1 blockade using Nivolumab," Oncol. Lett., vol. 9, no. 2,pp. 1559_1566, 2020.

[15]. Z. Zhou, S. Li, G. Qin, M. Folkert, S. Jiang, and J.Wang, ``Multi-objective- based radiomic feature selection for lesion malignancy classi_cation," IEEE J. Biomed. Health Informat., vol. 24, no. 1, pp. 194_204, Jan. 2020.

[16]. Z. Liu et al., ``Radiomics of multiparametric MRI for pretreatment pre- diction of pathologic complete response to neoadjuvant chemotherapy in breast cancer: A multicenter study," Clin. Cancer Res., vol. 25, no. 12, pp. 3538_3547, Jun. 2019.

[17]. Y. Xie et al., ``Knowledge-based collaborative deep learning for benign- malignant lung nodule classi_cation on chest CT," IEEE Trans. Med. imag., vol. 38, no. 4, pp. 991_1004, Apr. 2019.

[18]. Y. Boers and P. K. Mandal, ``Optimal particle-_lter-based detector," IEEE Signal Process. Lett., vol. 26, no. 3, pp. 435_439, Mar. 2019.

[19]. S. Hawkins et al., ``Predicting malignant nodules from screening CT scans," J. Thoracic Oncol., vol. 11, no. 12, pp. 2120_2128, Dec. 2016.

[20]. L. Tiancheng, B. Miodrag, and D. M. Petar, ``Resampling methods for particle _ltering: Classi_cation, implementation, and strategies," Signal Process. Mag., vol. 32, no. 3, pp. 70_86, May 2015.

[21]. K. Liu and G. Kang, ``Multiview convolutional neural networks for lung nodule classi_cation," Int. J. Imag. Syst. Technol., vol. 27, no. 1, pp. 12_22, Mar. 2017.

[22]. L. Tiancheng, B. Miodrag, and D. M. Petar, ``Resampling methods for particle _ltering: Classi_cation, implementation, and strategies," Signal Process. Mag., vol. 32, no. 3, pp. 70_86, May 2015.

[23]. D. R. Aberle et al., ``The national lung screening trial: Overview and study design," Radiology, vol. 258, no. 1, pp. 243_253, Jan. 2011.

[24]. T. Khan, P. Ramuhalli, and S. C. Dass, ``Particle-_lter-based multisensory fusion for solving low-frequency electromagnetic NDE inverse problems," IEEE Trans. Instrum. Meas., vol. 60, no. 6, pp. 2142_2153, Jun. 2011.

[25]. S. D. Chen and A. R. Ramli, ``Preserving brightness in histogram equalization-based contrast enhancement techniques," Digit Signal Pro- cess, vol. 14, no. 5, pp. 413_428, 2004.