Celal Bayar University Journal of Science

An Efficient and Lightweight Convolutional Neural Network Model for Lung Cancer Diagnosis

Çimen Uğur[1]*, Mahir Kaya[2]

[1,2] Tokat Gaziosmanpaşa University, Department of Computer Engineering, Tokat, Türkiye.
* cimenugur73@gmail.com
* Orcid No: 0009-0003-9970-2810

Received: January 15, 2025
Accepted: December 9, 2025
DOI: 10.18466/cbayarfbe.1620394

Abstract

Lung cancer stands as a major driver of mortality among cancer patients on a global scale, poses a substantial public health challenge due to its late diagnosis and the ambiguity of its early symptoms. This study proposes a novel, lightweight, and optimized Convolutional Neural Network architecture for detecting lung cancer types from Computed Tomography (CT) images. Initially, CT images of various lung cancer patients were combined for binary classification against normal patients. Experiments using different architectures and systematic hyperparameter optimizations resulted in the best model achieving a 93% accuracy rate. In the next phase, multi-class classification tasks were performed on the original dataset, and the performance of optimization algorithms such as Adamax, Adam, RMSprop, SGD, Adadelta, Nadam, and Adagrad were compared to determine the best optimizer. In these experiments, Adam algorithm achieved an 87% accuracy. The same study was repeated on an augmented dataset using data augmentation methods, reaching an accuracy of up to 96% with the RMSprop optimization algorithm. Furthermore, fine-tuned models were used and their test accuracies were evaluated using transfer learning methods (VGG16, ResNet50, InceptionV3, DenseNet121) on the available dataset. The primary contribution of this study is demonstrating that the proposed lightweight, optimized model achieves higher performance and efficiency on task-specific datasets, outperforming both established transfer learning methods and existing literature in accuracy. The proposed model will help healthcare professionals make quick and reliable decisions in lung cancer diagnoses.

Keywords: Convolutional Neural Networks, Lung Cancer, Transfer Learning, Hyperparameter Optimizations, Computed Tomography

1. Introduction

Lung cancer ranks among the leading causes of mortality, particularly in developed countries [1]. The disease often manifests with symptoms such as a persistent cough, shortness of breath, or unexplained fatigue [2]. According to 2020 statistics from the International Agency for Research on Cancer (GLOBOCAN), approximately 1.8 million deaths worldwide were attributed to lung cancer [3]. Despite significant advancements in medical science, lung cancer remains the most significant cause of cancer-related deaths globally [4]. To address this crucial public health issue, early diagnosis strategies are of paramount importance to reduce preventable deaths through timely intervention [5]. Detecting cancer in its early stages is vital for saving lives. Traditionally, cancer diagnosis relies on visual inspection and manual techniques. However, the manual analysis of medical images is both time-intensive and prone to errors [6]. In this context, automated diagnostic approaches have the potential to increase the efficiency of healthcare systems by reducing the workload on specialists and expediting diagnostic processes [7]. Deep learning has emerged as a promising tool in this field, with many researchers utilizing deep learning techniques for lung cancer diagnosis. Among these approaches, Convolutional Neural Networks (CNNs) have become one of the most widely used methods [8]. Similar to feed forward neural networks, CNNs consist of fundamental layers such as convolutional layers, pooling layers, and fully connected layers [9]. Existing studies in the literature often suffer from two key shortcomings. First, the hyperparameter optimizations of specially designed models are not investigated in sufficient depth. Second, in small and limited-labeled datasets, transfer learning models with a large number of parameters often exhibit low performance due to the overfitting problem. CNN

architectures demand thorough adjustment of hyperparameters to attain optimal performance. Therefore, this study presents a lightweight, enhanced with attention, and optimized custom CNN architecture with a minimum number of layers designed to accurately detect lung cancers from CT images. We conducted extensive experiments to evaluate different hyperparameters and various optimization algorithms to determine the configuration that provides the best performance. This study makes a noteworthy contribution to the literature by demonstrating that a carefully optimized light custom model can provide higher performance and efficiency compared to complex, pre-trained transfer learning models for specific medical imaging tasks. Furthermore, the proposed model can support healthcare specialists in making rapid and reliable diagnostic decisions for lung diseases.

2. Related Work

The literature includes many notable studies applying deep learning to lung cancer detection. Agnes et al. [10] studied lung tumor classification in their study. The authors conducted classification using a Convolutional Long Short-Term Memory (ConvLSTM) neural network on computed tomography images and achieved an accuracy rate of 0.93. This result is better than CNN and long short-term memory (LSTM) models alone. Rehman et al. [11] conducted a study on detecting lung cancer using machine learning methods applied to computed tomography images. The authors detected not only the cancer but also the type of cancer, and these are adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. The best results were obtained with Support Vector Machine (SVM) and K-Nearest Neighbor (k-NN) methods and the accuracy rates were 0.93 for SVM and 0.91 for k-NN. Lei et al. [12] classified lung tumors as good or bad and used Meta Ordered Weighted Network model using Meta Ordered Set. The authors obtained an accuracy rate of 0.69 with the Mow-Net model. Alakwaa et al. [13] proposed a CNN for classifying lung CT images in Data Science Bowl and Kaggle. They used the U-net CNN architecture and worked on the LUNA dataset and as a result, an accuracy rate of 0.87 was obtained with the method using the U-net CNN architecture. Liao et al. [14] investigated the use of 3D CNNs for the detection and classification of lung nodules from CT scans. The study focused more on the use of spatial information available in 3D scans, which was hypothesized to improve the accuracy of detecting nodules of various sizes, and the model achieved an accuracy of 0.86 as a performance measure. Anthimopoulos et al. [15] created a CNN model for classifying various tissue types in lung diseases, and their model includes 5 convolutional layers, 1 pooling layer, and 3 fully connected layers. In their study, the proposed CNN algorithm achieved better results in texture classification and detection with an accuracy of 0.86 compared to other algorithms such as LeNet, AlexNet and VGG Net. Song et al. [16] used the LIDC-IDRI dataset to classify CT images as benign and malignant in their study. The dataset was analyzed using architectures such as CNN, Deep Neural Networks (DNN), and Stacked Autoencoders (SAE). The CNN architecture outperformed the other proposed networks and showed the highest performance with 0.84 accuracy, 0.83 sensitivity and 0.84 specificity. Ashhar et al. [17] aimed to study the performance of five different Convolutional Neural Network architectures in the classification of lung tumors into malignant and benign using the LIDC-IDRI dataset, the architectures evaluated were DenseNet, MobileNetV2, ShuffleNet, SqueezeNet and GoogleNet. The performance of these architectures was measured based on accuracy, sensitivity, specificity and (AUC) performance values and showed that GoogleNet was the most successful CNN architecture with 0.95 accuracy, 0.99 specificity, 0.65 sensitivity and 0.86 AUC for lung tumor classification in CT images. Mohamed et al. [18] introduced a method that integrates a CNN model with a hybrid metaheuristic algorithm to enhance the accuracy of lung cancer diagnosis. In the proposed hybrid model, the Ebola Optimization Search Algorithm (EOSA) is applied to enhance the solution vector of the CNN architecture. First, a CNN architecture is designed and the solution vector is calculated. This solution vector is then passed to EOSA to determine the optimal combination of weights and biases to solve the classification problem. In experiments using a lung cancer dataset from the Iraqi Oncology Teaching Hospital/National Center for Cancer Diseases (IQ-OTH/NCCD), the EOSA-CNN model successfully proved its classification performance with an accuracy of 0.87. Tan et al. [19] employed a tailored VGG16-based 15-layer 2D CNN architecture utilizing transfer learning. This CNN model was trained and tested on CT image sets obtained from the National Lung Screening Trial and the National Institute of Allergy and Infectious Diseases TB Portals. The results show that the CNN with pre-trained lock-free weights achieves an F1 score of 0.90 and an accuracy of 0.90.

While these studies prove the effectiveness of deep learning models for lung cancer detection, the vast majority either do not detail important methods such as data augmentation and optimization, or report only limited metrics such as accuracy and F1 score. This leads to the neglect of more clinically meaningful metrics such as Precision, Recall, and Specificity. Furthermore, many of the transfer learning models used in existing studies can lead to overfitting problems in small and imbalanced datasets. To fill these gaps, our study proposes a lightweight, optimized custom model and presents a much more comprehensive performance evaluation with a wider range of metrics.

Table 1. Existing studies for lung cancer classification using CT images

Studies	Model	Year	Accuracy(%)	Classification
Agnes et al.	ConvLTSM	2020	93	Binary
Rehman et al.	SVM	2021	93	Multi-class
Lei et al.	MowNet	2021	69	Binary
Song et al.	CNN	2017	84	Binary
Ashhar et al.	GoogleNet	2021	95	Binary
Mohamed et al.	EOSA-CNN	2023	87	Multi-class
Tan et al.	DNN	2022	90	Binary
Alakwaa	U-net CNN	2017	87	Binary
Liao et al.	3D CNN	2019	86	Binary
Anthimopoulos et al.	CNN	2016	86	Multi-class

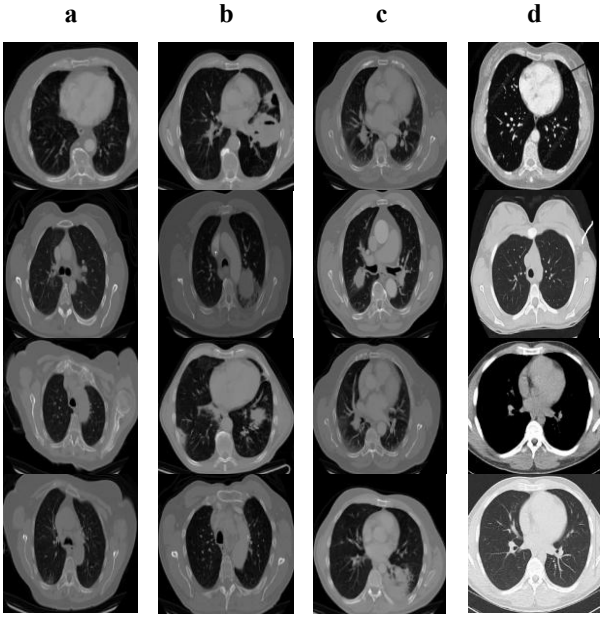
3. Materials and Methods

CNN models successfully solve a wide range of image-based classification problems, from medical image classification [20], image noise type detection [21] to waste classification [22]. They stand out from classical machine learning models, particularly due to their end-to-end and raw image as an input learning capabilities. CNNs extract feature maps using filters in each layer and reduce the image size through max pooling across blocks, extracting local features in the initial stages and object shapes in the final layers. To achieve successful results in CNN architecture based on image type and quantity, it is necessary to optimally adjust the model depth, i.e., the number of layers. In addition, to adjust the layer width, it is important to optimally determine the number of filters in each layer. If the models are very deep and wide, they can easily fall into overfitting when the dataset is small. Overfitting can cause the model to achieve high accuracy on the training dataset but perform poorly on the test dataset it has not seen before [23]. This study uses a dataset of lung CT images obtained from the Kaggle database [24]. The dataset includes subfolders for different types of lung cancer (squamous cell carcinoma, large cell carcinoma, adenocarcinoma) and a separate subfolder for normal CT scan images. The original Kaggle dataset was utilized as the primary dataset, whereas a more extensive secondary dataset was constructed by applying various data augmentation techniques to the original data. The datasets were split into 80% for training and 20% for testing. Data augmentation was applied to balance the classes and enhance the model's generalization ability. This process applied a range of transformations, including rescaling, rotation, horizontal and vertical shifts, shearing, zooming, and horizontal flipping.

Table 2. Comparison of the original and expanded datasets

	(Original)		(Expanded)	
Class	Training	Test	Training	Test
Normal	172	43	1000	250
Adenocarcinoma	271	67	1000	250
Large Cell Carcinoma	150	37	1000	250
Squamous Cell Carcinoma	208	52	1000	250

Table 3. Image samples in the dataset, **a)**Adenocarcinoma **b)** Large cell carcinoma **c)** Squamous cell carcinoma **d)** Normal lung CT images



These parameters made the model more robust against images in different positions and lighting conditions. The distributions of the data sets used are given in Table 2 and image examples in Table 3.
The methodological flow of the study is as follows:

i. **Binary Classification:** In the first stage, CT images of different lung cancer types were combined into a single "cancer" class and binary classification was performed against normal images. In this stage, different architecture and systematic hyperparameter optimization were used to determine the best model.
ii. **Multi-class Classification:** The dataset was returned to its original four-class structure, and the performance of different optimization algorithms (Adamax, Adam, RMSprop, Adadelta, SGD, Adagrad, Nadam) was compared for multi-class classification using the original and augmented datasets.
iii. **Transfer Learning:** Comparative experiments were performed using pre-trained transfer learning models such as VGG16, ResNet50, InceptionV3, and DenseNet121 on both the original and data-augmented datasets.

3.1. VGG16

VGG-16, proposed by the Visual Geometry Group (VGG) from the University of Oxford, features a 16-layer CNN architecture that includes 13 convolutional layers and 3 fully connected layers[25]. Each convolutional layer contains a rectified linear unit (ReLU) activation function to increase the model's capacity to learn complex patterns in images, while a “Softmax” activation function is applied in the output layer. Advanced models for transfer learning, including VGG16, have a large number of parameters that contribute to their adaptability and accuracy in complex image-based tasks.

3.2. Resnet50

The Residual Network (ResNet), developed by He et al. in 2015, is a widely used deep learning architecture [26]. A residual learning strategy was presented to ease the training of deeper networks, where each 2-layer block in the 34-layer configuration is replaced by a 3-layer bottleneck block, forming a 50-layer ResNet. Experimental results demonstrate that the 50-layer ResNet models outperform the 34-layer models in terms of accuracy. Moreover, no degradation problem was observed despite the increase in depth, which shows that increasing the network depth leads to significant accuracy gains in all evaluation criteria.

3.3. InceptionV3

Inception-v3 is a module proposed by Szegedy et al. [27], and it differs from inception-v2. The differences include the addition of not only convolution layers but also batch normalization and fully connected (FC) layers as auxiliary classifiers. As a result of this refinement, the framework has been named Inception- v3. Compared to the network defined by Ioffe and Szegedy [28] for single product evaluation in the ILSVRC 2012 classification and the top-quality version of Inception-v3, the presented Inception-v3 model is six times cheaper computationally and uses at least five times fewer parameters.

3.4. DenseNet121

The basic structure of the DenseNet121 model described by Huang et al. [29] consists of dense blocks each of which contains multiple convolutional layers, and these dense blocks provide deeper and more comprehensive feature learning and improve the efficiency of model parameters by establishing dense connections between layers and propagating feature maps directly to the next layers [29]. DenseNet121 includes a transition layer after each dense block that extends its depth up to 121 layers and a fully connected layer of 1000 units for ImageNet classification. In experiments, it tends to provide a consistent improvement in accuracy with an increasing number of parameters without any sign of performance degradation or overfitting, allowing the model to efficiently process large and detailed datasets.

3.5. Performance Measurement Parameters

The model's performance is evaluated on the test set using various metrics such as Recall, Accuracy, Precision, Confusion Matrix, and F1 Score. These metrics provide a comprehensive evaluation of the model's ability to accurately classify lung cancer images, and these values, true negative (TN), true positive (TP), false negative (FN), and false positive (FP), give us the number of predictions of the class labels.

3.5.1. F1 Score

The operation to be performed to find the F1 score value is the harmonic average of the Precision and Recall values [30]. The F1 score is employed to assess the effectiveness of diagnostic models in lung cancer detection, emphasizing the importance of accurately identifying positive cases (i.e., actual lung cancer patients). The formula of the performance parameter is given in Equation 3.1.

$$\text{F1 Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \qquad (\mathbf{3.1})$$

3.5.2. Recall

Recall evaluates how many of the truly positive instances are correctly classified, emphasizing its role in capturing positive predictions accurately [31]. The formula of the performance parameter is given in Equation 3.2.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad (\mathbf{3.2})$$

3.5.3. Accuracy

In the study, the classifier's accuracy was calculated by taking the ratio of the total correctly classified positives and negatives to the overall number of samples [32]. The performance parameter's formula is presented in Equation 3.3.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \quad (3.3)$$

3.5.4. Precision

The classifier's performance improved on the balanced dataset but declined on the unbalanced dataset [31]. Precision more clearly reveals performance differences than other metrics [31]. In the study, precision was used as an important performance metric. The formula of the performance parameter is given in Equation 3.4.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3.4)$$

3.5.5. Confusion Matrix

A confusion matrix serves as an effective and detailed tool for evaluating classifier performance, particularly in multi-class, single-label scenarios in which each instance belongs to a single class [33].

4. Results and Discussion

In the study, CT images containing normal lung tissue and different types of lung tumors were used. The experiments were carried out in the Google Colaboratory environment. In the initial phase of the study, extensive experiments were systematically performed on different CNN architectures and hyperparameters to determine the model that provides the highest accuracy for binary classification. In this process, the hyperparameters were tuned using a trial-and-error approach, and various optimization algorithms were evaluated. The experiments performed are given in Table 5.

4.1. Binary Classification and Model Optimization

To identify the most suitable architecture, different configurations were evaluated by varying hyperparameters such as the number of convolutional layers, epochs, batch size, and learning rate (see Table 4). The datasets were loaded using a data generator, with all images resized to a fixed input dimension and labeled for binary classification. For data preprocessing, all datasets (training, validation, and test) were normalized to the 0–1 pixel range using the parameter rescale is 1./255. To improve the model's generalization capability, data augmentation was applied to training set. Specifically, images were rotated up to ±5°, shifted horizontally and vertically up to 30%, subjected to shear transformations, and zoomed up to 30%. Horizontal flipping was also employed, and empty regions were filled using the nearest neighbor method. These transformations enhanced the model's robustness to variations in orientation, scale, and position. CNN architectures with an increased number of filters in successive layers were preferred, as they facilitate the extraction of more complex features [34].

Table 4. Hyperparameters and their abbreviations

Hyperparameters	Search Range Values
Number of Convolutional Layers (NCL)	5-15
Number of Filters (NF)	16-512
Optimization Algorithm (OA)	Adam, RMSprop
Learning Rate (LR)	0.00001, 0.0001, 0.001
Batch Size (BS)	16, 32
Input Image Size (IIS)	180, 224
Epoch (E)	10-100

Table 5. Experimental studies

Model No	NCL	NF	IIS	E	BS	OA	Acc (%)	F1 (%)	LR
1	7	32,32,64,64,128,128,512	224	10	16	RMSprop	87	85	0,0001
2	9	16,16,32,32,64,64,128,128,512	224	15	16	Adam	85	86	0,001
3	8	32,32,64,64,128,128,256,256	224	100	16	RMSprop	93	90	0,0001
4	7	32,32,64,64,128,128,512	224	15	32	Adam	80	71	0,001
5	7	32,32,64,64,128,128,512	224	10	16	Adam	80	82	0,001
6	9	32,32,64,64,128,128,256,256,512	224	30	16	Adam	67	65	0,00001
7	6	32,32,64,64,128,256	224	40	16	RMSprop	88	78	0,001
8	11	32,32,32,64,64,64,64,128,128,256,512	224	70	32	RMSprop	72	72	0,001
9	7	32,32,64,64,128,128,512	224	30	32	Adam	90	91	0,0001
10	8	32,32,64,64,128,128,256,512	224	35	32	RMSprop	65	64	0,001
11	11	32,32,32,64,64,64,64,128,128,256,512	224	25	32	RMSprop	80	71	0,0001
12	9	16,16,32,32,64,64,128,128,512	224	40	32	Adam	80	72	0,001
13	10	16,16,32,32,64,64,128,128,128,512	180	20	32	Adam	86	84	0,0001
14	9	16,16,32,32,64,64,128,128,512	180	40	32	Adam	70	82	0,001

15	7	32,32,64,64,128,128,512	224	10	32	Adam	53	48	0,001
16	8	16,16,32,32,64,64,128,512	180	50	32	Adam	80	71	0,001
17	7	32,32,64,64,128,128,512	180	20	16	RMSprop	64	66	0,001
18	10	16,16,32,32,64,64,128,128,128,512	224	25	32	RMSprop	77	78	0,0001
19	6	32,32,64,64,128,256	180	40	16	Adam	65	67	0,001
20	7	32,32,64,64,128,128,512	180	50	16	RMSprop	73	75	0,0001

Acc: Accuracy, F1: F1 Score

As a result of the experiments, the best model accomplished 93% accuracy and a 90% F1 score. This model was trained for 100 epochs with a learning rate of 0.0001 and a batch size of 16. The model's performance was observed using validation data, and the best-performing model was saved.

i. **Convolution Layers:** Each convolution layer uses a 3x3 filter with ReLU activation to stabilize the learning process, followed by Batch Normalization.
ii. **MaxPooling Layers:** MaxPooling layers with 2x2 filters were used after each convolutional block to decrease the spatial dimensions of the features.
iii. **Normalization Layer:** Normalization layers are used to normalize neuron activations within a layer to ensure they exhibit similar scales and distributions [35].
iv. **Dropout Layer:** Dropout layers in neural networks randomly deactivate some neurons at each training iteration, encouraging the network to learn redundant representations and reducing dependency on specific neurons [36].
v. **Flatten Layer:** The generated layers are tasked with extracting image features, producing a matrix-vector of low-level complex characteristics to represent the image uniquely [37]. To classify the output from the convolutional and pooling layers, a classification network is required. However, prior to classification, the output of these layers need to be merged into a single vector.
vi. **Dense Layer:** After flattening the layer of the feature maps, a fully connected layer with 512 units and ReLU activation was added.
vii. **Output Layer:** The output layer was run using sigmoid activation corresponding to two classes. The other dataset was run using softmax activation corresponding to four classes.

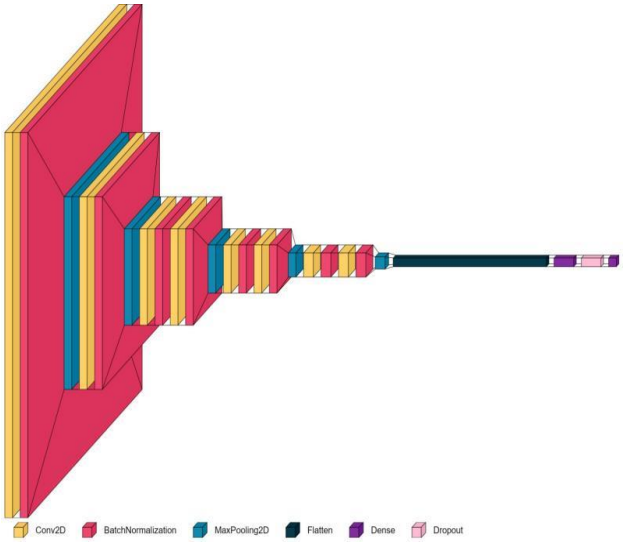


Figure 1. Proposed CNN architecture

The model's architecture is designed to capture hierarchical features from the input images. The architecture includes convolutional layers with ReLU activation followed by Batch Normalization layers, MaxPooling layers with 2×2 filters, Dropout layers, and a flattening layer. Finally, a fully connected layer with 512 units and the output layer (with sigmoid activation and softmax for multi-class) were added. The proposed model consists of eight convolutional layers, which can be organized into five blocks. The first two blocks each contain a convolutional layer, a batch normalization layer, and a max pooling layer, sequentially. The remaining three blocks each comprise two sequential convolutional layers with batch normalization in between, and each block also includes a max pooling layer. One of the defining characteristics of the model proposed in this study is its lightweight design. This lightness is directly related to our model having fewer parameters than other popular transfer learning models. The total number of parameters of our proposed model is 7,601,188. This number is quite low compared to complex models often used in the literature, which typically have many more parameters.

Celal Bayar University Journal of Science
Volume 22, Issue 1, 2026, p 95-105
Doi: 10.18466/cbayarfbe.1620394

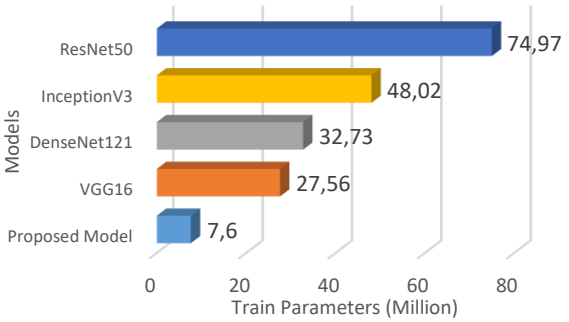


Figure 2. Comparison of training parameters between the proposed model and transfer learning models

In Figure 2, we compare the number of training parameters in our proposed model with state-of-the-art transfer learning models. Our proposed model demonstrates significantly better performance with fewer parameters.

4.2. Multi-class Classification and Algorithm Comparison

After determining the most suitable CNN architecture for binary classification, the effectiveness of different optimization algorithms was evaluated on a wider range of data. In the multi-class classification experiments performed on the original dataset, the performance of the Adamax, Adam, RMSprop, Adadelta, SGD, Adagrad, and Nadam algorithms was compared. In this phase, the model's performance was analyzed more comprehensively using metrics such as Precision, Recall, Specificity, and F1 score. The results obtained are shown in Table 6.

Table 6. Results obtained using different optimization algorithms with the original dataset obtained from Kaggle

Optimization Algorithm	Accuracy	F1 Score	Recall	Precision
Adam	87	87	87	88
SGD	63	61	61	61
RMSprop	73	71	72	69
Adadelta	78	76	76	77
Adagrad	60	63	63	63
Adamax	74	74	74	73
Nadam	62	58	61	57

In the study, the Adam algorithm, which gave the best results among the optimization algorithms, was run for 100 epochs and achieved an 87% accuracy and an 87% F1 score. The training process, accuracy, and losses were visualized graphically over the epochs in Figure 3. These graphs helped in understanding the model's learning behavior and identifying overfitting problems. We can say that throughout the epochs, training and validation data have mostly progressed in an overlapping manner. Figure 4 shows the confusion matrix obtained with the Adam optimization algorithm.

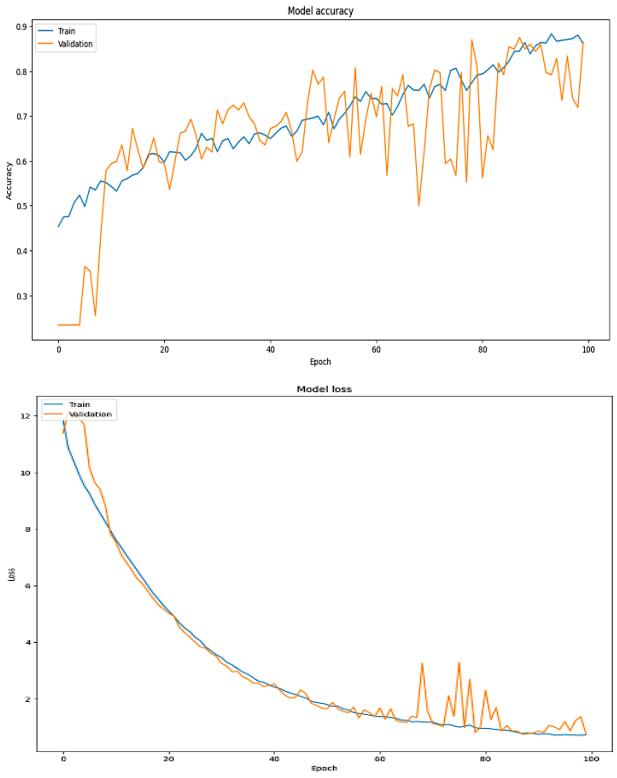


Figure 3. Accuracy and loss graph using Adam

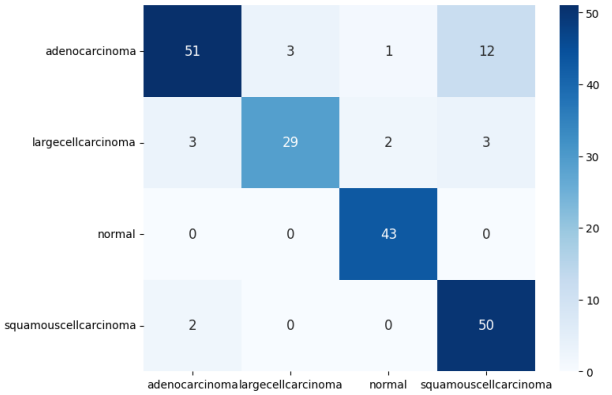


Figure 4. Confusion matrix using Adam

In Figure 3, fluctuations in the validation curve are generally observed in cases of dropout and small batch sizes. When lower dropout and larger batch sizes are used, these fluctuations decrease significantly, but when dropout is reduced, the model may overfit.
When examining Figure 4, we see that the model correctly classified 51 adenocarcinoma cases, but it also made the most errors here because it mistakenly classified 12 adenocarcinoma images as squamous cell carcinoma. It also incorrectly classified 3 images as large cell carcinoma and one as normal. Since our model made the most errors here, data augmentation methods that can

improve the distinction between these two classes can be applied.

In the experiments conducted on the data-augmented dataset, the RMSprop algorithm was run for 100 epochs (see Figure 5) and achieved 96% accuracy and a 95% F1 score. These findings demonstrate that data augmentation not only significantly increases performance but also enables the model to converge quickly. The results for the RMSprop algorithm are presented in Table 7, and the accuracy and loss curves in Figure 5, the confusion matrix in Figure 6, respectively.

Table 7. Results obtained using different optimization algorithms with the augmented dataset.

Optimization Algorithm	Accuracy	F1 Score	Recal	Precision
Adam	75	75	76	74
SGD	74	73	74	71
RMSprop	96	95	96	96
Adadelta	86	86	86	87
Adagrad	63	64	63	65
Adamax	90	90	90	91
Nadam	62	65	62	67

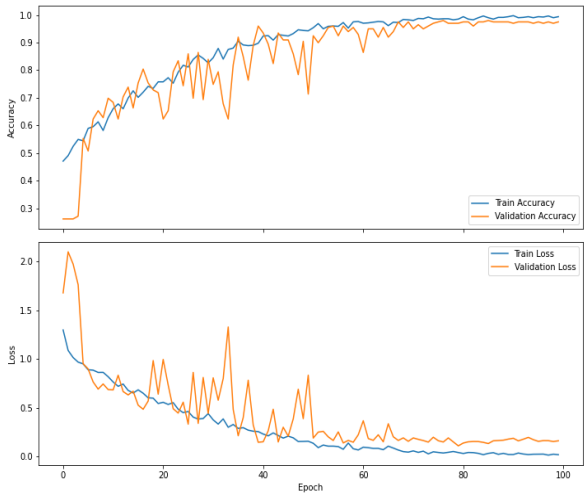


Figure 5. Accuracy and loss graph for 100 epochs using RMSprop.

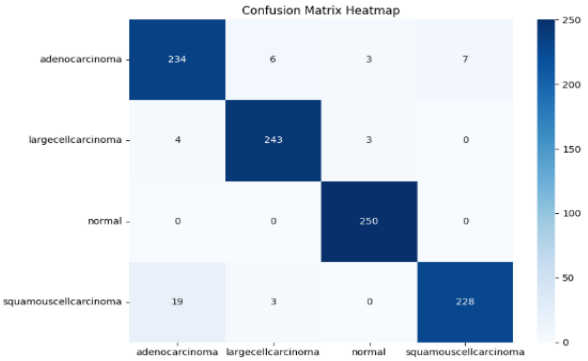


Figure 6. Confusion matrix using RMSprop

4.3. Transfer Learning and Comparison

Experiments were performed on both the original and data-augmented datasets to evaluate the performance of transfer learning models. The results obtained using VGG16, ResNet50, InceptionV3, and DenseNet121 on the original dataset from Kaggle are shown in Table 8. The highest accuracy was reached with ResNet50, providing 78% accuracy. The accuracy and loss curves for ResNet50 are shown in Figure 7 and Figure 8, respectively. The confusion matrix obtained from ResNet50 with transfer learning is presented in Figure 9.

Table 8. Results obtained using transfer learning with the original dataset.

Model	Accuracy	F1 Score	Recall
VGG16	67	65	67
Resnet50	78	78	78
InceptionV3	61	60	60
DenseNet121	76	74	74

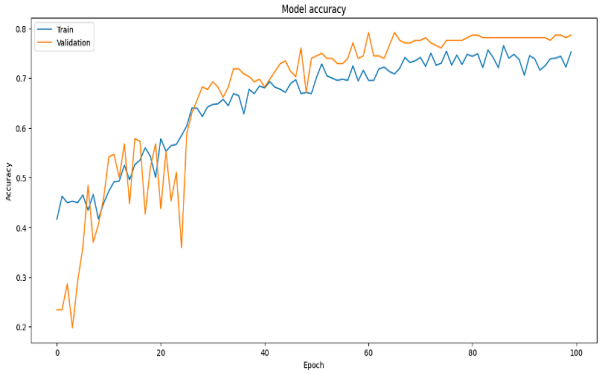


Figure 7. Accuracy graph implemented ResNet50.

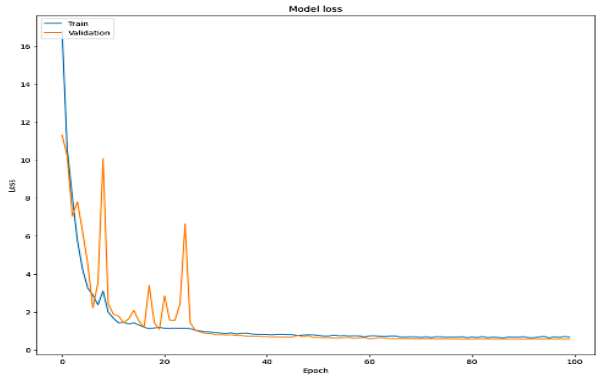


Figure 8. Loss graph implemented ResNet50.

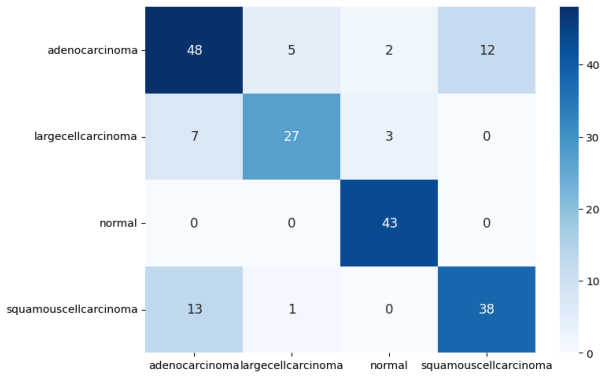


Figure 9. Confusion matrix implemented ResNet50.

The evaluations performed on the augmented dataset are shown in Table 9, revealing that the DenseNet121 architecture provided the best performance with 86% accuracy.

Table 9. Results obtained using transfer learning with the augmented dataset.

Model	Accuracy	F1 Score	Recall
VGG16	65	64	67
Resnet50	57	56	57
InceptionV3	60	61	61
DenseNet121	86	86	86

This comparison shows that the dataset created with data augmentation provides higher accuracy in transfer learning experiments as well. However, it was observed that the proposed custom CNN model (96% accuracy on the data-augmented dataset) performed much better than the transfer learning models (best at 86% accuracy). This proves that a properly optimized, lightweight architecture can be more suitable and effective for a specific task compared to large, pre-trained models.
Additionally, we used attention mechanisms. Specifically, after the last convolution block, we added both channel and spatial attention mechanisms to our model. Channel attention mechanisms enhance the impact of important channels, while spatial attention mechanisms enhance the impact of important regions on the feature map. By using these two attention mechanisms sequentially, we created a more stable model by first enhancing the impact of important channels and then enhancing the impact of important regions on the feature maps. With this approach, we have further improved the performance of the lightweight model.

5. Conclusion

Lung cancer remains one of the deadliest forms of cancer worldwide, and its early diagnosis is of critical importance for the timely and effective initiation of treatment. As traditional diagnostic methods can be prone to human error, the development of artificial intelligence-based systems that can directly analyze CT images is of great importance.
In this study, a lightweight, enhanced with attention and optimized CNN architecture was presented to classify lung cancers from CT images with high accuracy. The study conducted comparative experiments on the organized Kaggle dataset and augmented dataset via different optimization algorithms.
The experiments showed that the augmented dataset significantly improved the performance of both the proposed custom model and the transfer learning models. The proposed custom model, with the RMSprop algorithm on the augmented dataset, achieved 96% accuracy, a 95% F1 score, 96% Precision, and 96% Recall, surpassing the best-performing transfer learning model.
These results indicate that a lightweight and optimized custom architecture offers advantages such as high accuracy, efficiency, and low computational cost, suggesting it could be used as a potential decision support system for clinical applications. The low number of parameters in the model allows for less computational power and offers faster training and inference times. This is a great advantage, especially in systems with limited hardware resources or in real-time diagnostic applications. This lightness demonstrates that the model maintains its efficiency while achieving high accuracy rates.
For future work, it is proposed to extend the presented system into a Computer-Aided Diagnosis (CAD) interface that can not only detect cancerous tissues but also provide segmentation related features such as tumor size and volume.

Author's Contributions

Çimen Uğur: Drafted and wrote the manuscript, performed the experiment and result analysis.

Mahir Kaya: Assisted in analytical analysis on the structure, supervised the experiment's progress, result interpretation and helped in manuscript preparation.

Ethics

There are no ethical issues after the publication of this manuscript.

References

[1]. Retico A, Delogu P, Fantacci ME, Gori I, Preite Martinez A. *Lung nodule detection in low-dose and thin-slice computed tomography*. Comput Biol Med. 2008 Apr ;38(4):525–34.

[2]. *Akciğer Kanseri Belirtileri Nelerdir? Nedenleri ve Tedavisi | Anadolu Sağlık Merkezi Hastanesi [Internet]*. [cited 2024 Now 21]. Available from: https://www.anadolusaglik.org/saglik-rehberi/akciger-kanseri

[3]. *World Health Organization. Lung cancer [Internet*]. [cited 2024 Apr 21]. Available from: https://www.who.int/news-room/fact-sheets/detail/lung-cancer

[4]. Temurtaş F, Öztekin M, Yazdani M, Yörük YE, Aydemir F, Yonar D, et al. *Akciğer Kanseri Tanısı İçin Yeni Bir Yöntem*. Mühendislik Bilimleri ve Araştırma Dergisi. 2019;35–48.

[5]. Xie Y, Xia Y, Zhang J, Song Y, Feng D, Fulham M, et al. *Knowledge-based Collaborative Deep Learning for Benign-Malignant Lung Nodule Classification on Chest CT*. IEEE Trans Med Imaging. 2019 Apr 1;38(4):991–1004.

[6]. Munir K, Elahi H, Ayub A, Frezza F, Rizzi A. *Cancer Diagnosis Using Deep Learning: A Bibliographic Review*. 2019 Aug 23 ;11(9):1235.

[7]. Kaya M, Cetin-Kaya Y. *A novel ensemble learning framework based on a genetic algorithm for the classification of pneumonia*. Eng Appl Artif Intell. 2024 Jul 1; 133:108494

[8]. Krizhevsky A, Sutskever I, Hinton GE. *ImageNet classification with deep convolutional neural networks*. Commun ACM. 2017 May 24;60(6):84–90.

[9]. Çetin-Kaya Y, Kaya M. *A Novel Ensemble Framework for Multi-Classification of Brain Tumors Using Magnetic Resonance Imaging*. Diagnostics 2024;14(4):383.

[10]. Akila Agnes S, Alex Pandian Immanuel S, Anitha J, Arun Solomon A. *Classification of Lung nodules using Convolutional long short-term Neural Network*. Proceedings - 5th International Conference on Computing Methodologies and Communication, ICCMC 2021. 2021 Apr 8;1349–53.

[11]. Rehman A, Kashif M, Abunadi I, Ayesha N. *Lung Cancer Detection and Classification from Chest CT Scans Using Machine Learning Techniques*. In: 2021 1st International Conference on Artificial Intelligence and Data Analytics, CAIDA 2021. Institute of Electrical and Electronics Engineers Inc.; 2021. p.101–4.

[12]. Lei Y, Shan H, Zhang J. *Meta ordinal weighting net for improving lung nodule classification*. In: ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings. Institute of Electrical and Electronics Engineers Inc.; 2021. p. 1210–4.

[13]. Alakwaa W, Nassef M, Badr A. *Lung Cancer Detection and Classification with 3D Convolutional Neural Network (3D-CNN).* Int J Adv Comput Sci Appl. 2017;8(8):1- 6.

[14]. Liao F, Liang M, Li Z, Hu X, Song S. *Evaluate the Malignancy of Pulmonary Nodules Using the 3-D Deep Leaky Noisy-OR Network.* IEEE Trans Neural Netw Learn Syst. 2019 Nov 1;30(11):3484–95.

[15]. Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S. Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network. IEEE Trans Med Imaging. 2016 May 1;35(5):1207–16.

[16]. Song QZ, Zhao L, Luo XK, Dou XC. *Using Deep Learning for Classification of Lung Nodules on Computed Tomography Images*. J Healthc Eng. 2017 Jan 1;2017(1):8314740.

[17]. Ashhar SM, Mokri SS, Rahni AAA, Huddin AB, Zulkarnain N, Azmi NA, et al. *Comparison of deep learning convolutional neural network (CNN) architectures for CT lung cancer classification*. International Journal of Advanced Technology and Engineering Exploration. 2021;8(74):126–34.

[18]. Mohamed TIA, Oyelade ON, Ezugwu AE. *Automatic detection and classification of lung cancer CT scans based on deep learning and ebola optimization search algorithm*. PLoS One. 2023 Aug 17 ;18(8): e0285796.

[19]. Tan H, Bates JHT, Matthew Kinsey C. *Discriminating TB lung nodules from early lung cancers using deep learning*. BMC Med Inform Decis Mak. 2022 Dec 1 ;22(1):1–7.

[20]. Cheyi, J., & Kaya, Y. Ç. (2024). Advanced CNN-Based Classification and Segmentation for Enhanced Breast Cancer Ultrasound Imaging. Gazi University Journal of Science Part A: Engineering and Innovation, 11(4), 647-667.

[21]. Güneş, A., & Kaya, Y. Ç. (2024). Evrişimsel Sinir Ağları ile Görüntülerde Gürültü Türünü Saptama. Bilgisayar Bilimleri ve Mühendisliği Dergisi, 17(1), 75-89.

[22]. Saraçoğlu, Y. D., & Kaya, Y. Ç. (2025). Efficient Hyperparameter-Tuned Convolutional Neural Network for Waste Classification. Gazi University Journal of Science Part A: Engineering and Innovation, 12(3), 815-835.

[23]. Kaya M, Cetin-Kaya Y. *A Novel Deep Learning Architecture Optimization for Multiclass Classification of Alzheimer's Disease Level*. IEEE Access. 2024; 12:46562–81.

[24]. Chest CT-Scan images Dataset [Internet]. [cited 2023 Dec 11]. Available from: https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images

[25]. Simonyan K, Zisserman A. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings. 2014 Sep 4; Available from: https://arxiv.org/abs/1409.1556v6

[26]. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2015 Dec 10 ;2016-December:770–8. Available from: https://arxiv.org/abs/1512.03385v1

[27]. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. *Rethinking the Inception Architecture for Computer Vision.* Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2015 Dec 2;2016- December:2818–26. Available from: https://arxiv.org/abs/1512.00567v3

[28]. Ioffe S, Szegedy C. *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*. 32nd International Conference on Machine Learning, ICML 2015. 2015 Feb 11; 1:448–56. Available from: https://arxiv.org/abs/1502.03167v3

[29]. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. *Densely Connected Convolutional Networks. Proceedings* - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017. 2016 Aug 25;2017- January:2261–9. Available from: https://arxiv.org/abs/1608.06993v5

[30]. Goutte C, Gaussier E. *A Probabilistic Interpretation of Precision, Recall and F1- skore, with Implication for Evaluation*. Conference: Advances in Information Retrieval, 27th European Conference on IR Research. 2005 March 21-23.

[31]. Saito T, Rehmsmeier M. *The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets*. PLoS One. 2015 Mar 4;10(3): e0118432.

[32]. Memiş S, Enginoğlu S, Erkan U. A Data Classification Method in Machine Learning Based on Normalised Hamming Pseudo-Similarity of Fuzzy Parameterized Fuzzy Soft Matrices. Bilge International Journal of Science and Technology Research. 2019; 3:1–8.

[33]. Krstinić D, Braović M, Šerić L, Božić-Štulić D. Multi-label classifier performance evaluation with confusion matrix. 2020 Jan 14;1–14.

[34]. Kaya M. *Bayesian optimization based CNN framework for automatic detection of brain tumors.* Balkan Journal of Electrical and Computer Engineering. 2023;11(4):395-404.

[35]. Ba JL, Kiros JR, Hinton GE. *Layer Normalization*. 2016 Jul 21 Available from: https://arxiv.org/abs/1607.06450v1

[36]. Gal Y, Ghahramani Z. *Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning*. 33rd International Conference on Machine Learning, ICML 2016. 2015 Jun 6; 3:1651–60.

[37]. Jeczmionek E, Kowalski PA. *Flattening Layer Pruning in Convolutional Neural Networks*. Symmetry 2021. 2021 Jun 27;13(7):1147.