AI-Driven Histopathological Analysis for Early Lung Cancer Detection

Erika Severeyn[0000−0002−9500−3532]
Carlos Rangel
Alexandra La Cruz[0000−0001−6052−2933] and
Jesús Velásquez[0000−0002−0811−3320]

Abstract Lung cancer has experienced a significant increase in prevalence, leading to late diagnoses and putting pressure on health systems. This pathology affects the respiratory system through symptoms such as persistent cough, shortness of breath, chest pain, and other associated signs. The delay in diagnosis often results in advanced-stage detection, significantly reducing treatment options and survival rates. This research uses Artificial Intelligence algorithms to classify lung cancer through histopathological images, aiming to improve early detection. By leveraging advanced machine learning techniques, we seek to provide a more accurate and timely diagnosis. The public dataset 'Lung and Colon Cancer Histopathological Images' was utilized for this research. Four convolutional neural network models were developed to classify lung cancer images into the three categories presented in the data: benign lung cancer, lung adenocarcinoma, and squamous cell lung cancer. In addition, an interface was designed for better result visualization, interpretation, and automation, allowing quicker decision-making by medical professionals. This AI-driven decision support system aims to reduce the disease detection stage, thereby increasing survival chances and enhancing the user experience for medical specialists and contributing to a more effective and efficient healthcare system, ultimately improving patient outcomes.

Erika Severeyn
Department of Thermodynamics and Transfer Phenomena, Universidad Simón Bolívar, Venezuela, e-mail: severeynerika@usb.ve

Carlos Rangel
Computer Engineer and Biomedical Engineering, Universidad Internacional de Valencia, Valencia, España, e-mail: carlosrangelcascajosa@gmail.com

Alexandra La Cruz
Faculty of Engineering, Universidad de Ibagué, Ibagué, Tolima, Colombia, e-mail: alexandra.lacruz@unibague.edu.co

Jesús Velásquez
Department of Thermodynamics and Transfer Phenomena, Universidad Simón Bolívar, Venezuela, e-mail: jmvelasquezf@usb.ve

1 Introduction

Lung cancer persists as a formidable global health burden, with the World Health Organization (WHO) reporting a staggering 1.8 million fatalities attributed to the disease in 2020 alone. This alarming statistic underscores the pressing need for more efficacious strategies in early lung cancer detection, treatment, and monitoring [11].

The age of technological advancement has witnessed the emergence of artificial intelligence (AI) as a promising tool in medicine. Notably, advancements in convolutional neural networks (CNNs) have revolutionized medical image interpretation, including histopathological images crucial for lung cancer diagnosis [7]. Traditional methods relying on histopathological image analysis by pathologists can be time-consuming and prone to human error [1]. Several studies have explored the potential of deep learning, particularly CNNs, to address these limitations in lung cancer detection. In [9] developed a CNN model achieving 98.15% training accuracy and 98.07% validation accuracy in classifying lung cancer types (benign, adenocarcinoma and squamous cell carcinoma) from histopathological images. In [8]also employed a CNN model, achieving training and validation accuracies of 96.11% and 97.2% respectively, for similar lung cancer classification tasks on histopathological images. These studies demonstrate the promising capabilities of CNNs in achieving high accuracy for lung cancer detection from histopathological images.

It is important to note that other deep learning architectures beyond CNNs are also being explored for medical image analysis. For instance, in [2] investigated the use of Mask R-CNN for identifying pathological signs of respiratory diseases on chest CT scans. While achieving high sensitivity for ground-glass opacity detection, their model yielded lower performance for pulmonary nodule detection. These findings suggest that deep learning offers a powerful set of tools for computer-aided diagnosis of lung cancer and other respiratory diseases. However, further research is needed to refine and optimize these models for broader applicability and improved performance across various disease classifications.

This study utilizes a large dataset of lung cancer histopathological images obtained from [4] to develop an AI-driven support system for early and accurate detection of lung cancer. Through the application of AI algorithms, specifically CNNs, the images will be analyzed and classified. This novel approach combines advanced image processing techniques and powerful CNN models to significantly improve lung cancer detection. The goal of this research is to create a support system that can efficiently analyze histopathological images and classify them accurately, leading to a more precise and timely detection of lung cancer. This system has the potential to greatly improve the current standard of care for lung cancer patients by providing a reliable and accurate tool for early detection. By integrating such AI-driven decision support systems into clinical practice, we aim to enhance decision-making processes, optimize resource allocation, and ultimately improve patient outcomes in the healthcare industry.

The paper is structured as follows: Section 2 outlines the methodology, including the dataset description and its preparation, then the dataset spliting process for

training, validation and testing, the description of the different models implementd and metrics used for assesing the models. The findings of our study are presented in Section 3. Finally, section 5 provides an in-depth discussion of the results, their significance, and relevance to existing knowledge, and also proposes directions for future research.

2 Methodology

2.1 Database

A neural network model for classification tasks requires a database with a large database with labeled cases for each existing class [10]. After an evaluation of various database sources, the ”Lung and Colon Cancer Histopathological Images” database was selected [4]. This database contains 25,000 images, which were generated by augmenting 1,250 histology images. Out of the 25,000 images, 15,000 correspond to different types of lung cancer, with 5,000 images labeled as benign lung cancer, 5,000 as lung adenocarcinoma, and 5,000 as squamous cell lung cancer. A perfectly balanced dataset. The images in the dataset are high-resolution, with a size of 768x768 pixels in RGB mode, and are saved in JPEG format. Examples of images from the database for each cancer type used in the study are shown in Figure 1.

2.2 Data Preprocessing

To address the hardware limitations (memory and processing power) of the machines employed and the high resolution of the histopathological images, image preprocessing was indispensable for efficient model training. This involved applying compression filtering to reduce the image size while preserving essential information [12]. A systematic reduction of the image resolution was conducted, and the subsequent assessment of its effect on the model’s classification accuracy facilitated the identification of an optimal compression level. This optimal level substantially accelerated training speed and reduced memory requirements without compromising image integrity or model performance. Before using the images as input for a CNN, they were converted into NumPy arrays within the programming environment. Each original image was transformed into a 3D array with dimensions of 768x768x3, where the dimensions represent the following:

- The first two dimensions (768x768) denote the image’s spatial resolution, comprising the height and width (number of rows and columns).
- The third dimension (3) corresponds to the three color channels (Red, Green, and Blue) associated with each pixel, capturing the color information.

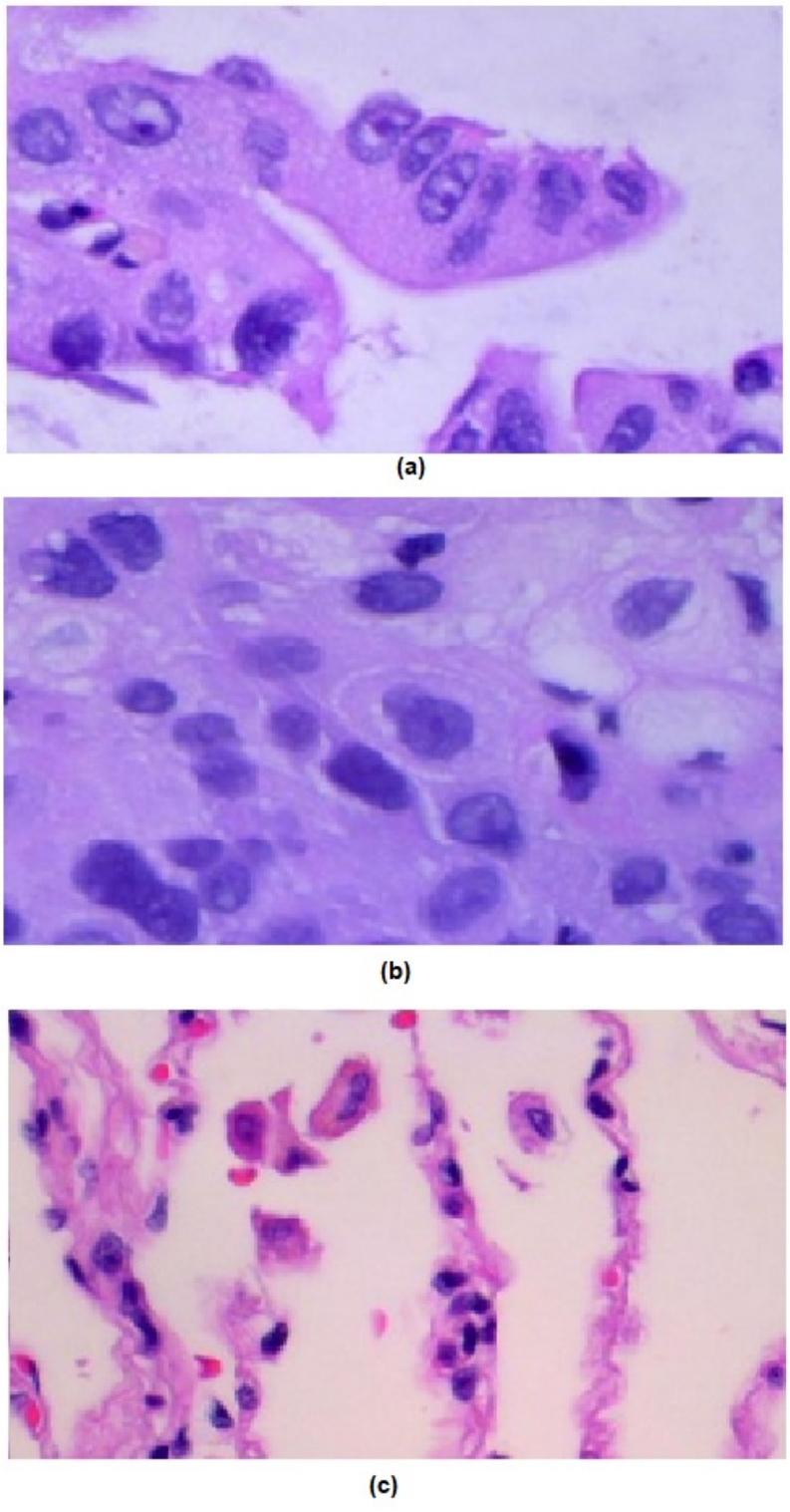


Fig. 1: Examples of images from the dataset for each cancer type used in the study. (a) Lung adenocarcinoma from the dataset, (b) Squamous cell lung cancer from the dataset, and (c) Benign tumor image from the dataset.

Each color channel is assigned an 8-bit binary value, which represents the intensity of that specific color, ranging from 0 (minimum intensity) to 255 (maximum intensity). To enhance model convergence, a normalization step was applied to the pixel intensity values. There are two prevalent normalization techniques [3]:

- Standardization: transforming the values to follow a normal distribution with a mean of 0 and a standard deviation of 1.
- Min-max scaling: rescaling the pixel intensities to the range [0, 1] by dividing each value by 255, the maximum possible value.

This study adopted the min-max scaling normalization approach. Following the application of compression filtering and normalization, each image was represented as a 100x100x3 NumPy array, wherein pixel values ranged from 0 to 1, rendering them suitable for input into a CNN.

2.3 Training, validation, and test set split

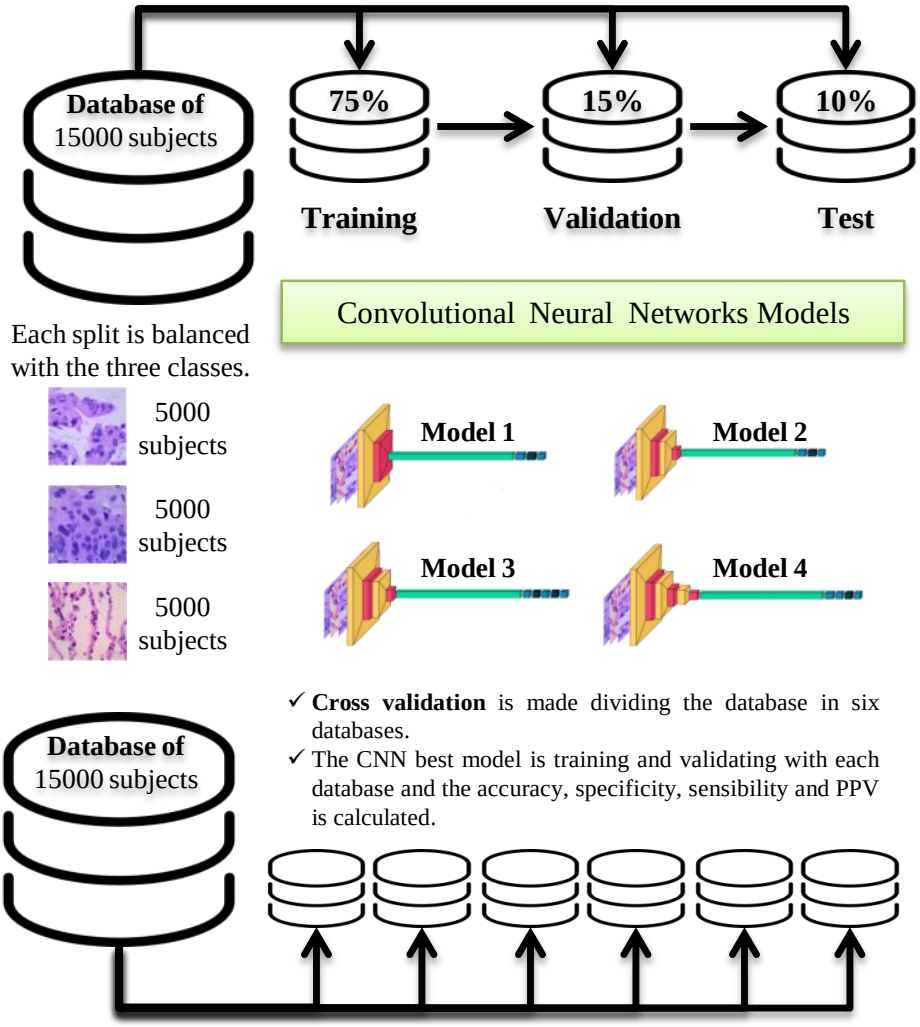
For a robust model training and evaluation, the dataset was randomly partitioned into three distinct sets: training, validation, and testing. This split ensured that the model learned from representative data and avoided biases. A split ratio of 75% for training, 15% for validation, and 10% for testing was employed [6]. The training set was used for training the models, the validation set for evaluating model performance during training, and the testing set for assessing the model's ability to generalize to new data (See Figure 2).

A training and testing step was performed for each model, and the results were used to identify the model with the best performance. The consistency of the top-performing model was handled by using the cross-validation technique (See Figure 2). The dataset was carefully balanced to maintain a consistent representation of each lung cancer type across all sets [5]. Table 1 presents the dataset breakdown for the different sets. By employing this three-way split and maintaining class balance, the model receives a diverse training set, learns effectively, and generalizes well to unseen data.

Table 1: Distribution of Images in Datasets by Cancer or Tumor Type used for this study

Number of Images per set	Total (100%)	Lung Adenocarcinoma	Benign Tumor	Squamous Cell Carcinoma
Training Set	11,251	3,780 (33.5%)	3,720 (33.1%)	3,751 (33.3%)
Validation Set	2,250	714 (31.7%)	780 (34.7%)	756 (33.6%)
Testing Set	1,499	506 (33.8%)	500 (33.3%)	493 (32.9%)

- ✓ Random selection for subjects in training, validation and test.
- ✓ With each model the accuracy, specificity, sensibility and PPV is calculated.



- ✓ **Cross validation** is made dividing the database in six databases.
- ✓ The CNN best model is training and validating with each database and the accuracy, specificity, sensibility and PPV is calculated.

Fig. 2: Training, validation and test split.

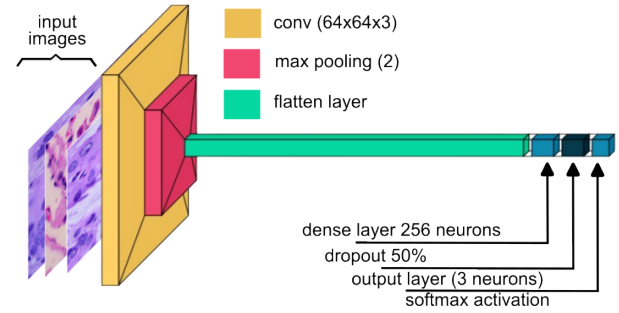
2.4 Convolutional Neural Network Models Training

Multiple CNN models were created and empirically tested to identify the most effective configurations, leading to improved metrics compared to existing literature. Among the eighteen models examined, four emerged as top performers, featuring a sequential architecture with convolutional and max pooling layers for feature extraction and dimensionality reduction, respectively. Each model consisted of a sequential architecture, where feature extraction was performed by a convolutional layer, followed by dimensionality reduction via a max pooling layer. Then, the data was flattened to prepare them for a dense layer that fed into the final output layer. The output layer consisted of three neurons using a softmax activation function to determine the probability distribution across three classification categories: benign tumor, adenocarcinoma, and squamous cell lung cancer.

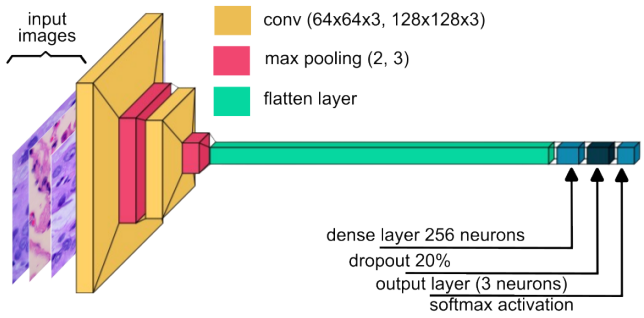
A thorough hyperparameter tuning process was conducted to optimize the neural network architecture's balance between performance and complexity. This process involved experimenting with various variables, including neuron numbers in each layer and kernel sizes in convolutional and pooling layers, utilizing four different models, described as follows:

- **Model 1**: Convolutional Layer with 64 kernels of size 3. Pooling Layer with filters of size 2. Flatten Layer. Dense Layer with 256 neurons and a sigmoid activation function. Dropout 50%. Output Layer with 3 neurons and softmax activation function (see Fig. 3(a)).
- **Model 2**: Convolutional Layer with 64 kernels of size 3. Pooling Layer with filters of size 2. Convolutional Layer with 128 kernels of size 3. Pooling Layer with filters of size 3. Flatten Layer. Dense Layer with 256 neurons and a relu activation function. Dropout 20%. Output Layer with 3 neurons and softmax activation function (see Fig. 3(b)).
- **Model 3**: Convolutional Layer with 64 kernels of size 3. Pooling Layer with filters of size 2. Convolutional Layer with 128 kernels of size 3. Pooling Layer with filters of size 3. Flatten Layer. Dense Layer with 128 neurons and a relu activation function. Dropout 20%. Dense Layer with 32 neurons and a relu activation function. Dropout 20%. Output Layer with 3 neurons and softmax activation function (see Fig. 3(c)).
- **Model 4**: Convolutional Layer with 64 kernels of size 3. Pooling Layer with filters of size 2. Convolutional Layer with 128 kernels of size 3. Pooling Layer with filters of size 2. Convolutional Layer with 256 kernels of size 3. Pooling Layer with filters of size 2. Flatten Layer. Dense Layer with 256 neurons and a relu activation function. Dense Layer with 32 neurons and a relu activation function. Dropout 20%. Output Layer with 3 neurons and softmax activation function (see Fig. 3(d)).

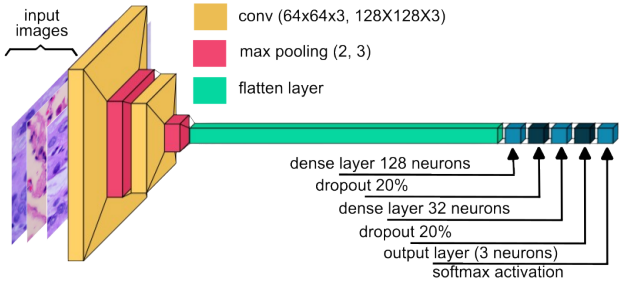
The models underwent training, validation, and testing using distinct datasets. The best model was selected and subjected to a six-fold cross-validation within the database to assess its robustness and detect biases from random data splitting. Sensitivity, specificity, positive predictive value, and accuracy metrics were



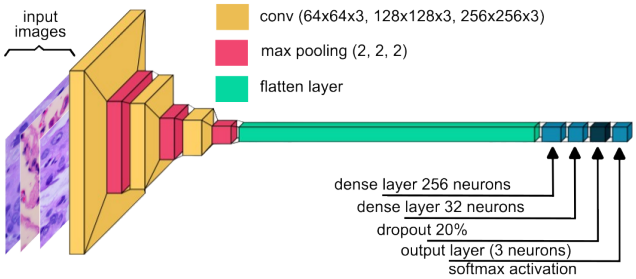
(a)



(b)



(c)



(d)

Fig. 3: CNN models architecture (a) model 1, (b) model 2, (c) model 3, and (d) model 4.

calculated for each cross-validation iteration. The outcomes are then reported as mean values and standard deviations. A Dropout layer was added to the models to reduce overfitting by randomly deactivating connections between neurons during training, enhancing the model's robustness and generalization capability. The network's output, generated using softmax activation in the output layer, indicates the probability of the input's classification. Performance assessment involves using Categorical Cross-Entropy loss, which measures loss based on predicted class probabilities and actual labels.

During training, the priority was to minimize the loss function and determine the optimal weights for the network's connections and convolutional kernels using the Adam optimizer. Adam dynamically adjusts the learning rate for each parameter based on historical gradients, controlling weight update magnitudes according to the error signal with a default learning rate of 0.001. Also, 32 samples per batch were used with the average error calculated over these mini-batches for 100 epochs. Models were trained on the training set and evaluated on the validation set to assess performance in classifying benign tumors, adenocarcinomas, and squamous cell lung cancers.

To ensure optimal performance during training, a mechanism was implemented to save the model weights whenever the prediction accuracy on the training set improved. This approach guaranteed that upon completion, a checkpoint file (.ckpt extension) is generated, containing the weights from the model iteration that achieved the highest accuracy on the training data. Additionally, accuracy and loss values for both the training and validation sets were logged in a comma-separated values (CSV) file after each epoch, facilitating a detailed analysis of the training process. This dual strategy of saving the best weights and logging performance metrics enhanced the model's reliability and provided valuable insights into its learning dynamics.

The evolution of accuracy values across training epochs was meticulously examined to assess the models' performance. The goal was to identify inflection points where continued training might lead to overfitting on the training data. To visualize these trends, helper functions were developed to plot the accuracy and loss values (retrieved from the CSV files) for both the training and validation sets, enabling effective monitoring of the learning process. Finally, the classification performance of each model was compared with their corresponding training times to identify potential trade-offs between accuracy and training efficiency. This approach provided valuable insights into optimizing both the performance and efficiency of the models.

2.5 Classifier Indexes

To evaluate the performance of each model, true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) were calculated. These metrics were used to visualize classification discrepancies in the models [13]. Additionally,

standard indexes such as accuracy, specificity, sensitivity, and positive predictive value (PPV) were employed for further assessment.

3 Results

Tables 2, 3, and 4 display the accuracy, sensitivity, PPV and specificity of the training and validation results for the four CNN models, categorized into three classes. Tables 5, 6 and 7 present the accuracy, sensitivity, specificity and PPV of the training and validation obtained from applying cross-validation for the model 4 which were the model with best performance. Finally, the Table 8 shows the results of testing of model 4.

Table 2: Results of training (T) and validation (V) for Lung Adenocarcinoma class for the model 1 (M1), model 2 (M2), model 3 (M3) and model 4 (M4).

Indexes	**T-M1**	**V-M1**	**T-M2**	**V-M2**	**T-M3**	**V-M3**	**T-M4**	**V-M4**
Accuracy	0.975	0.931	0.947	0.932	0.966	0.952	1.000	0.971
Sensibility	0.992	0.948	0.933	0.933	0.931	0.898	1.000	0.957
PPV	0.938	0.852	0.911	0.883	0.966	0.948	1.000	0.953
Specificity	0.967	0.923	0.954	0.944	0.984	0.977	1.000	0.978

Table 3: Results of training (T) and validation (V) for Benign tumor class for the model 1 (M1), model 2 (M2), model 3 (M3) and model 4 (M4).

Indexes	**T-M1**	**V-M1**	**T-M2**	**V-M2**	**T-M3**	**V-M3**	**T-M4**	**V-M4**
Accuracy	1.000	0.995	0.996	0.995	1.000	0.999	1.000	0.998
Sensibility	1.000	0.991	1.000	1.000	1.000	0.999	1.000	0.996
PPV	0.999	0.995	0.987	0.986	1.000	0.997	1.000	0.997
Specificity	0.967	0.997	0.954	0.993	0.984	0.999	1.000	0.999

Fig. 4 illustrates the evolution of classification accuracy for both training set and validation set samples throughout the training process. Specifically, this figure depicts the accuracy of the machine learning model on the training and validation datasets as a function of the number of epochs. Fig. 5 presents the interface designed to upload histopathological images and execute a prediction algorithm. This interface enables users to employ the classifier in a straightforward and practical manner, while also providing the accuracy of the classification.

These results underscore the model's potential for generating predictions from histology images without the need for healthcare professional involvement, thereby enhancing efficiency in the healthcare system and facilitating the detection of this disease. This can ultimately lead to improved disease development and treatment

Table 4: Results of training (T) and validation (V) for Squamous Cell Carcinoma class for the model 1 (M1), model 2 (M2), model 3 (M3) and model 4 (M4).

Indexes	**T-M1**	**V-M1**	**T-M2**	**V-M2**	**T-M3**	**V-M3**	**T-M4**	**V-M4**
Accuracy	0.976	0.936	0.951	0.936	0.966	0.953	1.000	0.973
Sensibility	0.934	0.852	0.908	0.908	0.967	0.955	1.000	0.959
PPV	0.993	0.951	0.944	0.922	0.933	0.910	1.000	0.962
Specificity	0.997	0.978	0.973	0.962	0.965	0.952	1.000	0.981

Table 5: Cross-validation results for Lung Adenocarcinoma class for model 4.

Indexes	**Training**	**Validation**
Accuracy	0.9903 ± 0.0035	0.9896 ± 0.0169
Sensibility	0.9875 ± 0.0038	0.9900 ± 0.0123
PPV	0.9839 ± 0.0107	0.9788 ± 0.0383
Specificity	0.9918 ± 0.0056	0.9895 ± 0.0193

Table 6: Cross-validation results for benign tumor class for model 4.

Indexes	**Training**	**Validation**
Accuracy	0.9990 ± 0.0012	0.9987 ± 0.0030
Sensibility	0.9992 ± 0.0007	1.0000 ± 0.0000
PPV	0.9978 ± 0.0040	0.9962 ± 0.0084
Specificity	0.9918 ± 0.0056	0.9980 ± 0.0046

Table 7: Cross-validation results for Squamous Cell Carcinoma class for model 4.

Indexes	**Training**	**Validation**
Accuracy	0.9912 ± 0.0029	0.9910 ± 0.0140
Sensibility	0.9845 ± 0.0112	0.9786 ± 0.0391
PPV	0.9898 ± 0.0055	0.9943 ± 0.0043
Specificity	0.9948 ± 0.0029	0.9972 ± 0.0020

Table 8: Indexes of test results for model 4 in each class.

Indexes	ACA	Benign Tumor	SCC
Accuracy	0.9840	1.0000	0.9840
Sensibility	0.9862	1.0000	0.9655
PPV	0.9671	1.0000	0.9855
Specificity	0.9829	1.0000	0.9930

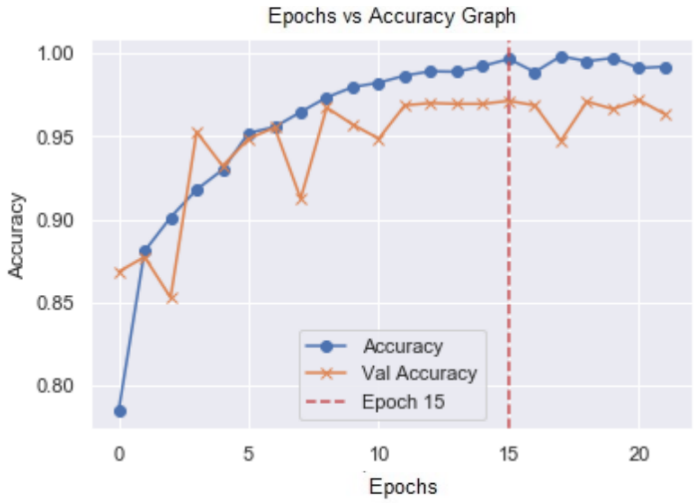


Fig. 4: Accuracy of a machine learning model on the training and validation datasets as a function of the number of epochs.

outcomes. To facilitate the model's use, a simple and intuitive interface has been designed, enabling any personnel to obtain predictions from images.

Figure 5 presents the interface designed to upload images and implement the prediction algorithm. This user-friendly interface allows users to easily employ the classifier. The interface features a button that opens the machine's file explorer, enabling users to select an image. Once selected, the image is displayed alongside the model's prediction for the type of cancer. Moreover, the interface displays the probability returned by the softmax activation function of the model's output layer for that class.

4 Discussion

This study focuses on the development of convolutional neural network (CNN) models for accurately classifying lung cancer images into different categories, including benign lung tumors, lung adenocarcinoma, and squamous cell lung cancer. The main findings highlight the high accuracy achieved by our models, supported

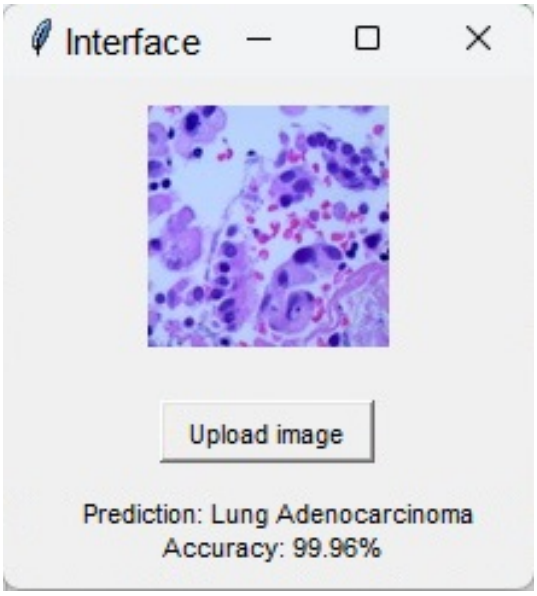


Fig. 5: Prediction result by the interface.

by robust performance metrics such as accuracy, sensitivity, specificity, and positive predictive value.

Methodologically, we employed meticulous data preprocessing techniques, careful dataset selection, innovative model architectures, rigorous training processes, and comprehensive evaluation methods. We stress the importance of data quality, preprocessing methodologies, and model optimization in achieving accurate and reliable results.

In general, the results show that all models have accuracy, sensitivity, positive predictive value, and specificity indexes above 90% for most types of cancer. This indicates that all models are quite good at detecting most types of lung cancer. However, Model 4 has the best overall results, with an accuracy of over 97% for all types of cancer. Model 4 also has the best sensitivity over 95.7%. The next best model is Model 3, and it also has good results, with an accuracy, sensitivity, PPV, and specificity of over 95% for all types of cancer.

All four models demonstrated excellent performance in classifying benign tumors. Model 4, in particular, achieved a remarkable 100% for all indexes in the test group (see Table 8). To further validate that Model 4's performance was indeed due to its intrinsic capabilities rather than the selection of subjects in the training and validation sets, cross-validation was conducted. The results of the cross-validation are presented in Tables 5, 6, and 7. The reported indexes include their standard deviations, which are all in the order of thousandths, indicating minimal differences in indexes across the different training and test groups. These findings strongly suggest that Model 4's

exceptional performance stems from its inherent ability to accurately classify benign tumors, scamous cell carcinoma, and lung adenocarcinoma.

The observed variations in the models performance could be attributed to several factors related to their architectural characteristics. Model 4 stands out with three convolutional layers, while the other models have only two. This additional layer allows Model 4 to extract more complex features from the histopathological images, potentially leading to improved cancer detection capabilities. Models 2, 3, and 4 employ the Rectified Linear Unit (ReLU) activation function in their hidden dense layers, while Model 1 utilizes the sigmoid function. ReLU is generally considered more effective for training deep neural networks, as it helps mitigate the vanishing gradient problem, a common challenge in training deep learning models. Finally, Models 2, 3, and 4 incorporate Dropout regularization into their dense layers. This technique helps prevent overfitting and enhances the model's generalization ability, allowing it to perform better on unseen data.

This research has significant implications for medical imaging and lung cancer diagnosis. By leveraging artificial intelligence algorithms and CNN models, we demonstrate how early detection of lung cancer can be improved, leading to better clinical outcomes for patients.

5 Conclusion

This study significantly contributes to advancing the field of medical imaging and artificial intelligence in healthcare as a direct application through a decision support system. By showcasing the efficacy of AI-driven approaches in lung cancer detection, we aim to facilitate early diagnosis of lung cancer and ultimately improve patient outcomes. This research underscores the importance of collaboration, innovation, and interdisciplinary research in driving scientific progress in healthcare.

In essence, the study objectives were successfully achieved. The employed methodology yielded promising results and reinforces the value of intelligent systems in medicine. We identified key challenges and limitations that offer avenues for future model improvement and broader applicability within decision support systems. This work contributes significantly to the field and aligns with the personal and academic goals of this aspiring Biomedical Engineer.

Future Works

Looking ahead, future challenges include improving model performance through the exploration of new techniques or algorithms, expanding dataset sizes, leveraging more powerful computational resources, and refining model architectures to handle higher resolution images. Additionally, modifications to existing architectures can facilitate multi-class classification of diseases with similar image characteristics, thereby broadening the spectrum of diseases detectable by our system. To carry

out the study, a biopsy is necessary to obtain the histology, avoiding the need for numerous tests that require the constant and cumulative application of radiation, which promotes the use of this system.

The implementation of the algorithms developed in this study with more powerful computational machines will allow the creation of architectures designed to work with higher resolution images. This will avoid the need to preprocess them by reducing their dimensionality and applying noise to them. In this way, by increasing the accuracy of the data, we will generate an improvement in the models, avoiding the loss of information necessary for the classification task.

The architectures used have been focused on the classification task according to the dataset used. Small variations in these architectures, by increasing the number of neurons in the output layer, can be used in the multi-class classification of different diseases with similar images, expanding the variety of diseases recognizable by the developed system.

Acknowledgements The authors would like to express their sincere gratitude to Universidad de Ibagué for their invaluable collaboration ad honorem of this research. Additionally, we extend our appreciation to Universidad Internacional de Valencia in Spain and Universidad Simón Bolívar in Caracas, Venezuela for their support and contributions.

Competing Interests The authors have no conflicts of interest to declare that are relevant to the content of this chapter.

Ethics Approval This study adheres to the principles outlined in the Declaration of Helsinki. Noting that the use of anonymized patient data without explicit consent is widely accepted in research practices, particularly when adherence to strict anonymization protocols and ethical standards is maintained. By prioritizing patient privacy and confidentiality, while upholding the integrity of scientific inquiry, this study endeavors to generate valuable insights for the benefit of both current and future patients.

References

1. Abdelsamea, M.M., Zidan, U., Senousy, Z., Gaber, M.M., Rakha, E., Ilyas, M.: A survey on artificial intelligence in histopathology image analysis. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery **12**(6), e1474 (2022)
2. Aguilar, E., La Cruz, A., Albertti, R., Carnier, M., Gavidia, L., Severeyn, E.: Detection of respiratory disease patterns using mask r-cnn. In: Proceedings of Seventh International Congress on Information and Communication Technology: ICICT 2022, London, Volume 2, pp. 739–750. Springer (2022)
3. Ballé, J., Laparra, V., Simoncelli, E.P.: Density modeling of images using a generalized normalization transformation. arXiv preprint arXiv:1511.06281 (2015)
4. Borkowski, A.A., Bui, M.M., Thomas, L.B., Wilson, C.P., DeLand, L.A., Mastorides, S.M.: Lung and colon cancer histopathological image dataset (lc25000). arXiv preprint arXiv:1912.12142 (2019)
5. Cui, Y., Jia, M., Lin, T.Y., Song, Y., Belongie, S.: Class-balanced loss based on effective number of samples. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 9268–9277 (2019)
6. Domingo, J.D., Aparicio, R.M., Rodrigo, L.M.G.: Cross validation voting for improving cnn classification in grocery products. IEEE Access **10**, 20913–20925 (2022)

7. Hamamoto, R., Suvarna, K., Yamada, M., Kobayashi, K., Shinkai, N., Miyake, M., Takahashi, M., Jinnai, S., Shimoyama, R., Sakai, A., et al.: Application of artificial intelligence technology in oncology: Towards the establishment of precision medicine. Cancers **12**(12), 3532 (2020)
8. Hatuwal, B.K., Thapa, H.C.: Lung cancer detection using convolutional neural network on histopathological images. Int. J. Comput. Trends Technol **68**(10), 21–24 (2020)
9. Karim, D.Z., Bushra, T.A.: Detecting lung cancer from histopathological images using convolution neural network. In: TENCON 2021-2021 IEEE Region 10 Conference (TENCON), pp. 626–631. IEEE (2021)
10. Mazurowski, M.A., Habas, P.A., Zurada, J.M., Lo, J.Y., Baker, J.A., Tourassi, G.D.: Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance. Neural networks **21**(2-3), 427–436 (2008)
11. Siddique, A., Otaibi, F.M., Khan, S.F.: Lung cancer in developing countries. In: Handbook of Medical and Health Sciences in Developing Countries: Education, Practice, and Research, pp. 1–28. Springer (2023)
12. Tang, S., Yuan, S., Zhu, Y.: Data preprocessing techniques in convolutional neural network based on fault diagnosis towards rotating machinery. IEEE Access **8**, 149487–149496 (2020)
13. Xu, J., Zhang, Y., Miao, D.: Three-way confusion matrix for classification: A measure driven view. Information sciences **507**, 772–794 (2020)