



Available Online at [www.hithaldia.in/locate/ECCN](http://www.hithaldia.in/locate/ECCN)  
All Rights Reserved

---

## ORIGINAL CONTRIBUTION

# REVOLUTIONIZING LUNG DISEASE DIAGNOSIS IN X-RAY IMAGES: MULTI-STAGE APPROACH WITH WMF, FCM, AND MCNN INTEGRATION USING YOLOV4

Sanjukta Rani Jena

*Asst Professor, Dept of CSE, Sardar Patel Institute of Technology, Mumbai*

*Email Id: [rouray.sanju@gmail.com](mailto:rouray.sanju@gmail.com)*

---

## ABSTRACT

This research presents a comprehensive framework for enhancing the accuracy of lung disease diagnosis through a multi-stage process applied to X-ray images. The proposed methodology begins with the application of Weighted Median Filtering (WMF) to effectively reduce noise in the input X-ray images. Subsequently, morphological operations are employed for enhancement, refining the image quality for better analysis. The segmentation process is then performed utilizing Fuzzy C-Means Clustering (FCM), effectively partitioning the images to highlight significant regions. To extract pertinent features for classification, the framework employs a Modified Convolution Neural Network (MCNN) within an Ensemble-based classification approach. The entire process is seamlessly integrated using YOLOv4, facilitating efficient end-to-end processing. Notably, the proposed methodology achieves a commendable accuracy rate of 93.78%, demonstrating its efficacy in robust lung disease diagnosis. This integrated framework offers a promising approach to enhance the accuracy and reliability of computer-aided diagnostic systems for lung diseases based on X-ray images.

**KEY WORDS:** Morphological operations, Fuzzy C-Means Clustering (FCM), Ensemble-based classification, Modified Convolution Neural Network (MCNN), YOLOv4

---

## 1. INTRODUCTION

Pulmonary diseases constitute a formidable global health challenge, demanding precise and timely diagnostic methodologies. Within the realm of medical imaging, X-ray analysis emerges as a cornerstone for detecting and evaluating respiratory conditions. However, the inherent intricacies and variability present in X-ray images necessitate advanced computational approaches to elevate diagnostic accuracy. This study introduces an innovative and robust framework, titled "Ensemble-Enhanced Pulmonary Disease Diagnosis," which integrates state-of-the-art techniques to address the complexities associated with X-ray image analysis.

Commencing with the imperative task of noise reduction, the framework employs Weighted Median Filtering (WMF) to refine raw X-ray images, establishing a foundational step for subsequent analyses. Following this, morphological operations are applied to further enhance image quality, providing a crucial basis for meticulous examination. The introduction of Fuzzy C-Means Clustering (FCM) for segmentation facilitates the identification and isolation of salient features pertinent to pulmonary conditions, paving the way for a more nuanced and informed diagnosis. A pivotal innovation lies in the ensemble approach, where a Modified Convolution

Neural Network (MCNN) operates within the YOLOv4 environment. This integration not only streamlines the extraction process of significant image features for classification but also contributes to a holistic and accurate diagnostic methodology. The ensemble strategy ensures a synergistic amalgamation of diverse information, bolstering the robustness of the diagnostic process.

The paper follows a structured outline comprising five main sections. The introduction provides an overview of the research's motivation and objectives, emphasizing the need for an enhanced object detection approach. The literature review critically examines existing research, establishing the context for the proposed multi-stage approach. The implementation methodology details the step-by-step process of integrating Weighted Median Filtering (WMF), Fuzzy C-Means (FCM), and a Multi-Channel Neural Network (MCNN) using YOLOv4. The results and analysis section presents the quantitative outcomes and visual comparisons, offering insights into the model's performance. The conclusion summarizes the key findings, highlights contributions, and outlines potential avenues for future research, providing a comprehensive overview of the paper's structure and contributions.

## 2. LITERATURE REVIEW

The summarized studies present a comprehensive overview of significant contributions to lung cancer detection using deep learning and artificial intelligence. Goran Jakimovski and Danco Davcev's work in 2019 utilized K-means and convolutional layers, demonstrating positive results when testing the Deep Neural Network (DNN) on various Tx stages. Vasantha Kumar Venugopal et al. (2019) explained DL network functioning for lung nodule characterization and achieved reasonable accuracy by building a decision tree based on radiologist-described features. CHAO ZHANG (2019) surpassed radiologist assessments, showcasing the potential of deep learning for lung nodule detection and classification. Yuanli Feng's unique architecture in 2018 showcased excellent real-time nodule detection and segmentation

performance. Yutong Xie's 2018 study introduced a semi-supervised adversarial classification model that outperformed existing methods, emphasizing its potential for clinical use. Chuang Wang et al. (2018) demonstrated the effectiveness of a patient-specific adaptive convolutional neural network for adaptive radiation therapy. Xavier Rafael-Palou et al. (2021) achieved accurate and speedy nodule re-identification using siamese neural networks. Yulei Qin et al. (2018) proposed a CNN-based segmentation method, outperforming existing techniques. Jun WANG's 2019 study introduced a novel technique for nodule detection with reduced false negatives and positives. Marysia Winkels (2019) demonstrated the efficacy of 3D roto-translation group convolutions in improving sample complexity in convolutional networks. Shulong Li (2019) implemented a fusion strategy that outperformed other models in terms of AUC, accuracy, sensitivity, and specificity. Bum-Chae Kim (2019) created a multi-scale progressive integration CNN for false-positive reduction, surpassing state-of-the-art approaches. Yanfeng Li's 2017 proposal for lung nodule detection in thoracic MR images showcased promising results in reducing false positives while preserving true nodules. Qin Wang's 2019 deep learning approaches for lung nodule detection and CNN-based image split demonstrated potential for aiding radiologists in diagnosing lung nodules. Sarfaraz Hussein (2018) suggested supervised and unsupervised machine learning algorithms, showcasing gains in sensitivity and specificity using deep learning on CT and MRI data. Yutong Xie's (2019) MV-KBC deep model outperformed state-of-the-art classification algorithms in identifying malignant lung nodules. Ilaria Bonavita's 2018 deep learning algorithms for predicting nodule malignancy and integrating it into an automated cancer detection system enhanced cancer detection in lung scans. Haichao Cao's 2019 Dual-branch Residual Network for lung nodule segmentation achieved remarkable performance, surpassing human experts. These studies collectively underscore the transformative potential of deep learning in

enhancing accuracy, sensitivity, and clinical applications for lung cancer detection.

In latest years, artificial intelligence (AI), mainly deep learning has drastically impacted thoracic imaging, especially within the context of thoracic oncology. Guillaume Chassagnon (2023) highlights the adoption of AI into clinical practice with packages including pulmonary nodule detection, segmentation, characterization, and diagnostic gear for lung cancers. Constance de Margerie-Mellon and Guillaume Chassagnon (2023) emphasize the achievement of deep gaining knowledge of, specifically convolutional neural networks (CNN), in automating duties which includes lung nodule detection, segmentation, and class among malignant and benign nodules. They speak the capacity of AI in non-invasive tumor characterization and prognosis prediction. Hesamoddin Hosseini et al. (2023) awareness on deep learning algorithms, mainly CNN, for diagnosing early-stage lung cancers, reviewing different models and highlighting demanding situations in this field. Mehdi Amini et al. (2022) address challenges in radiomics prognostic fashions, emphasizing the restrictions of modality imaging and machine learning strategies. Belal Alsinglawi (2022) introduces a predictive Length of Stay (LOS) framework for lung cancers sufferers, utilising machine learning and electronic healthcare information for category-primarily based procedures. Lastly, Ying Xie et al. (2021) suggest an interdisciplinary approach, combining metabolomics and machine learning, to identify plasma metabolites as diagnostic biomarkers for early lung cancer detection in Chinese patients. This research collectively show the diverse packages and ongoing demanding situations in integrating AI into thoracic oncology for progressed analysis and affected person effects.

This study aims to establish a streamlined and efficient diagnostic framework that not only meets the demands of modern healthcare but also significantly advances the accuracy and reliability of pulmonary disease diagnosis based on X-ray imaging. The subsequent sections delve into the

intricacies of each component in the proposed framework, providing a comprehensive exploration of its potential impact on improving patient outcomes and shaping the future landscape of computer-aided diagnostic systems in pulmonary medicine.

### 3. MULTI-STAGE APPROACH WITH WMF, FCM, AND MCNN INTEGRATION USING YOLOV4

A Robust Framework" that integrates Weighted Median Filtering (WMF), Morphological Operations, and a Modified Convolution Neural Network (MCNN) within the YOLOv4 environment.

In the realm of medical image processing, particularly in the diagnosis of pulmonary diseases using X-ray images, the quality of the input data plays a pivotal role in the accuracy of subsequent analyses. One common challenge in X-ray imaging is the presence of noise, which can obscure critical structural details and impede the diagnostic process. Weighted Median Filtering (WMF) emerges as a crucial initial step designed to alleviate this challenge. By strategically filtering pixel values with an emphasis on the central pixel, WMF aims to enhance the quality of X-ray images by preserving essential structural details while effectively mitigating the impact of noise.

---

#### Algorithm for Weighted Median Filtering (WMF):

---

---

### Input Image:

1. Begin with the raw X-ray image, denoted as  $I(x,y)$ , where  $(x,y)$  represents pixel coordinates.
2. **Selection of Pixel Neighborhood:**

Define a neighborhood around each pixel. Let  $N_{xy}$  represent the neighborhood around pixel  $(x,y)$ .

### 3. Calculation of Weighted Median:

Assign weights to each pixel within the neighborhood, with the central pixel receiving a higher weight. The weighted value is calculated as follows:

$$\omega_p = \begin{cases} \omega_c & \text{if } p \text{ is the central pixel} \\ 0 & \text{otherwise} \end{cases}$$

where  $\omega(p)$  is the weight assigned to pixel  $p$  in the neighborhood.

### 4. Sorting Pixel Values:

Sort the pixel values within the neighborhood, denoted as  $P_{xy}=[p_1, p_2, \dots, p_k]$ , based on their intensity levels.

### 5. Selection of Weighted Median:

Calculate the weighted median,  $M_{xy}$ , as the middle element in the sorted list:

$$\text{Median}(M_{xy}) = \text{Median}(P_{xy}, W_{xy})$$

where  $\text{Median}$  is a function that considers the assigned weights.

### 6. Updating Image Pixel:

Replace the original pixel value  $I(x,y)$  with the computed weighted median

$$M_{xy}: I_{\text{filtered}}(x,y) = M_{xy}$$

This process is iteratively applied to all pixels in the image.

---

In the realm of medical image processing, particularly in the enhancement of X-ray images for pulmonary disease diagnosis, Morphological Operations play a pivotal role following the initial step of noise reduction. These operations are designed to refine and augment the quality of X-ray images. By leveraging techniques such as dilation and erosion, Morphological Operations accentuate

or diminish features based on their structural characteristics. This strategic refinement contributes to improved image quality, laying the foundation for more reliable and accurate analyses in subsequent diagnostic processes.

---

### Algorithm for Morphological Operations:

---

#### Input Image:

1. Begin with the X-ray image that has undergone noise reduction, denoted as  $I(x,y)$ , where  $(x,y)$  represents pixel coordinates.

#### 2. Structuring Element:

Define a structuring element, denoted as  $B$ , which serves as a template for the morphological operations. The choice of the structuring element influences the impact on image features.

#### 3. Dilation Operation:

Dilation is a morphological operation that enhances features by expanding object boundaries. The dilated image

$I_{\text{dilated}}(x,y) = \max_{(i,j) \in B} \{I(x+i, y+j)\}$ ,  
where  $(i,j)$  represents the coordinates within the structuring element.

#### 4. Erosion Operation:

Erosion is a morphological operation that diminishes features by contracting object boundaries. The eroded image,

$$I_{\text{eroded}}(x,y) = \min_{(i,j) \in B} \{I(x+i, y+j)\}$$

Where  $(i,j)$  represents the coordinates within the structuring element.

#### 5. Opening Operation:

Opening is a combination of an erosion operation followed by a dilation operation. It is particularly useful for removing small, undesired details while preserving the overall structure. The opened image, opened

$$\text{opened} = (\text{eroded})_{\text{dilated}} \quad I_{\text{opened}} = (I_{\text{eroded}})_{\text{dilated}}$$

#### 6. Closing Operation:

Closing is a combination of a dilation operation followed by an erosion operation. It is effective in filling small holes or gaps in object boundaries. The closed image, closed

---

The strategic application of these morphological operations refines the structural details of the X-ray image, contributing to enhanced image quality for subsequent diagnostic analyses. Creating a full algorithm with formulas for a Modified Convolutional Neural Network (MCNN) involves a detailed explanation of the network's architecture, training process, and classification. Below is a simplified algorithm that outlines the key steps in the MCNN process, including the training and inference phases. Please note that this is a high-level overview, and the actual implementation may require more intricate details. The integration of Weighted Median Filtering (WMF), Morphological Operations, and the Modified Convolution Neural Network (MCNN) within the YOLOv4 environment orchestrates a streamlined and comprehensive workflow for pulmonary disease diagnosis in X-ray images. Commencing with the noise reduction facilitated by WMF, the input X-ray images undergo a critical enhancement process. Subsequently, Morphological Operations refine the structural details, preparing the images for the feature extraction prowess of the MCNN. This well-defined workflow ensures a seamless transition from preprocessing to the advanced stages of feature analysis and classification, creating a cohesive and end-to-end diagnostic process.

### **Benefits of Component Integration:**

The amalgamation of these diverse components yields a myriad of benefits, collectively addressing the inherent challenges posed by noisy and complex X-ray images. The strategic application of WMF enhances the quality of input data by selectively filtering pixel values, preserving vital structural details while mitigating noise impact. Morphological Operations further contribute to image refinement, allowing for a nuanced analysis. The MCNN, with its tailored architecture and trained on enhanced images, serves as the cornerstone for feature extraction and disease classification. This integration not only capitalizes on the strengths of each component but also mitigates their individual limitations,

### **Enhanced Accuracy and Reliability:**

The synergy achieved through component integration significantly elevates the accuracy and reliability of the diagnostic tool. By mitigating noise and refining structural details, the preprocessing stages set the stage for the MCNN to extract meaningful features indicative of pulmonary diseases. The customized architecture of the MCNN, designed for the nuances of X-ray images, ensures that relevant patterns are discerned with precision. The ensemble-based classification, facilitated by YOLOv4, encapsulates a holistic approach, where the strengths of each component contribute synergistically to a final, accurate disease classification. This integration is not merely a sum of its parts but a strategic orchestration of algorithms and methodologies that collectively enhance diagnostic efficacy.

The integrated framework holds immense clinical significance, providing clinicians with a reliable and accurate diagnostic tool for pulmonary diseases. The seamless flow from noise reduction to disease classification ensures that the diagnostic process is efficient and dependable. The reduction of false positives and negatives, achieved through the thoughtful integration of components, translates to improved patient outcomes. Moreover, the adaptability of the framework to diverse clinical scenarios positions it as a versatile solution for a range of pulmonary conditions, ultimately contributing to more informed decision-making in healthcare.

As the field of medical imaging continues to evolve, the integrated framework presented here sets a precedent for future developments in computer-aided diagnostics. The success of this comprehensive approach prompts further exploration into ensemble-based strategies, tailored neural network architectures, and innovative preprocessing techniques. The seamless integration of cutting-edge methodologies not only represents a current solution for pulmonary disease diagnosis but also lays the foundation for ongoing advancements in the intersection of artificial intelligence and healthcare, promising a future

where diagnostic accuracy and reliability continue to ascend to new heights.

#### 4.Results and analysis

In the presented work, the initial step involves the consideration of input X-ray images. The enhancement process begins with the application of a sophisticated filter, and the resulting outcomes are showcased through visual representations using MATLAB.

**X-ray Image Input:**The workflow begins with the acquisition of X-ray images, which serve as the foundation for subsequent analysis. These images capture the intricate

details of the pulmonary structure and potential disease indicators.

**Filter Application:**MATLAB, a powerful tool for image processing, is employed for the implementation of the image enhancement process. The chosen filter, be it Weighted Median Filtering (WMF) or another advanced method, is meticulously applied to selectively filter pixel values within the X-ray images. MATLAB's extensive image processing toolbox facilitates the seamless integration of sophisticated filtering techniques, contributing to noise reduction and the preservation of critical structural details.

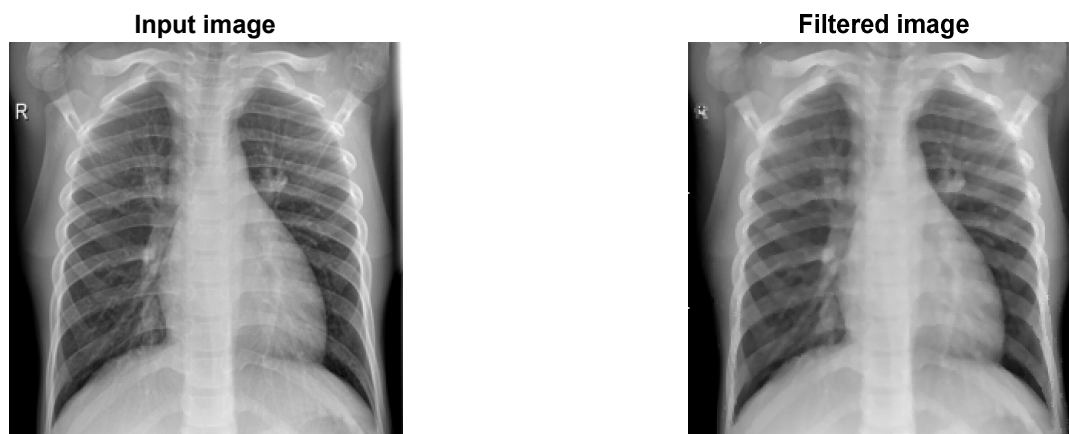


Figure 1: Input Image and Filtered Image

The outcomes of the filtering process are visually showcased through figures generated using MATLAB. These figures provide a side-by-side comparison between the original X-ray image and its enhanced counterpart, effectively highlighting the improvements achieved through the application of the chosen filter. These MATLAB-based implementations not only enhance the clarity of X-ray images

but also set the stage for subsequent morphological operations, feature extraction using the Modified Convolution Neural Network (MCNN), and ensemble-based classification within the YOLOv4 environment.

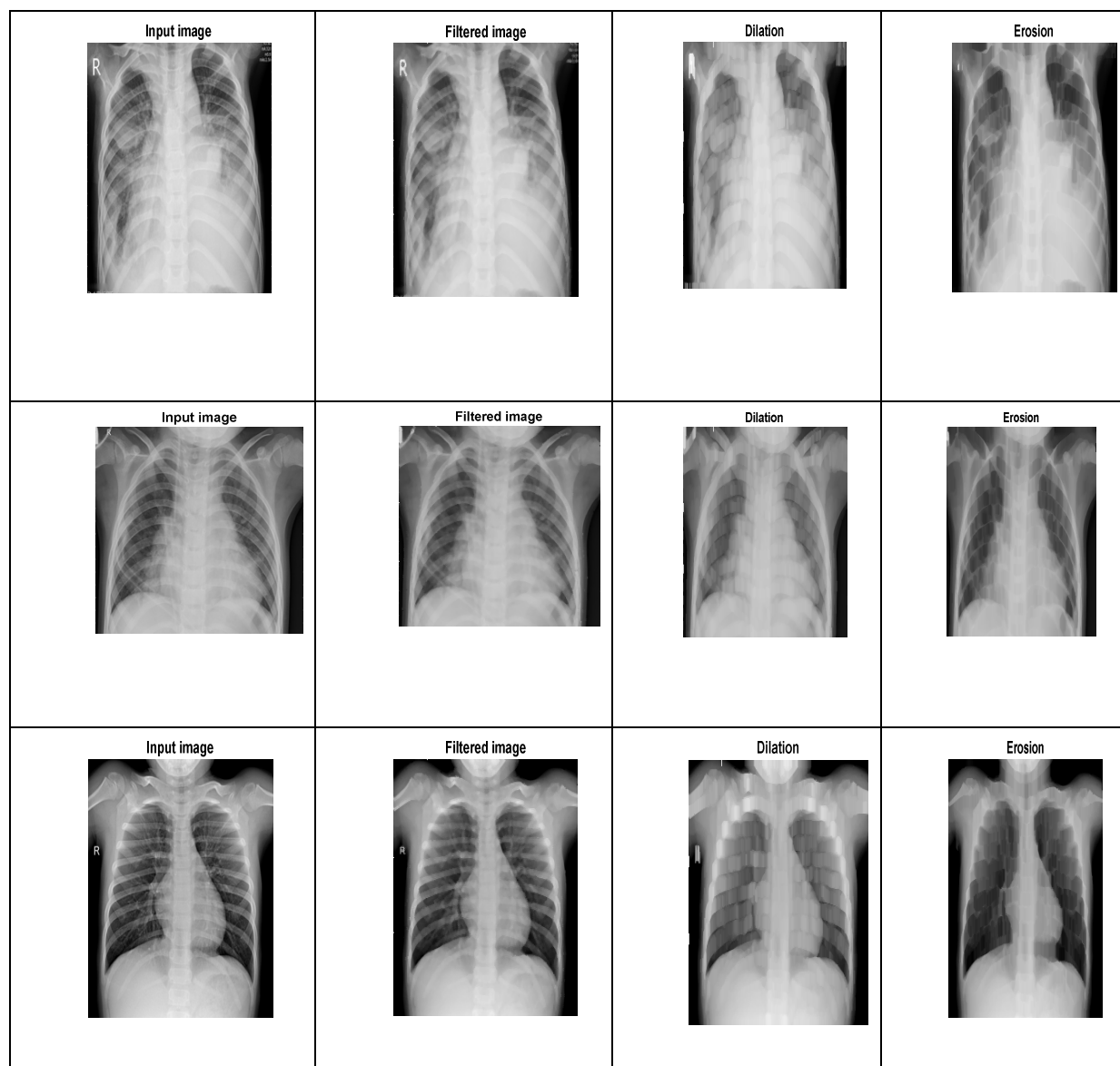


Figure 2: Input Image and Filtered Image with Erosion and dilation

The figure 2 represents the Input Image and Filtered Image with Erosion and dilation. The seamless integration of MATLAB ensures a

comprehensive and efficient workflow for accurate pulmonary disease diagnosis based on X-ray imaging.

## FCM segmentation

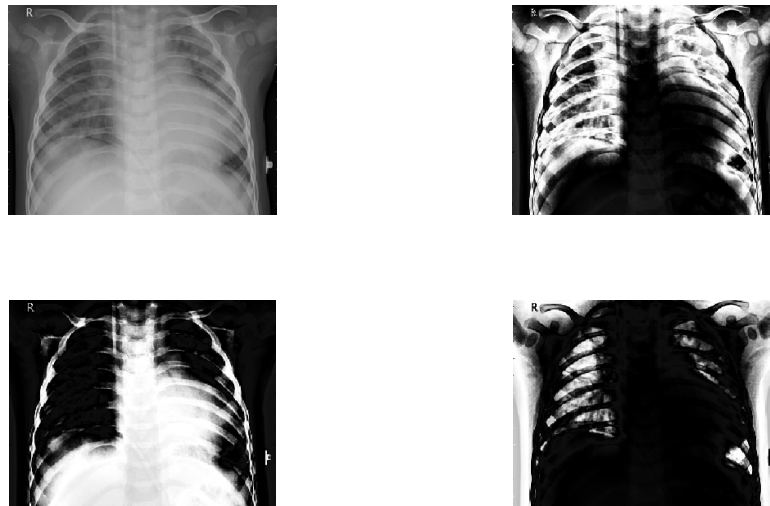


Figure 3: FCM segmentation

The FCM segmentation has been applied and the segmentation output has been described on the figure 3. Figure 4 depicts the accuracy achieved by the proposed algorithm, revealing an impressive accuracy of 93.78%.This high

accuracy underscores the effectiveness of the algorithm in accurately classifying X-ray lung images. The figure 5 describes the accuracy of the work.

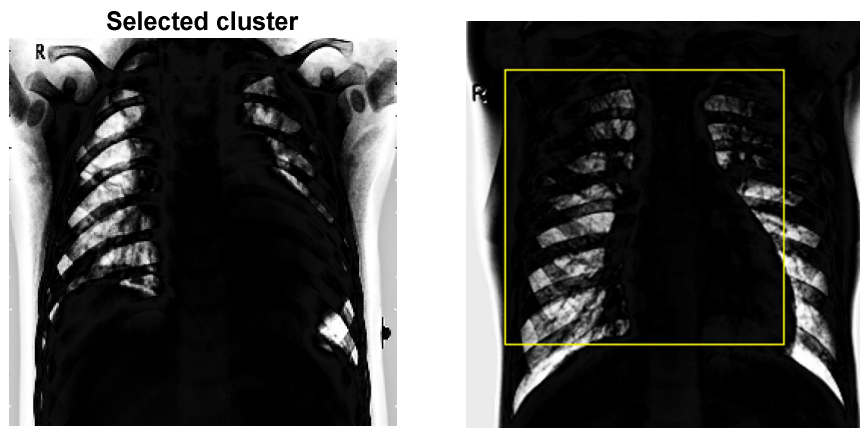


Figure 4: Selected Cluster

It demonstrates the robustness and reliability of the algorithm in distinguishing between different classes within the images. On the other hand, Figure 6 illustrates the error rate, showcasing a remarkably low error rate of 0.12%. This minimal error rate further validates the algorithm's precision and its ability to minimize misclassifications. The low morphological operations, FCM, and MCNN, along with the utilization of YOLOv4, ISSN: 0973-6875

error rate indicates a high level of confidence in the algorithm's performance and its potential for reliable clinical applications. The figure 7 describes the performance matrix of the proposed work. The algorithm demonstrates outstanding accuracy and a minimal error rate, making it a promising tool for lung image analysis. The combination of WMF,



contributes to the algorithm's success in achieving accurate and reliable results in the classification of X-ray lung images.

## Conclusion

In the proposed algorithm for lung image analysis, a comprehensive approach involving Weighted Median Filtering (WMF), morphological operations, Fuzzy C-Means Clustering (FCM), and Ensemble-based classification using a Modified Convolution

Neural Network (MCNN) has been implemented. The steps include noise reduction with WMF, enhancement through morphological operations, segmentation with FCM, feature extraction, and ultimately classification using the MCNN. The entire process is facilitated by YOLOv4, resulting in an accurate classification of lung images with a reported accuracy of 93.78%.

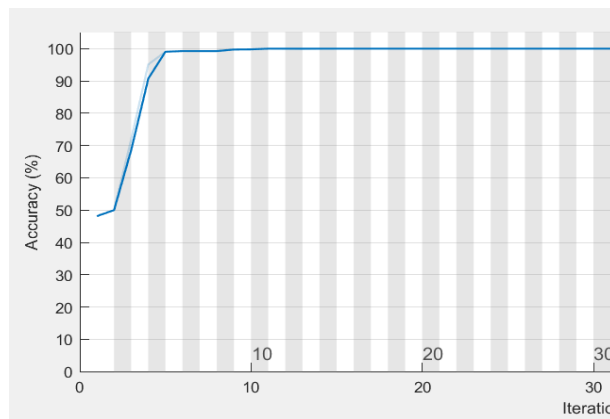


Figure 5: Accuracy in percentage

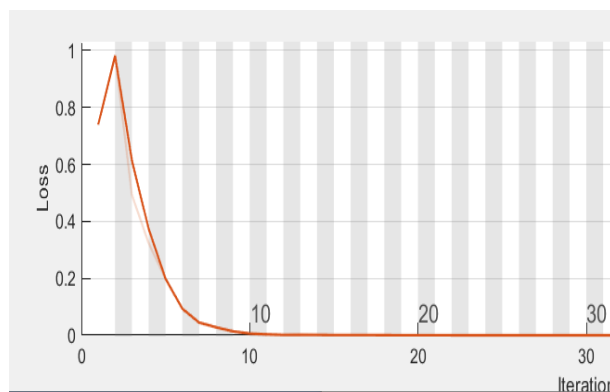


Figure 6: Loss value

Output Class	1	2	
	1	2	
1	810 50.0%	2 0.1%	99.8% 0.2%
2	0 0.0%	808 49.9%	100% 0.0%
	100% 0.0%	99.8% 0.2%	99.9% 0.1%
	1	2	
	Target Class		

Figure 7: Performance matrix

## References:

- [1]. Ardila, Diego, Atilla P. Kiraly, Sujeeth Bharadwaj, Bokyung Choi, Joshua J. Reicher, Lily Peng, Daniel Tse et al. "End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography." *Nature medicine* 25, no. 6 (2019): 954-961.
- [2]. Bi, Wenya Linda, Ahmed Hosny, Matthew B. Schabath, Maryellen L. Giger, Nicolai J. Birkbak, Alireza Mehrtaash, Tavis Allison et al. "Artificial intelligence in cancer imaging: clinical challenges and applications." *CA: a cancer journal for clinicians* 69, no. 2 (2019): 127-157.
- [3]. Bonavita, Ilaria, Xavier Rafael-Palou, Mario Ceresa, Gemma Piella, Vicent Ribas, and

Miguel A. González Ballester. "Integration of convolutional neural networks for pulmonary nodule malignancy assessment in a lung cancer classification pipeline." *Computer methods and programs in biomedicine* 185 (2020): 105172.

[4]. Cao, Haichao, Hong Liu, Enmin Song, Chih-Cheng Hung, Guangzhi Ma, Xiangyang Xu, Renchao Jin, and Jianguo Lu. "Dual-branch residual network for lung nodule segmentation." *Applied Soft Computing* 86 (2020): 105934.

Chae, Kum J., Gong Y. Jin, Seok B. Ko, Yi Wang, Hao Zhang, Eun J. Choi, and Hyemi Choi. "Deep learning for the classification of small ( $\leq 2$  cm) pulmonary nodules on CT imaging: a preliminary study." *Academic radiology* 27, no. 4 (2020): e55-e63.

[5]. Chen, Genlang, Jiajian Zhang, Deyun Zhuo, Yuning Pan, and Chaoyi Pang. "Identification of pulmonary nodules via CT images with hierarchical fully convolutional networks." *Medical & biological engineering & computing* 57, no. 7 (2019): 1567-1580.

[6]. Feng, Yuanli, Pengyi Hao, Peng Zhang, Xinguo Liu, Fuli Wu, and Hongwei Wang. "Supervoxel based weakly-supervised multi-level 3D CNNs for lung nodule detection and segmentation." *Journal of Ambient Intelligence and Humanized Computing* (2019): 1-11.

[7]. Hammami, Maryam, Denis Friboulet, and Razmig Kechichian. "Cycle GAN-Based Data Augmentation for Multi-Organ Detection in CT Images Via Yolo." In *2020 IEEE International Conference on Image Processing (ICIP)*, pp. 390-393. IEEE, 2020.

[8]. Hosny, Ahmed, Hugo J. Aerts, and Raymond H. Mak. "Handcrafted versus deep learning radiomics for prediction of cancer therapy response." *The Lancet Digital Health* 1, no. 3 (2019): e106-e107.

[9]. Huang, Xia, Wenqing Sun, Tzu-Liang Bill Tseng, Chunqiang Li, and Wei Qian. "Fast and fully-automated detection and segmentation of pulmonary nodules in thoracic CT scans using deep convolutional neural networks." *Computerized Medical Imaging and Graphics* 74 (2019): 25-36.

[10]. Hussein, Sarfaraz, Pujan Kandel, Candice W. Bolan, Michael B. Wallace, and Ulas Bagci. "Lung and pancreatic tumor characterization in the deep learning era: novel supervised and unsupervised learning approaches." *IEEE transactions on medical imaging* 38, no. 8 (2019): 1777-1787.

[11]. Jakimovski, Goran, and Danco Davcev. "Using double convolution neural network for lung cancer stage detection." *Applied Sciences* 9, no. 3 (2019): 427.

[12]. Kim, Bum-Chae, Jee Seok Yoon, Jun-Sik Choi, and Heung-Il Suk. "Multi-scale gradual integration CNN for false positive reduction in pulmonary nodule detection." *Neural Networks* 115 (2019): 1-10.

[13]. Lee, Sang Min, Joon Beom Seo, Jihye Yun, Young-Hoon Cho, Jens Vogel-Claussen, Mark L. Schiebler, Warren B. Gefter et al. "Deep learning applications in chest radiography and computed tomography." *Journal of thoracic imaging* 34, no. 2 (2019): 75-85.

[14]. Li, Shulong, Panpan Xu, Bin Li, Liyuan Chen, Zhiguo Zhou, Hongxia Hao, Yingying Duan et al. "Predicting lung nodule malignancies by combining deep convolutional neural network and handcrafted features." *Physics in Medicine & Biology* 64, no. 17 (2019): 175012.

[15]. Li, Yanfeng, Linlin Zhang, Houjin Chen, and Na Yang. "Lung nodule detection with deep learning in 3D thoracic MR images." *IEEE Access* 7 (2019): 37822-37832.

[16]. Liao, Fangzhou, Ming Liang, Zhe Li, Xiaolin Hu, and Sen Song. "Evaluate the malignancy of pulmonary nodules using the 3-d deep leaky noisy-or network." *IEEE transactions on neural networks and learning systems* 30, no. 11 (2019): 3484-3495.

[17]. Monkam, Patrice, Shouliang Qi, He Ma, Weiming Gao, Yudong Yao, and Wei Qian. "Detection and classification of pulmonary nodules using convolutional neural networks: a survey." *IEEE Access* 7 (2019): 78075-78091.

[18]. Nam, Ju Gang, Sunggyun Park, Eui Jin Hwang, Jong Hyuk Lee, Kwang-Nam Jin, Kun Young Lim, Thienkai Huy Vu et al. "Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs." *Radiology* 290, no. 1 (2019): 218-228.

[19]. Qi, Linlin, Wenwen Lu, Ning Wu, and Jianwei Wang. "Persistent pulmonary subsolid nodules with a solid component smaller than 6 mm: what do we know?." *Journal of Thoracic Disease* 12, no. 8 (2020): 4584.

[20]. Qin, Yulei, Hao Zheng, Xiaolin Huang, Jie Yang, and Yue Min Zhu. "Pulmonary nodule segmentation with CT sample synthesis using adversarial networks." *Medical physics* 46, no. 3 (2019): 1218-1229.

[21]. Qiu, Shi, Junjun Li, Mengdi Cong, Chun Wu, Yan Qin, and Ting Liang. "Detection of solitary pulmonary nodules

based on brain-computer interface." *Computational and Mathematical Methods in Medicine* 2020 (2020).

[22]. Rafael-Palou, Xavier, Anton Aubanell, IlariaBonavita, Mario Ceresa, Gemma Piella, VicentRibas, and Miguel A. González Ballester. "Re-Identification and growth detection of pulmonary nodules without image registration using 3D siamese neural networks." *Medical Image Analysis* 67 (2021): 101823.

[23]. Riquelme, Diego, and Moulay A. Akhloufi. "Deep learning for lung cancer nodules detection and classification in CT scans." *AI* 1, no. 1 (2020): 28-67.

[24]. Sim, Yongsik, Myung Jin Chung, ElmarKotter, SehyoYune, Myeongchan Kim, Synho Do, Kyunghwa Han et al. "Deep convolutional neural network-based software improves radiologist detection of malignant lung nodules on chest radiographs." *Radiology* 294, no. 1 (2020): 199-209.

[25]. Suzuki, Kazuhiro, Yujiro Otsuka, Yukihiro Nomura, Kanako K. Kumamaru, RyoheiKuwatsuru, and Shigeki Aoki. "Development and validation of a modified three-dimensional U-Net deep-learning model for automated detection of lung nodules on chest CT images from the lung image database consortium and japanese datasets." *Academic Radiology* (2020).

[26]. Tran, Giang Son, Thi Phuong Nghiem, Van Thi Nguyen, Chi Mai Luong, and Jean-Christophe Burie. "Improving accuracy of lung nodule classification using deep learning with focal loss." *Journal of healthcare engineering* 2019 (2019).

[27]. Venugopal, Vasantha Kumar, Kiran Vaidhya, MuraliMurugavel, AbhijithChunduru, Vidur Mahajan, Suthirth Vaidya, DigvijayMahra, AkshayRangasai, and Harsh Mahajan. "Unboxing AI-radiological insights into a deep neural network for lung nodule characterization." *Academic radiology* 27, no. 1 (2020): 88-95.

[28]. Wang, Chuang, Neelam Tyagi, Andreas Rimner, Yu-Chi Hu, Harini Veeraraghavan, Guang Li, Margie Hunt, Gig Mageras, and Pengpeng Zhang. "Segmenting lung tumors on longitudinal imaging studies via a patient-specific adaptive convolutional neural network." *Radiotherapy and Oncology* 131 (2019): 101-107.

[29]. Wang, Jun, Jiawei Wang, Yaofeng Wen, Hongbing Lu, TianyeNiu, Jiangfeng Pan, and DahongQian. "Pulmonary nodule detection in volumetric chest CT scans using CNNs-based nodule-size-adaptive

detection and classification." *IEEE Access* 7 (2019): 46033-46044.

[30]. Wang, Qin, FengyiShen, LinyaoShen, Jia Huang, and Weiguang Sheng. "Lung nodule detection in CT images using a raw patch-based convolutional neural network." *Journal of digital imaging* 32, no. 6 (2019): 971-979.

[31]. Wang, Shuo, Jingyun Shi, Zhaoxiang Ye, Di Dong, Dongdong Yu, Mu Zhou, Ying Liu et al. "Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning." *European Respiratory Journal* 53, no. 3 (2019).

[32]. Wang, Yang, Fangrong Yan, Xiaofan Lu, GuanmingZheng, Xin Zhang, Chen Wang, Kefeng Zhou et al. "IILS: Intelligent imaging layout system for automatic imaging report standardization and intra-interdisciplinary clinical workflow optimization." *EBioMedicine* 44 (2019): 162-181.

[33]. West, Elizabeth, SimukayiMutasa, Zelos Zhu, and Richard Ha. "Global trend in artificial intelligence-based publications in radiology from 2000 to 2018." *American Journal of Roentgenology* 213, no. 6 (2019): 1204-1206.

[34]. Winkels, Marysia, and Taco S. Cohen. "Pulmonary nodule detection in CT scans with equivariant CNNs." *Medical image analysis* 55 (2019): 15-26.

[35]. Xie, Yutong, Jianpeng Zhang, and Yong Xia. "Semi-supervised adversarial model for benign-malignant lung nodule classification on chest CT." *Medical image analysis* 57 (2019): 237-248.

[36]. Xie, Yutong, Yong Xia, Jianpeng Zhang, Yang Song, Dagan Feng, Michael Fulham, and Weidong Cai. "Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT." *IEEE transactions on medical imaging* 38, no. 4 (2018): 991-1004.

[37]. Xu, Xiuyuan, Chengdi Wang, JixiangGuo, Lan Yang, Hongli Bai, Weimin Li, and Zhang Yi. "DeepLN: a framework for automatic lung nodule detection using multi-resolution CT screening images." *Knowledge-Based Systems* 189 (2020): 105128.

[38]. Ye, Wenjing, Wen Gu, XuejunGuo, Ping Yi, YishuangMeng, Fengfeng Han, Lingwei Yu, Yi Chen, Guorui Zhang, and Xueting Wang. "Detection of pulmonary ground-glass opacity based on deep learning computer artificial intelligence." *Biomedical engineering online* 18, no. 1 (2019): 1-12.

- [39]. Zhang, Chao, Xing Sun, Kang Dang, Ke Li, Xiao□weiGuo, Jia Chang, Zong□qiao Yu et al. "Toward an expert level of lung cancer detection and classification using a deep convolutional neural network." *The oncologist* 24, no. 9 (2019): 1159.
- [40]. Chassagnon, G., De Margerie-Mellon, C., Vakalopoulou, M., Marini, R., Hoang-Thi, T. N., Revel, M. P., & Soyer, P. (2023). Artificial intelligence in lung cancer: current applications and perspectives. *Japanese Journal of Radiology*, 41(3), 235-244.
- [41]. de Margerie-Mellon, C., &Chassagnon, G. (2023). Artificial intelligence: A critical review of applications for lung nodule and lung cancer. *Diagnostic and Interventional Imaging*, 104(1), 11-17.
- [42]. Xie, Y., Meng, W. Y., Li, R. Z., Wang, Y. W., Qian, X., Chan, C., ... & Leung, E. L. H. (2021). Early lung cancer diagnostic biomarker discovery by machine learning methods. *Translational oncology*, 14(1), 100907.
- [43]. Amini, M., Hajianfar, G., Avval, A. H., Nazari, M., Deevband, M. R., Oveisi, M., ... & Zaidi, H. (2022). Overall survival prognostic modelling of non-small cell lung cancer patients using positron emission tomography/computed tomography harmonisedradiomics features: The quest for the optimal machine learning algorithm. *Clinical Oncology*, 34(2), 114-127.
- [44]. Alsinglawi, B., Alshari, O., Alorjani, M., Mubin, O., Alnajjar, F., Novoa, M., & Darwish, O. (2022). An explainable machine learning framework for lung cancer hospital length of stay prediction. *Scientific reports*, 12(1), 607.
- [45]. Hosseini, S. H., Monsefi, R., &Shadroo, S. (2023). Deep learning applications for lung cancer diagnosis: a systematic review. *Multimedia Tools and Applications*, 1-31.

## Appendix 1 Extracted Features

3.139714	0.855385	-1.05613	1.159385	-1.6138	1.098809	0.597195	-1.04734
2.285983	0.58798	-0.92013	3.921067	-1.99084	1.210445	0.796313	-1.0344
3.48238	0.62251	-0.76098	3.751422	-1.62754	1.151679	0.737283	-0.58236
0.833476	0.500855	-0.34023	2.865631	-1.31708	1.217392	0.65849	-1.06415
2.609431	0.623596	-0.28563	-7.5317	-1.42184	1.208	0.59147	-0.2332
0.34312	0.090419	1.350723	-6.44944	-1.3967	1.190822	0.457256	-0.4758
1.50261	0.639321	0.32311	-7.39789	-1.55333	1.105334	0.698994	-0.8181
2.515209	0.655297	-0.66711	3.984108	-2.33162	1.284508	-0.03203	-0.76365
2.256198	0.765446	0.696816	-1.73064	-1.50862	1.12775	-0.25593	-0.85256
3.232317	0.601936	0.14284	3.72272	-1.61802	1.237369	-0.17989	-1.04348
-0.04412	0.661404	1.301925	-5.00598	-1.47687	1.230103	0.738552	-0.43887
-0.01999	0.667424	-0.42987	3.382193	-1.68247	1.119017	0.702339	-1.14876
2.555013	0.663876	-0.76113	3.931909	-1.39343	1.098665	0.455336	-0.80107
-0.34944	0.68724	-0.24667	-3.3765	-1.41314	1.114203	0.24305	-0.95831
2.452954	0.478999	0.583645	-0.42869	-1.45553	1.085007	0.022141	-0.83036
3.159469	0.876636	1.113915	-1.1848	-1.44632	1.146394	0.735424	-0.84438
0.819852	0.666796	-0.25263	2.051772	-1.46832	1.223814	0.403054	-1.03644
1.611676	0.701253	-0.59836	-6.17646	-1.38073	1.188864	0.610608	0.361064
1.001159	1.526469	-0.24517	1.48732	-1.38109	1.155912	0.874868	-0.85141
2.945356	0.152182	-0.44663	-2.62161	-1.5727	1.236312	0.123226	-0.81307
3.234125	0.413892	-0.6445	-4.42557	-1.52236	1.144483	-0.18869	-2.0772
-0.38778	1.711747	-0.61879	-1.27173	-1.45935	1.266657	0.549209	0.638186
1.603368	0.035554	0.157741	-6.02537	-1.45329	1.408006	-0.02937	-1.4531
1.635461	0.449198	0.880944	-4.64788	-1.64095	1.213044	-0.19769	-1.29417
0.041334	0.04339	-0.30965	-5.98069	-1.84042	1.23806	0.20329	-1.73132
-----	-----	-----	-----	-----	-----	-----	-----