Volume 24, No 3 | International Journal of Radiation Research, July 2026

# Deep learning-based classification of lung nodules on CT images: A multicenter study of early lung cancer detection

**C. Qi[1] and Q. Jiang[2*]**

[1]*Baicheng Medical College, Baicheng, Jilin, 137000, China*
[2]*Baicheng Central Hospital, Baicheng, Jilin, 137000, China*

► ***Original article***

***Corresponding author:***
*Qing Jiang, Ph.D.,*
***E-mail:*** *77073199@bcmc.edu.cn*

**Received:** July 2025
**Final revised:** April 2025
**Accepted:** May 2025

***Int. J. Radiat. Res., July 2026; 24(3): 749-754***

**DOI:** 10.61186/ijrr.24.3.23

***Keywords:*** *Computed tomography, Deep learning, Lung cancer, Pulmonary nodule, Variational autoencoder.*

## ABSTRACT

***Background:*** Early lung cancer detection relies on accurate characterization of pulmonary nodules on computed tomography (CT). However, variability in CT acquisition and radiological interpretation may reduce reproducibility across institutions. This study evaluated a deep learning (DL) framework for benign-malignant lung nodule classification using multicenter CT datasets. ***Materials and Methods:*** Publicly available anonymized CT datasets from LUNA16 and LIDC-IDRI were used. Images were rescaled to a uniform resolution, denoised with Gaussian filtering, and normalized by Min-Max scaling. Nodule regions were segmented using a U-Net encoder-decoder network. Gray-Level Co-Occurrence Matrix (GLCM) texture features were extracted and combined with latent representations learned by an Effective Beetle Antennae Search-driven Variational Autoencoder (EBAS-VAE). Model performance was assessed using accuracy, precision, recall, F1-score, true-positive rate (TPR), false-positive rate (FPR), false discovery rate (FDR), and area under the receiver operating characteristic curve (AUC). ***Results:*** On LUNA16, EBAS-VAE achieved accuracy 0.9838, precision 0.94, recall 0.89, F1-score 0.96, and AUC 0.98. On LIDC-IDRI, accuracy was 0.97, TPR 0.97, FPR 0.04, and FDR 0.0001. EBAS-VAE outperformed 3D-CNN and SEOA-DBN comparators. ***Conclusion:*** The proposed EBAS-VAE pipeline improved lung nodule classification and reduced false-positive detection in multicenter CT images, supporting its potential as a radiologist-assistive screening tool.

## INTRODUCTION

Lung cancer remains one of the leading causes of cancer-related mortality worldwide, largely because many patients are diagnosed after the disease has progressed beyond an early, surgically curable stage [1]. Computed tomography (CT) is central to screening, detection, and characterization of pulmonary nodules, but image quality, slice thickness, nodule heterogeneity, and radiologist workload can affect diagnostic consistency [2-4]. Automated image analysis has therefore become an important adjunct to radiological interpretation, particularly for identifying small nodules, reducing false-positive alarms, and prioritizing clinically suspicious lesions [5-7].

Deep learning (DL), especially convolutional neural networks (CNNs), has shown strong performance in lung nodule detection and malignancy prediction because it can learn hierarchical image features from CT data without relying exclusively on hand-crafted descriptors [8-11]. Recent models have used CNNs, hybrid CNN-SVM systems, deep belief networks, attention modules, and multimodal fusion strategies for lung cancer detection [12-21]. Texture-based features such as Gray-Level Co-Occurrence Matrix (GLCM) descriptors remain useful because contrast, homogeneity, energy, dissimilarity, and correlation provide interpretable information about nodule heterogeneity, which is relevant to malignancy assessment [22-25].

Related work in International Journal of Radiation Research (IJRR) has also emphasized the importance of CT-based image analysis and quantitative prediction in lung oncology. Xu *et al.* demonstrated that low-dose CT combined with serum tumor markers improved early non-small cell lung cancer diagnosis [26]. Tanabe *et al.* used four-dimensional CT and neural network model fitting to improve prediction of lung tumor respiratory movement during stereotactic body radiotherapy [27]. Takano *et al.* showed that generative adversarial networks can generate CT images at different tube voltages, supporting the expanding role of artificial intelligence in radiological image processing [28]. These studies justify the clinical relevance of quantitative CT pipelines and support further evaluation of DL tools for lung cancer-related imaging tasks.

Despite these advances, several gaps remain. Many existing CNN-based models are sensitive to differences in scanner protocols and nodule morphology across institutions. False-positive detection remains a major limitation because benign nodules, vessels, and inflammatory lesions can

[ Downloaded from mail.ijrr.com on 2026-09-09 ]

[ DOI: 10.61186/ijrr.24.3.23 ]

resemble malignant lesions on CT. Some models are computationally intensive, poorly interpretable, or inadequately validated across heterogeneous datasets. These limitations restrict direct clinical translation, especially where the intended users are radiologists and clinicians rather than algorithm developers.

Novelty of the study: The present study proposes an Effective Beetle Antennae Search-driven Variational Autoencoder (EBAS-VAE) for lung nodule classification on multicenter CT data. The novelty lies in integrating U-Net segmentation, GLCM texture extraction, and EBAS-guided latent-space optimization within a VAE framework. This combination is intended to improve discriminative feature learning, reduce false-positive classifications, and enhance robustness across LUNA16 and LIDC-IDRI datasets while keeping the method clinically explainable through interpretable preprocessing, segmentation, feature extraction, and performance reporting steps.

## MATERIALS AND METHODS

### *Study design and data sources*

This retrospective computational study used only publicly available, anonymized imaging datasets. No new patient recruitment, drug administration, contrast injection, or laboratory testing was performed. The study followed a multicenter dataset approach by evaluating the model on LUNA16 and LIDC-IDRI CT image collections.

LUNA16 consists of 888 CT scans derived from the LIDC-IDRI database after exclusion of scans with slice thickness greater than 2.5 mm. LIDC-IDRI includes thoracic CT examinations with lesion annotations generated through a multireader radiologist annotation process. In this revised manuscript, the earlier incorrect description of LIDC-IDRI as chest X-ray data has been removed; the dataset is described as a CT-based resource.

### *Image preprocessing*

All CT images were rescaled to a uniform spatial resolution to minimize variation in image size before segmentation and classification. Gaussian filtering was applied to reduce high-frequency image noise while preserving clinically relevant lung parenchymal boundaries. Min-Max normalization scaled intensity values between 0 and 1 to stabilize model training across datasets (figure 1).

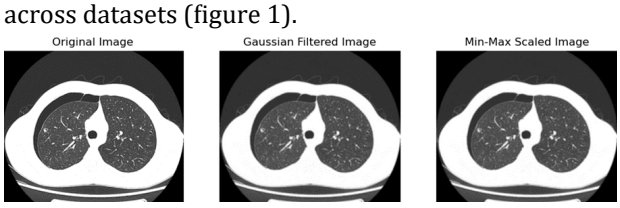


**Figure 1.** Representative preprocessing sequence showing original CT image, Gaussian-filtered image, and Min-Max scaled image.

### *Nodule segmentation*

Nodule segmentation was performed using a U-Net encoder-decoder architecture. The encoder pathway progressively reduced image dimensions while learning local and contextual features. The decoder pathway restored spatial resolution using up-sampling and skip connections, allowing preservation of nodule boundaries. Segmentation output was used to localize candidate nodule regions before feature extraction and classification.

The U-Net-based segmentation step was positioned before GLCM feature extraction in the revised workflow to reflect the intended image-processing sequence. This correction resolves the earlier interchange of segmentation and feature-extraction subsections.

### *Feature extraction*

Following segmentation, GLCM features were computed from candidate nodule regions. The retained texture descriptors were contrast, dissimilarity, homogeneity, energy, and correlation (figure 2). These features were selected because they summarize nodule texture and spatial gray-level organization in a clinically interpretable manner. Detailed mathematical derivations were removed and replaced with a concise description of the feature set.

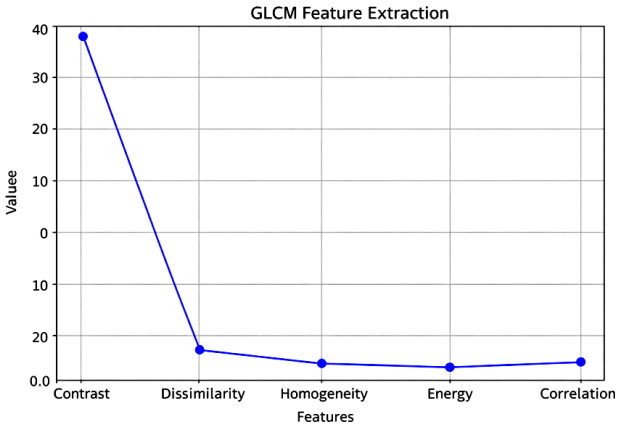


**Figure 2.** GLCM-based feature extraction profile showing contrast, dissimilarity, homogeneity, energy, and correlation values.

### *EBAS-VAE classification model*

The classification stage used an Effective Beetle Antennae Search-driven Variational Autoencoder (EBAS-VAE). The VAE learned latent representations from segmented nodule images and extracted feature vectors. The EBAS component optimized the latent representation to improve class separability between benign and malignant nodules. The final classifier generated the benign-malignant prediction for each candidate nodule.

The earlier Algorithm 1 section was deleted because it duplicated the narrative methods, contained incomplete variables, and was not informative for the intended medical readership. The algorithmic workflow is now summarized in

[ Downloaded from mail.ijrr.com on 2026-09-09 ]
[ DOI: 10.61186/ijrr.24.3.23 ]

procedural language.

### *Model training and validation*

Data were partitioned into training and testing subsets within each dataset. Preprocessing parameters were applied consistently to training and testing images. Model performance was evaluated separately for LUNA16 and LIDC-IDRI to assess whether the proposed framework retained performance across heterogeneous CT image sources.

Performance was assessed using accuracy, precision, recall, F1-score, TPR, FPR, FDR, and AUC. Definitions were described in text to avoid excessive equations. Accuracy represented the proportion of correctly classified cases. Precision represented the proportion of predicted malignant nodules that were true malignant cases. Recall/TPR represented the proportion of malignant nodules correctly identified. FPR represented benign cases incorrectly classified as malignant. FDR represented the proportion of false -positive results among positive predictions. AUC summarized ROC discrimination.

### *Hardware, software, and country of origin*

All appliances, software, and datasets used in this computational study are listed in table 1 with trade name/version and country of origin wherever applicable. No drugs, contrast agents, laboratory reagents, commercial diagnostic kits, or implantable devices were used because the study relied exclusively on anonymized public CT datasets.

**Table 1.** Hardware, software, datasets, and country of origin.

| Category | Trade name/version | Manufacturer/source | Country of origin | Purpose |
|---|---|---|---|---|
| Programming language | Python 3.10 | Python Software Foundation | United States | Model development and analysis |
| Numerical library | NumPy 1.24 | NumPy Developers / NumFOCUS-supported open-source project | United States | Array operations |
| Data library | Pandas 2.0 | Pandas Developers / NumFOCUS-supported open-source project | United States | Data handling |
| Image-processing library | OpenCV 4.8 | OpenCV.org / Intel-origin open-source library | United States | Image preprocessing |
| Deep learning framework | TensorFlow 2.13 with Keras API | Google LLC | United States | U-Net and VAE model implementation |
| Dataset platform | Kaggle Datasets | Google LLC | United States | Dataset access |
| Dataset | LUNA16 | Lung Nodule Analysis Challenge | Netherlands / international public challenge | Model training/testing |
| Dataset | LIDC-IDRI | National Cancer Institute-supported public database | United States | External dataset evaluation |
| Computer processor | Intel Core i3-4030U | Intel Corporation | United States | Model execution |
| Operating system | Microsoft Windows 8.1 | Microsoft Corporation | United States | Computing environment |

### *Ethical considerations*

This study used publicly available, fully anonymized CT datasets. No direct patient contact, intervention, identifiable health information, drugs, contrast media, or biological specimens were involved. Therefore, institutional ethical approval and informed consent were not required. Dataset use complied with the applicable public data-sharing policies.

## RESULTS

### *Experimental configuration*

The study was executed on a 64-bit computing platform with Intel Core i3 processor, 4 GB random-access memory, and 3 MB cache memory. The configuration was sufficient for proof-of-concept model training and evaluation, although higher-performance graphical processing hardware may be preferable for larger multicenter implementation. The experimental system configuration is summarized in table 2.

### *ROC performance*

The ROC curves demonstrated high discrimination for both datasets (figures 3 and 4)). The LUNA16 evaluation produced an AUC of 0.97 (figure 3), whereas LIDC-IDRI produced an AUC of 0.98 (figure 4). In both curves, the model line remained clearly above the random-classifier diagonal, indicating strong separation between benign and malignant nodules.

**Table 2.** Experimental system configuration.

| Component | Specification |
|---|---|
| Software | Python with NumPy, Pandas, OpenCV, TensorFlow/Keras |
| Operating system | Microsoft Windows 8.1 |
| CPU model | Intel Core i3-4030U |
| CPU architecture | 64-bit |
| RAM | 4 GB |
| Cache memory | 3 MB |

### *Feature extraction output*

GLCM feature extraction showed that contrast was the dominant texture descriptor, suggesting that local intensity variation contributed strongly to nodule discrimination (figure 5). Dissimilarity, homogeneity, energy, and correlation provided complementary descriptors of gray-level organization. These features supported the classifier by adding interpretable texture information to the latent representations generated by EBAS-VAE.

[ Downloaded from mail.ijrr.com on 2026-09-09 ]

[ DOI: 10.61186/ijrr.24.3.23 ]

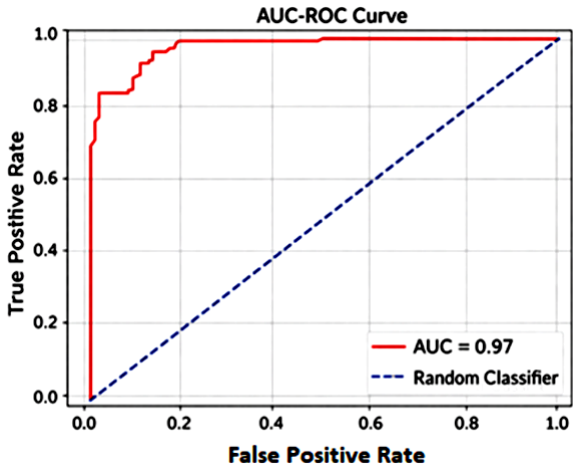


**Figure 3.** ROC curve for the LUNA16 dataset showing AUC = 0.97.

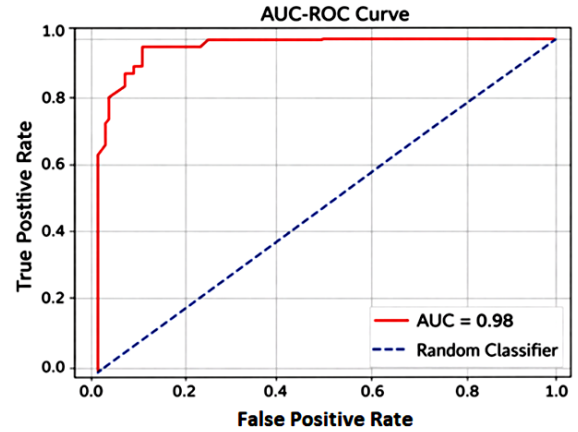


**Figure 4.** ROC curve for the LIDC-IDRI dataset showing AUC = 0.98.

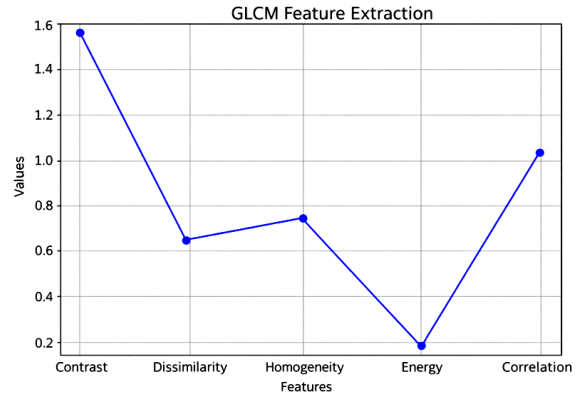


**Figure 5.** Dataset-wise feature extraction output for the additional CT dataset evaluation.

### *Comparative performance*

On LUNA16, EBAS-VAE outperformed the 3D-CNN comparator across all reported measures. The proposed model achieved AUC 0.98, accuracy 0.9838, precision 0.94, recall 0.89, and F1-score 0.96 (table 3). The gain was most evident in overall accuracy and F1-score, indicating improved balance between sensitivity and precision.

On LIDC-IDRI, EBAS-VAE achieved accuracy 0.97 and TPR 0.97, with lower FPR and FDR than the SEOA-DBN comparator (table 4). This indicates that the proposed approach improved malignant nodule detection while reducing false-positive predictions.

**Table 3.** Comparative performance on the LUNA16 dataset.

| Method | AUC | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|---|
| 3D-CNN [24] | 0.97 | 0.9717 | 0.92 | 0.87 | 0.94 |
| EBAS-VAE [Proposed] | 0.98 | 0.9838 | 0.94 | 0.89 | 0.96 |

**Table 4.** Comparative performance on the LIDC-IDRI dataset.

| Method | Accuracy | TPR | FPR | FDR |
|---|---|---|---|---|
| SEOA-DBN [25] | 0.96 | 0.96 | 0.05 | 0.0002 |
| EBAS-VAE [Proposed] | 0.97 | 0.97 | 0.04 | 0.0001 |

## DISCUSSION

The main observation was that EBAS-VAE produced high AUC and accuracy across two multicenter CT datasets. The ROC curves in figures 3 and 4 show that the classifier achieved strong discrimination, which is consistent with prior DL studies reporting improved classification when CT-based representation learning is combined with optimized feature extraction [11-21].

The preprocessing output in Figure 1 illustrates why standardized image preparation is necessary before model training. CT images collected from different sources may vary in matrix size, intensity distribution, noise level, and acquisition protocol. Rescaling, denoising, and normalization reduce this heterogeneity and help the network focus on nodule morphology rather than scanner-related variation. This is particularly important for multicenter work because models trained on a single imaging distribution may not generalize well to external datasets. The feature extraction output in figure 2 and figure 5 shows the importance of texture heterogeneity. Contrast was the highest GLCM descriptor, suggesting that local gray-level variation was prominent within segmented nodule regions. This supports the use of GLCM descriptors as complementary features, because malignant nodules often show heterogeneous internal architecture and irregular texture. While end-to-end CNN models can learn such features automatically, hand-crafted texture descriptors improve interpretability for medical readers and make the pipeline easier to audit.

Table 2 reports the system configuration and shows that the analysis was performed on modest computing resources. This is relevant for translational use in settings where advanced graphical processing units are not always available. Nevertheless, a larger prospective validation would require more powerful computing infrastructure, particularly if full-volume 3D CT input and real-time clinical processing are planned. Table 3 demonstrates that EBAS-VAE improved LUNA16 performance compared with 3D-CNN. The increase in accuracy from 0.9717 to 0.9838 and F1-score from 0.94 to 0.96 indicates that the proposed approach improved both correct classification and the balance between

[ Downloaded from mail.ijrr.com on 2026-09-09 ]
[ DOI: 10.61186/ijrr.24.3.23 ]

precision and recall. The improvement may be explained by the hybrid design: U-Net improves lesion localization, GLCM adds interpretable texture information, and EBAS-guided latent optimization improves feature separation in the VAE space. Table 4 demonstrates a similar advantage on LIDC-IDRI. EBAS-VAE increased accuracy and TPR while reducing FPR from 0.05 to 0.04 and FDR from 0.0002 to 0.0001 compared with SEOA-DBN. In clinical screening, even small reductions in false positives are important because false alarms may lead to additional imaging, invasive biopsy, patient anxiety, and increased healthcare cost. The current findings therefore suggest that the proposed model could assist radiologists by prioritizing suspicious nodules and reducing unnecessary follow-up for benign appearances.

The findings are consistent with recent CT-based and DL-based oncology studies. Xu *et al.* reported that low-dose CT combined with biomarkers improved early non-small cell lung cancer diagnosis, emphasizing the value of multimodal or enhanced CT -based decision support [26]. Tanabe *et al.* showed that neural network model fitting with four-dimensional CT could improve prediction of lung tumor movement during stereotactic body radiotherapy, demonstrating the usefulness of machine learning in lung-related radiation imaging workflows [27]. Takano *et al.* used a generative adversarial network to synthesize CT images at different tube voltages, further supporting the role of artificial intelligence in CT optimization and image translation [28]. Compared with these studies, the present work focuses specifically on lung nodule benign-malignant classification, but it shares the broader aim of improving quantitative CT interpretation.

The study has limitations. First, although LUNA16 and LIDC-IDRI are widely used public datasets, retrospective public data cannot replace prospective clinical validation. Second, patient-level clinical variables, smoking history, histopathology, and follow-up outcomes were not integrated into the model. Third, the proposed framework requires external testing on data from additional scanners and geographic regions. Fourth, explainability methods such as saliency maps or class activation maps were not included. Future work should include prospective validation, external hospital datasets, decision-curve analysis, explainability outputs, and comparison with radiologist performance.

## CONCLUSION

The EBAS-VAE framework improved lung nodule classification across multicenter CT datasets by combining standardized preprocessing, U-Net segmentation, GLCM texture extraction, and latent-space optimization. The model achieved high AUC and accuracy with lower false-positive detection than comparator models. These findings support EBAS-VAE as a promising radiologist-assistive approach for early lung cancer detection, although prospective external validation is required before clinical deployment.

***Acknowledgments*:** The authors thank the dataset contributors of LUNA16 and LIDC-IDRI for making anonymized CT images available for research use. Research on teaching reform of vocational education and adult education in Jilin Province; Research and practice of medical imaging technology specialty construction under the background of "Double-height (NO: 2022ZCZ013).

***Conflict of interest*:** The authors declare no conflicts of interest.

***Funding*:** This research received no external funding.

***Ethical considerations*:** All data were anonymized and publicly available; hence institutional ethical approval and informed consent were not required.

**Author contribution:** All authors were involved equally in all aspects of research and preparation of the manuscript. All authors read and approved final version of the manuscript for publication.

***AI assistance declaration*:** Used for language refining.

## REFERENCES

1. Wahab Sait AR (2023) Lung cancer detection model using deep learning technique. *Applied Sciences,* ***13****(22): 12510.*
2. Sajja TK, Devarapalli RM, Kalluri HK (2019) Lung cancer detection based on CT scan images by using deep transfer learning. *Traitement du Signal,* ***36****(4): 339-344.*
3. Gao S, Zhang G, Lian Y, Yan L, Gao H (2020) Exploration and analysis of the value of tumor-marker joint detection in the pathological type of lung cancer. *Cellular and Molecular Biology,* ***66****(6): 93-97.*
4. Kırca K and Kutlutürkan S (2021) The effect of progressive relaxation exercises on treatment-related symptoms and self-efficacy in patients with lung cancer receiving chemotherapy. *Complement Ther Clin Pract,* ***45****: 101488.*
5. Asuntha A and Srinivasan A (2020) Deep learning for lung cancer detection and classification. *Multimed Tools Appl,* ***79****(11): 7731-7762.*
6. Kummer S, Waller J, Ruparel M, Cass J, *et al.* (2020) Mapping the spectrum of psychological and behavioral responses to low-dose CT lung cancer screening offered within a lung health check. *Health Expect,* ***23****(2): 433-441.*
7. Nazir I, Haq IU, Khan MM, Qureshi MB, (2021) Efficient pre-processing and segmentation for lung cancer detection using fused CT images. *Electronics,* ***11****(1): 34.*
8. Yang Y, Yang J, Liang Y, Liao B, Zhu W, Mo X, *et al.* (2021) Identification and validation of the efficacy of immunological therapy for lung cancer from histopathological images based on deep learning. *Front Genet,* ***12****: 642981.*
9. Lu D, Shangguan Z, Su Z, Lin C, Huang Z, Xie H (2024) Artificial intelligence-based plasma exosome label-free SERS profiling strategy for early lung cancer detection. *Anal Bioanal Chem,* ***416*** *(23): 5089-5096.*
10. Chiet CC, How KW, Han PY, Yen YH (2024) A lung cancer detection with pre-trained CNN models. *J Informatics and Web Engineering,* ***3****(1): 41-54.*
11. Obayya M, Arasi MA, Alruwais N, Alsini R, *et al.* (2023) Biomedical image analysis for colon and lung cancer detection using tuna swarm algorithm with deep learning model. *IEEE Access,* ***11****: 94705 -94712.*
12. Wankhade S and Vigneshwari S (2023) A novel hybrid deep learning method for early detection of lung cancer using neural

[ Downloaded from mail.ijrr.com on 2026-09-09 ]

[ DOI: 10.61186/ijrr.24.3.23 ]

networks. *Healthcare Analytics, **3**: 100195*.
13. Sreeprada V and Vedavathi K (2023) Lung cancer detection from X -ray images using hybrid deep learning technique. *Procedia Comput Sci, **230**: 467-474*.
14. Alsheikhy AA, Said Y, Shawly T, Alzahrani AK, Lahza H (2023) A CAD system for lung cancer detection using hybrid deep learning techniques. *Diagnostics, **13**(6): 1174*.
15. Navaneethakrishnan M, Anand MV, Vasavi G, Rani VV (2023) Deep Fuzzy SegNet-based lung nodule segmentation and optimized deep learning for lung cancer detection. *Pattern Anal Appl, **26**(3): 1143-1159*.
16. Crasta LJ, Neema R, Pais AR (2024) A novel deep learning architecture for lung cancer detection and diagnosis from computed tomography image analysis. *Healthcare Analytics, **5**: 100316*.
17. Wang X, Sharpnack J, Lee T (2024) Improving lung cancer diagnosis and survival prediction with deep learning and CT imaging. *arXiv, **2024**: 2408.09367*.
18. Thanammal KK (2024) Lung cancer detection via deep learning-based pyramid network with honey badger algorithm. Measurement: *Sensors, **31**: 100993*.
19. Sangeetha SKB, Mathivanan SK, Karthikeyan P, Rajadurai H, S, *et al.* 2024An enhanced multimodal fusion deep learning neural network for lung cancer classification. *Systems and Soft Computing, **6**: 200068*.
20. Subash J and Kalaivani S (2024) Dual-stage classification for lung cancer detection and staging using hybrid deep learning techniques. *Neural Comput Appl, **36**(14): 8141-8161*.
21. Dawood H, Nawaz M, Ilyas MU, Nazir T, Javed A (2025) Attention-guided CenterNet deep learning approach for lung cancer detection. *Comput Biol Med, **186**: 109613*.
22. Diao S, Wan Y, Huang D, Huang S, *et al.* (2025) Optimizing Bi-LSTM networks for improved lung cancer detection accuracy. *PLoS One, **20**(2): e0316136*.
23. Ozdemir B, Aslan E, Pacal I (2025) Attention enhanced InceptionNeXt based hybrid deep learning model for lung cancer detection. *IEEE Access, **2025***.
24. Neal Joshua ES, Bhattacharyya D, Chakkravarthy M, Byun YC (2021) 3D CNN with visual insights for early detection of lung cancer using gradient-weighted class activation. *J Healthc Eng, **2021**: 6695518*.
25. Shanid M and Anitha A (2020) Lung cancer detection from CT images using salp-elephant optimization-based deep learning. *Biomed Eng Appl Basis Commun, **32**(01): 2050001*.
26. Xu R, Xu W, Liu D, Pan S, *et al.* (2025) Application value of low-dose computed tomography combined with serum tumor markers in diagnosis of early non-small cell lung cancer. *Int J Radiat Res, **23**(1): 77-81*.
27. Tanabe Y, Deguchi T, Kiritani M, Hira N, *et al.* (2024) Evaluation of statistical prediction model fitting by combining fiducial markers and lung volume for stereotactic body radiotherapy. *Int J Radiat Res, **22**(1): 17-22*.
28. Takano H, Kondo Y, Ishizawa Y, Onodera S, *et al.* (2025) Generation of computed tomography images with different tube voltages by generative adversarial network. *Int J Radiat Res, **23**(4): 873-879*.

[ Downloaded from mail.ijrr.com on 2026-09-09 ]

[ DOI: 10.61186/ijrr.24.3.23 ]