Computers in Biology and Medicine 186 (2025) 109613

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Attention-guided CenterNet deep learning approach for lung cancer detection

Hussain Dawood [a], Marriam Nawaz [b], Muhammad U. Ilyas [c], Tahira Nazir [d], Ali Javed [b,*]

[a] School of Computing, Skyline University College, Sharjah, United Arab Emirates
[b] Department of Software Engineering, University of Engineering and Technology-Taxila, 47050, Punjab, Pakistan
[c] School of Computer Science, University of Birmingham, Dubai, United Arab Emirates
[d] Department of Software Engineering and Computer Science, Riphah International University, Gulberg Green Campus Islamabad, Pakistan

ARTICLE INFO

Keywords:
Attention mechanism
Classification
CenterNet
Deep learning
Lung cancer
ResNet

ABSTRACT

Lung cancer remains a significant health concern worldwide, prompting ongoing research efforts to enhance early detection and diagnosis. Prior studies have identified key challenges in existing approaches, including limitations in feature extraction, interpretability, and computational efficiency. In response, this study introduces a novel deep learning (DL) framework, termed the Improved CenterNet approach, tailored specifically for lung cancer detection. The primary importance of this work lies in its innovative integration of ResNet-34 with an attention mechanism within the CenterNet architecture, addressing critical limitations identified in previous studies. By augmenting the base network with an attention mechanism, our framework offers improved feature extraction capabilities, enabling the model to learn relevant patterns associated with lung cancer amidst complex backgrounds and varying environmental conditions. This enhancement facilitates more accurate and interpretable predictions while reducing computational complexity and inference times. Through extensive experimental evaluations conducted on standard datasets, our proposed approach demonstrates promising results, highlighting its potential to advance the field of lung cancer detection and diagnosis. Specifically, we have acquired the precision, recall, and F1-Score of 99.89 %, 99.82 %, and 99.85 % on the LUNA-16 dataset, and 98.33 %, 98.02 %, and 98.17 % for the Kaggle data sample, respectively which is showing the efficacy of our approach. One limitation of the work is that it cannot effectively locate the samples with intense light variations. Therefore, future research work is focused on overcoming this challenge.

1. Introduction

Lung cancer is a severe category of cancer that is initiated by an uncontrolled progression of cells in the lung tissues [1]. Lung cancer is broadly divided into 2 groups: non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). NSCLC is a more prevalent type and typically exhibits slower growth and spread compared to SCLC [2]. Lung cancer ranks among the most widespread and lethal malignancies globally. According to global cancer statistics from the World Health Organization (WHO) and the International Agency for Research on Cancer (IARC), lung cancer consistently ranks as one of the leading causes of cancer-related deaths. In 2021, lung cancer ranked second globally in new cancer cases, with 2.2 million diagnoses following breast cancer, and of the 9.96 million cancer deaths, lung cancer was the major reason, accounting for 1.8 million fatalities [3]. While in 2022, about 2.5 million people suffered from this deadly disease, with 1.8 million demises [4]. These statistics emphasize the substantial impact of lung cancer on both incidence and mortality, underscoring the urgent need for effective prevention and treatment strategies. Motivated by the high mortality rate and the critical challenges in diagnosis, this research focuses on developing an innovative approach to lung cancer classification. By leveraging the cutting-edge DL model, this study aims to contribute to more reliable diagnostic solutions and improved patient outcomes. Lung nodules, small abnormal growths detected in lung imaging, often raise concerns due to their potential association with lung cancer [5]. These nodules, visible on chest X-rays or computed tomography (CT) scans, can be either benign or malignant. While benign nodules may result from various non-cancerous conditions, malignant nodules could be indicative of lung cancer, emphasizing the importance of thorough evaluation and monitoring. Prompt detection and

* Corresponding author.
E-mail address: ali.javed@uettaxila.edu.pk (A. Javed).

https://doi.org/10.1016/j.compbiomed.2024.109613
Received 1 August 2024; Received in revised form 13 December 2024; Accepted 21 December 2024
Available online 2 January 2025

characterization of lung nodules play a vital role in the overall management of lung cancer, allowing for timely intervention and improved treatment outcomes. Regular screening and follow-up assessments are essential components in the comprehensive approach to lung cancer prevention and care [6].

Various imaging scans are employed for lung cancer detection, with CT scans being particularly valuable. CT scans offer comprehensive cross-sectional views of the chest, assisting in the identification and characterization of lung cancer or other irregularities [7]. The high-resolution images produced by CT scans aid in distinguishing between benign and malignant lesions, enabling healthcare professionals to determine the nature and extent of lung abnormalities. CT scans are useful for lung cancer detection due to their ability to capture detailed three-dimensional images quickly, making them essential for early diagnosis and staging [8]. Additionally, CT scans are crucial in guiding further diagnostic procedures, such as biopsies, and are instrumental in treatment planning and monitoring the progress of lung cancer patients. The versatility and precision of CT scans contribute significantly to the detection and management of lung cancer [9]. While the manual interpretation of imaging scans, including CT scans for lung tumor recognition, has been a cornerstone in medical diagnostics, it comes with inherent limitations. The process of reviewing and analyzing numerous images manually can be inefficient, labor-intensive, and susceptible to human error. Radiologists may face challenges in consistently identifying subtle abnormalities or variations in images, and the subjective nature of manual interpretation introduces the possibility of inter-observer variability. Moreover, as the volume of medical imaging data increases with advancements in technology, the burden on healthcare professionals intensifies [10].

Existing research employs image scans of different modalities for automated lung cancer detection. AI and machine learning (ML) techniques have appeared as powerful tools in processing image-based data, and researchers increasingly leverage them for medical image analysis [11]. These technologies outperform in recalling complicated patterns and irregularities within medical samples, increasing diagnostic accuracy and efficiency. For instance, AI algorithms applied to CT scans can swiftly and systematically analyze large datasets, aiding radiologists in identifying subtle abnormalities that exist in tumors or nodules for lung cancer detection. The use of these advanced techniques not only expedites the diagnostic process but also contributes to the ongoing efforts to improve early detection and treatment outcomes in medical imaging, showcasing the transformative potential of AI in healthcare [12]. Automation also facilitates consistency in interpretation across different healthcare settings, reducing the impact of inter-observer variability. Implementing automated systems not only streamlines the diagnostic procedure but also assists medical specialists in making more informed decisions for timely intervention and improved patient outcomes in lung cancer recognition and management [13].

Initially, traditional ML techniques were employed in lung cancer detection, leveraging algorithms to analyze features extracted from medical imaging data [14]. Researchers used these techniques to identify patterns and correlations indicative of lung abnormalities in images such as CT scans. Features like shape, texture, and intensity of lesions were considered, and classifiers were trained to differentiate between benign and malignant conditions. While ML showed promise, it faced limitations in handling the complexity of medical imaging data and capturing complex patterns effectively [15]. The limitations of traditional ML in lung cancer detection included the need for manual feature engineering, which relied heavily on domain expertise and missed subtle and critical information. Additionally, traditional ML struggled with the hierarchical and nonlinear nature of complex medical image data. As a result, these methods were less adept at automatically learning sophisticated representations from raw data, potentially limiting their ability to achieve high accuracy in tasks like early lung cancer detection. The introduction of deep learning marked a transformative shift in lung cancer detection [16]. Deep learning (DL) techniques, particularly convolutional neural networks (CNNs), demonstrated extraordinary accomplishment in the automated computation of hierarchical features directly from raw images [17]. This eliminated the need for manual feature engineering and allowed the model to discern complex patterns independently [18–20]. In lung cancer detection, DL frameworks have shown superior performance in identifying subtle abnormalities, such as nodules or tumors, in medical images. The end-to-end learning approach in DL facilitated the development of models capable of handling the inherent complexity of medical image data, contributing significantly to the precise diagnosis of lung cancer [21–23].

Despite notable successes, DL approaches employed for lung cancer detection face challenges and limitations. One challenge is the potential for false positives and false negatives, where the model may misclassify non-cancerous structures or fail to detect actual tumors. Interpretability remains a significant issue, as DL models, especially complex NNs, often function as black boxes, hindering a clear understanding of how decisions are reached. Another limitation is the requirement for substantial computational resources, making the deployment of DL models challenging in real-time clinical settings. Moreover, DL models may struggle with handling variations in imaging quality, protocols, and devices, impacting their generalizability across diverse healthcare settings. The major objective of our work is to overcome these challenges to enhance model interpretability, optimize computational efficiency, and improve robustness to diverse clinical conditions for the widespread and effective implementation of DL in lung cancer detection. For this, an effort has been made to overwhelm the issues of existing works by proposing a novel DL framework called an improved CenterNet approach harmonizing ResNet-34 with attention guidance. Descriptively, an enhanced keypoints computation module ResNet-34, augmented with an attention mechanism, is proposed as the backbone network of the CenterNet model to recognize effective characteristics of lung cancer from input samples. The adapted feature extractor network is designed to enhance the model's capacity to capture essential keypoints amidst complex backgrounds and varying environmental conditions. Our approach facilitates concurrent identification and classification through an end-to-end training strategy, showcasing robust results in a real-world environment. The following are the main contributions of our work.

- The incorporation of an improved keypoints extraction module, coupled with an attention mechanism, represents a significant contribution. This enhancement allows the model to discern fine-grained characteristics associated with lung cancer, making it more adept at capturing subtle patterns even in complex backgrounds and varying environmental settings.
- The use of a single-stage object detection approach enhances efficiency by reducing the computational complexity associated with multi-stage methods. This leads to faster inference times, making the model more practical for real-time applications in clinical settings.
- Localization of the infected regions allows healthcare professionals to locate the region of interest within lung images that contribute to the model's decision-making process. This transparency aids in building trust and confidence in the model's predictions, addressing one of the key challenges in deep learning interpretability.
- The adoption of a single-stage object detection approach with the enhanced feature extraction module contributes to improved accuracy by minimizing false positives and false negatives. This is crucial in lung cancer detection, where misclassifications can have significant consequences for patient outcomes.
- Extensive experimental testing of the proposed approach is accomplished by employing a standard dataset indicating the effectiveness of our approach.

The remaining paper follows the given hierarchy: the existing works related to lung cancer detection are discussed in Section 2, while the proposed model details are provided in Section 3. The employed dataset, metrics along with an in-depth discussion of attained results are

presented in Section 4. Finally, the conclusion along with future work is mentioned in Section 5.

2. Related work

In this section, a thorough review of historical studies attempted to perform the classification of lung cancer has been investigated.

Some researchers have utilized the conventional ML approach for lung cancer recognition. Venkatesan et al. [24] suggested a model for effectively locating lung cancer. This study introduced a hybrid model to locate lung cancer from the CT samples, integrating a Discrete Local Binary Pattern (DLBP) and a Hybrid Wavelet Partial Hadamard Transform (Hybrid WPHT). The model performed preprocessing with adaptive median filtering, followed by feature extraction using DLBP and Hybrid WPHT. To optimize the feature selection process, an adaptive Harris-Hawk optimization (AHHO) approach was employed, effectively reducing feature dimensionality. The classification stage utilized the optimal SVM (OSVM) along with the improved weight-based beetle swarm (IW-BS) approach for parameter setting. The work [24] performs well for lung cancer lesion recognition, however, the model needs evaluation on a more complex and large data sample. Another framework was proposed in Ref. [25] where initially, a geometric mean filter was applied to advance the appearance of samples. Next, the K-mean technique approach was used to identify the areas of interest. In the next phase, the LDA algorithm was used to compute the sample features. Finally, various classifiers were used to perform the classification task. This work has reported the highest classification results with the ANN classifier, however, results are reported for a small dataset. The ML-based approaches show better results for this problem, however, lack to tackle complex sample transformations.

The advancements in DL approaches have inspired researchers to apply them in medical image analysis, like lung cancer recognition. Wani et al. [1] applied a DL framework called DeepXplainer for the recognition of lung cancer. The work utilized an interpretable hybrid methodology that applied both CNN and XGBoost techniques for locating lung cancer and offering transparent details for the predictions. In this methodology, CNN automatically learns intricate features from the input data through its numerous convolutional layers. Subsequently, XGBoost was utilized for determining output labels based on the acquired features. To enhance the interpretability of predictions, the work integrated an explainable AI method called "SHAP" (SHapley Additive exPlanations). The approach [1] enhances lung cancer classification results, however, the model suffers from a high computing cost. Swain et al. [26] employed both complex structures, such as the dense NNs like VGG-16 and ResNet-50, and less complicated structures, represented by the sparse NN Inception v3. DL models were used to learn keypoints from CT images, facilitating the accurate classification of non-small cell lung cancer. The evaluation involves 60 adenocarcinoma and squamous cell carcinoma victims, respectively. The Inception v3 network demonstrates impressive performance in comparison to the other two deep nets, however, the work requires performance evaluation on a challenging and standard dataset. Naseer et al. [27] proposed an approach for performing lung cancer classification. Initially, an improved ALexNet approach was applied to the input samples to extract the related set of sample information. In the next phase, the deep features were propagated to the SVM predictor to execute the classification task. The approach has reported a classification accuracy of 97.64 %, however, the approach lacks generalization ability. Pandian et al. [28] suggested a CNN approach in which the VGG-16 model along with the GoogleNet framework in an end-to-end manner, was used to compute a dense set of sample characteristics and accomplish the classification task. The approach has reported a classification score of 98 %, however, it is not effective in tackling the transformation alterations of samples. In Ref. [29], a CNN approach was proposed to classify lung cancer from the CT-Scan samples. The approach was evaluated on the Kaggle dataset and reported an accuracy value of 96.43 %. This model presents a lightweight approach for lung cancer classification, however, performance needs further improvements.

Many researchers employed the concept of using the segmentation approaches to locate the area of interest like in Ref. [30], a DL-oriented lung nodule segmentation approach was proposed termed Wavelet U-Net++. This approach integrated the U-Net++ framework with wavelet pooling to commendably compute high- and low-level details of the suspected sample, resulting in superior segmentation performance. Specifically, the Haar wavelet transform was applied to minimize the size of the feature keypoints in the encoder unit, facilitating the capture of intricate details in the image, which was later segmented in the decoder phase. The approach shows improved lung cancer nodule segmentation results, however, results need further enhancements. Siddiqui et al. [31] proposed a DL approach for recognizing lung cancer from the CT-Scan images using a DL approach. Descriptively, the model employed the Gabor filters in combination with an improved Deep Belief Network (E-DBN) to extract the dense features from the input samples. While performing the classification step, the SVM approach was adopted. The model performs well in recognizing the healthy and cancer samples of lung cancer images; however, performance results are affected by huge variations in the nodule size. Tyagi et al. [32] also presented a DL approach to classifying the lung cancer stages. First, an augmentation step was applied to increase the versatility of the input samples. After this, a Deep CNN approach was designed to extract the visual characteristics of the samples and accomplish the classification task. The work shows effective results in accomplishing lung cancer recognition, however, with a high processing demand.

Various works explored the employment of hybrid DL models for lung cancer detection. Gudur et al. [33] proposed a hybrid approach for recognizing lung cancer from X-ray images. First, a preprocessing stage was applied to advance the graphic appearance of the samples. After this, the MobileNet-v2 model was utilized in combination with the genetic algorithm (GA). The approach [33] shows improved lung cancer detection results, with an increased computing cost. Devi et al. [34] proposed a DL framework called the scalable attention mechanism (SAM) to segment lung cancer from X-ray images and attained the highest IoU of 0.90, however, the approach suffers from a high computing burden. Another such DL approach was presented in Ref. [35] where the contrast level of the input CT scan image was elevated through histogram equalization (HE). Next, the adaptive bilateral filter (ABF) was employed to enhance the CT scan images, effectively reducing noise. Following the pre-processing steps, an Ensemble Deep Convolutional Neural Network (EDNN) was utilized, leveraging the Modified Mayfly Optimization and Modified Particle Swarm Optimization (M2PSO) algorithm for the segmentation of lung cancer within the pre-processed CT images. The model performs well for noisy samples, however, it may not generalize well to real-world cases. Mothkura et al. [36] suggested a model to recognize lung cancer from the CT-Scan samples via utilizing various DL frameworks. First, a preprocessing phase was utilized to enhance the visual representation of samples. After this, three DL frameworks named vanilla 2D CNN, SqueezeNet, and MobileNet were utilized in an end-to-end training manner. The model attained the highest accuracy with the CNN model, with a value of 89.21 %, however, classification results need improvements. Lin et al. [37] proposed a DL approach called the fusion-based convolutional fuzzy neural network (F-CFNN) for performing lung cancer classification. The approach employed two convolutions and two pooling layers for extracting the relevant information from input images, which were later classified into various categories of lung cancer. The approach is effective in accomplishing the lung cancer classification task, however, it is unable to tackle the image transformational changes. In Ref. [38], the author has employed a 2DCNN approach using the idea of ensemble learning for computing a set of dense features from the input samples and performing the lung cancer classification task. The approach has reported a classification score of 95 %, however, with a computing burden.

Table 1
Analysis of the existing works.

Reference	Year	Technique	Dataset	Result	Limitation
[1]	2024	DeepXplainer a hybrid approach employing CNN with the XGBoost approach	Survey Lung Cancer	Accuracy = 97.43 %	Computationally inefficient
[24]	2024	DLBP, Hybrid WPHT along with OSVM	LUNA-16-I LUNA-16-II	Accuracy = 99.35 % (LUNA-16-I) Accuracy = 99.55 % (LUNA-16-I)	The model needs testing on a large image dataset
[26]	2024	Inception v3	Custom dataset	Accuracy = 98.92 %	The approach requires testing on a standard dataset
[30]	2024	U-Net++ framework with wavelet	LIDC-IDRI	mean IoU = 0.878	The model performance needs improvements
[33]	2024	MobileNet-v2 model along with GA	Custom dataset	AUC = 96.66 %	The model is computationally inefficient
[34]	2024	SAM	Kaggle X-ray dataset	IOU = 0.90	The approach suffers from a high computing burden
[35]	2024	HE along with the EDNN model	SHFSU	dice score = 9520 %	The approach may not generalize well to real-world cases
[31]	2023	Gabor filters along with the E-DBN	LIDC-IDRI LUNA-16	Accuracy = 98.97 % (LIDC-IDRI) Accuracy = 96.89 % (LUNA-16)	performance results are affected by huge variations in the nodule size.
[32]	2023	Deep CNN	Lung-PET-CT-Dx	Accuracy = 97 %	The model is computationally inefficient
[36]	2023	CNN	LIDC-IDRI	Accuracy = 89.21 %	The classification values need enhancement
[39]	2023	Swin Transformer	LUNA-16	Accuracy = 82.26 %	The classification score requires enhancement
[40]	2023	UNET++	Montgomery County X-ray set	Accuracy = 98 %	The model is computationally inefficient
[41]	2023	autoencoder	CGAD	Accuray = 99.71 %	The approach requires the evaluation on a larger data sample
[37]	2023	FCNN	SPIE-AAPM	Accuracy = 93.26 %	The model is unable to tackle the image transformations
[38]	2023	2DCNN	LUNA-16	Accuracy = 95 %	The approach is computationally inefficient
[27]	2023	AlexNet along with SVM	LUNA-16	Accuracy = 97.64 %	The model lacks generalization power
[28]	2022	GoogleNet along with vgg-16	Custom dataset	Accuracy = 98 %	The model is unable to tackle the image transformations
[25]	2022	K-mean, LDA along with the ANN classifier	Custom dataset	Sensitivity = 98 %	Results are reported for a small dataset.
[29]	2022	CNN	Kaggle dataset	Accuracy = 96.43 %	Performance needs further improvements

Now, various latest DL approaches like transformers-based techniques are being heavily explored by researchers. In Ref. [39], the authors employed the idea of using the swim-transformers for classifying the lung cancer samples. For this purpose, the work employed the pre-trained Swin-B approach and reported an accuracy rate of 82.26 %. The approach performs well for recognizing lung cancer images from healthy samples of CT-Scan images, however, classification performance requires improvements. Another work was discussed in Ref. [40] where the UNET++ was employed to segment various lung diseases. The approach reported a segmentation accuracy of 98 %, however, with a high computing burden. Pandit et al. [41] proposed an approach for classifying lung cancers by using a DL framework. First, the samples were preprocessed to minimize the sample resolution, and an autoencoder model was utilized to learn the related group of characteristics and accomplish the classification task. This approach performs well in performing lung cancer classification, however, the work should be evaluated on a big dataset. A detailed analysis of the work from the past has been shown in Table 1. The conducted analysis shows that a vast amount of work has been presented for the recognition of lung cancers from samples, however, still a gap exists for performance improvement.

3. Proposed method

In this work, an automated framework has been designed for the early recognition of lung cancer samples from the CT-Scan images by

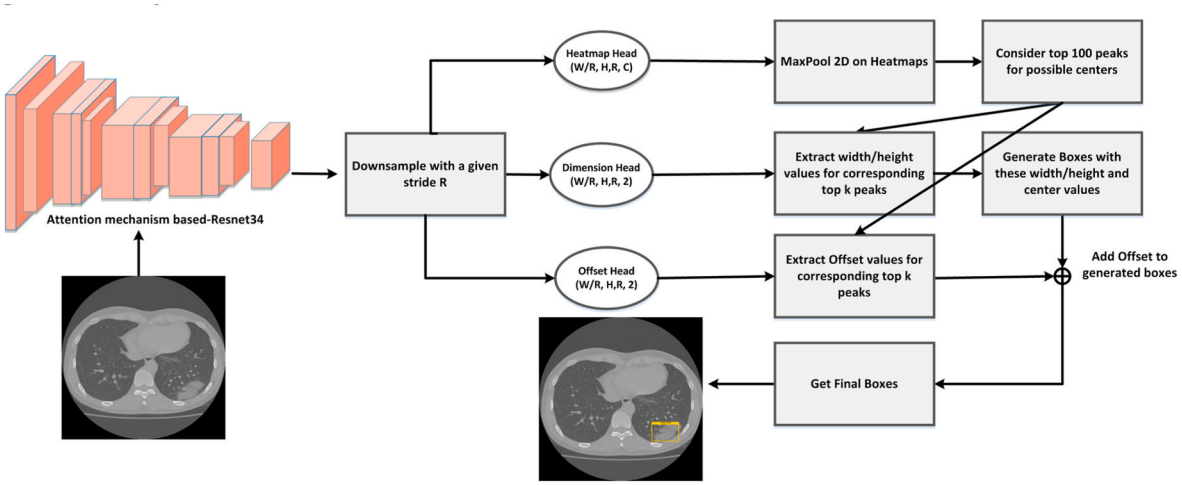

Fig. 1. Workflow representation of the suggested framework for lung cancer recognition.

proposing an improved DL approach. Descriptively, an enhanced keypoints computation module ResNet-34, augmented with an attention mechanism, is proposed as the base network of the CenterNet approach to recognize effective characteristics of lung cancer from input samples. This automated framework follows two steps to accomplish the designated task, which are recognized as the transfer learning and recognition steps, respectively. Fig. 1 displays the entire pipeline of the proposed work. Firstly, data sample preparation is performed in which sample annotations are created to determine the position of the diseased samples, which are later employed for model training. Next, we applied CenterNet approach with an attention mechanism (AM)- based ResNet-34 module as a feature extractor. The improved feature computation module accepts two inputs (processed image along with its annotation). The ResNet-34 computes a dense set of sample characteristics on which the AM strategy assists in focusing on the more relevant set of features. The computed features are passed to the recognition module of the CenterNet approach to localize and classify the diseased areas. In the last, model performance is computed by using the standard evaluation measures utilized in computer vision. A thorough elaboration of the introduced approach is provided in Algorithm 1.

Algorithm 1: Steps taken by the presented strategy for lung cancer recognition.

INPUT:
TS, AN
OUTPUT:
Localized Regions, ModifiedCenM, Classified samples
TS- samples used for model training.
AN – Position of diseased areas in the lung samples.
Localized Regions – Recognized lung cancer areas in investigated images.
ModifiedCenM –ResNet-34 with AM-based CenterNet model.
Classified samples – Output label for each identified diseased region.
Sample_Size ← [p q]
RoI Computation
***ê* ← RoIComputation(TS, AN)**
Modified Model
***ModifiedCenM* ← ResNet-34_With_AM_based_CenterNet (*Sample_Size, ê*)**
[T_r T_t] ← Distribution of data sample into train and test sets
Training Part
For sample *s* from T_r
Compute *AM-oriented- ResNet-34* feature →*ARf*
End For
Trained *ModifiedCenM* using *ARf*, and compute framework train time as ***t_res***
η_ res ← PredictDiseasedLoc(*ARf*)
***Ap_ res* ← Validate_AP (*AM-oriented- ResNet-34*, *η_ res*)**
Test Part
For sample *S* from T_t
i) Compute features with tuned model ϵ→βR
ii) [RoI, mAP, class] ←Predict (βR)
iii) Output image with RoI, mAP, and class
End For

3.1. Annotations

To begin our research, we carefully made detailed annotations that are crucial for training the proposed model. The annotation process involved the systematic identification and labeling of key areas within medical imaging datasets, specifically tailored for lung cancer recognition in our case. This sophisticated annotation scheme was designed to capture diverse instances of lung abnormalities, ensuring a robust and representative training dataset. The annotations, serving as the ground truth for the model, were crafted with precision to encompass the complex variations in lung cancer appearances. This crucial step in dataset preparation establishes a solid groundwork, facilitating the subsequent training of the AM-based ResNet34-oriented CenterNet model for accurate and context-aware lung cancer recognition. To achieve this, we employed the LabelImg [26] software to construct sample annotations. These annotations were generated and stored in an XML file, capturing essential information such as the output label for each affected region and the bbox scores necessary for outlining a rectangular box around the infected region. Subsequently, the XML file was used to create the training file, a pivotal component employed in training our model. This step ensures that the model learns from the annotated data, ultimately enhancing its ability to recognize and classify lung cancer accurately.

3.2. CenterNet

In the context of lung cancer recognition, obtaining precise and discriminative features is crucial for effective categorization. However, determining an optimal set of feature vectors poses challenges, as the model may face overfitting with large-sized keypoints vectors or miss important object behaviors with a small keypoints-set. Employing handcrafted feature calculation for models can result in reduced robustness, especially when dealing with extensive variations in magnitude, structure, shade, and location of lung abnormalities. To address these challenges, we have chosen the CenterNet model, a deep learning-based framework known for its capability to automatically compute dense and reliable features directly from input samples. The convolution filters within CenterNet analyze the structural aspects of suspected samples, providing a means to extract discriminative and robust features essential for accurate lung cancer identification and recognition.

The decision to adopt CenterNet over traditional approaches like RCNN [42], Fast-RCNN [43], and Faster-RCNN [44] for lung cancer recognition is driven by considerations of efficiency, simplicity, and adaptability. While R-CNN and its variants have demonstrated effectiveness, they suffer from a two-stage process involving region proposal and subsequent classification, introducing computational complexities and potential speed limitations. Fast R-CNN improved on this by combining region proposal and classification, yet Faster R-CNN further introduced the region proposal network. These additional steps can contribute to increased computational overhead, especially when dealing with large medical imaging datasets. Moreover, the reliance on anchor boxes in Faster R-CNN introduces challenges in handling diverse scales of lung abnormalities. The motivation to opt for CenterNet is rooted in its single-stage architecture, which streamlines the detection process, making it computationally efficient and suitable for nuanced tasks like lung cancer recognition.

In the context of lung cancer recognition, identifying key points of interest poses challenges due to several factors, including the precise localization of affected regions amid intense light and color variations, as well as discerning the category of each detected abnormality. CenterNet emerges as a solution with distinctive advantages for our task. By leveraging its heat maps and adopting a one-stage recognition algorithm, CenterNet excels in accurately identifying and categorizing affected lung regions across diverse categories. The incorporation of the Heat-map module, centered around the key points, enhances recognition performance, contributing to minimized extraction time for keypoints in our model. This strategic utilization of CenterNet addresses the complexities associated with lung cancer recognition, ensuring effective localization and categorization of abnormalities in medical imaging datasets.

3.3. Modified CenterNet

The conventional CenterNet approach, employing HourGlass-104, DLA-34, and ResNet-101 as feature extractors, possesses certain limitations that underscore the need for a custom approach. HourGlass-104, while effective in capturing multi-scale features, suffers from computational inefficiency and high memory requirements, hindering its scalability to larger datasets. Similarly, DLA-34, although renowned for its hierarchical feature extraction, lacks the depth necessary to recognize sophisticated patterns in complex medical imaging data, potentially compromising detection accuracy. Additionally, while ResNet-101 is widely utilized for its depth and robustness, its fixed architecture limits

its adaptability to the effective characteristics of lung abnormalities, potentially leading to suboptimal performance in certain scenarios. These limitations highlight the necessity of a custom CenterNet approach tailored to address the specific requirements and challenges of lung cancer recognition, offering improved efficiency, accuracy, and adaptability in medical imaging tasks.

To overcome the limitations of the conventional feature extractors in CenterNet, we have incorporated ResNet-34 with a Convolutional Block Attention Module (CBAM). This modification aims to enhance feature extraction by integrating attention mechanisms directly into the convolutional blocks of ResNet-34. The CBAM selectively emphasizes informative features while suppressing irrelevant ones, thereby improving the model's power to learn meaningful patterns in lung cancer images. By leveraging ResNet-34's depth and the attention mechanism's adaptability, our custom CenterNet approach offers a refined feature extraction process, leading to enhanced accuracy and robustness in lung cancer recognition tasks. In addition, our custom CenterNet approach also emphasizes a lightweight architecture to optimize computational efficiency without sacrificing performance. This lightweight design prioritizes resource-efficient operations, enabling faster inference times and reduced memory footprint, which is particularly advantageous for real-time applications and deployment on resource-constrained devices. By leveraging streamlined architectures and efficient operations, our custom CenterNet model offers a balance between computational efficiency and accuracy, making it well-suited for practical use cases in lung cancer recognition and medical imaging tasks.

The entire approach comprises 2 main steps to perform the detection and classification of lung cancer samples which are recognized as extracting relevant sample features and accomplishing the recognition task. The details are defined in the subsequent sections.

3.3.1. Keypoints computation: CBAM-based ResNet-34 approach

In our approach, we focus on utilizing a backbone network, typically a CNN, to extract keypoints maps that offer semantic and effective representations of input lung cancer images. These keypoint maps play a critical role in detecting target areas and facilitating classification tasks within object detection algorithms. The choice of backbone network greatly influences the detection performance, with more robust feature extraction leading to improved accuracy. To enhance feature extraction in our CenterNet approach, we adopted ResNet-34 [45] along with CBAM [46]. ResNet-34 is a state-of-the-art CNN model renowned for its use of identity shortcut links and residual mapping across layers, which contribute to its high accuracy. Unlike traditional deep networks, ResNet addresses the issue of gradient vanishing by introducing skip connections that bypass certain layers, allowing for better accuracy and more manageable training [47]. A comprehensive explanation of the feature extractor is mentioned in Table 2. The ResNet-34 model consists of 33 convolutional layers assembled into five units, each comprising several residual blocks. These residual blocks incorporate shortcut connections to reuse keypoint maps from previous layers, enhancing accuracy and facilitating training. The visual representation of the residual block is given in Fig. 2. Within the residual block, the layered components execute residual mapping through the establishment of shortcut links, which serve to locate mapping (j). These connections combine the outcomes with the residual function $F(x)$ of the stacked layers' output, yielding the resultant value of the residual block as follows:

$$T = F(j) + j \tag{1}$$

In equation (1), j indicates the input, F and T show the residual method and the outcome.

In our model, we have tailored the ResNet-34 architecture by incorporating a CBAM-based attention block at the beginning of the model. This modification is motivated by the desire to enhance feature representation through the utilization of an attention strategy. The inclusion of the attention block enables the model to prioritize affected areas while overwhelming unrelated background details, thereby refining recognition results even in complex scenarios such as variations in shades, luminance, and intensity. The CBAM unit enhances keypoints derived from the CNN by incorporating spatial and channel-wise attention, thereby augmenting the performance of the deep NN (DNN). Notably, the lightweight architecture of the CBAM block introduces minimal computational overhead, allowing for end-to-end training alongside base CNNs. The structural design of the custom network, featuring the attention block, is detailed in Table 2. Additionally, we have replaced the early 7x7 convolution and max-pooling layer with three stacked 3x3 convolution layers to eliminate the downsampling phase in the initial convolutional layer [48]. Further, to reduce the computing burden, the channel size for the newly added convolutional layers has been marked to 64.

3.3.2. Functional heads computation

The proposed approach computes several heads like heatmap, dimensions, and offset heads. These heads are computed to perform essential tasks for accurate object detection and recognition in the CenterNet approach, including classifying objects into different categories, precisely localizing objects within images, refining bounding box

Table 2
Architecture details of ResNet-34 original and modified.

Model Layer	Actual	Custom
Con_1	7×7, 64, 3×3 max pool	$[3 \times 3, 64] \times 3$, 7×7 Attention
Con_2_x	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$
Con_3_x	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 4$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 4$
Con_4_x	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 6$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 6$
Con_5_x	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$

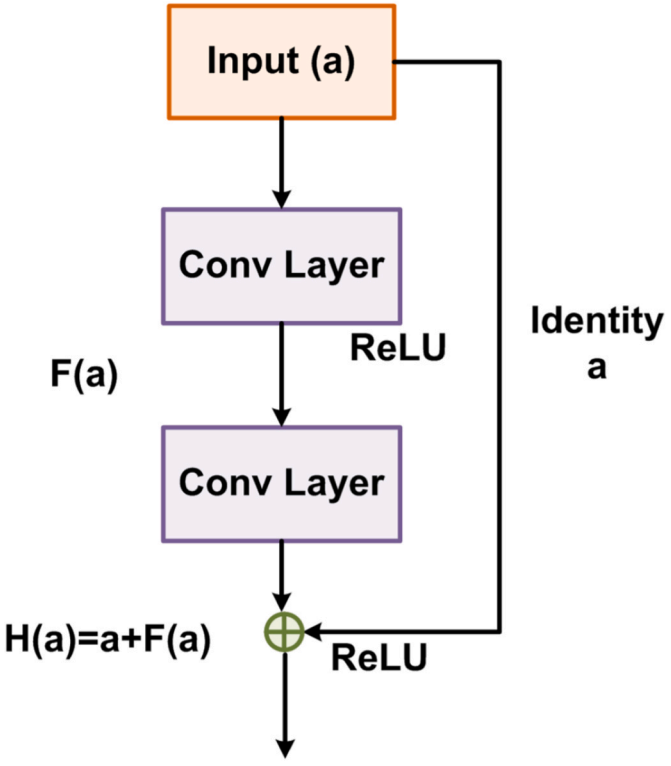

Fig. 2. A visual representation of the residual block.

predictions, providing supplementary information about object size and shape, facilitating tasks such as pose estimation or object manipulation, enabling detailed segmentation of objects from the background, and enhancing overall performance by prioritizing informative features while suppressing irrelevant or noisy information.

3.3.2.1. HeatMAP HEAD. The heatmap head in the CenterNet approach is responsible for generating heatmaps that highlight the presence and location of objects within an image. These heatmaps provide a visual representation of where objects are located and their spatial extent. Each heatmap corresponds to a specific class or category of objects that the model is trained to detect. During training, the heatmap head learns to produce high-intensity values at the pixels corresponding to the center of objects, while assigning lower values to surrounding pixels. This process enables the model to accurately localize objects within the image. Heatmaps serve as an intermediate representation that guides subsequent stages of object detection, such as bounding box regression and classification. Overall, the heatmap head plays a crucial role in facilitating precise object localization and recognition in the CenterNet approach. The mathematical interpretation of the heatmap is provided as:

$$\widehat{\boldsymbol{Q}}_{s,t,n} = \exp\left(-\frac{(\boldsymbol{s}-\widehat{\mathbf{e}}_{\mathbf{s}})^2+(\boldsymbol{t}-\widehat{\mathbf{e}}_{\boldsymbol{t}})^2}{2\sigma_p^2}\right) \tag{2}$$

In Equation 2, s and t are denoting the dimensions of the original candidate point, and $\widehat{e_i}$ and $\widehat{e_j}$ show estimated down-sampled candidate points. σ and p represents the object size-adaptive standard deviation. Total outputs are denoted by n and the $\widehat{Q}_{s,t,n}$ indicates the center against a detected feature with a score of 1, else, marked as background.

3.3.2.2. Dimension head. The Dimension Head in the CenterNet approach is responsible for predicting additional properties of detected objects, such as width, height, and depth. Unlike traditional object detection methods that focus solely on bounding box coordinates, the Dimension Head enriches the object representation by providing supplementary information about the size and shape of objects. By predicting dimensional attributes, the Dimension Head enables a more comprehensive understanding and characterization of detected objects. This additional information can be particularly useful in scenarios where precise object dimensions are crucial, such as in 3D object detection tasks or when dealing with objects of varying shapes and sizes. The estimation of the rectangular box containing a key object b with output label l having dimensions ($i1$, $i2$, $j1$, $j2$) is computed by using the $L1$ norm.

3.3.2.3. Offset head. The Offset Head in the CenterNet approach plays a crucial role in refining the initial bounding box predictions generated by the model. Its primary function is to predict the offsets or displacements required to align the estimated bounding boxes more accurately with the ground truth annotations of objects within the image. This refinement process helps to mitigate errors and inconsistencies in the initial bounding box predictions, resulting in more precise object localization. By adjusting the position and size of bounding boxes based on predicted offsets, the Offset Head boosts the model's aptitude to accurately delineate the boundaries of detected objects. This fine-tuning of bounding boxes contributes to improved object detection performance, particularly in scenarios where objects exhibit varying scales, orientations, or aspect ratios.

3.3.3. MULTI-LOSS function

The multi-loss function, also known as the multi-task loss function, is a key component in training neural network models, particularly in the context of multi-task learning where the network is skilled to execute numerous associated tasks simultaneously. In the context of the CenterNet approach, which involves multiple heads responsible for various tasks such as object detection, classification, and dimension prediction, the multi-loss function calculates the overall loss by aggregating individual losses from each task. The goal of the multi-loss function is to jointly optimize the model parameters across all tasks, ensuring that the model learns representations that are effective for all tasks simultaneously. By incorporating information from multiple tasks into a single loss function, the multi-loss function encourages the model to acquire collective representations that share common underlying patterns across tasks, assisting in enhanced generalization and performance.

The proposed approach uses the following multi-loss M over each mentioned head above as follows:

$$M_{CNet} = H_{loss} + c_{dim}D_{loss} + c_{off}O_{loss} \tag{3}$$

In Equation (2), M_{CNet} is the collective multi-loss computed by the CenterNet approach, while H_{loss}, D_{loss}, and O_{loss} are the loss values computed over the Heatmap, Dimension, and Offset heads, respectively. While, c_{dim} and c_{off} are constants having scores of 0.1 and 1, respectively.

The H_{loss} is calculated as:

$$H_{loss} = \frac{-1}{t}\sum_{s,t,n,}\begin{cases}(1-\widehat{\boldsymbol{p}}_{s,t,n})^{\alpha}\log(\widehat{\boldsymbol{p}}_{s,t,n}). & \text{if } \widehat{\boldsymbol{p}}_{s,t,n}=1\\ & \text{otherwise}\\ (1-\mathbf{p}_{s,t,n})^{\beta}(\widehat{\boldsymbol{p}}_{s,t,n})^{\alpha} & \\ \log(\mathbf{1}-\widehat{\boldsymbol{p}}_{s,t,n}) & \end{cases} \tag{4}$$

In Equation (4), t designates total key features, $\boldsymbol{p}_{s,t,n}$ shows the real center value of a key feature, and $\widehat{\boldsymbol{p}}_{s,t,n}$ the estimated center value of the candidate point. Further, α and β are the hyperparameters with scores of 2 and 4 for all our experiments, respectively.

The D_{loss} is computed as:

$$D_{loss} = \frac{1}{t}\sum_{a=1}^{t}|\widehat{\boldsymbol{c}}_a - c_a| \tag{5}$$

In Equation (5), $\widehat{\boldsymbol{c}}_a$ shows the estimated value of a rectangular box, and c_a indicates the actual dimensions. While O_{loss} is estimated as:

$$O_{loss} = \frac{1}{t}\sum_{p}\left|\widehat{V}_{\dot{k}} - \left(\frac{k}{R}-\widehat{k}\right)\right| \tag{6}$$

Here, $\widehat{V}$ shows the computed score of offset, while k and $\widehat{k}$ indicating the original and down-sampled key features.

3.3.4. Detection procedure

CenterNet represents a deep learning-based methodology that diverges from conventional approaches like selective search and proposal generation. In our work, the CenterNet framework operates on input images along with corresponding annotations generated for suspected lung cancer regions. Within this framework, CenterNet computes crucial parameters including the center points of lung abnormalities, offsets to the x and y coordinates, and the dimensions of bounding boxes, alongside identifying the target class associated with each anomaly. This approach streamlines the object detection process by directly inferring these parameters from the input data, thereby enhancing efficiency and accuracy in identifying lung cancer regions within medical imaging datasets.

Overall, our research offers significant societal benefits by improving the early detection and diagnosis of lung cancer which is critical for increasing survival rates. The proposed framework enhances diagnostic accuracy through advanced feature extraction and attention mechanisms, reducing false positives and negatives, thereby minimizing misdiagnoses. Its efficient single-stage detection approach ensures faster processing, making it suitable for real-time clinical applications and

reducing the burden on healthcare professionals. Additionally, the lightweight design of the framework promotes accessibility in resourced-limited medical facilities, enabling widespread adoption and improving outcomes for a larger population.

4. Result

Lung cancer detection poses several challenges, including the complexity of lung nodule structures, variations in size, shape, and texture of cancerous regions, and the influence of environmental noise in CT images. Traditional approaches often rely on multi-stage object detection frameworks that are computationally intensive and prone to inaccuracies. Key issues include inefficient feature extraction, limited interpretability of predictions, and susceptibility to false positives and negatives, which can lead to delayed or incorrect diagnoses. The results section highlights how the proposed framework addresses critical challenges in lung cancer detection, including complex nodule structures, environmental noise, and inefficiencies in traditional methods. By integrating ResNet-34 with a CBAM within the CenterNet architecture, the framework enhances feature extraction and focuses on relevant regions while suppressing background noise. This single-stage detection model offers faster inference and improved accuracy compared to multi-stage approaches like Faster R-CNN, with robust performance in noisy environments. The effective localization ability of the model provides interpretable predictions, enabling clinicians to identify key areas of interest, which fosters trust in its application. Experimental results on the two publicly available data samples, named the LUNA-16 and Kaggle lung cancer datasets, demonstrate superior performance compared to state-of-the-art methods. These contributions underline the framework's potential for early and accurate lung cancer detection, aiding timely clinical interventions and improving patient outcomes.

In short, this section holds the description of the employed data samples for network learning and testing along with the parameters employed for model results assessment. Further, we have provided a detailed results section indicating the effectiveness of the approach under diverse experimental scenarios.

4.1. Dataset

For model training and testing, we have employed two publicly available data samples named the LUNA-16 and Kaggle lung cancer datasets. The LUNA-16 dataset, a subset of the publicly available lung nodule dataset LIDC-IDRI, short for the Lung Nodule Analysis 2016 challenge data sample, is a widely used benchmark dataset for lung nodule detection and classification tasks [49]. It comprises computed tomography (CT) scans of the chest from various patients, annotated by radiologists to identify and characterize pulmonary nodules. The LUNA16 dataset encompasses images derived from 888 chest CT scans along with corresponding information regarding the locations of lung nodules. Notably, samples having a thickness of more than 2.5 mm in the LIDC-IDRI database were omitted from the LUNA16 dataset. Each image in the LUNA16 dataset was meticulously annotated by four domain experts, who confirmed the existence or nonappearance of nodules. For this study, a total of 1186 CT images were curated by processing the actual dataset. LUNA-16 provides detailed annotations for each nodule, including its location within the lung volume, size, shape, and density characteristics. With over a thousand annotated CT scans, the LUNA-16 dataset serves as a valuable resource for developing and evaluating algorithms aimed at improving the detection, diagnosis, and treatment of lung cancer. For the second dataset, we have collected normal and lung cancer-affected samples from two sources of Kaggle as provided in Refs. [50,51], respectively. This dataset accompanies a total of 1000 lung nodule images collected from Ref. [50] and 250 normal samples collected from Ref. [51]. Both datasets contain a diverse range of lung nodules, including both malignant and benign cases, captured across different imaging modalities and settings.

4.2. Parameters

To analyze the identification, and classification results of the proposed framework, several evaluators like precision value, recall measure, accuracy metric, F1-measure, and mean Average Precision (mAP) are used. The numeric elaboration of the accuracy metric is provided in Equation # 7.

$$Accuracy = \frac{TrP + TrN}{TrP + FaP + TrN + FaN} \tag{7}$$

Further, the formula to compute the mAP is discussed in Equation # 8, where s and S designate the analyzed test sample and a total number of images, respectively.

$$mAP : = \sum_{j=1}^{S} AP(s_j) \,/\, S \tag{8}$$

The pictorial representation of the precision and recall is given in Fig. 3.

4.3. Model evaluation

In this part, we have discussed the results of our method attained on both datasets namely LUNA-16, and Kaggle data samples with the help of various performance measuring indicators metrics like precision, recall, F1-Score, and accuracy.

First, we have reported our results on both data samples in the aspect of precision, recall, and F1-Score as it provides a comprehensive assessment of the robustness and generalization capability of our proposed approach. By analyzing performance metrics across different datasets, we gain an understanding of the framework's power to precisely detect lung cancer across diverse patient populations and imaging conditions. Precision estimates the amount of accurately recognized positive cases among total estimated positive cases, showing the network accuracy in locating lung cancer cases. Recall evaluates the network's aptitude to estimate all positive cases, thus indicating its sensitivity to detecting lung cancer. The F1-score, which takes both precision and recall, provides a stable estimation of the suggested framework's overall behavior. By presenting results using these metrics across both datasets, we demonstrate the reliability and effectiveness of our approach in lung cancer detection, facilitating its adoption in clinical practice and research settings. The attained values are provided in Fig. 4 from where it is quite visible that our approach shows significant values across all evaluation measures for both datasets. Descriptively, for the LUNA-16 dataset, we have acquired the precision, recall, and F1-Score of 99.89 %, 99.82 %, and 99.85 %, which are 98.33 %, 98.02 %, and 98.17 % for the other dataset, clearly depicting the effectiveness of our approach.

Next, we have reported the classification results on both datasets in the form of an accuracy measure as it is crucial for assessing the overall performance and generalization ability of our proposed approach. Accuracy measures the amount of appropriately classified test images among all samples in the test set, providing a comprehensive evaluation of the model's predictive power. By presenting accuracy values on multiple datasets, we can evaluate how well our model performs across different patient populations, imaging protocols, and data distributions. This allows us to assess the efficacy and reliability of our approach in practical situations and ensures that the model's performance is not biased towards specific datasets. The attained accuracy scores are provided in Fig. 5 which clarifies that our model shows high classification results on both datasets with scores of 99.63 % and 98.21 % on the LUNA-16 and Kaggle datasets, respectively.

Next, we have discussed the results in terms of the confusion matrix for both datasets. In the context of lung cancer detection, confusion matrices are instrumental in evaluating the performance of machine learning models tasked with classifying CT images such as benign and

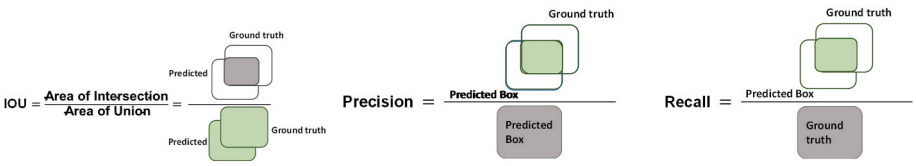

Fig. 3. Pictorial illustration of (a) Precision, (b) Recall, and (c) IOU evaluators.

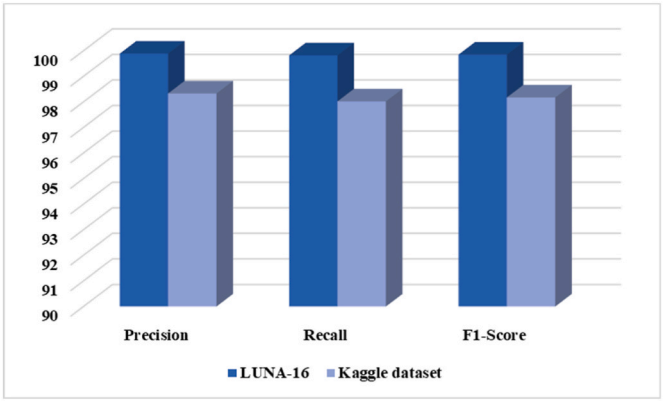

Fig. 4. Performance comparison of the introduced work in the form of precision, recall, and F1-Score across both datasets.

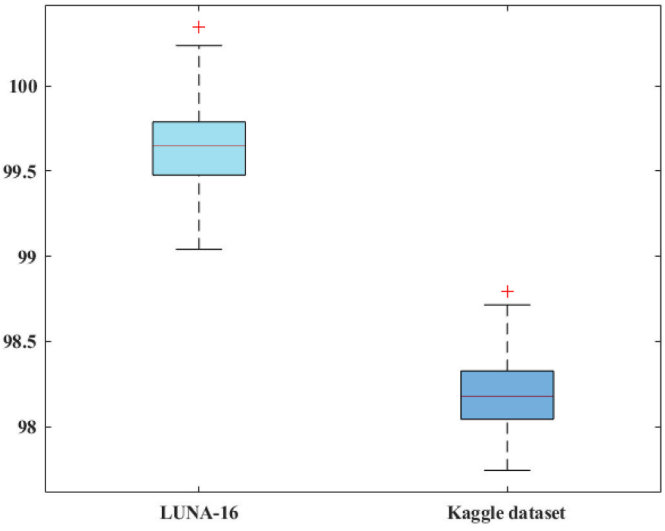

Fig. 5. Accuracy scores attained by our approach on both data samples.

malignant nodules. These matrices provide a thorough analysis of the model's forecasts, including true positives, false positives, true negatives, and false negatives. By examining the confusion matrix, domain experts and researchers can assess the model's accuracy in correctly identifying cancerous and non-cancerous nodules, as well as the prevalence of misclassifications. This information is crucial for optimizing model parameters, refining feature extraction techniques, and identifying areas for improvement in lung cancer detection algorithms. Moreover, confusion matrices enable the identification of specific types of errors, such as false negatives, which are particularly critical in the context of medical diagnosis as they may lead to missed cancer diagnoses and delayed treatments. The attained confusion matrixes for the LUNA-16 and Kaggle data samples are provided in Fig. 6 from where it can be visualized that our approach performs well for both datasets and shows minimum false positive rates.

Finally, we have presented the localized results for lung nodule recognition on the employed data samples. Localized results for lung nodules are essential as they offer visual confirmation of the model's predictions by overlaying bounding boxes or segmentation masks onto CT images. This visual feedback allows clinicians to verify the accuracy of the model's detections, instilling confidence in its diagnostic capabilities. Moreover, localized results enhance interpretability by providing insights into the features driving the model's decisions, aiding clinicians in understanding and trusting the underlying algorithm. The attained results are shown in Fig. 7, which clearly indicates that our approach can locate the lung cancer-affected areas with a high recognition rate. Descriptively, we have attained the mAP score of 0.9221 for the LUNA-16 dataset, which is 0.9140 for the Kaggle data sample, clearly proving the effectiveness of our approach.

4.4. Comparison with base models

In this part, we have compared our results with other object detection approaches like Faster-RCNN [52], YOLOv3 [53], YOLOv3 +BBO/EE [54], YOLO v3 +Adams [54], YOLOv6 + Advanced PSO [55], and UNet + RandomForest [56]. Comparing our results with other object detection approaches provides a crucial benchmark for evaluating the effectiveness and competitiveness of our proposed method. We have compared the results with other approaches in terms of accuracy metric, and the attained evaluation is provided in Table 3. The values in Table 3 prove the effectiveness of our approach as we acquired the highest accuracy results. The Faster-RCNN with adaptive anchor box has attained comparable results with an accuracy rate of 95.66 %. The YOLOv3

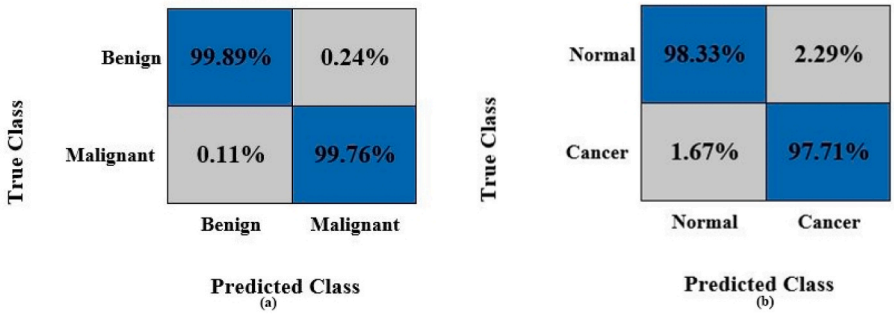

Fig. 6. Confusion matrix (a) LUNA-16 dataset, (b) Kaggle dataset, attained by our approach.

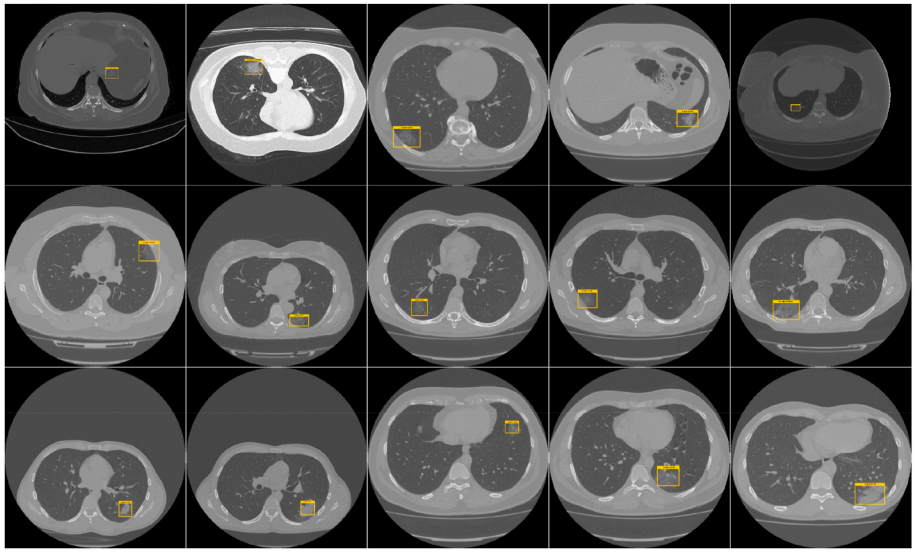

Fig. 7. Localized samples by the proposed work.

Table 3
Analysis of the proposed work with the base works.

Model	Accuracy (%)
Faster-RCNN with adaptive anchor box	95.66
YOLOv3	80.60
YOLOv3 +BBO/EE	91.20
YOLO v3 +Adams	90.55
YOLOv6 + Advanced PSO	82.79
UNet + RandomForest	74
Proposed	**99.63**

+BBO/EE also shows better results with an accuracy value of 91.20 %. The UNet approach with the RF classifier shows the lowest classification results with an accuracy value of 74 %. In comparison, the proposed approach outperforms other approaches with a performance gain of 13.83 %. Faster R-CNN, UNet + RF, and YOLO models face various challenges in the context of lung cancer detection. The main issue of the Faster-RCNN approach is its complexity in implementation and training. The multi-stage architecture, which includes separate components for region proposal and object detection, requires intricate optimization and tuning of hyperparameters. This complexity can lead to challenges in deployment and scalability, especially in resource-constrained environments or when dealing with large-scale datasets. Additionally, the training process for Faster-RCNN typically requires a substantial amount of computational resources and time, making it less practical for real-time applications or scenarios with limited computing capabilities. While UNet + RF struggles with accurately capturing fine details and boundaries of lung nodules. YOLO family, on the other hand, encounters difficulties in precisely localizing small and overlapping nodules. In contrast, our AM-based ResNet-34 with CenterNet approach addresses these issues effectively. By incorporating an attention mechanism into ResNet-34, we enhance the model's ability to focus on more representative features indicative of lung cancer, thereby improving detection accuracy. Additionally, CenterNet's one-stage detection architecture streamlines the inference process, resulting in faster and more efficient performance compared to Faster R-CNN. Furthermore, the attention mechanism assists in capturing fine details and complex boundaries of lung nodules, overcoming the limitations of UNet + RF. Moreover, our approach's robust feature extraction capabilities enable precise localization of small and overlapping nodules, enhancing detection performance relative to YOLO. Overall, our AM-based ResNet-34 with CenterNet approach offers a comprehensive solution that addresses the challenges faced by existing models in lung cancer detection, leading to improved accuracy and efficiency.

4.5. Comparison with DL models

In this section, the results of our approach are analyzed with several DL models. Comparing our results against other established methods is crucial for validating the efficacy and innovation of our approach. This comparative analysis serves as a means to benchmark the performance of our method against existing state-of-the-art DL in the field like the EfficientNet family [57], NASNetMobile [58], DenseNet121 [59], and MobileNetV2 [60], and MobileNet [61] as mentioned in Ref. [62]. By assessing how our results fare against those obtained by other approaches, we gain valuable insights into the relative strengths and weaknesses of our methodology. The attained comparison is given in Table 4, which indicates that our approach attains the highest results in comparison to the latest DL works with precision, recall, F1-score, and accuracy scores of 99.89 %, 99.82 %, 99.85 %, and 99.63 %. The second-highest classification scores are attained by the EfficientNetB2 with precision, recall, F1-score, and accuracy values of 95.50 %, 94.60 %, 95.10 %, and 95.40 %. The EfficientNetB3 also shows comparable results with precision, recall, F1-score, and accuracy values of 95.30 %, 94.60 %, 94.70 %, and 95.40 %. While, the NASNetMobile model achieves the lowest results with precision, recall, F1-score, and accuracy values of 88.25 %, 86.65 %, 87.65 %, and 87.45 %. Whereas, our work shows the highest classification scores in the aspect of all evaluation metrics. The main cause for the improved performance of the proposed approach in recognizing the healthy and cancerous affected samples of

Table 4
Comparison with DL models.

Model	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
EfficientNetB0	95.60	93.50	95	95.30
EfficientNetB1	95.70	94.40	94.90	95.20
EfficientNetB2	95.50	94.60	95.10	95.40
EfficientNetB3	95.30	94.60	94.70	95.40
EfficientNetV2B0	95.55	94.45	94.85	95.25
EfficientNetV2B1	95.65	94.35	94.75	95.15
EfficientNetV2B2	95.45	94.55	95.35	94.95
NASNetMobile	88.25	86.65	87.65	87.45
DenseNet121	86.80	89.10	88.50	87.50
MobileNetV2	87.30	88.80	87.90	89.00
MobileNet	87.90	88.50	87.40	88.80
Proposed	**99.89**	**99.82**	**99.85**	**99.63**

lung CT scans is the enhanced feature computation capability of the proposed approach. The EfficientNet family models are adept at learning hierarchical representations from images, including features at different levels of abstraction, and are unable to fully capture the complex characteristics of lung nodules relevant for accurate diagnosis. Lung nodules exhibit diverse morphological and textural variations, and discerning between benign and malignant nodules often relies on subtle visual cues, which is a challenging task. Moreover, NASNetMobile, DenseNet121, MobileNetV2, and MobileNet exhibit specific limitations when applied to lung cancer classification tasks. NASNetMobile, characterized by its compact architecture, struggles with learning effective structural variations within lung nodules due to its limited capacity for capturing fine-grained features and spatial relationships. DenseNet121, despite its efficient feature reuse, faces challenges in accurately characterizing the diverse texture differences present in lung nodules, potentially compromising its discriminative power for precise cancer classification. MobileNetV2's lightweight design, while advantageous for computational efficiency, hinders its ability to capture the dense details and spatial context crucial for distinguishing between benign and malignant nodules, resulting in reduced classification results. Similarly, MobileNet's simplified network architecture, though efficient, does not adequately capture the complex visual patterns inherent in lung nodules, limiting its effectiveness in accurately classifying lung cancer from medical images. Our approach overcomes the limitations of existing DL approaches for lung cancer recognition by employing a combination of innovative techniques. We enhance feature extraction using an improved keypoints computation module based on ResNet-34 augmented with an attention mechanism, allowing for more effective extraction of cancer-indicative characteristics from input samples. By integrating the CenterNet approach, our framework enables accurate localization and classification of diseased areas within lung scans, overcoming challenges associated with variations and complex patterns in lung nodules. Through comprehensive evaluation using standardized measures, we validate the effectiveness and reliability of our approach, ensuring its superiority in accurately detecting lung cancer from CT-Scan images and advancing the state-of-the-art in medical image analysis.

4.6. Analysis with state-of-the-arts

In this study, we benchmark our proposed method against state-of-the-art approaches in the domain of lung cancer detection. By comparing our results with established benchmarks, we validate the efficacy and innovation of our approach. This comparison serves to highlight the advancements made by our method and its potential to contribute meaningfully to the field. To perform this, we have nominated several new works [31,62–65] that performed lung cancer detection on the LUNA-16 dataset and compared our results against them. The analysis given in Table 5 shows that the suggested approach attains the highest results in contrast to the new framework in the aspect of all performance measuring metrics.

Table 5
Comparison with latest approaches.

Approaches	Year	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
Rehman et al. [62]	2024	95.80	94.69	95.24	95.40
Thangavel et al. [63]	2024	99.19	99.22	99.20	99.29
Naseer et al. [64]	2023	–	96.37	–	97.64
Siddiqui et al. [31]	2023	–	98.048	–	99.161
Alsheikhy et al. [65]	2023	99.88	99.76	99.82	99.42
Proposed	**2024**	**99.89**	**99.82**	**99.85**	**99.63**

Rehman et al. [62] introduced a CNN equipped with a dual attention mechanism tailored for analyzing lung nodule images. The CNN extracted informative features, while the attention module selectively emphasized significant elements using both channel and spatial attention mechanisms. Following the attention module, global average pooling was employed to consolidate spatial information. The approach attains an accuracy of 95.40 %. The work in Ref. [63] introduced an effective DL approach for classifying pulmonary nodules from CT images. It began with various pre-processing techniques to prepare the data, followed by segmentation of lung nodules using a TNet-based deep learning algorithm. Next, a CenterNet-based approach learned patterns and intensity characteristics from the segmented images. Finally, a NASNet-oriented predictor categorized the nodules as cancerous or non-cancerous based on the collected attributes. This approach attains an accuracy value of 99.29 %. Naseer et al. [64] suggested an approach named LungNet-SVM based on a modified AlexNet approach with dense keypoints computation. Next, the computed features were employed to train the SVM classifier to categorize the normal and cancerous areas. This approach attains an accuracy value of 97.64 %. Siddiqui et al. [31] presented a study for classifying lung cancer that utilized Gabor filters in conjunction with an enhanced Deep Belief Network (E-DBN) incorporating multiple classification techniques. The E-DBN comprised two cascaded Restricted Boltzmann Machines (RBMs). Among the applied methods, the SVM demonstrates optimal performance parameters with an accuracy rate of 99.161 %. The work in Ref. [65] employed a combination of DL techniques, including VGG-19 and long short-term memory networks (LSTMs), which were customized and integrated for lung cancer detection and classification. Additionally, image segmentation techniques were applied as part of a computer-aided diagnosis (CAD) system. Experimental results conducted in MATLAB demonstrate that the combined use of these tools achieves an accuracy of over 99.42 %. While our approach reports the top classification scores with an accuracy score of 99.63 % and provides a performance gain of 1.45 % in comparison to other approaches. Similarly, for the precision and recall scores, the nominative methods have attained average scores of 98.29 % and 97.62 %, which are 99.89 % and 99.82 % for our work, so we have provided performance gains of 1.6 % and 2.20 % for the mentioned metrics, respectively. Finally, for the F1-Score measure, the comparative approaches have attained an average score of 98.09 %, which is 99.85 % for our scenario, so we have provided a performance gain of 1.76 %, which indicates the effective recognition power of our approach. The major cause for this effective classification capability is that the comparative approaches [31,62–65] cannot learn the deep characteristics of the investigated samples and are unable to tackle the intense variations in the mass of the lung nodules. Further, these methods are unable to locate the exact region of interest. Comparatively, the modified CenterNet approach employs a more powerful feature extractor named CBAM-based ResNet-34, which directly utilizes the attention mechanism into the convolutional blocks of the feature extractor. The inclusion of the attention strategy empowers the model to rank affected areas while overwhelming unrelated background details, thereby upgrading recall power even in complex cases containing alterations in mass, color, light, and intensity of the samples, which in turn improves the recognition capability of the suggested approach.

4.7. Discussion

This research addresses the critical challenge of early lung cancer detection, a major issue that continues to impact global health. Lung cancer, being one of the most prevalent and deadly cancers, often presents at advanced stages, where treatment options are limited. Early detection, typically through medical imaging such as CT scans, is crucial for improving patient survival rates. However, the complexity of medical imaging, the presence of subtle nodules, and variations in imaging conditions pose significant challenges for automated detection systems. Previous methods like Faster-RCNN, YOLO, and UNet-based approaches,

while achieving varying degrees of success, still face limitations in terms of accuracy, efficiency, and interpretability. This work proposes an enhanced deep learning framework that combines the CenterNet approach with a ResNet-34 backbone network and an attention mechanism (CBAM), aiming to overcome these challenges and improve lung cancer detection. The novelty of the proposed approach lies in the integration of the CBAM attention mechanism with the ResNet-34 model within the CenterNet architecture. By incorporating this attention mechanism, the model can better focus on key features in the CT images, improving feature extraction and ensuring more accurate predictions, especially in the presence of complex backgrounds. The model's single-stage detection approach, compared to the multi-stage methods used by other architectures, reduces computational complexity and enhances efficiency, making it suitable for real-time applications in clinical environments. The extensive experimental results, obtained from both the LUNA16 and Kaggle datasets, show that the proposed model outperforms existing models in terms of both detection accuracy and inference speed. These results highlight the potential of the proposed approach to enhance the diagnostic capabilities of healthcare professionals, providing them with more reliable tools for detecting lung cancer at early stages.

In terms of contributions, the work introduces several significant advancements in lung cancer detection using deep learning. The incorporation of the attention mechanism improves the model's ability to identify subtle patterns in images, increasing both the accuracy and interpretability of predictions. The use of CenterNet's single-stage detection approach streamlines the process and offers significant computational benefits, making it a practical solution for real-time medical applications. The localization capability, allowing the model to highlight the regions of interest in the images, adds an element of transparency to the decision-making process, which is essential in clinical environments. This approach not only improves the detection rate but also aids clinicians in understanding the model's reasoning, which is crucial for establishing trust in AI-based tools. However, despite the promising results, this study also acknowledges several limitations. One of the key challenges is the model's reliance on the LUNA16 and Kaggle datasets, which, although widely used, may not fully represent the diversity and complexity of real-world clinical data. The narrow scope of these datasets may limit the generalizability of the model to a broader patient population, especially in terms of varying CT scan qualities, demographic diversity, and disease stages. Furthermore, while the proposed model achieves good computational efficiency, the depth of the ResNet-34 network may still lead to high training times and significant computational costs, especially when dealing with large-scale datasets. Although the use of a single-stage detection method helps mitigate some of these challenges, further optimization of the model architecture, such as reducing the number of parameters or using more lightweight networks, could improve performance. Additionally, incorporating multi-modal data, such as patient history or genetic information, could further enhance the model's accuracy, allowing for a more comprehensive approach to lung cancer detection. Looking ahead, future research directions could focus on expanding the dataset used for training, incorporating more diverse and comprehensive lung cancer imaging data from various sources. This would help improve the generalization ability of the model and make it more applicable to different populations. Moreover, exploring the integration of multi-modal data, such as combining CT images with patient demographic information, genetic data, and medical history, could significantly enhance the robustness and accuracy of the model. Future work could also focus on reducing the computational complexity of the model, perhaps through further model compression techniques, pruning, or using more efficient networks like EfficientNet or MobileNet. Another promising avenue is the exploration of federated learning to allow for privacy-preserving, multi-institutional collaboration, where data from multiple hospitals could be used to improve the model without the need to share sensitive patient data.

In conclusion, the proposed enhanced CenterNet approach represents a significant step forward in the field of automated lung cancer detection. By combining state-of-the-art deep learning architectures with innovative attention mechanisms, the model improves both the accuracy and efficiency of lung cancer detection. Despite some limitations, such as the reliance on specific datasets and computational costs, this study provides a foundation for future advancements in the field. With continued research and development, including addressing the identified limitations, this approach has the potential to become a valuable tool in the early detection of lung cancer, ultimately contributing to better patient outcomes.

5. Policy suggestions

The findings of this study can inform healthcare policies aimed at improving early lung cancer detection and diagnosis. Implementing the proposed automated framework in clinical settings could enhance diagnostic accuracy, reduce the reliance on subjective interpretations, and optimize the use of resources in radiology departments. Policymakers could prioritize the integration of such AI-based tools into standard diagnostic protocols, ensuring accessibility and affordability for healthcare providers. Additionally, investing in training programs for radiologists and clinicians to effectively use these technologies can facilitate their adoption and maximize their potential impact. Finally, establishing collaborative networks between AI researchers, medical professionals, and regulatory bodies can ensure ethical use, data privacy, and consistent performance of the system across diverse populations.

6. Conclusion

Lung cancer detection poses several challenges, including the complexity of lung nodule structures, variations in size, shape, and texture of cancerous regions, and the influence of environmental noise in CT images. Traditional approaches often rely on multi-stage object detection frameworks that are computationally intensive and prone to inaccuracies. Key issues include inefficient feature extraction, limited interpretability of predictions, and susceptibility to false positives and negatives, which can lead to delayed or incorrect diagnoses. This study has introduced an Improved CenterNet approach, a novel DL framework tailored for overcoming the challenges associated with lung cancer detection. By integrating ResNet-34 with an attention mechanism within the CenterNet architecture, our approach offers enhanced feature extraction capabilities, improving the model's ability to discern subtle patterns associated with lung cancer. Through extensive experimental evaluations on a standard dataset, our proposed approach demonstrates significant improvements in accuracy, interpretability, and computational efficiency compared to existing methods. The presented framework provides transparent and interpretable predictions, facilitating better understanding and trust in the model's decisions. The model is tested on the two publicly available data samples named the LUNA-16 and Kaggle lung cancer datasets to evaluate the robustness of the approach. The suggested work shows high classification results on both datasets, with scores of 99.63 % and 98.21 % on the LUNA-16 and Kaggle datasets, respectively.

While this study offers significant contributions to lung cancer detection through the improved CenterNet approach, it is important to acknowledge several limitations that impacted the overall performance and generalizability of the model. Firstly, the model's reliance on two specific datasets LUNA16 and a Kaggle dataset introduced challenges in terms of dataset diversity. Despite pre-processing techniques to normalize the data and enhance consistency, both datasets exhibited limitations in terms of variation in scanning protocols, resolution, and image quality. These factors potentially hindered the model's ability to generalize effectively across diverse clinical environments. Secondly, while the attention mechanism incorporated into the ResNet-34

backbone enhanced feature extraction, there are still challenges in distinguishing subtle nodules from complex backgrounds, particularly in low-contrast images. Finally, the computational complexity associated with training the model, especially when working with large datasets, remains a concern. Although we employed techniques such as transfer learning and efficient backbone networks like ResNet-34, the training process could still be time-consuming.

The future scope of this research involves enhancing the proposed model's ability to detect and classify lung cancer with even greater accuracy and robustness. Future studies could focus on expanding the dataset to include a wider variety of CT scans from different patient populations and institutions, ensuring that the model generalizes well across various clinical environments. Additionally, integrating multi-modality data, such as combining CT scans with patient demographics and other medical data, could help improve the model's decision-making process. One of the limitations faced during the study was the dependency on the LUNA16 and Kaggle datasets, which limited the diversity of the data, as both datasets are specific to certain types of CT scans. To address this, future work could include data augmentation techniques or transfer learning from broader, more diverse datasets. Another limitation was the computational cost associated with training the model on large datasets. Future research can focus on optimizing the architecture to improve training efficiency, such as applying pruning techniques or using more efficient backbone networks. By overcoming these limitations, the model could be made more suitable for deployment in real-world clinical applications, making a meaningful contribution to the early detection of lung cancer. Overall, the Improved CenterNet approach holds promise for advancing the area of lung cancer recognition and assisting in improving patient outcomes in clinical practice.

CRediT authorship contribution statement

Hussain Dawood: Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Marriam Nawaz:** Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Muhammad U. Ilyas:** Writing – review & editing, Writing – original draft, Validation, Investigation, Formal analysis. **Tahira Nazir:** Writing – review & editing, Visualization, Validation, Methodology, Formal analysis, Data curation. **Ali Javed:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Investigation, Conceptualization.

Ethics statement

The authors declare that this manuscript represents entirely original works, and/or if work and/or words of others have been used, that this has been appropriately cited or quoted. This material has not been published in whole or in part elsewhere. The manuscript is not currently being considered for publication in another journal.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] N.A. Wani, R. Kumar, J. Bedi, DeepXplainer: an interpretable deep learning based approach for lung cancer detection using explainable artificial intelligence, Comput. Methods Progr. Biomed. 243 (2024) 107879.
[2] D.R. Gomez, Z. Liao, Non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC), in: Target Volume Delineation and Field Setup: A Practical Guide for Conformal and Intensity-Modulated Radiation Therapy, Springer, 2012, pp. 87–103.
[3] A. Leiter, R.R. Veluswamy, J.P. Wisnivesky, The global burden of lung cancer: current status and future trends, Nat. Rev. Clin. Oncol. 20 (9) (2023) 624–639.
[4] R.L. Siegel, K.D. Miller, N.S. Wagle, A. Jemal, Cancer statistics, CA: a cancer journal for clinicians 73 (1) (2023) 17–48.
[5] H. Padinharayil, et al., Non-small cell lung carcinoma (NSCLC): implications on molecular pathology and advances in early diagnostics and therapeutics, Genes Diseases 10 (3) (2023) 960–989.
[6] R. Vadala, et al., A review on electronic nose for diagnosis and monitoring treatment response in lung cancer, J. Breath Res. 17 (2) (2023) 024002.
[7] M.G. Lanjewar, K.G. Panchbhai, P. Charanarur, Lung cancer detection from CT scans using modified DenseNet with feature selection methods and ML classifiers, Expert Syst. Appl. 224 (2023) 119961.
[8] S. Makaju, P. Prasad, A. Alsadoon, A. Singh, A. Elchouemi, Lung cancer detection using CT scan images, Procedia Computer Science 125 (2018) 107–114.
[9] U. Pastorino, et al., Early lung-cancer detection with spiral CT and positron emission tomography in heavy smokers: 2-year results, Lancet 362 (9384) (2003) 593–597.
[10] X. Pan, et al., Cost-effectiveness of volume computed tomography in lung cancer screening: a cohort simulation based on nelson study outcomes, J. Med. Econ. 27 (1) (2024) 27–38.
[11] E. Svoboda, Artificial intelligence is improving the detection of lung cancer, Nature 587 (7834) (2020).
[12] K. Sathyakumar, M. Munoz, J. Singh, N. Hussain, B.A. Babu, B. Babu, Automated lung cancer detection using artificial intelligence (AI) deep convolutional neural networks: a narrative literature review, Cureus 12 (8) (2020).
[13] K. Xu, et al., AI body composition in lung cancer screening: added value beyond lung cancer detection, Radiology 308 (1) (2023).
[14] E.S.N. Joshua, M. Chakkravarthy, D. Bhattacharyya, An extensive review on lung cancer detection using machine learning techniques: a systematic study, Rev. Intelligence Artif. 34 (3) (2020).
[15] G. Chassagnon, et al., Artificial intelligence in lung cancer: current applications and perspectives, Jpn. J. Radiol. 41 (3) (2023) 235–244.
[16] C. de Margerie-Mellon, G. Chassagnon, Artificial intelligence: a critical review of applications for lung nodule and lung cancer, Diagnostic Interventional Imaging 104 (1) (2023) 11–17.
[17] S.H. Hosseini, R. Monsefi, S. Shadroo, Deep Learning Applications for Lung Cancer Diagnosis: a Systematic Review, Multimedia Tools Applications, 2023, pp. 1–31.
[18] M. Huang, J. Zou, Y. Zhang, U.A. Bhatti, J. Chen, Efficient click-based interactive segmentation for medical image with improved Plain-ViT, IEEE J. Biomed. Health Inf. (2024) 1–12.
[19] M. Huang, X.S. Zhang, U.A. Bhatti, Y. Wu, Y. Zhang, Y.Y. Ghadi, An interpretable approach using hybrid graph networks and explainable AI for intelligent diagnosis recommendations in chronic disease care, Biomed. Signal Process Control 91 (2024) 105913.
[20] A.H. Nizamani, Z. Chen, A.A. Nizamani, U.A. Bhatti, Advance brain tumor segmentation using feature fusion methods with deep U-Net model with CNN for MRI data, Journal of King Saud University-Computer Information Sciences 35 (9) (2023) 101793.
[21] M. Cellina, et al., Artificial intelligence in lung cancer screening: the future is now, Cancers 15 (17) (2023) 4344.
[22] U.A. Bhatti, et al., MFFCG–Multi feature fusion for hyperspectral image classification using graph attention network, Expert Syst. Appl. 229 (2023) 120496.
[23] U.A. Bhatti, H. Tang, G. Wu, S. Marjan, A. Hussain, Deep learning with graph convolutional networks: an overview and latest applications in computational intelligence, Int. J. Intell. Syst. 2023 (1) (2023) 8342104.
[24] N. Venkatesan, S. Pasupathy, B. Gobinathan, An efficient lung cancer detection using optimal SVM and improved weight based beetle swarm optimization, Biomed. Signal Process Control 88 (2024) 105373.
[25] S. Nageswaran, et al., Lung cancer classification and prediction using machine learning and image processing, BioMed Res. Int. 2022 (2022).
[26] A.K. Swain, A. Swetapadma, J.K. Rout, B.K. Balabantaray, Classification of non-small cell lung cancer types using sparse deep neural network features, in: Biomedical Signal Processing Control, 87, 2024 105485.
[27] I. Naseer, T. Masood, S. Akram, A. Jaffar, M. Rashid, M.A. Iqbal, Lung cancer detection using modified AlexNet architecture and support vector machine, Comput. Mater. Continua (CMC) 74 (1) (2023).
[28] R. Pandian, V. Vedanarayanan, D.R. Kumar, R. Rajakumar, Detection and classification of lung cancer using CNN and Google net, Measurement: Sensors 24 (2022) 100588.
[29] A.A. Abd Al-Ameer, G.A. Hussien, H.A. Al Ameri, Lung cancer detection using image processing and deep learning, Indones. J. Electr. Eng. Comput. Sci 28 (2) (2022) 987–993.
[30] S.A. Agnes, A.A. Solomon, K. Karthick, Wavelet U-Net++ for accurate lung nodule segmentation in CT scans: improving early detection and diagnosis of lung cancer, Biomed. Signal Process Control 87 (2024) 105509.
[31] E.A. Siddiqui, V. Chaurasia, M. Shandilya, Detection and classification of lung cancer computed tomography images using a novel improved deep belief network with Gabor filters, Chemometr. Intell. Lab. Syst. 235 (2023) 104763.
[32] S. Tyagi, S.N. Talbar, LCSCNet: a multi-level approach for lung cancer stage classification using 3D dense convolutional neural networks with concurrent squeeze-and-excitation module, Biomed. Signal Process Control 80 (2023) 104391.

[33] A. Gudur, V. Patil, A. Jain, N. Garg, Hybrid genetic algorithm and deep learning approach for lung nodule detection and classification in chest X-rays, International Journal of Intelligent Systems Applications in Engineering 12 (3s) (2024) 577–587.
[34] K.J. Devi, S. Sudha, A novel panoptic segmentation model for lung tumor prediction using deep learning approaches, Soft Comput. (2024) 1–12.
[35] S. Poonkodi, M. Kanchana, Lung cancer segmentation from CT scan images using modified mayfly optimization and particle swarm optimization algorithm, Multimed. Tool. Appl. 83 (2) (2024) 3567–3584.
[36] R. Mothkur, B. Veerappa, Classification of lung cancer using lightweight deep neural networks, Procedia Computer Science 218 (2023) 1869–1877.
[37] C.-J. Lin, T.-Y. Yang, A fusion-based convolutional fuzzy neural network for lung cancer classification, Int. J. Fuzzy Syst. 25 (2) (2023) 451–467.
[38] A.A. Shah, H.A.M. Malik, A. Muhammad, A. Alourani, Z.A. Butt, Deep learning ensemble 2D CNN approach towards the detection of lung cancer, Sci. Rep. 13 (1) (2023) 2987.
[39] R. Sun, Y. Pang, W. Li, Efficient lung cancer image classification and segmentation algorithm based on an improved Swin transformer, Electronics 12 (4) (2023) 1024.
[40] S. Gite, A. Mishra, K. Kotecha, Enhanced lung image segmentation using deep learning, Neural Comput. Appl. 35 (31) (2023) 22839–22853.
[41] B.R. Pandit, et al., Deep learning neural network for lung cancer classification: enhanced optimization function, Multimed. Tool. Appl. 82 (5) (2023) 6605–6624.
[42] G. Aceto, D. Ciuonzo, A. Montieri, A. Pescapé, Toward effective mobile encrypted traffic classification through deep learning, Neurocomputing 409 (2020) 306–315.
[43] G.E. Hinton, S. Osindero, Y.-W. Teh, A fast learning algorithm for deep belief nets, Neural Comput. 18 (7) (2006) 1527–1554.
[44] G. Aceto, D. Ciuonzo, A. Montieri, A. Pescapè, MIMETIC: mobile encrypted traffic classification using multimodal deep learning, Comput. Network. 165 (2019) 106944.
[45] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[46] S. Woo, J. Park, J.-Y. Lee, I.S. Kweon, Cbam: convolutional block attention module, in: Proceedings of the European Conference on Computer Vision, ECCV), 2018, pp. 3–19.
[47] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, 2014 *arXiv preprint arXiv:1409.1556*.
[48] R. Zhu, et al., ScratchDet: training single-shot object detectors from scratch, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 2268–2277.
[49] E.S. Neal Joshua, D. Bhattacharyya, M. Chakkravarthy, Y.-C. Byun, 3D CNN with visual insights for early detection of lung cancer using gradient-weighted class activation, Journal of Healthcare Engineering 2021 (2021) 1–11.
[50] Lung CT Nodule/Lesion Segmenbtation, 2023.
[51] Chest CT-Scan Images Dataset, 2020. Available:.
[52] C.C. Nguyen, G.S. Tran, J.-C. Burie, T.P. Nghiem, Pulmonary nodule detection based on faster R-CNN with adaptive anchor box, IEEE Access 9 (2021) 154740–154751.
[53] J. Redmon, A. Farhadi, Yolov3: an Incremental Improvement, 2018 *arXiv preprint arXiv:.02767*.
[54] L. Goel, S. Mishra, A hybrid of modified YOLOv3 with BBO/EE optimizer for lung cancer detection, Multimed. Tool. Appl. (2023) 1–33.
[55] L. Goel, P. Patel, Improving YOLOv6 using advanced PSO optimizer for weight selection in lung cancer detection and classification, Multimed. Tool. Appl. (2024) 1–34.
[56] H. Cao, et al., Swin-unet: unet-like pure transformer for medical image segmentation, in: European Conference on Computer Vision, Springer, 2022, pp. 205–218.
[57] W. Albattah, A. Javed, M. Nawaz, S. Albahli, Artificial intelligence-based drone system for multiclass plant disease detection using an improved efficient convolutional neural network, Front. Plant Sci. 13 (2022) 808380.
[58] A.O. Adedoja, P.A. Owolawi, T. Mapayi, C. Tu, Intelligent mobile plant disease diagnostic system using NASNet-mobile deep learning, IAENG Int. J. Comput. Sci. 49 (1) (2022) 216–231.
[59] M. Chhabra, R. Kumar, A smart healthcare system based on classifier DenseNet 121 model to detect multiple diseases, in: Mobile Radio Communications and 5G Networks: Proceedings of Second MRCN 2021, Springer, 2022, pp. 297–312.
[60] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L.-C. Chen, Mobilenetv2: inverted residuals and linear bottlenecks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 4510–4520.
[61] H.-Y. Chen, C.-Y. Su, An enhanced hybrid MobileNet, in: 2018 9th International Conference on Awareness Science and Technology (iCAST), IEEE, 2018, pp. 308–312.
[62] Z. UrRehman, et al., Effective lung nodule detection using deep CNN with dual attention mechanisms, Sci. Rep. 14 (1) (2024) 3934.
[63] C. Thangavel, J. Palanichamy, Effective deep learning approach for segmentation of pulmonary cancer in thoracic CT image, Biomed. Signal Process Control 89 (2024) 105804.
[64] I. Naseer, et al., Lung cancer detection using modified AlexNet architecture and support vector machine 74 (1) (2023).
[65] A.A. Alsheikhy, Y. Said, T. Shawly, A.K. Alzahrani, H. Lahza, A CAD system for lung cancer detection using hybrid deep learning techniques, Diagnostics 13 (6) (2023) 1174.