Cognitive Computation (2025) 17:67
https://doi.org/10.1007/s12559-025-10408-2

REVIEW

Deep Learning Innovations in the Detection of Lung Cancer: Advances, Trends, and Open Challenges

Helena Liz-López[1] · Áurea Anguera de Sojo-Hernández[1] · Sergio D'Antonio-Maceiras[1] · Miguel Angel Díaz-Martínez[1] · David Camacho[1]

Received: 11 July 2024 / Accepted: 14 January 2025 / Published online: 17 February 2025

Abstract
Cancer is the second leading cause of death worldwide, and within this type of disease, lung cancer is the second most diagnosed, but the leading cause of death. Early detection is crucial to increase patient survival rates. One of the primary methods for detecting this disease is through medical imaging, which, due to its features, is well-suited for analysis by deep learning techniques. These techniques have demonstrated exceptional results in similar tasks. Therefore, this paper focusses on analyzing the latest work related to lung cancer detection using deep learning, providing a clear overview of the state of the art and the most common research directions pursued by researchers. We have reviewed DL techniques for lung cancer detection between 2018 and 2023, analyzing the different datasets that have been used in this domain and providing an analysis between the different investigations. In this state-of-the-art review, we describe the main datasets used in this field and the primary deep learning techniques used to detect radiological signs, predominantly convolutional neural networks (CNNs). As the impact of these systems in medicine can pose risks to patients, we also examine the extent to which explainable AI techniques have been applied to enhance the understanding of these systems, a crucial aspect for their real-world application. Finally, we will discuss the trends that the domain is expected to follow in the coming years and the challenges that researchers will need to address.

Keywords Lung cancer · Medicine · Medical imaging · Detection · Deep learning · Convolutional neural networks

Introduction

Cancer, the diverse group of diseases characterized by abnormal cell growth, is the second most common cause of death globally, accounting for 10 million deaths in 2020, representing one in six deaths. Following a recent report by the World Health Organization (WHO), this number is going to exponentially grow; by 2050, over 35 million new cancer cases are anticipated, marking a 77% rise from the estimated 20 million cases in 2022. This sharp increase in the global cancer burden is driven by population aging, population growth, and shifts in exposure to risk factors, many of which are linked to socioeconomic development [154]. In terms of cancer types, breast cancer has historically been the most frequently diagnosed; however, from 2022, lung cancer was the most commonly diagnosed cancer, representing 12.4% of new cases, and the leading cause of cancer-related deaths, accounting for 18.7% of fatalities. Moreover, by 2050, it is expected to remain the primary cause of cancer incidence (13.1% of new cases) and cancer mortality (19.2% of cancer deaths) [1]. Although 80% of cases for this type of cancer are caused by smoking and second-hand smoke, there is an increasing number of cases in people who have never smoked [155], showing that, even if smoking levels are reduced, lung cancer will continue to be present due to other factors [4].

✉ Helena Liz-López
helena.liz@upm.es

Áurea Anguera de Sojo-Hernández
aureamaria.angueradesojo@upm.es

Sergio D'Antonio-Maceiras
sergio.dantonio@upm.es

Miguel Angel Díaz-Martínez
m.diaz@upm.es

David Camacho
david.camacho@upm.es

[1] Computer Systems Department, Universidad Politécnica de Madrid, Madrid 28031, Spain

One of the most important aspects of reducing mortality is the early detection of the disease, which can make a huge difference between life and death [3], and here is where the use of computer science has been shown to be helpful to increase the speed and accuracy of diagnostics. Conventional medical techniques used for early detection are image-related; these include computed tomography (CT), magnetic resonance imaging (MRI), and chest radiographs (chest X-rays) [5]. In this area, deep learning (DL) algorithms can play a key role [6]. Over the last few years, major breakthroughs have been made in the field of DL for solving various visual-related tasks [7, 8]. Among these, convolutional neural networks (CNNs)—a type of artificial neural network inspired by biological neurons in the visual cortex [9–11]—and recurrent neural networks (RNNs) have shown practicality and efficiency in identifying and categorizing various diseases [12–14]. More recently, accuracy in this field has increased through the use of ensemble models [15–17]. This methodology integrates multiple learning algorithms to enhance the effectiveness of each individual component by merging their individual predictions to yield a consensus result. Advantages of this approach over single models include improved performance, as the combination of multiple models enhances their individual capabilities, thus allowing the final model to approximate the optimal solution more accurately [18–20]. Additionally, it offers robustness by reducing the variance of prediction errors generated by the contributing models and introducing bias, which helps prevent overfitting in the final model [21].

However, CNNs, much like numerous other deep learning and machine learning approaches, are often perceived as a "black-box" algorithm [22]. While users can easily analyze and interpret both the input and output, the inference process conducted by the algorithm remains opaque, thus eroding end-users' confidence in the results and adversely affecting decision-making [23]. This opaqueness renders the essential process of understanding "how" and "why" the algorithm arrived at a particular outcome inscrutable to humans [24]. This limitation can hinder its utility in fields such as medicine, where practitioners need information on how the algorithm reached a specific diagnosis for each patient [25]. This limitation can be addressed by implementing automatic explainable AI (XAI) systems [26], which help mitigate these issues [27]. However, the integration of DL models with medical knowledge facilitates the creation of novel clinical decision support systems (CDSS). Throughout this article, we will explore the differences between self-explanatory models, such as models focused on localization tasks, and classification models, which require the application of XAI techniques that allow specialists to understand the areas of imaging evidence most relevant to the DL system.

Decision support systems of this type have been validated in medical centers to test their performance and their ability to be applied in real-life situations. For example, a study [28] demonstrated that the combination of DL with medical expertise has proven to reduce the error rate from 3.5% in humans and 2.9% in DL models to a noteworthy 0.5%, signifying a 99.5% accuracy in correctly classifying cases. Another study by [29] analyzed the results of the system with senior and resident physicians. Senior physicians agreed with the system 70% of the time, and all residents improved their performance using the tool. For these reasons, these systems have been used in the medical field for years, yielding increasingly promising results [30].

This is not the first work to perform a literature review on this topic. A search performed in December 2023 in the prestigious *Web of Science* (WOS) Core Collection database gave three results, summarized in Table 1.

Table 1 Comparison of previous literature reviews

Authors	Year	Databases covered	Date range	Methodology	# Docs
Liu et al. [31]	2020	arXiv, Cochrane, Elsevier, Google Scholar, IEEE, MICCAI, PubMed, RSNA, SPIE Medical Imaging, Springer, and WOS	1988 - 2020	Not specified	21
Jassim and Jaber [32]	2022	Science direct, Scopus, WOS, and IEEE	2017-2021	PRISMA	55
Hosseini et al. [34]	2023	IEEE Xplore; ScienceDirect; Springer Link; Wiley Inter Science; Google Scholar	2016 to 2021	Own methodology based on SMS and SLR	32

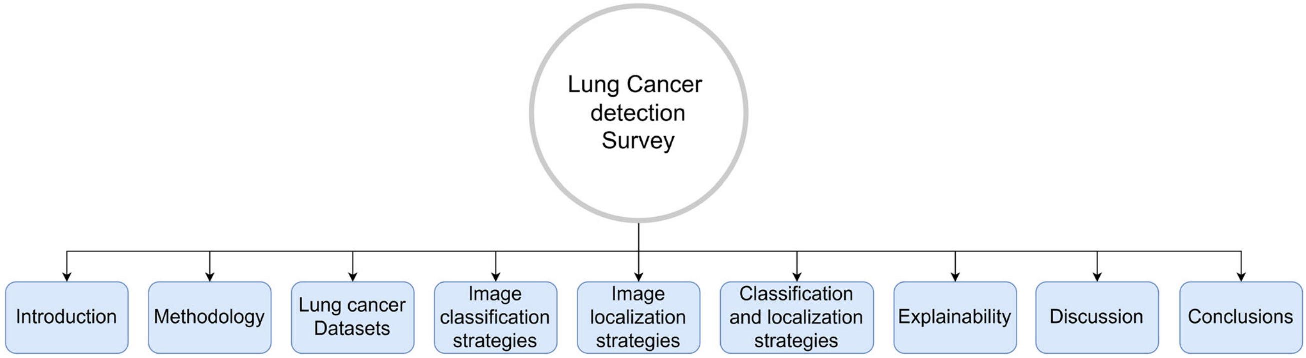

Fig. 1 Schematic representation of the structure of the survey

In 2020, Liu et al. [31] performed the first literature review analyzing documents from 1988 until 2020 searching in the most relevant databases. Their main goal was to conduct the first comprehensive review of computer-assisted detection and classification techniques, and although there have been several improvements in recent years, the challenges and opportunities stressed in this work are still valid. However, unfortunately in this work, the methodology is not specified, and therefore, the inclusion/exclusion criteria of the documents are not fully clear. The chapter dedicated to DL includes only eight documents.

The rapid growth of this topic is reflected in the difference from the second work. In 2022, Jassim and Jaber [32] also covered the main databases from 2017 to 2021. In this case, the authors followed the PRISMA guidelines [33] and analyzed a total of 55 documents focused on lung cancer detection and classification using ML and DL methods. Here, the authors also gathered the main proposals for future works and found that lack of data was the most common one, followed by the need for end-to-end explainable systems that integrate detection, segmentation, and prediction.

One year later, Hosseini et al. [34] conducted the first review that focused exclusively on DL methods. This work analyzed a total of 32 documents from 2016 to 2021 from various databases, but not from WOS. Excluding this prestigious database could be one reason for finding that low amount of documents; another one can be the strict search string that only looks for keywords in the titles Finally, while the methodology used is inspired by the systematic mapping study (SMS) and the systematic literature review (SLR), the steps are split along the paper and are not the same as PRISMA. To the best of our knowledge, our survey is the first systematic literature review focused on DL methods to tackle lung cancer that strictly follows the guidelines of the PRISMA methodology, is focused in the WoS database, and analyzes a larger number of papers and datasets used.

This review's contributions are organized around four formulated research questions.

- **RQ1**: What datasets are publicly available and accessible to researchers?
- **RQ2**: What main tasks or topics are investigated within this area of study?
- **RQ3**: Which DL algorithms are most commonly used by researchers?
- **RQ4**: Are explainability techniques applied to facilitate the application of these techniques in real-life situations?

The primary contribution of this article emerges from addressing the research questions and can be summarized as follows:

- It provides an updated picture of the status of lung cancer detection systems using DL techniques.
- It provides a list and analysis of available datasets that can be used by future researchers. These datasets may be of different types of imaging tests.
- The different tasks within this field of study are shown clearly.
- It shows a review of the main DL techniques for lung cancer detection in imaging.
- It describes the trends that are expected in the coming years and the challenges that researchers will have to face in the area of study.

Finally, this manuscript is organized as follows (see Fig. 1): The "Methodology" section explains the methodol-

Table 2 Documents selected in each screening phase

Phase	# articles
Search database	853
Article selection	571
Title/abstract screening	265
Article's content	115
Total	**89**

ogy used in this review and the selection criteria used to filter the articles. The "Lung Cancer Datasets" section describes publicly available datasets related to lung cancer. The "Image Classification Strategies" section discusses the various articles focused on the classification of medical imaging. The "Lung Nodule Localization Techniques" section explores the different articles that focus on the localization in medical imaging. The "Classification and Localization" describes work that combines classification and localization to detect lung cancer. The "Explainability in Lung Cancer Detection" section demonstrates the application level of XAI techniques within the different detection methods. The "Discussion" section addresses the research questions and highlights future trends and challenges associated with this topic. The "Conclusion" section presents the conclusions drawn from this review of the state of the art in lung cancer detection.

Methodology

In this section, we will describe the systematic process used to create this survey of articles related to DL techniques for the diagnosis and detection of lung cancer, based on the PRISMA protocol [33]. The process included retrieving articles from a prominent scientific database, Web of Science (WOS). The steps for manually selecting and reviewing the items are outlined below, with a summary of the different phases provided in Table 2 and Fig. 2.

1. *Search in databases*: This article focuses on the detection and diagnosis of cancer using DL techniques. Therefore, the thesaurus terms finally included in the research are: "Lung cancer" AND "Deep Learning" AND "Detection."
2. *First screening*: Second, only scientific articles in the field of computer science related to the detection or diagnosis of cancer using DL techniques, published between 2018 and 2023 (up to May), were selected. Repeated items were also removed in this step.
3. *Second screening*: Titles and abstracts were used to exclude articles that did not fit the topic of this survey, based on the following criteria:
   - The article must be written in English.
   - The article has to apply DL techniques to problems of cancer diagnosis or detection.
   - The methodology should provide a clear and complete description of all the details essential for its implementation, and the results should be clearly and accurately presented. Articles in which the methodology did not describe all the steps followed, the dataset was not specified, and the results were not presented or incomplete and were discarded.
   - Due to the scope of this review, and in conjunction with the previous criteria, pure qualitative research was discarded; only papers with quantitative results that could be compared in terms of metrics were included.
4. *Third screening*: A more exhaustive review of the articles was conducted, with a meticulous examination of the content of each document. Subsequently, articles that did not meet the specified criteria were excluded.
5. *Analysis of the selected articles*: The final step involved analyzing and comparing the information extracted from the selected articles on the diagnosis or detection of lung cancer.

Lung Cancer Datasets

One of the most critical issues in the development and training of deep learning models is the availability of high-quality data, particularly medical images. For developing diagnostic support systems, high-quality datasets with sufficient

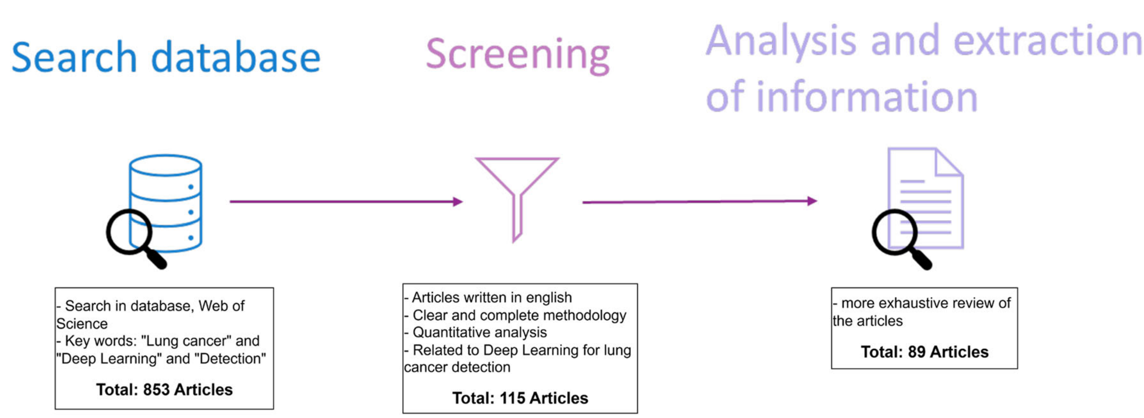

Fig. 2 Visual representation of methodology

variability are essential to train a robust model capable of performing adequately in real situations [35–38].

Within this domain, we can identify different types of medical images for lung cancer detection. The most common types used in the works analyzed are CT, X-ray, and histological images. CT (see Fig. 3) is a diagnostic imaging technique that uses X-rays to create detailed cross-sectional images of the body, enabling the visualization of internal structures in 3D. X-ray (see Fig. 4) is an imaging method that uses X-rays to view internal structures of the body in two dimensions, where dense tissues, such as bone, appear more prominently. Lastly, histological imaging (see Fig. 5) involves microscopic images of biological tissues that have been thinly sliced and stained to enhance structural details at the cellular level. Therefore, in this section, we will organize the available datasets according to the type of medical tests: CT, chest X-rays, histological images, and a combination of different tests.

Due to the protection of patients, the management of these data is a sensitive issue. Therefore, various organizations provide access to different image banks, which are crucial for researchers and experts. These image banks come from both hospitals and societies. In the context of lung cancer, there are several public datasets with images obtained by the aforementioned techniques (see Table 3).

CT

In this section, we will focus on datasets composed solely of CT samples. Computed tomography is a diagnostic imaging technique that uses X-rays and advanced computer processing to obtain detailed, cross-sectional images of the inside of the body. Unlike conventional X-rays, which produce two-dimensional images, CT produces three-dimensional images by capturing multiple cross-sectional slices from different angles. This method allows organs, tissues, bones, and blood vessels to be observed with great precision, facilitating the detection and evaluation of diseases, injuries, and abnormalities. It is a widely used tool in medicine because of its ability to provide detailed information quickly and non-invasively. In the following, we will describe each one, highlighting its properties.

Fig. 3 CT example from LUNA-16 dataset [39]

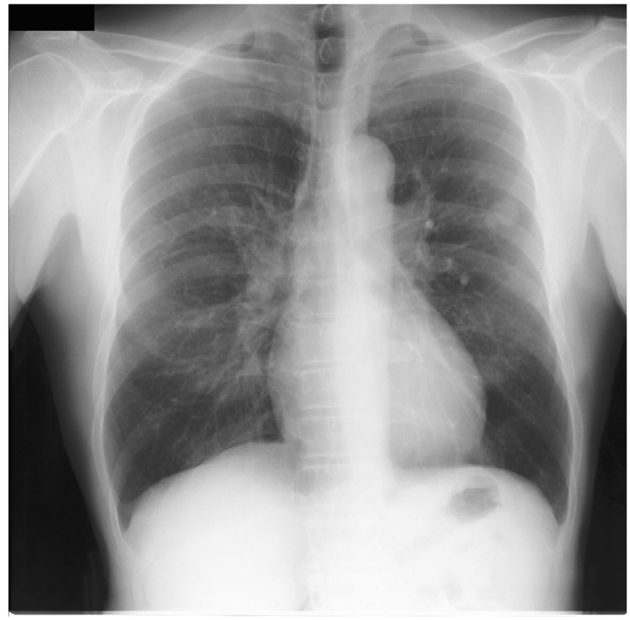

Fig. 4 Chest X-ray example from LUNA-16 dataset [39]

Automatic Nodule Detection 2009 (ANODE09) This database was part of a challenge in 2009 and is no longer available [50]. The database consisted of 55 CT scans from a lung cancer screening program, accompanied by a web-based framework

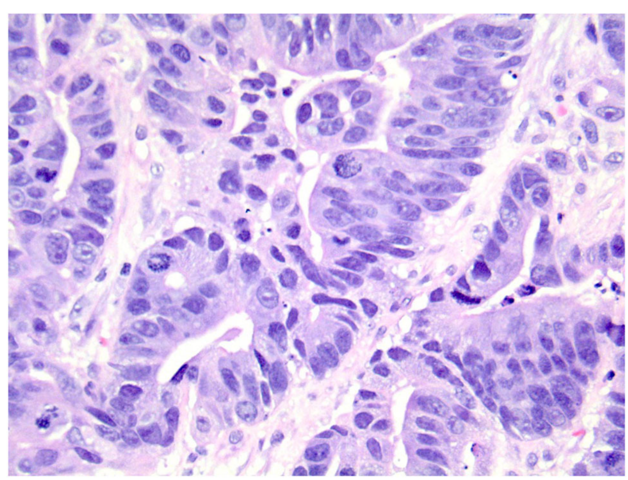

Fig. 5 Histological image example from LC25000 dataset [40]

Table 3 Datasets used in the analyzed works

Type of medical image	Dataset	Year	# images total	# Samples	# patients
CT	ANODE09 [41]	2011	22.000	55	–
	LIDC-IDRI [41]	2011	244.527	1.018	1010
	LUNA-16 [39]	2016	≈ 15.419	–	32
	DeepLesion [48]	2017	32,120	10.594	4.427
	LNDb [49]	2019	≈ 129.654	294	294
	LNPE1000 [51]	2022	≈ 400.000	1.000	1.000
	Chest CT-Scan images Dataset [54]	2020	–	1.000	–
	DSB [53]	2017	–	1.000	–
X-rays	RIDER [44]	2015	15.419	46	32
	JSRT [47]	2000	247	247	–
Histological images	LC25000 [40]	2019	25.000	–	–
Genetic data + CT	CIA [43]	2013	–	–	–
Genetic data + Histological images	TCGA-LUAD [45]	2016	48.931	–	69
X-ray + CT + Histological images	CPTAC-LSCC [46]	2018	53.100	–	248
CT + Histological images	NLST [52]	2011	–	54.000	54.000

for the objective evaluation of nodule detection algorithms. To address the issue of observer disagreement on nodule classification, the authors introduced a system that distinguishes between relevant and irrelevant findings. Irrelevant findings include nodules that are unlikely to be cancerous, such as calcified or very small nodules.

Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) This is one of the most widely used datasets in this area [41]. It consists of diagnostic and lung cancer screening thoracic computed tomography (CT) scans with markedup annotated lesions from 1018 cases. Each subject includes images from a clinical thoracic CT scan and an associated XML file that records the results of a two-phase image annotation process conducted by four experienced thoracic radiologists. In the initial blindedread phase, each radiologist independently reviewed each CT scan and marked lesions in one of three categories: nodule $\geq$3 mm, nodule $<$3 mm, and nonnodule $\geq$3 mm. In the subsequent unblindedread phase, each radiologist independently reviewed their own marks along with the anonymized marks of the three other radiologists to render a final opinion. The goal of this process was to identify as comprehensively as possible all lung nodules in each CT scan without requiring a forced consensus.

Lung Nodule Analysis (LUNA-16) This is a subset of the LIDC-IDRI database designed for a challenge and is also among the most widely used datasets [39]. It includes only the nodules with size >= 3 mm and accepted by at least 3 out of 4 radiologists, considering the rest as irrelevant. This selection criteria resulted in a total of 1186 lung nodules annotated in 888 CT scans.

DeepLesion (also Known as NIH) This is derived from over 32,000 bookmarked lesions on CT images of 4400 unique patients, and this dataset was created using these bookmarks and extensive retrospective medical data [48]. Unlike other medical image datasets focused on a single lesion type, DeepLesion encompasses a wide range of critical radiological findings across the body, including lung nodules, liver tumors, and enlarged lymph nodes.

Lung Nodule Database (LNDb) This includes 294 retrospectively collected CT scans from the Centro Hospitalar e Universitário de São João (CHUSJ) in Porto, Portugal, between 2016 and 2018 [49]. All data acquisition received approval from the CHUSJ Ethical Committee and underwent anonymization to remove personal information, retaining only the patient's birth year and gender. These scans were not specifically acquired for the LNDb project, and the inclusion criteria of the LIDC-IDRI dataset were followed.

Lung Nodules from Physical Examination-1000 (LNPE1000) It was collected from North China Petroleum General Hospital between August 2017 and March 2018 [51]; this comprises CT scans with at least one pulmonary nodule per sample, annotated by five experienced radiologists. The annotation process involved three rounds of evaluation, resulting in 4566 nodules being selected as ground truth, which included 3037 intrapulmonary nodules and 1529 pleural nodules. Nodules marked by fewer than three doctors were considered irrelevant findings. Additionally, the dataset includes 369 GGO nodules and 431 non-GGO nodules, aiding in the identification of these kinds of nodules. Notably, the nodule diameter distribution is distinct from the LUNA16 dataset, with 47% of nodules being less than 5 mm in diameter and 75.5% less than 6 mm, offering valuable data for pulmonary nodule research in CT images, particularly for nodule detection and characterization.

Kaggle It is a widely used platform for sharing all kinds of ML resources, including datasets and models. It is also well known for hosting several "Competitions" to solve concrete tasks. There have been several competitions related with lung cancer, and the list of datasets used is the following:

- ***Data Science Bowl 2017 (DSB)*** [53]: This dataset provided to each participant more than 1000 low-dose CT images taken from high-risk patients. Each image was a series comprising multiple axial slices of the chest cavity. The number of 2D slices per image varied, influenced by the scanning machine and patient specifics. Finally, all images included essential details, such as patient ID, and scan parameters like slice thickness.
- ***Chest CT-Scan Images Dataset*** [54]: This dataset is curated to detect chest cancer using ML and DL (CNN) techniques. It is composed of 1000 CT scans divided into training (70%), testing (20%), and validation (10%). The author also provides information about the type of cancer and the method of treatment and explains that the dataset was built using different, unspecified sources.

X-rays

In this section, we will focus on chest X-ray datasets, which, although less relevant than CT datasets, are still one of the most widely used radiological tests. X-rays are medical images obtained through the use of X-rays, a type of high-energy electromagnetic radiation that can pass through the human body. These images allow internal structures, such as bones, organs, and tissues, to be viewed in different shades depending on their density; bones appear lighter because they absorb more X-rays, while soft tissues appear darker. X-rays are widely used in medicine to diagnose fractures, infections, tumors, or other health conditions and are a fundamental tool in the field of radiology, due to their speed, simplicity, and lack of preparation.

Japan Society of Radiological Technology Dataset (JSRT) The database includes 154 conventional chest radiographs [47]: 100 with malignant nodules and 54 with benign lung nodules, along with 93 radiographs without nodules. It contains demographic data for the patient (age, sex), diagnostic results (malignant or benign), nodule coordinates (X and Y), and a basic diagram indicating the location of the nodule. Lung nodule images are categorized into five groups based on degrees of subtlety.

Histological Images

Another test used for the diagnosis of lung cancer is histological imaging. These are visual representations of biological tissues obtained through the use of microscopes, after a preparation process that includes fixation, cutting, and staining of the tissue. These images allow the organization, structure, and cellular composition of tissues to be observed, which is essential for studying their morphology and detecting alterations associated with diseases, such as cancers, inflammation, or infections. Histological imaging is a key tool in pathology and biology, providing essential information for diagnosis, biomedical research, and the development of medical treatments. Although unique sets of histological images are rare, several datasets have been found that combine them with other types of diagnostic tests.

Lung and Colon Cancer Histopathological Image Dataset (LC25000) This dataset contains 25,000 color images, distributed evenly across five classes [40]. These classes represent specific histologic entities: colon adenocarcinoma, benign colonic tissue, lung adenocarcinoma, lung squamous cell carcinoma, and benign lung tissue. To ensure privacy and compliance, all images were de-identified and are HIPAA-compliant.

Combination of Different Image Tests

Lastly, we will analyze datasets containing different diagnostic tests, such as CT scans, chest X-rays, genetic data, or histological images.

The Cancer Imaging Archive Dataset (CIA) This is a database organized into collections that is constantly updated with different types of images and lesions [43]. It also includes images from the LIDC-IDRI dataset. At the time of this writing, there was information from 7692 patients for lung cancer.

Several works analyzed use this dataset, and the complete list of specific datasets from this collection is the following:

- ***Image Database to Evaluate Therapy Response (RIDER)*** [44]: This dataset was developed during a study focused on assessing the variability of tumor measurements (unidimensional, bidimensional, and volumetric) on same-day repeat CT scans in patients with non-small cell lung cancer. This targeted data collection aimed to establish a consensus on standardizing data collection and analysis for quantitative imaging methods. The dataset consists of CT images and includes 32 participants with a total of 15,419 images.
- ***Cancer Genome Atlas Lung Adenocarcinoma (TCGA-LUAD)*** [45]: The Lung Adenocarcinoma dataset is part of an initiative that links cancer characteristics with genetic data. It includes clinical images from The Cancer Genome Atlas (TCGA) subjects, specifically CT, PET (positron emission tomography), NM (nuclear medicine), and histopathology images. Clinical, genetic, and pathological data are available in the Genomic Data Commons (GDC) Data Portal, while radiological data is housed in the Cancer Imaging Archive (CIA). Researchers can use matched patient identifiers to explore correlations between tissue genotypes, radiological features, and patient outcomes. Tissue samples were collected globally to meet specimen targets, resulting in a diverse set of image data, most of which were acquired as part of routine clinical care. The dataset includes 69 participants, 152 studies, 624 series, and 48,931 images.
- ***Clinical Proteomic Tumor Analysis Consortium Lung Squamous Cell Carcinoma Collection (CPTAC-LSCC)*** [46]: The dataset is composed of 248 subjects and a total of 53,100 images of radiology and pathology (CT, PT, histopathology). These images are part of an initiative by The Cancer Imaging Archive to provide researchers with publicly accessible data to explore potential correlations between cancer phenotypes, proteomic profiles, genomic information, and clinical data in the context of lung squamous cell carcinoma (LSCC).

The National Lung Screening Trial (NLST) The data was obtained by recruiting 53,456 participants aged between 55 and 74 from various medical centers across the United States [52]. These participants had a significant history of smoking but no present or past indications of lung cancer. Random assignment was implemented for participants to receive one of two screening protocols annually for 3 years: low-dose helical computed tomography (CT) or conventional chest X-ray screenings.

When developing a model, the researcher chooses an image dataset, which is used for training the neurons and subsequently testing to determine the performance of the model itself.

Image Classification Strategies

First, we will analyze articles that perform classification tasks for lung cancer detection in medical imaging. In this case, the various papers focus on classifying the samples as either "lung cancer" or normal, or distinguishing between different types of lung cancer. These techniques have shown promising results in a number of similar problems, such as detecting pneumonia or multiple diseases in medical images [58–61]. The papers have been grouped into four different research lines (see Fig. 6): (a) development and enhancement of CAD (computer-aided detection) and CADx (computer-aided diagnosis) systems, (b) innovations in neural network architectures and designs, (c) segmentation-based systems, and (d) strategies to improve accuracy and reduce false positives.

Improvements in Computer-Aided Detection and Computer-Aided Diagnosis

Advances in computer-assisted detection (CAD) and computer-assisted diagnosis (CADx) systems have been pivotal in leveraging deep learning for medical imaging. This subsec-

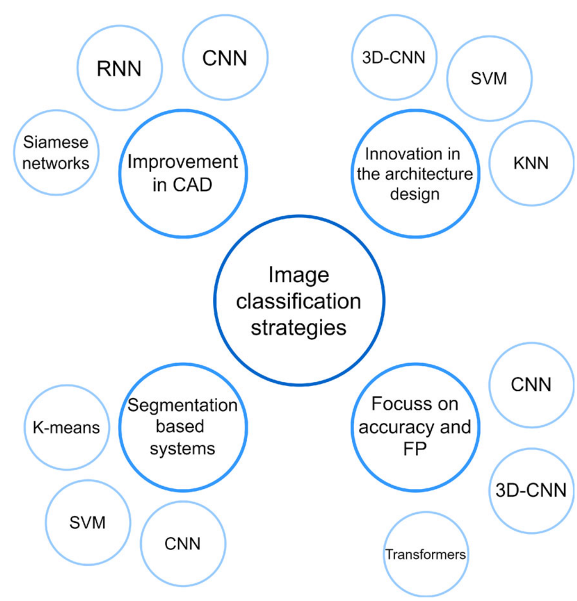

Fig. 6 Visual representation of main techniques for classification tasks

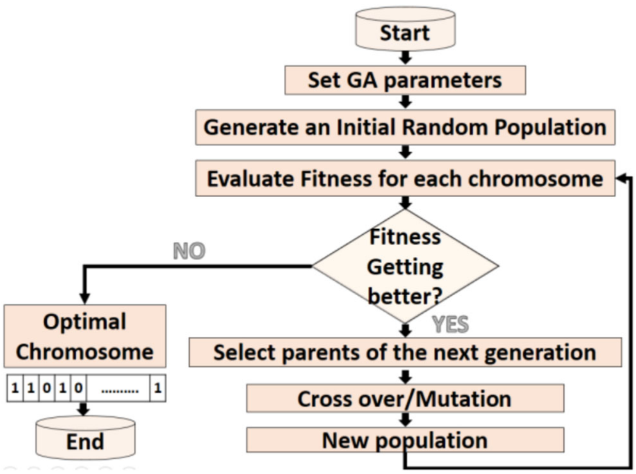

Fig. 7 Genetic algorithm used by Elnakib et al. [62]

tion explores various innovative approaches in this domain, emphasizing the integration of sophisticated convolutional neural network (CNN) architectures and learning techniques to enhance detection and classification accuracy.

Some authors choose to use CNNs trained with generic datasets, which are widely used in the state of the art. For example, Elnakib et al. [62] developed a CADe system using architectures like AlexNet, VGG16, and VGG19, where genetic algorithms played a key role in selecting vital features, as shown in Fig. 7, leading to high accuracy, sensitivity, and specificity. Alsheikhy et al. [63] combined the VGG-19 DCNN technique with LSTM networks, achieving remarkable accuracy and precision metrics. Their methodology included preprocessing steps like denoising and normalization, which are crucial for feature extraction in CAD systems.

Wang and Charkborty [64] propose a system capable of detecting nodules and determining their risk. For detection, they employ a 3D CNN network and RNN. For the assessment of cancer risk, several 3D-CNN models were trained to evaluate the features of the nodules. In this work, the LIDC and LUNA16 datasets were used, and the CNN network and RNN achieved an AUC value of 0.78.

Further enhancing nodule identification, Zhang and Kong [65] introduced a system with a four-channel CNN model, focusing on larger nodules from the LIDC/IDRI database. This system emphasized reducing false positives and included an innovative lung wall mending mechanism. Rafael-Palou et al. [66] presented a novel approach using a 3D Siamese neural network for nodule identification, demonstrating the system's efficacy without the need for extensive image recordings. Lastly, Manickavasagam et al. [67] introduced the CNN-5CL system for detecting pulmonary nodules. The system, characterized by its 5 convolutional layers, was tested using the LIDC/IDRI dataset and achieved notable accuracy and AUC values.

This section has addressed significant advances in computer-aided detection (CAD) and computer-aided diagnosis (CADx) systems, highlighting the use of advanced convolutional neural network (CNN) architectures and techniques to improve detection and classification accuracy in medical imaging. Techniques discussed include the use of traditional CNNs and more complex variants such as 3D CNNs and RNNs, as well as the integration of genetic algorithms and 3D Siamese networks to optimize feature selection and reduce false positives. Current research focuses on improving diagnostic accuracy, using supervised and recurrent learning algorithms together with preprocessing methods such as normalization and denoising. In Tables 4 and 5, a summary of the different articles and a comparison of these with the datasets used can be found.

Systems Based on Convolutional Neural Networks

This section explores a variety of innovative approaches ranging from 3D convolutional neural networks, residual networks, and capsule-based systems to reinforcement learning algorithms and dense networks, each applied to different datasets with the common goal of improving accuracy, sensitivity, and specificity in the detection of lung abnormalities. The papers presented address both the architectural design of the networks as well as training and validation strategies, highlighting the use of public and private databases to evaluate the performance of these advanced models. This section not only underscores the importance of technological innovation in early medical diagnosis but also illustrates the critical role of CNN in improving health outcomes through accurate and efficient detection of lung diseases. Tables 6 and 7 show

Table 4 Works and datasets focused on the improvement of CAD-CADx systems

	LUNA16	LIDC/IDRI	DSB	Own dataset	Others
Elnakib et al. [62]	X				X
Zhang and Kong [65]		X			
Wang and Charkborty [64]	X	X	X		
Rafael-Palou et al. [66]				X	
Manickavasagam et al. [67]		X			
Alsheikhy et al. [63]	X				X

Table 5 Summary of articles related to improvement of CAD on classification tasks

Reference	Year	# datasets	Methodology	Code available	Main contribution
Elnakib et al. [62]	2020	3	VGG-16/19, AlexNet	X	It uses genetic algorithms to select the most relevant features and changes the classifier of the archaeological sites
Zhang and Kong [65]	2020	1	Multi-Scene Deep Learning Framework (MSDLF)	X	System based on the Multi-Scene Deep Learning Framework (MSDLF) using a vesselness filter and four-channel CNNs
Wang and Charkborty [64]	2021	3	Ensemble of 3D-CNN and RNN, GBoost for regression task	X	Ensemble of heterogeneous models for lung nodule detection and cancer risk evaluation using GBoost
Rafael-Palou et al. [66]	2021	1	3D Siamese Neural Networks	X	Developed a 3D Siamese neural network for efficient pulmonary nodule re-identification and growth detection without image registration
Manickavasagam et al. [67]	2022		CNN from scratch	X	It shows the effectiveness of architectures from scratch in specific situation
Alsheikhy et al. [63]	2023	3	VGG-19 + LSTM	X	It uses CNN for spatial information and LSTM for temporal information

in summary form a comparison between the datasets and the different research and a summary of the methodology and main contributions of each work.

Some authors focus more on pre-processing than on the architectures used. For example, Sori et al. [68] introduced the DFD-Net, a "denoising first" two-path convolutional neural network (CNN) for detection from CT scans. This model first employs a residual learning denoising model (DR-Net) to remove noise, followed by a two-path CNN for cancer detection using the denoised image. The datasets used are from the Data Science Bowl 2017 challenge (DSB) and LUNA 16. Although the study outlined the methodology for evaluating the model using accuracy, recall, and specificity, the specific values or results of these metrics were not detailed in the document. Another example is Shakeel et al. [56], who introduced a model that employed an enhanced deep neural network (IDNN) for image segmentation after applying noise reduction to the images. The dataset used contains more than 5000 CT test images from the CPTAC-LSCC dataset. Regarding the IDNN network, they use a VGG network composed of 19 layers, where 3 of them are fully connected and the remaining 16 are convolutional. The reported metrics achieved were 96.2% accuracy, 98.4% specificity, 97.4% precision, 98% recall, and F1 values 98.4%.

Although it may seem surprising, given the progress in DL techniques for image processing, many authors within the domain of medicine still employ basic methods. For example, Eun et al. [69] developed a framework using 2D neural networks on the LUNA16 dataset, which yielded improved results compared to 3D networks, achieving a performance metric of 0.922. Ali et al. [70] introduced a model using a CNN network to distinguish between pulmonary and non-pulmonary nodules using the same dataset. In this model, 2D images extracted from 3D CT scans were used, with the 80th percentile of the images designated for training and the 20th percentile for testing. The final results included a specificity of 88.1%, a sensitivity of 68.66%, and an accuracy of 83%. Other authors opt to retrain architectures that have been extensively utilized in recent state-of-the-art research, such as ResNet, DenseNet, and Inception, among others. Guo et al. [71] utilized the ResNet-50 convolutional network to identify pulmonary nodules using an in-house dataset of 18 patients, divided into 12 for training, 4 for validation, and 2 for testing, achieving promising results with an AUC parameter exceeding 0.98. Lanjewar et al. [72] developed a system using a modified DenseNet201 model to differentiate among four types of lung cancer, employing the Kaggle dataset and two feature selection methods to enhance feature extraction, achieving a mean accuracy of 95%. Sajja et al. [73] designed a deep neural network based on GoogleNet, reducing costs by sparsifying the densely connected architecture, leaving 60% of all neurons in dropout layers. Using the LIDC dataset, this model achieved an accuracy of 100% in the validation stage and 99.03% in testing, surpassing the performance of contrasted architectures: AlexNet, GoogleNet, and ResNet50.

Other authors explore different types of classifiers to enhance the performance of CNNs. For example, SVM, despite being a simpler approach, has yielded competi-

Table 6 Works and datasets for systems based on convolutional neural networks

	LUNA 16	LNDb	LIDC-IDRI	SPIE-AAPM	Kaggle	CIA	NIH chest X-ray	LZ2500	NLST	NCI genomics	DSB2017	Own dataset	Other
Jin et al. [82]	X												
Eun et al. [69]	X		X										
Ali et al. [87]			X										
Xiao et al. [83]	X												
Sajja et al. [73]			X										
Nasrullah et al. [85]			X										
Causey et al. [84]	X		X					X		X			
Toğaçar et al. [76]													X
Sori et al. [68]	X				X								
Katase et al. [86]		X	X	X								X	
Guo et al. [71]													X
Xu et al. [78]					X	X						X	
Jain et al. [77]								X	X	X			
Shakeel et al. [56]													X
Arshad Choudhry and N. Qureshi [74]			X				X						
Shah et al. [81]	X												
Ali et al. [70]	X												
A.R. et al. [80]			X										
Lanjewar et al. [72]					X								
Siddiqui et al. [75]			X										

Table 7 Summary of different articles related to systems based on CNN

Reference	Year	# datasets	Methodology	Code available	Main contribution
Jin et al. [82]	2018	1	3D residual CNN	X	Spatial pooling and cropping layers (SPC) to extract multilevel contextual information
Eun et al. [69]	2018	1	Ensemble single-view 2D CNNs, automatic categorization	X	Achieved state-of-the-art performance in reducing false positives in lung nodule detection with low computational demands
Ali et al. [87]	2018	2	Deep reinforcement learning	X	Developed a novel RL model for detecting lung nodules in CT images
Xiao et al. [83]	2019	1	Multi-Scale Heterogeneous 3D CNN	X	Gradual multi-scale integration, heterogeneous feature extraction and fusion of machine learning weights
Sajja et al. [73]	2019	1	GoogleNet-based deep neural network	X	GoogleNet-based CNN architecture for efficient lung cancer detection from CT images with 60% dropout
Nasrullah et al. [85]	2019	1	CMixNet, Faster R-CNN, GBM	X	System combining CMixNet architecture, Faster R-CNN, and GBM
Causey et al. [84]	2020	4	Spatial Pyramid Pooling, 3D Convolution	X	Developed DeepScreener, a model combining Spatial Pyramid Pooling and 3D Convolution
Toğaçar et al. [76]	2020	1	AlexNet, mRMR, k-NN	X	Using AlexNet, mRMR feature selection, and k-NN classifier
Sori et al. [68]	2021	2	Denoising CNN, Two-path CNN, DCA	X	DFD-Net combining denoising, two-path CNN, and discriminant correlation analysis for lung cancer detection from CT images
Katase et al. [86]	2022	3	Faster R-CNN + 3D-CNN	X	The system, designed on the Faster R-CNN architecture with 3D convolution layers, aimed to distinguish nodules
Guo et al. [71]	2022	1	Zero shot learning	X	The methodology involves training a ResNet-50 CNN on original DICOM images from 20 patients
Xu et al. [78]	2022	1	CNN and attention mechanism	X	Improved NSCLC classification accuracy using ISANET on chest CT images
Jain et al. [77]	2022	3	KPCA-CNN and FDBNN	X	Kernel PCA for feature extraction combined with Convolution Neural Network and Fast Deep Belief Neural Networks on histopathological images
Shakeel et al. [56]	2022	1	Improved deep neural network, Ensemble classifier	X	Using an improved deep neural network for segmentation and an ensemble classifier for classification
Arshad Choudhry and N. Qureshi [74]	2022	1	Transfer learning, SVM, manual features	X	Hybrid CNN model incorporating manual features and SVM classifier. Using VGG and Inception architectures
Shah et al. [81]	2019	1	Ensemble of CNN	X	Combine different CNN trained from scratch into ensemble based system

Table 7 continued

Reference	Year	# datasets	Methodology	Code available	Main contribution
Ali et al. [70]	2023	1	End-to-End 2D CNN	X	End-to-end CNN, oversampling for data imbalance, image transformation, reducing false positives
A.R. et al. [80]	2023	1	LCD-CapsNet	X	Hybrid architecture combining CNN for feature extraction and CapsNet for classification
Lanjewar et al. [72]	2023	1	Modified DenseNet201 with ETC and MRMR	X	Lung cancer detection using modified DenseNet201, ETC, MRMR feature selection, and machine learning classifiers on CT images
Siddiqui et al. [75]	2023	2	Improved Gabor filter, Enhanced DBN, SVM	X	Classification method combining an improved Gabor filter, enhanced deep belief network, and support vector machine

tive results in certain studies. Arshad Choudhry and N. Qureshi [74] proposed a system that classifies lung cancer images by utilizing X-ray images from the National Institute of Health and applying transfer learning techniques with Inception models for image recognition, complemented by trained VGGs and SVMs for classification. This architecture achieved accuracy values of 96.87% and an AUC of 0.92. Siddiqui et al. [75] introduced a model that combines Gabor filters with an enhanced deep belief network (E-DBN) using multiple classification methods and different types of restricted Boltzmann machines. In addition, SVMs were applied to improve classification results of lung nodules. The model, tested using the LIDC-IDRI and LUNA16 datasets, achieved a precision of 99.42%, a sensitivity of 98.50%, a specificity of 98.32%, and an F1 score of 99.37%. Toğaçar et al. [76] employed a Le-Net network, as well as other DL models like AlexNet and VGG-16, for detecting lung cancer using a subset of the TCGA-LUAD dataset consisting of 100 images. The CNN networks were used for feature extraction and classification, with subsequent application of various techniques such as linear regression and decision trees. The use of a kNN classifier resulted in an accuracy of 99.51% (see Fig. 8). Jain et al. [77] developed a model to detect tumors in lungs by analyzing histological images, which involves de-noising, scaling, and feature extraction using a CNN network and a PCA Kernel. The classification was determined by a Fast Deep Belief Neural Network (FDBNN). Using the LZ2500, NLST, and NCI Genomic datasets, they achieved an accuracy, recall, and F-value consistently higher than 97%, 79%, and 77.9%, respectively.

Some authors are now exploring more sophisticated techniques, which have shown promising results in related fields, such as attention mechanisms. When integrated with CNNs, these mechanisms enable neural networks to selectively focus on pertinent parts of the input, thus enhancing accuracy. This capability is particularly beneficial in the analysis of medical imaging tests, where it is crucial to identify subtle yet key features within complex images for accurate diagnoses. For instance, Xu et al. [78] developed a model called ISANET, which employs a convolutional neural network (CNN) augmented by attention mechanisms to classify non-small cell lung cancer (NSCLC). Built upon the InceptionV3 architecture, ISANET incorporates both channel and spatial attention mechanisms to specifically target pathological areas in chest CT scans. The study utilized a dataset from the Affiliated Hospital of Hebei University, which included 30 cases each of lung squamous cell carcinoma, lung adenocarcinoma, and normal conditions, totaling 90 patients and 619 images.

Another line of research that is widely explored by researchers involves designing architectures from scratch. This approach offers several advantages, including reduced computational costs, as these models are usually smaller than state-of-the-art models. Furthermore, as demonstrated by Liz et al. [79], such architectures can surpass the performance of larger pre-trained models when combined into small ensembles. This feature is particularly advantageous for medical datasets, thereby making it an intriguing area of research with numerous possibilities.

Within this framework, we can distinguish two distinct approaches. The first focuses on 2D-CNNs, which are effective for analyzing X-rays and similar tests but fall short in extracting all available information from three-dimensional samples. This limitation has spurred the development of 3D-CNNs, which are capable of utilizing temporal or volumetric data from medical imaging tests. A.R. et al. [80] designed a model that combines CNN networks with a capsule neural network (CapsNet), dubbed LCD-CapsNet, using the LIDC dataset. This model achieved precision, recall, F1, specificity, AUC, and accuracy metrics of 95%, 94.5%, 94.5%, 99.07%, 0.989, and 94%, respectively. Shah et al. [81] developed

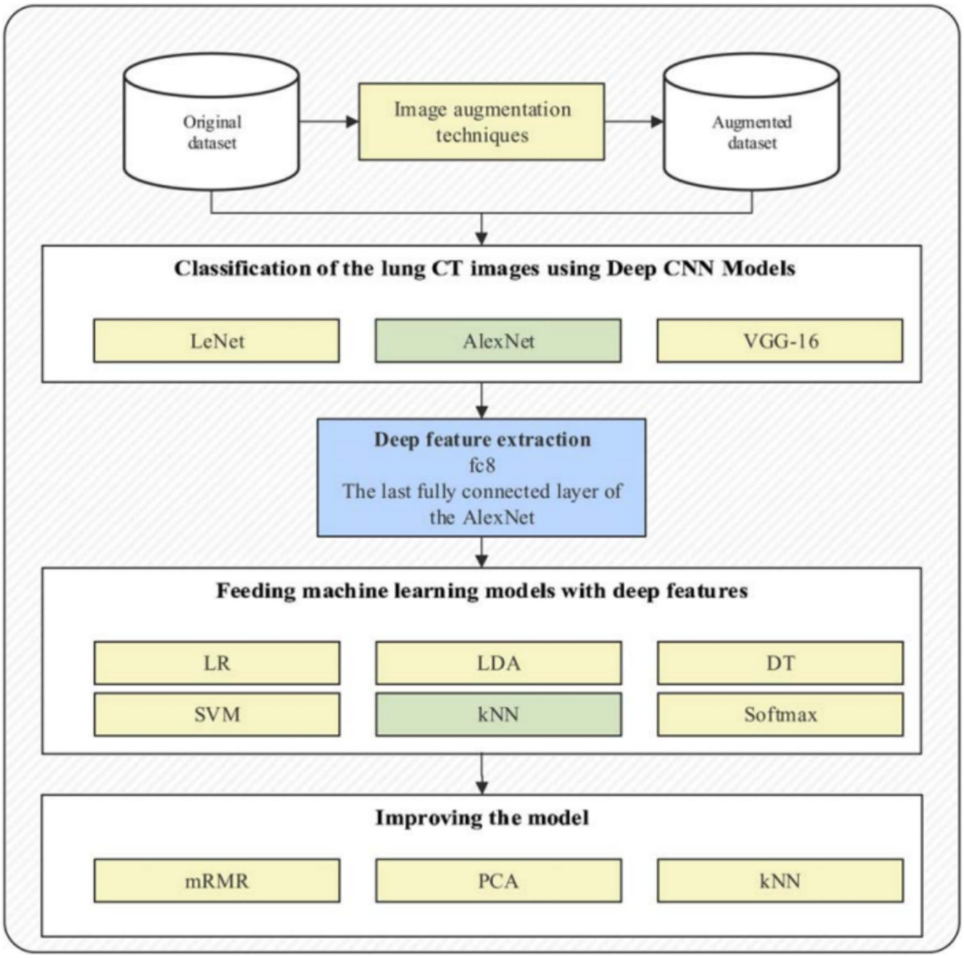

Fig. 8 The flowchart of Toğaçar et al. [76]

a CNN-based system to detect pulmonary nodules, which enhances accuracy by integrating two or more convolutional networks through ensemble techniques and employing a deep 2D assembler composed of three different CNN networks, achieving an accuracy of 95% on the LUNA16 dataset.

Additionally, some works focus on designing bespoke 3D-CNN architectures. For instance, Jin et al. [82] developed a deep 3D residual CNN to reduce false positives in pulmonary nodule detection from CT scans, featuring a novel 27-layer architecture with spatial pooling and cropping layers (SPC) for extracting multilevel contextual information. This network was evaluated using the LUNA16 Challenge dataset of 888 CT scans and achieved a high score of 0.924 in comparative evaluations. Xiao et al. [83] introduced a 3D convolutional network (MSH-CNN) based on chest CT scans, which employs multi-scale 3D nodule blocks to capture different levels of contextual information, extracts features through two distinct branches of 3D CNNs, and fuses these features using weights determined by back-propagation, achieving CPM values of 0.874 and a sensitivity of 91.7%. Causey et al. [84] proposed the DeepScreener algorithm, which utilizes Spatial Pyramid Pooling combined with 3D convolution to detect lung cancer from volumetric lung CT scans, evaluated using 1449 low-dose CT scans from the National Lung Screening Trial cohort, and employs a pseudo-3D imaging technique, where each CT slice is processed in three color channels, bypassing the complexity of full 3D convolutional networks, achieving an AUC of 0.892.

As noted, while 2D-CNNs, commonly known as CNNs, are limited to exploiting frame information, or two-dimensional arrays, some researchers have opted to combine CNNs with RNNs to fully analyze the available information in CT scans. Nasrullah et al. [85] introduced a model using two CMixNet architectures, where detection is performed by a Faster R-CNN and classification by a gradient boosting machine (GBM), utilizing the LIDC-IDRI dataset to achieve a sensitivity of 94% and a specificity of 91%. Katase et al. [86] proposed another model based on Faster R-CNN for detecting pulmonary nodules, employing 3D convolutional layers to exploit all available spatial information and achieving a sensitivity of 98% across different datasets such as LIDC-IDRI, SPIE-AAPM, and LNDb.

Lastly, we see authors employing less conventional approaches, such as reinforcement learning for classification tasks. An example of this innovative strategy is the work of Ali et al. [87], who developed and validated a reinforcement-based learning model for the early detection of temporal nodules in CT images, inspired by the AlphaGo system. They used the LIDC-IDRI dataset, which includes over 800 CT scans, and achieved an accuracy of 99%.

This section explores a variety of innovative approaches for diagnosing lung abnormalities using advanced neural network techniques and learning algorithms. It specifically highlights three-dimensional convolutional neural networks (3D-CNN), residual networks, capsule-based systems, and reinforcement learning algorithms, each employed across different datasets with the unified aim of enhancing accuracy, sensitivity, and specificity. This discussion not only underscores the significance of technological innovation in early medical diagnosis but also illustrates the pivotal role of CNNs in augmenting health outcomes through the accurate and efficient detection of lung diseases. Additionally, there is a growing trend towards the utilization of intensive pre-processing techniques before applying these complex architectural models, underscoring that the quality of the initial image is crucial for the effectiveness of the final diagnosis.

Systems Emphasizing Pre-processing Information

In this section, we will explore a collection of studies that emphasize pre-processing over detection models, specifically focusing on the application of segmentation techniques to eliminate irrelevant areas that may hinder the classification of samples. This analysis demonstrates how segmentation techniques help in classification tasks. The reviewed studies include the use of advanced techniques such as K-means, Cuckoo Search Optimization (CSO), and convolutional neural networks (CNN), coupled with specific segmentation methods and classifiers such as SVM. Approaches range from preprocessing and feature extraction to the accurate classification of lung nodules. Tables 8 and 9 summarize the most relevant information from the analyzed papers. In Table 8, a comparison between the datasets and different articles is presented; in Table 9, the methodology and main contributions of each work are outlined.

Some authors have focused on leveraging unsupervised learning techniques, such as K-means, for classification tasks. Ozdemir et al. [88] developed a comprehensive system incorporating both CADe and CADx components. They utilized a 3D CNN based on the V-Net architecture for segmentation and a variation of supervised learning known as MIL for classification, employing the LUNA and DSB datasets. Despite achieving a high sensitivity of 96.5% and an AUC mean of 0.87, they noted that dataset limitations impacted the results. Kanipriya et al. [89] also used this architecture for the classification of lung nodules, together with the K-means search algorithm, which groups objects into clusters based on their characteristics. They used various datasets such as the LIDC-IDRI and LungCT Diagnosis Dataset, with more than 100 cases. The highest accuracy value obtained was 99.84%, with an F1 value of 0.998 and an MCC value of 0.998. Prasad et al. [90] conducted research seeking a method to segment lung cancer. This system segments and classifies images into two categories, normal or abnormal, and consists of two phases. In the first phase, processing, feature extraction, classification, and segmentation are performed. In the second phase, the images are processed through the K-means algorithm. An SVM classifier is also used, achieving values of 96% in accuracy, 100% in specificity, and 99% in sensitivity. M and M [91] also used a modified K-means algorithm for their system based on four phases: pre-processing, segmentation, classification, and analysis. K-means was used for segmentation, while for classification, a CNN convolutional network together with the ATSO algorithm was employed. The dataset used was the LIDC, with 2610 tomography images, applying 80% of them

Table 8 Works and datasets comparison related to segmentation-based systems on classification tasks

	CIA	LIDC-IDRI	LungCT Diagnosis	LUNA-16	Kaggle	Own dataset
Jiang et al. [96]	X					
Shakeel et al. [92]	X					
Ozdemir et al. [88]				X	X	
Kanipriya et al. [89]		X	X			
Venkatesh and Bojja [93]						X
Venkatesh et al. [94]						X
Prasad et al. [90]		X				
Bhattacharjee et al. [95]		X				
M and M [91]		X				

Table 9 Summary of articles related to segmentation-based system on classification tasks

Reference	Year	# datasets	Methodology	Code available	Main contribution
Jiang et al. [96]	2018	1	Multigroup patch-based, deep learning network	X	Multigroup patch-based deep learning network and Frangi filter for improve metrics
Shakeel et al. [92]	2019	1	Profuse clustering technique (IPCT) + deep learning	X	Innovative approach combining improved profuse clustering and deep learning for lung cancer detection
Ozdemir et al. [88]	2020	2	V-Net, 3D-CNN	X	End-to-end system that integrates model uncertainty, improving calibration and probability interpretation in diagnostics
Kanipriya et al. [89]	2022	2	K-means, CNN-LSTM	X	Hybrid CNN-LSTM architecture, leveraging entropy-based K-means clustering and a Capuchin Search Algorithm
Venkatesh and Bojja [93]	2022	1	Cuckoo-search optimization with Otsu thresholding	✓	Cuckoo Search Optimization (CSO) algorithm for segmentation and Linear Binary Patterns (LBP) for the extraction process
Venkatesh et al. [94]	2022	1	Local binary pattern for feature extraction	X	Otsu thresholding segmentation and cuckoo search optimization for accurate nodule segmentation, local binary pattern for feature extraction
Prasad et al. [90]	2022	1	Fuzzy K-means + CNN	X	Integrating fuzzy K-means clustering and deep learning with SVM classifier for FP reduction
Bhattacharjee et al. [95]	2022	1	Adaptive Boost-based Grid Search Optimized Random Forest	X	Ada-GridRF utilizes Adaptive Boost-based Grid Search Optimized Random Forest for hyperparameter optimization in lung cancer detection
M and M [91]	2023	1	Segmentation based on k-means	X	Segmentation based on K-means and CNN with the integration of the adaptive optimization algorithm

for training and the remaining 20% for testing. The highest accuracy value was 96.5%, along with sensitivity values of 98%, specificity of 96.41%, and PPV of 97.07%. Shakeel et al. [92] explored other clustering techniques like IPCT to improve the quality of lung imaging and cancer diagnosis by reducing misclassification errors. The images were obtained from the CIA dataset, and during the process, noise was removed using the weighted mean histogram equalization method. For segmentation, an improved clustering technique called IPCT was used. After various tests, they achieved an accuracy of 98.42%.

Other authors focus on the optimization of hyperparameters using different techniques that allow us to achieve the best adjustment for DL models, improving performance. For example, Venkatesh and Bojja [93] developed a system using a CNN for classification, a Cuckoo Search Optimization (CSO) algorithm for segmentation, and linear binary patterns (LBPs) for the extraction process. For training and testing the model, images from a laboratory were used. These images consisted of more than 50 CT images. As final metrics, maximum accuracy values of 98% were obtained. In order to define the nodal features of the images, Venkatesh et al. [94] developed a system where the segmentation is performed by Otsu's method, by which an automatic thresholding of the images is performed. The CSO algorithm is used to define the features of the nodes. Finally, a CNN convolutional network is used to classify the nodule between benign and malignant. Again, the researchers do not specify the dataset, nor the quantity of images used. But this system obtained a 96.97% in the accuracy metric. Bhattacharjee et al. [95] propose a classifier called Ada-GridRF, which optimizes the parameters of the random forest model, to identify nodules as benign or malignant nodules on CT images. For this system, the LIDC-IDRI dataset, with a total of 7371 nodules, was used.

Table 10 Works and datasets used to improve accuracy and reduce false positive

	LUNA16	ANODE09	LIDC-IDRI
Saba et al. [98]	X		
Masood et al. [97]	X	X	X
Niu and Wang [100]	X		
Shen et al. [99]			X

As final metrics, accuracy, sensitivity, specificity, precision, and F1 values equivalent to 97.97%, 100%, 96%, 96.08%, and 98% were obtained.

Other techniques focus more on the preprocessing of the samples. For example, Jiang et al. [96] developed a scheme based on multigroup patches obtained from lung images (see Fig. 9), employing CNN networks for analysis. The LIDC-IDRI dataset, comprising over 1000 images, was utilized for this scheme. Within the CNN design, the ReLU function is used, achieving a sensitivity of more than 80%.

This section discusses how advanced segmentation and preprocessing techniques improve the classification of lung abnormalities in medical images. It highlights the use of methods such as K-means, Cuckoo Search Optimization (CSO), and convolutional neural networks (CNNs), which, together with classifiers such as SVM, form comprehensive systems for lung nodule detection and classification. The research reviewed demonstrates a significant focus on segmentation to remove irrelevant areas that may hinder sample classification.

The application of these techniques in public and private databases demonstrates their effectiveness, achieving high values in accuracy, sensitivity, and specificity. The developed systems show effective integration of preprocessing, segmentation, and classification phases, underlining the importance of proper data handling before final analysis. This approach not only improves the quality of diagnostics but also optimizes the performance of CAD and CADx systems in the medical context.

Strategies to Improve Accuracy and Reduce False Positives

In this section, several advanced methodologies implemented in lung cancer detection to optimize diagnostic accuracy and minimize false positive rates are examined. These studies highlight the use of decision support systems, such as 3D deep convolutional networks (3DDCNN) with automatic selection of regions of interest, and deep learning models that integrate optimized segmentation, feature extraction, and classification techniques. In addition, weakly supervised and self-supervised learning approaches for lung nodule detection and diagnosis are explored, demonstrating innovations in CT image processing through 3D nodule segmentation and nodular feature identification. Tables 10 and 11 show a summary of all the papers analyzed in this section. Table 10 makes a comparison between the different articles and the datasets used, and Table 11 shows the methodology and the main contributions of the articles.

Most authors focus on pre-processing. On the one hand, we have seen works focusing on feature selection, such as Masood et al. [97] who propose a 3DDCNN support system for radiologists' decision support (see Fig. 10). This system employs multiple RPNs for the automatic selection of regions of interest. This system has been evaluated and trained with LUNA16, ANODE09, and LIDC-IDR datasets. It obtains sensitivity, specificity, AUC, and accuracy values corresponding to 98.4%, 92%, 96%, and 98.51% respectively. Others use segmentation techniques, which will remove irrel-

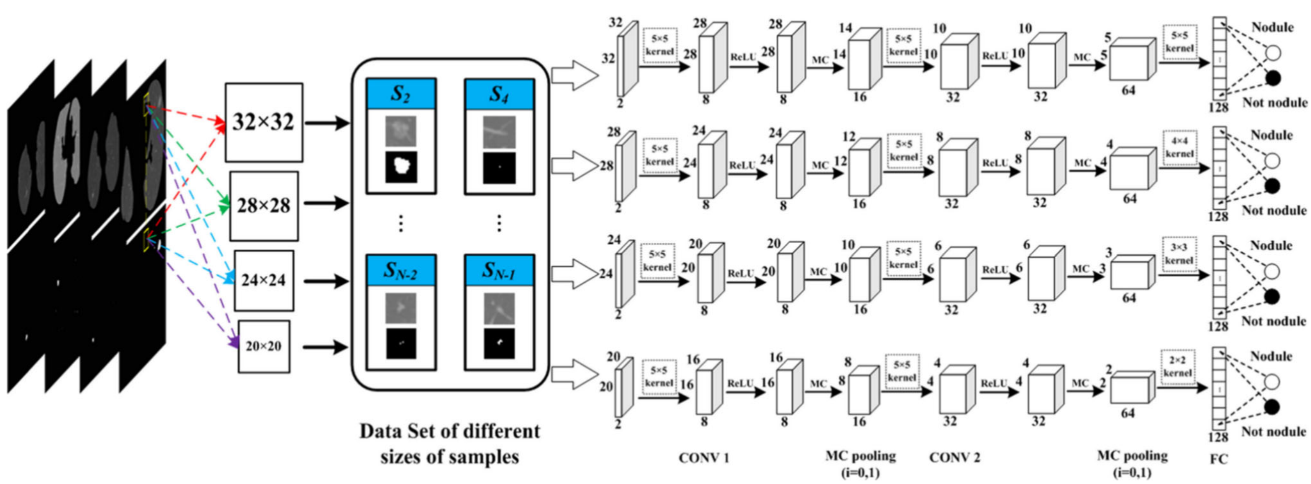

Fig. 9 Scheme of Jiang et al. [96]

Table 11 Summary of articles focuses on improving accuracy and reducing FP on classification tasks

Reference	Year	# datasets	Methodology	Code available	Main contribution
Saba et al. [98]	2019	1	Otsu thresholding + PCA	X	Otsu thresholding with morphological operations for lung nodule segmentation and extracting features using PCA
Masood et al. [97]	2020	4	3D-CNN + multi-Region Proposal Network	X	Multi-Region Proposal Network for automatic selection of potential regions-of-interest
Niu and Wang [100]	2022	1	3D transformer + contrastive learning	X	Develops a self-supervised 3D transformer model for lung nodule detection, utilizing region-based contrastive learning for effective training on a small dataset
Shen et al. [99]	2023	1	Weakly supervised network attention mechanism	X	Semi-supervised nodule detection (Semi-CADe) with a cross-nodule attention-based diagnosis network (CNA-CADx)

evant areas. For example, Saba et al. [98] develop a model for early stages of cancer, which is composed of segmentation, extraction, and classification phases. They use the Otsu threshold for segmentation, and the images used are obtained from the LUNA16 dataset. In this work, the authors propose an optimal CNN model in which Gaussian noise is removed to improve classification and certainty and obtain accuracy, precision, sensitivity, and specificity values of 0.9584, 97.4947, 91.3974, and 90.325. Shen et al. [99] introduce a weakly supervised lung cancer detection and diagnosis net-

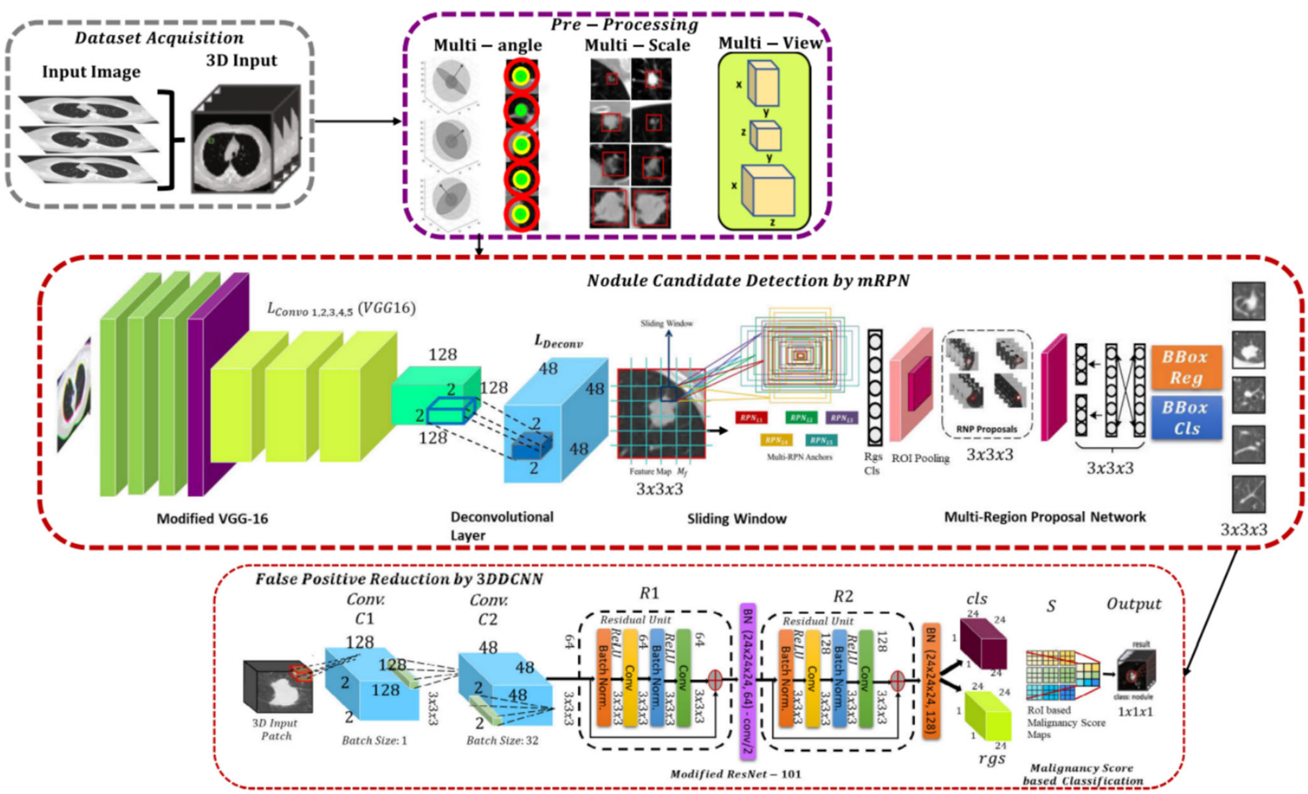

Fig. 10 Scheme of Masood et al. [97]

work called WS-LungNet. This model allows segmenting 3D lung nodules and determining how malignant the modules are. As a dataset, they use LIDC-IDRI and achieve a CPM of 82.99% and an AUC of 88.63%.

Lastly, the use of more novel techniques such as transformers or attention mechanisms is beginning to be observed. For example, Niu and Wang [100] presented in 2019 a transformer model based on self-supervised 3D regions, which was able to identify lung nodules between different regions. The 3D transformer divides images from CT scans into non-overlapping cubes and then extracts and analyzes the features. When performing the evaluation, they established two settings: one setting for nodules larger than 6 mm where they used the LUNA16 dataset, and in the second setting, they used the entire LUNA16 dataset without size limitations. This method improves the performance with respect to the 3D CNN networks.

Summarizing, this section illustrates the impact of deep learning techniques in improving lung cancer detection. The adoption of advanced models such as 3DDCNN, together with weakly supervised and self-supervised learning approaches, has enabled not only more effective segmentation and classification of lung nodules, but also a marked reduction in the incidence of false positives. These technological advances promise a significant improvement in radiological decision support systems, optimizing diagnostic protocols and contributing to earlier and more accurate detection of lung cancer.

Lung Nodule Localization Techniques

After discussing the classification of medical imaging for lung cancer detection, we will now focus on the task of localization. It, as opposed to classification, not only identifies the presence of lung cancer but also determines its specific location within the lungs [101–104]. This distinction is critical for planning precise and personalized treatments, allowing healthcare professionals to target interventions with greater accuracy. The precise localization of cancerous lesions is a complex challenge that has led to the development of advanced technologies, including 2D and 3D convolutional neural networks, and attention mechanisms that significantly improve the identification of affected areas, as shown in Fig. 11. Therefore, the papers have been organized following the same structure: from the application of CNN-2D and CNN-3D in cancer detection, through the use of attention mechanisms to improve localization accuracy, to the discussion of other innovative methodologies in the field.

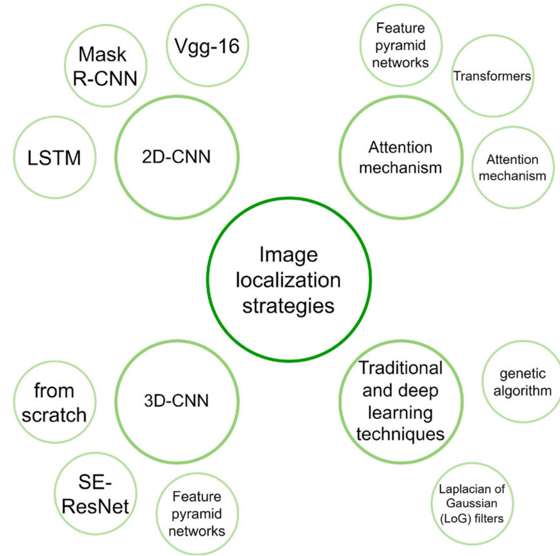

Fig. 11 Visual representation of main techniques of localization tasks

Use of 2D Convolutional Neural Network Architectures

In the subsection "2D CNN," we focus on the use of convolutional neural network (CNN) architectures for the accurate localization and detection of lung nodules in medical imaging. These studies range from histopathological image analysis to CT and MRI, using advanced CNN architectures such as VGG-16, DCNN, and Mask R-CNN, along with innovations in 3D lung nodule segmentation, detection, and visualization. Approaches include decision support systems for radiologists, detection of nodules in early stages of cancer, and automatic identification of suspicious lesions. Table 12 shows the datasets used in all the articles analyzed in this section, and Table 13 shows all the papers analyzed, their methodologies, and main contributions. It is worthwhile to note that these studies achieve remarkable results in terms of accuracy, sensitivity, and specificity, demonstrating the potential of 2D CNNs in improving lung diagnosis. However, it has a main limitation compared to other state-of-the-art techniques: it can only take advantage of information from independent frames and is not able to use all the information available in some medical images that present three dimensions.

Cai et al. [105] present an innovative system for detecting and segmenting pulmonary nodules and creating their 3D visualizations. The system uses a Mask R-CNN model with a ResNet50 backbone and a Feature Pyramid Network (FPN) to effectively extract multi-scale feature maps and propose candidate bounding boxes. The model's effec-

Table 12 Works and datasets used for developing CNN-2D architectures

	Kaggle	LUNA16	Other datasets*	Own datasets
Basha and Surputheen [110]	X			
Chi et al. [108]		X	X	
Zhang et al. [106]				X
Cai et al. [105]		X	X	
Akila Agnes et al. [111]		X		
Li et al. [109]				X
Xu et al. [107]		X		
Kvak et al. [112]				X

*Datasets for which information is not available, or is not available in English

tiveness is validated on the LUNA16 and the independent dataset from the Ali TianChi challenge, showcasing its capability to achieve high sensitivities of 88.1% and 88.7% at 1 and 4 false positives per scan, respectively. Additionally, the paper introduces a ray-casting volume rendering algorithm to generate detailed 3D models of pulmonary nodules, aiding in the comprehensive analysis of nodules' characteristics and improving diagnostic processes. According to the authors, the combination of Mask R-CNN for precise detection and segmentation, with advanced 3D visualization techniques, signifies a notable advance in the field of medical imaging and diagnostics. Zhang et al. [106] also develop a Mask R-CNN for detecting pulmonary nodules and examining lung function using CT and MRI scans. In this work, the authors included a R-FCN for faster detection. The research involves data from 56 patients who were diagnosed with pulmonary nodules through surgery or puncture. The CNN model is designed to improve the accuracy and efficiency of pulmonary nodule detection and lung function analysis. Although specific datasets are not named, the study

Table 13 Summary of articles that use 2D convolutional neural network architectures

Reference	Year	# datasets	Methodology	Code available	Main contribution
Cai et al. [105]	2020	2	Mask R-CNN + ray-casting	X	3D visualization for pulmonary nodule diagnosis using Mask R-CNN and ray-casting volume rendering
Chi et al. [108]	2020	3	Multi-step cascaded networks	X	Novel DCNN framework for pulmonary nodule detection in chest CT using cascaded U-Net-like networks
Li et al. [109]	2020	3	Multi-resolution fusion CNNs	X	Multi-resolution CNNs for enhanced lung nodule detection in chest X-rays with robust classification fusion
Zhang et al. [106]	2021	1	Mask-RCNN + R-FCN	X	Significantly improved accuracy and efficiency in pulmonary nodule detection using R-FCN
Basha and Surputheen [110]	2022	1	VGG-16 + CNN	X	High accuracy and effectiveness in distinguishing between lung cancer types
Akila Agnes et al. [111]	2022	2	Enhanced UNet +convolutional LSTM	X	Developed a two-stage CADe system for high sensitivity lung nodule detection using enhanced UNet and convolutional LSTM
Xu et al. [107]	2023	1	Improved Faster R-CNN + multi-scale training	X	Enhanced Faster R-CNN for lung nodule detection, improving precision and recall significantly
Kvak et al. [112]	2023	1	YOLOv5, convolutional neural networks	X	Developed a DLAD for CXR that significantly outperforms radiologists in detecting pulmonary lesions

utilizes actual patient data, ensuring the model's practical relevance. The research represents a significant advancement in the application of deep learning algorithms in the medical field, specifically in enhancing diagnostic procedures for lung-related diseases. Other authors that make use of R-CNN for faster detection are Xu et al. [107], who used this algorithm for feature extraction, using more than 1100 CT images from the LUNA16 dataset. The algorithm is divided into two paths: The first is used to obtain the regions of interest (ROI), while the second is employed to obtain a shared feature map. Regarding improvements, it presents better resolutions in the images, reaching an accuracy of 90.7%, being better in all indexes compared with the mainstream YOLOv3 and Cascade R-CNN detection algorithms.

In the case of Chi et al. [108], the authors developed a DCNN system for the detection of pulmonary nodules in images obtained by computed tomography. This system was composed of three cascading U-Net networks that segment, detect suspicious nodules, and finally detect real nodules. Images for the development of the system were obtained from the LUNA16 dataset and the independent ALIBABA Cloud Tianchi Medical Competition. Using LUNA16 dataset, they achieved metrics of accuracy, sensitivity, specificity, and AUC of 0.9390, 0.8988, 0.9476, and 0.9615, respectively. These metrics improve when performed with the Tianchi dataset, reaching an accuracy of 0.9475, sensitivity of 0.9036, specificity of 0.9655, and AUC of 0.9722. As shown in the work, the proposed model achieves better results in comparison with other architectures such as 3D-FCN, MR-CNN, 3D-UNET, PRN-HSN, DCNN, CLAHE-SVM, and MASK-RCNN. Li et al. [109] introduce a novel CNN-based approach for lung nodule detection in CXR images, reflected in Fig. 12. The method employs patch-based multi-resolution convolutional networks to extract features and utilizes four different fusion methods for classification. The system was evaluated using databases from three different sources: a publicly avail-

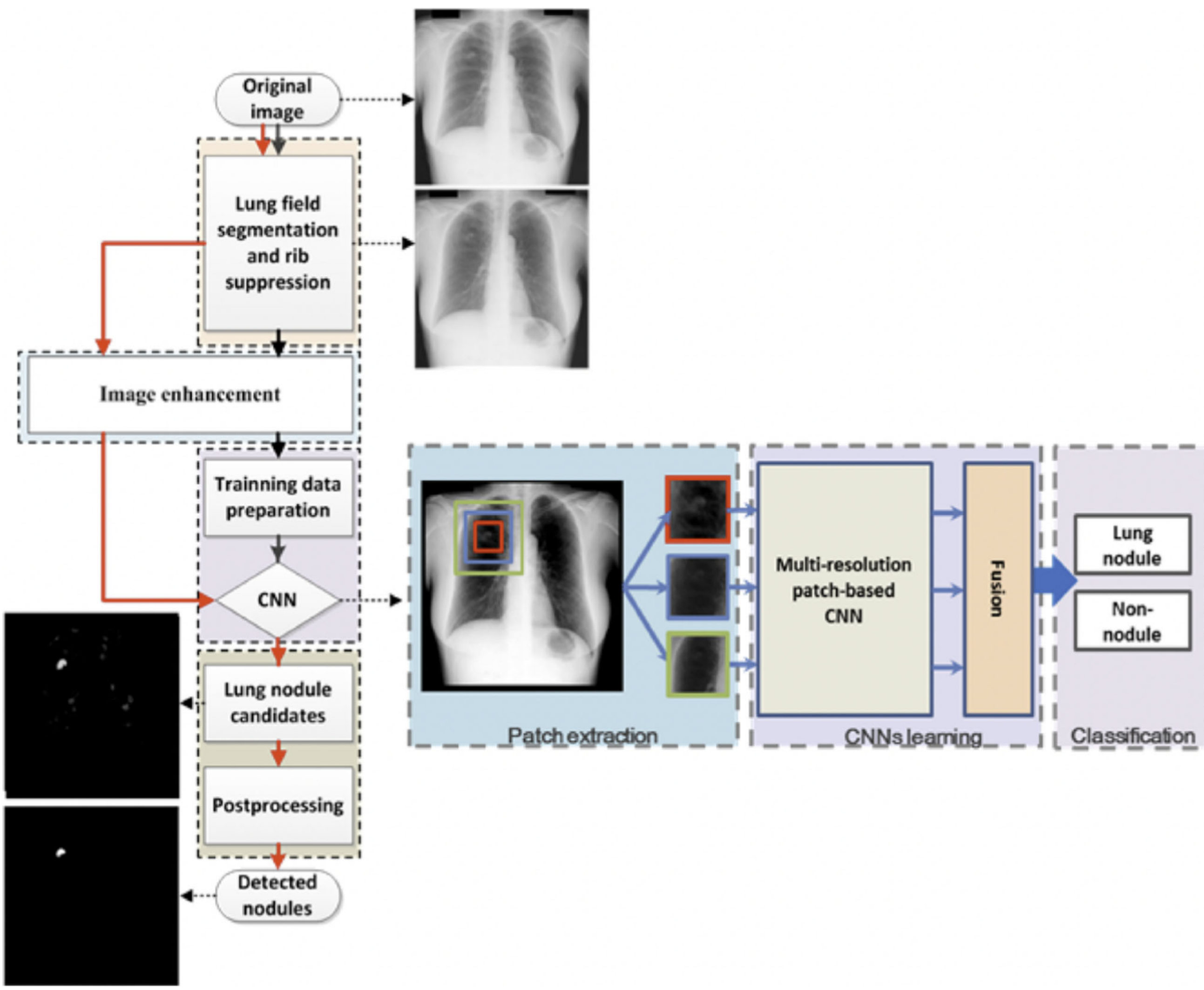

Fig. 12 Block diagram of Li's Li et al. [109] proposed method

able Japanese Society of Radiological Technology (JSRT) database and two additional databases from hospitals in Guangzhou (GDH) and Shenzhen (SZH). The JSRT database contains 247 digitized CXR images, while the GDH and SZH databases provide additional chest radiographs, contributing to the model's robustness and diversity in training and validation. In this work, the best results were obtained by the full-fusion method, with a Fairness-Area-Under-the-Curve (FAUC) of 0.982 and a Refined Competition Performance Metric (R-CPM) of 0.987. The results outperformed radiologist and non-full fusion methods. Given these results, and the low false positive per image rate reached (0.2), the authors conclude that this approach has the potential of applications in clinical practice.

In the study of Basha and Surputheen [110], a DL model leveraging the VGG-16 convolutional neural network architecture is used for the accurate classification and detection of lung cancer from histopathological images, specifically targeting non-smokers. The model, trained on 750 cases from the extensive dataset from Kaggle, achieved 97.99% of accuracy after ten iterations. The proposed model showed to be very effective in distinguishing between three types of lung cancer: benign, adenocarcinomas, and squamous cell carcinomas. Thereby, the authors stress the significant potential of specialized DL architectures in improving the precision and efficiency of lung cancer detection in medical diagnostics. On the other hand, Akila Agnes et al. [111] propose a computer-aided detection (CADe) system for lung nodule detection. The system utilizes a two-stage approach: initial nodule detection using an enhanced UNet model and false nodule elimination via a pyramid-dilated convolutional LSTM (PD-CLSTM) network. The first stage, Atrous UNet+, employs dilated convolutions and ensemble mechanisms to detect nodules of various sizes in the axial view of CT scans. The dilated convolution allows dense prediction without increasing computation and ensemble skip connections fuse multi-level semantic features for robust nodule detection. The second stage uses the PD-CLSTM network to categorize true from false nodules by learning inter-slice and intra-slice spatial features, applying pyramid dilation to capture multi-resolution spatial features without complicating the model. Tested on the LUNA16 dataset, the system demonstrates a high average sensitivity across multiple false positive rates, notably achieving sensitivities of 0.92 for small nodules (5–9 mm) and 0.93 for larger nodules (>10 mm). The combination of enhanced UNet for detailed detection and convolutional LSTM for precise classification establishes this system as a significant advancement in the detection of lung nodules, particularly due to its high detection rate across diverse nodule sizes. Finally, Kvak et al. [112] developed a deep learning-based automatic detection algorithm (DLAD) using the YOLOv5 architecture to identify pulmonary lesions from chest X-ray images. They trained the model on a custom dataset composed by 25.374 anonymized X-ray images collected across Europe, Asia, and North America, including 12.149 images with pathological findings and 13.225 with normal or insignificant findings. The main findings demonstrated that DLAD achieved a sensitivity of 91.0% and a specificity of 77.5%, which were superior in sensitivity to five assessing radiologists in a retrospective analysis of 300 images from a specialized oncology center. The study concluded that DLAD significantly enhances the detection of pulmonary lesions, suggesting its potential as a valuable decision-support tool in clinical settings despite its higher rate of false positives compared to human radiologists.

To conclude, the exploration of 2D CNNs in the subsection reveals significant advances in the localization and detection of pulmonary nodules, offering improvements in diagnostic accuracy and operational efficiency in medical practice. The adaptation of these deep learning technologies to the identification of lung pathologies shows a promising path towards the optimization of diagnostic processes, enabling early and personalized interventions. These studies underline the importance of 2D CNNs in the development of advanced diagnostic tools, marking a crucial advance in the fight against lung cancer and other lung diseases through the use of innovative technological solutions.

Use of 3D Convolutional Neural Network Architectures

This subsection focuses in the 3D CNNs. These architectures exploit more effectively the deep spatial information present in medical images, such as those obtained by computed tomography (CT), allowing for a more accurate interpretation of lung structures. These 3D models are able to capture the complexity and variability of lung nodules through multiple layers of convolution and clustering, which significantly improves nodule detection, false positive reduction, and accurate classification of nodules in terms of malignancy. Using large and diverse datasets, such as LUNA16 and LIDC-IDRI, the reviewed studies demonstrate remarkable advances in sensitivity and specificity, providing robust radiological decision support systems and highlighting the advantages of 3D CNNs over their 2D counterparts. Table 14 shows the comparison of the different papers and datasets used, and Table 15 the methodologies, and main contributions of these works.

The advantages of these kinds of architectures [113] designed a 3D CNN for automatic lung nodule detection. As for 2D CNNs, 3D convolution kernels and 3D pooling were used to obtain results. This system consisted of two steps, a first imaging step and a second discrimination step. This allowed the processing of large numbers of CT test images and reduced processing time. A total of 509 cases were used from the LIDC dataset following the LUNA16 rules. The

Table 14 Works and datasets used for developing 3D-CNN architectures

	LUNA16	LIDC-IDRI	LNPE1000	ANODE09	Own dataset
Pezeshk et al. [113]		X			
Gong et al. [114]	X				
Gu et al. [115]	X				
Han et al. [51]	X		X		
Xiao et al. [117]					X
Zheng et al. [116]		X			
Suzuki et al. [118]		X			
Lin et al. [120]	X				
Gürsoy Çoruh et al. [119]					X

Table 15 Summary of articles that use 3D convolutional neural network architectures

Reference	Year	# datasets	Methodology	Code available	Main contribution
Gu et al. [115]	2018	1	3D CNN + Multi-scale Prediction + Data Augmentation + Cube Clustering	x	Robust automatic lung nodule detection scheme using 3D CNNs and multi-scale prediction
Pezeshk et al. [113]	2019	1	FCN Conversion + Ensemble Learning	x	Two-stage 3D CNN system that works with CT scans with high sensitivity and minimal false positives
Gong et al. [114]	2019	1	3D SE-ResNet + U-Net Structure + Multitask Learning Loss + Online Hard Negative Mining	x	Effective 3D DCNN framework for automated pulmonary nodule detection, reducing false positives
Xiao et al. [117]	2021	1	Cascade FPN + Heterogeneous Neural Networks + Multi-scale Feature Integration	x	Cascade and heterogeneous neural network that significantly enhances the accuracy and reliability of pulmonary nodule detection in CT scans
Zheng et al. [116]	2021	1	Multiplanar Detection + Dense Training + Transfer Learning	x	Deep learning framework employing multiplanar views
Gürsoy Çoruh et al. [119]	2021	1	Fusion CNN Models + Feature Pyramid Networks	x	Proposed a fusion AI model combining state-of-the-art object detectors, significantly aiding in lung nodule detection and classification
Han et al. [51]	2022	2	3D CNN + Ensemble Learning + Multitask Learning	x	Developed an efficient and robust system for automated detection and classification of pulmonary nodules, enhancing accuracy and usability, with a deployed interface with hospitals
Suzuki et al. [118]	2022	1	Modified 3D U-Net + Soft Labeling	x	3D U-Net model validated across different datasets
Lin et al. [120]	2023	1	IR-UNet++ + Inception-ResNet + Multi-Scale Feature Integration	x	Advanced IR-UNet++ framework integrating Inception-ResNet blocks and SE attention for precise and efficient pulmonary nodule detection in CT images

training contains 833 nodules, and the tests have 104 nodules. This model achieves a sensitivity of 80% at 22.4 FP per case and a sensitivity of 95% at 563 FP per case. Therefore, a sensitivity of 95% is obtained. For the discrimination CNN, a sensitivity of 80% is obtained at 15.28 PF per case. Also, Gong et al.'s [114] works are based on this type of 3D CNN and the use of SE-ResNet residual networks, aimed at differentiating between nodules and reducing the false positive rate. In this work, the LUNA16 dataset is used, with more than 800 CT images included in the study. After performing the appropriate tests, sensitivity values greater than 93% and a competitive performance value (CPM) of 0.904 were obtained. In this line, a highly influential paper by Gu et al. [115] benefits from the fact that 3D convolutional networks take advantage of information in a more efficient way than 2D networks and use this type of network to develop a system capable of identifying pulmonary nodules. In this system, a five-phase composite, detailed in Fig. 13, segmentation is performed first, and then the CNN network is used. The method was evaluated in 888 thin-slice scans with 1186 nodules in the LUNA16 database, reaching a sensitivity greater than 87% and a CPM value of 0.7967.

In the study of Han et al. [51], a CAD system is introduced for the early detection and classification of pulmonary nodules using low-dose computer tomography (LDCT) images from physical examinations. The system integrates a 3D CNN-based model for identifying potential pulmonary nodules and providing detailed detection results with quantitative parameters. It employs 3D ResNet for the classification of detected nodules into intrapulmonary and pleural nodules, and a fully connected neural network (FCNN) for distinguishing ground-glass opacity (GGO) nodules from non-

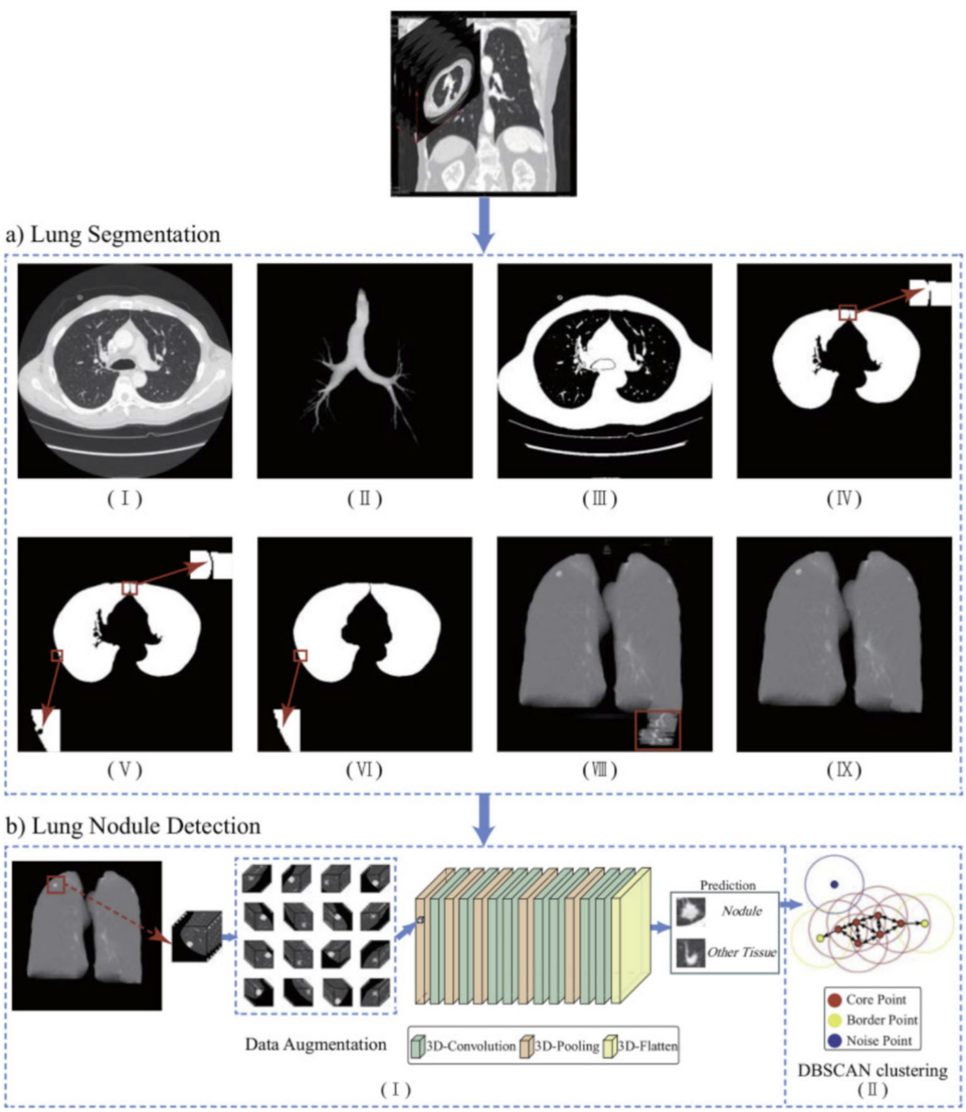

Fig. 13 Architecture of Gu et al. [115] proposed model. "Conv" and "Pool" denote the convolutional layer and pooling layer, respectively

GGO nodules. The system reaches high accuracy, achieving a detection accuracy of 0.879 and classification accuracies of 0.911 for intrapulmonary versus pleural nodules and 0.950 for GGO versus non-GGO nodules. In the study of Zheng et al. [116], another CAD system is proposed that uses a deeply supervised encoder-decoder network and a 3D multiscale dense CNN for multiplanar detection of pulmonary nodules. The system employs axial, coronal, and sagittal views for the detection of candidate nodules, merging the results to increase the sensitivity. In particular, the study highlights the proficiency of the system in detecting small nodules (<6 mm), a common challenge in the detection of pulmonary nodules. In this case, the authors use the LIDC-IDRI dataset with 888 CT scans and 1186 nodules. The system demonstrates high sensitivity, achieving 94.2% with 1.0 FP/scan and 96.0% with 2.0 FPs/scan. This research highlights the effectiveness of the multiplanar method and multiscale dense training in enhancing detection performance, marking a significant advancement in the field of medical imaging for the detection of lung nodules. In sum, these two works highlight the efficacy of the CAD system in supporting physicians by providing detailed information on suspected early lung cancer nodules, thus improving the detection efficiency in physical examination applications.

Other works focus on Feature Pyramid Networks (FPN). Xiao et al. [117] developed an automatic pulmonary nodule detection algorithm in this line. Initially, an FPN is used to detect candidate nodules with high sensitivity. Then, a second FPN is employed to refine the false positive results. Finally, a BasicNet is used alongside the second FPN to classify the final nodules and reduce false positives. Feature pyramids are used to find objects at different scales. However, the authors stress that due to limitations regarding computation and memory, small or large nodules may not be accurately identified. To address this limitation, these researchers decided to move to a 3D scenario instead of 2D, in addition to implementing FPNs. In cascade learning, a first FPN is trained with positive and negative patches/cuts regarding the nodules. Then, a threshold is calculated based on the minimum value that can be processed. Subsequently, negative cuts from false positives are selected, and along with the positive ones, the second FPN is trained. Therefore, the first FPN focuses on achieving high sensitivity, while the second focuses on reducing false positives. They used a total of 12,155 CT scans (one scan per patient) from multiple hospitals that were not specified and achieved a sensitivity of 88.8%, reducing the false positives.

In a different approach, Suzuki et al. [118] introduce a modified 3D U-net deep-learning model aimed at enhancing the accuracy of automated lung nodule detection on chest CT images. The model was trained using the LIDC-IDRI dataset, involving 888 CT scans, and externally validated using 450 chest CT scans from a Japanese university hospital, each case containing at least one nodule larger than 5 mm. The study focused on the model's accuracy, employing the competition performance metric (CPM) for evaluation. Internal validation showed a CPM of 94.7%, while external validation resulted in a CPM of 83.3%. The research concludes that the modified 3D U-net deep-learning model exhibits high performance in detecting lung nodules, marking a significant step forward in the application of deep learning in medical diagnostics for lung diseases.

Finally, Çoruh et al. [119] made a comparision study in which they developed a CNN model to detect nodules and their risk of being classified as malignant. This work developed a fusion model of various CNNs and, using the LUNA16 dataset, achieved a FROC value of 0.9513. These authors stressed the need to improve the joint work of physicians and AI tools. In this line, some works focus on building frameworks. Lin et al. [120] introduce a sophisticated 3D CNN framework, IR-UNet++, designed for the automatic detection of pulmonary nodules in lung CT images. The IR-UNet++ employs innovative elements such as Inception Net and ResNet blocks within a U-Net architecture, enhanced with squeeze-and-excitation (SE) mechanisms for improved feature extraction. The methodology involved training the model on the LUNA16 dataset and tested through systematic experiments. This framework achieved sensitivities of 90.13%, 94.77%, and 95.78% at 1, 4, and 8 false positives per scan, respectively. According to the authors, this framework enhances the detection of lung nodules, indicating a significant advancement over existing deep learning models in terms of accuracy and reliability in clinical diagnostics.

In conclusion, the implementation of 3D CNNs in lung nodule detection represents a significant advance in the field of medical imaging. The ability of these networks to process three-dimensional information offers a more complete understanding of lung anatomy, resulting in a marked improvement in detection accuracy and an effective reduction of false positives. The studies presented in this subsection highlight the importance of 3D deep learning techniques not only to overcome the inherent limitations of 2D CNNs, but also to enhance the early and accurate detection of lung cancer and, hence, other lung pathologies showing the same characteristics in terms of image classification.

Using Attention Mechanisms for Improving Detection

In the section “Attention mechanism for lung cancer detection,” we explore how attention mechanisms improve the accuracy and efficiency of neural networks in the detection and classification of lung nodules. These studies incorporate everything from DETR and 3D transformers to networks powered by attention mechanisms such as squeeze-and-excitation (SE) and attention channels, showing significant

Table 16 Works and datasets used to develop attention mechanisms for improving detection

	LUNA16	LIDC-IDRI	Own dataset
Naqi et al. [121]		X	
Barbouchi et al. [123]		X	
Mkindu et al. [125]	X		
Zhang et al. [122]	X		
Yang et al. [124]			X
Mkindu et al. [126]	X		

progress in the ability to highlight relevant features while minimizing false positives. The presented models demonstrate high performance on the two most used datasets: LUNA16 and LIDC-IDRI (see Table 16, with sensitivities up to 98. 65% and precision between 0.94 and 0.97. Table 17 is the summary of the papers and their main contributions. But most of all, this section underscores the importance of attention mechanisms in improving the representation of nodule features and assisting radiologists in a more accurate and efficient diagnosis.

In the first work in using this architecture, Naqi et al. [121] use geometric properties for nodule candidate detection and classification via a stacked autoencoder and softmax regression (see Fig. 14 for the proposed workflow). Employing the LIDC-IDRI dataset, their approach significantly reduced false positives to an average of 2.8 per scan, while achieving a high sensitivity of 95.6%. This demonstrated a substantial improvement in detecting lung nodules with high precision. The conclusion focuses on the potential of their method to support radiologists by enhancing the accuracy and efficiency of lung nodule diagnosis in clinical settings, making it a valuable tool for computer-aided diagnostic systems. Shortly after, Zhang et al. [122] proposed a different approach, developing a novel 3D feature pyramid network (3D FPN) enhanced by a squeeze-and-excitation (SE) attention mechanism for the detection of pulmonary nodules in CT scans. The model was aimed to improve detection performance by addressing the challenges of identifying small nodules and balancing the dataset's positive and negative samples. The 3D FPN captures spatial information and detects nodules at multiple scales, while the SE-attention block refines feature maps to emphasize relevant features and suppress irrelevant ones, improving the discriminative ability of the model. The network's effectiveness is demonstrated on the LUNA16 dataset, achieving a competition performance metric (CPM) value of 0.8934, indicating a significant improvement in detecting

Table 17 Summary of works using attention mechanisms for improving detection

Reference	Year	# datasets	Methodology	Code available	Main contribution
Naqi et al. [121]	2020	1	Geometric features+Autoencoder+softmax	x	Advanced lung nodule detection with significantly reduced false positives and high diagnostic sensitivity using DL and geometric analysis
Zhang et al. [122]	2021	1	3D Feature Pyramid Network (FPN) +squeeze-and-Excitation (SE) + CNN	x	Improved pulmonary nodule detection using a 3D FPN integrated with SE attention, enhancing sensitivity and reducing false positives significantly
Barbouchi et al. [123]	2023	1	Object detector based on an encoder-decoder transformer	x	Developed DETR, a transformer-based model achieving high accuracy in tumor localization and classification via PET/CT
Yang et al. [124]	2023	1	Lightweight neural network + YOLOv4 + spatial-temporal attention mechanism	x	Developed a highly effective and efficient lightweight neural network for lung nodule detection using an improved Ghost module
Mkindu et al. [126]	2023	1	3D U-shaped network + hybrid ECA modules + 3D regional proposal network	x	Excellent performance metric for lung nodule detection using a 3D CNN with integrated channel attention
Mkindu et al. [125]	2023	1	3D-MSViT + multi-scale patches + 3D RPN	x	Improved the state of the art using a 3D multi-scale vision transformer

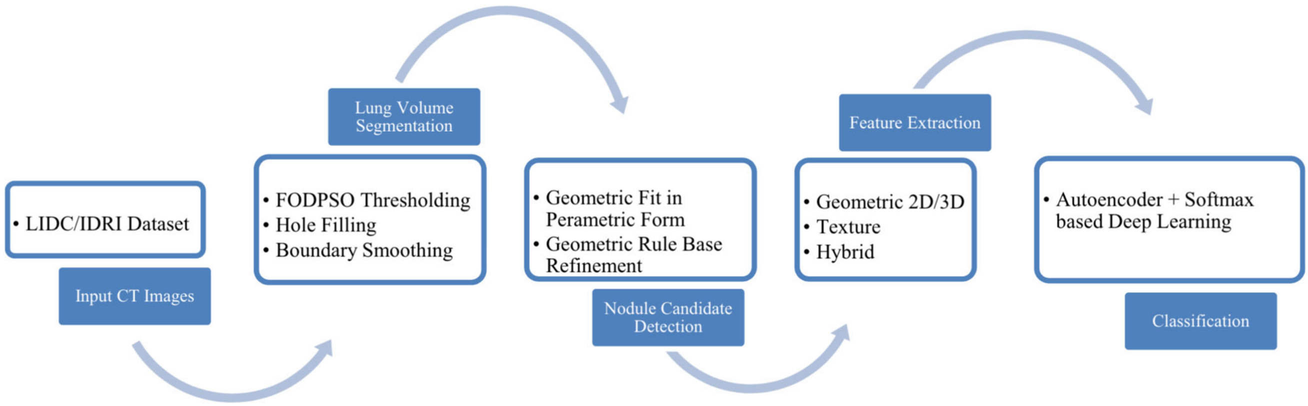

Fig. 14 Workflow model of the proposed method by Naqi et al. **(author?)** [121]

pulmonary nodules, especially smaller ones. The integration of 3D FPN with SE-attention and an optimized loss function showcases a promising approach for enhancing the accuracy and sensitivity of pulmonary nodule detection in medical imaging. By using another approach, Barbouchi et al. [123] present an approach to assess the ability of deep learning techniques to grade and detect lung cancer. They use the DETR model to detect tumors and assist experts. This model achieved an accuracy between 0.94 and 0.97.

Yang et al. [124] proposed an innovative and efficient deep learning model for the precise detection of lung nodules in CT images, using LUNA16 dataset. The model capitalizes on an enhanced version of the GhostNet structure, incorporating the structure from MobileNetV3 and enriched with a novel channel attention mechanism for refined weight distribution. Furthermore, a spatial-temporal attention mechanism is integrated, elevating the model's capability for in-depth feature extraction and analysis. This robust architecture not only refines the model's sensitivity to lung nodule characteristics but also maintains a lightweight framework, ensuring efficient processing. The model's effectiveness is underscored by its impressive performance metrics: achieving an F1-score of 0.87, precision of 86.34%, and recall of 86.69%, thereby marking a notable improvement over existing neural network models.

Finally, the architecture designed by Mkindu et al. [125], based on a multi-scale 3D vision transformer, has been validated on 888 CT images taken from the LUNA16 dataset. The highest sensitivity result obtained is 97.81% and the competition performance metric is 0.911. In this model, a CAD system is first used where all positive nodule candidates are marked, and then this number is reduced in a classification process. This architecture comprises different blocks: patch embedding, local transformer, global transformer, and detection module. The patch embedding block includes patch partitioning and linear embedding. Random image snippets are selected to be set as input values. The local transform block processes the features of the snippets, and then the feature maps are concatenated. Finally, a 3D RPN network is used for nodule detection on feature maps. On the other hand, Mkindu et al. [126] conducted a study employing CNNs with attention channel mechanisms to detect lung nodules on CT images. Efficient channel attention modules are also integrated to enhance the network representation. Using the LUNA16 dataset, a maximum sensitivity of 98.65% was achieved.

In conclusion, the integration of attention mechanisms in lung cancer detection through deep learning represents a remarkable progress in the field of diagnostic imaging. These approaches not only increase the sensitivity and accuracy of lung nodule identification but also optimize the interpretation of medical images, allowing a clearer distinction between benign and malignant nodules. By highlighting the most significant features and suppressing distractions, models with attention mechanisms offer a valuable tool to improve diagnostic outcomes and support clinical decision-making.

Merging Traditional and Deep Learning Techniques

This last subsection focuses on adding those innovative approaches to lung nodule detection that do not fit the rest of the classifications, especially highlighting methodologies that merge traditional image analysis and optimization techniques with the latest innovations in DL. On the one hand, we have the system that uses CNN together with the Marine Predator Algorithm (MPA) to optimize nodule detection, inspired by predator–prey dynamics in marine ecosystems. This approach achieves impressive accuracy and sensitivity in the RIDER dataset. On the other hand, NODULe, an algorithm that combines multiscale Gaussian Laplacian filters with shape and size constraints for initial nodule selection, followed by a densely dilated convolutional network for

Table 18 Summary of works of localization focuses on a combination of different deep learning techniques

Reference	Year	# datasets	Methodology	Code available	Main contribution
Zhang et al. [127]	2018	LUNA-16	Multi-scale LoG filters + 3D DCNN	x	Combined multi-scale LoG filters with a densely dilated 3D deep convolutional neural network (DCNN)
Lu et al. [128]	2021	RIDER	CNN + Marine Predators Algorithm	x	Achieved high detection accuracy compared with several state-of-the-art DL methods

accurate identification and estimation of nodule diameter, is presented, demonstrating its effectiveness with high scores in the LUNA16 challenge. Table 18 shows the datasets used in both papers and the summary information of the articles.

For the detection of pulmonary nodules on chest CT scans, Zhang et al. [127] propose a novel algorithm named NODULe. As can be seen in Fig. 15, the proposal integrates traditional image analysis methods with advanced deep learning techniques. Initially, it employs multiscale Laplacian of Gaussian (LoG) filters coupled with shape and size constraints to pinpoint potential nodule candidates. Subsequently, these candidates are processed through a densely dilated 3D deep convolutional neural network (DCNN), uniquely designed to increase receptive fields and extract more detailed and discriminative features. This network also estimates the diameters of the nodules. Using the LUNA16 dataset, NODULe achieved a high detection score, ranking 3rd on the LUNA16 Challenge leaderboard, and demonstrated a precise diameter estimation of the nodules. The combination of conventional methods for initial detection and deep learning for precise identification and measurement underscores the robustness and effectiveness of the algorithm in pulmonary nodule detection in CT images. Also, Lu et al. [128] developed a lung nodule detection system using

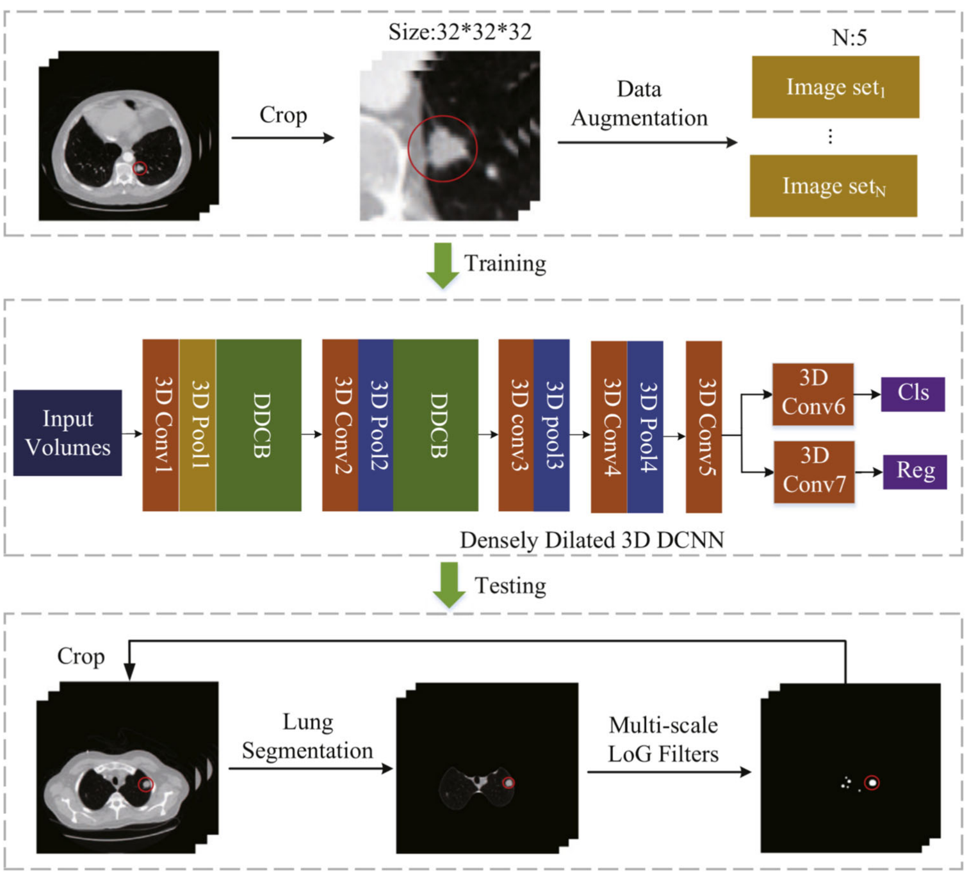

Fig. 15 Diagram of the proposed NODULe algorithm. "Conv," "Pool," "Cls," and "Reg" represent convolutional, pooling, classification, and regression layer, respectively. "DDCB" is the acronym of densely dilated convolutional layer [127]

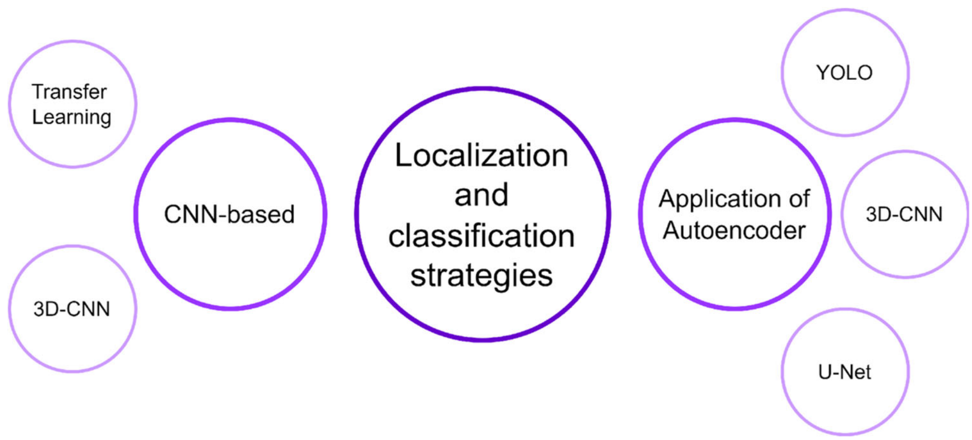

Fig. 16 Visual representation of main techniques for combination of classification and localization tasks

CNN and the Marine Predators Algorithm (MPA). This algorithm falls within the optimization group and is based on the predator–prey rate policy in marine ecosystems. Here, the search pattern changes depending on the number of prey in the environment. In this work, the RIDER dataset was used for the images with a precision of 93.4%, a sensitivity of 98.4%, and a specificity of 97.1%.

In summary, both approaches presented in this section illustrate the spectrum of innovations in lung nodule detection, from the application of optimization algorithms based on natural behaviors to the integration of traditional image processing techniques with advanced deep neural networks.

Classification and Localization

The last approach analyzed in this work is the combination of grading and localization techniques to detect radiological signs of lung cancer. By combining the classification of cancer—either by determining its presence or specifying its type—with the precise location of the lesion within the lung, these methodologies promise to improve diagnostic accuracy, optimize treatment plans, and increase the likelihood of therapeutic success [129–132]. We have identified two main lines of research: firstly, a group of authors has opted for the use of architectures based on CNNs; secondly, another approach is more focused on the feature selection, especially with the use of autoencoders. Figure 16 graphically represents a summary of the main techniques found in the research.

CNN-Based Methodologies for Localization and Classification

In the subsection "CNN-Based Methodologies for Localization and Classification," several strategies employing convolutional neural network (CNN) architectures for the detection and classification of lung nodules in computed tomography (CT) are reviewed. These studies stand out for their innovation in integrating techniques such as segmentation, false positive reduction, and adaptation of the networks to specific nodule characteristics, using large-scale datasets such as LIDC/IDRI and LUNA16. Table 19 details the datasets used in each research. Methodologies range from LeNet architecture to advanced systems like Faster R-CNN with adaptive anchor boxes and Elman neural networks. This section reflects the potential of CNNs to significantly improve diagnostic accuracy in lung cancer detection, highlighting the move towards more accurate and efficient techniques, as summarized in Table 20.

Table 19 Works and datasets used in CNN-based methodologies for classification and localization

	LUNA16	LIDC-IDRI	DSB	ANODE09	Own dataset
Li et al. [136]					X
Masood et al. [97]	X	X		X	
EL-Bana et al. [55]			X		
De Pinho Pinheiro et al. [135]		X			
Majidpourkhoei et al. [133]		X			
Nguyen et al. [137]	X				
Tiwari et al. [134]		X			
Su et al. [139]		X			
Chen et al. [138]		X			

Table 20 Summary of articles related to CNN based methodologies on classification and localization tasks combined

Reference	Year	# datasets	Methodology	Code available	Main contribution
Li et al. [136]	2019	1	Faster R-CNN, False Positive (FP) reduction	X	Deep learning-based lung nodule detection using Faster R-CNN, emphasizing scale independence and anatomical feature consideration for FP reduction in MR images
Masood et al. [97]	2020	3	Cloud-based system Multi-RPN	X	Cloud-based automated clinical decision support system using a modified VGG-16 for enhanced feature extraction and a multi-RPN structure
EL-Bana et al. [55]	2021	1	DeepLab-V3	X	Developed a two-stage framework using DeepLab-V3 for semantic segmentation followed by localization and classification
De Pinho Pinheiro et al. [135]	2020	1	Swarm intelligence, U-Net, Transfer learning	X	Novel approach using CNNs optimized with swarm intelligence
Majidpourkhoei et al. [133]	2021	1	Le-Net, transfer learning	X	Develops an advanced CNN based on Le-Net architecture for efficient feature extraction and classification in 3D CT lung scans
Nguyen et al. [137]	2021	1	Faster R-CNN, Mean-shift clustering	X	Enhanced Faster R-CNN for lung nodule detection employing adaptive anchor box sizes using Mean-shift clustering
Tiwari et al. [134]	2021	1	Mask-3 FCM, TWEDLNN	X	Novel deep learning and Mask-3 FCM based method for enhancing detection of lung nodules
Su et al. [139]	2021	1	Faster R-CNN, Adaptive anchor box sizes	X	Faster R-CNN for lung nodule detection by implementing adaptive anchor box sizes tailored through deep learning
Chen et al. [138]	2022	1	Multi-RPN	X	System using deep learning, significantly enhancing detection accuracy through innovative CNN architectures and advanced data enhancement techniques

As we have seen in the previous sections, many authors focus on applying techniques such as transfer learning to exploit the knowledge of CNNs in other domains to the problem of lung cancer detection. Majidpourkhoei et al. [133] designed a structure to identify lung nodules using the Le-Net architecture. Initially, they process the image and perform segmentation and identification. And finally, using a CNN, they classify the lung nodules. Using the LIDC/IDRI dataset, with more than 7000 images obtained from 300 CT scans, they achieved an accuracy of 90% and a sensitivity of 98.4%. Tiwari et al. [134] propose other deep learning methodologies using Elman neural networks (TWEDLNN) and Mask Unit (MU) which is based on the 3FCM algorithm. This system includes image segmentation using the Otsu method and other techniques. They employed the open dataset LIDC-IDRI. The accuracy of this system reached a value of 96%. De Pinho Pinheiro et al. [135] use swarm intelligence algorithms to develop different models for the classification and detection of nodules. The model was trained and validated using the public LIDC-IDRI dataset. Although transfer learning is not a particularly novel approach, it still performs well in problems such as the detection of lung cancer in imaging tests.

The main limitation of the above techniques is that they do not take advantage of all available information, as many medical imaging tests are three-dimensional, such as CT scans. To leverage this information, we observe two main lines: the use of RNN and the use of 3D-CNN. For example, Li et al. [136] developed a method to identify nodules in magnetic resonance imaging scans. To locate regions of lung nodules, they use Faster R-CNN networks with spatial three-channel input construction and transfer learning for fine-tuning. Furthermore, they designed a scheme to reduce the number of false positives. The dataset consisted of 142 T2-weighted MR scans from the First Affiliated Hospital of Guangzhou Medical University, with a total of 862 nodule regions. This work achieved a sensitivity of 85.2%, demonstrating that incorporating spatial information and anatomical characteristics significantly improves lung nodule detection in MRI images. Nguyen et al. [137] propose a system based on Faster R-CNN with adaptive anchor boxes derived from the ground-truth nodule sizes in the training dataset and a residual convolutional neural network for reducing false positives. After training and validation with LUNA16, a sensitivity of 95.64% is achieved. The false positive reduction network showed a sensitivity of 93.8%, specificity of 97.6%, and accuracy of 95.7%. With these results, the authors emphasize the effectiveness of adaptive anchor boxes and deep learning in medical imaging analysis. EL-Bana et al. [55] designed a system based on two stages. The first employs a segmentation technique using the DeepLab model, which improves the accuracy of nodule detection. Within segmentation, MobileNet-V2 and Xception are used. The second stage employs Faster R-CNN with Inception-V2. Using the LUNA16 dataset, a sensitivity of 96.4% was achieved, while a lung nodule classification accuracy of 95.99% was obtained with the DSB dataset.

Secondly, numerous studies are utilizing 3D-CNN to detect lung cancer through a combination of classification and localization techniques, yielding promising results. Masood et al. [97] propose a 3D-CNN support system to aid radiologists in decision-making. This system employs multiple RPNs for the automatic selection of regions of interest. It has been evaluated and trained with the LUNA16, ANODE09, and LIDC-IDRI datasets, achieving sensitivity, specificity, AUC, and accuracy values of 98.4%, 92%, 96%, and 98.51%, respectively. Chen et al. [138] proposed a model to address the challenges of localizing nodules and differentiating them from other possible injuries. For this purpose, they utilized an F-Net for detection and an MSS-Net for false positive reduction. The F-Net incorporates different scales along with Batch and ReLU activation functions, while the MSS-Net employs 3D convolutions to obtain better features, presenting 4 MS-Res blocks. They used 888 CT images with a total of 1186 nodules from the LIDC-IDRI dataset. In terms of metrics, they achieved an average sensitivity of 98.4% and a CPM value of 0.957%. Su et al. [139] employed R-CNN to develop a model capable of detecting lung nodules in computed tomography scans (see Fig. 17). Using the LIDC-IDRI dataset from 1018 patients, the dataset was augmented to include 7000 slices, emphasizing lung nodules of various types. With several parameter optimizations, the model achieved an average precision of 91.2%. The authors highlight the need to expand the dataset and incorporate small nodules and benign cases to further improve model sensitivity and detection capabilities.

In summary, this section of the article has explored several advanced deep learning techniques applied to the detection of pulmonary nodules in medical images. The main techniques identified include the use of convolutional neural networks (CNNs) and their three-dimensional variants (3D-CNNs), which have proven effective in CT image analysis. Additionally, transfer learning is highlighted as a key strategy to improve the accuracy of pre-existing models, although it is not a particularly novel approach, as it leverages prior knowledge from other domains. These methodologies have not only demonstrated high accuracy and sensitivity rates but have also enabled the implementation of innovative features such as adaptive anchor boxes and algorithms for false positive reduction. The convergence of these techniques underscores the ongoing evolution in the field of medical imaging, highlighting significant progress towards

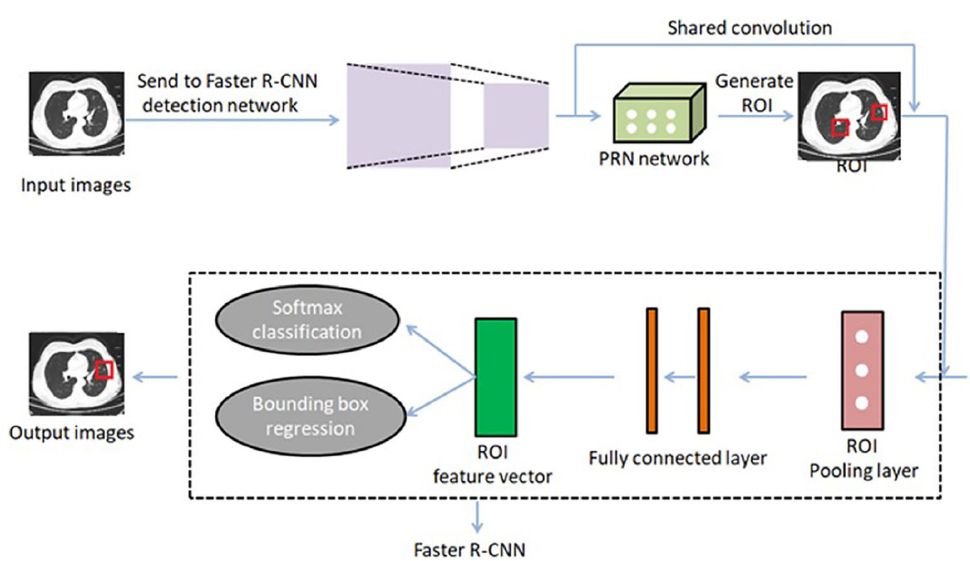

Fig. 17 Framework of Su et al. [139]

more accurate and efficient diagnostic tools for lung cancer detection.

Application of Autoencoder for Segmentation and Feature Extraction

In the “Application of Autoencoder for Segmentation and Feature Extraction” section, the use of autoencoder architectures, especially U-Net and encoder-decoder variants, to improve segmentation and feature extraction in lung nodule detection and classification is addressed. These studies explore different implementations of pre-trained 3D U-Net models and adapting mask architectures for accurate nodule identification, to employing swarm intelligence algorithms to develop robust models. With datasets ranging from private CT scans to chest X-rays (CXR) and the public LIDC-IDRI dataset, the presented methodologies demonstrate a variety of approaches to address nodule detection and classification challenges, achieving sensitivities up to 95.3% and overcoming labelling uncertainties with techniques such as the pseudo-labelling approach.

Some authors use 2D models, which means they do not take advantage of the sample’s depth information. For example, Chiu et al. [140] used a lung nodule detection model where they employed U-Net architecture for segmentation, and You Only Look Once version 4 (YOLO) for detection. For this model, they use chest radiography (CXR) images obtained from hospitals. They used 559 images for training and 100 for testing. According to the researchers, they did not obtain very high metrics, for example, a sensitivity of 82%, due to the small number of images used for training. However, other authors opt for 3D-CNNs, such as Ma et al. [141] that used a private dataset composed of 456 CT scans from the First Hospital of China Medical University, 80 benign and 406 malignant ground-glass opacity (GGO) nodules, classified based on pathological data and confirmed through expert medical consultation. This work proposes a two-stage model for the detection and classification of nodules. First, they employ a pre-trained 3D U-Net for feature extraction and then adapt the mask architecture. Finally, they use a false-positive elimination process, achieving a CPM value of 0.817. Chenyang and Chan [142] propose a joint model for the detection and classification of nodules (JNSC), see Fig. 18. It emphasizes on a joint approach for lung nodule detection, segmentation, and classification, considering the label uncertainty present in training data. The framework employs a 3D encoder-decoder architecture for both detection and classification, enhancing performance by leveraging nodule-specific features and addressing label uncertainties with a pseudo-label approach. They achieved a sensitivity of 0.953.

In conclusion, the application of autoencoders and encoder-decoder architectures in segmentation and feature extraction for the localization and classification of pulmonary nodules shows significant potential for the improvement of diagnostic processes in radiology. These techniques show excellent performance even in samples where there is uncertainty in labeling or variability. In Table 21, we can see a comparison between the different works analyzed, the used datasets, and a summary of their main contributions.

Explainability in Lung Cancer Detection

Artificial intelligence (AI) in healthcare has become a focus of innovation, offering new possibilities to improve diagnosis and treatment. However, its application in sensitive areas such as medicine requires careful compliance with legal, ethical, and technical standards. Medical data are classified as specially protected personal data under the General Data Protection Regulation (GDPR) in 2016, requiring robust protection to ensure their appropriate use. In addition, the European Health Data Space (2022) and the Artificial Intelligence Act (2023) introduce additional regulations addressing the use of AI systems in healthcare. These frameworks emphasize ethical principles and explainability, in particular

Table 21 Summary of articles related to the application of autoencoders on classification and localization tasks

Reference	Year	# datasets	Methodology	Code available	Main contribution
Chenyang and Chan [142]	2020	LUNA-16, LIDC-IDRI	3D-autoencoder	X	3D encoder-decoder structure and a pseudo-label approach, it effectively handles label uncertainty
Ma et al. [141]	2022	Own dataset	3D U-Net+Mask region-based convolutional neural networks	X	Two-stage 3D network, class-balanced loss, and FWC for false positive reduction
Chiu et al. [140]	2022	Own dataset	YOLOv4	X	U-Net for segmentation and YOLOv4 for nodule detection

for high-risk systems such as those used in medical diagnostics.

Medical data, as personal data, are subject to the regulations established by the GDPR (2016). This regulation establishes the category of specially protected personal data for this type of data. This classification implies certain requirements that data controllers must comply with before, during, and after processing. Thus, the GDPR establishes requirements such as risk analysis, the performance of impact assessments, or the appointment of a data protection officer. This is why companies or organizations involved in the development of AI systems that use this data must comply with the requirements of the GDPR.

The European Union has continued to strengthen these protections through additional measures. Following the adoption of the GDPR, the EU adopted the Regulation of the European Parliament and of the Council on the European Health Data Space (2022), which specifically refers to the secondary use of electronic health data such as images or diagnostic tests referred to in the various studies discussed in this article. This regulation includes the possibility of using this type of data for purposes related to scientific research related to the health sector, purposes linked to development and innovation activities of products or services that contribute to public health or social security, to the development of medicinal products or medical devices as well as to the training, testing, and evaluation of algorithms and AI systems and digital health applications, which contribute to public health or social security, or which ensure high levels of quality and safety of healthcare, medicinal products, or medical devices. Therefore, this legislation should also be taken into account when developing AI systems that are developed using this electronic health data. Procedures are established for individuals and companies to request access to these data for the development of the aforementioned activities for the implementation of systems, medicines, or medical devices.

It is also significant that the passage of the Artificial Intelligence Act (2023) has also brought new implications for artificial intelligence systems that use this type of data. In the classification of AI systems established by the AI Act, systems that use this type of data are considered high-risk systems, which implies that they have to comply with a series of requirements established by the regulation itself in order to be able to be developed and implemented within the EU. This standard defines a legal and ethical framework that must be especially respected by high-risk AI systems, such as those analyzed in this article.

Parameters such as anonymization are established for this type of data, in order to avoid damage to the fundamental rights and duties of data subjects. However, the regulation itself establishes certain protections, as it is aware that current anonymization systems are likely to allow the re-identification of the data subject. The AI Act requires that ethical parameters be met in the development of AI systems with the aim of creating a framework for AI that is lawful (complies with applicable regulations) and safe (does not cause harm, even indirectly).

Deep learning models, as explained in the introduction, are considered "black-box" algorithms, which means that we cannot understand how the system arrives at the final result. This limitation hinders their application in different domains, including medicine. An error in the system can cause harm to patients, so it is very important that the doctor can comprehend how the system has arrived at the final result.

In this article, we discuss how researchers address two different types of tasks for lung cancer detection: classification and localization. Classification focuses on identifying the category or categories to which an entire image belongs; however, localization not only identifies the presence of an object or objects within a sample but also pinpoints its location. These objectives are complementary in the detection of cancer in medical images, and several studies have been analyzed that use both techniques to enhance the accuracy and performance of the systems.

Localization tasks complement classification by providing spatial information, offering greater potential for Explainable AI (XAI) applications. While classification tasks provide a foundation for understanding what objects are present in an image, localization adds an additional layer by revealing where those objects are located. This dual approach enhances the interpretability of predictions, particularly in critical applications like lung cancer diagnosis.

Of all the papers analyzed in this survey, only two that developed models for sample classification applied XAI techniques. This underscores the urgent need to integrate XAI into medical AI applications. The use of these techniques enhances experts' understanding of the results, in this case, physicians. Guo et al. [71] employed class activation mappings (CAMs) to visualize the parts of the PET images that contribute most to the predictions of the lung cancer model, as shown in Fig. 19. These techniques help identify relevant features in the images that influence the model's decisions, allowing for a better understanding of how the model makes its predictions. Nam et al. [143] also utilized class activation mappings (CAMs), but in this case, for a chest X-ray classification system, as illustrated in Fig. 20.

Based on the above, we would like to emphasize the need for all proposed systems, and especially those intended for use in this medical field, to be explainable. In Europe, as a consequence of the General Data Protection Regulation and the recent Artificial Intelligence Regulation, the need for AI to be reliable is established to ensure its future development in all fields. This trustworthiness implies that AI systems used in the EU are safe, transparent, ethical, impartial, and under human control. In other words, trustworthiness implies the explainability of AI.

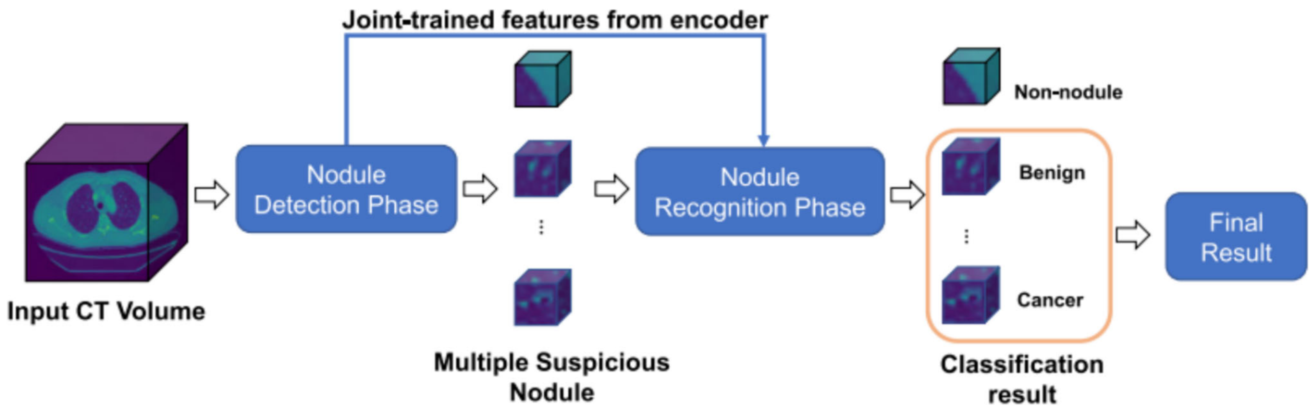

Fig. 18 Approach Chenyang and Chan [142]

As can be seen in this section and throughout the survey, the application of XAI techniques in classification tasks is extremely limited; however, there are numerous papers that show the capabilities and the variety of options they present to suit both the problem and the end users [144].

This shows that a very small percentage of the papers analyzed within the classification task use XAI techniques to understand the features that the system has used to arrive at the final result. This presents a serious problem that will have to be taken into account in future work in this area, or else, its application in real situations will be extremely limited.

Discussion

This survey has carried out a detailed analysis of the contributions of the application of DL techniques to lung cancer detection, more specifically this work has focused on the application of DL to analyze diagnostic images from two different approaches, classification, localization, and the combination of both.

Throughout this review, our research questions have been thoroughly analyzed, both descriptively and comparatively. This final section aims to complete the overview of this field by addressing three objectives: providing answers to the research questions, identifying the challenges and future trends associated with this issue, and finally offering a brief conclusion.

Answer to Research Questions

In the "Introduction" section, we outlined four research questions concerning the current status and trends in lung cancer screening research using deep learning techniques. These questions guided the methodology and literature analysis carried out in this review. Based on the analysis conducted in this article, we can now address these research questions with the insights obtained from our review process. Figure 21 presents

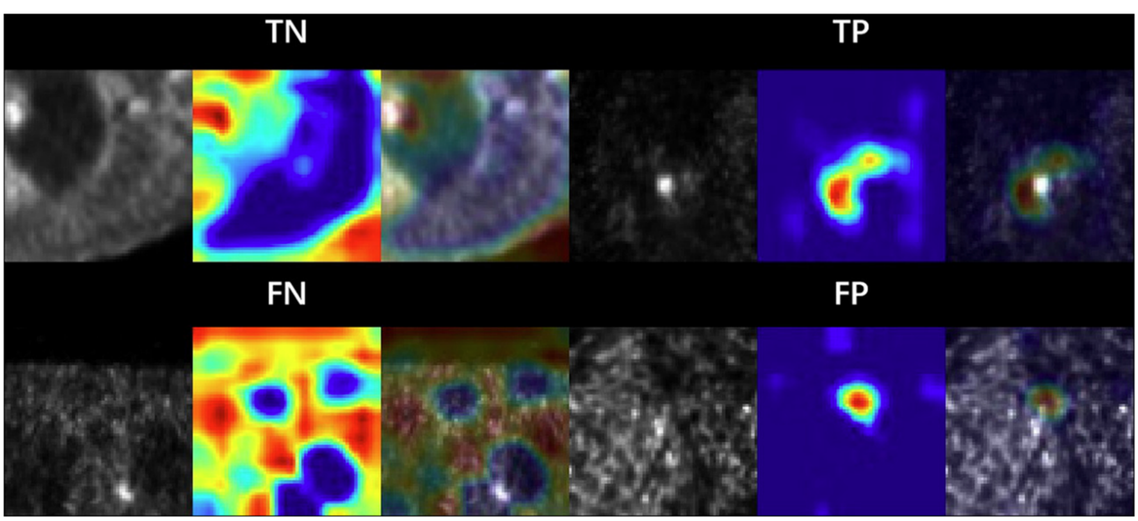

Fig. 19 Visualization from [71]

Fig. 20 Visualization from [143]

a summary of the main conclusions derived from this detailed analysis, emphasizing the answers to the research questions. A detailed explanation of these responses is provided below.

RQ1: What datasets are publicly available and accessible to researchers?

An interesting aspect to note is that most of the papers analyzed used open or freely accessible datasets that are still accessible online. Proof of this is that the LIDC-IDRI, which includes annotated computed tomography scans of 1,018 cases, has been used in 30 studies, being the most used dataset. The second most popular dataset is a subset of that, LUNA-16, that was specifically designed for challenges and is present in 29 works. As stated before, due to its precise and validated annotations, this dataset has proved to be particularly useful to perform the most complex and data-demanding techniques and architectures. The relevance of both datasets is highlighted by the fact that are the only ones that are present in all the analyzed tasks, and if we sum up the two, they are present in more than half of all jobs. The different datasets provided by the Cancer Imaging Archive dataset (CIA) are also among the most widely used: the RIDER dataset, which focuses on assessing tumor measurement variability with CT images from 32 participants, the TCGA-LUAD, containing images linked with genetic data, and CPTAC-LSCC, composed by radiology and pathology images from 248 subjects focused on tasks to explore the relations between cancer phenotypes and clinical data. Although we can see that there are a large number of datasets available, the most relevant, i.e., the most used in the articles analyzed in this survey are as follows:

- LIDC-IDRI
- LUNA16
- DSB 2017
- Kaggle
- The Cancer Imaging Archive (CIA)

But unfortunately, there is still a significative portion of works that makes use of insufficiently specified and unavailable data (marked as “other” in the tables, and accounting for a total of 6 papers) or use their “own” datasets (with 14 cases), that are not available. This poses a challenge and some risks, because while the “others” category is not present in the most complex architectures and tasks, using in-house own datasets has been present in each architecture/technique analyzed.

RQ2: What main tasks or topics are investigated within this area of study?

This paper was structured taking into account the main tasks involved in this area and ordered according to their increasing complexity. In this sense, each task builds on the previous one, and although they involve a higher degree of complexity, even today, they could be useful in different use cases.

Classification is the simplest task analyzed, involving categorizing detected nodules into benign or malignant based on their imaging characteristics. As in the other tasks, the complexity here lies in the need for a large annotated data set to train the model and the ability to capture subtle differences between benign and malignant nodules. Interestingly, it is in these tasks where we found the greatest variety of data sets, with a total of 15. One of the main advantages of this kind of architecture is that once trained, these models can be very quick. Here, increasing the performance of the models in terms of accuracy is critical, as the models included in this task can be very useful for early diagnosis and effective treatment planning.

The complexity of localization lies in the precise identification of nodules in imaging tests, not only the presence or absence of nodules in the whole image as was the case with the classification task. Also, the need to handle 3D images increases the difficulty of the task, requiring 3D-CNN algorithms, which in the vast majority of cases imply a much higher computational and data processing capacity. To meet the requirement of detecting the precise location of nodules, the models require a vast amount of high-quality data, labeled and contrasted by human professionals. That is why the LUNA-16 dataset excels and is the most used, unfortunately followed by in-house datasets that are not available. As these algorithms aim to pinpoint the exact coordinates of

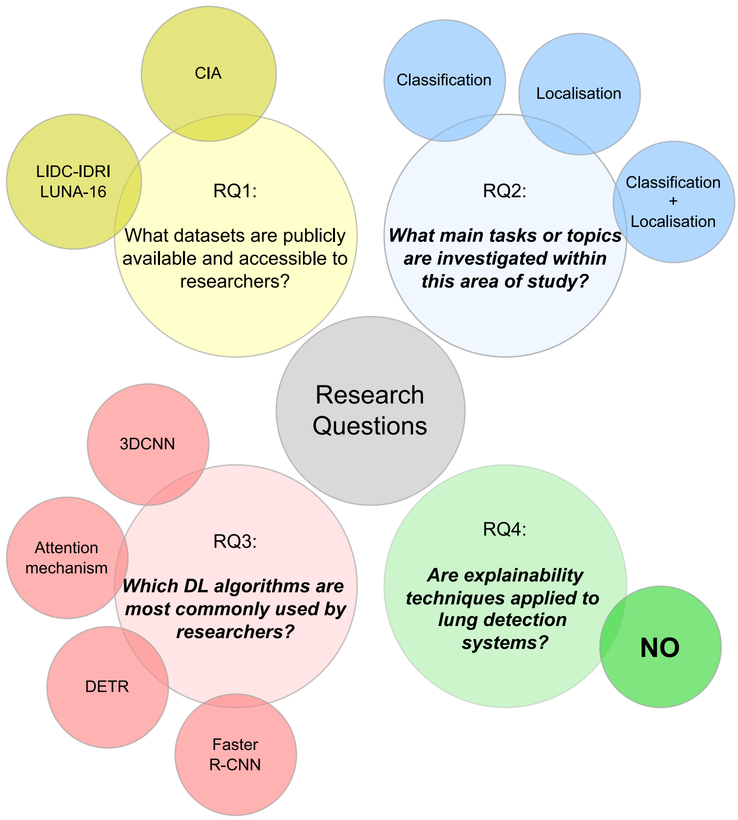

Fig. 21 Summary of key items of the RQ

nodules, the goal is to use them during medical procedures such as biopsies and surgeries, where precision is paramount.

Finally, classification and localization, the most complex task among the three, involve delineating the boundaries of the nodules with respect to the surrounding tissue, thus obtaining a much more accurate measurement of the size and evolution of the nodules. Here, pixel-level accuracy is mandatory to ensure that the segmented regions are precise. Not surprisingly, extremely complex architectures are used to accomplish this task, requiring significant processing power and sophisticated models to achieve high accuracy, using LIDC-IRC and LUNA-16 datasets in most of the cases. Finally, it is worthily to note that segmentation is essential for volumetric analysis, growth tracking, and treatment response evaluation, making it a critical component of comprehensive lung cancer detection systems.

RQ3: Which DL algorithms are most commonly used by researchers?

First, classification tasks in lung cancer detection primarily focus on identifying the nature of detected nodules, distinguishing between benign and malignant growths based on their imaging characteristics. It is not strange that CNNs are the most used algorithms in this approach due to their ability to automatically extract and learn complex features from medical images, and the VGG network is frequently combined with other architectures to increase classification performance and robustness in diagnostic outcomes. In addition to these individual models, ensemble and hybrid approaches are increasingly prevalent. For instance, ensembles of CNNs combine multiple networks, improving overall classification accuracy.

In the case of localization tasks, 3D-CNNs are widely used due to their ability to process volumetric data and capture spatial relationships within 3D medical images. Here, Deep 3D Residual CNN stands out for its effectiveness in reducing false positives and enhancing detection accuracy. Transformer-based models also show promise in this domain. For example, the DEtection TRansformers (DETR) model, excels in tumor localization and classification, particularly in PET/CT scans. The transformer architecture's ability to capture long-range dependencies and detailed spatial features enhances the accuracy and robustness of nodule localization and is proving to be, also in this field, an architecture with the capacity to make a qualitative leap in the way of managing the detection and localization of nodules.

Within the classification and localization architectures, one of the most prominent is the Faster R-CNN, which has been adapted with multi-scale training and adaptive anchor boxes to improve precision and recall in nodule detection. Enhanced versions of Faster R-CNN, incorporating mean-shift clustering for adaptive refinement of anchor boxes, further boost detection performance. In addition to these, Swarm intelligence optimized CNNs improves the feature extraction and classification processes. Transfer learning models, which leverage pre-trained networks on diverse datasets, facilitate the learning process and improve detection accuracy by adapting to the specific characteristics of lung nodule data.

RQ4: Are explainability techniques applied to lung detection systems?

Unfortunately, despite the importance, the application of XAI techniques in this field is limited. In this work, we only found a couple of papers that use grad-CAM algorithms. Explainable artificial intelligence (XAI) techniques are essential in the domain of deep learning applied to medicine due to the critical nature of the decisions derived from these systems. Medical data, such as radiological images, genetic sequences, or medical records, are often complex, multimodal, and high-dimensional, which makes deep learning models act as black boxes that are difficult to interpret. As these algorithms can have a direct impact on diagnoses, prognoses, and treatments, lack of interpretability can lead to mistrust among clinicians and hinder clinical adoption [156, 157]. XAI techniques can unravel model decisions, providing understandable explanations that help validate results, identify potential biases, and foster collaboration between AI systems and healthcare professionals, ensuring more ethical and safer use of these tools. In lung cancer detection from medical images, explainable artificial intelligence (XAI) techniques play a crucial role in ensuring the confidence and usefulness of deep learning models. Generic techniques, such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations), provide local or global interpretations of predictions by highlighting which features of the model (such as specific regions of the image) influence the outcome [158, 159]. On the other hand, image-specific techniques, such as Grad-CAM (Gradient-weighted Class Activation Mapping) [160] and saliency maps [161], allow to visualize critical areas in an image, identifying regions where the model focuses its attention when detecting possible malignant nodules. These tools not only help validate model performance, but also facilitate collaboration with medical specialists by providing visual explanations that can be integrated with clinical knowledge, thus improving accuracy and system integration.

Future Trends and Challenges

The preceding research questions and their corresponding answers provide an extensive and comprehensive overview of existing developments in the field of techniques and systems for the detection of lung cancer in medical images. However, the essential literature review conducted to answer these questions has also revealed several ideas and prospects for future exploration in this field. This section describes the upcoming trends that the literature is prepared to embrace, building on its current state. Additionally, it identifies the challenges that the research community will face and suggests possible ways to address them effectively. Figure 22 presents a schematic summary of these emerging trends and challenges.

Throughout this review of the state of the art, several trends have been observed, and potential directions for **future trends** in this field have been identified. Here, we will highlight the most promising research lines for the near future.

1. ***Improve the quality of the datasets***, several datasets analyzed in this article, such as Zhao et al. [44] and National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC) [46], have a very low number of patients compared to the number of samples. This will lead to a very ***low variability*** and will make it difficult for the model to generalize correctly in real situations, or different from the training samples. Therefore, one of the main lines will be the creation of quality datasets that allow the development of cancer detection systems that are able to generalize correctly and adapt to real situations.
2. An increasing number of authors are utilizing ***3D medical images***, such as CT scans, which provide significantly more information than other tests, such as chest X-rays. Consequently, more sophisticated detection systems, including *3D-CNNs* or combinations of *CNNs and RNNs*, are being employed to harness all the available information. This approach is emerging as one of the most promising lines of research because it offers a comprehensive view of the state of the lungs through the creation of ***new architectures***. In addition to CNNs, other architectures such as transformers are gaining traction in medical image analysis. ***Transformers*** [162–164], due to their capacity for capturing long-range dependencies and their adaptability to different data types, are increasingly being integrated into hybrid models or replacing traditional convolutional layers. Other promising architectures include graph neural networks (GNNs), which can represent complex spatial relationships, and ***Vision***

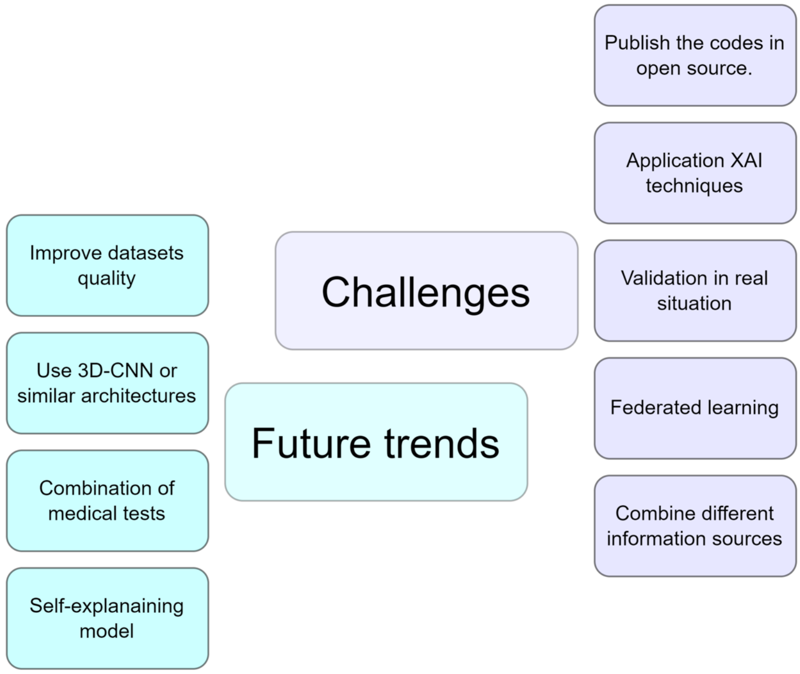

Fig. 22 Summary of future trends and challenges within lung cancer detection using DL

Transformers (ViTs), designed specifically for image processing. These advancements are expected to further enhance the accuracy and robustness of detection systems.

3. Doctors when analyzing imaging tests for disease diagnosis often use ***different imaging tests and views*** at the same time, as well as other medical tests, such as blood tests or respirometry, in the case of lung cancer. It is therefore expected that in the coming years, attempts will be made to ***combine different sources*** of information in the development of lung cancer detection systems, as it would provide high-quality information and would be close to the approach used by experts in the field. However, this approach will present a number of challenges for researchers, as Celik [145] or Bustos et al. [146] stated.
4. One of the main approaches for detecting radiological signs of lung cancer involves staging systems that do not utilize explainable artificial intelligence (XAI) techniques. Although this approach yields adequate results, its applicability is limited because it provides only the predicted class or the probability of belonging to that class. These systems are challenging to implement in fields such as medicine. Therefore, it is anticipated that in the coming years, researchers will increasingly focus on the ***localization*** approach or a combination of both ***classification and localization***. It will add a ***layer of explainability*** to the systems, facilitating their integration into real-world environments.

Secondly, we will examine the various challenges that the field of lung cancer detection will encounter. Throughout this review, we have identified several weaknesses that need to be addressed, presenting a **challenge** for future work in the area.

1. Throughout the article, it has been observed that the authors of most papers do ***not publish the code*** used in their research. This lack of transparency is likely one of the aspects most in need of revision, as it impedes the reproducibility of the studies and hinders progress in the field. Without access to the underlying code, it becomes challenging for other researchers to validate the results, identify potential flaws, or build upon previous work. This issue not only limits collaboration but also slows down the pace of innovation and improvement in medical DL applications. Future work in this area should aim to develop quality systems that are ***verifiable*** and ensure reliable performance. Additionally, encouraging a culture of openness and sharing within the research community could lead to more robust and trustworthy advancements, ultimately benefiting clinical practice and patient care.
2. One of the great challenges of AI, especially DL, in any field is the explainability of the decisions of these systems. In the medical field, the importance of the doctor understanding how the system has arrived at the final result is fundamental, since, among other things, it is the only way to explain the diagnosis and subsequent

treatment to the patient. Different methods have been developed to make deep learning or black-box algorithms more explainable or interpretable [147, 148]. However, in the case study of this article, in most of the texts analyzed, these methods are not applied. In many cases, this is because the use of these techniques may imply a lower performance of the system. It would be essential for the future development of AI systems applied to the detection of diseases such as cancer, to achieve this system explainability. As mentioned above, this would ensure the confidence of doctors and patients, who are the two actors most affected by the implementation of these systems. In addition, it would allow the widespread use of these systems by increasing users' confidence in them. Different methods have been developed to make deep learning or black box algorithms more explainable or interpretable. However, in the case study of this article, in most of the texts analyzed, these methods are not applied. It would be essential for the future development of AI systems applied to the detection of diseases such as cancer, to achieve this system explainability. As mentioned above, this would ensure the confidence of doctors and patients, who are the two actors most affected by the implementation of these systems. In addition, it would allow the widespread use of these systems by increasing users' confidence in them.

3. As explained in the section on future trends, it is expected that in the coming years, more authors will combine different medical tests to improve the performance of DL systems. This will force researchers to explore new lines of research such as ***information fusion***, a field that has great advantages [149, 150], and techniques such as ***ensembles***, which are widely used in other medical problems [79, 151]. Both approaches have shown promising results in other medical problems and in medical image analysis. Information fusion involves integrating data from various sources, such as imaging, lab results, and electronic health records, to provide a more comprehensive view of a patient's condition. This approach can lead to improved diagnostic accuracy and better-informed treatment decisions. Similarly, ensemble methods combine multiple models to enhance predictive performance, leveraging the strengths of each individual model to produce more robust and reliable results. As researchers continue to develop and refine these methods, we can expect significant advancements in the accuracy, reliability, and applicability of DL systems in clinical settings. This ongoing evolution will necessitate addressing challenges related to data integration, computational efficiency, and the handling of large-scale, heterogeneous datasets, ultimately leading to more effective and personalized healthcare solutions.
4. As explained in the section on future trends, it is expected that in the coming years, more authors will combine different medical tests to improve the performance of DL systems. This will force researchers to explore new lines of research such as ***information fusion***, a field that has great advantages [149, 150], and techniques such as ***ensembles***, which are widely used in other medical problems [79, 151]. Both approaches have shown promising results in other medical problems and in medical image analysis. Information fusion involves integrating data from various sources, such as imaging, lab results, and electronic health records, to provide a more comprehensive view of a patient's condition. This approach can lead to improved diagnostic accuracy and better-informed treatment decisions. Similarly, ensemble methods combine multiple models to enhance predictive performance, leveraging the strengths of each individual model to produce more robust and reliable results. Despite these advantages, several challenges remain to be addressed. One critical issue is the diversity of data sources, as medical imaging modalities (e.g., CT scans, X-rays, and MRIs) often have significant differences in resolution, format, and preprocessing requirements. In addition, patient demographics, such as age, ethnicity, and comorbidities, introduce variability that must be accounted for to avoid biases in the models. Furthermore, the quality of annotations is often inconsistent or incomplete. It can affect the reliability and generalizability of DL systems, highlighting the need for improved labeling tools and standardized annotation protocols. Integrating multimodal data poses additional computational challenges, requiring efficient frameworks capable of handling the complexity and scalability of large datasets. These frameworks must address the alignment and synchronization of data from disparate sources, as mismatched or missing data points can degrade model performance. Moreover, ensuring data security and compliance with healthcare regulations (e.g., HIPAA or GDPR) is critical when working with sensitive medical records. As researchers continue to develop and refine these methods, we can expect significant advancements in the accuracy, reliability, and applicability of DL systems in clinical settings. However, tackling these challenges will require interdisciplinary collaboration, involving not only AI specialists but also domain experts such as clinicians and radiologists.
5. Most authors develop cancer detection systems without ***testing models in real situations*** or with datasets different from those used during training, validation, and testing. Allowing these systems to be verified by clinicians in actual settings could significantly enhance their quality and facilitate their eventual application in real environments, such as hospitals or primary care centers.

Consequently, it is expected that there will be a trend towards more *multidisciplinary research* that includes adequate validation in the near future.

6. DL research in medicine must recognize that healthcare institutions do not operate in isolation, but are interconnected in a complex and dynamic system. It is therefore essential to adapt DL techniques to the particularities of this domain. A promising approach to overcome these challenges is ***Federated Learning*** [152, 153], which allows DL models to be trained using data distributed across multiple sites without the need for centralization, thus ensuring data privacy and security. Further research and development of these methods can facilitate the effective integration of DL into clinical practice.

Conclusion

Lung cancer is the most lethal form of cancer and the second leading cause of death worldwide. Therefore, the early detection of the disease is crucial for patient survival. DL systems, with their advanced image processing capabilities, can be a very useful tool for this purpose. Within this field, lung cancer detection can be approached from two main perspectives: classification and localization. However, we can also find numerous papers that combine both approaches with the aim of improving the performance of models and tools.

To develop DL systems that can detect lung cancer and be applied in real-life situations, quality datasets are essential. These datasets must be sufficiently varied, coming from different patients, to allow the system to generalize in real-life situations. Advances in the field of machine learning in recent years have provided us with the necessary techniques to aid in the rapid detection of lung cancer. This survey aims to achieve this objective by analyzing the following points:

1. The most widely used databases and datasets available.
2. The state-of-the-art screening techniques commonly employed.
3. The degree of application of XAI techniques within the domain.
4. An analysis of the evolution of the domain over the last 6 years.

Lung cancer detection techniques have been critically analyzed with the aim of defining the lines of research to be followed in the coming years and identifying the current challenges in the domain. From the information analyzed in this survey, we have been able to extract the main ideas of the domain to facilitate and stimulate research on lung cancer detection with DL techniques.

Appendix. Datasets download links

In this section, we have provided all the download links for the datasets mentioned (Table 3) in the survey in Table 22.

Table 22 Download links to the available datasets

Dataset	Link
LIDC-IDRI	www.cancerimagingarchive.net/collection/lidc-idri/
LUNA-16	https://luna16.grand-challenge.org/Data/
CIA	www.cancerimagingarchive.net/
RIDER	www.cancerimagingarchive.net/collection/rider-lung-ct/
CPTAC-LSCC	www.cancerimagingarchive.net/collection/cptac-lscc/
TCGA-LUAD	www.cancerimagingarchive.net/collection/tcga-luad/
JSRT	http://db.jsrt.or.jp/eng.php
DeepLesion	https://nihcc.app.box.com/v/DeepLesion
SPIE-AAPM-NCI	www.cancerimagingarchive.net/collection/spie-aapm-lung-ct-challenge/
LNDb	https://zenodo.org/records/6613714
ANODE09	https://anode09.grand-challenge.org/
NLST	https://wiki.cancerimagingarchive.net/display/NLST
DSB	https://www.kaggle.com/competitions/data-science-bowl-2017
Chest CT-Scan images Dataset	www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images/data
LC25000	https://academictorrents.com/details7a638ed187a6180fd6e464b3666a6ea0499af4af

Author Contribution HL: conceptualization, writing, methodology, review, supervision. AADH: conceptualization, writing, methodology, review. SDM: conceptualization, writing, methodology. MADM: writing, review, supervision. DC: resources, funding acquisition, supervision.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This work has been supported by the following projects: H2020 TMA-MSCA-DN TUAI project "Towards an Understanding of Artificial Intelligence via a transparent, open and explainable perspective" (HORIZON-MSCA-2023-DN-01-01, Grant agreement no: 101168344); by Strategic Networking and Development Program funded by the Ministry of Science and ICT through the National Research Foundation of Korea (RS-2023-00267476); by project PCI2022-134990-2 (MARTINI) of the CHISTERA IV Cofund 2021 program; by European Comission under IBERIFIER Plus - Iberian Digital Media Observatory (DIGITAL-2023-DEPLOY-04-EDMO-
HUBS 101158511); by EMIF managed by the Calouste Gulbenkian Foundation, in the project MuseAI; and by Comunidad Autonoma de Madrid, CIRMA-CAM Project (TEC-2024/COM-404).

Data Availability No datasets were generated or analyzed during the current study.

Declarations

Conflict of Interest The authors declare no competing interests.

References

1. Kratzer TB, Bandi P, Freedman ND, Smith RA, Travis WD, Jemal A, et al. Lung cancer statistics, 2023. Cancer. 2024;130(8):1330–48.
2. Ferlay J, Ervik M, Lam F, Colombet M, Mery L, Piñeros M, et al. Global cancer observatory: cancer today. International Agency for Research on Cancer. https://gco.iarc.fr/today.
3. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer J Clin. 2021;71(3):209–49. https://onlinelibrary.wiley.com/doi/pdf/10.3322/caac.21660. https://doi.org/10.3322/caac.21660.
4. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer J Clinicians. 2024;74(3):229–63.
5. Zhu Y, Yip R, Zhang J, Cai Q, Sun Q, Li P, et al. Radiologic features of nodules attached to the mediastinal or diaphragmatic pleura at low-dose CT for lung cancer screening. Radiology. 2024;310(1):e231219.
6. Hosseini SH, Monsefi R, Shadroo S. Deep learning applications for lung cancer diagnosis: a systematic review. Multimed Tools Appl. 2024;83(5):14305–35.
7. Martín A, Vargas VM, Gutiérrez PA, Camacho D, Hervás-Martínez C. Optimising convolutional neural networks using a hybrid statistically-driven coral reef optimisation algorithm. Appl Soft Comput. 2020;90:106144. https://doi.org/10.1016/j.asoc.2020.106144.
8. Archana R, Jeevaraj PE. Deep learning models for digital image processing: a review. Artif Intell Rev. 2024;57(1):11.
9. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436–44. Nature Publishing Group. https://doi.org/10.1038/nature14539.
10. Shwartz-Ziv R, Goldblum M, Bansal A, Bruss CB, LeCun Y, Wilson AG. Just how flexible are neural networks in practice? 2024. arXiv preprint arXiv:2406.11463.
11. Lee B, Lee J, Lee JO, Hwang Y, Bahn HK, Park I, et al. Breath analysis system with convolutional neural network (CNN) for early detection of lung cancer. Sens Actuators B Chem. 2024;409:135578.
12. Sirazitdinov I, Kholiavchenko M, Mustafaev T, Yixuan Y, Kuleev R, Ibragimov B. Deep neural network ensemble for pneumonia localization from a large-scale chest X-ray database. Comput Electr Eng. 2019;78:388–99. https://doi.org/10.1016/j.compeleceng.2019.08.004.
13. Litjens G, Sánchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, et al. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Sci Rep. 2016;6(1):26286. Nature Publishing Group. https://doi.org/10.1038/srep26286.
14. Jain A, Mandal SK, Abrol M. Exploring the potential of recurrent neural networks for medical image segmentation. In: 2024 International Conference on Optimization Computing and Wireless Communication (ICOCWC). IEEE; 2024. p. 1–6.
15. Li H, Lin Z, Shen X, Brandt J, Hua G. A convolutional neural network cascade for face detection. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2015. p. 5325–34. ISSN: 1063-6919. https://ieeexplore.ieee.org/document/7299170.
16. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: 3rd International Conference on Learning Representations (ICLR 2015). Computational and Biological Learning Society; 2015. p. 1–14. https://arxiv.org/abs/1409.1556.
17. Djenouri Y, Belhadi A, Yazidi A, Srivastava G, Lin JCW. Artificial intelligence of medical things for disease detection using ensemble deep learning and attention mechanism. Expert Syst. 2024;41(6):e13093.
18. Minetto R, Pamplona Segundo M, Sarkar S. Hydra: an ensemble of convolutional neural networks for geospatial land classification. IEEE Trans Geosci Remote Sens. 2019;57(9):6530–41. https://doi.org/10.1109/TGRS.2019.2906883.
19. Mian Z, Deng X, Dong X, Tian Y, Cao T, Chen K, et al. A literature review of fault diagnosis based on ensemble learning. Eng Appl Artif Intell. 2024;127:107357.
20. Ullah F, Ullah I, Khan RU, Khan S, Khan K, Pau G. Conventional to deep ensemble methods for hyperspectral image classification: a comprehensive survey. IEEE J Sel Top Appl Earth Obs Remote Sens. 2024.
21. Guidotti R, Monreale A, Ruggieri S, Turini F, Giannotti F, Pedreschi D. A survey of methods for explaining black box models. ACM Comput Surv. 2018;51(5):93:1-42. https://doi.org/10.1145/3236009.
22. Petch J, Bortesi JPT, Nelson W, Di S, Mamdani MH. Should I trust this model? Explainability and the black box of artificial intelligence in medicine. In: Artificial Intelligence for medicine. Elsevier; 2024. p. 265–73.
23. Marcus E, Teuwen J. Artificial intelligence and explanation: how, why, and when to explain black boxes. European J Radiology. 2024. p. 111393.
24. Holzinger A, Biemann C, Pattichis CS, Kell DB. What do we need to build explainable AI systems for the medical domain? 2017. arXiv:1712.09923. https://doi.org/10.48550/arXiv.1712.09923.
25. Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, et al. Explainable Artificial Intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf Fusion. 2020;58:82–115. https://doi.org/10.1016/j.inffus.2019.12.012.
26. Fong R, Vedaldi A. Interpretable explanations of black boxes by meaningful perturbation. In: 2017 IEEE International Conference on Computer Vision (ICCV); 2017. p. 3449–57. http://arxiv.org/abs/1704.03296.
27. Mahomed N, van Ginneken B, Philipsen RHHM, Melendez J, Moore DP, Moodley H, et al. Computer-aided diagnosis for World Health Organization-defined chest radiograph primary-

endpoint pneumonia in children. Pediatr Radiol. 2020;50(4):482–91. https://doi.org/10.1007/s00247-019-04593-0.
28. Samek W, Wiegand T, Müller KR. Explainable artificial intelligence: understanding, visualizing and interpreting deep learning models. 2017. arXiv:1708.08296. https://doi.org/10.48550/arXiv.1708.08296.
29. Domínguez-Rodríguez S, Liz-López H, Panizo-LLedot A, Ballesteros Á, Dagan R, Greenberg D, et al. Testing the performance, adequacy, and applicability of an artificial intelligence model for pediatric pneumonia diagnosis. Comput Methods Programs Biomed. 2023;242:107765.
30. Razzak MI, Naz S, Zaib A. Deep learning for medical image processing: overview, challenges and future. 2017. arXiv:1704.06825. https://doi.org/10.48550/arXiv.1704.06825.
31. Liu B, Chi W, Li X, Li P, Liang W, Liu H, et al. Evolving the pulmonary nodules diagnosis from classical approaches to deep learning-aided decision support: three decades' development course and future prospect. J Cancer Res Clin Oncol. 2020;146:153–85. https://doi.org/10.1007/s00432-019-03098-5.
32. Jassim MM, Jaber MM. Systematic review for lung cancer detection and lung nodule classification: taxonomy, challenges, and recommendation future works. J Intell Syst. 2022;31(1):944–64. Num Pages: 21 Place: Warsaw Publisher: De Gruyter Poland Sp Z O O Web of Science ID: WOS:000838249300001. https://doi.org/10.1515/jisys-2022-0062.
33. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, The PRISMA, et al. statement: an updated guideline for reporting systematic reviews. BMJ. 2020;2021:372. https://doi.org/10.1136/bmj.n71.
34. Hosseini SH, Monsefi R, Shadroo S. Deep learning applications for lung cancer diagnosis: a systematic review. Multimed Tools Appl. 2023;Num Pages: 31 Place: Dordrecht Publisher: Springer Web of Science ID: WOS:001023992100005. https://doi.org/10.1007/s11042-023-16046-w.
35. Räz T, Beisbart C. The importance of understanding deep learning. Erkenntnis. 2024;89(5):1823–40.
36. Damaševičius R, Jagatheesaperumal SK, Kandala RN, Hussain S, Alizadehsani R, Gorriz JM. Deep learning for personalized health monitoring and prediction: a review. Comput Intell. 2024;40(3):e12682.
37. Jones C, Castro DC, De Sousa Ribeiro F, Oktay O, McCradden M, Glocker B. A causal perspective on dataset bias in machine learning for medical imaging. Nat Mach Intell. 2024;6(2):138–46.
38. Lambert B, Forbes F, Doyle S, Dehaene H, Dojat M. Trustworthy clinical AI solutions: a unified review of uncertainty quantification in deep learning models for medical image analysis. Artif Intell Med. 2024. p. 102830.
39. Setio AAA, Traverso A, De Bel T, Berens MS, Van Den Bogaard C, Cerello P, et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge. Med Image Anal. 2017;42:1–13.
40. Borkowski AA, Bui MM, Thomas LB, Wilson CP, DeLand LA, Mastorides SM. Lung and colon cancer histopathological image dataset (LC25000). 2019. arXiv preprint arXiv:1912.12142.
41. The Cancer Imaging Archive (TCIA). Data from the Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans (LIDC-IDRI). https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=1966254.
42. Petrick N, Akbar S, Cha KH, Nofech-Mozes S, Sahiner B, Gavrielides MA, et al. SPIE-AAPM-NCI BreastPathQ Challenge: an image analysis challenge for quantitative tumor cellularity assessment in breast cancer histology images following neoadjuvant treatment. J Med Imaging. 2021;8(3):034501–034501. https://doi.org/10.1117/1.JMI.8.3.034501.
43. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. J Digit Imaging. 2013;26:1045–57. https://doi.org/10.1007/s10278-013-9622-7.
44. Zhao B, Schwartz LH, Kris MG. Data from RIDER lung CT. The Cancer Imaging Archive. https://www.cancerimagingarchive.net/collection/rider-lung-ct/.
45. Albertina B, Watson M, Holback C, Jarosz R, Kirk S, Lee Y, et al. The cancer genome atlas lung adenocarcinoma collection (TCGA-LUAD). The Cancer Imaging Archive. https://www.cancerimagingarchive.net/collection/tcga-luad/.
46. National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC). The Clinical Proteomic Tumor Analysis Consortium Lung Squamous Cell Carcinoma Collection (CPTAC-LSCC). The Cancer Imaging Archive. https://www.cancerimagingarchive.net/collection/cptac-lscc/.
47. Shiraishi J, Katsuragawa S, Ikezoe J, Matsumoto T, Kobayashi T, Komatsu K, et al. Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. Am J Roentgenology. 2000;174(1):71–4. https://doi.org/10.2214/ajr.174.1.1740071.
48. Yan K, Wang X, Lu L, Summers RM. DeepLesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning. J Med Imaging. 2018;5(3):036501–036501.
49. Pedrosa J, Aresta G, Ferreira C, Rodrigues M, Leitão P, Carvalho AS, et al. LNDb: a lung nodule database on computed tomography. 2019. arXiv preprint arXiv:1911.08434.
50. Van Ginneken B, Armato III SG, de Hoop B, van Amelsvoort-van de Vorst S, Duindam T, Niemeijer M, et al. Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: the ANODE09 study. Med Image Anal. 2010;14(6):707–22. https://doi.org/10.1016/j.media.2010.05.005.
51. Han Y, Qi H, Wang L, Chen C, Miao J, Xu H, et al. Pulmonary nodules detection assistant platform: an effective computer aided system for early pulmonary nodules detection in physical examination. Comput Methods Programs Biomed. 2022;217:106680. https://doi.org/10.1016/j.cmpb.2022.106680.
52. National Lung Screening Trial Research Team, Aberle DR, Berg CD, Black WC, Church TR, Fagerstrom RM, et al. The national lung screening trial: overview and study design. Radiology. 2011;258(1):243–53. https://doi.org/10.1148/radiol.10091808.
53. AJBuckeye, Kriss J, BoozAllen J, Sullivan J, O'Connell M, Nilofer, et al. Data science bowl 2017. Kaggle. https://kaggle.com/competitions/data-science-bowl-2017.
54. Many M. Chest CT-Scan images dataset. https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images/data.
55. EL-Bana S, Al-Kabbany A, Sharkas M. A two-stage framework for automated malignant pulmonary nodule detection in CT scans. Diagnostics. 2020;10(3):131. https://doi.org/10.3390/diagnostics10030131
56. Shakeel PM, Burhanuddin MA, Desa MI. Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier. Neural Comput Appl. 2022;34(12):9579–92. https://doi.org/10.1007/s00521-020-04842-6.
57. Bhat MA. Lung cancer. The data is collected from the website online lung cancer prediction system. 2024. https://www.kaggle.com/datasets/mysarahmadbhat/lung-cancer.
58. Yue Y, Li Z. Medmamba: vision mamba for medical image classification. 2024. arXiv preprint arXiv:2403.03849.
59. Helaly HA, Badawy M, Haikal AY. A review of deep learning approaches in clinical and healthcare systems based on medical image analysis. Multimed Tools Appl. 2024;83(12):36039–80.

60. El-Shafai W, Mahmoud AA, Ali AM, El-Rabaie ESM, Taha TE, El-Fishawy AS, et al. Efficient classification of different medical image multimodalities based on simple CNN architecture and augmentation algorithms. J Opt. 2024;53(2):775–87.
61. Cai L, Gao J, Zhao D. A review of the application of deep learning in medical image classification and segmentation. Ann Transl Med. 2020;8(11).
62. Elnakib A, Amer HM, Abou-Chadi FEZ. Early lung cancer detection using deep learning optimization. Int J Online Biomed Eng (iJOE). 2020;16(6):82. https://doi.org/10.3991/ijoe.v16i06.13657.
63. Alsheikhy AA, Said Y, Shawly T, Alzahrani AK, Lahza H. A CAD system for lung cancer detection using hybrid deep learning techniques. Diagnostics. 2023;13(6):1174. https://doi.org/10.3390/diagnostics13061174.
64. Wang W, Charkborty G. Automatic prognosis of lung cancer using heterogeneous deep learning models for nodule detection and eliciting its morphological features. Appl Intell. 2021;51(4):2471–84. https://doi.org/10.1007/s10489-020-01990-z.
65. Zhang Q, Kong X. Design of automatic lung nodule detection system based on multi-scene deep learning framework. IEEE Access. 2020;8:90380–9. https://doi.org/10.1109/ACCESS.2020.2993872.
66. Rafael-Palou X, Aubanell A, Bonavita I, Ceresa M, Piella G, Ribas V, et al. Re-identification and growth detection of pulmonary nodules without image registration using 3D siamese neural networks. Med Image Anal. 2021;67:101823. https://doi.org/10.1016/j.media.2020.101823.
67. Manickavasagam R, Selvan S, Selvan M. CAD system for lung nodule detection using deep learning with CNN. Med Biol Eng Comput. 2022;60(1):221–8. https://doi.org/10.1007/s11517-021-02462-3.
68. Sori WJ, Feng J, Godana AW, Liu S, Gelmecha DJ. DFD-Net: lung cancer detection from denoised CT scan image using deep learning. Front Comp Sci. 2021;15(2):152701. https://doi.org/10.1007/s11704-020-9050-z.
69. Eun H, Kim D, Jung C, Kim C. Single-view 2D CNNs with fully automatic non-nodule categorization for false positive reduction in pulmonary nodule detection. Comput Methods Programs Biomed. 2018;165:215–24. https://doi.org/10.1016/j.cmpb.2018.08.012.
70. Ali S, Asad S, Asghar Z, Ali A, Kim D. End-to-end 2D convolutional neural network architecture for lung nodule identification and abnormal detection in cloud. Computers, Materials & Continua. 2023;75(1):461–75. https://doi.org/10.32604/cmc.2023.035672.
71. Guo H, Wu J, Xie Z, Tham IWK, Zhou L, Yan J. Investigation of small lung lesion detection for lung cancer screening in low dose FDG PET imaging by deep neural networks. Front Public Health. 2022;10:1047714. https://doi.org/10.3389/fpubh.2022.1047714.
72. Lanjewar MG, Panchbhai KG, Charanarur P. Lung cancer detection from CT scans using modified DenseNet with feature selection methods and ML classifiers. Expert Syst Appl. 2023;224:119961. https://doi.org/10.1016/j.eswa.2023.119961.
73. Sajja T, Devarapalli R, Kalluri H. Lung cancer detection based on CT scan images by using deep transfer learning. Traitement du Signal. 2019;36(4):339–44. https://doi.org/10.18280/ts.360406.
74. Arshad Choudhry I, N Qureshi A. Detection of lung nodules on X-ray using transfer learning and manual features. Computers, Materials & Continua. 2022;72(1):1445–63. https://doi.org/10.32604/cmc.2022.025208.
75. Siddiqui EA, Chaurasia V, Shandilya M. Detection and classification of lung cancer computed tomography images using a novel improved deep belief network with Gabor filters. Chemom Intell Lab Syst. 2023;235:104763. https://doi.org/10.1016/j.chemolab.2023.104763.
76. Toğaçar M, Ergen B, Cömert Z. Detection of lung cancer on chest CT images using minimum redundancy maximum relevance feature selection method with convolutional neural networks. Biocybern Biomed Eng. 2020;40(1):23–39. https://doi.org/10.1016/j.bbe.2019.11.004.
77. Jain DK, Lakshmi KM, Varma KP, Ramachandran M, Bharati S. Lung cancer detection based on kernel PCA-convolution neural network feature extraction and classification by fast deep belief neural network in disease management using multimedia data sources. Comput Intell Neurosci. 2022;2022:1–12. https://doi.org/10.1155/2022/3149406.
78. Xu Z, Ren H, Zhou W, Liu Z. ISANET: non-small cell lung cancer classification and detection based on CNN and attention mechanism. Biomed Signal Process Control. 2022;77:103773. https://doi.org/10.1016/j.bspc.2022.103773.
79. Liz H, Sánchez-Montañés M, Tagarro A, Domínguez-Rodríguez S, Dagan R, Camacho D. Ensembles of convolutional neural network models for pediatric pneumonia diagnosis. Futur Gener Comput Syst. 2021;122:220–33.
80. A R B, R S VK, S S K. LCD-capsule network for the detection and classification of lung cancer on computed tomography images. Multimed Tools Appl. 2023. https://doi.org/10.1007/s11042-023-14893-1.
81. Shah AA, Malik HAM, Muhammad A, Alourani A, Butt ZA. Deep learning ensemble 2D CNN approach towards the detection of lung cancer. Sci Rep. 2023;13(1):2987. https://doi.org/10.1038/s41598-023-29656-z.
82. Jin H, Li Z, Tong R, Lin L. A deep 3D residual CNN for false-positive reduction in pulmonary nodule detection. Med Phys. 2018;45(5):2097–107. https://onlinelibrary.wiley.com/doi/pdf/10.1002/mp.12846. https://doi.org/10.1002/mp.12846.
83. Xiao Z, Du N, Geng L, Zhang F, Wu J, Liu Y. Multi-scale heterogeneous 3D CNN for false-positive reduction in pulmonary nodule detection, based on chest CT images. Appl Sci. 2019;9(16):3261. https://doi.org/10.3390/app9163261.
84. Causey J, Li K, Chen X, Dong W, Walker K, Qualls J, et al. Spatial pyramid pooling with 3D convolution improves lung cancer detection. IEEE/ACM Trans Comput Biol Bioinform. 2020. p. 1–1. https://doi.org/10.1109/TCBB.2020.3027744.
85. Nasrullah N, Sang J, Alam MS, Mateen M, Cai B, Hu H. Automated lung nodule detection and classification using deep learning combined with multiple strategies. Sensors. 2019;19(17):3722. https://doi.org/10.3390/s19173722.
86. Katase S, Ichinose A, Hayashi M, Watanabe M, Chin K, Takeshita Y, et al. Development and performance evaluation of a deep learning lung nodule detection system. BMC Med Imaging. 2022;22(1):203. https://doi.org/10.1186/s12880-022-00938-8.
87. Ali I, Hart GR, Gunabushanam G, Liang Y, Muhammad W, Nartowt B, et al. Lung nodule detection via deep reinforcement learning. Front Oncol. 2018;8:108. https://doi.org/10.3389/fonc.2018.00108.
88. Ozdemir O, Russell RL, Berlin AA. A 3D probabilistic deep learning system for detection and diagnosis of lung cancer using low-dose CT scans. IEEE Trans Med Imaging. 2020;39(5):1419–29. https://doi.org/10.1109/TMI.2019.2947595.
89. Kanipriya M, Hemalatha C, Sridevi N, SriVidhya SR, Jany Shabu SL. An improved capuchin search algorithm optimized hybrid CNN-LSTM architecture for malignant lung nodule detection. Biomed Signal Process Control. 2022;78:103973. https://doi.org/10.1016/j.bspc.2022.103973.
90. Prasad JMN, Chakravarty S, Krishna MV. Lung cancer detection using an integration of fuzzy K-means clustering and deep learning techniques for CT lung images. Int J Online Biomed Eng. 2022;70(3). https://doi.org/10.3991/ijoe.v16i06.13657.
91. Dl M, Dp M. An improved convolution neural network and modified regularized K-means-based automatic lung nodule detection

and classification. J Digit Imaging. 2023. https://doi.org/10.1007/s10278-023-00809-w.
92. Shakeel PM, Burhanuddin MA, Desa MI. Lung cancer detection from CT image using improved profuse clustering and deep learning instantaneously trained neural networks. Measurement. 2019;145:702–12. https://doi.org/10.1016/j.measurement.2019.05.027.
93. Venkatesh C, Bojja P. A dynamic optimization and deep learning technique for detection of lung cancer in CT images and data access through internet of things. Wireless Pers Commun. 2022;125(3):2621–46. https://doi.org/10.1007/s11277-022-09676-0.
94. Venkatesh C, Ramana K, Lakkisetty SY, Band SS, Agarwal S, Mosavi A. A neural network and optimization based lung cancer detection system in CT images. Front Public Health. 2022;10:769692. https://doi.org/10.3389/fpubh.2022.769692.
95. Bhattacharjee A, Murugan R, Soni B, Goel T. Ada-GridRF: a fast and automated adaptive boost based grid search optimized random forest ensemble model for lung cancer detection. Phys Eng Sci Med. 2022;45(3):981–94. https://doi.org/10.1007/s13246-022-01150-2.
96. Jiang H, Ma H, Qian W, Gao M, Li Y. An automatic detection system of lung nodule based on multigroup patch-based deep learning network. IEEE J Biomed Health Inform. 2018;22(4):1227–37. https://doi.org/10.1109/JBHI.2017.2725903.
97. Masood A, Yang P, Sheng B, Li H, Li P, Qin J, et al. Cloud-based automated clinical decision support system for detection and diagnosis of lung cancer in chest CT. IEEE J Transl Eng Health Med. 2020;8:1–13. https://doi.org/10.1109/JTEHM.2019.2955458.
98. Saba T, Sameh A, Khan F, Shad SA, Sharif M. Lung nodule detection based on ensemble of hand crafted and deep features. J Med Syst. 2019;43(12):332. https://doi.org/10.1007/s10916-019-1455-6.
99. Shen Z, Cao P, Yang J, Zaiane OR. WS-LungNet: a two-stage weakly-supervised lung cancer detection and diagnosis network. Comput Biol Med. 2023;154:106587. https://doi.org/10.1016/j.compbiomed.2023.106587.
100. Niu C, Wang G. Unsupervised contrastive learning based transformer for lung nodule detection. Phys Med Biol. 2022;67(20):204001. https://doi.org/10.1088/1361-6560/ac92ba.
101. Alaskar H, Hussain A, Almaslukh B, Vaiyapuri T, Sbai Z, Dubey AK. Deep learning approaches for automatic localization in medical images. Comput Intell Neurosci. 2022;2022(1):6347307.
102. Madhushree B, Basanth Kumar H, Chennamma H. An exhaustive review of authentication, tamper detection with localization and recovery techniques for medical images. Multimed Tools Appl. 2024;83(13):39779–821.
103. Cheng CT, Lin HH, Hsu CP, Chen HW, Huang JF, Hsieh CH, et al. Deep Learning for automated detection and localization of traumatic abdominal solid organ injuries on CT scans. J Imag Inf Med. 2024. p. 1–11.
104. Wang CH, Hwang T, Huang YS, Tay J, Wu CY, Wu MC, et al. Deep learning-based localization and detection of malpositioned endotracheal tube on portable supine chest radiographs in intensive and emergency medicine: a multicenter retrospective study. Crit Care Med. 2024;52(2):237–47.
105. Cai L, Long T, Dai Y, Huang Y. Mask R-CNN-based detection and segmentation for pulmonary nodule 3D visualization diagnosis. IEEE Access. 2020;8:44400–9. https://doi.org/10.1109/ACCESS.2020.2976432.
106. Zhang C, Li J, Huang J, Wu S. Computed tomography image under convolutional neural network deep learning algorithm in pulmonary nodule detection and lung function examination. J Healthc Eng. 2021;2021:1–9. https://doi.org/10.1155/2021/3417285.
107. Xu J, Ren H, Cai S, Zhang X. An improved faster R-CNN algorithm for assisted detection of lung nodules. Comput Biol Med. 2023;153:106470. https://doi.org/10.1016/j.compbiomed.2022.106470.
108. Chi J, Zhang S, Yu X, Wu C, Jiang Y. A novel pulmonary nodule detection model based on multi-step cascaded networks. Sensors. 2020;20(15):4301. https://doi.org/10.3390/s20154301.
109. Li X, Shen L, Xie X, Huang S, Xie Z, Hong X, et al. Multi-resolution convolutional networks for chest X-ray radiograph based lung nodule detection. Artif Intell Med. 2020;103:101744. https://doi.org/10.1016/j.artmed.2019.101744.
110. Basha BMF, Surputheen MM. Non-smokers' lung cancer bio-images detection using deep learning approach. Int J Comput Sci Netw Sec. 2022;22(3):417–24. https://doi.org/10.22937/IJCSNS.2022.22.3.53.
111. Akila Agnes S, Anitha J, Arun Solomon A. Two-stage lung nodule detection framework using enhanced UNet and convolutional LSTM networks in CT images. Comput Biol Med. 2022;149:106059. https://doi.org/10.1016/j.compbiomed.2022.106059.
112. Kvak D, Chromcová A, Hrubý R, Janů E, Biroš M, Pajdaković M, et al. Leveraging deep learning decision-support system in specialized oncology center: a multi-reader retrospective study on detection of pulmonary lesions in chest X-ray images. Diagnostics. 2023;13(6):1043. https://doi.org/10.3390/diagnostics13061043.
113. Pezeshk A, Hamidian S, Petrick N, Sahiner B. 3-D convolutional neural networks for automatic detection of pulmonary nodules in chest CT. IEEE J Biomed Health Inform. 2019;23(5):2080–90. https://doi.org/10.1109/JBHI.2018.2879449.
114. Gong L, Jiang S, Yang Z, Zhang G, Wang L. Automated pulmonary nodule detection in CT images using 3D deep squeeze-and-excitation networks. Int J Comput Assist Radiol Surg. 2019;14(11):1969–79. https://doi.org/10.1007/s11548-019-01979-1.
115. Gu Y, Lu X, Yang L, Zhang B, Yu D, Zhao Y, et al. Automatic lung nodule detection using a 3D deep convolutional neural network combined with a multi-scale prediction strategy in chest CTs. Comput Biol Med. 2018;103:220–31. https://doi.org/10.1016/j.compbiomed.2018.10.011.
116. Zheng S, Cornelissen LJ, Cui X, Jing X, Veldhuis RNJ, Oudkerk M, et al. Deep convolutional neural networks for multiplanar lung nodule detection: improvement in small nodule identification. Med Phys. 2021;48(2):733–44. https://doi.org/10.1002/mp.14648.
117. Xiao Y, Wang X, Li Q, Fan R, Chen R, Shao Y, et al. A cascade and heterogeneous neural network for CT pulmonary nodule detection and its evaluation on both phantom and patient data. Comput Med Imaging Graph. 2021;90:101889. https://doi.org/10.1016/j.compmedimag.2021.101889.
118. Suzuki K, Otsuka Y, Nomura Y, Kumamaru KK, Kuwatsuru R, Aoki S. Development and validation of a modified three-dimensional u-net deep-learning model for automated detection of lung nodules on chest CT images from the lung image database consortium and Japanese datasets. Acad Radiol. 2022;29:S11–7. https://doi.org/10.1016/j.acra.2020.07.030.
119. Gürsoy Çoruh A, Yenigün B, Uzun Ç, Kahya Y, Büyükceran EU, Elhan A, et al. A comparison of the fusion model of deep learning neural networks with human observation for lung nodule detection and classification. Br J Radiol. 2021;94(1123):20210222. https://doi.org/10.1259/bjr.20210222.
120. Lin J, She Q, Chen Y. Pulmonary nodule detection based on IR-UNet + +. Med Biol Eng Comput. 2023;61(2):485–95. https://doi.org/10.1007/s11517-022-02727-5.
121. Naqi SM, Sharif M, Jaffar A. Lung nodule detection and classification based on geometric fit in parametric form and deep

learning. Neural Comput Appl. 2020;32(9):4629–47. https://doi.org/10.1007/s00521-018-3773-x.
122. Zhang M, Kong Z, Zhu W, Yan F, Xie C. Pulmonary nodule detection based on 3D feature pyramid network with incorporated squeeze-and-excitation-attention mechanism. Concurr Comput Pract Exp. 2021. https://doi.org/10.1002/cpe.6237.
123. Barbouchi K, El Hamdi D, Elouedi I, Aïcha TB, Echi AK, Slim I. A transformer-based deep neural network for detection and classification of lung cancer via PET/CT images. Int J Imaging Syst Technol. 2023.
124. Yang L, Cai H, Luo X, Wu J, Tang R, Chen Y, et al. A lightweight neural network for lung nodule detection based on improved ghost module. Quant Imaging Med Surg. 2023;0(0):0–0. https://doi.org/10.21037/qims-21-1182.
125. Mkindu H, Wu L, Zhao Y. 3D multi-scale vision transformer for lung nodule detection in chest CT images. SIViP. 2023;17(5):2473–80. https://doi.org/10.1007/s11760-022-02464-0.
126. Mkindu H, Wu L, Zhao Y. Lung nodule detection of CT images based on combining 3D-CNN and squeeze-and-excitation networks. Multimed Tools Appl. 2023. https://doi.org/10.1007/s11042-023-14581-0.
127. Zhang J, Xia Y, Zeng H, Zhang Y. NODULe: Combining constrained multi-scale LoG filters with densely dilated 3D deep convolutional neural network for pulmonary nodule detection. Neurocomputing. 2018;317:159–67. https://doi.org/10.1016/j.neucom.2018.08.022.
128. Lu X, Nanehkaran YA, Karimi Fard M. A method for optimal detection of lung cancer based on deep learning optimized by marine predators algorithm. Comput Intell Neurosci. 2021;2021:1–10. https://doi.org/10.1155/2021/3694723.
129. Tian Y, Pu LZCT, Liu Y, Maicas G, Verjans JW, Burt AD, et al. Detecting, localizing and classifying polyps from colonoscopy videos using deep learning. In: Deep learning for medical image analysis. Elsevier; 2024. p. 425–50.
130. Khan MA, Muhammad K, Sharif M, Akram T, Kadry S. Intelligent fusion-assisted skin lesion localization and classification for smart healthcare. Neural Comput Appl. 2024;36(1):37–52.
131. Fan W, Yang Y, Qi J, Zhang Q, Liao C, Wen L, et al. A deep-learning-based framework for identifying and localizing multiple abnormalities and assessing cardiomegaly in chest X-ray. Nat Commun. 2024;15(1):1347.
132. Biswas J, Mustaquim S, Hossain SS, Siddique IM. Instantaneous classification and localization of eye diseases via artificial intelligence. Eur J Adv Eng Technol. 2024;11(3):45–53.
133. Majidpourkhoei R, Alilou M, Majidzadeh K, Babazadehsangar A. A novel deep learning framework for lung nodule detection in 3d CT images. Multimed Tools Appl. 2021;80(20):30539–55. https://doi.org/10.1007/s11042-021-11066-w.
134. Tiwari L, Raja R, Awasthi V, Miri R, Sinha GR, Alkinani MH, et al. Detection of lung nodule and cancer using novel Mask-3 FCM and TWEDLNN algorithms. Measurement. 2021;172:108882. https://doi.org/10.1016/j.measurement.2020.108882.
135. De Pinho Pinheiro CA, Nedjah N, De Macedo Mourelle L. Detection and classification of pulmonary nodules using deep learning and swarm intelligence. Multimed Tools Appl. 2020;79(21):15437–65. https://doi.org/10.1007/s11042-019-7473-z.
136. Li Y, Zhang L, Chen H, Yang N. Lung nodule detection with deep learning in 3D thoracic MR images. IEEE Access. 2019;7:37822–32. https://doi.org/10.1109/ACCESS.2019.2905574.
137. Nguyen CC, Tran GS, Nguyen VT, Burie JC, Nghiem TP. Pulmonary nodule detection based on faster R-CNN with adaptive anchor box. IEEE Access. 2021;9:154740–51. https://doi.org/10.1109/ACCESS.2021.3128942.
138. Chen Y, Hou X, Yang Y, Ge Q, Zhou Y, Nie S. A novel deep learning model based on multi-scale and multi-view for detection of pulmonary nodules. J Digit Imaging. 2022;36(2):688–99. https://doi.org/10.1007/s10278-022-00749-x.
139. Su Y, Li D, Chen X. Lung nodule detection based on faster R-CNN framework. Comput Methods Programs Biomed. 2021;200:105866. https://doi.org/10.1016/j.cmpb.2020.105866.
140. Chiu HY, Peng RHT, Lin YC, Wang TW, Yang YX, Chen YY, et al. Artificial intelligence for early detection of chest nodules in X-ray images. Biomedicines. 2023;10(11):2839. https://doi.org/10.3390/biomedicines10112839.
141. Ma H, Guo H, Zhao M, Qi S, Li H, Tian Y, et al. Automatic pulmonary ground-glass opacity nodules detection and classification based on 3D neural network. Med Phys. 2022;49(4):2555–69. https://doi.org/10.1002/mp.15501.
142. Chenyang L, Chan SC. A joint detection and recognition approach to lung cancer diagnosis from CT images with label uncertainty. IEEE Access. 2020;8:228905–21. https://doi.org/10.1109/ACCESS.2020.3044941.
143. Nam JG, Park S, Hwang EJ, Lee JH, Jin KN, Lim KY, et al. Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs. Radiology. 2019;290(1):218–28. https://doi.org/10.1148/radiol.2018180237.
144. Ali S, Abuhmed T, El-Sappagh S, Muhammad K, Alonso-Moral JM, Confalonieri R, et al. Explainable Artificial Intelligence (XAI): what we know and what is left to attain Trustworthy Artificial Intelligence. Inf Fusion. 2023;99:101805.
145. Celik G. Detection of Covid-19 and other pneumonia cases from CT and X-ray chest images using deep learning based on feature reuse residual block and depthwise dilated convolutions neural network. Appl Soft Comput. 2023;133:109906.
146. Bustos A, Pertusa A, Salinas JM, De La Iglesia-Vaya M. Padchest: a large chest x-ray image dataset with multi-label annotated reports. Med Image Anal. 2020;66:101797.
147. Liz-López H, Huertas-Tato J, Camacho D. Transparency in medicine: how eXplainable AI is revolutionizing patient care. In: 2023 International Conference on Network, Multimedia and Information Technology (NMITCON). IEEE; 2023. p. 1–6.
148. Arrieta AB, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, et al. Explainable Artificial Intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf Fusion. 2020;58:82–115.
149. Wu Y, Dravid A, Wehbe RM, Katsaggelos AK. DeepCOVID-Fuse: a multi-modality deep learning model fusing Chest X-rays and clinical variables to predict COVID-19 risk levels. Bioengineering. 2023;10(5):556.
150. Kim D. CheXFusion: effective fusion of multi-view features using transformers for long-tailed chest X-Ray classification. In: Proceedings of the IEEE/CVF international conference on computer vision; 2023. p. 2702–10.
151. Liz H, Huertas-Tato J, Sánchez-Montañés M, Del Ser J, Camacho D. Deep learning for understanding multilabel imbalanced Chest X-ray datasets. Futur Gener Comput Syst. 2023;144:291–306.
152. Guan H, Yap PT, Bozoki A, Liu M. Federated learning for medical image analysis: a survey. Pattern Recognit. 2024. p. 110424.
153. Rauniyar A, Hagos DH, Jha D, Håkegård JE, Bagci U, Rawat DB, et al. Federated learning for medical applications: a taxonomy, current trends, challenges, and future research directions. IEEE Internet Things J. 2023.
154. World Health Organization. Global cancer burden growing, amidst mounting need for services. https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing--amidst-mounting-need-for-services. Accessed 10 Dec 2024.
155. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. CA: A Cancer J Clin. 2023;73(1):17–48.

156. Ihongbe E, Fouad S, Mahmoud TF, Rajasekaran A, Bhatia B. Evaluating Explainable Artificial Intelligence (XAI) techniques in chest radiology imaging through a human-centered lens. PLoS ONE. 2024;19(10):e0308758.
157. Hamida SU, Chowdhury MJM, Chakraborty NR, Biswas K, Sami SK. Exploring the landscape of explainable artificial intelligence (XAI): a systematic review of techniques and applications. Big Data Cogn Comput. 2024;8(11):149.
158. Vimbi V, Shaffi N, Mahmud M. Interpreting artificial intelligence models: a systematic review on the application of LIME and SHAP in Alzheimer's disease detection. Brain Inf. 2024;11(1):10.
159. Ahmed S, Kaiser MS, Hossain MS, Andersson K. A comparative analysis of LIME and SHAP interpreters with explainable ML-based diabetes predictions. IEEE Access. 2024.
160. Idroes GM, Noviandy TR, Emran TB, Idroes R. Explainable deep learning approach for Mpox skin lesion detection with Grad-CAM. Heca J Appl Sci. 2024;2(2):54–63.
161. Cerekci E, Alis D, Denizoglu N, Camurdan O, Seker ME, Ozer C, Hansu MY, Tanyel T, Oksuz I, Karaarslan E. Quantitative evaluation of Saliency-Based Explainable artificial intelligence (XAI) methods in Deep Learning-Based mammogram analysis. Eur J Radiol. 2024;173:111356.
162. Azad R, Kazerouni A, Heidari M, Aghdam EK, Molaei A, Jia Y, Jose A, Roy R, Merhof D. Advances in medical image analysis with vision transformers: a comprehensive review. Med Image Anal. 2024;91: 103000.
163. Rahman MM, Marculescu R. Multi-scale hierarchical vision transformer with cascaded attention decoding for medical image segmentation. In: Medical imaging with deep learning. PMLR; 2024. p. 1526–44.
164. Thirunavukarasu R, Kotei E. A comprehensive review on transformer network for natural and medical image analysis. Comput Sci Rev. 2024;53:100648.