



Review

Deep Machine Learning for Medical Diagnosis, Application to Lung Cancer Detection: A Review

Hadrien T. Gayap and Moulay A. Akhloufi *

Perception, Robotics, and Intelligent Machines (PRIME), Department of Computer Science, Université de Moncton, Moncton, NB E1A3E9, Canada; ehg5435@umoncton.ca

* Correspondence: moulay.akhloufi@umoncton.ca

Abstract: Deep learning has emerged as a powerful tool for medical image analysis and diagnosis, demonstrating high performance on tasks such as cancer detection. This literature review synthesizes current research on deep learning techniques applied to lung cancer screening and diagnosis. This review summarizes the state-of-the-art in deep learning for lung cancer detection, highlighting key advances, limitations, and future directions. We prioritized studies utilizing major public datasets, such as LIDC, LUNA16, and JSRT, to provide a comprehensive overview of the field. We focus on deep learning architectures, including 2D and 3D convolutional neural networks (CNNs), dual-path networks, Natural Language Processing (NLP) and vision transformers (ViT). Across studies, deep learning models consistently outperformed traditional machine learning techniques in terms of accuracy, sensitivity, and specificity for lung cancer detection in CT scans. This is attributed to the ability of deep learning models to automatically learn discriminative features from medical images and model complex spatial relationships. However, several challenges remain to be addressed before deep learning models can be widely deployed in clinical practice. These include model dependence on training data, generalization across datasets, integration of clinical metadata, and model interpretability. Overall, deep learning demonstrates great potential for lung cancer detection and precision medicine. However, more research is required to rigorously validate models and address risks. This review provides key insights for both computer scientists and clinicians, summarizing progress and future directions for deep learning in medical image analysis.



Citation: Gayap, H.T.; Akhloufi, M.A. Deep Machine Learning for Medical Diagnosis, Application to Lung Cancer Detection: A Review. *BioMedInformatics* **2024**, *4*, 236–284. <https://doi.org/10.3390/biomedinformatics4010015>

Academic Editor: Hans Binder

Received: 30 October 2023

Revised: 6 January 2024

Accepted: 10 January 2024

Published: 18 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: lung cancer; deep learning; diagnosis; database; computed tomography; chest-ray

1. Introduction

Cancer represents a predominant cause of mortality globally [1], and early detection is pivotal for enhancing patient outcomes. Conventional techniques for cancer diagnosis, including visual examination and biopsy [2], can be time-consuming, subjective, and susceptible to mistakes. Deep machine learning (ML) constitutes a potent recent instrument with the capability to transform cancer diagnosis.

Deep ML algorithms are trained on large datasets of medical images, text, and other data. This allows them to learn to identify patterns that are invisible to the human eye. In recent years, deep ML algorithms have been shown to achieve state-of-the-art performance (SOAP) in cancer detection, often outperforming human experts [3]. This has led to a number of benefits, including: Increased accuracy of cancer detection: Deep ML algorithms can identify cancer with greater accuracy than traditional methods. This can lead to earlier detection and treatment, which can improve patient outcomes. Reduced cost of cancer diagnosis: Deep ML algorithms can analyze large datasets of medical images and data, reducing the cost of cancer diagnosis. Improved patient experience: Deep ML algorithms can automate cancer diagnosis, reducing the time and stress that patients experience.

Despite the numerous advantages of deep machine learning, certain challenges remain to be tackled. These challenges include: Data availability: Deep ML algorithms necessitate

substantial datasets of medical imagery and information to train. This can pose a difficulty, as not every medical institution has access to such datasets. Bias: Deep ML algorithms may exhibit bias if trained on datasets not illustrative of the population. This can result in imprecise cancer detection. Explainability: Explaining the decision-making of deep ML algorithms can be arduous. This can engender distrust in the outcomes of deep ML algorithms.

However, the field of deep learning has witnessed remarkable advancements, driven by the availability of vast amounts of data, significant computational power, and breakthroughs in neural network architectures. This progress has paved the way for the application of deep learning techniques in various domains, including medical diagnosis [4]. Within the realm of medical imaging, deep learning algorithms have demonstrated exceptional capabilities in detecting and classifying diseases, particularly in cancer diagnosis. The unique ability of deep learning algorithms to automatically learn intricate patterns and features from complex datasets has enabled the development of robust and accurate models for cancer detection. By training on large-scale datasets of medical images, these algorithms can discern subtle nuances indicative of cancerous growth, facilitating early detection and intervention.

Moreover, deep learning techniques have transcended traditional image-based approaches by incorporating multi-modal data sources. By integrating medical images, clinical reports, genomics, and other patient-related information, deep learning algorithms can extract comprehensive features and provide a holistic view of cancer diagnosis. This integration of diverse data modalities not only enhances the accuracy of detection but also contributes to a more personalized and precise approach to cancer management.

Despite the rapid progress in this field, several challenges need to be addressed to ensure the widespread adoption and reliability of deep learning in cancer diagnosis. These challenges include the need for standardized protocols, validation on diverse populations, interpretation of deep learning model outputs, and integration into clinical workflows.

Several recent reviews have explored the application of AI in lung cancer detection. A comprehensive study by [5] underscored the use of AI in lung cancer screening via CXR and chest CT, highlighting the FDA-approved AI programs that are revolutionizing detection methods. Dodia et al. [6] provides an overview of lung cancer, along with publicly available benchmark datasets for research purposes. It also compares recent research performed in medical image analysis of lung cancer using deep learning algorithms, considering various technical aspects such as efficiency, advantages, and limitations. Ref. [7] provides an overview of recent state-of-the-art deep learning algorithms and architectures proposed as computer-aided diagnosis (CAD) systems for lung cancer detection in CT scans. The authors divide the CAD systems into two categories: nodule detection systems and false positive reduction systems. They discuss the main characteristics of the different techniques and analyze their performance. Another review [8] provided an in-depth examination of AI in improving nodule detection and classification, emphasizing the significant role of neural networks in early-stage detection. A further review [9] summarized various AI algorithm applications, including Natural Language Processing (NLP), Machine Learning, and Deep Learning, elucidating their value in early diagnosis and prognosis. Qureshi et al. [10], highlighted specific advances in deep learning, such as AlphaFold2 and deep generative models of deep learning with in understanding drug resistance mechanisms in the treatment of non-small cell lung cancer (NSCLC). Al-Tashi et al. [11], reviewed the works that have been carried out to highlight the contribution of deep learning, thanks to its ability to process large quantities of data and identify complex patterns for the distinction between prognostic, and predictive biomarkers in personalized medicine.

The main contributions of this paper diverge from existing literature in several key aspects:

- **A Unique Grouping of Articles:** Unlike traditional reviews that often categorize research based on methods or results, this review adopts a distinctive approach by grouping articles based on the types of databases utilized, offering a fresh perspective

on the research landscape. This approach allows for a more comprehensive and nuanced understanding of the research landscape. By examining the types of databases that researchers are using, it's therefore possible to gain insights into the availability, quality, and representativeness of data and the limitations of the current research. This information can be used to identify gaps in the literature, suggest new avenues for research, and develop more robust, and reliable research methods.

- **Detailed Presentation of Widely-Used Databases:** This review provides an extensive examination of the most commonly used databases in lung cancer detection, shedding light on their features, applications, and significance.
- **Integration of Recent Research:** In a rapidly evolving field, this review incorporates recent articles on the subject, ensuring a contemporary understanding of the latest developments and trends.

This paper is organized into distinct sections to provide a comprehensive review of AI in lung cancer diagnosis. The article begins in Section 2 with a detailed presentation of the methodology used for the selection and collection of articles, and specifically outlines the methodological approach, presenting the steps taken, the search strategies used, and the criteria for inclusion and exclusion. Following that, Section 3 describes the metrics used for the evaluation of the results models in the selected articles. Section 4 focuses on the public databases used in these selected works. Section 5, is divided into two sub-sections. Section 5.1 presents articles that have employed public databases in their research. Each of these studies is described in detail, highlighting their methods, findings, and implications for the field. Meanwhile, Section 5.2 focuses on articles that have utilized private databases. Emphasis is placed on those that have incorporated segmentation techniques, classification methods, and Natural Language Processing (NLP). Finally, Section 6 critically analyses how these different databases and techniques have influenced the results and how they can potentially contribute to advancements in the field of AI for lung cancer diagnosis.

2. Methodology

The main objective was to gather relevant articles encompassing the keywords “lung cancer”, “deep learning”, “transformer”, “NLP”, “diagnosis”, “machine learning”, “chest” and “computed tomography”.

To perform this review of literature on deep machine learning for medical diagnosis, with a specific focus on its application to lung cancer detection, a systematic methodology was followed. The search process commenced by utilizing various academic databases, including PubMed, IEEE Xplore, and Google Scholar. A combination of the above keywords was used to retrieve a wide range of articles. In particular, these combinations were: “Lung cancer + deep learning + transformer + diagnosis”, “lung cancer + Machine learning + chest + Computed Tomography + diagnosis”, “Lung cancer + deep learning + diagnosis + chestray”, “Lung cancer + deep learning + diagnosis + NLP”, “Deep Learning lung cancer”. Also, the search was not constrained by date range, encompassing both recent and older publications to provide a comprehensive overview of the topic.

The initial search produced a large number of articles, which were then refined by scrutinizing their titles, abstracts and keywords for relevance. The selection process involved eliminating articles that were not directly related to deep machine learning for lung cancer diagnosis or that did not use relevant techniques, such as transformers and machine learning algorithms.

Following this selection process, a final set of articles was obtained for the literature review, ensuring a diverse representation of studies addressing the intersection of deep learning, lung cancer and medical diagnosis. The PRISMA diagram [12] in Figure 1 summarizes the successive stages of our literature search, from the initial identification of articles to the final selection of studies included in this review.

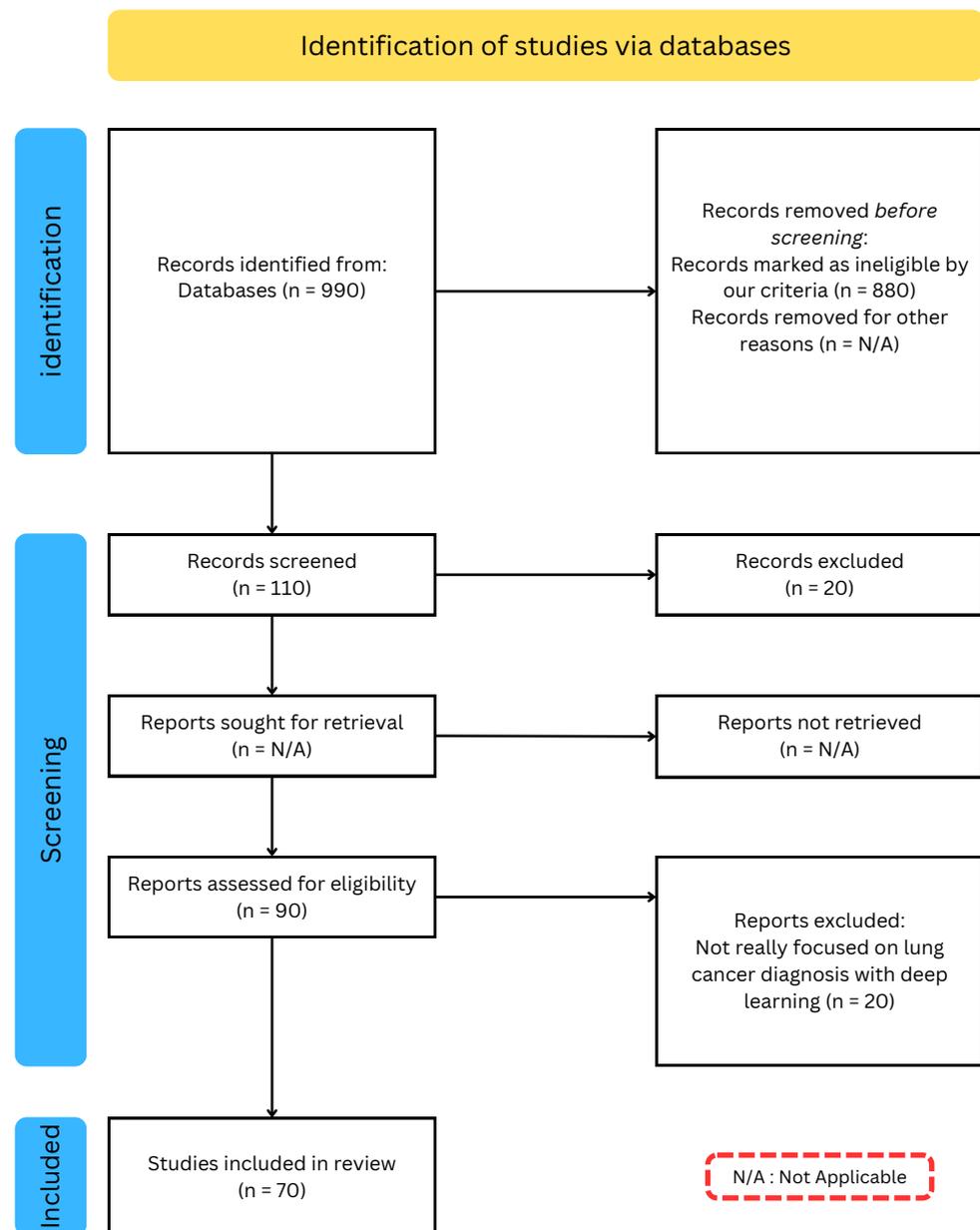


Figure 1. PRISMA diagram: Systematic Selection Process for our Literature Review.

3. Performance Metrics

Evaluation metrics play an important role in measuring the performance of these algorithms and determining their clinical utility [13]. Some commonly used evaluation metrics include Accuracy, Sensitivity [14], Specificity [14], Receiver Operating Characteristic curves (ROC) [15], Precision and Recall [16], F1-score [17] and Area Under -ROC-Curve (AUC) [18], Dice Similarity Coefficient (DSC) [19], Intersection over Union (IOU). Appendix A Table A1 presents metrics commonly used in lung cancer diagnosis for different tasks.

Therefore, the choice of metrics plays a crucial role in the development and evaluation of deep learning models, particularly in the medical domain. Selecting the appropriate metrics can help ensure that models are accurate, reliable, and clinically relevant. Inaccurate or poorly chosen metrics can lead to flawed models that are unsuitable for clinical use. In the context of lung cancer diagnosis, the use of segmentation metrics such as Dice Similarity Coefficient and Intersection over Union can help assess the accuracy of segmentation masks and improve the localization of lung tumors. Similarly, classification metrics such as

Classification Accuracy and F1-score can be used to evaluate the overall performance of models in detecting lung tumors. However, it is important to note that different metrics may have different strengths and weaknesses and may be better suited to specific applications or datasets. Finally, it is essential to carefully consider the selection of metrics in the development and evaluation of deep learning models for lung cancer diagnosis and other medical applications.

4. Datasets

In the field of deep learning for lung cancer diagnosis, having access to high-quality datasets is crucial for developing accurate and reliable models [20]. These datasets typically consist of medical images, as well as associated clinical data, such as patient demographics and medical history. By training deep learning models on these datasets, researchers and clinicians can improve their ability to accurately detect and diagnose lung cancer, ultimately improving patient outcomes. By understanding the strengths and limitations of each dataset, researchers and clinicians can make informed decisions about which dataset is most appropriate for their specific research or clinical needs. Table 1 summarizes the publicly available databases used for lung cancer diagnosis in the reviewed articles.

Table 1. Databases used for lung cancer diagnosis.

Ref.	Dataset Description	Number of Images	Data Type	File Format	Label Type
[21]	LIDC-IDRI is a comprehensive collection of lung CT scans and annotations designed to support the development of computer-aided diagnostic systems for lung cancer.	1010 patients	CT	Nifti, DICOM	Detection
[22]	LUNA16 derived from the LIDC database, specializing in lung nodule detection.	888 patients, 1186 nodules	CT	DICOM	Detection Segmentation
[23]	ChestX-Ray8 aims to enable the detection and localization of diseases, by providing a large-scale database annotated with NLP and specialists for clinical challenges.	108,948 front view X-ray images of 32,717 patients	Chest X-ray	DICOM	Classification Localization
[24]	JSRT is composed of digital radiographs of Japanese patients with a resolution of 2048×2048 pixels, provides detailed annotations of lung nodules.	247 chest radiographs	CT	Big-endian raw	Segmentation Classification
[25]	NLST contains data from a large clinical trial conducted by the National Cancer Institute to assess low-dose CT screening for lung cancer.	Information on 53,000 individuals	CT, Chest X-ray	DICOM	Classification
[26]	MSKCC contains digitized histopathology slide images with resolution of 512×512 pixels and an 8-bit color depth from patients with different cancer types that have been manually annotated by pathologists to identify regions of cancerous tissue.	25,000 digitized slides, age, sex, cancer type	CT	TSV, SEG	Prediction

Table 1. Cont.

Ref.	Dataset Description	Number of Images	Data Type	File Format	Label Type
[27]	TCGA includes subtypes such as lung adenocarcinoma and lung squamous cell carcinoma, and is designed to support cancer research and diagnostic and therapeutic development.	Data for over 11,429 patient sample	CT	DCM	Classification
[28]	Managed by the National Cancer Institute, the SEER dataset includes more than 40 years of cancer incidence and survival data from registries across the U.S., analyzing cancer trends, risk factors and treatment outcomes to inform cancer research and guide health policy decisions.	9.5 million cancer cases	CT	N/A	Segmentation Classification
[29]	MSD contains 3D medical images covering 10 imaging modalities and 10 anatomical structures to support research and evaluation of segmentation algorithms in medical image analysis. It was used to develop the MSD Challenge, which tests the ability of machine learning algorithms to generalize to various semantic segmentation tasks.	2633 scans	CT, MRI, US	Nifti	Segmentation
[30]	TCIA (Lung) has been created to promote research in the field of medical image analysis, in particular in the development and evaluation of computer-assisted diagnostic systems for cancer offering a wide variety of images.	48,723 scans	CT	Nifti, DICOM	Segmentation Classification
[31]	Tianchi encourages early detection of cancer, by providing various medical images with radiologist labels indicating the presence, location, size and malignancy of the nodule. The dataset contains various imaging parameters and patient demographics to extract lung nodule characteristics.	800 scans	CT	MHD	Detection
[32]	DSB contained more 256×256 pixel grayscale microscope images of cells, labeled with annotations indicating the presence of cancer, to challenge participants to develop image recognition algorithms for cancer detection. The various images of cell cultures, tissues and blood samples presented complexities such as variation in cell size, shape and texture for the classification task.	800 scans	CT	DICOM	Classification

5. Deep Learning Approach for Lung Cancer Diagnosis

As deep learning techniques continue to revolutionize the field of medical imaging, researchers have increasingly turned to large-scale databases to train and validate their algorithms. Many studies have been done to diagnose lung cancer using different datasets, both public and private. Each dataset has its own unique characteristics and challenges. To provide a comprehensive overview of the state-of-the-art in this field, this section presents a review of the relevant literature organized by the databases used in each study. By examining the approaches and results of each study in turn, we aim to identify common trends, best practices, and areas for future research in the use of deep learning for lung cancer diagnosis.

5.1. Deep Learning Techniques Using Public Databases

This section of the review aims to provide a comprehensive overview of research studies focused on the segmentation, classification, and detection of Regions of Interest (ROIs) in lung cancer using data from public databases. Appendix A Table A2 summarizes the reviewed methods for lung cancer diagnosis using Public Datasets.

5.1.1. Deep Learning Techniques for Lung Cancer Using LIDC Dataset

One of the most widely used databases for evaluating deep learning algorithms in the context of lung cancer diagnosis is the LIDC dataset. Since its release in 2011, the LIDC dataset has been used in numerous studies to develop and test deep learning algorithms for automated nodule detection, classification, and segmentation. In this section, we review a selection of studies that have utilized the LIDC database to train and validate deep learning models for lung cancer diagnosis. By examining the methods and results of these studies, we aim to identify key insights and challenges in using the LIDC dataset.

Da et al. [33] explored the performance of deep transfer learning for the classification of lung nodules malignancy. The study utilized CNN including VGG16, VGG19, MobileNet, Xception, InceptionV3, ResNet50, Inception-ResNet-V2, DenseNet169, DenseNet201, NASNetMobile, and NASNetLarge, and then classified the deep features returned using various classifiers including Naive Bayes, MultiLayer Perceptron (MLP), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest (RF). The best combination of deep extractor and classifier was found to be CNN-ResNet50 with SVM-RBF achieving an accuracy of 88.41% and an AUC of 93.19%. These results are comparable to related works even using only a CNN pre-trained on non-medical images. The study therefore proved that deep transfer learning is a relevant strategy for extracting representative imaging biomarkers for the classification of lung nodule malignancy in thoracic CT images.

Song et al. [34] addressed the problem of inaccurate lung cancer diagnosis due to the experience of physicians. To improve the accuracy of lung cancer diagnosis, deep learning techniques were used in medical imaging. Specifically, the study compared the prediction performance of three deep neural networks (CNN, DNN, and SAE) for classifying benign and malignant pulmonary nodules on CT images. The results showed that the CNN network outperformed the other two networks with an accuracy of 84.15%, sensitivity of 83.96%, and specificity of 84.32%. The study indicated that the proposed method can be generalized for other medical imaging tasks to design high-performance CAD systems in the future.

Zhang et al. [35] aimed to tackle the problem of classifying lung nodules in 3D CT images for computer-aided diagnosis (CAD) systems. The early detection of lung nodules is critical for improving the survival rate of lung cancer patients. The proposed approach involved applying DCNN to perform an end-to-end classification of raw 3D nodule CT patches, eliminating the need for nodule segmentation and feature extraction in the CAD system. State-of-the-art CNN models such as VGG16, VGG19, ResNet50, DenseNet121, MobileNet, Xception, NASNetMobile, and NASNetLarge were modified to 3D-CNN models for this study. Experimental results showed that DenseNet121 and Xception achieved the

best results for lung nodule diagnosis in terms of accuracy (87.77%), specificity (92.38%), precision (87.88%), and AUC (93.79%).

Shetty et al. [36] presented a new technique for accurate segmentation and classification of lung cancer using CT images by applying optimized deformable models and deep learning techniques. The proposed method involved pre-processing, lung lobe segmentation, lung cancer segmentation, data augmentation, and lung cancer classification. In the pre-processing step, median filtering was used, while Bayesian fuzzy clustering was applied for segmenting the lung lobes. The lung cancer segmentation was carried out using Water Cycle Sea Lion Optimization (WSL_{NO}) based deformable model. To improve the classification accuracy, the data augmentation process was used, which involved augmenting the size of the segmented region. The lung cancer classification was done effectively using Shepard Convolutional Neural Network (ShCNN), which was trained by WSL_{NO} algorithm. The proposed WSL_{NO} algorithm was designed by incorporating Water cycle algorithm (WCA) and Sea Lion Optimization (SL_{NO}) algorithm. The proposed technique showed improved performance in terms of accuracy, sensitivity, specificity, and average segmentation accuracy. The average segmentation accuracy achieved was 0.9091, while the accuracy, sensitivity, and specificity values were 0.9303, 0.9123, and 0.9133, respectively. The combination of optimized deformable model-based segmentation and deep learning techniques proved to be effective in accurately detecting and classifying lung cancer using CT images.

Brocki et al. [37] highlight the limitations of deep neural networks (DNNs) in clinical applications for cancer diagnosis and prognosis due to their lack of interpretability. To address this issue, the authors propose ConRad, an interpretable classifier that combines expert-derived radiomics and DNN-predicted biomarkers for CT scans of lung cancer. The proposed model is evaluated using CT images of lung tumors, and compared to CNNs acting as a black box classifier. The ConRad models using nonlinear SVM and logistic regression with Lasso outperforms the others in five-fold cross-validation, with the interpretability of ConRad being its primary advantage. The increased transparency of the ConRad model allows for better-informed diagnoses by radiologists and oncologists and can potentially help in discovering critical failure modes of black box classifiers. However, the study's limitations include the focus on a single dataset, and the lack of external validation, which warrants further investigation. Nonetheless, the proposed model demonstrates the potential for the broader incorporation of explainable AI into radiology and oncology, with the code available on [38] for reproducibility. In Figure 2 presents the different steps used to obtain these results.

Hua et al. [39] aim to simplify the image analysis pipeline of conventional CAD with deep learning techniques for differentiating a pulmonary nodule on CT images. The authors introduce two deep learning models, namely, a deep belief network (DBN) and a CNN, in the context of nodule classification, and compare them with two baseline methods that involve feature computing steps. The LIDC dataset is used for classification of malignancy of lung nodules without computing the morphology and texture features. The experimental results indicate that the proposed deep learning framework outperforms conventional hand-crafted feature computing CAD frameworks.

Khademi et al. [40] proposed a novel hybrid discovery Radiomics framework that integrates temporal and spatial features extracted from non-thin chest CT slices to predict Lung Adenocarcinoma (LUAC) malignancy with minimum expert involvement. The proposed hybrid transformer-based framework consisted of two parallel paths: The first was the Convolutional Auto-Encoder (CAE) Transformer path, which extracted and captured informative features related to inter-slice relations via a modified Transformer architecture, and the second the Shifted Window (SWin) Transformer path, which extracted nodules related spatial features from a volumetric CT scan. The extracted temporal and spatial features were then fused through a fusion path to classify LUACs. The proposed CAET-SWin model combined spatial and temporal features extracted by its two constituent parallel paths (the CAET and SWin paths) designed based on the self-attention mechanism. The

experimental results on a dataset of 114 pathologically proven Sub-Solid Nodules (SSNs) showed that the CAET-SWin significantly improved reliability of the invasiveness prediction task while achieving an accuracy of 82.65%, sensitivity of 83.66%, and specificity of 81.66% using 10-fold cross-validation. The CAET-SWin significantly improved reliability of the invasiveness prediction task compared to its radiomics-based counterpart while increase the accuracy by 1.65% and sensitivity by 3.66%. Figure 3 presents the pipeline proposed by Khademi et al.

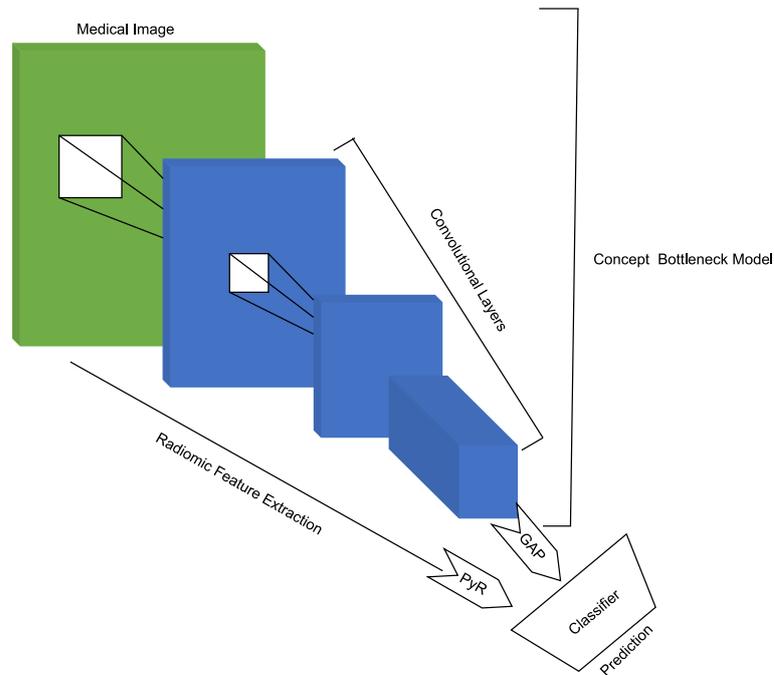


Figure 2. Scheme of image processing, feature extraction, and t-distributed stochastic neighborhood embedding (t-SNE) visualization in [37].

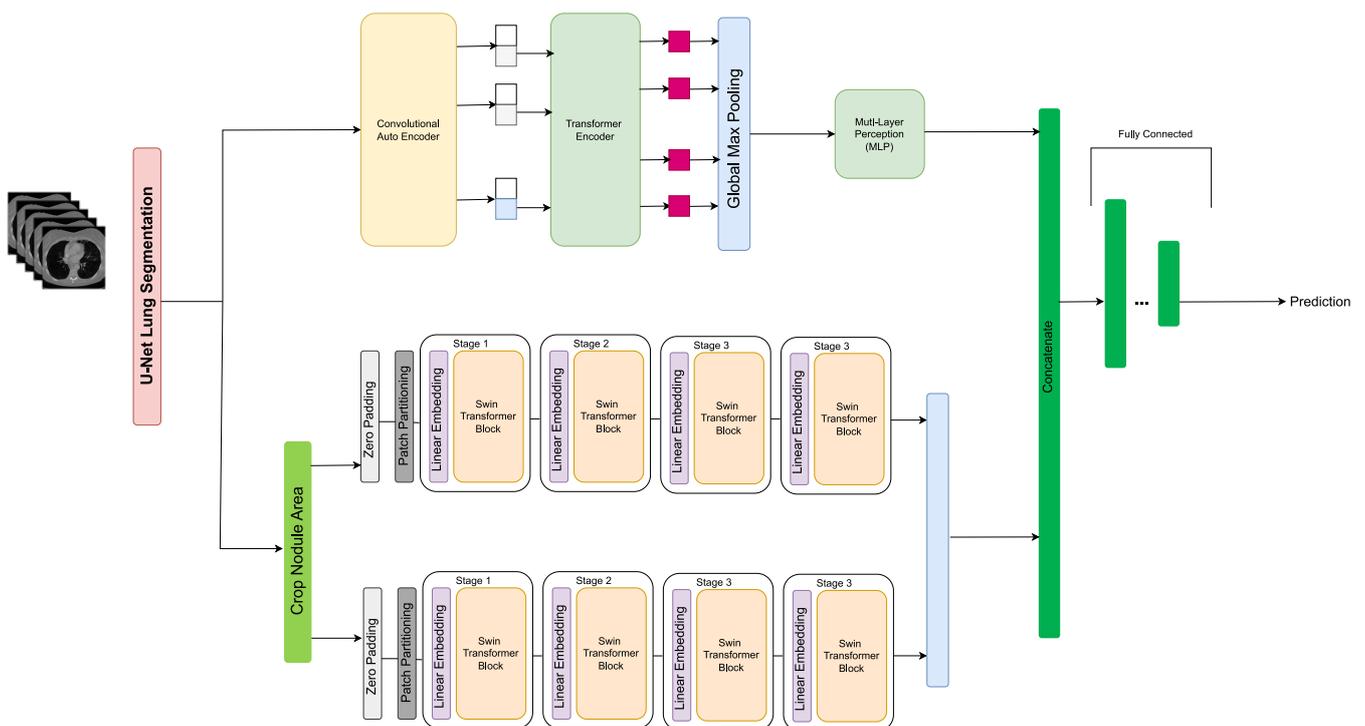


Figure 3. Pipeline of the CAET-SWin Transformer in [40].

Mukherjee et al. [41] developed LungNet, a shallow CNN to predict the outcomes of patients with NSCLC. LungNet was trained and evaluated on four independent cohorts of patients with NSCLC from different medical centers. The results showed that the outcomes predicted by LungNet were significantly associated with overall survival in all four independent cohorts, with concordance indices of 62%, 62%, 62%, and 08% on cohorts 1, 2, 3, and 4, respectively. Additionally, LungNet was able to classify benign versus malignant nodules on the LIDC dataset with an improved performance (AUC = 85%) compared to training from scratch (AUC = 82%) via transfer learning. Overall, the results suggest that LungNet can be used as a non-invasive predictor for prognosis in patients with NSCLC, facilitating the interpretation of CT images for lung cancer stratification and prognostication. The Figure 4 shows the process of making predictions using LungNet. The input to LungNet is a CT image of a lung tumor. The network then extracts features from the image using a series of convolutional layers. These features are then passed through a fully connected layer to make a prediction about the patient's prognosis. The code for LungNet is available at [42].

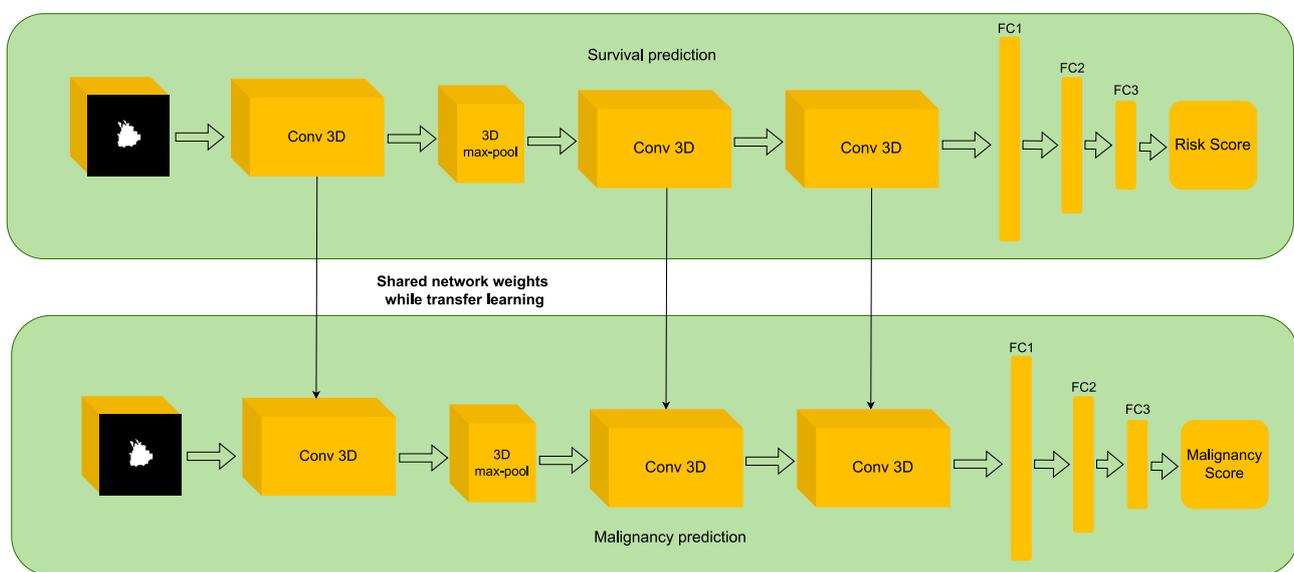


Figure 4. The proposed LungNet architecture in [41].

Da et al. [43] addressed the problem of early detection in lung cancer. To improve early detection and increase survival rates, the study explored the performance of deep transfer learning from non-medical images on lung nodule malignancy classification tasks. The authors preprocessed the data by resizing and normalizing the images and extracting patches around the nodules. Using various convolutional neural networks trained on the ImageNet dataset, the study achieved the highest AUC value of 93.10% using the ResNet50 deep feature extractor and the SVM RBF classifier. In addition to comparing different convolutional neural network architectures and classifiers, the authors also performed ablation experiments to investigate the contribution of different components in their method such as the use of data augmentation and different types of pretraining. They found that data augmentation and using a more complex pretraining dataset ImageNet-21k [44] improved the results.

Han et al. [45] proposed a data augmentation method to boost sensitivity in 3D object detection. The authors suggested using 3D conditional GANs to synthesize realistic and diverse 3D images as additional training data. Specifically, they proposed the use of 3D Multi-Conditional GAN (MCGAN) to generate $32 \times 32 \times 32$ nodules naturally placed on lung CT images. The MCGAN model employed two discriminators for conditioning: The context discriminator learned to classify real versus synthetic nodule/surrounding pairs with noise box-centered surroundings while the nodule discriminator attempted

to classify real versus synthetic nodules with size/attenuation conditions. The results of the study indicated that 3D CNN-based detection could achieve higher sensitivity under any nodule size/attenuation at fixed False Positive rates. The use of MCGAN-generated realistic nodules helped overcome the medical data paucity and even expert physicians failed to distinguish them from the real ones in a Visual Turing Test. The bounding box-based 3D MCGAN model could generate diverse CT-realistic nodules at desired position/size/attenuation blending naturally with surrounding tissues. The synthetic training data boosted sensitivity under any size/attenuation at fixed FP rates in 3D CNN-based nodule detection. This was attributed to the MCGAN's good generalization ability which came from multiple discriminators with mutually complementary loss functions along with informative size/attenuation conditioning. In Figure 5 MCGAN generated realistic and diverse nodules naturally on lung CT scans at desired position/size/attenuation based on bounding boxes and the CNN-based object detector used them as additional training data.

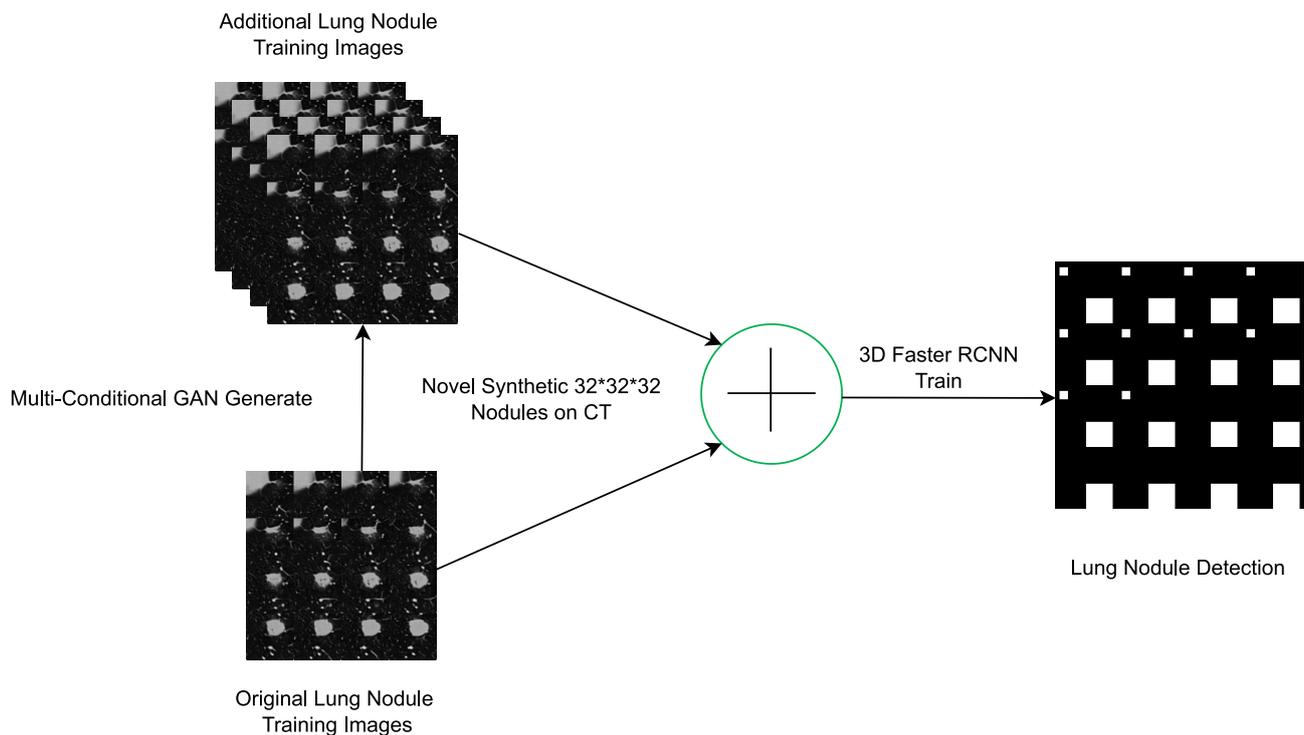


Figure 5. 3D MCGAN-based DA for better object detection in [45].

Katase et al. [46] worked on the development of a CAD system that automatically detects lung nodules in CT images. The main challenge faced by radiologists is identifying small nodule shadows from 3D volume images which can often result in missed nodules. To address this issue, the researchers used deep learning technology to design an automated lung nodule detection system that is robust to imaging conditions. To evaluate the detection performance of the system, the researchers used several public datasets including LIDC-IDRI and SPIE-AAPM as well as a private database of 953 scans and 1177 chest CT scans from Kyorin University Hospital. The system achieved a sensitivity of 98.00/96.00% at 3.1/7.25 false positives per case on the public datasets and sensitivity did not change within the range of practical doses for a study using a phantom. To investigate the clinical usefulness of the CAD system a reader study was conducted with 10 doctors including inexperienced and expert readers. The study showed that using the CAD system as a second reader significantly improved the detection ability of nodules that could be picked up clinically ($p = 0.026$). The analysis was performed using the Jackknife Free-Response Receiver Operating Characteristic (JAFROC). Figure 6 shows the Feature extraction layers extract characteristics from 3D image data by 3D convolution, and region proposal layers

output multiple candidate regions, region classification layers determine whether each candidate region is a nodule, and make this the final output.

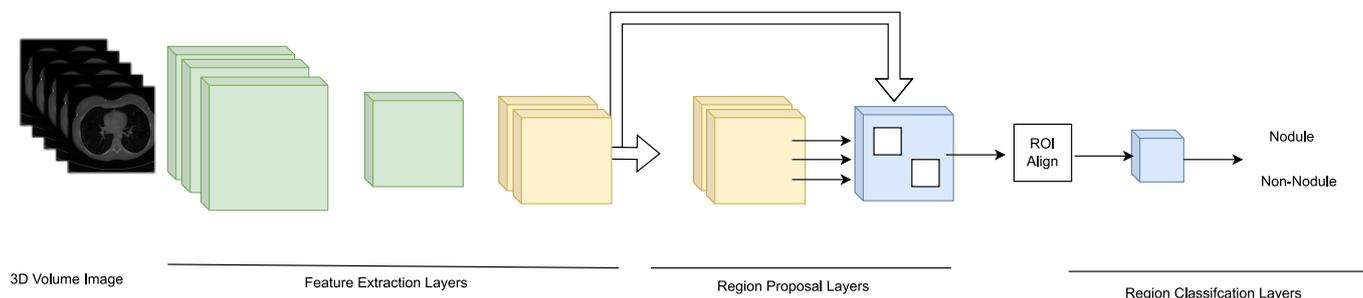


Figure 6. Overview of the Feature extraction layers network in [46].

Tan et al. [47] proposed a new approach to automated detection of juxta-pleural pulmonary nodules in chest CT scans. The article highlighted the challenge of using CNN with limited datasets which can lead to overfitting and presented a novel knowledge-infused deep learning-based system for automated detection of nodules. The proposed CAD methodology infused engineered features, specifically texture features into the deep learning process to overcome the dataset limitation challenge. The system significantly reduced the complications of traditional procedures for pulmonary nodules detection while retaining and even outperforming the state-of-the-art accuracy. The methodology utilized a two-stage fusion method (early fusion and late fusion) which enhanced scalability and adaptation capability by allowing for the easy integration of more useful expert knowledge in the CNN-based model for other medical imaging problems. The results demonstrated that the proposed methodology achieved a sensitivity of 88.00% with 1.9 false positives per scan and a sensitivity of 94.01% with 4.01 false positives per scan. The methodology showed high performance compared to both existing CNN-based approaches and engineered feature-based classifications achieving an AUC of 0.82 with an end-to-end voting-based CNN method for lung nodule detection.

Feng et al. presented in [48] a novel weakly-supervised method for accurately segmenting pulmonary nodules at the voxel-level using only image-level labels. The objective was to extend a CNN model originally trained for image classification to learn discriminative regions at different resolution scales and identify the true nodule location using a candidate-screening framework. The proposed method employed transfer learning from a CNN trained on natural images and adapted the VGG16Net architecture to incorporate GAP operations. The authors demonstrated that their weakly-supervised nodule segmentation framework achieved competitive performance compared to a fully-supervised CNN-based segmentation method, with accuracy values of 88.40% for 1-GAP CNN, 86.60% for 2-GAP model and 84.40% for 3-GAP model on the test set. Furthermore, the proposed method exhibited smaller standard deviations, indicating fewer large mistakes. The proposed method was based on the Nodules Activation Maps (NAM) framework. Figure 7 illustrate the process of the method: In the Training part (A) a CNN model is trained to classify CT slices and generate NAMs; In Segmentation (B) for test slices classified as “nodule slice”, nodule candidates are screened using a spatial scope defined by the NAM for coarse segmentation. Residual NAMs (R-NAMs) are generated from images with masked nodule candidates for fine segmentation.

Aresta et al. [49] introduced iW-Net, a deep learning model that allowed for automatic and interactive segmentation of lung nodules in computed tomography images. iW-Net was composed of two blocks: the first one provided automatic segmentation, and the second one allowed the user to correct it by analyzing two points introduced in the nodule’s boundary. The results of the study showed that iW-Net achieved SOAP with an intersection over union score of 55.00%, compared to the inter-observer agreement of 59.00%. The model also allowed for the correction of small nodules which is essential for proper patient referral decisions and improved the segmentation of challenging non-solid nodules, thus

increasing the early diagnosis of lung cancer. iW-Net improved the segmentation of more than 75.00% of the studied nodules especially those with radii between 1–4 mm, which are crucial for referral. In Figure 8 the network used a $3 \times 3 \times 3 \times N$ convolution followed by batch normalization and rectified linear unit activation where N is the number of feature maps indicated on top of each layer. It also used a $3 \times 3 \times 3 \times N$ convolution with a $2 \times 2 \times 2$ stride followed by batch normalization and rectified linear unit activation. The network then employed a $2 \times 2 \times 2$ nearest neighbor up-sample and a $3 \times 3 \times 3 \times N$ convolution with sigmoid activation. The source code for iW-Net is available at [50].

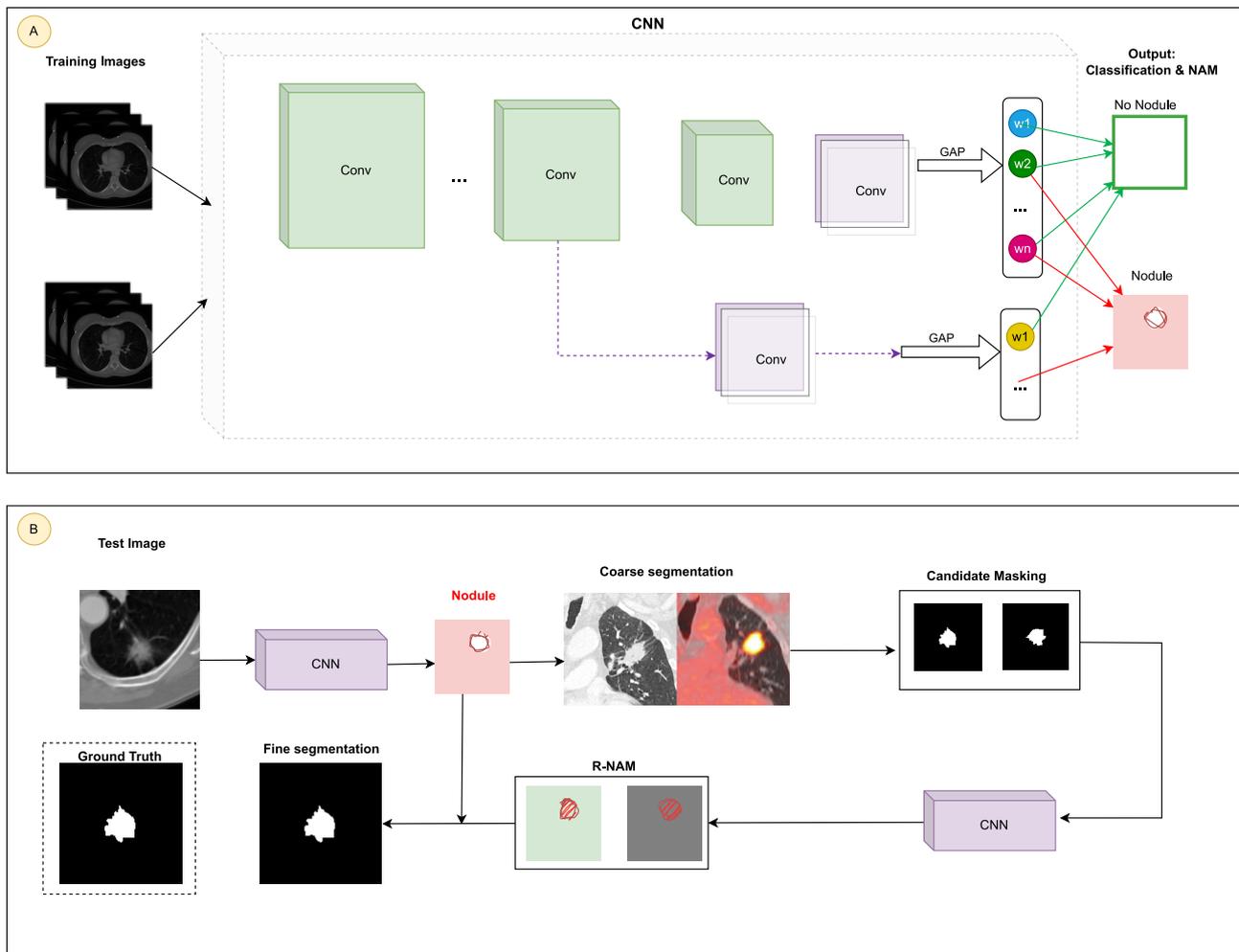


Figure 7. Architecture of the automated detection of juxta-pleural pulmonary nodules in [48]: (A) Training: a CNN is trained to identify CT images and create nodule activation mappings (NAMs); (B) Segmentation: test images identified as containing nodules are subjected to potential nodule filtering, based on a spatial delineation established by the NAM for initial segmentation. Residual NAMs (R-NAMs) are then generated using images in which potential nodules are masked, enabling more precise segmentation.

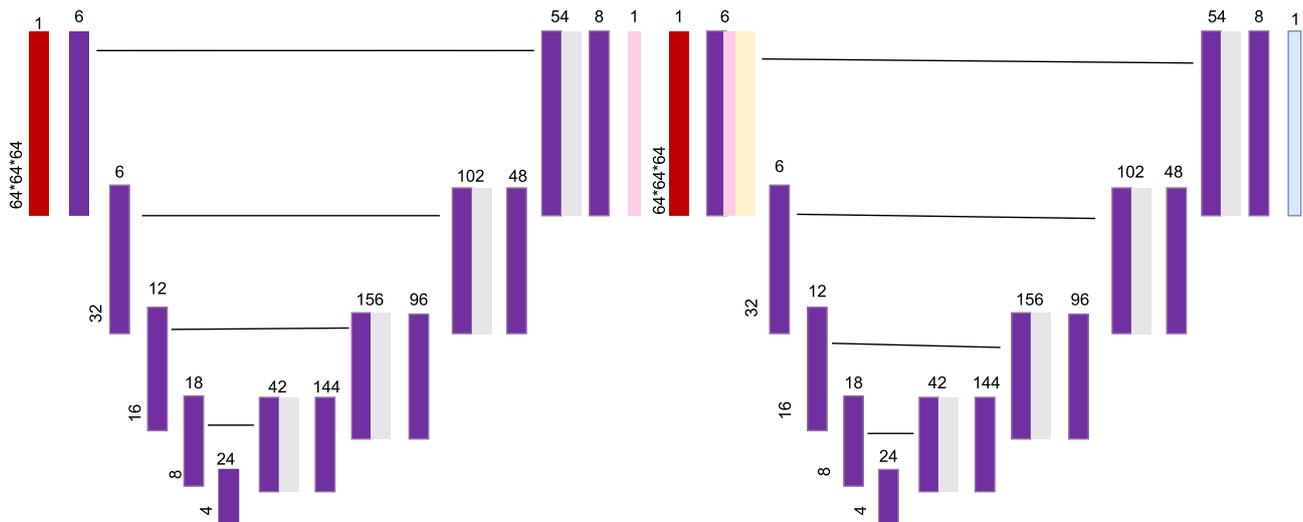


Figure 8. iW-Net: a network for guided segmentation of lung nodules as proposed in [49].

Joana Rocha et al. [51] proposes three distinct methodologies for pulmonary nodule segmentation in CT scans. The first approach is a conventional one that implements the Sliding Band Filter (SBF) to estimate the filter’s support points and match the border coordinates. The other two approaches are Deep Learning based and use the U-Net and a novel network called SegU-Net to achieve the same goal. The study aims to identify the most promising tool to improve nodule characterization. The authors used a database of 2653 nodules from the LIDC database and compared the performance of the three approaches. The results showed Dice scores of 66.30%, 83.00%, and 82.30% for the SBF, U-Net, and SegU-Net, respectively. The U-Net based models yielded more identical results to the ground truth reference annotated by specialists making it a more reliable approach for the proposed exercise. The novel SegU-Net network revealed similar scores to the U-Net while at the same time reducing computational cost and improving memory efficiency. The Figure 9 illustrate SegU-Net. SegU-Net adds a few modifications to the U-Net architecture. Firstly, it uses reversible convolutions in the ascending part. This enables SegU-net to preserve image detail during segmentation. Secondly, it uses a fusion layer to combine information from both parts of the network. This enables SegU-net to achieve more accurate segmentations.

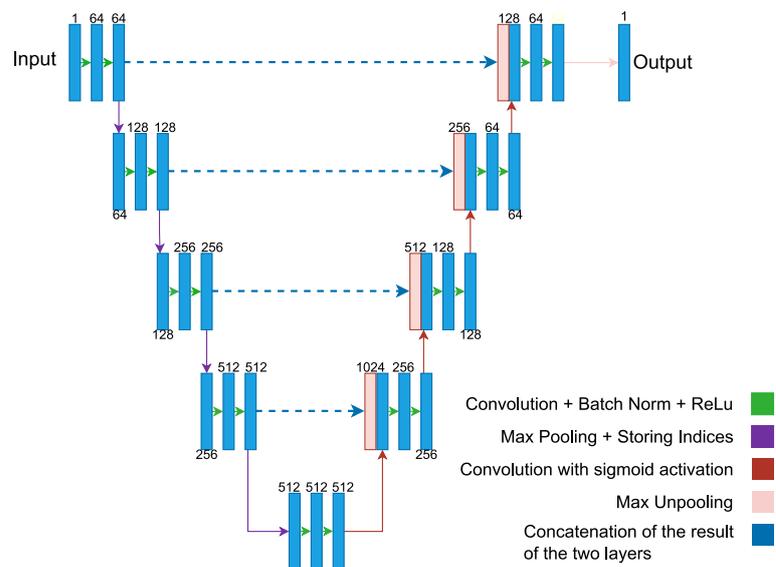


Figure 9. SegU-Net’s model in [51].

Wang et al. [52] present an approach called Multi-view Convolutional Neural Networks (MVCNN) that captures a diverse set of nodule-sensitive features from axial, coronal and sagittal views in CT images simultaneously. The objective is to segment various types of nodules, including juxta-pleural, cavitory, and nonsolid nodules. The methodology consists of three CNN branches each with seven stacked layers that take multi-scale nodule patches as input. These branches extract features from three orthogonal image views in CT which are then integrated with a fully connected layer to predict whether the patch center voxel belongs to the nodule. The approach does not involve any nodule shape hypothesis or user-interactive parameter settings. The study uses 893 nodules from the public LIDC-IDRI dataset where ground-truth annotations and CT imaging data were provided. The results show that MVCNN achieved an average dice similarity coefficient (DSC) of 77.67% and an average surface distance (ASD) of 24.00%, outperforming conventional image segmentation approaches.

Tang et al. [53] proposed a novel approach to solve nodule detection, false positive reduction, and nodule segmentation jointly in a multi-task fashion. The authors presented a new end-to-end 3D deep convolutional neural net (DCNN) called NoduleNet. The goal was to improve the accuracy of nodule detection and segmentation on the LIDC dataset. To avoid friction between different tasks and encourage feature diversification, the authors incorporated two major design tricks in their methodology. Firstly, decoupled feature maps were used for nodule detection and false positive reduction. Secondly, a segmentation refinement subnet was used to increase the precision of nodule segmentation. The authors used the LIDC dataset as their base dataset for training and testing their model. They showed that their model improves the nodule detection accuracy by 10.27%, compared to the baseline model trained only for nodule detection. The cross-validation results on the LIDC dataset demonstrate that the NoduleNet achieves a final CPM score of 87.27% on nodule detection and a DSC score of 83.10% on nodule segmentation, which represents the current state-of-the-art performance on this dataset. The code of NoduleNet is available on [54].

The early identification and classification of pulmonary nodules are crucial for improving lung cancer survival rates. This is considered a key requirement in computer-assisted diagnosis. To address this challenge, ref. [55] proposed a method for predicting the malignant phenotype of pulmonary nodules. The method is based on weighted voting rules. Features of the pulmonary nodules were extracted using Denoising Auto Encoder, ResNet-18, and modified texture and shape features. These features assess the malignant phenotype of the nodules. The results showed a final classification accuracy of $93.10 \pm 2.4\%$, highlighting the method's feasibility and effectiveness. This method combines the robust feature extraction capabilities of deep learning with the use of traditional features in image representation. The study successfully identified multi-class nodules which is the first step in lung cancer diagnosis. The study also explored the importance of various features in classifying the malignant phenotype of pulmonary nodules. It found that shape features were most crucial followed by texture features and deep learning features.

5.1.2. Deep Learning Techniques for Lung Cancer Using LUNA16 Dataset

Xie et al. [56] introduce a new approach to the complex task of automated pulmonary nodule detection in CT images. The aim of this work is to assist in the CT reading process by quickly locating lung nodules. A two-stage methodology is used, consisting of nodule candidate detection and false positive reduction. To accomplish this, the authors present a detection framework based on Faster Region-based CNN (Faster R-CNN). The Faster R-CNN structure is modified with two region proposal networks and a deconvolutional layer for nodule candidate detection. Three models are trained for different kinds of slices to integrate 3D lung information, and then the results are fused. For false positive reduction, a boosting 2D CNN architecture is designed. Three models are sequentially trained to handle increasingly difficult mimics. Misclassified samples are retrained to improve sensitivity in nodule detection. The outcomes of these networks are fused to

determine the final classification. The Luna16 database serves as the evaluation benchmark, achieving a sensitivity of 86.42% for nodule candidate detection. For false positive reduction, sensitivities of 73.40% and 74.40% are reached at 1/8 and 1/4 FPs/scan, respectively.

Sun et al. [57], address the problem of low accuracy in traditional lung cancer detection methods particularly in realistic diagnostic settings. To improve accuracy, the authors propose the use of the Swin Transformer model for lung cancer classification and segmentation. They introduce a novel visual converter that produces hierarchical feature representations with linear computational complexity related to the input image size. The LUNA16 dataset and the MSD dataset are used for segmentation to compare the performance of the Swin Transformer with other models. These include Vision Transformer (ViT), ResNet-101 and data-efficient image transformers (DeiT)-S [58]. The findings reveal that the pre-trained Swin-B model achieves a top-1 accuracy of 82.26% in classification tasks, outperforming ViT by 2.529%. For segmentation tasks, the Swin-S model shows an improvement over other methods with a mean Intersection over Union (mIoU) of 47.93%. These results indicate that pre-training enhances the Swin Transformer model's accuracy.

Agnes et al. [59], address the challenge of manually examining small nodules in computed tomography scans, a process that becomes time-consuming due to human vision limitations. To address this, they introduce a deep-learning-based CAD framework for quick and accurate lung cancer diagnosis. The study employs a dilated SegNet model to segment the lung from chest CT images and develops a CNN model with batch normalization to identify true nodules. The segmentation model's performance is evaluated using the Dice coefficient, while sensitivity is used for the nodule classifier. The discriminative power of features learned by the CNN classifier is further validated through principal component analysis. Experimental outcomes show that the dilated SegNet model achieves an average Dice coefficient of $89.00 \pm 23.00\%$ and the custom CNN model attains a sensitivity of 94.80% for classifying nodules. These models excel in lung segmentation and 2D nodule patch classification within the CAD system for CT-based lung cancer diagnosis. Visual results substantiate that the CNN model effectively classifies small and complex nodules with high probability values. Figure 10 highlights the dilated SegNet for lung segmentation.

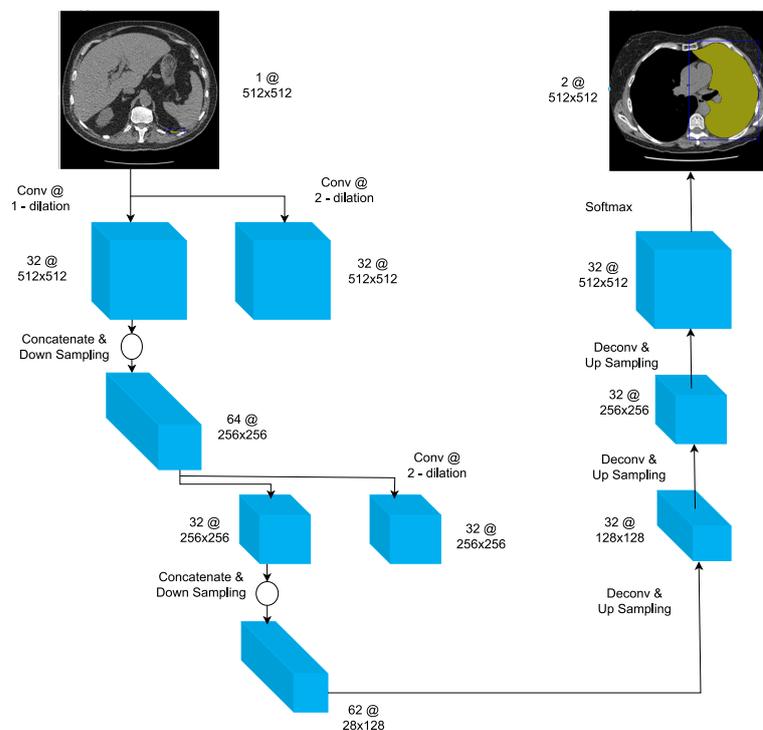


Figure 10. The proposed dilated SegNet for lung segmentation proposed in [59].

Yuan et al. [60] introduce a novel method for early diagnosis, and timely treatment

of lung cancer through the detection and identification of malignant nodules in chest CT scans. The proposed multi-modal fusion multi-branch classification network merges structured radiological data with unstructured CT patch data to differentiate benign from malignant nodules. This network features a multi-branch fusion-based effective attention mechanism for 3D CT patch unstructured data. It also employs a 3D ECA-ResNet, inspired by ECA-Net [61], to dynamically adjust the features. When tested on the LUNA16 and LIDC-IDRI databases, the network achieves the highest accuracy of 94.89%, sensitivity of 94.91% and F1-score of 94.65% along with the lowest false positive rate of 5.55%.

Hassan Mkindu et al. [62] propose a computer-aided diagnosis (CAD) scheme for lung nodule prediction based on a 3D multi-scale vision transformer (3D-MSViT). The goal of this scheme is to improve the efficiency of lung nodule prediction from 3D CT images by enhancing multi-scale feature extraction. The 3D-MSViT architecture uses a local-global transformer block structure where the local transformer stage processes each scale patch separately and then merges multi-scale features at the global transformer level. Unlike traditional methods, the transformer blocks rely solely on the attention mechanism without including CNNs to reduce network parameters. The study uses the Luna16 database to evaluate the proposed scheme. The results demonstrate that the 3D-MSViT algorithm achieved the highest sensitivity of 97.81% and competition performance metrics of 91.10%. However, the proposed scheme is limited to single image modality (CT images) and does not include a stage for false-positive reduction. 3D-ViTNet is an architecture that relies on a single-scale vision transformer encoder without using CNNs. Experimental results show that the integration of 3D ResNet with the attention module improves the detection sensitivity in all tenfold cross-validations compared to plain 3D ResNet. The introduction of 3DViTNet slightly reduces the sensitivities in each experiment fold. The optimal sensitivity is achieved using the multi-scale architecture 3D-MSViT.

5.1.3. Deep Learning Techniques for Lung Cancer Using NLST Dataset

Ardila et al. [63] develop an end-to-end deep learning algorithm for lung cancer screening using low-dose chest CT scans. The goal is to predict a patient's risk of developing lung cancer using their current and prior CT volumes, while demonstrating the potential for deep learning models to increase accuracy, consistency and adoption of lung cancer screening worldwide. The methodology consists of four components, all trained using the TensorFlow platform: Lung segmentation: This component uses Mask-RCNN, which is a 2D object detection algorithm, to segment the lungs in the CT images and compute the center of the bounding box for further processing. Cancer ROI detection: This component uses RetinaNet, which is a 3D object detection algorithm, to detect cancer ROIs in the CT images; Full-volume model: This component uses 3D inflated Inception V1 to predict whether the patient has cancer within 1 year; Cancer risk prediction model: This component uses 3D Inception to extract features from the output of the previous two models and predict the patient's individual malignancy score. The algorithm achieves a higher performance with an area under the curve of 94.40% on 6716 scans. The model outperforms all six radiologists with absolute reductions of 11% in false positives and 5% in false negatives when prior CT imaging was not available.

Than et al. [64], address the challenge of distinguishing between lung cancer and lung tuberculosis (LTB) without invasive procedures which can have significant risks. The study proposes using transfer learning on early convolutional layers to mitigate the challenges posed by limited training datasets. The methodology used a customized 15-layer VGG16-based 2D DNN architecture trained and tested on sets of CT images extracted from the NLST and the NIAID TB Portals [65]. The performance of the DNN was evaluated under locked and step-wise unlocked pretrained weight conditions. The results indicate that the DNN achieved an accuracy of 90.40% with an F score of 90.10%, supporting its potential as a noninvasive screening tool capable of reliably detecting, and distinguishing between lung cancer and LTB.

Li et al. [66], address the problem of interpreting temporal distance between sparse,

irregularly sampled spatial features in longitudinal medical images. They propose two interpretations of a time-distance ViT using vector embeddings of continuous time and a temporal emphasis model to scale self-attention weights. The proposed methods were evaluated on the NLST dataset and the experiments showed a fundamental improvement in classifying irregularly sampled longitudinal images compared to standard ViTs. The two interpretations of the time-distance ViT are as follows: Time embedding strategy: This strategy uses a vector embedding of continuous time to represent the temporal distance between two spatial features. The vector embedding is learned during training; Temporal emphasis model: This strategy learns a separate temporal emphasis model in each attention head. The temporal emphasis model assigns a weight to each spatial feature, depending on its temporal distance to the other spatial features. The experiments on the NLST dataset showed that the time-distance ViTs with both the time embedding strategy and the temporal emphasis model achieved the best performance. They achieved AUC scores of 78.50% and 78.60%, respectively which significantly outperformed the standard ViTs (AUC score of 73.40%) and the cross-sectional approach (AUC score of 77.90%). The results of this study suggest that the time-distance ViTs have the potential to improve the classification of longitudinal medical images. The study also provides a new way to interpret temporal distance in longitudinal medical images. The Figure 11 explains the process of the proposed method. A region proposal network (RPN) is used to propose five local regions in each image volume. Then, the features of these regions are extracted and embedded as five feature tokens. The time distance between the repeated images is then integrated into the model using the time embedding strategy or the temporal emphasis model. The code for the proposed method is available on [67].

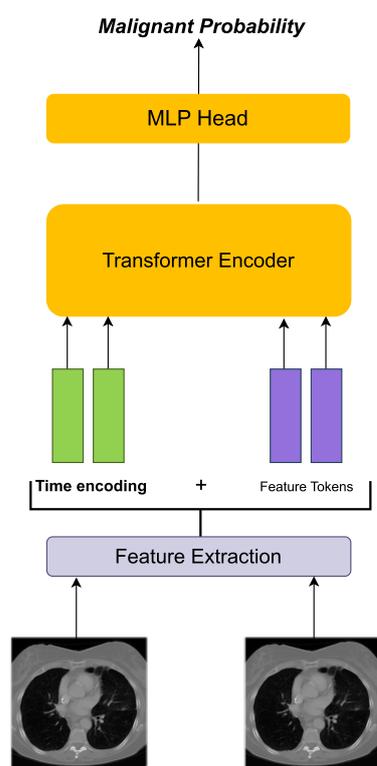


Figure 11. Feature extraction of 2D data was done by taking uniform patches and linearly projecting them in to a common embedding space proposed in [66].

Lu et al. [68] describe the development and validation of a CNN called CXR-LC that predicts long-term incident lung cancer using data commonly available in the electronic medical record (EMR) 12 years before. The CXR-LC model was developed in the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial and validated in additional PLCO and NLST smokers. The results showed that the CXR-LC model had

better discrimination for incident lung cancer than CMS eligibility, with an AUC of 75.50% vs. 63.40%, respectively ($p < 0.001$). The CXR-LC model's performance was similar to that of PLCOM2012 a state-of-the-art risk score with 11 inputs, in both the PLCO dataset (CXR-LC AUC of 75.50% vs. PLCOM2012 AUC of 75.10%) and the NLST dataset (65.90% vs. 65.00%). When compared in equal-sized screening populations, CXR-LC was more sensitive than CMS eligibility in the PLCO dataset (74.90% vs. 63.80%; $p = 0.012$) and missed 30.70% fewer incident lung cancers. On decision curve analysis, CXR-LC had higher net benefit than CMS eligibility and similar benefit to PLCOM2012. The CXR-LC model identified smokers at high risk for incident lung cancer, beyond CMS eligibility and using information commonly available in the EMR.

5.1.4. Deep Learning Techniques for Lung Cancer Using TCGA Dataset

Khan et al. [69] propose an end-to-end deep learning approach called Gene Transformer to address the complexity of high-dimensional gene expression data in the classification of lung cancer subtypes. The study investigates the use of transformer-based architectures which leverage the self-attention mechanism to encode gene expressions and learn representations that are computationally complex and parametrically expensive. The Gene Transformer architecture is inspired by the Transformer encoder architecture and uses a multi-head self-attention mechanism with 1D convolution layers as a hybrid architecture to assess high-dimensional gene expression datasets. The framework prioritizes features during testing and outperforms existing SOTA methods in both binary and multiclass problems. The authors demonstrate the potential of the multi-head self-attention layer to perform 1D convolutions and suggest that it is less expensive than ordinary 2D convolutional layers.

The accurate identification of lung cancer subtypes in medical images is crucial for their proper diagnosis and treatment. Despite the progress made by existing methods, challenges remain due to limited annotated datasets, large intra-class differences, and high inter-class similarities. To address these challenges, Cai et al. [70] proposed a dual-branch deep learning model called the Frequency Domain Transformer Model (FDTrans). FDTrans combines image domain and genetic information to determine lung cancer subtypes in patients. To capture critical detail information a pre-processing step was added to transfer histopathological images to the frequency domain using a block-based discrete cosine transform. The Coordinate-Spatial Attention Module (CSAM) was designed to reassign weights to the location information and channel information of different frequency vectors. A Cross-Domain Transformer Block (CDTB) was then designed to capture long-term dependencies and global contextual connections between different component features. Feature extraction was performed on genomic data to obtain specific features and the image and gene branches were fused. Classification results were output through a fully connected layer. In 10-fold cross-validation, the method achieved an AUC of 93.16% and overall accuracy of 92.33% which is better than current lung cancer subtypes classification detection methods.

In [71], Primakov et al. address the importance of detecting and segmenting abnormalities on medical images for patient management and quantitative image research. The authors present a fully automated pipeline for the detection and volumetric segmentation of NSCLC using 1328 thoracic CT scans. They report that their proposed method is faster and more reproducible compared to expert radiologists and radiation oncologists. The authors also evaluate the prognostic power of the automatic contours by applying RECIST criteria [72] and measuring tumor volumes. The results show that segmentations by their method stratify patients into low and high survival groups with higher significance compared to those methods based on manual contours. Additionally, the authors demonstrate that on average, radiologists and radiation oncologists preferred automatic segmentations in 56% of cases. The code is available at [73].

5.1.5. Deep Learning Techniques for Lung Cancer Using JSRT Dataset

Ausawalaithong et al. [74], aim to develop an automated system for predicting lung cancer from chest X-ray images using deep learning. They explore the use of a 121-layer convolutional neural network, DenseNet121 and a transfer learning scheme to classify lung cancer. The model is trained on a lung nodule dataset before being fine-tuned on the lung cancer dataset to address the issue of a small dataset. The JSRT and ChestX-ray8 datasets are used to evaluate the proposed model. The results indicate a mean accuracy of $74.43 \pm 6.01\%$, mean specificity of $74.96 \pm 9.85\%$ and mean sensitivity of $74.68 \pm 15.33\%$. Additionally, the model provides a heatmap for identifying the location of the lung nodule. These findings are promising for further development of chest X-ray-based lung cancer diagnosis using the deep learning approach.

Gordienko et al. [75] aim to leverage advancements in deep learning and image recognition to automatically detect suspicious lesions and nodules in CXRs of lung cancer patients. The study preprocesses the CXR images using lung segmentation and bone shadow exclusion techniques before applying a deep learning approach. The original JSRT dataset and the BSE-JSRT dataset, which is the same as the JSRT dataset but without clavicle and rib shadows were used for analysis. Both datasets were also used after segmentation, resulting in four datasets in total. The results of the study demonstrate the effectiveness of the preprocessing techniques, particularly bone shadow exclusion in improving accuracy and reducing loss. The dataset without bones had significantly better results than the other preprocessed datasets. However, the study notes that pre-processing for label noise during the training stage is crucial because the training data was not accurately labeled for the test set. The study also identifies potential areas for improvement, such as increasing the size and number of images investigated and using data augmentation techniques for lossy and lossless transformations. These improvements would enable a wider range of CXRs to be analyzed and increase the accuracy and efficiency of the deep learning approach.

5.1.6. Deep Learning Techniques for Lung Cancer Using Kaggle DSB Dataset

Yu et al. [76] address the problem of divergent software dependencies in automated chest CT evaluation methods for lung cancer detection, which makes it difficult to compare and reproduce these methods. The study aims to develop reproducible machine learning modules for lung cancer detection and compare the approaches and performances of the award-winning algorithms developed in the Kaggle Data Science Bowl. The authors obtained the source codes of all award-winning solutions and evaluated the performance of the algorithms using the log-loss function and the Spearman correlation coefficient of the performance in the public and final test sets. The low-dose chest CT datasets in DICOM format from the Kaggle Data Science Bowl website were used. The datasets consisted of a training set with ground truth labels and a public test set without labels. Most solutions implemented distinct image preprocessing, segmentation and classification modules. Variants of U-Net, VGGNet and residual net were commonly used in nodule segmentation and transfer learning was used in most of the classification algorithms.

Tekade et al. [77] propose a solution to the problem of detecting and classifying lung nodules, as well as predicting the malignancy level of these nodules using CT scan images. The authors introduce a 3D multipath VGG-like network, which is evaluated on 3D cubes and combined with U-Net for final predictions. The study uses the LIDC-IDRI, the LUNA16, and the Kaggle DBS 2017 datasets. The proposed approach achieves an accuracy of 95.60% and a logloss of 38.77%, with a dice coefficient of 90%. The results are useful for predicting whether a patient will develop cancer in the next two years. The study concludes that Artificial Neural Networks play an important role in better analyzing the dataset, extracting features and classification. The proposed approach is effective for lung nodule detection and malignancy level prediction using lung CT scan images.

5.1.7. Deep Learning Techniques for Lung Cancer Using Decathlon Dataset

Said et al. [78] proposed a system using deep learning architectures for the early diagnosis of lung cancer in CT scan imaging. The proposed system consists of two parts: segmentation and classification. Segmentation is performed using the UNETR network [79], while classification is performed using a self-supervised network. The segmentation part aims to identify the ROI in the CT scan images. This is done by first projecting the 3D patches from the volumetric image into an embedding space. A positional embedding is then added to these patches. The Transformer encoder captures global and long-range dependencies in the image through attention mechanisms. It extracts high-level representations of the ROI. The classification part is then used to classify the output of the segmentation part, either as benign or malignant. This part is developed on top of the self-supervised network and aims to classify the identified regions as either cancerous or non-cancerous. The proposed system uses 3D-input CT scan data, making it a powerful tool for early diagnosing and combating lung cancer. The system shows promising results in diagnosing lung cancer using 3D-input CT scan data, achieving a segmentation accuracy of 97.83% and a classification accuracy of 98.77%. The Decathlon dataset was used for training and testing experiments.

Guo et al. [80] propose a solution to the anisotropy problem in 3D medical image analysis. This problem occurs when the slice spacing varies significantly between training and clinical datasets, which can degrade the performance of machine learning models. The authors propose a transformer-based model called TSMUNet, which is adaptable to different levels of anisotropy and is computationally efficient. TSMUNet is based on a 2D U-Net backbone consisting of a downsampling stream, upsampling stream and a transformer block (Figure 12). The downsampling stream takes 3D CT scans as input and extracts features at multiple resolutions. The upsampling stream then reconstructs the image from the extracted features. The transformer block is used to encode inter-slice information using a self-attention mechanism. This allows the model to adapt to variable slice spacing and to capture long-range dependencies in the image. The authors evaluated TSMUNet on the MSD database, which includes 3D lung cancer segmentation data. The results showed that TSMUNet outperforms baseline models such as the 3D U-Net and LSTMUNet. TSMUNet achieved a segmentation accuracy of 87.17% which is significantly higher than the 77.44% achieved by the 3D U-Net and the 85.73% achieved by the LSTMUNet.

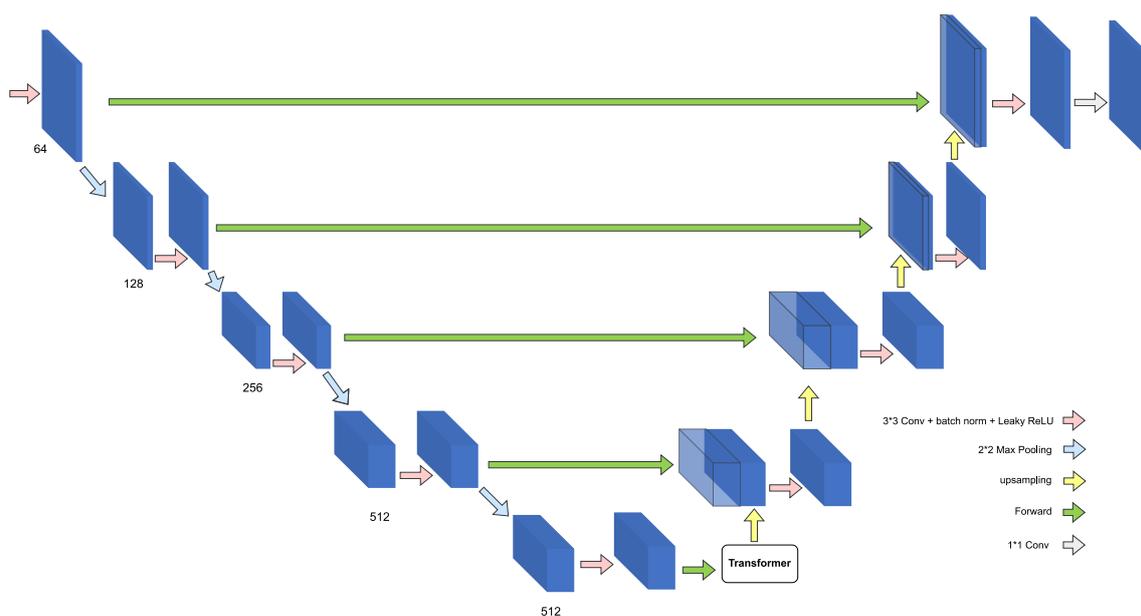


Figure 12. The propose TSMUNet in [80].

5.1.8. Deep Learning Techniques for Lung Cancer Using Tianchi Dataset

Hao Tang et al. [81] propose a novel end-to-end framework for pulmonary nodule detection, that integrates nodule candidate screening and false positive reduction into one model. The objective is to improve nodule detection by jointly training the two stages of nodule candidate generation and false positive reduction. The proposed framework follows a two-stage strategy: (1) Generating nodule candidates using a 3D Nodule Proposal Network. (2) Classifying the nodule candidates to reduce false positives. The nodule candidate screening branch uses a 3D Region Proposal Network (RPN) adapted from Faster R-CNN. The RPN predicts a set of bounding boxes around potential nodule candidates. The predicted bounding boxes are then used to crop features of the nodule candidates using a 3D ROI Pool layer as shown in Figure 13. These features are then fed as input to the nodule false positive reduction branch. The nodule false positive reduction branch uses a CNN to classify the nodule candidates as either true positives or false positives. The CNN is trained to minimize the loss between the predicted labels and the ground truth labels. Convolution blocks are then built using residual blocks and maxpooling to further reduce the spatial resolution. This allows the model to learn more abstract features from the image while reducing the amount of memory required. The authors evaluated the proposed framework on the Tianchi competition dataset. The results showed that the end-to-end system outperforms the two-step approach by 3.88%. The end-to-end system also reduces model complexity by one third and cuts inference time by 3.6 fold.

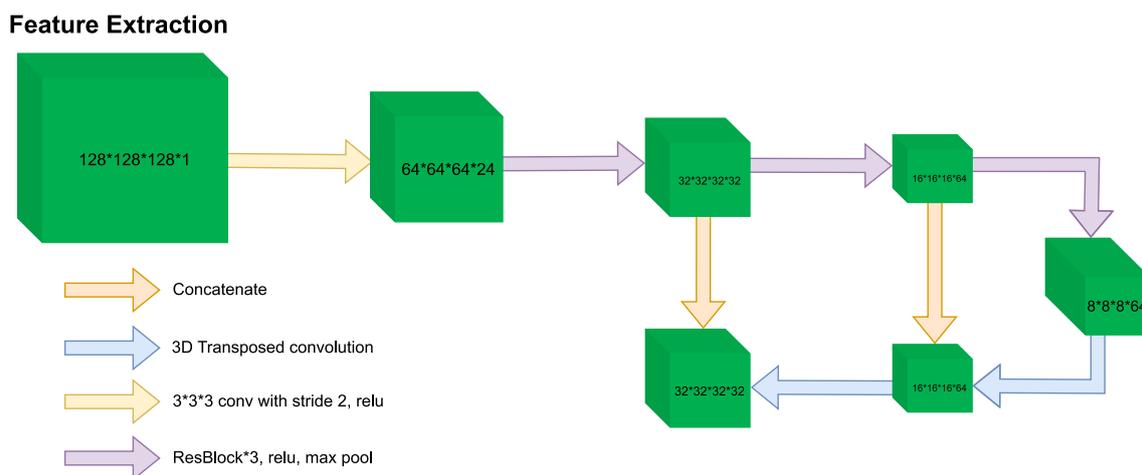


Figure 13. End-to-end pulmonary nodule detection framework in [81]. (*) is equivalent to the times (\times) sign.

Huang et al. [82] propose an improved CNN framework for the more effective detection of pulmonary nodules. The framework consists of three 3D CNNs, namely CNN-1, CNN-2 and CNN-3 which are fused into a new Amalgamated-CNN model to detect pulmonary nodules. To detect nodules, the authors first use an unsharp mask to enhance the nodules in CT images. Then, CT images of 512×512 pixels are segmented into smaller images of 96×96 pixels and the plaques corresponding to positive and negative samples are segmented. CT images segmented into 96×96 pixels are then downsampled to 64×64 and 32×32 sizes respectively. The authors discard nodules less than 5 mm in diameter and use the AdaBoost classifier to fuse the results of CNN-1, CNN-2, and CNN-3. They call this new neural network framework the Amalgamated-Convolutional Neural Network (A-CNN). The authors evaluated the proposed A-CNN model on the LUNA16 dataset where it achieved sensitivity scores of 81.70% and 85.10% when the average false positives number per scan was 0.125 FPs/scan and 0.25 FPs/scan respectively. These scores were 5.40% and 0.50% higher than those of the current optimal algorithm.

Tang et al. [83] address the challenge of detecting nodules in three-dimensional medical imaging data using DCNN. While previous approaches have used 2D or 2.5D components

for analyzing 3D data, the proposed DCNN approach is fully 3D end-to-end and utilizes SOTA object detection techniques. The proposed method consists of two stages: candidate screening and false positive reduction. In the first stage, a U-Net-inspired 3D Faster R-CNN is used to identify nodule candidates while preserving high sensitivity. In the second stage, 3D DCNN classifiers are trained on difficult examples produced during candidate screening to finely discriminate between true nodules and false positives. Models from both stages are then ensembled for final predictions allowing for flexibility in adjusting the trade-off between sensitivity and specificity. The proposed approach was evaluated using data from Alibaba's 2017 TianChi AI Competition for Healthcare. The classifier was trained for 300 epochs using positive examples from the Faster R-CNN detector balanced with hard negative samples. The input candidates for test set predictions were provided by the detector and the checkpoint with the highest CPM on the validation set was used for prediction on the test set. The proposed approach achieved superior performance with a CPM of 81.50%.

5.1.9. Deep Learning Techniques for Lung Cancer Using Peking University Cancer Hospital Dataset

Clinical staging is crucial for treatment decisions and prognosis evaluation for lung cancer. However, inconsistencies between clinical and pathological stages are common due to the free-text nature of CT reports. In [84], Fischer et al. developed an information extraction (IE) system to automatically extract staging-related information from CT reports using three components: NER [85], relation classification (RC) [86] and postprocessing (PP). The NER component was used to identify entities of interest such as tumor size, lymph node status and metastasis. The RC component was used to classify the relationships between these entities. The PP module was used to correct errors and inconsistencies in the extracted information. The IE system was evaluated on a clinical dataset of 392 CT reports. The BERT model outperformed the ID-CNN-CRF model and the Bi-LSTM-CRF model for NER, with macro-F1 scores of 80.97%, 90.06% and 90.22% respectively. The BERT-RSC model outperformed the baseline methods for RC, with macro-F1 and micro-F1 scores of 97.13% and 98.37% respectively. The PP module achieved macro-F1 and micro-F1 scores of 94.57% and 96.74% respectively for all 22 questions related to lung cancer staging. The experimental results demonstrated the system's potential for use in stage verification and prediction to facilitate accurate clinical staging.

Zhang et al. [87] proposed a novel deep learning approach for extracting clinical entities from Chinese CT reports for lung cancer screening and staging. The free-text nature of CT reports poses a significant challenge to effectively using this valuable information for clinical decision-making and academic research. The proposed approach utilizes the BERT-based BiLSTM-Transformer network (BERT-BTN) with pre-training to extract 14 types of clinical entities. The BERT-BTN model first uses BERT [88] to generate contextualized word embeddings. This is followed by a BiLSTM layer to capture local sequential information. Then a Transformer layer is added to model global dependencies between words regardless of distance. The BiLSTM provides positional information while the Transformer draws long-range dependencies. After the BERT-BiLSTM-Transformer encoder a CRF layer is added to further incorporate constraints from the labels. The proposed approach was evaluated on a clinical dataset consisting of 359 CT reports collected from the Department of Thoracic Surgery II of Peking University Cancer Hospital. The results showed that the BERT-BTN model achieved an 85.96% macro-F1 score under exact match scheme which outperforms the benchmark BERT-BTN, BERT-LSTM, BERT-fine-tune, BERT-Transformer, FastText-BTN, FastText-BiLSTM, and FastText-Transformer models. The results indicate that the proposed approach efficiently recognizes various clinical entities for lung cancer screening and staging and holds great potential for further utilization in clinical decision-making and academic research. The proposed BERT-BTN is shown in Figure 14.

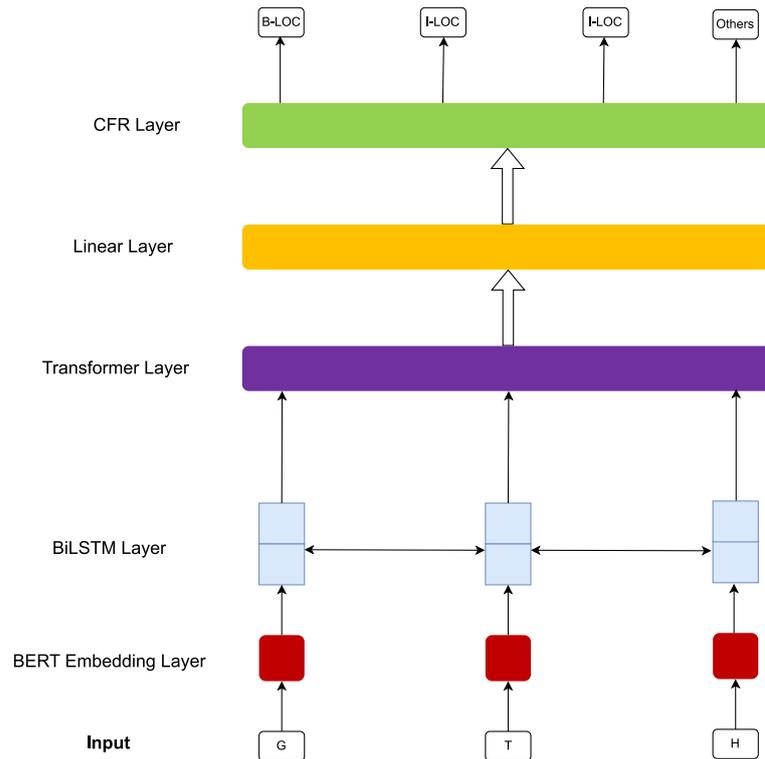


Figure 14. The propose BERT-BTN in [87].

5.1.10. Deep Learning Techniques for Lung Cancer Using MSKCC Dataset

Kipkogei et al. [89] introduced the Clinical Transformer a variation of the transformer architecture that is used for precision medicine to model the relationship between molecular, and clinical measurements and the survival of cancer patients. The Clinical Transformer first uses an embedding strategy to convert the molecular and clinical data into vectors. These vectors are then fed into the transformer which learns long-range dependencies between the features. The transformer also uses an attention mechanism to focus on the most important features for predicting survival. The authors proposed a customized objective function to evaluate the performance of the Clinical Transformer. This objective function takes into account both the accuracy of the predictions and the interpretability of the model. The authors evaluated the Clinical Transformer on a dataset of 1661 patients from the MSKCC. The results showed that the Clinical Transformer outperformed other linear and non-linear methods currently used in practice for survival prediction. The authors also showed that initializing the weights of a domain-specific transformer with the weights of a cross-domain transformer further improved the predictions. The attention mechanism used in the Clinical Transformer successfully captures known biology behind these therapies.

5.1.11. Deep Learning Techniques for Lung Cancer Using SEER Dataset

Doppalapudi et al. [90] aimed to develop deep learning models for lung cancer survival prediction in both classification and regression problems. The study compared the performance of ANN, CNN and RNN models with traditional machine learning models using data from the SEER. The deep learning models outperformed traditional machine learning models achieving a best classification accuracy of 71.18% when patients' survival periods were segmented into classes of " ≤ 6 months", "0.5–2 years" and ">2 years". The RMSE of the regression approach was 13.5% and the R2 value was 0.5. In contrast, the traditional machine learning models saturated at 61.12% classification accuracy and 14.87% RMSE in regression. The deep learning models provide a baseline for early prediction and could be improved with more temporal treatment information collected from treated patients. Additionally, the feature importance was evaluated to investigate the model

interpretability and gain further insight into the survival analysis models and the factors that are important in cancer survival period prediction.

5.1.12. Deep Learning Techniques for Lung Cancer Using TCIA Dataset

Barbouchi et al. [91] present a new approach for the classification and detection of lung cancer using deep learning techniques applied to PET/CT images. Early detection is crucial for increasing the cure rate and this approach aims to fully automate the anatomical localization of lung cancer from PET/CT images and classify the tumor to determine the speed of progression and the best treatments to adopt. The authors used the DETR model based on transformers to detect the tumor and assist physicians in staging patients with lung cancer. The TNM staging system and histologic subtype classification were taken as a standard for classification. The proposed approach achieved an IoU of 0.8 when tested on the Lung-PET-CT-Dx dataset, indicating a high level of accuracy in detecting tumors. It also outperformed SOTA T-staging and histologic classification methods, achieving classification accuracy of 0.97 and 0.94 for T-stage and histologic subtypes respectively.

In [91], Barbouchi et al. present a new approach for the classification and detection of lung cancer using deep learning techniques applied to positron emission tomography/computed tomography (PET/CT) images. Early detection is crucial for increasing the cure rate and this approach aims to fully automate the anatomical localization of lung cancer from PET/CT images and classify the tumor to determine the speed of progression and the best treatments to adopt. The authors used the DETR model which is a transformer-based model to detect the tumor and assist physicians in staging patients with lung cancer. The TNM staging system and histologic subtype classification were used as a standard for classification. The proposed approach achieved an IOU of 80% when tested on the Lung-PET-CT-Dx dataset, indicating a high level of accuracy in detecting tumors. It also outperformed SOTA T-staging and histologic classification methods achieving classification accuracy of 97% and 94% for T-stage and histologic subtypes respectively. An IOU of 80% indicates that the predicted region overlaps with the ground truth region by 80%.

5.2. Deep Learning Technics Using Proprietary Datasets

In addition to publicly available datasets, many studies on DL-based lung cancer diagnosis have utilized private databases which may be proprietary or collected specifically for research purposes. While private datasets may offer advantages such as greater size or more detailed annotations they are often not publicly accessible and may be subject to confidentiality agreements. In this section, we review a selection of studies that have utilized private databases for the development and validation of deep learning algorithms for lung cancer diagnosis. These studies demonstrate the potential benefits of working with large and detailed datasets but also raise questions about reproducibility and generalizability of results. By examining the methods and results of these studies, we aim to identify key insights and challenges in using private datasets for deep learning-based lung cancer diagnosis and consider the implications of this approach for the wider research community. Appendix A Table A3 summarizes the reviewed methods for lung cancer diagnosis using proprietary datasets.

Weikert et al. [92] aimed to develop and test a Retina U-Net algorithm for detecting primary lung tumors and associated metastases of all stages on FDG-PET/CT. The methodology involved evaluating detection performance for all lesion types, assigning detected lesions to categories T, N, or M using an automated anatomical region segmentation and visually analyzing reasons for false positives. The study used a dataset of 364 FDG-PET/CTs of patients with histologically confirmed lung cancer, which was split into a training, validation and internal test dataset. The results showed that the Retina U-Net algorithm had a sensitivity of 86.2% for T lesions and 94.3% accuracy in TNM categorization based on the anatomical region approach. The Figure 15 shows that Retina U-Net architecture resembles a standard U-Net with an encoder-decoder structure. It is a segmentation model that is complemented by additional detection network branches in the lower (coarser) decoder

levels for end-to-end object classification and bounding box regression. This allows the detection network to exploit higher-level object features from the segmentation model. The segmentation model provides high-quality pixel-level training signals that are back-propagated to the detection network. This enables Retina U-Net to leverage segmentation labels for object detection in an end-to-end fashion. The study's performance metrics had wide 95% confidence intervals as the internal test set was only a small portion of the whole internal data set.

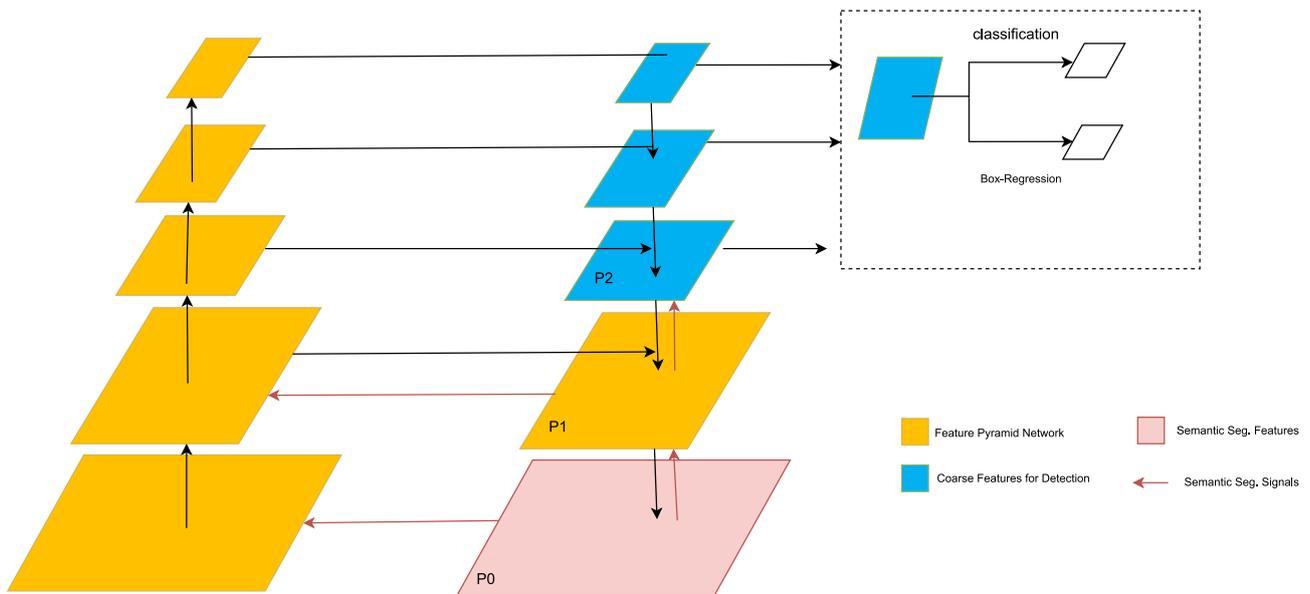


Figure 15. Retina U-Net architecture presented in [92]. The encoder-decoder structure resembles a U-Net.

Nishio et al. [93] presented a CADx method for the classification of lung nodules into benign nodule, primary lung cancer, and metastatic lung cancer. The study evaluated the usefulness of DCNN for CADx in comparison to a conventional method (hand-crafted imaging feature plus machine learning), the effectiveness of transfer learning, and the effect of image size as the DCNN input. To perform the CADx, the authors used a previously-built database of CT images and clinical information of 1236 patients out of 1240. The CADx was evaluated using the VGG-16 convolutional neural network with and without transfer learning. The hyperparameter optimization of the DCNN method was performed by random search. For the conventional method, CADx was performed using rotation-invariant uniform-pattern local binary pattern on three orthogonal planes with a support vector machine. The study found that DCNN was better than the conventional method for CADx and the accuracy of DCNN improved when using transfer learning. Additionally, the authors discovered that larger image sizes as inputs to DCNN improved the accuracy of lung nodule classification. The best averaged validation accuracies of CADx were 55.90%, 68.00% and 62.40% for the conventional method, the DCNN method with transfer learning and the DCNN method without transfer learning respectively. For image size of 56, 112, and 224, the best averaged validation accuracy for the DCNN with transfer learning were 60.70%, 64.70%, and 68.00%, respectively. The study demonstrates that the 2D-DCNN method is more useful for ternary classification of lung nodule than the conventional method for CADx and transfer learning enhances the image recognition for CADx by DCNN when using medium-scale training data.

Lakshmanaprabu et al. [94] presented a novel automated diagnostic classification method for CT images of lungs with the aim of enhancing lung cancer classification accuracy. The methodology comprises two phases: in the first phase, selected features are extracted and reduced using Linear Discriminant Analysis (LDA); in the second phase, an Optimal Deep Neural Network (ODNN) is employed, incorporating the Modified

Gravitational Search Algorithm (MGSA) optimization algorithm for classifying CT lung cancer images. The study was conducted on a dataset of 50 low-dosage and recorded lung cancer CT images. The results demonstrate that the proposed classifier achieves a sensitivity of 96.20%, specificity of 94.20% and accuracy of 94.56%.

Shin et al. [95] present a new method for accurately diagnosing early-stage lung cancer using deep learning-based surface-enhanced Raman spectroscopy (SERS) of exosomes. The objective of this study was to analyze the SERS spectra of human plasma-derived exosomes based on features of cell-derived exosomes using the deep learning algorithm with the hypothesis that the Raman signal of cancer cell exosomes can be detected from the plasma exosomes. The deep learning model was trained with SERS signals of exosomes derived from normal and lung cancer cell lines. Exosomes were first isolated from cell culture supernatant and human plasma samples, and their SERS signals were collected using a gold nanoparticle (GNP)-coated plate. The spectral dataset of exosomes from cell culture supernatant was used to train the deep learning models for binary classification of cell types. The total dataset consisted of 2150 cell-derived exosome data. For binary classification, the normal and cancer cell exosomes were respectively labeled with 0 and 1. The entire data were shuffled randomly before training to make the dataset have both normal cell- and cancer cell-derived exosome data in each batch. The authors found that their deep learning model could classify exosomes with an accuracy of 95%. In 43 patients including stage I and II cancer patients, the deep learning model predicted that plasma exosomes of 90.7% of patients had higher similarity to lung cancer cell exosomes than the average of the healthy controls. Notably, the model predicted lung cancer with an AUC of 91.20% for the whole cohort and stage I patients with an AUC of 91.0%. The Figure 16 shows that the deep learning model utilizes a ResNet architecture. The input data passes through an initial convolutional layer with 64 filters that expands the channel depth while reducing the data length via a pooling layer with a 3×3 kernel. After several basic ResNet blocks, the model connects two fully connected layers with ReLU activation and 40% dropout. The basic blocks output feature maps of varying lengths and channel depths. During training, the loss and accuracy were monitored at each epoch with dropout disabled. The training loss decreased and accuracy increased over iterations, indicating convergence. After full training, the model generated output scores for a representative sample of 200 data points.

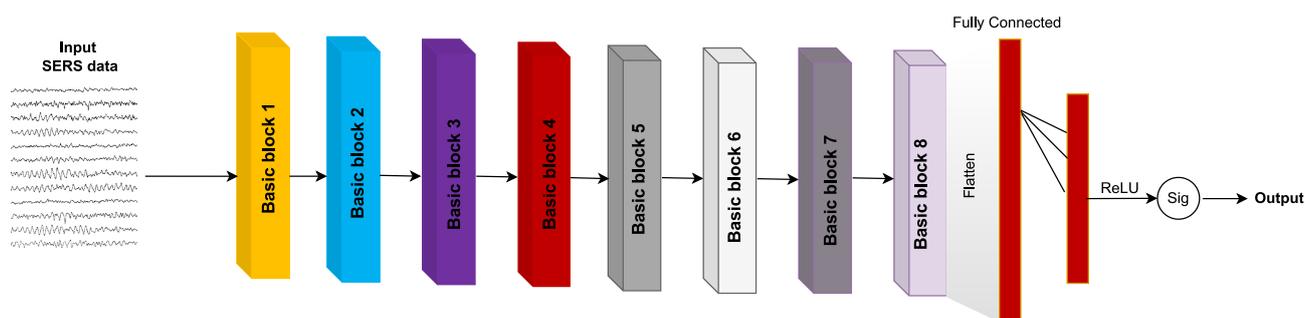


Figure 16. Deep learning-based cell exosome classification in [95].

Shao et al. [96] approached the challenges of lung cancer screening in resource-limited settings through the use of mobile CT scanning coupled with deep learning techniques. The researchers enrolled over 12,000 participants who underwent scans using a mobile CT vehicle. It was found that 9511 (76.95%) of the participants had pulmonary nodules detected. The authors developed a deep learning system for nodule detection and malignancy evaluation which achieved SOTA performance as measured by recall, FROC, accuracy, and AUC. After 1-year of follow-up, 86 patients were diagnosed with lung cancer of which 80 (93.03%) cases were adenocarcinoma and 73 (84.88%) were early stage. For nodule detection, the deep learning system attained a recall of 95.07% and FROC of 64.70%. For lung cancer risk stratification it achieved an accuracy of 86.96% and macro-AUC of 85.16%.

Traditional Chinese medicine (TCM) has been proven effective in managing advanced

lung cancer [97]. Accurate syndrome differentiation is crucial for successful TCM treatment. Intelligent models for TCM syndrome differentiation have been developed by leveraging documented TCM treatment cases and advancements in AI technology. In [98], Liu et al. aimed to establish an end-to-end TCM diagnostic model for lung cancer syndrome differentiation using unstructured medical records. The approach taken was to treat lung cancer TCM syndrome differentiation as a multilabel text classification problem. First, entity representation was conducted using BERT and conditional random fields (CRF) models. Five deep learning-based text classification models were then employed to construct a multilabel classifier for medical records. Two data augmentation strategies were employed to mitigate overfitting concerns. Lastly, a fusion model approach was employed to enhance model performance. The dataset utilized in the experiment comprised 1206 clinical records of patients diagnosed with non-small cell lung cancer. The resulting models were anticipated to exhibit greater efficiency compared to approaches relying on structured TCM datasets. The RCNN model with data augmentation achieved an F1 score of 86.50%, demonstrating a 2.41% improvement over the unaugmented model. The text-hierarchical attention network (Text-HAN) model achieved the highest F1 scores, 86.76% and 87.51% respectively. The mean average precision for the word-encoding-based RCNN was 10% higher than that of the character-encoding-based representation. A fusion model comprising the text-convolutional neural network text-recurrent neural network and Text-HAN models achieved an F1 score of 0.8884 showcasing superior performance among the models.

Wang et al. [99] aimed to develop an automated method for lung cancer segmentation using deep learning and dual-modality imaging and to evaluate the clinical performance of the method. The methodology involved constructing a 3D neural network with dual inputs from diagnostic PET and simulation CT based on U-Net. The performance of the 3D dual-modality network was compared against that of a CT-only network and the results were evaluated using a dataset of 290 pairs of PET and CT from lung cancer patients with manual physician contours as the ground truth. The proposed 3D network with a novel GTV volume-based stratification strategy generated clinically useful lung cancer contours that were quantitatively similar to the ground truth and highly acceptable in physician review. The GTV volume-based stratification strategy divides the dataset into two subsets: a large GTV subset and a small GTV subset. The model for the large GTV subset is trained with GTVs of all sizes and the model for the small GTV subset is trained with small GTVs only. This strategy was found to improve the performance of the network for both large and small GTVs. The results showed that the dual modality inputs delivered better results than the CT-only inputs with a mean DSC, HD, and BLD of $79 \pm 1\%$, 5.8 ± 3.2 mm, and 2.8 ± 1.5 mm respectively.

Park et al. [100] propose a two-stage U-Net architecture for automatic lung cancer segmentation in [18F]FDG PET/CT scans. The proposed method involves a global U-Net in Stage 1, which receives a 3D PET/CT volume as input and extracts the preliminary tumor area generating a 3D binary volume as output. In Stage 2, a regional U-Net receives eight consecutive PET/CT slices around the slice selected by the global U-net in Stage 1 and generates a 2D binary image as output. The proposed method was evaluated using a dataset of 887 patients with lung cancer which was randomly partitioned into training, validation and test sets. The results showed that the proposed two-stage U-Net architecture outperformed the conventional one-stage 3D U-Net in primary lung cancer segmentation. The two-stage U-Net model successfully predicted the detailed margin of the tumors as confirmed by quantitative analysis using the Dice similarity coefficient. The proposed method is expected to reduce the time and effort required for accurate lung cancer segmentation in [18F]FDG PET/CT. This study highlights the potential of deep learning approaches such as U-Net for automatic lung cancer segmentation in medical imaging.

Li et al. [101] present the results of the Automatic Cancer Detection and Classification in Whole-slide Lung Histopathology (ACDC@LungHP) Grand Challenge [102] which was designed to evaluate different CAD methods for the automatic diagnosis of lung cancer.

The focus of the challenge was on segmentation (pixel-wise detection) of cancer tissue in whole slide imaging (WSI) using a dataset of 150 training images and 50 test images from 200 patients. The article reviews the challenge and summarizes the top 10 submitted methods for lung cancer segmentation. All methods were based on deep learning and categorized into two groups: multi-model methods and single-model methods. In general, multi-model methods were significantly better than single-model methods with mean Dice coefficients of 79.66% and 75.44% respectively. The DC of the best method was near the inter-observer agreement of 83.98%.

Chen et al. [103] developed a novel deep learning-based architecture for lung cancer segmentation in CT images called MAU-Net. The methodology involves applying a Dual Attention Module [104] at the bottleneck of the U-Net architecture which models the semantic interdependencies in spatial and channel dimensions. A novel multiple attention gate module is proposed to adaptively recalibrate and fuse multiscale features from the dual attention module the previous decode feature maps and the corresponding features from the encoder. Extensive ablation studies were conducted on a clinical dataset comprising 322 CT images to evaluate the performance of the proposed architecture. MAU-Net achieved average DSC, Haus95, and RAJD values of 86.67%, 13.00 and 15.52% respectively. The performance gains for all three metrics illustrate the effectiveness of the proposed MAU-Net. The study also highlights the importance of the attention mechanism for segmentation as the MAU-Net significantly outperformed the base U-Net. Furthermore, the study showed that MAU-Net achieved better segmentation results for smaller cancer areas compared to the base U-Net and DAU-Net. MAU-Net is based on a 3D U-Net as the backbone segmentation network. It incorporates two main attention mechanisms: a dual attention module (DAM) and a multiple attention gated module (MAGM). The DAM module contains a spatial attention block (SAB) and channel attention block (CAB) applied at the bottleneck of the U-Net encoder to model interdependencies between spatial locations and feature channels. The MAGM modules are inserted in the U-Net decoder before concatenation with encoder features. They enable MAU-Net to selectively emphasize informative features and semantic contexts across various scales thereby enhancing lung tumor segmentation. Combining dual attention modeling and multi-scale feature fusion with attention gates enhances the representational capabilities of MAU-Net.

Pan et al. [105] developed a deep learning-based system to automatically measure bone mineral density (BMD) for opportunistic osteoporosis screening utilizing LDCT scans acquired during lung cancer screening. The system was trained and tested on 200 annotated LDCT scans to segment and label all vertebral bodies (VBs). This achieved a mean Dice coefficient of 86.60% for VB segmentation and an accuracy of 97.50% for VB labeling. The mean CT numbers of the trabecular region within the target VBs were derived using the segmentation mask through geometric operations. A linear function was established to correlate the trabecular CT numbers of the target VBs with their corresponding BMDs collected from approved software utilized for osteoporosis diagnosis. The diagnostic performance of the system was assessed using an independent dataset of 374 LDCT scans with established BMD values and osteoporosis diagnoses. The results revealed strong concordance between the predicted BMD values and the actual ground truth. The AUC was 92.7% for osteoporosis detection and 94.2% for discerning low BMD. The developed system holds promise as an automated tool for quantifying vertebral BMD in opportunistic osteoporosis screening through LDCT scans acquired during lung cancer screening.

Shimazaki et al. [106] developed and validated a deep learning-based model to detect lung cancer on chest radiographs through the application of a segmentation technique. The objective was to assess the model's capability in identifying lung cancer nodules/masses and to evaluate its sensitivity and mean false positive indications per image (mFPI) using an independent test dataset. The training dataset encompassed 629 radiographs containing 652 nodules/masses while the test dataset comprised 151 radiographs featuring 159 nodules/masses both collected between January 2006 and June 2018 at the hospital. The DL-based model exhibited a sensitivity of 73% along with an mFPI of 13% in the test

dataset. Nevertheless, the sensitivity was reduced in instances where lung cancers overlapped with challenging areas like pulmonary apices, pulmonary hila, chest wall, heart and subdiaphragmatic space (ranging from 50% to 64%) in contrast to non-overlapping regions (87%). On average, the Dice coefficient for the 159 malignant lesions was 52%. The DL-based model exhibited robust performance in lung cancer detection on chest radiographs characterized by a low mFPI.

Feng et al. [107] aimed to evaluate the diagnostic value of deep learning-optimized chest CT in patients with lung cancer. The study used a Mask-RCNN model for end-to-end image segmentation and a DPN for nodule detection. The study included 90 patients diagnosed with lung cancer through surgery or puncture. The accuracy of the DPN algorithm model in detecting lung lesions in lung cancer patients was 88.74%, and the accuracy of CT diagnosis was 88.37%, with a sensitivity of 82.91% and specificity of 87.43%. When combining deep learning-based CT examination with serum tumor detection, the accuracy improved to 97.94%, sensitivity to 98.12%, and specificity to 100%, showing significant differences ($p < 0.05$).

Gil et al. [108] propose a new approach for using DL in medical imaging analysis by employing two maximum intensity projection (MIP) images generated from whole-body FDG PET volume as inputs to pre-trained models based on 2D images. The methodology involves extracting image features using a pre-trained CNN model, ResNet-50 and depicting the relationship between the images on a parametric 2D axes map using t-distributed stochastic neighbor embedding (t-SNE) with clinicopathological factors. The results show that the DL-based feature map extracted by two MIP images was embedded by t-SNE and the PET images were clustered by clinicopathological features. The clustering pattern showed a difference between the clusters of PET patterns according to the posture of the patient.

Yan et al. [109] investigated the effectiveness of a deep learning algorithm-based CT image segmentation in the diagnosis of lung cancer for perioperative rehabilitation nursing. They constructed a hybrid feature fusion model (HFFM) by fusing a 2D and 3D CNN. Sixty patients with lung cancer were randomly divided into control and intervention groups to receive perioperative routine nursing and rehabilitation nursing. The HFFM showed higher Dice coefficients (87.60%), sensitivities (0.84.90%) and positive predictive values (PPVs) (87.50%) than other models for lung cancer CT image segmentation. The accuracy rate for lung cancer diagnosis was 96.70%. After the nursing intervention, the pulmonary function indexes of the intervention group were significantly improved compared to the control group as reflected by increased PaO₂ (partial pressure of oxygen) levels and decreased PaCO₂ levels.

Chen et al. [110] aimed to establish an auxiliary diagnosis model for lung cancer based on lung CT image scores using the SegNet approach [111] and explore its value in distinguishing benign and malignant lung CT images. They collected CT images from 240 patients, half of whom were diagnosed with early-stage lung cancer and half with benign lung nodules. The authors proposed a SegNet recognition technology to segment images and compared its performance with DeepLab v3 [112]), VGG19 and manual image segmentation. SegNet recognition technology is a deep learning-based approach used for image segmentation, which separates an image into multiple segments or regions based on its content. In the context of the article, the SegNet approach was applied to lung CT images to establish an auxiliary diagnosis model for lung cancer. The SegNet model had the closest overlap rate to manual segmentation and showed a sensitivity of 98.33%, specificity of 86.67%, accuracy of 92.50% and a total segmentation time of 30.42 s which is shorter than manual segmentation.

Choe et al. [113] investigated the effect of different reconstruction kernels on radiomic features and assessed whether image conversion using a CNN could improve the reproducibility of radiomic features between different kernels. The CNN model was developed using residual learning and an end-to-end approach. Kernel-converted images were generated from B30f to B50f and from B50f to B30f. Semi-automatically segmented pulmonary

nodules or masses were analyzed and 702 radiomic features (tumor intensity, texture, and wavelet features) were extracted. The study involved 104 patients with pulmonary nodules or masses (mean age, 63.2 years \pm 10.5), including 54 women and 50 men. The Concordance Correlation Coefficient (CCC) between two readers using the same kernel was 92% and 592 of 702 (84.30%) of the radiomic features were reproducible (CCC \geq 85%). Using different kernels, the CCC was 38% and only 107 of 702 (15.20%) of the radiomic features were reproducible. Texture features and wavelet features were predominantly affected by reconstruction kernel (CCC, from 88% to 61% for texture features and from 0.92 to 0.35 for wavelet features). After applying image conversion, the CCC increased to 84% and 403 of 702 (57.40%) radiomic features were reproducible (CCC 0.85 for texture features and 0.84 for wavelet features).

Yu et al. [114] addressed the limited research on social and behavioral determinants of health (SBDoh) factors in clinical outcomes due to the absence of structured SBDoh information within current electronic health record (EHR) systems. The authors propose utilizing SOTA transformer-based NLP models such as BERT and RoBERTa to extract SBDoh concepts from clinical narratives. The most effective model was utilized to extract SBDoh concepts from a lung cancer screening patient cohort comprising 864 patients and 161,933 diverse clinical notes. The results of the study demonstrated that significantly more detailed information regarding smoking, education and employment was exclusively captured within clinical narratives and that utilizing both clinical narratives and structured EHRs is essential for extracting comprehensive SBDoh information. The NLP model based on BERT achieved the highest strict/lenient F1-scores of 87.91% and 89.99% respectively, demonstrating its effectiveness in extracting SBDoh concepts. Comparing the SBDoh information extracted through NLP with structured EHRs further underscored the limitations of structured EHRs in comprehensively capturing SBDoh information.

Hwang et al. [115] examined the potential improvement in diagnostic yield for newly identified lung metastases on chest radiographs in cancer patients using a deep learning-based CAD system. The investigation employed a regulatory-approved CAD system designed for lung nodules aiming to interpret chest radiographs from patients referred by the medical oncology department in clinical practice. A total of 2916 chest radiographs from 1521 patients underwent analysis in the CAD-assisted interpretation group while 5681 chest radiographs from 3456 patients were subjected to analysis in the conventional interpretation group. The CAD-assisted interpretation group demonstrated a higher diagnostic yield for newly identified metastases (0.86% vs. 0.32%; $p = 0.004$).

6. Discussion

In recent years, the field of cancer detection particularly lung cancer utilizing deep machine learning has seen remarkable advancements. Deep learning's capabilities in representation learning and pattern recognition have demonstrated unparalleled performance compared to traditional diagnostic methods. Yet, the integration of this technology into medical practice presents numerous challenges and opportunities as discussed below.

(1) Utilization of Public and Private Databases:

The development of cancer detection models heavily relies on diverse and rich datasets. The use of both public and private databases has been instrumental in training robust models. Challenges such as data privacy, accessibility and standardization hinder progress. It is advisable to promote collaboration between institutions to build comprehensive datasets, covering various types and stages of cancer adhering to ethical guidelines. The diagnosis and detection of lung cancer through deep learning models heavily rely on the quality and diversity of the datasets used for training. Public databases, such as LIDC-IDRI [21] with 1010 CT scans and LUNA16 [22] with 888 patients and 1186 nodules, ChestX-Ray8 [23] with 108,948 front view X-ray images of 32,717 patients have played a pivotal role in advancing the field. These public databases offer a wealth of data for detection, characterization and segmentation tasks. Private databases, have also contributed to the development of models tailored to specific characteristics and tasks such as classification, detection and localization.

The utilization of both public and private databases ensures a robust and comprehensive evaluation of deep learning models fostering innovation and enhancing performance. The results obtained from different databases reveal important trends in terms of algorithm efficiency and prediction quality. As shown in Table A2, public databases such as LIDC-IDRI, LUNA16 and NLST are frequently used due to their accessibility and standardization which favors the comparability of studies. For example, ref. [33] using CNN-ResNet50 on LIDC achieved an accuracy of 88.41% while [36] with SchCNN achieved an accuracy of 93.03%, illustrating the effectiveness of specific optimization methods despite the use of the same dataset. Data quantity also plays a crucial role in model performance. In private databases, where the number of data may be limited but more specific as in [94] with only 50 lung cancer records a high sensitivity of 96.20% was obtained, suggesting that targeted high-quality data can sometimes compensate for the lack of quantity. However, it should also be noted that large databases do not always guarantee superior performance, but do show the robustness of the model trained to generalize. For example, ref. [96] with 12,360 participants achieved a recall of 95.05% which is comparable to studies using smaller datasets. This may suggest that the quality of data labeling and the specific characteristics of the lung cancer cases in the dataset may have a more significant impact than database size alone. In addition, research advances in network architectures, such as Transformers in [57] or attention methods as in [103] show that algorithmic improvements can also improve performance independently of the data source even though these architectures can benefit from large amounts of data for optimal training. Finally, the choice of datasets whether public or private, as well as their size has a notable impact on the performance of deep learning models for lung cancer diagnosis. Studies using public databases offer a basis for standardized comparison of model performance, but high-quality private datasets can also provide very good results. It is essential to strike a balance between data quantity and quality and to continue developing more robust network architectures to process these data efficiently.

(2) Enhancing Model Interpretability and Ethical Considerations:

Deep learning models often suffer from the “black box” phenomenon where the decision-making process is obscured. Ensuring transparency and ethical considerations in model deployment is paramount. Research should focus on explainable AI techniques and ethical frameworks to ensure patient confidentiality and unbiased decision-making. The “black box” phenomenon often associated with deep learning models poses challenges in medical diagnosis particularly in the field of lung cancer detection. Highlighting this issue, Brocki et al. [37] emphasized the limitations of DNNs advocating for enhanced interpretability. Interpretability is crucial for ensuring trust and transparency in the decision-making process. Research in this area has led to the development of models like ConRad which combines expert-derived radiomics and DNN-predicted biomarkers for CT scans of lung cancer providing a more interpretable classifier compared to traditional CNNs. Efforts to combine various architectures such as CNN-ResNet50 and DenseNet121 with Xception [35] further exemplify the push towards models that not only perform well but also provide insights into their decision-making processes. Ethical considerations like patient confidentiality and unbiased decision-making must also be at the forefront of model deployment. Targeted research into explainable AI techniques and ethical frameworks is essential to address these issues. By adding interpretability, the deep learning community can increase confidence in these advanced models and potentially achieve better results in lung cancer detection and diagnosis.

(3) Development of Lightweight Models for Real-time Applications:

Deep learning’s computational demands may hinder real-time applications in clinical settings. Researchers should explore lightweight models and edge computing to enhance efficiency without sacrificing accuracy. Techniques such as model pruning, quantization and optimization should be explored. While achieving high accuracy is important in medical diagnosis, there exists a critical balance between performance and computational efficiency. Several works in the literature have demonstrated remarkable accuracy levels, exceeding

90% in lung cancer detection tasks. For instance, a study utilizing DAU-Net [103] reported a high performance in terms of DSC, with a 95% Hausdorff distance. Another work [43] showed AUC reaching 93.10% highlighting the capabilities of deep learning models in this domain. However, these high-performing models often utilize complex architectures that are not suited for real-time applications. Techniques such as model pruning, quantization, and optimization should be explored to enhance efficiency without sacrificing accuracy. The success stories of high accuracy must be tempered by the fact that not all architectures can be easily deployed in real-time clinical environments. Developing lightweight models that can be deployed on edge devices offers the potential to revolutionize real-time lung cancer detection and diagnosis making it more accessible and timely. It emphasizes the importance of not only striving for optimal performance but also considering the practical constraints of deployment in a real-world medical environment.

(4) Integration with Existing Medical Practices:

The seamless integration of deep learning models with existing medical practices is a complex undertaking that requires careful consideration of various factors. The integration of diverse data modalities, model outputs and 3D-ResNet with clinical practices has been explored and shown promising results [60]. Challenges in generalization across datasets, interoperability, standardization and technology adaptation must be addressed to ensure that deep learning technologies augment rather than disrupt existing practices. Works like [113] highlight the potential of integrating radiomics with deep learning for clinical diagnostics. Moreover, the development of models that allow for easy integration of useful features and the application of segmentation in clinical practice demonstrates the evolving landscape of integration in lung cancer diagnosis [116]. The convergence of deep learning with traditional medical practices holds great promise but requires a concerted effort in terms of workflow design [98], infrastructure planning, training and policy development. Collaborative initiatives and tailored solutions are needed to foster integration ensuring that the transformative potential of deep learning in lung cancer detection is fully realized.

(5) Research on Personalized Treatment and Predictive Analytics:

The application of deep learning in personalized treatment and predictive analytics is an emerging frontier. Utilizing patient-specific data, genetic information and lifestyle factors can lead to personalized treatment plans and early intervention strategies. Works like the study by Doppalapudi et al. [90] which developed deep learning models for lung cancer survival prediction demonstrate the potential of deep learning in this domain. However, a significant challenge in this area lies in the focus of public databases on known tasks such as segmentation and classification. This limitation often restricts the exploration of personalized treatment methodologies and hinders the development of models tailored to individual patient characteristics. Efforts to curate and standardize public databases that encompass diverse patient profiles and treatment outcomes are essential to unleash the full potential of personalized medicine. The ability to tailor treatments and predict outcomes based on individual patient characteristics offers a pathway towards more personalized and effective care outperforming traditional machine learning models. Continued research and collaboration between oncologists, data scientists and technologists are needed to realize the full potential of personalized treatment and predictive analytics.

7. Conclusions

Lung cancer detection is central to the modern medical landscape where accurate and timely diagnosis can mean the difference between life and death. Traditional diagnostic methods though firmly rooted in medical practice come with their own limitations and constraints. In a transformative shift, the advent of deep learning has begun to reshape the field offering new avenues for detection and personalization. This study looked at the promising applications of deep learning in lung cancer detection exploring five key dimensions in particular: data utilization, interpretability, lightweight model development, integration with existing practices, and personalized treatment. By presenting a range of methodologies, successes and challenges, the study paints a nuanced picture of a rapidly

evolving field breaking traditional frontiers in some areas while addressing new issues in others. The exploration of commonly used datasets, evaluation metrics and innovative architectures provided a broad overview presenting both the potential and limitations of this frontier. Therefore, the study underscored the importance of ethical considerations, technological finesse and collaborative research in steering the future of deep-learning-based lung cancer detection. A notable observation in the current landscape is the relative underutilization of transformer architectures. Transformers, renowned for their ability to capture long-range dependencies and complex patterns, have revolutionized many domains, particularly in natural language processing and computer vision. Unlike conventional convolutional and recurrent neural networks, transformers do not rely on sequential processing, allowing for more parallelism and scalability. This unique capability could be exploited for more complex analyses in lung cancer detection, including the interpretation of complex medical images and multimodal data. The adaptability and flexibility of transformers offer a promising avenue for exploration, potentially enhancing both accuracy and efficiency. Additionally, the development of large-scale Foundation Models (FM), Large language and vision models presents new opportunities for lung cancer detection. These models are pre-trained on massive datasets to learn general world knowledge before being fine-tuned on specialized medical tasks. Their ability to transfer learned knowledge could boost performance on lung cancer datasets which tend to be small and scarce. Models like GPT-3, GPT-4, GPT-4V(ision) and PaLM which demonstrate strong language understanding could aid in reading radiology reports and analyzing patient history. Multimodal FM that combine vision, language and clinical data like Med-PaLM Multimodal may also unlock new diagnostic capabilities. While challenges of explicability and bias must be overcome, FM has great potential to improve deep learning for personalized and accurate lung cancer detection.

Author Contributions: Conceptualization, M.A.A. and H.T.G.; methodology, H.T.G. and M.A.A.; validation, H.T.G. and M.A.A.; formal analysis, H.T.G. and M.A.A.; writing—original draft preparation, H.T.G.; writing—review and editing, M.A.A.; funding acquisition, M.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been supported in part by the New Brunswick Health Research Foundation (NBHRF) and the New Brunswick Innovation Foundation (NBIF), New Brunswick Priority Occupation Student Support Fund (NBPOSS) POF2021-006.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript

ANN	Artificial Neural Network
AUC	Area Under the (ROC) Curve
BERT	Bidirectional Encoder Representation from Transformers
CT	Computed Tomography
DCNN	Deep Convolutional Neural Network
DL	Deep Learning
DNN	Deep Neural Network
GAN	Generative Adversarial Networks
IoU	Intersection Over Union
ML	Machine Learning
PET	Positron Emission Tomography
RNN	Recurrent Neural Network
SEA	Sparse AutoEncoder
ViT	Vision Transformer

Appendix A. Details informations about Metrics and Results of Works Using Private/Public Datasets

Table A1. Metrics commonly used for lung cancer diagnosis.

Metric	Definition	Note	Task
Accuracy	$Acc = \frac{TP+TN}{TP+TN+FP+FN}$	Quantifies the model's effectiveness in correctly classifying both positive and negative instances	Classification
Sensitivity (Recall)	$Sens = \frac{TP}{TP+FN}$	Measures the proportion of actual positive cases that are correctly identified	Classification
Specificity	$Spec = \frac{TN}{TN+FP}$	Measures the proportion of actual negative cases that are correctly identified	Classification
Precision	$Prec = \frac{TP}{TP+FP}$	Measures the proportion of positive identifications that are actually correct	Classification
F1-score	$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$	Harmonic mean of Precision and Recall, provides a balanced measure	Classification, Segmentation
ROC	Receiver Operating Characteristic Curve	Graphical representation of Sensitivity vs (1-Specificity)	Classification
Dice (DSC)	$Dice = \frac{2 \times A \cap B }{ A + B }$	Measures the similarity of two sets, commonly used for image segmentation	Segmentation
IOU	$IOU = \frac{ A \cap B }{ A \cup B }$	Intersection over Union, measures the overlap between two sets	Segmentation

TP (True Positives) represents instances correctly classified as positive, TN (True Negatives) represents instances correctly classified as negative, FP (False Positives) represents instances incorrectly classified as positive and FN (False Negatives) instances incorrectly classified as negative. Also A and B are two sets (for example, the sets of pixels in a segmented image), $|A|$ and $|B|$ are the sizes of these sets, and $|A \cap B|$ is the number of elements common to both sets.

Table A2. Results of work using Public datasets.

Ref.	Methodology	Dataset	Results(%)	Tasks
[33]	Application of CNN-ResNet50 combined with RBF SVM to 11 datasets generated by different deep extractors.	LIDC	Accuracy = 88.41, AUC = 93.19	Classification
[34]	CNN has been trained with a learning rate of 0.01 and a batch size of 32. There are two convolution operations, each with 32 filters and a kernel size of 5. An aggregate layer with a kernel size of 2 is used to prevent excessive motion.	LIDC	Accuracy = 84.15, Sensitivity = 83.96	Classification
[35]	DenseNet121 uses identity connections between layers, giving each layer access to the characteristics of all previous layers. This increases the use of information from all layers without increasing the complexity of the model.	LIDC	Accuracy = 87.67, Specificity = 93.38, Precision = 87.88, AUC = 93.79	Classification
[37]	ConRad is designed to extract various features from cancer images using both biomarkers predicted by CBM and radiomic features. CBM predicts biomarkers such as subtlety, calcification, sphericity, margin, lobulation, spiculation, texture and diameter.	LIDC	AUC = 96.10	Classification

Table A2. Cont.

Ref.	Methodology	Dataset	Results(%)	Tasks
[39]	DBN models the joint distribution of lung nodule images and hidden neural network layers. It is built iteratively using RBMs stacked on top of each other. Each MBR consists of a visible and a hidden layer. RBMs are trained using stochastic gradient descent and the contrastive divergence algorithm.	LIDC	Sensitivity = 73.40, Specificity = 82.20	Classification
[40]	CAET-SWin (Transformer) combines spatial and temporal features extracted using two parallel self-attention mechanisms to perform malignancy prediction based on CT images. It takes advantage of the 3D structure of unthin CT scans by simultaneously extracting inter-slice and intra-slice features. The extracted features are then merged to form the final output.	LIDC	Accuracy = 82.65, Sensitivity = 83.66, Specificity = 81.66	Classification
[41]	LungNet extracts distinctive features from CT scans. It includes three 3D convolution layers with dimensions $16 \times 3 \times 3$ and a 3D maximum aggregate layer with kernel size = 2, stride = 2. Three fully connected layers with decreasing feature vectors (128, 64, and 64) were combined to reduce the dimensionality of the learning convergence characteristics of the model.	LIDC	AUC = 85.00	Classification
[43]	ResNet50 was used as a feature extractor. The extracted features were converted into feature vectors. Each nodule image was transformed into a digital vector representing the features extracted by the model. These feature vectors served as input data for the RBF SVM classifier.	LIDC	AUC = 93.10	Detection
[45]	3D-CGAN generates realistic images of lung nodules with various conditions such as size, ground attenuation and presence of surrounding tissue. CGAN consists of a generator and two discriminators (context and node). The generator takes as input noise regions (noise bins) of fixed size and generates realistic lung nodes adhering to specified size and damping conditions. It also uses context information, such as surrounding tissue.	LIDC	CPM = 55.00	Detection
[46]	A recognition system based on the Faster R-CNN model with a regional recommendation network consisting of 27 convolutional layers is proposed. The model uses 3D convolutional layers to extract three-dimensional information from chest CT images.	LIDC	Sensitivity = 96.00	Detection
[47]	CNN classifies INCs into lung nodules through convolutional layers. Each layer applies filters to extract important information from the input image. These filters detect certain patterns in an image, such as edges, shapes or textures.	LIDC	Sensitivity = 94.01, AUC = 82.00	Detection

Table A2. Cont.

Ref.	Methodology	Dataset	Results(%)	Tasks
[36]	ShCNN is trained using a WSLnO-based deformable model for lung region segmentation. The results show that the ShCNN model based on the proposed method outperforms several other existing methods, including CNN, IPCT+NN, dictionary-based segmentation+ShCNN and the deformable model based on WCBA+ShCNN.	LIDC	Precision = 93.03	Segmentation
[48]	A CNN based on the VGG16Net architecture is trained to classify thoracic CT slices. It uses a full convolutional layer structure and a convolutional layer structure + a global average clustering layer (Conv + GAP), resulting in a FC layer. A nodule activation map (NAM) generated by a weighted average of the activation maps with weights learned in the FC layer.	LIDC	Accuracy = 86.60	Segmentation
[49]	iW-Net works in 2 ways. Automatic segmentation: iW-Net receives as input a cube of fixed dimensions centered on the nodule, identified either manually by the user or automatically by the system. The network proposes an initial segmentation of the nodule. Interactive segmentation: If the user is not satisfied with the proposed segmentation, he can adjust it by manually inserting the ends of a line representing the nodule's diameter. iW-Net then integrates this information to refine the segmentation.	LIDC	IoU = 55.00	Segmentation
[51]	U-Net's symmetrical architecture, consisting of encoder and decoder blocks with jump connections, enables detailed low-level information to be retained and combined with high-level information. This improves the accuracy of nodular pixel localization in the image.	LIDC	DSC = 83.00	Segmentation
[53]	NoduleNet combines node detection, false positive reduction and node segmentation in a unified framework trained for multiple tasks. This unified approach improves model performance by solving several aspects of the node detection problem.	LIDC	CPM = 87.27, DSC = 83.10	Segmentation
[52]	MV-CNN through Multivue structure incorporates three branches processing axial, coronal and sagittal views of CT images separately. This multiview approach enables the model to capture 3D information without requiring the input of a full 3D volume, thus reducing data redundancy and increasing efficiency.	LIDC	DSC = 77.67, ASD = 24.00	Segmentation
[62]	3D-MSViT processes information at different scales, capturing both fine details and global characteristics of nodules through the Patch Embedding Block. It processes the patch features at each scale dimension of CT images individually using the Local Transformer Block. Feature maps of different scales are scaled to uniform resolution and merged into a unified representation using the Global Transformer Block.	LUNA16, LIDC	Sensitivity = 97.81	Classification

Table A2. Cont.

Ref.	Methodology	Dataset	Results(%)	Tasks
[60]	3D ECA-ResNet introduces residual connections (skip connections), effectively alleviates the problem of gradient disappearance, enabling feature reuse and faster information transmission. It emphasizes channel information by explicitly modeling the correlation between them. It adaptively adjusts the feature channel, thereby strengthening the feature extraction capability of the network.	LUNA16, LIDC-IDRI	Accuracy = 94.89	Classification
[57]	Swin Transformer transforms CT images into non-overlapping blocks through patching operations, including embedding, merging, and masking patches. This method allows efficient processing of images that are not naturally sequential, as is the case with CT images. Introducing connections between non-overlapping windows in consecutive Swin Transformer blocks improves network modeling capabilities.	LUNA16	Accuracy = 82.26	Classification
[59]	Dilated SegNet helps enlarge the receptive field of filters with its dilated convolution layers. This helps capture broader features of CT images, which is effectively useful for detecting smaller nodules.	LUNA16	DSC = 89.00	Segmentation
[56]	The framework uses an adapted version of Faster R-CNN which includes 2 RPNs and a deconvolution layer, to detect nodule candidates. Multiple CNNs are trained sequentially each handling more difficult cases than the previous model. This boosting approach helps increase sensitivity for detection of such small lung nodules.	LUNA16	Sensitivity = 86.42	Segmentation
[63]	A 3D CNN model is used to analyze full CT volumes end-to-end. This in-depth analysis can detect cancer candidate regions in CT volumes. This allows potentially cancerous areas to be precisely located, facilitating a more targeted assessment. An additional CNN model is developed to predict cancer risk based on the outputs of ROI detection models and full volume analysis. This model can also incorporate regions from previous CT scans of the patient, allowing for longitudinal comparison and more accurate risk assessment.	NLST	AUC = 94.40	Classification
[66]	Time-distance ViT is proposed to interpret temporal distances in longitudinal and irregularly sampled medical images. The method uses continuous time vector embeddings to integrate temporal information into image analysis. Time encoding is performed with sinusoids at different frequencies, allowing a linear representation of temporal distances. TEM is used to modulate self-attention weights based on the time elapsed between image acquisitions.	NLST	AUC = 78.60	Classification

Table A2. Cont.

Ref.	Methodology	Dataset	Results(%)	Tasks
[68]	CXR-LC based on a CNN, it combines radiographic images with basic information (age, gender, smoking status). It also leverages transfer learning from the Inception v4 network to predict all-cause mortality in the PLCO trial.	NLST	AUC = 75.50	Classification
[64]	The proposed model is based on the VGG16 2D CNN architecture using transfer learning which made it possible to take advantage of a rich and varied knowledge base for a specific task, thus improving the accuracy and speed of learning of the model. The adaptation of this architecture made it possible to benefit from its 13 convolutional layers and its 4 pooling layers for efficient extraction of features from lung images.	NLST	Accuracy = 90.40, F1-score = 90.10	Classification
[70]	FDTrans implements a preprocessing process to convert histopathological images to YCbCr color space and then to spatial spectrum via DCT. This step captures relevant frequency information, essential for distinguishing the subtle details of cancerous tissue. CSAM reallocates weights between low- and high-frequency information channels. CDTB processes features of Y, Cb and Cr channels, capturing long-term dependencies and global contextual connections between different components of images.	TCGA	AUC = 93.16	Classification
[69]	Gene Transformer combines multi-head attention mechanics with 1D layers. Multi-head attention allows the model to simultaneously process complex genomic information from thousands of genes from different patient samples. the attention mechanism sequentially selects subsets of genes and reveals a set of scores defining the importance of each gene for subtype classification, focusing only on genes relevant to a task.	TCGA	Accuracy = 100, Precision = 100, Recall = 100, F1-Score = 100	Classification
[74]	The study explores the use of DenseNet121 combined with transfer learning. The structure of DenseNet121 solves the problem of the disappearing gradient. The model also provides CAMs to identify the location of lung nodules. This allows visualization of the most salient regions on the images used to identify the output class.	JSRT	Specificity = 74.96	Classification
[75]	U-Net is used to accurately segment the left and right lung fields in standard CXRs. This segmentation allows the lungs to be isolated from the heart and other parts in the images, thus leading to a more focused analysis of suspicious lesions and lung nodules.	JSRT	Accuracy = 96.00	Segmentation

Table A2. Cont.

Ref.	Methodology	Dataset	Results(%)	Tasks
[76]	The study compares the approaches and performances of award-winning algorithms developed during the Kaggle Data Science Bowl. U-Net has been commonly used for lung nodule segmentation. The study found substantial performance variations in the public and final test sets. Transfer learning has been used in most classification algorithms, highlighting U-Net's ability to adapt and learn from pre-existing data, which is essential when working with limited datasets or specific.	DSB	Logloss = 39.97	Segmentation
[77]	The VGG-like 3D multipath network takes advantage of multiple paths to process 3D volumetric data which allows a better understanding of the spatial structure of lung nodules. The network is able to distinguish not only the presence of lung nodules but also classify their level of malignancy, an essential step in the early diagnosis and treatment planning of lung cancer.	LIDC, LUNA16, DSB	Accuracy = 95.60, Logloss = 38.77, DSC = 90.00	Segmentation
[80]	TSMUNet integrates a transformer module to process anisotropic data in CT images. This integration allows the model to more effectively adapt to variations in slice spacing, a common challenge in medical images. The model encodes information along the Z axis by representing the feature maps of a cut as a weighted sum of these maps and the feature maps of neighboring cuts.	MSD	Dice = 87.17	Segmentation
[78]	The system consists of two main components: a segmentation part based on UNETR and a classification part based on a self-supervised network. UNETR uses transformers as an encoder to efficiently capture global multiscale information, thereby learning sequential representations of the input volume. For the classification of segmented nodules, the system uses a self-supervised architecture. This architecture focuses on predicting the same class for two different perspectives of the same sample, allowing labels and representations to be learned simultaneously in a single end-to-end process.	MSD	Accuracy = 98.77, Accuracy = 97.83	Classification, Segmentation
[81]	Proposes an integrated framework for the detection of pulmonary nodules from low-dose CT scans using a model based on a 3D CNN and a 3D RPN network. This approach combines the steps of nodule screening and false positive reduction into a single jointly trained model. 3D RPN is adapted from the Faster-RCNN model for generating nodule candidates.	Tianchi	CPM = 86.60	Detection

Table A2. Cont.

Ref.	Methodology	Dataset	Results(%)	Tasks
[82]	Amalgamated-CNN before segmenting CT images uses a non-sharpening mask to enhance the signal from nodules. It includes three separate CNN networks (CNN-1, CNN-2, CNN-3) with different input sizes and number of layers. This multi-layer design allows the model to process and analyze images at different scales. It uses AdaBoost classifier to merge the results.	Tianchi, LUNA16	Sensitivity = 85.10	Segmentation
[83]	DCNN consists of two main steps: detection of nodule candidates and reduction of false positives. The model uses a 3D version of the Faster R-CNN network, inspired by U-Net. A 3D DCNN classifier is then used to finely discriminate between true nodules and false positives.	Tianchi	CPM = 81.50	Segmentation
[87]	BERT-BTN is used for clinical entity extraction from Chinese CT reports. BERT is used to learn deep semantic representations of characters which is essential for understanding the complex context of CT reports. A BiLSTM layer is used after BERT to capture nested structures and latent dependencies of each character in reports.	PUCH	Macro-F1 = 85.96	Named Entity Recognition
[89]	Clinical Transformer is an adaptation of the Transformer architecture for precision medicine aimed at modeling the relationships between molecular and clinical measurements and survival of cancer patients. It aims to model how a biomarker in the context of other clinical or molecular features can influence patient survival particularly in immunotherapy treatment.	MSKCC	C-Index = 61.00	Predicting survival in cancer patients
[90]	ANN demonstrated superior ability to model complex, nonlinear decision boundaries although it had difficulty clearly distinguishing between survival classes of less than 6 months and 0.5 to 2 years.	SEER	Accuracy = 71.18	Classification
[91]	DETR is based on Transformers and aims to fully automate the anatomical localization of lung cancer in PET/CT images. It integrates global attention to the entire image allowing better localization and classification of tumors.	TCIA	IoU = 80.00, Accuracy = 97.00	Segmentation Classification

Table A3. Deep Learning Methods for Lung Cancer Diagnosis in Private Databases.

Ref.	Method	Dataset Size	Performance Metrics	Task
[92]	Retina U-Net is suitable for the detection of primary lung tumors and associated metastases at all stages on FDG-PET/CT images. It has been adapted and trained to specifically detect T, N, and M lesions on these images. It contains additional branches in the lower levels of the decoder for end-to-end object classification and bounding box regression.	364 FDG-PET/CTs	Sensitivity = 86.20	Classification

Table A3. Cont.

Ref.	Method	Dataset Size	Performance Metrics	Task
[93]	DCNN coupled with transfer learning with VGG16 and finetuning were used by converting 3D images into 2D. The optimization of the hyperparameters of the DCNN was carried out by random search which made it possible to identify the most effective parameters for this specific classification task.	1236 patients	Accuracy = 68.00	Classification
[94]	ODNN combined with LDA was used to extract deep features from lung CT images and reduce their dimensionality. It was optimized using MGSA to improve lung cancer classification. This optimization made it possible to refine the structure of the network and improve its classification performance.	50 lung cancer CT images	Sensitivity = 96.20	Classification
[95]	ResNet is used to analyze exosomes in blood plasma via SERS. The model was trained to distinguish between exosomes derived from normal cells and those from lung cancer cell lines.	2150 cell-derived exosome data	Accuracy = 95.00 AUC = 91.20	Classification
[96]	A system including a deep learning model is developed to improve lung cancer screening in rural China, using mobile CT scanners. The model was designed to identify suspicious nodule candidates from LDCT images. After nodule detection, the model evaluates the probability of malignancy of each nodule.	12,360 participants	Recall = 95.07	Classification
[98]	RCNN combined with transformers is applied to syndrome differentiation in traditional Chinese medicine (TCM) for lung cancer diagnosis. The model was designed to process unstructured medical records. The use of data augmentation helped improve the performance of the RCNN model. The integration of transformers made it possible to efficiently process long and complex text sequences.	1206 clinical records	F1-score = 86.50	Classification
[99]	3D U-Net with dual inputs is used for both PET and CT. It features two parallel convolution paths for independent feature extraction from PET and CT images at multiple resolution levels, thereby optimizing the analysis of features specific to each type of imaging. The features extracted from the convolution arms were concatenated and fed the deconvolution path via skip connections.	290 pairs of PET and CT	Mean DSC = 92.00	Segmentation
[100]	Dual-input U-Net implements a two-step approach. In the first step a global U-Net processes the 3D PET/CT volume to extract the preliminary tumor area while in the second step a regional U-Net refines the segmentation on slices selected by the global U-Net.	887 patients	N/A	Segmentation

Table A3. Cont.

Ref.	Method	Dataset Size	Performance Metrics	Task
[103]	3D MAU-Net is an adaptation of U-Net. At the bottleneck of the U-Net, a dual attention module was integrated to model semantic interdependencies in the spatial and channel dimensions respectively. The model proposes a multiple attention module to adaptively recalibrate and fuse the multi-scale features from the dual attention module of the previous feature maps of the decoder and the corresponding features of the encoder.	322 CT images	DSC = 86.67	Segmentation
[105]	3D U-Net is intended for automatic measurement of bone mineral density (BMD) using LDCT scanners for opportunistic osteoporosis screening. Using 3DU-Net combined with dense connections achieved a balance between performance and computation.	200 annotated LDCT scans, 374 independent LDCT scans	AUC = 92.70	Segmentation
[106]	Unlike detection methods that provide a bounding box or classification methods that determine malignancy from a single image, this CNN model is based on an encoder-decoder architecture that reduces the resolution of the feature map and improves the robustness of the model to noise and overfitting.	629 radiographs with 652 nodules/masses	DSC = 52.00	Segmentation
[107]	Mask-RCNN is an evolution of the Faster-RCNN model which replaces the ROI pooling layer of Faster-RCNN with a more efficient ROI Alignment layer. This modification allows a more precise correspondence between the output pixels and the input pixels thus effectively preserving the spatial data contained in the image.	90 patients	Sensitivity = 82.91	Segmentation
[108]	ResNet-50 is used to analyze maximum intensity projection (MIP) images from PET in lung cancer patients by extracting features from MIP images of anterior and lateral views. These MIP images are generated from 3D PET volumes, ResNet-50 being a 2D model avoids the difficulty of directly processing 3D volumetric data.	N/A	chi-square = 23.6	Segmentation
[109]	HFFM combines the advantages of 2D and 3D neural networks. 3D CNN is capable of learning three-dimensional information from CT images while 2D CNN focuses on obtaining detailed semantic information. This combination allows the model to capture both the complex spatial features and fine details of CT images.	60 patients	Dice = 87.60	Segmentation
[110]	SegNet uses a decoding process that is based on the pooling indices calculated during the maximum pooling stage of the corresponding encoder. This approach allows for efficient nonlinear sampling and eliminates the need to learn upsamples.	240 participants	Sensitivity = 98.33	Segmentation

Table A3. Cont.

Ref.	Method	Dataset Size	Performance Metrics	Task
[113]	The CNN used provided a significant improvement in the reproducibility of radiomic characteristics between the different reconstruction kernels. While the concordance of the radiomic features between the B30f and B50f cores was initially low, the application of the CNN improved them thus making the radiomic features much more reliable and comparable between the different cores.	104 patients	CCC = 92.00	Radiomic
[114]	BERT and RoBERTa were developed to extract information on social and behavioral determinants of health (SBDoH) from unstructured clinical text in electronic health records (EHR).	161,933 clinical notes	Strict-F1-score = 87.91 Lenient-F1-score = 89.99	Extract SBDoH
[115]	The deep learning-based CAD system significantly increased the diagnostic yield for detecting new visible lung metastases. The actual positive detection rate was higher in the CAD-assisted interpretation group than in the conventional interpretation group.	2916 chest radiographs from 1521 patients	80.00	Evaluate deep learning-based CAD system

References

- World Health Organization. Cancer. 2023. Available online: <https://www.who.int/news-room/fact-sheets/detail/cancer> (accessed on 18 September 2023).
- Freitas, C.; Sousa, C.; Machado, F.; Serino, M.; Santos, V.; Cruz-Martins, N.; Teixeira, A.; Cunha, A.; Pereira, T.; Oliveira, H.P.; et al. The role of liquid biopsy in early diagnosis of lung cancer. *Front. Oncol.* **2021**, *11*, 634316. [[CrossRef](#)] [[PubMed](#)]
- Ali, S.; Li, J.; Pei, Y.; Khurram, R.; Rehman, K.u.; Rasool, A.B. State-of-the-Art challenges and perspectives in multi-organ cancer diagnosis via deep learning-based methods. *Cancers* **2021**, *13*, 5546. [[CrossRef](#)] [[PubMed](#)]
- Hosny, A.; Parmar, C.; Quackenbush, J.; Schwartz, L.H.; Aerts, H.J. Artificial intelligence in radiology. *Nat. Rev. Cancer* **2018**, *18*, 500–510. [[CrossRef](#)] [[PubMed](#)]
- Chiu, H.Y.; Chao, H.S.; Chen, Y.M. Application of artificial intelligence in lung cancer. *Cancers* **2022**, *14*, 1370.
- Dodia, S.; Annappa, B.; Mahesh, P.A. Recent advancements in deep learning based lung cancer detection: A systematic review. *Eng. Appl. Artif. Intell.* **2022**, *116*, 105490.
- Riquelme, D.; Akhloufi, M.A. Deep learning for lung cancer nodules detection and classification in CT scans. *Ai* **2020**, *1*, 28–67.
- Zareian, F.; Rezaei, N. *Application of Artificial Intelligence in Lung Cancer Detection: The Integration of Computational Power and Clinical Decision-Making*; Springer International Publishing: Edinburgh, UK, 2022; pp. 1–14.
- Huang, S.; Yang, J.; Shen, N.; Xu, Q.; Zhao, Q. Artificial intelligence in lung cancer diagnosis and prognosis: Current application and future perspective. *Semin. Cancer Biol.* **2023**, *89*, 30–37. [[CrossRef](#)]
- Qureshi, R.; Zou, B.; Alam, T.; Wu, J.; Lee, V.; Yan, H. Computational methods for the analysis and prediction of EGFR-mutated lung cancer drug resistance: Recent advances in drug design, challenges and future prospects. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2022**, *20*, 238–255.
- Al-Tashi, Q.; Saad, M.B.; Muneer, A.; Qureshi, R.; Mirjalili, S.; Sheshadri, A.; Le, X.; Vokes, N.I.; Zhang, J.; Wu, J. Machine Learning Models for the Identification of Prognostic and Predictive Cancer Biomarkers: A Systematic Review. *Int. J. Mol. Sci.* **2023**, *24*, 7781. [[CrossRef](#)]
- Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Int. J. Surg.* **2021**, *88*, 105906. [[CrossRef](#)]
- Cheng, P.M.; Montagnon, E.; Yamashita, R.; Pan, I.; Cadrin-Chenevert, A.; Perdigón Romero, F.; Chartrand, G.; Kadoury, S.; Tang, A. Deep learning: An update for radiologists. *Radiographics* **2021**, *41*, 1427–1445. [[PubMed](#)]
- Swift, A.; Heale, R.; Twycross, A. What are sensitivity and specificity? *Evid.-Based Nurs.* **2020**, 2–4. [[CrossRef](#)]
- Mandrekar, J.N. Receiver operating characteristic curve in diagnostic test assessment. *J. Thorac. Oncol.* **2010**, *5*, 1315–1316. [[CrossRef](#)] [[PubMed](#)]
- Davis, J.; Goadrich, M. The relationship between Precision-Recall and ROC curves. In Proceedings of the 23rd International Conference on Machine Learning, 2006; pp. 233–240.

17. Goutte, C.; Gaussier, E. A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In Proceedings of the European Conference on Information Retrieval, Santiago de Compostela, Spain, 21–23 March 2005; pp. 345–359.
18. Janssens, A.C.J.; Martens, F.K. Reflection on modern methods: Revisiting the area under the ROC Curve. *Int. J. Epidemiol.* **2020**, *49*, 1397–1403. [[CrossRef](#)] [[PubMed](#)]
19. Zou, K.H.; Warfield, S.K.; Bharatha, A.; Tempany, C.M.; Kaus, M.R.; Haker, S.J.; Wells III, W.M.; Jolesz, F.A.; Kikinis, R. Statistical validation of image segmentation quality based on a spatial overlap index1: Scientific reports. *Acad. Radiol.* **2004**, *11*, 178–189. [[CrossRef](#)]
20. Sun, C.; Shrivastava, A.; Singh, S.; Gupta, A. Revisiting unreasonable effectiveness of data in deep learning era. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 843–852.
21. Prior, F.; Smith, K.; Sharma, A.; Kirby, J.; Tarbox, L.; Clark, K.; Bennett, W.; Nolan, T.; Freymann, J. The public cancer radiology imaging collections of The Cancer Imaging Archive. *Sci. Data* **2017**, *4*, 1–7. [[CrossRef](#)]
22. Setio, A.; Traverso, A.; De Bel, T.; Berens, M.S.; Bogaard, C.v.d.; Cerello, P.; Chen, H.; Dou, Q.; Fantacci, M.E.; Geurts, B.; et al. Lung Nodule Analysis 2016 (LUNA16) Dataset. 2016. Available online: <https://luna16.grand-challenge.org/> (accessed on 27 March 2023).
23. Wang, X.; Peng, Y.; Lu, L.; Lu, Z.; Bagheri, M.; Summers, R.M. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In Proceedings of the the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017.
24. Yanagita, S.; Imahana, M.; Suwa, K.; Sugimura, H.; Nishiki, M. Image Format Conversion to DICOM and Lookup Table Conversion to Presentation Value of the Japanese Society of Radiological Technology (JSRT) Standard Digital Image Database. *Nihon Hoshasen Gijutsu Gakkai Zasshi* **2016**, *72*, 1015–1023. [[CrossRef](#)]
25. Team, N.L.S.T.R. National Lung Screening Trial (NLST) dataset. *N. Engl. J. Med.* **2011**, 395–409.
26. Samstein, R.M.; Lee, C.H.; Shoushtari, A.N.; Hellmann, M.D.; Shen, R.; Janjigian, Y.Y.; Barron, D.A.; Zehir, A.; Jordan, E.J.; Omuro, A.; et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat. Genet.* **2019**, *51*, 202–206.
27. Tomczak, K.; Czerwińska, P.; Wiznerowicz, M. Review The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. *Contemp. Oncol. Onkol.* **2015**, *2015*, 68–77. [[CrossRef](#)]
28. National Cancer Institute. Surveillance, Epidemiology, and End Results Program (SEER) Database. 2023. Available online: <https://seer.cancer.gov/data/> (accessed on 28 March 2023).
29. Simpson, A.L.; Antonelli, M.; Bakas, S.; Bilello, M.; Farahani, K.; Van Ginneken, B.; Kopp-Schneider, A.; Landman, B.A.; Litjens, G.; Menze, B.; et al. A large annotated medical image dataset for the development and evaluation of segmentation algorithms. *arXiv* **2019**, arXiv:1902.09063.
30. Cancer Imaging Archive. A Large-Scale CT and PET/CT Dataset for Lung Cancer Diagnosis (Lung-PET-CT-Dx). 2013. Available online: <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=70224216> (accessed on 28 March 2023). [[CrossRef](#)]
31. Tianchi, A. Tianchi Medical AI Competition [Season 1]: Intelligent Diagnosis of Pulmonary Nodules. 2017. Available online: <https://tianchi.aliyun.com/competition/entrance/231601/information> (accessed on 6 April 2023).
32. Booz Allen, K. Kaggle Data Science Bowl 2017. 2017. Available online: <https://www.kaggle.com/c/data-science-bowl-2017> (accessed on 6 April 2023).
33. Da Nóbrega, R.V.M.; Peixoto, S.A.; da Silva, S.P.P.; Rebouças Filho, P.P. Lung nodule classification via deep transfer learning in CT lung images. In Proceedings of the 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS), Karlstad, Sweden, 18–21 June 2018; pp. 244–249.
34. Song, Q.; Zhao, L.; Luo, X.; Dou, X. Using deep learning for classification of lung nodules on computed tomography images. *J. Healthc. Eng.* **2017**, *2017*, 8314740. [[CrossRef](#)] [[PubMed](#)]
35. Zhang, Q.; Wang, H.; Yoon, S.W.; Won, D.; Srihari, K. Lung nodule diagnosis on 3D computed tomography images using deep convolutional neural networks. *Procedia Manuf.* **2019**, *39*, 363–370. [[CrossRef](#)]
36. Shetty, M.V.; Tunga, S. Optimized Deformable Model-based Segmentation and Deep Learning for Lung Cancer Classification. *J. Med. Investig.* **2022**, *69*, 244–255.
37. Brocki, L.; Chung, N.C. Integration of Radiomics and Tumor Biomarkers in Interpretable Machine Learning Models. *Cancers* **2023**, *15*, 2459.
38. Brocki, L.; Chung, N.C. ConRad. Available online: <https://github.com/lenbrocki/ConRad> (accessed on 2 May 2023).
39. Hua, K.L.; Hsu, C.H.; Hidayati, S.C.; Cheng, W.H.; Chen, Y.J. Computer-aided classification of lung nodules on computed tomography images via deep learning technique. *Oncotargets Ther.* **2015**, *8*, 2015–2022.
40. Khademi, S.; Heidarian, S.; Afshar, P.; Naderkhani, F.; Oikonomou, A.; Plataniotis, K.; Mohammadi, A. Spatio-Temporal Hybrid Fusion of CAE and SWIn Transformers for Lung Cancer Malignancy Prediction. *arXiv* **2022**, arXiv:2210.15297.
41. Mukherjee, P.; Zhou, M.; Lee, E.; Schicht, A.; Balagurunathan, Y.; Napel, S.; Gillies, R.; Wong, S.; Thieme, A.; Leung, A.; et al. A shallow convolutional neural network predicts prognosis of lung cancer patients in multi-institutional computed tomography image datasets. *Nat. Mach. Intell.* **2020**, *2*, 274–282. [[CrossRef](#)]

42. Mukherjee, P.; Zhou, M.; Lee, E.; Gevaert, O. LungNet: A Shallow Convolutional Neural Network Predicts Prognosis of Lung Cancer Patients in Multi-Institutional CT-Image Data. 2020. Available online: <https://codeocean.com/capsule/5978670/tree/v1> (accessed on 18 March 2023).
43. da Nobrega, R.V.M.; Reboucas Filho, P.P.; Rodrigues, M.B.; da Silva, S.P.; Dourado Junior, C.M.; de Albuquerque, V.H.C. Lung nodule malignancy classification in chest computed tomography images using transfer learning and convolutional neural networks. *Neural Comput. Appl.* **2020**, *32*, 11065–11082.
44. Ridnik, T.; Ben-Baruch, E.; Noy, A.; Zelnik-Manor, L. Imagenet-21k pretraining for the masses. *arXiv* **2021**, arXiv:2104.10972.
45. Han, C.; Kitamura, Y.; Kudo, A.; Ichinose, A.; Rundo, L.; Furukawa, Y.; Umemoto, K.; Li, Y.; Nakayama, H. Synthesizing diverse lung nodules wherever massively: 3D multi-conditional GAN-based CT image augmentation for object detection. In Proceedings of the 2019 International Conference on 3D Vision (3DV), Quebec, QC, Canada, 16–19 September 2019; pp. 729–737.
46. Katase, S.; Ichinose, A.; Hayashi, M.; Watanabe, M.; Chin, K.; Takeshita, Y.; Shiga, H.; Tateishi, H.; Onozawa, S.; Shirakawa, Y.; et al. Development and performance evaluation of a deep learning lung nodule detection system. *BMC Med. Imaging* **2022**, *22*, 203. [[CrossRef](#)] [[PubMed](#)]
47. Tan, J.; Huo, Y.; Liang, Z.; Li, L. Expert knowledge-infused deep learning for automatic lung nodule detection. *J.-Ray Sci. Technol.* **2019**, *27*, 17–35.
48. Feng, X.; Yang, J.; Laine, A.F.; Angelini, E.D. Discriminative localization in CNNs for weakly-supervised segmentation of pulmonary nodules. In Proceedings of the Medical Image Computing and Computer Assisted Intervention- MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, 11–13 September 2017; Proceedings, Part III 20; pp. 568–576.
49. Aresta, G.; Jacobs, C.; Araújo, T.; Cunha, A.; Ramos, I.; van Ginneken, B.; Campilho, A. iW-Net: An automatic and minimalistic interactive lung nodule segmentation deep network. *Sci. Rep.* **2019**, *9*, 11591. [[CrossRef](#)] [[PubMed](#)]
50. Aresta, G.; Jacobs, C.; Araújo, T.; Cunha, A.; Ramos, I.; van Ginneken, B.; Campilho, A. iW-Net: Source Code. 2019. Available online: <https://github.com/gmaresta/iW-Net> (accessed on 25 March 2023).
51. Rocha, J.; Cunha, A.; Mendonça, A.M. Conventional filtering versus u-net based models for pulmonary nodule segmentation in ct images. *J. Med. Syst.* **2020**, *44*, 81.
52. Wang, S.; Zhou, M.; Gevaert, O.; Tang, Z.; Dong, D.; Liu, Z.; Jie, T. A multi-view deep convolutional neural networks for lung nodule segmentation. In Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Republic of Korea, 11–15 July 2017; pp. 1752–1755.
53. Tang, H.; Zhang, C.; Xie, X. Nodulenet: Decoupled false positive reduction for pulmonary nodule detection and segmentation. In Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, 13–17 October 2019; Proceedings, Part VI 22; pp. 266–274.
54. Tang, H.; Zhang, C. LungNet Code. Github. 2019. Available online: <https://github.com/uci-cbcl/NoduleNet> (accessed on 20 March 2023).
55. Xiao, N.; Qiang, Y.; Bilal Zia, M.; Wang, S.; Lian, J. Ensemble classification for predicting the malignancy level of pulmonary nodules on chest computed tomography images. *Oncol. Lett.* **2020**, *20*, 401–408. [[CrossRef](#)]
56. Xie, H.; Yang, D.; Sun, N.; Chen, Z.; Zhang, Y. Automated pulmonary nodule detection in CT images using deep convolutional neural networks. *Pattern Recognit.* **2019**, *85*, 109–119. [[CrossRef](#)]
57. Sun, R.; Pang, Y.; Li, W. Efficient Lung Cancer Image Classification and Segmentation Algorithm Based on an Improved Swin Transformer. *Electronics* **2023**, *12*, 1024. [[CrossRef](#)]
58. Touvron, H.; Cord, M.; Douze, M.; Massa, F.; Sablayrolles, A.; Jégou, H. Training data-efficient image transformers & distillation through attention. In Proceedings of the International Conference on Machine Learning, Virtual, 18–24 July 2021; pp. 10347–10357.
59. Agnes, S.A.; Anitha, J. Appraisal of deep-learning techniques on computer-aided lung cancer diagnosis with computed tomography screening. *J. Med. Phys.* **2020**, *45*, 98.
60. Yuan, H.; Wu, Y.; Dai, M. Multi-Modal Feature Fusion-Based Multi-Branch Classification Network for Pulmonary Nodule Malignancy Suspiciousness Diagnosis. *J. Digit. Imaging* **2022**, *36*, 617–626. [[CrossRef](#)]
61. Wang, Q.; Wu, B.; Zhu, P.; Li, P.; Zuo, W.; Hu, Q. ECA-Net: Efficient channel attention for deep convolutional neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 11534–11542.
62. Mkindu, H.; Wu, L.; Zhao, Y. 3D multi-scale vision transformer for lung nodule detection in chest CT images. *Signal Image Video Process.* **2023**, *17*, 2473–2480. [[CrossRef](#)]
63. Ardila, D.; Kiraly, A.P.; Bharadwaj, S.; Choi, B.; Reicher, J.J.; Peng, L.; Tse, D.; Etemadi, M.; Ye, W.; Corrado, G.; et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat. Med.* **2019**, *25*, 954–961. [[CrossRef](#)] [[PubMed](#)]
64. Tan, H.; Bates, J.H.; Matthew Kinsey, C. Discriminating TB lung nodules from early lung cancers using deep learning. *Bmc Med. Inform. Decis. Mak.* **2022**, *22*, 1–7.
65. National Institute of Allergy and Infectious Diseases. National Institute of Allergy and Infectious Disease (NIAID) TB Portal. 2021. Available online: <https://tbportals.niaid.nih.gov/> (accessed on 27 March 2023).
66. Li, T.Z.; Xu, K.; Gao, R.; Tang, Y.; Lasko, T.A.; Maldonado, F.; Sandler, K.; Landman, B.A. Time-distance vision transformers in lung cancer diagnosis from longitudinal computed tomography. *arXiv* **2022**, arXiv:2209.01676.

67. Li, T. Time Distance Transformer Code. 2023. Available online: <https://github.com/tom1193/time-distance-transformer> (accessed on 18 September 2023).
68. Lu, M.T.; Raghu, V.K.; Mayrhofer, T.; Aerts, H.J.; Hoffmann, U. Deep learning using chest radiographs to identify high-risk smokers for lung cancer screening computed tomography: Development and validation of a prediction model. *Ann. Intern. Med.* **2020**, *173*, 704–713.
69. Khan, A.; Lee, B. Gene transformer: Transformers for the gene expression-based classification of lung cancer subtypes. *arXiv* **2021**, arXiv:2108.11833.
70. Cai, M.; Zhao, L.; Hou, G.; Zhang, Y.; Wu, W.; Jia, L.; Zhao, J.; Wang, L.; Qiang, Y. FDTrans: Frequency Domain Transformer Model for predicting subtypes of lung cancer using multimodal data. *Comput. Biol. Med.* **2023**, *158*, 106812. [[CrossRef](#)]
71. Primakov, S.P.; Ibrahim, A.; van Timmeren, J.E.; Wu, G.; Keek, S.A.; Beuque, M.; Granzier, R.W.; Lavrova, E.; Scrivener, M.; Sanduleanu, S.; et al. Automated detection and segmentation of non-small cell lung cancer computed tomography images. *Nat. Commun.* **2022**, *13*, 3423. [[CrossRef](#)]
72. Padhani, A.; Ollivier, L. The RECIST criteria: Implications for diagnostic radiologists. *Br. J. Radiol.* **2001**, *74*, 983–986.
73. Primakov. DuneAI-Automated-Detection-and-Segmentation-of-non-Small-Cell-Lung-Cancer-Computed-Tomography-Images. 2022. Available online: <https://github.com/primakov/DuneAI-Automated-detection-and-segmentation-of-non-small-cell-lung-cancer-computed-tomography-images> (accessed on 15 March 2023).
74. Ausawalaithong, W.; Thirach, A.; Marukatat, S.; Wilaiprasitporn, T. Automatic lung cancer prediction from chest X-ray images using the deep learning approach. In Proceedings of the 2018 11th Biomedical Engineering International Conference (BMEiCON), Chiang Mai, Thailand, 21–24 November 2018; pp. 1–5.
75. Gordienko, Y.; Gang, P.; Hui, J.; Zeng, W.; Kochura, Y.; Alienin, O.; Rokovyi, O.; Stirenko, S. Deep learning with lung segmentation and bone shadow exclusion techniques for chest X-ray analysis of lung cancer. In *Advances in Computer Science for Engineering and Education 13*; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 638–647.
76. Yu, K.H.; Lee, T.L.M.; Yen, M.H.; Kou, S.; Rosen, B.; Chiang, J.H.; Kohane, I.S. Reproducible machine learning methods for lung cancer detection using computed tomography images: Algorithm development and validation. *J. Med. Internet Res.* **2020**, *22*, e16709. [[CrossRef](#)]
77. Tekade, R.; Rajeswari, K. Lung cancer detection and classification using deep learning. In Proceedings of the 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, 6–18 August 2018; pp. 1–5.
78. Said, Y.; Alsheikhy, A.; Shawly, T.; Lahza, H. Medical Images Segmentation for Lung Cancer Diagnosis Based on Deep Learning Architectures. *Diagnostics* **2023**, *13*, 546. [[CrossRef](#)]
79. Hatamizadeh, A.; Tang, Y.; Nath, V.; Yang, D.; Myronenko, A.; Landman, B.; Roth, H.R.; Xu, D. Unetr: Transformers for 3d medical image segmentation. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 3–8 January 2022; pp. 574–584.
80. Guo, D.; Terzopoulos, D. A transformer-based network for anisotropic 3D medical image segmentation. In Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15 January 2021; pp. 8857–8861.
81. Tang, H.; Liu, X.; Xie, X. An end-to-end framework for integrated pulmonary nodule detection and false positive reduction. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 8–11 April 2019; pp. 859–862.
82. Huang, W.; Xue, Y.; Wu, Y. A CAD system for pulmonary nodule prediction based on deep three-dimensional convolutional neural networks and ensemble learning. *PLoS ONE* **2019**, *14*, e0219369. [[CrossRef](#)]
83. Tang, H.; Kim, D.R.; Xie, X. Automated pulmonary nodule detection using 3D deep convolutional neural networks. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018; pp. 523–526.
84. Fischer, A.M.; Yacoub, B.; Savage, R.H.; Martinez, J.D.; Wichmann, J.L.; Sahbaee, P.; Grbic, S.; Varga-Szemes, A.; Schoepf, U.J. Machine learning/deep neuronal network: Routine application in chest computed tomography and workflow considerations. *J. Thorac. Imaging* **2020**, *35*, S21–S27. [[PubMed](#)]
85. Mohit, B. Named entity recognition. In *Natural Language Processing of Semitic Languages*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 221–245.
86. Zeng, D.; Liu, K.; Lai, S.; Zhou, G.; Zhao, J. Relation classification via convolutional deep neural network. In Proceedings of the COLING 2014, The 25th International Conference on Computational Linguistics: Technical Papers, Dublin, Ireland, 23–29 August 2014; pp. 2335–2344.
87. Zhang, H.; Hu, D.; Duan, H.; Li, S.; Wu, N.; Lu, X. A novel deep learning approach to extract Chinese clinical entities for lung cancer screening and staging. *BMC Med. Inform. Decis. Mak.* **2021**, *21*, 214.
88. Tenney, I.; Das, D.; Pavlick, E. BERT rediscovers the classical NLP pipeline. *arXiv* **2019**, arXiv:1905.05950.
89. Kipkogei, E.; Arango Argoty, G.A.; Kagiampakis, I.; Patra, A.; Jacob, E. Explainable Transformer-Based Neural Network for the Prediction of Survival Outcomes in Non-Small Cell Lung Cancer (NSCLC). *medRxiv* **2021**. [[CrossRef](#)]
90. Doppalapudi, S.; Qiu, R.G.; Badr, Y. Lung cancer survival period prediction and understanding: Deep learning approaches. *Int. J. Med. Inform.* **2021**, *148*, 104371.

91. Barbouchi, K.; El Hamdi, D.; Elouedi, I.; Aicha, T.B.; Echi, A.K.; 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**. [CrossRef]
92. Weikert, T.; Jaeger, P.; Yang, S.; Baumgartner, M.; Breit, H.; Winkel, D.; Sommer, G.; Stieltjes, B.; Thaiss, W.; Bremerich, J.; et al. Automated lung cancer assessment on 18F-PET/CT using Retina U-Net and anatomical region segmentation. *Eur. Radiol.* **2023**, *33*, 4270–4279.
93. Nishio, M.; Sugiyama, O.; Yakami, M.; Ueno, S.; Kubo, T.; Kuroda, T.; Togashi, K. Computer-aided diagnosis of lung nodule classification between benign nodule, primary lung cancer, and metastatic lung cancer at different image size using deep convolutional neural network with transfer learning. *PLoS ONE* **2018**, *13*, e0200721. [CrossRef]
94. Lakshmanaprabu, S.; Mohanty, S.N.; Shankar, K.; Arunkumar, N.; Ramirez, G. Optimal deep learning model for classification of lung cancer on CT images. *Future Gener. Comput. Syst.* **2019**, *92*, 374–382.
95. Shin, H.; Oh, S.; Hong, S.; Kang, M.; Kang, D.; Ji, Y.g.; Choi, B.H.; Kang, K.W.; Jeong, H.; Park, Y.; et al. Early-stage lung cancer diagnosis by deep learning-based spectroscopic analysis of circulating exosomes. *ACS Nano* **2020**, *14*, 5435–5444. [PubMed]
96. Shao, J.; Wang, G.; Yi, L.; Wang, C.; Lan, T.; Xu, X.; Guo, J.; Deng, T.; Liu, D.; Chen, B.; et al. Deep learning empowers lung cancer screening based on mobile low-dose computed tomography in resource-constrained sites. *Front.-Biosci.-Landmark* **2022**, *27*, 212. [CrossRef]
97. Su, X.L.; Wang, J.W.; Che, H.; Wang, C.F.; Jiang, H.; Lei, X.; Zhao, W.; Kuang, H.X.; Wang, Q.H. Clinical application and mechanism of traditional Chinese medicine in treatment of lung cancer. *Chin. Med. J.* **2020**, *133*, 2987–2997.
98. Liu, Z.; He, H.; Yan, S.; Wang, Y.; Yang, T.; Li, G.Z. End-to-end models to imitate traditional Chinese medicine syndrome differentiation in lung cancer diagnosis: Model development and validation. *JMIR Med. Inform.* **2020**, *8*, e17821.
99. Wang, S.; Mahon, R.; Weiss, E.; Jan, N.; Taylor, R.; McDonagh, P.; Quinn, B.; Yuan, L. Automated Lung Cancer Segmentation Using a Dual-Modality Deep Learning Network with PET and CT Images. *Int. J. Radiat. Oncol. Biol. Phys.* **2022**, *114*, e557–e558. [CrossRef]
100. Park, J.; Kang, S.K.; Hwang, D.; Choi, H.; Ha, S.; Seo, J.M.; Eo, J.S.; Lee, J.S. Automatic Lung Cancer Segmentation in [18F] FDG PET/CT Using a Two-Stage Deep Learning Approach. *Nucl. Med. Mol. Imaging* **2022**, *57*, 86–93. [CrossRef] [PubMed]
101. Li, Z.; Zhang, J.; Tan, T.; Teng, X.; Sun, X.; Zhao, H.; Liu, L.; Xiao, Y.; Lee, B.; Li, Y.; et al. Deep learning methods for lung cancer segmentation in whole-slide histopathology images—the acdc @ lunghp challenge 2019. *IEEE J. Biomed. Health Inform.* **2020**, *25*, 429–440. [CrossRef]
102. Li, Z. Automatic Cancer Detection and Classification in Whole-Slide Lung Histopathology Challenge. 2023. Available online: <https://acdc-lunghp.grand-challenge.org/> (accessed on 18 September 2023).
103. Chen, W.; Yang, F.; Zhang, X.; Xu, X.; Qiao, X. MAU-Net: Multiple attention 3D U-Net for lung cancer segmentation on CT images. *Procedia Comput. Sci.* **2021**, *192*, 543–552. [CrossRef]
104. Fu, J.; Liu, J.; Tian, H.; Li, Y.; Bao, Y.; Fang, Z.; Lu, H. Dual attention network for scene segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 3146–3154.
105. Pan, Y.; Shi, D.; Wang, H.; Chen, T.; Cui, D.; Cheng, X.; Lu, Y. Automatic opportunistic osteoporosis screening using low-dose chest computed tomography scans obtained for lung cancer screening. *Eur. Radiol.* **2020**, *30*, 4107–4116. [CrossRef]
106. Shimazaki, A.; Ueda, D.; Choppin, A.; Yamamoto, A.; Honjo, T.; Shimahara, Y.; Miki, Y. Deep learning-based algorithm for lung cancer detection on chest radiographs using the segmentation method. *Sci. Rep.* **2022**, *12*, 727. [CrossRef]
107. Feng, J.; Jiang, J. Deep learning-based chest CT image features in diagnosis of lung cancer. *Comput. Math. Methods Med.* **2022**, *2022*, 4153211. [CrossRef] [PubMed]
108. Gil, J.; Choi, H.; Paeng, J.C.; Cheon, G.J.; Kang, K.W. Deep Learning-Based Feature Extraction from Whole-Body PET/CT Employing Maximum Intensity Projection Images: Preliminary Results of Lung Cancer Data. *Nucl. Med. Mol. Imaging* **2023**, *57*, 216–222. [PubMed]
109. Yan, S.; Huang, Q.; Yu, S.; Liu, Z. Computed tomography images under deep learning algorithm in the diagnosis of perioperative rehabilitation nursing for patients with lung cancer. *Sci. Program.* **2022**, *2022*, 8685604. [CrossRef]
110. Chen, X.; Duan, Q.; Wu, R.; Yang, Z. Segmentation of lung computed tomography images based on SegNet in the diagnosis of lung cancer. *J. Radiat. Res. Appl. Sci.* **2021**, *14*, 396–403. [CrossRef]
111. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef]
112. Chen, L.C.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking atrous convolution for semantic image segmentation. *arXiv* **2017**, arXiv:1706.05587.
113. Choe, J.; Lee, S.M.; Do, K.H.; Lee, G.; Lee, J.G.; Lee, S.M.; Seo, J.B. Deep learning-based image conversion of CT reconstruction kernels improves radiomics reproducibility for pulmonary nodules or masses. *Radiology* **2019**, *292*, 365–373. [CrossRef]
114. Yu, Z.; Yang, X.; Dang, C.; Wu, S.; Adekanattu, P.; Pathak, J.; George, T.J.; Hogan, W.R.; Guo, Y.; Bian, J.; et al. A study of social and behavioral determinants of health in lung cancer patients using transformers-based natural language processing models. In Proceedings of the AMIA Annual Symposium Proceedings, San Diego, CA, USA, 30 October–3 November 2021; p. 1225.

115. Hwang, E.J.; Lee, J.S.; Lee, J.H.; Lim, W.H.; Kim, J.H.; Choi, K.S.; Choi, T.W.; Kim, T.H.; Goo, J.M.; Park, C.M. Deep learning for detection of pulmonary metastasis on chest radiographs. *Radiology* **2021**, *301*, 455–463. [[CrossRef](#)]
116. Schlemper, J.; Oktay, O.; Schaap, M.; Heinrich, M.; Kainz, B.; Glocker, B.; Rueckert, D. Attention gated networks: Learning to leverage salient regions in medical images. *Med. Image Anal.* **2019**, *53*, 197–207. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.