Kerbala Journal for Engineering Science

https://kjes.uokerbala.edu.iq/

A Review of Automatic Nodule Detection Algorithms of Lung Cancer in CT-Scan Images

Zahraa Hussain Abd Al Rahman [a]*, Elham Mohammed Thabit [a]

[a] Department of Information Technology, College of Computer Science and Information Technology, University of Kerbala, Karbala, Iraq.

* Corresponding author, Email: zahraa.abd@s.uokerbala.edu.iq

Received: 29 March 2022; Revised: 29 June 2022; Accepted: 30 June 2022; Published: 30 June 2022

Abstract

Lung cancer has repeatedly surfaced as among the most fatal diseases that humanity has ever known. It is also one of the highest frequent malignancies and one of the leading causes of mortality. Lung cancer cases are quickly expanding. The disorder has a desire to keep asymptomatic in its initial phases, the ability to manifest is exceedingly difficult. As a result, early tumor identification is critical in healing. The sooner a patient is diagnosed, the better his or her prospects of healing and survival. In order to effectively diagnose the condition, technology is crucial. Based on their observations, several researchers have come up with solutions. In the latest years, a number of computer-aided diagnostic (CAD) procedures and systems have been proposed, implemented, and produced in order to use digital transmission to tackle this issue. Such algorithms employ a variety of machine learning and deep learning approaches, along with multiple methodologies based on image processing-based approaches for predicting cancer malignant levels. The purpose of this research is to find, compare, and evaluate a variety of image categorization, semantic segmentation, and other methodologies for categorizing and identifying lung cancer in its early phases.

Keywords: CAD System, Lung cancer, Machine learning.

1. Introduction

When it comes to categories and percentages, lung cancer is at the top of the priority list. Lung cancer is estimated to occur in roughly 2.09 million people worldwide, with 1.76 million deaths accounting for about 84% of all deaths [1]. In lung cancer, tumors are formed by the growth of aberrant cells [2]. Owing to the existence of blood streams and lymph fluid in lung tissue, cancer cells rapidly spread. Cancer cells commonly move to the center of the chest associated with normal lymph circulation [3]. Malignant tumor occurs when cancer cells travel to different organs. Screening methods include computed tomography (CT), positron emission tomography (PET), magnetic resonance imaging (MRI), and X-ray capture images of the lungs for assessment [4]. Because of its capacity to provide a scan without overlapped features, the CT image approach is the most commonly used of the approaches listed. Lung cancer could be diagnosed with high accuracy using CT scans. Image analysis and artificial neural networks can be used to detect lung cancer [5] [6]. CT scanning can be used to create three-dimensional imaging of the chest by capturing images of the lungs in various facets. This three-dimensional image will be used to look for cancers [7]. A CT image is typically used by a physician or other field professional to identify malignancy. It is challenging for a physician or technician to diagnose cancer rapidly and effectively because of the vast amount of CT scans. Figure 1 shows the basic stages in a computer-aided diagnostic Model [8].

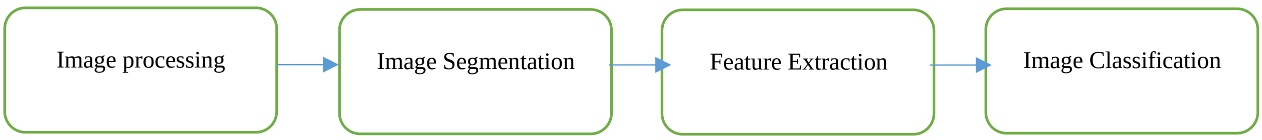

Figure 1 The basic stages in a computer-aided diagnostic Model.

1.1. Imaging Pre-Processing

CT images can not be immediately used by a CAD model. They must be thoroughly pre-processed prior to usage [9]. To improve the image quality and render images that are acceptable for usage, many image pre-processing strategies are used [10] [11]. This improves the overall network efficiency and, as a result, reliability. Table 1 lists a variety of image pre-processing approaches.

Table 1 Variations of image pre-processing methods.

Author	Technique	Usage
Teramoto, A. et al., (2017) [12]	Filtration with Gaussian and Convolutional edge detection	Improves the image's local inconsistencies near the margins of different items
Vas, M. and A. Dessai, (2017) [13]	Median filtering	Getting rid of the salt and pepper noise
Ozdemir, O., R.L et al,. (2019) [14]	Adaptive Gaussian Filtering	Disposal of Gaussian distribution
Asuntha, A. and A. Srinivasan et al., (2020) [15]	Adaptive bilateral filtering	Improvement of clarity and noise reduction

1.2. Imaging Segmentation

The technique of segmenting an image into multiple portions is described as imaging segmentation [16]. It has been used to identify limitations in input images. As the complexity of the image is decreased through segmentation, the process of evaluating the image becomes simpler [17]. Table 2 lists the different segmentation approaches used, as well as the number of sample images examined.

According to the study done by Vas, M. and A. Dessai [13], the author used morphological Operations as a segmentation method with 216 CT scans; 128 images for training and 88 images for testing. In the study done by Alam, J., S and colleagues [18], watershed transform was used as a segmentation method with 500 infected lung CT images for prediction role as well as 500 typical lung CT images dataset, malignancy has been detected then for predicting purposes. The suggested method had a 97% accuracy rate, identifying 126 scans as malignant and 4 scans as non-cancerous. Šarić, M et al. [19] used Region of Interest as a segmentation technique with 33 CT scanning; 25 for training and 8 for testing. On the other hand, Asuntha, A et al. [15] applied K-Means, FCM, and Ant Colony algorithms with nearly 1000 CT images as shown in Table 2.

Table 2 Differences in image segmentation slgorithms.

Author	Segmentation methods	Number of sample images
Vas, M. and A. Dessai, (2017) [13]	Morphological Operations	216 Computed Tomography Scans (128-train & 88- test)
Alam, J., S. et al., (2018) [18]	Watershed transform	500 Computed Tomography Scans
Rahman, M.S. et al., (2019) [20]	Otsu Thresholdings	1000 CT Scans
Šarić, M. et al., (2019) [19]	Region of Interests	33 Computed Tomography Scans (25-train & 8-test)
Asuntha, A. and A. Srinivasan et al., (2020) [15]	K-Means, FCM, and Ant Colony algorithms	1000 Computed Tomography Scans

1.3. Feature Extraction

Feature Extraction is a way of reducing the number of variables in a raw dataset so that it may be processed more quickly and grouped into useful categories [21]. Huge amounts of datasets have a variety of features that necessitate the use of computer systems to conduct and produce results [22]. To simplify data and ensure that no information is lost, feature extraction algorithms are used [23]. These methods are in charge of picking and combining pieces in order to limit the amount of data that needs to be processed [24] [25]. Table 3 provides a variety of feature extraction methods as well as the characteristics which should be assessed when utilizing each strategy.

In 2017, the study performed by Vas, M. and A. Dessai [13], applied the Gray Level Co-occurrence matrix technique for extracting features with Haralick as a selected feature. The same technique was used in the study done by Alam, J., S. et al. [18] with GLCM (Gray Level CoOccurrence technique) by which it could arrange a huge number of pixel brightness levels occurring in a sample. In the studies mentioned in table 3, by using distance and direction, the gray-level co-occurrence matrix (GLCM) examines the co-occurrence of grey levels among linked pairs of cells. The GLCM findings enable the computation of a number of properties, comprising texturing, energy, contrasting, and correlations. According to its surroundings, each pixel in a Local Binary Pattern (LBP) is given a label (a binary number). The spatial pattern of the pixels in an image or the ROI is described by Statistical Moments (SM), which are vector magnitudes. Among the most often used techniques for extracting image features is SM. Times centered on areas are referred to as central moments (CM) [26] [27] [28].

Table 3 Comparison of feature extraction methods.

Author	Technique for Extracting Features	Selected Elements/Features
Vas, M. and A. Dessai, (2017) [13]	Gray Level Co-occurrence Matrix	Haralick features
Alam, J., S. et al., (2018) [18]	Gray Level Co-occurrence Matrix	IDM, kurtosis, Energy, relations, frequencies, homogenization, RMS, SD, Mean, fluctuations, entropy, softness
Rodrigues, M.B. et al., (2018) [26]	Structural Co-Occurrence Matrix (SCM)	Statistical Analysis, Data, Mean
Xie, Y. et al., (2018) [27]	Multi-View Knowledge-Based Collaborative (MVKBC)	Cross-entropy
Chen, W. et al., (2019) [28]	Hybrid features fusion model in three-dimensional and two-dimensional Convolutional Neural Networks	Volumetric and two-dimensional features
Asuntha, A. and A. Srinivasan, (2020) [15]	Region-Of-Interest Extraction	Features such as volumetrics (Zernike moments, Scale-Invariant Feature Transform), texturing (Wavelet & Local binary patterns), intensity (Histogram of Oriented Gradients), and geometrics (Eccentricity & Curvature descriptors)

1.4. Image Classification

Image classification is a vital activity that seeks to comprehend an image as a whole. The objective of offering it to specific labelling is to describe the image [29]. According to major instances, image classification applies on scans in which only one item is viewed and assessed [30]. Cancer identification, on the other hand, requires the achievement of both classification and segmentation assignments. It is employed to explore more realistic scenarios where an image could have numerous items. [31] [32]. Identifying a lung nodule's malignant or non-cancerous status is the aim of this procedure [33]. Table 4 clarifies different classification approaches as well as the outcomes.

Image processing techniques and accuracy vary according to studies implemented in table 4. Studies number [34] and [26] applied decision tree and SVM technique which shows acceptable accuracy for each technique. Serj, M.F. et al. [35] and Kirienko, M. et al. [36] applied Convolutional Neural Network technique with an accuracy score of 95 and 80 respectively. Schwyzer, M. et al. [35] used transfer learning with an accuracy of 97.10. On the other hand, Xie, Y. et al [27] applied Knowledge-based Collaborative Deep Learning with an accuracy value of 91.60.

Table 4 Contrast of image processing methodologies.

Author	Classifying Techniques	Findings
Poreva, A. et al., (2017) [34]	Decision Tree and Support Vector Machine	Accuracy: Decision Time - 72 Support Vector Machine – 75
Kirienko, M. et al., (2018) [36]	Convolutional Neural Network	Accuracy: Validation - 69 Testing - 69 Training - 87 Dice Scores: Training - 82 Testing – 80
Rodrigues, M.B. et al., (2018) [26]	MLP, Support Vector Machine, k-Nearest Neighbors	Accuracy: MLP - 95.40 Support Vector Machine - 96.70 KNN - 95.30
Schwyzer, M. et al., (2018) [37]	Transfer learnings	Accuracy - 97.10 Sensitivity - 95.90 Specificity - 98.10
Serj, M.F. et al., (2018) [35]	Convolutional Neural Network	Sensitivity – 87 Specificity - 99.1 F1 Score – 95
Xie, Y. et al., (2018) [27]	Knowledge-based Collaborative Deep Learning	Accuracy - 91.60
Tran, G.S. et al., (2019) [38]	Two Dimensional Deep Convolutional Network	Accuracy - 97.20 Sensitivity - 96.00 Specificity - 97.30
Nasser, I.M. and S.S. Abu-Naser (2019) [39]	Artificial Neural Network (ANN)	Accuracy - 96.67
Asuntha, A. and A. Srinivasan	Fuzzy Particle Swarm Optimizations	Accuracy - 95.62

(2020) [15]	(FPSO) and Convolutional Neural Network	Sensitivity - 96.23 Specificity - 95.89
Huang, X. et al., (2020) [40]	Deep Transfer Convolutional Neural Network (DTCNN) and Extreme Learning Machine (ELM)	Accuracy - 94.57
Shanthi, S. and N. Rajkumar, (2021) [41]	Stochastic diffusion search algorithm and Neural Networks (SDS-NN)	Accuracy - 89.63

Literature Survey

1. Convolutional Neural Networks (CNN)

In 2019, Moradi and co-workers suggested more methods to distinguish lung cancer nodules from non-nodules. The nodules were split into four categories according to their size. To improve the results, researchers integrated all four classification algorithms. In an attempt to get superior results, all four algorithms were integrated. Each CNN was created using a mixture of Max pooling and convolutional layers. ReLU is the perceptron used here. Ultimately, a Softmax layer with a completely associated layer is employed to generate the result [42]. Because nodule sizes range from 3mm to 3cm, utilizing only one layer may result in incorrect predictions for either extremely small nodules or quite big results. So researchers combined the output values (expected results) of all four CNNs and submitted them to the last classification model [43]. They selected a logistic regression predictor that accepts four CNN parameters and generates prediction accuracy. As an output, they discovered that the merged classifier's outcome is superior to each of the individual classifiers [44].

In 2019, Pouria Moradi and colleagues suggested the usage of three-dimensional CNN to decrease false positives. Xavier weight initialization is used to initiate the network parameters. Optimization algorithm may be used to develop weight values with a detection rate of 0.01 and a mutation rate of 5-10 per session with 0.9 linear momentum. A Meta classification was created by combining three decision branches that had been tested. The LUNA 16 data was used to develop and assess the algorithm. For 3.09 false positives, the algorithm has a 91.23% accuracy rate [44].

Applying deeper CNN-based approaches, Mehdi Fatan Serj et al. reported a process to determine lung tumors effectively in 2018. The researchers created a network of two max-pooling levels, three convolutions, softmax (binary) layers, and a dense layer. The approach has been evaluated using Kaggle's database for the Kaggle Machine Learning Bow l 2017 competition [45]. A deeper CNN-based strategy performed better than that of other CNN-based algorithms. To eliminate the multivariable logistic regression goal and therefore the frequency of patients with lung cancer, cross-entropy had been employed for the output layer. The researchers were able to attain a sensitivity of 87% and a specificity of 99.1% [35].

In 2019, Ruchita Tekade et al. devised an approach that uses two configurations: one for nodular separation and the other for determining the malignant degree. CNN is utilized for categorization along with features extracting to determine the malignant degree; max pooling is utilized for sub pooling; ReLU has been used as an input signal, and softmax is the classifier used to do the categorization and allocate malignant standard. In convolutional kernels, the Adam classification is used to optimize weighting distribution [46] [47]. Simple median filter, obvious boundary, morphological eroding, morphological closure, and morphological raising are used in the pre-processing of CT scanned images for classification. For lungs Computed tomography pictures, masses are created utilizing U-Net classification, and lung nodules are differentiated. The datasets used in this investigation were LIDC-IDRI, LUNA16, and Data Science Bowl2017. This technique had a 95.66 % accuracy, a loss of 0.09, a dice coefficient of 90%, and a log loss prediction accuracy of 38% when utilizing U-Net to segment and thus further anticipate malignant growth stages [48].

In 2020, Patra et al. proposed a deep learning-based strategy for recognizing and categorizing malignant cells. CT images from LIDC and private databases were considered as inputs, and the contrasting intensity was increased using Histogram Equalization. The CT images were denoised using the adaptable Bilateral filtering approach. To obtain the ROI, the image was segmented using the artificial Bee Colony segmentation technique. Applying Local Binary Pattern and several wavelet approaches, 180 characteristics were recovered (20 Zernike, 1 Curvature, 18 SIFT, 1 Eccentricity, 26 wavelets, and 18 HOG). The most essential parameter was selected using Fuzzy Particle Swarm Optimization, which was then utilized to minimize the difficulty of the CNN model, that it was being used to categorize the retrieved nodules as malignancy. The proposed subject's accuracy rate, specificity, and sensitivity are 95.62%, 95.89%, and 96.23%, respectively [21].

Mesut Togacar et al. presented a CNN-based lung cancer detection system in 2020. They received a total of 100 pictures from 69 different treatments (50 malignant and 50 non-cancerous). Because there were fewer photos, enhancement was utilized to produce a healthier database. The CNNs AlexNet, LeNet, and VGG-16 were used in the research. To adjust the parameters for each training dataset, Stochastic Gradient Descent was utilized as optimization approach (for AlexNet and VGG-16). Aside from that, the refinement techniques RMSProp and ADAM were also applied (for LeNet). The characteristics were extracted using the mRMR technique. Classic machine learning methods including LR, LDA, SVM, KNN, and DT are utilized. Using the Principal

Component Analysis approach, the efficiency was enhanced. By combining KNN with CNN and mRMR, 99.51% accuracy was achieved [49].

Lastly, the average score of all the images from the testing set was subtracted to zero-center the images. Rather than immediately feeding the separated images into the classification, a U-Net was developed using the LUNA16 database and then utilized to locate the exact placement of the nodule for image preprocessing. Further categories such as linear classifiers, three-dimensional CNN, and three-dimensional Googlenet templates have been utilized to reduce false-positive outcomes. Including an efficiency of 75.1%, a sensitivity of 77%, a specificity of 74.1%, and an AUC of 75.7%, three-dimensional Googlenet outperformed the other two. The primary conclusion was that the algorithm was built on a fewer group of labeled data, allowing it to be applied to all types of tumors [50].

2. Others

In 2018, Aicha Majda and colleagues introduced and compared four distinct feature extraction techniques: CNN, PCA, Restricted Boltzmann Machines (RBM), and two-dimensional DFT. Three hidden levels of CNN architecture were utilized to further evaluate which strategy offered the highest results. These neurological networks were trained using the LIDC-IDRI database. Using a description document, tumor segmentation locations are recovered in chunks from the CT image, associated with data supplementation to increase the volume of the given dataset. In this trial, CNN outperformed the other approaches in terms of accuracy [51]. Apart from the fact that the findings of CNN two-dimensional DFT were very similar in terms of accuracy, it had a lot of volatility and bias [52]. Owing to relationships being ignored between segmented images and outcomes maintained as objective, this bias and variability in two-dimensional DFT finally increased the number of imbalanced datasets [53].

P. Mohamed Shakeel et al. suggested two approaches for lung cancer diagnosis using CT scans in 2019. The Carcinoma Incidence Archive dataset was included in this investigation. In this research, Machine Learning educated neural networks and Enhanced Excessive grouping are applied. CT images have reduced images and clutter, thus pre-processing is used to eliminate all of this. Image histogram methods can increase image quality since they are a highly effective strategy for various photos. With the use of an enhanced CT image and IPCT, cancer-affected areas are segmented. From the upgraded lung CT image, the enhanced excessive clustering algorithm is used to separate cancer-affected areas. Two methods of enhanced filtering methods operate to identify inconsistencies in image pixels by checking the image pixel and grouping related superpixels together [54]. The pixel frequency is used to forecast information similarities during the segmentation method, once the pixels are continuously evaluated. Unique wavelength characteristics such as standard error, 3rd-moment skewness, average, and 4th-moment kurtosis are obtained from the detected image and transmitted to the segmentation stage because they are particularly helpful in detecting lung cancer with related features. The algorithm ensures 98.42% accuracy, with a minimal classifying algorithm of 0.038 [55].

S. Shanthi et al. suggested a method to identify lung cancer in 2020 that included a random displacement algorithm-based and classification techniques including Machine learning, Decisions orchards, and Naive Bayes. A total of 270 images (140 normal and 130 aberrant) were obtained and used from the TCGA collection. To identify texture characteristics, the grey level co-occurrence matrix (GLCM) was used. For shape-based characteristics, the Gabor filter was utilized. For image segmentation, the SDS algorithm was utilized. Initialization stage (assigning units to a certain randomized assumptions), Assessment stage (evaluating the optimal solution to determine the optimum), Testing stage: (if Activity Member presented, the owner's significantly affected randomized owner's as an optimal solution, otherwise, Non active Agent is preferrable), and Spreading stage [56]. Following the application of SDS, many classification techniques were used. Following comparing the levels of accuracy of all the different classifiers, it was discovered that the Learning Algorithm combined with the SDS method (SDS-NN) performed better than the others. An evaluation was done suggesting that better data preprocessing enhances image categorization [41].

3. Support Vector Machine (SVM)

Nidhi S. Nadakarni and colleagues suggested an image retrieval approach based on image features extraction and subsequently, classification in 2019. Various filtering has been used in image pre-processing to eliminate undesirable interference and sustain the photo. Shape-based FETs (Length, Boundary, Median, Average, and Variation) and vividness FETs (Comparison, Homogeneity, and Uniformity) are employed in the image retrieval section. The linearization sequence is then utilized to identify textures [57]. LBP surpasses other literary structures in terms of reliability. After that, and to categorize the input, a Support vector machine (SVM) was used. A hyperplane is selected so that the margins are maximized [58].

Nidhi S. Nadakarni et al. presented an adaptive method for an early diagnosis of lung cancer in 2019. CT scans in DICOM format have been taken out from Cancer Image Archive Collection.

To eliminate complexity and increase pixel density, these scans were prepared utilizing different image improvement methods including Median Filtration, Straightening, and Color Correction. After converting the monochrome scale image into a binary for background subtraction, structural breaking procedures were done. Area, radius, and irregularity (curviness) are all viewed in the extracting features approaches [41]. The SVM tested predictor is used to sort scans into regular and irregular imaging and employing these qualities in a useful way. The researchers claim that the suggested technique effectively diagnoses cancer in its earlier phases [59].

4. Artificial Neural Networks (ANN)

In 2019, Ibrahim M. Nasser and his co-workers suggested employing artificial neural networks to indicate the existence or lack of lung cancer. Indicators were utilized to diagnose the condition. After 1418105, the ANN method showed the existence of lung cancer with 96.67% authenticity and less than 1% learning failure velocity. Indeed it has been discovered that "Age" has the greatest influence on the outcomes [39].

Moritz Schwyzer et al. developed an adaptive Convolutional Systems technique for lung cancer detection utilizing ultralow dosage PET/CT in 2018. The data used in this study consists of 100 patient registrations, 50 of which are cancer sufferers and the other 50 are not. The statistical categorization was done on sections where the individuals' lung tumors were visible and slicing where the patients did not have any lung cancer [60]. By categorizing lung cancer, the leftover neural net was used for testing purposes. The results were 97.1% accuracy, 95.9% sensitivity, and 98.1% specificity [37].

Table 5 Comparing of various techniques and their outcomes.

Author	Database Utilized	Method	Findings
Rebouças Filho, P.P. et al., (2017) [61]	40 chest Computerized Tomography images	Three-dimensional Adaptive Crisp Active Contour Methodology (Three-dimensional ACACM)	F-measure - 99.22 ± 0.14
Shen, W. et al., (2017) [62]	Lung Image Dataset Consortium scans gatherings	Multi-crop Convolutional Neural Network (MC-Convolutional Neural Network)	Accuracy - 87.14
Song, Q. et al., (2017) [63]	Lung Image Dataset Consortium scans gatherings	Convolutional Neural Network, Deep neural networks, SAE	Accuracy-84.15%
Chapaliuk, B. and Y. Zaychenko (2018) [64]	Data Science Bowl 2017	CT three dimensional, Three-dimensional Dense Convolutional Network	-
Choi, H. and K.J. Na	NCBI GEO (Gene Expression	Weighted Gene	< C-index-0.709±0.042>

Author	Database Utilized	Method	Findings
(2018) [65]	Dataset) [11] microarray database	Coexpression Network Analysis, Cox regression, Convolutional Neural Network, Kaplan Meier Method,	
Jiang, J. et al., (2018) [66]	The Cancer Imaging Archive, Memorial Sloan Kettering Cancer Center, Lung Image Database Consortium image collection	U-Net SegNet FRRN Increment Michigan Recruitment & Retention Network Desne Michigan Recruitment & Retention Network	Sensitivity: 0.80 0.77 0.76 0.85 0.82
Khosravan, N. and U. Bagci (2018) [67]	LUng Nodule Analysis16	Morphological operations, ADAM optimizer Semi-Supervised Multi-Task Learning	DSC - 91 Sensitivity - 98
Li, X.-X. et al., (2018) [68]	Lung Image Database Consortium image collection, General Hospital Of Guangzhou Military Command	Anisotropic nonlinear diffusion filter, Random Walker (RW), Random Forest (RF), GLCM, LBP, Gabor Filter	Sensitivity- 0.92 Specificity-0.83 Accuracy-0.90 AUC-0.95
Makaju, S. et al., (2018) [69]	Lung Image Database Consortium image collection	Median Filter, Gaussian Filter Watershed Segmentation Support Vector Machine	Accuracy-92%, Specificity-50%, Sensitivity-100%
Nishio, M. et al., (2018) [70]	The Cancer Imaging Archive (TCIA)	Support Vector Machine or XGBoost	AUC-0.850 Accuracy-0.797
Rodrigues, M.B. et al., (2018) [26]	Lung Image Database Consortium image collection	Laplace, Gaussian & Sobel filtering, multilayer perceptron, Support Vector Machine, KNN, SCM Mean HU	Accuracy (SCM Mean HU) - 96.70
Skourt, B.A. et al, (2018) [53]	Lung Image Database Consortium image collection	U-Net	Dice Coefficient-0.9502
Xie, Y. et al., (2018) [27]	Lung Image Database Consortium image collection	Knowledge-based Collaborative Deep Learning, U-Net, Three dimensional-GLCM-Support Vector Machine	Accuracy-91.60% Sensitivity86.52% Specificity-94% AUC-95.70%
Zhu, W. et al. (2018) [71]	LUng Nodule Analysis16	Three-dimensional DPN, 10-fold cross-validation, Three dimensional Faster R- Convolutional Neural Network	Accuracy-81.42%
Baek, S. et al., (2018) [72]	96 PET/CT scans of NSCLC patients	U-Net	-
Barros, A.C. et al., (2019) [73]	Walter Cantidio Hospital	Spatial Interdependence Matrix, Visual Information Fidelity Optimum-path forest (OPF) classification	Accuracy - 98.2% F-score - 95.2%
Chen, W. et al., (2019) [28]	134 CT images from Shandong Cancer Hospital	Three-dimensional and two-dimensional Convolutional Neural Network, Hybrid	Dice score - 88.8% Sensitivity - 87.2% Precision - 90.9%

Author	Database Utilized	Method	Findings
		features fusion module (HFFM)	
Dabeer, S., M.M. et al, (2019) [50]	LUng Nodule Analysis16, Kaggle Data Science Bowl 2017	Three-dimensional Convolutional Neural Network,	Sensitivity-87% Specificity- 99.1%
Xie, H., et al. (2019) [74]	LUng Nodule Analysis 16 (Testing),	Two-dimensional CNN R- Convolutional Neural Network (Detecting Of Nodules)	AUC-0.954
Liu, Z., et al. (2019) [75]	Private Dataset	DQN, H-DQN, Convolutional Neural Network	
Pehrson, L.M. M.B., et al, (2019) [76]	Lung Image Database Consortium image collection	Feature-Based Framework, Support Vector Machine, GLMR, Elaboration likelihood model, A probabilistic neural network, Artificial Neural networks, Deep Belief Network, D Architecture	Accuracy-90%
Rahman, M.S., P.C. et al, (2019) [20]	The Cancer Imagings Archive (TCIA)	Gaussian Blur, Otsu Threshold, Mobile Net, Inception-V3, VGG-8	Best Achieved among three Neural Networks Accuracy-97%, Specificity97.85%, Sensitivity-96.26%
Shakeel, P.M., M.A., (2019) [55]	Deep-learning instantaneously trained neural networks Improved profuse clustering (IPCT)	Cancer Imaging Archive (TCIA)	Accuracy-98.42%
Singh, G.A.P. and P. Gupta, (2019) [2]	Private	Gerbil Lung Cell Conditioned Medium, Support Vector Machine, k-Nearest Neighbors, Decision Tree, MLP, Stochastic gradient descent, Stochastic Gradient, RF classifier, Bayes Classifier	-
Tran, G.S. et al., (2019) [38]	Lung Image Database Consortium image collection	Two-dimensional Deep Convolutional Network	Accuracy-97.2% Sensitivity-96.0% Specificity-97.3%
Wei, H. et al., (2019) [77]	SCLC patients Shandong Cancer Hospital (database of 134 case)[	Neighborhood gray-tone difference matrices, Spatial gray-level dependence matrices, Gray Level Histogram Analysis	AUC-0.797
Huang, X. et al., (2019) [78]	Lung Image Database Consortium image collection	Faster R- Convolutional Neural Network	Accuracy - 91.4%
Lu, M. et al., (2020) [79]	Database of cases from Third Affiliated Hospital of Soochow University.	Min-Redundancy Max-Relevance (mRMR), Risk Ovarian Malignancy Algorithm, Logistic Regression, Decision Tree	-
Lu, M. et al., (2020)	Cancer Imaging Archive (TCIA)	1D Convolutional	Accuracy-96 ± 3%

Author	Database Utilized	Method	Findings
[80]		Neural Network	
Huang, X. et al., (2020) [40]	LIDC-IDRI, First Affiliated Hospital of Guangzhou Medical University in China (FAMGMU) Amount of comments =115	Extreme Learning Machine (ELM) and Deep Transfer Convolutional Neural Network (DT Convolutional Neural Network)	Accuracy-94.57%
Yu, K.-H. et al., (2020) [51]	Kaggle Science Bowl dataset	Lung mask, lung segmentation two dimensional & Three dimensional A residual neural network, U-Net, Visual Geometry Group-Net Convolutional Neural Network, tree-based classifiers	-

Conclusion

Lung cancer is among the most deadly diseases that have ever occurred. Regrettably, if this condition has progressed to a significant amount or reached a dangerous phase, it is harder to treat. Computer-Aided Detection (CAD) is considered a vital technique of the rapidly evolving methods which assist in the diagnosis and screening by bringing in patient-related data including CT scans, X-rays, MRI scans, odd complaints in sufferers, indicators, and so on. SVM, CNN, ANN, Watershed Classification, Augmentation of images, and Image analysis are just a few of the approaches utilized to increase performance and speed up the process. The most often used resources for learning are LUNA16, Super Bowl Dataset 2016, and LIDC-IDRI.

References:

[1] Linning E, Lu L, Li L, Yang H, Schwartz LH, Zhao B. Radiomics for classification of lung cancer histological subtypes based on nonenhanced computed tomography. Academic radiology. 2019;26(9):1245-52.

[2] Singh GAP, Gupta P. Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans. Neural Computing and Applications. 2019;31(10):6863-77.

[3] Lakshmanaprabu S, Mohanty SN, Shankar K, Arunkumar N, Ramirez G. Optimal deep learning model for classification of lung cancer on CT images. Future Generation Computer Systems. 2019;92:374-82.

[4] Sujitha R, Seenivasagam V. Classification of lung cancer stages with machine learning over big data healthcare framework. Journal of Ambient Intelligence and Humanized Computing. 2021;12(5):5639-49.

[5] Chaunzwa TL, Hosny A, Xu Y, Shafer A, Diao N, Lanuti M. Deep learning classification of lung cancer histology using CT images. Scientific reports. 2021;11(1):1-12.

[6] Radhika P, Nair RA, Veena G, editors. A comparative study of lung cancer detection using machine learning algorithms. 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT); 2019: IEEE.

[7] Jenipher VN, Radhika S, editors. A study on early prediction of lung cancer using machine learning techniques. 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS); 2020: IEEE.

[8] Vijh S, Gaur D, Kumar S. An intelligent lung tumor diagnosis system using whale optimization algorithm and support vector machine. International Journal of System Assurance Engineering and Management. 2020;11(2):374-84.

[9] ALzubi JA, Bharathikannan B, Tanwar S, Manikandan R, Khanna A, Thaventhiran C. Boosted neural network ensemble classification for lung cancer disease diagnosis. Applied Soft Computing. 2019;80:579-91.

[10] Bicakci M, Ayyildiz O, Aydin Z, Basturk A, Karacavus S, Yilmaz B. Metabolic imaging based sub-classification of lung cancer. IEEE Access. 2020;8:218470-6.

[11] Luna JM, Chao H-H, Diffenderfer ES, Valdes G, Chinniah C, Ma G. Predicting radiation pneumonitis in locally advanced stage II–III non-small cell lung cancer using machine learning. Radiotherapy and Oncology. 2019;133:106-12.

[12] Teramoto A, Tsukamoto T, Kiriyama Y, Fujita H. Automated classification of lung cancer types from cytological images using deep convolutional neural networks. BioMed research international. 2017;2017.

[13] Vas M, Dessai A, editors. Lung cancer detection system using lung CT image processing. 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA); 2017: IEEE.

[14] Ozdemir O, Russell RL, Berlin AA. A 3D probabilistic deep learning system for detection and diagnosis of lung cancer using low-dose CT scans. IEEE transactions on medical imaging. 2019;39(5):1419-29.

[15] Asuntha A, Srinivasan A. Deep learning for lung Cancer detection and classification. Multimedia Tools and Applications. 2020;79(11):7731-62.

[16] Kadir T, Gleeson F. Lung cancer prediction using machine learning and advanced imaging techniques. Translational lung cancer research. 2018;7(3):304.

[17] Pradhan K, Chawla P. Medical Internet of things using machine learning algorithms for lung cancer detection. Journal of Management Analytics. 2020;7(4):591-623.

[18] Alam J, Alam S, Hossan A, editors. Multi-stage lung cancer detection and prediction using multi-class svm classifie. 2018 International conference on computer, communication, chemical, material and electronic engineering (IC4ME2); 2018: IEEE.

[19] Šarić M, Russo M, Stella M, Sikora M, editors. CNN-based method for lung cancer detection in whole slide histopathology images. 2019 4th International Conference on Smart and Sustainable Technologies (SpliTech); 2019: IEEE.

[20] Rahman MS, Shill PC, Homayra Z, editors. A new method for lung nodule detection using deep neural networks for CT images. 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE); 2019: IEEE.

[21] Patra R, editor Prediction of lung cancer using machine learning classifier. International Conference on Computing Science, Communication and Security; 2020: Springer.

[22] Joshua ESN, Chakkravarthy M, Bhattacharyya D. An Extensive Review on Lung Cancer Detection Using Machine Learning Techniques: A Systematic Study. Rev d'Intelligence Artif. 2020;34(3):351-9.

[23] Bębas E, Borowska M, Derlatka M, Oczeretko E, Hładuński M, Szumowski P. Machine-learning-based classification of the histological subtype of non-small-cell lung cancer using MRI texture analysis. Biomedical Signal Processing and Control. 2021;66:102446.

[24] Yuan F, Lu L, Zou Q. Analysis of gene expression profiles of lung cancer subtypes with machine learning algorithms. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease. 2020;1866(8):165822.

[25] Yu L, Tao G, Zhu L, Wang G, Li Z, Ye J. Prediction of pathologic stage in non-small cell lung cancer using machine learning algorithm based on CT image feature analysis. BMC cancer. 2019;19(1):1-12.

[26] Rodrigues MB, Da Nobrega RVM, Alves SSA, Rebouças Filho PP, Duarte JBF, Sangaiah AK, et al. Health of things algorithms for malignancy level classification of lung nodules. IEEE Access. 2018;6:18592-601.

[27] Xie Y, Xia Y, Zhang J, Song Y, Feng D, Fulham M. Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT. IEEE transactions on medical imaging. 2018;38(4):991-1004.

[28] Chen W, Wei H, Peng S, Sun J, Qiao X, Liu B. HSN: hybrid segmentation network for small cell lung cancer segmentation. IEEE Access. 2019;7:75591-603.

[29] Raoof SS, Jabbar MA, Fathima SA, editors. Lung Cancer Prediction using Machine Learning: A Comprehensive Approach. 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA); 2020: IEEE.

[30] Baskar S, Shakeel PM, Sridhar K, Kanimozhi R, editors. Classification system for lung cancer nodule using machine learning technique and CT images. 2019 International Conference on Communication and Electronics Systems (ICCES); 2019: IEEE.

[31] Li J, Wang Y, Song X, Xiao H. Adaptive multinomial regression with overlapping groups for multi-class classification of lung cancer. Computers in biology and medicine. 2018;100:1-9.

[32] Abdullah DM, Abdulazeez AM, Sallow AB. Lung cancer prediction and classification based on correlation selection method using machine learning techniques. Qubahan Academic Journal. 2021;1(2):141-9.

[33] Hyun SH, Ahn MS, Koh YW, Lee SJ. A machine-learning approach using PET-based radiomics to predict the histological subtypes of lung cancer. Clinical nuclear medicine. 2019;44(12):956-60.

[34] Poreva A, Karplyuk Y, Vaityshyn V, editors. Machine learning techniques application for lung diseases diagnosis. 2017 5th IEEE Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE); 2017: IEEE.

[35] Serj MF, Lavi B, Hoff G, Valls DP. A deep convolutional neural network for lung cancer diagnostic. arXiv preprint arXiv:180408170. 2018.

[36] Kirienko M, Sollini M, Silvestri G, Mognetti S, Voulaz E, Antunovic L. Convolutional neural networks promising in lung cancer T-parameter assessment on baseline FDG-PET/CT. Contrast Media & Molecular Imaging. 2018;2018.

[37] Schwyzer M, Ferraro DA, Muehlematter UJ, Curioni-Fontecedro A, Huellner MW, Von Schulthess GK. Automated detection of lung cancer at ultralow dose PET/CT by deep neural networks–initial results. Lung Cancer. 2018;126:170-3.

[38] Tran GS, Nghiem TP, Nguyen VT, Luong CM, Burie J-C. Improving accuracy of lung nodule classification using deep learning with focal loss. Journal of healthcare engineering. 2019;2019.

[39] Nasser IM, Abu-Naser SS. Lung cancer detection using artificial neural network. International Journal of Engineering and Information Systems (IJEAIS). 2019;3(3):17-23.

[40] Huang X, Lei Q, Xie T, Zhang Y, Hu Z, Zhou Q. Deep transfer convolutional neural network and extreme learning machine for lung nodule diagnosis on CT images. Knowledge-Based Systems. 2020;204:106230.

[41] Shanthi S, Rajkumar N. Lung cancer prediction using stochastic diffusion search (SDS) based feature selection and machine learning methods. Neural Processing Letters. 2021;53(4):2617-30.

[42] Cherezov D, Goldgof D, Hall L, Gillies R, Schabath M, Müller H. Revealing tumor habitats from texture heterogeneity analysis for classification of lung cancer malignancy and aggressiveness. Scientific reports. 2019;9(1):1-9.

[43] Hirsch FR, Scagliotti GV, Mulshine JL, Kwon R, Curran Jr WJ, Wu Y-L. Lung cancer: current therapies and new targeted treatments. The Lancet. 2017;389(10066):299-311.

[44] Moradi P, Jamzad M, editors. Detecting lung cancer lesions in CT images using 3D convolutional neural networks. 2019 4th International Conference on Pattern Recognition and Image Analysis (IPRIA); 2019: IEEE.

[45] Rossetto AM, Zhou W, editors. Deep learning for categorization of lung cancer ct images. 2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE); 2017: IEEE.

[46] Rahane W, Dalvi H, Magar Y, Kalane A, Jondhale S, editors. Lung cancer detection using image processing and machine learning healthcare. 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT); 2018: IEEE.

[47] Khawaja A, Bartholmai BJ, Rajagopalan S, Karwoski RA, Varghese C, Maldonado F. Do we need to see to believe?—Radiomics for lung nodule classification and lung cancer risk stratification. Journal of Thoracic Disease. 2020;12(6):3303.

[48] Tekade R, Rajeswari K, editors. Lung cancer detection and classification using deep learning. 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA); 2018: IEEE.

[49] Toğaçar M, Ergen B, Cömert Z. Detection of lung cancer on chest CT images using minimum redundancy maximum relevance feature selection method with convolutional neural networks. Biocybernetics and Biomedical Engineering. 2020;40(1):23-39.

[50] Dabeer S, Khan MM, Islam S. Cancer diagnosis in histopathological image: CNN based approach. Informatics in Medicine Unlocked. 2019;16:100231.

[51] Yu K-H, Lee T-LM, Yen M-H, Kou S, Rosen B, Chiang J-H. Reproducible machine learning methods for lung cancer detection using computed tomography images: Algorithm development and validation. Journal of medical Internet research. 2020;22(8):e16709.

[52] Bartholomai JA, Frieboes HB, editors. Lung cancer survival prediction via machine learning regression, classification, and statistical techniques. 2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT); 2018: IEEE.

[53] Skourt BA, El Hassani A, Majda A. Lung CT image segmentation using deep neural networks. Procedia Computer Science. 2018;127:109-13.

[54] Wang C, Long Y, Li W, Dai W, Xie S, Liu Y. Exploratory study on classification of lung cancer subtypes through a combined K-nearest neighbor classifier in breathomics. Scientific reports. 2020;10(1):1-12.

[55] Shakeel PM, Burhanuddin MA, Desa MI. Lung cancer detection from CT image using improved profuse clustering and deep learning instantaneously trained neural networks. Measurement. 2019;145:702-12.

[56] Wei H, Xie L, Liu Q, Shao C, Wang X, Su X. Automatic classification of label-free cells from small cell lung cancer and poorly differentiated lung adenocarcinoma with 2D light scattering static cytometry and machine learning. Cytometry Part A. 2019;95(3):302-8.

[57] Banerjee N, Das S, editors. Prediction Lung Cancer–In Machine Learning Perspective. 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA); 2020: IEEE.

[58] Jena SR, George T, Ponraj N, editors. Texture analysis based feature extraction and classification of lung cancer. 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT); 2019: IEEE.

[59] Nadkarni NS, Borkar S, editors. Detection of lung cancer in CT images using image processing. 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI); 2019: IEEE.

[60] Günaydin Ö, Günay M, Şengel Ö, editors. Comparison of lung cancer detection algorithms. 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT); 2019: IEEE.

[61] Rebouças Filho PP, Cortez PC, da Silva Barros AC, Albuquerque VHC, Tavares JMR. Novel and powerful 3D adaptive crisp active contour method applied in the segmentation of CT lung images. Medical image analysis. 2017;35:503-16.

[62] Shen W, Zhou M, Yang F, Yu D, Dong D, Yang C. Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification. Pattern Recognition. 2017;61:663-73.

[63] Song Q, Zhao L, Luo X, Dou X. Using deep learning for classification of lung nodules on computed tomography images. Journal of healthcare engineering. 2017;2017.

[64] Chapaliuk B, Zaychenko Y, editors. Deep learning approach in computer-aided detection system for lung cancer. 2018 IEEE First International Conference on System Analysis & Intelligent Computing (SAIC); 2018: IEEE.

[65] Choi H, Na KJ. A risk stratification model for lung cancer based on gene Coexpression network and deep learning. BioMed research international. 2018;2018.

[66] Jiang J, Hu Y-C, Liu C-J, Halpenny D, Hellmann MD, Deasy JO. Multiple resolution residually connected feature streams for automatic lung tumor segmentation from CT images. IEEE transactions on medical imaging. 2018;38(1):134-44.

[67] Khosravan N, Bagci U, editors. Semi-supervised multi-task learning for lung cancer diagnosis. 2018 40th Annual international conference of the IEEE engineering in medicine and biology society (EMBC); 2018: IEEE.

[68] Li X-X, Li B, Tian L-F, Zhang L. Automatic benign and malignant classification of pulmonary nodules in thoracic computed tomography based on RF algorithm. IET Image Processing. 2018;12(7):1253-64.

[69] Makaju S, Prasad P, Alsadoon A, Singh A, Elchouemi A. Lung cancer detection using CT scan images. Procedia Computer Science. 2018;125:107-14.

[70] Nishio M, Nishizawa M, Sugiyama O, Kojima R, Yakami M, Kuroda T. Computer-aided diagnosis of lung nodule using gradient tree boosting and Bayesian optimization. PloS one. 2018;13(4):e0195875.

[71] Zhu W, Liu C, Fan W, Xie X, editors. Deeplung: Deep 3d dual path nets for automated pulmonary nodule detection and classification. 2018 IEEE winter conference on applications of computer vision (WACV); 2018: IEEE.

[72] Baek S, He Y, Allen BG, Buatti JM, Smith BJ, Tong L. Deep segmentation networks predict survival of non-small cell lung cancer. Scientific reports. 2019;9(1):1-10.

[73] Barros AC, Ramalho GL, Pereira CR, Papa JP, de Albuquerque VHC, Tavares JMR. Automated recognition of lung diseases in CT images based on the optimum-path forest classifier. Neural Computing and Applications. 2019;31(2):901-14.

[74] Xie H, Yang D, Sun N, Chen Z, Zhang Y. Automated pulmonary nodule detection in CT images using deep convolutional neural networks. Pattern Recognition. 2019;85:109-19.

[75] Liu Z, Yao C, Yu H, Wu T. Deep reinforcement learning with its application for lung cancer detection in medical Internet of Things. Future Generation Computer Systems. 2019;97:1-9.

[76] Pehrson LM, Nielsen MB, Ammitzbøl Lauridsen C. Automatic pulmonary nodule detection applying deep learning or machine learning algorithms to the LIDC-IDRI database: a systematic review. Diagnostics. 2019;9(1):29.

[77] Wei H, Yang F, Liu Z, Sun S, Xu F, Liu P. Application of computed tomography-based radiomics signature analysis in the prediction of the response of small cell lung cancer patients to first-line chemotherapy. Experimental and Therapeutic Medicine. 2019;17(5):3621-9.

[78] Huang X, Sun W, Tseng T-LB, Li C, Qian W. Fast and fully-automated detection and segmentation of pulmonary nodules in thoracic CT scans using deep convolutional neural networks. Computerized Medical Imaging and Graphics. 2019;74:25-36.

[79] Lu M, Fan Z, Xu B, Chen L, Zheng X, Li J. Using machine learning to predict ovarian cancer. International Journal of Medical Informatics. 2020;141:104195.

[80] Moitra D, Mandal RK. Classification of non-small cell lung cancer using one-dimensional convolutional neural network. Expert Systems with Applications. 2020;159:113564.

مراجعة البحوث في مجال اكتشاف سرطان الرئة بالصور السينية

الخلاصة: *يعتبر سرطان الرئة من أكثر الأمراض فتكًا على حياة البشرية، حيث يعتبر من أكثر الأورام الخبيثة انتشارًا وأحد الأسباب الرئيسية للوفاة. تزداد حالات سرطان الرئة بسرعة، ولكن في المراحل الاولية للمرض يكون بدون أعراض، حيث يكون اكتشافه صعب للغاية. لذلك يعد التعرف المبكر على اورام الرئة أمرًا بالغ الأهمية في انقاذ حياة المريض. كلما تم تشخيص المرض مبكرًا ، كانت فرصة الشفاء أفضل. ومع تقدم التكنولوجيا في جميع مجالات الحياة بما في ذلك المجالات الطبية، تم اقتراح وتنفيذ طرق متقدمة لمعالجة سرطان الرئة وذلك من خلال استخدام إجراءات وأنظمة لتشخيص المرض بمساعدة الكمبيوتر حيث تم استخدام عدد من الخوارزميات المتنوعة بالاعتماد على أساليب التعلم الآلي والتعلم العميق ، إلى جانب استخدام منهجيات متعددة تعتمد على الأساليب القائمة على معالجة الصور للتنبؤ بمستويات السرطان الخبيثة. الغرض من هذا البحث هو إيجاد ومقارنة وتقييم مجموعة متنوعة من تصنيفات الصور والتجزئة الدلالية والمنهجيات الأخرى لتصنيف وتحديد سرطان الرئة في مراحله المبكرة.*