

Automated Lung Cancer Detection a Comparison amongst Physicians: A Literature Review

Kaviya Sathyakumar, Michael Munoz, Snehal Bansod, Jaikaran Singh, Jasmin Hundal, B Benson A. Babu*

Hospital Medicine, Northwell Health, New York, USA

ABSTRACT

Introduction: Lung cancer is the number one cause of cancer-related deaths in the United States as well as worldwide. Radiologists and physicians experience heavy daily workloads thus are at high risk for burn-out. To alleviate this burden, this literature review compares the performance of four different AI models in lung nodule cancer detection, as well as their performance to physicians/radiologists.

Methods: 648 articles were extracted from 2008 to 2019. 4/648 articles were selected. Inclusion criteria: 18-65 years old, CT chest scans, lung nodule, lung cancer, deep learning, ensemble and classic methods. Exclusion criteria: age greater than 65 years old, PET hybrid scans, CXR and genomics. Outcomes analysis: Sensitivity, specificity, accuracy, sensitivity-specificity ROC curve, Area under the curve (AUC). Data bases: PubMed/MEDLINE, EMBASE, Cochrane library, Google Scholar, Web of science, IEEEExplore, DBLP.

Conclusion: Hybrid Deep-learning architecture is state-of-the-art architecture, with a high-performance accuracy and low false-positive reports. Future studies, comparing each model accuracy in depth, would be valuable. Automated physician-assist systems such as this hybrid architecture, may help preserve a high-quality doctor-patient relationship and reduce physician burn out.

KEYWORDS: Lung Cancer; Lung Nodule; Lung Tumor; Malignancy; Metastasis

INTRODUCTION

Lung cancer is the number one cause of cancer-related deaths in the United States and worldwide [1]. Furthermore, lung cancer has amongst the highest public burden of cost worldwide. Healthcare cost to Medicare beneficiaries were analyzed [2]: the highest costs were related to surgery and an estimated \$30,000 over a 15-year period. Similarly, patients receiving chemotherapy and radiation therapy faced a cost of \$4000-\$8000 per month, with an average life expectancy of 14 months from the time of diagnosis [2]. Europe's incidence of lung cancer is estimated to be 60 per 100,000 inhabitants. Its costs of healthcare and management for the patient post-intervention are estimated to be 17,000 Euros per year [3].

The National Lung Screening Trial (NLST) found that examination with Low-Dose Computed Tomography (LDCT) instead of the standard chest X-ray, in a high-risk population, led to a 20% reduction in mortality rate [8]. Additionally, the detection rate of lung cancer screening with low-dose CT is 2.6 to tenfold higher than that with chest radiography [3]. The key to reducing lung-cancer related deaths is early diagnosis and this relies on fast

and accurate detection of lung nodules and careful examination of chest CT scans to determine malignancy: a process which requires considerable time and effort on behalf of radiologists and physicians.

According to a recent study, physicians spend 75% of each patient visit on activities other than face-to-face patient encounter [4], including working with the EMR. Studies also found that physicians from various specialties spend up to 2 hours on administrative duties for each hour that they see patients in the office, followed by an additional 1 to 2 hours of work after clinic, mostly devoted to the EMR [5]. It is likely, although not investigated, that these are much higher for physicians screening patients at risk for lung-cancer, due to the time required for the initial examination and evaluation of CT scans.

During the 18th World Conference on Lung Cancer (WCLC), Dr Flanou confirmed that oncologists were at highest risk from burn-out compared to other physicians as well as other oncology care staff (nurses, psychologists and social workers), with a reported prevalence between 35-60%. Amongst individuals who suffer burn-

Correspondence to: Benson A. Babu, Hospital Medicine, Northwell Health, New York, USA, E-mail: bensonbabu@gmail.com

Received: February 03, 2021, **Accepted:** February 17, 2021, **Published:** February 24, 2021

Citation: Sathyakumar K, Munoz M, Bansod S, Singh J, Hundal J, et al. (2021) Automated Lung Cancer Detection a Comparison amongst Physicians: A Literature Review. J Vasc Med Surg. 9:407.

Copyright: © Babu BA, et al. 2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

out there is a risk of mental health issues in 20-35%, moreover in physicians it is associated with a decrease in empathy towards patients and reduced quality of care [6]. It is therefore of utmost importance that all ways in which the burden of work on physicians may be reduced, should be explored, for the wellbeing of both the patients and physicians.

One such solution is AI automated CT lung cancer detection, which can be used to assist physicians: thereby reducing their burden of work; optimizing hospital operational workflow; and providing more time to develop a high-quality doctor-patient relationship. A computer-aided detection (CAD) system was first introduced by Niki et al. (2001) as a means to extract and analyze data from CT scans, classify benign and malignant lung cancer changes, and for the purpose of screening patients using 3D CT scans [7]. Since then, numerous studies have found improved detection of lung nodules on CT scans when examination by a physician/radiologist is combined with the use of a CAD system [9,10]. Improved radiologist performance with CAD was noted especially in the detection of small lung nodules, <5 mm in size, which are often easily overlooked by visual inspection alone [1]. Thus, CAD and its associated AI models help not only to reduce the burden of work on physicians, and subsequently fatigue-related errors of judgement, but to improve detection of nodules particularly in the early stages of lung cancer, which are more likely to be missed.

METHODS

PICO Framework, Problem: Lung Cancer, Intervention: Machine and Deep Learning, Comparison: Deep learning Ensemble CNN vs Classic Machine Learning Model performance, Outcomes: Sensitivity, measures how well the algorithm recognizes the type of nodule correctly, Specificity measures the ability of the algorithm to remove the false positives, and a high specificity value means a low rate of misdiagnosis, Accuracy measures the proportion of data that was classified correctly. Sensitivity-specificity ROC curve and Area under the curve (AUC).

Data bases: PubMed/MEDLINE, EMBASE (or Scopus), Cochrane Library, Google Scholar, Web of Science, IEEEExplore, DBLP. Searched terms strategies used are Boolean and fuzzy logic, truncated terms, and wild card.

648 articles were extracted. Two independent reviewers selected 4/648 studies: article year range 2008-2019.

Inclusion criteria: 18-65 years old, CT chest scans, Lung Nodule, Lung Cancer, Deep learning, ensemble and classic methods. Exclusion criteria: Greater than 65 years old, PET hybrid scans, CXR, genomics.

In this experiment, a hybrid model was proposed: for this specific task, LeNet, AlexNet, and VGG-16 were used. In addition, the features obtained from the last fully-connected layer of CNNs were applied as input for the following machine learning/classification models: linear regression (LR), linear discriminate analysis (LDA), decision tree (DT), support vector machine (SVM), k-nearest neighbor (kNN) and softmax. All the machine learning classifiers were tested at the end and examined separately by comparing their performance. In order to increase the classification accuracy, image augmentation techniques were used during the training of the models. In this scope, approximately 20 additional images were obtained from each original sample in the dataset. Lastly, the mRMR feature selection method was used to find the most

efficient features, which were then applied as the input in the above-mentioned method.

RESULTS AND DISCUSSION

The main reason the Minimum redundancy, maximum relevance feature selection method with CNN performed better than the methods described in the three other papers, is the use of additional techniques such as image augmentation, principal component analysis (PCA), mRMR and appropriate feature selection.

In this method, during the last couple of iterations, the dimensions of the feature set obtained using image augmentation techniques were reduced using PCA before the classification task. The KNN classifier was then fed with the reduced feature set, resulting in an accuracy of 97.92 %. Then, the KNN classifier was fed using the mRMR algorithm with the 1000 features obtained from the fc8 layer of AlexNet architecture. 33, 50, 100, 150 and 200 of the most efficient features were determined and ranked, respectively. The extracted features were reclassified with KNN. A 10-fold cross-validation method was used for testing.

PCA decreases the classification accuracy from 98.74% to 97.92 %. The PCA method obtained this level of success with only 33 features and consumed less time when training the model, due to the use of fewer features. In addition, the performance results of the KNN with and without PCA method were close.

Next, the most efficient features were selected by the mRMR method of 1000 features, obtained from the last layer of AlexNet without using the PCA method. The best rate of success obtained was 99.51 % with 200 features provided by mRMR. It was found that using 100, 150 or 200 features from the mRMR algorithm, was more successful than using all 1000 features obtained from the fc8 layer of AlexNet.

After this point, the experiment was extended by focusing on the KNN classifier. In this scope, the k value corresponding to the number of the nearest neighbors was searched in the range of 100 and 102 considering various distance functions by using the Bayesian optimization method. Notably, the classification success decreased relatively and gradually as the k value increased. The most efficient results were ensured for KNN when the k was set to 1 and the distance function was adjusted to Correlation. In this experiment, the 10-fold cross-validation was also used for evaluation. The model achieved an accuracy of 99.51 %, sensitivity of 99.32%, specificity of 99.71 % and F-score of 99.51 %.

Numerous studies assessing the performance of radiologists in lung nodule detection show low inter-observer agreement, varying sensitivities ranging from 30-97%, and false positive counts of 0.6-2.1 per patient, depending on the input data, method and criteria for identification [2]. A study from the NLST, assessed CAD retrospectively in 134 subjects and found an improved inter-observer agreement (kappa increase from 0.53- 0.66): results confirmed by similar studies [2]. As well as reducing inter-observer variation, one of the greatest advantages of CAD remains the detection of smaller lung nodules that are easily missed by radiologists/physicians [1]. The use of CAD by 2 radiologists in an emergency clinic study, did find improved reading time when CAD was used (Radiologist 1 94.6 s vs. 102.7 s, P>0.05; Radiologist 2 61.1 s vs. 76.5 s, P<0.05). Although this decrease in reading time was not statistically significant for both radiologists, they did get a significantly improved rate of nodule detection: 34% and 27% for

Radiologists 1 and 2 respectively when CAD was reviewed after the CT images, but not when it was reviewed before the scans [10].

An observer performance study compared the performances of 10 radiologists without and with the use of CAD, in 50 CT examination cases [3]. Alternative free-response ROC curves for each output (with and without CAD) were calculated by plotting the true-positive fraction against the likelihood of obtaining an image with false-positive findings (i.e. with one or more false-positive lesions) at each confidence level. Using the area under each alternative free-response ROC curve (A_z) to compare the observers' performances, they found that the performance of all observers was significantly improved with the use of CAD. Routine use of CAD by radiologists and physicians, especially in high-pressure environments, is justified due to improved rates of lung nodule detection, inter-observer agreement, interpretation speed, higher true-positive to false-positive ratios and for detection of small (<5 mm) nodules.

CONCLUSION

The experiment conducted here performs well but it uses a very small dataset thus, may not perform well on a large production scale. Ideally, the models should be tested on a larger dataset to ensure they work on large, real production data. Also, the image augmentation method was used here to increase the number of images: these techniques may create very correlated images which can lead to overfitting. Another indication of correlation might be the KNN algorithm, which relies on the nearest neighbor, as it performed best on this dataset. It would be beneficial to further test these models on new dataset, which is relatively large and from a different data source. Also, the test dataset should not undergo image augmentation, but be tested in its original form.

REFERENCES

1. Herbst RS, Heymach JV, Lippman SM. Lung cancer. *N Engl J Med*. 2008;359:1367-1380.
2. Hruby A, Sweany H. Primary carcinoma of the lung. *Arch Intern Med*. 1933;52:497-498.
3. Eagan RT, Fleming TR, Schoonover V. Evaluation of response criteria in advanced lung cancer. *Cancer*. 1979;44:1125-1128.
4. Rivera MP, Mehta AC. Initial diagnosis of lung cancer: ACCP evidence-based clinical practice guidelines. *Chest*. 2007;132(13):131-148.
5. Kulkarni A, Panditrao A. Classification of lung cancer stages on CT scan images using image processing. *Int Conf Adv Techno*. 2014;65(10):1384-1388.
6. Kimura T, Matsuura K, Murakami Y, Hashimoto Y, Kenjo M, et al. CT appearance of radiation injury of the lung and clinical symptoms after stereotactic body radiation therapy (SBRT) for lung cancers: are patients with pulmonary emphysema also candidates for SBRT for lung cancers? *Int J Radiat Oncol Biol Phys*. 2006; 66: 483-491.
7. Munshi A, Krishnatry R, Banerjee S, Agarwal JP. Stereotactic conformal radiotherapy in non-small cell lung cancer - an overview. *Clinical oncology*. 2012;24: 556-568.
8. Sunyach MP, Falchero L, Pommier P, Perol M, Arpin D, et al. Prospective evaluation of early lung toxicity following three-dimensional conformal radiation therapy in non-small-cell lung cancer: preliminary results. *Int J Radiat Oncol Biol Phys*. 2000;48: 459-463.
9. Takeda A, Ohashi T, Kunieda E, Enomoto T, Sanuki N, et al. Early graphical appearance of radiation pneumonitis correlates with the severity of radiation pneumonitis after stereotactic body radiotherapy (SBRT) in patients with lung tumors. *Int J Radiat. Oncol Biol Phys* 2010;77: 685-690.
10. Makimoto T, Tsuchiya S, Hayakawa K, Saitoh R, Mori M. Risk factors for severe radiation pneumonitis in lung cancer. *Japanese journal of clinical oncology*. 2000;29: 192-197.