Smart Health Care Management System for Diagnosis of Lungs Cancer

Ahed Abugabah[1]*, Farah Shahid[2], Amjad Alafeef[3] and Rizwan Khan[4]

College of Technological Innovation, Zayed University, Abu Dhabi, UAE

Department of Computer Science, University of Agriculture, Sub Campus (Burewala-Vehari), Faisalabad, Punjab, Pakistan

Ministry Of culture, Jordan, Amman alafeef2006@yahoo.com

School of Computer Science and Mathematics, Zhejiang Normal University, Jinhua 321004, China
Ahed.abugabah@zu.ac.ae

Abstract

Lung cancer has a mortality rate that is significantly greater than other forms of cancer, making it the second biggest cause of death worldwide after cardiovascular disease. Detecting lung cancer in its earliest stages has been the focus of a significant amount of research and development over the past few years, leading to the development of several unique computer-aided diagnostic tools that use deep learning. On the other hand, deep learning models are readily susceptible to overfitting, and this issue invariably results in decreased performance. In this study, we proposed a CNN-based approach for the classification of lung cancer and attained 95.62% accuracy. When applied to classifying lung cancer, the solution achieves the most outstanding performance possible throughout the entire dataset. The overfitting issue that arises during lung cancer classification tasks may be solved with the help of the proposed framework, which also outperforms existing methods that are considered to be state-of-the-art.

Keywords: Convolutional neural network, lungs cancer, classification

1. Introduction

Lung cancer is the most lethal form of the disease, as evidenced by the fact that it has the most excellent fatality rate. Many lives can be saved with early detection. Following prostate cancer in males and breast cancer in females as the most frequent type of cancer, lung cancer ranks second. Lung cancer can develop from pulmonary glands, which are tiny, spherical lesions seen inside the lungs [1]. Because of its small size and placement in the glands, lung cancer in its initial stages is often undetectable on CT scans; symptoms only arise later in the condition. Computed tomography (CT) can be used to detect lung nodules, which presents a difficulty for doctors. Both in the course of ordinary clinical care and the context of diagnosis, a growing percentage of these pulmonary nodules are being discovered due to technological advancements in CT scanners [2]. About half of smokers will develop a lung nodule at some point. Both computed tomography and magnetic resonance imaging are standard diagnostic tools used in modern medicine for early discovery, which increases the likelihood of a patient surviving their condition [3].

Several recent research has suggested utilizing AI, and in particular Deep Learning, to describe pulmonary nodules to reduce the number of scans required to determine whether or not they are benign or cancerous. The accuracy of these techniques is relatively moderate at the moment, with the majority of these investigations encompassing more prominent nodules with diameters of up to 30 millimeters [4]. Because lung cancer nodules are typically more extensive than benign ones, big nodule sizes lead to skewed data sets. As a result, whether such Deep Learning approaches discover nodule characteristics that are unique to lung cancer or whether they primarily only stratify nodules based on size is still open for debate. According to the statistics, how lung cancer is discovered in its earlier stages is the primary factor responsible for disease and death brought on by the condition. The purpose of this study is to determine whether or not the Epidermal Growth Factor Receptor is mutagenic [5].

Radiologists widely use screening to assess the quality of CT images across various specialties. While automated algorithmic solutions have the potential to be helpful, the interaction between such tools and medical professionals presents its own set of difficulties. Low-dose CT scans are a solution that has been presented. It has been reported that a significant number of lives can be preserved if lung cancer is detected at an early stage [6]. Radiologists are faced with a duty that is not only difficult but also time-consuming and repetitive: the early detection of lung cancer nodules. They suggested using a method known as a "Multi-Scene Deep Learning Framework (MSDLF)" in conjunction with the "vesselness filter" to raise the level of accuracy while also lowering the number of false positives. The primary objective of this study is to identify significant nodes with diameters of more than 3 millimeters [7]. The Experience and understanding Assessment of Premature Prediction Network, or KAMP-Net, was proposed to determine the likelihood of mortality in lung cancer patients. Data augmentation is the method that is employed in this approach to training Convolutional Neural Networks (CNN) [8].

They hypothesized that adding more data would improve CNN's capabilities, and this was one of their working hypotheses. The medical data are used to train a classifier called a Support Vector Machine (SVM), and then the findings from the CNN and the SVM are combined to generate a probability of death [9]. All of the clinical evaluations were obtained through manual acquisition. Multi-Channel Image Coding, Network Development and Implementation, Incorporation of Deep Learning, and Clinical Knowledge are the steps that are involved in this method. This strategy uses the data collected during the National Lung Screening Trial (NLST)[10].

As a result, the application of the deep learning method to categorizing cancer classes is investigated in this work. It was hypothesized that deep learning may help in the detection of cancer and that it could acquire better outcomes than conventional methods of machine learning, given that the use of deep learning had previously been shown to be superior to all AI systems [11]. The convolutional neural network, also known as a CNN, was the deep learning technology that demonstrated the most consistency. As a result, a CNN was developed to differentiate cancer classes. The numerous methods of pooling utilized by CNN were examined, tested, and contrasted. This study, in contrast to previous work, investigates a method known as deep learning with ten convolutional layers by using max and average pooling size, which is the primary distinction between the two. The CNN can automatically extract task-related features, which are referred to as cancer features in this study. It does not require the actions of users. Additionally, CNN is renowned for its high level of accuracy, which is comparable to or even superior to human specialists.

2. Material and Methods

We introduced the CNN-based three models in the proposed study, which have ten layers. It is an optimal model which finalized after different experiences with data samples. In the proposed models, we add the two types of pooling layers and change the activation function. Here we used two types of lung cancer such as adenocarcinoma (AC) and lung squamous cell carcinoma (LSCC) [12]. The first is the standard type of cancer in the USA, and the second is primarily located in the center of the lungs on the central airway. Here we used histopathological images. The total number of images for classes is 10000. Initially, we resize the data samples according to the model requirements and then provide the data samples to the Model to get the more critical features from the input samples.

2.1CNN Model

The traditional approaches to computer vision can be broken down into three distinct stages. The first stage is extraction, the second stage is reducing features, and the third stage is categorizing. Despite this, researchers typically merge these phases into a single convolutional neural network (CNN). That eliminates the requirement for CNN to configure the function manually. On the other hand, the weights that fit the initial layers were used for feature extraction, and their values were acquired through iterative learning. Additionally, CNN can achieve a higher performance level than its contemporaries, such as feedforward neural networks [13]. In our proposed models, we used the seven convolutional layers and three fully connected layers with different activation functions. For these activation functions, we used such as Sigmoid, ReLU and likely ReLU. Figure. One shows the proposed model flow chart.

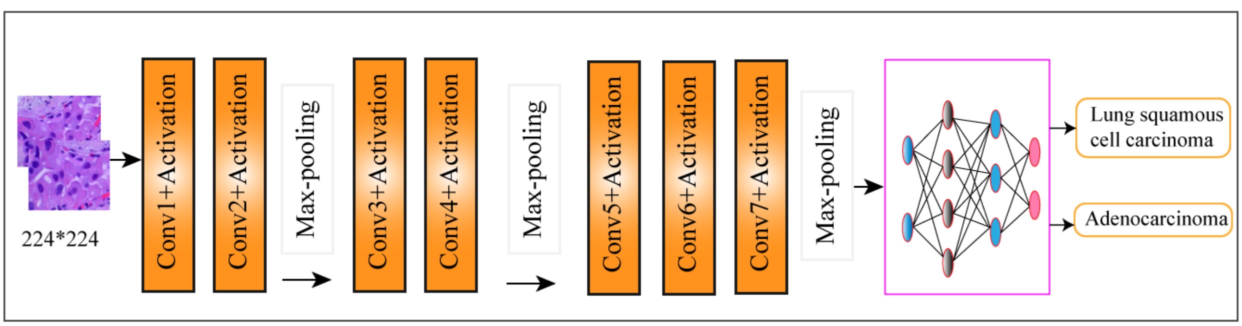

Figure 1. Proposed CNN-based Model for the classification of lung cancer.

3. Results and Discussion

In the proposed study, we used 10000 images. We used the Keras library with 64 GB RAM during the experimental process. In this work, we proposed three models with different activation functions. We also saw the overall behavior of the Model during training. For classification purposes, we split the data, such as 80% for training and 20% for testing.

Table 1. The proposed CNN-based model results for classifying AC vs. LSCC with different activation functions.

Classes	Activation Function	Accuracy	Sensitivity	Specificity
AC vs. LSCC	ReLU	93.05	91.56	94.55
AC vs. LSCC	LReLU	95.62	93.38	96.87
AC vs. LSCC	Sigmoid	90.36	89.37	91.36

In Table 1. We can see the proposed models produce promising results in terms of accuracy. LReLU activation function-based Model obtained 95.62% accuracy, 93.38% sensitivity and 96.87% specificity, respectively. In Table 2. We can see the comparative analysis with other CNN-based approaches, but our technique produced promising results in terms of accuracy.

Table 2. Proposed model comparison with other techniques.

Approaches	Accuracy
Bukhari et al. [14]	93.91%
Hlavcheva et al. [15]	94.61%
Proposed Model	95.62%

Conclusion

In recent years, lung cancer has emerged as one of the diseases with the highest incidence rate. According to the study that has been done in this area, each year, there are more than 200,000 cases discovered in the United States. When cell division and growth in the lung are not adequately regulated, the result is the production of cancerous tumors. In this study, we proposed the CNN based for the classification of lung cancer. Primarily CNN based Model have an issue related to overfitting which directly impacts the model performance in terms of accuracy. In this work, we overcame the overfitting issue and produced the best accuracy results for the classification of lung cancer and attained 95.62%.

Acknowledgement

This work is supported by Zayed University, Office of research, Abu Dhabi UAE. Ref:

RIF R22050, 2022

References

[1] M. A. Heuvelmans *et al.*, "Lung cancer prediction by Deep Learning to identify benign lung nodules," *Lung Cancer*, vol. 154, no. November 2020, pp. 1–4, 2021.

[2] P. Nanglia, S. Kumar, A. N. Mahajan, P. Singh, and D. Rathee, "A hybrid algorithm for

lung cancer classification using SVM and Neural Networks," *ICT Express*, vol. 7, no. 3, pp. 335–341, 2021.

[3] Y. Gu *et al.*, "A survey of computer-aided diagnosis of lung nodules from CT scans using deep learning," *Computers in Biology and Medicine*, vol. 137, no. August. 2021.

[4] F. Kanavati *et al.*, "A deep learning model for the classification of indeterminate lung carcinoma in biopsy whole slide images," *Sci. Rep.*, vol. 11, no. 1, pp. 1–14, 2021.

[5] K. Pradhan, P. Chawla, and S. Rawat, "A deep learning - based approach for detection of lung cancer using self adaptive sea lion optimization algorithm (SA - SLnO)," *J. Ambient Intell. Humaniz. Comput.*, no. 0123456789, 2022.

[6] A. Asuntha and A. Srinivasan, "Deep learning for lung Cancer detection and classification," *Multimed. Tools Appl.*, vol. 79, no. 11–12, pp. 7731–7762, 2020.

[7] H. Yang *et al.*, "Deep learning-based six-type classifier for lung cancer and mimics from histopathological whole slide images: a retrospective study," *BMC Med.*, vol. 19, no. 1, pp. 1–14, 2021.

[8] A. Mehmood, A. Abugabah, A. A. AlZubi, and L. Sanzogni, "Early Diagnosis of Alzheimer's Disease Based on Convolutional Neural Networks," *Comput. Syst. Sci. Eng.*, vol. 43, no. 1, pp. 305–315, 2022.

[9] A. Abugabah, A. Mehmood, S. Almotairi, and A. A. L. Smadi, "Health care intelligent system: A neural network based method for early diagnosis of Alzheimer's disease using MRI images," *Expert Syst.*, no. March, pp. 1–13, 2022.

[10] G. S. Tran, T. P. Nghiem, V. T. Nguyen, C. M. Luong, J. C. Burie, and Y. Levin-Schwartz, "Improving Accuracy of Lung Nodule Classification Using Deep Learning with Focal Loss," *J. Healthc. Eng.*, vol. 2019, 2019.

[11] R. Khan, A. Akram, and A. Mehmood, "Multiview Ghost-Free Image Enhancement for In-the-Wild Images with Unknown Exposure and Geometry," *IEEE Access*, vol. 9, pp. 24205–24220, 2021.

[12] A. A. Borkowski, M. M. Bui, L. B. Thomas, C. P. Wilson, L. A. DeLand, and S. M. Mastorides, "Lung and Colon Cancer Histopathological Image Dataset (LC25000)," *arXiv 2019, arXiv:1912.12142.*, pp. 1–2, 2019.

[13] A. Mehmood, M. Maqsood, M. Bashir, and Y. Shuyuan, "A deep siamese convolution neural network for multi-class classification of alzheimer disease," *Brain Sci.*, vol. 10, no. 2, 2020.

[14] S. U. K. Bukhari, A. Syed, S. K. A. Bokhari, S. S. Hussain, S. U. Armaghan, and S. S. H. Shah, "The Histological Diagnosis of Colonic Adenocarcinoma by Applying Partial Self Supervised Learning," *medRxiv*, p. 2020.08.15.20175760, 2020.

[15] D. Hlavcheva, V. Yaloveha, A. Podorozhniak, and H. Kuchuk, "Comparison of CNNs for Lung Biopsy Images Classification," *2021 IEEE 3rd Ukr. Conf. Electr. Comput. Eng. UKRCON 2021 - Proc.*, pp. 1–5, 2021.