IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES
ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

Lung Cancer Detection Using Deep Learning

Mr.Shaik Wasim Akram[1], Assistant Professor, Department of CSE,
Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt., Andhra Pradesh.

Kolakaluri Tejaswini[2], **Kumbha Nagarjuna[3]**, **Magam Gayathri[4]**, **Godavarthi Bala Sai Teja[5]**
[2,3,4,5] UG Students, Department of CSE,
Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt., Andhra Pradesh.
shaikwasimakram585@gmail.com,kolakaluritejaswini@gmail.com,gayathri.magam01@gmail.com,nagarjunakumbhacse@gmail.com,balasaiteja1020@gmail.com

DOI:10.48047/IJFANS/V11/I12/189

Abstract
Cancer is a disease in which the body's cells multiply uncontrollably. When referring to cancer that starts in the lungs, the term "lung cancer" is used. Lung disease is among the most prevalent illnesses that are impacted the initial phases of treatment to improve patient survival rates. Lymph nodes or other bodily organs may become affected by lung cancer once it has progressed from the lungs. Pulmonary cancer, also referred to as lung carcinomas the main reason for cancer-related deaths worldwide. By identifying malignancy that has not yet spread, computed tomography (CT) has the potential to avert hundreds of thousands of fatalities each year. A powerful and emerging technique for feature learning and pattern identification is deep learning. In this research, we developed a convolutional neural network model for the detection of lung cancer. According to our findings, Lung cancer is treated with convolutional neural networks diagnosis more often than other deep learning techniques.

Keywords: Pulmonary Cancer Detection, Lymph node, Deep Learning Network, Convolutional neural network.

Introduction
Pulmonary tumours are now regarded being among the most lethal illnesses. Lung tissue's erratic and uncontrolled cell proliferation is the primary contributor to lung cancer. Among the causes is smoking. Early detection increases the likelihood of a successful recovery. CT pictures are one type Today, one of the worst diseases is regarded as being pulmonary cancer. Lung cancer is primarily caused by the irregular and unchecked cell growth of lung tissue. Smoking is one of the factors. A successful recovery is more likely when it is discovered early. New statistics each year show that there are currently over 1.6 million lung cancer patients in the United States. Both men and women are affected by the same dangerous illness. One form of filtering method that uses appealing areas is CT imaging.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

capture images in films.Therefore, In order to get an accurate answer right away, we can apply modern methodologies that put to use the processing of images with machine learning domains. The main objective is usually to automatically determine if lung nodules are cancerous (Malignant) or Non-Cancerous(benign). Convolutional neural network layers are used in this paper to achieve good accuracy.

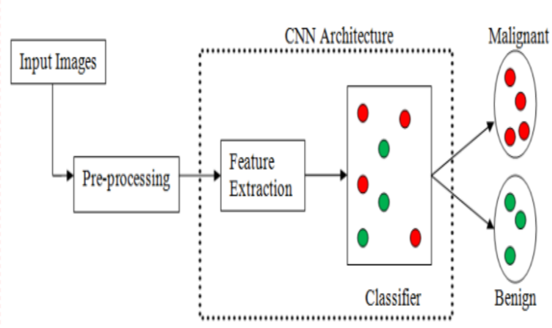

Figure 1: Description

Literature Review

Several studies have been conducted on the detection and classification of lung cancer using different algorithms; some of them are addressed in this publication:

The Writer The first paper examines stage progression of lung cancer and makes use of a variety ofmethods for image processing and machine learning, includingBinarization, noise reduction, and grayscale conversion.The source photos were obtained from the ELCAP Database of Public Lung Images, which comprises over 200 lung images from the CT scan results of cancer- and non-cancer-patients. The Amazon cloud that's where the data is stored. The framework development system employs MySQL as the database andThe scripting language used to save the images from the CT scan report is Java Script Pages (JSP).Using SVM, image processing methods, the research has provided a viable way for identifying stage progression of lung cancer.Also, it aims to deliver outcomes that are more precise[1].Data from Data World and the UCI Machine Learning Repository were used to build the datasets for this study on lung cancer. They arrived at a result after examining the accuracy percentages of each categorization model they employed. This study compares the classification approaches of SVM, Logistic Regression, Naive Bayes, and Decision Tree. For the best results, the SVM formula categorizedthe highly dimensioned observation. For The UCI Machine Learning Repository[2].combining data discovery with biological image processing techniques, the objective of this project is to increase specificity and define a value for the early identification of lung cancer. Both the 25 image dataset for testing and the 75 picture dataset for training were collected from medical photographs. Throughout the process, two algorithms—using the SVM classification technique both the random forest approach and were employed. With the help of the given CT scan images, this method aims to pinpoint the lung's cancerous region. If new sets of photos can be provided with the help

 UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

of the logic regression approach or SVM algorithm, the accuracy of the entire system can be improved[3]. This paper proposes aapproach for automatically viewing CT scan images to look for lung cancer.The pictures were obtained for use in study from the database of the Cancer Imaging Archive.It is advised to utilize image preprocessing techniques likemedian filtering to identify lung regions of interest before segmenting those regions using mathematical morphological algorithms. To classify lung CT scans into normal and malignant, From the segmented tumor region, three geometrical features—area, perimeter, and eccentricity—were obtained and fed to the classifier.Hence, the proposed methodology helps identify lung cancer precisely and at an early stage[4].Using image processing methods of an artificial neural network classifier, this proposed system was created to create a computer-based detection system for lung cancer and to diagnose lung cancer at an early stage. This technique uses a total of 20 images of malignancy and 20 images of healthy tissue for testing. As compared to The accuracy of the ANN classifier with specific features used in this system outperforms SVM, KNN, and DT classifiers superior.In contrast to SVM classifier, the ANN classifier requires more data inputs, but it also produces outputs that are more exact and accurate[5].

Problem Identification

Lung cancer ranks among the leading causes of cancer-related fatalities in the globe. The likelihood of effective treatment and greater survival rates for lung cancer can be significantly increased by early diagnosis and identification. However, there are a number of difficulties with lung cancer detection, lack of early symptoms, imaging limitations, difficulty in obtaining tissue samples, Researchers and healthcare professionals are investigating new technologies and methods for detecting lung cancer in an effort to overcome these challenges. These methods include advanced imaging methods and machine learning algorithms that can analyze imaging data and find patterns suggestive of lung cancer. [7-15]

Methodology

Using a flowchart, Figure 2 demonstrates the suggested technique for using CT scans to find lung cancer.The technique consists of five fundamental steps, each of which is described in more detail in the section below.

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

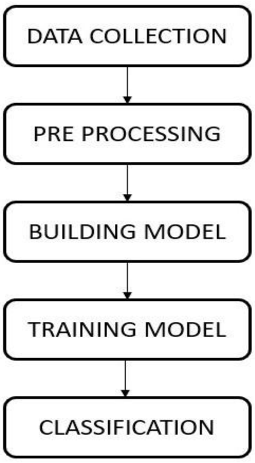

Figure 2 :Flowchart

Data collection:

The initial step is to obtain lung CT images of cancer patients. The images were downloaded for research from the lung cancer dataset at the National Center for Cancer Diseases at Iraq-Oncology Teaching Hospital (IQ-OTH/NCCD). It comprises CT images of healthy individuals as well as lung cancer patientswho are at various stages of the disease. In malignant tumors, cancer is present (ie., they invade other sites). In the lymphatic or circulatory systems, they can spread to far-off locations.

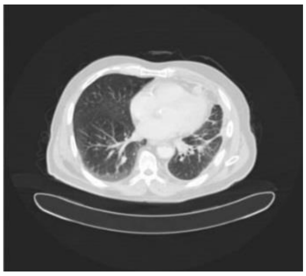

Figure 3: Malignant Tumor

A tumor is said to as "benign" if it does not spread from its initial place to other regions of the body. They don't spread to nearby structures or to far-off regions of the body. Benign tumors frequently have clearly defined borders and develop gradually. Normal benign tumors do not pose a threat.

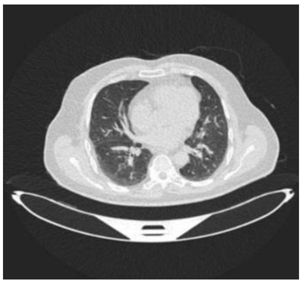

Figure 4: Benign Tumor

 UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

Preprocessing:

In image processing and computer vision tasks, resizing is a typical preprocessing method. It entails resizing an image to a particular width and height. When working with images of various sizes, such as in a dataset of lung CT scans, resizing is frequently necessary.

Rescale is the process of increasing an image's pixel values to a predetermined range. By dividing each pixel value by the highest pixel value present in the image, the most typical range of values is [0, 1].

Classification:

We created a model using the CNN architecture and trained it with data. Determine whether a nodule is benign or malignant using this trained classifier.

Implementation

CNN Model Implementation:

Convolutional neural networks (CNNs) are implemented utilising a variety of deep learning techniques for lung cancer diagnosis, including:

Data Preparation:

assembling and putting together the lung cancer image collection from IQ-OTH/NCCD. The dataset contains a balanced mixture of cancerous and non-cancerous pictures.

Data Pre processing:

Performing picture preprocessing to make sure the images' size and brightness are standardised. Regular preprocessing methods include contrast amplification, resizing and rescaling.

Model Architecture:

Convolutional and pooling layers are frequently used in multiples when designing the CNN architecture(Figure 5). The amount and complexity of the dataset will determine the exact architecture, which can then be adjusted using methods like transfer learning.Three sets-one for training, one for validation, and one for testing-were created from the dataset. To train the model, utilise the training set.

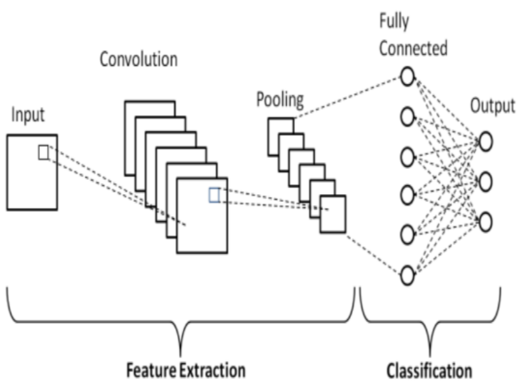

Figure 5: CNN Architecture

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

Use of the validation set is for examine How accurately the model performed during training.After training and adjusting, the model's performance is evaluated using the test set.On the architecture of CNN, we constructed the model. Computer vision applications including image categorization, object identification, and segmentation typically use convolutional neural networks (CNN). CNNs process images through a number of layers of convolutional filters, pooling layers, and fully connected layers in order to identify patterns in the data.

Model Training:

Adam optimizer is the optimizer that we defined. An optimization approach called Adam (Adaptive Moment Estimation) is used to adjust a neural network's parameters while it is being trained.

We have utilized sparsecategorical cross-entropy in the compilation model. By calculating the cross-entropy between them, it calculates the difference between the genuine label distribution and the anticipated label distribution.

Sparse categorical cross-entropy has the following formula:$-\sum_i y_i \log(\hat{y}_i)$

After that, with a batch size of 62, fit the model on the training data for 15 epochs. That is the validation set also used to monitor the model's performance while it is being taught.

Results

Utilising a different set of validation data to assess the effectiveness of the trained model.Using the history object generated by Keras' fit method, we visualised Fig. 6 shows the accuracy and loss of a trained model to assess the model's performance. Each epoch's training and validation accuracy and loss are represented in the history object.

The confusion matrix, which is depicted in Fig.7, displays the system's overall accuracy.The true class and the expected class are connected to the confusion matrix's rows and columns, respectively. We could generate five distinct metrics quantifying the validity of our model using the terms true positive (TP), true negative (TN), false positive (FP), and false negative (FN) from our confusion matrix (FN).

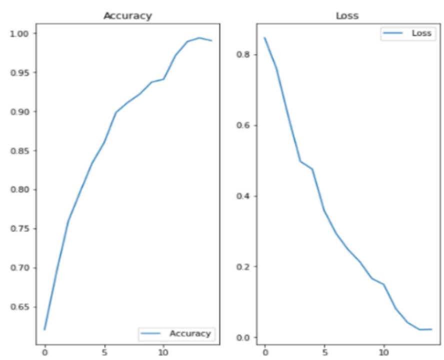

Figure 6: Accuracy Graph

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

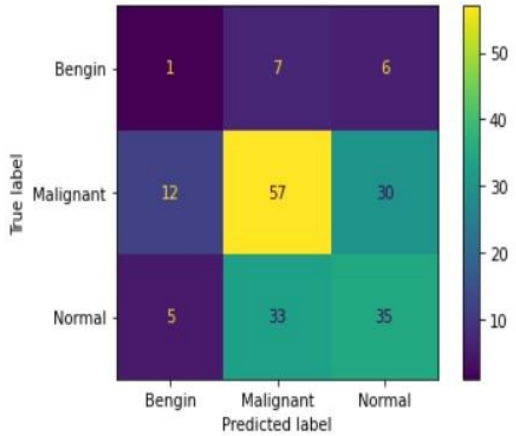

Figure 7: Confusion Matrix

Accuracy =TP + TN / TP + TN + FP + FN

Precision = TP / TP + FP

Sensitivity = TP / TP + FN

Specificity =TN / TN + FP

The classification report, which includes the precision, recall, f1-score, and support, is shown in the figure 8.

	precision	recall	f1-score	support
0	0.83	1.00	0.91	15
1	1.00	0.95	0.97	102
2	0.93	0.96	0.94	69
accuracy			0.96	186
macro avg	0.92	0.97	0.94	186
weighted avg	0.96	0.96	0.96	186

Figure 8 : Classification report

Our model's test accuracy for the classification task utilising CNN's basic architecture was 97.31%, and its training accuracy was 99.06%.Figures 9 and 10 illustrate the results after we deployed the model.The findings demonstrated that our model was highly accurate during both thetraining and validation phases.According to table 1, our model is effective.

Architecture	**Training Acc.**	**Testing Acc.**
CNN	99.06%	97.31%

Table 1: Results

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

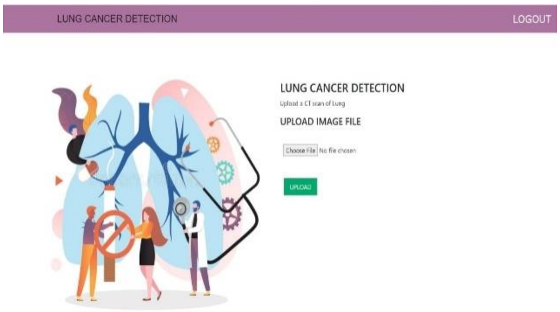

Figure 9: Image upload Page

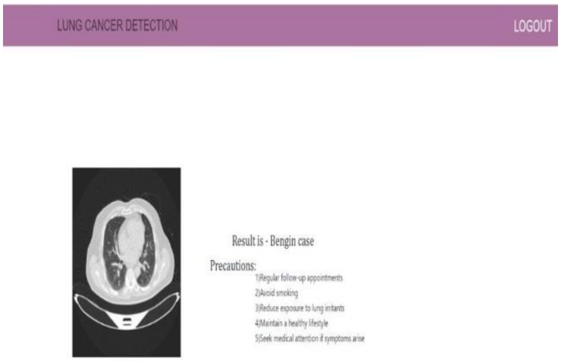

Figure 10: output page

Conclusion

Early lung cancer detection is vital for successful treatment and enhanced survival rates.In conclusion, convolutional neural networks (CNNs) have achieved promising results in the detection of lung cancer.In this research, we created a CNN model to distinguish between benign and malignant lung tumor tissues using CT scan images. Our model generated precise results with 97% accuracy.As a result, we are effective in accomplishing our objectives.

Future Scope

In the future ,this model can be extended for detection of other cancerous tumors.Researchers can investigate novel treatments that can enhance patient outcomes, such as immunotherapy and targeted medicines. In general, future research in lung cancer detection should focus on enhancing screening test precision, identifying high-risk individuals, and creating efficient lung cancer treatment options.

References

[1] W. Rahane, H. Dalvi, Y. Magar, A. Kalane and S. Jondhale, "Lung Cancer Detection Using Image Processing and Machine Learning HealthCare," 2018 International

Conference on Current Trends towards Converging Technologies (ICCTCT), Coimbatore, India, 2018,pp.15,doi:10.1109/ICCTCT.2018.8551008.

[2] R. P.R., R. A. S. Nair and V. G., "A Comparative Study of Lung Cancer Detection using Machine Learning Algorithms," 2019 IEEE International Conference on Electrical,Computer and Communication Technologies (ICECCT), Coimbatore, India, 2019,pp.14,doi:10.1109/ICECCT.2019.8869001.

[3] K. Roy et al., "A Comparative study of Lung Cancer detection using supervised neural network," 2019 International Conference on Opto-Electronics and Applied Optics (Optronix), Kolkata, India, 2019, pp.15,doi:10.1109/OPTRONIX.2019.8862326.

[4] N. S. Nadkarni and S. Borkar, "Detection of Lung Cancer in CT Images using Image Processing," 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2019, pp. 863-866, doi: 10.1109/ICOEI.2019.8862577.

[5] S. S. Bhise and S. R. Khot, "Early Stage Lung Cancer Diagnosis using ANN Classifier," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Coimbatore, India,2021,pp.22-27,doi:10.1109/ICAIS50930.2021.9395952.

[6] S. Das and S. Majumder, "Lung Cancer Detection Using Deep Learning Network: A Comparative Analysis," 2020 Fifth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), Bangalore, India, 2020, pp.3035,doi:10.1109/ICRCICN50933.2020.9296197.

[7] Sri Hari Nallamala, et al., “A Literature Survey on Data Mining Approach to Effectively Handle Cancer Treatment”, (IJET) (UAE), ISSN: 2227 – 524X, Vol. 7, No 2.7, SI 7, Page No: 729 – 732, March 2018.

[8] Sri Hari Nallamala, et.al., “An Appraisal on Recurrent Pattern Analysis Algorithm from the Net Monitor Records”, (IJET) (UAE), ISSN: 2227 – 524X, Vol. 7, No 2.7, SI 7, Page No: 542 – 545, March 2018.

[9] Sri Hari Nallamala, et.al, “Qualitative Metrics on Breast Cancer Diagnosis with Neuro Fuzzy Inference Systems”, International Journal of Advanced Trends in Computer Science and Engineering, (IJATCSE), ISSN (ONLINE): 2278 – 3091, Vol. 8 No. 2, Page No: 259 – 264, March / April 2019.

[10] Sri Hari Nallamala, et.al, “Breast Cancer Detection using Machine Learning Way”, International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume-8, Issue-2S3, Page No: 1402 – 1405, July 2019.

[11] Sri Hari Nallamala, et.al, “Pedagogy and Reduction of K-nn Algorithm for Filtering Samples in the Breast Cancer Treatment”, International Journal of Scientific and Technology Research, (IJSTR), ISSN: 2277-8616, Vol. 8, Issue 11, Page No: 2168 – 2173, November 2019.

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

[12] Kolla Bhanu Prakash, Sri Hari Nallamala, et al., "Accurate Hand Gesture Recognition using CNN and RNN Approaches" International Journal of Advanced Trends in Computer Science and Engineering, 9(3), May – June 2020, 3216 – 3222.

[13] Sri Hari Nallamala, et al., "A Review on 'Applications, Early Successes & Challenges of Big Data in Modern Healthcare Management'", Vol.83, May - June 2020 ISSN: 0193-4120 Page No. 11117 – 11121.

[14] Nallamala, S.H., et al., "A Brief Analysis of Collaborative and Content Based Filtering Algorithms used in Recommender Systems", IOP Conference Series: Materials Science and Engineering, 2020, 981(2), 022008.

[15] Nallamala, S.H., Mishra, P., Koneru, S.V., "Breast cancer detection using machine learning approaches", International Journal of Recent Technology and Engineering, 2019, 7(5), pp. 478–481.