IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 12, December 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.625

© 2024 IJIRCCE | Volume 12, Issue 12, December 2024| DOI: 10.15680/IJIRCCE.2024.1212011

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Recognition of Lung Cancer using Convolutional Neural Network

Devashish Sharma[1], Shanu Gour[2]

M. Tech Scholar, Department of Computer Science and Engineering, Bharti Vishwavidyalaya, Bhilai, Chhattisgarh, India[1]

Assistant Professor, Department of Computer Science and Engineering, Bharti Vishwavidyalaya, Bhilai, Chhattisgarh, India[2]

ABSTRACT: In the present life digital image processing plays important job. Digital image processing involving in clinical field changes the clinical world definitely. Clinical imaging is becoming quickly because of advancements in picture processing methods including picture acknowledgment, examination and upgrade. Presently a day's various kinds sicknesses that have confronting people in day-to-day existence. Cancer in the lungs is one of the significant reasons for death in people. Lung Cancer is ordinarily completed via prepared experts and these are significantly useful in beginning phase of identification. This recognition technique presents the chance of human blunders which outcomes a mechanized cycle. In this paper it is for the most part expects to identify the cellular breakdown in the lungs at a beginning phase through a mechanized cycle and decreases the human mistakes and comes by the precise outcomes. There are various kinds of handling procedures are presently utilizing. Beginning phase of identification will save the human existence. With the usage of a convolutional neural network technique architecture, a PC supported technique was created through this exploration to assist with the ID of patient cases as normal, benign, or malignant. 550 CT scan images are utilized as dataset. High accuracy is provided by the suggested model, reaching 99.7%.

KEYWORDS: Chest CT scan image, Lung Cancer, Artificial Intelligence, Convolutional Neural Network

I. INTRODUCTION

Cancer in the lungs is one of the sicknesses and we can likewise call it as lung carcinoma and it's a malign lung tumor. It is predominantly recognized by the unconstrained cell development in lung tissues. These are gathered of malignant growth cells that are grouped into two kinds. One is dangerous growth and second one is harmless cancer. On the off chance that the distinguishing proof requires some investment, it causes passing of human existence. There are essentially two kinds of cellular breakdowns in the lungs are in particular- a) SCLC, b) NSCLC. In this SCLC is alluded as small cell lung carcinoma and NSCLC is alluded as non-small cell lung carcinoma. Overall NSCLC offers over 84% of generally speaking cellular breakdowns in the lungs and SCLC contributes almost 16% of cellular breakdowns in the lungs. The development pace of growth is for the most part sorted into grades design. In Grade 1: cells which are in lungs are seems to be ordinary cells. In grade 1 cell is normally sluggish spreading. Where as in Grade 2: cells are seeming to be in strange position and these cells are extremely quick spreading in human body. Grade 2 cells spreading multiple times than grade 1 cell. Another two grades are additionally present and these spreading strangely in human body contrast with different grades. The ethical quality pace of cellular breakdown in the lungs is lower and it never been higher. World cancer research fund (WCRF) gave first position to the Cancer in the lungs or respiratory framework cancer. This exploration distributed consequence of people how they are influencing by disease. In the year 2017 almost 22,65,600 new cases are recognized in human bodies. It is nearly being 20% of all diseases in around the world. It shows more danger to the world. In the year 2017 absolute 17,65,600 men were analysed around the world. This is almost 21% of all tumor in male populace. It is most perilous sickness in male. While 5,00,000 ladies were analysed in that year which is 10% of ladies' populace on the planet. As per WCRF it is third most hazardous illness in ladies. Finding or handling the wild illness is dependably a troublesome cycle however by utilizing trend setting innovations it is not difficult to identify cells and reduction is likewise happening. The expense of trend setting innovations utilization is more costly.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. RELATED WORK

Table 1: Comparisons of various techniques and methodology utilized in existing system

Ref.	Year	Datasets	Preprocessing	Methods	Results
[1]	2023	LIDC-IDRI & LUNA-16 dataset	Utilizes an improved Gabor filter	AlexNet SVM-classification model	97.98% of accuracy, 98.84% of sensitivity, 97.47% of specificity, 97.53% of precision, and 97.70% of F1
[2]	2023	chest X-ray pictures	Utilizes Histogram Equalization	OCNN-SVM approach	98% of accuracy, 98.34% of recall, 98.49% of precision, and 98.76% of F1
[3]	2021	CT Scan pictures	Uses Median Filtering. Histogram Equalization and morphological operation	ANN Classifier	95.6% of accuracy, 93.1% of sensitivity, 100% of specificity
[4]	2020	Lung Cancer CT Scan pictures	GLCM features on pictures	SVM classification algorithm	83.33% of accuracy
[5]	2020	Dataset of Chest X-Ray and Lung cancer (LIDC-IDRI)	Morphological segmentation and watershed segmentation are utilizes for automated nodule segmentation	MAN is used to classify chest X-Rays images and EFT is used to classify the lung CT pictures.	DL accuracy is 96% for X-Ray images while the exactness is 97.27% for CT pictures
[6]	2020	Database of cancer imaging archive (CIA) dataset	Multilevel brightness-preserving approach	Improved deep neural network and ensemble classifier.	The proposed system recognized the cancer with maximum exactness.
[7]	2019	Image was collected from Cancer imaging Archive (CIA) dataset	The noise is removed using weighted mean histogram equalization approach. In addition, improved profuse clustering method (IPCT) is applied for segmenting the affected region.	Deep learning instantaneously trained neural network (DITNN) is utilized.	exactness 98.42% & minimum classification error of 0.038.
[8]	2019	Database obtained from UCI repository	Picture securing, pre-handling, binarization, thresholding, division, feature extraction are applied.	The fuzzy neural system is used to test the neural system with machine learning approaches.	Exactness 96.67 %.
[9]	2019	Database of Lung Image Dataset Consortium pictures collection (LIDC-IDRI)	This step involves of segmentation is followed by normalization and zero centering.	A number of classifiers like XGBoost and Random-Forest are utilized.	Exactness 84%
[10]	2018	Dataset of Lung Image Dataset Consortium (LIDC)	Median filter and Gaussian filter are applied on the CT images.	Watershed segmentation for detection and SVM for classification of nodule as Malignant or benign	Exactness 92%.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

[11]	2018	Dataset obtained from UCI repository	Pre-processing for data cleaning is applied	A number of classifiers including: MLP, Neural Network, Decision Tree, Naïve Bayes, Gradient Boosted Tree, and SVM are assessed	Exactness 90%.
[12]	2018	Database of Lung Image Database Consortium (LIDC)	Converting to grayscale image, applying denoising methods such as median blur, Gaussian blur, and bilateral blur, then applying thresholding methods for converting the grayscale image into a binary image.	K-nearest neighbors classifier, support vector machine classifier, decision tree classifier, multinomial naïve Bayes classifier, stochastic gradient descent classifier, random forest classifier, and multi-layer perceptron (MLP) classifier are applied	Exactness 88.55%.
[13]	2018	Used UCI machine learning dataset	Image enhancement and segmentation has been done. Image scaling, color space transformation and contrast enhancement	multi-class SVM classifier	Precision of 97% for cancer recognition and 87% for cancer prediction
[14]	2016	Used CT-scan Images are from VIA and ELCAP dataset	Image enhancement using Gabor filter is applied	Region Growing, Marker Controlled Watershed, and Marker Controlled Watershed with Masking are applied.	Watershed with masking method has highest exactness and robustness.

III. PROPOSED ALGORITHM

Our proposed system architecture provides lung-cancer detection system which utilized CT-scan pictures to recognize lung pictures have Cancerous or non- Cancerous. Modules for the proposed architecture are discussed below in the Fig. 1.

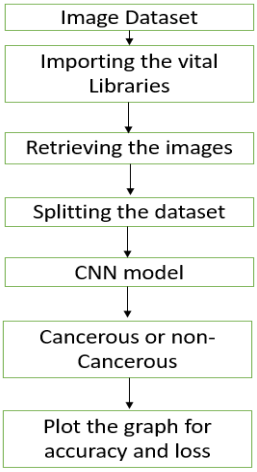


Fig. 1 Proposed system architecture for CT scan image/ dataset for lung cancer recognition

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

a. Image Dataset

CT lung cancers prediction using CNN information set given in model folder. The dataset consists of round 550 CT lung cancers pics. In Fig.2 Displays the CT scan image of lung cancer detection.

Fig 2. Displays the Lung CT scan image as dataset sample

(b) Importing the necessary libraries

We are utilizing the Python language for the system architecture. Firstly, we import the important libraries which include Keras for constructing the main model, sklearn for splitting the training and take a look at data, PIL for changing the pictures into array of numbers and other libraries consisting of pandas, NumPy, matplotlib and TensorFlow**.**

(c) Retrieving the CT images

We will retrieve the pictures and their labels. Then resize the pictures to (224,224) as all pictures have to have equal size for reputation. Then convert the photographs into NumPy array.

(d) Splitting the dataset- Splits the data into 3 categories

- Training (utilizes for generation of model)
- Private test (utilizes for generation of model)
- Public test (utilizes for evaluating the model)
- Split the database into trained and test. Our system finds 80% train data and 20% test facts.

(e) Conventional Neural Network

The targets behind the first module of the path four are:

- To recognize the convolution operation
- To understand the pooling operation
- Remembering the vocabulary utilized in convolutional neural networks (padding, stride, filter, and many others.)
- Building a convolutional neural network for multi-magnificence category in pictures.

Convolution Network

In a convolutional network (ConvNet), there are essentially 3 types of layers:

a) Convolution layer
b) Pooling layer
c) Fully connected layer

(f) Building the model

For building the we will utilize sequential version from Keras library. Then we will add the layers to make convolutional neural community. In the primary 2 Conv2D layers we've got utilized 32 filters and the kernel length is (5,5). In the MaxPool2D layer we've saved pool size (2,2) because of this it will pick the most of each 2 x 2 place of the pictures. By doing this dimension of the photo will reduce through factor of 2. In dropout layer we've got kept dropout charge = 0.25 which means 25% of neurons are removed randomly. We apply these three layers once more with a few changes in parameters. Then we follow flatten layer to transform 2-D records to one-D vector. This layer is

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

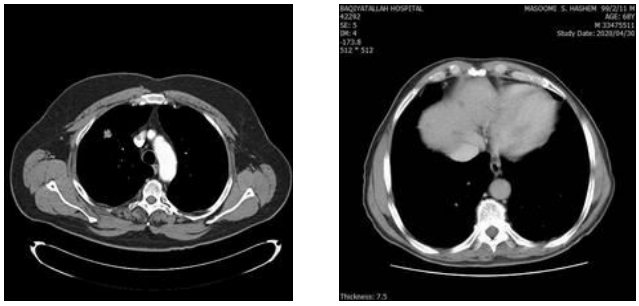
accompanied through dense layer, dropout layer and dense layer again. The ultimate dense layer outputs 2 nodes because the brain tumor or no longer. This layer uses the softmax activation feature which provides opportunity price and predicts which of the 2 options has the best chance.

(g) Apply the model and plot the graphs for accuracy and loss

We will compile the model and observe it using in shape function. The batch size may be 1. Then we will plot the graphs for accuracy and loss. The model creates during the training process comprise of pre-trained weights and values, which can be utilized for implementation of a new lung cancer detection problem.

IV. RESULT AND DISCUSSION

We have analysed pictures of CT scan Lung cancer database. We have considered about 550 various CT-Scan pictures of lungs as shown in Fig. 3 (a) and (b).



(b) (b)

Fig. 3: Lung CT Scan images (a)Cancerous image and (b) non-Cancerous image

STEP 01: Run program and then open in localhost server- Fig. 4 shows the local server snapshot. A home page is appeared.

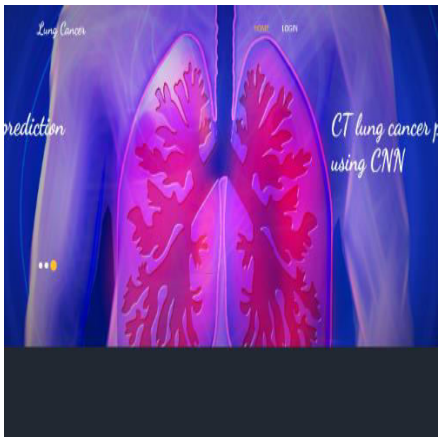


Fig. 4 Displays the screenshot of local server

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Step 2: Login with username and password- After clicking on login button, username and password is required as shown in Fig. 5.

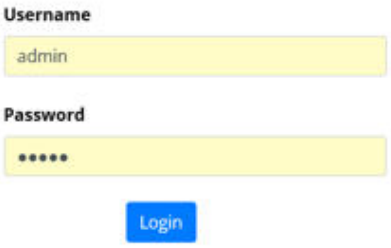


Fig. 5: Displays the login form.

Step 3: Upload an Image- Select the picture from the trained database and then click on the submit button as shown in Fig. 6.

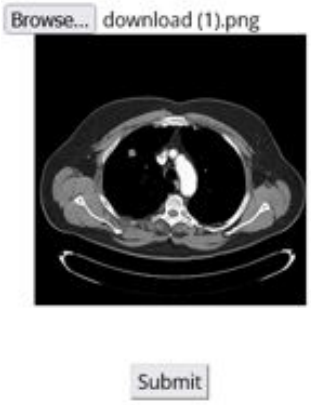


Fig. 6. Displays the browsed trained image from the database

Step 4: Prediction outcome of dataset using image processing.

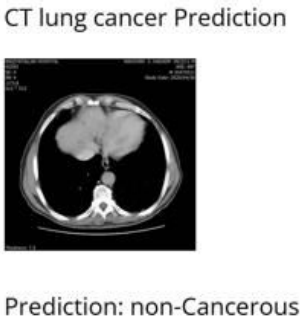


Fig. 7. Shows the of prediction of lung Cancer detection

This prediction shows the resultant as the CT pictures is cancerous or non-cancerous.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Step 5: Performance analysis of the system- Performance analysis gives the output in the form of Accuracy, precision, recall and F-measure. Our system provides 99.7% accuracy as shown in fig. 8.

Fig: 8 Shows the performance analysis of the system

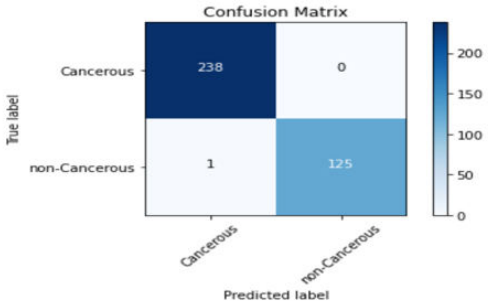


Fig. 9 Shows the confusion matrix

The confusion matrix shows the true label and predicted label of cancerous and non-cancerous images as shown in fig. 9. From the analysis, true label gives the correctly classified lung cancer images and predicted label gives the prediction of lung cancer image.

Step 6: Graphical representation

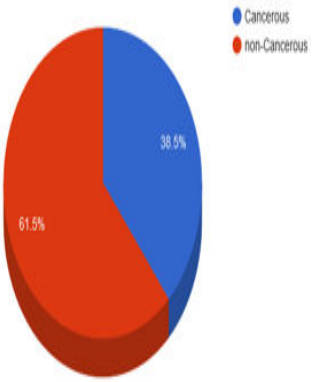


Fig 10: Shows the graph for cancerous and non- cancerous image

Graph shows the graphical representation of cancerous and non-cancerous images present in the database as shown in fig. 10. There are 38.5% images are cancerous and 61.5% images are non-cancerous in our 550 CT scanned images dataset.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

TABLE II. Comparison Table of Accuracy obtained from various frameworks

S.No.	Author	Year	Technique Used	Accuracy
1	Naseer et al. [1]	2023	AlexNet SVM-classification model	97.98%
2	Vs et al. [2]	2023	OCNN-SVM approach	98.7%
3	Shailesh et al. [3]	2021	ANN Classifier	95.6 %
4	Proposed System	2024	Conventional Neural Network	99.7%

V. CONCLUSION AND FUTURE WORK

A conventional neural network-based architecture was deployed to recognized the malignancy cell available in the CT scan lung cancer images. CT lung cancer images consist of various shapes and size of the cancerous cells presented at the input training set of the system. Our proposed system is eligible to recognize the presence and absence of cancerous tissue with accuracy about 99.7%. In the near future, the system will be trained with large number of datasets to recognize the type of lung cancer with its size and shape.

REFERENCES

[1] Naseer, Iftikhar & Akram, Sheeraz & Masood, Tehreem & Rashid, Muhammad & Jaffar, Arfan. (2023). Lung Cancer Classification Using Modified U-Net Based Lobe Segmentation and Nodule Detection. IEEE Access. PP. 1-1. 10.1109/ACCESS.2023.3285821.

[2] Vs, Sreeprada & Vedavathi, Dr. (2023). Lung Cancer Detection from X-Ray Images using Hybrid Deep Learning Technique. Procedia Computer Science. 230. 467-474. 10.1016/j.procs.2023.12.102.

[3] Shailesh, M., Bhise, S., & Khot, P.S. (2021). Early Stage Lung Cancer Diagnosis using ANN Classifier. 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), 22-27.

[4] Q. Firdaus, R. Sigit, T. Harsono and A. Anwar, "Lung Cancer Detection Based On CT-Scan Images With Detection Features Using Gray Level Co-Occurrence Matrix (GLCM) and Support Vector Machine (SVM) Methods," 2020 International Electronics Symposium (IES), IEEE, 2020, pp. 643-648, doi: 10.1109/IES50839.2020.9231663.

[5] Bhandary, A., Prabhu, G. A., Rajinikanth, V., Thanaraj, K. P., Satapathy, S. C., Robbins, D. E., ... & Raja, N. S. M. (2020). Deep-learning framework to detect lung abnormality–A study with chest X-Ray and lung CT scan images. Pattern Recognition Letters, 129, 271-278.

[6] Shakeel, P. M., Burhanuddin, M. A., & Desa, M. I. (2020). Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier. Neural Computing and Applications, 1-14.

[7] Shakeel, P. M., Burhanuddin, M. A., & Desa, M. I. (2019). Lung cancer detection from CT image using improved profuse clustering and deep learning instantaneously trained neural networks. Measurement, 145, 702-712.

[8] Reddy, U., Reddy, B., & Reddy, B. (2019). Recognition of Lung Cancer Using Machine Learning Mechanisms with Fuzzy Neural Networks. Traitement du Signal, 36(1), 87-91.

[9] Bhatia, S., Sinha, Y., & Goel, L. (2019). Lung cancer detection: A deep learning approach. In Soft Computing for Problem Solving (pp. 699-705). Springer, Singapore.

[10] Makaju, S., Prasad, P. W. C., Alsadoon, A., Singh, A. K., & Elchouemi, A. (2018). Lung cancer detection using CT scan images. Procedia Computer Science, 125, 107-114.

[11] Faisal, M. I., Bashir, S., Khan, Z. S., & Khan, F. H. (2018, December). An evaluation of machine learning classifiers and ensembles for early stage prediction of lung cancer. In 2018 3rd International Conference on Emerging Trends in Engineering, Sciences and Technology (ICEEST) (pp. 1-4). IEEE.

[12] Singh, G. A. P., & Gupta, P. K. (2019). Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans. Neural Computing and Applications, 31(10), 6863-6877,

[13] Alam, J., Alam, S., & Hossan, A. (2018, February). Multi-stage lung cancer detection and prediction using multi-class svm classifie. In 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2) (pp. 1-4). IEEE

[14] Abdillah, B., Bustamam, A., & Sarwinda, D. (2017, October). Image processing-based detection of lung cancer on CT scan images. In Journal of Physics: Conference Series (Vol. 893, No. 1, p. 012063). IOP Publishing.

.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details