INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

ISSN: 2057-5688

Detection of lung cancer using CNN

CH.TEJASREE

Assistant Professor Department Electronics & Communication Engineering, kamala institute of technology and science huzurabad (TS).India.
EMAIL: chtejasreechityala@gmail.com

ABSTRACT

Lung diseases are actually diseases affecting the lungs that impair the respiratory mechanism. Lung cancer is one of the leading causes of death in men worldwide. Early detection can improve people's chances of survival. If the disease is diagnosed in time, the average survival rate of people with lung cancer increases from 14% to 49%. Although computed tomography (CT) is much more effective than X-rays, a thorough diagnosis involves several imaging approaches to support each other. A deep neural network for lung cancer detection from CT images was developed and evaluated. To classify lung imaging as normal or malignant, a densely connected complex neural network (DesceNt). A dataset of 201 lung images was used, of which 85% were used for training and 15% were used for testing and classification. Experimental results show that the proposed method has an accuracy of 90.85%.

KEYWORDS: CNN, Artificial Intelligence (AI), LUNG CANCER

I .INTRODUCTION

Lung cancer is identified as one of the most serious causes of death in the world. It is one of the most malignant tumors that can affect human life. Its mortality rate ranks among all tumor deaths and is also the leading cause of cancer death in both men and women. There are nearly 1.8 million new cases of lung cancer each year (13% of all cancers), and 1.6 million deaths worldwide (19.4% of all cancers). Lung cancer is the proliferation of abnormal cells that expand and develop into a tumor. Among other forms of cancer, lung cancer mortality is the highest. Tobacco smoke causes about 85% of lung cancer in men and 75% in women. Lung cancer is one of the most devastating diseases in the developing world, with a mortality rate of 19.4%. Lung cancer is one of the deadliest cancers in the world, with the lowest rate of success after diagnosis, with the number of victims steadily increasing each year. The benefit of fuzzy logic in pre-predictions leads to result-oriented analysis. Lung cancer survival after diagnosis is directly related to its progression. However, individuals have higher success rates than those found in early life. Cancer cells are distributed in the blood of the lungs, the lymph fluid that coats the lung tissues. Lymph travels through lymphatic vessels and is discharged through lymph nodes in the lungs and chest area. The examination and treatment of lung diseases has become one of the biggest obstacles facing humanity in recent years. A reliable early diagnosis of a tumor will promote its survival for a large number of lives around the world. This paper presents a method using complex neural networks (CNN) to identify lung tumors as malignant / benign.

II. LITERATURE SURVEY

2.1 An Automatic Detection System of Lung Nodule Based on Multigroup Patch-Based Deep Learning Network

AUTHORS: Jiang, H., Qian, W., Gao, M., Li, Y.

The highly efficient detection of lung nodules contributes significantly to the assessment of lung cancer risk. Rapid identification of the exact locations of pulmonary nodules is an important and difficult task. Extensive work has been done

ISSN: 2057-5688

by researchers in this field for about two decades. However, previous computer-aided detection (CADe) schemes were mostly complex and time-consuming because they might require more image processing modules, such as image conversion. CT image, lung nodule segmentation and feature extraction image processing, to build a complete CADe. System. It is very difficult for these modes to process and analyze the huge data as the medical images continuously multiply. Also, some modern deep learning programs can be strict with the database standard. This study proposes an efficient lung nodule detection scheme based on multi-group patches excised from lung images, enhanced by the Franti filter. By combining two groups of images, a four-channel agglomeration neural network model is designed to learn the knowledge of radiologists to detect nodules at four levels. This CADe scheme can achieve a sensitivity of 80.06% with 4.7 false positives per scan and a sensitivity of 94% with 15.1 false positives per scan. The results demonstrate that the multi-group patch-based learning system is effective in improving lung nodule detection performance and significantly reducing false positives under the huge amount of imaging data.

2.2 Identifying Lung Cancer Using Image Processing Techniques

AUTHORS: Disha Sharma, Gagandeep Jindal

The computer-aided automatic diagnosis (CAD) system is proposed in this paper to detect lung cancer from the analysis of CT images. To create a successful computer-aided diagnostic system, several problems must be resolved. In recent years, image processing has been widely used in a number of medical fields to improve the earlier stages of detection and treatment, where time is crucial for detection. Disease in patients as soon as possible, especially in various cancerous tumors such as lung cancer, breast cancer. This system typically first segments the area of interest (lungs) and then scans the resulting area separately to detect nodules to diagnose disease. Initially, basic image processing techniques such as erosion, median filtering, dilation, monitoring, and lung border extraction were applied to CT images to detect lung regions. Then, a segmentation algorithm is applied to detect cancerous nodules from extracted lung images. After segmentation, a rule-based technique is applied to classify cancerous nodules. Finally, a set of diagnostic rules is generated from the extracted features. To test the proposed technique, CT images were collected from the NIH / NCI Lung Imaging Database Consortium (LIDC) dataset, providing the possibility to perform the proposed study. DICOM [9] (Digital Imaging and Communication in Medicine) has become a standard for medical imaging. Its goal is to standardize digital medical images and data for easy access and sharing. Many commercial viewers support the DICOM image format and can read metadata. The main objective of the project is to develop a CAD (computer aided diagnosis) system to find early lung cancer nodules using CT images of the lungs and classify the nodules as benign or malignant.

2.3 Lung Nodule diagnosis from CT images Based on Ensemble Learning

AUTHORS: Farzad Vasheghani Farahani

ABSTRACT: Early detection of cancer is the most promising way to increase a patient's chances of survival. This paper presents a computational classification method using computer tomography (CT) images of the

lungs, based on an ensemble of three classifiers, including MLP, ANN, and SVM. In this study, the entire lung is first segmented from a CT image, and specific features such as roundness, roundness, compactness, Ellipticity, and eccentricity are calculated from the segmented image. These morphological features are used in the classification process in such a way that each classifier makes its own decisions. Finally, the majority decision method is used to combine the decisions of this ensemble system. The performance of this system was assessed using 60 CT scans collected by the Lung Imaging Database Consortium (LIDC), and the results show excellent improvements in the diagnosis of lung nodules.

2.4 Lung cancer prediction using machine learning and advanced imaging techniques

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

ISSN: 2057-5688

AUTHORS: Fergus Gleesona

ABSTRACT: Machine learning-based lung cancer prediction models have been proposed to help clinicians treat uncertain lung nodules detected accidentally or by screening. Such systems may be able to reduce nodule classification variability, improve decision making, and ultimately reduce the number of benign nodules that are unnecessarily tracked or processed. This article outlines the major lung cancer prediction approaches proposed so far and focuses on some of their relative strengths and weaknesses. It describes some of the challenges in developing and validating such technologies and outlines the path to clinical application.

2.5 Detection of Cancer in lung with K-NN classification using genetic algorithm

AUTHORS: P. Bhuvaneswari

ABSTRACT: This article focuses on early detection of lung cancer. A nonparametric K-nearest neighbor (GKNN) genetic algorithm has been proposed for detection. This optimization algorithm allows doctors to identify early nodules present on CT lung images, or lung cancer. Manual interpretation of lung cancer CT images is time consuming and very important, so to overcome this difficulty, the genetic algorithm can be combined with the K-nearest neighbor (K-NN) algorithm to accelerate the cancer. Sort effectively. Image. Implementations based on the MATLAB Image Processing Toolbox are performed on CT lung images and classification of these images is performed. Key performance indicators such as classification rate and false positive rate are analyzed. In the traditional K-NN algorithm, the distance between all test and training patterns is calculated first, and the higher distance K-nearest neighbors is used for classification. This proposed method uses a genetic algorithm to select a sample size of A (50-100) per iteration, achieving 90% classification accuracy as a fit. The highest accuracy is recorded every time.

III.EXISTING SYSTEM

Cancer cells spread from the lungs, the lymph that lines the lung tissue, into the blood. Lymph vessels enter the lymph vessels and are drained through the lymph

ISSN: 2057-5688

nodes in the lungs and chest. Research and treatment of lung disease has become one of the greatest obstacles facing humankind in recent years. Early diagnosis of tumors ensures the survival of many people around the world. Various laboratory and clinical stages are used, including chemical therapies, targeted therapies, and even radiation therapy to destroy or stop the replication of malignant cells. All of these cancer identification and detection methods are long, costly and painful for the patient. Therefore, to overcome all these problems, we used appropriate machine learning techniques to process these medical images consisting of CT scan images.

Disadvantages of Existing System:

Detecting and diagnosing cancer is time-consuming, costly, and painful for the patient.

IV PROPOSED SYSTEM:

Convolutional neural networks are designed to minimize the number of parameters in the image classification network architecture. A convolutional neural network consists of many layers organized according to its characteristics and functions. The architecture of Convent is very similar to that of the human brain. Data expansion is a process that increases the amount and complexity of data. Collect new data instead of converting the data that is already available. Data expansion is a deep learning phase. This data extension works because deep learning requires large amounts of data and, in some cases, cannot capture thousands or millions of images. This allows you to maximize the dataset size and add uncertainty within the dataset.

Advantages of Proposed System:

Image capabilities to achieve the highest accuracy.

To design and achieve better classification accuracy on CNN.

IV.SYSTEM ARCHITECTURE

ISSN: 2057-5688

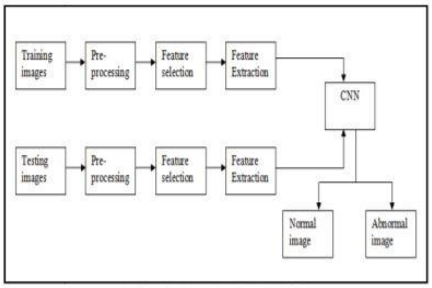
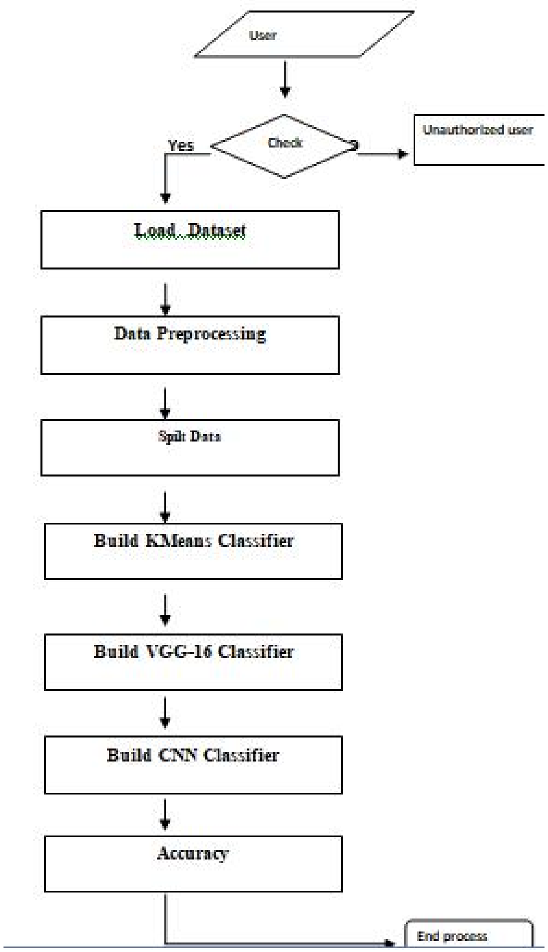


Fig1: Architecture of face expression recognition system.



ALGORITHMS:

CNN:

This article describes the concept of automating government services using artificial intelligence techniques such as deep learning algorithms called convolutional neural networks (CNNs). Governments launch new systems on the Internet. People can read news and news about such systems, people can write opinions about such systems, and those opinions are better for the government. It helps make decisions. In order to automatically recognize public opinion about a system, we need software that can easily understand whether the opinions people are writing are positive or negative, like the human brain.

To build such automated opinion recognition, the author proposes to build a CNN model that can function like the human brain. This CNN model can be generated for any service and designed to act like automated decision making without

ISSN: 2057-5688

human intervention. To propose this approach, the authors authored multiple that one model could recognize or recognize human handwritten numbers and the second model could recognize emotions from textual sentences that humans could give about the government system. We have already explained the concept of implementing the model. In the extended model, we added another model that can detect mood from a person's face image. People's facial expressions can express emotions more than words and sentences. Therefore, our extended work can predict mood from images of people's faces.

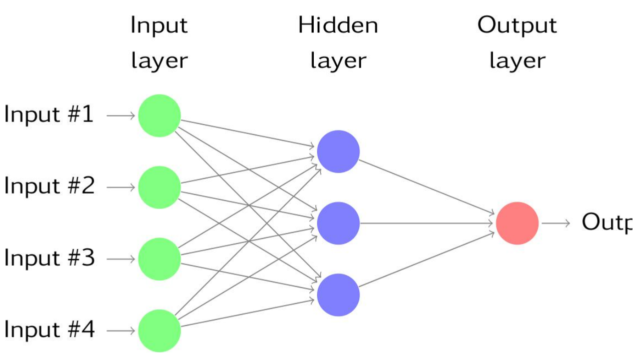
To show how to create a convolutional neural network-based image classifier, we will create a 6-layer neural network that identifies and separates one image from another. This network to be built is a very small network that can be executed by the CPU. Traditional neural networks, which are very good at image classification, have more parameters and will take a lot of time to train on a regular CPU. However, our goal is to show how to use TENSORFLOW to build a real convolutional neural network.

Neural networks are essentially mathematical models used to solve optimization problems. They consist of neurons, which are the basic computing units of neural networks. The neuron receives an input (e.g. x), performs some calculations on that input (e.g. multiplies the variable w, adds another variable b), and values (e.g. z = wx + b). To generate. This value is passed to a non-linear function called the activation function (f) to generate the final output (activation) of the neuron. There are many types of activation functions. One of the most popular activation features is Sigmoid. Neurons that use the sigmoid function as the activation function are called sigmoid neurons. Neurons are named according to the activation function, and there are many types, such as KELU and TanH.

When you stack neurons in a row, it's called a layer. It is the next component of the neural network. See image below with

ISSN: 2057-5688

layers.



Multiple layers work together to get the best layer to predict the image class. This process continues until there are no improvements.

MODULES:

•Load Dataset

•Data Preprocessing

•Spilt Data

•Build KMeans Classifier

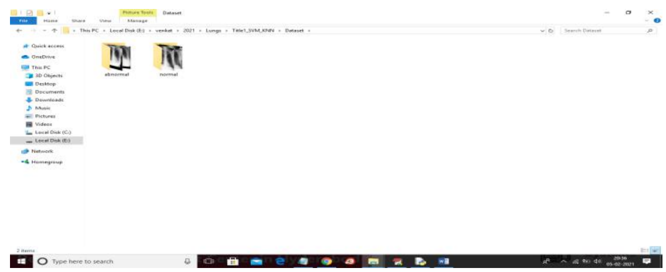
•Build VGG-16 Classifier

•Build CNN Classifier

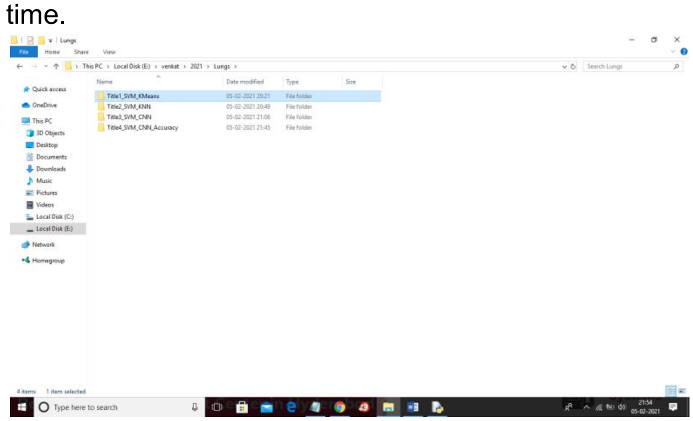
•Accuracy

RESULT:

Lung Cancer Nodule Classification and K-means Accuracy Comparison to Monitor Patient Health Using Neural Network Topology Using SVM Algorithm This project uses a CT scan lung cancer nodule dataset to predict patient health using SVM and KMeans algorithms and compare the prediction accuracy between them. To implement this project, we will use a dataset containing lung cancer images and the following screen showing the details of the dataset. This dataset is stored in the dataset folder.



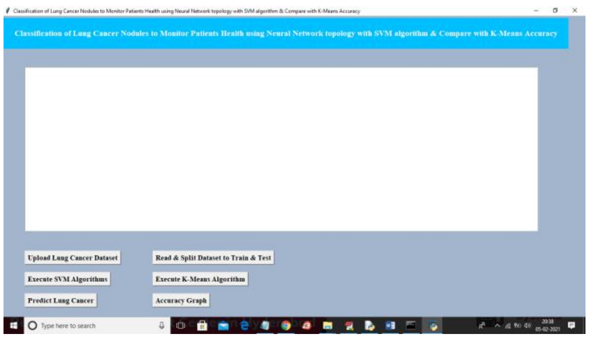
There are two types of images on the screen above the dataset. B. Normal and abnormal, then SVM and KMEANS are trained on the above dataset and when you upload a new image, SVM predicts whether the new image is normal or abnormal. To implement the four titles, you can create four folders with separate algorithms for each title and run them one at a time.



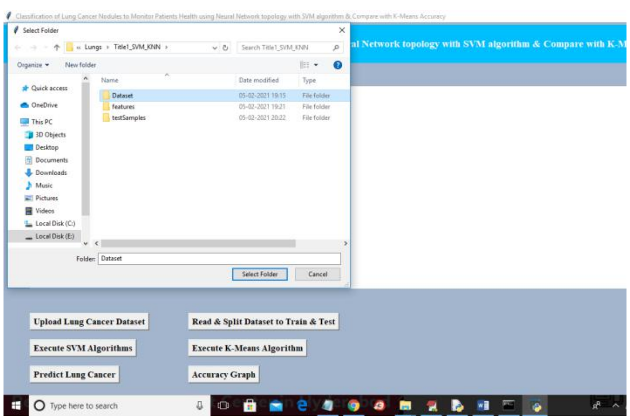
SCREENS SHOT

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE
ISSN: 2057-5688

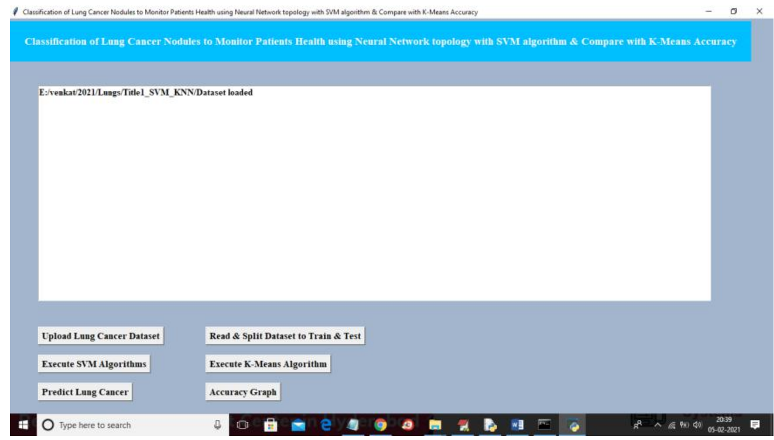
To run the project, double-click the run.bat file in the Title1_SVM_MEans folder to display the following screen.



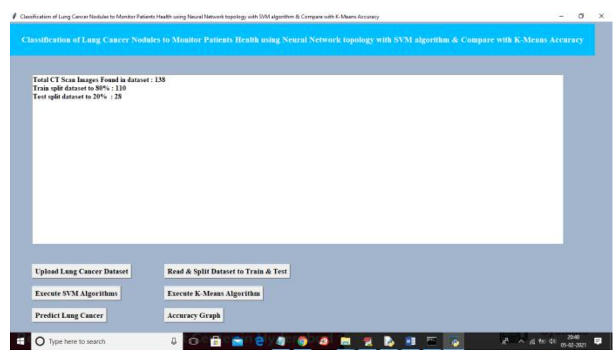
On the screen above, click the Upload Lung Cancer Dataset button, then upload the dataset folder.



On the screen above, select the "Datasets" folder to upload. Then click the Select Folder button to load the dataset and go to screen below.



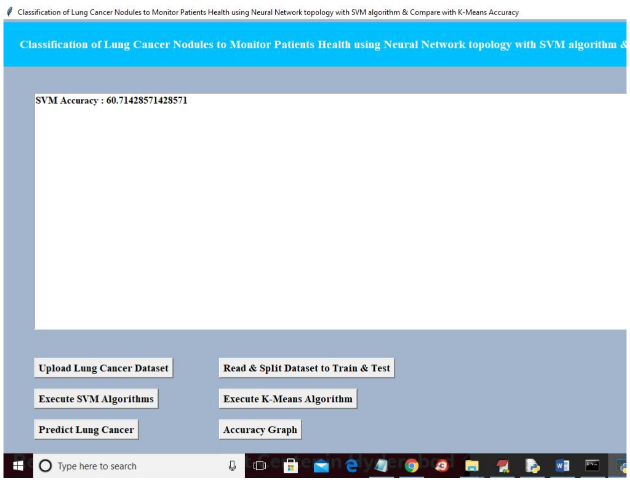
To run the project, double-click the run.bat file in the Title1_SVM_MEans folder to display the following screen.



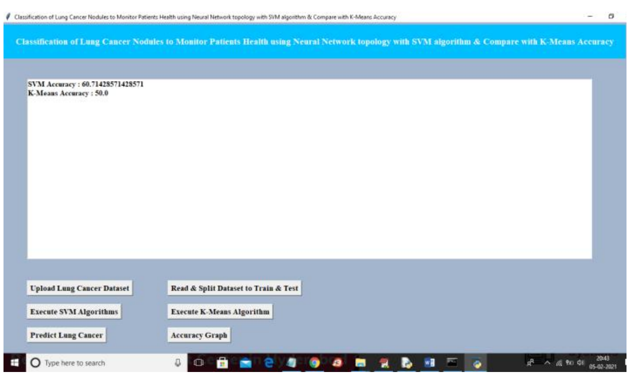
On the screen above, you can see that the dataset contains a total of 138 images, and the application is using 110 images for training and 28 images for testing. Now the data is ready and you can click the Run SVM Algorithm button to run the SVM on the loaded and inaccurate datasets.

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

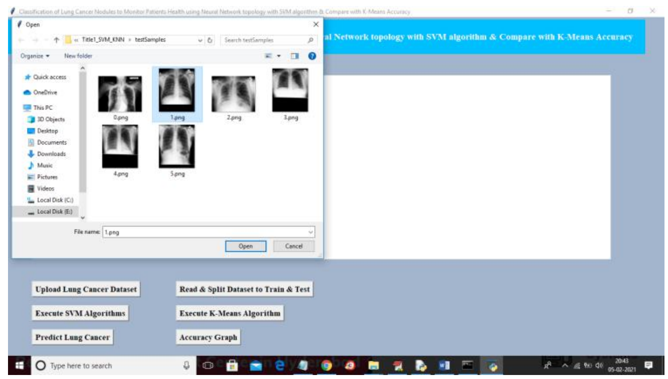
ISSN: 2057-5688



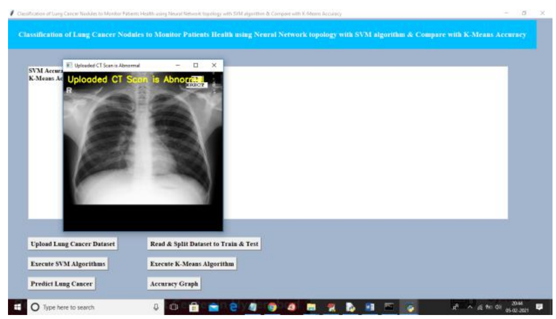
In the screen above, the SVM accuracy is 60%. Then click the Run K-Means Algorithm button to run the KMEANS algorithm on the loaded dataset and go to the screen
below.



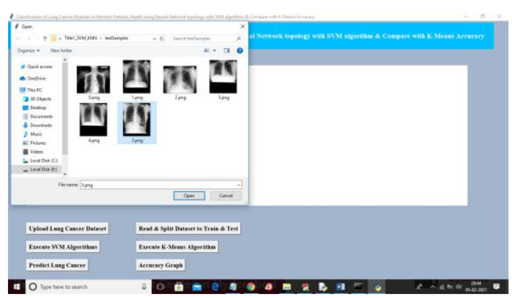
In the screen above, the SVM accuracy is 60%. Then click the Run K-Means Algorithm button to run the KMEANS algorithm on the loaded dataset and go to the screen below.



On the above screen, select the "1.png" file and upload it. Then click the "Open" button to get the following
results



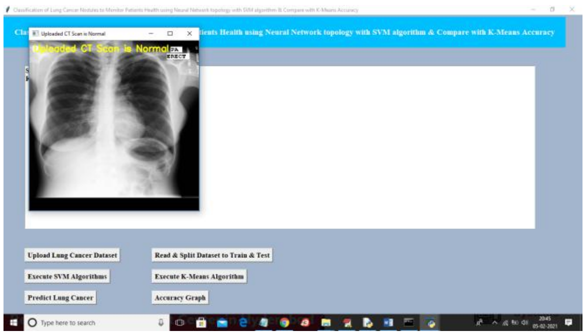
In above screen uploaded image predicted as Abnormal and now test with another image



In above screen uploading '5.png' and below is the result

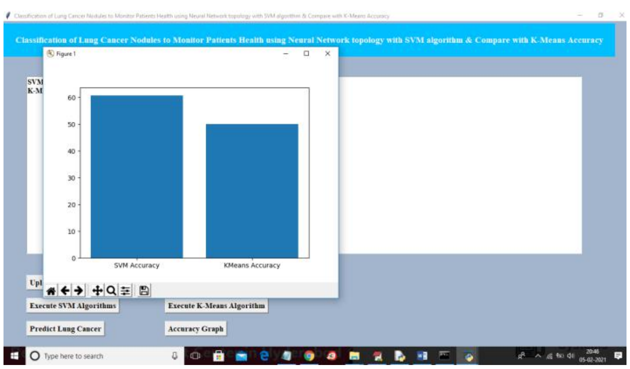
INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

ISSN: 2057-5688



The image above is expected as usual. Similarly, you can upload any image to make a prediction and click the Precision

Graph button to get the graph below.



In the screen above, the x-axis represents the names of the algorithms and the y-axis represents the accuracy of those algorithms. From the graph above, we can conclude that SVM is superior to KMEANS in prediction.

TITLE 2

A hybrid model that helps clinicians manage disease through the SVM algorithm and compares error rates with the k-nearest neighbor method by classifying lung nodules with continuous, time-varying data.

In this project, we will train the SVM and ANN algorithms using the same dataset as above and calculate the error rate between these two algorithms. This error rate indicates the rate of misclassification. For example, if the application correctly predicted 18 out of 20 records, the error rate would be (1- (18/20) = 0.1.

VII. CONCLUSION

The biggest advantage of deep learning over other machine learning algorithms is the ability to perform feature engineering itself. This allows you to examine the data for relevant properties and integrate them for faster learning. Take advantage of the spatial coherence of the input. Image training and testing takes place where the image is preprocessed and feature selection and feature extraction is performed from the image. When the training and testing parts are completed successfully, the CNN algorithm classifies the input lung image as normal or abnormal and displays the output. Therefore, the Deep CNN network is used to classify lung images and detect cancer.

IX. REFERENCES

1. Jiang, H., Qian, W., Gao, M., Li, Y. " An automatic detection system of lung nodule

based on multigroup patch-based deep learning network" IEEE Journal of Biomedical and Health Informatics,22(4):1227-1237.

2. Disha Sharma, Gagandeep Jindal, International Conference on Lung Cancer Identification, Computational and Artificial Intelligence Using Image Processing Techniques, pp. 116-120, 2011.

3. Farzad Vasheghani Farahani "Diagnosis of lung nodules from CT images based on ensemble learning". IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology, 2015.

4. Atsushi Teramoto of Ell. "Automatic classification of benign and malignant cells from lung cytology images using deep convolutional neural networks" Computer Science in Medicine unlocked 16, 2019, 100205elsvier.

5. Xin-Yu Jin, Yu-Chen Zhang, Qi-Liang Jin "Detection of lung nodules based on CT images using convolutional neural networks" International symposium on computational intelligence and design. 2016

6. Ryota Shimizu, Shusuke Yanagawa, Yasutaka Monde, Hiroki Yamagishi, Mototsugu Hamad, Toru Shimizu and Tadahiro Kuroda "Deep Learning Application Trial to Lung Cancer Diagnosis for Medical Sensor Systems", International Symposium on Com puter and Communication, 2016.

7. Po-Whei Huang, Phen-Lan Lin, Cheng-Hsiung Lee, CH Kuo, "Lung Nodule Classification System for CT Images Based on Fractional Brownian Motion Model", IEEE International Conference on System Science and Engineering, July 2016.

8. Vaishali C. Patil, Shrinivas R. Dhotre, "Detection of Lung Cancer from Images from Computer Tomography Scans", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 3, No. 5, Ausgabe 7, July 2016.

9. Milton Felix, Marcelo Oliveira, Aydano Machado, Jose Raniery, "Use of 3D Textures and Margin Sharpness Features in Classification of Small Lung Nodules", 29th SIBGRAPI Conference on Graphics, Patterns and Images, 2016.

10. Rotem Golan, Christian Jacob, Jorge Denzinger, "Detection of Lung Nodules on CT Images Using Deep Convolutional

Neural Networks", International Joint Conference on Neural Networks, 2017.

11. Sri Widodo, Ratnasari Nur Rohmah, Bana Handaga, "Classification of Pulmonary Nodules and Arteries in Computer Tomography Scanned Images Using Principal Analysis", 2nd International Conference on Information Technology, Information Systems, and Electrical Engineering, 2017.

12. Ravindranath K, K Somashekar, "Early Detection of Lung Cancer by Nodule Extraction-Investigation", International Conference on Electrical, Electronic, Communication, Computational and Optimization Techniques, 2017.