ISSN 2515-8260 Volume 07, Issue 09, 2020

DEEP BELIEF NETWORK (DBN) CLASSIFICATION FOR LUNG CANCER PREDICTION USING KNN CLASSIFIER

V. Shanthi[1], G. Sridevi[2], R. Charanya[3], J. Josphin Mary[4]

[1]Professor, Faculties of Humanities Science, Meenakashi Academy of Higher Education Research Chennai, TamilNadu, India

[2,3,4]Assistant Professor, Faculties of Humanities Science, Meenakashi Academy of Higher Education Research Chennai, TamilNadu, India

E-mail: principal@maherfhs.ac.in[1], sridevig@maherfhs.ac.in[2], charanyar@maherfhs.ac.in[3], josphinmaryj@maherfhs.ac.in[4]

Abstract:

Lung cancer is the one of the deadliest disease in the world. Cancer is a disease in which cells develop out of balance in the body. Lung cancer is the primary cause of cancer mortality among both men and women in the United States, and the second most diagnosed disease. After decades of growth, lung cancer levels decrease nationwide, with fewer cigarettes being consumed. The most frequent source of lung cancer is cigarette-smoking. Lung cancer may also be induced by the usage of certain kinds of cigarettes. In this paper we propose a classification of predictive analysis through Deep Belief Network and K-Nearest neighbor (KNN) classifier.

Keywords: Deep belief Network (DBN), cigarettes, nationwide, KNN classifier.

Introduction:

Lung cancer is the most frequent uninformed cancer which causes late-health care death. Currently, CT is used to assist doctors in early-stage diagnosis of lung cancer. Sometimes, the treatment of lung cancer identification relies on the physicians' expertise, which may miss some people and trigger certain complications. In several areas of medical imaging deep learning has been shown to be a common and powerful tool. In this paper, [1] three separate forms of deep neural networks for calcification of lung cancer (for example, CNN, DNN and SAE) are discussed.

Deep learning is seen in pattern recognition and classification, as a popular and powerful method. However, in the area of diagnostic medical imaging there are not many highly structured applications, since large datasets are not always available for medical imagery. The feasibility of using Lung Picture Database Consortium (LIDC) cases for deep learning algorithms for diagnosing pulmonary cancer is explored in this report. Nodules were broken according to the marks given by the radiologists on any computed topographical (CT) slice. We have collected

174412 specimens with 52 by 52 pixels each and the relevant truth files after sampling and rotating [2].

One of the world's leading deaths is lung cancer. For the differential diagnosis of the pulmonary cancer, it is currently important to have an exact classification of cancers (adenocarcinoma, squamous cell carcinomas and small cell carcinoma). [3] However, testing performance and reliability are difficult to improve. In this research we have created an artificial classification scheme for lung cancers, a deep neural network (DCNN) provided on the microscopic image and an important tool for profound learning.

Early diagnosis of lung cancer would lead to a major reduction in the incidence of lung cancer mortality, comprising some 17% of the estimated deaths associated with cancer. Radiologists undergo most cases of initial appraisal on a regular basis. CAD [4] devices will have a second outlook for radiologists and make the whole cycle much quicker. We offer a CAD program which uses automatically extracted deep encoder functions to classify the lung nodules as malignant or benign. The introduction of lung cancer programs will produce an immense number of chest CT scans in order to be able to rely on a surveillance plan, and radiologists could view them in the near future. The function of the screen-detected nodules is highly dependent on the nodule[9] size and nodule form, according to current guidelines. In this paper, [5]we present Dee learning method focused on multifunctional networks, which identify all nodule forms related to nodule operations automatically.

Through the use of deep learning methods, we attempted to simplify the pipeline of image processing for traditional CAD [6]. We introduced models of a deep-faith network and innovative neural networks, in particular in the context of a nodule classification in computer tomography images. Two conventional methods of functional computing measures have been applied as a reference.

In order to provide a segmentation of the pulmonary CT-picture, [7] the U-net architecture, one of the most used for profound learning, is used. The architecture consists of a mediated method of collecting knowledge at a high level and a symmetrically growing way of acquiring information.

Methods and Materials:

Deep Learning is a branch of machine learning, a subset of artificial intelligence, on the other hand. Artificial intelligence is a general term for techniques that allow computers to emulate the actions of human beings. Deep Learning, on the other hand, is just a form of machine learning inspired by a human brain structure. Machine learning is a group of data-trained, all-possible algorithms. Deep learning algorithms are used for constantly analyzing knowledge using a logical framework to create identical assumptions to human beings. Deep learning employs a multi-layered algorithm framework called neural networks for this purpose.

ISSN 2515-8260 Volume 07, Issue 09, 2020

Pre-processing:

Pre-processing applies to our data's transformations before passing it to the algorithm. Data Preprocessing is a technique used to transform raw data to a clean collection of data. This is, as the data is obtained from different outlets, it is extracted in raw format and cannot be analyzed. In this process, we use Deep Belief Network (DBN) to train and validate the images of lung disease in dataset directory.

Dataset Exploration:

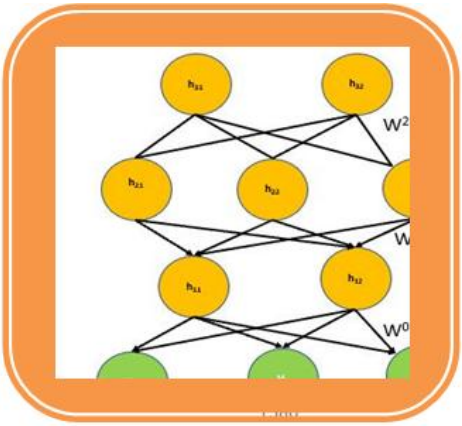
The dataset is designed such that various research approaches can be used to analyze the patterns in the usage of contrast and patient age image details in CT. The basic concept is to define the picture textures, statistics and features that closely correlate to such characteristics and to build easy resources for classifying such images [8] automatically, when misclassified (or for identifying outliers that may be suspect, defective or improperly calibrated).

Deep Belief Network (DBN) for Features Extraction:

For machine learning a Deep-Belief Network (DBN) is a generative graphical model or alternatively deep-neural networking groups consisting of several layers of latent variables (the "hidden units").The two top levels of DBN, which are associative memories, are symmetrically connected. The relations to all the lower layers are guided with the arrows points to the layer nearest to the data. In lower layers acyclic relations have been guided to convert Associative Memory to observed variables.

The input data is obtained in the lowest layer or the observable units. Integer or actual input details should be. The intra layer connections like the RBM are not provided. Hidden units are software which gathers correlations in the data. The symmetrical weight matrix W connects two layers. In each layer each device in each next layer is connected to another level.

A Greedy learning algorithm is used to pre-train DBN. Greedy learning algorithm uses the top-down generative weights layer by layer strategy. Such weights specify whether factors in one layer are calculated by the above layer factors. In DBN we analyze Gibbs in multiple measures on the top two secret layers. This stage basically draws a sample of the RBM identified by the two top clocked layers. The Systematic architecture has been shown in fig (1.1).A basic bottom-up transfer can be obtained from the values of latent variables in any row. The preprocessed input images are extracted by the features of deep belief network (DBN)

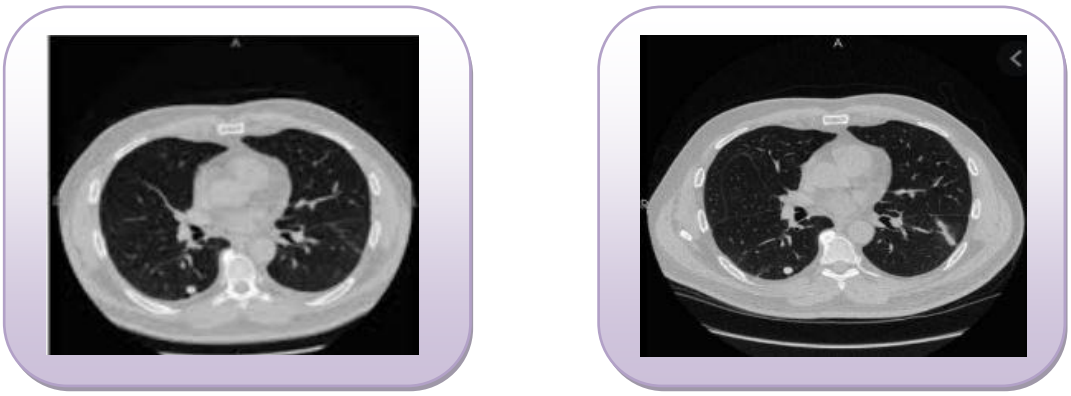


Classification of K-Nearest Neighbor (KNN):

K-nearest neighbor (KNN) is an algorithm for both classification and regression estimation problems that is tracked in ML. However, it is used primarily for statistical problems of classification in industry**.** In this Study the lung CT images are classified by KNN classifier for better prediction and analysis. The algorithm of K nearest neighbors (KNN) uses 'feature similarity' to estimation the values of new data points, which also implies that a value is given to the new data point depending on how well it fits the points in the training collection. The distance between the test data and each training row shall be determined using one specific method, namely Euclidean, Manhattan and Hamming. Euclidean is the most commonly known method of estimating size.

Result and discussion:

By the Pre-processing and feature Extraction of lung image through deep Belief Network and KNN classifier is Comes out with better accuracy and result. Sample predicted images has been shown in fig (1.2)



(A) (B)

Fig (1.2) Sample CT images in dataset directory

The Classification of KNN algorithm is to predict the result of Accuracy, Precision and Specificity value as shown in table-1.

Table-1 Analysis report for CT images for lung cancer

State	Specificity value	Precision value	Accuracy value
Normal	0.68%	0.76	0.86
Abnormal	0.75%	0.70	0.89
Average	0.71	0.73	0.87

Conclusion:

The process of lung cancer is starts form lung. Cancer occurs as cells in the body are uncontrollable. In this study we use CT images as dataset. It contains 100 images in directory. It is split up in to 60 images in normal directory folder and 40 images in abnormal directory. Then Pre-processed the images as input and features extracted images by Deep belief Network (DBN).Then the final prediction is used by K-nearest Neighbor (KNN) to classify the prediction value of precision , specificity , Average and Accuracy. The Best accuracy value is predicted in abnormal value of CT images dataset. In future we try various neural Networks and classifier is to predict more information about lung CT images.

Reference:

1. Chen, Y., Wang, Y., Hu, F. and Wang, D., 2020. Lung Deep Thick Neural Network for the Segmentation of Robust Lung Parenchyma. IEEE Access.
2. Ort, V. and Kabelková, E., 2019, November. Better distribution of minute ventilation and lungs energy load in the neonatal inhomogenic lungs in multi-frequency ventilation: statistical simulation. In 2019 E-Health and Bioengineering Conference (EHB) (pp. 1-4). IEEE.
3. Maulana, R., Kurniawan, D. and Ichsan, M.H.H., 2020, June. Intelligent Lung Diagnostic Instruments based on body temperature and fingernail colour. In 2020 2nd World Symposium on Artificial Intelligence (WSAI) (pp. 113-117). IEEE.
4. Sousa, P., Galdran, A., Costa, P. and Campilho, A., 2019, April. CT-based Lung Volume Segmentation Learning from Semi-Auter Scans. In 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) (pp. 1202-1206). IEEE.
5. Sreekumar, A., Nair, K.R., Sudheer, S., Nayar, H.G. and Nair, J.J., 2020, July. Detection of malignant lung nodule by deep analysis. In 2020 International Conference on Communication and Signal Processing (ICCSP) (pp. 0209-0212). IEEE.
6. Nitta, J., Nakao, M., Imanishi, K. and Matsuda, T., 2020, July. Deep Learning Lung Segmenting with Generative Adversary Nets Data Preprocessing. In 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) (pp. 1278-1281). IEEE.
7. Jafari, P., Hoover, D.A., Yaremko, B.P., Parraga, G., Samani, A. and Sadeghi-Naini, A., 2019, July. Incorporate pathological-induced heterogeneity in a patient-specified biomechanical lung model for specific tumor movement estimates. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 6964-6967). IEEE.
8. Xie, Y., Xia, Y., Zhang, J., Song, Y., Feng, D., Fulham, M. and Cai, W., 2018. Collaborative and knowledge-based learning for the classification of benign-malignant lung nodules in chest CT. IEEE transactions on medical imaging, 38(4), pp.991-1004.
9. Xu, X., Hou, R., Zhao, W., Teng, H., Sun, J. and Zhao, J., 2020, July. A Poor Automated Lung Cancer Classification Supervisory System for the Entire Slide Picture. In 2020

ISSN 2515-8260 Volume 07, Issue 09, 2020

42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) (pp. 1372-1375). IEEE.

10. Akbar, J., Shahzad, M., Malik, M.I., Ul-Hasan, A. and Shafait, F., 2019, December. Aerial images with deep learning Runway Detection & Location. In 2019 Digital Image Computing: Techniques and Applications (DICTA) (pp. 1-8). IEEE.