WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Lung Cancer Detection Using CT Image Processing.

Mallesh Malagi[1], Hemant Patil[2], Basavaraj Malagi[3], Bheemangoud Nagaral[4], Dr Nandkishor Deshpande[5]

[1]Computer Science Dept. AITM Belagavi, India.
[2]Computer science dept AITM Belagavi, India.
[3]Computer science dept AITM Belagavi, India.
[4]Computer science dept AITM Belagavi, India
[5]Prof Dept of CSE AITM, Belagavi, India

ABSTRACT—

In this generation many people's sufferings from varieties of health issues and there are many more reasons such as climate change, food issues and irregular exercise etc. and now a day's most of us suffering from health issues like diabetes, aids and cancer and in this group, cancer is the one of the popular and most dangerous health issue. There are many types of cancer exists in medical system such as blood cancer and lung cancer. Lung cancer is one of the most common cancers out there in humanity and it is too danger too as compared to other types of cancer. Doctors will find occurrence of cancer in our body using some different technologies like MRI scan, CT scan and bone scan etc. but CT scan is the one of the best technologies to detect cancer because when we use CT scan for scanning our body it will scan our body in close section and the output of this images shows bone, tissue and even some organs too. In this paper we are proposing system by using this system we will going to detect the lung cancer without going to the hospital and here we are using machine learning technologies in order to predict the detection of cancer occurrence in our lungs. In this project our main motto is too detect occurrence of cancer in lungs using CT Images and we are set up 4 steps in order to detect cancer portion in lungs and the first step is to collect CT images from any sources and then in second step we have to preprocess that images by using noise reduction algorithm in order to remove noise and unnecessary data from the CT image then in third step we are measuring some parameters by using some existing features then in last step we are classifying the portions of cancer and non-cancer portions.

Keywords-DIAMO, pre-processing, SSD, CT images, objects.

INTRODUCTION

When we think about cancer then lung cancer is very common and deadly and it will occur in both gender (Male and Female) in simple lung cancer means formation of malignant due uncontrolled growth cells. The main causes for lungs cancer are smoking and eating tobacco because when anyone person smoke then it will produce some cancer-causing substances in our body such as carcinogens and heavy use of this substances lead to cancer in our lungs. In terms of cancer patients time is the one of the important factors, do you know why? Because by time cancer spreads across the body and it also increases its size and it is to danger for survival of cancer patient but when we find cancer in early stage then it is increasing the survival nature of cancer patient. There are two types of cells are available when any person suffering from cancer those are non-small cells and small cell and difference in between small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC) are SCLC is more aggressive than NSCLC that means SCLC is have the capacity of spreading cancer to the other parts of the body and SCLC is less common as compared NSCLC and it may be NSCLC 80-85% and SCLC 10-15%. We can assume different stages of the cancer such as stage-1, stage-2, stage-3 and stage-4 based on the size of the malignant, tumor and portion.

There are many ways to diagnose cancer under the care of doctors and doctors are able to detect cancer by seeing images are taken with help of CT scan , x-ray , MRI scan PET scan but here CT scan is best because it scans from the closest position of body and it will depth representation of bone , tissue and even some organ too but this process of detecting cancer will takes more time and it will lead to more errors so in order to avoid this problem we are proposing system on lung.

LITERATURESURVEY

1. A Title: Title: Tata clustering method for discovering clusters in spatial cancer data sets. Author: Ritu Chauhan, Harleen Kaur, M.Afshar Alam Journal: International Journal of Computer Application. Year: 2010 • In this study, the researchers used the data analytical tools and data mining techniques to cluster medical spatial datasets, includes discovery and extraction of hidden, interesting, and useful patterns from this large database by grouping the objects into clusters. • In this paper both K-means and HAC (Hierarchical aggolometric clustering) algorithm and SOM.

2. A Title: Breast cancer data analysis for survivability studies and prediction. Author: Shukla N, Hagenbuchner, Yang J Journal: Computer Methods and Programs in Bio-Medicine. Year: 2018 • Unsupervised data mining methods are applied such as the SOM algorithm. • DBSCAN algorithm is able to detect area of high density. • These two algorithms are created new way of data driven approach is developed
3. Title: Artificial Neural Networks for automatic segmentation And Identification of nasopharyngeal carcinoma Author: Mohamad M.A, Ghani M.K.A, Hamed, Ibrahim D.A, Abdullah M.K Journal: International Journal of Engineering and Advanced Technology (IJEAT). Year: 2019 • In this paper researchers used data mining to generate new information by analyzing large dataset. • Here they also used artificial neural networks for automatic segmentation of objects in an image. • It is 91% accurate.

METHODOLOGY

1. DATA COLLECTION

Our proposed system is totally based on image CT processing so in this step we need to collect some of the CT images in order to detect cancer and CT lung images are available in online and we can it for research work and we need to save this CT images in the form of DIACOM and the long form of Digital Imaging and Communication In Medicine and there are two types of DICOM versions are available those are single and multi and in terms of lung cancer we are using DICOM.

2. IMAGE PRE-PROCESSING

Image pre-processing is the one of the crucial and important steps before building the actual model and usually this term will do not increase the quality of a CT image but it will remove distortion and noise from the image and it will also enhance some image features those are help to gain well results and in terms of image pre-processing, we have one steps that is PBT and that is nothing but Pixel Brightness Transformation. Brightness transformations modify pixel brightness and the transformation depends on the properties of a pixel itself. In PBT output pixel's value depends only on the corresponding input pixel value. Examples of such operators include brightness and contrast adjustments as well as color correction and transformations. Contrast enhancement is an important area in image processing for both human and computer vision. It is widely used for medical image processing and as a pre-processing step in speech recognition, texture synthesis, and many other image/video processing applications

IMAGE SEGMENTATION

Image segmentation is a method in which a digital image is broken down into various subgroups called Image segments which helps in reducing the complexity of the image to make further processing or analysis of the image simpler. When we preprocess the CT images then we have to apply image segmentation methods for creating subgroups in that images in order to extract the same features and by using this technique we are able to get less complex representation of an image. Image segmentation also helps to find objects of our own interest and it helps for further processing such as recognition and detection and in this proposed system we are using non contextual thresholding. In thresholding process, we simply convert color image or grey color image to binary image. Binary image means those are images that have only two colors (black and white) on their pixels.

1. FEATURE EXTRACTION

Feature extraction is nothing but getting useful features from existing data and this term also known as extracting. In practice, there are two types of mainstream object detection algorithms. Algorithms like R-CNN and Faster R-CNN use a two-step approach - first to identify regions where objects are expected to be found and then detect objects only in those regions using convnet. On the other hand, algorithms like YOLO (You Only Look Once) and SSD (Single-Shot Detector) use a fully convolutional approach in which the network is able to find all objects within an image in one pass through the convnet. The region proposal algorithms usually have slightly better accuracy but slower to run, while single-shot algorithms are more efficient and has as good accuracy and that's what we are going to focus on in this section. In the proposed system we are using single short detection algorithm for detection. In image classification computer vision (CV) takes an image and predicts the object and in other hand object detection method detect both objects as well its location. SSD mainly contain two parts those are backbone model and SSD head model and here backbone model is pre-trained model for image classification network and work as feature extractor and SSD head is the one more convolutional layer to help backbone model.

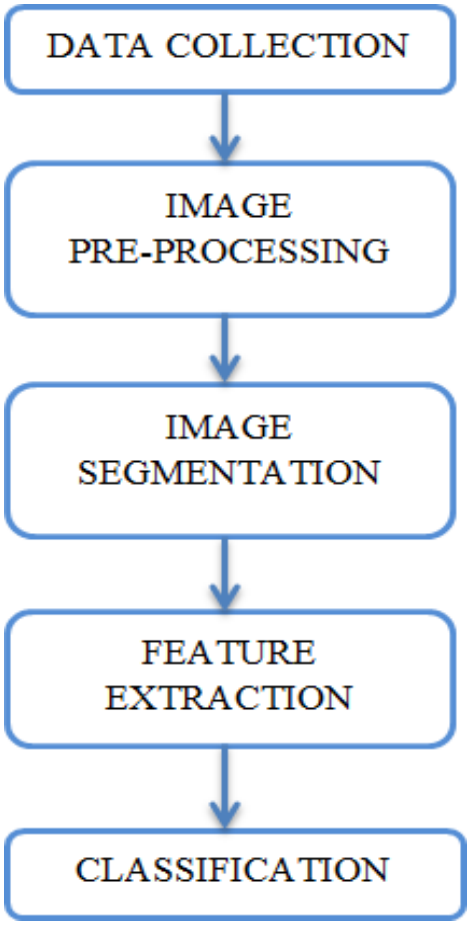

Fig 1.1 Overall methodology.

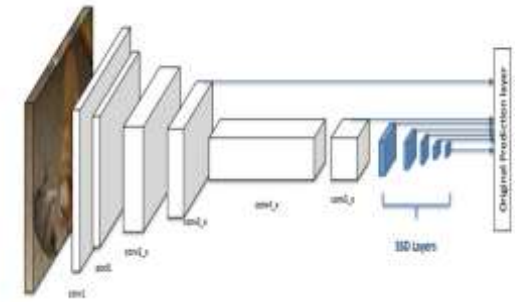

Fig 1.2 Brief view of SSD algorithm architecture

In this figure white layers are backbone models and blue layers are SSD head models. SSD divides the image using a grid and have each grid cell be responsible for detecting objects in that region of the image. Detection objects simply means predicting the class and location of an object within that region. If no object is present, we consider it as the background class and the location is ignored. For instance, we could use a 4x4 grid in the example below. Each grid cell is able to output the position and shape of the object it contains.

Each grid cell in SSD can be assigned with multiple anchor/prior boxes. These anchor boxes are pre-defined and each one is responsible for a size and shape within a grid cell.

SSD uses a matching phase while training, to match the appropriate anchor box with the bounding boxes of each ground truth object within an image. Essentially, the anchor box with the highest degree of overlap with an object is responsible for predicting that object's class and its location. This property is used for training the network and for predicting the detected objects and their locations once the network has been trained. In practice, each anchor box is specified by an aspect ratio and a zoom level.

CLASSFICATION

The Classification stage involves labeling the CT scan images as normal and abnormal. In our method SSD algorithm will be used for detection of lung cancer in CT images. It has no delegated region proposal network and predicts the boundary boxes and the classes directly from feature maps in one single pass. To improve accuracy, SSD introduces: small convolutional filters to predict object classes and offsets to default boundary boxes.

CONCLUSION

Lung cancer detection using image processing project helps to detect the cancer in both x-ray and CT scan images and this process is lesss expensive as compared to the hospitality checking and with help of this system peoples can easily detect the cancer with the help of trained models.

REFERENCES

- Ritu Chauhan, Harleen Kaur, M.Afshar Alam (2010) Data Clustering Method for Discovering Clusters in Spatial Cancer Databases, International Journal of Computer Applications (0975 – 8887) Volume 10– No.6.
- Shukla, N., Hagenbuchner, M., Win, K. T., & Yang, J. (2018) Breast cancer data analysis for survivability studies and prediction. Computer Methods and Programs in Biomedicine, 155, 199–208. [26]. Self-Organizing Maps. Springer-Verlag New York, Inc.; 2001.
- Mohammed, M.A., Ghani, M.K.A., Hamed, R.I., Ibrahim, D.A., Abdullah, M.K., (2017) Artificial neural networks for automatic segmentation and identification of nasopharyngeal carcinoma. J. Comput. Sci. 21, 263–274
- Wu CF, Wu YJ, Liang PC, Wu CH, Peng SF, Chiu HW (2017) Disease free survival assessment by artificial neural networks for hepatocellular carcinoma patients after radiofrequency ablation. J Formos Med Assoc; 116:765e73.
- K. Punithavathy, M.M. Ramya, Sumathi Poobal, "Analysis of statistical texture features for automatic lung cancer detection in PET/CT images", International Conference on Robotics, Automation, Control and Embedded systems (RACE), IEEE ,18-20 February 2015.
- Badrul Alam Mia, Mohammad Abu Yusuf, "Detection of lung cancer from CT image using image processing and neural network", International conference on Electrical Engineering and Information Communication Technology (ICEEICT), IEEE, May, 2015.
- Anita Chaudhary, Sonit Sukhraj Singh “Lung cancer detection on CT images using image processing”, computing sciences 2012 international conference, IEEE, 2012.
- Nooshin Hadavi, Md Jan Nordin, Ali Shojaeipour , “Lung cancer diagnosis using CT-scan images based on cellular learning automata”, International conference on Computer and Information Sciences(ICCOINS), IEEE , 2014.