**CANCER DETECTION USING MACHINE LEARNING TECHNIQUES**

**Abstract:**

Lung cancer is one of the most lethal cancer types; thousands of peoples are infected with this type of cancer, and if they do not discover it in the early stages of the disease, then the chance of surviving of the patient will be very poor. For the suggested reasons above and to help in overcoming this terrible, early diagnosis with the assistance of artificial intelligence procedures most needed. Through this research, a Computer-aided system introduced for detecting lung cancer in a dataset collected from the Iraqi hospitals by using a convolutional neural network technique with AlexNet architecture for helping with the diagnosis of the patient's cases: normal, benign, or malignant. The proposed model gives high accuracy ups to 93.548%. The other performance metrics comes with high values such as 95.714% for sensitivity and 95% for Specificity.

**CHAPTER-1**

INTRODUCTION:

Lung cancer is one of the most well-known life-threatening illnesses in the globe(1). The up-to-date estimations are given by the "World Health Organization" (WHO) say that approximately 7.6 million mortality worldwide every year because of lung cancer. Furthermore, mortality due to this type of cancer is assumed to continue growing, to become almost 17 million worldwide in 2030(2, 3). According to the statistics of “The American cancer society”, lung cancer is the head cancer killer in people in the United States.(4) The overall number of the estimated new cases of all types of cancer in 2013 was 1660290 (854790 for men and 805500 for women), the number of lung cancer was 228199 incidences (118080 for men and 110110 for women). Where the total number of estimated death cases of cancer are 580350 cases (306920 for men and 273430 for women), in the case of the lung cancer the number of death cases was 159480(87260 for men and 732220 for women)(5). According to the Iraqi ministry of health in 2016, lung cancer is the second most widespread cancer type in Iraq. There are 2123 people who have lung cancer from two genders. This number represents about 8.31% of the total infections in the country. This portion indicates a small increase as compared to the ratio of the past year, which is about 8.1%. The rate of lung cancer represents approximately 13.27% of the total cancer cases, and this shows that lung cancer is the leading type in males. Also, there is a rise when compares to the ratio recorded in 2015, which approximately reaches 12.7%.(6, 7)**.** For females, lung cancer not the leading cancer type, it ranked fifth between other cancer types, in 2016 there is only 638 woman who has this disease, there are 638 women who diagnosed with lung cancer at 2016, this represents about 4.44% from the total cancer types infections. There is a slight rise if compared with the previous year, which represents approximately about 4.2%.(6, 7) Cancer is the fourth cause of death in the eastern Mediterranean region and is the third cause of mortality in Iraq,

and this rate is growing continuously. The chief and most notable of this raise is smoking. Other factors include pollution, unhealthy diet, endless exposure to manufacturing and agricultural carcinogens, and lowered physical motion(6). Total cancer mortality in Iraq in 2014 is (8211), approximately (4525) in males and (3959) in females. The most cancer sort deaths were the lung cancer with a total number of estimated deaths about (1339), (918) of them was for men and (421) of the total calculated estimate was for women, the entire portion of lung cancer

mortality amongst all other kinds of malignancy was 16.31%. During 2016, the whole number of cancer losses fell to 7568 cases, where the most significant portion of them was lung cancer, approximately 1257, which implies a total percentage of 16.61% from the total predicted mortality(6, 7). So for the reasons explained above, there is a necessity for implementing a CAD system for helping doctors in diagnosing lung cancer as possible as when in it is early stage, not only detecting the nodule but with high accuracy. Several studies applied artificial intelligence techniques for this purpose, for examples: using artificial neural network for detecting lung cancer as in (8, 9), or using support vector machine technique as in (10-12)*,* or applying

K-nearest neighbor as in (13), or using genetic algorithm for this operation as in (14, 15), also, fuzzy techniques are efficient when using to detect lung cancer as in (16-18), convolutional neural network can be used for this purpose as in (19-21). Artificial intelligence not only used in the area of lung cancer diagnosing, but it applied in all fields of biomedical engineering such as: diagnosis of breast cancer in (22-24), diagnosis of Heart disease in (25-27), also diagnosing and classification of diabetes in (28). In order to apply the above machine learning techniques, there is a need for using data as input to the algorithms which have been applied. Various methods are convenient for diagnosing lung cancer, particularly MRI, isotope, Xray, and CT. X-ray chest radiography and Computer Tomography (CT) are the two well-known imaging modalities

that are commonly utilized in the identification of different lung diseases(2, 29). In addition, there are many publicly databases used for the purposes of scientific research such as: ELCAP Public Lung Image Database, LIDC Database, and Data Science Bowl 2017.

The aim of our study is to implement a CAD system used as an assistant to doctors while deciding and diagnosing lung cancer, this system used for detecting and classifying the lung cancer cases if it normal, benign, or malignant with high accuracy. This done by applying convolutional neural network technique to a data set of lung cancer CT scans collected and diagnosed at the Iraqi hospitals.

**OVERVIEW**

Computer-aided system introduced for detecting lung cancer in a dataset collected from the Iraqi hospitals by using a convolutional neural network technique with AlexNet architecture for helping with the diagnosis of the patient's cases: normal, benign, or malignant.

**RELATED WORK**

In the proposed procedure, CNNs are applied to detect and classify lung cancer CT scans of the patients collected from hospitals. **Convolutional Neural Networks** is a sort of deep learning paradigm applied for processing data which has a grid pattern like images (31), it is all about using Deep Learning with Computer Vision. A good way to gain foreknowledge about this technique is to imagine a Neural Network Architecture also how it is practiced to

visual tasks i.e. Video and Images. Furthermore, the Convolutional Neural Networks is an important technique used for Object Recognition, create Facial Recognition, Self-Driving Cars. A **Convolutional Neural Network** is a Deep Learning algorithm that can take in image as input, with assigning importance learnable weights and biases to various objects inside this image and be capable of differentiating one from the other. In addition, the pre-processing

required for this technique is much lower if comparing with other classification algorithms. The role of the CNN is for reducing the images to a form that is easier to process but without losing features that are important for getting a good prediction (32). A typical CNN consists of three types of operation layers: the convolutional layer (CONV), the pooling layer (POOL), and finally the classifier layer (FC).

**PROJECT STATEMENT**

Lung cancer is the most perilous and widely spread cancer in the world according to stage of discovery of the cancer cells in the lungs. The motivation of this paper is to discover the cancer cells in the lungs at earlier stage. The lung cancer is detected using CAD System which is an interdisciplinary approach based on the techniques of Image Processing and Machine Learning. The forecasting of lung cancer is the most challenging problem, because of the structure of cancer cells, where most of the cells are superimposed. Recently, the image processing techniques are widely used in several medical areas for detection and treatment levels. The time factor is most important to determine the abnormality issues in the targeted image. Image quality and accuracy are the significant factors for quick identification of diseases. Image quality assessment and advancement are depending on the enhancement stage. For the suggested reasons above and to help in overcoming this terrible, early diagnosis with the assistance of artificial intelligence procedures most needed. Through this research, a Computer-aided system introduced for detecting lung cancer in a dataset collected from the Iraqi hospitals by using a convolutional neural network technique with AlexNet architecture for helping with the diagnosis of the patient's cases: normal, benign, or malignant. The proposed model gives high accuracy ups to 93.548%. The other performance metrics comes with high values such as 95.714% for sensitivity and 95% for Specificity.

**SCOPE OF THE PROJECT**

The scope of this project is to use Deep Learning techniques for early detection of lung cancer from CT scans and use the results in clinical diagnostics and cancer screening applications to support radiologist's diagnosis.

**PROPOSED SYSTEM**

Lung cancer is one of the most well-known life-threatening illnesses in the globe(1). The up-to-date estimations are given by the "World Health Organization" (WHO) say that approximately 7.6 million mortality worldwide every year because of lung cancer. Through this research, a Computer-aided system introduced for detecting lung cancer in a dataset collected from the Iraqi hospitals by using a convolutional neural network technique with AlexNet architecture for helping with the diagnosis of the patient's cases: normal, benign, or malignant. The proposed model gives high accuracy ups to 93.548%. The other performance metrics comes with high values such as 95.714% for sensitivity and 95% for Specificity.

**Hardware and Software requirements**

**REQUIREMENT ANALYSIS**

The project involved analyzing the design of few applications so as to make the application more users friendly. To do so, it was really important to keep the navigations from one screen to the other well ordered and at the same time reducing the amount of typing the user needs to do. In order to make the application more accessible, the browser version had to be chosen so that it is compatible with most of the Browsers.

**3.2.REQUIREMENT SPECIFICATION**

**3.2.1.Functional Requirements**

* Graphical User interface with the User.

**3.3.Software Requirements**

For developing the application the following are the Software Requirements:

1. Python
2. tkinter

**3.4.Operating Systems supported**

1. Windows 7
2. Windows XP
3. Windows 8

**3.5.Technologies and Languages used to Develop**

1. Python

**3.6.Debugger and Emulator**

* Any Browser (Particularly Chrome)

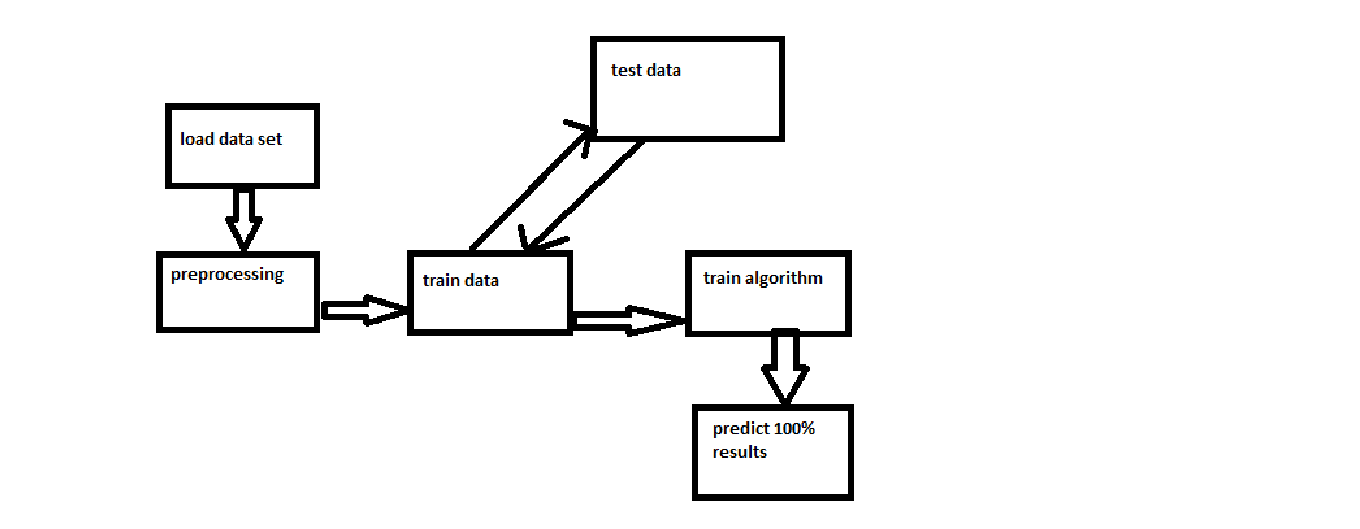
**3.7.Hardware Requirements**

For developing the application the following are the Hardware Requirements:

* Processor: Pentium IV or higher
* RAM: 256 MB
* Space on Hard Disk: minimum 512MB

**Detailed Design**

**System Architecture**



CONCLUSION AND FUTURE SCOPES.

This proposed study tries to defeat the problems faced in the early detection of lung cancer nodules before it gets worst. For this purpose, this study develops an effective computer-aided diagnosis scheme for early detecting of this lethal cancer. Chest tomography scans have been employed here as data input to the proposed model. This study's goal was to improve a CNN deep learning model able to detecting and classifying lung cancer nodules successfully. The obtained model gives high accuracy reaches 93.548% while applying on the dataset collected. In the future, this study can be improved by includes regularly enhancing the accuracy of the model by training it on bigger and more extensive datasets. Besides, various machine learning models can be combined to compare