RESEARCH ARTICLE

AN INTELLIGENT CBMIR SYSTEM FOR DETECTION AND LOCALIZATION OF LUNG DISEASES

Nora I. Yahia[1], Dr. Ayman I. Al-Dosouki[2], Dr. Sahar A. Mokhtar[3] and Dr. Hany M. Harb[4]

1. PhD. Student, Computer and system Department, Electronics Research Institute, Cairo, Egypt.
2. Professor, Computer and system Department, Electronics Research Institute, Cairo, Egypt.
3. Professor Assistance, Computer and system Department, Electronics Research Institute, Cairo, Egypt.
4. Professor, Computer and system Engineering Department, Faculty of Engineering, Al-Azhar University, Cairo, Egypt.

……………………………………………………………………………………………………......

Manuscript Info

……………………….

Abstract

………………………………………………………………

The diagnosis of lung diseases is a complicated and time-consuming task for radiologists. Often radiologists struggle with accurately diagnosing lung diseases, They use Commonly CT imaging signs (CISs) which common appear in CT lung nodules in the diagnosis of lung diseases. Computer-aided diagnosis systems (CAD) can automatically diagnose and detect these signs by analyzing CT scans, which will reduce radiologist's workload. The diagnosis and recognition efficiency and accuracy can be improved by using content-based medical image retrieval (CBMIR). This paper proposes a new intelligent CBMIR method to retrieve CISs helping in diagnosing and recognize lung diseases by using deep Convolutional Neural Network (CNN). Fine-tuned YOLOv4 (You Only Look Once) object detector model are proposed to fast detect and efficiently localize signs in real-time. The proposed CBMIR system can be applied as a useful and accurate medical instrument for diagnostics. The experimental results show an average detection accuracy of CT signs lung diseases as high as 92% and a mean average precision (MAP) of 0.92 is achieved using the proposed technique. Also, it takes only 0.1 milliseconds for the retrieval process. The proposed system presents high improvement as compared to the other system. It achieved better precision of retrieval results and the fastest run of the retrieval time.

……………………………………………………………………………………………………......

Introduction:-

Nowadays many hospitals and clinics are generating enormous medical images with a variety of medical imaging types such as computed tomography (CT), magnetic resonance imaging (MRI), X-ray, and many more, which cause a large collection of databases. To efficiently access, search and retrieve such a massive amount of medical images, there are two approaches in image retrieval processes which are text-based Image retrieval (TBIR) and Content-Based Image Retrieval (CBIR). In a diagnostic system, CBMIR is considered an efficient medical tool to help the physician in diagnosing the diseases. The main role of CBMIR is to analyze and retrieve similar images for the input image. It is based on an image's visual features such as color, shape, and texture, or any other features. The traditional CBMIR usually utilizes low-level features or high-level semantic knowledge separately. So, there is a semantic gap

Corresponding Author:-Nora I. Yahia
Address:-PhD. Student, Computer and system Department, Electronics Research Institute, Cairo, Egypt.

651

between low-level features and high-level semantic descriptions. The challenge of CBMIR systems is to eliminate the semantic gap. In terms of their features, part of the data is missed when describing medical images in terms of their features, i.e., from high-level to low-level features [1]. As a promising method in artificial intelligence, deep learning techniques have been proven to be successful in bridging the semantic their features gap, also, they have powerful techniques for improving medical image analysis which is used in the recognition and diagnosis of diseases. CBMIR has established its importance in medical diagnosis by analyzing large medical databases based on visual and semantic features. Furthermore, deep learning is a current successful technique in image analysis, so, it achieved significant results for CBMIR. Also, CBMIR based on the matching and similarity measurements in which the query image compared with the database images by using distance metrics such as Euclidean distance, Manhattan distance, and Cosine similarity, and others are used [3], [4], [5], and [2].

This paper proposes a new deep learning technique to predict, detect, localize, and retrieve CISs, which helps in the diagnosis of lung diseases through a transfer-learning approach. The proposed framework uses the transfer learning by configuring a pre-trained model. A pre-trained model is a network which has been previously trained on a larger dataset and can be reused in a new model with different datasets. Also, the proposed deep learning CNN overcomes the semantic gap depending on a low visual level and semantic features extraction of CT scan. Also, this approach proposes a new methodology for retrieving similarities between CISs to enhance the precision of the CBMIR framework. This method can assist the physician by comparing the relevant historical cases with their diagnostic reports, which can enhance the precision of diagnosis. In addition, the proposed system can teach medical students how to detect and localize lung nodules, which helps in disease diagnosis.

This work includes six sections. The investigation of CBMIR systems is listed in section2. Section 3 presents the new proposed CBMIR system in detail and introduces the new proposed fine-tuned deep learning methodology. Also, the retrieval process is discussed in section 3. In section 4, experimental results are discussed. The evaluation of the proposed system and its comparison with other systems is given in section 5. Finally, the conclusion and possible future directions are introduced in section 6.

Related work:-
This section discusses the previous work of the various CBMIR systems for Lung CT images.

Ling Ma and Xiabi Liu [6] proposed a fused context-sensitive similarity method of CBMIR applying the SVM technique to build an SVM classifier of semantic and visual features. They constructed a weighted graph whose nodes represent the images and edges to measure their pairwise similarity by using the shortest path algorithm. The shortest path computation consumes a lot of the running time because they didn't cluster the datasets.

Biswas and S. Roy [7] described the CBMIR system for extracting pulmonary nodules from CT images. The preprocessing of images is the first step of this system. Then segmentation of the preprocessed images is performed by cropping the region of interest (ROI), The gray level co-occurrence matrix (GLCM) and shape features are used for feature extraction by applying Fast Discrete Curvelet Transform (FDCT) to extract the co-efficient from images. The irrelevant features are deleted from extracted features and limited features are selected using the Enhanced Moth Flame Optimization algorithm. Then, detection of the various types of lung nodules is achieved by a deep learning algorithm. Finally, retrieving the nodules is performed.

Biswas, S. Roy, and A. Biswas [8] presented a Triplet-CBMIR method to retrieve the lung nodules. A Triplet CBMIR based on shape and texture low-level features, semantic features, and relevance feedback techniques applying Hybrid wavelet decomposition. CNN is used for feature extraction and the mutual information based Neighborhood Entropy algorithm for feature selection.

Ling Ma and Xiabi Liu [9] applied the local binary pattern, the wavelet features, the bag of visual words based on the HOG, and the histogram of CT on a low-level visual scale from the ROI of CISLs. In addition, they used the auto-encoder neural networks for high-level semantic scale with distribution of the optimized features on the mid-level attribute scale. The sum of all similarities at multiple levels is calculated.

Juan Zh., Ling p., Peng Fei [10] proposed a retrieval method to recognize signs of lung nodules. Semantic image features are extracted using CNN which is interpreted into binary codes with supervised hashing.

Qayyum et al. [11] proposed the Convolutional neural network method to classify Interstitial Lung Diseases. The evaluation results proved the performance of classification reached a value of 85.5% in lung pattern classification and achieved an average precision of 0.69. They applied the Boltzmann method of deep learning to convolutional classification.

Ash.Dasare and Harsha S [12] proposed a Content-Based Image Retrieval (CBIR) to analyze malignant and benign nodules using nine visual and shape-based features extraction. They used Minkowski distance for the similarity computation of retrieval

Diego Riquelme and Moulay A. Akhloufi [13] presented a review of the deep learning algorithms and architectures for lung cancer detection to improve the performance and overcome some limitations of the standard CNN, such as Residual-Networks (ResNets), Inception, Xception, or Dense Networks.

Hongtao, DongbaoY, Nannan, Zhineng, Yongdong [14] used a Faster Region-based Convolutional Neural Network with two regions proposed networks and a deconvolution layer to detect pulmonary nodules. The boosting architecture of CNN is used as a classifier to reduce false positives.

Kashif, Gulistan Raj. and f. Shaukat [15] applied a sum of descriptors such as local ternary pattern, local phase quantization, and discrete wavelet transform with joint mutual information for feature selection which apply the same method of [6].

The next section will introduce the proposed CBMIR system in detail.

The proposed CBMIR system:-
In this section, the CBMIR and object detector models based on deep learning have been developed. The main aim of this system is to efficiently detect and localize signs of lung disease in CT scans and efficiently retrieve similar images.

The proposed CBMIR system can incorporate visual and semantic information together to address their limitations, and also use the fundamental database structure for better retrieval performance. The proposed CBMIR system consists of two main modules: the training module and the retrieval module. In the training module, the images are annotated and augmented to train the object detector model, as shown in figure 1. However, in the retrieval module, a query image is used as input for the trained model to recover the lung CT images from the large medical image dataset as shown in figure 2. In the next subsection, each module of the proposed method will be explained in detail.

Training Module:-
In this module, a deep neural network model is trained on the dataset to detect and localize signs in CT scans of lung diseases.

As shown in figure 1, in order to train the model, the images are annotated to mark ROI. Then the annotated images are augmented based on an augmentation algorithm. Finally, the Yolov4 object detection model is trained on this augmented image to detect the disease. In the next, each step will be clarified in detail.
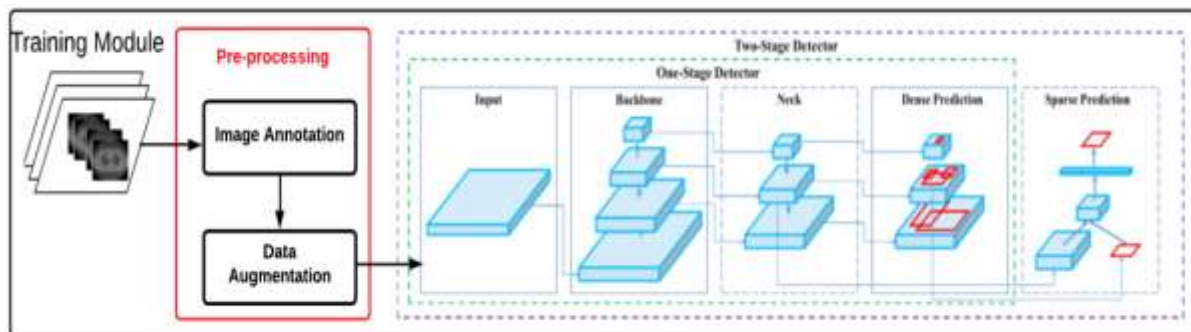

Figure 1:- The proposed training module.

Image processing:-

Image preprocessing is the first stage in the proposed training module; it provides better image visualization to improve the image quality. Also, this step involves image conversion from Digital Imaging and Communications in Medicine (DI-COM) to other formats such as PNG, JPG, or TIFF, or NIFTI for ease of distribution. In the proposed system, the image is transformed from DICOM format to JPG format and it is resized. There are two new proposed preprocessing steps which are described in the following subsections.

Image Annotation:-

Image annotation is the process of labeling images, including the relevant regions and class labels, in order to make the database very useful. In the proposed image annotation step, a qualified senior radiologist marks ROI for each JPG image by bounding box, then every image file is accompanied by a label file indicating the disease boundaries coordinates using an automatic labeling annotation tool.

Image data augmentation:-

Data augmentation is applied as a pre-processing step which helps inartificially increasing the size of data and adding diversity to training datasets while avoiding overfitting. In this paper, the following augmentation techniques are applied to generate more CISs for training the model:

1. Flipping: both horizontally and vertically.
2. 90° Rotate: rotate image 90 degrees both clockwise and counter-clockwise.
3. Rotation: with a rotation range of -21° and $+21^{o}$.
4. Brightness: with a brightness range of -14% and $+14\%$.

These four proposed methods are randomly applied with one of the methods, two combined or all together. Rotating larger than 90 degrees would not add much variance.

The next section shows how to extract the features of the augmented signs from CT scans. Also, it introduces the details of the proposed training model, prediction, and localization of CISs

Proposed CNN model:-

In this section, a deep convolutional neural network algorithm is applied to the proposed CBMIR to retrieve CT lung signs. CNN is one of the most widely used models for deep learning. The simple network consists of an input layer and an output layer, as well as a hidden layer. The hidden layer consists of a series of convolutional layers, a Rectified Linear Unit (RLU) activation feature, pooling layers, fully connected layers, and normalization layers. There are different pre-trained architecture models available on CNN [16], such as AlexNet [17], GoogleNet [18], VGGNet [19], ResNet50 [20], DarkNet, etc. In pre-trained deep learning models, a model is trained on a large volume of the dataset, and during training it learns model weights and bias, which are transferred to smaller datasets.

The YOLOv4 [21] model is a CNN architecture consisting of CSPDarknet53 as a backbone, spatial pyramid pooling, path-aggregation neck, and the head of YOLOv3. It is one of the most accurate, fast-operating and achieves optimal average precision models in object detection. The new proposed method fine-tunes the YOLOv4 model in order to train CT lung signs and predict diseases effectively and efficiently in real-time. The YOLO model [22] is an object detection model that extracts features from input images and then feeds these features through a prediction system to recognize objects and predict their classes rapidly and precisely. The main idea of YOLO is to segment the input image into a grid of several cells and estimate the probability of getting an object for each cell using anchor boxes. The YOLO output is a list that contains the coordinates of the detected bounding box and probability classes. All of the YOLO models are typically trained on the COCO dataset, which includes a wide variety of 80 object classes. The YOLOv4 [21] network consists of four main pieces.Input: This is for input images.

Backbone: a network is like a VGG16, Resnet-50, Darknet52, or ResNext50 which extracts the feature map from input images.

Neck: is a sub-set of the backbone to enhance the feature discriminability and robustness using the Feature Pyramid Network, Path aggregation network, and Receptive field block.

Head: is a sub-set of the backbone, which handles the prediction using a one-stage detector for dense prediction or a two-stage detector for sparse prediction.

Also, the new fine-tuned YOLO v4 uses Bag of Freebies (BoF) and Bag of Specials (BoS) techniques for both the backbone and detector to achieve fast speed with high accuracy.

The retrieval module:-
The retrieval process is the last stage of CBMIR to search and retrieve similar images matched to the given query. After the training process is executed, and having the pre-trained weights, all classes to which each image belongs are recognized. In order to retrieve the similar CT lung nodule after classes' predictions, diagnostic and detecting processes are implemented, a new method of retrieval is proposed in this paper as shown in the retrieval process of figure 2. First, the new query image is resized to fit the trained YOLOv4 [21] detector, then the detector can recognize the predicted class of an input image. The predicted label of the input image is stored as the resulting label in the buffer before sending it to the search engine. The search engine searches for the predicted result and queries the diagnostic directory in the database of images, returning the whole matched directory and retrieving similar images.
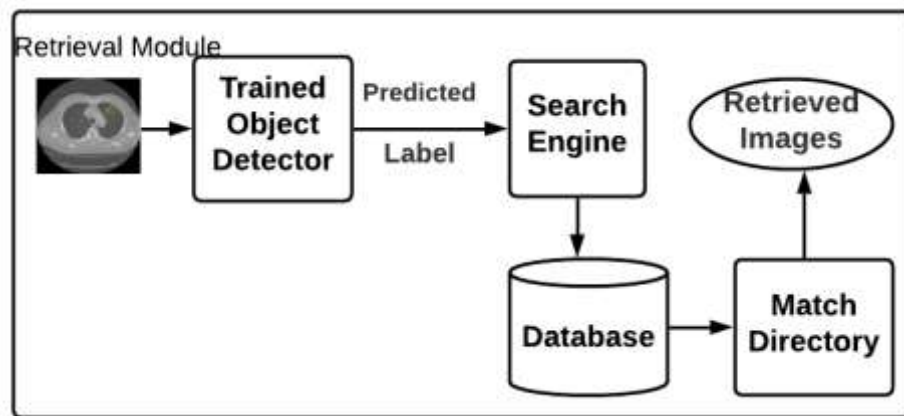

Figure 2:- the proposed retrieval module.

The experimental results of applying the proposed system on the LISS database of CISs are introduced in the next section.

Experimentation and results:-
In order to test the proposed intelligent CBMIR system, a dataset was assembled from the LISS database of CISs [23] and used as an example. The nine categories of CISs are ground glass opacity (GGO), lobulation, calcification, cavity &vacuolus, spiculation, pleural indentation, air bronchogram, bronchial mucus plugs, and obstructive pneumonia as shown in figure 3. All CT scans are collected in the DICOM format with a slice thickness of 5 mm and the image size is 512*512 pixels. The 511 rectangular ROIs are labeled and annotated by an experienced radiologist manually.
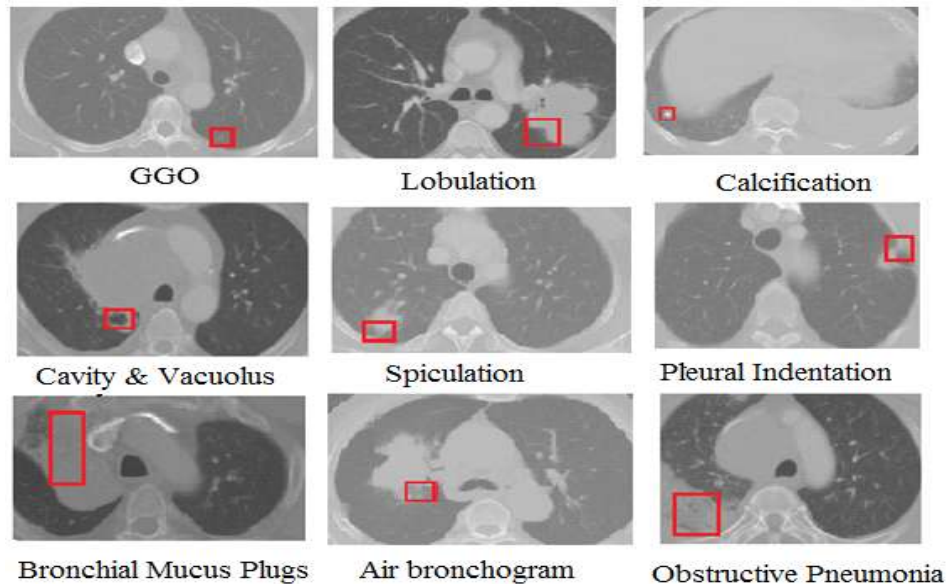
**Figure 3:-** The sample of annotated nine categories CISs.

As discussed in section 3.1, the first step of the proposed framework is to transform the DICOM input images to JPG format and annotate the transformed data to identify ROI for each image. The sample of annotated images looks as shown in figure 4.

Figure 4:- The sample of annotated images.

As discussed in section 3.1.2, in order to train the proposed model, the training data should be large, so that the model may augment the images. Figure 5 shows a sample of the generated image after the proposed augmentation method is applied.
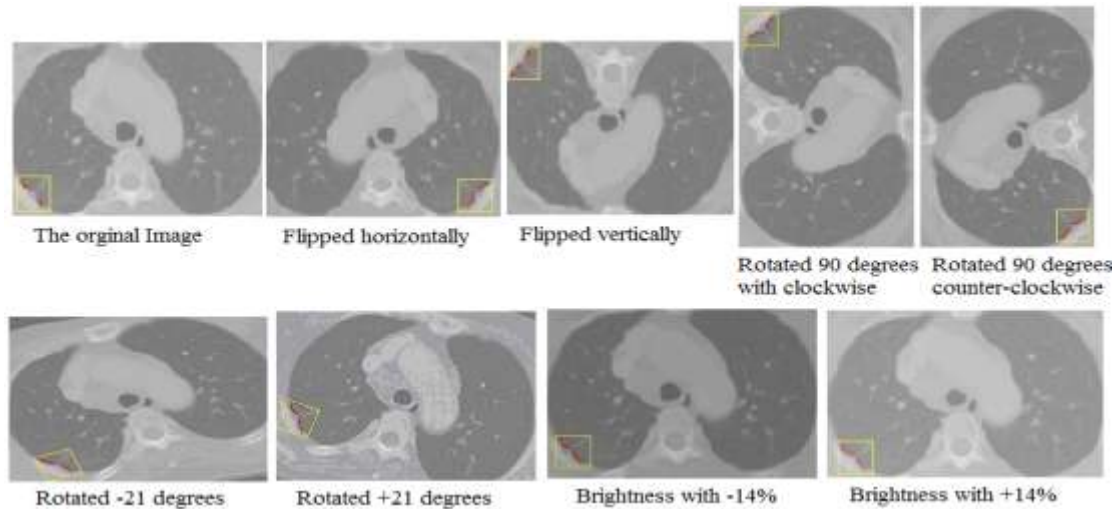
Figure 5:- The applied data augmentation on the same image.

After augmentation, the training images reached a number of 804 images ready for the training step. The data is split into 80:20, 80% for training, and 10% for testing and the same ratio to validate the model. The number of images is 34. Also, the number of validating images is 34 images. In order to train YOLOv4 on the Darknet with 804 images of CISLs, all images are resized to 416x416 and exported to the new annotation files on the Darknet. After that, configure a custom new YOLO v4 training by setting a series of parameters, such as:

1. Set the batch size to 64, where the batch size is the number of images per iteration.
2. Set subdivisions to 16, subdivisions are the number of mini-batches which the batch is broken into for GPU memory, where mini-batches = batch/ subdivisions.
3. Pre-trained weights will be updated for each batch sample of training iteration processes.
4. Update the number of classes from 80 to 9 classes in all of the YOLO layers.
5. Set maximum batches to 2000 * 9.
6. Steps to 80% and 90% of maximum batches.
7. Change filters in all of the YOLO layers to 42 filters according to the number of classes. Where filters = (no of classes+5)*3.

After setting weights that achieved the highest MAP on the validation set and making inferences on a test image, the proposed framework achieved an optimal accuracy of 92%. As shown in figure 6, mAP equals 0.92 after 280 epochs. The heighest precision reached 0.77 at epoch 380. The recall is as high as 0.92 at 400, which loss was converged.
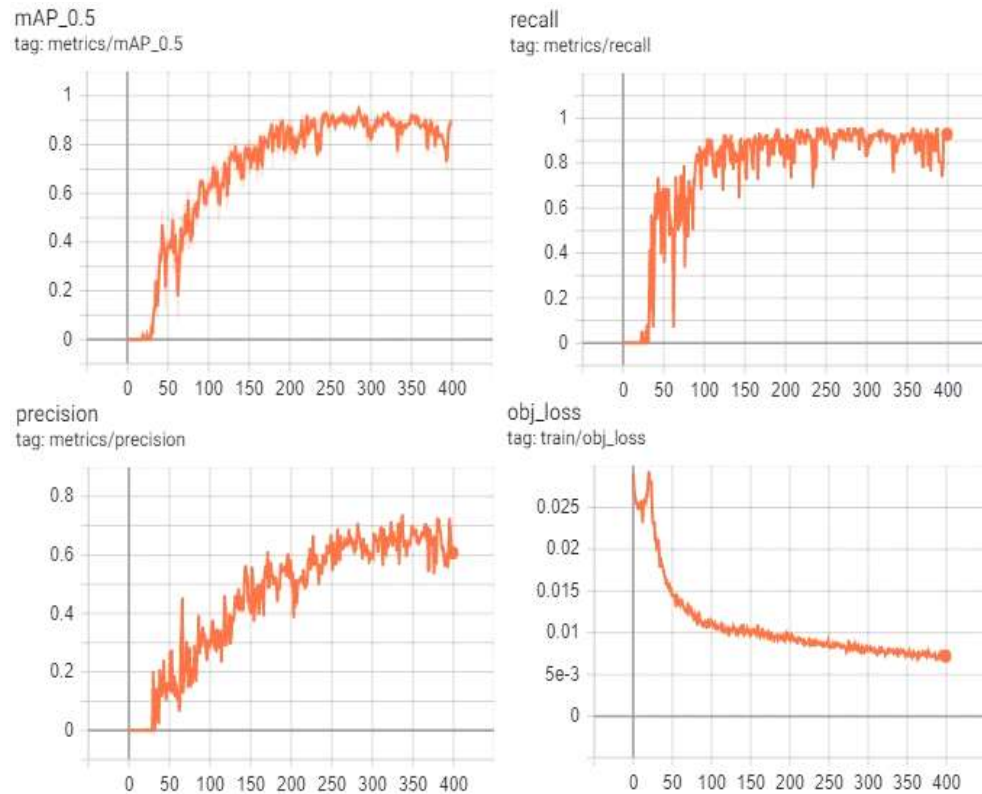
Figure 6:- Graphical representation of performance evaluation of object detector.

As discussed in section 3.2, in order to test the new proposed method of retrieval, when the query image is entered as an input to the proposed system, the output will retrieve all relevant images for the query image according to the diagnosis of this image. Figure 7 shows a sample of the retrieved images which are similar to the input image. The new retrieval method obtains an average precision of 92%. Therefore, the proposed method is effective.
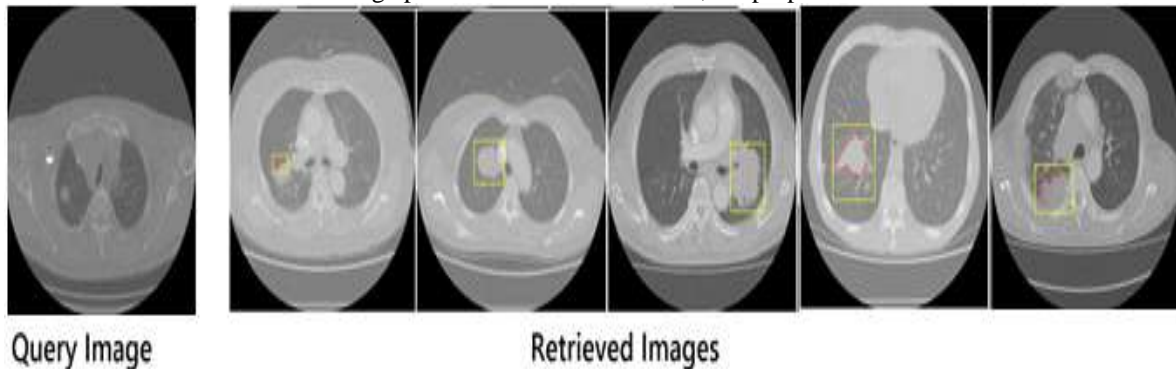
Figure 7:- The sample images of retrieval results.

Comparison:-
To evaluate this proposed system, we compared the performance of this system with other systems from previous research. The experimental results show the proposed method's accuracy is 92% compared with FCSS [6], MLS [9], FDCT [7], TRIPLET [8], and the method by Gulistan Raj [15]. In figure 9, the average precision of the different methods is rated and compared with the proposed method.The proposed method is more performance than the other methods. Also, we compared the proposed retrieval method with the other methods according to the time of retrieval. It noticed that the running retrieval time of the proposed method is 0.1 milliseconds as shown in Table 1. The proposed method is the speediest one as compared to some other methods; because it depends on the matching

after all deep learning stages have finished the training process. All techniques are implemented using the Python framework with a Gtx 1080 ti GPU, an IntelCore i5–2.33 GHz CPU, and 4 GB of memory on a PC.

Figure 9:- The average precision values from proposed method and compared methods.

Table 1:- Comparison of running retrieval time between some approaches and the proposed method.

Methods	MLS	FCSS	TRIPLET	The proposed Method
Retrieval time(ms)	3.6	46.61	0.114	0.1

Conclusion:-
The goal of this paper is to develop a CBMIR system that can detect, localize, and retrieve signs of lung diseases from CT scans of patients. The experimental results showed that the proposed fine-tuned pertained object detector has achieved good results for the diagnosis of CT images of lung signs. The proposed retrieval method is effective and simply searches and retrieves similar images from a large database that are matched with a predicted class label of the input image. The proposed CBMIR system is an intelligent and efficient instrument for helping radiologists in decision-making for the diagnosis of CT images of lung diseases. The evaluation results show an improvement in terms of precision and recall.

Future work:-
In future work, the size of the dataset can be increased by adding a number of classes and evaluating them in real-time on other big datasets. Better training models may be developed for further improvement of retrieval results. The various deep learning methods can be developed by experimenting with the number of layers, and by changing the hyper parameters.

References:-
1. K. K. Kumar and T. V. Gopal, "A novel approach to self-order feature reweighting in CBIR to reduce semantic gap using Relevance Feedback," in Circuit, Power and Computing Technologies (ICCPCT), 2014 International Conference on, 2014, pp. 1437-1442.
2. M. G. Ayadi, R. Bouslimimi, J.Akaichi,"A medical image retrieval scheme with relevance feedback through a medical social network. Social Network Analysis and Mining", Vol. 6, Issue.1, 2016.
3. D. Perez and Y. Shen, Joan Dayanghirang and Jiang Li, Shengli Wang and Zezhong Zheng, "Deep Learning for Pulmonary Nodule CT Image Retrieval —An Online Assistance System for Novice Radiologists", IEEE International Conference on Data Mining Workshops (ICDMW), 2017.
4. R. Marcos. Nesso-Jr., Mirela T. Cazzolato, Lucas C. Scabora, Paulo H. Oliveira,GabrielSpadon, Jessica A. de Souza, Willian D. Oliveira, Daniel Y. T. Chino,Jose F. Rodrigues-Jr., Agma J. M. Traina, Caetano Traina-Jr.RAFIKI: Retrieval-Based Application for Imaging and Knowledge Investigation.IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS), 2018.
5. M.ViniciusN. Bedol & D.Pereira dos Santos&Marcelo Ponciano-Silva&Paulo Mazzoncini de Azevedo-Marques&André Ponce de León Ferreira de Carvalho&Caetano TrainaJr, "Endowing a Content-Based Medical Image Retrieval System with Perceptual Similarity Using Ensemble Strategy. Journal of Digital Imaging", Vol. 29, Issue. 1, PP. 22–37, 2015.
6. Ma L, Liu X, Gao Y et al (2017) A new method of content based medical image retrieval and its applications to CT imaging sign retrieval. J Biomed Inform 66:148–158.
7. R. Biswas, S. Roy Content Based CT Image Sign Retrieval using Fast Discrete Curvelet Transform and Deep Learning, Volume 8, No.3, May -June 2019International Journal of Advanced Trends in Computer Science and Engineering.
8. Ranjit Biswas, Sudipta Roy, Abhijit Biswas, Triplet Contents based Medical Image Retrieval System for Lung Nodules CT Images Retrieval and Recognition Application, international Journal of Engineering and Advanced Technology (IJEAT)ISSN: 2249 –8958,Volume-8 Issue-6, August 2019.
9. Ma L, Liu X, Gao Y et al. A multi-level similarity measure for the retrieval of the common CT imaging signs of lung diseases. Medical & Biological Engineering & Computing (2020) 58:1015–1029.
10. Zhao JJ, Pan L, Zhao PFet al. Medical sign recognition of lung nodules based on image retrieval with semantic features and supervised hashing. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 32(3): 457–469 May 2017.
11. A. Qayyum, S. M. Anwar, M. Awais, and M. Majid, "Medical image retrieval using deep convolutional neural network," Neuro computing, vol. 266, pp. 8-20, 2017.

12. Ashwini Dasare, Harsha S, CBIR System for Lung Nodule Retrieval in CT Scans, International Journal of Innovative Technology and Exploring Engineering (IJITEE)ISSN: 2278-3075,Volume-9Issue-2S, December 2019.

13. Diego Riquelme and Moulay A. Akhloufi, Deep Learning for Lung Cancer Nodules Detection and Classification in CT Scans, MDPI/ journal AI,8 January 2020.

14.] Hongtao, DongbaoY, Nannan, Zhineng, Yongdong, Automated pulmonary nodule detection in CT images using deep convolutional neural networks, Pattern Recognition 85 (2019) 109–119.

15. M. Kashif, Gulistan Raj.,f. Shaukat, An Efficient Content-Based Image Retrieval System for the Diagnosis of Lung Diseases, Journal of Digital Imaging (2020):971–987.

16. Alom,Z.,Taha,T.,Yakopcic,C.,Westberg,S.,Sidike,P.,Nasrin,M.,Essen,B.,Awwal,A.,Asari,   V.   (2018).The history began from AlexNet: A comprehensive survey on deep learning approaches.arXiv preprint arXiv:1803.01164.

17. rizhevsky A, Sutskever I, Hinton GE: Imagenet classification withdeep convolutional neural networks. Adv Neural Inf Process Syst:1097–1105, 2012.

18. Simonyan K, Zisserman A: Very deep convolutional networks forlarge-scale image recognition. arXiv preprint arXiv:1409.1556,2014.

19. zegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, ErhanD, Vanhoucke V, Rabinovich A: Going deeper with convolutions.Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit, 2015.

20. He K, Zhang X, Ren S, Sun J: Deep residual learning for imagerecognition. Proc IEEE Comput Soc Conf Comput Vis PatternRecognit 770–778, 2016.

21. YOLOv4: Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao," Optimal Speed and Accuracy of Object Detection", April 2020.

22. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," inCVPR, 2016.

23. G. Han et al. The LISS—a public database of common imaging signs of lung diseases for computer-aided detection and diagnosis research and medical education, IEEE Trans. Biomed. Eng. 62 (2) (2015) 648–656.