MINAR International Journal of Applied Sciences and Technology

Article type : **Research Article**

Date Received : **05/07/2021**
Date Accepted : **12/08/2021**
Date published : **01/09/2021**

: **www.minarjournal.com**

http://dx.doi.org/10.47832/2717-8234.3-3.10

COMPARISON BETWEEN DIFFERENT METHODS FOR IDENTIFYING LESION IN PULMONARY X-RAY IMAGES

Rana M. HASAN[1] & Ielaf O. Abdul MAJJED[2]

Abstract

The image retrieval system is one of the most prevalent and challenging systems of deep learning. To perform image retrieval for lung disease radiography systems, three methods were used: Firstly, we built a convolution neural network from scratch to extract and classify features by using six convolution layers and two fully connected layers. Secondly, it trained the feature patterns and classified categories by transfer learning techniques (Resnet_50). Thirdly, by training feature patterns by (inception V3) and classifying them by Support Vector Machine (SVM). After the system retrieval set of images depending on the class labels, these methods were compared to find the most accurate and fastest method among them. The concluded from the results of our proposed system that the accuracy of CNN from scratch was better than the learning methods (96.2%), but Resnet-50 was faster than the other methods and had good accuracy (94.81%).

Keywords: CBMIR, CMIR, Deep learning, Transfer learning, CNN.

[1] Mosul University, Iraq, rana.csp45@student.uomosul.edu.iq, https://orcid.org/0000-0002-8440-0508
[2] Mosul University, Iraq, ie_osamah@uomosul.edu.iq, https://orcid.org/0000-0002-8520-0233

Introduction

Content-Based Image Retrieval (CBIR) is a computer vision system which enables big database image search algorithms. The search is based on the picture properties (texture, shape, color). The main factors for a central photo recovery system are these characteristics[1][2].Several layers of the deep learning technique are utilized to extract characteristics and classify them to simulate the human brain and include much information. Complex learning methods are a simple approach to learning deep structural levels[1]. In recent research, DL methods were effective for theContent-Based Medical Image Retrieval(CBMIR)challenged. In the first phase, the neural network model was trained in the classification of medical images, and in the second phase, the features gained in the CBMIR are represented. It offered a deep analysis of the method for the quality of recovery of various taken photos. The semantic gap-deeper learning seeks to address this gap by direct learning of the visual properties of pictures without needing handmade functions. A new technique in machine learning has offered a fresh approach.Deep neural networks with enough neurons and layers can simulate any non-linear, multivarian connection. This allows the processing of complicated data, such as digital signals (images, sounds, videos), to cluster and generate processes [3] [4]. CNN's main feature is the convolutionary layer. The input is combined with a few filters. Each filter can record a specific characteristic (e.g. corners and edges) [1] and the filter outputs a higher value each time it is detected in the images.The filter outputs are aggregated to create a new input and additional convolutionary layers [5]. The initial layers are capable ofcapturing fundamental geometric properties, whereas the later levels can model high-level semantics and complicated forms. Current layers will normally be supplemented by a pooling layer in order to reduce the size of the input [6][7] and non-linear function (sigmoid, straightened linear, hypersensitive tangent) to provide modeling capability. The purpose of the CNN model (training CNN from scratch, retrained Inception V3 Net, and ResNEt-50) in this system is to classify and retrieve lung disease imagesinto four classes (pneumonia, covid-19, lung cancer, and normal lung).

1- Related work

Since medical images have been frequently used in recent years, medical images are returned using query, as these systems are among the necessary systems to assist doctors in diagnosing and quickly accessing the required information by searching for similar images in a huge database of medical images, and this matter is not manually efficient, so the need arose to retrieve medical images based on content (CBMIR) [9].We describe some prior research on our approach utilized in this article (2017-2020) in this sector in order to understand the development of this area:

A- Convolution Neural network From Scratch

The CNN was utilized in the pneumonia detection in (2020) by the searchers (Rajaraman S. et al.). They have observed the information provided by studying relevant features helps to enhance classification in a particular way. Reduced prediction variability and sensitivity to training data variations [3]. A. Qayyum et al (2017) uses DCNN for the extraction of medical images, using five convolution layers for extracting maps that contain 64 kernels (11*11) with strands of four (cl1) and then used filter size (5*5) for the kernels first to (192) and (384) of size (95*5) to cl2 and cl3 and (256) kernels of size (3*3) for the lats two layers [4], In 2019, it also used radiological images to identify cancer lung. [5].

B- Resnet-50

Haikal Abdulah et al In utilized Res-CR-Net (2020). For extraction and classification, they employed a new form of completely convolutionary. They determined that ~97 % was obtained utilizing JMS dataset. It was 94% in the V7 dataset [6]. In 2019, 2,048 GPUs were utilized by authors to achieve their test findings. The self-optimized deep learning framework of AI Bridging Cloud Infrastructure (ABCI) has succeeded, the validation accuracy of ResNet-50 on ImageNet is 75.08% in 74.7 seconds for 81,920 mini-batch size [2].

C- Inception-V3

In 2018, D.S. Kermany et al. applied transfer learning on 108309 pictures on image.net with qualified-weight starting CNN v3, when compared to human pneumonia detection specialists, the approach of CNN was highly sensitive, but of poor specificity, while the algorithms of the first V3 were both. Their technique, after testing it on a small dataset of around 5000 chest X-ray pneumonia images [7], has an average accuracy of 92,8 percent with sensitivity of 93,2 percent and a 90,1 percent specificity. V. P. Vianna(2018) utilized the X-Ray dataset chest in two types: pneumonia and normal case where CNN models are employed on the PyTorch platform PyTorch available Alex net, ResNet, Squeez Net, and Inception [9].

2- Methodology

MIRS performs many operations, including feature extraction and training, then classifies, and finally retrieval a set from the radiographs that are of the same type as the one used to diagnose the disease. Our system data chart (fig. 1) shows the main steps of the system operation :

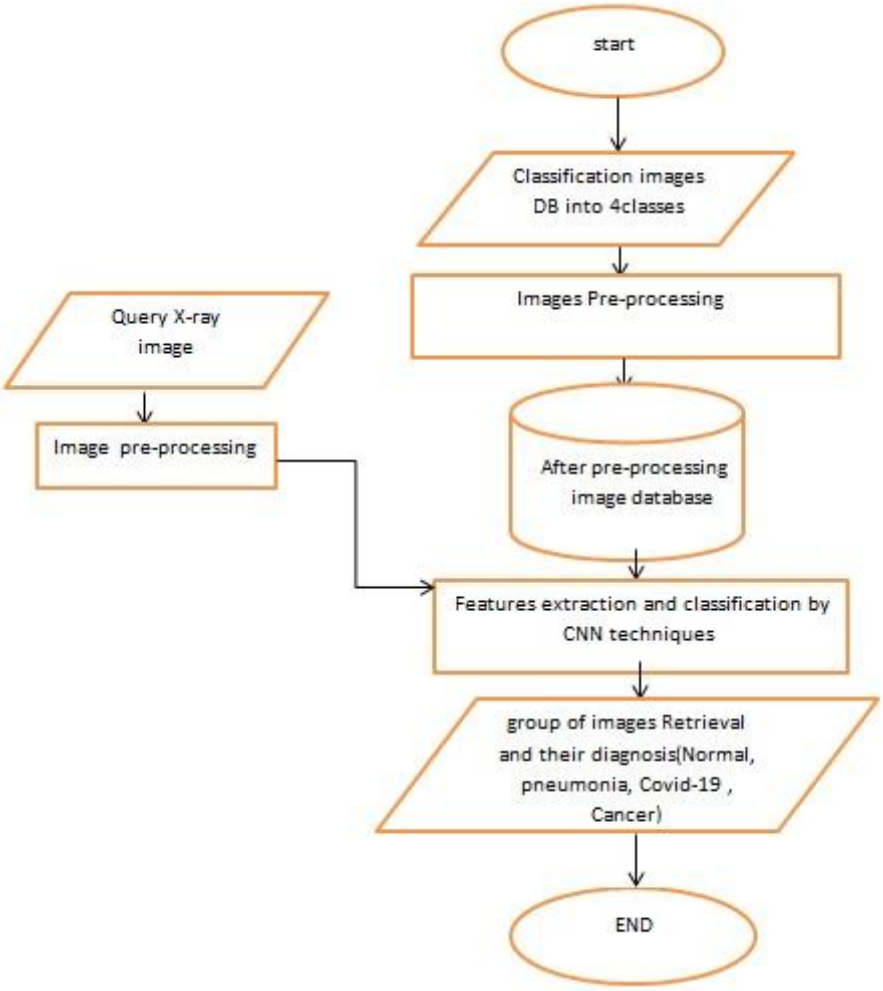

Fig 1: General system of DMIRS

3-1 Databases Image Collection

The accuracy of classification procedures is assessed in lung disease databases to be as realistic as possible. For decades, researchers in this field have been compiling databases of photos of lung disorders. We used three of these datasets (Nodules in chest X-rays data set (JSRT) in the Chest (XSRT) [10], Covid-19 chest X-ray dataset [11], and Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification [12])collected (5957) chest x-ray images divided as follows: (265) lung cancer images, (460) coronavirus images, (3883)

pneumonia images, and (1349) healthy lung images.We mention here the images databases in our thesis:

a) Nodules in chest X-rays data set (JSRT)

Dataset originating from Japan, year ~2000 The chest X-rays that are given are scanned films from a high-quality digital camera. Each nodule case comprises a single nodule that is assessed by 20 independent radiologists, with AUC ranging from 0.71 to 0.89. This is an ideal dataset to test nodule detection performance on different levels of nodule subtlety. Database's images are available in png with resolution 2048 by 2048 gray level scale value. From this database, we used only radiographs of malignant nodules to diagnose Lung cancer[10].

b) Covid-19 chest x-ray dataset

We used a database COVID-19 (coronavirus disease 2019) is an infectious illness caused by the coronavirus strain severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The initial instances were discovered in late December 2019 in Wuhan, China, before spreading around the world. On March 11, 2020, the World Health Organization (WHO) declared the current epidemic a pandemic.

- Currently Reverse transcription polymerase chain reaction (RT-PCR) is used for diagnosis of the COVID-19. X-ray machines are widely available and provide images for diagnosis quickly so chest X-ray images can be very useful in early diagnosis of COVID-19.
- The dataset is divided into two files (train and test), each with three subfolders (COVID19, PNEUMONIA, and NORMAL). The DataSet comprises a total of 6432 x-ray pictures, with test data accounting for 20% of the total images.
- Acknowledgements
- Database's image are available jpg with multiples values resolution gray scale value.
- Images are gathered from a variety of publically available sources.
- Sources:
- https://github.com/ieee8023/covid-chestxray-dataset.
- https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia
- https://github.com/agchung

We used only covid-19 radiography cases from this dataset[11].

c) Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification

We used a database of pneumonia the source of this data is Kermany, Daniel; Zhang, Kang; Goldbaum, Michael (2018), “Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification”, Mendeley Data, V2, doi: 10.17632/rscbjbr9sj.2.

Database's images are available in jpeg with resolution multiple values gray level scale value. This dataset contains pneumonia cases and normal cases, we used both of them[12].

3-2 Image pre-processing

The collection of images that used in the training have been generated from different models of X-ray imaging pulmonary illnesses, which vary in size, color and background. Thus, image processing requires preparatory treatment, in order to accurately differentiate between diseases.

a) The process for CNN consists of the following:

- Converting images from color (RGB) to grayscale(Fig. 2).

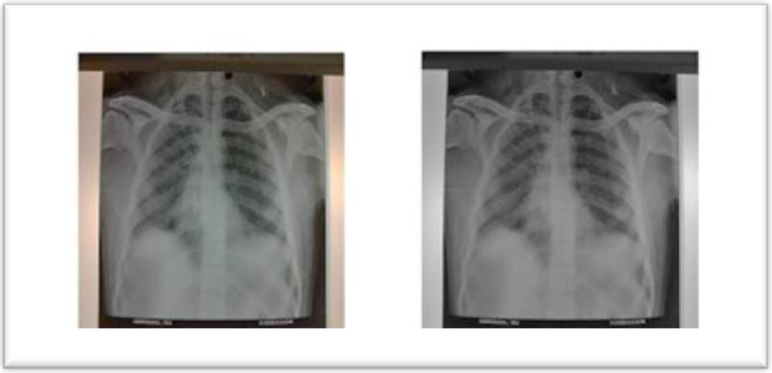

Fig. 2. Converting color (RGB) to grayscale

- Crop the Patient information from X-ray (Fig. 3).

Fig. 3. Images after and before cropping

- Standardization of all used pictures Interpolation:Bicubic interpolation is one of the algorithms used to change the image size or resolution. The pixels that are closest to the one being estimated are given larger weights than those that are further away. The image is resized to 201x173 for comparison. (Fig. 4).

Fig. 4. Images after and before resize.

- For image enhancement and to solve the contrast problems, Contrast Limited Adaptive Histogram Equalization (CLAHE) was used (Fig.5).

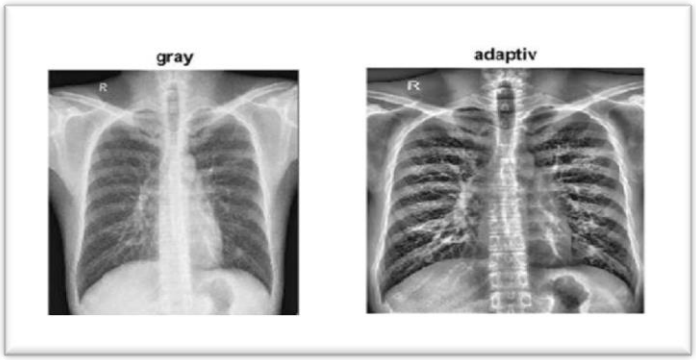

Fig.5images before and after CLAHE

b) The process for Resnet-50

The images are converted to color images, resized to (224x224), augmented image data store was used, and Resnet-50 model is installed.

c) The process for Inception v3

All images are converted to color images, resized to (299x299) and inception v3 model is installed.

3-3 Define Network architecture

Training of Convolution Neural network (Fig.6) consisted of eight layers, six of which were convolutional layers (CL) and two of which were fully connected layers (FCL), where the signifies the layer number, such as CL1 for the first convolutional layer. The latest FCL (FCL2) output has been provided with a softmax function with 4 outputs, which creates a distribution of probability for each class label. Hence, a probabilities vector of size (1×4) where each vector element corresponds to a class of the dataset is obtained. The network accepts grayscale images of dimension (201×173) as inputs. The CL1 filters the input image with 8 kernels of size (5×5) with stride equal to 2 pixels. A stride is the distance between the centers of receptive fields of neighboring neurons. The CL1 filters the input image in the kernel map, the output of the first convolutional layer is input into a non-linearity (RelU), which is subsequently used to sum neighboring neurons using the spatial max pooling layer. This network with ReLUs can be trained several times quicker than a network with tanh units, and it can also solve vanishing gradient problems in greater depth. CL2 filters the output of CL1 with 16 kernels of size (5x5) that have been processed by ReLU non-linearity and spatial max pooling layers. The CL3 is made up of 32 kernels of size (5x5) that obtain their input from the second convolutional layers' pooled outputs. CL4 has 64 kernels of size (5x5) that obtain their input from the third convolutional layers' pooled outputs, CL5 has 128 kernels of size (5x5) that obtain their input from the fourth convolutional layers' pooled outputs and CL6 has 256 kernels of size (5x5) that obtain their input from the fifth convolutional layers' pooled outputs. A pooling layer was utilized after the convolutional layers. The model can't overfit during training because of the overlapping pooling operation. The FCL2 output is fed into the log softmax input, which has N Inputs and four outputs. With Resnet-50 there are 50 depth layers and 48 depth layers in the inception v3.

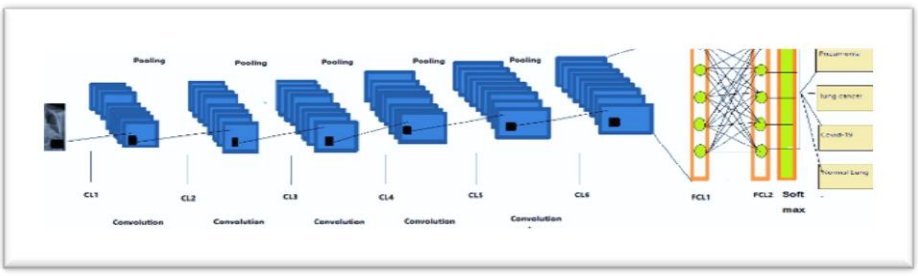

Fig. 6 architecture of proposed CNN

3-4 Training networks and results

Lung disorders are diagnosed and separated using the CNN algorithm with two implementation techniques: the CNN training method from scratch and the transfer learning method, where we used two TL (start-up) approaches (inseption V3 and Resnet-50).

In CNN from scratch, we have a good number of images for training, (80% for training was taken randomly and 20% for testing). Depending on the images used, the number of convolutional layers and the parameters of each layer are calculated to avoid over-processing and network performance was evaluated at each level. Then the most efficient combination has been selected. A random regression algorithm with momentum algorithm (SGDM) was used to train the network, with a learning rate of 0.001 and a maximum period of 30. The six-layer deep model with feature mappings of 8, 16, 32, 64, 128, 256 in CL1, CL2,..., CL6 with a filter size (5×5) is our model in terms of training and validation accuracy. This model was the most accurate compared to transfer learning methods, as the accuracy was surprising (95.2%) (Fig. 7). While in the Resnet-50 dataset was split 30% for testing and 70% for training, while in Inception-v3 with SVM classifier the dataset was split into 20% for testing and 70% for training. The network accepts an image as input and outputs a label for each object in the image, as well as probabilities for each type of object. With the help of the fitcecoc function, the proposed system used SVM as aclassifier. In both ways, contacted with (avg_pool) layer for retrained new data using a single GPU.The accuracy was 94.8% for Resnet-50 (Fig.8) and 84.4% for Inception v3 (Fig. 9).

Fig. 7 Confusion matrix of CNN

The accuracy and loss obtained by the training process in our CNN model were (96.2%), (3.8%) with a learning rate of (0.001) and 30 epochs. Whereas lung cancer disease achieved a positive predictive (precision) value and negative predictive value rate (98.1%) and (1.9%), was sensitivity and specificity (96.2%) and (3.8%), Covid-19 disease got a positive predictive and negative predictive value rate (94.5%) and (5.5%), got sensitivity and specificity (98.1%) and (1.9%). pneumonia disease got a positive predictive and negative predictive (96.2%) and (3.8%), got sensitivity and specificity (94.3%) and (5.7%), and normal lung got a positive predictive and negative predictive (96.2%) and (3.8%), got sensitivity and specificity (96.2%) and (3.8%).

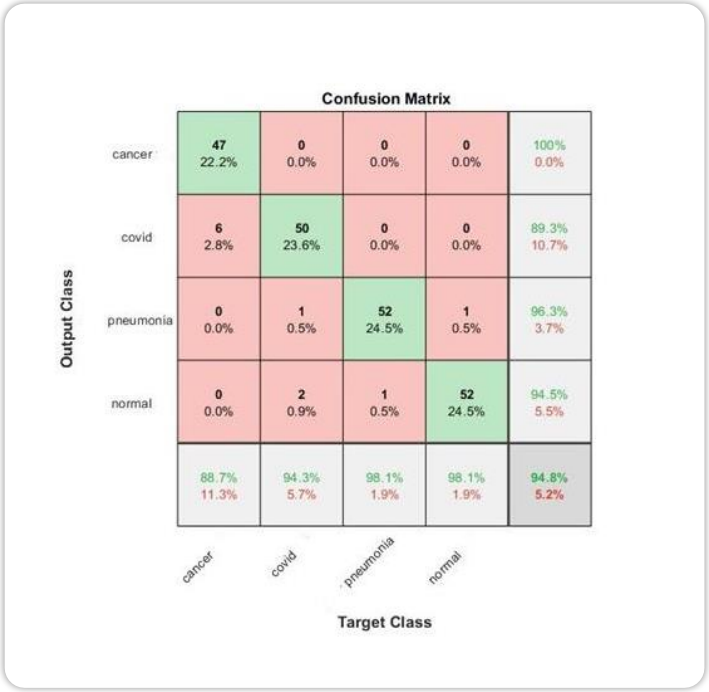

Fig.8 Resnet-50 confusion matrix

The accuracy and loss obtained by the training process in Resnet-50 model were (94.8%), (5.2%). Where lung cancer disease got a positive predictive value (100%) and negative predictive rate (0%) and got (88.7%) sensitivity and (11.3%) specificity, Covid-19 disease got (89.3%) positive predictive value and negative value rate (10.7%), and got (94.3%) sensitivity and specificity (5.7%). Pneumonia disease got (96.3%) positive predictive value and negative predictive (3.7%), and got (98.1%) sensitivity and specificity (1.9%), and normal lung got (94.5%) positive predictive value and negative predictive (5.5%) and got (98.1%) sensitivity and specificity sensitivity and specificity sensitivity and specificity (1.9%).

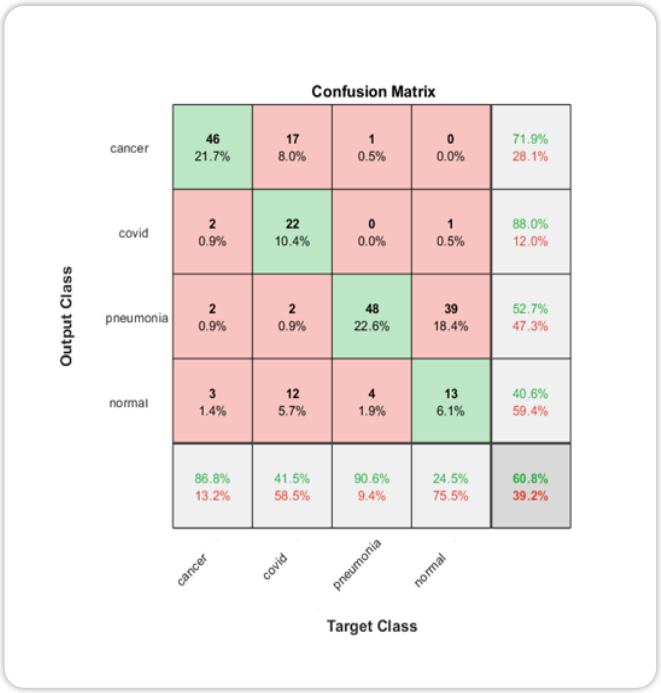

Fig.9 inception v3 confusion matrix.

The accuracy and loss obtained by the training process in the Inception V3 model were 60.8%, 39.2%. Where lung cancer disease got (71.9%) positive predictive value and negative predictive value rate (28.1%), gotthe sensitivity and specificity (86.8%) and (13.2%), the Covid-19 disease got (88%) positive predictive value and negative predictive value (12%), got the sensitivity and specificity (41.5%) and (58.5%). Pneumonia disease got (52.7%) positive predictive value and negative predictive rate (47.3%), got the sensitivity and specificity (90.6%) and (9.4%), and normal lung got (40.6%) positive predictive value and negative predictive (59.4%) and got the sensitivity and specificity (24.5%) and (75.5%).

A complete confusion matrix for each disease (table 1) reveals that, with Resnet-50, the accuracy rate for cancer is greater (100%), while with inception V3 is the lowest (67.9 %) (Fig. 10).

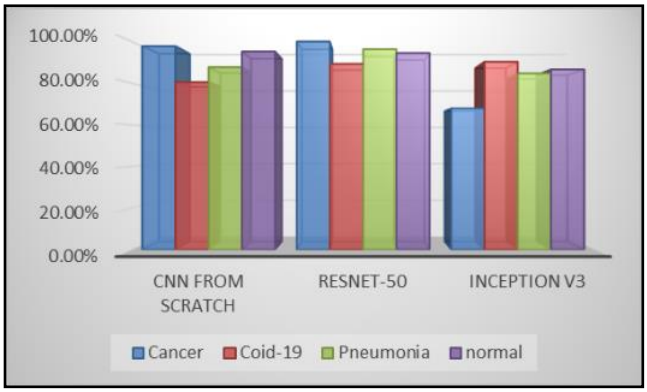

Fig.10positive predictive ratio for each disease and all methods

Table 1: The accuracy rate for each disease by each method

Methods	Precision			
	Cancer	Covid-19	Pneumonia	Normal
CNN From Scratch	97.8%	80.5%	87.7%	95.2%
Resnet-50	100%	89.3%	96.3%	94.5%
Inception V3	67.9%	90.3%	84.7%	86.6%

This table shows each diseases, positive predictive in all methods (CNN from Scratch, Resnet-50, Inception V3). Denoted the best Precision for cancer in Resnet-50, Covid-19 in CNN, pneumonia in Resnet-50, and normal lung in CNN.

3- Conclusion

We conclude that the training of gray images needs lower layers than color images and reduces computer time while building convolutionary neural networks from scratch. Resent 50 has greater accuracy than linked to the fc 1000 layer when we are connected to the "avg pool" layer. When connected to the "avg pool" layer, the accuracy at inception v3 is greater than when attached to "projections." Given that the data are split up for training and testing to the size of the lowest classification set, we have discontinued a number of images about a small lung disease that has unfavorable effects on network training accuracy.

References

S. Hijazi, R. Kumar, and Chris Rowen, “Using Convolutional Neural Networks for Image Recognition,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 8925, pp. 572–578, 2015, doi: 10.1007/978-3-319-16178-5_40.

M. Yamazaki et al., “Yet Another Accelerated SGD: ResNet-50 Training on ImageNet in 74.7 seconds,” 2019,[Online].Available:http://arxiv.org/abs/1903.12650.

S. Rajaraman and S. K. Antani, “Modality-Specific Deep Learning Model Ensembles Toward Improving TB Detection in Chest Radiographs,” IEEE Access, vol. 8, pp. 27318–27326, 2020, doi: 10.1109/ACCESS.2020.2971257.

A. Qayyum, S. M. Anwar, M. Awais, and M. Majid, “Medical image retrieval using deep convolutional neural network,” arXiv, 2017.

Tami freeman, “Deep learning helps radiologists detect lung cancer on chest X-rays,” Phys. World, 2019.

H. Abdulah, B. Huber, S. Lal, H. Abdallah, H. Soltanian-Zadeh, and D. L. Gatti, “Lung segmentation in chest x-rays with res-CR-net,” arXiv, pp. 1–8, 2020.

D. S. Kermany et al., “Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning,” Cell, vol. 172, no. 5. pp. 1122-1131.e9, 2018, doi: 10.1016/j.cell.2018.02.010.

V. V. Estrela and A. E. Herrmann, ―Content-Based Image Retrieval (CBIR) in Remote Clinical Diagnosis and Healthcare,‖ Encycl. E-Health Telemed., pp. 495–520, 2016, doi: 10.4018/978-1-4666-9978-6.ch039.

V. P. Vianna, “Study and development of a Computer-Aided Diagnosis system for classification of chest x-ray images using convolutional neural networks pre-trained for ImageNet and data augmentation.” 2018, [Online]. Available: http://arxiv.org/abs/1806.00839.

Raddar, ―Nodules in Chest X-rays (JSRT) _ Kaggle.‖ Kaggle repostory, japan, 2000, [Online]. Available: https://www.kaggle.com/raddar/nodules-in-chest-xrays-jsrt?select=jsrt_metadata.csv.

J. P. Cohen, ―ieee8023/covid-chestxray-dataset: We are building an open database of COVID-19 cases with chest X-ray or CT images.,‖ GitHub. [Online]. Available: https://github.com/ieee8023/covid-chestxray-dataset.

D. Kermany, ―Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification,‖ Mendeley Data, vol. 2. 2018.