Lung cancer classification with Convolutional Neural Network Architectures

Shivan H. M. Mohammed
Computer Science
Duhok University
Duhok, Iraq
ORCID: 0000-0002-9173-4426
shivan.cs@uod.ac

https://doi.org/10.48161/qaj.v1n1a33

Ahmet ÇINAR
Computer Engineering
Firat University
Elazig, Turkey
ORCID: 0000-0001-5528-2226
acinar@firat.edu.tr

Abstract— one of the most common malignant tumors in the world today is lung cancer, and it is the primary cause of death from cancer. With the continuous advancement of urbanization and industrialization, the problem of air pollution has become more and more serious. The best treatment period for lung cancer is the early stage. However, the early stage of lung cancer often does not have any clinical symptoms and is difficult to be found. In this paper, lung nodule classification has been performed; the data have used of CT image is SPIE AAPM-Lung. In recent years, deep learning (DL) was a popular approach to the classification process. One of the DL approaches that have used is Transfer Learning (TL) to eliminate training costs from scratch and to train for deep learning with small training data. Nowadays, researchers have been trying various deep learning techniques to improve the efficiency of CAD (computer-aided system) with computed tomography in lung cancer screening. In this work, we implemented pre-trained Convolutional Neural Networks include: AlexNet, ResNet18, Googlenet, and ResNet50 models. These networks are used for training the network and CT image classification. CNN and TL are used to achieve high performance resulting and specify lung cancer detection on CT images. The evaluation of models is calculated by some matrices such as confusion matrix, precision, recall, specificity, and f1-score.

Keywords— Deep learning; Transfer learning; Lung cancer; pre-trained network; CNN.

I. INTRODUCTION

In recent years, in the field of image recognition and deep learning, especially convolutional neural networks have proved highly successful. It defeated other traditional machine learning methods when it first appeared and won the championship of the ImageNet large-scale image recognition challenge[1] in one fell swoop, significantly reducing image recognition Error rate. However, most of the deep learning models are only applied in the field of natural image recognition, and there are few applications in the areas of medical image diagnosis. The use of deep learning technology for lung cancer CT image diagnosis can significantly reduce the diagnosis time of doctors, improve hospitals' efficiency, effectively alleviate the shortage of medical resources and other problems, and early diagnosis,

early treatment, and even save lives. There are two strategies for applying deep learning to medical image diagnosis: first, Using medical images, training convolutional neural network models from scratch. Second, Transfer learning[2], using a pre-trained convolutional neural network model and weight parameters to extract features. However, in the diagnosis of lung cancer CT images, there is currently only a way to train from scratch, and no one has ever used transfer learning. So this study proposes a strategy for transfer learning in lung cancer CT image diagnosis and conducts experiments. The experimental results verify the transfer learning method has delighted results. Through learning, the model selects the accurate features from the training data so that when testing new data, it can make correct decisions. Therefore, deep learning plays a crucial role in medical image processing.

In recent years, deep learning has continued to make significant progress, mainly due to the continuous improvement of computing power and the continuous increase in the amount of available data, as well as the continuous improvement of deep learning models and algorithms. Since 2006, many convolutional Neural Network architectures have developed to overcome the problems encountered earlier. Some of CNN architectures proposed in this paper such as AlexNet[3], ResNet18[4], Googlenet[5] and Resnet50.

II. LITERATURE REVIEW

In the literature on lung cancer diagnosis, there are several studies.

E. Cengil et al. [6] implemented deep learning methods to classify lung cancer detection that has implemented TensorFlow libraries and 3D-CNN architecture from deep learning. It is also used a SPIE-AAPM-LungX dataset of CT images. The accuracy of this model after evaluated it is 70%.

Z. Shi et al.[7] has proposed a deep Convolutional Neural Network (DCNN) based on transfer learning for detecting a pulmonary nodule on CT slices, it utilized one of the pre-trained networks in Convolutional Neural Network, VGG-16 is used for feature extraction, and this work used support vector machine (SVM) to classify a nodule.

Rushil Anirudh et al. [8] proposed the three-dimensional convolutional neural network CNN architecture for lung nodule detection. To train the CNN, they used unsupervised segmentation to enlarge the 3D region. Evaluation of the results is done with data from SPIE-LUNGx.

Q. Z. Song et al. [9] preferred deep learning with three deep neural networks such as CNN, Deep Neural Network (DNN), and The Sparse Autoencoder (SAE); they are configured for classifying lung cancer. These networks are applied with some modification to the CT image classification task for the benign and malignant nodules of the lungs.

In R. M. Devarapalli et al. [10] Different image processing methods have been innovated for detecting cancer and Implemented as a median-wiener filter in the preprocessing step. To detect whether the nodule is cancerous or not, the classification network has been used, such as Support Vector Machines (SVM), Forward Neural Networks, Back Propagation model, and Convolution Neural Networks (CNN).

V. Makde et al. [11] proposed two frameworks of CNN (Alexnet and ZFNet), to train the system for tumor detection in lung nodules as well as brain tumor. This study used two different datasets such as Lung CT image and Brain MRI image with obtaining more than 97 % of the accuracy in classification training data.

Through literature studies have been conducted with lots of different algorithms for deep learning; owing to its popularity, academic and science environments have recently adopted deep learning. In the research, the lung cancer classification is presented, demonstrating early detection of lung cancer. For processing Convolutional neural networks have favored. Some of the popular pre-trained networks are used, and Transfer Learning [2] for implementing our work.

The studies structure has described in; section 3, illustrates the structure of dataset and Methods. Section 4 describes the Result and discussion. Finally, section 5 defines the conclusion.

III. METHODOLOGY

A. Dataset

Diagnosis of lung cancer has based on clinical databases. The LIDC-IDRI dataset is the most commonly used data from this kind of datasets [8]. Dataset consists of a screening of lung cancer and CT scans. Seven research centers and eight medical imaging companies cooperated to build a dataset of 1018 cases. Furthermore, it also organizes competitions to improve the accuracy of classification. The SPIE-AAPM Lung Challenge dataset is a subset of the medical imaging conference at SPIE in 2015 with the support of American Association of Physicists in Medicine (AAPM) and the National Cancer Institute (NCI). To identify the pulmonary nodules as benign or malignant, the use of a standard dataset has proposed to test the competitors more precisely. SPIE-AAPM dataset [12] is utilized in our research. The dataset includes 70 patient CT images. Ten of these cases have been used for training and the remaining 70 for testing. In this implementation, we used the data augmentation technique to increase the number of CT image artificially from a small dataset to thousands of images. Figure 1 shows some examples of our dataset.
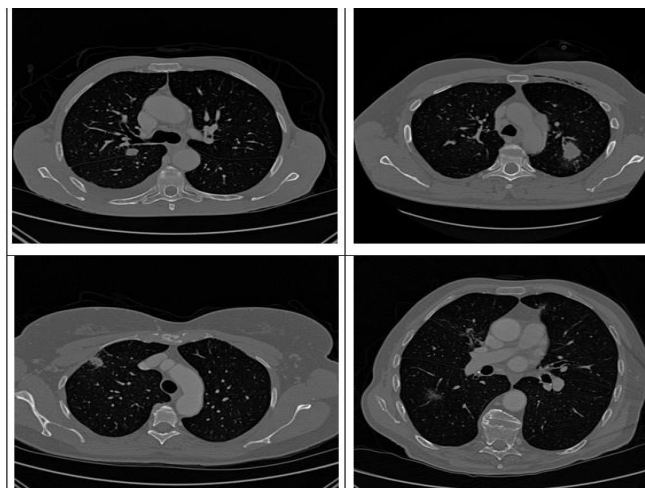

Figure.1 Lung cancer and non-cancer samples

B. Transfer Learning and CNN architectures

Transfer learning [2] is a prevalent approach in the area of computer vision because it can construct accurate models, and less time is needed. Using transfer learning is not to start learning from scratch, but to start from the model learned when solving various problems. In this way, we can use previous learning results and avoid starting from scratch. Deep CNNs are still commonly used in present-day research. They provide creative assistance to overcome many challenges relating to classification. Lack of data from training is a common issue when using deep CNN models that need a significant number of data to perform well.

Furthermore, it is tedious to collect a vast dataset, and it continues even now. Thus, the transfer learning approach has generally used to solve the limited data collection issue [2]. Transfer learning is a method where CNN models have trained on datasets with a massive number of data, and afterward, the models have fine-tuned to train on a small required dataset. Transfer learning is a useful technique where we might use a pre-trained model like (AlexNet, ResNet18, Googlenet, and ResNet50) and adjust the network for the next application by making specific improvements in the architecture of the network [13]. AlexNet and the other pre-trained networks have trained for 1000 classes of real-world images. We can use this network to identify some other classes with some adjustment to the network (i.e. in our research, we used two classes Benign and Malignant, instead of 1000 classes; furthermore, we modified the last layer and fully connected layer to adjust for our work.

The mechanism of using any pre-trained models will be described in figure 2. The common pre-trained models like (AlexNet, ResNet18, Googlenet, and ResNet50) have proposed for well-organized classification. In this study, all models have altered the last three layers to adjust the new image classification. We modified each model as in the following:

- AlexNet: the last three layers of the constructed network with a group of layers are changed fully connected layer (FC), softmax layer, and output layer (classification) to classify images into relevant classes.

- ResNet18: The network layers (fc1000, prob, and ClassificationLayer_predictions) are replaced with fully connected layer, softmax layer, and classification output layer. Afterward, the last remaining transferred layer on the network (pool5) is linked to the novel layers.

- GoogleNet: also, the last three layers of the network are modified. The layers loss3-classifier, prob, and classification output layer are adjusted with a fully connected layer, softmax layer, and an output layer. Later, the last transferred layer still existing on the network (pool5_drop7x7_s1) is connected to the new layers.

- ResNet50: The network's (fc1000, fc1000_softmax, and ClassificationLayer_fc1000) layers are replaced with fully connected layer, softmax layer, and classification output layer. Afterward, the last remaining transferred layer on the network (avg_pool) is linked to the novel layers.
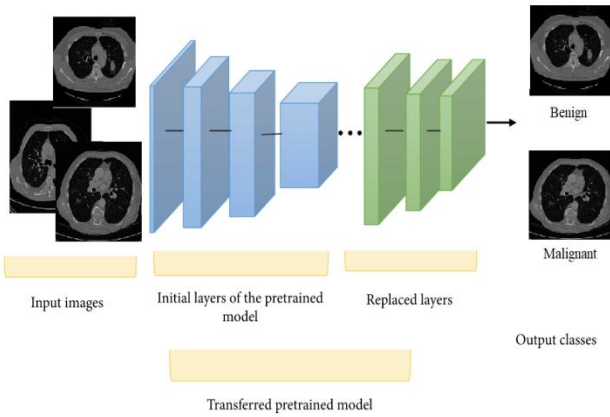

Figure.2 Transfer Learning Process using pre-trained models

Table 1 is described as the pre-trained models characteristics.

TABLE 1. PROPERTIES OF PRE-TRAINED NETWORKS

Network	Depth	Parameters (Millions)	Image Input Size
AlexNet	8	60	227×227
ResNet 18	18	11.7	224×224
GoogleNet	22	7	224×224
ResNet 50	50	25.6	224×224

permit a reasonable comparison among the experiments, an initiative also is made to optimize the hyper parameters throughout the experiments, the following hyper-parameters described in table 2 has been used.

TABLE 2. PRE-TRAINED NETWORK'S HYPER-PARAMETERS AND VALUES

Hyper-Parameters	Value
Optimization algorithm	SGDM *
Momentum	0.9000
Learning rate	0.0001
Epochs	20
Batch size	64

* Stochastic Gradient Descent with Momentum (SGDM) [14].

C. Data augmentation

Deep learning is one of the best choices in the field of image processing. However, its medical imaging application has limited the massive need for high-quality labeled images as training samples. It is costly to collect medical images, which needs a specialist to label the images. To address this problem, we took advantage of Transfer Learning and data augmentation. Data Augmentation is a technology that artificially expands the training data set by allowing limited data to generate more comparable data. It is an effective means to overcome the lack of training data and avoid over-fitting issues [15]. In this study, we have used some image processing operations such as the rotation technique to enhance the training data. Rotating has implemented by applying angles of 45, 90, 135, 180, 225, 270, 315 degrees, as shown in fig.3 for two reasons[16]; we have chosen with rotation over other techniques. Firstly, doctors can examine histological images of Lung cancer from different angles without impacting the process of diagnosis. Therefore, utilizing data augmentation using rotation moreover improves the dataset. Secondly, the rotation technique has enlarged the dataset's size without impacting the quality of the input images [17].
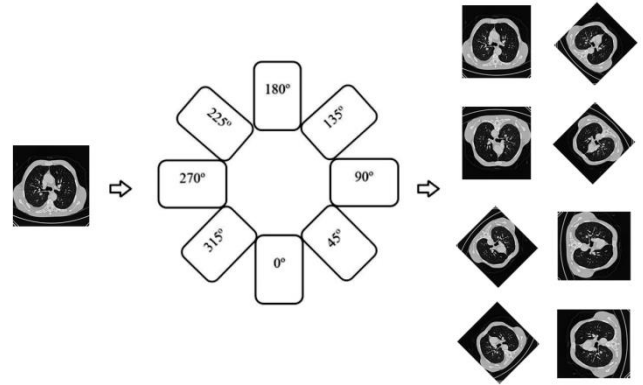

Figure.3 Rotation Process in different angels

D. Evaluation matrix

The classification process's performance was evaluated by several matrices such as confusion matrix (accuracy) equ1, recall (sensitivity) equ2, precision equ3, specificity equ4, and F1-score equ5. The confusion matrix (accuracy of the model) is the most basic, intuitive, and easiest way to measure models' accuracy [18]. The parameters and equations have defined as:

True Positive (TP): predict a complimentary class as a positive class number.

True Negative (TN): predict a negative class as a hostile class number.

False Positive (FP): predict a negative class as a positive class number.

False Negative (FN): Predict the complimentary class as a hostile class number.

$$accuracy = (TP + TN)/ (TN + TP + FP + FN) \quad (1)$$

$$Recall\ (sensitivity) =\ TP\ /\ (TP + FN) \qquad (2)$$

$$Precision\ =\ TP\ /\ (TP + FP) \qquad (3)$$

$$Specificity\ =\ TN\ /\ (TN + FP) \qquad (4)$$

$$F1 - score =\ 2 * (Precision * Recall)\ /\ (Precision + Recall) \qquad (5)$$

IV. RESULTS AND DISCUSSION

Our research SPIE-AAPM dataset has been used; the dataset is a part of the medical imaging conference at SPIE in 2015. It Includes 22,489 pulmonary CT images with the real data that is the position of the tumor and other details of which patient has a cancer tumor and which patient has a non-cancer tumor. Deep learning needs a massive amount of data to correctly work on the training dataset, so that Data augmentation has been used to increase the number of labeled data artificially. Transfer learning can decrease the required training data and reduce the training time for Deep Learning.

In this phase, the dataset used are presents two classes of Lung infection; the dataset was divided into 70% for training and 30% for testing. Therefore, four pre-trained models, namely AlexNet, ResNet18, GoogleNet and ResNet50 were trained by the concept of fine-tuning, which consists of replacing the last three layers, where the final output layer has to be well-matched with the number of classes.

These pre-trained models are evaluated by different metrics such as accuracy, precision, recall, and F1-Score. As shown in Table 3, which offers a general result performance, AlexNet achieved better results than the other architectures. Besides, Googlenet, shows low performance.

An overview of pre-trained networks (AlexNet, ResNet18, Googlenet and ResNet50) for the role of classifying Lung cancer infections using CT images was done. The purpose of this study is to compare the Convolutional Neural Network evaluating the accuracy, recall, precision, and f1-score by fine-tuning. The comparison results are shown in table 3.

TABLE 3. PRE-TRAINED NETWORKS WITH COMPARISON RESULTS

Performance Measures	Alexnet	Resnet18	Googlenet	Resnet50
Accuracy	98.52	97.97	94.10	97.05
Recall	99.63	98.89	94.10	98.52
Precision	97.47	97.10	94.10	95.70
Specificity	97.41	97.04	94.09	95.57
F1-score	98.54	97.99	94.10	97.09
Time (min.)	67	192	209	525

All the models demonstrated similar and statistically significant performance. Starting with the accuracy matrix, Alexnet obtained a higher result with 98.52 %, followed closely by Resnet18, Goolglenet, and Resnet50 with 97.97 %, 94.10 %, and 97.05 %, respectively. On the other hand, the precision matrix, the higher percentage, is obtained by Alexnet with 97.47 %, and the lowest one is by Googlenet with 94.10 %, while the result of Resnet18 is 97.10 % and the result of Resnet50 is 95.70 %. Finally, in the f1-score, specificity, and sensitivity measurements, the performance of googlenet was poor, with 94.10 %, 94.09 %, and 94.10 %, respectively. In contrast to Alexnet that obtained the finest percentage in the evaluation matrix with 98.54 %, 97.41 % and 99.63 %, respectively. Moreover, according to the processing time for every CNN architectures to perform the classification methods, Alexnet presented the best performance in taking the shortest time with 67 minutes, followed closely by Resnet18 with 192 minutes, while the extended classification of processing time was for Googlenet and Resnet50 with 209 and 525 minutes respectively. Additionally, figure 4, 5, 6 and 7 demonstrate the confusion matrix of the pre-trained models respectively with the best consequence based on the performance measures, which is Alexnet. Depending on the results, it is possible to visually evaluate the classifier's performance and determine which classes are spotlighted by the neurons of the Alexnet model. The rows are present with the output class, while the columns are present to the actual class.
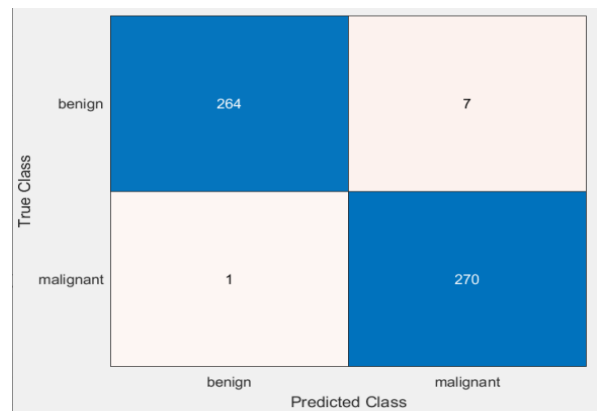

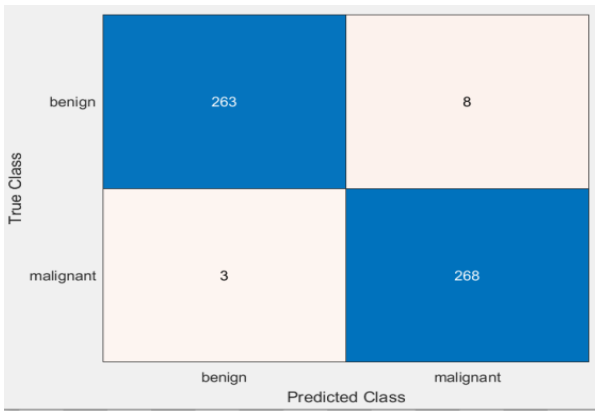Figure 4. Alexnet confusion matrix

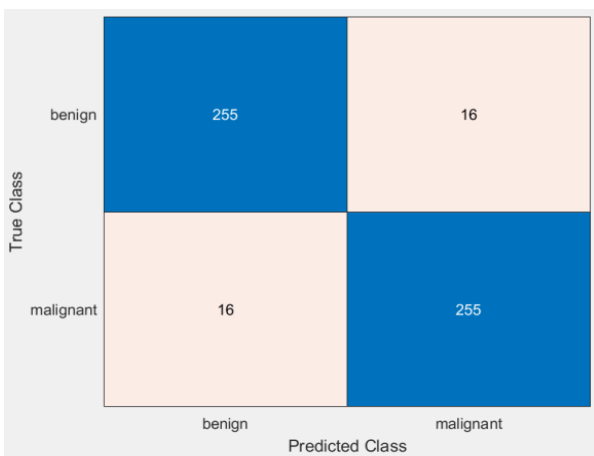Figure 5. Resnet18 confusion matrix
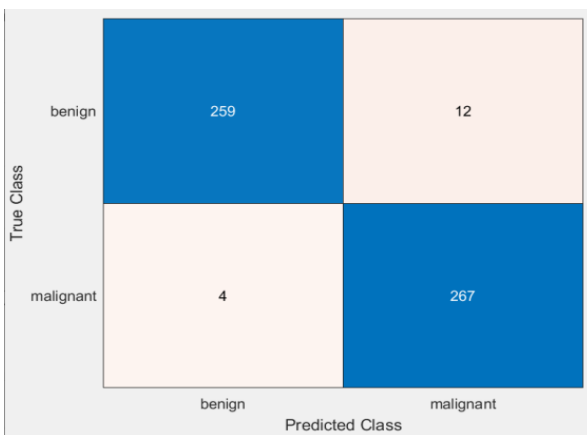

Figure 6. Googlenet confusion matrix

Figure 7. Resnet50 confusion matrix

Finally, the learning curve is a widely used diagnostic tool in deep learning, an algorithm for incremental learning from training data sets. After each update during training, the model can be evaluated on the training data set and the persistence validation data set, and a graph of the measured performance can be created to show the learning curve. Our CNN pre-trained models Figure 8, 9, 10 and 11 (AlexNet,

ResNet18, Googlenet, and ResNet50) respectively are used learning curve and loss function i.e., the loss function is a fault prediction of the model's performance; as we notice in Figure 8, 9, 10 and 11, the loss curve tends to zero through every epoch.
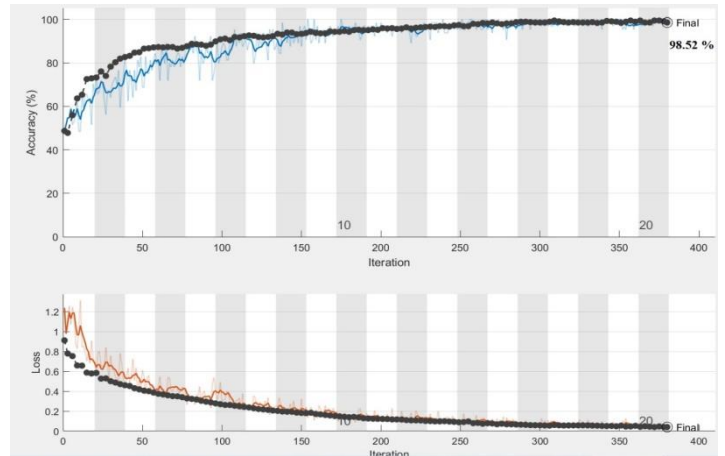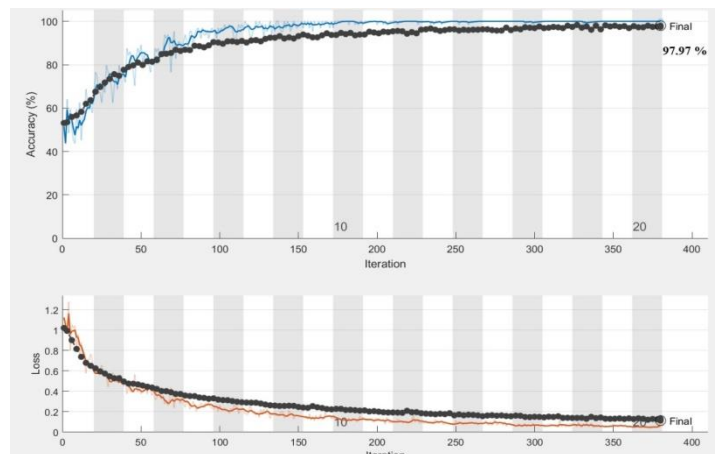

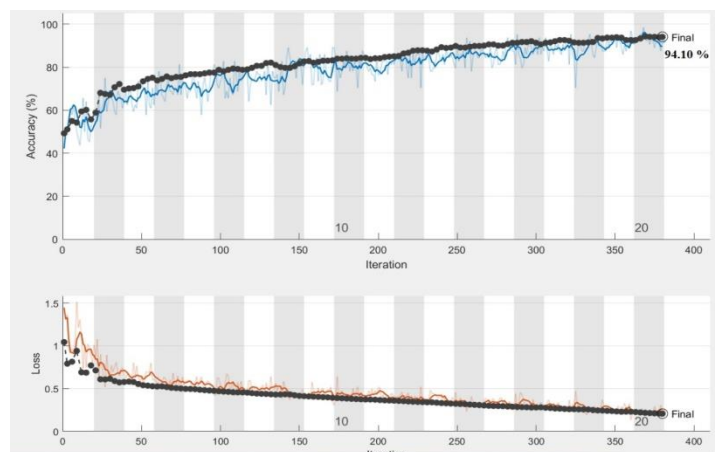Figure 8. Alexnet network's curve

Figure 9. Resnet18 network's curve

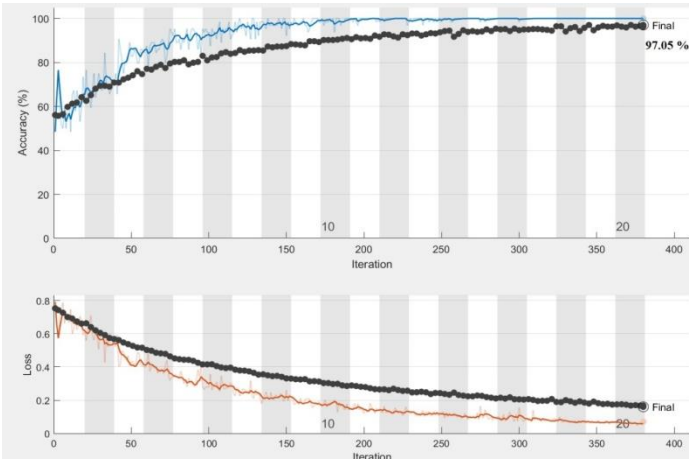
Figure 10. Googlrnet network's curve

Figure 11. Resnet50 network's curve

Raul V. et al.[20] Proposed CNN-ResNet50 with SVM-RBF methods; used 1536 CT samples with 20% testing and 80% training obtained 88.41% accuracy. Our pre-trained networks (Alexnet, Resnet18, Googlenet and Resnet50) obtained (98.52%, 97.97%, 94.10% and 97.05%) accuracy correspondingly.

T. K. Sajja et al.[21] Used AlexNet GoogleNet ResNet50 and densely connected architecture they tested on 80% training and 20% testing dataset samples they got 100 %, 98.84 %, 100 %, and 100 % accuracy respectively.

TABLE 4. PROPOSED MODELS AND EXISTING APPROACHES COMPARISON

S. No	Dataset information			Existing Method(other methods)			Our Proposed Method			
	No. of samples	Training	Testing	Authors	Methods Used	Results	AlexNet	ResNet18	GoogleNet	ResNet50
1	1006	90 %	10 %	H. Jiang et al. [19]	Filter, Convolution Neural Networks	Sensitivity 94 %	Sensitivity 99.63 %	Sensitivity 98.89 %	Sensitivity 94.10 %	Sensitivity 98.52 %
2	1536	80%	20%	Raul V. et al.[20]	CNN-ResNet50 with SVM-RBF	accuracy 88.41 %	Accuracy 98.52 %	Accuracy 97.97 %	Accuracy 94.10 %	Accuracy 97.05 %
3	1006	80%	20%	T. K. Sajja et al. [21]	AlexNet GoogleNet ResNet50 and densely connected architecture	Accuracy 100 %, 98.84 %, 100 %, and 100 %	Accuracy 98.52 %	Accuracy 97.97 %	Accuracy 94.10 %	Accuracy 97.05 %

According to Table 4, the existing methods and our pre-trained networks are compared.

H. Jiang et al. [19] nominated Frangi filter and Convolution Neural Networks approach which are tested on 1006 CT scan of LIDC dataset with 10% testing and 90% training they obtained 94% sensitivity. The pre-trained networks (Alexnet, Resnet18, Googlenet and Resnet50) are tested on 1804 CT images with 70% training and 30% testing they got (99.63%, 98.89%, 94.10% and 98.52%)sensitivity, respectively.

V. CONCLUSION

As we all know, Deep learning the training of the model relies on a large amount of data. Therefore, the two techniques used in our study's first technique are Transfer Learning. The second is a data augmentation technique to tackle the lack of training data and over-fitting issues and not build a network from scratch. Through using dataset with fine-tuning models (AlexNet, Resnet18, Googlenet, and Resnet50), it could classify lung cancer data, i.e., Classification of Benign and Malignant. Our model's learning in 20 Epoch has obtained a high accuracy and very fit learning curve with suitable

training time. By applying these techniques and obtaining a high performance of the model evaluation, we early confirmed lung cancer detection to protect patients from death. In the future, we would compare our model to some other models and improved the data with the proposed model.

REFERENCES

[1] O. Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge," Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, 2015, doi: 10.1007/s11263-015-0816-y.

[2] S. J. Pan and Q. Yang, "A survey on transfer learning," IEEE Trans. Knowl. Data Eng., vol. 22, no. 10, pp. 1345–1359, 2010, doi: 10.1109/TKDE.2009.191.

[3] G. E. H. Alex Krizhevsky, Ilya Sutskever, "Handbook of approximation algorithms and metaheuristics," Handb. Approx. Algorithms Metaheuristics, pp. 1–1432, 2007, doi: 10.1201/9781420010749.

[4] V. Sangeetha and K. J. R. Prasad, "Syntheses of novel derivatives of 2-acetylfuro[2,3-a]carbazoles, benzo[1,2-b]-1,4-thiazepino[2,3-a]carbazoles and 1-acetyloxycarbazole-2- carbaldehydes," Indian J. Chem. - Sect. B Org. Med. Chem., vol. 45, no. 8, pp. 1951–1954, 2006, doi: 10.1002/chin.200650130.

[5] C. Szegedy et al., "Going deeper with convolutions," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 07-12-June, pp. 1–9, 2015, doi: 10.1109/CVPR.2015.7298594.

[6] E. Cengil and A. Çinar, "A Deep Learning Based Approach to Lung Cancer Identification," 2018 Int. Conf. Artif. Intell. Data Process. IDAP 2018, 2019, doi: 10.1109/IDAP.2018.8620723.

[7] Z. Shi et al., "A deep CNN based transfer learning method for false positive reduction," Multimed. Tools Appl., vol. 78, no. 1, pp. 1017–1033, 2019, doi: 10.1007/s11042-018-6082-6.

[8] R. Anirudh, J. J. Thiagarajan, T. Bremer, and H. Kim, "Lung nodule detection using 3D convolutional neural networks trained on weakly labeled data," Med. Imaging 2016 Comput. Diagnosis, vol. 9785, no. November 2017, p. 978532, 2016, doi: 10.1117/12.2214876.

[9] Q. Z. Song, L. Zhao, X. K. Luo, and X. C. Dou, "Using Deep Learning for Classification of Lung Nodules on Computed Tomography Images," J. Healthc. Eng., vol. 2017, 2017, doi: 10.1155/2017/8314740.

[10] R. M. Devarapalli, H. K. Kalluri, and V. Dondeti, "Lung cancer detection of ct lung images," Int. J. Recent Technol. Eng., vol. 7, no. 5, pp. 413–416, 2019.

[11] V. Makde, J. Bhavsar, S. Jain, and P. Sharma, "Deep neural network based classification of tumourous and non-tumorous medical images," Smart Innov. Syst. Technol., vol. 84, no. Ictis 2017, pp. 199–206, 2018, doi: 10.1007/978-3-319-63645-0_22.

[12] L. P. Armato III, Samuel G.; Hadjiiski, Lubomir; Tourassi, Georgia D.; Drukker, Karen; Giger, Maryellen L.; Li, Feng; Redmond, George; Farahani, Keyvan; Kirby, Justin S.; Clarke, "SPIE-AAPM-NCI Lung Nodule Classification Challenge Dataset," 2015, [Online]. Available: https://doi.org/10.7937/K9/TCIA.2015.UZLSU3FL.

[13] H. C. Shin et al., "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning," IEEE Trans. Med. Imaging, vol. 35, no. 5, pp. 1285–1298, 2016, doi: 10.1109/TMI.2016.2528162.

[14] X. Cui, W. Zhang, Z. Tüske, and M. Picheny, "Evolutionary stochastic gradient descent for optimization of deep neural networks," Adv. Neural Inf. Process. Syst., vol. 2018-Decem, no. Nips, pp. 6048–6058, 2018.

[15] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," J. Big Data, vol. 6, no. 1, 2019, doi: 10.1186/s40537-019-0197-0.

[16] L. Alzubaidi, O. Al-Shamma, M. A. Fadhel, L. Farhan, J. Zhang, and Y. Duan, "Optimizing the performance of breast cancer classification by employing the same domain transfer learning from hybrid deep convolutional neural network model," Electron., vol. 9, no. 3, 2020, doi: 10.3390/electronics9030445.

[17] L. Perez and J. Wang, "The Effectiveness of Data Augmentation in Image Classification using Deep Learning," 2017, [Online]. Available: http://arxiv.org/abs/1712.04621.

[18] H. M and S. M.N, "A Review on Evaluation Metrics for Data Classification Evaluations," Int. J. Data Min. Knowl. Manag. Process, vol. 5, no. 2, pp. 01–11, 2015, doi: 10.5121/ijdkp.2015.5201.

[19] H. Jiang, H. Ma, W. Qian, M. Gao, and Y. Li, "An Automatic Detection System of Lung Nodule Based on Multigroup Patch-Based Deep Learning Network," IEEE J. Biomed. Heal. Informatics, vol. 22, no. 4, pp. 1227–1237, 2018, doi: 10.1109/JBHI.2017.2725903.

[20] R. V. M. Da Nóbrega, S. A. Peixoto, S. P. P. Da Silva, and P. P. R. Filho, "Lung Nodule Classification via Deep Transfer Learning in CT Lung Images," Proc. - IEEE Symp. Comput. Med. Syst., vol. 2018-June, pp. 244–249, 2018, doi: 10.1109/CBMS.2018.00050.

[21] T. K. Sajja, R. M. Devarapalli, and H. K. Kalluri, "Lung cancer detection based on CT scan images by using deep transfer learning," Trait. du Signal, vol. 36, no. 4, pp. 339–344, 2019, doi: 10.18280/ts.360406.