Pneumonia detection using deep learning

by Sumit Prajapati

Analysis and Prediction of Pneumonia at the Earliest Stage Using Deep Learning Techniques

Sumit Prajapati
Computer Science and Engineering
Lovely Professional University
Phagwara, India
prajapatisumit7801@gmail.com

Pragya Sinha
Computer Science and Engineering
Lovely Professional University
Phagwara, India
pragyasinharps123@gmail.com

Gaurav Kumar
Computer Science and Engineering
Lovely Professional University
Phagwara, India
kumar9122gaurav@gmail.com

Dr. Soni Singh
Computer Science and Engineering
Lovely Professional University
Phagwara, India

Abstract— **The health and well-being of those who contract pneumonia are seriously at risk due to the disease's rapid spread. The use of numerous diagnostic instruments and the assessment of numerous clinical aspects are necessary for an accurate biological diagnosis of pneumonia, but these processes are hampered by the dearth of expertise and resources. Based on the findings described here, a web application that classifies patients as having pneumonia or not is being developed using deep learning algorithms. It was intended that this project will result in the development of a web application prototype for neural network-based pneumonia detection. This procedure is made simpler and problems like how many layers neural network should have been eliminated by using a high-level tool like Create ML etc. to deal and identify it. We have taken various sample from different places and trained our model with great precision.**

Keywords— Chest radiography (CXR), AlexNet, VGG16, ResNet, Data Preprocessing, Convolutional Neural Network (CNN)

I. INTRODUCTION

In Iraqi healthcare, respiratory illnesses account for between 30 and 60% of hospital admissions and between 50 and 70% of consultations. According to estimates, adults with pneumonia have a hospitalisation rate of between 22 and 42%, an intensive care unit requirement of between 5 and 10%, and a lethality rate of between 5 and 50%, depending on how severe the illness is. These rates are higher in older and immunocompromised patients.

Pneumonia is difficult to diagnose because the diagnosis involves a review of chest radiography (CXR) by a highly qualified specialist, laboratory testing, vital signs, and clinical history. Typically, it appears as a region of increased opacity in the CXR. Nevertheless, it can be difficult to diagnose pneumonia because Nevertheless, because pneumonia can coexist with other pulmonary disorders such haemorrhages, lung cancer, postsurgical alterations, pulmonary edoema, atelectasis, or collapse, diagnosing pneumonia can be challenging. Diagnosis depends on comparing CXR results taken at various dates and determining how they relate to the clinical history.

Some studies concentrate on techniques based on convolutional neural networks (CNN) to determine if a patient has pneumonia or not since they learn and select functions automatically, taking use of the vast quantity of images generated by digital processing. Additional studies emphasise the use of artificial neural networks, hidden Markov models, Deep learning models, CNN and other convolutional models.

II. LITERATURE REVIEW

Zhou et al. suggested a three-phase learning strategy that makes use of a fusion framework and deep features [1]. Fan has demonstrated an image registration technique based on convolutional neural networks.

Xu et al. investigation and construction of the CNN model used a new hierarchical loss function [2]. Yates et al. [3] used a DNN model's last layer. After reconstruction, the Inception model was categorized as gazing in binary mode. By contrasting the ResNet50 and RBF classifications developed by Nobrega et al., scientists have detected malignancies in lung nodules [4]. K et al. used an ANN with self-learning techniques using X-ray image descriptors (spatial distribution of HSB) as image descriptors. Khormouji Behzadi et al. Zhang et al. [5] examined a variety of brain signal types and the deep learning concepts that go along with them in order to assess brain signals. The primary justification for the VGG16 model [6] is its propensity to accurately classify pneumonia.

Yao L. et al. use dense net and LSTM (long-term memory network) to exploit the dependencies between anomalies [7]. Earth Mover's Distance is suggested by Khatri et al. as a means of differentiating between diseased and uninfected lungs. The CNN model was used to classify pneumonia by Rahab [8], Okeke [9], and colleagues; however, Rahman et al.

Xiol designed a three-dimensional heterogeneous deep convolutional neural network (DCNN) (MSHCNN) on a higher dimensional scale. Nada M. Elshenawy [10] offered four distinct models to change pre-existing models: a convolutional neural network (CNN), a long short-term memory (LSTM), and two pre-trained models (MobileNetV2 and ResNet152v2). Over 91% of the results were correctly represented by LSTM-CNN. Medical image segmentation, which uses the UNet model to extract regions of interest from X-ray pictures, yields results for Montgomery and JavaScript Runtime (JSRT) datasets that are 97.1% and 97.7%, respectively. Its architecture is improved by utilizing the entirety of the CNN SENet architecture.

Rajpurkar et al. developed a CheXNeXt architecture [11] for the detection of several illnesses, including pneumonia. X-ray pictures with a spatial distribution of HSBs were proposed by Keet [12]. A brighter picture as a result increases approach accuracy. The element measuring the whole Pareto was generated by the generic technique NSGANetV1 [13], and multiple generic functions were then used in tandem to progressively replicate and converge each design model. Yu et al. created an automated CNN-based pipeline for the analysis of prostate cancer.

TABLE 1. Review of Various DL Models

Feature	VGG16	DenseNet	SqueezeNet
Architecture	CNN	CNN	CNN
Year of release	2014	2017	2016
Depth	16 layers(13 Convolutional, 3 Fully connected)	Variable (typically DenseNet-121 has 121 layers)	18 layers
Pattern of Connection	Traditional skip connections between consecutive layers	Dense connectivity between all the layers.	Fire modules(Squeeze and expand)
Reusable Feature	Due to traditional architecture its less reuse.	Due to dense connectivity its high reuse	Due to fire modules its moderate reuse.
Computational Effeciency	VGG16 tends to have lower computational efficiency due to its larger model size and lack of parameter sharing.[29]	DenseNet achieves high computational efficiency by promoting parameter sharing through dense connectivity.[23]	SqueezeNet achieves high efficiency by reducing the number of parameters, making it suitable for deployment on resource-constrained devices.[24]
Memory consumption	It typically consumes more memory consumption.	Moderate memory consumption.	Small memory consumption.
Accuracy	It is known for their high accuracy on various image classification tasks.[21]	It is also known for their high accuracy on various image classification tasks.[23]	SqueezeNet achieves moderate to high accuracy while maintaining a smaller model size.[17]
Consumption cost	VGG16 has a higher computational cost due to its higher computational complexity and more parameters.[18]	DenseNet's computational cost is moderate to high depending on the depth of the model.[29]	SqueezeNet has lower to moderate computational cost due to its smaller model size and efficient architecture.[22].

III. MATERIALS AND METHODS

Combining data and characteristics from several sources has been shown in recent deep learning research to improve model performance, especially for picture classification. In light of this, we created an attention-based deep convolutional neural network (CNN). This method improves the model's capacity to identify and categorize pneumonia from chest X-ray pictures by combining the strengths of several deep learning architectures, including CNN, Alexnet, DenseNet and VGG16. We might determine whether our approach is useful for interpreting chest X-ray pictures and identifying conditions like pneumonia by doing these tests. This enables us to assess how well our concept is functioning and whether patients and physicians may benefit from it. As. More precise and reliable predictions are made possible by the network's ability to dynamically focus attention on pertinent portions of the input data thanks to attention combining methods.
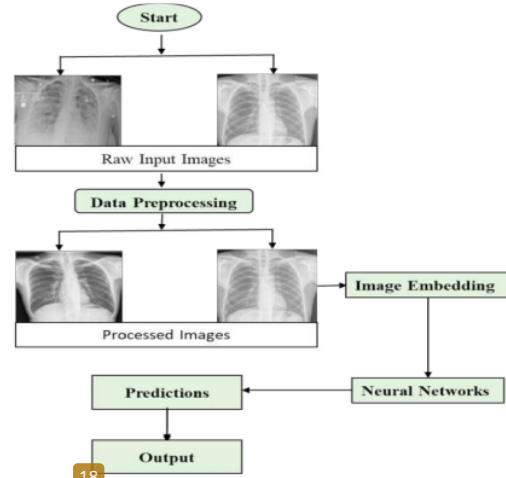

Fig. 1. Flow diagram of the proposed framework

A. Dataset:

The dataset has three main folders: training, testing, and validation, with a total of 5836 images. These folders are divided into pneumonia and normal subfolders. The images show the chests of children aged 1 to 6 years from both the front and back. The goal of the experiment was to improve validation accuracy and reduce validation loss in classifying pneumonia images from this chest X-ray dataset.

B. Data Preprocessing:

Data preprocessing is like getting your ingredients ready before cooking. When we collect data from different sources, it's usually in a raw form, just like collecting groceries from the store. But raw data isn't ready for analysis yet. However, raw data is not yet suitable for analysis. Preprocessing entails cleaning, dicing, and arranging the components in preparation for cooking. Preprocessing is the process of changing the raw data in a number of ways to prepare it for analysis. When working with photos, for instance, we might resize them to the same size or change their color. This facilitates consistency and computer comprehension of the data. By making minor adjustments to the current data, such as flipping or rotating

photographs, we may also be able to expand the number of samples in our dataset. This makes our dataset bigger and more diverse, which is helpful for training our deep learning algorithm.

C. Data Augmentation:

Data augmentation techniques were applied to the photos in order to accomplish this. The original dataset was modified, allocating 2208 photos to the validation set and 3628 images to the training set, in order to preserve data balance. This modification significantly increased validation accuracy. Overall, preprocessing is all about getting our data ready for analysis by cleaning it up, making it consistent, and adding more examples to work with. We carefully looked at and prepared all the chest X-rays to make sure they were clear and easy to read. We removed any scans that were blurry or hard to understand. This helps us make sure our analysis of the chest X-rays is accurate. Several data augmentation techniques were used to artificially improve the quantity and quality of radiographs, which helped to raise the model's productivity during training [15]. The value by which images will be increased or lowered for the enhancement process is represented by the rescale operation. The 40-degree rotation range represents the images' random rotation during the training phase. For random picture translation in both horizontal and vertical orientations, variations in width and height are restricted to 0.2 percent.

We made slight angle adjustments to the images, tilting them counterclockwise by a small amount (0.2 percent). We also zoomed in randomly on the images by a tiny fraction (0.2 percent) and flipped them horizontally. These actions improved the accuracy of the system because the input images could be taken from various angles and have different sizes. Some images might be far away, while others are close, so these preprocessing techniques help to handle these variations.

TABLE II. SETTINGS USED FOR IMAGE AUGMENTATION

Method	Setting
Rescale	1/255
Rotation range	40
Width shift	0.2
Height shift	0.2
Shear range	0.2
Zoom range	0.2
Horizontal flip	True
Fill mode	Nearest

D. Deep Learning Methods:

Deep learning algorithms are now used to identify pneumonia in the majority of cases since 2016. Deep learning techniques like CNN, AlexNet, ResNet and VGG16 are the most researched. Only a select few tactics are selected based on precision and excellent outcomes.

a. *AlexNet:* Alex Net classifies more than 1 thousand good training samples using deep layers with 650,000 neurons and 60 million boundaries. The convolution layers are made up of the output of the soft max layer, three pooling layers, five convolutional layers, two fully connected

layers (FCL), and five convolutional layers [14]. To transfer the information image to the second layer, the main convolution layer of Alex Net uses 96 kernels, each with a size of 11×11×3 and a stride of 4 pixels. Alex Net requires an input image of dimensions 227 × 227 x 3.

b. *ResNet:* ResNet is particularly renowned for its capacity to train remarkably deep networks efficiently. Its design diverges from conventional convolutional neural networks (CNNs) by incorporating residual connections, which enable the direct flow of information through the network. This innovation tackles the challenge of vanishing gradients encountered in deep networks by facilitating the learning of residual functions with respect to the input. With ResNet, each layer has the ability to access both the loss function and gradient values from the real image data, enhancing its learning capability. This unique architecture allows ResNet to effectively learn intricate features and hierarchical representations, ultimately leading to superior performance in various image recognition tasks, including medical imaging applications such as pneumonia detection from chest X-ray images.
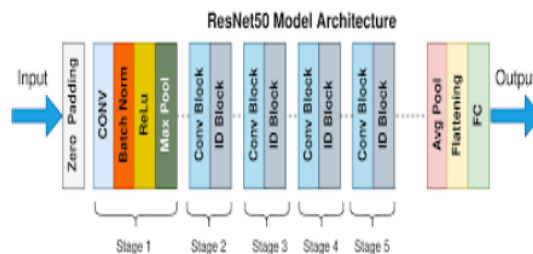

Fig. 2 Resnet Architecture

c. *VGG16*: The CNN design created by Simonyan and Zisserman of the VGG16 Visual Geometry Group took first place in the 2014 ImageNet competition. The simplicity of its design is what sets it apart. It employs a 2x2 kernel for the max pooling layer and a 3x3 kernel for the convolution layer in place of convolutional layers. VGG16 is frequently used for transfer learning in pneumonia classification because, when adjusted for various tasks, it produces good results. Analogously, other models, such ResNet152V2 and Inception-Resnet-v2, exhibit increasing accuracy with time.
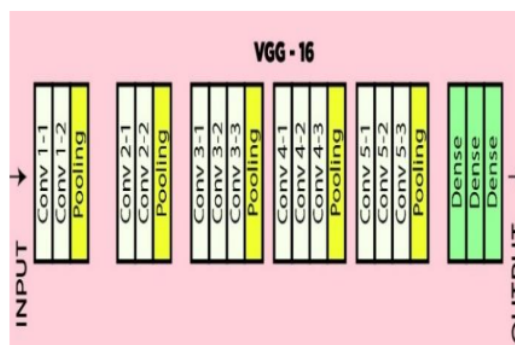
Fig.3 Architecture of VGG 16

d. *CNN Model :* Convolutional Neural Networks (CNN) is one kind of deep learning model that is made especially for problems requiring images and other structured grid data, such as audio and time series. CNNs have demonstrated remarkable success in various computer vision tasks, including image classification, object detection, segmentation, and image generation.
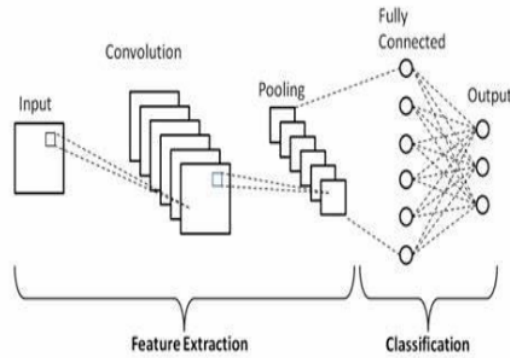

Fig. 4 Convolutional neural network (CNN)

Different components and functioning of CNN are discussed in detail below-'

i)Convolutional Layer :

This layer is the core blocks of CNNs. In order to extract different features from input images, they apply learnable filters, also known as kernels. From the input data, each filter extracts particular features or patterns. To create a feature map, the convolution operation involves swiping a filter over the input image, calculating element-wise multiplication, and then summarizing the results. Multiple filters are applied simultaneously, to create multiple feature maps that capture different features of the input image.

ii)Pooling Layers :

Convolutional layers are separated by pooling layers, which lower the spatial dimensions of feature maps without sacrificing the most important information. A common pooling process is called "max pooling," in which the maximum value in a small neighbourhood (such as a 2x2 window) is kept and the remaining values are discarded.

iii)Fully connected Layers :

There are one or more completely connected layers added towards the conclusion of the CNN architecture. Like a regular neural network, these layers link every neuron in one layer to every other layer's neuron. Convolutional and pooling layers teach high-level features, which are then aggregated and mapped to the desired output classes by fully linked layers. To acquire class probabilities in classification problems, the output layer usually consists of neurons equal to the number of classes, followed by a softmax activation function.

iv)Flattening :

The feature maps are flattened into a one-dimensional vector before being passed to fully linked layers as the outputs of the convolution and pooling layers. The output is completely consistent with related layers thanks to this flattening procedure.

v)Training:

CNNs use gradient descent optimisation and back-propagation to automatically extract pertinent features from the input data during training. To reduce the discrepancy between the expected outputs and the actual labels in the training data, the network modifies the weights and biases of each layer.

CNNs are used in many different fields, including item identification, facial recognition, medical picture analysis, autonomous vehicles, and image classification, and natural language processing jobs utilising structured grid data.

In summary, CNNs are strong deep learning models that are especially well-suited for tasks involving images and structured grid data. They are made to automatically learn hierarchical representations from input data. Their capacity to capture invariant characteristics and spatial hierarchies makes them extremely useful in resolving a variety of computer vision issues.

vi)Experimental Result :

Binary Prediction: This initializes an empty list 'binary prediction' that store the binary prediction "0 or 1" based on a range. Threshold Selection: The precision score is used to choose the range. From the list of thresholds where the accuracy is more than or equal to 0.80, the highest threshold is chosen. By using this method, the model's accuracy is guaranteed to be at least 0.80, which means that 80% of the positive predictions will come true. Binary classification: Every prediction probability is added to binary prediction and classed as positive (1) if it exceeds or equals the threshold. It is categorised as negative (0) if the expected probability is less than the cut-off. Set thresholds for our model, we want the results to be precise while not sacrificing too much recall.

Below describe the accuracy , precision and recall of CNN model.

TABLE III. Result of CNN model

Testing Set	Score
Accuracy	91%
Precision	91%
Recall	79%
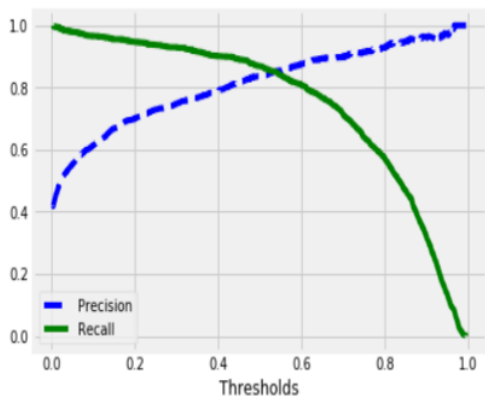
Fig.5 Precision vs Recall

is used for the classification of image, they are also different in architecture, computational efficiency, memory consumption, model size, and accuracy. The choice between these models depends on the specific requirements of the application, such as computational resources, memory constraints, and desired accuracy. The comparison table are discussed below-

TABLE IV. COMPARISION OF MODELS WITH CNN MODEL

Method	Alexnet	Resnet	VGG16		
Accuracy	89%	76%	87%		
Precision	87%	73%	85%		
Recall	71%	70%	72.03%		

Accuracy: The percentage of accurate predictions among all predictions produced – both true positives and true negatives – is measured by CNN model on the test set. The accuracy is 91% .

Precision: The percentage of accurate predictions among all predictions produced – both true positives and true negatives – is measured by CNN model on the test set. The precision is 91% .

Recall: Recall is the percentage of accurate predictions made for all true positive cases in the test set. With a recall of 79% . The test set's true positive cases are completely captured by the model.
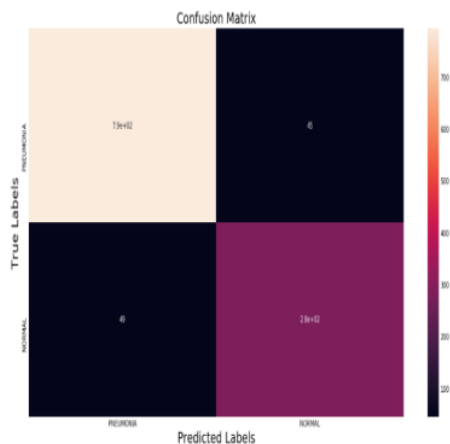

Fig.6 Confusion Matrix

IV. COMPARISON OF VGG16 , ALEXNET AND RESNET

In this section the comparison of CNN model with another DL model such as VGG16, ResNet and AlexNet are discussed. These are all deep convolutional neural networks models that

V. CONCLUSION

The aim of this study was to develop a web application prototype that uses neural networks to identify pneumonia from a chest X-ray image. A high-level programme called develop ML was used to develop the model. It streamlines the process and removes obstacles like deciding which algorithms to employ, how many layers to put in a neural network, and how to initialise the model's parameters. Anyone can use the model because it has been implemented as a web application. A more comprehensive and diverse data set can be used in future research to categorise various lung illnesses. Other picture modalities, including MRIs or mammograms, can also be analysed to determine whether a person is having pneumonia disease or not.

VI. REFERENCES

[1] Zhu, X.; Shen, D. Zhou, T.; Thung, K.;Effective feature learning and fusion of multimodality data using stage-wise deep neural network for dementia diagnosis. Hum. Brain Mapp. 2018, 40, 1001–1016

[2] Harvey, H. Yates, E.J.Yates, L.C.;Machine learning "red dot": Open-source, cloud, deep convolutional neural networks in chest radiograph binary normality classification. Clin. Radiol. 2018, 73, 827–831

[3] Yap, P.T.; Shen, Fan, J.; Cao, X.; D. BIRNet: Dual-supervised fully convolutional networks for brain image registration. 2019 Med. Image Anal

[4] Zhang, Wo ´zniak, M.; Ko´smider, L.; Damasevi´´cius, R.; Wei, W.; Połap, D.; Ke, Q. a neuro-heuristic method for identifying lung conditions from X-ray pictures. 2019, 126, 218–232; 54, 193-206. Expert Syst. Appl.

[5] Tawsifur Rahman, Muhammad E. H. Chowdhury2,Amith Khandakar2, Saad Kashem5, Muhammad A. Kadir1, Zaid B. Mahbub4, Khandaker R. Islam3, Khandaker F. Islam2, and Amith Khandakar2 Utilizing Transfer Learning and Deep Convolutional Neural Network (CNN) for Chest X-ray-Based Pneumonia Detection

[6] H.; Rostami, H.; Assadi, M.; Batouli, A.; Masoumi, M.; Salemi, S.; Keshavarz, A.; Gholamrezanezhad, A.; Salehi, Behzadi-khormouji, S.; Derakhshande- Rishehri, T. Consolidation detection on chest X-ray images using deep learning, reusable, and problem-based architectures. 2020, 185, 105162; Comput. Methods Programs Biomed..

[7]   S.M. Jadhav, S.S. Yadav,"Deep convolutional neural network- based medical image classification for disease diagnosis," Journal of Big Data, 6(1), 113, 2019, doi:10.1186/s40537-019- 0276-2.

[8]   Satyarth Katiyar 2,†, Avinash G Keskar 3,†, Neeraj Dhanraj Bokde 4,† and Zong Woo Geem 5, Mohammad Farukh HashmiEfficient Pneumonia Detection in Chest Xray Images Using Deep Transfer Learning.

[9]   E.; Dagunts, DYao, L.; Poblenz.; Covington, B.; Bernard, D.; Lyman, K. Learning to diagnose from scratch by exploiting dependencies among labels. arXiv 2017, arXiv:1710.10501.

[10]   C. Roberto,B. Vijay, K. Alex, "Segnet: Deep convolutional encoder-coder architecture for image segmentation," 2015, http://arxiv.org/abs/1511.00561.

[11] . Felix Bragman Sebastien Zach Eaton-Rosen,Ourselin M. Jorge Car do so proving Data Augmentation for Medical Image Segmentation.

[12] . K. Rashed, R. Kawal, B.M. Aliasghar, Jeremy, and B. Ulas, Cardiacnet, Segmentation of Left Atrium and Proximal Pulmonary Veins from MRI using Multi-View CNN, MICCAI, Springer, New York, NY, USA, 2017.

[13] . In Proceedings of the International Conference of AI and Information Technology (ICAIIT), Ouargla, Algeria, 4–6 March 2019; pp. 486–489.

[14] . Wang , X.; Monaghan,Zhang , X.; Yao, L.; J.; McAlpine , D. A Survey on Deep Learning based Brain-Computer Interface: Recent Advances and New Frontiers. arXiv 2019 ,arXiv:1905.04149.

[15]   Stirenko S, Kochura Y, Alienin O, Rokovyi O, Gordienko Y, Gang P, Zeng W. Chest X-ray analysis of tuberculosis by deep learning with segmentation and augmentation. In 2018 IEEE

Pneumonia detection using deep learning

1   www.researchgate.net
Internet Source   **3**%

2   A. Beena Godbin, S. Graceline Jasmine. "Analysis of pneumonia detection systems using deep learning-based approach", 2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), 2022
Publication   **2**%

3   Submitted to University of Northumbria at Newcastle
Student Paper   **2**%

4   Devvret Verma, Chitransh Bose, Neema Tufchi, Kumud Pant, Vikas Tripathi, Ashish Thapliyal. "An efficient framework for identification of Tuberculosis and Pneumonia in chest X-ray images using Neural Network", Procedia Computer Science, 2020
Publication   **2**%

5   www.ncbi.nlm.nih.gov
Internet Source   **2**%

| 6 | Submitted to Lovely Professional University
Student Paper | 1% |

| 7 | export.arxiv.org
Internet Source | 1% |

| 8 | Submitted to National School of Business Management NSBM, Sri Lanka
Student Paper | 1% |

| 9 | Submitted to Liverpool John Moores University
Student Paper | <1% |

| 10 | Md.Momenul Haque, Subrata Kumer Paul, Rakhi Rani Paul, Nurnama Islam, Mirza A. F. M. Rashidul Hasan, Md.Ekramul Hamid. "Chapter 10 Improving Performance of a Brain Tumor Detection on MRI Images Using DCGAN-Based Data Augmentation and Vision Transformer (ViT) Approach", Springer Science and Business Media LLC, 2023
Publication | <1% |

| 11 | Sakrapee Paisitkriangkrai, Jamie Sherrah, Pranam Janney, Anton van den Hengel. "Semantic Labeling of Aerial and Satellite Imagery", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2016
Publication | <1% |

| 12 | gala.gre.ac.uk | |

Internet Source <1%

13 Mohammad Abdolahi, Mohammad Salehi, Iman Shokatian, Reza Reiazi. "Artificial intelligence in automatic classification of invasive ductal carcinoma breast cancer in digital pathology images", Medical Journal of The Islamic Republic of Iran, 2020
Publication <1%

14 www.ijitee.org
Internet Source <1%

15 www.mdpi.com
Internet Source <1%

16 Swati Hira, Anita Bai, Sanchit Hira. "An automatic approach based on CNN architecture to detect Covid-19 disease from chest X-ray images", Applied Intelligence, 2020
Publication <1%

17 link.springer.com
Internet Source <1%

18 "Data Intelligence and Cognitive Informatics", Springer Science and Business Media LLC, 2021
Publication <1%

19   "Proceedings of 3rd International Conference on Computer Vision and Image Processing", Springer Science and Business Media LLC, 2020

Publication

<1 %

Exclude quotes   On   Exclude matches   Off
Exclude bibliography   On

Pneumonia detection using deep learning