

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

Sumit Prajapati
Computer Science and Engineering
Lovely Professional University
Phagwara, India
prajapatsumit7801@gmail.com

Pragya Sinhaline
Computer Science and Engineering
Lovely Professional University
Phagwara, India
pragyasinharp123@gmail.com

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

line 1: 4th Given Name Surname
Computer Science and Engineering
Lovely Professional University
Phagwara, India
line

Abstract— The health and well-being of those who contract pneumonia are seriously at risk due to the disease's rapid spread. A correct biological diagnosis of pneumonia requires the application of multiple diagnostic tools and the evaluation of multiple clinical factors, 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 process is streamlined, and issues such as the optimal number of layers for a neural network are resolved by employing advanced tools such as Create ML, among others, to address and recognize them. We have taken various sample from different places and trained our model with great precision.

Keywords— Chest radiography (CXR), DenseNet, Data Preprocessing, Convolutional Neural Networks (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.

Diagnosing pneumonia is challenging because it requires a highly qualified professional to study chest radiography (CXR), as well as laboratory tests, vital signs, and clinical history. It usually shows up in the CXR as an area of increased opacity. However, pneumonia can be challenging to detect because Diagnosing pneumonia can be difficult, though, because it can occur with other pulmonary conditions such as hemorrhages, lung cancer, postsurgical changes, pulmonary edema, atelectasis, or collapse. Comparing CXR results obtained at different times and figuring out how they connect to the clinical history are necessary for the diagnosis.

In order to ascertain if a patient has pneumonia or not, some research focuses on methods utilizing convolutional neural networks (CNNs), which learn and choose functions automatically by utilizing the enormous volume of images produced 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.'s investigation and construction of the CNN model used an innovative 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 comparing the ResNet50 and RBF classifications 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 main argument in favor of the VGG16 model [6] is that it tends to classify pneumonia correctly.

Yao L. et al. use dense net and using a long-term memory network (LSTM) to take advantage of the relationships between abnormalities [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, 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, to extract regions from X-ray pictures, yields results for Montgomery and Javacript Runtime (JSRT) datasets that are 97.1% and 97.7%, respectively. Its architecture is improved by utilizing the entirety of the CNN SENet architecture.

A CheXNeXt architecture was created by Rajpurkar et al. [11] to identify a number of diseases, including pneumonia. Keet [12] proposed X-ray images with a spatial distribution of HSBs. Approach accuracy is therefore increased by a brighter picture. The generic method NSGANetV1 [13] produced the element measuring the entire Pareto, and several generic functions were subsequently applied simultaneously to

gradually reproduce and converge each design model. An automated CNN-based pipeline for prostate cancer analysis was developed by Yu et al.

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 Efficiency	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 has a reputation for having excellent accuracy on a range of picture classification tasks.[21]	It is renowned for having excellent accuracy in a range of picture categorization 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].

Studies that used CNN chest data set by Stephen et al. Trained to diagnose pneumonia were more successful than those that used transfer learning. To increase the amount of information gathered, he employed information expansion strategies such flat flip, zoom, and modification in width and height.

This CNN model has four layers where each layer processes information using a 3x3 grid. After each layer, there's a max pooling step to simplify the information. The

activation function used is "relu", and the final two layers are fully connected to make sense of all the data. Sometimes, the model may not fully understand the image after applying these layers, but overall, the model's accuracy has been tested with different types of images, including color images with a size of 200x200 pixels, and it performed really well.

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.

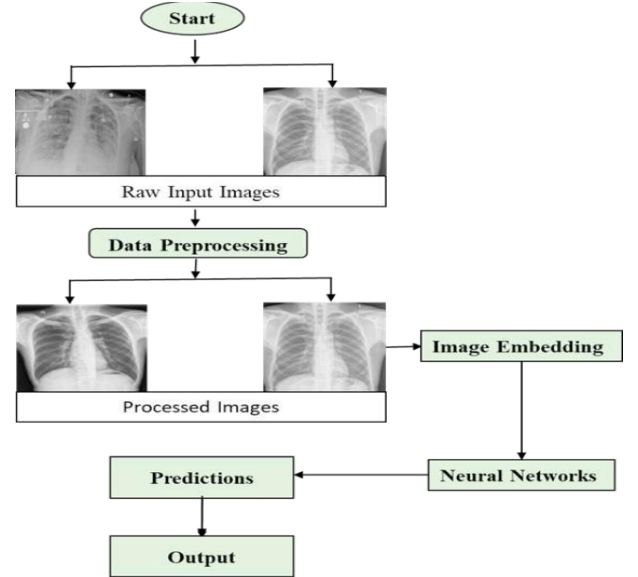


Fig. 1. Flow diagram of the proposed framework

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 project aimed to decrease validation loss and increase validation accuracy in the classification of 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.

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.

C. Data Augmentation

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. The number and quality of radiographs were artificially increased using a number of data augmentation approaches, which increased the model's output during training [15]. The rescale operation represents the value by which images will be decreased or raised during the enhancing process. 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). These actions improved the accuracy 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 I. 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.

- I. *Alex Net*: Alex Net uses deep layers with 650,000 neurons and 60 million borders to classify over a thousand high-quality training examples. The output of the soft max layer, three pooling layers, two fully connected layers (FCL), five convolutional layers, and three convolutional

layers comprise the convolution 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 $11 \times 11 \times 3$ and a stride of 4 pixels. Alex Net requires an input image of dimensions $227 \times 227 \times 3$.

- II. *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 X-ray images of chest.

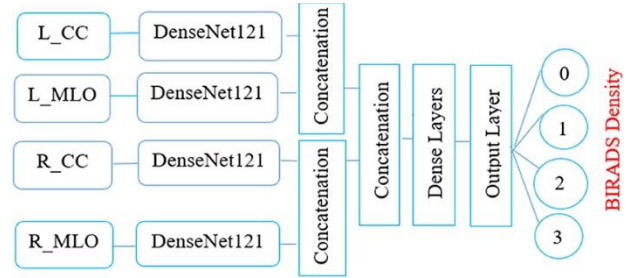


Fig. 2. Denesenet Architecture

- III. *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 2×2 kernel for the max pooling layer and a 3×3 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 as ResNet152V2 and Inception-Resnet-v2, exhibit increasing accuracy with time.

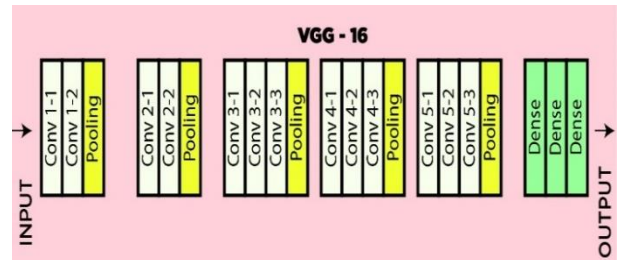


Fig. 3. Architecture of VGG 16

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.[25]. Across a range of computer vision applications, such as segmentation, object identification, image synthesis, and image classification, CNNs have shown impressive performance[26].

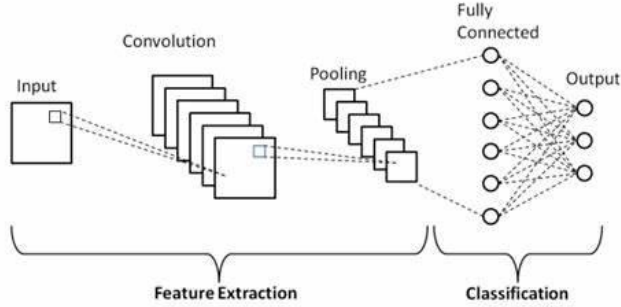


Fig. Convolutional neural networks (CNN)

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

A. 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.[27] 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.[25] Multiple filters are applied simultaneously, to create multiple feature maps that capture different features of the input image.[27].

B. 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.[28]

C. 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.[26]. Convolutional and pooling layers teach high-level features, which are then aggregated and mapped to the desired output classes by fully linked layers.[29]. 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.[30]

D. Flattening :

As the outputs of the convolution and pooling layers, the feature maps are first flattened into a one-dimensional vector and then sent to fully linked layers. The output is completely

consistent with related layers thanks to this flattening procedure.[27]

E. Training:

CNNs use gradient descent optimisation and back-propagation to automatically extract pertinent features from the input data during training.[17] The network adjusts the weights and biases of each layer in order to lessen the difference between the expected outputs and the actual labels in the training data.

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.[25]

In summary, CNNs are strong deep learning models that are especially well-suited for tasks involving images and structured grid data.[25] 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.

IV. EXPERIMENTAL RESULT

Binary Prediction: This initializes an empty list 'binary prediction' that store the binary prediction "0 or 1" based on a range.[20]. 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.[21]. 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.[22]. 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.[17]

Table 3. and Fig. 4. Below describe the result of precision and recall of the prediction model.

TABLE 3. Result of CNN model

Testing Set	Score
Accuracy	91%
Precision	91%
Recall	79%

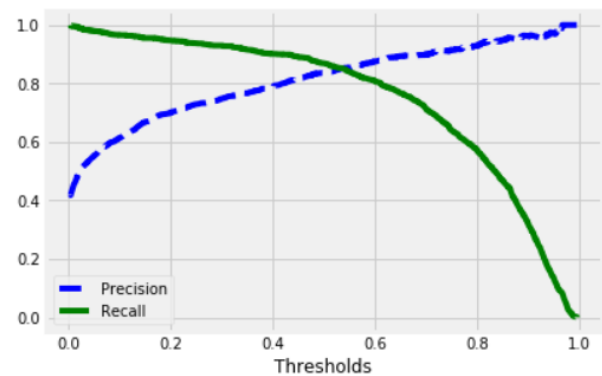


Fig. 4. Precision vs Recall

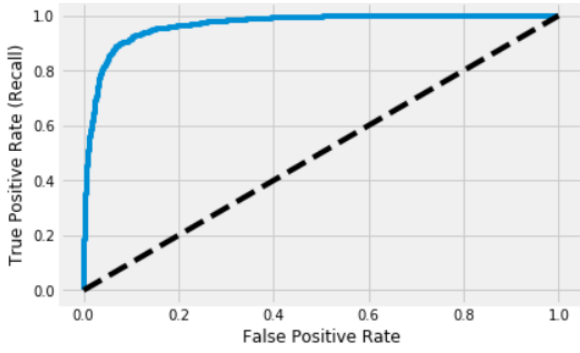
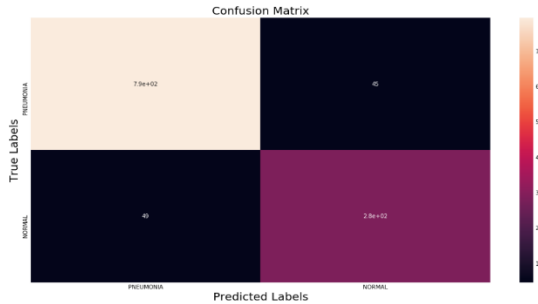


Fig. 5. FRP(False Positive rate) vs TRP (True Positive Rate)

Accuracy: The percentage of accurate predictions among all predictions produced – true negatives as well as true positives – is measured by accuracy on the test set. In this example, the accuracy is 91.04%, that is, 91.04% of the model's predictions are realised.[21]

Precision: Accuracy on the test set indicates what proportion of the model's positive predictions are actually true. With an accuracy of 91.49% in this instance, the model is 91.49% accurate when predicting a positive case.[16] This is consistent with the cutoff point established to obtain a minimum of 80% accuracy.

Recall: Recall is the percentage of accurate predictions made for all true positive cases in the test set. With a recall of 79.63%, the test set's true positive cases are completely captured by the model.[22] This indicates that the model can detect about 79.63% of instances of pneumonia without significantly missing any.



V. COMPARISON RESULT

In this section the comparison of CNN model with another DL model such as VGG16, DenseNet and SqueezeNet are discussed. These are all deep convolutional neural networks models that is used for the classification of image, they are also different in architecture, computational efficiency, memory consumption, model size, and accuracy.[11] The exact needs of the application, such as memory limits, processing power, and required precision, determine which of these models to use. The comparison table are discussed below-

TABLE II. COMPARISON RESULT OF DEEP LEARNING MODELS WITH CNN MODEL

Method	Alexnet	Resnet	VGG16	CNN
Accuracy	89%	76	87	91
Precision	87%	73	85	91
Recall	71%	70	72.03%	79
Accuracy	0.2			

CONCLUSION

This study's objective was to create 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.

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.; Polap, 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] Tawisfur 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- Rishchri, 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 Hashmi Efficient Pneumonia Detection in Chest X-ray Images Using Deep Transfer Learning.
- [9] E.; Dagunts, DYao, L.; Poblentz.; 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-decoder 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 (ICAIT), 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, Rokovy O, Gordienko Y, Gang P, Zeng W. Chest X-ray analysis of tuberculosis by deep learning with segmentation and augmentation. In 2018 IEEE
- [16] 38th International Conference on Electronics and Nanotechnology (ELNANO) 2018 Apr 24 (pp. 422-428). IEEE. G. Eason, B. Noble,

- and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, April 1955. (*references*)
- [17] [16] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9).
- [18] [17] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533-536.
- [19] [18] Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks?. In *Advances in neural information processing systems* (pp. 3320-3328).
- [20] [19] Shin, H. C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., ... & Summers, R. M. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE transactions on medical imaging*, 35(5), 1285-1298.
- [21] [20] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.
- [22] [21] Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427-437.
- [23] [22] Powers, D. M. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies*, 2(1), 37-63.
- [24] [23] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
- [25] [24] Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., & Keutzer, K. (2016). SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size. *arXiv preprint arXiv:1602.07360*.
- [26] [25] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- [27] [26] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems* (pp. 1097-1105).
- [28] [27] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning* (Vol. 1). MIT press Cambridge.
- [29] [28] Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., & LeCun, Y. (2013). OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks. *arXiv preprint arXiv:1312.6229*.
- [29] [29] Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv preprint arXiv:1409.1556*