DETECTING PNEUMONIA IN A CHEST X-RAY USING DEEP LEARNING

Khaled A. Alsheshtawi, PhD[1]
Head of Artificial Intelligence Dept.
College of Information Technology,
Jan Youssef[2], Bishoy Samir[3], Mohamed Osama[4], Mohamed Tarek[5]

[1]Misr University for Science and Technology
Email: khaled.alsheshtawi@must.edu.eg

[2]Misr University for Science and Technology
Email: 89477@must.edu.eg

[3]Misr University for Science and Technology
Email: 89534@must.edu.eg

[4]Misr University for Science and Technology
Email: 89402@must.edu.eg

[5]Misr University for Science and Technology
Email: 89610@must.edu.eg

ABSTRACT

Our motivation for this project is to address the global shortage of trained radiologists [1], specifically in Egypt, which faces a steep increase in healthcare prices and a decline in availability especially for low-income Egyptians. Our objective is to create an AI model which is capable of classifying Chest X-Rays into those positive for Pneumonia and those negative, with a higher focus on sensitivity or the *"true-positive rate"*, as recommended by the WHO [2], these guidelines were chosen because the highest chances of saving lives through detecting cases which are very likely to be positive. Our methodology consisted of dividing the data into 80% for training, 10% for testing, and 10% for validation, following that we tuned our model to give equal weights to each of the RGB channels in the image, as we found one channel tended to dominate the others. During the training, the model was instructed to learn much more heavily from the negative cases, as our data was comprised of 90% positive cases and 10% negative cases which was biasing the model towards the positive case. We found that our model achieved ~88% accuracy, compared with [3] which achieved 76.4% on the same dataset. This was accomplished with minimal training on the free tier provided by the Kaggle platform, further fine-tuning may be possible with better hardware.

KEYWORDS: Artificial Intelligence; Pneumonia; X-rays; Detection; Screening

1. INTRODUCTION

Is it possible to replace trained radiologists with trained AI models? Multiple previous studies have attempted to use AI for the detection of pneumonia in chest x-rays, [3] found that 50,000 people die from pneumonia annually. They developed a novel mechanism using feature extraction and dimensionality reduction, outperforming Mobile Net, Exception Net, and Resnet. Multiple deep learning architectures were used, with the proposed mechanism achieving the highest accuracy. Well-known architectures are unsuitable for pneumonia detection in low computational power environments. However, they were only able to reach an accuracy of 76.4%. While previous research attempted to tackle the problem of radiologist shortage using AI, their results were not strong enough to motivate decision makers, in this paper we believe our approach can be used in environments where computing resources are not abundant, thus making it perfect for developing countries and groups with access to incomplete data for training. Our paper argues that AI models must be used extensively by Egypt and other countries suffering from a shortage of medical professionals, to alleviate the pressure on low-income groups suffering from healthcare inaccessibility. It first establishes the functional and non-functional requirements for such a project, then goes on to discuss the scope and algorithms used, and finally the results and potential objections.

2. METHODOLOGY

2.1 Data loading and preprocessing: The dataset is loaded using a class from the torchvision library. Data transformations such as resizing images to a specific size and converting them to tensors are applied using the transforms module.

2.2 The directory structure of the dataset is explored, which returns the class labels.

2.3 Data loaders for the training and testing datasets are created using the DataLoader class. The dataset can be iterated over in batches, with the data shuffled and loaded in parallel. The datasets.ImageFolder class and the previously defined transformations are used to create the training and testing data loaders.

2.4 The mean and standard deviation of the training dataset are calculated. These statistics are used to normalize the image data.

2.5 We used Transfer Learning technique with ResNet18 model as the base model, and a new fully connected layer is substituted to adapt it to the binary classification task. The model parameters are optimized using the Adam optimizer, and a learning rate scheduler is used to adjust the learning rate during training. The binary cross-entropy loss function is defined using nn.BCEWithLogitsLoss.

2.6 A weight tensor is created based on based on quantity of each class in the training dataset. This tensor is utilized to assign higher weights to minority class samples during training.

2.7 The model is trained using a loop that iterates over the training dataset. In each iteration, the model parameters are updated using backpropagation and the calculated gradients. The model weights are adjusted based on the computed loss value. The learning rate scheduler is also updated according to the validation loss.

2.8 The test dataset is evaluated by iterating over the test data loader. Predictions are generated using the trained model, and the performance metrics are calculated.

3. RESULT
3.1 Training vs Testing Accuracy
Our model achieved a commendable accuracy rate of 88%, showcasing its robust performance and proficiency in the task at hand.
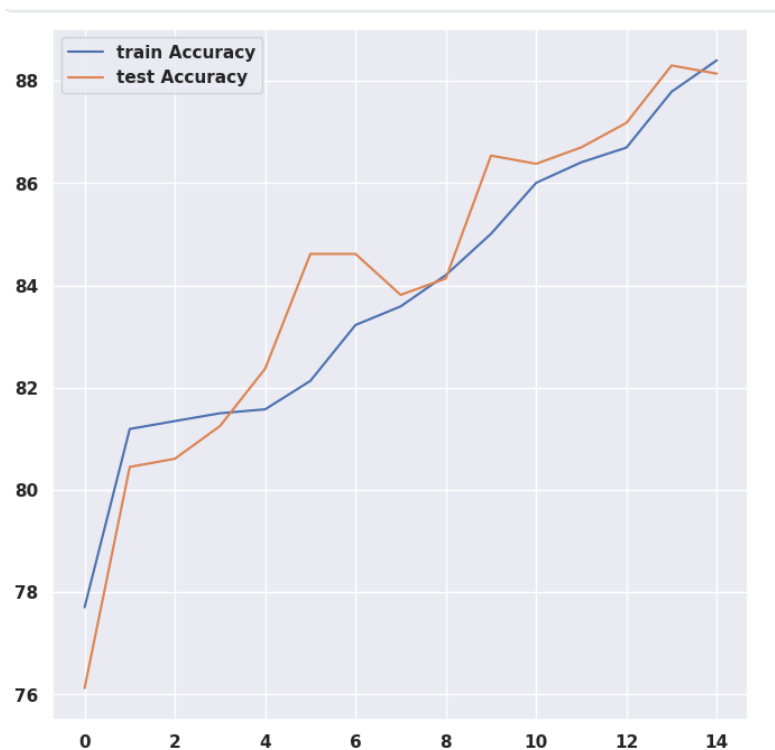

Figure 1. Train vs Test Accuracy.
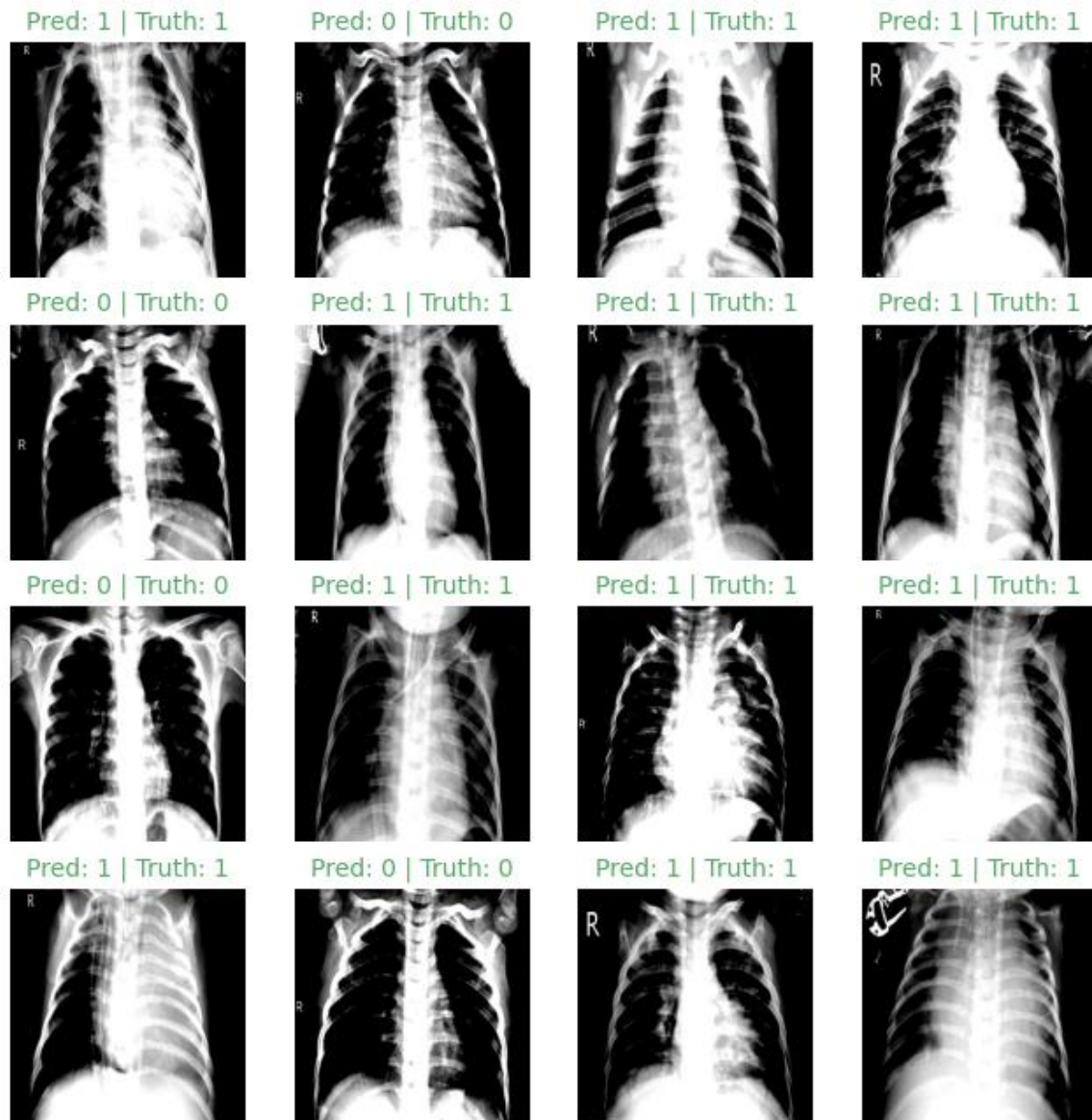
3.2 Predicted Vs Actual Value for A Sample of X-rays

Figure 2. a 4x4 plot of predicted and actual value for a sample of CXRs.

3.3 The Response From Our API When a CXR is Uploaded

Figure 3. The response from our API

3.4 A Comparison of Algorithms Used and Accuracy

Table 1. A Comparison of Algorithms Used and Accuracy

Study	Published on	Dataset	Algorithms used	Accuracy
Hayelny	2022-23	5,863 Chest X-ray images	Channel-wise normalization Weighted loss function Weight decay Learning scheduler Data Augmentation CNN Architecture: Resnet18	88%
[4]	December 2020	Montgomery dataset (138 images) Shenzhen dataset (662 images)	Canny Algorithm [13]	93.59%
[5]	July 2021	3000 images for normal chest X-rays were selected from different public image databases. 623 chest X-ray COVID-19 images were collected from the GitHub repository, expanded to 2000 by image augmentation	VGG16 [14] ResNet50 [15] InceptionV3[16] Transfer learning [17]	>98%
[6]	May 2021	550,297 images	Mask RCNN28 with a ResNet-101-FPN. Single Shot MultiBox Detector EfficientNet-B7 with an attention pooling layer and a fully-connected layer.	Sensitivity: 88% Specificity: 79%
[3]	2021	5,863 Chest X-ray images	SIFT Feature Scale-space extreme value detection Construct scale space Keypoint positioning and feature description. Vector normalization to generate descriptors. Feature clustering algorithm: - K-means ++	76.4%

			- Bag of Visual Words	
			- Support Vector Machine	
			- K-Nearest Neighbor	
			- Naive Bayes	
			- Random Forest	
			Deep Neural Networks:	
			- VGGNet	
			- GoogleNet	
			- ResNet	
			- DenseNet	
			- Inception	
			- MobileNet	
			- Xception	
[7]	Jan 2021	6749 two-view chest radiographs	Resnet18 pre-trained with ImageNet CNN	74%

4. DISCUSSION

We Managed to achieve an accuracy of 88% despite a bias in the dataset, this places us firmly as the top performing study published which used the same dataset, for example, [3] only managed to reach 76.4%, despite their complex model architecture and resources. Criticisms to our work are the limited nature of our model, due to the limited size, and the bias in the dataset, coupled with the limitations which face any AI model when moving from the lab to the real world. Our model bundles together all types of pneumonia under one label for the purposes of simplification, this subtracts from its real-life value.

5. CONCLUSION

In conclusion, we find that AI provides a strong opportunity for use in low-income countries as an alternative to trained medical personnel, not ignoring the inherent limitations of our model and all AI models in general such as the trust problem and ethical considerations. Lastly we recommend that more work and resources be allocated to this type of research, and future research to attempt to tack the limitations found in this paper.

ACKNOWLEDGMENT

6. REFRENCES

[1] "Radiology Facing a Global Shortage." https://www.rsna.org/news/2022/may/Global-Radiologist-Shortage (accessed Nov. 09, 2022).

[2] "WHO consolidated guidelines on tuberculosis: module 2: screening: systematic screening for tuberculosis disease." https://www.who.int/publications-detail-redirect/9789240022676 (accessed Nov. 09, 2022).

[3] Z. Meng, L. Meng, and H. Tomiyama, "Pneumonia Diagnosis on Chest X-Rays with Machine Learning," *Procedia Comput. Sci.*, vol. 187, pp. 42–51, Jan. 2021, doi: 10.1016/j.procs.2021.04.032.

[4] S. Hwa, A. Bade, M. Hijazi, and M. Jeffree, "Tuberculosis detection using deep learning and contrastenhanced canny edge detected X-Ray images," *IAES Int. J. Artif. Intell. IJ-AI*, vol. 9, p. 713, Dec. 2020, doi: 10.11591/ijai.v9.i4.pp713-720.

[5] S. Guefrechi, M. B. Jabra, A. Ammar, A. Koubaa, and H. Hamam, "Deep learning based detection of COVID-19 from chest X-ray images," *Multimed. Tools Appl.*, vol. 80, no. 21–23, pp. 31803–31820, 2021, doi: 10.1007/s11042-021-11192-5.

[6] S. Kazemzadeh *et al.*, "Deep Learning Detection of Active Pulmonary Tuberculosis at Chest Radiography Matched the Clinical Performance of Radiologists," *Radiology*, p. 212213, Sep. 2022, doi: 10.1148/radiol.212213.

[7] J. D. Schroeder *et al.*, "Prediction of Obstructive Lung Disease from Chest Radiographs via Deep Learning Trained on Pulmonary Function Data," *Int. J. Chron. Obstruct. Pulmon. Dis.*, vol. 15, p. 3455, 2020, doi: 10.2147/COPD.S279850.