International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 OPEN ACCESS
Available Online at : www.ijsrcseit.com
doi : https://doi.org/10.32628/CSEIT2410116

Machine Learning-Based Pneumonia Detection in Chest X-rays: A Comprehensive Study

Ardon Kotey[1], Hariaksh Pandya, Mithil Kadam[2], Vedant Jamthe, Reeve Gonsalves, Lalith Samanthapuri, Kushagra Bande[3], Udit Srinivasan[4]

[1]Department of Information Technology, Dwarkadas J. Sanghvi College of Engineering, Mumbai, Maharashtra, India
[2]Department of Artificial Intelligence and Data Science, Thakur College of Engineering and Technology, Mumbai
[3]Department of Artificial Intelligence and Machine Learning, Vellore Institute of Technology, Bhopal
[4]Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Vadapalani Campus

ARTICLEINFO

Article History:

Accepted: 01 Feb 2024
Published: 12 Feb 2024

Publication Issue
Volume 10, Issue 1
January-February-2024

Page Number
160-165

ABSTRACT

In recent years, artificial intelligence and machine learning has proved to be remarkable in the medical field. The medical sector, however, requires a high level of accountability and thus transparency. Explanations for machine decisions and predictions are thus needed to justify their reliability. This requires greater interpretability, which often means we need to understand the mechanism underlying the algorithms. Unfortunately, the blackbox nature of deep learning is still unresolved, and many machine decisions are still poorly understood. The reason radiologists are weary of using AI is because they do not trust a model to predict ailments without any form of explainability. Thus, we aim to create a system that not only focuses on interpretability and explainability but also has a high enough accuracy to make it reliable enough to be trusted and used by medical practitioners.

Keywords : Pneumonia Detection, X-Ray Imaging, Machine Learning, Deep Learning, Diagnostic Accuracy, Clinical Decision Support

I. INTRODUCTION

The objective of the research and development domain known as Explainable AI (XAI) is to render artificial intelligence (AI) systems human-comprehensible, interpretable, and transparent. XAI aims to develop artificial intelligence systems capable of elucidating their reasoning, actions, and decision-making processes in a manner that is accessible to non-experts. XAI programmes strive to produce models that are more comprehensible while preserving a solid level of performance. In the domain of medical imaging, for instance, XAI implementation could significantly improve patient care. Increasingly, AI algorithms are being implemented in medical imaging to aid in treatment and diagnosis decisions. However, medical professionals may occasionally find it challenging to interpret and comprehend the results generated by these algorithms, especially when complex pattern recognition or deep learning models are involved. XAI can facilitate healthcare professionals' comprehension of the rationale behind AI-generated diagnoses, thereby empowering them to render more informed decisions, by increasing the explainability of AI-powered tools.

For instance, XAI can be utilised in radiology to assist in the explanation of the decisions rendered by AI algorithms tasked with the analysis of medical images, including X-rays, MRI scans, and CT scans. Through the implementation of XAI methodologies, such as feature attribution and model visualisation, these algorithms can generate justifications for their identification of a specific condition or disease, elucidating the precise image characteristics that contributed to the diagnosis.

XAI can be utilised to increase the accountability and transparency of AI medical imaging systems. Averting apprehensions regarding the capacity of AI to sustain preexisting biases or inaccuracies in medical diagnoses can be accomplished through the provision of rationales by XAI.

II. LITERATURE REVIEW

A. Literature Review

CheXNet is an algorithm that was introduced in the paper "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning" by Pranav Rajpurkar and Jeremy Irvin. An algorithm was devised to identify pneumonia from chest X-rays with an accuracy surpassing that of radiologists in practice. The ChexNet algorithm is trained using the largest publicly accessible chest X-ray dataset, ChestX-ray14. It comprises more than 100,000 frontal X-ray images of more than 30,000 distinct patients. CheXNet, a 121-layer convolutional neural network, processes a chest X-ray image as input and generates a heatmap delineating the regions of the image that are most suggestive of pneumonia and the probability of the ailment. Randomly dividing the dataset into training (70%), testing (20%), and validation (10%) subsets ensured that there was no overlap in the patient counts. To facilitate comparison with radiologists, 420 frontal chest X-rays were gathered as a test set. Having 4, 7, 25, and 28 years of experience, the four radiologists were tasked with labelling each X-ray. The model and the labels assigned to the radiologists were inputted into a standardised data entry programme. Both the model and the radiologist were not granted any access to patient information. As ground truth, the F1 score (with 95% CI) of individual radiologists and CheXNet was calculated against each of the other four classifications. Consequently, CheXNet attains an F1 score of 0.435, surpassing the mean score of 0.387 achieved by radiologists (0.383, 0.356, 0.365, 0.442). In recent years, deep learning models has made a lot of progress, giving us a lot of insights. The reliability in these models has also increased, due to high levels of accuracy. But a few sectors, like healthcare, need transparency as well as reliability. Unfortunately, the black box mechanism, while it ensures high reliability and accuracy, fails to be transparent towards the user. As these models grow more and more complex, understanding the working of the model becomes more and more difficult- almost impossible. To combat this issue, explainability is implemented to make the model more transparent in its working. For explainability in images, 3 algorithms were selected - Grad-CAM, LIME, and SHAP. In the paper, ‘Evaluating the performance of the LIME and Grad-CAM explanation methods on a LEGO multi-label image classification task’, they found that Grad-CAM provides a more detailed insight than LIME from the view of core performance and people seemed to trust Grad-CAM more than LIME. SHAP makes use of certain valves calculated by complex mathematics called Shapley values. SHAP is also not the best suited for images among the three algorithms.

B. Methodology

1. Dataset

Mr. Paul Mooney is the creator of the dataset that was used. The dataset contains chest X-ray pictures that were chosen from retrospective cohorts of paediatric patients aged one to five years old who were treated at the Guangzhou Women and Children's Medical Centre in Guangzhou. These images were taken in an anterior-posterior orientation. All of the chest X-ray imaging was carried out as a standard part of the clinical treatment that patients received. Initially, all chest radiographs were checked for quality control, and any scans that were of poor quality or were illegible were removed. This was done in preparation for the analysis of chest x-ray pictures. After that, the diagnosis for the photographs were evaluated by two highly qualified medical professionals before being given the go-ahead to be trained by the AI system. In addition, the assessment set was examined by a third expert in order to take into account any grading mistakes that may have occurred.

2. Our Approach

During the nascent phases of development, the algorithm under consideration was Local Interpretable Model-agnostic Explanations (LIME). This widely used XAI technique furnishes explanations for the predictions generated by black-box machine learning models. By approximating the predictions of complex models locally with an interpretable model, such as a decision tree or linear model, LIME aims to render the predictions of such models interpretable. It operates by perturbing the input data at random to produce a set of samples that are similar but subtly distinct. By utilising the black-box model, the class or outcome for each of these samples is subsequently predicted. Subsequently, an interpretable model is trained using the resulting predictions to approximate the behaviour of the black-box model in the immediate vicinity of the input data. One benefit of utilising LIME is its applicability in elucidating the predictions generated by machine learning models, irrespective of their internal architecture or structure. By offering explanations in relation to the input features, LIME facilitates practitioners' comprehension of the rationale behind a specific prediction.

But LIME does possess a few disadvantages. The purpose of LIME is to furnish explanations by analysing the input characteristics. The resulting explanations may be influenced, nonetheless, by the interpretable model selected for approximation. An illustration of this can be seen in the potential bias of explanations towards linear associations between the outcome and the features when a linear model is employed. The term for this is explanation bias. Furthermore, the performance of LIME is contingent upon the selection of hyperparameters, including the magnitude of the local neighbourhood and the interpretable model employed for approximation. The quality of the explanations produced can be substantially influenced by the hyperparameters selected and determining the most effective hyperparameters can present a formidable task.

To address a portion of these obstacles, GradCam, an alternative algorithm under consideration, was selected as the ultimate algorithm implemented. GradCam is a method utilised to visually represent the significance of distinct image regions in relation to the forecasts generated by a deep neural network. By means of a heatmap, which emphasises the areas of an image that make the greatest contribution to a specific prediction, it enables professionals to acquire comprehension regarding the decision-making process of the network. GradCam operates by calculating the gradient of the forecast in relation to the activations of the network's final convolutional layer. The prediction is then backpropagated to the activations of the final convolutional layer, which represent distinct regions of the input image, using this gradient. Subsequently, the heatmap is generated by assigning weights to the activations of each region, thereby emphasising the areas that hold the utmost significance in the prediction.

3. Data Pre-processing and Augmentation

As two fundamental preprocessing stages, all incoming X-ray images are resized to 224x224x3 and normalised by converting each pixel value from 0-255 to a float pixel value in the range 0-1.

An extensive amount of knowledge is required for deep learning to produce dependable results. Obtaining sufficient information can be challenging in certain circumstances, especially when it comes to medical matters, due to the expensive and time-consuming nature of the process. Data augmentation addresses this issue through the more efficient utilisation of pre-existing data. It is implemented during the training phase subsequent to the pre-processing and dividing of the datasets in order to augment the training data, mitigate the potential for overfitting, and enhance the accuracy.

4. Classification Method

The features are extracted from the EfficientNetB0 model that has been pre-trained. These characteristics are transmitted to the linear SVM classifiers to predict pneumonia chest X-ray images and classify the output. Except for the last twenty layers, the EfficientNetB0 architecture freezes the layers in order to update the weights using the utilised dataset. In EfficientNetB0, the completely connected layers are eliminated and four additional layers are incorporated. The initial layer is a collapsed layer in which the feature matrix is transformed into a single feature vector. Detection-wise, the second layer comprises 512 neurons. The Dropout layer in the third layer has a 0.2-percentage-point value, while the linear SVM classifier in the fourth layer is based on the Hinge loss function. Support Vector Machines (SVMs) are binary classifiers that differentiate between the two classes in lieu of the Sigmoid function.

Hyperparameters of Hybrid model

Learning Rate	0.00001
Optimizer	Adam
Batch Size	32
Max Epoch	100

Fig 1: Hinge Loss SVM Based Classifier

III.RESULTS AND DISCUSSION

A. Technology Used

The following technology was used:

1. **Machine learning frameworks,** such as scikit-learn, TensorFlow, and PyTorch, are widely utilised in the development and training of machine learning models.
2. **Deep Learning Models:** Convolutional Neural Networks (CNNs) are frequently utilised for feature extraction and image classification. This includes pre-trained models such as VGG, ResNet, and Inception.
3. **Data Preprocessing Tools:** For image preprocessing, libraries such as OpenCV and Pillow are utilised to perform operations such as noise reduction, resizing, and normalisation.
4. **Medical Image Libraries:** To manage medical images in formats such as DICOM, libraries such as DICOM (Digital Imaging and Communications in Medicine) are utilised.
5. **Python Programming:** Python's extensive library and versatility make it the language of choice for implementing machine learning algorithms and libraries.
6. **Data Augmentation Techniques:** To enhance the diversity of the dataset, various data augmentation techniques are implemented, including random cropping, rotation, and turning.
7. Jupyter notebooks offer a collaborative and interactive setting that facilitates the exploration of data, development of models, and visualisation of results.

8. **Cloud computing:** AWS, Google Cloud, and Microsoft Azure are examples of cloud platforms that are utilised to execute distributed and scalable machine learning tasks.
9. Model evaluation tools, such as scikit-learn, offer functionalities to assess the performance of models by utilising metrics including accuracy, precision, recall, and F1-score.
10. **Visualisation Libraries:** For data visualisation and model performance analysis, libraries such as Matplotlib and Seaborn are utilised.

11. **Medical Imaging Databases:** For training and evaluating the models, databases containing chest X-ray images, such as the NIH Chest X-ray dataset, are utilised.

By leveraging machine learning techniques on X-ray images, these technologies facilitate the creation of pneumonia detection systems that are both resilient and precise.

B. Classification Report

	Precision	Recall	F1 Score	Support
0.0	0.97	0.95	0.96	235
1.0	0.98	0.99	0.99	644
Macro-Average	0.98	0.97	0.97	879
Weighted -Average	0.98	0.98	0.98	879

Fig 2- Classification Report

The efficacy of the GradCam model was subjected to a rigorousevaluation in our project, which produced noteworthy outcomes. The model's proficiency was evidenced by its extraordinary accuracy rate of 97.95% in the identification of pneumonia from chest X-ray images. Furthermore, the loss function, which is an essential metric in assessing the performance of the model, was computed to be 5.73%. This value signifies the model's resilience in effectively reducing errors throughout the training process. The results highlight the effectiveness of our methodology in utilising sophisticated machine learning methods to identify pneumonia with accuracy and dependability. This provides encouraging potential for improving diagnostic precision in clinical environments.

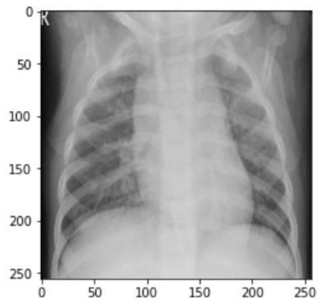
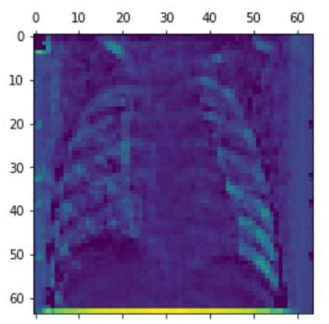
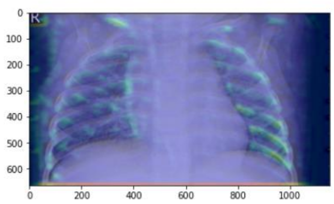
Input Image	
Generated heatmap	
Output after superimposing the heatmap on the original image	

Fig 3- Heatmap Generation from Xrays

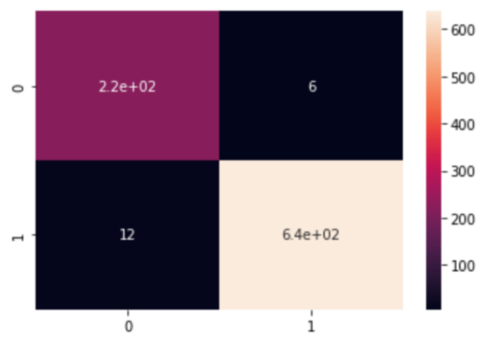


Fig 4- Confusion Matrix

IV. CONCLUSION

In brief, the model we have presented not only exhibits exceptional precision but also furnishes significant elucidation, a critical element in healthcare implementations specifically within the domain of chest radiography. By bridging the divide in access to expertise for interpreting medical images, this system has the potential to fundamentally transform the diagnostic process for diseases. In addition to its ability to detect pneumonia, our model demonstrates considerable potential for the future. By being expanded to diagnose a wide range of additional medical conditions and fractures, it significantly improves its utility within the healthcare domain. Additionally, there is considerable potential for further improvement, as the integration of supplementary features and functionalities can yield an even more comprehensive and invaluable resource for healthcare practitioners. Anticipating the future, our research unveils potential breakthroughs that have the capacity to significantly influence the healthcare sector and enhance patient results.

V. FUTURE SCOPE

The prospective implications of our research are far-reaching and paradigm-shifting. In addition to its notable precision and comprehensibility, our model possesses the capacity to fundamentally transform healthcare through its ability to identify a wide range of musculoskeletal abnormalities, fractures, and pulmonary diseases in X-ray images, thereby augmenting its diagnostic applicability. It can contribute to large-scale screening programmes for infectious diseases, facilitate telemedicine consultations in remote areas, and serve as the foundation for real-time decision support in radiology. Moreover, the explainable AI functionalities of our model facilitate opportunities for personalised medication and ongoing enhancement. Ensuring ethical conduct and adhering to regulatory requirements will be of paramount importance in our pursuit of universal healthcare accessibility and impact. Looking ahead, our research establishes a foundation for a future in which sophisticated AI technologies improve healthcare results and revolutionise the way medical knowledge is imparted.

REFERENCES

[1] Rajpurkar, Pranav & Irvin, Jeremy & Ball, Robyn & Zhu, Kaylie & Yang, Brandon & Mehta, Hershel & Duan, Tony & Ding, Daisy & Bagul, Aarti & Langlotz, Curtis & Patel, Bhavik & Yeom, Kristen & Shpanskaya, Katie & Blankenberg, Francis & Seekins, Jayne & Amrhein, Timothy & Mong, David & Halabi, Safwan & Zucker, Evan & Lungren, Matthew. (2018). Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. PLOS Medicine. 15. e1002686. 10.1371/journal.pmed.1002686.

[2] Erico Tjoa, and Cuntai Guan, Fellow, IEEE "A Survey on Explainable Artificial Intelligence (XAI): towards Medical XAI", JOURNAL OF LATEX CLASS FILES, VOL. 14, NO. 8, AUGUST 201

[3] Feiyu Xu, Hans Uszkoreit, Yangzhou Du, Wei Fan "Explainable AI: A Brief Survey on History, Research Areas, Approaches and Challenges", Natural Language Processing and Chinese Computing

[4] Arun Das, Graduate Student Member, IEEE, and Paul Rad, Senior Member, IEEE "Opportunities and Challenges in Explainable Artificial Intelligence (XAI): A Survey"

[5] Shaw-Hwa Lo, Yiqiao Yin "A novel interaction-based methodology towards explainable AI with better understanding of Pneumonia Chest X-ray Images

[6] Bas H.M. van der Velden, Hugo J. Kuijf, Kenneth G.A. Gilhuijs, Max A. Viergever "Explainable artificial intelligence (XAI) in deep learning-based medical image analysis", Image Sciences Institute, University Medical Center Utrecht, Utrecht, The Netherlands.

Cite this article as :

Ardon Kotey, Hariaksh Pandya, Mithil Kadam, Vedant Jamthe, Reeve Gonsalves, Lalith Samanthapuri, Kushagra Bande, Udit Srinivasan, "Machine Learning-Based Pneumonia Detection in Chest X-rays: A Comprehensive Study", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 10, Issue 1, pp.160-165, January-February-2024. Available at doi : https://doi.org/10.32628/CSEIT2410116

Journal URL : https://ijsrcseit.com/CSEIT2410116