PNEUMONIA DETECTION

Ajay Kumar[1], Nikhil Bansal[2], Srishti[3], Muskan Maheshwari[4], Nikhil Verma[5]

Assistant Professor, Inderprastha Engineering College[1]

Student, Inderprastha Engineering College[2-3-4-5]

ABSTRACT

Pneumonia is a bacterial or viral respiratory infection that affects many people, especially in developing and poor countries where pollution, unsanitary living conditions and overpopulation, and inadequate health care infrastructure are common. increase.

Pneumonia causes pleural effusion, a condition in which fluid fills the lungs and causes breathing problems. Early detection of pneumonia is critical to ensure curative treatment and improve survival. A chest x-ray is the most common way to detect pneumonia. However, the examination of chest radiographs is a highly subjective and difficult task.

In this study, we developed a computer-assisted diagnostic method for automatic detection of pneumonia using chest radiographs. We created an ensemble of three convolutional neural network models using GoogLeNet, ResNet-18, and DenseNet-121.

The weights given to elementary school students were calculated using a proprietary method in the weighted average ensemble method. A weight vector is generated by fusing the values of four typical metrics: accuracy, recall, f1 value, and area under the curve. This is an approach that is often experimentally established in literature studies and prone to imprecision.

Using a 5-fold cross-validation scheme, the proposed approach was tested on his two published pneumonia X-ray datasets published by Kermany et al. provided. The sensitivity rates are 98.80% and 87.02% respectively. The results surpassed state-of-the-art methods and our method outperformed widely used ensemble techniques. The robustness of our approach was demonstrated by statistical analysis of the data set using McNemar's and ANOVA tests.

Keywords:- *Diseases, Lung, Feature Extraction, X-Rays, Support Vector Machine, Task Analysis, Convolutional Neural Networks*

---

1. PROLOGUE

Pneumonia is acute pneumonia caused by bacteria, viruses, or fungi that infect the lungs, causing inflammation of the air sacs and fluid filling of the lungs, pleural effusion. More than 15% of deaths in children under the age of 5 are due to this disease [1]. Most cases of pneumonia occur in poor and developing countries where medical resources are limited and environmental factors such as overpopulation, pollution, and unclean environments exacerbate the problem. Early detection can prevent disease become fatal. Diagnosis of lung disease often requires radiographic examination using computed tomography (CT), magnetic resonance imaging (MRI), or radiography (X-rays). X-rays are a non-invasive and inexpensive method of examining the lungs. A case of pneumonia and his radiograph of a healthy lung are shown in Figure 1. On a chest X-ray known as an infiltrate. white spots. However, chest radiography for the detection of pneumonia is subject to subjective variability [2,3]. Therefore, pneumonia detection should be automated. Using an ensemble of deep transfer learning models, this work developed a computer-aided diagnosis (CAD) system that accurately diagnoses chest radiographs.

2. PURPOSE

• Conducted a comprehensive literature review on pneumonia detection using medical imaging to identify gaps in current research.

• Evaluate algorithm performance on large data sets of medical images such as: B. Chest X-ray or CT scan.

• Provide visualizations or descriptions of features that the pneumonia detection algorithm uses to improve interpretability.

• Discussion of the practical implications of algorithms in clinical practice, such as: B. Early detection and prompt treatment of pneumonia.

3. BACKGROUND RESEARCH

This paper references the following papers from 1984 to 2022, which can be summarized as follows:

1. Pneumonia is a common and potentially life-threatening lung infection that can be caused by bacteria, viruses, or fungi. A chest x-ray or CT scan is often used to diagnose pneumonia.

2. Deep learning algorithms show promising results in automating pneumonia detection from medical images. Convolutional neural networks (CNNs) are often used for this purpose.

3. Several studies have reported high accuracy of CNN-based algorithms in detecting pneumonia using chest X-rays and CT scans. However, there is still room for improvement in terms of speed, interpretability, and generalizability.

4. Some of the challenges associated with detecting pneumonia using medical imaging include variability in image quality, the complexity of lung anatomy, and the presence of other lung diseases that can mimic pneumonia. included.

5. Availability of large datasets of labeled medical images. B. Chest X-ray14 and the NIH chest X-ray data set facilitated the development and evaluation of pneumonia detection algorithms.

6. Interpretability of deep learning algorithms is an important aspect in clinical settings. Techniques such as Grad-CAM and LIME can be used to generate visualizations of image regions that are most important for algorithmic decisions. 7. The ethical and regulatory implications of the use of medical imaging for research and clinical purposes, including privacy and patient safety, should be carefully considered and addressed.

8. Future research directions in pneumonia detection from medical images include developing algorithms that can detect pneumonia in its early stages, integrating clinical data into the algorithms, and evaluating the performance of the algorithms in real clinical settings. there is a possibility.

9. The incidence of pneumonia is higher in certain population groups. B. Older people, children, and people with underlying medical conditions. Therefore, it is important to evaluate the performance of pneumonia detection algorithms in different patient populations.

4. METHODOLOGY

1. Client Interface (Front end): This is the user interface where the user will use to interact with the system and perform automated pneumonia detection using medical images. It is designed using a web framework and will allow the user to upload medical images, view the results of the pneumonia detection algorithm, and visualize the regions in the image that are most important for the algorithm's decision.

2. Upload Feature: This module enables the user to upload medical images for pneumonia detection. The system will accept medical images in DICOM or other commonly used medical image formats.

3. Core Functionality Modules:

3.1 Pre-processing: In this module, the medical images uploaded by the user will undergo pre-processing to improve the quality of the images and remove noise. This may include techniques such as resizing, cropping, and normalization.

3.2 Pneumonia Detection: In this module, a deep learning algorithm such as a CNN will be used to detect pneumonia in the medical images uploaded by the user. The algorithm will output a probability score for each image, indicating the likelihood of pneumonia being present.

3.3 Interpretability: In this module, techniques such as Grad-CAM and LIME will be used to generate visualizations of the regions in the image that are most important for the algorithm's decision. This will help the user to understand why the algorithm made a particular diagnosis.

Fig -1: Information Page

Fig -2: Image Upload Page

4. Database: This module will store the medical images uploaded by the user and the results of the pneumonia detection algorithm. This will allow the user to view the results of previous pneumonia detection analyses and track the progress of the disease over time.

5. Login/Register Page: This module enables a new user to register themselves on the system and then login into the system to perform pneumonia detection in a more secured and customized manner. The user will also be able to view their past pneumonia detection analyses and access their stored medical images.

5. EXPERIMENTAL SETUP

Pandas is a free Python library for data analysis and data processing. Pandas provides a variety of powerful and easy-to-use data structures and operations for working with data in the form of numerical tables and time series.

NumPy is a free Python software library for numerical computations on data. Numerical computations can be in the form of large arrays or multidimensional matrices. These multidimensional matrices are the main objects in NumPy, their dimensions are known as axes and the number of axes is known as rank.

Matplotlib is a data visualization and 2D plotting library for Python. Matplotlib can be used to create many visuals such as charts, graphs, histograms, power spectra, stem charts and any other desired visualization charts.

Seaborn is a Python data visualization library based on Matplotlib and tightly integrated with NumPy and Pandas data structures.

Streamlit is a Python-based open source app framework. This allows him to build great data science and machine learning web apps in less time. Compatible with major Python libraries such as Scikit-learn, Keras, PyTorch, Latex, NumPy, Pandas, Matplotlib.

6. IMPLEMENTATION

The client interface, i.e. the frontend designed for viewer interaction with the system, was developed in Python using Stremlite. This means that the user's first interaction with the system is through the Streamlit-designed user interface. All buttons such as data upload, data analysis user preference dropdown menu and other his UI buttons are designed in Steamlit. For example; the button for selecting the type of activity in the system is designed with a "selection field" and a "save" function. Similarly, the option to upload a file was created using Streamlit's file_uploader function.

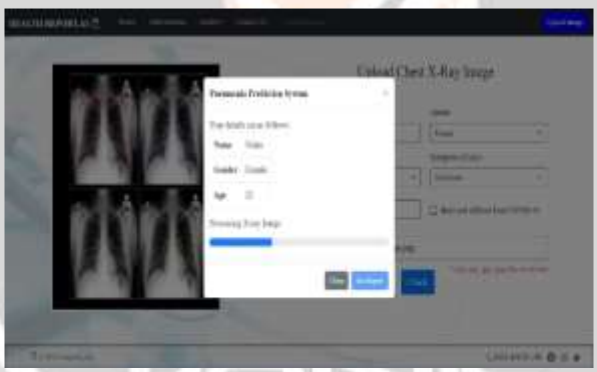

Fig -3: Processing Image

7. THREATS TO VALIDATE

Limited dataset size: The accuracy of the pneumonia detection algorithm may be limited by the size of the dataset used to train it. While the system is suitable for small to medium sized datasets, the accuracy may decrease if larger datasets are used.

Restricted dataset formats: The system may have a restriction on the types of medical image formats that can be used for pneumonia detection. This may limit the usefulness of the system for certain medical institutions or researchers who use different image formats.

Limited automated techniques: The interpretability module of the system may use a limited number of automated techniques for visualizing the regions of the medical image that are important for the pneumonia detection algorithm's decision. This may limit the user's ability to understand why the algorithm made a particular diagnosis.

Biased dataset: The accuracy of the pneumonia detection algorithm may be influenced by the bias in the dataset used to train it. If the dataset is not representative of the population, the algorithm may not perform as well on new, unseen medical images.

Generalizability to other populations: The pneumonia detection algorithm may perform differently on medical images from populations that are different from the population used to train the algorithm. This may limit the generalizability of the system to other populations.

8. CONCLUSION

Early detection of pneumonia is crucial for determining the appropriate treatment of the disease and preventing it from threatening the patient's life. Chest radiographs are the most widely used tool for diagnosing pneumonia; however, they are subject to inter-class variability and the diagnosis depends on the clinicians' expertise in detecting early pneumonia traces. To assist medical practitioners, an automated CAD system was developed in this study, which uses deep transfer learning-based classification to classify chest X-ray images into two classes "Pneumonia" and "Normal." An ensemble framework was developed that considers the decision scores obtained from three CNN models, Google Net, ResNet-18, and DenseNet-121, to form a weighted average ensemble.

The weights assigned to the classifiers were calculated using a novel strategy wherein four evaluation metrics, precision, recall, f1-score, and AUC, were fused using the hyperbolic tangent function. The framework, evaluated on two publicly available pneumonia chest X-ray datasets, obtained an accuracy rate of 98.81%, a sensitivity rate of 98.80%, a precision rate of 98.82%, and an f1-score of 98.79% on the Kermany dataset and an accuracy rate of 86.86%, a sensitivity rate of 87.02%, a precision rate of 86.89%, and an f1-score of 86.95% on the RSNA challenge dataset, using a five-fold cross-validation scheme. It outperformed state-of-the-art methods on these two datasets. Statistical analysis of the proposed model using McNemar and ANOVA tests demonstrate the feasibility of the approach. Moreover, the proposed ensemble model is domain-independent and therefore applicable to a wide variety of computer vision tasks.

However, as mentioned above, in some cases the Ensemble framework failed to generate correct predictions. In the future, techniques such as image contrast enhancement and other preprocessing steps may be explored to improve image quality. We can also consider segmenting the lung images before classification so that the CNN model can achieve improved feature extraction. Moreover, since three CNN models are required to train the proposed ensemble, the computational cost is higher than that of CNN baselines developed in literature studies. In the future, we may try to reduce computational requirements using methods such as snapshot ensembles.

9. REFERENCES

[1] WHO Pneumonia. World Health Organization. (2019)

[2] Neuman M., Lee E., Bixby S., Diperna S., Hellinger J., Markowitz R., et al. Variability in the interpretation of chest radiographs for the diagnosis of pneumonia in children. Journal Of Hospital Medicine

[3] Williams G., Macaskill P., Kerr M., Fitzgerald D., Isaacs D., Codarini M., et al. Variability and accuracy in interpretation of consolidation on chest radiography for diagnosing pneumonia in children under 5 years of age.

[4] Kermany D., Zhang K. & Goldbaum M. Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification. (Mendeley,2018)

[5] Lal S., Rehman S., Shah J., Meraj T., Rauf H., Damaševičius R., et al. Adversarial Attack and Defence through Adversarial Training and Feature Fusion for Diabetic Retinopathy Recognition. Sensors. 21, 3922 (2021) pmid:34200216

[6] Rauf H., Lali M., Khan M., Kadry S., Alolaiyan H., Razaq A., et al. Time series forecasting of COVID-19 transmission in Asia Pacific countries using deep neural networks. Personal And Ubiquitous Computing. pp. 1–18 (2021) pmid:33456433

[7] Deng J., Dong W., Socher R., Li L., Li K. & Fei-Fei, L. Imagenet: A large-scale hierarchical image database. 2009 IEEE Conference On Computer Vision And Pattern Recognition. pp. 248-255 (2009)

[8] Dalhoumi S., Dray G., Montmain J., Derosière, G. & Perrey S. An adaptive accuracy-weighted ensemble for inter-subjects classification in brain-computer interfacing. 2015 7th International IEEE/EMBS Conference On Neural Engineering (NER). pp. 126-129 (2015)

[9] Albahli S., Rauf H., Algosaibi A. & Balas V. AI-driven deep CNN approach for multi- label pathology classification using chest X-Rays. PeerJ Computer Science. 7 pp. e495 (2021) pmid:33977135

[10] Rahman T., Chowdhury M., Khandakar A., Islam K., Islam K., Mahbub Z., et al. Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray. Applied Sciences. 10, 3233 (2020)

[11] Liang G. & Zheng L. A transfer learning method with deep residual network for pediatric pneumonia diagnosis. Computer Methods And Programs In Biomedicine. 187 pp. 104964 (2020) pmid:31262537

[12] Ibrahim A., Ozsoz M., Serte S., Al-Turjman F. & Yakoi P. Pneumonia classification using deep learning from chest X-ray images during COVID-19. Cognitive Computation. pp. 1–13 (2021) pmid:33425044

[13] Zubair S. An Efficient Method to Predict Pneumonia from Chest X-Rays Using Deep Learning Approach. The Importance Of Health Informatics In Public Health During A Pandemic. 272 pp. 457 (2020)

[14] Rajpurkar P., Irvin J., Zhu K., Yang B., Mehta H., Duan T., et al. & Others Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. ArXiv Preprint ArXiv:1711.05225. (2017)

[15] Albahli S., Rauf H., Arif M., Nafis M. & Algosaibi A. Identification of thoracic diseases by exploiting deep neural networks. Neural Networks. 5 pp. 6 (2021)

[16] Chandra T. & Verma K. Pneumonia detection on chest X-Ray using machine learning paradigm. Proceedings Of 3rd International Conference On Computer Vision And Image Processing. pp. 21-33 (2020)

[17] Kuo K., Talley P., Huang C. & Cheng L. Predicting hospital-acquired pneumonia among schizophrenic patients: a machine learning approach. BMC Medical Informatics And Decision Making. 19, 1–8 (2019) pmid:30866913

[18] Yue H., Yu Q., Liu C., Huang Y., Jiang Z., Shao C., et al. & Others Machine learning-based CT radiomics method for predicting hospital stay in patients with pneumonia associated with SARS-CoV-2 infection: a multicenter study. Annals Of Translational Medicine. 8 (2020) pmid:32793703

[19] Sharma H., Jain J., Bansal P. & Gupta S. Feature extraction and classification of chest x-ray images using cnn to detect pneumonia. 2020 10th International Conference On Cloud Computing, Data Science & Engineering (Confluence). pp. 227-231 (2020)

[20] Stephen O., Sain M., Maduh U. & Jeong D. An efficient deep learning approach to pneumonia classification in healthcare. Journal Of Healthcare Engineering. 2019 (2019) pmid:31049186