



Pneumonia Classification Using Sparse Auxiliary Pneumonia Detection Network

Leena Vinod Patil, Bindu Garg

Bharati Vidyapeeth Deemed to be university, Pune

Abstract:

Pneumonia is a serious pulmonary condition that impacts millions globally. An early analysis of pneumonia is essential for effective treatment and improved survival rates. This has created an urgent requirement for rapid detection and classification methods for pneumonia to facilitate efficient treatment and prompt recovery of affected individuals. The advancement and cost-effectiveness of chest X-ray technology, coupled with the progress in artificial intelligence, have led to increased global interest in pneumonia identification utilising deep learning and chest X-ray imaging. This study examines the use of Sparse Auxiliary Pneumonia Detection Networks (SAPD-net) for classifying pneumonia through chest X-ray images. Consequently, SAN is an efficient system developed to address the challenge of adapting to recent discrimination in pneumonia images. The objective is to develop an interpretable evaluation framework for pneumonia infection classification, utilising deep classification and transfer learning techniques.

The proposed method's experimental results are validated using the benchmark chest X-rays database. The simulation results indicated the superior performance of the proposed method compared to alternative techniques. The results indicate that deep features yield precise and reliable attributes for pneumonia detection. The proposed method enables radiologists to evaluate pneumonia patients and provide a swift diagnosis.

Keywords-Pneumonia detection, Convolutional Neural Network, Chest X-ray, Deep Learning networks, Sparse, Auxiliary

INTRODUCTION:

Pneumonia is a significant respiratory condition characterised by inflammation of one or both lungs, leading to symptoms such as fever, coughing, and respiratory distress. According to the World Health Organisation, 2.5 million deaths due to pneumonia were reported in 2019, with 14% occurring in children aged 0–5 years [1]. Pneumonia is the leading cause of mortality in children globally, responsible for 22% of deaths among those aged 1 to 5 years, according to the World Health Organisation (WHO) in 2019 [2]. The elevated mortality rate necessitates the diagnosis of pneumonia to prevent the failure of human bodily functions. Pneumonia is a prevalent disease resulting from various microbial species, including bacteria, viruses, and fungi. The term "Pneumonia" is derived from the Greek word "Pneumonia," which refers to the lungs. The term pneumonia pertains to a disease of the lungs. Pneumonia is defined as a disease characterised by inflammation of one or both lung parenchyma [3]. The condition



impacts one or both lungs, leading to inflammation of the lung parenchyma, which encompasses the lung tissue responsible for gas exchange, including the pulmonary alveoli. Inflammation leads to the accumulation of pus or fluid in the lung's alveoli, resulting in symptoms such as shortness of breath, cough, chest pain, and fever. Early diagnosis is essential for the effective treatment of pneumonia and the improvement of patient outcomes. Identifying pneumonia remains a complex task. The condition may be easily mistaken for other pulmonary diseases, and the interpretation of X-ray images can differ markedly among radiologists, complicating consistent and accurate diagnosis [4]. Common diagnostic tools for pneumonia include chest X-rays, chest computed tomography (CT) scans, magnetic resonance imaging (MRI) of the chest, and chest ultrasound scans. Conventional imaging techniques commonly utilised include computed tomography (CT), magnetic resonance imaging (MRI), and chest radiographs. Chest X-rays are preferred due to their non-invasive nature and cost-effectiveness [6].

The current pneumonia tests based on chest X-rays necessitate a substantial volume of clinical data for training and learning purposes. Additionally, the diagnosis necessitates the involvement of specialised radiologists and clinicians. The diagnosis relies on markers, clinical manifestations, and physiological test results. The use of imaging for pneumonia detection is an emerging research area. In contrast to conventional image processing algorithms in computer vision, deep learning algorithms eliminate the laborious task of manual feature extraction, directly yielding detection results through input-output mapping. Deep learning methods have become essential for enhancing the accuracy and effectiveness of numerous diagnostic procedures. Advancements in deep learning have notably enhanced the ability of medical experts to diagnose pneumonia by supporting their decision-making processes. Deep learning models have proven to be highly effective in the detection and diagnosis of various diseases and conditions within medical imaging. The efficacy of DCNNs in the field of computer vision has led researchers to utilise these networks for the segmentation of COVID-19 infections in CT images. Utilising deep learning models enables healthcare professionals to improve diagnostic accuracy and facilitate informed treatment decisions for patients with suspected pneumonia. These models can autonomously learn and extract intricate features from medical images for the purposes of disease identification and classification. Convolutional Neural Networks (CNNs) have demonstrated efficacy in the detection of pneumonia through chest X-ray imaging.

The remainder of the paper is organised into the following sections: Section II reviews related work, presenting the literature on current research and technologies. Section III outlines the methodology, detailing the research approach, data collection methods, and data analysis techniques employed. Section IV presents the results and discussion of the proposed approach. Section V outlines the supporting arguments and situates them within a wider framework.



LITERATURE SURVEY:

Parthasarathy et al. [9] introduce a novel Computer Aided Diagnosis utilising the Harris Hawks Optimiser in conjunction with Deep Learning (CAD-HHODL) for the detection of pneumonia in chest X-ray images. The CAD-HHODL method analyses chest X-ray images for the identification and classification of pneumonia. Asnake et al. [10] propose a method for the identification and classification of chest X-ray images pertaining to normal individuals and those infected with pneumonia. The developed deep learning model initially preprocesses X-ray images to extract relevant features, subsequently segments them through a threshold segmentation technique, and identifies individuals with normal conditions and pneumonia infections. M. Ali et al. [11] implemented and evaluated six deep learning models: CNN, InceptionResNetV2, Xception, VGG16, ResNet50, and EfficientNetV2L. The Adam optimiser is also incorporated, effectively adjusting the epoch for all models.

Nalluri, S et al. [12] present a deep learning model, NASNet (Neural Architecture Search Network), pre-trained on ImageNet, designed to predict pneumonia at a very early stage using chest X-rays of patients. R. Das et al. [13] diagnose pneumonia through chest X-ray (CXR) imaging by employing VGG-19, a component of the CNN model, to develop an autonomous pneumonia detection system. H. Wu et al. [14] introduce a deep learning approach utilising Retinanet for pneumonia detection, employing a novel predicted boxes fusion algorithm termed Fuzzy Non-Maximum Suppression (FNMS). This method enhances the robustness of predicted boxes by merging overlapping detection boxes. Wang, T., et al. [15] propose a Vision Transformer (ViT)-based model for accurate diagnosis utilising channel-based attention on X-ray images of the lung, applying multi-head attention to channel patches instead of feature patches. Chen Y. et al. [16] propose a lung image classification model utilising synthetic image data augmentation within the framework of the medical image multi-domain translation algorithm MI-GAN, which is based on the key migration branch. Haozhe J. et al. [17] present a segmentation network incorporating a novel pixel-wise sparse graph reasoning (PSGR) module aimed at segmenting COVID-19 infected regions in CT images. The PSGR module, positioned between the encoder and decoder of the network, enhances the modelling of global contextual information. Kalane P. et al. [18] proposed an automated Covid-19 detection system that utilises indications from Computer Tomography (CT) images to train a novel deep learning model based on the U-Net architecture. Xia, X. et al. proposed an automated framework for lung CT image infection segmentation, termed Probabilistic Graphical Model U-Net (PGM-U-Net). The entire framework employs iterative training in an end-to-end manner utilising a general back-propagation mechanism, incurring minimal computational overhead from the PGM component. Chandra V. et al. [20] present a two-stage pipeline for the automated detection of lung diseases and the identification of infection regions. A multi-class classification approach is employed to accurately identify lung illnesses utilising a Convolutional Hidden Markov Neural Network (CHMNN) classifier. Ren H. et al. developed and tested a multimodal model named CheXMed, which integrates clinical notes with image data to enhance pneumonia detection performance in older adults.



METHODOLOGY:

This section presents the experimental dataset and methodology for the detection and classification of pneumonia using x-ray images. This study employed a network of convolutional neural networks for the categorisation and identification of pneumonia. The 22-layer architecture enables the learning of more complex features than shallower networks, thereby enhancing performance in pneumonia detection. The design concept allows for execution on individual devices with limited computing capabilities. The design incorporates two auxiliary classifier layers connected to the output of the Inception layers.

- The architectural features of auxiliary classifiers are as follows:
- A pooling layer utilising an average filter of size 5x5 with a stride of 3. A 1x1 convolution employing 128 filters for sparse representation, accompanied by ReLU activation.
- A fully connected layer with 1025 outputs and ReLU activation.
- Regularisation of dropouts with a dropout ratio of 0.8.
- A softmax classifier that produces two classes, similar to the fundamental softmax classifier.

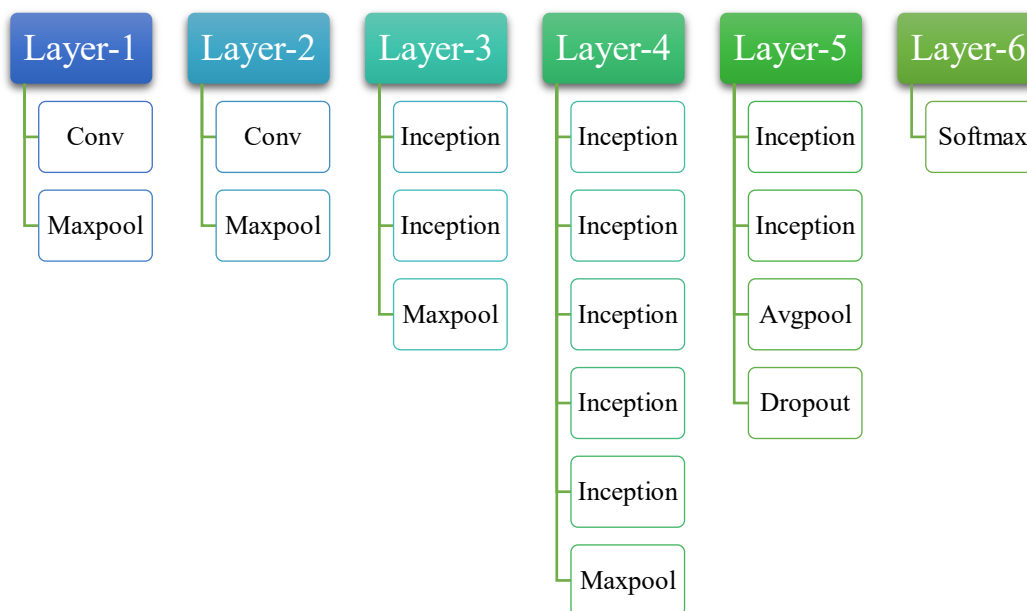


Figure 1: Proposed Architecture

The Inception modules process inputs at multiple scales concurrently, enabling the network to capture a broader spectrum of features. The application of 1x1 convolutions effectively manages the parameter count and mitigates overfitting, particularly in deep neural networks. The system architecture serves as the foundational element of the proposed strategy



for the categorisation and identification of pneumonia cases utilising deep learning, as illustrated in Fig. 1. The initial hidden layer receives data from the input layer and subsequently transmits categorisation scores to subsequent layers. Only active nodes transmit categorisation scores to the subsequent layer. The output layer categorises target values (genuine or false) according to the final classification scores, indicating that the output represents the expected class. This process is referred to as feed-forward or forward-propagation.

To calculate node outputs, H_i , first compute node nets, Z_i , which are the sum of the input weights (W_{ij}) multiplied by their corresponding inputs (I_{ij}) and biases b_i .

$$Z_i = \sum(W_{ij} * I_{ij} + b_i) \quad (1)$$

Then, if the target classes cannot be separated linearly, the activation functions as well as node outputs are as follows:

$$H_i = \frac{1}{1 + E^{-Y_i(H_i)}} \quad (2)$$

where E is the overall error and H_i denotes the H_i Hidden nodes.

a) Layer 1 & 2:

It begins with a series of convolutional layers that extract low-level features (e.g., edges, textures). These layers apply filters to the input, generating feature maps.

b) Layer 3:

It consists of multiple branches:

1x1 Convolutions: These are used to reduce the dimensionality of the feature maps.

3x3 and 5x5 Convolutions: These capture spatial features at different scales.

Pooling Layers: Max pooling is applied to down-sample feature maps.

Each branch processes the input in parallel, and their outputs are concatenated along the depth dimension, allowing the network to learn rich feature representations.

c) Layer 4:

This intermediate layers help combat the vanishing gradient problem during training by providing additional gradient signals.

Back-Propagation:

As seen in Equation 2, back-propagation involves calculating gradients with regard to the weights. Back-propagation is used to minimise the loss function whereas forward-



propagation is used for calculating outputs. The loss function is a measure of how much predictions deviate from the desired value. To do this, we must reverse the weights and biases. Before making any changes to the weights, the total error (E) is calculated

$$E = \sum(0.5 * (T_i - H_i)^2) \quad (3)$$

where T_i is target values. Next, the learning rate (L_i) multiplied with the derivative of the total error w.r.t weights are subtracted from the given weights:

$$\frac{dE}{dW_i} = \frac{dE}{dH_{i+1}} * \frac{dH_{i+1}}{dH_i} * \frac{dH_i}{dW_i} \quad (4)$$

$$W_i = W_i - (L_i * \frac{dE}{dW_i}) \quad (5)$$

Either choose a learning rate that results in a severe loss function, or decrease it gradually between training iterations.

d) Layer 5:

After passing through several Inception modules, the feature maps are flattened. A final fully connected layer is applied, typically followed by a softmax activation function to produce class probabilities.

The input image is resized to 512x512 and normalized before being fed into the network.

$$X = [x_1, x_2, \dots, x_k] \quad (6)$$

K denotes the input image pixels. Then normalization carried out by adjusting and scaling the activations. It computes the mean and variance of the mini-batch and uses these to normalize the activations. This helps to reduce internal covariate shift and allows for higher learning rates. Normalization layers are essential for stabilizing and accelerating the training of DNNs.

$$x = \left[\frac{x - \min}{\max - \min} \right] \quad (7)$$

Convolutional layers allow to learn features at various levels of abstraction. Each Layer 3,5, &5 module includes multiple convolutional filters of varying sizes (1x1, 3x3, 5x5) that operate simultaneously. This multi-scale approach enables the network to learn diverse features efficiently. 1x1 filters are particularly significant for reducing dimensionality before applying larger convolutions, helping to control the number of parameters and computational load.



After convolution layer we got estimated the weight(w), bias (b_j).

$$x_i^{l,j} = \sigma[b_j + \sum_{a=1}^m w_a^j x_{i+a-1}^{l-1,j}] \quad (8)$$

Activation function is indicated by the variable σ . The proposed architecture helps to maintain a healthy gradient flow through the network during training, reducing issues like vanishing gradients. This is partly due to the structure of the convolutional layers and the use of auxiliary classifiers.

Max pooling reduces the spatial dimensions of the feature maps, which decreases the number of parameters and computations in the network. It captures the most prominent features within a pooling window (typically 2×2) by taking the maximum value. This helps retain important spatial information while discarding less significant details, effectively summarizing the feature maps. By focusing on the maximum value within a region, max pooling provides some invariance to small translations and distortions in the input image. This means that slight changes in the input will not drastically affect the output, enhancing the model's robustness.

$$x_i^{l,j} = \max_{n=1}^r (x_{(i-1)*T+n}^{l-1,j}) \quad (9)$$

where n is pooling size and T is pooling stride.

Reducing the size of feature maps through max pooling leads to faster training times and lower memory usage, which is critical when dealing with large datasets and complex models.

Following equation models the Hidden layer to output .Proposed method have this capability

$$h_t = g(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \quad (10)$$

$$z_t = g(W_{hz}h_t + b_z) \quad (11)$$

here, g indicates elementwise nonlinearity(it can be sigmoid or hyperbolic tangent), x_t is the input $h_t \in \mathbb{R}^N$ is the hidden state having hidden units equals to N . Output is denoted by Z_t at instant t .

pixel sequence (x_1, x_2, \dots, x_T) having T number of coefficient, then h_1 (letting $h_0 = 0$), $z_1, h_2, z_2, \dots, h_t, z_t$.



Time Complexity of the Proposed architecture:

The overall time complexity of proposed model can be approximated as the sum of the complexities from all the convolutional layers, pooling layers, and fully connected layers. Given its depth (22 layers) and the complex structure of Inception modules, the actual computation may vary, but it remains efficient due to the parallel nature of operations.

$$O(\sum_{l=1}^d n_{l-1} \cdot s_1^2 \cdot n_l \cdot m_1^2) \quad (12)$$

above equation calculate the time complexity of the convolutional layer. Layer index is indicated by l which is d in number. Total filter number is n_l in l th layer and their spatial size is denoted by s_1 . Channel number of is n_{l-1} at l th layer. Output have spatial size of m_1 . Initial Convolutional layer required 7% of the execution time.

EXPERIMENT AND RESULTS:

By comparing with the above classical detection frameworks, we can find that the proposed framework can achieve the best results. The assessment of pneumonia detection models utilizing deep learning with chest X-ray images encompasses the evaluation of various metrics, including accuracy, recall, precision and the F1 score. Accuracy gauges the model's proficiency in accurately predicting positive cases, while specificity evaluates its aptitude in correctly identifying negative cases. Precision quantifies the fraction of accurately classified positive cases out of all predicted positive cases. The F1 score presents a harmonized measure by balancing both precision and recall, thereby offering healthcare practitioners a comprehensive tool to inform their decisions and ultimately enhance patient care. The following equations are used for accuracy, precision, recall and F1 score:

Accuracy:

This is the most important metric for evaluating a model and is defined as the proportion of correct predictions to the total number of predictions made by the model. It is evaluated using Eq.

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

The evaluation of the model's accuracy involves the examination of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN).

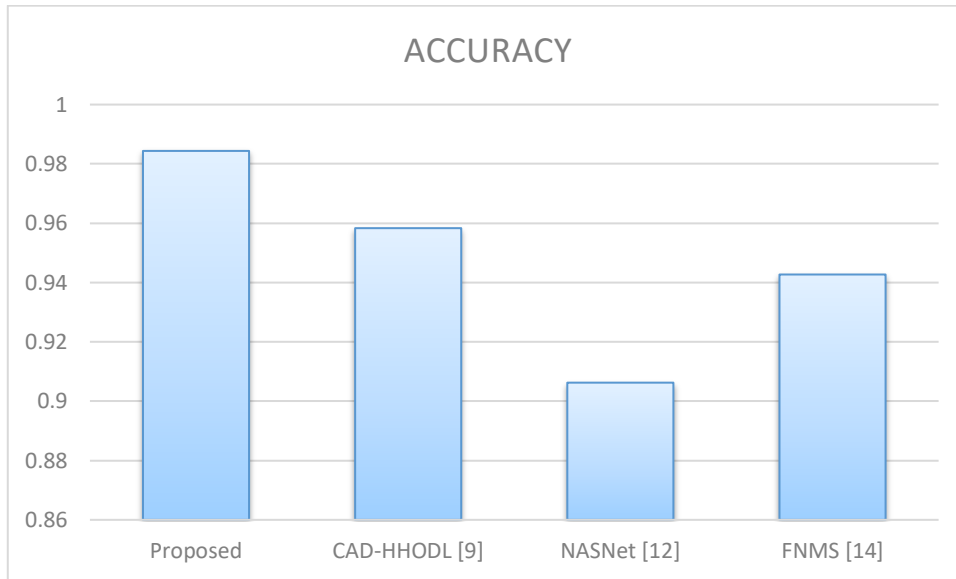


Fig: 2 Comparison of Accuracy

Precision:

Higher-precision classifiers produce fewer false positives. High accuracy reduces the likelihood of misclassifying negative instances as positive in numerous applications where false positives have severe consequences. Precision is calculated by Eq

$$\text{Precision} = \frac{(\text{TP})}{(\text{TP} + \text{FP})}$$

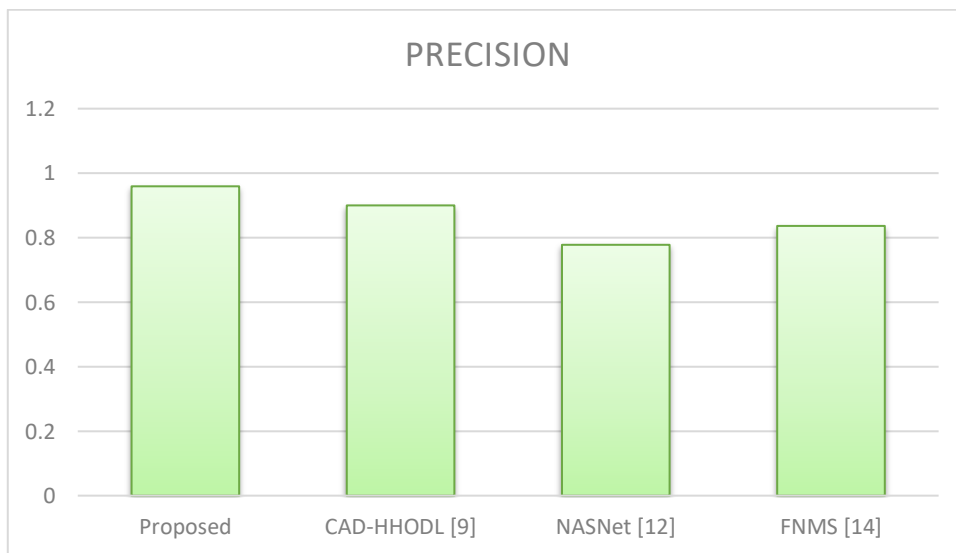


Fig: 3 Comparison of Precision



Recall (sensitivity or true positive rate) :

Classifiers with higher recall have fewer false negatives. The classifier captures positive cases and reduces false negatives. A classifier with lower recall has more false negatives. The recall is determined by Eq.

$$\text{Recall} = \frac{(\text{TP})}{(\text{TP} + \text{FN})}$$

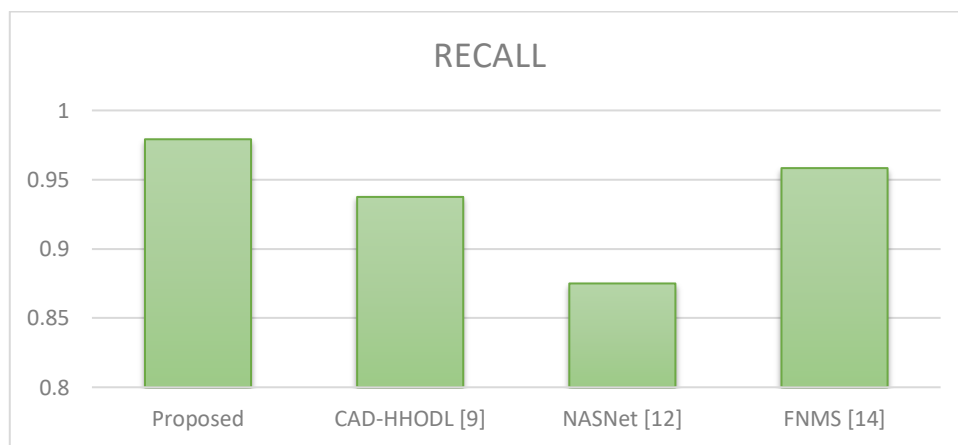


Fig: 4 Comparison of Recall

F1 Score :

The F1 Score is the harmonic mean of precision and recall, indicating patterns between them and calculated using Eq.

$$\text{F1 Score} = \frac{2 * \text{Precision} * \text{Recall}}{(\text{Precision} + \text{Recall})}$$

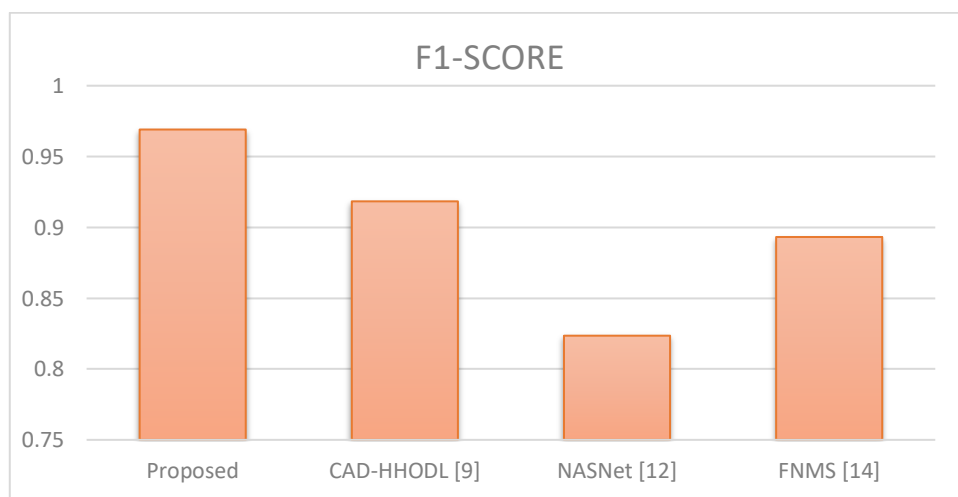


Fig: 5 Comparison of F1-Score



Discussion:

Pneumonia is the primary cause of infant mortality globally. Pneumonia is a respiratory infection caused by microbes and other environmental factors. This infection affects the lungs, resulting in fluid accumulation and respiratory distress, and is the primary cause of mortality in children under five years of age. Early diagnosis is the only feasible strategy for mitigating the disease's impact on the patient.

Deep learning techniques can assist doctors to detect the areas of pneumonia in the chest X-rays images. Existing methods do not adequately account for the significant variation scale and the indistinct boundaries of the pneumonia area. This study introduces a reinforcement learning approach utilising a sparse dense network for the detection of pneumonia. Interpretable deep classification offers a thorough analysis of transparency and facilitates transfer learning, resulting in enhanced accuracy. This paper addresses the issues of inadequate attention in current translation models, limited capacity for key transfer and generation, subpar quality of generated images, and insufficient detailed features by investigating pneumonia classification using a Sparse Dense Network. Ensemble learning methods were employed to construct the classification model for optimisation purposes. This research employs hyperparameter fine-tuning to examine the effects of learning rates, batch sizes, and other critical parameters on the accuracy of the classification model, thereby enhancing the understanding of convolutional neural networks in medical analysis. A comprehensive evaluation of convolutional neural network architectures reveals significant accuracies. The findings underscore the importance of Deep Learning Networks, specifically the proposed Sparse Dense model, in improving pneumonia detection accuracy and decreasing mortality rates.

CONCLUSION:

Pneumonia is an infectious disease that can be life-threatening. Diagnosis is typically achieved through physical examinations and diagnostic imaging techniques, including chest X-rays, ultrasounds, or lung biopsies. This study examines the use of Sparse Dense Networks for classifying pneumonia through chest X-ray images. The study employs rigorous experimentation and data analysis to demonstrate the model's learning capabilities, achieving an accuracy of 96% in pneumonia classification.

REFERENCES:

- 1] Sharma, S., et.al. A systematic literature review on deep learning approaches for pneumonia detection using chest X-ray images. *Multimed Tools Appl* 83, 24101–24151 (2024). <https://doi.org/10.1007/s11042-023-16419-1>
- 2] World Health Organization: WHO. Pneumonia [EB/OL]. Who.int. World Health Organization:WHO2019-08-02. <https://www.who.int/en/news-room/fact-sheets/detail/pneumonia>



- 3] Obaro SK, Madhi SA. Bacterial pneumonia vaccines and childhood pneumonia: are we winning, refining, or redefining? *Lancet Infect Dis.* 2006;6(3):150–61.
- 4] H. P. Action, “Key facts: Poverty and poor health,” <https://www.healthpovertyaction.org/news-events/key-facts-poverty-and-poor-health/> , last update: 2018-01.
- 5] Yaseliani, M.; Hamadani, A.Z.; Maghsoodi, A.I.; Mosavi, A. Pneumonia Detection Proposing a Hybrid Deep Convolutional Neural Network Based on Two Parallel Visual Geometry Group Architectures and Machine Learning Classifiers. *IEEE Access* 2022, 10, 62110–62128.
- 6] RadiologyInfo, “Chest x-ray,” <https://www.radiologyinfo.org/en/info/chestrad> , last update: 2022-11 .
- 7] Yang, F. et al. Pneumoconiosis computer aided diagnosis system based on X-rays and deep learning. *BMC Med Imaging* 21, 1–7 (2021).
- 8] Khatri, A., et al. Pneumonia identification in chest X-ray images using EMD. In *Trends in Communication, Cloud, and Big Data: Proceedings of 3rd National Conference on CCB, 2018* (Springer Singapore, 2020).
- 9] Parthasarathy, V., Saravanan, S. Computer aided diagnosis using Harris Hawks optimizer with deep learning for pneumonia detection on chest X-ray images. *Int. j. inf. tecnol.* 16, 1677–1683 (2024). <https://doi.org/10.1007/s41870-023-01700-1>
- 10] Asnake, N.W., Salau, A.O. & Ayalew, A.M. X-ray image-based pneumonia detection and classification using deep learning. *Multimed Tools Appl* 83, 60789–60807 (2024). <https://doi.org/10.1007/s11042-023-17965-4>
- 11] M. Ali et al., "Pneumonia Detection Using Chest Radiographs With Novel EfficientNetV2L Model," in *IEEE Access*, vol. 12, pp. 34691-34707, 2024, <https://doi.org/10.1109/ACCESS.2024.3372588>
- 12] Nalluri, S., Sasikala, R. A deep neural architecture for SOTA pneumonia detection from chest X-rays. *Int J Syst Assur Eng Manag* 15, 489–502 (2024). <https://doi.org/10.1007/s13198-022-01788-x>
- 13] R. Das, D. S. K. Nayak, C. P. Rout, L. Jena and T. Swarnkar, "Deep Learning Techniques for Identification of Pneumonia: A CNN Approach," 2024 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC), Bhubaneswar, India, 2024, pp. 1-5, doi: 10.1109/ASSIC60049.2024.10507933.
- 14] H. Wu, H. Lu, M. Ping, W. Zhu and Z. Li, "A Deep Learning Method for Pneumonia Detection Based on Fuzzy Non-Maximum Suppression," in *IEEE/ACM Transactions on*



- Computational Biology and Bioinformatics, vol. 21, no. 4, pp. 902-911, July-Aug. 2024, doi: 10.1109/TCBB.2023.3247483.
- 15] Wang, T., Nie, Z., Wang, R. et al. PneuNet: deep learning for COVID-19 pneumonia diagnosis on chest X-ray image analysis using Vision Transformer. *Med Biol Eng Comput* 61, 1395–1408 (2023). <https://doi.org/10.1007/s11517-022-02746-2>
- 16] Chen, Yunfeng & Lin, Yalan & Ding, Jinzhen & Li, Chuzhao & Zeng, Yiming & Xie, Weifang & Huang, Jianlong. (2022). Multi-domain medical image translation generation for lung image classification based on Generative Adversarial Networks. *Computer Methods and Programs in Biomedicine*. 229. 107200. 10.1016/j.cmpb.2022.107200.
- 17] Haozhe, Jia & Tang, Haoteng & Ma, Guixiang & Cai, Weidong & Huang, Heng & Zhan, Liang & Xia, Yong. (2023). A convolutional neural network with pixel-wise sparse graph reasoning for COVID-19 lesion segmentation in CT images. *Computers in Biology and Medicine*. 155. 106698. 10.1016/j.compbimed.2023.106698 .
- 18] Kalane, Prasad & Patil, Sarika & Patil, B P & Sharma, Davinder. (2021). Automatic Detection of COVID-19 Disease using U-Net Architecture Based Fully Convolutional Network. *Biomedical Signal Processing and Control*. 67. 102518. 10.1016/j.bspc.2021.102518.
- 19] Xia, X. et al. (2022) ‘Fully automated segmentation of pneumonia infection based on Probabilistic Graphical Model and U-Net blend network’, *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 10(6), pp. 608–615. doi: 10.1080/21681163.2021.2015447.
- 20] Chandra, Vinod & R., Nitha. (2024). Novel CEFNet framework for lung disease detection and infection region identification. *Biomedical Signal Processing and Control*. 10.1016/j.bspc.2024.106624.
- 21] Ren, Hao & Jing, Fengshi & Chen, Zhurong & He, Shan & Zhou, Jiandong & Liu, Le & Jing, Ran & Lian, Wanmin & Tian, Junzhang & Zhang, Qingpeng & Xu, Zhongzhi & Cheng, Weibin. (2024). CheXMed: a Multimodal Learning Algorithm for Pneumonia Detection in the Elderly. *Information Sciences*. 654. 119854. 10.1016/j.ins.2023.119854.