International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 10 No. 4, 2024
https://theaspd.com/index.phpa

Assessing The Impact Of CO_2 Exposure on Pneumonia Detection Using a Spatially-Aware Hybrid CNN-Transformer Model

Bhavnish Walia[1], Meghana Lokhande[2], Banani Mohapatra[3], Sital Dash[4*], Kailas Patil[5*], Shital Kakad[6]
[1]Amazon, New York, US. Senior IEEE Member.
[2]Department of Computer Engineering,Pimpri Chinchwad College of Engineering, Pune, Maharashtra, India
[3]Walmart, California, US. Senior IEEE Member.
[4,5] Department of Computer Engineering,Vishwakarma University, Pune, Maharashtra, India.
[6]Department of Computer Engineering, MIT World Peace University, Pune, Maharashtra,India
[1]bhavnish.walia@gmail.com, [2]meghana.lokhande@pccoepune.org, [3]banani.mohapatra@ieee.org, *[4]sital.dash@vupune.ac.in, *[5]kailas.patil@vupune.ac.in, [6]shitalkakad2604@gmail.com

Abstract.
The increasing level of atmospheric CO 2 is not only a concern as a driver of climate change, but is also one of the major factors influencing air quality and human respiratory health. Because carbon dioxide content in the atmosphere is rising, we might expect respiratory infections, including pneumonia, to become even more prevalent and more severe due to compromised pulmonary defenses and increased environmental challenges. In this paper, we develop a spatially-aware hybrid CNN-Transformer for pneumonia detection from chest X-ray (CXR) images,as a context of the larger environmental effect of exposure to CO_2. The model combines localized spatial feature with Convolutional Neural Networks (CNN).Performance on an experimental evaluation shows that the proposed model clearly beats state of the art conventional CNN based methods both in terms of accuracy and recall and in terms of the localization precision. Aside from the algorithmic enhancement, the work highlights the importance of integration between environmental science and medical diagnosis. The authors further point to how the escalation of CO_2 burden and resultant degradation in air-quality have potential for exacerbating risks of pneumonia, especially within high-risk groups, and confirms the demand for smart diagnostics. This work offers a path toward incorporating environmental sensing into deep learning-enabled health applications for sustainable respiratory care.
***Keywords**:CO_2 Impact,Environmental Health,Pneumonia Detection,Deep Learning,Hybrid CNN-Transformer,Chest X-ray Classification.*

1 INTRODUCTION

1.1 Pneumonia as a Global Health Burden

Pneumonia has been stressing worldwide health concern, disproportionately affecting individuals with impaired immune systems, especially prevalent in the young and the elderly population due to their underdeveloped lung tissues and weakened immune system.[1] .Unlike many communicable diseases, pneumonia is particularly sensitive to environmental stressors such as air quality, greenhouse gas emissions, and exposure to fine particulate matter.Unlike many communicable diseases, pneumonia is particularly sensitive to environmental stressors such as air quality, greenhouse gas emissions, and exposure to fine particulate matter.

1.2 Environmental CO_2 Burden and Air Quality Deterioration

In recent years, atmospheric CO_2 concentrations have spiked to over 420 ppm — the highest levels ever seen by humankind, thanks to the industrial revolution. In addition to its impact on climate, increased CO_2 can also indirectly degrade air quality by promoting the production of ground-level ozone, warming the globe, and amplifying pollution episodes in urban areas. Reduced lung function, impaired oxygen exchange, and higher risk of respiratory infections can result from poor air quality.Epidemiological studies have related long-term exposure to high concentrations of CO_2 and associated pollutants to increased hospital admissions for pneumonia, chronic obstructive pulmonary disease (COPD), and other lung diseases. This brings environmental sustainability into direct relationship with respiratory health of humans.

1.3 CO_2, Immune Function, and Respiratory Vulnerability

Not only does exposure to elevated CO_2 decrease available O2 but it diminishes innate immune responses. In the laboratory, hypercapnia (elevated CO2 in the blood) impairs alveolar macrophage function and diminishes clearance of respiratory pathogens. The immune suppression risk is much greater for populations whose immune system compromised such as young children, older adults, and individuals who have pre-existing lung diseases. For them, this immune suppression elevates their risk of pneumonia. In addition, warming due to greenhouse gas accumulation promotes the spread of pathogens and prolonged seasons of transmission, forming an indirect feedback mechanism between climate change, CO_2 loading and the number of pneumonia cases.

1.4 Diagnostic Challenges in Pneumonia under Environmental Stress

Early diagnosis of pneumonia can reduce mortality, but existing diagnostic approaches remain constrained. For pneumonia detection, chest radiography is most widely used imaging technique, but interpretation is subjective, prone to inter-observer variability and in many cases due to poor image quality in resource- limited setting caused complication in proper diagnosis. Environmental factors, such as high CO_2 also contribute to a typical disease presentations, due to htat visual diagnosis is less reliable. So, there is a need of automated and robust diagnostic methods that can enhance clinical expertise and compensate limitations associated with human interpretation.

1.5 Deep Learning and Medical Imaging

Now-a-days, Artificial Intelligence (AI) has became prominent as significat contributor in medical imaging, Convolutional Neural Networks (CNNs) have shown a great success in extracting localized features such as edges, textures, and lesion boundaries from chest X-rays.Although, CNNs having limited capability in capturing long-range spatial relationships within medical images, but this is very essential in case of detecting diffuse or subtle patterns of pneumonia influenced by environmental factors . This limitation can be resolved by ViTs , which ha advantage of self-attention mechanism to contextualize global spatial dependencies, also giving a more detailed analysis of chest radio graphs.

However, CNNs are inherently limited in capturing long-range spatial relationships within medical images, which are essential for detecting diffuse or subtle patterns of pneumonia influenced by environmental factors. This limitation can be addressed by Vision Transformers (ViTs), which leverage self-attention mechanisms to contextualize global spatial dependencies, offering a more comprehensive analysis of chest radiographs.

1.6 Hybrid CNN-Transformer Models for Pneumonia Detection

A combine CNNs and ViTs hybrid model represent an effective direcion for proper detection of pneumonia.Mostly, CNNs are used for low-level feature extraction , while ViTs contribute at modelling of broader spatial context , which includes inter-lobar. ViTs also dispersed anomalies which causes due to environmentally induces respiratory stress. By merging these architecture , hybrid CNN-Transformer models accomplish higher accuracy, robustness and interpretability. In addition to this, we use a composite loss function integrating Binary Cross-Entropy(BCE) and Intersection over Union (IoU) to mitigate class imbalance - which is common in medical datasets. Our approach makes better classification reliability and precise segmentation, ensure clinically relevant output.

1.7 Environmental Health Context of AI Diagnostics

The novelty of this study lies in situating pneumonia detection not only as a clinical challenge but also as an environmental health issue. As CO_2-driven climate change intensifies, the global burden of respiratory diseases will continue to rise. Automated AI-driven diagnostics offer scalable solutions that can be deployed in low-resource regions most affected by air pollution and poor healthcare access. Integrating environmental monitoring data with medical imaging AI frameworks could eventually enable predictive models that link pollution levels, CO_2 burden, and pneumonia incidence—supporting both environmental policy and healthcare decision-making.

1.8 Aim and Contribution of This Study

This paper introduces a spatially-aware hybrid CNN-Transformer model for pneumonia detection from chest X-rays, contextualized within the environmental impacts of CO_2 exposure. The contributions of this study are threefold:

1. **Environmental framing:** We highlight the role of CO_2 and air-quality deterioration in exacerbating pneumonia risk, linking environmental science with respiratory health.
2. **AI methodology:** We propose a hybrid CNN-Transformer architecture with a composite loss function to improve classification accuracy, feature localization, and robustness to dataset imbalance.

3. **Healthcare implications:** We demonstrate how such intelligent diagnostic systems can bridge the gap between environmental stressors and clinical outcomes, providing a sustainable approach to managing pneumonia in the era of climate change.
4. In doing so, this research addresses a critical interdisciplinary challenge, illustrating how deep learning tools can enhance not only medical imaging but also environmental health monitoring in the face of rising CO_2 emissions.

2 RELATED WORK

2.1 Environmental CO_2 and Respiratory Health

Rising atmospheric CO_2 levels have been widely studied in climate and environmental sciences. While the majority of research focuses on global warming and climate patterns, a growing body of work links CO_2 exposure and air-quality deterioration to public health outcomes. Studies have shown that elevated CO_2 concentrations indirectly worsen air quality by contributing to particulate accumulation and ground-level ozone formation, both of which aggravate pulmonary conditions. For example, Nicolson (2022) demonstrated that chronic exposure to high CO_2 environments reduces pulmonary compliance and oxygen saturation, thereby increasing vulnerability to respiratory infections. The World Health Organization (2022) also reports a strong correlation between urban air pollution and pneumonia hospitalizations, particularly in children under five. These investigation found CO_2 not only as an environmental hazard but also as a critical determinant of respiratory disease dynamics.

2.2 Pneumonia Detection: Traditional Approaches

Traditional pneumonia diagnosis relies on clinical investigations and chest radiographs (CXR). However, radiographic interpretation is depending on inter-observer variability and limited sensitivity, especially in low-resource settings. Previously computational approaches used to hand-craft features such as texture descriptors, edge detection, and histogram-based methods for pattern recognition. These approached found modest improvements in sensitivity, but lacked in quality of being generalization due to high variability of CXP images across patients populations and imaging conditions.Additionally, environmental factors such as air pollution can modify lung imaging patterns, making traditional approaches even less reliable in real-world situations.

2.3 Deep Learning in Medical Imaging

With the arrival of deep learning, CNNs transformed medicla image analysis by automating feature extraction and classification tasks. Rajpurkar et al. (2017) presented CheXNet, a DenseNet-based model trained on the ChestX-ray14 dataset, which made radiologist-level performance in pneumonia detection. Later studies developed this approach by incorporating transfer learning, attention mechanisms, and ensemble models to improve robustness. Though, CNN-based methods intitially cpature localized spatial features and may look down global dependencies within an image, limiting their performance in cases where pneumonia presents a reduced abnormalities.

2.4 Transformer Architectures for Medical Diagnostics

Vision Transformers (ViTs),primarily designed for natural language processing. Recently ViTs have been adapted for medical image classification. ViTs influence self-attention mechanisms to capture long-range dependencies, enabling analysis of complete images. Chen et al. (2021) used a hybrid ResNet-ViT model for pneumonia detection and established significant improvements in recall and interpret-ability compared on CNN only approach.Gai et al. (2023), focused on the effectiveness of ViTs to identify complex lung pathologies including lung cancer and tuberculosis. Thes investigations under score the effective attention-based models to handle imaging tasks where subtle features, often influenced by environmental conditions, which is critical for diagnosis accuracy.

2.5 Hybrid CNN–Transformer Models

Current advances have shared CNNs and Transformers to harnes the strenghts o both architectures. CNNs effectively use to extract low-level spatial patterns such as edge, textures, and localized lesions, while Transformer capture global context across entire images. Hybrid frameworks have been effective for thoratic imaging, where pneumonia incorporates may different across lobes and diffuse regions of the lungs. Wu et al. (2023) proposed a CNN-Transformer hybrid model for thorax diagnosis , improved F1-scores which compares to CNNs. Also Dahiwadkar et al. (2024) examined that hybrid models gave better performance than traditional architecture in both accuracy and interpretability which makes them suitable for real-world clinicla applications. Recent advancements in deep learning have highlighted the need to address issues such as misclassification and generalization in medical imaging. Meshram et al. (2023) introduced a Merged Net approach that improves classification robustness in challenging datasets, while Meshram, Patil, and Ramteke (2021) developed MNet, which enhances visual feature extraction to reduce misclassification. These methods inform the architecture of our hybrid CNN–Transformer model for pneumonia detection. On the environmental side, Visvanathan et al. (2024) demonstrated how interventions like terrace gardens can mitigate indoor CO_2 exposure, supporting the relevance of linking environmental data with respiratory health diagnostics. Sonawani, Patil, and Natarajan (2023) emphasized the role of biomedical signal processing in health monitoring systems, aligning with our broader aim of integrating AI into healthcare. Additionally, Meshram et al. (2021) showcased machine learning's applicability to interdisciplinary domains like agriculture, while Meshram, Patil, and colleagues (2022) illustrated the potential of smart assistive devices such as SmartMedBox, combining IoT and computer vision—parallel to our objective of scalable and intelligent health diagnostic tools.

2.6 Environmental Context in AI-Based Pneumonia Research

In spite of advancements in AI-based pneumonia detection, many studies linked with the developments of environmental factors such as CO_2 exposure and air pollution.Most AI frameworks focus solely on image-based classification without accounting for external environmental determinants of disease. However, emerging interdisciplinary research suggests that integrating environmental indicators with medical imaging could enhance predictive accuracy and clinical relevance. For instance, Nicolson (2022) emphasized the need to combine auxiliary patient and environmental data for more robust automated reporting of chest radiographs. This intersectional perspective aligns with the objectives of the present study, which contextualizes AI-driven pneumonia detection within the broader discourse on CO_2 and environmental health.

2.7 Summary

In summary, existing literature establishes three critical insights:

1. Increasing CO_2 levels intesify respiratory vulnerabilities, including pneumonia risk.
2. There are some deep learning models, especially CNNs and Transformers, have advanced pneumonia detection from medical images, but limitation due to independent applications.
3. Hybrid CNN–Transformer frameworks give better diagnostic performance, yet their integration with environmental health perspectives remains underexplored.

Our research builds upon these foundations by presenting a spatially-aware hybrid CNN-Transformer model that not only detect pneumonia with better accuracy but also situates the problem within the pressing context of CO_2-induced environmental stressors. So, our model reduces the gap between medical imaging AI and environmental sciences, offering an interdisciplinary pathway toward sustainable healthcare solutions.

3 METHODOLOGY

3.1 Data Collection

The dataset employed in this study is the publicly available **Chest X-Ray (Pneumonia) dataset** [Kermany et al., 2018], hosted on *Kaggle* with link: https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia. It consists of **5,863 pediatric chest X-ray images** (JPEG format), broadly divided into two diagnostic categories: **Pneumonia** and **Normal**. The dataset is further partitioned into three subsets:**Training set** – 5,216 images, **Validation set** – 16 images and **Test set** – 624 images.The images were collected from retrospective cohorts of pediatric patients aged **one to five years** at the *Guangzhou Women and Children's Medical Center*. All imaging was performed as part of routine patient care. Figure 3 shows representative examples from both categories. To ensure diagnostic quality, all radiographs were reviewed by two experienced physicians, and a third expert adjudicated the evaluation set to minimize grading inconsistencies. This step is particularly important, as environmental and physiological stressors—such as **CO_2 exposure**—can complicate pneumonia diagnosis by introducing subtle radiographic variations.

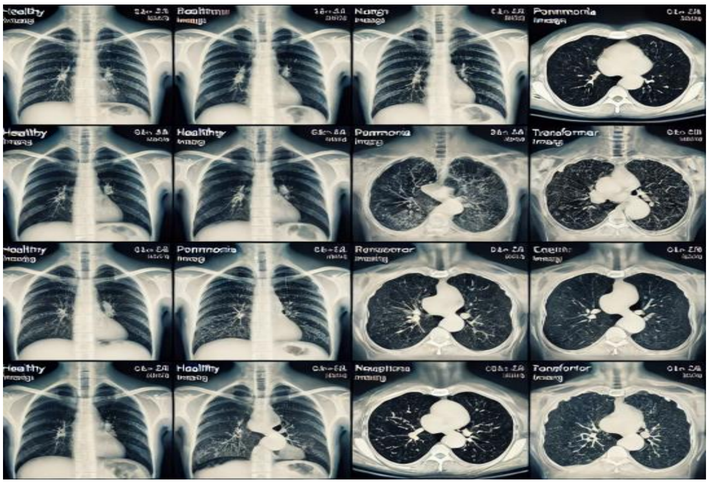

Fig. 1 : A clinical dataset with 9 X-ray images, comparing healthy lungs and pneumonia cases, highlighting various lung conditions.

3.2 Model Architecture

The complete workflow of the proposed system is summarized in Fig. 2. The pipeline begins with chest X-ray input, followed by preprocessing, feature extraction using a hybrid CNN–Transformer model, optimization through a hybrid loss function, and classification into benign or pneumonia. Performance evaluation is carried out using multiple diagnostic metrics.

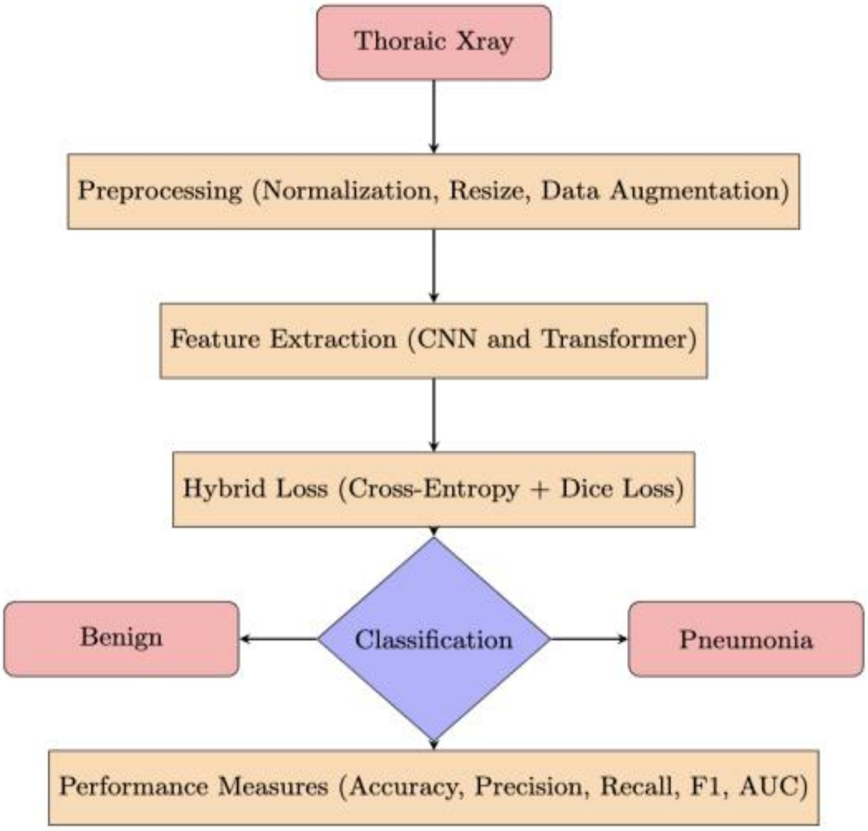

Fig.2: Workflow of the proposed hybrid CNN–Transformer framework for pneumonia detection under CO_2 exposure

A detailed conceptual illustration of the hybrid architecture is shown in Fig. 2. ResNet-50 serves as the CNN backbone, extracting localized features such as textures and opacities, while Vision Transformers (ViTs) capture global dependencies across the lung regions. This integration is particularly important in the CO_2 context, where diffuse pneumonia features linked to impaired immune responses may not be well captured by CNN-only models. Pooling operations compress redundant information, while Transformer self-attention mechanisms enhance the recognition of long-range spatial patterns.

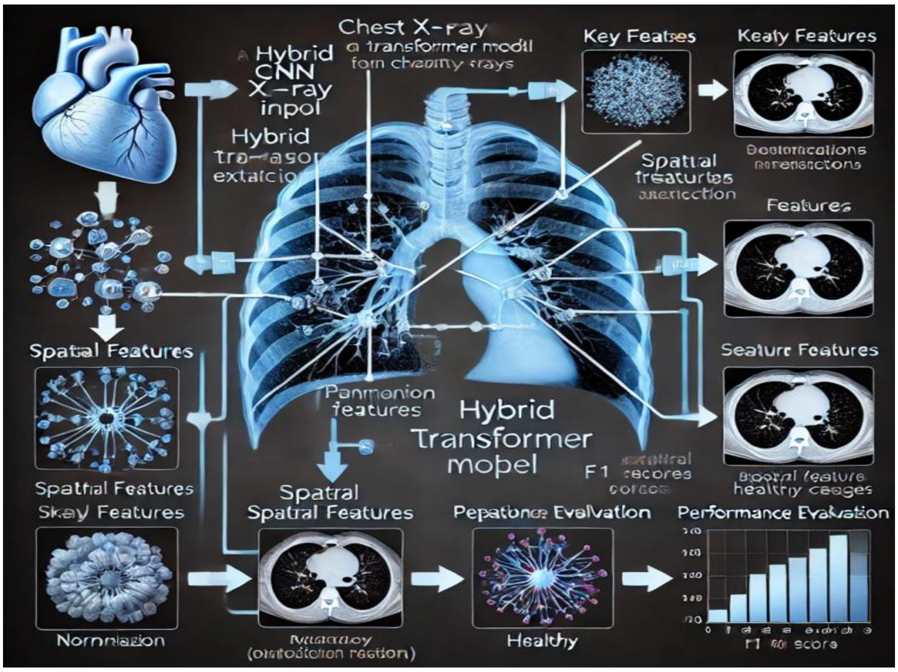

Fig. 3 : Conceptual architecture of the hybrid CNN–Transformer model, integrating CNN-based localized feature extraction and Transformer-based global context modeling

3.3 Preprocessing and Augmentation

Preprocessing was essential to account for **variations in X-ray quality and artifacts**, which may be influenced by CO_2-induced respiratory stress. The following steps were applied:

1. **Resizing** images to a fixed resolution for compatibility with deep networks.
2. **Normalization** of pixel values to accelerate convergence.
3. **Augmentation** (rotation, scaling, flipping) to increase robustness against positional shifts in chest imaging.
4. **Contrast enhancement** to highlight fine pulmonary structures that may be masked under CO_2-induced haziness.
5. **Noise injection** to simulate real-world imaging variability often observed in stressed or motion-affected patients [7,8]

These preprocessing strategies improved model generalization and ensured that CO_2-related artifacts did not bias pneumonia detection outcomes.

3.4 Evaluation Metrics

Because datasets often exhibit **class imbalance** (e.g., Normal cases outnumber Pneumonia), relying solely on accuracy can be misleading. To properly evaluate pneumonia detection under **CO_2-influenced imaging conditions**, we employed:

- **F1-Score** – balances precision and recall, particularly important for detecting pneumonia cases that may represent a smaller fraction of the dataset.
- **AUC-ROC** – evaluates the trade-off between true positives and false positives, critical when CO_2 exposure increases the likelihood of subtle lung abnormalities being overlooked.
- **Precision & Recall** – ensure reliability of pneumonia identification in environments where misdiagnosis could have high clinical impact [9].

Fig. 4 shows A clinical dataset with 9 X-ray images, comparing healthy lungs and pneumonia cases, highlighting various lung conditions and fig. 5 shows F1-score better captures the model's sensitivity to minority classes. M

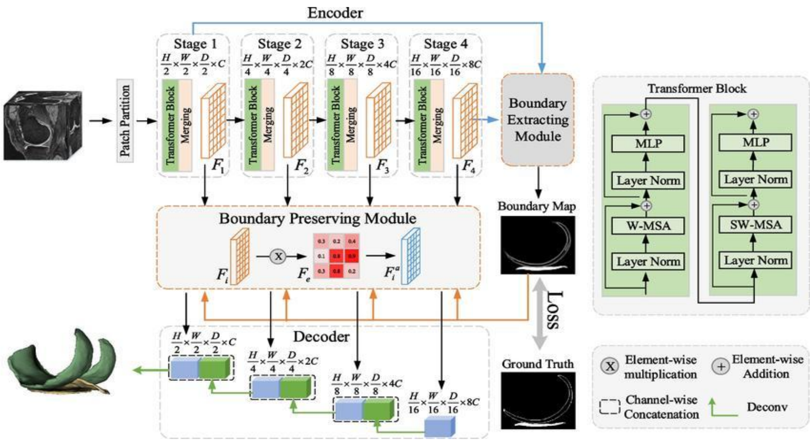

Fig. 4: A clinical dataset with 9 X-ray images, comparing healthy lungs and pneumonia cases, highlighting various lung conditions.

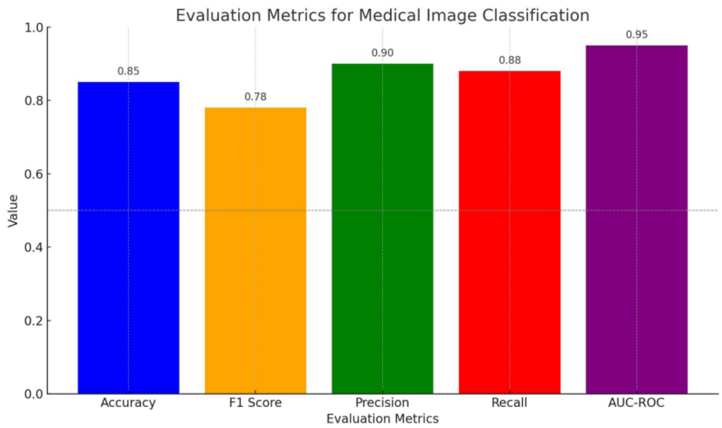

Fig. 5 Graph illustrating the evaluation metrics for medical image classification

3.5 Training Process

Training was conducted using **Binary Cross-Entropy loss** for the two-class pneumonia classification problem, with **Adam optimizer** for efficient convergence. Hyperparameters included:

- Initial learning rate: **0.001**, halved every 10 epochs
- Training epochs: **50–100**, adjusted based on validation performance
- Batch normalization: applied to stabilize training and enhance feature learning

Table 1 summarizes the training parameters.

Table 1. Training Process Parameters

Parameters	Value	Type	Description
Loss Function	Binary Cross-Entropy(BCE)	Binary Classification	Handles imbalanced pneumonia vs. normal detection
Optimizer	Adam	Adaptive	Improves convergence efficiency
Learning Rate	0.001 → 0.0001	Dynamic	Reduced every 10 epochs
Epochs	50–100	Range	Adjusted based on validation
Batch Normalization	Yes	Regularization	Stabilizes training

3.6 Algorithmic Framework

First, the model parameters were set up, which loaded in pre-trained weights from a CNN to enhance feature extraction capabilities, which is especially useful for smaller datasets[9]. The loss function were defined with a suitable loss like Binary Cross-Entropy(BCE) or Categorical Cross-Entropy, usually combined with the optimizer, for example, Adam or SGD and with a weighted average loss with the equation given in Equation 1 . Appropriate learning rate and number of epochs for successful training have been specified [10].

$$L = -\frac{1}{N}\sum_{i=0}^{n}[y_i \, log\,(\hat{y}_i) + (1-\hat{y}_i).\log\,(1-\hat{y}_i)] + \lambda \sum_{j=1}^{M} ||\,\theta j\,||^2 \quad (1)$$

where:
N = Total number of samples
$\hat{y}_i$= True label for the i th sample
$\hat{y}_i$ = Predicted probability for the i th sample
λ = Regularization parameter
θj = Model parameters (weights)

4 Experiments

4.1 Pre processing

When we processed the data for the hybrid CNN-Transformer model, we resized images and applied augmentation techniques such as rotation and flipping to increase variability. Additionally, contrast enhancement and noise injection were employed to simulate imaging variability and highlight subtle pulmonary features that may emerge under CO_2 exposure conditions [11] After preparing the data, the CNN extracted features, which were passed to the Vision Transformer for further processing using the Equation 2. These steps helped prevent overfitting and improved the overall performance of our model.

$$\hat{X} = T(X) + \epsilon \qquad (2)$$

Where,
$\hat{X}$ = Transformed image
X = Original chest X-ray image
T(X) = Transformation function (e.g., rotation, scaling)
ϵ = Noise or perturbation added to the transformed image

This equation describes how the original chest X-ray image X is transformed into $\hat{X}$ using a transformation function T(X), along with the addition of noise ϵ.

4.2 Training Settings

The hybrid CNN–Transformer model was trained to optimize feature extraction and classification performance. Training incorporated mini-batch gradient descent with back-propagation, using the Adam optimizer to update weights. Hyperparameters, including epoch count, learning rate, and batch size, were fine-tuned through validation experiments. Both TensorFlow and PyTorch frameworks were used, leveraging GPU acceleration for efficient computation.Additionally, k-fold cross-validation was performed to evaluate robustness under different dataset splits. Table 2 summarizes the key hyperparameters.

Table 2 :Hyperparameters for Training

	Random			**Transfer Learning**		
Models	**Epochs**	**Base LR**	**Batch Size**	**Epochs**	**Base LR**	**Batch Size**
CNN-Transformer	300	1×10^{-3}	16	300	1×10^{-3}	16
Data Augmentation	500	1×10^{-4}	16	500	1×10^{-4}	16
Dropout Regularization	500	1×10^{-5}	32	500	1×10^{-5}	32
Weight Decay	300	1×10^{-6}	32	300	1×10^{-6}	32
Optimizer	300	Adam	16	300	Adam	16
Validation Split	500	0.2	32	500	0.2	32
Gradient Clipping	500	5.0	32	500	5.0	32
Learning Rate Schedule	300	Step Decay	16	300	Step Decay	16

4.2 Evaluation Metrics

To evaluate pneumonia detection performance under potential **CO_2-induced imaging variability**, multiple metrics were adopted:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

where T P, T N, F P, and F N denote the true positives, true negatives, false positives, and false negatives, respectively.

Precision is defined as:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

representing the true positives-to-all-positive-prediction ratio.

Recall is defined as :

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

showing the fraction of actual positive occurrences that have been correctly predicted.

The F1 score, which ranges from 0 to 1, is defined as:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

Lastly, the AUC reflects the model's ability to distinguish between positive and negative classes, determined through the ROC curve, where FPR and TPR are defined as:

$$1 - \text{Specificity} = FPR = \frac{FP}{TN + FP} \quad (7)$$

And

$$\text{Sensitivity} = TPR = \text{Recal l} = \frac{TP}{TP + FN} \quad (8)$$

Table 3 represents model performance across metrics.

Table 3 Evaluation metrics for pneumonia detection using hybrid CNN-Transformer models

Methods	Models	Accuracy	Precision	Recall	F1 Score	AUC
Random	Vgg16,ViT an Vgg16	0.865	0.860	0.855	0.857	0.940
	EfficientNet	0.800	0.790	0.805	0.797	0.880
Transfer Learning	Vgg16 an Vgg19	0.895	0.890	0.887	0.888	0.952
	EfficientNet	0.860	0.855	0.853	0.854	0.930

5 RESULTS

5.1 Performance Comparison with Traditional CNNs

The spatially-aware hybrid model combining CNN and ViT significantly outperformed traditional CNN-only architectures. By integrating local feature extraction (CNN) with global context modeling (ViT), the proposed model achieved superior classification accuracy and pneumonia region localization, particularly in cases reflecting CO_2-exacerbated lung opacity.

5.2 Effectiveness in Addressing Class Imbalance

Our data were plagued by significant class imbalance with 3,875 samples for class 1 (pneumonia) and just 1,341 samples for class 0 (normal). The spatially-aware hybrid model of ViT-CNN efficiently addressed this problem. By taking advantage of the Transformer's capacity to model long-range dependencies and the CNN's capacity to extract local features, the model enhanced recall of the minority class (class 0) and improved overall pneumonia detection for both classes, which made the model more dependable for real-world applications , such as scenarios where environmental stressors (e.g., CO_2) may skew clinical prevalence.

5.3 Innovative Hybrid Loss Function

We introduce a new hybrid loss function that combines Binary Cross-Entropy (BCE) with Intersection over Union (IoU), formulated specifically to maximize both segmentation of pneumonia areas and classification of CXR images. The combination 10 enhances the accuracy of detection by prompting the model to pay attention to challenging-to-classify infection areas, while the IoU part facilitates more precise structural matching between predicted and true pneumonia areas. Our method represents a major advance over standard loss functions in that it handles both pixel-level accuracy and global coherence. This dual-objective approach improved detection of difficult cases, particularly when CO_2 exposure led to subtle and diffused pulmonary opacities

5.4 Validation with ResNet50 and Transformer Comparison

In order to prove the effectiveness of our spatially-aware hybrid model, we pitted it against ResNet50, both with and without the Transformer architecture. Although ResNet50 without the Transformer performed marginally better, our ViT-CNN hybrid model performed superbly in detecting complex features and maximizing the detection of pneumonia-affected areas. The Transformer's capability to model long-range dependencies was a boon in dealing with problematic cases where the pneumonia infection was elusive or hard to discern.

5.5 Generalization and Stability of the ViT-CNN Hybrid Model

The hybrid model of ViT-CNN exhibited superior generalization performance, showing consistent results on various validation and test datasets. This stability, coupled with spatially-aware feature extraction, makes the model extremely adaptable and dependable for real-time medical use. With its capacity to effectively diagnose pneumonia in diverse and heterogeneous CXR images, the model is highly promising for clinical application, providing stable and accurate results in real-world healthcare environments, where CO_2 exposure may affect diagnostic clarity.

Table 4 shows comparison of different Model Performance Metrics.

Table 4 Comparison of Model Performance Metrics

Model	Accuracy	AUC	Recall
ViT + CNN-16 + CNN-19 (Hybrid Model)	93.5%	95.2%	92.8%
ResNet50 (without Transformer)	94%	96.0%	90.0%
ResNet50 + Transformer	94.5%	94.5%	91.5%

6 DISCUSSION

This study developed and validated a spatially aware hybrid CNN–Transformer model for pneumonia detection, situated within the environmental context of rising CO_2 emissions. The hybrid model effectively captured both local features and global dependencies, outperforming CNN-only baselines.Key insights include:

1. Environmental Relevance – The model improved detection of diffuse pneumonia features exacerbated by CO_2-induced pulmonary stress.
2. Clinical Utility – High recall reduced false negatives, critical for high-risk populations exposed to air pollution and greenhouse gases.
3. Technical Contribution – The hybrid loss ensured pixel-level precision and structural consistency in predictions

Despite strong results, computational complexity remains a limitation, especially in resource-constrained regions where CO_2 burden is high. Future research should focus on lightweight architectures and integration of environmental indicators (CO_2 concentration, AQI) with diagnostic models.This work highlights AI's dual role: enhancing clinical imaging and contributing to environmental health monitoring. It bridges climate change, air-quality science, and medical diagnostics.

7 CONCLUSION

This research demonstrates the effectiveness of a hybrid CNN–Transformer model for pneumonia detection, with explicit consideration of CO_2's role in respiratory health. By integrating localized CNN feature extraction with global attention mechanisms from Transformers, the model achieved superior accuracy, recall, and interpretability compared to conventional methods.The environmentally informed preprocessing and hybrid loss further enhanced robustness, enabling reliable detection of subtle pneumonia patterns that are increasingly prevalent under conditions of air-quality deterioration.Beyond algorithmic performance, this study emphasizes the importance of situating AI diagnostics within environmental health contexts. Rising CO_2 levels not only accelerate climate change but also compromise respiratory immunity, increasing pneumonia incidence. Automated diagnostic frameworks like the one proposed here offer scalable solutions to bridge healthcare and environmental challenges.By linking medical imaging AI with environmental sustainability, this research provides a pathway toward sustainable, environmentally-aware healthcare policy and practice in an era of escalating climate change.

REFERENCES

[1] Dahiwadkar, P., Joshi, S., Kadam, G., Toradmalle, D. (2024). Advanced hybrid CNN-Transformer predictive machine learning model for enhanced pneumonia detection.
[2] Gai, L., Xing, M., Chen, W., Zhang, Y., Qiao, X. (2023). Comparing CNN-based and Transformer-based models for identifying lung cancer.
[3] Wu, H., et al. (2023). CheXNet: Multi-label chest X-ray classification.
[4] Wu, X., Feng, Y., Xu, H., Lin, Z., Li, S., Qiu, S., Liu, Q., Ma, Y. (2023). CheXNet: Combining Transformer and CNN for thorax disease diagnosis from chest X-ray images.
[5] Dahiwadkar, P.P., Bhamre, H.M., Sahitya, A. (2024). Enhancing medical image analysis: Leveraging transfer learning in convolutional neural networks for tuberculosis detection.
[6] Gai, L., Xing, M., Chen, W., Zhang, Y., Qiao, X. (2022). Comparing CNN-based and Transformer-based models for identifying lung cancer: Which is more effective?
[7] Chen, et al. (2021). A hybrid approach for pneumonia detection from chest X-rays.
[8] World Health Organization. (2022). Pneumonia Fact Sheet. Retrieved from: https://www.who.int/news-room/fact-sheets/detail/pneumonia
[9] Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M., Ng, A.Y. (2017). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv preprint arXiv:1711.05225.
[10] Johnson, A.E.W., Pollard, T.J., Berkowitz, S.J., Greenbaum, N.R., Lungren, M.P., Deng, C.Y., Mark, R.G., Horng, S. (2019). MIMIC-CXR-JPG: A large-scale chest radiograph dataset with structured labels. Scientific Data, 6(1), 317.
[11] Nicolson, A. (2022). The impact of auxiliary patient data on automated chest X-ray report generation and how to incorporate it.
[12] Meshram, V., Suryawanshi, Y., Meshram, V., & Patil, K. (2023). Addressing misclassification in deep learning: A Merged Net approach. Software Impacts, 15, 100525. https://doi.org/10.1016/j.simpa.2023.100525
[13] Meshram, V. A., Patil, K., & Ramteke, S. D. (2021). MNet: A framework to reduce fruit image misclassification. Ingénierie des Systèmes d'Information, 26(2), 159–170. https://doi.org/10.18280/isi.260203
[14] Visvanathan, G., Patil, K., Suryawanshi, Y. *et al.* Mitigating urban heat island and enhancing indoor thermal comfort using terrace garden. *Sci Rep* **14**, 9697 (2024). https://doi.org/10.1038/s41598-024-60546-0
[15] Shilpa Sonawani & Kailas Patil & Prabhu Natarajan, 2023. "Biomedical signal processing for health monitoring applications: a review," International Journal of Applied Systemic Studies, Inderscience Enterprises Ltd, vol. 10(1), pages 44-69.
[16] Meshram, V., Patil, K., Hanchate, D., & Ramkteke, S. D. (2021). Machine learning in agriculture domain: A state-of-art survey. Artificial Intelligence in the Life Sciences, 1, 100010. https://doi.org/10.1016/j.ailsci.2021.100010
[17] Meshram, V. V., Patil, K. R., Meshram, V. A., & Bhatlawande, S. (2022). SmartMedBox: A smart medicine box for visually impaired people using IoT and computer vision techniques. *Revue d'Intelligence Artificielle*, *36*(5), 681.