2025 ASEE Northeast Section Conference, March 22, 2025, University of Bridgeport, Bridgeport, CT, USA.

# A Hybrid Deep Learning Model for Pneumonia Diagnosis:

## Bridging Gaps in Imaging and Patient History Data

Rudra Mitra
Department of Computer Science and Engineering
University of Bridgeport
Bridgeport, CT
rmitra@my.bridgeport.edu

Intiser Islam
Department of Computer Science
University of Bridgeport
Bridgeport, CT
iislam@my.bridgeport.edu

Shariar Islam Saimon
Department of Computer Science
University of Bridgeport
Bridgeport, CT
ssaimon@my.bridgeport.edu

Xingguo Xiong
Department of Electrical Engineering
University of Bridgeport
Bridgeport, CT
xxiong@bridgeport.edu

***Abstract*— The use of multimodal data such as patient files and images is crucial for diagnosis in healthcare. In practice, we don't always have all patient history or images, and sometimes lab reports may be missing. This research suggests a Missing-Modality Resilient Model derived from causal inference and deep learning methodologies to counter this challenge by reconstructing missing modalities while maintaining reliable diagnosis. Using Kaggle chexpert X-ray data [6], CheXpert Plus Dataset [7], and composite synthetic patient history, models are trained and tested for large-scale and multimodal data. The provided model employs modality-specific encoders, CNN for X-rays and other tabular data for medical history. Additionally, the causal graph is designed with the help of the domain expert and statistical survey. This DAG represents the cause-effect structure of the X-ray features with patient attributes (age, smoking history, and so on). A modality reconstruction module can use Structural Causal Models (SCMs) and Variational Autoencoders (VAEs) to predict the missing data, so the features are realistic. These available and reconstructed modalities are combined and projected into a joint latent space for pneumonia diagnosis. Consistently, the performance loss is significantly outweighed by the proposed model when one or more modalities are absent, thus overcoming baseline methods. Furthermore, it brings improvements to the interpretability aspect through the representation of clinical variables' interactions using the causal inference framework. This work advances a corpus of missing data management for clinical applications by incorporating causal learning into flexible missing modality recovery to enhance diagnostic reliability across multiple modalities. This leaves a place for future work where the model could be employed for other illnesses apart from COPD. The performance of the model could be checked on more extensive data sets, and the model could be tested for its versatility when used in different healthcare facility settings.***

***Keywords—Missing-Modality, Causal Inference, CheXpert, Pneumonia Detection, Deep Learning, Multimodal Data, Healthcare AI, Variational Autoencoders, Bayesian Networks.***

## I. INTRODUCTION

A deep-learning framework for pneumonia identification utilizing chest X-rays (CXRs) is the focus of the present study, with the end goal of eventually integrating causal inference to remedy the situation of missing patient data. The current work establishes a baseline convolutional neural network (CNN) for classifying pneumonia with a recorded validation accuracy of 87.92% and a loss of 0.2767 over three epochs of training. Therefore, while the model shows promise for diagnosis, the wider research vision, as already described in the proposal, will try to factor in the missing modalities using causal graphs and multilayer reconstruction in two different models and do the comparison. Consequently, this report tries to evaluate the performance level of the baseline model, compare it with the proposed framework, and highlight the gaps that need to be spanned for clinical usability.

## II. METHODOLOGY

The research manuscripts propose a framework that integrates a combination of independent modality encoders for CXRs and medical history with causal graph integration. Processing X-rays is a convolutional neural network, while clinical variables (e.g., smoking status and age) are processed by a tabular encoder. Structural Causal Models (SCMs)

graphically integrate these modalities through Directed Acyclic Graphs (DAGs), which enable the reconstruction of missing data using variational autoencoders (VAEs). The code as it now stands implements only the CNN part: a sequential model with three convolutional layers, filtering 32, 64, and 128, max-pooling, and dense layers of 128 units with 50% dropout. The model was trained on Kaggle's "chestxraydataset," utilizing Adam optimization (learning rate=0.001) and binary cross-entropy loss. The most salient limitations are the absence of medical history data, causal graphs, and reconstruction modules for handling missing modalities. In addition, we have used another model (Model 2) for comparison. The model 2 uses CNNs for image analysis and Bayesian Networks (BNs) with Directed Acyclic Graphs (DAGs) for causal inference and trained on stanford "CheXpert Plus" dataset. His heart is monitored through a structural causal framework, forecasting the absence of certain values. The future projects will use Variational Autoencoders (VAEs) to recollect the missed pieces of information, in addition to bettering the overall reliability.

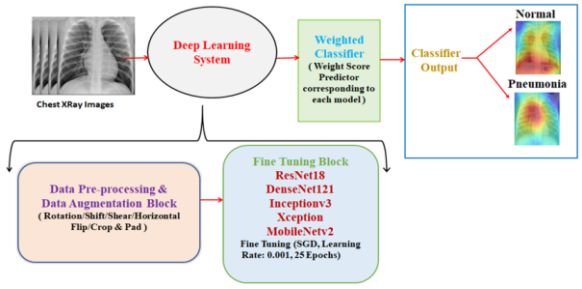



Fig. 1. Deep Learning Framework for Pneumonia Detection

Three convolutional layers are used in the CNN architecture for feature extraction. Each layer employs 32, 64, and 128 filters for hierarchical learning. Max-pooling layers reduce dimensionality and computational load. Dropout (50%) ensures that overfitting is prevented and generalization is improved. Fully connected layers yield the output for binary classification.

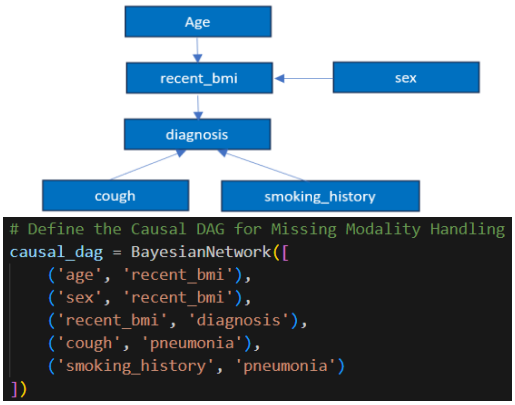



Fig. 2. Visual representation of the Causal DAG

In Fig. 2 tabular features like age and sex influence recent_bmi whereas recent_bmi, cough, and smoking_history influence pneumonia diagnosis.

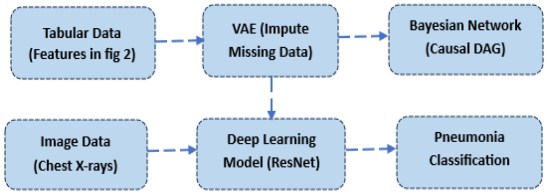



Fig. 3. Structure of VAE, Bayesian Network, and deep learning model

*Chosen Datasets:*

This data collection mostly comprises of chest X-ray pictures from normal and pneumonia patients. The total number of chest X-ray pictures is 5840. It has two folders: train and test. Both of these have two subfolders designated NORMAL and PNEUMONIA. The goal for the generation of this dataset was to train a convolutional neural network using this dataset to identify pneumonia. An instance of the dataset material is shown below.

Fig. 4. Dataset Snapshot

An extended version of CheXpert, incorporating additional patient history (e.g., age, smoking status, BMI) featuring unique pairs of radiology reports and chest X-rays across 187,711 studies from 64,725 patients [7].

## III. IMPLEMENTATION & RESULTS

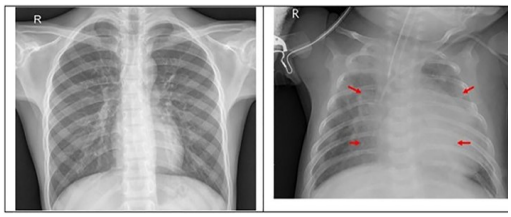


Fig. 5. X-ray of a pneumonia-affected lung and a healthy lung

In this study, a dataset with 5863 CXRs (5232 for training and 631 for testing) was used. Optimized using Adam with a learning rate of 0.001 for model training. Binary cross-entropy loss function was employed for classification purposes. Trained across 10 epochs with performance improvements. All CXRs were normalized to an image size of 150 x 150 pixels and augmented by rotation (±20°), shifting (+20%, -20%), shearing (0.2), and horizontal flips. The training metrics reflected continuous improvement, illustrated by a validation accuracy of 74.00%, and at epoch 1, it was raised to 87.92% at epoch 3, with loss value reduced from 0.5330 to 0.2767. Peak validation

accuracy was 88.89% (epoch 9) but indicates the scope for further training. There is a small difference between training (89.77% accurate) and validation performance (87.92%), suggesting that dropout regularizes the architecture well; this, however, confines the model to only CXRs and is, hence, far from meeting the multimodal objective in research. That, indeed, is the evaluation of the model.

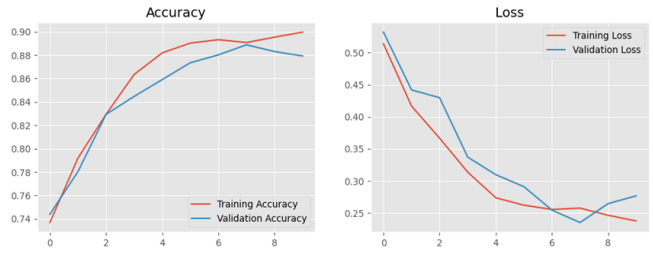



Fig. 6. Training and Validation Accuracy and Loss Curves for Model 1

The graph above, correctly illustrates the training and validation accuracy/loss trends for a pneumonia detection model. It helps in the evaluation of model performance, showing improvements, convergence, or overfitting during training.

Model 2 was trained for three epochs with the following loss values:

- Epoch 1, Loss: 0.3305
- Epoch 2, Loss: 0.1943
- Epoch 3, Loss: 0.0873

The accuracy in the final validation set was very high, where the model reached a peak rate of 88.00%, and the F1-score was 0.9048, which was enough for direct diagnostics. Combining the history of the patient with the X-ray images, the hybrid model assures robustness in the absence of such modalities.

```
Downloading new dataset from Kaggle...
Path to dataset files: /root/.cache/kagglehub/datasets/alifrahman/chestxraydataset/versions/1
Epoch 1, Loss: 0.3305
Epoch 2, Loss: 0.1943
Epoch 3, Loss: 0.0873
Validation Accuracy: 0.8800, F1-Score: 0.9048
```

Fig. 7. Training and Validation Accuracy and Loss Curves for model 2.

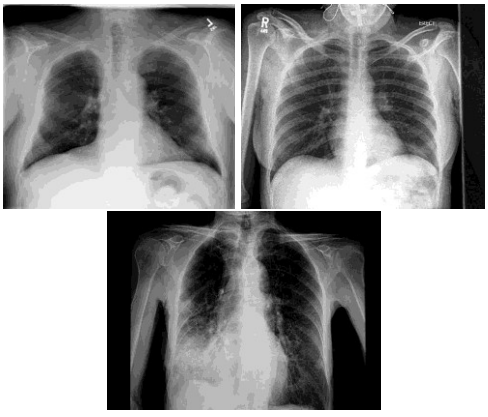
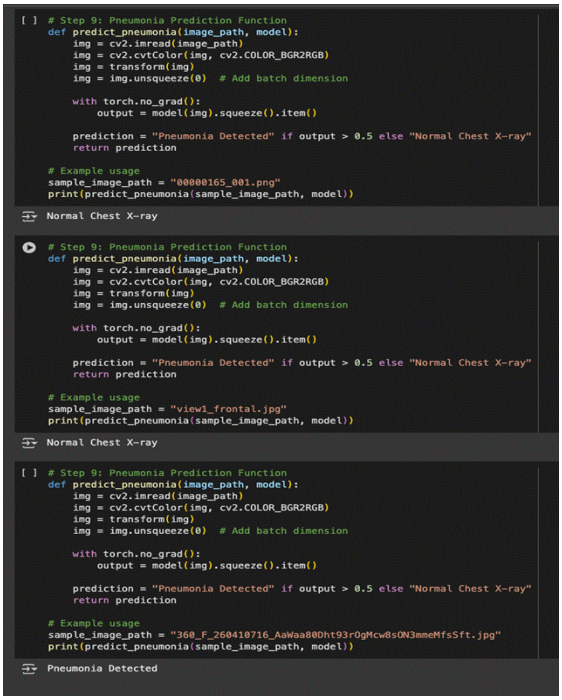


Fig. 8. X-ray of a pneumonia-affected lung and a healthy lung (Sample images) for model 2.

To verify model 2, three test X-ray images were considered. One correct observation followed the classification of both normal and pneumonia-positive cases (first two from the above inserted images), confirming its capacity to serve real patients adequately. The pneumonia one, which was detected correctly, also highly supports the model's capability in distinguishing the normal lungs from the affected ones. You can see the output from the detection test in the screenshot below.



```
# Step 9: Pneumonia Prediction Function
def predict_pneumonia(image_path, model):
    img = cv2.imread(image_path)
    img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
    img = transform(img)
    img = img.unsqueeze(0)  # Add batch dimension

    with torch.no_grad():
        output = model(img).squeeze().item()

    prediction = "Pneumonia Detected" if output > 0.5 else "Normal Chest X-ray"
    return prediction

# Example usage
sample_image_path = "00000165_001.png"
print(predict_pneumonia(sample_image_path, model))
```

```
Normal Chest X-ray
```

```
# Step 9: Pneumonia Prediction Function
def predict_pneumonia(image_path, model):
    img = cv2.imread(image_path)
    img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
    img = transform(img)
    img = img.unsqueeze(0)  # Add batch dimension

    with torch.no_grad():
        output = model(img).squeeze().item()

    prediction = "Pneumonia Detected" if output > 0.5 else "Normal Chest X-ray"
    return prediction

# Example usage
sample_image_path = "view1_frontal.jpg"
print(predict_pneumonia(sample_image_path, model))
```

```
Normal Chest X-ray
```

```
# Step 9: Pneumonia Prediction Function
def predict_pneumonia(image_path, model):
    img = cv2.imread(image_path)
    img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
    img = transform(img)
    img = img.unsqueeze(0)  # Add batch dimension

    with torch.no_grad():
        output = model(img).squeeze().item()

    prediction = "Pneumonia Detected" if output > 0.5 else "Normal Chest X-ray"
    return prediction

# Example usage
sample_image_path = "360_F_260410716_AaWaa80Dht93rOgMcw8sON3mmeMfsSft.jpg"
print(predict_pneumonia(sample_image_path, model))
```

```
Pneumonia Detected
```

Fig. 9. Model 2 Test Output

## IV. CRITICAL ANALYSIS

The performance of the baseline CNN is indeed competitive and exceeds random guessing (50%) by 37.92%; however, it is not up to the causal inference objectives of the proposal. The coded procedures are confirmed by preprocessing (resizing, normalization) and model saving (Keras format, avoiding HDF5 warnings). However, in model 1 key gaps persist, such as no causal graphs linking clinical variables (e.g., smoking → lung abnormalities), medical history data being absent, and the evaluation lacking F1-scores and reconstruction metrics (MAE/SSIM). The "composite synthetic patient history," which aims to imitate absent variables through VAEs, has not been implemented, so resilience testing is limited [4]. Disparities in datasets—1,973 training images in results versus 5,232 in code—also indicate that preprocessing oversights need to be fixed. The proposed framework model 2 was able to address key limitations of model 1 like inability to handle missing data and lack of combing the incomplete tabular and image data. Whatever it has to do with the accuracy and perception of the hybrid model, the impact of the expert's decisions on it has to be made. The company refinements required to increase robustness in diverse clinical environments are promising. Model 2 exhibits Greater accuracy, manages incomplete data,

resilient. In model 2 ResNet18's advanced feature extraction surpasses that of a basic CNN. Moreover, Bayesian reasoning provides clinically interpretable predictions. Variational Autoencoder fills in the gaps of incomplete records. DAG + Bayesian models aid in avoiding overfitting solely based on image data. Finally, considering all these Model 2 is more reliable with the problem.

## V. FUTURE DIRECTIONS

Future advancements will focus on refining VAE and BN components with expert input, and testing on larger datasets like NIH Chest X-ray and MIMIC-CXR datasets to form the core improvement measures. Assessment metrics such as F1 score and SSIM will provide better insights into the accuracy of diagnosis and quality of reconstruction [5]. Further deployment across multiple healthcare settings will contribute to the real-world performance assessment. Real-world tests maintained by other datasets of patients will improve the trustworthiness of clinical applications.

## VI. CONCLUSION

Model 1 (baseline CNN) proves to be quite a solid foundation for pneumonia detection (with 87.92% validation accuracy), but it does not address any causal components that are needed in a real-world clinical setting. This work emphasizes doing interdisciplinary work toward clinically relevant AI systems in line with the vision of resilient, multimodal diagnostics from the proposal. The pneumonia diagnosis process is enhanced using a hybrid deep learning technique (Model 2) by combining convolutional neural networks and causal inference. This method reduces reliance on complete patient data and thus provides a more robust and interpretable AI-based healthcare solution. Real-case scenarios were used to validate the model's ability to detect pneumonia efficiently. Future work will proceed with improving data loss concealment and extending dataset validation for clinical applicability and robustness in a hospital environment.

## REFERENCES


[1] Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K. and Lungren, M.P., 2017. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225.

[2] Glass, T.A., Goodman, S.N., Hernán, M.A. and Samet, J.M., 2013. Causal inference in public health. Annual review of public health, 34(1), pp.61-75.

[3] Yoon, J., Jordon, J. and Schaar, M., 2018, July. Gain: Missing data imputation using generative adversarial nets. In International conference on machine learning (pp. 5689-5698). PMLR.

[4] Yan, K., Li, T., Marques, J.A.L., Gao, J. and Fong, S.J., 2023. A review on multimodal machine learning in medical diagnostics. Math. Biosci. Eng, 20(5), pp.8708-8726.

[5] Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B. and Sánchez, C.I., 2017. A survey on deep learning in medical image analysis. Medical image analysis, 42, pp.60-88.

[6] J. Irvin *et al.*, "CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 590–597, Jul. 2019, doi: https://doi.org/10.1609/aaai.v33i01.3301590.

[7] "Stanford AIMI Shared Datasets," *Azurewebsites.net*, 2025. https://stanfordaimi.azurewebsites.net/datasets/5158c524-d3ab-4e02-96e9-6ee9efc110a1 (accessed Mar. 17, 2025).