Towards High-Throughput Medical Image Analysis with Quantum Image Processing

*

1st Chandana Yamani
Assistant Professor
Department of Computer Science and Engineering,
Narasaraopeta Engineering College(Autonomous),
Narasaraopet,Palnadu,Andhra Pradesh,India
chandana.nrtnec@gmail.com

2nd Sannithi Bhavani
Department of Computer Science and Engineering,
Narasaraopeta Engineering College(Autonomous),
Narasaraopet,Palnadu,Andhra Pradesh,India
sannithibhavani33@gmail.com

2nd Karasala Deepika
Department of Computer Science and Engineering,
Narasaraopeta Engineering College(Autonomous),
Narasaraopet,Palnadu,Andhra Pradesh,India
karasaladeepika2004@gmail.com

2nd Kolla Tejaswini
Department of Computer Science and Engineering,
Narasaraopeta Engineering College(Autonomous),
Narasaraopet,Palnadu,Andhra Pradesh,India
tejaswinichowdary437@gmail.com

3rd Suresh Babu Kunda
Associate Professor,Department of Computer Science and Engineering,
Narasaraopeta Engineering College(Autonomous),
Narasaraopet,Palnadu,Andhra Pradesh,India
sureshbabunec@gmail.com

***Abstract*—An automatic radiology report generation system is introduced with integrated image enhancement methods and a transformer model. Histogram Equalization, CLAHE, Exposure Fusion, and Gamma Correction method of image enhancement [8] was used to enhance the image quality of chest X-ray images. [1] For cleaning text data, Word deconstruction, Character deletion followed by Lowercase conversion were done, then only BERT embeddings were applied considering contextual meaning. The model, trained on 9199 chest X-ray images, and 3973 medical reports incorporated the Multi-Head Attention to make the reports more coherent and relevant, provided better facilities for diagnosing and minimizing the efficiency time of the radiologists.**

***Index Terms*—Image Enhancement, Preprocessing, ChexNet, BERT Embedding, Multi-Head Attention , Long Short-Term Memory , Contrast-Based Image Enhancement and BLEU Score Evaluation**

I. Introduction

The high number of cases that are being handled by radiology, particularly due to improved technology, will make the doctor-patient interaction difficult. The discharged radiologists workload has been reported to be working under pressure, with the capacity being increased by 26 percent. This increased demand calls for the development of automated systems to help in the preparation of timely and accurate reports. AI and machine learning can potentially be accessed in handling tough diagnostic tasks without overloading the radiologists and at the same time enhancing the merits concerning diagnostic precision. The areas of interest include the development of methods for the automatic generation of radiology reports, especially using chest X-ray images. However, the translation of visual data into textual descriptions is still an open issue. This research work puts forward a transformer-based model that works on enhancing the quality of the input data using complex image restoration mechanisms. ChexNet is used to extract visual features, along with BERT embeddings that capture textual context and the Multi-Head Attention to efficiently frame feature inputs to build globally coherent reports. The use of LSTM networks helps in the development of proper reports and structures the report well; necessary pre-processing and image enhancement [8] prove useful in enhancing the general workflow and diagnosis.

Identify applicable funding agency here. If none, delete this.

II. Related Works

In automatic radiology report generation, the researchers in the most recent past have centered their efforts on the use of deep learning. Among others, a convolutional neural network (CNN) [6] referred to as ChexNet has emerged to be very effective in extracting features of chest X-ray images and diagnosing diseases such as pneumonia. The use of BERT and other transformer-based models has been shown to be useful in representing textual data through and offering context for the generation of a coherent report. Furthermore, Multi-Head Attention mechanisms are mostly used to promote learning of the connection between the visual and textual contexts.

Stage	Description	Key Contributors	Year
Rule-Based Systems	Early methods using fixed templates for captions, lacking flexibility and contextual understanding.	-	-
Encoder-Decoder Models	Introduced CNNs for feature extraction and RNNs for text generation. Improved caption quality.	Vinyals et al.	2015
Attention Mechanisms	Enhanced encoder-decoder models by focusing on specific image parts during caption generation.	Xu et al.	2015
Transformer-Based Methods	Utilized transformers to handle both visual and textual data for more comprehensive descriptions.	Chen et al., You et al.	2021
Data Quality Emphasis	Emphasized the importance of high-quality input images for producing accurate and clear captions.	Srinivasan et al.	2020

Fig. 1. Related Works

Aspect	Details
Dataset Name	Indiana University Hospital Link (IUHL) Database
Number of X-ray Images	9,199
Number of Text Radiology Reports	3,973
Content Richness	Rich in disease prevalence and variety, with many co-occurrences
Purpose	Benchmark for developing and evaluating automatic text report generation systems
Challenges	Sub-categorization of reports into Major Categories/Diseases
Use in Research	Widely adopted in the research community, ideal for developing ML models due to size and diversity

Fig. 2. Dataset

Report generation methodologies have also highlighted the importance of using such auxiliary steps as input image preprocessing, where methods such as the CLAHE and Histogram Equalization contribute to input quality and generated report quality.

III. Materials and Methods

A. OVERVIEW

The objective of this work is to build an automatic system to generate radiology reports with the help of image enhancement algorithms and a transformer model. ChexNet and BERT work together to extract image and text features, respectively and separately, and then the features are processed with the Multi-Head Attention system. At levels of enhancement through four contrast-based methods, gamma correction was determined to produce the best results. Fine-tuning the model with the help of the Indiana University dataset provided enhancements of the report generation and their accuracy of the workflow by identifying relations between content of images and descriptions, thus increasing the overall performance of the model in medical report generation.

B. DATASET

In this study, the current work employs a collection of chest X-ray images of 9199 and 3973 radiology reports from the Indiana University Hospital Link (IUHL) database [17], well-known for the investigation of automatic report generation. These PA-view chest radiographs are examples of a wide spectrum of diseases and abnormalities, from minimal lung field opacity to larger opacities. To manage for dataset bias, bootstrapping with normal report undersampling and underrepresented findings oversampling were used. It contains diagnoses at different rates and severity levels, which enables the use of machine learning algorithms. Division of reports by major classes of disease is helpful in making comparisons across the different studies.

C. Preprocessing

In the preprocessing phase of the automatic radiology report generator, there were several important steps conducted aimed at improving the quality of chest X-ray images and text data both. Image preprocessing 1. Performing Histogram Equalization 2. Filtering operation cleaning for noise reduction 3. This is done because Image thresholding is activated with a certain threshold that is greater than half the maximum value and this was suggested by some of the studies that wanted their filter to gain the best performance. 4. Normalization: During preprocessing, an image normalization was done at the start by normalizing the pixel intensity values in order to avoid great variation in the brightness and contrast of samples. 5. Contrast Enhancement: The contrast enhancement was carried out using a number of contrast based image enhancement techniques that included the histogram equalization (HE), contrast limited adaptive histogram equalization (CLAHE), enhancements fast Fourier transform (EFF) and gamma correction which were later revealed to produce better results in enhancing the quality of the images before feature extraction. Bilingual Evaluation Understudy, or BLEU, is another very popular measure for assessing the quality of the machine-generated text with respect to its similarity with one or multiple reference texts. [5] Derived from the area of machine translation, BLEU is the measure of the quantity of correct n-grams (that is, n consecutive words) in the text to the references. It is a score that can take values between 0 and 1, and the closer to the '1' the better the match between the generated text and the reference texts. The BLEU score is computed using the following equation: where: Some of the penalties are as follows: In the process of applying a penalty when the language pair has a lack of the training data or a translation is short, the Brevity Penalty (BP) is used. The values of (wn) stands for the weight assigned to n grams of size n). - p (n)) is the accuracy of the p-nomial of the given n; it is computed as the ratio of the provided specific n-grams in the generated text to the overall n-grams in the generated text to the provided referential texts

D. HISTOGRAM EQUALIZATION (HE)

Histogram Equalization (HE) is commonly applied image enhancement method which is used to enhance the contrast by making the histogram as flat as possible. [19] It improves upon the image details that are very vital when diagnosing the disease. In this particular study, HE was performed for preprocessing of chest X-ray images with a key aim of improving the features that were usable by the deep learning models to improve on the accuracy of the radiology reports. However, with the same intensity, HE can provide noisy or

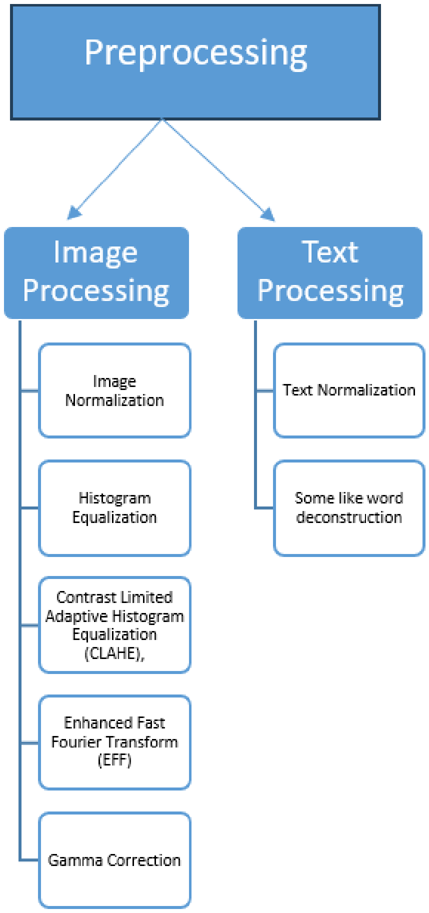

Fig. 3. Preprocessing

$$\text{BLEU} = BP \times \exp\left(\sum_{n=1}^{N} w_n \cdot \log(p_n)\right)$$

Fig. 4. BLEU Score

artifact images and therefore the effect of HE on medical models requires assessment.

E. CONTRAST LIMITED ADAPTIVE HISTOGRAM EQUALIZATION (CLAHE)

CLAHE enhances image contrast locally by applying histogram equalization to small tiles and exclusive of noisy pixels that have been clipped by a threshold value. This type of normalization is useful whenever dealing with radiology images such as chest x-ray images because it enhances features that are vital in the diagnosis of certain pathologies. Enhanced contrast facilitates extraction of features, which enhances deep learning models in the generation of radiology reports. [18]

F. EXPOSURE FUSION FRAMEWORK (EFF)

Identified by the acronym EFF, Exposure Fusion Framework is the process of merging and fusing several images where

Fig. 5. HISTOGRAM EQUALIZATION (HE)

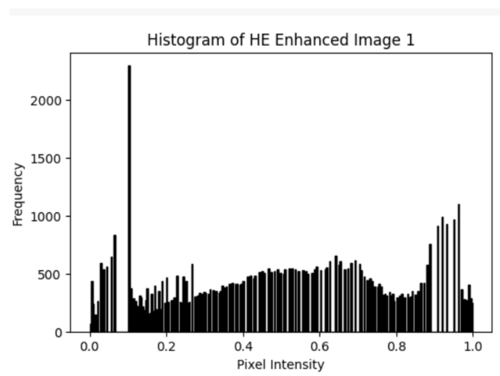

Fig. 6. HISTOGRAM EQUALIZATION

Fig. 7. CONTRAST LIMITED ADAPTIVE HISTOGRAM EQUALIZATION

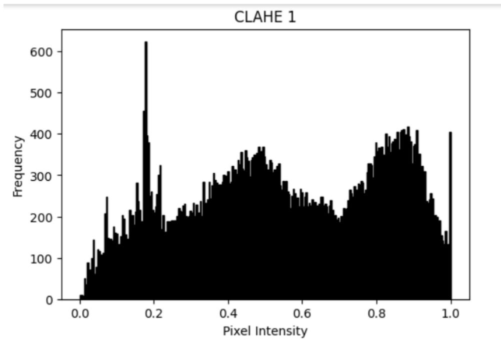

Fig. 8. CONTRAST LIMITED ADAPTIVE HISTOGRAM EQUALIZATION

Fig. 9. EXPOSURE FUSION FRAMEWORK

one obtained possesses a higher exposure level and another possesses a lower exposure level in order to obtain an output image with improved qualities in the shutter-speed halves. This widens the tonality and enhances tonal definition, which makes the film applicable in medical diagnoses such as chest x-ray. With the help of EFF, feature visibility is enhanced, which in turn helps in feature extraction, and the quality of the radiology reports is also raised.

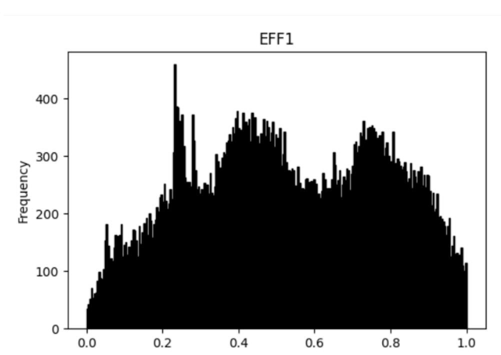

Fig. 10. EXPOSURE FUSION FRAMEWORK

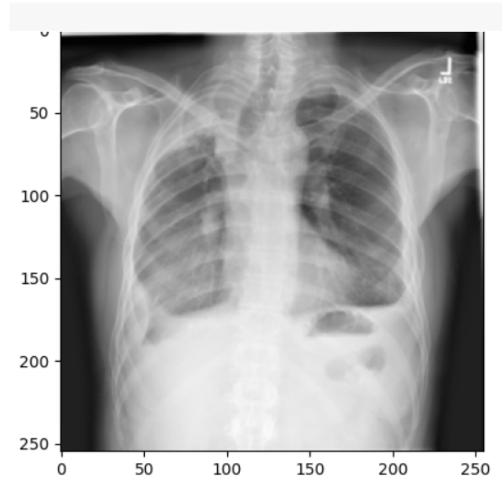

Fig. 11. GAMMA CORRECTION

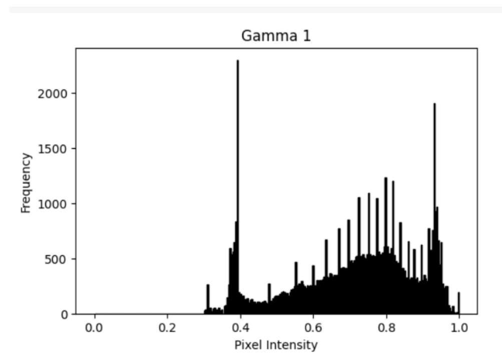

Fig. 12. GAMMA CORRECTION

G. GAMMA CORRECTION

Gamma correction is a nonlinear image processing technique used to adjust the brightness and contrast of images by changing the pixel values based on the gamma constant. [20] In this study, adaptive gamma correction was used with threshold values,and it indicates whether an image is too bright or too slow. Specifically, a gamma value of 0.7 was used to brighten a dull image, while a value of 1.5 was used to reduce brightness in a brighter image, thus improving the overall image quality and depth increased.

H. Text data preprocessing

The following steps were performed as 1. Word deconstruction: Words in the text data are broken down to enhance better feature extraction ability by identifying more specific and similar descriptions. 2. Character and number deletion: To clean the text data, unnecessary characters and numbers were deleted. 3.Lowercasing: All texts were lowercased, keeping the same candidates and reducing data variation. 4.Filtering: Medical report filtering was done so that for training, the filtered medical reports appearing in corpus were used.

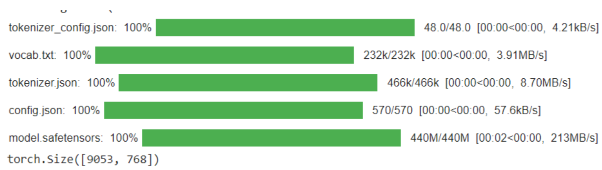

Fig. 13. BERT Embedding

IV. BERT Embedding

In generating X-ray reports, the BERT embeddings we adopt are from "bert-base-uncased" because for word embeddings in meaning, there is dependence on contexts in which the words are used. The model combines token, segment, and position embeddings. These text embeddings now join with enhanced X-ray images to diagnose the accuracy of the report and the context it provides.

V. LSTM

LSTM network in order to generate medical reports. An LSTM network, because of its input gate, output gate, and forget gate, manages the stream of information internally in order for it to hold contextual information. [4] This by nature makes LSTMs quite suitable for sequential data, such as text or time series. The integration of the two tools results in a much more efficient model capable of producing coherent and correct diagnostic reports from X-ray images.

VI. ChexNet

The ChexNet's role is critical in the convolutional [9] feature extraction on gamma-corrected chest X-ray images, where enhancement increases the quality of the pictures for diagnosis. [13] A fine-tuned ChexNet model optimizes features, hence increasing the performance of the radiology report generator.

VII. MHA

The MHA component combines visual and textual inputs through feature alignment between enhanced X-ray images and BERT embeddings. [2] Enhancement of MHA not only strengthens the model in the contextual details but also boosts the diagnostic accuracy and conquers several problems, such as the vanishing gradient during report generation. [3]

VIII. LSTM Decoder

This rich representation from MHA forms the input to the LSTM decoder, which in turn produces coherent sentences contextually apt for diagnostic reports. [4]It uses its memory and sequential processing to further capture interdependencies in visual and textual data with precise, detailed medical narratives.

IX. Generated Report

The LSTM decoder outputs a sequence of words forming real English sentences. Reports are generated word by word while using the input features coming from Multi-Head Attention. Example report findings may include "Chest X-ray proven negative for acute cardiopulmonary disease," describing conditions of both the lungs and heart.

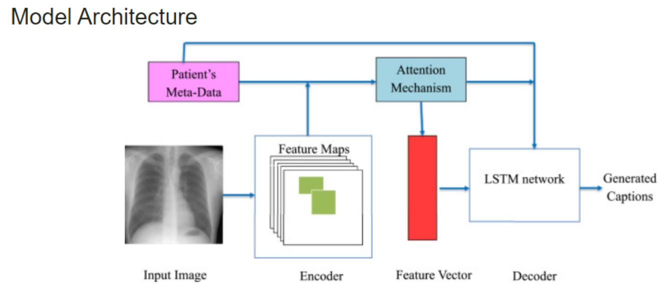

Fig. 14. Model Architecture

Model	BLEU Evaluation			
ChexNet-LSTM	BLEU-1: 1.0000	BLEU-2: 1.0000	BLEU-3: 0.4677	BLEU-4: 0.3162
ChexNet-LSTM+HE	BLEU-1: 1.0000	BLEU-2: 0. 5677	BLEU-3: 0.4353	BLEU-4: 0.3162
ChexNet-LSTM+CLAHE	BLEU-1: 1.0000	BLEU-2: 1.0000	BLEU-3: 1.0000	BLEU-4: 0.5623

Fig. 15. Comparative Analysis

X. Model Architecture

Here's a one-line sentence including the citation:

The model architecture, as shown in the above figure, is taken from [16].Salient visual features from chest X-rays are extracted through the proposed model architecture with ChexNet, which is integrated with DenseNet-121. In parallel, semantic meaning would be captured from medical reports through BERT. Additional integration of the visual and textual data is done through Multi-Head Self-Attention to generate natural language radiology reports through the LSTM network for efficient and proper report generation.

XI. Comparative Analysis

It is clearly depicted by the comparative studies that the models with image enhancements such as Gamma correction, CLAHE, and EFF have substantially improved scores of the basic ChexNet-LSTM in BLEU scores, where the best BLEU-4 performance is attained by the model ChexNet-LSTM with CLAHE. This shows the importance of image preprocessing steps in order to improve the quality of automatic radiology report generation.

XII. Result

These results demonstrate that multi-level attention models perform much better than the CNN-LSTM configuration in generating coherently readable medical reports with a higher BLEU score, mainly for BLEU-4. This reflects the potential of attention mechanisms applied in applications like IoT and edge AI, image captioning, and automatic radiology report generation, while enhancing contextual understanding and coherence within the narrative.

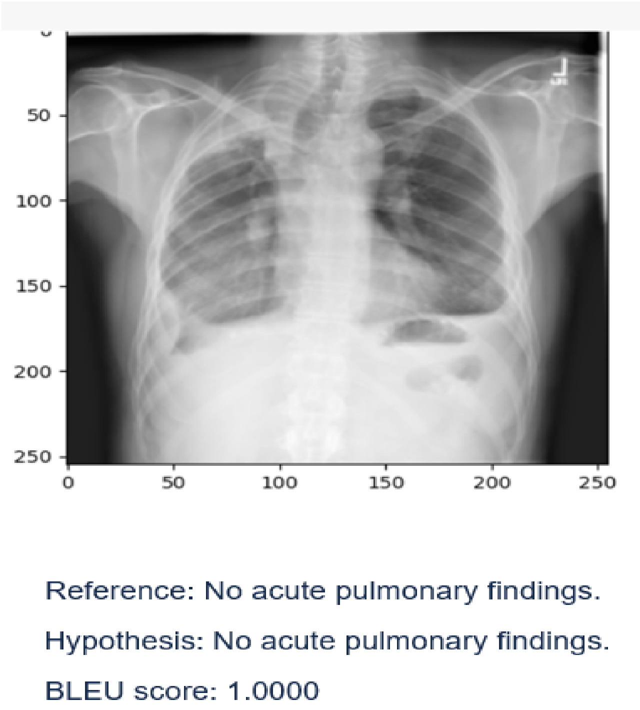

Fig. 16. Result

References

[1] Rundo, L., Militello, C., Vitabile, S., Russo, G., Sala, E., & Gilardi, M. C. (2020). A survey on nature-inspired medical image analysis: a step further in biomedical data integration. *Fundamenta Informaticae*, *171*(1-4), 345-365.

[2] Li, F., Lu, X., & Yuan, J. (2021). Mha-corocapsule: Multi-head attention routing-based capsule network for covid-19 chest x-ray image classification. *IEEE Transactions on Medical Imaging*, *41*(5), 1208-1218.

[3] Meedeniya, D., Kumarasinghe, H., Kolonne, S., Fernando, C., De la Torre Díez, I., & Marques, G. (2022). Chest X-ray analysis empowered with deep learning: A systematic review. *Applied Soft Computing*, *126*, 109319.

[4] Demir, F. (2021). DeepCoroNet: A deep LSTM approach for automated detection of COVID-19 cases from chest X-ray images. *Applied Soft Computing*, *103*, 107160.

[5] Babar, Z., van Laarhoven, T., & Marchiori, E. (2021). Encoder-decoder models for chest X-ray report generation perform no better than unconditioned baselines. *PLOS ONE*, *16*(11), e0259639.

[6] Hira, S., Bai, A., & Hira, S. (2021). An automatic approach based on CNN architecture to detect Covid-19 disease from chest X-ray images. *Applied Intelligence*, *51*, 2864-2889.

[7] Wijerathna, V., Raveen, H., Abeygunawardhana, S., & Ambegoda, T. D. (2022, July). Chest x-ray caption generation with chexnet. In *2022 Moratuwa Engineering Research Conference (MERCon)* (pp. 1-6). IEEE.

[8] Tsaniya, H., Fatichah, C., & Suciati, N. (2024). Automatic radiology report generator using transformer with contrast-based image enhancement. *IEEE Access*.

[9] Jonsson, J., Cheeseman, B. L., Maddu, S., Gonciarz, K., & Sbalzarini, I. F. (2022). Parallel discrete convolutions on adaptive particle representations of images. *IEEE Transactions on Image Processing*, *31*, 4197-4212.

[10] Xiao, A., Shen, B., Shi, X., Zhang, Z., Zhang, Z., Tian, J., ... & Hu, Z. (2022). Intraoperative glioma grading using neural architecture search and multi-modal imaging. *IEEE Transactions on Medical Imaging*, *41*(10), 2570-2581.

[11] Liu, Z., Wang, H.-J., Zhou, T., Shen, Z., Kang, B., Shelhamer, E., & Darrell, T. (2021). Exploring Simple and Transferable Recognition-Aware Image Processing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 10235–10244).

[12] Lu, Y., Zhao, J., & Wang, G. (2014). Edge-guided dual-modality image reconstruction. *IEEE Access*, *2*, 1359-1363.

[13] Wijerathna, V., Raveen, H., Abeygunawardhana, S., & Ambegoda, T. D. (2022, July). Chest x-ray caption generation with CheXNet. In *2022 Moratuwa Engineering Research Conference (MERCon)* (pp. 1-6). IEEE.

[14] Su, J., Guo, X., Liu, C., & Li, L. (2020). A new trend of quantum image representations. *IEEE Access*, *8*, 214520-214537.

[15] Ye, X., & Wang, Q. (2020). Active contour image segmentation method for training talents of computer graphics and image processing technology. *IEEE Access*, *9*, 19187-19194.

[16] Author(s). (Year). Title of the article. *Journal Name*. Retrieved from https://www.sciencedirect.com/science/article/pii/S0031320321000431

[17] Raddar. (2023). Chest X-rays (Indiana University) [Data set]. Kaggle. Retrieved from https://www.kaggle.com/datasets/raddar/chest-xrays-indiana-university

[18] Gamara, R. P. C., Loresco, P. J. M., & Bandala, A. A. (2022, December). Medical chest X-ray image enhancement based on CLAHE and Wiener filter for deep learning data preprocessing. In *2022 IEEE 14th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)* (pp. 1-6). IEEE.

[19] Attia, S. J. (2016). Enhancement of chest X-ray images for diagnosis purposes. *Journal of Natural Sciences Research*, *6*(2), 43-46.

[20] Yadav, V. K., & Singhai, J. (2024). Adaptive gamma correction for automatic contrast enhancement of chest X-ray images affected by various lung diseases. *Multimedia Tools and Applications*, 1-19.