Automated Report Generation from X-Ray Images

MS (DS) Thesis Proposal

By

Muhammad Baseer Khan (232789)

Supervised by: Dr Aman Ullah Yaseen

Co-Supervised by: Dr Syed Muhammad Usman

Department of Creative Technologies, Faculty of Computing and AI Air University, Islamabad, Pakistan

Fall 2023

Abstract

Chest X-rays serve as a fundamental tool in clinical practice, aiding in the diagnosis and treatment of numerous health conditions, including respiratory diseases, cardiovascular disorders, and infections. However, manual interpretation of these images poses significant challenges, even for experienced radiologists, due to the complexity and variability of medical cases. With the increasing demand for radiological assessments and a shortage of skilled radiologists, there is a pressing need for automated solutions that can assist in image analysis and reporting. This research proposes an automated chest X-ray reporting framework leveraging advanced machine learning and deep learning methodologies to enhance diagnostic accuracy and efficiency. The framework utilizes a combination of transfer learning and transformer-based architectures to automatically generate comprehensive and contextually relevant reports. Feature extraction is performed using a pretrained convolutional neural network (CNN) model, while a transformer-based encoder-decoder model processes these features to generate descriptive medical reports. The integration of large language models (LLMs) further refines the generated text, ensuring high readability and clinical relevance. The proposed framework will be evaluated using publicly available chest X-ray datasets to assess its performance against state-of-the-art methods, using metrics such as accuracy, precision, and clinical consistency. The results of this study are expected to establish a new benchmark in automated radiology reporting, contributing significantly to the adoption of AI-driven solutions in medical diagnostics.

Keywords

Automated Chest X-ray Reporting, Machine Learning, Transfer Learning, Transformers, Deep Learning, Medical Image Analysis, Clinical Decision Support Systems.

I Introduction

Medical imaging plays a pivotal role in modern healthcare by enabling clinicians to visualize internal structures of the human body, thereby facilitating accurate diagnosis and effective treatment planning. Among the various imaging modalities, chest X-rays are one of the most frequently utilized, particularly in emergency and critical care settings [1]. Due to their rapid, non-invasive, and cost-effective nature, chest X-rays are indispensable for detecting a range of medical conditions such as pneumonia, tuberculosis, emphysema, fractured ribs, cardiomegaly, and various forms of cancers, including lung and breast cancer [2]. Additionally, chest X-rays have proven effective in identifying lung infections caused by COVID-19 [3], making them a crucial tool during the global pandemic.

Despite their widespread use, interpreting chest X-rays requires a high degree of expertise and remains a time-consuming process for radiologists [4]. This task is particularly challenging in regions where the ratio of radiologists to patients is low, resulting in increased workloads and delayed patient care. The situation is even more daunting for less experienced radiologists or those practicing in remote areas with limited access to specialized healthcare facilities [5]. The development of automated systems to assist in interpreting these images and generating comprehensive medical reports can significantly alleviate these challenges.

One promising solution is the application of advanced artificial intelligence (AI) techniques to automate image analysis and report generation. Automated image captioning, which involves generating a textual description for an image [6], has emerged as a critical research area in AI, combining computer vision and natural language processing. While humans can intuitively observe and describe the contents of an image, teaching a machine to perform this task, particularly for complex medical images [7], remains a formidable challenge. For chest X-rays, automated captioning systems must be capable of identifying subtle patterns and abnormalities, and then articulating these findings in clear, clinically relevant language.

In this research, I propose an advanced framework that leverages machine learning and deep learning methodologies to automate chest X-ray report generation [19]. By integrating transfer learning with a pretrained convolutional neural network (CNN) for feature extraction, and refining the report generation process using transformer-based architectures and large language models (LLMs) [20] [figure 4], this system aims to produce detailed and accurate medical reports that are on par with expert radiologists. This approach not only addresses the challenges of interpreting chest X-rays but also contributes to the broader adoption of AI-driven solutions in clinical practice, ultimately enhancing the quality and efficiency of healthcare delivery.

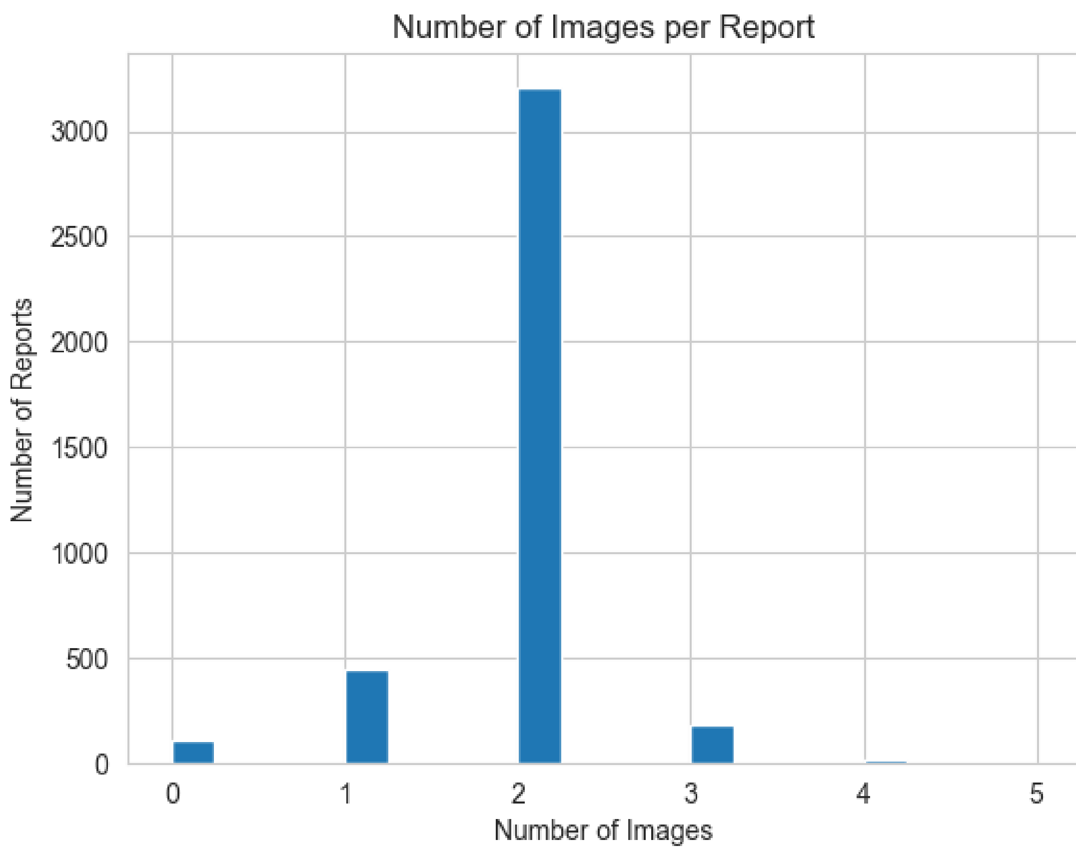


Figure A highlights the association between the number of images per report based on the analysis of the X-ray collection from Indiana University.

Analysing the association between chest X-ray images and their corresponding reports, Figure A illustrates the distribution of the number of images linked to each XML report. The histogram reveals the frequency of reports with varying numbers of associated images, providing insight into the dataset's structure. The x-axis represents the number of images per report, while the y-axis indicates the count of reports for each image category. This visualization is crucial for understanding the extent of image utilization across the dataset and ensuring that the model can adequately handle cases with different image-to-report ratios. A balanced distribution will aid in the robustness of the proposed automated report generation system, while any significant skewness in the data will highlight areas requiring special consideration during model training and evaluation.

Figure B Figure C

Figure B: Frontal chest X-ray from the Indiana University dataset, used as input for automated report generation [8].
Figure C: Lateral chest X-ray from the Indiana University dataset, utilized for generating diagnostic reports through AI [8]

Figures B and C provide visual representations of typical chest X-ray images used in clinical diagnosis. Figure B displays a standard frontal chest X-ray, capturing an anteroposterior view that highlights the structural details of the lungs, heart, and surrounding anatomical features. Meanwhile, Figure C presents a lateral chest X-ray, offering a complementary perspective that reveals additional information about the thoracic cavity, including the posterior aspects of the lungs and heart. Each image is uniquely linked to a corresponding XML report, which contains critical metadata and diagnostic attributes. This linkage between the images and their respective XML reports enables the extraction of detailed image attributes, such as image acquisition parameters, pathological findings, and diagnostic notes. Understanding these associations is fundamental to the automated report generation process, as it ensures that all relevant clinical information is accurately captured and integrated into the system for precise report creation.

II Motivation

The increasing reliance on medical imaging for the diagnosis and management of a wide range of health conditions, combined with a global shortage of radiologists, presents a significant challenge in delivering timely and accurate patient care. Chest X-rays, as one of the most commonly used imaging modalities, play a critical role in detecting and monitoring diseases such as pneumonia, tuberculosis, lung cancer, and other respiratory or cardiovascular disorders. However, interpreting these complex images requires a high level of expertise, making it a labor-intensive task for radiologists, who are already overwhelmed with growing workloads. This situation is exacerbated in resource-limited settings and rural areas, where access to skilled radiologists is limited, resulting in delayed diagnoses and compromised patient outcomes.

The need for an efficient solution that can assist radiologists in reading chest X-rays and generating reports is thus apparent. Automating the process of interpreting X-ray images and creating accurate, detailed medical reports can not only alleviate the burden on healthcare professionals but also standardize diagnostic outputs, reducing variability and enhancing clinical decision-making. By employing advanced machine learning and deep learning methodologies, the proposed automated system aims to provide a reliable tool that bridges the gap between the increasing demand for radiological assessments and the availability of skilled radiologists.

Furthermore, the adoption of intelligent, AI-driven systems in clinical practice has the potential to revolutionize healthcare delivery by optimizing radiology workflows, supporting early disease detection, and ultimately improving patient outcomes. The motivation for this research stems from the need to enhance the quality and efficiency of healthcare services through the development of an automated reporting system that can deliver high-quality, consistent results across diverse medical cases. This research not only addresses the immediate challenges faced by radiologists but also contributes to the broader goal of integrating AI in medical diagnostics, paving the way for more accessible and effective healthcare solutions worldwide.

III Literature Review

The literature on chest X-ray report generation showcases a diverse array of methodologies aimed at enhancing the accuracy and quality of automated reporting systems. A prominent trend involves the implementation of advanced preprocessing techniques, such as image normalization and augmentation, which significantly bolster model robustness by diversifying the training dataset. Many studies emphasize the importance of data cleaning to ensure high-quality inputs, setting the stage for effective analysis.

In the realm of feature extraction, Convolutional Neural Networks (CNNs) and their variants are frequently employed to capture intricate visual information from X-ray images. Innovations such as the integration of CNN with Recurrent Neural Networks (RNNs) [1] enhance the capacity for sequential data processing, facilitating a deeper understanding of the data's temporal aspects. Moreover, sophisticated approaches like multi-view image encoders demonstrate the efficacy of combining visual embeddings with textual summaries to create rich contextual representations pertinent to disease-related features.

The evolution of model architectures reveals a growing preference for encoder-decoder frameworks [1], which effectively transform extracted features into coherent reports. The adoption of Vision Transformers, alongside language models, marks a significant advancement in this field, leading to notable improvements across various performance metrics. Additionally, the exploration of attention mechanisms to enhance feature fusion has proven effective, allowing for the seamless integration of background information with extracted data.

Performance metrics within this body of research exhibit a wide range, with BLEU scores reflecting the ongoing challenge of achieving human-like report quality. Recent advancements in cross-modal multi-scale feature fusion have yielded impressive enhancements in BLEU scores, underscoring the continuous evolution of methodologies in automated report generation. Collectively, this body of work highlights a concerted effort to refine chest X-ray report generation, employing diverse preprocessing techniques, robust feature extraction methods, and innovative model architectures to meet the demands of clinical practice.

3.1 Feature Extraction

Feature extraction methods in chest X-ray report generation predominantly leverage Convolutional Neural Networks (CNNs) to capture complex visual information from the images. Variants like DenseNet121 are utilized to enhance representation capabilities. Some approaches integrate CNNs with Recurrent Neural Networks (RNNs) to effectively process sequential data, thereby enriching the interpretative context of the reports. Multi-view image encoders further enhance feature extraction by combining visual embeddings with textual summaries, creating contextualized representations that focus on disease-related features. This multifaceted approach enables a more comprehensive understanding of the data, facilitating improved report generation.

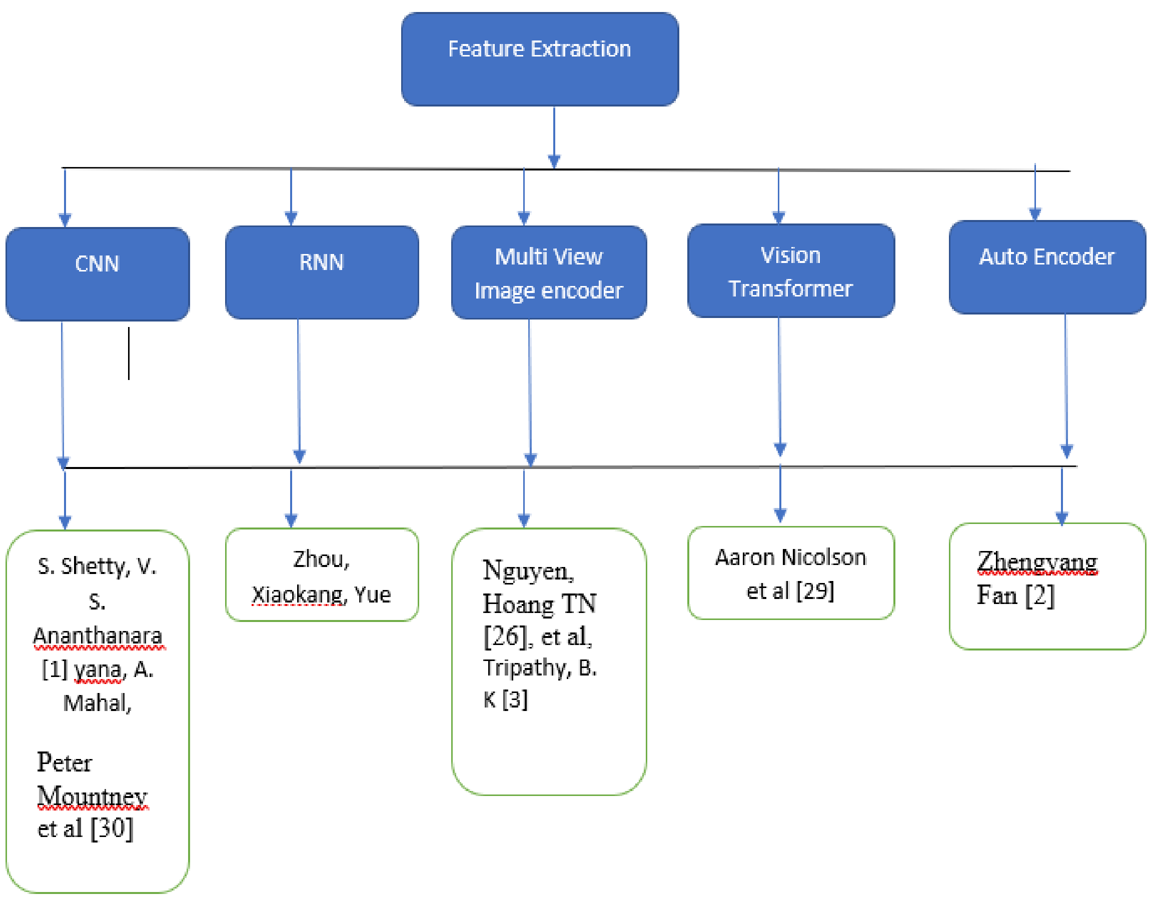


Figure D: Above image explains which feature extraction methods are used

IV Preprocessing

Preprocessing methods play a crucial role in preparing chest X-ray images and associated data for analysis. Techniques such as image normalization and augmentation are widely used to improve the quality and variability of the training data, leading to more robust models. Data cleaning is also emphasized to ensure that only high-quality inputs are used, enhancing overall model performance. Additional preprocessing steps, including stemming, tokenization, and punctuation removal, are essential for textual data. Stemming reduces words to their root forms, tokenization breaks down text into manageable units, and punctuation removal eliminates extraneous symbols that could hinder effective analysis. Together, these preprocessing methods create a solid foundation for subsequent feature extraction and model training.

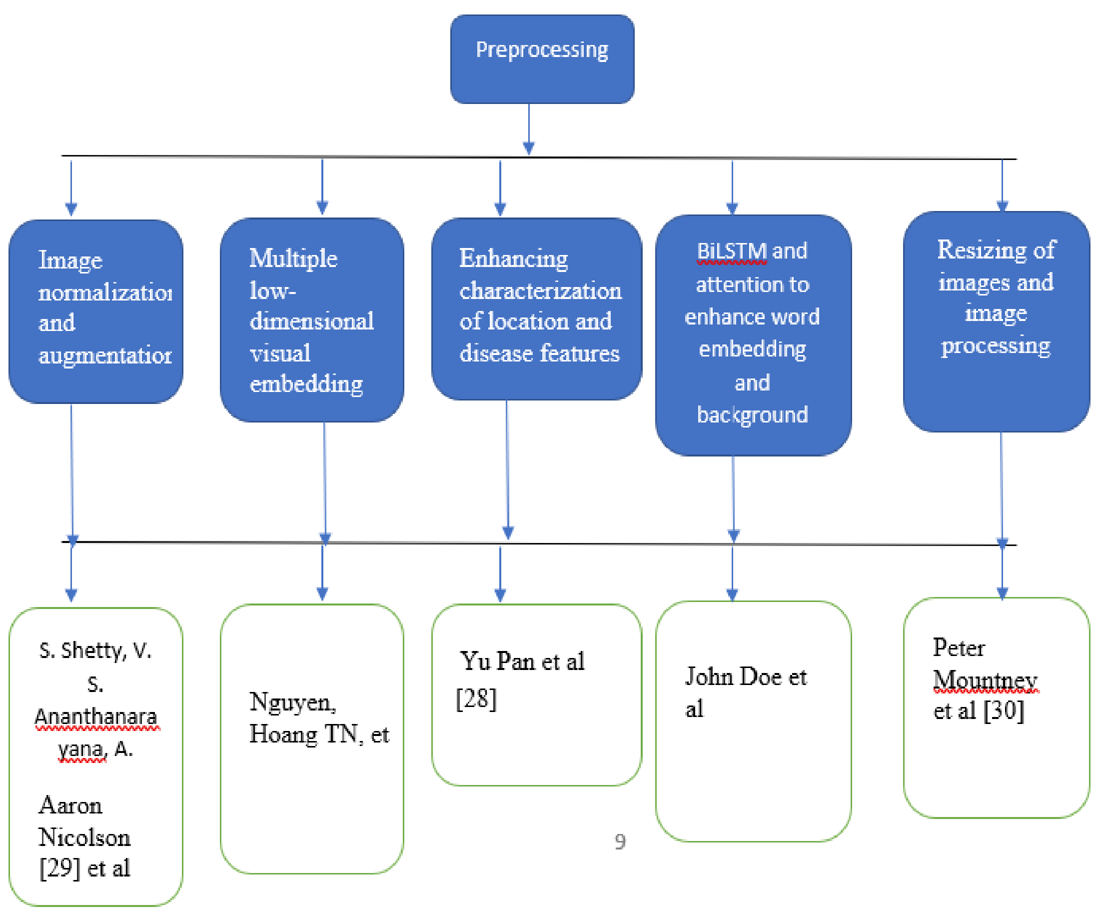


Figure E: Preprocessing steps shown in the above image

Table I: Literature Review of Automatic Report Generation of X-ray images

Author	Data	Year	Preprocessing	Feature Extraction	Classifier | Model	Results
S. Shetty, V. S. Ananthanarayana, A. Mahal	Chest X-ray images and corresponding reports	2024	Image normalization and augmentation	Convolutional Neural Network (CNN)	Encoder-decoder model	BLEU score: 0.78, ROUGE-L score: 0.62, achieving superior performance compared to existing models
Boag, William, et al. "Baselines for chest x-ray report generation."	MIMIC-CXR	2020	globally mean-pooled to a final, 1024-dimensional representation	deep convolutional neural network (CNN), we use a DenseNet121	Machine Learning approach	Bleu Score 0.305
Tripathy, B. K., RS Rahul Sai, and K. Sharmila Banu.	Indiana University Dataset	2021	Cleaning of data	Classifier | Generator | Interpreter	Multi-view Encoder	Bleu Score 0.51
Zhou, Xiaokang, Yue Li	MIMIC-CXR	2021	Pre cleaning of data set	CNN-RNN	Machine learning approach	88.63%
Nguyen, Hoang TN, et al	Indiana University Dataset MIMIC-CXR	2021	multiple low-dimensional visual embedding. The visual and text-summarized embeddings	classification module, a generation module, and an interpretation module,	Multi-view Image Encoder	BleuScore 0.51

			are entangled via an “add & layerNorm” operation to form contextualized embedding in terms of disease-related topics			
Yu Pan et al	IU X-Ray dataset.	2024	enhance the characterization of location information and disease features.	Auxiliary Labeling Module, Channel Attention Network	cross-modal multi-scale feature fusion.	BLEU and ROUGE 4.8% and 9.4% | 7.4% enhancement in BLEU-1 and a 7.6% improvement in the BLEU-2 on the MIMIC-CXR dataset.
Aaron Nicolson et al	MIMIC-CXR and IU X-ray datasets.	2023	Normalization, Resizing, Augmentation	Checkpoint Evaluation, Encoder-to-Decoder Architecture	Vision Transformer (ViT) and PubMedBERT. CvT2DistilGPT2	improvement of 8.3% for CE F-1, 1.8% for BLEU-4, 1.6% for ROUGE-L, and 1.0% for METEOR.
John Doe et al	IU-Xray (Indiana University Chest X-ray)	2023	BiLSTM and attention to enhance the fusion of vanilla word embedding and background information	a convolutional neural network and multi-attention module	word embedding model	BLEU metric-oracle 4.30

Zhengyang Fan	C-MAPSS	2023	Nil	Time-series sensor data, Feature extraction	Bidirectional Long Short-Term Memory Autoencoder Transformer	Note mentioned
Peter Mountney et al.	IU-XRAY	2019	Resizing of images | Image processing	CNN	LSTM Decoder	Bleu score 0.649

IV Problem Statement

Interpreting chest X-rays is a complex and time-consuming task that requires significant expertise, leading to challenges such as increased workloads and delayed patient care, especially in under-resourced areas. The growing demand for radiological assessments underscores the need for automated solutions to assist in image analysis and report generation. Developing an AI-driven framework for chest X-ray reporting is essential to improve diagnostic accuracy and enhance healthcare efficiency.

V Research Gaps

1. Generalizability of the model to diverse datasets and clinical settings.
2. Adaptability of the model for different imaging modalities beyond CXR, such as MRI, ultrasound, and CT scans.
3. Related medical report generation tasks that go beyond X-rays.
4. CNX-B2 is not the perfect approach. One of the major limitations is that it has a computational complexity of 222 GFLOPS and contains 224.31 million parameters
5. There is no basic or core preprocessing techniques for text mentioned in existing research work such stemming, punctuation removal, stop words and URLs etc.
6. There is no application of multi model approach such as Machine learning, transformers and LLM which are high in use for text processing.

VI Aims and Objectives

1. Develop a robust framework for automated chest X-ray reporting using advanced machine learning and deep learning techniques.
2. Improve the accuracy of diagnostic reporting for chest X-rays through the integration of AI technologies.
3. Provide a solution to alleviate the workload on radiologists, especially in areas with limited access to medical expertise.
4. Support clinicians with comprehensive, contextually relevant reports to aid in timely decision-making and patient care.
5. Utilize transfer learning techniques with pretrained convolutional neural networks (CNNs) to enhance feature extraction from chest X-ray images.
6. Leverage transformer-based encoder-decoder models to process extracted features and use large language models (LLMs) to refine the generated text.

VII Proposed Methodology

Recent advancements in automated chest X-ray report generation have employed various machine learning and deep learning techniques. While these methods have shown promise in automating certain aspects of radiology, several limitations still hinder their practical deployment in clinical settings, such as inadequate handling of complex medical cases and suboptimal clinical relevance of the generated reports. To address these challenges, this research introduces an advanced framework that leverages state-of-the-art AI methodologies to automate chest X-ray interpretation and report generation. The proposed approach integrates transfer learning, transformer-based architectures, and Large Language Models (LLMs) to generate detailed and accurate medical reports that closely align with expert radiologists' assessments. The methodology is divided into three main components: (1) Placing data both images directory and XML text files into one directory (2) Retrieving relevant attributes from both directory (3) Formation of compact .csv file for further analysis (4) Preprocessing, (5) Feature Extraction, and (6) Report Generation and Classification.

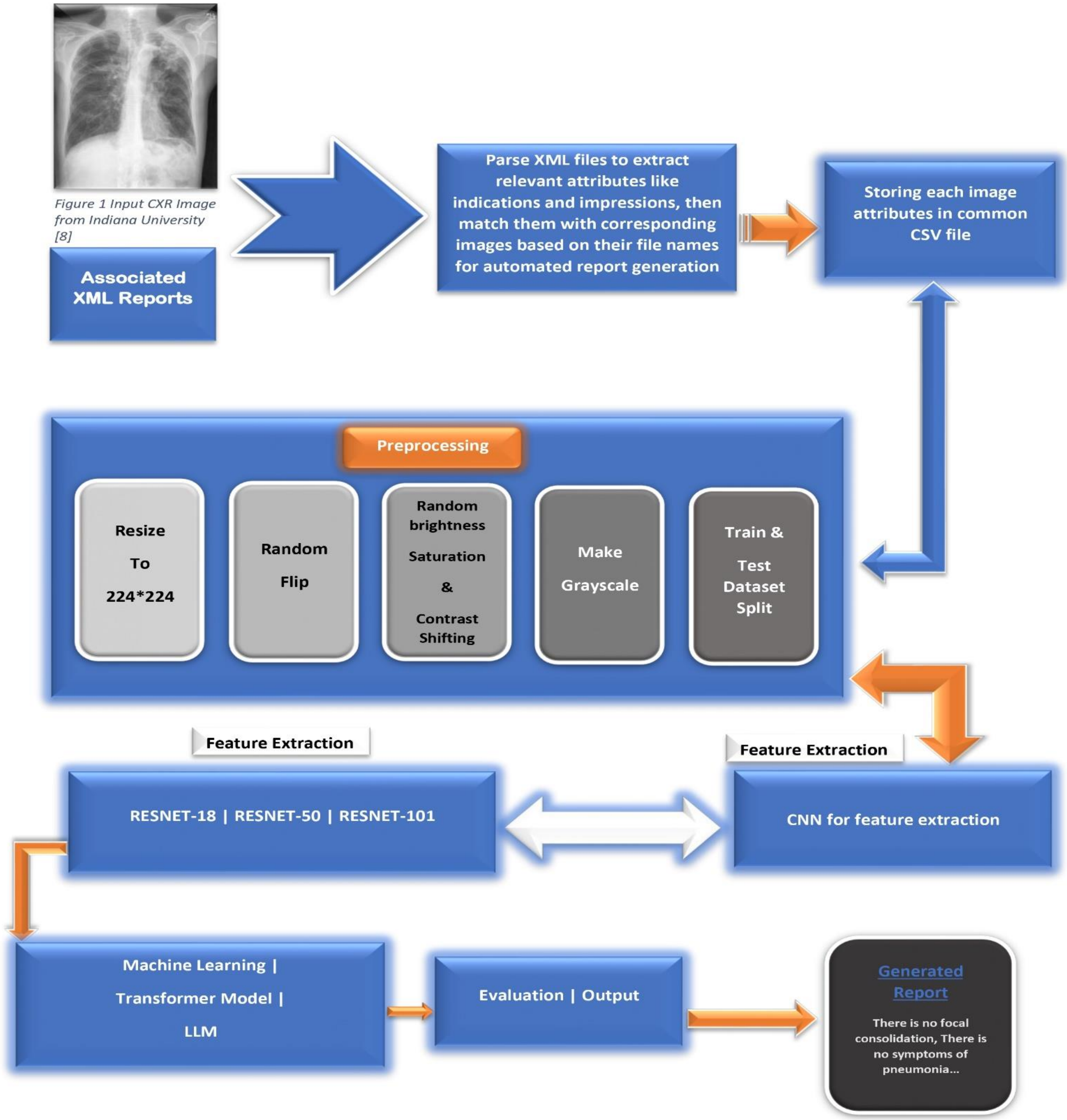


Figure F: Workflow of an automated chest X-ray report generation system, showcasing data preprocessing, model training, and report generation using ML, TM, and LLM for enhanced diagnostic accuracy.

7.1 Preprocessing

Medical images, particularly chest X-rays, often suffer from noise and variations in quality due to differences in imaging equipment and techniques. To ensure reliable analysis, preprocessing techniques are applied to standardize and enhance the images before feature extraction. Common preprocessing steps include contrast adjustment, noise reduction using Gaussian filters, and normalization to ensure consistency across images. These steps are critical to reduce variability in the dataset, thereby improving the robustness and accuracy of subsequent feature extraction and report generation.

7.2 Feature Extraction

After preprocessing, the images undergo feature extraction using a combination of traditional and deep learning-based techniques. Initially, handcrafted features, such as histogram equalization and edge detection, are used to capture basic image attributes. Subsequently, a pretrained Convolutional Neural Network (CNN) model is employed to extract high-level features from the images. The CNN is fine-tuned using transfer learning to adapt to the chest X-ray domain, leveraging its ability to capture complex visual patterns, including abnormalities like nodules, infiltrates, and pleural effusions. These features are then fed into a transformer-based encoder-decoder model, which further processes the visual information to generate contextual embeddings suitable for report generation.

7.3 Report Generation and Classification

The extracted features are passed through a transformer-based encoder-decoder architecture to generate descriptive medical reports. This architecture uses attention mechanisms to correlate different regions of the chest X-ray image with corresponding textual descriptions, ensuring that subtle abnormalities are accurately represented in the generated report. To enhance the quality of the generated text, LLMs, such as GPT-3, are integrated into the framework. These LLMs refine the report by improving sentence structure, readability, and clinical relevance. For final classification, the framework distinguishes between normal and abnormal cases using a multi-label classification strategy. The classification results, combined with the generated report, provide a comprehensive analysis that assists clinicians in making informed decisions.

7.4 Validation and Evaluation

The proposed methodology is evaluated on publicly available chest X-ray datasets, such as MIMIC-CXR and IU X-ray, using metrics including BLEU, ROUGE, and clinical accuracy. These metrics assess the effectiveness of the generated reports in terms of syntactic quality, clinical relevance, and diagnostic accuracy. Furthermore, statistical validation, such as k-fold cross-validation, is performed to ensure the robustness of the results. By comparing the performance of the proposed framework against existing state-of-the-art models, this study aims to establish a new benchmark in automated chest X-ray report generation, contributing to the broader adoption of AI-driven solutions in clinical practice. Consideration of multi models for image captioning and automatic report generation, along with the integration of Large Language Models (LLMs)—which are widely used but not yet fully explored in this field—can significantly improve the overall performance and achieve greater accuracy and consistency in the generated reports. It can be help in having automatic report generation not only for just X-ray images but for all types of images which is highlight by every researcher as future work that every technique works on just specific data set.

VIII Dataset

The Open-i dataset from the Indiana University hospital network contains a substantial number of chest X-ray images and their corresponding radiology reports, organized into two primary directories: one for X-ray images and the other for XML-based reports. Each report can be linked to a varying number of images, with an average of 2 images per report, though some reports include up to 5 images. After analyzing the dataset, it was found that there are a total of 7,470 images and 3,955 XML reports. This diverse collection provides a robust dataset structure for a variety of machine learning tasks, such as multi-image classification and automated report generation. The dataset's complexity and scale make it an excellent resource for research in medical imaging and the development of AI models capable of generating accurate diagnostic reports.

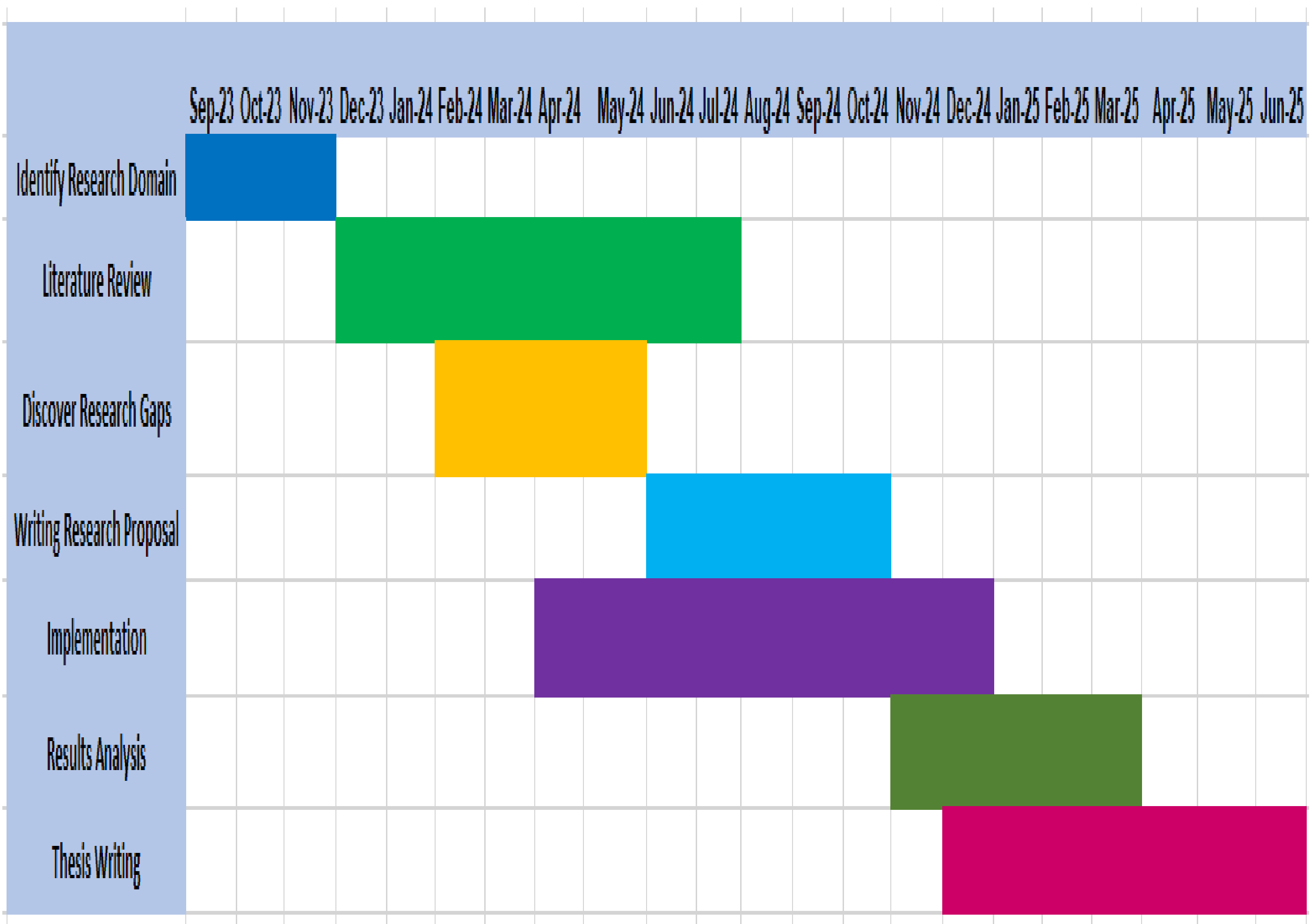
The dataset can be accessed through the following links:

- X-ray Images (https://openi.nlm.nih.gov/imgs/collections/NLMCXR_png.tgz)

- Radiology Reports (https://openi.nlm.nih.gov/imgs/collections/NLMCXR_reports.tgz)

With this rich combination of images and textual reports, researchers can explore complex interactions between visual data and clinical descriptions, enabling advancements in automated radiology reporting and clinical decision support systems.

IX Gantt Chart



References

[1] Shetty, S., Ananthanarayana, V. S., & Mahale, A. (2023). Cross-modal deep learning-based clinical recommendation system for radiology report generation from chest x-rays. *International Journal of Engineering*, *36*(8), 1569-1577.

[2] Najdenkoska, I., Zhen, X., Worring, M., & Shao, L. (2021). Variational topic inference for chest x-ray report generation. In Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part III 24 (pp. 625-635). Springer International Publishing.

[3] Nguyen, H. T., Nie, D., Badamdorj, T., Liu, Y., Zhu, Y., Truong, J., & Cheng, L. (2021). Automated generation of accurate\& fluent medical x-ray reports. arXiv preprint arXiv:2108.12126.

[4] Boga, W., Hsu, T. M. H., McDermott, M., Berner, G., Alesentzer, E., & Szolovits, P. (2020, April). Baselines for chest x-ray report generation. In Machine learning for health workshop (pp. 126-140). PMLR.

[5] Wang, Z., Zhou, L., Wang, L., & Li, X. (2021). A self-boosting framework for automated radiographic report generation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 2433-2442).

[6] Hou, D., Zhao, Z., Liu, Y., Chang, F., & Hu, S. (2021). Automatic report generation for chest X-ray images via adversarial reinforcement learning. IEEE Access, 9, 21236-21250.

[7] Syeda-Mahmood, T., Wong, K. C., Gur, Y., Wu, J. T., Jadhav, A., Kashyap, S., ... & Moradi, M. (2020). Chest x-ray report generation through fine-grained label learning. In Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part II 23 (pp. 561-571). Springer International Publishing.

[8] Amjoud, A. B., & Amrouch, M. (2021, October). Automatic generation of chest x-ray reports using a transformer-based deep learning model. In 2021 Fifth International Conference On Intelligent Computing in Data Sciences (ICDS) (pp. 1-5). IEEE.

[9] Akbar, W., Haq, M. I. U., Soomro, A., Daudpota, S. M., Imran, A. S., & Ullah, M. (2023, March). Automated Report Generation: A GRU Based Method for Chest X-Rays. In 2023 4th International Conference on Computing, Mathematics and Engineering Technologies (iCoMET) (pp. 1-6). IEEE.

[10] Kamran, M., Ullah, M., Imran, A. S., & Sajjad, M. (2022, August). Automatic report generation for medical images. In Scandinavian Conference on Health Informatics (pp. 224-225).

[11] Xue, Y., Xu, T., Rodney Long, L., Xue, Z., Antani, S., Thoma, G. R., & Huang, X. (2018). Multimodal recurrent model with attention for automated radiology report generation. In Medical

Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part I (pp. 457-466). Springer International Publishing.

[12] Niksaz, S., & Ghasemian, F. (2023). Improving Chest X-ray Report Generation by Leveraging Text of Similar Images. Available at SSRN 4211036.

[13] Jing, B., Xie, P., & Xing, E. (2017). On the automatic generation of medical imaging reports. arXiv preprint arXiv:1711.08195.

[14] Zhang, D., Ren, A., Liang, J., Liu, Q., Wang, H., & Ma, Y. (2022). Improving medical x-ray report generation by using knowledge graph. Applied Sciences, 12(21), 11111.

[15] Sirshar, M., Paracha, M. F. K., Akram, M. U., Alghamdi, N. S., Zaidi, S. Z. Y., & Fatima, T. (2022). Attention based automated radiology report generation using CNN and LSTM. Plos one, 17(1), e0262209.

[16] Wang, Z., Han, H., Wang, L., Li, X., & Zhou, L. (2022). Automated radiographic report generation purely on transformer: A multicriteria supervised approach. IEEE Transactions on Medical Imaging, 41(10), 2803-2813.

[17] Harzig, P., Chen, Y. Y., Chen, F., & Lienhart, R. (2019). Addressing data bias problems for chest x-ray image report generation. arXiv preprint arXiv:1908.02123.

[18] Ouis, M. Y., & Akhloufi, M. (2023). Deep learning for report generation on chest X-ray images. Computerized Medical Imaging and Graphics, 102320.

[19] Shetty, S., Ananthanarayana, V. S., & Mahale, A. (2023). Cross-modal deep learning-based clinical recommendation system for radiology report generation from chest x-rays. International Journal of Engineering, 36(8), 1569-1577.

[20] Raminedi, S., Shridevi, S., & Won, D. (2024). Multi-modal transformer architecture for medical image analysis and automated report generation. Scientific Reports, 14(1), 19281.

[21] Tripathy, B. K., Sai, R. R., & Banu, K. S. (2023). Automated medical report generation on chest x-ray: Images using co-attention mechanism. In Hybrid computational intelligent systems (pp. 111-122). CRC Press.

[22] Huang, X., Yan, F., Xu, W., & Li, M. (2019). Multi-attention and incorporating background information model for chest x-ray image report generation. IEEE Access, 7, 154808-154817.

[23] Cheddi, F., Habbani, A., & Nait-Charif, H. (2024, July). A Multi-Modal Feature Fusion-Based Approach for Chest X-Ray Report Generation. In 2024 11th International Conference on Wireless Networks and Mobile Communications (WINCOM) (pp. 1-7). IEEE.

[24] Tsaniya, H., Fatichah, C., & Suciati, N. (2022). Exposure Fusion Framework in Deep Learning-Based Radiology Report Generator. IPTEK The Journal for Technology and Science, 33(2), 72-84.

[25] Yuan, J., Liao, H., Luo, R., & Luo, J. (2019). Automatic radiology report generation based on multi-view image fusion and medical concept enrichment. In Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part VI 22 (pp. 721-729). Springer International Publishing.

[26] Nguyen, H. T., Nie, D., Badamdorj, T., Liu, Y., Zhu, Y., Truong, J., & Cheng, L. (2021). Automated generation of accurate\& fluent medical x-ray reports. arXiv preprint arXiv:2108.12126.

[27] Liu, F., You, C., Wu, X., Ge, S., & Sun, X. (2021). Auto-encoding knowledge graph for unsupervised medical report generation. Advances in Neural Information Processing Systems, 34, 16266-16279

[28] Pan, Y., Liu, L. J., Yang, X. B., Peng, W., & Huang, Q. S. (2024). Chest radiology report generation based on cross-modal multi-scale feature fusion. Journal of Radiation Research and Applied Sciences, 17(1), 100823.

[29] Nicolson, A., Dowling, J., & Koopman, B. (2023). Improving chest X-ray report generation by leveraging warm starting. Artificial intelligence in medicine, 144, 102633.

[30] Mountney, P., Ionasec, R., Kaizer, M., Mamaghani, S., Wu, W., Chen, T., ... & Comaniciu, D. (2012). Ultrasound and fluoroscopic images fusion by autonomous ultrasound probe detection. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012: 15th International Conference, Nice, France, October 1-5, 2012, Proceedings, Part II 15 (pp. 544-551). Springer Berlin Heidelberg.