Study of X-Ray interpretation Techniques

Patel Ketul(ketvpate@iu.edu), and Patil Harshwardhan(hrpatil@iu.edu)
Indiana University, Bloomington

***Abstract*—This paper presents a study of different deep learning approaches for automatically generating succinct report summaries from X-ray images. This is a comparative study of the models trained on the Indiana University - Chest X-Rays dataset. In this project, we have conducted in-depth literature reviews and suggested improvements to existing deep learning models, all geared towards optimizing the interpretation of Chest X-ray images.**

I. Introduction

Radiology is a branch of medicine that uses imaging technology to diagnose and treat diseases. Doctors specializing in this field, known as radiologists, are equipped with years of training and expertise to interpret X-rays, CT scans, MRIs, nuclear medicine scans, PET scans, ultrasounds, etc. As the number of medical images continues to grow, so too does the workload of radiologists responsible for interpreting them. The sheer volume can overwhelm even the most skilled radiologists, potentially impacting their ability to analyze and interpret them effectively. One possible option is to automate this process.

Machine learning stands out as the most effective way to automate the analysis and diagnosis of medical images. Its diverse applications include:

- **Image segmentation:** Precisely segmenting anatomical structures like the brain, spine, lungs, liver, kidneys, and colon for accurate measurements and analysis.[6]
- **Computer-aided diagnosis and detection:** Detecting and highlighting suspicious lesions in CT or MRI scans, like mammograms and CT colonography, to assist radiologists in making faster and more accurate diagnoses.[10]
- **Brain function analysis:** Analyzing fMRI data to understand brain activity and diagnose neurological diseases, providing valuable insights into brain function and behavior.
- **Content-based image retrieval:** Searching large medical image databases efficiently based on specific features.
- **Text analysis of radiology reports**: Utilizing NLP and NLU to extract key information, identify trends, and even assist in generating reports, streamlining workflow and improving patient care efficiency.

These applications demonstrate machine learning's transformative potential in radiology.

Out of all theses applications, we caught our eye on the topic of medical image interpretation and reporting. Many factors can contribute to faulty report generation - lack of knowledge, staff shortage, excessive workload, etc. This process can be error-prone, even for experienced radiology specialists. To reduce the error occurrences, an automated system is needed. [2]

Existing literature on automated medical report generation based on radiological images centers around the recurrent neural network with attention mechanisms[5]. Notably, research concentrating on the Indiana University X-Ray dataset has sequentially addressed the classification of frontal and lateral views [9], the grading of cardiomegaly severity [10], and the computer-aided diagnosis (CAD) of lung diseases [11]. These studies commonly employed CNN and Variational Topic Inference for classification and exploratory analysis. Additionally, recent work in image camptioniong explores using GRU models [12], knowledge graphs [13], and Multimodal Large Language Models for automated report generation [14].

In this paper, we will present comparative study of the different techniques used for automated report generation. Our primary focus is on the comparative study of distinct visual feature extraction models, namely DenseNet121, ResNet101, InceptionV3, and Xception. The contributions of our work lie in evaluating and comparing the performance of these models to discern their effectiveness in the automated generation of radiology reports. For comparing performance we will use BLEU [17], METEOR [18], ROUGE[19], CIDEr[20] metrics.

II. Data

In this study, we employ the Indiana University Chest X-Ray collection [15] as our primary dataset for evaluation. This dataset encompasses a total of 7,471 X-ray images and there are about 3955 patients text reports available in .XML format. Notably, some reports are associated with more than one image. Each report is uniquely linked to a pair of images representing the frontal and lateral views.

The content of these reports is structured in XML format, necessitating XML parsing to extract and convert the information into a more accessible CSV format. The report sections include essential medical information such as comparison, indication, findings, and impression.

For our model, we consider both image and text inputs. The image data consists of the X-ray images associated with each report. Additionally, we extract textual information from the abstract, comparison, indication, and findings sections of the reports to serve as text input features.

The target variable for our analysis is the ”Impression” section of the medical reports, which is treated as a text feature. The combination of image and text inputs allows the model to learn complex patterns and relationships essential for automatic report generation in the field of medical image interpretation.

Comparison: None.

Indication: Positive TB test

Findings: The cardiac silhouette and mediastinum size are within normal limits. There is no pulmonary edema. There is no focal consolidation. There are no XXXX of a pleural effusion. There is no evidence of pneumothorax.

Impression: Normal chest x-XXXX.

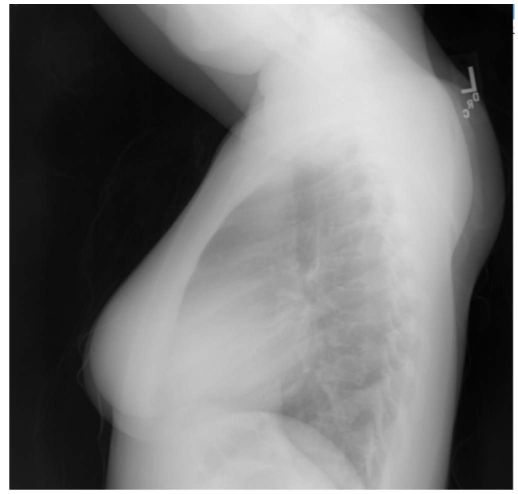

Fig. 1. A sample report for the chest x-ray where the comparision, indication, findings and impressions are mentioned. [16]

III. Methodology

A. Data Pre-processing and Augmentation

Mention what data preprocessing techniques you have utilized and how you augmented the data

B. Page Numbers

Do not type any page numbers. Page numbers will be added in the digest editing/production process.

C. Equations

Equations should be centered in the column and numbered sequentially. Place the equation number to the right of the equation within a parenthesis, with right justification within its column. An example would be

D. Figures

Most of the following applies to Microsoft Word. Figures should utilize as much of the column width as possible in order to maximize legibility. Use a sans serif font, such as Helvetica. Helvetica is larger and much easier to read than Times New Roman. Using 8- to 10-point Helvetica usually results in a legible figure.

Place figure captions directly below each figure, as in the example, Fig. 1 below. Use

IV. Related Work

Use the examples provided [1] - [2] as a guide.

V. Conclusion

Write your conclusions here.

VI. Future Work

As a potential avenue for future research, extending our models to include the PadChest[ZXX] dataset could be valuable. The PadChest dataset stands out as one of the largest publicly available chest X-ray databases, providing substantial data for training supervised models specifically tailored to radiographs. Exploring the integration of this dataset into our model training pipeline could enhance the generalization and performance of our models, fostering a broader and more robust application in the field of medical image interpretation.

References

[1] Wang, Shijun, and Ronald M. Summers. Machine Learning and Radiology. Medical Image Analysis 16, no. 5 (2012): 933-951. https://doi.org/10.1016/j.media.2012.02.005.

[2] Y. Xue, T. Xu, L.R. Long, Z. Xue, S. Antani, G.R. Thoma, X. Huang Multimodal recurrent model with attention for automated radiology report generation Proceedings of the international conference on medical image computing and computer-assisted intervention (MICCAI), Springer, Granada, Spain (2018), pp. 457-466

[3] Mazurowski, M.A., Buda, M., Saha, A. and Bashir, M.R. (2019), Deep learning in radiology: An overview of the concepts and a survey of the state of the art with focus on MRI. J. Magn. Reson. Imaging, 49: 939-954. https://doi.org/10.1002/jmri.26534

[4] Sahiner, B., Pezeshk, A., Hadjiiski, L.M., Wang, X., Drukker, K., Cha, K.H., Summers, R.M. and Giger, M.L. (2019), Deep learning in medical imaging and radiation therapy. Med. Phys., 46: e1-e36. https://doi.org/10.1002/mp.13264

[5] Cho, Kyunghyun, Bart van Merrienboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation." 2014. arXiv preprint, 1406.1078, cs.CL.

[6] Hesamian, Mohammad Hesam, Wenjing Jia, Xiangjian He, and Paul Kennedy. "Deep learning techniques for medical image segmentation: achievements and challenges." Journal of digital imaging 32 (2019): 582-596.

[7] Doi, Kunio. "Computer-aided diagnosis in medical imaging: historical review, current status and future potential." Computerized medical imaging and graphics 31, no. 4-5 (2007): 198-211.

[8] Chen, Chung-Ming, Yi-Hong Chou, Norio Tagawa, and Younghae Do. "Computer-aided detection and diagnosis in medical imaging." Computational and mathematical methods in medicine 2013 (2013).

[9] Z. Xue et al., "Chest X-ray Image View Classification," 2015 IEEE 28th International Symposium on Computer-Based Medical Systems, Sao Carlos, Brazil, 2015, pp. 66-71, doi: 10.1109/CBMS.2015.49.

[10] S. Candemir, S. Rajaraman, G. Thoma and S. Antani, ”Deep Learning for Grading Cardiomegaly Severity in Chest X-Rays: An Investigation,” 2018 IEEE Life Sciences Conference (LSC), Montreal, QC, Canada, 2018, pp. 109-113, doi: 10.1109/LSC.2018.8572113.

[11] S. V. G, N. Ponraj and D. P. L, ”Study on Public Chest X-ray Data sets for Lung Disease Classification,” 2021 3rd International Conference on Signal Processing and Communication (ICPSC), Coimbatore, India, 2021, pp. 54-58, doi: 10.1109/ICSPC51351.2021.9451726.

[12] Akbar, Wajahat, Muhammad Inam Ul Haq, Abdullah Soomro, Sher Muhammad Daudpota, Ali Shariq Imran, and Mohib Ullah. ”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, 2023.

[13] Lee, Seowoo, Jiwon Youn, Mansu Kim, and Soon Ho Yoon. ”CXR-LLaVA: Multimodal Large Language Model for Interpreting Chest X-ray Images.” arXiv preprint arXiv:2310.18341 (2023).

[14] Liu, Weihua, Youyuan Xue, Chaochao Lin, and Said Boumaraf. ”Dynamic Multi-Domain Knowledge Networks for Chest X-ray Report Generation.” arXiv preprint arXiv:2310.05119 (2023).

[15] Demner-Fushman, D., Kohli, M.D., Rosenman, M.B., Shooshan, S.E., Rodriguez, L., Antani, S., Thoma, G.R., McDonald, C.J.: Preparing a collection of radiology examinations for distribution and retrieval. J. Am. Med. Inform. Assoc. 23(2), 304–310 (2015)

[16] Indiana University. ”Indiana University - Chest X-Rays (XML Reports).” 2023. Radiology Reports Collection. Accessed [Date]. Available at: https://i.imgur.com/PWo3x47.png.

[17] Papineni, K., Roukos, S., Ward, T., and Zhu, W. ”Bleu: A Method for Automatic Evaluation of Machine Translation.” In Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, July 6-12, 2002, Philadelphia, PA, USA. ACL (2002), pp. 311-318.

[18] Banerjee, S., and Lavie, A. ”METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments.” In Proceedings of the Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation andor Summarization at ACL 2005, Ann Arbor, Michigan, USA, June 29, 2005. Association for Computational Linguistics (2005), pp. 65-72.

[19] Lin, C.-Y. ”ROUGE: A Package for Automatic Evaluation of Summaries.” In Text Summarization Branches Out, Barcelona, Spain. Association for Computational Linguistics (July 2004), pp. 74-81.

[20] Vedantam, R., Zitnick, C.L., Parikh, D. ”Cider: Consensus-Based Image Description Evaluation.” In IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, Boston, MA, USA, June 7-12, 2015. IEEE Computer Society (2015), pp. 4566-4575.