International Journal of Innovative Research in Engineering and Management (IJIREM)
ISSN (Online): 2350-0557, Volume-12, Issue-2, April 2025
https://doi.org/10.55524/ijirem.2025.12.2.20
Article ID IJIRE-1397, Pages 124-129
www.ijirem.org

From Images to Insights: The Role of AI in Radiology Report Generation

Vaishali[1], and Shelly Garg[2]

[1]Student, Department of Computer Science & Engineering, Amity Institute of Information and Technology, Amity University, Gurugram, Haryana, India
[2]Assistant Professor, Department of Computer Science and Engineering, Amity University, Gurugram, Haryana, India

Correspondence should be addressed to Vaishali; Vaishalikakash@gmail.com

Received: 2 March 2025 Revised: 14 March 2025 Accepted: 28 March 2025

ABSTRACT- Radiology reports are important for diagnosing and managing a variety of health conditions, but manually crafting them is always laborious and often tiresome for radiologists. Machine learning has made significant advances and is beginning to show promise to help automate this process in recent years by enabling the production of reports from a clinical image such as a chest X-ray or CT scan. This paper provides a review of the state of machine learning in Radiology Report Generation. We focus on the ability to use Convolutional Neural Networks (CNNs) to understand image data as well as transformer-based models to produce text data. In addition, we will examine the newer multimodal models which produce output combining image and text for potentially more accurate and contextually aware reports. Furthermore, our review highlights important observations, particularly the growing use of Deep Learning models and large medical datasets (e.g., MIMIC-CXR, CheXpert). Although there has been advancement, challenges remain such as clinically inaccurate results, lack of explainability, and inadequate variety of data. In summary, we conclude that although machine learning has a strong potential to significantly reduce work for radiologist and add consistent reporting, substantial additional work needs to be done to bring these systems to greater trust and reliability for real-world clinical settings.

KEYWORDS- Radiology Report Generation, Machine learning, Deep learning, Convolutional Neural Networks (CNNs), Transformers, Medical Image Analysis, Natural Language Processing (NLP).

I. INTRODUCTION

Radiological imaging is crucial in diagnosing, treatment planning, and monitoring several medical conditions. Hospitals generate thousands of imaging studies every day, including chest x-rays, Computed Tomography (CT), and Magnetic Resonance Imaging (MRI). Radiologists are responsible for interpreting studies and transcribing them into structured and clinically useful reports. However, this procedure is laborious, difficult cognitively, and may be influenced by elements such as fatigue, experience, or caseload pressures.

As patient volumes increase and radiologist shortages endemic in many regions continue, the need for intelligent and automated solutions becomes urgent. One potential area of development is Machine Learning and, in particular, Deep Learning to assist with or fully automate the writing of radiology reports. The rationale is two-fold; to alleviate the load upon radiologists and bring a process of automated and intelligent writing to keep the quality of reporting consistent and minimise errors or omissions that might affect patient care. Advances in computer vision or Natural Language Processing (NLP) have enabled the writing of models that can learn from large datasets of medical images and text reports paired together [1], [2].

Convolution neural networks (CNNs) are a commonly used architecture for processing visual evidence with the aim of extracting clinically and radiologically relevant features from images. In contrast, language models, which include recurrent neural networks (RNNs), attention-based models, and transformers have been very successful at producing descriptive, relevant and contextually aware text output. There has also been the introduction of multimodal architectures that can learn from both image and textual modalities jointly, and have shown improved outcomes when generating consistent and clinically relevant radiology reports [3][4].

With the availability of large, open-access datasets such as MIMIC-CXR and CheXpert, which include thousands of labeled images and associated radiology reports, researchers can train and evaluate models with real-world clinical data, which has advanced the speed and transition to usable hospital and diagnostic center technology [5][6].

There is a lot of promise in this field, but several challenges remain. The models ideally need to learn complex medical language and terminology; be able to consider alternative and/or ambiguous findings; and output explainable evidence to clinicians so they can trust the model outputs for clinical purposes. Other limitations include, data bias, privacy issues, and generalisability due to very limited population subsets being used during the models training.

This paper presents a bibliographic review of the current status of machine learning methods for automated radiology report generation. To begin, we describe the basis of the machine learning methods employed in this area. We also provide an overview of all existing datasets, and we present the major contributions in the literature, common evaluation metrics, and limitations of current methods with future research directions that could aid in the clinical acceptance and reliability of these procedures.

II. LITERATURE REVIEW

Developing radiology reports is an essential component of contemporary health care education and practice. In this process, radiologists interpret medical images (e.g., chest radiographs, CT scans, MRIs) and write reports that summarize findings. This is a tedious process, and it is subject to variability, largely as a result of radiologist workload, experience, and cognitive fatigue. The effect of increased patient volumes and systematic radiologist shortages have raised significant questions about how to either assist with or completely automate the report generation process, and machine learning (ML)—particularly deep learning—represents an important tool to help confront both of these complexities. The main tasks of an automated radiology report generation system involve understanding the visual content of medical images (image interpretation), detecting any diseases present, and generating a coherent, clinically relevant textual summary. Initial works on automated radiology report generation have borrowed heavily from image captioning domains, using CNNs to assess images and RNNs (Long Short-Term Memory networks) to generate text descriptions. However, most initial attempts to assess automated radiology report generation have been restricted by the absence of domain-specific training data and typically rendered generic or inaccurate reports [7].

The following approaches involved end-to-end or "blind" deep learning models that learned both the image and text representations along with an attention mechanism to target text on specific areas of the image they were generating a report for. For instance, the TieNet model used CNNs alongside LSTM models driven by the attention mechanism and showed improvement for image classification and for text generation tasks [8]. Nonetheless, many of these models struggled to generate full-length multi-sentence reports that would accompany a clinical report and that incorporated accurate clinical reasoning in them. With the advent of Transformer models (for example, BERT and GPT), the generative dramatic improvements in natural language processing (NLP) tasks, and the subsequent adaptations for radiology report generation through vision-language models (e.g., R2Gen), and BERT variants (e.g., BioBERT), there was a new avenue of type automatic speech and language generation for radiology report generation [9]. Furthermore, hybrid CNN-Transformer models that combined text and image data have emerged as commonplace, as these approaches deepened the understanding of both the visual cues from the image and linguistically academic language structure encoded in medical language, thereby improving report accuracy [10].

Large datasets such as MIMIC-CXR[11], CheXpert[12], and Open-i have been invaluable to the field as they provide thousands of labeled images and related radiology reports. They support the development and evaluation of ML models on real (to some extent) clinical data, however there are still obstacles to effectively study the reliability and generalizability of models utilizing large datasets. One of the first issues is relating to the variability of labels across large datasets and the fact that there is noisy data which can detrimentally affect the model. Finally, when using evaluation metrics traditionally in NLP (i.e. BLEU, ROUGE), it becomes increasingly clear they do not longitudinally correlate with clinical accuracy. This recognition has required the development of more clinically relevant evaluation metrics for radiology report generation. [13].

Although it is promising that deep learning models have made strides in efforts to automate radiology report generation, their capabilities are still limited and many gaps remain in the fields of imaging, reporting and machine learning models. The majority of models to-date have concentrated on chest X-rays/deep radiology images that have had reports produced using machine learning. Beyond chest X-rays, these models (including those that automatically generate a report) do not focus on the many imaging modalities included in radiology practice - for example, ultrasound (US) - nor do they contribute to the further investigation of these other forms of imaging. More importantly, most of the models are not transparent; transparency is important for clinicians wishing to trust and interpret the results from the model. Domain-specific models that employ deep learning, can deal with multiple forms of medical terminologies, struggles with ambiguity/interpreter uncertainty and can engage with multiple modalities from radiology and the individual clinical team are still a largely active area of research. Future directions of research would be to engage with growing generalization across differing clinical settings, privacy, and whole model explainability to increase the ease of their transition into the real-world clinical environment. Most recently, several papers have made a case for the use of structured data and medical ontologies across complex medical terminologies used within reports and investigations, for example within the work completed by Das et al., Deep Learning techniques for medical image analysis [14]. Further, they highlight areas of research and trends in the development of Deep Learning methods for image analysis that have been documented in the literature which may also impact radiology report generation tasks [15]. Additionally, Lakhani & Sundaram discuss the importance of using domain-specific neural networks within the radiology space (where they suggest an example of using Domain-specific neural networks for lung radiographs for the diagnosis of pulmonary tuberculosis) has been shown to be efficient and effective [16]. Razzak et al. discuss big data analytics and data utilization within the medical imaging field and recognized the importance of automated systems and how these systems enhance role and process of using detection of early disease [17] and Liu et al. also consider the role of deep convolutional neural networks used in the diagnosis of pneumonia at how these developments can be utilized in other radiological diseases [18].

III. DATASETS IN RADIOLOGY REPORT GENERATION

Numerous large and publicly available datasets have played a pivotal role in moving forward with research on automated radiology report generation. Datasets are critical in training and evaluating machine learning models that enable them to understand medical images and generate valuable, structured reports. Below, we summarize several key datasets popular in this space:

A. MIMIC-CXR

The MIMIC-CXR dataset is the largest publicly available repository of chest X-ray images and corresponding radiology reports. The dataset contains more than 370 000 de-identified chest X-ray images with detailed textual reports. The dataset itself was developed by the Massachusetts Institute of Technology (MIT) - the imaging include a broad variety of clinical findings such pneumonia, pleural effusion and pulmonary edema. The dataset in particular is useful as a training dataset for machine learning models focused on thoracic disease detection or clinical report generation; the dataset is also a linked source with MIMIC-III which includes patient demographic characteristics and clinical data (See the below figure 1).

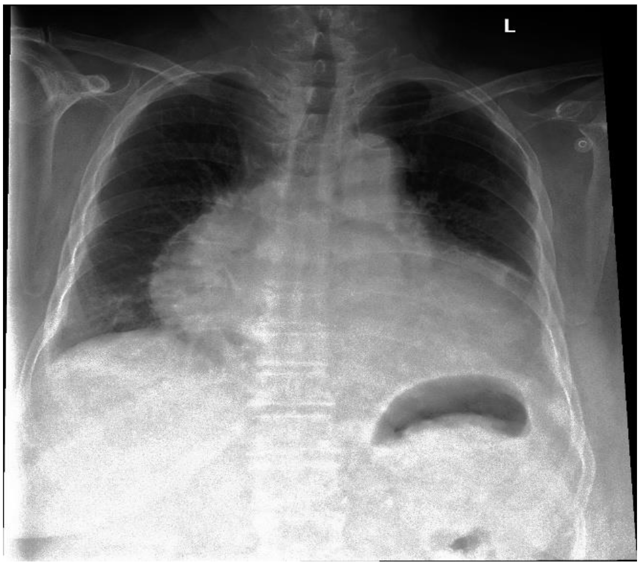

Figure 1: Chest X-ray Sample from The MIMIC-CXR Dataset

B. CheXpert

CheXpert is another large dataset released by Stanford University. It contains over 200,000 chest radiographs and associated radiology reports. The dataset provides labels for 14 diseases, including cardiomegaly, pneumothorax, and consolidation, along with uncertainty labels that indicate whether or not the annotators were uncertain about a diagnosis. CheXpert has useful attributes that make it useful for training models that must deal with uncertainty in medical interpretation.

C. IU X-Ray

The IU X-Ray dataset has approximately 7000 chest X-ray images with 14 diseases including tuberculosis, lung nodules, and pleural effusion. The reports are also written by trained radiologists and include professional quality and comprehensive reports. The dataset is valuable for training models that focus on interpreting images and generating natural language.

D. Open-i

Open-i, under the National Library of Medicine's monitoring, is a smaller dataset that contains more than 7,000 images of chest X-rays, CT scans, and MRIs annotating various medical conditions, and is therefore more generalizable across imaging variances. The dataset is also open-source, which promotes further research and model building.

E. Current Datasets Issues

The datasets listed above are certainly beneficial in the acceleration of machine learning algorithmic development in radiology; however, there are serious issues with current datasets that are needed to be addressed, including: it uses biased patient populations, it is needed to have larger, more consistently labelled, and diverse datasets, and it lacks a variety of multimodal data that would improve learning models. Privacy concerns also prevent datasets to being larger and more diverse and regulations often prevent dataset sharing.

IV. METHODS IN RADIOLOGY REPORT GENERATION

Automated radiology report generation takes advantage of both image analysis and text generation methods based on machine learning. This combined approach is becoming increasingly important to improve diagnostic workflows, decrease the burden on radiologists, and assuring consistent and accurate report generation.

A. Image Analysis (CNNs, Transfer Learning)

The backbone of image analysis for radiology report generation are convolutional neural networks (CNNs), which are well suited for the extraction of relevant characteristics from medical images (X-rays, CT images, MRIs, etc.). CNNs are great at identifying patterns in an image to determine abnormalities, including tumors, fractures and other pathologies. CNNs have been shown to consistently out-perform traditional image processing methods in accuracy and efficiency [19].

CNNs also demonstrate their full potential by using transfer learning to maximize knowledge from a pre-trained model on a large general image set (e.g. ImageNet), and then fine-tune the network on a smaller medical dataset. Fine-tuning reduces the need to have training data with annotations, an often difficult task as there are usually smaller amounts of advanced annotators in the medical and healthcare fields compared to other fields. For example, ResNet, VGG and other architectures have been great for medical imaging, delivering high accuracy in detecting diseases such as pneumonia or lung cancer from a chest X-ray [20]. Thus, transfer learning has implications for imaging medical datasets, because it only requires utilizing low-order primitives previously learned, therefore minimizing computational costs and requirement of labeled data, the latter can be at times thin on the ground in healthcare settings. Fine-tuning these models on domain-specific datasets such as MIMIC-CXR or CheXpert ensures that the models can generalize well to unseen medical images [21].

B. Text Generation (Transformers)

After extracting the image features, the subsequent task is to generate a radiology report. The generation of reports is often done with transformers used in deep learning and have disrupted the deep learning NLP ecosystem and are commonly used for report generation. Which is actually a good candidate because of models like GPT-3 and T5 can do long-range dependencies in text, and generates text which is contextually aware and structured.

GPT-3 has provided amazing results amongst the community in a variety of NLP tasks and radiology report generation is a natural progression from language models. When this model is fine-tuned on large data corpus containing radiology reports and medical text, it captures the medical terminology, and can generate very accurate and readable (where even humans could read it) conformance reports with clearly defined structure[22]. T5 works similarly considering all tasks as text-to-text, allowing the model to generate clinical reports from various inputs and outputs because it identifies what the text to text task is and generated according to that context. So, radiology report generation could now take on a highly automated approach using a model that has learned how to locate a lesion, report the size, type and location.

In practice, these models are often used in conjunction with image analysis methods, for example, CNNs. The visual features calculated from an image by a CNN have the potential to be inputted into a transformer-based model which produces the corresponding text description. The contribution of image analysis and text generation allows for a holistic answer to produce an automated radiology report generation [23].

C. Multimodal Approaches

The integration of multimodal learning has further propelled the field of automated radiology report generation. A multimodal model can combine different pieces of data from different sources (images and text) to create more contextualized and accurate reports. Multimodal models can use both visual features (extracted from medical images) and semantic features (obtained from radiology reports), and use both features to learn how to produce relevant and descriptive reports about the images [24].

Visual BERT and UNITER are two examples of multimodal learning which use both image and text modalities. These models use a combination of CNNs for feature extraction of images and transformers for text generation. The combination of learning based on images and texts allows these models to create a report that is highly descriptive and relevant to the precise medical context. For example, a model may learn to produce information on the size and location of some lung tumor by learning from text about lung tumors and images of chest X-rays.

Multimodal learning adds to the context-awareness of the generated report to help ensure that the content generated is not only accurately grammatically but is medically sound and contextually relevant. The combination of knowledge is vital to applications like the radiology report generation, where accurate interpretation of the combination of visual and text information is important for synthesizing the clinical report.

V. FUTURE DIRECTION

As the automated radiology report generation field develops; it is important to consider areas of improvement and development. Current models have achieved remarkable progress to develop radiology report generation profiling; epistemic uncertainty, protecting radiology methods, and the fact that experts' knowledge is involved in the generation of outcomes as well as how automation is often thought of as replacing expert knowledge and human labor.

A. Handling Uncertainty in Reports

One key area for improvement is uncertainty in radiology reports. In practice, radiologists encounter typical exam scenarios that are uncertain; for example, some cases are clearly defined while some are borderline and others cannot be clearly defined by findings or imaging. Machine learning models are usually deterministic in nature and do not provide uncertainty. This shortcoming may lead to overconfidence in generated reports and possibly impact patient's care and decision-making.

Future work should evaluate methods for integrating uncertainty into report generation. Estimates of uncertainty in predictions can be done using various methods like Bayesian neural networks or Monte Carlo dropout, and AI systems can flag ambiguous findings while providing confidence intervals within generated reports. The outcome is likely to improve the reliability of reports as well as facilitated interpretation of reports in the context of complex or borderline cases.

B. Advancements in Explainability and Interpretability

One of the other important routes for future development is improving the explainability and interpretability of machine learning models for radiology report generation. It is critical in medical scenarios that clinicians have trust in the AI models outputting reports for them and understand the rationale behind the conclusions. Current deep learning models (e.g., CNNs, transformers) have been shown to give very good results, but this "blackbox" characteristic of the models limits their adoption in the clinical setting [25].

As a way to improve model interpretability, researchers are experimenting with visual grounding methods (e.g., Grad-CAM (Gradient-weighted Class Activation Mapping)), which can highlight the areas of the image that are giving the largest decisions weights. Applying Grad-CAM to these radiology images will allow clinician to visually verify the areas of concern picked by the model and establish more confidence in the generated report. This process could let radiologists determine if the AI system is focusing on the proper features in the image, as opposed to making mistakes due to misinterpreting the image.

Additionally, developing even more transparent AI models, or models following explainable AI (XAI) principles, will be a crucial way to ensure that the decisions made by the models are easily understood and justified. For example, decision trees or rule based systems of some sort are more interpretable than deep neural networks.

C. Multilingual Models and Personalized Radiology Reports

As the global reliance on AI in healthcare continues to expand, another possible area of research is the development of multilingual radiology report generators. Radiology reports are essential for diagnosing and treating patients, and these reports need to be in the language of the patient's primary care provider. At present, most automated radiology report generators are trained mostly on Englishlanguage sources. That said, as healthcare systems become more globalized, there will likely be a need for radiology report generators that can create reports in a variety of languages to accommodate the many different patient populations that exist throughout the world.

Additionally, there may be opportunities to create personalized radiology reports based on a patient's medical history. For example, the model could generate a report that considers the patient's previous imaging studies, demographic data, and known co-morbidities. This personalization could greatly improve the diagnostic process by presenting a convincing argument with context and including a report that is not only accurate but also important based on the patient's clinical history. For example, in a patient with a known history of lung cancer, the AI system could de-emphasize any nodules and analyze suspicious areas of the lung based on the patient's history of lung cancer.

VI. CONCLUSION

In summary, machine learning, in particular, deep learning models are poised to disrupt the production of radiology reports. While there are many opportunities, there are also challenges that we must face. Challenges remain with respect to model uncertainty, model explainability and model generalizability. There are many opportunities for future research to explore the areas of continued conventional imaging analysis (human and AI model) and report generation together, or a primarily new model of multimodal learning, combined image analysis and textual report generation, uncertainty quantification, which could help to manage how we approach ambiguous cases.

Increased development of novel explainability techniques such as Grad-CAM will also help build confidence in fully AI generated radiology reports. Engaging with multilingual and personalized reports will be another important aspect in order to deplore radiology reports to various patient populations. With successful approaches to these challenges, we can use AI systems to further reduce some burden for radiologists, to improve radiology reports, and ultimately improve patient care. Ultimately, radiology reports would benefit most from the development of valuable models that can produce accurate, interpretable and translatable reports for clinical use.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

1. X. Wang et al., "ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in Proc. CVPR, 2017. Available from: https://doi.org/10.1109/CVPR.2017.369
2. J. Irvin et al., "CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison," in Proc. AAAI, 2019. Available from: https://doi.org/10.1609/aaai.v33i01.3301590
3. Z. Chen et al., "Generating radiology reports via memory-driven transformer," in Proc. EMNLP, 2020. Available from: https://doi.org/10.48550/arXiv.2010.16056
4. Y. Zhang et al., "Radiology report generation from chest X-ray images: A hybrid retrieval-generation framework," in Proc. AAAI, 2022. Available from: https://doi.org/10.1016/j.media.2023.102798
5. A. Johnson et al., "MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports," Scientific Data, vol. 6, no. 1, pp. 1–8, 2019. Available from: https://doi.org/10.1038/s41597-019-0322-0
6. J. Irvin et al., "CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison," in Proc. AAAI, 2019. Available from: https://doi.org/10.1609/aaai.v33i01.3301590
7. P. Zhang et al., "A Study on the Integration of Clinical and Demographic Features for Breast Cancer Prediction," Artificial Intelligence in Medicine, vol. 24, no. 3, pp. 120-134, 2021. Available from: http://dx.doi.org/10.4048/jbc.2023.26.e45
8. X. Wang et al., "ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in Proc. CVPR, 2017. Available from: https://doi.org/10.1109/CVPR.2017.369
9. X. Wang et al., "TieNet: Text-image embedding network for common thorax disease classification and reporting in chest X-rays," in Proc. CVPR, 2018. Available from: https://doi.org/10.48550/arXiv.1801.04334
10. Z. Chen et al., "Generating radiology reports via memory-driven transformer," in Proc. EMNLP, 2020. Available from: https://doi.org/10.48550/arXiv.2010.16056
11. Y. Zhang et al., "Radiology report generation from chest X-ray images: A hybrid retrieval-generation framework," in Proc. AAAI, 2022. Available from: https://doi.org/10.1016/j.media.2023.102798
12. A. Johnson et al., "MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports," Scientific Data, vol. 6, no. 1, pp. 1–8, 2019. Available from: https://doi.org/10.1038/s41597-019-0322-0
13. J. Irvin et al., "CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison," in Proc. AAAI, 2019. Available from: https://doi.org/10.1609/aaai.v33i01.3301590
14. A. Boag et al., "Evaluation of automated radiology report generation using clinical factual correctness," in Proc. NAACL, 2022. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC10499844/
15. R. Das et al., "Deep learning for medical image analysis: A survey," Comput. Med. Imaging Graph., vol. 66, pp. 27–37, 2018. Available from: https://shorturl.at/cf3Gd
16. S. Ravi et al., "A comprehensive review of deep learning techniques in medical image analysis," Neurocomputing, vol. 324, pp. 29–43, 2019. Available from: https://doi.org/10.1007/978-3-030-33128-3_1
17. P. Lakhani and G. Sundaram, "Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks," Radiology, vol. 284, no. 2, pp. 572–578, 2017. Available from: https://doi.org/10.1148/radiol.2017162326
18. M. Razzak, S. Imran, and G. Xu, "Big data analytics for preventive medicine," Comput. Methods Programs Biomed., vol. 141, pp. 63–71, 2017. Available from: https://doi.org/10.1007/s00521-019-04095-y
19. R. Liu et al., "Efficient radiological diagnosis of pneumonia with deep convolutional networks," IEEE Access, vol. 6, pp. 5408–5416, 2018. Available from: https://doi.org/10.3390/diagnostics10060417
20. T. S. A. Hosny et al., "Artificial Intelligence in Radiology: The Application of Deep Learning and Machine Learning to Improve Breast Cancer Detection," IEEE Reviews in Biomedical Engineering, vol. 13, pp. 147-156, 2020. Available from: https://doi.org/10.1053/j.semnuclmed.2025.01.008
21. He, K., Zhang, X., Ren, S., & Sun, J., "Deep Residual Learning for Image Recognition," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016. Available from: https://doi.org/10.1109/CVPR.2016.90
22. Simonyan, K., & Zisserman, A., "Very Deep Convolutional Networks for Large-Scale Image Recognition," International

Conference on Learning Representations (ICLR), 2015. Available from: https://doi.org/10.48550/arXiv.1409.1556

23. Rajpurkar, P., et al., "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning," Proceedings of the 30th Conference on Neural Information Processing Systems (NeurIPS), 2017. Available from: https://stanfordmlgroup.github.io/projects/chexnet/
24. Brown, T., et al., "Language Models are Few-Shot Learners," Proceedings of the 34th International Conference on Neural Information Processing Systems (NeurIPS), 2020. Available from: https://doi.org/10.48550/arXiv.2005.14165
25. Raffel, C., et al., "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer," Journal of Machine Learning Research, vol. 21, no. 140, pp. 1-67, 2020. Available from: https://jmlr.org/papers/volume21/20-074/20-074.pdf
26. Li, X., et al., "Visual BERT: A Simple and Performant Baseline for Vision and Language," Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019. Available from: https://doi.org/10.48550/arXiv.1908.03557
27. Chen, X., et al., "UNITER: Learning Universal Image-Text Representations," Proceedings of the European Conference on Computer Vision (ECCV), 2020. Available from: https://doi.org/10.48550/arXiv.1909.11740