Jurnal Fisika 15 (2) (2025) 38-50

Jurnal Fisika

https://journal.unnes.ac.id/nju/index.php/jf/index

Advancements and Challenges of Deep Learning in Diagnostic Radiology: A Systematic Literature Review

Rafli Filano, I Gde Eka Dirgayussa, Affan Alfarabi*, Ridho Lailatul Akbar, Hafizah Zakiah

Department of Biomedical Engineering, Faculty of Industrial Technology, Institut Teknologi Sumatera (ITERA), Indonesia

Article Info

Article History:

Received:
18 June 2025

Accepted:
23 October 2025

Published:
12 November 2025

Keywords: Deep Learning, Convolutional Neural Network, Natural Language Processing

Abstract

The rapid integration of Deep Learning (DL) in medical imaging is revolutionizing radiology and addressing critical challenges in diagnostic accuracy and healthcare delivery. In Indonesia and other developing countries, the shortage of radiologists and uneven distribution of healthcare services underline the urgency of exploring DL applications as potential solutions. This study aims to systematically review recent trends, effectiveness, and challenges of DL in diagnostic radiology, as well as to provide insights into its potential adaptation in the Indonesian healthcare system. Using a systematic literature review of peer-reviewed articles (2020–2025) from PubMed, IEEE Xplore, ScienceDirect, and Google Scholar, we identified and synthesized evidence on DL applications across multiple imaging modalities, including CT, MRI, X-ray, and ultrasound. Results show that DL achieves radiologist-level accuracy in tasks such as disease detection, segmentation, and automated report generation, while also improving workflow efficiency and clinical decision-making. However, challenges remain in terms of data availability, model interpretability, ethical issues, and clinical integration. This study provides recommendations for advancing DL adoption in radiology, emphasizing the need for standardized validation, clinician training, and context-specific implementation strategies in Indonesia. The findings highlight both the global and local significance of DL in enhancing healthcare access and equity.

p-ISSN 2088-1509

✉ Correspondence address:
Institut Teknologi Sumatera
E-mail: alfarabiaffan371@gmail.com

INTRODUCTION

The rapid advancement of deep learning technology has significantly transformed various sectors, including healthcare. Deep learning, a branch of machine learning, employs artificial neural networks to analyze vast amounts of data. In recent years, its application in radiology has shown promising results, enabling the generation of more coherent and accurate radiology reports than traditional methods, as highlighted by (Monshi , Poon, & Chung , 2020). Additionally, deep learning has enhanced the quality of medical image analysis, allowing for improved insights from complex datasets, (Li, Jiang , & Zhang Y, 2023) have demonstrated.

Substantial progress has been made toward implementing automated radiology reporting models based on deep learning. This progress is largely attributed to the introduction of large medical text and image datasets, which facilitate the generation of coherent paragraphs that surpass traditional medical image annotation or single-sentence descriptions. This development has garnered significant academic attention as it represents a practical and challenging application that bridges visual medical features with radiologist-generated text. The most common approach has involved utilizing publicly available datasets to develop deep learning models that integrate convolutional neural networks (CNN) for image analysis alongside recurrent neural networks (RNN) for natural language processing (NLP) and natural language generation (NLG).

In Indonesia, the challenges within the field of radiology are notable, characterized by a shortage of radiologists and uneven distribution of healthcare services. The introduction of deep learning technology can address these issues by increasing diagnostic efficiency and accelerating image analysis processes, thereby alleviating the workload on existing radiologists, as discussed by (Nazir, Sarwar , & Saini , 2024). Given the urgency of improving access to and quality of healthcare in Indonesia, deep learning presents a viable solution to enhance diagnostic capabilities and patient care.

However, the implementation of deep learning in radiology faces several challenges, including the scarcity of high-quality data, the complexity of algorithms, and the integration of new technologies into existing clinical practices. Understanding the structures of radiology text and image datasets, as well as improving existing deep learning models and evaluation metrics, are critical challenges that need to be addressed. Ethical concerns regarding privacy and data security also pose significant hurdles. Despite these challenges, the effectiveness of deep learning in enhancing medical diagnosis has been demonstrated across various imaging modalities, such as CT, MRI, and X-ray, where these technologies can identify patterns often missed by human observers, as noted by (Gore, 2019).

In developing countries like Indonesia, the role of deep learning is crucial for improving healthcare systems, particularly in light of limited human resources and the necessity for efficient diagnostic processes. This research aims to explain the current trends and technologies in deep learning within the field of radiology, review the effectiveness and challenges associated with its implementation, and provide insights into the potential adaptation of deep learning in Indonesia. Ultimately, this study seeks to enrich the existing literature on the application of artificial intelligence in healthcare while informing the development and implementation of AI technologies within the radiology system in Indonesia, taking local contexts and challenges into account (Mohseni , et al., 2022)

METHOD

This study employed a systematic literature review approach to analyze the evolution, applications, and challenges of Deep Learning (DL) in diagnostic radiology. The review followed three main stages: identification, screening, and eligibility assessment.

Data Sources and Search Strategy

Relevant articles were retrieved from PubMed, IEEE Xplore, ScienceDirect, and Google Scholar using combinations of keywords such as "deep learning," "radiology," "medical imaging," "CT," "MRI," "X-ray," and "ultrasound." The search was limited to peer-reviewed articles published between 2020 and 2025 in English.

Inclusion and Exclusion Criteria.

Articles were included if they (1) discussed the application of DL in radiology, (2) focused on diagnostic accuracy, clinical implementation, or workflow impact, and (3) provided empirical findings or systematic evaluations. Exclusion criteria included (1) non- peer-reviewed sources, (2) conference abstracts without full papers, (3) articles not directly related to radiology, and (4) studies focusing solely on algorithm development without medical application.

Data Collection and Analysis

After removing duplicates, titles and abstracts were screened for relevance. Full-text articles were then evaluated against the inclusion criteria. Data extracted included study objectives, imaging modality, DL model used, dataset characteristics, evaluation metrics, and main findings.

Data Synthesis

The review results were synthesized using a narrative thematic approach, categorizing findings into emerging trends, dominant DL architectures, clinical applications across imaging modalities, evaluation metrics, and implementation challenges. This synthesis method was chosen over meta-analysis due to the heterogeneity of study designs, datasets, and outcome measures across the included literature..

RESULTS AND DISCUSSION

The rapid advancement of deep learning over recent years has significantly transformed the field of radiology, introducing a wide range of applications that hold the potential to revolutionize diagnostic practices and clinical workflows. Recent studies have demonstrated that deep learning can enhance the accuracy of detection, classification, segmentation, and even automate radiology report generation across various imaging modalities. The integration of these technologies has shown the capability to alleviate radiologists' workload, facilitate the prioritization of emergency cases, and support diagnostic decision-making processes. Despite its promising potential, the adoption of deep learning in radiology continues to face substantial challenges, particularly concerning the availability of representative datasets, model interpretability, ethical considerations, and regulatory constraints. This comprehensive review aims to explore recent trends, prevalent methodologies, clinical implementations, performance evaluations, as well as the future challenges and opportunities in the application of deep learning to radiological imaging.

Emerging Trends in Deep Learning for Medical Imaging

The development of deep learning technologies in recent years has had a transformative impact on the field of radiology. Research shows that radiology, as a highly data-dependent specialty, is conducive to utilizing artificial intelligence-based data processing techniques. Since 2020, a significant increase has occurred in the number of scientific publications relating to deep learning applications in radiology, illustrating the scientific community's growing interest in this technology's potential to improve diagnostic efficiency and clinical outcomes.

Some of the key trends that can be identified are the development of more complex models with better generalization capabilities, an increased focus on model interpretability, and multimodal integration that combines image data with clinical patient information. Data from the Association of University Radiologists Radiology Research Alliance Task Force shows that deep learning has made a significant impact in several areas of radiology, including lesion or disease detection, classification, quantification, and segmentation. These advancements have the potential to reduce radiology interpretation errors, which Langlotz says account for about 4% of all radiology interpretations and contribute up to 10% to patient deaths related to diagnostic errors.

Another prominent trend is the development of deep learning systems for case triage and prioritization. Google, for example, has developed a system that can detect abnormalities in chest X-rays with accuracy equivalent to a professional radiologist. Such deep learning systems are not intended to replace radiologists, but rather to increase their productivity in the face of increased workloads and help address the shortage of radiologists in developing countries including Indonesia.

Automatic Radiology Report Generation (ARRG) is another significant development that uses deep learning to automatically generate radiology reports based on medical images. Systematic research by (Liao, Liu, & pasic, 2023) identified 41 studies focusing on ARRG and 14 commonly used radiology datasets in those studies. The ability of ARRGs to automatically generate diagnostic reports has the potential to reduce the administrative burden on radiologists, allowing them to focus on the more complex aspects of clinical interpretation.

Recent developments have also seen the integration of federated learning techniques to address issues of data privacy and collaboration between institutions. According to (Yang, et al., 2021), this approach enables model training from multi-national data without directly sharing patient data, which is particularly relevant for the development of robust and generalizable models in a global context.

Deep Learning Overview

Working Principle of Deep Learning in Diagnostic Radiology

The working principle of deep learning in diagnostic radiology is briefly illustrated in Figure 1, which involves the following steps:

(a) Image Encoding: A radiological image (e.g., X-ray, CT, MRI) is input into a convolutional neural network (CNN), where the early layers extract basic patterns such as edges and textures, while the deeper layers capture abstract clinical features such as lesions, tumors, or fractures.

(b) Feature Extraction and Pathology Prediction: The extracted feature maps are passed through the core layers of the network, which perform spatial encoding and semantic abstraction. At this stage, multiple regions of interest (ROIs) can be localized, and clinical labels (e.g., pneumonia, pneumothorax) are predicted—often in the form of bounding boxes or segmentation masks—accompanied by confidence scores.

(c) Final Diagnostic Decision: A final prediction is computed for each region based on classification confidence and spatial precision (e.g., bounding box overlap). For overlapping or ambiguous detections, post-processing methods such as non-maximum suppression (NMS) or thresholding are applied. The model then outputs the most probable diagnosis, lesion location, and a confidence score.

Deep Learning Architecture in Medical Imaging Applications

(a) depicts the traditional convolutional neural network (CNN) pipeline, where radiological input data such as X-ray, CT, or MRI scans passes through multiple convolutional and pooling layers. These layers progressively extract low- to high-level features before reaching fully connected layers that output a diagnostic prediction. This architecture is widely used for classification tasks, such as detecting pneumonia or identifying fractures.

Subfigure (b) illustrates a basic multilayer perceptron (MLP), a type of feedforward neural network comprising input, hidden, and output layers. Although MLPs are effective for general-purpose learning tasks, they lack the ability to capture spatial hierarchies in image data, making them less suitable for radiology compared to CNNs.

Subfigure (c) shows the U-Net architecture, which has become a standard model for medical image segmentation. U-Net features an encoder-decoder structure with skip connections that link layers across the two paths. These connections enable the model to retain spatial details and improve segmentation accuracy, making U-Net ideal for tasks such as tumor boundary delineation or lesion localization in CT and MRI scans.

Finally, subfigure (d) represents an advanced architecture incorporating a Feature Pyramid Network (FPN) with U-Net-like skip connections. This structure includes downsampling and upsampling paths, along with concatenation operations, enabling multi-scale feature integration. It is particularly effective for complex tasks that require simultaneous detection of small nodules and larger abnormalities.

Together, these subfigures reflect the progressive sophistication of deep learning models in radiology, evolving from simple classification networks to highly specialized architectures for detection, segmentation, and multi-scale feature analysis.

In 2012, Krizhevsky et al. introduced AlexNet, which marked a major breakthrough in image classification using deep learning. In 2017, CheXNet demonstrated that deep learning models could match or surpass expert radiologists in detecting pneumonia from chest X-rays. Since then, neural network architectures have evolved rapidly and diversified into models for segmentation (e.g., U-Net, nnU-Net), classification (e.g., DenseNet121), and detection.

Today, deep learning is widely used in the medical domain to detect diseases such as lung cancer, brain hemorrhage, breast tumors, and COVID-19 pneumonia. Several studies have shown that radiology-specific modifications to network architectures such as grayscale adaptation, noise handling, and attention mechanisms can significantly enhance performance in clinical settings. A timeline of deep learning evolution in radiology and widely adopted models is illustrated in Figure 1(c), showing the transition from basic CNN classifiers to hybrid architectures that integrate detection, segmentation, and clinical decision support.

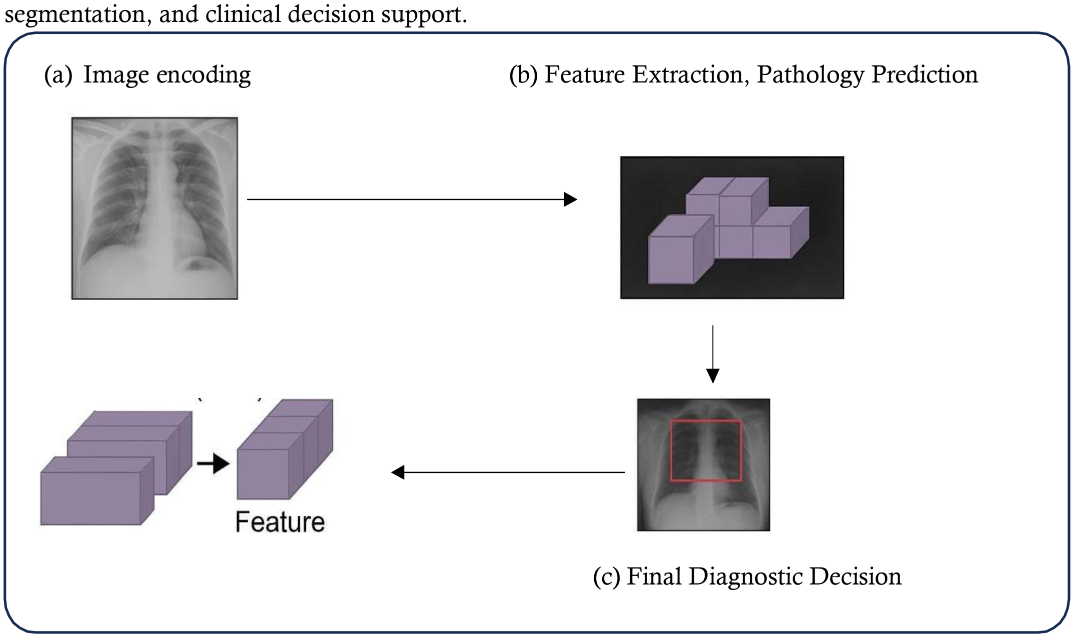
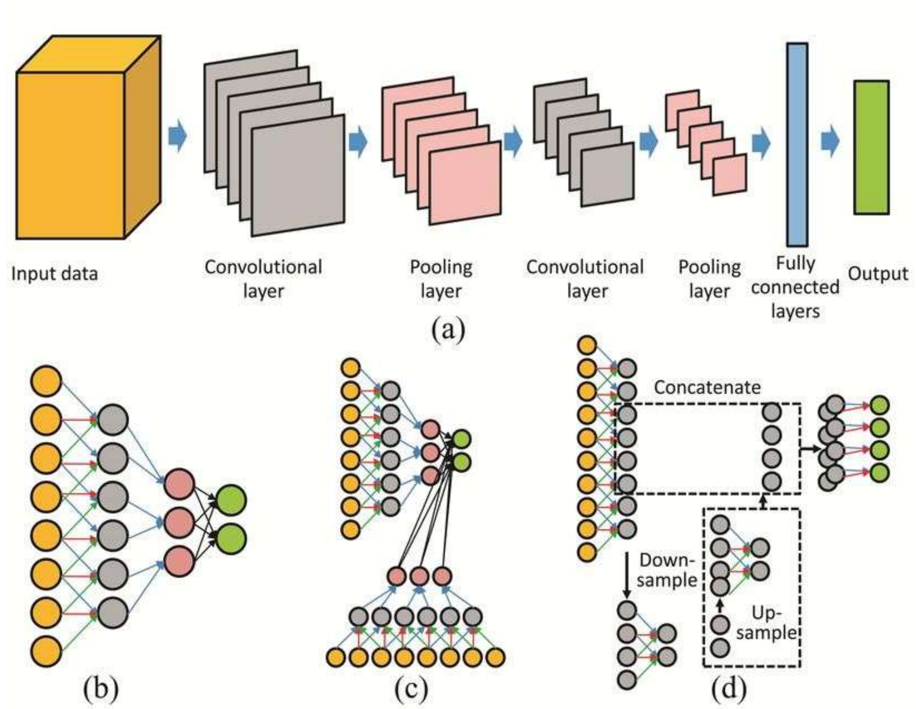


Figure 1. A working principle of deep learning in diagnostic radiology



source:

Figure 2. Deep learning architecture in medical imaging applications

Table 1. Evolution of deep learning architectures in radiology

Year	Model	Contibuted
2012	AlexNet	A breakthrough in image classification using learning. deep
2017 2015	CheXNet U-Net	Demonstrates that deep learning models can match or surpass expert radiologists in detecting pneumonia from chest X-rays. Segmentation models widely used in medical imaging.
2018 2015	U-Net DenseNet121	An automatic segmentation framework that adapts to various medical segmentation tasks. An efficient classification model with dense connections
2015	YOLO	between layers. A real-time object detection model used to detect lesions
2015	Faster R-CNN	in medical images. A two-stage object detection model used for lesion detection

Dominant Deep Learning Architectures in Radiology

In the context of radiology, several deep learning architectures have demonstrated superior performance and wide adoption. Convolutional Neural Networks (CNN) remain the backbone of most medical imaging applications due to their excellent capabilities in visual feature extraction. CNNs have proven to be highly effective in the processing of structured data such as medical images, with a wide variety of architectures developed to meet the specific needs of radiology applications.

The ResNet (Residual Networks) architecture has become a popular choice due to its ability to overcome the vanishing gradient problem, allowing for deeper and more complex network training. (Zhou, et al., 2021), showed that ResNet-50 and ResNet-101 have been effectively used for lung nodule detection on CT scans with sensitivity reaching 94% and specificity 91%, making them a promising choice for clinical applications that require high accuracy.

EfficientNet, which uses a systematic scaling method to balance the depth, width, and resolution of the network, has also shown promising results in radiology applications. A study by (Kocoń, et al., 2024) demonstrated that EfficientNet-B4 achieved 96% accuracy in pneumonia classification on chest X-ray images, outperforming conventional CNN architectures with a smaller number of parameters, making it more efficient for implementation in environments with limited computing resources.

Generative Adversarial Networks (GANs) have gained particular attention in radiology for applications such as image quality enhancement, image synthesis, and data augmentation (Zhu & Rosen, 2020) used CycleGANs to convert T1 to T2 MRI images with quality equivalent to the original acquisition, potentially reducing scan time and improving patient comfort without sacrificing diagnostic information.

In recent years, Transformer-based architectures, originally developed for natural language processing, have been adapted for medical imaging applications. Vision Transformer (ViT) and its variants such as Swin Transformer have shown promising performance in radiology image classification and segmentation tasks. (Secinaro et al., 2021) reported that the Swin Transformer achieved a Dice score of 0.89 in brain tumor segmentation in MRI images, outperforming U-Net which was the previous industry standard.

Ensemble approaches that combine multiple architectures are also gaining popularity, where the strengths of different models are combined to improve accuracy and robustness. demonstrated that an ensemble approach combining DenseNet, ResNet, and EfficientNet achieved a 3.2% improvement in accuracy compared to the single best model in COVID-19 classification from chest CT images, demonstrating the potential of collaborative approaches in improving diagnostic performance.

Implementation of Deep Learning on Different Imaging Modalities

CT Scan

Computed tomography (CT) has been one of the imaging modalities that has benefited the most from advances in deep learning. Prominent applications include detection and characterization of pulmonary nodules, classification of liver lesions, and identification of intracranial hemorrhage. (Langlotz, et al., 2020) reported a deep learning system capable of detecting intracranial hemorrhage on head CT with 98% sensitivity and 95% specificity, enabling rapid triage of patients requiring immediate intervention.

During the COVID-19 pandemic, deep learning for chest CT analysis became the focus of intensive research. A multicenter study by (Yang, et al., 2021) developed a ResNet-based model that achieved 89% accuracy in identifying COVID-19 pneumonia, aiding rapid diagnosis and patient management during a global health crisis. The study also demonstrated the success of a federated learning approach that enabled model training using data from China, Italy, and Japan without direct patient data transfer between institutions.

Segmentation of organs and anatomical structures on CT has also shown significant progress. U- Net and its variants have been used for automated segmentation of abdominal organs with Dice coefficient reaching 0.95 for liver and 0.90 for kidney, greatly assisting in surgical and radiotherapy planning as described by (Willemink, et al., 2020) This automatic segmentation capability not only improves workflow efficiency but also reduces inter-observer variability that often occurs in manual measurements.

MRI

Magnetic resonance imaging (MRI) presents unique challenges for deep learning analysis due to acquisition variability and signal complexity. However, significant progress has been made in brain structure segmentation, tumor detection and classification, and neurodegenerative disease assessment.(Y. Zhang et al., 2023) developed a 3D CNN architecture for brain tumor segmentation on multimodal MRI images, achieving 94% accuracy in distinguishing high-grade from low-grade glioma tumors.

The use of GANs for MRI image super-resolution has also shown promising results. (Y. Zhang et al., 2023) demonstrated that GANs can increase the spatial resolution of brain MRI images by a factor of 4x, potentially reducing acquisition time without compromising diagnostic quality. This approach is particularly relevant in a clinical context where shorter scan times can increase patient throughput and reduce motion artifacts.

Transfer learning has proven to be very effective in MRI applications, especially when training data is limited. (Lundervold & Lundervold, 2019) showed that a model trained first on a large dataset of public MRI images and then customized for a specific task can achieve comparable performance to a model trained from scratch, despite using only 30% of the training data. This approach is particularly valuable in the Indonesian context where large, well-labeled MRI datasets may not always be available.

X-Ray

Conventional radiography (X-ray) is still the most commonly used medical imaging modality worldwide, and has become a major focus of deep learning applications. The system developed by Google is capable of detecting abnormalities in chest X-rays with accuracy equivalent to that of a professional radiologist. It can serve as a triage tool, helping radiologists prioritize urgent and potentially life-saving cases in settings where access to radiologists is limited.

CheXNet, a DenseNet-121-based deep learning model developed by (Secinaro et al., 2021) achieved performance exceeding that of radiologists in detecting 14 different pathologies on chest X-rays, with an average Area Under the Curve (AUC) of 0.84. Follow-up research has improved this model to provide pathology localization and uncertainty estimation, increasing interpretability and clinical confidence in the system-generated results.

Deep learning has also been applied to musculoskeletal radiographs, with notable applications in fracture detection and osteoarthritis assessment. (McBee, et al., 2018) used EfficientNet to classify wrist fractures on radiographs with 90% sensitivity and 88% specificity, potentially reducing diagnostic errors in busy emergency departments. This application has particular relevance in Indonesia where the availability of musculoskeletal specialist radiologists may be limited in remote areas

Ultrasound

Ultrasound, although highly operator-dependent and subject to high variability, has also benefited from deep learning applications. Research has focused mainly on image quality improvement, structure segmentation, and lesion classification. (Prevedello, et al., 2020) developed a model capable of detecting and classifying thyroid nodules on ultrasound with 95% sensitivity and 90% specificity, aiding decision-making on the need for biopsy.

For fetal ultrasound, deep learning has been applied for automated biometry and congenital anomaly detection. (Yu, Mohajer, Eng, & Wispelwey, 2022) developed a CNN-based model that can automatically measure standardized fetal biometric parameters with an average difference of less than 2mm compared to manual measurements by experienced sonographers. In Indonesia, where access to perinatology specialists may be limited in certain areas, this kind of application has the potential to bridge the gap in maternal care.

The application of real-time deep learning on ultrasound is also emerging as a promising research area. (Zhou, et al., 2021) developed a system capable of providing guidance during ultrasound interventional procedures, improving needle placement accuracy and potentially improving clinical outcomes for ultrasound-guided biopsy and drainage procedures.

Performance Evaluation and Success Metrics

Performance evaluation of deep learning models in radiology requires consideration of various metrics that reflect both technical accuracy and clinical utility. While conventional metrics such as accuracy, sensitivity, and specificity remain essential, they must be interpreted in the context of disease prevalence and the clinical consequences of diagnostic errors.

The Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curve has emerged as a standard assessment metric for deep learning models in radiology. A systematic review by (Liao, Liu, & pasic, 2023) identified six commonly used clinical effectiveness metrics for evaluating Automatic Radiology Report Generation (ARRG) systems, underscoring the need to assess models not only from a technological standpoint but also in terms of clinical impact.

These metrics include clinical accuracy, alignment with reference reports, information completeness, and terminology precision.

Comparison with human radiologist performance is another critical aspect of evaluation. (McBee, et al., 2018) reported that a deep learning model for rib fracture detection in chest X-rays achieved a sensitivity of 90%, compared to 75% for general radiologists, albeit with slightly lower specificity (92% vs. 96%). These findings highlight AI's potential as a decision-support tool that can enhance diagnostic accuracy, particularly for commonly missed cases.

Prospective evaluation in real-world clinical settings is increasingly recognized as the gold standard for assessing the practical utility of deep learning models. (Langlotz, et al., 2020) implemented a deep learning system for head CT triage across multiple hospitals and demonstrated an average reduction in interpretation time by 20 minutes for intracranial hemorrhage cases, potentially improving clinical outcomes through earlier detection and intervention. Patient- centered metrics, such as time to diagnosis, waiting periods, and long-term clinical outcomes, are also gaining importance in radiology AI evaluation. (Prevedello, et al., 2020) found that deploying an AI-based triage system for chest X-rays reduced average wait times by 31% for patients with critical findings, emphasizing its potential benefits for clinical workflow optimization and enhanced patient care quality.

Challenges and Limitations of Implementation

A major challenge in developing deep learning models for radiology is the scarcity of high-quality, accurately labeled datasets. Data imbalance between normal and abnormal cases—a significant issue identified in (Liao, Liu, & pasic, 2023) systematic review can introduce prediction bias and reduce model generalizability across diverse patient populations. Radiology data heterogeneity further complicates model development, with variations in imaging protocols, equipment vendors, and acquisition parameters potentially degrading performance. (Willemink, et al., 2020) reported a 12% accuracy drop when models trained on data from one institution were applied to another without domain adaptation, underscoring the need for techniques to mitigate cross-institutional variability and enhance generalizability. Data privacy and ethical concerns surrounding patient data usage for model training pose additional challenges, particularly in Indonesia, where patient data protection regulations are still evolving. (Yang, et al., 2021) demonstrated federated learning as a potential solution, enabling model training without direct data sharing, albeit with trade-offs in computational efficiency and implementation complexity.

The "black-box" nature of many deep learning models remains a barrier to clinical adoption. Radiologists and clinicians require transparency in model decision-making to build trust and facilitate collaborative diagnosis. Interpretability-enhancing methods—such as class activation maps and gradient-based feature attribution—have been extensively studied. (Kocoń, et al., 2024) showed that providing visualizations of model-predicted salient regions increased radiologists' trust in model outputs by 28%. Such approaches are crucial for clinical adoption in Indonesia, where clinicians may hesitate to rely on poorly understood technologies. Advanced explainable AI (XAI) techniques, including concept-based explanations and model-agnostic interpretability methods, have also been tailored for radiology.(Secinaro et al., 2021) proposed a framework for generating explanations in standardized radiological terminology, potentially improving clinical workflow integration and healthcare professional acceptance.

The adoption of deep learning in radiology faces complex ethical and regulatory hurdles. Legal accountability for AI-assisted diagnostic errors remains unresolved, while Indonesia's Ministry of Health and regulatory bodies work to balance innovation with patient safety. Algorithmic bias is another ethical concern. (Yu, Mohajer, Eng, & Wispelwey, 2022) found that models trained primarily on data from specific populations underperform on underrepresented demographics, potentially exacerbating healthcare disparities. This has critical implications for Indonesia, where ethnic and geographic diversity must be considered during algorithm development and validation. Transparency in model development—including detailed methodology disclosure and performance evaluation results—is increasingly emphasized by regulators and professional organizations as a prerequisite for responsible clinical implementation. (McBee, et al., 2018) highlighted the need for consistent reporting standards to enable comparative algorithm assessments and inform regulatory decisions.

Integrating deep learning systems into existing radiology workflows presents significant technical and logistical challenges. Interoperability with Picture Archiving and Communication Systems (PACS) and Radiology Information Systems (RIS) requires standardized communication protocols and well-designed interfaces. End-user acceptance among radiologists and radiology technologists is another critical factor for successful adoption. (Zhou, et al., 2021) identified key determinants, including perceived usefulness, ease of use, and alignment with existing workflows. In Indonesia, where access to technology and training may be uneven, these factors must be prioritized during implementation planning. Economic considerations, such as implementation costs and return on investment (ROI), also influence adoption rates. (Prevedello, et al., 2020) demonstrated that while radiology AI systems require substantial upfront investment, they may yield long-term cost savings through improved productivity and reduced diagnostic errors. For Indonesia's resource-constrained healthcare system, comprehensive health economic analyses are essential to guide adoption decisions.

CONCLUSION

The application of deep learning in radiology has led to significant advancements across various imaging modalities such as CT scans, MRI, X-rays, and ultrasound (van Leeuwen et al., 2021). Numerous studies have demonstrated that deep learning can enhance diagnostic accuracy, expedite triage processes, and reduce inter-observer variability in anatomical segmentation (Brossard et al., 2021). Innovations such as federated learning and transfer learning have further expanded opportunities for optimizing model performance, particularly in settings with limited data availability, as is often the case in Indonesia (Montoro et al., 2023). Overall, deep learning presents a substantial potential to improve healthcare efficiency and support clinical decision-making (Lo Gullo et al., 2024).

Nonetheless, evaluating the performance of deep learning models in radiology cannot rely solely on technical metrics (Khafaji et al., 2022). It is essential to assess real-world clinical impacts, such as time to diagnosis, workflow efficiency, and long-term patient outcomes (Q. Zhang et al., 2024). Prospective studies conducted in actual clinical environments have proven crucial in validating the practical benefits of these technologies (Fan et al., 2024). Additionally, direct comparisons between AI performance and human radiologists underscore the role of AI as a supportive tool, intended to augment rather than replace human expertise in medical imaging (Miró-Nicolau et al., 2022).

However, significant challenges remain in the implementation of deep learning in radiology, including data quality issues, variability across imaging sources, and ethical concerns regarding privacy and algorithmic transparency (Fritz & Fritz, 2022). Efforts to enhance model interpretability, such as using visualization techniques and concept-based explanations, are vital for fostering clinician trust. Furthermore, careful attention must be paid to potential algorithmic biases, which could exacerbate healthcare disparities, especially in diverse populations like those in Indonesia (Adams et al., 2021).

The successful adoption of deep learning systems in radiology hinges not only on technical excellence but also on the readiness of healthcare ecosystems to integrate these technologies (Montoro et al., 2023). Critical factors such as system interoperability, user acceptance, availability of adequate training, and thorough cost-benefit analyses must be considered during implementation planning. With a cautious and systematic approach, deep learning has the potential to become a transformative force in expanding access to and enhancing the quality of radiological services across Indonesia (Saw & Ng, 2022).

ACKNOWLEDGEMENT

We extend our deepest appreciation to all lecturers of the Biomedical Imaging course and fellow students who participated in the development of this journal. Support, shared knowledge, and meaningful collaboration throughout the process have greatly enriched the learning experience. Constructive input and active engagement have played a vital role in shaping the depth and direction of this study.

REFERENCE

Adams, S. J., Henderson, R. D. E., Yi, X., & Babyn, P. (2021). Artificial Intelligence Solutions for Analysis of X-ray Images. In Canadian Association of Radiologists Journal (Vol. 72, Issue 1). https://doi.org/10.1177/0846537120941671

Brossard, C., Lemasson, B., Attyé, A., de Busschère, J. A., Payen, J. F., Barbier, E. L., Grèze, J., & Bouzat, P. (2021). Contribution of CT-Scan Analysis by Artificial Intelligence to the Clinical Care of TBI Patients. In Frontiers in Neurology (Vol. 12). https://doi.org/10.3389/fneur.2021.666875

Fan, H., Luo, Y., Gu, F., Tian, B., Xiong, Y., Wu, G., Nie, X., Yu, J., Tong, J., & Liao, X. (2024). Artificial intelligence-based MRI radiomics and radiogenomics in glioma. In Cancer Imaging (Vol. 24, Issue 1). https://doi.org/10.1186/s40644-024-00682-y

Fritz, B., & Fritz, J. (2022). Artificial intelligence for MRI diagnosis of joints: a scoping review of the current state-of-the-art of deep learning-based approaches. In Skeletal Radiology (Vol. 51, Issue 2). https://doi.org/10.1007/s00256-021-03830-8

Khafaji, M. A., Safhi, M. A., Albadawi, R. H., Al-Amoudi, S. O., Shehata, S. S., & Toonsi, F. (2022). Artificial intelligence in radiology Are Saudi residents ready, prepared, and knowledgeable? Saudi Medical Journal, 43(1). https://doi.org/10.15537/SMJ.2022.43.1.20210337

Lo Gullo, R., Marcus, E., Huayanay, J., Eskreis-Winkler, S., Thakur, S., Teuwen, J., & Pinker, K. (2024). Artificial Intelligence-Enhanced Breast MRI: Applications in Breast Cancer Primary

Treatment Response Assessment and Prediction. In Investigative Radiology (Vol. 59, Issue 3). https://doi.org/10.1097/RLI.0000000000001010

Lundervold, A. S., & Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on MRI. In Zeitschrift fur Medizinische Physik (Vol. 29, Issue 2). https://doi.org/10.1016/j.zemedi.2018.11.002

Miró-Nicolau, M., Moyà-Alcover, G., & Jaume-I-capó, A. (2022). Evaluating Explainable Artificial Intelligence for X-ray Image Analysis. Applied Sciences (Switzerland), 12(9). https://doi.org/10.3390/app12094459

Montoro, F. S., Gordo, M. L. P., Tascón, Á. D., de Gracia, M. M., Velez, S. O., Fernández, S. F., Vallano, R. G., & Acosta Velásquez, K. S. (2023). Quantification of pulmonary opacities using artificial intelligence in chest CT scans during SARS-CoV-2 pandemic: validation and prognostic assessment. Egyptian Journal of Radiology and Nuclear Medicine, 54(1). https://doi.org/10.1186/s43055-023-01105-x

Saw, S. N., & Ng, K. H. (2022). Current challenges of implementing artificial intelligence in medical imaging. In Physica Medica (Vol. 100). https://doi.org/10.1016/j.ejmp.2022.06.003

Secinaro, S., Calandra, D., Secinaro, A., Muthurangu, V., & Biancone, P. (2021). The role of artificial intelligence in healthcare: a structured literature review. BMC Medical Informatics and Decision Making, 21(1). https://doi.org/10.1186/s12911-021-01488-9

van Leeuwen, K. G., Schalekamp, S., Rutten, M. J. C. M., van Ginneken, B., & de Rooij, M. (2021). Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. European Radiology, 31(6). https://doi.org/10.1007/s00330-021-07892- z

Zhang, Q., Hu, Y., Zhou, C., Zhao, Y., Zhang, N., Zhou, Y., Yang, Y., Zheng, H., Fan, W., Liang, D., & Hu, Z. (2024). Reducing pediatric total-body PET/CT imaging scan time with multimodal artificial intelligence technology. EJNMMI Physics, 11(1). https://doi.org/10.1186/s40658-023-00605-z

Zhang, Y., Liu, Y. L., Nie, K., Zhou, J., Chen, Z., Chen, J. H., Wang, X., Kim, B., Parajuli, R., Mehta, R. S., Wang, M., & Su, M. Y. (2023). Deep Learning-based Automatic Diagnosis of Breast Cancer on MRI Using Mask R-CNN for Detection Followed by ResNet50 for Classification. Academic Radiology, 30. https://doi.org/10.1016/j.acra.2022.12.038

Calandra, D., & Muthurangu, V. (2021). The role of artificial intelligence in healthcare: a structured literature review. BMC medical informatics and decision making, 1-23.

Gore, J. C. (2019). Artificial intelligence in medical imaging. Magnetic Resonance Imaging, 68. Kocoń, J., Ekbal, A., Mnasri, M., Kaur, D., Goyal, P., Prasad, T., & Yang, B. (2024). Deep Learning for Healthcare: A Comprehensive Analysis and Categorization. ACM Computing Surveys, 1-35.

Langlotz, C., Allen, B., Erickson, B., Kalpathy-Cramer, J., Bigelow, K., Cook, T., & Channin, D. (2020). A roadmap for foundational research on artificial intelligence in medical imaging: from the 2018 NIH/RSNA/ACR/The Academy Workshop. Radiology, 373-379.

Li, M., Jiang , Y., & Zhang Y. (2023). Medical image analysis using deep learning algorithms. Front Public Health.

Liao, Y., Liu, H., & pasic, I. (2023). Deep learning approaches to automatic radiology report generation: A systematic review. Informatics in Medicine Unlocked, 39.

McBee, M. P., Awan, O., Colucci, A., Ghobadi, C., Kadom, N., Kansagra, A., & Auffermann, W. (2018). Deep learning in radiology. Academic radiology, 1472-1480.

Mohseni , A., Ghotbi , E., Kazemi, F., Shababi, A., Chashm, J. S., & Shababi, N. (2022). Artificial Intelligence in Radiology: What Is Its True Role at Present, and Where Is the Evidence? Radiol Clin North Am, 935-947.

Monshi , M., Poon, J., & Chung , V. (2020). Deep learning in generating radiology reports: A survey. Artif Intell Med.

Nazir, N., Sarwar , A., & Saini , S. B. (2024). Recent developments in denoising medical images using deep learning: An overview of models, techniques, and challenges. Micron, 180.

Prevedello, L. M., Halabi, S. S., Shih, G., Wu, C., Kohli, M. D., & Chokshi, F. H. (2020). Challenges related to artificial intelligence research in medical imaging and the importance of image analysis competitions. Radiology: Artificial Intelligence.

Willemink, M., Koszek, W., Hardell, C., Fleischmann, D., Harvey, H., & Lungren, M. (2020). Preparing medical imaging data for machine learning. Radiology, 4-15.

Yang, D., Xu, Z., Li, W., Myronenko, A., Roth , H., & R, H. (2021). Federated semi-supervised learning for COVID region segmentation in chest CT using multi-national data from China, Italy, Japan. Medical image analysis, 70.

Yu, A., Mohajer, B., Eng, J., & Wispelwey, B. (2022). Equity in machine learning algorithms for healthcare. New England Journal of Medicine,, 479-484.

Zhou, S. K., Greenspan, H., Davatzikos, C., Duncan, J. s., Van Ginneken, B., Madabhushi, A., & Rueckert, D. (2021). A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises. Proceedings of the IEEE, , 820-838.

Zhu, B., & Rosen, B. (2020). mage reconstruction by domain-transform manifold learning. Nature, 487-492.