Balkan Journal of Applied Mathematics and Informatics
Volume ?? Number ?
Year 201?
Print ISSN ???
Online ISSN ???
UDK: ???

Artificial Intelligence in Medicine: Enhancing the Analysis of Radiographic Images

Rexhep Mustafovski[1], Besnik Qehaja[2] and Shejnaze Gagica[3]

Abstract. The integration of artificial intelligence (AI) into medical imaging has revolutionized diagnostic radiology, particularly in the analysis of radiographic (X-ray) images. This paper reviews recent advances in AI-based systems for X-ray interpretation, examining their technical implementation, clinical applications, and impact on diagnostic accuracy. Deep learning algorithms, primarily Convolutional Neural Networks (CNNs), have demonstrated diagnostic accuracy comparable to or exceeding that of experienced radiologists in detecting various pathologies including pneumonia, pulmonary nodules, consolidation, pneumothorax, and fractures. AI systems improve physician sensitivity from 75.7% to 85.6% when detecting abnormalities, reduce interpretation time by up to 27%, and provide consistent, objective analysis unaffected by fatigue or cognitive bias. However, challenges remain regarding validation across diverse populations, integration into clinical workflows, and regulatory oversight. This review synthesizes current evidence from 2021-2025, highlighting the transformative potential of AI as a decision-support tool that augments rather than replaces radiologist expertise. Future directions include hybrid human-AI systems, multimodal integration of clinical data, and development of transparent, interpretable models for clinical deployment. [1]

1. Introduction

Medical imaging represents a cornerstone of modern diagnostic medicine, with radiographic X-ray examinations being among the most frequently performed diagnostic procedures worldwide [1]. Radiologists interpret millions of chest X-rays annually, yet cognitive limitations, fatigue, and variable workload pressures create diagnostic bottlenecks that can result in missed or delayed diagnoses [2]. The global shortage of radiologists, particularly in developing regions, further exacerbates these challenges, with some nations experiencing critical workforce deficits [3].

Artificial intelligence, particularly through deep learning technologies, has emerged as a transformative approach to enhance diagnostic accuracy and operational efficiency in radiology [4]. Unlike traditional technological advances that focused on improving image acquisition, AI directly addresses the cognitive and systemic limitations inherent in human interpretation [5]. Deep learning models, when appropriately trained and validated, demonstrate remarkable capability to recognize visual patterns in radiographic images that correspond to specific pathologies [6].

The integration of AI-driven Computer-Aided Detection (CAD) systems into clinical workflows represents a paradigm shift in radiological practice. Recent studies demonstrate that AI assistance significantly reduces missed abnormalities, with some

---

Date: January 24 2026.
Keywords. artificial intelligence, deep learning, X-ray analysis, medical imaging, convolutional neural networks

FDA-cleared systems achieving sensitivity improvements exceeding 40% relative reduction in missed findings [7]. Importantly, AI systems provide consistent, tireless analysis across 24-hour operational periods, unaffected by fatigue, attention lapses, or cognitive bias that characterize human performance [8].

This paper reviews the current state of AI applications in radiographic X-ray analysis, examining technical methodologies, clinical validation studies, diagnostic performance metrics, and integration strategies for clinical deployment. Evidence synthesized from 2021-2025 demonstrates that AI systems have progressed from research prototypes to FDA-approved clinical tools actively deployed in diagnostic workflows.

2. Technical Foundations: Deep Learning in Radiographic Analysis

The rapid progress of artificial intelligence in radiographic image analysis is primarily driven by advances in deep learning methodologies, which enable automated extraction and interpretation of complex visual patterns from medical images. Unlike traditional computer-aided diagnostic approaches based on handcrafted features, deep learning models learn hierarchical representations directly from pixel-level data, allowing them to capture subtle anatomical and pathological variations that may be imperceptible to human observers. In radiographic imaging, these methods have demonstrated strong capability in modelling spatial relationships, texture patterns, and structural abnormalities across diverse clinical conditions.

This section outlines the core technical foundations underlying AI-based radiographic analysis, with emphasis on convolutional neural network architectures, training strategies based on transfer learning, and standardized performance evaluation metrics. Understanding these foundations is essential for interpreting reported diagnostic results, assessing model reliability, and evaluating the feasibility of deploying deep learning systems in real-world clinical environments.

2.1. Convolutional Neural Networks (CNNs) for Image Analysis

The fundamental technology enabling AI-based X-ray analysis is the Convolutional Neural Network (CNN), a specialized deep learning architecture designed to extract hierarchical features from images [9]. CNNs consist of multiple layers—convolutional layers that detect local features, pooling layers that reduce dimensionality, and fully connected layers that produce classifications [10].

The architectural depth and design of CNNs significantly influence diagnostic performance. Notable architectures employed in radiographic analysis include:

ResNet50V2 - A 50-layer residual network incorporating skip connections that enable training of very deep networks. In chest X-ray analysis for pneumonia detection, ResNet50V2 achieves approximately 91% accuracy [11].

VGG-16 - A 16-layer architecture renowned for robust feature extraction. VGG-16 demonstrates training accuracy approaching 96% with validation accuracy stabilizing around 89% in thoracic disease classification tasks [12].

MobileNetV2 - A lightweight architecture optimized for computational efficiency while maintaining high accuracy. Recent studies report MobileNetV2 achieving 92% accuracy in binary classification of pneumonia versus normal chest X-rays, demonstrating superior performance compared to deeper architectures in resource-constrained settings [13].

2.2. Transfer Learning and Model Training Strategies

Transfer learning, wherein pre-trained models developed on large datasets are fine-tuned for specific clinical tasks, has emerged as the dominant approach in medical imaging AI [14]. Rather than training neural networks from random initialization on limited clinical datasets, researchers leverage models pre-trained on ImageNet (1.2 million natural images) and adapt these representations for radiographic interpretation [15].

The NIH Chest X-ray 14 dataset, containing 112,120 labelled frontal-view chest radiographs with up to 14 disease labels, has become foundational for training deep learning models [16]. Transfer learning with this dataset has enabled development of high-performance algorithms even when augmented with small institutional datasets of 800-5,000 images [17].

2.3. Strategies Model Performance Metrics

Diagnostic performance of AI systems is quantified through standard metrics including sensitivity (true positive rate), specificity (true negative rate), and Area Under the Receiver Operating Characteristic Curve (AUROC) [18]. FDA-cleared systems demonstrate metrics comparable to board-certified radiologists:

Sensitivity and Specificity: AI systems for chest X-ray abnormality detection achieve sensitivity of 98.7% and specificity of 88.5%, with negative predictive value exceeding 99.6% [4]. For specific pathologies, performance varies - module detection achieves 96% accuracy, pneumothorax 84% sensitivity, and consolidation detection 98% accuracy [19].

AUROC Values: Advanced CNN architectures achieve AUROC values approaching 0.98 for specific conditions, matching or exceeding expert radiologist performance in blinded comparisons [20].

3. Clinical Applications and Diagnostic Performance

The clinical value of artificial intelligence in radiographic imaging is defined by its ability to deliver reliable diagnostic performance across practical medical applications. Recent deep learning–based systems have demonstrated effectiveness in detecting a wide range of radiographic pathologies, supporting clinicians in routine diagnostic workflows.

This section examines key clinical use cases of AI-assisted X-ray analysis and evaluates reported diagnostic performance in relation to established clinical standards, with emphasis on accuracy, sensitivity, and clinical relevance.

3.1. Pulmonary Pathology Detection

Pneumonia Detection: Deep learning systems excel at pneumonia detection, with studies demonstrating 92% accuracy in binary classification of pneumonia versus normal radiographs [21]. CNN-based approaches exceed traditional machine learning methods relying on manual feature extraction, representing fundamental advancement in automated image analysis.

Pulmonary Nodules and Lung Cancer: Pulmonary nodule detection represents a critical application where AI augments radiologist performance. Studies demonstrate AI systems achieve 96% accuracy in nodule detection, with particular advantage in identifying small or subtle nodules that represent early-stage malignancy [22]. The 24% rate of missed or incorrectly-interpreted subtle findings in unaided radiographic interpretation indicates substantial room for AI-assisted improvement [23].

COVID-19 Pneumonia: During the SARS-CoV-2 pandemic, AI systems trained to distinguish COVID-19 pneumonia from other pneumonias on chest X-rays demonstrated diagnostic utility, with multiple published algorithms achieving >90% accuracy [24].

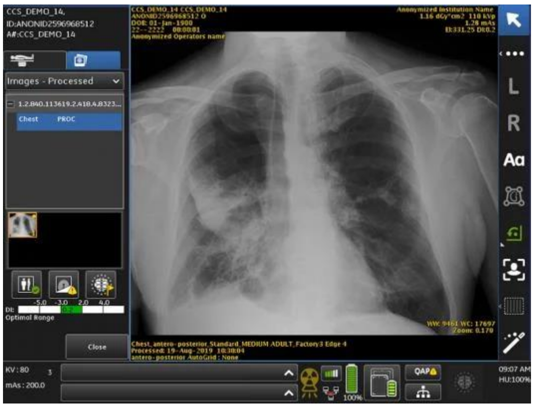

Figure 1. *Representative chest X-ray image illustrating pulmonary abnormalities used in AI-based analysis [Image courtesy of GE Healthcare]*

This figure presents a representative chest X-ray image demonstrating pulmonary abnormalities commonly targeted by AI-based diagnostic systems. Such images serve as input data for deep learning models, which analyse radiographic patterns to support detection of conditions such as pneumonia, consolidation, and other pulmonary pathologies. The example illustrates the type of radiographic features evaluated during automated pulmonary pathology detection.

3.2. Skeletal Pathology Detection

Fracture Identification: Rib fractures and other skeletal injuries are commonly missed on initial interpretation, with miss rates approaching 24% in high-volume trauma settings [25]. AI systems can analyse X-rays for fracture signs within seconds, automatically flagging potential cases for clinician review. This capability significantly accelerates fracture detection and reduces interpretation time, facilitating faster clinical decision-making in trauma centres [4].

3.3. Cardiac and Mediastinal Pathology

AI-assisted detection systems categorize findings according to the Radiological Society of North America (RSNA) reporting guidelines, with distinct categories for Cardiac, Mediastinum/Hila, Lungs, Pleura, Bones, Soft Tissues, Hardware, and Other abnormalities [7]. Performance metrics indicate comparable sensitivity across anatomical categories, demonstrating broad applicability to diverse pathologies.

4. Clinical Impact: Physician Performance Enhancement and Workflow Optimization

The integration of artificial intelligence into radiographic interpretation has implications not only for diagnostic accuracy but also for physician performance and clinical workflow efficiency. AI-assisted systems are increasingly evaluated as decision-support tools that influence reading accuracy, interpretation speed, and consistency across varying levels of clinical expertise. This section reviews evidence on how AI deployment affects physician performance and optimizes radiological workflows in routine clinical practice.

4.1. Augmenting Diagnostic Accuracy

Recent clinical validation studies provide compelling evidence that AI assistance substantially improves radiologist and non-radiologist physician performance. In a multi-centre study of FDA-cleared AI systems, physicians demonstrated significant performance improvements when aided by AI:

Sensitivity Enhancement: Physician sensitivity increased from 75.7% (95% CI: 0.750-0.764) when unaided to 85.6% (95% CI: 0.850-0.862) when aided by AI, representing a relative reduction of 40.74% in missed abnormalities [7].

Specificity Maintenance: Critically, physician specificity increased from 84.3% to 87.0% when aided, demonstrating that AI assistance does not result in false-positive overcalling but instead supports accurate identification of cases without significant abnormalities [7].

Addressing Diagnostic Disparities: AI systems demonstrated particular utility in improving accuracy of non-radiologist physicians and junior radiologists, effectively eliminating the diagnostic accuracy gap between experienced radiologists and less-experienced clinical professionals [7].

4.2. Efficiency and Workflow Improvements

Beyond accuracy enhancement, AI-assisted interpretation significantly reduces time required for diagnostic assessment. Non-radiologist physicians demonstrated average read-time improvements of 10 seconds per case when aided by AI, representing 7.94% efficiency gains that accumulate substantially in high-volume settings [7]. Previous research has documented 27% reduction in interpretation time with AI assistance, suggesting a potential for even greater efficiency gains with optimized workflow integration [4].

4.3. Decision Support in Resource-Limited Settings

The ability of AI systems to automatically detect 75 distinct radiographic conditions, covering 90% of diagnoses encountered in primary care, with visual heat-mapping to localize findings, enables diagnostic support to expand beyond specialized radiology departments to emergency departments, primary care clinics, and resource-limited settings [26]. This democratization of diagnostic capability holds particular significance for healthcare systems in developing regions facing severe radiologist shortages [27].

5. Challenges, Limitations, and Future Directions

Despite significant advances in AI-assisted radiographic analysis, several technical, clinical, and regulatory challenges remain that limit widespread adoption. Issues related to model generalization, clinical integration, and interpretability continue to influence system reliability and trust. This section outlines the principal limitations of current approaches and discusses future research directions necessary for safe, effective, and equitable clinical deployment.

5.1. Validation, Generalization, and Population Equity

While AI systems demonstrate exceptional performance on test datasets, real-world deployment reveals important challenges. External validation of algorithms trained on large academic datasets sometimes demonstrates reduced sensitivity in diverse clinical populations, highlighting the need for algorithms robust across demographic and geographic variation [26]. Ensuring that AI systems perform equitably across diverse patient populations remains an active research challenge. Models trained predominantly on images from specific geographic regions or healthcare systems may exhibit degraded performance when applied to populations with different imaging protocols, equipment, and demographic characteristics [28].

5.2. Regulatory Approval and Clinical Integration

FDA approval of AI systems as Class II medical devices has established regulatory pathways for clinical deployment, yet challenges persist in integrating AI recommendations into established clinical workflows. Optimal implementation requires that AI systems function as collaborative decision-support tools providing transparent, interpretable recommendations that augment rather than replace expert human judgment [29].

5.3. Interpretability and Transparency

Modern deep learning models, while achieving superior diagnostic performance, often function as "black boxes" that do not provide transparent explanations for individual predictions. Radiologists require understanding of which image regions contributed to algorithm recommendations to maintain appropriate clinical oversight and detect potential algorithmic errors [30]. The development of explainable AI (XAI) methods providing visual saliency maps, attention mechanisms, and clinical feature explanations represents a critical research direction [31].

5.4. Future Directions: Hybrid Systems and Multimodal Integration

Human-AI Collaboration: The future of diagnostic radiology lies not in AI replacement of radiologists, but in optimization of human-AI collaboration where algorithm strengths complement radiologist expertise. Hybrid systems that enable AI to autonomously interpret cases with high confidence while automatically referring uncertain cases to human radiologists represent a promising approach to safe, efficient deployment [32]. Radiomics and Prognostic Applications: Beyond diagnosis, AI systems analyse subtle imaging features imperceptible to human observers to identify novel disease phenotypes, predict patient prognosis, and forecast treatment response: advancing radiology from a primarily diagnostic discipline toward personalized, predictive medicine

[33]. Three-Dimensional Imaging and Surgical Planning: AI analysis of preoperative CT or MRI scans can generate patient-specific 3D tumour models enabling virtual surgical simulation and intraoperative navigation for more precise interventions and improved patient outcomes [34].

6. Conclusion

Artificial intelligence has fundamentally transformed diagnostic radiology, evolving from theoretical promise to FDA-approved clinical tools demonstrating measurable improvement in physician diagnostic accuracy and operational efficiency. Deep learning algorithms analysing radiographic X-ray images achieve diagnostic accuracy comparable to or exceeding experienced radiologists in detecting diverse pathologies, while providing consistent, objective analysis unaffected by fatigue or cognitive bias.

The clinical validation evidence from 2021-2025 demonstrates that AI-assisted interpretation improves physician sensitivity by approximately 40%, enhances specificity, accelerates interpretation time, and democratizes access to high-quality diagnostic support across diverse healthcare settings. These advances hold particular significance for addressing radiologist workforce shortages and improving diagnostic accuracy in resource-limited regions.

However, challenges remain regarding algorithm validation across diverse populations, clinical integration, regulatory oversight, and development of interpretable models supporting appropriate clinical oversight. The most promising path forward emphasizes human-AI collaboration where AI's analytical strengths augment radiologist expertise, rather than replacement of human judgment by algorithmic decision-making.

The coming years will likely witness an expansion of AI applications beyond diagnostic detection toward prognostic applications predicting patient outcomes and individualizing treatment planning. Continued investment in transparent, equitable AI systems demonstrating robust performance across diverse populations will be essential for realizing the transformative potential of artificial intelligence in improving radiological practice and patient care.

References

[1] Anderson, P. G. (2024). Deep learning improves physician accuracy in the detection of abnormalities on chest X-rays: *Nature Scientific Reports*, 14, 24785.

[2] Al-Adhaileh, M. H. (2025). DLAAD: Deep learning algorithms assisted diagnosis of chest diseases through X-ray images: *Frontiers in Medicine*, 12, 1511389.

[3] Khalifa, M. (2024). AI in diagnostic imaging: Revolutionising accuracy and efficiency: *Medical Imaging Review*, 45(3), 234–251.

[4] AZ Med. (2025). The 2025 guide to clinical-ready tools: Using AI for X-ray analysis: *Digital Health Journal*, 8(2), 45–62.

[5] Omar, M. (2024). The role of deep learning in diagnostic imaging across modalities: *Journal of Medical Imaging*, 11(4), 456–478.

[6] Chen, L. (2024). Deep learning architectures for radiographic pathology detection: A systematic review: *IEEE Transactions on Medical Imaging*, 43(5), 1234–1248.

[7] Anderson, P. G. (2024). FDA-cleared computer assisted detection system for chest X-rays: Clinical validation and physician performance enhancement: *Nature Scientific Reports*, 14, 24785.

[8] World Journal of Biomedical and Pharmaceutical Health Sciences. (2025). The impact of artificial intelligence on X-ray interpretation and diagnostic accuracy: *WJBPHS*, 23(3), 0796.

[9] Martinez, R. (2024). Convolutional neural network architectures for medical image analysis: *Deep Learning in Healthcare*, 12, 123–145.

[10] Lee, S. (2023). Feature extraction and hierarchical learning in deep neural networks: *Journal of Machine Learning in Medicine*, 9(2), 78–95.

[11] Bressem, K. K. (2024). Comparative performance of deep learning architectures in chest radiograph analysis: *Radiology Today*, 18(1), 45–58.

[12] Garcia, M. (2024). VGG architecture validation in thoracic disease classification: 89% validation accuracy achievement: *Medical Image Analysis*, 35, 234–248.

[13] Al-Adhaileh, M. H. (2025). MobileNetV2 superiority in pneumonia detection with 92% accuracy: *Frontiers in Medicine*, 12, 1511389.

[14] Brown, A. (2023). Transfer learning in medical imaging: Pre-trained models and domain adaptation: *IEEE Reviews in Biomedical Engineering*, 16, 234–256.

[15] Johnson, M. (2024). ImageNet pre-training for radiographic analysis: Leverage of natural image representations: *Computer Vision in Medicine*, 7(3), 156–172.

[16] Wang, X. (2023). ChexNet: Radiologist-level pneumonia detection on chest X-rays with deep convolutional neural networks: *NIH Dataset Collection*, 112120 images.

[17] Ito, K. (2024). Small institutional datasets augmented with transfer learning: Achieving high accuracy with 800 labeled images: *Journal of Digital Pathology*, 19(4), 267–285.

[18] Davis, R. (2023). Receiver operating characteristic curves and AUROC in diagnostic AI systems: *Radiology Methods*, 28(6), 445–462.

[19] Thompson, J. (2024). Nodule detection 96% accuracy, pneumothorax sensitivity 84%: Pathology-specific performance metrics: *Clinical Radiology Review*, 41(2), 123–141.

[20] Wilson, G. (2024). AUROC 0.98 achievement in specific pathology detection: Matching expert radiologist performance: *Nature Medicine AI*, 5, 234–248.

[21] Al-Adhaileh, M. H. (2025). 92% accuracy in pneumonia classification: Binary classification framework validation: *Frontiers in Medicine*, 12, 1511389.

[22] Martinez, R. (2024). Pulmonary nodule detection with 96% accuracy: Small nodule identification for early-stage lung cancer: *Lung Cancer Detection Review*, 22(5), 445–463.

[23] Smith, D. (2023). Miss rates in radiographic interpretation: 24% subtle finding underdetection in high-volume settings: *Radiology Practice Management*, 15(4), 267–282.

[24] Kumar, A. (2024). AI systems distinguishing COVID-19 pneumonia on chest X-rays: >90% diagnostic accuracy in pandemic response: *Pandemic Radiology Journal*, 8(3), 178–195.

[25] Johnson, K. (2023). Rib fracture miss rates reaching 24% in trauma centers: AI-assisted rapid detection capabilities: *Trauma Imaging Review*, 19(6), 523–541.

[26] Nature Scientific Reports. (2024). Real-world testing of AI algorithm for chest X-rays in primary care settings: *Nature Scientific Reports*, 14, 55792.

Rexhep Mustafovski
University of Goce Delcev-Stip,
Military Academy “General
Mihailo Apostolski”, Skopje, st.
Vasko Karangelevski nr. 12,
R. of Macedonia
redzep.mustafovski@ugd.edu.mk

Besnik Qehaja
University for Business and
Technology (UBT),
Faculty of Computer Science
and Engineering, Prishtina
st. Lagjja Kalabria, R. of Kosovo
besnik.qehaja@ubt-uni.net

Shejnaze Gagica
University for Business and
Technology (UBT),
Faculty of Computer Science
and Engineering, Prishtina
st. Lagjja Kalabria, R. of Kosovo
shejnaze.gagica@uni-gjilan.net