

17

Chapter

Role of Medical Imaging and Machine Learning Techniques in Assistive Disease Diagnosis

Sourav Dey Roy¹ & Arpan Datta Roy²

¹Assistant Professor, School of Cyber Security and Digital Forensics,
NFSU Tripura Campus, Agartala, Tripura

²Assistant Professor, School of Forensic Science,
NFSU Tripura Campus, Agartala, Tripura

Abstract

Healthcare is undergoing a transformative evolution driven by advances in digital technologies, with medical imaging and machine learning (ML) emerging as pivotal tools for improving disease diagnosis, treatment, and patient outcomes. Medical imaging modalities enable visualization of internal body structures, while ML algorithms facilitate automated, accurate, and efficient analysis of complex imaging data. The integration of ML with medical imaging has demonstrated remarkable success in early disease detection, personalized treatment planning, predictive prognosis, and continuous patient monitoring. This chapter provides a comprehensive overview of medical imaging techniques and ML approaches applied in healthcare, emphasizing their synergistic role in assistive disease diagnosis and management. The insights presented aim to guide researchers, clinicians, and policymakers in leveraging ML-driven medical imaging for improved, patient-centered care.

Keywords: Healthcare Technology; Medical Imaging; Machine Learning (ML) in Healthcare; Disease Diagnosis and Detection; Explainable Machine Learning.

1. Introduction

Healthcare is a fundamental pillar of human well-being, continually evolving with advancements in science and technology to improve the quality, accessibility, and efficiency of medical services. In recent decades, the integration of digital technologies into healthcare systems has transformed traditional practices, enabling more precise, data-driven, and patient-centered approaches. Among these advancements, medical imaging has emerged as a cornerstone of modern diagnosis and treatment, providing non-invasive visualization of internal body structures and functions. Techniques such as X-ray, computed tomography (CT), magnetic resonance imaging (MRI), and ultrasound play a crucial role in detecting and monitoring a wide range of diseases, including cancer, cardiovascular disorders, and neurological conditions.

The rapid advancement of digital technologies has significantly transformed the landscape of modern healthcare, giving rise to the concept of intelligent healthcare. Despite the remarkable capabilities of medical imaging, the increasing volume and complexity of imaging data present significant challenges for healthcare professionals. Manual interpretation of images is often time-consuming and subject to variability depending on the expertise of radiologists. This is where machine learning (ML) techniques have begun to play a transformative role. Fig. 1 presents a block diagrammatic representation of the diverse applications of machine learning in healthcare, with a central node labeled “Applications of Machine Learning in Healthcare.” This central block acts as the core system, interconnected with multiple surrounding functional modules, each representing a key application domain. These include diagnosis and disease detection, where machine learning models assist in early identification of diseases; medical imaging, which leverages advanced algorithms for image analysis and interpretation; and predictive analytics, used for forecasting disease progression and patient outcomes. Additional blocks highlight patient monitoring through wearable devices and real-time data analysis, drug discovery for accelerating the development of new medicines, and personalized treatment that tailors therapies based on individual patient data. Another important component, healthcare operations, focuses on optimizing administrative and clinical workflows. Arrows connecting

these blocks emphasize the integration and continuous data flow between different components, illustrating how machine learning serves as a unifying framework to enhance efficiency, accuracy, and decision-making across the healthcare ecosystem.

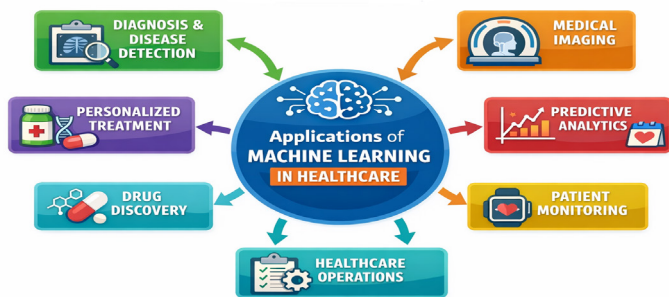


Fig. 1 Block Diagram Representing Applications of Machine Learning in Healthcare

Consequently, the integration of medical imaging and machine learning has demonstrated remarkable success in the early detection and diagnosis of critical diseases such as cancer, cardiovascular disorders, neurological conditions, and pulmonary infections. For instance, machine learning models can identify minute abnormalities in imaging scans, such as tumors, lesions, or tissue irregularities, often at stages earlier than conventional diagnostic methods. This early and accurate diagnosis plays a crucial role in improving patient outcomes and reducing healthcare costs. Detecting diseases at an initial stage allows for timely intervention, which can prevent disease progression, reduce complications, and increase survival rates. For example, early detection of cancers or cardiovascular diseases can significantly enhance the effectiveness of treatment strategies. Moreover, these technologies assist clinicians in reducing diagnostic errors, improving workflow efficiency, and handling the increasing demand for imaging-based evaluations. Beyond diagnosis, machine learning-driven imaging analysis plays a vital role in treatment planning, prognosis assessment, and disease monitoring. Predictive analytics can support personalized medicine by tailoring treatment strategies based on individual patient data, imaging biomarkers, and disease progression patterns.

This chapter aims to provide a comprehensive overview of the role of medical imaging and machine learning techniques in assistive disease diagnosis and treatment. It explores the fundamental concepts of medical imaging modalities and machine learning algorithms, highlighting their integration in healthcare applications. The insights presented in this chapter are intended to benefit researchers, practitioners, and policymakers interested in leveraging technology for improved healthcare delivery using the applications of machine learning.

2. Medical Imaging for Disease Diagnosis

Medical imaging modalities are essential tools in modern healthcare, enabling non-invasive visualization of internal body structures for diagnosis, treatment planning, and disease monitoring. Each modality is based on different physical principles and is suited for specific clinical applications. It supports early detection, accurate diagnosis, treatment planning, and continuous monitoring of a wide range of diseases, including cancer, cardiovascular disorders, neurological conditions, and musculoskeletal abnormalities. Over the years, advances in imaging technologies have significantly improved the precision and reliability of medical diagnosis. This subsection describes the medical imaging modalities and their properties of structural appearance for abnormality detection towards disease diagnosis.

X-ray Imaging. X-ray imaging is one of the oldest and most widely used medical imaging modalities, relying on ionizing radiation to produce images of internal body structures. When X-rays pass through the body, different tissues absorb varying amounts of radiation depending on their density. Dense structures such as bones absorb more radiation and appear white, while softer tissues appear in shades of gray, and air-filled spaces appear black. X-rays are commonly used for diagnosing fractures, infections, lung diseases such as pneumonia, and dental issues. The technology is relatively fast, cost-effective, and widely accessible, making it suitable for emergency situations. However, it provides limited soft tissue contrast compared to advanced modalities like MRI. Digital radiography has improved image quality and reduced radiation exposure compared to traditional film-based systems. Despite its advantages, repeated exposure to ionizing radiation can pose health risks, so its use is carefully controlled. Advances

such as computed radiography and portable X-ray devices have enhanced clinical workflows, especially in rural and critical care settings.

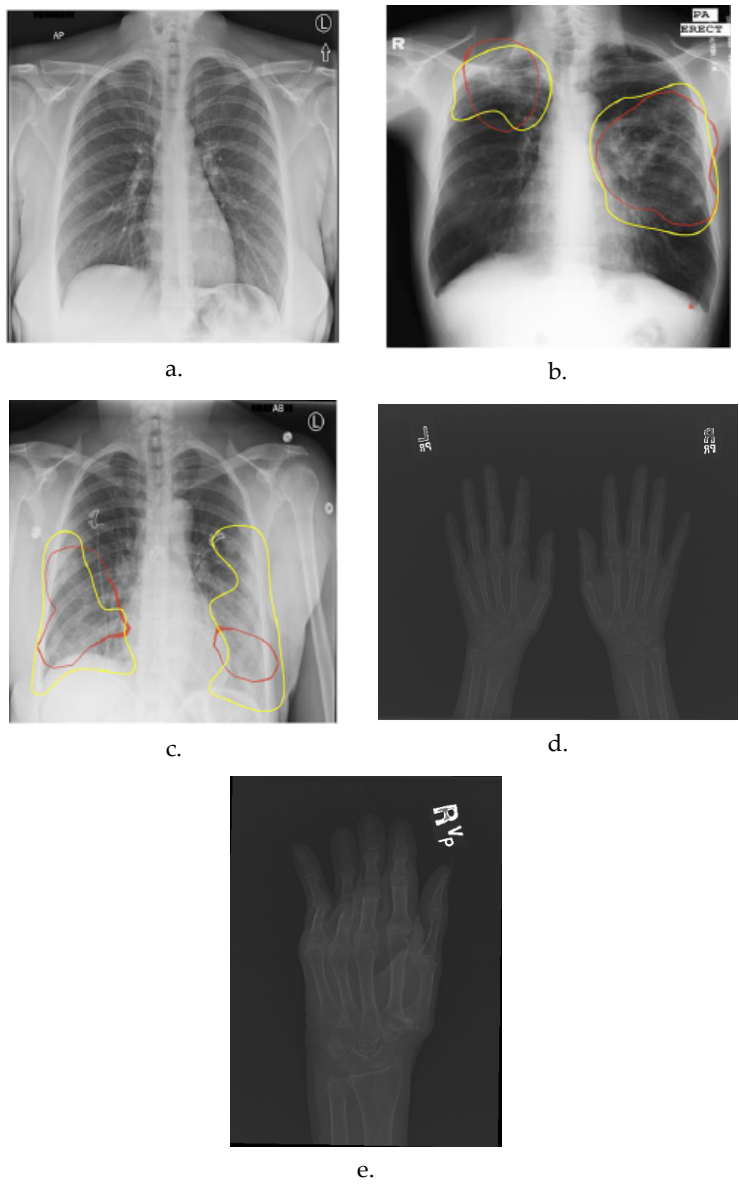


Fig. 2 X-Ray Imaging for Disease Diagnosis

Considering an example, a normal chest X-ray shows clear lung fields without opacities, indicating the absence of infection or fluid accumulation. The rib cage appears as bright white structures due to high calcium content, while the lungs appear dark because they are air-filled. The heart shadow is visible in the central region, with a normal cardiothoracic ratio. The diaphragm forms a dome-shaped boundary at the base of the lungs. This type of image is often used as a baseline reference for detecting abnormalities such as tumors, infections, or structural deformities [1][2]. On the otherhand, bone fracture X-rays reveal discontinuities or breaks in the bone structure, appearing as dark lines or irregularities across otherwise bright white bone regions [3]. Depending on the severity, fractures may be simple (clean break), comminuted (multiple fragments), or displaced (misaligned bones). X-ray imaging is the primary diagnostic tool for fractures due to its high sensitivity to bone density differences. These images are crucial for treatment planning, including casting or surgical intervention. Consequently, Tuberculosis X-ray images typically show upper-lobe opacities, cavitations, or nodular lesions in the lungs. In advanced cases, fibrosis and lung volume loss may be observed [4]. These abnormalities appear as irregular white patches or hollow spaces in lung regions. Chest X-rays play a critical role in early TB screening, especially in regions with limited access to advanced imaging. AI-based systems increasingly assist in automated TB detection from such radiographs.

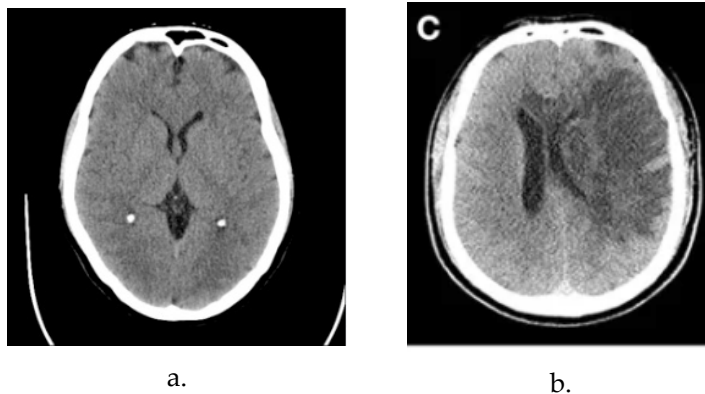


Fig. 3 Computed Tomography (CT) for Disease Diagnosis

Computed Tomography (CT). Computed Tomography (CT) is an advanced imaging technique that combines multiple X-ray measurements taken from different angles to produce cross-sectional images of the body. These images are reconstructed using computational algorithms to provide detailed visualization of bones, soft tissues, and blood vessels. CT scans are particularly useful for detecting tumors, internal bleeding, complex fractures, and lung diseases. The modality offers higher spatial resolution than conventional X-rays and can generate 3D reconstructions, aiding in surgical planning and diagnosis. Modern CT systems employ spiral (helical) scanning and multi-detector arrays to improve speed and accuracy. Contrast agents are often used to enhance the visibility of specific organs or vascular structures. However, CT scans involve higher radiation doses compared to standard X-rays, which is a concern for repeated imaging. Despite this, CT remains a cornerstone in emergency medicine due to its rapid imaging capability and diagnostic precision. Continuous advancements in low-dose CT and AI-based reconstruction techniques are improving safety and image quality.

For example reference, axial CT images of the brain are widely used in emergency diagnosis, especially for detecting intracranial hemorrhage, stroke, tumors, and traumatic brain injury. In these scans, dense structures like bone appear bright white, while cerebrospinal fluid appears dark. Acute hemorrhage is typically hyperdense (bright), making CT the first-line modality in suspected stroke cases. The symmetry of brain structures is also assessed to identify abnormalities such as edema or mass effect. Consequently, Chest CT scans provide detailed visualization of lung parenchyma and are extensively used for diagnosing lung infections, pulmonary nodules, fibrosis, and COVID-19-related abnormalities [5]. The lung window setting enhances contrast between air-filled spaces and soft tissue, allowing detection of ground-glass opacities, consolidations, and interstitial patterns. CT is far more sensitive than standard X-rays for early-stage lung disease. For instance, Contrast-enhanced CT scans of the abdomen help visualize organs such as the liver, kidneys, pancreas, and spleen, as well as vascular structures. Contrast agents improve differentiation between tissues, aiding in the detection of tumors, cysts, inflammation, and internal bleeding. These scans are commonly used in oncology and trauma cases [6].

Magnetic Resonance Imaging (MRI). Magnetic Resonance Imaging (MRI) is a non-invasive imaging modality that uses strong magnetic fields and radiofrequency waves to generate detailed images of internal body structures. Unlike X-rays and CT scans, MRI does not use ionizing radiation, making it safer for repeated use. MRI is particularly effective for imaging soft tissues such as the brain, spinal cord, muscles, and ligaments. It works by aligning hydrogen protons in the body and detecting the signals emitted as they return to their original state after excitation. Different tissues produce distinct signal patterns, allowing for high-contrast imaging. MRI is widely used in neurology, oncology, and musculoskeletal diagnostics. Functional MRI (fMRI) can even map brain activity by measuring changes in blood flow. However, MRI scans are relatively expensive and time-consuming compared to other modalities. Patients with metal implants or claustrophobia may face limitations. Recent innovations include faster scanning techniques, open MRI systems, and AI-assisted image reconstruction, which enhance diagnostic accuracy and patient comfort.

T1-weighted MRI images provide high anatomical detail of brain structures. In these images, white matter appears brighter than gray matter, while cerebrospinal fluid (CSF) appears dark. These scans are widely used to assess brain anatomy, detect tumors, structural abnormalities, and neurodegenerative changes. T1 images are particularly useful for evaluating normal anatomical structures and post-contrast enhancement in tumor diagnosis. Conversely, T2-weighted MRI images highlight fluid-containing regions, where CSF appears bright. These images are highly sensitive to edema, inflammation, infections, and lesions. They are commonly used in diagnosing conditions such as multiple sclerosis, stroke, and brain tumors. The contrast difference helps clinicians identify abnormal tissue more clearly compared to T1-weighted images. For instance, Functional MRI (fMRI) measures brain activity by detecting changes in blood oxygenation (BOLD signal). Active brain regions appear highlighted during specific tasks such as movement, speech, or visual processing. fMRI is widely used in neuroscience research, brain mapping, and pre-surgical planning to identify critical functional areas [7].

Ultrasound Imaging. Ultrasound imaging, also known as sonography, uses high-frequency sound waves to produce real-time images of internal organs and tissues. A transducer

emits sound waves that reflect off tissues, and the returning echoes are used to construct images. Ultrasound is widely used in obstetrics for monitoring fetal development, as well as in cardiology, abdominal imaging, and vascular studies. It is safe, non-invasive, and does not involve ionizing radiation, making it ideal for repeated use and sensitive populations such as pregnant women. Doppler ultrasound can assess blood flow and detect abnormalities such as blockages or clots. The modality is portable and relatively inexpensive compared to CT and MRI. However, image quality can be operator-dependent and may be limited in patients with obesity or excessive gas. Despite these limitations, ultrasound remains a versatile and widely used diagnostic tool, especially in point-of-care and emergency settings. Advances in 3D/4D ultrasound and AI-based interpretation are further enhancing its capabilities.

Obstetric Ultrasound (Fetal Imaging) shows typical prenatal ultrasound scans used to monitor fetal growth and development. The fetus appears as a bright structure within the darker amniotic fluid. Key anatomical features such as the head, spine, and limbs can be visualized. These scans are essential for assessing gestational age, detecting congenital anomalies, and monitoring fetal health. Real-time imaging allows clinicians to observe fetal movement and heartbeat [8]. Consequently, Doppler ultrasound images show blood flow within vessels using color mapping (typically red and blue). These colors indicate the direction and velocity of blood flow relative to the transducer. This technique is crucial for detecting vascular disorders such as stenosis, thrombosis, and abnormal circulation. Spectral Doppler waveforms can also quantify flow velocity [9].

Positron Emission Tomography (PET). Positron Emission Tomography (PET) is a functional imaging modality that provides insights into metabolic and biochemical processes within the body. It involves the injection of a radioactive tracer, typically fluorodeoxyglucose (FDG), which accumulates in metabolically active tissues. As the tracer decays, it emits positrons that interact with electrons, producing gamma rays detected by the scanner. PET images are often combined with CT or MRI to provide both functional and anatomical information. This modality is widely used in oncology for detecting cancer, monitoring treatment response, and identifying metastases.

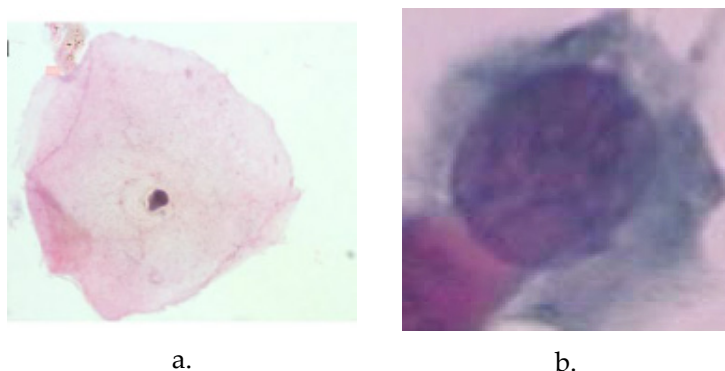


Fig. 4 Pap-Smear (Cytological Imaging) for Disease Diagnosis

It is also used in neurology and cardiology to study brain activity and heart function. PET scans are highly sensitive but relatively expensive and involve exposure to radioactive materials. The resolution is lower than CT or MRI, but the functional information is invaluable for diagnosis and treatment planning. Advances in hybrid imaging systems like PET/CT and PET/MRI have significantly improved diagnostic accuracy. Brain PET imaging shows metabolic activity across different brain regions. In neurodegenerative diseases like Alzheimer's, reduced FDG uptake (hypometabolism) is observed in the temporal and parietal lobes. PET is valuable for early diagnosis and differentiation of dementia types [10]. Cardiac PET evaluates myocardial perfusion and viability. Areas with reduced tracer uptake indicate ischemia or infarction. It helps in assessing coronary artery disease and determining whether damaged heart tissue can recover after treatment [11].

Pap-Smear (Cytological Imaging). Pap-smear is a cytological screening technique used primarily for the early detection of cervical cancer. Unlike other imaging modalities, it involves microscopic examination of cells collected from the cervix. The sample is stained and analyzed to identify abnormal or precancerous cells. This method has significantly reduced cervical cancer mortality through early detection and intervention. In modern healthcare, Pap-smear analysis is increasingly supported by digital imaging and machine learning techniques for automated cell classification. High-resolution microscopy images are processed to detect abnormalities such

as changes in cell shape, size, and nucleus structure. While the procedure is simple, cost-effective, and minimally invasive, it requires skilled interpretation. False negatives can occur if abnormal cells are missed. Integration with HPV testing has improved diagnostic accuracy. Recent advancements include AI-based cytology systems that enhance screening efficiency and reduce human error, making Pap-smear a critical component of preventive healthcare. Pap-smear imaging begins with the identification of normal cervical epithelial cells, which serve as a baseline for detecting abnormalities. Normal samples typically contain superficial and intermediate squamous epithelial cells with small, uniform nuclei and abundant cytoplasm [12]. The nucleus-to-cytoplasm (N:C) ratio is low, and chromatin is evenly distributed. These cells appear flat and polygonal under light microscopy after staining using the Papanicolaou stain, which enhances nuclear detail and cytoplasmic transparency. In abnormal cases, cytological samples show marked irregular nuclear shapes, and prominent nucleoli. Cells lose their normal architecture and appear disorganized. Highly irregular cells with enlarged, hyperchromatic nuclei and distorted shapes. Background may contain necrotic debris and inflammatory cells, indicating tumor invasion [13][14].

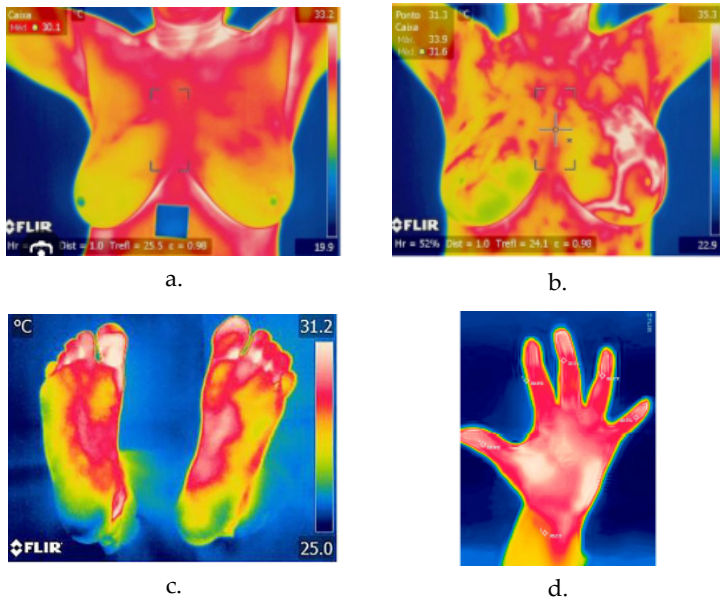


Fig. 5 Infrared Thermography for Disease Diagnosis

Infrared Imaging (Thermography). Infrared imaging, also known as thermography, detects heat patterns and blood flow in body tissues using infrared radiation. It captures temperature variations on the skin surface, which can indicate underlying physiological changes. Areas with increased blood flow, inflammation, or abnormal metabolic activity often appear warmer in thermographic images. This modality is non-invasive, radiation-free, and painless, making it suitable for frequent monitoring. Infrared imaging has applications in breast cancer screening, vascular disorders, musculoskeletal injuries, and inflammation detection. However, it does not provide structural details like CT or MRI and is often used as a complementary tool rather than a standalone diagnostic method. The accuracy can be influenced by environmental conditions and patient preparation. Advances in image processing and AI have improved pattern recognition and diagnostic reliability. Despite its limitations, thermography is valuable for early detection and monitoring of conditions associated with temperature changes. Breast thermography images typically display temperature distributions across both breasts using a color-coded scale, where warmer regions appear in red, orange, or white, and cooler regions appear in blue or green. In normal cases, the thermal pattern is bilaterally symmetrical. However, in abnormal cases—such as suspected breast cancer—localized hotspots or asymmetric vascular patterns are observed (as shown in Fig. 5 (b)). These hotspots indicate increased metabolic activity and angiogenesis associated with tumor growth. Studies have shown that thermography can detect physiological changes even before structural abnormalities become visible in modalities like mammography. However, due to variability in sensitivity and specificity, it is primarily used as an adjunct screening tool rather than a standalone diagnostic method [15][16]. For instance, in diabetic patients (as shown in Fig. 5 (c)), thermography can reveal early signs of diabetic foot complications by detecting subtle temperature differences before visible ulcers develop [17]. Further, thermographic imaging is highly effective in identifying musculoskeletal injuries, inflammation, and pain-related conditions. Injured or inflamed tissues exhibit elevated temperatures due to increased blood flow and metabolic activity, which appear as bright red or white regions in thermograms. For example, in cases of arthritis [18], tendonitis, or sports injuries, thermography can highlight affected joints or muscles even before swelling becomes clinically evident.

Histopathological Imaging. Histopathological imaging involves the microscopic examination of biological tissues to study the manifestations of disease, particularly cancer. Tissue samples are collected through biopsy or surgical excision, processed, and stained using techniques such as Hematoxylin and Eosin (H&E) staining to highlight cellular structures. These stained slides are then observed under a microscope or digitized using whole-slide imaging systems for computational analysis. Histopathology is considered the gold standard for definitive diagnosis of many diseases, as it provides detailed insights into cellular morphology, tissue architecture, and pathological abnormalities. In recent years, digital pathology has enabled the integration of machine learning and deep learning algorithms for automated detection, segmentation, and classification of abnormal cells. Convolutional Neural Networks (CNNs) are widely used to identify cancerous patterns, grade tumors, and assist pathologists in decision-making. Despite its high diagnostic accuracy, the process is time-consuming and requires expert interpretation. Variability in staining and imaging conditions can also affect analysis. However, advancements in AI-assisted pathology, cloud-based platforms, and high-resolution scanners are transforming histopathological imaging into a scalable and efficient diagnostic tool. In normal tissue, cells exhibit uniform shape, well-organized, and clearly defined nuclei stained in blue (hematoxylin), while cytoplasm appears pink (eosin). In contrast, cancerous tissue shows pleomorphism (variation in cell size and shape), enlarged nuclei, hyperchromasia (dark staining), and disrupted tissue architecture. These morphological abnormalities are critical indicators used by pathologists to identify malignancy. Machine learning models, particularly CNNs, learn these patterns for automated classification of benign vs malignant tissues [19][20].

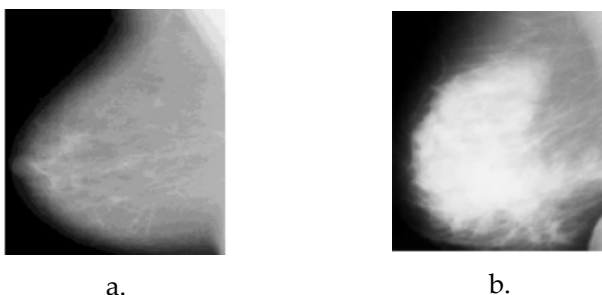


Fig. 6 Mammography for Disease Diagnosis

Mammography. Mammography is a specialized medical imaging modality that uses low-dose X-rays to examine breast tissue, primarily for the early detection and diagnosis of breast cancer. It is capable of identifying microcalcifications, masses, and structural abnormalities that may not be palpable during physical examination. There are two main types of mammography: screening mammography, used for routine check-ups in asymptomatic individuals, and diagnostic mammography, used when abnormalities are suspected. Digital mammography has largely replaced traditional film-based systems, offering improved image quality, enhanced contrast, and easier storage and transmission. Advanced techniques such as Digital Breast Tomosynthesis (DBT), also known as 3D mammography, provide layered images of the breast, improving lesion detection and reducing false positives. Mammography plays a critical role in reducing breast cancer mortality through early detection. However, it involves exposure to low levels of ionizing radiation and may have reduced sensitivity in dense breast tissues. Recent developments include AI-based computer-aided detection (CAD) systems that assist radiologists in identifying suspicious regions, improving diagnostic accuracy and efficiency in large-scale screening programs. Here, a normal mammogram typically shows homogeneous fatty or fibroglandular tissue with no suspicious masses, distortions, or calcifications. In contrast, abnormal mammograms may reveal irregular masses, architectural distortion, or clustered microcalcifications [21]. Radiologists assess these features based on shape, margin, and density to classify findings using BI-RADS (Breast Imaging Reporting and Data System). Consequently, masses in mammography are evaluated based on shape (round, oval, irregular), margins (circumscribed, obscured, spiculated), and density. Malignant masses often exhibit irregular shapes and spiculated margins due to invasive growth patterns. Benign tumors such as fibroadenomas usually present as smooth, well-defined lesions. Mammography enables early visualization of these masses, even when they are too small to be clinically detected, thus playing a crucial role in early diagnosis and treatment planning [22].

3. Challenges in Medical Imaging Data for Building Decision Support Systems Towards Diagnosis of the Disease

Medical imaging data presents a wide range of challenges that significantly affect the performance, reliability, and generalization of machine learning models. Below is a detailed discussion of key challenges beyond noise, artifacts, class imbalance, and annotation issues:

Noise and Signal Variability. Medical images often suffer from modality-specific noise, such as Gaussian noise in MRI or Poisson noise in low-dose CT scans. Variability in acquisition parameters (scanner type, resolution, patient movement) further introduces inconsistencies. This degrades image quality and complicates feature extraction, often requiring sophisticated denoising techniques or robust model architectures.

Imaging Artifacts. Artifacts arise due to hardware limitations, patient motion, or reconstruction errors. Examples include motion blur in MRI, beam hardening in CT, and speckle noise in ultrasound. These distortions may mimic pathological patterns, leading to false positives or reduced diagnostic confidence. Artifact removal remains a complex preprocessing challenge.

Class Imbalance. Medical datasets are typically skewed, with significantly fewer abnormal/pathological samples compared to normal ones. This imbalance biases models toward majority classes, reducing sensitivity to rare but critical conditions. Techniques like weighted loss functions, oversampling (e.g., SMOTE), and focal loss are often used to address this.

Data Annotation and Labeling Issues. Accurate annotation requires domain expertise from radiologists or pathologists, making it time-consuming and expensive. Inter-observer variability leads to inconsistent labels, especially in borderline cases. Additionally, pixel-level annotations (e.g., segmentation masks) are labor-intensive, limiting the availability of high-quality labeled datasets.

Limited Data Availability and Privacy Constraints. Medical data is highly sensitive and governed by strict privacy regulations (such as HIPAA and GDPR). This restricts data sharing across institutions, resulting in small, fragmented datasets. Deep learning models, which require large-scale data, often struggle to generalize under such constraints.

Heterogeneity Across Imaging Devices and Protocols. Images acquired from different hospitals or scanners vary in resolution, contrast, and acquisition protocols. This domain shift leads to poor cross-institutional generalization. Models trained on one dataset may fail when applied to another due to this distribution mismatch.

High Dimensionality and Computational Complexity. Medical images, especially 3D modalities like CT and MRI, have high spatial resolution and volumetric depth. This leads to large data sizes, increasing memory requirements and computational cost. Efficient architectures and dimensionality reduction techniques are necessary to manage this complexity.

Multi-Modality Data Integration Challenges. Combining information from multiple imaging modalities (e.g., MRI, CT, PET) can improve diagnostic accuracy, but aligning and fusing such heterogeneous data is non-trivial. Differences in spatial resolution, orientation, and intensity distributions complicate multimodal learning.

Interpretability and Explainability Issues. Deep learning models in medical imaging are often black boxes. Lack of transparency reduces trust among clinicians, especially in high-stakes decision-making. Developing explainable AI (XAI) methods, such as saliency maps or attention mechanisms, is critical but still evolving.

Overfitting and Poor Generalization. Due to limited and imbalanced datasets, models tend to memorize training data rather than learn generalizable features. This results in high training accuracy but poor performance on unseen data. Regularization techniques, cross-validation, and data augmentation are commonly used to mitigate this issue.

Ethical and Bias Concerns. Datasets may not represent diverse populations, leading to biased models that perform poorly on underrepresented groups. This raises ethical concerns, especially when AI systems are used in clinical decision-making. Ensuring fairness and inclusivity is a major ongoing challenge.

4. Machine Learning Foundations for Medical Imaging

Machine Learning is playing a transformative role in healthcare by augmenting the power of medical imaging. These techniques can process vast amounts of imaging data with remarkable speed

and accuracy, assisting in early detection, disease classification, and prognosis prediction. Machine Learning systems can identify subtle patterns and anomalies that may be overlooked by human observers, thereby enhancing diagnostic precision. Furthermore, AI-driven tools support personalized medicine by tailoring treatment plans based on individual patient data. As these technologies continue to evolve, the integration of early diagnosis, digital imaging, and machine learning is paving the way for a smarter, more efficient, and patient-centric healthcare ecosystem.

Machine Learning Techniques. Machine learning techniques in medical imaging are broadly categorized into supervised, unsupervised, and reinforcement learning paradigms, each addressing different clinical and computational challenges. Supervised learning is the most widely used paradigm in medical imaging, where models are trained using labeled datasets consisting of input images and corresponding annotations (e.g., tumor boundaries, disease labels). These methods are particularly effective in tasks such as classification, segmentation, and detection. Deep supervised models, especially convolutional neural networks (CNNs), have demonstrated high accuracy in identifying pathological patterns due to their ability to learn hierarchical feature representations directly from imaging data [23]. However, their performance heavily depends on the availability of large, well-annotated datasets, which are often expensive and time-consuming to obtain in healthcare settings. Unsupervised learning addresses the limitation of labeled data by discovering intrinsic patterns in unlabeled datasets. Techniques such as clustering [24], autoencoders [25], and generative models [26] are used to learn data distributions and latent representations. In medical imaging, unsupervised learning is useful for anomaly detection, image reconstruction, and feature extraction, particularly when annotated data is scarce [27]. It plays a critical role in radiomics, where high-dimensional imaging features are extracted to characterize disease phenotypes. Despite its potential, unsupervised learning may produce less clinically interpretable results due to the absence of explicit supervision. Reinforcement learning (RL) is an emerging paradigm in medical imaging, where an agent learns optimal decision-making strategies through interaction with an environment. RL has been applied to tasks such as landmark

detection, image navigation, and adaptive image acquisition [28]. Although less commonly used compared to supervised methods, RL shows promise in optimizing sequential decision processes in imaging workflows. Recent studies highlight its application in improving diagnostic accuracy and automation, though challenges such as complexity and limited clinical adoption remain.

Deep Learning Architectures. On the otherhand, deep learning architectures form the backbone of modern medical image analysis systems, enabling automated feature extraction and high-level representation learning. Convolutional Neural Networks (CNNs) are the most dominant architecture in medical imaging due to their ability to capture spatial hierarchies through convolutional filters [23]. CNNs are extensively used for image classification, segmentation, and detection tasks. Architectures such as U-Net and ResNet have significantly improved performance in clinical applications by enabling precise localization and robust feature extraction [29]. Their success lies in exploiting local spatial correlations and translation invariance in images. Consequently, Recurrent Neural Networks (RNNs) are designed for sequential data processing and are particularly useful when temporal or contextual information is involved, such as in longitudinal imaging studies or video-based diagnostics (e.g., ultrasound or endoscopy). Variants like Long Short-Term Memory (LSTM) networks address issues like vanishing gradients and enable learning long-term dependencies [31]. In medical imaging, RNNs are often combined with CNNs to incorporate both spatial and temporal features. Transformers represent a newer class of architectures that rely on self-attention mechanisms to capture global contextual relationships in data. Unlike CNNs, which focus on local receptive fields, transformers model long-range dependencies effectively. Their application in medical imaging has grown rapidly, particularly in segmentation, classification, and multimodal learning tasks. Recent surveys highlight that transformers are increasingly integrated with CNNs to leverage both local and global feature representations.

2D vs 3D Deep Learning Models. Medical imaging data can be processed using either 2D or 3D deep learning models, depending on the modality and application. 2D models operate

on individual image slices and are computationally efficient, making them suitable for large-scale datasets and real-time applications [32]. They are widely used in modalities such as X-rays and histopathology images. However, 2D models may fail to capture volumetric context, which is critical in many diagnostic tasks. 3D models, on the other hand, process volumetric data (e.g., CT, MRI scans) and can capture spatial dependencies across multiple slices. These models provide more comprehensive anatomical information and often achieve higher accuracy in segmentation and detection tasks. For example, 3D CNNs have been successfully applied to tumor characterization and volumetric analysis in clinical studies. However, they require significantly higher computational resources and memory, limiting their scalability [33][34].

Feature Engineering vs Representation Learning. Feature extraction is a critical step in medical image analysis, traditionally performed using handcrafted methods but increasingly replaced by automated representation learning. Feature engineering (handcrafted features) involves manually designing features based on domain knowledge, such as texture, shape, intensity, and statistical descriptors [35][36]. Radiomics is a prominent example, where quantitative features are extracted from medical images to support clinical decision-making [37]. These features can reveal subtle patterns not visible to the human eye, aiding in diagnosis and prognosis. However, handcrafted features are often task-specific and may lack generalization across datasets. Representation learning (deep features), enabled by deep learning, automatically learns hierarchical feature representations directly from raw data. CNNs and other deep architectures eliminate the need for manual feature design by learning discriminative features during training. These learned features are more robust, scalable, and adaptable to different tasks. Studies have shown that deep features outperform handcrafted features in various medical imaging applications, particularly in complex tasks such as tumor detection and multi-class classification. The transition from handcrafted to deep features represents a paradigm shift, allowing models to capture complex, non-linear patterns and improving overall diagnostic performance.

5. Explainable Machine Learning Techniques for Building Medical Imaging Guided Decision Support Systems

The integration of machine learning (ML) into medical imaging has significantly advanced computer-aided diagnosis, enabling automated detection, classification, segmentation, and prognosis of diseases. However, the adoption of ML models in clinical practice is often hindered by their “black-box” nature, particularly in deep learning architectures such as convolutional neural networks (CNNs) and transformers. In high-stakes environments like healthcare, clinicians require transparency, interpretability, and trust in automated decisions. This has led to the emergence of Explainable Artificial Intelligence (XAI), a subfield within Machine Learning, focused on making model predictions understandable and justifiable as shown in Fig. 7. Medical imaging-guided decision support systems (DSS) rely on explainable ML techniques to provide not only accurate predictions but also interpretable insights into the reasoning behind those predictions [38][39]. These systems are critical in domains such as radiology, histopathology, and cardiology, where decisions directly impact patient outcomes. Explainability is essential in medical imaging DSS for several reasons. First is the clinical trust and adoption where clinicians are more likely to trust systems that provide visual or feature-based explanations rather than opaque predictions. Second is the regulatory compliance. Healthcare regulations increasingly require transparency in AI-driven decision-making. Third is the error analysis and debugging. Explainability helps identify model biases, data leakage, and failure modes. Fourth is the ethical considerations. Transparent models reduce the risk of biased or unfair decisions, particularly in sensitive populations. Explainability bridges the gap between complex computational models and human-understandable reasoning, facilitating collaborative decision-making between AI systems and clinicians. Explainable Machine Learning (XML) techniques can be broadly categorized based on how interpretability is achieved either inherently through model design or externally through post-hoc analysis. These categories are particularly important in medical imaging applications, where transparency and trustworthiness are critical for clinical decision-making.

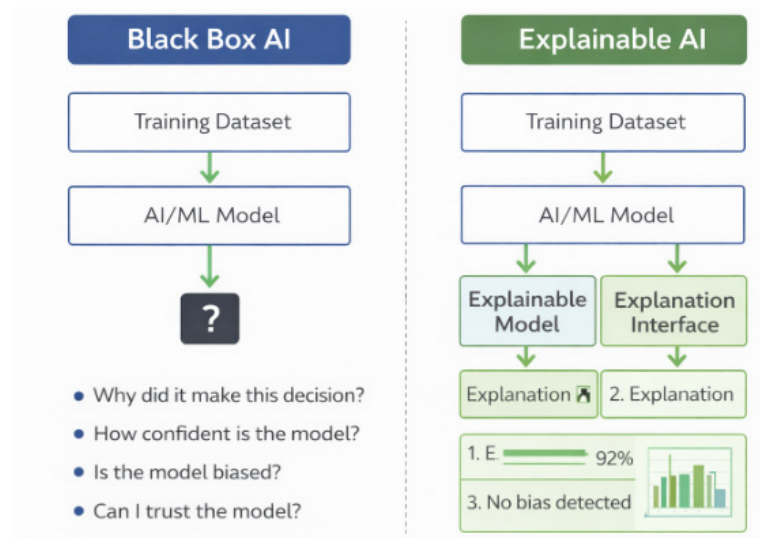


Fig. 7 Need for Explainable Machine Learning in Disease Diagnosis

Intrinsic Interpretability (White-box Models). Intrinsic interpretability refers to models that are inherently transparent in their structure and decision-making processes [40]. Such models, often termed white-box models, include decision trees, linear regression, logistic regression, and rule-based systems. These approaches enable direct understanding of how input features contribute to predictions through explicit mathematical relationships or decision rules. For instance, in linear models, feature coefficients provide a clear indication of feature importance, while decision trees visually represent hierarchical decision paths. Despite their interpretability advantages, these models often struggle to capture the complex, high-dimensional patterns present in medical imaging data, such as those found in MRI, CT, or histopathological images. Consequently, their predictive performance may be inferior compared to more complex models like deep neural networks.

Post-hoc Explainability Methods. Post-hoc explainability methods are designed to interpret predictions after a model has been trained, particularly when dealing with complex and opaque models such as deep learning architectures [41][42][43]. These techniques do not alter the original model but instead provide insights into its behavior, making them highly valuable

in medical imaging tasks where accuracy and interpretability must coexist. Feature attribution methods constitute a major class of post-hoc techniques. These methods assign importance scores to input features or image pixels to quantify their contribution to a model's prediction. Techniques such as SHAP (SHapley Additive exPlanations) [44] leverage principles from cooperative game theory to fairly distribute the contribution of each feature, offering both global and local interpretability. Similarly, LIME (Local Interpretable Model-agnostic Explanations) [45] approximates the behavior of complex models in the vicinity of a specific instance using simpler, interpretable surrogate models. In medical imaging, these methods are extensively used to highlight regions of interest that influence classification outcomes, thereby aiding clinicians in understanding model reasoning. Another important category is gradient-based visualization techniques, which are particularly suited for deep neural networks. These methods utilize gradients flowing through the network to identify influential regions in the input data. For example, Grad-CAM (Gradient-weighted Class Activation Mapping) [46] generates class-specific heatmaps that localize important areas in an image, while techniques like guided backpropagation and integrated gradients provide fine-grained pixel-level explanations. Such spatial visualizations are especially useful in radiology and histopathology, where identifying disease-relevant regions is essential. Perturbation-based methods offer an alternative approach by analyzing how systematic modifications to the input affect the model's output. Techniques such as occlusion sensitivity involve masking different parts of an image to observe changes in prediction confidence, while permutation importance evaluates the impact of shuffling feature values. These approaches help identify critical regions or features that significantly influence model decisions, providing intuitive and model-agnostic explanations. Overall, the choice of explainability technique depends on the application requirements, model complexity, and the level of interpretability needed, especially in sensitive domains like medical imaging where both accuracy and transparency are paramount. Further, the explainability module bridges the gap between complex model outputs and clinical interpretation, enhancing reliability, accountability, and adoption of AI-driven medical imaging systems.

6. Conclusion

The convergence of medical imaging and machine learning represents a significant advancement in modern healthcare, enabling more accurate, efficient, and patient-centered diagnosis and treatment. By leveraging advanced imaging modalities along with machine learning algorithms, clinicians can detect diseases at earlier stages, monitor progression, and personalize treatment strategies based on individual patient data. This integration not only enhances diagnostic accuracy and workflow efficiency but also supports predictive analytics, treatment planning, and outcome assessment, thereby improving overall patient care and reducing healthcare costs. Looking forward, ongoing research in explainable AI, multimodal data integration, and sustainable machine learning promises to further transform healthcare systems, enabling proactive, precise, and personalized medical care. The collaboration between medical imaging and machine learning holds immense potential to improve global health outcomes and redefine the future of disease diagnosis and management.

7. References

- Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2017). Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2097-2106).
- Nabulsi, Z., Sellergren, A., Jamshy, S., Lau, C., Santos, E., Kiraly, A. P., ... & Shetty, S. (2021). Deep learning for distinguishing normal versus abnormal chest radiographs and generalization to two unseen diseases tuberculosis and COVID-19. *Scientific reports*, 11(1), 15523.
- Rajpurkar, P., Irvin, J., Bagul, A., Ding, D., Duan, T., Mehta, H., ... & Ng, A. Y. (2017). Mura: Large dataset for abnormality detection in musculoskeletal radiographs. *arXiv preprint arXiv:1712.06957*.
- Jaeger, S., Candemir, S., Antani, S., Wang, Y. X. J., Lu, P. X., & Thoma, G. (2014). Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. *Quantitative imaging in medicine and surgery*, 4(6), 475.
- Kazerooni, E. A. (2001). High-resolution CT of the lungs. *American Journal of Roentgenology*, 177(3), 501-519.
- Lee, J. K. (Ed.). (2006). *Computed body tomography with MRI correlation* (Vol. 1). Lippincott Williams & Wilkins.
- Huettel, S. A., Song, A. W., & McCarthy, G. (2004). *Functional magnetic resonance imaging*.

- Rumack, C. M., & Levine, D. (Eds.). (2023). Diagnostic ultrasound E-book. Elsevier Health Sciences.
- Zwiebel, W. J., & Pellerito, J. S. (2005). Introduction to vascular ultrasonography (pp. 19-89). Philadelphia: Elsevier Saunders.
- Mosconi, L., Berti, V., Guyara-Quinn, C., McHugh, P., Petrongolo, G., Osorio, R. S., ... & Brinton, R. D. (2017). Perimenopause and emergence of an Alzheimer's bioenergetic phenotype in brain and periphery. *PloS one*, 12(10), e0185926.
- Dorbala, S., Di Carli, M. F., Delbeke, D., Abbara, S., DePuey, E. G., Dilsizian, V., ... & Shreve, P. (2013). SNMMI/ASNC/SCCT Guideline for Cardiac SPECT/CT and PET/CT 1.0. *Journal of Nuclear Medicine*, 54(8), 1485-1507.
- Prieto Bohórquez, S. E., Velásquez, J. O. R., Correa Herrera, S. C., & Soracipa Muñoz, M. Y. (2014). Diagnosis of cervical cells based on fractal and Euclidian geometrical measurements: Intrinsic Geometric Cellular Organization. *BMC Medical Physics*, 14(1), 2.
- Solomon, D., Davey, D., Kurman, R., Moriarty, A., O'Connor, D., Prey, M., ... & Young, N. (2002). The 2001 Bethesda System: terminology for reporting results of cervical cytology. *Jama*, 287(16), 2114-2119.
- Fekri-Ershad, S. (2019). Pap smear classification using combination of global significant value, texture statistical features and time series features. *Multimedia Tools and Applications*, 78(22), 31121-31136.
- Gautherie, M. (1980). Thermopathology of breast cancer: measurement and analysis of in vivo temperature and blood flow. *Annals of the New York Academy of Sciences*, 335(1), 383-415.
- Arora, N., Martins, D., Ruggerio, D., Tousimis, E., Swistel, A. J., Osborne, M. P., & Simmons, R. M. (2008). Effectiveness of a noninvasive digital infrared thermal imaging system in the detection of breast cancer. *The American Journal of Surgery*, 196(4), 523-526.
- Saminathan, J., Sasikala, M., Narayanamurthy, V. B., Rajesh, K., & Arvind, R. J. I. P. (2020). Computer aided detection of diabetic foot ulcer using asymmetry analysis of texture and temperature features. *Infrared Physics & Technology*, 105, 103219.
- Avila-Camacho, F. J., Moreno-Villalba, L. M., Cortes-Altamirano, J. L., Alfaro-Rodríguez, A., Lara-Figueroa, H. N., Herrera-López, M. E., & Romero-Morelos, P. (2025). Infrared Thermography and Deep Learning Prototype for Early Arthritis and Arthrosis Diagnosis: Design, Clinical Validation, and Comparative Analysis. *Technologies*, 13(10), 447.
- Gurcan, M. N., Boucheron, L. E., Can, A., Madabhushi, A., Rajpoot, N. M., & Yener, B. (2009). Histopathological image analysis: A review. *IEEE reviews in biomedical engineering*, 2, 147-171.
- Spanhol, F. A., Oliveira, L. S., Petitjean, C., & Heutte, L. (2015). A dataset for breast cancer histopathological image classification. *Ieee transactions on biomedical engineering*, 63(7), 1455-1462.

- Novitasari, D. C. R., Lubab, A., Sawiji, A., & Asyhar, A. H. (2019). Application of feature extraction for breast cancer using one order statistic, GLCM, GLRLM, and GLDM. *Advances in Science, Technology and Engineering Systems Journal (ASTESJ)*, 4(4), 115-120.
- Sickles, E. A., Wolverton, D. E., & Dee, K. E. (2002). Performance parameters for screening and diagnostic mammography: specialist and general radiologists. *Radiology*, 224(3), 861-869.
- Kshatri, S. S., & Singh, D. (2023). Convolutional neural network in medical image analysis: a review. *Archives of Computational Methods in Engineering*, 30(4), 2793-2810.
- Fayez, M., Safwat, S., & Hassanein, E. (2016, July). Comparative study of clustering medical images. In *2016 SAI Computing Conference (SAI)* (pp. 312-318). IEEE.
- Rais, K., Amroune, M., Benmachiche, A., & Haouam, M. Y. (2024). Exploring variational autoencoders for medical image generation: a comprehensive study. *arXiv preprint arXiv:2411.07348*.
- Celard, P., Iglesias, E. L., Sorribes-Fdez, J. M., Romero, R., Vieira, A. S., & Borrajo, L. (2023). A survey on deep learning applied to medical images: from simple artificial neural networks to generative models. *Neural Computing and Applications*, 35(3), 2291-2323.
- Armya, R. E., & Abdulazeez, A. M. (2021). Medical images segmentation based on unsupervised algorithms: a review. *Qubahan Academic Journal*, 1(2), 71-80.
- Hu, M., Zhang, J., Matkovic, L., Liu, T., & Yang, X. (2023). Reinforcement learning in medical image analysis: Concepts, applications, challenges, and future directions. *Journal of Applied Clinical Medical Physics*, 24(2), e13898.
- Çalışan, M., Gündüzalp, V., & Olgun, N. (2023). Evaluation of U-net and ResNet architectures for biomedical image segmentation. *International Journal of 3D Printing Technologies and Digital Industry*, 7(3), 561-570.
- Yin, C., Qian, B., Wei, J., Li, X., Zhang, X., Li, Y., & Zheng, Q. (2019, November). Automatic generation of medical imaging diagnostic report with hierarchical recurrent neural network. In *2019 IEEE international conference on data mining (ICDM)* (pp. 728-737). IEEE.
- Zhang, H., & Qie, Y. (2023). Applying deep learning to medical imaging: a review. *Applied Sciences*, 13(18), 10521.
- Habib, A. B., Akhter, M. E., Sultaan, R., Zahir, Z. B., Arfin, R., Haque, F., ... & Palit, R. (2020, January). Performance analysis of different 2D and 3D CNN model for liver semantic segmentation: a review. In *International Conference on Medical Imaging and Computer-Aided Diagnosis* (pp. 166-174). Singapore: Springer Singapore.
- Singh, S. P., Wang, L., Gupta, S., Goli, H., Padmanabhan, P., & Gulyás, B. (2020). 3D deep learning on medical images: a review. *Sensors*, 20(18), 5097.

- Tiwari, S., Jain, G., Shetty, D. K., Sudhi, M., Balakrishnan, J. M., & Bhatta, S. R. (2023). A comprehensive review on the application of 3d convolutional neural networks in medical imaging. *Engineering Proceedings*, 59(1), 3.
- Vishraj, R., Gupta, S., & Singh, S. (2022). A comprehensive review of content-based image retrieval systems using deep learning and hand-crafted features in medical imaging: Research challenges and future directions. *Computers and Electrical Engineering*, 104, 108450.
- Ma, J., Jiang, X., Fan, A., Jiang, J., & Yan, J. (2021). Image matching from handcrafted to deep features: A survey. *International Journal of Computer Vision*, 129(1), 23-79.
- Lambin, P., Leijenaar, R. T., Deist, T. M., Peerlings, J., De Jong, E. E., Van Timmeren, J., ... & Walsh, S. (2017). Radiomics: the bridge between medical imaging and personalized medicine. *Nature reviews Clinical oncology*, 14(12), 749-762.
- Purwono, P., Wulandari, A. N. E., & Nisa, K. (2025). Explainable Artificial Intelligence (XAI) in Medical Imaging: Techniques, Applications, Challenges, and Future Directions. *Advanced Mechanical and Mechatronic Systems*, 1(1), 52-66.
- Ahmed, F., Naz, N. S., Khan, S., Rehman, A. U., Ismael, W. M., & Khan, M. A. (2026). Explainable artificial intelligence (XAI) in medical imaging: a systematic review of techniques, applications, and challenges. *BMC Medical Imaging*.
- Ayyar, M. P., Benois-Pineau, J., & Zemhari, A. (2021). Review of white box methods for explanations of convolutional neural networks in image classification tasks. *Journal of Electronic Imaging*, 30(5), 050901-050901.
- Vale, D., El-Sharif, A., & Ali, M. (2022). Explainable artificial intelligence (XAI) post-hoc explainability methods: Risks and limitations in non-discrimination law. *AI and Ethics*, 2(4), 815-826.
- Bhati, D., Amiruzzaman, M., Zhao, Y., Guercio, A., & Le, T. (2025). A survey of post-hoc xai methods from a visualization perspective: Challenges and opportunities. *IEEE Access*.
- Slack, D., Hilgard, A., Singh, S., & Lakkaraju, H. (2021). Reliable post hoc explanations: Modeling uncertainty in explainability. *Advances in neural information processing systems*, 34, 9391-9404.
- Walia, S., Kumar, K., Agarwal, S., & Kim, H. (2022). Using xai for deep learning-based image manipulation detection with shapley additive explanation. *Symmetry*, 14(8), 1611.
- Kumarakulasinghe, N. B., Blomberg, T., Liu, J., Leao, A. S., & Papapetrou, P. (2020, July). Evaluating local interpretable model-agnostic explanations on clinical machine learning classification models. In *2020 IEEE 33rd international symposium on computer-based medical systems (CBMS)* (pp. 7-12). IEEE.
- Selvaraju, R. R., Das, A., Vedantam, R., Cogswell, M., Parikh, D., & Batra, D. (2016). Grad-CAM: Gradient-weighted class activation mapping. *Arxiv*.