The Original

MEDICAL IMAGE ANALYSIS IN KNEE OSTEOARTHRITIS USING GENERATIVE AI: A COMPREHENSIVE REVIEW

Mrs. Radhika Mahesh Mane*[1], Dr. Pankaj Jain[2], Ganesh S Wedpathak[3]

[1*]Assistant Professor, Department of Computer Science and Engineering, Bharati Vidyapeeth's College of engineering, Kolhapur, India and Research Scholar, Symbiosis Institute of Technology, Pune Campus, Symbiosis International(Deemed University), Pune, India
Email: radhikajadhav2525@gmail.com Orcid ID-0009-0003-0120-9279
[2]Assistant Professor, Symbiosis Institute of Technology, Pune Campus, Symbiosis International (Deemed University), Pune, India
Email: pankaj.jain@sitpune.edu.in Orcid ID-0000-0002-6206-497X
[3]Assistant Professor, Department of Computer Science and Engineering, Kasegaon Education Society's Rajarambapu Institute of Technology, affiliated to Shivaji University, Sakharale, MS-415414, India, Email: ganesh.wedpathak@ritindia.edu Orcid ID- 0009-0006-6364-557X

ABSTRACT
Knee osteoarthritis (KOA) is a prevalent degenerative disease characterized by cartilage deterioration, joint space narrowing, osteophyte formation, and progressive functional decline. Medical imaging modalities such as radiographs, MRI, CT, and ultrasound play a central role in KOA diagnosis, severity grading, and monitoring. Traditional computer-aided diagnosis (CAD) approaches and early deep learning models face limitations due to small datasets, annotation challenges, and variability across imaging centers. Recent advances in Generative Artificial Intelligence (GenAI), including GANs, VAEs, and diffusion models, have enabled realistic synthetic image generation, domain adaptation, super-resolution enhancement, and automated grading. This review provides a comprehensive overview of GenAI applications in KOA imaging, discusses state-of-the-art developments, highlights clinical challenges, and outlines open research problems. Our survey indicates that GenAI has strong potential to improve diagnostic consistency, mitigate data scarcity, and support precision orthopedics.

KEYWORDS: Knee Osteoarthritis, Generative AI, Medical Imaging, GAN, Diffusion Models, KL Grading

1. INTRODUCTION
Knee osteoarthritis (KOA) is one of the most prevalent musculoskeletal disorders and a leading cause of chronic pain, disability, and reduced quality of life among the aging population worldwide. It is a complex, multifactorial degenerative joint disease characterized by the progressive deterioration of articular cartilage, remodeling of subchondral bone, synovial inflammation, and structural alterations in surrounding soft tissues. The disease is influenced by multiple risk factors, including aging, obesity, joint injury, genetic predisposition, mechanical stress, and metabolic disorders. With the continuous increase in life expectancy and obesity rates, the global prevalence of KOA is rising, creating a substantial socioeconomic burden on healthcare systems and highlighting the need for early diagnosis and effective disease management. The pathological progression of KOA begins with biochemical changes in the extracellular matrix of articular cartilage, including degradation of collagen fibers and depletion of proteoglycans, which reduce the cartilage's ability to absorb mechanical loads. Progressive cartilage degeneration increases friction within the joint, leading to cartilage thinning, fissures, and focal defects. Simultaneously, the subchondral bone undergoes adaptive remodeling, resulting in sclerosis, cyst formation, and osteophyte development. Inflammatory mediators released by damaged cartilage and synovial tissues, including cytokines, matrix metalloproteinases (MMPs), and prostaglandins, further accelerate cartilage destruction and contribute to pain sensitization. Chronic synovitis, joint effusion, and meniscal degeneration collectively impair joint stability and function. These pathological changes eventually manifest as radiographic features such as joint space narrowing (JSN), osteophyte formation, bone marrow lesions (BMLs), cartilage defects, and subchondral bone abnormalities, which serve as important imaging biomarkers for disease diagnosiand progression assessment. Medical imaging plays a pivotal role in the diagnosis, severity assessment, monitoring, and treatment planning of KOA. Conventional X-ray remains the most widely used imaging modality because of its low cost, rapid acquisition, and ability to evaluate structural abnormalities such as joint space narrowing, osteophytes, subchondral sclerosis, and bone deformities. Radiographic findings are commonly assessed using the Kellgren–Lawrence (KL) grading system, which serves as the clinical standard for evaluating disease severity. However, radiographs provide only indirect information about cartilage damage and are relatively insensitive to early-stage osteoarthritic changes. Magnetic resonance imaging (MRI) offers superior soft-tissue visualization and enables comprehensive assessment of articular cartilage, menisci, ligaments, synovium, bone marrow lesions, and joint effusion. Owing to its excellent tissue contrast, MRI facilitates the detection of early structural abnormalities before they become apparent on radiographs, making it valuable for disease monitoring and research applications. Computed tomography (CT) provides high-resolution visualization of osseous structures and is particularly useful for evaluating subchondral bone architecture and osteophytes, although its routine clinical use is limited by ionizing radiation exposure. Ultrasound imaging offers a dynamic, real-time,

and cost-effective approach for assessing synovitis, joint effusion, and periarticular soft tissues, while also supporting image-guided therapeutic interventions. Recent advances in artificial intelligence (AI), particularly deep learning and generative AI, have significantly transformed medical image analysis by enabling automated feature extraction, disease classification, severity grading, image reconstruction, data augmentation, image synthesis, and multimodal image translation. Generative AI techniques, including Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), diffusion models, and transformer-based generative architectures, have demonstrated remarkable capability in addressing challenges such as limited annotated datasets, class imbalance, low image quality, missing imaging modalities, and privacy-preserving data generation. These models facilitate the creation of realistic synthetic medical images that enhance model training, improve diagnostic robustness, support image restoration and super-resolution, and enable cross-modal image synthesis between different imaging modalities.

Given the rapid evolution of generative AI technologies and their growing impact on musculoskeletal imaging, a comprehensive understanding of their applications in KOA is both timely and necessary. This review presents a systematic overview of recent developments in generative AI for knee osteoarthritis medical image analysis. It critically examines state-of-the-art generative models, imaging modalities, publicly available datasets, evaluation metrics, clinical applications, and emerging trends while discussing current challenges, limitations, and future research directions. By consolidating recent advances, this review aims to provide researchers, clinicians, and healthcare professionals with a comprehensive reference for understanding the transformative role of generative AI in advancing early diagnosis, severity assessment, and personalized management of knee osteoarthritis.

2. Comparison of Imaging Modalities

Traditional computer-aided diagnosis (CAD) systems for knee osteoarthritis (KOA) relied heavily on handcrafted features and classical machine learning algorithms. Early studies extracted texture-, shape-, and intensity-based descriptors such as GLCM, LBP, HOG, and morphological measurements to capture joint space narrowing, osteophyte presence, and bone texture patterns. These features were typically fed into classifiers including support vector machines (SVM), k-nearest neighbors (k-NN), random forests, and logistic regression to differentiate between healthy and osteoarthritic knees or to perform Kellgren–Lawrence (KL) grading. Although these methods demonstrated moderate performance, their accuracy was limited by handcrafted feature dependence, sensitivity to imaging variations, and inability to capture deeper anatomical and pathological representations. The introduction of deep learning improved performance substantially, but traditional CAD approaches remain important for understanding the evolution of KOA imaging analysis and the challenges—such as data scarcity and variability—that motivated the shift toward more advanced generative and deep learning models.[1] Generative AI (GenAI) in knee osteoarthritis (KOA) imaging stems from several persistent challenges that limit the effectiveness of traditional and deep learning–based diagnostic systems. KOA datasets are often small, imbalanced, and heterogeneous, with limited representation of early-stage disease and severe KL grades. High-quality MRI and X-ray annotations require expert radiologists, making large-scale labeling expensive and time-consuming. Furthermore, imaging variations caused by different scanners, acquisition protocols, and patient positioning introduce domain shifts that reduce model generalizability. GenAI techniques—such as GANs, VAEs, and diffusion models—address these limitations by generating realistic synthetic images, balancing minority classes, enhancing resolution, and enabling cross-modality translation without the need for extensive paired datasets. These capabilities not only reduce reliance on large annotated datasets but also improve robustness, augment diagnostic accuracy, and support the development of more reliable and clinically deployable KOA analysis models. As a result, GenAI has emerged as a powerful tool for overcoming data scarcity, improving model performance, and advancing precision orthopedics. Table 1 shows a comparison of Imaging techniques used in KOA.

Table 1. Comparison of Imaging Techniques Used in KOA

Modality	Strengths	Limitations	Use in KOA
X-ray	Low cost, good for KL grading	Limited soft tissue detail	Osteophytes, JSN
MRI	Excellent soft tissue contrast	Expensive, long scan time	Early diagnosis, BMLs
CT	High bone resolution	Radiation exposure	Bone remodeling analysis
Ultrasound	Real-time, low cost	Operator dependent	Synovitis, effusion

3. LITERATURE SURVEY

Recent advances in Artificial Intelligence (AI) and deep learning have significantly improved the analysis of Knee Osteoarthritis (KOA) medical images. Initially, research primarily focused on conventional convolutional neural networks (CNNs) for disease classification and severity grading. More recently, the emergence of Generative Artificial Intelligence (GenAI), including Generative Adversarial Networks (GANs), diffusion models, and transformer-based architectures, has enabled realistic image synthesis, disease progression modeling, image enhancement, and improved diagnostic performance. El-Ghazaly et al. developed one of the earliest applications of Generative Adversarial Networks (GANs) for synthetic knee radiograph generation. Their study produced approximately 320,000 synthetic (DeepFake) knee X-ray images using 5,556 real radiographs as training data. Validation by fifteen medical experts demonstrated that the generated images were frequently indistinguishable from authentic radiographs. Furthermore, incorporating synthetic images into the training dataset significantly improved KOA severity classification in data-scarce scenarios while preserving patient privacy through synthetic image replacement [2]. Li et al. proposed a deep learning reconstruction (DLR)-based synthetic MRI framework capable of generating both high-quality morphological knee MRI and quantitative T2 maps from a single acquisition. Their method accurately detected early cartilage degeneration while significantly reducing MRI acquisition time. The reported T2 measurements were comparable to conventional MRI, demonstrating the clinical feasibility of synthetic MRI for quantitative cartilage assessment [3]. Zhang et al. conducted the first systematic review and meta-analysis evaluating deep learning algorithms for radiographic KOA grading using the Kellgren–Lawrence (KL) classification system. Their findings indicated that deep learning models consistently achieved high diagnostic sensitivity and specificity, confirming their effectiveness in automated KOA diagnosis while emphasizing the need for standardized evaluation protocols and multicenter validation [4]. Runhaar et al. introduced the KNOAP2020 Challenge, which established the first international benchmark for predicting incident symptomatic radiographic knee osteoarthritis. The challenge compared multiple machine learning and deep learning methods using hidden test datasets collected over a 78-month longitudinal follow-up, providing a standardized framework for evaluating disease progression prediction models [5]. Li et al. proposed the OA-MEN framework, a multi-scale ensemble deep learning architecture integrating ResNet and MobileNet with multi-scale feature fusion. By effectively capturing both local and global semantic information, the proposed model achieved a classification accuracy of 84.88% for KOA detection and Kellgren–Lawrence grading [6]. Wang et al. presented a Vision Transformer (ViT)-based end-to-end framework that jointly analyzes knee radiographs and clinical risk factors. The model effectively captured long-range contextual relationships while maintaining interpretability through attention maps. Furthermore, it demonstrated strong cross-cohort generalization beyond the Osteoarthritis Initiative (OAI), highlighting its potential for clinical deployment and disease progression prediction [7]. Touahema et al. developed MedKnee, an automated software platform for radiographic KOA prediction based on deep learning techniques. Their work demonstrated the feasibility of integrating AI algorithms into user-friendly clinical software, facilitating automated KOA diagnosis in routine orthopedic practice [8]. Yeoh et al. comprehensively reviewed 74 studies involving 2D and 3D convolutional neural networks for KOA diagnosis, classification, and segmentation. Their survey highlighted the transition from traditional 2D CNNs toward 3D deep learning models capable of extracting richer anatomical information while discussing challenges such as computational complexity, data scarcity, and clinical validation [9]. Recent research has increasingly focused on Generative AI for medical image synthesis and disease progression modeling. Al-Mouqdad et al. combined conditional diffusion models with conditional GANs for unsupervised segmentation of bone marrow lesions in knee MRI, demonstrating the effectiveness of hybrid generative frameworks for structural abnormality detection [10]. Similarly, Wang et al. proposed a diffusion-based morphing model capable of generating continuous disease progression sequences from healthy knees to advanced osteoarthritis. The proposed framework enabled realistic simulation of longitudinal structural changes and demonstrated considerable potential for prognosis prediction and disease monitoring [11]. Issa et al. introduced a diffusion-based patella shape synthesis framework capable of preserving patient-specific anatomical characteristics while generating realistic pathological variations. This approach provides promising opportunities for personalized disease progression analysis and prognostic assessment [12]. A recent CycleGAN-based framework further extended generative modeling by synthesizing both historical and future KOA states without requiring paired longitudinal datasets. This approach demonstrated the capability of image-to-image translation techniques for realistic simulation of disease progression and regression, supporting longitudinal KOA analysis and synthetic dataset generation [13]. Overall, the literature demonstrates a clear transition from conventional machine learning and CNN-based diagnostic systems toward advanced Generative AI architectures. While traditional deep learning methods primarily focused on disease classification and segmentation, recent Generative AI approaches have expanded the scope of KOA medical image analysis through realistic synthetic image generation, disease progression modeling, image enhancement, multimodal data augmentation, and privacy-preserving data synthesis. Nevertheless, challenges related to anatomical realism, explainability, computational complexity, standardized evaluation protocols, and clinical validation remain significant barriers to clinical translation, indicating important directions for future research [2]–[13].

4. METHODOLOGY

This review presents a comprehensive analysis of the recent advancements in Generative Artificial Intelligence (Generative AI) for medical image analysis in Knee Osteoarthritis (KOA). The adopted methodology follows a systematic literature review approach consisting of six sequential phases: literature search, study selection, data extraction, classification of Generative AI techniques, comparative analysis, and identification of research gaps and future directions. The primary objective of this methodology is to provide a structured overview of existing research, identify emerging trends, and highlight the potential of Generative AI for improving KOA diagnosis, severity assessment, image synthesis, and clinical decision support. The figure shows the methodology for making a comprehensive review for Medical Image Analysis in Knee Osteoarthritis Using Generative AI .

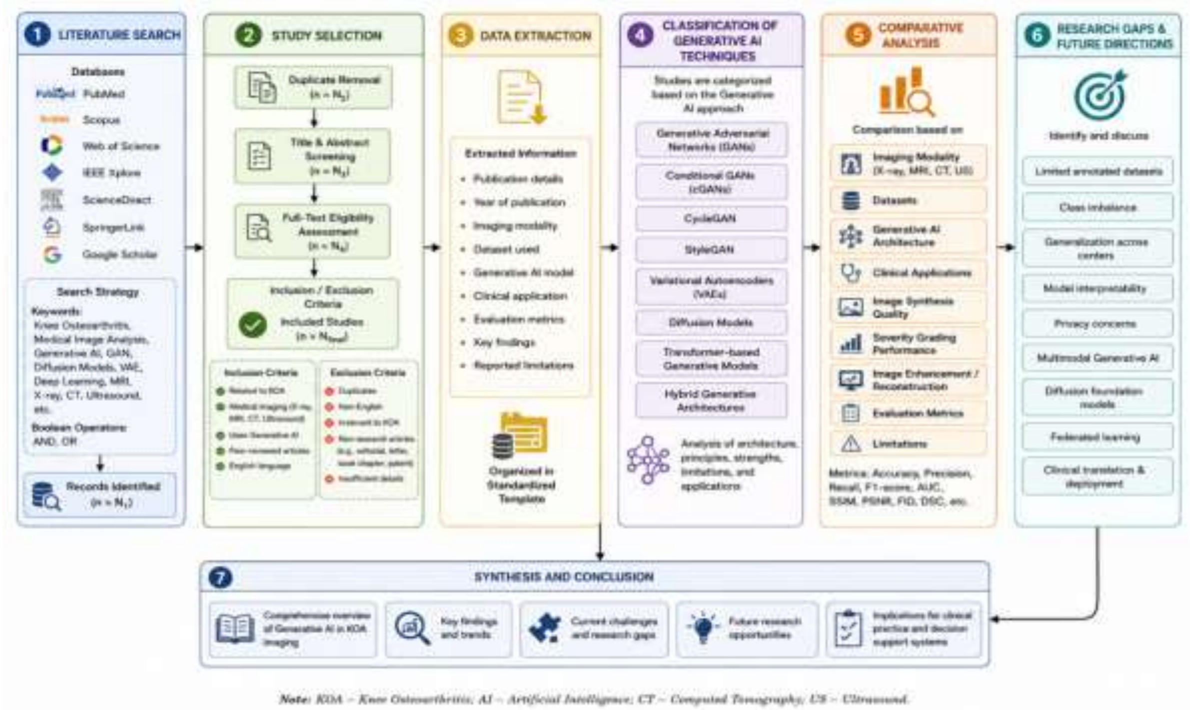

Figure 1. Methodology of Medical Image Analysis in Knee Osteoarthritis Using Generative AI : Comprehensive Review

4.1 Literature Search

A comprehensive literature search was conducted using major scientific databases, including PubMed, Scopus, Web of Science, IEEE Xplore, ScienceDirect, SpringerLink, and Google Scholar. The search focused on peer-reviewed journal articles and conference papers published between 2018 and 2025. Various keyword combinations such as Knee Osteoarthritis, Medical Image Analysis, Generative Artificial Intelligence, Generative Adversarial Networks (GANs), Diffusion Models, Variational Autoencoders (VAEs), Medical Image Synthesis, Deep Learning, X-ray, MRI, and Computed Tomography (CT) were employed using Boolean operators (AND, OR) to retrieve the most relevant publications. Duplicate records were removed before further screening to ensure the uniqueness of the selected studies.

4.2 Study Selection

The retrieved publications were systematically screened based on predefined inclusion and exclusion criteria to ensure the relevance and quality of the reviewed literature. Studies were included if they focused on Knee Osteoarthritis, employed medical imaging modalities such as X-ray, MRI, CT, or Ultrasound, utilized Generative AI techniques for image analysis or synthesis, were published in English, and appeared in peer-reviewed journals or conference proceedings. Publications such as editorials, book chapters, patents, abstracts without full text, duplicate articles, and studies unrelated to KOA or Generative AI were excluded. After title and abstract screening, full-text articles were carefully reviewed to determine their eligibility for inclusion in the final analysis.

4.3 Data Extraction

Relevant information from each selected publication was extracted using a standardized data collection framework to facilitate systematic comparison. The extracted information included publication details, year of publication, imaging modality, dataset used, Generative AI architecture, clinical application, evaluation metrics, key findings, and reported limitations. The collected data were organized into comparative tables to provide a clear overview of existing research and to identify similarities, differences, and research trends among various studies.

4.4 Classification of Generative AI Techniques

The selected studies were categorized according to the underlying Generative AI methodologies employed for KOA medical image analysis. The reviewed approaches included Generative Adversarial Networks (GANs), Conditional GANs (cGANs), CycleGAN, StyleGAN, Variational Autoencoders (VAEs), Diffusion Models, Transformer-based Generative Models, and hybrid generative architectures. Each category was analyzed with respect to its architectural design, working principles, advantages, limitations, and applications such as image synthesis, image enhancement, super-resolution, image reconstruction, cross-modal image translation, data augmentation, and disease severity grading.

4.5 Comparative Analysis

A detailed comparative analysis of the selected studies was performed to evaluate the effectiveness of different Generative AI techniques in KOA medical image analysis. The comparison considered multiple factors, including imaging modality, publicly available datasets, Generative AI architecture, clinical application, image synthesis quality, disease severity grading capability, image enhancement performance, reconstruction accuracy, computational complexity, and clinical applicability. Additionally, commonly reported evaluation metrics such as Accuracy, Precision, Recall, F1-score, Area Under the ROC Curve (AUC), Structural Similarity Index Measure (SSIM), Peak Signal-to-Noise Ratio (PSNR), Fréchet Inception Distance (FID), and Dice Similarity Coefficient (DSC) were summarized to provide a comprehensive assessment of the reported methodologies.

4.6 Research Gaps and Future Directions

Finally, the reviewed literature was critically analyzed to identify existing challenges, research gaps, and future opportunities in the application of Generative AI for KOA medical image analysis. Particular attention was given to limitations such as the scarcity of annotated medical datasets, class imbalance, lack of external validation, limited model interpretability, privacy concerns associated with medical imaging data, and insufficient clinical deployment. Emerging research directions, including multimodal Generative AI, diffusion foundation models, explainable AI, federated learning, privacy-preserving synthetic data generation, and clinically deployable decision-support systems, were also discussed. This systematic methodology provides a comprehensive foundation for understanding the current state of Generative AI in Knee Osteoarthritis imaging and offers valuable insights for future research and clinical translation.

5. Benchmark Datasets

The development and evaluation of Artificial Intelligence (AI) and Generative AI models for Knee Osteoarthritis (KOA) heavily rely on the availability of high-quality, annotated medical imaging datasets. Publicly available benchmark datasets provide standardized imaging data, clinical information, and expert annotations that facilitate the development, validation, and comparison of automated diagnostic algorithms. These datasets include various imaging modalities, such as radiographs (X-rays) and magnetic resonance imaging (MRI), along with clinical metadata, enabling comprehensive assessment of disease progression, severity grading, and structural abnormalities.

Among the available resources, the Osteoarthritis Initiative (OAI) is the most extensively used longitudinal dataset, containing thousands of knee radiographs and MRI scans collected over multiple follow-up visits. It provides expert annotations, including Kellgren–Lawrence (KL) grades, joint space narrowing, cartilage morphology, and clinical outcomes, making it a valuable resource for disease progression analysis and deep learning research. Similarly, the Multicenter Osteoarthritis Study (MOST) offers a large collection of X-ray and MRI images with detailed annotations such as KL grades, MRI Osteoarthritis Knee Score (MOAKS), cartilage lesions, bone marrow lesions, and patient-reported clinical outcomes. This dataset is widely utilized for risk prediction, biomarker discovery, and AI-based disease severity assessment. The Cohort Hip and Cohort Knee (CHECK) dataset primarily focuses on patients with early-stage osteoarthritis and provides longitudinal radiographic data that support studies on early diagnosis and disease progression. The Japanese Osteoarthritis Initiative (JOINT/JOG) dataset contributes valuable imaging data from an Asian population, enabling investigations into ethnic variations, demographic differences, and population-specific disease characteristics. In addition to public datasets, Hospital Picture Archiving and Communication System (PACS) datasets are increasingly employed for external validation and real-world clinical evaluation of AI models. These institutional datasets contain routine clinical images with radiologist annotations and facilitate the assessment of model generalizability across diverse patient populations. With the rapid advancement of Generative AI, these benchmark datasets have become indispensable for developing synthetic medical images, addressing class imbalance, enhancing data augmentation, improving image reconstruction, and enabling multimodal image synthesis. Their standardized annotations and longitudinal follow-up information provide a reliable foundation for training, validating, and benchmarking Generative AI models aimed at improving the accuracy, robustness, and clinical applicability of KOA diagnosis and severity assessment. Consequently, benchmark datasets continue to play a crucial role in accelerating research and fostering the translation of AI-based solutions into clinical practice. Table 2. Shows Comparison of Benchmark Knee Osteoarthritis (KOA) Datasets.

Table 2. Comparison of Benchmark Knee Osteoarthritis (KOA) Datasets

Dataset	Approx. Size	Imaging Modality	Annotations / Labels	Primary Applications
Osteoarthritis Initiative (OAI)	~9,000 participants (millions of X-ray and MRI images across follow-up visits)	X-ray, MRI	Kellgren–Lawrence (KL) grades, joint space narrowing (JSN), cartilage morphology, clinical outcomes	Disease progression analysis, severity grading, deep learning, longitudinal studies
Multicenter Osteoarthritis Study (MOST)	~3,000 participants	X-ray, MRI	KL grades, MOAKS, cartilage lesions, bone marrow lesions, pain scores	Risk prediction, disease progression, biomarker discovery, AI model development

Dataset	Approx. Size	Imaging Modality	Annotations / Labels	Primary Applications
Cohort Hip and Cohort Knee (CHECK)	~1,000 participants	X-ray	KL grades, clinical assessments, longitudinal follow-up	Early KOA diagnosis, progression analysis, epidemiological studies
Japanese Osteoarthritis Initiative (JOINT/JOG)	~2,500 participants	X-ray, MRI	KL grades, demographic information, imaging biomarkers	Ethnic variation analysis, severity assessment, AI-based diagnosis
Hospital PACS Datasets	Varies by institution	X-ray, MRI, CT	Radiologist annotations, KL grades, clinical reports	Clinical validation, external testing, transfer learning, real-world deployment

6. Comparative Analysis of Generative AI Models

Recent advances in Generative Artificial Intelligence (GenAI) have significantly transformed medical image analysis by enabling realistic image synthesis, data augmentation, image enhancement, disease progression modeling, and cross-modal image translation. In Knee Osteoarthritis (KOA), Generative AI has become an important tool for addressing challenges such as limited annotated datasets, class imbalance, privacy concerns, and variability in disease severity. Among the various generative architectures, Generative Adversarial Networks (GANs), Diffusion Models, and Autoencoder-based approaches have emerged as the most widely adopted techniques. Each model exhibits distinct characteristics in terms of image quality, computational complexity, interpretability, training stability, and clinical applicability. A comprehensive comparison of these approaches is presented in the following sections.

6.1 GAN-Based Models

Generative Adversarial Networks (GANs) are among the earliest and most extensively studied Generative AI models for KOA medical image analysis. A GAN consists of two competing neural networks, namely a generator and a discriminator, which are trained simultaneously through an adversarial learning process. The generator aims to produce realistic synthetic knee radiographs, while the discriminator distinguishes between real and generated images. Through continuous competition, the generator progressively learns to synthesize high-quality images that closely resemble clinical radiographs. Several GAN variants, including Conditional GAN (cGAN), CycleGAN, and StyleGAN, have been successfully applied to KOA imaging. Conditional GANs generate images conditioned on specific Kellgren–Lawrence (KL) grades, enabling disease-specific image synthesis. CycleGAN facilitates image-to-image translation between different disease stages without requiring paired datasets, while StyleGAN produces high-resolution radiographs with enhanced anatomical detail. GAN-based approaches have demonstrated considerable success in data augmentation, class imbalance reduction, and synthetic image generation, thereby improving the performance of deep learning classifiers. However, these models are often affected by training instability, mode collapse, sensitivity to hyperparameter selection, and the generation of subtle anatomical artifacts, which may limit their direct application in clinical practice.

6.2 Diffusion Models

Diffusion models have recently emerged as a powerful alternative to GANs for medical image synthesis due to their superior image quality and stable training process. These models generate synthetic images by progressively removing noise from a random distribution through a sequence of denoising steps, resulting in highly realistic and anatomically consistent knee radiographs. Conditional diffusion models further enable the generation of images corresponding to different KL grades, allowing continuous simulation of disease progression from healthy joints to severe osteoarthritis. This capability is particularly valuable for studying longitudinal disease evolution and generating balanced datasets for underrepresented severity levels. Recent studies have demonstrated that diffusion-generated images substantially improve classifier performance, disease severity prediction, and image reconstruction quality while producing fewer visual artifacts than GAN-generated images. Furthermore, identity-preserving diffusion models can maintain patient-specific anatomical characteristics while modifying only pathological changes, making them highly suitable for personalized prognosis and clinical decision-support systems. Despite these advantages, diffusion models require significantly greater computational resources, longer training times, and higher memory consumption than GAN-based approaches.

6.3 Autoencoder-Based and Hybrid Generative Models

Autoencoder-based generative models, particularly Variational Autoencoders (VAEs), provide an alternative approach by learning compact latent representations of knee images rather than focusing solely on photorealistic image generation. These models encode medical images into a lower-dimensional latent space and subsequently reconstruct them while preserving clinically relevant information. Hybrid generative frameworks further combine autoencoders with adversarial learning, contrastive learning, or diffusion mechanisms to improve image quality and latent feature representation simultaneously. Unlike GANs and diffusion models, autoencoder-based methods emphasize representation learning and interpretability, making them particularly suitable for early disease detection, biomarker discovery, anomaly detection, and explainable artificial intelligence. Although the generated images are generally less realistic than those produced by

GANs or diffusion models, the structured latent space enables better understanding of disease characteristics and supports downstream clinical applications requiring transparent decision-making.

6.4 Comparative Discussion

The selection of an appropriate Generative AI model largely depends on the intended clinical application. GAN-based models are highly effective for generating realistic synthetic radiographs and expanding limited training datasets through data augmentation. However, their susceptibility to unstable training and anatomical inconsistencies necessitates careful quality assessment before clinical use. Diffusion models currently represent the state of the art in medical image synthesis because of their exceptional image fidelity, stable optimization process, and ability to model continuous disease progression while preserving anatomical structures. These characteristics make diffusion models particularly attractive for disease progression analysis, personalized prognosis, and synthetic dataset generation, although their computational requirements remain relatively high. In contrast, autoencoder-based and hybrid generative approaches prioritize interpretability and structured feature learning over photorealistic image generation, making them valuable for explainable AI, representation learning, and early KOA diagnosis. Overall, no single Generative AI architecture is universally optimal. GANs remain a practical choice for efficient data augmentation, diffusion models offer superior image realism and progression modeling, and autoencoder-based methods provide enhanced interpretability and feature representation. Future research is expected to focus on hybrid generative architectures that combine the strengths of these approaches to develop robust, explainable, and clinically deployable AI systems for Knee Osteoarthritis medical image analysis.

7 Performance Evaluation

Recent studies have demonstrated that the integration of Generative Artificial Intelligence (GenAI) substantially enhances the performance of artificial intelligence models for Knee Osteoarthritis (KOA) diagnosis and severity assessment. Compared with conventional deep learning models trained solely on real medical images, GenAI-based approaches improve dataset diversity, reduce class imbalance, and generate realistic synthetic images that strengthen model generalization. These advantages are particularly important for underrepresented disease stages, where limited annotated data often restrict the performance of supervised learning algorithms.

Among the various Generative AI techniques, Generative Adversarial Networks (GANs) have been widely employed for medical image augmentation and synthetic radiograph generation. The literature indicates that GAN-generated images improve the classification performance of KOA models by increasing the representation of minority KL grades and enhancing feature learning. However, GAN-based methods are often affected by unstable adversarial training, mode collapse, and occasional anatomical inconsistencies, which may reduce the reliability of the generated images and limit their clinical applicability.

More recently, Diffusion Models have emerged as the state-of-the-art approach for medical image synthesis. Studies consistently report that diffusion-generated knee radiographs exhibit superior anatomical realism, improved structural consistency, and fewer visual artifacts compared with GAN-generated images. Consequently, classifiers trained using diffusion-augmented datasets generally demonstrate better diagnostic accuracy, improved robustness, and enhanced disease severity prediction. Furthermore, conditional and identity-preserving diffusion models enable realistic simulation of disease progression while maintaining patient-specific anatomical characteristics, making them highly suitable for longitudinal KOA analysis and personalized clinical decision support.

Autoencoder-based and hybrid generative models have also been explored for KOA image analysis. Although these methods generally produce images with lower visual realism than GANs or diffusion models, they provide structured latent representations that improve feature learning, anomaly detection, and model interpretability. Hybrid architectures combining autoencoders with adversarial or diffusion-based learning have shown promising results by balancing image quality with explainable feature extraction.

Overall, the reviewed literature indicates that diffusion-based Generative AI models currently provide the most consistent performance for synthetic medical image generation and disease progression modeling, whereas GAN-based approaches remain effective for data augmentation due to their computational efficiency. Autoencoder-based models contribute primarily through improved representation learning and explainability. The selection of an appropriate Generative AI architecture therefore depends on the intended application, available computational resources, and the desired balance between image realism, interpretability, and computational complexity .Table 3. Shows summary of major research articles on Knee Osteoarthritis using deep learning and generative AI.

Table 3. Descriptive summary of major research articles on Knee Osteoarthritis using deep learning and generative AI.

Sr.	Article (Reference)	Methods Used	Key Results	Limitations
1	Panwar et al., Sci Rep (2024)	Deep Stacking Ensemble combining CNN, AlexNet, ResNet34 and ResNet50 for KOA MRI grading; augmented MRI dataset.	Achieved 99.71% accuracy in KOA severity prediction; ensemble outperformed all individual models.	MRI-only study; lacks multi-centre validation and comparison with radiologists.

Sr.	Article (Reference)	Methods Used	Key Results	Limitations
2	El-Ghazaly et al., Sci Rep (2022)	WGAN-GP–based DeepFake KOA X-ray synthesis for augmentation; expert survey + transfer learning classification.	DeepFake images fooled experts (61.35% accuracy). Augmentation (+200%) improved classification accuracy to 75.76%.	Synthetic images contain artifacts; unpaired GAN training may reduce anatomical realism.
3	Li et al., Insights Imaging (2025)	DL-based reconstruction for Synthetic MRI (SyMRI + Deep Learning Reconstruction).	High agreement with conventional MRI T2 mapping; improved ability to detect cartilage changes.	Single-centre dataset; no arthroscopy ground truth; limited generalisability.
4	Zhang et al., Eur Radiol (2024)	Meta-analysis of DL-based KL grading from radiographs (multi-class classification).	Pooled sensitivities: KL0–86.7%, KL1–64%, KL2–75%, KL3–84.7%, KL4–90.3%. DL shown effective overall.	High heterogeneity across studies; limited comparison with human experts.
5	Runhaar et al., Osteoarthritis Cartilage (2022) — KNOAP2020	Multi-modal ML challenge using X-ray, MRI and clinical data to predict incident KOA.	Best ROC AUC 0.636; prediction of KOA onset remains difficult even with MRI + clinical risk factors.	Models show low discriminative power; datasets restricted to specific demographics.
6	Li et al., Front Bioeng Biotechnol (2024) — OA-MEN	Hybrid multi-scale ensemble combining ResNet + MobileNet; KOA severity grading.	Accuracy 84.88%; strong performance on KL1 grading (typically difficult).	KL1 still inconsistently classified; limited external testing.
7	Wang et al., Diagnostics (2025)	Vision Transformer (ViT) with radiographs + clinical factors to predict KOA progression.	AUROC 0.808 on OAI dataset; external dataset AUROC 0.709. Good sensitivity for progression prediction.	Performance decreased on external cohort; population bias exists.
8	Touahema et al., Diagnostics (2024) — MedKnee	Xception-based automated KL grading software with GUI; trained on OAI dataset.	Test accuracy 97.20%; high precision for KL3; performed well across two external datasets.	Requires manual ROI selection; trained only on pre-processed fixed-flexion X-rays.
9	Yeoh et al., CIN (2021) — Review	Survey of 74 DL-based studies on KOA segmentation and classification (2D/3D CNN).	Found DL performance comparable to human graders; 3D CNNs promising for early detection.	Search limited to specific databases; potential selection bias.
10	Al-Mouqdad et al., Bioengineering (2022)	Unsupervised Bone Marrow Lesion (BMEL) segmentation using conditional diffusion models + cGANs.	Best DICE ≈ 0.22; shows feasibility of generative models for multi-sequence MRI segmentation.	Very low segmentation accuracy; limited annotator agreement; no external validation.
11	Wang et al., arXiv (2024) — Diffusion Morphing Model	DDPM-based morphing along KL0→KL4 progression path; intermediate X-ray generation.	Improved classification (KL1–KL2 +3.84%); radiologists rated synthetic frames 3.9/5.	Not validated beyond OAI dataset; imaging variability affects stability.
12	Issa et al., MOST (2025) — SynPatNet	Diffusion model for predicting 60-month patella shape evolution; CNN-based prognostic scoring.	AUC 0.909 for patellofemoral OA onset; AUC 0.823 for knee replacement prediction.	Underrepresents fast-progressing cases; relies on uniform imaging; U.S.-specific cohort.

Sr.	Article (Reference)	Methods Used	Key Results	Limitations
13	CycleGAN KOA Progression (2025)	CycleGAN to model past/future KOA radiographs; evaluated using EfficientNetV2 classifier.	Future-state predictions correctly classified 83.7% of mild cases; demonstrated disease trajectory mapping.	GAN may hallucinate structures; unpaired training reduces anatomical reliability.

8. Challenges and Open Research Issues

Despite the remarkable progress achieved by Generative Artificial Intelligence (GenAI) in Knee Osteoarthritis (KOA) medical image analysis, several technical, clinical, and ethical challenges continue to hinder its widespread adoption in routine healthcare practice. Although Generative AI has demonstrated significant potential for synthetic image generation, data augmentation, disease progression modeling, and image enhancement, important limitations remain regarding data availability, model robustness, interpretability, clinical validation, and regulatory compliance. Addressing these challenges is essential for developing reliable, explainable, and clinically deployable AI-based decision-support systems.

8.1 Limited Availability of High-Quality Annotated Data

One of the primary challenges in KOA medical image analysis is the limited availability of large-scale, high-quality annotated datasets. Although publicly available datasets such as the Osteoarthritis Initiative (OAI) and the Multicenter Osteoarthritis Study (MOST) provide extensive collections of knee radiographs and MRI scans, their annotations are primarily limited to Kellgren–Lawrence (KL) grades and a small number of structural biomarkers. Detailed annotations describing cartilage lesions, osteophytes, meniscal degeneration, bone marrow lesions, and synovial abnormalities remain scarce because they require considerable time and expertise from experienced musculoskeletal radiologists. Furthermore, advanced disease stages (KL Grade 3 and Grade 4) are relatively underrepresented, resulting in severe class imbalance that limits the effectiveness of supervised learning algorithms and Generative AI models.

8.2 Anatomical Realism of Synthetic Images

The generation of anatomically accurate synthetic medical images remains a significant challenge for Generative AI models. Although GANs and diffusion models can produce visually realistic knee radiographs, they may occasionally generate subtle anatomical inconsistencies, distorted bone boundaries, unrealistic cartilage morphology, or artificial texture patterns. These inaccuracies can adversely affect downstream diagnostic models and reduce clinician confidence in synthetic data. Therefore, ensuring anatomical fidelity and preserving clinically meaningful structural characteristics remain essential requirements for the safe application of Generative AI in KOA diagnosis.

8.3 Patient Identity Preservation and Disease Progression Modeling

Another important challenge involves preserving patient-specific anatomical characteristics while simulating disease progression. Existing generative models often modify anatomical structures beyond the intended pathological changes, thereby reducing their suitability for longitudinal disease monitoring. Identity-preserving image synthesis is particularly important for predicting disease progression, generating follow-up examinations, and supporting personalized treatment planning. Developing generative models capable of maintaining individual anatomical identity while accurately simulating gradual structural deterioration remains an active area of research.

8.4 Lack of Standardized Evaluation Protocols

The evaluation of synthetic medical images continues to be an open research problem. Conventional image quality metrics, including Fréchet Inception Distance (FID), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and Learned Perceptual Image Patch Similarity (LPIPS), primarily measure visual similarity but do not adequately assess clinical realism or diagnostic utility. Consequently, many studies rely on subjective evaluation by expert radiologists, which is time-consuming, expensive, and difficult to standardize. The development of standardized evaluation protocols that integrate quantitative image quality measures with clinically meaningful assessment criteria remains a major research priority.

8.5 Generalization Across Clinical Centers

Most Generative AI models are trained using images collected from specific institutions or publicly available datasets, limiting their ability to generalize across different healthcare environments. Variations in imaging equipment, acquisition protocols, patient demographics, image resolution, and clinical practice often lead to significant domain shifts that reduce model performance when applied to external datasets. Improving domain adaptation, cross-institutional robustness, and federated learning techniques will be essential for developing clinically reliable and widely deployable Generative AI systems.

8.6 Computational Complexity and Resource Requirements

Although diffusion models currently achieve state-of-the-art performance in medical image synthesis, they require substantial computational resources, large memory capacity, and extended training times. These requirements restrict their implementation in resource-constrained healthcare facilities and real-time clinical applications. Future research should focus on lightweight generative architectures, efficient sampling strategies, model compression, and knowledge distillation techniques to improve computational efficiency without compromising image quality.

8.7 Explainability and Clinical Trust

The limited interpretability of deep generative models remains a significant barrier to clinical adoption. While visualization techniques such as Grad-CAM, attention maps, and saliency methods provide partial explanations of model predictions, they do not fully explain how synthetic images are generated or how disease progression is simulated. Improving model transparency, interpretability, and explainability is essential for increasing clinician confidence, supporting regulatory approval, and facilitating the integration of Generative AI into routine clinical workflows.

8.8 Ethical, Privacy, and Regulatory Considerations

The increasing use of synthetic medical images also raises important ethical and legal concerns. Although synthetic data can enhance privacy by reducing reliance on real patient images, there remains a risk of generating misleading images, introducing algorithmic bias, or creating synthetic data that could be misused. Ensuring patient privacy, fairness, transparency, and compliance with healthcare regulations is therefore essential. Future Generative AI systems should incorporate privacy-preserving learning techniques, secure data-sharing mechanisms, and robust governance frameworks to promote responsible clinical deployment.

8.9 Integration of Multimodal Clinical Information

Current Generative AI research primarily focuses on single-modality imaging, particularly radiographs and MRI scans. However, clinical decision-making often requires the integration of multiple data sources, including imaging findings, electronic health records, laboratory biomarkers, genetic information, gait analysis, and patient-reported outcomes. Developing multimodal Generative AI models capable of simultaneously learning from heterogeneous clinical data represents an important research direction that could significantly improve diagnostic accuracy, disease severity assessment, and personalized treatment planning.

8.10 Limited Clinical Validation and Real-World Deployment

Although numerous Generative AI models have demonstrated promising results on publicly available benchmark datasets, relatively few have undergone rigorous clinical validation. Most published studies remain limited to retrospective analyses using curated research datasets, with only a small number evaluating performance in real-world hospital environments. Prospective clinical trials, multi-center validation studies, radiologist-in-the-loop evaluations, and regulatory assessments are required before Generative AI systems can be safely integrated into routine orthopedic and radiological practice. Establishing clinically validated, trustworthy, and explainable AI systems remains one of the most important future objectives in KOA medical image analysis.

9. CONCLUSION

Generative AI (GenAI) has emerged as a powerful paradigm in medical image analysis, offering significant advancements in the study and management of Knee Osteoarthritis (KOA). This review examined the progression of GenAI models, ranging from traditional Generative Adversarial Networks (GANs) to the latest diffusion-based and hybrid architectures. The literature demonstrates that GenAI enables the synthesis of anatomically realistic knee radiographs, supports balanced data augmentation, and enhances downstream tasks such as KL grade classification, disease progression prediction, and multimodal fusion.

Despite these promising capabilities, several challenges must be addressed before GenAI can be translated into routine clinical practice. These include ensuring strict anatomical fidelity in synthetic images, maintaining patient-specific identity during progression modeling, and establishing standardized evaluation frameworks tailored to KOA imaging. Moreover, the computational cost of diffusion models, limited availability of richly annotated datasets, and gaps in explainability and clinical trust present additional barriers. Ethical considerations such as privacy, data misuse, and regulatory compliance further emphasize the need for responsible deployment of GenAI technologies.

Overall, the rapid developments in GenAI indicate its strong potential to reshape KOA diagnosis, monitoring, and personalized decision support. Future research will benefit from closer collaboration between clinicians and AI researchers, improved dataset curation, and development of interpretable, efficient, and clinically validated generative models. As these challenges are addressed, GenAI is poised to become a cornerstone technology in the next generation of KOA imaging research and clinical applications.

REFERENCE

1. . Panwar, et al., "Optimizing knee osteoarthritis severity prediction on MRI images using deep stacking ensemble technique," Scientific Reports, vol. 14, no. 1, 2024.
2. S. El-Ghazaly, et al., "DeepFake X-rays from generative adversarial networks deceive medical experts and improve classification with scarce data," Scientific Reports, vol. 12, 2022.
3. Y. Li, et al., "Deep learning N reconstruction in synthetic MRI for knee cartilage assessment," Insights into Imaging, vol. 16, 2025.

4. J. Zhang, et al., "Diagnostic performance of deep learning in radiographic grading of knee osteoarthritis: A systematic review and meta-analysis," European Radiology, 2024.
5. J. Runhaar, et al., "KNOAP2020 challenge: Predicting incident symptomatic radiographic knee osteoarthritis," Osteoarthritis and Cartilage, vol. 30, no. 12, pp. 1617–1628, 2022.
6. Z. Li, et al., "OA-MEN: A multi-scale ensemble network for knee osteoarthritis detection and KL grading," Frontiers in Bioengineering and Biotechnology, vol. 12, 2024.
7. C. Wang, et al., "A vision transformer-based deep learning approach for knee osteoarthritis prognosis," Diagnostics, vol. 15, 2025.
8. Touahema, et al., "MedKnee: A deep learning-based software for automated prediction of radiographic knee osteoarthritis," Diagnostics, vol. 14, 2024.
9. D. Yeoh, et al., "Emergence of deep learning in knee osteoarthritis diagnosis," Computational Intelligence and Neuroscience, vol. 2021, Article ID 4931437, 2021.
10. Al-Mouqdad, et al., "Unsupervised bone marrow lesion segmentation in knee MRI using conditional diffusion models and cGANs," Bioengineering, vol. 11, 2024.
11. H. Wang, et al., "Temporal evolution of knee osteoarthritis: A diffusion-based morphing model for X-ray medical image synthesis," arXiv preprint arXiv:2408.00891, 2024.
12. S. Issa, et al., "Diffusion-based patella shape synthesis for prognostic assessment of knee osteoarthritis," 2025.
13. "CycleGAN-based synthesis of past and future knee osteoarthritis states," Medical Image Analysis, 2025.
14. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in Proc. MICCAI, 2015, pp. 234–241.
15. Oktay, et al., "Attention U-Net: Learning where to look for the pancreas," arXiv preprint arXiv:1804.03999, 2018.
16. Dosovitskiy, et al., "An image is worth 16×16 words: Transformers for image recognition at scale," in International Conference on Learning Representations (ICLR), 2021.
17. J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in Advances in Neural Information Processing Systems (NeurIPS), vol. 33, pp. 6840–6851, 2020.
18. J. Song, C. Meng, and S. Ermon, "Denoising diffusion implicit models," in International Conference on Learning Representations (ICLR), 2021.Goodfellow, et al., "Generative adversarial nets," in Advances in Neural Information Processing Systems (NeurIPS), vol. 27, 2014.
19. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 4401–4410.
20. D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," in International Conference on Learning Representations (ICLR), 2014.