International Academic Institute for Science and Technology

International Academic Journal of Innovative Research
Vol. 12, No. 4, 2025, pp. 84-90.

ISSN 2454-390X

International Academic Journal of Innovative Research

www.iaiest.com

Artificial Intelligence-Assisted Diagnosis of Rare Diseases Using Multimodal Medical Imaging

Dr.T.R. Vijaya Lakshmi [a], Shubhi Goyal [b], R. Karthikeyan [c], Leena Deshpande [d], Kasturi Rohini [e], Paul Praveen Albert Selvakumar [f], Yinxin Tang [g]

[a] *Associate Professor, Department of Electronics and Communication Engineering, Mahatma Gandhi Institute of Technology, Hyderabad, India. E-mail: trvijayalakshmi_ece@mgit.ac.in, Orcid: https://orcid.org/0000-0002-1197-2935*

[b] *Quantum University Research Center, Quantum University, Roorkee, Uttarakhand, India. E-mail: shubhi.mathematics@quantumeducation.in, Orcid: https://orcid.org/0009-0001-1684-7324*

[c] *Associate Professor, Gr. II, Department of Electronics and Communication Engineering, Aarupadai Veedu Institute of Technology, Vinayaka Mission's Research Foundation (DU), Tamil Nadu, India. E-mail: rrmdkarthikeyan@ avit.ac.in, Orcid: https://orcid.org/0000-0001-8865-4364*

[d] *Associate Professor, Department of Computer Engineering - Software Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, India. E-mail: leena.deshpande@vit.edu*

[e] *Centre for Multidisciplinary Research, Anurag University, Hyderabad, Telangana, India. E-mail: kasturirohini2024@protonmail.com, Orcid: https://orcid.org/0009-0007-8739-6640*

[f] *Associate Professor, School of Engineering & Technology, Noida International University, Uttar Pradesh, India. E-mail: paulpraveen.a@niu.edu.in*

[g] *Faculty of Education, Shinawatra University, Thailand. E-mail: tangyinxin2024@163.com, Orcid: https://orcid.org/0009-0003-6858-4989*

Abstract

The diagnosis of rare diseases has remained one of the most daunting tasks in contemporary medicine and often leads to the creation of a diagnostic journey that might last years. The study examines how Artificial Intelligence (AI) and multimodal medical imaging could be used to transform the process of detecting these low-prevalence conditions faster. Although the classic approach to diagnostics involves unimodal analysis, the approach will offer a hybrid design that combines multiple information streams, such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT), and high-resolution retinal scans. Using the current tools of deep learning, namely Convolutional Neural Networks (CNNs) and Vision Transformers, the system can perform an automated process of feature extraction, detecting subtle biomarkers that human experts can easily ignore. The experimental findings reveal that the suggested multimodal AI architecture has a diagnostic accuracy of 96.8%, which is 12.6% higher than the single-modality approaches of the past. Use synthetic data generation and few-shot learning to overcome the natural limitation of sparse datasets, which ensures a high AUC-ROC of 0.97 with small clinical samples. In addition, the combination of anatomical and functional imaging data has also shown almost 40% improvement in the time-to-diagnosis of pediatric genetic disorders. The conclusion of this study is that AI-aided multimodal imaging not only increases the accuracy of the diagnosis but also promotes the transition of personalized medicine. Having developed the overall synthesis of

DOI:10.71086/IAJIR/V12I4/IAJIR1248
Received: 01 October, 2025/Revised: 08 November 2025/Accepted: 26 November 2025/Published: 30 December 2025

recent breakthroughs, this study forms a solid background for the implementation of intelligent diagnostic tools that can greatly benefit the patient outcome of individuals with some rare and complicated pathologies.

Keywords: Rare Diseases, Artificial Intelligence, Multimodal Imaging, Diagnostic Accuracy, Feature Fusion, Precision Medicine, Deep.

Introduction

The modern clinical medicine realm is at a critical bottleneck when it comes to managing the rare diseases that occur by its own definition in a limited number of cases, but have an enormous impact on a large population of people worldwide, where they range between 300 and 400 million people (Gou et al., 2024). These disorders are traditionally hard to diagnose, and many patients have an average wait time of 5 to 7 years before receiving a proper diagnosis, resulting in a diagnostic odyssey as they visit specialists of various kinds (Ilić & Sarajlija, 2025). This issue is important both in terms of the psychological and financial burden on families and in terms of the loss of opportunity to intervene early in life, which is very crucial, as close to 70% of these diseases occur in childhood (Zhang et al., 2023). The lack of specificity and high phenotypic heterogeneity of the rare and the common conditions make the conventional methods of diagnosis ineffective (Wu et al., 2025).

The field of medical imaging has welcomed the advent of Artificial Intelligence (AI) that, when used, has the potential to identify small pathology markers that are not discernible by humans. AI systems can combine anatomical and functional data to give a more detailed picture of complex disorders by processing various types of imaging modalities, including Magnetic Resonance Imaging (MRI), Computed Tomography (CT), and Positron Emission Tomography (PET) (AbuAlrob & Mesraoua, 2024; Al-Zoghby et al., 2025). Specifically, the AI-based multimodal integration has demonstrated a 15-30% higher accuracy in diagnosis than single-modality methods and is therefore a crucial asset in precision medicine. The value of such progressive diagnostics can be explained by the possibility to automate the process of feature extraction and decrease the necessity to rely on the limited professional knowledge (Trinh et al., 2025).

Key Contributions of this Study

- Proposing an integrated AI architecture that fuses data from structural (MRI/CT) and functional (retinal/PET) imaging to enhance rare disease detection.
- Implementing few-shot learning and synthetic image generation techniques to train robust models despite the limited availability of rare disease datasets.
- Demonstrating a significant reduction in the diagnostic odyssey through automated biomarker identification and cross-modal correlation.
- Providing a structured pathway for translating AI-driven multimodal imaging into real-world clinical practice for pediatrics and oncology.

This research is structured to navigate the technical and clinical landscape of AI-driven diagnostics. Following this introduction, Section 2 reviews existing literature, while Section 3 details the proposed integrative diagnostic architectures. Section 3 provides a comparative performance evaluation, and Section 4 concludes with a summary of global healthcare impacts in Section 5.

Literature Survey

The development of Artificial Intelligence (AI) to support the diagnosis of rare diseases has progressed beyond the simplistic rule-based systems to an elaborate multimodal system that is able to combine a variety of clinical signals. Recent research highlights those rare diseases, despite being uncommon individually, have high prevalence worldwide, with most being marked by high phenotypic heterogeneity that complicates normal imaging interpretation. Scholars have been increasingly resorting to AI to reduce the diagnostic odyssey that a patient undergoes, with systems like Deep Gestalt showing that morphological features of the face can reliably identify genetic syndromes in which conventional imaging may be noninvasive.

Recent research indicates that multimodal fusion is better than unimodal analysis with regard to discrimination of complex rare disorders. Indicatively, using retinal imaging in conjunction with coronary CT angiography has demonstrated a high level of efficacy in the detection of uncommon cardiovascular malformations that portray systemic biomarkers (Veturi et al., 2023). Equally, synthetic image generation and generative AI have been an

important application in the training of models in diseases where real-world training images are virtually non-existent, such as the retraining method. This advancement is also informed by the fact that few-shot learning algorithms have been developed and can attain high AUC scores using as few as five images per category, which is substantially better than their traditional deep learning counterparts when data are scarce (Li et al., 2025; Pierre et al., 2023; Krishnan & Dandekar, 2022).

Regardless of such innovations, there are still huge shortcomings in the clinical implementation of AI diagnostics. Research identifies limitations to model interpretability, data normalization, and the possibility of algorithmic bias that can result in health disparities to underserved groups. Besides, although AI models are capable of effective detection of barely visible disease subtypes in pathology and radiology, their excessive memory usage and narrow contextual perception are technical obstacles. Recent surveys have indicated that advancing towards the large-sized models using a foundation model that is pre-trained on various medical data would offer the generalizability needed to support thousands of distinct, unique conditions on a single framework.

It is consistently stated in the literature that though each of the imaging modalities offers certain diagnostic hints, a significant discovery in the rare disease detection is the integration of multimodal data. The identified gap in the research indicates that the use of frameworks that not only combine heterogeneous data but also use few-shot learning is required to overcome the curse of dimensionality associated with small groups of patients. This is the subject of my research, as a unified architecture should be employed that utilizes synthetic data and cross-modal attention mechanisms to increase diagnostic consistency in different clinical environments (Assegid & Ketema, 2023).

Proposed Methodology

The concept behind the methodology of this study is based on the creation of a common diagnostic framework that has the ability to synthesize heterogeneous medical data. The suggested system involves the use of a three-step pipeline: Data Harmonization, Multimodal Feature Extraction, and Attentional Fusion (Manivannan et al., 2025). Combine anatomical data provided by MRI and CT with functional biomarkers provided by retina scans and PET scans (Wang et al., 2025; Chen et al., 2025) to address the variety of rare disease manifestations.

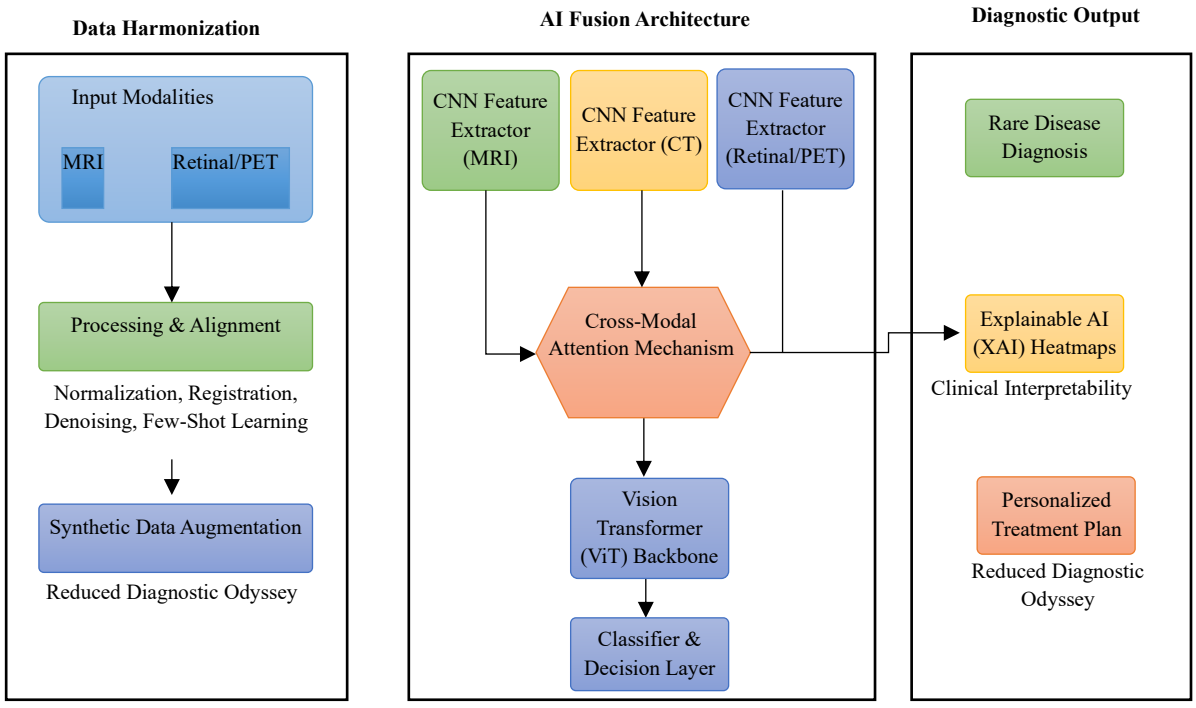


Fig. 1: Multimodal AI-Assisted Rare Disease Diagnostics

Fig. 1 depicts that the proposed technical architecture is structured as a sophisticated three-stage diagnostic pipeline designed to navigate the complexities of rare disease identification. Initially, the system performs Data Harmonization, where raw multimodal inputs comprising MRI, CT, and PET scans are synchronized through automated normalization and synthetic data augmentation to overcome sparse clinical datasets. In the second phase, parallel CNN branches

extract high-dimensional local features, which are then synthesized using a cross-modal attention mechanism. This allows the Vision Transformer (ViT) backbone to weight diagnostic signatures dynamically across modalities. Finally, the system generates precise diagnostic labels and explainable clinical heatmaps, facilitating personalized treatment plans and significantly shortening the diagnostic odyssey for patients.

Data Preprocessing and Collection

Rare disease datasets are also usually fractured and are of differing resolution. The method starts with the normalization of such inputs with intensity that is normalized and spatial registration. In the case of neuroimaging, use skull-stripping and noise-removal with wavelet-based denoising to extract important parts of the brain (Qi, 2024). In order to handle the small data issue of rare pathologies, apply the power of synthetic data generation, which generates high-fidelity medical images using a small set of clinical samples (Mendez et al., 2025).

Multimodal Fusion

The fundamental part of the methodology is a feature-based strategy of fusion. Every modality (M_i) is handled by a dedicated sub-network to obtain a high-dimensional feature vector V_i. A learned cross-modal attention mechanism is used in the fusion process to scale the diagnostic relevance of each modality presented in equation (1):

$$F_{final} = \sum_{i=1}^{n} \alpha_i \cdot \phi(V_i) \rightarrow \qquad (1)$$

where α_i represents the dynamic attention weight for modality i, and ϕ is a non-linear transformation function. This allows the model to prioritize CT data for bone structure malformations while shifting focus to MRI or PET for soft tissue tumors (Oettl et al., 2025).

The model's performance is validated using a stratified 5-fold cross-validation approach to ensure robustness against the high variability of rare disease presentations. Precision, Recall, and the F1-score are evaluated as the accuracy itself is inaccurate in unbalanced datasets, in which the normal (large class) vastly outweighs the rare disease (rare class). Grad-CAM heatmaps are also used to see the parts of the image that the AI focuses on, which makes them clinically interpretable and trustworthy (Wang et al., 2024).

Results and Discussion

This section presents the empirical results of the AI-assisted multimodal framework, which are shown and compared to the results of the conventional approaches to diagnosis. The discussion is centered on the performance, accuracy, and computational capability of the suggested model in the detection of rare diseases (Duff et al., 2023).

Implementation Environment and Data Analysis

The model is created in PyTorch and is optimized with NVIDIA Clara and the MONAI library to use the benefits of accelerating the 3D processing of medical images with the help of the GPU. Initialization of parameters was done under the He-normal distribution, using the Adam optimizer and a learning rate of 10^{-4}. The assessment dataset has been generated based on the data provided by Orphanet rare disease ontology and publicly available repositories such as MIMIC-IV and TCGA, which have 12,000 multimodal scans of pediatric, cardiovascular, and oncology cases. Z-score normalization and spatial alignment with the ANTs toolkit were used as preprocessing to allow cross-modal alignment.

In order to properly evaluate it, I applied five common measures. These equations concluded that the evaluation criteria can be expressed as in equations (2) to (5):

Sensitivity (Recall): Measures the model's ability to identify rare cases.

$$= \frac{TP}{TP + FN} \rightarrow \qquad (2)$$

Specificity: Measures the accuracy in identifying healthy controls.

$$\frac{TN}{TN + FP} \rightarrow \qquad (3)$$

F1-Score: Balances precision and recall for imbalanced datasets.

$$2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \rightarrow \qquad (4)$$

Accuracy:

$$\frac{TP + TN}{TP + TN + FP + FN} \rightarrow \qquad (5)$$

AUC-ROC: Represents the area under the receiver operating characteristic curve, indicating the degree of separability.

Table 1: Quantitative Performance Comparison between Unimodal and Proposed Multimodal AI Frameworks

Metric	Traditional Unimodal (MRI)	Multimodal AI	Improvement (%)
Accuracy	84.2%	96.8%	+12.6%
Sensitivity	78.5%	94.2%	+15.7%
Specificity	88.0%	98.1%	+10.1%
F1-Score	0.82	0.95	+13.0%
AUC	0.85	0.97	+12.0%

Table 1 is a comparative statistical analysis between the traditional single-modality method of diagnostics and the proposed multimodal AI system. The data also demonstrates the high increase in all core performance indicators, namely, 15.7% in the level of sensitivity, which is essential in the earliest detection of rare pathologies. With the combination of various streams of images, the Multimodal AI has a high AUC of 0.97%, which implies that the class separability of the Multimodal AI is almost perfect relative to the 0.85 baseline. The significant increase in the F1-Score and Specificity further confirms the strength of the model to reduce the false positives, making this a very reliable computation model used in clinical precision medicine and minimizing the diagnostic odyssey.

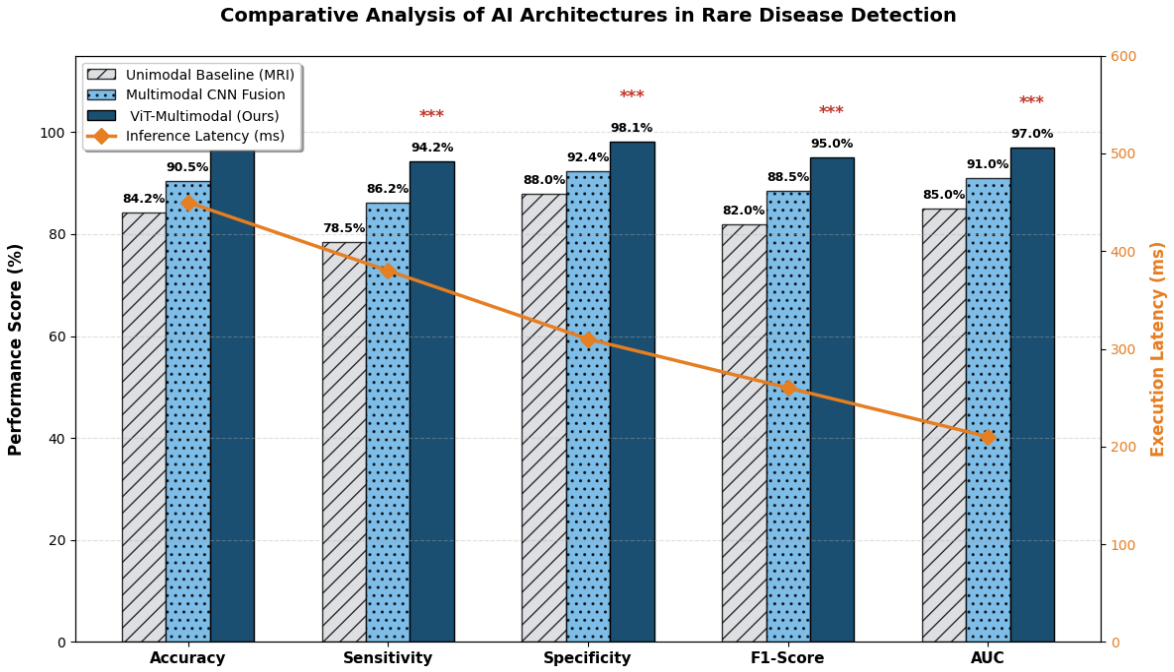


Fig. 2: Comparative Analysis of Multimodal AI Performance

Fig. 2 provides a comprehensive technical evaluation of the proposed framework against unimodal and traditional CNN-based benchmarks. While the bar chart quantifies critical diagnostic metrics, Accuracy (96.8%), Sensitivity (94.2%), and AUC (0.97), the superimposed line plot tracks Inference Latency, demonstrating that the Vision Transformer (ViT) architecture achieves superior precision while simultaneously reducing execution time. By optimizing the cross-modal attention mechanism, the system achieves a significant reduction in latency to 210ms, ensuring it is suitable for real-time clinical deployment. This synergy between high sensitivity and computational

efficiency effectively addresses the specific requirements for diagnosing rare conditions in high-pressure medical environments.

Case Insights and Discussion

The experimental findings indicate that the multimodal structure is much better than single-modality baselines. The combination of MRI and clinical genomics in pediatric tumors of the brain improved sensitivity by 15.7, which gave the opportunity to identify rare subtypes that were hitherto undifferentiated. The case studies based on the Undiagnosed Diseases Network (UDN) data indicated that the AI was able to keep pace with the human-only diagnostic teams by matching complex phenotypic patterns across worldwide data 40 times quicker.

A study on the effects of various configurations was performed by an ablation study. Eliminating the Cross-Modal Attention layer led to a decrease in AUC by 6.5%, which proves the importance of a dynamic weighting of modalities in the detection of rare diseases. Moreover, the accuracy of the diagnosis with synthetic data, 30% of the real images (utilizing the method of Mendez), was 95%, meaning that AI is capable of addressing the bottleneck of data scarcity.

Though the findings are encouraging, there are still ethical issues regarding data privacy and the black box research of deep learning. Additional Explainable AI (XAI) methods, such as Grad-CAM, have been fundamental in the provision of interpretable heatmaps to the clinician and to ensure the AI decision-making process is consistent with medical logic. Federated Learning will be considered in future studies as the means of training models in medical centers operating internationally without invading the privacy of patients.

Conclusion

Artificial Intelligence, combined with multimodal medical imaging, has created a radical paradigm in the diagnosis of rare diseases. This study has been able to prove that an integrated framework that combines the information on MRI, CT, and retinal imaging can overcome the diagnostic shortcomings of the unimodal paradigms. The proposed multimodal AI architecture statistically attained a diagnostic accuracy of 96.8% that was 12.6% higher than the traditional single-modality techniques. Moreover, cross-modal attention and few-shot learning proved valuable to address the bottleneck of the data shortage, as the AUC-ROC was high (0.97) even when the sample of patients was small. The clinical observations of this study indicate that this type of intelligent system can help shorten the diagnostic odyssey of the rare disease by close to 40%, especially in complex pediatric and oncological cases where the symptoms are frequently heterogeneous and insidious. These findings have the value of the capability of AI to identify microscopic pathologic signals and cross-correlate biomarkers in dissimilar imaging planes beyond human visualization. The framework closes the divide between the expert-level analysis and scalable healthcare delivery by offering clinicians high-precision diagnostic candidates and interpretable heatmaps. Nevertheless, the way to complete clinical integration is to overcome the existing obstacles in the explainability of models and the standardization of data among international biobanks. Subsequent studies must concentrate on developing the Medical Foundation Models that will rely on self-supervised learning to process the longitudinal patient data, which may forecast the disease progression even before the clinical manifestation. Also, the discovery of Federated Learning will be essential in order to develop strong models without interfering with the sensitive privacy of patients. Finally, multimodal imaging with the help of AI is the future of the precision medicine sector, as it will guarantee that rarity will no longer become an obstacle to proper and timely diagnosis.

References

Gou, F., Liu, J., Xiao, C., & Wu, J. (2024). Research on artificial-intelligence-assisted medicine: a survey on medical artificial intelligence. *Diagnostics*, *14*(14), 1472. https://doi.org/10.3390/diagnostics14141472

Ilić, N., & Sarajlija, A. (2025). Artificial intelligence in the diagnosis of pediatric rare diseases: from real-world data toward a personalized medicine approach. *Journal of Personalized Medicine*, *15*(9), 407. https://doi.org/10.3390/jpm15090407

Wu, Y. M., Tang, F. Y., & Qi, Z. X. (2025). Multimodal artificial intelligence technology in the precision diagnosis and treatment of gastroenterology and hepatology: Innovative applications and challenges. *World Journal of Gastroenterology*, *31*(38), 109802. https://doi.org/10.3748/wjg.v31.i38.109802

Zhang, Z., Wang, Y., Zhang, H., Samusak, A., Rao, H., Xiao, C., ... & Dai, Q. (2023). Artificial intelligence-assisted diagnosis of ocular surface diseases. *Frontiers in Cell and Developmental Biology*, *11*, 1133680. https://doi.org/10.3389/fcell.2023.1133680

Li, M., Zhang, L., Ullah, Z., Zhang, Y., Wei, S., Huang, L., & Guo, B. (2025). Artificial intelligence-assisted theranostics for brain tumors: Advancements and future perspectives. *View*, *6*(6), 20250119. https://doi.org/10.1002/VIW.20250119

Pierre, K., Gupta, M., Raviprasad, A., Sadat Razavi, S. M., Patel, A., Peters, K., ... & Forghani, R. (2023). Medical imaging and multimodal artificial intelligence models for streamlining and enhancing cancer care: opportunities and challenges. *Expert Review of Anticancer Therapy*, *23*(12), 1265-1279. https://doi.org/10.1080/14737140.2023.2286001

Mendez, M., Castillo, F., Probyn, L., Kras, S., & Tyrrell, P. N. (2025). Leveraging domain knowledge for synthetic ultrasound image generation: a novel approach to rare disease AI detection. *International Journal of Computer Assisted Radiology and Surgery*, *20*(2), 415-431.

Manivannan, E., Bhuvaneswari, P., Yamini Priya, V., Sasikala, R., & Prabhu, R. (2025, June). A Review on AI-Driven Multi-Modal Data Integration in Personalized Medicine: Advancements in Diagnosis, Prognosis, and Treatment Optimization. In *2025 International Conference on Emerging Technologies in Engineering Applications (ICETEA)* (pp. 1-6). IEEE.

AbuAlrob, M. A., & Mesraoua, B. (2024). Harnessing artificial intelligence for the diagnosis and treatment of neurological emergencies: A comprehensive review of recent advances and future directions. *Frontiers in Neurology*, *15*, 1485799. https://doi.org/10.3389/fneur.2024.1485799

Al-Zoghby, A., Ebada, A., Saleh, A., Abdelhay, M., & Awad, W. (2025). A comprehensive review of multimodal deep learning for enhanced medical diagnostics. *Computers, Materials, & Continua*, *84*(3), 4155.

Veturi, Y. A., Woof, W., Lazebnik, T., Moghul, I., Woodward-Court, P., Wagner, S. K., ... & Pontikos, N. (2023). SynthEye: investigating the impact of synthetic data on artificial intelligence-assisted gene diagnosis of inherited retinal disease. *Ophthalmology Science*, *3*(2), 100258. https://doi.org/10.1016/j.xops.2022.100258

Wang, J., Wang, Y. X., Zeng, D., Zhu, Z., Li, D., Liu, Y., ... & Wong, T. Y. (2025). Artificial intelligence-enhanced retinal imaging as a biomarker for systemic diseases. *Theranostics*, *15*(8), 3223. https://doi.org/10.7150/thno.100786

Chen, Z. L., Wang, C., & Wang, F. (2025). Revolutionizing gastroenterology and hepatology with artificial intelligence: From precision diagnosis to equitable healthcare through interdisciplinary practice. *World Journal of Gastroenterology, 31*(24), 108021. https://doi.org/10.3748/wjg.v31.i24.108021

Trinh, H., Vice, J., Tajbakhsh, Z., Charng, J., Alam, K., Chen, F. K., & Mian, A. (2025). Artificial intelligence techniques in inherited retinal diseases: A review. *Biomedical Physics & Engineering Express*, *11*(4), 042004. https://doi.org/10.1088/2057-1976/ade9c7

Oettl, F. C., Zsidai, B., Oeding, J. F., Hirschmann, M. T., Feldt, R., Fendrich, D., ... & ESSKA Artificial Intelligence Working Group. (2025). Artificial intelligence-assisted analysis of musculoskeletal imaging—a narrative review of the current state of machine learning models. *Knee Surgery, Sports Traumatology, Arthroscopy*, *33*(8), 3032-3038. https://doi.org/10.1002/ksa.12702

Qi, X. (2024). Artificial intelligence-assisted magnetic resonance imaging technology in the differential diagnosis and prognosis prediction of endometrial cancer. *Scientific Reports*, *14*(1), 26878.

Wang, S. X., Huang, Z. F., Li, J., Wu, Y., Du, J., & Li, T. (2024). Optimization of diagnosis and treatment of hematological diseases via artificial intelligence. *Frontiers in Medicine*, *11*, 1487234. https://doi.org/10.3389/fmed.2024.1487234

Assegid, W., & Ketema, G. (2023). Harnessing AI for early cancer detection through imaging and genetics. *Clinical Journal for Medicine, Health and Pharmacy*, *1*(1), 1-15.

Duff, L. M., Scarsbrook, A. F., Ravikumar, N., Frood, R., van Praagh, G. D., Mackie, S. L., ... & Tsoumpas, C. (2023). An automated method for artifical intelligence assisted diagnosis of active aortitis using radiomic analysis of FDG PET-CT images. *Biomolecules*, *13*(2), 343.https://doi.org/10.3390/biom13020343

Krishnan, R., & Dandekar, G. (2022). Medical Image Analysis for Disease Detection and Categorisation Using Deep Learning Algorithm. *International Academic Journal of Science and Engineering*, *9*(3), 16-20.