Gen-AI For Medical Diagnosis and Drug Discovery Using UNETR

Aryan A. Dalvi[1], Pruthvi R. Patel[2], Prasad V. Reddy[3], Prof. Swati R. Bhatt[4]

[1,2,3]*Student, Manjara Charitable Trust's Rajiv Gandhi Institute of Technology*

[4]*Professor, Manjara Charitable Trust's Rajiv Gandhi Institute of Technology*

***Abstract*—Generative Artificial Intelligence (GenAI) is poised to revolutionize medicine, yet its adoption is challenged by significant barriers in clinical diagnosis and drug discovery. Deep learning has shown promise, but two major issues limit its clinical application: the "black-box" nature of AI predictions and concerns regarding their reliability and accuracy. This review paper surveys an integrated framework that connects high-fidelity medical image segmentation for diagnosis with de novo molecule generation for therapeutic design. We analyze advanced architectures like the UNETR (UNEt TRansformers) for volumetric segmentation and explore the role of specialized platforms like MONAI in their implementation. To address model opacity, the importance of Explainable AI (XAI) techniques in building trust and validating computational drug discovery is examined. This integrated approach surveys a path forward that addresses not only performance but also the critical needs for efficiency, transparency, and interpretability across the diagnosis-to-treatment pipeline.**

***Index Terms*—Generative AI, Medical Diagnosis, Drug Discovery, Deep Learning, UNETR, Medical Image Segmentation, MONAI, Explainable AI (XAI), De Novo Molecule Generation.**

I. INTRODUCTION

Generative Artificial Intelligence (GAI) is emerging as a transformative technology in healthcare, poised to revolutionize critical areas like medical diagnosis and drug discovery [9]. Traditionally, these fields are notoriously slow, costly, and resource-intensive [3]. Drug development, for instance, faces high failure rates [2, 3], while early and accurate medical diagnosis often relies on the subjective, manual interpretation of complex data like histopathological images [7] or 3D scans [5]. Consequently, GAI models, which can generate novel text, images, and data, are being explored to enhance efficiency, accuracy, and patient outcomes [6, 8].

In medical diagnostics, GAI is increasingly applied to medical imaging tasks [6]. Architectures like the UNETR (UNet Transformer) leverage transformers to perform complex 3D segmentation of MRIs and CTs [5]. Hybrid deep learning models, such as those using DenseNet201, have achieved remarkable accuracy (99.68%) in the early detection of lung cancer from histological images [7]. In parallel, GAI is accelerating de novo drug discovery. Models are being developed to generate novel molecular structures, often represented as SMILES strings, directly from biological data such as gene expression profiles [2]. This progress is supported by open-source, collaborative frameworks like MONAI, which provides a standardized PyTorch-based platform for medical AI, including implementations of networks like UNETR [4].

Despite these advances, significant gaps remain. The "black box" nature of these complex models is a primary barrier to clinical trust and adoption [3, 9]. This lack of transparency has fueled the need for Explainable AI (XAI), a field dedicated to making model predictions interpretable [3].

Furthermore, model robustness is a critical, often-overlooked challenge. Foundation models that perform well on in- distribution (ID) data often experience significant performance degradation when applied to out-of- distribution (OOD) data from different clinical settings, making uncertainty quantification essential [10].

Other major hurdles include the risk of GAI generating inaccurate information, ensuring patient data privacy, and the potential for algorithmic bias [8, 9]..

This review synthesizes the current landscape of Generative AI in both medical diagnosis and drug discovery. We examine key applications, from AI-enabled histological analysis [7] and 3D segmentation [5] to de novo molecular generation [2]. This work provides a critical analysis of the primary limitations facing the field, focusing on the essential role of XAI [3], the need for robust uncertainty quantification for foundation models [10], and the practical and ethical challenges of integrating GAI into clinical services [1, 8]. The goal is to provide a comprehensive overview of GAI's transformative potential while grounding it in the challenges that must be addressed for its responsible implementation in healthcare.

II. HISTORY & BACKGROUND

The traditional pathways for medical diagnosis and drug discovery have long been characterized as laborious, costly, and high-risk [2, 3]. In drug development, identifying promising therapeutic candidates is a resource-intensive process with high failure rates [2, 3]. In diagnostics, clinicians often rely on the manual interpretation of complex data, such as histological images [7] or 3D volumetric scans [5], which can be time-consuming and subject to variability. This landscape created a clear need for advanced computational methods to improve efficiency and accuracy [3, 9].

The advent of Artificial Intelligence (AI) and Machine Learning (ML) promised to meet this need by analyzing enormous datasets to discover patterns and generate predictions [3]. In medical diagnosis, AI models were increasingly applied to medical imaging for tasks like disease detection, classification, and screening [1, 6, 7]. A key development in this area was the establishment of open-source, domain-specific frameworks. MONAI, a consortium-led PyTorch-based framework, was introduced to provide standardized, "domain-optimized foundational capabilities" and unify the fragmented field of healthcare AI, especially for imaging [4].

As AI in medicine matured, model architectures evolved. A significant advancement was the adaptation of transformers, which had been highly successful in natural language processing, for vision tasks [5]. This led to novel architectures like the UNETR (UNet Transformer), which reformulates 3D segmentation as a sequence-to-sequence prediction problem. By using a transformer as the encoder, UNETR can "effectively capture the global multi-scale information" from volumetric data, moving beyond the limitations of older convolutional networks [5]. This new class of "foundation models" began to show strong performance on complex segmentation tasks [10].

Alongside these analytical models, Generative AI (GAI) emerged as a subfield focused on creating new content rather than just classifying it [6, 8, 9]. This opened two major frontiers. In drug discovery, GAI enabled de novo molecular generation, with models (e.g., VAEs and LSTMs) capable of producing novel, syntactically valid SMILES strings based on inputs like gene expression profiles [2].

In medical diagnosis, GAI is used to synthesize realistic medical images (e.g., X-rays, MRIs) to augment limited datasets [6, 9] or for image-to-image translation [6].

High Performance vs. Interpretability: Initial works often prioritized raw predictive power but offered little insight into model reasoning. In contrast, XAI approaches offer interpretability but must be integrated without compromising performance.
Fragmentation vs. Integration: Many AI tools address isolated tasks (diagnosis or drug discovery) . Integrating these into end-to-end pipelines offers synergy but remains a significant architectural challenge .
Data Dependency & Bias: High-performing models require large, diverse datasets, but medical data can be limited or biased, affecting generalization and fairness.

Research Gaps:

Despite progress, major gaps remain:
Limited Clinical Integration: Many GenAI applications are preliminary, mostly assisting rather than automating clinical workflows.
Robust Explainability: Developing XAI methods that provide clinically meaningful and reliable explanations for complex models like Transformers is crucial.
Bias and Reliability: Systematically evaluating and

mitigating bias, ensuring data privacy, and validating AI outputs for clinical safety are underexplored.
Integration Challenge: Building truly unified, robust, and validated "diagnosis-to-treatment" pipelines remains a central, unaddressed problem.

III. METHODOLOGY

Methodologies for Generative AI in healthcare are bifurcated into two primary domains: medical diagnosis and drug discovery [9]. Diagnostic approaches typically involve complex deep learning models for image segmentation and classification [5, 7], while drug discovery focuses on generating novel molecular structures from biological data [2, 3]. A common challenge across both domains is the "black box" nature of the models [3, 9] and ensuring robust performance on diverse clinical data [10]. Thus, a comprehensive methodology also integrates (i) model architecture, (ii) data processing pipelines, (iii) explainability (XAI) techniques [3], and (iv) uncertainty quantification [10].

Diagnostic Model Architectures and Data:
For medical diagnosis, the methodology centers on analyzing medical images such as CT, MRI, and histopathology slides [5, 6, 7]. Datasets like the Medical Segmentation Decathlon (MSD) or the Multi Atlas Labeling Beyond The Cranial Vault (BTCV) are used for training 3D segmentation models [5]. For 2D histological analysis, datasets like LC25000 are used [7]. To standardize these complex workflows, open-source frameworks like MONAI are often employed, providing a PyTorch-based ecosystem with domain-specific transforms, network architectures, and data loaders [4].
A key architectural trend is the use of transformers [5]. The UNETR (UNet Transformer) model, for example, reformulates 3D segmentation as a sequence-to-sequence problem. It uses a transformer as an encoder to learn sequence representations of the input volume, capturing both global and local features. These transformer layers are then connected to a CNN-based decoder via skip connections to compute the final segmentation output [5]. For 2D image classification, hybrid approaches combining CNNs like DenseNet201 with feature extraction methods like color histograms are also utilized [7].
De Novo Drug Discovery Models:
In drug discovery, the methodology shifts from image analysis to molecular generation [3]. The goal is to create novel, syntactically valid molecules, often represented as SMILES strings [2]. These models are trained on datasets that pair biological data with molecular structures, such as gene expression profiles from the LINCS database [2].
A common approach is a hybrid neural network. For instance, a Variational Autoencoder (VAE) can be used as a feature extractor to learn a low-dimensional latent representation of the high-dimensional gene expression data. This latent vector then serves as a condition for a chemical generator, such as a Long Short-Term Memory (LSTM) network, which iteratively produces the SMILES string for a new molecule that should exhibit the desired biological activity [2].

Validation, Explainability, and Robustness: Validating these generative models is a critical methodological step. For diagnosis, standard metrics like the Dice score and Hausdorff Distance (HD95) are used to evaluate segmentation accuracy on in-distribution (ID) test sets [5, 10]. However, a crucial part of the methodology is assessing robustness by testing on out-of-distribution (OOD) datasets, which represent concept or distribution shifts (e.g., images from different scanners or with different diseases) [10]. On OOD sets, metrics such as entropy (to quantify model uncertainty) and volume occupancy (to measure mis-detected tumor volume) are vital [10]. To address the "black box" problem [3, 9], Explainable AI (XAI) techniques are integrated into the methodology. These methods provide transparency by showing how a model reached its conclusion [3]. This includes visualization techniques like Grad-CAM and occlusion sensitivity [4], or feature-importance methods like LIME and SHAP, which help build clinical trust by justifying the model's predictions [3].
This systematic review of preliminary evidence suggests that GAI is still in its initial stages, primarily assisting clinicians with knowledge access rather than fully guiding or automating clinical services [1, 8].

IV. RESULT AND ANALYSIS

Generative AI (GAI) models have demonstrated high technical accuracy in specialized healthcare tasks, though these results are tempered by significant challenges in robustness, explainability, and practical clinical integration [1, 3, 10].

In medical diagnosis, GAI and related deep learning models show expert-level performance. For instance, a hybrid model combining DenseNet201 achieved 99.68% accuracy in classifying lung cancer from histological images [7]. For complex 3D medical imaging, transformer-based architectures like UNETR (UNet Transformer) have set new state-of-the-art performance benchmarks on multi-organ segmentation tasks by effectively capturing global, multi-scale information [5]. This work is supported by open-source frameworks like MONAI, which provide the specialized, domain-specific tools required for medical AI development [4].In drug discovery, GAI is successfully accelerating the de novo design of novel molecules [9]. Models such as HVL2Mol can now generate syntactically valid SMILES strings for new molecules directly from gene expression profiles. These generated molecules show high validity (88.6%), uniqueness (83.0%), and strong drug-like properties, confirming GAI's ability to explore the chemical space effectively [2].

Despite these promising results, critical challenges remain. A primary issue is robustness; foundation models that perform well on in-distribution (ID) data often fail when applied to out-of-distribution (OOD) datasets, highlighting the need for better uncertainty quantification metrics beyond standard accuracy [10]. Furthermore, the "black box" nature of these complex models hinders clinical trust [3, 9], making Explainable AI (XAI) essential for providing transparency [3, 4]. Finally, a systematic review of GAI's role in clinical services found its use is still in an "initial stage" [1, 8]. The vast majority of applications (87.6%) are used to assist clinicians with knowledge access, while very few (4.3%) are used to automate clinical functions, indicating a significant gap between GAI's technical potential and its real-world implementation [1],

V. CONCLUSION

This review synthesizes the current landscape of Generative AI, detailing its dual application in medical diagnosis and *de novo* drug discovery. The findings confirm that GAI models are achieving significant technical milestones. In diagnostics, advanced architectures like the transformer-based UNETR [5] and hybrid DenseNet201 models [7] demonstrate high accuracy in complex tasks like 3D segmentation and histological cancer detection. In drug discovery, GAI is successfully generating novel, valid molecular structures (SMILES) from biological data like gene expression profiles [2], accelerating a traditionally slow research pipeline.However, this review identifies three critical challenges that limit the clinical translation of these powerful models. First is the "black box" problem; the opaque nature of deep learning models necessitates the integration of Explainable AI (XAI) to ensure transparency and build clinical trust [3, 9]. Second, model robustness remains a primary concern. Foundation models that excel on in-distribution (ID) data often show significant performance degradation on out-of-distribution (OOD) data, making rigorous uncertainty quantification essential before deployment [10]. Finally, a systematic review of GAI in practice reveals its role is still in an "initial stage" [8], predominantly *assisting* clinicians with knowledge access rather than fully *automating* clinical functions [1].The integration of GAI into healthcare, therefore, depends on addressing these gaps. Future work must focus on creating models that are not only accurate but also robust, interpretable, and seamlessly integrated into clinical workflows [4]. By unifying the assessment of model performance with explainability [3] and robustness [10], this work contributes to bridging the disconnect between GAI's technical potential and its responsible, effective implementation in medicine.

REFERENCES

[1] D. Yim, J. Khuntia, V. Parameswaran, and A. Meyers, "Preliminary Evidence of the Use of Generative AI in Health Care Clinical Services: Systematic Narrative Review," JMIR Med Inform, vol. 12, p. e52073, 2024. [Online]. Available: https://doi.org/10.2196/52073

[2] C. Li and Y. Yamanishi, "De Novo Generation of Hit-like Molecules from Gene Expression Profiles via Deep Learning," arXiv preprint arXiv:2412.19422, 2024. [Online]. Available: https://arxiv.org/abs/2412.19422

[3] R. Alizadehsani et al., "Explainable Artificial Intelligence for Drug Discovery and Development: A Comprehensive Survey," IEEE Access, vol. 12, pp. 35796–35813, 2024. [Online].

Available:
https://doi.org/10.1109/ACCESS.2024.3373195
preprint arXiv:2403.13113, 2024. [Online]. Available: https://arxiv.org/abs/2403.13113

[4] M. J. Cardoso et al., "MONAI: An open- source framework for deep learning in healthcare," arXiv preprint arXiv:2211.02701, 2022. [Online]. Available: https://arxiv.org/abs/2211.02701

[5] A. Hatamizadeh et al., "UNETR: Transformers for 3D Medical Image Segmentation," arXiv preprint arXiv:2103.10504, 2021. [Online].Available: https://arxiv.org/abs/2103.10504

[6] K. Manubolu and R. Peeriga, "Generative AI in Disease Diagnosis," Preprints.org, 2024. [Online]. Available:https://doi.org/10.20944/preprints202410.1555.v 1

[7] N. F. Noaman, B. M. Kanber, A. Al Smadi, L. Jiao, and M. K. Alsmadi, "Advancing Oncology Diagnostics: AI-Enabled Early Detection of Lung Cancer Through Hybrid Histological Image Analysis," IEEE Access, vol. 12, pp. 64396–64415, 2024. [Online]. Available: https://doi.org/10.1109/ACCESS.2024.3397040

[8] R. Xu and Z. Wang, "Generative artificial intelligence in healthcare from the perspective of digital media: Applications, opportunities and challenges," Heliyon, vol. 10, p. e32364, 2024. [Online]. Available: https://doi.org/10.1016/j.heliyon.2024.e32364

[9] S. Said et al., "Generative AI for Transformative Healthcare: A Comprehensive Study of Emerging Models, Applications, Case Studies, and Limitations," IEEE Access, vol. 12, pp. 31078–31107, 2024. [Online]. Available: https://doi.org/10.1109/ACCESS.2024.3367715

[10] A. Rangnekar, N. Nadkarnia, J. Jiang, and H. Veeraraghavan, "Quantifying uncertainty in lung cancer segmentation with foundation models applied to mixed-domain datasets," arXiv