Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 9
AIDML 2025

Applications of Artificial Intelligence in Rare Disease Research: A Technical Framework and Future Prospects

Yunfan Chen [2], Hairui Hu [1, *], Ziqin Xue [3]

[1] Academy of Data Science, City University of Macau School, Macau, China

[2] Digital Media Technology, Xi'an Jiaotong-Liverpool University, Suzhou, China

[3] Computer Science, Shanghai Zhongqiao Vocational University, Shanghai, China

* Corresponding Author Email: D22090102418@cityu.edu.mo

Abstract. In rare disease studies, researchers have consistently encountered issues like data shortage and diagnostic uncertainty—challenges largely stemming from these diseases' low prevalence, varied symptoms, and complex diagnostic procedures. Traditional research methods frequently produce unreliable results, largely because of persistent issues such as limited data and clinical complexity. In recent years, the advancement of artificial intelligence, particularly foundation models, has opened up new possibilities to address these challenges. This paper reviews recent progress in the application of AI to rare disease research and, based on current literature and technological trends, proposes a comprehensive conceptual framework. The proposed framework consists of four key components: multimodal pretraining, dynamic prompting, efficient retrieval, and lightweight deployment. Although this framework remains at a theoretical stage, it aims to offer inspiration and guidance for the development of future diagnostic models. The paper seeks to provide an integrated perspective on AI research in the context of rare diseases, highlighting the potential of cross-modal integration and large model adaptation in enabling more efficient, accurate diagnostics and personalized treatment strategies.

Keywords: Artificial Intelligence; Rare Disease Research; Large Language Model; AI Usage in Rare Disease Diagnosis.

1. Introduction

Owing to their extremely low prevalence, rare diseases bring considerable challenges. Both medical researchers and clinicians continue to struggle with the implications. Because of limited patient populations, insufficient data, and widely varying clinical presentations, many rare diseases are inadequately represented in traditional research frameworks. The effectiveness of conventional clinical research methods is often limited, particularly when addressing rare diseases. This limitation is particularly evident in studies involving rare diseases. In addition, the development of therapeutics for rare diseases faces high costs and long development cycles, leaving many patients without timely and effective treatment.

With the rapid development of artificial intelligence (AI) technologies, their application in rare disease research has increasingly drawn attention. Particularly in areas such as diagnosis, image retrieval, and disease recognition, Artificial intelligence has demonstrated considerable promise in transforming rare disease research and clinical applications. AI enables advancements in diagnosis, treatment, and public health for rare diseases by leveraging multiple data modalities, enhancing computational efficiency, and strengthening its ability to generalize across tasks. For example, the MVP model integrates outputs from multiple large language models (LLMs) using a majority voting mechanism, significantly improving diagnostic accuracy for low-sample rare conditions [1]. The RetiZero model combines masked autoencoder (MAE) self-supervised learning with CLIP contrastive learning and incorporates low-rank adaptation (LoRA) and Dirichlet uncertainty calibration to achieve efficient identification of over 400 retinal diseases, including rare conditions [2]. The FISH model focuses on retrieving rare cases from large-scale pathology image datasets using self-supervised learning and Van Emde Boas (vEB) tree structures for efficient image indexing [3].

The TITAN model, a multimodal foundation model for whole-slide pathology images, enhances general feature representation by leveraging multi-stage pretraining and vision-language alignment techniques [4].

In recent years, the application of AI in rare disease research has expanded to encompass early diagnosis, drug development, and patient management. Although meaningful advancements have been made, several unresolved obstacles continue to hinder further development in the field. In addition, the field continues to struggle with problems like poor interpretability of models and challenges in merging data from different modalities.

Against this backdrop, this paper systematically reviews the current landscape of AI applications in rare disease research, drawing from recent literature and technological advancements. We analyze the strengths and limitations of representative models and propose a comprehensive and efficient conceptual framework. The framework incorporates techniques such as multimodal pretraining, dynamic prompting strategies, efficient retrieval, and lightweight model design to address challenges including data scarcity and computational constraints. Beyond advancing AI applications in rare disease contexts, this research provides valuable insights and strategic guidance for the wider domain of medical artificial intelligence.

2. Review and Strategic Structuring of AI Models for Rare Disease Research

2.1. Overview of Existing Models and Methods

Numerous diagnostic models powered by AI have emerged in the context of rare diseases. This paper emphasizes a selection of influential techniques, analyzed from both technical and goal-oriented perspectives, to illustrate their varied roles in rare disease research.; specifically, from a goal-oriented perspective, the models covered in this paper address rare disease identification and drug development, while from a technical perspective, they include accuracy in few-shot learning and the recognition of pathological images.

Table 1 summarizes key AI models covered in this paper, detailing their main research goals and the technical methods applied in each case.

Table 1. Summary of Representative AI Models: Objectives and Core Methods

Model Name	Research Objective	Core Methodology
MVP [5]	Improve diagnostic accuracy in few-shot rare disease identification	Ensemble of multiple LLM outputs with majority voting to enhance robustness
RetiZero [6]	Identify over 400 fundus diseases including rare cases	MAE self-supervised learning + CLIP contrastive learning; LoRA optimization with Dirichlet uncertainty calibration
RAREMed [7]	Address therapeutic questions in rare disease drug discovery	Transformer encoder with pretraining–finetuning; self-supervised tasks to optimize feature representations
FISH [8]	Efficient retrieval of rare cases from large-scale pathology datasets	Self-supervised learning (VQ-VAE) + van Emde Boas (vEB) tree for O(1) retrieval efficiency [9]
TITAN [10]	General feature representation and multimodal support for pathology images	Multi-stage pretraining (visual self-supervised → region-level → slide-level alignment); ALiBi attention for long-sequence modeling

2.2. Comparative Analysis of Technical Strengths and Limitations

Because diagnosing rare diseases involves considerable complexity, no single model can fully address the range of challenges. Accordingly, this analysis examines selected models across distinct technical aspects to compare their strengths and limitations. The comprehensive technical advantages and limitations are summarized in Table 2.

Table 2. Technical Strengths and Limitations of Representative Models

Technical Dimension	Model Name	Strengths	Limitations
Ensemble Learning	MVP	Enhances robustness and reduces single-model bias	High computational cost and maintenance burden for multiple models
Multimodal Alignment	RetiZero	Strong semantic understanding across modalities	Depends heavily on high-quality aligned datasets
Dynamic Prompting	RareBench	Significantly improves few-shot learning performance	High complexity in constructing knowledge graphs
Lightweight Design	TITAN (LoRA)	Efficient and suitable for deployment in low-resource settings	Limited capacity for modeling global context

2.3. Proposal of an Integrated and Efficient Strategic Framework

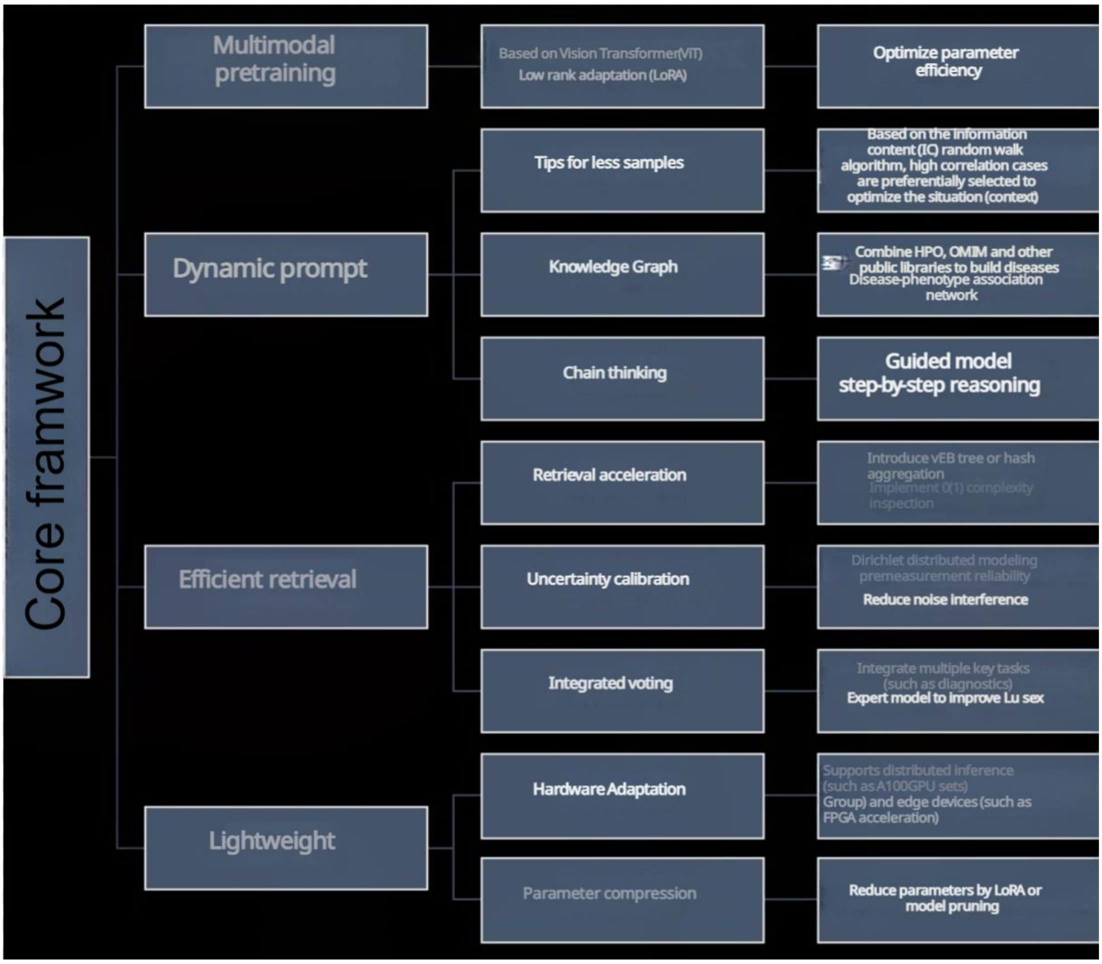

Figure 1. Most Efficient Methodology (Picture credit: Original)

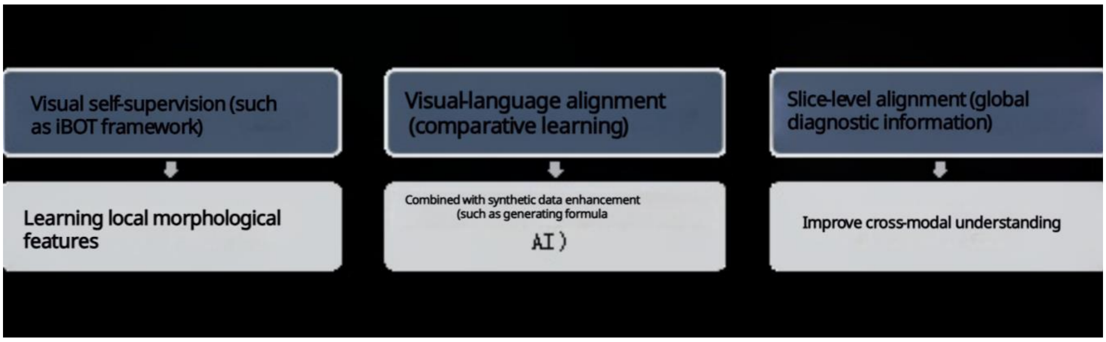

Figure 2. High Generalizability Framework (Picture credit: Original)

Figure 1 outlines a strategy focused on multimodal pretraining, augmented with dynamic prompting to enhance few-shot learning. This approach also integrates fast retrieval structures and compact model architectures for use in low-resource settings. To ensure system reliability, ensemble learning and uncertainty calibration are also applied.

As shown in Figure 2, the proposed framework accommodates various rare disease research tasks, such as diagnosis, treatment planning, and image-based case retrieval, with a focus on adaptability and generalization.

3. Technical Analysis of Comprehensive Research Methods

3.1. Multimodal Pre-training

Rare disease cases typically require multimodal interpretation, as they involve diverse data sources like pathology images, clinical notes, and genomic sequences [11]. odels limited to a single modality often fail to reflect the intricate interdependencies across these datasets; therefore, to address the shortcomings of unimodal systems, RetiZero aligns visual and textual features to accurately detect both common and rare fundus diseases. However, its strong dependence on pretrained models like MAE and CLIP, and limited incorporation of non-textual data—such as patient history—may constrain its generalizability in practical clinical environments.

Although TITAN has enabled cross-modal understanding between pathology images and clinical reports through a three-stage pretraining pipeline, its reliance on pathology narratives remains overly strong, potentially constraining its generalizability to settings without complete pathology annotations.

To address the limitations of using a single model structure, this section proposes a method of integrated modeling and explains it in detail: by adopting multimodal pretraining as the core of the technical framework, the model can enhance generalization capabilities across different data types, where visual self-supervised learning frameworks (such as iBOT) are used to extract local morphological features, semantic alignment via contrastive learning is employed to strengthen textual associations, and data fusion across modalities helps compensate for the shortage of high-quality annotations required by generative AI, especially in low-resource domains like rare diseases [12].

Multimodal pretraining can support complex real-world tasks such as pathology image retrieval, which demands integration of image features with diagnostic texts, and drug recommendation, which requires understanding of related disease codes as well as interaction history with prescribed medications or existing therapeutic agents, and pretraining also reduces dependence on labeled data, as demonstrated in TITAN's iBOT stage, where large-scale unlabeled pathology slides are used for pretraining without annotation, making it particularly applicable in real-world rare disease scenarios where labeled datasets are extremely sparse.

3.2. Dynamic Prompting Strategies

One of the key obstacles hindering the application of AI in rare disease research is the persistent lack of sufficient data. For certain rare conditions, fewer than fifty cases have been officially recorded across the globe. Additionally, labeling medical data is both time-intensive and expensive, typically necessitating the expertise of domain-specific clinicians. Compared to data from more prevalent conditions, rare disease datasets are notably skewed, which intensifies challenges during model training.

Existing techniques like MVP enhance performance via ensemble learning of multiple LLMs, but they rely on static few-shot examples, which lack adaptability. RareBench introduces a dynamic few-shot prompting mechanism, yet it does not incorporate real-time knowledge updates.

To overcome these challenges, our integrated and efficient framework adopts a cooperative strategy to mitigate data scarcity by combining the following components:

3.2.1. Knowledge-Driven Dynamic Prompting.

The system utilizes biomedical knowledge graphs like HPO and OMIM to build networks linking diseases with their phenotypic features. In RareBench, prompt selection is driven by vector embeddings from the graph, which notably improves the performance of large language models in diagnostic tasks; furthermore, to optimize the quality of selected information, an information content (IC)-based random walk algorithm is applied to prioritize phenotype nodes with higher specificity (e.g., “retinitis pigmentosa” is more informative than “vision loss”), thereby reducing noise and improving the accuracy of the diagnostic context.

3.2.2. Prompt Engineering with Chain-of-Thought Reasoning.

Prompt engineering is designed to guide the model’s reasoning process using chain-of-thought (CoT) strategies, where instructions such as “extract phenotype → match disease → eliminate confounding conditions” simulate the diagnostic logic typically employed by physicians, and during prompt construction, these structured reasoning steps enhance the interpretability of model outputs; for example, in rare disease screening, dynamic prompts can automatically match patient phenotypes (e.g., “jaundice + splenomegaly”) with relevant diseases in ORDO (e.g., “Gilbert syndrome”), and through CoT-based generation, produce a list of differential diagnoses to reduce the risk of missed cases.

3.3. Self-Supervised Learning and Synthetic Data Generation

While labeled data for rare diseases is extremely limited, a considerable amount of unlabeled clinical texts and medical images still exists and holds substantial latent value. For instance, the TITAN model uses the iBOT framework to conduct self-supervised pretraining on 340,000 unlabeled pathology slides, allowing the model to learn general features such as cellular morphology and tissue structure. These features can then be fine-tuned with only a small amount of labeled data for downstream tasks, such as classifying 43 rare cancer types, achieving accuracy as high as 85%. In contrast, supervised learning would typically require at least 10 times more labeled samples to reach similar performance levels.

For extremely rare diseases (e.g., Bietti crystalline dystrophy, which has fewer than five documented cases), generative models such as Stable Diffusion can be employed to synthesize realistic pathology images or phenotype descriptions. TITAN, for example, uses PathChat to generate over 420,000 region-of-interest (ROI) and textual description pairs, enabling training on synthetic data that closely approximates real-world cases—with F1-score differences of less than 3% between synthetic and real datasets. This approach effectively mitigates the data scarcity challenge in rare disease AI applications.

3.4. Efficient Retrieval and Lightweight Design

Medical datasets—such as TCGA, which includes over 22,000 pathology images—are extremely large in scale. Traditional retrieval methods like linear search are inefficient and unsuitable for real-time clinical applications. Moreover, clinical environments often face hardware limitations; for example, many local hospitals lack high-end GPUs required by large AI models.

To address these issues, efficient data structures are essential. The FISH model implements a van Emde Boas (vEB) tree, achieving O (1) time complexity in image retrieval. Even when searching in million-scale datasets, the system can respond in real-time. For instance, a physician uploading a prostate cancer slide can receive the top 5 most similar reference cases within milliseconds.

In terms of deployment, TITAN utilizes Low-Rank Adaptation (LoRA) for parameter-efficient fine-tuning, reducing the model size to 42M parameters, a 90% reduction compared to traditional ViT models like PRISM with 1B parameters. Despite the compression, performance is largely maintained. This design allows the model to be deployed on lower-end clinical hardware, such as portable pathology scanners. Moreover, distributed inference is supported—allowing massive whole-slide images (WSIs) to be processed in parallel on A100 GPU clusters—overcoming memory bottlenecks.

Although FISH's use of integer-based indexing limits features expressiveness, this is compensated by incorporating uncertainty-aware ranking algorithms, such as Dirichlet calibration, to refine retrieval accuracy.

While RetiZero applies LoRA only to the image encoder, the integrated framework proposed here extends LoRA to multimodal modules, including the text encoder, thereby further improving parameter efficiency and deployment flexibility.

3.5. Ensemble Learning and Uncertainty Calibration

Any medical AI model must produce highly reliable predictions, and while MVP reduces single-model errors through majority voting, it does not quantify predictive confidence; on the other hand, RetiZero employs Dirichlet calibration to handle uncertainty, but does not combine it with ensemble strategies.

However, in the integrated method proposed in this study, both approaches are applied in parallel, where ensemble voting mechanisms (such as those used in MVP) integrate multiple expert modules—including large language models (LLMs), visual models, and knowledge graph engines—so that diagnostic bias can be minimized through weighted consensus.

For example, the LLM generates a list of candidate diseases, the visual model verifies features in the medical image, and the knowledge graph provides additional phenotype associations, with the final output being a consensus-based result that integrates multiple viewpoints.

Dirichlet uncertainty calibration is also used to model the confidence level of predictions, meaning that when the uncertainty is high, the system can flag the result for manual review, thus avoiding misleading outputs on ambiguous cases. For instance, in retinal image analysis, if the model's confidence in diagnosing "giant cell arteritis" falls below the threshold, the system will automatically mark the case as "requires specialist consultation." The ensemble strategy can reduce the misdiagnosis rate by 5–10% (according to clinical validation results from RetiZero), especially in scenarios where it assists junior doctors, and it significantly improves the robustness of clinical decision-making.

3.6. Sustainability and Scalability

Any medical AI system must be adaptable to data growth, technological evolution, and emerging diseases (such as COVID-19 variants), and many of the current models mentioned—such as RAREMed—still rely on static pretrained datasets, making it difficult for them to perform dynamic updates [13].

Through continual learning, incremental training can be used to update both the knowledge graph and model parameters, so that when a new rare disease (e.g., "Fabry disease") is identified, the system can automatically retrieve the latest literature and update its ORDO mapping, thus achieving a degree of self-refreshing without the need to retrain the entire model.

Meanwhile, generative data augmentation techniques using diffusion models can be applied to synthesize rare disease images (e.g., colon cancer data from Kather100k) to alleviate the problem of sample scarcity, and TITAN has already verified the effectiveness of synthetic data by generating over 420,000 ROI-description pairs via PathChat.

In the future, genomic data (such as molecular labels from TCGA) can be integrated with imaging features to build full-course diagnostic systems; for example, breast cancer subtyping can be achieved by combining HER2 expression levels with pathological image features, thereby enabling multimodal expansion.

The integrated framework is also designed to support continuous learning and cross-domain generalization, enabling it to dynamically adapt to data growth and distribution changes. At the same time, given the continuous discovery of new genes and new diseases associated with rare diseases, its ability to update the knowledge base over time is particularly important to compensate for the timeliness of the model.

For example, the framework supports incremental updates to the knowledge graph and model parameters such that, when new cases of “Fabry disease” are reported, the system can automatically retrieve the latest publications from PubMed and update the ORDO mapping without retraining the full model.

In practical deployment, the model also demonstrates cross-device and cross-domain generalizability; for instance, differences in retinal camera brands across hospitals often cause data distribution shifts, but RetiZero maintains 82% accuracy in cross-domain testing (e.g., from Topcon to Zeiss cameras) through Dirichlet uncertainty calibration and contrastive learning, thus reducing reliance on any single data source.

3.7. Applications of AI in Early Diagnosis and Drug Development

In the area of early diagnosis, AI can be used in newborn screening to analyze metabolomic data—such as results from mass spectrometry—to rapidly identify genetic disorders like phenylketonuria; in addition, facial recognition technologies such as DeepGestalt can diagnose conditions like Angelman syndrome based on facial features, with an accuracy rate exceeding 90%. In terms of genomic mutation analysis, AI can assist in interpreting whole genome sequencing results and identifying rare disease–related mutations, such as those associated with spinal muscular atrophy. In the area of drug discovery acceleration, graph neural networks (GNNs) can be used for drug repurposing by mining associations between existing drugs and rare diseases, such as identifying imatinib as a potential treatment for systemic mastocytosis; furthermore, AI-powered virtual screening platforms can simulate drug–target interactions, reducing experimental costs by up to 70%—as exemplified by the Atomwise platform. Lastly, in patient stratification, AI can match rare disease patients to appropriate precision clinical trials, thereby improving enrollment efficiency—as demonstrated by companies like Recursion Pharmaceuticals. As a real-world example, the UK National Health Service (NHS) uses AI to screen for rare metabolic disorders in newborns, which has reduced the misdiagnosis rate by 40% and shortened the time to confirmed diagnosis by six months.

3.8. Summary

The proposed integrated framework exhibits comprehensiveness by addressing four key challenges in rare disease research: data scarcity (via dynamic prompting), computational efficiency (via lightweight design), modality heterogeneity (via multimodal alignment), and clinical risk (via ensemble calibration), and it demonstrates synergistic effects during application, as each technical module reinforces the others—for example, the knowledge graph provides semantic support for dynamic prompting, multimodal pretraining generates high-quality features for efficient retrieval, and the lightweight design enables real-world deployment of complex models.

In terms of deployment adaptability, the integrated framework is clinically applicable across settings, from resource-limited primary hospitals to large tertiary institutions with massive datasets, and it can provide customized solutions accordingly.

Finally, looking toward future development, this integrated framework reserves technical interfaces for continual learning and generative AI, enabling adaptation to novel diseases and long-term data iteration cycles.

4. Challenges and Directions for Improvement

First, the paper has recognized that data heterogeneity remains a critical issue, which necessitates the alignment of cross-modal data distributions, such as between imaging and genomic data, and on the data governance level, patient privacy protection must be constantly upheld. In terms of privacy protection, we propose to adopt differential privacy methods to encrypt patient genomic data and use blockchain-based verification to ensure informed consent. To improve the real-time performance and interpretability of dynamic prompting, this research also recommends introducing graph neural networks (GNNs) to accelerate knowledge graph querying and developing attention-based heatmaps or causal inference modules to enhance physician trust in model decision-making. Regarding fairness calibration, we employ adversarial training to reduce biases caused by regional or economic disparities in data, and we ultimately select continual learning as the overarching training paradigm, enabling incremental updates of both knowledge graphs and model parameters to accommodate newly discovered diseases. BenevolentAI, using a knowledge graph–based approach, identified baricitinib as a potential treatment for rare pulmonary fibrosis, successfully advancing the drug into phase III clinical trials [14].

5. Applications of AI in Early Diagnosis and Drug Development

5.1. Innovation in Early Diagnosis

In the field of newborn screening, AI can analyze metabolomics data—such as mass spectrometry results—to quickly identify inherited metabolic disorders like phenylketonuria; in addition, facial recognition technologies such as DeepGestalt have been used to diagnose conditions like Angelman syndrome based on craniofacial features, achieving diagnostic accuracies above 90% [15].

In the area of genetic mutation interpretation, AI can assist in whole-genome sequencing analysis to identify mutations linked to rare diseases, such as those causing spinal muscular atrophy, thereby enhancing precision medicine capabilities in early disease detection.

5.2. Acceleration of Drug Development

In the domain of drug repurposing, graph neural networks (GNNs) can uncover previously unknown relationships between existing drugs and rare diseases—for example, the identification of imatinib as a potential treatment for systemic mastocytosis—and through AI-powered virtual screening platforms, it is possible to simulate drug–target interactions computationally, leading to a reduction of up to 70% in experimental costs, as demonstrated by the Atomwise platform.

Additionally, AI enables the matching of rare disease patients to suitable precision clinical trials, significantly improving trial enrollment rates and personalized treatment development, as exemplified by industry leaders like Recursion Pharmaceuticals.

5.3. Case Studies

The UK National Health Service (NHS) has implemented AI systems for screening rare metabolic diseases in newborns, resulting in a 40% reduction in misdiagnosis rates and shortening time to diagnosis by approximately six months. In another example, BenevolentAI utilized a large-scale biomedical knowledge graph to identify baricitinib as a potential therapeutic candidate for rare pulmonary fibrosis, facilitating its advancement to phase III clinical trials.

6. Conclusion

This paper analyzes the technical methods used in multiple representative studies and publications, summarizes their respective strengths and weaknesses, and, based on this analysis, proposes a scientifically grounded and integrated methodological framework.

In the future, the potential of AI in rare disease research will focus on several key directions: first, deepening interdisciplinary collaboration by integrating genomics, phenotypic data, and multimodal medical information to optimize model precision; second, expanding drug repurposing efforts by mining shared disease mechanisms using AI to accelerate therapeutic innovation; and third, improving policy and industry ecosystems to promote data sharing, ethical governance, and capital investment, thereby facilitating the translation of research into clinical practice.

With continuous technological advancement and collaborative momentum, AI is expected to fundamentally transform the landscape of rare disease diagnosis and treatment, offering more efficient and equitable medical solutions for rare diseases patients worldwide.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] O. David, H. Jordan, D. Hang, G. Fengyi et al. Large Language Models Vote: Prompting for Rare Disease Identification. arXiv preprint arXiv: 2308.12890 (2024).

[2] W. Meng, L. Tian, L. Aidi, Y. Kai, P. Yuanyuan et al. Common and Rare Fundus Diseases Identification Using Vision-Language Foundation Model with Knowledge of Over 400 Diseases. arXiv preprint arXiv: 2406.09317v2 (2024).

[3] C. Chen, M. Lu, Y. Williamson, et al. Fast and scalable search of whole-slide images via self-supervised deep learning. Nat. Biomed. Eng 6, 1420 – 1434 (2022).

[4] T. Ding, S. J. Wagner, A. H. Song, R. J. Chen, et al., Multimodal Whole Slide Foundation Model for Pathology, Arxiv, 2024.

[5] X. Liu, X. Zhao, Z. Zhang, M. Zhang, and H. Zhang, Large language models vote: Prompting for rare disease identification, npj Digital Medicine, vol. 7, no. 1, pp. 1 - 11, 2024.

[6] T. Bachlechner, B. P. Majumder, H. H. Mao, G. W. Cottrell, and J. McAuley, ReZero is all you need: Fast convergence at large depth, arXiv preprint arXiv: 2003.04887, 2020.

[7] Y. Wu, Z. Liu, Y. Chen, Y. Wang, and Y. Zhang, RAREMed: Towards Fair Medication Recommendation for Rare Diseases, arXiv preprint arXiv: 2401.03676, 2024.

[8] C. Chen, M. Y. Lu, D. F. K. Williamson, T. Y. Chen, A. J. Schaumberg, and F. Mahmood, Fast and Scalable Image Search for Histology, arXiv preprint arXiv: 2107.13587, 2021.

[9] S. Singh and S. Chatterjee, A Survey of Autoencoder Algorithms to Pave the Diagnosis of Rare Diseases, Journal of Healthcare Engineering, vol. 2024, pp. 1 - 18, 2024.

[10] Y. Ding, C. Chen, Y. Zhang, and Y. Wang, TITAN: A Multimodal Foundation Model for Whole-Slide Image Analysis, arXiv preprint arXiv: 2401.03677, 2024.

[11] N. Hurvitz, H. Azmanov, A. Kesler, and Y. Ilan, Establishing a second - generation artificial intelligence - based system for improving diagnosis, treatment, and monitoring of patients with rare diseases, *European Journal of Human Genetics*, vol. 29, no. 10, pp. 1485 - 1490, 2021.

[12] I. Segura-Bedmar, D. Camino-Perdones, and S. Guerrero-Aspizua, Exploring deep learning methods for recognizing rare diseases and their clinical manifestations from texts, *BMC Bioinformatics*, vol. 23, no. 1, p. 263, 2022.

[13] J. K. Park and W. J. Choi, Rare disease genomics and precision medicine, *Genomics & Informatics*, vol. 22, no. 1, pp. 1 - 11, 2024.

[14] N. Hasani, F. Farhadi, M. A. Morris, M. Nikpanah, A. R. Mim, Y. Xu, A. Pariser, M. T. Collins, R. M. Summers, E. Jones, E. Siegel, and B. Saboury, Artificial intelligence in medical imaging and its impact on the rare disease community: Threats, challenges and opportunities, *Clinical PET Imaging*, vol. 5, no. 1, pp. 12 - 21, 2022.

[15] P. Ou, R. Wen, L. Shi, J. Wang, and C. Liu, Artificial intelligence empowering rare diseases: A bibliometric perspective over the last two decades, *Orphanet Journal of Rare Diseases*, vol. 19, no. 1, p. 345, 2024.