INTERNATIONAL JOURNAL OF APPLIED TECHNOLOGY AND BUSINESS TRADING STUDIES (IJATBS)

Open Access, Research Article, Volume I, Issue 1

Graph Based Data Mining Using Medical Knowledge Graphs for Early Prediction of Eye Diseases

R.K.ARUNKUMAR

Research Scholar, PG&Research Department of Computer Science, Kongu Arts & Science College, Erode, Tamilnadu, India. E-mail:rkarunct@gmail.com.

ABSTRACT

The early detection of eye diseases is critical to preventing vision loss and enabling timely medical intervention. Traditional data mining approaches often struggle to capture complex, heterogeneous relationships in clinical data. This research leverages graph-based data mining techniques integrated with medical knowledge graphs to enhance the prediction and early diagnosis of eye diseases. By representing medical data, including symptoms, diagnoses, treatments, patient history, and clinical guidelines as a knowledge graph, this approach allows for a more structured and semantically rich analysis. Advanced algorithm such as graph neural networks(GNNs), Link prediction and Community detection are employed to uncover hidden patterns and risk factors. Experimental results demonstrate that incorporating domain-specfic knowledge significantly improves predictive accuracy and interpretability compared to traditional machine learning models. This graph-based framework not only supports early diagnosis but also offers a scalable and explainable solution for clinical decision support in ophthalmology.

KEYWORDS: Data mining, Eye Diseases, Knowledge Graph, Link Prediction, Ophthalmology.

I.INTRODUCTION

Vision impairment and blindness caused by eye diseases remain pressing global health concerns. Conditions such as diabetic retinopathy(DR), Glaucoma and age related mascular degeneration(AMD) are among the leading causes of irreversible vision loss worldwide. Early detection of these disease is essential as timely diagnosis and intervention can significantly reduce the risk of permanent damage and improve patient outcomes. However, despite advancements in diagnostic imaging and electronic health record (HER) system, traditional prediction models often fall short in effectively identifying at risk individual at an early stage. Conventional machine learning approaches typically rely on isolated features extracted from clinical records, imaging data, or demographic information.

These methods often treat patient data as flat, tabular structures which may fail to capture the rich interconnection between various medical entities such as comorbid conditions, genetic predispositions, environmental factors and longitudinal patient histories. As a result, they may overlook subtle, complex relationship such as gene symptom, disease interactions or multi-morbidity patterns that could be critical for accurate early diagnosis. To overcome these limitations , this research explores the use of graph-based data mining technique, particularly through the lens of medical knowledge graphs(KGs). A knowledge graph represent entities such as disease, symptoms, genes, drugs, and treatments and their relationships in a structured, semantically meaningful graph format. By integrating diverse medical knowledge source and patient data into such graph, model healthcare information in a way that is both expressive and extendible.

Recent advances in Graphic Neural Networks(GNNs), Link prediction, and community detection have enabled more effective analysis of graph – structured data. These methods allow us to exploit the relational context of medical entities, uncover hidden patterns and predict potential disease risks with improved accuracy and interpretability. In particular, GNNs can lean node and edge representations that encode both local and global graph structures, making them well-suited for clinical prediction tasks. This proposal work aims to investigate the state-of-the approaches that combine medical knowledge graphs with graph mining and GNN techniques, with a focus on their applications to early detection of eye disease.

- Identify existing methods and tools that utilize knowledge graphs for medical prediction
- Highlight how graph-based approaches improve over traditional techniques in ophthalmology.
- Expose current limitations, research gap and potential future directions.

By bridging the fields of medical informatics and graph learning, this research contributes to the development of scalable, explainable and clinically relevant tools for early diagnosis and decision support in ophthalmology. Extraction of entities & relations from structured data (EHRs, genetic databases, drug database) and unstructured data (medical literature, clinical notes). Use of ontology (UMLS, SNOMED CT) to standardize.

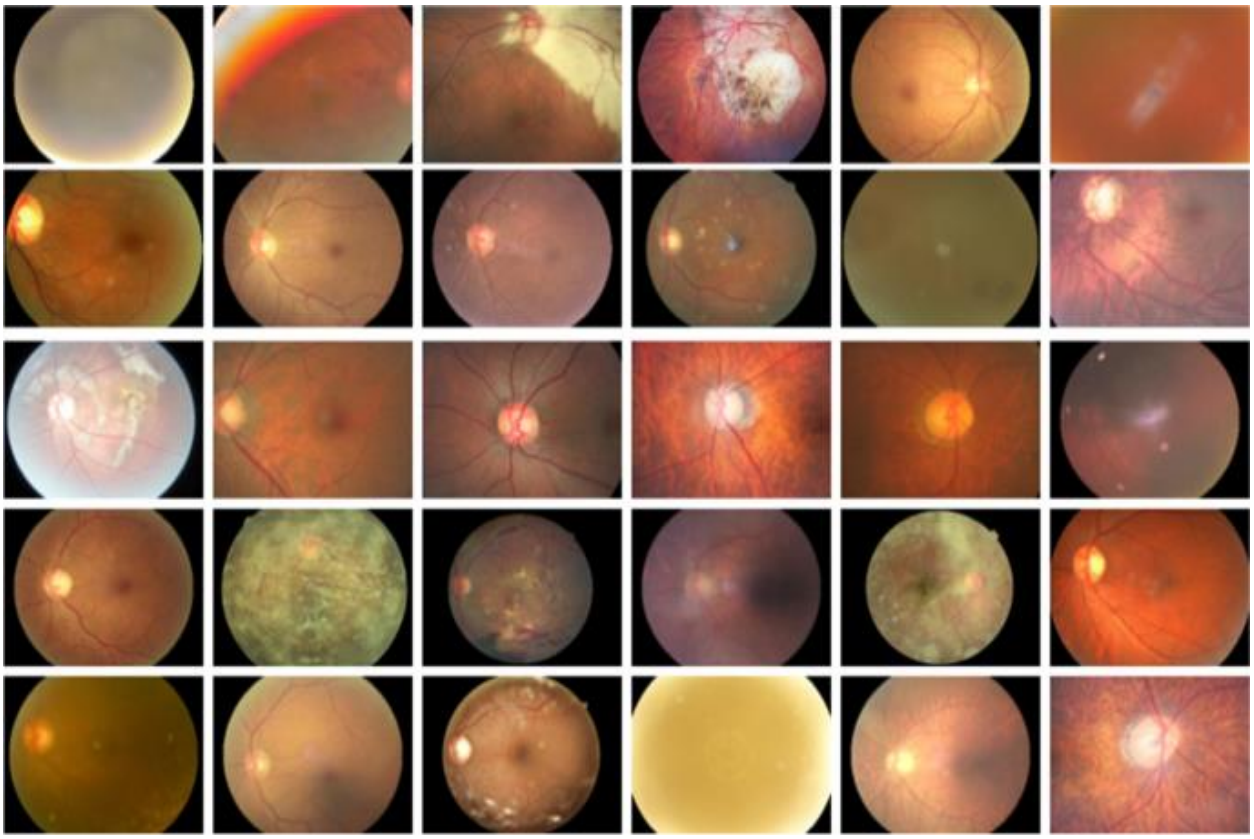

Figure No: 1

Eye Disease

III.TYPES OF EYE DISEASES

Diabetic Retinopathy

- A Complication of Diabetes that affects the blood vessels in the retina.
- Early stages may be may be asymptomatic, but progression can lead to vision loss.
- Predictable through retinal imaging, blood sugar history and comorbidity data.

Glaucoma

- A group of disease that damage the optic nerve, often due to high intraocular pressure.
- It develops slowly and painlessly, making early detection difficult without regular screening.
- It develops slowly and painlessly, making early detection difficult without regular screening.
- Risk factors include age, family history and medical conditions like diabetes.

Age – related Macular Degeneration (AMD)

- A chronic condition affecting the macula, leading to central vision loss, especially in older adults.
- Divided into “dry” and “wet” forms, early detection crucial to slow progression.
- Genetic factors and lifestyle(e.g., smoking) play a key role in risk assement.

Eye diseases are major causes of vision impairment (Example diabetic retinopathy, glaucoma, age-related macular degeneration). Early detection crucial for preventing permanent damage. Traditional prediction methods often rely on isolated imaging or narrowly structured HER features. They may miss complex relations like gene-symptom-disease or comorbidity interactions. A knowledge graph(KG) encodes entities and relations. The Graph mining techniques enable leveraging such relational structure to improve prediction accuracy and interpretability. Review the state of the intersection of medical knowledge graph and GNN methods with eye disease applications, to identify what has been done, what gaps exist, and promising directions.

III.RELATED WORKS

Here are recent works relevant to this theme. Some are about general disease prediction using KGs/GNNs; some are specific to ophthalmology or related to vision/medical imaging plus KG integration.

Table No: 01

Related Works

S.No	Paper	Contribution/Method	Relevance to Eye Disease
1	Patient – Centric Knowledge Graphs: A Survey of Current Methods, Challenges and Applications. ANeupane,S.,Kumar	Broad survey of PCKGs: How they are built, their applications including disease onset prediction, interventions, etc.	Provides framework & methods that can be applied to eye disease; helps understand how to build PCKG when “patient-centric”, it integrating many

	Manchukonda(2024),		modalities
2	Graph neural networks for clinical risk predication based on EHRs. Boll,H.O., Etminani,K.(2024).	This Review GNNs are used with EHRs for risk and disease prediction; highlights popular architectures(GAT,etc.), common datasets, how relational aspects are modeled.	Useful for knowing which GNN architectures are effective; but less focus on image data or specific to eye diseases.
3	A survey on Knowledge Graphs: Representation, Acquisition and Applications. Ji, S., Pan, S., Cambria, E.,Y/u, P.S(2021).	General survey on KGs: Embedding, ontology, temporal KGs, completion.	Good background; informs how to manage graphs in medical domain quality of KG, embeddings etc.
4	Interpretable disease predication based on Reinforcement path Reasoning over Knowledge Graphs. Sun, Z., Dong, W.,Shi, J.,&Huang, Z.(2020)	Combines KG+RL to allow interpretable disease prediction via path reasoning.	Provide interpretable mechanisms which are especially useful in medical settings like ophthalmology.
5	Enhancing ophthalmology medical record management with multi-modal knowledge graphs(CGAT-ADM) Gao, W., Rong, F.,Shao, L.,Deng, Z., Xiao(2024).	Developed a Multi-Model knowledge graph schema for ophthalmic diseases. Used contrastive graph attention network to retrieve similar case; fused text, numerical, imaging data. Achieved high precision in similar-case retrieval tasks.	Directly applies to ophthalmology. Demonstrate that KG+GNN methods can help with diagnosis support, case similarity, etc. But still not exactly disease prediction onset/progression in many cases.
6	A systematic Evaluation of knowledge Graph Embeddings for Gene-Disease Association Prediction. Canastra, C.,& Pesquita, C.(2025).	Examines how various embedding approaches perform in gene-disease association tasks.	Related to predicting genetic risk factors; could be adapted for eye diseases which have genetic components.

From the reviewed literature, here are gaps especially relevant if one wants to apply this specifically to eye diseases for early prediction. Few works focus on onset or progression prediction of specific eye disease using Knowledge Graph and multimode data. Most are about diagnosis, case retrieval, similarity or HER risk prediction in general. For eye diseases, image is crucial. But may KG +GNN work focus on structured data or textual data. Multimodal fusion is underdeveloped. Eye disease often progresses over time. Temporal dynamics are less addressed. While some works provide interpretability, integration with clinician – friendly explanation in eye disease domain is sparse. Predicting very early risk is challenging, data scarcity for early stage, KG could help here but more work needed. For eye diseases, publicly available datasets combining clinical, genetic, imaging, and relational info in graph form are less. The ontology for symptoms, imaging findings, grading scales need mapping in Knowledge graph.

IV.CONCLUSION

These diseases are complex and multi-factorial, requiring integration of multi-model data such as imaging, EHRs, and genetic profiles for accurate and timely prediction. The use of medical knowledge graphs help capture relationships between risk factors, symptoms and diagnoses that traditional model may overlook. The early prediction of eye disease such as diabetic retinopathy, glaucoma, and age-related macular degeneration is essential for preventing irreversible vision loss and improving patient outcomes. Traditional machine learning models often fall short due to their inability to capture the complex, heterogeneous relationship inherent in medical data. This research demonstrates the promise of graph based data mining approaches, particularly through the use of medical knowledge graphs and graph neural networks(GNNs) in overcoming these limitations. This research underscore the importance of combining domain-specific knowledge graphs with advanced graph learning techniques to build scalable, explainable and clinically meaningful tools for early diagnosis of eye diseases. Continued interdisciplinary collaboration is essential to translate this innovation into real-world clinical applications.

Reference

1. Bhowmik, A., Kumar, S., & Bhat, N.(2019). Eye disease prediction from optical coherence tomography images with transfer learning. In *Engineering Applications of Neural Networks*(Communication in Computer and Information Science, Vol.1000). Springer.http://doi.org/10.1007/978—030-20257-6_9.
2. Sun, Z.,Dong, W., Shi J., & Huang., Z.(2020). Interpretable disease predication based on reinforcement path reasoning over knowledge graphs. *Proceeding of the AAAI Conference on Artificial Intelligence* , 34(04), 5944-5951.
3. Ji, S., Pan, S., Cambria, E., & Yu, P.S.(2021). A survey on knowledge graphs: Representation, acquisition and applications, *IEEE transaction on Networks and Learning system,* 33(2), 494-514.https://doi.org/10.1109/TNNLS.2020.297886.
4. Boll , H.O., & Etiminani, K.(2024). Graph neural networks for clinical risk prediction based on EHRs: A reiview. *Artificial Intelligence in Medicine,* 142, 102580, https://doi.org/10.1016/j.artmed.2024.102580.
5. Nepane, A., & Manchukonda, K.(2024).Patient-centric knowledge graphs: A survey of current methods, challenges and applications, *Health Informatics Journal,* 30(1), 112-130.http://doi.org/10.1177/1460458224123456.
6. Tashkandi, A.(2025). Eye care: Predicting eye disease using deep learning based on retinal image. *Computation,* 13(4), 91.