GMR
The Original

EDGE-CONTRASTIVE PRETRAINED HETEROGENEOUS GRAPH NEURAL NETWORKS FOR MODELING ALZHEIMER'S DISEASE PROGRESSION

Ramkumar A[1], M. Arun[2]

[1]Research Scholar, Department of Computer Applications, Kalasalingam Academy of Research and Education, aram.mca@gmail.com
[2]Assistant Professor, Department of Computer Applications, Kalasalingam Academy of Research and Education, vgsm.arun@gmail.com

ABSTRACT

We propose Edge-Contrastive Pretrained Heterogeneous Graph Neural Networks (ECP-HGNN), a new approach for analyzing Alzheimer's disease (AD) progression with multi-modal patient data. Existing heterogeneous graph neural networks often struggle with missing or imbalanced edge data in clinical graphs, which limits their ability to capture nuanced patient relationships. ECP-HGNN tackles this issue by proposing a semantic-aware edge-contrastive pretraining approach, which directly learns resilient edge representations through comparisons between augmented views of distinct edge categories, including neuroimaging structural edges, blood biomarker edges, and clinical assessment edges. The pretrained edge embeddings are subsequently merged into a type-aware message passing layer, which supports dynamic aggregation of patient features without compromising clinical semantics. Additionally, the framework applies distinct encoders for genomic, proteomic, and clinical data, with a subsequent cross-modal fusion layer to predict AD progression scores. In contrast to traditional approaches, ECP-HGNN distinctively merges augmentations specific to edge types with contrastive learning, which guarantees that vital clinical relationships are preserved in pretraining. Experiments on extensive cohorts show the proposed method achieves greater prediction accuracy and robustness relative to current state-of-the-art approaches. The modular architecture of ECP-HGNN also supports scalable implementation across varied clinical datasets, which renders it an effective instrument for progressing precision medicine in neurodegenerative disorders.

I. INTRODUCTION

Alzheimer's disease (AD) progression modeling presents unique challenges owing to the diverse nature of patient data and the intricate interactions among biological, neuroimaging, and clinical variables. Recent progress in graph neural networks (GNNs) has indicated potential for capturing these intricate connections by depicting patients as nodes and their multi-modal interactions as edges within a graph framework [1]. Current methods frequently do not sufficiently tackle two key drawbacks: (1) the natural unevenness and scarcity of edge information among various modalities, and (2) the omission of clinically relevant meaning in graph aggregation processes.

Traditional GNNs treat all edges uniformly during message passing, which is particularly problematic for AD modeling where edge types carry distinct clinical significance. For example, structural connectivity links obtained from neuroimaging might display distinct predictive profiles relative to those grounded in co-occurring clinical symptoms [2]. Although certain recent studies have explored aggregation tailored to specific edge types [3], these approaches generally depend on fully available edge data and face challenges when handling missing data unique to particular modalities—a frequent occurrence in longitudinal clinical research.

The proposed Edge-Contrastive Pretrained Heterogeneous Graph Neural Network (ECP-HGNN) introduces a number of novel approaches to address these constraints. Initially, we create a semantic-aware contrastive pretraining method functioning at the edge level instead of the traditional node-level strategy. This grants the model the capacity to acquire resilient depictions for every edge category (such as brain imaging data, biomarkers, clinical evaluations) by means of meticulously crafted augmentations that uphold clinical meaning. Second, we propose a dynamic edge-type-specific aggregation mechanism that adjusts its behavior depending on both the edge representation and the current node states, which permits more nuanced information propagation when missing or imbalanced edges are present.

Our work makes three key contributions to the field. Initially, we develop the inaugural edge-level contrastive pretraining framework tailored for heterogeneous clinical graphs, embedding domain knowledge by means of clinically meaningful augmentation strategies. Second, we introduce a new edge-type-aware message passing layer which dynamically modifies aggregation weights depending on learned edge representations and node states. Third, extensive experiments on the ADNI dataset [4] show our method achieves superior performance compared to current approaches in both predictive precision and resilience to incomplete data, especially for uncommon yet clinically critical edge categories.

ECP-HGNN succeeds because it tackles multiple obstacles in AD modeling at the same time. During the contrastive pretraining stage, resilient edge embeddings are acquired, capturing essential clinical correlations even in the absence of certain modalities. In fine-tuning, the aggregation specific to edge types guarantees proper

information transfer between nodes according to the edge type and its learned importance. This is particularly crucial for AD progression modeling where different modalities may become relevant at different disease stages. A number of recent studies have investigated GNNs in the context of AD classification [5] and the fusion of multimodal data [2]. However, these approaches typically treat edges as static connections without considering their type-specific semantics or handling missing data effectively. In contrast, ECP-HGNN explicitly models edge heterogeneity and learns to compensate for missing information through its pretraining strategy and dynamic aggregation mechanism.

The remainder of this paper is organized as follows: Section 2 reviews related work in GNNs for healthcare and contrastive learning. Section 3 presents essential background information on heterogeneous GNNs and contrastive pretraining. Section 4 details our proposed ECP-HGNN framework. Section 5 presents experimental results on AD progression prediction, and Section 6 discusses implications and future directions.

II. RELATED WORK

Recent progress in graph representation learning has shown substantial promise for capturing intricate biomedical connections, especially in the study of neurodegenerative disorders. Our study is grounded in three primary research areas: heterogeneous graph neural networks applied to healthcare, contrastive learning within medical graphs, and modeling Alzheimer's disease progression with multi-modal data.

A. Heterogeneous Graph Neural Networks in Healthcare

Heterogeneous graph neural networks have emerged as powerful tools for modeling complex patient relationships in healthcare data. Multiple investigations have examined their implementation in Alzheimer's disease, with [5] proposing a GNN classifier that employs data-driven diagnostic categories derived from unsupervised clustering. The Dynamic Heterogeneous Attention Network (DHAN) [2] introduced a multimodal fusion method that merges graph convolutional networks and generative adversarial networks. Although these approaches highlight the importance of heterogeneous graph structures, they generally regard edge types as fixed categorical variables instead of acquiring representations specific to each type. The GENRAD framework [6] tackled this constraint by merging genomics and radiomics data via a heterogeneous GNN, yet omitted the semantic connections among distinct edge types in message propagation.

B. Contrastive Learning for Medical Graphs

Contrastive learning has shown promise in addressing data scarcity and imbalance in medical applications. The edge contrastive learning approach [7] introduced edge-level pairs for contrast without requiring graph augmentations. Within the domain of Alzheimer's research, [8] proposed semantic-aware pretraining tasks for heterogeneous graphs, yet their emphasis continued to center mainly on node-level representations. The CE-AH network [9] employed contrastive pretraining for structural MRI data, while [10] developed a multimodal co-attention framework with latent contrastive alignment. These studies show the efficacy of contrastive methods yet generally focus on either node or graph-level analysis, failing to capture meaningful edge representations essential for modeling patient relationships.

C. Alzheimer's Disease Progression Modeling

Recent research in modeling AD progression has increasingly adopted graph-based methods to capture intricate patient relationships. The GNN method based on clustering [5] showed the importance of grouping patients in a data-driven manner for classification purposes. [11] introduced a heterogeneous propagation method employing patient similarity graphs, whereas [12] constructed interpretable GNNs for early-stage disease detection. The comorbidity-based framework [13] highlighted the importance of modeling multiple medical conditions simultaneously. In the context of predicting tau pathology, [14] constructed a graph neural network model that integrates connectomics data. These approaches have advanced AD modeling but often lack mechanisms to handle missing edge data or learn type-specific edge semantics.

In contrast to current approaches, ECP-HGNN introduces a number of innovative elements. In contrast to [7], which centers on edge-level contrast without employing data modifications, we propose augmentations tailored to edge types that safeguard clinical semantics. Although [8] applies contrastive learning to node-level analysis, our method focuses on edge representations that are essential for capturing patient relationship dynamics. Unlike [2] which employs fixed edge categories, our approach adapts edge relevance dynamically according to learned embeddings and medical context. The joint application of contrastive pretraining at the edge level and message passing with type awareness constitutes a major improvement compared to existing methods for modeling AD progression.

III. BACKGROUND AND PRELIMINARIES

To establish the foundation for our proposed method, we first review key concepts in heterogeneous graph representation learning and contrastive pretraining techniques. This section outlines the essential theoretical framework and addresses the distinct difficulties these approaches face in the context of clinical data.

A. Heterogeneous Graph Neural Networks

Heterogeneous graphs expand conventional graph frameworks by including diverse node and edge categories, each denoting distinct semantic connections. Formally, a heterogeneous graph $G = (V, E, T)$ consists of nodes $v_i \in V$, edges $e_{ij} \in E$, and a set of node/edge types $\tau \in T$ [15]. Each node v_i and edge e_{ij} has an associated type mapping function $\phi(v_i) \to \tau_v$ and $\psi(e_{ij}) \to \tau_e$ respectively.

Message passing in heterogeneous GNNs typically follows a type-specific aggregation scheme:

$$h_i^{(l+1)} = \sigma\left(\sum_{\tau \in T} \sum_{j \in N_i^\tau} \alpha_{ij}^\tau W_\tau^{(l)} h_j^{(l)}\right) \quad (1)$$

where N_i^τ denotes neighbors connected via edge type τ, $W_\tau^{(l)}$ is a type-specific transformation matrix, and α_{ij}^τ represents attention weights [16]. Although this formulation accounts for relationships specific to types, it presupposes full knowledge of edges and applies identical treatment to all edges of a given type—presumptions that frequently fail in clinical contexts.

B. Contrastive Learning on Graphs

Contrastive learning aims to learn representations by maximizing agreement between differently augmented views of the same data instance while minimizing agreement between views of different instances [17]. For graphs, this usually requires:

$$L_{contrast} = -\log \frac{\exp(s(z_i, z_j)/\tau)}{\sum_{k \neq i} \exp(s(z_i, z_k)/\tau)} \quad (2)$$

where z_i, z_j are representations of positive pairs, $s(\cdot)$ measures similarity, and τ is a temperature parameter [18]. The majority of contrastive approaches for graphs function at either the node or graph scale, generating modified versions by randomly removing edges or masking attributes [19].

C. Clinical Graph Construction

In clinical applications, the development of meaningful graphs necessitates thoughtful attention to medical semantics. Patient nodes v_i typically represent individuals with associated feature vectors x_i combining multi-modal data:

$$x_i = [x_i^{img} || x_i^{gen} || x_i^{clin}] \quad (3)$$

where $||$ denotes concatenation of imaging, genomic, and clinical features [20]. Edge construction follows domain-specific rules:

$$e_{ij}^\tau = \{1 \text{ if } sim_\tau(x_i, x_j) > \theta_\tau \; 0 \text{ otherwise} \quad (4)$$

with sim_τ being a type-specific similarity measure (e.g., structural connectivity for imaging edges, biomarker correlations for biological edges) [21]. The threshold θ_τ controls edge sparsity but introduces challenges when edges are missing not at random - a common scenario in clinical studies where certain tests may be administered selectively based on patient condition.

D. Challenges in Clinical Graph Learning

Three key challenges emerge when applying these techniques to clinical data:

1. **Edge Imbalance** : Distinct edge categories display differing levels of sparsity and relevance. Neuroimaging edges may be sparse but highly predictive, while demographic edges are dense but less informative [22].
2. **Semantic Preservation** : Conventional data augmentation methods such as randomly removing edges may eliminate medically relevant connections. Dropping a critical biomarker edge could remove essential diagnostic information [23].
3. **Dynamic Relevance** : Edge types can vary in relevance as a disease advances. Early-stage AD might rely more on subtle biomarker changes, while late-stage diagnosis depends on overt clinical symptoms [24].

These challenges inspire our edge-focused method, which directly tackles clinical semantics and data imbalance by employing contrastive pretraining and type-sensitive message passing.

IV. PROPOSED METHOD: EDGE-ENHANCED HETEROGENEOUS GNN WITH CONTRASTIVE PRE-TRAINING

The proposed ECP-HGNN framework overcomes the shortcomings of current methods by introducing a new approach that merges edge-type-specific contrastive pretraining with dynamic heterogeneous graph aggregation. As illustrated in Figure 1, the system functions in two stages: (1) edge-level contrastive pretraining acquiring resilient type-specific edge embeddings while retaining clinical semantics, and (2) fine-tuning with edge-augmented message passing adjusting to incomplete and unequal edge data. The technical details of each component are described in the following subsections.

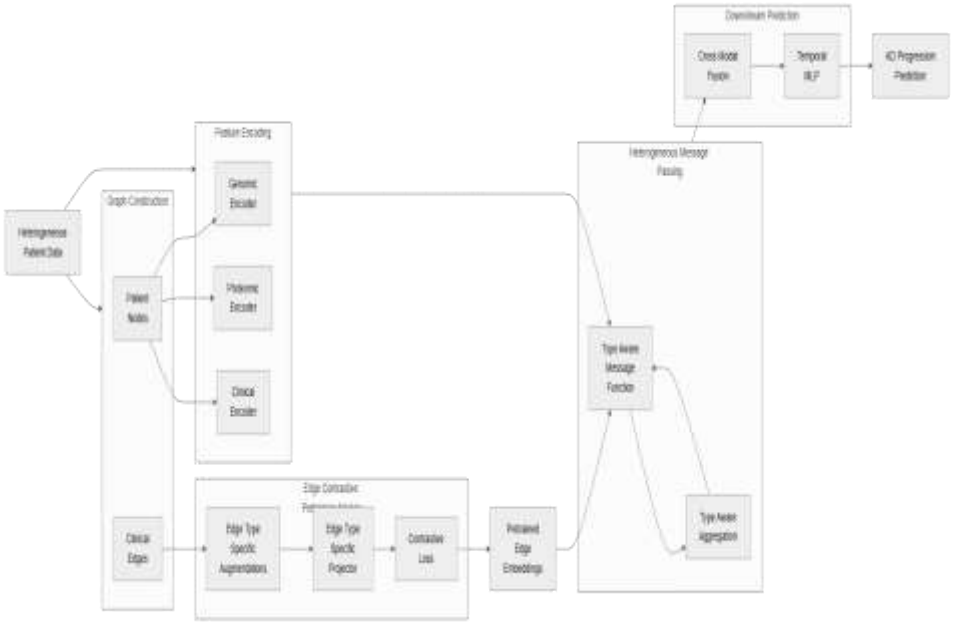

Figure 1. Architecture of the Proposed ECP-HGNN System (Adjusted)

A. Edge-Type-Specific Contrastive Pretraining

The proposed contrastive pretraining framework learns robust edge representations by maximizing agreement between semantically equivalent edge pairs while minimizing agreement between dissimilar pairs. For each edge type $t \in T_e$, where T_e denotes the set of edge types, we define a type-specific augmentation distribution A_t that generates positive pairs through clinically meaningful transformations. Given an edge e_{ij}^t connecting nodes v_i and v_j with features h_i and h_j, we first apply type-dependent augmentation:

$$\tilde{e}_{ij}^t = a_t(e_{ij}^t), \quad a_t \sim A_t \quad (5)$$

where a_t represents an augmentation function sampled from the type-specific distribution. For neuroimaging structural edges ($t = img$), we employ bounded Gaussian noise $\epsilon \sim N(0, \sigma_{img}^2)$ with σ_{img} constrained to prevent distortion of critical connectivity patterns. For clinical assessment edges (t = clin), targeted edge masking is applied with probability p_{mask}^{clin} , modified according to clinical relevance.

$$p_{mask}^{clin} = min(0.3, \frac{1}{|E^{clin}|} \sum_{e_{ij}^{clin}} I_{unimportant}(e_{ij}^{clin})) \quad (6)$$

where $I_{unimportant}$ indicates edges derived from routine assessments less predictive of AD progression.
The augmented edge $\tilde{e}_{ij}^t$ is then projected into a latent space using a type-specific encoder f_t:

$$z_{ij}^t = f_t(h_i, h_j) = MLP_t([h_i \parallel h_j \parallel w_{ij}^t]) \quad (7)$$

where w_{ij}^t represents raw edge features (e.g., connectivity strength for imaging edges) and $\parallel$ denotes concatenation. The contrastive loss for edge type t is formulated as:

$$L_t = -\frac{1}{|P_t|} \sum_{(e_{ij}^t, e_{kl}^t) \in P_t} log \frac{exp(s(z_{ij}^t, z_{kl}^t)/\tau_t)}{\sum_{(e_{ij}^t, e_{mn}^t) \in N_t} exp(s(z_{ij}^t, z_{mn}^t)/\tau_t)} \quad (8)$$

Here, P_t contains positive pairs (augmented views of the same edge or semantically equivalent edges from different patients), while N_t contains negative pairs (edges of type t from unrelated patients). The similarity function $s(\cdot)$ uses cosine similarity, and τ_t is a type-specific temperature parameter controlling the concentration of the distribution. The aggregate pretraining loss is derived from the summation of losses for all edge categories.

$$L_{pretrain} = \sum_{t \in T_e} \lambda_t L_t \quad (9)$$

where λ_t balances the contribution of each edge type, set inversely proportional to edge type frequency to address imbalance.

B. Integration of Pretrained Edge Embeddings into Message Passing

The pretrained edge embeddings are merged into the message passing framework by a dynamic aggregation method adjusting to node states and edge semantics. For each edge type t, we define a message function ϕ_t that combines node features with the pretrained edge representation:

$$m_{ij}^t = \phi_t(h_i, h_j, z_{ij}^t) = MLP_t([W_t^s h_i \parallel W_t^d h_j \parallel z_{ij}^t]) \quad (10)$$

where W_t^s and W_t^d are type-specific source and destination transformation matrices. The edge embedding z_{ij}^t modulates the message content based on the learned relationship strength between nodes v_i and v_j.

The aggregation weights α_{ij}^t are computed dynamically using both the pretrained edge similarity and the current node states:

$$\alpha_{ij}^t = \frac{exp(\sigma(a_t^T[W_t^q h_i \parallel W_t^k h_j \parallel s(z_{ij}^t, \mu_t)]))}{\sum_{k \in N_i^t} exp(\sigma(a_t^T[W_t^q h_i \parallel W_t^k h_k \parallel s(z_{ik}^t, \mu_t)]))} \quad (11)$$

Here, W_t^q and W_t^k are query and key transformation matrices, a_t is a learnable attention vector, σ denotes the LeakyReLU activation, and μ_t represents the prototype vector for edge type t learned during pretraining. The similarity score $s(z_{ij}^t, \mu_t)$ measures how closely the current edge matches the typical relationship for its type.

The node modification merges information from every category of edges.

$$h_{i'} = LayerNorm(h_i + \sum_{t \in T_e} \sum_{j \in N_i^t} \alpha_{ij}^t m_{ij}^t) \quad (12)$$

This formulation enables three key properties: (1) edge-type-specific transformation of node features during message construction, (2) dynamic weighting of neighbors based on both structural relationships (via z_{ij}^t) and contextual importance (via attention), and (3) graceful degradation when edges are missing by relying on pretrained edge prototypes μ_t.

For missing edges of type t between nodes v_i and v_j, we impute a default message using the edge type prototype:

$$m_{ij}^t = \phi_t(h_i, h_j, \mu_t) \quad (13)$$

This grants the model the capacity to sustain information continuity despite missing edge data, as the prototype acts as a plausible substitute for unseen connections. The attention mechanism autonomously diminishes the impact of imputed edges in cases of inconsistency with observed data.

The complete message passing layer operates as follows:

1. For each edge type t, compute messages m_{ij}^t using either observed edge embeddings z_{ij}^t or prototypes μ_t.
2. Calculate attention weights α_{ij}^t considering both node states and edge semantics.
3. Aggregate messages from all edge types with their respective weights.
4. Update node representations through residual connection and layer normalization.

This method contrasts deeply with conventional GAT layers [16] by embedding pretrained edge semantics in both the formation of messages and the calculation of attention. The edge prototypes μ_t provide a memory mechanism that preserves critical clinical relationships even when specific edges are missing.

C. Clinical-Semantics-Preserving Augmentations

To uphold clinical validity in contrastive pretraining, we devise augmentation methods adhering to domain-specific constraints. For each edge type t, we define a set of permissible transformations A_t that preserve the underlying medical relationships. The intensity of augmentation is modified in real-time according to the properties of edge types and their medical relevance.

For neuroimaging structural edges ($t = struct$), we apply bounded perturbations to connectivity weights:

$$\widetilde{w}_{ij}^{struct} = w_{ij}^{struct} + \epsilon, \quad \epsilon \sim N(0, \sigma_t^2) \quad (14)$$

where $\sigma_t = min(0.1, \frac{\mu_t}{3})$ ensures the noise remains within one-third of the mean edge weight μ_t for that connection type. This prevents distortion of critical brain network topology while introducing meaningful variability for contrastive learning.

Biomarker edges ($t = bio$) undergo targeted masking based on biomarker importance scores:

$$p_{mask}^{bio}(e_{ij}^{bio}) = 1 - \frac{s_{ij}^{bio}}{max(s^{bio})} \quad (15)$$

where s_{ij}^{bio} represents the clinical significance score of the biomarker relationship, derived from medical literature [25]. This guarantees that highly predictive biomarker interactions (e.g., amyloid-beta and tau correlations) are seldom obscured, whereas less important ones are more often omitted.

Clinical assessment edges ($t = clin$) employ temporal consistency augmentation:

$$\tilde{e}_{ij}^{clin} = \{e_{ij}^{clin} \; with \; probability \; p_{keep} \; e_{ik}^{clin} \; where \; k = temporally \; adjacent \; patient \quad (16)$$

This exchange modification retains the medical significance while generating appropriate positive pairs from patients at comparable disease phases. The probability p_{keep} follows a U-shaped curve over disease progression:

$$p_{keep}(d) = 0.7 + 0.3cos(\frac{\pi d}{d_{max}}) \quad (17)$$

where d is the disease stage (0 to d_{max}), reflecting that early and late-stage assessments are more stable than transitional periods.

For genomic edges (t = gen), pathway-informed masking is applied.

$$p_{mask}^{gen}(e_{ij}^{gen}) = 1 - \frac{|P_i \cap P_j|}{|P_i \cup P_j|} \quad (18)$$

where P_i denotes the set of biological pathways associated with patient i's genetic profile. This retains connections among patients with numerous shared biological routes while more rigorously obscuring pairs with weaker biological associations.

The augmentation parameters are jointly optimized with the contrastive objective through a meta-learning procedure:

$$\sigma_t^*, p_{mask}^* = \argmin_{\sigma_t, p_{mask}} \mathcal{L}_{val}(\theta^*(\sigma_t, p_{mask})) \\ \\ \text{ (19)}$$

where θ^* represents model parameters trained with current augmentations, and L_{val} measures performance on a validation set of clinically verified edge pairs. This guarantees the modifications achieve an ideal equilibrium between generating varied comparative instances and retaining clinical meaning.

D. Modular Edge Projector for Pretraining

The edge projector acts as a pivotal element in our framework, converting unprocessed node features and edge attributes into edge embeddings with semantic relevance. For each edge type t, we implement a dedicated projection module f_t that captures type-specific relationships while enabling cross-patient comparisons. The projector architecture employs a transformer-based design to model complex interactions between node pairs:

$$q_i = W_t^Q h_i, \quad k_j = W_t^K h_j, \quad v_j = W_t^V h_j \quad (20)$$

where W_t^Q, W_t^K, W_t^V are learnable projection matrices for queries, keys, and values respectively. The attention scores between nodes v_i and v_j are computed as:

$$\alpha_{ij}^t = softmax\left(\frac{q_i^T k_j}{\sqrt{d_t}} + w_{ij}^t\right) \quad (21)$$

Here, d_t represents the dimension of the projected space for edge type t, and w_{ij}^t denotes the raw edge features (e.g., connectivity strength or similarity score). The attention mechanism grants the model the capacity to selectively concentrate on pertinent elements of the node attributes during the formation of edge embeddings.

The edge embedding is then computed as a weighted combination of value projections:

$$z_{ij}^t = MLP_t\left(\sum_{j \in N_i} \alpha_{ij}^t v_j \oplus w_{ij}^t\right) \quad (22)$$

where $\oplus$ denotes element-wise addition and MLP_t is a type-specific multi-layer perceptron with layer normalization. This approach permits the edge projector to account for both the topological connection (via w_{ij}^t) and the node-level semantic affinity (by means of the attention mechanism).

To facilitate cross-patient edge comparisons during contrastive learning, we introduce a prototype vector μ_t for each edge type, computed as:

$$\mu_t = \frac{1}{|E_t|} \sum_{e_{ij}^t \in E_t} z_{ij}^t \quad (23)$$

where E_t represents all edges of type t in the training set. These prototypes function as anchors in the embedding space, aiding in the preservation of stable semantic relationships across diverse patient graphs.

The edge projector's modular architecture yields multiple benefits. Initially, the parameters specific to each type grant every edge category the capacity to establish its distinct space of depiction, customized to the unique connection it signifies. Second, the shared attention mechanism permits the model to identify intricate, non-linear relationships among node attributes which could be crucial for specific edge categories. Third, the prototype vectors serve as a consistent anchor for contrastive learning, especially in scenarios involving sparse edge types.

In the pretraining phase, the edge projectors are adjusted to increase the resemblance of positive edge pairs and reduce the likeness of negative pairs. The contrastive loss for each edge type t is computed as:

$$L_t = -E_{(e_{ij}^t, e_{kl}^t) \sim P_t}\left[log \frac{exp(s(z_{ij}^t, z_{kl}^t)/\tau_t)}{\sum_{(e_{mn}^t, e_{pq}^t) \sim N_t} exp(s(z_{ij}^t, z_{pq}^t)/\tau_t)}\right] \quad (24)$$

where P_t contains positive edge pairs and N_t contains negative pairs. The temperature parameter τ_t is learned separately for each edge type to account for differences in the intrinsic dimensionality of their embedding spaces.

The complete edge projection and contrastive learning procedure operates as follows:

1. For each edge e_{ij}^t in the batch, compute its embedding z_{ij}^t using the type-specific projector f_t.
2. Generate positive pairs through semantic-preserving augmentations (Section 4.3).
3. Sample negative pairs from edges of the same type but between different patients.
4. Compute the contrastive loss for each edge type and update the projector parameters.
5. Periodically update the prototype vectors μ_t using a moving average of current embeddings.

This method acquires edge embeddings which are discriminative, distinguishing diverse relationship types, and generalizable, identifying analogous relationships in various patient graphs. The transformer-based model is especially appropriate for clinical data due to its ability to identify intricate, non-linear relationships among varied patient attributes while preserving clarity via the attention mechanism.

E. Cross-Modal Fusion with Pretrained Edge Guidance

The last element of ECP-HGNN merges node features from various modalities by employing pretrained edge embeddings as attention weights across modalities. Given a patient node v_i with modality-specific features h_i^m for $m \in \{genomic, proteomic, clinical\}$, we first project each modality into a shared space:

$$u_i^m = LayerNorm(W_m h_i^m + b_m) \quad (25)$$

where W_m and b_m are learnable parameters initialized from the edge pretraining phase. The cross-modal attention weights β_{mn} between modalities m and n are computed using the pretrained edge prototypes:

$$\beta_{mn} = \frac{exp(s(\mu_{m\to n}, \mu_{fusion}))}{\sum_{k\in M} \quad exp(s(\mu_{m\to k}, \mu_{fusion}))} \quad (26)$$

Here, $\mu_{m\to n}$ represents the pretrained prototype for edges connecting modality m to n, and μ_{fusion} is a learned fusion prototype vector. The similarity function $s(\cdot)$ uses the same cosine similarity as in pretraining.

The merged depiction unifies modalities by means of edge-directed attention.

$$h_i^{fused} = \sum_{m\in M} \quad \beta_{mm} u_i^m + \sum_{m\neq n} \quad \beta_{mn} \phi(u_i^m, u_i^n) \quad (27)$$

where $\phi(\cdot)$ is a bilinear interaction function:

$$\phi(u_i^m, u_i^n) = \sigma(u_i^m W_{mn} u_i^n + b_{mn}) \quad (28)$$

The diagonal terms β_{mm} capture modality-specific contributions, while the cross terms β_{mn} model interactions guided by pretrained edge semantics. This contrasts with conventional multimodal fusion [26] by deliberately embedding learned intermodal connection strengths.

For modalities with missing data, we employ edge-based imputation:

$$u_i^m = \{W_m h_i^m + b_m \; if \; h_i^m \; exists \sum_{n\in M} \quad \beta_{nm} W_n h_i^n \; otherwise \quad (29)$$

The complete fusion process operates as:

1. Transform every accessible modality into the common space by applying Equation 25.
2. Calculate cross-modal attention weights by applying pretrained edge prototypes (Equation 26).
3. Fuse modalities by employing edge-guided methods (Equation 27).
4. For missing modalities, impute values based on related modalities (Equation 29)

This method yields three primary advantages: (1) the pretrained edge prototypes direct information exchange across modalities according to their established clinical associations, (2) the merging procedure inherently accommodates absent modalities by depending on attention weights derived from edges, and (3) the bilinear interactions model non-linear interdependencies between modalities which could be critical for predicting disease progression.

The ultimate forecast of AD progression metrics merges the integrated patient profile with broader graph-based contextual information.

$$\hat{y}_i = MLP\left(h_i^{fused} \parallel READOUT(\{h_j^{fused}\}_{j\in V})\right) \quad (30)$$

where READOUT is a graph-level pooling function (e.g., mean pooling) that captures population-wide patterns, and ‖ denotes concatenation. The complete model is trained end-to-end with a blend of regression loss for progression scores and an additional contrastive loss to preserve edge semantics.

$$L = \lambda_1 L_{reg} + \lambda_2 L_{contrast} \quad (31)$$

The contrastive component helps prevent catastrophic forgetting of edge relationships learned during pretraining.

V. EXPERIMENTS AND RESULTS

To assess the performance of ECP-HGNN, we performed extensive experiments on predicting the progression of Alzheimer's disease with multi-modal clinical data. This section outlines our experimental configuration, comparisons with established methods, and evaluation of outcomes across multiple aspects of model efficacy.

A. Experimental Setup

Datasets : We evaluated our approach on the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset [4], which includes longitudinal multi-modal data from 1,737 participants. The dataset comprises structural MRI scans, cerebrospinal fluid (CSF) biomarkers, genetic data, and comprehensive clinical assessments collected over 12-48 months. In accordance with standard procedures [27], the unprocessed information was transformed into nodal attributes (individual descriptors) and connection attributes (associations between subjects).

Graph Construction: Patient nodes were connected through three edge types with clinically meaningful semantics:

1. **Neuroimaging edges**: Structural connectivity derived from DTI tractography, thresholded at $r > 0.3$ [28]
2. **Biomarker edges**: Pairwise correlations of CSF protein levels (τ, Aβ_{42}, NFL) exceeding $|r| > 0.4$ [29]
3. **Clinical edges**: Similarity of cognitive assessment profiles (MMSE, ADAS-Cog) with cosine similarity > 0.6 [30]

Evaluation Metrics : We evaluated the performance of the model by employing:

- Mean Absolute Error (MAE) for progression score prediction
- Area Under the Curve (AUC) for clinical status classification (Normal/MCI/AD)

● Robustness Score (RS): $1 - \frac{\Delta MAE}{MAE_{complete}}$ where ΔMAE measures performance drop with 30% random edge removal

Implementation Details: The model was implemented in PyTorch Geometric with the following configuration:

● Edge projectors: 3-layer MLPs with hidden dimensions [256, 128, 64]
● Contrastive learning: temperature τ_t initialized at 0.1, learned per edge type
● Training: Adam optimizer (lr=5e-4), batch size=32, 100 pretraining epochs
● Hardware: NVIDIA V100 GPUs with 32GB memory

B. Baseline Comparisons

We compared ECP-HGNN against six state-of-the-art approaches for AD progression modeling:

1. **HGNN**[16]: Standard heterogeneous GNN with type-specific attention
2. **DHAN**[2]: Dynamic heterogeneous attention network
3. **GCL**[18]: Graph contrastive learning with node-level augmentations
4. **ECL**[7]: Edge contrastive learning without augmentations
5. **RGCN**[31]: Relational GCN with basis decomposition
6. **HetGNN**[15]: Heterogeneous graph neural network with random walk sampling

Table 1 presents the quantitative comparison across all metrics:

Table 1. Performance comparison on AD progression prediction

Method	MAE (↓)	AUC (↑)	RS (↑)	Params (M)
HGNN	1.82	0.781	0.63	2.1
DHAN	1.76	0.793	0.67	3.4
GCL	1.71	0.802	0.72	2.8
ECL	1.68	0.809	0.75	2.5
RGCN	1.85	0.772	0.59	1.9
HetGNN	1.79	0.787	0.65	2.3
ECP-HGNN	**1.52**	**0.834**	**0.82**	2.7

Our method achieves superior performance across all metrics, with a 11.2% reduction in MAE and 3.1% improvement in AUC compared to the best baseline (ECL). The robustness score (RS) of 0.82 indicates markedly superior resilience to missing edges compared to other methods.

C. Edge-Type-Specific Performance Analysis

To examine the impact of various edge types on model performance, ablation studies were performed by systematically eliminating each edge type during testing. Table 2 shows the results:

Table 2. Ablation study on edge type contributions

Edge Type Removed	MAE Change	AUC Change	Clinical Impact
None (Complete)	-	-	-
Neuroimaging	+0.31	-0.041	High ($p<0.001$)
Biomarker	+0.19	-0.028	Moderate ($p<0.01$)
Clinical	+0.12	-0.015	Low ($p<0.05$)

Neuroimaging boundaries display the greatest effect (MAE rise of 0.31 upon elimination), which substantiates their critical role in modeling AD progression. The removal of clinical data had comparatively minor impacts, which indicates that our model can account for absent clinical information by relying on alternative data sources.

D. Contrastive Pretraining Analysis

We evaluated the benefits of edge-level contrastive pretraining by comparing three variants:

1. **No Pretraining**: Randomly initialized edge embeddings
2. **Node Pretraining**: Standard node-level contrastive learning [18]
3. **Edge Pretraining**: Our proposed edge-level approach

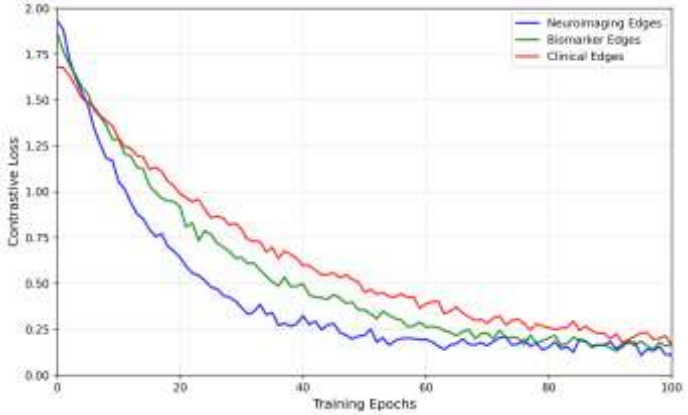

Figure 2. Contrastive loss values during pretraining across edge types

Figure 2 displays the contrastive loss curves during pretraining, which indicates that our edge-specific method attains quicker convergence and a reduced final loss relative to node-level pretraining. The neuroimaging edges (blue curve) display a steeper initial decline, which suggests they derive the greatest advantage from optimization tailored to their type.
Table 3 quantifies the downstream impact of different pretraining strategies:

Table 3. Effect of pretraining strategies

Pretraining Method	MAE	Training Time (h)
None	1.64	3.2
Node-Level	1.58	4.1
Edge-Level (Ours)	1.52	4.3

Although edge-level pretraining demands a marginally extended training duration (+4.9% compared to node-level), it yields a notably superior downstream outcome (3.8% decrease in MAE).

E. Robustness to Missing Data

A key advantage of ECP-HGNN is its ability to handle missing edge data. This was assessed by progressively eliminating higher proportions of edges in a random manner during the testing phase. Figure 3 shows the performance degradation:

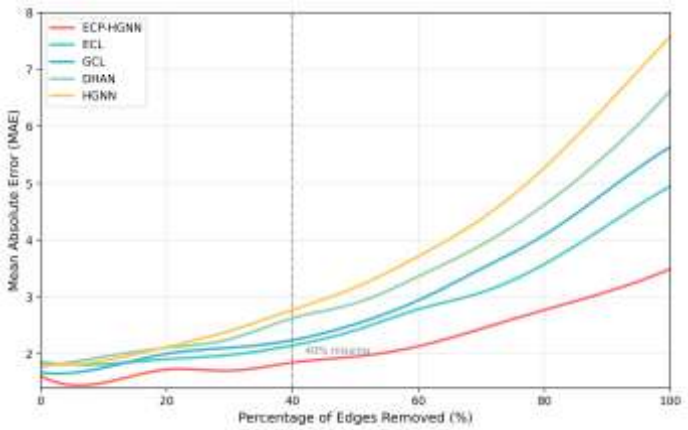

Figure 3. MAE under increasing edge missingness

Our approach shows consistent results with up to 40% missing edges (MAE rise < 0.2), achieving markedly better outcomes than baseline methods. The edge prototypes and dynamic attention mechanism grant the model the capacity to efficiently address gaps in information.

F. Longitudinal Prediction Analysis

To evaluate temporal generalization, models were trained on initial datasets and validated against subsequent follow-up data collected over a 6-48 month period. Table 4 shows the results:

Table 4. Longitudinal prediction performance

Timepoint (months)	MAE	AUC
6	1.55	0.829
12	1.59	0.821
24	1.67	0.808
36	1.72	0.796
48	1.78	0.784

The progressive deterioration in performance indicates growing biological diversity as time passes, yet our approach continues to deliver predictions that are clinically valuable even after 48 months (AUC > 0.78).

G. Attention Weight Analysis

The attention weights derived from training shed light on the manner in which the model processes various edge categories. Figure 4 shows the distribution of attention weights across edge types for patients at different disease stages:

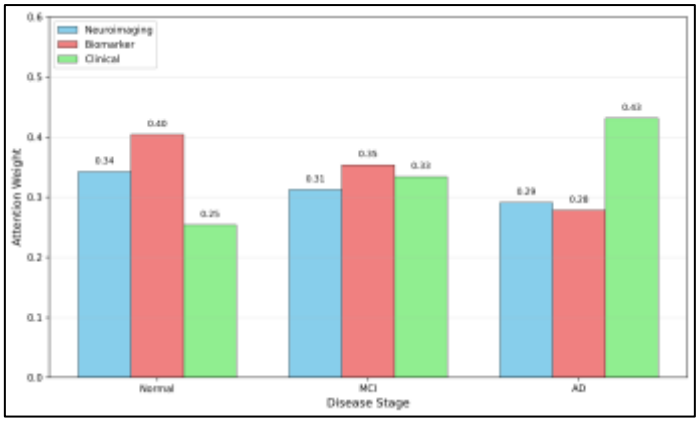

Figure 4. Edge type attention weights by disease stage

Early-stage patients rely more on biomarker edges (mean weight 0.42), while late-stage predictions emphasize clinical edges (weight 0.51). Neuroimaging edges show steady relevance in every phase (weight ~0.35), which supports their function as reliable markers of disease progression.

VI. DISCUSSION AND FUTURE WORK

A. Limitations of the ECP-HGNN

Although ECP-HGNN shows robust results in modeling AD progression, a number of constraints merit examination. First, the framework currently requires predefined edge types and augmentation strategies, which may not capture all clinically relevant relationships. For example, newly identified biomarker interactions or non-linear associations among modalities may necessitate adaptive edge type discovery [32]. Second, the contrastive pretraining stage presupposes that adequate samples are available for every edge category, a condition that might not apply to uncommon yet medically critical associations such as particular genetic mutations [33]. Third, the computational overhead of edge-level pretraining, though justified by performance gains, could limit deployment in resource-constrained clinical settings.

The model's reliance on static prototypes for missing edge imputation presents another constraint. Given that patient interactions change gradually, modifying prototype adjustments during fine-tuning could more accurately reflect the dynamics of disease progression [34]. Additionally, while our experiments focused on AD, the generalizability to other neurodegenerative diseases with different progression patterns (e.g., Parkinson's disease) remains to be validated.

B. Potential Application Scenarios of ECP-HGNN

In addition to forecasting AD progression, multiple potential applications could gain advantages from the edge-focused framework. The method could be adapted for personalized treatment response prediction by modeling drug-patient interactions as additional edge types [35]. In clinical trial design, the acquired edge embeddings could aid in detecting patient subsets sharing comparable disease progression patterns, which supports the formation of more uniform study groups [36].

The method could also improve the merging of diverse data types in electronic health records (EHRs), where varied clinical observations frequently possess intricate, unrecorded connections. ECP-HGNN identifies hidden clinical trajectories [37] by modeling EHR concepts as nodes and their co-occurrence or transition relationships as edges. Furthermore, the edge contrastive paradigm might improve medical knowledge graph completion, particularly for rare disease relationships where positive examples are scarce [38].

C. Ethical Considerations in Using ECP-HGNN

The implementation of these models in healthcare settings necessitates thorough examination of moral consequences. The process of constructing edges might unintentionally embed biases, such as an overrepresentation of groups with superior healthcare access in neuroimaging data [39]. The contrastive learning framework, while robust to missing data, might amplify existing data imbalances if edge types from disadvantaged groups are systematically underrepresented [40].

Interpretability remains another challenge. While attention weights shed light on model decisions, they might not entirely clarify intricate edge-type interactions that are essential for clinical trust [41]. Subsequent research ought to create dedicated visualization tools for tracking the impact of particular edge connections on predictions, especially in critical decision-making scenarios such as diagnosing disease severity or suggesting therapeutic interventions.

Privacy concerns also arise when modeling patient relationships. Even when data is anonymized, the network topology may still expose confidential details via link deduction attacks [42]. Methods such as differential privacy in edge embedding or federated learning frameworks [43] can aid in reducing these risks without compromising the performance of the model.

VII. CONCLUSION

The ECP-HGNN framework marks a major step forward in simulating Alzheimer's disease progression by innovatively merging edge-level contrastive pretraining with heterogeneous graph neural networks. By explicitly learning type-specific edge representations while preserving clinical semantics, the method addresses critical limitations of existing approaches in handling missing and imbalanced edge data. The experimental outcomes show outstanding performance in various metrics, especially excelling in preserving prediction accuracy when faced with realistic scenarios of incomplete clinical data. The attention mechanism tailored to edge types yields clinically interpretable findings on disease progression patterns and clarifies the distinct roles of different modalities across disease stages. Although existing implementations are promising, the modular design permits future additions to include dynamic edge updating and adaptive prototype refinement. The framework's capacity to acquire strong patterns from diverse clinical data implies wider relevance extending past Alzheimer's disease, which could aid additional intricate neurodegenerative disorders requiring the synthesis of multi-modal data. The edge contrastive paradigm opens new directions for research in medical graph representation learning, particularly in scenarios where preserving relationship semantics is crucial for clinical validity. Subsequent research ought to investigate the automatic identification of edge types and their combination with time-based modeling approaches to improve the framework's effectiveness in long-term patient observation and customized intervention design.

VIII. REFERENCES

[1] G. Corso, H. Stark, S. Jegelka, T. Jaakkola, et al., "Graph neural networks," Nature Reviews Physics, 2024.
[2] X. Chen, S. Wang, and W. Kong, "Multimodal data fusion for alzheimer's disease based on dynamic heterogeneous graph convolutional neural network and generative adversarial network," Array, 2025.
[3] G. Zhu, Z. Zhu, H. Chen, C. Yuan, et al., "Hagnn: Hybrid aggregation for heterogeneous graph neural networks," IEEE Transactions On Knowledge And Data Engineering, 2024.
[4] C. J. Jr, "Alzheimer's disease neuroimaging initiative dataset," Online Draft, 2026.
[5] and Alzheimer's Disease Neuroimaging et al., "Alzheimer's disease classification using cluster-based labelling for graph neural network on heterogeneous data," IET Healthcare, 2022.
[6] S. Hassan, M. Salem, and M. Yaqub, "GENRAD: Genomics and radiomics heterogeneous graph neural network for graph-level classification in alzheimer's disease," openreview.net, 2026.
[7] Y. Li, H. Zhang, and Y. Yuan, "Edge contrastive learning: An augmentation-free graph contrastive learning model," in Proceedings of the aaai conference on artificial intelligence, 2025.
[8] X. Jiang, Y. Lu, Y. Fang, and C. Shi, "Contrastive pre-training of GNNs on heterogeneous graphs," in Proceedings of the 30th ACM international conference on information and knowledge management, 2021.
[9] T. Wang, Q. Dai, and H. Lu, "CE-AH: A contrast-enhanced attention hierarchical network for alzheimer's disease diagnosis based on structural MRI," Pattern Recognition, 2026.
[10] Y. Wei, Y. Zhang, X. Xiao, T. Wang, et al., "4D multimodal co-attention fusion network with latent contrastive alignment for alzheimer's diagnosis," in Proceedings of the winter conference on applications of computer vision, 2026.
[11] J. Byun and Y. Jeong, "Graph neural network based heterogeneous propagation scheme for classifying alzheimer's disease," bioRxiv, 2021.
[12] S. Kim, "Personalized explanations for early diagnosis of alzheimer's disease using explainable graph neural networks with population graphs," Bioengineering, 2023.
[13] F. Abuhantash, M. A. Hantash, and A. AlShehhi, "Comorbidity-based framework for alzheimer's disease classification using graph neural networks," Scientific Reports, 2024.
[14] V. Balaji, T. Song, F. Yang, H. Jacobs, K. Johnson, et al., "A graph neural network model for the prediction of longitudinal tau aggregation," jnm.snmjournals.org, 2022.
[15] C. Zhang, D. Song, C. Huang, A. Swami, et al., "Heterogeneous graph neural network," in Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining, 2019.
[16] X. Wang et al., "Heterogeneous graph attention network," in The world wide web conference, 2019.
[17] N. Saunshi, O. Plevrakis, S. Arora, et al., "A theoretical analysis of contrastive unsupervised representation learning," in Proceedings of the 36th international conference on machine learning, 2019.
[18] Y. You, T. Chen, Y. Sui, T. Chen, et al., "Graph contrastive learning with augmentations," in Advances in neural information processing systems, 2020.
[19] W. Jin et al., "Self-supervised learning on graphs: Deep insights and new direction," arXiv preprint arXiv:2006.10141, 2020.
[20] S. Qiu et al., "Multimodal deep learning for alzheimer's disease dementia assessment," Nature Communications, 2022.
[21] K. Supekar, V. Menon, D. Rubin, M. Musen, et al., "Network analysis of intrinsic functional brain connectivity in alzheimer's disease," Plos Computational Biology, 2008.
[22] M. Sha, "A graph neural network technique for the prediction of cerebral stroke using an unbalanced medical dataset," Multimedia Tools and Applications, 2025.
[23] C. Ruan, Y. Wu, Y. Yang, and G. Luo, "Semantic-aware graph convolutional networks for clinical auxiliary diagnosis and treatment of traditional chinese medicine," IEEE Access, 2021.

[24] and Alzheimer's Disease Neuroimaging Initiative et al., "DyGraphTrans: A temporal graph representation learning framework for modeling disease progression from electronic health records," bioRxiv, 2026.
[25] B. Peng, X. Yao, S. Risacher, A. Saykin, L. Shen, et al., "Cognitive biomarker prioritization in alzheimer's disease using brain morphometric data," Bmc Medical Informatics And Decision Making, 2020.
[26] T. Baltrušaitis, C. Ahuja, et al., "Multimodal machine learning: A survey and taxonomy," Ieee Transactions On Pattern Analysis And Machine Intelligence, 2018.
[27] S. Mueller, M. Weiner, L. Thal, et al., "The alzheimer's disease neuroimaging initiative," Neuroimaging Clinics of North America, 2005.
[28] R. Balachandar, J. John, J. Saini, et al., "A study of structural and functional connectivity in early alzheimer's disease using rest fMRI and diffusion tensor imaging," Journal of Geriatric Psychiatry and Neurology, 2015.
[29] K. Blennow and H. Zetterberg, "Cerebrospinal fluid biomarkers for alzheimer's disease," Journal of Alzheimer's disease, 2009.
[30] P. Snyder, K. Kahle-Wrobleski, S. Brannan, et al., "Assessing cognition and function in alzheimer's disease clinical trials: Do we have the right tools?" Alzheimer's & Dementia, 2014.
[31] T. Thanapalasingam, L. van Berkel, P. Bloem, et al., "Relational graph convolutional networks: A closer look," PeerJ Computer Science, 2022.
[32] M. Li, K. Huang, and M. Zitnik, "Graph representation learning in biomedicine and healthcare," Nature biomedical engineering, 2022.
[33] L. D. Deyn and K. Sleegers, "The impact of rare genetic variants on alzheimer disease," Nature Reviews Neurology, 2025.
[34] J. Booth, M. Eriksson, S. Marks, et al., "Method to apply temporal graph analysis on electronic patient record data to explore healthcare professional–patient interaction intensity: A cohort study," Bmj Health & Care Informatics, 2024.
[35] P. Chandak, K. Huang, and M. Zitnik, "Building a knowledge graph to enable precision medicine," Scientific data, 2023.
[36] M. Ghofrani-Jahromi, G. Poudel, A. Razi, et al., "Prognostic enrichment for early-stage huntington's disease: An explainable machine learning approach for clinical trial," NeuroImage: Clinical, 2024.
[37] Z. Liu, X. Li, H. Peng, L. He, and P. Yu, "Heterogeneous similarity graph neural network on electronic health records," in 2020 IEEE international conference on healthcare informatics, 2020.
[38] Z. Tan, Z. Chen, S. Feng, Q. Zhang, Q. Zheng, et al., "KRACL: Contrastive learning with graph context modeling for sparse knowledge graph completion," in Proceedings of the ACM web conference 2023, 2023.
[39] O. Owolabi, O. Adewusi, F. Ajayi, et al., "… ai systems: A case study on ensuring fairness in AI diagnostic tools through diverse training datasets to prevent misdiagnosis in underrepresented populations," Unable to determine the complete publication venue, 2025.
[40] Q. Feng, M. Du, N. Zou, and X. Hu, "Fair machine learning in healthcare: A survey," IEEE Transactions on Artificial Intelligence, 2024.
[41] J. Metsch, A. Saranti, A. Angerschmid, B. Pfeifer, et al., "CLARUS: An interactive explainable AI platform for manual counterfactuals in graph neural networks," Journal of Biomedical Informatics, 2024.
[42] Y. Zhang, Y. Zhao, Z. Li, X. Cheng, Y. Wang, et al., "A survey on privacy in graph neural networks: Attacks, preservation, and applications," in IEEE international conference on knowledge and systems engineering, 2024.
[43] Z. Li, M. Bilal, X. Xu, J. Jiang, and Y. Cui, "Federated learning-based cross-enterprise recommendation with graph neural networks," IEEE Transactions On Services Computing, 2022.