

# 图基础模型：迈向通用图智能的探索与进展

## 摘要

近年来，基础模型在自然语言处理和计算机视觉领域取得了革命性进展，展现出强大的跨任务、跨领域泛化能力。受此启发，研究界开始积极探索图基础模型（Graph Foundation Models, GFM）的可能性，旨在构建能够处理多样化图数据、适应不同图学习任务的通用模型。然而，图数据的内在复杂性——包括结构异质性、节点/边特征的多样性、任务范式的差异性以及缺乏统一的“图词汇”——为 GFM 的构建带来了巨大挑战。本文聚焦于近期 GFM 领域的五项代表性工作（OFA, UniGraph, GOFA, AnyGraph, GFT），系统梳理它们在应对上述挑战、推动图学习通用化方面所做的核心贡献与创新思路。本文首先概述了构建 GFM 面临的关键挑战，随后深入剖析了这五项工作在统一特征表示、任务范式、整合结构与语义、定义可迁移模式以及实现模型自适应性等方面的具体方法与技术路径。通过对比分析，本文揭示了当前 GFM 研究的不同技术流派及其优劣，并指出了该领域存在的局限与未来发展方向，旨在为后续研究提供有益的参考与启示。

## 1. 引言

基础模型（Foundation Models）[7]通过在超大规模数据集上进行预训练，学习通用的、可迁移的知识表示，并通过微调或提示学习等方式适应广泛的下游任务，已成为人工智能领域的研究热点[4, 5, 48, 50, 112]。自然语言处理领域的 LLMs（如 GPT 系列 [1]）和计算机视觉领域的 LVMs（如 Sora[46], CLIP[40]）的成功，极大地激发了将基础模型理念拓展至图学习领域的兴趣。图结构数据无处不在，从社交网络、生物分子结构到知识图谱、推荐系统，开发能够理解和推理多样化图数据的 GFM，有望解锁前所未有的应用潜力[50, 93]。

然而，与文本和图像数据相比，图数据具有显著的异质性和复杂性。**首先，特征异质性（Feature Heterogeneity）**：不同来源的图（如引文网络、分子图、知识图谱）其节点和边可能拥有截然不同的特征空间（数值、类别、文本、图像等多模态信息）和维度[45, 25]，难以直接使用单一模型进行统一表征。**其次，结构异质性（Structure Heterogeneity）**：图的拓扑结构千差万别，从同配性到异配性，从稀疏到稠密，从简单图到异质图、动态图，对模型的结构感知能力提出了极高要求[49, 104]。**再次，任务范式多样性（Task Diversity）**：图学习任务涵盖节点级别（分类、回归）、边级别（链接预测、关系预测）和图级别（图分类、图生成）等多种范式，需要模型具备灵活的任务处理能力[35, 68]。**最后，缺乏通用图词汇（Lack of Universal Graph Vocabulary）**：与文本的词语或图像的视觉令牌（visual tokens）不同，图领域长期以来缺乏公认的、能够跨领域、跨任务迁移的基础模式或“词汇”[50, 93]。这些挑战共同构成了构建通用、高效、可扩展的 GFM 的主要障碍。

面对这些挑战，近期涌现了一系列开创性的研究工作。本文选取了五篇具有代表性的论文——OFA [35], UniGraph [20], GOFA [Kong et al., Preprint], AnyGraph [Xia & Huang,

Preprint], GFT [Wang et al., NeurIPS 2024]——进行综述。这些工作从不同角度切入,探索了 GFM 的构建路径,分别在统一特征与任务、融合语言模型与图神经网络、利用混合专家架构处理异质性、以及定义可迁移图模式等方面做出了重要贡献。

## 2. 核心方法与贡献

### 2.1 基于文本属性图与提示学习的统一 (OFA, UniGraph)

为了解决特征异质性问题, **OFA (One-For-All)** [35] 创造性地提出将所有图数据视为**文本属性图 (Text-Attributed Graphs, TAGs)**。它利用大型语言模型 (LLM) 将不同来源图的节点和边特征 (即使是原始非文本特征) 统一描述为自然语言文本, 并将其编码到同一语义嵌入空间。这一思想极大地简化了跨领域图数据的表示统一问题。在任务统一方面, OFA 聚焦于分类任务, 引入了**兴趣节点 (Nodes-of-Interest, NOI)** 的概念来标准化不同层级 (节点、边、图) 的分类目标, 并通过精心设计的图提示范式 (**Graph Prompting Paradigm, GPP**), 将任务信息 (如类别描述) 有机地注入图结构中, 通过在提示节点上进行预测, 实现了单一模型对多种分类任务的处理, 并具备了一定的零样本学习能力。

**UniGraph** [20] 同样基于 TAGs, 但其目标更为宏大, 旨在构建一个能够进行**自监督预训练**的跨领域 GFM。它采用**级联的 LM+GNN 架构**进行端到端的表示学习, 提出了基于掩码图建模 (Masked Graph Modeling) 的自监督预训练目标, 旨在从大规模无标签 TAGs 中学习通用的图表示。与 OFA 依赖有监督联合训练不同, UniGraph 侧重于预训练-微调/提示的范式。为了实现零样本预测, UniGraph 进一步引入了**图指令微调 (Graph Instruction Tuning)**, 利用 LLM 理解自然语言指令, 将预训练的图嵌入与任务指令相结合, 从而统一不同图的标签空间, 实现对未见图和任务的零样本泛化。

### 2.2 融合 GNN 与 LLM 的生成式建模 (GOFA)

在 OFA 和 UniGraph 探索如何利用 LLM 辅助 GNN 处理图数据 (LLM as enhancer) 的同时, **GOFA (Generative One-For-All)** [Kong et al., Preprint] 则探索了更深层次的融合, 旨在构建一个生成式 (Generative) 的 GFM, 赋予模型类似 LLM 的**任务流动性 (Fluidity in tasks)** 和**图理解能力 (Graph understanding)**。GOFA 的关键创新在于其独特的架构设计: 它将随机初始化的 GNN 层交错 (interleave) 插入到冻结的预训练 LLM (如 ICAE[14]的压缩器和解码器) 中。这种设计使得 LLM 在保持其强大的文本生成和语义理解能力的同时, 能够通过 GNN 层感知和处理图的结构信息。GOFA 提出了图层面的**下一词预测 (next-word prediction)**、**问答链 (QA-Chain)** 和**结构推理 (structural tasks)** 等预训练任务, 旨在让模型联合学习图的语义和结构知识。通过这种方式, GOFA 不仅能处理分类、回归等判别式任务, 还能执行自由形式的文本生成和问答任务, 显著提升了模型的通用性和灵活性。

### 2.3 基于混合专家架构的异质性处理与自适应 (AnyGraph)

真实世界的图数据不仅跨领域异质, 领域内部也常常存在显著的分布漂移

(distribution shift)。AnyGraph [Xia & Huang, Preprint] 认为单一的“庞大”模型可能难以有效应对如此复杂的异质性，并可能导致模型在适应新数据时产生干扰或灾难性遗忘。为此，AnyGraph 引入了**图混合专家（Graph Mixture-of-Experts, MoE）架构**。该架构包含多个（轻量级的）图专家模型，每个专家负责处理具有特定特征或结构属性的图数据子集。AnyGraph 设计了一个基于图自监督任务损失的**自动化图专家路由（Graph Expert Routing）机制**，能够根据输入图的特性，动态地选择最“胜任”的专家进行处理。这种 MoE 设计不仅能有效处理结构和特征层面的异质性，还能通过仅激活和更新部分专家网络，实现对新数据和新领域的**快速自适应（Fast Adaptation）**。此外，AnyGraph 还实验验证了其模型的**规模法则（Scaling Law）**，表明增加模型（专家数量或大小）和数据量能够持续提升模型的泛化能力和涌现能力。

## 2.4 基于计算树词汇的可迁移模式学习（GFT）

探索图的“通用词汇”是构建 GFM 的根本性问题。**GFT (Graph Foundation model with transferable Tree vocabulary)** [Wang et al., NeurIPS 2024] 提出了一种全新的视角：将图的可迁移模式重新定义为**计算树（Computation Trees）**——即从 GNN 消息传递过程中展开得到的树状结构。GFT 认为计算树能够有效捕获图的局部结构信息，并且这种结构模式在不同图任务和领域中具有一定的可迁移性。基于此，GFT 提出在预训练阶段执行**计算树重构（Computation Tree Reconstruction）任务**，包括重构根节点特征、节点连接性以及树的整体语义。更重要的是，GFT 利用向量量化（Vector Quantization, VQ）技术，将连续的计算树嵌入空间量化为离散的树词汇表（Tree Vocabulary）。这个词汇表封装了跨领域、跨任务的基本图模式。在微调阶段，GFT 将所有下游图任务（节点、边、图级别）统一\*\*重构为计算树分类（Computation Tree Classification）\*\*任务，利用学习到的树词汇进行预测。这种基于计算树词汇的方法不仅为定义图的可迁移模式提供了新思路，还有效地缓解了负迁移风险，提升了模型的泛化性能。

## 3. 对比分析与讨论

这五项工作代表了当前 GFM 研究的不同技术路径和侧重点：

- **特征与任务统一策略**：OFA 和 UniGraph 都利用 LLM 处理文本属性图来统一特征，但在任务统一上有所不同。OFA 通过 GPP 和 NOI 统一分类任务，而 UniGraph 通过指令微调实现更广泛的零样本任务泛化。GOFA 则通过生成式建模，天然地具备处理自由形式任务的能力。GFT 通过计算树分类统一了所有任务范式。AnyGraph 则通过 MoE 架构隐式地处理任务多样性。
- **LLM 与 GNN 的融合方式**：OFA 主要将 LLM 作为特征编码器。UniGraph 采用级联架构。GOFA 将 GNN 层嵌入 LLM 内部，实现了更深度的融合。AnyGraph 则将 GNN 作为 MoE 中的专家基础。GFT 主要在预训练阶段利用 GNN 编码计算树，LLM 的作用相对较小。

- **预训练范式：**UniGraph 采用了自监督的掩码图建模。GOFA 提出了图层面的语言建模和结构推理任务。GFT 设计了多方面的计算树重构任务。OFA 则依赖于多任务的有监督联合训练。AnyGraph 的预训练与其 MoE 路由机制紧密相关，基于自监督损失进行专家选择和训练。
- **核心创新点：**OFA 的创新在于 GPP 和 NOI。UniGraph 在于其端到端的 TAG 自监督学习和指令微调。GOFA 在于其生成式图语言建模和 GNN/LLM 交错架构。AnyGraph 在于图 MoE 架构和路由机制。GFT 在于将计算树定义为可迁移词汇并通过 VQ 学习。

这些方法各有优劣。基于 LLM 嵌入和提示的方法（OFA）相对简单，易于实现，但在任务类型和深度图结构理解上可能受限。生成式模型（GOFA）潜力巨大，任务灵活，但训练复杂度和稳定性要求高。MoE 架构（AnyGraph）在处理异质性和自适应性上表现突出，但可能增加模型复杂度和路由开销。基于计算树词汇的方法（GFT）从更根本的层面思考图模式的可迁移性，但其效果依赖于计算树假设的有效性和 VQ 的质量。

#### 4. 局限性与未来展望

尽管上述工作取得了显著进展，但 GFM 的研究仍处于早期阶段，面临诸多挑战与机遇：

- **更深度的结构与语义融合：**如何设计更优的架构（如 GOFA 的探索）来无缝融合 GNN 的结构推理能力和 LLM 的语义理解与生成能力，仍是核心问题。
- **更鲁棒的预训练范式：**需要探索更通用、更有效的图自监督预训练目标，使其能够捕获跨越不同尺度、不同领域的共性知识。计算树重构（GFT）和图语言建模（GOFA）是值得深入的方向。
- **探索更丰富的图词汇：**计算树（GFT）提供了一种可能的图词汇。未来可以探索其他类型的结构基元（如 motif、graphlet）或结合多尺度信息来构建更丰富的图词汇表。
- **动态与时序图基础模型：**当前工作主要集中在静态图上，如何构建能够处理动态、时序演化图数据的 GFM 是一个重要的未来方向。
- **可扩展性与效率：**训练 GFM 需要处理大规模、多样化的图数据，对计算资源和算法效率提出了极高要求。MoE 架构（AnyGraph）提供了一种潜在的解决方案，但仍需优化。
- **可解释性与可信赖性：**随着模型能力的增强，理解 GFM 的决策过程、确保其公平性和鲁棒性变得愈发重要。

#### 5. 结论

图基础模型是图学习领域一个充满挑战和机遇的前沿方向。本文综述了 OFA,

UniGraph, GOFA, AnyGraph, GFT 等五项近期代表性工作，它们分别从统一特征表示、任务范式、融合架构、异质性处理和可迁移模式定义等角度，为构建通用图智能模型进行了富有成效的探索。这些工作展示了利用大型语言模型、自监督学习、混合专家架构、计算树等先进技术构建 GFM 的巨大潜力。尽管挑战依然存在，但这些开创性的研究无疑为实现真正意义上的图基础模型铺平了道路，预示着图学习领域即将迎来新的突破。未来的研究需要在结构与语义的深度融合、预训练范式的创新、图词汇的探索、动态图处理以及模型效率与可信赖性等方面持续深耕。