



面向人工智能领域的组合泛化方法研究综述

李昌原, 程恺, 郝文宁, 邵天浩, 张金鹏

引用本文:

李昌原, 程恺, 郝文宁, 等. 面向人工智能领域的组合泛化方法研究综述[J]. 智能系统学报, 2026, 21(4): 815–833.
LI Changyuan, CHENG Kai, HAO Wenning, et al. Comprehensive survey on compositional generalization methodologies in artificial intelligence[J]. *CAAI Transactions on Intelligent Systems*, 2026, 21(4): 815–833.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202507001>

您可能感兴趣的其他文章

一种卷积神经网络集成的多样性度量方法

Diversity measuring method of a convolutional neural network ensemble

智能系统学报. 2021, 16(6): 1030–1038 <https://dx.doi.org/10.11992/tis.202011023>

多智能体分层强化学习综述

A survey on multi-agent hierarchical reinforcement learning

智能系统学报. 2020, 15(4): 646–655 <https://dx.doi.org/10.11992/tis.201909027>

人工智能中的封闭性和强封闭性——现有成果的能力边界、应用条件和伦理风险

Criteria of closeness and strong closeness in artificial intelligence——limits, application conditions and ethical risks of existing technologies

智能系统学报. 2020, 15(1): 114–120 <https://dx.doi.org/10.11992/tis.202001001>

图像情境下的数字序列逻辑学习

Number sequence logic learning in image context

智能系统学报. 2019, 14(6): 1189–1198 <https://dx.doi.org/10.11992/tis.201905044>

弹性网络核极限学习机的多标记学习算法

Multi-label learning algorithm of an elastic net kernel extreme learning machine

智能系统学报. 2019, 14(4): 831–842 <https://dx.doi.org/10.11992/tis.201806005>

多标记学习自编码网络无监督维数约简

Unsupervised dimensionality reduction of multi-label learning via autoencoder networks

智能系统学报. 2018, 13(5): 808–817 <https://dx.doi.org/10.11992/tis.201804051>

DOI: 10.11992/tis.202507001

网络出版地址: <https://link.cnki.net/urlid/23.1538.tp.20260622.1812.002>

面向人工智能领域的组合泛化方法研究综述

李昌原, 程恺, 郝文宁, 邵天浩, 张金鹏

(陆军工程大学 指挥控制工程学院, 江苏 南京 210007)

摘要: 组合性是现代人工智能系统所缺失的、人类智能的一个重要方面, 能够增强系统泛化和解决新颖且复杂任务的能力。针对组合泛化概念内涵模糊化、方法分类碎片化、机理阐释表层化等问题, 明确了组合泛化的基本概念, 从系统性、可组合性、可解释性和动态适应性等 4 个方面分析了其特征, 提出了基于架构演进的组合泛化方法分类体系, 按照基础表征架构、关系建模架构与学习机制架构对现有组合方法进行了分类, 详细阐述了每种方法的模型结构与组合泛化之间的机理, 进一步比较了不同方法之间的优缺点及场景适用性。针对组合泛化面临的动态性、知识依赖、计算复杂、异质关系等问题, 提出了可行的解决思路, 为该领域的深入研究提供了指导。

关键词: 组合性; 组合泛化; 人工智能; 基础表征架构; 关系建模架构; 学习机制架构; 机器学习; 元学习

中图分类号: TP183; TP181; TP391.1 **文献标志码:** A **文章编号:** 1673-4785(2026)04-0815-19

中文引用格式: 李昌原, 程恺, 郝文宁, 等. 面向人工智能领域的组合泛化方法研究综述 [J]. 智能系统学报, 2026, 21(4): 815-833.

英文引用格式: LI Changyuan, CHENG Kai, HAO Wenning, et al. Comprehensive survey on compositional generalization methodologies in artificial intelligence[J]. CAAI transactions on intelligent systems, 2026, 21(4): 815-833.

Comprehensive survey on compositional generalization methodologies in artificial intelligence

LI Changyuan, CHENG Kai, HAO Wenning, SHAO Tianhao, ZHANG Jinpeng

(College of Command and Control Engineering, Army Engineering University of PLA, Nanjing 210007, China)

Abstract: Compositionality, a critical aspect of human intelligence, is often lacking in modern artificial intelligence systems. Its integration can considerably enhance a system's generalization capabilities and its ability to solve novel and complex tasks. This study addresses the current challenges of a vague understanding of compositional generalization, fragmented method classification, and superficial explanations of underlying mechanisms. The study begins by clarifying the fundamental concept of compositional generalization, analyzing its characteristics across four dimensions: systematicity, compositionality, interpretability, and dynamic adaptability. A classification system for compositional generalization methods based on architectural evolution is then proposed, in which existing compositional methods are categorized according to their foundational representation architecture, relational modeling architecture, and learning mechanism architecture. The mechanism linking each method's model structure to compositional generalization is elaborated in detail. The advantages and disadvantages, as well as the scene applicability of different methods, are further compared across various scenarios. Finally, feasible solutions are proposed to address persistent challenges in compositional generalization, such as dynamics, knowledge dependence, computational complexity, and heterogeneous relationships, thereby providing guidance for in-depth research in this field.

Keywords: compositionality; compositional generalization; artificial intelligence; foundational representation architecture; relational modeling architecture; learning mechanism architecture; machine learning; meta-learning

随着新一轮科技革命和产业变革的深入推进, 人工智能作为引领未来发展的战略性技术,

正以前所未有的速度重塑科技创新格局。2022 年 11 月, OpenAI 发布的 ChatGPT 标志着生成式人工智能技术取得重大突破^[1], 该技术通过构建大规模预训练语言模型, 并创新性地引入基于人类反馈的强化学习 (reinforcement learning with hu-

收稿日期: 2025-07-01. 网络出版日期: 2026-06-23.

基金项目: 国家自然科学基金项目 (61806221).

通信作者: 程恺. E-mail: chengkai911@126.com.

©《智能系统学报》编辑部版权所有

man feedback, RLHF)^[2]机制,有效解决了人工智能生成模型的核心技术难题。这一突破性进展不仅推动了人工智能从专用型(narrow AI)向通用型(artificial general intelligence, AGI)的范式转变^[3],更引发了学术界和产业界对通用人工智能发展路径的深入思考。

人类智力最显著的特征之一是能够从简单元素组合出越来越复杂的概念,从而实现快速学习和应用知识。当前的人工智能系统在这方面还存在不足,通常无法解决超出其训练分布的任务^[4]。在这一背景下,组合泛化(compositional generalization)作为人工智能领域的核心科学问题地位日益凸显,其重要性体现在以下方面:

1) 迈向通用人工智能(AGI)的核心阶梯^[5]

当前许多人工智能系统虽然在特定任务上表现优异,但它们往往依赖于模式匹配和统计关联,在遇到全新组合时可能表现不佳。组合泛化能力则要求模型理解各个组成部分的语义和功能并能够进行逻辑组合,更接近人类的智能方式。因此,提升组合泛化能力是克服当前人工智能局限、迈向能够举一反三的通用人工智能的核心阶梯。

2) 突破当前人工智能数据效率瓶颈和泛化能力限制的关键

要让模型可靠地泛化到所有可能的组合,所需的训练数据量可能与词汇表大小成二次方甚至更高比例增长,单纯增加数据量成本高昂且不可持续。缺乏组合泛化能力意味着模型的表现上限很大程度上受限于训练数据中出现的组合。

3) 实现人工智能可靠性和安全性的关键

现实世界是无限复杂的,充满了稀有和未知情况的组合。一个不具备组合泛化能力的人工智能系统无法安全地部署在开放世界中,因为它无法可靠地处理未知情况。例如,一个在训练集上表现完美的自动驾驶系统,可能会因为遇到一个从未见过的“下雨天+太阳强光”的组合场景而完全失效。

自Fodor与Pylyshyn提出“系统性”概念^[6]以来,组合泛化的研究已从认知科学延伸至深度学习^[7]、自然语言处理^[8]、医疗影像分析^[9]等多个领域,形成了跨学科的探索路径^[10]。本文旨在对现有研究成果进行全面回顾和深入分析,系统梳理该领域的研究脉络,采用基于架构演进的思想对现有方法进行体系化分类,深刻阐释不同模型结构与组合泛化之间的机理,推动组合泛化在人工智能领域的发展和应用。本文的创新点主要体现在:1)针对组合泛化概念内涵不清、方法分类分散

和机理分析不足等问题,明确了组合泛化的基本概念及其系统性、可组合性、可解释性和动态适应性特征;2)提出了基于架构演进的组合泛化方法分类体系,从基础表征架构、关系建模架构和学习机制架构3个层面系统梳理典型方法;3)进一步分析了不同方法的组合泛化机理、适用场景及局限性,并对未来提升人工智能系统组合泛化能力的研究方向进行了展望。

1 组合泛化概述

1.1 组合泛化概念

在深入探讨组合泛化之前,有必要首先明确泛化这一核心概念。在人工智能领域,泛化能力(generalization ability)是指机器学习模型通过训练数据集学习内在规律,并能够将该规律有效迁移至未见数据,从而实现准确预测的性能指标。组合泛化作为分布外泛化(out-of-distribution generalization)的重要子类^[11],其概念最初源于语言学领域,随后被引入机器学习研究范畴。组合泛化特指智能系统能够基于已知的组件(如词汇、规则或概念)推理出新组合的语义^[12]或行为模式,进而实现对未知任务的系统性泛化。

组合泛化的理论基础建立在复合性原则之上,该原则强调通过3个关键机制实现模型的组合泛化能力:1)基本元素的识别与表征;2)组合规则的归纳与学习;3)新组合的泛化与应用。为便于理解,我们可以将组合泛化类比为积木构建过程:每个积木单元对应于基本元素,掌握单元间的组合方式即相当于理解组合规则,而面对新型结构时,能够基于已有知识体系推导出合适的组合方案。

组合泛化的内涵,强调其作为实现通用人工智能的关键路径:它使模型超越数据驱动的浅层泛化,进入基于规则和结构的深度泛化,从而处理开放世界的不确定性问题。组合泛化的外延涵盖多个人工智能子领域和具体任务,其范围从基础研究扩展到实际应用,体现了从“内涵属性”到“具体对象”的映射。组合泛化的应用场景覆盖自然语言处理、计算机视觉、数学推理、机器人控制等多个领域,体现为从翻译、生成到规划等具体任务^[13]。

1.2 组合泛化特征

1.2.1 系统性(systematicity)

系统性是语言组合性的核心特征,其核心机制在于复杂表达式的语义由其组成成分的语义及其句法结构通过系统化、规则化的方式共同决

定。这种系统性使得语言系统具备了计算上的可递归性和可生成性:通过有限的词汇集和语法规则的计算组合,可以生成并理解无限数量的合法句子,包括从未在训练数据中出现过的句子实例。这种特性与图灵完备性具有深层的理论联系,为自然语言的形式化建模提供了计算基础。

1.2.2 可组合性 (compositionality)

可组合性作为人类语言认知与人工智能系统实现创造性表达的核心机制,构建了从离散元素到复杂语义的生成范式。该机制通过有限的结构单元和运算规则实现无限组合可能,其本质在于建立形式化结构单元与语义空间的系统性映射关系。例如,在自然语言处理中,这一机制体现为基于递归组合的生成-解析双向过程,通过有限的语法结构和语义约束,实现无限的语言生成与理解能力。在认知层面,人类借助符号操作与语义框架实现递归组合,而人工智能模型则通过符号主义^[14]、连接主义^[15]及神经符号融合^[16]等路径模拟这一过程。

1.2.3 可解释性 (interpretability)

在组合泛化中,模型的可解释性需满足透明性与因果可追溯性的双重约束,其核心目标是确保模型在组合新概念或结构时,其推理逻辑既能被人类直观理解,又能通过因果机制验证逻辑的稳健性。透明性要求模型通过显式结构对组合逻辑进行形式化表征。因果可追溯性则需定位输出

结果与输入元素间的因果链条,支持反事实干预以验证逻辑稳健性,并能在错误发生时精准定位失效环节。透明性与因果可追溯性的协同机制为组合泛化模型提供了可审计性基础。

1.2.4 动态适应性 (dynamic adaptability)

组合泛化中的动态适应性是人工智能领域应对未知组合推理问题的核心研究方向,其目标是通过模型内部机制的实时调整,实现对新组合结构的快速泛化。该框架将组合泛化的本质挑战,即基于有限原子组件生成或理解无限组合模式,与动态计算范式相结合,通过元学习、模块化条件计算^[17]和强化策略组合等技术路径,使模型能够根据上下文动态重构推理逻辑。研究表明,动态适应性机制在自然语言处理^[18]、机器人任务规划^[19]等场景中显著提升了模型的泛化性能。

2 组合泛化方法分类

组合泛化作为实现通用人工智能的核心能力之一^[3],在学术界存在多种实现方式的分类标准。现有分类标准主要依据实现方法的特性或应用场景的差异,局限性在于缺乏对模型本质结构的深入分析,难以揭示模型结构设计对泛化的本质影响^[20]。组合泛化的核心挑战在于模型架构能否对由已知元素构成的新组合进行有效推理,因而本文基于架构演进的思想对现有组合泛化方法进行分类,如图1所示。

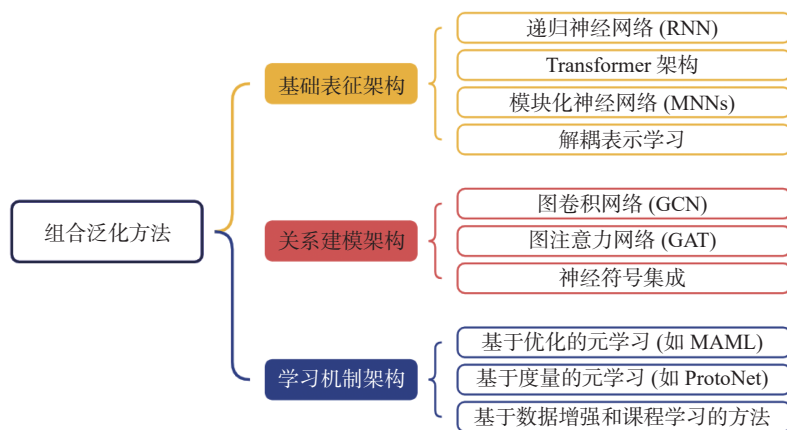


图1 组合泛化方法分类体系

Fig. 1 Taxonomy of compositional generalization methods

组合泛化的底层认知机制在根本上由模型的结构特性所决定,因而本文根据模型架构的差异将现有组合泛化方法分为3类:基础表征架构、关系建模架构和学习机制架构,体现了组合泛化领域的递进研究脉络。

2.1 基础表征架构

基础表征架构是人工智能系统的核心框架,

旨在将多源异构数据转化为结构化、可计算的高阶语义表征,并通过分层设计与多模态融合实现高效推理与应用,支持组合推理的泛化能力。基于基础表征架构的方法,可根据结构特征和实现机理的差异划分为4大技术分支:基于序列建模的递归神经网络、基于注意力机制的Transformer架构、具有组合式结构的模块化神经网络,以及

旨在分离数据中隐含的、可解释的生成因子的解耦表示学习。

2.1.1 递归神经网络

递归神经网络(recurrent neural network, RNN)通过其动态时序计算与参数共享机制^[21],显著提升了组合泛化能力。RNN的核心机制是递归连接,这种递归连接可以保留上一时间步的信息,使信息能够在不同时间步之间传递。如图2所示,左侧为递归神经网络的原始结构。在闭环之外,从下到上就是“输入层→隐藏层→输出层”的三层结构, x 、 h 、 o 分别表示输入、隐藏状态、输出, W_h 和 W_x 是序列中共享的权重矩阵。右侧为递归展开之后的结构。

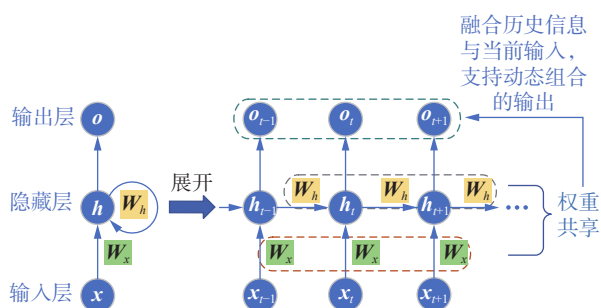


图 2 递归神经网络

Fig. 2 Hierarchical recursive architecture

RNN 引入了时间维度的循环连接,使隐藏状态 h_t 同时依赖当前输入 x_t 和上一时刻的隐藏状态 h_{t-1} ,其数学表达为

$$h_t = \sigma(W_h h_{t-1} + W_x x_t + b)$$

式中: σ 为激活函数(如 $\tanh$), o_t 为时刻 t 的输出。参数 W_h 、 W_x 、 b 分别表示隐藏层到隐藏层的循环权重矩阵、输入层到隐藏层的权重矩阵、隐藏层的偏置向量,3个参数在时间步间共享,实现序列数据的动态建模。权重共享和递归处理是RNN的两个重要机制,权重共享的本质是用有限的参数表达指数级的组合可能性,这正是人类语言泛化能力的核心特征之一;递归处理确保信息能够在不同时间步之间传递。RNN通过权重共享和递归处理,将组合规则编码为共享的权重参数,从而支持层级化结构的泛化学习,而非仅记忆表层模式。在对话系统中,RNN通过语义槽填充,进行上下文感知的组合理解,并通过预训练词向量的迁移,适应新领域组合^[22]。Tu等^[23]通过理论分析RNN的泛化边界并提出正则化方法,发现了RNN在处理复杂时间序列任务时的分解与重组能力,从而为增强模型的组合泛化能力提供了理论支持和实践指导。虽然RNN通过一系列

机制为组合泛化提供了基础支持,尤其在需要逐步构建层次化表示的任务中表现突出,但是在面对长序列问题时,长序列会导致梯度指数级衰减(消失)或增长(爆炸),难以学习长期依赖。针对此问题,Hochreiter & Schmidhuber 提出的LSTM(long short-term memory)通过门控单元控制信息流,显著缓解了梯度问题^[24-25]。例如,在机器翻译任务中,LSTM可跨越1000+时间步的依赖关系,实现对复杂句法结构的组合泛化^[26]。LSTM在增加了门控单元之后,虽然缓解了RNN面临的梯度问题,但也增加了更多参数,导致了计算复杂度高和内存消耗大的问题^[27]。

2.1.2 Transformer 架构

Transformer 架构^[28]最初由 Google 在 2017 年提出,其关键组成部分包括多头注意力、前馈神经网络、位置编码等。Transformer 架构核心设计围绕多头注意力机制展开:多头注意力将查询、键、值矩阵拆分为多个子空间,并行执行多组注意力计算,然后拼接计算结果,增强模型的组合泛化能力。前馈神经网络由两层全连接网络构成,中间使用 ReLU 激活函数,通过非线性变换操作,增强模型的表达能力。位置编码为序列位置信息生成唯一的编码向量,弥补自注意力本身不包含序列位置信息的缺陷。

如图3所示,Transformer 架构由多个编码器-解码器层堆叠而成,二者均主要由多头注意力层和前馈神经网络组成。其设计突破了传统模型的局部感知限制,能够从有限数据中抽象出可推广的规则,从而处理未见过的复杂组合结构。

Transformer 架构的层级堆叠结构逐步整合信息,从基础单元到复杂结构进行多粒度抽象,实现了对未知组合的泛化。Transformer 动态绑定和上下文敏感的特性,使其在语言、代码等任务中,能高效重组已有知识解决新问题,突破传统模型对固定模板的依赖,实现跨结构、跨模态的组合泛化能力。在视觉问答(visual question answering, VQA)任务中,Russin等^[29]通过分离对齐和翻译机制,在SCAN数据集中验证了仅使用标准递归网络即可提升模型组合泛化能力。Meta & CMU 采用 MT-ACT 架构(multi-task action chunking Transformer)将动作序列分块建模,在38个机器人任务中实现了12种技能的组合泛化,同时将L3级(新技能-物体组合)任务成功率提升了22%^[30]。尽管Transformer架构在某些场景性能优异,但仍会出现长序列任务泛化不稳定的现象,导致在面对训练中未出现的长序列时,泛化性能不佳^[28]。

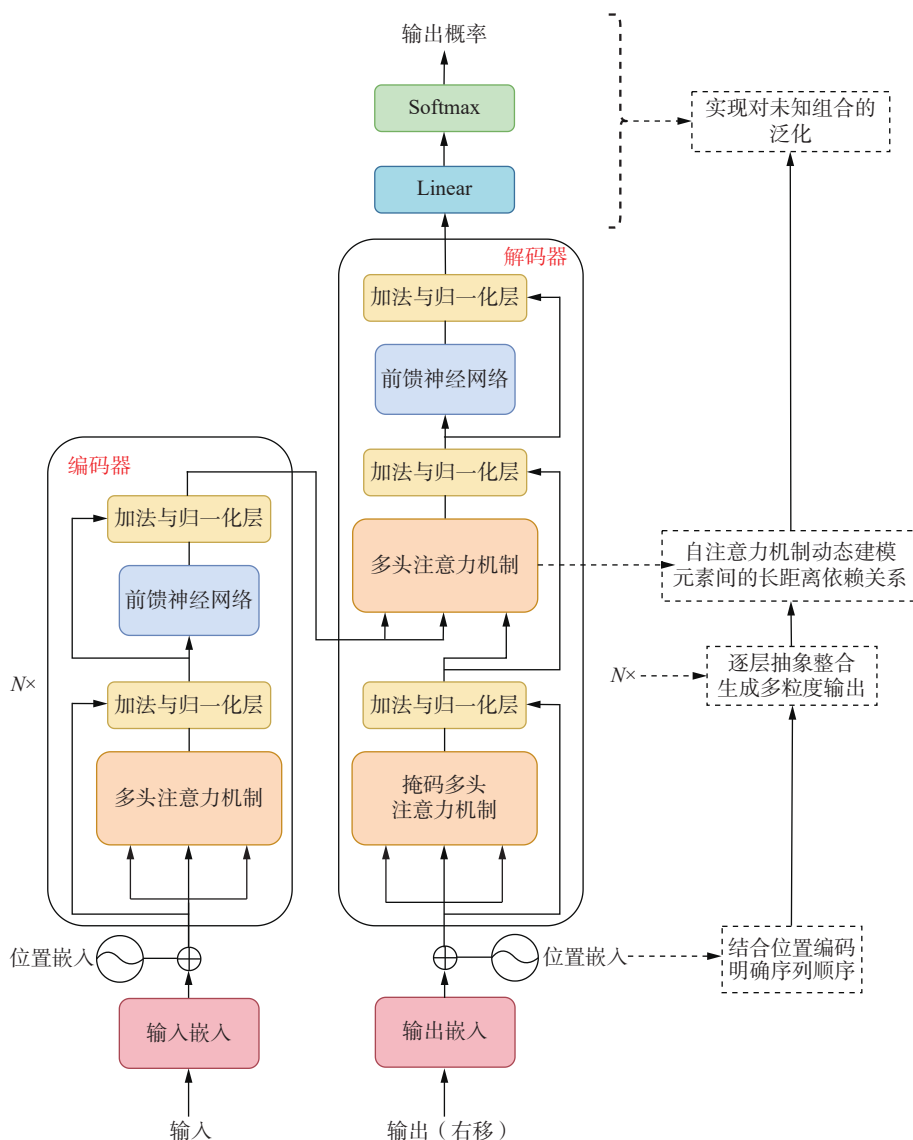


图3 Transformer 架构

Fig. 3 Transformer architecture

2.1.3 模块化神经网络

模块化神经网络(modular neural networks, MNNs)^[31]是由多个独立子网络(模块)组成的系统,这些模块分别处理不同的任务或输入,然后通过某种方式整合结果。模块化神经网络通过显式解耦功能模块与组合规则,模仿人类“分治-组合”的认知机制,已成为当前提升组合泛化能力的前沿方法^[32]。其通过训练大量不同的网络,构成一个神经“模块”的集合,这些模块是预先训练好的独立子网络,功能固定,每一个模块都实现推理的一个步骤。当需要在新问题上使用已训练的模型时,可以动态地组合这些模块,使之成为一个针对该问题的新的网络结构,即模块是静态预定义的,但组合是动态的。如图4所示,对于“蓝色的圆柱体和红色金属物体右侧的大块材料是一样的吗?”这一问题,模型首先将问题信息拆分成

单个模块,但不会将模块作为独立单元。每个模块都是推理到问题最终答案的一个步骤,通过将所有模块动态组合解决未知问题。

模块化神经网络的设计架构围绕功能解耦和动态组合展开^[33]。功能解耦的方法是将网络划分为 K 个功能独立的子模块 $\{M_i\}_{i=1}^K$,每个模块学习特定子任务。参数隔离是实现功能解耦的核心手段,能够有效避免任务间的知识混淆和冗余。参数隔离通过约束模块间连接稀疏性实现,公式为

$$\mathcal{L}_{\text{decouple}} = \sum_{i \neq j} \|\mathbf{W}_{M_i}^T \mathbf{W}_{M_j}\|_F^2$$

式中: $\mathbf{W}_{M_i}$ 为模块 M_i 的参数矩阵,正交约束减少知识冗余。

动态组合的核心目标为整合模块输出,即将激活模块的结果按需组合为最终输出。输出公式为

$y = \sum \alpha_i M_i(x)$ 或 $y = f(\{M_i(x)\})$
功能解耦确保了模块的高效学习与跨任务复

用, 动态组合实现了模块的灵活重组, 二者协同作用实现了模块化神经网络对未见组合的泛化能力。

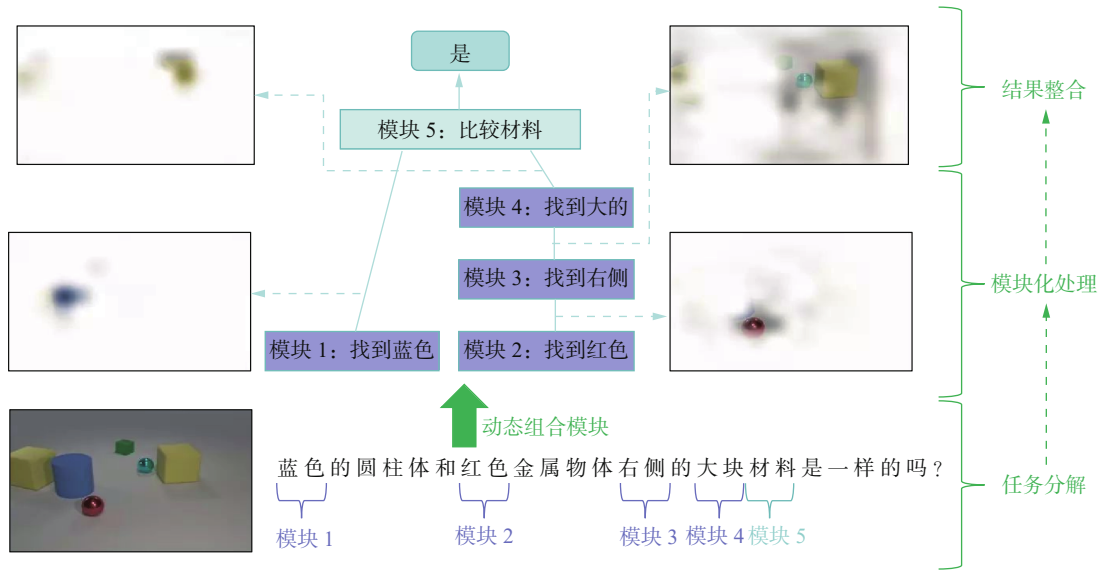


图 4 模块化神经网络流程示意

Fig. 4 Schematic of modular neural network processing workflow

与模块化神经网络相比, 神经元路由 (如 mixture of experts, MoE) 和胶囊网络虽在模块化实现机制上有所不同, 但同样关注模块化与动态组合, 具有一定的组合泛化潜力。在神经元路由 (如 MoE) 中, 模块是预定义的, 但路由网络动态选择激活哪些专家。这与模块化神经网络类似, 但模块化神经网络更强调功能解耦, 而 MoE 的路由更隐式。MoE 通过稀疏激活实现计算效率, 但模块功能可能重叠, 泛化能力依赖路由网络的训练^[34]。而在胶囊网络中, 胶囊是动态创建的, 模块 (胶囊) 本身是隐式表征, 通过向量输出和动态路由实现组合。胶囊网络擅长表征部分与整体的关系, 但对复杂逻辑推理的处理不如模块化神经网络直观。

相较于其他模块化方法, 模块化神经网络通过显示功能解耦和动态组合, 更直接地提升了组合泛化能力, 尤其适合需要多步骤推理的复杂任务。Andreas 等^[35]提出的 NMN (neural module networks), 通过将问题解析为子任务模块 (如“查找”“比较”“计数”), 动态组合子任务模块生成问题的推理路径。在 CLEVR 数据集上, NMN 的组合泛化准确率显著高于端到端模型, 且可解释性明显增强。Zheng 等^[36]设计了包含属性、位置、关系等专用模块的 MGA-Net, 通过门控图网络整合多个子模块输出, 实现了视觉关系推理。MGA-Net 在 RefCOCO+ 等数据集上表现优异, 为模块化在视觉组合泛化领域的应用指明了前景。然而, 模

块化神经网络的组合灵活性较差, 模型无法处理超出预设模块的未知组合, 这在一定程度上削弱了模型的组合泛化能力^[20]。

2.1.4 解耦表示学习

解耦表示学习 (disentangled representation learning) 是一种机器学习技术, 旨在从复杂数据中分离出隐含的、可解释的生成因子, 形成彼此独立的表示, 每个生成因子仅捕捉数据的一个方面, 从而提升模型的可解释性、泛化能力和操作灵活性^[37]。解耦表示学习试图让模型可以熟练地将数据中的各类变化因素分离编码, 使得潜在表示中的每个维度或子集都对应一个独立且具有明确物理含义的生成因子, 因而模型能够像搭积木一样, 将这些解耦后的因子进行重新组合, 进一步理解和生成训练时从未见过的新组合。这也是解耦表示学习实现组合泛化的核心原理, 即通过分离数据中的独立变化因子, 使模型能够独立操作和重组这些因子, 从而支持对新组合的泛化。在实现解耦表示的基础上, 结合其他机制的辅助, 模型可以较好地实现更为复杂的组合泛化。如图 5 所示, 在实现解耦表示学习的基础上, 结合组合操作实现组合泛化。

1) 解耦表示学习

使用变分自编码器 (variational autoencoder, VAE) 框架实现解耦表示学习。目标是将输入数据 x 映射到潜在空间 z , 其中每个维度 z_i 独立且对应一个解释因子。

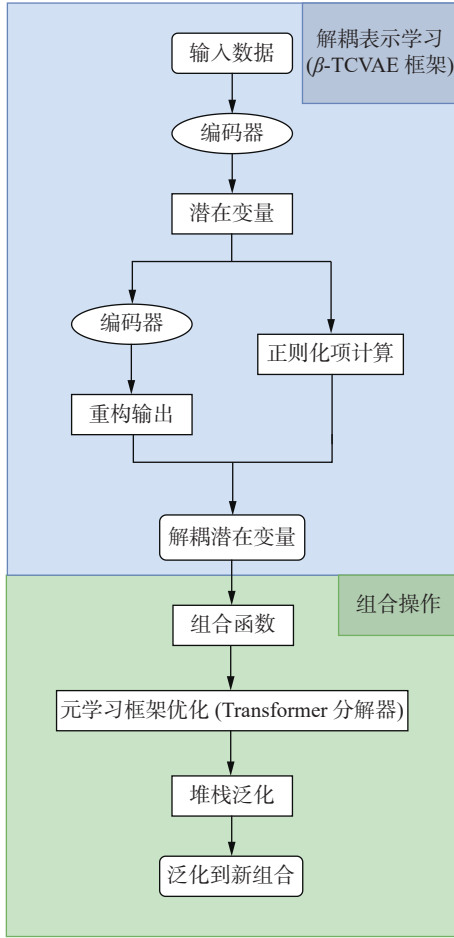


图 5 解耦表示学习和组合操作示意

Fig. 5 Schematic of disentangled representation learning and compositional operation

采用 β -TCVAE, 其目标函数在标准 VAE 的 ELBO 基础上, 添加解耦正则项:

$$\mathcal{L}_{\beta\text{-TC}} = \mathbb{E}_{q_{\phi}(\mathbf{z}|\mathbf{x})} [\log p_{\theta}(\mathbf{x}|\mathbf{z})] - \alpha I_q(\mathbf{x}; \mathbf{z}) - \beta T_C(\mathbf{z}) - \gamma \sum_{j=1}^d K_L(q(\mathbf{z}_j) \| p(\mathbf{z}_j))$$

式中: $\mathbb{E}_{q_{\phi}(\mathbf{z}|\mathbf{x})} [\log p_{\theta}(\mathbf{x}|\mathbf{z})]$ 是重构项, 确保解码器 p_{θ} 能准确重建输出。 $\alpha I_q(\mathbf{x}; \mathbf{z})$ 控制输入与潜在变量的互信息, 防止过拟合 ($\alpha > 0$)。 $\gamma \sum_{j=1}^d K_L(q(\mathbf{z}_j) \| p(\mathbf{z}_j))$ 是各潜在维度的 KL 散度, 约束其接近先验分布 $p(\mathbf{z}_j)$, 通常为标准正态分布。

关键项解耦: $\beta T_C(\mathbf{z})$ 是总相关, 定义为

$$T_C(\mathbf{z}) = K_L(q(\mathbf{z}) \| \prod_{i=1}^d q(\mathbf{z}_i))$$

式中: T_C (total correlation, TC) 用于衡量潜在变量间的依赖性; 最小化 $T_C(\mathbf{z})$ 迫使 $\mathbf{z}_i$ 相互独立, 即解耦。

2) 组合操作

在解耦潜在空间 $\mathbf{z}$ 上, 定义组合函数 $C(\cdot)$, 将原子因子组合为新结构, 使用元学习框架实现系统性组合。设输入序列 S 被解析为组合树, 其中

叶节点为原子因子, 非叶节点为组合操作。组合函数定义为

$$\mathbf{z}_{\text{composed}} = C(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k; \Theta)$$

Θ 为可学习参数, 通过元学习优化以支持单步分解, 将深层子树压缩为浅层表示。

系统性组合: 在解耦空间中, 组合等价于因子间的代数运算, 以命令 “dax twice” 的组合为例:

$$\mathbf{z}_{\text{composed}} = \mathbf{W} \cdot (\mathbf{z}_{\text{dax}} \otimes \mathbf{z}_{\text{twice}})$$

式中: $\otimes$ 表示张量积, $\mathbf{W}$ 为投影矩阵。

堆栈泛化: 组合多个基础模型的输出, 通过元模型融合, 提升泛化能力, 定义为

$$f_{\text{ensemble}}(\mathbf{x}) = g\left(\sum_{k=1}^K w_k f_k(\mathbf{x})\right)$$

式中: f_k 是第 k 个基础模型的预测, $g(\cdot)$ 是元组合函数。

系统性组合和堆栈泛化是实现组合泛化的关键, 使模型能够有效处理未见组合。解耦表示学习在多个领域被应用于提升组合泛化能力, 尤其在分布偏移和零样本场景下。在自然语言处理中, 通过解耦序列到序列模型在编码器端自适应重新编码源输入, 以目标上下文为条件, 改善了组合泛化。例如, 在 COGS 和 GeoQuery 数据集上处理新指令组合, 或在机器翻译中处理新词汇搭配, 提升 BLEU 分数^[38]。虽然解耦表示促进组合泛化的理论基础极具吸引力, 但解耦可能只是必要条件之一, 并非充分条件。学习到的表示在解耦指标上得分高, 并不一定能转化为在下游任务中更好的性能^[39]。多项研究系统性地测试了这一假设, 发现解耦模型未能支持更困难的组合泛化形式。

2.2 关系建模架构

关系建模架构是实现组合泛化的核心支撑, 其定义了实体间交互的逻辑与结构规则。关系建模架构通过符号逻辑与动态交互的协同, 实现了从静态结构到动态行为的覆盖, 其核心价值在于平衡模型的确定性及灵活性。根据关系建模架构基础理论和促进组合泛化机制的差异, 主要划分为图卷积网络 (graph convolutional network, GCN)、图注意力网络 (graph attention network, GAT) 和神经符号集成, 这种分类方式清晰地反映了关系建模架构从纯神经网络到融合符号逻辑的技术演进路径。

2.2.1 图卷积网络

图卷积网络 (GCN)^[40] 是一种将卷积运算从传统数据 (如图像) 推广到图数据的神经网络模

型。GCN 的核心思想是邻域聚合^[41], 即每个节点的表示不仅由自身特征决定, 还通过与其邻居节点的信息交互来更新。它能够将图的拓扑结构与节点特征相结合, 通过图卷积操作来提取节点之间的关系和特征信息, 直接处理不规则的图结构。图 6 为 GCN 示意, 表示 GCN 在半监督节点分类任务中的架构示意。输入通道 (C 个), 每个节点通过特征矩阵 $\mathbf{X} \in \mathbf{R}^{N \times C}$ 表示, N 为节点数, C 为特征维度。输出特征图 (F 个), 最后一层通过 Softmax 激活生成类别概率分布 $\mathbf{Z} \in \mathbf{R}^{N \times F}$, F 为分类任务的类别数。少量节点的标签 Y_i 参与监督损失计算 (交叉熵), 模型通过图结构将标签信息传播到未标记节点。

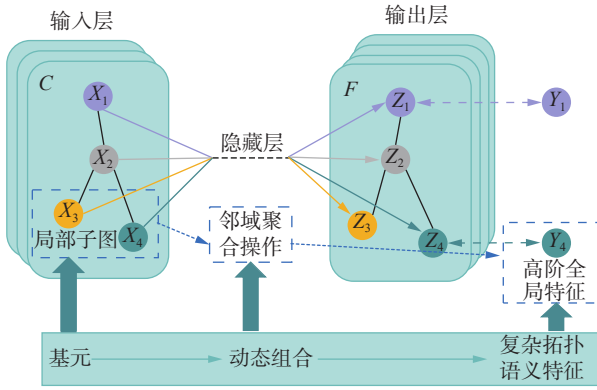


图 6 GCN 示意

Fig. 6 GCN schematic

GCN 基于图结构的谱域或空间域卷积操作, 通过邻接矩阵和节点特征矩阵实现信息传播, 其核心公式为

$$\mathbf{H}^{(l+1)} = \sigma \left(\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \mathbf{H}^l \mathbf{W}^l \right)$$

式中: $\mathbf{H}^{(l+1)}$ 表示第 $l+1$ 层的节点特征矩阵; $\tilde{\mathbf{D}}$ 是 $\tilde{\mathbf{A}}$ 的度矩阵; $\tilde{\mathbf{A}}$ 为加入自连接后的邻接矩阵; $\mathbf{H}^l$ 表示第 l 层的节点特征矩阵; $\mathbf{W}^l$ 表示第 l 层可学习权重矩阵; σ 为激活函数, 通常可以采用 ReLU、Sigmoid 等。

GCN 通过层级化的邻域信息聚合机制, 显著提升了图数据的组合泛化能力。其核心在于将局部子图模式通过可微的邻域聚合操作逐层组合为高阶全局特征, 同时利用参数共享策略学习与位置无关的通用结构规律。这种设计使模型能够从已知的原子化图模式中解耦出可复用的基元^[42], 并通过动态组合推理出未见过的复杂拓扑语义特征。KSII 团队^[43] 提出 GCNXSS (graph convolutional network-based cross-site scripting detection approach) 方法, 将统一资源定位符 (uniform resource locator, URL) 中的词序列转换为词同质图, 利用 GCN 挖掘潜在攻击模式, 使泛化未见过的攻

击变体的能力相比传统规则引擎提高了 32%。吴琳等^[44] 提出基于结构误差的 GCN 模型, 通过改进支持向量机的损失函数, 约束同类样本的类内紧密度和类间间隔。在 NTURGB+D120 数据集上, 模型对复杂动作的组合泛化准确率提升至 89.7%。在社交网络谣言检测任务中, Kipf 和 Welling^[41] 提出的经典 GCN 模型通过邻接矩阵传播标签信息, 泛化到未标注节点的新组合, 实现了 81.5% 的分类准确率。但在实际应用中, 多层 GCN 容易产生过平滑的问题, 反而降低了区分复杂组合的能力^[16,45]。

2.2.2 图注意力网络

图注意力网络 (GAT)^[46] 是一种基于注意力机制的图神经网络, 其核心思想是通过注意力机制来计算节点之间的关系, 确定节点与其邻居节点之间的重要性, 实现对邻居节点的加权聚合。这种注意力机制使模型能够根据节点之间关系的重要性获得相应的权重, 并依据这个权重来更新节点的状态, 达到更好地捕捉图中结构信息的目的。在 GAT 的实际运算中, 注意力系数主要由“注意力头”结构来计算, 每个注意力头都能单独计算注意力系数, 并通过将所有注意力头的计算结果拼接起来而得到最终结果。下面主要对计算注意力系数和加权求和的过程进行介绍。

如图 7 所示, 左侧为计算注意力系数的过程, 对于节点对 (i, j) , $\mathbf{h}_i$ 、 $\mathbf{h}_j$ 分别表示节点 i 、 j , $\mathbf{W}$ 为一个投影矩阵。通过注意力网络 a 得到注意力值 e_{ij} , 然后经过 Softmax 进行正则化处理, 得到注意力权重 α_{ij} ; 右侧为加权求和的示意图, 节点 1 对其邻域进行多头注意力 ($K=2$, K 表示注意力头数量) 的示例, 不同的颜色表示独立的注意力计算, 每个头的聚合特征被拼接或平均以获得 $\mathbf{h}'_1$ 。

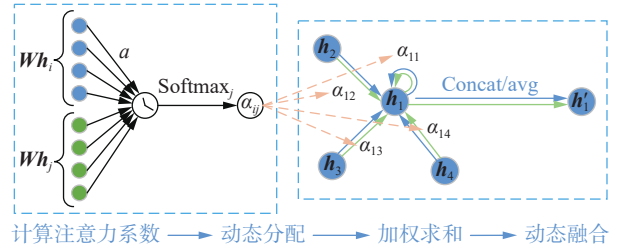


图 7 图注意力网络动态融合过程

Fig. 7 Graph attention network dynamic fusion process

1) 计算注意力系数

对于顶点 i , 逐个计算它与它的邻居们 ($j \in \mathcal{N}_i$) 之间的相似系数。

$$e_{ij} = a(\mathbf{W}\mathbf{h}_i, \mathbf{W}\mathbf{h}_j)$$

为了使得互相关系数更容易计算和便于比较, 引入了 Softmax 对所有的 i 的相邻节点 j 进行

正则化:

$$\alpha_{ij} = \text{Softmax}_j(e_{ij}) = \frac{\exp(e_{ij})}{\sum_{k \in N_i} \exp(e_{ik})}$$

注意力机制为

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(\mathbf{a}^T [\mathbf{W}\mathbf{h}_i] [\mathbf{W}\mathbf{h}_j]))}{\sum_{k \in N_i} \exp(\text{LeakyReLU}(\mathbf{a}^T [\mathbf{W}\mathbf{h}_i] [\mathbf{W}\mathbf{h}_k]))}$$

2) 加权求和

根据计算得到的注意力系数,把特征加权求和:

$$\mathbf{h}'_i = \sigma \left(\sum_{j \in N_i} \alpha_{ij} \mathbf{W}\mathbf{h}_j \right)$$

式中: $\mathbf{h}'_i$ 是 GAT 输出的对于每个顶点 i 的新特征(融合了邻域信息)。为了增加注意力网络的稳定性,使用多头注意力机制,同时用多个不同的注意力机制 $\mathbf{a}$ 来进行计算,再对最后的结果进行平均:

$$\mathbf{h}'_i = \sigma \left(\frac{1}{K} \sum_{k=1}^K \sum_{j \in N_i} \alpha_{ij}^k \mathbf{W}^k \mathbf{h}_j \right)$$

GAT 在促进组合泛化中的作用主要体现在其动态注意力机制对局部子结构的敏感捕捉与灵活组合能力。通过为每个节点自适应地分配邻居权重, GAT 能够识别图中关键的子结构,并在新输入中将这子结构以未见过的组合方式动态融合。在 SKEMPI 蛋白质结合亲和力预测任务^[47]中,基于 GAT 的多构象蛋白质复合物建模框架通过集成不同构象下的子图表征,实现了跨突变体的组合泛化能力。Poursaeed 等^[48]提出的对抗性训练 GAT (graph adversarial training), 通过生成无限制对抗样本,在语义分割和目标检测任务中首次实现了跨任务泛化。Zhang 等^[49]提出的权重约束 GAT (weights-bounded GAT), 通过限制权重矩阵的 Frobenius 范数,减少了模型的过拟合风险并

提升了泛化能力。然而,由于在不同时间对不同注意力矩阵进行乘法运算, GAT 的过度平滑更为复杂,在一定程度上直接削弱了模型对未见组合的推理能力^[50]。

2.2.3 神经符号集成

神经符号集成 (neuro-symbolic integration)^[51] 是一种将神经网络的感知学习能力与符号系统的逻辑推理能力相结合的人工智能范式,旨在解决传统神经网络缺乏可解释性和组合泛化能力的问题。神经符号集成的核心逻辑是“神经处理+符号推理”的协同:神经网络部分负责从原始数据中提取特征和学习复杂的输入输出映射,而符号系统部分则负责将这些特征和映射结果转换为有意义的符号表示,并进一步抽取出数据中隐含的结构化关系或模式,最终通过逻辑规则和推理过程进行处理。神经符号集成正是通过结合神经网络强大的感知与模式识别能力,以及符号系统强大的逻辑推理与可解释性,优化目标确保系统能够处理新组合,解决了传统神经网络缺乏组合泛化能力的问题。神经符号递归机 (neural-symbolic recursive machine, NSR)^[52] 作为神经符号集成的多种实现方式之一,因其递归结构天然支持组合性,从而在促进组合泛化方面具有先天优势。如图 8 所示, NSR 将组合泛化建模为概率生成过程,包含 3 个模块。1) 神经感知模块:将原始输入 x 映射为符号表示 s , 概率模型为 $p(s|x; \theta_p)$ 。2) 依赖解析模块:生成结构化表示 e , 概率模型为 $p(e|s; \theta_s)$ 。3) 程序归纳模块:输出可执行程序 v , 概率模型为 $p(v|e, s; \theta_t)$ 。整体推导通过最大化联合概率 $p(v, e, s|x)$ 实现, 目标函数为

$$\max_{\theta_p, \theta_s, \theta_t} E_{(x,v)} [\log p(v, e, s|x)]$$

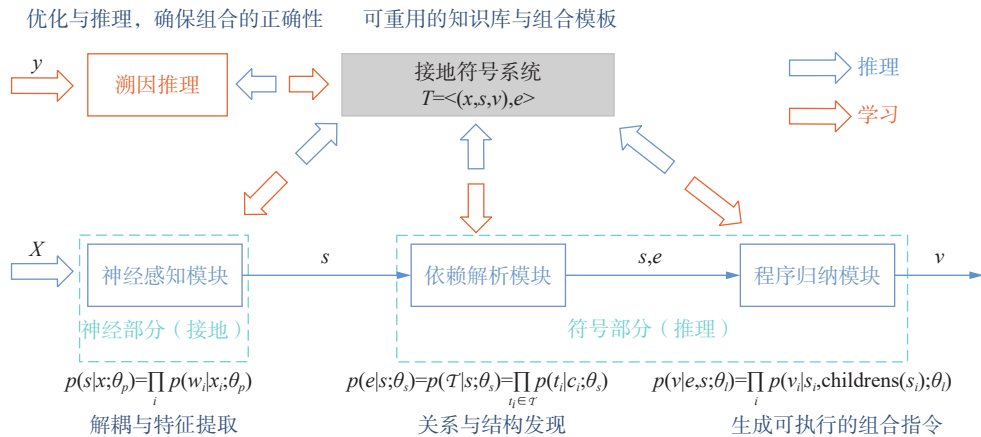


图 8 NSR 概率建模框架

Fig. 8 Probabilistic modeling framework diagram for neural-symbolic reasoning

联合概率描述了从原始输入 x 生成最终输出 v 的完整概率过程。整个系统的核心任务,就是理解抽象的指令 y ,并在感知到的具体环境 x 中,生成一个可执行的、能完成该指令的具体程序 v 。简单来说,指令 y 是系统的输入目标和任务,程序 v 是系统的输出解决方案。系统训练的目标是最大化联合概率,即通过调整参数 θ ,使得对于训练数据,生成正确答案 v 的概率最大。

原始输入 x 代表具体的、高维的、原始的感知数据, y 代表一个抽象的指令、目标或问题。溯因推理是逻辑推理的一种形式,意为推理出最佳解释,它接收指令 y ,结合知识库 T 中的背景知识,推测出完成这个指令所需要的抽象规划或假设。神经感知模块的关键作用在于实现高效的特征提取与表示学习,将复杂的感知数据 x 转化为更抽象、更符号化的中间表征 s 。依赖解析模块接收来自神经感知的符号化表征 s ,并进一步分析这些元素之间的结构关系和依赖关系,输出更丰富的结构化信息,其核心任务是分析并显式地抽取出这些符号之间的依存关系 e 。程序归纳模块是实现组合泛化的最核心模块,接收带有结构的 (s, e) ,并将其“编译”或“归纳”为一个可执行的程序 v ,程序 v 本质上是一系列可组合的操作指令。接地符号系统是一个框架,各模块作为其核心组成部分,共同协作来实现“接地符号系统”的功能。其包含符号接地和双向桥接两个核心作用:符号接地将来自顶层的抽象符号与底层神经网络感知到的具体实例绑定起来,赋予符号实际的意义;双向桥接是指上下两层之间建立了一个闭环,自上而下将溯因推理产生的抽象规划分解和具体化为一系列符号指令(如 e),传递给程序归纳模块,指导它如何生成程序;自下而上是指接地符号系统接收并整合来自底层的符号化感知信息(如 s),用以更新和维护当前的世界状态模型,确保顶层推理所依据的信息是实时且准确的。接地符号系统是动态处理架构,知识库 T 是其内部核心,定义了数据流($x \rightarrow s \rightarrow v$)和规则库(e),共同实现感知到行动的智能转换。推理(蓝色箭头)是一个执行过程,利用系统当前已有的知识和规则,对新的、未知的输入进行理解和解释;学习(橙色箭头)是一个优化过程,根据推理结果的反馈,来更新和修正系统内部的知识库和模型参数,从而使系统在未来表现得更好。

NSR通过解耦,结构化、可执行的程序归纳以及接地符号的存储与重用,较好地实现了组合泛化,不仅能处理见过的例子,更能通过重新组

合已知的、已理解的元素和规则,理解和创造无限的新组合^[53]。在视觉推理扩展任务中,将复杂的视觉场景映射为符号谓词,并利用神经符号方法进行程序归纳,以解决需要多步推理的视觉任务^[53]。NSR在多个基准测试中表现出色,例如在SCAN语义解析、HINT(handwritten arithmetic with INTEGERS)算术推理和组合机器翻译任务中,NSR实现了100%的通用性,超越了现有神经网络和混合模型^[52]。但NSR在处理多变量问题时,基于Transformer的模型可能存在记忆偏差,导致模型倾向于生成训练数据中已有的表达式,难以进行组合生成并探索新表达式^[54]。

2.3 学习机制架构

学习机制架构是组合泛化的核心驱动力,旨在通过优化模型的训练范式,从有限的数据中提取通用规则,并将已知概念动态组合以解决新问题。作为学习机制架构的重要核心,元学习通过“学会如何学习”的机制,克服了传统机器学习需要大量有标签的训练数据的限制,能够让模型利用现有经验和知识快速适应新的任务和环境,为实现组合泛化提供了核心支持^[55]。在学习机制架构分类中,根据核心实现原理与策略导向的不同,可划分为基于优化的元学习(如 model-agnostic meta-learning, MAML)、基于度量的元学习(如 prototypical networks, ProtoNet)以及基于数据增强和课程学习的方法。

2.3.1 基于优化的元学习

基于优化的元学习(optimization-based meta-learning)是一类通过调整模型的优化过程来快速适应新任务的元学习方法。该方法有多种,MAML作为首选的基线方法,可通过双层优化结构提升模型的泛化能力,因而主要以MAML方法为例进行介绍。MAML的核心思想是通过优化初始模型参数,使模型在面对新任务时,只需经过少量训练数据和梯度更新,就能快速收敛到理想的解决方案。其特点在于对模型结构没有过多依赖,因此可以应用于任何基于梯度的机器学习模型。MAML的关键数学原理是通过双层优化学习一个通用的初始化参数 θ ,使得在少量梯度更新后能快速适应新任务^[56]。其计算步骤如下:

1) 内层优化

对每个任务 $\mathcal{T}_i$,从初始化参数 θ 出发,通过一步或多步梯度下降更新得到任务特定参数 θ'_i :

$$\theta'_i = \theta - \alpha \nabla_{\theta} \mathcal{L}_{\mathcal{T}_i}(f_{\theta})$$

式中: α 是内层学习率, $\mathcal{L}_{\mathcal{T}_i}$ 是任务 $\mathcal{T}_i$ 的损失函数。

2) 外层优化

外层更新通过梯度下降进行:

$$\theta \leftarrow \theta - \beta \nabla_{\theta} \sum_{\mathcal{T}_i} \mathcal{L}_{\mathcal{T}_i}(f_{\theta_i})$$

其中 β 是外层学习率。

优化初始参数 θ ,使得所有任务在更新后的参数 θ_i 上损失最小化:

$$\theta^* = \arg \min_{\theta} \sum_{\mathcal{T}_i \sim p(\mathcal{T})} \mathcal{L}_{\mathcal{T}_i}(f_{\theta_i})$$

MAML 中执行了两次梯度下降,分别作用在基学习器和元学习器上。图9给出了MAML中特定任务参数 θ_i 和初始化参数 θ 的更新过程。灰色线表示特定任务的损失函数的梯度值(方向),每个梯度向量指向特定任务损失函数的最陡上升方向,实现快速适应;黑色线表示初始化参数选择更新的方向(黑色线方向是几个特定任务产生方向的平均值);虚线代表跨任务泛化,表示在各自任务上进行单步梯度下降后得到的适应参数,不同的方向代表不同任务更新的方向。

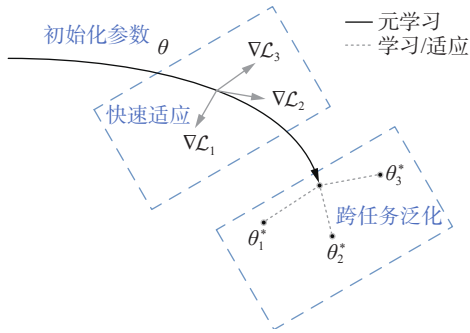


图9 MAML 示意

Fig. 9 MAML schematic

MAML 初始化参数 θ 对任务损失的梯度方向高度敏感,使得内层更新能快速调整参数以适应新任务的组合结构。MAML 在元训练阶段接触大量基础任务,其目标是学会一个“万能”的初始化参数 θ 。当遇到新组合时,模型从这个初始化参数出发,仅需少量样本就能快速调整参数,从而解决这个新任务。组合泛化所要求的“用旧元素组合出新含义”的能力,在数学上对应了MAML参数空间中“从初始点出发快速到达新点”的能力。因此,MAML“学会学习”的能力,使其在机制上天然适合于处理组合泛化问题。

MAML 通过提高模型参数的初始化敏感性与跨任务泛化能力,在组合泛化任务中表现出独特优势。例如,在强泛化评估集^[57]中,MAML 的准确率为42.87%,显著优于传统序列模型(<30%)。在组合动作序列^[58]任务中,MAML 仅需1~2次策略更新即可适应新环境,表明了其对物理规则组合的适应性。在少样本迁移学习^[59]中,MAML 通

过参数敏感初始化,使CLIP(contrastive language-image pre-training)模型在组合概念任务中的零样本性能提升了15%。但在组合泛化任务中,新组合可能会影响原有参数空间的连续性,导致元参数难以通过梯度更新来获取新规则^[60]。

基于优化的元学习在解决组合泛化问题上也存在相当的局限性。首先,MAML的初始化参数 θ 是通过元训练任务分布 $p(\mathcal{T})$ 优化得到的,模型只能良好地适应那些与元训练任务在组合模式上相似的新任务。如果新组合的模式完全超出了元训练分布的覆盖范围,那么从 θ 出发的梯度更新方向可能无法有效收敛到正确的解,因而其组合泛化能力被严格限制在元训练所见的组合模式之内。其次,MAML的有效性依赖于一个假设:所有任务(包括新组合任务)的最优参数都分布在初始化参数 θ 的一个连续邻域内,因此可以通过梯度下降这种局部搜索方法到达。当某些复杂的、离散的组合泛化可能需要模型参数发生非连续的、跳跃性的改变时,基于局部梯度更新的MAML机制可能会失效。最后,MAML依赖于“少量”梯度更新来适应新任务。对于某些组合泛化任务,可能需要更深入的特征重构或推理,此时有限的几步更新无法完成任务,从而导致泛化失败^[61]。

2.3.2 基于度量的元学习

基于度量的元学习(metric-based meta-learning)是一类通过学习样本间的相似性度量来快速适应新任务的元学习方法。该方法以原型网络ProtoNet^[62]为例介绍。相比于其他基于度量的元学习方法,ProtoNet具有高效的原型聚合机制、对噪声的鲁棒性以及架构轻量化的优势,因而得到了更加广泛的应用。ProtoNet是一种经典的、简单有效的小样本方法,其核心在于通过原型向量度量样本间的相似性,并通过欧氏距离分类新样本,使其在组合泛化任务中展现出独特优势。ProtoNet通过度量学习构建一个高维嵌入空间,使得同类样本靠近,异类样本远离。其核心步骤如下。

1) 原型计算

对每个类别 k ,计算支持集 $\mathcal{S}_k$ 的均值作为原型向量 $\mathbf{c}_k$:

$$\mathbf{c}_k = \frac{1}{|\mathcal{S}_k|} \sum_{(x_i, y_i) \in \mathcal{S}_k} f_{\phi}(x_i)$$

式中: f_{ϕ} 为嵌入函数, ϕ 为嵌入网络的参数。

2) 分类决策

对查询样本 $\mathbf{x}_q$,计算其与所有原型的欧氏距

离, 并通过 Softmax 生成概率分布:

$$p(y = k | \mathbf{x}_q) = \frac{\exp(-d(f_\varphi(\mathbf{x}_q), \mathbf{c}_k))}{\sum_j \exp(-d(f_\varphi(\mathbf{x}_q), \mathbf{c}_j))}$$

其中 $d(\cdot, \cdot)$ 为欧氏距离的平方。

3) 损失函数

最小化查询集的负对数似然损失:

$$\mathcal{L} = - \sum_{(\mathbf{x}_q, y_q)} \log p(y = y_q | \mathbf{x}_q)$$

如图 10 所示, ProtoNet 利用聚类思想, 将支持集投影到一个度量空间, 在欧氏距离度量的基础上获取向量均值, 对查询样本计算到每个原型的距离, 实现分类。

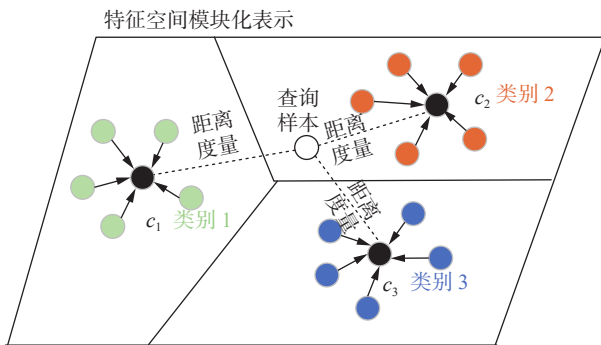


图 10 ProtoNet 示意

Fig. 10 ProtoNet schematic

ProtoNet 通过嵌入空间的特征解耦和原型聚合机制促进组合泛化: 其嵌入网络隐式学习数据中的基础特征, 并在低维空间中形成模块化表示, 使得新组合可通过特征相似性度量进行分类。模型在元训练阶段接触大量基础任务, 学习如何从支持集中提取原型。在测试时, 面对新组合, 模型只需计算其嵌入向量与由基础特征组合而成的各个原型的距离即可快速分类。例如, 当新类别由已知特征组合构成时, 查询样本的嵌入向量与对应原型的距离可能更小^[63], 从而利用已有特征的线性或非线性组合实现泛化。因此, 解决组合泛化问题被转化为一个元学习问题^[64]。元学习“学会”的特征解耦和相似性度量能力, 直接服务于组合泛化所要求的基础元素的重新组合与推理。

在语义解析任务中, ProtoNet 通过原型映射处理未见过的组合指令。例如, 训练数据包含“打开灯”和“关闭风扇”, 测试时需要理解“关闭灯+打开风扇”的新组合。实验显示, 结合 BERT (bidirectional encoder representations from Transformers) 嵌入的 ProtoNet 在 CFQ 数据集上 F_1 分数达到 72.1%, 优于传统序列模型^[65]。在视觉问答中, ProtoNet 被用于组合属性推理, 如在“黄色方块”

和“绿色圆形”训练样本中, 泛化到“绿色方块”的查询。通过局部原型对齐, 模型在 CLEVR 数据集上的组合泛化准确率提升至 68.3%。但由于 ProtoNet 的类原型对支持集样本数量高度敏感, 导致在少样本场景下, 原型无法充分覆盖类内特征的多样性, 从而使面对新组合时组合泛化能力下降^[66]。

基于度量的元学习在解决组合泛化问题上也存在相当的局限性。首先, ProtoNet 对元训练任务分布极度依赖, 其性能核心在于嵌入网络和原型计算, 这些都是在元训练阶段从支持集样本中学到的, 因而模型只能处理那些与元训练任务在组合方式上相似的新任务^[67]。其次, 由于 ProtoNet 的类原型对支持集样本数量高度敏感, 因而其组合泛化能力不是一个绝对能力, 而是一个相对能力。如果支持集样本少或有偏差, 其组合泛化能力自然会下降。最后, ProtoNet 的分类决策基于欧氏距离, 可能无法学会处理需要复杂逻辑推理而非简单距离度量的组合泛化任务^[68]。

2.3.3 基于数据增强和课程学习的方法

数据增强 (data augmentation) 和课程学习 (curriculum learning) 是两种既独立又常结合使用的技术, 它们的核心思想不同, 但组合后能产生“1+1 > 2”的效果。数据增强和课程学习都通过引入结构化的变化和循序渐进的学习过程, 来提升模型系统性泛化和动态适应能力^[69]。数据增强提供“素材”, 通过生成已知元素的多种变体和新组合, 为模型提供了大量练习组合泛化的“训练场”; 课程学习提供“教学计划”, 智能地调度包括原始数据和增强数据这些“素材”的学习顺序, 确保模型先掌握基础、核心的概念和特征, 然后再去应对更复杂的组合和变化。二者结合有助于模型通过由易到难的渐进过程逐步构建起应对未知组合的能力。如图 11 所示, 整个流程可分为数据增强 (蓝)、课程学习 (绿) 和训练过程 (黄) 三模块, 整个流程的核心是数据增强和课程学习的协同。

原始训练数据在数据增强模块中进行各种变换, 生成大量多样化、具有不同难度的新样本^[70]。简单增强对数据做轻微、不改变本质的扰动, 让模型学会对微小的、不重要的变化不敏感, 专注于核心特征; 中等增强引入更显著的变化和部分遮挡, 提升模型的鲁棒性, 使其在目标被部分遮挡或信息不完整时也能作出正确判断; 复杂增强生成接近真实世界极限情况的、变化最大的样本, 极大扩展训练数据的分布, 是提升泛化能力的关键。课程学习调度器是整个流程的“大脑”,

不直接处理数据,而是通过接收监控模型训练状态的反馈信号来智能地决定在什么时间、给模型提供什么难度的数据。在课程学习调度器的指挥下,模型的训练过程被系统地分为3个阶段:初始训练主要使用原始数据和简单增强数据,让模型快速掌握任务中的基本规律和核心特征;中期训练混合使用简单增强和中等增强数据,提高模型抗干扰能力;后期训练引入复杂增强数据,极大化模型的泛化能力,这是实现组合泛化的关键。

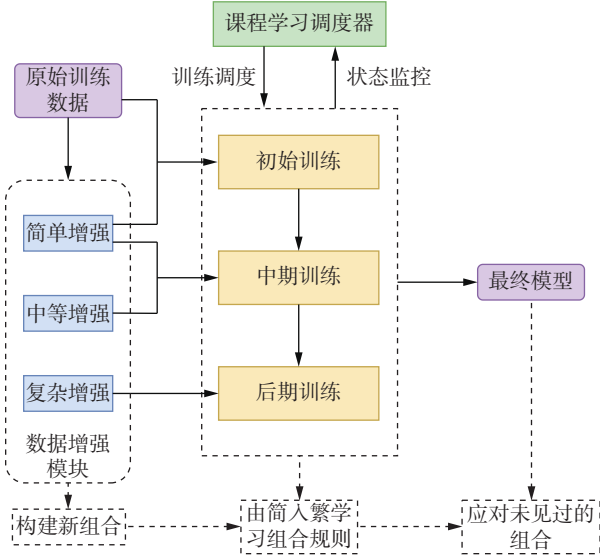


图 11 基于数据增强和课程学习的方法示意

Fig. 11 Schematic of methods based on data augmentation and curriculum learning

从概率框架看,组合泛化要求模型学习真实数据分布 $p_{\text{data}}(x,y)$ 。数据增强通过变换 T 近似地拓宽了训练数据的分布:

$$p_{\text{aug}}(x',y) = \frac{1}{N} \sum_{i=1}^N \delta(x' - T_i(x)) \cdot p_{\text{data}}(x,y)$$

式中: δ 是狄拉克函数, $T_i(x)$ 代表对原始数据样本 x 进行第 i 种数据增强变换, N 为数据增强变换的总数, $p_{\text{aug}}(x',y)$ 为增强后的数据概率分布。理想的增强应使 p_{aug} 更接近真实未知的分布 p_{real} 。

课程学习则可以看作是在训练过程中逐步优化一个随时间变化的目标分布:

$$p_{\text{curr}}^{(t)}(x,y) \propto p_{\text{aug}}(x,y) \cdot \mathbb{I}[D(x,y) \leq \text{threshold}(t)]$$

式中: $p_{\text{curr}}^{(t)}(x,y)$ 为模型在训练时刻 t 实际面对的训练目标分布, $p_{\text{aug}}(x,y)$ 为第一个公式得到的、增强后的完整数据分布。 $\mathbb{I}[\cdot]$ 为指示函数,判断方括号内的条件满足时,函数值为 1; 条件不满足时,函数值为 0。 $D(x,y)$ 为难度度量函数,给每个数据样本 (x,y) 分配一个标量值,代表模型学习这个样本的“难度”。 $\text{threshold}(t)$ 随时间 t 增加而提高,逐渐覆盖更难的样本,这使模型能逐步拟合更复杂的

数据分布。 $D(x,y) \leq \text{threshold}(t)$ 筛选出难度不大于当前阈值的样本。这种数据增强和课程学习相结合的模式使模型不仅能见到“更多”数据,还能以“更优的顺序”学习这些数据,从而更有效地学习和泛化到未知的元素组合上,应对真实世界中复杂多变的情境。

Akyürek 等^[71] 提出的 R&R (recombination and resampling) 方法,通过原型驱动的生成式数据重组与重采样,在指令遵循 (SCAN) 和形态分析 (SIGMORPHON 2018) 任务中,仅用 8 个初始样本即可学习新结构和时态,其性能媲美甚至优于当前最先进的神经符号方法,显著提升了神经序列模型的组合泛化能力。Qiu 等^[72] 提出 CSL (compositional structure learner) 模型,通过准同步语法骨架重组训练样本并微调 T5 模型。在语义解析和数学推理任务中,该方法将组合偏置转移至序列模型,在 CFQ 和 GeoQuery 数据集上超越基线模型 10% 以上。但 Keyzers 等构建的 CLEVR-CoGenT (CLEVR compositional generalization test) 基准表明:当训练-测试集的复合差异增大时,增强模型的准确率呈负相关,揭示了当前方法对数据分布偏移的脆弱性。

2.4 组合泛化方法对比分析

本节主要依据模型结构的差异介绍了现有的 3 类组合泛化方法,描述了这些方法的主要原理、相关计算过程和具体应用。各类方法与组合泛化特征映射如表 1 所示;相关方法、优缺点、复杂度变化、组合泛化机理、适用场景及应用领域举例如表 2 所示。

表 1 系统梳理了基础表征架构、关系建模架构和学习机制架构与组合泛化的系统性、可组合性、可解释性和动态适应性等特征之间的映射关系,显示了各类方法与组合泛化特征之间的紧密联系。

如表 2 所示,各类方法的优缺点、主要应用领域已较为详细,在这里不过多赘述。下面着重对各类方法的复杂度变化、组合泛化机理进行详细讨论。

在复杂度变化方面,递归神经网络通过隐状态递归传递序列信息,是基础的序列模型,复杂度低;Transformer 架构通过注意力权重可视化组合关系,复杂度中;模块化神经网络人工定义模块结构和交互方式,复杂度高;解耦表示学习的学习解耦表征本身就是一个难题,复杂度高;GCN 的复杂度与图的规模以及特征的维度线性相关,对所有邻居节点一视同仁,采用等权重的聚合方式,复杂度中;GAT 的复杂度同样与图的

规模线性相关,但由于需要为每条边计算注意力,其常数因子和实际计算开销通常高于 GCN,复杂度高;神经符号集成需要设计两种系统接口和协作框架,复杂度最高;基于优化的元学习需要内循环的多次梯度更新和外循环的二阶优化,

复杂度高;基于度量的元学习通过一个高效的特征提取器将问题转化为简单的最近邻分类问题,复杂度低;基于数据增强和课程学习的方法整体实现较为直接,但设计好的增强策略和课程需要大量领域知识,复杂度可变。

表 1 基于模型结构划分的组合泛化方法的组合泛化特征映射

Table 1 Compositional generalization feature mapping in compositional generalization methods based on model architecture taxonomy

方法		系统性	可组合性	可解释性	动态适应性
基础表征架构	递归神经网络	权重共享与递归连接支持泛化到新结构	递归处理逐步组合元素	层次化语义编码提供部分解释	时序计算实时处理序列, 参数共享适应可变量长输入
	Transformer架构	通过多头注意力和层级堆叠结构, 系统化泛化到新组合	自注意力动态绑定组件, 支持组件重组	注意力权重可视化	根据上下文动态调整注意力, 位置编码处理变长序列
	模块化神经网络	通过功能解耦和参数隔离实现结构化、规则化的推理流程	模块静态预定义但动态组合	模块对应可解释的子任务, 推理路径透明	动态组合, 根据输入实时组合模块
	解耦表示学习	分离独立的变化因子, 使潜在表示具有系统性	解耦后的因子灵活组合, 生成新的结构或表示	每个潜在维度对应一个“具有明确物理含义的生成因子”	通过元学习框架, 优化组合函数, 支持单步分解和快速适应
关系建模架构	图卷积网络(GCN)	通过邻域聚合和权重共享系统化处理未知子图组合	通过多层消息传递层级化组合特征	节点决策可部分追溯至多跳邻居特征	通过归纳学习泛化到新节点或子图
	图注意力网络(GAT)	通过标准化注意力计算与多头集成, 构建模块化、分层级的图数据处理框架	动态注意力权重支持局部子结构的灵活识别与组合	注意力系数显式量化节点间重要性	根据节点特征自适应计算关系权重, 无需预设固定结构, 适应异构与演化图数据
	神经符号集成	通过模块化设计和概率生成模型, 构建结构化的推理流程	程序归纳模块支持元素灵活重组以生成新组合	符号表示与逻辑规则提供透明决策路径	通过溯因推理和知识库反馈动态调整组合策略, 适应新问题
学习机制架构	基于优化的元学习(如MAML)	通过双层优化构建标准元学习流程, 确保结构化学习	敏感初始化参数支持快速组合已有能力解决新组合任务	较弱, 参数变化与语义概念缺乏显式关联	通过1~2次梯度更新快速适应新任务环境
	基于度量的元学习(如ProtoNet)	通过原型计算、分类决策、损失优化构建标准小样本学习流程	通过相似性度量实现对新特征组合的分类	决策过程基于与原型距离的比较, 透明且易于理解	通过计算新支持集原型快速适应新类别
	基于数据增强和课程学习的方法	通过数据增强、课程学习、训练监控三模块构建标准化渐进训练框架, 流程结构化	数据增强生成元素新组合, 课程学习调度由易到难的学习顺序	课程学习基于难度函数与阈值规则调度, 符合人类认知逻辑, 增强过程可追溯	课程学习动态调整数据难度, 使模型逐步适应复杂分布

在组合泛化机理方面,递归神经网络无论序列多长,都使用同一套参数(参数共享)按时间步递归处理输入,能利用已学到的通用规则进行动态计算和推理,从而泛化到新的序列上;Transformer的自注意力机制使其能够直接计算序列中任意两个元素之间的关系权重,实现了真正的“全局建模”,模型能通过注意力权重灵活地聚焦于相关部分,从而理解 and 处理全新的组合;模块化神经网络将

复杂的整体功能解耦为多个专门化的、可重复使用的功能模块,根据当前输入动态地选择和组织已有的模块泛化到新的组合场景;解耦表示学习通过约束模型,将数据中的不同变化因素分离到表征空间中相互独立的维度上,通过对这些独立的维度进行组合操作来生成新的表征;GCN通过层级化地聚合邻居节点的信息来生成节点的嵌入表示,实现对未知图结构的推理;GAT通过动态

注意力机制为不同的邻居节点计算不同的权重,从而更精准地实现泛化;神经符号集成结合了神经网络和符号系统的优势,该系统能天然地处理新的概念组合,实现可解释且可靠的组合泛化;基于优化的元学习让模型学会一个“易于适应”的通用表示空间,当面对由已知基础元素组合而成的新任务时,它能极快地调整到最佳状态,实现组合泛化;基于度量的元学习通过特征解耦和原型聚合机制,对于新组合出的类别,模型只需计算其支

持样本的原型位置,查询样本就会根据与这个新原型的距离远近被正确分类,无需重新训练,实现了“按例类比”的泛化;基于数据增强和课程学习的方法中,数据增强通过人工规则或生成模型,系统地创建包含新组合的训练数据,让模型在训练阶段就“见过”各种可能的组合,课程学习则通过从简单样本到复杂样本的循序渐进的训练策略,引导模型先学好基础概念,再逐步学习如何组合它们,两者结合有效提升了模型对未知组合的泛化能力。

表2 基于模型结构划分的组合泛化方法比较

Table 2 Comparative analysis of compositional generalization methods based on model architecture taxonomy

方法	优缺点	复杂度变化	组合泛化机理	适用场景及应用领域举例
基础表征架构	递归神经网络 优点: 时序建模灵活、局部泛化强 缺点: 长程依赖弱、噪声敏感、组合泛化不稳定	低	动态时序计算与参数共享机制	适用场景: 实时控制、短序列语言处理 应用领域举例: 自然语言处理、金融预测、物联网
	Transformer架构 优点: 长程依赖强、结构可扩展、训练稳定 缺点: 分布外泛化弱、数据依赖性强	中	自注意力机制	适用场景: 大规模预训练、跨模态任务、复杂语义组合 应用领域举例: 大语言模型、搜索与广告
	模块化神经网络 优点: 显式组合建模、高维泛化优、训练高效 缺点: 模块设计依赖先验、设计复杂、协调成本高	高	模块功能解耦和动态组合	适用场景: 视觉问答、分步推理、增量学习场景 应用领域举例: 自动驾驶、家用机器人
	解耦表示学习 优点: 可解释性强、数据效率高、迁移性好 缺点: 难以完美解耦、训练不稳定	高	解耦表示学习和组合操作	适用场景: 图像生成与编辑、少样本学习与域适应 应用领域举例: 娱乐与电商、推荐系统
关系建模架构	图卷积网络(GCN) 优点: 适合处理图结构数据、多层堆叠捕捉高阶组合关系 缺点: 对动态图或异构图适应性差、过平滑问题	中	层级化的邻域信息聚合机制	适用场景: 节点分类、分子属性预测 应用领域举例: 推荐系统、欺诈检测与风控、智能问答
	图注意力网络(GAT) 优点: 灵活捕捉重要邻居节点、可处理异质关系 缺点: 图训练效率低、注意力权重缺乏可解释性	高	动态注意力机制	适用场景: 推荐系统、知识图谱补全 应用领域举例: 交通预测、精准推荐、药物研究
	神经符号集成 优点: 兼具感知与推理、可解释性极高、系统性强 缺点: 系统设计复杂、灵活性受限	最高	“神经处理+符号推理”的协同	适用场景: 知识图谱问答、机器人任务规划 应用领域举例: 金融风控、医疗辅助诊断、智能客服
学习机制架构	基于优化的元学习(如MAML) 优点: 任务适应性强、模型无关性、泛化性能突出 缺点: 任务相似性依赖、优化稳定性弱	高	双层优化结构	适用场景: 少样本分类、机器人控制、强化学习 应用领域举例: 工业视觉检测、仿真训练、个性化推荐
	基于度量的元学习(如ProtoNet) 优点: 计算高效、可解释性强、对小样本敏感度低 缺点: 对异常值敏感、忽略类内多样性、依赖嵌入空间质量	低	特征解耦和原型聚合机制	适用场景: 少样本图像分类、增量学习、零样本学习 应用领域举例: 工业异常检测、生物识别与安全认证、医疗影像辅助诊断
	基于数据增强和课程学习的方法 优点: 实现简单、通用性强、鲁棒性强 缺点: 依赖先验知识、可能引入偏差	可变	数据增强和课程学习技术	适用场景: 图像分类、自然语言处理 应用领域举例: 计算机视觉、自动驾驶

总体而言,各类方法在组合泛化能力上呈现互补性,需根据具体任务场景、任务的数据结构、计算资源与泛化需求等进行权衡选择。

3 结束语

本文系统介绍了当前组合泛化领域的相关研究情况,对组合泛化的内涵进行了深入探讨,并从系统性、可组合性、可解释性和动态适应性的角度介绍了组合泛化的特点。基于架构演进的思想将组合泛化方法分为:基础表征架构、关系建模架构和学习机制架构3大类,体现了组合泛化模型结构发展的递进研究脉络。每个大类又分别进行划分,分析了其核心思想和核心步骤,对方法的实现做了详细讲解,并分析了相关应用和优缺点。

综上所述,基于本文对组合泛化领域的研究情况,认为组合泛化未来发展还存在以下问题。

1) 模型静态结构与动态适应性问题

组合泛化面临的核心挑战之一是系统难以动态适应复杂多变的场景。现有方法中,Transformer在长程依赖建模、动态图结构处理或分布外任务泛化中表现不足,GCN对动态图或异构图的适应性弱,限制了组合泛化在时序推理和复杂关系建模中的应用。

为解决这一问题,可以通过引入自适应机制与混合架构。自适应机制能使模型根据实际变化自动优化结构和参数,可通过元学习框架实现,即元学习器通过多任务预训练,提炼跨任务的通用优化策略,使模型仅需少量微调即可高效迁移至新场景。混合架构(如Transformer与图网络融合)可整合全局感知与结构推理能力,提升对动态组合关系的建模效率,以推动开放环境下智能系统的结构化推理能力向人类水平逼近。

2) 模块化设计与先验知识依赖问题

组合泛化在模块化设计和先验知识依赖方面存在较大挑战。模块化神经网络和GCN在设计时依赖先验知识,设计复杂且协调成本高,局部过拟合风险较大。

为解决这一问题,可以采用自动机器学习(automated machine learning, AutoML)技术,自动搜索最优的模块组合和网络结构,减少对先验知识的依赖。其实现技术主要依靠神经架构搜索(neural architecture search, NAS)和端到端管道:神经架构搜索可通过贝叶斯优化或进化算法自动探索最优网络结构;端到端管道可整合模块组合决策。此外,结合图神经网络和注意力机制,利用

图结构数据和关键信息捕捉能力,通过注意力机制自动识别冗余模块,进一步优化模块设计效率。

3) 模型训练效率与高计算复杂度问题

组合泛化方法的训练效率和高计算复杂度成为限制组合泛化能力的重要因素。Transformer架构和GAT的大规模数据训练效率低下,计算复杂度较高,影响模型的实用性和可扩展性。模块化神经网络和基于度量的元学习方法虽然在计算效率上有一定优势,但在处理高维数据和复杂任务时仍面临挑战。

为解决这一问题,可以利用分布式计算和并行训练技术,通过多设备协作加速训练过程,解决大模型训练的计算瓶颈,提高模型的训练效率。同时,采用模型压缩和加速技术,如剪枝、量化和知识蒸馏,降低模型的计算复杂度。

4) 异质关系建模与可解释性问题

组合泛化方法在处理异质关系和可解释性方面存在较大挑战。GAT在处理异质关系时存在注意力权重的语义模糊性的局限性,即GAT的注意力系数仅反映节点特征的相似性,无法明确对应具体关系类型。虽然GAT能够灵活捕捉重要邻居节点和异质关系,但注意力权重缺乏可解释性,影响了模型的透明度和可信度。

为解决这一问题,可以采用可解释性学习的方法,如注意力可视化和特征重要性分析,通过可视化神经网络中注意力权重的分布,揭示模型关注的关键输入区域,识别冗余特征,提升模型的可解释性和透明度。同时,通过探索动态结构生成与因果推理的深度耦合,可突破静态模型的局限,以弥合数据驱动模型与人类认知逻辑的鸿沟。

组合泛化是构成下一代人工智能系统的核心特征。本研究为克服当前人工智能模型的限制提供了新的方向,并为实现更具可组合性的人工智能系统打下了坚实基础。通过构建可以学习和组合多个预训练模型的人工智能系统,充分释放了人工智能系统的潜力,不断提高其性能并有效处理了动态世界中的新任务。通过模型组合和持续学习,未来我们距离充分发挥人工智能的潜力又近了一步。

参考文献:

- [1] MITROU I N, STAMATOPOULOS P. Generative artificial intelligence: models, benefits, dangers and detection of AI-generated text on specialized domains[D]. Athens: National and Kapodistrian University of Athens, 2024: 16.
- [2] STIENNON N, OUYANG L, WU J, et al. Learning to

- summarize with human feedback[C]//Advances in Neural Information Processing Systems 33. Vancouver: Curran Associates, Inc., 2020: 3008–3021.
- [3] BUBECK S, CHANDRASEKARAN V, ELDAN R, et al. Sparks of artificial general intelligence: early experiments with GPT-4[EB/OL]. (2023–03–22)[2025–01–01]. <https://arxiv.org/abs/2303.12712>.
- [4] VIJAYARAGHAVAN P, QUEIBER J F, FLORES S V, et al. Development of compositionality through interactive learning of language and action of robots[J]. *Science robotics*, 2025, 10(98): eadp0751.
- [5] MORRIS M R, SOHL-DICKSTEIN J N, FIEDEL N, et al. Position: levels of AGI for operationalizing progress on the path to AGI[C]//International Conference on Machine Learning. Vienna: PMLR, 2024: 36308–36321.
- [6] FODOR J A, PYLYSHYN Z W. Connectionism and cognitive architecture: a critical analysis[J]. *Cognition*, 1988, 28(1/2): 3–71.
- [7] 柯琳, 杨笑笑, 陈智斌. 一种带泛化性能的动态混合模型求解大范围 TSP 问题[J]. *系统科学与数学*, 2024, 44(1): 31–44.
KE Lin, YANG Xiaoxiao, CHEN Zhibin. A dynamic hybrid model with generalization performance to solve large-scale TSP[J]. *Journal of systems science and mathematical sciences*, 2024, 44(1): 31–44.
- [8] 吴国栋, 秦辉, 胡全兴, 等. 大语言模型及其个性化推荐研究[J]. *智能系统学报*, 2024, 19(6): 1351–1365.
WU Guodong, QIN Hui, HU Quanxing, et al. Research on large language models and personalized recommendation[J]. *CAAI transactions on intelligent systems*, 2024, 19(6): 1351–1365.
- [9] 潘登, 毕晓君. 基于 Transformer 模型的自闭症功能磁共振图像分类[J]. *智能系统学报*, 2025, 20(2): 400–406.
PAN Deng, BI Xiaojun. Classification of functional magnetic resonance images for autism based on Transformer model[J]. *CAAI transactions on intelligent systems*, 2025, 20(2): 400–406.
- [10] YAGCIOGLU S, İNCE O B, ERDEM A, et al. Sequential compositional generalization in multimodal models [C]//Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Kerrville: Association for Computational Linguistics, 2024: 5591–5611.
- [11] KOH P W, SAGAWA S, MARKLUND H, et al. WILDS: a benchmark of in-the-wild distribution shifts[EB/OL]. (2020–12–14)[2025–01–01]. <https://arxiv.org/abs/2012.07421>.
- [12] LAKE B M, BARONI M. Human-like systematic generalization through a meta-learning neural network[J]. *Nature*, 2023, 623(7985): 115–121.
- [13] SINHA S, PREMSRI T, KORDJAMSHIDI P. A survey on compositional learning of AI models: theoretical and experimental practices[EB/OL]. (2024–06–13)[2025–01–01]. <https://arxiv.org/abs/2406.08787>.
- [14] 汤欧, 郑琪. 论符号主义与联结主义人工智能的发展[J]. *哈尔滨学院学报*, 2024, 45(7): 20–24.
- TANG Ou, ZHENG Qi. On the development of symbolism and connectionism artificial intelligence[J]. *Journal of Harbin University*, 2024, 45(7): 20–24.
- [15] 丁林. 人工智能如何走到今天?[N]. *北京科技报*, 2024–10–14(006). DOI:10.28030/n.cnki.nbjkj.2024.000087.
- [16] 刘超, 梁婷婷, 刘小洋, 等. 融合多角度信息和图卷积网络的社交网络节点分类模型[J]. *重庆理工大学学报 (自然科学)*, 2022, 36(5): 147–160.
LIU Chao, LIANG Anting, LIU Xiaoyang, et al. Social networks node classification model based on multi-angle information fusion and graph convolutional networks[J]. *Journal of Chongqing University of Technology (natural science)*, 2022, 36(5): 147–160.
- [17] 李玲, 雷宏友. 结合 MoE 与 Transformer 的生态翻译模型优化研究[J]. *自动化与仪器仪表*, 2025(4): 178–181, 186.
LI Ling, LEI Hongyou. Research on ecological translation model optimization combining MoE and Transformer[J]. *Automation & instrumentation*, 2025(4): 178–181, 186.
- [18] SHIN R, LIN C, THOMSON S, et al. Constrained language models yield few-shot semantic parsers[C]//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Kerrville: Association for Computational Linguistics, 2021: 7699–7715.
- [19] KIPF T, ELSAYED G F, MAHENDRAN A, et al. Conditional object-centric learning from video[EB/OL]. (2021–11–24)[2025–01–01]. <https://arxiv.org/abs/2111.12594>.
- [20] BAHDANAU D, MURTY S, NOUKHOVITCH M, et al. Systematic generalization: what is required and can it be learned? [EB/OL]. (2018–11–30)[2025–01–01]. <https://arxiv.org/abs/1811.12889>.
- [21] 王雨嫣, 廖柏林, 彭晨, 等. 递归神经网络研究综述[J]. *吉首大学学报 (自然科学版)*, 2021, 42(1): 41–48.
WANG Yuyan, LIAO B L, PENG Chen, et al. Research review of recurrent neural networks[J]. *Journal of Jishou University (natural science edition)*, 2021, 42(1): 41–48.
- [22] ZHANG C, PENG Benji, SUN Xintian, et al. From word vectors to multimodal embeddings: techniques, applications, and future directions for large language models [EB/OL]. (2024–11–06)[2025–01–01]. <https://arxiv.org/abs/2411.05036>.
- [23] TU Z, HE F, TAO D. Understanding generalization in recurrent neural networks[C]//International Conference on Learning Representations. Addis Ababa: International Conference on Learning Representations, 2020: 1–17.
- [24] ZHU Jiancong. Comparative study of sequence-to-sequence models: from RNNs to transformers[J]. *Applied and computational engineering*, 2024, 42(1): 67–75.
- [25] CABRERA-LEÓN Y, BÁEZ P G, FERNÁNDEZ-LÓPEZ P, et al. Neural computation-based methods for the early diagnosis and prognosis of Alzheimer’s disease not using neuroimaging biomarkers: a systematic review[J]. *Journal of Alzheimer’s disease*, 2024, 98(3): 793–823.
- [26] SUTSKEVER I, VINYALS O, LE Q V. Sequence to sequence learning with neural networks[C]//Proceedings of

- the 28th International Conference on Neural Information Processing Systems-Volume 2. New York: ACM, 2014: 3104–3112.
- [27] GREFF K, SRIVASTAVA R K, KOUTNÍK J, et al. LSTM: a search space odyssey[J]. *IEEE transactions on neural networks and learning systems*, 2017, 28(10): 2222–2232.
- [28] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems 30. Long Beach: Curran Associates, Inc., 2017: 5998–6008.
- [29] RUSSIN J, JO J, O'REILLY R, et al. Compositional generalization by factorizing alignment and translation [C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop. Online: ACL, 2020: 313–327.
- [30] BHARADHWAJ H, VAKIL J, SHARMA M, et al. Robo-Agent: generalization and efficiency in robot manipulation via semantic augmentations and action chunking [C]//2024 IEEE International Conference on Robotics and Automation. Yokohama: IEEE, 2024: 4788–4795.
- [31] 陈金磊. 基于聚类分析的模块化神经网络研究与应用 [D]. 郑州: 郑州大学, 2019: 1–54.
CHEN Jinlei. Research and application of modular neural network based on cluster analysis[D]. Zhengzhou: Zhengzhou University, 2019: 1–54.
- [32] WANG Liyuan, ZHANG Xingxing, LI Qian, et al. Incorporating neuro-inspired adaptability for continual learning in artificial intelligence[J]. *Nature machine intelligence*, 2023, 5(12): 1356–1368.
- [33] LIU Ziming, KHONA M, FIETE I R, et al. Growing brains: co-emergence of anatomical and functional modularity in recurrent neural networks[EB/OL]. (2023–10–11)[2025–01–01]. <https://arxiv.org/abs/2310.07711>.
- [34] FEDUS W, ZOPH B, SHAZEER N. Switch transformers: scaling to trillion parameter models with simple and efficient sparsity[J]. *Journal of machine learning research*, 2022, 23(1): 5232–5270.
- [35] ANDREAS J, ROHRBACH M, DARRELL T, et al. Neural module networks[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 39–48.
- [36] ZHENG Yihan, WEN Zhiqian, TAN Mingkui, et al. Modular graph attention network for complex visual relational reasoning[M]//Computer Vision. Cham: Springer International Publishing, 2021: 137–153.
- [37] WANG Xin, CHEN Hong, TANG Si'ao, et al. Disentangled representation learning[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2024, 46(12): 9677–9696.
- [38] ZHENG Hao, LAPATA M. Disentangled sequence to sequence learning for compositional generalization[C]//Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics. Kerrville: Association for Computational Linguistics, 2022: 4256–4268.
- [39] WANG Jindong, LAN Cuiling, LIU Chang, et al. Generalizing to unseen domains: a survey on domain generalization[J]. *IEEE transactions on knowledge and data engineering*, 2023, 35(8): 8052–8072.
- [40] 谢娟英, 张建宇. 图卷积神经网络综述[J]. 陕西师范大学学报 (自然科学版), 2024, 52(2): 89–101.
XIE Juanying, ZHANG Jianyu. The review of the graph convolutional neural networks[J]. *Journal of Shaanxi Normal University (natural science edition)*, 2024, 52(2): 89–101.
- [41] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[EB/OL]. (2016–09–09)[2025–01–01]. <https://arxiv.org/abs/1609.02907>.
- [42] HAMILTON W L, YING R, LESKOVEC J. Inductive representation learning on large graphs[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM, 2017: 1025–1035.
- [43] PAN H, FANG Y, HUANG C, et al. GCNXSS: an attack detection approach for cross-site scripting based on graph convolutional networks[J]. *KSII transactions on internet and information systems*, 2022, 16(12): 4008–4023.
- [44] 吴琳, 许茹玉, 粟兴旺, 等. 基于结构误差的图卷积网络 [J]. 计算机应用研究, 2023, 40(1): 155–159.
WU Lin, XU Ruyu, SU Xingwang, et al. Graph convolutional networks based on structural errors[J]. *Application research of computers*, 2023, 40(1): 155–159.
- [45] PASA L, NAVARIN N, SPERDUTI A. Polynomial-based graph convolutional neural networks for graph classification[J]. *Machine learning*, 2022, 111(4): 1205–1237.
- [46] 吴誉兰, 舒建文. 基于自适应差异化图卷积的图注意力网络表示学习算法[J]. *现代电子技术*, 2025, 48(2): 51–54.
WU Yulan, SHU Jianwen. Graph attention network representation learning algorithm based on adaptive differentiation graph convolution[J]. *Modern electronics technique*, 2025, 48(2): 51–54.
- [47] ZHOU Zhiyuan, YIN Yueming, HAN Hao, et al. ProAffinity-GNN: a novel approach to structure-based protein-protein binding affinity prediction via a curated data set and graph neural networks[J]. *Journal of chemical information and modeling*, 2024, 64(23): 8796–8808.
- [48] POURSAEED O, JIANG Tianxing, YANG H, et al. Robustness and generalization via generative adversarial training[C]//2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2022: 15691–15700.
- [49] ZHANG Zhi, LI Weijian, LIU Han. Multivariate time series forecasting by graph attention networks with theoretical guarantees[C]//International Conference on Artificial Intelligence and Statistics. Valencia: PMLR, 2024.
- [50] BRODY S, ALON U, YAHAV E. How attentive are graph attention networks? [EB/OL]. (2021–05–30)[2025–01–01]. <https://arxiv.org/abs/2105.14491>.
- [51] YU Dongran, YANG Bo, LIU Dayou, et al. A survey on neural-symbolic learning systems[J]. *Neural networks*, 2023, 166: 105–126.
- [52] LI Qing, ZHU Yixin, LIANG Yitao, et al. Neural-symbolic recursive machine for systematic generalization [EB/OL]. (2022–10–04)[2025–01–01]. <https://arxiv.org/>

- [abs/2210.01603](https://arxiv.org/abs/2210.01603).
- [53] GUPTA T, KEMBHAVI A. Visual programming: compositional visual reasoning without training[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2023: 14953–14962.
- [54] SATO S, SATO I. Can test-time computation mitigate reproduction bias in neural symbolic regression? [EB/OL]. (2025–05–28)[2025–06–01]. <https://arxiv.org/abs/2505.22081>.
- [55] HOSPEDALES T M, ANTONIOU A, MICAELLI P, et al. Meta-learning in neural networks: a survey[J]. IEEE transactions on pattern analysis and machine intelligence, 2022, 44(9): 5149–5169.
- [56] FALLAH A, MOKHTARI A, OZDAGLAR A. Generalization of model-agnostic meta-learning algorithms: recurring and unseen tasks[EB/OL]. (2021–02–07)[2025–01–01]. <https://arxiv.org/abs/2102.03832>.
- [57] DEB B, AWADALLAH A H, ZHENG Guoqing. Boosting natural language generation from instructions with meta-learning[C]//Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2022: 6792–6808.
- [58] VETTORUZZO A, BOUGUELIA M R, VANSCHOREN J, et al. Advances and challenges in meta-learning: a technical review[J]. IEEE transactions on pattern analysis and machine intelligence, 2024, 46(7): 4763–4779.
- [59] FINN C, ABBEEL P, LEVINE S. Model-agnostic meta-learning for fast adaptation of deep networks[C]//Proceedings of the 34th International Conference on Machine Learning. New York: ACM, 2017: 1126–1135.
- [60] WANG Zhenhailong, YU Hang, LI Manling, et al. Rethinking task sampling for few-shot vision-language transfer learning[C]//Proceedings of the First Workshop on Performance and Interpretability Evaluations of Multimodal, Multipurpose, Massive-Scale Models. Online: ACL, 2022: 7–14.
- [61] NICHOL A, ACHIAM J, SCHULMAN J. On first-order meta-learning algorithms[EB/OL]. (2018–03–04)[2025–01–01]. <https://arxiv.org/abs/1803.02999>.
- [62] 马怡琳. 基于改进 ProtoNet 的小样本关系抽取方法研究[D]. 天津: 河北工业大学, 2021: 1–67.
- MA Yilin. Research on few-shot relation extraction method base on improved ProtoNet[D]. Tianjin: Hebei University of Technology, 2021: 1–67.
- [63] CHEN CHAOFAN, LI OSCAR, TAO CHAOFAN, et al. This looks like that: deep learning for interpretable image recognition[C]//Proceedings of the 33rd International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2019: 8930–8941.
- [64] SNELL J, SWERSKY K, ZEMEL R S. Prototypical networks for few-shot learning[EB/OL]. (2017–03–15)[2025–01–01]. <https://arxiv.org/abs/1703.05175>.
- [65] MANNEKOTE A. Towards a neural era in dialogue management for collaboration: a literature survey[EB/OL]. (2023–07–18)[2025–01–01]. <https://arxiv.org/abs/2307.09021>.
- [66] GÜLCÜ A, ALKAN M. Az örnekle öğrenme problemleri için MAML ve ProtoNet algoritmalarının incelenmesi[J]. European journal of science and technology, 2021(21): 113–121.
- [67] CHANG H, PARK J, CHO H, et al. The coverage principle: a framework for understanding compositional generalization[EB/OL]. (2025–05–26)[2025–06–01]. <https://arxiv.org/abs/2505.20278>.
- [68] ZHANG Baoquan, LI Xutao, YE Yunming, et al. Prototype completion with primitive knowledge for few-shot learning[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2021: 3753–3761.
- [69] CHOI J, KIM Y. Colorful cutout: enhancing image data augmentation with curriculum learning[EB/OL]. (2024–03–29)[2025–06–01]. <https://arxiv.org/abs/2403.20012>.
- [70] CHEN P, LIU S, ZHAO H, et al. Gridmask data augmentation[EB/OL]. (2024–03–26)[2025–06–01]. <https://arxiv.org/abs/2001.04086>.
- [71] AKYÜREK E, AKYÜREK A F, ANDREAS J. Learning to recombine and resample data for compositional generalization[EB/OL]. (2020–10–08)[2025–06–01]. <https://arxiv.org/abs/2010.03706>.
- [72] QIU Linlu, SHAW P, PASUPAT P, et al. Improving compositional generalization with latent structure and data augmentation[C]//Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Kerrville: Association for Computational Linguistics, 2022: 4341–4362.

作者简介:



李昌原, 硕士研究生, 主要研究方向为组合泛化。E-mail: 1744014975@qq.com。



程恺, 副教授, 主要研究方向为智能任务规划、数据分析挖掘。获军队科技进步奖二等奖 2 项、三等奖 3 项、指控学会二等奖 1 项。发表学术论文 60 余篇, 出版学术专著 2 部。E-mail: chengkai911@126.com。



郝文宁, 教授, 主要研究方向为大规模高维数据压缩与作战效能评估。发表国际期刊论文 20 余篇。E-mail: hwnbox@aeu.edu.cn。

[责任编辑: 孟玮]