数字孪生架构下基于GAN增强的多智能体深度强化学习边缘推理与异构资源协同优化

袁晓铭[1),2)] 田汉森[1)] 黄锟达[1)] 邓庆绪[1)] 康嘉文[3)]
李长乐[2)] 段续庭[4)]

[1)](东北大学秦皇岛分校河北省海洋感知网络与数据处理重点实验室 河北 秦皇岛 066004)
[2)](西安电子科技大学空天地一体化综合业务网全国重点实验室 西安 710071)
[3)](广东工业大学自动化学院 广州 510006)
[4)](北京航空航天大学车路一体智能交通全国重点实验室 北京 100191)

摘 要 边缘侧大模型应用正成为推动智能健康、智慧城市等领域智能化与数字化进程的关键驱动力。然而,大模型海量智能任务异构性和高动态网络的不可预测性,使得边缘设备有限的算力资源难以满足复杂推理任务对高效且可靠服务质量(Quality of Service, QoS)的需求。因此本文提出了一种基于生成对抗网络(Generative Adversarial Network, GAN)增强的多智能体深度强化学习(Multi-Agent Deep Reinforcement Learning, MADRL)的边缘推理与异构资源协同优化方法,以实现数字孪生(Digital Twin, DT)驱动的边缘侧大模型赋能系统中异构资源的动态负载均衡,确保推理任务高效性与可靠性。首先,本文构建并分析了DT驱动的边缘侧大模型系统中的物理网络层和孪生网络层,并采用GAN实现对物理实体的孪生映射,从而对海量异构边缘数据进行分布式处理、生成与优化。接着,利用MADRL算法来对系统中的异构资源进行综合量化与协同优化,并将边缘推理数据反馈至MADRL算法中以减少集中式训练过程中的数据通信开销。同时,借助于联邦学习,该架构能够实现多方知识共享,从而有效提升模型训练速度与性能。最后,仿真结果表明,该算法能够在动态复杂大模型赋能边缘系统环境中有效降低推理任务的时延和能耗,充分利用有限的系统资源,确保推理任务的高效性,并提升智能服务的质量。

关键词 边缘侧大模型;数字孪生;移动边缘计算;多智能体深度强化学习;生成对抗网络;联邦学习
中图法分类号 TP393 **DOI号** 10.11897/SP.J.1016.2025.01763

A GAN-Enhanced Multi-Agent Deep Reinforcement Learning for Digital Twin-Driven Edge Inference and Heterogeneous Resource Co-Optimization

YUAN Xiao-Ming[1),2)] TIAN Han-Sen[1)] HUANG Kun-Da[1)] DENG Qing-Xu[1)]
KANG Jia-Wen[3)] LI Chang-Le[2)] Duan Xu-Ting[4)]

[1)](Hebei Key Laboratory of Marine Perception Network and Data Processing, Northeastern University, Qinhuangdao, Hebei 006004)
[2)](State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071)
[3)](School of Automation, Guangdong University of Technology, Guangzhou 510006)
[4)](State Key Lab of Intelligent Transportation System, Beihang University, Beijing 100191)

Abstract The deployment of large models at the edge has emerged as a pivotal enabler for the intelligent and digital transformation across various domains, including smart healthcare and urban

收稿日期:2025-02-14;在线发布日期:2025-05-19。本课题得到车路一体智能交通全国重点实验室开放基金课题(2024-B008)、国家自然科学基金面上项目(62371116)资助。**袁晓铭**(通信作者),博士,副教授,中国计算机学会(CCF)会员,主要研究领域为5G/6G网络关键技术和边缘智能等。E-mail:yuanxiaoming@neuq.edu.cn。**田汉森**,硕士,主要研究领域为计算卸载、多智能体强化学习。**黄锟达**,硕士,主要研究领域为数字孪生、资源调度。**邓庆绪**,博士,教授,中国计算机学会(CCF)杰出会员,主要研究领域为实时嵌入式系统和智能物联网。**康嘉文**,博士,教授,中国计算机学会(CCF)会员,主要研究领域为元宇宙、物联网等。**李长乐**,博士,教授,主要研究领域为智能交通系统和无线传感器网络等。**段续庭**,博士,教授,主要研究领域为车联网与智能协作理论等。

中国知网 https://www.cnki.net

systems. However, the heterogeneity of massive intelligent tasks and the unpredictability of high-dynamic networks pose significant challenges in the limited computational resources of edge devices to meet the demands of complex inference tasks for efficient and reliable Quality of Service (QoS). Therefore, this paper proposes an edge inference and heterogeneous resource collaborative optimization method based on Generative Adversarial Network (GAN)-enhanced Multi-Agent Deep Reinforcement Learning (MADRL), aiming to achieve dynamic load balancing of heterogeneous resources in Digital Twin (DT)-driven edge large model-enabled systems, ensuring the efficiency and reliability of inference tasks. First, this paper analyzes the physical network layer and twin network layer of the DT-driven edge large model system, and leverages GAN for twin mapping, enabling distributed processing, generation, and optimization of massive heterogeneous data. Next, the MADRL algorithm is applied to the comprehensive quantification and collaborative optimization of heterogeneous resources, with the edge inference data being fed back into the MADRL algorithm to reduce the data communication overhead during centralized training. Meanwhile, the framework utilizes federated learning, allowing for multi-party knowledge sharing, effectively improving model training speed and performance. Finally, simulation results demonstrate that the proposed algorithm reduces inference task delay and energy consumption, optimally utilizes system resources, and improves intelligent service quality in dynamic edge environments.

Keywords edge-based large models; digital twin; mobile edge computing; multi-agent deep reinforcement learning; generative adversarial network; federated learning

1 引 言

近年来，边缘侧大模型在提升智能推理服务效率、加速数字化进程并增强系统实时响应能力等方面发挥着至关重要的作用[1]。在可靠服务质量(Quality of Service，QoS)保障下，边缘侧大模型能够有效应对数字化转型过程中计算资源短缺和分布不均问题。随着大模型在医疗、交通等领域广泛应用，基于边缘侧大模型的智能系统在数据分析和决策支持方面展现出巨大潜力。

然而，边缘侧大模型的复杂性及其对计算资源的高需求为智能系统带来了诸多挑战。首先，在动态且高度异构的移动边缘计算(Mobile Edge Computing，MEC)环境中，系统必须高效处理并利用海量异构数据，从而支持大模型推理任务。其次，亟需对推理过程中所涉及的边缘侧计算与通信资源进行有效管理与优化，以保障智能服务高效性与稳定性。此外，还需在降低时延和能耗的同时，提升个性化服务体验和任务处理效率，以满足智能服务持续发展的需求。

为了解决边缘侧大模型中的关键挑战，数字孪生技术[2](Digital Twin，DT)与MEC[3]的融合为上述的问题提供了具有广阔前景的解决方案。DT通过实时数据感知和信息交互，能够高效提取多模态信息特征，虚拟映射边缘侧大模型系统中物理实体和服务功能，从而实现智能服务全生命周期的建模、仿真与预测。同时，基于DT构建的个性化虚拟模型，能够精准模拟边缘计算用户的个性化需求及场景差异，有效弥补大模型在多样化边缘侧应用中训练数据不足的问题，显著提升其泛化能力与场景适应性。

与此同时，MEC所提供的实时智能边缘推理服务显著提升了数据处理与响应的速度与效率。面对边缘推理任务对计算资源的高需求，DT与MEC的协同机制能够有效模拟并优化系统资源的分配与调度，进而显著增强边缘侧的数据分析与决策支持能力，为大模型驱动下的实时、可靠的边缘侧智能推理服务提供了坚实保障。

因此，本文提出了一种DT驱动的大模型赋能边缘侧系统，旨在确保边缘侧智能推理任务高效性与可靠性。该系统架构聚焦于两个关键问题：其一是如何实现DT的虚拟映射，即如何精确地将物理实体和服务功能转化为数字模型，并利用这些数字模型为边缘推理提供高质量的数据支撑；其次，如何处理并利用海量异构数据以及DT虚拟映射，优化

中国知网 https://www.cnki.net

推理任务的资源调度与决策策略，以高效利用系统资源，确保推理任务的及时响应与高效执行。目前，虽然已有大量研究[4-6]集中探讨边缘侧大模型场景中资源协同调度问题，尤其是通过利用深度强化学习算法（Deep Reinforcement Learning，DRL）实现资源调度策略的优化[7-9]，进而提升边缘侧大模型系统的服务性能。然而，尽管DT在解决边缘侧大模型挑战中展现出巨大潜力，目前研究大多侧重于DT理论框架的构建与分析，对于如何结合DT模型优化算法，尤其是在推理任务资源调度中的实际应用，仍缺乏深入的研究与实践探索。

首先，针对动态边缘侧大模型多用户场景下推理任务协同处理问题，基于分布式执行集中式训练框架的多智能体深度强化学习算法（Multi-Agent DRL，MADRL）[10]为系统的协同优化提供了有效途径。MADRL支持各智能体根据当前所感知的局部状态信息自主执行决策，并通过引入时序神经网络（Recurrent Neural Network，RNN）结构有效缓解多智能体环境中的感知受限问题，从而增强系统对环境动态变化的适应能力。此外，集中式训练机制可在策略学习阶段整合全局信息，实现跨智能体的协同优化，显著增强资源调度与推理任务分配的协调性与执行效率。

其次，生成对抗网络（Generative Adversarial Network，GAN）能够通过有效处理并利用智能体所感知的多模态状态信息，生成高质量且逼真的数据，从而实现边缘侧大模型系统的虚拟映射。相较于扩散模型，GAN在生成速度上更快且计算资源消耗更低，使其更适用于边缘侧大模型推理任务，尤其在资源受限场景下更具优势。基于GAN的DT虚拟映射不仅能丰富DRL训练所需的数据集，还能减少多智能体算法信息通信需求，进而提升MADRL智能体模型的泛化能力和整体性能。此外，DRL训练过程中产生的反馈信息可用于持续优化DT的映射模型，从而提升推理任务预测的准确性与时效性，进一步优化智能任务的服务质量，确保系统决策更具个性化、精准性与实时响应能力。

与此同时，引入个性化联邦学习（Federated Learning，FL）[11]机制，有助于在保护数据隐私的前提下实现智能体间模型的高效知识共享，从而加速GAN映射模型与MADRL智能体策略的联合训练与优化。FL不仅能够有效降低数据传输、缓存与处理带来的系统负担，还可以显著减少隐私泄露的风险[12]。

综上，本文提出了一种条件生成对抗网络（Conditional GAN，CGAN）增强的多智能体深度确定性策略梯度（Multi-Agent Deep Deterministic Policy Gradient，MADDPG）算法，以支持DT驱动边缘侧大模型系统的边缘推理与资源协同优化。该算法面向动态且复杂的MEC环境，能够感知并整合多源异构状态信息，协同生成高效边缘推理策略与资源调度方案，并保障个性化智能服务的安全性、高效性与可靠性，同时增强系统适应性与泛化能力。具体而言，本文贡献总结如下：

（1）DT驱动的边缘侧大模型架构

构建了一个基于DT的边缘侧智能架构，该架构通过就近计算服务和高质量数据支撑以构建全局知识共享平台，实现多智能体间的信息同步与数据交互，有效支持智能体的高效协作，进而显著提升智能服务质量与系统整体性能。

（2）基于CGAN的DT虚拟映射

利用智能体历史经验数据，CGAN提取并融合多源异构的感知信息，构建智能体的DT虚拟映射模型。通过CGAN生成虚拟数据，可以丰富计算卸载算法的训练数据，并且减少集中式训练中涉及隐私数据传输，从而提升DRL算法泛化能力和模型性能。

（3）基于知识共享的MADDPG计算卸载

采用MADDPG算法实现多用户低时延、低能耗、高可靠性的计算卸载。各智能体根据本地感知的状态信息独立进行决策，并且借助知识共享平台通过DT虚拟映射同步信息与FL跨节点模型参数共享，以加速模型训练并缓解感知受限问题。

本文其余部分的结构安排如下：第2节介绍相关工作；第3节构建DT驱动的边缘侧大模型资源协同优化系统模型；第4节提出联合优化时延与能耗的系统成本最小化问题，并构建与分析相应的DRL模型；第5节详细描述所设计的核心算法流程与机制；第6节通过仿真实验进行对比分析，评估所提算法性能表现；第7节总结本文，并提出未来改进方向。

2　相关工作

目前，针对边缘侧大模型系统中的边缘推理与资源协同优化问题，已有大量研究工作取得了显著进展。Deng等人[13]以最小化任务处理时延为目标，设计了一种用于多用户MEC系统中时延敏感计算任务的自主部分卸载机制。Li等人[4]提出了一种无

中国知网　https://www.cnki.net

人机辅助MEC中按需部署多无人机的协作计算卸载策略，通过预测用户轨迹并在最小路径损耗条件下动态部署多架无人机，以提升系统响应效率。Hou等人[5]研究了一种面向时延容忍与时延敏感任务的分层计算卸载方法，基于MADDPG算法，实现更快的任务处理、动态资源分配以及更低系统成本。Zhang等人[14]设计了一种结合元强化学习与DRL的计算卸载算法，其中元强化学习用于增强模型的迁移能力，DRL则借助多个并行神经网络从历史卸载场景中进行策略学习。Liu等人[15]则提出一种分布式在线不稳定性感知计算卸载启发式算法，最大限度地减少服务中断时间。此外，还有部分研究致力于基于博弈论的卸载与资源分配策略的设计[16-17]，通过建模多用户之间的策略互动过程，实现系统资源的优化调度。

在数据隐私安全保护方面，现有研究主要集中于区块链技术与FL融合应用[18-19]。Liu等人[16]提出一种面向区块链驱动的MEC数据共享框架，在保障数据隐私前提下实现多方数据可信交互。Zhang等人[20]提出了一种基于联邦双深度Q网络的任务卸载和资源分配算法。其中DRL在本地设备上进行训练，FL用于聚合本地模型参数并更新全局策略，从而有效降低任务执行时延与能耗。Wang等人[21]基于联邦区块链架构，设计了一种MEC资源分配机制，以最低服务成本实现最优服务质量。在具体场景中，尹等人[22]提出一种异步鲁棒FL方法，在实现车辆数据隐私保护的同时，提高模型协同训练的效率。Yuan等人[23]将FL与广义学习系统相结合，有效提升了车联网下的数据共享效率与精度。Seid等人[24]提出了一种多智能体联邦强化学习算法，用于联合优化任务卸载与资源分配，以实现最低时延与能耗，并保障系统QoS。Yang等人[25]提出一种联合任务调度与资源管理的无人机FL方案，有效提升了数据隐私保护能力与任务处理效率。

然而，在边缘侧的大模型智能服务系统中，尤其是在多智能体场景下，算法需要具备良好的分布式执行能力和高效的数据处理能力。在现有研究中，一方面，许多工作主要聚焦于多用户的协作机制，通常将整个系统简化为单智能体进行研究，忽视了环境中大规模、复杂智能体之间的决策竞争问题；另一方面，集中式的训练过程在多智能体场景下常常面临数据同步延迟以及通信瓶颈等挑战，难以满足MEC环境下对低时延和高隐私的需求。因此，本文考虑结合CGAN生成虚拟映射数据与FL技术，通过CGAN实现高质量的虚拟映射，强化智能体之间的知识迁移与共享；同时利用FL保障数据隐私安全，减少通信开销，以构建一个高效的分布式协同学习平台。该方法能够有效提升推理任务的卸载效率与灵活性。

另外，近年来越来越多的工作关注将DT与MEC相结合[26-28]，以提升系统的智能感知与决策能力。Zhang等人[27]提出一种DT驱动的协作MEC智能任务卸载框架，其中DT用于构建物理-虚拟空间的映射模型，从而优化任务卸载策略，实现更高效的资源利用。Zhao等人[28]设计了一种基于DT与DRL的智能部分卸载方案，其中引入基于卸载聚类结果的DT反馈机制，动态调整算法参数，以提升虚拟空间与物理空间的协同效率。He等人[29]提出了一种将DT和MEC技术与FL框架相结合的架构，并通过DRL算法来解决联合优化计算卸载和资源分配问题，并有效缓解FL中的落后者效应。Yuan等人[30]构建了一种DT驱动的车辆任务卸载和智能反射面配置框架，通过DT实现物理运行环境的实时数据收集和数字化表示，以更好地支持决策。为进一步优化系统性能，该工作还提出了一种融合DRL与迁移学习的两阶段优化算法，旨在降低整体系统处理时延与能耗。

然而，现有的大多数DT驱动的边缘侧大模型架构研究仍停留在理论层面，尚未实现与实际算法的深度融合。尤其在边缘侧大模型智能体协同场景下，如何高效构建DT虚拟映射并提升计算卸载执行效率，仍是亟待解决的关键问题。为此，本文将利用CGAN实现DT虚拟映射，并将生成的虚拟映射数据作为支撑，以进一步提升边缘侧大模型推理任务计算卸载算法的适应性和效率。此外，DT架构可为多智能体系统构建全局知识共享平台，进一步优化整体算法性能，提升多智能体资源协同优化的智能化水平与系统响应能力。

3 系统模型

本节将围绕边缘侧大模型多用户场景，构建一个面向资源协同优化的系统模型。该模型融合DT与MEC技术，能够实时感知系统状态，精准模拟任务特征与资源动态，支撑高效的计算卸载与资源调度策略，从而实现边缘推理过程中对高质量数据与实时计算服务的双重支持。在此基础上，针对

中国知网 https://www.cnki.net

多用户、多边缘节点的动态环境，系统进一步深入分析推理任务分配与资源协同管理中的关键优化策略，为构建高效、可靠的边缘智能服务提供理论与方法支撑。

系统模型分为两个层次，如图1所示，分别为表示实际应用场景的物理网络层与实现DT虚拟映射的孪生网络层。其中，物理网络层对应边缘侧资源协同优化真实环境，MEC服务器为终端用户提供就近推理服务，以满足推理任务低时延、低能耗的QoS要求。孪生网络层则通过创建虚拟模型，实时模拟和映射物理实体以及计算卸载过程，辅助进行性能优化与策略反馈，以提升推理任务的处理效率与智能决策能力。上述两个层次将在接下来的两个小节中进行详细阐述。

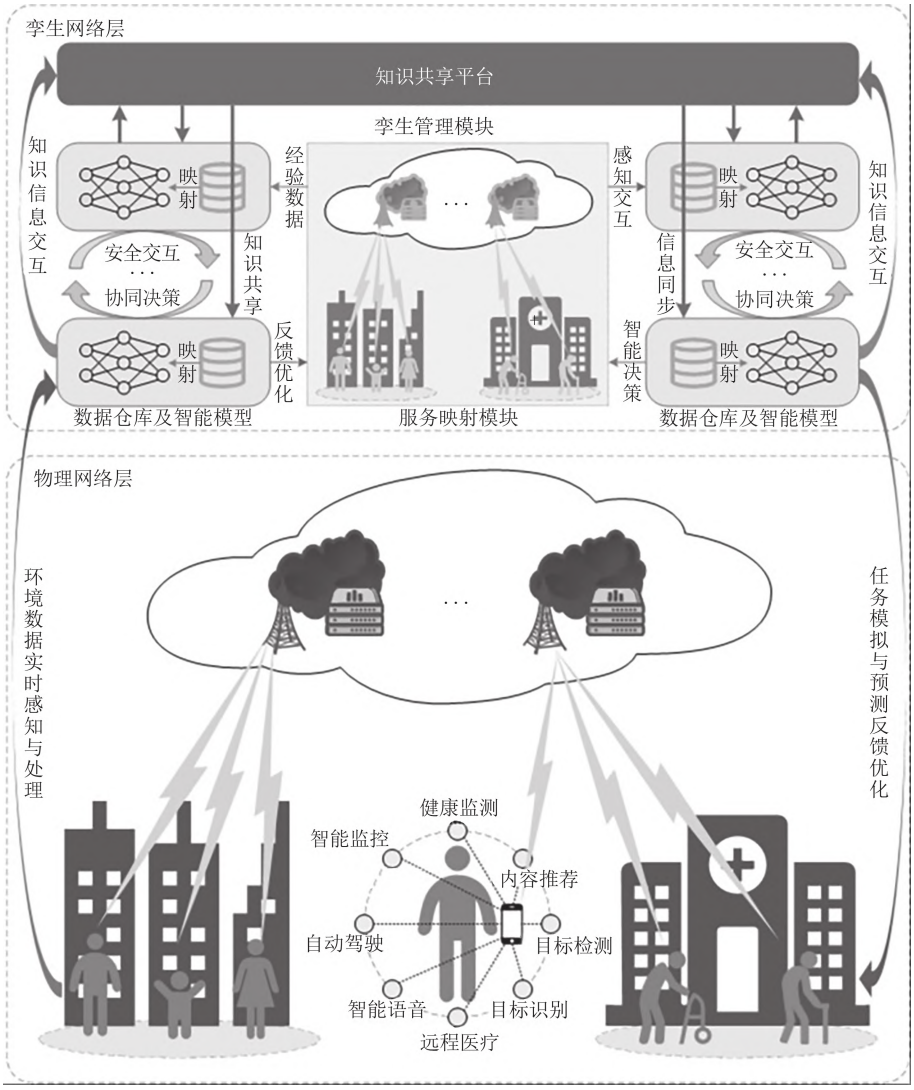

图1　系统模型图

3.1　物理网络层

物理网络层负责物理实体之间的数据感知、智能推理任务的生成与处理，以及支持实时高效任务响应的系统资源动态分配。其组成涵盖边缘智能服务中涉及的所有关键物理实体，包括用户终端设备、MEC服务器，以及支持实时计算服务的资源协同与优化过程。在物理网络层中，首先每个用户$i\in\mathcal{U}$的感知终端能够持续采集环境数据并实时生成N_i^{t}个计算密集型智能推理任务，定义用户集合及其任务集合分别为$\mathcal{U}=\{i|1,2,\cdots,N^{\mathrm{u}}\}$和$\mathcal{T}_i=\{j|1,2,\cdots,N_i^{\mathrm{t}}\}$。为了满足低时延和低能耗的QoS要求，这些推理任务可以在本地直接处理，或者卸载至合适的MEC服务器，以缓解本地计算负载，MEC服务器的集合记为$\mathcal{M}=\{k|1,2,\cdots,N^{\mathrm{m}}\}$。对于用户$i$的任意推理任务$task_{i,j}$，$f_{i,j}$可以表示为本地设备或者MEC服务器$k\in\mathcal{M}$为其分配的计算资源。据此，任务$task_{i,j}$的

中国知网　https://www.cnki.net

计算时延 $T_{i,j}^{c}$ 和能耗 $E_{i,j}^{c}$ 可以分别表示为 $T_{i,j}^{c}=X_{i,j}/f_{i,j}$ 和 $E_{i,j}^{c}=\kappa f_{i,j}^{2}X_{i,j}$，其中 $X_{i,j}$ 表示任务计算复杂度，κ 表示能耗相关系数。

其次，针对需卸载至边缘侧处理的推理任务 $task_{i,j}$，需综合考虑无线传输环境以及目标MEC服务器可用计算与通信资源，以实现高效的任务卸载与资源分配。设MEC服务器 k 总带宽为 B_k，对于每个选择将任务卸载至服务器 k 的用户 $i\in\mathcal{U}_k$ 为其分配的通信资源为 b_i。同时卸载至同一MEC服务器的不同任务，用户设备 i 采用频分多址进行并行传输，其传输功率记为 P_i^{tr}。根据香农公式，$task_{i,j}$ 的最大传输速率可以表示为

$$rate_{i,k}=b_{i,j}\log_2\left(1+\frac{g_{i,k}P_i^{tr}d_{i,k}^{-\lambda}}{b_{i,j}\mathcal{N}_0}\right) \tag{1}$$

在上述公式中，$g_{i,k}$ 和 $d_{i,k}$ 分别表示MEC服务器 k 与用户 $i\in\mathcal{U}_k$ 之间的信道增益和距离。λ 和 $\mathcal{N}_0$ 分别表示路径损耗和高斯噪声。因此，可以得到卸载推理任务 $task_{i,j}$ 的传输时延 $T_{i,j}^{tr}=D_{i,j}/rate_{i,k}$ 以及能耗 $E_{i,j}^{tr}=P_i^{tr}T_{i,j}^{tr}$，其中 $D_{i,j}$ 表示其任务数据量。

综上所述，定义 $s_{i,j}\in\{0\}\cup\mathcal{M}$ 为推理任务 $task_{i,j}$ 的卸载策略，其中当 $s_{i,j}=0$，该任务将在用户 i 的智能设备上进行本地处理。因此，$task_{i,j}$ 的总处理时延和能耗可以分别表示为如下公式：

$$T_{i,j}=\left(s_{i,j}\in\mathcal{M}\right)T_{i,j}^{tr}+T_{i,j}^{c} \tag{2}$$

$$E_{i,j}=\left(s_{i,j}\in\mathcal{M}\right)E_{i,j}^{tr}+E_{i,j}^{c} \tag{3}$$

3.2 孪生网络层

孪生网络层旨在对物理网络层中的物理实体及其计算卸载过程进行高保真映射，以提升系统的数据分析能力，实现更快速、准确且智能化的卸载决策以及资源分配策略。通过对物理系统的实时映射与反馈，孪生网络层能够辅助边缘侧智能体更加高效地进行计算与通信资源协同优化，最大化系统资源利用率，并显著增强智能体处理复杂与动态任务场景的自适应能力。

在孪生网络层中，需要分别对每个用户及其智能设备以及MEC服务器进行建模映射，以构建完整的DT系统，其对应的DT孪生实体集合分别表示为 $\hat{\mathcal{U}}$ 和 $\hat{\mathcal{M}}$。为全面映射资源协同优化过程，还需要对推理任务卸载过程中所依赖的无线通信环境进行建模，综合考虑系统中的通信状态、设备运行状态以及空间位置信息等多维动态因素。

首先，孪生网络层需构建对节点 $l\in\mathcal{U}\cup\mathcal{M}$（包括智能设备以及MEC服务器）的精确映射。由于真实物理环境感知误差与动态变化的存在，孪生实体可能与物理实体存在偏差。因此，引入计算资源分配偏差 $\Delta f_{i,j}$ 与通信带宽分配 Δb_i 偏差来衡量节点设备在DT系统中的映射与实际物理设备之间的差异。每个DT映射节点可表示为

$$DT_l=\mathrm{DT}\left((b_i,\Delta b_i),\left\{\left(f_{i,j},\Delta f_{i,j}\right)\right\}_{i\in\mathcal{T}_l}\right) \tag{4}$$

其中，当 $i=l$ 时表示 $task_{i,j}$ 在本地处理，此时为其分配的带宽 b_i 和带宽偏差 Δb_i 都为0。

其次，由于MEC网络具有复杂的设备交互与动态通信环境，即使具备完整的节点映射结构，孪生空间中的通信状态仍可能与物理层存在一定误差。为了准确描述这种偏差，可通过用户设备 i 与其目标MEC服务器 k 之间的信道增益偏差 $\Delta g_{i,k}$ 及其传输距离偏差 $\Delta d_{i,k}$ 进行建模，从而量化孪生网络层中资源协同优化所依赖的映射通信环境的映射精度。具体映射关系可表示为

$$DTN=\mathrm{DT}\left(\left\{\left(g_{i,k},\Delta g_{i,k}\right),\left(d_{i,k},\Delta d_{i,k}\right)\right\}_{i\in\mathcal{U},k\in\mathcal{M}}\right) \tag{5}$$

进一步地，对于每个用户以及MEC服务器，期望其在孪生空间中的映射实体能够提供与物理网络层相一致的服务能力。针对每个映射的推理任务 $task'_{i,j}$，其在DT空间中的估计时延 $\hat{T}_{i,j}$ 与估计能耗 $\hat{E}_{i,j}$ 应满足以下条件：

$$T_{i,j}=\mathbb{E}\left(\hat{T}_{i,j}\right),\forall j\in\hat{\mathcal{T}}_i,\forall i\in\hat{\mathcal{U}} \tag{6}$$

$$E_{i,j}=\mathbb{E}\left(\hat{E}_{i,j}\right),\forall j\in\hat{\mathcal{T}}_i,\forall i\in\hat{\mathcal{U}} \tag{7}$$

因此，在孪生网络层的构建过程中，通过与DT虚拟映射的实时数据交互和反馈，需确保映射的推理任务处理时延和能耗尽可能与预期一致，并最小化公式(4)和公式(5)中所涉及的映射误差。

综上所述，基于DT驱动的边缘侧大模型系统架构中，通过构建高精度、高质量的孪生映射，孪生网络层能够为物理网络层实现资源协同优化提供准确可靠的数据支撑，以显著提升计算卸载策略的决策质量与系统资源分配的效率。同时，借助DT的实时感知与反馈机制，该架构还能够为后续多智能体DRL协同决策的训练建立全局知识共享平台。通过此平台，算法可以在动态且高度异构的环境中做出更为精准的决策，有效满足不同用户在差异化任务中的个性化QoS需求。

中国知网 https://www.cnki.net

4　问题定义

本节重点围绕联合优化系统模型中推理任务时延与能耗展开研究，旨在构建系统总成本最小化问题。在此基础上，进一步将该优化问题形式化为马尔可夫决策过程（Markov Decision Process，MDP）模型，为下一节基于DRL的资源协同优化策略提供理论支撑与建模基础。

4.1　问题公式化

在边缘侧大模型资源协同优化场景中，推理任务处理时延 $T_{i,j}$ 和能耗 $E_{i,j}$ 对用户而言同样重要。时延直接影响服务的实时性与响应质量，尤其在如医疗等对时效性要求极高的场景中，服务超时可严重降低用户体验，甚至危及用户生活安全。同时，较低的能耗不仅有助于延长智能设备的续航时间，减少充电或更换频率，还能提高计算与通信过程的稳定性和可靠性。

因此，在满足任务时延容忍度 $T^{\max}$ 的前提下，需尽可能降低能耗，以保障系统提供高可靠性、低时延的QoS。为了实现这一目标，本文将任务成本 $C_{i,j}$ 定义为推理任务 $task_{i,j}$ 的综合开销，以联合优化任务的处理时延与能耗，$C_{i,j}$ 定义如下：

$$C_{i,j}=\mu_i^{\mathrm{T}}T_{i,j}+\mu_i^{\mathrm{E}}E_{i,j} \tag{8}$$

此处，任务的时延与能耗 $T_{i,j}$ 和 $E_{i,j}$ 需要进行归一化处理，然后分别通过单位时间成本 μ_i^{T} 与单位能耗成本 μ_i^{E} 进行线性加权。最后，可以得到系统成本最小化问题 P 如下所示：

$$P:\quad \min_{\mathbf{s,f,b}}\sum_{i\in\mathcal{U}}\sum_{j\in\mathcal{T}_i}C_{i,j}$$

$$s.t.\quad \begin{cases} C_1: s_{i,j}\in\{0\}\cup\mathcal{M} \\ C_2: \sum_{j'=1}^{N_{i,i}^{\mathrm{l}}} f_{i,j'}\leqslant F_i,\ \forall i\in\mathcal{U} \\ C_3: \sum_{i'=1}^{N_k^{\mathrm{u}}}\sum_{j'=1}^{N_{i',k}^{\mathrm{l}}} f_{i,j'}\leqslant F_k,\ \forall k\in\mathcal{M} \\ C_4: \sum_{i'=1}^{N_k^{\mathrm{u}}} b_{i'}\leqslant B_k,\ \forall k\in\mathcal{M} \\ C_5: \sum_{j'=1}^{N_i^{\mathrm{l}}} E_{i,j'}\leqslant E_i^{-},\ \forall i\in\mathcal{U} \\ C_6: T_{i,j}\leqslant T^{\max},\ \forall i\in\mathcal{U},\forall j\in\mathcal{T}_i \end{cases} \tag{9}$$

其中，C_1 表示卸载决策的取值范围。C_2 和 C_3 分别确保推理任务本地处理或卸载处理时，所分配的计算资源不超过目标处理设备的最大计算能力 F_i 或者 F_k。C_4 确保卸载至同一MEC服务器任务分配的通信资源不会超出该服务器最大通信能力 B_k。C_5 确保每个用户设备处理任务时，所消耗的能量不超过设备的剩余电量 E_i^{-}。最后，C_6 则表示每个任务的时延不超过设定的最大容忍时延 $T^{\max}$。

4.2　问题分析

考虑到问题 P 的NP难度，以及在动态边缘侧大模型系统中决策环境的复杂性与高维性，采用基于DRL的方法是应对此类问题的有效手段。

为了准确建模MEC环境下多用户推理任务的计算卸载与资源协同优化过程，本文将每个用户建模为一个自主智能体，并引入分布式部分可观测MDP模型（Decentralized Partially Observable MDP，Dec-POMDP）作为系统的决策框架。该框架适用于多智能体间存在感知受限和通信不完全的边缘侧场景，特别是当涉及用户隐私保护时，智能体只能在具有不完全信息的动态博弈中进行决策的协同优化。

在该框架下，每个智能体 i 的POMDP可形式化表示为元组：$M_i=(S_i,A_i,T_i,R_i,O_i,Z_i)$，其中 S_i 表示系统环境的状态空间，A_i 表示智能体 i 的动作空间，T_i 表示状态转移概率函数，R_i 表示即时奖励函数，O_i 是观测空间，表示每个状态下一组可能的观测结果。Z_i 表示观测概率分布，用于描述在某一状态下可获得观测的概率。具体定义如下：

（1）观测空间 O_i

由于在边缘侧大模型资源协同优化过程中的用户设备数据的隐私性，每个智能体对全局环境状态的感知能力受限。智能体 i 在时间 t 的观测 $o_{t,i}$ 可以表示为：

$$o_{t,i}=\left\{\mathcal{T}_{t,i},E_{t,i}^{-},F_{t,i},\left\{F_{t,k},B_{t,k}\right\}_{k\in\mathcal{M}}\right\} \tag{10}$$

其中，包括时间 t 下智能体 i 的本地任务集合 $\mathcal{T}_{t,i}$，本地处理设备的剩余能量 $E_{t,i}^{-}$，本地计算频率 $F_{t,i}$，以及可感知所有MEC服务器 $k\in\mathcal{M}$ 的计算频率 $F_{t,k}$ 与通信资源 $B_{t,k}$。

（2）动作空间 A_i

动作空间包括其每个任务 $j\in\mathcal{T}_{t,i}$ 的卸载策略 $s_{t,i,j}$，以及分配的计算资源 $f_{t,i,j}$ 和通信资源 $b_{t,i,j}$。

$$a_{t,i}=\left\{\left\{s_{t,i,j},f_{t,i,j},b_{t,i,j}\right\}|j\in\mathcal{T}_{t,i}\right\} \tag{11}$$

（3）收益函数 R_i

智能体 i 的目标是尽量减少任务时延和能耗。任务的时延和能耗越低，智能体的奖励就越高。因此，智能体奖励 r_t^i 可通过任务成本 $C_{t,i,j}$ 以联合优化时延与能耗，具体如下：

$$r_{t,i}=-\sum_{j\in\mathcal{T}_{t,i}}\left(\mu_i^{\mathrm{T}}T_{t,i,j}+\mu_i^{\mathrm{E}}E_{t,i,j}\right) \tag{12}$$

中国知网　https://www.cnki.net

5 优化算法

本节将详细介绍所提出的MADDPG-CGAN算法，即基于CGAN增强的MADDPG算法及其具体实现细节，并对其在优化边缘侧推理任务卸载策略和资源协同分配中的作用进行深入分析。

如图2所示，MADDPG-CGAN算法模型主要由两个核心的模块构成：一是基于CGAN的DT映射模块，二是融合知识共享机制的MADDPG算法。前者通过实现对真实边缘场景的有效建模，为后续决策提供增强数据支持；后者则专注于解决多智能体间的策略博弈与协同决策问题。两者协同运行，共同提升了多智能体在边缘侧复杂任务中的推理效率与资源分配能力。

5.1 基于CGAN的DT映射

边缘侧资源协同优化场景中，由于各智能体在感知能力和观测范围方面存在一定限制，实现智能体间的实时状态信息共享成为关键。然而，以“集中式训练、分布式执行”为范式的MADRL算法在实际部署过程中，面临多方面的挑战。

一方面，智能体间状态信息的频繁共享显著增加

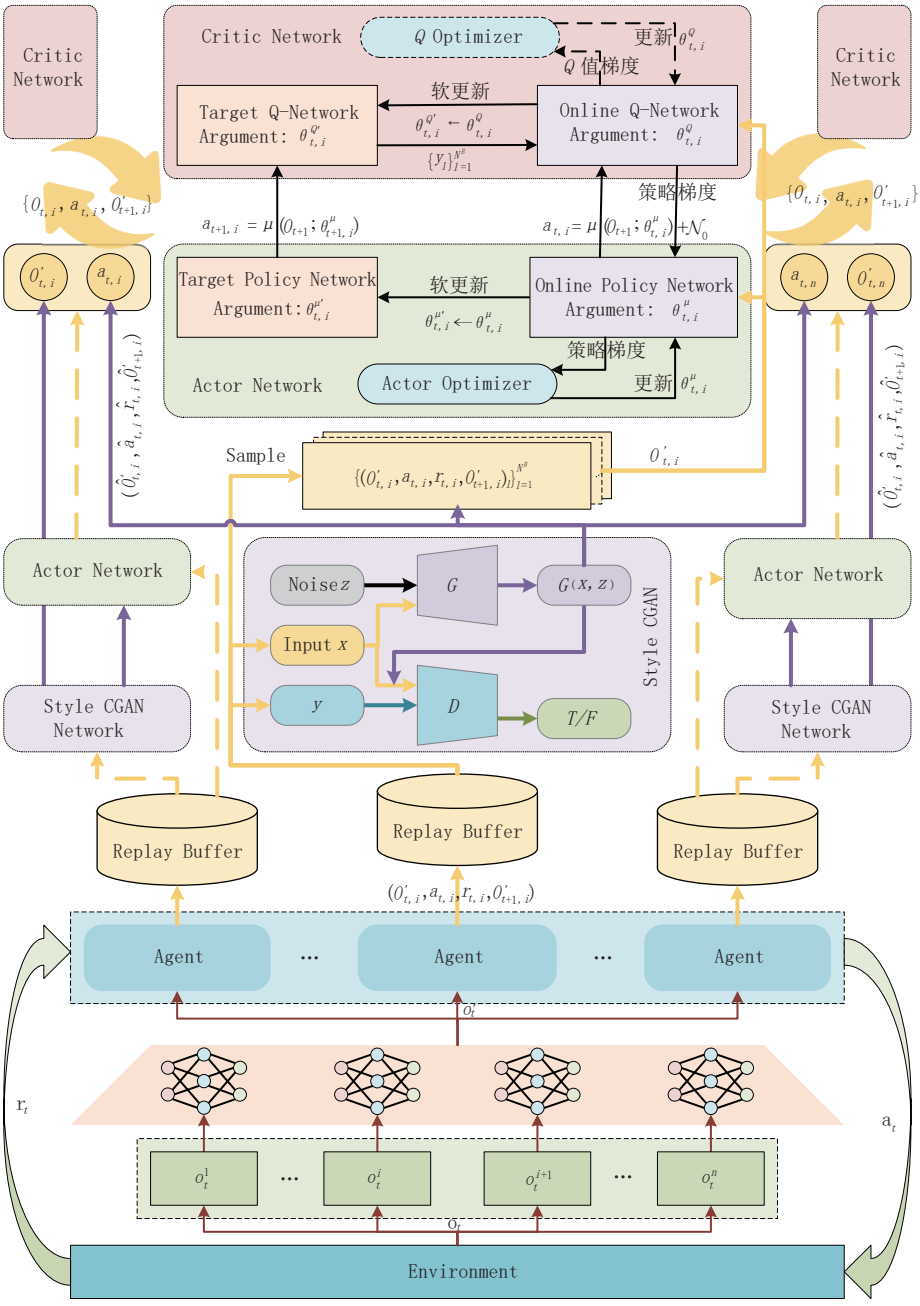

图2 MADDPG-CGAN算法模型图

中国知网 https://www.cnki.net

了通信开销与存储负担，进一步加剧了边缘设备在计算和能耗方面的本地压力。另一方面，这些共享信息往往包含用户的敏感数据，在传输、缓存和处理过程中存在潜在的隐私泄露风险，尤其是在缺乏端到端加密或隐私保护机制的情况下。此外，在边缘侧大模型的复杂环境中，由于数据分布异构、网络可达性差或策略隔离等因素，不同用户之间的训练数据难以直接交互。这不仅抑制了智能体之间的有效协作和知识迁移，也限制了策略泛化能力和模型整体性能。

为了应对上述问题，本文基于DT驱动的资源协同优化模型，借助CGAN以实现物理网络层中智能体与其虚拟孪生体之间的映射关系。该映射机制使得每个孪生体能在虚拟环境中模拟、预测并优化其对应智能体协同决策。特别是在MADRL算法的集中式训练的过程中，智能体可通过访问其他智能体的DT虚拟映射，间接获取其他智能体的经验数据，从而避免直接共享敏感数据。

如图3所示，所提出的CGAN算法采用改进StyleGAN架构作为生成器基础，并在其生成流程中引入条件状态信息，使得虚拟映射模块能够根据输入智能体当前的环境状态、动作历史和观测信息，有条件地生成潜在智能体训练数据分布。

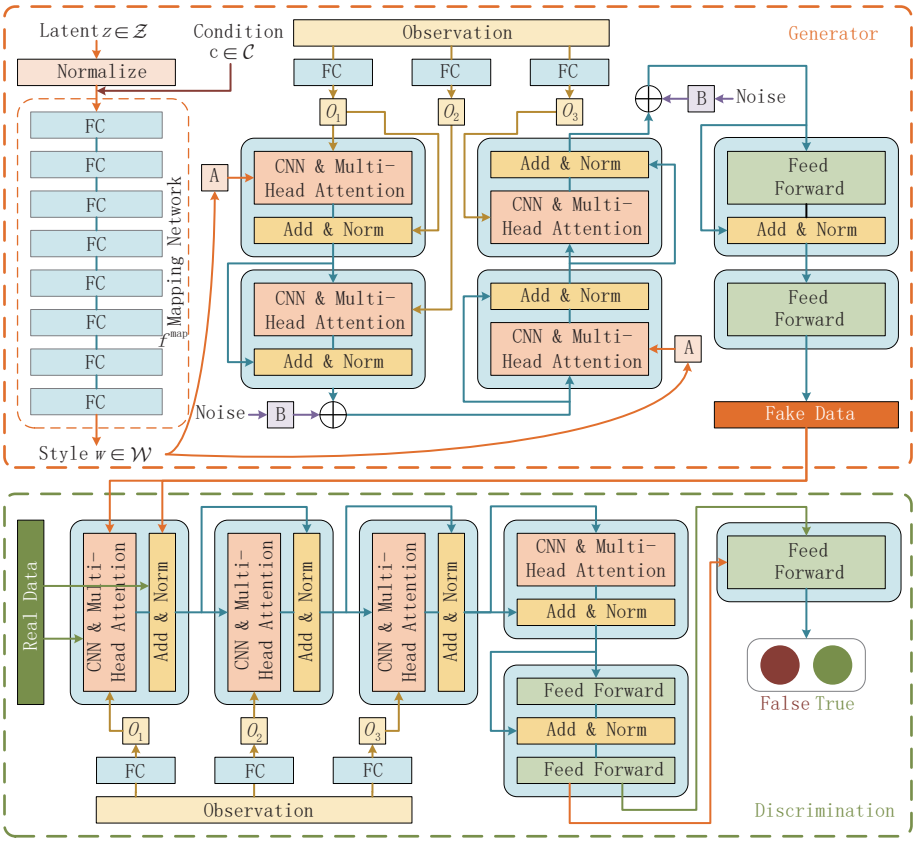

图3　条件StyleGAN网络模型图

首先，所设计的条件Style生成器G旨在给定条件状态信息$c\in\mathcal{C}$的约束下，从潜变量$z\sim\mathcal{N}(0,I)$中生成逼近真实样本分布的数据$\hat{x}=G(z,c)\approx x$。该生成器主要由条件风格映射网络、AdaIN模块和生成网络模块构成。

其中，风格映射网络将输入潜在变量z映射到风格向量空间，生成中间风格表示。该中间风格向量w不直接参与生成过程，而是作为调控信号注入到各层生成网络中，控制图像的多尺度结构、纹理细节及语义分布。为了增强生成样本对条件状态信息c的响应与控制，本文通过在风格映射网络中引入交叉注意力机制，构建了具备条件感知能力的条件风格映射网络。

在该网络中，条件状态首先经过嵌入编码，随后与潜变量z通过交叉注意力模块进行交互，并逐层引导风格映射网络输出具有状态语义的条件风格向量$\tilde{w}=W(z,c)$。该机制确保生成器在每一层风格注入中均具备对环境状态的感知能力，实现条件可控的样本生成。

为了在生成图像的条件一致性与风格多样性之

中国知网　https://www.cnki.net

间实现更优权衡，本文进一步在条件风格映射网络中引入残差结构，通过残差通路将条件风格向量与原始风格向量进行加权融合，实现风格引导强度的动态调节，提升风格表达在不同语义层级下的稳定性与泛化能力。最终输出的风格向量表示如下：

$$\tilde{\boldsymbol{w}}=(1-\lambda)\boldsymbol{w}+\lambda\boldsymbol{w}^{c} \tag{13}$$

其中，$\lambda\in(0,1)$表示平衡因子，$\boldsymbol{w}$和$\boldsymbol{w}^{c}$分别为由潜在变量$\boldsymbol{z}$映射得到的基础风格向量以及由条件状态c引导的条件风格向量。该设计使得生成器能够根据任务需求，在样本多样性与状态一致性间实现可控调节，从而支持智能体DT映射过程中的高效泛化能力与个性化行为建模。

其次，通过风格映射网络获得的风格向量$\tilde{\boldsymbol{w}}$将通过自适应实例归一化模块（Adaptive Instance Normalization，AdaIN）引入到生成网络中，作为连接风格信息与图像生成过程的关键桥梁。AdaIN的核心作用是将风格映射变量与生成网络得到的卷积特征$\mathcal{F}$进行融合，从而在保持图像生成质量的同时，有效控制样本的多样性与风格表达。

具体而言，AdaIN对每一层卷积特征$\mathcal{F}$进行标准化处理，并通过风格向量生成的缩放因子与偏移因子进行调节，其计算公式如下：

$$\mathrm{AdaIN}(\mathcal{F},\tilde{\boldsymbol{w}})=\alpha(\tilde{\boldsymbol{w}})\frac{\mathcal{F}-\mu(\mathcal{F})}{\sigma(\mathcal{F})}+\beta(\tilde{\boldsymbol{w}}) \tag{14}$$

其中，$\mu(\mathcal{F})$和$\sigma(\mathcal{F})$分别为特征图在通道维度上的均值和标准差，$\alpha(\tilde{\boldsymbol{w}})$和$\beta(\tilde{\boldsymbol{w}})$是由风格向量$\tilde{\boldsymbol{w}}$映射得到的缩放与偏移参数。

生成网络模块由多层卷积结构与Transformer Encoder模块共同构成，协同完成图像的逐层建构与全局语义建模。在前端阶段，生成网络通过一系列上采样卷积层逐步提升特征图的空间分辨率。同时，风格向量$\tilde{\boldsymbol{w}}$通过AdaIN机制注入至各卷积层中，引导特征生成过程，确保生成结果在满足条件约束的同时具备细致的局部纹理与结构信息。为了进一步提升生成图像的全局一致性与语义表达能力，在卷积模块之后引入Transformer Encoder模块。该模块通过多头自注意力机制建模特征图中不同区域之间的长程依赖关系，使生成网络在保持局部风格控制的基础上，能够捕捉跨区域的上下文关系与高层语义结构。最终，融合后的高维特征经由卷积输出层解码为完整图像，实现从潜在表示到高质量、条件一致性强的样本生成过程。

另一方面，判别器D结构在整体上借鉴了生成器G生成网络模块的设计，同样由卷积模块与Transformer Encoder模块构成，旨在判定输入样本与真实样本之间的差异性。与生成器G区别在于：判别器D在卷积模块提取出局部特征$\mathcal{F}^{\mathrm{D}}$后，直接将其与条件变量c通过交叉注意力机制进行融合，以显式建模样本与条件之间的对应关系。融合后的特征进一步输入至Transformer Encoder模块以捕捉全局上下文信息，最终通过输出模块计算得到输入样本为真实或生成的概率值，实现对条件样本的一致性与真实性的综合判定。

从生成效率来看，该StyleGAN网络结构在边缘侧场景中具有显著优势。其推理阶段仅需一次前向传播，从而避免扩散模型中多步去噪过程带来的高时延，生成速度更快。同时，相较于存在模糊和细节丢失问题的变分自编码器，GAN在图像细节保真度与视觉质量方面表现更优，能够满足边缘侧对低时延与高保真度的实时推理需求。

另一方面，相较于传统的无条件生成方法，该条件StyleGAN结构在样式映射阶段引入了状态嵌入作为调控信号，使生成过程具备更强的语义控制能力。这不仅提升了模型对多智能体行为策略的解释能力，也增强了其在不同任务和环境状态下的泛化性能，为DT虚拟映射提供了高保真、状态相关的行为数据支撑。

在条件StyleGAN模型的训练过程中，传统GAN通常采用生成器G与判别器D之间的最小-最大对抗博弈损失函数进行优化，其基本目标为

$$\min_{G}\max_{D}\Big\{\mathbb{E}_{\boldsymbol{x}\sim P_{data}}\big[\log D(\boldsymbol{x},\boldsymbol{c})\big]+\mathbb{E}_{\boldsymbol{z}\sim P_{\boldsymbol{z}}}\Big[\log\big(1-D\big(G(\boldsymbol{z},\boldsymbol{c})\big)\big)\Big]\Big\} \tag{15}$$

然而，该损失形式存在训练不稳定、梯度消失等问题，特别是在高分辨率图像生成及多条件边缘侧环境建模中表现不佳。为缓解上述问题，WGAN（Wasserstein GAN）基于Wasserstein距离，通过优化生成样本分布P_g与真实数据分布P_{data}之间的Wasserstein距离，以提升生成过程稳定性与生成质量。其基本对抗目标函数可表示为

$$\min_{G}\max_{D}\mathbb{E}_{\boldsymbol{x}\sim P_{data}}\big[D(\boldsymbol{x},\boldsymbol{c})\big]-\mathbb{E}_{\hat{\boldsymbol{x}}\sim P_{g}}\big[D(\hat{\boldsymbol{x}},\boldsymbol{c})\big] \tag{16}$$

为进一步确保判别器满足1-Lipschitz连续性要求，WGAN-GP在上述损失基础上引入了梯度惩罚项（Gradient Penalty），以替代传统的权重裁剪方式，有效避免梯度爆炸或消失问题，进一步提高训练稳定性。引入惩罚项后的判别器优化目标函数表示为：

中国知网 https://www.cnki.net

$$\mathcal{L}_D=\mathbb{E}_{\hat{x}\sim P_g}\left[D(\hat{x},c)\right]-\mathbb{E}_{x\sim P_{data}}\left[D(x,c)\right]+\rho_{\mathrm{GP}}\cdot\mathcal{L}_{\mathrm{GP}} \tag{17}$$

$$\mathcal{L}_{\mathrm{GP}}=\mathbb{E}_{\hat{x}\sim P_g}\left[\left\|\nabla_{\hat{x}}D(\hat{x},c)\right\|_2-1\right]^2 \tag{18}$$

其中，ρ_{GP}表示梯度惩罚系数。公式(18)表示惩罚项约束$\mathcal{L}_{\mathrm{GP}}$的计算公式，通过强制判别器$D$在插值样本上的梯度接近1，从而实现对Lipschitz条件的软约束，进而有效提升了对抗训练的稳定性。

同时，为了增强生成器G对条件变量c的响应能力与控制能力，本文在生成器损失中引入了一个条件一致性约束项$\mathcal{L}_{\mathrm{CE}}$，以分类交叉熵(Cross-Entropy，CE)损失形式对生成样本进行条件标签监督，引导其生成更加一致且语义明确的样本。改进后的生成器的优化目标函数表示为

$$\mathcal{L}_G=\mathbb{E}_{z\sim P_z}\left[D\left(G(z,c),c\right)\right]+\rho_{\mathrm{CE}}\cdot\mathcal{L}_{\mathrm{CE}}\left(C\left(G(z,c)\right),c\right) \tag{19}$$

其中，ρ_{CE}用于调节对抗损失与条件约束的权重。$\mathcal{L}_{\mathrm{CE}}$旨在衡量生成样本的预测结果$C\left(G(z,c)\right)$与真实条件标签$c$之间的一致性。

最终，整个训练过程采用判别器D与生成器G交替优化的方式进行。判别器D不断提升其区分真实与生成样本的能力，而生成器G则在对抗损失与条件一致性损失的双重驱动下，持续优化生成质量与条件控制能力，最终实现高质量、状态一致且语义可控的样本生成过程。

5.2　基于知识共享的MADDPG

根据Dec-POMDP模型，该边缘侧大模型场景属于多智能体环境，且智能体的动作包括连续的资源分配动作，因此，本文采用MADDPG算法来解决多智能体场景下智能体间的信用分配问题，以实现高效边缘智能推理卸载和资源协同分配。

MADDPG算法是DDPG算法在多智能体场景中的扩展，能够有效缓解智能体感知空间受限的问题，并应对智能体间竞争与合作挑战。如模型图2所示，MADDPG算法延续Actor-Critic架构，基于集中式训练和分布式执行框架实现多智能体资源协同分配决策以及模型训练。在该框架下，每个智能体$i\in\mathcal{U}$的Actor网络仅需接受本地的状态感知$o_{t,i}$作为输入，并根据$o_{t,i}$独立地执行策略$a_{t,i}$，而无需考虑其他智能体决策。用于集中式训练的Critic网络则需要接收所有智能体当前时刻t下的观测$\boldsymbol{o}_t=\left\{o_{t,1},o_{t,2},\cdots,o_{t,N^u}\right\}$和动作$\boldsymbol{a}_t=\left\{a_{t,1},a_{t,2},\cdots,a_{t,N^u}\right\}$，以便使每个智能体在制定策略时考虑到其他智能体的影响，从而更准确地评估智能体的价值，并有效解决信用分配问题。

为了支持MADDPG算法的集中式训练，每个智能体的经验池$\mathcal{D}_i$不仅存储自身历史经验数据，还应包括所有其他智能体经验数据。$\mathcal{D}_i$中每一条经验表示为$\left((\boldsymbol{o}_t,s'_t),\boldsymbol{a},\boldsymbol{r},(\boldsymbol{o}_{t+1},s'_{t+1})\right)$元组，其中$s'$表示部分可观测的全局状态信息。通过时序差分法，可以计算得到真实的状态-动作值，如下所示：

$$y_{t,i}=r_{t,i}+\gamma Q'_i\left((\boldsymbol{o}_t,s'_t),\boldsymbol{a}_t\right)\big|_{a_{t,i}=\mu_i\left(o_{t,i},s'_t\right)} \tag{20}$$

接着，Critic网络通过最小化真实状态-动作值$\boldsymbol{y}$与在线Critic网络的估计Q值之间的均方误差来进行训练，具体如下：

$$\mathcal{L}(\theta_i)=\mathbb{E}\left[y_{t,i}-Q_i\left((\boldsymbol{o}_t,s'_t),\boldsymbol{a}_t\right)\right]^2 \tag{21}$$

Actor网络则通过最大化Critic网络的Q值来进行训练，即Actor的损失函数可以表达如下：

$$\mathcal{L}(\omega_i)=-\mathbb{E}\left[Q'_i\left((\boldsymbol{o}_t,s'_t),\boldsymbol{a}_t\right)\right] \tag{22}$$

此外，目标网络的参数$\boldsymbol{\theta}'$和$\boldsymbol{\omega}'$基于在线网络进行软更新，以提高MADDPG稳定性和收敛性。

如第5.1节所述，传统的MADDPG算法在执行集中式训练时，通常面临智能体之间状态信息共享与同步困难的问题。如何在保障数据安全与隐私的前提下，实现高效、可靠的集中式训练机制，已成为边缘侧资源协同优化的关键问题。

为此，本节基于上一节提出的基于条件StyleGAN的DT虚拟映射方法。如图2中间部分所示，每个智能体通过其本地经验池完成个体训练并构建其对应的DT虚拟映射模型。在集中式训练阶段，智能体无需直接交换原始观测数据，而是通过访问其他智能体的DT映射结果来获取经验特征，实现间接的数据共享。

具体而言，在当前时刻t下，智能体独立执行决策$a_{t,i}$并转移到下一状态$o_{t+1,i}$后，该状态$o_{t+1,i}$将在DT虚拟映射空间中继续迭代n次。智能体的迭代信息$\left\{\left((\hat{o}_{t,i},\hat{s}_t),\hat{a}_{t,i},\hat{r}_{t,i},(\hat{o}_{t+1,i},\hat{s}_{t+1})\right)\middle|t\in(t_0,t_0+n]\right\}$将以时间序列形式存储在对应DT经验池$\mathcal{D}'_i$中。时间$t_0$下的$\mathcal{D}'_i$中的数据表示如下：

$$\mathcal{D}'_i(t_0)=\begin{bmatrix}\delta_{t_0}^{t_0+1},\delta_{t_0}^{t_0+2},\cdots,\delta_{t_0}^{t_0+n}\\ \delta_{t_0-1}^{t_0+1},\delta_{t_0-1}^{t_0+2},\cdots,\delta_{t_0-1}^{t_0+n}\\ \cdots\\ \delta_{t_0-n'}^{t_0+1},\delta_{t_0-n'}^{t_0+2},\cdots,\delta_{t_0-n'}^{t_0+n}\end{bmatrix} \tag{23}$$

其中，δ_p^q表示在时间$p\in[0,n']$，孪生映射体在时间

中国知网　https://www.cnki.net

$q\in[1,n]$的预测结果。

随着迭代次数的增加，$\mathcal{D}_i'$中预测结果与真实结果偏差越来越大。由于预测数据的准确性无法完全保证，不能直接将DT经验池$\mathcal{D}_i'$的预测数据用于丰富智能体经验池。因此，t_0时刻下可以为$\mathcal{D}_i'$中的之后时刻数据进行决策，并通过$\varepsilon-greedy$算法将预测列表$\left\{\hat{\delta}_t^i|t\in[t_0,t_0+n]\right\}$添加到真实经验池$\mathcal{D}_i$中。即对于每个时刻$t$，以概率$\varepsilon$从预测序列$\left\{\hat{\delta}_t^i|t\in[t_0,t_0+n]\right\}$中选择$Q(\hat{s}_{t,i},\hat{a}_{t,i})$最大的预测，否则根据时间间隔权重$1-(p+q)/(n+n')$随机选择预测。此外，概率$\varepsilon$随着时间间隔$q$的增加逐渐减小，以降低DT映射误差的累积，确保经验池中的数据更加贴近实际环境，从而提升模型的训练效果和决策精度。

通过引入条件StyleGAN，MADDPG算法在集中式训练过程中的时间复杂度可以得到显著优化。通过生成虚拟数据有效减少了智能体间的实时数据同步需求，通信复杂度可以从$O\left((N^u)^2\right)$降至$O(N^u)$。此外，条件StyleGAN可以通过提供更加多样化的训练数据，增强了模型的训练效率，减少了训练过程中的迭代次数，从而进一步降低了整体算法的时间复杂度。

与此同时，基于DT驱动的边缘侧大模型架构提供的知识共享平台，本文引入FL来实现智能体之间的知识共享。通过在该平台上共享各智能体Actor和Critic网络参数，不仅能够有效解决数据碎片化和信息隔离问题，还在确保边缘侧用户隐私安全的基础上，加速模型的训练过程并提升其效果。在FL框架下，每个智能体只需维护本地的DRL模型。经过一段时间的训练后，智能体i将在线网络的参数$(\theta_{t,i},\omega_{t,i})$进行广播，并接收其他智能体参数。这些参数随后被聚合为$\left(\hat{\theta}_{t,i},\hat{\omega}_{t,i}\right)$，以更新智能体$i$的目标网络。以Critic网络为例，目标Critic网络的参数更新公式如下所示：

$$\omega'_{t+1,i}=\tau^{\omega}\left((1-\tau_t)\omega_{t,i}+\tau_t\hat{\omega}_{t,i}\right)+\left(1-\tau^{\omega}\right)\omega'_{t,i} \tag{24}$$

其中，在线Critic网络参数聚合时，权重$\tau_t\in(0,1)$随着迭代次数的增加逐渐减小。

此外，在参数聚合过程中，不同智能体对状态的感知存在差异，可能导致参数聚合后逼近全局模型准确性下降。因此，考虑到边缘侧大模型用户的异质性，本文根据智能体当前价值$Q_{t,i}$和经验池大小$N_{t,i}^d$来调整共享参数，以获取聚合参数$\hat{\omega}_{t,i}$，如下所示：

$$\hat{\omega}_{t,i}=\frac{\sum_{i'\in\mathcal{U}}\left(Q_{t,i'}\omega_{t,i'}/N_{t,i'}^d\right)}{N^{\mathrm{u}}Q_{t,i}/N_{t,i'}^d} \tag{25}$$

综上，MADDPG-CGAN算法结合了基于知识共享的MADDPG算法以及CGAN增强模块，旨在优化边缘侧多智能体的资源协同优化问题。其中MADDPG基于分布式执行集中式训练架构来缓解感知受限并促进智能体间协作，同时基于FL框架借助DT架构实现智能体间知识共享。基于CGAN实现DT虚拟映射，通过生成高质量的经验特征数据，减少MADDPG集中式训练中数据传输以保障用户隐私，并丰富训练集以增强算法整体性能。算法伪代码如算法1所示：

算法1. CGAN增强的MADDPG算法

输入：初始化DT驱动的边缘侧大模型系统，初始化所有智能体经验池$\{\mathcal{D}_i\}_{i\in\mathcal{U}}$以及DT经验池$\{\mathcal{D}_i'\}_{i\in\mathcal{U}}$，初始化智能体的在线网络$\theta$和目标网络参数$\theta'$，初始化每个智能体DT虚拟映射，即CGAN算法中生成器G和辨别器D的网络模型。

输出：所有智能体策略$\{\pi_i\}_{i\in\mathcal{U}}$包括卸载决策$s$，通信资源分配$b$，计算资源分配$f$

For episode=1 do：
  重置DT驱动的边缘侧大模型架构；
  智能体感知状态$o_{0,i}$以及受限全局状态s_0'；
  For t=1 do：
    智能体根据当前的策略$\pi_{t,i}$执行决策$a_{t,i}$；
    通过与环境进行交互得到当前收益$r_{t,i}$以及下一状态信息$(o_{t+1,i},s_{t+1}')$；
    将当前经验$\left((o_{t,i},s_t'),a_{t,i},r_{t,i},(o_{t+1,i},s_{t+1}')\right)$存储到本地经验池$\mathcal{D}_i$中；
    For 每个智能体$i\in\mathcal{U}$ do：
      从本地经验池中随机抽样一批经验数据；
      每个智能体通过Actor网络独立地得到决策$a_{t,i}=\mu_{t,i}(o_{t,i},s_t')$；
      每个智能体通过知识共享平台来访问其他智能体的DT虚拟映射，进而获得当前时刻下其他智能体的经验特征。
      根据所获取的经验数据$(\boldsymbol{o}_t,\ \boldsymbol{a}_t)$通过Critic得到当前的动作-状态价值$Q_i((\boldsymbol{o}_t,\ s_t'),\boldsymbol{a}_t)$。
      根据公式(21)和(22)分别训练Actor和Critic网络，并通过软更新方式更新目标网络。
      智能体基于本地经验数据根据公式(17)和(19)来分别训练CGAN算法生成器和辨别器；
      通过CGAN网络来预测当前状态下后续的经验特征序列，并填充DT经验池$\mathcal{D}_i'$；

中国知网 https://www.cnki.net

16.　通过ε－*greedy*算法选择丰富本地经验池；
17.　End For
18.　每经过进行几次迭代，智能体通过FL共享本地智能模型，并通过公式(24)来更新参数；
19.　End For
20.　End For

由算法伪代码可以看出，基于CGAN增强的MADDPG算法相比于传统MADDPG在减少数据通信和决策同步上有显著优势。在传统方法中，智能体间需要在11步时频繁交换实际经验数据，将会导致高通信开销。而CGAN通过生成虚拟经验数据，有效减少了智能体间的数据交换，降低通信频率。借助于CGAN增强的MADDPG智能体可以独立生成并优化策略，避免全局同步的高延迟。这些优化可以有效提升训练效率，特别是在动态和分布式边缘侧环境中，可以显著增强系统的可扩展性和响应速度。

6　性能评估

本节构建了一个DT驱动的边缘侧仿真环境，并通过对比其他基准算法来分析所提出的MADDPG-CGAN算法的性能表现。首先，仿真实验环境搭建在一台配备了Intel i7-10700K处理器、32 GB内存和NVIDIA RTX 3070显卡的计算机上进行，该设备可以支持高效的深度学习训练和仿真环境的实时模拟。软件环境方面，实验使用Python 3.9并结合gym包构建仿真环境，同时使用了PyTorch框架进行深度学习模型的训练与评估。

为了更真实地模拟实际边缘侧资源协同优化环境，仿真实验中引入了随机生成的推理任务，以体现真实场景中的任务多样性与复杂性。这些任务基于MedMNIST数据集构建，该数据集涵盖多个医学图像子数据集，分别面向不同类型的医学图像分类任务，如皮肤病图像分类、心电图分析、眼底图像识别等，任务形式涵盖二分类与多分类，具备良好的任务多样性和代表性。针对上述数据集，实验选用多种复杂度深度学习模型，包括CNN、ResNet以及DenseNet等，用以模拟实际边缘侧大模型场景中因任务异构性而产生的多样化推理需求。具体的子数据集及其推理任务配置详见表2，其中训练集占总样本数量的70%。

为进一步增强仿真实验的现实性，实验环境设置了用户终端与MEC之间异构的计算资源和网络带宽配置，并引入高斯噪声以模拟通信信道中的干扰与波动。该设置有助于还原边缘侧资源协同优化场景中复杂、受限的网络环境，更准确地评估所设计模型在实际部署条件下的鲁棒性与性能表现。主要仿真参数配置详见表1。

表1　模型参数

参数	值
用户设备的计算频率	1.0～2.0 GHz
用户设备的发射功率	1.0～2.0 kW
MEC服务器的计算频率	5.0～10.0 GHz
MEC服务器的带宽	100～200 kHz
用户与服务器之间的距离	200～400 m
信道增益	10^{-14}～2×10^{-14} W
高斯噪声	10^{-13} W

首先，本文对基于条件StyleGAN的DT虚拟映射能力进行了性能验证。以MNIST数据集为测试基准，分别引入最小-最大对抗博弈损失(Minimax GAN)、WGAN、WGAN-GP，以及本文提出的条件WGAN-GP(Conditional WGAN-GP)进行对比实验。在实验中，为每类标签生成100张图像，累计20轮次，总计生成20 000张DT虚拟样本，并基于这些样本的识别准确率作为评估指标。四种模型对应的平均分类准确率分别为0.980 19(Minimax GAN)、0.919 89(WGAN)、0.979 10(WGAN-GP)以及0.982 79(Conditional WGAN-GP)。

图4展示了各类GAN在条件控制下的图像生成效果。结果表明，尽管Minimax GAN在准确率方面较高，但其生成样本在多样性方面表现不足，存在模式坍塌的风险；WGAN在提升样本多样性方面效果较好，但其对梯度的依赖较强，训练过程缓慢，且最终模型难以收敛至最优状态，导致准确率偏低。相比之下，WGAN-GP和Conditional WGAN-GP在引入梯度惩罚后，训练更为稳定并能加速收敛；特别是本文提出的Conditional WGAN-GP，在利用标签条件引导生成器的前提下，不仅提升了图像生成质量，还显著增强了生成样本的多样性与类内一致性。

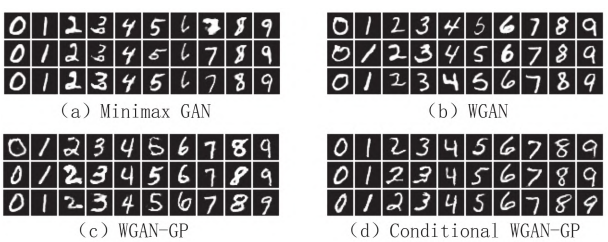
(a) Minimax GAN　(b) WGAN
(c) WGAN-GP　(d) Conditional WGAN-GP

图4　条件StyleGAN生成效果对比

中国知网　https://www.cnki.net

进一步地，本文在MedMNIST数据集上对所提出的DT虚拟映射方法进行了扩展评估，其结果汇总于表2最后一列。从结果可以看出，基于StyleGAN的DT虚拟映射方法能够有效模拟真实推理任务的特征分布，尤其在边界清晰、类别区分度高的数据集上（如PneumoniaMNIST和OrganMNIST）表现尤为显著，生成样本在结构清晰度和语义一致性方面具有较高质量。然而，对于类间差异较小、类别边界模糊或者具有高度相似性的任务（如BloodMNIST和RetinaMNIST）以及多标签任务（如ChestMNIST），虚拟样本的判别性仍存在一定提升空间，生成模型在细粒度特征区分上的能力有待进一步优化。

表2 基于MedMNIST数据集的推理任务配置以及DT映射准确率

数据集	任务类别	样本数量	模型结构	参数量	FLOPs	ACC	AUC	**Gen ACC**
PneumoniaMNIST	二分类	5856	CNN	0.62 M	40.37 M	0.851	0.944	**0.845**
BreastMNIST	二分类	780	CNN	1.14 M	40.89 M	0.819	0.957	**0.829**
OCTMNIST	多分类(4)	109 309	EfficientNet	3.97 M	31.98 M	0.761	0.946	**0.736**
OrganMNIST-Axial	多分类(11)	58 830	EfficientNet	3.98 M	31.99 M	0.925	0.996	**0.902**
OrganMNIST-Coronal	多分类(11)	23 583	ResNet	11.18 M	148.87 M	0.897	0.993	**0.887**
OrganMNIST-Sagittal	多分类(11)	25 211	DenseNet	6.88 M	231.31 M	0.752	0.974	**0.749**
BloodMNIST	多分类(8)	17 092	ResNet+NL-Attention	11.19 M	150.96 M	0.931	0.997	**0.907**
DermaMNIST	多分类(7)	10 015	EfficientNet+SEBlock	3.57 M	30.38 M	0.780	0.964	**0.764**
TissueMNIST	多分类(8)	236 386	ResNet	11.18 M	148.86 M	0.764	0.942	**0.719**
PathMNIST	多分类(9)	107 180	ResNet+DenseNet	18.15 M	385.29 M	0.791	0.976	**0.684**
ChestMNIST	多标签(14)二分类	112 120	DenseNet+NL-Attention	9.32 M	245.06 M	0.945	0.739	**0.823**
RetinaMNIST	序数回归(5)	1600	DenseNet+NL-Attention	9.06 M	244.81 M	0.548	0.786	**0.481**

此外，该DT虚拟映射方法在小样本场景中展现出较强的泛化能力。例如在BreastMNIST数据集上，生成器可基于有限的训练样本合成高质量伪样本，从而有效扩展原始样本空间，增强数据分布的覆盖度，进而提升分类模型的训练效果与预测准确率。

其次，针对基于知识共享的MADDPG资源协同优化算法的性能评估。如图5所示引入CGAN增强模块后，MADDPG算法在训练过程中展现出更平滑且快速的损失收敛趋势。CGAN生成的高质量伪专家数据有效扩充了训练样本空间，显著加快了MADRL算法的收敛速度，并提升了整体训练过程的稳定性与鲁棒性。图5的实验结果验证了虚拟映射机制不仅具备优越的样本生成质量，还能够为集中式训练架构下的策略优化过程提供高质量的数据支持，显著增强多智能体协同决策任务中的学习效率与协同性能，充分体现了该方法在实际应用场景中的实用性与可行性。

此外，表3通过对比传统MADDPG和CGAN增强的MADDPG算法，展示了MADDPG-CGAN算法在引入虚拟映射生成后的显著优势。引入CGAN后，MADDPG-CGAN算法数据同步时延得到了有效减少，训练过程加速，数据传输需求显著降低，同时训练效果得到了提升。

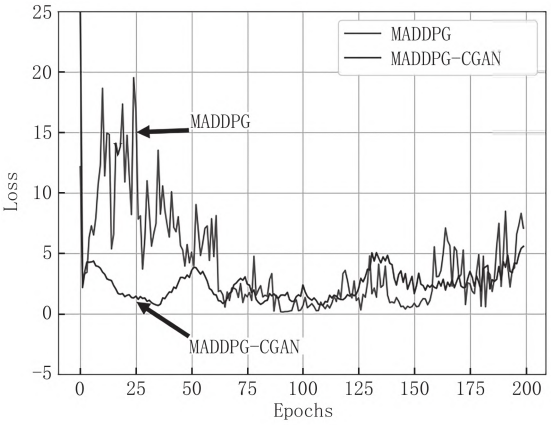

图5 MADDPG算法的损失曲线

表3 MADDPG-CGAN算法的算法复杂度对比

量化指标	传统MADDPG	CGAN增强的MADDPG
平均收敛时间(训练轮次)	38	60
数据同步时延(毫秒/步)	5.31	2.59
最终平均奖励(Reward)	50.74	55.49
模型性能提升(%)	8.55%	11.04%

接下来，为了全面评估所提出的MADDPG-CGAN算法的性能，本文将其与三种基准算法——Greedy、DQN及DDPG进行了对比分析。其中，Greedy算法在多智能体场景下体现为个体理性决策，缺乏上层协调机制，各智能体倾向于最大化自身

中国知网 https://www.cnki.net

利益，往往忽略系统整体性能；而DQN与DDPG作为主流DRL算法，借助集中式训练与分布式执行的范式（如VDN结构），也能在一定程度上实现多智能体的策略协调与全局优化。

实验首先对不同用户数量下，各算法在任务平均处理时延与能耗方面的表现进行了分析。图6和图7展示了用户数量从3增加至12时，不同算法下医疗任务的处理时延与能耗变化趋势，MEC服务器数量固定为6。随着用户数量增加，MEC服务器的计算能力逐渐成为瓶颈，Greedy算法由于缺乏策略探索与智能体间交互建模，性能下降明显；DQN算法在处理连续动作空间方面存在局限，难以有效满足医疗场景中差异化的服务需求。而相比之下，MADDPG-CGAN算法通过引入CGAN生成的高精度伪专家数据，提升了训练样本的多样性与质量，增强了智能体间的协作能力，有效缓解了由局部信息不完全导致的性能瓶颈，显著提升了模型的全局优化能力与适应性。

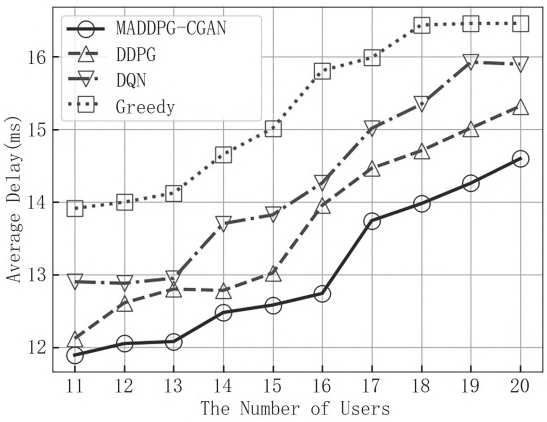

图6　不同用户数量下的不同算法的任务平均处理时延

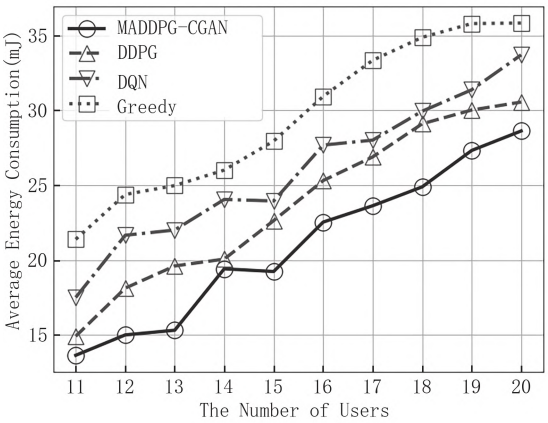

图7　不同用户数量下的不同算法的任务平均处理能耗

进一步地，图8和图9展示了在用户数量固定为8的前提下，随着MEC服务器数量的增加，系统在时延与能耗方面的性能变化趋势。综合对比结果显示，MADDPG-CGAN在原DDPG架构基础上，融合了DT映射与模型参数聚合机制，不仅通过精确的虚拟任务映射提升经验数据共享质量，还通过优化资源分配与智能体协作策略，在资源协同优化场景中实现显著的时延与能耗优化，展现出优于传统算法的整体性能与环境适应性。

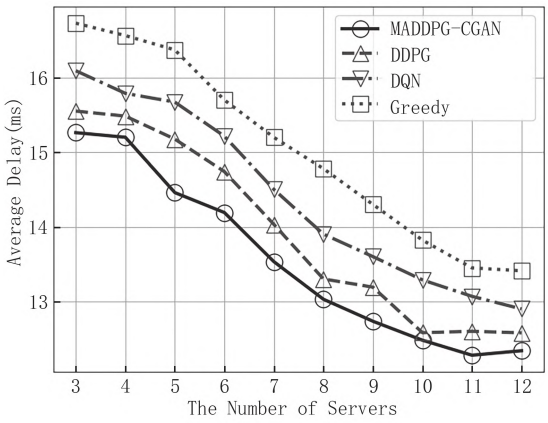

图8　不同服务器数量下的不同算法的任务平均处理时延

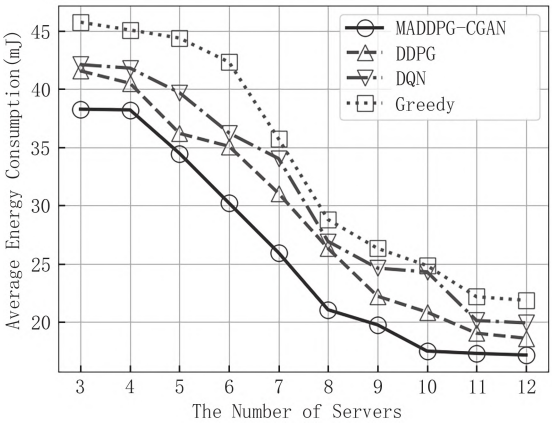

图9　不同服务器数量下的不同算法的任务平均处理能耗

最后，本文进一步评估了所提出的个性化FL在边缘侧资源协同优化场景中知识共享的效果。通过与集中式训练方式以及传统联邦平均算法（FedAvg）进行对比，验证所提出的个性化FL机制在协同策略学习中的性能优势。如图10所示，所提出的个性化FL方法在收敛速度与最终性能方面均优于FedAvg，表现出更强的稳定性与泛化能力。此外，个性化FL机制有效规避了集中式训练中对大规模数据传输的依赖，降低了通信开销，并在一定程度上可以缓解因数据集中带来的隐私泄露风险。

中国知网　https://www.cnki.net

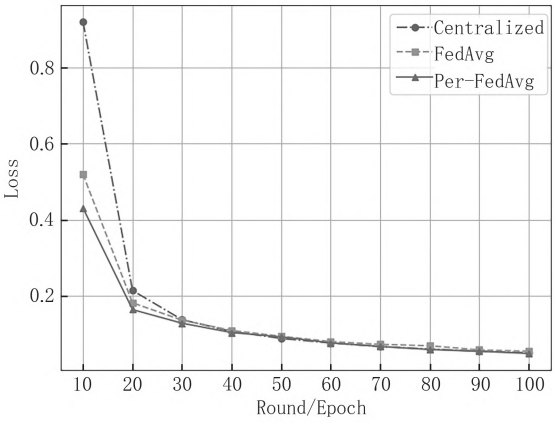

图10 不同FL聚合机制的收敛曲线对比

7 结论及未来工作

本文构建了一个融合实时信息交互与动态反馈机制的DT驱动边缘侧大模型系统，实现了MEC模型与异构推理任务之间的高效虚拟映射，并支持多用户环境下的资源协同调度。算法中引入的条件式StyleGAN模块可以用于生成高质量、多样化且语义一致的智能体经验样本，实现了DT空间中对真实任务环境的有效虚拟映射；同时，所提出的MADDPG-CGAN协同优化算法则可在多用户任务卸载过程中兼顾系统性能与安全性，有效降低任务处理延迟与能耗，并增强整个系统的鲁棒性与适应性。仿真实验结果表明，在基于MedMNIST数据集的边缘侧智能任务场景中，相较于基准方法，MADDPG-CGAN算法可将任务平均时延降低约12%，能耗下降约25%。

未来工作将首先聚焦于在多模态与高复杂度任务数据集扩展方面，将引入更具挑战性的多模态医疗数据集，或涵盖智能交通、工业物联网等典型场景边缘协同推理任务数据集，以提升系统的通用性与不同应用场景下的适应能力。其次，研究将结合更先进的MADRL算法，以进一步提升智能体间的协同感知、信息共享与分布式决策能力，从而更高效地应对任务异构性与资源动态变化所带来的挑战。最后，在生成式AI与DT的深度融合方向，将进一步挖掘生成式AI与DT技术的协同潜力，通过增强表征学习能力，强化模型在高动态性、复杂网络拓扑以及大规模实时数据处理环境下的泛化能力与资源调度优化表现。

参 考 文 献

[1] Chen W, Lin X, Lee J, et al. 5G-advanced toward 6G: Past, present, and future. IEEE Journal on Selected Areas in Communications, 2023, 41(6): 1592-1619

[2] Tang F, Chen X, Rodrigues T K, et al. Survey on digital twin edge networks (DITEN) toward 6G. IEEE Open Journal of the Communications Society, 2022, 3: 1360-1381

[3] Ji B, Wang Y, Song K, et al. A Survey of computational intelligence for 6G: Key technologies, applications and trends. IEEE Transactions on Industrial Informatics, 2021, 17(10): 7145-7154

[4] Li C, Gan Y, Zhang Y, et al. A cooperative computation offloading strategy with on-demand deployment of multi-UAVs in UAV-aided mobile edge computing. IEEE Transactions on Network and Service Management, 2024, 21(2): 2095-2110

[5] Hou W, Wen H, Song H, et al. Multiagent deep reinforcement learning for task offloading and resource allocation in cybertwin-based networks. IEEE Internet of Things Journal, 2021, 8(22): 16256-16268

[6] Liu W, Li B, Xie W, et al. Energy efficient computation offloading in aerial edge networks with multi-agent cooperation. IEEE Transactions on Wireless Communications, 2023, 22(9): 5725-5739

[7] Yao Z, Xia S, Li Y, et al. Cooperative task offloading and service caching for digital twin edge networks: A graph attention multi-agent reinforcement learning approach. IEEE Journal on Selected Areas in Communications, 2023, 41(11): 3401-3413

[8] Liu F, Yu H, Huang J, et al. Joint service migration and resource allocation in edge IoT system based on deep reinforcement learning. IEEE Internet of Things Journal, 2023, 11(7): 11341-11352

[9] Suzuki A, Kobayashi M, Oki E. Multi-agent deep reinforcement learning for cooperative computing offloading and route optimization in multi cloud-edge networks. IEEE Transactions on Network and Service Management, 2023, 20(4): 4416-4434

[10] Feriani A, Hossain E. Single and multi-agent deep reinforcement learning for AI-enabled wireless networks: A tutorial. IEEE Communications Surveys & Tutorials, 2021, 23(2): 1226-1252

[11] Wu C, Li L, Zhang L, et al. Efficient GAN-based federated optimization for vehicular task offloading with mobile edge computing in 6G network. IEEE Internet of Things Journal, 2025, 12(3): 2736-2748

[12] Zhao J, Li B, Yang T. Hierarchical federated learning in inland waterways via dataset distillation and resource allocation. IEEE Transactions on Vehicular Technology, 2025, 74(3): 3695-3707

[13] Deng X, Yin J., Guan P, et al. Intelligent delay-aware partial computing task offloading for multiuser industrial internet of things through edge computing. IEEE Internet of Things Journal, 2023, 10(4): 2954-2966

中国知网 https://www.cnki.net

[14] Zhang Z, Wang N, Wu H, et al. MR-DRO: A fast and efficient task offloading algorithm in heterogeneous edge/cloud computing environments. IEEE Internet of Things Journal, 2023, 10(4): 3165-3178

[15] Liu F, Chen J, Zhang Q, et al. Online MEC offloading for V2V networks. IEEE Transactions on Mobile Computing, 2023, 22(10): 6097-6109

[16] Djigal H, Xu J, Liu L, et al. Machine and deep learning for resource allocation in multi-access edge computing: A survey. IEEE Communications Surveys & Tutorials, 2022, 24(4): 2449-2494

[17] Wang M, Zhang L, Gao P, et al. Stackelberg game-based intelligent offloading incentive mechanism for a multi-UAV-assisted mobile edge computing system. IEEE Internet of Things Journal, 2023, 10(17): 15679-15689

[18] Zhang R, Xie Z, Yu D, et al. Digital twin-assisted federated learning service provisioning over mobile edge networks. IEEE Transactions on Computers, 2023, 73(2): 586-598

[19] Lee J, Solat F, Kim T Y, et al. Federated learning-empowered mobile network management for 5G and beyond networks: From access to core. IEEE Communications Surveys & Tutorials, 2024, 26(3): 2176-2212

[20] Zhang R, Pan C, Wang Y, et al. Federated deep reinforcement learning for multimedia task offloading and resource allocation in MEC networks. IEICE Transactions on Communications, 2024, E107-B(6): 446-457

[21] Wang Z, Hu Q, Xiong Z, et al. Resource optimization for blockchain-based federated learning in mobile edge computing. IEEE Internet of Things Journal, 2024, 11(9): 15166-15178

[22] Yin H, Wang S, Zhang K, et. al. Research on asynchronous robust federated learning method in vehicle computing power network. Chinese Journal on Internet of Things, 2024, 8(4): 14-22. (in Chinese)
(尹宏博, 王帅, 张科, 等. 车辆算力网络中异步鲁棒联邦学习方法研究. 物联网学报, 2024, 8(4): 14-22.)

[23] Yuan X, Chen J, Zhang N, et al. A federated bidirectional connection broad learning scheme for secure data sharing in Internet of vehicles. China Communications, 2021, 18(7): 117-133

[24] Seid A M, Erbad A, Abishu H N, et al. Multi-agent federated reinforcement learning for resource allocation in UAV-enabled internet of medical things networks. IEEE Internet of Things Journal, 2023, 10(22): 19695-19711

[25] Yang H, Zhao J, Xiong Z, et al. Privacy-preserving federated learning for UAV-enabled networks: learning-based joint scheduling and resource management. IEEE Journal on Selected Areas in Communications, 2021, 39(10): 3144-3159

[26] Liu T, Tang L, Wang W, et al. Digital-twin-assisted task offloading based on edge collaboration in the digital twin edge network. IEEE Internet of Things Journal, 2021, 9(2): 1427-1444

[27] Zhang Y, Hu J, Min G. Digital twin-driven intelligent task offloading for collaborative mobile edge computing. IEEE Journal on Selected Areas in Communications, 2023, 41(10): 3034-3045

[28] Zhao L, Zhao Z, Zhang E, et al. A digital twin-assisted intelligent partial offloading approach for vehicular edge computing. IEEE Journal on Selected Areas in Communications, 2023, 41(11): 3386-3400

[29] He Y, Yang M, He Z, et al. Computation offloading and resource allocation based on DT-MEC-assisted federated learning framework. IEEE Transactions on Cognitive Communications and Networking, 2023, 9(6): 1707-1720

[30] Yuan X, Chen J, Zhang N, et al. Digital twin-driven vehicular task offloading and IRS configuration in the Internet of vehicles. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(12): 24290-24304

YUAN Xiao-Ming, Ph. D., associate professor. Her research interests are key technologies of 5G/6G networks, Internet of vehicles, and edge intelligence.

TIAN Han-Sen, M. S. candidate. His research interests are computing offloading, and multi-agent reinforcement learning.

HUANG Kun-Da, M. S. candidate. His research interests are digital twin, and resource allocation.

DENG Qing-Xu, Ph. D., professor. His research interests are real-time embedded systems, and AIoT.

KANG Jia-Wen, Ph. D., professor. His research interests are metaverse, and IoT.

LI Chang-Le, Ph. D., professor. His research interests are intelligent transportation systems, and wireless sensor networks.

DUAN Xu-Ting, Ph. D., professor. His research interests are the theories of V2X communications and collaborative intelligence.

Background

In the context of edge-side large models, the computational offloading and resource coordination optimization for intelligent inference tasks have broad application prospects. With the rapid development of artificial intelligence and big data technologies, particularly in fields such as the IoMT, smart manufacturing, and intelligent transportation, MEC has become a key technology for enhancing system performance and response speed. MEC enables

中国知网 https://www.cnki.net

the migration of computational resources from data centers to edge devices closer to data sources, thereby reducing delay, minimizing network bandwidth consumption, and improving data processing efficiency. However, as the scale and complexity of edge-side large models continue to increase, efficiently performing computational offloading and resource coordination optimization to meet the real-time, energy efficiency, and security requirements of intelligent inference tasks has become a critical challenge.

Intelligent inference tasks for edge-side large models typically involve significant data processing and complex computational demands, which creates resource bottlenecks for traditional MEC. To address this challenge, computational offloading strategies have emerged. By offloading part of the computational tasks to the cloud or more powerful edge nodes, the computational pressure on local devices can be effectively alleviated. However, the offloading process must contend with issues such as computation delay, bandwidth consumption, and device energy consumption. How to rationally allocate tasks and optimize computational and communication resources has become an important topic in edge computing.

This paper proposes a computational offloading algorithm based on GAN-enhanced MADDPG to address challenges in edge-side large models for intelligent inference tasks. The method optimizes computational efficiency, reduces delay, and lowers energy consumption through effective offloading and resource scheduling. A multi-layered edge computing architecture, utilizing DT technology, combines physical simulation and real-time data feedback to optimize resource allocation and offloading strategies between edge devices and the cloud. By leveraging DRL and multi-agent collaboration, the MADDPG algorithm dynamically adjusts computational load and communication strategies, achieving low latency and high energy efficiency. GAN enhances the algorithm's robustness by simulating complex network environments, enabling better adaptation to real-world uncertainties. Additionally, DT technology provides real-time environmental feedback, optimizing offloading decisions, improving security, and reducing sensitive data transmission to mitigate privacy risks. This solution enhances edge computing performance in large-scale inference tasks, addresses resource configuration issues in dynamic environments, and offers efficient, reliable, and secure solutions for intelligent inference in fields like smart cities, industrial automation, and healthcare.

This work is jointly supported by the National Natural Science Foundation of China (No. 62371116), the State Key Laboratory of Intelligent Transportation System (2024-B008), and the Key Project of Scientific and Technological Research of Higher Education of Hebei Province (ZD2022164).

中国知网 https://www.cnki.net