医疗健康数字孪生网络应用场景及需求

（2022 年）

算网融合产业及标准推进委员会

2022年12月

版权声明

参与编写单位

中国科学院计算技术研究所、中国电信集团有限公司、亚信科技（中国）有限公司、中兴通讯股份有限公司、新华三技术有限公司、中国信息通信研究院

主要撰稿人

史红周、曾辉、吴艳芹、吕田田、杨刚刚、郭建超、谢宝国、李培、段世惠，韩淑君，穆琙博

目 录

图 目 录

1 目标

随着人工智能技术的不断发展，医疗健康行业也积极探索应用人工智能技术以提高生产效能。在国务院颁布的《关于促进和规范健康医疗大数据应用发展的指导意见》的推动下，各大医疗机构积极响应，大力推进医疗数据化和信息化工作，促进了该领域的快速发展。尽管医疗机构在信息化方面取得了显著成果，但在人工智能技术的应用方面，该领域面临以下三个主要挑战：

首先是“数据墙”问题：AI 模型的训练需要大量数据，但单个机构所拥有的数据是远远不足的，因此需要集合大多数医疗机构的数据才能构建尽可能优化的 AI 模型。但是，数据共享难度大，数据集合后的体量也是一个重要难题。

其次是“模型墙”问题：不同的 AI 健康医疗应用场景需要不同的 AI 模型，导致模型种类繁多且各模型之间缺乏联系，难以实现跨场景应用。

最后是“算力墙”问题：随着医疗技术的发展，数据的种类和体量将会不断增大，这最终将会体现在计算资源的消耗上。另外，医疗数据需要隐私性和安全性的保证，隐私计算和加密计算也将进一步增大医疗健康领域的计算资源需求。根据 OpenAI 的报告[1]，最大的 AI 训练运行中使用的计算量呈指数级增长，加倍时间为 3.4 个月。因此，随着对人工智能能力的不断挖掘，训练模型所需的计算资源也将呈指数增长。

综上所述，人工智能技术的应用给医疗健康行业带来了生产效能的提升，但在实际应用过程中，医疗机构所面临的“数据墙”、“模型墙”和“算力墙”等问题限制了人工智能技术的应用效果。然而，这些问题的本质仍然是医疗领域数字化时代联合协作的问题。数字孪生网络技术的出现为解决这些问题提供了一种新的思路，它可以构建医疗健康行业应用场景的虚拟网络孪生体，并通过多个机构的数据共享、跨场景应用和隐私保护等支持，实现具体医疗健康 AI 技术创新以及医疗健康应用的落地。因此，数字孪生网络技术将成为医疗领域数字化时代的重要推动力量，促进医疗健康行业的快速发展和进步。

2 术语说明

数据模型：孪生网络层功能模块中的模型，它们是应用层各个应用需求的功能映射。当应用层提出需求和初始配置信息时，数据模型接收应用需求并利用功能模块中的辅助模型在数字孪生网络中构建和模拟与应用需求相对应的解决方案。

辅助模型：孪生网络层中功能模块中的模型，它们作为数据模型功能的扩展。数据模型负责应用需求的主要功能，辅助模型负责应用需求的附加功能。

策略配置：孪生网络层与物理网络层之间的数据交互。不同于常规的数字孪生网络，物理网络层中的各个网络节点（医疗机构）都有绝对的独立自主权，孪生网络层无权修改机构的配置信息。因此，孪生网络层得到的优化配置只能以建议的形式推送到物理节点，并且由

机构自身决定是否采用。

模型融合：一般的模型融合是将多个不同的机器学习模型的预测结果结合起来，以提高整体预测的准确性和稳定性的技术。另外，联邦学习也能够看作是一种模型融合的变体，它通过将多个参与方的本地模型进行融合，从而构建一个更加准确、安全和隐私保护的全局模型。

模型即服务：将训练完成的模型作为服务提供者，对外提供预测推理服务。通过对外提供计费网络 API 的方式，接收服务使用者发来的数据特征，然后将数据特征输入模型以得到预测（推理）结果，最后将预测结果返回给服务使用者。

模型安全：指保护机器学习模型的完整性、可用性和保密性，以防止针对模型的恶意攻击和滥用。

数据隐私保护：是采取各种技术手段，以保护个人或机构数据的机密性、完整性和可用性，防止数据泄露或滥用的过程。

3 范围

现在，医疗机构正在积极地利用 AI 技术的发展机遇开发各自的医疗健康应用。然而，由于数据和算力等方面的限制，早期的“闭门造车”发展模式已经无法满足更高精度模型的需求。因此，多机构合作开发大型医疗模型成为医疗健康应用的发展趋势。数字孪生网络为多机构之间的合作提供了理论基础，使得参与合作方能够利用虚拟网络进行快速且低成本地构建医疗应用大模型以及尝试构建新型合作

模式。

本文针对医疗健康应用进行研究和分析，重点对数字孪生网络对医疗健康应用的需求、架构和关键技术进行深入探讨。主要研究内容包括：

- 应用场景：研究适用于医疗健康数字孪生网络的应用场景；
- 需求分析：医疗健康应用对数字孪生技术的需求分析；
- 关键问题及技术：描述实现医疗健康数字孪生网络架构所面临的难题以及现有的技术。

4 研究背景

4.1 数字孪生

“数字孪生”一词最初由 Grieves 在 2003 年的一次演讲中提出，并被记录在一份白皮书中，为“数字孪生”概念的发展奠定了基础[2]。2012 年，美国国家航空航天局（NASA）发布了一份题为“未来 NASA 和美国空军飞行器的数字孪生范式”的文件[3]，其中定义“数字孪生”为：“数字孪生是一个综合的多物理、多尺度、概率模拟的成车或系统，利用可用的物理模型、传感器更新、行车历史等来反映其相应的飞行器孪生体的生命周期。”数字孪生包括三个方面：物理对象、虚拟对象和交互数据。

在数字孪生中，物理对象是指与数字孪生相关的实际物理系统或实体，而虚拟对象则是与之对应的数字化的、基于模型的系统或实体。

交互数据是物理对象和虚拟对象之间的数据传输和交互。数字孪生可用于物理对象的建模、仿真、监测和控制，并可在其生命周期中提供数据驱动的支持，从而提高其效率和性能。综合来看，数字孪生技术为各种物理系统的仿真、监测和控制提供了新的解决方案，并在许多领域具有广泛应用前景。

4.2 医疗健康研究现状

随着人工智能技术的深入研究，越来越多的行业将其纳入战略发展规划之中。这一趋势不仅推动着产业进步，也对社会变革起到了重要作用。在医疗健康领域，保障人民健康是维护社会稳定的重要因素之一，因此需要不断提升自身能力以满足人们对健康生活的向往需求。人工智能技术因其高效性和准确性被视为实现这一目标的重要手段之一，因此其研究与应用数量呈指数形式增长，也助力了医疗健康领域的进步。

从医疗健康领域的角度来看，人工智能的应用可分为“临床诊断中的应用”、“临床决策中的应用”和“临床治疗中的应用”[4]。目前，临床诊断方面的研究较为广泛，涉及各种需要利用医疗图像进行疾病诊断的应用场景，例如基于胸部 CT 的辅助肺癌诊断[5]，以及利用内窥镜 TM 图像辅助耳部感染诊断[6]等。在临床决策方面，研究人员还探索了根据患者的 HER 以及其他信息进行患者住院决策的方法[7]。在临床治疗方面，研究者们也在积极探索手术辅助智能[8]和化疗辅助智能[9]等应用。

从人工智能的视角考虑，医疗健康领域的应用可以根据所用的训

练集数据划分为不同类型，包括文本、图像、视频和音频等数据。随着以卷积神经网络为主的各种图像人工智能模型的快速发展，医学图像诊断应用在医疗领域中得到了广泛应用，如肺癌诊断[5]、耳部感染诊[6]、皮肤病诊断[10]等。而以文本为主要数据的人工智能应用则包括住院决策[7]和患者生存状态预测[11]等。另外，以音频为主要数据的人工智能应用也被广泛研究，例如心脏病诊断[12]。

综上所述，医疗领域的人工智能研究已有许多进展。但是在采用特定疾病诊断等人工智能模型时，不同机构需要考虑模型的准确率等因素。由于所拥有的数据分布不同，不同的数据分布也会导致模型在不同数据上的准确率等因素有所不同。这种情况导致各机构对于相同疾病所使用的人工智能模型存在差异，而且模型的推理能力也受训练数据量的限制。然而数据孪生网络能够帮助各机构之间构建虚拟信息交换网络，通过在虚拟网络中尝试各种模型融合和数据交换创新模式，从而使各机构能够获得具有更强推理能力的医疗模型。

4.3 数字孪生网络

根据 ITU-T 的《Digital Twin Network - Requirements and Architecture》项目，该项目定义了数字孪生网络[13]，它是“一个物理网络的虚拟表示形式，有助于对物理网络进行分析、诊断、模拟和控制，实现物理网络和虚拟孪生网络之间的实时互动映射”，其中涵盖了数据、模型和接口。该定义主要适用于物理网络，即承载互联网的底层网络，由各种物理设备（如主机、路由器、交换机等）和介质（光缆、电缆等）连接起来形成的网络。然而，Wu 在其数字孪生网

络综述文献[14]中提出数字孪生网络的定义更倾向于广义的网络，他认为数字孪生网络是“由多个一对一的数字孪生构建的多对多的映射网络，即数字孪生网络使用先进的通信技术来实现物理对象与其虚拟孪生体、虚拟孪生体与其他虚拟孪生体、物理对象与其他物理对象之间的实时信息交互”。在这个定义中，物理对象和虚拟孪生体可以相互沟通、协作、共享信息、完成任务，并通过连接多个数字孪生节点形成一个信息共享网络。

在医疗健康数字孪生网络中，各孪生体以及物理对象之间的信息交互体现了 Wu 提出的数字孪生网络的定义[14]。各物理节点之间的通信在虚拟孪生体中模拟，并且各个虚拟孪生体之间也会由于模型的要求进行必要信息的传输。医疗数字孪生网络对代表各医疗机构的网络节点、以及连接节点之间的底层物理网络设备、各机构内计算节点等基础设施进行孪生，使得能够以数字的形式显示出各虚拟孪生体的状态信息并且同时能够采集其中的数据。

4.4 医疗健康与数字孪生网络

数字孪生网络为医疗健康行业带来了新的可能性，尤其是在医疗机构之间的联合协作和联合诊疗方面。目前，由于医疗机构之间的数据孤岛和技术壁垒等问题，医疗机构很难进行大规模的人工智能应用和联合协作。而数字孪生网络可以将不同医疗机构之间的数据和信息进行整合和共享，实现联合协作和联合诊疗。

数字孪生网络不仅可以为医疗机构提供更加精准、高效的医疗诊疗服务，同时也可以为医疗健康行业带来更多的商业机会和发展前景。

随着数字孪生网络技术的不断发展和应用，未来数字孪生网络将会在医疗健康行业中将会发挥越来越重要的作用，推动医疗健康行业的数字化转型和升级。

5 医疗健康数字孪生网络架构

参考 ITU-T 《Digital Twin Network - Requirements and Architecture》，医疗健康数字孪生网络的架构仍然是三层结构，分别是物理网络层、孪生网络层以及网络应用层。在孪生网络层，根据医疗健康的特点做出了一些独特的修改，具体内容如下：

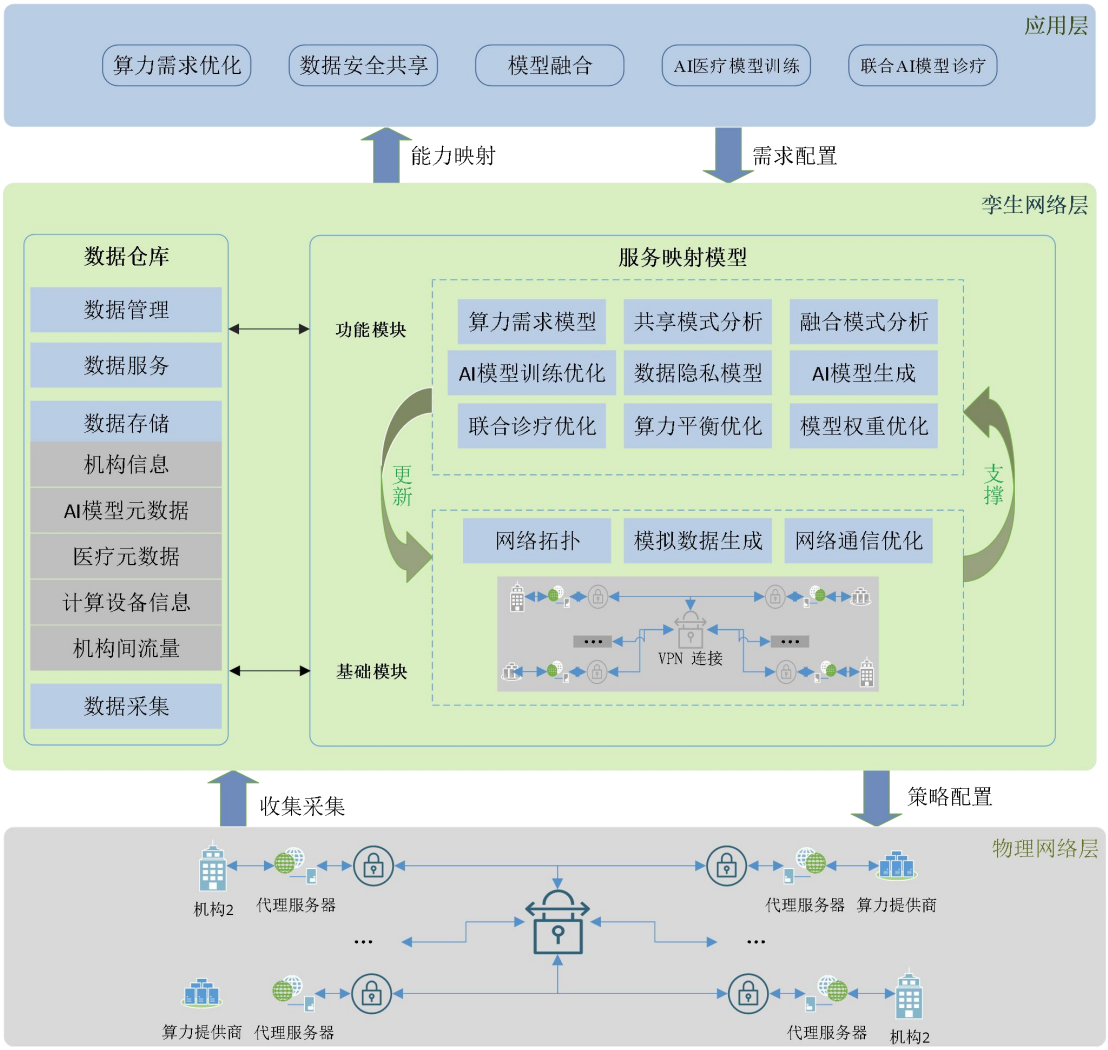

图 1 医疗健康数字孪生网络应用场景架构图

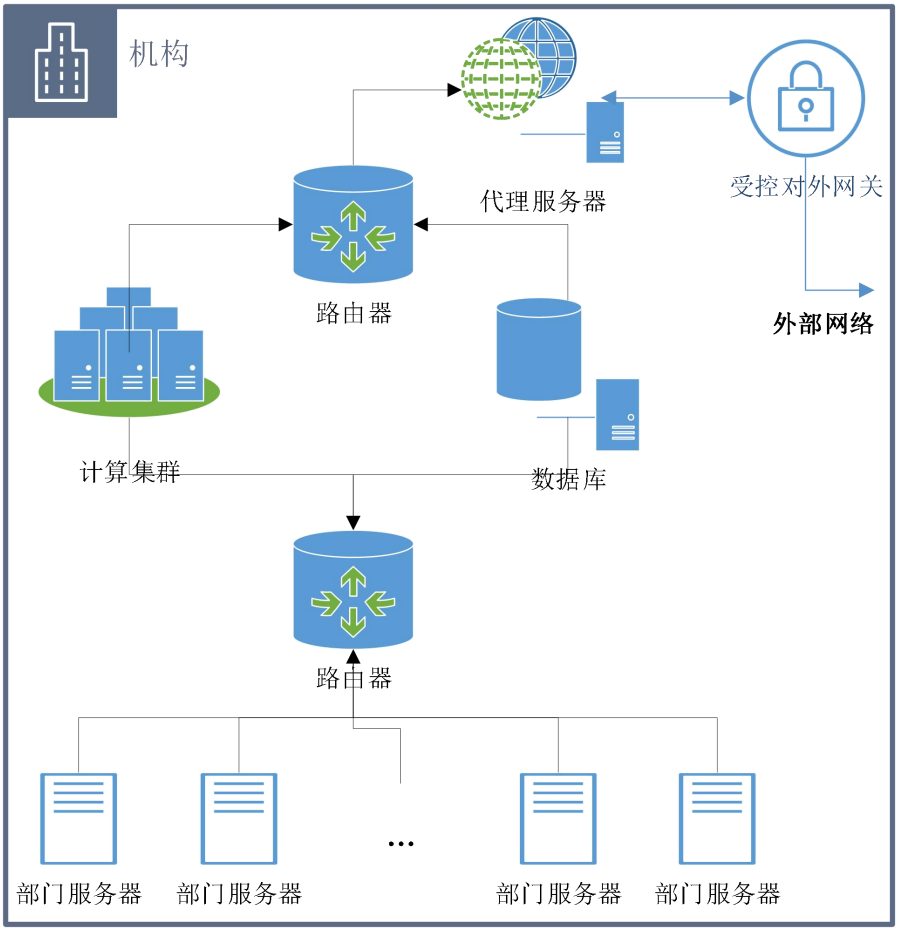

图 2 机构内部对象关联关系示意图

5.1 物理网络层

物理网络层由各医疗机构实体和算力提供商实体组成，并且这些实体通过 VPN 链接，如图 1 所示。该网络连接方式确保了网络中各机构之间能够实现机密通信，并且能够进行统计机构之间的网络流量等信息交换。图 2 展示了机构内部实体之间的关联，机构在该物理网络中发送的信息都需要经过代理服务器和 VPN 网关。其中，代理服务器承担“数据过滤网”的作用，所有代理服务器都采用相同的、经过所有参与方共识的数据披露策略检查对外发送的数据。所谓的数据披露策略本质上是确定通信数据的格式、可公开的医疗数据范围、医疗

数据元数据的定义以及可公开本地网络设备信息范围等其他在孪生网络层使用到的可公开数据定义。

5.2 孪生网络层

架构中间层为孪生网络层，该层是医疗健康数字孪生网络应用场景架构中的核心。它由两个子系统组成：数据仓库和服务映射。

1)数据仓库：不仅需要收集物理网络层中与算力、AI 医疗模型、医疗数据相关的信息，还需要实时更新物理网络层中机构之间网络流量相关信息。并且还要有提供信息整合能力为服务映射中的模型提供准确和完备的信息。数据仓库主要功能如下：

- 数据采集：使用 ETL 技术从物理网络层获取数据并进行数据清洗。
- 数据存储：根据数据类型，利用不同数据库技术对数据进行高效存储。其中数据通常为服务映射中模型所需的基础数据，主要包括：网络流量数据（延迟、丢包、数据包大小等信息）、计算设备信息（GPU、内存、GPU 以及缓存等信息）、医疗元数据（医疗数据的结构、类型、格式、编码等）、AI 模型元数据（架构、性能、训练和优化、应用场景等）、机构信息（名称、类型、级别、IP 地址等）。
- 数据服务：为服务映射中的模型提供统一的数据访问接口，例如快速检索、条件检索、批量服务等。
- 数据管理：主要包括数据安全管理、数据质量管理以及元数据管理。

2)服务映射：其中的上层功能模型通过数据仓库和应用层的需求构建模型实例，模型实例再利用下层的基础模型实现数字孪生网络中的模拟、预测、优化、配置和验证。在数字孪生网络中验证了有效性后再将策略或配置发送到物理网络层中的各个机构。在机构的代理服务器中设置了两种策略配置接口，一种是通过预先设定的策略过滤规则自动决定策略在本地实施；另一种接口则暂存接收到的所有策略配置信息，再由专门的人员进行在线配置。

- 基础模型是将物理网络中的基本配置、网络环境、运行状态、链路拓扑等信息实现虚拟化，构建与物理网络层对应的孪生网络层。该模型能够辅助模拟和验证各种医疗健康领域中应用场景需求的解决方案，以确保方案在物理机构网络之间的有效性和可靠性。
- 功能模型包括基于数据和应用需求建立的数据模型以及支撑这些数据模型的辅助型模型。数据模型包括算力需求模型、共享模式分析、融合模式分析、AI 医疗模型训练优化以及联合诊疗优化等。辅助型模型包括数据隐私模型（数据加密、数据脱敏、数据访问、匿名化等）、AI 模型生成（决策树、神经网络）、算力平衡优化（负载均衡、资源预测、资源池化）以及模型权重优化（梯度优化、误差加权、交叉验证、Stacking）等。

5.3 应用层

架构顶层是应用层。该层将应用需求和对应的初始配置信息输入

到孪生网络层的数据模型并在孪生网络中实例化相关服务。待通过孪生网络层的验证和优化后，孪生网络层将优化方案的相关策略和配置下发到物理网络层的各个机构中。算力需求优化、数据安全共享、AI 医疗模型训练、联合 AI 模型诊疗等医疗健康在数字孪生网络中的应用能够在多个医疗机构中低成本、高效率的部署，并且能够尽可能减小对机构原有业务的影响。

6 医疗健康数字孪生网络应用场景需求

6.1 算网需求描述

6.1.1 场景描述

医疗健康领域的计算能力是实现医疗健康目标的基础，包括数据收集、聚合、清洗、隐私计算、AI 医疗模型训练和推理等计算过程。随着区块链[16]、联邦学习[19]和同态加密技术的不断发展，医疗领域数据孤岛的问题逐渐得到缓解。但是，处理海量且复杂的医疗数据需要大量的计算资源，因此多节点并行计算已成为必然选择。此外，在多机构合作的环境中，不同机构的计算过程和计算资源需求也可能不同。因此，医疗健康数字孪生网络需要提供算力需求描述模型，标准化度量各机构不同计算过程所需的算力资源，并提供算力资源交易服务，以充分利用各机构中的闲置资源。机构可以发布无法满足的算力资源需求，获取其他机构或第三方算力资源供应商提供的“算力商品”。

网络通信的要求是数字孪生网络在医疗健康领域中的另一个重要问题。在医疗健康领域，数字孪生网络需要提供稳定的网络通信支持，确保医疗机构之间的数据和信息能够稳定地传输和共享。上述算力资源交易和后续将提到的数据安全共享以及联邦学习，对网络通信有着较高的要求。因此医疗健康数字孪生网络需要网络需求描述模型，该模型不仅能够在孪生网络模拟过程中评估应用所需的基本网络需求，还能在现有的网络基础上进行优化。

6.1.2 算网需求描述对数字孪生网络的技术要求

算网需求描述模型需要网络通信优化模型以保证数据在网络中高效传递。为了实现这一目标，网络拓扑模型需要将物理网络中所有设备及其连接关系和状态数字化，以便实时反馈网络整体状态。

实现算力需求描述模型需要与各种计算过程相关的数据，例如计算对象基础运算的拆分、基础运算的顺序、计算过程所需的时间以及计算所应用的硬件设备频率等信息。通过这些数据建立计算过程与所需计算资源的对应关系，使得各机构能够利用该模型度量特定计算过程的算力需求描述。

6.2 数据安全共享

6.2.1 需求描述

随着医疗机构信息化的不断推进，各机构内部已逐渐构建了电子健康记录系统，其中存储着具有一定价值的医疗数据。各医疗机构也积极地开展了针对医疗数据的数据价值挖掘研究。由于医疗数据种类

繁多，既有文本类数据，又有图片、音视频类数据，利用传统的机器学习算法已经很难满足智慧医疗的需求。因此，目前大多数研究已经转向以复杂程度更高的深度学习算法为代表的一系列人工智能技术，即医疗健康。然而，对单个医疗机构而言，其拥有的数据量不足以支持深度学习算法模型训练到期望的准确度，因此多个组织之间的信息共享成为了开展医疗健康工作的前提条件。目前，数据共享还停留在自主收集、无偿提供、自愿公开三种方式，这几种方式具备极其不稳定的缺点。

当前医疗数据共享性差的 4 点原因为：无法确定数据集合的主导权、数据泄露风险大、难以精准授权以获得收益、缺乏激励机制[15]。虽然现今有多种研究性的医疗数据共享方式，但由于各医疗机构对于研究成果的质疑，大多数研究性的成果都没有能够实现落地实施。而医疗健康数字孪生网络能够基于多机构数字孪生网络构建各种创新性数据共享模式的虚拟实现。在模拟创新性的共享方案时，需要由隐私计算模型对各类传输数据进行加密，保证数据的隐私性。还需要加入利益计量模型，对各参与机构进行利益最终分配，得到各参与机构预期的最终经济效益信息。

数据共享离不开各机构之间的信息的交换，因此医疗健康数字孪生网络还需要保证数据在各机构之间能够高效传输。由于许多种类的医疗数据受到精度的影响，传输的数据体量大种类多，因此保证数据在各组织之间高效传递也成了其中的一项需求。这需要使用到常见的数字孪生网络的功能——网络流量路由（路径规划）。利用从物理设

备实时收集到的数据进行机器学习并预测，从而实现智能化的网络流量规划，减轻在传输大体量数据时网络拥塞的情况。

6.2.2 数据安全共享对数字孪生网络的技术要求

高效数据传输的要求：利用历史网络流量数据训练网络路由优化模型，能够实时同物理网路节点设备获取网络流量数据（设备缓存、带宽、节点吞吐量等）并进行推理获取最优网络路由。

数据安全传输的要求：支持各种数据加密算法，例如 SHA1、MD5、RSA 等。能够结合各种医疗数据的基本特征（bit 长度）、数据加密后的特征以及加密算法等因素构建模拟加密模型，能够直接根据不同的医疗数据类型和加密算法直接生成孪生网络层应用的加密数据。

构建共享模式的要求：目前，区块链是医疗数据共享研究的基础[16]、[17]、[18]。因此，数字孪生网络需要能够根据应用层的输入构建特定类型的区块链医疗数据共享方案。

6.3 AI 医疗模型构建

6.3.1 场景描述

医疗健康数字孪生网络的核心功能是为医疗机构提供更优质的 AI 医疗模型构建服务，并整合机构的 AI 医疗服务能力，从而提高社会医疗服务水平，为人民的健康生活提供基础的保障。该网络从两个方面提升 AI 医疗模型。首先，医疗健康提供以深度学习为代表的一系列机器学习技术支持，以支持医疗模型的构建。其次，数字孪生网络为医疗模型构建提供基础的数据、算力和网络通信服务，进一步提

升 AI 医疗模型的可用性。以下是医疗健康数字孪生网络在 AI 医疗模型构建上的应用场景描述。

为了构建更优秀的 AI 医疗模型，各机构都会对一些常用的深度神经网络进行不同程度的改进，但这些改进后的模型结构是机构机密，对于公共的孪生网络层是不可见的。但是，机构可以在网络应用层输入公开的基础深度神经网络（例如 DenseNet、VGG 等）、训练迭代次数、医疗数据类型等信息，孪生网络层会根据网络应用层的输入信息预估所需的算力资源，同时提供医疗数据来源机构的推荐。医疗机构可以根据孪生网络层提供的算力资源信息以及数据来源推荐等反馈信息规划 AI 医疗模型的构建工作。

6.3.2 AI 医疗模型构建对数字孪生网络的技术要求

模拟医疗数据生成的要求：由于医疗数据具有很强的隐私性，作为公共的孪生网络层，不适合将真实的医疗数据用于算力资源描述，因此需要有医疗数据生成模型。该模型能够根据所需的医疗数据类型模拟其数据特征，并直接生成虚构的医疗数据，保证模拟数据和真实数据在长度等方面一致。

基础深度神经网络生成：首先需要准备各种基础的深度神经网络结构，并且能够根据网络应用层的需求构建符合输入输出要求的指定神经网络结构。

神经网络训练算力评估模型：根据模拟神经网络训练所消耗的算力，推测特定轮次训练所需的算力资源。

6.4 模型融合模式创新

6.4.1 场景描述

在通用模型技术尚未完全成熟的情况下，对于各种特定领域下的智能医疗诊断模型，各医疗机构都是用对应疾病的数据进行 AI 建模。由于数据类型的不同（例如视频、音频、文本、图片等数据），再者针对相同类型数据建模的 AI 模型具体结构层次也不相同（例如针对图像数据的卷积神经网络模型，仅文献研究就涉及 DenseNet、R-CNN、VGG-16、ResNet 等多种主流的处理图像的深度神经网络），这导致各医疗机构采用各自的方式实现医疗服务模型。对于这些机构中存在的各式各样的 AI 医疗模型，如何进一步挖掘其中的医疗价值也是医疗健康数字孪生网络需要解决的问题。

模型即服务是一种可能的解决方案，利用医疗健康数字孪生网络为各机构提供展示 AI 医疗服务模型功能的接口，同时还需要提供模型融合策略服务，该服务对应于模型融合数字孪生体，模型融合策略服务能够自动地为用户提供最佳的模型选择以及权重分配，并且尽可能保证融合模型在效率和经济收益上达到最优值以尽量满足用户期望。

医疗健康数字孪生还需要提供“模型动”、“数据动”的动态解决方案。不同的 AI 医疗服务模型由于结构不同，参数大小也各异，而对于使用 AI 医疗模型服务接口的用户来说，使用模型时数据量大小也有所不同。当模型大小和数据量达到一定规模时，考虑是传输网络

模型还是数据是非常必要的，因为不恰当的选择将会影响服务的使用效率。此外，不同的模式选择还需要相应的加密方案，即加密模型和加密数据。模式的选择需要考虑到加密后模型和数据两者体积的变化。

除了上述提出的利用基础的模型融合方式解决医疗健康中“模型墙”问题的方式外，目前联邦学习也是医疗健康中训练 AI 模型的一种主要研究方向，联邦学习的研究遍布医疗领域中的各个方向[20]。联邦学习的主要特点在于数据不离开本地，模型训练的主要工作也分布在各个节点，这同样对本地计算能力有较高的要求，而模型训练过程中也伴随着大量关于梯度值的通信。要使得这种新型的理论研究方式实现落地还有一段距离，但是医疗健康数字孪生网络能够提供验证的环境，包括各机构相互连接的网络模拟，中央节点的模拟，数据传输的模拟以及本地模型训练的模拟。在孪生网络层的融合模型模式中构建联邦学习的模拟机制，当接收到网络应用层输入特定需求时，构建对应的模型，实现新型模型融合创新技术的验证。

6.4.2 模型融合模式创新对数字孪生网络的技术要求

数据隐私的要求：上述提到的两种模式中，都要考虑数据隐私，数字孪生网络层需要支持加密算法、脱敏以及匿名等技术要求。

模型安全的要求：在模型即服务的模式中，需要考虑“数模调度”，其中“模型动”需要保证模型结构与参数的安全，使用者只能得到推理结果，无法获得模型的信息。

构建参数预测模型的要求：由模型即服务模式的要求，需要能够从各机构代理节点获取模型融合期望数据、加密的用户测试数据集等

数据，为机构选择最优的模型组合以及各模型的权重占比。

构建联邦学习网络的要求：由联邦学习这类模型融合模式，数字孪生网络能够从网络应用层获得具体的联邦学习构建请求以及参数，在虚拟网络中构建各机构互联的联邦学习网络。

7 医疗健康数字孪生网络应用场景关键问题及技术

7.1 概述

应用在医疗健康领域的数字孪生网络技术不仅保障了医疗数据在机构间的网络中能够安全高效的传输，还为各医疗机构提供了AI医疗模型服务接口以及为之服务的算力需求描述功能。

数据高效安全的传输为解决医疗健康领域中的“数据墙”问题奠定了结实的硬件基础，能够支撑起各种医疗数据共享方案的模拟实施。其采用经典的数字孪生网络技术，通过硬件设备通信数据的采集等信息实现提前预测、事先规划，从而提高网络整体的通信效率。对于数据的加密，只需要通过隐私计算孪生体模拟数据的加密形式即可，无需参与真正的加密计算，这满足了数据的高效安全传输要求。

AI医疗模型即服务为医疗健康领域中“模型墙”问题提供了功能性的解决方案，其中涉及的数字孪生网络偏向于广义上的网络。虽然在模型即服务的应用场景架构中，也利用了经典数字孪生网络中常见的网络流量规划等典型用例，但该应用场景的重点在于融合模型的选择和参数的预测。由于随时可能有新的机构加入或退出网络，现有机

构也在不断更新AI医疗模型，因此需要一个能够实时计算最优组合的模型，以承担模型选择和参数预测的工作。

算力需求描述不仅保障了数据高效安全传输中隐私计算的算力需求，还保证了机构AI模型训练以及模型融合所需的算力。此外，还可能存在机构内部计算资源不足而向外部调用算力的情况。这涉及到模型和数据的隐私计算以及在隐私计算下额外的算力需求额度计算。

医疗健康数字孪生网络虽然为机构间合作共建大型AI医疗模型提供了理论框架，但是其中的具体实现还存在一些难题需要攻克。以下将描述医疗健康数字孪生网络架构中存在的关键问题以及可能应对的解决技术。

7.2 建立医疗机构间的共识

传统的合作需要在互相信任的基础上进行，但是在医疗健康数字孪生网络提出的架构中，任何具备资质的医疗机构都可以参与共建AI医疗模型的进程。而且，这些机构之间可能互不相识，甚至还存在竞争关系。因此，医疗健康数字孪生网络需要提供机制来建立机构之间的共识，以保证后续的合作进程能够顺利进行。

在架构的物理网络层提及的数据披露策略的制定是机构间建立共识的一个具体体现。数据披露策略涉及到不同孪生层功能的数据，例如数据共享模式中的医疗数据、算力需求描述模型中的设备信息以及这些数据的元数据，即数据格式。如果没有数据披露策略的制约，不同机构按照自己的习惯披露数据范围以及发送不同格式的数据，这

将导致数字孪生网络层无法构建有效的功能模型。因此，在机构之间建立共识是整个架构中的关键问题。

区块链作为一种分布式账本，对其上的记录需要经过所有成员的验证，且记录无法篡改。区块链的本质就是为互不信任的参与方创建共识，因此区块链能够作为解决上述问题的一种技术。以数据披露策略为例，在保证数据格式一致和数据隐私的前提下，向区块链提出机构各自的策略条例，只有绝大多数的成员验证通过的条例能够登记在区块链中。然后所有的参与方的代理设备都从区块链中下载统一的数据披露策略来验证所有出入数据。这就保证了所有参与方能够得到可信的统一的策略。此外，它还能够发现不符合策略的数据。

7.3 保护医疗机构数据隐私

随着公众对隐私保护意识的提高和政府部门对医疗数据隐私与安全指导的明确，医疗机构越来越保守地使用数据。作为一种突破行业“数据墙”的技术，医疗健康数字孪生网络需要在数据隐私方面提供相应的方案。机构在数据共享和模型融合过程中需要调用真实的医疗数据，但由于这些数据可能被传输到其他机构，因此需要进行加密保护。然而，在验证不同的数据共享创新技术和模型融合方案时，全程加密可能会影响验证效率。因此，快速模拟加密数据在数字孪生网络中成为一个必须解决的关键问题。本文提出以下解决框架：

不同的加密算法能够将原始数据变换成具备特定数据结构的加密数据。因此，可以利用不同算法和不同长度的原始数据对应的加密数据结构这些数据构建一个模型，它能够映射加密算法和加密数据结

构之间的关系。模型的输入是医疗数据的元数据（加密算法类型、原始数据长度），模型根据对应的加密数据结构直接生成类似的加密数据。并且数字孪生网络还需要存储所有输入的元数据以及对应生成的加密数据，以供解密端使用。在孪生网络验证新型技术时，机构只需要使用医疗数据的元数据就能使数字孪生网络的模拟过程顺利运行。

7.4 保护医疗机构模型安全

在前文的模型融合模式中提到“模型动”的可能，已经训练好的模型属于医疗机构的数字化资产。由于AI模型数字化的特性，其参数在网络中传输可能会被窃取或篡改。AI模型参数被窃取将导致机构资产受到损害，而参数被篡改也可能会对使用方造成潜在的风险。因此，保护模型安全也是医疗健康数字孪生网络面临的关键问题。

同态加密支持数据在加密态进行计算，得到的结果经过解密后正是原始数据进行相同运算的结果。这项技术无论是运用在“模型动”还是“数据动”的方案上，都能保护机构的模型安全和数据隐私。例如在“模型动”方案中运用多密钥同态加密技术，首先使用同态加密算法将需要传输的模型参数加密，AI模型的计算也服从同态加密算法，得到的预测结果需要拥有私钥的模型拥有者才能解密，因此保证了模型安全且限制了加密模型的使用能力。

7.5 建立算力需求描述模型

当下人工智能发展趋势之一是构建更大更全面的模型，在医学领域也开始采用人工智能技术关注不同疾病之间的关联，例如与糖尿病

有关的眼部疾病的检测。由于OpenAI指出大型人工智能模型的算力需求目前呈指数形式增长，因此算力保障是医疗健康发展的重要因素。而算力需求描述作为算力保障的第一步，对于整个算力保障过程也是至关重要的。

对事物的描述都少不了度量单位，而用于算力需求描述的度量自然也是必不可少的。由于“”医疗健康是以构建大型AI医疗模型为中心的，因此算力的度量可以采用OpenAI提出的单位，即 petaflops/s-day，它表示以每秒执行10^{15}次神经网络操作维持一天的计算量（10^{20}次）。OpenAI定义神经网络中一次乘法或者加法为一次操作。以此为基础，通过拆分不同AI模型以计算总体操作量，就能够得到大致的算力需求总量。配合不同服务器的硬件设备频率等信息，模型能够得到更为精确的算力需求描述。

7.6 建立算力资源交易网络

算力保障关键步骤在于充分利用机构中的闲置计算资源。当机构通过算力需求描述模型得到算力需求总量高于机构自身计算能力上限时，需要能够获得外部的援助来保证所有计算过程的稳定运行。因此建立算力资源交易网络，为机构之间互相交换算力资源提供基础设施。目前能够采用众包的形式，各机构利用统一的算力度量单位向网络中发布计算委托，而有空闲计算资源的机构能够接取其中的委托，最后以智能合约保障双方的利益。另外如果没有额外的医疗机构接取计算委托，算力供应商能够提供足够的算力资源。

7.7 建立数据安全共享实验框架

数据安全共享创新验证，作为医疗健康数字孪生网络架构的主要应用之一，它能够为为机构探索出行之有效的数据安全共享方案。通过将文献中的理论研究带入到虚拟孪生网络层中验证，比较不同数据安全共享研究方案在具体的应用场景下的优劣，从而选择最适合的方案。

由于目前大多数医疗领域的数据安全共享研究都采用区块链为基础技术。当孪生网络层的数据安全共享模式接收到来自网络应用层的验证数据共享方案的需求时，它结合内置的区块链构建模块和来自应用层的输入参数构建虚拟区块链，模拟理论方案中以区块链为基础的数据安全共享过程。除了模拟全过程的数据共享方案，还需要有统一的评价模型来区别不同方案的优劣。评价模型统计不同方案的网络流量、计算资源开销、时间等数值，最终以可视化的形式展示给机构，机构根据这些数据选择符合其条件的优势方案。

7.8 建立模型融合实验框架

利用数字孪生网络低成本验证理论方案，模型融合实验框架的建立为不同模型融合方案在数字孪生网络中模拟提供了基础设施。目前成体系的相关研究是联邦学习技术，该项技术不仅能够保证参与训练模型的隐私，还能充分地融合不同机构数据特征从而构建泛化能力更强的AI医疗模型。因此该框架要求孪生网络在接收来自网络应用层验证模型融合的请求和相关参数后，孪生网络层能够构建联邦学习的模

拟运行机制，以达到验证不同联邦学习方案的目的。同时框架也需要具备评价模型，该模型能够根据机构关注的指标进行可视化展示，以供机构选择合适的模型融合方案。

7.9 建立联邦学习应用框架

由于各医疗机构仅通过代理节点与其他机构通信，这种通信方式虽然保护了机构内部的信息安全，但是也使得构建公共智能模型变得更加困难，例如算力需求描述模型的构建就可能需要使用到机构内部计算机硬件运行状态等信息。联邦学习技术正好能够在保证各参与方数据隐私的同时构建AI模型。利用联邦学习技术，机构能够在孪生网络层联合构建可用的AI模型为网络应用层提供服务。

联邦学习技术也是模型融合模式的一种形式。各机构利用自身的医疗数据构建统一的本地模型，然后提交到中央节点聚合得到兼备各方数据特征的AI模型，各机构再利用聚合后的模型进一步在本地训练。如此循环，最终能够得到泛化性极高的AI模型。而联邦学习技术在医疗机构中的应用涉及到机构本地算力、网络流量、数据隐私等方面的问题，这些问题能够在孪生网络层模拟而具体地显现。因此建立联邦学习应用框架也是实现医疗健康数字孪生网络的关键问题。

参考文献

[1] “AI and Compute,” *OpenAI*, May 16, 2018. https://openai.com/blog/ai-and-compute/ (accessed Oct. 12, 2022).

[2] M. Grieves, “Digital Twin: Manufacturing Excellence through Virtual Factory Replication,” Mar. 2015.

[3] E. Glaessgen and D. Stargel, “The digital twin paradigm for future NASA and U.S. air force vehicles,” Apr. 2012. doi: 10.2514/6.2012-1818.

[4] 王家庆 and 王光锁, “人工智能在医学中的应用进展,” 山东医药, vol. 61, no. 04, pp. 112–115, 2021.

[5] A. Masood *et al.*, “Computer-Assisted Decision Support System in Pulmonary Cancer detection and stage classification on CT images,” *Journal of Biomedical Informatics*, vol. 79, pp. 117–128, Mar. 2018, doi: 10.1016/j.jbi.2018.01.005.

[6] M. A. Khan *et al.*, “Automatic detection of tympanic membrane and middle ear infection from oto-endoscopic images via convolutional neural networks,” *Neural Networks*, vol. 126, pp. 384–394, Jun. 2020, doi: 10.1016/j.neunet.2020.03.023.

[7] “A Machine Learning Approach to Predicting Need for Hospitalization for Pediatric Asthma Exacerbation at the Time of Emergency Department Triage - Patel - 2018 - Academic Emergency Medicine - Wiley Online Library.” https://onlinelibrary.wiley.com/doi/10.1111/acem.13655 (accessed Jul. 26, 2022).

[8] Y. Mintz and R. Brodie, “Introduction to artificial intelligence in medicine,” *Minim Invasive Ther Allied Technol*, vol. 28, no. 2, pp. 73–81, Apr. 2019, doi: 10.1080/13645706.2019.1575882.

[9] H. G. Moussa, G. A. Husseini, N. Abel-Jabbar, and S. E. Ahmad, “Use of Model Predictive Control and Artificial Neural Networks to Optimize the Ultrasonic Release of a Model Drug From Liposomes,” *IEEE Transactions on NanoBioscience*, vol. 16, no. 3, pp. 149–156, Apr. 2017, doi: 10.1109/TNB.2017.2661322.

[10] A. Romero Lopez, X. Giro-i-Nieto, J. Burdick, and O. Marques, “Skin lesion classification from dermoscopic images using deep learning techniques,” in *2017 13th*

IASTED International Conference on Biomedical Engineering (BioMed), 2017, pp. 49–54. doi: 10.2316/P.2017.852-053.

[11] Y. Wang *et al.*, “Mortality Prediction in ICUs Using A Novel Time-Slicing Cox Regression Method,” *AMIA Annu Symp Proc*, vol. 2015, pp. 1289–1295, 2015.

[12] S. Shinde and J. C. Martinez-Ovando, “Heart Disease Detection with Deep Learning Using a Combination of Multiple Input Sources,” in *2021 IEEE Fifth Ecuador Technical Chapters Meeting (ETCM)*, 2021, pp. 1–3. doi: 10.1109/ETCM53643.2021.9590672.

[13] “Y.3172 : Architectural framework for machine learning in future networks including IMT-2020.” https://www.itu.int/rec/T-REC-Y.3172-201906-I/en (accessed Oct. 14, 2022).

[14] Y. Wu, K. Zhang, and Y. Zhang, “Digital Twin Networks: A Survey,” *IEEE INTERNET OF THINGS JOURNAL*, vol. 8, no. 18, p. 16, 2021.

[15] 张振, 杨翠湄, 徐静, 李琳, and 周毅, “健康医疗大数据应用发展现状与数据治理,” 医学信息学杂志, vol. 43, no. 07, pp. 2–8, 2022.

[16] 曹萌, 余孙婕, 曾辉, and 史红周, “基于区块链的医疗数据分级访问控制与共享系统,” 计算机应用, pp. 1–11.

[17] K. Peterson, R. Deeduvanu, P. Kanjamala, and K. Boles, “A Blockchain-Based Approach to Health Information Exchange Networks,” p. 10.

[18] A. Tandon, A. Dhir, A. K. M. N. Islam, and M. Mäntymäki, “Blockchain in healthcare: A systematic literature review, synthesizing framework and future research agenda,” *Computers in Industry*, vol. 122, p. 103290, Nov. 2020, doi: 10.1016/j.compind.2020.103290.

[19] J. Passerat-Palmbach *et al.*, “Blockchain-orchestrated machine learning for privacy preserving federated learning in electronic health data,” in *2020 IEEE International Conference on Blockchain (Blockchain)*, 2020, pp. 550–555. doi: 10.1109/Blockchain50366.2020.00080.

[20] V. A. Patel *et al.*, “Adoption of Federated Learning for Healthcare Informatics: Emerging Applications and Future Directions,” *IEEE Access*, vol. 10, pp. 90792–90826, 2022, doi: 10.1109/ACCESS.2022.3201876.

算网融合产业及标准推进委员会（TC621）

地址：北京市海淀区花园北路 52 号

邮编：100191

电话：010-6230XXXX

传真：010-62304980

网址：www.ccnis.org.cn