

UNICON: 统一的医学基础模型持续学习

Mohammad Areeb Qazi, Munachiso S Nwadike, Ibrahim Almakky,
Mohammad Yaqub and Numan Saeed

Mohamed bin Zayed University of Artificial Intelligence, Abu Dhabi, UAE
{firstname.lastname}@mbzuai.ac.ae

Abstract. 基础模型通过在广泛的数据库上进行训练来捕捉某一领域的总体趋势。然而，在医学成像中，数据的稀缺性使得为每个领域、模式或任务预训练变得具有挑战性。连续学习提供了一个解决方案，即通过对不同领域或任务顺序微调一个模型，使其能够整合新知识而无需每次训练阶段都使用大型数据库。在本文中，我们提出了统一持续学习医学基础模型（[联合国统一国际惯例编纂处](#)），该框架使基础模型能够无缝适应不同的领域、任务和模式。与传统适应方法将这些变化孤立处理不同，UNICON 提供了一个统一的、不断扩展的框架。通过仔细整合，我们展示了基础模型可以在成像模式、解剖区域和临床目标之间动态扩展而不发生灾难性遗忘或任务干扰。在经验上，我们通过将最初用于分类的基础胸部 CT 模型适应于预后和分割任务来验证我们的方法。结果显示，在额外的任务中性能得到了提升。此外，我们不断纳入了 PET 扫描，并且与各自的基线相比，Dice 分数提高了 5%。这些发现表明基础模型并不局限于其初始训练范围，而是可以进化，为医学成像领域的通用人工智能模型铺平道路。

Keywords: 持续学习 · 效率医学成像 · 多模态适应

1 介绍

医学成像在现代医疗保健中扮演着至关重要的角色，提供了诊断、监测和治疗各种疾病所必需的视觉洞察力 [14,7,18]。尽管它非常重要，但收集和整理医学成像数据却非常耗费资源，大型医院每年会产生大约 100 太字节的成像数据，这些数据需要大量的存储空间、处理能力和基础设施才能有效管理。[25,11,15]。因此，分享医疗数据变得困难。

如今，随着基于大规模数据预训练的医疗基础模型（FM）的引入，深度学习系统在其训练领域内展现了卓越的表现 [32,27,8,13,31]。这些模型通常通过利用专业化的数据集、架构和培训目标来针对其特定领域，以符合医

学影像任务的需求。例如，它们在放射学报告、CT 扫描、MRI 或病理切片上进行训练，使它们能够捕捉到通用模型可能忽略的领域特定模式和特征 [28,29,26]。然而，医疗影像模态的多样性对 FM 构成了重大挑战。例如，CT 提供了详细的解剖结构横截面图像，而 PET 则可视化代谢过程，突出组织的功能方面。每种模态都捕捉了人体解剖学和生理的不同方面，导致了不同的数据分布。这种变异性可能会阻碍 FM 在所有影像类型上的泛化能力，因为它们可能难以在其训练数据中未代表的模态上保持良好的表现 [22]。

训练每种医学成像模式的 FM 由于需要大量的数据而不可行，这些数据在医疗领域中通常无法获得。因此，已经采用适应技术来扩展模型跨越不同的域、类别和任务 [19,16,30,4,24]。研究表明，分类模型可以适应以纳入新类别同时保持其原有任务 [17,30]，分割模型可以在不同解剖区域之间进行泛化 [4]，报告生成模型可以为各种医学发现生成结构化的输出 [24]。此外，研究还表明，模型可以适应以在同一解剖区域内集成多种模式 [23,19]。

尽管取得了这些进展，现有的方法主要集中在单一任务或领域内的持续适应上，缺乏一个统一的框架来将基础模型扩展到新的任务、成像模式和解剖区域。此外，与在精心策划的数据集上进行预训练不同，现实世界的医学影像数据经常在分辨率和质量上存在差异，这给将 FM 适应多样化的任务和领域带来了额外的挑战。这突显了需要一个全面的解决方案来实现跨所有维度的无缝适应。为此，我们提出了统一持续学习医疗基础模型（**统一符号**），这是一个框架，用于持续适应最初为特定模式、任务和解剖区域训练的基础模型，使其能够有效地推广到新的模式、任务和解剖区域。与传统的持续学习和常规领域自适应技术不同，我们的方法促进了在各种场景中的多步骤适应，确保了医学影像应用的可扩展性和鲁棒性。我们的贡献可以总结如下：

- 我们引入了框架**统一图标**，它使得基础模型能够在不降低性能的情况下持续扩展到新任务、模态和解剖区域。
- 我们提出了一种将基础模型适应于任意分辨率大小的方法，使得其能够超越原始训练限制进行无缝泛化。
- 我们证明了医学基础模型可以顺序整合新知识而不会发生灾难性遗忘，从而减少了对单独专业模型的需求。

2 方法论

问题陈述：我们解决了医学影像基础模型的持续学习问题，旨在扩展其在多种模式 (CT-PET (CP) 和文本报告 (R))、任务 (预后预测、分割 (M))

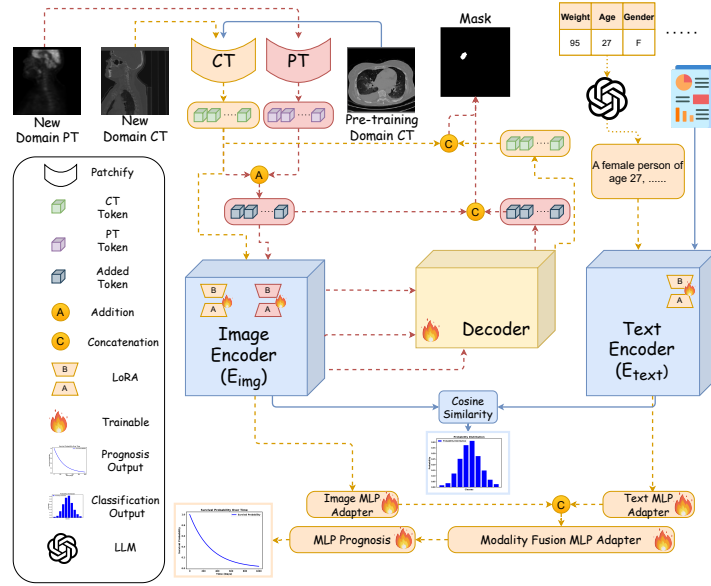


Fig. 1. 我们提出的 UNICON 框架概述。基础模型（蓝色块）处理预训练域的 CT 扫描及其相应的报告，通过计算图像和文本嵌入之间的相似性来进行分类。该框架引入了三个关键适应：(a) 预后适应（黄色流程）：大型语言模型（LLM）首先将 EHR 数据重构为类似报告的格式。然后，集成 MLP 适配器（图像和文本）、模态融合适配器以及 LoRA 微调以增强新域 CT 扫描的预后预测。(b) 分割适应（仅 CT）：添加 LoRA 权重和解码器以使用新域 CT 扫描进行分割。(c) 分割适应（CT + PET）（红色流程）：通过集成新域的 PT 扫描进一步使模型适应多模态分割。这涉及添加特定于 PT 的令牌和融合机制来处理两种模态。

等)和解剖区域(腹部、头颈部等)中的能力,同时保持基础知识的完整性。给定一个图像和文本编码器被冻结的基础模型 $E_{\text{img}}, E_{\text{text}}$, 挑战在于使这个强大的模型具备无缝适应来自不同模式和任务的新数据集 $\mathcal{X}_m$ 的能力, 其中 $m = 1, 2, \dots, M$ 不改变其基本参数。目标是使适应后的模型 $\mathcal{F}_\Psi$ 能够灵活地将多模态输入 (CP, R) 映射到相应的输出 $\mathcal{Y}_{\text{Prognosis}}$, 用于预后和 $\mathcal{Y}_{\mathcal{M}}$ 的分割, 从而转换为复杂医学影像任务中的**静态基础转变为动态、不断演进的助手**。所有符号术语定义如下: CP 指代共注册的 CT-PET 图像, R 指代相应的文本报告, M 指代分割掩模。输出用 $\mathcal{Y}_{\text{Prognosis}}$ 表示预后标签, 用 $\mathcal{Y}_{\mathcal{M}}$ 表示分割掩模。在统一框架中, $\mathcal{Y}_m$ 代表任意给定任务 m 的输出空间。

2.1 标准学习任务

预后预测：多模态数据的预后预测涉及从 CP 扫描和相关 R [20] 中估计患者结果。令 $\mathcal{X}_{CP, R} = \{(x_i, r_i, y_i)\}$ ，其中 x_i 是一个 CP 卷， r_i 是一份报告， y_i 是预后标签，可能被审查即为 NULL。我们定义： $\mathcal{F}_{\Theta_1}^{\text{Prog}} : \mathcal{X}_{CP, R} \rightarrow \mathcal{Y}_{\text{Prog}}$ ，其中 Θ_1 表示预后适配器的可训练参数，并通过跨模态注意力整合嵌入标记 T^{Img} 和 T^{Text} 。

CT 和 PET 体数据分割：分割任务生成体素级标签 $\mathcal{M}_i$ （分割掩码）用于解剖结构 [5]。对于 $\mathcal{X}_{CP} = \{(x_i, \mathcal{M}_i)\}$ ： $\mathcal{F}_{\Theta_2}^{\text{Seg}} : \mathcal{X}_{CP} \rightarrow \mathcal{Y}_{\mathcal{M}}$ ，其中 Θ_2 表示分割适配器的可训练参数，通过 transformer 层处理 patch 嵌入 T^{Seg} 。

分类：我们实验中的基础模型最初是通过 C 张图像和 R 训练进行疾病分类的。令 $\mathcal{X}_{C, R} = \{(x_i, r_i, y_i)\}$ ： $\mathcal{F}_{\Theta_0}^{\text{Class}} : \mathcal{X}_{CP, R} \rightarrow \mathcal{Y}_{\text{Class}}$ ，其中 Θ_0 表示为分类训练的基础模型参数。挑战在于将基础模型扩展到诸如预后预测 (Θ_1)、分割 (Θ_2) 以及多达 N 个不同的任务 ($\Theta_1, \Theta_2, \dots, \Theta_N$)，这些任务跨越多个领域且不会导致灾难性遗忘。

2.2 统一持续学习：UNICON

医学成像基础模型需要适应多种任务、模式和解剖区域，每种都需要不同的参数集 $\Theta_1, \Theta_2, \dots, \Theta_N$ 。为了解决这一问题，我们采用了一种参数高效的方案，在该方案中，基础模型 $\mathcal{F}_{\Psi}$ 是通过轻量级适配器 $\mathcal{A}^{(m)}$ 进行逐步适应的。这些适配器被实现为低秩 LoRA 模块、非线性 MLP 层或用于分割重建的解码器模块。每个模块的选择都基于即将进行的任务。我们的适应框架结构分为两个主要部分：模型内适应和模型后适应。

模型内适应 (WMA)。对此，我们利用 LoRA[10] 高效地更新预训练模型参数。该技术将低秩矩阵 $\Phi_1 \in \mathbb{R}^{d \times r}$ 和 $\Phi_2 \in \mathbb{R}^{r \times h}$ 引入原始权重矩阵，从而得到一个适应的权重矩阵： $W' = W + \Phi_1 \Phi_2$ ，其中 $r \ll \min(d, h)$ ， d 表示输入维度， h 表示输出维度。这种表述方式确保了高效的微调同时保留了核心模型的学习知识。

模型后适应 (PMA)。在此阶段，我们专注于通过使用非线性 MLP 层来优化预训练编码器提取的特征，从而从编码器嵌入中学习鲁棒表示。这里的关键组成部分是多模态融合 MLP 层，它整合并学习来自多种模式的组合表示。

对于特定任务的要求，例如分割，我们也包含一个专门的解码器模块，该模块根据学习到的嵌入重建分割掩码。

分辨率适应。预训练模型通常针对固定图像分辨率进行了优化，这在扫描分辨率多变的实际场景中可能会成为一个问题。为了高效地将我们的模型适应到新的模态或领域，我们通过动态纳入一个补丁嵌入层来解决这种分辨率不匹配的问题，该补丁嵌入层位于每个任务的变压器架构内。在训练过程中，补丁嵌入层以及位置嵌入和适应模块都会进行微调。这种方法使模型能够无缝地调整到不同的分辨率，同时保持跨各种领域的稳健性能。

图 1 展示了我们提出的方案。UNICON 通过在每个学习步骤中逐步集成新的适应模块，实现了跨任务和模态的持续学习统一。令 $\mathcal{X}$ 表示包含 CP 和 R 的输入空间，并且 $\mathcal{Y}_m$ 为任何任务 m 的输出空间。在每个步骤 $m, m = 1, 2, \dots, M$ ，都会训练一个新的适应模块 $\mathcal{A}^{(m)}$ 而不修改冻结的基础参数。统一的持续学习函数是：

$$\mathcal{F}_{\Psi}^{\text{UNICON}} : \bigcup_{m=1}^M \mathcal{X}_m \rightarrow \{\mathcal{Y}_m\}, \quad (1)$$

其中 $\mathcal{X}_m$ 表示任务 m 的输入数据。参数集 Ψ 包含冻结的 FM 编码器 $\mathbf{E}_{\text{img}}, \mathbf{E}_{\text{text}}$ ，以及所有适应模块 $\{\mathcal{A}^{(m)}\}$ 。每个适应模块在推理过程中动态选择。因此：

$$\Psi = \{\mathbf{E}_{\text{img}}, \mathbf{E}_{\text{text}}\} \cup \{\mathcal{A}^{(m)} : m = 1, 2, \dots, M\}.$$

图像编码器 $\mathbf{E}_{\text{img}}$ 将 CP 转换为潜在表示，而 $\mathbf{E}_{\text{text}}$ 嵌入 R 。适配模块 $\mathcal{A}^{(m)}$ 对这些嵌入进行优化以适应特定任务和模式，确保知识保留和无缝持续学习。

推理。我们的适应模块作为基础模型的动态扩展，在推理过程中根据输入和目标任务有选择地激活相关组件。这种设计使最初为特定任务和模态训练的模型能够无缝扩展其在多个任务和模态中的能力。通过高效地将数据路由到特定模态和融合适配器——以及用于分割的解码器——该框架确保了灵活且可扩展的推理，而无需进行大量的重新训练。

3 实验设置

数据集描述。我们使用了公开可用的 HECTOR 数据集 (HEad and neCK TumOR)，因为它为数不多的包含多种模态的数据集之一，即来自七个中心和不同机器类型的 488 名患者的 CT 和 PET 扫描、分割掩码和电子健康记

Table 1. 我们在实验中的渐进适应学习序列。

Order→	基础模型	步骤 1	步骤 2	步骤 3
身体区域	Chest →	头颈部 →	Head & Neck	→ Head & Neck
模态	CT →	CT →	CT →	CT & PET
任务	Classification →	Prognosis →	Segmentation →	Segmentation

录 (EHR) [2]。该综合数据集包括无复发生存期 (RFS) 信息, 详细说明了事件发生时间和审查状态。

配置。在本研究中, 我们将 CT CLIP 模型作为我们的基础模型 [8]。CT CLIP 是在胸部区域的 CT 扫描上进行预训练以完成分类任务的。表 1 详细描述了 CT CLIP 模型在各种任务和领域的顺序适应过程。我们的流程始于将基础模型调整为用于预后任务, 接着是仅使用 CT 扫描进行分割, 最后扩展到同时使用 PET 和 CT 扫描进行分割。具体任务配置如下。

对于患者结局预测, 我们整合了两个生存模型—DeepHit[12] 和 MTLR[6]。为了将 CT CLIP 模型定制化以适应生存分析, 我们在模型中加入了模块内适配器和模型后适配器模块。这种方法在减少参数开销的同时增强了特定任务的微调。预测性能使用一致性指数 (C-指数) 进行评估。模型通过 AdamW 优化器进行优化, 学习率为 $3e-4$, 权重衰减为 $1e-5$, 以批量大小为 16 训练了 50 个周期。CT 扫描被调整到 CT CLIP 模型接受的分辨率, 并保持比例不变。然后, 将值裁剪至范围 $(-1024, 1024)$ 并随后归一化到 $[-1, 1]$ 。对于文本报告, 通过在 GPT-4 模型 [1] 上使用提示工程将 EHR 数据转换为医疗报告。最佳模型是根据最高的验证 C-指数来选择的。

对于分割, 我们使用流行的 UNETR 3D 模型作为基线, 并对其进行了适应 [9]。模型的最大训练步数为 25k, 并使用 MONAI 库 [3]。所有 CT 和 PET 扫描都调整到 96X96X96 的尺寸, 并进行了裁剪和归一化。PET 图像通过 Z-score 标准化进行预处理。所有的分割实验均采用了 AdamW 优化器, 学习率为 $1e-4$, 权重衰减为 $1e-5$, 批量大小为 1。在验证集中平均 Dice 得分最高的模型被选为最优模型。

4 结果与讨论

我们评估了在 HECKTOR 数据集上使用五折交叉验证的适应型 CT CLIP 模型的预测性能。基线性能是通过两个知名的生存模型——DeepHit[12] 和 MTLR[6] 建立的。最初, 我们通过对文本编码器进行内部和后

Table 2. 预后适应性能在不同输入模式下使用 C 指数。我们使用了 DeepHit[12] 和 MTLR[6] 生存模型在不同输入模式下的性能。

Input Modality ↓	适应		C 指数	
	PMA	WMA	多任务学习关系	深度
基线 [21]	✗	✗	0.634	0.634
文本	✓	✗	0.668	0.686
	✓	✓	0.652	0.683
图像	✓	✗	0.546	0.566
	✓	✓	0.603	0.626
图像和文本	✓	✗	0.658	0.693
	✓	✓	0.670	0.721

Table 3. 不同输入医学模态下分割适应的 Dice 分数 (D-分数) 表现。我们使用 UnetR[9] 模型作为基线并对其进行实验改编。

Medical Modality ↓	适应		D-Score
	PMA	WMA	
CT (基线) [9]	✗	✗	0.609
CT	✓	✗	0.615
	✓	✓	0.628
CT & PET	✓	✓	0.657

模型适应来将 CT CLIP 模型调整为预测任务。如表 2 所示，这些适应模块相较于基线特定任务模型提高了 CT CLIP 基础模型的性能，C-index 提高了大约 2%。此外，当同时应用各种适应方法时，可以获得最佳结果。特别是，使用这些模块对文本和图像编码器进行双重适应与 DeepHit 结合产生了最高的 C-index，为 0.721。值得注意的是，仅采用图像编码器并没有带来显著的改进，这可以归因于预测任务严重依赖通过文本模态处理的 EHR 数据。

对于分割，我们在相同的数据和配置上训练了 UNETR[9] 作为基线。我们首先通过仅使用 CT 扫描并结合后模型适应模块（特别是解码器）来持续调整模型。然后，我们单独和联合地利用了适应模块。表 3 展示了 CT 适应在分割上的结果。仅仅使用 CT 使 dice 分数提高了 2%。此外，结合 PET 与 CT 扫描产生了最高的 dice 分数 65.7%。表 4 呈现了基线 CT-CLIP 模型与其 UNICON 调整版本之间的比较。该比较基于模型的功能能力和它们在不同适应情况下的表现。

这些结果突显了几个关键方面。首先，它们证明了将基础模型作为知识库在各种下游应用中的有效性。通过利用大规模预训练过程中从多样化数据集中学到的丰富表示，我们的方法能够高效地将知识转移到新任务、领域和模态上——涵盖了 CT 和 PET 扫描中的预测、分类和分割。这种知识转移由 LoRA 和适配器模块促进，这些模块在需要极少量额外计算资源的情况下实现了高效的适应性调整。这在医学成像中尤为宝贵，因为标注数据的稀缺构成了重大挑战。其次，与传统持续学习不同的是，后者要求为适应设定预定义边界，我们的方法允许模型动态地适应新领域和任务而无需修改基础模

Table 4. CT-CLIP 和 UNICON 修改版本在不同任务上的比较。该表格比较了 FM 与 UNICON 修改模型在不同任务上的能力：分类 (Cls)、预后 (Prog)、CT 上的分割 (Seg (C)) 以及 CT & PET 上的分割 (Seg (CP))。

Task→	类	程序	段 (C)	段 (CP)
CT-CLIP	✓	✗	✗	✗
UNICON 适应版	✓	✓ (72.1 C-Index)	✓ (62.8 D-Score)	✓ (65.7 D-Score)

型的参数。这种灵活性至关重要，因为在现实世界中的顺序数据并不严格局限于类、任务、领域或模态的变化，而是可能包括它们之间的任何组合。

第三，我们的研究结果表明，仅微调一小部分参数就能使基础模型保留其原有的语义理解能力，同时有效适应新数据的特定特征。这在医学成像领域尤为重要，在该领域中获取大规模、标注良好的数据集通常不可行。最后，本框架有效地解决了领域迁移问题，确保了基础模型即使应用于超出原始训练范围的新模式或任务时仍保持鲁棒性。通过持续微调并扩展模型以纳入新的模式和任务，我们的方法减轻了灾难性遗忘的风险，并支持增量学习，使其非常适合不断发展的医学应用。

5 结论

在这项工作中，我们介绍了 UNICON，一个扩展基础医学成像模型能力的新连续训练框架，使其超越单一领域或任务的适应。我们的方法成功地统一了传统类型的持续学习，使得单个模型能够处理多种模式和任务。我们的实验表明，经过调整的基础模型在执行特定任务时与专门的任务特定模型相当，并且计算开销最小。这项工作强调了持续学习范式在克服数据稀缺性和提高医学成像系统灵活性方面的潜力。这项工作的潜在局限性在于量化适应新领域或任务所需的数据量。未来的研究将专注于进一步完善这些技术，并探索它们在更广泛临床任务和成像模式中的适用性。

References

1. Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F.L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., et al.: Gpt-4 technical report. arXiv preprint arXiv:2303.08774 (2023)

2. Andrearczyk, V., Oreiller, V., Boughdad, S., Rest, C.C.L., Elhalawani, H., Jreige, M., Prior, J.O., Vallières, M., Visvikis, D., Hatt, M., et al.: Overview of the hecktor challenge at miccai 2021: automatic head and neck tumor segmentation and outcome prediction in pet/ct images. In: 3D head and neck tumor segmentation in PET/CT challenge, pp. 1–37. Springer (2021)
3. Cardoso, M., Li, W., Brown, R., Ma, N., Kerfoot, E., Wang, Y., Murrey, B., Myronenko, A., Zhao, C., Yang, D., et al.: Monai: An open-source framework for deep learning in healthcare. arxiv preprint arxiv: 221102701. Published online (2022)
4. Chen, Q., Zhu, L., He, H., Zhang, X., Zeng, S., Ren, Q., Lu, Y.: Low-rank mixture-of-experts for continual medical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 382–392. Springer (2024)
5. Foster, B., Bagci, U., Mansoor, A., Xu, Z., Mollura, D.J.: A review on segmentation of positron emission tomography images. *Computers in biology and medicine* **50**, 76–96 (2014)
6. Fotso, S.: Deep neural networks for survival analysis based on a multi-task framework. arXiv preprint arXiv:1801.05512 (2018)
7. Giger, M.L.: Machine learning in medical imaging. *Journal of the American College of Radiology* **15**(3), 512–520 (2018)
8. Hamamci, I.E., Er, S., Almas, F., et al.: Developing generalist foundation models from a multimodal dataset for 3d computed tomography. arXiv preprint arXiv:2403.17834 (2024)
9. Hatamizadeh, A., Yang, D., Roth, H., Xu, D.U.: Transformers for 3d medical image segmentation. arxiv 2021. arXiv preprint arXiv:2103.10504 (2021)
10. Hu, E.J., Shen, Y., Wallach, P., Lipton, Z.C., Li, Y.: Lora: Low-rank adaptation of large language models. arXiv preprint arXiv:2106.09685 (2021)
11. Langer, S.G., Langlotz, C.P.: Challenges in multi-site imaging trials: The need for standardization and data integration. *Journal of Digital Imaging* **24**(6), 1175–1183 (2011). <https://doi.org/10.1007/s10278-011-9364-1>
12. Lee, C., Zame, W., Yoon, J., Van Der Schaar, M.: Deephit: A deep learning approach to survival analysis with competing risks. In: Proceedings of the AAAI conference on artificial intelligence. vol. 32 (2018)
13. Liang, J., Cao, P., Yang, W., Yang, J., Zaiane, O.R.: 3d-sautomed: Automatic segment anything model for 3d medical image segmentation from local-global perspective. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 3–12. Springer (2024)

14. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B., Sánchez, C.I.: A survey on deep learning in medical image analysis. *Medical image analysis* **42**, 60–88 (2017)
15. OpenMedScience: Big data in medical imaging: Storage and computational challenges. *OpenMedScience Reports* (2023), <https://www.openmedscience.com/big-data-in-medical-imaging>
16. Qazi, M.A., Almakky, I., Hashmi, A.U.R., Sanjeev, S., Yaqub, M.: Dynammo: Dynamic model merging for efficient class incremental learning for medical images. In: *Annual Conference on Medical Image Understanding and Analysis*. pp. 245–257. Springer (2024)
17. Qazi, M.A., Hashmi, A.U.R., Sanjeev, S., Almakky, I., Saeed, N., Yaqub, M.: Continual learning in medical imaging from theory to practice: A survey and practical analysis. *arXiv preprint arXiv:2405.13482* (2024)
18. Ranschaert, E.R., Morozov, S., Algra, P.R.: *Artificial intelligence in medical imaging: opportunities, applications and risks*. Springer (2019)
19. Saadi, N., Saeed, N., Yaqub, M., Nandakumar, K.: Pemma: Parameter-efficient multi-modal adaptation for medical image segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 262–271. Springer (2024)
20. Saeed, N., Al Majzoub, R., Sobirov, I., Yaqub, M.: An ensemble approach for patient prognosis of head and neck tumor using multimodal data. In: *3D Head and Neck Tumor Segmentation in PET/CT Challenge*, pp. 278–286. Springer (2021)
21. Saeed, N., Ridzuan, M., Maani, F.A., Alasmawi, H., Nandakumar, K., Yaqub, M.: Survnc: Learning ordered representations for survival prediction using rank-n-contrast. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 659–669. Springer (2024)
22. Shi, C., Rezai, R., Yang, J., Dou, Q., Li, X.: A survey on trustworthiness in foundation models for medical image analysis. *arXiv preprint arXiv:2407.15851* (2024)
23. Sobirov, I., Nazarov, O., Alasmawi, H., Yaqub, M.: Automatic segmentation of head and neck tumor: How powerful transformers are? In: *International Conference on Medical Imaging with Deep Learning*. pp. 1149–1161. PMLR (2022)
24. Sun, Y., Khor, H.G., Wang, Y., Wang, Z., Zhao, H., Zhang, Y., Ma, L., Zheng, Z., Liao, H.: Continually tuning a large language model for multi-domain radiology report generation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 177–187. Springer (2024)
25. System, D.U.H.: Healthcare data volume and storage challenges. *DAIC* (2015), <https://www.daic.com/article/100tb-per-year-medical-imaging>

26. Taher, M., et al.: Foundational artificial intelligence models and modern medical diagnostics: The paradigm of the eden ai model. *Journal of Clinical Medicine* **12**(20), 6500 (2023)
27. Wu, C., Zhang, X., Zhang, Y., Wang, Y., Xie, W.: Towards generalist foundation model for radiology by leveraging web-scale 2d&3d medical data. *arXiv preprint arXiv:2308.02463* (2023)
28. Wu, C., Zhang, X., Zhang, Y., Wang, Y., Xie, W.: Towards generalist foundation model for radiology by leveraging web-scale 2d&3d medical data. *arXiv preprint arXiv:2308.02463* (2023)
29. Yu, K.H.: Ai breakthrough raises hopes for better cancer diagnosis. *Financial Times* (2024), <https://www.ft.com/content/0a8f2c61-77f4-43ce-87d2-a7b421bbda85>
30. Zhang, W., Huang, Y., Zhang, T., Zou, Q., Zheng, W.S., Wang, R.: Adapter learning in pretrained feature extractor for continual learning of diseases. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 68–78. Springer (2023)
31. Zhang, X., Ou, N., Basaran, B.D., Visentin, M., Qiao, M., Gu, R., Ouyang, C., Liu, Y., Matthews, P.M., Ye, C., et al.: A foundation model for brain lesion segmentation with mixture of modality experts. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 379–389. Springer (2024)
32. Zhao, T., Gu, Y., Yang, J., Usuyama, N., Lee, H.H., Naumann, T., Gao, J., Crabtree, A., Abel, J., Moungh-Wen, C., et al.: Biomedparse: a biomedical foundation model for image parsing of everything everywhere all at once. *arXiv preprint arXiv:2405.12971* (2024)