

内容列表可在[科学直接](#)

医学图像分析

期刊主页: www.elsevier.com/locate/media



调查论文

用于医学图像分析的主动学习和人机循环深度学习综述



塞缪尔·巴德^A, 艾玛·C·罗宾逊^B, 伯恩哈德·凯恩斯^A

^A英国伦敦帝国理工学院计算系
^B英国伦敦国王学院影像科学系

文章信息

文章历史:
2019年10月7日收稿 2021年3月
26日修订 2021年3月30日接受
2021年4月9日在线发布

关键词:
医学图像分析
深度学习
人在环
主动学习

抽象的

全自动深度学习已成为许多任务的最先进技术,包括图像采集、分析和解释,以及提取临床有用信息以进行计算机辅助检测、诊断、治疗计划、干预和治疗。然而,医学图像分析带来的独特挑战表明,在任何支持深度学习的系统中保留人类最终用户将是有益的。在这篇综述中,我们研究了人类在开发和部署支持深度学习的诊断应用程序中可能扮演的角色,并重点关注能够保留人类最终用户重要输入的技术。由于医疗领域工作的安全性至关重要,我们认为人在环计算是未来研究中日益重要的一个领域。**主动学习**选择最佳数据进行注释以获得最佳模型性能;(2) **与模型输出的交互**-使用迭代反馈来引导模型针对给定的预测进行优化,并提供有意义的方法来解释和响应预测;(3) **实际考虑**-开发全面的应用程序以及部署前需要考虑的关键因素;(4) **未来的前瞻性和未解答的问题**-随着人类在环计算的发展,知识差距和相关研究领域将受益。我们就最有前途的研究方向以及如何将每个领域的各个方面统一起来以实现共同目标提供意见。

© 2021 Elsevier BV 保留所有权利。

一、简介

医学成像是临床决策的主要支柱,也是许多患者旅程中不可或缺的一部分。从医学图像中提取的信息在计算机辅助检测、诊断、治疗计划、干预和治疗等许多领域具有临床用途。虽然医学成像仍然是众多临床任务的重要组成部分,但解释复杂医学图像的合格放射科医生的日益短缺表明,显然需要可靠的自动化方法来减轻医疗保健从业者日益增长的负担。[皇家放射科医生学院, 2017](#))。

与此同时,医学成像科学也受益于用于分析图像等结构化数据的新颖计算技术的发展。图像采集、分析和解释算法的开发正在推动创新,特别是在配准、重建、跟踪、分割和建模领域。

医学图像本质上难以解释,需要事先具备专业知识才能理解。生物医学图像可能存在噪声,并且包含许多特定于模态的伪影,这些伪影是在各种采集条件下使用不同协议采集的。因此,一旦经过训练的模型就无法从一个临床任务或地点无缝转移到另一个临床任务或地点,因为领域差距经常扩大([卡姆尼察斯等人, 2017](#); [本-大卫等人, 2010](#))。监督学习方法需要大量的重新标记才能在不同的工作流程中恢复初始性能。

处理此类数据所需的经验和先验知识意味着在注释医学数据时通常存在观察者之间和观察者内部的巨大差异。这不仅引发了关于什么是黄金标准地面事实注释的问题,而且还导致了对于地面事实到底是什么的分歧。这些问题导致与医学图像数据集的注释和重新标记相关的巨大成本,因为我们需要大量的专家注释者(预言机)来执行每个注释并达成共识。

近年来,深度学习(DL)已成为执行许多医学图像分析的最先进技术。

*通讯作者。

电子邮件地址: 塞缪尔.budd13@imperial.ac.uk (S.巴德)。

sis任务 (泰巴赫什等人, 2020; Tizhoosh 和潘塔诺维茨, 2018; 沉等人, 2017; Litjens 等人, 2017; 铃木, 2017)。计算机视觉领域的发展在转移到医学图像分析方面显示出了巨大的前景, 并且多种技术已被证明可以与人类观察者一样准确地执行 (汉斯勒等人, 2018; 三月和索耶, 2018)。然而, 迄今为止, 深度学习方法在临床实践中的采用受到限制, 这主要是由于处理复杂的医疗数据、监管合规问题和对训练模型的信任所面临的独特挑战。

在开发用于临床环境中的医学图像分析的 DL 应用程序时, 我们确定了三个关键挑战:

1. 缺乏训练数据: 传统上, 有监督的深度学习技术依赖于大量且均匀分布的精确注释数据点, 虽然越来越多的医学图像数据集变得可用, 但注释此类数据集所需的时间、成本和精力仍然很大。
2. 最后的百分比: 深度学习技术已经在医学图像分析任务中实现了最先进的性能, 但在安全关键领域, 即使是最小的错误也可能导致下游的灾难性结果。实现临床上可信的输出可能需要对预测 (来自预言机) 进行交互式解释, 以便在实践中发挥作用, 即用户必须有能力纠正和覆盖自动预测, 以满足所需的任何验收标准。
3. 透明度和可解释性: 目前, 大多数深度学习应用程序被认为是一个“黑匣子”, 用户解释、理解或纠正模型预测方式的有意义的方式有限。可信度对于医疗应用来说是一个不利的特征, 因为必须评估来自各种来源的信息才能做出临床决策。需要进一步表明模型如何得出预测结论, 以培养对支持深度学习的系统的信任, 并允许用户适当权衡自动预测。

医学图像分析研究界正在共同努力将深度学习方法应用于各种医学图像分析任务, 并且这些方法显示出了巨大的前景。我们建议读者参考一些关于医学成像中深度学习的评论 (赫萨米安等人, 2019; 伦德沃德和伦德沃德, 2019; 山下等人, 2018)。这些工作主要侧重于开发特定任务的预测模型, 并展示该任务的最先进性能。

本综述旨在概述人类将继续参与用于医学图像分析的 DL 系统的开发、部署和实际使用。我们专注于医学图像分割技术, 以探索人类最终用户的支持深度学习的系统中的作用。

自动化图像解释任务 (例如图像分割) 存在上述医学图像数据所带来的所有缺点。有许多新兴技术试图减轻使用医学图像数据来执行图像自动分割所增加的复杂性。分割旨在将图像划分为语义上有意义的区域 (像素集), 以便执行许多下游任务, 例如生物识别测量。手动为图像的每个像素分配标签是一项艰巨的任务, 因此自动分割方法在实践中很重要。主动学习 (AL) 和应用于分割问题的人机循环计算等深度学习技术的进步, 在克服上述关键挑战方面取得了进展, 这些是本综述重点研究的内容。

第2节引入了主动学习, 它是机器学习 (ML) 和人在环计算的一个分支, 旨在寻找最信息丰富的接下来要注释的来自未标记分布的样本。通过对信息最丰富的样本子集进行训练, 相关工作可以实现最先进的性能, 同时减少与注释医学图像数据相关的昂贵注释负担。

第三节评估用于根据用户反馈改进模型预测的技术, 指导模型实现更准确的每图像预测。我们评估旨在提高自动预测的可解释性的技术, 以及模型如何提供对其自身输出的反馈以指导用户做出更好的决策。

第4节评估了在实践中开发和部署支持深度学习的人在环系统的关键实际考虑因素, 并概述了在这些领域正在开展的工作, 以解决上述三个关键挑战。这些领域以人为本, 并评估人类最终用户如何与这些系统交互。

在第5节我们讨论了 ML 和 DL 研究的相关领域, 这些领域对 AL 和人在环计算产生影响, 并开始影响概述的三个关键挑战。我们对人在环深度学习研究的未来方向以及评估的技术有多少可以结合起来以实现共同目标提出意见。

2. 主动学习

在本节中, 我们假设一个场景, 其中有大量未注释的数据 U 我们可以使用, 并且我们有一个预言机 (或一组预言机), 我们可以从中请求每个未注释的数据点的注释 X_U 添加到带注释的集合 L 。我们希望训练一些模型 $f(x|L)$ 在哪里 $L \subseteq L$ 并考虑依赖注释数据来执行此操作的方法。这个问题一个强力解决方案是要求预言机注释每个 X_U 这样 $L'=L$, 但这很少是一个实用或具有成本效益的解决方案, 因为与注释生物医学图像数据相关的独特挑战。据推测, 有一些 L' 达到同等性能 $L, \mathbb{E}f(x|L') \approx f(x|L)$ 。在某些最佳子集上训练的模型 L' 数据集的模型可能会获得与在整个带注释的数据集上训练的模型相当的性能。主动学习 (AL) 是机器学习的一个分支, 旨在找到这个最佳子集 L' 给定当前模型 $F(X|L')$, 在哪里 L' 是中间带注释的数据集, 和未注释的数据集 U 。AL 方法旨在迭代地寻找信息最丰富的数据点 X 用于训练模型, 假设模型和未注释的数据集都会随着时间的推移而演变, 而不是选择一次固定的子集用于训练。在更广泛的背景下, 在深度学习出现之前, 安定下来 (2009) 将该领域视为最先进的机器学习方法。

我

典型的 AL 框架, 如图 1, 包括一种评估方法信息性每个未注释的数据点 X_U 给定 $F(X|L')$ 与选择密切相关查询类型, 之后需要对所有选定的数据点进行注释。一旦获得新的注释, AL 框架必须使用新数据来改进模型。这通常由以下任一者完成再培训使用所有可用的注释数据的整个模型 L' , 或通过微调使用最近注释的数据点的网络 X 。使用这种方法, 可以使用更少的注释来完成多个生物医学图像分析任务, 从而实现最先进的性能, 如本节讨论的方法所示, 从而扩大注释瓶颈并降低与开发支持深度学习的系统相关的成本来自未注释的数据。

我

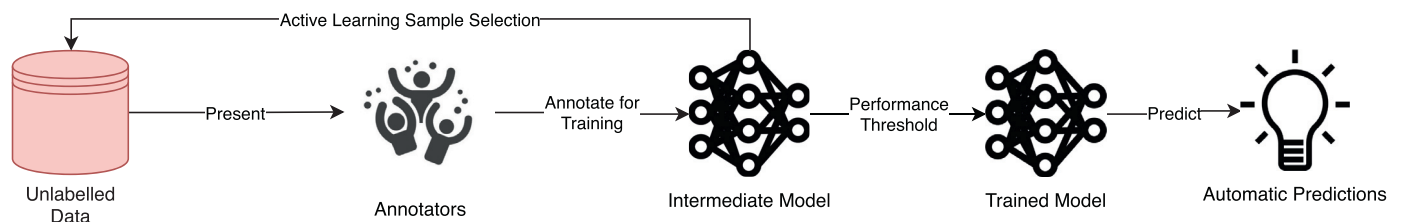


图. 1。主动学习框架概述。

2.1. 查询类型

在每个 AL 框架中，首先要选择的是什么类型查询？我们希望使用模型和未注释的数据集。目前有三种主要选择，每种选择都适合特定的场景，具体取决于我们可以访问的未注释数据的类型以及我们希望向预言机提出什么问题。

基于流的选择性采样假设传入的未注释数据点是连续的流 X_U （阿特拉斯等人，1990；科恩等人，1994）。当前模型和 **信息性措施** $I(x_u)$ 用于决定对于每个传入的数据点是否向预言机请求注释（达甘和恩格尔森，1995）。这种查询类型通常计算成本低廉，但由于每个决策的孤立性质，提供的性能优势有限：不考虑底层分布的更广泛上下文，因此与其他查询类型相比，平衡分布的探索和利用的捕获效果较差。这种查询类型的另一个缺点是校准用于所选信息量度量的阈值，这样我们就不会请求每个传入数据点的注释，并且我们不会拒绝太多数据点的注释，从而导致有价值的信息丢失。

会员查询综合假设我们不是从数据点的真实分布中提取，而是生成

一个数据点 x_u 需要注释的（安格鲁因，1988）。这生成的数据点是当前模型“认为”对其自身最具信息性的数据点。然后，该数据点由预言机注释（安格鲁因，2001），这在有限域中非常有效。这种方法可能会遇到与以下相同的缺点：**基于流**的作为模型的方法可能不知道分布的未见区域，因此无法请求这些区域的注释。当查询请求对人类预言机毫无意义的数据点进行注释时，可能会出现（朗和鲍姆，1992），并且不能代表正在建模的实际分布，提出了基于流和基于池的采样方法来克服这些问题（安顿下来，2009）。尽管如此，最近的进展**生成对抗网络**（GAN）在生成模拟许多不同类型数据（包括生物医学图像）的真实世界分布的数据点方面表现出了巨大的希望，这可能会在某种程度上解决使用复杂分布的查询合成的关键问题，我们将在第2.2.3节。在完全理解要生成的分布的情况下，或者在自动获取注释而不是从人类那里获取注释的领域中，这种查询类型可能是有利的（金等人，2004；2009年）。

基于池的采样假设有一个大型的未注释的真实数据集 U 从中抽取样本并寻求选择一批 N 个样品 x_1, \dots, x_N 从分发请求标签（刘易斯和卡特莱特，1994）。基于池的方法通常使用当前模型对每个未注释的数据点进行预测，以获得排序的度量 $I(x_u|P(x_u|L))$ 对于未注释集中的每个数据点，然后选择顶部 N 使用此指标的样本由预言机进行注释。水池

在深度学习出现之前，基于采样已应用于多个现实世界的任务（刘易斯和卡特莱特，1994；麦卡勒姆和尼加姆，1998；塞特尔斯和克雷文，2008；张和陈，2002；豪普特曼等人，2006）。这些方法的计算成本可能很高，因为每次迭代都需要对分布中的每个数据点进行度量评估。然而，这些方法在与深度学习方法结合时被证明是最有前途的，深度学习方法本质上依赖于基于批量的训练方案。除非另有说明，本节其余部分讨论的大多数方法都使用基于池的抽样。虽然基于池的方法在找到要获取的信息最丰富的注释方面比其他方法具有优势，但基于流或基于合成的查询具有优势的场景也很常见，例如当内存或处理能力有限时，例如在移动或嵌入式设备（安顿下来，2009）。

2.2. 评估信息量

在开发 AL 框架时，一旦选择了查询类型，下一个要问的问题是如何衡量信息量 $I(x_u)$ 每个数据点？人们已经采取了许多不同的方法来量化给定模型和基础分布的样本的信息量。在这里，我们根据这些指标提供的人类可解释性水平对它们进行排序。

传统上，AL 方法采用手工设计的启发式方法来量化我们人类认为的信息量。各种模型特定指标试图量化使用样本进行训练对模型的影响，例如模型参数的最大变化。然而，由于将这些方法应用于具有大量参数的深度学习模型的计算挑战，这些方法不如人类设计的启发式方法普遍。最后，出现了一些方法，这些方法与人类对信息的可解释性完全无关，而是寻求从可用数据和先前的迭代中学习最佳选择策略，如详细讨论的第2.2.4节。

2.2.1. 不确定

信息量测量的主要系列属于计算不确定性。有人认为，越不确定预测是，通过在训练集中包含该样本的基本事实，我们可以获得的信息越多。

有多种方法可以计算不同 ML/DL 模型的不确定性。当考虑深度学习进行分割时，最简单的衡量标准是给定图像分割中每个像素的最低类别概率之和。有人认为，更确定的预测将具有更高的逐像素类别概率，因此图像中每个像素的最小类别概率之和越低，预测就越确定：

$$X_{\text{最上色}} = \underset{X}{\operatorname{argmax}} 1 - \text{磷}(\hat{y}|X)$$

在哪里 $\hat{y} = \underset{y}{\operatorname{argmax}} \theta(y|X)$ 。这是一种相当直观的考虑不确定性的方法，并提供了一种对分布内样本的不确定性进行排序的方法。我们将上面的方法称为

最不自信抽样，其中选择不确定性最高的样本进行标记（安顿下来，2009）。一个缺点是 最不自信采样的特点是它只考虑最可能的标签信息，而丢弃剩余标签分布的信息。已经提出了两种替代方法来减轻这种担忧。第一个，叫做 边际抽样（安顿下来，2009），可以在多类设置中使用，并考虑模型下第一和第二最可能的标签并计算它们之间的差异：

$$X_{\text{中}} = \underset{X}{\text{argmax}} \theta(y_1|X) - \theta(y_2|X)$$

在哪里 y_1 和 y_2 分别是当前模型下第一和第二最可能的标签。这里的直觉是，两个最可能的标签之间的间隔越大，模型分配该标签的信心就越大。第二种更流行的方法是使用 熵（香农）作为不确定性度量：

$$X_{\text{中}} = \underset{X}{\text{argmax}} - \sum_{y \in \mathcal{Y}} P(y|X) \log P(y|X)$$

在哪里 $y \in \mathcal{Y}$ 范围涵盖所有可能的注释。熵用于测量对分布进行编码所需的信息量，因此通常被认为是机器学习中的不确定性的度量。对于二元分类，所有三种方法都简化为查询类后验最接近 0.5 的数据点。的能力 熵易于推广到概率多类注释以及更复杂的结构化数据点的模型，使其成为基于不确定性的查询策略的最受欢迎的选择 塞特尔斯和克雷文 (2008)。

使用上述措施之一，对未注释的样本进行排名，并选择最“不确定”的案例进行下一轮注释。最近，在 DL 领域的 AL 方法中，基于不确定性的采样得到了许多应用，接下来将对此进行讨论。

王等人。(2017)提出用于深度图像分类的成本有效的主动学习 (CEAL) 方法。CEAL 方法使用一组未标记的样本进行初始化你，初始标记样本 L ，泳池大小的选择 K ，高置信度样本选择阈值 ω ，阈值衰减率 β ，最大迭代次数 t_{max} 和微调间隔 t_{micro} 。初始化后，CNN 权重 W 初始化为 L 该模型用于对每个数据点进行预测 U 。CEAL 探索使用上述三种不确定性方法来获得 K 需要手动注释和添加不确定的数据点 D_L 。到目前为止，CEAL 方法非常密切地遵循上述传统主动学习方法中概述的方法，但它们引入了一个额外的训练步骤，其中最有可能的样本（其 熵小于 ω ）从 U 被添加到 D_H 。 D_L 和 D_H 然后用于微调 W 为了 t_{micro} 迭代。CEAL 随后更新 ω 之前 伪标签从 D_H 被丢弃，每个数据点被添加回你，尽管 D_L 被添加到 L 。这个过程重复进行 t_{max} 迭代。作者将这种同时从最不确定注释的手动标签和最不确定注释的预测标签中学习的方法描述为 补充抽样。CEAL 方法表明，使用两个非医学数据集 (CADC 和 Caltech-256) 的不到 60% 的可用数据即可实现人脸识别和对象分类的最先进性性能。

文等人。(2018)提出一种主动学习方法，使用不确定性采样来支持病理图像中细胞核分割的质量控制。他们的工作比较了通过主动学习三种不同算法系列所实现的性能改进：支持向量机 (SVM)、随机森林 (RF) 和卷积神经网络 (CNN)。他们表明 CNN 可以达到最高的准确度，需要信号

显着减少迭代次数以达到与 SVM 和 RF 相同的精度。

估计信息量的另一种常见方法是衡量执行同一任务的多个模型之间的一致性。有人认为，对同一数据点的预测之间发现的分歧越大，意味着不确定性越高。这些方法被称为 共识查询 并且一般应用在以下情况： 合奏用于提高性能 - 即训练多个模型在略有不同的参数/设置下执行相同的任务 安定下来 (2009)。集成方法已被证明可以很好地测量信息量，但以计算资源为代价——需要训练和维护多个模型，并且每个模型都需要在新选择的训练样本存在的情况下进行更新。

尽管如此，贝鲁奇·Bcai 等人。(2018)展示集成对于主动学习的力量，并与集成的替代方案进行比较。他们专门比较了主动学习的采集函数和不确定性估计方法与用于图像分类任务的 CNN 的性能，并表明基于集成的不确定性优于其他不确定性估计方法，例如“MC Dropout”。他们发现主动学习性能的差异可以通过模型容量下降和 MC dropout 集合多样性降低来解释。在糖尿病视网膜病变诊断任务中表现出良好的性能。

科纽什科娃等人。(2019)提出一种主动学习方法，利用图像空间中的几何平滑先验来帮助分割过程。他们使用传统的不确定性度量来估计接下来应该注释哪些像素，并在多类设置中引入新的不确定性标准。他们通过在给定邻居预测的情况下估计属于某一类的超像素概率的熵来利用几何不确定性，并将这些结合起来以鼓励在类之间非平滑过渡的区域中选择不确定区域。他们展示了从 EM 图像进行线粒体分割以及针对二元和多类分割的 MRI 肿瘤分割任务的最先进性性能。

加尔等人。(2017)介绍贝叶斯 CNN 在主动学习中的使用以及“贝叶斯主动学习不一致”或 BALD，并表明在主动学习的背景下，贝叶斯 CNN 的使用优于确定性 CNN，并通过使用新的获取函数来利用这一点选择期望最大化获得的有关模型参数的信息的数据点，即最大化预测和模型后验之间的互信息。该方法使用贝叶斯 CNN（在推理过程中使用 Dropout 诱导） 加尔和加拉玛尼，2016），使用网络的所有参数为每个未标记的数据点生成单个预测，并为每个未标记的数据点生成一组随机预测，在启用 dropout 的情况下生成。然后将 BALD 采集函数计算为平均预测的熵与随机预测的平均熵之间的差。直观上，该函数选择模型平均不确定的数据点，但存在会产生高度确定性不一致的预测注释的模型参数。他们展示了从皮肤病变图像诊断皮肤癌的方法，与使用 BALD 方法进行样本选择的均匀采样相比，性能显著提高。虽然这种方法已被证明对于 AL 特别有效，但在查询批量数据时 -

点, 当以贪婪方式使用时, 通常会导致获取许多非常相似的冗余数据点, 因此引入 BatchBALD 来缓解这个问题[基尔希等人。(2019)。BatchBALD 方法不再计算单个样本预测与模型后验之间的互信息, 而是计算一批样本与模型后验之间的互信息来对这批样本进行联合评分, 使得 BatchBALD 能够更准确地评估联合结果相互信息并选择样本批次进行注释, 从而在采集的批次中一起选择较少的冗余数据点。这个扩展是背后动机的一个例子第2.2.2节其中我们讨论了超越纯粹基于不确定性的方法并开始测量选定样本之间的多样性以减少冗余注释的方法。

2.2.2. 代表性

许多 AL 框架扩展了选择策略, 以包括一些衡量标准 代表性除了不确定性测量之外。包括代表性度量背后的直觉是, 方法只涉及 不确定有可能只关注分布的小区域, 并且对来自同一分布区域的样本进行训练将为选择策略引入冗余, 或者可能使模型偏向分布的特定区域。增加代表性度量旨在鼓励选择策略从分布的不同区域进行采样, 并增加样本的多样性, 从而提高 AL 性能。具有高代表性的样本涵盖了分布同一区域中许多图像的信息, 因此不需要包含许多被代表性图像覆盖的样本。

为此, 杨等人。(2017)目前的建议注释是一种用于医学图像分割的深度主动学习框架, 它使用不确定性采样的替代公式与代表性密度加权的形式相结合。他们的方法包括训练多个模型, 每个模型都排除一部分训练数据, 这些模型用于计算基于集成的不确定性度量。他们将选择最具代表性的示例表述为最大集合覆盖问题 (NP Hard) 的广义版本, 并提供了一种贪婪的方法使用模型中的特征向量选择最具代表性的图像。他们使用 MICCAI 腺体分割挑战和淋巴结分割任务的 50% 可用数据展示了最先进的性能。

斯迈拉吉克等人。(2018)提出 勋章, 一个用于医学图像分割的主动学习框架。他们提出了一种结合不确定性和特征描述符之间的距离的采样方法, 从未标记的数据集中提取信息最丰富的样本。一旦训练了初始模型, MedAL 方法就会通过首先过滤掉预测熵低于阈值的未标记数据点来选择要标记的数据点。从该集合中, 正在训练的 CNN 用于通过获取 CNN 中间层的输出出来生成每个数据点的特征描述符, 然后使用各种距离函数 (例如 “欧几里德”、'Russellrao', 'Cosine'), 以便找到彼此最近的特征描述符。选择与其他未标记数据点 (高于熵阈值) 平均距离最高的数据点进行注释。通过这种方式, MedAL 采集函数可以找到对模型提供信息的数据点集, 并通过从彼此最远的输入分布区域进行采样, 在它们之间产生最少的冗余。MedAL 方法以一种新颖的方式初始化模型, 利用现有的计算机视觉图像描述符来查找彼此最不相似的图像, 从而覆盖图像分布的更大区域, 以用作注释后的初始训练集。他们表现得很好 并通过从彼此距离最远的输入分布区域进行采样, 在它们之间产生最少的冗余。MedAL 方法以一种新颖的方式初始化模型, 利用现有的计算机视觉图像描述符来查找彼此最不相似的图像, 从而覆盖图像分布的更大区域, 以用作注释后的初始训练集。他们表现得很好

三种不同医学图像分析任务的结果, 与基于随机或纯不确定性的方法相比, 用更少的训练数据实现了基线精度。

奥兹德米尔等人。(2018)提出基于博尔达计数的不确定性和代表性度量的组合来选择下一批样本。不确定性通过模型中使用 MC dropout 的 N 预测的体素方差来衡量。他们引入了新的代表性度量, 例如 “内容距离”, 定义为预训练分类网络的层激活响应之间的均方误差。他们通过最大熵编码代表性来扩展这一贡献, 以使用新颖的熵损失函数优化网络权重。

苏拉蒂等人。(2018)首次在 CNN 中提出了一种通过计算 Fisher 信息 (FI) 来确保查询样本多样性的新方法。这里, 通过传播的梯度计算实现高效计算, 从而允许在 CNN 的大参数空间上计算 FI。他们展示了他们的方法在两种不同类型的任务上的性能: (a) 特定主题的半自动分割 (来自原始训练数据中不存在的不同组/不同病理学), 其中迭代地标记由 AL 实现了对该主题的准确分割; (b) 使用 AL 构建可推广到给定数据集中所有图像的模型。他们表明, 在这两种情况下, 基于 FI 的 AL 在标记一小部分体素后提高了性能,

2.2.3. 信息量的生成对抗网络基于生成对抗网络 (GAN) 的方法已应用于医学成像的多个领域, 例如去噪、模态转移、异常检测以及图像合成, 可直接应用于 AL 场景。这为用于扩展有限数据集的许多数据增强技术提供了替代 (或补充) 易等人。(2019)以及深度学习方法委员会查询综合。

最后等人。(2018)提出一种基于条件 GAN (cGAN) 的主动学习方法, 其中使用判别器输出作为所提出的分割的不确定性度量, 并使用该度量对未标记数据集中的样本进行排名。从这个排名中, 最不确定的样本被提交给预言机进行分割, 并且最不确定的图像被包含在标记数据集中, 如下所示 伪真实情况标签。他们表明, 随着交互式注释样本百分比的增加, 他们的方法的准确性不断提高, 仅使用 80% 的标签即可达到完全监督的基准方法的性能。这项工作推动了使用 GAN 判别器分数作为预测不确定性的衡量标准。

玛哈帕特拉等人。(2018)还使用 cGAN 生成以真实图像为条件的胸部 X 射线图像, 并使用贝叶斯神经网络评估每个生成样本的信息量, 决定每个生成样本是否应用作训练数据。如果是这样, 则用于微调网络。他们证明, 该方法可以实现与在完全注释的数据上进行训练相当的性能, 使用训练集中只有 33% 的像素进行注释的数据集, 为注释者节省大量时间、精力和成本。

赵等人。(2019)通过使用学习到的转换, 提出了一种替代 GAN 的数据合成方法。他们从单个手动分割的图像中, 以类似 SSL 的方法利用其他未注释的图像, 从图像中学习转换模型, 并使用该模型和标记数据来合成其他带注释的样本。变换包括空间变形和强度变化, 以便能够合成复杂的效果, 例如解剖和图像采集变化。他们以监督的方式训练模型

MRI 脑图像分割，并显示出比其他一次性生物医学图像分割方法最先进的改进。

基于 GAN 的方法在 AL 场景中的实用性超出了单模态图像合成的范围。许多工作已经证明了 GAN 执行跨模态图像合成的能力，它不仅直接解决了训练数据有限的问题，而且还解决了多模态分析场景中出现的模态丢失问题。可以生成缺失模态来填充缺失数据点的方法，从而使全套 AL 方法能够应用于多模态分析问题。

王等人。(2019)介绍一种基于 GAN 的方法，用于跨不同显微镜模式的超分辨率。这项工作使用 GAN 将衍射受限输入图像转换为超分辨率图像，提高使用低数值孔径物镜获取的宽视场图像的分辨率，以匹配使用高数值孔径物镜获取的分辨率。这项工作扩展了这种方法，以展示跨模态超分辨率，将共焦显微镜图像转换为受激发射损耗显微镜获得的分辨率。这种方法使得以较低分辨率获取的多种类型的图像能够实现超分辨率，以匹配较高分辨率的图像，从而使 AL 及其他领域的多模态图像分析方法具有更高的性能。

王等人。(2018)介绍一种基于 GAN 的方法，用于生成高质量 PET 图像，通常需要全剂量放射性示踪剂才能获得。这项工作使得低剂量示踪剂能够用于获取低质量 PET 图像，然后使用 3D 条件 GAN（以低剂量图像为条件）从中生成高质量 PET 图像。除此之外，还引入了基于 3D c-GAN 的渐进细化方案，以进一步提高估计图像的质量。通过这项工作，获取高质量 PET 图像所需的放射性示踪剂剂量大大减少，减少了对患者的危害，并使低剂量 PET 图像能够与高剂量图像一起用于下游分析。

余等人。(2019)扩展现有的基于 GAN 的方法，以改进在不同扫描参数下采集的 MR 图像的跨模态合成。他们的工作引入了边缘感知生成对抗网络（Ea-GAN），该网络专门集成了反映图像内容纹理结构的边缘信息来描绘图像中不同对象的边界，这超越了仅专注于最小化像素的方法。明智的强度差异。他们使用两种学习策略将边缘信息引入生成器诱导的 Ea-GAN（gEa-GAN）和判别器诱导的 Ea-GAN（dEa-GAN），分别通过生成器以及生成器和判别器合并边缘信息，因此边缘相似性也是对抗性学习的。

潘等人。(2020)探索使用 GAN 从相应的 MR 图像中估算缺失的 PET 图像，以使用基于 GAN 的方法识别脑部疾病，以避免丢弃数据缺失的受试者，从而增加可用的训练样本数量。使用混合 GAN 生成缺失的 PET 图像，然后使用空间约束的 Fisher 表示网络提取神经图像的统计描述符以进行疾病诊断。三个数据库的结果表明，与其他最先进的方法相比，该方法可以合成合理的神经图像，并在脑部疾病识别方面取得有希望的结果。

上述工作展示了使用以极少量注释数据为条件的合成数据来生成的能力

使用可用于高精度训练模型的新训练样本，这对于 AL 方法非常有价值，在这种方法中，我们通常需要初始训练集来训练模型，然后才能采用数据选择策略。这些方法还展示了标记数据的有效使用，并允许我们从单独注释的图像生成多个训练样本，这可能允许通过生成多个在 AL/Human-in-the-Loop 方法中获得的注释数据来更有效地使用为单个请求的注释训练样本，进一步减少训练最先进模型所需的注释工作。

2.2.4. 学习主动学习

到目前为止讨论的大多数方法都采用手工设计的信息启发式方法，但已经出现了一些工作，试图通过以前的样本选择结果的经验来了解信息最丰富的样本是什么。这提供了一种更有效地选择样本的潜在方法，但代价是所采用的启发式方法的可解释性。许多因素会影响使用手工启发式数据选择的性能和优化性。科纽什科娃等人。(2017)提出“学习主动学习”，其中回归模型根据先前 AL 结果的经验来学习数据选择策略。认为无法预见所有因素的影响，例如类别不平衡、标签噪声、异常值和分布形状。相反，他们的回归模型会根据问题“调整”其选择，而无需明确说明具体规则。巴赫曼等人。(2017)将这个想法更进一步，提出一个模型，利用来自不同但相关任务的标记实例来学习选择策略，同时调整其数据表示和预测函数。

强化学习 (RL) 是机器学习的一个分支，它使“代理”能够在交互式环境中通过反复试验、使用来自其自身行为和经验的反馈来学习，努力实现系统的既定目标。主动学习最近被认为是强化学习的一个潜在用例，并且一些工作已经开始探索这一领域。

伍德沃德等人。(2017)提出一种与强化学习相结合的一次性学习方法，让模型在推理过程中决定哪些例子值得标记。呈现图像流并做出决定，要么预测标签，要么付费接收正确的标签。通过选择 RL 奖励函数，他们能够实现比纯粹监督任务更高的预测精度，或者用预测精度换取更少的标签请求。

方等人。(2017)将数据选择过程重新构建为强化学习问题，并明确学习数据选择策略。这与 AL 框架中常见的数据选择启发法无关，提供了一种更通用的方法，展示了实体识别的改进，但这尚未应用于医学图像数据。

强化学习方法为解决 AL 和人在环问题提供了一种不同的方法，该方法非常适合帮助支持 DL 的应用程序与其最终用户之间的实时反馈，但是它需要特定于任务的目标，而这些目标可能无法在不同的任务中通用。医学图像分析任务。

2.3. 微调与再训练

每个 AL 框架的最后一步是使用新获得的注释来改进模型。有两种主要方法用于根据新注释训练模型。这些是使用包括新获取的注释在内的所有可用数据重新训练模型，或者仅使用新注释或新注释加上现有注释的子集来微调模型。

泰巴赫什等人。(2016)研究在几个医学图像分析任务中使用迁移学习和微调并进行演示

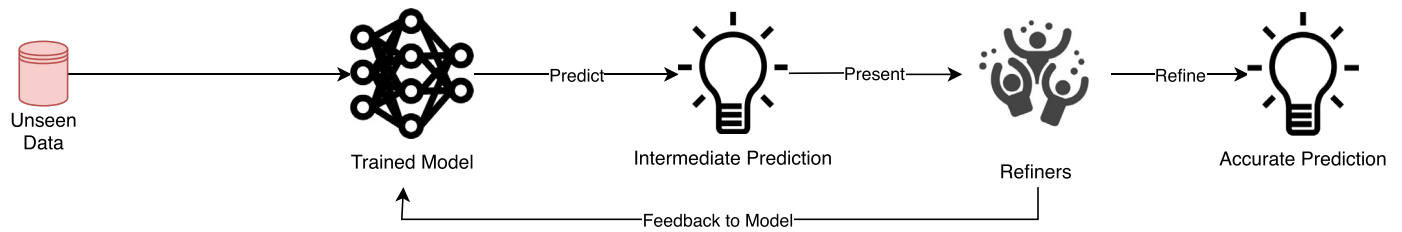


图 2. 细化框架概述。

表明使用经过微调的预训练 CNN 的性能优于从头开始训练的 CNN，并且这些微调的 CNN 对于训练集的大小更加稳健。他们还表明，浅层调整和深度调整都不是特定应用的最佳选择，并提出了一种分层训练方案，该方案可以提供一种实用的方法，根据可用数据量为所选任务实现最佳性能。这项工作中采用的方法执行一次性微调，其中预训练模型仅使用可用训练样本微调一次，但这不适应主动选择过程或连续微调。

周等人。(2017)提出了一种连续微调方法，可以使用逐渐增大的数据集微调预训练的 CNN，并证明该方法比反复微调预训练的 CNN 收敛得更快。他们还发现，仅使用新获取的注释进行持续微调需要仔细的元参数调整，这使得它在许多不同的任务中不太实用。

受上述两种主要方法启发，另一种从新数据进行重新训练的方法是使用所有可用数据重新训练模型，但使用之前的参数作为初始化，但是在作者的任何作品中，这种方法尚未应用于 AL 都知道。

再训练的计算成本比微调更高，但它提供了评估 AL 框架性能的一致方法。微调用于许多不同的 ML 领域，例如一次或几次学习以及迁移学习，而实现这一点的最佳方法仍然是一个悬而未决的问题，因此在 AL 框架中不太普遍，随着微调的改进，我们可能会看到它在 AL 框架中的使用的转变。重要的是建立基线微调 and 再训练方案，以有效比较应用它们的 DL/AL 方法，以便将这些方案的效果与其他领域的改进隔离开来。

3. 最终百分比：模型输出的交互式细化

到目前为止，我们已经考虑了人类在注释用于训练模型的数据中的作用，但是一旦训练了模型，我们仍然需要一个人在循环中来解释模型预测，并可能对其进行改进以获得未见数据的最准确结果，如中所述图2。在人机交互场景中，模型对看不见的输入进行预测，并且根据接受标准，自动预测可能需要手动调整才能满足这些接受标准。有关预测的信息交流对于自信地满足验收标准并形成对自动预测局限性的理解非常重要。这种通信有两个方面，即用户必须能够与用于指导预测获得更准确的结果或纠正错误预测的模型进行通信，并且模型必须能够与用户进行通信以提供对模型预定义的有意义的解释。

词典，使用户能够在与模型输出交互时采取最佳行动方案，并减轻人类的不确定性。这将创建反馈循环，如下所示图2。

3.1. 互动细化

即使我们能够为医学图像应用开发准确、稳健和可解释的模型，我们仍然无法保证自动预测满足呈现给模型的每个看不见的数据点的接受标准。泛化到看不见的输入的能力是深度学习应用的基石，但在现实世界的分布中，泛化很少是完美的。因此，纠正这些差异的方法必须内置到用于医学图像分析的应用程序中。与执行手动注释或纯手动校正相比，这种迭代细化必须节省最终用户的时间和精力。人们已经提出了许多交互式图像分割系统，最近这些系统都建立在深度学习的进步之上，允许用户细化模型输出并将更准确的结果反馈给模型以进行改进。

阿姆雷恩等人。(2017)推出了 UI-Net，它建立在流行的 U-Net 架构之上，用于医学图像分割 罗尼伯格等人。(2015)。UI-Net 经过训练活跃用户模型，并允许用户通过提供与建议的细分进行交互 涂鸦通过在图像上指示应包含或不包含的区域，网络使用模拟用户交互进行训练，并因此响应迭代用户涂鸦以细化分割以获得更准确的结果。

条件随机场已用于各种任务中以鼓励分割同质性。郑等人。(2015)提出 CRF-CNN，一种循环神经网络，具有 CNN 和 CRF 的理想特性。王等人。(2019)提出 DeepIGeoS，一种用于医学图像分割的交互式测地线框架。该框架使用两个 CNN，第一个执行初始自动分割，第二个执行初始分割以及用户与初始分割的交互以提供精炼结果。他们通过测地距离变换将用户交互与 CNN 结合起来 克里米尼西等人。(2008)，这些用户交互作为硬约束集成到条件随机场中，灵感来自 郑等人。(2015)。他们将这两个网络称为 P-Net（初始分割）和 R-Net（用于细化）。与全自动 CNN 相比，他们展示了从 2D 胎儿 MRI 分割胎盘和从 3D FLAIR 图像分割脑肿瘤的优异结果。获得这些分割结果的时间大约是使用 GeoS 或 ITK-SNAP 等传统交互式方法执行相同分割所需时间的三分之一。

图割也被用于分割以合并用户交互——用户提供种子点到算法（例如，将一些像素标记为前景，将另一个像素标记为背景），并据此计算分割。王等人。(2018)提出 BIFSeg，一个交互式分割框架，其灵感来自于

图形切割。他们的工作通过将 CNN 与边界框和基于涂鸦的分割管道相结合，引入了用于交互式分割的深度学习框架。用户在他们感兴趣的分割区域周围提供一个边界框，然后将其输入到 CNN 中以生成初始分割预测，然后用户可以提供涂鸦来将图像区域标记为错误分类 - 这些用户输入然后使用基于图割的算法在精细分割的计算中给予很大的权重。

布雷德尔等人。(2018)提出了 BIFSeg 的替代方案，其中训练两个网络，一个用于执行初始分割（他们使用 CNN，但该初始分割可以使用任何现有算法执行），第二个网络被他们称为 inter-CNN，将图像作为输入，一些用户涂鸦和初始分割预测并输出精细分割，他们表明，通过对多个用户输入进行多次迭代，分割的质量比初始分割有所提高，并与其他交互式相比实现了最先进的性能方法。

到目前为止，上面讨论的方法主要涉及对单个图像或切片进行分割，但是许多分割任务寻求提取特定感兴趣区域 (ROI) 的 3D 形状/表面。库岑多夫等人。(2017)提出了一种基于智能画笔 2D 分割的双重方法来生成 3D 分割，用户引导用户进行良好的 2D 分割，并且在分割几个切片后，使用 Hermite 径向基函数将其转换为 3D 表面形状，从而实现高准确性。虽然这种方法不使用深度学习，但它是交互式分割可用于生成用于深度学习应用的高质量训练数据的有力示例 - 他们的方法是通用的，可以为大量任务生成分割。有可能将深度学习纳入其流程中，以改善结果并加速交互式注释过程。

张和金 (2019)提出了一种交互式分割方案，该方案可推广到任何先前训练的分割模型，该模型接受有关目标对象和背景的用户注释。通过测量每个像素到带注释的地标的距离，将用户注释转换为交互图，之后前向传递输出初始分割。用户注释的点可能在初始分割中被错误分割，因此他们提出了 BRS（反向传播细化方案）来纠正错误标记的像素。他们证明，他们的算法在多个数据集上优于传统方法，并且 BRS 可以通过将现有 CNN 转换为用户交互版本来推广到医学图像分割任务。

廖等人。(2020)提出将迭代交互细化的动态建模为马尔可夫决策过程 (MDP)，并使用多智能体强化学习来解决这个问题。将每个体素视为具有共享体素级行为策略的代理，以这种方式使体素预测变得易于处理。多智能体方法成功地捕获了分割任务的体素之间的依赖关系，并且通过将先前分割的预测不确定性通过状态空间传递可以得到更精确和更精细的分割。使用这种方法，它们的交互次数更少，收敛速度更快，显著优于现有的最先进方法。

在本节中，我们重点关注与迭代细化细分以获得所需输出质量相关的应用程序。在上述场景中，这是在最终用户提供的未见过的图像上执行的，但是没有理由不能采用相同的方法来迭代地生成用于训练的更准确的注释，例如，使用主动学习来选择接下来要注释哪些样本，并迭代地细化预测

直到获得足够准确的注释为止。这有可能加速训练注释，而无需任何额外的实现开销。AL 中所做的许多工作都忽略了预言机的作用，只是假设我们可以在需要时获得准确的标签，但实际上这提出了更重大的挑战。我们预计，随着 AL 研究改进对提供注释的预言机的考虑，AL 和 HITL 计算将变得更加紧密地耦合。

用户如何细化医学图像的分割是相当直观的，但其他医学图像分析任务的情况并非如此。涉及分类和回归的临床任务预测的改进比分割方面的发展要少，并且仍然是一个开放的研究领域。以下工作已采取措施解决分类和回归任务的交互式细化策略。

廉等人。(2020)探索使用 CNN 方法自动诊断阿尔茨海默病，并发现许多最先进的方法依赖于结构 MRI (sMRI) 中信息位置的预先确定。这一阶段的判别性定位与特征提取和分类器构建的后期阶段是隔离的。他们的工作提出了一种分层全卷积 CNN (H-FCN) 来自动识别全脑 sMRI 中具有辨别力的局部斑块和区域，从中可以联合学习和融合多尺度特征表示以构建分类模型。这项工作可以实现补丁选择和分类器构建的交互式细化，如果人类最终用户进行干预，可以引导网络走向更具辨别力的兴趣区域，从而获得更有效的分类器。

相似地，刘等人。(2018)引入基于地标的深度多实例学习 (LDML) 框架用于脑部疾病诊断。首先，通过采用数据驱动的方法来发现大脑 MR 图像中与疾病相关的解剖标志以及附近的图像块。其次，该框架学习端到端 MR 图像分类器，用于捕获所选地标块中的局部结构信息，以及从所有检测到的地标中导出的全局结构信息。通过分割地标检测和分类器构建的步骤，可以引入人机交互来干预选定的地标并引导网络走向信息量最大的图像区域。因此，可以通过更新图像的哪些区域用作输入来细化所得到的分类器。

3.2. 互动口译

在上一节中，我们讨论了人机交互系统的用户可以与预测模型进行通信的方法，在本节中，我们考虑模型可以与用户进行通信的方法，从而完成反馈循环图2。根据上下文，“解释”可能意味着许多不同的事情，因此在这里我们重点关注模型输出的解释，目的是在下游分析中适当加权自动预测（例如预测的不确定性），并使用户能够做出最明智的更正或手动调整模型预测（例如“注意力门控”奥克斯等人。(2018)）。

虽然深度学习已成为许多医学图像分析任务的标准最先进方法，但它们在很大程度上仍然是黑盒方法，最终用户解释模型预测的有意义的方法有限。深度学习方法的这一特性是将支持深度学习的应用程序部署到医学图像分析等安全关键领域的一个重大障碍。我们希望模型高度准确和稳健，但同时也是可解释和可解释的。这种可解释性对于减轻人类的不确定性并培养对在具有现实世界后果的下游任务中使用自动预测的信任至关重要。

最近的欧盟法律¹导致了“解释权”，即任何主体都有权针对他们做出的自动决策进行解释。这进一步凸显了对透明算法的需求，我们可以对此进行推理[古德曼与弗拉克斯曼 (2017), 爱德华兹和维尔 (2017a), 爱德华兹和维尔 (2017b)]。

对于用户来说，了解模型如何做出特定决策非常重要，因为即使是最准确、最稳健的模型也并非万无一失，必须识别错误或不确定的预测，以便培养对模型的信任和预测在临床决策过程中适当权衡。最终用户、监管机构和审计人员都有能力将深度学习模型生成的自动化决策置于情境中，这一点至关重要。在这里，我们概述了一些不同的方法，用于提供 DL 模型及其预测的可解释的推理方式。

通常，深度学习方法可以提供有关模型输出不确定性的统计指标，其中讨论的许多不确定性指标第2节也用于帮助可解释性。虽然不确定性度量很重要，但这些不足以培养对深度学习模型的完全信任，该模型应该为其输出提供人类可理解的理由，以便能够得出见解，阐明模型的内部工作原理。查克拉博蒂等人。(2017)讨论围绕模型不可解释性的许多核心问题，并重点介绍各种研究成果，这些研究成果展示了使深度学习模型在整个深度学习领域可解释的复杂方法。在这里，我们评估了一些已应用于医学图像分割的工作，并建议读者参考[斯托亚诺夫等人。(2018), 霍尔金格等人。(2017)]进一步阅读有关医学成像领域其他部分的可解释性。

奥克斯等人。(2018)和施伦珀等人。(2019)引入“注意力门控”来引导网络以视觉上可解释的方式对某些图像区域给予更多“注意力”——可能有助于随后的注释细化。流行的 U-Net 架构中引入了注意力门 (Attention Gates) (罗尼伯格等人, 2015)，其中从粗尺度层提取的信息用于选通，以在连接之前消除跳跃连接中不相关和噪声响应的歧义，以仅合并相关层激活。这种方法消除了图像分割和回归任务中应用外部对象定位模型的需要。注意力门层系数指示图像特征激活将被允许传播到最终预测的位置，为用户提供模型在进行预测时高度加权的图像区域的视觉表示。

在巴德等人。(2019)我们提出了一种视觉方法来解释超声图像中的自动头围测量值，在测试时使用 MC Dropout 来获取 N 个头部分割，以实时计算头围测量值的上限和下限。这些边界显示在图像上，以引导超声检查人员查看模型最有信心预测的视图。该上下限是作为未见图像的模型合规性的度量而不是不确定性的度量。最后，提出了方差启发法来量化预测的置信度，以便接受或拒绝头围测量结果，并且表明，一旦删除“拒绝”图像，这些可以改善整体性能测量。

米莱塔里等人。(2019)提出将强化学习应用于超声护理，指导可能缺乏经验的用户进行正确的操作

声波窗口并使他们能够获得感兴趣的解剖结构的临床相关图像。这种人机交互应用程序是将深度学习/强化学习与实时系统相结合而实现的新颖应用程序的一个示例，使用户能够响应模型反馈以获取最准确的可用信息。

王等人。(2019b)建议使用测试时间增强来获取任意（基于图像）不确定性的度量，并将其方法与认知（模型）不确定性度量进行比较，并表明他们的方法比单独基于测试时间丢失的模型不确定性提供更好的不确定性估计，并且减少过度自信的错误预测。

琼戈和雷耶斯 (2019)评估应用于医学图像分割的几种不同的体系不确定性估计方法的可靠性和局限性，并表明当前的不确定性估计方法表现相似。他们的结果表明，虽然不确定性估计可以在数据集级别上得到很好的校准（捕获认知不确定性），但它们往往在主题级别上被错误校准（任意不确定性）。这损害了这些不确定性估计的可靠性，并凸显了开发主题不确定性估计的必要性。他们表明辅助网络是常见不确定性方法的有效替代方法，因为它们可以应用于任何先前训练的分割模型。

开发透明的系统将能够更快地应用于临床实践，并将人类纳入深度学习临床管道中，这将缩短当前最佳实践与深度学习所提供的可能增强的广度之间的过渡期。

我们认为，正在进行的提高 DL 模型可解释性的工作也将对 AL 产生积极影响，因为大多数提高可解释性的方法都集中在为模型预测提供不确定性度量，这些相同的不确定性度量可用于 AL 选择策略代替目前采用的现有不确定性措施。随着不可解释性和不确定性度量的改进，我们预计 AL 框架也会有类似的改进，因为它们包含了最有前途的不确定性度量。

中讨论的方法第3节仍然是研究兴趣的开放领域，对 AL 开发的进展以及在临床实践中更多地采用 DL 和 HITL 方法具有重大影响。用户和模型之间交互的研究变得越来越重要，并且对深度主动学习系统的功效及其在实际应用中的部署产生重大影响，特别是在临床场景中（比德等人, 2020；阿姆雷恩等人, 2019）。对机器学习可解释性的更广泛研究和人机交互的研究可能看起来截然不同且存在分歧，但随着人类用户和机器模型之间的反馈循环变得越来越重要，我们预计这两个研究领域将通过主动学习实现融合。

4. 实际考虑

到目前为止，我们已经讨论了 AL 背后的核心工作，即模型解释和预测细化，虽然上面讨论的工作在涵盖大多数正在进行的研究方面大有帮助，但开发和部署支持 DL 的应用程序有几个实际考虑因素我们必须考虑这一点。在本节中，我们概述了影响 DL 应用程序开发流程的主要实践研究领域，并建议我们下一步可能关注的方向（图3）。

4.1. 喧闹的神谕

医学图像数据的黄金标准注释是通过聚合来自多个专家预言机的注释来取得的，但作为

¹关于在个人数据处理和此类数据自由流动方面保护自然人的条例 (EU) 2016/679，并废除指令 95/46/EC（一般数据保护条例）[2016] OJ L119/1

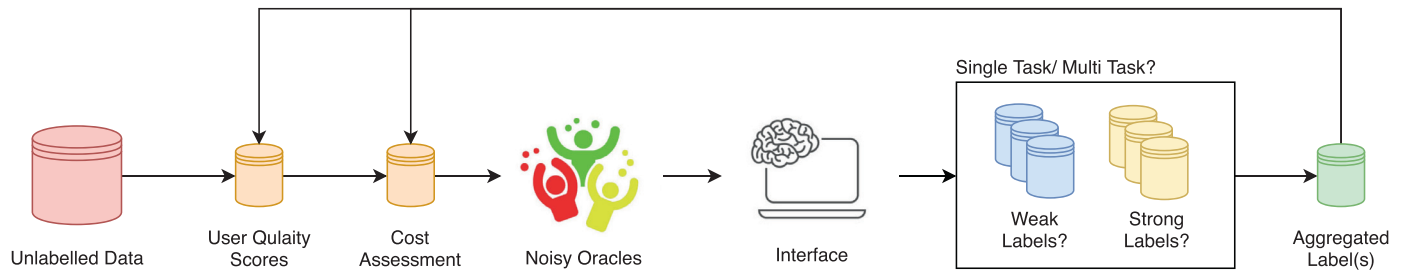


图 3. 实际考虑因素概述。

正如前面所讨论的，由于执行此类注释所需的专业知识，对于大型复杂数据集来说，这几乎是不可行的。在这里，我们问如果我们从没有领域专业知识的预言机获取标签，可能会对性能产生什么影响，以及我们可以使用哪些技术来减轻使用非专家预言机时可疑的注释质量下降，以避免任何潜在的准确性损失。

张和乔胡里 (2015) 提出主动学习方法，假设数据将由一群非专家或“弱”注释者进行注释，并提供减少将不良标签引入数据集的方法。他们同时了解各个注释者的质量，以便最强的注释者可以标记信息最丰富的示例。

李等人。(2019)提出了两种场景下的众包学习方法。首先，他们的目标是通过从两个不同的角度对人群的专业知识和标签相关性进行建模，在给定人群注释的情况下推断实例的基本事实：首先，他们基于个体标签对专业知识进行建模，基于标记者对相似实例的注释应该相似的想法，并且其次，通过对人群的专业知识进行建模来区分标签对之间的相关性。他们将方法扩展到主动范式，并提供标准，例如串联选择的标签和贴标机，以最大限度地降低注释成本。

切普利吉娜等人。(2016)探索使用 Amazon 的 MTurk 收集 CT 图像中气道的注释。结果表明，新手预言机能够解释图像，但提供的指令过于复杂，导致许多注释无法使用。一旦删除了不良注释，这些注释确实显示出与专家注释的中等到高度相关性，特别是在聚合注释的情况下。

罗德里格斯和佩雷拉 (2018) 描述一种评估人群中注释者可靠性的方法，以及用于从多个注释者的噪声标签训练深度模型的人群层，在内部捕获不同注释者的可靠性和偏差，以实现多个注释者的最先进结果众包数据集任务。

我们可以看到，通过使用预言机注释质量的学习模型，我们可以减轻低质量注释的影响，并向最有能力的预言机呈现最具挑战性的案例。通过提供清晰的指令，我们可以降低非专家预言机执行准确注释的障碍，但这不是普遍适用的，并且是我们希望执行的每个新注释任务所必需的。

4.2. 弱监督学习

大多数分割任务需要逐像素注释，但这并不是我们可以为图像提供的唯一注释类型。可以使用“弱”注释来执行分割，其中包括图像级标签（例如模式、存在的器官等）和注释（例如边界框、椭圆或涂鸦）。有人认为，使用“较弱”的注释公式可以使人类预言机的任务变得更容易，从而产生更准确的注释。

站。“弱”注释已被证明在多个分割任务中表现良好，拉杰切尔等人。(2016a) 演示在给定具有“弱”边界框注释的图像数据集的情况下获得像素级分割。他们提出了 DeepCut，一种将 CNN 与迭代密集 CRF 公式相结合的架构，以实现良好的准确性，同时大大减少所需的注释工作。在后来的研究中，拉杰切尔等人。(2017) 检查不同“弱”注释类型所需的专业知识对肝脏分割准确性的影响。结果显示，正如预期的那样，所有注释类型的准确性都会随着专业知识的减少而降低。尽管如此，当使用弱标记数据集进行异常值校正时，分割精度与最先进的性能相当。他们的方法的稳健性能表明，来自非专家群体的“弱”注释可用于在许多不同的任务上获得准确的分割，但是他们对数据集的使用使得这种方法的通用性低于预期。

在拉杰切尔等人。(2016b) 他们研究使用超级像素来加速注释过程。该方法使用预处理步骤来获取每个图像的超像素分割，然后使用非专家通过选择哪些超像素是目标区域的一部分来执行注释。结果表明，该方法很大程度上减少了用户的注释负担。12 名注释者在一小时内完成了 5000 个切片的非专家注释，而专家需要三个工作日才能通过高级界面建立相同的注释。非专家界面基于网络，展示了分布式注释收集/众包的潜力。本文的一个令人鼓舞的方面是，与专家注释性能相比，结果显示了所讨论的分割任务的高性能，

事实证明，我们可以使用弱注释数据开发高性能模型，并且由于弱注释需要较少的专业知识来执行，因此可以更快地从非专家群体中获取它们，并且与黄金标准注释相比，准确性损失更小。这对于未来的研究来说非常有希望，因为弱注释数据的数据集可能更容易管理，也更具成本效益。

4.3. 多任务学习

许多工作旨在一次训练模型或获取多个任务的训练数据，有人认为这可以节省成本，因为补充信息可能会在多个不同任务上带来更高的性能（莫斯科普斯等人，2016）。王等人。(2019a) 提出了一种用于联合分割和检测任务的双网络，用于 CT 图像中的肺结节分割和耳蜗分割，其中只有一部分数据被密集注释，其余数据被边界框弱标记，利用这一点，他们表明他们的架构优于执行多个基线。目前这项工作仅处理两种不同的情况

标签类型，但他们建议扩展真正的多任务场景的框架。

这是一个充满希望的领域，但到目前为止，它尚未被纳入主动学习环境中。因此，当模型同时训练多个任务时，分析不同 AL 方法选择的样本之间的差异可能会有所说明。然而，洛厄尔等人。(2019)由于主动学习选择策略和正在训练的模型之间的固有耦合，引起了对主动获取的数据集到未来模型的可转移性的担忧，并表明在主动获取的数据集上训练后继模型通常会导致比随机采样更差的性能。他们认为，随着数据集开始比在其上训练的模型寿命更长，人们会担心主动学习的有效性，因为获取的数据集可能不利于训练后继模型。可能需要探索主动获取的数据集如何在多个模型上执行，以解释主动获取的数据集与一个模型相结合对相关模型性能的影响。

4.4. 注释接口

到目前为止，大多数人机交互方法都假设预言机与数据注释和模型预测之间存在显著的交互，但很少有人考虑预言机可能与这些图像交互的接口的性质。在提出分布式在线平台来执行此类注释时，需要特别注意医学图像的性质。虽然到目前为止讨论的大多数技术都使用预先存在的数据标签代替新获取的标签来展示其性能，但重要的是要考虑实际界面可能产生的注释准确性的影响。

纳利斯尼克等人。(2015)提出了一个对组织的全幻灯片图像 (WSI) 进行在线分类的框架。他们的界面使用户能够使用主动学习过程快速构建分类器，最大限度地减少标记工作，并展示其解决方案对神经胶质瘤肿瘤定量测量的有效性。

科斯拉万等人。(2017)提出了一种用于图像分割的新颖界面，该界面跟踪用户的注视以启动用于分割感兴趣对象的种子点，作为与图像交互的唯一手段，从而实现高分割性能。斯坦伯等人。(2019)扩展这个想法，并将使用眼动追踪生成的训练样本与传统的手工注释训练样本进行比较，以训练深度学习模型。他们表明，使用通过眼动追踪生成的注释可以实现几乎相同的性能，并表明这种方法可能适用于快速生成训练数据。他们承认，仍然需要改进，以更快、更自然和更少分散注意力的方式将眼动追踪集成到典型的临床放射学工作流程中。

蒂纳蒂等人。(2017)评估 Eye-Wire 背后的玩家动机，这是一款在线游戏，要求一群玩家帮助分割小鼠大脑中的神经元。这项任务的游戏化已经有超过 500,000 名玩家注册，并且获得的细分已被用于多项研究工作中 [金等人。(2014)]。游戏化最令人兴奋的事情之一是，在调查中，用户最受激励的是做出科学贡献，而不是任何潜在的金钱奖励。然而，这是非常专门针对这个特定任务的，并且很难应用于其他类型的医学图像分析任务。

开发注释界面有许多不同的方法，我们上面考虑的只是已应用于医学图像分析的几种方法。随着开发的增加，我们期望看到更多的在线工具用于医学图像分析，并且所选的界面格式将在这些应用程序的可用性和整体成功中发挥重要作用。

4.5. 可变的学习成本

当从各种类型的预言机获取训练数据时，值得考虑与查询特定预言机类型以获取该注释相关的相对成本。我们可能希望从专家预言机获得更准确的标签，但这可能比从非专家预言机获得更昂贵。当然，代价是所获得的标签的准确性 - 预言机的专业知识较少可能会导致注释质量较低。已经提出了几种方法来对此进行建模，并允许开发人员在成本和所获取注释的整体准确性之间进行权衡。

郭等人。(2018)提出一种用于颅内出血检测的成本敏感的主动学习方法。由于示例之间的注释时间可能会有很大差异，因此他们对注释时间进行建模并优化投资回报。他们展示了他们的方法选择了一组多样化且有意义的样本来进行注释，相对于统一的成本模型，该模型主要选择大量流失的样本，而这些样本的注释非常耗时。

沙阿等人。(2018)通过密集分割、边界框和地标，在混合监督设置（强注释和弱注释）中提出基于预算的成本最小化框架。他们的框架使用不确定性和代表性排名策略来选择接下来要注释的样本。它们以显著减少的培训预算展示了最先进的性能，强调了注释类型的选择对获取培训数据的成本的重要作用。

上述工作均显示了对整理训练数据时产生的经济负担的改进考虑。一个有价值的研究方向是在统一的框架中评估预言机专业水平、注释类型和图像注释成本的影响，因为这三个因素密切相关，并且可能相互产生深远的影响。

5. 未来的前瞻性和未解答的问题

在第 2 节和 3 我们讨论了用户可以收集训练数据来构建模型、使用模型来预测新数据并接收反馈以迭代地改进模型输出以获得更准确的结果的方法。这些技术中的每一种都假设一些人类最终用户在初始注释、解释和细化时在场与系统交互。这些领域中的每一个都试图实现一个共同的目标，即通过尽可能少的注释数据实现最高性能的模式 - 并采用适当权衡模型预测结论的方法。

AL 并不是唯一一个旨在从有限数据中学习的研究领域。半监督学习和迁移学习都为从有限的标记数据中提取最大价值做出了重大贡献。

在存在大型数据集但缺乏标签的情况下，无监督和半监督方法提供了一种无需为所有数据点添加标签即可提取信息的方法。这可能会对经常出现这种情况的医学图像分析领域产生巨大影响。

在半监督学习 (SSL) 场景中，我们可能有一些标记数据，但这通常非常有限。然而，我们确实有大量未注释的实例（很像主动学习）可以从中获取信息，目标是使用未标记的实例改进模型（仅在标记的实例上进行训练）。由此我们得出两个不同的目标：a) 预测未来数据的标签（归纳式 SSL）和 b) 预测可用的未注释数据的标签（传导式 SSL）（切普利吉娜等人，2019；夏等人，2020）。SSL 方法提供了一种从未注释的图像数据中提取有用信息的强大方法，我们

相信这一领域的进展将有利于 AL 系统，因为它们需要更准确的初始化模型来指导数据选择策略。

迁移学习 (TL) 是深度学习的一个分支，旨在使用预先训练的网络作为新应用程序的起点。给定针对特定任务训练的预训练网络，事实证明，该网络可以根据有限的训练数据针对目标任务进行“微调”。我们建议读者参考莫里德等人。(2021);拉古等人。(2019);切普利吉娜等人。(2019)了解医学成像中迁移学习的更一般概述，并在下面重点介绍 TL 在 AL 场景中的使用。泰巴赫什等人。(2016)证明了 TL 对于各种医学图像分析任务的适用性，并表明，尽管自然图像和医学图像之间存在很大差异，但在自然图像上进行预训练并在医学图像上进行微调的 CNN 可以比从头开始训练的医学 CNN 表现更好。当可用的目标任务训练示例较少时，这种性能提升会更大。到目前为止讨论的许多方法都是从对自然图像数据进行预训练的网络开始的。

周等人。(2018a)提出 AFT^{*}，一个结合 AL 和 TL 来减少标注工作的平台，旨在解决 AL 中的几个问题。非洲金融时报从完全空的标记数据集开始，不需要种子样本。预先训练的 CNN 用于寻找“有价值”的样本进行注释，并通过持续微调逐步增强 CNN。我们采取了许多措施来尽量减少灾难性遗忘的风险。他们之前的作品周等人。(2017)对几个医学图像分析任务应用了类似但功能较少的方法，以证明通过大幅减少训练数据集可以达到相同的性能。然后，他们使用这些任务来评估网络表现出的几种预测模式以及这些模式与 AL 选择标准的选择有何关系。

周等人。(2018b)已经开始使用他们的 AFT 框架来注释 CIMT 视频，这是一种用于表征心血管疾病的临床技术。他们向视频领域的扩展提出了自己独特的挑战，因此他们提出了注释单元的新概念 - 将 CIMT 视频注释减少到只需 6 次用户鼠标点击，并将其与 AFT 框架相结合，相对降低注释成本 80% 从头开始训练，相对于随机选择要注释的新样本（并用于微调），提高了 50%。

库什巴尔等人。(2019)使用 TL 进行监督域适应，以最少的用户交互进行皮下层下大脑结构分割。他们通过利用预训练网络显著减少来自不同 MRI 成像领域的训练图像数量，并通过减少 CNN 中可训练参数的数量来提高训练速度。他们表明，他们的方法在使用目标域中的极少量图像的情况下取得了与基线相似的结果，并且表明，即使使用目标域中的一张图像也足以优于他们的基线。

上述方法以及本综述中讨论的更多方法证明了 TL 的适用性，可以减少从有限的训练数据中训练新任务模型所需的带注释样本的数量。通过使用在带注释的自然图像数据（有大量）上训练的预训练网络，我们可以提高模型性能并进一步减少实现最先进行性能所需的注释工作。

TL 的一个值得探索的相关子领域是域适应 (DA)。许多用于医学图像分析的深度学习技术都面临着由于源数据和目标数据之间的不同分布而引起的域转移问题，这通常是由于医学图像是在各种不同的扫描仪、扫描参数和受试者群体上获取的。DA 被提出作为一种特殊类型的迁移学习，其中域特征空间和任务保持不变，而源的边缘分布

和目标域不同。我们建议读者参考关和刘 (2021);乔杜里等人。(2020)概述了用于医学图像分析的 DA 方法，希望未来能够看到 DA 方法在 AL 场景中有更多的应用。

在本次评论中描述的许多场景中，模型不断接收新的注释用于训练，理论上我们可以继续无限期地重新训练或微调模型，但这是否实用且具有成本效益？重要的是要量化使用新数据训练模型的长期影响，以评估模型如何随时间变化以及性能是否有所改善，或者更糟，是否下降。事实证明，从连续的数据流中学习比预期更困难，常常导致“灾难性遗忘”或“干扰”帕里西等人。(2019)。我们面临着稳定性-可塑性-困境。当从连续的数据流中学习时，避免神经网络中的灾难性遗忘可以大致分为三种概念策略：(a) 在正则化的同时重新训练整个网络（以防止忘记以前学习的任务）。(b) 有选择地训练网络并在需要时扩展网络来表示新任务，以及 c) 保留以前的经验，以便在没有新输入的情况下使用记忆重放进行学习。我们建议读者参考帕里西等人。(2019)有关这些方法的更详细概述。

巴韦贾等人。(2018)研究使用神经网络持续学习两个 MRI 分割任务，以防止学习新任务时对第一个任务的灾难性遗忘。他们研究了弹性权重巩固，这是一种基于费舍尔信息的方法，顺序学习正常大脑结构的分割，然后分割白质病变，并证明该方法减少了灾难性遗忘，但承认对于连续学习的挑战性环境，还有很大的改进空间。学习。

量化模型在其生命周期的每个阶段的性能和稳健性非常重要。考虑停止的一种方法是评估继续训练的成本何时超过当前模型所犯错误的成本。试图量化医疗干预经济价值的现有衡量标准是质量调整生命年 (QALY)，其中一个 QALY 相当于一年的健康生活尼斯 (2013)。这个指标可以合并到模型中吗？目前，我们无法量化深度学习医学成像应用程序所造成的错误成本，但这样做可以让我们更深入地了解深度学习模型的准确度。

随着模型在更多最终用户自己的数据上进行训练，这是否会导致网络在来自该用户系统的数据上表现更好，尽管在模型最初训练的数据上表现较差？灾难性遗忘表明情况将会如此，但这是一件坏事吗？对于模型来说，逐渐偏向于最终用户自己的数据的高性能可能是有益的，即使这会导致模型变得不太可转移到其他数据。法夸尔等人。(2021)探索偏差在 AL 方法中的作用。引入偏差是因为训练数据不再遵循 AL 中的总体分布。作者提供了一种通用方法，通过该方法可以使用新颖的校正权重来构建无偏 AL 估计器以消除偏差。除此之外，还对忽略这种偏差的现有 AL 方法的经验成功进行了解释。结果表明，在训练过度参数化模型（例如数据相对较少的神经网络）时，AL 方法引入的偏差会非常有帮助。这进一步促使未来的工作更好地理解 AL 引入的偏差何时会对 AL 方法的性能产生积极影响，从而损害对其他数据源的普遍性。

主动学习假设存在执行注释的用户界面，但只关心要注释哪些数据。细化假设我们可以通过生成注释

与当前模型预测的迭代交互。因此，在未来的工作中将这两者结合起来是可取的。如果我们可以使用少量训练数据来训练模型，然后要求注释者改进模型预测以获得更准确的标签，那么我们可以通过减少初始注释工作量并减少与未见数据一起使用的额外接口工作来加快注释过程。这将与用于创建训练注释的界面相同。通过将主动学习和迭代细化的努力结合到一个统一的框架中，我们可以快速生成注释来训练我们的模型，并从一开始就从我们的模型中获得高质量的结果。这还应该具有额外的副作用，即使用与预测相同的分布的数据来训练模型，

通过在模型生命周期的每个阶段纳入我们的最终用户，我们还可以使用对模型性能的人类反馈来添加更“人类可解释”的模型置信度指标，因为每个用户都可以对每个输入的性能进行排名它可能会根据人类对模型输出的解释给出一个置信度量。这当然需要专家使用该系统。有人可能会争辩说，模型的初始预测可能会对人类用户产生一些影响，但通过将初始注释众包给不太专业的多标签人群，我们可以减少这种偏差。

不确定性量化的发展将有利于 AL 选择启发式和模型输出的解释，但不能保证用于选择要注释的新样本的最佳性能不确定性指标将与人类用户最容易解释的指标相同。

图 4 概述了本综述中讨论的每篇论文在人机循环计算中使用的核心方法。该图表明，人机循环计算的许多领域的研究目标存在显著重叠，但为了理解不同方法之间的关系以及这些方法可能如何影响其性能，还需要填补很大的空白。

随着深度学习研究的许多领域趋向于共同目标，即利用有限的训练数据来实现最先进的结果，我们期望看到更多的系统出现，利用此处描述的机器学习子领域范围内取得的进步。我们已经看到了将几种方法组合到单独的框架中，但到目前为止还没有任何作品将讨论的所有方法组合到一个框架中。随着不同的方法组合开始出现，重要的是要考虑我们评估其绩效的衡量标准，因为隔离个人发展变得更加困难。开发用于比较的基线人机交互方法对于评估每个领域中各个工作的贡献以及更好地了解这些领域的竞争性改进的影响至关重要。

6. 结论

在这篇综述中，我们探讨了以人类最终用户为中心的大量新兴医学图像分析工作。深度学习具备引发我们处理大量临床任务的方法范式转变的所有要素。人类的直接参与将在这一转变中发挥核心作用。本次评论中提出的每项工作都提供了自己的将人类纳入循环的方法，我们建议许多方法有足够的重叠，可以在人机循环计算的同一标题下考虑它们。我们希望看到新的方法论出现，将 AL 和 HITL 计算的优势结合到端到端系统中，以开发可用于临床实践的深度学习应用程序。虽然如上所述存在一些实际限制，但

针对此类问题提出了许多解决方案，并且随着这些方向的研究继续进行，深度学习应用发展成为用于日常任务的成熟、准确和强大的系统只是时间问题。我们正处于医学图像分析的一个激动人心的时代，拥有无限的机会来创新和改进当前的最先进技术，并利用深度学习的力量对医疗保健产生全面的真正影响。通过勤奋的研究和开发，我们应该会看到越来越多由深度学习功能推动的应用程序进入市场，使用户能够以比以前更少的专业知识、更快地获得更好的结果，从而腾出专家的时间来用于最重要的事情。具有挑战性的案例。人机循环计算领域将在实现这一目标方面发挥至关重要的作用。

竞争利益声明

作者声明，他们没有已知的可能影响本文报告工作的相互竞争的经济利益或个人关系。

CRediT 作者贡献声明

塞缪尔·巴德：概念化、形式分析、调查、写作 - 初稿、写作 - 审查和编辑、可视化。**艾玛·C·罗宾逊**：概念化、资源、写作 - 审查和编辑、监督。**伯恩哈德·凯恩斯**：概念化、资源、写作 - 审查和编辑、监督、项目管理、资金获取。

致谢

SB 由以下机构支持**电力系统研究委员会**智能医学影像博士培训中心**EP/S022104/1**。这项工作得到了 EP/S013687/1、Intel 和 Nvidia 的部分支持。我们感谢创新英国：伦敦医学影像和人工智能中心基于价值的医疗保健 [104691] 共同资助这项研究。

补充材料

与本文相关的补充材料可以在在线版本中找到，网址为[10.1016/j.media.2021.102062](https://www.sciencedirect.com/journal/medical-image-analysis/issue/S022104102062)

参考

- Amrehn, M., Gaube, S., Unberath, M., Schebesch, F., Horz, T., Strumia, M., Steidl, S., Kowarschik, M., Maier, A., 2017. UI-Net: 基于用户模型的迭代图像分割的交互式人工神经网络。见：Bruckner, S., Hennemuth, A., Kainz, B., Hotz, I., Merhof, D., Rieder, C. (编辑)，欧洲生物医学和医学视觉计算研讨会。欧洲图形协会 doi:[10.2312/vcbm.20171248](https://doi.org/10.2312/vcbm.20171248)。
- Amrehn, M., Steidl, S., Kortekaas, R., Strumia, M., Weingarten, M., Kowarschik, M., Maier, A., 2019. 交互式图像分割系统的半自动可用性评估框架。国际。J.生物医学。成像 2019.doi:[10.1155/2019/1464592](https://doi.org/10.1155/2019/1464592)。网址：<https://pubmed.ncbi.nlm.nih.gov/31582963/>。
- Angluin, D., 1988. 查询和概念学习。马赫。学习。2 (4), 319–342。土井：[10.1023/A:1022821128753](https://doi.org/10.1023/A:1022821128753)。
- Angluin, D., 2001. 重新审视查询。在：计算机科学讲义（包括-荷兰国际集团子系列人工智能讲义和生物信息学讲义）。施普林格出版社，第 12–31 页。土井：[10.1007/3-540-45583-3_3](https://doi.org/10.1007/3-540-45583-3_3)。Atlas, LE, Cohn, DA, Ladner, RE, 1990. 训练联结网络-适用于查询和选择性采样。见：Touretzky, DS (编辑)，神经信息处理系统的进展，2. Morgan-Kaufmann，页数 566–573。网址：[http://papers.nips.cc/paper/261-带有查询和选择性采样的训练连接网络.pdf](https://papers.nips.cc/paper/261-带有查询和选择性采样的训练连接网络.pdf)。Bachman, P., Sordani, A., Trischler, A., 2017. 主动学习算法。技术报告。网址：[http://会议记录.mlr.press/v70/bachman17a/bachman17a.pdf](https://会议记录.mlr.press/v70/bachman17a/bachman17a.pdf)。
- Baweja, C., Glocker, B., Kamnitsas, K., 2018. 走向医学持续学习校准成像。技术报告。网址：https://www.doc.ic.ac.uk/~bglocker/public/mednips2018/med-nips_2018_paper_82.pdf。

- Beede, E., Baylor, E., Hersch, F., Iurchenko, A., Wilcox, L., Ruamviboonsuk, P., Var-doulakis, LM, 2020. 对临床中部署的用于检测糖尿病视网膜病变的深度学习系统进行以人为本的评估。见: 2020 年 CHI 计算系统人为因素会议论文集。计算机协会 (ACM), 美国纽约州纽约市, 第 1-12 页。土井: [10.1145/3313831.3376718](https://doi.org/10.1145/3313831.3376718)。
- Beluch Bcaí, WH, Nürnberger, A., Bcaí, JMK, 2018. 集成对于 Ac- 的力量
图像分类中的主动学习。技术报告。网址: http://openaccess.thecvf.com/content_cvpr_2018/CameraReady/1487.pdf。
- Ben-David, S., Blitzer, J., Crammer, K., Kulesza, A., Pereira, F., Vaughan, JW, 2010. A 从不同领域学习的理论。马赫。学习。79 (1-2), 151-175。Bredell, G., Tanner, C., Konukoglu, E., 2018. 细分市场的迭代交互训练
站编辑网络。见: 计算机科学讲义 (包括人工智能讲义子系列和生物信息学讲义)。施普林格出版社, 第 363-370 页。土井: [10.1007/978-3-030-00919-9_42](https://doi.org/10.1007/978-3-030-00919-9_42)。巴德, S., 辛克莱, M., 卡纳尔, B., 马修, J., 劳埃德, D., 戈麦斯, A., 杜桑, N.,
Robinson, EC, Kainz, B., 2019. 通过超声波进行可靠的头围测量, 并为超声检查人员提供实时反馈。见: 计算机科学讲义 (包括人工智能讲义子系列和生物信息学讲义)。施普林格, 第 683-691 页。土井: [10.1007/978-3-030-32251-9_75](https://doi.org/10.1007/978-3-030-32251-9_75)。
- Chakraborty, S., Tomsett, R., Raghavendra, R., Harborne, D., Alzantot, M., Cerutti, F., Srivastava, M., Preece, A., Julier, S., Rao, RM, Kelley, TD, Braines, D., Sensoy, M., Willis, CJ, Gurram, P., 2017. 深度学习模型的可解释性: 结果调查。于: 2017 IEEE SmartWorld、普适智能与计算、先进与可信计算、可扩展计算与通信、云与大数据计算、人联网与智慧城市创新 (Smart-World/SCALCOM/UIIC/ATC/CBDCom/IOP/SCI)。IEEE, 第 1-6 页。土井: [10.1109/UIIC-ATC.2017.8397411](https://doi.org/10.1109/UIIC-ATC.2017.8397411)。
- Cheplygina, V., de Bruijne, M., Pluim, JP, 2019. 不受监督: 一项调查
医学图像分析中的半监督、多实例和迁移学习。医学。图像肛门。54, 280-296。土井: [10.1016/j.media.2019.03.009](https://doi.org/10.1016/j.media.2019.03.009)。Cheplygina, V., Perez-Rovira, A., Kuo, W., Tiddens, HA, de Bruijne, M., 2016 年。早期
胸部 CT 气道注释众包的经验。见: 计算机科学讲义 (包括人工智能讲义子系列和生物信息学讲义)。施普林格出版社, 第 209-218 页。土井: [10.1007/978-3-319-46976-8_22](https://doi.org/10.1007/978-3-319-46976-8_22)。
- Choudhary, A., Tong, L., Zhu, Y., Wang, MD, 2020. 推进医学影像信息化
通过基于深度学习的领域适应来实现医学。年。医学。通知。29 (1), 129-138。土井: [10.1055/s-0040-1702009](https://doi.org/10.1055/s-0040-1702009)。
- Cohn, D., Atlas, L., Ladner, R., 1994. 通过主动学习提高泛化能力。
马赫。学习。15 (2), 201-221。土井: [10.1007/bf00993277](https://doi.org/10.1007/bf00993277)。
- Criminisi, A., Sharp, T., Blake, A., 2008. GeoS: 测地线图像分割。
施普林格, 柏林, 海德堡, 第 99-112 页。土井: [10.1007/978-3-540-88682-2_9](https://doi.org/10.1007/978-3-540-88682-2_9)。Dagan, I., Engelson, SP, 1995. 基于委员会的概率训练抽样
分类器。在: 机器学习论文集。爱思唯尔, 第 150-157 页。土井: [10.1016/b978-1-55860-377-6.50027-x](https://doi.org/10.1016/b978-1-55860-377-6.50027-x)。
- Edwards, L., Veale, M., 2017. 奴役算法: 从权利到解释
并有权做出更好的决定? SSRN 电子。J. doi: [10.2139/ssrn.3052831](https://doi.org/10.2139/ssrn.3052831)。Edwards, L., Veale, M., 2017. 算法的奴隶? 为什么解释权是
可能不是您正在寻找的补救措施。SSRN 电子。J. doi: [10.2139/ssrn.2972855](https://doi.org/10.2139/ssrn.2972855)。
- Fang, M., Li, Y., Cohn, T., 2017. 学习如何主动学习: 深度强化
学习方法。见: 2017 年自然语言处理经验方法会议论文集。计算语言学协会, 丹麦哥本哈根, 第 595-605 页。土井: [10.18653/v1/D17-1063](https://doi.org/10.18653/v1/D17-1063)。
- Farquhar, S., Gal, Y., Rainforth, T., 2021. 关于主动学习中的统计偏差: 如何
以及何时修复它。见: 国际学习代表会议。Gal, Y., Ghahramani, Z., 2016. Dropout 作为
贝叶斯近似: 表示
深度学习中的模型不确定性。见: 第 33 届国际机器学习会议 (ICML-16) 会议记录。
- Gal, Y., Islam, R., Ghahramani, Z., 2017. 利用图像进行深度贝叶斯主动学习
数据。见: 第 34 届国际机器学习会议论文集
- 第 70 卷。JMLR.org, 第 1183-1192 页。
- Goodman, B., Flaxman, S., 2017. 欧盟关于算法决策的规定 -
制作和“解释权”。人工智能杂志。38 (3), 50-57。土井: [10.1609/aimag.v38i3.2741](https://doi.org/10.1609/aimag.v38i3.2741)。
- 关华, 刘明, 医学图像分析领域适应: 一项调查。arXiv:
2102.09508
- Haenssle, HA, Fink, C., Schneiderbauer, R., Toberer, F., Buhl, T., Blum, A., Kallou, A.,
与 58 名皮肤科医生相比, 深度学习卷积神经网络在皮肤黑色素瘤识别方面的诊断性能。
安。安科尔。29 (8), 1836-1842。土井: [10.1093/annonc/mdy166](https://doi.org/10.1093/annonc/mdy166)。
- Hesamian, MH, Jia, W., He, X., Kennedy, P., 2019. 医学领域的深度学习技术
图像分割: 成就与挑战。J. 数字。成像 32 (4), 582-596。土井: [10.1007/s10278-019-00227-x](https://doi.org/10.1007/s10278-019-00227-x)。
- Holzinger, A., Malle, B., Kieseberg, P., Roth, PM, Müller, H., Reihs, R., Zatloukal, K., 走向增强病理学家: 数字病理学中可解释人工智能的挑战。arXiv:1712.06657
- Jang, W.-D., Kim, C.-S., 2019. 通过反向传播进行交互式图像分割
细化方案。见: IEEE 计算机视觉和模式识别会议论文集。
- Jungo, A., Reyes, M., 2019. 评估可靠性和不确定性挑战
医学图像分割的系统。技术报告。
- 卡姆尼察斯, K., 鲍姆加特纳, C., 莱迪格, C., 纽科姆, V., 辛普森, J., 凯恩, A., Menon, D., Nori, A., Criminisi, A., Rueckert, D. 等人, 2017 年。对抗性网络病灶分割
中的无监督域适应。见: 医学影像信息处理国际会议。施普林格, 第 597-609 页。
- Khosravan, N., Celik, H., Turkbey, B., Cheng, R., McCreedy, E., McAuliffe, M., Bednarova, S., Jones, E., Chen, X., Choyke, P., Wood, B., Bagci, U., 2017. Gaze2Segment: 将眼动追踪技术集成到医学图像分割中的试点研究。查姆·施普林格, 第 94-104 页。土井: [10.1007/978-3-319-61188-4_9](https://doi.org/10.1007/978-3-319-61188-4_9)。
- Kim, JS, Greene, MJ, Zlateski, A., Lee, K., Richardson, M., Turaga, SC, Purcaro, M., Balkam, M., Robinson, A., Behabadi, BF, Campos, M., Denk, W., Seung, HS, Eye-Wirers, t., 2014. 时空布线特异性支持视网膜的方向选择性。自然 509 (7500), 331-336。
土井: [10.1038/自然13240](https://doi.org/10.1038/自然13240)。
- 金, RD, 罗兰, J., 奥克利, SG, 杨, M., 奥布里, W., 伯恩, E., 利亚卡塔, M., Markham, M., Pir, P., Soldatova, LN, Sparkes, A., Whelan, KE, Clare, A., 2009 年。科学的自动化。科学 324 (5923), 85-89。土井: [10.1126/科学.1165620](https://doi.org/10.1126/科学.1165620)。
- 金, RD, 惠兰, KE, 琼斯, FM, 赖瑟, PG, 科比, CH, 马格尔顿, SH, Kell, DB, Oliver, SG, 2004. 机器人科学家的功能基因组假设生成和实验。自然 427 (6971), 247-252。土井: [10.1038/自然02236](https://doi.org/10.1038/自然02236)。
- Kirsch, A., van Amersfoort, J., Gal, Y., 2019. Batchbald: 高效且双向
用于深度贝叶斯主动学习的诗歌批量获取。见: 神经信息处理系统的进展 32。Curran Associates, Inc., 第 7026-7037 页。网址: <http://papers.nips.cc/paper/8925-batchbald-efficient-and-diverse-batch-acquisition-for-deep-bayesian-active-learning.pdf>。Konyushkova, K., Snitman, R., Fua, P., 2017. 从数据中学习主动学习。
网址: <https://papers.nips.cc/paper/7010-learning-active-learning-from-data>。
- Konyushkova, K., Snitman, R., Fua, P., 2019. 二进制主动学习中的几何
和多类图像分割。计算。维。图像理解。182, 1-16。土井: [10.1016/J.CVIU.2019.01.007](https://doi.org/10.1016/J.CVIU.2019.01.007)。
- Kuo, W., Häne, C., Yuh, E., Mukherjee, P., Malik, J., 2018. 成本敏感型主动
学习颅内出血检测。查姆·施普林格, 第 715-723 页。土井: [10.1007/978-3-030-00931-1_82](https://doi.org/10.1007/978-3-030-00931-1_82)。
- Kurzenhofer, T., Fischer, P., Mirshahzadeh, N., Pohl, T., Brost, A., Steidl, S., Maier, A., 2017. 使用径向基函数插值对 3D 医学图像进行快速交互式直观分割。见: 医学图像理解
与分析年会, 第 11-13 页。土井: [10.3390/jimaging3040056](https://doi.org/10.3390/jimaging3040056)。库什巴尔, K., 巴尔韦德, S., 冈萨雷斯-维拉, J., 伯纳尔, J., 卡贝萨斯, M., 奥克利, A.,
Lladó, X., 2019. 以最少的用户交互进行自动皮下大脑结构分割的监督域适应。科学。
众议院 9 (1), 6742。doi:10.1038/s41598-019-43299-z。
- Lang, K., Baum, E., 1992. 当人类预言机出现时, 查询学习的效果可能会很差
用来。见: IEEE 国际神经网络联合会议论文集。IEEE, 第 335-340 页。
- Last, F., Klein, T., Ravanbakhsh, M., Nabi, M., Batmanghelich, K., Tresp, V., 2018. Hu-
医学图像分割的人机协作。技术报告。Lewis, DD, Catlett, J., 1994. 监督的异质不确定性
抽样
学习。见: 第十一届国际机器学习会议论文集。摩根·考夫曼, 第 148-156 页。
- Li, S., Jiang, Y., Chawla, NV, Zhou, Z., 2019. 群体多标签学习。IEEE
跨。知道。数据工程 31 (7), 1369-1382。土井: [10.1109/TKDE.2018.2857766](https://doi.org/10.1109/TKDE.2018.2857766)。Lian, C., Liu, M., Zhang, J., Shen, D., 2020. 分层全卷积网络
使用结构 MRI 进行关节萎缩定位和阿尔茨海默病诊断。IEEE 传输。模式肛门。马赫。英特尔。
尔。42 (4), 880-893。土井: [10.1109/TPAMI.2018.2889096](https://doi.org/10.1109/TPAMI.2018.2889096)。
- 廖X., 李W.-H., 徐Q., 王X., 金B., 张X., 张Y., 王Y., 2020.
通过多智能体强化学习迭代细化交互式 3D 医学图像分割。见: 2020 年 IEEE/CVF 计算机视觉
和模式识别会议 (CVPR), 第 9391-9399 页。
- Litjens, G., Kooi, T., Bejnordi, BE, Setio, AAA, Ciampi, F., Ghafoorian, M., van der Laak, JA, van Ginneken, B., Sánchez, CI, 2017 年。医学图像分析中深度学习的调查。
医学。图像肛门。42, 60-88。土井: [10.1016/J.media.2017.07.005](https://doi.org/10.1016/J.media.2017.07.005)。
- Liu, M., Zhang, J., Adeli, E., Shen, D., 2018. 基于Landmark的深度多实例
学习脑部疾病诊断。医学。图像肛门。43, 157-168。土井: [10.1016/j.media.2017.10.005](https://doi.org/10.1016/j.media.2017.10.005)。
- Lowell, D., Lipton, ZC, Wallace, BC, 2019 年。部署主动部署的实际障碍
学习。见: 2019 年自然语言处理经验方法会议和第九届自然语言处理国际联合会会议
(EMNLP-IJCNLP) 论文集。计算语言学协会, 中国香港, 第 21-30 页。土井: [10.18653/v1/D19-1003](https://doi.org/10.18653/v1/D19-1003)。
- Lundervold, AS, Lundervold, A., 2019. 医学深度学习概述
成像以 MRI 为主。Z. 医学。物理。29 (2), 102-127。土井: [10.1016/J.ZEMEDI.2018.11.002](https://doi.org/10.1016/J.ZEMEDI.2018.11.002)。
- Hauptmann, A., Lin, W.-H., Yan, R., Yang, J., Chen, M.-y., 2006. 极端视频重新
trieval: 人类和计算机性能的联合最大化。第 385-394 页。10.1145/1180639.1180721

- Mahapatra, D., Bozorgtabar, B., Thiran, J.-P., Reyes, M., 2018. 高效主动学习使用样本选择和条件生成对抗网络进行图像分类和分割。查姆·施普林格, 第 580–588 页。土井: [10.1007/978-3-030-00934-2_65](https://doi.org/10.1007/978-3-030-00934-2_65)。
- Mar, VJ, Soyer, HP, 2018 年。用于黑色素瘤诊断的人工智能: 如何我们兑现承诺了吗? 安·安科尔。29 (8), 1625–1628。土井: [10.1093/annonc/ndy193](https://doi.org/10.1093/annonc/ndy193)。
- McCallum, A., Nigam, K., 1998. 采用 em 和基于池的主动学习文本分类。见: 第十五届国际机器学习会议论文集。Morgan Kaufmann Publishers Inc., 美国加利福尼亚州旧金山, 第 350–358 页。土井: [10.5555/645527.757765](https://doi.org/10.5555/645527.757765)。
- Millitari, F., Birodkar, V., Sofka, M., 2019. 开门见山: 强化学习-为超声用户提供指导。见: 计算机科学讲义 (包括人工智能讲义子系列和生物信息学讲义)。施普林格, 第 3–10 页。土井: [10.1007/978-3-030-32875-7_1](https://doi.org/10.1007/978-3-030-32875-7_1)。Moeskops, P., Wolterink, JM, van der Velden, BHM, Gilhuijs, KGA, Leiner, T., Viergever, MA, Išgum, I., 2016. 多模态多任务医学图像分割的深度学习。查姆·施普林格, 第 478–486 页。土井: [10.1007/978-3-319-46723-8_55](https://doi.org/10.1007/978-3-319-46723-8_55)。
- Morid, MA, Borjali, A., Del Fiol, G., 2021. 对反式的范围界定审查使用 ImageNet 进行医学图像分析的学习研究。10.1016/j.combiomed.2020.104115
- Nalisnik, M., Gutman, DA, Kong, J., Cooper, LA, 2015. 交互式学习用于可扩展的病理图像分类框架。在: 会议记录: ... IEEE 国际大数据会议。IEEE 国际大数据会议, 2015 年, 第 928–935 页。土井: [10.1109/大数据.2015.7363841](https://doi.org/10.1109/大数据.2015.7363841)。
- NICE, 2013. 判断公共卫生干预措施是否物有所值| 指导和指南| 好的。网址: <https://www.nice.org.uk/advice/lgb10>。Oktay, O., Schlemper, J., Folgoc, L., Lee, M., Heinrich, M., Misawa, K., Mori, K., Mc-Donagh, S., Hammerla, N., Kainz, B., Glocker, B., Rueckert, D., 注意力 u-net: 学习在哪里寻找胰腺。arXiv:1804.03999
- Ozdemir, F., Peng, Z., Tanner, C., Fuernstahl, P., Goksel, O., 2018. 主动学习-通过优化内容信息以获得最大熵来进行分割。查姆·施普林格, 第 183–191 页。土井: [10.1007/978-3-030-00889-5_21](https://doi.org/10.1007/978-3-030-00889-5_21)。
- Pan, Y., Liu, M., Lian, C., Xia, Y., Shen, D., 2020. 空间受限的渔民代表对不完整的多模态神经图像进行脑部疾病识别的不满。IEEE 传输。医学。成像 39 (9), 2965–2975。土井: [10.1109/TMI.2020.2983085](https://doi.org/10.1109/TMI.2020.2983085)。
- Parisi, GI, Kemker, R., Part, JL, Kanan, C., Wermter, S., 2019. 终身持续使用神经网络学习: 评论。10.1016/j.neunet.2019.01.012 Raghu, M., Zhang, C., Kleinberg, J., Bengio, S., 2019. 输出: 理解医学成像的迁移学习。神经信息处理系统。
- Rajchl, M., Koch, LM, Ledig, C., Passerat-Palmbach, J., Misawa, K., Mori, K., Rueckert, D., 采用弱注释来解决医学图像分析问题。arXiv: 1708: 06297v1
- Rajchl, M., Lee, MCH, Oktay, O., Kamnitsas, K., Passerat-Palmbach, J., Bai, W., Damodaram, M., Rutherford, MA, Hajnal, JV, Kainz, B., Rueckert, D., 2016. Deepcut: 使用卷积神经网络从边界框注释中进行对象分割。IEEE 传输。医学。成像 36, 674–683。
- Rajchl, M., Lee, MCH, Schrans, F., Davidson, A., Passerat-Palmbach, J., Tarroni, G., Alansary, A., Oktay, O., Kainz, B., Rueckert, D., 分布式弱监督下的学习。arXiv:1606.01100v1
- Rodrigues, F., Pereira, FC, 2018. 来自人群的深度学习。在: 三十秒-第二届 AAAI 人工智能会议 (AAAI), 2018 年。AAAI 出版社, 第 1611–1618 页。
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: 卷积网络生物医学图像分割。查姆·施普林格, 第 234–241 页。土井: [10.1007/978-3-319-24574-4_28](https://doi.org/10.1007/978-3-319-24574-4_28)。
- Schlemper, J., Oktay, O., Schaap, M., Heinrich, M., Kainz, B., Glocker, B., Rueckert, D., 2019. 注意力门控网络: 学习利用医学图像中的显著区域。医学。图像肛门。53, 197–207。土井: [10.1016/j.media.2019.01.012](https://doi.org/10.1016/j.media.2019.01.012)。Settles, B., 2009 年。主动学习文献调查。技术报告。大学威斯康星-麦迪逊计算机科学系。
- Settles, B., Craven, M., 2008. 序列主动学习策略分析标记任务。见: 自然语言处理经验方法会议论文集。美国计算语言学会, 第 1070–1079 页。土井: [10.5555/1613715.1613855](https://doi.org/10.5555/1613715.1613855)。
- Shah, MP, Bhalgat, YS, Awate, SP, 2018. 医疗注释成本最小化使用建议混合监督全卷积网络进行图像分割。技术报告。
- 香农, CE, 。关于通信的数学理论。技术报告。Shen, D., Wu, G., Suk, H.-I., 2017. 医学图像分析中的深度学习。Annu。生物医学教师。工程师。19, 221–248。土井: [10.1146/annurev-bioeng-071516-044442](https://doi.org/10.1146/annurev-bioeng-071516-044442)。
- Smailagic, A., Noh, HY, 科斯塔, P., Walawalkar, D., Khandelwal, K., Mirshekari, M., Fagert, J., Galdran, A., Xu, S., 2018. MedAL: 医学图像分析的深度学习主动学习采样方法。未定义, URL: <https://www.semanticscholar.org/paper/MedAL%3A-Smailagic-Noh的深度学习主动学习采样方法/fa23dc7a8b3927953d83f5ce46e0b622b7cac456>。
- Sourati, J., Gholipour, A., Dy, JG, Kurugol, S., Warfield, SK, 2018. 主动深度学习-结合渔民信息进行补丁式语义分割。在: 医学图像分析中的深度学习和用于临床决策支持的多模态学习: 第四届国际研讨会, DLMIA 2018, 和第八届国际研讨会, ML-CDS 2018, 与 MICCAI 2018 联合举行, 格拉纳达, 西班牙, S..., 11045, 第 83–91 页。土井: [10.1007/978-3-030-00889-5_10](https://doi.org/10.1007/978-3-030-00889-5_10)。
- Stember, JN, Celik, H., Krupinski, E., Chang, PD, Mutasa, S., Wood, BJ, Lignelli, A., Moonis, G., Schwartz, LH, Jambawalikar, S., Bagci, U., 2019. 使用卷积神经网络进行深度学习分割的眼动追踪。J.数字。成像 32 (4), 597–604。土井: [10.1007/s10278-019-00220-4](https://doi.org/10.1007/s10278-019-00220-4)。
- , 2018. 理解和解释医学图像计算应用中的机器学习。见: Stoyanov, D., Taylor, Z., Kia, SM, Oguz, I., Reyes, M., Martel, A., Maier-Hein, L., Marquand, AF, Duchesnay, E., Löfstedt, T., Landman, B., Cardoso, MJ, Silva, CA, Pereira, S., Meier, R. (编辑)。计算机科学讲义, 11038. Springer International Publishing, Cham doi: [10.1007/978-3-030-02628-8](https://doi.org/10.1007/978-3-030-02628-8)。
- Suzuki, K., 2017. 医学影像深度学习概述。放射线。物理。科技-不。10 (3), 257–273。土井: [10.1007/s12194-017-0406-5](https://doi.org/10.1007/s12194-017-0406-5)。
- Tajbakhsh, N., Jeyaseelan, L., Li, Q., Chiang, JN, Wu, Z., Ding, X., 2020. 拥抱不完美的数据集: 医学图像分割深度学习解决方案的回顾。医学。图像肛门。63, 101693.doi:[10.1016/j.media.2020.101693](https://doi.org/10.1016/j.media.2020.101693)。Tajbakhsh, N., Shin, JY, Gurudu, SR, Hurst, RT, Kendall, CB, Gotway, MB, Liang, J., 2016. 用于医学图像分析的卷积神经网络: 全面训练还是微调? IEEE 传输。医学。成像 35 (5), 1299–1312。土井: [10.1109/TMI.2016.2535302](https://doi.org/10.1109/TMI.2016.2535302)。
- 英国皇家放射科医学院, 2017 年。英国临床放射学劳动力普查 2017 年报告。技术报告。
- Tinati, R., Luczak-Roesch, M., Simperl, E., Hall, W., 2017. 对玩家的调查 Eyewire (一个游戏化的公民科学项目) 的动机。计算。呼。行为。73, 527–540。土井: [10.1016/j.chb.2016.12.074](https://doi.org/10.1016/j.chb.2016.12.074)。
- Tizhoosh, HR, Pantanowitz, L., 2018. 人工智能和数字病理学: 挑战与机遇。J. Pathol. 通知。9, 38. doi: [10.4103/jpi.jpi_53_18](https://doi.org/10.4103/jpi.jpi_53_18)。Wang, D., Li, M., Ben-Shlomo, N., Corrales, CE, Cheng, Y., Zhang, T., Jayender, J., 2019. 用于医学图像分割的混合监督双网络。见: 计算机科学讲义 (包括人工智能讲义子系列和生物信息学讲义)。施普林格, 第 192–200 页。土井: [10.1007/978-3-030-32245-8_22](https://doi.org/10.1007/978-3-030-32245-8_22)。
- Wang, G., Li, W., Aertsen, M., Deprest, J., Ourselin, S., Vercauteren, T., 2019. 任意使用卷积神经网络进行医学图像分割的测试时间增强的不确定性估计。神经计算 338, 34–45。土井: [10.1016/J.NEUCOM.2019.01.103](https://doi.org/10.1016/J.NEUCOM.2019.01.103)。
- Wang, G., Li, W., Zuluaga, MA, Pratt, R., Patel, PA, Aertsen, M., Doel, T., David, AL, Deprest, J., Ourselin, S., Vercauteren, T., 2018. 使用深度学习和图像特定微调的交互式医学图像分割。IEEE 传输。医学。成像 37 (7), 1562–1573。
- Wang, G., Zuluaga, MA, Li, W., Pratt, R., Patel, PA, Aertsen, M., Doel, T., David, AL, Deprest, J., Ourselin, S., Vercauteren, T., 2019. Deepigeos: 用于医学图像分割的深度交互式测线框架。IEEE 传输。模式肛门。马赫。英特尔。41 (7), 1559–1572。
- Wang, H., Rivenson, Y., Jin, Y., Wei, Z., Gao, R., Günaydin, H., Bentolila, LA, Kural, C., Ozcan, A., 2019. 深度学习使荧光显微镜中的跨模态超分辨率成为可能。纳特。方法 16 (1), 103–110。土井: [10.1038/s41592-018-0239-0](https://doi.org/10.1038/s41592-018-0239-0)。
- 王凯、张东、李勇、张瑞、林丽, 2017. 经济有效的主动学习用于深度图像分类。IEEE 传输。电路系统。视频技术。27 (12), 2591–2600。土井: [10.1109/TCSVT.2016.2589879](https://doi.org/10.1109/TCSVT.2016.2589879)。
- 王Y., 于B., 王L., 祖C., Lalush DS, 林W., 吴X., 周J., 沉D., Zhou, L., 2018. 用于低剂量高质量 PET 图像估计的 3D 条件生成对抗网络。神经影像 174, 550–562。土井: [10.1016/j.neuroimage.2018.03.045](https://doi.org/10.1016/j.neuroimage.2018.03.045)。
- Wen, S., Kurc, TM, Hou, L., Saltz, JH, Gupta, RR, Batiste, R., 赵, T., Nguyen, V., Samaras, D., Zhu, W., 2018. 比较具有主动学习的不同分类器以支持病理图像核分割的质量控制。2017 年, 227–236。网址: <http://www.ncbi.nlm.nih.gov/pubmed/29888078> Woodward, M., Finn, C., Research, BA, 2017. 主动一次性学习。技术的报告。
- 夏Y., 刘F., 杨D., 蔡J., 于L., 朱Z., 徐D., Yuille, A., 罗斯H., 2020. 具有不确定性感知多视图协同训练的 3D 半监督学习。见: 2020 年 IEEE 计算机视觉应用冬季会议 (WACV), 第 3635–3644 页。
- Yamashita, R., Nishio, M., Do, RKG, Togashi, K., 2018. 卷积神经网络著作: 放射学概述和应用。洞察成像 9 (4), 611–629.doi:[10.1007/s13244-018-0639-9](https://doi.org/10.1007/s13244-018-0639-9)。
- 杨L., 张Y., 陈J., 张S., 陈DZ, 2017. 暗示性注释: A 用于生物医学图像分割的深度主动学习框架。查姆·斯普林格, 第 399–407 页。土井: [10.1007/978-3-319-66179-7_46](https://doi.org/10.1007/978-3-319-66179-7_46)。
- Yi, X., Walia, E., Babyn, P., 2019. 医学成像中的生成对抗网络: 回顾。医学。图像肛门。58, 101552.doi:[10.1016/j.media.2019.101552](https://doi.org/10.1016/j.media.2019.101552)。Yu, B., Zhou, L., Wang, L., Shi, Y., Fripp, J., Bourgeat, P., 2019. Ea-GAN: 边缘感知用于跨模态 MR 图像合成的生成对抗网络。IEEE 传输。医学。成像 38 (7), 1750–1762。土井: [10.1109/TMI.2019.2895894](https://doi.org/10.1109/TMI.2019.2895894)。张, C., Chaudhuri, K., 2015. 向弱标签者和强标签者主动学习。在: 第 28 届神经信息处理系统国际会议论文集-第 1 卷。麻省理工学院出版社, 美国马萨诸塞州剑桥, 第 703–711 页。土井: [10.5555/2969239.2969318](https://doi.org/10.5555/2969239.2969318)。
- 张, C., 陈, T., 2002. 基于内容的信息的主动学习框架化检索。IEEE 传输。多媒体。4, 260–268。
- 赵, A., 巴拉克里希南, G., 杜兰德, F., 古塔格, JV, 达尔卡, AV, 2019. 数据增强-使用学习到变换进行一次性医学图像分割。见: IEEE 计算机视觉和模式识别会议论文集, 第 8543–8553 页。

郑 S.、Jayasumana, S.、Romera-Paredes, B.、Vineet, V.、苏 Z.、杜 D.、黄 C.、Torr, PHS, 2015。条件随机场作为循环神经网络。技术报告。

Zhou, Z.、Shin, J.、Feng, R.、Hurst, RT、Kendall, CB、Liang, J., 2018。积分主动颈动脉内膜中层厚度视频解读的学习和迁移学习。J.数字。成像 doi:[10.1007/s10278-018-0143-2](https://doi.org/10.1007/s10278-018-0143-2)。

Zhou, Z.、Shin, J.、Feng, R.、Hurst, RT、Kendall, CB、Liang, J., 2018。积分主动颈动脉内膜中层厚度视频解读的学习和迁移学习。J.数字。成像 doi:[10.1007/s10278-018-0143-2](https://doi.org/10.1007/s10278-018-0143-2)。

Zhou, Z.、Shin, J.、Zhang, L.、Gurudu, S.、Gotway, M.、Liang, J., 2017。微调卷积用于生物医学图像分析的视觉神经网络：主动和增量。在：诉讼程序。IEEE 计算机学会计算机视觉和模式识别会议, 2017 年, p. 第4761章: [10.1109/CVPR.2017.506](https://doi.org/10.1109/CVPR.2017.506)。