

使用机器学习对阿尔兹海默症进行诊断

方桂安, 刘梦莎, 钟意, 张焯霖, 黄杰辉[†], 唐振超[†], 陈语谦^{*}

摘要—阿尔兹海默病 (AD) 是一种急性神经性疾病, 会导致脑细胞退化, 从而逐渐导致记忆损失。这是一种致命的脑疾病, 主要影响老年人。它导致认知和生物学功能的下降, 进而导致脑萎缩。要对阿尔兹海默病进行准确的诊断, 现代化的方法如机器学习是必不可少的。近年来, 机器学习在医疗行业中引起了很大的关注和流行。随着疾病的进展, 患有阿尔兹海默病的人在进行甚至最基本的任务时会遇到更多困难, 最严重的情况下, 他们的大脑完全停止工作。采用机器学习方法可以确定一个人是否有早期阿尔兹海默病。在我们完成科创项目综述论文的过程中, 通过在 Google Scholar 上查找 2008 年至 2023 年间关于基于深度学习技术和强化学习的阿尔兹海默病诊断的文献, 选取了 60 篇相关论文进行分析。这些论文基于 AD 的生物标志物和使用的机器学习技术进行了分析。分析结果显示, 深度学习方法具有很强的特征提取和 AD 分类的能力, 并且具有较高的准确性。在图像处理领域, 强化学习方法 (DRL) 的应用相对较少。深度学习和强化学习的比较结果表明, 深度强化学习在痴呆症检测中的应用前景有待进一步探索。

关键词—阿尔兹海默症, CNN, RNN, DRL

1. 概述

AD, 即阿尔兹海默症, 是在众多类型的痴呆症中最常见的一种, 可能对记忆造成严重损害。AD 由局部脑萎缩所伴随。它会损害神经元的关键生物学功能, 如通讯、代谢、修复、重塑和再生。根据最新研究, 预计 2021 年将有约 620 万 65 岁及以上的人被诊断为 AD。AD 影响的认知能力包括思维、推理和记忆。AD 的前驱阶段是轻度认知障碍 (MCI), 导致累积性记忆丧失和行为问题。根据报告显示, 约 15-20% 的老年人 (65 岁或以上) 患有 MCI, 其中 30-40% 的人在 5-6 年内发展为 AD。在大多数情况下, 这种转化的时间周期为 18 个月, 但可能在 6 到 36 个月之间有所变化。65 岁以上的女性比男性更容易患 AD, 比例高出 6%。只有在尸检中通过组织病理学检查 淀粉样斑块和神经原纤维缠结才能得出 AD 的最终诊断。这些斑块是由收集

淀粉蛋白在脑神经元之间形成, 从而干扰神经元功能。另一方面, 神经原纤维缠结是一种异常积累的 Tau 蛋白。显然, 需要除尸检外的其他标准来确认 AD。2011 年, 美国国家老龄化研究所和阿尔兹海默病协会将生物标志物分为两类: 基于 PET 成像的生物标志物和基于脑脊液以及 MRI 的生物标志物。AD 的诊断标准每 3 到 4 年更新一次, 以便吸纳新的知识以追踪 AD。迄今为止, 尚未发现 AD 的完全治愈方法。因此, 早期检测这种疾病变得至关重要。阿尔兹海默病患者的平均寿命为 3 至 9 年。AD 的早期诊断可以帮助减缓患者症状加速恶化的速度, 从而减缓最严重的痴呆症状。关键步骤是检测 MCI 转化为 AD, 而不仅仅是对 AD 患者进行分类。目前可用的计算机辅助机制不能取代该领域的医学专家, 但可以作为一种辅助工具来加强他们的临床诊断。一些关于使用机器学习技术进行 AD 诊断的评论研究已经发表。这些研究基于不同类型的生物标志物、特征选择方法、单模态和多模态方法。

研究人员对阿尔兹海默病 (AD) 的诊断兴趣逐年增加, 相较于其他神经疾病, 如图 1 所示。虽然已经开发了多种用于 AD 诊断的机器学习模型, 但本项目进行的文献综述表明, 仍需要探讨专门的机器学习方法, 结合深度学习和强化学习的优势, 用于早期 AD 诊断。

A. 体征和症状

1) 早期阶段: 阿尔兹海默病是一种随着时间推移逐渐恶化的神经疾病, 分为三个阶段。这三个阶段中, 尤其是早期阶段, 很难被察觉, 常常与老年化或压力混淆。根据医生使用的测试类型, 阿尔兹海默病治疗的问题可能需要长达八年才能显现出来。在初期阶段, 患者的日常活动变得越来越具有挑战性。最常见的症状是短期记忆丧失, 这使得人们难以记住最近学到的东西, 也让学习新事物变得困难。患有阿尔兹海默病的人可能在早期阶段出现缺乏能量和普遍的悲伤感。疾病最常见的症状是冷漠或缺乏关心或兴趣。轻度认知障碍是一种可能会影响一个人日常任务、学习新信息和理解环境的神经疾

方桂安, 20354027, (e-mail: fanggan@mail2.sysu.edu.cn)。

刘梦莎, 20354091, (e-mail: liumsh6@mail2.sysu.edu.cn)。

钟意, 20320229, (e-mail: zhongy276@mail2.sysu.edu.cn)。

张焯霖, 20354156, (e-mail: zhangzhlin8@mail2.sysu.edu.cn)。

本项目由陈语谦^{*}老师和黄杰辉[†]、唐振超[†]师兄进行指导。

Percentage of the number of Articles indexed in Google Scholar (2010-2021)

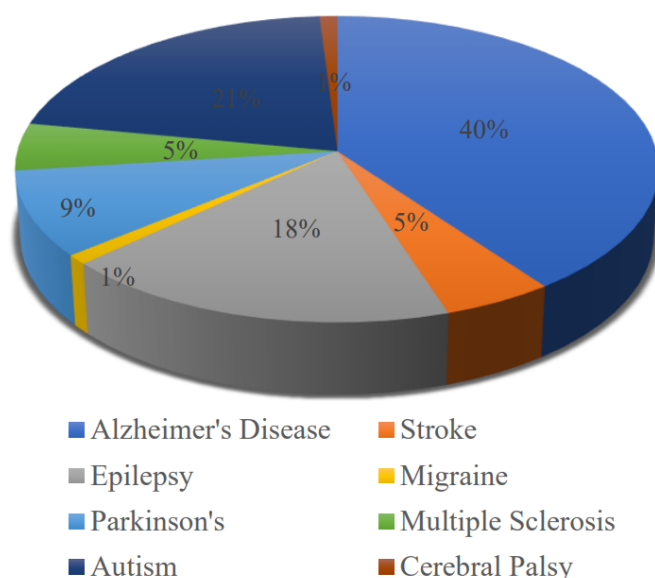


Fig. 1. 在 Google Scholar 上发布的不同神经疾病的文章 (2010-2021 年)

病。记忆障碍（轻度认知障碍）表现为短期记忆下降，通常在阿尔茨海默病的早期阶段出现。研究表明，患有失忆性轻度认知障碍的人有 90% 的机会患上阿尔茨海默病。语言问题的主要表现是在处理较难的词汇时出现困难。阿尔茨海默病患者在阅读、写作、穿衣、安排或计划事物等活动中都会遇到困难，但他们无法意识到这些问题。在依赖他人的帮助下，患有阿尔茨海默病的人仍然可以完成日常任务。

2) 中期阶段: 阿尔茨海默病患者在无法记住单词时，会在交流时遇到困难，导致他们使用词语不准确。还有些患者失去了辨认亲近亲属的能力。随着时间的推移，个体辨认亲近亲属的能力逐渐恶化。患者完成一件事的方式会发生变化，使得他们更难以集中精神进行某项任务。30% 的阿尔茨海默病患者倾向于认为其他人、地方和事物的身份发生了变化。渐渐地，患者开始难以确定自己是谁，这让照顾患者的压力越来越大。

3) 晚期阶段: 到了这个阶段，阿尔茨海默病患者几乎完全依赖于照顾者。虽然患者无法正常说话，但他仍然可以通过信号和标志进行交流。在这个阶段，患者完全忘记了如何独自完成日常事务。患者将无法自己进食，还可能通过自残造成伤口。最终，阿尔茨海默病会导致器官功能衰竭和其他身体伤害。

B. AD 诊断

AD 的诊断通常使用生物标志物测试、神经生理学测试和基因测试。由于难以追踪随时间变化的行为变化，阿尔茨海默病的研究需要耗费大量人力资源。然而，临床设置通常不使用生物标志物测试。Mini-Mental State Examination (MMSE) 是一种用于获取患者心理状态有用的临床结果的测试。该测试评估了个体认知能力和记忆力的多个方面，包括这些能力随时间的变化情况、记忆力如何以及短期记忆能力如何。医生会询问一系列旨在评估患者常见心理能力的问题。Mini-Mental State Examination (MMSE) 的最高可能得分为 30 分。患者的年龄和 MMSE 得分都会被考虑在内。医学领域广泛使用机器学习进行不同的诊断。“决策树”被用于选择患者最佳的治疗方案。决策树是对在权衡可用选项时可能采取的各种路径的图形表示。可以使用监督式机器学习模型来找到患者疾病的根本原因。监督式学习模型被用于做出决策并对给定情况进行深入分析。疾病检测最有效的方法以及它与其他患者相比的优劣也是研究的课题。医学领域的研究对于了解阿尔茨海默病并使世界变得更美好发挥了积极作用。在医学诊断技术中应用了人工智能技术和长期记忆监测技术。

II. AD 诊断中的生物标志物

不同的生物标志物已经被用于诊断阿尔茨海默病，如图 2 所示。通过对淀粉样蛋白 (A) 进行检测，可以揭示阿尔茨海默病的早期阶段。阿尔茨海默病病因的主要生物标志物是长期以来淀粉样前体蛋白基因。科学家们已经发现，位于第 14 号染色体上的载脂蛋白 E 基因的第四个变体与某些家族和人群中阿尔茨海默病的高发性有关。曾经认为高血压会导致阿尔茨海默病。研究表明，神经生长因子前体蛋白 (proNGF) 的受损影响了阿尔茨海默病诊断领域的进展。以认知正常的被试者的语言表现作为生物标志物，可以预测阿尔茨海默病的发病。一种基于脑脊液（直接接触人脑的液体）作为生物标志物，结合近红外 (NIR) 拉曼光谱和机器学习分析的新方法已经被开发用于诊断阿尔茨海默病。血液中基因的表达谱被作为生物标志物。近年来，神经影像学生物标志物如磁共振成像 (MRI) 和正电子发射断层显像 (PET) 在 AD 的诊断中变得不可或缺。通过磁共振成像 (MRI)，更容易了解与 AD 相关的功能和解剖学上的脑部变化。许多研究人员开发了多种使用人工智能算法和包含 MRI 图像的数据集的技术，描述了脑部结

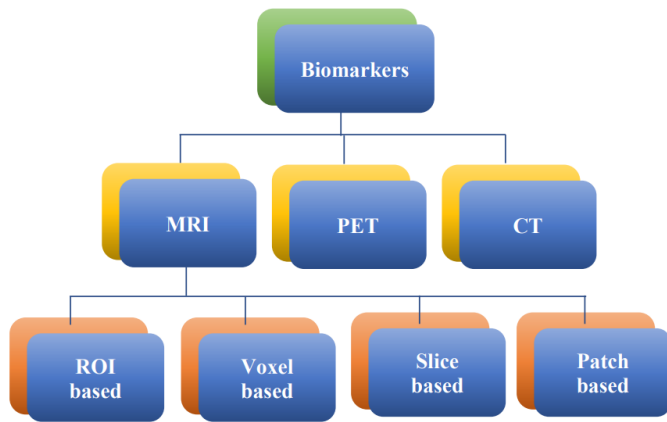


Fig. 2. 阿尔茨海默病诊断中的不同生物标志物和特征提取方法

构变化，如皮层和皮下萎缩、总脑容积的变化、颞内皮质和海马萎缩，因为这些是神经纤维缠结的目标位置。其中一些方法以整个脑部 MRI 图像作为输入，而其他方法则使用基于兴趣区（ROI）的图像、基于体素的图像、基于切片的图像和基于块的图像。

III. 输入数据

A. 基于 ROI 的特征提取

在阿尔茨海默病的早期阶段，海马、颞叶内侧叶、杏仁核和颞旁回等大脑区域的功能和结构开始丧失。在研究中，使用了 45 个派生的 ROI 和 2 个组合 ROI 来训练和测试模型。结果表明，海马、颞旁回和颞叶内侧叶在诊断中获得了最佳效果。然而，由于采用了 2D 区域特征提取方法，难以从小的 ROI 中获取足够的信息。使用 3D 特征提取方法可以解决这个问题，ROI 分析同时涵盖了白质和灰质区域。

B. 基于体素的特征提取

在基于体素 (Voxel) 的特征提取方法中，通过 MRI 图像中的白质或灰质来获取体素的强度值。灰质的提取，也称为组织分割，是在 MRI 图像上进行的。在相关研究中，使用了 649 个基于体素形态学 (Voxel-based Morphometry, VBM) 图像，提出了一种新的基于体素的特征检测框架。这种方法减轻了基于 ROI 的方法的局限性，但高维数据的影响仍然较大。在相关研究中，使用了极限学习模型 (Extreme Learning Model, ELM)，该模型使用了 VBM 的特征。该研究使用了简单的特征选择方法来提取 VBM 特征。通过使用其他方法和大型数据集，这种方法可以进一步改进。首先，将

大脑根据其解剖结构划分为不同的区域，然后在这些区域之间推导出复杂的非线性关系。在这种方法中，体素的强度值表示了每个区域的体积测量值。

C. 基于切片的特征提取

在基于切片 (Slice-Based) 的特征提取方法中，从三维体积脑数据中提取二维特征，以生成所需的三维特征。为了减少特征的数量，超参数的数量被削减。一些使用二维图像切片的研究考虑了神经影像的标准投影，例如矢状面（中矢面）、冠状面（前矢面）或轴向面（水平面）。在相关研究中，每隔 3.0 毫米在冠状方向采样的切片被用于确保特征的多样性。为了提高分类准确性，这种方法可以采用更好的方式来选择要使用的切片。在相关研究中，每个三维 MRI 图像被拆分成二维切片，并在卷积神经网络 (CNN) 之后使用循环神经网络 (RNN)。考虑图像序列之间的关系，从而可以根据所有切片而不仅仅是一个切片来做出决策。研究得出结论，通过将二维图像切片与其他特征（如记忆测试）结合起来，可以提高准确性。在相关研究中，基于结构性 MRI (MRI) 灰质切片和注意力机制的深度学习框架被实现，从而使 AD 的诊断更加准确。

D. 基于分块的特征提取

在基于图像分块 (Patching) 的特征提取方法中，通过在图像上选择任意形状的区域，可以捕捉与疾病相关的模式。整个脑图像被分割成小的区域，从每个区域中提取多个 3D 图像块，并使用 K-Means 聚类方法将其聚集成簇。然后，使用 Densenet 学习每个簇的图像块特征，并将其组合进行分类。虽然在这种方法中很难直观地展示用于分类的学习到的特征，但可以通过结合多模态的脑图像分析方法来进一步改进。在相关研究中，针对局部图像块的不同部分构建了多个个体分类器，然后将它们组合在一起以获得更准确的分类结果。这种方法可以扩展到其他多种模态的图像数据，以进一步提高分类的准确性。

IV. 研究方法

本文中讨论了在阿尔茨海默病 (AD) 诊断中常用的深度学习模型，包括深度神经网络 (DNN)、卷积神经网络 (CNN)、循环神经网络 (RNN) 和受限玻尔兹曼机 (RBM)。如图 3 所示，越来越多的研究已经在使用深度学习技术进行 AD 诊断方面展开。

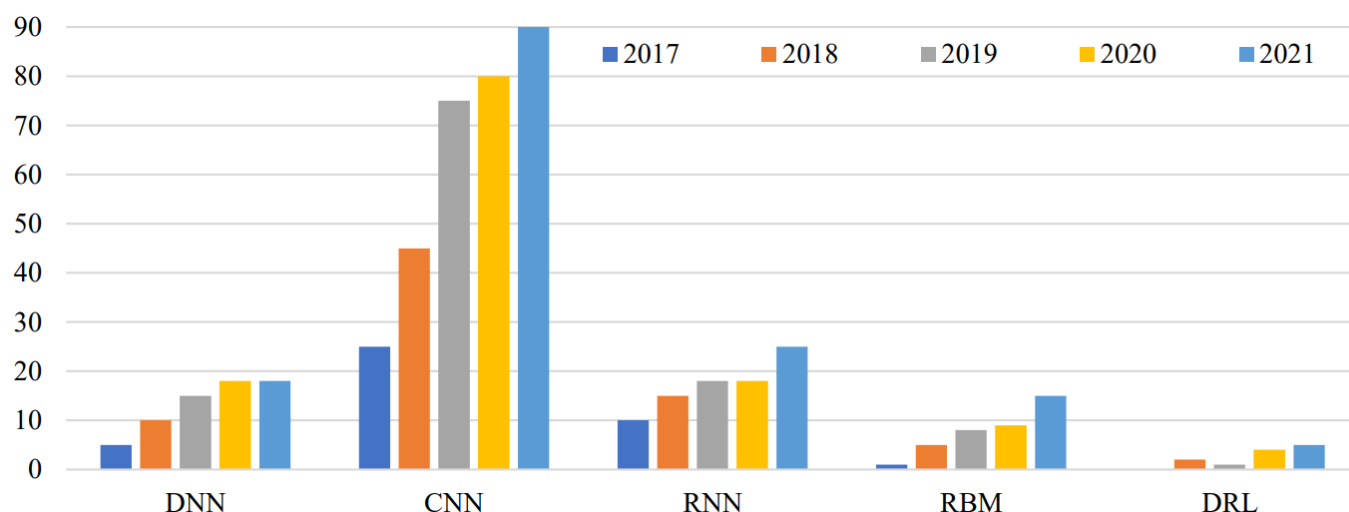


Fig. 3. 过去 5 年中发表的研究结果表明了深度学习技术在阿尔茨海默病 (AD) 诊断中的应用

A. 深度神经网络 (DNN)

DNN 是一种具有选择性和适应性的神经网络，其配置与多层感知器 (MLP) 相同，但具有多于两层的隐藏层。DNN 通常包含一个输入层、一个输出层和多个隐藏层。它们主要用于分类和回归，因为 DNN 处理未标记和非结构化的数据。DNN 利用神经网络的前馈特性来处理数据。在论文中，一种常用的基于 DNN 的模型 LeNet 使用了最大池化和最小池化层的串联。在相关研究中，使用钐材料的 DNN 模型比机器学习算法获得了更高的准确性。在另一研究中，使用从血样中获得的 AD 相关基因的 DNN 模型显示出大于 0.80 的 AUC (曲线下面积) 值。在论文中，作者通过在每层中增加更多的隐藏单元并调整学习率超参数来扩展 DNN 模型，以提高学习算法的性能。研究表明，用于阿尔茨海默病分类的最佳 DNN 模型使用 11 层的神经元，每层的神经元数量逐层减少。

B. 卷积神经网络 (CNN)

卷积神经网络 (CNNs) 属于深度神经网络家族，受到了大脑视觉皮层层次结构的启发。相比于人工神经网络 (ANNs)，CNNs 更具优势，因为它们可以同时进行特征提取和分类。在过去的几年中，CNN 在使用医学影像诊断疾病方面，尤其是使用图像进行诊断方面，变得非常流行。卷积神经网络的主要部分包括输入层、特征提取层和分类层。特征提取层由卷积层和池化层的迭代组成。卷积层包含了滤波器和激活映射/层。全连接层用于计算类别分数。研究旨在检查 CNN 模型在磁共振成像 (MRI) 中对阿尔茨海默病诊断和护理的分割效

果。在论文中，三个遵循密集连接模式的深度 CNNs 集成获得了比其他基线深度 CNNs 更高的准确性。在相关研究中，提出了一种基于 CNN 的多模态统一框架，用于联合分类神经影像和遗传数据，其中 CNN 用于从 MRI、PET 和 SNP 图像中学习特征。一些研究使用了 2D CNNs，而另一些研究则使用了 3D CNNs 进行神经影像处理。这些研究使用了具有三个卷积层的 2D CNNs，用于单模态方法来寻找 AD。在另一研究中，通过将 CNN、RNN 和 LSTM 模型结合在一起，以获得更加准确的模型。

C. 递归神经网络 (RNN)

自然语言处理和语音识别中使用的神经网络是循环神经网络 (RNN)。在 RNN 中，输入、输出和隐藏层之间存在关联。每一步的输入是上一步的输出，并且同样的权重在所有层之间共享。RNN 中的隐藏层可以学习序列信息，从而有助于处理顺序性的信息。

在相关研究中，建议使用 RNN 模型基于个体的临床诊断、认知状态和脑室体积来填补未来每个月的丢失数据。卷积神经网络 (CNNs) 和循环神经网络 (RNNs) 被提出作为检测和分类阿尔茨海默病的模型的一部分。通过卷积神经网络 (CNN) 学习 MR 图像的空间特性。CNN 的输出作为输入传递给由三个级联的双向门控循环单元 (BGRU) 组成的循环神经网络 (RNN)，用于提取纵向特征。研究人员建议使用结构性 MRI 数据的混合卷积和循环神经网络，以深入研究海马。在这项研究中，RNN 被级联使用，通过结合左右海马的信息来学习更高级的 AD 分类特征。

在另一研究中,通过改进的多对一 LSTM RNN 模型,包括一个输入层、一个输出层和两个隐藏层,验证了 RNN 可以有效处理患者过去预约中获得的内在时间和医学模式,从而处理 AD 的进展问题。使用数据驱动的技术和 RNN 对 AD 进行分类,开发了一种新的特征序列表示方法。

D. 受限玻尔兹曼机 (RBM)

限制玻尔兹曼机 (RBM) 是一种随机生成神经网络,具有明显和隐藏两层。在这种设计中,输入层神经元与隐藏层神经元相互通信,但它们不会与同一层中的其他神经元相互作用。研究表明,RBM 可以学习内部表示,因此已被应用于阿尔茨海默病的诊断。

在特征提取方面,一组 RBM (也称为深度置信网络, DBN) 和支持向量机 (SVM) 用于分类。这些 RBM 在这项工作中经历了无监督的预训练。在 AD 检测方面,研究人员转向 DBN,这是一种用于获取训练数据深层次分层表示的图形模型。

E. 深度强化学习 (DRL)

深度强化学习 (DRL) 结合了深度学习 (DL) 和强化学习 (RL) 的特点。强化学习是机器学习 (ML) 的一个分支,其中代理 (agent) 被用于学习环境,并通过从环境中获得的奖励来学习。

在相关研究中,所提出的方法与一些现有的强化学习模型相比,产生了更准确的结果。这个基于强化学习和神经网络的系统可以生成不平衡的数据集段类别。

V. 研究方向

在过去的 20 年里,出现了许多诊断阿尔茨海默病的方法。他们从不同角度切入,根据阿尔茨海默病对患者的影响做出诊断。

A. 认知测试和临床评估

MMSE(Mini-Mental State) 是一种经过简化的,以评分为形式的认知心理状态的测试,该测试仅包含 11 个问题,能够在 5-10 分钟内完成,并且对可证明可靠性和有效性的认知状态的进行量化评估,用以判断患者是否患有 AD。MoCA(The Montreal Cognitive Assessment) 是一种新的认知筛查工具,这项工具具有高灵敏度和特异性,对认知障碍的轻度阶段非常有用(包括 MCI 和轻度 AD),能够快速评估不同程度的认

知障碍。对上述两种方法以 CERAD-NB 为标准进行对比,认为 MoCA 作为一种全局评估工具优于 MMSE,尤其是在识别认知能力下降的早期阶段。并且提供了 MoCA 到 MMSE 评分的可靠且简单的转换。尽管两种方法都对临床实践有极高的价值,但都不能单独承担诊断的责任,仅能作为诊断的辅助。

与其他神经退行性疾病一样,阿尔茨海默症的诱因是大脑中的蛋白质错误折叠,而生物标志物的研究可以表明 AD 的病理学可以在临床前检测。结合临床方法和认知测试,生物标志物更有助于完成准确的评估认知障碍及其原因,甚至能够让临床医生在认知障碍发生前进行识别并检测到由 AD 引起的病理。生物标志物,是对生物或致病过程的客观测量,目的是评估疾病风险或预后,指导临床诊断或监测治疗干预措施。发展至今,经过证实有用的生物标志物变化包括磁共振成像 (MRI) 上的海马萎缩、18F-氟脱氧葡萄糖正电子发射断层扫描 (FDG-PET) 上的皮层低代谢、脑脊液 (CSF) 中 $A\beta_{42}$ 的降低以及 tau 蛋白和磷酸化 tau 蛋白的增加。同样地,目前也没有任何一种生物标志物能够单独准确诊断 AD

B. 声音

AD 带来的认知障碍常常与声音方面也有联系,语言障碍被认为是 AD 最典型的症状之一,它是认知障碍所带来的直接的、不可避免的结果。有研究证明,AD 患者在不同语言测试中表现不佳,他们更频繁地重复相同的想法,使用更长的空单词,使用更简单的语言结构,并且在一些声学特征、韵律特征上表现得与常人不同,这都表明了使用声音模态来表征阿尔茨海默氏症痴呆症有可行性。

语义言语流利度测试 (SVF)、自发言语测试 (SS) 是两种大众化且表现良好的语言任务,这两种任务均使用了与单词检索、语义和声学损伤以及错误率相关的语言特征,能够取得期望的结果。其中,针对 SS 的研究在区分 AD 和健康状况方面取得了最高的准确性。然而许多使用上述测试的研究往往都集中在患者所述内容,从而忽略了语音本身发生的变化。这些人类无法自然感知的参数,如韵律、语音强度等都是专家常用的诊断工具,并且韵律是神经退行性疾病患者最常研究的参数之一。ESA(Emotional Speech Analysis) 的研究综合了声学特征、语音质量特征和持续时间特征,在 AZTHIAO 数据集上区分 AD 患者和健康受试者达到

了 90.7% 的正确率。

近年来, 基于声音领域诊断 AD 发展迅猛, *LinLiu* 等利用从语音数据中提取的谱图特征来识别 AD, 并且供了一个关于老年人语音数据 (*VBSD*) 的新数据集。*FasihHaider* 等首次将 *eGeMAPS*、*emobase* 和 *ComParE* 特征集用作阿尔茨海默病的“数字生物标志物”的实证尝试, 介绍和评估一种表示这些声学特征的新方法 (*ADR*), *SaturninoLuz* 等描述了 *ADReSS* 挑战, 并为其任务设置了简单的基线, 证明了在阿尔茨海默病诊断和 *MMSE* 预测的背景下, 自发语音的声学 and 语言特征与认知障碍检测的相关性。

作为一种非侵入式、快速的诊断方法, 相比于难以获取的医学图像, 基于患者语音数据的识别技术能够有效降低医疗成本, 特别是近年来自然语言处理技术、信号处理技术、深度学习技术得到了显著发展, 基于语音信号记录自动处理的技术逐渐趋于成熟。

C. 步态

步行被认为是一项需要对涉及到注意力、计划、记忆和其他运动、知觉、认知过程有更高程度控制的任务。并且步态已被证明与认知有着密切的关系, 而 AD 患者的认知障碍时时常也表现在步态上, 特别是当认知资源受到压力时, AD 患者的步态会进一步恶化。

在该方向研究中, 不少研究者关注使用双重任务范式研究步态功能, 即执行认知任务的同时保持行走, 这是为了能够模拟日常生活中那些对于执行力和注意力都比较高的场景。发现了在轻度 AD 患者中, 同时进行认知任务似乎对步态功能有明显影响, 进一步发现了 AD 患者的步态受到认知双重任务的干扰比受到运动双重任务的影响更大。因此, 双任务步态, 即在执行需要注意力的任务时步行, 在有认知障碍的个体中比在没有认知缺陷的个体中受到的影响更大, 这都表现出了步态在诊断早期 AD 的潜力。

而利用以下八种步态特征的均值、标准差 (SD) 和不对称性: 步态时间、步时间、单支撑时间、摆动时间、双支撑时间、站立时间、步幅和步长, 以及这些特征衍生的总共 108 种步态特征。开发了一种不需要认知测试、仅基于步态评估的模型。该模型摆脱了由认知任务带来的种种不确定性, 如不同的教育水平和文化背景, 作为认知障碍的筛查可以促使临床医生对 MCI 和 AD 的诊断进行进一步评估。

同时, AD 似乎伴随着运动活动特征的变化, 加速度测量已被证明是一种可行的评估方法。结合生物特征

和 IMU 动力学变量, 建立了一个能够为 ANN 配置参数的 AD 诊断工具。开发了一种能够从加速度计捕获患者活动数据中识别 AD 阶段的方法, 且这种方法能够应用在手机等方便的加速传感器中, 对于促进 AD 进展的监测非常有价值, 而不需要专门的设备或专业的医疗人士监督获取。

VI. 结果和讨论

不同的机器学习方法已经被用于 AD 分类, 并且在综述论文中对其中一些方法进行了比较。在图 4 中, 我们可以看到每种机器学习技术与其他技术相比的表现情况。在分析用于 AD 分类的各种深度学习技术的比较时, 发现在使用 DNN 技术时, ADNI 数据集上使用 Lenet 的准确率为 96.64%, 而使用 OASIS 数据集的 20 层隐藏层的变种 DNN 的准确率为 91.00%, ADNI 数据集上的前馈 DNN 的准确率为 79.3%。在关注第二种技术 CNN 时, 使用 MRI 和 OASIS 数据集的变种深度 CNN 的准确率为 93%。其次, 使用包含多模态 MRI、PET 和 SNP 的 ADNI 数据集的变种监督 2-D CNN 的准确率为 98.22%。通过检查使用 ADNI 的 MRI 数据集的 CNN-RNN-LSTM 的集成变种, 得到了 92.22% 的准确率。使用 ADNI 的 MRI 数据集的 3 级双向门控循环单元 (BGRU) 的 RNN 的准确率为 91.33%, 而使用国家阿尔茨海默病协调中心的多模态 CDR、GDS、FAQ 数据集的 LSTM RNN 的准确率为 99%。与 DNN 和 RNN 相比, CNN 在曲线下面积 (AUC) 方面取得了更高的准确率。深度强化学习在 OASIS 数据集上显示出比强化学习更高的准确性。此外, 多模态方法比单模态方法显示出更好的结果。

VII. 研究总结

阿尔茨海默病是一种影响人的心智和行为的可怕疾病。因此, 早期诊断对于这种疾病非常重要。近年来, 研究人员尝试了多种深度学习算法和方法来检测 AD, 其中大多数研究使用 CNN 将数据分类。在 90% 的研究中, ADNI 是主要的数据集。ADNI 是一个多中心纵向研究, 旨在通过临床、成像、遗传和生化指标的早期诊断来寻找 AD 的迹象。使用先进的深度学习算法, 这些数据集可以通过识别最佳组合的多样性生物标志物, 改善和提前预测 AD。阿尔茨海默病 (AD) 是目前最常见的痴呆症类型之一。根据 2018 年的《世界阿尔茨海默报告》, 这种疾病在 2018 年影响了约 5000 万人, 预计到 2050 年将增加三倍。轻度认知障碍 (MCI) 是阿尔

Fig. 4. 用于 AD 分类的各种 ML 技术的比较

Technique	Variants	Modality	Dataset Source	AUC	Accuracy
DNN	DNN with two hiddenlayers	AD-related genes fromblood samples	ADNI, ANM1, ANM2	0.80	
	LeNet	MRI	ADNI		96.64%
	DNN with 20 hidden layers	MRI	OASIS		91.00%
	Feed-Forward DNN	MRI	ADNI		79.3%
CNN	Improved CNN	MRI	IGMC hospital		
	Deep CNN	MRI	OASIS		93%
	Supervised 2D CNN	MRI, PET, SNP	ADNI		98.22%
	CNN	fMRI	Huashan Hospital of Fudan University		95.59%
	Deep 3D CNN	MRI, PET	ADNI	0.92	
	Ensemble of CNN-RNN-LSTM	MRI	ADNI		92.22%
RNN	Minimal RNN	MRI, PET, CSF measures	ADNI		
	RNN with 3 cascaded BGRU	MRI	ADNI		91.33%
	Hybrid RNN with CNNmodel	MRI	ADNI	0.91	
	LSTM RNN	CDR, GDS, FAQ	National Alzheimer'sCoordinating Centre		99%
	Data-Driven RNN	Speech and Neuropsychological test		0.83	
RBM	Stacked RBM	MRI, PET	ADNI		91.4%
	DBN	MRI	OASIS		91.76%
	Multimodal DBM	MRI, PET	ADNI		95.35%
DRL	RL (Reinforcement Learning)	MRI	OASIS		82.23%
	DRL (Deep Reinforcement Learning)	MRI	OASIS		84.34%
	DRL-XGBOOST	MRI	OASIS		90.23%

茨海默病进展的初始阶段。然而，并不是每个患有 MCI 的人都会发展成阿尔茨海默病。因此，了解从 MCI 到 AD 的转变是当前研究的主要焦点。这些转变可以通过医学成像或其他技术如血浆光谱学来测量。多个开源阿尔茨海默病研究数据集的可用性加速了该领域的研究。ADNI (adni.loni.usc.edu)、AIBL (aibl.csiro.au) 和 OASIS 只是最广泛使用的数据库之一 (oasis.csiro.au)。首次向公众开放的 J-ADNI 数据库包含了在日本进行的纵向研究的数据。

此外，MRI 图像处理相当耗时。Wellcome 人类神经影像学中心发布了免费工具，如统计参数映射 (SPM)，以便于 MRI 扫描的分析。在使用 MRI 数据进行基于体素的形态测量学 (VBM) 时，通常使用 SPM。许多科学家还使用广泛应用的开源应用程序 Freesurfer 进行基于体积的形态测量学。在过去的十年里，机器学习技术已被发现对诊断阿尔茨海默病非常有用。常见的分类方法包括支持向量机 (SVM)、人工神经网络 (ANN) 和深度学习 (DL)。SVM 和 ANN 之间设置优化问题的

方式是主要的区别。SVM 和 ANN 都有一个重要的步骤叫做“特征提取”。而深度学习则将提取特征的步骤融入到学习模型本身。深度学习对于大数据集，特别是图像数据，非常有帮助。一些研究人员还使用称为“集成学习”的方法来提高阿尔茨海默病分类的准确性。深度学习结合强化学习 (RL) 可以用于各种 MRI 处理技术。目前很少有研究采用深度强化学习进行 AD 预测。与 RL 相比，深度强化学习表现出更好的预测结果。因此，基于深度强化学习的阿尔茨海默病诊断仍有更多的研究空间。

VIII. 阶段性成果

A. 研究进展

目前小组对已经收集的文献进行了详细的阅读并着手综述的写作，小组以多模态角度切入阿尔兹海默症的诊断，对过去十数年来使用多种模态的文献进行总结。从认知测试、临床评估、声音、步态、图像等领域，描述了人们诊断阿尔兹海默症的进展，并了解了多个模型下所使用的方式方法，以表格，文字等多种方式呈现，以各个模态为基，阐述了该模态在该领域下的突出工作和研究进展，并探究多种模态相互结合所带来的积极作用，为 AD 的早期诊断提供更多，更好，更廉价，更方便的解决策略。小组对认知测试、临床评估、声音、步态部分已经完成了初稿，正在积极地研读其他方向的文献，争取尽快完成整体的初稿，并进行润色修改。

B. 问题与解决

小组在研究中也遇到了一些问题，并积极采取措施进行解决。首先，在文献收集和综述写作过程中，由于阿尔兹海默症诊断领域涉及多个模态，包括认知测试、临床评估、声音、步态和图像等，涉及的文献较为广泛且繁杂，需要小组成员耗费大量时间和精力进行详细阅读和整理。为了克服这一问题，小组采取了分工合作的方式，每个成员负责一个或多个模态的文献收集和总结，然后进行交流和讨论，确保综述文章的全面性和准确性。

其次，小组在研究中还面临着多模态数据的融合和分析的挑战。不同模态的数据具有不同的特点和形式，如何将它们进行有效融合，提取有用的信息并进行综合分析，对小组来说是一项复杂的工作。为了解决这一问题，我们查阅大量参考文献，对不同模态的数据进行处理和集成，从而得出更全面和深入的研究结论。

此外，小组还缺乏如何将多模态诊断方法转化为实际临床应用的问题。虽然多模态诊断方法在实验室研究中显示出了潜在的优势，但其在临床实践中的应用和推广仍存在一定的难度和限制。我们希望提高这篇综述论文的整理，探索了将多模态诊断方法引入临床实践的途径，帮助相关专业人员进行深入交流和讨论，以解决实际应用中的问题。

C. 未来展望

尽管在研究中遇到了一些问题，但小组对于阿尔兹海默症诊断多模态方法的研究仍然充满信心，并对未来的研究方向有了明确的展望。首先，小组将继续深入研究不同模态在阿尔兹海默症诊断中的优势和应用，包括认知测试、临床评估、声音、步态和图像等，探索更加精确和可靠的多模态诊断策略，并进一步探究多模态相互结合所带来的协同效应，以提高阿尔兹海默症早期诊断的准确性和效率。

此外，小组还计划进一步深入研究多模态诊断方法的生物学机制和病理生理学基础，探讨不同模态在阿尔兹海默症发生发展中的作用机制，从而更好地理解阿尔兹海默症的病理生理学过程，并为早期诊断和干预提供更深入的科学依据。

最后，小组还将关注多模态诊断方法在临床实践中的经济效益和社会效益，探讨其在成本效益、资源利用和社会影响等方面的可行性和可持续性。通过全面评估多模态诊断方法的综合效益，为其在临床实践中的推广和应用提供更加全面和可靠的支持。

综上所述，小组将继续深入研究阿尔兹海默症多模态诊断方法，在解决问题和优化方法的基础上，不断推动其在临床实践中的应用，并为阿尔兹海默症早期诊断提供更多、更好、更廉价、更方便的解决策略，为阿尔兹海默症的防治工作做出积极的贡献。

D. 分工

在综述论文写作中，由张焯霖负责引言、相关工作的写作，刘梦莎负责每个模态的数据简介与特征提取介绍，钟意负责评价指标、实验结果分析，对比表格与流程图的绘制，方桂安负责传统机器学习和多模态深度学习的研究方法介绍。

在中期报告中，由张焯霖负责摘要概述与生物标志物的写作，刘梦莎负责输入数据与研究方法的写作，钟意负责研究方向与结果讨论的写作，方桂安负责研究总结与阶段性成果的写作。