

# 基于深度学习医学多视角的阿尔茨海默病辅助筛查系统研究与设计

**【摘要】**阿尔茨海默病（Alzheimer's Disease, AD）是一种不可逆的神经退行性疾病，以渐进性认知功能衰退为主要病征。在深度学习辅助诊断 AD 研究领域，传统的卷积神经网络和简单的特征拼接方法无法有效利用多模态数据的互补信息，简单的特征拼接方法容易导致模态融合过程中关键信息丢失。近年来，深度学习技术的发展为攻克多模态特征如何进行有效融合这一难题带来了新可能。本文提出一种新型的深度学习算法框架以辅助医疗专业人员进行 AD 诊断。通过融合大脑的氟脱氧葡萄糖正电子发射断层成像(PET)、核磁共振成像(MRI)、遗传基因数据和临床数据医学多视角准确检测 AD、轻度认知障碍（Mild Cognitive Impairment, MCI）、认知正常（Cognitively Normal, CN）的存在。算法的创新点在于使用了非对称跨模态交叉注意力机制，这种机制可以有效捕捉不同数据模态特征之间相互作用的关键信息。本文将非对称跨模态交叉注意力机制与传统的单模态、多模态深度学习模型诊断 AD 的算法框架进行了对比，对非对称跨模态交叉注意力机制的重要性进行了评估。模型在测试集上实现了 94.88% 的准确率；同时设计了基于智能交互式的蒙特利尔认知评估（MoCA）量表的 AD 辅助筛查系统，可以实现简化认知障碍筛查的流程和提升医疗专业人员筛查的效率。

**【关键词】**阿尔茨海默病；非对称跨模态交叉注意力机制；深度学习；蒙特利尔认知评估量表；辅助筛查系统

# **Research and Design of Alzheimer's Disease Auxiliary Screening System Based on Multi-perspectives of Deep Learning in Medicine**

**[Abstract]** Alzheimer's Disease (AD) is an irreversible neurodegenerative disease, which is mainly characterized by progressive cognitive decline. In the field of deep learning aided diagnosis of AD research, traditional convolutional neural networks and simple feature stitching methods cannot effectively use the complementary information of multi-modal data, and simple feature stitching methods are easy to lead to the loss of key information in the process of modal fusion. In recent years, the development of deep learning technology has brought new possibilities to solve the problem of how to effectively fuse multimodal features. This thesis proposes a novel deep learning algorithm framework to assist medical professionals in AD diagnosis. By fusing fluorodeoxyglucose positron emission tomography (PET), magnetic resonance imaging (MRI), genetic data and clinical data of the brain, we can accurately detect AD and Mild (Cognitive Impairment from multiple perspectives. MCI) and Cognitively Normal (CN). The innovation of the algorithm is the use of asymmetric cross-modal cross-attention mechanism, which can effectively capture the key information of the interaction between the features of different data modalities. In this study, the asymmetric cross-modal cross-attention mechanism was compared with the traditional algorithm models of single-modal and multi-modal models for diagnosing AD, and the importance of the asymmetric cross-modal cross-attention mechanism was evaluated. The model achieves an accuracy of 94.88% on the test set. At the same time, an intelligent interactive Montreal Cognitive Assessment (MoCA) scale based AD assistant screening system was designed, which could simplify the process of cognitive disorders screening and improve the efficiency of screening for medical professionals.

**[Keywords]** Alzheimer's Disease Asymmetric Cross-Modal Cross-Attention Mechanism Deep Learning Montreal Cognitive Assessment Scale Auxiliary Screening System

# 目 录

<b>1 前言 .....</b>	<b>1</b>
<b>1.1 选题背景 .....</b>	<b>1</b>
<b>1.2 国内外研究现状 .....</b>	<b>1</b>
<b>1.3 论文研究内容与创新点.....</b>	<b>3</b>
<b>1.4 论文组织结构.....</b>	<b>4</b>
<b>1.5 小结.....</b>	<b>4</b>
<b>2 数据与方法.....</b>	<b>5</b>
<b>2.1 数据集与预处理 .....</b>	<b>5</b>
<b>2.1.1 标签预处理.....</b>	<b>6</b>
<b>2.1.2 临床数据预处理.....</b>	<b>6</b>
<b>2.1.3 影像数据预处理.....</b>	<b>7</b>
<b>2.1.4 基因数据预处理.....</b>	<b>8</b>
<b>2.1.5 确定训练数据集.....</b>	<b>8</b>
<b>2.2 ACMCA 模型架构设计 .....</b>	<b>9</b>
<b>2.2.1 浅层特征提取模块.....</b>	<b>10</b>
<b>2.2.2 多模态特征融合模块.....</b>	<b>11</b>
<b>2.2.3 深层特征提取模块.....</b>	<b>13</b>
<b>2.2.4 输出分类模块.....</b>	<b>14</b>
<b>3 实验与结果 .....</b>	<b>15</b>
<b>3.1 参数实验和选择 .....</b>	<b>15</b>
<b>3.2 评价指标 .....</b>	<b>17</b>
<b>3.3 多模态有效性实验 .....</b>	<b>18</b>

3.4 性能对比实验 .....	20
3.5 消融实验 .....	22
<b>4 系统设计与实现 .....</b>	<b>25</b>
4.1 研究背景 .....	25
4.2 系统解决痛点问题 .....	26
4.3 系统技术方案 .....	27
4.3.1 微信小程序端.....	27
4.3.2 Web 应用程序端 .....	29
4.3.3 ACMCA 模型中 CAM 技术的嵌入 .....	30
4.3.4 ACMCA-CAM 模型的训练.....	31
4.3.5 ACMCA-CAM 模型的应用 .....	32
<b>5 总结.....</b>	<b>34</b>
5.1 临床意义 .....	34
5.2 未来展望 .....	35
<b>参考文献.....</b>	<b>37</b>
<b>致谢.....</b>	<b>40</b>

# 1 前言

## 1.1 选题背景

阿尔茨海默症（Alzheimer's Disease, AD）是一种神经退行性疾病，全球已有超过 4000 万人患有该疾病<sup>[1]</sup>。目前针对 AD 的治疗方法以对症治疗为主，通过治疗减缓疾病的进展。轻度认知障碍（MCI）通常被认为是衰老认知变化和早期痴呆之间的边界。在这一阶段，对 AD 的有效治疗将使患者的认知功能保持在尽可能高的水平<sup>[2]</sup>。对 AD 以及 MCI 有效的诊断也有助于确认患者的病情进展情况，方便医疗专业人员对不同阶段的患者进行定制化治疗。

## 1.2 国内外研究现状

对 AD 研究的需要促成了阿尔茨海默症神经影像计划（Alzheimer's Disease Neuroimaging Initiative, ADNI）的诞生。深度学习在利用该数据进行临床决策支持方面展现了巨大的潜力。早期研究者通常使用单种生物标志物数据作为深度学习模型的输入。随着多模态深度学习技术的进步，越来越多的研究表明，在 AD 诊断中，利用多种模态数据进行的诊断通常比单一模态数据诊断的准确率更高<sup>[3]</sup>。从临床治疗反馈中也发现了由于 AD 的病因复杂，由 AD 引起的痴呆症状也容易与其他疾病引起的痴呆症状相混淆。因此，医生在诊断 AD 时往往需要综合患者的医学影像资料、神经系统评估数据以及实验室检验结果等多模态数据进行判断，其中 PET 影像、MRI 影像、临床数据等被常用作评估 AD 的临床指标。近年来，随着 ADNI 研究数据库的不断补充发展，更多支持临床决策的数据被纳入深度学习算法诊断 AD 的研究领域中<sup>[4]</sup>。在临床诊断 AD 实际应用场景中，基因数据亦是诊断 AD 的关键性指标之一<sup>[5]</sup>。

在计算机辅助 AD 诊断的研究领域，早期的研究者大多使用单模态数据进行 AD 分类诊断任务<sup>[6-8]</sup>。Zhang J 等人<sup>[9]</sup>将 DenseNet 中的连接方式改造为加权求和，使用 MRI 影像进行 AD、CN、MCI 三分类诊断任务。Basheera S 等人<sup>[10]</sup>将经过头骨剥离预处理的 MRI 成像输入到卷积神经网络，同样用于三分类诊断。尽管这些研究推动了深度学习技术在 AD 研究领域的发展，但使用单一模态数据对深度学习诊断 AD 模型进行训练，会造成模型精确度局限性。准确诊断 AD 往往需要综合分析患者的多方面信息。

在早期的多模态诊断 AD 研究任务中，多模态数据的特征融合通常是将不同模态数据提取的特征进行简单拼接，然后直接应用于后续的分类诊断任务。如 Swamy K R 等人<sup>[11]</sup>使用了两种分别基于 Attention 和 Inception 架构的子深度学习网络对患者的 MRI 影像进行处理，该研究中将这两种处理方式得到的图像特征与临床数据进行简单拼接合并，作为模型分类层的输入。然而，这种模型算法直接将拼接合并后的特征输入至分类模型中的网络设计，忽略了这些特征的深层次信息，没有对特征进行更深层次的提取，并且，这种简单的融合方式忽略了不同模态之间的重要程度，从而，可能导致多模态数据特征融合后，跨模态数据之间的关键性关联信息丢失，限制了模型的性能提升。近年来，在深度学习技术诊断 AD 研究领域中，如何有效融合多模态数据才能充分发挥不同模态数据之间的相互作用是研究热点。

因此，部分研究者尝试利用注意力机制改善多模态数据特征融合方法，以提升模型诊断 AD 的精确度。J Zhang 等人<sup>[12]</sup>先将 MRI 影像特征与 PET 影像特征进行拼接，随后使用交叉注意力机制将该拼接后的图像特征与脑脊液（CSF）生物标志物特征进行融合。H Chen 等人<sup>[13]</sup>将 MRI 与 PET 影像特征合并为一个成像特征，再将成像特征与医学检查和评估数据的编码特征输入到通道注意力模块并结合空间注意力机制以获取模态间互补信息。然而，传统两两模态间

的交叉注意力机制存在局限性，其仅针对于单一模态间的局部交互，难以全面捕捉成像数据（如 MRI、PET 影像）与非成像数据（如临床数据、基因检测指标）间的关联信息，使用注意力机制进行跨模态数据特征的融合方面仍有改进的空间。

### 1.3 论文研究内容与创新点

近年来，有研究者尝试使用频域转换器代替多头自注意力机制以减少模型的参数量。B Patro 等人<sup>[14]</sup> 将 GFNet 与 Transformer 进行堆叠，提出了 SpectFormer，其证明了同时使用频域转换器与传统的多头注意力机制能获得更好的效果。本文受 SpectFormer 的启发，创新性的提出基于非对称跨模态交叉注意力机制的阿尔茨海默症多分类诊断框架（Asymmetric Cross - Modal Crossover Attention - based Multiclass Diagnostic Framework for Alzheimer's Disease，简写 ACMCA），旨在改进传统深度学习模型算法在 AD 诊断研究领域的多模态数据特征的有效融合问题，该算法模型结合了频域转换器和多头注意力机制，以挖掘深层次多模态特征内部的关系，增强多模态特征联合表达。在本文的 ACMCA 模型中设计了一种 Fnet 与 Transformer 并行的网络结构。实验部分已证明其有效性。

为了更好地研究 AD 在深度学习领域的分类任务，在本文的工作中将诊断分为 AD、MCI、CN 三种类型。而以往早期绝大多数利用深度学习技术对 AD 诊断分类的工作中仅对样本进行二分类（AD 和 CN）诊断任务。本文工作尝试对样本进行三分类（AD、CN 和 MCI）的诊断任务。与 CN 测试者相比，MCI 测试者表现出更多的异质性特征，并且更有可能显示出 AD 病理的负面指标，这使得三分类比二分类更具挑战性，对模型细粒度特征的学习能力也提出了更高要求。同时，在利用多模态信息诊断 AD 的研究中，传统的研究者仅使用

PET 影像、MRI 影像和临床数据这三种医学成像和临床信息数据作为模型的参考指标。有研究表明，在多种模态的数据驱动下可以使得模型的准确度得到提升，在本文的实验部分得到了验证。本文创新性的将 PET 影像、MRI 影像、临床数据、基因数据四种不同类型的数据相结合输入到模型中进行训练，对比传统运用三种模态数据对 AD 进行诊断分类的模型，可以得到更高的精确度。

## 1.4 论文组织结构

本文的第一章前言，论述的是本论文的选题背景，在阿尔茨海默病深度学习研究领域的单模态、多模态相关性工作，现有问题及本文的解决方案和创新点；本文的第二章方法与材料，论述数据集来源、数据预处理、模型架构、实验设计；本文的第三章实验与结果，论述的是模型在训练过程中超参数调优、多模态有效性实验、性能对比实验、消融实验的效果。有效验证了多模态性能、非对称跨模态交叉注意力机制的贡献、以及不同模态数据对模型的重要性；本文的第四章系统的设计与实现，论述的是实现智能交互式辅助筛查系统的思路与方法，其中包括蒙特利尔（MoCA）认知量表数据集来源、交互式MoCA量表筛查系统设计思路以及web应用程序端模型与系统的架构设计；本文的第五章讨论与结论，论述的是本文模型与系统的临床意义以及未来的发展方向。

## 1.5 小结

本文研究围绕“基于深度学习医学多视角的阿尔茨海默病辅助筛查系统研究与设计”展开，结合国际前沿的多模态医学数据诊断 AD 分类发展趋势，针对临床诊断中的实际需求，提出一种融合MRI影像、PET影像、基因遗传数据和临床数据的创新性非对称跨模态交叉注意力机制算法框架。研究严格遵循深度学习项目的开发流程，通过研究背景、模型设计、算法实现与实验验证、算法应用五个核心阶段展开论述，全文章节依此划分并系统阐述。

## 2 数据与方法

### 2.1 数据集与预处理

本文中使用的数据来自 ADNI 数据库，该数据库提供了 4 项研究计划（ADNI1、ADNI2、ADNI GO 和 ADNI3）中 2220 多名患者的影像学、临床数据和遗传基因数据。ADNI 数据库建立目的是测试这些数据相结合是否有助于判断 MCI 和 AD 的病情进展，并利用这些被整理后的数据为推动生物学、医学和计算机等相关学科研究 AD 的发展提供基础。本文对来自于 ADNI1、2 和 GO 患者的多模态数据参照了以往研究者的处理方法。为了结合各种模态的数据，本文研究专注于具有可用影像学、遗传学和临床数据的患者。在模型的训练中只使用了那些记录了所有四种模态的患者（称为重叠数据集），如表 2-1 所示。

表 2-1 ADNI 数据集统计

	人数(名)	CN	MCI	AD
临床数据	2384	942	796	646
MRI 影像	551	278	123	150
PET 影像	482	225	98	132
基因数据	805	241	318	246
数据重叠	239	165	39	35

注：表中显示每种数据的参与者人数及其诊断。下表显示了数字的参与者，并进一步将参与者分为他们的诊断类型。数据重叠部分是指记录了所有四种模式的测试者。

### **2.1.1 标签预处理**

自 2003 年 ADNI 研究计划启动以来，越来越多的参与者被加入到 ADNI 的研究计划中，并且持续追踪了最初参与者的 AD 进展。本文通过数据分析发现一些最初被标记为 CN 和 MCI 的患者随着时间的推移和病情的发展，其诊断结果发生了变化，例如，随着时间的推移部分 MCI 患者发展成为 AD 患者。同时发现部分患者进行了多达 16 次 MRI 扫描，而临床数据和基因数据的收集频率要低得多，因此，将每位患者的四种数据进行有效结合处理过程非常繁琐，有时会出现诊断结果相互矛盾的情况。因此，对于本文中重叠数据集，使用了每位患者最新的临床数据，以及采用最新的诊断结果。

### **2.1.2 临床数据预处理**

对于临床数据，本文选择使用 2384 名患者的神经系统检测数据和人口统计学数据。神经系统检测数据反映 AD 中在不同中枢神经系统区域的受影响情况<sup>[15]</sup>。而人口统计学数据也有助于AD的诊断，有研究表明，年龄和性别可能会对大脑状态产生影响<sup>[16]</sup>；临床数据预处理：本文将参与者的人口统计数据与神经系统检测数据结合在一起作为该参与者的临床数据信息，临床数据总共包含 7 项不同数据，分别是性别、年龄、蒙特利尔认知测试分数（MoCA）、简易精神状态检查分数（MMSE）、临床痴呆评级（CDR）、老年抑郁量表（GDS）、功能活动问卷分数（FAQ），代表着参与者的 7 种不同的属性或医疗检测结果。

对临床数据的处理主要包括两步：处理缺失值、特征缩放。前者采用了中位数策略，即用中位数填补缺失值。后者会对数据进行均值标准差归一化，即

将数据缩放到均值为 0、标准差为 1 的标准正态分布中。

### 2.1.3 影像数据预处理

MRI 影像以实现非侵入性检测的优势，能够精准呈现出脑组织内部的解剖结构与形态学特征，可以通过图像直观的显示 AD 患者脑部萎缩的临床表现。以往的研究表明<sup>[17]</sup>，即使在阿尔茨海默症早期阶段，MRI 影像也可以清晰的展现 AD 患者病情进展，本文使用 MRI 影像为其中一种多模态模型训练数据。此外，本文还选择使用  $^{18}\text{F}$ -AV-45 作为示踪剂的淀粉样 PET 影像作为另一种多模态模型训练数据。这种成像技术允许对淀粉样斑块进行无创的体内检测，而淀粉样斑块是阿尔茨海默症的主要神经病理学标志之一<sup>[18]</sup>，其形成方式与基因表达有一定的关系<sup>[19]</sup>；本文使用的图像数据包含参与者的横断面 MRI 影像和 PET 影像。为了减少模型参数数量和资源消耗，本文只选择了参与者 MRI 和 PET 影像的中间位置切片，如图 2-1、2-2 所示。中间位置的脑部切片包含更多的脑部结构，选择相同位置的 MRI 切片和 PET 切片可以直观地反映不同参与者的脑部萎缩和淀粉样蛋白沉积情况，为模型的训练提供更丰富的信息。



图 2-1 MRI 影像切片示例图

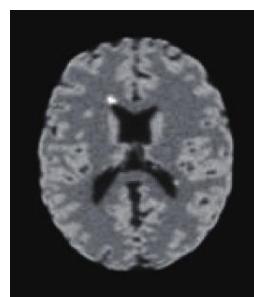


图 2-2 PET 影像切片示例图

## 2.1.4 基因数据预处理

基因数据由 805 名 ADNI 参与者的全基因组测序（WGS）数据构成，这些数据由 Illumina 的非临床实验室改进修正案（non-CLIA）实验室在 2012 年和 2013 年进行测序。2014 年，ADNI 使用 Burrows-Wheeler 比对器和基因组分析工具包生成了相应的变异调用文件（VCFs）。每个受试者的原始 VCF 文件中大约有 300 万个单核苷酸多态性（SNPs）。然而，并非所有检测到的 SNPs 都有助于预测阿尔茨海默病。

本文按照文献<sup>[20]</sup>中详细介绍的既定预处理步骤来减少 SNPs 的数量，仅保留相关的遗传成分。经过筛选，每位患者剩下 547,863 个 SNPs。由于本文只有 805 名拥有基因数据的患者，这使得数据矩阵非常稀疏。

等位基因是指位于一对同源染色体相同位置上控制同一性状不同形态的基因。某些特定的等位基因被视为 AD 的风险因素。例如，载脂蛋白 E（APOE）基因存在多个等位基因，其中 APOE ε4 等位基因是目前公认的晚发性 AD 的重要遗传风险因子<sup>[5]</sup>。本文使用随机森林分类器作为有监督的特征选择方法，以确定哪些是最重要的特征，将特征空间减少到大约 15,000 个 SNPs，最终减少为 500 个 SNPs。最终数据被分为三类：无等位基因、有任意一个等位基因或两个等位基因都存在。

## 2.1.5 确定训练数据集

本文使用了 239 名患者的数据训练 ACMCA 模型，这些患者的四种类型数据均可用。重叠数据集是从上述原始数据（MRI 影像数据 551 名患者、PET 影像数据 482 名患者、单核苷酸多态性（SNP）数据 805 名患者和临床数据 384 名患者）中选取的，用于预测 AD 的诊断类型。

对于每位患者 SNP 数据只会采集一次，但临床数据和影像学数据却会多次采集。为了确保正确合并数据，避免出现数据重复使用的情况，研究中使用临床数据中的时间戳，并将其与最接近的 MRI 影像扫描日期进行匹配。这样的数据预处理，可以最大限度的避免患者数据的重复使用。患者的人口统计信息如表 2-2 所示。

表 2-2 女性参与者的人口统计信息

组别	人数(名)	女性占比(%)	平均年龄 (岁)
CN	165	53.9	77.8
MCI	39	34.2	76.6
AD	35	31.4	78.1

注：此表展示了各类型的参与者数量，以及每组中女性的占比和平均年龄。

## 2.2 ACMCA 模型架构设计

本文构建了 ACMCA 模型以充分挖掘 MRI 影像、PET 影像、基因数据与临床数据之间的跨模态数据关联信息以进行对三种 AD 类型的诊断任务。

ACMCA 模型由浅层特征提取模块、多模态特征融合模块、深层特征提取模块以及输出分类四大模块构成。在浅层特征提取模块，由于不同模态数据的特性，本文采用了差异化提取特征的方法，主要分为医学影像特征提取模块与结构性数据特征提取模块。

完成浅层特征提取之后，将提取特征输入至多模态特征融合模块，在该模块中运用非对称跨模态交叉注意力机制，实现四种类型数据特征融合，并生成融合后的多模态特征。

然而，这些融合后的特征并不直接用于分类，而是输入至由自注意力机制和频域转换器构成的深层特征提取模块中。

在这深层特征提取模块中，模型能够聚合多模态特征信息，进而，深入的学习到不同模态特征之间的深层联系，随后将特征信息输入到分类模块中。具体模型架构见图 2-3。

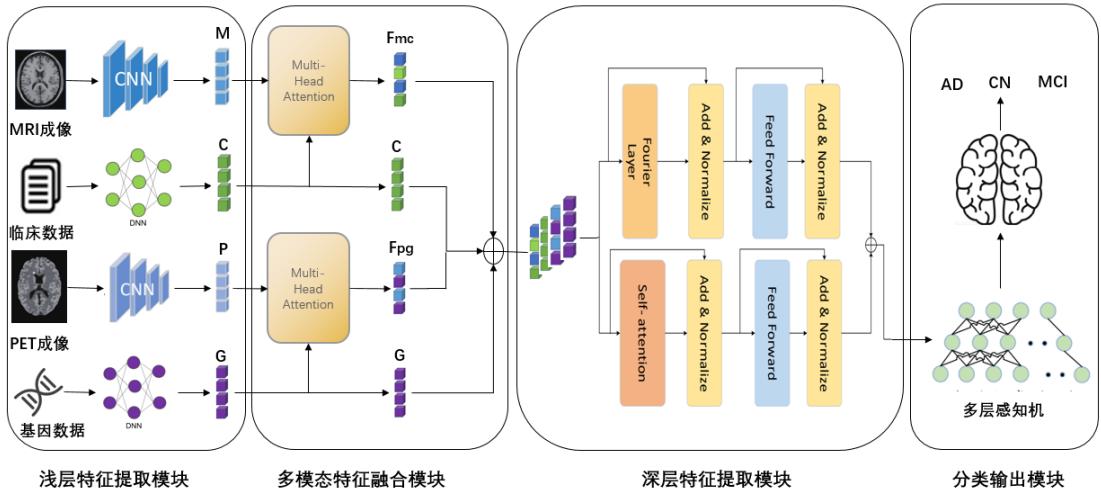


图 2-3 非对称跨模态交叉注意力机制 (ACMCA) 框架模型图

## 2.2.1 浅层特征提取模块

本文借助预训练模型 ResNet50 进行图像模态的特征提取，以获取更丰富的图像特征。由于 MRI 影像和 PET 影像是医学成像数据，使用 ResNet50 能够有效提取其中的高级图像语义特征。对于基因数据，将三种类型的数据（无等位基因、有任意一个等位基因、两个等位基因都存在）转换为深度神经网络（Deep Neural Network, DNN）输入的格式，将其表示为编码向量，并且同步执行降维操作。对于临床数据（包含患者的 7 项检测指标）运用 DNN 对编码后的临床数据进行特征提取，并同步执行增维操作。以确保临床数据特征、基因数据特征、MRI 影像特征以及 PET 影像特征在维度上保持一致。统一这些特征的维度，可以使它们在特征空间中具备相似的表现形式，为后续的特征融合模块提供基础。

## 2.2.2 多模态特征融合模块

在深度学习技术 AD 的分类诊断研究领域中，医学影像直观反映脑部结构变化，临床数据涵盖患者认知功能及症状表现，基因数据则揭示与疾病相关的蛋白质细胞表达，三者结合能够提供更全面的诊断依据。鉴于临床数据与基因数据作为结构化数据，与医疗成像数据在数据结构和特征分布上存在较大差异因此，在模态融合时，更应关注于不同数据间的显著差异关系。

本文提出一种非对称跨模态交叉注意力机制，用于捕捉临床数据、基因数据与 MRI 影像数据、PET 影像数据之间的跨模态信息交互，具体模块设计见图 2-4。该机制通过位置编码，将不同模态的特征分别映射为查询 (Query)、键 (Key) 和值 (Value) 向量。接着，通过计算查询向量与键向量之间的相似度以确定各个位置之间的关联。最后，对值向量进行加权求和，得到包含不同模态关注度信息的表示向量，实现跨模态信息的有效融合。

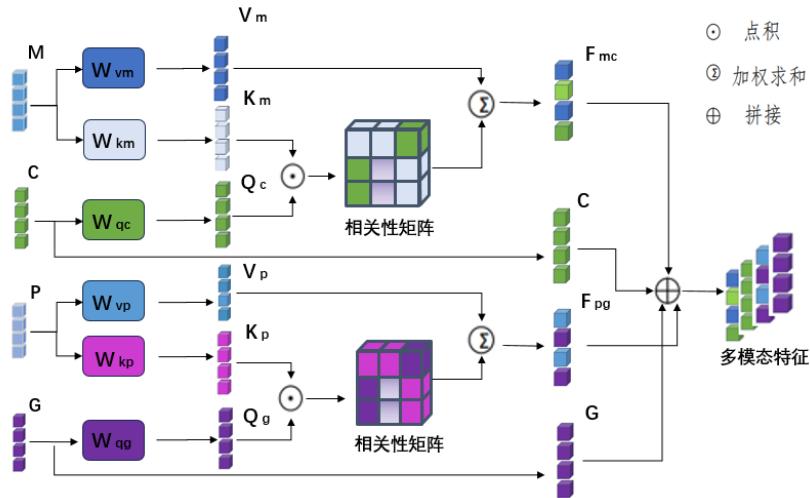


图 2-4 非对称跨模态交叉注意力机制 (ACM) 多模态融合图

在 ACMCA 模型的多模态融合过程中，临床数据、基因数据和 MRI 影像数据、PET 影像数据的融合被称为 Image-numerical 融合分支。在这融合分支中，操作过程基于特定的原理与公式。首先，对经过位置编码的临床数据特征、基

因数据特征进行处理，使其成为查询向量，这一过程遵循公式（1）以及公式（2）。而 MRI 特征与 PET 特征则会分别映射为对应的键向量与值向量，对应着公式（3）与公式（4）。接下来，针对每一个查询向量，都要计算它和所有键向量的点积，再将这个点积结果除以向量维度的平方根，之后运用 Softmax 函数，把得到的结果转化为概率向量。通过这样的操作，就能获取不同位置特征之间的相关性矩阵。最后，利用这个相关性矩阵与值向量进行运算，进而得到经过注意力机制处理后的成像表示，这一结果可由公式（5）和公式（6）表示。

$$Q_c = W_{qc} C \quad (1)$$

$$Q_g = W_{qg} G \quad (2)$$

$$K_m = W_{km} M , \quad V_m = W_{vm} M \quad (3)$$

$$K_p = W_{kp} P , \quad V_p = W_{vp} P \quad (4)$$

$$F_{mc} = CMA(Q_c, K_m, V_m) = \text{Softmax}\left(\frac{Q_c K_m^T}{\sqrt{d_k}}\right) V_m \quad (5)$$

$$F_{pg} = CMA(Q_g, K_p, V_p) = \text{Softmax}\left(\frac{Q_g K_p^T}{\sqrt{d_k}}\right) V_p \quad (6)$$

在本文的多模态融合模块中，从临床数据特征提取模块输出的特征标记为 C，它包含了患者的临床信息，神经系统检测结果、人口统计学数据等信息。此外模块中，从基因数据特征提取出的特征标记为 G，M 代表从 MRI 影像中提取的特征，P 则是从 PET 影像获取的特征，C、M、G 和 P 这四者的特征维度都设定为 100。这意味着它们都是 100 维的向量，用于后续的特征融合。在 ACMCA 模型训练过程中，W 发挥着重要作用，它是可迭代优化的转换矩阵。在模型训练阶段，W 会不断调整自身的矩阵数值值，实现对不同模态特征的线性变换。在注意力机制的框架下， $Q_c$  和  $Q_g$  是由 C 与 G 经过  $W_{qc}$ 、 $W_{qg}$  线性转换

得到的查询向量。其作用是在 MRI 影像 和 PET 影像的特征中学习相关联的特征信息，帮助模型学到跨模态数据的关键信息。而 MRI 影像特征  $M$  和 PET 影像特征  $P$ ，会分别借助转换矩阵转化为键向量  $K_m$ 、 $K_p$  以及值向量  $V_m$ 、 $V_p$ 。键向量用于查询向量计算相似度，以确定不同位置特征之间的关联紧密程度；值向量则依据相似度计算的结果进行加权求和操作，最终生成多模态融合特征。这些多模态融合特征中蕴含了不同模态间的关联信息。其中， $F_{mc}$  是 MRI 影像特征与临床数据特征融合后得到的特征， $F_{pg}$  则是 PET 影像特征与基因数据特征融合后得到的特征。在经过一系列复杂且有序的计算后，该模块的输出由公式（7）来表示。

$$\text{concat}(F_{mc}, C, F_{pc}, G) \quad (7)$$

这种基于非对称跨模态交叉注意力的融合方法，能够引导模型着重学习医疗图像与数值数据间的对齐关系，进而高效地理解成像数据块和临床数据之间的内在联系。通过该融合方式所产出的混合特征，被定义为多模态特征，其张量形态为(批次大小，通道数，特征维度)，具体表现为(32, 3, 100)。

### 2.2.3 深层特征提取模块

深层特征提取模块的作用是对多模态注意力模块输出后融合特征进行深层次的提取。其作用是提取出多模态特征内部的有效关联信息。从而，让不同位置的特征能够更好地交互。该模块由自注意力模块和傅里叶模块组成。Patro BN 等研究者把 transformer 和频域转换器相结合<sup>[14]</sup>，利用频域分析提升特征提取效果。其研究已表明这种结合方式可以加强模型对特征的提取性能，并且能有效提升模型的精确度，本文也会在实验部分验证其有效性。

傅里叶模块利用 Fnet 结构，在傅里叶变换层（Fourier Layer）运用离散傅里叶变换来传递多模态特征内部信息。对于长度为 N 的序列，快速傅里叶变换的公式为：

$$X[k] = \sum_{n=0}^{N-1} e^{-i\frac{2\pi}{N}nk} x[n] \quad k = 0, 1, \dots, N-1 \quad (8)$$

具体来说，对于特征的每个位置  $k$ ，Fnet 结构会沿着维度方向对原始多模态特征进行傅里叶变换，将其转化为新的表示形式。这种新形式包含了该特征其他位置的信息。把序列的全局信息转换到频域后，模型就能捕捉到序列中不同位置之间的关系。和 Transformer 结构类似，频域表示输出后会进入全连接的前馈神经网络，并且在傅里叶变换层和前馈神经网络层之后，都会进行残差连接和层归一化操作。

## 2.2.4 输出分类模块

深层特征提取模块输出的特征会被输出到分类模块。分类模块由两层多层感知机构成。首先，全连接层对深层多模态特征进行线性变换，然后使用 ReLU 激活函数引入非线性变换，最后输出三种分类结果，分别是 CN、MCI、以及 AD。在训练模型时，使用交叉熵（Cross Entropy Loss）作为损失函数，其定义为：

$$L = \frac{1}{N} \sum_i L_i = -\frac{1}{N} \sum_i \sum_{c=1}^M y_{ic} \log(p_{ic}) \quad (9)$$

注：  $i$  表征样本的总体数量， $M$  代表分类任务中的类别总数，鉴于本文聚焦于阿尔茨海默症的三分类诊断任务，此处  $M$  取值为 3。其中， $y$  对应样本的真实标签，直观反映样本所属的实际类别； $p$  则为模型输出的预测值，体现模型对样本类别的判断结果。

通过不断减小这个损失函数的值，对模型的参数进行优化，使模型的预测结果更接近真实情况，以此提高模型在 AD 分类任务中的表现。

### 3 实验与结果

#### 3.1 参数实验和选择

为了确定模型训练的最优参数组合，开展了超参数实验研究。在实验过程中，固定部分参数，分别探究不同训练周期（Epoch）、批次大小（Batch Size）以及特征维度对最终模型分类准确率的影响。当批次大小设定为 32，且特征维度为 100 时，观察不同训练周期下模型的分类准确率变化，实验结果通过图 3-1 呈现。从图 3-1 中可以清晰看到，随着训练周期的增加，模型分类准确率呈现先升后降的变化趋势。

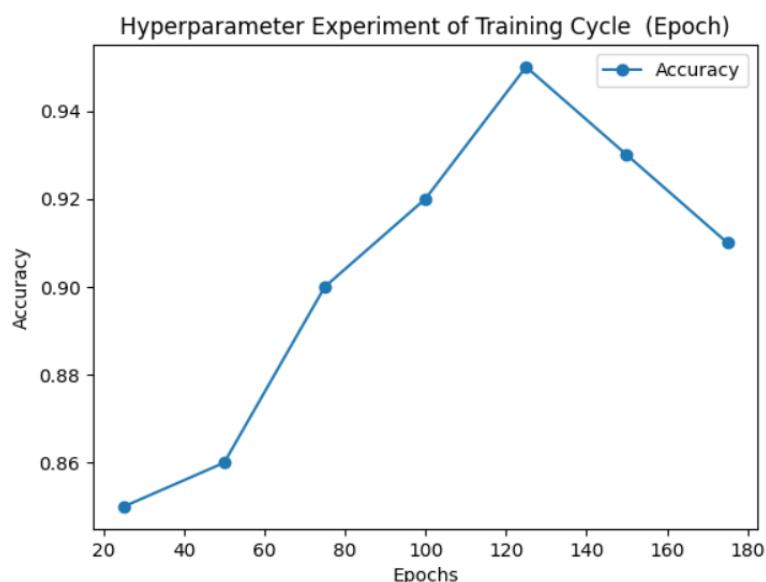


图3-1 训练周期（Epoch）超参数实验

在另一组实验中，将训练周期固定为 125，特征维度保持 100 不变，研究不同批次大小对模型分类准确率的影响，相关结果展示在图 3-2 中。通过对图 3-2 的分析，能够直观了解到批次大小的对模型精确度的影响。

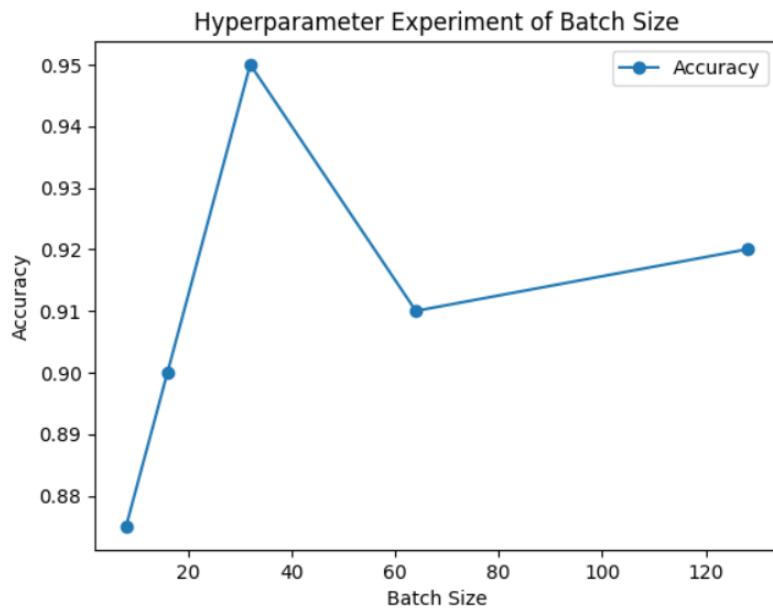


图3-2 批大小（Batch Size）超参数实验

同时，还进行了特征维度的实验，在训练周期为 125，批次大小为 32 的条件下，研究不同特征维度对模型分类准确率的影响，实验结果见图 3-3。从图 3-3 中可以观察到不同特征维度下模型分类准确率的变化趋势。

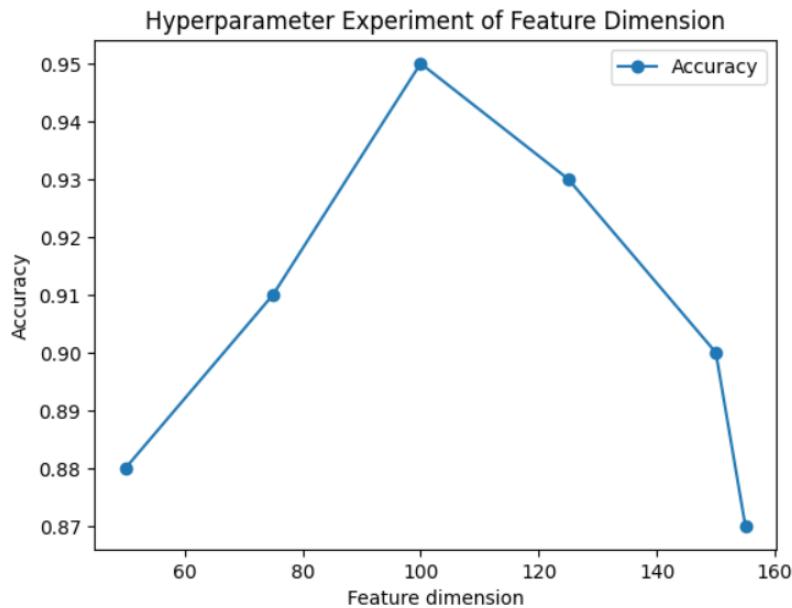


图3-3 特征维度（Feature Dimension）超参数实验

通过综合比较各个超参数设置下模型的表现，筛选出了表现最佳的一组超参数用于本文模型的训练。最终确定的训练模型使用的超参数如下表 3-1 所示。

表 3-1 模型最佳参数设置

参数名	数值
成像特征维度	100
非成像特征维度	100
图片大小	$224 \times 224$
学习率	1e-3
批量大小 (Batch size)	32
迭代次数 (Epochs)	125

### 3.2 评价指标

为评估 ACMCA 模型在 AD 多分类预测任务中的性能，本研究以 ADNI 数据库中的 MRI 影像、PET 影像、基因数据及临床数据为基础展开实验比对，并将其与其他基线模型进行比对。在性能评价方面，选用了准确率 (Accuracy)、召回率 (Recall)、特异性 (Specificity) 和 F1 分数 (F1 - score) 这四个典型指标，各指标的计算方式如下：

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (11)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (12)$$

$$F_1 = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (13)$$

在本文的模型评估体系中，真正例（TP）表示模型正确识别的阳性样本数值；真反例（TN）为模型正确识别的阴性样本数；假正例（FP）指被误判为阳性的阴性样本；假反例（FN）则是被误判为阴性的阳性样本。为系统评价模型性能，采用受试者工作特征曲线（ROC）作为评估工具。该曲线通过动态描绘真正例率（TPR，即灵敏度，反映模型对阳性样本的识别能力）与假正例率（FPR，等于 1 减去特异度，体现模型对阴性样本的识别精度）之间的关系，直观展现分类器的判别性能。同时，借助 ROC 曲线下面积（AUC）对分类器性能进行量化分析，该指标能够综合反映模型在不同阈值下的分类表现，该平均 ROC 曲线能够反映模型的敏感性和特异性，其曲线下面积越大，意味着模型在识别阳性和阴性样本方面的综合性能就越出色。

### 3.3 多模态有效性实验

为探究多模态数据在AD诊断中是否具有互补信息的作用，本文开展了一系列基于不同模态数量数据的实验，对 AD、MCI、CN 进行三分类诊断。

在单模态实验中，对于 MRI 影像和 PET 影像分别进行单独测试。运用 ResNet50 模型对成像数据进行特征提取，所提取的成像特征不经过多模态注意力模块，而是直接输入至深层特征提取模块。通过这一操作，能够得到带有自注意力权重信息的单模态特征表示。之后，该单模态特征输入分类模块处理，完成 AD 的分类任务。当仅采用临床数据、基因数据时，先对临床数据、基因数据进行编码，再利用深度神经网络（DNN）提取临床数据、基因数据特征，后续流程与仅使用成像模态数据的实验保持一致。实验结果整理于表 3-2。

在开展单模态实验的过程中，仅依赖基因数据数据进行实验所呈现的效果是最为突出的。这一现象意味着，包含了患者多种SNPs基因表达数据，与单一的二维医学影像数据相比，蕴含着更为丰富多样的信息。

表 3-2 模态性能比较

模态	准确率	召回率	特异性	F1 分数
<b>基因数据</b>	<b>0.838</b>	<b>0.712</b>	<b>0.894</b>	<b>0.791</b>
临床数据	0.810	0.553	0.853	0.773
MRI	0.823	0.551	0.843	0.562
PET	0.834	0.637	0.881	0.673
临床数据+MRI	0.845	0.809	0.912	0.824
基因数据+PET	0.839	0.794	0.903	0.832
<b>临床数据</b>				
<b>+MRI+基因数 据+PET</b>	<b>0.865</b>	<b>0.821</b>	<b>0.923</b>	<b>0.872</b>

相较于使用单一模态数据，运用多模态数据对 AD 进行诊断展现出了更为卓越的性能。当把临床数据和MRI影像数据相结合时，诊断的准确率能够达到 84% - 85%，F1 分数也会超过 82%，这些数据均高于所有单一模态实验所获得的结果。当把基因数据和PET影像数据相结合时，诊断的准确率达到 83.9%，F1 分数也会超过 83%，这些数据也均高于所有单一模态实验所获得的结果。而同时使用基因数据、临床数据、MRI 成像数据以及 PET 成像数据这四种模态数据时，模型的分类性能最佳。这清晰地表明，源自不同模态的生物标志物能够从不同层面捕捉与疾病相关的信息，并且这些信息之间可以相互补充，从而显著提升了模型的分类能力。

### 3.4 性能对比实验

为验证ACMCA模型在AD诊断分类方面的卓越性能，本研究将其与近年来的多种基线模型展开对比分析。以下是对各基线模型的简要介绍：

模型 1：J Venugopalan 等人<sup>[20]</sup>采用简单拼接的方式整合提取的成像与非成像特征，随后直接运用随机森林分类器开展AD多分类诊断。

模型 2（MCAD）：J Zhang 等人运用交叉注意力机制融合成像与非成像特征并对融合后的特征进行降维处理，并借助交叉熵损失和模态对齐损失优化模型网络架构。

模型 3：H Chen 等人先后将成像特征与非成像特征输入到通道注意力模块和空间注意力模块进行模态融合，融合后的特征经卷积模块进一步提取，最终送入分类头。

模型 4（MADDi）：M Golovanevsky 等人<sup>[21]</sup>让三种不同模态特征先经过一个头自注意力模块，之后对处理后的特征两两进行交叉注意力操作，最后拼接融合特征用于分类。

模型 5（MADDi-ACM）：该模型将 MADDi中的多模态融合方式替换为本研究模型中的非对称交叉注意力机制。

模型 6（ACMCA -CM）：此模型将 MAMDF 模型中的非对称交叉注意力模块替换为对称交叉注意力模块。

模型 7（ACMCA-MCAD）：该模型仿照 J Zhang 等人的研究方法，先拼接MRI与PET影像像特征，再利用交叉注意力机制融合非成像特征。

模型 8（ACMCA）：本文中的跨模态交叉注意力机制模型，将本文模型与上述基线模型进行比较，实验结果如表 3-3 所示。

表 3-3 性能对比实验结果

模型	准确率	召回率	特异性	F1
模型 1	0.780	0.780	0.770	0.780
MCAD	0.735	0.639	0.820	0.619
模型 3	0.841	0.847	0.847	0.812
MADDi	0.859	0.761	0.924	0.769
MADDi-ACM	0.885	0.798	0.934	0.813
ACMCA-CM	0.865	0.810	0.933	0.810
ACMCA-MCAD	0.859	0.787	0.928	0.793
<b>ACMCA</b>	<b>0.948</b>	<b>0.865</b>	<b>0.935</b>	<b>0.855</b>

从表 3-3 与图 3-4 的统计数据结果能够表明，注意力机制可以显著增强模型在多分类任务中的性能。J Venugopalan 等人构建的模型 1，只是单纯地把不同模态的特征拼接起来，这种粗略的特征提取方式，使得模型难以学习到多模态特征之间的关联信息。本文性能对比实验中 MADDi 模型、H Chen 等人开发的模型 3 以及本文提出的 ACMCA 模型都借助了注意力机制处理不同模态的特征，相较于模型 1 的实验结果，这些模型在分类准确率上分别实现了 16.8%（ACMCA 对比模型 1）、6.1%（模型 3 对比模型 1）和 7.9%（MADDi 对比模型 1）的提升。这准确率的提升充分说明了，在模态融合过程中运用注意力机制能更好地学习到多模态特征数据之间的内在关联。

实验结果图 3-4 还清晰地显示出，ACMCA 模型所采用的非对称跨模态交叉注意力机制在融合成像特征与非成像特征方面具有显著优势。MADDi 模型在进行多模态融合时，采用的是对两两模态直接进行对称交叉注意力处理，然后再

进行拼接特征的方法。当把 MADDi 模型的对称交叉注意力机制替换为本研究的非对称跨模态交叉注意力机制后，模型对成像特征与非成像特征之间联系的关注程度明显提高，分类准确度也随之提升了 2.6%。J Zhang 等人同样认识到成像特征与非成像特征之间存在较强的异构性，他们在模型 2 中先对成像特征进行拼接，然后对拼接后的成像特征与非成像特征实施对称交叉注意力处理。本研究将 ACMCA 模型原本的非对称跨模态交叉注意力机制，替换为与模型 2 相同的处理方式后，发现模型的分类性能出现了一定程度的下降。

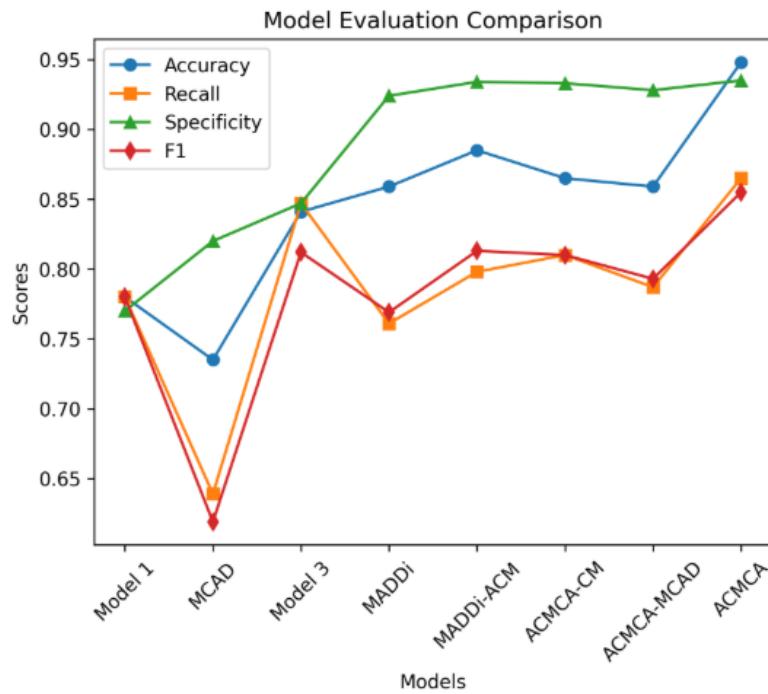


图3-4 性能对比实验结果图

### 3.5 消融实验

为了验证模型各个子结构的有效性，在本节实验中，将本文的 ACMCA 模型与四种经过消融处理的模型，置于同一重叠数据集上开展 AD 的三分类诊断任务，并对它们的性能加以对比。这四种消融模型具体如下：

**ACMCA-WCM:** 该模型是在本文的ACMCA模型基础上，移除交叉注意力模块，让成像特征与非成像特征直接进入深层特征提取模块。

**ACMCA-WDE:** 在本文的ACMCA模型里，把非对称跨模态交叉注意力融合模块输出的特征予以拼接（移除深层特征提取模块），紧接着直接执行分类任务。

**ACMCA-WFnet:** 此模型是在本文 ACMCA模型中，对深层特征提取模块里的傅里叶模块进行屏蔽操作，其余模块维持原样。

**ACMCA-WT:** 这个模型是从本文的 ACMCA模型出发，将深层特征提取模块中的自注意力模块加以屏蔽，而其他模块保持不变。

表 3-4 消融实验结果

模型	准确率	召回率	特异性	F1	AUC
ACMA-WCM	0.785	0.724	0.932	0.722	0.910
ACMA-WDE	0.805	0.751	0.939	0.751	0.927
ACMA-WFnet	0.825	0.778	0.945	0.776	0.926
ACMA-WT	0.826	0.782	0.946	0.778	0.938
<b>ACMCA</b>	<b>0.946</b>	<b>0.865</b>	<b>0.935</b>	<b>0.855</b>	<b>0.948</b>

实验成果表 3-4 呈现出，当在 ACMCA 模型中去除非对称跨模态交叉注意力模块之后，其性能出现大幅下滑。这一现象充分说明，非对称跨模态交叉注意力机制能够充分挖掘到不同模态数据之间的内在联系，进而提升模型诊断分类的性能。与原始的 ACMCA 模型相对比，若将深层特征提取模块移除，模型的性能表现亦同样出现下滑。此外，当分别对深层特征提取（DE）模块里的 Fnet 模块和 Transformer 模块进行屏蔽操作时，最终获取的实验结果显示两者之间模

型的性能较为近似，这一情况证实了频域转换器能够发挥出与自注意力机制相似的作用效果。

在图 3-5 中，清晰展示了消融实验里各模型的 ROC 曲线。ROC 曲线下的面积大小，能够从更为全面的视角去展现模型的分类能力。相较于原始模型，移除深层特征提取模块，该模型在 AD 多分类任务中的 AUC 值降低了 0.021；若仅仅去掉自注意力模块，AUC 值会下降 0.01；而仅仅去掉傅里叶模块时，AUC 值则下降 0.022。这些数据充分表明，深度特征提取模块并非单纯通过网络堆叠就能实现理想效果，其设计的合理性极为关键。

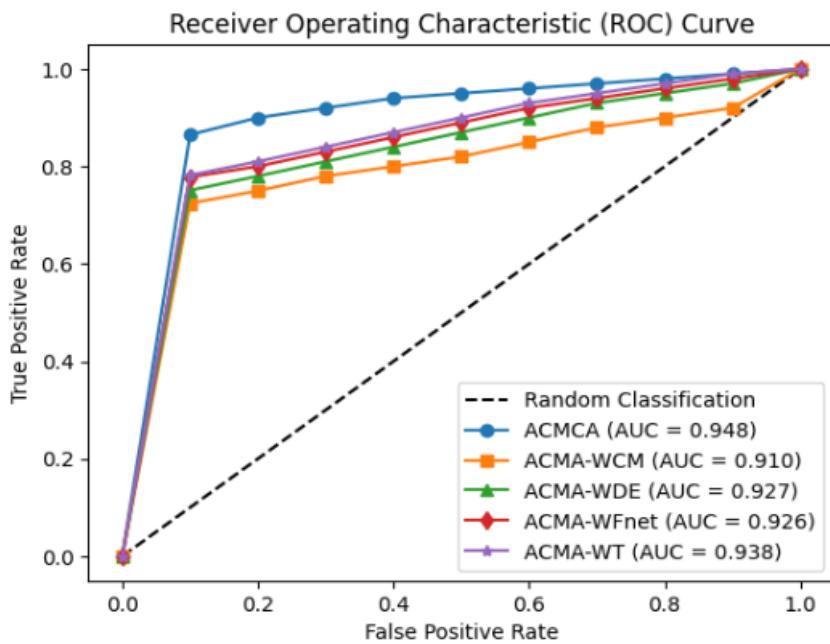


图3-5 非对称跨模态交叉注意力模块消融实验 ROC 曲线

## 4 系统设计与实现

### 4.1 研究背景

AD 是痴呆症中最为常见的类型，据统计，全球约有 5000 万人深受其害。随着人口老龄化的加剧，预计到 2050 年，患者数量将激增两倍，平均每 85 人中就可能有一人患病<sup>[22]</sup>。在中国，截至 2019 年，AD 患病人数累计已超过 1000 万<sup>[23]</sup>，患病人数位居全球之首，由 AD 引发的认知、功能和行为退化，严重影响患者生活质量，也给家庭和社会带来了巨大的经济负担与精神压力。

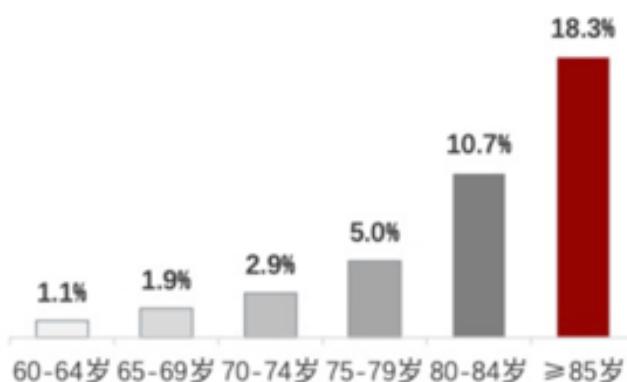


图4-1 年龄段合并患病率

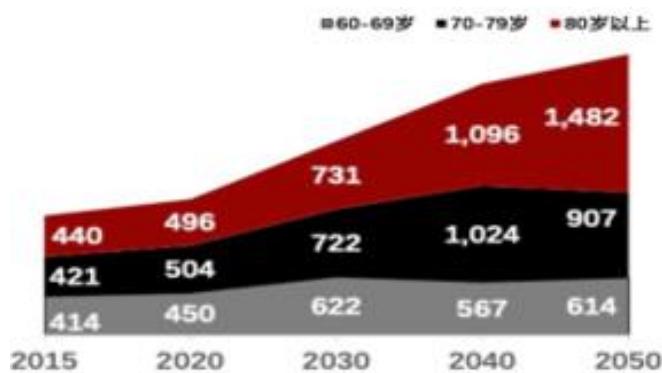


图4-2 阿尔兹海默症发展趋势(万人)

临幊上常用蒙特利尔认知评估（Montreal Cognitive Assessment, MoCA）测幊对 AD 进行初步筛查<sup>[24]</sup>。但传统纸质 MoCA 测幊存在诸多弊端，临幊测试效率较低，测幊过程需医生与患者一对一互动，且依赖医生主观判断手动评分，

导致测试结果的置信度和客观性较差，单位时间内筛查人数有限，无法满足日益增长的筛查需求。因此，开发一款测试方式简便、评价客观科学的阿尔茨海默病筛查系统对于缓解社会医疗压力、提升患者生活质量具有重要意义。

## 4.2 系统解决痛点问题

传统的 AD 诊断方式需要患者多次往返医院，在医院进行的纸质版人工 MoCA 量表的筛查测试。在测试过程中，医生需要引导患者进行一对一互动，传统的筛查方式效率较低，且会受到医生主观因素的影响，例如不同医生对评分标准的理解差异等因素，都可能导致测试结果置信度不高。

“阿尔茨海默病智能辅助筛查系统”旨在解决这些问题。借助深度学习技术构建了图像识别模型，该模型在实际应用中能够优化手绘图像识别任务，这是 MoCA 量表中的一项重要测试内容。同时，依托百度飞桨（Paddle）框架的语音识别转换技术，搭建了智能语音交互系统。在筛查过程中，可以实现无需医护人员在场的情况下，测试者可以通过语音交互的方式完成 MoCA 测试，提高了测试的便捷性。

系统利用 ACMCA 模型进行深度学习训练，将收集到的脑部 MRI 影像、MoCA 测试分数与简易精神状态检查分数等临床医学数据相结合输入到模型中（考虑到基因数据和 PET 影像检测费用高且预处理复杂，在实际应用场景中，仅使用 MRI 影像与临床数据对模型进行训练，这样既贴合实际需求，又能保证模型的实用性和可靠性）最终输出测试者的患病风险评估，可以提高专业医疗机构人员的筛查效率。

## 4.3 系统技术方案

本系统为阿尔茨海默病辅助筛查系统，主要由微信小程序端（交互式 MoCA 量表）以及 Web 端（非对称跨模态交叉注意力类激活预测模型）两部分组成。

### 4.3.1 微信小程序端

MoCA 测试包含视觉空间与执行能力、语言、记忆、注意力、执行功能、抽象思维、定向力等多个认知领域的测试任务。其中在执行视觉空间与执行能力测试中，要求患者绘制特定几何图形；执行语言测试时包括命名物体、重复短语等；执行记忆测试时涉及单词和数字的重复与回忆；执行注意力测试时通过数列回忆和逆序记忆等任务进行评估。这些任务综合起来，能有效帮助医生判断患者是否存在认知功能受损迹象，是临床诊断 AD 重要的神经心理学测试项目。

在微信小程序平台上，系统利用百度飞桨（Paddle）框架的语音识别转换技术，系统能够以语音的形式智能播报 MoCA 量表题目。同时，将测试者的语音回答自动转换为文本进行回答评分。目前，系统不仅支持普通话作答，还兼容粤语，满足了不同区域语言习惯的用户需求，拓展了系统的适用范围。

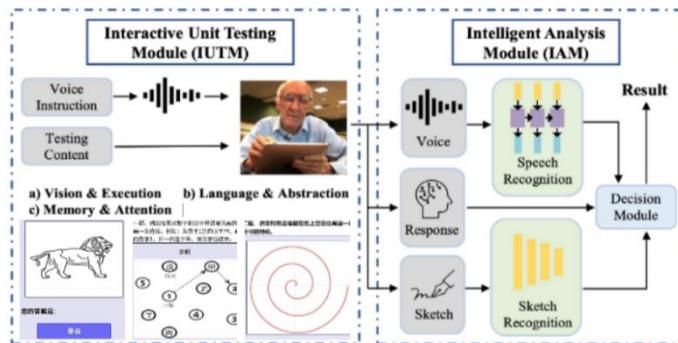


图 4-3 智能辅助筛查系统网络架构

系统基于深度残差神经网络（ResNet50）构建手绘几何图形智能评分模型可以实现 MoCA 量表中图形绘制任务的自动化评分，如图 4-4 所示。以 ResNet50 为基础模型运用迁移学习技术在 Sketch4IAS 数据集上训练模型。该数据集由宁波康复医院联合 QiKe Zhao、Sen Yang 等研究团队开发<sup>[25]</sup>，包含 400 张屏幕截图与 1850 张相机拍摄图像，如图 4-5 所示，其作用是评估测试者的空间结构能力与执行能力。训练时，通过调整模型的超参数提升模型的性能，对模型的输出结合 MoCA 评分标准，使模型能自动对手绘图形测试项目评分。利用模型进行评分可以减少人工评分的主观性。



图 4-4 自动化评分应用示例

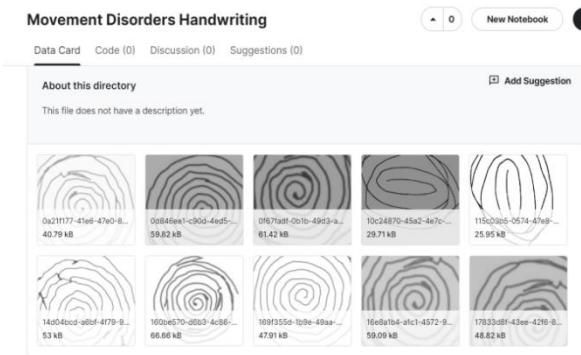


图 4-5 Sketch4IAS 数据集展示

系统利用前端技术改进了传统测试者需要在纸上绘图的答案输入方式，通过建立数字化绘画板，升级了交互界面。智能交互式蒙特利尔认知评估（MoCA）量表共集成了 17 项标准化测试项目。能帮助测试者发现认知功能衰退早期迹象，为治疗 AD 争取时间，延缓 AD 的进程，并提供了客观化的认知障碍筛查依据。

### 4.3.2 Web 应用程序端

Web 应用端是为专业医疗机构人员设计的应用程序，作用是辅助医疗机构人员评估患者患病风险。医疗机构人员输入受试者的 MRI 影像及相关检查数据资料，借助 ACMCA 基础模型，可实现自动化判断受试者患病风险。在实际应用中，受试者可通过微信小程序端自主完成 MoCA 量表自我检测，初步了解自身患病风险。专业医疗机构人员在临床诊断时，将 MoCA 量表评分数结果作为神经心理学测试数据输入到模型，实现 Web 端与小程序端的数据交互。

为了帮助专业医疗机构人员更直观地观察 MRI 影像结构的变化，系统引入了可视化技术来展示模型的注意力聚焦位置。应用类激活图（CAM, class activation map）技术进行可视化处理。类激活图是一张与原始图像尺寸相同的热力图，它反映了输入图像中各区域对模型预测输出的贡献分布情况。CAM 利用特征图权重叠加的原理获得热力图，公式与原理如下：

$$L_{\text{CAM}}^c = \sum_k \underbrace{w_k^c A^k}_{\text{linear combination}} \quad (14)$$

注： $L_{\text{CAM}}^c$  代表类别 c 的类激活映射得分。

在类激活图中， $k$  是一个索引值，用于遍历求和操作中涉及的不同项。在卷积神经网络的情境下，通常对应着不同的特征图，通过  $k$  可以对各个特征图相关项进行区分和计算； $w_k^c$  表示针对类别 c，第  $k$  个特征图的权重系数。它衡量了第  $k$  个特征图在类别 c 的分类决策中所占的重要性，反映了该特征图对类别 c 分类贡献的大小； $A^k$  即第  $k$  个特征图，是卷积神经网络中间层输出的结

果。特征图包含了图像在不同抽象层次和不同方面的特征信息，是后续进行类别决策的重要依据；求和符号，其作用是将所有基于不同  $k$  值计算得到的  $w_k^c A^k$  项累加起来，从而得到最终的类别  $c$  的类激活映射得分；linear combination（线性组合）：这里强调  $w_k^c A^k$  是对特征图  $A^k$  按照权重进行的线性组合操作，是一种基本的数学运算形式，用于融合不同特征图对类别  $c$  的贡献。

在处理 MRI 影像时，这一技术可以显示出大脑结构中哪些部分在模型判断患病风险时起到了主导作用，类激活图会高亮显示与疾病相关的脑部位置，如海马体等区域，让专业医疗人员能快速捕捉到可能存在萎缩病变的位置。

### 4.3.3 ACMCA 模型中 CAM 技术的嵌入

本文尝试将 CAM 技术嵌入到 ACMCA 模型中，通过对 ACMCA 模型中浅层特征提取模块的修改，将 CAM 技术嵌入其中，构建出非对称跨模态交叉注意力类激活模型（Asymmetric Cross - Modal Cross - Attention Mechanism - based Class Activation Model 简写 ACMCA-CAM）。在浅层特征提取模块，通过修改 ResNet50 网络结构，获取其最后一个卷积层输出的特征数据。这些特征数据蕴含了丰富的图像特征信息，如图 4-6 所示。

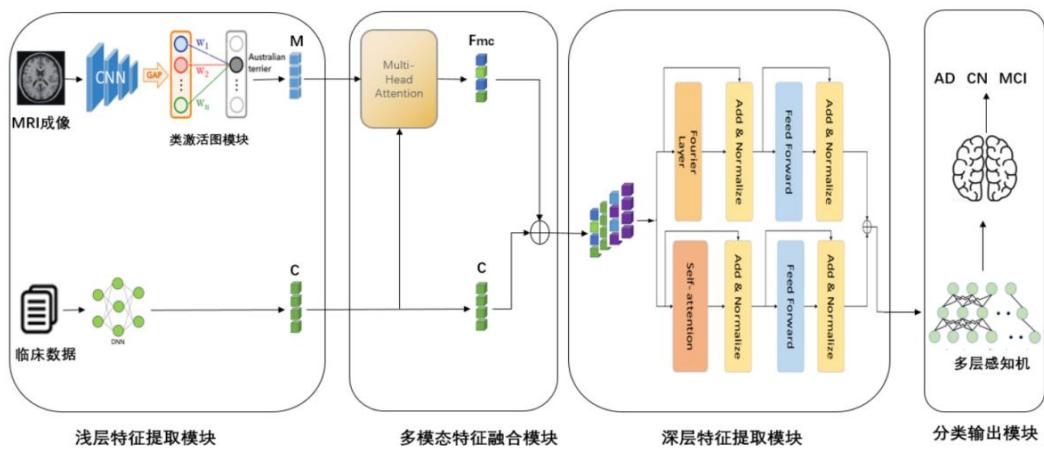


图 4-6 非对称跨模态交叉注意力类激活模型（ACMCA-CAM）架构图展示

接下来是权重获取模块。从 ResNet50 的全连接层中提取连接最后卷积层特征图的权重，这些权重在模型的分类决策中重要数据。其每一个权重值都反映

了最终分类结果的贡献程度。通过提取这些权重，为后续的 CAM 计算提供相关参数，从而确保生成的热力图能够真实地反映模型的决策逻辑。

在完成特征图和权重的参数后，进入 CAM 计算阶段。由公式（14），将特征图和权重进行加权求和操作，后将数据输出到多模态特征融合模块中。以直观的脑部区域热力图的形式展示了模型在进行特征提取时，对图像中各个区域的关注程度。不同颜色的区域代表了不同的重要性，颜色越深的区域表明该区域对分类决策的影响越大。

#### 4.3.4 ACMCA-CAM 模型的训练

在对临床 ADNI 数据的训练过程中，模型历经 100 次训练迭代。如图 4-7 所示，随着训练轮数的增加，模型的损失（Loss）总体呈持续下降趋势，与此同时，测试集中模型的精度稳步攀升，并且趋于稳定。这一现象表明，模型在训练过程中不断优化，有效学习到了数据中的特征模式。最终，模型在 ADNI 数据集上达到了 92.5% 的精度，有力地验证了模型在临床数据处理任务中的有效性与可靠性。

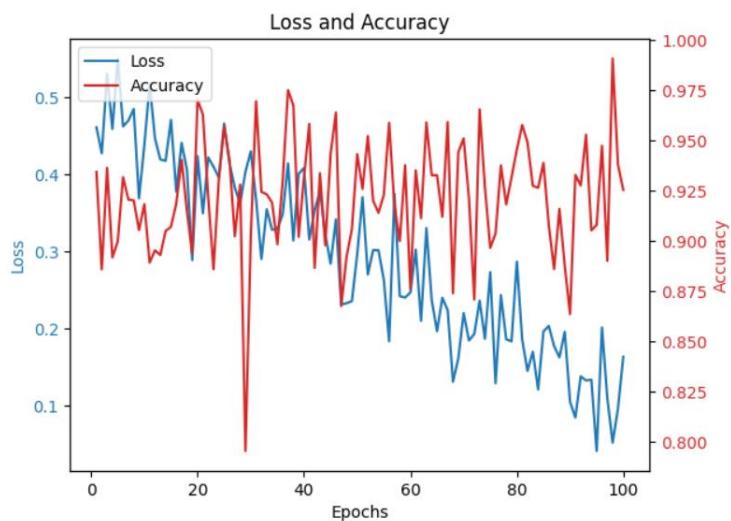


图 4-7 ACMCA-CAM 模型训练过程

综上所述，将 CAM 技术嵌入到 ACMCA 模型中，为医学影像分析提供了一种有效的可视化手段，有助于提升 ACMCA 模型在医学诊断领域的应用价值。

#### 4.3.5 ACMCA-CAM 模型的应用

在实际应用中，本文对某位女性患者的数据进行了分析，以展现了本模型的应用效果。临床数据显示，该患者年龄为 66 岁，MoCA 量表测试分数为 21 分，简易精神状态检查得分为 25 分，临床痴呆评级为 0.5，老年抑郁量表分数是 17 分，功能活动问卷分数同样为 17 分，如图 4-8 所示。

临床神经心理学测试项目	数值
病例号	600603
姓名	李四
性别	女
年龄	66
摩卡量表测试分数(MOCA)	21
简易精神状态检查 (MMSE)	25
临床痴呆评级 (CDR)	0.5
老年抑郁量表分数 (GDS)	17
功能活动问卷分数 (FAQ)	17

图 4-8 测试者的临床数据信息

结合其脑部 MRI 图像（见图 4-9），模型输出结果如图 4-10 所示，模型在测 MCI 类别上输出值高达 98%，表明模型判断该患者处于 MCI 状态具有较高置信度。

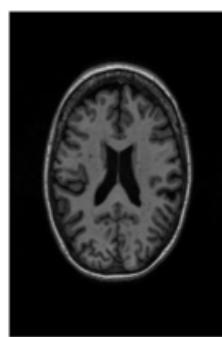


图 4-9 测试者的 MRI 影像

进一步利用模型生成的分析热力图和脑部区域热力图（见图 4-11），在图中能够直观地看到模型在分析 MRI 图像时关注的重点区域，其萎缩区域部分被模型精准的识别。

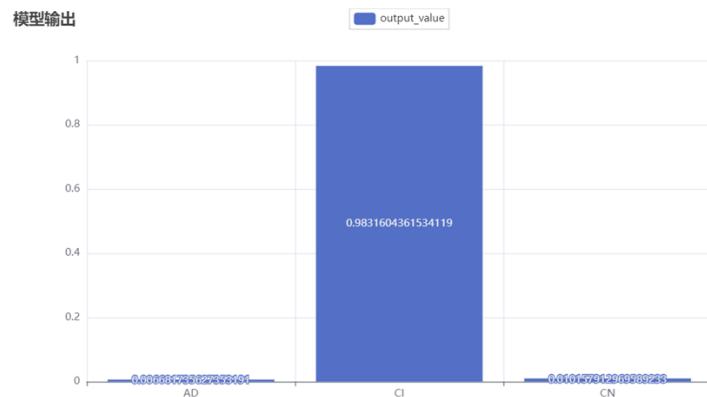


图 4-10 ACMA-CAM 预测 MCI 类型概率

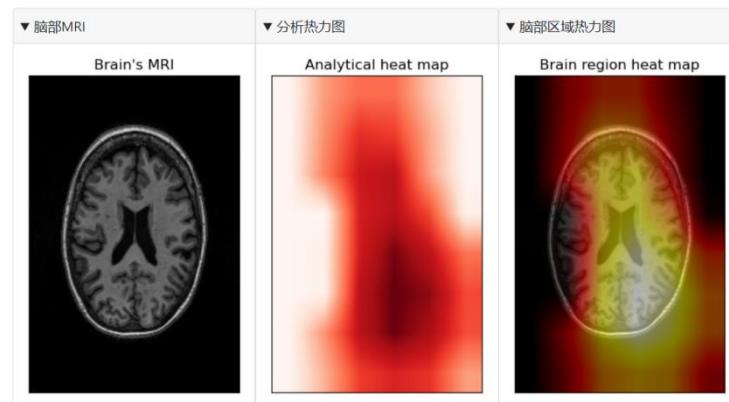


图 4-11 ACMA-CAM 输出注意力热力图

## 5 总结

### 5.1 临床意义

AD 作为一种常见的神经退行性疾病，其早期诊断对于临床干预和治疗具有重要的意义。本文研究成果展现了多方面的应用价值。

从疾病的早期发现与干预角度来看，AD 的病理进程在临床症状出现之前就早已开始，MCI 阶段作为 AD 的前期阶段，因此 MCI 阶段是进行有效干预 AD 病情发展的关键窗口期。本文中，基于多模态数据构建的 ACMCA 模型能够有效的对 AD 类型进行分类诊断。通过整合 PET 影像、MRI 影像、遗传基因数据和临床数据，对 AD 实现多模态数据联合的预测。

在利用多模态数据参与模型的训练中，如何选取有效的生物学指标数据输入到模型中训练亦是精确诊断 AD 的关键。其中，PET 影像可反映大脑特定区域的代谢变化，在 AD 早期，大脑颞叶、顶叶等区域会出现葡萄糖代谢降低的现象。MRI 影像能够清晰呈现大脑的结构特征，例如海马体萎缩等结构改变。遗传基因数据携带了个体患 AD 的遗传信息，例如载脂蛋白 E (APOE) 基因的 ε4 等位基因。临床数据则采用年龄、性别、MoCA量表分数、MMSE分数等，能够直接反映测试者的认知水平。在多种模态数据相互融合作用下，使得模型能够更全面地评估患者的患病风险，有效提高了诊断的精确度。

在医疗资源优化配置方面，传统的 AD 诊断流程中患者需要多次前往医院并接受一系列复杂的检查和评估。本文设计的智能交互式 MoCA 量表筛查系统可以实现在微信小程序端患者的自主筛查。患者在家中即可借助语音交互和图像识别功能完成 MoCA 测试，降低了医疗成本的同时，也提升了筛查效率。同时，Web 应用程序端的 ACMCA-CAM 模型，能够快速处理患者的 MRI 图像和相关检查数据，为医生提供准确的诊断建议，提高了医疗系统的运行效率。

## 5.2 未来展望

文中所涉及的阿尔茨海默病辅助筛查系统，为计算机技术在AD研究领域开拓出了新的思考路径。要想让该系统更加妥帖地在实际临床当中得以应用，往后其实还存在着诸多有待改进之处，同时也有着不少能够进一步拓展的方向。

就模型优化来讲，当下的ACMCA模型在多模态数据融合这块已然做出了改进举措，使得模型的精确程度有所提升。而随着深度学习技术持续发展，能够把更多那些经实践验证确实可提升模型性能的模块添加进ACMCA模型里。

从一方面来看，可以运用更为高效的注意力机制。比如说，动态注意力机制能够依照输入数据所具备的特点来动态地对注意力分配加以调整，以更为灵活的方式聚焦在关键信息以及细节之上；分层注意力机制则能够在不同的层次针对数据展开注意力的计算工作，从而更好地捕捉到数据的层次结构以及长短期依赖关系；还有类似Transformer的一些变体，像基于位置的自注意力机制就能够更为妥善地处理数据的位置信息，对于医学影像以及临床数据当中的特征提取有着极大的优势。通过实施这些改进措施，可以进一步增强模型针对多模态数据的学习能力，进而提升AD诊断的准确性能。

多模态数据的进一步拓展已然成为未来研究领域当中极为重要的方向。就当下的情况而言，本文所运用的数据大多是从ADNI数据库获取而来的，如此一来，便在数据规模方面呈现出了一定的局限性。在未来的研究进程当中，可以在全球范围之内广泛地去收集来自不同地区以及不同种族人群的数据，通过这样的方式来切实有效地扩大样本量，进而促使模型的泛化能力得以提升。毕竟不同种族人群在遗传背景方面、生活环境层面以及饮食习惯等诸多方面均存在着较为明显的差异，而这些差异因素很有可能会对AD的发病机制以及临床表现产生相应的影响。截止到目前为止，在中国境内还并没有相关的研究机构去建

立起关于AD的大型国产化数据集，从我国医疗行业长远发展的角度来考量，去着手建立国内首个大型AD数据集无疑是有着极为重要且不容忽视的意义的。与此同时，还能够把更多类型的生物标志物数据纳入其中，以此来为AD诊断提供有力的支持。除了现有的影像学以及基因数据之外，像血液当中的神经递质、炎症标志物等这类生物标志物同样也可以被纳入其中。

就系统完善程度来讲，交互式辅助筛查系统是能够朝着更为‘智能’的方向发展的。在微信小程序这边，提升语音识别以及图像识别的准确率，这对于模型性能的提升而言极为关键。不妨采用端到端语音识别模型，同时结合大量的语音数据展开训练工作，如此一来便能够提升针对不同口音、不同语速以及各类语言环境下语音的识别本领。在图像识别这个方面，通过引入生成对抗网络技术，是能够让模型在手绘图形以及医学影像的识别能力上得以增强的。在Web应用程序这边，除了现有的类激活图（CAM）可视化功能之外，还能够进一步去开发更多的数据可视化的方式。比如说，利用3D可视化技术来展示大脑结构，这样能够助力医生以更为直观的方式去判断患者MRI影像所发生的结构性变化。与此同时，还能够着手建立起患者医疗数据管理系统，进而实现对患者数据展开长期的跟踪以及分析工作。另外，医疗信息管理系统的安全性以及隐私保护，这在临床应用当中属于必须要去解决的重要事项。

综合来看，往后在探讨怎样改进AD辅助筛查系统时，能够从对模型加以优化、对数据予以拓展以及对系统进行完善等诸多方面来展开相关工作。可以想见，伴随计算机技术持续发展，AD辅助筛查系统肯定会变得更加完善，进而能给AD的早期诊断与治疗给予更为有力的支撑，也能为减轻社会所面临的医疗负担贡献更多的力量。

## 参考文献

- [1] 范文凭,齐金蕾,刘江美,等.2013-2021 年中国阿尔茨海默病和其他痴呆死亡水平及趋势分析[J].疾病监测,2025,40(02):245-250.
- [2] 黄文斌.阿尔茨海默病——正确防治,无问早晚[J].健康向导,2025,31(02):55-56.
- [3] Qiu S, Miller M I, Joshi P S, et al. Multimodal deep learning for Alzheimer's disease dementia assessment[J]. Nature communications, 2022, 13(1): 3404.
- [4] Veitch D P, Weiner M W, Miller M, et al. The Alzheimer's Disease Neuroimaging Initiative in the era of Alzheimer's disease treatment: A review of ADNI studies from 2021 to 2022[J]. Alzheimer's & Dementia, 2024, 20(1): 652-694.
- [5] Bettens K, Sleegers K, Van Broeckhoven C. Genetic insights in Alzheimer's disease[J]. The lancet neurology, 2013, 12(1): 92-104.
- [6] Cuingnet R, Gerardin E, Tessieras J, et al. Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database[J]. neuroimage, 2011, 56(2): 766-781.
- [7] 陈金雨,张肖南,武志芳,等.基于 ADNI 中 PET/CT 的阿尔茨海默病致病机制及诊断的最新研究[J].神经损伤与功能重建,2024,19(11):669-672.DOI:10.16780/j.cnki.sjssgncj.20230615.
- [8] Jack Jr C R, Bernstein M A, Fox N C, et al. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods[J]. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine, 2008, 27(4): 685-691.
- [9] Zhang J, Zheng B, Gao A, et al. A 3D densely connected convolution neural network with connection-wise attention mechanism for Alzheimer's disease classification[J]. Magnetic Resonance Imaging, 2021, 78: 119-126.
- [10] Basheera S, Ram M S S. Convolution neural network-based Alzheimer's disease classification using hybrid enhanced independent component analysis based segmented gray matter of T2 weighted magnetic resonance imaging with clinical valuation[J]. Alzheimer's & Dementia: Translational Research & Clinical Interventions, 2019, 5: 974-986.

- [11]SWAMY K R, SENTHILKUMAR S. DENSE INCEPTION ATTENTION NEURAL NETWORK FOR ALZHEIMER'S DISEASE CLASSIFICATION USING MAGNETIC RESONANCE IMAGES[J]. Journal of Theoretical and Applied Information Technology, 2023, 101(22).
- [12]Zhang J, He X, Liu Y, et al. Multi-modal cross-attention network for Alzheimer's disease diagnosis with multi-modality data[J]. Computers in Biology and Medicine, 2023, 162: 107050.
- [13]Chen H, Guo H, Xing L, et al. Multimodal predictive classification of Alzheimer's disease based on attention - combined fusion network: Integrated neuroimaging modalities and medical examination data[J]. IET Image Processing, 2023, 17(11): 3153-3164.
- [14]Patro B N, Namboodiri V P, Agneeswaran V S. Spectformer: Frequency and attention is what you need in a vision transformer[C]//2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). IEEE, 2025: 9543-9554.
- [15]冒春燕,房荣华,丛辉.阿尔茨海默病的发病机制及诊疗的研究进展[J].江苏医药,2025,51(03):300-304.DOI:10.19460/j.cnki.0253-3685.2025.03.021.
- [16]Wood Alexander M, Honer W G, Saloner R, et al. The interplay between age at menopause and synaptic integrity on Alzheimer's disease risk in women[J]. Science Advances, 2025, 11(10): ead0757.
- [17]Arafa D A, Moustafa H E D, Ali H A, et al. A deep learning framework for early diagnosis of Alzheimer's disease on MRI images[J]. Multimedia Tools and Applications, 2024, 83(2): 3767-3799.
- [18]Fruhwürth S, Zetterberg H, Paludan S R. Microglia and amyloid plaque formation in Alzheimer's disease—Evidence, possible mechanisms, and future challenges[J]. Journal of Neuroimmunology, 2024: 578342.
- [19]Jain U, Johari S, Srivastava P. Current insights of Nanocarrier-mediated gene therapeutics to treat potential impairment of amyloid Beta protein and tau protein in Alzheimer's disease[J]. Molecular Neurobiology, 2024, 61(4): 1969-1989.

- [20] Venugopalan J, Tong L, Hassanzadeh H R, et al. Multimodal deep learning models for early detection of Alzheimer's disease stage[J]. *Scientific reports*, 2021, 11(1): 3254.
- [21] Golovanevsky M, Eickhoff C, Singh R. Multimodal attention-based deep learning for Alzheimer's disease diagnosis[J]. *Journal of the American Medical Informatics Association*, 2022, 29(12): 2014-2022.
- [22] Meikandan P V. Global Challenges in Accessing Mental Health Services and Addressing the Impact of Alzheimer's Disease and Depression[D]. Marquette University, 2024.
- [23] 任汝静,殷鹏,王志会,等. 中国阿尔茨海默病报告 2021[J]. 诊断学理论与实践,2021,20(4):317-337. DOI:10.16150/j.1671-2870.2021.04.001.
- [24] 邢思雨,鞠彦秀,马云影,等.蒙特利尔认知评估量表与老年阿尔茨海默病患者焦虑状态的相关性[J].中国老年学杂志,2024,44(20):4953-4956.
- [25] 宁波市康复医院.一种基于图像识别技术的 AD 症状识别方法与系统:202410857049.2[P].2024-10-29.

## 致谢

行文至此，落笔为终。时光荏苒，岁月如梭，大学四年的光阴转瞬即逝，我的本科生涯即将随着这篇学士毕业论文的交付而终结。仅以此段文字表达对大学时光的亲切留念。

时光清浅处，一步一安然。那个十七岁夏天的蝉鸣，比哪一年都聒噪。犹记得初入校园的晚训时光——暮色浸染操场时，我躺在草坪上看银河碎钻般洒落，隔壁班教官“一二一，正步走”的口令突然穿透晚风，惊起我手忙脚乱戴帽子归队的慌张。那个带着青草香的夏夜，成了我与广药大故事的温柔序章。

专业探索的旅程并非坦途。大一、大二时，我曾在计算机专业课的迷宫里反复徘徊：晦涩的代码像未解的密码，算法逻辑如缠绕的线团，常常让我在深夜图书馆的台灯下凝眉思索。幸而遇见李沐、李宏毅、吴恩达等老师的网络课程，他们以深入浅出的讲解为我推开 AI 世界的窗——那些关于深度学习的精妙阐释、机器学习的生动案例，如同暗夜中的灯塔，指引我在代码的海洋里锚定方向。当第一个独立设计的项目终于运行成功，键盘敲击声里藏着多少个破晓前的坚持，此刻都化作了眼底的星光。

在这里我要特别感谢本文的指导老师周苏娟教授，周苏娟教授是我学术道路上的引路人。“经师易得，人师难求”，她不仅在论文写作中给予指导，更以大创项目为舟，载我驶向东南大学中国大学生计算机设计大赛的赛场；她用渊博学识与温和鼓励，让我懂得科研不仅是代码的堆砌，更是对真理的赤诚追寻与坚持探索。感谢我的父母，含辛茹苦的将我抚养成人。最后感谢同学们四年以来的陪伴。

不问花开几许，只向浅笑安然。站在毕业的渡口回望，广药大的操场仍回荡着蝉鸣与口号，图书馆的落地窗依然收藏着日出与星辰。愿未来的自己，如

四年前那个追着星光奔跑的少年，永远保持对代码的热忱、对世界的好奇；更愿能在人生的急流与深潭间，守住内心的浅笑安然——不为花期所困，只为热爱奔赴，永远拥有在细碎时光里采摘幸福的能力。