

# 多模态深度学习中的决策级融合方法及其应用研究

田傲翔, 张长伦

北京建筑大学理学院, 北京

收稿日期: 2025年8月24日; 录用日期: 2025年9月14日; 发布日期: 2025年9月25日

## 摘要

单模态深度学习强大的学习能力已经在诸多应用领域取得了优异的成效。然而, 单模态深度学习不能学习到某一现象完整的信息。为了缓解这一问题, 多模态深度学习被提出并受到广泛关注。多模态深度学习旨在构建能够关联多个模态信息的模型。本文首先介绍多模态深度学习技术并引出多模态决策级融合, 然后对多模态决策级融合中的决策算法进行详细介绍, 并介绍决策算法在多模态深度学习领域中的应用现状, 最后对多模态决策级融合进行总结和展望。

## 关键词

多模态, 深度学习, 决策级融合, 决策融合算法

# Research on Decision-Level Fusion Method and Application in Multimodal Deep Learning

Aoxiang Tian, Changlun Zhang

School of Science, Beijing University of Civil Engineering and Architecture, Beijing

Received: August 24, 2025; accepted: September 14, 2025; published: September 25, 2025

## Abstract

Deep learning has achieved excellent results in the field of single-modal applications due to its powerful learning ability. However, single-modal deep learning cannot learn the complete information of a phenomenon. In order to alleviate this problem, multimodal deep learning has been proposed and received widespread attention. Multimodal deep learning aims to build models that can correlate

文章引用: 田傲翔, 张长伦. 多模态深度学习中的决策级融合方法及其应用研究[J]. 图像与信号处理, 2025, 14(4): 412-421. DOI: 10.12677/jisp.2025.144038

multiple modal information. This paper first introduces multi-modal deep learning technology and multi-modal decision-level fusion. Then, the decision-making algorithm in multi-modal decision-level fusion is introduced in detail, and the application status of decision-making algorithm in multi-modal deep learning is introduced. Finally, the multi-modal decision-level fusion is summarized and prospected.

## Keywords

Multimodal, Deep Learning, Decision-Level Fusion, Decision Fusion Algorithm

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

在数据领域, 模态用来表示数据形式, 多模态用来表示不同数据的不同形式或同一数据的不同格式, 包括文本、图像、音频、视频等[1]。单模态深度学习在文本、图像、音视频处理领域取得了重大进展, 但是单模态深度学习不能学习到某一现象的完整信息。多模态深度学习旨在让机器能从文本、图像、音视频数据中提取更加完善的表征, 并利用模态之间的相关性和互补性, 完成对信息的融合, 达到更好的学习效果。近年来, 随着多模态深度学习的发展, 已经在音视频生成[2][3]、自动驾驶[4]、机器人[5][6]、智能医疗[7][8]等诸多领域取得了重大进展。

多模态融合是多模态学习的核心技术, 通过将多个模态数据的信息进行整合, 并通过多模态深度学习模型对整合的信息进行学习、训练, 可以充分利用模态之间的互补性, 减少模态之间的信息冗余, 完成多种复杂的任务。多模态融合方法可以分为三类: 特征级融合、决策级融合和混合融合[9], 如图 1 所示。特征级融合集成各模态的特征表示, 训练一个共同的模型完成任务; 决策级融合训练多个独立模型, 每种模态输出结果后才进行融合; 混合融合尝试在一个框架内利用特征级融合和决策级融合两种方法的优点。

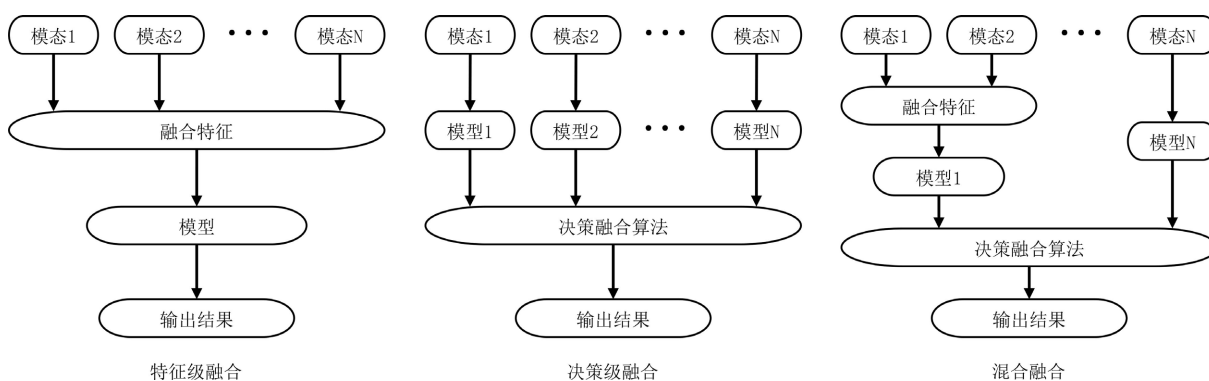


Figure 1. Multimodal fusion

图 1. 多模态融合

多模态决策级融合方法也称晚期融合方法, 针对不同模态的数据, 使用不同的深度学习模型进行训练, 多个深度学习模型输出结果之后, 使用决策融合算法将多个输出结果进行决策融合。决策级融合的

融合过程与特征无关, 而且使用不同深度学习模型对不同模态数据进行单独训练, 使得多个输出结果的错误通常是不相关的, 可以更好地利用数据之间的互补性, 剔除冗余性, 因此, 多模态决策融合方法普遍受到关注。

多模态决策融合方法的核心就在于决策融合算法的选取。本文重点讨论决策融合算法, 全面深入研究常用的多模态决策级融合算法的核心概念, 并详细介绍决策融合算法的计算原理。然后介绍决策融合算法在当前多模态深度学习领域中的应用现状, 并总结多模态决策级融合面临的挑战以及未来的发展方向。

## 2. 多模态决策融合算法

目前, 常用的决策融合算法主要有四种: 加权平均法、贝叶斯推理融合、基于证据理论的方法和基于集成学习的方法。

### 2.1. 加权平均法

加权平均法[10]是最简单的决策融合算法, 考虑不同模型的能力不同, 对最终结果的贡献也有差异, 因此将不同模态的输出结果按照权重进行加权平均, 得到一个综合决策结果, 加权平均法计算公式为:

$$P = \sum_i \omega_i \cdot P_i \quad (1)$$

其中  $P_i$  为每个独立模型的输出结果;  $\omega_i$  是分配给的权重, 并且  $\sum_i \omega_i = 1$ ;  $i$  是模型数量;  $P$  为综合决策结果。

### 2.2. 贝叶斯推理融合

贝叶斯推理融合算法是基于贝叶斯理论的决策融合算法, 最早由英国数学家 Thomas Bayes 提出[11], 其主要思想是通过测得的数据对先验概率进行更新而得到后验概率。贝叶斯推理融合的基础为条件概率公式和全概率公式[12] [13]。条件概率公式为:

$$P(A|B) = \frac{P(AB)}{P(B)} \quad (2)$$

全概率公式为:

$$P(B) = P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A}) \quad (3)$$

由条件概率公式和全概率公式可推导得到贝叶斯公式为:

$$P(A|B) = \frac{P(A) \cdot P(B|A)}{P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A})} \quad (4)$$

其中,  $P(A|B)$  称为后验概率;  $P(A)$  称为先验概率, 根据经验和已有信息对  $A$  发生概率的估计;  $\frac{P(B|A)}{P(B)}$  称为标准似然度。

对于多模态决策融合任务, 假设多模态融合任务目标框架为  $H = \{H_1, H_2, \dots, H_r\}$ , 识别框架中的子元素互斥不相交且目标识别框架为完备样本空间, 多个深度学习模型的输出信息为  $D_1, D_2, \dots, D_N$ , 并且由于不同深度学习模型的输出信息满足条件独立, 根据贝叶斯定理,  $H_i (H_i \in H)$  的后验概率为:

$$P(H_i|D_1, D_2, \dots, D_N) = a_b P(H_i) \prod_{j=1}^N P(D_j|H_i) \quad (5)$$

其中,  $a_b = \frac{1}{\sum_{m=1}^r \left( \prod_{j=1}^N P(D_j | H_i) \right)}$ 。

贝叶斯推理融合算法在多模态决策融合中的应用过程如图2所示。

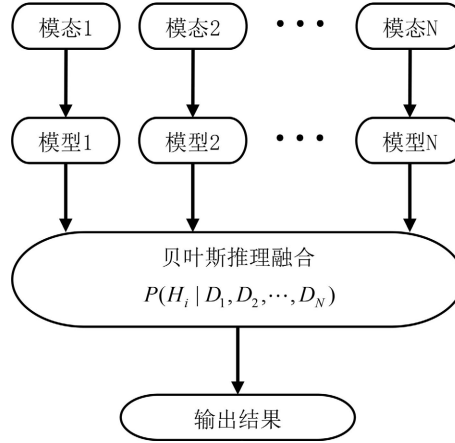


Figure 2. Bayesian inference fusion process  
图2. 贝叶斯推理融合过程

### 2.3. 证据理论

证据理论又称 Dempster-Shafer 理论(D-S), 是一种处理不确定性信息的有效方法, 在多个信息源提供证据时, 可以通过证据理论组合规则合成数据, 提高决策的准确性[14]。

设多模态融合任务目标为非空集合的样本空间, 样本空间内元素互斥不相交, 则称  $H$  为识别框架。幂集定义为  $2^H$ , 是识别框架  $H$  所有子集的集合。对于  $2^H$  中的任意命题  $h$ , 定义函数  $D(h): 2^H \rightarrow [0,1]$ ,  $D(h)$ 满足以下条件:  $H = \{H_1, H_2, \dots, H_r\}$

$$\begin{cases} 0 \leq D(h) \leq 1 \\ D(\emptyset) = 0 \\ \sum_{h \subseteq H} D(h) = 1 \end{cases} \quad (6)$$

其中, 用  $D(h)$ 表示为识别框架  $H$  上的基本概率分配(BPA), 反映了证据对命题  $h$  的信任程度。  $\emptyset$  为空集,  $D(\emptyset)=0$  反映了证据对空间没有任何信任度。  $\sum_{h \subseteq H} D(h)=1$  表示识别框架所有子集的信任度总和为1。

定义函数  $Bel: 2^H \rightarrow [0,1]$  为识别框架  $H$  下的信任函数,  $Bel(h)$ 表示对命题  $h$  为真的信任总和, 是信任区间的下限函数,  $Bel(h)$ 具有以下性质:

$$\begin{cases} Bel(\emptyset) = D(\emptyset) = 0 \\ Bel(h) = \sum_{g \subseteq h} D(g) \end{cases} \quad (7)$$

定义函数  $pl: 2^H \rightarrow [0,1]$  为识别框架  $H$  下的似然函数,  $pl(h)$ 表示对命题  $h$  的非假的信任总和, 是信任区间的上限函数,  $pl(h)$ 满足:

$$pl(h) = \sum_{g \cap h = \emptyset} D(g) = 1 - Bel(\bar{h}), \forall h \subseteq \Theta \quad (8)$$

则命题  $h$  的不确定区间为  $[Bel(h), pl(h)]$ , 用  $pl(h)-Bel(h)$  表示命题  $h$  的不确定度, 即无法完全信任部分。

由于多模态数据来源的多样性, 相同的证据可能产生多个不同的基本概率分配函数, 导致信息之间的不一致性, 为了整合这些不一致的信息, Dempster-Shafer 提出了一种合成规则, 允许将多个概率分配函数进行合并。这种合成方法可以在不考虑数据源的前提下将它们合并统一。Dempster-Shafer 合成规则表达式为:

$$(D_1 \oplus D_2 \oplus \cdots \oplus D_N)(h) = \frac{1}{K} \sum_{D_1 \cap \cdots \cap D_N} D_1(h_1) \cdots D_N(h_N) \quad (9)$$

其中  $K$  为归一化常数, 取值为:

$$K = \sum_{D_1 \cap \cdots \cap D_N \neq \emptyset} D_1(h_1) \cdots D_N(h_N) = 1 - \sum_{D_1 \cap \cdots \cap D_N = \emptyset} D_1(h_1) \cdots D_N(h_N) \quad (10)$$

在多模态决策级融合任务, 证据理论的应用过程如图 3 所示。

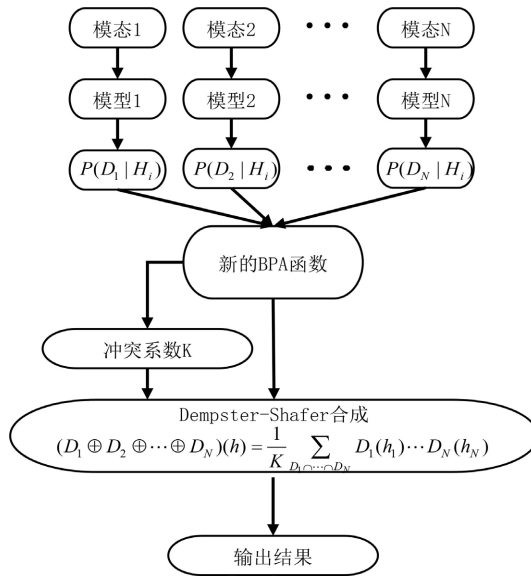


Figure 3. D-S theory fusion process  
图 3. 证据理论融合过程

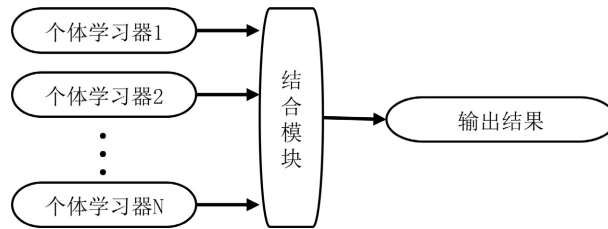
## 2.4. 集成学习

集成学习方法是一种独特的机器学习算法, 通过构建并结合多个机器学习来完成学习任务[15]。训练若干个个体学习器, 通过一定的结合策略, 就可以最终形成一个强学习器, 以达到博采众长的目的, 集成学习示意图如图 4 所示。目前常见的集成学习算法大致可分为两类: 个体学习器间不存在强依赖关系、可同时生成的并行化方法, 代表算法为基于 Bagging 的算法; 个体学习器间存在强依赖关系, 必须串行生成的序列化方法, 代表算法为基于 Boosting 的算法[16]。

### 2.4.1. 基于 Bagging 的算法

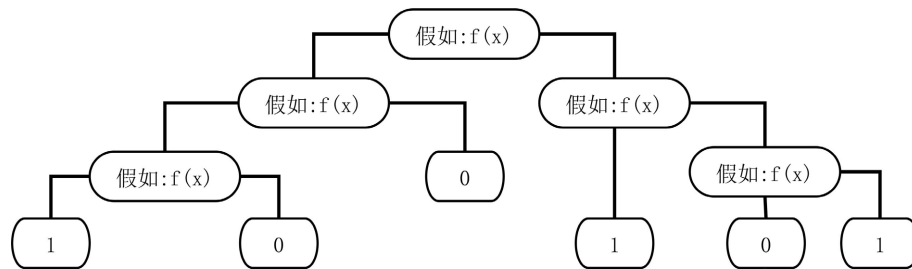
Bagging 算法是并行式集成学习方法最著名的代表, 其基本思想基于自助采样法, 给定包含  $m$  个样本的数据集, 随机取出一个样本放入采样集中, 再把该样本放回初始数据集中, 使得下次采样时该样本仍有可能被选中, 这样, 可采样出  $T$  个含  $m$  个训练样本的采样集, 然后基于每个采样集训练出一个基学

习器, 再将这些基学习器进行结合。



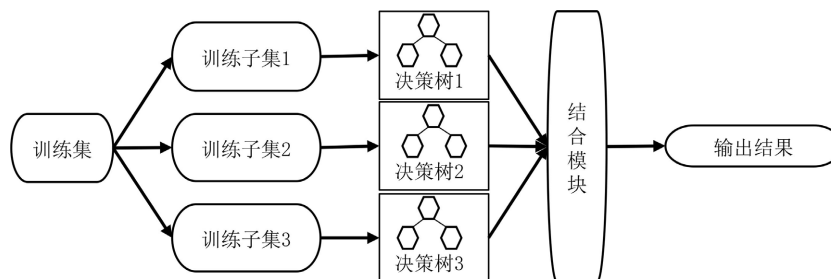
**Figure 4.** Ensemble learning  
**图 4.** 集成学习

随机森林[17]是最常用的基于 Bagging 算法的集成学习算法, 随机森林的模型由多棵决策树组成, 树型结构模型因为组织形式简单, 一般具有可解释性。树型结构示意图如图 5 所示。



**Figure 5.** Tree structure model  
**图 5.** 树型结构模型

随机森林的构建开始于从原始数据集中随机抽取的多个子集, 每个子集用来训练一个决策树。在训练每棵树的过程中, 随机森林不使用所有特征作为分裂节点的特征, 而是从所有可用的特征中随机选择一个子集, 每棵树独立成长到最大化。最后, 所有决策树的预测结构通过结合模块进行聚合, 形成最终的模型预测。随机森林算法示意图如图 6 所示。



**Figure 6.** Random forest algorithm  
**图 6.** 随机森林算法图

### 2.4.2. 基于 Boosting 的算法

Boosting 是一种将弱学习器提升为强学习器的算法。AdaBoost 算法[18]是基于 Boosting 算法中最具代表性的算法, 其核心思想是针对同一个训练集训练不同的弱学习器, 然后把这些弱学习器集合, 构成一个强学习器。首先初始化样本的权值分布, 每个样本具有相同的权重; 然后训练弱学习器, 如果样本

分类正确, 则在下一个训练集中, 它的权值就会被降低, 反之提高, 用更新后的样本集去训练下一个学习器; 最后, 将所有弱学习器组合成强学习器, 各个弱学习器的训练过程结束后, 加大分类误差率小的弱学习器的权重, 降低分类误差率大的弱学习器的权重。AdaBoost 算法示意图如图 7 所示。

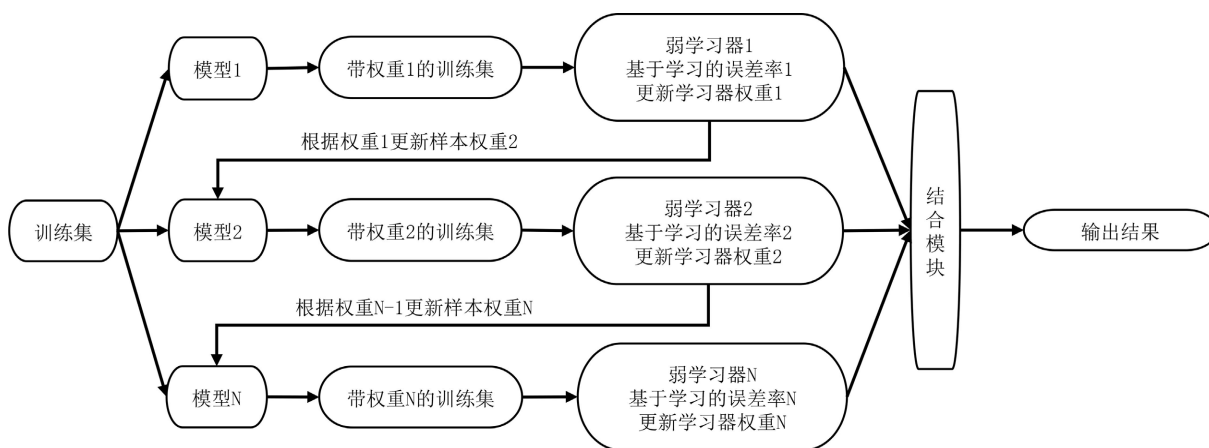


Figure 7. AdaBoost algorithm

图 7. AdaBoost 算法

### 3. 多模态决策融合的应用

随着多模态深度学习的发展, 多模态决策级融合开始被应用到多种研究领域, 典型的四种决策融合算法被广泛应用到决策级融合中, 从而克服单一模态深度学习模型的缺点。

#### 3.1. 加权平均法

加权平均法算法简单清晰, 应用较为广泛。在 MER2023 大会上, Zong 等人[19]提出了一种数据增强和加权监督信号融合的多模态训练策略, 巧妙地应对了单一模型训练存在的不足, 实现了多模态情绪分析。Wang 等人[20]提出一种自适应模态加权多模态信息融合方法, 根据每种模态对情感识别的贡献为其分配适当的权重, 更有效地实现了多模态信息的决策融合。Lu 等人[21]开发了一种阿尔兹海默症多模态融合预测框架, 通过注意力机制学习每种模态适当的权重, 并通过分层注意力机制理解模态之间的相互作用, 完成多模态数据之间的决策融合。

加权平均法可以充分利用不同模态在不同任务中的优势, 根据对任务做出的不同贡献赋予不同的权重, 可以有效地提高任务完成率。但是加权平均法是对多模态原始数据进行线性叠加, 无法有效处理模态的异构性与数据冲突。

#### 3.2. 贝叶斯推理融合

贝叶斯推理融合算法可以有效整合多源异构数据, 通过概率分布显示建模多模态数据的异构性, 同时后验分布保留了加权平均法中可以追溯各模态贡献度的优点。

Xie 等人[22]使用贝叶斯融合对单模态深度学习模型学习到的结果进行决策融合, 以获得用于情感预测的多模态分布。Wang 等人[23]为了组合和目标响应相关的冗余信息源, 并考虑多模态数据的相关质量, 同时适应不同数据的不同采样率, 提出一种正则化贝叶斯融合, 用于生成具有预期平滑度水平的目标变量预估值, 可以兼容多种信息源。Zhou [24]等人使用脑电图和心电图信号数据构建原始生理数据集, 通过贝叶斯融合进行疲劳状态识别, 实现了不同格式数据之间的多模态决策融合。郭家星[25]提出一种基于

贝叶斯融合对 SHM 系统多通道不统一检测结果数据进行融合的方法, 排除了故障数据和随机数据的干扰, 有效整合了多通道不统一数据。

但是在贝叶斯推理融合算法中, 错误的先验会导致后验失真, 并且当先验知识不足时, 会导致后验退化为无信息先验, 失去决策价值。同时贝叶斯推理融合算法需要多模态数据严格时空对齐。

### 3.3. 证据理论

相较于贝叶斯推理融合, 证据理论决策融合算法通过基本概率函数(BPA)将信任度分配给样本空间的子集, 这一操作可以表达不确定性, 并且独特的 Dempster-Shafer 合成规则可以抵制似然度证据的影响, 可以有效使多源异构数据完成多模态决策融合。

耿彦涛[26]使用 Hellinger 距离和 Shannon 熵衡量冲突程度和信息量, 以对 BPA 函数进行加权处理, 然后使用 Dempster-Shafer 合成规则进行融合处理, 实现了音频和图像数据间的互相描述。Yang 等人[27]考虑机器算法和人类智能之间的互补性, 提出一种基于证据理论的融合策略, 对机器算法和人类智能对文本和图像的评论进行多模态融合。Wang 等人[28]提出了一种基于 Dempster-Shafer 证据理论的多模态推荐算法, 根据 Dempster-Shafer 证据理论, 将用户对不同模态的非交互式项目的偏好视为证据, 将这些证据于不同的模态偏好融合在一起, 得到了更好的推荐结果。Huang 等人[29]提出深度证据融合框架, 利用 Dempster-Shafer 证据理论和深度神经网络分割多模态医学图像, 通过上下文信息, 量化每个信息源对每个类的信任度来校正基本概率分配函数, 然后使用 Dempster-Shafer 合成规则将每个信息源的证据整合。

但证据理论的计算复杂度为  $O(2^n)$ , 难以满足实时需求; 在证据冲突系数较大时, Dempster-Shafer 合成规则可能产生反直觉的结果, 并且基本概率函数(BPA)的构造主观性强, 容易引入人为误差。

### 3.4. 集成学习

集成学习方法是多模态决策级融合核心思想的体现, 不同基学习器在各自模态中具有显著的异构性, 可以很好的保持多模态数据的独有特征, 并且支持多层级的融合策略。

Li [30]等人通过 ELM 网络对现有的诗歌翻译进行学习, 得到多个弱预测器, 然后使用 AdaBoost 进行分类迭代, 得到强预测器, 利用决策融合实现了多模态翻译。Susmita Palmal 等人[31]提出一种具有校准随机森林的多模态图卷积网络。使用图卷积神经网络从多模态数据中提取到特征, 将级联起来的特征使用随机森林进行最终预测, 实现了对乳腺癌患者术后生存期的精准预测。Varun 等人[32]采用两个独立模型完成语音情感识别和面部情感识别, 在决策层使用随机森林算法将两个模型的结果结合, 实现了多模态决策融合情感识别。Khandakar [33]等人提出一种具有交叉权重采样方法的多源 AdaBoost, 在更新抽样权重分布时, 考虑多源数据的影响, 以重新对高质量观测值进行抽样, 实现了多传感器虚拟计量的决策融合。

集成学习中多个并行的训练模型, 会导致计算资源消耗成倍增长。同时, 多个模型的集成难以追溯各模态的贡献。

## 4. 研究展望

多模态决策级融合是多模态深度学习领域的研究热点之一, 建模对象不再是原始数据或特征, 而是来自多个单模态模型的输出结果。随着多模态决策级融合的发展, 多模态任务愈发复杂, 单模态模型在复杂任务中无法保持高精度检测性能, 容易出现检测错误。单模态模型的检测错误通过决策融合机制, 传递给决策结果, 导致决策结果失真。如何避免单模态模型检测错误的干扰, 提高多模态决策级融合的任务性能, 是具有挑战和有价值的研究方向。

### 1) 多模态信息交叉反馈修正单模态模型决策

多模态决策级融合对单模态模型输出结果进行建模, 忽略了单模态模型在训练过程中学习的信息, 导致单模态模型检测结果中的错误也被当作建模对象。不同的单模态模型在训练过程中会学习到不同的有效信息, 多种模态信息的互相反馈可以提高单模态模型的检测性能。比如将视觉模态的语义特征, 反馈给时序模型的隐藏层, 可以有效修正时序模型的检测错误。因此如何使用单模态模型学习到的信息, 进行信息交叉反馈, 使得单模态模型检测结果的错误得到修正, 是具有挑战和价值的研究方向。

### 2) 单模态模型输出结果互相验证

尽管证据理论通过基本概率分配对单模态模型输出结果的不确定性进行建模, 但多种输出结果互相独立, 使得不确定性仍然通过合成规则传递给了决策结果。如何在证据理论的基础上, 创新决策融合机制, 使得多种单模态模型的输出结果可以互相验证, 并且根据上下文信息, 自适应调整概率分配函数, 从而避免单模态模型输出结果检测错误的传递, 干扰决策结果, 是值得深入探索的问题。

## 参考文献

- [1] 任泽裕, 王振超, 柯尊旺, 等. 多模态数据融合综述[J]. 计算机工程与应用, 2021, 57(18): 49-64.
- [2] Zhang, D.J., Li, D., Le, H., Shou, M.Z., Xiong, C. and Sahoo, D. (2025) Moonshot: Towards Controllable Video Generation and Editing with Motion-Aware Multimodal Conditions. *International Journal of Computer Vision*, **133**, 3629-3644. <https://doi.org/10.1007/s11263-025-02346-1>
- [3] Marmon, A., Schindler, G., Lezama, J., Kondratyuk, D., Seybold, B. and Essa, I. (2024) CamVIG: Camera Aware Image-to-Video Generation with Multimodal Transformers. arXiv: 2405.13195.
- [4] Hwang, J. J., Xu, R., Lin, H., Hung, W. C., Ji, J., Choi, K., Tan, M., *et al.* (2024) Emma: End-to-end Multimodal Model for Autonomous Driving. arXiv: 2410.23262.
- [5] Wang, T., Zheng, P., Li, S. and Wang, L. (2023) Multimodal Human-Robot Interaction for Human-centric Smart Manufacturing: A Survey. *Advanced Intelligent Systems*, **6**, Article ID: 2300359. <https://doi.org/10.1002/aisy.202300359>
- [6] Yue, Y., Wang, Y., Kang, B., Han, Y., Wang, S., Song, S., Huang, G., *et al.* (2024) DeeR-VLA: Dynamic Inference of Multimodal Large Language Models for Efficient Robot Execution. arXiv: 2411.02359.
- [7] Karbout, K., El Ghazouani, M., Lachgar, M. and Hrimech, H. (2024) Multimodal Data Fusion Techniques in Smart Healthcare. 2024 *International Conference on Global Aeronautical Engineering and Satellite Technology (GAST)*, Marrakesh, 24-26 April 2024, 1-6. <https://doi.org/10.1109/gast60528.2024.10520803>
- [8] Krones, F., Marikkar, U., Parsons, G., Szmul, A. and Mahdi, A. (2025) Review of Multimodal Machine Learning Approaches in Healthcare. *Information Fusion*, **114**, Article ID: 102690. <https://doi.org/10.1016/j.inffus.2024.102690>
- [9] 何俊, 张彩庆, 李小珍, 等. 面向深度学习的多模态融合技术研究综述[J]. 计算机工程, 2020, 46(5): 1-11.
- [10] 费业泰. 误差理论与数据处理[M]. 北京: 机械工业出版社, 2004.
- [11] Dale, A.I. (2003) *Most Honourable Remembrance: The Life and Work of Thomas Bayes*. Springer Science & Business Media.
- [12] Klein, L.A. (2004) *Sensor and Data Fusion: A Tool for Information Assessment and Decision Making*. SPIE Press. <https://doi.org/10.1117/3.563340>
- [13] Vanik, M.W., Beck, J.L. and Au, S.K. (2000) Bayesian Probabilistic Approach to Structural Health Monitoring. *Journal of Engineering Mechanics*, **126**, 738-745. [https://doi.org/10.1061/\(asce\)0733-9399\(2000\)126:7\(738\)](https://doi.org/10.1061/(asce)0733-9399(2000)126:7(738))
- [14] 郑柳姜. 基于 D-S 证据理论的船舶碰撞风险评估[D]: [硕士学位论文]. 大连: 大连海事大学, 2024.
- [15] 竺承远. 基于集成学习的在线支付欺诈检测研究[D]: [硕士学位论文]. 海口: 海南师范大学, 2024.
- [16] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.
- [17] Breiman, L. (2001) Random Forests. *Machine Learning*, **45**, 5-32. <https://doi.org/10.1023/a:1010933404324>
- [18] Schapire, R.E. (2013) Explaining AdaBoost. In: Schölkopf, B., Luo, Z. and Vovk, V., Eds., *Empirical Inference*, Springer Berlin Heidelberg, 37-52. [https://doi.org/10.1007/978-3-642-41136-6\\_5](https://doi.org/10.1007/978-3-642-41136-6_5)
- [19] Zong, D., Ding, C., Li, B., Zhou, D., Li, J., Zheng, K., *et al.* (2023) Building Robust Multimodal Sentiment Recognition via a Simple Yet Effective Multimodal Transformer. *Proceedings of the 31st ACM International Conference on Multimedia*, Ottawa, 29 October-3 November 2023, 9596-9600. <https://doi.org/10.1145/3581783.3612872>

- [20] Wang, Y., He, J., Wang, D., Wang, Q., Wan, B. and Luo, X. (2024) Multimodal Transformer with Adaptive Modality Weighting for Multimodal Sentiment Analysis. *Neurocomputing*, **572**, Article ID: 127181. <https://doi.org/10.1016/j.neucom.2023.127181>
- [21] Lu, P., Hu, L., Mitelpunkt, A., Bhatnagar, S., Lu, L. and Liang, H. (2024) A Hierarchical Attention-Based Multimodal Fusion Framework for Predicting the Progression of Alzheimer's Disease. *Biomedical Signal Processing and Control*, **88**, Article ID: 105669. <https://doi.org/10.1016/j.bspc.2023.105669>
- [22] Xie, Z., Yang, Y., Wang, J., Liu, X. and Li, X. (2024) Trustworthy Multimodal Fusion for Sentiment Analysis in Ordinal Sentiment Space. *IEEE Transactions on Circuits and Systems for Video Technology*, **34**, 7657-7670. <https://doi.org/10.1109/tcsvt.2024.3376564>
- [23] Strelet, E., Wang, Z., Peng, Y., Castillo, I., Rendall, R. and Reis, M.S. (2024) Regularized Bayesian Fusion for Multimodal Data Integration in Industrial Processes. *Industrial & Engineering Chemistry Research*, **63**, 20989-21000. <https://doi.org/10.1021/acs.iecr.4c02956>
- [24] Zhou, Y., Chen, P., Fan, Y. and Wu, Y. (2024) A Multimodal Feature Fusion Brain Fatigue Recognition System Based on Bayes-gcForest. *Sensors*, **24**, Article 2910. <https://doi.org/10.3390/s24092910>
- [25] 郭家星. 基于深度学习与贝叶斯融合的结构健康监测地震事件可靠识别[D]: [硕士学位论文]. 昆明: 昆明理工大学, 2024.
- [26] 耿彦涛. 基于深度学习与 D-S 证据理论的多模态数据融合方法研究[D]: [硕士学位论文]. 西安: 西安理工大学, 2023.
- [27] Yang, Y., Zhang, L., Xu, G., Ren, G. and Wang, G. (2024) An Evidence-Based Multimodal Fusion Approach for Predicting Review Helpfulness with Human-AI Complementarity. *Expert Systems with Applications*, **238**, Article ID: 121878. <https://doi.org/10.1016/j.eswa.2023.121878>
- [28] Wang, X. and Qin, J. (2023) Multimodal Recommendation Algorithm Based on Dempster-Shafer Evidence Theory. *Multimedia Tools and Applications*, **83**, 28689-28704. <https://doi.org/10.1007/s11042-023-15262-8>
- [29] Huang, L., Ruan, S., Decazes, P. and Deneux, T. (2025) Deep Evidential Fusion with Uncertainty Quantification and Reliability Learning for Multimodal Medical Image Segmentation. *Information Fusion*, **113**, Article ID: 102648. <https://doi.org/10.1016/j.inffus.2024.102648>
- [30] Li, S. (2023) Construction of a Multimodal Poetry Translation Corpus Based on AdaBoost Model. *Applied Mathematics and Nonlinear Sciences*, **9**. <https://doi.org/10.2478/amns.2023.1.00102>
- [31] Palmal, S., Arya, N., Saha, S. and Tripathy, S. (2024) Integrative Prognostic Modeling for Breast Cancer: Unveiling Optimal Multimodal Combinations Using Graph Convolutional Networks and Calibrated Random Forest. *Applied Soft Computing*, **154**, Article ID: 111379. <https://doi.org/10.1016/j.asoc.2024.111379>
- [32] Avabratha, V.V., Rana, S., Narayan, S., Raju, S.Y. and S, S. (2024) Speech and Facial Emotion Recognition Using Convolutional Neural Network and Random Forest: A Multimodal Analysis. 2024 *Asia Pacific Conference on Innovation in Technology (APCIT)*, Mysore, 26-27 July 2024, 1-5. <https://doi.org/10.1109/apcit62007.2024.10673495>
- [33] Ahmed, K.R., Hossain, A., Asif, M.A.B., Mohammad, M., Rahaman, M. and Dewan, M.A. (2025) Optimizing Production and Inventory Management for Defective Items in Manufacturing Systems. 2025 *3rd International Conference on Intelligent Systems, Advanced Computing and Communication (ISACC)*, Silchar, 27-28 February 2025, 86-91. <https://doi.org/10.1109/isacc65211.2025.10969370>