

中图分类号: TP391

单位代号: 10280

密 级: 公开

学 号: 22721545

上海大学



硕士学位论文

SHANGHAI UNIVERSITY
MASTER'S DISSERTATION

题 目	面向功能材料数据的特征处理与 预测方法研究
--------	--------------------------

作 者 包胜奇

学科专业 计算机应用技术

导 师 韩越兴

完成日期 二〇二五年四月

姓名：包胜奇

学号：22721545

论文题目：面向功能材料数据的特征处理与预测方法研究

上海大学

本论文经答辩委员会全体委员审查，确认符合上海大学硕士学位论文质量要求。

答辩委员会签名：

主席：

委员：

导师：

答辩日期：2025年6月10日

姓名：包胜奇

学号：22721545

论文题目：面向功能材料数据的特征处理与预测方法研究

上海大学学位论文原创性声明

本人郑重声明：所呈交的学位论文是本人在导师指导下，独立进行研究工作所取得的成果。除了文中特别加以标注和致谢的内容外，论文中不包含其他人已发表或撰写过的研究成果。参与同一工作的其他研究者对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名：包胜奇

日期：2025年6月10日

上海大学学位论文使用授权说明

本人完全了解上海大学有关保留、使用学位论文的规定，即：学校有权保留论文及送交论文复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容。

(保密的论文在解密后应遵守此规定)

学位论文作者签名：包胜奇

导师签名：韩继光

日期：2025年6月10日

日期：2025年6月10日

上海大学工学硕士学位论文

面向功能材料数据的特征处理与预测 方法研究

作者: 包胜奇
导师: 韩越兴
学科专业: 计算机应用技术

计算机工程与科学学院

上海大学

2025年4月

A Dissertation Submitted to Shanghai University for the
Degree of Master in Engineering

Research on Feature Processing and Prediction Methods for Functional Material Data

Candidate: Shengqi Bao

Supervisor: Yuexing Han

Major: Technology of Computer Application

School of Computer Engineering and Science

Shanghai University

April, 2025

摘要

功能材料是驱动现代科技革命性进步的物质基础，在电子器件、能源系统、生物医学等战略性产业中扮演着不可或缺的角色。深度学习作为一种强大的数据驱动技术，为功能材料数据的特征处理与预测提供了新的解决方案。然而，深度学习在功能材料中的有效应用仍然面临若干关键挑战，一方面是如何有效提取功能材料图像中存在的多层次复杂特征；另一方面是如何实现功能材料多模态异构数据特征的深度融合。本论文围绕上述挑战展开研究，主要的工作如下：

(1) 针对功能材料图像多层次特征复杂的问题，本论文提出一种基于单模态特征解耦的轻量级深度学习方法，以解决钙钛矿 PUF 防伪标签真实性预测任务。特征解耦策略将宏观与微观特征的学习分离为独立阶段，减少了不同层级特征间的干扰。为满足高效率与低延迟需求，设计了轻量化分流卷积网络，其核心的分流卷积机制显著降低了计算复杂度。同时，采用额外角度边界损失函数，通过增大类间距离、减小类内差异，增强了特征的判别能力。为实现未知样本的有效拒绝，提出了测地度量方法，在高维流形空间根据测地距离对样本真实性进行预测。实验结果表明，本论文提出的方法在钙钛矿 PUF 防伪标签数据集上展现出全面的性能优势，在预测准确率、计算效率方面均表现最佳。

(2) 针对功能材料多模态异构数据特征难以融合的问题，本论文提出了一种基于多模态特征融合的轻量级深度学习方法，以解决海藻酸钙/石墨烯复合材料性能预测任务。通过设计表格驱动特征融合网络，有效整合了图像与表格两种异构数据源。该网络包含两大核心模块：表格引导视觉语义增强模块和门控特征融合模块。表格引导视觉语义增强模块利用表格信息引导视觉特征学习，以实现深度跨模态对齐与增强；门控特征融合模块通过交叉注意力与门控机制，进行模态间的有效交互、对齐及自适应加权融合。同时，为确保计算效率，网络采用了分流卷积、Mamba 等轻量化设计。实验结果表明，相较于其他对比方法，本论文提出的方法在海藻酸钙/石墨烯复合材料数据集上展现出最优的预测性能。

综上所述，本论文针对功能材料数据特征处理与预测所面临的挑战，提出了两种轻量级方法，其中基于特征解耦的单模态方法实现了对钙钛矿 PUF 防伪标签真实

性的预测，而基于特征融合的多模态方法则实现了海藻酸钙/石墨烯复合材料的性能预测。实验结果均表明，两项工作在准确性和计算效率方面展现出显著优势。

关键词：功能材料；轻量级模型；特征解耦；多模态融合

ABSTRACT

Functional materials are the material foundation driving the revolutionary progress of modern technology, playing an indispensable role in strategic industries such as electronic devices, energy systems and biomedicine. Deep learning, as a powerful data-driven technique, provides new solutions for the feature representation of functional materials data. However, there are still several challenges for the effective application of deep learning in functional materials. On the one hand, how to effectively extract the multi-level complex features present in functional materials images; on the other hand, how to achieve the deep fusion of multimodal heterogeneous data features of functional materials. This thesis focuses on the above challenges and the main work is as follows:

(1) Aiming at the problem of complex multi-level features in functional material images, this thesis proposes a lightweight deep learning method based on single-modal feature decoupling for the authenticity prediction of perovskite PUF anti-counterfeiting labels. The feature decoupling strategy separates macroscopic and microscopic feature learning into independent stages, thereby reducing interference among features at different hierarchical levels. To meet the efficiency requirements, a lightweight convolutional network is designed, which significantly reduces the computational cost. Concurrently, an additional angular margin loss function is utilised to enhance feature discriminability by enlarging inter-class margins and minimising intra-class variation. To realise the effective rejection of unknown samples, the geodesic measurement is proposed to predict the authenticity of samples based on the geodesic distance in the high-dimensional manifold space. The experimental results show that the method proposed in this thesis demonstrates comprehensive performance advantages on the perovskite PUF anti-counterfeiting labels, with the best performance in terms of prediction accuracy and computational efficiency.

(2) Aiming at the problem of difficult fusion of multimodal heterogeneous data features for functional materials, this thesis proposes a lightweight deep learning method based on multimodal feature fusion to solve the task of property prediction of calcium al-

ginate/graphene composites. By designing a table-driven feature fusion network, two heterogeneous data sources, namely image and table, are effectively integrated. The network includes two core modules: table-guided visual semantic enhancement module and gated feature fusion module. The table-guided visual semantic enhancement module uses table information to guide visual feature learning for deep cross-modal alignment and enhancement; the gated feature fusion module performs effective inter-modal interaction, alignment and adaptive weighted fusion through cross-attention and gating mechanisms. Meanwhile, to ensure computational efficiency, the network adopts lightweight designs such as PaCoBlock and Mamba. The experimental results show that the proposed method exhibits optimal prediction performance on the calcium alginate/graphene composite dataset compared to other comparative methods.

In summary, this thesis proposes two lightweight approaches to address the challenges of feature processing and prediction of functional material data, including a unimodal approach based on feature decoupling to achieve authenticity prediction of perovskite PUF anti-counterfeiting labels, and a multimodal approach based on feature fusion to achieve prediction of the property of calcium alginate/graphene composites. The experimental results show that both works exhibit significant advantages in terms of accuracy and computational efficiency.

Keywords: Functional materials; Lightweight model; Feature decoupling; Multimodal fusion

目 录

摘 要	I
ABSTRACT	III
第一章 绪论	1
1.1 课题来源	1
1.2 课题背景概述	1
1.3 国内外研究现状.....	3
1.3.1 传统的功能材料特征处理与预测方法.....	3
1.3.2 基于数据驱动的功能材料特征处理与预测方法.....	5
1.4 论文主要工作	8
1.5 论文组织架构	10
第二章 相关理论与技术概述	11
2.1 卷积神经网络相关技术	11
2.1.1 深度残差网络	11
2.1.2 深度可分离卷积.....	12
2.1.3 混合注意力模块.....	13
2.2 序列模型相关技术.....	14
2.2.1 自注意力机制	14
2.2.2 结构化状态空间模型	15
2.3 特征融合	17
2.3.1 特征融合方式	17
2.3.2 特征融合层级	18
2.4 开放集识别	20
2.5 评价指标	21
2.6 本章小结	23
第三章 基于单模态特征解耦的轻量级防伪标签真实性预测方法	24
3.1 方法概述	24

3.1.1	分流卷积网络	25
3.1.2	损失函数	27
3.1.3	测地度量	29
3.2	实验分析	31
3.2.1	数据集介绍	31
3.2.2	实验设置	33
3.2.3	对比实验	33
3.2.4	消融实验	36
3.2.5	不同防伪体系对比	42
3.3	本章小结	43
第四章	基于多模态特征融合的轻量级复合材料性能预测方法	45
4.1	方法概述	45
4.1.1	表格引导视觉语义增强模块	47
4.1.2	门控特征融合模块	50
4.2	实验分析	51
4.2.1	数据集介绍	52
4.2.2	实验设置	53
4.2.3	对比实验	54
4.2.4	消融实验	57
4.3	本章小结	64
第五章	总结与展望	65
5.1	总结	65
5.2	展望	66
参考文献	67
攻读硕士学位期间取得的研究成果	76
致 谢	77

第一章 绪论

1.1 课题来源

本课题得到国家自然科学基金（面上，编号：52273228），云南省科技计划重点项目（编号：202302AB080022），上海大学硅酸盐文物保护教育部重点实验室开放课题项目（编号：SCRC2023ZZ07TS）的资助。

1.2 课题背景概述

功能材料作为现代科学技术的物质基础，通过其独特的物理或化学性质实现特定功能，广泛应用于电子器件、能源系统以及生物医学等重要领域^[1-3]，其发展水平深刻影响着前沿技术的进步。然而，材料的性质往往受到化学组成、微观结构及制备工艺等多重因素共同调控。这不仅构成了极其广阔的材料设计空间，也意味着材料数据本身蕴含着丰富、复杂且多维度的特征。面对这些复杂的数据和巨大的设计空间，传统的研发模式，无论是依赖实验试错还是计算能力有限的模拟方法，都难以进行高效的探索，尤其是在有效处理和利用这些复杂数据特征以实现精准的性能预测并加速特定应用的开发进程方面，传统方法普遍面临着研发周期冗长、成本高昂的严峻挑战^[4-5]。

人工智能，特别是深度学习技术的快速发展，为突破上述瓶颈提供了强大的计算工具，正推动材料科学向数据驱动的“第四范式”——材料信息学（Materials Informatics, MI）加速转型^[6]，如图1.1所示。通过从海量的实验数据与计算模拟结果中学习复杂的潜在关联并识别精细的特征模式，人工智能能够有效揭示材料的“成分-组织-性能”内在联系，从而显著提升材料的发现与优化效率，并为开发材料应用开辟了新的途径^[7]。

高安全性防伪技术是功能材料应用开发的一个重要方向^[9]。钙钛矿等具有特殊光学响应的材料，因其优越的光电特性和工艺灵活性，在构建光学物理不可克隆函数（Physical Unclonable Function, PUF）防伪标签方面具有突出优势。然而，如何高效且精准地预测这些光学 PUF 标签的真实性是一个新的技术挑战。因为 PUF 标签的形貌特征通常表现为微观、随机并且高度相似，使得传统的图像比对或特征点匹配算

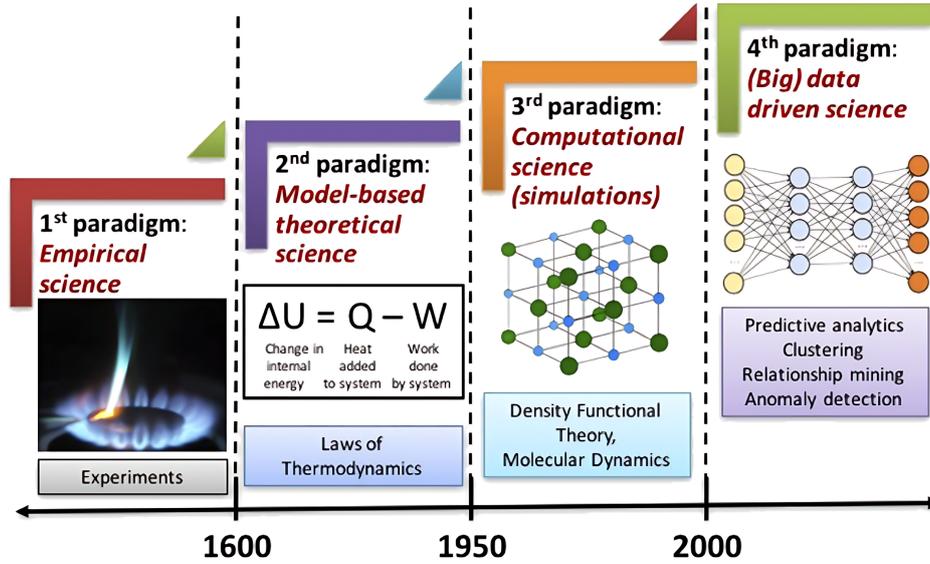


图 1.1 科学研究的四种范式：经验型、理论型、计算型和数据驱动型^[8]。

法在处理速度及区分细微差异的精度上常常难以满足大规模实际应用的需求^[10]。深度学习模型利用其强大的复杂模式认知与自动化特征提取能力，可以从 PUF 标签中发掘并学习那些人眼难以察觉的细微判别信息，对标签实现快速且准确的真实性预测，突破传统方法的性能瓶颈，为高可信度防伪技术的广泛应用奠定关键技术基础。

与此同时，在功能材料的性能预测方面，准确预测材料的关键物理或化学性质，例如导热性能、电学性能、光学特性或力学性能，对于实现材料的按需设计至关重要。然而，许多先进材料，特别是那些组分多样且结构复杂的复合材料，其性能往往是多因素非线性耦合作用的结果，导致使用传统的物理模型或经验公式难以进行精确预测^[11]。以海藻酸钙/石墨烯这类复合材料为例，其导热性能受到微观结构和填料含量等多种复杂因素的综合影响。深度学习方法能够从材料的结构或组成信息中进行挖掘，捕捉高维数据中的非线性依赖关系，从而建立起从结构组分到性能的精准映射模型。这使得对复杂材料体系性能的预测不再依赖简化假设或繁琐的特征工程，大幅提升了预测的精度和效率，进而显著加速了新型功能复合材料的筛选与优化进程。

将深度学习应用于功能材料研究，无论是开发面向特定需求的应用，还是精准预测性能以加速材料发现，都展现出克服传统方法局限性的巨大潜力，这种结合不仅具有深远的科学价值，也预示着广阔的应用前景。

1.3 国内外研究现状

本节概述了功能材料特征处理与预测方法的国内外研究进展，主要围绕两方面展开：一是基于物理模型推导、计算机视觉的经典方法；二是利用机器学习与深度学习从数据中学习规律的数据驱动方法。本节将分别探讨这两类方法在防伪标签真实性预测、材料性能预测中的现状、优势与不足。

1.3.1 传统的功能材料特征处理与预测方法

在材料科学领域，传统的功能材料特征处理与预测方法主要依赖于基于物理原理的模型推导和计算机视觉方法分析。然而，随着材料体系日益复杂以及对高性能、高精度预测需求的不断提升，这些传统方法在处理大规模高维数据时逐渐显现出局限性。

具有物理不可克隆函数(Physical Unclonable Function, PUF)特性的功能材料^[12]因其能产生随机、独特且难以复制的物理“指纹”，为开发高安全性的先进防伪技术提供了理想的物质基础^[13]。防伪系统通过将这些独特的物理结构特征转化为可供处理的数字化特征信息，实现可靠的防伪标签真实性预测。传统的 PUF 标签真实性预测方法通常基于经典计算机视觉方法，其中一类方式是分析材料的特定响应图谱。例如 Hu 等人^[14]提出一种基于拉曼光谱的方法，通过提取拉曼映射图，将其二值化为二进制矩阵作为特征表示，最后计算汉明距离进行匹配验证。

相较于全局图谱分析，局部特征点检测与匹配是更为普遍采用的特征处理方法。其中，尺度不变特征变换(SIFT)^[15]是这类方法的代表，通过检测图像中对尺度、旋转及部分光照变化具有不变性的显著关键点，为每个关键点生成一个高维特征描述符，该描述符构成了对关键点邻域信息的特征表示。Murataj 等人^[16]采用 SIFT 算法提取图像的关键点特征，通过匹配与几何校正对齐图像后，计算像素级的汉明距离来识别 PUF 标签。尽管 SIFT 特征因其不变性被认为较为可靠，但其特征提取与描述的计算复杂度较高，限制了其在追求高效率的应用场景中的部署^[17]。因此，为了平衡特征处理的精度与效率，计算速度更快的 ORB (Oriented FAST and Rotated BRIEF)^[18]算法应运而生。ORB 提供了一种计算成本更低的特征点检测与特征表示方案。Kiremitler 等人^[19]选用了 ORB 算法处理多色荧光显微图像，他们首先提取 ORB 特征点及其描述符，然后进行特征匹配，依据匹配程度计算相似度比率，最终通过与阈值比较来判定 PUF 样本的真伪。同样采用 ORB 算法，Esidir 等人^[20]在处理其研制的食品级可

食用 PUF 标签时，结合了 RANSAC 技术优化特征匹配过程，并依据成功匹配的特征点对数量进行判定。而在处理信息更复杂的石墨烯 PUF 拉曼图像时，Lee 等人^[21]在使用 ORB 算法提取各颜色层的关键点特征的基础上，进一步结合了特征袋 (Bag of Features) 模型进行匹配，并根据匹配点数量完成认证。

虽然上述基于经典计算机视觉的方法提供了一系列 PUF 防伪标签真实性预测方案，但它们在应用于复杂多变的实际环境时，普遍面临着一些共有的局限性。首先，这些方法对图像采集过程中的噪声干扰、光照不均、离焦模糊等因素通常较为敏感，可能直接影响特征描述符的准确性，进而降低特征匹配的可靠性；其次，局部特征点检测与匹配算法往往难以完全鲁棒地应对标签在实际使用中可能出现的微小位移、旋转和尺度变化，这些变化会干扰特征点的位置和特征描述符的计算，导致特征匹配失败或错误匹配。另外，诸如 SIFT 等算法的计算耗时限制了在高通量认证场景中的应用潜力，而对大规模 PUF 标签数据库进行逐一匹配也带来了显著的时间开销。更为关键的是，这些方法普遍依赖于人工设计的基于梯度、强度、角点等底层图像信息的特征提取器和描述符，导致特征表示能力受限，无法最优地捕捉材料内在的微观结构细节。

在复合材料性能预测领域，传统的特征处理与预测方法主要依赖基于物理原理的模型，其中有效介质理论 (Effective Medium Theory, EMT) 或有效介质近似 (Effective Medium Approximation, EMA) 在填料浓度较低、几何结构较为简单的体系中，能够提供合理的预测结果，因此至今仍常被用作理论分析的基准和初步性能评估的手段^[22-23]。

然而，EMT 方法在高精度预测和处理复杂体系方面存在局限性。由于 EMT 通常基于平均场假设和高度简化的微观结构，导致在复杂材料系统和高填料浓度下往往难以提供精确的预测。针对这些问题，研究者们也在尝试发展或改进 EMT 模型。例如，为克服传统 EMT 在处理复杂填料分布上的难题，Ding 等人^[24]针对含多分散氮化硼填料的聚酰亚胺复合材料，开发并运用了一种迭代 EMT 模型，该模型能够结合填料的分布，从而更深入地研究材料的导热机理。类似地，为更准确地考虑界面效应对导热性能的影响，Ruan 等人^[25]在研究胺化石墨烯/聚酰亚胺复合薄膜时，结合了改进的 EMT 模型与拉曼光谱表征，证实了胺化处理能够优化界面、有效降低界面热阻与声子散射，进而提升导热性能。

有限元方法 (Finite Element Method, FEM) 作为一种强大的数值计算工具，擅长

处理具有复杂几何形状和物理特性的问题，在复杂材料系统的模拟与分析中得到了广泛应用。其优势在于能够通过网格离散精确描述复杂的材料微观结构，并允许在模型中引入界面热阻（ITR）等关键物理参数，从而具备实现高保真度预测的能力。Li 等人^[26]利用 FEM 系统研究了多种因素对 WC-Co 复合材料导热性的影响并进行了实验验证，Liu 等人^[27]通过结合扩展有限元与广义多尺度有限元来处理包含大量随机缺陷的异质材料热传导问题。然而，FEM 也面临一定局限性，首先其计算成本通常较为高昂，特别是对于复杂的三维精细模拟；同时，精确的几何建模与高质量网格生成过程既困难又耗时；此外，其预测精度高度依赖于详细且往往难以准确获取的输入参数。这些因素共同限制了 FEM 在需要高效率的大规模材料筛选和设计优化中的实际应用价值。

总的来说，无论是应用于 PUF 标签真实性预测的经典计算机视觉算法，还是用于复合材料性能预测的经典物理模型和数值模拟方法，虽然它们为各自领域的研究奠定了重要基础，但在处理日益复杂的材料系统和大规模数据时，其特征处理能力的不足，直接制约了预测精度和计算效率。因此，需要发展更为有效的特征处理与预测工具，为后续探索数据驱动的方法提供了动机。

1.3.2 基于数据驱动的功能材料特征处理与预测方法

数据驱动方法，特别是机器学习和深度学习技术，正在深刻改变功能材料的研究范式和应用前景^[28]。这些方法能够从大量的实验或模拟数据中学习特征表示以及复杂的结构组分-性能关系，实现对功能材料的准确预测。

为克服经典计算机视觉方法在特征处理上的局限性，深度学习技术被引入防伪领域，并展现出巨大潜力。在研究初期，研究者主要尝试利用相对简单的卷积神经网络进行 PUF 标签的特征提取与真实性预测。Liu 等人^[29]使用经典的 AlexNet 模型处理量子点荧光 PUF，证明了通过学习到的特征，模型可以有效应对成像条件变化，并完成快速的预测；Chen 等人^[30]同样使用 AlexNet，成功从包含复杂信号的硅基纳米杂化物 PUF 中提取了用于有效预测的特征。尽管这些早期工作展示了深度学习应用于 PUF 标签真实性预测的可行性，但也暴露了相应问题。像 AlexNet 这样相对较浅的网络，特征表示能力有限；当面对内部结构复杂且具有跨尺度特征的 PUF 标签时，它们往往难以学习到能够有效区分细微随机“指纹”信息的最优特征，最终限制了预测的准确性。

随着 PUF 标签的材料与结构设计的复杂化，研究者开始探索更强大的深度学习架构以提升预测性能。Jing 等人^[31]针对具有复杂多尺度特征的二维材料褶皱边 PU 标签，开发了名为 DeepKey 的深度学习认证系统。该系统通过专门设计的特征处理、分类和验证机制，实现了对这类复杂图案的快速准确且耐极端环境的认证，展示了深度学习处理复杂结构化特征的能力。

为了进一步提升模型对实际成像条件变化的鲁棒性并兼顾效率，研究者探索了数据增强与先进 CNN 架构相结合的特征学习策略。Chen 等人^[32]通过对原始 PUF 标签施加旋转、噪声等模拟变换构建增强数据库，并选用 ResNeXt 模型进行训练，提升了模型在高精度、快速预测方面的表现。与此同时，更深层的卷积模型也被引入研究。Pradhan 等人^[33]将 InceptionV3 模型应用于蛋白质自组装图像，从中提取特征生成加密密钥，探索了生物材料 PUF 标签与深度学习结合的低成本方案。

除了在卷积架构上进行优化外，研究人员也开始探索基于 Transformer 的架构的方法，希望利用其注意力机制处理 PUF 标签中长距离的特征依赖关系。其中，局部特征变换器^[34] (LoFTR) 算法便是代表之一。LoFTR 通过自注意力和交叉注意力机制进行像素级密集特征匹配，适合处理传统稀疏特征检测器难以有效工作的低纹理区域。基于 LoFTR 的这些优势，Zhang 等人^[35]将其应用于通过纳米打印辅助快速合成法制备的、具有荧光和形貌双通道读出的碳点纳米膜 PUF 标签真实性预测中，他们通过在图像对之间进行像素级的密集匹配，并基于匹配点的数量进行相似性分析，实现了高安全性、高可靠性的预测；Zhang 等人^[36]在其激光打印的不可克隆电子 PUF 研究中，验证了 LoFTR 算法处理光学显微镜图像的有效性，实现了精准可靠且用户友好的真实性预测。

此外，研究者还发展了其他不同的策略。Sun 等人^[37]提出了一种混合策略，他们使用 ResNet50 模型在可扩展数据库中进行快速检索，然后结合 FSIM 相似度指数进行预测；Millwood 等人^[38]提出了一种防御性的预测思路，他们基于“PUF 表型”概念，采用改进 CNN 提取特征并结合多种分类器进行认证，其目标并非直接比对原始 PUF 响应，而是通过预测 PUF 响应的来源来抵抗模型攻击，实现了基于模型的群组预测。

尽管上述基于深度学习的工作在防伪领域取得了一系列进展，但是这些方法大多采用一体化的特征提取网络，倾向于学习到一个整体性的、高度耦合的特征表示。对于包含多层次复杂特征的 PUF 标签而言，在统一的特征学习过程中，不同层级或

不同性质的关键视觉特征可能产生交叉干扰。这种潜在的特征耦合与层级间的干扰，抑制了各自独立信息的充分有效表达，最终限制了模型预测的泛化能力与鲁棒性。

在功能材料的性能预测方面，基于人工提取的特征描述符结合机器学习模型的方法首先得到了应用并展现了一定优势^[39]。研究者依据领域知识或模拟计算，提取与目标性能密切相关的关键物理化学特征，并以此为基础构建预测模型。Yang 等人^[40]利用分子动力学模拟获得的关键物理特征，通过机器学习技术成功预测了石墨烯/六方氮化硼异质结构的热导率。Mishra 等人^[41]同样基于活化能、晶格参数等特性，采用随机森林回归器等集成模型，实现了对固态电解质离子电导率的可靠预测。与深度学习模型相比，机器学习的方法通常具有更好的可解释性。Kanti 等人^[42]应用多种机器学习算法，预测了包含不同纳米颗粒的水基纳米流体的热物理性质，并借助可解释性分析找出了关键影响因素。然而，机器学习方法在特征工程方面的高度依赖性成为了其主要瓶颈。有效特征的选择与提取过程不仅需要深入的专业知识和大量的人工介入，而且预先定义的有限特征集往往无法全面捕捉材料性能的所有复杂因素，特别是那些包含在微观结构中的空间排布或拓扑结构。这些信息难以通过少量的人工设计描述符完全概括，从而限制模型的预测潜力和适用范围。

为了解决这一问题，基于图像的深度学习方法，特别是卷积神经网络成为了有效的替代方案。卷积神经网络的核心优势在于其能够端到端地从原始图像中学习特征，避免了手动设计描述符的需求。Choudhary 等人^[43]利用卷积神经网络分析了模拟的 STM 图像，实现了对二维材料的晶格的识别；Yang 等人^[44]通过深度残差网络，基于表面应变分布图像，预测了纳米晶体的微观缺陷程度；Liu 等人^[45]以微观结构图像为输入，利用卷积神经网络准确预测了多孔材料的有效弹性性能；Chen 等人^[46]设计了多输入卷积神经网络模型，通过整合不同视图的板材图像，提高了刨花板力学性能的预测精度。尽管基于图像的卷积神经网络方法在提取空间特征方面具有一定优势，但其信息来源于单一的图像模态也带来了局限性。材料性能通常是多种因素相互作用的结果，单一模态的数据无法全面捕捉这些相互作用，非图像类信息同样对性能预测起着关键作用。因此，仅依赖图像数据的模型，由于信息表示不完全，在处理复杂材料体系时，会显著限制模型的预测精度和泛化能力。

为克服单一模态数据在预测功能材料性能时因信息不完整而产生的局限性，研究者开始探索多模态数据融合技术。多模态学习的核心优势在于其能够有效整合来自不同来源、不同形式的数据流。通过融合这些包含互补信息的异构数据，模型可以

构建出对材料系统更完整、更丰富的特征表示,从而捕捉到单一数据源无法体现的跨模态关联与协同效应。Farizhandi 等人^[47]构建了一个混合输入的深度神经网络,通过融合微观结构图像和相关数值数据的特征,反向预测了合金的化学成分与加工历史;He 等人^[48]提出一种多模态融合神经网络,采用门控融合策略整合图像与表格数据特征,以预测复合气凝胶的电磁干扰(EMI)屏蔽性能;Das 等人^[49]针对 GNN 预测晶体属性缺乏全局信息的问题,提出了多模态框架 CrysMMNet,该方法融合晶体图结构与材料文本描述,生成更鲁棒的联合表示,显著提升了多种晶体属性的预测精度;Wang 等人^[50]为提升聚合物性能预测,提出了多模态多任务预训练框架 MMPolymer,该方法融合聚合物 1D 序列与 3D 结构信息,通过多种预训练任务学习联合表示,显著提升了下游聚合物性能预测任务的表现。

尽管多模态学习在功能材料性能预测中展现了巨大潜力,但是也面临一些问题。首先,现有的方法大多缺乏有效的跨模态交互与引导机制,未能在特征提取阶段捕捉到模态间的关联。其次,现有的融合策略未能实现特征的充分融合,通常存在信息丢失或语义不对齐的现象。最后,虽然多模态融合可以提升预测效果,但现有的方法通常计算量大、模型复杂,导致计算效率低,难以应用于高通量筛选或资源受限的实际场景。

1.4 论文主要工作

深度学习作为一种强大的数据驱动技术,正为功能材料数据的特征处理与预测带来革命性的机遇。然而,在将深度学习应用于多样化的实际材料科学问题时,仍需面对并克服多项关键性的挑战。

一方面,对于功能材料图像数据,其信息通常呈现出多层次的复杂特征,不仅包括反映整体轮廓与几何结构的宏观形状,还包含体现细微纹理与随机性特征的微观信息。这些不同层次的信息在尺度及细节上差异较大,给特征提取带来了一定挑战。现有的一体化特征提取网络在设计上缺乏对这种多层次特征的针对性设计,导致在学习过程中难以平衡全局与局部的关注点,造成不同层级信息间的交叉干扰,最终难以充分捕捉到最为重要的关键特征,从而限制了模型的预测精度与鲁棒性。

另一方面,功能材料的性能通常由多种因素共同作用而成,如微观结构、化学组分和制备工艺等。不同因素对应的数据来源多样,包含图像、表格等不同形式的

记录。这些异构数据在维度、尺度、以及数据结构上存在显著差异。如何设计能够有效对齐这些不同来源、不同维度、不同数据结构特征的代表空间，并实现它们之间有意义的信息交互与融合，以充分挖掘其互补性与协同性，是当前多模态学习在材料领域应用的核心难题。现有的多模态融合方法在多数情况下较为简化，容易出现信息损失或模态偏见，致使难以全面挖掘和利用各数据来源中包含的全部有效信息。

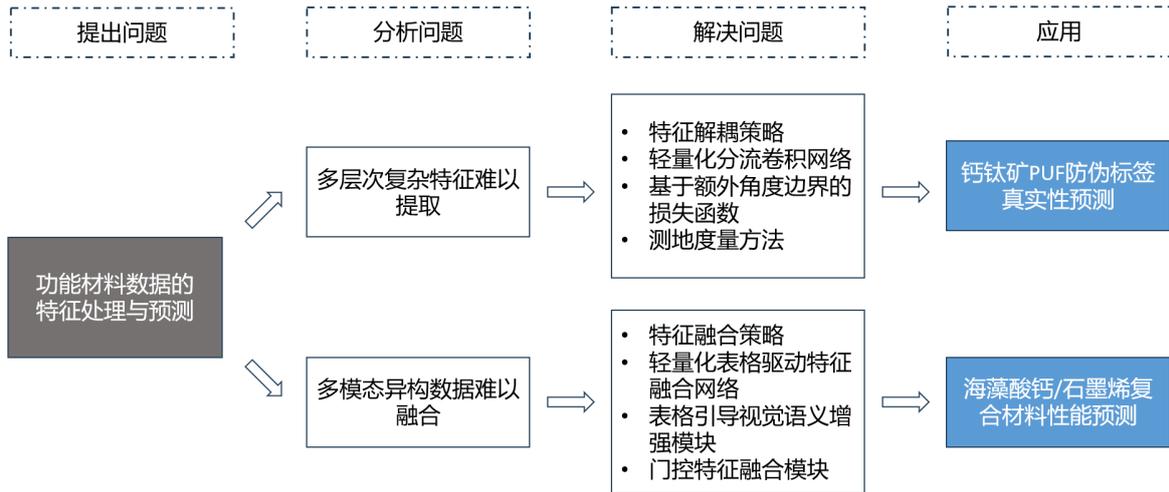


图 1.2 本论文的技术路线图。

为了解决上述挑战，本论文针对功能材料数据的特征处理与预测提出了一系列方法，如图1.2所示。主要的工作和创新如下：

(1) 针对功能材料图像多层次特征复杂的问题，本论文提出了一种基于单模态特征解耦的轻量级深度学习方法，以解决钙钛矿 PUF 防伪标签真实性预测任务。首先，采用特征解耦策略，将特征提取过程显式地解耦为两个独立阶段，分别优化不同层级信息的捕捉能力，从而提升特征表示的有效性，解决了单模态图像中包含的多层次复杂特征难以通过一体化网络有效提取且易产生交叉干扰的问题。其次，设计了一种轻量化分流卷积网络架构，其核心的分流卷积模块通过仅对输入特征图的部分通道进行计算，保留有效信息的同时显著降低了运算量，满足了实际应用对高效率 and 低延迟的需求。再者，在损失函数中加入额外角度边界，以约束模型学习到判别力更强的特征表示。最后，为实现未知样本的有效拒绝，提出了测地度量方法，将学习到的特征映射到高维流形上，利用测地距离对进行样本判别，从而有效识别已知样本并拒绝未在训练中出现的未知伪造样本。

(2) 针对功能材料多模态异构数据特征难以融合的问题，本论文提出了一种基于

多模态特征融合的轻量级深度学习方法，以解决海藻酸钙/石墨烯复合材料性能预测任务。考虑到复合材料的性能通常由微观结构图像信息、化学组分及加工工艺等多种因素协同作用决定，而单一模态数据难以全面捕捉这种复杂的相互依赖关系，本论文构建了一个表格驱动的特征融合网络，以有效整合两种异构数据源。该网络包含两个模块，其中表格引导视觉语义增强模块利用表格数据动态调控图像特征提取过程，实现跨模态信息的深层对齐与语义增强；门控特征融合模块则通过交叉注意力机制促进不同模态特征向量间的有效信息交互与对齐，并结合门控前馈机制实现自适应加权融合。为确保计算效率，网络采用了分流卷积、Mamba 等轻量化设计，有效削减了计算开销和参数规模。

1.5 论文组织架构

本论文基于作者在攻读硕士学位期间的课题研究和工作成果，探讨了面向功能材料数据的特征处理与预测方法的研究工作。本文共分为五个章节，结构安排如下：

第一章首先介绍了功能材料数据特征处理与预测任务的研究背景和意义，然后从传统方法和基于数据驱动方法两个角度梳理和讨论了功能材料特征处理与预测任务的发展、优点和局限性，最后概述了本论文的主要工作和文章的组织架构。

第二章介绍了与本文相关的理论和方法，包括卷积神经网络相关技术、序列模型相关技术、开放集识别相关知识和特征融合相关知识以及相关任务的评价指标。

第三章介绍了基于单模态特征解耦的轻量级深度学习方法，并将其应用于钙钛矿 PUF 防伪标签真实性预测任务。首先对整体方法进行概述，包括使用的网络、损失函数和测地度量方法；其次，展示本方法在钙钛矿 PUF 防伪标签数据集上的对比实验结果，并通过消融实验验证各个模块的有效性和部分超参数对性能的影响；最后和其他类型的防伪技术体系进行了比较分析。

第四章介绍了基于多模态特征融合的轻量级深度学习方法，并将其应用于海藻酸钙/石墨烯复合材料性能预测任务。首先对方法的整体流程进行了概述，包括提出的表格引导视觉语义增强模块和门控特征融合模块；其次，在海藻酸钙/石墨烯复合材料数据集上进行了对比试验；最后通过消融实验验证了各个模块的有效性。

第五章对全文的工作进行了总结，分析了本论文方法的优势和不足，并对未来的研究进行了展望。

第二章 相关理论与技术概述

2.1 卷积神经网络相关技术

本节概述了若干重要的卷积神经网络相关技术，将讨论包括深度残差网络、深度可分离卷积及混合注意力模块在内的代表性方法，分析它们对现代 CNN 架构设计、效率优化与性能提升的主要贡献。

2.1.1 深度残差网络

如何优化深度网络是一个长期存在的关键问题，随着网络层数的持续增长，这一问题愈发突出。为解决深度网络训练中的梯度传播困难与性能退化难题，He 等人^[51]提出了深度残差网络 (ResNet) 及创新的残差学习框架。此框架让网络学习残差映射 $F(x) = H(x) - x$ ，而非直接拟合目标映射 $H(x)$ 。这种方法基于的关键假设是优化残差映射 $F(x)$ 比优化目标映射 $H(x)$ 更简单。若目标映射 $H(x)$ 接近恒等映射，理想的残差映射 $F(x)$ 的输出应为零；而让网络将 $F(x)$ 的权重学习到零，通常比用复杂的非线性层拟合恒等映射 $H(x) = x$ 更为可行。

ResNet 通过构建残差块来实现残差学习。每个残差块包含若干卷积层、批量归一化层和激活函数 ReLU，其关键设计在于引入了一条快捷连接，它将输入 x 跳过变换层，直接与变换层计算后得到的输出 $F(x)$ 进行逐元素加法操作。加和结果通常再经过一次 ReLU 激活，得到最终输出为 $y = \text{ReLU}(F(x) + x)$ 。这种结构为信息和梯度提供了一条直接通路，有效缓解了梯度消失问题，并促进了深度网络的训练。

出于网络深度和计算效率的考量，ResNet 采用了两种主要的残差块设计。在较浅的网络，如 ResNet-18/34 中，通常使用包含两个 3×3 卷积层的基础块。而在更深的网络，如 ResNet-50/101/152 中，则采用瓶颈块设计。瓶颈块由 1×1 、 3×3 、 1×1 三个卷积层构成，其中首个 1×1 卷积用于降低通道维度，随后的 3×3 卷积在较低维度上进行空间特征提取，最后的 1×1 卷积再将通道维度恢复或提升。这种设计在保持相似表示能力的同时，显著减少了参数量和计算复杂度。

ResNet 的整体架构遵循分层模式。网络的初始部分通常由一个较大卷积核、较长步长的卷积层和一个最大池化层组成，用于快速减小图像的初始尺寸并提取低级

特征。接下来，网络由多个阶段堆叠而成，每个阶段包含若干个残差块。在每个阶段之间，特征图的尺寸通常减半，而通道数则会加倍，下采样操作通常发生在每个新阶段的第一个残差块内部。在最后一个残差阶段之后，通常使用全局平均池化层将每个特征图的空间维度压缩为一个单一值，随后连接一个全连接层用于分类输出。此外，批量归一化层通常在每个卷积操作后、激活函数之前使用，以稳定训练过程、支持更高的学习率并加速网络的收敛。

2.1.2 深度可分离卷积

深度可分离卷积^[52]是一种通过分解标准卷积操作来显著减少计算量和参数量的卷积方法。标准的卷积操作通常使用一组卷积核，每个卷积核同时处理输入特征图的所有通道，通过加权求和产生输出特征图的一个通道。深度可分离卷积则将此过程分解为两步：深度卷积和逐点卷积。首先，深度卷积对输入的每个通道应用一个独立的空间卷积核，仅在各自通道内部进行空间特征提取，不混合通道信息。然后，逐点卷积利用一系列 1×1 的卷积核，对深度卷积产生的特征图进行跨通道的线性组合，将这些通道信息有效地整合并映射到新的输出通道空间。

标准卷积的计算成本为：

$$Cost_{std} = D_K^2 \times M \times N \times D_F^2, \quad (2.1)$$

其中 $Cost_{std}$ 表示标准卷积的计算成本， D_K 表示卷积核的尺寸， M 表示输入通道数， N 表示输出通道数， D_F 表示特征图的尺寸。

对于深度卷积来说，它通过对每个输入通道分别进行独立的卷积操作，并将这些信息融合来生成输出特征，因此计算成本分为分深度卷积和逐点卷积两部分，具体如下所示：

$$Cost_{sep} = D_K^2 \times M \times D_F^2 + M \times N \times D_F^2, \quad (2.2)$$

其中 $Cost_{sep}$ 表示深度可分离卷积的计算成本。

深度可分离卷积与标准卷积的计算量比值表示为：

$$\frac{Cost_{sep}}{Cost_{std}} = \frac{D_K^2 \times M \times D_F^2 + M \times N \times D_F^2}{D_K^2 \times M \times N \times D_F^2} = \frac{1}{N} + \frac{1}{D_K^2}, \quad (2.3)$$

从比值可以看出，深度可分离卷积相较于标准卷积大大减少了计算量，特别是在输出通道数 N 较大或者卷积核尺寸 D_K 较大时，计算量节省的效率优势更加明显。

2.1.3 混合注意力模块

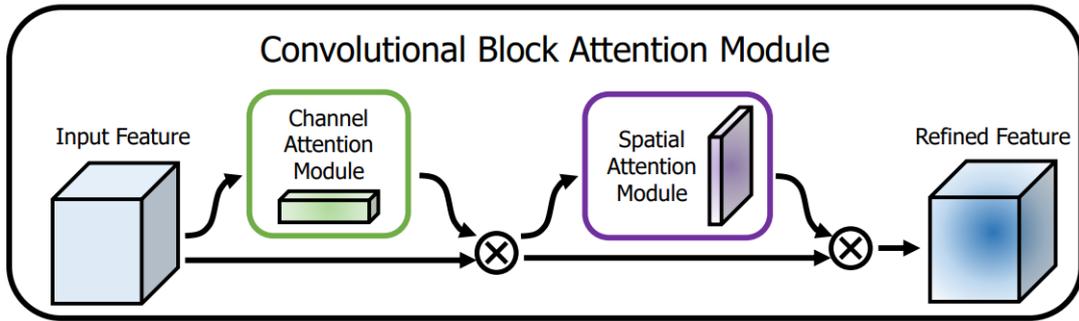


图 2.1 混合注意力模块^[53]。

混合注意力模块 (Convolutional Block Attention Module, CBAM)^[53] 是一种融合了通道和空间注意力机制的模块，通过自适应地聚焦于重要特征来增强卷积神经网络的表达能力，如图2.1所示。该模块依次串联通道注意力与空间注意力，首先执行通道注意力，对全局语义信息进行重校准；然后执行空间注意力，以突出关键的局部细节。

在通道注意力阶段，模块首先使用全局平均池化和全局最大池化，分别捕捉每个通道的整体统计分布与显著响应信息，生成两个通道描述符。接着，这两个描述符被送入一个共享权重的多层感知机进行处理，并通过非线性激活与 Sigmoid 归一化，最终计算出一组范围在 0 到 1 之间的通道注意力权重。通过这些权重逐通道地乘回原始输入特征图，实现对各通道响应的自适应调整，有效强化重要通道特征并抑制冗余或噪声信息。

在空间注意力阶段，模块首先对经过通道注意力调整后的特征图，沿着通道维度进行压缩，同样采用平均池化和最大池化操作计算每个空间位置的概要统计特征，产生两个二维空间描述图。这两个描述图随后被拼接起来，并通过一个具有较大感受野的卷积层进行空间上下文信息的进一步整合。最后，通过 Sigmoid 函数归一化处理，得到最终的二维空间注意力图。该图中的值反映了对应空间位置的重要性，将其与输入特征图逐元素相乘，引导模型在空间上聚焦于关键区域，同时削弱无关区域的影响。

2.2 序列模型相关技术

本节介绍序列模型的相关技术，主要包括自注意力机制和以 Mamba 为代表的选择性结构化状态空间模型。

2.2.1 自注意力机制

自注意力^[54]是一种能够动态地捕捉输入序列内部元素间的长距离依赖关系的计算方式。它通过显式计算序列中任意两个位置间的相关性，为序列中的每个元素生成融合了全局上下文信息的表示向量。

给定一个包含 n 个元素的输入序列，其中每个元素 x_i 由一个 d 维向量表示，即 $x_i \in \mathbb{R}^d$ 。整个序列可表示为矩阵 $X \in \mathbb{R}^{d \times n}$ ，其列向量对应序列中的各个元素：

$$X = [x_1, x_2, \dots, x_n]. \quad (2.4)$$

自注意力机制通过计算元素间的关联度来捕捉依赖关系。为此，它首先利用三个可学习的权重矩阵 $W_Q, W_K, W_V \in \mathbb{R}^{d' \times d}$ 将输入序列 $X \in \mathbb{R}^{d \times n}$ 分别线性投影，生成查询、键和值矩阵：

$$Q = W_Q X, \quad K = W_K X, \quad V = W_V X. \quad (2.5)$$

其中 Q, K, V 维度均为 $\mathbb{R}^{d' \times n}$ ， d' 代表投影后的向量维度。矩阵 Q, K, V 的第 i 列 q_i, k_i, v_i 分别代表原始输入 x_i 对应的查询、键和值向量。

随后，通过计算查询向量与所有键向量的点积来评估它们之间的相关性。具体而言，第 i 个位置的查询向量 q_i 与第 j 个位置的键向量 k_j 之间的原始注意力得分 s_{ij} 计算如下：

$$s_{ij} = \frac{q_i^T k_j}{\sqrt{d'}}. \quad (2.6)$$

引入缩放因子 $\frac{1}{\sqrt{d'}}$ 是为了缓解当 d' 较大时点积结果可能出现的数值过大问题，从而确保后续 Softmax 函数计算的数值稳定性。所有这些得分构成了注意力得分矩阵 $S \in \mathbb{R}^{n \times n}$ 。

为将得分转化为表示概率分布的注意力权重，对得分矩阵 S 的每一行独立应用

Softmax 函数进行归一化：

$$a_{ij} = \frac{\exp(s_{ij})}{\sum_{k=1}^n \exp(s_{ik})}, \quad (2.7)$$

得到了注意力权重矩阵 $A = \text{Softmax}(S) \in \mathbb{R}^{n \times n}$ ，其中 a_{ij} 表示在计算第 i 个位置的输出时，应赋予第 j 个位置的值向量 v_j 的权重。

最终，每个位置 i 的输出表示 $z_i \in \mathbb{R}^d$ 通过对所有值向量 v_j 进行加权求和得到，权重由对应的注意力系数 a_{ij} 确定：

$$z_i = \sum_{j=1}^n a_{ij} v_j. \quad (2.8)$$

2.2.2 结构化状态空间模型

为应对长序列建模中的效率与性能权衡问题，Gu 等人^[55]基于结构化状态空间模型（State Space Model, SSM）设计了 Mamba 架构。Mamba 通过引入选择机制，赋予模型动态的、依赖于输入内容的信息处理能力。同时，通过专门设计的硬件感知高效并行计算算法，确保了这种选择性模型能够以线性时间复杂度高效处理长序列数据。

在选择性 SSM 模型的核心部分，使用了连续时间的线性微分方程来描述输入序列 $x(t)$ 和输出序列 $y(t)$ 之间的映射，其形式可表示为：

$$\dot{h}(t) = A h(t) + B x(t), \quad y(t) = C h(t) + D x(t), \quad (2.9)$$

其中 $h(t) \in \mathbb{R}^N$ 表示隐藏状态， $\dot{h}(t)$ 表示隐藏状态 $h(t)$ 随时间变化的速率，而 A, B, C, D 分别为模型的连续参数。为了在离散时刻进行序列预测，需要对该连续系统进行离散化，常见的方法是零阶保持（Zero-Order Hold, ZOH）。ZOH 的离散化公式可写为：

$$A = \exp(\Delta A), \quad B = A^{-1}(\exp(\Delta A) - I) B, \quad (2.10)$$

其中 Δ 为时间步长，可视为可学习的时间尺度。对应的离散状态更新方程为

$$h_{t+1} = A h_t + B x_t, \quad y_{t+1} = C h_t + D x_t. \quad (2.11)$$

这类离散状态空间模型由于具有线性时不变（LTI）的结构，能够在训练过程中通过

卷积或递归两种方式等价实现，若采用卷积形式，能在时间上实现并行展开并大幅提升计算效率；若采用递归方式则更符合自回归的生成流程。此前的大多数结构化 SSM 皆是遵从这一 LTI 思路，从而在长序列上取得了非常好的速度与准确率。

虽然 LTI 带来了良好的计算优势，但它也意味着模型对输入序列内容的选择或关注能力受限。为此 Gu 等人提出了选择性 SSM 的思路，使与序列相关的参数对当前输入 x_t 敏感，从而在每个时间步上根据内容来动态地调整模型行为。具体形式可写为：

$$B_t = \text{Linear}_B(x_t), \quad C_t = \text{Linear}_C(x_t), \quad \Delta_t = \text{softplus}(\text{Linear}_\Delta(x_t)). \quad (2.12)$$

其中 $\text{softplus}(z) = \log(1 + e^z)$ ， $\text{Linear}(\cdot)$ 表示一个可学习的线性投影层。这样一来，离散化后的参数 A_t, B_t, C_t 就能够随时间和输入而变化，打破了原本 LTI 的限制。虽然此时模型无法简单地通过全局卷积来并行计算，但通过并行扫描的技巧，在硬件层面结合内存层级优化与重计算等手段，依旧实现了接近线性复杂度且极其高效的 GPU 实现。

在网络整体结构上，Gu 等人将选择性 SSM 与门控多层感知机的思路相结合，设计了 Mamba 模块。该模块整体处理流程如图 2.2 所示：

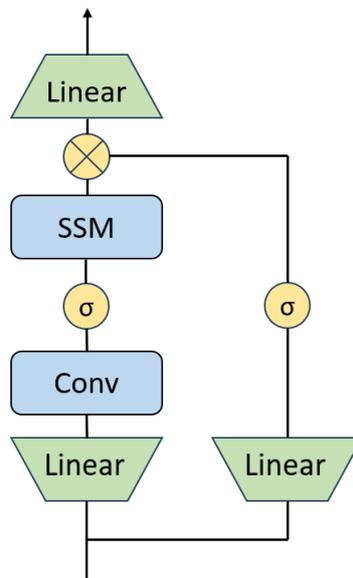


图 2.2 Mamba 模块^[55]。

在 Mamba 模块内部，输入序列首先经由线性投影扩展至较高维度，随后该高维表示被并行送入两个处理分支。一个分支作为核心的序列变换路径，其输入依次通

过因果卷积、非线性激活函数以及关键的选择性 SSM 层，以深度捕捉序列依赖关系。另一分支则通过非线性激活函数直接生成一个门控信号。随后，由第二分支产生的门控信号，与第一分支经过 SSM 处理后的序列表示进行逐元素相乘。最终，该门控调制后的表示通过线性投影层映射回原始维度，完成模块的计算流程并输出结果。

选择性 SSM 以及 Mamba 模块的提出，标志着在序列建模领域的一次重要进展，它兼具了 RNN 的常数时间步推理优势与卷积/注意力模型的强表示能力，为长序列任务中如何有效地对上下文进行筛选并保留关键信息给出了全新的理论与工程实现路径。

2.3 特征融合

特征融合是多模态学习的核心问题，通过结合不同来源的信息以提升模型性能。本节将介绍主要的融合方式和融合层级，并探讨其特点。

2.3.1 特征融合方式

多模态学习通过特征融合，整合不同来源的信息，以利用其异构表示和丰富语义提升模型性能。由于现实世界信息的复杂性，单一模态往往不足以全面描述事物，有效的特征融合能够生成更完整、鲁棒及具判别力的统一表示，是提升模型能力的关键步骤。实现有效特征融合的核心在于设计能够促进不同模态特征进行交互的机制。实践表明，不存在普遍最优的融合策略，针对不同的任务需求、数据特点及模态关系，研究者已提出并验证了多种融合方法，这些方法在交互方式、计算代价和适用性上各有侧重。

基础的融合方法包括元素级运算。此类操作如逐元素的加法、加权平均或乘法，通常应用于已被映射至相同维度空间的不同模态特征向量。例如元素级加法假定特征具有互补性或可加性，而元素级乘法允许一个模态在特征维度上对另一个模态进行调制。这类方法计算较为高效且通常不改变特征维度，但其交互模式较为简单，仅限于对应维度之间，并且假设了特征间已经进行语义对齐的前提。

拼接是另一种广泛应用的策略，它将不同模态的特征向量沿特征维度连接，形成一个更高维的组合特征向量。拼接操作的核心优势在于保留了所有输入模态的完整信息，并将学习跨模态交互的任务通过后续的网络层完成。拼接策略虽然实现简单且不需预设模态关系，但会导致特征维度显著增加，可能加重后续计算负担，并

需要更强大的后续网络来学习复杂的模态间关联。

外积运算提供了一种捕捉复杂二阶及高阶跨模态特征交互的有效途径。这类方法显式地模拟不同模态特征向量中任意元素对之间的乘积关系。其优势在于能够生成高度丰富的交互特征表示，从而精确捕捉细粒度的非线性依赖关系。但是这类模型往往具有较高的计算复杂度与较大的参数规模。

交叉注意力机制通过一种动态加权的方式整合来自不同模态的信息。其核心思想是利用一个模态的特征表示来引导对另一个模态特征表示的选择与聚合。具体而言，该机制通常先通过可学习的投影矩阵将源模态特征映射为查询，并将目标模态特征映射为键和值。接着，计算查询与所有键之间的相似度得分。随后，将这些原始得分通过如 Softmax 等归一化函数处理，生成一组非负且和为 1 的注意力权重。最终，利用这些动态计算出的权重对值向量进行加权运算，得到融合后的输出表示。注意力权重的大小直观地体现了模型在处理源模态特定部分时，认为目标模态的哪些部分最为相关，从而使模型能够显式地学习跨模态间的软对齐关系。这种基于内容计算相关性并进行加权的方式，赋予了模型强大的能力来处理模态间复杂的对应关系、指代关系以及上下文依赖，并能够动态调整不同信息片段的重要性权重。

门控机制通过利用一个模态的信息生成门控信号向量，进而调节另一个模态的特征表示。具体而言，门控信号通常通过将控制模态特征输入一个小型神经网络进行计算获得，该网络通常包含一个或多个线性层，并结合非线性激活函数。常见的激活函数为 Sigmoid 函数，其输出范围介于 0 和 1 之间，从而使得生成的门控信号可视为一种软开关或信息流的通过率，接近 1 时允许信息通过，接近 0 时则阻止信息流动。计算得到的门控信号随后逐元素地作用于目标模态的特征表示。通过这种机制，目标模态中每个特征维度的信息贡献程度由控制模态特征计算出的门控信号动态调控。门控机制的主要优势在于其直观的概念和相对较高的计算效率，为实现模态间的条件依赖建模提供了一种高效的途径。然而，门控机制的局限性也较为明显，其主要在元素级别进行缩放或筛选交互，因此在捕捉特征元素间更复杂的组合性关联方面，表达能力较为有限。

2.3.2 特征融合层级

在多模态学习中，特征融合发生的时机与层级的选择，对模型架构及其最终性能具有决定性影响。依据融合发生的阶段不同，通常可以将特征融合策略划分为原

始数据级融合、层次特征融合和决策级融合三种主要方式^[56]。

原始数据级融合发生在模型处理流程的初始阶段，在输入层对各模态的原始数据或经过独立基础预处理后的数据进行整合。最常见的整合方式是数据拼接，通过将不同模态的数据向量或张量连接起来，从而形成一个维度扩展的统一输入表示。整合后的表示随后被传入一个共享参数的编码器网络，该网络负责从混合输入中同时学习跨模态的高层次抽象特征。这种融合方法的核心优点在于能够从一开始便利用所有模态的原始信息，理论上能够保留最为完整和底层的数据细节。然而，原始数据级融合也面临诸多挑战。首先，随着模态数量的增加，输入维度会迅速扩展，这不仅增加了后续处理的计算负担，还可能导致维度灾难的问题；其次，原始数据级融合主要适用于同质数据，即数据类型和结构相似的模态。对于异质数据如图像、文本、音频等，在输入层直接进行融合较为困难，通常需要设计复杂的对齐机制以及共享表示空间的映射方法；最后，要求单一编码器能够同时有效地从性质各异的多种输入数据流中学习是一项具有挑战性的任务，可能导致模型难以在同等程度上捕捉各模态的细微特征。

层次特征融合在模型的中间层对不同模态的特征进行融合。该策略下，每个模态首先通过各自的编码器提取初步的、带有更丰富语义信息的中高层特征表示。融合操作基于这些经过预处理和抽象的特征进行，并且可以灵活地设置在网络的一个或多个不同深度层次。层次特征融合的主要优势在于其高度的灵活性和强大的交互建模能力。研究人员可以根据任务需求以及对数据内在关联的理解，选择最合适的融合层级。该方法使得模型能够在信息的不同抽象粒度上学习跨模态的深度交互关系，捕捉模态间的协同与互补信息，从而构建信息丰富的联合多模态表示。然而，实现层次特征融合需要精心设计网络架构、特征对齐机制以及具体的融合模块。同时，其计算复杂度不仅随着模态数量的增加而增加，还依赖于所选融合操作的复杂性以及融合发生的网络深度。

决策级融合发生在模型处理流程的末端。此时，各个模态通常已独立完成端到端处理并生成了各自的输出。该融合策略的目标是将这些独立的单模态预测结果或置信度进行整合，以形成统一的最终系统决策。常用的整合技术包括简单投票、对预测概率进行平均、选择置信度最高的预测，或训练额外的元学习器来最优地组合各单模态输出。决策级融合的主要优点在于实现简单，且能保持各模态处理分支的高度独立性与模块化，这不仅便于利用预训练的单模态模型，也易于分析各模态对

最终决策的独立贡献。然而，其缺点在于，由于融合发生在最终决策层，它很大程度上忽略了在特征层面进行跨模态深度交互的潜力。因为各模态在大部分处理流程中相互独立，模态间潜在的复杂协同关系或互补信息可能无法被充分学习和利用。此外，最终的预测性能可能严重依赖于其中表现最佳的单模态子网络的性能。如果某个子网络效果不佳或存在特定偏差，可能会显著影响整体结果，从而对系统的鲁棒性造成影响。

2.4 开放集识别

在传统的监督式学习，尤其是分类任务中，研究者通常默认一种封闭集的设定。该设定基于的假设是训练阶段所学习和涵盖的类别集合 $C_{train} = \{1, 2, \dots, K\}$ 与模型在后续测试阶段所遇到的类别集合 C_{test} 是完全一致的。因此，在封闭集识别框架下，模型的主要目标是将任何输入样本判别并分配到这 K 个预先定义的已知类别中的某一个，无需考虑任何训练阶段未曾遇见过的新类别的可能性。

然而，众多现实世界的应用场景，如防伪识别、安全监控、医疗诊断等，其环境本质上是动态和开放的，这意味着系统会不可避免地遭遇其训练数据覆盖范围之外的未知对象。这使得封闭集的假设在这些场景下不再适用，从而引出了开放集识别这一更具挑战性但也更贴近实际需求的研究问题。

开放集识别的核心设定在于模型在测试阶段面对的样本不仅可能来自于训练时学习过的 K 个已知类别集合 C_{train} ，还可能来自于一个或多个在训练阶段完全未见过的未知类别集合 $C_{unknown}$ ，这两个类别集合是互斥的。在此开放集设定下，对识别模型的要求不再局限于对已知类别的区分，而是扩展为一项双重任务。一方面，模型需要保持对确实属于已知类别 $k \in C_{train}$ 的样本具有高分类准确性。另一方面，模型必须具备有效检测并拒绝未知类别的能力。具体而言，当输入样本 x 实际上源自某个未知类别 $u \in C_{unknown}$ 时，模型应能识别出其不属于任何已知类别，并拒绝将其错误地归类到 C_{train} 中的任何一个类别。

开放集问题的根本挑战在于如何使模型恰当地处理这些未知输入。为封闭集设计的标准分类器，例如基于 Softmax 输出的模型缺乏内在的拒绝机制，其决策空间被已知类别所完全覆盖，倾向于将所有输入强制映射到某个已知类别上，可能导致严重的误判。因此，开放集识别要求模型不仅要学习区分已知类别间的界限，更要

能有效界定整个已知空间的边界，并识别出落在该边界之外的样本。

2.5 评价指标

在评估深度学习模型的计算效率时，通常采用以下几个关键指标：参数量 (Params)、浮点运算量 (FLOPs)、吞吐量 (Throughput) 以及延迟 (Latency)。这些指标从不同维度衡量了模型的复杂度、计算需求与实际运行速度。

参数量指的是模型中所有需要学习的参数的总数量。它关系到模型的存储大小，并在一定程度上反映了模型的内存占用。参数量的计算方式是累加模型所有层中权重张量和偏置向量的元素总数。

浮点运算量衡量模型执行一次前向传播所需的浮点运算总次数，是衡量模型理论计算复杂度的指标。浮点运算量的计算方式是计算通过的每层模型的数学运算如卷积、全连接层的乘加次数并累加得到。

吞吐量表示模型在单位时间内能够处理的输入数据数量，常以每秒处理的样本数 (Instances Per Second, IPS) 来衡量。其计算公式可定义为：

$$T = \frac{N}{t}, \quad (2.13)$$

其中 T 代表吞吐量， N 代表在给定时间内处理的样本总数， t 代表处理这些样本所需的总时间。

延迟是指模型处理单个输入样本所需的总时间，通常以毫秒 (ms) 为单位。其计算公式可定义为：

$$L = \frac{t}{N}, \quad (2.14)$$

其中 L 代表延迟， N 代表在这段时间内处理的样本总数， t 代表处理 N 个样本所需的总时间。

在分类任务中，模型性能通常基于混淆矩阵衍生的若干关键指标进行评估。混淆矩阵包含四个基本计数，包括：真阳性；真阴性；假阳性以及假阴性，如表2.1所示。基于这些基本计数，常用的评价指标及其定义如下。

准确率 (Acc.) 衡量的是模型整体的分类正确程度，定义为正确分类的样本总数

表 2.1 混淆矩阵。

	预测为正例	预测为负例
实际为正例	真阳性 (TP)	假阴性 (FN)
实际为负例	假阳性 (FP)	真阴性 (TN)

与样本总数的比值:

$$\text{Acc.} = \frac{TP + TN}{TP + TN + FP + FN}. \quad (2.15)$$

精确率 (Prec.) 关注的是所有被模型预测为正类的样本中, 实际为正类的比例, 其计算方式为:

$$\text{Prec.} = \frac{TP}{TP + FP}. \quad (2.16)$$

召回率 (Rec.) 衡量的是所有实际为正类的样本中, 被模型成功预测出来的比例, 计算方式为:

$$\text{Rec.} = \frac{TP}{TP + FN}. \quad (2.17)$$

F1 分数 (F1) 是精确率和召回率的调和平均数, 计算方式为:

$$\text{F1} = 2 \times \frac{\text{Prec.} \times \text{Rec.}}{\text{Prec.} + \text{Rec.}}. \quad (2.18)$$

真负类率 (TNR) 用于评估模型正确识别负类样本的能力, 其计算方式为:

$$\text{TNR} = \frac{TN}{TN + FP}. \quad (2.19)$$

对于回归任务, 常采用的评价指标包括平均绝对误差 (MAE)、均方误差 (MSE) 以及决定系数 (R^2)。平均绝对误差用于衡量预测值与真实值之间的平均绝对差异, 其计算公式为

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|, \quad (2.20)$$

其中, N 表示样本的总数量, y_i 表示第 i 个样本的真实值, 而 \hat{y}_i 表示模型对第 i 个样本的预测值。

均方误差则通过对每个预测误差的平方求平均来度量误差，其公式为

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2. \quad (2.21)$$

决定系数 R^2 用以衡量模型对数据方差的解释能力，其定义为

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}, \quad (2.22)$$

其中 \bar{y} 表示真实值的平均值。 R^2 的取值范围在 0 到 1 之间，值越接近 1 表示模型对数据的解释能力越强，反之则说明模型的拟合效果较弱。

2.6 本章小结

本章系统介绍了本论文所涉及的关键理论与技术。内容首先包含了卷积神经网络中的重要进展，如深度残差网络、深度可分离卷积和卷积块注意力模块。接着探讨了序列建模的核心方法，包括自注意力机制和以 Mamba 为代表的结构化状态空间模型。随后讨论了开放集识别问题及其对模型鲁棒性的意义，并对多模态特征融合的多种方法与不同层级进行了梳理和分析。最后，总结了衡量性能的关键评价指标。

第三章 基于单模态特征解耦的轻量级防伪标签真实性预测方法

随着全球贸易的扩展和品牌价值保护需求的提升，开发先进的防伪解决方案成为关键议题。利用功能材料如钙钛矿薄膜制备物理不可克隆函数 (Physical Unclonable Function, PUF) 标签，为实现这一目标提供了新的技术路径。近年来，深度学习凭借其强大的自动特征学习能力，在 PUF 防伪标签真实性预测方面展现出超越传统方法的潜力，正迅速成为该领域的研究热点。

然而现有深度学习方法应用于钙钛矿 PUF 防伪标签真实性预测时，仍面临若干关键挑战。首先是特征复杂与耦合干扰问题，PUF 标签同时包含宏观形状信息与微观纹理信息等多层次复杂特征。当前主流的深度学习模型通常采用一体化网络提取这些特征，容易导致不同层级信息的相互耦合与干扰，难以充分地表示各层次的关键特征。其次是计算效率与实时性需求难以平衡的挑战，许多高性能的深度模型具有庞大的参数量和密集的计算需求；虽然可能提升预测精度，但其推理速度往往难以满足工业检测场景所需的低延迟要求，这限制了模型的实际部署。最后是开放集识别能力欠缺问题，真实的防伪场景是一个开放环境，要求模型不仅能识别真实样本，还要能有效拒识未知伪造样本。当前绝大多数基于深度学习方法是在封闭集假设下训练和评估的，缺乏可靠的未知样本拒识机制。

3.1 方法概述

针对现有方法在处理钙钛矿 PUF 防伪标签真实性预测任务时面临的挑战，本章提出了一种基于单模态特征解耦的轻量级深度学习的方法。该方法通过特征解耦策略，将宏观形状和微观纹理的学习分离为两个独立阶段，有效提取多层次复杂特征。具体来说，首先利用一个轻量化的分流卷积网络 (Partial Convolution Network, PaCoNet) 处理输入的待测 PUF 标签，专注于捕捉其整体轮廓等宏观形状特征，并据此进行初步真实性预测。随后，根据宏观特征的预测结果，从预先构建的纹理模型库中选取与形状对应的一个分流卷积网络，用于深入挖掘 PUF 标签的微观纹理特征。为进一步提升微观特征的区分度，在损失函数中添加了额外角度边界 (Additive Angular

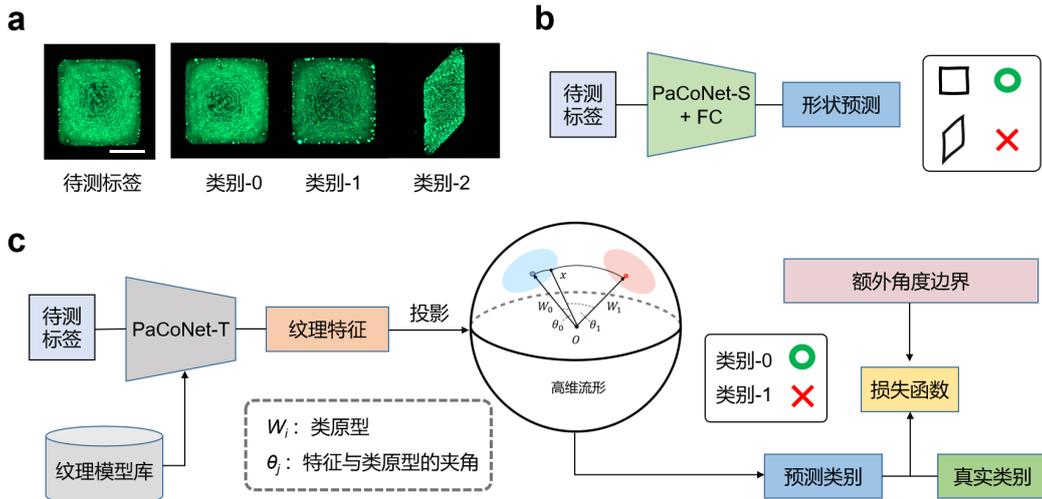


图 3.1 钙钛矿 PUF 防伪标签真实性预测方法流程图。(a) 钙钛矿 PUF 防伪标签。(b) 形状阶段。(c) 纹理阶段。图中比例尺为 100 μm 。

Margin, AAM) 进行优化。最后,为解决开放集识别问题,本章提出了测地度量方法 (Geodesic Measurement, GM), 在高维流形空间中根据微观纹理特征与其对应类原型之间的测地距离进行真实性预测,并实现对未知样本的有效拒识。整体流程如图 3.1所示,本节后续将详细介绍各组成部分及其具体实现细节。

3.1.1 分流卷积网络

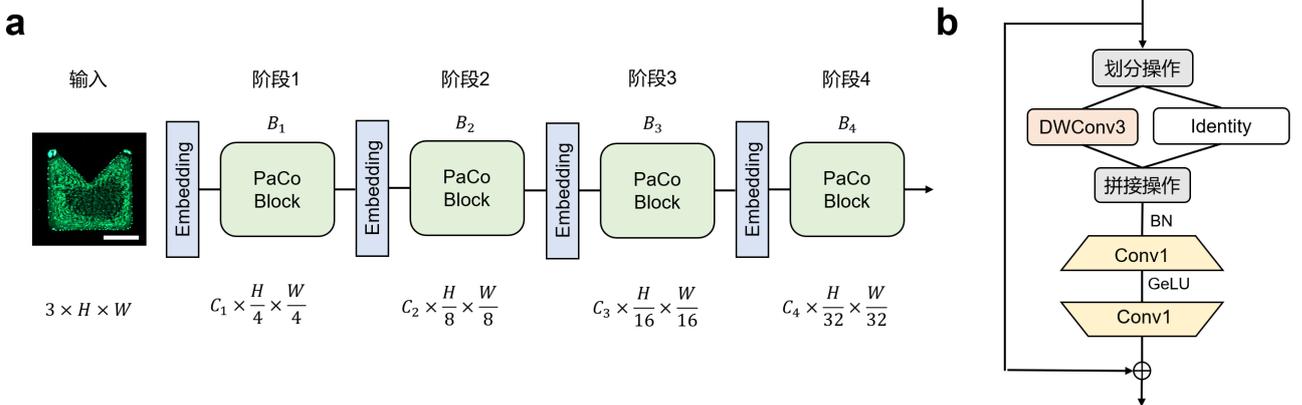


图 3.2 分流卷积网络示意图。(a) 分流卷积网络架构。(b) 分流卷积模块 (PaCoBlock) 结构图。图中比例尺为 100 μm 。

本章提出的轻量化分流卷积网络,其整体架构如图 3.2 (a) 所示,采用包含四个阶段的层级化设计。每个阶段的起始处设置了嵌入层,负责执行下采样操作,以实现特征图空间尺寸的压缩以及通道维度的扩展。在各个阶段内部,分别堆叠了 B_1 至 B_4 个分流卷积模块 (Partial Convolution Block, PaCoBlock)。通过动态调整每个阶段

的模块数量 $B_{i \in \{1,2,3,4\}}$ ，可以灵活构建出不同深度和计算复杂度的可扩展网络架构。

分流卷积模块的具体实现如图 3.2(b) 所示，其高效的卷积方式，大量减少了计算冗余。具体来说，对于给定的输入特征图 $X \in \mathbb{R}^{H \times W \times C}$ ，该模块首先沿通道维度将其划分为更新组 X_p 和保留组 X_i ，其中更新组的通道占比为 α ，即更新组的大小为 $X_p \in \mathbb{R}^{H \times W \times (\alpha C)}$ 。在传统卷积的计算过程中，对输入特征图的所有 C 个通道执行卷积操作，在输出通道数与输入保持一致的情况下，其参数总量为 $K^2 C^2$ ，其中 K 为卷积核大小，计算复杂度为 $K^2 C^2 H W$ 。相比之下，本方法仅对更新组 X_p 应用深度可分离卷积 (Depthwise Separable Convolution, DWConv)^[57]。首先对更新组中的每个通道独立应用一个 $K \times K$ 的卷积核，其参数量为 $K^2(\alpha C)$ ，计算复杂度为 $K^2(\alpha C) H W$ ；然后使用 1×1 的卷积实现通道间的特征交互，其参数量为 $(\alpha C)^2$ ，计算复杂度为 $(\alpha C)^2 \times H \times W$ ；最后将经过深度卷积处理后的更新组输出与保留组 X_i 沿通道维度重新拼接，生成最终的输出特征图 $X' \in \mathbb{R}^{H \times W \times C}$ 。通过这种仅处理部分通道的方式，相比于标准卷积，分流卷积模块显著减少了参数量和运算量。

分流卷积模块设计的理论基础在于对深度卷积网络中普遍存在的特征图通道冗余性的观察。GhostNet^[58] 和 SPConv^[59] 等研究已经证明，传统卷积生成的特征图中，许多通道之间常常存在高度的相关性或相似性，这意味着其中包含了相当程度的信息冗余。分流卷积模块的设计正是利用这一特性来提升计算效率，通过只计算和更新部分通道，同时保留其余通道的信息，不仅能够在很大程度上保留输入的代表性信息，还能避免对冗余信息的重复计算所带来的计算资源消耗。

为了改善特征分布并加速训练收敛，分流卷积模块采用了一系列优化策略。首先，通过应用批量归一化^[60]，对每一层的输入进行标准化处理，加速了训练过程并提高了模型的稳定性。模块结构上采用倒置瓶颈结构^[61]，通过先扩展后压缩特征维度，增强了模型的非线性表达能力，同时优化了计算效率。为了进一步提升性能，模块选择了 GELU^[62] 激活函数，替代了经典的 ReLU^[63]。与 ReLU 相比，GELU 具有更平滑的非线性变换特性，能够有效缓解梯度消失问题，并提升模型的泛化能力。最后，受到 ResNet^[51] 残差机制的启发，在模块内引入了快捷连接以缓解深层网络中的梯度消失问题，进一步增强训练的稳定性和模型的鲁棒性。

针对防伪标签中宏观形状与微观纹理特征在复杂度上的显著差异，为平衡模型计算效率与特征提取能力，本章基于分流卷积网络的四阶段可扩展层级架构设计了两不同规模的网络变体。具体而言，对于复杂度相对较低的宏观形状特征提取，构

建了一个轻量化网络 PaCoNet-S，其各阶段的模块配置设定为较为精简的 $[1, 1, 2, 1]$ 。而对于需要捕捉精细、高复杂度信息的微观纹理特征，则设计了规模更大的 PaCoNet-T 网络。相较于 PaCoNet-S，PaCoNet-T 不仅增加了各阶段分流卷积模块的堆叠数量，将配置扩展为 $[2, 3, 4, 3]$ ，还在网络的第三和第四阶段，扩展了特征通道数，分别从 48 扩展至 80 以及从 96 扩展至 160，以增强其深度特征学习与表达能力，具体配置详见表 3.1。其中，“Conv_K_C”与“DWConv_K_C”分别表示卷积核大小为 K 、输出通道数为 C 的标准卷积与深度可分离卷积； H 和 W 表示输入图像的尺寸，而 B_i 表示第 i 阶段中分流卷积模块的堆叠数量。尽管采用更大规模的网络配置会增加参数量和计算量，从而在一定程度上影响计算效率。然而，考虑到防伪标签真实性预测任务对高精度的核心要求，在纹理特征提取阶段投入更多计算资源以换取更高的预测准确率是必要的。

3.1.2 损失函数

Softmax 损失函数在多类别分类任务中被广泛采用，它将神经网络最后一层的原始输出转换为概率分布，并通过最大化样本真实类别的后验概率来指导模型参数的更新。设训练数据集包含 N 个样本，其中第 i 个样本 ($i = 1, \dots, N$) 由特征向量 x_i 和其对应的真实类别标签 y_i 组成，标签 y_i 从 K 个可能的类别 $\{1, 2, \dots, K\}$ 中取值。Softmax 损失函数 L_S 的公式如下所示：

$$L_S = -\frac{1}{N} \sum_{i=1}^N \log \frac{\exp(W_{y_i}^T x_i + b_{y_i})}{\sum_{j=1}^K \exp(W_j^T x_i + b_j)}, \quad (3.1)$$

其中 W_j 表示与类别 j 对应的权重向量， b_j 为对应的偏置项。

然而，Softmax 损失主要关注样本是否被正确分类，而没有显示约束同一类别内样本的聚集以及不同类别之间的分离。在类别相似度较高的情况下，这往往会导致决策边界模糊，进而影响整体识别性能。例如，在防伪领域中，PUF 标签的内部微观纹理具有独特且不可复制的特征，但这种细微差异在视觉上并不明显，这往往导致学习到的特征决策边界模糊，进而影响整体的预测性能。近年来，研究者们提出了引入边缘间隔^[64]、角度间隔^[65]等改进方法，通过对特征空间中的角度和距离进行显式约束，从而提升所学特征的区别能力。

在此基础上，本章在损失函数中添加了额外角度边界 (Additive Angular Margin, AAM)，以进一步提升特征的判别性。具体而言，首先通过归一化将每个类别的权重

表 3.1 分流卷积网络在不同规模下的配置细节。

模块名称	特征尺寸	具体配置		PaCoNet	
				S	T
Embedding	$\frac{H}{4} \times \frac{W}{4}$	Conv_3_C ₁ , BN	#Channel C ₁	24	24
阶段 1		$\left[\begin{array}{c} \text{DWConv}_3\text{-C}_1, \\ \text{BN}, \\ \text{Conv}_1\text{-4C}_1, \\ \text{Acti}, \\ \text{Conv}_1\text{-C}_1 \end{array} \right] \times B_1$	#Blocks B ₁	1	2
Embedding	$\frac{H}{8} \times \frac{W}{8}$	Conv_3_C ₂ , BN	#Channel C ₂	40	40
阶段 2		$\left[\begin{array}{c} \text{DWConv}_3\text{-C}_2, \\ \text{BN}, \\ \text{Conv}_1\text{-4C}_2, \\ \text{Acti}, \\ \text{Conv}_1\text{-C}_2 \end{array} \right] \times B_2$	#Blocks B ₂	1	3
Embedding	$\frac{H}{16} \times \frac{W}{16}$	Conv_3_C ₃ , BN	#Channel C ₃	48	80
阶段 3		$\left[\begin{array}{c} \text{DWConv}_3\text{-C}_3, \\ \text{BN}, \\ \text{Conv}_1\text{-4C}_3, \\ \text{Acti}, \\ \text{Conv}_1\text{-C}_3 \end{array} \right] \times B_3$	#Blocks B ₃	2	4
Embedding	$\frac{H}{32} \times \frac{W}{32}$	Conv_3_C ₄ , BN	#Channel C ₄	96	160
阶段 4		$\left[\begin{array}{c} \text{DWConv}_3\text{-C}_4, \\ \text{BN}, \\ \text{Conv}_1\text{-4C}_4, \\ \text{Acti}, \\ \text{Conv}_1\text{-C}_4 \end{array} \right] \times B_4$	#Blocks B ₄	1	3
参数量 (M)				1.10	2.58
计算量 (G)				0.07	0.17

向量 W_j 固定为 1，偏置项 b_j 置为 0，同时对特征向量 x_i 进行归一化处理，并引入尺度因子 s 进行重新缩放。原始的损失函数可以重写为：

$$L'_S = -\frac{1}{N} \sum_{i=1}^N \log \frac{\exp(s \cos \theta_{y_i})}{\sum_{j=1}^K \exp(s \cos \theta_j)}, \quad (3.2)$$

其中 θ_{y_i} 表示样本 x_i 与真实类别 y_i 对应权重向量之间的夹角，而 θ_j 表示 x_i 与类别 j

对应权重向量之间的夹角。

为了使类内聚集、类间分离，本章在真实类别对应的角度 θ_{y_i} 上引入了一个固定的角度边界 m 。引入额外角度边界后的损失函数表示为：

$$L_{AAM} = -\frac{1}{N} \sum_{i=1}^N \log \frac{\exp(s \cos(\theta_{y_i} + m))}{\exp(s \cos(\theta_{y_i} + m)) + \sum_{j \neq y_i} \exp(s \cos \theta_j)}. \quad (3.3)$$

改进后的损失函数使在训练过程直接优化特征与类原型之间的余弦距离，并通过添加额外角度边界，要求学习到的特征不仅能被正确分类，还需要在特征空间中满足更优的分布特性，一方面有效增大了不同类别特征之间的距离，提升了类间的可区分度；另一方面则促使同一类别内部的样本特征更加紧密地聚集在其类原型附近，提高了类内的紧凑度。在面对 PUF 标签真实性预测这样需要区分高度相似样本的困难场景下，显著提升了网络的整体预测性能。

3.1.3 测地度量

防伪标签真实性预测作为一种开放集任务，其难点不仅在于对已知正样本的准确识别，更在于有效拒识训练阶段未曾见过的、形式多样的伪造样本。开放集任务的挑战在于数据分布的不完全可知性，在训练阶段仅能获得有限的正样本类别，而在实际应用中可能出现多样化的伪造样本。传统深度学习模型大多基于闭集假设进行训练和决策，倾向于在已知类别构成的特征空间内划分决策边界。这使得它们在面对分布之外的未知样本时，往往会将其强制或错误地归类为某个已知类别。

针对上述问题，本章提出了测地度量方法 (Geodesic Measurement, GM)，通过在高维流形上基于测地距离定义已知类别的分布边界来实现开放集识别。首先将特征映射到高维流形上，在本章的具体实现中，通过归一化操作将分流卷积网络提取的特征向量映射到单位超球面。选择超球面的原因在于，它能够解耦特征表示中的方向信息与幅度信息，将分析重点放在更能体现类别差异的方向性上，同时提供了一个几何结构清晰且有界的空间。其次，采用测地距离进行相似性度量，在超球面上通过计算待测样本 x 与每个已知类别 j 的归一化类原型 \mathcal{W}_j 之间的测地距离来评估相似性，测地距离的计算如下所示：

$$gd(x, \mathcal{W}_j) = \cos^{-1} \left(\frac{x}{\|x\|} \cdot \mathcal{W}_j \right), \quad (3.4)$$

其中 $\|\cdot\|$ 表示 Frobenius 范数， x 表示待识别标签的特征向量， \mathcal{W}_j 表示归一化后的

类原型。测地距离作为流形上两点间最短路径的长度，是一种内在的几何度量，在超球面上使用测地距离，量化了特征向量的方向一致性，并且相对欧氏距离，其对特征向量的幅度变化不敏感，因此更为稳健。

为实现拒识功能，测地度量方法依据已知类别在高维流形上的分布来估计拒识边界。该方法设定了一个判别阈值 λ ，用于界定已知类别在高维流形上的分布区域估计。该阈值基于训练数据中所有已知样本到其最近邻类原型的最大测地距离来确定已知类别的经验边界。通过训练数据的内在分布特性来界定已知范围，为区分显著偏离所有已知类别分布的未知样本提供量化依据，具体计算公式如下所示：

$$\lambda = k \cdot \max_{i \in \{1, \dots, N\}} \left\{ \min_{j \in \{1, \dots, K\}} \{gd(x_i, \mathcal{W}_j)\} \right\}, \quad (3.5)$$

其中， x_i 表示训练集中第 i 张图像的特征向量， N 表示训练样本总数， K 表示类别总数，超参数 k 则提供了对该判别边界范围进行量化调整的机制，允许根据具体应用对误拒率和未知样本接受风险进行权衡。

在真实性预测过程中，首先计算待测样本的归一化特征 x' 与所有已知的归一化类原型 \mathcal{W}_j 之间的测地距离，从中确定最小值 $d_{min} = \min_j \{gd(x', \mathcal{W}_j)\}$ ，随后，将这个最小测地距离 d_{min} 与预先设定的判别阈值 λ 进行比较。若 d_{min} 小于等于 λ ，则表明待测样本 x' 足够接近其最近的已知类原型 \mathcal{W}_j ，落在了可接受的分布区域内，样本将归类为类别 j 的真实样本。反之，若 d_{min} 大于 λ ，则说明即使是到最近的已知类原型，该样本的距离也超出了边界，显著偏离所有已知类别分布，因此将其判定为未知类别或伪造品。

测地度量方法的有效性在于将开放集识别问题转化为在高维流形上的几何距离判别问题。传统的欧氏距离仅度量两点间的直线距离，因此在处理分布于高维流形中的特征时，往往会因忽略流形本身的弯曲而导致对距离的判断产生偏差，难以有效捕捉数据点之间复杂的非线性关联。相比之下，测地距离通过沿着流形表面计算最短路径，不仅能够更真实地反映样本在特征空间中的内在结构与非线性关系，还能有效聚焦于特征的方向信息，降低对可能无关或易变的特征幅度的敏感性。此外，测地度量基于训练数据在经过额外角度边界损失优化后的流形上的经验分布，提供了一种可通过超参数 k 调节的拒识边界，能够通过估计已知类别在流形上的经验边界，为拒识未知样本提供了数据驱动的依据，这些特性使其相比依赖 Softmax 输出置信度或在欧氏空间中易受尺度影响的方法，更具鲁棒性。

3.2 实验分析

本节介绍了使用的数据集与实验设置，并进行了对比实验和消融实验，以验证模型的有效性和鲁棒性。

3.2.1 数据集介绍

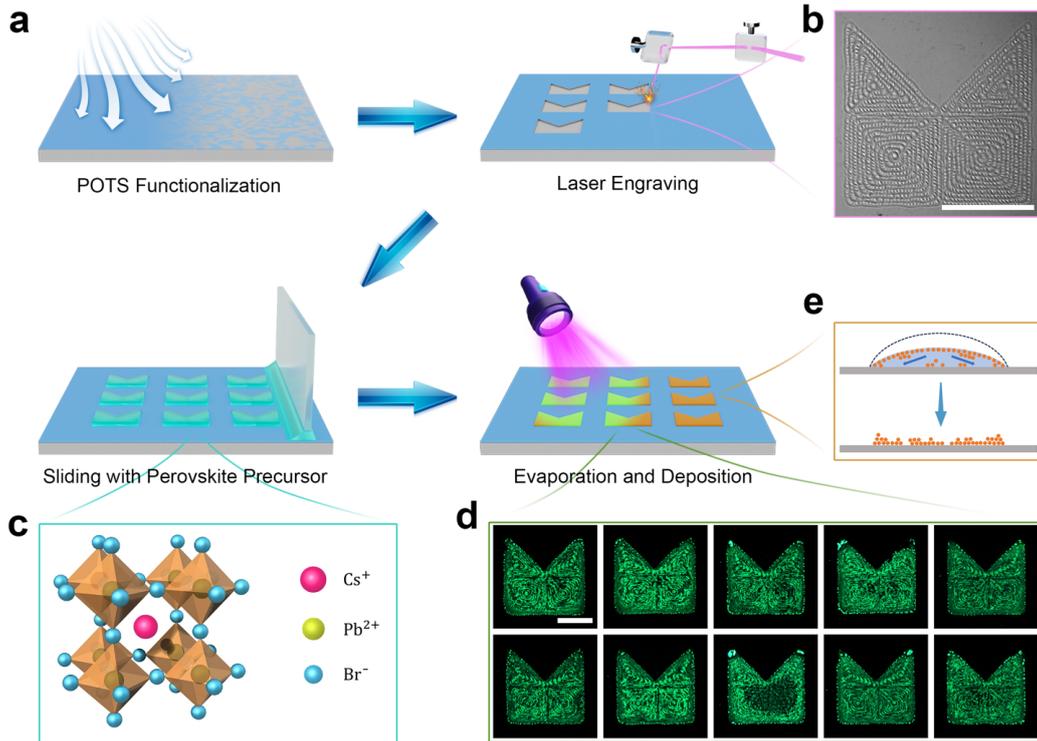


图 3.3 钙钛矿纳米晶体薄膜的制备流程。图像的比例尺为 $100 \mu\text{m}$ ^[9]。

本章构建了一个基于钙钛矿晶体薄膜图像的 PUF 防伪标签数据集。防伪标签的制备过程如图 3.3 所示，首先利用高通量液滴阵列生成技术，通过精确控制表面张力，制备出具有特定形貌与尺寸的钙钛矿晶体薄膜。随后，通过钙钛矿纳米晶体的随机沉积，最终构建出具备荧光特性与物理不可克隆功能的微标签。物理不可克隆性是该标签的主要防伪机制，其独特的微观纹理来源于激光刻蚀后基底的不规则性以及钙钛矿纳米晶体结晶过程的随机性。这种随机性在物理层面上难以精确控制和预测，确保了每个标签都拥有独一无二、难以伪造的纹理特征。此外，钙钛矿材料固有的荧光特性为标签提供了一个易于检测的鉴别维度，通过简单的荧光激发与观测，即可对标签的真伪进行初步判断。本章所提出的 PUF 防伪标签的尺寸小于 5.3×10^{-2} 平方毫米，微小的体积便利了标签在各类产品及其包装中的集成应用。同时，标签

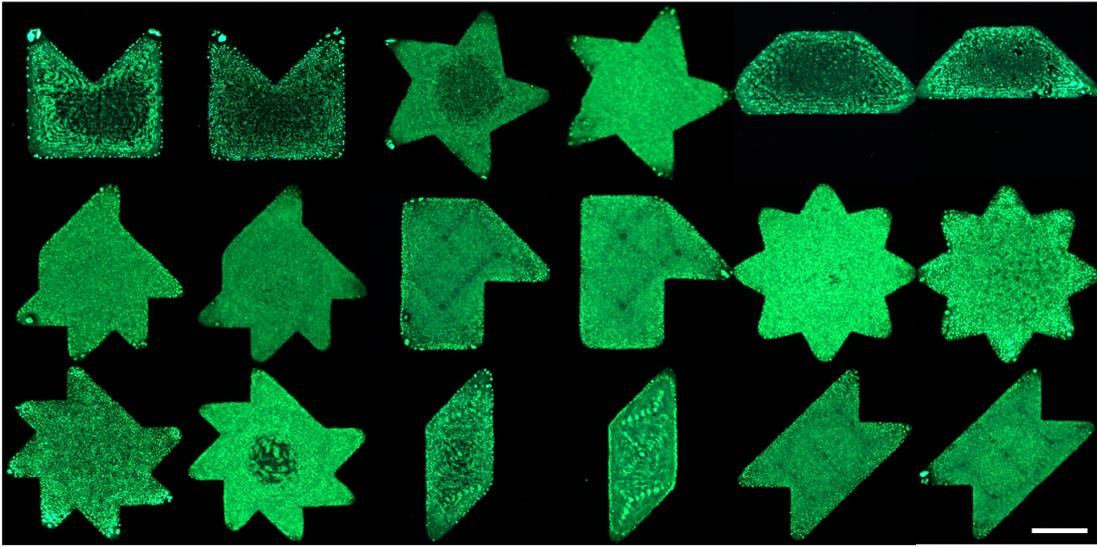


图 3.4 钙钛矿纳米晶体薄膜图像示例。图像的比例尺为 $100\ \mu\text{m}$ 。

的制备成本低廉，低于 3×10^{-4} 美元，为大规模商业化应用提供了经济可行性，使得广泛的产品能够以较低的成本获得有效的防伪保护。此外，该标签的理论编码容量高达 1.7×10^{1956} ，能够生成大量唯一性标签，避免标签之间重复性。

本章的数据集包含 60 种不同的形状类别，每个类别下包含 64 张分辨率为 960×960 像素的原始图像，这些图像均通过专业荧光显微镜采集，如图 3.4 所示。为增强数据多样性以及模拟实际应用场景中可能出现的环境干扰，本章对原始图像施加了不同角度的旋转和亮度扰动，将每张样本扩展至 4680 个子样本。

考虑到伪造样本在现实中的多样性与不可穷尽性，本章将从未在训练中出现的类别标签整体视为伪造样本。数据集的具体划分过程如下：首先，从全部 60 个形状类别中，随机选取 50 个类别作为已知类别，其数据将用于模型的训练与验证；剩余的 10 个类别则作为未知类别，它们的数据完全不参与模型训练，仅在最终测试时用于评估模型的拒识能力。接着，对于这 50 个已知类别的数据，将其样本按照 7:3 的比例随机划分为训练集和验证集，分别用于模型参数学习和超参数调整。最终构建的测试集则由两部分构成：一部分是从已知类别预留样本中随机抽取的 5000 张图像，代表真实样本；另一部分是从未知类别的样本中随机抽取的 5000 张图像，代表伪造样本。

3.2.2 实验设置

本章在 Ubuntu 20.04 操作系统环境下搭建实验平台，采用 PyTorch 1.12 深度学习框架构建网络模型，硬件配置搭载 NVIDIA RTX 3090 GPU (24 GB 显存)。实验中统一将图片大小调整为 224×224 作为模型输入。针对提出的 PaCoNet-S 和 PaCoNet-T 不同的模型规模，制定差异化训练策略：PaCoNet-S 采用 AdamW 优化器，迭代轮次设置为 60，初始学习率设为 0.003，每 10 个迭代轮次进行学习率衰减，批量大小设置为 128；PaCoNet-T 采用 AdamW 优化器，迭代轮次设置为 120，初始学习率设为 0.002，每 15 个迭代轮次进行学习率衰减，批量大小设置为 128。

3.2.3 对比实验

表 3.2 不同模型在钙钛矿 PUF 防伪标签数据集上的计算效率与准确率对比。最好和次好的结果分别以粗体和下划线标出。

网络模型	网络类型	参数量 (M)	计算量 (G)	吞吐量 (IPS)	延迟 (ms)	准确率 (%)	时间消耗 (s)		
							形状	纹理	总计
Mixer-S/32 ^[66]	MLP	18.6	1.00	6016	5.29	79.40	0.373	0.371	0.744
ResNet18 ^[51]	Conv	11.2	1.82	4847	<u>3.23</u>	86.65	0.252	0.257	0.509
DenseNet×1 ^[67]	Conv	6.9	2.80	1389	20.30	83.82	0.397	0.390	0.787
MobileNetV2 ^[61]	Conv	2.3	0.30	4309	6.13	83.57	0.145	0.140	0.285
MnasNet×1 ^[68]	Conv	3.1	0.31	4586	5.86	83.12	0.166	0.170	0.336
EfficientNet-B0 ^[69]	Conv	4.0	0.38	3299	9.60	82.15	0.201	0.199	0.400
GhostNet×1 ^[58]	Conv	3.9	<u>0.15</u>	5700	11.06	77.68	0.223	0.207	0.430
RegNetX-200MF ^[70]	Conv	2.3	0.20	<u>10719</u>	5.91	80.44	0.140	0.135	0.275
ConvNeXt-Atto ^[71]	Conv	3.4	0.55	2540	5.20	81.78	0.132	0.124	0.256
ViT-T ^[72]	Hybrid	5.5	1.07	3168	9.84	79.68	0.225	0.226	0.451
DeiT-T ^[73]	Hybrid	5.5	1.08	3166	9.74	81.28	0.228	0.232	0.460
Swin-T ^[74]	Hybrid	27.5	4.38	798	14.32	50.00	0.626	0.620	1.246
MobileViT-XXS ^[75]	Hybrid	1.0	0.26	2854	11.64	83.08	0.140	0.134	0.274
PVTv2-B0 ^[76]	Hybrid	3.4	0.53	1837	11.11	50.00	0.149	0.143	0.292
EdgeNeXt-XXS ^[77]	Hybrid	1.2	0.20	2130	9.87	86.57	<u>0.107</u>	0.108	<u>0.215</u>
PaCoNet(Ours)	Conv	<u>1.1</u>	0.07	30223	2.67	99.79	0.054	<u>0.123</u>	0.177

为了验证本章提出的网络 PaCoNet 在计算效率和准确率方面的优势，本节将其与多种主流卷积神经网络以及代表性的混合模型进行了对比，实验结果如表 3.2 所示。PaCoNet 在计算效率方面相较于传统卷积神经网络如 DenseNet^[67] 和 ResNet^[51]，展现出了显著的优势。以 DenseNet 为例，PaCoNet 的参数量仅为其六分之一，运算量则为其四十分之一。DenseNet 由于其密集连接的设计，虽然在特征复用和梯度流动方面有优势，但由于其较高的参数量和浮点运算量，使得计算开销较大，从而影响了其在实时性要求较高的应用中的表现。轻量级卷积网络如 MobileNetV2^[61] 和

MnasNet^[68]，在计算效率方面展现了一定的优势。MobileNetV2 采用深度可分离卷积，通过将传统卷积操作拆分为深度卷积和逐点卷积，将参数量和计算量降低至 2.3M 和 0.3G，从而提升了计算效率。MnasNet 则通过网络架构搜索和调整卷积核大小，进一步优化了计算流程，减少了不必要的计算开销，将参数量和计算量降低至 3.1M 和 0.31G。

混合模型通常结合了卷积神经网络和自注意力机制，以充分利用两者的优势。ViT-T^[72]和 DeiT-T^[73]通过引入自注意力机制来捕捉全局信息，从而提升复杂场景时的特征提取能力。然而自注意力机制通常带来较高的计算成本。ViT-T 的浮点运算量为 1.07G，吞吐量为 3168 IPS，但其推理时间相对较长，总时间消耗为 0.451 秒。DeiT-T 的浮点运算量和吞吐量稍低，但计算效率仍受限于自注意力机制的高复杂度，导致延迟为 9.74 毫秒，整体时间消耗为 0.460 秒。

针对混合模型计算成本高昂的问题，研究者们进一步探索了轻量化设计策略，这方面代表性的工作包括 MobileViT-XXS^[75]和 EdgeNeXt-XXS^[77]。MobileViT-XXS 巧妙地结合了轻量化卷积操作与局部化的自注意力机制，在尝试保留全局信息捕捉能力的同时，有效地将参数量和计算量分别降低至 1.0M 和 0.26G。而 EdgeNeXt-XXS 则采用了不同的策略，通过自适应地调整卷积核大小以及选择性地使用特征通道来进一步优化计算过程、减少冗余，将参数量和计算量降低至 1.2M 和 0.20G。

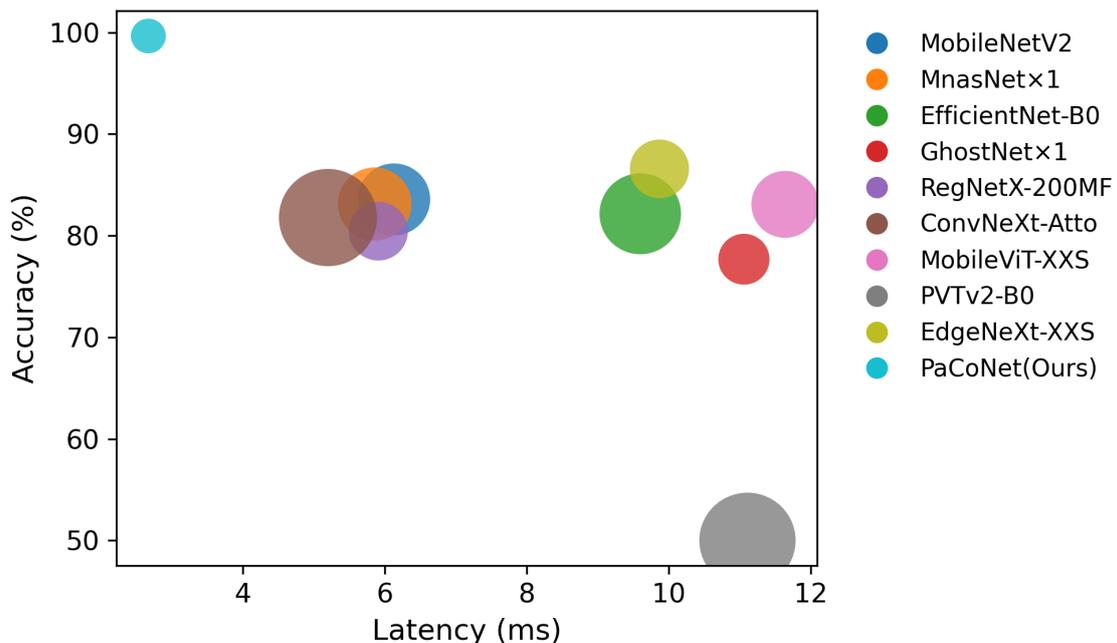


图 3.5 不同模型准确率、延迟与计算量的对比结果。

为了更直观地探讨不同模型在准确率、延迟与计算量方面的表现，本章选取了与 PaCoNet 参数量相近的多个模型并绘制了散点图，具体如图3.5所示。图中横轴表示延迟，纵轴表示准确率，点的大小代表了模型的计算量，各模型通过不同颜色进行区分。结果表明，MobileNetV2 和 MnasNet 等模型在延迟方面表现出一定的竞争力，其延迟大约在 5 至 6 毫秒之间。然而，这些模型的准确率未能满足防伪标签真实性预测任务的需求，分别为 83.57% 和 83.12%。防伪标签采用物理不可克隆函数技术生成，其内部微观纹理具有独特且无法复制的特征。由于这些细微差异在视觉上难以察觉，容易导致相似类别的样本发生误分类。MobileNetV2 和 MnasNet 在设计时侧重于降低计算量和提高实时性，没有关注如何有效捕捉材料图像中的微小纹理和复杂结构特征，从而导致在准确率上的不足。相比之下，本章通过在损失函数中添加额外角度边界进行约束，使类内样本更加聚集，类间样本更加分离。同时提出测地度量方法对防伪标签进行真实性预测，有效应对了高维非线性数据分布的挑战，从而取得了 99.79% 的准确率。

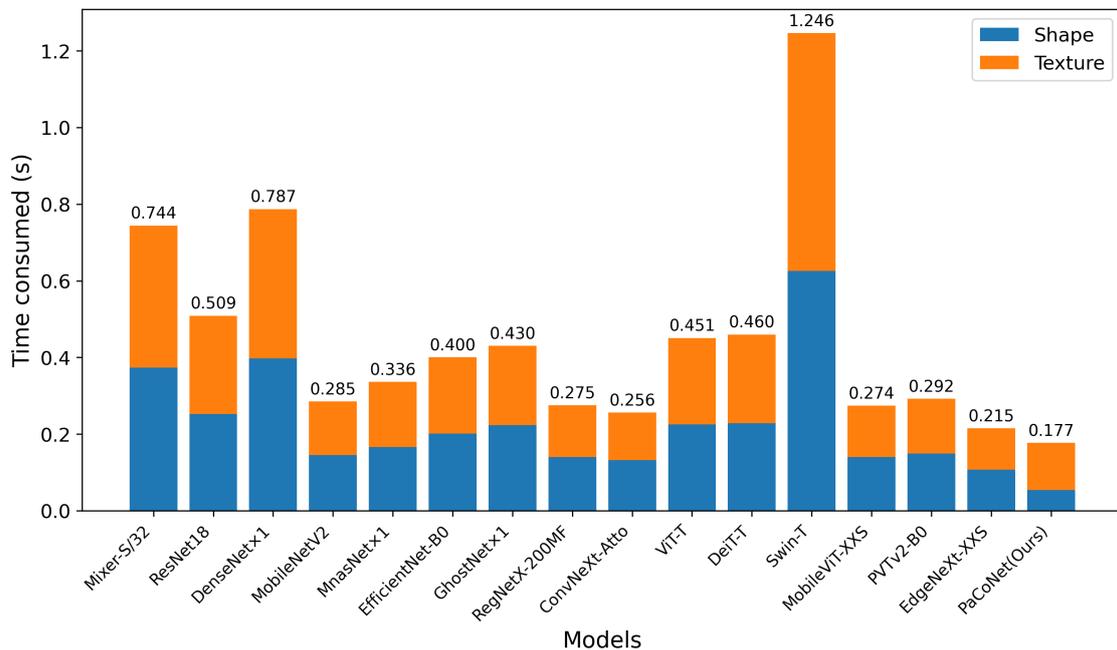


图 3.6 不同模型在形状识别和纹理识别任务中消耗时间的对比结果。

此外，图 3.6 比较了不同模型在执行形状与纹理特征提取时所需的时间。实验结果显示了各模型在处理效率上的差异，MobileViT-XXS 和 EdgeNeXt-XXS 作为先进的轻量化模型，展现出较为优异的处理效率，其总时间消耗分别约为 0.274 秒和 0.215 秒；相比之下，结构相对更复杂的 ResNet18 和 DenseNet 则耗时较长，分别达到了

0.509 秒和 0.787 秒，这反映了它们较高的计算复杂度，难以满足对实时性要求较高的应用场景。值得注意的是，本章提出的 PaCoNet 在所有参与比较的模型中表现最为出色，效率最高，其完成任务的总时间消耗仅为 0.177 秒，显著低于其他所有模型。

3.2.4 消融实验

为了探究本章提出的各个模块的有效性以及模块中超参数的影响，本章设计了一系列消融实验。

表 3.3 不同模块在分流卷积网络上的消融结果。

额外角度边界	测地度量	准确率	精确率	召回率	F1
×	×	55.85	42.12	15.59	22.69
✓	×	58.59	51.15	13.58	21.46
×	✓	90.22	81.46	99.91	89.59
✓	✓	99.79	99.92	99.59	99.75

表 3.3 通过消融实验，展示了额外角度边界和测地度量方法对分流卷积网络预测性能的影响。实验结果表明，未加入额外角度边界和测地度量方法的基准模型的准确率为 55.85%，精确率为 42.12%，召回率为 15.59%，F1 分数为 22.69%，说明基础网络结构所学习到的特征判别性有限，难以对视觉特征高度相似的 PUF 防伪标签进行准确预测。当单独引入额外角度边界损失进行训练后，模型性能没有获得显著的的提升。额外角度边界损失通过在高维流形中为样本与其对应类原型间的夹角添加一个固定的角度边界，直接优化特征间的角度距离，拉大类间间隔并压缩类内分布。然而，尽管额外角度边界损失优化了特征空间的区分度，但在缺乏适配判别机制的情况下，其潜力无法得到充分发挥。相比之下，单独引入测地度量方法使网络的性能得到显著改善。准确率提升至 90.22%，精确率达到 81.46%，召回率达到 99.91%，F1 分数达到 89.59%，证明了测地度量方法本身的有效性。该方法通过将特征投影到高维流形并运用测地距离进行度量判别，建立了一种对特征方向性敏感且对幅度变化鲁棒的判别方式。正是这种内在的判别特性，使得该方法即便作用于未经特定角度优化训练的特征空间，也能够达到良好的分类性能。

在此基础上进一步结合在额外角度边界损失与测地度量方法，模型的准确率提升至 99.79%，达到了理想的预测效果。这主要归因于测地度量方法与额外角度损失的优化目标相一致以及测地度量方法本身的优势。一方面，测地度量方法将特征投

影至高维流形并运用测地距离进行度量，这与额外角度边界损失在流形空间优化角度距离的目标相一致，使得基于测地距离的判别过程能充分利用额外角度边界损失优化得到的特征分布特性。另一方面，测地度量方法采用学习到的类原型作为几何判别中心，依据样本到其最近类原型的测地距离来实现对已知类别的精准分类。在此基础上，还通过将样本与最近类原型的最小测地距离与一个基于训练数据分布设定的阈值进行比较，从而能够有效拒识未知样本。相比于传统线性分类器或仅依赖 Softmax 置信度的方法，测地度量方法不仅能够更精确、鲁棒地反映样本与类原型的几何接近程度，而且还具备未知样本的拒识能力。

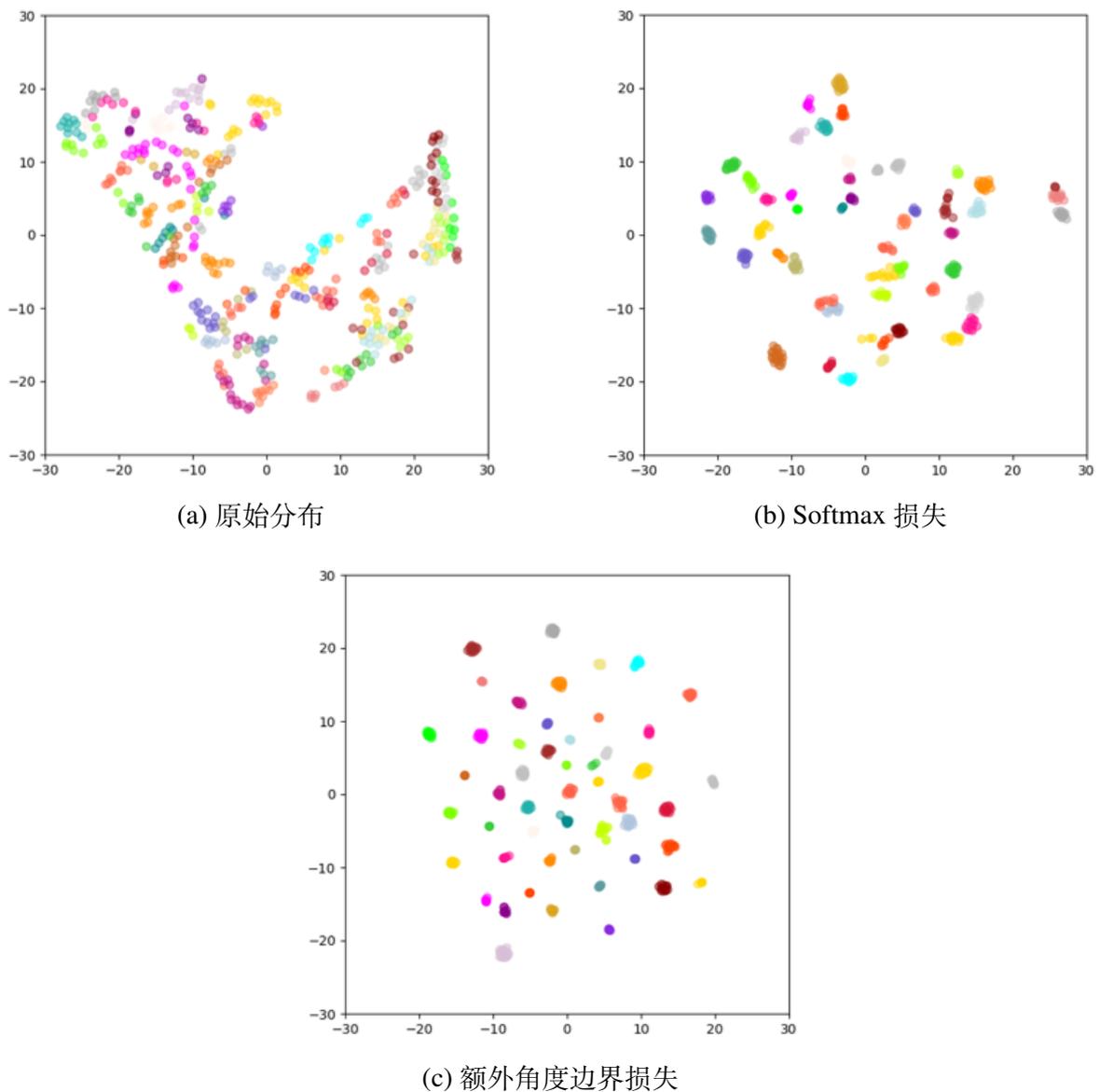


图 3.7 特征分布可视化对比。

为进一步探究额外角度边界在提升模型特征提取能力方面的具体作用，本章使用 t-SNE 方法^[78]，可视化分析了不同处理方法下特征的分布情况，如图 3.7所示。图 3.7(a) 展示了模型未经训练时提取到的原始特征的分布情况，此时数据点分布较为分散，类别间间隔较小，导致不同类别之间的重叠现象较为明显，使得原始特征不能被很好的区分。图 3.7(b) 呈现了模型在经过 Softmax 损失函数训练后生成的特征分布，经过 Softmax 处理，不同类别的特征之间的边界更加清晰，特征点的分布表现出更高的区分度；然而，尽管类别间的重叠有所缓解，仍存在部分交集，可能在分类任务中引发误分类风险。图 3.7(c) 进一步展示了在损失函数中引入额外角度边界约束后生成的特征分布，结果显示，特征点在特征空间中实现了更优的聚集效果，类间样本分离度显著提升，特征的可区分性得到进一步优化。额外角度边界直接优化特征与类原型之间的余弦距离，有效扩大了不同类别间的决策边界，同时促使同一类别内的样本在特征空间中更加紧密地聚合，显著提升了模型在高相似度场景下的判别能力，为特征提取的鲁棒性和分类精度提供了有力支持。

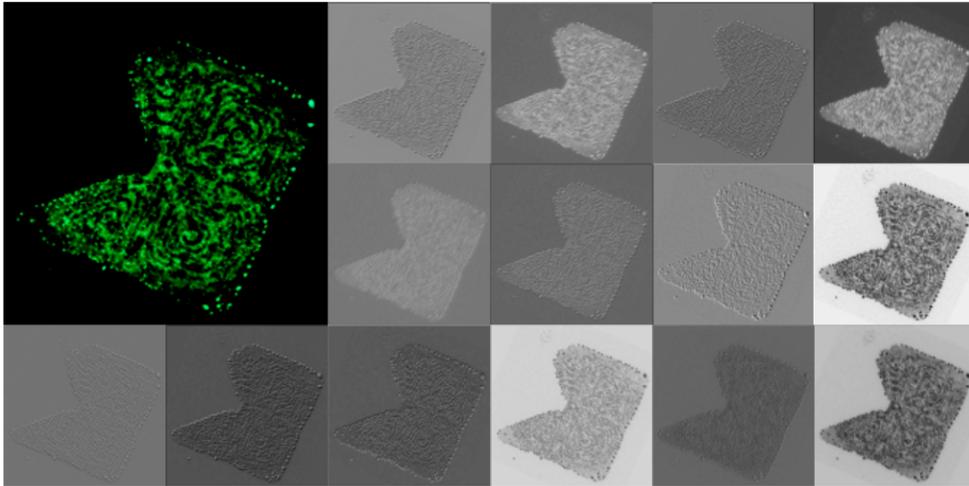


图 3.8 不同通道的中间特征可视化。

本章提出的分流卷积模块，对于输入的特征图沿通道维度将其划分为更新组和保留组，仅对更新组进行计算，其中更新组与保留组的通道之比被定义为分流比。这种设计的动机是跳过对潜在冗余通道的计算，从而优化模型整体的运算效率。图3.8展示了不同通道的中间特征可视化结果，其中左上角为原始图像，其他每张图代表了模型在阶段 1 中不同通道的特征输出。这些特征图展示了每个通道在提取图像不同方面的特征，如边缘、纹理和形状等细节信息。从结果可以看出，某些通道的灰度分

布高度相似，表明这些通道可能提取了相似或冗余的信息。冗余通道不仅没有为模型提供额外的有价值信息，还导致计算资源的浪费。

从更广泛的角度看，特征通道冗余是深度卷积网络中的一个普遍现象。随着网络层数和通道数量的增加，虽然模型的表达能力增强，但也更容易出现多个通道学习到功能相近或信息重叠特征的情况。这种冗余直接降低了计算效率，因为部分计算并未带来相应的性能增益，同时增加了模型的参数量和内存需求。

表 3.4 不同分流比的消融实验结果。

分流比	准确率	精确率	召回率	F1
1/2	89.18	80.95	99.84	88.94
1/4	97.37	94.31	99.86	96.93
1/8	99.79	99.92	99.59	99.75
1/16	97.60	94.79	99.84	97.21

表3.4探讨了不同分流比对模型预测性能指标的影响。本章提出的分流卷积模块，对于输入的特征图沿通道维度将其划分为更新组和保留组，仅对更新组进行计算，其中更新组通道数与总通道数之比被定义为分流比。实验结果表明，模型性能并非随通道参与比例的增加而单调提升。当分流比设置为 1/8 时，模型的准确率达到 99.79%，F1 分数达到 99.75%，为实验中的最优值。值得注意的是，采用更高的分流比，让更大比例的通道参与运算时，性能反而呈现下降趋势；例如分流比为 1/4 时，准确率降至 97.37%，而在分流比为 1/2 时，准确率降至 89.18%。

这种现象的原因在于信息有效性与潜在干扰。并非所有提取到的特征通道都包含对当前预测任务的有效信息。较高的分流比设置将信息含量较低、冗余甚至包含噪声的通道也加入了计算之中。这些非关键通道的加入干扰了模型对核心判别特征的有效学习，降低最具信息量通道的相对贡献度。因此，1/8 的分流比代表了在信息有效性与通道数量间取得最佳平衡的通道子集。同时，实验结果也表明过低分流比的局限性。当分流比为 1/16 时，准确率为 97.60%，因为过少的通道数量会限制模型的特征表征能力，无法捕获足够全面的信息。

表3.5针对额外角度边界的大小进行了消融实验，以分析不同取值对模型性能产生的影响。实验结果表明，当边界 m 设置为 0.2 时，模型取得了最佳的综合性能指标，准确率达到 99.79%，F1 分数达到了 99.75%。这表明 0.2 的边界提供了一个有效的正则化强度，引导模型学习到了既具有类内紧凑性又具有足够类间间隔的判

表 3.5 不同额外角度边界大小的消融实验结果。

边界大小	准确率	精确率	召回率	F1
0.2	99.79	99.92	99.59	99.75
0.4	98.47	96.73	99.89	98.24
0.6	71.55	59.44	99.90	74.53
0.8	68.04	67.80	51.79	54.65

别性特征。

将边界从 0.2 增大至 0.4 时，召回率微增至 99.89%，但准确率和 F1 分数出现了小幅下降，精确率降低至 96.73%。更大的角度边界 m 要求类别间在角度空间上被推开更远的距离。虽然这理论上有利于提升区分度，但也可能使得损失函数的优化目标变得更难达成。网络在优化过程难以收敛到全局最优状态。

当边界值 m 增大至 0.6 及以上时，模型的性能呈现显著下降。例如， m 等于 0.6 时准确率降至 71.55%，而 m 等于 0.8 时准确率降低至 68.04%。较大的 m 值对网络施加了更严格的几何约束，要求学习到的特征表示必须同时满足同类样本的高度聚集以及不同类别间存在一个较大的固定角度边界。过强的约束一方面增大了优化过程的复杂性，使得优化算法难以找到能同时满足所有苛刻条件的解；另一方面，较大的角度边界可能超出了模型在给定特征维度下的有效表征能力范围，在类别数量较多或数据本身存在较高相似性的情况下，几何上满足所有类别间存在大间隔变得困难。这两方面共同限制了有效的特征学习和模型的稳定收敛过程。

表 3.6 不同超参数 k 的消融实验结果。

k 的取值	准确率	精确率	召回率	F1	真负类率
0.1	91.93	100.00	80.90	88.95	100.00
0.2	98.55	100.00	96.59	98.24	100.00
0.3	99.50	100.00	98.83	99.41	100.00
0.4	99.73	99.99	99.36	99.67	100.00
0.5	99.79	99.92	99.59	99.75	99.94
0.6	99.71	99.56	99.73	99.64	99.69
0.7	99.24	98.38	99.80	99.09	98.84
0.8	98.19	95.96	99.85	97.86	97.01
0.9	96.61	92.58	99.90	96.08	94.29
1.0	94.55	88.56	99.92	93.86	90.75

表3.6探讨了超参数 k 对实验结果的影响。可以看出，当 $k = 0.1$ 时，模型表现出

较高的保守性。虽然真负类率达到了 100.00%，表明模型能够完全识别出伪造样本，但召回率仅为 80.90%，准确率为 91.93%，F1 分数为 88.95%。这意味着模型会将较多的真实样本误判为伪造样本，导致召回率和准确率偏低。随着 k 增加至 0.4，模型的性能显著提升。召回率大幅提高至 99.36%，准确率也提升至 99.73%，F1 分数为 99.67%。这表明，模型能够识别绝大多数真实样本，并且误判伪造样本的概率较低。在 $k = 0.5$ 时，召回率达到 99.59%，真负类率为 99.75%，准确率为 99.79%，F1 分数为 99.75%。此时，模型在识别真实和伪造样本方面都取得了较高的准确率，达到了一个较优的平衡点。当 $k = 0.8$ 时，召回率达到了 99.85%，表明几乎所有真实样本都能被正确识别，然而真负类率略有下降至 97.86%，准确率为 98.19%，F1 分数为 97.86%，表明虽然模型能够识别绝大部分真实样本，但误判伪造样本的风险增加。后续随着 k 值的增大，可以发现真负类率呈现持续下降的趋势，说明误判伪造样本的风险在不断增加。因此，在选择合适的 k 值时，需要在识别真实样本的能力和避免误判伪造样本的风险之间进行权衡。

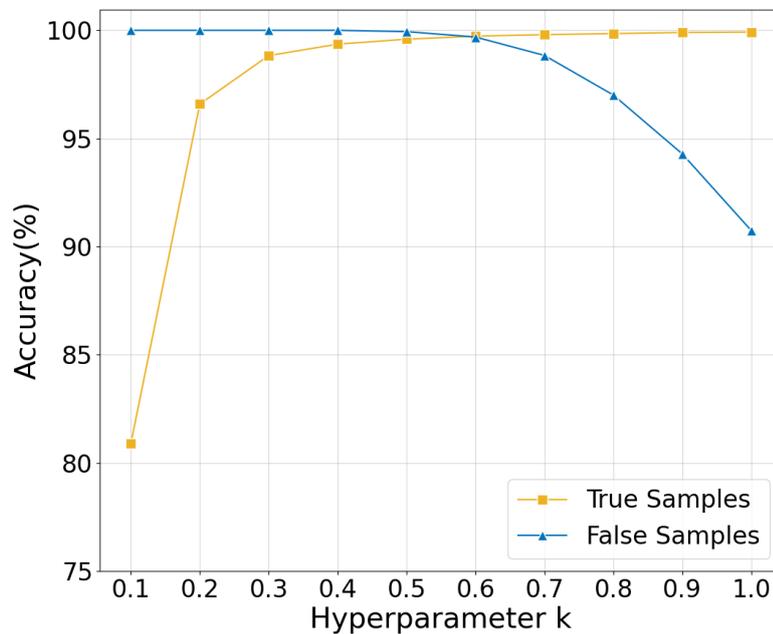


图 3.9 不同超参数 k 对准确率的影响。

图3.9展示了测地度量方法中关键超参数 k 不同取值对模型准确率的影响。其中蓝色曲线代表伪造样本，黄色曲线代表真实样本，横坐标代表超参数 k 的取值范围，纵坐标表示准确率，以百分比呈现。从黄色曲线可以看出，当 k 值较小时，存在较多

真实样本被漏检。随着 k 值增大, 该曲线迅速上升, 在 k 接近 1.0 时, 达到近 100% 的准确率。观察蓝色曲线可以看出, 在 k 小于等于 0.3 时, 伪造样本的预测准确率高达 100%, 表明模型能够完全预测出伪造样本。然而, 随着 k 值的增加, 蓝色曲线呈现出缓慢下降的趋势。在 $k = 0.5$ 时, 伪造样本的预测准确率为 99.94%, 而在 $k = 1.0$ 时, 该准确率下降至约 90.75%。这说明, 随着 k 值的增大, 模型将伪造样本错误预测为真实样本的风险逐渐增加。

从变化趋势上可以看出, 召回率对 k 值变化的高敏感性与真负类率的相对不敏感性, 造成这种不对称现象的原因在于模型特征空间与阈值判定的相互作用。增大 k 值会放宽接受样本为真实的标准, 使得更多真实样本落入接受域, 从而大幅提升召回率。而真负类率之所以能基本维持在较高水平, 关键在于本章在损失函数中添加了额外角度边界。额外角度边界促使模型学习到了判别力更强的特征表示, 类内特征高度聚集, 类间特征显著分离。因此, 即使判定为真实的边界有所扩大, 绝大多数伪造样本的特征点在高维流行中仍然远离所有已知类原型, 不易被错误接纳, 使得真负类率下降缓慢。综合考量最大化真实标签预测能力与最小化伪造标签误判风险这两个目标, 需要在两者之间进行权衡。当 k 取值 0.5 时, 召回率达到了 99.59%, 同时真负类率仍保持在 99.94%, 能够满足 PUF 防伪标签真实性预测任务对准确性与可靠性的双重需求。

3.2.5 不同防伪体系对比

表 3.7 不同防伪体系对比结果。

材料类型	验证设备	编码容量	处理时间 (秒/张)
长余辉荧光粉 ^[79]	智能手机 + 便携式显微镜	2.3×10^{442}	1.20
钙钛矿 SMFP ^[80]	/	1×2^{900000}	/
镧系元素掺杂剂 ^[81]	扫描共聚焦显微镜	6×10^{104}	/
滴铸间隙增强型拉曼标签 ^[82]	共聚焦拉曼系统	3×10^{15051}	12.00
喷墨印花花状图案 ^[29]	智能手机 + 便携式显微镜	4.7×10^{202}	/
二维皱缩材料 ^[31]	扫描电子显微镜	1×10^{144494}	0.12
随机 Au 网络 ^[37]	智能手机 + 便携式显微镜	1×10^{348}	6.40
响应性胶体光子晶体 ^[83]	特殊认证器	1×10^{2500}	2.00
钙钛矿荧光薄膜 (本章方法)	智能手机 + 便携式显微镜	1.7×10^{1956}	0.18

为评估钙钛矿 PUF 防伪体系的性能水平, 本章与多种不同防伪体系, 从防伪材料类型、验证设备、编码容量及处理时间等维度上进行了比较, 如表3.7所示。在验

证设备的要求方面，镧系元素掺杂剂、液滴间隙增强型拉曼标签及二维皱缩材料等防伪标签，其验证过程依赖于扫描共聚焦显微镜、共聚焦拉曼系统或扫描电子显微镜等大型、昂贵且通常限于实验室环境的专业设备，而响应性胶体光子晶体则需要专门设计的认证器。这些条件无疑限制了相关技术的广泛普及与应用。相比之下，本章提出的方法与长余辉荧光粉、喷墨印花花状图案、随机 Au 网络等技术，都采用了“智能手机 + 便携式显微镜”的组合方案，降低了设备门槛与部署成本，提高了系统的便携性与易用性，为现场验证及终端用户的广泛应用提供了可行性。

在编码容量方面，依据 Carro-Temboury 模型的评估方法^[81]，本章 PUF 标签的有效编码容量达到了 1.7×10^{1956} 。该数值显著超过了长余辉荧光粉、镧系元素掺杂剂、喷墨印花图案以及随机 Au 网络等多种技术。虽然与钙钛矿 SMFP、滴铸间隙增强型拉曼标签等少数几种技术的理论编码容量相比并非最高，但 1.7×10^{1956} 本身已是一个极其庞大的数值，足以满足大规模应用中确保标签唯一性的严苛需求。需要指出的是，不同研究可能采用不同的编码容量计算范式，在直接进行数值比较时需持谨慎态度。

在处理时间方面，本章的方法表现出显著的高效性，处理单个样本仅需约 0.18 秒，这一速度明显优于长余辉荧光粉、液滴间隙增强型拉曼标签、随机 Au 网络及响应性胶体光子晶体等多种技术。

综合来看，本章提出的钙钛矿 PUF 防伪体系在验证便捷性、编码安全性和处理效率这三个关键评价维度之间取得了良好的平衡。它结合了易于部署且成本相对较低的验证方式、足以保障大规模应用安全的高编码容量以及优于多数同类技术的快速处理能力，展现出作为一种适用于广泛实际部署场景、有竞争力的防伪解决方案的巨大潜力。

3.3 本章小结

为解决钙钛矿 PUF 防伪标签真实性预测任务中处理多层次特征、满足高效率和开放集识别方面的挑战，本章提出了一个基于单模态特征解耦的轻量级深度学习识别方法。通过特征解耦分离防伪标签宏观形状与微观纹理提取，以避免信息干扰并提升表示质量；设计基于分流卷积机制的 PaCoNet 轻量化网络，降低计算复杂度；在损失函数中添加额外角度边界以增强特征判别力；并提出测地度量方法实现开放集

环境下的未知样本的有效拒识。实验结果表明，所提出的 PaCoNet 模型在识别准确率与计算效率的综合表现上优于多种主流模型，消融研究也证实各项设计对模型性能的关键贡献。此外，通过与现有防伪体系的比较，进一步突显了本章构建的钙钛矿 PUF 技术体系在验证便捷性、高编码容量及处理速度等方面的实际优势。

第四章 基于多模态特征融合的轻量级复合材料性能预测方法

上一章探讨了针对单模态功能材料数据的特征处理与预测方法，本章的研究在此基础上进行拓展，将视角从单模态扩展至多模态场景。相较于第三章侧重于单一视觉模态内部复杂特征的解耦与提取，本章的核心目标转变为实现多模态异构数据的深度融合，通过结合材料的微观结构图像与包含其组分、工艺参数等的表格数据，以实现复合材料性能的精准预测。

材料性能预测是加速新材料发现、优化材料设计和缩短研发周期的关键技术环节。由于材料性能往往是其化学组成、微观结构、加工工艺等多种因素复杂作用的综合体现，有效的预测必须综合分析这些来源多样、性质各异的已知信息。近年来，多模态深度学习方法为材料性能预测提供了新的解决方案，通过融合不同来源的互补数据，模型能够从不同数据模态中捕捉它们之间的复杂关联与协同效应，从而构建对材料系统全面理解，以获得更准确的预测能力。然而，目前材料领域使用的多模态方法仍存在一些局限。首先，现有方法通常在特征提取阶段对各模态信息进行独立处理，直到后期才引入跨模态交互，无法充分挖掘互补信息。其次，现有的注意力融合策略缺少根据输入样本动态调整各模态信息重要性的自适应加权能力，导致融合效果不够理想或缺乏灵活性。最后，为了提升性能，一些先进的多模态融合模型往往参数量巨大，计算成本高昂，使得它们难以应用在需要快速预测或资源受限的实际材料研发场景中，实用性受到限制。

4.1 方法概述

本章提出了表格驱动特征融合网络 (Table-driven Feature Fusion Network, TFFN)，如图 4.1 所示，实现对海藻酸钙/石墨烯复合材料性能的高效精准预测。设计该网络的动机在于克服单模态预测的信息局限性，并解决现有融合方法中普遍存在的跨模态交互不足和计算效率低下的问题。表格驱动特征融合网络通过深度整合互补的图像与表格两种异构信息源，并结合创新的交互机制与轻量化设计来实现上述目标。

具体架构上，表格驱动特征融合网络包含两个并行的编码通路和一个最终的融

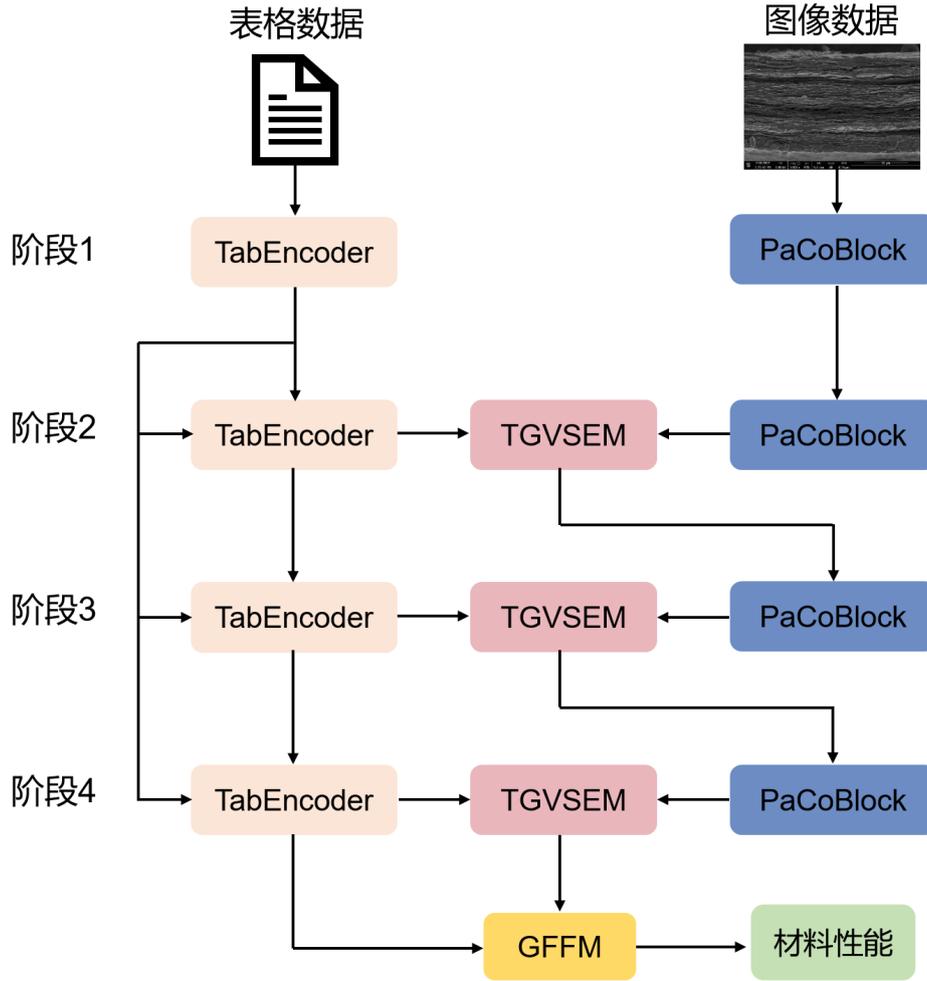


图 4.1 表格驱动特征融合网络示意图。

合模块。图像通路采用上一章提出的轻量化分流卷积模块（PaCoBlock）多阶段处理输入的微观结构图像，以保证计算效率。表格通路则使用多层全连接网络构成的表格编码器（TabEncoder）来编码包含材料成分、工艺参数等的表格数据。

表格驱动特征融合网络的创新与优势体现在其多阶段交互与最终融合机制上。为实现模态间早期且深度的交互，避免传统后期融合丢失过多早期信息的弊端，该网络在图像编码通路的中间阶段引入了表格引导视觉语义增强模块（Table-Guided Visual Semantic Enhancement Module, TGVSEM）。该模块接收来自同阶段图像编码器与表格编码器的输出特征作为输入，并利用表格特征作为引导信号，动态地调制与增强该阶段的视觉特征。随后，表格引导视觉语义增强模块的输出被送入下一阶段的图像编码器，这种设计确保了图像特征在逐层传递和学习中，能够尽早融入表格信息的上下文约束，从而获得更具目标导向性和语义相关性的视觉表征。

在最终的融合阶段，为实现两种模态深度特征的自适应整合，网络采用了门控

特征融合模块 (Gated Feature Fusion Module, GFFM)。该模块接收最后一个表格引导视觉语义增强模块输出的、经过多层交互增强的视觉特征与最后一个表格编码器输出的深度编码表格特征作为输入。门控特征融合模块负责对这两种深度特征进行最终的自适应融合：它根据数据特性动态调整融合权重，以有效捕捉跨模态协同效应，并避免简单融合方式可能带来的信息损失或偏重，最后输出性能预测结果。本节后续将详细介绍各组成部分及其具体实现细节。

4.1.1 表格引导视觉语义增强模块

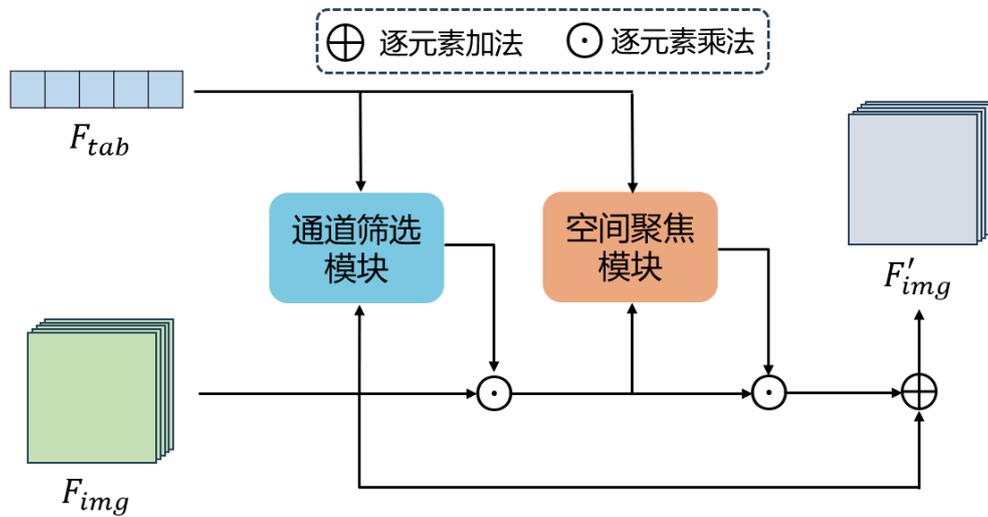


图 4.2 表格引导视觉语义增强模块。

为实现图像与表格特征之间的深层关联，本章提出了表格引导视觉语义增强模块，其结构如图4.2所示，由通道筛选模块 (Channel Filtering Module, CFM) 和空间聚焦模块 (Spatial Focusing Module, SFM) 构成。表格引导视觉语义增强模块利用表格数据对图像特征的通道响应和空间注意力分布进行动态调控，从而实现图像特征与表格语义的精准对齐。具体而言，通道筛选模块以图像特征 F_{img} 和表格特征 F_{tab} 作为输入。首先，对图像特征 F_{img} 应用两条并行的全局空间信息提取路径。如图4.3所示，一条路径计算空间平均池化 $AP_s(F_{img})$ ；另一条路径计算空间最大池化 $MP_s(F_{img})$ 。池化操作将输入图像特征压缩，生成包含每个通道全局信息的通道描述符。随后，将两个通过不同池化方法得到的通道描述符经过逐元素加法进行融合，得到聚合后的池化特征 F_{agg_pool} 。与此同时，原始的表格特征 F_{tab} 经过线性变化 F'_{tab} 。接下来，将聚合后的池化特征 F_{agg_pool} 与处理后的表格特征 F'_{tab} 沿通道维度进行拼接。构成了

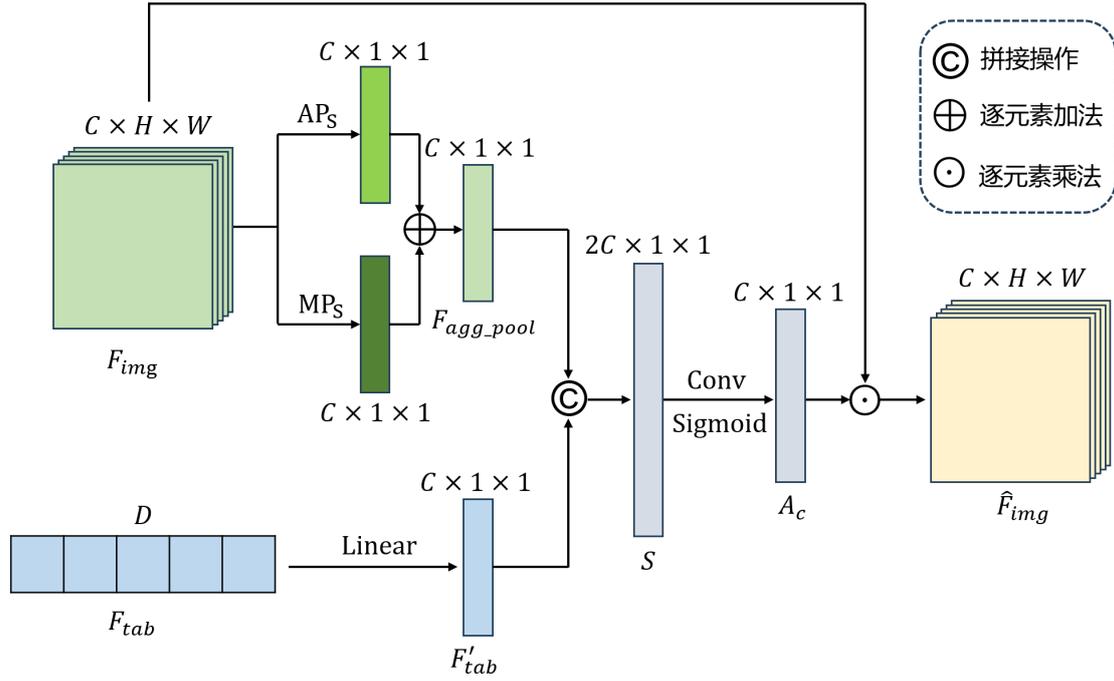


图 4.3 通道筛选模块。

一个新的混合特征向量 S 。

为了生成通道注意力权重，混合特征 S 通过一个 1×1 卷积层，对特征进一步处理并调整其通道数。卷积层的输出接着通过 Sigmoid 函数，Sigmoid 函数将输出值压缩到 0 到 1 的范围内，构成了通道注意力权重向量 A_c ，其计算过程可表示为：

$$A_c = \sigma(\text{Conv}(S)), \quad (4.1)$$

其中 Conv 表示卷积计算， σ 表示 Sigmoid 函数。最后，将计算得到的通道注意力权重 A_c 应用于原始的图像特征 F_{img} ，通过逐元素乘法将 A_c 中的每个权重值乘以 F_{img} 中对应通道的所有特征值，根据融合的表格信息和图像全局上下文动态地调整了每个特征通道的重要性，对通道进行了筛选和增强，其计算过程表示为：

$$\hat{F}_{img} = F_{img} \odot A_c, \quad (4.2)$$

其中 \odot 表示逐元素乘法。

空间聚焦模块以通道筛选后的特征 \hat{F}_{img} 以及表格特征 F_{tab} 作为输入。首先，将表格特征 F_{tab} 与图像特征 \hat{F}_{img} 沿通道维度进行拼接，生成中间特征 \tilde{F} 。接下来，对中间特征 \tilde{F} 并行应用通道平均池化 (AP_c) 和通道最大池化 (MP_c)。通道池化沿着 \tilde{F} 的通道维度聚合信息，生成一个单通道的空间图谱。随后，将通道平均池化的结果

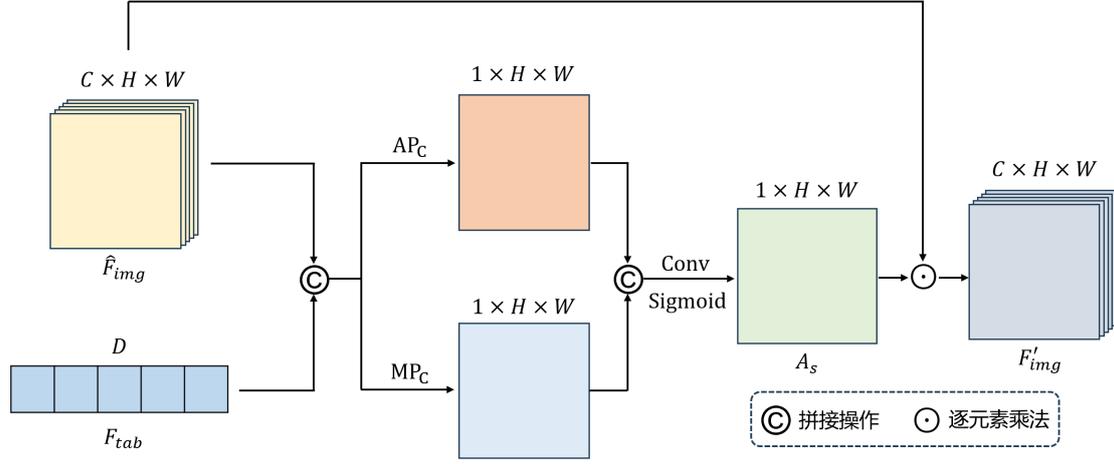


图 4.4 空间聚焦模块。

$AP_c(\tilde{F})$ 与通道最大池化的结果 $MP_c(\tilde{F})$ 再次沿通道维度进行拼接，形成一个包含两种池化信息的特征图。

为了生成空间注意力权重，将拼接后的特征图通过一个 7×7 卷积层，卷积层的输出结果再经过 Sigmoid 函数，计算得到单通道的空间注意力权重 A_s ，其中每个位置的值反映了对应空间位置的重要性。其计算过程可表示为：

$$A_s = \sigma\left(\text{Conv}\left([\text{AP}_c(\tilde{F}); \text{MP}_c(\tilde{F})]\right)\right), \quad (4.3)$$

其中 AP_c 表示通道平均池化， MP_c 表示通道最大池化，Conv 表示卷积计算， σ 表示 Sigmoid 函数。

最后，将计算得到的空间注意力权重 A_s 应用于图像特征 \hat{F}_{img} ，通过逐元素乘法，将空间权重 A_s 与 \hat{F}_{img} 的每个通道相乘。这一步骤调整了原始图像特征在不同空间位置的响应强度，实现了空间的聚焦和特征的调制。其计算过程表示为：

$$F'_{img} = \hat{F}_{img} \odot A_s, \quad (4.4)$$

其中 \odot 表示逐元素乘法。

与传统仅依赖图像内部统计信息来推断注意力权重的机制^[53]相比，本章提出的表格引导视觉语义增强模块引入了跨模态视角。传统注意力方法在缺乏外部信息指导时，可能无法准确识别出图像中哪些微观结构特征与表格数据所代表的特定组分或工艺条件存在关联，从而导致注意力分配的不准确。表格引导视觉语义增强模块利用表格数据作为先验知识，通过其内部的通道筛选与空间聚焦模块，对图像特征

进行目标导向的选择性增强。这意味着模型被引导去关注那些根据表格信息判断最为关键的细粒度结构特征，同时抑制可能无关或冗余的视觉信息。这种机制不仅能有效捕捉对性能预测至关重要的图像细节，更重要的是，它在特征提取的早期阶段就促进了不同模态信息的初步交互与语义对齐，能够更准确地反映出不同模态间的深层内在关联。

此外，该模块设计的快捷连接允许原始的视觉特征信息直接传递到模块输出端，有效缓解了深层网络中信息经过多重变换后可能出现的细节丢失或梯度消失问题，确保了特征表示的丰富性和训练的稳定性。

4.1.2 门控特征融合模块

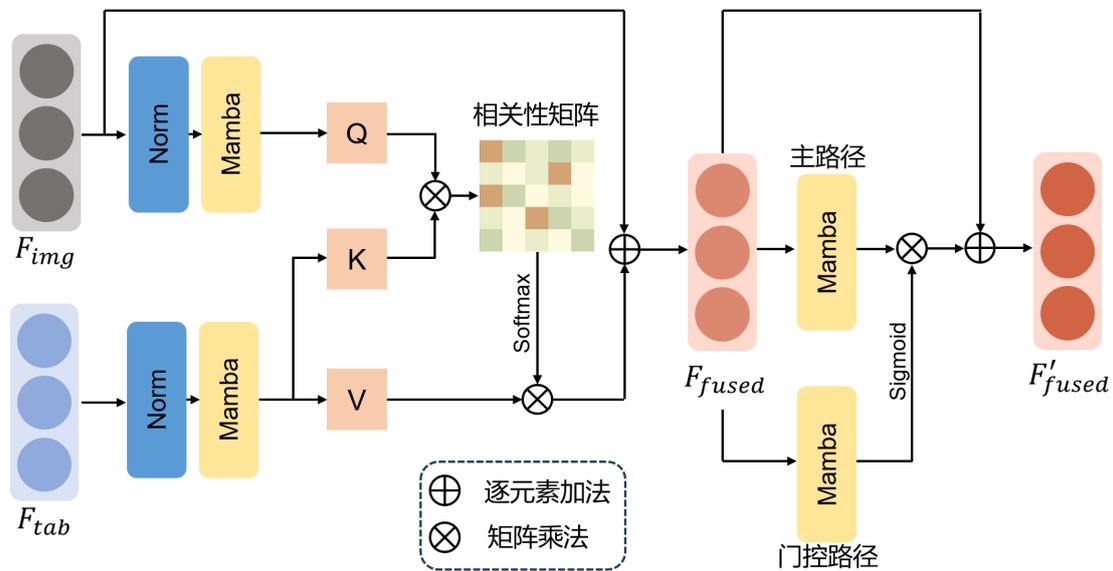


图 4.5 门控特征融合模块。

为了实现功能材料性能的准确预测，需要有效融合其异构的多模态数据，例如结合图像的微观结构与表格的组分工艺信息，以克服单一数据源的局限性。然而，这些数据在结构、维度和语义上的差异为融合带来挑战，简单的拼接或相加等基础策略难以捕捉关键的跨模态关联，无法充分发挥多源信息的潜力。

为应对异构数据深度融合的挑战并实现高效整合，本章提出了门控特征融合模块，模块的实现过程如图4.5所示，该模块主要由两部分构成：跨模态特征交互和门控前馈融合。在跨模态特征交互阶段，首先对图像特征 F_{img} 和表格数据特征 F_{tab} 分别进行归一化处理，并通过 Mamba^[55] 的轻量级变换进行初步特征提取。Mamba 作

为替代 Transformer 中标准自注意力机制的一种新架构，其优势在于计算效率，它借鉴状态空间模型并结合了输入依赖的选择性扫描机制，在有效建模序列数据中长距离依赖的同时，将计算复杂度从自注意力的二次方级别降低到了线性或近线性级别。特征交互的核心环节基于交叉注意力机制构建，将图像特征 F_{img} 通过线性变换得到查询向量 $Q = W_q F_{img}$ ，将表格数据特征 F_{tab} 经线性投影转换为键 $K = W_k F_{tab}$ 和值 $V = W_v F_{tab}$ 。计算查询 Q 与键 K 之间的相关性得分，生成注意力权重图，随后通过 Softmax 函数进行归一化，具体计算过程如下：

$$\text{Attention}(Q, K, V) = \text{Softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V, \quad (4.5)$$

其中 d_k 为缩放因子。接着，将归一化的注意力权重与值向量 V 进行矩阵乘法，得到经过注意力加权的特征表示 F_{fused} 。

在完成跨模态特征交互后，本章进一步采用门控前馈融合机制对得到的融合特征 F_{fused} 进行动态调控。该机制包含两个分支：主路径和门控路径。主路径对融合特征 F_{fused} 进行上下文相关的特征变换，生成表示 h ，其表达式为：

$$h = \text{Mamba}(F_{fused}). \quad (4.6)$$

门控路径则将 Mamba 的输出依次经由 GELU 激活函数和 Sigmoid 函数处理，从而生成动态权重序列 g ，其计算过程如下：

$$g = \sigma \left(\text{GELU}(\text{Mamba}(F_{fused})) \right). \quad (4.7)$$

最终，通过快捷连接将调制后的特征与原始输入融合，输出序列 F'_{fused} 表示为：

$$F'_{fused} = g \odot h + F_{fused}, \quad (4.8)$$

其中 \odot 表示逐元素乘法。

4.2 实验分析

本节介绍了使用的数据集与实验设置，并进行了对比实验和消融实验，以验证模型的有效性和鲁棒性。

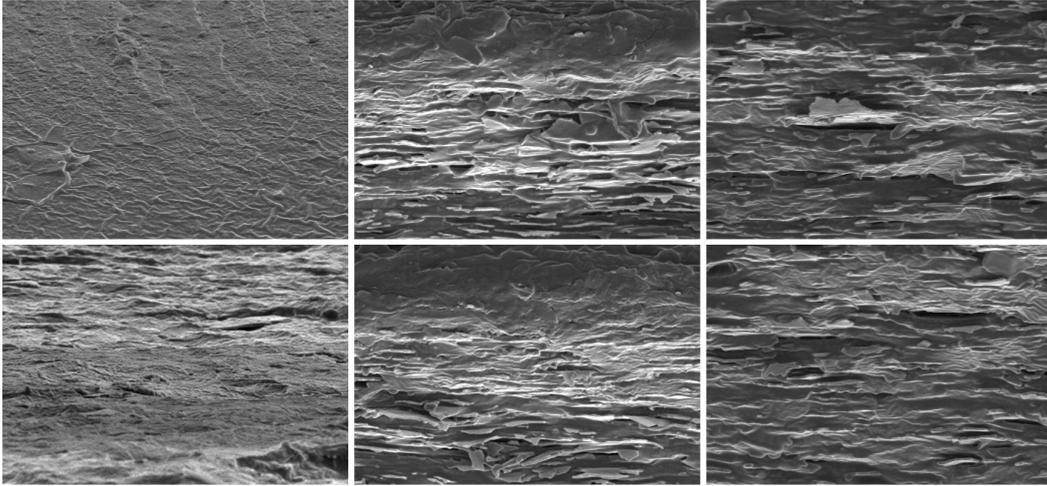


图 4.6 石墨烯/海藻酸钙复合材料图像示例。

4.2.1 数据集介绍

本章的数据集来自石墨烯/海藻酸钙复合材料。该材料采用超铺展策略制备，利用硅油与反应溶液间的界面张力差异，诱导反应溶液在硅油中产生剪切流力以实现定向铺展。反应溶液由石墨烯与海藻酸钠在水中的混合物组成，其中石墨烯的平均尺寸为 $2.5 \mu\text{m} \times 3.5 \mu\text{m}$ 。为探究不同石墨烯含量对复合材料性能的影响，其质量浓度分别设定为 0 wt%, 5 wt%, 9 wt%, 16 wt% 和 20 wt%。在制备过程中，反应溶液从距离基底 1 cm 的高度滴出，液滴先穿过 10 cm 高、粘度为 $1.79 \times 10^{-2} \text{ mPa} \cdot \text{s}$ 的空气介质，再进入 5 cm 高、粘度为 20 cSt 的硅油介质。复合材料的显微组织形貌通过扫描电子显微镜采集，同时还收集了样品的制备方式、石墨烯含量 (wt%)、厚度 (m)、密度 (g/cm^3)、比热容 ($\text{J}/(\text{kg} \cdot \text{K})$)、热扩散系数 (m^2/s) 以及取向因子等数据，并以表格形式呈现。

表格中共记录了 7 条数据，每条数据分别对应一个复合材料样品。针对每个复合材料，在不同位置拍摄了 1 至 3 张图像。这些图像的分辨率均为 1536×1103 像素，共计采集了 14 张高分辨率的灰度图像，如图 4.6 所示。

由于采集到的图像总量十分有限，为扩充数据集规模并提高模型的泛化能力，本章对图像数据进行了数据增强处理。具体而言，首先对每幅原始高分辨率图像进行裁切，将一幅图像分割成三幅子图；随后采用 Improved HP-VAE-GAN 模型^[84]进一步扩充每个子图，每幅子图生成 20 幅新图像，从而显著增加整体数据量，构建出一个更加丰富的数据集。

Improved HP-VAE-GAN 利用图像内部固有的局部重复性和自相似性特征，通过卷积神经网络将图像划分为大量细小补丁。在样本数量极为有限的情况下，该方法仍能充分提取出丰富的局部特征信息，从而为后续生成过程提供充分的数据支撑。

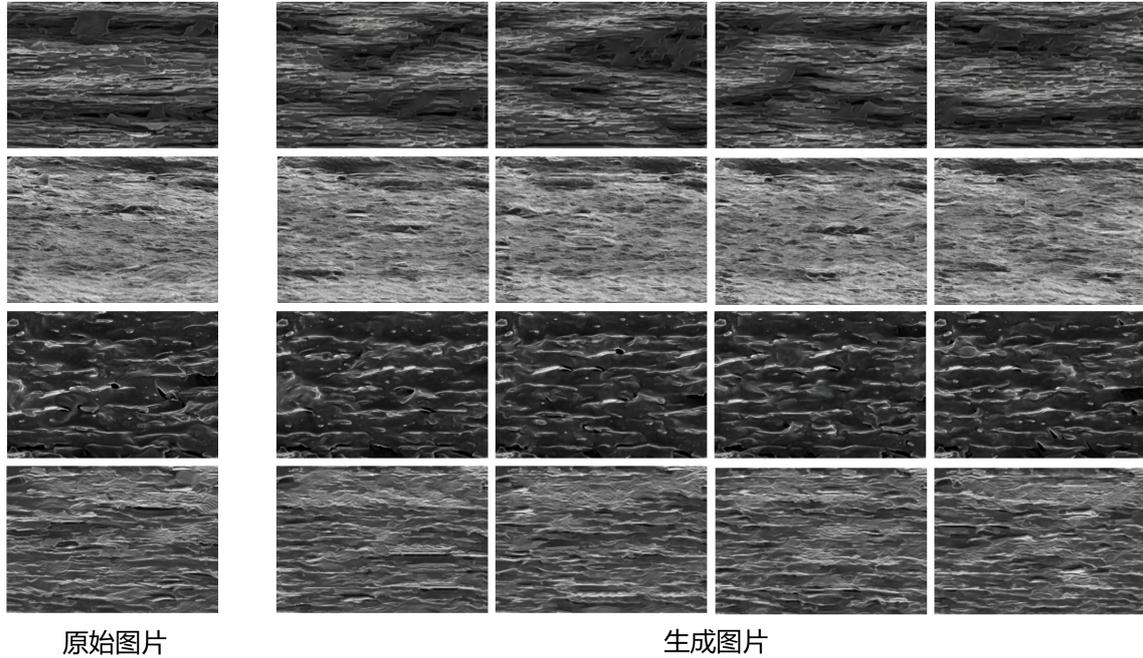


图 4.7 生成数据展示。

从图4.7可以观察到，生成图像在细节、纹理以及整体结构上均与原图保持较高的相似性，同时呈现出多样化的局部变化，较好地保留了原始复合材料的微观组织特征，例如石墨烯与海藻酸钠混合物形成的纹理及铺展特性。需要指出的是，生成的数据仅作为辅助训练使用，不参与最终测试。

4.2.2 实验设置

本实验在配备 24 GB 显存的 NVIDIA GeForce RTX 3090 GPU 平台上进行，操作系统为 Ubuntu 20.04，并采用 PyTorch 框架实现模型训练。为避免因显存限制引起的溢出，所有训练图像均在输入模型前统一裁剪至固定尺寸 224×224 。批量大小设置为 128。训练过程中选用优化器 AdamW，学习率采用余弦退火调度策略，初始学习率设定为 0.001。

4.2.3 对比实验

表 4.1 不同模型在石墨烯/海藻酸钙复合材料数据集上的性能对比。最好和次好的结果分别以粗体和下划线标出。

网络模型	图像	表格	MAE	RMSE	R ²	延迟 (ms)
Trompt ^[85]	×	✓	0.040	0.051	0.440	7.444
ResNet-Table ^[86]	×	✓	0.043	0.073	-0.161	1.344
MLP ^[87]	×	✓	0.065	0.107	0.432	0.129
MNv4-Conv ^[88]	✓	×	0.053	0.078	-0.333	8.639
ConvNeXt V2-L ^[89]	✓	×	0.079	0.096	-1.010	30.303
RDNet-T ^[90]	✓	×	0.069	0.091	-0.813	10.029
MMCL ^[91]	✓	✓	<u>0.025</u>	<u>0.030</u>	<u>0.805</u>	141.660
MLP-CNN ^[92]	✓	✓	0.036	0.054	0.360	<u>0.896</u>
MFMPPM ^[93]	✓	✓	0.033	0.047	0.519	2.857
MDFM ^[48]	✓	✓	<u>0.025</u>	0.034	0.742	3.246
TFFN(Ours)	✓	✓	0.018	0.024	0.873	6.059

为充分验证所提出的表格驱动特征融合网络的有效性，本章与多个当前最先进的模型进行了比较分析。这些对比模型包含了不同的模态输入类型，具体而言，可以划分为三类：第一类是仅利用表格数据进行预测的表格模型，第二类是仅依赖图像数据进行预测的视觉模型，第三类是同时整合表格和图像信息进行预测的多模态模型，如表4.1所示。为了对所有对比模型的性能进行量化评估，本章采用了被广泛认可的评估指标，包括平均绝对误差（MAE）、均方根误差（RMSE）以及决定系数（R²）。此外，还记录了每个模型的延迟，以评估其在实际应用中的效率。

在仅采用表格数据作为输入的模型中，本章对比了包括 Trompt^[85]、ResNet-Table^[86]与 MLP^[87]在内的模型。这些模型仅通过结构化的表格数据完成预测任务。例如，Trompt 模型在 R² 指标上表现为 0.440，延迟为 7.444 毫秒，表明其对表格数据的拟合能力有限。ResNet-Table 模型则表现出较差的预测性能，R² 值为-0.161，表明其更适合处理空间结构的数据。相比之下，MLP 模型取得了 0.432 的 R² 值，且延迟最低，仅为 0.129 毫秒，显示出在速度上的优势。然而，上述模型的整体预测精度普遍不佳，表明了仅依赖表格数据预测海藻酸钙/石墨烯复合材料导热性能的局限性。尽管表格数据能够精确反映填料量、制备工艺等关键的定量信息，但复合材料的导热性能还取决于材料内部由填料形成的导热通路微观结构及其连通性。这些空间结构特征和分布信息是表格数据无法充分描述的，通常需要借助微观组织结构图

像等视觉信息进行表示。因此，仅依据表格数据中的定量信息，难以准确预测复合材料的导热性能。

在仅采用图像作为输入进行预测的模型中，本章对比了包括 MNv4-Conv^[88]、ConvNeXt V2-L^[89]与 RDNet-T^[90]。实验结果显示，MNv4-Conv 模型的 R^2 值为-0.333，延迟为 8.639 毫秒，性能表现欠佳。ConvNeXt V2-L 模型的 R^2 值为-1.010，延迟达到 30.303 毫秒，表明其在此任务上性能较差且计算成本较高。RDNet-T 模型的 R^2 值为-0.813，延迟为 10.029 毫秒，性能虽优于 ConvNeXt V2-L，但仍显著低于理想水平。这些结果表明，仅依赖图像数据进行预测同样面临挑战，不同视觉模型间的性能差异可能源于其网络结构和学习策略，但总体而言，它们都难以有效提取足够的相关信息以实现准确预测。尽管图像数据能够提供关于复合材料微观结构的信息，例如石墨烯的分布和导热通路结构，但是通常很难仅从二维显微图像中准确地反推出填料量等对于导热性能非常重要的精确数值。关键定量信息的缺失使得仅基于图像的预测模型难以建立特征与材料性能之间的深层关联，从而根本上限制了其预测精度。

在同时整合表格和图像信息进行预测的多模态模型中，本章选取了多种的模型进行对比分析，包括在材料领域的多模态模型 MLP-CNN^[92]、MFMPMM^[93]、MDFM^[48]，以及源自计算机领域、基于对比学习策略的 MMCL 模型^[91]。为更加直观地展示模型的预测值与实际值间的差异，本章对结果进行了可视化，如图4.8所示。其中横坐标为真实导热率，纵坐标为模型预测值，散点在 1:1 线周围的分布越紧密，说明模型在该区间的拟合能力越好。MLP-CNN 将表格特征和 CNN 提取的图像特征通过向量拼接的方式进行融合，能够在中等导热率范围内取得一定程度的精确预测，但对于导热率较高的样本普遍存在低估，导致散点在高值区间明显偏离 1:1 线，从而拉低了整体 R^2 。MFMPMM 借助注意力机制在单一模态内部进行更深度的特征挖掘。相比 MLP-CNN，其中在低导热率区间的误差分布有所改善，散点更贴近基准线。然而在高值样本上，仍能观察到某些偏离明显的散点，说明对部分样本分布的适应性仍有欠缺。MDFM 利用门控机制对表格与图像特征的融合进行了优化，因而其散点分布更为紧凑，尽管如此，对高值样本预测偏低的问题依然存在。MMCL 利用跨模态对比损失函数来学习图像与其他模态特征之间的相似性与差异性，从而实现特征对齐。从散点图中可以观察到，预测点的整体分布并不十分紧凑，甚至在部分区间表现出明显的离散性，但模型取得了相对较高的 R^2 值。这种差异源于 R^2 指标本身的计算特性，它主要衡量预测值与真实值之间的线性相关程度以及模型对总体数据

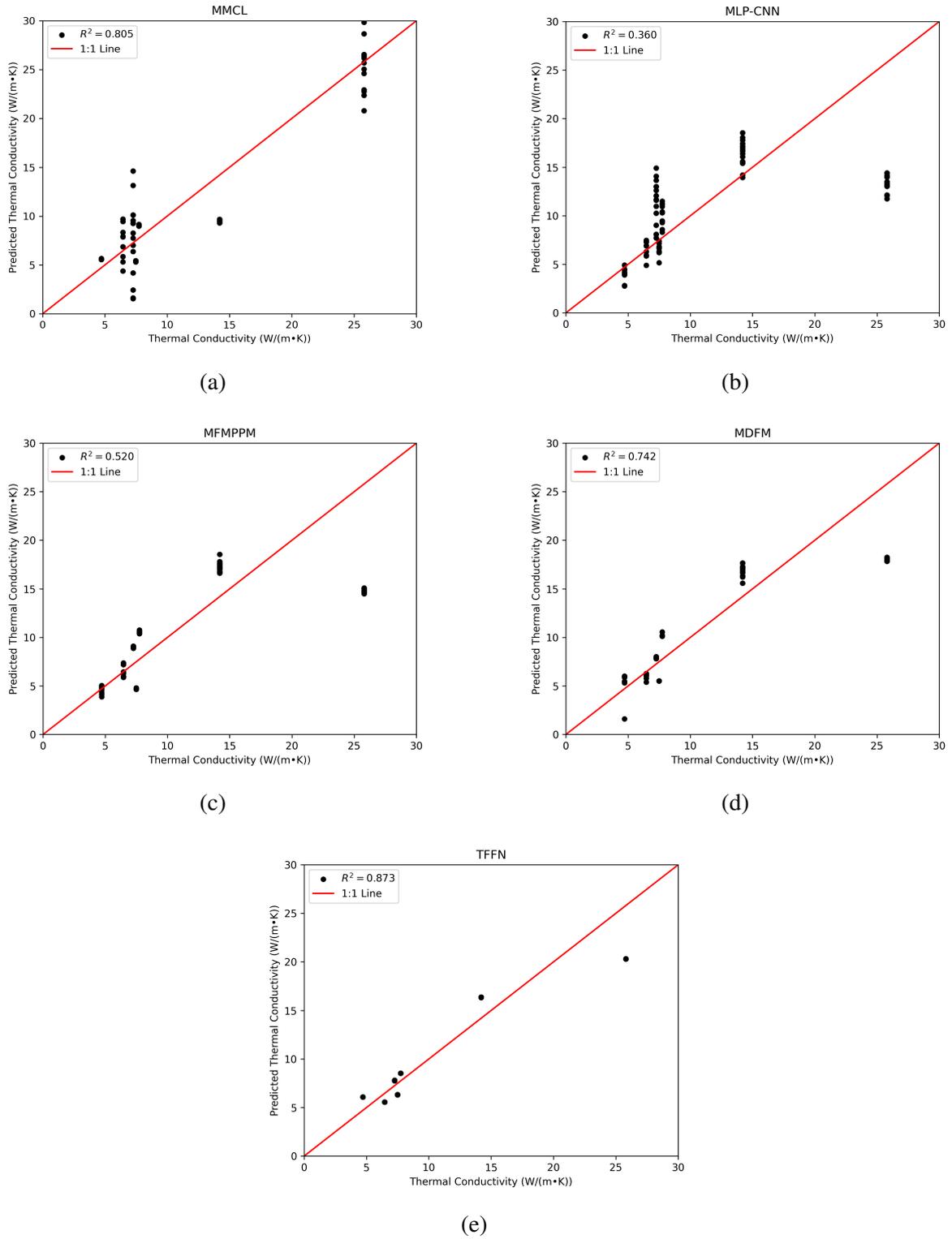


图 4.8 模型导热率预测性能评估图。

方差的解释能力，而非直接反映单个样本的绝对预测误差大小。换言之，MMCL 较好地捕捉了数据的整体变化趋势从而得到较高的 R^2 ，但其个体预测的精确度相对不足。这也表明，在评估模型性能时，仅依赖 R^2 可能会掩盖个别样本预测偏差较大的问题，因此需要结合其他误差指标以进行更为全面的分析。

本章提出的表格驱动特征融合网络通过在图像编码器的不同层级中加入表格引导视觉语义增强模块，将表格信息逐层地融入图像特征的提取过程，实现了跨模态引导与语义增强；同时，在门控特征融合模块中，通过动态调控融合权重来实现两类模态深度特征的自适应互补整合，最终构建出信息丰富的联合特征用于性能预测。得益于上述设计，表格驱动特征融合网络在性能上表现突出， R^2 达到了 0.873，在 MAE 和 RMSE 指标上也表现出最低的预测误差。表格驱动特征融合网络的散点图显示，各导热率区间内的预测点均分布在理想的 1:1 基准线附近，证明了模型对个体样本误差的有效控制。计算效率方面，表格驱动特征融合网络的延迟虽高于结构简单的 MLP-CNN，但需注意的是 MLP-CNN 的预测精度严重不足。因此，考虑到精度优势，表格驱动特征融合网络的计算效率处于实用且可接受的范围内，体现了其在精度与效率之间的良好平衡。

4.2.4 消融实验

表 4.2 不同模块在表格驱动特征融合网络上的消融结果。

表格引导视觉语义增强模块	门控特征融合模块	MAE	RMSE	R^2
×	×	0.024	0.034	0.742
✓	×	0.021	0.029	0.817
×	✓	0.024	0.033	0.753
✓	✓	0.018	0.024	0.873

为了验证所提出的表格引导视觉语义增强模块 (TGVSEM) 和门控特征融合模块 (GFFM) 的有效性，本节进行了消融实验，以评估每个模块对模型整体性能的独立贡献以及它们之间的协同作用，如表4.2所示。

对比不包含 TGVSEM 和 GFFM 的基础模型和仅加入 TGVSEM 的模型，可以看出 TGVSEM 的加入显著提升了模型性能，MAE 从 0.024 降低至 0.021，RMSE 从 0.034 降低至 0.029， R^2 从 0.742 提升至 0.817，证明了 TGVSEM 的有效性。上述性能的提升归因于 TGVSEM 的机制设计，通过其内部的通道筛选和空间聚焦组件，模

块能够有效利用表格数据对图像特征进行动态调制与选择性增强。这不仅在特征提取的阶段实现了跨模态语义信息的有效对齐，而且能够引导模型将注意力精确地集中在那些与当前表格信息上下文最相关的视觉特征通道和关键的空间区域上。

对比不包含 TGVSEM 和 GFFM 的基础模型和仅加入 GFFM 的模型，GFFM 单独带来的性能提升较为有限。这表明，GFFM 的有效性依赖于输入特征的质量和相关性。尽管 GFFM 本身设计了跨模态交互和门控融合机制，但如果直接处理未经 TGVSEM 增强和引导的原始特征，这些特征可能缺乏足够的跨模态相关性与对齐性，导致 GFFM 的融合机制难以有效运作，导致其融合效果受到限制。

当 TGVSEM 和 GFFM 两个模块结合使用时，模型取得了最佳的整体预测性能，所有评价指标均达到了最优值，不仅显著优于不包含这两个模块的基础模型，也大幅超越了仅使用 TGVSEM 或仅使用 GFFM 的单一模块配置，体现出 TGVSEM 与 GFFM 之间的协同作用。TGVSEM 利用表格数据优化了视觉特征，使其携带了更明确的语义信息并具备了更好的潜在对齐性。这为后续的 GFFM 模块提供了理想的输入基础，使得 GFFM 的交叉注意力与门控融合机制能够更精准、更高效地运作，以捕捉深层跨模态关联并实现最佳的自适应整合效果。

表 4.3 TGVSEM 子模块消融实验结果。

通道筛选模块	空间聚焦模块	MAE	RMSE	R^2
×	×	0.024	0.033	0.753
✓	×	0.023	0.028	0.819
×	✓	0.024	0.031	0.783
✓	✓	0.018	0.024	0.873

为了验证 TGVSEM 中通道筛选模块 (CFM) 和空间聚焦模块 (SFM) 的有效性，本章进行了更细致的消融实验。实验结果如表4.3所示。

仅采用 CFM 的模型对比未使用 CFM 和 SFM 的基础模型，性能呈现出一定的提升，MAE 降低至 0.023，RMSE 降低至 0.028， R^2 提升至 0.819。这是因为 CFM 通过对图像特征进行空间池化，随后将池化结果与表格特征进行拼接，并生成通道注意力权重。此过程使得模型能够依据表格数据的重要性，对图像特征的不同通道赋予相应的权重，从而筛选出与表格信息更为相关的视觉特征通道，使得模型能够更加关注携带关键信息的视觉通道。

仅采用 SFM 的模型对比未使用 CFM 和 SFM 的基础模型，性能亦有所提升，但

提升幅度相对较小, RMSE 降低至 0.031, R^2 提升至 0.783, 而 MAE 与基础模型持平, 为 0.024。SFM 将经 CFM 处理后的特征和原始表格特征在通道维度上进行拼接, 随后通过通道池化和卷积操作生成空间注意力权重。根据表格数据, 对图像特征的不同空间区域赋予相应的权重, 使得模型能够聚焦于与表格信息相关的特定区域, 从而辅助模型关注图像中关键的局部结构信息。而提升幅度相对较小的原因在于, 影响海藻酸钙石墨烯复合材料的导热性能的信息, 例如导热通路的排布方式更多地体现在图像的全局性特征通道响应上, CFM 能够更有效地捕捉这类信息。虽然 SFM 关注的局部空间结构也对性能有影响, 但其独立作用不如直接进行通道筛选的 CFM 显著。

当 CFM 和 SFM 同时被采用时, 模型展现出最优的性能。MAE 进一步降低至 0.018, RMSE 降低至 0.024, R^2 提升至 0.873。这表明 CFM 和 SFM 能够实现有效的协同工作。CFM 首先在通道层面筛选出重要的视觉信息, 随后 SFM 在空间层面进一步聚焦于这些通道中与表格数据相关的区域。这种先通道筛选后空间聚焦的策略, 使得模型能够更为精确地对齐图像和表格数据之间的语义信息, 从而更有效地提取跨模态的联合特征, 最终提升预测性能。此外, 模块中采用了快捷连接, 保留了原始的视觉特征信息, 有助于缓解深层网络可能出现的信息丢失问题, 进一步增强了模型的鲁棒性。

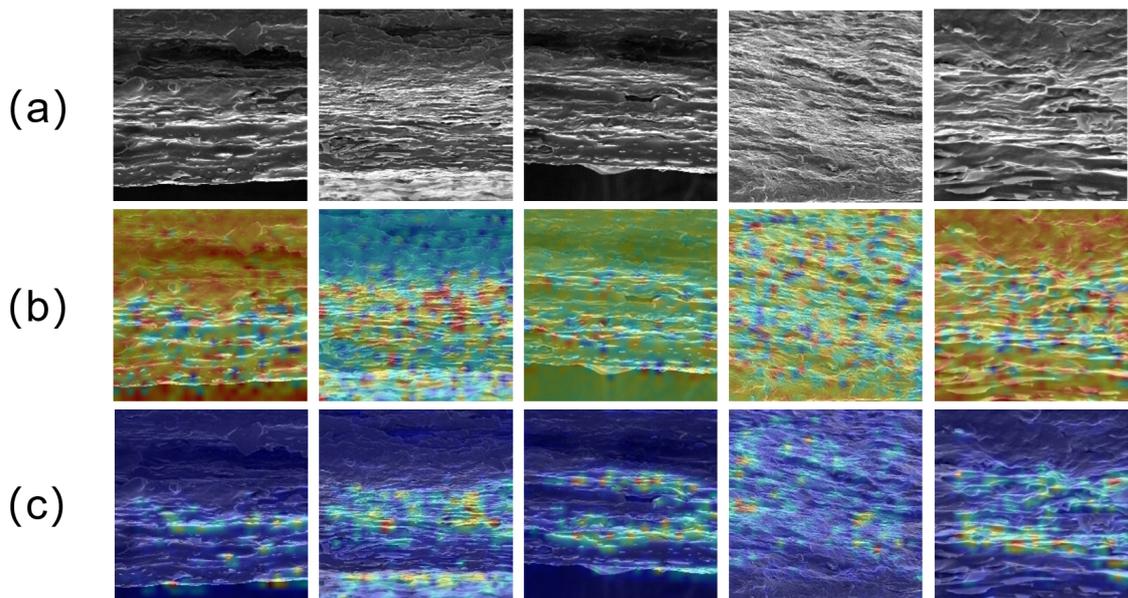


图 4.9 注意力可视化结果图。(a) 原始图像。(b) 未加入 TGVSEM。(c) 加入 TGVSEM。

为了更直观地理解 TGVSEM 的作用机制，本章可视化了表格驱动特征融合网络阶段二输出的特征注意力图，如图4.9所示。图4.9 (b) 可以看出，在未加入 TGVSEM 时，阶段二提取的特征图的注意力分布呈现出较为弥散的状态，主要响应图像的边缘、纹理等视觉模式，由于缺乏与表格数据中定量信息的关联，这些特征可能无法有效地聚焦于对材料性能预测至关重要的微观结构。图4.9 (c) 展示了加入 TGVSEM 后的注意力图，当加入 TGVSEM 后，特征图在通道维度上通过通道筛选模块实现了对与表格数据更相关信息的增强，并在空间维度上通过空间聚焦模块将注意力集中于图像中关键的微观结构区域，例如石墨烯的分布和导热通路，从而使得特征表达更具判别性。TGVSEM 利用表格数据作为先验知识，引导模型从复杂的背景中识别并聚焦于最具信息量的细粒度视觉特征，最终为提升性能预测的准确性提供了关键支撑。

表 4.4 不同阶段加入 TGVSEM 模块的结果对比。

配置	浅层阶段	中层阶段	深层阶段	MAE	RMSE	R ²
Config1	×	×	×	0.024	0.033	0.753
Config2	✓	×	×	0.022	0.029	0.810
Config3	×	✓	×	0.021	0.028	0.822
Config4	×	×	✓	0.023	0.031	0.794
Config5	✓	✓	×	0.022	0.029	0.818
Config6	✓	×	✓	0.023	0.031	0.789
Config7	×	✓	✓	0.018	0.024	0.873
Config8	✓	✓	✓	0.025	0.033	0.757

为探究表格驱动特征融合网络中采用表格引导视觉语义增强模块 (TGVSEM) 的最优策略，本章对不同阶段组合进行了消融实验。表格驱动特征融合网络采用上一章提出的分流卷积网络作为图像编码器。分流卷积网络采用层级设计，可以划分为四个阶段：阶段 1、阶段 2、阶段 3 以及阶段 4。在此基础上，本章进一步将阶段 1 定义为浅层阶段，阶段 2 与阶段 3 定义为中层阶段，而阶段 4 则定义为深层阶段。基于这一阶段划分，通过在不同单一阶段以及阶段组合中引入 TGVSEM，以评估其对模型性能的影响。具体的配置及对应的性能指标如表4.4所示，同时在图4.10中对各项指标进行了可视化对比。

实验结果表明，在任何单个阶段引入 TGVSEM，都能在一定程度上提升模型性能，其中在中层引入时的性能提升最为显著，MAE 下降至 0.021，RMSE 下降至 0.028，

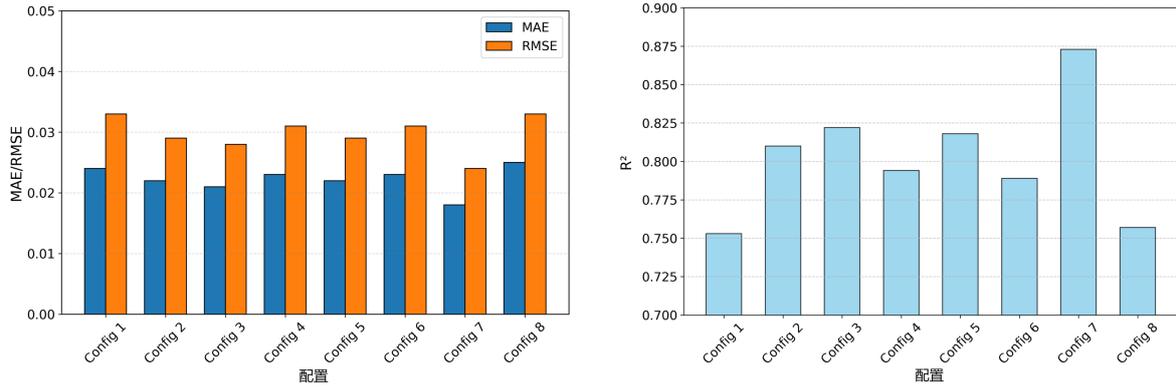


图 4.10 不同阶段加入 TGVSEM 模块的结果可视化对比。

R^2 达到了 0.822。这一结果表明，在浅层引入时，由于该阶段主要捕捉的是如边缘、纹理等低级细节，这些特征的信息量较大但语义表达较弱，因此虽然 TGVSEM 能对低级信息进行动态调控，但往往会受到噪声和冗余低级特征的干扰，表格语义引导作用有限。相比之下，深层特征主要反映全局高阶语义信息；在深层引入 TGVSEM 会导致有效的表格信息融入前破坏了编码器学习到的高级全局语义结构，进而影响预测性能。而中层特征则兼具局部细节和一定的语义表达能力，从而更有效地支持图像特征与表格信息之间的动态交互与融合。因此，仅在中层引入 TGVSEM 能实现更精准的注意力调控和特征增强，最终带来最佳的性能提升。

进一步的阶段组合实验中，在中层和深层引入 TGVSEM 的配置表现显著优于其他组合，MAE 下降至 0.018，RMSE 下降至 0.024， R^2 达到了 0.873。中层阶段不仅保留了丰富的局部细节，还具备较强的语义表达能力，使其成为图像特征与表格数据交互与融合的理想阶段；而深层则负责对中层筛选后的特征进行全局语义再聚焦，通过进一步优化整体特征表达，增强模型对关键信息的捕捉能力。

在浅层与中层同时引入 TGVSEM 的配置未能取得相同的提升效果。这主要是由于浅层阶段主要提取低级纹理、边缘等信息，尽管这些信息对初步特征提取有一定帮助，但在浅层引入 TGVSEM 后，由于对低级特征的调控较为粗糙，放大了信息冗余或噪声，从而对后续中层的特征学习产生负面影响。

相比之下，其他多阶段组合策略的效果不如“中层 + 深层”的配置。例如，“浅层 + 中层”的配置未能超越“中层 + 深层”的性能，且与仅在中层引入的配置相比提升有限。在浅层引入 TGVSEM 会传播语义信息不足或含噪声较多的特征，对后续中层阶段的特征学习产生不利影响。类似地，“浅层 + 深层”的配置的表现也逊于“中

层 + 深层”。此外，“浅层 + 中层 + 深层”的配置导致了性能的显著下降，MAE 达到了 0.025，RMSE 达到了 0.033， R^2 降低至 0.757，这也表明，错误的应用 TGVSEM 会导致早期噪声的放大以及对编码器内部特征表达的潜在破坏。

表 4.5 数据增强结果对比。

网络模型	原始数据			原始数据 + 生成数据		
	MAE	RMSE	R^2	MAE	RMSE	R^2
MMCL	0.041	0.045	0.557	0.025	0.030	0.805
MLP-CNN	0.046	0.055	0.344	0.036	0.054	0.360
MFMPPM	0.070	0.084	-0.537	0.033	0.047	0.519
MDFM	0.060	0.084	-0.543	0.025	0.034	0.742
TFFN(Ours)	0.041	0.057	0.279	0.018	0.024	0.873

由于原始图像样本量仅为 14 张，训练过程中模型极易出现过拟合现象，从而限制了泛化能力。为扩充训练数据、缓解模型潜在的过拟合风险并提升其泛化能力，本章采用了基于 Improved HP-VAE-GAN^[84] 的图像生成方法进行数据增强，从而使模型能够学习到更为全面的特征信息。为了定量评估数据增强策略的实际效果，设计了消融实验，对比了在“仅原始数据”与“原始数据 + 生成数据”两种训练策略下，各模型性能的变化情况，结果如表所示。可以看出，引入 Improved HP-VAE-GAN 生成的数据后，各模型的性能均得到了显著改善。

以本章提出的 TFFN 模型为例，在“原始数据 + 生成数据”的混合数据集上训练时，其 MAE 从 0.041 显著降低至 0.018，RMSE 从 0.057 降低至 0.024，同时 R^2 值由 0.279 上升至 0.873，说明预测精度和拟合优度均有大幅提升。在原始数据上表现不佳的 MFMPPM 和 MDFM 模型，其 R^2 值分别从 -0.537 和 -0.543 提升至 0.519 和 0.742，同时 MAE 与 RMSE 均明显下降。这表明原始数据中缺乏足够的多样性，而的生成的图像对这些信息进行了补充，帮助模型学习到更泛化的模式。

MMCL 模型虽然在原始数据上已有相对不错的表现， R^2 达到了 0.557，但在混合数据集上训练后依然取得了明显增益，MAE 从 0.041 降至 0.025，RMSE 从 0.045 降至 0.030， R^2 值也提升至 0.805。MLP-CNN 的 R^2 虽然仅从 0.344 稍微提高到 0.360，MAE 从 0.0462 减少至 0.036，说明数据增强对其也具备一定的帮助作用，只是程度略小于其他模型。数据增强作为一种正则化技术：通过引入更多样化的生成图像，让模型不再过度依赖于原始训练集中可能存在的特定模式，从而降低过拟合风险并提

升模型的泛化能力。

总的来说，本章所采用的数据增强方法，不仅增加了训练数据的数量与多样性，而且提升了包括 TFFN 在内的多种模型的预测精度与泛化能力。这表明数据增强对于本章场景下构建鲁棒且具备较强泛化能力的预测模型的重要性，对于那些在原始数据上表现不佳的模型，生成数据的加入补充了原始数据的不足，使这些模型能够学习到更有代表性的特征模式，从而实现性能的提升。

表 4.6 不同融合方式结果对比。

融合方式	MAE	RMSE	R ²
加法	0.030	0.044	0.577
拼接	0.030	0.045	0.573
交叉注意力	0.021	0.029	0.817
门控特征融合模块	0.018	0.024	0.873

为评估不同特征融合方式对模型预测性能的影响，本章设计了对应的消融实验，对比了包括加法、拼接、交叉注意力以及门控特征融合模块在内的多种融合方式，如表4.6所示。

本章首先评估了两种基础的特征融合方法：加法与拼接。前者对特征向量进行逐元素相加，后者对特征向量进行拼接。实验结果表明，这两种策略展现出相似且有限的预测性能。造成性能局限的原因在于这两种方法缺乏对图像特征与表格特征之间内在关联性和互补性的显式建模能力，也没有根据特征重要性进行自适应调整的机制。这种对跨模态信息交互的简单处理方式，导致了模型难以有效利用两种数据源，从而限制了预测精度，造成了相对较高的预测误差和较低的可解释性值。

交叉注意力机制通过利用一种模态的信息来引导对另一种模态特征的关注度分配，使得模型能够聚焦于与表格信息更相关的图像区域特征。引入交叉注意力后，模型的 MAE 降低至 0.021, RMSE 降低至 0.029, R² 达到了 0.817。这表明，相较于基础的特征级融合方法，交叉注意力具备了跨模态信息交互能力，能够识别并关联两种模态间的相关特征。然而，它虽然解决了特征对齐的问题，但是缺乏对融合后特征进行进一步的细粒度调控。

本章提出的门控特征融合模块采用两阶段架构，整合了模态特征交互与门控前馈融合，实现图像与表格数据的深度、自适应整合。实验结果表明，GFFM 在所有融合方法中表现最佳，MAE 为 0.018, RMSE 为 0.024, R² 达到 0.873，优于基础融合

方法和交叉注意力。门控特征融合模块的优势归因于其两阶段的融合方式，充分利用图像与表格这两种异构且互补的数据源。在模态特征交互阶段，利用交叉注意力机制。这实现了跨模态信息的有效对齐，使模型能够识别并聚焦于与给定表格数据上下文相关的关键图像特征区域。在门控前馈融合阶段，对融合后的特征进行动态自适应的调控，选择性地放大那些来自不同模态、能够相互补充、对预测目标最为关键的信息组合。同时，门控特征融合模块还结合了快捷连接，确保原始交互阶段的有价值信息得以有效保留和传递。

4.3 本章小结

针对功能材料多模态性能预测中现有方法在跨模态交互深度、特征融合有效性及计算效率方面的普遍局限，本章以海藻酸钙/石墨烯复合材料的性能预测为具体应用场景，提出了表格驱动特征融合网络。为解决特征提取阶段缺乏跨模态引导的问题，本章设计了表格引导视觉语义增强模块，通过其内部的通道筛选与空间聚焦组件，利用表格先验知识在早期动态调控图像特征的通道与空间注意力，实现深度跨模态对齐与语义增强。为应对融合阶段交互对齐不足和加权机制单一的挑战，本章设计了门控特征融合模块，通过交叉注意力机制促进模态间的有效对齐与信息交换，并结合门控前馈网络，实现特征贡献度的自适应加权与融合。同时，网络采用分流卷积及 Mamba 等轻量化设计，提升了计算效率与实用性。实验结果表明，本章提出的网络在各项评价指标上均优于代表性的单模态模型及多种先进的多模态融合模型。消融实验进一步证实了各模块对性能提升的重要贡献，为基于多源异构数据的功能材料性能预测提供了一种精准、高效且实用的方法。

第五章 总结与展望

5.1 总结

深度学习技术为功能材料的特征处理与性能预测带来了新的机遇，但在将其有效应用于实际材料科学问题时，仍面临若干关键挑战，一方面是如何有效提取功能材料图像中复杂的多层次特征；另一方面是如何实现图像、表格等多模态异构数据的深度融合与协同利用。针对上述挑战，本论文主要开展了以下两方面的工作：

(1) 针对功能材料图像多层次特征复杂的问题，本论文提出了一种基于特征解耦的轻量级深度学习方法，以解决钙钛矿 PUF 防伪标签真实性预测任务。针对多层次特征难以有效提取的问题，该方法采用了特征解耦策略，通过将特征提取过程显式地分解为两个独立阶段，分别优化对不同层级视觉信息的捕捉，提升了特征表示的有效性并减少了信息交叉干扰。为满足高效率和低延迟的应用需求，设计了一种轻量化的分流卷积网络，采用分流卷积机制，通过优化卷积计算方式，在有效保留信息的同时降低了运算量和参数规模。为增强特征的判别能力，在损失函数中加入了额外角度边界约束，通过显式地促进类内特征聚集与类间特征分散，引导模型学习更易于区分的深度表示。最后，针对防伪任务的开放集识别特性，提出了测地度量方法，通过将学习到的特征映射至高维流形，利用测地距离进行样本判别，有效识别已知样本并拒绝未知伪造样本。

(2) 针对功能材料多模态异构数据特征难以融合利用的问题，本论文提出了一种基于特征融合的轻量级深度学习方法，以解决海藻酸钙/石墨烯复合材料性能预测任务。为解决特征提取阶段缺乏跨模态引导与对齐的问题，提出了一种轻量级的表格驱动特征融合网络。其中的表格引导视觉语义增强模块，通过内部的通道筛选与空间聚焦组件，利用表格先验知识动态调控图像特征的通道与空间注意力，提升了跨模态特征表示的融合度与语义相关性。在特征融合层面，该网络设计了门控特征融合模块，通过采用交叉注意力机制促进模态间有效对齐与信息交换，并结合门控前馈网络实现特征贡献度的自适应加权，有效整合了异构信息的互补优势。此外，为提升计算效率与实用性，网络采用了分流卷积、Mamba 等轻量化设计，在保证性能的同时显著降低了计算开销和参数规模，以适应实际应用需求。

5.2 展望

尽管本论文提出的两个功能材料特征处理与预测方法都展现了优异的性能和应用潜力，为人工智能技术在功能材料领域的深入应用提供了有价值的参考方案，但现有工作仍存在一定的局限性。为了进一步拓宽方法的应用范围并深化相关理论与技术研究，未来可在以下几个方面继续开展相关工作：

(1) 本论文提出的高效性能预测方法能够准确、快速地建立材料的“结构/组分/工艺”与“性能”之间的复杂前向映射关系，为深入理解材料的构效关系提供了关键的数据驱动视角。未来的核心发展方向是将这些精准的预测模型作为性能评估工具或优化过程中的目标函数，并与先进的生成式模型如生成对抗网络 GANs、扩散模型等深度结合，实现从“给定材料预测性能”到“给定性能设计材料”的关键范式转变。通过从期望的性能指标出发，反向推导在广阔设计空间中具有目标性能的新型材料组分或微观结构，加速定制化新材料的发现与创新，满足特定应用需求，并探索具有全新功能的非直觉材料设计方案。

(2) 本论文提出的表格驱动特征融合网络有效地融合了图像与表格数据，但材料系统的完整描述及其性能的精确预测通常依赖于更广泛、更多样化的信息来源。未来的研究应致力于将这一融合框架扩展至能够处理和整合更多模态数据，包括但不限于光谱数据、衍射图谱、三维显微结构数据、时序相关的过程监控数据，甚至从科学文献或实验报告中提取的非结构化文本信息。实现这一目标不仅需要发展能够处理这些异构、高维数据的新型特征表示学习方法，还需研究能够有效捕捉更多模态之间复杂高阶交互作用的先进融合架构与机制，从而构建对材料系统更为全面精准的理解。

(3) 本论文提出的方法在钙钛矿 PUF 防伪标签和海藻酸钙/石墨烯复合材料数据上表现出了良好的性能，但其在更广泛材料体系和多样化数据条件下的泛化能力仍是未来研究的关键方向。材料科学领域的数据具有高度的多样性，不同材料类别的物理化学规律往往存在显著差异。因此，当将现有方法应用于新的数据集时，可能会因领域漂移等问题而导致性能下降。未来的研究可以着眼于提升网络的自适应能力与可迁移性，开发具有更强泛化能力、对标注数据集依赖性更低的材料分析模型，从而使其能够更加可靠高效地服务于复杂多变的材料科学研究和工业应用需求。

参考文献

- [1] KATIYAR A K, HOANG A T, XU D, et al. 2d materials in flexible electronics: recent advances and future perspectives[J]. *Chemical Reviews*, 2023, 124(2): 318-419.
- [2] ABDELHAMID H N. An introductory review on advanced multifunctional materials [J]. *Heliyon*, 2023, 9(7): e18060.
- [3] KUMAR Y, PASWAN K K, NAYAN K, et al. Introduction to functional materials: synthesis, properties, environmental sustainability, and general applications[M]// *Functional materials for the oil and gas industry*. CRC Press, 2023: 1-22.
- [4] MENON D, RANGANATHAN R. A generative approach to materials discovery, design, and optimization[J]. *ACS omega*, 2022, 7(30): 25958-25973.
- [5] 赵瑞, 石竞琛, 刘霏凝, 等. 人工神经网络加速计算材料科学的研究进展[J]. *吉林师范大学学报 (自然科学版)*, 2023, 44(01): 21-33.
- [6] SIVAN D, SATHEESH KUMAR K, ABDULLAH A, et al. Advances in materials informatics: a review[J]. *Journal of Materials Science*, 2024, 59(7): 2602-2643.
- [7] VU T S, HA M Q, NGUYEN D N, et al. Towards understanding structure–property relations in materials with interpretable deep learning[J]. *npj Computational Materials*, 2023, 9(1): 215.
- [8] AGRAWAL A, CHOUDHARY A. Perspective: Materials informatics and big data: Realization of the “fourth paradigm” of science in materials science[J]. *APL Materials*, 2016, 4(5): 053208.
- [9] HAN Y, BAO S, SHI B, et al. Fast and accurate recognition of perovskite fluorescent anti-counterfeiting labels based on lightweight convolutional neural networks[J]. *ACS Applied Materials & Interfaces*, 2024, 16(31): 41107-41118.
- [10] LIN Y, ZHANG H, FENG J, et al. Unclonable micro-texture with clonable micro-shape towards rapid, convenient, and low-cost fluorescent anti-counterfeiting labels [J]. *Small*, 2021, 17(30): 2100244.

- [11] ALADE A, IBRAHIM A. Application of finite element method for mechanical characterization of wood and reconstituted lignocellulosic-based composites—a review[J]. *Recent Progress in Materials*, 2022, 5(01): 003.
- [12] YADAV A, KUMAR S, SINGH J. A review of physical unclonable functions (pufs) and its applications in iot environment[J]. *Ambient Communications and Computer Systems: Proceedings of RACCCS 2021*, 2022: 1-13.
- [13] LI Y, BIDMESHKI M M, KANG T, et al. Genetic physical unclonable functions in human cells[J]. *Science Advances*, 2022, 8(18): eabm4106.
- [14] HU Y W, ZHANG T P, WANG C F, et al. Flexible and biocompatible physical unclonable function anti-counterfeiting label[J]. *Advanced Functional Materials*, 2021, 31(34): 2102108.
- [15] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. *International journal of computer vision*, 2004, 60: 91-110.
- [16] MURATAJ I, MAGOSSO C, CARIGNANO S, et al. Artificial fingerprints engraved through block-copolymers as nanoscale physical unclonable functions for authentication and identification[J]. *Nature Communications*, 2024, 15(1): 1-12.
- [17] INDHUMATHI S, CLEMENT J C. Comparison of feature detection using sift and orb detector[C]//TENCON 2022-2022 IEEE Region 10 Conference. IEEE, 2022: 1-3.
- [18] RUBLEE E, RABAUD V, KONOLIGE K, et al. Orb: An efficient alternative to sift or surf[C]//2011 International conference on computer vision. Ieee, 2011: 2564-2571.
- [19] KIREMITLER N B, ESIDIR A, DRAKE G A, et al. Tattoo-like multi-color physically unclonable functions[J]. *Advanced Optical Materials*, 2024, 12(12): 2302464.
- [20] ESIDIR A, KAYACI N, KIREMITLER N B, et al. Food-grade physically unclonable functions[J]. *ACS applied materials & interfaces*, 2023, 15(35): 41373-41384.
- [21] LEE S, PEKDEMIR S, KAYACI N, et al. Graphene-based physically unclonable functions with dual source of randomness[J]. *ACS applied materials & interfaces*, 2023, 15(28): 33878-33889.

- [22] TU Y, LIU B, YAO G, et al. A review of advanced thermal interface materials with oriented structures for electronic devices[J]. *Electronics*, 2024, 13(21): 4287.
- [23] BREITKOPF C. Theoretical characterization of thermal conductivities for polymers—a review[J]. *Thermo*, 2024, 4(1): 31-47.
- [24] DING D, ZOU M, WANG X, et al. Thermal conductivity of polydisperse hexagonal bn/polyimide composites: Iterative emt model and machine learning based on first principles investigation[J]. *Chemical Engineering Journal*, 2022, 437: 135438.
- [25] RUAN K, GUO Y, LU C, et al. Significant reduction of interfacial thermal resistance and phonon scattering in graphene/polyimide thermally conductive composite films for thermal management[J]. *Research*, 2021.
- [26] LI Z, XIAO W, RUAN X. A finite element analysis of the effects of graphene and carbon nanotubes on thermal conductivity of co phase in wc-co carbide[J]. *Materials*, 2021, 14(24): 7656.
- [27] AL-MASRI A, KHANAFER K, VAFAI K. Multiscale homogenization of aluminum honeycomb structures: Thermal analysis with orthotropic representative volume element and finite element method[J]. *Heliyon*, 2024, 10(2): e24166.
- [28] CHOUDHARY K, DECOST B, CHEN C, et al. Recent advances and applications of deep learning methods in materials science[J]. *npj Computational Materials*, 2022, 8(1): 59.
- [29] LIU Y, HAN F, LI F, et al. Inkjet-printed unclonable quantum dot fluorescent anti-counterfeiting labels with artificial intelligence authentication[J]. *Nature communications*, 2019, 10(1): 2409.
- [30] WANG J, ZHANG Q, CHEN R, et al. Triple-layer unclonable anti-counterfeiting enabled by huge-encoding capacity algorithm and artificial intelligence authentication [J]. *Nano Today*, 2021, 41: 101324.
- [31] JING L, XIE Q, LI H, et al. Multigenerational crumpling of 2d materials for anti-counterfeiting patterns with deep learning authentication[J]. *Matter*, 2020, 3(6): 2160-2180.

- [32] CHEN H, HU H, SUN B, et al. Dynamic anti-counterfeiting labels with enhanced multi-level information encryption[J]. *ACS Applied Materials & Interfaces*, 2022, 15(1): 2104-2111.
- [33] PRADHAN S, RAJAGOPALA A D, MENO E, et al. Deep-learning-based digitization of protein-self-assembly to print biodegradable physically unclonable labels for device security[J]. *Micromachines*, 2023, 14(9): 1678.
- [34] SUN J, SHEN Z, WANG Y, et al. Loftr: Detector-free local feature matching with transformers[C]//*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2021: 8922-8931.
- [35] ZHANG J, LIU Y, NJEL C, et al. An all-in-one nanoprinting approach for the synthesis of a nanofilm library for unclonable anti-counterfeiting applications[J]. *Nature Nanotechnology*, 2023, 18(9): 1027-1035.
- [36] ZHANG J, TAN R, LIU Y, et al. Printed smart devices for anti-counterfeiting allowing precise identification with household equipment[J]. *Nature Communications*, 2024, 15(1): 1040.
- [37] SUN N, CHEN Z, WANG Y, et al. Random fractal-enabled physical unclonable functions with dynamic ai authentication[J]. *Nature Communications*, 2023, 14(1): 2185.
- [38] MILLWOOD O, MISKELLY J, YANG B, et al. Puf-phenotype: a robust and noise-resilient approach to aid group-based authentication with dram-pufs using machine learning[J]. *IEEE Transactions on Information Forensics and Security*, 2023, 18: 2451-2465.
- [39] JUAN Y, DAI Y, YANG Y, et al. Accelerating materials discovery using machine learning[J]. *Journal of Materials Science & Technology*, 2021, 79: 178-190.
- [40] YANG Y, YANG R C, YANG J, et al. Machine learning for thermal conductivity prediction in graphene/hexagonal boron nitride van der waals heterostructures[J]. *The Journal of Physical Chemistry C*, 2025, 129(5): 2764-2774.
- [41] MISHRA A K, RAJPUT S, KARAMTA M, et al. Exploring the possibility of machine learning for predicting ionic conductivity of solid-state electrolytes[J]. *ACS omega*, 2023, 8(18): 16419-16427.

- [42] KANTI P K, PARAMASIVAM P, WANATASANAPPAN V V, et al. Experimental and explainable machine learning approach on thermal conductivity and viscosity of water based graphene oxide based mono and hybrid nanofluids[J]. Scientific Reports, 2024, 14(1): 30967.
- [43] CHOUDHARY K, GARRITY K F, CAMP C, et al. Computational scanning tunneling microscope image database[J]. Scientific data, 2021, 8(1): 57.
- [44] YANG Z, PAPANIKOLAOU S, REID A C, et al. Learning to predict crystal plasticity at the nanoscale: Deep residual networks and size effects in uniaxial compression discrete dislocation simulations[J]. Scientific reports, 2020, 10(1): 8262.
- [45] LIU C, GUO R, SU Y. A deep learning based prediction model for effective elastic properties of porous materials[J]. Scientific Reports, 2025, 15(1): 6707.
- [46] CHEN S, SAKAI S, MATSUO-UEDA M, et al. Investigation of multi-input convolutional neural networks for the prediction of particleboard mechanical properties[J]. Scientific Reports, 2025, 15(1): 4162.
- [47] FARIZHANDI A A K, BETANCOURT O, MAMIVAND M. Deep learning approach for chemistry and processing history prediction from materials microstructure[J]. Scientific reports, 2022, 12(1): 4552.
- [48] HE C, YU L, JIANG Y, et al. Deep-learning approach for developing bilayered electromagnetic interference shielding composite aerogels based on multimodal data fusion neural networks[J]. Journal of Colloid and Interface Science, 2025, 688: 79-92.
- [49] DAS K, GOYAL P, LEE S C, et al. Crysmmnet: multimodal representation for crystal property prediction[C]//Uncertainty in Artificial Intelligence. 2023: 507-517.
- [50] WANG F, GUO W, CHENG M, et al. Mmpolymer: A multimodal multitask pre-training framework for polymer property prediction[C]//Proceedings of the 33rd ACM International Conference on Information and Knowledge Management. 2024: 2336-2346.
- [51] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.

- [52] CHOLLET F. Xception: Deep learning with depthwise separable convolutions[C]// Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 1251-1258.
- [53] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]// Proceedings of the European conference on computer vision. 2018: 3-19.
- [54] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]// NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2017: 6000–6010.
- [55] GU A, DAO T. Mamba: Linear-time sequence modeling with selective state spaces [C]//First Conference on Language Modeling. 2024.
- [56] ZHAO F, ZHANG C, GENG B. Deep multimodal data fusion[J]. ACM computing surveys, 2024, 56(9): 1-36.
- [57] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications[A]. 2017. arXiv: 1704.04861.
- [58] HAN K, WANG Y, TIAN Q, et al. Ghostnet: More features from cheap operations[C]// Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 1580-1589.
- [59] ZHANG Q, JIANG Z, LU Q, et al. Split to be slim: An overlooked redundancy in vanilla convolution[C]//Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence. 2020: 3195-3201.
- [60] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]//International conference on machine learning. pmlr, 2015: 448-456.
- [61] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 4510-4520.
- [62] HENDRYCKS D, GIMPEL K. Gaussian error linear units (gelus)[A]. 2023. arXiv: 1606.08415.

- [63] NAIR V, HINTON G E. Rectified linear units improve restricted boltzmann machines [C]//Proceedings of the 27th international conference on machine learning. 2010: 807-814.
- [64] WANG H, WANG Y, ZHOU Z, et al. Cosface: Large margin cosine loss for deep face recognition[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018: 5265-5274.
- [65] DENG J, GUO J, XUE N, et al. Arcface: Additive angular margin loss for deep face recognition[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 4685-4694.
- [66] TOLSTIKHIN I O, HOULSBY N, KOLESNIKOV A, et al. Mlp-mixer: An all-mlp architecture for vision[J]. Advances in neural information processing systems, 2021, 34: 24261-24272.
- [67] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4700-4708.
- [68] TAN M, CHEN B, PANG R, et al. Mnasnet: Platform-aware neural architecture search for mobile[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 2820-2828.
- [69] TAN M, LE Q. Efficientnet: Rethinking model scaling for convolutional neural networks[C]//International conference on machine learning. 2019: 6105-6114.
- [70] RADOSAVOVIC I, KOSARAJU R P, GIRSHICK R, et al. Designing network design spaces[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 10428-10436.
- [71] LIU Z, MAO H, WU C Y, et al. A convnet for the 2020s[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022: 11976-11986.
- [72] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[C]//International Conference on Learning Representations. 2021.

- [73] TOUVRON H, CORD M, DOUZE M, et al. Training data-efficient image transformers & distillation through attention[C]//International conference on machine learning. 2021: 10347-10357.
- [74] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2021: 10012-10022.
- [75] MEHTA S, RASTEGARI M. Mobilevit: Light-weight, general-purpose, and mobile-friendly vision transformer[C]//International Conference on Learning Representations. 2022.
- [76] WANG W, XIE E, LI X, et al. Pvt v2: Improved baselines with pyramid vision transformer[J]. Computational visual media, 2022, 8(3): 415-424.
- [77] MAAZ M, SHAKER A, CHOLAKKAL H, et al. Edgenext: efficiently amalgamated cnn-transformer architecture for mobile vision applications[C]//European conference on computer vision. 2022: 3-20.
- [78] VAN DER MAATEN L, HINTON G. Visualizing data using t-sne.[J]. Journal of machine learning research, 2008, 9(11): 2579-2605.
- [79] CHEN H, HU H, SUN B, et al. Dynamic anti-counterfeiting labels with enhanced multi-level information encryption[J]. ACS Applied Materials & Interfaces, 2023, 15(1): 2104-2111.
- [80] LIU Y, ZHENG Y, ZHU Y, et al. Unclonable perovskite fluorescent dots with fingerprint pattern for multilevel anticounterfeiting[J]. ACS Applied Materials & Interfaces, 2020, 12(35): 39649-39656.
- [81] CARRO-TEMBOURY M R, ARPPE R, VOSCH T, et al. An optical authentication system based on imaging of excitation-selected lanthanide luminescence[J]. Science Advances, 2018, 4(1): e1701384.
- [82] GU Y, HE C, ZHANG Y, et al. Gap-enhanced raman tags for physically unclonable anticounterfeiting labels[J]. Nature communications, 2020, 11(1): 516.

- [83] WU J, LIU X, LIU X, et al. A high-security mutual authentication system based on structural color-based physical unclonable functions labels[J]. *Chemical Engineering Journal*, 2022, 439: 135601.
- [84] HAN Y, LIU Y, CHEN Q. Data augmentation in material images using the improved hp-vae-gan[J]. *Computational Materials Science*, 2023, 226: 112250.
- [85] CHEN K Y, CHIANG P H, CHOU H R, et al. Trompt: towards a better deep neural network for tabular data[C]//*Proceedings of the 40th International Conference on Machine Learning*. 2023.
- [86] GORISHNIY Y, RUBACHEV I, KHRULKOV V, et al. Revisiting deep learning models for tabular data[J]. *Advances in neural information processing systems*, 2021, 34: 18932-18943.
- [87] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors[J]. *nature*, 1986, 323(6088): 533-536.
- [88] QIN D, LEICHNER C, DELAKIS M, et al. Mobilenetv4: universal models for the mobile ecosystem[C]//*European Conference on Computer Vision*. 2024: 78-96.
- [89] WOO S, DEBNATH S, HU R, et al. Convnext v2: Co-designing and scaling convnets with masked autoencoders[C]//*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2023: 16133-16142.
- [90] KIM D, HEO B, HAN D. Densenets reloaded: paradigm shift beyond resnets and vits [C]//*European Conference on Computer Vision*. 2024: 395-415.
- [91] HAGER P, MENTEN M J, RUECKERT D. Best of both worlds: Multimodal contrastive learning with tabular and imaging data[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023: 23924-23935.
- [92] LI M, LI S, TIAN Y, et al. A deep learning convolutional neural network and multi-layer perceptron hybrid fusion model for predicting the mechanical properties of carbon fiber[J]. *Materials & Design*, 2023, 227: 111760.
- [93] SONG L, WANG D, LIU X, et al. Prediction of mechanical properties of composite materials using multimodal fusion learning[J]. *Sensors and Actuators A: Physical*, 2023, 358: 114433.

攻读硕士学位期间取得的研究成果

一、论文

1. Han Y, Bao S, Shi B, et al. Fast and Accurate Recognition of Perovskite Fluorescent Anti-counterfeiting Labels Based on Lightweight Convolutional Neural Networks[J]. ACS Applied Materials & Interfaces, 2024, 16(31): 41107-41118. (已发表, 导师一作, 本人二作, 中科院二区)

二、软著

1. 软件名称: 基于深度学习的防伪标签识别平台 V1.0, 开发人: 韩越兴、包胜奇、王冰。登记号: 2023SR1430101, 登记日期: 2023.11.14, 申请人: 上海大学。(导师一作, 本人二作)

2. 软件名称: 基于多模态模型的高分子材料导热性能预测平台 V1.0, 开发人: 韩越兴、包胜奇、王冰。登记号: 2024SR1572005, 登记日期: 2024.10.21, 申请人: 上海大学。(导师一作, 本人二作)

致 谢

时光荏苒，我与这座校园的初见，还定格在许多年前那个蝉鸣不止的初夏。七载春秋，从懵懂新生到硕士毕业，这里早已是我青春的注脚与心安的归处。又是一个绿意盎然的初夏，当初的起点，如今成了告别的渡口。回望这段研究生岁月，恰如这漫长求学路上一场最绚烂的盛放，其间离不开每一位师长与同学的阳光雨露。在此，请允许我致以最诚挚的谢意。

首先我要感谢韩越兴老师，从论文选题到实验设计再到终稿的完成，都离不开韩老师的悉心指导和帮助。科研上，韩老师严谨求是的治学态度和对学术的一腔热忱，深深地感染着我。正是这种言传身教的力量，激励着我在面对科研挑战时保持耐心、迎难而上，也让我懂得了何为学者的坚韧与执着。同时，我还要感谢陈侨川老师精准的指点和创新的思路。在此衷心地祝愿各位老师身体健康、学术长青。

其次，我要感谢我的家人。感恩我的父母，是他们给予我动力，去探索无限可能，也始终做我最坚实的后盾。正是这份守护与支持，托举我站上他们的肩膀，去看见一个比他们所见更广阔的世界。也要感谢汤，在这段充满挑战的岁月里，给予我最长情的陪伴与最温柔的包容，分担我的焦虑，分享我的喜悦。未来的路，愿我们继续携手并肩，一同前行。最后，感谢所有关心我的亲人们。唯愿你们长安久乐，身体康健，余生皆是平安顺遂。

学贵得师，亦贵得友，我要感谢我的同门。感谢赵哥，用动感舞姿，为沉闷的科研注入活力；感谢阮哥，用一杯匠心手冲，让我提前品尝到了人生的真味；感谢帆哥，用酒馆中的智谋，为我解开科研难题之局；感谢慧子姐，用“万物皆可冲会员”的生活哲学，让我领悟到对生活的热爱就应这般纯粹直接；感谢一姐，用她的严谨认真，为我树立了学术的标杆。同时，也要感谢实验室所有其他师兄姐妹们，这份同窗情谊，是我求学路上最宝贵的财富。

再回首，轻舟已过万重山。再次感谢所有伸出援手的人，愿天高海阔，万事胜意。愿自己，身怀赤诚，于万物众生中磊落做人。