

中图分类号: TP391

单位代号: 10280

密 级: 公开

学 号: 21721561

上海大学



硕士学位论文

SHANGHAI UNIVERSITY
MASTER'S DISSERTATION

题 目	基于形状空间理论的图像特征增强及其在材料性能预测中的应用
--------	------------------------------

作 者 万冠新

学科专业 计算机应用技术

导 师 韩越兴

完成日期 二〇二四年六月

姓 名：万冠新

学号：21721561

论文题目：基于形状空间理论的图像特征增强及其在材料性能预测中的应用

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：付剑锋

委 员：付剑锋

刘 伟

王宜敏

导 师：韩越兴

答辩日期：2024 年 8 月 9 日

姓 名：万冠新

学号：21721561

论文题目：基于形状空间理论的图像特征增强及其在材料性能预测中的应用

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：万冠新

日期：2024 年 8 月 9 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：万冠新

导师签名：韩越兴

日期：2024 年 8 月 9 日

日期：2024 年 8 月 9 日

上海大学工学硕士学位论文

基于形状空间理论的图像特征增强及
其在材料性能预测中的应用

作者: 万冠新

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

2024 年 6 月

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

Image Feature Augmentation Based on Shape Space Theory and Application in Predicting Material Properties

Candidate: Guanxin Wan

Supervisor: Yuexing Han

Major: Technology of Computer Application

School of Computer Engineering and Science

Shanghai University

June, 2024

摘 要

随着机器学习和深度学习的迅速发展，越来越多的学习方法被引入到图像处理的相关任务场景中以取得更加出色的性能表现。大部分机器学习和深度学习模型依赖于大规模的数据量和高质量的标注样本，但是在实际应用场景中获取大量图像数据非常困难且成本高昂。在小样本图像的数据规模下，由于样本数量有限，导致模型难以充分学习到数据的特征分布，从而影响模型的性能和泛化能力。因此，充分挖掘小样本图像的内在信息并扩充样本数据规模变得至关重要。本论文针对小样本图像的数据规模，结合机器学习、深度学习和形状空间理论，提出高效的图像特征增强方法，并应用于图像分类和基于材料图像的材料性能预测中，主要研究内容如下：

(1) 针对小样本图像场景下可用数据少和多样性不足的特点，本论文提出一种基于预形状空间测地曲线的图像特征增强方法，简称 **FAGC-PSS** (**Feature Augmentation on Geodesic Curves in Pre-Shape Space**)。首先使用深度学习模型来提取小样本图像的特征；根据形状空间理论提升图像特征的维度并将其投影至预形状空间；分别对每一类别的特征数据构造出对应的测地曲线；最后沿着最优的测地曲线生成特征数据，用于图像处理模型的训练。该方法的创新包括以下三点：第一，实现小样本图像特征增强，有助于模型全面地理解训练样本分布及规律，提高模型的鲁棒性和可靠性；第二，提出的 **FAGC-PSS** 特征增强方法可以应用于多个下游任务，比如与机器学习的分类模型结合能获得较好的结果；第三，在小样本图像分类任务的交叉熵损失函数中设计随机概率函数和影响因子，可以平衡生成特征和图像特征对模型的影响。

(2) 针对材料图像存在小样本和材料性能预测精度不佳的特点，本论文提出一种基于 **FAGC-PSS** 的材料性能预测方法。该方法通过设计 **FAGC-PSS** 的下游任务框架结构，结合伪标签机制，实现在小样本材料图像上的材料性能预测任务。具体流程包括四个步骤：材料图像特征的提取，通过 **FAGC-PSS** 生成特征，利用伪标签机制为生成特征标注性能值标签，最后利用增强特征对材料性能预测模型进行训练。该方法的创新包括以下：在预测材料性能模型中引入 **FAGC-PSS** 模块，以增强特征数据的多样性和复杂性；对生成的特征数据设计伪标签机制进行标注。实验结果表明，

对于不同种类的材料性能预测任务，该方法能够展现出良好的有效性和普适性。

本论文基于公共图像数据集和多种材料图像进行实验，分别验证本论文方法在小样本图像分类和材料性能预测的有效性和普适性。此外，本研究对其他学科领域内的小样本问题同样具有借鉴意义，为小样本图像的研究提供了新思路。

关键词：小样本图像，特征增强，形状空间理论，测地曲线

ABSTRACT

The rapid development of machine learning and deep learning has led to the introduction of an increasing number of learning methods into image processing-related tasks, with the aim of achieving enhanced performance. The majority of machine learning and deep learning models depend on extensive data volumes and high-quality labeled samples. However, acquiring such data in practical applications is often challenging and costly. In the context of small sample images, the limited number of samples presents a challenge for the model in fully learning the data's feature distribution, which in turn affects the model's performance and generalization ability. Therefore, it is vital to fully exploit the intrinsic information of small sample images and expand the sample data size. This dissertation proposes an efficient image feature enhancement method for small sample images, combining machine learning, deep learning, and shape space theory. This method is applied to image classification and material performance prediction based on material images. The main research contents are as follows:

(1) In response to the challenges of limited and diverse data in the context of small sample images, this dissertation proposes an image feature augmentation method based on Geodesic curves in pre-shape space, abbreviated as FAGC-PSS (Feature Augmentation on Geodesic Curves in Pre-Shape Space). First, deep learning models are employed to extract features from small sample images. These features are then projected into pre-shape space by enhancing their dimensions according to shape space theory. For each category of feature data, corresponding Geodesic curves are constructed. Finally, feature data are generated along the optimal geodesic curves for training image processing models. The innovations of this method include three key points: first, the enhancement of small sample image features, which helps models to comprehensively understand the distribution and patterns of training samples, thereby improving robustness and reliability; second, the proposed FAGC-PSS feature augmentation method can be applied to various downstream tasks, such as achieving

better results when combined with machine learning classification models; third, a random probability function and influence factor are designed in the cross-entropy loss function of small sample image classification tasks, which can balance the impact of generated features and image features on the model.

(2) In light of the shortcomings associated with the prediction of material properties using small sample sizes and the limitations of existing material property prediction models, this dissertation proposes a method for predicting material performance from material images based on FAGC-PSS. The method employs the FAGC-PSS framework to predict material performance based on small sample material images, integrating the pseudo-labelling mechanism to achieve this objective. The specific process comprises four steps: the initial extraction of material image features, the subsequent generation of features by FAGC-PSS, the labelling of the generated features with performance value labels using the pseudo-labelling mechanism, and finally, the training of the material performance prediction model using the augmented features. The method's key innovations include the introduction of the FAGC-PSS module into the predictive material performance model, which enhances the diversity and complexity of the feature data, and the design of a pseudo-labelling mechanism for labelling the generated feature data. The experimental results demonstrate that the method is effective and generalisable for different material performance prediction tasks.

This dissertation conducts experiments based on public image datasets and various material images to validate the effectiveness and generalizability of the proposed method in small sample image classification and material performance prediction. Furthermore, this research holds significant implications for addressing small sample issues in other disciplinary fields, offering new perspectives for the study of small sample images.

Keywords: Small Sample Image; Feature Augmentation; The Shape Space Theory; The Geodesic Curves

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 课题研究目的与意义	2
1.4 国内外研究现状.....	3
1.4.1 生成式图像增强研究现状.....	3
1.4.2 非生成式图像增强研究现状.....	6
1.4.3 材料性能预测研究现状	8
1.5 论文主要工作	9
1.6 论文组织结构	10
第二章 相关理论与技术概述	12
2.1 特征增强	12
2.2 深度学习	14
2.3 形状空间理论	17
2.4 伪标签设计机制.....	20
2.5 数据增强评价指标.....	21
2.6 本章小结	22
第三章 基于预形状空间测地曲线的图像特征增强方法	23
3.1 方法概述	24
3.1.1 图像特征的提取	24
3.1.2 投影图像特征至预形状空间	24
3.1.3 构造测地曲线拟合特征	25
3.1.4 改进的损失函数	26

3.2	实验分析	29
3.2.1	数据集和实验环境设置	29
3.2.2	相关方法对比	30
3.2.3	方法有效性分析	35
3.2.4	消融实验	39
3.3	本章小结	40
第四章	基于 FAGC-PSS 的材料性能预测方法	42
4.1	方法概述	42
4.1.1	材料图像特征提取	42
4.1.2	标注模型	44
4.1.3	预测模型	46
4.2	实验分析	46
4.2.1	材料图像数据集	46
4.2.2	特征提取模型的对比实验	47
4.2.3	CuCrZr 图像的材料性能预测	48
4.2.4	CoCr 图像的材料性能预测	53
4.2.5	生成特征数量与预测模型的相关性分析	54
4.2.6	标注模型与预测模型的关联性实验	54
4.3	本章小结	56
第五章	总结与展望	58
5.1	结论	58
5.2	工作展望	59
参考文献		61
攻读硕士学位期间取得的研究成果		73
攻读硕士学位期间参与的科研项目		74
致 谢		75

第一章 绪论

1.1 课题来源

本课题得到国家重点研发计划（编号：2018YFB0704400，2018YFB0704402，2020YFB0704503），上海市自然科学基金项目（编号：20ZR1419000）资助，之江实验室科研攻关项目资助（编号：2021PE0AC02），国家自然科学基金（面上，编号：52273228），云南省科技重点项目（编号：202302AB080022），硅酸盐质文物保护教育部重点实验室（上海大学）资助（课题编号：SCRC2023ZZ07TS）。

1.2 课题背景概述

近年来，随着人工智能和深度学习的迅猛发展，图像处理和计算机视觉技术已成为科学研究和工业应用中的重要组成部分。在众多领域，如医学影像分析、智能监控、自动驾驶等，图像数据的获取和分析发挥着至关重要的作用^[1-4]。在图像分类任务场景下，机器学习模型和深度学习模型通常需要大量的训练数据才能充分学习数据的特征和表示。然而，在实际应用中，常常由于训练数据量不足，导致模型难以从有限的标注样本中学习到的特征分布，从而产生过拟合或欠拟合现象，降低模型的泛化能力。在基于材料图像预测材料性能任务中，模型需要从数据中学习材料微观结构与性能之间的关系，然后用于预测未知材料的性能。由于材料图像的获取难度大和图像质量不可控等因素，极端情况下可能只有几张图像，这些图像的数量和质量是无法满足模型训练的需求，导致样本无法覆盖可能的特征分布，从而影响了模型的预测精度和稳定性^[5-7]。此外，直接根据特定图像数据设计算法模型被认为是简单且有效的处理方式，已有许多研究者针对特定图像或任务场景来设计算法^[8]。然而，当这些算法模型应用到其他图像任务中，模型的性能表现往往并不理想，算法模型缺乏普适性问题尤为突出。

针对小样本图像分类或材料性能预测等任务场景，目前主要的解决方案包括：基于模型微调的小样本学习方法，基于迁移学习的小样本学习方法以及基于迁移学习的小样本学习方法。基于模型微调的方法通常在大规模数据集上预训练模型，再根

据目标数据集修改网络模型输出层参数进行微调训练，但此类方法的局限性在于目标数据集和源数据集的分布要足够相似。基于迁移学习的方法^[9]主要目标是将已经学会的知识有效地转移到一个新的领域中，即利用旧知识学习新知识，但是过度迁移可能会导致模型过度依赖预训练知识，而忽略目标任务的关键特征。基于数据增强的方法^[10-12]利用已有数据产生更多具有相似性的数据，实现扩充样本集，避免少量数据带来的模型鲁棒性低、过拟合等问题的发生。

在众多方法中，数据增强策略由于其独特的优势备受关注。首先，通过扩充样本集，数据增强显著降低了模型的过拟合风险。过拟合是小样本学习中常见的问题，因为模型容易记住少量的训练样本而不能泛化到处理新的数据。通过增加数据量，模型在训练过程中可以看到更多多样化的样本，从而更好地学习到数据的本质特征，提升其在新数据上的表现。其次，数据增强还能够降低对复杂算法的依赖。传统上，为了提高模型的性能，研究人员往往需要设计复杂的模型架构。然而，数据增强技术通过增加训练数据的多样性，使得即使是简单的模型也能取得优异的性能。这不仅降低了模型的训练和推理时间，还简化了模型的部署和维护。此外，数据增强技术的应用范围广泛，不仅在图像处理领域得到了广泛应用，还在自然语言处理、语音识别等领域展现出其有效性^[13-15]。例如，在图像处理领域，常用的数据增强方法包括随机裁剪、色彩变换、对比度调整等^[16]；在自然语言处理领域，可以通过同义词替换、随机插入、随机删除等方法进行数据增强^[17]；在语音识别领域，则可以通过改变音频速度、音调以及添加背景噪声等方式来增强数据^[18]。因此，如何有效利用数据增强技术扩充样本数量，已然成为小样本图像任务处理的首要解决方案。

1.3 课题研究目的与意义

随着人工智能和大数据技术的迅速发展，数据的质量和数量成为制约机器学习和深度学习模型性能的关键因素。在这一背景下，数据增强技术（Data Augmentation）作为一种重要的方法，得到了广泛关注和应用。数据增强技术通过对原始数据进行各种变换或扩展，生成新的样本以增加数据的多样性和丰富性，从而在一定程度上缓解了数据不足和模型过拟合的问题。尽管数据增强技术在图像数据上取得了一定的成功，但是如何设计小样本图像的数据增强方法，仍然面临许多问题。首先，如何

在小样本数据上进行有效的数据增强并保持数据的多样性和代表性；其次，数据增强需要考虑到图像的特征和结构，以保证增强后的样本仍然具有可解释性和有效性；最后，数据增强可能引入新的噪声和偏差，导致模型学习到不真实的特征。

为了克服上述问题，本论文提出一种适用于小样本图像的数据增强方法，以快速且便捷地增强图像的特征数据为目标，该方法也可以称为基于预形状空间测地曲线的图像特征增强方法（Feature Augmentation on Geodesic curves in Pre-Shape Space, FAGC-PSS）。该方法可以降低模型对样本量的依赖和专家标注图像数据的成本，通过挖掘图像数据中包含的语义特征和结构信息来扩充高质量的特征数据，增加数据的多样性，以提高模型的鲁棒性和适应性，训练出更优秀的网络模型，实现对小样本图像的有效处理以供后续研究与分析工作。

材料图像往往是小样本数据，这导致大多数机器学习方法效果不佳，基于 FAGC-PSS，本论文提出一种从小样本材料图像预测材料性能的方法。通过该方法可有效地增加数据特征的多样性，使得模型在训练中能够更全面地学习到材料图像的特征分布，从而提高模型的性能和预测精度。

综上，本论文针对小样本图像，结合机器学习、深度学习以及形状空间理论，提出 FAGC-PSS 方法并分别实验于图像分类和材料性能预测中。本论文提出的特征增强方法可提高模型的性能和泛化能力，同时该方法可以推广到多种图像处理场景中。

1.4 国内外研究现状

本节介绍生成式图像增强、非生成式图像增强以及材料性能预测研究现状。

1.4.1 生成式图像增强研究现状

图像增强是计算机视觉和图像处理领域中的重要研究方向，旨在通过提升图像的视觉质量和信息表现，使其在各类应用中表现得更加出色。近年来，生成式方法逐渐成为图像增强的研究热点，该方法通过构建一个生成模型来学习样本的潜在分布，进而生成具有高质量和多样性的图像。常见的生成模型包括生成对抗网络 (GAN)^[19]、变分自编码器 (VAE)^[20]、流模型^[21]和扩散模型^[22]等。

GAN 自从被提出以来就一直备受关注，迄今为止其相关的衍生模型已达上百种之多。GAN 的基本结构包括生成器和判别器两个网络，它们通过对抗训练的方式

相互促进,生成器试图生成逼真的图像,判别器试图区分真实样本和生成样本。深度卷积生成对抗网络(DCGAN)是GAN的一种变体,它引入了深度卷积神经网络来构建生成器和判别器,使用卷积层和批量归一化技术,使得模型更加稳定,能够生成高分辨率的图像。然而,DCGAN在处理类别信息不足的情况下可能会出现模式崩溃的问题,而且对输入图像的分辨率和大小较为敏感^[23]。有条件生成对抗网络(CGAN)^[24]在原始GAN的基础上引入了条件信息,如类别标签,以实现有控制的图像生成。这种结构能够生成特定类别的图像,但需要额外的标签信息,并且CGAN对类别信息的准确性和完整性要求较高。Tripe-GAN^[25]则是引入额外的分类器网络,使得生成器不仅能够根据条件信息(如类别标签)生成图像,还能够生成与条件信息相关的真实分布。StackGAN^[26]则采用逐层递进式的生成方式,将文本描述逐渐映射到不同细节层次的图像上。WGAN^[27]以 Wasserstein 距离替代 JS 距离来衡量两个概率分布之间的差异,从而指导生成器和判别器的优化方向。WGAN-GP^[28]则是引入梯度惩罚技术来缓解梯度爆炸和梯度消失的问题。除此之外,GAN在图像到图像的风格迁移应用领域中也扮演着重要角色。比如 CycleGAN^[29]是一种无需配对数据即可进行图像转换的生成对抗网络,它通过引入循环一致性损失来学习两个不同域之间的映射关系。InstaGAN^[30]通过共享生成器和判别器实现多域图像转换任务中的高效性和鲁棒性,从而提高了转换效果和计算效率。IcGAN^[31]通过在潜在空间中进行插值操作,保持条件信息不变的情况下生成中间过渡图像。尽管GAN在图像生成、风格转换等领域取得了显著的成就,但也存在一些缺陷和局限性,比如图像生成过程中可能出现模式崩溃、训练不稳定、伪影、失真或者不真实的问题,特别是在小样本图像环境下,由于数据量十分受限,通常导致模型过拟合或者生成图像质量缺乏多样性和创造性等现象。

VAE是另一种重要的生成图像模型,和GAN一样,它也是从图像数据中学习潜在分布,并生成新的数据样本。VAE与GAN相比,两者的设计理念和训练方法大不相同,VAE本质是通过变分推断将图像数据映射到潜在空间,并从该空间进行重建,接着使用解码器生成具有相似分布的新样本。VAE可以与GAN进行结合,比如条件变分自编码器(CVAE)^[32]通过引入条件变量,如标签信息,以实现在给定条件下生成具有多样性和高质量的图像。InfoVAE^[33]使用最大化重构误差和最小化信息论度

量（如互信息）之间的差异作为优化目标，生成更具有丰富信息的图像，其训练过程相对复杂并且需要过多的计算资源。 β -VAE^[34]是在标准 VAE 框架下引入了额外的超参数 β ，通过调节 β 的值，平衡潜在空间的规范性和表示能力，从而实现更好地控制生成图像的多样性和重建质量，但这种方法需要大量实验来调整参数。除此之外，还有许多 VAE 的变体模型能够生成质量较高的图像，比如 AdaIN^[35]、IntroVAE^[36]和 IAF-VAE^[37]等。虽然 VAE 能够对数据进行有效压缩和信息表示，但同样存在一些缺点，诸如非常依赖训练数据、生成图像存在模糊、失真以及训练过程相对复杂等现象^[38]。此外，对于小样本图像数据，由于对数据分布的建模偏差，模型性能表现可能不佳，导致 VAE 生成的图像存在逼真度和细节不足的问题。

与 GAN 和 VAE 等隐变量生成模型不同，流模型直接建模数据的密度函数，并通过变换将输入数据映射到潜在空间中。在流模型中，变换通常由一系列可逆函数组成，这些函数被称为流变换。流模型的核心思想是通过连续且可逆的变换来学习数据的分布，从而实现数据的生成和密度估计。流模型的优势之一在于其生成过程的可解释性和可控性。由于流变换是可逆的，因此可以直观地理解生成过程，并且能够实现对生成样本的精细控制，特别是在处理高维数据时，流模型的可解释性更加突出。Dinh 等人提出的 RealNVP 展示了流模型在生成过程中的可解释性和效率^[39]。相较于其他生成模型，流模型可以直接建模数据的密度函数，如 Papamakarios 等人提出 Masked Autoregressive Flow(MAF) 模型^[40]。然而，流模型也面临一些挑战和局限性，比如复杂的结构导致训练和推断过程更加繁琐和耗时、甚至处理高维数据时面临维度灾难等问题^[41]。此外，流模型更多应用于大规模数据集的场景，对于小样本图像任务而言，生成的图像质量可能会低于其他生成模型。

扩散模型作为一种新兴的生成模型，采用了一种完全不同的生成策略，即通过连续的扩散过程逐步生成数据样本^[42]。在扩散模型中，数据样本从一个简单的噪声分布开始，经过多个扩散步骤，逐渐转化为目标数据分布。这种生成方式与自然界中许多现象相似，如热扩散、颗粒度扩散等，因此具有较好的物理意义和直观性。常见的扩散模型有去噪扩散隐式模型（Denoising diffusion implicit model, DDIM）^[43]和去噪扩散概率模型（Denoising diffusion probabilistic models, DDPM）^[44]，分别引入了可逆操作和马尔可夫噪声来模拟数据的扩散过程，对于视频生成和处理领域有较好的

应用潜力,但是它们的训练和推断过程较为繁琐,同时模型的复杂度较高,可能导致过拟合和泛化性能不佳等问题。**ARDM**^[45]通过将扩散过程与自回归模型相结合,实现了高效的图像生成和修复。该方法利用自回归模型逐像素地建模条件分布,不仅生成高质量图像,而且具备灵活性和可控性。**ADD**^[46]结合了扩散过程和对抗训练的思想,旨在提高生成图像的质量和多样性。该方法通过引入对抗损失,利用学生和教师网络结构,从复杂的高斯混合模型中学习生成图像的潜在表示,以实现更加逼真和多样的图像生成。尽管扩散模型在生成高质量图像方面取得了一定的成果,但在处理复杂图像内容、多模态数据以及小样本图像数据时,模型容易产生模糊、失真或不真实的图像样本,影响模型的实际应用效果。

1.4.2 非生成式图像增强研究现状

在非生成式图像增强方法中,注重通过直接操作图像像素或特征来改善图像的质量和多样性,而不依赖于模型生成图像。其中包括几何变换、图像混合、特征空间转换等技术^[47]。几何变换技术涉及对图像进行旋转、缩放、平移等操作,以改变其外观和结构。图像混合技术则通过融合或叠加多张图像来生成新的图像样本。特征空间转换技术则利用深度学习模型学习图像的特征表示,然后通过特定的变换操作来改变图像的外观和结构。非生成式图像增强方法的优势在于其操作的直接性和多样性,可以灵活地应用于不同类型和大小的图像数据。此外,它们具有较低的计算复杂度,能够在较短的时间内实现图像增强。同时,非生成式方法通常能够保持图像的真实性和细节,避免生成式方法可能出现的模糊或失真问题。

基于几何变化的图像数据增强方法相对简单且易于实现,该方法包括对图像平移、旋转、缩放、镜像翻转和仿射变换等操作。其中图像平移、旋转和缩放是通过调整图像的角度、位置和尺寸比例实现的,从而提高图像的视觉质量和多样性。这些增强策略可以用在很多深度学习网络,如 **AlexNet**^[48]、**VGG-16**^[49]和 **ResNet**^[50]。然而,对于尺寸较大的图像,这种处理方式可能会导致图像失真,进而影响模型的性能。镜像翻转和仿射变换是几何变换类中特别的数据增强手段,它们通过改变图像的形状和结构以增加图像的多样性和复杂性。进一步研究中,**Jaderberg** 等人提出的空间转换网络 (**Spatial Transformer Networks, STN**)^[51]可以自动学习执行几何变换的网络模块,从而改善模型的性能。**Deng** 等人^[52]实验验证仿射变换操作可以完成各种

复杂的图像变换，如图像校正、图像配准和图像融合等，有助于改善图像的质量和视觉效果。尽管基于几何变换的图像增强方法简单易用，但仍存在一些缺点。首先，这些方法容易导致图像标签数据的改变。以手写数字分类任务为例，对于 MNIST 数据集^[53]的图像进行翻转或旋转操作，可能会导致某些图像的标签不再准确^[54]。另外，几何变换类方法所产生的图像样本相对单一，未考虑这些样本是否具有实际价值，因此适用场景和操作范围相对有限。大多数几何变换类方法并未改变图像本身的内容，仅对局部图像或像素进行重新分布，导致增强后的图像多样性不高。此外，这些方法的效果在小样本图像环境中十分有限，因为它们的提升能力受到数据量和样本多样性的限制。

图像混合增强技术通过将多个图像样本进行混合或合成，产生新的训练样本，以丰富数据的多样性和增加模型的鲁棒性。该方法的本质在于通过引入图像间的关联信息，提高模型对数据的理解和学习能力，从而改善模型在实际应用场景中的性能表现。其中 Mixup^[55]作为一种简单而有效的图像混合方法，通过线性插值的方式将两个图像按一定比例混合，生成新的训练样本。Cutout^[56]是另一种常见的图像混合方法，通过随机选一个区域并将该区域的像素值置零来获取新的混合图像。在这些图像混合方法的基础上，可以进一步考虑数据空间图像增强的效果和多样性。例如，CutMix^[57]方法采用局部区域替换的方式来生成新的混合图像，以增强模型对局部信息的学习能力。AugMix^[58]方法通过堆叠多种数据增强操作来生成新的混合图像，从而提高数据的多样性和复杂度。GridMask^[59]方法则引入了网络遮挡的概念，通过在图像上添加网络遮挡来生成新的混合图像，以增强模型对全局和局部信息的学习能力。Context Encoders 是一个典型的图像补全模型^[60]，该模型通过在图像中随机遮挡一部分区域，然后利用自编码器网络来预测遮挡区域的内容，从而学习图像的特征表示。Gidaris 等人提出了一种基于图像旋转预测的图像增强方法。该方法通过随机旋转图像一定角度，然后利用模型预测旋转角度，从而实现图像增强效果^[61]。尽管图像混合方法简易且有效，但仍然存在一些缺陷和局限性，比如样本之间内容或特征差异过大会导致生成样本信息不够清晰或含糊不清，图像混合过程中可能引入额外的噪声和干扰以及混合的计算成本较高等。特别在处理小样本图像时，可能导致混合样本偏向于某些类别或类型。

特征空间增强方法不直接作用于原始图像数据，而是在特征表示空间中生成具有丰富特征的新样本，从而提高模型的泛化能力和鲁棒性。特征空间插值是一种常见的特征增强方法，典型代表为自适应特征插值^[62]，该方法利用特征空间的平坦化效应，通过最近邻特征插值生成新的图像样本，从而增加图像的多样性。另一种特征增强方法则是特征空间转换，将图像特征空间映射到更具有判别性和可分性的流形空间中，以提高特征的代表能力。比如拉普拉斯特征映射的流形学习方法通过图的拉普拉斯矩阵进行特征分解，从而实现数据的降维和流形学习^[63]。彭亚新等人对基于流形假设的骨架序列动作识别算法进行了分析总结，相较于传统欧氏度量下的识别方法，流形更有利于研究非线性结构下的识别^[64]。形状空间理论也是一种流形理论，主要关注数据的几何属性、结构和排列。该理论由于其显著的增强方式便捷性和对小样本数据的不敏感性，在图像处理中得到广泛应用^[65-67]。Na Liu 等人提出了一种基于形状空间的漫画面部表情外推方法，为漫画生成夸张的表情。不同于现有的从原始图像到目标图像的表情插值方法，此方法通过引入在形状空间中去除刚性变换（平移、缩放、和旋转）的表情模型，自动外推漫画中夸张的面部表情，与原始图像相比，结果具有高度的一致性和保真性^[68]。这类方法利用形状空间中点的分布信息来描述形状的局部几何特征。此外，形状空间的信息可以用来指导特征提取和模型训练过程，从而优化数据增强效果。通过准确捕捉形状的几何属性，这些方法显著提升了模型的性能和鲁棒性。并且这类方法对数据集不敏感，面对小样本图像环境时也有不错的表征能力。

上述工作中的数据增强方法都在各自特定领域中发挥重要作用，但面向小样本图像环境，大多数方法对模型提升的效果十分有限并且算法缺乏普适性。因此，如何在小样本图像的任务场景中使用有效且鲁棒的数据增强算法是急需解决的问题，而特征层面的数据增强能够有效避免图像样本失真、模糊等情况，并且特征数据往往具有更好的稳定性，能够更准确地反映数据的本质样本的真实分布，从而有助于提高模型的性能表现，使得该方法具有更广泛的应用场景。

1.4.3 材料性能预测研究现状

材料性能预测是材料科学领域的一个重要研究方向，旨在准确预测材料的各种性能参数，如力学性能、导电率性能等。在这一领域，图像作为一种丰富信息的数据

来源, 被广泛运用于材料性能的挖掘工作中。通过将图像或图像特征作为输入, 并运用机器学习方法来建立图像与性能之间的关系模型。这种方法能够帮助研究人员更好地理解材料的性能特征, 并为新材料的设计和 optimization 提供有价值的参考。例如, Li 等人提出一种基于决策树的集成学习算法, 该算法能够以较高精度来预测化学机械平坦化的材料去除率^[69]; Du 等人采用基于卷积神经网络的深度学习方法, 对烧结银 (sintered silver) 的有效热导率进行预测。该研究包含 6156 个烧结银的真实微观结构数据, 实验结果表明改进的 CNN 优于机器学习的回归模型^[70]; Wei 等人提出一种改进 CNN 结构的模型来预测复合材料和多孔介质的有效热导率^[71]; 根据 6061 铝合金激光振荡焊接实验的金相显微图, Ai 等人采用图像识别技术计算焊缝横截面积, 利用线性回归、多项式回归等机器学习算法, 建立了不同工艺条件下焊缝面积的预测模型^[72]; Song 等人使用卷积神经网络提取材料的微观结构特征, 使用多层感知器提取成分特征, 并使用神经网络对自适应融合的高级特征完成复合材料的力学性能预测^[73]; Ren 等人构建了一个深度学习框架耦合成分的数值信息和多源显微组织的图像信息, 用于预测双相钢的拉伸性能^[74]。然而, 上述方法大多针对特定材料设计相应的性能预测模型, 对于其他材料种类来说, 算法模型存在较大的局限性。此外, 材料图像通常呈现出小样本的特点, 这对于算法模型的训练提出了挑战。为了克服数据量不足的问题, 研究者往往需要借助数据增强技术来扩充数据集规模, 以提高模型的预测精度和泛化能力。因此, 数据增强技术在处理材料性能预测时具有重要的意义。

1.5 论文主要工作

为了解决小样本图像数据规模下的图像处理任务, 本文针对模型处理小样本图像存在的数据信息少、数据多样性不足和算法缺乏普适性三个难点, 通过结合机器学习、深度学习以及形状空间理论, 提出适用于小样本图像的数据增强方法, 并应用于图像分类和基于材料图像的材料性能预测中, 实现对模型性能的显著增强。本文的主要工作和创新如下:

(1) 为了充分挖掘小样本图像数据集的可用信息, 以获取处理图像分类任务时学习模型的最佳性能提升, 本论文设计了基于预形状空间测地曲线的图像特征增强算

法 (FAGC-PSS)。本方法采用拟合小样本图像特征数据, 在预形状空间中构造对应类别的测地曲线, 并沿着曲线生成高质量特征的方式实现数据集的扩充, 有助于模型全面地理解数据的特征和规律, 提高模型的鲁棒性和可靠性。同时 FAGC-PSS 克服特征提取器和特征分类器的模型限制, 并缓解了小样本图像分类时模型性能较差的问题。此外, FAGC-PSS 在交叉熵损失函数中设计随机概率函数和影响因子, 以平衡生成特征和图像特征对模型的作用程度。

(2) 为了缓解材料性能预测任务中小样本的影响, 本论文设计了基于 FAGC-PSS 的材料性能预测方法。本方法通过重新设计测地曲线特征增强模块的下游任务框架结构, 以适应材料图像的性能预测任务。利用伪标签机制对测地曲线的生成特征进行标注, 使得回归模型在有限材料图像样本的情况下, 表现出良好的预测精度和泛化能力。对于不同种类的材料性能预测任务, 该方法能够展现出良好的有效性和普适性。

1.6 论文组织结构

本论文以本文攻读硕士学位期间承担的相关项目研究为基础, 针对模型处理图像小样本存在的数据信息少、数据多样性不足和算法缺乏普适性三个难点。结合机器学习方法、深度学习方法以及形状空间理论, 运用特征增强和伪标签机制等手段, 研究和设计了基于预形状空间测地曲线的图像特征增强方法, 并通过结构设计和实验验证了该方法在图像分类和材料性能预测中的有效性和普适性。

本论文的其他各章节内容安排如下:

第二章介绍了与本论文研究相关的理论与技术。首先介绍了特征增强相关原理与应用, 并对传统特征提取和深度学习特征提取进行分析; 接着, 阐述了深度学习的基本概念, 并介绍了两种新颖的图像处理模型; 然后, 对形状空间理论的基本原理、应用场景和研究进展进行了说明; 接着对伪标签的概念和应用进行了阐述; 最后, 对本论文中所使用的数据增强评价指标进行了描述。

第三章提出一种基于预形状空间测地曲线的图像特征增强方法 (FAGC-PSS)。首先介绍方法的整体流程, 进而详细说明如下内容: 基于深度学习模型的图像特征提取方式; 根据形状空间理论处理特征数据的过程; FAGC-PSS 测地曲线构造方法和特

征生成方式。最后对本特征增强在图像分类任务中进行实验，以验证方法的有效性和普适性。

第四章提出一种基于 **FAGC-PSS** 的材料性能预测方法。首先概述了整体方法，其次介绍材料图像特征提取的方式，接着详细说明 **FAGC-PSS** 模块如何生成特征数据，然后介绍伪标签机制如何对生成特征进行标注，最后说明性能预测模型对生成特征进行训练的方式。通过结合多种机器学习的回归模型，在不同种类的材料性能预测实验中，验证本方法良好的有效性和稳定性。

第五章对全文进行总结与展望。讨论并分析本论文研究成果的优势与不足，并对未来的研究方向提出展望与设想。

第二章 相关理论与技术概述

随着深度学习技术的迅速发展，计算机视觉领域取得了显著进展。然而，深度学习模型通常需要大量标注数据才能发挥出良好性能，在实际应用中，数据稀缺是一个普遍存在的问题。为了充分利用有限的的数据资源，数据增强技术应运而生。数据增强通过对原始数据进行各种变换和扭曲，生成更多样的训练样本，从而提高了模型的性能和泛化能力。本论文的研究目的是针对小样本图像处理任务中存在的三大难点：数据可用信息少、数据多样性不足和算法缺乏普适性，探索适用于图像分类、材料性能预测等小样本图像环境下的特征增强方法。因此，本章节回顾了与本论文相关的理论与技术，为后续章节的方法介绍提供铺垫。

2.1 特征增强

在图像小样本环境下，大部分机器学习和深度学习方法都需要数据增强这一重要技术手段来丰富数据的多样性，从而提升模型性能。然而，面向小样本图像的数据增强方法可能存在以下两个方面问题。一方面，对于图像样本极其稀少或者特定领域的场景中，常规的图像像素级增强方式效果十分有限，甚至产生噪声或者干扰性质的样本数据。另一方面，基于深度学习的生成模型本身就需要大量的图像数据作为支撑，在面临数据多样性不足的环境时容易出现过拟合现象，导致模型生成同质化或无序的图像样本，进一步增加模型学习图像信息的负担。为了解决上述问题，近期的研究尝试从特征层面增强数据的表征能力，即特征增强。

特征增强的核心方法是将图像样本映射到特征空间，并施加一系列数据处理操作，以提取更具代表性和区分性的信息。在处理图像数据时，传统的特征增强方法主要依赖于特征算子来提取图像信息，以提供更具价值的特征表示给机器学习模型。例如，尺度不变特征变换（**Scale-invariant feature transform, SIFT**）特征算子^[75]，其主要思想是在不同尺度和方向上寻找关键点，并对每个关键点周围的图像区域计算梯度直方图，从而描述局部特征。**SIFT** 算子通常由关键点检测、关键点描述和关键点匹配三个主要步骤组成。在特征描述阶段，**SIFT** 将每个关键点周围的图像区域划分为子

区域，并计算每个子区域的梯度幅值和方向，最终将这些信息组合成一个高维向量，作为该关键点的特征描述符。另一个常用的特征算子是方向梯度直方图（Histogram of Oriented Gradient, HOG）^[76]，它主要用于目标检测和图像分类任务。HOG 算子通过计算图像中局部区域的梯度方向直方图来描述图像的纹理和形状信息。具体而言，HOG 将图像分割为多个细胞，并计算每个细胞内像素的梯度方向，然后将这些梯度方向组合成一个直方图，作为该细胞的特征表示。但是，它也存在许多缺点，比如特征算子的手动设计和选择，因为不同的特征算子适用于不同的图像任务和场景，而这可能需要大量的实验和评估工作。其次，特征算子在小样本图像环境下往往表现较差，这是因为小样本数据通常不能充分反映图像的多样性和复杂性，导致特征算子提取的特征无法具备足够的判别性和区分度。

随着深度学习技术的不断发展，卷积神经网络可以不需要手动设计特征提取器就能自动学习图像的特征表示，并表现出更强大的泛化能力和性能。因此，越来越多的特征增强方法倾向于使用深度学习的网络模型进行特征提取和增强。例如，Li 等人提出一种简单的随机特征增强（Simple Feature Augmentation, SFA）^[77]方法用于解决领域泛化下的模型性能问题。通过在训练过程中使用高斯噪声扰动特征嵌入并迭代地估计完整的类条件特征协方差矩阵，验证提出的方法如何朝着训练一个健壮的可泛化模型的方向发展。同样，传统上训练图像识别模型时，潜在特征的矩（如均值和标准差）通常被视为噪声而被移除，以提高稳定性和减少训练时间。然而，在图像生成领域，这些矩扮演着更为重要的角色，能够捕获图像的风格和形状信息。Li 等人提出一种矩交换（Moment Exchange, MoEx）的特征增强方法^[78]，MoEx 通过交换训练图像的特征矩，以及插值目标标签的方式，强制模型从矩信息中提取训练信号，从而提高了模型的泛化能力。

变分自编码器（Variational Autoencoders, VAE）也适用于特征增强，因为解码器可以将编码器提取的低维表示重构为原始图像^[20]。类似地，还可以通过改变网络的输出层，使输出为低维特征而不是类别标签，从而在特征空间中增强卷积神经网络的特征表示。这些特征表示可以训练机器学习模型，例如朴素贝叶斯^[79]、支持向量机^[80]以及全连接层网络^[81]等，从而提高下游任务的模型性能。因此，本论文提出的特征增强方法也借鉴了这一思想，利用深度学习的网络模型进行特征提取，然后

对提取的特征进行一系列的增强处理。这种方法在特征空间内操作，速度快，且与现有的数据增强方法相辅相成，可有效提高网络模型的性能。

2.2 深度学习

近年来，深度学习以其强大的模式识别和表示学习能力，以及在大规模数据集上的训练效果而备受瞩目。随着图形处理单元（GPU）等硬件的快速发展，深度学习模型的训练速度大幅提升，为深度学习技术的广泛应用奠定了基础。此外，深度学习算法的创新也推动了深度学习模型的性能不断提升，使其在各种复杂任务中展现出了强大的能力。

卷积神经网络（Convolution neural network, CNN）是深度学习中流行且广泛使用的算法。它已广泛应用于不同的应用中，例如计算机视觉^[48]、语音处理^[82]和NLP^[83]等。据文献^[84]指出，CNN具备参数共享、稀疏交互和等效表示三个主要优点。为了充分利用输入数据（例如图像信号）的二维结构，CNN采用了局部连接和共享权重的策略，而非传统的全连接网络。这一过程使得网络参数显著减少，从而让网络训练速度更快、训练过程更加容易。在典型的CNN中，有许多卷积层，后面是池化（子采样）层，在最后阶段的层中，通常使用全连接层（Multi-layer perceptron, MLP）。图2.1显示了用于图像分类的CNN架构示例。在CNN中，通常将每个层的输入 x 排列为三个维度，即 $m \times m \times r$ ，其中 m 表示输入的高度和宽度， r 表示深度或通道数（例如，对于RGB图像， $r = 3$ ）。在每个卷积层中，存在多个大小为 $n \times n \times q$ 的滤波器 k 。这里， n 应小于输入图像的尺寸，而 q 可以小于或等于 r 。如前所述，滤波器是局部连接的基础，它们与输入进行卷积并共享相同的参数（权重 W^k 和偏差 b^k ）以生成 k 个特征图（ h^k ），每个特征图的大小为 $m - n - 1$ 。对于MLP，卷积层计算权重与其输入之间的点积，但输入是原始输入体积的小区域。然后，将激活函数 f 或非线性函数应用于卷积层的输出：

$$h^k = f(W^k * x + b^k), \quad (2.1)$$

此后，在子采样层中，对每个特征图进行下采样，以减少网络中的参数，加快训练过程。最后一层（例如，Softmax或SVM）可用于生成分类分数，其中每个分数是给定

实例的特定类别的概率。

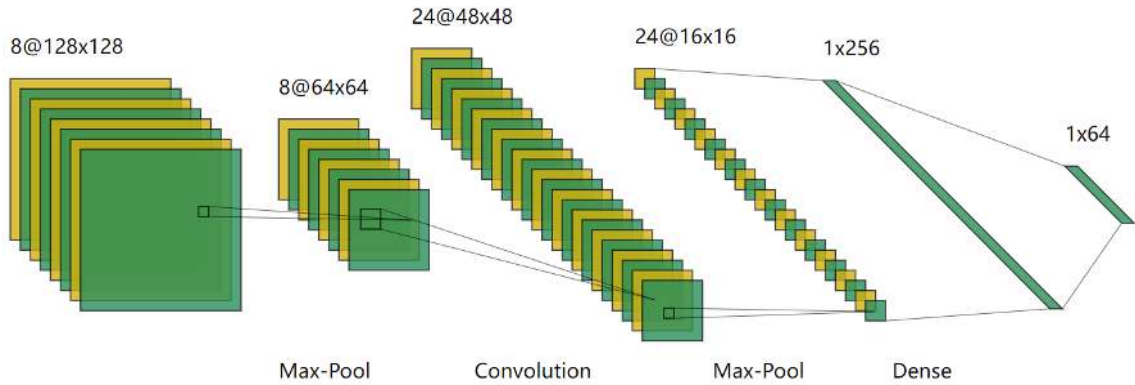


图 2.1 一个简单的卷积神经网络实例架构

在 CNN 模型的发展过程中，提高模型准确度的常见方法包括增加网络的深度或宽度，以及使用更高分辨率的输入图像进行训练和评估。然而，任意增加网络的维度虽然可以提高准确度，但需要手动微调，并且性能提升有限。为了解决这一问题，Tan 在其工作中提出了 EfficientNet 网络架构^[85]。EfficientNet 通过一组固定的缩放系数来调整网络的宽度、深度和分辨率，使得模型在保持精度的同时，减少了计算复杂度和参数量。在网络深度方面，使用了复合函数来表示网络的深度，通过缩放系数来调整每一层的深度。在网络宽度方面，采用了通道缩放的方法来调整每一层的通道数，以实现模型参数量的控制。而在网络分辨率方面，通过对输入图像进行不同程度的缩放来调整模型的输入尺寸，从而适应不同大小的输入图像。其核心思想是，随着输入图像尺寸的增加，网络需要更多的层和通道来有效地处理更高分辨率的图像。此外，EfficientNet 的卓越性能得益于其核心网络构建块，Mobile Inverted Bottleneck Convolution(MBConv)^[86]，它是一种改进的卷积单元如图2.2所示。其次，EfficientNet 还采用了一系列轻量级的模型设计策略，包括使用深度可分离卷积、轻量级的激活函数和批归一化层等。这些设计策略有效地降低了模型的计算复杂度和参数量，使得 EfficientNet 在资源受限的环境下仍能够保持高效的推理和训练能力。

随着图像处理任务的不断发展和复杂化，CNN 的局限性逐渐凸显，尤其是在处理全局信息和长程依赖性方面的能力受到限制。Transformer 架构作为一种在自然语言处理领域取得显著成功的模型^[87]，引起了广泛关注。Transformer 以其自注意力机制为核心，能够有效地捕获序列数据中的全局依赖关系，这使得模型在处理长序列

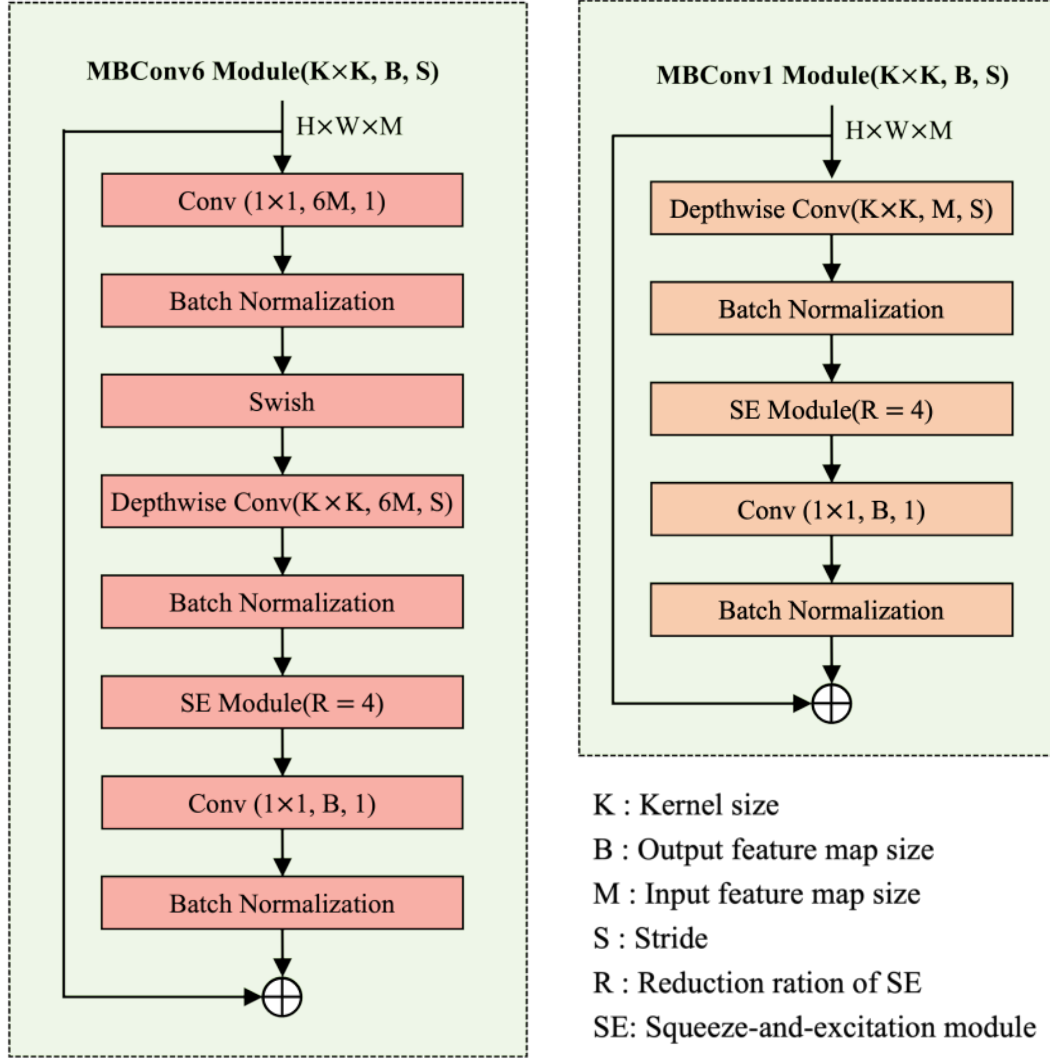


图 2.2 MBConv1 和 MBConv6 的模块架构

数据时表现出色。基于 Transformer 的成功经验，Dosovitskiy 等人于 2020 年提出了 Vision Transformer (ViT) 模型^[88]，其核心思想是将图像视为一个序列，通过自注意力机制来实现全局信息的交互和处理。ViT 的架构主要包含图像的编码器和分类器两个组件，其网络模型和模块结构分别如图2.3和图2.4所示。首先，图像被分割成大小固定的小块（通常为 16×16 像素），每个小块被视为一个 patch。然后，每个 patch 的像素值被展平成一个向量，并加上位置嵌入（position embeddings）以表示其在图像中的位置信息。这些 patch 向量序列被传递给 Transformer 编码器进行处理。Transformer 编码器由多个 Transformer 块组成，每个块包含了多头自注意力层和前馈全连接层。在自注意力层中，每个 patch 向量都与其他 patch 向量进行交互，以捕获全局的图像语义信息。通过多头机制，模型可以同时关注图像的多个方面和局部特征。在前馈

全连接层中，每个 **patch** 向量经过非线性变换和组合，以生成更高级别的特征表示。最后，ViT 的输出被传递给分类器，通常是一个全连接层，用于将图像的特征表示映射到类别概率分布。在训练过程中，ViT 通过最小化分类误差来优化模型参数，从而使得模型能够学习到图像的有效特征表示。

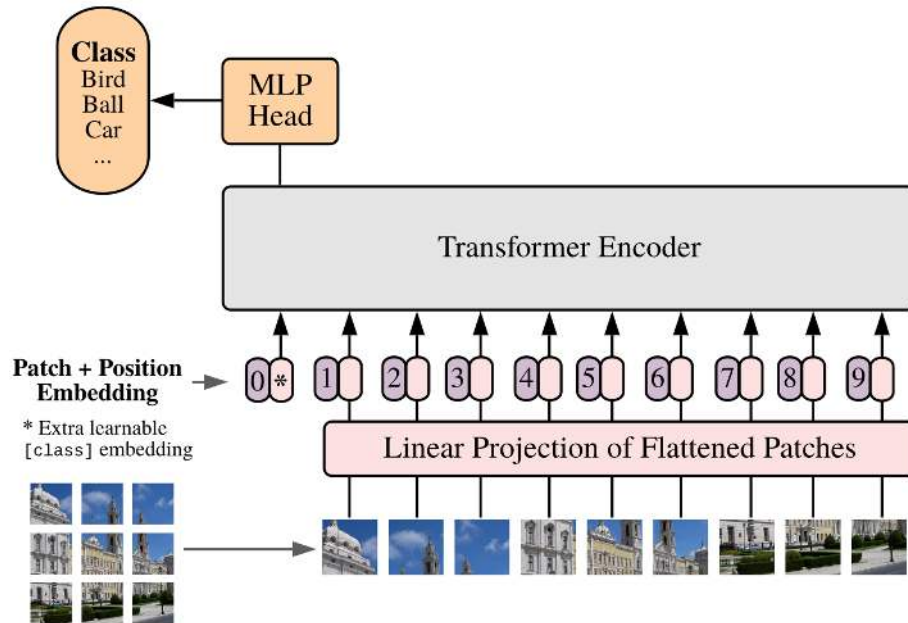


图 2.3 ViT 网络结构图^[88]

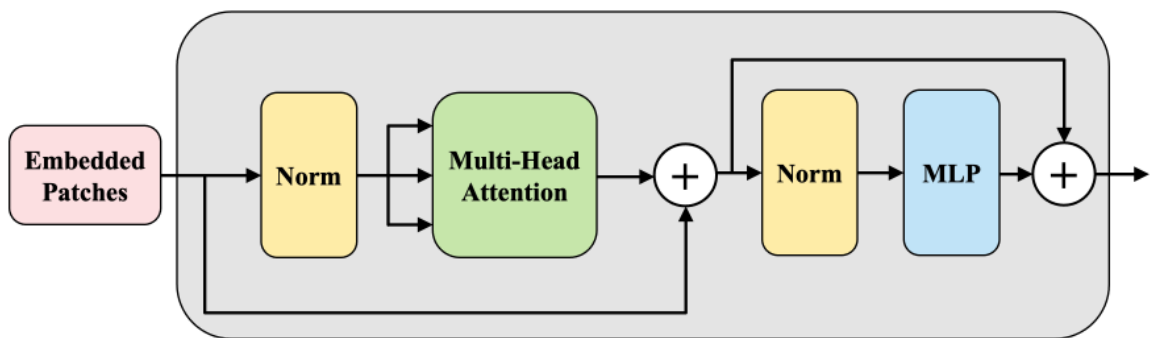


图 2.4 Transformer 编码器模块架构图^[88]

2.3 形状空间理论

形状空间理论是一种用于描述物体形状的重要理论，主要思想是将一个对象及其所有变化表示为一个等价类，并将其映射到非欧几里得空间中的一个点^[89]。在形状空间中，形状的位置、尺度和方向可以忽略不计。形状空间理论经常用于物体识别

场景，通过计算形状空间中观测物体与数据模型之间的测地距离来实现。如果用二维空间的一组坐标来表示二维形状，假设坐标在二维欧氏空间中用特征点表示，即：

$$P = \{p_1(x_1, y_1), p_2(x_2, y_2), \dots, p_n(x_n, y_n)\}, \quad (2.2)$$

其中 n 表示坐标的数量， p_i 表示第 i 个坐标的位置。为了排除形状的位置影响，向量表示为：

$$P' = \{p'_i = (x_i - \bar{x}, y_i - \bar{y})\}, \quad i = 1, \dots, n \quad (2.3)$$

其中 $\bar{x}$ 和 $\bar{y}$ 分别表示 $\{x\}$ 和 $\{y\}$ 的平均值。然后，为了移除尺度的影响，得到在预形状空间的向量 v ：

$$v = \frac{P'}{\|P'\|}, \quad (2.4)$$

其中 $\|\cdot\|$ 表示欧几里得范数，那么含有 p 个坐标的特征向量的所有形状都被嵌入到一个单位的超球体中^[90-92]，记为 S_*^{2p-3} 。任何形状都是这个超球体上的一个点或者向量，而形状的所有变化，即位置、缩放和二维旋转，都会产生一个新的形状，在 S_*^{2p-3} 的一个大圆上，记为 $O(v)$ 。其中 v 表示该形状在 S_*^{2p-3} 。单位超球 S_*^{2p-3} 中所有形状集合构成轨道空间 $\sum_2^p$ ，即如下描述形状空间：

$$\sum_2^p = \{O(v) : v \in S_*^{2p-3}\}, \quad (2.5)$$

图2.5展示了预形状空间和形状空间的一个示例，其中同一个多边形在不同位置、尺度和方向上的形状结构都在 S_*^{2p-3} 的同一个大圆上，该大圆作为形状空间中的一个点或者向量。

与欧几里得距离不同，测地距离被用于定义预形状空间 S_*^{2p-3} 中两点之间的最短距离。假设 v_1 和 v_2 是预形状空间中的两个特征，则 v_1 和 v_2 之间的测地距离定义为^[93]：

$$d(v_1, v_2) = \cos^{-1}(\langle v_1, v_2 \rangle), \quad (2.6)$$

其中 $\langle v_1, v_2 \rangle$ 表示 v_1 和 v_2 之间的内积。形状空间 $\sum_2^p$ 中两个形状之间的距离可以

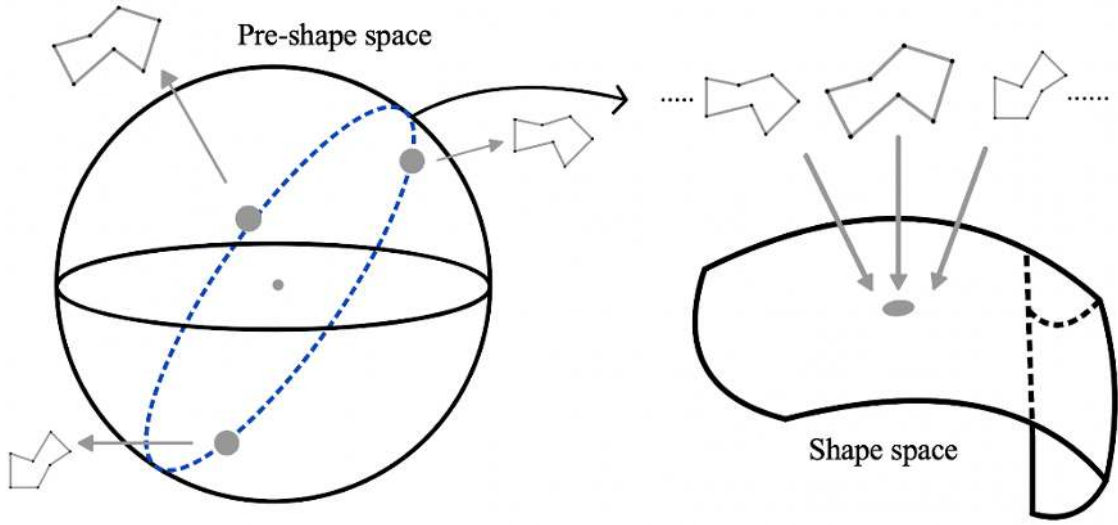


图 2.5 预形状空间和形状空间示意图

定义如下^[91,94]：

$$d_p[O(v_1), O(v_2)] = \inf[\cos^{-1}(\langle \alpha, \beta \rangle) : \alpha \in O(v_1), \beta \in O(v_2)], \quad (2.7)$$

然而，公式中的 d_p 计算非常复杂，通常被投影到复数域空间中，以便公式 (2.7) 可以用 Procrustean 距离描述^[95]：

$$d_p[O(v_1), O(v_2)] = \cos^{-1} \left(\left| \sum_{j=1}^p v_{1j} v_{2j}^* \right| \right), \quad (2.8)$$

其中 v_{1j} 和 v_{2j} 分别表示 v_1 和 v_2 的第 j 个复坐标， v_{2j}^* 表示 v_{2j} 的复共轭。随后，研究人员发现，基于形状空间理论可以减少数据冗余，用于快速目标识别^[96]。进一步的工作包括使用对应的特征点计算两幅图像之间的测地线距离，从而计算出两幅图像中关键对象的相似程度^[97]。

在预形状空间中，测地曲线上的点可用来描述一组变化的形状。**Evans** 等人提出了一种拟合预形状空间中数据的方法，即在预形状空间中获取一组正交点来构建测地曲线^[98]。在此基础上，**Han** 等人提出了一种构建测地曲线的方法，从而利用少量数据扩展预形状空间中的数据，其中曲线的两个端点可以是非正交的^[99]。假设 v_1 和 v_2 表示预形状空间中的两个点，则测地曲线公式可推导如下：

$$\Gamma_{(v_1, v_2)}(s) = (\cos s) \cdot v_1 + (\sin s) \frac{v_2 - v_1 \cdot \cos \theta_{(v_1, v_2)}}{\sin \theta_{(v_1, v_2)}}, \quad (0 \leq s \leq \theta_{(v_1, v_2)}) \quad (2.9)$$

其中 $\theta_{(v_1, v_2)}$ 表示 v_1 和 v_2 之间的测地距离, s 是表示新生成点 v 与 v_1 之间的测地距离的角度。如果 $0 \leq s \leq \theta_{(v_1, v_2)}$, 那么 $v_1 \leq \Gamma_{(v_1, v_2)}(s) \leq v_2$ 。沿着测地曲线 $\Gamma_{(v_1, v_2)}(s)$, 可以在预形状空间中生成一系列新的数据或点, 所有这些点都是 v_1 和 v_2 之间等价变体, 这对于扩展数据集的大小非常有帮助。也就是说通过特征空间转换的方式, 利用形状空间的测地距离和测地曲线对特征数据进行度量和扩充, 是解决小样本图像的潜在有效方法。因此, 本论文设计在预形状空间中测地曲线的构造方法, 以拟合多幅图像的特征, 并沿着曲线生成新的特征, 从而实现特征增强。

2.4 伪标签设计机制

伪标签学习的概念源于半监督学习, 其核心思想是利用相对容易获得的无标签数据来改进有限标签样本条件下模型的性能。伪标签之所以能够提升模型的性能, 主要是因为大多数情况下数据样本的分布具有一定的规律性。在样本空间中, 不同类别样本的分类边界通常位于样本分布的低密度区域。因此, 假设处于同一高密度区域的数据样本对应于相同的标签是合理的。基于这一假设规则, 在小样本图像环境下, 对额外补充的无标签数据赋予的伪标签是相对可靠的, 而相应的图像处理模型的性能也能够获得显著的提升。

然而, 使用伪标签监督训练模型仍然存在一些问题。首先, 由于伪标签未必与真实标签的类别相符, 即使样本的分布符合假设条件, 仍然可能出现类别错误的伪标签, 这会在监督训练过程中引入噪声干扰, 误导模型的学习。此外, 为了确保伪标签的准确性, 可能只对有标注数据极为接近的少数无标签数据赋予伪标签, 这可能导致使用伪标签标注的数据无法为模型提供额外有效的监督信息, 从而使模型只学习已知信息, 引发过拟合问题。因此, 为了实现模型性能的提升, 伪标签的准确性和差异性至关重要。

伪标签的使用通常可以分为以下五个步骤: 首先, 使用有标签数据训练模型; 其次, 使用训练得到的模型对无标签数据进行预测; 然后, 从无标签数据的预测结果中筛选出高置信度的结果作为伪标签; 接着, 使用有标签数据和伪标签数据训练得到一个新的模型; 最后, 不断重复第二到第四步, 直至预测结果趋于稳定^[100]。图2.6展示了伪标签的应用流程图。

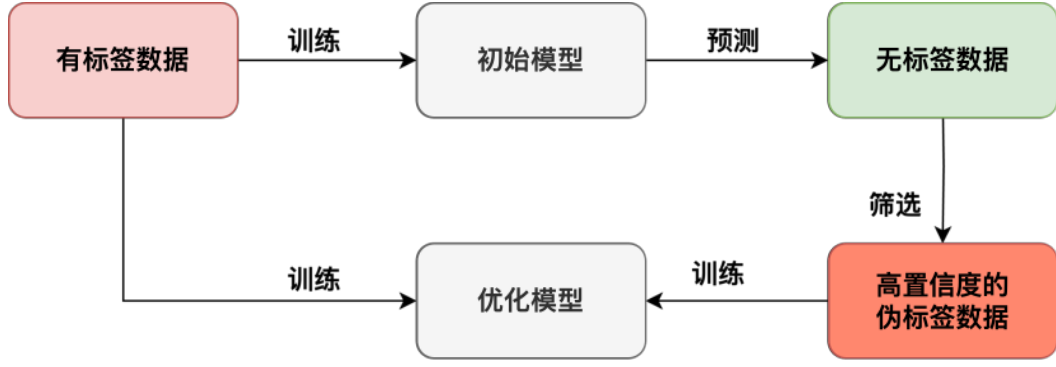


图 2.6 伪标签应用流程示意图

伪标签的思想和技术从诞生至今已被应用于多种任务并取得了显著的效果。在早期研究中，伪标签主要用于解决图像分类工作。例如，Lee 等人利用伪标签机制，在 MNIST 手写数据集^[53]上以极少的标注数据取得了良好的分类效果。近期研究中，FixMatch^[101]、ConMatch^[102]以及 FlexMatch^[103]等方法进一步发展了伪标签的思想，提升了半监督图像分类问题的效果。

2.5 数据增强评价指标

数据增强方法作为学习模型在训练过程中的重要技术，在评估其方法的有效性时，不仅需要考虑方法本身的性能，还需要结合下游图像任务的类型和目标进行判断。本论文提出的数据增强主要应用于图像识别和材料图像性能预测两种任务场景中。对于图像识别任务，通常关注模型对图像进行分类的准确性，因此准确率（Accuracy）是一种常用的评价指标。而对于材料图像性能预测任务中，则更关心模型对图像特征的精确预测能力和泛化能力，常用的评价指标包括平均绝对误差（Mean Absolute Error, MAE）、均方误差（Mean Squared Error, MSE）、均方根误差（Root Mean Squared Error, RMSE）和 R^2 分数（Coefficient of Determination）。上述四个指标可用于衡量模型预测结果与真实值之间的差异程度，从而评估模型的性能。假设 y_i 是第 i 个材料图像样本的真实观测值， $\hat{y}_i$ 是对应的模型预测值。材料图像性能预测结果的平均绝对误差（MAE）由模型预测值与真实观测值之间绝对差异的平均值计算得出，其计算公式如下：

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (2.10)$$

其中 n 表示样本数量。材料图像性能预测结果的均方误差 (**MSE**) 由模型预测值与真实观测值之间差异的平方的平均值计算得到, 其计算公式为:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (2.11)$$

材料图像性能预测结果的均方根误差 (**RMSE**) 为 **MSE** 的平方根:

$$\text{RMSE} = \sqrt{\text{MSE}}, \quad (2.12)$$

材料图像性能预测结果的 R^2 分数衡量模型预测值能够解释目标变量方差的比例, 其取值范围为 $(-\infty, 1]$, 越接近于 1 表示模型拟合效果越好, 其计算公式如下:

$$R^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (2.13)$$

其中 $\bar{y}$ 表示模型预测值的平均值。上述图像性能预测的评价指标能够全面评估模型的性能, 在大多数材料图像性能预测任务中, **MAE** 和 **RMSE** 用于衡量模型的预测误差大小, 而 R^2 分数则用于评估模型的拟合程度。

2.6 本章小结

本章首先介绍了特征增强相关原理与应用, 并对传统特征提取和深度学习特征提取进行分析。接着, 阐述了深度学习的基本概念, 并介绍了两种新颖的图像处理模型 **EfficientNet** 和 **ViT**。然后, 对形状空间理论的基本原理、应用场景和研究进展进行了说明, 并展示了形状空间与预形状空间示意图。接着, 对伪标签学习的概念和应用进行了阐述, 并展示了伪标签的应用流程图。最后, 对本论文中所使用的数据增强评价指标进行了描述。

第三章 基于预形状空间测地曲线的图像特征增强方法

在许多实际的图像处理场景中，数据量往往有限。在图像小样本条件的限制下，数据增强已成为提高模型性能和泛化能力的首要策略，然而，数据增强方法也面临许多挑战。首先，由于原始数据样本较少，生成的增强数据可能缺乏多样性，导致模型容易过拟合或欠拟合。其次，在小样本条件下很难生成高质量的图像样本。此外，数据增强方法还需考虑计算复杂度和时间成本，在保证增强数据质量的前提下，提高数据增强过程的效率。本章针对模型处理图像小样本存在的数据信息少、数据多样性不足和算法缺乏普适性三个难点，以快速地扩充任一类别的图像特征数据集，提出一种基于预形状空间中测地曲线的图像特征增强方法（Feature Augmentation on Geodesic Curves in Pre-Shape Space, FAGC-PSS）。FAGC-PSS 通过在预形状空间中拟合训练样本的特征向量构造测地曲线，以生成更多的特征向量，结合机器学习模型实现对小样本的图像分类任务，其步骤可以分为四个部分：图像样本特征的提取、投影图像特征至预形状空间、构造测地曲线拟合特征向量以及沿曲线生成新特征。

上述过程如图3.1所示。

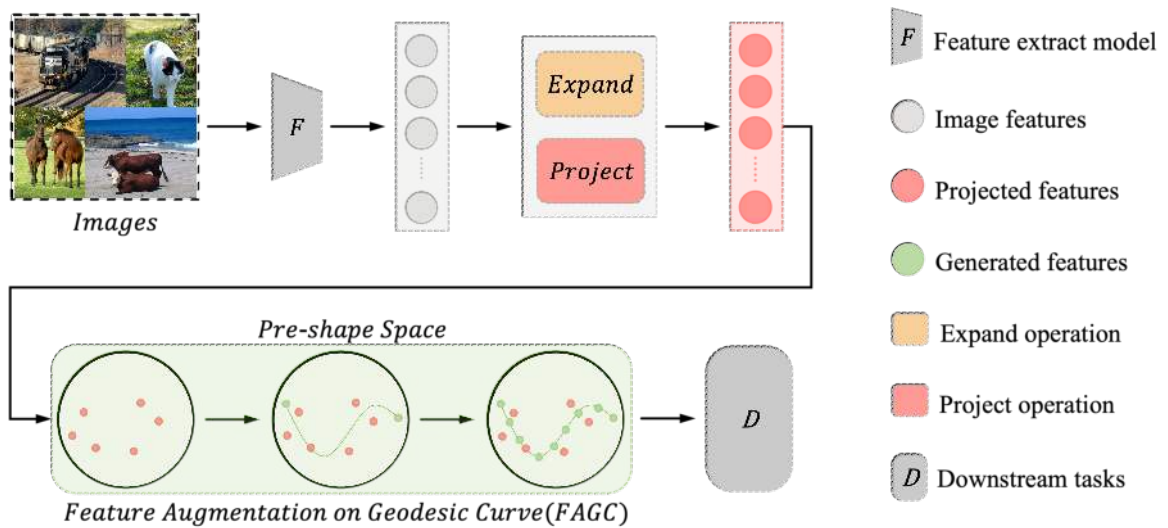


图 3.1 FAGC-PSS 方法流程图

3.1 方法概述

3.1.1 图像特征的提取

本论文采用深度学习方法中的神经网络进行图像特征提取。深度学习模型与传统机器学习方法不同，不需要预先设计特征提取器，而是通过多层次的非线性变换，从原始图像数据中提取到高质量的特征。通常来说，越复杂的神经网络模型，提取到的图像特征就越有更高的判别性，从而提升下游任务中模型的性能。在本章中，采用 ViT^[88] 作为骨干网络对图像数据进行特征提取，该模型以其独特的自注意力机制和位置编码方式在各种复杂的图像场景中表现出色，能够有效地理解图像的整体结构和语义信息。我们在 ImageNet 数据集上对 ViT 进行预训练，并将训练结果作为后续处理的权重参数。然后，把小样本训练图像输入到预训练好的 ViT 网络中对其进行微调，直至损失不再变化时停止迭代，此时选取 ViT 网络的输出数据作为提取到的特征应用到下一步处理中。

3.1.2 投影图像特征至预形状空间

根据形状空间的理论，一个形状结构的顶点通常用二维坐标或三维坐标进行表示，本论文采用的是二维坐标来表示。因此，从 ViT 网络提取的特征在投影到预形状空间之前，需要提升到原始维度的两倍。上述步骤可以通过插值或转换等技术，来调整形状结构所需的维度空间。本方法升维方式是将一个特征复制为两个相同的特征。假设一张图像的特征被表示为 $V_0 = \{x_1, x_2, \dots, x_n\}$ ，其中 n 为该图像特征的维度数。因此，升维函数表示为

$$y = f(x) = x, \quad (3.1)$$

其中图像特征 V_0 转换为 $V_1 = \{x_1, y_1, x_2, y_2, \dots, x_n, y_n | y_i = x_i, i \in [1, n]\}$ 。为了减轻特征位置的影响，特征 V_1 的平均值被移除并得到新的特征 $V_2 = \{x'_1, y'_1, x'_2, y'_2, \dots, x'_n, y'_n | x'_i = x_i - X_\mu, y'_i = y_i - Y_\mu, i \in [1, n]\}$ ，其中 X_μ, Y_μ 分别表示 $\{x_i\}$ 和 $\{y_i\}$ 的平均值。然后，采用公式 $z = V_2 / \|V_2\|$ 排除特征 V_2 的缩放影响，其中 $\|V_2\|$ 表示 V_2 的欧几里得范数。经过上述步骤，图像特征 V_0 被投影到一个预形状空间中，记为 z 。

3.1.3 构造测地曲线拟合特征

在预形状空间中构造一条测地曲线时，通常需要至少两个或两个以上的图像样本。如果为了扩充不同类别的图像特征，则需要构造多条测地曲线与每个类别进行对应。如上小节所述，通过 ViT 网络从多张图像中提取特征向量，并投影至预形状空间中，训练特征集表示为 $Z = \{z_1^1, \dots, z_1^{M_1}, \dots, z_N^1, \dots, z_N^{M_N}\}$ ，其中 N 表示图像总的类别数， M_i 表示第 i 类的图像样本数量以及 z_i^j 表示第 i 类图像集中的第 j 个样本。在预形状空间中多类别图像的测地曲线集表示为 $\{\Gamma_{(v_i^*, w_i^*)} | i \in (1, N)\}$ ，其中 v_i^* 和 w_i^* 表示为对应类别 i 的最优测地曲线的两个端点，对应公式(2.9)中的 v_1 和 v_2 。为了获取最优测地曲线 $\Gamma_{(v_i^*, w_i^*)}(\cdot)$ 中的 v_i^* 和 w_i^* ，首要步骤是初始化 v_i^* ，通过计算特征向量集 Z 中所有特征向量之间的测地距离，选取与其它向量距离总和最大的特征向量作为初始化参数 v_i^* 的样本点，定义如下：

$$v_i^* = \underset{z_i^k}{\operatorname{argmax}} \sum_{j=1}^{M_i} d(z_i^k, z_i^j), k \in \{1, \dots, M_i\} \quad (3.2)$$

其中 $d(z_i^k, z_i^j)$ 表示使用公式(2.8)计算 z_i^k 和 z_i^j 之间的测地距离。假设 w_i^0 和 w_i^1 是距离 v_i^* 最远的两个特征向量，如图3.2(c)所示，以 v_i^* 作为起始点， w_i^0 和 w_i^1 由以下公式计算：

$$w_i^0 = \underset{Q_i - \{v_i^*\}}{\operatorname{argmax}} (d(v_i^*, Q_i - \{v_i^*\})), \quad (3.3)$$

和

$$w_i^1 = \underset{Q_i - \{v_i^*, w_i^0\}}{\operatorname{argmax}} (d(v_i^*, Q_i - \{v_i^*, w_i^0\})), \quad (3.4)$$

其中 $Q_i = \{z_i^1, z_i^2, \dots, z_i^{M_i}\}$ 表示第 i 类的特征集。 $Q_i - \{\#\}$ 表示如果特征向量 $\{\#\}$ 存在于特征集 Q_i ，那么该特征将从 Q_i 中移除。然后，使用公式(2.9)获取 w_i^0 和 w_i^1 之间的测地曲线 $\Gamma_{(w_i^0, w_i^1)}(\cdot)$ 。接着沿着测地曲线 $\Gamma_{(w_i^0, w_i^1)}(\cdot)$ 上生成一组特征 $\hat{w}_i^{\{1, \dots, S\}}$ ，这些特征可以表示为 $\{\Gamma_{(w_i^0, w_i^1)}(s_i^1), \Gamma_{(w_i^0, w_i^1)}(s_i^2), \dots, \Gamma_{(w_i^0, w_i^1)}(s_i^S)\}$ ，其中 S 表示测地曲线 $\Gamma_{(w_i^0, w_i^1)}(\cdot)$ 上临时特征的数量。随后， v_i^* 分别与 $\hat{w}_i^{\{1, \dots, S\}}$ 组合得到多个测地曲线 $\Gamma_{(v_i^*, \hat{w}_i^{\{1, \dots, S\}})}(\cdot)$ ，接着分别计算第 i 类特征 $\{z_i^1, z_i^2, \dots, z_i^{M_i}\}$ 到测地曲线 $\Gamma_{(v_i^*, \hat{w}_i^{\{1, \dots, S\}})}(\cdot)$

的测地距离平方和，其中最小的距离平方和对应的特征 w_i^* 计算如下：

$$w_i^* = \underset{\hat{w}_i^k}{\operatorname{argmin}} \sum_{j=1}^{M_i} \left(d \left(z_i^j, \Gamma_{(v_i^*, \hat{w}_i^k)}(\cdot) \right) \right)^2, k \in \{1, \dots, S\} \quad (3.5)$$

其中 $d \left(z_i^j, \Gamma_{(v_i^*, \hat{w}_i^k)}(\cdot) \right)$ 为特征向量 z_i^j 到测地曲线 $\Gamma_{(v_i^*, \hat{w}_i^k)}(\cdot)$ 的测地距离，此时 w_i^* 为该阶段最优测地曲线的端点，并且该步骤如图3.2(f)和图3.2(g)所示。随后，将特征向量 w_i^* 赋值给 v_i^* ，同时计算出距离 v_i^* 最远的两个特征，如图3.2(h)所示。重复前面步骤，即图3.2(i)到(j)，直至 v_i^* 和 w_i^* 不再发生变化时停止迭代。最后， $\Gamma_{(v_i^*, w_i^*)}(\cdot)$ 表示为拟合特征集 z_i^* 分布的最优测地曲线，如图3.2(k)所示。算法3.1概述了拟合特定类别特征的 FAGC-PSS 的详细过程。

不同类别图像的测地曲线 $\{\Gamma_{(v_i^*, w_i^*)}(\cdot) | i \in \{1, N\}\}$ 都可以通过上述方法进行构造，并且利用公式(2.8)计算出所有组 v_i^* 和 w_i^* 之间的测地距离 $d(v_i^*, w_i^*)$ ，表示在测地曲线函数 $\Gamma_{(v_i^*, w_i^*)}(\cdot)$ 中角度的大小。为了增加生成特征向量的多样性，采用均匀分布 $s \sim U(0, d(v_i^*, w_i^*))$ 对一组角度值 $\{s_i^j | j \in \{1, K\}\}$ 进行随机抽样，其中 K 表示第 i 类测地曲线采样角度值的数量。新特征 $Z' = \{z_1'^1, \dots, z_1'^K, \dots, z_N'^1, \dots, z_N'^K\}$ 是通过将 s_i^j 输入到第 i 类的测地曲线 $\Gamma_{(v_i^*, w_i^*)}(\cdot)$ 获得的。这些特征向量和训练图像的原始特征，构成了扩充的特征向量集 $\{z_i^j, z_i'^k | i \in (1, N), j \in (1, M_i), k \in (1, K)\}$ 。FAGC-PSS 能够使模型更全面地学习到数据的特征表示，获得的特征可以用于其它机器学习模型的训练，从而提高了模型的泛化能力和鲁棒性。此外，FAGC-PSS 模块在小样本环境下，可以很好的改善了小样本带来的局限，增加了数据的多样性和变化性，使得下游任务的模型能够在数据较少的情况下仍然取得良好的性能。进而提高模型的性能和泛化能力。

3.1.4 改进的损失函数

一般来说，使用 FAGC-PSS 模块生成的特征数量会远远超过原始训练特征的数量，为了平衡两者数量差距和减弱模型偏向性学习的影响，本方法提出一种适用于 FAGC-PSS 的交叉熵损失函数，这里选择交叉损失函数作为基本损失函数是因为它可以更敏感地反映预测类别和真实类别之间的差异。具体而言，采用标准的交叉熵损失函数，对机器学习模型在原始特征向量和生成特征向量上的性能进行评估，分

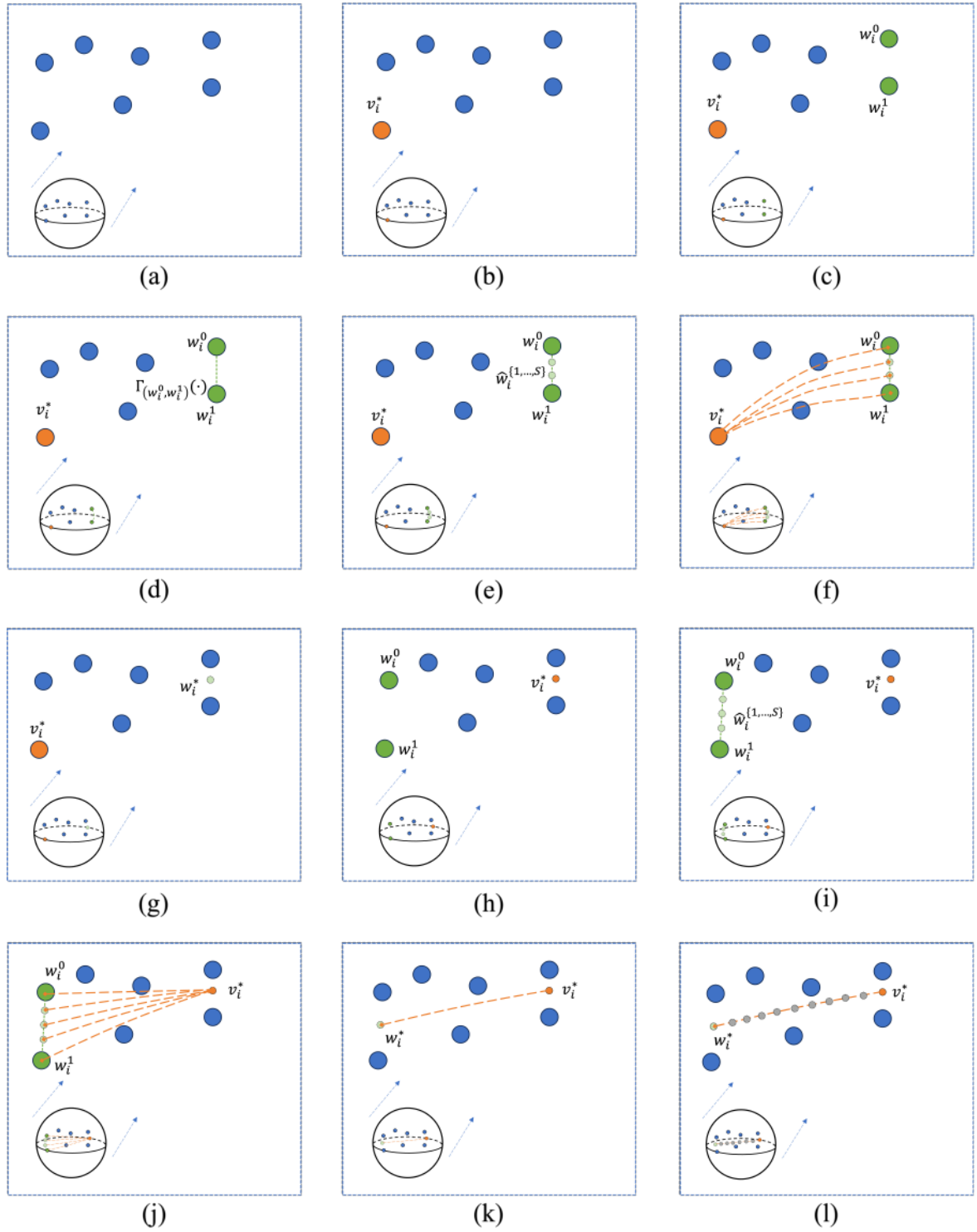


图 3.2 FAGC-PSS 方法的处理流程。超球体是一个预形状空间；蓝色实心点表示预形状空间中的特征；橙色实心点代表参数样本点 v_i^* ；绿色实心点表示距离特征 v_i^* 最远的两个特征 w_i^0 和 w_i^1 。

算法 3.1: FAGC-PSS 算法流程

Input: 训练特征集 $Q_i = \{z_i^1, \dots, z_i^{M_i}\}$, 第 i 类的特征样本数 M_i , 生成特征数量 K 和测地曲线上临时特征数量 S

- 1: $v_i^* = \underset{z_i^k}{\operatorname{argmax}} \sum_{j=1}^{M_i} d(z_i^k, z_i^j), k \in \{1, \dots, M_i\};$
- 2: **while** v_i^*, w_i^* not converge **do**
- 3: $w^0 = \underset{Q_i - \{v_i^*\}}{\operatorname{argmax}} (d(v_i^*, Q_i - \{v_i^*\}));$
- 4: $w^1 = \underset{Q_i - \{v_i^*, w^0\}}{\operatorname{argmax}} (d(v_i^*, Q_i - \{v_i^*, w^0\}));$
- 5: $\hat{w}_i^{\{1, \dots, S\}} = \{\Gamma_{(w^0, w^1)}(s \cdot 1), \Gamma_{(w^0, w^1)}(s \cdot 2), \dots, \Gamma_{(w^0, w^1)}(s \cdot S)\};$
- 6: $w_i^* = \underset{\hat{w}_i^k}{\operatorname{argmin}} \sum_{j=1}^{M_i} (d(z_i^j, \Gamma_{(v_i^*, \hat{w}_i^k)}(\cdot)))^2, k \in \{1, \dots, S\};$
- 7: Assign w_i^* to v_i^* ;
- 8: **end while**
- 9: Calculate $d(v_i^*, w_i^*)$ with Formula(2.8);
- 10: **for** $k = 1$ to K **do**
- 11: Sample $\hat{s}_i^k \sim U(0, d(v_i^*, w_i^*));$
- 12: $\hat{z}_i^k = \cos(\hat{s}_i^k) \cdot v_i^* + \sin(\hat{s}_i^k) \frac{w_i^* - v_i^* \cdot \cos \theta_{(v_i^*, w_i^*)}}{\sin \theta_{(v_i^*, w_i^*)}};$
- 13: **end for**

Output: 最佳测地曲线 $\Gamma_{(v_i^*, w_i^*)}(\cdot)$ 和生成特征向量集 $Z' = \{z_i'^1, \dots, z_i'^K\}$

别记为 L_r 和 L_g 。同时引入两个关键的控制因素, 即随机概率因子 P_g 和损失影响因子 λ 。这样, 在每个训练批次中, 分类模型的损失函数 L_{cls} 定义如下:

$$L_{cls} = L_r + P_g \times \lambda \times L_g, \quad (3.6)$$

其中 P_g 是一个分段函数, 可以表示为

$$P_g(p) = \begin{cases} 0, & 0 \leq p < 0.5 \\ 1, & 0.5 \leq p \leq 1 \end{cases} \quad (3.7)$$

其中 p 为随机数, 取值范围为 $[0, 1]$ 。在某个训练批次中, 利用 P_g 可以随机选择是否使用生成特征进行训练, 如果 $P_g = 1$, 那么分类模型的损失函数 L_{cls} 需要结合生成特征的交叉熵损失 L_g , 否则, 如果 $P_g = 0$, 生成特征将在这一批次训练中被忽略。 λ 作为超参数用于调整生成的特征在整体损失函数中的贡献程度。因此, 分类模型通过 P_g 和 λ 来控制原始特征和生成特征对模型的影响程度, 以达到最佳的训练效果。

3.2 实验分析

3.2.1 数据集和实验环境设置

为了在公共图像数据集上模拟小样本图像场景, 本文对五个图像数据集的训练样本进行抽样, 这五个数据集包括 CIFAR-10^[104]、CIFAR-100^[104]、STL-10^[105]、EuroSAT^[106-107] 和 TinyImageNet^[108]。具体抽样方案如下:

- CIFAR-10 和 CIFAR-100 图像数据集由 50000 张训练样本和 10000 张测试样本组成的, 图像种类分别为 10 种和 100 种, 每张图像尺寸为 32*32 像素大小。抽样方案为每个类别的训练图像样本数为 4 张、20 张和 100 张, 即从 CIFAR-10 中抽出的训练样本数分别是 40 张、200 张和 1000 张, 从 CIFAR-100 抽出的训练样本数分别是 400 张、2000 张和 10000 张。
- STL-10 图像数据集由 5000 张标记样本和 8000 张测试样本组成, 抽样方案为每个类别的训练样本数为 4 张、10 张和 40 张, 即从 STL-10 中抽出的训练样本数分别是 40 张、100 张和 400 张。
- EuroSAT 数据集, 它由 10 种类别组成并包含 27000 张有标记的地理参考图像, 抽样方案为每个类别的训练样本数为 2 张、4 张和 20 张, 即从 EuroSAT 中抽出的训练样本数分别是 20 张、40 张和 200 张。
- TinyImageNet 数据集, 它包含 200 种类别且每个类别有 500 张训练图像共有 100000 张图像, 抽样方案为每个类别的训练样本数为 4 张、10 张和 25 张, 即从 TinyImageNet 中抽出的训练样本数分别是 400 张、1000 张和 2500 张。

所有实验均在 Intel(R) Core(TM) i7-11700 CPU 2.9GHz*32、64G RAM 和 RTX 3090 GPU 的计算机上执行。

3.2.2 相关方法对比

为了充分验证 FAGC-PSS 对小样本图像环境下分类模型的增强效果, 本文针对 CIFAR-10、CIFAR-100、STL-10、Euro-SAT 和 TinyImageNet 共五个公共图像数据集, 采用 FAGC-PSS 方法对图像进行特征增强, 通过多轮迭代寻找最优测地曲线来生成高质量特征数据, 验证提出的方法在小样本图像分类问题上的有效性和普适性。其中选取 CIFAR-10、CIFAR-100 数据集, 通过详细的实验与其他数据增强方法进行对比, 以验证 FAGC-PSS 的灵活性。此外, 本文将特征提取模型的分类结果作为基准, 对比应用 FAGC-PSS 后的分类精度, 以验证特征提取模型与 FAGC-PSS 之间的关联程度。

在实验中, 本文共选择 8 种不同的神经网络模型作为全监督机器学习模型用于图像的特征提取及图像分类, 包括 ResNet101^[50]、ResNet152^[50]、EfficientNet-B0^[85]、EfficientNet-B7^[85]、ViT-T、ViT-S、WRN-28-2^[109]和 WRN-28-8^[109]。

表 3.1 基于 CIFAR-10 抽样数据集对比是否采用 FAGC-PSS 的图像分类准确率 (%)

训练样本数		40		200		1000	
网络模型	增强方案	/	+FAGC-PSS	/	+FAGC-PSS	/	+FAGC-PSS
ResNet101		64.32	67.17(2.85 ↑)	86.32	88.15(1.83 ↑)	90.78	91.34(0.56 ↑)
ResNet152		65.16	68.42(3.26 ↑)	88.94	90.17(1.23 ↑)	91.45	92.63(1.18 ↑)
EfficientNet-B0		66.53	69.54(3.01 ↑)	87.42	89.23(1.81 ↑)	90.13	91.43(1.30 ↑)
EfficientNet-B7		68.41	71.63(3.22 ↑)	89.04	91.17(2.13 ↑)	93.54	94.37(0.83 ↑)
ViT-T		79.32	82.83(3.51 ↑)	90.13	91.77(1.64 ↑)	96.21	96.35(0.14 ↑)
ViT-S		90.00	93.31(3.31 ↑)	96.45	97.12 (0.67 ↑)	98.19	98.26 (0.07 ↑)
WRN-28-2		92.89	93.77(0.88 ↑)	93.92	94.17(0.25 ↑)	94.30	94.58(0.28 ↑)
WRN-28-8		93.29	94.51 (1.22 ↑)	93.32	94.53(1.21 ↑)	95.32	95.68(0.36 ↑)

通过观察表3.1、表3.2、表3.3、表3.4和表3.5所展示的各网络模型分类准确率结果可以发现, 引入 FAGC-PSS 后图像分类结果有着良好的精度提升。在大多数情况下, 使用样本数量较少的抽样数据集时, 比如训练样本数为 40 的 CIFAR-10 图像数据集中, WRN-28-2 网络模型结合 FAGC-PSS 能够获得最佳分类结果; 而在训练样本数较

表 3.2 基于 CIFAR-100 抽样数据集对比是否采用 FAGC-PSS 的图像分类准确率 (%)

训练样本数		400		2000		10000	
网络模型	增强方案	/	+FAGC-PSS	/	+FAGC-PSS	/	+FAGC-PSS
ResNet101		52.31	56.72(4.41 ↑)	62.16	64.24(2.08 ↑)	72.41	73.15(0.74 ↑)
ResNet152		53.57	57.32(3.75 ↑)	63.72	65.12(1.40 ↑)	73.15	74.07(0.92 ↑)
EfficientNet-B0		57.53	59.63(2.10 ↑)	65.41	67.65(2.24 ↑)	75.72	76.42(0.70 ↑)
EfficientNet-B7		58.32	61.23(2.91 ↑)	66.46	68.15(1.69 ↑)	76.28	77.32(1.04 ↑)
ViT-T		63.02	66.69(3.67 ↑)	74.39	76.54(2.15 ↑)	80.96	82.61(1.65 ↑)
ViT-S		72.65	75.36(2.71 ↑)	81.20	83.43 (2.23 ↑)	86.28	87.39 (1.11 ↑)
WRN-28-2		68.92	72.09(3.17 ↑)	74.73	76.33(1.60 ↑)	75.59	76.21(0.62 ↑)
WRN-28-8		75.73	77.21 (1.48 ↑)	78.82	79.56(0.74 ↑)	81.23	82.12(0.89 ↑)

表 3.3 基于 STL-10 抽样数据集对比是否采用 FAGC-PSS 的图像分类准确率 (%)

训练样本数		40		100		400	
网络模型	增强方案	/	+FAGC-PSS	/	+FAGC-PSS	/	+FAGC-PSS
ResNet101		63.17	66.15(2.98 ↑)	73.74	75.49(1.75 ↑)	79.32	80.13(0.81 ↑)
ResNet152		64.75	67.28(2.53 ↑)	74.83	76.19(1.36 ↑)	80.17	80.98(0.81 ↑)
EfficientNet-B0		67.24	70.58(3.34 ↑)	76.53	78.38(1.85 ↑)	82.65	83.96(1.31 ↑)
EfficientNet-B7		69.52	72.42(2.90 ↑)	78.73	80.05(1.32 ↑)	84.73	85.03(0.30 ↑)
ViT-T		75.41	78.53(3.12 ↑)	86.69	88.43(1.74 ↑)	92.58	93.47(0.89 ↑)
ViT-S		79.54	82.43 (2.89 ↑)	88.42	90.14 (1.72 ↑)	94.32	95.03 (0.71 ↑)
WRN-28-2		74.54	77.52(2.98 ↑)	85.41	87.17(1.76 ↑)	91.24	92.42(1.18 ↑)
WRN-28-8		78.53	80.62(2.09 ↑)	87.52	89.18(1.66 ↑)	93.52	94.32(0.80 ↑)

多的情况下，比如训练样本数为 1000 的 CIFAR-100、训练样本数为 400 的 STL-10、训练样本数为 200 的 EuroSAT 以及训练样本数为 5000 的 TinyImageNet，ViT-S 网络模型结合 FAGC-PSS 则表现出最佳分类性能。在比较五个数据集的图像分类任务中，FAGC-PSS 在训练样本数量较少的情况下提升分类精度最为显著。随着原始训练样本数量的增加，基于深度学习的网络模型会学习到更全面的特征分布，从 FAGC-PSS

表 3.4 基于 EuroSAT 抽样数据集对比是否采用 FAGC-PSS 的图像分类准确率 (%)

训练样本数		20		40		200	
网络模型	增强方案	/	+FAGC-PSS	/	+FAGC-PSS	/	+FAGC-PSS
ResNet101		62.75	65.16(2.41 ↑)	73.69	74.84(1.15 ↑)	81.63	82.27(0.64 ↑)
ResNet152		63.27	65.97(2.70 ↑)	74.19	75.27(1.08 ↑)	82.18	83.41(1.23 ↑)
EfficientNet-B0		65.58	67.37(1.79 ↑)	76.19	78.48(2.29 ↑)	85.31	86.42(1.11 ↑)
EfficientNet-B7		66.95	68.56(1.61 ↑)	77.53	79.23(1.70 ↑)	86.42	87.18(0.76 ↑)
ViT-T		70.47	73.19(2.72 ↑)	79.53	81.28(1.75 ↑)	91.47	92.17(0.70 ↑)
ViT-S		72.15	74.58(2.43 ↑)	81.69	83.17 (1.48 ↑)	93.79	94.24 (0.45 ↑)
WRN-28-2		69.54	72.37(2.83 ↑)	79.11	82.09(2.98 ↑)	89.67	90.26(0.59 ↑)
WRN-28-8		74.57	76.49 (1.92 ↑)	81.58	82.96(1.38 ↑)	92.17	93.45(1.28 ↑)

表 3.5 基于 TinyImageNet 抽样数据集对比是否采用 FAGC-PSS 的图像分类准确率 (%)

训练样本数		800		2000		5000	
网络模型	增强方案	/	+FAGC-PSS	/	+FAGC-PSS	/	+FAGC-PSS
ResNet101		36.42	39.17(2.75 ↑)	48.14	50.57(2.43 ↑)	55.32	56.26(0.94 ↑)
ResNet152		37.52	39.85(2.33 ↑)	49.32	51.13(1.81 ↑)	56.73	57.41(0.68 ↑)
EfficientNet-B0		41.24	44.38(3.14 ↑)	52.17	54.28(2.11 ↑)	61.23	62.43(1.20 ↑)
EfficientNet-B7		42.51	45.19(2.68 ↑)	53.48	55.28(1.80 ↑)	62.16	63.07(0.91 ↑)
ViT-T		53.17	56.59(3.42 ↑)	64.79	66.27(1.48 ↑)	70.58	71.16(0.58 ↑)
ViT-S		54.28	57.57(3.29 ↑)	65.83	67.72 (1.89 ↑)	71.29	72.05 (0.76 ↑)
WRN-28-2		51.84	53.59(1.75 ↑)	62.18	64.25(2.07 ↑)	68.78	69.23(0.45 ↑)
WRN-28-8		55.78	58.74 (2.96 ↑)	64.86	66.11(1.25 ↑)	70.36	71.28(0.92 ↑)

扩充的特征数据上学习效果会逐渐减弱,进而性能提升的幅度也相应减少。因此实验验证 FAGC-PSS 更适用于小样本图像的数据集规模。

为了深入展示 FAGC-PSS 的增强效果,本文对比一些图像分类算法,包括了十五种半监督的图像分类算法,半监督图像分类算法包括 CoMatch 图像算法^[110]、Pi-Model 图像算法^[111]、MeanTeacher 图像分类算法^[112]、Pseudo Labeling 图像分类算

法^[100]、MixMatch 图像分类算法^[113]、VAT 图像分类算法^[114]、Dash 图像分类算法^[115]、CRMatch 图像分类算法^[116]、UDA 图像分类算法^[117]、FixMatch 图像分类算法^[101]、FlexMatch 图像分类算法^[103]、AdaMatch 图像分类算法^[118]、FreeMatch 图像分类算法^[119]、SimMatch 图像分类算法^[120]和 SoftMatch 图像分类算法^[121]。本实验保持了数据集的训练样本数量与第3.2.1节的实验设置一致，相较于全监督图像分类算法直接以抽取后的数据集进行训练，半监督学习图像分类算法还将未抽取的图像作为无标签样本一并进行训练。

表 3.6 基于 CIFAR-10 使用多种学习算法结合 FAGC-PSS 得到的模型分类精度 (%)

训练样本数		40		200	
学习算法	增强方案	/	+FAGC-PSS	/	+FAGC-PSS
CoMatch		88.24	90.95(2.71 ↑)	93.64	94.39(0.75 ↑)
II-Model		80.66	82.56(1.90 ↑)	92.90	92.77(0.13 ↓)
Mean Teacher		81.21	84.33(3.12 ↑)	92.64	92.43(0.21 ↓)
Pseudo Labeling		81.92	82.44(0.52 ↑)	94.15	94.57(0.42 ↑)
MixMatch		81.81	85.07(3.26 ↑)	93.83	94.31(0.48 ↑)
VAT		89.92	90.21(0.29 ↑)	95.28	95.14(0.14 ↓)
Dash		93.20	93.55(0.35 ↑)	95.30	95.90(0.60 ↑)
CRMatch		92.79	93.31(0.52 ↑)	95.41	95.89(0.48 ↑)
UDA		91.74	92.13(0.39 ↑)	95.33	95.78(0.45 ↑)
FixMatch		91.40	91.84(0.44 ↑)	95.08	95.35(0.27 ↑)
FlexMatch		92.57	93.08(0.51 ↑)	95.75	95.63(0.12 ↓)
AdaMatch		90.95	92.96(2.01 ↑)	95.12	95.58(0.46 ↑)
FreeMatch		93.57	94.07(0.50 ↑)	95.73	96.11(0.38 ↑)
SimMatch		92.78	93.18(0.40 ↑)	95.49	95.77(0.28 ↑)
SoftMatch		93.27	93.68(0.41 ↑)	95.72	95.55(0.17 ↓)
平均精度		88.45	89.76 (1.31 ↑)	94.46	94.81 (0.35 ↑)

对于 CIFAR-10 和 CIFAR-100 数据集，分别采用 ViT-T 和 ViT-S 作为模型的骨干网络结构进行训练。通过观察表3.6和表3.7所展示的各学习算法分类结果评价指标可

表 3.7 基于 CIFAR-100 使用多种学习算法结合 FAGC-PSS 得到的模型分类精度 (%)

训练样本数		400		2000	
学习算法	增强方案	/	+FAGC-PSS	/	+FAGC-PSS
	CoMatch	62.37	68.10(5.73 ↑)	80.71	82.84(2.13 ↑)
	II-Model	54.04	69.29(15.25 ↑)	82.90	84.12(1.22 ↑)
	Mean Teacher	70.50	75.96(5.46 ↑)	81.95	83.68(1.73 ↑)
	Pseudo Labeling	73.51	76.29(2.78 ↑)	82.61	84.42(1.81 ↑)
	MixMatch	73.87	76.06(2.19 ↑)	84.07	84.58(0.51 ↑)
	VAT	77.33	80.40(3.07 ↑)	83.91	85.43(1.52 ↑)
	Dash	78.74	82.69(3.95 ↑)	86.46	87.17(0.71 ↑)
	CRMatch	79.84	80.54(0.70 ↑)	86.12	87.16(1.04 ↑)
	UDA	80.51	82.52(2.01 ↑)	86.31	87.50(1.19 ↑)
	FixMatch	80.59	82.96(2.37 ↑)	86.76	87.46(0.70 ↑)
	FlexMatch	80.78	82.86(2.08 ↑)	86.12	87.16(1.04 ↑)
	AdaMatch	81.93	83.78(1.85 ↑)	85.84	87.13(1.29 ↑)
	FreeMatch	82.22	83.34(1.12 ↑)	86.27	87.56(1.29 ↑)
	SimMatch	82.48	83.57(1.09 ↑)	86.98	87.88(0.90 ↑)
	SoftMatch	82.72	83.11(0.39 ↑)	86.83	87.81(0.98 ↑)
平均精度		75.88	79.18 (3.30 ↑)	84.69	85.96 (1.27 ↑)

以发现，在引入 FAGC-PSS 后，大多数情况下模型的性能表现均得到不同程度的提高。特别的，在标签样本数为 400 的 CIFAR-100 抽样数据集中，FAGC-PSS 在多个半监督学习算法的平均精度上提高约 4%。在 CIFAR-10 数据集中，针对不同数量的标签样本，FAGC-PSS 算法在仅使用 40 张有标签样本的情况下，就表现出显著的性能提升。这一结果验证了 FAGC-PSS 在处理小样本图像数据时的有效性和适应性，表明其对小样本规模并不敏感。此外，如果训练图像样本的数量较大，比如在含有标签样本数为 200 的 CIFAR-10 抽样数据集中，FAGC-PSS 对模型的性能提升并不显著。这是因为主干网络模型在对 200 张图像进行训练后已经得到了较为充分的特征分布，并且分类准确率已经达到 90.13%。在这种情况下，增加 FAGC-PSS 生成的特征数据，网络模型的性能得到约 1% 的提高并且分类精度达到 91.77%，表明

FAGC-PSS 不仅在小样本图像数据规模下表现出色，而且在较大样本规模下也有一定的性能提升。

3.2.3 方法有效性分析

为了充分验证 FAGC-PSS 在实际应用中的有效性和普适性，本文从三个关键方面展开探究。首先，本文着眼于 FAGC-PSS 与各种数据增强方法组合的有效性分析，因为各种数据增强方法及组合是提升机器学习性能的常用手段。面对小样本数据规模的图像分类任务时，经常组合使用多种数据增强方法提升模型的性能，那么 FAGC-PSS 与众多数据增强方法之间的独立性和关联性显得尤为重要。其次，本文同样关注 FAGC-PSS 对分类模型的影响。通过探究不同分类模型上 FAGC-PSS 的应用效果，可以更好地理解其在特征提取与分类过程中的作用。最后，因为生成特征数量的变化会影响模型的性能表现，所以本文还实验分析 FAGC-PSS 生成特征的数量与分类模型性能之间的相关性。

(a)FAGC-PSS 与数据增强方法组合的有效性分析

由于数据增强方法十分丰富，并且大多数增强方法都具有良好的独立性，在机器学习任务中往往组合使用多种增强方法，这导致其中一些增强方法仅适用于某些特定图像任务场景下模型性能的增强。为了验证 FAGC-PSS 的灵活性和普适性，本文选取了六种具有代表性的增强方法，包括水平翻转 (horizontal flip)、旋转 (rotate)、随机裁剪 (random crop)、颜色抖动 (color jitter)、Cutout^[56]和 Augmix^[58]。本文逐一将上述六种增强方法与 FAGC-PSS 进行对比实验，并且根据相关文献^[122]和实验经验组合多个数据增强方法与 FAGC-PSS 进行比较，以验证 FAGC-PSS 的独立性和普适性。同时为了模拟小样本数据规模下使用不同数据增强方法组合对分类模型性能的影响，选取 ViT-S 作为主干网络模型用于图像分类实验，并以 CIFAR-100 为实验数据集，其中训练样本数仅为 400 张和 2000 张，即每个类别的图像样本数分别为 4 张和 20 张。

通过观察表3.8和表3.9的分类结果可以发现，FAGC-PSS 与大部分数据增强方法组合后能够在图像分类任务上取得正向积极的性能效果，同时具有较好的普适性并与大部分数据增强方法之间相互独立。之所以 FAGC-PSS 能够与大多数数据增强方法组合使用，这得益于 FAGC-PSS 生成的特征质量和增强的方式不同。值得关注的

表 3.8 在训练样本数为 400 的 CIFAR-100 抽样数据集上使用不同数据增强方法以及 FAGC-PSS 的模型分类精度 (%)

ID	数据增强方法						分类精度	
	水平翻转	旋转	随机裁剪	颜色抖动	Cutout	Augmix	/	+FAGC-PSS
1	✓						70.32	72.24(1.92 ↑)
2		✓					70.84	72.75(1.91 ↑)
3			✓				70.21	73.39(3.18 ↑)
4				✓			71.12	72.13(1.01 ↑)
5					✓		72.14	74.93(2.79 ↑)
6						✓	71.89	73.46(1.57 ↑)
7	✓	✓					70.23	72.34(2.11 ↑)
8	✓		✓				72.65	75.36 (2.71 ↑)
9		✓			✓		70.11	72.28(2.17 ↑)
10	✓		✓			✓	72.48	73.54(1.06 ↑)
11	✓	✓	✓	✓	✓	✓	66.38	68.27(1.89 ↑)

表 3.9 在训练样本数为 2000 的 CIFAR-100 抽样数据集上使用不同数据增强方法以及 FAGC-PSS 的模型分类精度 (%)

ID	数据增强方法						分类精度	
	水平翻转	旋转	随机裁剪	颜色抖动	Cutout	Augmix	/	+FAGC-PSS
1	✓						80.24	82.17(1.93 ↑)
2		✓					80.32	82.05(1.73 ↑)
3			✓				80.21	82.36(2.15 ↑)
4				✓			80.12	82.13(2.01 ↑)
5					✓		82.14	83.93 (1.79 ↑)
6						✓	82.25	83.62(1.37 ↑)
7	✓	✓					80.41	82.73(2.32 ↑)
8	✓		✓				81.20	83.43(2.23 ↑)
9		✓			✓		80.53	82.37(1.84 ↑)
10	✓		✓			✓	81.43	83.14(1.71 ↑)
11	✓	✓	✓	✓	✓	✓	79.62	81.76(2.14 ↑)

是，所有数据增强方法的组合使用并没有让学习模型取得最佳的分类结果，正如 ID 11 所示。相比之下，在训练样本数为 400 的 CIFAR-100 数据集中，水平翻转、随机裁剪和 FAGC-PSS 的组合使用可以取得最佳的分类精度。

(b)FAGC-PSS 对分类模型的影响

除了不同数据增强方法之间的组合对算法模型的性能影响，本文还通过三组实验数据集了解 FAGC-PSS 对不同分类模型的影响。实验数据集由从 CIFAR-10 数据集中分别抽样 40 张、200 张和 1000 张图像构成。本文选取特征提取模型为 ViT-S，通过实验对比不同基于机器学习模型的性能影响，分类模型包括 K 最近邻（K-Nearest Neighbor, KNN）算法^[123]、支持向量机分类（Support Vector Classification, SVC）算法^[124]、决策树（Decision tree, DT）算法^[125]、极限树（Extra tree, ET）算法^[126]、随机森林（Random forest, RF）算法^[127]、Bagging 算法^[128]、梯度提升（Gradient boosting, GT）算法^[129]和多层感知机（Multilayer perceptron, MLP）算法^[130]。

表 3.10 基于 CIFAR-10 使用不同分类模型的分类精度 (%)

训练样本数		40		200		1000	
分类模型	增强方案	/	+FAGC-PSS	/	+FAGC-PSS	/	+FAGC-PSS
KNN		79.54	91.11(11.57 ↑)	96.29	96.75(0.46 ↑)	96.79	96.98(0.19 ↑)
SVC		89.02	91.29(2.27 ↑)	96.72	96.85(0.13 ↑)	96.82	97.35(0.53 ↑)
DT		56.76	58.51(1.75 ↑)	85.62	86.03(0.41 ↑)	87.32	88.43(1.11 ↑)
ET		51.37	66.07(14.70 ↑)	78.68	85.89(7.21 ↑)	84.53	87.31(2.78 ↑)
RF		91.36	91.23(0.13 ↓)	96.09	96.37(0.28 ↑)	96.24	96.64(0.40 ↑)
Bagging		74.52	77.87(3.35 ↑)	94.38	94.52(0.14 ↑)	94.46	94.73(0.27 ↑)
GT		71.32	76.41(5.09 ↑)	88.48	89.81(1.33 ↑)	88.86	90.21(1.35 ↑)
MLP		90.00	93.31 (3.31 ↑)	96.45	97.12 (0.67 ↑)	98.19	98.26 (0.07 ↑)

通过观察表3.10所呈现的八种分类模型的结果可见，FAGC-PSS 对机器学习模型性能的提升是显著的。不同分类模型的性能提升幅度存在差异，这种差异主要取决于模型本身的性能以及训练样本的数量。值得注意的是，结合 FAGC-PSS 和 MLP 模型的组合表现出最佳的分类精度。这主要归因于 MLP 模型利用了特征提取器 ViT-S

的输出层权重参数进行了初始化。实验结果强调了 FAGC-PSS 在增强模型性能方面的有效性以及在选取分类模型上的灵活性，并为进一步研究和应用提供了重要的参考依据。

(c)FAGC-PSS 生成特征的数量与特征分类器的相关性分析

为了评估生成特征数量对模型性能的影响，本文在仅有 400 张图像样本的 CIFAR-100 上进行训练，选取多种机器学习模型作为分类器，包括 KNN、SVC、DT、ET、RF、Bagging、GT 和 MLP。特征提取网络模型为 ViT-S，每个类别的生成特征数量分别设置为 10、100、400、1000 和 2000。根据图3.3展示的分类精度随着生成特征数量变化的趋势图可以发现，生成特征的数量与模型的分类准确率评价指标之间不存在严格的单调关系。过多的生成特征还会导致机器学习模型过度关注这些特征的分布，而忽略真实特征的分布。其中 MLP 模型在分类准确率方面取得了最佳的表现。

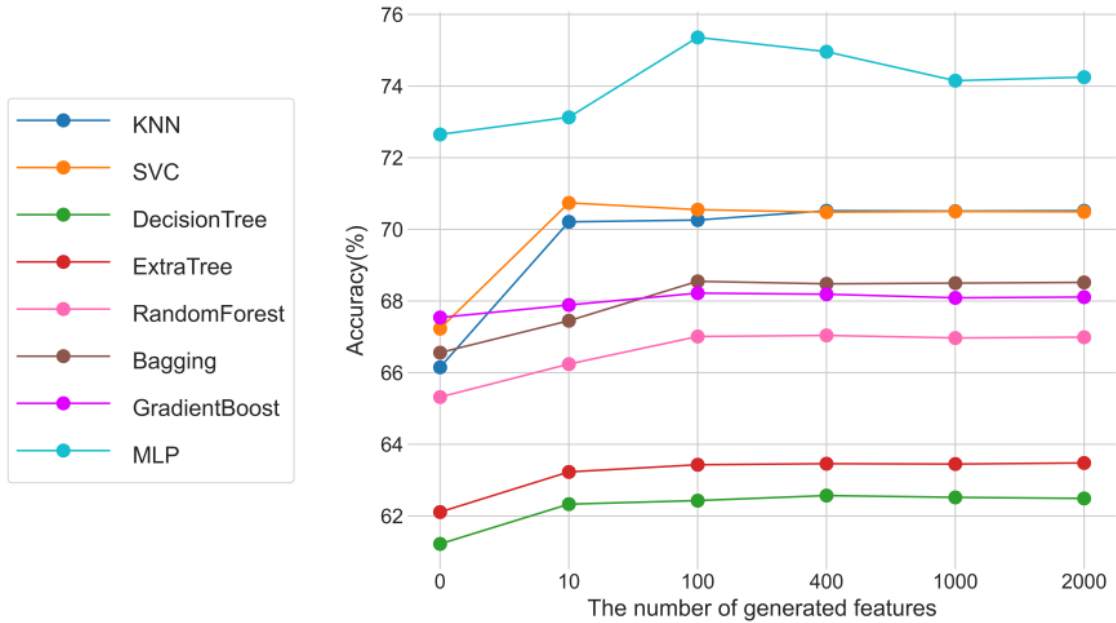


图 3.3 基于 FAGC-PSS 的不同生成特征数量对多种机器学习模型分类精度的影响

此外，本文利用不同的特征提取模型进行了补充实验，包括 ViT-S、ViT-T、WRN-28-2 和 WRN-28-8。从图3.4展示出不同特征提取器模型随生成特征数量变化的趋势图可以看出，不同的机器学习算法取得最佳性能时所需要的生成特征数量不尽相同。在仅有 400 张训练样本的 CIFAR-100 场景中，对于 MLP 模型而言，每个类别最佳生

成特征的数量为 100 至 400 的区间。因此，在选择生成特征数量时，需要考虑到模型的复杂度和数据的特点，以避免过拟合和欠拟合的问题。

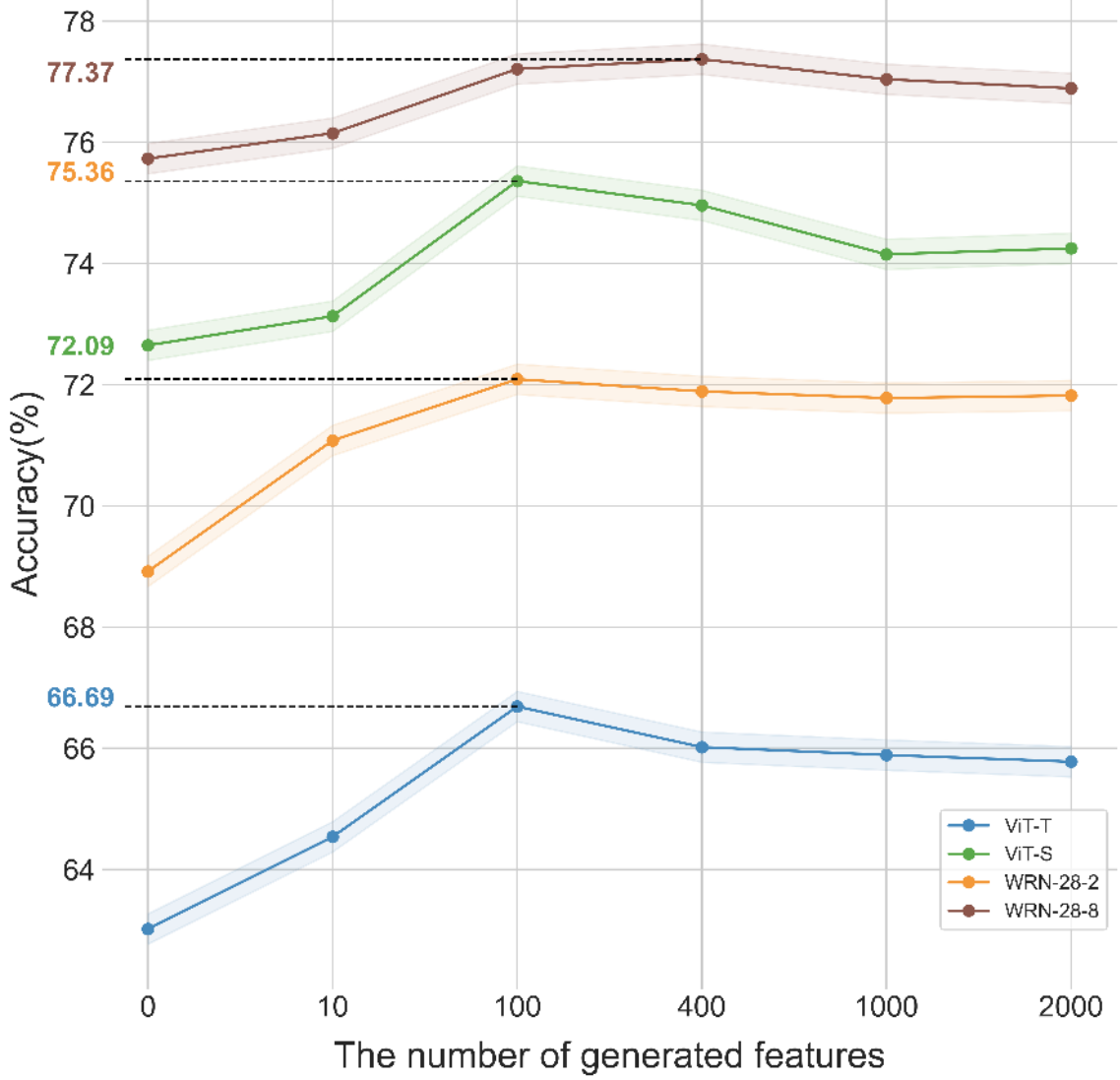


图 3.4 基于 FAGC-PSS 的不同生成特征数量对 MLP 模型分类精度的影响

3.2.4 消融实验

为了评估测地曲线生成的特征质量对模型性能的影响，本文选取仅有 40 张训练图像样本的 CIFAR-10 数据集进行了实验，同时选取 ViT-T、ViT-S、WRN-28-2 和 WRN-28-8 分别作为特征提取模型，并以 MLP 为最终的特征分类器模型。对分类损失 L_{cls} 中的参数 λ 进行消融实验，如公式(3.6)所示。值得注意的是，如果 $\lambda = 0$ ，则将当前批次的生成特征排除在模型分类训练过程之外；另一方面，如果是 $\lambda = 1$ ，则

将生成特征视为真实特征放入学习模型中进行训练。从表3.11中展示不同 λ 取值对 CIFAR-10 消融实验结果可以发现，随着影响因子 λ 的增加，学习模型的分类准确率起初会有上升，然后略有下降。下降的趋势是由于生成的特征对于损失函数的影响越来越大，导致模型倾向于学习生成特征而忽略了真实的特征。这说明 FAGC-PSS 在小样本数据集下对模型分类精度有良好的提升效果，同时需要控制生成特征放入训练过程的占比和迭代批次，这样学习模型才能最大程度利用 FAGC-PSS 生成的特征数据。

表 3.11 超参数 λ 的取值在 CIFAR-10 数据集上的消融实验结果

网络模型 超参数 λ	ViT-T	ViT-S	WRN-28-2	WRN-28-8
0	79.32	90.00	92.89	93.29
0.1	80.95	92.34	93.48	93.89
0.3	81.86	93.12	93.54	94.17
0.45	81.86	93.33	93.74	94.43
0.5	82.92	93.31	93.77	94.51
0.55	82.83	93.42	93.81	94.49
0.7	82.63	93.37	93.74	94.41
0.9	82.46	93.23	93.68	94.38
1	82.32	93.11	93.65	94.35

3.3 本章小结

本章针对小样本图像数据集，提出一种基于预形状空间中测地曲线的图像特征增强方法（FAGC-PSS）。FAGC-PSS 为机器学习模型提供额外的特征数据，实现了在小样本图像分类时学习模型性能的显著提升。本章所提出的 FAGC-PSS 方法具有以下三方面的贡献：首先 FAGC-PSS 针对小样本图像规模下数据增强困难这一问题，采用拟合特征数据来构造测地曲线并生成高质量特征数据的方式，既有效增强了图像特征数据的多样性和复杂性，还可以让使用者根据实际应用场景扩充所需的特征样本；其次，FAGC-PSS 不受到特征提取模型和特征分类模型的限制，采用 MLP 作为特征分类器就能取得不错的实验结果，同时 FAGC-PSS 的特征增强模块有着良好

的便捷性和灵活度；最后，FAGC-PSS 在小样本图像分类的任务场景中，对分类损失函数引入随机概率和影响因子来平衡生成特征和真实特征对模型的作用程度。

FAGC-PSS 相较于其他数据增强算法，能够在小样本图像规模下进一步扩充数据的特征信息，这得益于形状空间理论的测地距离和测地曲线。测地距离可以更好地衡量预形状空间中特征样本之间的相似度，而本文设计的构造测地曲线方法能高效地拟合图像的特征分布。但是 FAGC-PSS 仍有潜在的改进空间，例如，如何排除孤立特征样本点的影响；如果图像特征样本过于分散，测地曲线无法精准地反映其特征分布。这些改进的方式将进一步提高 FAGC-PSS 对机器学习模型的帮助效果。此外，本章所使用的公共图像数据集主要由自然场景的图像数据组成，与专业科学领域的材料图像存在较大差异。同时材料图像大多具有小样本、复杂的纹理信息以及多样性不足的特点。因此，本文将提出一种基于 FAGC-PSS 的材料性能预测方法，以创新性地解决小样本图像规模中材料性能预测问题。

第四章 基于 FAGC-PSS 的材料性能预测方法

在上一章中,提出基于预形状空间测地曲线的图像特征增强方法,面对小样本图像数据集规模的分类任务场景,可以有效地根据对应类别的特征样本来构造测地曲线,然后沿着曲线生成所需的特征样本实现数据的扩充。本章进一步将所提的 FAGC-PSS 方法应用于材料科学领域,通过对材料图像的处理与分析预测材料性能。材料图像与自然场景图像差异性较大,并且具有小样本、纹理信息复杂以及多样性不足等特点。本章结合材料图像的特点,通过设计 FAGC-PSS 的下游任务框架结构,实现了对小样本材料图像的性能预测任务。该方法加强了模型在预测精度和泛化能力方面的性能表现,对于不同材料种类的性能预测任务具有良好的普适性,同时有助于模型挖掘了材料图像中微观结构与性能之间的内在联系。

4.1 方法概述

为应对材料性能预测过程中常见的小样本问题,本方法引入了伪标签机制,并且重新设计了 FAGC-PSS 的下游任务框架结构。具体步骤分为以下四个关键阶段:首先,通过深度神经网络模型对材料图像进行特征提取;其次,扩展特征的维度,并将其投影至预形状空间中,以更好地捕获材料图像的几何形状信息;然后,通过拟合预形状空间下的特征数据来构造测地曲线,经过多轮迭代,直至核心参数收敛,生成新的特征数据;最后,利用机器学习模型对生成的特征进行预测,并为其赋予伪标签,将其与原始特征样本组合,一并输入到性能预测模型中进行训练。这一过程经过多次迭代,直至回归损失趋于稳定,得到可靠的性能预测模型,用于材料性能预测任务。通过本方法,能够充分挖掘材料图像中的信息,提高预测精度和泛化能力,为材料科学领域的进一步研究提供新的思路和方法。本方法的执行过程如图4.1所示。

4.1.1 材料图像特征提取

在应用 FAGC-PSS 方法之前,需要通过特征提取的方式将输入的材料图像从像素空间映射到特征空间。通常来说,基于深度学习的神经网络结构复杂度与模型性

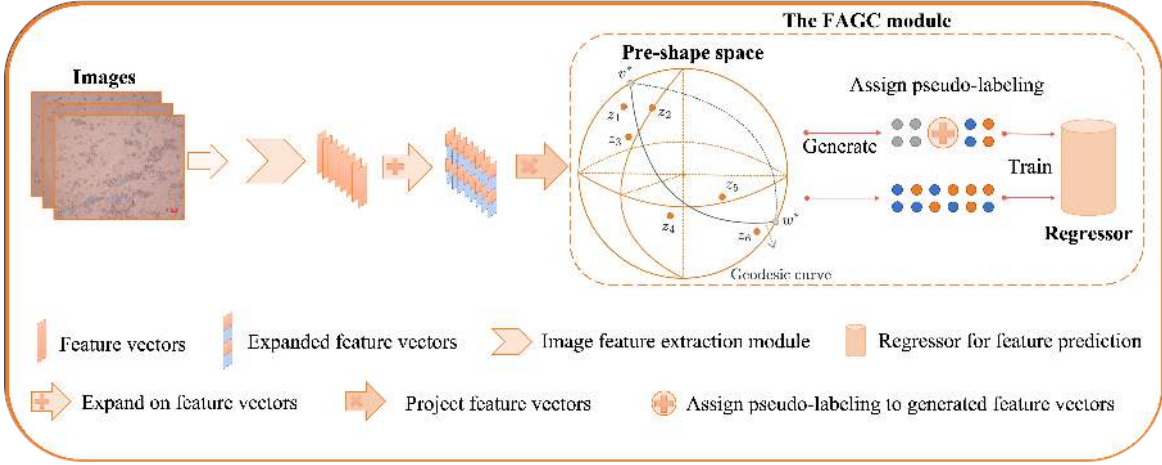


图 4.1 基于 FAGC-PSS 的材料性能预测方法流程图

能存在正向相关性，往往神经网络的层数越多模型的性能就越出色，但在处理小样本规模的材料图像时，冗余繁重的网络架构容易出现过拟合和训练不稳定的现象。虽然 ViT 能够通过自注意力机制捕捉全局特征，但在处理具有明显局部特征的材料图像时，卷积神经网络往往更具优势。EfficientNet 首次通过复合系数将网络的深度、宽度和分辨率三个方面均匀地缩放，平衡模型的参数量与计算成本，使得模型在资源受限下也能达到出色的性能表现。因此，本章选取 EfficientNet-B6^[85]作为主干网络并修改输出层结构以适应材料图像的性能预测任务，网络结构如图4.2所示。首先，MBConv 块引入扩展卷积（Expansion Convolution）技术，通过使用 1×1 卷积核扩展特征图的深度，旨在保持原始空间分辨率的同时增加特征维度。这一过程的关键在于通过深度扩展卷积提高了模型对输入数据的表征能力，使其能够更加全面捕获图像中的复杂结构和抽象特征。接着，MBConv 块还采用了深度可分离卷积（Depthwise Separable Convolution），它包含两个关键阶段。第一阶段是深度卷积，对每个通道都单独应用空间滤波器。这种方法的优势在于显著降低计算量，因为它避免了对所有通道的合并特征进行处理，而是独立地处理每个通道。在第二个阶段是 MBConv 块使用 1×1 卷积核将独立的通道信息重新组合，这不仅恢复了特征图的深度，还将不同通道的信息融合在一起。为了增加网络的非线性特性，激活函数采用 Swish^[131]，它是一种非线性且平滑的激活函数，能够帮助网络学习更复杂的特征表示：

$$\text{swish}(x) = x\sigma(\beta x), \quad (4.1)$$

其中 $\sigma(\cdot) \in (0, 1)$ 为 **Logistic** 函数，可以看作一种软性的门控机制。 β 为可学习的参数或者一个手工定义的超参数。当 $\sigma(\beta x)$ 接近于 1 时，激活函数的输出近似于 x 本身，反之，当 $\sigma(\beta x)$ 接近于 0 时，激活函数的输出近似于 0。本章通过修改模型全连接层的输出维度，设计一个基于 **EfficientNet-B6** 的回归预测模型。同时，改进的 **EfficientNet-B6** 模型已在 **ImageNet** 数据集上进行预训练，并且网络模型训练过程采用均方误差（**Mean square error, MSE**）作为损失函数，直至损失值趋于稳定时终止迭代。此时，使用训练好的 **EfficientNet-B6** 网络的输出数据作为输入材料图像的特征样本，以提供后续特征数据处理的使用。

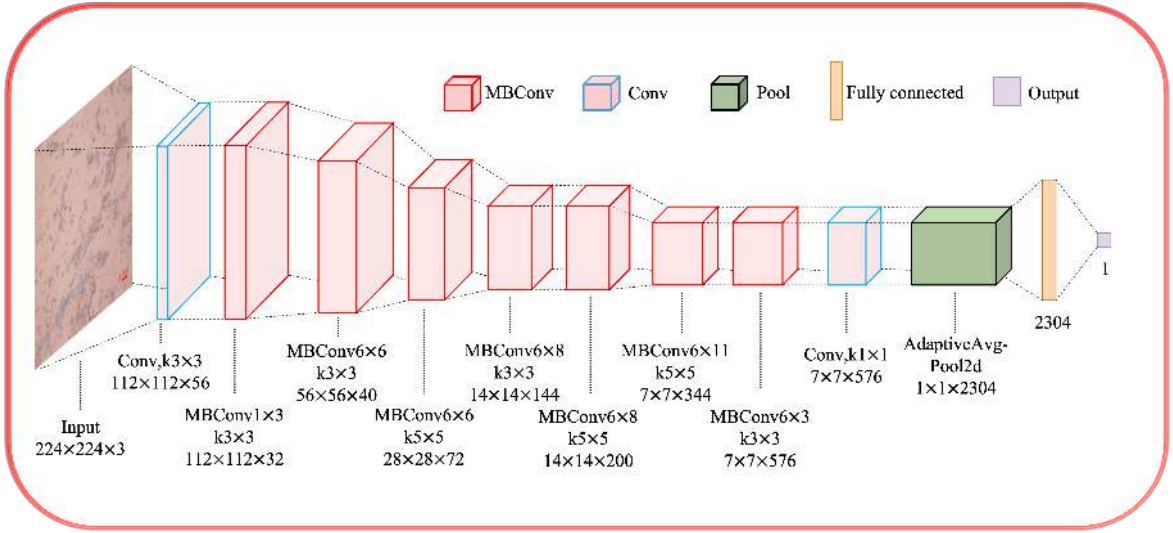


图 4.2 基于 **EfficientNet-B6** 的材料性能预测的网络架构（引用文献^[85]）

4.1.2 标注模型

在进行材料图像特征标注之前，根据第三章的 **FAGC-PSS** 方法对材料图像进行特征增强，首先将材料图像投影至预形状空间中，如图4.3所示，其中 A 和 B 表示特征空间中的两幅材料图像。蓝色虚线描述了材料图像 A 的测地曲线，向量 z^i, z^j, z^k 表示材料图像特征 A 的三种形状构造。接着，对材料图像特征进行测地曲线的构造。相较于自然场景的多种类别图像，特定材料图像属于单一类别图像，仅需构造一条测地曲线，其中求解测地曲线两个端点 v^* 和 w^* 公式如下：

$$v^* = \operatorname{argmax}_{z^i} \sum_{j=1}^M d(z^i, z^j), i \in \{1, \dots, M\} \quad (4.2)$$

和

$$w^* = \underset{\hat{w}^k}{\operatorname{argmin}} \sum_{j=1}^M \left(d(z^j, \Gamma_{(v^*, \hat{w}^i)}(\cdot)) \right)^2, i \in \{1, \dots, S\} \quad (4.3)$$

其中 $Z = \{z^1, z^2, \dots, z^M\}$ 表示为材料图像特征向量集, M 表示材料图像特征的数量, S 表示临时特征的数量, $\Gamma_{(v^*, w^*)}$ 是最优的测地曲线并且良好地拟合特征向量集 Z 的分布。随后, 沿着测地曲线 $\Gamma_{(v^*, w^*)}(\cdot)$ 构建了一系列生成的特征。同时, 为了丰富特征的多样性, 采用均匀分布 $s \sim U(0, d(v^*, w^*))$ 对一组角度数值 $\{s^1, s^2, \dots, s^K\}$ 进行抽样, 其中 K 既表示角度的数量, 也代表生成特征的数量。然后, 这些角度数值被代入测地曲线函数 $\Gamma_{(v^*, w^*)}(\cdot)$ 中, 得到在这条测地曲线上的一系列特征数据, 表示为 $G = \{g^1, g^2, \dots, g^K\}$ 。

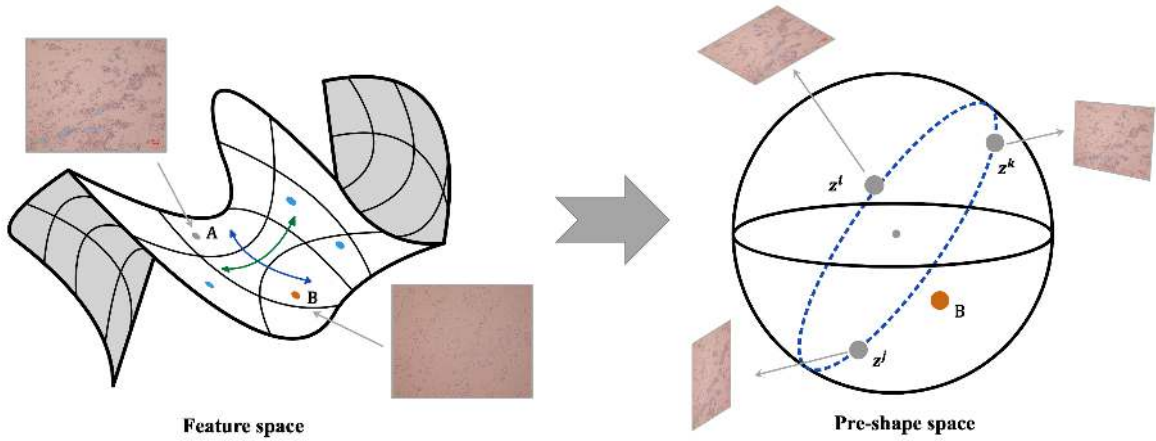


图 4.3 材料图像从特征空间到预形状空间的投影过程示意图

为了充分利用生成特征中的信息, 本章设计了一个伪标签机制用于标注生成特征数据。通过基于机器学习的标注模型 R_a 来预测测地曲线生成特征集 G 中的特征对应的性能, 表示为 $R_a(G) = p_i$, 其中 $\{p_i\}$ 是相应的预测性能值, 这些预测值被标注为生成的特征集 G 的伪标签。通过扩充原始特征集, 可以获得更多样化和复杂化的特征数据集。这个扩充的数据集由两部分组成: 一部分是原始图像特征集 Z , 另一部分是带有伪标签的生成特征集 G 。将这两部分组合在一起得到的扩充数据集表示为 $Z' = \{z^1, \dots, z^M, g^1, \dots, g^K\}$, 其中 M 和 K 分别表示原始特征集和生成特征集的数量。

4.1.3 预测模型

为了充分学习扩充后的特征集 Z' ，采用未初始化的机器学习模型作为材料性能预测模型 R_b ，预测模型 R_b 的训练目标是 minimized 均方误差，直至损失值趋近于稳定时终止迭代。通过对不同材料性能的特征集 Z' 进行预测训练，获得到最佳预测模型 R_b ，它能够更准确地预测材料图像的相关性能（例如导电率、硬度等特性）。图4.4全面说明了从生成的特征集到训练最终性能预测模型的步骤顺序。

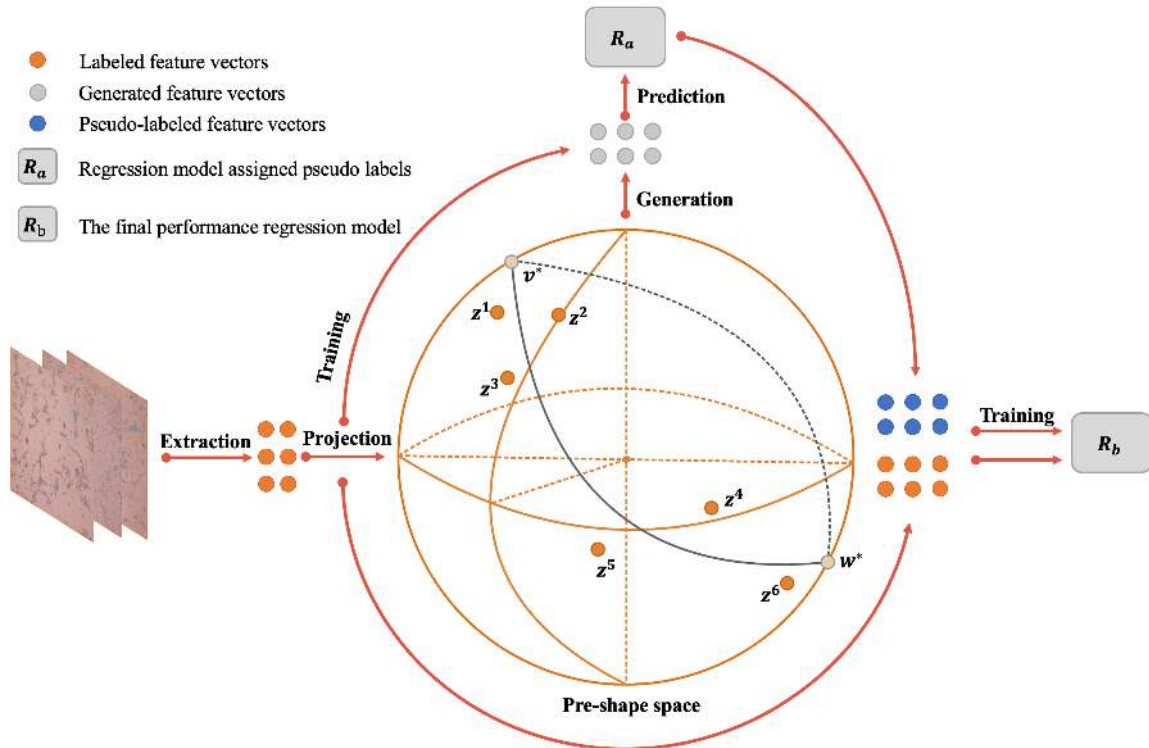
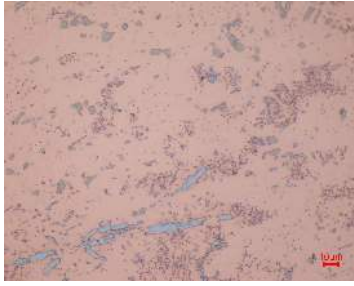


图 4.4 基于 FAGC-PSS 的伪标签标注工作流程图

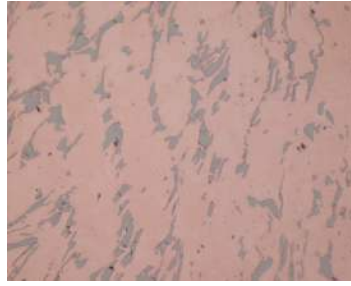
4.2 实验分析

4.2.1 材料图像数据集

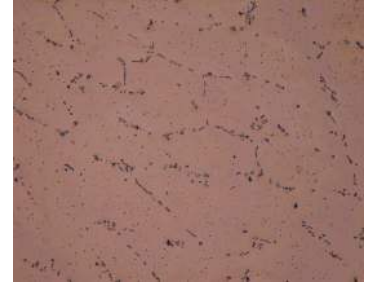
为了充分验证本方法对小样本材料图像性能预测任务的增强效果，本章选取铜铬锆（CuCrZr）和经过王水腐蚀处理后的钴铬（CoCr）合金图像进行了测试以验证所提出的性能预测方法。CuCrZr 材料图像数据集包含 18 幅图像样本，其中性能值为导电率和硬度，CoCr 合金图数据包含 21 幅图像样本，性能值为硬度，部分 CuCrZr 图像和 CoCr 合金图像实例分别如图4.5和图4.6所示。



(a) 导电率 64.68; 硬度 229.86

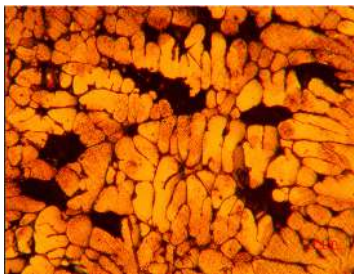


(b) 导电率 67.35; 硬度 211.51

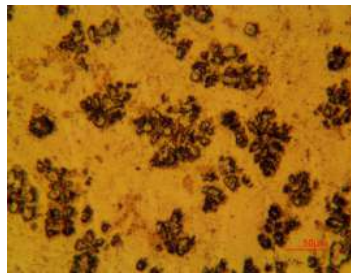


(c) 导电率 87.15; 硬度 174.16

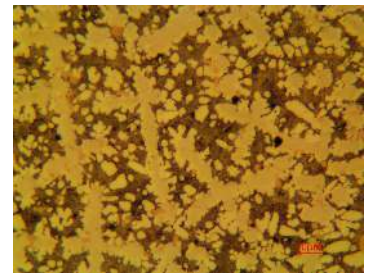
图 4.5 CuCrZr 部分图像实例图



(a) 硬度 374



(b) 硬度 517



(c) 硬度 777

图 4.6 经过王水腐蚀处理的 CoCr 合金部分图像实例图

4.2.2 特征提取模型的对比实验

从图像样本的数量来说，CuCrZr 材料图像和 CoCr 合金图像均属于小样本图像数据规模，在传统的训练模式中会直接使用神经网络模型进行回归预测。然而，当面对小样本数据规模时，往往难以获得令人满意的预测效果。本章采用了直接生成新的特征数据的策略，可以克服特征提取模型和性能预测模型架构的限制，从而能够有效应对小样本数据的挑战。为了确定特征提取能力最佳的 EfficientNet 模型结构，本章测试了 EfficientNetB0 到 EfficientNet-B7 在 CuCrZr 材料上性能预测的表现。在实验设置方面，网络模型使用 ImageNet 数据进行预训练，以初始化权重参数。选择均方误差作为模型训练的损失函数，优化器采用自适应矩估计（Adaptive Moment Estimation, Adam）算法^[132]，初始学习率为 0.0001，总训练批次为 200。同时将 R^2 分数、MAE 和 RMSE 作为模型的性能评价指标，并采用六折交叉验证方法对 CuCrZr 数据集进行划分。在每一折的迭代中，选择 3 张图像组成本次迭代的测试集，剩余的图像样本用于训练性能预测模型的网络权重。实验结果如表 4.1 所示。从 EfficientNet-B3 到 EfficientNet-B7 观察到的 R^2 分数、MAE 和 RMSE 差异较小，考虑到性能指标结果，

选择 EfficientNet-B6 模型作为最佳的特征提取模型。

表 4.1 不同 EfficientNet 架构预测 CuCrZr 材料导电率和硬度的预测结果

网络模型	导电率 (Conductivity)			硬度 (Hardness)		
	R ²	MAE	RMSE	R ²	MAE	RMSE
EfficientNet-B0	0.326	9.674	10.348	0.389	32.532	35.428
EfficientNet-B1	0.334	9.452	10.287	0.427	31.975	34.317
EfficientNet-B2	0.487	8.375	9.031	0.496	29.863	32.191
EfficientNet-B3	0.559	8.164	8.376	0.559	28.613	30.079
EfficientNet-B4	0.556	7.963	8.395	0.553	27.462	30.291
EfficientNet-B5	0.570	7.682	8.263	0.557	27.521	30.172
EfficientNet-B6	0.606	7.415	7.918	0.619	26.883	27.981
EfficientNet-B7	0.591	6.872	8.061	0.598	28.947	28.763

4.2.3 CuCrZr 图像的材料性能预测

为了全面评估测地曲线生成的特征对性能预测模型的增强效果，本章选取 EfficientNet-B6 网络模型用于 CuCrZr 材料图像的特征提取及材料性能预测，使用 FAGC-PSS 模块结合不同预测模型与基准模型进行比较，预测模型包括线性回归 (Linear Regressor, LR) 算法、K 最近邻 (K-Nearest Neighbor, KNN) 算法^[123]、自适应提升 (Adaptive boosting, AdaBoost) 算法^[133]、决策树 (DecisionTree, DT) 算法^[125]、极限树 (ExtraTree, ET) 算法^[126]和 Bagging 算法^[128]。图4.7展示了不同机器学习性能预测模型对 CuCrZr 导电率的性能预测结果，其中红色虚线对应的是真实导电率与预测值之间的完美相关性，蓝色点由预测值和真实值构成的二维坐标。如果蓝色点越接近于红色虚线，则说明预测的性能值越接近真实的材料性能值。实验结果明确展示 FAGC-PSS 模块在提升模型预测性能方面的显著效果。这归因于 FAGC-PSS 模块通过拟合预形状空间中的特征数据来构建测地曲线。在多轮迭代优化过程中，测地曲线逐渐逼近材料图像的特征分布，生成了新的特征数据。这些新特征数据不仅包含了原始特征的信息，还结合了材料图像的几何形状信息，从而显著增强了特征的表示能力。测地曲线的构建过程有效地捕捉了材料微观结构的复杂特性，使得模

型能够在预测中利用这些信息，从而提高了预测的准确性。其次，生成的特征数据被赋予伪标签后，与原始特征样本一起输入到性能预测模型中进行训练。多次迭代训练有效地避免了小样本数据集带来的过拟合问题，显著提升了模型的泛化能力。此外，伪标签机制不仅提供了额外的监督信息，还增强了模型在小样本情况下的鲁棒性，使得模型能够在不同的数据分布和任务需求下依然表现出色。特别是决策树预测模型，其 R^2 分数和 RMSE 分别达到 0.978 和 1.882。详细的实验结果数据如表4.2。

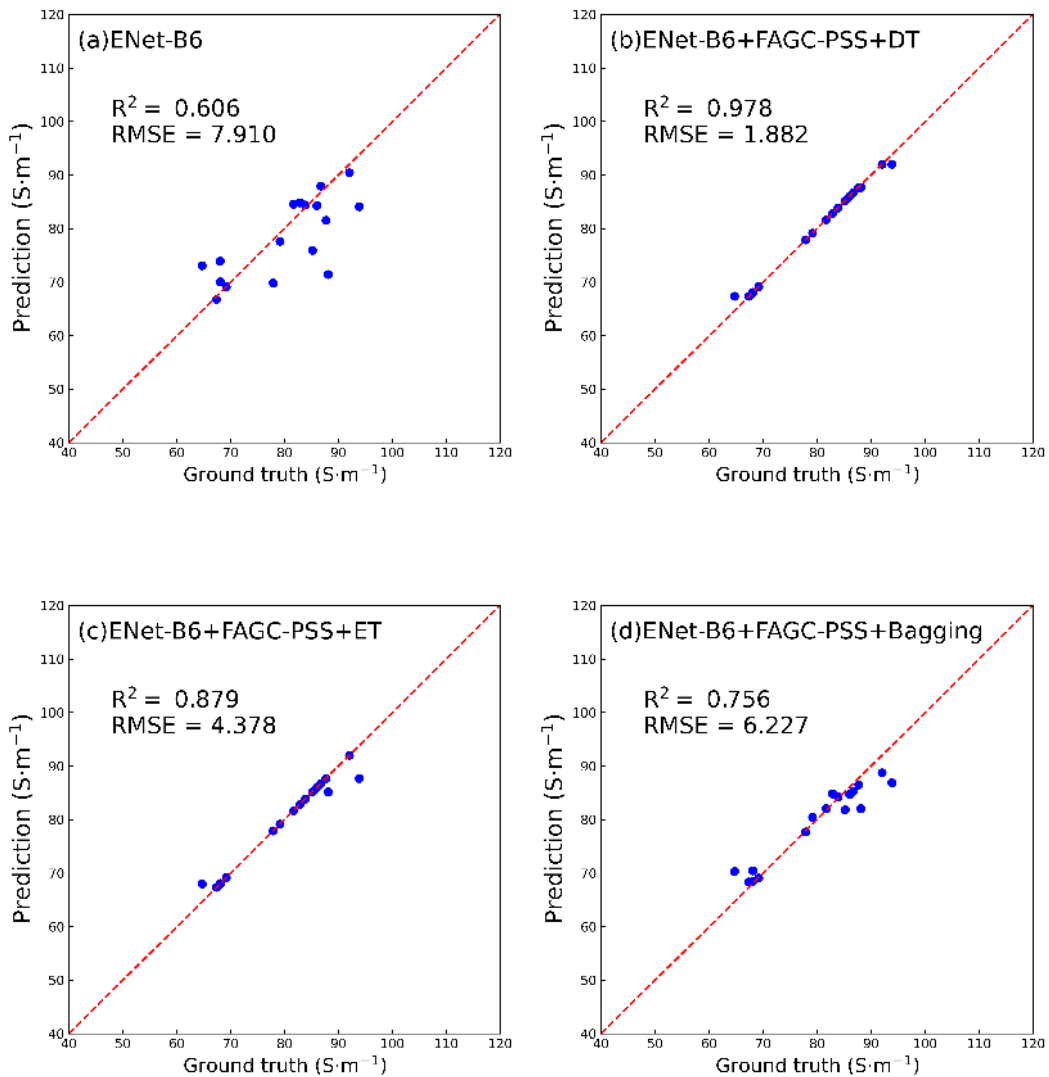


图 4.7 基于 FAGC-PSS 和 EfficientNet-B6(ENet-B6) 特征提取器使用不同性能预测模型对 CuCrZr 导电率的预测值和实际性能的比较

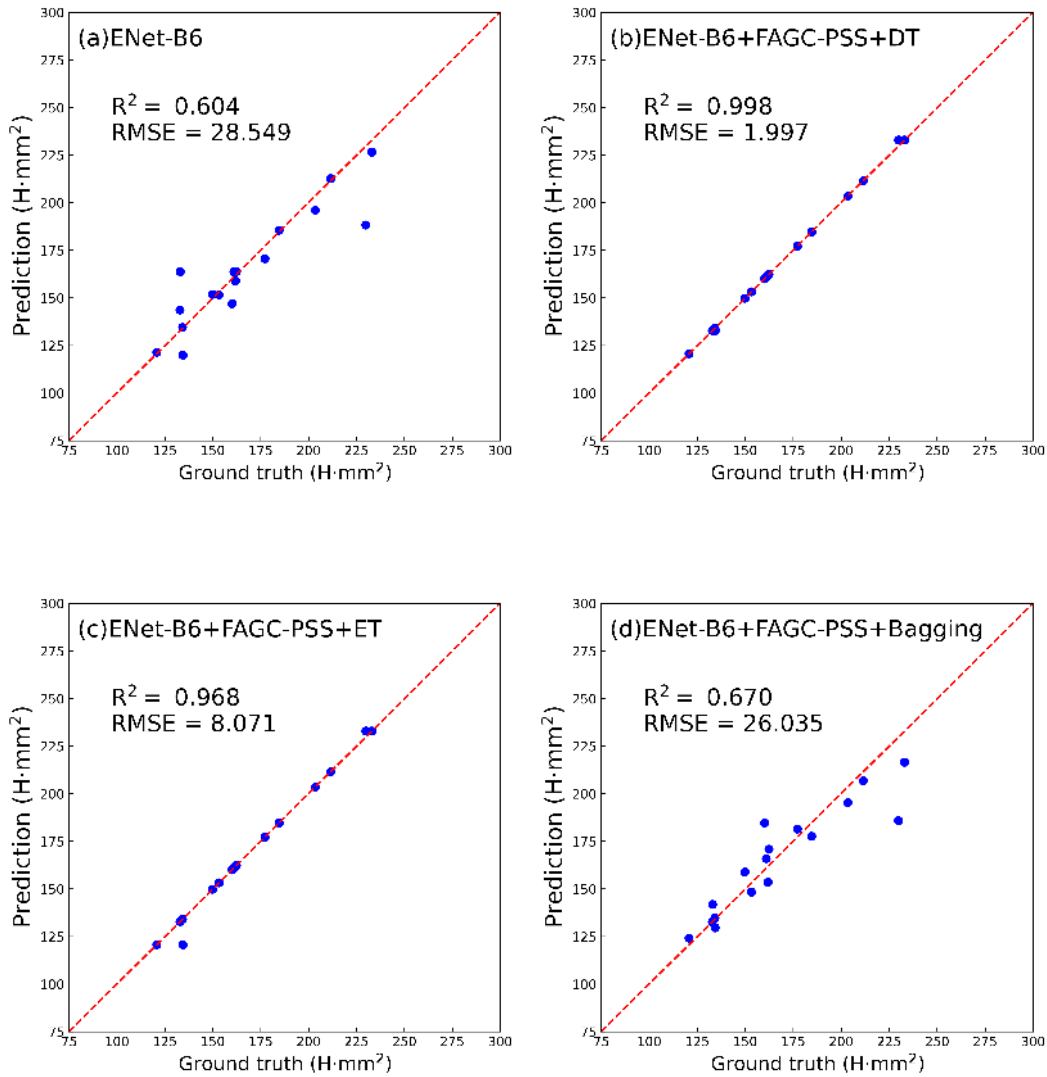


图 4.8 基于 FAGC-PSS 和 EfficientNet-B6(ENet-B6) 特征提取器使用不同性能预测模型对 CuCrZr 硬度的预测值和实际性能的比较

此外,为了验证 FAGC-PSS 模块在不同性能预测任务中的通用性,本章对 CuCrZr 的硬度性能进行性能预测实验。同样选取 EfficientNet-B6 用于 CuCrZr 材料图像的特征提取及材料性能预测,使用 FAGC-PSS 模块结合不同预测模型与基准模型进行比较,相关实验设置与上文所述保持一致。图4.8展示了预测结果的对比分析,对于 CuCrZr 的硬度性能预测, EfficientNet-B6 性能预测模型的 R^2 分数为 0.604、MAE 为 26.883、MSE 为 815.018 以及 RMSE 为 28.549。模型预测精度上的提高归因于 FAGC-PSS 模块拟合特征分布的有效性,即使是在特征数据集有限的情况下也是如此。其次,通过伪标签机制实现生成特征数据的有效利用和数据集的扩展,有助于丰富数

表 4.2 基于 FAGC-PSS 模块的性能预测模型对 CuCrZr 导电率性能的评价指标

性能预测模型	R^2	MAE	MSE	RMSE
LR	0.548	8.136	71.843	8.476
KNN	0.605	7.423	62.647	7.915
AdaBoost	0.663	6.957	53.568	7.319
ExtraTree	0.763	4.910	37.601	6.132
DecisionTree	0.609	7.192	62.142	7.883
Bagging	0.689	6.829	49.365	7.026
EfficientNet-B6	0.606	7.415	62.695	7.918
EfficientNet-B6 + FAGC-PSS + LR	0.510	8.740	77.881	8.825
EfficientNet-B6 + FAGC-PSS + KNN	0.660	7.021	54.052	7.352
EfficientNet-B6 + FAGC-PSS + AdaBoost	0.710	6.505	46.036	6.785
EfficientNet-B6 + FAGC-PSS + ExtraTree	0.879	4.134	19.167	4.378
EfficientNet-B6 + FAGC-PSS + DecisionTree	0.978	1.632	3.542	1.882
EfficientNet-B6 + FAGC-PSS + Bagging	0.756	6.202	38.776	6.227

表 4.3 基于 FAGC-PSS 模块的性能预测模型对 CuCrZr 硬度性能的评价指标

性能预测模型	R^2	MAE	MSE	RMSE
LR	0.483	28.632	1060.023	32.558
KNN	0.539	26.891	947.716	30.785
AdaBoost	0.504	23.207	1018.759	31.918
ExtraTree	0.550	18.118	925.133	30.416
DecisionTree	0.638	19.675	743.762	27.272
Bagging	0.430	24.971	1171.419	34.226
EfficientNet-B6	0.604	26.883	815.045	28.549
EfficientNet-B6 + FAGC-PSS + LR	0.477	28.779	1074.528	32.780
EfficientNet-B6 + FAGC-PSS + KNN	0.557	25.970	911.617	30.193
EfficientNet-B6 + FAGC-PSS + AdaBoost	0.915	11.057	174.742	13.219
EfficientNet-B6 + FAGC-PSS + ExtraTree	0.968	5.632	65.141	8.071
EfficientNet-B6 + FAGC-PSS + DecisionTree	0.998	1.598	3.988	1.997
EfficientNet-B6 + FAGC-PSS + Bagging	0.670	19.188	677.821	26.035

据集的多样性和复杂性。特别是对于决策树模型和极限树模型， R^2 分数分别达到了 0.998 和 0.968。表4.3详细呈现不同性能预测模型在硬度性能预测下的评价指标，实

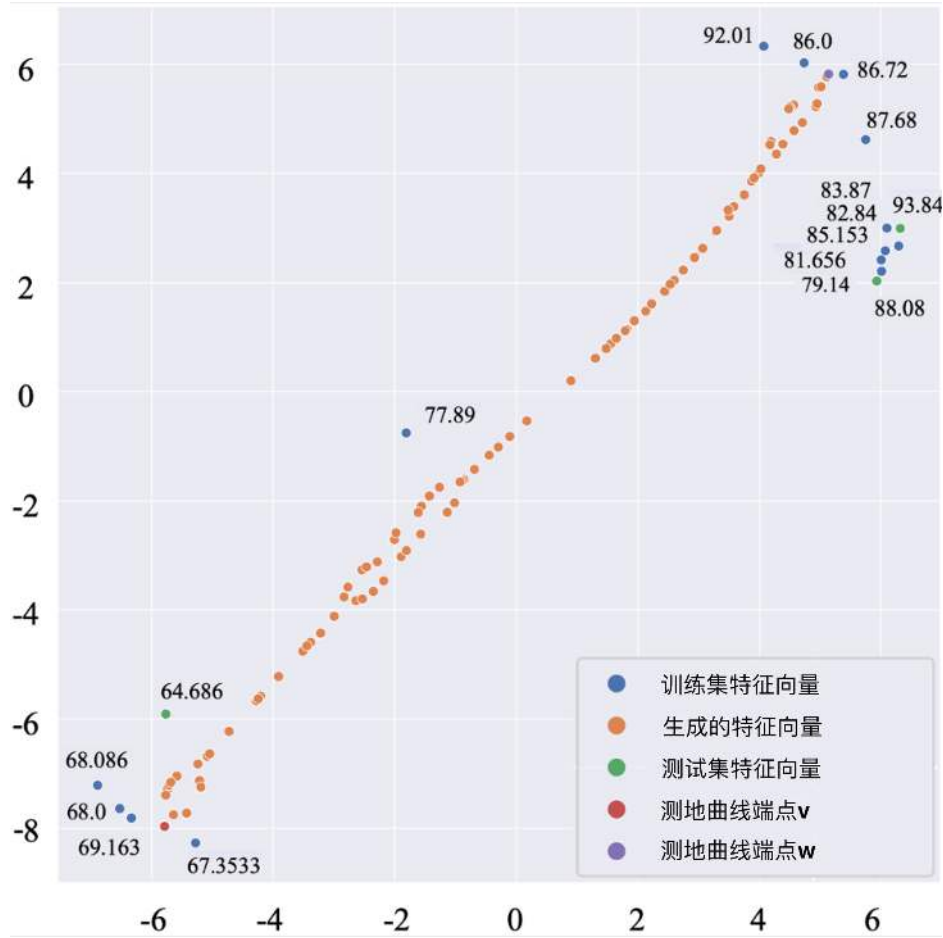


图 4.9 CuCrZr 特征和生成特征的 t-SNE 可视化结果图

验结果进一步验证 FAGC-PSS 模块在预测 CuCrZr 性能方面的有效性和通用性。

为了研究生成的特征与 CuCrZr 材料特征性能之间的相关性，本章还使用了 t 分布随机领域嵌入 (t-SNE) 方法^[134]对特征数据进行降维可视化处理。通过将特征数据映射到二维空间上，可以观察生成特征和 CuCrZr 材料特征之间的分布关系。图4.9展示了生成特征与 CuCrZr 材料特征之间的分布概况，其中蓝点 (*train_features*) 和绿点 (*test_features*) 分别代表 CuCrZr 材料图像的训练集和测试集特征，橙色点 (*generate_features*) 代表生成的特征，红色点 (*params_v*) 和紫色点 (*params_w*) 代表测地曲线的两个端点。从图中可以看出，测地曲线出色地拟合 CuCrZr 的特征分布区域，进一步验证测地曲线拟合特征的有效性。此外，本章还在图中标注了 CuCrZr 材料的导电率性能，可以发现 CuCrZr 材料的导电率性能从左下角到右上角呈上升趋势，这与生成特征的分布区域较为接近，表明赋予生成特征伪标签的合理性和可靠性。

4.2.4 CoCr 图像的材料性能预测

由于材料图像的特点各不相同，导致某些方法仅适用于单一种类的材料性能预测工作，为了充分验证本方法的通用性，本章选取经过王水腐蚀处理过的 CoCr 合金图像进行硬度性能预测实验。CoCr 合金图像数据集共包含 21 张图像且属于小样本图像数据规模，采用七折交叉验证对数据集进行划分，在每一折的迭代中，选择 3 张图像组成本次迭代的测试集，剩余图像被用作训练集，用以训练性能预测模型。同样选取 EfficientNet-B6 网络模型用于 CoCr 合金图像的特征提取及材料性能预测，并使用 FAGC-PSS 模块结合不同预测模型，实验设置同第4.2.3小节保持一致。从表4.4展示的模型预测结果可以发现，通过引入 FAGC-PSS 模块，大多数机器学习模型的预测精度得到明显提高，特别是决策树预测模型，其 R^2 分数和 RMSE 分别达到 0.932 和 27.725。此外，实验结果也证实 FAGC-PSS 模块不局限于单一种类的材料性能预测，并且对小样本图像的性能预测任务有着实际应用价值。

表 4.4 基于 FAGC-PSS 模块的性能预测模型对 CoCr 合金硬度性能的评价指标

性能预测模型	R^2	MAE	RMSE
LR	0.296	82.012	89.065
KNN	0.352	84.933	85.474
AdaBoost	0.417	80.711	81.102
ExtraTree	0.461	67.402	78.025
DecisionTree	0.478	75.827	76.726
Bagging	0.453	75.933	78.536
EfficientNet-B6	0.486	75.316	76.077
EfficientNet-B6 + FAGC-PSS + LR	0.533	62.947	72.524
EfficientNet-B6 + FAGC-PSS + KNN	0.562	63.334	70.239
EfficientNet-B6 + FAGC-PSS + AdaBoost	0.691	55.175	58.966
EfficientNet-B6 + FAGC-PSS + ExtraTree	0.886	35.333	35.842
EfficientNet-B6 + FAGC-PSS + DecisionTree	0.932	25.333	27.725
EfficientNet-B6 + FAGC-PSS + Bagging	0.758	45.757	52.192

4.2.5 生成特征数量与预测模型的相关性分析

为了评估生成特征数量对材料性能预测模型的影响，本章采用多种机器学习模型来预测 CuCrZr 的导电率和硬度性能。特征提取模型采用 EfficientNet-B6，生成特征的数量设置为 10、20、40、100、400 和 1000。从图4.10和图4.11展示了性能预测模型在不同特征数量下的差异，增加生成特征的数量可以显著提高模型的预测精度。除此之外，生成特征的数量与性能预测模型的 R^2 分数之间并非始终保持严格的单调关系。过多的生成特征会导致性能预测模型过度关注其特征分布，而忽略真实特征的分布区域。因此，选取合理的生成特征数量才能实现性能预测模型的最佳性能表现。在 CuCrZr 材料图像数据集中，经实验发现，生成约 100 个特征较为理想，即数量上约为原始 CuCrZr 图像数据集大小的五倍。在这样的条件下，大多数基于 FAGC-PSS 的性能预测模型都能表现出最佳的预测性能。

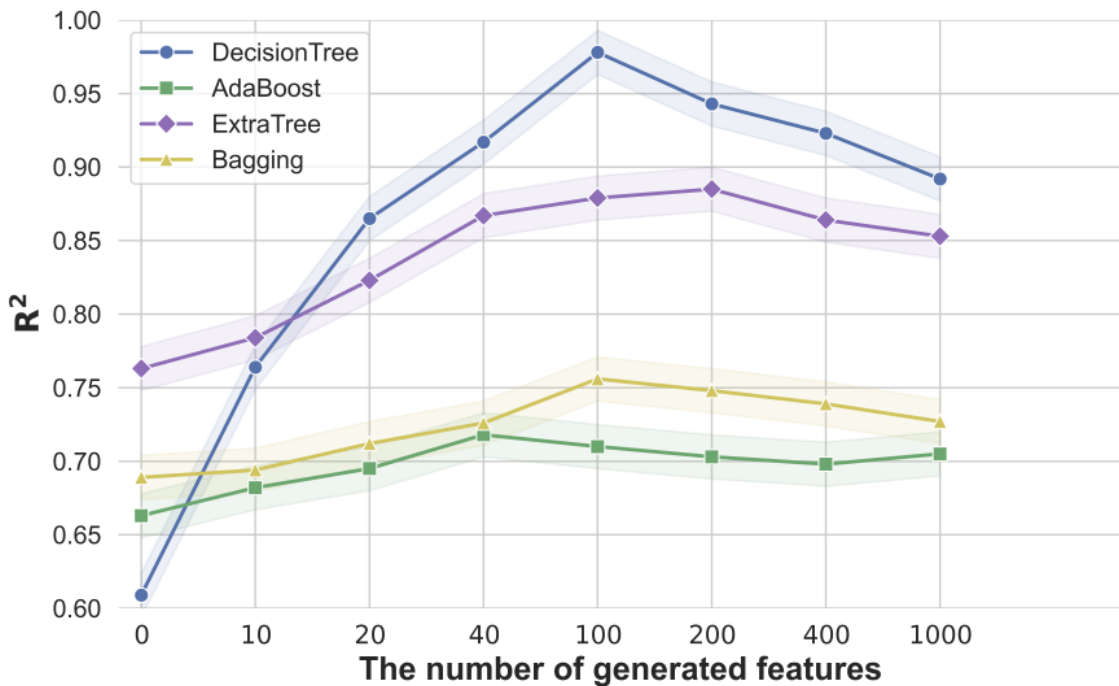


图 4.10 基于 FAGC-PSS 生成不同特征数量时预测 CuCrZr 导电率的 R^2 分数趋势图

4.2.6 标注模型与预测模型的关联性实验

为了探究本方法中标注模型与性能预测模型之间的相关性，本章对伪标签的标注模型与预测模型之间进行对比实验，选取决策树、极限树、AdaBoost 和 Bagging

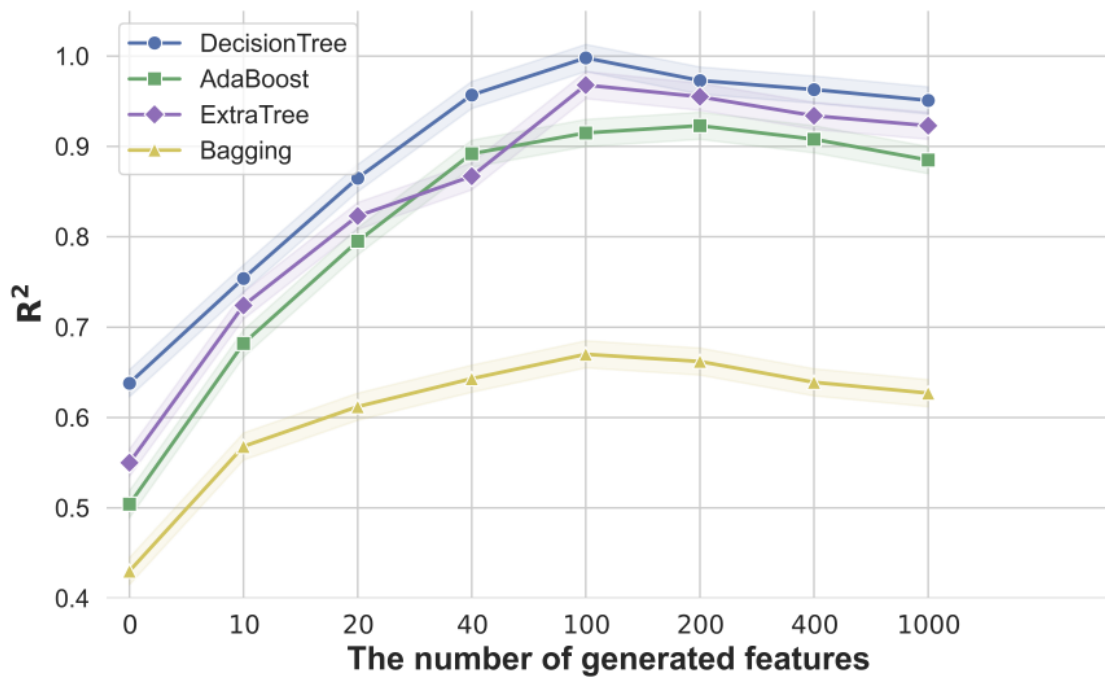


图 4.11 基于 FAGC-PSS 生成不同特征数量时预测 CuCrZr 硬度的 R^2 分数趋势图

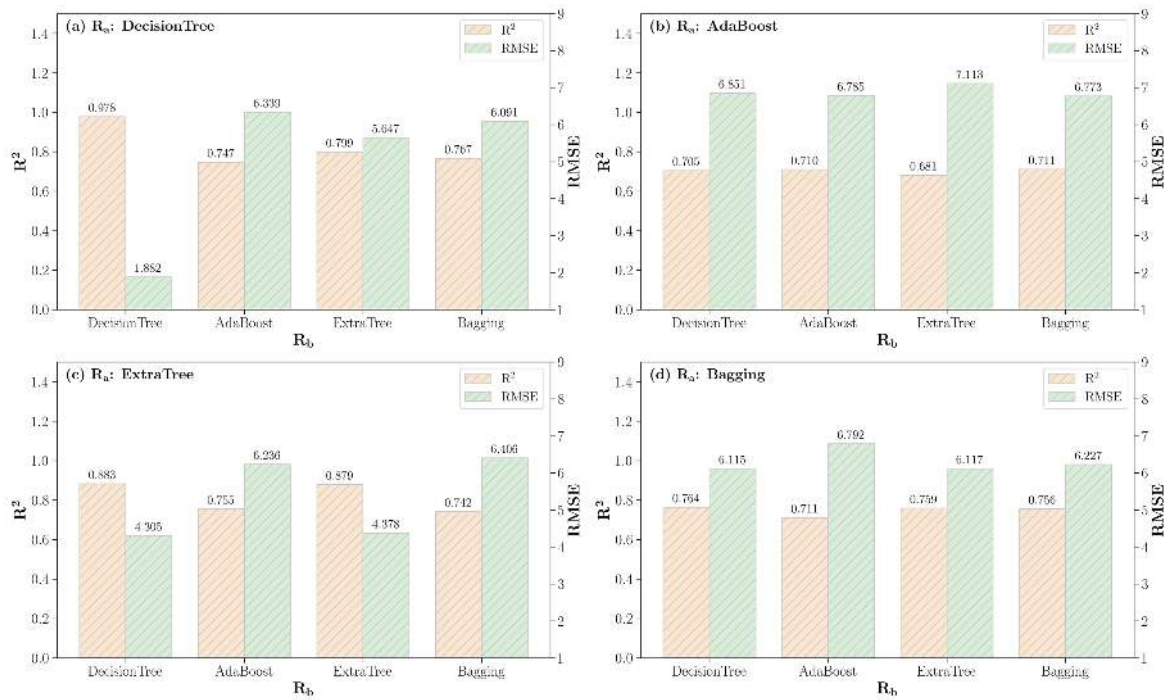


图 4.12 性能预测模型和标注模型预测 CuCrZr 导电率的对比实验结果

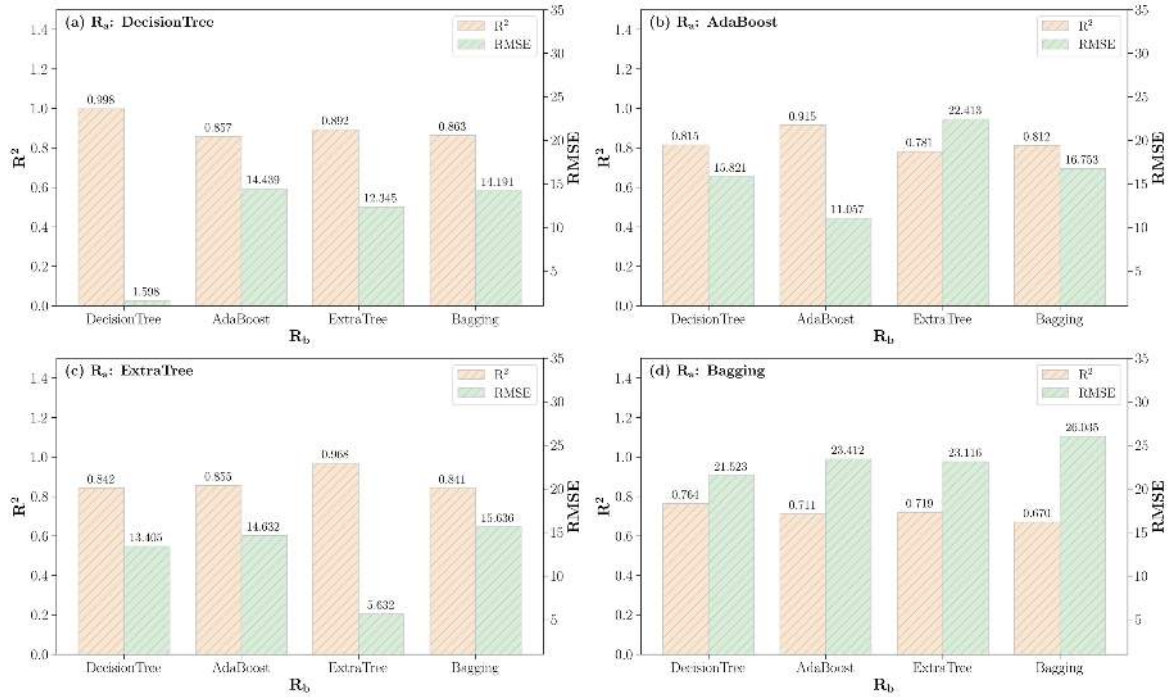


图 4.13 性能预测模型和标注模型预测 CuCrZr 硬度的对比实验结果

分别作为伪标签的标注模型 R_a 和性能预测模型 R_b 。同样选择 EfficientNet-B6 作为特征提取模型，用于预测 CuCrZr 材料图像的导电率和硬度性能。从图4.12和图4.13展示的结果可以表明，伪标签的质量对性能预测模型的性能有着显著影响，并且决策树模型在这方面表现出了最佳性能。当使用该模型向生成的特征数据标注伪标签之后，性能预测模型通过学习由决策树模型生成的伪标签数据，都实现了显著的性能提升，提升的效果超过了通过自身标注伪标签时所取得的性能。因此，当使用性能一般的性能预测模型进行伪标签标注时，从扩充的数据集中获得的性能提升相对有限。此外，大部分性能预测模型在性能指标上均超越了 EfficientNet-B6，进一步验证 FAGC-PSS 模块对于提升性能预测模型的有效性和通用性。

4.3 本章小结

本章针对小样本图像的材料性能预测任务中存在的可用信息不足、数据多样性不足以及算法缺乏普适性三个问题，提出一种基于 FAGC-PSS 的材料性能预测方法。该方法通过 FAGC-PSS 模块进行特征增强，提高了材料性能预测结果，并且设计对应的伪标签机制为生成特征赋予性能值，然后对扩充后的特征数据集进行训练，最终得到具有良好的性能预测模型。本章所提出的方法具有以下贡献：首先，本章

设计了基于 **FAGC-PSS** 模块的材料性能预测方法，使得模型可以在图像小样本进行训练并取得良好的材料性能预测；其次，设计了针对测地曲线生成特征的伪标签机制。该方法不依赖于特定的网络结构，同时在不改变 **FAGC-PSS** 模块的情况下，能够适用于实现不同种类的材料性能预测任务。

第五章 总结与展望

5.1 结论

小样本图像数据集通常由样本数相对较少的图像组成，其主要的特点包括数据稀缺性、数据分布不均、标注成本高以及样本噪声多等，这些特点就导致小样本数据规模下图像处理任务存在诸多问题。这些问题可以被概括为数据信息少、数据多样性不足和算法缺乏普适性三个难点。具体表现为：大多数实际应用场景中图像样本难以被批量获取，导致算法模型无法在有限数据规模下取得良好的性能表现；样本数量有限时，训练的生成模型往往会出现过拟合问题，导致生成的图像样本缺乏多样性和真实性；现今的数据增强方法为了更出色地提高模型的性能，通常根据特定的图像处理任务和图像样本本身设计改进算法，导致算法缺乏普适性。因此，为了充分挖掘小样本图像的内在信息和增加特征数据的多样性和复杂性，本文针对小样本数据规模的图像处理任务，结合机器学习、深度学习和形状空间理论，提出分别适用于图像识别和材料性能预测的特征增强方法，即：

1. 基于预形状空间测地曲线的图像特征增强方法。首先，对小样本图像进行深度特征的提取，包括对图像进行预处理和通过神经网络模型进行特征提取；然后，根据形状空间理论提升图像特征的维度并将其投影至预形状空间；接着，根据类别对特征数据进行拟合并构造对应的测地曲线，直至测地曲线权重不在变化时终止迭代，最后沿着训练好的测地曲线生成该类别下新的特征。该方法在小样本图像环境下实现了特征数据的增强和扩充，增加了特征数据的多样性，使得模型更好地学习到数据的真实分布，从而提高模型的泛化能力和分类精度，同时提供了一种解决小样本图像分类问题的新途径。
2. 基于 FAGC-PSS 的材料性能预测方法。该方法设计一个基于 FAGC-PSS 模块的材料性能预测框架，以克服材料图像的性能预测任务中存在的不足、纹理信息复杂以及算法单一等问题。该材料性能预测框架由四个子模块组成：特征提取模块、FAGC-PSS 模块、伪标签标注模块和性能预测模块。特征提取模块通过网络模型提取材料图像特征，FAGC-PSS 模块通过升维提取的图像特征，并

投影特征数据到预形状空间以及拟合特征数据并构建最优的测地曲线，伪标签标注模块用于为生成特征标注性能值，性能预测模型使用扩充数据实现材料性能预测任务。该方法通过 **FAGC-PSS** 模块和伪标签机制对材料性能预测任务进行特征数据扩充，实现基于小样本材料图像的数据规模下，训练出具有高预测精度和泛化能力的材料性能预测模型。

本论文对以上两个方法进行了大量的实验验证，实验表明上述方法在解决图像识别和材料性能预测问题时，尤其是在小样本数据规模的限制下，具有出色的性能表现和普适性，体现出较高的实用价值。同时能够有效减少了标注成本和人力资源消耗，特别是对于应用人工智能在数据稀缺或样本获取困难的领域具有重要意义。

5.2 工作展望

尽管本论文提出的方法能够在小样本数据规模下，提升机器学习模型的泛化能力和预测精度，并针对图像处理领域的图像识别和材料性能预测两种下游任务中设计相应的模型结构，但是目前方法仍然存在一定的局限性。为了进一步完善本论文方法的增强效果和有效性，本论文可以在如下几个方面进一步深入研究：

1. 基于自适应的特征选取和特征聚类，获取高质量或者高识别度的特征数据。在本文的特征提取方法中，采用深度学习的网络模型提取小样本图像特征，特征数据的质量影响后续生成特征的增强效果，但是小样本图像本身数据规模受限，现有的复杂网络模型也很难获取到高质量的特征样本。因此构建自适应的特征选取方法，有助于算法模型排除孤立点和噪声等干扰，同时结合特征聚类的方式，将同类别下的特征数据聚类分组来构建多条拟合的测地曲线，这样生成的特征数据具有更丰富的多样性和复杂性。
2. 进一步提高基于 **FAGC-PSS** 材料性能预测模型的精度。在本论文的机器学习方法中，回归模型在对材料图像的性能预测中取得了良好的结果。然而，通过对比伪标签的质量与模型性能之间的实验可以发现，伪标签的标注模型很大程度影响后续的性能预测结果，如果采用集成多个回归模型的方式有可能进一步提高预测精度。同时材料图像种类复杂繁多，构建一条测地曲线可能无法适用于特殊材料的性能预测任务，更适合的方式是构建多条测地曲线或者是生成测地

曲面，以满足预测模型对数据的数量和多样性的需求。因此如何在小样本图像环境下，进一步改善 FAGC-PSS 模块来提升生成特征的质量具有较高的研究价值与实用意义。

尽管上述研究方向取得了一定进展，但仍需大量的探索 and 实验，才能取得相应的研究成果。未来之路充满挑战，任重而道远。

参考文献

- [1] LI N, MA L, XING T, et al. Automatic design of machine learning via evolutionary computation: A survey[J]. *Applied Soft Computing*, 2023: 110412.
- [2] CHANG T A, BERGEN B K. Language model behavior: A comprehensive survey [J]. *Computational Linguistics*, 2024: 1-58.
- [3] ZHANG P, ZHONG Y, DENG Y, et al. A survey on deep learning of small sample in biomedical image analysis[A]. 2019. arXiv preprint arXiv:1908.00473.
- [4] 葛轶洲, 刘恒, 王言, 等. 小样本困境下的深度学习图像识别综述[J]. *软件学报*, 2022, 33(1): 18.
- [5] ZHANG Y, SHEN Z, JIAO R. Segment anything model for medical image segmentation: Current applications and future directions[J]. *Computers in Biology and Medicine*, 2024: 108238.
- [6] PANDURI B, RAO O S. A survey on brain tumour segmentation techniques in deep learning[J]. *International Journal of Intelligent Systems and Applications in Engineering*, 2024, 12(7s): 412-425.
- [7] HUANG P, DONG J, HAN X, et al. Prediction of mechanical properties of composite materials based on convolutional neural network-long and short-term memory neural network[J]. *Metalurgija*, 2024, 63(3-4): 369-372.
- [8] GARCEA F, SERRA A, LAMBERTI F, et al. Data augmentation for medical imaging: A systematic literature review[J]. *Computers in Biology and Medicine*, 2023, 152: 106391.
- [9] SUN Q, LIU Y, CHUA T S, et al. Meta-transfer learning for few-shot learning[C]// *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019: 403-412.
- [10] MIKOŁAJCZYK A, GROCHOWSKI M. Data augmentation for improving deep learning in image classification problem[C]// *2018 international interdisciplinary PhD workshop (IIPhDW)*. IEEE, 2018: 117-122.

- [11] DING K, XU Z, TONG H, et al. Data augmentation for deep graph learning: A survey [J]. ACM SIGKDD Explorations Newsletter, 2022, 24(2): 61-77.
- [12] BAYER M, KAUFHOLD M A, REUTER C. A survey on data augmentation for text classification[J]. ACM Computing Surveys, 2022, 55(7): 1-39.
- [13] XU M, YOON S, FUENTES A, et al. A comprehensive survey of image augmentation techniques for deep learning[J]. Pattern Recognition, 2023, 137: 109347.
- [14] PELLICER L F A O, FERREIRA T M, COSTA A H R. Data augmentation techniques in natural language processing[J]. Applied Soft Computing, 2023, 132: 109803.
- [15] PHAM N T, DANG D N M, NGUYEN N D, et al. Hybrid data augmentation and deep attention-based dilated convolutional-recurrent neural networks for speech emotion recognition[J]. Expert Systems with Applications, 2023, 230: 120608.
- [16] KHALIFA N E, LOEY M, MIRJALILI S. A comprehensive survey of recent trends in deep learning for digital images augmentation[J]. Artificial Intelligence Review, 2022, 55(3): 2351-2377.
- [17] CHEN J, TAM D, RAFFEL C, et al. An empirical survey of data augmentation for limited data learning in nlp[J]. Transactions of the Association for Computational Linguistics, 2023, 11: 191-211.
- [18] DING K, LI R, XU Y, et al. Adaptive data augmentation for mandarin automatic speech recognition[J]. Applied Intelligence, 2024: 1-14.
- [19] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[J]. Advances in neural information processing systems, 2014, 27.
- [20] KINGMA D P, WELLING M. Auto-encoding variational bayes[A]. 2013. arXiv preprint arXiv:1312.6114.
- [21] KOBYZEV I, PRINCE S J, BRUBAKER M A. Normalizing flows: An introduction and review of current methods[J]. IEEE transactions on pattern analysis and machine intelligence, 2020, 43(11): 3964-3979.
- [22] YANG L, ZHANG Z, SONG Y, et al. Diffusion models: A comprehensive survey of methods and applications[J]. ACM Computing Surveys, 2023, 56(4): 1-39.

- [23] GAO F, YANG Y, WANG J, et al. A deep convolutional generative adversarial networks (dcgans)-based semi-supervised method for object recognition in synthetic aperture radar (sar) images[J]. Remote Sensing, 2018, 10(6): 846.
- [24] MIRZA M, OSINDERO S. Conditional generative adversarial nets[A]. 2014. arXiv preprint arXiv:1411.1784.
- [25] LI C, XU T, ZHU J, et al. Triple generative adversarial nets[J]. Advances in neural information processing systems, 2017, 30.
- [26] ZHANG H, XU T, LI H, et al. Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks[C]//Proceedings of the IEEE international conference on computer vision. 2017: 5907-5915.
- [27] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein generative adversarial networks[C]//International conference on machine learning. PMLR, 2017: 214-223.
- [28] GULRAJANI I, AHMED F, ARJOVSKY M, et al. Improved training of wasserstein gans[J]. Advances in neural information processing systems, 2017, 30.
- [29] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//Proceedings of the IEEE international conference on computer vision. 2017: 2223-2232.
- [30] MO S, CHO M, SHIN J. Instagan: Instance-aware image-to-image translation[A]. 2018. arXiv preprint arXiv:1812.10889.
- [31] PERARNAU G, VAN DE WEIJER J, RADUCANU B, et al. Invertible conditional gans for image editing[A]. 2016. arXiv preprint arXiv:1611.06355.
- [32] SOHN K, LEE H, YAN X. Learning structured output representation using deep conditional generative models[J]. Advances in neural information processing systems, 2015, 28.
- [33] ZHAO S, SONG J, ERMON S. Infovae: Balancing learning and inference in variational autoencoders[C]//Proceedings of the aaai conference on artificial intelligence: Vol. 33. 2019: 5885-5892.

- [34] HIGGINS I, MATTHEY L, PAL A, et al. beta-vae: Learning basic visual concepts with a constrained variational framework.[J]. ICLR (Poster), 2017, 3.
- [35] HUANG X, BELONGIE S. Arbitrary style transfer in real-time with adaptive instance normalization[C]//Proceedings of the IEEE international conference on computer vision. 2017: 1501-1510.
- [36] BERTHELOT D, RAFFEL C, ROY A, et al. Understanding and improving interpolation in autoencoders via an adversarial regularizer[A]. 2018. arXiv preprint arXiv:1807.07543.
- [37] KINGMA D P, SALIMANS T, JOZEFOWICZ R, et al. Improved variational inference with inverse autoregressive flow[J]. Advances in neural information processing systems, 2016, 29.
- [38] GIRIN L, LEGLAIVE S, BIE X, et al. Dynamical variational autoencoders: A comprehensive review[A]. 2020. arXiv preprint arXiv:2008.12595.
- [39] DINH L, SOHL-DICKSTEIN J, BENGIO S. Density estimation using real nvp[A]. 2016. arXiv preprint arXiv:1605.08803.
- [40] PAPAMAKARIOS G, NALISNICK E, REZENDE D J, et al. Normalizing flows for probabilistic modeling and inference[J]. Journal of Machine Learning Research, 2021, 22(57): 1-64.
- [41] DINH L, KRUEGER D, BENGIO Y. Nice: Non-linear independent components estimation[A]. 2014. arXiv preprint arXiv:1410.8516.
- [42] CROITORU F A, HONDRU V, IONESCU R T, et al. Diffusion models in vision: A survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023.
- [43] SONG J, MENG C, ERMON S. Denoising diffusion implicit models[A]. 2020. arXiv preprint arXiv:2010.02502.
- [44] HO J, JAIN A, ABBEEL P. Denoising diffusion probabilistic models[J]. Advances in neural information processing systems, 2020, 33: 6840-6851.
- [45] HOOGEBOOM E, GRITSENKO A A, BASTINGS J, et al. Autoregressive diffusion models[A]. 2021. arXiv preprint arXiv:2110.02037.

- [46] SAUER A, LORENZ D, BLATTMANN A, et al. Adversarial diffusion distillation [A]. 2023. arXiv preprint arXiv:2311.17042.
- [47] YANG S, XIAO W, ZHANG M, et al. Image data augmentation for deep learning: A survey[A]. 2022. arXiv preprint arXiv:2204.08610.
- [48] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[J]. Advances in neural information processing systems, 2012, 25.
- [49] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[A]. 2014. arXiv preprint arXiv:1409.1556.
- [50] 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.
- [51] JADERBERG M, SIMONYAN K, ZISSERMAN A, et al. Spatial transformer networks[J]. Advances in neural information processing systems, 2015, 28.
- [52] DENG J, DONG W, SOCHER R, et al. Imagenet: A large-scale hierarchical image database[C]//2009 IEEE conference on computer vision and pattern recognition. Ieee, 2009: 248-255.
- [53] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [54] DENG L. The mnist database of handwritten digit images for machine learning research[J]. IEEE signal processing magazine, 2012, 29(6): 141-142.
- [55] ZHANG H, CISSE M, DAUPHIN Y N, et al. mixup: Beyond empirical risk minimization[A]. 2017. arXiv preprint arXiv:1710.09412.
- [56] DEVRIES T, TAYLOR G W. Improved regularization of convolutional neural networks with cutout[A]. 2017. arXiv preprint arXiv:1708.04552.
- [57] YUN S, HAN D, OH S J, et al. Cutmix: Regularization strategy to train strong classifiers with localizable features[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2019: 6023-6032.

- [58] HENDRYCKS D, MU N, CUBUK E D, et al. Augmix: A simple data processing method to improve robustness and uncertainty[A]. 2019. arXiv preprint arXiv:1912.02781.
- [59] CHEN P, LIU S, ZHAO H, et al. Gridmask data augmentation[A]. 2020. arXiv preprint arXiv:2001.04086.
- [60] PATHAK D, KRAHENBUHL P, DONAHUE J, et al. Context encoders: Feature learning by inpainting[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 2536-2544.
- [61] GIDARIS S, SINGH P, KOMODAKIS N. Unsupervised representation learning by predicting image rotations[A]. 2018. arXiv preprint arXiv:1803.07728.
- [62] DAI M, HANG H, GUO X. Adaptive feature interpolation for low-shot image generation[C]//European Conference on Computer Vision. Springer, 2022: 254-270.
- [63] BELKIN M, NIYOGI P. Laplacian eigenmaps for dimensionality reduction and data representation[J]. Neural computation, 2003, 15(6): 1373-1396.
- [64] 彭亚新, 赵倩. 基于流形假设的骨架序列动作识别算法[J]. 上海大学学报(自然科学版), 2022, 28(2): 179-200.
- [65] KUME A, DRYDEN I L, LE H. Shape-space smoothing splines for planar landmark data[J]. Biometrika, 2007, 94(3): 513-528.
- [66] CHEN D, LI J, WANG Z, et al. Learning canonical shape space for category-level 6d object pose and size estimation[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 11973-11982.
- [67] LV C, LIN W, ZHAO B. Kss-icp: point cloud registration based on kendall shape space[J]. IEEE Transactions on Image Processing, 2023, 32: 1681-1693.
- [68] LIU N, ZHANG D, RU X, et al. Caricature expression extrapolation based on kendall shape space theory[J]. IEEE Computer Graphics and Applications, 2021, 41(3): 71-84.
- [69] LI Z, WU D, YU T. Prediction of material removal rate for chemical mechanical planarization using decision tree-based ensemble learning[J]. Journal of Manufacturing Science and Engineering, 2019, 141(3): 031003.

- [70] DU C, ZOU G, ZHANWEN A, et al. Highly accurate and efficient prediction of effective thermal conductivity of sintered silver based on deep learning method[J]. International Journal of Heat and Mass Transfer, 2023, 201: 123654.
- [71] WEI H, ZHAO S, RONG Q, et al. Predicting the effective thermal conductivities of composite materials and porous media by machine learning methods[J]. International Journal of Heat and Mass Transfer, 2018, 127: 908-916.
- [72] AI Y, LEI C, CHENG J, et al. Prediction of weld area based on image recognition and machine learning in laser oscillation welding of aluminum alloy[J]. Optics and Lasers in Engineering, 2023, 160: 107258.
- [73] 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.
- [74] REN D, WANG C, WEI X, et al. Building a quantitative composition-microstructure-property relationship of dual-phase steels via multimodal data mining[J]. Acta Materialia, 2023, 252: 118954.
- [75] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International journal of computer vision, 2004, 60: 91-110.
- [76] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR): Vol. 1. 2005: 886-893.
- [77] LI P, LI D, LI W, et al. A simple feature augmentation for domain generalization[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 8886-8895.
- [78] LI B, WU F, LIM S N, et al. On feature normalization and data augmentation[C]// Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021: 12383-12392.
- [79] CHEN S, WEBB G I, LIU L, et al. A novel selective naïve bayes algorithm[J]. Knowledge-Based Systems, 2020, 192: 105361.

- [80] VAPNIK V. The nature of statistical learning theory[M]. Springer science & business media, 2013.
- [81] AHLAWAT S, CHOUDHARY A. Hybrid cnn-svm classifier for handwritten digit recognition[J]. Procedia Computer Science, 2020, 167: 2554-2560.
- [82] DAHL G E, YU D, DENG L, et al. Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition[J]. IEEE Transactions on audio, speech, and language processing, 2011, 20(1): 30-42.
- [83] ZHAO Z, WU Y. Attention-based convolutional neural networks for sentence classification.[C]//Interspeech: Vol. 8. 2016: 705-709.
- [84] GOODFELLOW I, BENGIO Y, COURVILLE A. Deep learning[M]. MIT press, 2016.
- [85] TAN M, LE Q. Efficientnet: Rethinking model scaling for convolutional neural networks[C]//International conference on machine learning. PMLR, 2019: 6105-6114.
- [86] 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.
- [87] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.
- [88] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[A]. 2020. arXiv preprint arXiv:2010.11929.
- [89] KENDALL D G. Shape manifolds, procrustean metrics, and complex projective spaces[J]. Bulletin of the London mathematical society, 1984, 16(2): 81-121.
- [90] KENDALL D G, BARDEN D, CARNE T K, et al. Shape and shape theory[M]. John Wiley & Sons, 2009.
- [91] SMALL C G. The statistical theory of shape[M]. Springer Science & Business Media, 2012.

- [92] DRYDEN I L, MARDIA K V. Statistical shape analysis: with applications in r: Vol. 995[M]. John Wiley & Sons, 2016.
- [93] HAN Y, WANG B, IDESAWA M, et al. Recognition of multiple configurations of objects with limited data[J]. Pattern Recognition, 2010, 43(4): 1467-1475.
- [94] KENDALL D G, LE H, BARDEN D, et al. Shape and shape theory[M]. Wiley, 1999.
- [95] SMALL C G, SMALL C G. Distributions of random shapes[J]. The Statistical Theory of Shape, 1996: 149-172.
- [96] ZHANG J, ZHANG X, KRIM H. Invariant object recognition by shape space analysis [C]//Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No. 98CB36269). IEEE, 1998: 581-585.
- [97] HAN Y. Recognize objects with three kinds of information in landmarks[J]. Pattern Recognition, 2013, 46(11): 2860-2873.
- [98] KENOBI K, DRYDEN I L, LE H. Shape curves and geodesic modelling[J]. Biometrika, 2010, 97(3): 567-584.
- [99] HAN Y, KOIKE H, IDESAWA M. Recognizing objects with multiple configurations [J]. Pattern Analysis and Applications, 2014, 17: 195-209.
- [100] LEE D H, et al. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks[C]//Workshop on challenges in representation learning, ICML: Vol. 3. Atlanta, 2013: 896.
- [101] SOHN K, BERTHELOT D, CARLINI N, et al. Fixmatch: Simplifying semi-supervised learning with consistency and confidence[J]. Advances in neural information processing systems, 2020, 33: 596-608.
- [102] KIM J, MIN Y, KIM D, et al. Conmatch: Semi-supervised learning with confidence-guided consistency regularization[C]//European Conference on Computer Vision. Springer, 2022: 674-690.
- [103] ZHANG B, WANG Y, HOU W, et al. Flexmatch: Boosting semi-supervised learning with curriculum pseudo labeling[J]. Advances in Neural Information Processing Systems, 2021, 34: 18408-18419.

- [104] KRIZHEVSKY A. Learning multiple layers of features from tiny images: TR-2009 [R/OL]. University of Toronto, 2009. <https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf>.
- [105] COATES A, NG A, LEE H. An analysis of single-layer networks in unsupervised feature learning[C]//Proceedings of the fourteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings, 2011: 215-223.
- [106] HELBER P, BISCHKE B, DENGEL A, et al. Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019, 12(7): 2217-2226.
- [107] HELBER P, BISCHKE B, DENGEL A, et al. Introducing eurosat: A novel dataset and deep learning benchmark for land use and land cover classification[C]//IGARSS 2018-2018 IEEE international geoscience and remote sensing symposium. IEEE, 2018: 204-207.
- [108] LE Y, YANG X. Tiny imagenet visual recognition challenge[J]. CS 231N, 2015, 7 (7): 3.
- [109] ZAGORUYKO S, KOMODAKIS N. Wide residual networks[A]. 2016. arXiv preprint arXiv:1605.07146.
- [110] LI J, XIONG C, HOI S C. Comatch: Semi-supervised learning with contrastive graph regularization[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 9475-9484.
- [111] RASMUS A, BERGLUND M, HONKALA M, et al. Semi-supervised learning with ladder networks[J]. Advances in neural information processing systems, 2015, 28.
- [112] TARVAINEN A, VALPOLA H. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results[J]. Advances in neural information processing systems, 2017, 30.
- [113] BERTHELOT D, CARLINI N, GOODFELLOW I, et al. Mixmatch: A holistic approach to semi-supervised learning[J]. Advances in neural information processing systems, 2019, 32.

- [114] MIYATO T, MAEDA S I, KOYAMA M, et al. Virtual adversarial training: a regularization method for supervised and semi-supervised learning[J]. IEEE transactions on pattern analysis and machine intelligence, 2018, 41(8): 1979-1993.
- [115] XU Y, SHANG L, YE J, et al. Dash: Semi-supervised learning with dynamic thresholding[C]//International Conference on Machine Learning. PMLR, 2021: 11525-11536.
- [116] FAN Y, KUKLEVA A, DAI D, et al. Revisiting consistency regularization for semi-supervised learning[J]. International Journal of Computer Vision, 2023, 131(3): 626-643.
- [117] XIE Q, DAI Z, HOVY E, et al. Unsupervised data augmentation for consistency training[J]. Advances in neural information processing systems, 2020, 33: 6256-6268.
- [118] BERTHELOT D, ROELOFS R, SOHN K, et al. Adamatch: A unified approach to semi-supervised learning and domain adaptation[A]. 2021. arXiv preprint arXiv:2106.04732.
- [119] WANG Y, CHEN H, HENG Q, et al. Freematch: Self-adaptive thresholding for semi-supervised learning[A]. 2022. arXiv preprint arXiv:2205.07246.
- [120] ZHENG M, YOU S, HUANG L, et al. Simmatch: Semi-supervised learning with similarity matching[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 14471-14481.
- [121] CHEN H, TAO R, FAN Y, et al. Softmatch: Addressing the quantity-quality trade-off in semi-supervised learning[A]. 2023. arXiv preprint arXiv:2301.10921.
- [122] LEWY D, MAŃDZIUK J. An overview of mixing augmentation methods and augmentation strategies[J]. Artificial Intelligence Review, 2023, 56(3): 2111-2169.
- [123] ZHANG S, CHENG D, DENG Z, et al. A novel knn algorithm with data-driven k parameter computation[J]. Pattern Recognition Letters, 2018, 109: 44-54.
- [124] CERVANTES J, GARCIA-LAMONT F, RODRÍGUEZ-MAZAHUA L, et al. A comprehensive survey on support vector machine classification: Applications, challenges and trends[J]. Neurocomputing, 2020, 408: 189-215.

- [125] CHARBUTY B, ABDULAZEEZ A. Classification based on decision tree algorithm for machine learning[J]. Journal of Applied Science and Technology Trends, 2021, 2 (01): 20-28.
- [126] GEURTS P, ERNST D, WEHENKEL L. Extremely randomized trees[J]. Machine learning, 2006, 63: 3-42.
- [127] GENUER R, POGGI J M, GENUER R, et al. Random forests[M]. Springer, 2020.
- [128] ALTMAN N, KRZYWINSKI M. Ensemble methods: bagging and random forests[J]. Nature Methods, 2017, 14(10): 933-935.
- [129] BENTÉJAC C, CSÖRGŐ A, MARTÍNEZ-MUÑOZ G. A comparative analysis of gradient boosting algorithms[J]. Artificial Intelligence Review, 2021, 54: 1937-1967.
- [130] POPESCU M C, BALAS V E, PERESCU-POPESCU L, et al. Multilayer perceptron and neural networks[J]. WSEAS Transactions on Circuits and Systems, 2009, 8(7): 579-588.
- [131] RAMACHANDRAN P, ZOPH B, LE Q V. Searching for activation functions[A]. 2017. arXiv preprint arXiv:1710.05941.
- [132] KINGMA D P, BA J. Adam: A method for stochastic optimization[A]. 2014. arXiv preprint arXiv:1412.6980.
- [133] SHAHRAKI A, ABBASI M, HAUGEN Ø. Boosting algorithms for network intrusion detection: A comparative evaluation of real adaboost, gentle adaboost and modest adaboost[J]. Engineering Applications of Artificial Intelligence, 2020, 94: 103770.
- [134] CAI T T, MA R. Theoretical foundations of t-sne for visualizing high-dimensional clustered data[J]. Journal of Machine Learning Research, 2022, 23(301): 1-54.

攻读硕士学位期间取得的研究成果

攻读硕士期间论文发表情况：

1. Yuxing Han, Guanxin Wan, Yi Liu, Bing Wang. "Conductivity and Hardness Prediction of Cu-Cr-Zr via Feature Augmentation on Geodesic Curve," 2024 5th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT), Nanjing, China, 2024, pp. 1937-1941, doi: 10.1109/AINIT61980.2024.10581478. (导师一作, 本人二作, 已见刊, EI 会议)
2. Han Y, Wan G, Wang B. FAGC: Feature Augmentation on Geodesic Curve in the Pre-Shape Space. (导师一作, 本人二作, submitted: IEEE Transactions on Image Processing)
3. Han Y, Wan G, Han T, Wang B, Liu Y. Revealing the structure-property relationships of copper alloys with FAGC. (导师一作, 本人二作, submitted)

攻读硕士期间专利获取情况：

1. 专利名称：一种数据特征增强方法、装置、电子设备及介质，发明人：韩越兴、万冠新、王冰。申请号或专利号：2023111837377，申请公布号：CN 117076910 A，申请日：2023 年 9 月 14 日，公开日：2023 年 11 月 17 日。(导师一作, 本人二作)

攻读硕士期间软著获取情况：

1. 软件名称：基于形状空间理论和机器学习方法的材料性能预测平台 V1.0，开发者：韩越兴、万冠新、王冰。登记号：2023SR1427901，申请人：上海大学，开发完成日期：2023 年 6 月 2 日，登记日期：2023 年 7 月 1 日。

攻读硕士学位期间参与的科研项目

1. 项目来源：国家自然科学基金面上项目

项目名称：基于小样本机器学习的 Mg-Gd 系合金图像-组织-性能的内禀关系研究

项目编号：52273228

执行期限：2023.01-2026.12

2. 项目来源：科学技术部

项目名称：知识与数据双驱动的树脂基复合材料智能设计共性关键技术研发与空间应用

项目编号：2022YFB3707800

执行期限：2022.11-2025.10

3. 项目来源：科技部国家重点研发计划

项目名称：材料大数据技术研究

项目编号：2020YFB0704500

执行期限：2020.09-2022.08

4. 项目来源：上海市自然科学基金

项目名称：小样本环境下物体自适应识别方法研究

项目编号：20ZR1419000

执行期限：2020.07-2023.06

5. 项目来源：科技部国家重点研发计划

项目名称：材料复杂显微结构高精度图像处理方法研究

项目编号：2018YFB0704400 2018YFB0704402

执行期限：2018.7-2022.6

致 谢

光阴荏苒，岁月如梭。转眼间已度过研究生的三年宝贵时光，即将告别学校，心中充满感慨与不舍。回首在校学习研究时光，我深深感受到了身边帮助过我的老师、同学和家人的无私支持与关怀，在此特别对所有给予我帮助的人们表达最诚挚的感谢。

首先，我要感谢我的导师韩越兴老师。韩老师不仅在学术上给予我无微不至的指导和帮助，更在人生道路上给予我悉心的关怀。他拥有丰富的学术和严谨的治学态度，无论是论文撰写、数据分析还是实验设计等方面，都给予了我大量的宝贵建议和指导。在我遇到困难和挑战的时候，韩老师总是耐心倾听，给予我鼓励和支持，让我在迷茫的学术研究中找到前进的方向。此外，韩老师对于科研的热爱和对细节的追求精神，深深地感染和激励着我。他的严谨治学态度、对科研工作的热情和耐心，是我学习的楷模和榜样。在他身上，我学到了如何对待科研工作，如何面对困难和挑战，如何保持谦逊和勤奋。在学习工作之余，韩老师还总是关心我们的日常生活，疫情爆发期间，韩老师时刻关系我们的身体健康和心理状态，他的关心和帮助让我们感受到在校生活的温暖和安心。在此祝愿韩老师身体健康、前程万里、桃李满天下。

其次，我要诚挚地感谢我的父母。是他们给予了我在校求学研究的机会，让我得以拥有这段宝贵的学习时光。我更要感激他们二十余年来无私的养育之恩。言语难以表达我内心的感激之情，但我真诚地祝愿你们身体健康，福寿无疆，愿您们年年都能享受丰盛的喜悦和幸福。

接着，我还要衷心感谢陈侨川老师、王冰老师以及所有合作老师对我在研究生期间的悉心指导与无私支持。没有你们的帮助，我无法顺利完成研究工作。在这三年的学习生涯中，每一次与你们的交流都是我宝贵的学习经历，让我受益匪浅，收获颇丰。在我迷茫时，你们给予了耐心的解答和指导；在我遇到困难时，你们伸出援手，给予了及时的支持。在我取得成绩时，你们的欣慰与鼓励更激励着我不断前行。在此，我衷心祝愿各位老师幸福安康，心想事成。

然后，我还要感谢课题组的每一位小伙伴，经常分享美食的王璐学姐、专注科研的胡干学长、跟我一起挑灯夜战的韩思凡同学、喜欢旅游和滑雪的池洳婷同学、思维敏捷的阮礼恒学弟、可靠稳健的包胜奇学弟、热爱跳舞的赵宸学弟、羽球人的凌晨帆学弟以及一直陪伴和鼓励我的小慧同学，正是你们让我的研究生生活更加丰富和精彩。祝愿每位小伙伴的未来充满了温暖和阳光，前程一片美好景象。

最后，向百忙中抽出时间的盲审专家和诸位答辩老师致以最诚挚的感谢，谢谢你们。