

中图分类号:

单位代号: 10280

密 级:

学 号: 17721731

上海大学



硕士学位论文

SHANGHAI UNIVERSITY
MASTER'S DISSERTATION

题 目	基于材料小样本数据的分割与识别方法研究
--------	---------------------

作 者 宋磊磊

学科专业 计算机科学与技术

导 师 韩越兴

完成日期 2020 年 2 月

姓 名： 宋磊磊

学号： 17721731

论文题目： 基于材料小样本数据的分割与识别方法研究

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：

主任：

委员：

导 师：

答辩日期：

姓 名：宋磊磊

学号：17721731

论文题目：基于材料小样本数据的分割与识别方法研究

原创性声明

本人声明：所呈交的论文是本人在导师指导下进行的研究工作。除了文中特别加以标注和致谢的地方外，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同志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。

签 名：_____日 期：_____

本论文使用授权说明

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）

签 名：_____导师签名：_____日期：_____

上海大学工学硕士学位论文

基于材料小样本数据的分割与识别 方法研究

姓 名： 宋磊磊

导 师： 韩越兴

学科专业： 计算机科学与技术

上海大学计算机工程与科学学院

2020 年 2 月

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

Research on Segmentation and Recognition Method Based on Small Sample Data of Materials

MA Candidate: Leilei Song

Supervisor: Yuexing Han

Major: Computer science and Technology

School of computer science and engineering

Shanghai University

02, 2020

摘 要

根据材料数据的统计规律探究材料的宏观性能,对完善材料的制备工艺、开发新型材料等具有极其重要的作用。因材料数据具有复杂、多样、不均匀分布和小样本等特性,给材料数据的统计、处理和分析工作带来了巨大挑战。目前,对材料数据统计、处理和分析工作大分部是由材料相关领域内的专家完成,对人力成本的要求极其昂贵。随着计算机技术的发展,利用计算机技术处理材料数据已成为必然趋势,但受限于材料制备工艺的复杂、材料数据获取的人力成本和时间成本等,难以有大量的材料数据供于计算机处理和分析。因此,只有解决材料小样本的问题,才能加快材料科学与计算机科学的交融,加速材料的制备和研发。为了分析和解决典型的材料小样本问题,本文分别对三种材料小样本数据进行处理和分析,并利用机器学习、深度学习以及复杂网络等相关理论知识有针对性地解决材料小样本问题。本文处理的三种小样本数据分别是金属拉升形变过程中的晶体结构数据、原子力显微镜(AFM)下的纳米材料图像和陶瓷的微观结构图像数据。针对晶体结构的数据,本文提出了一种基于机器学习的晶体结构识别方法,该方法通过对样本数据进行扩增,然后根据原子在空间中的分布特点定义特征表示以及特征选择,最后利用机器学习实现对晶体结构的识别。针对 AFM 图像数据,本文提出一种基于迁移学习的识别方法,该方法通过预处理实现图像去噪,然后改进分水岭分割方法实现对重叠目标的分割,最后利用迁移学习实现对目标物体的识别。针对陶瓷的微观结构图像数据,本文提出了一种基于复杂网络理论的图像分割方法,该方法根据图像的像素值分布来生成网络空间的节点集,然后定义节点之间的相似性并生成网络拓扑结构,最后实现网络拓扑结构的优化以及图像分割。本文提出的三种小样本处理方法不仅能够处理本文给出的小样本材料数据,也能够给材料科学中其他小样本问题的研究带来启示。

关键词: 材料数据; 小样本; 机器学习; 迁移学习; 复杂网络

ABSTRACT

Exploring the macroscopic properties of materials according to the statistical laws of data plays an important role in improving the preparation process of materials and developing new materials. Material data is complex, diverse, messy distributed and small samples, which brings huge challenges to the statistics, processing and analysis of data. The traditional material analysis and processing work is performed by experts, and the demand for labor costs is extremely high. With the development of computer technology, the use of computers to process and analyze material data has become an inevitable trend. However, due to the complexity of the preparation, the cost of labor and time, the data used for computer processing are insufficient. Solving the problem of small samples of materials is conducive to accelerating the integration of materials science and computer science, and accelerating the development and preparation of materials. To research the problem of small samples of materials, this thesis analyzes and processes three typical small samples of materials. The machine learning, deep learning, and complex networks are used to solve the problem of small samples of materials. The sample data processed in this thesis are crystal structure data, nanomaterial images under atomic force microscope (AFM), and microstructure images of ceramics. For the data of crystal structure, this thesis proposes a recognition method based on machine learning. This method amplifies data according to the small samples. Then, the features are defined based on the distribution characteristics of atoms in space. Finally, Machine learning is used to identify crystal structures. For AFM image, this thesis proposes a recognition method based on transfer learning. This method implements image denoising through preprocessing, and uses an improved watershed segmentation method to segment overlapping targets. Finally, it uses transfer learning to learn the segmented targets. For the microstructure image of ceramics, this thesis

proposes a segmentation method based on a complex network. First, the node set of the network is generated according to the pixel distribution. Then, the similarity is defined for the nodes to generate the network topology. Finally, the network structure is optimized and segmentation is implemented. The three proposed methods can not only process the material data given in the thesis, but also enlighten the other small sample problems.

Keywords: material data, small samples, machine learning, transfer learning, complex network

目 录

摘 要.....	V
ABSTRACT.....	VI
目 录.....	VIII
第一章 绪 论.....	1
1.1 课题来源.....	1
1.2 课题背景概述.....	1
1.3 课题研究的目的和意义.....	2
1.4 国内外研究概况.....	3
1.4.1 小样本研究现状.....	3
1.4.2 材料中的小样本研究现状.....	5
1.5 论文的主要研究内容.....	6
第二章 基于机器学习的晶体结构识别方法.....	7
2.1 晶体结构识别背景介绍.....	7
2.2 基于机器学习的分类方法.....	10
2.2.1 数据预处理.....	10
2.2.2 特征表示.....	11
2.2.3 特征选择.....	14
2.2.4 晶体结构识别.....	16
2.3 实验与讨论.....	18
2.3.1 实验环境和评估指标.....	18
2.3.2 特征选择的比较分析.....	19
2.3.3 与共近邻方法比较分析.....	22
2.4 本章小结.....	24
第三章 基于迁移学习的 AFM 图像分割识别方法.....	26
3.1 纳米材料 AFM 图像背景介绍.....	26
3.2 图像识别方法.....	28

3.2.1 AFM 图像预处理	28
3.2.2 基于距离的分水岭分割方法	29
3.2.3 多尺度的分水岭分割方法	32
3.2.4 AFM 图像物体识别	34
3.3 实验与分析	36
3.3.1 迁移学习的数据准备	36
3.3.2 实验比较与分析	38
3.4 本章小结	42
第四章 基于复杂网络理论的材料图像分割方法	44
4.1 复杂网络背景介绍	44
4.2 材料微观结构图像分割方法	45
4.2.1 预处理	45
4.2.2 网络拓扑结构建立	46
4.2.3 参数优化	48
4.2.4 轮廓提取与分割	49
4.3 实验	50
4.3.1 实验数据处理	50
4.3.2 网络构建与图像分割	51
4.4 总结	52
第五章 总结与展望	54
5.1 结论	54
5.2 工作展望	55
参考文献	56
作者在攻读硕士学位期间公开发表的论文与专利	64
致 谢	65

第一章 绪 论

1.1 课题来源

本课题来源国家重点研发计划：材料基因工程专用数据库平台建设与示范应用（科技部）中的“陶瓷涂层材料数据自动采集、整合与应用”的课题。项目编号：2018YFB0704400, 2018YFB0704402；上海市科学技术委员会浦江人才项目：“基于形状空间理论和复杂网络理论”中的“纳米材料图像处理方法的研究”，编号：17PJ1402900。

1.2 课题背景概述

随着近现代科学技术的不断发展，学术界、工业界对高性能材料的要求越来越高，技术的突破、行业的发展和经济的繁荣都依赖于先进材料的研发。然而新型材料的研发过程漫长且具有高成本，无法满足各行各业对材料的需求。如何加快先进材料的研发是 21 世纪亟需解决的问题，也是材料科学所面临的巨大难题和挑战。2011 年，美国提出一项跨越机构的研究计划，即“材料基因组计划”（Materials Genome Initiative, MGI）【1】，意在开创政策、资源和基础设施共同发展的新时代，制造和部署先进材料。2012 年，我国也推出了《材料科学系统工程发展战略研究——中国版材料基因组计划》重大项目，这对提升我国材料科学领域的核心技术水平以及工业制造能力具有重要战略意义。

随着计算机科学技术的不断突破和发展，计算机技术的应用越来越广，各研究领域对其依赖性也越来越强。近年来，材料基因的迅速发展得益于计算机技术的支持。计算机技术可以帮助处理材料数据、统计数据规律和分析材料宏观性能与材料数据的相关性，例如挖掘材料图像属性与材料微观结构之间的关系【2】【3】。然而由于材料制备工艺复杂、人工成本高、材料数据提取困难，致使可用于计算机进行处理、统计和分析的数据稀少。计算机擅长对大量数据进行统计、分析和归纳，但面对材料数据的小样本问题计算机该如何起到数据分析的作用，这是计算机科学技术应用于材料科学的一大难点与挑战。

如今身处大数据时代，但这并不意味着所有数据都充足的。大数据往往是指收集到的有效数据总量是巨大的，然而材料科学中有效数据往往非常有限。目前针对材料小样本数据的处理工作，很多情况下还是由人工完成，利用计算机技术对材料小样本数据进行自动化的分析、处理还停留在初级的数据增强方式，对材料小样本问题的探究还处在初级阶段，解决方法还在不断探索中【4-5】。

计算机对小样本数据的归纳和分析能力有限导致计算机在面临小样本数据时往往难以发现有效的规律进行分析和处理。随着计算机技术的发展，数据分析能力的提升，算法不断优化，而且在很多研究领域内小样本数据的问题阻碍着其学科的发展，小样本数据问题开始吸引广大学者和专家们的关注。针对小样本数据如何有效地进行处理、归纳、分析是我们亟需不断研究和探索的。

1.3 课题研究的目的是和意义

材料数据往往较为复杂，种类繁多，不同的材料因制备工艺、参数等差异而产生不同的结果，这给人工分析和处理数据带来巨大的困难。另一方面，材料体系庞大，不同数据的分析工作需要不同技术的支持，而人工能力有限，往往难以面面俱到，不同数据的分析工作可能需要不同研究领域的专家进行处理【6-7】，耗费着大量的人力。为了解决这一问题，需利用计算机技术对材料小样本数据进行自动化地分析和处理，以降低对人工的依赖性。此外，计算机本身具备着很高的数据分析能力，但这依赖于庞大数据的支撑。对于材料科学，材料的种类繁多、制备工艺复杂等因素，导致获取的数据往往不足以支撑计算机的分析工作。为了加速计算机科学与材料科学的交融，加快材料基因组的发展，针对材料小样本问题的研究亟需提到日程上来。

将计算机技术应用于材料科学中来，利用计算机技术代替人工对材料数据进行处理和分析，对提高材料科学研究的效率、加快材料研发的速度都具有极其重要的作用。通过计算机技术可以降低人工的工作量、避免主观因素的影响，进而提高数据分析的效率和准确性。另外，利用计算机技术处理、分析材料小样本数据才能加快材料科学与计算机科学的交叉融合、加速材料科学的发展、促进材料基因组计划的实施。此外，针对材料小样本数据的研究给其他学科的小样本数据

研究提供了借鉴意义，并为小样本研究奠定坚实基础。

1.4 国内外研究概况

1.4.1 小样本研究现状

材料数据往往因材料本身的稀缺，材料制备工艺的复杂等因素导致数据量有限，这严重阻碍着计算机技术对材料数据的处理、统计和分析工作，进而严重影响了材料科学的研究与发展。目前，针对小样本数据的研究已广受关注，其主要方法大概可以分四大类。

第一类是传统方法，这些方法无需大量数据。基于小样本环境下的传统目标识别，例如，著名的 SIFT【8-9】和 SURF【10-11】方法可以利用一个样本图片来实现相同目标的图像匹配和识别。但这些方法对目标物体的三维旋转过于敏感。AFoVs 方法采用少数图像数据进行旋转来识别不同角度的同一种物体【12-13】，该方法主要针对有规律的几何外形。Ren 把带有关节的物体外形用线条表示出来，通过研究这些线条来理解该物体的变化趋势，从而实现在少量数据情况下对物体形态的识别【14】。尽管这些方法都可以在一定程度上实现小样本环境下的目标识别，但是，它们都有严格的使用条件，难以达到很高的准确度。

第二类是数据增强，最直观的方法之一。数据增强方法通常使用增加噪声、裁剪、缩放、翻转和旋转等方式增加数据量。此外还可以对已有数据进行分析以生成相似的数据。自编码器（AutoEncoder, AE）是一种有效的、成熟的表示学习方法，属于无监督学习的一种【15】。VAE（Variational AutoEncoder）是最广泛使用的一种自编码器，VAE 能够把被关注物体比较清晰地描述出来【16】，从而生成相似的图像数据。此外，近年来随着以生成对抗网络（GAN, Generative Adversarial Networks）【17】为代表的大量基于深度学习的生成模型不断涌现，为数据增强提供了新思路，该模型为生成式网络模型，其优化过程借鉴于博弈论，通过交替训练判别模型 D(discriminative model)和生成模型 G(generative model)使得优化目标达到纳什均衡【18】。GAN 自从提出以来短短的几年时间里得到了迅猛发展，并取得了广泛的应用【19-20】。尽管这些生成模型能够从很少的样本

中生成更多的数据，但本质上都需要一定的基本数据进行模型的训练。当训练数据量过小时，得到的模型很难客观地反映现实物体的特性。

第三类是迁移学习，即在大规模数据集上学习到的知识或者模型应用于数据量有限的相关领域内。在机器学习或者深度学习中，经常将已经在其它数据集上训练好的模型，如将 ImageNet 数据集上训练好的模型，在新的数据集上进行微调，使模型适应新的数据集。例如，Socher 等设计了一种迁移学习模型，能够在缺乏学习数据的情况下识别出图像中的目标【21】。基于 Bayesian 模型，Cao 等结合 Kullback-Leibler 散度和相似性函数来构建最大似然估计函数，并利用主成分分解，有效地融合了不同目标感兴趣域的有限数据样本【22】。Gao 等直接采用高斯过程回归来融合现有的检测跟踪方案，在目标跟踪上取得了有效的提升【23】。Zheng 等利用迁移学习对人体运动进行识别【24】。尽管迁移学习在科学研究和工程项目上取得了一定成功，但如果用相似度较低的数据进行学习，会造成因过度拟合样本分布的无关信息而导致学习模型识别效率低下。

第四类是小样本学习 (Few-shot Learning)。小样本学习方法可以分为三种，分别为基于模型结构的方法、基于度量的方法和基于优化的方法。基于模型结构的方法是通过优化模型的结构，使得模型可以快速在少量样本上学习到特征，完成参数的更新，例如，Santoro 等人【25】采用记忆的神经网络模型解决小样本学习问题。基于度量的方法借助最近邻的思想，计算样本与样本之间的距离来设计网络模型，例如孪生网络【26】、原型网络【27】、相关网络【28】。孪生网络根据样本之间的距离建立并训练模型，使得同类样本靠近，异类样本疏远，然后利用网络所提取的特征进行小样本的学习；在原型网络中类别都有一个原型表达，即数据集在向量空间中的均值，因此将分类问题转换为向量空间中的最近邻问题；相关网络通过建立和训练网络模型来学习距离的度量方式。基于优化的方法认为普通的梯度下降方法难以在小样本环境下拟合数据，因此从参数优化方法方面解决小样本的问题，例如，Finn【29】提出使用少量的迭代步骤就能够获得较好的泛化能力，而且模型很容易进行微调。但是，小样本学习在复杂程度高、模态不同的任务上难以泛化，面对复杂图像识别、问答和逻辑命题结合的任务时还存在一定困难。此外，小样本学习对硬件设备的要求非常高，即使是处理简单的数据

集，小样本学习也需要在最新的 GPU 上运行数十甚至数百小时。

1.4.2 材料中的小样本研究现状

材料领域的数据处理往往都面临小样本的问题，但由于材料数据本身的复杂性，最新的一些小样本研究成果并无法成功应用于材料数据上来。目前，针对材料数据的小样本问题还偏向于使用一些传统方法。在材料图像的处理任务中，大多数还在使用一些传统算法，例如阈值分割算法【30】、图割法【31】、区域生长算法【32】等。这些算法都是直接针对像素处理，可以从一个像素点开始进行扩展，也可以从整个图像开始进行处理。这种处理方式对数据量没有要求，一张图像也可以处理，但其准确性却不高。在材料领域，一般先用这些方法进行简单处理，而后再由人工处理。例如，汤力琨【33】等人利用阈值分割法实现对材料图像的分割，裘俊彦等人【34】在阈值法的基础上结合数学形态学实现对不同形态的材料微观组织图像进行分割，Ananyev 等人【35】和 Lopez 等人【36】使用了形态学滤波方法对材料图像进行分割并定量分析。机器学习的迅猛发展，其优越性不断展现，在材料领域中也应用越来越广泛。目前，应用较为广泛的是传统机器学习算法，这些方法通过自定义特征，然后利用随机森林、神经网络、支持向量机等机器学习算法，实现对材料数据的分割或者识别。例如，Lai 等人【37】通过为材料图像上每个像素点生产一组特征向量，然后利用随机森林对每个像素点分类，实现了小样本数据下的图像分割，Wang 等人【38】提出了利用支持向量机算法实现对少数几张 AFM 图像分割的方法。还有一些无监督的机器学习算法，在材料用也是应用广泛，例如，Chen 等人【39】采用模糊聚类对材料图像中的像素进行划分从而实现图像分割，赵墨等人【40】利用谱聚类算法实现对陶瓷的背散射电子成像（Back scattered Electron Imaging, BSE）图像的分割，Han 等人【41】提出了一种基于区域和聚类的图像分割方法，可在图像样本极其稀少的情况下完成分割任务。近些年，深度学习发展十分迅猛，其主要优势在于无需人工定义特征，但因需要大规模的数据以及标签进行训练，这给小样本材料数据带来难题。因材料图像数据基本都是显微镜下获得的高分辨率微观组织结构图像，因此 Seyed 等人【42】对高分辨的图像进行重叠切分，从而得到很多低分辨率的

图像数据，用于深度学习训练，然后再将预测的结果进行组合。

总的来说，近些年关于材料小样本数据的研究越来越多，但是还没有有效的、统一的、通用的方法来处理和分析材料中的小样本数据。目前，针对材料小样本数据的分析工作仍具有很大的挑战，也是非常迫切的。

1.5 论文的主要研究内容

本论文以本人攻读硕士学位期间承担的相关项目研究作为基础，针对不同种类的小样本材料数据，分别研究、设计了相应的小样本数据解决方法，并通过很多的对比实验来验证本文所提方法的有效性。这些小样本数据分别是金属拉升形变过程中的晶体结构数据、原子力显微镜（AFM）下的纳米材料图像和陶瓷的微观结构图像。

论文其它章节安排如下：

第二章介绍了在晶体结构数据不足的情况下，结合数据增强、特征表示与选择和多层感知机等多种方法，提出了基于机器学习的晶体结构识别方法，并进行相关对比实验和验证实验。

第三章介绍了在 AFM 图像数据只有几张的情况下，结合高斯拉普拉斯（LOG）、分水岭和迁移学习等方法，从预处理、分割到识别，逐层递进的实现对 AFM 图像的识别，并进行了对比和验证实验。

第四章介绍了在材料微观结构图像极其匮乏的情况下，结合复杂网络中的相关理论方法，提出一种基于复杂网络的图像分割方法，并在 $\text{TiB}_2\text{-TiC}$ 的微观结构图像中进行了验证。

第五章总结了本文的研究内容、方法、成果以及各方法的优缺点，并阐述了未来工作。

第二章 基于机器学习的晶体结构识别方法

由于原子运动使得晶体结构发生变化,进而导致塑性形变。晶体材料的塑性形变与材料性能具有密切相关,研究材料形变机理对分析材料的性能以及理解形变机制具有极其重要的作用。然而,目前晶体结构识别常常面临样本数据不足的问题,且现有的一些方法并不能识别出所有的晶体结构。针对晶体结构中的小样本问题以及现存方法的不足,本章提出了一种基于机器学习的晶体结构识别方法。该方法解决了晶体结构的小样本的问题,且给晶体结构的研究带来新的突破,并首次实现对每个晶体结构类型的概率判断。本章所述方法已在期刊 *Modeling and Simulation in Materials Science and Engineering* 上发表。

2.1 晶体结构识别背景介绍

分析晶体结构有利于对材料性能和形变机制的研究,晶体结构一般可分为三种,即体心立方结构 BCC (Body-Centered Cubic),面心立方结构 FCC (Face Centered Cubic),密排六方结构 HCP (Hexagonal Close-packed)【43-44】。详细介绍如下:

1) BCC (Body-Centered Cubic) 结构: 一个中心原子位于立方体结构中心,且有八个原子分布于立方体的八个角上。每一个角原子同时也是其他体心立方体结构的角原子,因此角原子是共享的。BCC 结构如图 2.1 所示。

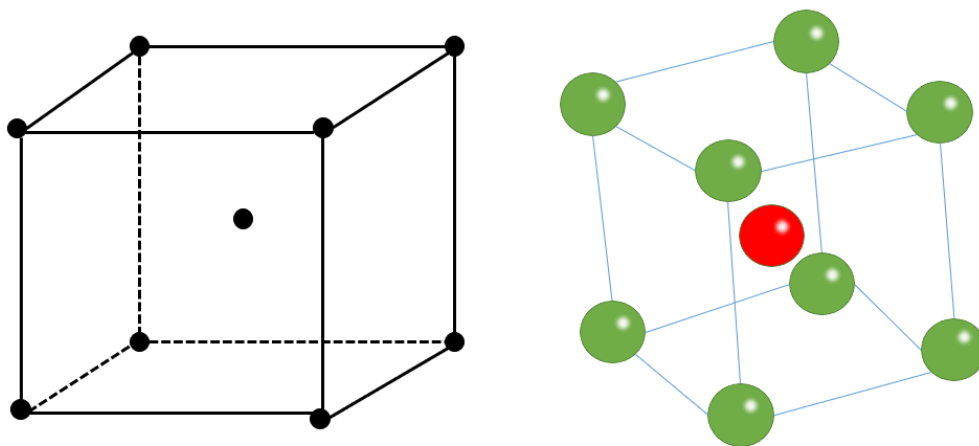


图 2.1 BCC 结构示意图。

BCC 结构不允许原子紧密排列在一起，常见于高温金属材料中，其堆积因子（packing factor，表示每单元体积内原子个数）约为 0.68，常见于钡、 α -铁、钾、铬、锂、钨和钠等金属中。

2) FCC (Face Centered Cubic) 结构：八个角原子与六个面中心原子构成面心立方体结构，即八个角原子构成立方体，立方体的每个面都有一面中心原子。与 BCC 结构类似，每个角原子也是另一个面心立方体的角原子，面中心原子同时也是另外面心立方体的面中心原子，因此角原子和面中心原子都是共享的。FCC 结构如图 2.2 所示。

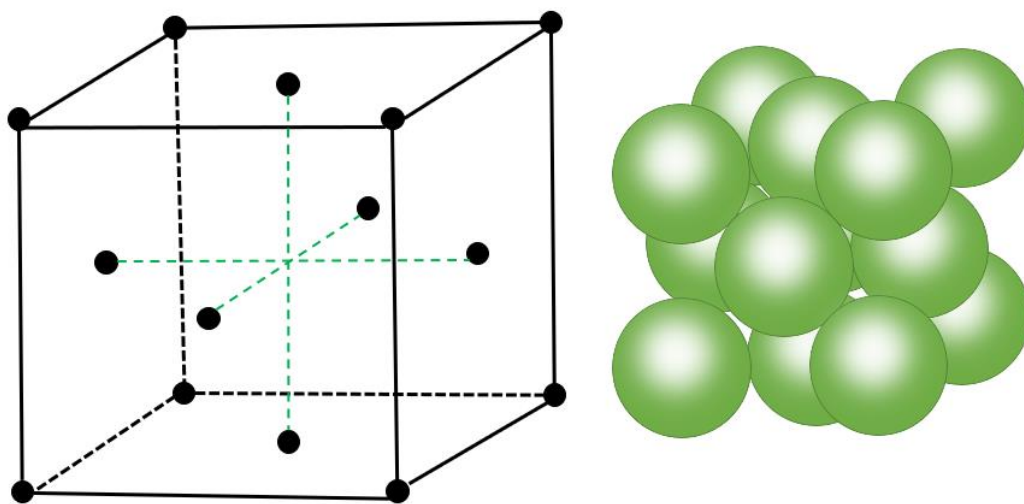


图 2.2 FCC 结构示意图。

相比于 BCC 结构 FCC 结构中原子的排布更加紧密，其堆积因子为 0.74，常见于银铝、铜、铂、铅、铀、金和镍等金属中。

3) HCP (Hexagonal Close-packed) 结构：六角晶包的 12 个顶角原子，两个底面中心原子以及晶包内部的三个共面原子构成了六排密方结构。交替层的六角结构发生位移，因此其原子与上一层的间隙对齐，与 FCC 结构类似，同一层中的原子将自己嵌套在相邻层原子中间的空白位置。HCP 结构如图 2.3 所示。

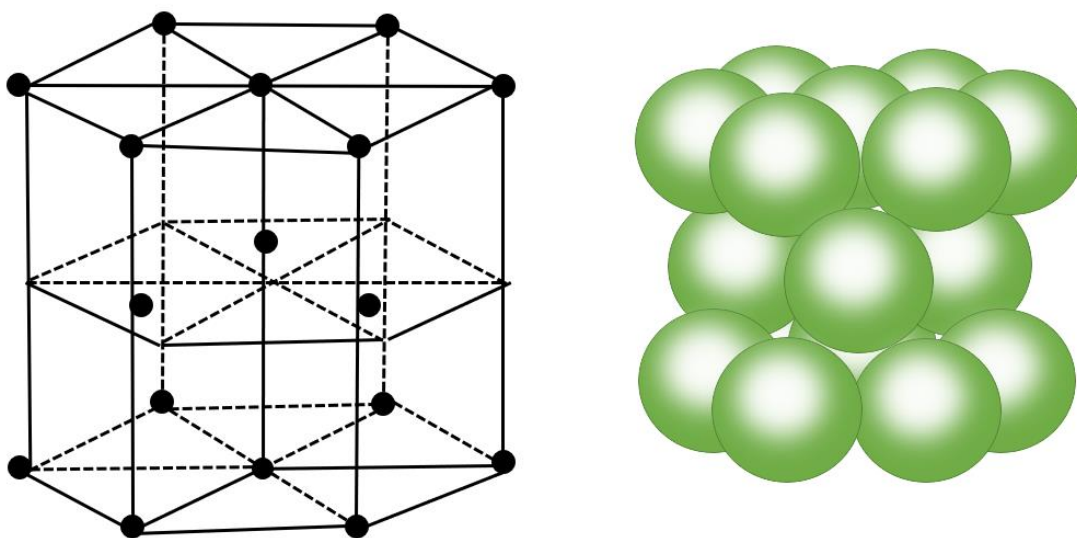


图 2.3 HCP 结构示意图。

由图 2.3 可见，HCP 结构有三层，在每个顶层和底层中，各有六个原子以六边形的形式排列，结构也较为紧密，其堆积因子为 0.74，常见于铅、镁、镉、钛、锌和铍等金属中。

为了探究晶体结构，许多方法被提出用于分析晶体的有效信息并识别出晶体结构。常用的方法有键序参数 BOOP【45】、键角分析 BAA【46】、共近邻分析 CNA【47】、中心对称参数 CSP【48】，其中 BOOP、BAA 和 CSP 方法从相邻原子位置导出序参数来表征不同晶体结构的特征，CNA 方法则根据原子邻域内的原子分布构建并分析邻居拓扑来得到参数表征不同晶体结构的特征。这些方法都试图通过原子的排列找到一些特征对其进行描述，从而实现对晶体结构的分类。虽然 CNA 方法适用于高度结晶的原子，但实际系统中热波动会干扰标准的固定截断半径的取值，因此，Stulowski 提出另一个自适应的 CNA 方法，即 A-CNA(Adaptive Common Neighbor Analysis)。

共近邻方法根据原子在空间中的邻居分布情况实现对晶体结构的识别，因准确性很高，在晶体结构识别领域内颇受欢迎。

1) CNA 方法：首先定义截断半径 R ，然后根据截断半径为每个原子生成一个邻居列表，即，与中心原子之间的距离小于截断半径的原子都被划为中心原子的邻居。然后，根据邻居列表为每个原子计算邻居节点个数、连接边的个数、最长路径三个特征，再将这三个特征组成的三元组作为原子的特征向量。最后，将

该特征向量作为识别标志用于晶体结构的分类。

2) A-CNA 方法: 根据每个原子与其所有一阶邻居节点之间的平均距离计算一个截断半径, 因此每个原子的截断半径都有所不同, 与其一阶邻居相关。A-CNA 主要用于识别多相金属, 其中每个原子具有很多的邻居且离得很近。在低密度液体或气体相中, 具有低一阶邻居(配位数)会分配很高的截断半径, 这会干扰 A-CNA 的识别, 因为在稀疏区域中原子可能被认为与他们距离更加远的邻居具有虚假的高连通性。

虽然 CNA 和 A-CNA 方法在识别精度很高且操作方便, 但是因为该方法只能根据其领域内原子分布判断结构类型, 因此受到一定的限制, 导致在材料表面的晶体结构无法被识别出来。另外, 这两种方法缺乏对中间结构类型进行概率上的判断。本工作针对 CNA 和 A-CNA 方法的缺点进行改进, 不仅可以识别材料表面的晶体结构, 而且以概率形式输出晶体结构结果, 对晶体结构研究具有重大意义。

2.2 基于机器学习的分类方法

通过分析原子在空间中的分布信息实现对晶体结构的分类, 即晶体结构识别。对于小样本和数据分布不均匀的晶体结构数据, 利用有监督的机器学习方法进行模型训练会导致效果极差。因此, 本研究首先针对这样的数据进行预处理, 弥补数据不足并且消除数据不平衡的问题, 然后再利用机器学习方法实现晶体结构的分类。该方法主要包括以下四个步骤: 第一步根据原子的空间分布特点, 生成原子数据, 从而弥补数据不足并消除数据不平衡的问题。第二步根据原子在空间中的分布, 设计合理的特征表示, 并转化为特征向量的形式。第三步利用互信息对特征进行选择, 一方面选择出对晶体结构识别有用的特征, 另一方面降低特征维度可以加快模型的训练。最后一步训练多层感知机模型实现对晶体结构的分类。

2.2.1 数据预处理

在拉伸材料使其形变过程中, 获取到的模拟数据是有限的, 且不同材料拉伸形变得到的晶体结构数量分布可能非常不均匀。在本章中使用 Cu 拉伸形变过程

中的模拟数据，根据 A-CNA 方法对其晶体结构进行识别，结果统计如表 2.1 所示。

表 2.1 A-CNA 方法识别结果统计

晶体结构	数据集	
	原子数量	原子数量占比
FCC 结构	34590	0.598
HCP 结构	8085	0.140
BCC 结构	213	0.004
未知结构	14912	0.258

从表 2.1 中我们可以看到，不同晶体结构下的原子数量具有很大的差异性，与 FCC 和 HCP 结构相比，BCC 结构的原子数量过少，因此，数据分布极其不平衡。而且 BCC 结构下的原子数量只有 200 多个，对于机器学习任务是极其少的，是典型的小样本问题。

针对数据的问题，我们利用原子在空间中的分布以及在拉伸形变中原子分布的变化特征，对原子空间随机扰动从而生成新的数据样本。原子空间是三维，每个原子可以在上、下、左、右、前和后六个方向上移动，因此，对原子数据在这六个方向上都做一定的随机扰动，便可以获得新的样本数据以解决数据不平衡以及小样本数据的问题。设置扰动概率 P_a 表示空间中每个原子有多大的概率扰动，若 $P_a = 0$ 表示空间原子未被扰动， $P_a = 1$ ，表示空间中所以原子被扰动了，也就是说， P_a 越大表示扰动概率越大，即空间中移动的原子数量越多。在本章研究中将 P_a 设置为 0.9。

2.2.2 特征表示

在基于机器学习方法中，特征是机器学习模型得以训练成功的关键因素之一，好的特征才能很好的体现数据的特点，使得模型获得到较高的准确率。在不同的研究领域内，针对不同的数据存在不同的特征表示方法，例如，在自然语言处理

领域内有词袋模型【49】、one-hot 编码等特征表示方法。近年来，随着深度学习的发展，因卷积神经网络（CNN）可以自动的从图像中提取到深层次的特征，现在基于卷积神经网络的特征表示方法已经成为了主流和热门。以上这些特征表示方法都无需人工，其输入形式都是向量、矩阵或者张量等。对于晶体结构数据，从空间中原子分布信息得到有用的晶体结构特征是本章研究的重点之一，常用的晶体结构识别方法根据原子空间坐标定义一些特征值用于原子特征表示【50-52】，这些特征表示不仅不够直观、难以理解，而且需要人工定义参数，例如截断半径、高斯参数等。

对于晶体结构中的原子，我们需要根据原子邻域信息对晶体结构进行分类。直接将原子的空间坐标信息输入到神经网络系统中，虽然可以实现对晶体结构的分类，但是每个原子局部空间中会移动，其坐标信息过于庞大，耗时巨大且难以训练。实际上，每个原子在晶体结构中的位置信息是相对的，因此，可为每个原子提取有效的局部邻域信息。在本研究中，基于中心原子的局部邻域内的原子分布特点对中心原子进行特征描述，定义为 LSC(Local Spatial Charctersistic)特征，LSC 特征描述了原子的三维信息。不同于 CNA 等方法，LSC 将坐标特征映射成为一维向量的形式。为了简单的理解，首先对二维空间中的 LSC 特征进行简单描述。如图 2.4 中的 (a) 所示，用实心圆来表示原子。这些实心圆可分为两种类型的结构，如图 2.4 (a) 中的黑框 A 和 B 标记所示。将二维空间离散化，从而形成一个由许多单元网格构成了离散网格空间，如图 2.4 (b) 所示。为了数值化原子在这个离散网格空间内的分布情况，则统计每个网格内是否存在原子。如果实心圆（原子）的中心位于一个单元格中，则该单元格标记为 1；否则，该单元格标记为 0。图 2.4 (c) 中标记的黑色部分作为一个例子，其中红色实心圆表示当前关心的原子，也称为中心原子，其领域是黑色边框所围成的正方形内部。中心原子的 LSC 矩阵描述如下：

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}, \quad (2.1)$$

其中，矩阵中的 0, 1 分别表示离散单元格内不存在原子和存在原子。三维空间

中原子的 LSC 获取过程与二维空间中原子的 LSC 相似。首先，为了离散化三维空间，创建了许多立方体以离散原子所在的三维空间。其次，判断每一个立方体内是否包含原子的中心，若包含则就将其标记为 1，否则将其标记为 0。这里立方体的长度 L 需要认为设置为一个合理的数值，如果数值太小，则无法很好地描述原子之间的相对位置；如果数值太大，则会导致数据的维度非常大，这将花费大量时间进行计算。为了获得合理的离散三维空间，根据经验每个立方体的长度 L 定义如下：

$$L = \frac{1}{6}\bar{D} + \frac{1}{2}, \quad (2.2)$$

其中 \bar{D} 是原子间的平均最短距离。三维空间示例如图 2.4 (d) - (f) 所示，红色球表示当前关心原子，也就是中心原子，蓝色球表示其局部空间内的邻居原子。标记的红色立方体和蓝色立方体分别表示中心原子及其相邻原子在 $5L * 5L * 7L$ 子空间中的位置。在将原子的 LSC 输入到神经网络之前，我们需要将其转换为向量形式，即将矩阵化的 LSC 的多个列向量取出合并为一个向量。为了简化提取过程，使用 $gLSC(i; H, l, w, h)$ 表示提取 LSC 的过程，该过程包括两个子过程：第一，将原子的三维空间离散化，每个单元用一个立方体表示，立方体的长度可通过公式 (2.2) 计算出；第二，以每个原子作为中心原子获取其局部原子分布的 LSC 矩阵表示，并将其转化为一维向量形式。我们将原子 i 的长度为 F 的 LSC 表示为 R_i^F ，描述如下：

$$R_i^F = g_{LSC}(i; H, l, w, h), \quad (2.3)$$

其中 H 表示离散化的空间，即图 2.4 (e) 中的 3D 网格； l 、 w 、 h 分别是原子附近局部空间的长度、宽度以及高度。

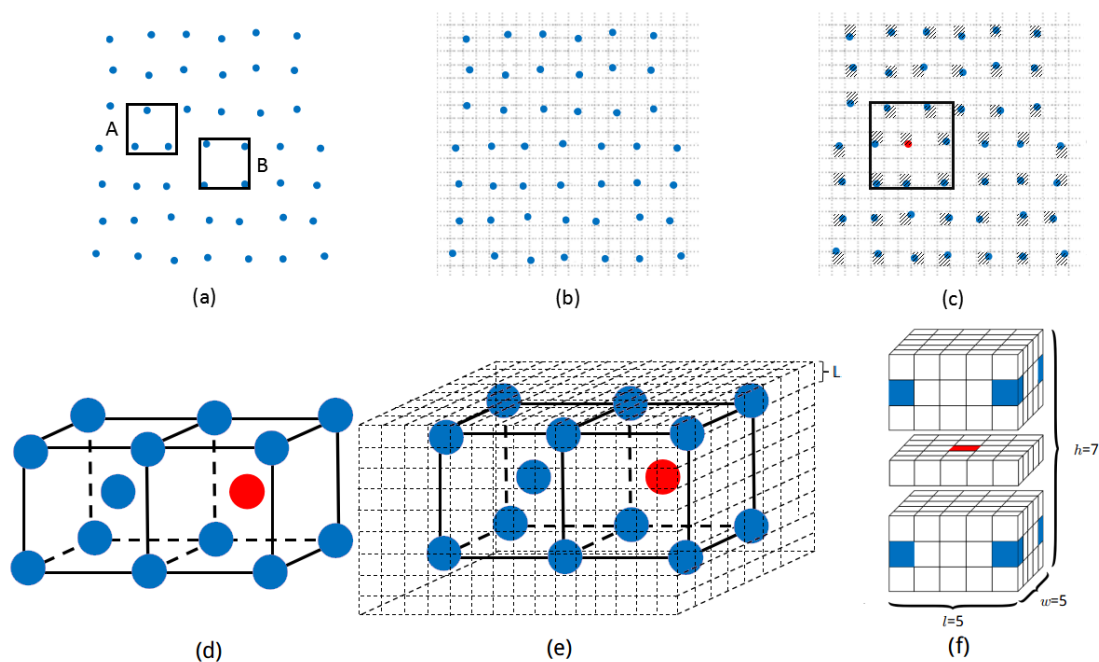


图 2.4 LSC 特征提取过程。(a) 2D 空间下的原子分布；(b) 2D 空间离散化；(c) 黑色框内的红色原子的 LSC；(d) 3D 空间下的原子分布；3D 孔家离散化；(e) 3D 孔家离散化；(f) 红色原子的 LSC。

2.2.3 特征选择

对于神经网络模型，如果输入节点太多的话可能会花费很长的模型训练时间和大量的计算内存，因此，数据需要被简化或者降维。例如，在将分辨率大图像输入到网络框架之前，可以先将其拆分为分辨率小的图像，使用 FCNN 将大小为 $7000 * 8000$ 的素材图像裁剪为 $1000 * 1000$ 【42】。在本研究中，需要将向量 R_i^F 作为输入数据放到多层感知机神经网络模型中以识别晶体结构。在本模型中 R_i^F 的维度是 13145，这造成了占用过多的内存和运行时间，需要对其降维，但是如果 R_i^F 降维后维度太小，则无法准确的描述局部原子空间分布特征信息。因此需要进行特征选择【53-55】，对于输入数据 D ，包含具有 M 个特征维度的 N 个数据样本及其标签 Y ，特征选择是从 M 个维度的特征中找到重要的 m 个子特征，通过调整 m ，使其较高的准确性和精度，又可节省存储空间和计算时间。

特征选择不同于特征降维，例如 PCA【56-57】、Isomap【58】和 LLE【59】等特征降维方法都难以从原始特征中选择重要的子特征。从数据中选择重要特征的最受欢迎的方法之一是特征选择的最大相关性 (Max-Relevance)【60-61】，即

最接近目标类别的特征。互信息 (Mutual Information, MI) 是选择关联特征的方法, 度量了两个随机变量的之间的关联性, 且成功应用于许多领域, 包括一些性能良好的系统【62-63】。

如公式 (2.3) 所示, 一个原子周围的立方体数量被设置为 $M = l * w * h$ 。互信息用于从 M 维的特征中选择的 m 维特征。如图 2.5 (a) 所示, N 原子数据用于训练, 记为 $A_i, i = 1, \dots, N$; $X^j, j = 1, \dots, M$ 表示原子的局部空间分布特征, 取值为 0 或 1; 标签 Y 描述了这些原子的所属的晶体结构类型, 其值 1, 2, 3 分别表示 FCC, HCP 和 BCC 三种晶体结构, 因此, 互信息 $I(X^j, Y)$ 可定义为:

$$I(X^j, Y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log \left(\frac{p(x, y)}{p(x)p(y)} \right), \quad (2.4)$$

其中 $p(x, y)$ 是联合概率分布, $p(x), p(y)$ 是边缘概率分布。图 2.5 (b) 中显示了使用公式 (2.4) 计算后的示例, 并以降序给出特征的索引及其 MI 值。利用最大相关性原理【64】, 选择前 m 个具有较大互信息的特征。如果 m 太小, 将会有严重信息损失; 如果 m 太大, 则原子的 LSC 对于神经网络来说依然通过庞大, 这里 m 选择为 2000。特征选择后的特征记为 m-LSC。

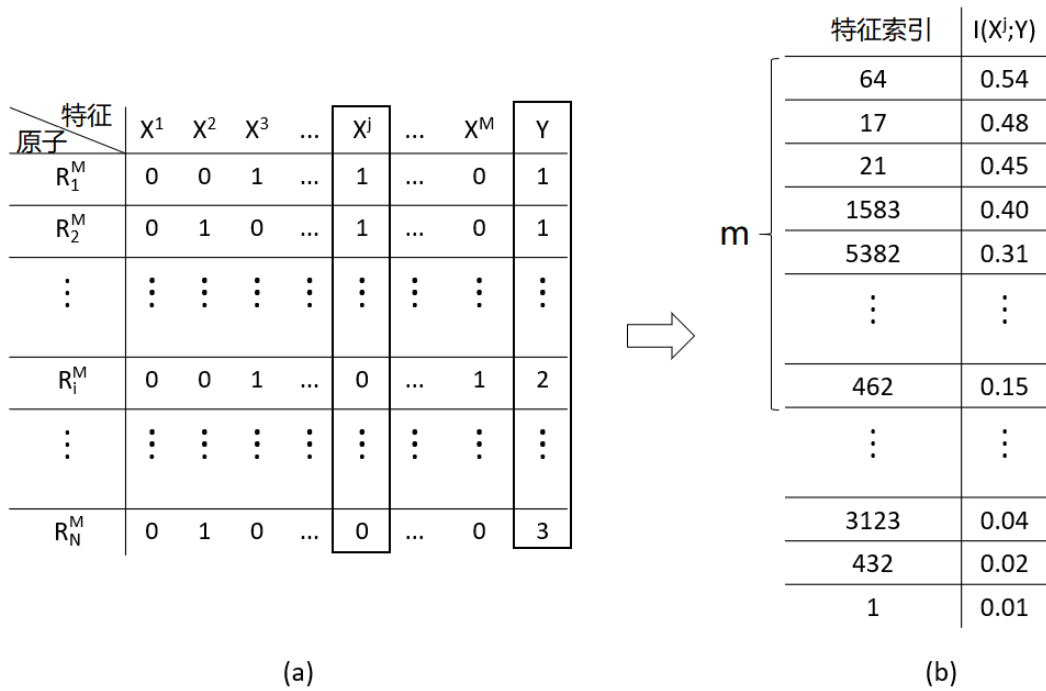


图 2.5 使用互信息方法选择重要特征过程。左表是一个虚构的训练数据，其中包含 N 个原子。每个原子具有包含在标签数据 Y 和 M 个特征。右表是不同特征对应的互信息值，并选取前 m 个特征。

2.2.4 晶体结构识别

对于上一节中获得的降维后的晶体结构特征向量 m -LSC，可以通过多层感知神经网络模型进行分类和识别。多层感知机是人工神经网络（Artificial Neural Network, ANN）的经典模型之一【65】，包含了输入层，隐藏层和输出层，结构如图 2.6 所示。每个层包含不同数量的神经元，输入层中神经元的数量等于 m -LSC 中的特征维度，输出层是由对应的三个节点表示的最终的分类结果三种晶体结构，在输入层和输出层之间有 5 个隐藏层，其节点数分别为 20、32、32、64 和 128，多层感知机可以使用多种激活函数，例如 logistic 和 tanh【66】，此处我们采用 ReLu【67】作为激活函数，因此，第 l 层中的任何一个神经元节点 $z_j^{(l)}$ 都与上一层，即 $l - 1$ 层中的所有节点相连，公式如下所示：

$$z_j^{(l)} = \max\left(0, \sum_i z_i^{(l-1)} w_{ij}^{(l-1)} + b_j^{(l-1)}\right), \quad (2.5)$$

其中 $z_i^{(l-1)}$ 是 $l-1$ 层的第 i 个神经元； $w_{ij}^{(l-1)}$ 是从 $l-1$ 层的第 i 个神经元到 l 层的第 j 个神经元的连接权重； $b_j^{(l-1)}$ 是 $l-1$ 中的偏置。在输出层中使用 softmax 函数， $O_k^{(l)}$ 用于给出输入矢量在不同类别 k 上的概率， $k = 1,2,3$ ，公式如下所示：

$$O_k^{(l)} = \frac{e^{\sum_j z_j^{(l-1)} w_{jk}^{(l-1)}}}{\sum_{i=1}^m e^{\sum_j z_j^{(l-1)} w_{jk}^{(l-1)}}}, \quad (2.6)$$

其中 $O_k^{(l)}$ 以概率的形式输出(FCC, HCP, BCC)的预测值。 $O_k^{(l)}$ 的最大值对应于最可能的晶体结构类型，例如，如果向量为[0.1, 0.1, 0.8]，则 BCC 结构的预测概率最大，这个晶体结构为 BCC 结构。将概率向量简化为 one-hot 格式，则晶体结构（FCC、HCP 和 BCC）可分别表示为[1, 0, 0]、[0, 1, 0]和[0, 0, 1]。

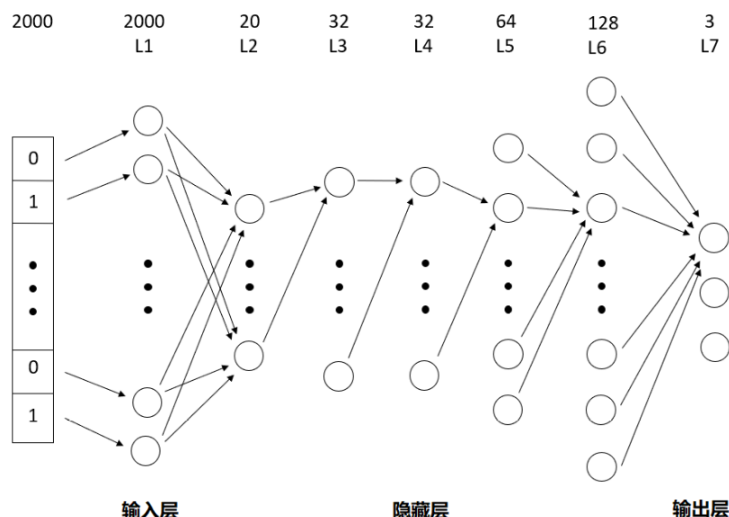


图 2.6 多层感知神经网络模型。

为了使得模型收敛，采用反向传播算法来优化参数，这里采用 Keras 模型中使用了 ADAM 梯度优化器来优化多层感知机模型，其动量和权重衰减分别定义为 10^6 、 10^5 。另一方面，L2 正则化【68】和 Dropout【69】被使用来防止过拟合。L2 正则化被添加到损失函数中，以使权重更接近原点，这里 L2 权重设置为 10^{-3} 。Dropout 以指定的概率使得神经元失去作用，这里设置为 0.5。

2.3 实验与讨论

2.3.1 实验环境和评估指标

在实验中使用 2.0 GHz * 4 和 GTX 760 的 i7 处理器的计算机以及 python 2.7 和 ovito 2.9 开发工具。MLP NN 是使用 Keras 库实现的，它是用 python 编写的高级神经网络库，该库使我们可以使用 CPU 和 GPU 轻松运行深度学习方法。在第 2.2.2 节中，根据经验和实验结果，将 l , w , h 设置为 $25L$, $25L$, $21L$ ，其中 L 由公式 (2.2) 计算。

在机器学习算法中，需要对预测的结果加以评估。预测结果分为四种，实际为真被预测为真、实际为假被预测为假、实际为假被预测为真和实际为真被预测为假，分别表示为真正例 (TP)、真反例 (TN)、假正例 (FP) 和假反例 (FN)。常用的评估指标有准确率、精确率、召回率、 F_{score} 、ROC、AUC 和混淆矩阵。公式如下：

准确率：表示被分类器正确划分的样本数量占所有样本数量的比值

$$Accuracy = \frac{TP + TN}{TP + FN + FP + FN}, \quad (2.7)$$

精确率：表示预测为正的样本数据中预测正确的样本所占比重

$$Precision = \frac{TP}{TP + FP}, \quad (2.8)$$

召回率：表示实际的正样本中被正确预测的样本所占比重

$$Recall = \frac{TP}{TP + FN}, \quad (2.9)$$

F_{score} ：将精确率和召回率进行结合，并通过参数 β 调节精确率和召回率的占比，公式描述如下：

$$F_{score} = (1 + \beta^2) \cdot \frac{precision \cdot Recall}{\beta \cdot Precision + Recall}, \quad (2.10)$$

ROC (Receiver operating characteristic) 曲线：这是一种功率度量方法，用于说明具有不同判别阈值的方法的性能。将该曲线绘制在一个平面中，该平面的 x 轴表示真正例 (TP)， y 轴表示假正例 (FP)，即 TPR 和 FPR ，描述如下：

$$TPR = \frac{TP}{TP + FN}, \quad (2.11)$$

$$FPR = \frac{FP}{TN + FP}, \quad (2.12)$$

AUC: 是 ROC 曲线下面积。由于 ROC 曲线无法清楚地表明哪种分类方法效果更好, 因此将 AUC 用作评估标准。通常, AUC 的值越大, 分类器越好。

混淆矩阵: 是由真正例 (TP)、真反例 (TN)、假正例 (FP) 和假反例 (FN) 构成的矩阵形式, 其中列和行分别代表预测结果和实际结果。

以上的准确率、精确度、召回率、ROC 曲线、AUC 等评估指标常用于机器学习二分类任务中, 而本工作的分类任务中有三个类别, 即 FCC, HCP 和 BCC。为了迎合这样的多分类任务, 一个加权平均的评估指标如下所示:

$$W_{\Delta} = \frac{F_{\Delta} \times N_F + H_{\Delta} \times N_H + B_{\Delta} \times N_B}{N_F + N_H + N_B}, \quad (2.13)$$

其中, Δ 表示准确率、精确率、召回率、 F_{score} 或者 AUC; F_{Δ} 、 H_{Δ} 、 B_{Δ} 分别为对应的 FCC、HCP、BCC 下的评估指标值 (准确率、精确率、召回率、 F_{score} 、AUC); N_F 、 N_H 、 N_B 分别表示 FCC、HCP、BCC 样本数量。

为了说明多少比例的晶体结构能够被识别出来, 我们定义了识别率, 公式如下:

$$P_r = \frac{N_r}{N}, \quad (2.14)$$

其中 N_r 和 N 分别表示了被识别出的原子样本数量和原子样本总数。

2.3.2 特征选择的比较分析

利用互信息 (MI) 进行特征选择是极其重要的一步。我们比较了混淆矩阵、精确率、召回率、 F_{score} 、AUC 以及 ROC 曲线, 用于评估在相同测试数据, 不使用互信息和使用互信息的特征训练的多层感知机模型在预测是的差异性。表 2.2 和表 2.3 中给出了混淆矩阵的比较结果, 其他评估指标的详细结构显示在表 2.4 和表 2.5 中, 可以发现使用互信息进行特征选择, 既可以降低特征维度, 同时又可以保持识别精度。为了更直观的评估所提出的模型的效率, 分别绘制了 ROC 曲线, 计算宏平均和微平均, 使用互信息和没有使用互信息的结果分别如图 2.7 和图 2.8 所示, 可以发现, 微平均和宏平均在使用 MI 和没有 MI 时基本相同。

根据上述指标的结果, 使用互信息和不使用互信息的模型泛化能力几乎相同,

然而，使用互信息对输入特征的维度可以大大减少了数量，加快了模型的训练。在对同一测试集进行测试时，使用互信息的模型预测时间为 1.23s，而不使用的需要 1.97s，详细的比较如表 2.6 所示，因此，训练精度和测试精度的两个种模式几乎相同，但模型所花费时间上使用互信息更占优势。

表 2.2 不使用互信进行特征选择的混淆矩阵。

		实际类别		
		FCC	HCP	BCC
预测类别	FCC	26829	8	5
	HCP	37	7475	19
	BCC	0	27	74

表 2.3 使用互信息进行特征选择的混淆矩阵。

		实际类别		
		FCC	HCP	BCC
预测类别	FCC	26836	5	1
	HCP	8	7504	19
	BCC	0	19	82

表 2.4 没有使用互信息的模型在测试集上的评价结果。

类别	精确率	召回率	F-Score	AUC
FCC	0.99862	0.99951	0.99906	0.99998
HCP	0.99533	0.99257	0.99394	0.99974
BCC	0.75510	0.73267	0.74371	0.97924
权值平均	0.99718	0.99721	0.99719	0.99986

表 2.5 使用互信息的模型在测试集上的评价结果。

类别	精确率	召回率	F_{score}	AUC
FCC	0.99971	0.99977	0.99973	0.99999
HCP	0.99681	0.99641	0.99661	0.99989
BCC	0.80392	0.81188	0.80788	0.99691
权值平均	0.99851	0.99848	0.99848	0.99995

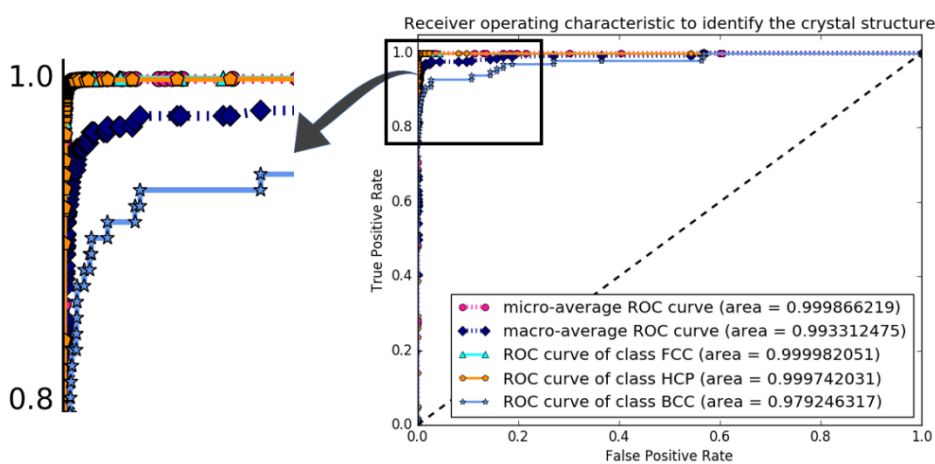


图 2.7 没有使用互信息降维的 ROC 曲线。

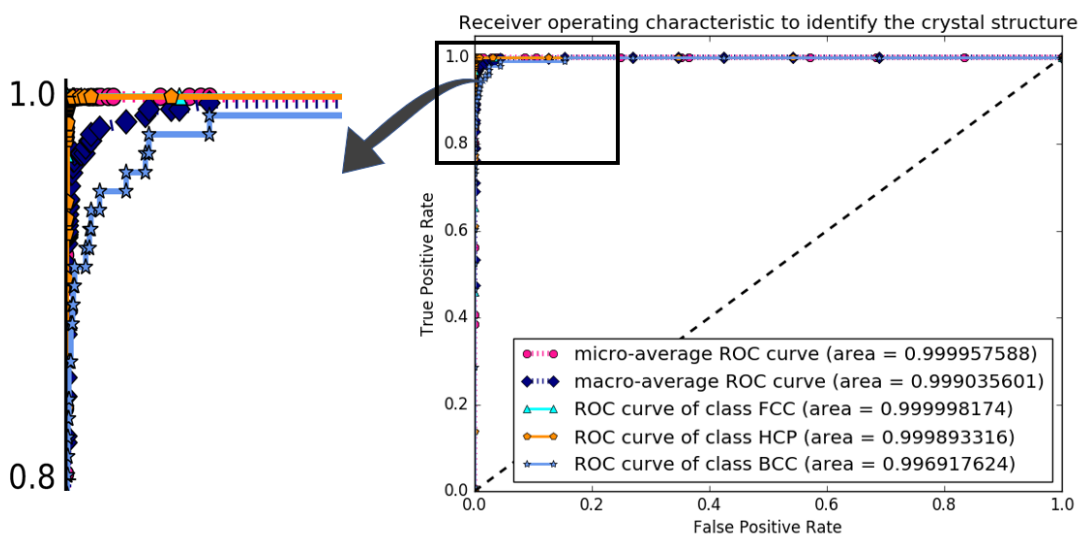


图 2.8 使用互信息降维的 ROC 曲线。

表 2.6 使用互信息和不使用互信息的比较。

	特征数	训练准确率	测试准确率	速度 (s)
未使用互信息	13125	0.99979	0.99721	1.93
使用互信息	2000	0.99996	0.99649	1.23

2.3.3 与共近邻方法比较分析

在多层感知机模型中，输出层采用 softmax 函数实现对晶体结构的分类，即通过 softmax 函数给出某一样本所属 FCC、HCC 和 BCC 三种晶体结构的概率。通过实验计算输出，我们发现两种有趣的现象，结果显示在表 2.7 中。第一种，大部分的原子晶体结构是以高概率的形式预测，也就是说 softmax 给出的三种类别的概率值中其中有一个值非常大接近 1，而另外两个类别的概率接近于 0。例如，表 2.7 (a) 中 softmax 预测值 (0.999879, 0.000001, 0.000119) 的结果为 FCC 结构。第二种，晶体结构是以一种相对较低的概率被预测。例如，表 2.7 (b) 中 softmax 预测值 (0.579238, 0.420573, 0.000189) 的结果依然是 FCC，但是却以 57% 的概率。对于这种低概率的预测的结果，我们认为这些结构可能是处在晶体结构转换过程中的中间状态。这两种状态的可视化结果，如图 2.9 所示。在图 (a) 中为分类后的结果，通过三种颜色代表三种晶体结构，分别以 (1, 0, 0)、(0, 1, 0)、(0, 0, 1) 代表红色、蓝色和绿。图 2.9 (b) 为 softmax 输出的概率结果，若输出结果接近于 (1, 0, 0) 则越接近红色。图中使用黑色箭头标记具有差异性的一些结果，通过结果的可以发现，大部分晶体结构都是以高概率预测的，只有少部分原子以低概率预测。这少部分低概率预测的晶体结构或许对新结构的发现具有非常大的意义。

表 2.7 晶体结构的两种预测现象示例。

		高概率预测			低概率预测		
		FCC	HCP	BCC	FCC	HCP	BCC
预 测 类 别	FCC	0.999879	0.000001	0.000119	0.579238	0.420573	0.000189
	HCP	0.000346	0.999654	0.0	0.439824	0.556608	0.003569
	BCC	0.000113	0.000158	0.999729	0.497449	0.000001	0.502550

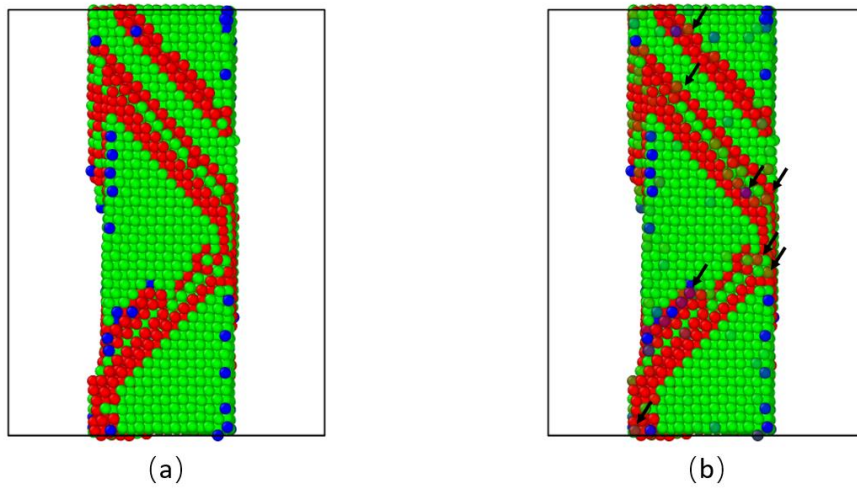


图 2.9 概率结果与分类结果可视化对比。

在本章工作中，将 A-CNA 的识别出的晶体结构作为标签训练我们的模型。为了比较我们与共近邻结构的差异性，我们分别比了识别率，准确率，时间，结果如表 2.8 所示。结果表明，我们方法的准确率高达99%，几乎与 CNA 和 A-CNA 相同。此外，我们的方法可以识别所有晶体结构，包括 CNA 和 A-CNA 方法不能识别出的表面的晶体结构。因此，本方法的识别率为100%，但是，我们方法的执行时间相较于 CNA 和 A-CNA 是慢的。

另一方面，为了与 CNA 方法和 A-CNA 方法进行定性比较，图 2.10 显示了 CNA 方法，A-CNA 方法和我们的方法在模拟数据中的可视化结果，这是更直观的显示。显然，表面原子的晶体结构不能用 CNA 方法和 A-CNA 方法识别，而可以我们所提出的识别方法则可以，并且我们的方法可以给出概率预测的结果，

这对新晶体结构的发现具有极大的意义。

表 2.8 我们方法与 CNA、A-CNA 方法的比较。

	识别率	准确率	时间
CNA	73.04%	100%	0.21
A-CNA	73.22%	100%	0.25
未使用互信息	100%	99.721%	1.93
使用互信息	100%	99.849%	1.23

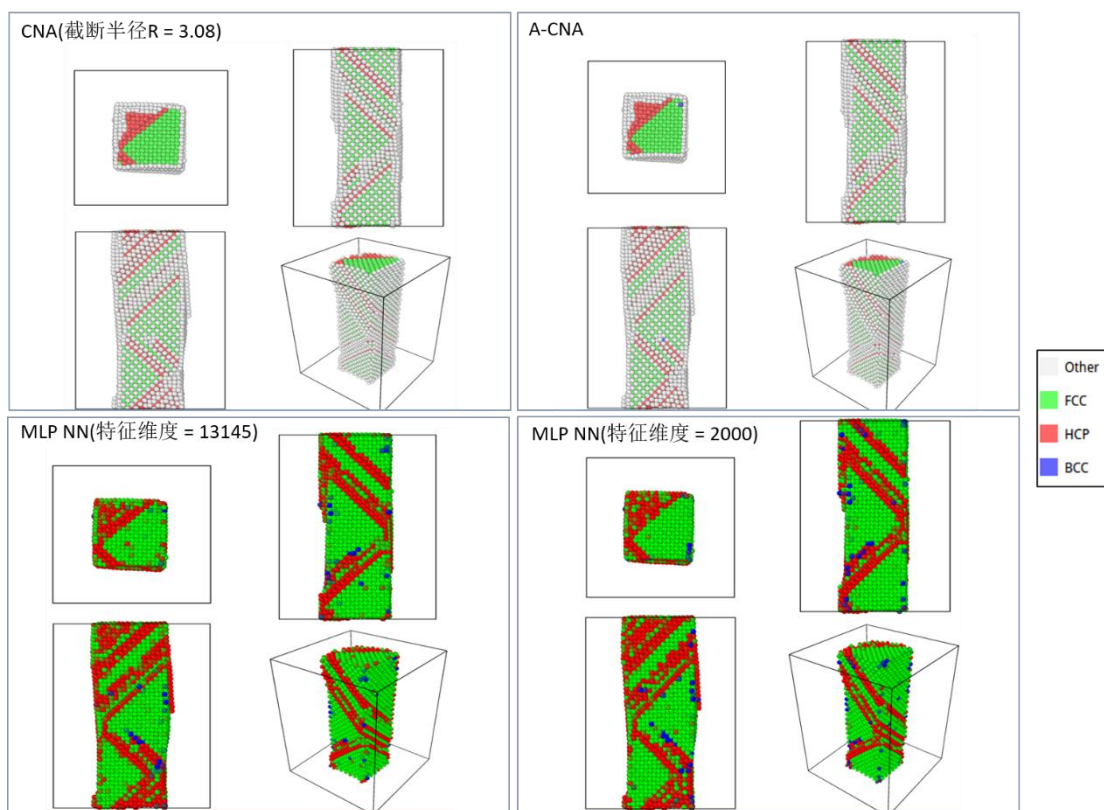


图 2.10 我们方法与 CNA、A-CNA 方法的可视化结果对比。

2.4 本章小结

本章提出了一种基于机器学习的晶体结构识别方法，该方法通过对样本数据扩增来解决小样本问题，并根据原子空间分布信息定义特征并进行特征选

择，最后训练多层感知机模型实现对晶体结构的识别。本研究方法不仅解决了小样本问题，而且可以识别出表面原子的晶体结构并获得每个晶体结构的预测概率，对晶体结构的研究以及新结构的发现具有重大意义。

然而本研究仍有一定的不足需要改进。一方面，由于 LSC 特征不具有特征旋转不变性，导致扩展性不强；另一方面，本方法的计算成本较为昂贵。因此，还需要进一步的优化和研究。

第三章 基于迁移学习的 AFM 图像分割识别方法

近些年，纳米结构材料因具有较好的性能，吸引了广大研究人员的关注。为了设计和制备纳米结构材料，常常需要使用原子力显微镜（Atomic Force Microscope, AFM）等现代设备获取清晰的纳米结构物体的 AFM 图像数据。然而在研究中获取到的 AFM 图像数据量往往很少，因此本研究提出利用迁移学习实现对少量纳米材料结构的 AFM 图像的识别，不仅能有效地区分重叠的纳米物体而且能解决纳米物体中的小样本问题。本研究工作将有助于加深度学习方法在纳米结构科学技术领域内的发展以及加速材料小样本研究的进程。本章内容已投稿至 Pattern Recognition 期刊。

3.1 纳米材料 AFM 图像背景介绍

近年来，利用纳米技术制造了许多新型材料，并应用于诸多领域，例如，石墨烯材料作为一种用于环境的新型材料而受到了越来越多的关注【70】。无机半导体的异质结构也受到广大研究者的关注【71】。【72】概述了 TMP（Transition Metal Phosphides）纳米材料作为催化剂的最新进展，该催化剂用于高活性和稳定性的制氢。【73】描述了纳米结构的 TiO₂ 基生物材料在骨组织工程，血管内支架和生物传感器方面的最新研究进展。【74-75】中总结了 ZnO 和 GLNs（Graphene-like 2D Layered Nanomaterials）纳米材料在生物医学应用中的用途，例如生物医学成像，药物输送和癌症治疗。

要设计和构建良好的纳米结构，需要使用原子力显微镜（AFM）等现代设备直接对其进行观察，并通过处理和分析获取的 AFM 图像来研究纳米材料的性能。AFM 自带了简单的图像处理软件，但是不同材料得到的 AFM 图像差异性很大，很难有统一的方法处理 AFM 图像。此外，AFM 图像本身数据自带很多噪声，且目标物体边界不清晰，一些纳米材料图像数据有限，液体中的纳米物体经常重叠，这些问题都给 AFM 图像处理和分析工作带来了巨大的挑战和难题。

为了解决 AFM 图像中的噪声问题，一些常规技术【76】（例如形态转换和阈值去噪）已被应用来消除白点和水印但它们可能会损坏 AFM 图像中目标物体的

形状。利用传统的噪声过滤技术【77】来去除 AFM 图像中的多种噪声，往往无法消除小区域的某些噪音，有时，目标物体的形状也可能会改变。

专家和学者们已对 AFM 图像中目标重叠问题进行了很多研究，例如分水岭方法及其变体已被用于分割生物细胞微颗粒场中的重叠物体和区域【78-79】。Wang 等人【38】利用支持向量机 (SVM) 实现对 AFM 图像中的重叠物体分割。对于 AFM 图像中的微纳米物体识别，已经有一些类似的工作，例如 Park 等人【80】对显微镜图像中的纳米颗粒进行了分割和进一步识别。机器学习方法也已用于纳米结构的识别工作【81】。近年来，深度学习方法已成为最流行的图像处理方法之一，但是它们需要大量的大数据来训练模型【82】。在大多数情况下，很难收集到如此多的注释数据。

在本章中所使用的数据是 DNA 分子机器人的 AFM 图像。DNA 分子机器人的几何结构如图 3.1 所示，包含三种类型结构：逆平行、交叉和平行【83】。由于 DNA 分子机器人是由 DNA 大分子制成并存在于液体中，因此，他们的形状是多种多样的，既扭曲变形又彼此重叠，这严重阻碍了 AFM 图像处理。

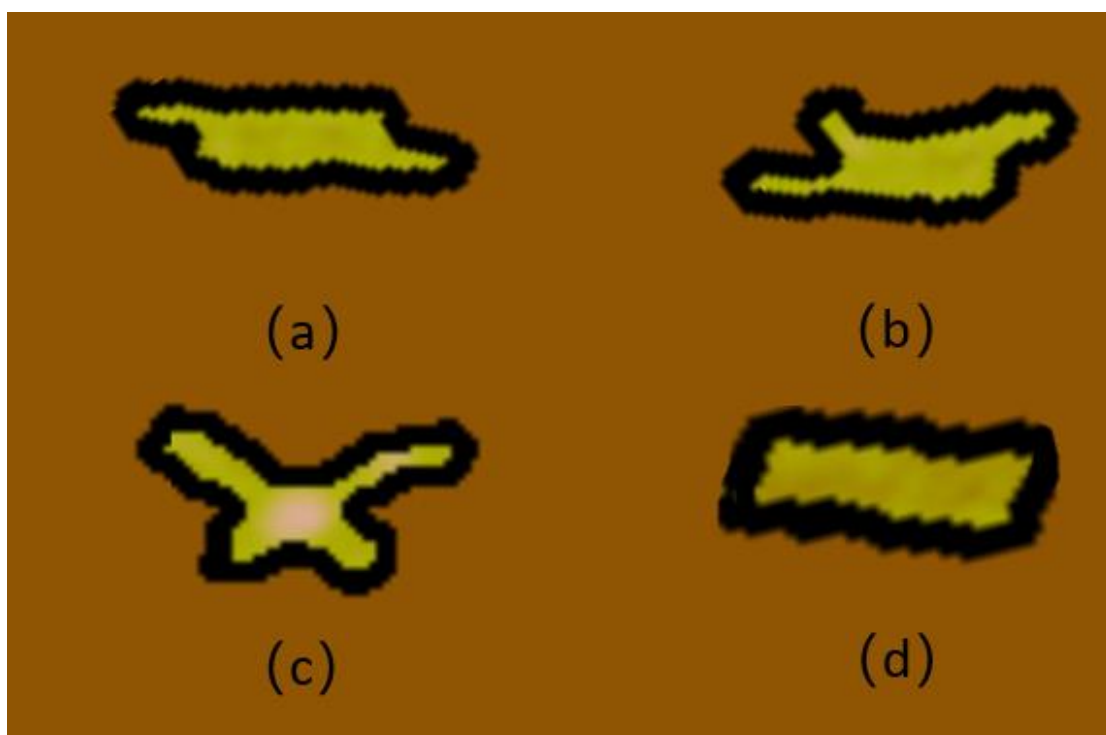


图 3.1 四种理想情况下的 DNA 分子机器人几何形状；(a)和(b)表示逆平行类型；(c)表示交叉类型；(d)表示平行类型。

3.2 图像识别方法

针对小样本环境下的 AFM 图像处理问题，本研究分为三个步骤，分别是预处理，分割和识别。在预处理阶段，利用高斯拉普拉斯算子（LoG）去除 AFM 图像中的噪声点；在分割阶段，改进分水岭分割方法实现 AFM 图像中重叠目标的分割；在识别阶段利用结合迁移学习的卷积神经网络（CNN）方法完成对目标物体的识别。

3.2.1 AFM 图像预处理

AFM 图像包含了三维信息，这里将三维高度转换为彩色图像中的像素值，因此，本章随后图像处理中不考虑 AFM 图像的三维高度信息。AFM 图像中的噪声常常有斑点、水印、小噪声点、条纹噪声等噪声，而这些噪声对后续的 AFM 图像处理具有很大的影响，因此先对 AFM 图像进行一定预处理以去除噪声是非常有必要的。本章的方法是先对图像进行高斯平滑，然后采用高斯的拉普拉斯（Laplacian of the Gaussian）检测并去除 AFM 图像中的主要噪声，拉普拉斯算子是使用二阶导数的一种图像二维空间中的各向同性度量，可以通过尺度空间理论检测图像中形状的边缘【84】。这里我们使用 I 表示 AFM 图像，其为二维空间中的离散像素集合，灰度化后的 AFM 图像，记为 I_{gray} ，LoG 算子描述如下：

$$LoG = \frac{\partial^2}{\partial x^2} G_\sigma(x, y) + \frac{\partial^2}{\partial y^2} G_\sigma(x, y) = \frac{x^2 + y^2 - 2\sigma^2}{2\pi\sigma^6} e^{-(x^2+y^2)/2\sigma^2}, \quad (3.1)$$

其中 $G_\sigma(x, y)$ 高斯函数是关于方差 σ 的函数，不同的方差 σ 可得到不同的 LoG 算子。LoG 算子与图像都是关于空间坐标 (x, y) 的函数。因此，灰度图像的高斯拉普拉斯变换是将图像与 LoG 算子做卷积运算：

$$S(x, y) = LoG(x, y, \sigma) * I_{gray}(x, y), \quad (3.2)$$

$S(x, y)$ 为灰度图像的高斯拉普拉斯变换结果，该结果表示灰度图像每个像素点与 LoG 算子的相似性，即卷积结果越大，表示图像与 LoG 算子之间的相似性越高，反之卷积结果越小，则图像与 LoG 算子相似性越低。当斑点形状与 LoG 算子的形状一致时，则灰度图像的拉普拉斯响应最大，因此通过判断图像高斯拉普拉斯

变换后的极值点则可确认斑点，如图 3.2 (b) 中所示，每个斑点被看作是半径 $r = \sqrt{2}\sigma$ 的圆。但是，如果斑点的形状不规则，圆很难将其覆盖，而不能完全消除斑点噪声。为了解决这个问题，将检测到的圆中每一个点最为初始点，对其 8 邻域内的广度优先搜索，然后对检测出来的斑点使用背景像素进行替换，结果如图 3.2 (c) 所示。为了获得目标物体的形状，我们使用 OTSU 方法【85】将 AFM 图像分为前景和背景，成为二值图像。然后，采用腐蚀和膨胀操作除去一些小的噪声点，描述在图 3.2 (d) 和 (e)。最后，将膨胀处理后的白色部分作为背景，用原图像素平均值进行填充，以消除原图背景中的条纹噪声。同时，黑色部分作为目标形状，使用去除斑点后图像（即图 3.2 (c)）中相同位置的像素进行填充，从而实现图像颜色和形状的还原，结果如图 3.2 (f) 所示。

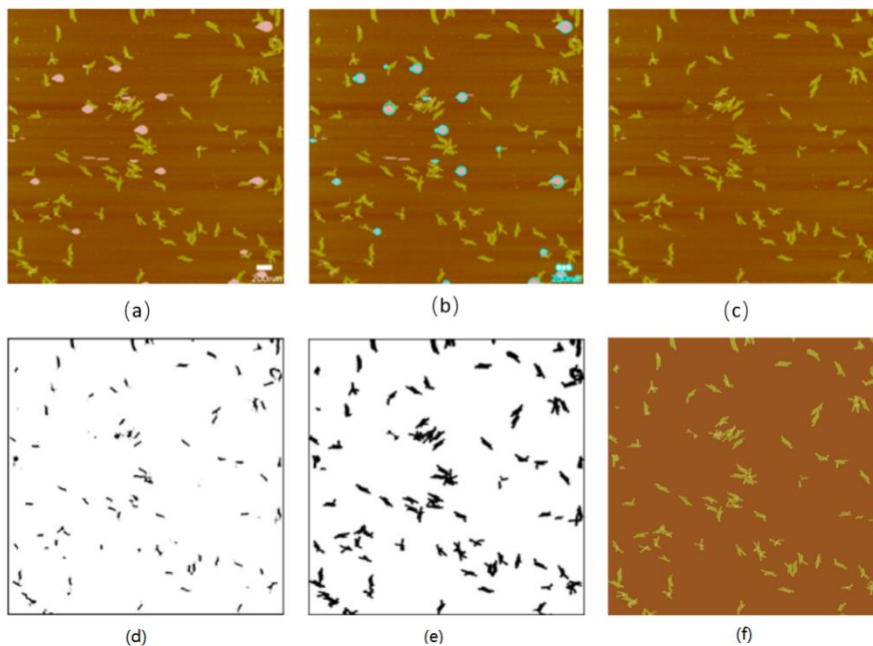


图 3.2 AFM 图像去噪声过程；(a) 含有噪声的 AFM 原图；(b) 检测出的斑点结果；(c) 背景填充斑点的结果；(d) 腐蚀后的结果；(e) 膨胀后的结果；(f) 去噪后的图像。

3.2.2 基于距离的分水岭分割方法

经过 3.3.2 章节处理，我们已经得到了去除噪声的 AFM 图像，使用符号 I_s 表示。但是图像 I_s 中的目标物体具有交叉重叠的问题。针对重叠目标物体问题，我

们改进了原有的分水岭方法【86】，提出了新的算法用于对图像中重叠目标进行分割。基于数学形态学理论的分水岭方法把图像看成是测地学中的拓扑地貌，像素值看成海拔高度，将像素的极小值区域作为集水区，而分水岭就是集水区的边界。分水岭方法对为微弱的边缘都具有很好的响应，但是图像中噪声，或者物体表面的一些微弱的像素变化也会被检测出来，很容易产生过分割现象。因此本研究提出一种基于距离的分水岭方法，不仅可以解决目标重叠问题，而且可以减少过分割问题，算法描述如表 3.1 所示。

基于距离的分水岭分割方法首先通过识别标记找出分割目标的主要部分，然后对主要目标进行扩展得到更精确的分割结果。为了标记极值区域需要知道每个目标区域内的像素值。DNA 分子机器人在 AFM 图像中的像素几乎相同，如果直接使用 AFM 图像中的像素值来区分区域，则无法分割重叠的两个 DNA 分子机器人形状，因此，将每个像素点到形状边缘的距离作为该点的像素值【87】。首先，将去噪后的彩色图像 I_s 转换为灰度图像，记为 I_b ，结果如图 3.3 (b) 所示。然后在 I_b 上执行距离变换得到 I_g ，使得距离边缘轮廓越近的像素点其像素值越大，结果如图 3.3 (c) 所示。为了消除奇异值的影响，需要将图像 I_g 中的像素值归一化到区间[0,1]内，并归一化后的 I_g 转化为二值图像，即设定一个阈值 $\xi(0 < \xi < 1)$ ，若像素值小于 ξ ，则为黑色，否则为白色。为了减少欠分割的情况，这里采用了开运算来断开一些小的连接，结果如图 3.3 (d) 所示。保留下来的黑色区域作为分水岭的标记部分。最后使用传统的分水岭方法来进一步扩展标记，结果如图 3.3 (e) 所示。

表 3.1 基于距离的分水岭分割方法的算法描述。

输入：图像 I_s ，阈值 ξ

输出：分割后的结果 I_{seg} ，面积参数 A_{max}

1. Function CutImageByScale (I_s, ξ)
 2. I_s 二值化后变成 I_b
 3. I_b 经过距离变换和归一化变成 I_g
 4. 获取 I_g 行数 rows 和列数 cols
 5. for i < rows do
 6. for j < cols then
 7. if I_g 中(i, j)处的像素值大于等于 ξ then
 8. I_g 中(i, j)处的像素值修改为 255
 9. else
 10. I_g 中(i, j)处的像素值修改为 0
 11. end if
 12. end for
 13. end for
 14. 对 I_g 执行开运算得到 I_{label}
 15. 根据 I_b , I_{label} 执行传统分水岭得到 I_{seg}
 16. 计算 A_{min} 和 A_{max}
 17. for $object_j \in I_{seg}$ do
 18. if $object_j$ 的面积小于等于 A_{min}
 19. 将 $object_j$ 进行融合
 20. 输出 I_{seg} 、 A_{max}
-

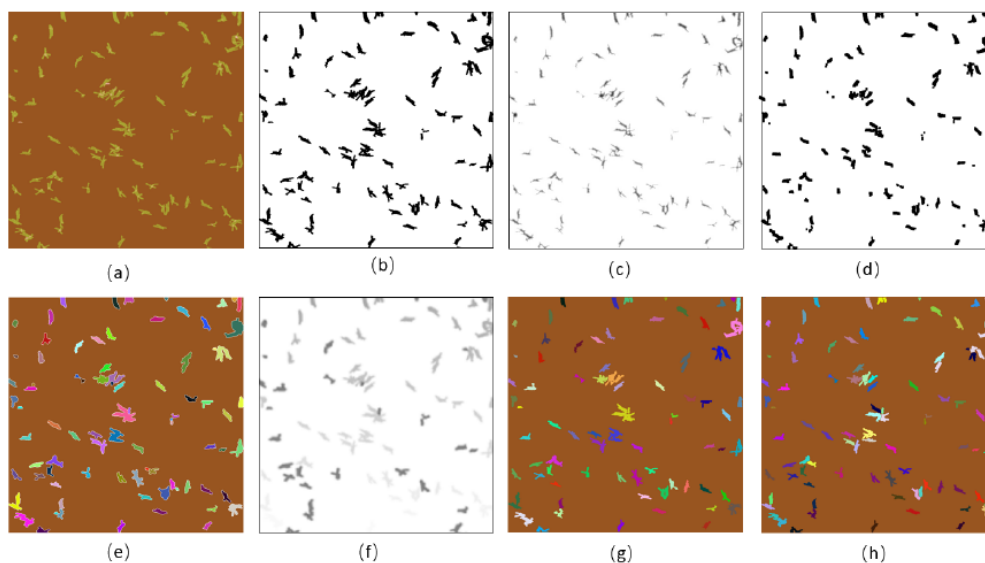


图 3.3 AFM 图像的分割流程；(a)原始图像；(b)二值化图像；(c)距离变换后图像；(d)分水岭的标记图像；(e)传统分水岭分割后结果；(f)需要进一步融合的区域结果；(g)区域融合后的结果；(h)多尺度的分水岭方法分割结果。

经过以上的处理经常会出现过分割或者欠分割的问题，因此需要进一步处理。计算初步分割后的每一个目标物体大小，若分割形状面积非常小的，则认为是目标过分割，需要将其进行融合；若分割形状面积非常大，则认为是目标欠分割，需要进一步分割。定义两个面积阈值，用于区分形状面积的大小，这两个阈值分别定义为 $A_{min} = \frac{1}{2}A$ 和 $A_{max} = \frac{3}{2}A$ ，其中 A 是分割后的所有分割形状面积的平均值。若分割形状面积小于 A_{min} 则表示过分割，若目标物体面积大于 A_{max} 则表示欠分割。当形状是过分割的时候，则考虑将其与邻居区域合并，具体内容参考算法 1。需要融合的结果如图 3.3 (f) 所示，其中黑色区域表示待合并的区域，融合后的结果如图 3.3 (g) 所示。

3.2.3 多尺度的分水岭分割方法

在基于距离的分水岭算法中，用于转化二值图像的阈值 ξ 对算法的分割结果起到至关重要的影响。如果 ξ 取值过小，会导致过分割，反之，如果 ξ 取值过大，会导致欠分割。此外，AFM 图像因分辨率的不同，目标大小存在差异，因此不

同的图像则需要不同的 ξ 。另外，在基于距离的分水岭算法中还存在尚待解决的欠分割问题。因此在基于距离的分水岭分割算法的基础上提出了多尺度的分水岭分割方法，算法的详细描述过程如表 3.2 所示。

表 3.2 多尺度分水岭分割方法的算法描述。

输入：图像 I_s ，初始阈值 ξ_0

输出：分割后的结果 I_{result}

1. Function CutAFMImage(I_s, ξ_0)
2. $I_{seg}, A_{max} \leftarrow \text{CutImageByScale}(I_s, \xi_0)$
3. $step \leftarrow 0.01, upper \leftarrow 0.6, I_{result} \leftarrow \emptyset$
4. for $S_i \in I_{seg}$ do
5. if S_i 面积大于等于 A_{max} , then
6. $\xi \leftarrow \xi_0 + step$
7. while ξ 小于等于 $upper$ do
8. $I_{seg}^{S_i} \leftarrow \text{CutImageByScale}(I_{S_i}, \xi)$
9. if $I_{seg}^{S_i}$ 数量大于等于 2 then
10. $I_{result} = I_{result} \cup I_{seg}^{S_i}$
11. else
12. $\xi \leftarrow \xi + step$
13. else
14. $I_{result} = I_{result} \cup S_i$
15. end if
16. end for
17. 输出 I_{result}

在算法 2 中，初始化阈值参数 $\xi_0 \in [0.1, 0.25]$ ，将其输入到算法 1 中则可获得一个初始分割结果 I_{seg} 和参数 A_{max} 。在 I_{seg} 中，若某一目标区域 S_i 的面积大于参数 A_{max} ，则表示该目标区域需要进一步的分割。在进一步分割的时候，新的参数 ξ 基于 ξ_0 迭代累加，累加步长使用 $step$ 表示。如果一个目标区域进一步分割后得到

两个以上区域或者参数 ξ 达到预设的最大值 $upper$ ，则停止分割。参数 $upper$ 的取值依赖于图像数据中的目标物体大小，如果图像数据中的目标物体偏小，则 $upper$ 应该设置的小一点，反之，如果目标物体较大，则需要调大 $upper$ 。针对 DNA 分子机器人的 AFM 数据中，步长 $step$ 设置为0.01， $upper$ 设置为 0.6。多尺度的分水岭方法的分割结果如图 3.3 (h) 所示。在图 3.4 中还给出了其他一些 AFM 图像分割结果，可以看出欠分割比较少，过分割居多，其主要原因是在 AFM 图像中的目标大小不一，形状各异。

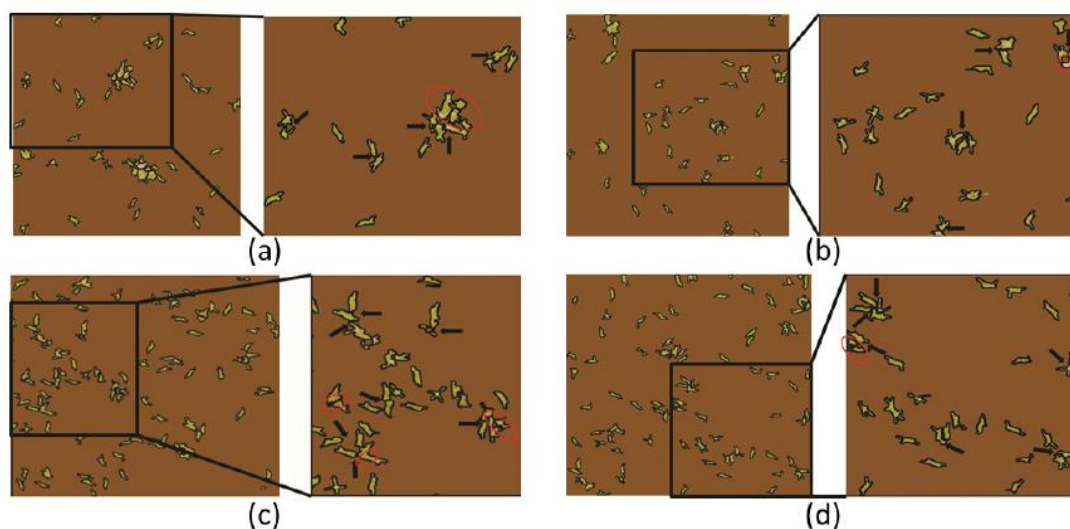


图 3.4 多尺度分水岭方法分割 AFM 图像的结果。

3.2.4 AFM 图像物体识别

分割完成后，需要对每个目标物体进行识别。一些常用的机器学习方法，例如支持向量机 (Support Vector Machine, SVM)、朴素贝叶斯、神经网络等方法可以用来识别目标，但这些方法往往需要人工定义一些特征用于模型的学习，因此主观因素比较强。卷积神经网络 (CNN) 会自动的根据图像信息提出深层语义特征用于识别目标物体，在大型图像数据库上具有突出的表现，但是训练深度网络所需的大量数据的时间和计算复杂性始终无法令人满意。幸运的是，一旦在大型数据库上训练了 CNN 模型，就可以使用微调使其适应相似的新数据集，即通过迁移学习解决相关的视觉任务。

CNN 作为近些年深度学习的新宠，其强大的特征提取能力使其在视觉任务

中超越了以往很多技术。CNN 模型可以识别目标变化性较大的图片，例如手写字符，并且对简单几何变换具有很强的鲁棒性。LeNet-5【88】是有 Lecun 提出的用于手写数据集识别的卷积神经网络模型。基于 LeNet-5 网络结构，我们提出了一种用于 DNA 纳米物体识别的特定 CNN 网络结构，如图 3.5 所示。模型结构一共 7 层，包含了一个输入层、5 个隐藏层和一个输出层。输入层为待识别图像。隐藏层包括 3 层卷积“C1”、“C3”、“C5”和 2 层最大池化“S2”、“S4”，其中“C1”包含 32 个卷积内核，大小都为 5*5；“C3”包括 64 个内核，大小为 5*5；“C5”是包含 1024 个节点的全连接的层；“S2”的核为 2*2，“S4”是核为 4*4。最后，输出层包含三个节点，表示三种 DNA 目标类型,分别是平行，交叉，和反平行结构。在 CNN 模型有多种激活函数，这里我们采用 Relu 作为激活函数。

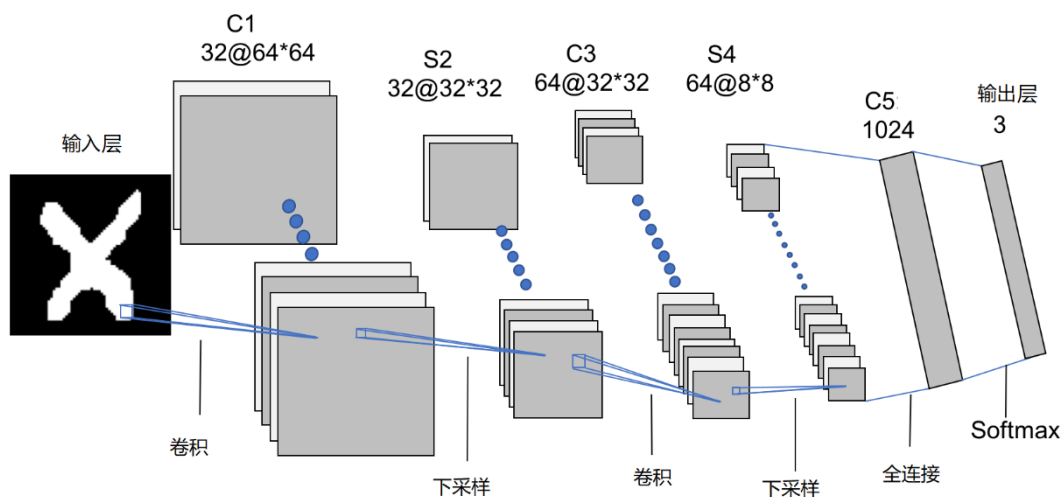


图 3.5 CNN 网络模型结构。

为了缩小预测结果与真实标签的差异，需要根据损失函数优化模型参数，损失函数度量了预测结果与实际标签之间的差异性。常见的损失函数有平方损失、0-1 损失、交叉熵损失等，这里我们采用交叉熵损失。假设有 N 个样本，则 $y_n = [y_n^1, y_n^2, y_n^3]$ 表示第 n 个样本的 one-hot 编码，即标签； $\hat{y}_n = [\hat{y}_n^1, \hat{y}_n^2, \hat{y}_n^3]$ 表示模型的预测结果，即交叉熵损失如下：

$$J = -\frac{1}{N} \sum_{n=1}^N \sum_{i=1}^3 [y_n^i \log \hat{y}_n^i + (1 - y_n^i) \log (1 - \hat{y}_n^i)], \quad (3.3)$$

其中 J 为交叉熵损失，衡量预测值 \hat{y}_n 和标签 y_n 之间的差异性。 J 越小，则标签 y_n 与

预测值 \hat{y}_n 越相似，否则 J 越大，标签 y_n 与预测值 \hat{y}_n 越不相似。通过梯度下降算法不断的更新权重，使得损失函数 J 越来越小，直至收敛至小于一个阈值。

迁移学习的过程包含了两个阶段。首先，在大规模的相似数据集上对CNN模型进行预训练，并保存模型。然后，将预训练阶段保存的基础模型用于真实 DNA 纳米物体图像数据上，进行微调训练。预训练阶段的初始学习率为 10^{-4} ，微调阶段的学习率为 10^{-5} ，在两个阶段中，batchsize大小都设置为100，模型的优化器使用Adma，为了防止模型过拟合使用dropout操作，其中dropout率设定为0.5。这里对输出层的 softmax 分类器的预测概率结果加一个阈值判定，即若预测结果 $\hat{y}_n \geq 0.6$ ，才认定为 i 类别。迁移学习的示意图如图 3.6 所示。

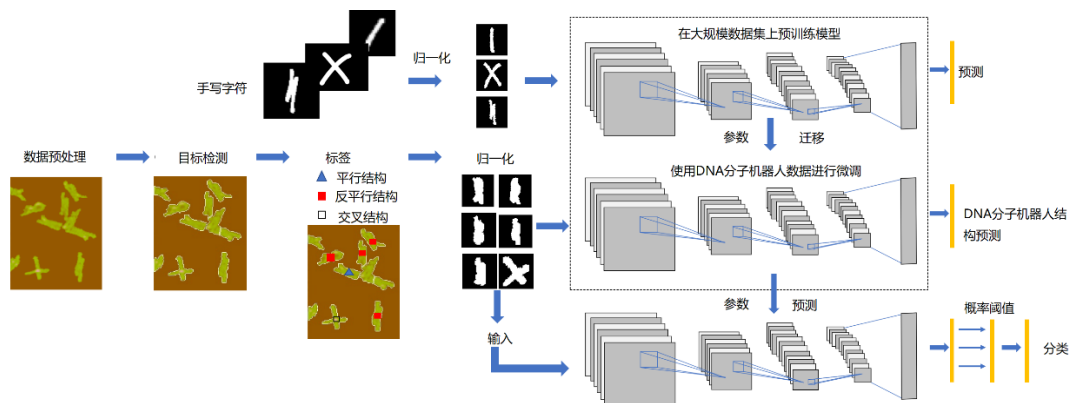


图 3.6 迁移学习过程示意图。

3.3 实验与分析

3.3.1 迁移学习的数据准备

要利用迁移学习，第一步是找到与 DNA 纳米物体相似的其它物体。因某些 DNA 分子机器人的形状类似于一些笔迹，我们使用手写字体库进行迁移学习。如图 3.1 中所示，交叉类型的 DNA 目标是相似于手写数据的“X”，平行类型的 DNA 目标与手写字符集中的“1”相似。对于交叉类型，首先提取来自 UCI 机器学习存储库的 46 个大写字母“X”，然后使用 Python Imaging Library (PIL) 中的 ImageDraw 和 ImageFont 模块创建另外 1954 个“X”。对于平行类型，2000 个手写字母“1”来自 MNIST，这是 NIST（美国国家标准技术研究院数据库）提供的一大套子集。但是逆向平行的 DNA 物体则没有相似的字符，因此我们邀请了 8

个人写了逆向平行的数据将其作为手写数据加入训练数据中，共获得了 2000 个逆平行形式的手写字符。图 3.7 中给出了“X”、“1”和手写逆平行结构的一些示例。

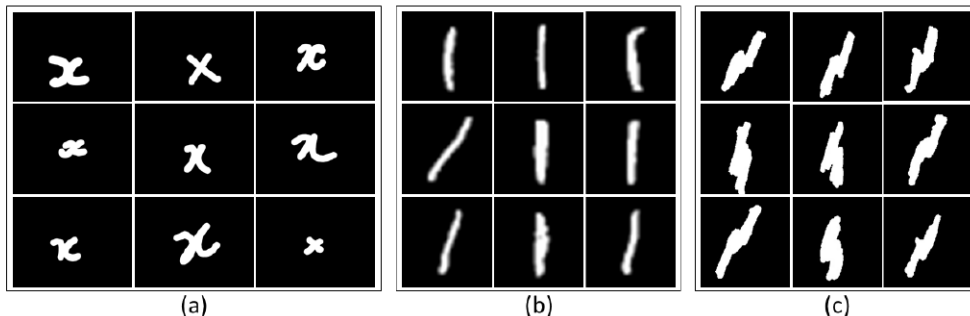


图 3.7 “X”、“1”和逆平行结构的一些数据示例。

因为这些图像数据大小格式存在差异，需要一定的预处理。首先，将所有图像的分辨率调整至 64×64 ，然后，为消除角度的影响，将字符在二维空间中旋转。在这里，将覆盖字符的最小边框的宽度和高度分别定义为 w 和 h ，如图 3.8 所示。如果将最小边界框和 x 轴形成的夹角定义为 ϕ ，则将 ϕ 的补角为 $\theta = 90^\circ - \phi$ 。因此，字符被逆时针旋转 θ 度垂直于 x 轴，旋转的详细过程如图 3.9 所示。为了统一数据个数且保证图像目标不发生形变，还需要进一步就进行规则化，将字符图像数据大小调整为 $1.25h \times 1.25h$ 。最后，再将图像分辨率恢复为 64×64 。数据预处理的过程如图 3.8 所示。

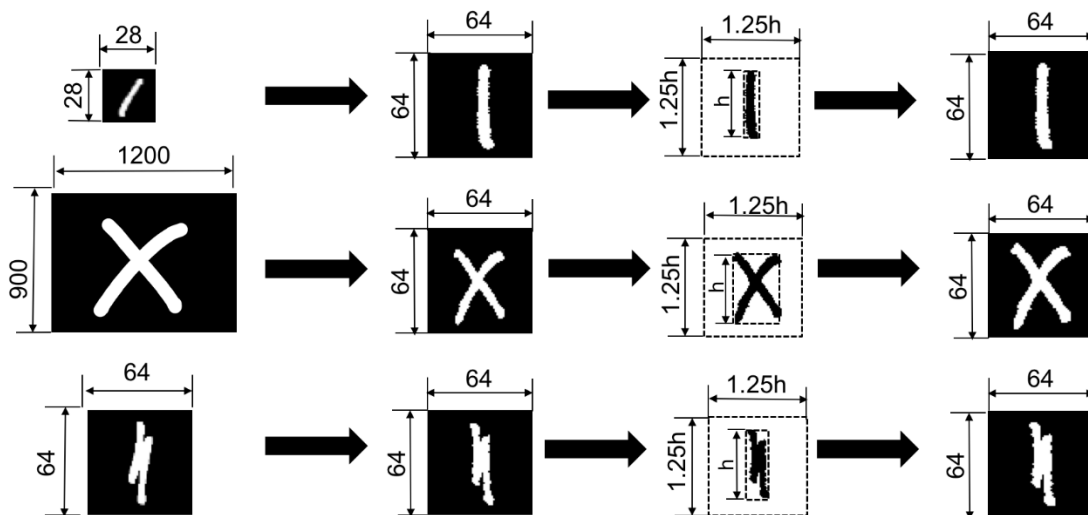


图 3.8 数据预处理过程示意图。

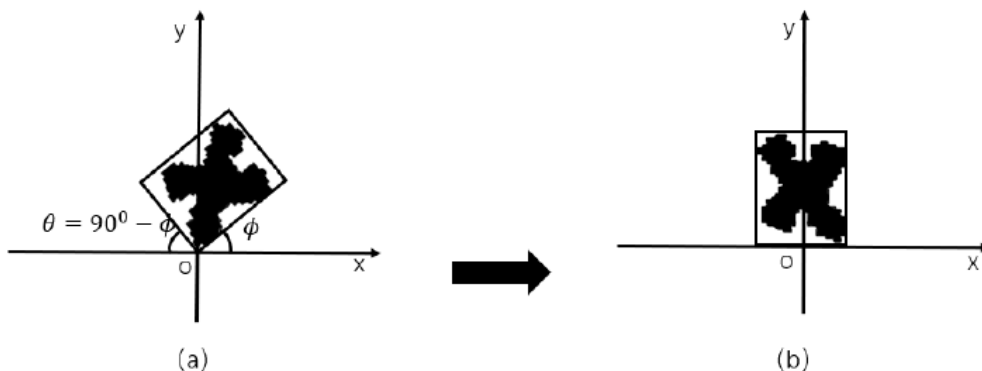


图 3.9 旋转目标示意图。

本研究中，DNA 分子机器人的 AFM 图像作为实验材料用于模型的微调。DNA 分子机器人的 AFM 图像共有 12 张，其中 8 张包含了 255 个目标的图像用于微调卷积神经网络，剩余 4 张包含共包含 168 个目标的图像用于测试。此外，训练时采用交叉验证，每次选取 20 个交叉类型、20 个平行类型和 40 个逆平行类型作为训练数据，剩余的 175 个作为验证数据。

3.3.2 实验比较与分析

DNA 分子机器人识别的结果受分割结果的影响较大，为了验证分割方法的有效性，将基于改进的分水岭分割方法与支持向量机的分割方法【38】进行比较。比较结果如图 3.10 所示，图中用红色圆圈标记的一些示例可以用我们的方法对某些重叠的形状进行分割，而使用支持向量机的方法则不能分割。

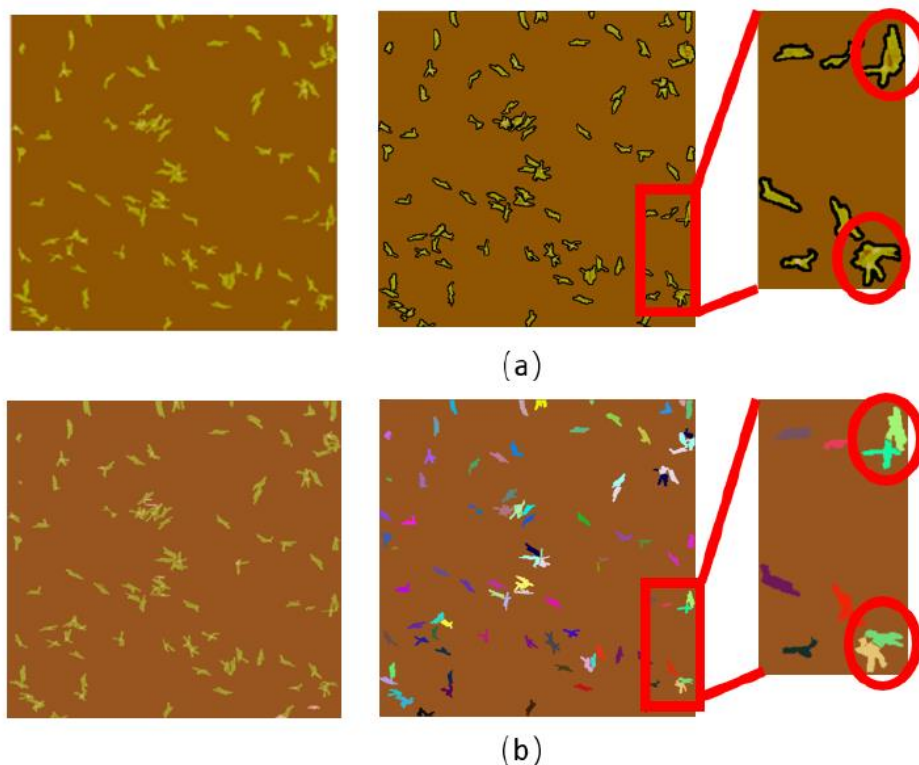


图 3.10 分割结果的比较。(a) SVM 方法的分割结果；(b) 我们方法的分割结果。

下面对识别方法的验证和比较，本研究提出的深度学习模型与VGG16和ResNet50进行了比较。在对比实验中，进行10次交叉验证。实验结果显示VGG16为平均准确率为74.28%、ResNet50平均准确率为62.57%、我们模型的平均准确率为84.22%。详细结果如表 3.3 所示，该表显示了我们的方法模型在准确性上明显优于VGG16和ResNet50。

表 3.3 与 VGG16、ResNet50 模型十次实验准确率对比。

实验	我们模型	VGG16	ResNet50
1	80.57%	75.42%	59.42%
2	78.85%	74.85%	60.57%
3	84.00%	74.28%	58.28%
4	87.42%	74.28%	60.00%
5	86.28%	73.71%	59.42%
6	82.85%	74.28%	58.28%
7	84.57%	74.28%	58.28%
8	85.71%	73.71%	68.00%
9	85.14%	73.71%	72.00%
10	86.85%	74.28%	71.42%
平均	84.22%	74.28%	71.42%

4张测试图像的识别结果如图 3.11 和表 3.4 所示。图 3.11 (a) 中没有重叠 DNA 分子机器人，因此，其识别准确率最高。图 3.11 (b) 的具有大量的 DNA 分子机器人且具有一定重叠，使得识别准确率较低。图 3.11 (c) 和图 3.11 (d) 包含许多复杂且粘着的 DNA 分子机器人，因此它们比图 3.11 (a) 和 (b) 更难以处理且有错误识别，其主要的错误来自一些无法识别的不规则形状。

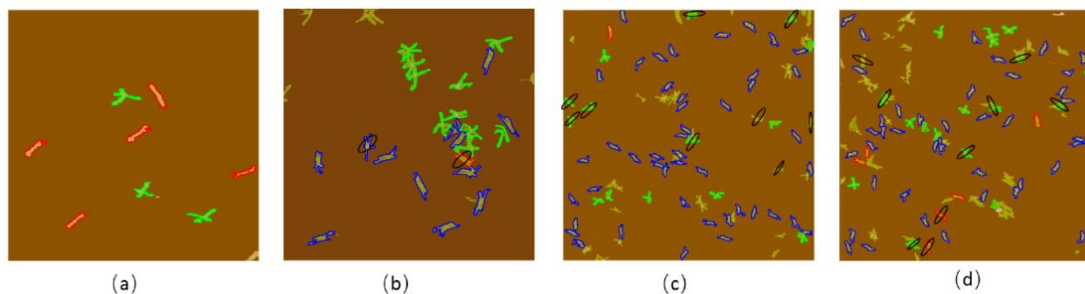


图 3.11 四张 AFM 图像的分割结果，红色、蓝色和绿色分别表示了识别出的平行类型、反平行类型和交叉类型；黑色椭圆表示未能正确分类和识别出的结果。

表 3.4 图 3.11 中四张 AFM 图像的认识结果。

图像数据	所含目标数	过滤后的目标数	正确识别	错误识别	未识别	准确率
图 3.11(a)	9	8	8	0	0	100%
图 3.11(b)	28	23	21	2	0	91.3%
图 3.11(c)	86	67	51	9	7	76.1%
图 3.11(d)	99	70	48	14	8	68.6%

为了进一步验证我们的方法的性能，将我们的方法与 Han【83】提出的方法进行了比较，该方法结合了曲率尺度空间方法和凸凹检测技术来识别 DNA 分子机器人。我们方法识别 Han 文章中所用图像数据的结果如图 3.12 所示，详细的对比结果如表 3.5 所示。

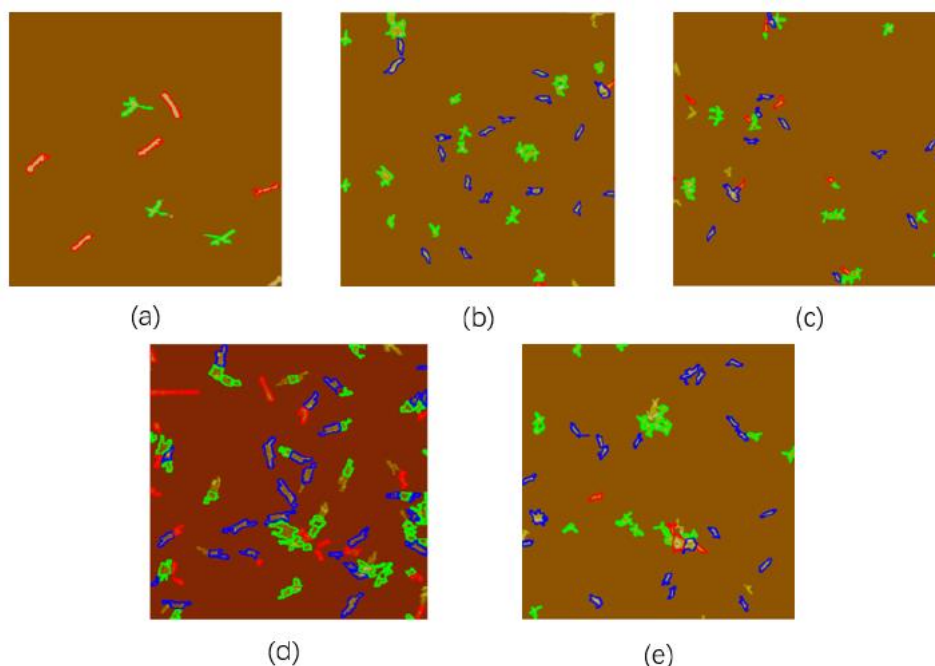


图 3.12 我们方法对 Han 文章中所用数据的识别结果。

如表 3.3 所示，我们方法的平均召回率高于 Han 方法，其主要原因是 Han 的方法无法分解重叠的纳米物体，并且无法识别大量的重叠的纳米物体。此外，由于我们的方法消除了很多噪声，因此我们的方法的平均精度高于 Han 的方法。

表 3.5 与 Han 方法在图 3.12 中的五张 AFM 图像的认可准确率对比。

图像数据	我们方法的精确率	我们方法的召回率	Han 方法的精确率	Han 方法的召回率
图 3.12(a)	100%	100%	83.3%	62.5%
图 3.12(b)	82.9%	78.4%	60.0%	32.4%
图 3.12(c)	78.8%	70.3%	85.7%	37.5%
图 3.12(d)	86.5%	78.0%	70.6%	29.3%
图 3.12(e)	75.9%	71.9%	68.2%	26.3%
平均	84.8%	79.7%	73.6%	37.6%

此外，我们将【83】中所使用三张 AFM 图像作为实验材料，与 Han 的方法进行比较，结果如图 3.13 所示，可以发现，本章所提出的方法更加有效。

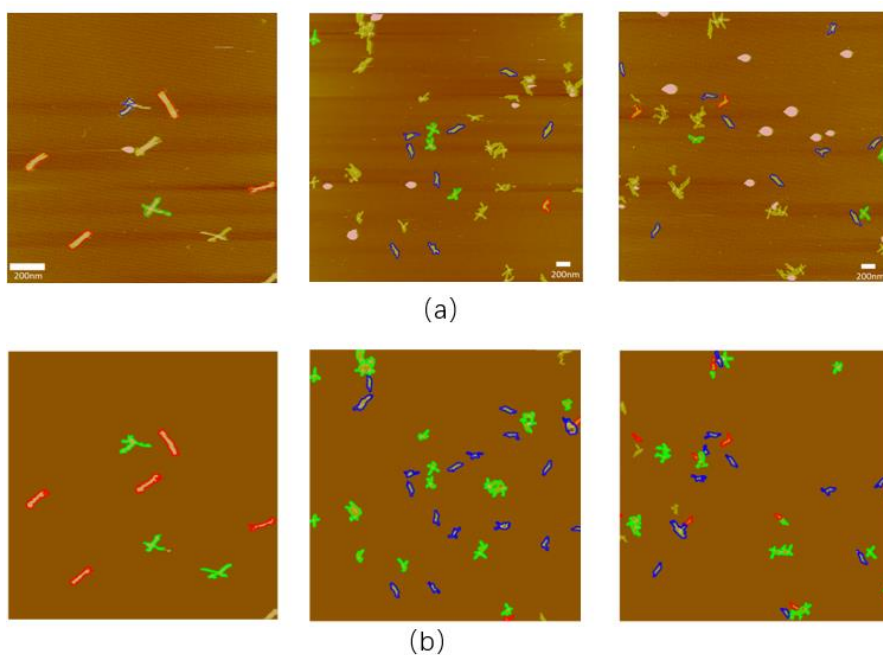


图 3.13 与 Han 方法的识别结果对比。(a)为 Han 方法的识别结果；(b)我们方法的识别结果。

3.4 本章小结

本章在小样本环境下提出了一种分割和识别纳米物体 AFM 图像的新方法，

该方法包括了去噪、分割和识别三步。利用高斯拉普拉斯 (LoG) 去除图像中的噪声可以获得更清晰的图像, 以便进一步观察纳米物体, 改进分水岭算法实现对重叠物体的分割, 利用迁移学习解决小样本问题并实现对纳米物体的识别。本研究不仅解决了小样本环境下的 DNA 分子机器人的识别任务, 其思想也可用于其他类型的小样本数据识别, 实验结果证明了本方法的有效性, 因此该研究可为小样本下的 AFM 图像处理开辟新的道路。

尽管本研究提供了一种有效的纳米结构图像处理方法, 但仍存在需要改进的地方和不足。首先, 由于重叠区域较大, 使用改进的多尺度分水岭算法无法正确分割某些形状; 其次, 一些纳米物体由于其形状扭曲而被错误识别; 最后, 目标的识别过分依赖于图像分割结果, 针对以上缺点和不足还需进一步研究。

第四章 基于复杂网络理论的材料图像分割方法

在材料微观结构研究中,结合计算机图像处理能力对材料微观结构图像进行分析是非常重要的研究方法。材料微观结构图像的分割是材料微观结构分析工作中最为关键的一步,通过对材料微观结构图像的分割不仅可以反映不同微观组织之间的关系,同时根据图像分割后的结果可以关联材料的性能。但是,材料制备的复杂、材料种类的庞杂和材料微观结构获取的难度,导致可用的微观结构图像很少。近些年,图像分割技术逐渐成熟,但面对材料图像这种纹理复杂、分辨率高和数据极其有限的情况,往往都难以有效应用。本研究另辟蹊径,结合复杂网络的相关理论,提出了无需大量数据就能实现对材料微观结构图像分割的方法。该章节内容已经申请中国专利,专利号为:202010083808.6。

4.1 复杂网络背景介绍

在复杂网络理论的研究热潮下,复杂网络方法处理和分析图像数据的优势逐渐显现,也吸引了越来越多的学者们的注意。研究人员发现,利用复杂网络可以结合图像局部和整体属性,为图像处理和分析提供更加全面的信息,用于解决图像处理和分析的问题,比如轮廓识别【89】、纹理分析【90】等。

由节点和节点之间的关系(边)构成了复杂网络,其为一种非欧几里德结构。研究人员通常只关注节点与节点之间是否存在连边,因此,不依赖于节点的特定位置和连边形式的性质被称为复杂网络的拓扑属性,相对应的网络结构被称为网络拓扑结构,对其描述时,常常借助图来表示。图常被表示为 $G = (V, E)$,其中 $V = \{v_1, v_2, \dots, v_n\}$ 为节点的集合, $E = \{e_1, e_2, \dots, e_n\}$ 为边的集合。特别的,当网络中的边具有权重时,图 G 也可以表示成 $G = \{V, E, W\}$,其中 $W = \{w_1, w_2, \dots, w_3\}$ 表示为边的权重集。复杂网络理论中有多个拓扑属性用于描述拓扑结构,例如,邻接矩阵、节点的度、节点强度等,描述如下:

- 1) 邻接矩阵: $n * n$ 的方阵,表示网络节点一阶连接关系。矩阵中的 (i, j) 位置的取值根据节点 i 与节点 j 之间是否存在连边来判断,若有则为1,没有则为0;
- 2) 节点的度: 表示与节点相连边的个数;
- 3) 节点的强度: 对于带权图,为节点所有连边的权值之和;

社区结构是复杂网络的拓扑理论中非常重要属性之一，指网络中的节点根据社区属性进行划分，社区内部的节点之间有着较大的相似性，而不同社区中的节点之间相似性很低，其揭示了网络拓扑结构中隐藏的规律和行为特征。针对社区结构这一拓扑特性，可以与图像分割任务关联起来，因为在图像分割任务中，是找出不同的目标区域，而在社团结构检测任务中，是找出不同的社团结构，两者具有很强的关联性。如果将图像表示成网络拓扑结构，那么网络中相互连接紧密的节点，即一个社团结构，对应着图像中同一目标区域内的像素点集，连接稀疏的节点表示不同目标区域内的像素点，从而可以实现图像的分割。

4.2 材料微观结构图像分割方法

4.2.1 预处理

预处理分为降噪、聚类、节点集生成三步。图像获取时的操作或是材料本身原因经常引入一些噪声，对后续图像处理具有极大影响，因此通过预处理去除图像中的噪声，常见有均值滤波、中值滤波、高斯滤波等【91】方法。这里采用高斯滤波，其为一种线性平滑滤波，对整幅图像进行加权平均，每一个像素点的新值都是由其本身和领域内的其他像素值经过加权平均后得到的，其权值随着与中心像素点的距离增加而呈现高斯衰减。这样的好处在于，离算子中心很远的像素点的作用很小，从而能在一定程度上保持图像的边缘特征。

滤波降噪之后，对图像进行像素级别上的聚类，以实现针对不同像素级上目标的初步分割。将图像按照像素进行聚类，是指将像素值相近的像素点归为一个区域，因为同一目标其像素值是近似的。常见的聚类算法有 K-means 算法【92】、DBSCAN 算法【93】、Meanshift 算法【94】等。由于 K-means 简单高效，解释性强且可以自定义聚类簇的数量，因此本研究采用 K-means 算法。

生成网络空间中的节点集的直观做法是对聚类后的图像，每个目标像素点定义为网络中的一个节点，但图片分辨率很大时，时间、空间存储以及计算消耗太大。为了解决这一问题，本章对图像进行网格化，让一个网格内聚集的多个像素点构成一个网络节点。网格化处理是将图像划分成一个个大小为 $d*d$ 的方格，当 d 取 1 时表示一个网格为一个像素点，计算每个方形网格内中提取的类别的像素

点数，若大于设定阈值，将其记录并标号，并将网格中心点坐标表示节点位置，一次扫描下来则可以获得网络中的所有节点。这样处理不仅使得网络中节点的数量比直接使用像素点作为节点数量少了很多，而且也过滤了大量的噪声点，示意图如图 4.1 所示。

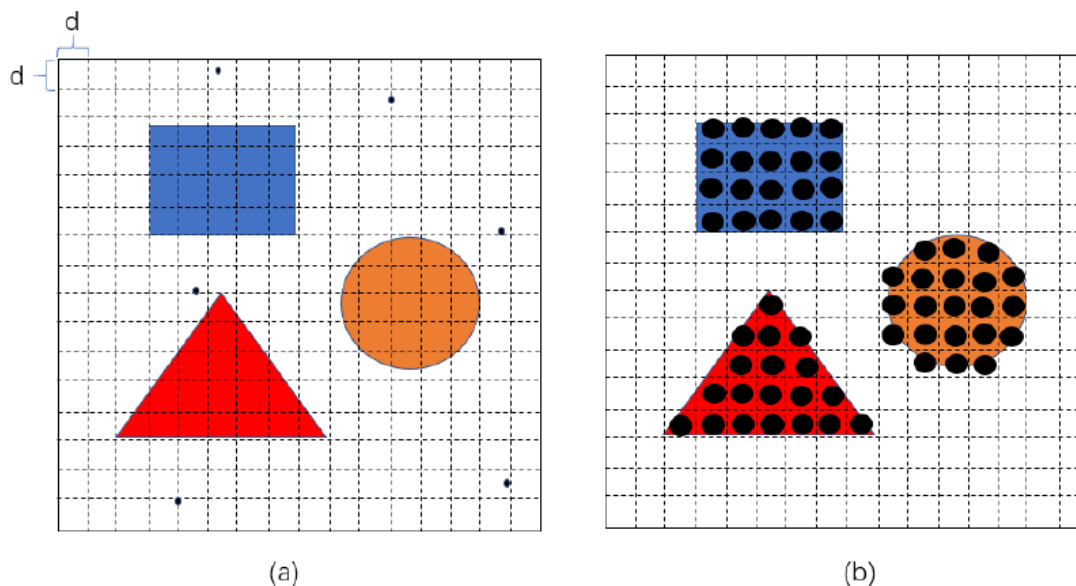


图 4.1 网格化操作示意图。(a) 中的蓝色正方形、红色三角形和橙色圆形表示三个目标物体，小黑点表示噪声；(b) 网格化的结果，黑色实心圆表示节点。

4.2.2 网络拓扑结构建立

在预处理之后已获得节点集 V ，其中每个节点都有一个坐标值，即 $v_i(x_i, x_j)$ 。那么在这样一个二维坐标空间中，节点之间的相似性则可以使用节点之间的欧式距离来度量，即 $\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ 。为了简化后续的运算，将节点之间的距离进行归一化处理，即使得节点之间的距离取值在 $[0,1]$ 内，公式如下所示：

$$\Delta = \max \left(dis(v_i, v_j) \right), \quad (4.1)$$

$$dis(v_i, v_j) = \frac{dis(v_i, v_j)}{\Delta}, \quad (4.2)$$

其中 $dis(v_i, v_j)$ 表示归一化后的节点 v_i 和 v_j 之间的距离。

有了任意两个节点之间归一化的距离之后，引入半径阈值 $r \in [0,1]$ ，用来判

断两节点之间是否应该存在连边，公式如下所示：

$$e(v_i, v_j) = \begin{cases} r - \text{dis}(v_i, v_j), & \text{dis}(v_i, v_j) < r \\ 0, & \text{其它} \end{cases}, \quad (4.3)$$

若节点 v_i 与节点 v_j 之间的距离小于阈值 r ，则认为节点 v_i 和节点 v_j 之间存在连边。因节点之间的相似性与节点距离成反比，故将阈值 r 与 $\text{dis}(v_i, v_j)$ 的差作为节点之间的相似性。

经过上述处理之后的拓扑结构存在的问题，如图 4.2 所示，(a) 和 (b) 具有相同的节点数量，但图 4.2 (b) 的连接紧密性明显要比图 4.2 (a) 高。为了度量网络中节点的连接紧密型，引入节点的连接密度，公式如下所示：

$$d(v_i) = \frac{s(v_i)}{k(v_i)}, \quad (4.4)$$

其中 $s(v_i)$ 为节点强度， $k(v_i)$ 为节点度。为了避免奇异值的影响，将节点的连接密度归一化到区间 $[0,1]$ 内，公式如下所示：

$$d(v_i) = \frac{d(v_i)}{\max(d(v_i))}, \quad (4.5)$$

通过节点的连接密度断开一些连接不紧密的节点，使得网络拓扑结构更加紧密。因此，一个密度阈值 t 被定义，若两个节点的密度同时大于阈值时，则保留原本连接，否则断开，公式如下：

$$e(v_i, v_j) = \begin{cases} e(v_i, v_j), & d(v_i) > t \text{ 且 } d(v_j) > t \\ 0, & \text{其它} \end{cases}, \quad (4.6)$$

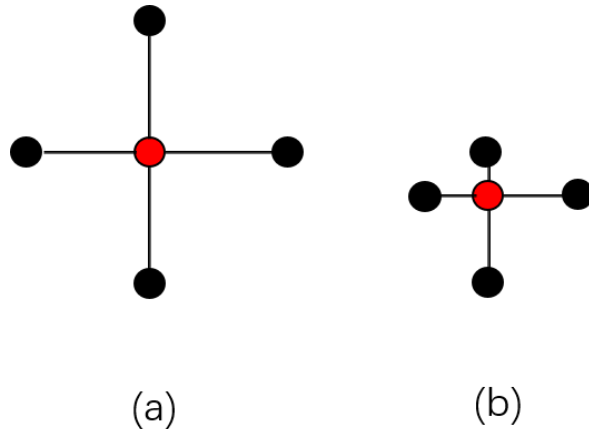


图 4.2 相同节点数，不同节点连接密度示意图。

4.2.3 参数优化

最终的网络拓扑结构取决于半径阈值 r 和密度阈值 t 的取值，其中半径阈值 r 控制节点的连接领域，对于任意一个节点，其 r 领域内的所有节点都会成为其邻居节点，即存在连接边，而密度阈值 t 会断开一些不紧密的连接。当 r 取值很小， t 取值很大时，节点是孤立的，不存在连接边；当 r 取值很大， t 取值很小的时候，网络中所有节点之间都有连边，即为完全连通图。随着 r 和 t 的不断变化，网络拓扑从一个个孤立节点到完全连通。因此，我们需要选取适当的 r 和 t ，使得节点相似性高的聚集在一起，即有连接关系；而相似性低的节点之间则没有连接，从而形成一个密集连接的社区结构。因此，选取最优的 r 和 t ，使得产生的一个个连通区域对应着分割目标。为了评估拓扑结构划分的是否合理，我们采用模块度进行衡量。模块度由 Newman 提出【95】，也称为模块化度量 Q ，其公式如下所示：

$$Q = \frac{1}{2m} \sum_{i,j} [A_{ij} - \frac{k_i k_j}{2m}] \delta(C_i, C_j), \quad (4.7)$$

$$\delta(C_i, C_j) = \begin{cases} 1, & C_i = C_j \\ 0, & C_i \neq C_j \end{cases} \quad (4.8)$$

其中 m 表示网络中的总边数， A 表示网络的邻接矩阵， k_* 表示节点*的度， C_* 表示节点*所属的社区。

从公式（4.7）可以看到，模块度的大小定义为社团内部的总边数和网络中总边数的比值与一个数值的差，该数值表示将网络设定为随机网络时，同样的社区分配所形成的社团内部总边数和网络中总边数的比值大小。因为随机网络不具备社团结构性质，所以公式（4.7）中的 Q 越大，即差值越大表示社团划分越合理。模块度是基于连通图的，而这里得到的网络拓扑并不是连通的，因此为了衡量这里的非连通网络中的每一个子网络划分是否合理，我们基于模块度进行调整。首先，将半径阈值 r 和密度阈值 t 确定的网络拓扑结构中的子网络，即连通域，作为一个社团划分结果。然后，在这种社团分配下，将每个节点进行全连通，保证权值与节点之间的距离成反比，并用公式（4.9）进行归一化。考虑权值的模块度公式如下：

$$w(v_i, v_j) = \frac{\max(dis(v_i, v_j)) - dis(v_i, v_j)}{\max(dis(v_i, v_j)) - \min(dis(v_i, v_j))}, \quad (4.9)$$

$$Q = \frac{1}{2W} \sum_{v_i, v_j} [D(v_i, v_j) - \frac{s(v_i)s(v_j)}{2W}] \delta(C_{v_i}, C_{v_j}), \quad (4.10)$$

其中， W 是网络总强度之和， C_{v_i} 和 C_{v_j} 分别表示节点 v_i 、 v_j 所属的社区， $D(v_i, v_j)$ 满足以下公式：

$$D(v_i, v_j) = \begin{cases} w(v_i, v_j), & e(v_i, v_j) \in E \\ 0, & \text{其它} \end{cases}, \quad (4.11)$$

利用公式（4.10），将节点度改成节点强度，将邻接矩阵改成年值矩阵，因此具有模块度的性质，该值越高表示划分越合理。将公式（4.10）作为目标函数优化半径阈值 r 和密度阈值 t ，通过选择最优的 r 和 t 使得模块度最大。

4.2.4 轮廓提取与分割

根据章节 4.2.1 可知网络中每个节点对应的都是图像中某一网格内像素点的集合。在章节 4.2.2 节中，定义了网络拓扑结构，将相似的节点进行连接，即将相似像素点都放在一个社团结构内并且彼此之间具有连接关系。因此，只要将拓扑图像中每个连通区域的轮廓提取出来便得到了图像中每个目标区域。首先，二值化网络拓扑图像，使用 0 表示黑色背景，用 1 表示白色目标。其中像素点 1 构成的区域成为 1 连通域，由像素点 0 构成区域称为 0 连通域。对于一个像素点定义 4 邻域和 8 邻域，其中 4 邻域表示像素位置的上、下、左、右四个近邻位置，8 邻域表示对应像素位置的上、下、左、右、左上、左下、右上和右下 8 个近邻位置。如果一个 1 像素点在它的 4 邻域或者 8 邻域内存在一个 0 像素点，则表示该点为边界点。因此，外边界由任意 1 连通域和直接包围它的 0 连通域之间的边界点集合构成。为了提取网络拓扑的外轮廓，只需找到对应的外边界即可。

4.3 实验

4.3.1 实验数据处理

在本研究中，选择 $\text{TiB}_2\text{-TiC}$ 的微观结构图像，如图 4.3 所示。

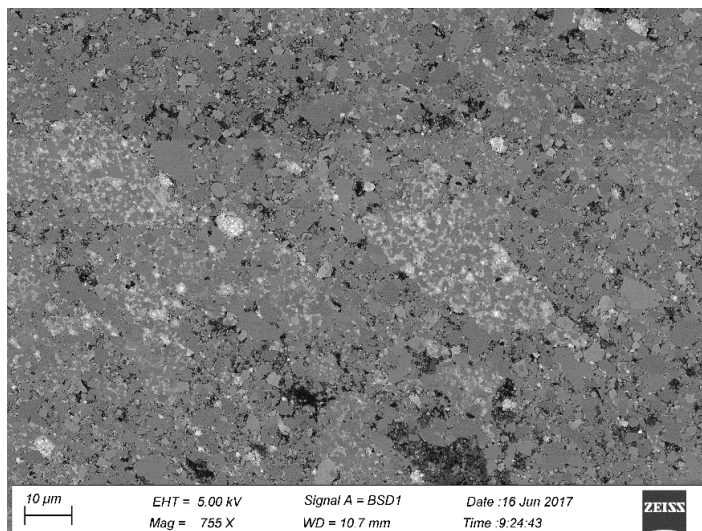


图 4.3 $\text{TiB}_2\text{-TiC}$ 的微观结构图像。

因 $\text{TiB}_2\text{-TiC}$ 的纹理非常复杂，现存的图像分割算法很难对其分割，分析 $\text{TiB}_2\text{-TiC}$ 分别是黑、深灰、灰白、浅灰、和白，因此在预处理过程中将 K-means 聚类的类别数 k 设置为 5，聚类结果如图 4.4 (a) 所示，用五种不同的颜色表示不同的相。这里，我们选择其中一种类别的区域作为实验例子，即 TiC 相，在图中紫色表示，网格化生成的节点如图 4.4 (b) 所示。

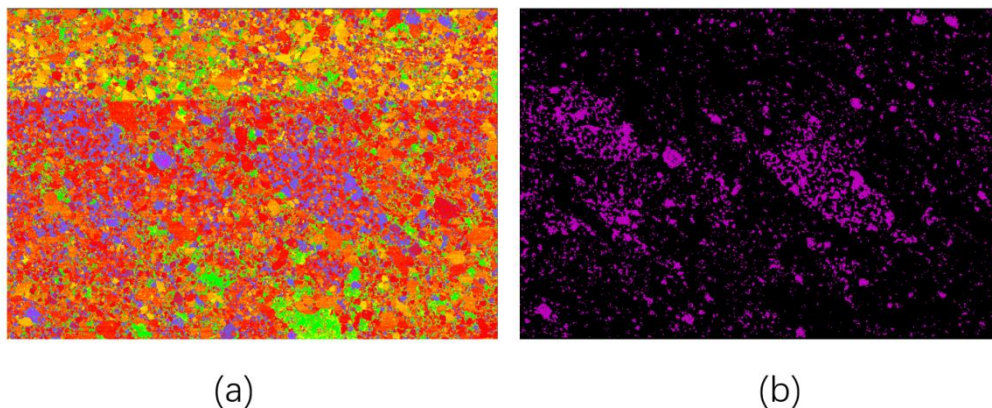


图 4.4 $\text{TiB}_2\text{-TiC}$ 微观结构 K-means 聚类结果。

4.3.2 网络构建与图像分割

在构建网络拓扑时,选取不同半径阈值 r 和密度阈值 t 的情况下可以得到不同的拓扑结构,结果如图 4.5 所示。结果显示,半径阈值 r 越大,网络连接范围变得越广,而节点密度阈值 t 越大,网络连接变得越分散。为了选取合适的参数,选择模块度最大值对应的 r 和 t 作为最终参数。实验中,设置 r 的取值范围为 $[0.001, 0.009]$,步长为 0.001, t 的取值范围为 $[0.1, 0.9]$,步长为 0.1。在这个参数区间外的取值会导致网络没有连接,所有节点都是孤立的,因此不予考虑。在设定的半径阈值 r 和密度阈值 t 的取值区间内,绘制模块度变化曲面,如图 4.6 所示。结果显示,函数在大部分区域为凸,当半径阈值 r 和密度阈值 t 取值不合理时,模块度为 0,因为这种情况下网络拓扑中没有连边存在,每个节点都是孤立存在的。网络会随着半径阈值 r 和密度阈值 t 的变化,形成社区结构。在实验结果中,当 r 取值为 0.006、 t 取值为 0.4 时,模块度取值最大,说明此时社团划分最好,然后根据得到的拓扑图,进行轮廓提取,分割目标区域,结果如图 4.7 所示。

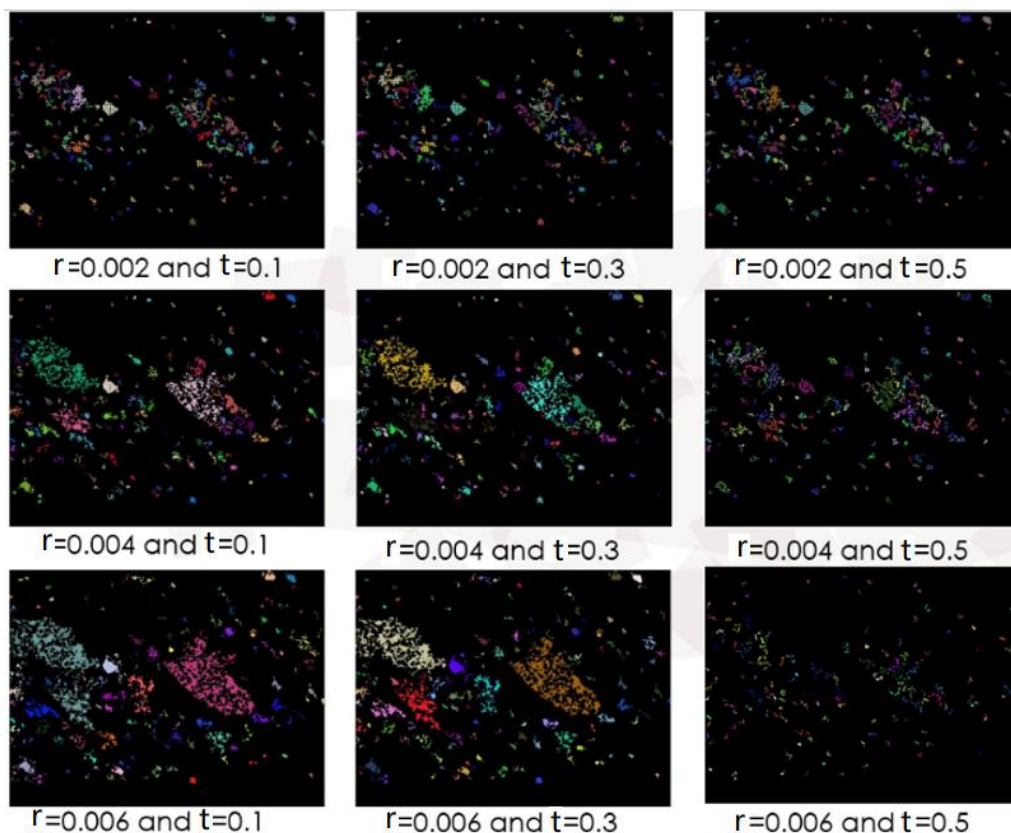


图 4.5 不同阈值下的网络拓扑结构图。

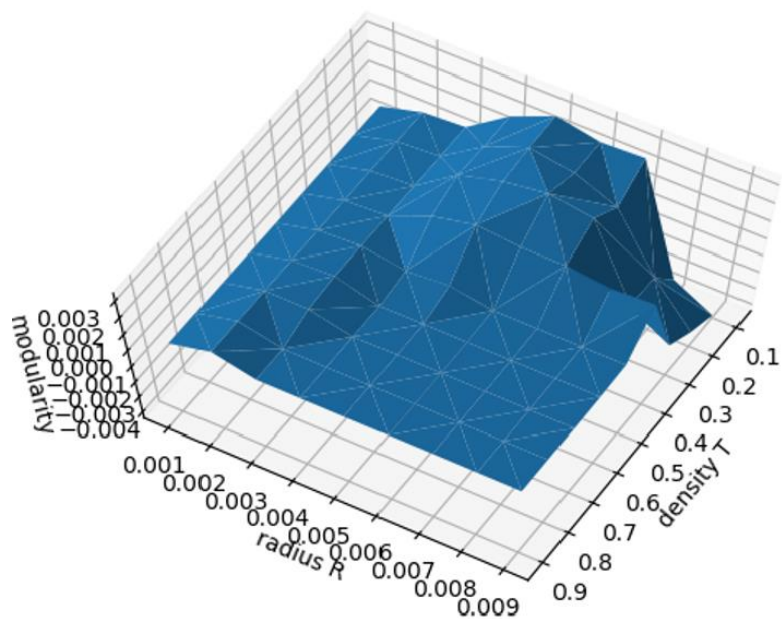


图 4.6 模块度随着半径阈值 r ，密度阈值 t 变换曲面。

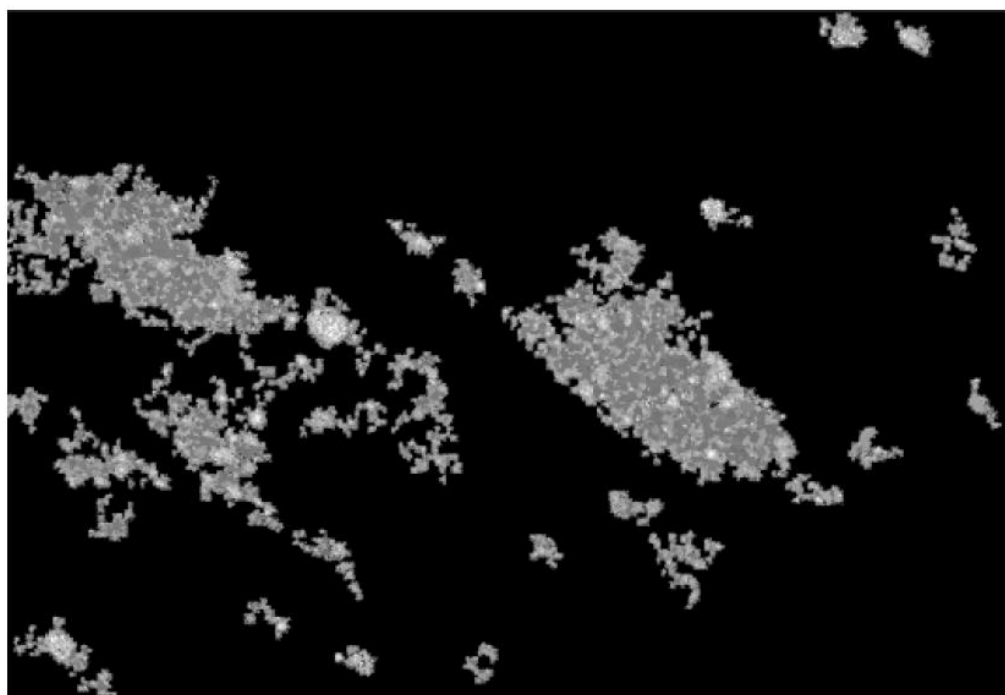


图 4.7 TiB₂-TiC 微观结构图像分割结果。

4.4 总结

本章提出一种基于复杂网络理论的图像分割方法，通过将图像映射为网络空间中的拓扑结构，划分最优社区来实现图像分割，该方法对数据量没有要求，即

使只有一张图像也可以分割。本章方法为小样本问题提供了新思路、新方向。此外，本章提出的方法可以自动化找出合适参数，无需人工定义参数，避免因调参导致的分割错误。为了验证提出方法的有效性，将 $\text{TiB}_2\text{-TiC}$ 的微观组织结构图像作为实验材料进行测试，结果显示了本章方法的有效性。

本章方法仍具有一些不足之处。一方面，将图像映射到网络空间中会有一些程度上的细节信息损失，从而导致部分边缘分割效果不好。因此，如何高效的，无损失的完成从图像空间到网络空间的映射是需要深入研究。另一方面，在网络空间下进行社区检测需要很大的计算量，所费时间往往较长，因此，如何优化模型，加快分割速度需要进一步研究。

第五章 总结与展望

5.1 结论

计算机技术的快速发展为材料科学领域的相关研究提供了有效的技术支持和强大的分析工具。加快材料科学与计算机科学的交叉融合,利用计算机技术辅助材料的研发与设计,可以促进学科的发展、科技的创新和人类的进步。材料科学与计算机科学交叉融合的困难点之一是材料数据的匮乏,即材料小样本问题,解决材料小样本问题,才能极大限度的提高材料研发的效率。本文结合机器学习、深度学习、迁移学习以及复杂网络等知识,着重研究小样本数据的处理问题,且分别针对三种不同的数据,即晶体结构数据、AFM 图像数据、材料微观组织结构图像数据,提出了相应的解决方案,具体内容总结如下:

- 1) 针对数据量小且类别分布不均匀的晶体结构数据,提出了基于机器学习的晶体结构识别方法。该方法根据数据本身的特点对数据进行扩增,并定义特征表示和特征选择,最后训练多层感知机实现对晶体结构的识别。该方法很好的解决了材料小样本问题,且在效率、准确率、复杂度上都具有一定的优越性,但是该方法时间花费比较大,还需进一步提升。
- 2) 针对小样本的 AFM 图像,提出基于迁移学习的图像识别方法。该方法从预处理,分割和识别三步出发,逐层递进,在预处理阶段去除 AFM 自带的噪声,在分割阶段提出改进的分水岭分割方法解决了目标重叠的问题,最后通过迁移学习实现 AFM 图像的识别。该方法很好的解决了小样本问题,也具有很高的扩张性,但是识别结果对分割结果的依赖性太强,还需要进一步研究。
- 3) 针对小样本的材料微观组织结构图像数据,提出了基于复杂网络的图像分割方法。该方法发展了复杂网络中相关理论,通过将图像数据映射到网络空间中,在网络空间中构建网络拓扑并分析拓扑属性,实现对材料微观结构图像的分割。该方法很好的解决了小样本的问题,但是复杂度较高,还需进一步优化。

本文对以上三个方法都进行了实验验证与分析,实验表明这三种方法在面对

小样本问题时能取得较好的分割和识别结果，具有很高的准确性以及使用价值，可以在一定程度上解决材料面临的小样本问题，加快材料的研发，为材料科学的发展提供技术支持。

5.2 工作展望

利用计算机技术辅助材料设计已成为材料发展的重要环节，然而材料的小样本问题还处在初步探索的阶段，未来的研究工作任重而道远。未来工作可以从以下方面进一步探究：

- 1) 对已有数据进行扩增，生成相似的数据是非常有效的、直接的解决小样本问题的方式。以生成对抗网络为基础的数据生成模型，生成的样本数据与真实数据之间具有很高的相似性，在一定程度上解决了图像领域内的一些小样本数据问题。因此面对材料小样本数据，结合材料数据的特征，利用生成对抗网络的数据生成模型对少量已有材料数据进行扩增是值得进一步研究的。
- 2) 无监督学习方法无需大量数据的训练，但其准确性往往没有监督学习高。若将无监督学习方法与其他相关学科领域，例如统计学理论、复杂网络理论进行结合，分析小样本材料数据本身的内在属性特征，有望提升数据处理准确性。
- 3) 机器学习和深度学习的许多模型都依赖于大数据的训练，面对小样本问题时，困难点是数据有限，模型难以拟合。元学习的出现为小样本学习提供了新思路，其将以往的经验用于新任务的指导。将元学习模型应用于材料小样本数据中，把在常见的、数据量大的材料数据中得来的经验，应用于小样本的新型材料数据应该是合理有效的。

以上的内容能否完成，需要我们不断的探索，不管对材料科学还是计算机科学都具有很大研究价值，为科学的进步、人类的发展带来正面且极大的意义。

参考文献

- 【1】 Jain A, Ong S P, Hautier G, Chen W, Richards W D, Dacek S, Cholia S, Gunter D, Skinner D, Ceder G, Persson K A. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation[J]. *APL Materials*, 2013, 1(1):011002.
- 【2】 Korniy V, Rusyn B, Lysak Y, Lutsyk O. Quantitative Analysis of Metallographic Images[C]. *Modern Problems of Radio Engineering, Telecommunications, and Computer Science*, 2006. TCSET 2006. International Conference. IEEE, 2006:251-252.
- 【3】 Broderick S, Suh C, Nowers J, Vogel B, Mallapragada S, Narasimhan B, Rajan K. Informatics for combinatorial materials science[J]. *JOM*, 2008, 60(3):56-59.
- 【4】 Alpaydin, E. *Introduction to machine learning* (2nd ed.). Cambridge, MA: MIT Press, (2010).
- 【5】 LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* 521, 436–444 (2015).
- 【6】 Kalidindi S R, Stephen R N, Ayman A S. Microstructure informatics using higher-order statistics and efficient data-mining protocols[J]. *Jom*, 2011, 63(4): 34-41.
- 【7】 DeCost B L, Elizabeth A H. A computer vision approach for automated analysis and classification of microstructural image data[J]. *Computational materials science*, 2015, 110:126-133.
- 【8】 Lowe D G. Object recognition from local scale-invariant features. *International Conference on Computer Vision (ICCV)*, Corfu, Greece, 1999, 1150-1157.
- 【9】 Lowe D G. Distinctive image features from scale-invariant key points. *International Journal of Computer Vision (IJCV)*, 2004, Vol. 60(2), 91–110.
- 【10】 Bay H, Tuytelaars T, Gool L V. SURF: speeded up robust features. *ECCV*, 2006.
- 【11】 Bay H, Ess A, Tuytelaars T, Gool L V. Speeded-up robust features (SURF), *Computer Vision and Image Understanding (CVIU)*, 1999, Vol. 110(3), 346–359.
- 【12】 Li W, Bebis G, and Bourbakis N. G. 3D Object Recognition Using 2D Views. *IEEE Transactions on Image Processing*, 2008, Vol. 17, No. 11, 2236-2255.
- 【13】 Bebis G, Louis S, Varol Y and Yfantis A. Genetic Object Recognition Using Combinations of Views. *IEEE Transactions on Evolutionary Computation*, 2002, Vol. 6, No. 2, 132-146.

- 【14】 Ren X. Learning and Matching Line Aspects for Articulated Objects. *Computer Vision and Pattern Recognition, CVPR 07. 2007*, 1-8.
- 【15】 Rumelhart DE, Hinton GE, Williams R J. Learning representations by back-propagating errors [J]. *Nature*, 1986, 323: 533-536.
- 【16】 Kingma D P, Welling M. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.
- 【17】 Goodfellow I J, Pougetabadie J, Mirza M, et al. Generative Adversarial Nets[C]. *Neural Information Processing Systems*, 2014: 2672-2680.
- 【18】 Ratliff L J, Burden S A, Sastry S, et al. On the Characterization of Local Nash Equilibria in Continuous Games[J]. *IEEE Transactions on Automatic Control*, 2016, 61(8): 2301-2307.
- 【19】 Erpek T, Sagduyu Y E, Shi Y. Deep Learning for Launching and Mitigating Wireless Jamming Attacks[J]. *Networking and Internet Architecture*, 2018.
- 【20】 Iqbal, Talha & Ali, Hazrat. Generative Adversarial Network for Medical Images (MI-GAN). *Journal of Medical Systems*. 2018, 42. 10.1007/s10916-018-1072-9.
- 【21】 Socher R, Ganjoo M, Manning C D, et al. Zero-shot learning through cross modal transfer. *Advances in Neural Information Processing Systems*, 2013, 1: 935-943.
- 【22】 Cao X, Wipf D, Wen F, et al. A practical transfer learning algorithm for face verification. *ICCV*, 2013, 1: 3208-3215.
- 【23】 Gao J, Ling H, Hu W, et al. Transfer learning based visual tracking with Gaussian processes regression. *ECCV*, 2014, 1: 188-203.
- 【24】 Wang J, Zheng V W, Chen Y, et al. Deep Transfer Learning for Cross-domain Activity Recognition[J]. *ACM Press the 3rd International Conference-Singapore, Singapore (2018.07.28-2018.07.31), Proceedings of the 3rd International Conference on Crowd Science and Engineering, (ICCSE'18)*. 2018:1-8.
- 【25】 Santoro A, Bartunov S, Botvinick M, et al. One-shot learning with memory-augmented neural networks[J]. *arXiv preprint arXiv:1605.06065*, 2016.
- 【26】 Koch, Gregory, Richard Zemel, and Ruslan Salakhutdinov. "Siamese neural networks for one-shot image recognition." *ICML Deep Learning Workshop. Vol. 2*. 2015.

- 【27】 Snell, Jake, Kevin Swersky, and Richard Zemel. "Prototypical networks for few-shot learning." *Advances in Neural Information Processing Systems*. 2017.
- 【28】 Sung, Flood, et al. "Learning to compare: Relation network for few-shot learning." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.
- 【29】 Finn, Chelsea, Pieter Abbeel, and Sergey Levine. "Model-agnostic meta-learning for fast adaptation of deep networks." *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. JMLR. org, 2017.
- 【30】 SharmaN, MishraM, ShrivastavaM. Color image segmentation techniques and issues: an approach[J]. *International Journal of Scientific & Technology Research*, 2012, 1(4):9–12.
- 【31】 FelzenszwalbP F, HuttenlocherD P. Efficientgraph-based image segmentation[J]. *International journal of computer vision*, 2004,59(2):167–181.
- 【32】 Ning J, Zhang L, Zhang D, Wu C. Interactive image segmentation by maximal similarity-based region merging[J]. *Pattern Recognition*, 2010,43(2):445–456.
- 【33】 汤力琨, 罗代升, 王正勇, 龙建忠. 图像分割技术在金相分析中的应用[J]. *理化检验-物理分册*, 2005, 41(5):236-239.
- 【34】 裘俊彦, 潘安霞. 基于形态学的球墨铸铁金相组织分割算法研究[J]. *铸造*, 2014, 63(9):906-909.
- 【35】 Ananyev M, Medvedev D, Gavriilyuk A, Mitri S, DeminA, Malkov V, Tsiakaras P. Cu and Gd co-doped BaCeO₃ proton conductors: Experimental vs SEM image algorithmicsegmentation results[J]. *Electrochimica Acta*, 2014.125:371-379.
- 【36】 Lopez P, Lira J, Hein I. Discrimination of Ceramic Types Using Digital Image Processing by Means of Morphological Filters[J]. *Archaeometry*, 2015, 57(1):146-162.
- 【37】 Lai C, Song L, Han Y, Li Q, Gu H, Wang B, Qian Q, Chen W. Material image segmentation with the machine learning method and complex network method [OL]. *MRS Advances*, 2019, 1-6.
- 【38】 Qian Wang, Yuexing Han, Qing Li, Bing Wang, and Akihiko Konagaya. Segmenting overlap360 ping nano-objects in atomic force microscopy image. *Journal of Nanophotonics*, 12(1):016003, 2018.
- 【39】 Chen L, Han Y, Cui B, Guan Y, Luo Y. Two-Dimensional Fuzzy Clustering Algorithm (2DFCM)

- for Metallographic Image Segmentation Based on Spatial Information[C]. International Conference on Information Science and Control Engineering. IEEE, 2015:519-521.
- 【40】 赵墨, 丁广太, 樊明磊, 张惠然, 王路, 陈琳. 融合 LBP 纹理和局部灰度特征的材料图像分割[J]. 计算机技术与发展, 2016, 26(10):11-16.
- 【41】 Han Y, Lai C, Wang B, GuH. Segmenting images with complex textures by using hybrid algorithm[J]. Journal of Electronic Imaging, 2019, 28(1):013030.
- 【42】 S.M. Azimi, D. Britz, M. Engstler, M. Fritz, F. Mücklich. Advanced steel microstructure classification by deep learning methods. Sci. Rep., 8 (1) (2018).
- 【43】 Stukowski A 2012 Structure identification methods for atomistic simulations of crystalline materials Modelling Simul. Mater. Sci. Eng. 20 045021.
- 【44】 J.L. Priedeman, C.W. Rosenbrock, O.K. Johnson, E.R. Homer. Quantifying and connecting atomic and crystallographic grain boundary structure using local environment representation and dimensionality reduction techniques. Acta Mater., 161 (2018), pp. 431-443.
- 【45】 Steinhardt P J and Nelson D R and Ronchetti M 1983 Bond-orientational order in liquids and glasses Phys. Rev. B 28 784–805.
- 【46】 Ackland G J and Jones A P 2006 Applications of local crystal structure measures in experiment and simulation. Phys. Rev. B 73 054104.
- 【47】 Honeycutt J D and Andersen H C 1987 Molecular dynamics study of melting and freezing of small Lennard-Jones. clusters J. Phys. Chem. 91 4950–63.
- 【48】 Kelchner C L, Plimpton S J and Hamilton J C 1998 Dislocation nucleation and defect structure during surface. indentation Phys. Rev. B 58 11085.
- 【49】 Zhang Y., Jin R., Zhou Z.H. Understanding bag-of-words model: A statistical framework Int. J. Mach. Learn. Cybern., 1 (1–4) (2010), pp. 43-52.
- 【50】 Jorg Behler and Michele Parrinello, Generalized neural-network representation of high-dimensional potential-energy surfaces," Physical Review Letters, vol. 98, no. 14, pp. 146401, 2007.
- 【51】 Jorg Behler, "Perspective: Machine learning potentials for atomistic simulations," The Journal of Chemical Physics, vol. 145, no. 17, pp. 170901, 2016.
- 【52】 GPP Pun, R Batra, R Ramprasad, and Y Mishin, "Physically informed artificial neural networks

- for atomistic modeling of materials.," *Nature Communications*, vol. 10, no. 1, pp. 2339-2339, 2019.
- 【53】 Keinosuke Fukunaga, *Introduction to statistical pattern recognition*, Elsevier, 2013.
- 【54】 Anil K Jain, Robert PW Duin, and Jianchang Mao, "Statistical pattern recognition: A review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4-37, 2000.
- 【55】 Nojun Kwak and Chong-Ho Choi, "Input feature selection by mutual information based on parzen window," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 12, pp. 1667-1671, 2002.
- 【56】 Karl Pearson, "Liii. on lines and planes of closest fit to systems of points in space," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2, no. 11, pp. 559-572, 1901.
- 【57】 Harold Hotelling, "Analysis of a complex of statistical variables into principal components.," *Journal of Educational Psychology*, vol. 24, no. 6, pp. 417, 1933.
- 【58】 Joshua B Tenenbaum, Vin De Silva, and John C Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319-2323, 2000.
- 【59】 Sam T Roweis and Lawrence K Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323-2326, 2000.
- 【60】 Yvan Saeys, Inaki Inza, and Pedro Larrañaga, "A review of feature selection techniques in bioinformatics," *Bioinformatics*, vol. 23, no. 19, pp. 2507-2517, 2007.
- 【61】 Isabelle Guyon and Andre Elisseev, "An introduction to variable and feature selection," *Journal of Machine Learning Research*, vol. 3, no. Mar, pp. 1157-1182, 2003.
- 【62】 Pedro M Morgado, Margarida Silveira, Alzheimer's Disease Neuroimaging Initiative, et al., "Minimal neighborhood redundancy maximal relevance: Application to the diagnosis of alzheimer's disease," *Neurocomputing*, vol. 155, pp. 295-308, 2015.
- 【63】 Saurav Mallik, Tapas Bhadra, and Ujjwal Maulik, "Identifying epigenetic biomarkers using maximal relevance and minimal redundancy-based feature selection for multi-omics data," *IEEE Transactions on Nanobioscience*, vol. 16, no. 1, pp. 3-10, 2017.
- 【64】 Hanchuan Peng, Fuhui Long, and Chris Ding, "Feature selection based on mutual information

- criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226-1238, 2005.
- 【65】 Frank Rosenblatt, "Principles of neurodynamics. perceptrons and the theory of brain mechanisms," Tech. Rep., CORNELL AERONAUTICAL LAB INC BUFFALO NY, 1961.
- 【66】 Simon Haykin, *Neural networks: a comprehensive foundation*, Prentice Hall PTR, 1994.
- 【67】 Vinod Nair and Georey E Hinton, "Rectified linear units improve restricted boltzmann machines," pp. 807-814, 2010.
- 【68】 Giorgio Gnecco and Marcello Sanguineti, "Weight-decay regularization in reproducing kernel hilbert spaces by variable-basis schemes", *Computational Management Science*, vol. 8, pp. 625–634, 2009.
- 【69】 Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting", *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.
- 【70】 Francois Perreault, Andreia Fonseca De Faria, and Menachem Elimelech. Environmental applications of graphene-based nanomaterials. *Chemical Society Reviews*, 44(16):5861-5896, 2015.
- 【71】 Nan Gao and Xiaosheng Fang. Synthesis and development of graphene-inorganic semiconductor nanocomposites. *Chemical Reviews*, 115(16):8294-8343, 2015.
- 【72】 Yanmei Shi and Bin Zhang. Recent advances in transition metal phosphide nanomaterials: synthesis and applications in hydrogen evolution reaction. *Chemical Society Reviews*, 45(6):1529- 1541, 2016.
- 【73】 ShuilinWu, ZhengyangWeng, Xiangmei Liu, KWK Yeung, Paul Chu, et al. Functionalized TiO₂-based nanomaterials for biomedical applications. *Advanced Functional Materials*, 24(35):5464-5481, 2014.
- 【74】 Yin Zhang, T R Nayak, Hao Hong, and Weibo Cai. Biomedical applications of zinc oxide nanomaterials. *Current Molecular Medicine*, 13(10):1633-1645, 2013.
- 【75】 Guohai Yang, Chengzhou Zhu, Dan Du, Junjie Zhu, and Yuehe Lin. Graphene-like two-dimensional layered nanomaterials: applications in biosensors and nanomedicine. *Nanoscale*,

- 7(34):14217-14231, 2015.
- 【76】 Liu S P, Fang Y. A Contourlet domain image denoising method based on mathematical morphology[J]. *Guangzi Xuebao/acta Photonica Sinica*, 2008, 37(1):197-201.
- 【77】 Wang Haojun, Zheng Chongxun, and Ma Dong. Applying watershed algorithm to segmentation of color images from blood and bone marrow smears. *JOURNAL-XIAN JIAOTONG UNIVERSITY*, 35(12):1296-1299, 2001.
- 【78】 Jierong Cheng, Jagath C Rajapakse, et al. Segmentation of clustered nuclei with shape markers and marking function. *IEEE Transactions on Biomedical Engineering*, 56(3):741-748, 2009.
- 【79】 Qian Wang, Yuexing Han, Qing Li, Bing Wang, and Akihiko Konagaya. Segmenting overlapping nano-objects in atomic force microscopy image. *Journal of Nanophotonics*, 12(1):016003, 2018.
- 【80】 Chiwoo Park, Jianhua Z Huang, Jim X Ji, and Yu Ding. Segmentation, inference and classification of partially overlapping nanoparticles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(3):1-1, 2012.
- 【81】 Wang M, Wang T, Cai P, et al. Nanomaterials Discovery and Design through Machine Learning[J]. *Small Methods*, 2019.
- 【82】 Tsung-Han Chan, Kui Jia, Shenghua Gao, Jiwen Lu, Zinan Zeng, and Yi Ma. Pcanet: A simple deep learning baseline for image classification? *IEEE Transactions on Image Processing*, 24(12):5017–5032, 2015.
- 【83】 Yuexing Han, Akito Hara, Akinori Kuzuya, Ryosuke Watanabe, Yuichi Ohya, and Akihiko Konagaya. Automatic recognition of dna pliers in atomic force microscopy images. *New Generation Computing*, 33(3):253-270, 2015.
- 【84】 Joao C. Neves, Helena Castro, Ana Tomas, Miguel Coimbra, and Hugo Proença. Detection and separation of overlapping cells based on contour concavity for leishmania images. *Cytometry395 Part A*, 85(6):491–500, 2014.
- 【85】 Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62-66, 1979.
- 【86】 N Malpica, C. O. de Solorzano, J. J. Vaquero, A Santos, I Vallcorba, J. M. Garcia-Sagredo, and

- Pozo F Del. Applying watershed algorithms to the segmentation of clustered nuclei. *Cytometry Part B Clinical Cytometry*, 28(4):289-297, 2015.
- 【87】 Ron Kimmel, Nahum Kiryati, and Alfred M Bruckstein. Sub-pixel distance maps and weighted distance transforms. *Journal of Mathematical Imaging and Vision*, 6(2-3):223-233, 1996.
- 【88】 Yann LeCun, Leon Bottou, Yoshua Bengio, and Patrick Hader. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278-2324, 1998.
- 【89】 Andre Ricardo Backes, Dalcimar Casanova, and Odemir Martinez Bruno, "A complex network-based approach for boundary shape analysis", *Pattern Recognition*, vol. 42, no. 1, pp. 54-67, 2009.
- 【90】 Andr Ricardo Backes, Dalcimar Casanova, and Odemir Martinez Bruno, "Texture analysis and classification: A complex network-based approach", *Information Sciences*, vol. 219, no. 1, pp. 168-180, 2013.
- 【91】 Rafael C Gonzalez, Richard E Woods, et al., "Digital image processing [m]," Publishing house of electronics industry, vol. 141, no. 7, 2002.
- 【92】 Rui Xu and Donald C Wunsch, "Survey of clustering algorithms," 2005.
- 【93】 Jorg Sander, Martin Ester, Hans-Peter Kriegel, and Xiaowei Xu, "Density-based clustering in spatial databases: The algorithm gbscan and its applications", *Data mining and knowledge discovery*, vol. 2, no. 2, pp. 169-194, 1998.
- 【94】 Yizong Cheng, "Mean shift, mode seeking, and clustering", *IEEE transactions on pattern analysis and machine intelligence*, vol. 17, no. 8, pp. 790-799, 1995.
- 【95】 M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks", *Physical Review E Statistical Nonlinear Soft Matter Physics*, vol. 69, no. 2, pp. 026113, 2004.

作者在攻读硕士学位期间公开发表的论文与专利

- 【1】 Yuexing Han, Leilei Song, Bing Wang*, Sheng Sun*, Quan Qian, Qian Wang. AtomicNet: a novel approach to identify the crystal structure of each simulated atom. *Modelling and Simulation in Materials Science and Engineering*, 28(2020) 035005 (19pp), 2020. DOI: 10.1088/1361-651X/ab6da7
- 【2】 专利名称：基于复杂网络理论的图像分割方法，发明人：韩越兴、宋磊磊、王冰，申请号或专利号：202010083808.6

致 谢

硕士论文书写至此，近二十载求学生涯也邻近尾声。毕业在即，念起过往，心中涌起万千感慨，谨在此向帮助过我的老师同学、家人朋友致以诚挚的感谢！

首先，衷心感谢我的导师韩越兴老师，韩老师温婉谦和，平易近人，虽然工作繁忙，仍旧在科研工作以及生活上予以我们指导和关心，韩老师严谨治学，具有敏锐的学术洞察力，常叮嘱我们在学习上要对自己有更高的要求。韩老师的为人处事，对工作的忘我精神，对科研的孜孜不倦，对学生的真诚友善，都是我学习的榜样也激励着我不断向前。

同时，感谢王冰老师在课题研究上的帮助和指导，王老师总是能抓住问题的关键点，一针见血地指出我们科研工作中思虑不足的地方，并用易懂而恰当的比喻向我们讲解思路和方法。

另外，特别感谢上海大学材料基因组研究院孙升老师、顾辉老师研究团队提供的材料数据，以及在材料相关知识给予的指导和帮助。

还要感谢我的家人在我学业上，生活上支持，鼓励与帮助，是他们给予了我更多的信心和力量，攻坚克难，一步步走到今天。

最后感谢我们组内的所有师兄，同门，师弟和师妹们给予我的在生活中，科研工作中的指导和帮助。

再次感谢所有关心过我的老师、同学、家人、朋友们，谢谢你们。