

分类号: _____

密 级: _____

U D C: _____

单位代码: _____

安徽工业大学

硕士学位论文

论文题目: 基于深度学习的煤岩镜质组分类



学 号: _____

作 者: _____

专 业 名 称: _____ 控制工程

2021 年 10 月 12 日

安徽工业大学硕士学位论文

论文题目：基于深度学习的煤岩镜质组分类

Classification of Vitrinite in Coal Based on Deep Learning

作 者： _____ 学院： _____

指 导 教 师： _____ 单位： _____

协助指导教师： _____ 单位： _____

_____ 单位： _____

论文提交日期： 2021 年 10 月 12 日

学位授予单位： 安 徽 工 业 大 学

安徽马鞍山 243002

独 创 性 说 明

本人郑重声明：所呈交的论文是我个人在导师指导下进行的研究工作及取得研究成果。尽我所知，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经发表或撰写过的研究成果，也不包含为获得安徽工业大学或其他教育机构的学位或证书所使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中做了明确的说明并表示了谢意。

签名：_____日期：_____

关于论文使用授权的说明

本人完全了解安徽工业大学有关保留、使用学位论文的规定，即：学校有权保留送交论文的复印件，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，可以采用影印、缩印或其他复制手段保存论文，保密的论文在解密后应遵循此规定。

签名：_____导师签名：_____日期：_____

摘要

镜质组作为煤岩显微结构的三大组别之一，其粘结性好且挥发分产率较高使其成为炼焦的优质原料，并且根据镜质组的反射率可以初步判断煤种以鉴别混煤。因此，实现煤岩镜质组显微组分高效、自动的分类具有重要意义。然而，由于镜质组中部分显微亚组分类别间特征差异较小易混淆、局部特征甚至有交错，使得采用特征分类方法对镜质组显微组分的识别较为困难。深度学习具有强大的特征表达能力，能够自动且有效的提取特征，但是深度学习模型一般规模较大、计算量较大，对训练样本数据量要求高，对硬件设备要求较高，难以应用到实际的工业生产场景如移动设备或嵌入式设备上。因此本文在深度学习的基础上，构建轻量化的煤岩镜质组分类模型，并采取有效措施提升模型准确率，完成对煤岩镜质组显微图像的自动分类。本文的主要工作如下。

(1) 在充分调研相关国内外文献的基础上，详细分析了煤岩显微组分、图像分类算法及深度学习理论的研究现状，阐述本文的研究背景和意义。

(2) 针对数据集样本数不充足的情况，在分析煤岩显微组分样本特征的基础上，对煤岩镜质组显微图像进行数据扩充操作，增加样本数量，为后续深度学习模型的训练提供数据基础。

(3) 针对深度卷积神经网络模型一般规模较大、参数较多，难以应用于内存有限的设备中等问题，本文以经典的深度卷积神经网络模型 DenseNet121 为基础进行改进，采用深度可分离卷积代替标准卷积并优化增长率，以降低模型大小，实现模型轻量化。同时，在实现模型轻量化的基础上，保持或提升模型准确率，本文在该模型中引入了注意力机制，以获取更为关键的特征信息，提高模型分类准确率，构建出适用于煤岩镜质组分类的轻量化模型。

(4) 为提高经典轻量级模型的分类准确率，本文以轻量级网络为基础网络，结合煤岩镜质组显微组分的特点，对模型进行改进。引入迁移学习及改进的注意力机制，以提高轻量级模型的分类准确率。通过各类评价指标对模型进行评估，验证模型的有效性，实现煤岩镜质组显微组分的高准确率分类。

本文的特色与创新之处在于：将深度学习引入煤岩镜质组显微图像分类任务中，避免了传统分类方法的特征抽取与降维等问题；构建基于深度卷积神经网络轻量化的煤岩镜质组分类模型，保持分类准确率的同时实现模型轻量化，解决深度学习模型规模大难以应用于内存有限设备中的问题；在轻量级模型的基础上进行改进，引入迁移学习和改进的注意力机制，提高了经典轻量级模型的分类准确率。本文的研究对于煤岩镜质组显微图像的分类具有一定的参考价值和实用价值。

关键词：镜质组，深度学习，注意力机制，轻量化模型，分类

Abstract

Vitrinite, as one of the three major components of coal microstructure, its good cohesiveness and high volatile component yield makes it be high-quality raw materials for coking. According to the reflectivity of the vitrinite maceral, the coal type can be preliminarily judged to identify mixed coal. Therefore, it is of great significance to realize the efficient and automatic classification of vitrinite macerals. However, due to the small difference in features between some maceral categories in the vitrinite, it is easy to be confused, and local features are even intertwined, which makes the identification of vitrinite macerals a challenge. Deep learning has powerful feature expression capabilities, which can automatically and effectively extract features. However, deep learning models are generally large in scale and computationally expensive, and require huge size sample set and high hardware equipment, are difficult to apply to actual industrial production scenarios such as mobile devices or embedded devices. Therefore, this dissertation builds a lightweight coal vitrinite classification model on the basis of deep learning, and improves the accuracy of the model by some effective measures, and completes the automatic classification of the vitrinite microscopic images of coal. The main work of this dissertation is as follows:

(1) Based on the full investigation of relevant domestic and foreign literature, this dissertation analyzes the research status of coal macerals, image classification algorithm and deep learning theory in detail, and expounds the research background and significance of this dissertation.

(2) In view of the insufficiency of sample data of the set, based on the analysis of the characteristics of the coal macerals samples, data expansion of the coal vitrinite microscopic image is performed to increase the number of samples and provide a data basis for subsequent model training.

(3) In view of the fact that deep convolutional neural network models are generally large in scale and have many parameters, they are difficult to apply to devices with limited memory. Based on the classical deep convolutional neural network model DenseNet121, this dissertation improved the model by using depth separable convolution to replace the standard convolution and optimize the growth rate, so as to reduce the model size and realize the model lightweight. At the same time, on the basis of the realization of model lightweight, the model accuracy was maintained or improved. Besides, attention mechanism was introduced into the model to obtain more critical feature information, improve the model classification accuracy, and build a lightweight model suitable for coal vitrinite classification.

(4) In order to improve the classification accuracy of the classic lightweight model, this dissertation uses the lightweight network as the basic network and combines the characteristics of the vitrinite macerals of coal to improve the model. Transfer learning

and improved attention mechanism are introduced to improve the classification accuracy of lightweight models. The model is evaluated by various evaluation indicators to verify the validity of the model and realize the classification of vitrinite macerals with high accuracy.

The special and innovation of the dissertation lies in: Deep learning is introduced into the microscopic image classification of coal vitrinite, which avoids some problems such as feature extraction and dimension reduction. In order to solve the problem that the large scale of deep learning models is difficult to apply to devices with limited memory, a lightweight coal vitrinite classification model based on deep convolutional neural networks is constructed to maintain classification accuracy while achieving lightweight model; In order to improve the classification accuracy of the classic lightweight model, a transfer learning and improved attention mechanism are introduced based on the lightweight model. The research in this dissertation has certain reference value and practical value for the classification of vitrinite microscopic images of coal.

Key Words: vitrinite, deep learning, attention mechanism, lightweight model, classification

目 录

| | |
|--|----|
| 摘 要..... | I |
| Abstract..... | II |
| 目 录..... | IV |
| 第一章 绪论..... | 1 |
| 1.1 课题的研究背景及意义..... | 1 |
| 1.2 课题研究对象..... | 1 |
| 1.3 国内外研究现状..... | 2 |
| 1.3.1 煤岩显微组分研究现状..... | 2 |
| 1.3.2 传统图像分类算法研究现状..... | 3 |
| 1.3.3 深度学习在图像分类中的研究现状..... | 4 |
| 1.4 论文的主要内容与安排..... | 5 |
| 第二章 镜质组显微图像分析与数据集构建..... | 7 |
| 2.1 镜质组显微图像分析..... | 7 |
| 2.1.1 镜质组显微组分..... | 7 |
| 2.1.2 镜质组显微图像特征分析..... | 9 |
| 2.2 镜质组显微图像数据集构建..... | 9 |
| 2.2.1 镜质组显微图像初始数据集..... | 9 |
| 2.2.2 镜质组显微图像数据增强..... | 10 |
| 2.3 本章小结..... | 11 |
| 第三章 基于 DenseNet121 模型轻量化的煤岩镜质组分类..... | 12 |
| 3.1 卷积神经网络结构..... | 12 |
| 3.2 DenseNet 模型结构..... | 16 |
| 3.3 DenseNet121 模型的轻量化改进..... | 18 |
| 3.3.1 深度可分离卷积..... | 18 |
| 3.3.2 注意力机制..... | 20 |
| 3.3.3 改进后的模型结构..... | 23 |
| 3.4 实验与结果讨论..... | 25 |
| 3.4.1 实验环境与参数设置..... | 25 |
| 3.4.2 评价指标..... | 26 |
| 3.4.3 实验结果与分析..... | 27 |
| 3.4.3.1 改进的 DenseNet121 模型实验结果与分析..... | 27 |
| 3.4.3.2 增长率优化实验结果与分析..... | 29 |

| | |
|---|----|
| 3.4.3.3 消融实验结果与分析 | 30 |
| 3.4.3.4 不同模型对比实验结果与分析 | 30 |
| 3.5 本章小结 | 33 |
| 第四章 基于注意力机制的轻量级网络煤岩镜质组分类 | 34 |
| 4.1 轻量级网络模型 | 34 |
| 4.2 ShuffleNet V2 模型结构 | 35 |
| 4.3 ShuffleNet V2 模型的改进 | 36 |
| 4.3.1 迁移学习 | 36 |
| 4.3.2 改进的注意力机制 | 37 |
| 4.3.3 改进后的模型结构 | 38 |
| 4.4 实验与结果讨论 | 38 |
| 4.4.1 改进的 ShuffleNet V2 模型实验结果与分析 | 39 |
| 4.4.2 模型收敛性对比实验结果与分析 | 39 |
| 4.4.3 不同注意力机制的模型对比实验结果与分析 | 40 |
| 4.4.4 不同模型对比实验结果与分析 | 41 |
| 4.5 本章小结 | 43 |
| 第五章 总结与展望 | 44 |
| 5.1 总结 | 44 |
| 5.2 展望 | 44 |
| 参考文献 | 46 |
| 在学研究成果 | 52 |
| 致 谢 | 53 |
| 附录 A 插图清单 | 54 |
| 附录 B 插表清单 | 55 |

第一章 绪论

1.1 课题的研究背景及意义

煤炭在不同化石燃料中成本较低且资源丰富,因此煤炭将长期保持在一次能源生产和使用中的重要地位^[1]。由于煤炭资源不可再生且优质炼焦煤资源短缺,清洁高效的利用煤炭资源对我国能源发展策略的实现具有重要意义^[2-3]。相关研究表明,显微组分不仅与煤的反应性、粘结性、结焦性等工艺性质密切相关,而且影响着配煤结构的优化和焦炭质量的预测^[4-7]。因此,实现煤岩显微组分的自动、准确的识别对煤炭资源的合理高效利用具有重要意义。

煤岩显微组分的分类和命名标准各个国家略有不同,目前我国主要以《烟煤显微组分分类■GB/T15588-2013》作为煤岩显微组分的分类标准,该标准按照成因和性质的不同,将煤划分为3个主要的显微组分:镜质组、壳质组和惰质组^[8]。其中镜质组因含量最高成为三大组分中最主要的显微组分。与其他两种组分相比,镜质组具有较高的氧含量、中等的氢含量和较低的碳含量,在加热过程中能熔融粘结,因此其粘结性好且挥发分产率较高,是炼焦的优质原料^[9-10]。镜质组的反射率是表征煤变质程度的重要指标,根据镜质组反射率可以初步判断煤种鉴别混煤^[11-12]。因此实现煤岩显微组分镜质组的自动分类具有重要意义。目前已有大量的学者采用人工分析的方法识别煤岩显微组分,但这种方法存在着对测定者专业要求高、劳动强度大和分析成本高等诸多局限。随着人工智能技术的迅猛发展,在图像分类领域深度学习获得了突破性的进展。深度学习中的卷积神经网络可以自动提取图像特征并分类,实现端到端的自动识别,本课题采用深度学习的方法实现煤岩显微组分的自动分类和识别。

本文研究内容来源于国家自然科学基金项目:炼焦煤显微组分自动识别及显微组分组成对焦炭质量的影响(51574004)。

1.2 课题研究对象

本课题主要对煤的三大显微组分之一的镜质组显微组分进行自动识别研究。根据分类标准,镜质组可分为结构镜质体-1、结构镜质体-2、团块镜质体、均质镜质体、基质镜质体、胶质镜质体和碎屑镜质体共7类显微组分和显微亚组分^[8]。在显微镜观察下,各显微组分及显微亚组分表现出不同的形态结构和纹理特征。本文以煤岩镜质组显微图像为研究对象,采用深度学习的方法对煤岩显微组分镜质组显微图像进行分类,并通过各类评价指标分析模型性能和分类结果,最终实现煤岩镜质组显微组分端到端的自动分类。

1.3 国内外研究现状

1.3.1 煤岩显微组分研究现状

煤岩的化学、物理和工艺性质与煤的显微组分密切相关^[13-14], 因此研究煤的显微组分可以更加深入地了解煤的性能和成因, 进而实现煤的高效合理的应用。根据煤岩显微组分性质的不同, 煤岩显微组分常见的测定方法可分为性质指标法、浮沉离心分离技术、国标数点法和图像分析法^[15]。性质指标法通过构建显微组分含量与煤质指标之间的表达式对显微组分进行测定, 该方法易操作, 但结果精度不高; 浮沉离心分离显微组分技术, 其采用等密度梯度分离法实现, 该方法实施方便、操作简单, 但该方法不能充分分离各组分, 准确度不高; 国标数点法是目下使用最普遍的煤岩显微组分含量测定的方法, 该方法虽然结果较精确, 但存在着对测定者专业要求高、劳动强度大和分析成本高等诸多局限。

针对以上方法存在的缺陷, 国内外研究人员提出并采用图像分析法测定煤岩显微组分, 以降低操作人员劳动强度, 提高测定速度, 减少分析成本等问题, 以上优势使图像分析法成为煤岩显微组分分析的重要发展方向。二十世纪七十年代, England 等^[16]使用 Quantimet720 图像分析仪测定煤岩显微组分的反射率, 并用反射率数据区分显微组分, 该方法是图像分析法在煤岩显微组分分析的初步探索, 但因当时条件限制并未实现完全的自动化。Riepe 等^[17]采用 Leitz 图像分析仪测定煤岩显微组分组成, 将图像的像素按 0-255 的灰度范围划分, 由于显微组分不同其灰度也不同, 该方法根据这种灰度差异进行区分, 其缺点是壳质组和树脂在透射光下灰度值相近, 导致识别时存在一定的缺陷。Lester 等^[18]开发了一个图像分析系统可以根据形态学特征和灰度特征对主要显微组分分析, 该系统对壳质组和树脂实现了很好的识别, 但对其他较复杂的显微图像分析效果并不明显。虽然以上工作取得了一定的成果, 但由于不同显微组分的反射率存在细微差异, 仅使用灰度值进行区分并不合理。

为了更全面的描述煤岩显微组分图像的信息, 需要结合更多的特征进行分析, 如形状, 纹理等。王素婷等^[19]将 RILBP 和 GLCM 算法融合提取煤岩显微组分的纹理特征, 再构建多分类支持向量机完成煤岩显微组分的识别。Skiba 等^[20]采用直方图和灰度共生矩阵确定惰质组图像的统计参数, 并利用主成分分析提取主要特征, 结合多层感知机实现煤岩惰质组的有效识别。Wang 等^[21]采用 k-means 聚类算法对煤岩显微图像进行分割并分析特征, 再使用随机森林完成煤岩显微组分的分类识别。近年来, 本课题组对煤岩显微组分图像分析进行了深入的研究, 并取得了一定的成果: 如构建圆形 LBP(Local Binary Pattern)及其均匀模式提取镜质组的纹理特征量, 并分析镜质组各显微组分的纹理特征^[22]; 提出离散小波变换的方法提取煤岩显微图像频域的特征信息, 并完成惰质组的识别^[23];

基于灰度共生矩阵和灰度分布统计对煤岩显微图像进行特征提取,然后采用 PCA 降维,最后搭建支持向量机完成镜质组中三个显微组分的分类^[24];采用 Contourlet 变换提取煤岩显微图像特征量,再采用超限学习机完成对煤岩壳质组的分类识别^[25];提出曲波变换和压缩感知的特征提取及降维的方法构成特征量集,并用支持向量机实现对煤岩惰质组显微组分的分类^[26];采用多重分形去趋势波动分析算法提取多重分形特征,分析惰质组图像复杂的纹理特征^[27]。以上研究工作对煤岩显微组分的识别与分类均做出了一定的贡献,但这些传统方法在对煤岩显微组分进行特征量提取时需要人为干预,存在着提取信息量冗余且泛化能力弱等缺点,难以实现特征提取和分类器紧密联系的端对端的完整自动分类。

1.3.2 传统图像分类算法研究现状

图像分类是一种根据某些已知图像信息反映的特征来划分图像类别的有效方法。图像分类是实现检测、定位和分割等应用的基础,也是计算机视觉领域的主要研究方向之一^[28]。目前国内外主流的传统分类算法包括支持向量机、K 最近邻分类器、朴素贝叶斯分类器以及随机森林等。

(1) 支持向量机是通过最大化分离超平面和数据之间的裕度,从而最小化泛化误差的上界,该算法被广泛应用于小样本分类的任务^[29]。王宏漫等^[30]通过使用 PCA/ICA 提取人脸面部图像特征,提出分阶段淘汰的支持向量机分类机制实现人脸识别,获得了较高的识别精度。张浩等^[31]通过分析遥感图像的几何特征、空间特征以及纹理特征,提出基于支持向量机的建筑物图像分类方案,得到了较高的准确率。

(2) K 最近邻分类器是根据已知数据通过其最邻近的样本划分未知样本类别的算法^[31]。Ameur 等^[33]通过融合 LBP 和 Gabor 小波完成特征提取,降维后结合 K 近邻和 SCR 分类器实现人脸识别,提出了一种非受控环境下的人脸识别方法,获得了较好的识别结果。Sun 等^[34]使用 K 最近邻分类器和 Hadoop 在云环境中进行面部图像的标记与分类,并评估处理时间和识别率,证明了该方法在性能上显著提高。

(3) 朴素贝叶斯分类器是基于最大概率原理的分类算法。主要是通过贝叶斯准则转换待分类样本在各类别中的后验概率来计算先验概率,最后得到概率最大的一类作为该样本的类别。石洪波等^[35]为改善朴素贝叶斯的性能,通过分析朴素贝叶斯模型及其变异形式模型的结构特点,提出了一种基于贝叶斯定理的新分类模型,取得了较好的分类效果。Zhang 等^[36]提出一种朴素贝叶斯最近邻和协作表示框架下的图像分类算法,该算法较好缓解了分类时产生的参数过拟合和训练数据稀缺的问题。

(4) 随机森林是包含多个决策树的分类器,其采用组合多棵树对样本进行

训练并预测,通过投票得到最终的预测结果。郭金鑫等^[37]为提高人脸识别率,首先通过梯度直方图提取特征,然后采用特征降维算法对特征降维,最后构建随机森林进行分类。刘毅等^[38]针对遥感影像分类中难以选择合适分类器的问题,将集成学习的思想以及随机森林的方法引入到小卫星遥感影像分类,结果获得了较高的分类精度。

传统的图像分类算法主要包括特征提取和分类器的构建部分。特征提取部分需要人工选择特征,包括颜色、纹理以及形状等特征,针对不同的图像分类任务需要人工选取不同的特征。对于一些复杂的图像,人工选取合适的特征和分类器较为困难,因此分类结果受人为因素影响较大,具有一定的局限性。

1.3.3 深度学习在图像分类中的研究现状

深度学习是机器学习领域一个新的研究方向,与传统的图像分类算法相比,深度学习不需要人工的提取和描述图像特征,而是通过神经网络自动地学习图像中更丰富、抽象的特征,并且这些特征与分类器紧密联系,解决了人工特征提取和分类器选择困难的问题,是一种端到端的自动分类模型^[39]。并且深度学习善于从原始数据中,挖掘出越来越抽象的特征,使学习到的特征更具表达力。

深度学习源于人工神经网络,由多个多层感知机组成。多层感知机^[40]、卷积神经网络^[41]以及递归神经网络^[42]等是深度学习的代表性网络。其中卷积神经网络作为计算机视觉领域的重要结构,被广泛应用于图像分类任务中。1962年,神经学家 Hubel 和 Wiesel^[43]提出了感受野的概念,说明了视觉是由从视网膜输入,后经过各层感受野的激发而形成的,这成为卷积神经网络的研究基础。1980年, Fukushima^[44]提出了卷积神经网络的早期模型即神经认知机。20世纪90年代贝尔实验室的 LeCun 等^[45]提出了基于梯度学习的 CNN 模型 LeNet-5 网络模型,该模型通过卷积层和池化层提取图像的特征,借助全连接层结构对提取到的特征进行分类,但其在大规模数据集分类任务中并未取得很好的效果。之后研究人员提出诸多方法对卷积神经网络进行改进。2012年 Krizhevsky 等^[46]提出了 AlexNet 网络,该网络模型结果领先于之前所有方法。随着该网络的成功,卷积神经网络被广泛的应用于图像分类领域,深度学习迅猛发展,并在之后的几年中不断有各种新的网络模型被提出。如 VGGNet^[47]、ResNet^[48]、GoogleNet^[49]和 DenseNet^[50]等网络模型的提出,使模型性能更加优越,模型层数也越来越深,图像识别准确率也得到了显著的提升。近年来,深度学习模型在图像分类领域发挥着重要的作用并被广泛的应用于图像分类的各个领域。何雪英等^[51]为提高乳腺癌病理图像分类精度,构建深度卷积神经网络模型,最终分类准确率可达 91%。刘大伟等^[52]针对高分辨率遥感影像分类任务,构建深度信念网络(DBN)实现分类,与浅层学习算法相比,该算法分类效果更好。张泽中等^[53]对 GoogLeNet 模型进行的优化的

基础上融合不同结构的模型，实现对胃癌病理图像的分类。李彦枝等^[54]提出将 VGG16 和稠密链接思想融合，设计出一种 Dense-VGG 网络，完成极光图像的分类。

尽管这些经典的网络模型实现了分类准确率的提升，但一般规模较大、训练时间较长，难以应用到内存有限的设备上。因此研究规模小、速度快的轻量级架构至关重要。目前轻量级网络已成为 CNN 领域的研究热点，一系列轻量级网络被提出，包括 SqueezeNet^[55]、MobileNet V2^[56]、ShuffleNet V2^[57]等。李坤伦等^[58]利用 SqueezeNet 及特征融合策略实现了对车辆颜色的识别。Rabano 等^[59]使用 MobileNet 和迁移学习策略实现了垃圾分类。刘梦伦等^[60]针对肝硬化识别问题，结合 MobileNet V2 与超限学习机进行识别，结果表明该方法具有较快的运行速度。虽然这些轻量级模型在模型尺寸和速度方面有很大的优势，但是相较于深度卷积神经网络模型其准确率有待进一步提高，因此研究精度高且规模小的模型至关重要。

综上所述，目前深度学习已在图像分类中应用广泛，并取得了良好的效果。由于镜质组显微图像特征信息丰富，深度学习可以自主的学习到镜质组显微图像中更高维、更抽象的特征信息，且解决了人工特征提取和分类器选择困难等问题，实现了端到端的自动分类模型。因此，本文选择引入深度学习的方法对镜质组显微图像进行分类。

1.4 论文的主要内容与安排

本课题以实现镜质组显微图像分类识别为目的，构建深度学习分类模型并根据镜质组显微图像样本对模型进行改进，实现镜质组显微图像各组分更理想的分类效果。本文的章节安排如下。

第一章 绪论。首先阐述了煤岩镜质组分分类的选题背景和研究意义，然后说明了本文的研究对象，最后介绍了相关的国内外研究现状，包括煤岩显微组分研究现状、传统图像分类算法研究现状以及深度学习在图像分类任务中的研究现状。

第二章 镜质组显微图像分析与数据集构建。该部分详细介绍了本课题的研究对象即镜质组各显微组分并对显微图像特征进行了分析，接着对镜质组显微图像数据集的构建进行分析，最后对镜质组显微图像进行数据扩充以增加样本集数量，为本文后续研究提供数据基础。

第三章 基于 DenseNet121 模型轻量化的煤岩镜质组分分类。本章为实现深度神经网络模型的轻量化，以 DenseNet121 模型为主干网络进行改进，设计出 DE-DenseNet121 煤岩镜质组分分类模型。首先介绍 DenseNet 模型的结构，并分析其模型特点。接着将深度可分离卷积和注意力机制引入 DenseNet121 模型，使改进后的模型在保证一定分类准确率的基础上，进一步减小模型的参数量和计算量，实

现模型的轻量化。最后介绍了实验环境、参数设置和评价指标等，并对该分类模型实验结果进行相关的实验分析，验证 DE-DenseNet121 模型的有效性。

第四章 基于注意力机制的轻量级网络煤岩镜质组分类。本章为提高经典轻量级模型的分类准确率，在现存的经典轻量级模型 ShuffleNet V2 上进行改进，提出 TA-ShuffleNet V2 煤岩镜质组分类模型。首先介绍并分析了 ShuffleNet V2 模型的结构及特点。接着引入迁移学习和改进的高效轻量化注意力机制对模型进行改进，使改进后的模型准确率大幅度提升。最后通过实验验证了该方法在煤岩镜质组数据集上的可行性。

第五章 总结与展望。本章对论文的研究内容进行总结，归纳并分析了论文的创新性与局限性，同时对下一阶段的研究方向进行展望。

第二章 镜质组显微图像分析与数据集构建

深度学习模型对数据的依赖性较强,通常需要一定规模及完备性较好的数据集才能达到很好的识别效果。因此本章对煤岩镜质组显微图像进行介绍和分析,并针对数据集样本数不充足的情况,在分析煤岩显微组分样本特征的基础上,对煤岩镜质组显微图像进行数据扩充操作,以增加样本集数量为之后的研究工作提供数据基础。

2.1 镜质组显微图像分析

2.1.1 镜质组显微组分

镜质组中包含着不同的显微组分,它们是经腐殖化作用和胶凝化作用转化的植物木质纤维组织而形成的^[8]。低煤级到高煤级中,在光学显微镜下观察时,镜质组在油浸反射光下的反射色渐渐变浅,表现为深灰色到灰白色变化。镜质组中各显微组分都具有特定的工艺性质和化学性质,表现出不同的纹理特征和形态结构。根据植物细胞结构的大小、形状、保存情况以及成因等特点,镜质组可分为3个显微组分,分别为结构镜质体、无结构镜质体和碎屑镜质体,这3个显微组分又可进一步细分为若干显微亚组分,具体分类如表2.1所示。

表 2.1 镜质组显微组分

| 显微组分组 | 显微组分 | 显微亚组分 |
|-------|--------|---------|
| 镜质组 | 结构镜质体 | 结构镜质体-1 |
| | | 结构镜质体-2 |
| | 无结构镜质体 | 均质镜质体 |
| | | 基质镜质体 |
| | | 团块镜质体 |
| | | 胶质镜质体 |
| | 碎屑镜质体 | |

(1) 结构镜质体

通过显微观察,结构镜质体可以直接看到植物细胞结构。由于显微镜下,结构镜质体的植物细胞结构的保存情况不同,故将其进一步细分为两个显微亚组分即结构镜质体-1和结构镜质体-2:

1) 结构镜质体-1:细胞结构规则排列且保存完好,细胞壁表现为原始状态或微膨胀状态,胞腔清晰可见,多表现为椭圆形、方形以及纺锤形等,胞腔中空或伴有填充物。典型显微图像如图2.1(a)所示。

2)结构镜质体-2: 细胞结构保存情况较差,细胞壁完全膨胀,胞腔全部变形,仅能观察到细胞结构残留的痕迹,闭合后的细胞腔通常表现为线条状。典型显微图像如图 2.1(b)所示。

(2) 无结构镜质体

通过显微观察,无结构镜质体无法直接看到细胞结构。依据其成因和形态特征差异,无结构镜质体可以进一步细分为以下 4 个显微亚组分:

1) 均质镜质体: 主要形成于植物的树皮、树叶以及木质部等组织的强凝胶化作用,观察不到植物细胞结构。表现为透镜状或宽窄不等的条带状,成分纯净均一。其典型显微图像如图 2.1(c)。

2) 基质镜质体: 主要由植物组织经受强烈的凝胶化作用形成,观察不到任何细胞结构痕迹。该显微组分含量在镜质组中占比最多,其形态不固定,经常与其他显微组分或矿物质结胶。如图 2.1(d)。

3) 团块镜质体: 形成于植物细胞壁分泌的鞣质或腐植凝胶,主要呈纺锤形、圆形和椭圆形单独分布或集合状分布。大小与细胞腔大小近似,多为 20~100 μm ,较大者约为 150~200 μm 。其内部均一,边界轮廓清晰。如图 2.1(e)。

4) 胶质镜质体: 由腐植凝胶凝聚变化而成,常填充于植物细胞腔、裂隙或其他空腔中。无确定形态,均一纯净。典型显微图像如图 2.1(f)。

(3) 碎屑镜质体:

即镜屑体,是粒径小于 10 μm 的镜质组碎屑,其形态大多数表现为不规则状或粒状,少数表现为棱角状。与碎屑惰质体等混淆聚集时易于辨认,被基质镜质体胶结时不易辨认。典型显微图像如图 2.1(g)。

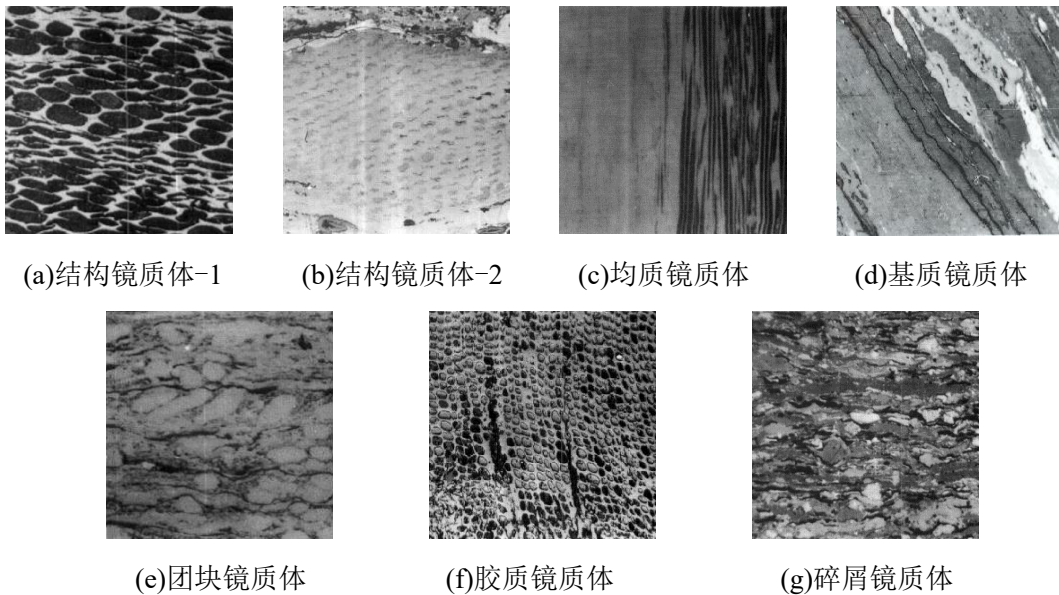


图 2.1 部分镜质组显微组分典型图像

本文将对以上分析的镜质组中 7 类显微组分及显微亚组分进行自动分类研

究。

2.1.2 镜质组显微图像特征分析

通过观察图 2.1 油浸反射光下典型显微图像，可以看出镜质组显微图像具有多样性和复杂性，以下主要从亮度、形状以及纹理三个方面进行特征分析。

亮度是指图像的明暗程度。在油浸反射光下，镜质组反射色表现为深灰色到灰白色渐渐变浅。其中均质镜质体在油浸反射光下表现为深灰色至灰色，基质镜质体油浸反射光下表现为暗灰色，部分显微组分的亮度特征相近。因此，只分析亮度特征并不能直观的区分镜质组显微图像。

形状作为图像的基本特征，因其具备旋转、尺度变化等不敏感特性，可以有效地表达图像的直观信息^[61]。观察镜质组显微图像形状特征，均质镜质体表现为宽窄不等的条带状，团块镜质体主要表现为圆形、纺锤形和椭圆形，碎屑镜质体多表现为不规则状或粒状，部分显微组分没有固定形态，常与其他显微组分或矿物质结胶。因此，仅分析镜质组显微组分的形状特征并不充分。

纹理描绘了图像中反复呈现的局部模式与他们的排列规则^[62-63]。观察镜质组显微图像表现出明显的纹理特征，结构镜质体-1 的细胞结构保存完好且细胞排列规则，团块镜质体呈纺锤形、圆形和椭圆形单独排列或集合状排列，均质镜质体以透镜状或宽窄不等的条带状排列。但是纹理仅仅可以描述物体的表面特性，不可以表达物体的本质特性，所以只用纹理特征对图像进行分析并不合理。

通过特征分析可知，镜质组中不同显微组分的亮度、形状以及纹理等特征具有一定差异，但又存在一定的相关性，且部分显微亚组分类别间特征差异较小易混淆、局部特征甚至有交错。若使用传统的人工分析方法鉴别镜质组各显微组分，不仅耗时长、劳动强度大，而且需要专业能力强的测定者。针对传统人工分析方法的缺陷与不足，基于图像分析的方法应运而生，目前主要是采用传统的图像分类方法对煤岩显微组分进行分类。传统的图像分类方法主要包括特征提取与分类识别两部分，在特征提取阶段通常采用以上分析的亮度、形状、纹理等方面提取图像特征信息，但诸多特征量简单的联合会产生特征冗余，影响分类准确率，且这些特征需要先验知识和人为的选择，缺乏特征的自学习能力。相对于传统的图像分类方法，深度学习中的卷积神经网络能自动提取特征并分类，并且具有很强的特征表达能力，能够完成端到端的自动识别。基于以上优点，本文采用深度学习的方法构建分类模型实现煤岩镜质组的自动分类。

2.2 镜质组显微图像数据集构建

2.2.1 镜质组显微图像初始数据集

本文镜质组显微图像数据集在油浸反射光下通过光学显微镜获取，由结构镜

质体-1、结构镜质体-2、均质镜质体、基质镜质体、团块镜质体、胶质镜质体和碎屑镜质体^[8]共 7 种不同类别的镜质组显微图像组成，每个类别有 60 幅的显微图像，共 420 幅有效图像。由于本文后续构建多种分类模型进行对比，为公平比较将图片输入大小统一设置为 224×224。数据集是否包含足够丰富的样本是深度学习中最基础最重要的部分，如果数据集过少往往难以训练参数量巨大的大型深度卷积神经网络。由于模型训练数据不足，会导致训练出现过拟合现象，最终难以达到较好的分类效果^[64]。虽然本文所提出的模型均是轻量化网络模型，相较于深度卷积神经网络而言模型参数量较少，较为适用于样本较少的分类任务，但是本文数据集的数据规模仍然较小，因此需要对训练数据进行数据扩充。

2.2.2 镜质组显微图像数据增强

本文以 2:1 的比例将煤岩镜质组显微组分样本分为训练集和测试集，即将其中 280 幅显微图像划分为训练集，140 幅划分为测试集。并对训练集采用数据增强方式进行数据扩充，即利用现有训练集的数据，通过增强变换获得更丰富的信息。因此，为了有较多的样本参与训练，使得网络模型对待识别图像的观察角度以及噪声影响因素不敏感，提高所设计的煤岩镜质组分类模型的泛化能力。本文采用以下几种数据增强的方法来增加训练数据，主要使用随机旋转、水平翻转以及添加椒盐噪声操作完成对训练数据集的增强。

(1) 图像角度旋转

图像旋转是指绕图像中心点，旋转任意角度。本文实验为了防止图像出现角度变换问题，选取顺时针 90 度、180 度以及 270 度三个角度对图像进行旋转。

(2) 图像水平翻转

图像翻转与旋转不同，是指对图像进行镜面翻转。本文将图像进行左右水平镜面翻转以扩充数据集。

(3) 添加椒盐噪声

椒盐噪声类似于椒盐的黑白点随机在图像上分布，在采集图像时可能因为各种各样的干扰因素而引入图像噪声，为了使镜质组显微图像更加多样，有效提高深度学习模型的泛化性能，在图像中随机添加椒盐噪声以扩充数据集。

增强后的训练集包含 1260 幅图像，共 1400 幅图像。由于煤岩显微图像是具有旋转不变性，煤岩学专家可以从不同角度分析煤岩显微图像辨别其组分。而旋转和加噪后的图像对神经网络模型而言却是一幅与之前不一样的图像，但其标签一致，因此增加了样本的多样性。在数据增强后，镜质组分类模型的泛化能力更强。以煤岩镜质组显微组分的碎屑镜质体为例，经数据增强后的样本如图 2.2 所示。

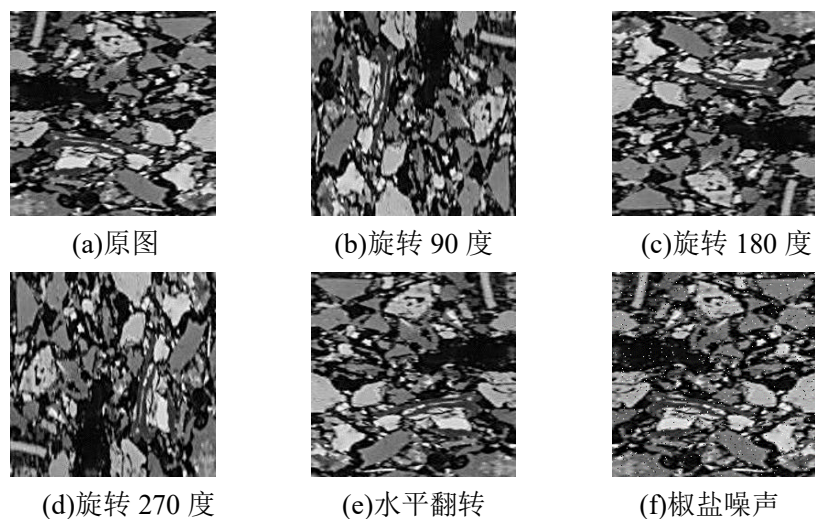


图 2.2 数据增强样本

2.3 本章小结

本章主要对煤岩镜质组显微图像以及数据集的构建进行详细的阐述。首先详细介绍了本文的研究对象镜质组各显微组分的特点；然后对各显微组分的特征进行分析以阐明采用深度学习的必要性；最后对数据集进行数据增强操作，并阐明了数据增强对于网络训练的重要性。通过对镜质组显微组分的数据扩充，为后续研究及创新提供了数据基础。

第三章 基于 DenseNet121 模型轻量化的煤岩镜质组分类

目前深度卷积神经网络模型在图像分类领域表现较好，但模型一般规模较大、参数较多，难以应用于某些真实的应用场景如移动设备或嵌入式设备上，故研究规模小的卷积神经网络模型至关重要。因此，本章在经典的深度卷积神经网络模型 DenseNet121 的基础上进行改进，在保持准确率的情况下降低模型的参数量和计算量，实现深度卷积神经网络模型的轻量化，构建出适用于煤岩镜质组分类的轻量化模型。

3.1 卷积神经网络结构

卷积神经网络(CNN)^[65]的结构主要由三部分组成，分别是卷积层、池化层和全连接层，其结构如图 3.1 所示。通常采用卷积层和池化层交替堆叠的设计思路，其具体工作过程为：首先卷积层和池化层用来提取图像不同语义信息的特征图，并将信息逐层传递，通过连续的卷积、池化操作后提取样本的高语义特征信息，然后将提取到的这些特征图转换成一维特征向量，最后将这些特征向量输入到全连接层，并经过 Softmax 分类器完成分类。

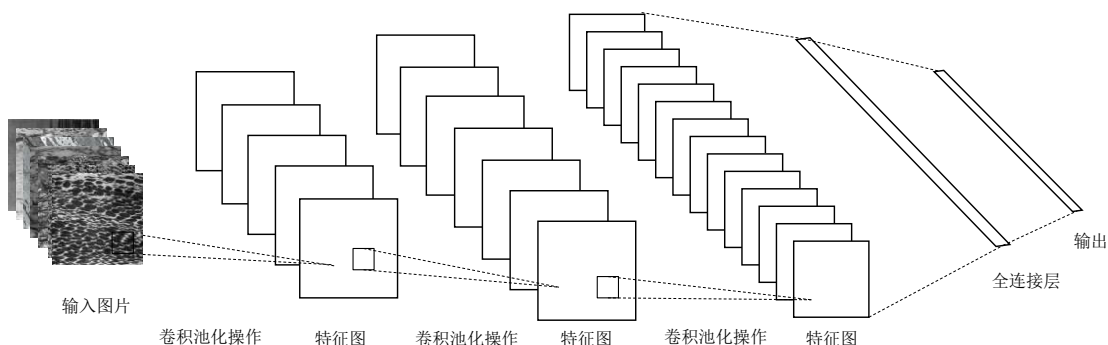


图 3.1 卷积神经网络的结构

(1) 卷积层

卷积层是卷积神经网络的重要组成部分，主要用于学习输入图像的特征表示^[66]。卷积层实现特征提取的具体操作步骤为：卷积核采用滑动窗口的模式对输入信息进行遍历，然后将图像对应位置元素与卷积核中的权值相乘并累加求和，最终得到变换后的特征图。卷积操作如图 3.2 所示。

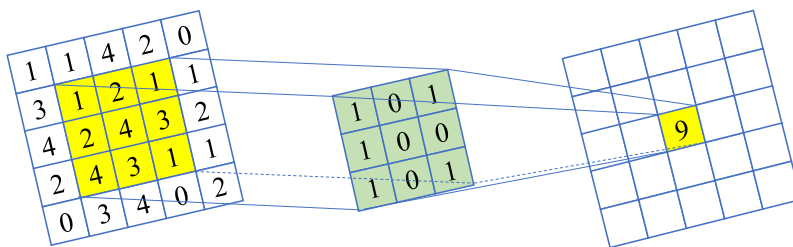


图 3.2 卷积操作示例

在卷积运算过程中，输入图像、卷积核和输出特征图之间相互依赖。当输入图像的大小为 $W_I \times H_I \times C_I$ ，卷积核大小为 $K_W \times K_H$ 、数量为 K_C ，对输入图像边缘填充圈数为 P ，卷积核移动的步长为 S ，卷积操作后得到的特征图大小为 $W_O \times H_O \times C_O$ 时，输出特征图的计算表达式如下：

$$\begin{aligned} W_O &= \frac{W_I - K_W + 2P}{S} + 1 \\ H_O &= \frac{H_I - K_H + 2P}{S} + 1 \\ C_O &= K_C \end{aligned} \quad (3.1)$$

(2) 池化层

池化层也称下采样层，通常位于卷积层之后。由于神经网络的不断加深导致网络参数规模增加，池化层的主要作用是特征筛选和信息过滤，使主要特征得到保留的同时，减少网络模型的参数规模并提高计算速度。同时池化操作能够对输入数据在一定程度上保持平移、缩放和旋转的不变性。

由于采样方式的不同，可分为最大池化和平均池化。最大池化是选取池化区域中最大的元素作为输出，平均池化将池化区域内的平均值作为输出。两种池化在实际应用中的侧重点并不相同，最大池化侧重于对纹理信息的选择，平均池化则更侧重于对背景信息的选择。图 3.3 为池化窗口大小为 2×2 ，步长为 2 池化操作示意图。

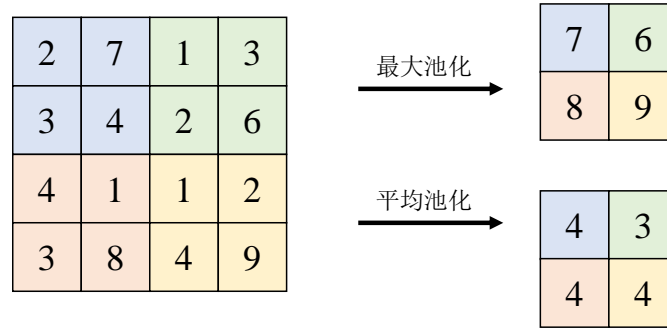


图 3.3 池化操作示意图

(3) 全连接层

通常在 CNN 的末端是全连接层，全连接层将卷积层和池化层等之前层学到的特征信息映射到样本标记空间中^[67]，并将特征信息转换成一维向量通过分类器函数得到最终输出。全连接层中神经元与上层的所有神经元均相连，所以参数量占比较大。全连接层的结构如图 3.4 所示。

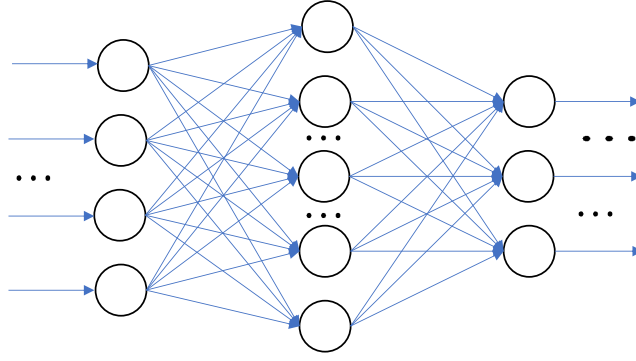


图 3.4 全连接层的结构

由于全连接层具有分类器的作用，因此需要在全连接层后添加一个分类层完成前层特征提取的分类。常用 Softmax 函数^[68]作为分类层，其将全连接层输出的特征向量压缩到(0,1)区间内，输出值即为输入样本属于该类别的概率，值越大说明属于该类别的概率越大。Softmax 函数的计算表达式如下：

$$y_k = \frac{e^{z_k}}{\sum_{i=1}^n e^{z_i}} \quad (3.2)$$

其中， z_k 表示全连接层映射到 softmax 函数第 k 个输入信号，分母表示所有输入信号 z_i 的指数和， y_k 表示第 k 个输入信号的输出概率值，且所有输出概率值相加之和为 1。

(4) 激活函数

由于卷积池化层和全连接层的运算过程是线性过程，但实际上多数应用场景都是非线性的任务，因此使用非线性函数作为神经网络的激活函数。激活函数的主要作用是引入非线性因子，有效解决一些非线性任务，增强网络对特征的表达能力，逼近更复杂的表达和函数。神经网络中常见的激活函数图像如图 3.5 所示。

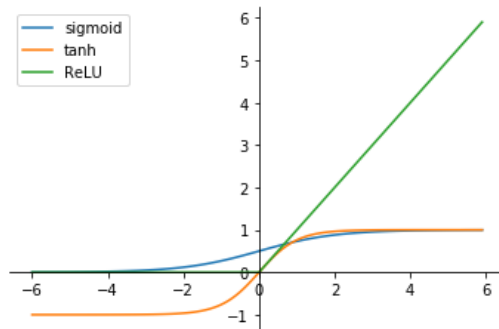


图 3.5 常见激活函数图像

1) Sigmoid 函数表达式如下：

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3.3)$$

Sigmoid 函数可以将输入任意数值压缩到 0 到 1 之间。由图 3.5 可以看出，

Sigmoid 函数严格单调递增，求导简单。但其也存在一些缺点：在反向传播时，由于软饱和性，使函数的导数逼近于 0，网络出现梯度消失现象，不利于深度模型的训练；Sigmoid 函数的输出均值非零，随着网络的训练会发生内部偏移现象，影响模型的训练。

2) Tanh 函数表达式如下：

$$g(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (3.4)$$

该函数可将输入值映射到 $-1 \sim 1$ 之间。Tanh 函数虽然解决了 Sigmoid 函数输出均值非零的问题，但其与 Sigmoid 函数一样存在着软饱和性，易产生梯度消失，导致模型训练困难。

3) ReLU 函数表达式如下：

$$z(x) = \max(0, x) \quad (3.5)$$

为缓解梯度消失现象，ReLU 激活函数被提出。如图 3.5 所示，当输入大于 0 时，函数是线性的，输出为其本身，很好地缓解了梯度消失现象的产生；当输入小于 0 时，输出恒为 0，神经元失活，使神经网络变得稀疏，降低了过拟合现象的产生。由 ReLU 函数的特点可知，该函数收敛速度较快，计算简单复杂度低，同时解决了梯度消失的问题。

(5) 批量归一化

批量归一化(Batch Normalization, BN)^[69]也是神经网络重要的组成部分。在网络训练过程中，会出现内部协变量偏移现象，该现象是由参数分布的变化导致网络隐藏层输入输出分布的变化而产生的，且网络越深这种现象越明显。内部协变量偏移现象会给卷积神经网络的训练带来一些问题，如网络的学习过程变得不稳定、收敛速度降低等。批量归一化可以有效的解决上述问题，可以保证即使网络中的各个层参数发生了变化，其输入输出数据分布也不会产生较大变化，从而避免内部协变量偏移现象的产生。批量归一化的过程如表 3.1 所示。

表 3.1 批量归一化过程

| |
|---|
| 算法： 批量归一化算法 (BN) |
| 输入： 小批量样本数据(mini-batch)输入 $\mathcal{B} = \{x_{1..m}\}$ ， x_i 为激活函数的输入。可学习参数 γ 和 β 。 |
| 输出： 归一化处理后的响应值 $\{y_i = \text{BN}_{\gamma, \beta}(x_i)\}$ |
| 1: $\mu_{\mathcal{B}} \leftarrow \frac{1}{m} \sum_{i=1}^m x_i$ //计算 mini-batch 的均值 |

- 2: $\sigma_B^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2$ //计算 mini-batch 的方差
- 3: $\hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}$ //mini-batch 的归一化处理
- 4: $y_i \leftarrow \gamma \hat{x}_i + \beta \equiv \text{BN}_{\gamma, \beta}(x_i)$ //尺度变换和偏移
- 5: **return** 学习的参数 γ 和 β

3.2 DenseNet 模型结构

经典的卷积神经网络模型有 AlexNet、VGGNet、ResNet 和 DenseNet 等，各个模型特点不同。AlexNet 网络结构简单，层数较少但每层卷积核数量较多，参数量较多。VGGNet 使用小卷积核代替大卷积核，增加了网络深度，但该模型的参数量和计算量都非常大，模型训练较为困难，而且当数据集规模不大时，训练易产生过拟合现象。ResNet 解决了网络层次过深导致的梯度消失问题，实现了更深层网络的构建，其在大规模数据集上可以达到很好的识别效果，但在小数据集上易产生过拟合^[70]。DenseNet 的基本思想是缩短输入与输出之间的路径，使层与层之间最大程度的信息流通，同时实现低层特征和高层特征的融合，加强了特征重用，具有很强的抗过拟合能力，因此该模型在数据集规模不大的情况下能提取更鲁棒的特征，使分类任务具有较好的效果，适用于本文煤岩镜质组显微组分的分类。本节详细阐述 DenseNet 模型的结构及其特点。

密集卷积网络(DenseNet)是由 huang 等^[50]人提出的一种卷积神经网络。该网络借鉴了 ResNet 网络的优点，ResNet 网络的核心是建立前层与后层间的短路连接，从而解决深层网络梯度消失的问题，训练出更深的卷积神经网络。与 ResNet 的基本思路相同，DenseNet 也是利用了前层与后层短路连接的优势，与 ResNet 不同的是，DenseNet 是建立前面所有层与后面层的密集连接，即前面每个层的输出都输入到后面每个层中，后面每个层的输入都是前面每个层的输出。这种密集连接更有效地利用了每层提取的特征图，实现了层与层之间最大程度的信息流通，缓解了梯度消失问题。

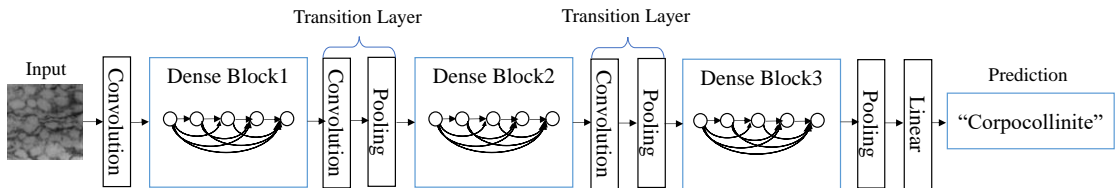


图 3.6 DenseNet 网络基本结构

DenseNet 网络的基本结构如图 3.6 所示，由图可以看出 DenseNet 主要由两

部分构成,即密集连接块(Dense Block)和过渡层(Transition Layer)。其中 Dense Block 模块的基本结构如图 3.7 所示,Dense Block 模块由多个 Dense Layer 构成,Dense Layer 的结构如图 3.10(a)所示,每个 Dense Layer 由批量归一化(BN)、ReLU 激活函数、 1×1 的卷积和 3×3 的卷积构成。Dense Block 模块的核心是每个层从前面所有层获得输入,并将其自身产生的特征图传递给后面所有层。假设 Dense Block 模块有 L 个卷积层,则第 l 层的有 l 个输入,包括前面所有层提取的特征图,而它自己提取的特征图也被作为后面 $L-l$ 层的输入。因此在 L 层网络中引入了 $L(L+1)/2$ 个连接,而不是像传统的卷积神经网络结构只引入 L 个连接。

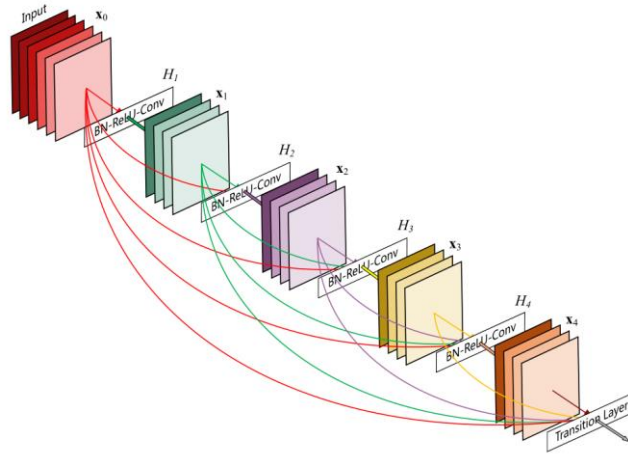


图 3.7 Dense Block 模块结构

传统的卷积神经网络前层与后层之间只有一个直接连接,因此传统的卷积神经网络在第 l 层的输出为:

$$x_l = H_l(x_{l-1}) \quad (3.6)$$

与传统的卷积神经网络不同的是,DenseNet 采用密集连接的方式,前面每个层的输出都是后面每个层的输入,后面每个层的输入都是前面每个层的输出,因此 DenseNet 在第 l 层的输出为:

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}]) \quad (3.7)$$

其中, x_l 表示第 l 层的输出; x_{l-1} 表示第 $l-1$ 层的输出; $H_l(\cdot)$ 表示非线性复合函数,该函数包括批量归一化(BN)、ReLU 激活函数和卷积操作; $[x_0, x_1, \dots, x_{l-1}]$ 表示将第 l 层之前的第 $0, 1, \dots, l-1$ 层中产生的特征图按通道的维度进行连接。

Dense Block 模块的设计使得 DenseNet 比传统的卷积神经网络拥有更少的参数,因为每个卷积层输出特征图的数量都非常小。除了减少参数量,Dense Block 模块还可以使梯度消失现象得以改善,让模型训练更加容易。梯度消失现象的产生是由于网络层加深时,梯度信息和输入信息在多层间传递而导致的。而 Dense Block 模块中密集连接的方式使每一层都可以直接访问来自损失函数和原始输入

信号的梯度，从而减缓梯度消失现象。除此之外，这种密集连接方式还具有正则化的效果，可以有效减缓因训练集较小产生的过拟合。

此外，DenseNet 还通过引入增长率、添加瓶颈层和过渡层压缩的方式实现网络优化。

(1)Growth rate 是网络的增长率，即增长率 k ，每个密集连接块(Dense Block)中每层输出的特征图个数可以用增长率 k 表示。假设 k_0 为 Dense Block 的输入层的通道数，则第 l 层有 $k_0 + k(l-1)$ 个输入通道数。每层输出通道数呈逐渐增加趋势，故称为增长率。DenseNet 将特征图视为网络的全局状态，每层都会添加自身的 k 个特征图到该状态。当全局状态形成，就可以从网络中的任何地方访问，不需要像传统的网络那样一层层地复制。DenseNet 与传统网络结构的最重要区别是模型可以很窄，即 k 的取值可以较小，DenseNet 作者的实验也证明 k 的取值较小可以有更好的效果。因此，本章使用较小的增长率 k 设计 DenseNet 模型，并针对煤岩镜质组数据集设计了增长率 k 的最优化实验。

(2)Bottleneck layer 是网络的瓶颈层。根据密集连接的设计，后面的层可以得到前面所有层的输入，这就导致了拼接后的输入较大，网络计算量增加。因此为减少输入特征图的数量，实现降维和减少计算量同时融合各通道的特征，DenseNet 在 3×3 的卷积之前引入 1×1 的卷积作为瓶颈层。

(3)Transition layer 是网络的过渡层。如图 3.6 所示，置于两个 Dense Block 模块中间，它包括一个批量归一化层、一个 1×1 卷积层、一个 2×2 平均池化层和压缩率 θ 。主要实现降采样操作，当特征图大小发生变化时，式(3.7)中的操作将无法使层与层之间的特征图进行连接，而降采样可以改变特征图的大小，实现特征图大小的统一。同时为了提高模型紧凑性，DenseNet 在 Transition layer 中引入了压缩率 θ 。其中 $0 < \theta \leq 1$ ，DenseNet 默认 $\theta = 0.5$ ，表示 Transition layer 输出特征图数量减少为原来的一半，这样传递到下一个 Dense Block 的特征图数量也减少一半。

3.3 DenseNet121 模型的轻量化改进

3.3.1 深度可分离卷积

DenseNet 网络虽然相较于传统卷积神经网络而言参数量和计算量较少，但对于一些移动设备和嵌入式设备而言所占存储空间仍然较大。受轻量级网络 MobileNet^[71]的启发，本章采用深度可分离卷积代替标准卷积。与标准卷积相同，深度可分离卷积也是用来提取特征的卷积操作，但它可以显著地减少模型的参数量和计算量。由于镜质组显微组分数据集样本较少，而深度可分离卷积比传统卷积需要调整的参数量更少，因此使用较少的数据就能学到较好的特征表达。

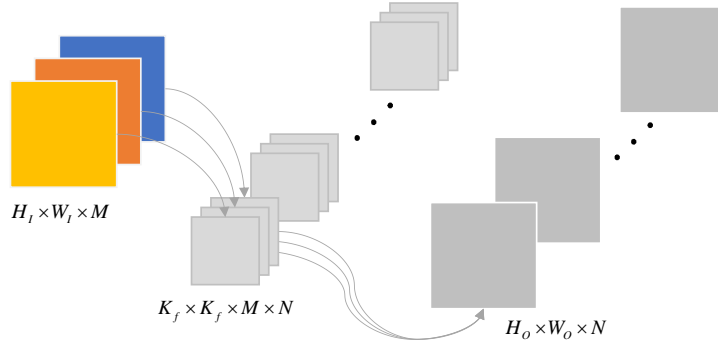


图 3.8 标准卷积

标准卷积如图 3.8 所示。假设输入特征图为 $H_I \times W_I \times M$ ，输出特征图为 $H_O \times W_O \times N$ 。其中， H_I 和 W_I 是输入特征图的高和宽， M 是输入通道数， H_O 和 W_O 是输出特征图的高和宽， N 是输出通道数。标准卷积层使用大小为 $K_f \times K_f \times M \times N$ 卷积核，其中 K_f 为卷积核的高和宽， M 和 N 分别为输入输出通道数，假设标准卷积步长为 1，Same Padding 填充。

标准卷积的参数量为：

$$K_f \times K_f \times M \times N \quad (3.8)$$

标准卷积的计算量为：

$$K_f \times K_f \times M \times N \times H_I \times W_I \quad (3.9)$$

标准卷积以卷积核为基础，过滤特征并合并特征，从而生成新的特征图；深度可分离卷积则将过滤和组合分为两个步骤执行以减少计算成本。因此深度可分离卷积由两个部分组成，一部分是深度卷积(depthwise convolution, DW)，另一部分是逐点卷积(pointwise convolution, PW)。其卷积过程如图 3.9 所示。

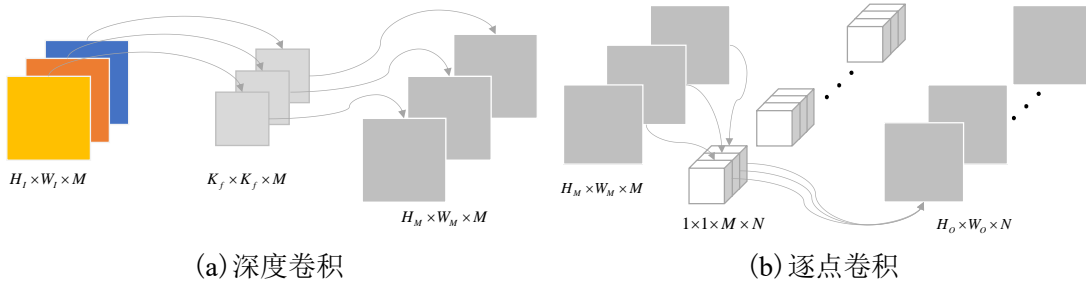


图 3.9 深度可分离卷积

深度卷积对每个输入通道采用单个滤波器过滤特征，其参数量为 $K_f \times K_f \times M$ ，计算量为 $K_f \times K_f \times M \times H_I \times W_I$ ；逐点卷积由 1×1 的卷积组成，用于实现深度层输出的线性组合，其参数量为 $1 \times 1 \times M \times N$ ，计算量为 $1 \times 1 \times M \times N \times H_I \times W_I$ 。

则深度可分离卷积的参数为：

$$K_f \times K_f \times M + 1 \times 1 \times M \times N \quad (3.10)$$

深度可分离卷积的计算量为：

$$K_f \times K_f \times M \times H_l \times W_l + 1 \times 1 \times M \times N \times H_l \times W_l \quad (3.11)$$

由式(3.8)和式(3.10)可得：

$$\frac{K_f \times K_f \times M + 1 \times 1 \times M \times N}{K_f \times K_f \times M \times N} = \frac{1}{N} + \frac{1}{K_f^2} \quad (3.12)$$

由式(3.9)和式(3.11)可得：

$$\frac{K_f \times K_f \times M \times H_l \times W_l + 1 \times 1 \times M \times N \times H_l \times W_l}{K_f \times K_f \times M \times N \times H_l \times W_l} = \frac{1}{N} + \frac{1}{K_f^2} \quad (3.13)$$

式中， N 是通道数通常大于 1， K_f 是卷积核的尺寸通常为 3×3 ， 5×5 ， 7×7 。

因此，式(3.12)和式(3.13)的值均小于 1，即深度可分离卷积的参数量和计算量比标准卷积少。

由以上对深度可分离卷积的卷积过程和原理分析可知，深度可分离卷积可以显著地减少模型的参数量和计算量。因此本章将 Dense Block 模块中 3×3 的标准卷积替换为深度可分离卷积，以减少计算成本构建更加轻量化的 DenseNet 模型。

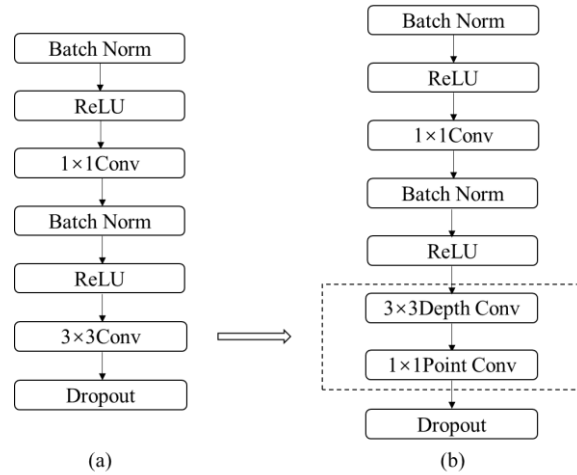


图 3.10 Dense Layer 结构

改进前后的 Dense Layer 结构如图 3.10 所示，图 3.10(a)表示改进前的 Dense Layer 结构，图 3.10(b)表示使用深度可分离卷积替换标准卷积后的 Dense Layer 结构。使用深度可分离卷积改进后的 Dense Layer 结构参数量和计算量大大减少，同时减缓小数据集产生的过拟合现象。

3.3.2 注意力机制

注意力机制模拟了人类大脑对视觉信号的处理过程^[72]。在观察一幅图像时，人类通常首先扫描全局图像，得到感兴趣的区域，然后投入更多的注意力到该区域，以关注该区域的细节信息，最后过滤掉其他无关的信息^[73]。注意力机制目前已被广泛应用于深度学习的各个领域，如图像分类^[74-75]、语言识别^[76-77]、自然语

言处理等^[78-79]。

(1) 注意力机制概述

注意力机制包括两种形式：软注意力机制和硬注意力机制^[80]。其中软注意力机制更关注特征图的每一通道或空间位置的权重信息，是确定性的注意力机制，而且软注意力是可微分的，因此可以使用神经网络计算更新梯度，并通过前向传播和反向传播以更新注意力权重。硬注意力机制只关注特征图的某一位置，是一种不确定性的动态过程，作用的对象是离散的位置信息，因此硬注意力作用后的区域是不可微分的，从而无法实现网络端到端的训练。由于软注意力机制可以端对端的进行训练，因此深度学习中广泛使用软注意力机制。本文也将引入软注意力机制实现煤岩显微组分镜质组的分类研究。

假设注意力网络的输入特征为 $X_{1:N} = [x_1, x_2, \dots, x_N]$, $i = 1, \dots, N$ ，注意力网络学习到每个区域的注意力权重值，权重值可以代表当前任务对该区域的关注程度。根据注意力机制变量 $Z \in [1, N]$ 选择输入特征信息的相应位置信息，即 $Z = i$ 表示选择了序号为 i 的特征信息为与当前任务相关查询信息 q 最匹配的特征信息，其中与当前任务相关查询信息 q 是指权重值的集合。使用软注意力机制时，注意力权重以概率形式表达，即表示该特征为该任务中关键特征的概率值，采用 α_i 表示选定该特征的概率值，则 α_i 的计算公式为：

$$\begin{aligned}\alpha_i &= p(Z = i | X_{1:N}, q) \\ &= \text{softmax}_i(s(x_i, q)) \\ &= \frac{\exp(s(x_i, q))}{\sum_{j=1}^N \exp(s(x_j, q))}\end{aligned}\quad (3.14)$$

其中， $s(x_i, q)$ 为注意力的打分函数，需要根据不同注意力任务、样本特点以及网络模型等因素进行选择。通用的打分函数可以表示为以下形式：

$$s(x_i, q) = v^T \tanh(Wx_i + U_q) \quad (3.15)$$

式中 v, W, U 是网络可训练的参数。

在得到注意力权重 α_i 后，为了获得注意力加权后的特征，需要将注意力权重 α_i 与输入特征进行加权运算，从而实现对特征的评分进而选择更重要的特征。对于软注意力机制，输入特征进行注意力加权的過程可表示为如下公式：

$$\text{attention}(X_{1:N}, q) = \sum_{i=1}^N \alpha_i x_i = E_{z \sim p_q(z|x_i)}[X] \quad (3.16)$$

(2) ECA 注意力机制

注意力机制在改善深度卷积神经网络的性能方面具有巨大的潜力，其可以从众多信息中选择出对当前任务目标更关键的信息，而抑制其他无关信息，极大地

提高了分类的效率与准确性。由于镜质组中部分显微亚组分类别间特征差异较小易混淆、局部特征甚至有交错，通过注意力机制可以关注到目标中更为关键的信息，快速提取重要的特征以解决图像难以区分的问题，因此本章将注意力机制引入到煤岩镜质组显微组分分类任务。

经典的注意力机制包括 SENet^[81]、CBAM^[82]、SCSE^[83]等，但这些注意力机制的复杂度仍然较高，会导致模型参数量和计算量的显著增加，不利于实现模型的轻量化。因此，Wang^[84]等提出了轻量级的架构单元 ECA，并证明了该模块可以用较低的计算成本来提高性能。为实现 DenseNet 模型的轻量化，本章将高效轻量化注意力机制 ECA 引入 DenseNet 网络，其可以有效地利用不同通道的相关性，且仅添加极少的计算量和参数量。DenseNet 仅在空间域上实现了特征融合，而 ECA 可以学习到不同通道间关系及不同通道的重要程度。因此本章将 ECA 引入 DenseNet 网络，使网络既可以在空间上特征融合也可以学习到特征通道间的关系，进一步提升网络性能。ECA 结构如图 3.11 所示。

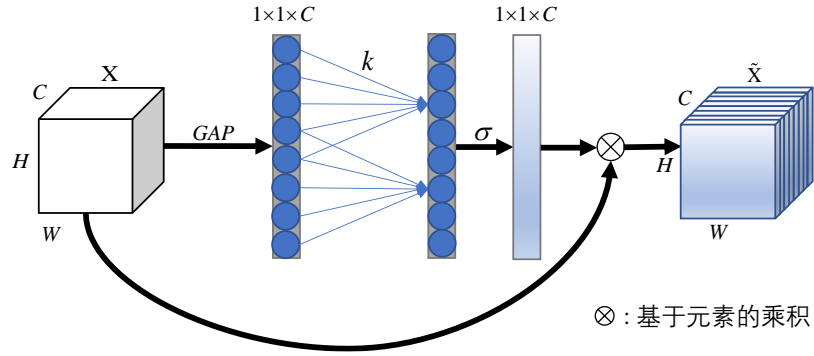


图 3.11 ECA 结构

注意力机制 ECA 是对传统的通道注意力机制 SE 模块的改进。SE 模块主要是首先通过全局平均池化对特征图进行通道域维度压缩以实现特征聚合，然后引入两个全连接层先降维后升维，最后非线性 Sigmoid 函数来生成通道权重，其中两个全连接层的目的是捕获全局通道信息交互，降维的目的是降低模型的复杂度，而降维的过程会损失一部分的特征表达，并且对所有通道间的信息进行交互的方式并不高效。因此 ECA 模块在不降低维度的情况下，首先对输入特征 X 的每个通道使用全局平均池化(GAP)来聚合特征图；然后使用大小为 k 的快速一维卷积完成局部的跨通道信息交互，再通过 Sigmoid 函数生成通道权重，其中使用一维卷积实现局部跨通道信息交互的同时降低了模型复杂度；最后将每个特征通道生成的权重作用于原始输入特征图，实现了对无效特征的抑制和对有效特征的增强。

由于本章的 ECA 通道注意力机制的目的是获得适当跨通道信息交互，因此需要确定交互的覆盖范围，也就是确定一维卷积的内核 k 的大小。对于不同通道数以及不同 CNN 架构的卷积块来说， k 值可能会有所不同。虽然可以手动调整

交互的覆盖范围,但是手动调整会造成计算资源的大量消耗。因此,ECA 提出了自适应选择一维卷积核大小的方法,以确定局部跨通道交互的覆盖范围。交互作用的覆盖范围与通道维数 C 相关,即一维卷积的内核大小 k 和 C 之间可能存在某种映射 ϕ ,如式(3.17)所示。

$$C = \phi(k) \quad (3.17)$$

最简单的映射是线性映射,即 $\phi(k) = \gamma * k - b$ 。但线性映射过于局限,又因通道维数 C ,即滤波器的数量一般是 2 的指数。因此,将线性函数 $\phi(k) = \gamma * k - b$ 扩展成非线性函数,如式(3.18)。

$$C = \phi(k) = 2^{(\gamma * k - b)} \quad (3.18)$$

在给定通道维度 C 的情况下,一维卷积的内核大小 k 可以自适应的确定,如式(3.19)所示。

$$k = \psi(C) = \left\lceil \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rceil_{\text{odd}} \quad (3.19)$$

式(3.19)中: $\lceil x \rceil_{\text{odd}}$ 表示最近 x 的奇数, γ 和 b 为 2 和 1。

3.3.3 改进后的模型结构

由于 DenseNet 的密集连接块中每一层输出都与后面所有层相连,这种方式会产生特征冗余的问题^[85]。为解决此问题,同时提高模型分类准确率,本章将 ECA 模块引入 DenseNet 网络,通过注意力机制 ECA 模块对提取到的特征重新分配权重,增强有用信息,抑制无用信息,从而减缓 DenseNet 产生的特征冗余的影响,并提升模型准确率,使模型得到优化。分析 DenseNet 模型可知,DenseNet 仅实现了空间上的特征融合,而 ECA 模块可以学习到不同通道间的关联度和重要程度,因此,本章将 ECA 模块引入 DenseNet 网络使得网络既可以融合通道特征信息也可以融合空间特征信息,使 DenseNet 网络模型得到优化。同时,ECA 模块既高效又轻量,引入此注意力机制仅增加极少的参数量。为了进一步减少模型参数量,本章采用深度可分离卷积替换普通卷积,同时优化特征图通道增长率 k ,以实现 DenseNet 模型的轻量化。

DenseNet121 是 DenseNet 模型中最经典且应用最为广泛的模型,该模型层数为 121 层,原理和结构与 DenseNet 模型相同。本章使用 DenseNet121 作为主干网络,将深度可分离卷积和 ECA 注意力机制引入到 DenseNet121 模型,提出一种基于 DenseNet121 轻量化的镜质组分类模型(DE-DenseNet121),该模型的网络结构如图 3.12 所示。首先输入待分类的镜质组显微组分图像,经过一个卷积层提取浅层特征;然后经过 Dense Block,并在 Dense Block 中每个 Dense layer 之后添加 ECA 模块,之后经过 Transition layer,并在其后添加 ECA 模块,为特征

重新分配权重；最后进入分类层以实现镜质组显微组分的分类。

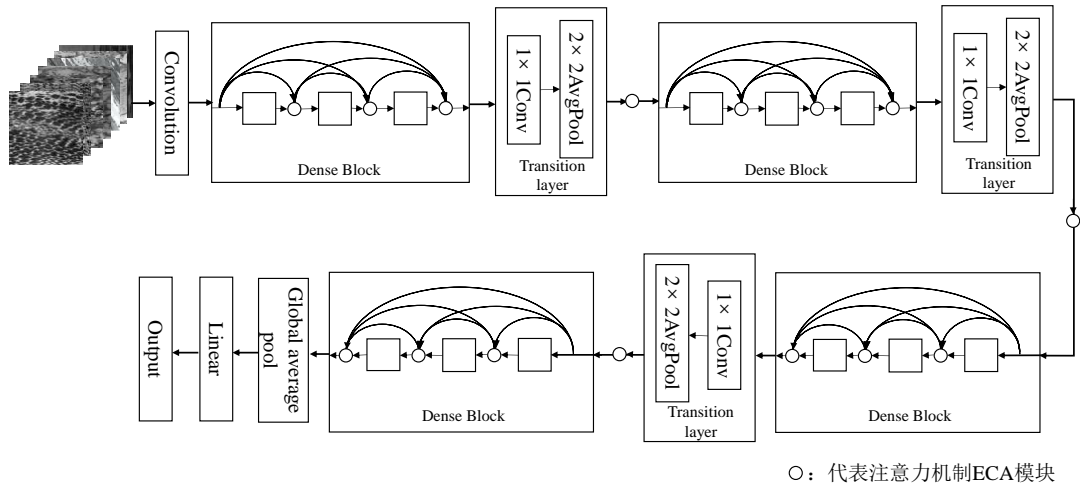


图 3.12 DE-DenseNet121 网络结构图

本章在 DenseNet121 模型的基础上引入 ECA 模块，并将标准卷积替换为深度可分离卷积，同时优化特征图通道增长率 k ，构建 DE-DenseNet121 卷积神经网络。表 3.2 为 DE-DenseNet121 和原 DenseNet121 的网络结构配置，网络模型采用经典的 DenseNet121 内部密集连接块数量和内部网络层数，即两个网络均含有 4 个 Dense Block，每个 Dense Block 内部的 Dense layer 使用个数分别为 6，12，24，16。且均包含一个 7×7 的初始卷积层、一个最大池化、4 个 Dense Block 和 3 个 Transition Layer 交替堆叠，最后一个分类层包括全局平均池化、全连接以及 Softmax 函数实现分类。其中 DE-DenseNet121 的每个 Dense Block 中 3×3 的标准卷积均替换为深度可分离卷积并在其后添加 ECA 模块，每个 Transition Layer 中平均池化后均添加 ECA 模块，以提升模型性能。由章节 3.2 可知 k 是网络特征图通道的增长率， k 的取值较小也可以有更好的效果，本章使用较小的增长率 k 设计改进的 DenseNet121 模型，即 $k=12$ ，并在章节 3.4.3.2 中针对煤岩镜质组数据集设计了增长率 k 的最优化实验。

表 3.2 DenseNet121 和 DE-DenseNet121 网络结构配置信息表

| Layers | DenseNet121 | DE-DenseNet121 |
|----------------------|---|---|
| Convolution | 7×7 conv, stride 2 | 7×7 conv, stride 2 |
| Pooling | 3×3 max pooling, stride 2 | 3×3 max pooling, stride 2 |
| Dense Block (1) | $\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$ | $\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ Depth wise Conv} \\ 1 \times 1 \text{ Point wise Conv} \\ \text{ECA module} \end{bmatrix} \times 6$ |
| Transition Layer (1) | 1×1 conv 2×2 average pooling, stride 2 | 1×1 conv 2×2 average pooling, stride 2 ECA module |

| | | |
|----------------------|--|--|
| Dense Block (2) | $\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$ | $\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ Depth wise Conv} \\ 1 \times 1 \text{ Point wise Conv} \\ \text{ECA module} \end{bmatrix} \times 12$ |
| Transition Layer (2) | $1 \times 1 \text{ conv}$ $2 \times 2 \text{ average pooling,}$ $\text{stride } 2$ | $1 \times 1 \text{ conv}$ $2 \times 2 \text{ average pooling, stride } 2$ ECA module |
| Dense Block (3) | $\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$ | $\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ Depth wise Conv} \\ 1 \times 1 \text{ Point wise Conv} \\ \text{ECA module} \end{bmatrix} \times 24$ |
| Transition Layer (3) | $1 \times 1 \text{ conv}$ $2 \times 2 \text{ average pooling,}$ $\text{stride } 2$ | $1 \times 1 \text{ conv}$ $2 \times 2 \text{ average pooling, stride } 2$ ECA module |
| Dense Block (4) | $\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$ | $\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ Depth wise Conv} \\ 1 \times 1 \text{ Point wise Conv} \\ \text{ECA module} \end{bmatrix} \times 16$ |
| Classification Layer | $7 \times 7 \text{ global average pool}$ | |
| | fully connected, softmax | |

3.4 实验与结果讨论

3.4.1 实验环境与参数设置

本文实验基于 64 位的 Windows10 操作系统，通过搭建 Pytorch 深度学习框架，采用 Python 编程语言实现，其详细配置如表 3.3 所示。

表 3.3 本文实验环境配置

| 名称 | 参数/版本号 |
|---------|-----------------------------|
| GPU | Nvidia GeForce RTX 3080 |
| CPU | Intel(R) Core(TM) i9-10850K |
| 内存 | 32GB |
| CUDA | 10.1 |
| Python | 3.6 |
| Pytorch | 1.4.0 |

本文实验还涉及到各种超参数的设置，结合实验环境和实验对象，训练时各个参数的设置如表 3.4 所示。在模型的训练过程中采用 Adam 优化算法作为网络优化器，它对于不同的参数具有不同的自适应学习率，并可以基于煤岩镜质组的训练图像不停地更新权重。设置批次大小为 24，分类函数为 Softmax 函数，分类

损失函数为交叉损失函数。在训练过程使用学习率衰减策略，初始学习率设为 0.0001，第 50 个 epoch 对学习率乘以 0.5，一共训练 100 个 epoch。

表 3.4 训练参数设置

| 参数名称 | 参数值 |
|------------|--------------------------|
| Batch_size | 24 |
| 优化器 | Adam |
| 分类函数 | Softmax |
| 损失函数 | Categorical_crossentropy |

3.4.2 评价指标

为了便于实验结果比较和分类模型性能的评估，实验选用了多个在深度学习领域和分类领域中常用的评价指标，包括混淆矩阵(Confusion Matrix)、准确率(Accuracy)、精确率(Precision)、召回率(Recall)、F1 值(F1-Score)以及复杂度评价指标。模型的复杂度评价指标包括模型参数量(Parameters)和浮点运算数(FLOPs)。其中 Parameters 是指模型中包含的参数数量，直接决定模型文件的大小；FLOPs 指浮点运算数，理解为计算量，可以衡量模型的复杂度。

表 3.5 二分类混淆矩阵

| 实际情况 | 预测情况 | |
|------|---------|---------|
| | 正例 | 负例 |
| 正例 | TP(真正例) | FN(假反例) |
| 负例 | FP(假正例) | TN(真反例) |

混淆矩阵是用来表现分类模型性能可视化结果的一种矩阵。对于二分类任务，混淆矩阵中的四种判断，即 TP、FP、TN、FN，这四种情况分别表示预测为正、实际为正；预测为正、实际为负；预测为负、实际为负；预测为负、实际为正。

根据混淆矩阵可以计算出各个类别的准确率、精确率、召回率以及 F1 值，计算公式如下：

(1) 准确率(Accuracy)

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \quad (3.20)$$

准确率表示分类正确的样本占总样本数量的比例，它是分类任务中最简单最直观的评价指标。

(2) 精准率(Precision)

$$Precision = \frac{TP}{TP + FP} \quad (3.21)$$

精确率表示的是正确预测为正样本的数量占全部预测为正样本数量的比例。

(3) 召回率(Recall)

$$Recall = \frac{TP}{TP + FN} \quad (3.22)$$

召回率又称查全率，表示分类正确为正样本的数量占真实为正样本的比例。

(4) F1 值(F1-Score)

$$F1 = 2 \times \frac{precision \times recall}{precision + recall} \quad (3.23)$$

F1 值是兼顾精确率和召回率的综合指标，通常 F1 值越大，模型性能越好。

由于镜质组显微组分有 7 个类别，是一种多分类任务。以上评价指标用于评价多分类任务时，通常采用宏平均的度量方法，包括平均准确率、宏精确率、宏召回率以及宏 F1 值。各评价指标计算公式如下：

$$Average - accuracy = \frac{\sum_{i=1}^n accuracy_i}{n}, n = 7 \quad (3.24)$$

$$Macro - precision = \frac{1}{n} \sum_{i=1}^n \frac{TP_i}{TP_i + FP_i}, n = 7 \quad (3.25)$$

$$Macro - recall = \frac{1}{n} \sum_{i=1}^n \frac{TP_i}{TP_i + FN_i}, n = 7 \quad (3.26)$$

$$Macro - F1 = \frac{2 \times macro - precision \times macro - recall}{macro - precision + macro - recall} \quad (3.27)$$

其中， i 表示镜质组显微组分的某一类别， n 表示总类别数， FP_i 表示实际为其他类但被预测为第 i 类的样本数量， TN_i 表示实际为其他类被预测其他类的样本的数量， TP_i 表示实际为第 i 类被预测为第 i 类的样本数量， FN_i 表示实际为第 i 类但被预测为其他类的样本的数量。

3.4.3 实验结果与分析

3.4.3.1 改进的 DenseNet121 模型实验结果与分析

由于本章旨在实现 DenseNet121 模型的轻量化，所以除了分类性能评价指标平均准确率外，还需要考虑模型的复杂度指标。评价深度学习模型复杂度的指标为参数量(Parameters)和浮点运算次数(FLOPs)。其中 FLOPs 理解为模型的计算量。表 3.6 显示了改进前后分类模型的参数量、FLOPs 和平均准确率。

表 3.6 DenseNet121 与 DE-DenseNet121 模型性能对比

| 模型 | 参数量(M) | FLOPs(M) | 平均准确率(%) |
|-------------|--------|----------|----------|
| DenseNet121 | 7.979 | 2865.672 | 95.71 |

续表 3.6

| | | | |
|----------------|-------|---------|-------|
| DE-DenseNet121 | 1.210 | 443.145 | 96.43 |
|----------------|-------|---------|-------|

由表 3.6 可以看出，与 DenseNet121 模型相比，本文构建的 DE-DenseNet121 模型不仅平均准确率高，而且模型参数量和 FLOPs 大幅度减少。其模型参数量和 FLOPs 仅为 1.210M 和 443.144M，均约为 DenseNet121 原模型的 1/6。证明了深度可分离卷积和优化增长率操作降低了模型的大小，注意力机制提高了模型的准确率。表明了 DE-DenseNet121 模型在保持较高分类准确率的同时有效地降低了模型的复杂度，实现了模型的轻量化。

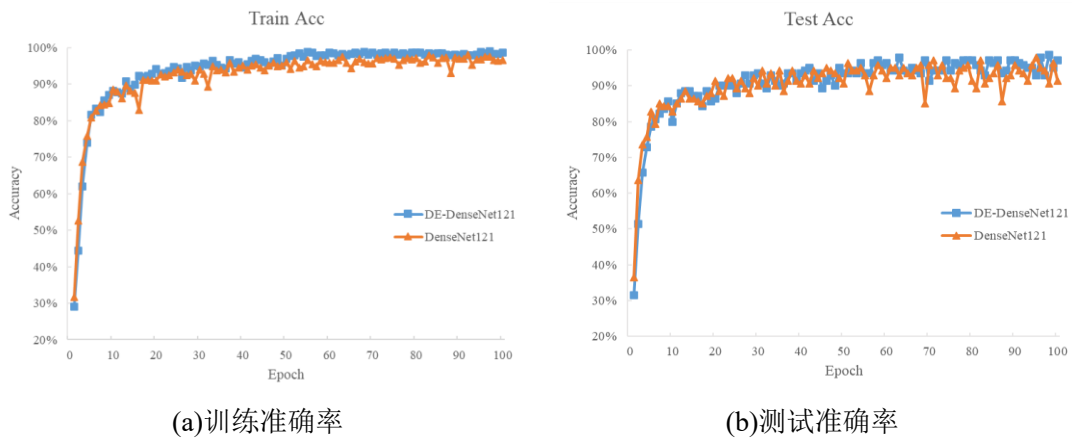
为了综合评估 DE-DenseNet121 模型的整体分类性能，除了使用平均准确率外，还使用宏精确率、宏召回率与宏 F1 值共四种评价指标进一步分析模型分类性能。模型的分类结果对比如表 3.7 所示。

表 3.7 DenseNet121 与 DE-DenseNet121 模型分类结果对比

| 模型 | 平均准确率(%) | 宏精确率(%) | 宏召回率(%) | 宏 F1 值(%) |
|----------------|----------|---------|---------|-----------|
| DenseNet121 | 95.71 | 96.46 | 95.71 | 95.83 |
| DE-DenseNet121 | 96.43 | 96.59 | 96.43 | 96.41 |

由表 3.7 可以看出，DE-DenseNet121 模型的四种评价指标均高于 DenseNet121 原模型。与 DenseNet121 模型相比，本章提出的模型其平均准确率、宏精确率、宏召回率和宏 F1 值分别高于 DenseNet121 模型 0.72%、0.13%、0.72% 和 0.58%，表明本章提出的 DE-DenseNet121 模型分类结果相对较好。

为验证 DE-DenseNet121 模型的稳定性及收敛性，本章实验使用煤岩镜质组数据集对构建的 DE-DenseNet121 模型和 DenseNet121 模型进行训练和测试，图 3.13 展示了两种模型训练和测试的准确率及损失值迭代曲线。



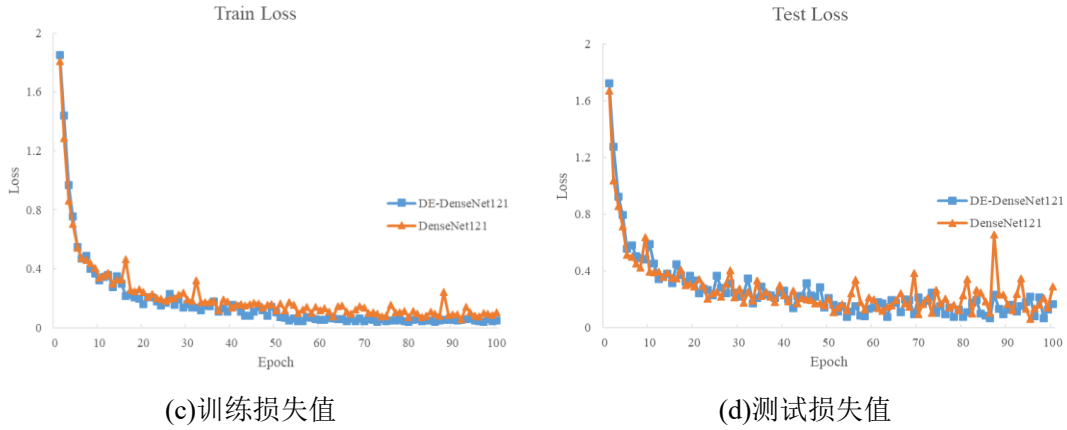


图 3.13 DenseNet121 和 DE-DenseNet121 模型收敛图

从图 3.13 可以看出，DE-DenseNet121 模型的训练和测试准确率在第 20 个 epoch 后达到 90%以上并趋于稳定，损失值也在第 20 个 epoch 后趋于稳定。相较于 DE-DenseNet121 模型，DenseNet121 模型稳定性相对较差，且准确率略低。表明本章提出的 DE-DenseNet121 模型的收敛速度较快且稳定性较高，对于煤岩镜质组显微组分的分类任务是有效的。

3.4.3.2 增长率优化实验结果与分析

由于增长率 k 的取值会对模型的性能造成一定的影响，本节设计了一组增长率 k 的最优化实验，增长率 k 表示每个 Dense Block 模块中每层输出的特征图个数，由于 DenseNet121 与传统网络结构的一个重要区别就是网络模型可以非常窄，即 k 的取值可以较小，因此综合考虑选择 k 的取值为 6、8、12、18、24、32 六种方案对模型进行训练和测试，具体结果如表 3.8 所示。

表 3.8 不同增长率配置实验对比

| 模型 | 参数量(M) | FLOPs(M) | 平均准确率(%) |
|--------------------------|--------|----------|----------|
| DE-DenseNet121($k=6$) | 0.449 | 239.278 | 90.71 |
| DE-DenseNet121($k=8$) | 0.666 | 296.733 | 93.57 |
| DE-DenseNet121($k=12$) | 1.210 | 443.145 | 96.43 |
| DE-DenseNet121($k=18$) | 2.302 | 741.587 | 94.29 |
| DE-DenseNet121($k=24$) | 3.731 | 1134.896 | 95.71 |
| DE-DenseNet121($k=32$) | 6.154 | 1806.493 | 96.43 |

由表 3.8 可以看出，随着 k 值的增长，模型的参数量和计算量不断增加，这是因为 k 值表示每个 Dense Block 模块中每层输出的特征图个数， k 值越大模型参数量和计算量越大。当增长率 k 取 12 和 32 时，模型的准确率最高均达到 96.43%，但是当增长率 k 的取值为 32 时，其模型参数量和计算量比 k 取 12 时分别高出 5

倍和 4 倍。当增长率 k 取 6 时,其模型参数量和计算量仅为 0.449M 和 239.278M,模型复杂度最低,但其准确率仅为 90.71%,明显低于 k 取 12 时模型的准确率,但是以牺牲准确率的代价来降低模型的参数量和计算量是不可取的。以上实验结果表明本章选定的增长率 k 取 12 时,模型性能最优,最为适用于煤岩镜质组显微组分的分类。

3.4.3.3 消融实验结果与分析

为了更好地分析本章提出的 DE-DenseNet121 模型的优越性,证明深度可分离卷积和 ECA 注意力机制模块的有效性,设计了一组消融实验,实验结果如表 3.9 所示。该组实验中固定各个模型增长率 k 的取值均为 12,其中增长率优化实验见表 3.8。

表 3.9 DE-DenseNet121 模型消融实验

| 深度可分离卷积 | ECA 注意力机制 | 平均准确率(%) | 参数量 |
|---------|-----------|----------|-----------|
| × | × | 94.29 | 1,448,376 |
| √ | × | 94.29 | 1,209,648 |
| × | √ | 96.43 | 1,448,559 |
| √ | √ | 96.43 | 1,209,831 |

从表 3.9 可以看出,仅将标准卷积替换为深度可分离卷积时,模型参数量减少,证明了深度可分离卷积能够有效减少参数量;仅使用 ECA 注意力机制时,模型仅增加极少量参数,但准确率提高了 2.14%,证明了引入 ECA 注意力机制对于镜质组显微组分的分类具有一定的效果;当同时使用深度可分离卷积和 ECA 注意力机制时,模型参数量降低且准确率提高,表明本章提出的模型具有一定的优越性。

3.4.3.4 不同模型对比实验结果与分析

为了验证本章所提出模型的在镜质组显微组分的识别中优于其他模型,选用轻量级网络模型、经典的卷积神经网络与 DE-DenseNet121 模型进行对比,共 8 组实验。其中轻量级网络模型包括 SqueezeNet、MobileNet V2 和 ShuffleNet V2,经典的卷积神经网络模型包括 AlexNet、VGG16、ResNet50 和 DenseNet121。不同模型分类结果对比如表 3.10 所示。

表 3.10 不同模型分类结果对比

| 模型 | 平均准确率(%) | 宏精确率(%) | 宏召回率(%) | 宏 F1 值(%) |
|--------------|----------|---------|---------|-----------|
| SqueezeNet | 93.57 | 93.56 | 93.57 | 93.55 |
| MobileNet V2 | 93.57 | 94.09 | 93.57 | 93.64 |

续表 3.10

| | | | | |
|----------------|-------|-------|-------|-------|
| ShuffleNet V2 | 91.43 | 92.20 | 91.43 | 91.28 |
| AlexNet | 94.29 | 94.97 | 94.29 | 94.32 |
| VGG16 | 95.71 | 96.06 | 95.71 | 95.68 |
| ResNet50 | 95.00 | 95.30 | 95.00 | 94.84 |
| DenseNet121 | 95.71 | 96.46 | 95.71 | 95.83 |
| DE-DenseNet121 | 96.43 | 96.59 | 96.43 | 96.41 |

表 3.10 显示了 8 种分类模型的平均准确率、宏精确率、宏召回率和宏 F1 值的实验结果。由表 3.10 可以看出,本章提出的 DE-DenseNet121 模型的平均准确率、宏精确率、宏召回率和宏 F1 值均高于其他几个轻量级网络及经典卷积神经网络,表明本章所提出的模型分类效果较好,可以有效地应用于煤岩镜质组显微组分的分类。

由于本章提出的模型旨在实现 DenseNet121 模型的轻量化及高准确率,因此为了更全面的描述深度学习模型,本章将使用模型的复杂度评价指标对各个模型进行评估。表 3.11 显示了每个模型的参数量、FLOPs 和平均准确率。

表 3.11 不同模型性能对比

| 模型 | 参数量(M) | FLOPs(M) | 平均准确率(%) |
|----------------|---------|-----------|----------|
| SqueezeNet | 1.235 | 351.911 | 93.57 |
| MobileNet V2 | 3.505 | 314.130 | 93.57 |
| ShuffleNet V2 | 1.367 | 42.524 | 91.43 |
| AlexNet | 61.101 | 714.692 | 94.29 |
| VGG16 | 138.358 | 15483.862 | 95.71 |
| ResNet50 | 25.557 | 4111.515 | 95.00 |
| DenseNet121 | 7.979 | 2865.672 | 95.71 |
| DE-DenseNet121 | 1.210 | 443.145 | 96.43 |

通过分析表 3.11 可知,本章提出的 DE-DenseNet121 模型的平均准确率虽然仅高于 DenseNet121 模型 0.72%,但是参数量和 FLOPs 仅为 DenseNet121 模型的 1/6,表明本章所提出的模型在提高准确率的同时大幅度缩减了模型的规模。具体分析,DE-DenseNet121 模型与轻量级模型 SqueezeNet、MobileNet V2 和 ShuffleNet V2 相比,虽然 FLOPs 高于三个轻量级模型,但是其平均准确率分别提高了 2.86%、2.86%和 5.00%,模型的参数分别减少了 0.025M、2.295M 和 0.157M,表明了 DE-DenseNet121 模型不仅与目前主流的轻量化模型大小相当,而且其准确率高于主流的轻量化模型;与经典的卷积神经网络 AlexNet、VGG16、ResNet50

和 DenseNet121 相比,其平均准确率分别提高了 2.14%、0.72%、1.43%和 0.72%,模型的参数分别约为经典卷积神经网络的 1/50、1/114、1/21 和 1/6,模型的 FLOPs 分别约为经典卷积神经网络的 3/5、1/35、1/9 和 1/6,准确率提升的同时参数量和计算量大幅度减少。

为了更直观地观察不同分类模型性能,采用不同模型的准确率、参数量和计算量对比图进行观察分析,如图 3.14、3.15 所示,通过图 3.14 可以直观看出 DE-DenseNet121 模型的准确率最高,参数量也最少;通过图 3.15 可以直观看出 DE-DenseNet121 模型的准确率最高,FLOPs 虽略高于轻量级模型但明显低于经典卷积神经网络模型。说明本章所提出的模型不仅与经典的卷积神经网络相比拥有更小的模型规模和较高准确率,而且与轻量级网络模型相比也拥有相当的模型规模和更高的准确率。以上分析表明本章所提出的模型性能更优,且实现了 DenseNet121 模型的轻量化及高准确率。

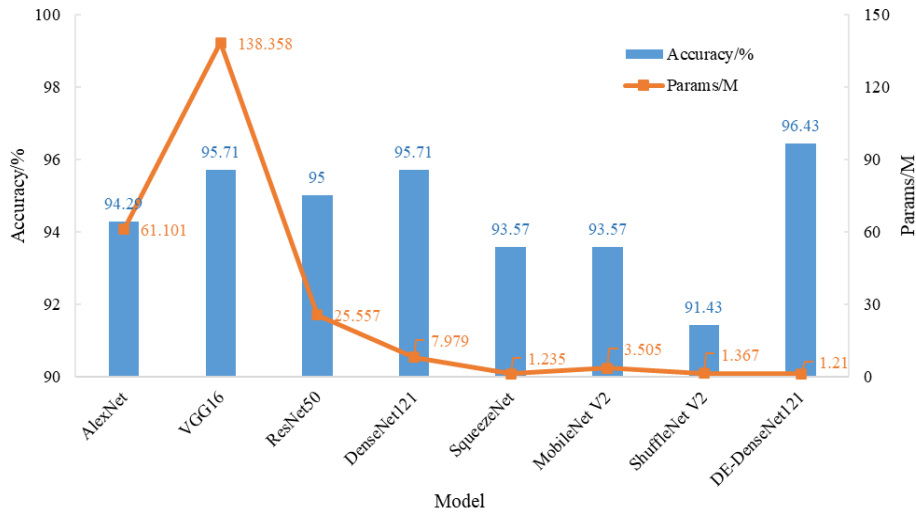


图 3.14 不同模型准确率和参数量对比图

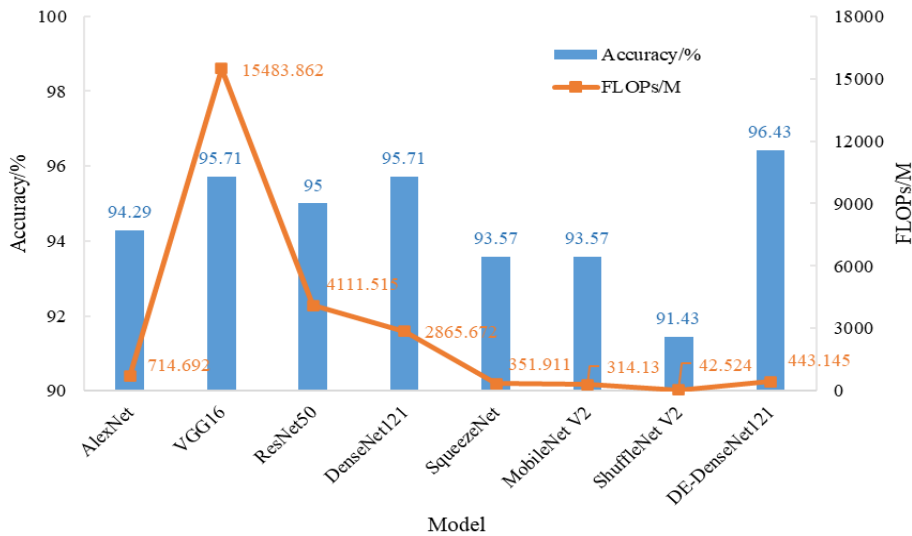


图 3.15 不同模型准确率和计算量对比图

3.5 本章小结

为实现经典卷积神经网络模型的轻量化,本章提出了一种基于模型轻量化的煤岩镜质组分类方法。首先,在 DenseNet121 模型中引入注意力机制 ECA 模块,通过注意力机制对提取到的特征重新分配权重,增强有用信息,抑制无用信息,减缓 DenseNet121 模型产生的特征冗余的影响;然后,将 DenseNet121 模型中的标准卷积替换为深度可分离卷积,同时优化模型增长率,以减少模型参数量和计算量;最后,对改进后模型的实验结果进行分析,结果表明与经典的卷积神经网络模型、轻量级模型及原模型相比,DE-DenseNet121 模型分类效果最佳且实现了模型的轻量化。本章提出的模型在保持分类准确率的情况下实现了模型的轻量化,为煤岩镜质组显微组分的识别提供了一种可行的方法。

第四章 基于注意力机制的轻量级网络煤岩镜质组分类

在第三章的研究中,在对经典卷积神经网络模型进行优化时,减少了深度卷积神经网络模型参数量和计算量,实现了深度卷积神经网络模型的轻量化,但模型分类准确率并未显著提高。为保持模型轻量化的基础上,进一步提高模型分类准确率。本章以轻量级网络为基础进行改进,虽然现存的轻量级模型在模型大小方面有很大的优势,但是准确率仍有较大提升空间,因此本章引入迁移学习及改进的注意力机制,以提高轻量级模型分类准确率,实现煤岩镜质组显微组分的有效分类。

4.1 轻量级网络模型

深度卷积神经网络在计算机视觉领域飞速发展并取得很好的成果,但是常规的卷积神经网络模型在达到较高的精度的同时,一般需要占用较大的存储空间和消耗较多的计算量。目前小型移动设备和嵌入式设备不断发展,为了更好的应用到这些日常的移动应用设备中,研究规模小、速度快的轻量级架构至关重要。因此,深度学习中轻量级网络模型的设计成为了深度学习领域的研究热点。在轻量级网络模型的研究中,较为经典的轻量级模型有 SqueezeNet、MobileNet V2 和 ShuffleNet V2。

(1) SqueezeNet 模型提出了包含压缩层(Squeeze)和扩展层(Expand)的 Fire 模块为核心。其中 Squeeze 层是由大小为 1×1 卷积核组成的卷积层,对上层输出的特征图进行卷积,以压缩特征图通道数降低维数,减少计算量和参数量。Expand 层包含了两个并行的卷积即 1×1 的卷积核和 3×3 的卷积核,后将卷积运算结果 concat 得到输出,使用 1×1 卷积核和 3×3 卷积核,既能降低模型参数量,又能获取到不同层次的特征以增强模型表达能力。

(2) MobileNet V2 模型是一种新型的轻量级卷积神经网络模型。其核心思想在于采用反向残差结构(Inverted residual block)和深度可分离卷积。其中反向残差结构主要借鉴 ResNet 的残差思想,该结构先使用 1×1 的卷积将输入特征图维度扩展,再采用 3×3 的深度卷积(depthwise convolution, DW)提取特征,最后用 1×1 的卷积来缩小维度,减少模型的参数量和计算量,实现模型的轻量化,提升模型性能。

(3) ShuffleNet V2 模型的设计借鉴了 ResNet, MobileNet 等模型的核心思想,引入了深度可分离卷积、通道拆分(channel split)、通道混洗(channel shuffle)等操作,并提出了一些设计高效轻量级网络的准则,满足了模型低计算复杂度和高性能的要求。

4.2 ShuffleNet V2 模型结构

目前轻量级网络已成为 CNN 领域的研究热点，一系列轻量级网络被提出，包括 4.1 节介绍的 SqueezeNet、MobileNet V2、ShuffleNet V2 等。虽然这些轻量级模型在模型大小方面有很大的优势，但是准确率仍有较大提升空间。其中，ShuffleNet V2 提出了一些设计高效轻量级网络的准则，并因其低计算复杂度已成为轻量化架构设计领域的重大突破，引起了广泛关注。本章使用 ShuffleNet V2 作为主干网络构建煤岩镜质组分类网络，并对该模型进行改进以提升准确率，本节详细介绍 ShuffleNet V2 模型的结构及其特点。

ShuffleNet V2 提出了四条设计高效轻量级网络的准则：过多的分组卷积会增加内存访问操作所消耗的计算；卷积层的输入和输出特征通道数相等时内存访问操作所消耗的计算最小；过多的网络分支会降低并行性；减少元素级操作。针对以上准则，提出以下方法来设计 ShuffleNet V2 网络。如图 4.1(a)基本单元所示，使用通道拆分(channel split)操作代替分组操作，在每个基本单元的开头，将 c 个特征通道的输入分为两个分支，分别具有 $c-c'$ 和 c' 个通道。拆分后的一个分支不做任何操作等同于恒等映射向下传递，另一个分支有三个卷积操作，且满足输入和输出特征通道数相等的准则。同时卷积操作使用 1×1 的卷积代替逐点组卷积，以减少过多使用组卷积带来的缺陷。三个卷积操作完成后，左右两个分支采用拼接(concat)而不是相加(add)的操作进行合并，减少了元素级操作，合并后使用通道混洗(channel shuffle)操作实现信息交流。由于池化操作对分类网络模型影响较大，如果没有池化操作对特征图进行下采样，网络模型的复杂度会较大，ShuffleNet V2 针对池化操作的缺点，设计了一个空间下采样单元。对于空间下采样单元，如图 4.1(b)所示，为保证在下采样时增加整体输出通道的数量，去掉了通道拆分操作，因此输出通道的数量翻了两倍。图 4.1 详细展示了 ShuffleNet V2 模型的基本单元和空间下采样单元。

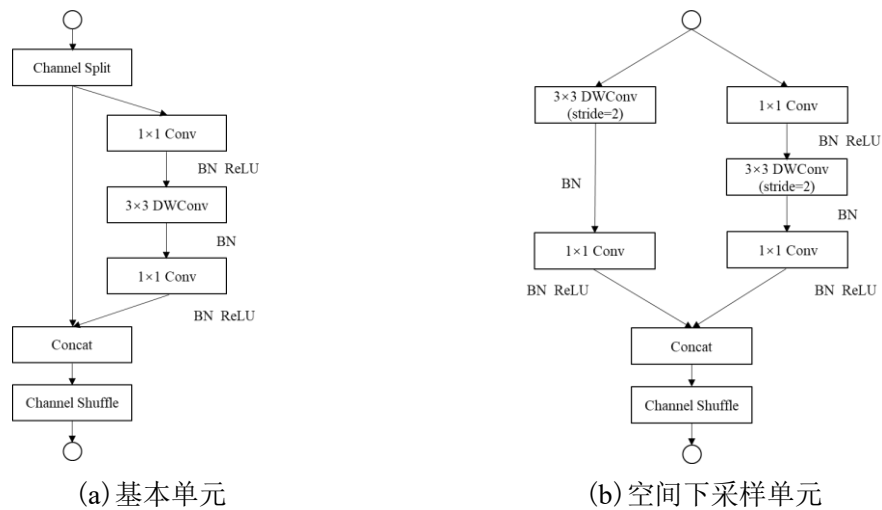


图 4.1 ShuffleNet V2 结构图

表 4.1 是 ShuffleNet V2 的网络结构, 如表 4.1 所示, 先对输入图片进行一次卷积和最大池化操作, 然后堆叠 Stage2、Stage3、Stage4, 其中 Stage2 和 Stage4 由 1 个空间下采样单元和 3 个基本单元组成, Stage3 由 1 个空间下采样单元和 7 个基本单元组成。最后通过全局池化操作后, 进行全连接, 用于分类。

表 4.1 ShuffleNet V2 网络结构

| Layer | Output size | KSize | Stride | Repeat | Output channels |
|------------|------------------|--------------|--------|--------|-----------------|
| Image | 224×224 | —— | —— | —— | 3 |
| Conv1 | 112×112 | 3×3 | 2 | 1 | 24 |
| MaxPool | 56×56 | 3×3 | 2 | | |
| Stage2 | 28×28 | —— | 2 | 1 | 48 |
| | 28×28 | | 1 | 3 | |
| Stage3 | 14×14 | —— | 2 | 1 | 96 |
| | 14×14 | | 1 | 7 | |
| Stage4 | 7×7 | —— | 2 | 1 | 192 |
| | 7×7 | | 1 | 3 | |
| Conv5 | 7×7 | 1×1 | 1 | 1 | 1024 |
| GlobalPool | 1×1 | 7×7 | —— | —— | —— |
| FC | —— | —— | —— | —— | 1000 |

4.3 ShuffleNet V2 模型的改进

4.3.1 迁移学习

迁移学习是通过从源域数据中提取有用信息并将其用于目标域来解决这种跨领域学习的问题^[86]。用深度学习模型实现图像分类, 通常需要较为复杂的调参、较长的训练时间以及大量带有标注的数据, 而数据的获得和标注代价昂贵。迁移学习可以有效解决这些问题, 降低了时间的消耗, 且只需少量数据和较低的算力。深度网络的微调是最简单的深度网络迁移学习方法, 就是针对自己的任务对已经训练好的网络再进行调整^[87]。迁移学习方法中的微调优势明显, 其在训练数据不足的情况下加载在 ImageNet 数据集上预训练的模型权重, 经过 ImageNet 数据集预训练的模型具有更好的泛化能力, 同时避免了从头开始训练模型, 减少了时间消耗。

本文使用的煤岩镜质组显微图像数据集样本数较少, 且与 ImageNet 数据集存在较大差别, 但这些图像所包含的基本纹理结构信息具有普适性, 即网络靠近输入的层所学习到的图像浅层特征也具有通用性。因此, 本章采用迁移学习微调的方法将预训练网络模型靠近输入的若干层冻结, 仅对网络模型靠近输出的层的权值进行微调以完成本章的分类任务。通过迁移学习的微调策略, 可以加快网络

模型的收敛速度，同时提高模型的分类准确率。

4.3.2 改进的注意力机制

注意力机制可根据注意力域的不同分为空间注意力、通道注意力及混合注意力。因为浅层特征通常包含空间细节信息，一般选择空间注意力机制关注有效的浅层特征；而深层特征通常包含全局上下文感知信息，不同通道之间有一定的语义差异，通常选择通道注意力机制关注有效的深层特征^[88]。由于本章采用迁移学习将模型靠近输入的若干层冻结，仅微调靠近输出层的权重，而神经网络靠近输出的层提取到的是深层特征，因此迁移学习 ShuffleNet V2 模型将提取得到深层特征，根据深层特征通常包含较高的抽象语义，无需过滤空间信息，本章使用通道注意力机制关注深层特征。由章节 3.3.2 可知 ECA 模块提供了一种轻量级的架构模块，以较低的计算成本来提高准确率。但这种通道注意力机制使用平均池化的方式对每个通道特征进行压缩，未能充分提取纹理特征，而煤岩镜质组显微组分纹理特征较为丰富。因此本章结合 ECA 的思想、迁移学习后 ShuffleNet V2 模型特点及煤岩显微图像的特点，对这种通道注意力机制进行了改进，采用全局最大池化和全局平均池化输出特征对应位置相加的方法，充分提取纹理特征和丰富的高层次特征，实现更有效的特征聚合。

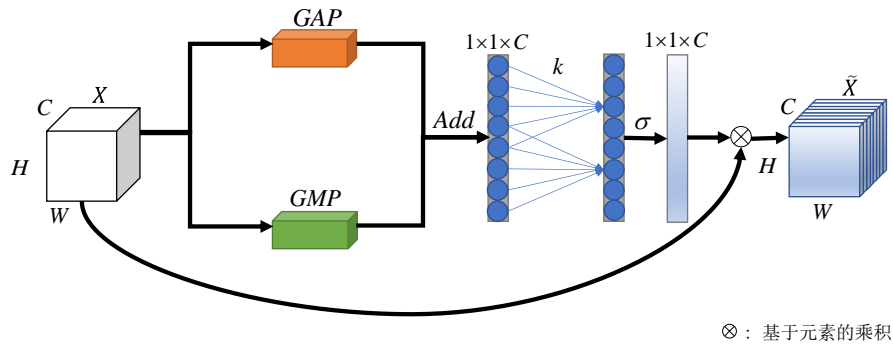


图 4.2 本章改进的注意力机制模块

本章采用的改进的通道注意力机制模块，其结构如图 4.2 所示。首先将输入的特征图采用全局平均池化和全局最大池化相加融合的方法来聚合特征；然后通过大小为 k 的快速一维卷积实现适当的跨通道信息交互，再通过 Sigmoid 函数生成通道权重，其中自适应选择一维卷积核大小 k 的方法同章节 3.3.2，上述过程如式(4.1)所示；最后将每个特征通道生成的权重作用于原始输入特征图，如式(4.2)所示。

$$w = \sigma(CID_k(add(GAP, GMP))) \quad (4.1)$$

式(4.1)中： GMP 和 GAP 分别表示全局最大池化和全局平均池化； add 表示将全局最大池化和全局平均池化输出特征对应位置相加； CID 表示一维卷积； k 表示一维卷积内核的大小； σ 表示 Sigmoid 激活函数； w 的维度为 $1 \times 1 \times C$ ，它表示每

个通道生成的特征图的权重。

$$\tilde{x}_c = F(x_c, w_c) = x_c w_c \quad (4.2)$$

式(4.2)中: x_c 是一个二维矩阵, 即第 c 个通道的特征图; w_c 是一个数, 也就是权重; F 表示基于元素的乘积操作, 相当于把二维矩阵中的每个值都乘 w_c 。

4.3.3 改进后的模型结构

ShuffleNet V2 网络是一种高效的轻量级网络, 其参数量和计算量很少, 易于应用到某些真实的应用场景如移动设备和嵌入式设备中。虽然 ShuffleNet V2 网络规模很小, 但是准确率有待进一步提高。而注意力机制在改善深度卷积神经网络的性能方面具有巨大的潜力, 其可以从众多信息中选择出对当前任务更关键的信息, 而抑制其他无关信息, 极大地提高分类的效率与准确率。由于本文使用的镜质组数据集样本数较少, 而迁移学习可以有效地解决训练数据不足的问题, 并加快网络模型的收敛速度。因此, 为提高煤岩镜质组显微组分的分类准确率, 提升深度学习分类模型的性能, 本章以 ShuffleNet V2 为主干网络构建煤岩镜质组分类网络, 并引入改进的注意力机制和迁移学习。

针对煤岩显微组分镜质组分类问题中数据集样本数量不充足, 部分显微亚组分类别间特征差异较小、局部特征甚至有交错较难区分, 轻量级网络准确率有待提高等问题, 本文提出一种基于注意力机制的轻量级煤岩镜质组分类网络(TA-ShuffleNet V2), 其网络模型结构如图 4.3 所示。首先采用数据增强的方法对训练数据增强, 提高模型泛化能力; 然后构建 ShuffleNet V2 模型提取特征, 采用迁移学习的方法对模型进行预训练和微调, 提取深层特征; 最后在 ShuffleNet V2 中嵌入注意力机制, 并根据煤岩镜质组显微图像的特点及网络结构特点对注意力机制进行改进, 构建轻量级端到端的分类网络模型, 实现煤岩显微组分镜质组的自动分类。

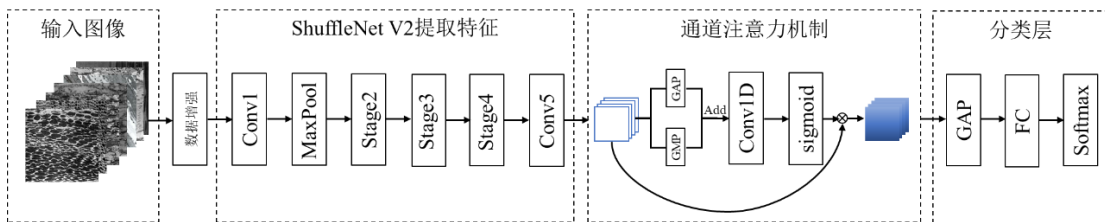


图 4.3 基于注意力机制的轻量级分类模型结构图

4.4 实验与结果讨论

本章构建模型使用的实验数据、实验环境、参数设置及模型评价指标均已在第二、三章进行了详细地介绍, 此处便不再赘述。

4.4.1 改进的 ShuffleNet V2 模型实验结果与分析

本章旨在实现轻量级模型的分类性能的提升,所以为了综合评估模型的整体分类性能,使用平均准确率、宏精确率、宏召回率与宏 F1 值四种评价指标进一步分析模型分类性能。改进前后模型的各项评价指标结果对比如表 4.2 所示。

表 4.2 ShuffleNet V2 与 TA-ShuffleNet V2 模型分类结果对比

| 模型 | 平均准确率(%) | 宏精确率(%) | 宏召回率(%) | 宏 F1 值(%) |
|------------------|----------|---------|---------|-----------|
| ShuffleNet V2 | 91.43 | 92.20 | 91.43 | 91.28 |
| TA-ShuffleNet V2 | 97.86 | 97.96 | 97.86 | 97.86 |

从表 4.2 可以看出,TA-ShuffleNet V2 模型的四项评价指标均明显高于 ShuffleNet V2 原模型。与 ShuffleNet V2 模型相比,本章提出的模型其平均准确率、宏精确率、宏召回率和宏 F1 值分别高于 ShuffleNet V2 模型 6.43%、5.76%、6.43%和 6.58%。说明在 ShuffleNet V2 模型中引入本章改进的注意力机制和迁移学习对镜质组显微组分的分类性能提升具有一定的效果。表明本章提出的 TA-ShuffleNet V2 模型整体分类性能更好。

为了更全面地描述深度学习模型,除了分类性能评价指标,还需要考虑模型的复杂度。评价深度学习模型复杂度的指标为参数量和 FLOPs。改进前后分类模型性能对比如表 4.3 所示。

表 4.3 ShuffleNet V2 与 TA-ShuffleNet V2 模型性能对比

| 模型 | 参数量 | FLOPs | 平均准确率(%) |
|------------------|-----------|------------|----------|
| ShuffleNet V2 | 1,366,792 | 42,524,256 | 91.43 |
| TA-ShuffleNet V2 | 1,366,795 | 42,578,528 | 97.86 |

从表 4.3 可以看出,与 ShuffleNet V2 原模型相比,本章构建的 TA-ShuffleNet V2 模型其参数量和 FLOPs 几乎不变,而平均准确率可达 97.86%,较原模型提升了 6.43%。表明本章构建的 TA-ShuffleNet V2 模型在保持轻量化的基础上显著地提高了模型的准确率。

4.4.2 模型收敛性对比实验结果与分析

为了更全面的验证本章提出的分类模型的有效性,本节对改进前后模型的稳定性及收敛性进行实验和分析。实验使用煤岩镜质组数据集对构建的 TA-ShuffleNet V2 模型和 ShuffleNet V2 进行训练和测试,图 4.4 展示了两种模型训练和测试的准确率及损失值迭代曲线。

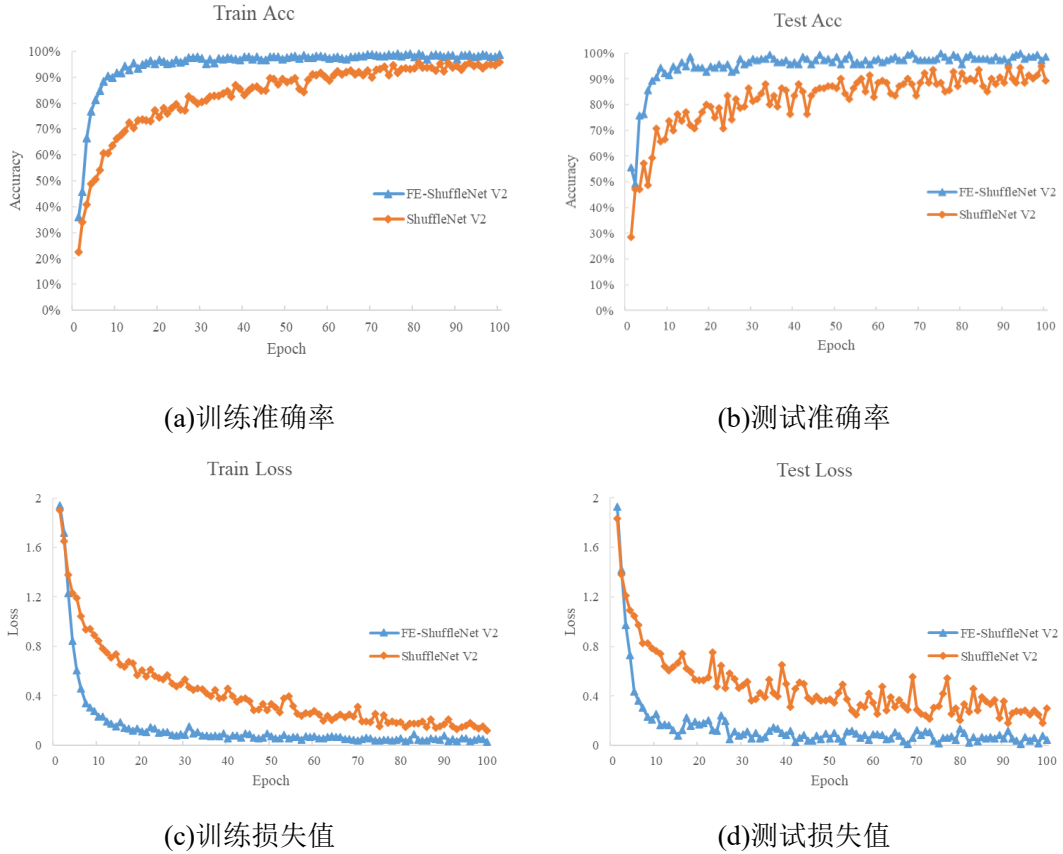


图 4.4 ShuffleNet V2 与 TA-ShuffleNet V2 模型收敛图

通过分析图 4.4 可知，TA-ShuffleNet V2 模型损失值较低，准确率高，且迭代 20 个 epoch 之后模型收敛基本趋于稳定，说明迁移学习加快了模型的收敛速度，改进的注意力机制提高了模型的准确率。而 ShuffleNet V2 原模型损失值较高，准确率较低，且迭代第 60 个 epoch 之后模型收敛，测试准确率曲线和测试损失值曲线较为震荡，稳定性较差。结合模型复杂度和分类性能的实验结果分析，验证了本章提出的 TA-ShuffleNet V2 模型不仅分类效果较好、模型复杂度低，而且模型收敛速度更快、稳定性更好。

4.4.3 不同注意力机制的模型对比实验结果与分析

为了验证本章所提出的改进的通道注意力机制模块的优越性，本节设计了一组实验，该组实验将 TA-ShuffleNet V2 模型中改进的通道注意力机制模块依次换成 SE^[81]、CBAM^[82]、SCSE^[83]和 ECA^[84]这四种不同的注意力机制模块，并对模型进行训练和测试实现分类。不同注意力机制模块性能对比结果如表 4.4 所示。

表 4.4 注意力机制性能对比

| 模型 | 参数量(M) | FLOPs(M) | 平均准确率(%) |
|-------|--------|----------|----------|
| +SE | 1.498 | 42.707 | 97.14 |
| +SCSE | 1.500 | 42.757 | 97.14 |

续表 4.4

| | | | |
|-------|-------|--------|-------|
| +CBAM | 1.498 | 42.842 | 97.14 |
| +ECA | 1.367 | 42.579 | 97.14 |
| 本章模型 | 1.367 | 42.579 | 97.86 |

由表 4.4 可以得知, 本文提出的改进的通道注意力机制平均准确率最高, 与 SE、SCSE 和 CBAM 相比参数量和 FLOPs 最少, 与 ECA 相比参数量和 FLOPs 基本一致但平均准确率有所提高。因此, 本章提出的改进的通道注意力机制模块有较为明显的优越性。

4.4.4 不同模型对比实验结果与分析

为了进一步验证本章所提出模型的在镜质组显微组分的识别中的优越性, 选用轻量级网络模型、经典的卷积神经网络以及 DE-DenseNet121 模型与本章提出的 TA-ShuffleNet V2 模型进行对比。其中轻量级网络模型包括 SqueezeNet、MobileNet V2 和 ShuffleNet V2, 经典的卷积神经网络模型包括 AlexNet、VGG16、ResNet50 和 DenseNet121。不同模型分类结果对比如表 4.5 所示。

表 4.5 不同模型分类结果对比

| 模型 | 平均准确率(%) | 宏精确率(%) | 宏召回率(%) | 宏 F1 值(%) |
|------------------|----------|---------|---------|-----------|
| SqueezeNet | 93.57 | 93.56 | 93.57 | 93.55 |
| MobileNet V2 | 93.57 | 94.09 | 93.57 | 93.64 |
| ShuffleNet V2 | 91.43 | 92.20 | 91.43 | 91.28 |
| AlexNet | 94.29 | 94.97 | 94.29 | 94.32 |
| VGG16 | 95.71 | 96.06 | 95.71 | 95.68 |
| ResNet50 | 95.00 | 95.30 | 95.00 | 94.84 |
| DenseNet121 | 95.71 | 96.46 | 95.71 | 95.83 |
| DE-DenseNet121 | 96.43 | 96.59 | 96.43 | 96.41 |
| TA-ShuffleNet V2 | 97.86 | 97.96 | 97.86 | 97.86 |

表 4.5 显示了 9 种模型的平均准确率、宏精确率、宏召回率和宏 F1 值的实验结果。由表 4.5 可以看出, 本章提出的 TA-ShuffleNet V2 模型的平均准确率、宏精确率、宏召回率和宏 F1 值均不仅高于其他几个轻量级网络和经典卷积神经网络, 而且高于第三章提出的 DE-DenseNet121 模型。表明本章所提出的模型分类效果更好, 能够更有效地实现煤岩镜质组显微组分的分类。

由于本章提出的模型旨在轻量级模型的基础上提高模型的高分类准确率, 因此为了更全面的描述本章模型的性能, 本章将轻量级网络模型、经典的卷积神经

网络以及 DE-DenseNet121 模型的参数量和计算量与 TA-ShuffleNet V2 模型进行对比。表 4.6 显示了 9 种不同模型的参数量、FLOPs 和平均准确率。

表 4.6 不同模型性能对比

| 模型 | 参数量 (M) | FLOPs(M) | 平均准确率(%) |
|------------------|---------|-----------|----------|
| AlexNet | 61.101 | 714.692 | 94.29 |
| VGG16 | 138.358 | 15483.862 | 95.71 |
| ResNet50 | 25.557 | 4111.515 | 95.00 |
| DenseNet121 | 7.979 | 2865.672 | 95.71 |
| SqueezeNet | 1.235 | 351.911 | 93.57 |
| MobileNet V2 | 3.505 | 314.130 | 93.57 |
| ShuffleNet V2 | 1.367 | 42.524 | 91.43 |
| DE-DenseNet121 | 1.210 | 443.145 | 96.43 |
| TA-ShuffleNet V2 | 1.367 | 42.579 | 97.86 |

通过表 4.6 可知，TA-ShuffleNet V2 模型与经典的卷积神经网络 AlexNet、VGG16、ResNet50 和 DenseNet121 相比，模型的参数分别约为经典的卷积神经网络的 1/45、1/101、1/19 和 1/6，模型的 FLOPs 分别约为其 1/17、1/364、1/97 和 1/67，其平均准确率也分别提高了 3.57%、2.15%、2.86%和 2.15%，准确率提升的同时参数量和计算量大幅度降低；与轻量级模型 SqueezeNet、MobileNet V2 和 ShuffleNet V2 相比，虽然 TA-ShuffleNet V2 模型的参数量和 FLOPs 不是最少的但仍然相对较少，且其平均准确率分别提高了 4.29%、4.29%和 6.43%；与 DE-DenseNet121 模型相比，TA-ShuffleNet V2 模型虽然参数量少量增加，但是 FLOPs 仅约为 DE-DenseNet121 模型的 1/10，且平均准确率也提高了 1.43%。

为了更直观地观察不同分类模型性能，采用不同分类模型的准确率、参数量和计算量对比图进行观察分析，如图 4.5、4.6 所示，由图 4.5 可以看出 TA-ShuffleNet V2 模型的准确率最高，参数量也相对较少；由图 4.6 可以看出 TA-ShuffleNet V2 模型的准确率最高，FLOPs 仅略高于 ShuffleNet V2 原模型。表明了本章提出的 TA-ShuffleNet V2 在保持模型轻量化的基础上大幅度地提高了模型的准确率，模型性能更优，实现了煤岩镜质组显微组分的高效分类。

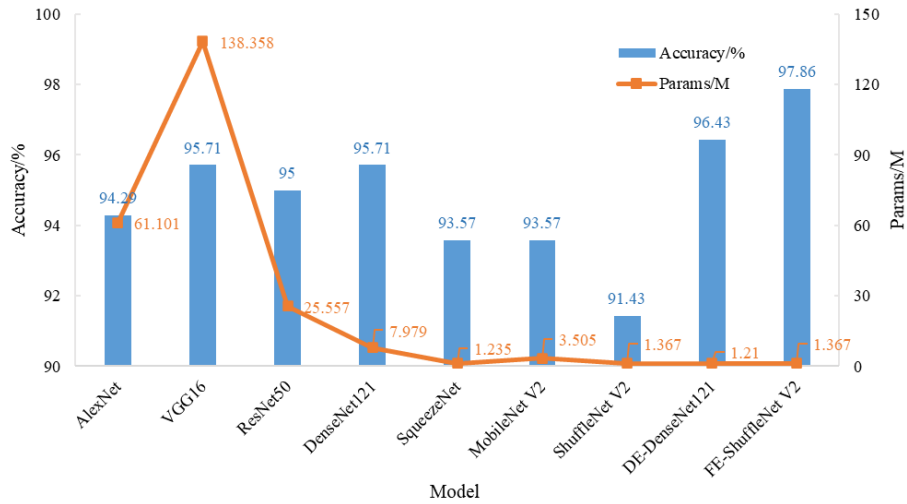


图 4.5 不同模型准确率和参数量对比图

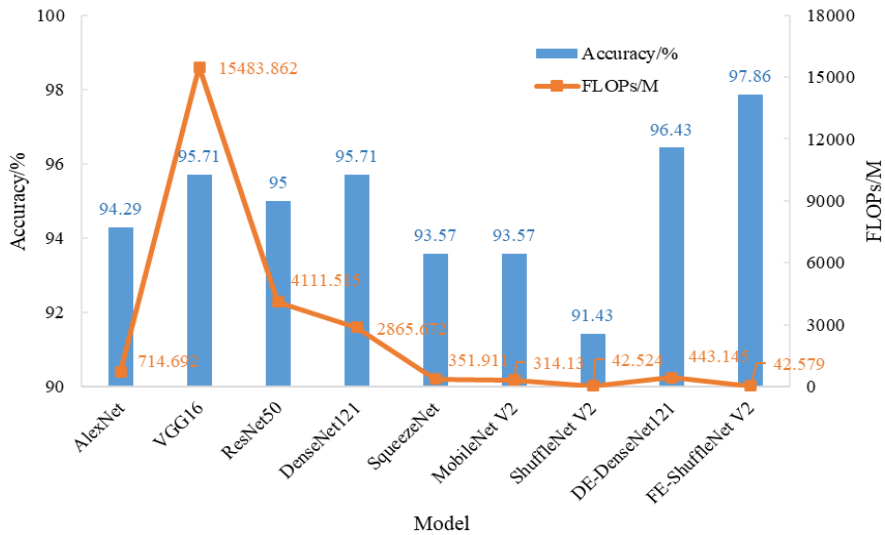


图 4.6 不同模型准确率和计算量对比图

4.5 本章小结

本章以保持模型轻量化的基础上，提高模型分类准确率为目标，构建了 TA-ShuffleNet V2 模型。首先该模型以 ShuffleNet V2 模型为主干网络提取特征，采用迁移学习的方法对模型进行预训练和微调，提取深层特征。接着在 ShuffleNet V2 中嵌入注意力机制，并根据煤岩镜质组显微图像的特点对注意力机制进行改进，构建轻量级端到端的分类网络模型，实现对煤岩显微组分镜质组 7 类显微组分的自动分类。最后使用各分类指标和模型复杂度评价指标对模型进行评价和分析，结果表明本章所构建的 TA-ShuffleNet V2 模型分类准确率显著提高，且模型仍保持轻量化，实现了煤岩镜质组的高效分类。

第五章 总结与展望

5.1 总结

本文的主要研究是围绕着构建基于深度学习的煤岩镜质组分类模型展开的。针对经典的深度卷积神经网络规模大且训练耗时，难以应用于移动设备或嵌入式设备上等问题，本文提出了基于深度卷积神经网络轻量化的分类方法完成对煤岩镜质组的分类，实现了模型的轻量化；针对轻量级模型准确率有待进一步提高，本文对轻量级模型进行改进，在保持模型轻量化的基础上，提高模型的分类准确率，实现对煤岩镜质组的高准确率分类。本文的主要研究内容如下：

（1）本文充分调研了煤岩显微组分、图像分类算法及深度学习的研究现状，引出了基于深度学习的煤岩镜质组分类研究。详细分析了本文的研究对象煤岩镜质组显微组分的特点，并对镜质组样本进行了预处理，采用数据增强的方式增加样本数量，以缓解因样本不充足而产生的模型过拟合现象。

（2）针对经典的深度卷积神经网络模型规模大，难以应用于内存有限的设备中的问题，提出了一种基于 DenseNet121 模型轻量化的煤岩镜质组分类方法。该模型以 DenseNet121 模型为基础，使用深度可分离卷积并优化增长率以减少模型参数量，同时引入了 ECA 注意力机制以提高模型分类准确率，对模型进行了重新构建。通过煤岩镜质组显微组分数据集对不同模型进行实验验证，证明了本文所构建的分类模型不仅与经典的卷积神经网络相比具有较少的参数量，与轻量级网络相比仍然具有较少的参数量，实现了深度卷积神经网络的轻量化。

（3）为进一步提升煤岩镜质组显微组分的分类准确率，本文以轻量级网络 ShuffleNet V2 为基础，引入迁移学习及改进的注意力机制，提出了基于注意力机制的轻量级网络煤岩镜质组分类方法。实验结果显示该模型平均分类准确率可达 97.86%，相较于原模型提高了 6.43%，且模型参数量和计算量基本保持不变。表明本文提出模型在保持模型轻量化的基础上，提高了模型的分类准确率，实现了煤岩镜质组显微组分的有效分类。

（4）本文所设计的深度学习模型均在 Windows10 平台下，搭建 Pytorch 深度学习框架，采用 Python 接口完成对网络模型的训练。并且本文所构建的模型能够自动地学习煤岩镜质组显微组分的特征信息，实现端到端的自动分类。

5.2 展望

本文将深度学习引入煤岩镜质组显微图像的分类任务中，所构建的分类模型实现了煤岩镜质组显微组分高精度的自动识别，同时模型的规模更小准确率更高，更易于应用到内存有限的设备中。本文的研究方法可以辅助煤岩学工作者更快速、

准确地辨别煤岩镜质组显微组分，具有一定的参考价值和实际应用意义。然而由于实验条件和时间的限制，更加深入的研究工作可从以下几个方面展开：

（1）增加煤岩镜质组显微组分样本数量。由于样本的数量对于深度学习模型的训练具有一定的影响，因此可以添加一些更具多样性的样本以丰富煤岩镜质组数据集，提高模型泛化能力。

（2）本文提出的卷积神经网络模型的轻量化仅对 DenseNet121 模型进行了轻量化研究，并没有实现其他经典卷积神经网络模型如 ResNet、VGG 和 AlexNet 等模型的轻量化，下一步可以尝试对其他经典卷积神经网络模型进行优化和改进，在降低模型大小同时提高煤岩镜质组显微组分的分类精度。

（3）本文提出的基于深度学习的煤岩镜质组分类方法具有良好的分类效果，并且模型规模较小适用于嵌入式设备和移动设备，未来可以尝试开发一个跨平台软件，实现算法理论有效地应用于实际工业领域。

（4）本文仅完成了煤岩镜质组显微组分的分类，未来在其他煤岩显微组分的分类研究中可以参考本文的方法进行改进和优化。

参考文献

- [1] 柳君波, 高俊莲, 徐向阳. 中国煤炭供应行业格局优化及排放[J]. 自然资源学报, 2019, 34(03):473-486.
- [2] 林伯强, 吴微. 中国现阶段经济发展中的煤炭需求[J]. 中国社会科学, 2018(02):141-161+207-208.
- [3] Wang G, Xu Y, Ren H. Intelligent and ecological coal mining as well as clean utilization technology in China: Review and prospects[J]. International Journal of Mining Science and Technology, 2019, 29(2):161-169.
- [4] 郭治, 杜铭华, 曲思建. 焦炭反应性及反应后强度预测模型研究与分析[J]. 煤炭学报, 2005(01):113-117.
- [5] 张代林, 李伟锋, 曾涛, 等. 炼焦煤的热解过程研究[J]. 煤炭学报, 2012, 37(02):323-327.
- [6] 张秀仪, 龚至从, 门桂珍, 等. 烟煤的成因——工业分类[J]. 煤炭学报, 1981(02):30-37.
- [7] 叶道敏, 肖文钊, 罗俊文, 等. 煤岩配煤和焦炭强度的预测[J]. 燃料与化工, 1998(05):233-236.
- [8] 全国煤岩标准化技术委员会. 烟煤显微组分分类: GB/T 15588-2013[S]. 北京: 中国标准出版社, 2014.
- [9] 陈家良, 邵震杰, 秦勇. 能源地质学[M]. 徐州: 中国矿业大学出版社, 2004:32-33.
- [10] Singh P K, Singh V K, Singh M P, et al. Petrographic characteristics and Paleoenvironmental history of Eocene lignites of Cambay basin, Western India[J]. International Journal of Coal Science & Technology, 2017, 4(3):214-233.
- [11] 李文华, 白向飞, 杨金和, 等. 烟煤镜质组平均最大反射率与煤种之间的关系[J]. 煤炭学报, 2006(03):342-345.
- [12] 白向飞. 焦化生产中煤岩学应用现状及对策分析[J]. 煤质技术, 2009(02):1-3.
- [13] Jin L, Han K, Wang J, et al. Direct liquefaction behaviors of Bulianta coal and its macerals[J]. Fuel Processing Technology, 2014, 128:232-237.
- [14] Feng J, Li J, Li W. Influences of chemical structure and physical properties of coal macerals on coal liquefaction by quantum chemistry calculation[J]. Fuel processing technology, 2013, 109:19-26.
- [15] 张磊, 刘文礼, 门东坡. 煤岩显微组分测定方法的研究现状及几点建议[J]. 煤炭工程, 2013, 45(04):97-99+103.
- [16] England B M, Mikka R A, Bagnall E J. Petrographic characterization of coal using automatic image analysis[J]. Journal of Microscopy, 1979, 116(3):329-336.

- [17] Riepe W, Steller M. Characterization of coal and coal blends by automatic image analysis[J]. Fuel, 1984, 63(3):313-317.
- [18] Lester E, Allen M, Cloke M, et al. An automated image analysis system for major maceral group analysis in coals[J]. Fuel, 1994, 73(11):1729-1734.
- [19] 王素婷, 朱宪坤, 吕青. 基于 RILBP-GLCM 算法的煤岩显微组分识别[J]. 煤炭技术, 2017, 36(03):142-144.
- [20] Skiba M, Młynarczuk M. Identification of macerals of the inertinite group using neural classifiers, based on selected textural features[J]. Archives of Mining Sciences, 2018, 63(4).
- [21] Wang H, Lei M, Chen Y, et al. Intelligent Identification of Maceral Components of Coal Based on Image Segmentation and Classification[J]. Applied Sciences, 2019, 9(16):3245.
- [22] 王培珍, 董双, 张代林, 等. 基于圆形 LBP 均匀模式的煤镜质组显微组分纹理分析[J]. 安徽工业大学学报(自然科学版), 2014, 31(02):147-151.
- [23] 王培珍, 殷子皖, 丁海涛, 等. 基于小波变换的煤岩惰质组显微图像分类[J]. 安徽工业大学学报(自然科学版), 2016, 33(03):278-283.
- [24] 王培珍, 殷子皖, 王高, 等. 一种基于 PCA 与 RBF-SVM 的煤岩显微组分镜质组分类方法[J]. 煤炭学报, 2017, 42(04):977-984.
- [25] 王培珍, 刘婕梅, 汪文艳, 等. 基于轮廓波变换的煤壳质组显微组分分类[J]. 煤炭学报, 2018, 43(S2):641-645.
- [26] 王培珍, 翟羽佳, 王慧, 等. 基于曲波变换和压缩感知的煤岩惰质组分类[J]. 中国矿业大学学报, 2019, 48(05):1119-1125.
- [27] LIU M, WANG P, CHEN S, et al. The Classification of Inertinite Macerals in Coal Based on the Multifractal Spectrum Method[J]. Applied Sciences, 2019, 9(24):5509.
- [28] 苏赋, 吕沁, 罗仁泽. 基于深度学习的图像分类研究综述[J]. 电信科学, 2019, 35(11):58-74.
- [29] 张学工. 关于统计学习理论与支持向量机[J]. 自动化学报, 2000(01):36-46.
- [30] 王宏漫, 欧宗瑛. 采用 PCA/ICA 特征和 SVM 分类的人脸识别[J]. 计算机辅助设计与图形学学报, 2003(04):416-420+431.
- [31] 张浩, 赵云胜, 陈冠宇, 等. 基于支持向量机的遥感图像建筑物识别与分类方法研究[J]. 地质科技情报, 2016, 35(06):194-199.
- [32] 毋雪雁, 王水花, 张煜东. K 最近邻算法理论与应用综述[J]. 计算机工程与应用, 2017, 53(21):1-7.
- [33] Ameur B, Masmoudi S, Derbel A G, et al. Fusing Gabor and LBP feature sets for

- KNN and SRC-based face recognition[C]// 2016 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP). IEEE, 2016.
- [34]Sun K, Kang H, Park H H. Tagging and Classifying Facial Images in Cloud Environments Based on KNN using MapReduce[J]. Optik - International Journal for Light and Electron Optics, 2015, 126(21):S0030402615006324.
- [35]石洪波, 王志海, 黄厚宽, 等. 一种限定性的双层贝叶斯分类模型[J]. 软件学报, 2004, (02):193-199.
- [36]Zhang X, Hao S, Xu C, et al. Image classification based on low-rank matrix recovery and Naive Bayes collaborative representation[J]. Neurocomputing, 2015, 169(dec.2):110-118.
- [37]郭金鑫, 陈玮. 基于 HOG 多特征融合与随机森林的人脸识别[J]. 计算机科学, 2013, 40(10):279-282+317.
- [38]刘毅, 杜培军, 郑辉, 等. 基于随机森林的国产小卫星遥感影像分类研究[J]. 测绘科学, 2012, 37(04):194-196.
- [39]苏赋, 吕沁, 罗仁泽. 基于深度学习的图像分类研究综述[J]. 电信科学, 2019, 35(11):58-74.
- [40]Homik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators[J]. Neural Networks, 1989, 2(5):359-366.
- [41]卢宏涛, 张秦川. 深度卷积神经网络在计算机视觉中的应用研究综述[J]. 数据采集与处理, 2016, 31(01):1-17.
- [42]Hush D, Abdallah C, Horne B. The recursive neural network and its applications in control theory[J]. Computers & Electrical Engineering, 1993, 19(4):333-341.
- [43]Hubel D H, Wiesel, T. Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex[J]. Journal of Physiogoly, 1962, 160:106-154.
- [44]Fukushima K. Neocognitron: A self-organizing neural network for a mechanism of pattern recognition unaffected by shift in position[J]. Biological Cybernetics. 1980, 36(4):193-202.
- [45]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.
- [46]Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]//International Conference on Neural Information Processing Systems. New York: Curran Associates, 2012:1097-1105.
- [47]Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer Science, 2014.

- [48]He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]//In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016:770-778.
- [49]Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015:1-9.
- [50]Huang G, Liu Z, Laurens V, et al. Densely Connected Convolutional Networks[C]//In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2017:4700-4708.
- [51]何雪英, 韩忠义, 魏本征. 基于深度学习的乳腺癌病理图像自动分类[J]. 计算机工程与应用, 2018, 54(12):121-125.
- [52]刘大伟, 韩玲, 韩晓勇. 基于深度学习的高分辨率遥感影像分类研究[J]. 光学学报, 2016, 36(04):306-314.
- [53]张泽中, 高敬阳, 吕纲, 等. 基于深度学习的胃癌病理图像分类方法[J]. 计算机科学, 2018, 45(S2):263-268.
- [54]李彦枝, 陈昌红, 谢晓芳. 基于改进卷积神经网络的极光图像分类算法研究[J]. 南京邮电大学学报(自然科学版), 2019, 39(06):86-93.
- [55]Iandola F N, Han S, Moskewicz M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size[J]. arXiv preprint arXiv:1602.07360, 2016.
- [56]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.
- [57]Ma N, Zhang X, Zheng H T, et al. Shufflenet v2: Practical guidelines for efficient cnn architecture design[C]//Proceedings of the European conference on computer vision (ECCV). 2018:116-131
- [58]李坤伦, 魏泽发, 宋焕生. 基于 SqueezeNet 卷积神经网络的车辆颜色识别[J]. 长安大学学报(自然科学版), 2020, 40(04):109-116.
- [59]Rabano S L, Cabatuan M K, Sybingco E, et al. Common Garbage Classification Using MobileNet[C]//International Conference on Humanoid. IEEE, 2018.
- [60]刘梦伦, 赵希梅, 魏宾. 基于 MobileNet V2-ELM 的肝硬化识别[J]. 青岛大学学报(自然科学版), 2019, 32(04):17-21.
- [61]杨舒, 王玉德. 基于 Contourlet 变换和 Hu 不变矩的图像检索算法[J]. 红外与激光工程, 2014, 43(01):306-310.
- [62]Skilansky J. Image segmentation and feature extraction[J]. IEEE Transactions on

- Systems, Man, and Cybernetics, 1978, 8(5):237-247.
- [63]刘丽, 匡纲要. 图像纹理特征提取方法综述[J]. 中国图象图形学报, 2009, 14(04):622-635.
- [64]高友文, 周本君, 胡晓飞. 基于数据增强的卷积神经网络图像识别研究[J]. 计算机技术与发展, 2018, 28(08):62-65.
- [65]周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.
- [66]陈超, 齐峰. 卷积神经网络的发展及其在计算机视觉领域中的应用综述[J]. 计算机科学, 2019, 46(03):63-73.
- [67]Cheng X, Zhang Y, Chen Y. Pest identification via deep residual learning in complex background[J]. Computers and Electronics in Agriculture, 2017, 141:351-356.
- [68]R. Memisevic, C. Zach, G. E. Hinton, et al. Gated Softmax Classification[C]. Proceedings of the 23rd International Conference on Neural Information Processing Systems, Vancouver, 2010, 1603-1611.
- [69]Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]. Proceedings of the International Conference on Machine Learning. 2015:448-456.
- [70]张荣磊, 田爱奎, 谭浩, 等. 基于卷积神经网络的图像识别算法研究[J]. 山东理工大学学报(自然科学版), 2018, 32(01):48-50.
- [71]Howard A G, Zhu M, Chen B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications[J]. arXiv preprint arXiv:1704.04861, 2017.
- [72]Wang N, Song Y, Ma C, et al. Unsupervised Deep Tracking[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020.
- [73]Luong M T, Pham H, Manning C D. Effective Approaches to Attention-based Neural Machine Translation[J]. Computer Science, 2015.
- [74]Lee J B, Rossi R A, Kim S, et al. Attention Models in Graphs: A Survey[J]. ACM Transactions on Knowledge Discovery from Data, 2019, 13(6):1-25.
- [75]Seo Y, Shin K S. Image classification of fine-grained fashion image based on style using pre-trained convolutional neural network[C]// 2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA). IEEE, 2018.
- [76]Sun B, Yang L, Zhang W, et al. Demonstration of applications in computer vision and nlp on ultra power-efficient cnn domain specific accelerator with 9.3 tops/watt[C]//2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). IEEE, 2019:611-611.
- [77]Nidadavolu P S, Iglesias V, Villalba J, et al. Investigation on neural bandwidth

- extension of telephone speech for improved speaker recognition[C]//ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2019:6111-6115.
- [78]Wang D, Wang B, Zhou Y. Twinsadvnet: Adversarial learning for semantic segmentation[C]//2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP). IEEE, 2019:1-4.
- [79]Lertpiya A, Chaiwachirasak T, Maharattanamalai N, et al. A preliminary study on fundamental thai nlp tasks for user-generated web content[C]//2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP). IEEE, 2018:1-8.
- [80]Xu K, Ba J, Kiros R, et al. Show, Attend and Tell: Neural Image Caption Generation with Visual Attention[J]. Computer Science, 2015:2048-2057.
- [81]Hu J, Shen L, Sun G, et al. Squeeze-and-Excitation Networks[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018:7132-7141.
- [82]Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module[C]//Proceedings of the European conference on computer vision (ECCV). 2018:3-19.
- [83]Roy A G, Navab N, Wachinger C. Concurrent spatial and channel ‘squeeze & excitation’ in fully convolutional networks[C]//International conference on medical image computing and computer-assisted intervention. Springer, Cham, 2018:421-429.
- [84]Wang Q, Wu B, Zhu P, et al. ECA-net: Efficient channel attention for deep convolutional neural networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020:11534-11542.
- [85]Chen Y, Li J, Xiao H, et al. Dual path networks[C]//Advances in Neural Information Processing Systems. 2017:4467-447
- [86]Shao L, Zhu F, Li X. Transfer Learning for Visual Categorization: A Survey[J]. IEEE Transactions on Neural Networks & Learning Systems, 2015, 26(5): 1019.
- [87]Weiss K, Khoshgoftaar T M, Wang D D. A survey of transfer learning[J]. Journal of Big Data, 2016, 3(1):1-40.
- [88]Zhao T, Wu X. Pyramid feature attention network for saliency detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019:3085-3094.

在学研究成果

发表论文：

- [1] 期刊论文, 发表于《煤炭科学技术》(中国知网 检索)
- [2] IEEE国际会议论文, 2021年“The 11th International Conference on Image and Graphics(ICIG)”(已录用)

申请专利：

- [1] 中国发明专利, (实审中).
- [2] 中国发明专利, (实审中).

致 谢

附录 A 插图清单

| | |
|---|----|
| 图 2.1 部分镜质组显微组分典型图像..... | 8 |
| 图 2.2 数据增强样本..... | 11 |
| 图 3.1 卷积神经网络的结构..... | 12 |
| 图 3.2 卷积操作示例..... | 12 |
| 图 3.3 池化操作示意图..... | 13 |
| 图 3.4 全连接层的结构..... | 14 |
| 图 3.5 常见激活函数图像..... | 14 |
| 图 3.6 DenseNet 网络基本结构..... | 16 |
| 图 3.7 Dense Block 模块结构..... | 17 |
| 图 3.8 标准卷积..... | 19 |
| 图 3.9 深度可分离卷积..... | 19 |
| 图 3.10 Dense Layer 结构..... | 20 |
| 图 3.11 ECA 结构..... | 22 |
| 图 3.12 DE-DenseNet121 网络结构图..... | 24 |
| 图 3.13 DenseNet121 和 DE-DenseNet121 模型收敛图..... | 29 |
| 图 3.14 不同模型准确率和参数量对比图..... | 32 |
| 图 3.15 不同模型准确率和计算量对比图..... | 32 |
| 图 4.1 ShuffleNet V2 结构图..... | 35 |
| 图 4.2 本章改进的注意力机制模块..... | 37 |
| 图 4.3 基于注意力机制的轻量级分类模型结构图..... | 38 |
| 图 4.4 ShuffleNet V2 与 TA-ShuffleNet V2 模型收敛图..... | 40 |
| 图 4.5 不同模型准确率和参数量对比图..... | 43 |
| 图 4.6 不同模型准确率和计算量对比图..... | 43 |

附录 B 插表清单

| | |
|---|----|
| 表 2.1 镜质组显微组分..... | 7 |
| 表 3.1 批量归一化过程..... | 15 |
| 表 3.2 DenseNet121 和 DE-DenseNet121 网络结构配置信息表 | 24 |
| 表 3.3 本文实验环境配置..... | 25 |
| 表 3.4 训练参数设置..... | 26 |
| 表 3.5 二分类混淆矩阵..... | 26 |
| 表 3.6 DenseNet121 与 DE-DenseNet121 模型性能对比 | 27 |
| 表 3.7 DenseNet121 与 DE-DenseNet121 模型分类结果对比 | 28 |
| 表 3.8 不同增长率配置实验对比..... | 29 |
| 表 3.9 DE-DenseNet121 模型消融实验 | 30 |
| 表 3.10 不同模型分类结果对比..... | 30 |
| 表 3.11 不同模型性能对比..... | 31 |
| 表 4.1 ShuffleNet V2 网络结构 | 36 |
| 表 4.2 ShuffleNet V2 与 TA-ShuffleNet V2 模型分类结果对比 | 39 |
| 表 4.3 ShuffleNet V2 与 TA-ShuffleNet V2 模型性能对比 | 39 |
| 表 4.4 注意力机制性能对比..... | 40 |
| 表 4.5 不同模型分类结果对比..... | 41 |
| 表 4.6 不同模型性能对比..... | 42 |