

电 子 科 技 大 学
UNIVERSITY OF ELECTRONIC SCIENCE AND TECHNOLOGY OF CHINA

专业学位硕士学位论文

MASTER THESIS FOR PROFESSIONAL DEGREE



论文题目 基于 CT 图像的新冠肺炎检测研究

专业学位类别	<u>工程硕士</u>
学 号	<u>201922140220</u>
作者姓名	<u>沈昊钧</u>
指导教师	<u>张会雄 副教授</u>
学 院	<u>生命科学与技术学院</u>

分类号 R563.1;TP183 密级 公开
UDC 注 1

学 位 论 文

基于 CT 图像的新冠肺炎检测研究

(题名和副题名)

沈昊钧

(作者姓名)

指导教师 张会雄 副教授
电子科技大学 成 都

(姓名、职称、单位名称)

申请学位级别 硕士 专业学位类别 工程硕士
专业学位领域 生物医学工程
提交论文日期 2022 年 4 月 26 日 论文答辩日期 2022 年 5 月 27 日
学位授予单位和日期 电子科技大学 2022 年 6 月
答辩委员会主席 廖伟
评阅人

注 1: 注明《国际十进分类法 UDC》的类号。

Research on the Detection of COVID-19 Based on CT Images

A Master Thesis Submitted to
University of Electronic Science and Technology of China

Discipline **Master of Engineering**

Student ID **201922140220**

Author **Shen Haojun**

Supervisor **Prof. Zhang Huixiong**

School **School of Life Science and Technology**

独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。据我所知，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，也不包含为获得电子科技大学或其它教育机构的学位或证书而使用过的材料。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作了明确的说明并表示谢意。

作者签名： 沈昊钧

日期： 2022年 5月 18日

论文使用授权

本学位论文作者完全了解电子科技大学有关保留、使用学位论文的规定，有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。本人授权电子科技大学可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后应遵守此规定）

作者签名： 沈昊钧

导师签名： 张会雄

日期： 2022年 5月 18日

摘 要

计算机断层扫描是目前临床上胸部影像学检查最有效的辅助诊断手段之一。然而人工分析 CT 影像往往对专业知识的要求较高，且对影像特征的分析较为费时费力。利用深度学习进行 CT 影像的识别与诊断，不仅可以提高诊断的效率，还可以在在一定程度上避免误诊和漏诊。

本文希望通过深度学习的方法，得到对新冠肺炎 CT 图像分类效果较好的模型，来帮助临床医生尽早识别、诊断新冠肺炎。本文研究的主要内容是通过对卷积神经网络结构进行改进以及采用迁移学习的方式，来提高对新冠肺炎 CT 图像分类准确率。

在对卷积神经网络进行改进的方式中，本文采用了对卷积神经网络分类层结构进行改进的方式以及在卷积神经网络中嵌入 SE 模块、CBAM 模块的方式。其中对卷积神经网络分类层结构进行改进的实验中 InceptionV3、ResNet50、DenseNet169 三种模型都体现出较好的性能，在对 SARS-CoV-2 CT 数据集的分类中，准确率上都达到了 95%左右，相对于原模型都有所提升。对于在改进后的 ResNet50 模型中，嵌入了 SE 模块和 CBAM 模块，其分类准确率得到了进一步的提升。

对于数据量不够充足的情况下，迁移学习往往能体现出更好的学习性能。本文研究了基于卷积神经网络的迁移学习方式在新冠肺炎 CT 图像分类的效果。其中 InceptionV3、ResNet50、DenseNet169 三种模型微调方式的迁移学习训练准确率达到 97%以上，VGG19 模型微调方式的迁移学习训练分类准确率也有 91.54%。通过对 InceptionV3、ResNet50、DenseNet169 三种模型进行集成学习，集成模型得到了本文的最佳分类效果，分类准确率为 98.59%。

针对数据量更小的 COVID-CT 图像数据集，本文提出了基于微调和特征提取相结合的迁移学习方法，将 SARS-CoV-2 CT 数据集作为源域进行迁移学习，并与直接从 ImageNet 数据集上微调的迁移学习模型对比，实验结果表明了本方法的有效性。

最后本文用 Flask 框架搭建了新冠肺炎检测系统，对上传的新冠肺炎 CT 图像进行检测，提供了一种简便快捷的新冠肺炎检测方式。

关键词：新冠肺炎，CT 图像分类，卷积神经网络，迁移学习

ABSTRACT

Computed tomography is currently one of the most effective auxiliary diagnostic methods for chest imaging in clinical practice. However, manual analysis of CT images often requires high professional knowledge, and the analysis of image features is time-consuming and labor-intensive. Using deep learning to identify and diagnose CT images can not only improve the efficiency of diagnosis, but also avoid misdiagnosis and missed diagnosis to a certain extent.

This thesis hopes to obtain a model with good classification effect on CT images of COVID-19 through deep learning methods, so as to help clinicians identify and diagnose new coronary pneumonia as soon as possible. The main content of this thesis is to improve the classification accuracy of COVID-19 CT images by improving the convolutional neural network structure and transfer learning.

In the way of improving the convolutional neural network, this thesis adopts the way of improving the classification layer structure of the convolutional neural network and the way of embedding the SE module and the CBAM module in the convolutional neural network. Among them, the three models of InceptionV3, ResNet50 and DenseNet169 in the experiments of improving the classification layer structure of the convolutional neural network all showed good performance. In the classification of the SARS-CoV-2 CT dataset, the accuracy rate has reached about 95%, which is improved compared to the original model. For the improved ResNet50 model, SE module and CBAM module are embedded, and its classification accuracy has been further improved.

When the amount of data is not sufficient, transfer learning can often reflect better learning performance. This thesis studies the effect of the transfer learning method based on convolutional neural network in the classification of Corona Virus Disease 2019 CT images. Among them, the fine-tuned transfer learning training accuracy of InceptionV3, ResNet50, and DenseNet169 model has reached more than 97%, and the classification accuracy of fine-tuned transfer learning training of VGG19 model is also 91.54%. Through ensemble learning of InceptionV3, ResNet50, and DenseNet169 models, the ensemble model obtains the best classification effect in this thesis, and the classification accuracy rate is 98.59%.

For the COVID-CT image dataset with a smaller amount of data, this thesis proposes a transfer learning method based on a combination of fine-tuning and feature extraction. The SARS-CoV-2 CT dataset is used as the source domain for transfer

learning, and the transfer learning method is compared directly with ImageNet data. The experimental results show the effectiveness of this method.

Finally, this thesis uses the Flask framework to build a COVID-19 detection system to detect the uploaded COVID-19 CT images, providing a simple and fast COVID-19 detection method.

Keywords: COVID-19, CT image classification, convolutional neural network, transfer learning

目 录

第一章 绪论.....	1
1.1 研究背景.....	1
1.2 国内外研究现状.....	1
1.3 研究意义.....	4
1.4 本文主要工作.....	5
1.5 论文结构.....	5
第二章 相关技术与理论.....	7
2.1 卷积神经网络结构.....	7
2.1.1 卷积层.....	7
2.1.2 池化层.....	8
2.1.3 激活函数.....	9
2.1.4 全连接层.....	11
2.1.5 输出层.....	11
2.1.6 反向传播.....	12
2.2 卷积神经网络经典架构.....	14
2.2.1 AlexNet.....	14
2.2.2 VGGNet.....	15
2.2.3 GoogLeNet.....	15
2.2.4 ResNet.....	16
2.2.5 DenseNet.....	18
2.3 迁移学习.....	18
2.3.1 迁移学习介绍.....	18
2.3.2 迁移学习的分类.....	19
2.4 本章小结.....	20
第三章 基于卷积神经网络的新冠肺炎 CT 图像分类	21
3.1 数据集.....	21
3.2 实验环境.....	21
3.3 数据预处理.....	22
3.4 基于卷积神经网络改进的新冠肺炎 CT 图像分类	23
3.4.1 改进方法.....	23
3.4.2 模型训练.....	25

3.5 引入注意力机制的新冠肺炎 CT 图像分类	25
3.5.1 SE 模块	25
3.5.2 CBAM 模块	27
3.5.3 ResNet 引入注意力机制的新冠肺炎 CT 图像分类方法	27
3.5.4 模型训练	29
3.6 实验结果和分析	29
3.6.1 评价指标	29
3.6.2 卷积神经网络改进的实验结果与分析	31
3.6.3 ResNet 引入注意力机制的实验结果与分析	35
3.7 本章小结	36
第四章 基于迁移学习的新冠肺炎 CT 图像分类	37
4.1 数据集	37
4.2 迁移学习方法的选择	37
4.3 集成学习概述	38
4.4 基于迁移学习的新冠肺炎 CT 图像分类方法	39
4.4.1 微调方式的迁移学习训练	39
4.4.2 集成分类模型的构建	40
4.4.3 微调与特征提取相结合的迁移学习训练	41
4.5 基于迁移学习的新冠肺炎 CT 图像分类结果	41
4.5.1 微调方式的迁移训练和集成模型分类结果	41
4.5.2 微调与特征提取相结合的迁移训练分类结果	45
4.6 可视化研究	45
4.7 本章小结	47
第五章 新冠肺炎检测系统的实现	48
5.1 系统需求分析	48
5.1.1 系统编写目的	48
5.1.2 功能需求分析	48
5.2 Flask 框架	48
5.3 系统结构设计	49
5.4 系统功能实现	50
5.3.1 图像上传模块	50
5.3.2 图像检测模块	51
5.3.3 新冠肺炎检测系统演示	52
5.5 本章小结	52

第六章 总结与展望.....	53
6.1 总结.....	53
6.2 展望.....	54
致谢.....	55
参考文献.....	56

第一章 绪论

1.1 研究背景

新冠肺炎，是由 2019 新型冠状病毒（SARS-CoV-2）感染引起，官方将其命名为 COVID-19(Corona Virus Disease 2019)。第一批病例于 2019 年 12 月底在中国武汉发现，然后在全世界迅速传播^[1-3]。世界卫生组织官网的数据显示，截至 2022 年 3 月 4 日，全球新冠肺炎累计确诊病例达 440807756 例，累计死亡病例 5978096 例。新冠肺炎的常见症状是发热、咳嗽、呼吸短促、咽喉痛、头痛、腹泻^[4]，严重的会发生呼吸衰竭甚至休克。

目前官方认定的 SARS-CoV-2 检测方法是通过逆转录聚合酶链式反应 (Reverse transcription polymerase chain reaction, RT-PCR) 进行核酸检测^[5]。因为 SARS-CoV-2 在感染的早期就可以传播，所以通过 RT-PCR 的核酸检测非常适合对于无症状感染者的早期筛查。由于病毒感染一般从鼻咽部开始，COVID-19 核酸检测的标本采集方式主要是鼻拭子和咽拭子。然而核酸检测的方式存在试剂敏感性不佳导致漏诊误诊、检测时间较长、检测结果假阴性较高等问题，因此为提高检测成功率，往往需要结合其它检测手段多次进行病毒检测。

以医学影像学检查作为辅助诊断手段，在 SARS-CoV-2 早期诊断、疾病分期分级、治疗指导、疗效估计等过程中起着重要作用^[6]。胸部 X 光和胸部 CT 扫描等放射成像有助于及时隔离感染者并控制疫情^[7]。这些影像学技术可以很好地检测 COVID-19 的影像学特征。放射科医生的首选一般是胸部 X 光检查，因为大多数医院都配备了 X 光机。然而，从 X 光机获得的胸部图像无法准确区分软组织^[8]，检查早期病变的患者漏诊率较高，为了消除这个问题，一般以胸部 CT 扫描为主要的辅助诊断手段。《新型冠状病毒肺炎影像学诊断指南（2020 第一版）》建议使用 CT 扫描对新冠肺炎患者进行检查，利用原始数据进行高分辨 CT 扫描有利于胸部病变细节的观察和显示（如小叶间隔及磨玻璃影），对于辅助诊断具有重要价值；胸部 CT 检查方便易行，对于有高风险地区旅行史、确诊患者密切接触者及发热患者都是较好的筛查手段^[9]。

医疗人员可以根据肺部 CT 影像的不同特征进行病毒感染的分期，如早期主要表现为多发小结节状和间质改变^[10,11]，进展期病灶融合，境界不清，多发磨玻璃影、间质水肿增厚^[12,13]。

1.2 国内外研究现状

2019 新型冠状病毒在世界范围内的迅速蔓延，严重危害了人类的生命安全，

健康生存环境持续恶化。也给生产生活、国际贸易、公共卫生、国际交往等领域造成巨大损害。新冠病毒非常隐蔽，其潜伏期较长，无症状感染者的人数在不断增加，给病毒检测和疫情防控带来了很大的困难。目前用于新冠肺炎检测的方法和技术有核酸检测、抗体检测、胸部影像学检查等，其中核酸检测是目前最主要也是最广泛的新冠肺炎检测手段。然而核酸检测的方式存在漏诊误诊、检测结果假阴性较高等问题，因此为提高检测成功率，往往需要结合其它检测手段多次进行病毒检测。利用计算机辅助诊断技术来检测新冠肺炎成为当下热门的话题，已有相当多的国内外学者对此展开研究。

在深度学习还未引发人们的关注之前，计算机辅助诊断大部分采用的是传统机器学习的方法。在关于医疗数据分析的疾病辅助诊断领域，许多传统的机器学习模型仍被广泛使用，如支持向量机^[14]、随机森林^[15]、决策树^[16]等。由于目前新型冠状病毒核酸检测存在检测时间较长、检测结果假阴性较高等问题，高瞻^[17]等人设计了一种基于 XGBoost^[18]的新冠肺炎智能检测系统。他们对包含 169 例新冠肺炎确诊患者的 791 例样本，以生化、血液等指标为特征，借助于 XGBoost 算法进行模型训练和预测。该模型对新冠肺炎检测的准确度高达 94.34%。

胸部影像学检查可以发现新冠肺炎患者的典型影像学特征，不仅可以检测新冠肺炎，对疑似病例进行区分；还可以根据不同的影像学表现进分期，判断病情及其发展，对制定治疗方案和疗效评价等过程起着重要的指导作用。胸部影像学检查已成为诊断新冠肺炎的重要方式。其中胸部 X 射线（CXR）和计算机断层扫描（CT）是在新冠肺炎诊断中发挥重要作用的成像技术。然而疫情爆发期间医院每天会产生大量的胸部影像学检查图像，影像科医生的人工诊断压力过重。随着人工智能技术的发展，其在医疗领域的应用越来越广泛，使得利用计算机辅助医生进行诊断成为可能。同时大量的医学影像数据也吸引了众多学者进行胸部影像学的计算机辅助诊断研究。

一些学者尝试用传统机器学习的方法对胸部影像学图像进行新冠肺炎检测，也取得了不错的效果。如 Barstugan 等人^[19]利用 GLCM、LDP、GLRLM、GLSZM 和 DWT 等五种不同的特征提取技术从胸部 CT 图像中提取特征。这些提取的图像通过支持向量机（SVM）对 COVID-19 病例进行分类。他们使用了 10 折交叉验证，该方法的分类准确率为 99.68%。Prabir 等人^[20]使用基于深度特征加支持向量机(SVM)的方法来对 X 射线图像进行检测，其数据来自 GitHub 和 Kaggle 网站，包括 COVID-19、普通肺炎、正常三种；他们使用 13 种不同的卷积神经网络模型（AlexNet、DenseNet201、GoogleNet、InceptionResNetV2、InceptionV3、MobileNetV2、ResNet18、XceptionNet、ResNet101、ShuffleNet、VGG16、VGG19 和 ResNet50）来提取图像的深度特征，然后将卷积神经网络全连接层提取到的特

征送到 SVM 分类器进行分类。他们将每个模型独立训练 20 次，记录平均值，其中 ResNet50 得到了最好的实验结果，准确率达 95.33%。

相对于传统的计算机辅助诊断方法，深度学习可以自动学习样本的特征，大大减少了人工选择特征带来工作量。特别是卷积神经网络，由于其强大的特征提取能力，在图像识别和图像分类任务中被广泛使用。近年来，随着深度学习和各类卷积神经网络的不断发展，深度学习在医疗领域取得的效果越来越显著。大多数的研究者选择使用深度学习和各类卷积神经网络进行新冠肺炎识别。

Wang 等人^[21]使用 InceptionNet 检测 COVID-19 CT 图像。他们在 1065 张 CT 图像上测试了 InceptionNet 模型，其中 325 张 CT 图像为新冠肺炎阳性，740 张 CT 图像为新冠肺炎阴性，识别准确率为 89.50%。Xu 等人^[22]利用 3D CNN 模型在 CT 扫描图像中区分新冠病毒和甲型流感病毒性，使用 ResNet 提取特征。CNN 模型的准确率为 86.70%。Li 等人^[23]开发了新冠肺炎检测系统（CVNET），从胸部 CT 图像中提取出特征，用于检测冠状病毒感染。利用 COVNet 对 3322 名患者的 4356 张胸部 CT 图像进行训练。从 COVNet 获得的 COVID-19 的 AUC 值为 96%。出于隐私考虑，用于新冠病毒研究工作的 CT 图像未能公开。

Zheng 等人^[24]提出了一种弱监督的深度学习技术，用于诊断 COVID-19 患者的 3DCT 图像。他们使用预先训练过的 U-Net 技术对 3D 肺部图像进行分割，然后对感染区域进行预测。他们的模型得到的准确度为 95.9%。Das 等人^[25]开发了一种基于自动深度迁移学习的方法，通过使用 Xception 模型来检测胸部 X 光片中的 COVID-19 感染。其训练集和测试集分类准确率分别为 99.52%和 97.40%。

Pathak^[26]等人使用深度迁移学习技术来技术来构建 COVID-19 感染患者的分类模型，他们使用 ResNet50 网络提取胸部 CT 图像的潜在特征，并用十折交叉验证来防止过拟合问题。所提出的模型分别实现了高达 96.2264%和 93.0189%的训练和测试准确率。Jaiswa 等人^[27]借助卷积神经网络和预训练的 DenseNet201，设计了一种深度迁移学习模型。该模型对胸部 CT 扫描图像进行分类，其训练集、验证集、测试集的分类准确率分别为 99.82%、97.4%、96.25%。他们还将该模型与 VGG16、Inception ResNet、ResNet 152V2 等深度迁移模型进行对比，得出仍是 DenseNet201 表现更好。

Transformer 架构在自然语言处理任务中得到广泛应用，但在计算机视觉领域中的应用仍然非常有限。Vision Transformer(ViT)的出现打破了这一限制，它通过将图像拆分为块，并将这些图像块组合成序列作为输入，能够很好的执行图像分类任务；通过实验发现 Vision Transformer 与 CNN 相比，可获得更优异的结果，同时仅需更少的训练资源^[28]。因此也有学者尝试用 Vision Transformer 对新冠肺炎图像进行分类。如 Gao 等人^[29]用 Vision Transformer 和卷积神经网络中的

DenseNet 对 MIA-COV19 竞赛数据集中的胸部 CT 图像进行分类,得出 Vision Transformer 的准确率为 76.6%, 而 DenseNet 的准确率为 73.7%。

1.3 研究意义

新型冠状病毒肺炎的全球大流行给世界多个国家和地区带来了极大的危害与挑战。大多数新型冠状病毒肺炎患者都有典型的临床表现,但随着人们对 COVID-19 的进一步研究以及 SARS-CoV-2 的检测技术进一步提高,发现无症状感染者的人数在不断增加。许多 COVID-19 无症状感染者虽然没有明显的临床表现,但由于体内存在 SARS-CoV-2,可将病毒传播给他人,从而加速了此次疫情在世界范围内的大流行,并给全球疫情防控带来了极大的难度^[30]。因此,尽早发现并通过各种检测手段确诊无症状感染者是控制新冠疫情大流行的有效措施。

目前用于新冠肺炎检测的方法和技术有核酸检测、抗体检测、胸部影像学检查等。其中胸部影像学检查可以发现新冠肺炎患者的典型影像学特征,不仅可以检测新冠肺炎,对疑似病例进行区分;还可以根据不同的影像学表现进分期,判断病情及其发展,对制定治疗方案和疗效评价等过程起着重要的指导作用。目前临床中上使用的胸部医学成像方式主要有胸部 X 射线 (CXR)、计算机断层扫描 (CT)、核磁共振(MRI)。其中核磁共振的价格高昂, X 射线成像的 X 胸片无法准确区分软组织,检查早期病变的患者漏诊率较高,一般以价格相对便宜且成像质量较好的 CT 为主要的新冠肺炎检测和辅助诊断手段。

但是,人工分析 CT 影像往往对专业知识的要求较高,且对影像特征的分析较为费时费力。我国医疗资源分别失衡较为严重,部分地区存在医生专业技术和经验不足的情况,可能造成新冠肺炎的漏诊和误诊。尤其是在疫情爆发期间,病患数量迅速增长的情况下医疗资源十分紧缺,仅靠人工分析往往不足以应对疫情。近年来,人工智能(Artificial Intelligence)技术^[31]的快速发展对医疗影像分析带来了巨大变革,利用深度学习(Deep learning)算法^[32]自动学习图像数据,提取各层次抽象特征来分析医学影像,并给出辅助诊断结论已成为现代临床影像分析工作中的重要发展趋势。大量研究表明,利用深度学习对 CT 影像进行筛选和分类的辅助诊断技术能为临床医生提供有价值的参考意见^[33,34]。利用深度学习算法对 CT 影像进行新冠肺炎辅助诊断,能在一定程度上避免新冠肺炎其他检测方式的漏诊和误诊,同时也能帮助一些缺乏专业技术和经验的医生做出更准确的诊断。

在训练样本较少的情况下,使用深度学习往往容易出现过拟合的情况。使用迁移学习能较好的解决这一问题。

1.4 本文主要工作

本文分析了新冠肺炎图像识别在国内外研究现状,本课题旨在探究不同卷积神经网络和迁移学习方法对新冠肺炎 CT 图像的分类效果,并提出如何提升基于卷积神经网络和迁移学习方法在新冠肺炎 CT 图像上的分类性能。用来找到相对较佳的分类方法对新冠肺炎进行辅助诊断。本文的主要研究工作可以分为以下三个方面:

1. 本文探索了基于 VGGNet、GoogLeNet、ResNet、DenseNet 等经典卷积神经网络模型对新冠肺炎 CT 图像的分类效果。并在此基础上,通过改进分类层结构、嵌入 SE 模块、嵌入 CBAM 模块等方式来提升网络的分类性能。通过对照实验,对比原模型和改进后模型的分类效果。实验结果证明改进分类层结构、嵌入 SE 模块、嵌入 CBAM 模块等方式都能在一定程度上提升模型的分类准确率;

2. 本文探索了基于卷积神经网络的不同迁移学习方式对新冠肺炎 CT 图像的分类效果,通过对照实验证明基于微调方式的迁移学习方法可以有效的提高模型在新冠肺炎 CT 图像分类的准确率。通过对不同模型进行集成,达到了本文的最佳分类效果,分类准确率为 98.59%。并且在数据集较小的情况下,本文设计了微调与特征提取相结合的迁移学习方式,进一步提高迁移学习模型的分类性能。

3. 本文用 Flask Web 开发框架搭建了新冠肺炎检测系统,包含图片上传和 CT 图像检测等功能,通过加载已训练好的模型来预测用户上传的图像,为用户提供了简便快捷的新冠肺炎 CT 图像检测方式。

1.5 论文结构

本篇论文总共分为六章,章节结构与安排如下:

第一章,绪论。本章首先介绍新冠肺炎检测的研究背景,总结了新冠肺炎检测的国内外研究现状,并介绍了本文的研究意义与主要工作。

第二章,相关技术与理论。本章对卷积神经网络的基本结构进行介绍,并总结了几种卷积神经网络的经典架构,最后介绍了迁移学习以及迁移学习不同的分类方式。

第三章,基于卷积神经网络新冠肺炎 CT 图像分类。探索了基于 VGGNet、GoogLeNet、ResNet、DenseNet 等经典卷积神经网络模型对新冠肺炎 CT 图像的分类效果。并在此基础上,通过改进分类层结构、嵌入 SE 模块、嵌入 CBAM 模块等方式来提升网络的分类性能。

第四章,基于迁移学习的新冠肺炎 CT 图像分类。研究了基于微调方式的迁移学习对新冠肺炎 CT 图像的分类。通过对不同模型的集成进一步提高了模型的分类效果。对于图像数量较少的小数据集,提出通过微调与特征提取相结合的迁

移学习方法，来进一步提高迁移学习模型的分类性能。最后，通过基于梯度的类激活映射可视化方法，得到对模型分类过程中关注的特征区域进行标记的热力图，对模型的分类结果做出一定的解释。

第五章，新冠肺炎检测系统的实现。搭建了基于 **Flask Web** 开发框的新冠肺炎检测系统，首先进行系统结构设计，分别搭建图像上传和图像检测模块，完成系统功能的实现。

第六章，总结与展望。

第二章 相关技术与理论

卷积神经网络在图像识别、语音识别、自然语言处理等许多领域都展现了强大的能力。卷积神经网络的三大核心特征是局部感知、权值共享和池化，对应了卷积神经网络中的卷积层和池化层。这些特征不仅缩小了网络的参数量，还使网络拥有了很强的特征提取能力，促进了卷积神经网络的发展和应用。本章主要介绍了 CNN 的基本组成部分以及几种经典的卷积神经网络模型，最后还介绍了迁移学习和迁移学习按不同方式的分类。

2.1 卷积神经网络结构

上世纪 60 年代,生物学家 Hubel 等人在对猫的视觉皮层细胞的研究过程中,最早提出了感受野的概念^[35],到 80 年代, Fukushima 在感受野概念的基础上提出了神经认知机的概念^[36],可以看作是卷积神经网络的首次实现。1986 年 Rumelhart 等人在人工神经网络的基础上提出了反向传播算法^[37],引起了神经网络在机器学习中的应用热潮。1998 年纽约大学的 YannLecun 设计提出了第一个卷积神经网络 (Convolutional Neural Networks) 网络结构 LeNet-5^[38],并在手写数字识别任务上取得了良好的效果。随后 AlexNet、VGGNet、GoogLeNet、Resnet、DenseNet 等神经网络结构相继被提出。卷积神经网络的结构比早期的人工神经网络有很大的改进,人工神经网络的结构只包括输入层、隐藏层和输出层,而卷积神经网络对隐藏层进行了细化,加入了卷积层、池化层和全连接层,大大增强了网络的学习能力。

近年来, CNN 在图像识别、目标检测、图像分割、医学图像分析等诸多计算机视觉与生物识别任务上取得了突破性的进展。卷积神经网络的层数在不断增加,网络结构也是越来越复杂。

2.1.1 卷积层

卷积层的作用是对输入数据进行特征提取,输入为图像时,经过卷积层后会得到对应的特征图。卷积层内部一般包含多个卷积核,卷积核的大小有 1×1 , 3×3 , 5×5 和 7×7 等。严格意义上说,卷积层上进行的并不是卷积运算,而是互相关运算。在计算卷积的过程中需要做卷积核翻转,为了减少开销,我们会省略卷积核翻转的过程,用互相关运算来代替卷积运算。

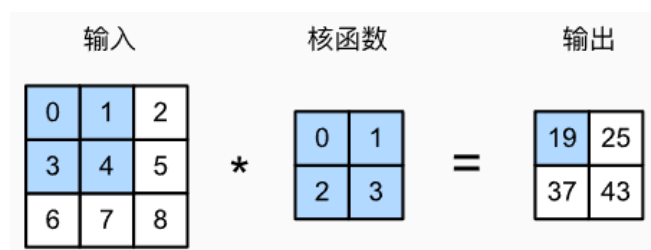


图 2-1 互相关运算示例图

互相关运算从图 2-1 中的输入张量的左上角开始计算，卷积窗口从左到右，从上到下移动。卷积窗口每移动到一个位置，将该窗口中的数值和卷积核中的数值按元素相乘并求和作为输出，例如 $0 \times 0 + 1 \times 1 + 3 \times 2 + 4 \times 3 = 19$ 作为卷积后对应位置的输出。卷积的输出形状取决于输入形状和卷积核的形状。假设输入形状为 $oh \times ow$ ，卷积核形状为 $ch \times cw$ ，那么输出形状是 $(oh - ch + 1) \times (ow - cw + 1)$ 。

此外卷积的输出形状还取决于填充和步长。填充是为了解决多层卷积时边缘像素丢失的问题。在输入图像的边界填充元素（通常填充元素是 0），可以分别设置图像高度和宽度的填充值。设置步长是为了减少采样次数来提高计算效率，卷积窗口可以直接跳过中间的多个元素，每次跳过元素的数量称为步长。步长也可以同时设置垂直步长和水平步长。如图 2-2 是以 2 行 2 列填充，并设置垂直步长为 3 水平步长为 2 的卷积运算结果。

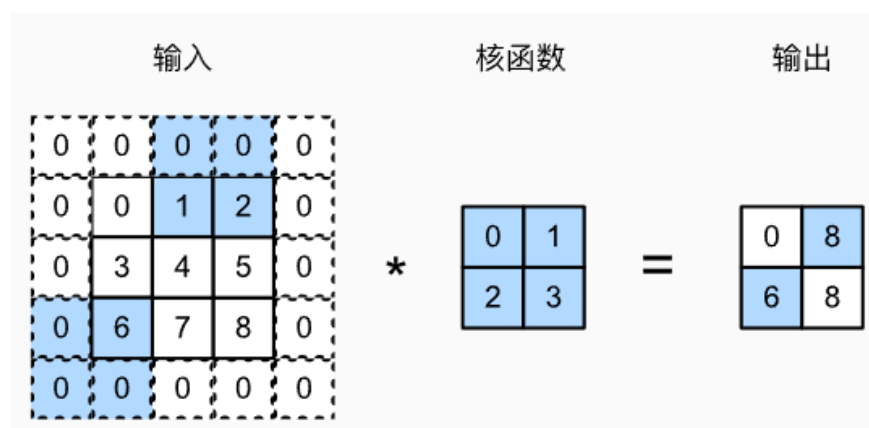


图 2-2 卷积填充和按步长滑动卷积运算示例图

2.1.2 池化层

图像中相邻像素的值相似的可能性非常高，因此卷积层虽然可以减少特征图维度，但是卷积层的输出中依然会包含大部分的冗余信息。这可能会因为计算量太大，显存占用超出 gpu 显存上限，而导致出现模型停止训练，而且还容易出现过拟合的现象。为了进一步压缩数据，在卷积层对特征进行提取后，输出的结果

需要通过池化层,对特征进行选择并适当的过滤信息。池化操作的本质是下采样,池化层对于输入的特征图像选择某种方式对其进行压缩,从而减少参数量和计算量并降低内存使用量。因此池化层能在一定程度上解决过拟合问题。池化层的操作主要有两种,一种是计算池化窗口内元素的平均值,另一种是直接取池化窗口内元素的最大值,分别叫做平均池化和最大池化。最大池化容易获得与相邻像素差距较大的值,能够保留图像的纹理特征,而平均池化能够保留图像整体的数据特征。最大池化和平均池化运算过程如图 2-3 所示,其中池化步长为 2。和卷积层类似,池化层输出的特征图像也由池化大小、步长和填充控制。

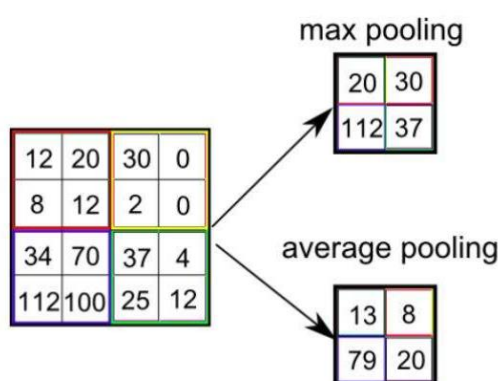


图 2-3 对 4×4 矩阵做步长为 2 的最大池化和平均池化运算示例

2.1.3 激活函数

神经网络的计算一般是线性的,如果没有激活函数,无论神经网络有多少层,输出都是输入的线性组合,就像最原始的神经感知机一样,网络的逼近能力非常有限,拟合能力很差,无法拟合任何复杂的函数。因此激活函数在神经网络中十分重要,只有引入激活函数,神经网络才能把线性运算映射为非线性分布的数据,深层神经网络表达能力就更加强大。这样,神经网络拟合的函数也能适应复杂函数的波动,拟合效果更好。常用的激活函数包括 Sigmoid、Tanh、ReLU 等,接下来将对它们进行介绍。

(1) Sigmoid 函数

Sigmoid 函数可以将输入的连续实值转换为 0 到 1 之间的输出,并且能把大的负数转换成 0,把大的正数转换成 1。Sigmoid 函数的表达式如公式 2-1 所示。

$$f(x) = \frac{1}{1+e^{-x}} \quad (2-1)$$

Sigmoid 函数图形如图 2-4 所示：

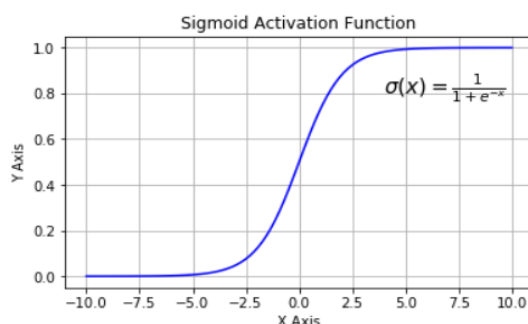


图 2-4 Sigmoid 函数曲线图

从上图可以看出当输入 x 接近 ± 5 ，输出已经接近 0 或 1 了，而输出的映射区间在 $(0,1)$ 内，当输入继续远离原点时导数趋于 0。在反向传播的过程中，这些神经元的权值就无法更新，同时影响相邻神经元的权值更新减慢，会造成梯度消失的问题。此外 Sigmoid 激活包含幂运算，计算量较大。Sigmoid 的输出不是 0 均值，随着网络的加深，会改变数据的原始分布。

(2) Tanh 函数

Tanh 函数又叫作双曲正切激活函数，Tanh 函数的表达式如公式 2-2 所示。对比 Sigmoid 函数可以发现两者十分相似，区别在于 Tanh 函数的输出值范围为 $(-1,1)$ ，Sigmoid 函数的输出值范围为 $(0,1)$ ，我们可以把 Tanh 函数近似看做是 sigmoid 平移和拉伸的结果。Tanh 函数不存在 Sigmoid 函数的不是 0 均值输出问题。此外，Tanh 函数计算量较大仍然较大，相对于 Sigmoid 函数，梯度消失问题得到了改善但并未彻底解决。

$$\tanh(x) = \frac{2}{1 + e^{-2x}} - 1 \quad (2-2)$$

Tanh 函数图形如图 2-5 所示：

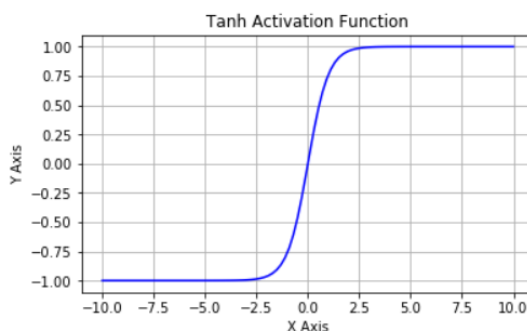


图 2-5 Tanh 函数曲线图

(3) ReLU 函数

ReLU 函数的表达式如公式 2-3 所示。

$$f(x) = \max(0, x) \quad (2-3)$$

ReLU 函数其实就是一个取最大值函数。ReLU 在 $x > 0$ 时的导数为常数 1，所以不存在梯度饱和的问题。由于 ReLU 函数只有线性关系，其计算速度和收敛速度都远快于 Sigmoid 和 Tanh 函数。但 ReLU 函数同样也存在问题，前向传导过程中如果输入的是负数，ReLU 是完全不被激活的，神经元也会保持非激活状态，且在后向传导过程中梯度就会完全变成 0，这样权重无法得到更新。针对这些问题，ReLU 函数还被提出了一些改进型，如 ELU、LeakyReLU，然而，并没有充分证明它们在实际应用中总是优于 ReLU。ReLU 函数同样不是以 0 为中心的函数。尽管存在这些问题，ReLU 函数仍然是最常用的激活函数。

ReLU 函数图形如图 2-6 所示：

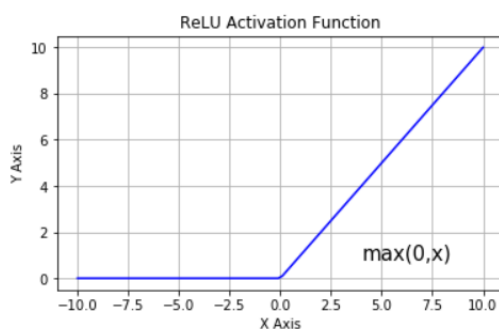


图 2-6 ReLU 函数曲线图

2.1.4 全连接层

全连接层上一层的每个神经元都与下一层的所有神经元相连,与卷积层的局部连接不同。全连接层位于卷积神经网络输出结果之前的部分,输入的数据在经过全连接层之前的卷积层、池化层和激活函数层等操作后,原数据的特征将会被提取出来。而全连接层的作用就是把前面提取的特征做加权和,最后通过输出层转化为计算每个类别的概率,在卷积神经网络的整体结构在中相当于起分类器的作用。

2.1.5 输出层

输出层在卷积神经网络中位于最后一层，负责输出最后结果。对于图像分类问题，输出层的输出结果为目标任务的分类标签，一般采用 softmax 函数作为类别预测输出，softmax 的输出表征不同类别之间的相对概率。二分类问题也可以用公式 2-1 中的 sigmoid 函数预测输出，两者在二分类问题中可以看作等价。对

于输入 x_i , softmax 函数的表达式如公式 2-4 所示。

$$\text{softmax}(x_i) = \frac{\exp(x_i)}{\sum_{k=1}^n \exp(x_k)} \quad (2-4)$$

从公式 2-4 中可以理解 softmax 如何将分类输出转换为各个类别的预测概率，具体可以分为两步。首先分子通过指数函数将输出映射到零到 $(0, +\infty)$ ，以确保概率的非负性，然后分母将所有映射结果相加进行归一化。因此每个预测结果的概率之和等于 1，模型预测分类的结果取其中概率值最高的类别。

2.1.6 反向传播

反向传播算法，是训练卷积神经网络最常用且最有效的算法。反向传播相对于正向传播而言，正向传播即卷积神经网络输入某个向量，这个输入向量通过每一层隐含层，直到输出层，然后获得网络的一个输出。反向传播则计算输出结果与实际值之间的误差，并将该误差从输出层向隐藏层反向传播，直至传播到输入层。在神经网络初始化的时候，每个神经元的权重和偏置值都是随机赋值的。反向传播算法通过计算输出层结果与真实值之间的偏差来进行逐层调节参数，对网络中所有权重计算损失函数的梯度，这个梯度会反馈给最优化方法（如梯度下降法），用来更新权值并损失函数值最小。

损失函数是用来计算神经网络输出值与实际值之间的误差的。常用的损失函数有均方差损失函数和交叉熵损失函数。均方差损失函数的表达式如公式 2-5 所示。

$$C = \frac{1}{2n} \sum_x \|y(x) - a^L(x)\|^2 \quad (2-5)$$

交叉熵损失函数的表达式如公式 2-6 所示。

$$C = -\frac{1}{n} \sum_x [y \ln a + (1 - y) \ln(1 - a)] \quad (2-6)$$

在池化窗口内的元素知道了神经网络输出值和实际值之间的误差后，反向传播调整各层神经元的权值和偏置值方式如下。下图 2-7 是一个简单的神经网络，layer1 到 layer3 依次是输入层、隐藏层和输出层。

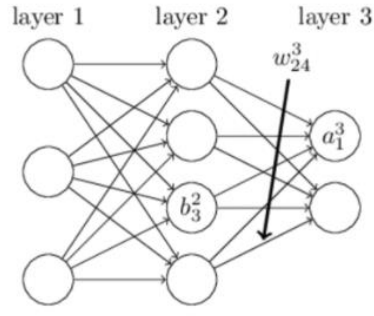


图 2-7 简单神经网络示意图

我们先进行参数的定义。 w_{jk}^l 表示第 $(l-1)$ 层的第 k 个神经元连接到第 l 层的第 j 个神经元的权重， b_j^l 表示第 l 层第 j 个神经元的偏置， z_j^l 表示第 l 层第 j 个神经元的输入， a_j^l 表示第 l 层第 j 个神经元的输出。由定义可得如下，其中 σ 表示激活函数。

$$z_j^l = \sum_k w_{jk}^l a_k^{l-1} + b_j^l \quad (2-7)$$

$$a_j^l = \sigma(z_j^l) \quad (2-8)$$

首先将第 l 层第 j 个神经元中产生的误差定义为：

$$\delta_j^l = \frac{\partial C}{\partial z_j^l} \quad (2-9)$$

则输出层的第 j 个神经元中产生的误差可以表示为：

$$\delta_j^L = \frac{\partial C}{\partial a_j^L} \cdot \frac{\partial a_j^L}{\partial z_j^L} \quad (2-10)$$

由此可得输出层的误差方程如公式 2-11 所示。其中 \odot 表示 Hadamard 乘积。

$$\delta^L = \nabla_a C \odot \sigma'(z^L) \quad (2-11)$$

在神经网络中下一层神经元的输入是上一层神经元输出的线性组合，因此可以由下一层神经元的误差来表示上一层神经元的误差。第 l 层的误差可以由 $l+1$ 层反馈过来，这就是反向传播的过程。从后向前计算每一层的误差方程如公式 2-12 所示。

$$\delta^l = ((w^{l+1})^T \delta^{l+1}) \odot \sigma'(z^l) \quad (2-12)$$

使用梯度下降更新每一层的权值 w 和偏置值 b 的表达式如公式 2-13 和公式 2-14 所示。

$$w^l \rightarrow w^l - \frac{\eta}{m} \sum_x \delta^{x,l} (a^{x,l-1})^T \quad (2-13)$$

$$b^l \rightarrow b^l - \frac{\eta}{m} \sum_x \delta^{x,l} \quad (2-14)$$

2.2 卷积神经网络经典架构

卷积神经网络可以自动学习样本数据的特征表示，提取样本的特征，不再需要人工提取特征。卷积神经网络早期的研究发现，卷积神经网络的结构越复杂，越能够更好地拟合非线性数据，通过增加网络的深度可以提高模型的准确率。但是之后的研究发现，到达一定程度后继续加深卷积神经网络，准确率反而下降了。其原因是模型越大，网络参数就越大，就越容易产生过拟合现象，还会出现梯度消失的问题。这种情况下往往需要用更大的数据集来进行训练，然而构建大型数据集需要花费大量的人力物力，实际研究中采取这一方法很不现实。同时模型越来越大，往往容易超出已有计算机的计算性能。因此许多研究者尝试改进卷积神经网络模型，试图在不断增加网络深度的同时减少网络参数。本节将详细介绍 AlexNet、VGGNet、GoogleNet、ResNet、DenseNet 等五种经典的卷积神经网络模型。

2.2.1 AlexNet

LeNet5 是第一个 CNN 网络结构，但是 AlexNet 却是第一个引起大家注意的网络。Alex 等人提出的 AlexNet 网络在 2010 年 ImageNet 比赛上获得冠军，且成绩远远超过第二名^[39]，因此卷积神经网络甚至深度学习再次引起了广泛关注。AlexNet 首次在卷积神经网络中使用 Dropout 机制^[40]及局部响应归一化技术，将 ReLU 作为激活函数，并且在数据量有限的情况下通过数据增强扩充训练数据，这些方法和技术推动了卷积神经网络在图像处理领域的迅速发展。AlexNet 主要由 8 层组成，包括五个卷积层、三个全连接层，其中最后一个全连接层为输出层。AlexNet 的网络结构图如图 2-8 所示。

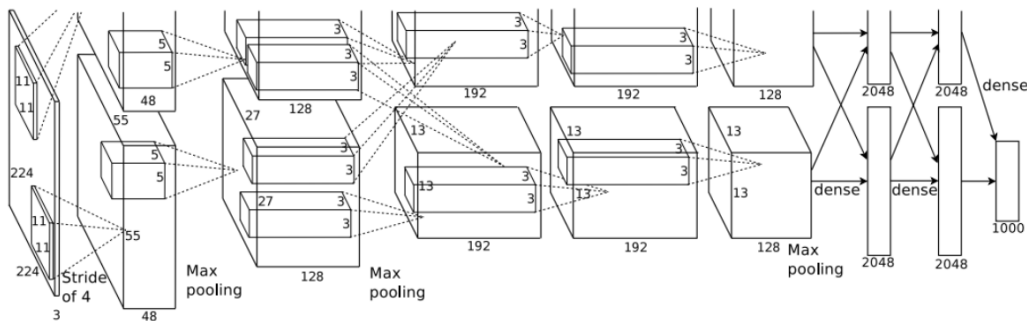


图 2-8 AlexNet 网络结构图^[39]

AlexNet 使用两块 GPU 进行计算，在当时 GPU 显存有限的情况下很大程度上提高了运算效率。这在现在来说并没有必要，用 AlexNet 训练模型一般都一块 GPU 进行计算，如图 2-9 所示。

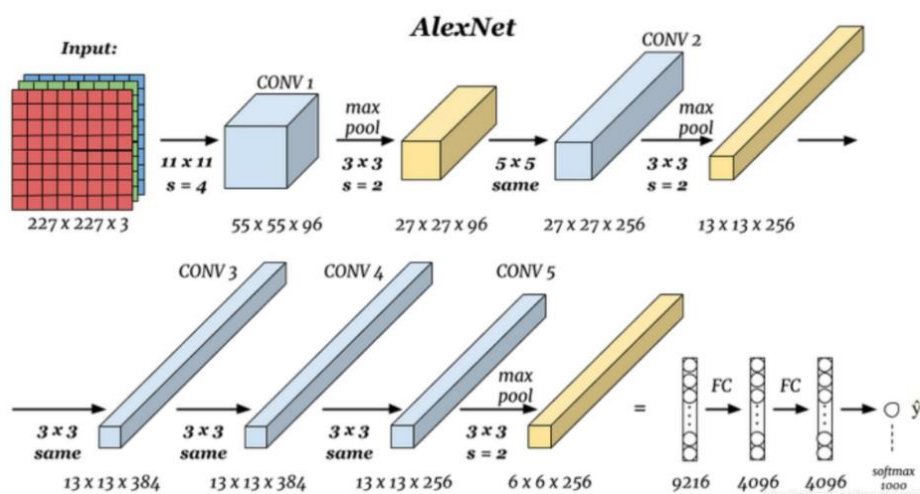


图 2-9 AlexNet 单 GPU 网络结构图

2.2.2 VGGNet

VGGNet 是由 Simonyan 和 Zisserman^[41]开发的，在 2014 年 ImageNet 大赛上取得了亚军。VGGNet 与 AlexNet 类似，都是由卷积层、全连接层两大部分构成。相对于 AlexNet，VGGNet 主要是从深度方面对卷积神经网络做改进，想要探究卷积神经网络的深度越深其效果是否越好。此外，VGGNet 整个网络都使用了更小的 3×3 卷积核和 2×2 最大池化尺寸，使网络更加简洁。卷积核的尺寸虽然减小了，但却使用了更多的卷积层数。两个 3×3 卷积核的堆叠相对于 5×5 卷积核的视野，三个 3×3 卷积核的堆叠相当于 7×7 卷积核的视野。这样做的好处是可以有效减少网络参数，而且相当于进行了更多的非线性映射，增加了卷积神经网络对复杂函数的拟合能力和对特征的学习能力。VGGNet 网络共有 5 个卷积模块，每个卷积模块后面跟一个池化层，最后是 3 个全连接层，可以通过调整卷积模块中的卷积层数量和卷积核大小来实现不同深度的网络，深度包括 11 层、13 层、16 层和 19 层。

2.2.3 GoogLeNet

GoogLeNet 是 google 推出的基于 Inception 模块的深度神经网络模型，InceptionV1 在 2014 年的 ImageNet 竞赛中摘得桂冠，并在后来被不断改进，相继形成了 InceptionV2、InceptionV3、InceptionV4 等版本。

Inception 模块的思想就是把多个卷积或池化操作,放在一起组装成一个网络模块,再用这些网络模块去组装整个网络结构。这种模块化的思想就好像搭积木一样。如图 2-10 是 Inception 模块的最初版本,由 1×1 卷积, 3×3 卷积, 5×5 卷积, 3×3 最大池化 4 个结构组成。使用不同大小的卷积核表示获得不同大小的感受野,最后将不同大小的卷积拼接,让一个模块可以得到多个尺度的信息。然而由于最后合并的方法是将各个部分输出的特征图相加,导致卷积层中数据量的和计算量很大。于是作者借鉴了 NIN (Network in Network) [42], 采用 1×1 卷积核来进行降维,产生了 InceptionV1 模块,有效地降低了计算量,如图 2-11 所示。

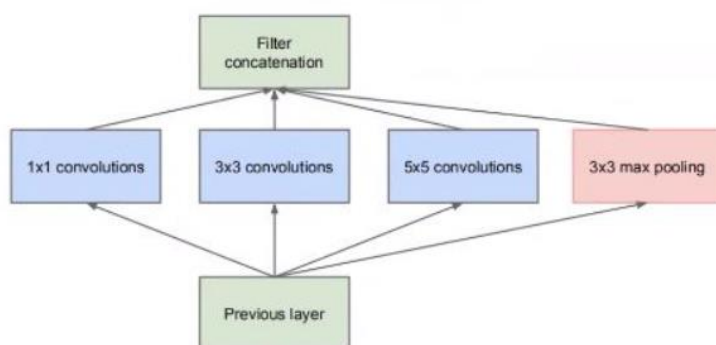


图 2-10 Inception 模块最初结构图

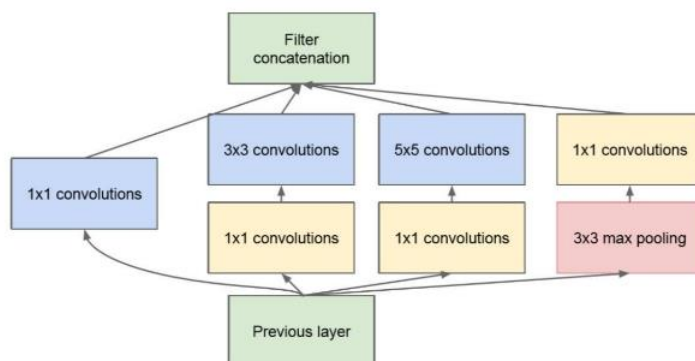


图 2-11 InceptionV1 模块结构图

GoogLeNet 通过 Inception 方法,丢弃全连接层,在增加网络深度和宽度的同时减少了网络的参数。

2.2.4 ResNet

ResNet 由 KaimingHe 等[43]四名华人提出,通过使用残差结构成功训练出了 152 层的神经网络,在 2015 年 ImageNet 竞赛中的分类任务和目标检测都获得第一名。

VGGNet 曾试图探索卷积神经网络的深度与模型性能之间的关系，然而 VGGNet 的网络深度也十分有限，最深的网络模型也只有 19 层。网络越深，能获取的特征信息也越丰富。但是实验表明，随着网络的加深，性能会变得停滞，甚至训练数据和测试数据的准确率反而都降低了。这是因为在网络的深度到达一定程度后，继续加深网络会造成梯度爆炸和梯度消失的问题。ResNet 的残差单元在一定程度上解决了这一问题^[44]。

ResNet 网络在 VGG19 网络结构的基础上进行了修改，通过加入残差单元实现了短路机制。通过对比 VGG19 和 ResNet34 可以发现，ResNet 除了第一层卷积层使用了 7×7 卷积核，其余所有卷积层都保持 3×3 卷积核，ResNet 最大的区别在于每两层间增加了短路机制，形成残差单元。

ResNet 使用两种残差单元，一种是以两个 3×3 的卷积网络串接在一起作为一个残差单元，另外一种是由 1×1 、 3×3 、 1×1 的 3 个卷积网络串接在一起作为一个残差单元，如图 2-12 所示。

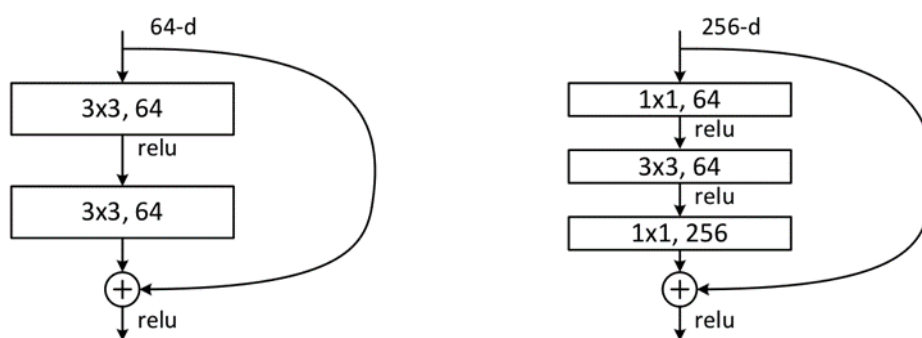


图 2-12 两种残差单元

不同残差单元的堆叠可以构成不同网络层数的 ResNet，如图 2-13 所示。例如 ResNet50 是从一层 7×7 的卷积层开始，经过一个 3×3 的最大池化层，之后分别由 3、4、6、3 个三层的残差单元组合而成，再经过一个全局平均池化层和全连接层后输出分类结果。

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	$7 \times 7, 64, \text{stride } 2$				
		$3 \times 3 \text{ max pool, stride } 2$				
conv2_x	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

图 2-13 ResNet 的不同结构

2.2.5 DenseNet

DenseNet^[45]是 Gao Huang 等人在 2017 年的 CVPR 会议上提出的。DenseNet 提出了密集连接机制，每个层都会与前面所有层在通道维度上连接在一起。DenseNet 的网络结构主要由稠密块(Dense Block)和过渡层(Transition Layer)组成。稠密块是 DenseNet 最突出的设计，与残差块中的相加跳转连接不同，稠密块是直接进行连接，这一设计使得稠密块中每个卷积层输出的特征图都很小。因此 DenseNet 相对于其它卷积神经网络更窄，参数也更少。稠密块的结构如图 2-14 所示。

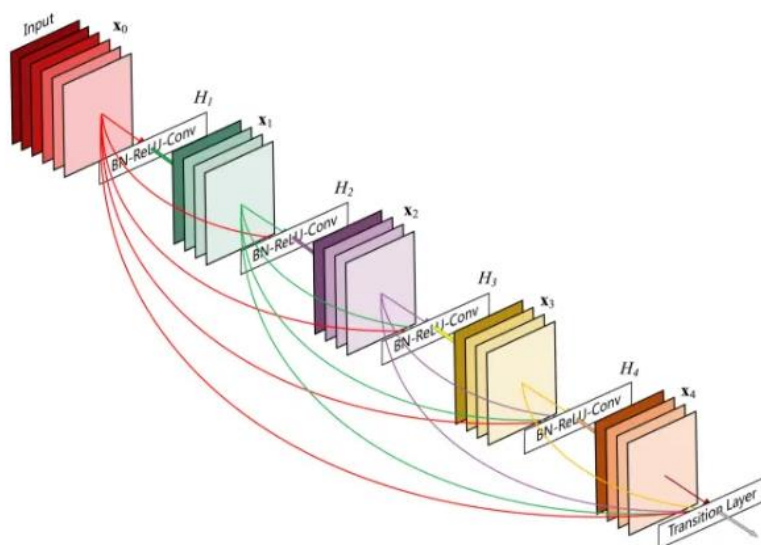


图 2-14 稠密块网络结构图^[45]

DenseNet 的优点主要有：更好地解决了梯度消失问题、网络的参数更少、有效地进行了特征的重用、网络自带一定的正则化能力、易于训练等。DenseNet 通过对特征的极致利用，使其在网络参数更少的同时获得了更好的效果。

2.3 迁移学习

2.3.1 迁移学习介绍

迁移学习 (Transfer Learning)^[46]最早是一种机器学习方法，随着卷积神经网络的发展发现其浅层网络的特征具有通用性，可以将这些特征从其他已经训练好的网络中迁移过来，迁移学习开始逐渐应用于深度学习中。迁移学习是通过从一个相关领域转移信息来提高在另一个领域的学习能力。先前建立的模型可以通过迁移学习适应其他问题。比如说，一个人可以根据自己骑自行车的经验来学习如

何驾驶摩托车。这并不意味着一个人不会骑自行车就不能学习如何骑摩托车，而是说借鉴自己骑自行车的经验可以加快学骑摩托车的效率。

2.3.2 迁移学习的分类

在迁移学习中，有两个基本的概念：领域(Domain)和任务(Task)^[47]。

(1)领域：是进行学习的主体。领域主要由两部分构成：数据和生成这些数据的概率分布。领域在迁移过程中分为源领域和目标领域。源领域就是有知识、有大量数据标注的领域，是我们要迁移的对象；目标领域就是我们最终要赋予知识、赋予标注的对象。

(2)任务：是学习的目标。任务主要由两部分组成：标签和标签对应的函数。

根据源领域和目标领域及任务之间的不同情况，可以将迁移学习分为三类：归纳式迁移学习、直推式迁移学习、无监督迁移学习^[47]。

表 2-1 迁移学习按源领域和目标领域及任务之间的不同情况分类

迁移学习种类	源领域和目标领域	源任务和目标任务
归纳迁移学习	相同	不同但相关
无监督迁移学习	不同但相关	不同但相关
直推式迁移学习	不同但相关	相同

迁移学习按学习方式可以分为四类：基于样本的迁移学习、基于特征的迁移学习、基于模型的迁移学习、基于关系的迁移学习^[47]。

基于样本的迁移：重新对源域中的一些数据进行重加权，以便在目标域中使用。

基于特征的迁移：既基于特征映射的迁移学习方法，通过对特征进行变换，将源域和目标域的特征映射到同一个空间里面。

基于模型的迁移：既共享模型的参数，将源域中已经训练好的模型全部或者部分迁移到目标域当中。比如从像 ImageNet 这样的大型图像集预训练的模型可以高度迁移到各种图像识别任务中。这是目前使用最为广泛的一种迁移学习方式。由于不需要针对新任务从头开始训练网络，可以节省大量的训练时间。

基于关系的迁移：指当源域和目标域存在某种相似的关系时，可以利用关系进行类比迁移。比如将师生关系迁移到上司下属的关系。目前利用这种迁移学习方法的研究比较少。

基于卷积神经网络使用迁移学习中最常用的是基于模型的迁移和基于特征的迁移，分别对应两种常见的迁移学习场景：微调(fine-tune)卷积网络 and 把卷积神经网络作为特征提取器。第一种微调卷积神经网络，保持模型原有的结构和初始

权重不变，用自己需要学习的数据来微调预训练好的网络模型的参数。在微调的时候可以选择冻结前面几层的网络权重保持不变，仅对后面几层网络进行微调训练，也可以选择对全部层采用微调的方式进行训练。第二种把卷积神经网络作为特征提取器，是把预训练好的卷积神经网络模型提取最后一层卷积层的输出或者提取除去最后一个全连接输出层后的倒数第二层的输出作为卷积神经网络的特征，然后按自己的需求构建分类器或者直接使用 SVM、贝叶斯或 softmax 等分类器来完成新的任务。此时将之前预训练好的每一层网络参数冻结，仅对最后增加的分类器进行训练。

2.4 本章小结

本章首先介绍了卷积神经网络结构和网络中每个组成模块的作用，然后对 AlexNet、VGGNet、GoogLeNet、ResNet、DenseNet 等经典网络模型的结构及其设计理念进行了较为细致的讲解。最后对迁移学习的概念进行介绍，归纳了迁移学习按不同方式的分类，并介绍了基于卷积神经网络使用迁移学习通常用到的两种方法。

第三章 基于卷积神经网络的新冠肺炎 CT 图像分类

本章将用各种卷积神经网络对新冠肺炎 CT 图像进行分类，介绍了分类模型所使用的数据集、数据集的预处理方法、实验环境等，在 VGG19、InceptionV3、ResNet50、DenseNet169 等网络模型的基础上进行改进，通过修改分类层结构提高分类性能。之后在改进后的 ResNet50 的基础上通过加入 SE 模块和 CBAM 模块进一步提升模型的分类性能。

3.1 数据集

本实验使用的数据集为 SARS-CoV-2 CT^[48]数据集，数据集共有 2481 张肺部 CT 图像，其中包括 1252 张感染了 SARS-CoV-2 的 CT 扫描，以及 1229 张没有感染 SARS-CoV-2 的 CT 扫描。这些图像数据来自巴西圣保罗医院的真实患者。该数据集的创建者使用了一种可解释的深度学习方法（xDNN），并获得 97.31% 的 F1 分数，作为该数据集的基线结果。

将数据集划分为训练集、验证集、测试集，先从数据集中随机取出感染新冠肺炎的 CT 图像 249 张和未感染新冠肺炎的 CT 图像 248 张作为测试集，将剩下的图像数据按 3:1 分为训练集和验证集，其中训练集 1488 张，验证集 496 张。训练集、验证集、测试集的比例为 6:2:2。

3.2 实验环境

本实验用的框架是 Keras 和 Tensorflow。为加快网络模型的训练速度，本实验在 GPU 环境下进行模型的训练。实验环境配置详细信息如表 3-1 所示。

表 3-1 实验环境配置

名称	型号/参数
CPU	Intel i7-6850K
GPU	Nvidia Geforce GTX1080
硬盘	2TB
系统	Ubuntu16.04
内存	32G
JupyterNotebook 版本	1.0.0

表 3-1 实验环境配置（续）

名称	型号/参数
Python 版本	3.7.13
Keras 版本	2.8.0
Tensorflow 版本	2.8.0

Keras 是由纯 python 编写的基于 theano/Tensorflow 的深度学习框架。Keras 是一个高度模块化的神经网络库，支持快速实验，能够把你的想法迅速转换为结果。Keras 能提供很多经典的模型，如 VGG19/ResNet/InceptionV3/Xception。

TensorFlow^[49]是最受欢迎的深度学习框架之一，本文使用的是 2.8.0 版本。TensorFlow 提供了所有开源的深度学习框架里，最全的算法库，并且在不断的添加新的算法库。在谷歌的支持下，TensorFlow 周围形成了一个强大的社区，TensorFlow 的效率不断提高。TensorFlow 可以同时多个 CPU，GPU 或者两者混合运行。但是 TensorFlow 在提供经典图像识别模型上，却做得不好，且 fine-tuning 的流程相对 Keras 比较复杂。

目前 Keras 的 API 已经整合到 Tensorflow 中，在 Tensorflow 中通过 tf.keras 就可以调用 Keras。为了提高模型的效率，本文将两个框架联合起来使用。在训练时使用 GPU 提升网络的训练速度。

3.3 数据预处理

SARS-CoV-2 CT 数据集中的图像大小都不相同。在一个分类模型中，如果输入图像尺寸太小容易造成信息丢失，如果输入图像尺寸太大计算量就大，对于计算资源的需求以及模型训练的时间也越大。因此我们使用 OpenCV(cv2)库中的 resize 方法，将所有图像统一调整为 224×224 的尺寸。OpenCV 库是一个跨平台的计算机视觉库，可在 Linux、Windows、Android 和 Mac OS 等操作系统上运行，它拥有丰富的图像处理函数库，能够快速实现一些图像处理任务。OpenCV 提供了五种插值计算输出图像像素值的方法：最近邻插值法(INTER_NEAREST)、双线性插值法(INTER_LINEAR)、基于局部像素的重采样(INTER_AREA)、基于 4x4 像素邻域的 3 次插值法(INTER_CUBIC)，和基于 8x8 像素邻域的 Lanczos 插值法(INTER_LANCZOS4)。由于对于绝大多数的图像数据调整为 224x224 的尺寸都是缩放，这里我们选用 INTER_AREA 插值方法。之后对图像数据进行归一化处理，使得数据的分布满足高斯分布。

卷积神经网络有很强的特征表达能力，但是在训练数据量较少的情况下某些样本的抽样误差会被当作特征被过度学习，容易出现过拟合的情况，往往表现为

模型在训练集上的分类准确率较高，而在验证集上的分类准确率较低。通常用来解决过拟合问题的方法有 Dropout、正则化、数据增强等。Dropout 方法是通过在每一轮的训练过程中按照指定概率随机丢弃部分网络节点来实现的，正则化是在损失函数中加入惩罚系数，两者都是在训练模型的过程中防止过拟合的方法。数据增强则是通过图像变换的方式扩增图像数据，可以在无法获得更多标记数据的情况下，通过镜像、旋转、尺度变换、色彩抖动等方式，让有限的的数据产生更多的数据。数据增强的好处是增加了噪声数据提升网络模型的鲁棒性，同时随机改变训练样本可以降低模型对某些特征的依赖，能够提高模型的泛化能力。

本实验使用 Keras 提供的 ImageDataGenerator 类可以直接实现上述操作。ImageDataGenerator 是 keras.preprocessing.image 模块中的图片生成器，可以对每一个批次的样本数据进行增强，以扩大数据集的规模，增强模型的泛化能力。本实验采用的数据扩增方式是水平和垂直偏移图像、水平和垂直翻转图像、旋转图像这三种方式。

3.4 基于卷积神经网络改进的新冠肺炎 CT 图像分类

3.4.1 改进方法

VGGNet 及其之前的神经网络模型基本上都用全连接层加 softmax 层的结构作为分类层，全连接层相当于起到分类器的作用。全连接就是一种无损的卷积的过程，会大大占用计算机的计算力。模型里面通常有两层全连接层，因为一个全连接卷积层无法使特征数据聚集到一种可以准确判断的程度，再加一层使特征数据更加集中，例如 VGG16 使用了两层 4096 个节点的全连接层。虽然搭建多层全连接层有可能提高卷积神经网络模型分类准确率，但全连接层会带来大量的参数，会导致模型的计算量更大，训练时间更长，且更容易导致过拟合问题。

为了降低网络模型全连接层可能带来的问题影响，在 NIN (Network in Network) 论文中提出了使用全局平均池化层(GAP)代替全连接层的思路^[42]，其实实验结果证明用全局平均池化层代替全连接层不仅减少了大量的网络参数，还能在一定程度上解决过拟合的问题。在 GoogLeNet 的 Inception 系列网络和 DenseNet 网络中都使用全局平均池化层加 softmax 层的结构作为分类层。

然而用全局平均池化层代替全连接层的方法在一定程度上损失了全连接层提取特征值的能力，且用全局平均池化层后的参数较为固化，尤其不利于迁移学习。在一些实验过程中还发现全局平均池化层可能会造成收敛速度减慢。因此，本实验提出用全局平均池化层加一个节点数量较少的全连接层再加 softmax 层作为分类器。这样设计的好处是在减少大量训练参数的同时保留全连接层提取特征

值的能力。为了提高收敛速度，在全局平均池化层和全连接层后都加一层 Batch Normalization(BN)层，最后在两个 BN 层后面都加入 Dropout 层，防止过拟合问题的发生。

本实验中我们选取 VGG19、InceptionV3、ResNet50、DenseNet169 四种神经网络作为实验对象，对模型进行详细的分析。在搭建好四种神经网络结构后，保留除最后分类层之外的其余所有层，删去网络原来的分类层结构，再加入全局平均池化层+BN 层+Dropout 层+节点数较少的全连接层+BN 层+Dropout 层+2 节点的 softmax 层作为新的分类层。改进分类层后的网络模型结构如图 3-1 所示。

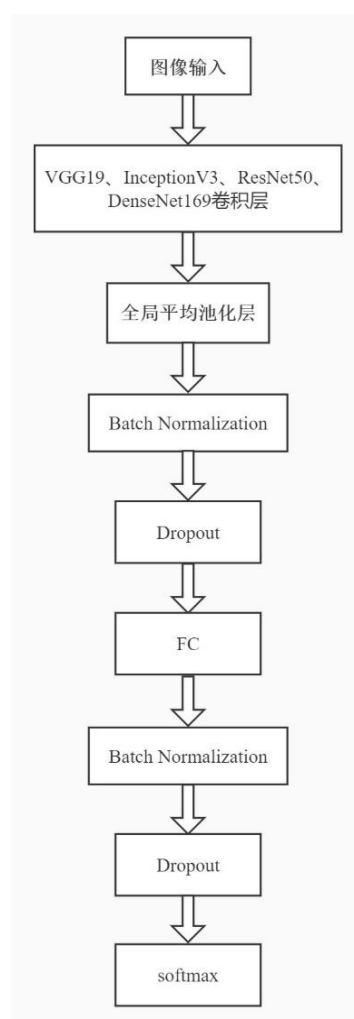


图 3-1 改进分类层后的网络模型结构

为确定全连接层节点数量为多少时模型的分类效果最佳，我们对 DenseNet 169 网络的分类层按上述方法进行改进后，设置了全连接层节点数量为 128、256、512 作为分组实验，分别进行训练。

3.4.2 模型训练

SARS-CoV-2 CT 数据预处理后，将改进后的网络在 SARS-CoV-2 CT 数据集上进行训练，设置 SARS-CoV-2 CT 阳性图像的标签为 0，SARS-CoV-2 CT 阴性图像的标签为 1。实验中损失函数选择交叉熵损失函数 `CrossEntropyLoss`，优化器使用 Adam 优化器并设置参数为：lr=0.003, betas=(0.9, 0.999)。同时在整个的训练过程中通过 `ReduceLROnPlateau` 动态调整学习率，并设置初始参数 factor=0.70, patience=5, verbose=1, min_lr=1e-4, 其意义为模型在每训练 5 个 epoch 后，如果验证集的损失函数值没有提升则将学习率改为原来的 70%，并设置最低学习率为 1e-4。训练的模型用 .h5 格式的文件保存。

为验证改进方法的有效性，我们用四种模型原来的网络结构在 SARS-CoV-2 CT 数据集上进行训练作为对照实验，实验参数设置与上面保持一致。

3.5 引入注意力机制的新冠肺炎 CT 图像分类

3.5.1 SE 模块

SE 模块即 SENet，SE 的全称为 Squeeze and Excitation，在 2017 年 ImageNet 竞赛中拿到了分类比赛的冠军^[50]。SE 模块结构清晰且易于实现，可以嵌入现有的网络模型框架中来提高模型的性能。SE 模块的提出主要是考虑到模型通道之间的相互依赖性。在传统的卷积池化过程中，默认特征图的每个通道是同等重要的，而在实际问题中，不同通道的重要性是有差异的。SE 模块能够通过自动学习的方式来获取不同通道的重要程度，重要程度越大表示该特征通道对当前任务越有效，然后依照这个重要程度去增加效果好的特征通道的权重，减小效果差的特征通道的权重。SE 模块并不是像卷积操作一样进行特征通道间的融合，而是采用了一种特征重标定的策略。SE 模块通过 Squeeze、Excitation 和 Scale 三个操作来实现上述过程，其中 Squeeze 和 Excitation 是最主要的操作。SE 模块的结构如图 3-2 所示。

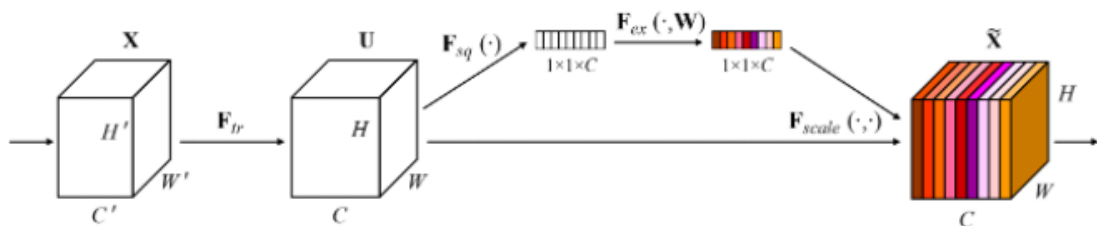


图 3-2 SE 模块结构^[50]

首先进行 F_{tr} 转换操作，把输入 X 映射到特征图 U ，其定义如公式 3-1 所示。

$$F_{tr}: X \rightarrow U, X \in R^{H' \times W' \times C'}, U \in R^{H \times W \times C} \quad (3-1)$$

F_{tr} 在 SE 模块其实就是一个卷积操作，设卷积核为 $V=[v_1, v_2, \dots, v_c]$ ， v_c 表示第 c 个卷积核的参数。则输出 $U=[u_1, u_2, \dots, u_c]$ 的表达式如公式 3-2 所示。其中 $*$ 表示卷积， v_c^s 表示一个 s 通道的二维卷积核， x^s 表示第 s 个输入。

$$u_c = v_c * X = \sum_{s=1}^{C'} v_c^s * x^s \quad (3-2)$$

Squeeze 操作将全局空间信息压缩到一个实数中，该实数相当于获得了全局感受野，其输出的维度等于特征通道数。原始特征图的维度为 $H \times W \times C$ ，其中 H 是高度(Height)， W 是宽度(width)， C 是通道数(channel)。Squeeze 其实就是把 $H \times W \times C$ 压缩为 $1 \times 1 \times C$ ，相当于把 $H \times W$ 压缩成一维了，这个过程一般是用全局平均池化实现的。将 $H \times W$ 压缩成一维的计算如公式 3-3 所示。

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (3-3)$$

Excitation 操作旨在完全捕获通道相关性，它是一个类似于循环神经网络中门的机制，通过参数 w 来为每个特征通道生成权重，其中参数 w 被学习用来显式地建模特征通道间的相关性^[50]。在得到 Squeeze 操作维度为 $1 \times 1 \times C$ 的结果 z 后，先进行全连接层操作，用 W_1 乘以 z ， W_1 的维度是 $C/r \times C$ ，其中 r 是一个缩放参数，在原文中取的是 16，设置该参数的目的是为了通过减少通道个数来降低计算量。 $W_1 z$ 的结果是 $1 \times 1 \times C/r$ ，然后再经过一个 ReLU 激活函数再和 W_2 相乘，这也是一个全连接层的过程， W_2 的维度是 $C \times C/r$ ，因此输出的维度就是 $1 \times 1 \times C$ ；最后再经过 sigmoid 函数，得到 s 。其具体计算如公式 3-4 所示。其中 σ 表示 sigmoid 激活函数， δ 表示 ReLU 激活函数。

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \quad (3-4)$$

最后是一个 Scale 操作，在 Excitation 操作后获得了输入特征图 U 的各个通道的权重，只需要将权重和原始的特征进行融合。 s_c 即是 Excitation 输出的特征权重。

$$\tilde{x}_c = F_{scale}(u_c, s_c) = s_c u_c \quad (3-5)$$

3.5.2 CBAM 模块

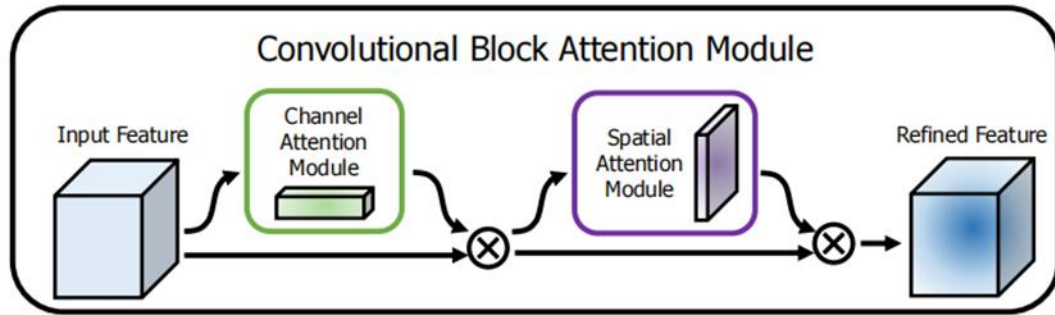


图 3-3 CBAM 模块示意图^[51]

CBAM 模块是在 SE 模块的基础上改进的，CBAM 模块的基本结果如图 3-3^[51]所示。前面是一个使用通道注意力机制的 SE 模块，后面是一个空间注意力模块。空间注意力模块的设计借鉴了 SE 模块，它通过在通道上使用全局平均池化获得了一个二维的空间注意力系数矩阵。CBAM 通过通道注意力和空间注意力的结合，使注意力机制的效果得到进一步加强。

3.5.3 ResNet 引入注意力机制的新冠肺炎 CT 图像分类方法

上文中介绍了 SE 模块的三个基本操作和其完成特征重标定的具体过程。SE 模块从特征通道的角度出发，通过学习不同通道的重要程度，将这些重要程度归一化并将权重重新分配到每个特征通道中，以此来完成特征的重标定。特征重标定后，对于任务中效果好的特征得到了加强，效果较差的特征则被抑制，从而增强了网络提取特征的能力。因此，本章节将 SE 模块嵌入到 ResNet 中，通过增强网络提取特征的能力来进一步提高 ResNet 网络的分类效果。

本章节提出了基于 SE-ResNet 改进的新冠肺炎 CT 图像分类方法。该分类方法是在 ResNet 网络的基础上，首先对网络的分类层结构进行了改进，再加入 SE 模块。先对 ResNet 网络进行了改进，改进方法和上文中的一样，把 ResNet 网络的最后几层分类层改为全局平均池化层+BN 层+全连接层+BN 层+softmax 层的结构。全连接层有 256 个神经元，softmax 层有 2 个神经元。

然后在 Resnet 网络的残差单元块中插入 SE 模块，构成 SE-Resnet 模块，SE-Resnet 模块如图 3-4 所示。先将经过残差单元块后每个通道上的特征图使用全局平均池化操作，将该特征图压缩成一个具有其全局感受野的实数，即 Squeeze 操作。再通过一个全连接层，一个 ReLU 激活函数和另一个全连接层，获得各个通道的重要程度，再通过 sigmoid 激活函数将所有通道的重要程度进行归一化，即

Excitation 操作。最后再将归一化后的各个通道重要程度的权重融合到每个通道上的特征图上，即 Scale 操作。这样就完成了原始特征的重标定。

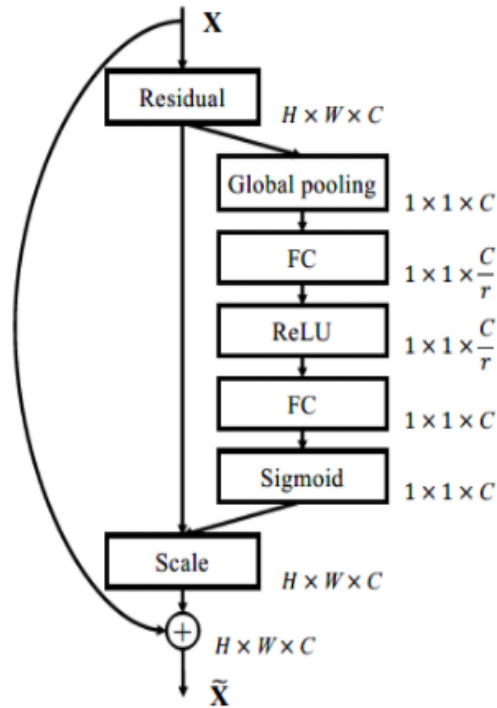


图 3-4 SE-Resnet 模块

我们也将 CBAM 模块加入到 ResNet 网络中。在分类层改进后的 ResNet50 网络中分别加入了两个 CBAM 模块，第一个 CBAM 模块位于 ResNet50 卷积层结构之前，并在 CBAM 模块与图像输入之间加入了一个 3×3 的卷积层，第二个 CBAM 模块位于 ResNet50 卷积层结构之后。CBAM-ResNet50 的具体结构如图 3-5 所示。

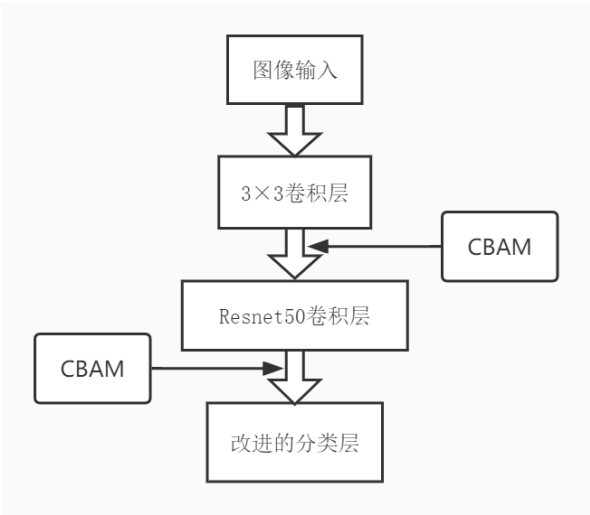


图 3-5 CBAM-ResNet50

3.5.4 模型训练

SARS-CoV-2 CT 数据预处理后,将改进后的 SE-Resnet50 和 CBAM-ResNet50 网络在 SARS-CoV-2 CT 数据集上进行训练,设置 SARS-CoV-2 CT 阳性图像的标签为 0, SARS-CoV-2 CT 阴性图像的标签为 1。实验中,损失函数选择交叉熵损失函数 CrossEntropyLoss, 优化器使用 Adam 优化器并设置参数为: lr=0.003, betas=(0.9, 0.999)。通过 ReduceLROnPlateau 来调整学习率, 并设置初始参数为 factor=0.70, patience=5, verbose=1, min_lr=1e-4。训练的模型用.h5 格式的文件保存。

3.6 实验结果和分析

3.6.1 评价指标

为准确衡量本文提出方法的有效性,本文采用常用的二分类评估指标。通过计算分类结果的准确率,查准率,召回率, F1 值, ROC/AUC^[52,53]等对实验结果进行评估。

我们借助二分类的混淆矩阵(如表 3-2 所示),对上述概念进行解释。

表 3-2 二分类混淆矩阵表

	预测正类 (P)	预测反类 (N)
实际正类 (T)	TP (真阳性)	TN (假阴性)
实际反类 (F)	FP (假阳性)	FN (真阴性)

准确率（Accuracy）表示正确预测样本的数量占总样本数的比率，其表达式如公式 3-6 所示。准确率越高体现出分类效果越好，但在正类反类样本数量不平衡的情况下准确率反映出的分类效果可能不够客观。

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+FP+TN} \quad (3-6)$$

查准率（Precision）表示正确预测的正类个数占所有预测为正类个数的比率，其表达式如公式 3-7 所示。

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3-7)$$

召回率（Recall）表示预测正确的正类个数占所有实际正类个数的比率，其表达式如公式 3-8 所示。召回率体现正类个数分类正确的比率，也称为查全率。

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3-8)$$

F1 值（F1score）是查准率和召回率的调和平均数，它更接近于查准率和召回率中较小的数。由于查准率和召回率往往互相影响，F1 值是对这两者的混合评价，在正类反类样本数量不平衡的情况下更能体现出它的效果。其表达式如公式 3-9 所示。

$$F1 = \frac{2*precision*recall}{precision+recall} \quad (3-9)$$

由二分类的混淆矩阵表还可以得到真正类率和假正类率，如下所示：

$$\text{TPR} = \frac{TP}{TP+FN} \quad (3-10)$$

$$\text{FPR} = \frac{FP}{TN+FP} \quad (3-11)$$

ROC/AUC：ROC 是以假正类率 FPR 为横坐标，真正类率 TPR 为纵坐标构造的曲线图。AUC（Area under ROC Curve）表示 ROC 曲线的下面积，AUC 值越大越好，能够直观地反映模型的分类效果。由于样本比例改变不会引起 ROC 曲线的改变，所以 AUC 在正类反类样本数量不平衡的情况下也能比较好地反映出模型分类效果的好坏。

3.6.2 卷积神经网络改进的实验结果与分析

实验首先确定对分类层改进后,将全连接层节点数设置为多少时模型的分类效果最佳。在对 DenseNet169 网络的分类层按上述方法进行改进后,设置了全连接层节点数量为 128、256、512 作为分组实验,其结果如表 3-3 所示。

表 3-3 DenseNet169 分类层改进后设置全连接层不同节点数的测试结果

节点数	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)
128	90.94	93.33	89.81	91.53
256	95.37	94.14	96.78	95.44
512	93.96	91.01	97.59	94.17

由上表可知, DenseNet169 分类层改进后设置全连接层节点数为 256 时,测试结果的准确率、查准率和 F1 值都是最高的,且测试集总体准确率达到分类实验最佳效果 95.37%。为保证后续实验的一致性,对 VGG19、InceptionV3、ResNet50 等模型的分类层改进后,统一设置全连接层节点数为 256。

使用改进分类层后的 VGG19、InceptionV3、ResNet50、DenseNet169 四种神经网络对 SARS-CoV-2 CT 数据集迭代 100 次训练,训练过程中的训练集和验证集的准确率和损失函数值变化如图 3-6 至图 3-9 所示。

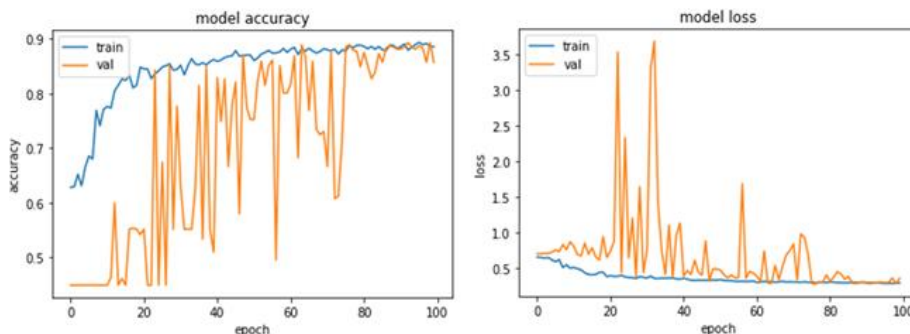


图 3-6 分类层改进后的 VGG19 训练集和验证集准确率和损失

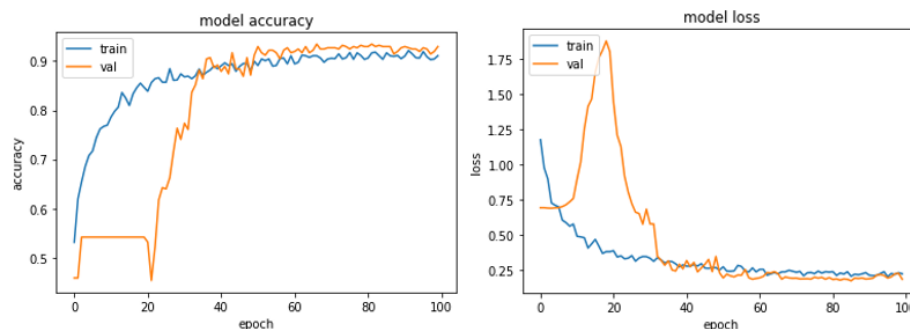


图 3-7 分类层改进后的 InceptionV3 训练集和验证集准确率和损失

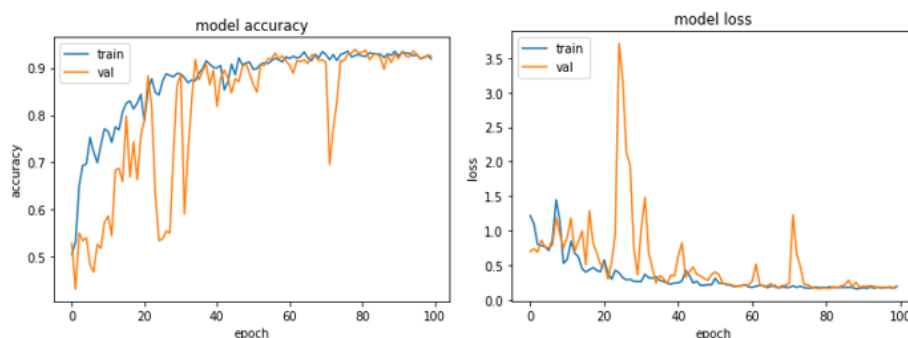


图 3-8 分类层改进后的 ResNet50 训练集和验证集准确率和损失

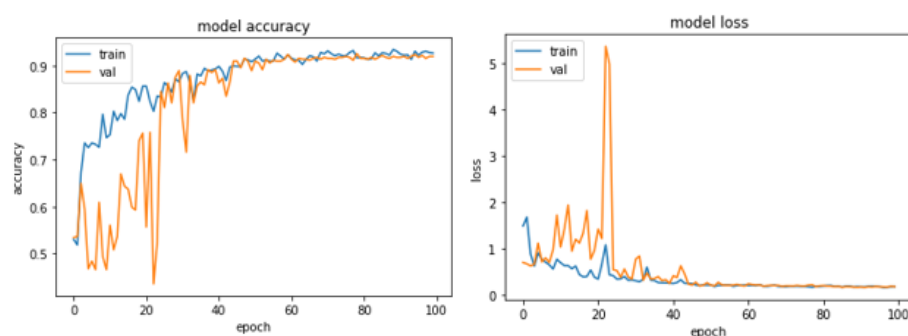


图 3-9 分类层改进后的 DenseNet169 训练集和验证集准确率和损失

本节对 VGG19、InceptionV3、ResNet50、DenseNet169 四种模型分类层改进后的训练模型用测试集进行验证，其实验结果如表 3-4 所示。

表 3-4 卷积神经网络分类层改进后各个模型的实验结果

Module	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)	AUC (%)
VGG19	88.93	91.45	85.94	88.61	95.24
InceptionV3	94.76	95.51	93.97	94.73	98.78
ResNet50	94.36	94.02	94.77	94.39	98.55
DenseNet169	95.37	94.14	96.78	95.44	99.15

从表 3-4 中可见 InceptionV3、ResNet50、DenseNet169 三种网络改进分类层后的总体分类准确率较为相近，相对来说 DenseNet169 的分类效果最好，其准确率、召回率和 F1 值都高于其他三种网络。VGG19 的分类效果最差，且与其他三种网络的差距较大。我们将不同网络分类结果的混淆矩阵进行可视化，上表的各

评价指标均可根据下面的混淆矩阵计算得出，不同网络分类结果的混淆矩阵如图 3-10 所示。

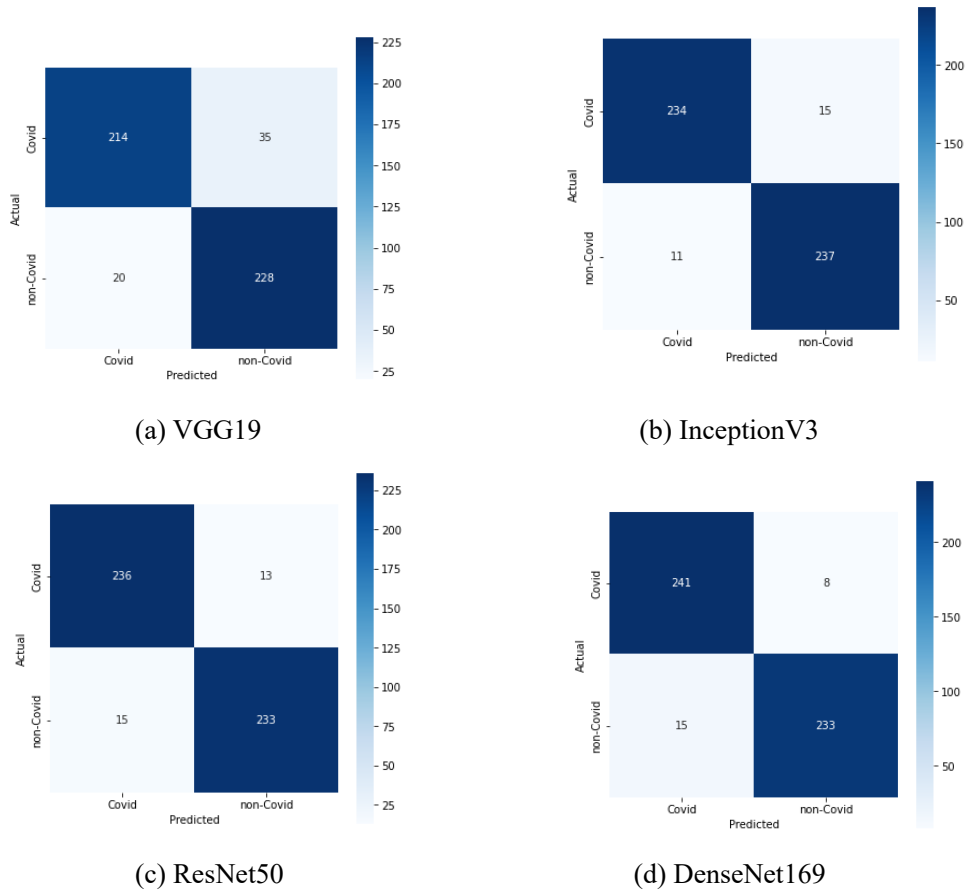


图 3-10 卷积神经网络分类层改进后各个模型的测试集混淆矩阵结果

其中，横轴代表模型预测类别，纵轴代表真实类别。每个数字表示该真实类别所有图像中预测为横轴上对应模型预测类别的图像数量，数字越大方块的颜色越深。容易得到主对角线上的方块是预测正确的，副对角线上的方块是预测错误的。

为验证在分类层结构改进中加入 Batch Normalization 层是否可以加快网络模型的收敛速度，我们在 DenseNet169 改进的分类层结构中未加入 Batch Normalization 层作为对照实验。其训练过程中准确率和损失变化如图 3-11 所示。

对比图 3-11 和 3-9 可以发现，DenseNet169 改进的分类层结构中加入 Batch Normalization 层在 50 个 epoch 左右时有较明显的收敛趋势，而未加入 Batch Normalization 层在 80 个 epoch 左右时才有较明显的收敛趋势。实验验证了加入 Batch Normalization 层确实可以加快模型的收敛速度。

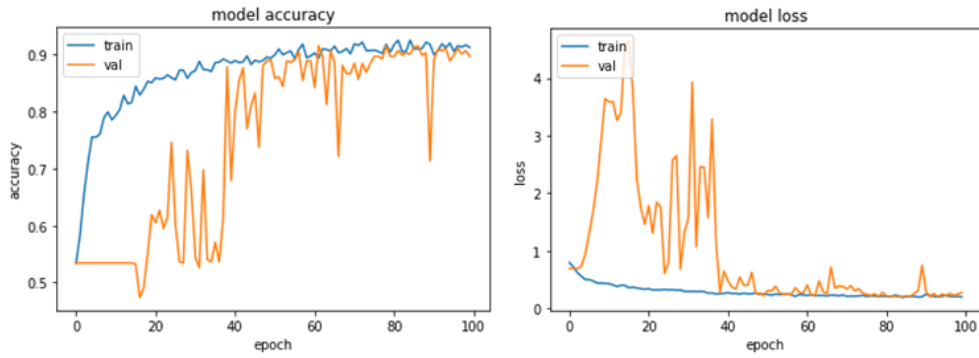


图 3-11 DenseNet169 未加入 Batch Normalization 层时训练集和验证集准确率和损失

为验证模型分类层结构改进的有效性，我们使用 VGG19、InceptionV3、ResNet50、DenseNet169 四种模型原来的网络结构在 SARS-CoV-2 CT 数据集上进行训练，并用测试集验证的结果如下。

表 3-5 卷积神经网络原模型各个模型的实验结果

Module	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)	AUC (%)
VGG19	64.85	66.82	58.23	62.22	73.25
InceptionV3	92.95	92.80	93.17	92.98	98.26
ResNet50	92.15	93.75	90.36	92.02	98.10
DenseNet169	93.36	93.54	93.17	93.35	98.38

对比表 3-5 和表 3-4 中 VGG19、InceptionV3、ResNet50、DenseNet169 四种网络原模型和改进分类层结构后的实验结果，可以看出改进分类层结构后四种网络的总体分类准确率以及其他各项评价指标都有不同程度的提升，说明改进分类层结构可以增强模型的分类性能。其中 VGG19 改进分类层结构后的分类性能提升较大，相对于原模型的总体分类准确率提升了 24.08%，ResNet50、InceptionV3、DenseNet169 三种模型改进分类层结构后的分类性能提升较小，都只有 2% 左右。

我们对比 VGG19、InceptionV3、ResNet50、DenseNet169 四种网络原模型分类层结构和本章中提出的改进后的分类层结构。InceptionV3 和 DenseNet169 原模型的分层结构都是由全局平均池化层加 softmax 层组成，而本章提出的分类层结构改进方法相当于在全局平均池化层和 softmax 层之间加上一层节点数量较小的全连接层。虽然加上一层全连接层后网络的计算量有所提升，但实验结果表明改进后网络模型的分类性能也有一定程度的提升。ResNet50 原模型的分层是由一个 2×2 的平均池化层经过 Flatten 之后接一个 softmax 层组成，其实类似于用全局平均池化层代替全连接层的思想。VGG19 原模型的分层结构是

由两个 4096 节点的全连接层接 softmax 层组成，用本章提出的分类层结构改进方法后，模型的分类性能有了很大的提升。

3.6.3 ResNet 引入注意力机制的实验结果与分析

本实验主要是在 ResNet50 网络分类层改进的基础上，通过加入 SE 模块和 CBAM 模块来进一步提高 ResNet50 网络的分类效果，并与没加入 SE 模块和 CBAM 模块的 ResNet50 分类层改进模型进行对比。本实验采用 SARS-CoV-2 CT 数据集，主要用的评价指标有准确率、查准率、召回率和 F1 分数。用改进分类层结构并加入 SE 模块后的 SE-ResNet50 网络和加入 CBAM 模块后的 CBAM-ResNet50 网络在 SARS-CoV-2 CT 数据集上进行训练，并用测试集验证的结果如下。

表 3-6 ResNet50 引入注意力机制的实验结果

Module	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)
ResNet50	94.36	94.02	94.77	94.39
SE-ResNet50	95.17	94.82	95.58	95.20
CBAM-ResNet50	94.56	94.40	94.78	94.59

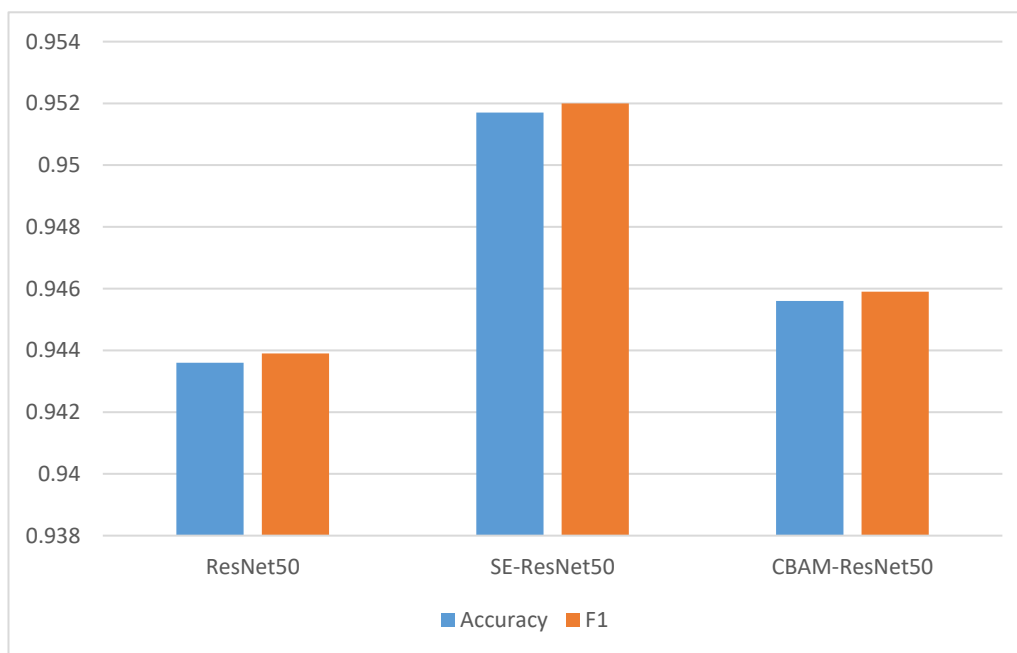


图 3-11 ResNet50 引入注意力机制的实验结果比较

通过以上表格和统计图中的 ResNet50 的实验结果可以发现，加入 SE 模块和加入 CBAM 模块的 ResNe50t 改进模型比没加入注意力机制的 ResNet 改进模型在准确率、查准率、召回率和 F1 分数上都略有提升。这说明了嵌入 SE 模块和 CBAM 模块提升了 ResNet 的分类性能。相对来说 SE 模块的表现略好于 CBAM 模块。

3.7 本章小结

本章先是介绍了实验的数据集，和数据预处理方式以及实验环境，然后主要介绍了对 VGG19、InceptionV3、ResNet50、DenseNet169 等卷积神经网络分类层改进和在 Resnet50 中加入 SE 模块和 CBAM 模块的方法，设置对照实验并进行训练。最后介绍了实验的评价指标，并对实验结果进行分析。实验结果表明，对卷积神经网络分类层进行改进和加入 SE 模块、CBAM 模块的方法提升了对新冠肺炎 CT 图像的分类准确率。

第四章 基于迁移学习的新冠肺炎 CT 图像分类

本章基于 VGG19、InceptionV3、ResNet50、DenseNet169 四种卷积神经网络进行迁移学习训练对新冠肺炎 CT 图像分类。首先介绍了数据集大小不同和该数据集与源域数据集相似程度不同等情况下, 如何选择更合适的迁移学习方法。由于 SARS-CoV-2 CT 数据集和 ImageNet 数据集相似程度很小, 我们选择用微调方式的迁移学习对 SARS-CoV-2 CT 数据集进行分类。针对 COVID-CT 数据集数据量少的问题, 本章提出了微调与特征提取相结合的迁移学习方法。先对 ImageNet 预训练模型用微调的迁移学习方式, 获得对 SARS-CoV-2 CT 数据集的分类模型, 然后利用这个模型的参数采用特征提取的迁移学习方法, 冻结所有卷积层, 用 COVID-CT 数据集仅对剩下的分类层进行训练。最终得到对 COVID-CT 数据集的分类模型。

4.1 数据集

在本实验中仍然使用第三章中的 SARS-CoV-2 CT^[48]数据集, 同时还采用了 COVID-CT^[54]数据集 (<https://github.com/UCSD-AI4H/COVID-CT>)。

COVID-CT^[54]数据集是由加州大学圣地亚哥分校的 XingyiYang 等人构建的, 包含从 216 例新冠肺炎患者获取的 349 张新冠肺炎阳性病历图片, 和 397 张新冠肺炎阴性图片。XingyiYang 等人将收集到的原始 CT 图像, 交给放射科医师进行病变损伤标记, 并基于多任务学习和自监督学习的方法来训练 COVID-19 的诊断模型。该模型的 F1 分数为 0.90, AUC 为 0.98, 精度为 0.89。

4.2 迁移学习方法的选择

基于卷积神经网络使用迁移学习能够利用预先学习的类似任务的知识来提高对新任务的性能。它克服了数据稀缺问题并节省了时间和硬件资源, 为医学图像分析做出了重大贡献。

基于卷积神经网络使用迁移学习通常有两种方法: 第一种微调卷积神经网络, 保持模型原有的结构和初始权重不变, 用自己需要学习的数据来微调预训练好的网络模型的参数。在微调的时候可以选择冻结前面几层的网络权重保持不变, 仅对后面几层网络进行微调训练, 也可以选择对全部层采用微调的方式进行训练。第二种把卷积神经网络作为特征提取器, 是把预训练好的卷积神经网络模型提取

最后一层卷积层的输出或者提取除去最后一个全连接输出层后的倒数第二层的输出作为卷积神经网络的特征，然后按自己的需求构建分类器或者直接使用 SVM、贝叶斯或 softmax 等分类器来完成新的任务。此时将之前预训练好的每一层网络参数冻结，仅对最后增加的分类器进行训练。

如何选择这两种迁移学习方法一般根据目标数据集的规模大小和该数据集和源域数据集的相似程度来决定，可以分为以下四种情况：

(1)目标数据集较小并且和源域数据集相似度较高。由于数据集较小，用微调的方式进行迁移学习训练容易造成过拟合的现象。而由于目标数据集和源域数据集相似度较高，两个数据集图像的特征比较接近。此时把整个卷积神经网络作为特征提取器会是更好的选择；

(2)目标数据集较大且和源域数据集相似度较高。由于目标数据集较大，用微调的方式进行迁移学习训练并不容易发生过拟合现象，且采用预训练模型效率更高，此时保持模型原有的结构和初始权重不变，将网络结构的全部层进行微调更好；

(3)目标数据集较小且和源域数据集相似度很低。同样的由于目标数据集较小，用微调的方式进行迁移学习训练容易造成过拟合的现象，一般不会选择微调的方式。但两个数据集的相似度较低，图像特征差距也会较大。这种情况下的解决方案一般是仅冻结卷积层中较低的层，对较高层的网络仍然进行训练，用低层卷积层进行特征提取，因为卷积神经网络浅层网络的特征具有通用性；

(4)目标数据集较大但和源域数据集相似度很低。这种情况可以从头训练卷积神经网络。但为了提升网络的训练速度，仍会使用迁移学习的方法，且更偏向于将网络结构的全部层进行微调。

4.3 集成学习概述

集成学习通常被用在监督机器学习任务中，在计算智能和机器学习领域中受到了越来越多的关注。集成系统最初是为了减少自动决策系统的方差而开发的，后来成功地用于解决各种机器学习问题，如特征选择、置信度估计等。

集成学习的主要思想是，将多个学习算法获得的预测结果进一步聚合，以提高预测性能；即使某个算法预测错误，其他的算法也可以将错误纠正回来。这些不同的算法一般被称为基学习器或弱学习器，集成学习中包含同类型的弱学习器称为同质，包含不同类型的弱学习器称为异质。通过集成多个同质或异质弱学习器，可以有效提高单个学习器的泛化能力和鲁棒性^[55]。许多研究表明，集成学习比单个弱学习器具有更好的预测性能，并已广泛用于不同的问题场景。基于不同的集成策略，集成学习主要可分为三类^[56]：Bagging、Boosting 和 Stacking。

Bagging^[57]策略是一种并行方法。先从整个数据集中采用有放回的随机抽样方式,得到 N 个子数据集,在每个子数据集上学习训练得到一个弱学习器,对这 N 个弱学习器的预测结果用投票或者平均的方式获得最终的预测结果。每个弱学习器彼此独立,因此可以同时训练。一般对分类问题采用投票的方式,对回归问题采用取平均值的方式。

Boosting^[58]集成学习方法是一个顺序学习的过程,每个弱学习器不能独立训练。前一个弱学习器训练好后,会根据这个弱学习器的表现来更新训练样本的权重,预测错误的样本会在下一次训练中受到更多的关注,然后用更新权重后的样本训练下一个弱学习器,重复进行这个过程直到弱学习器数量达到指定的数目,再对这些弱学习器进行结合获得强学习器。

Stacking 策略是通过组合其他各个模型来训练一个模型。首先训练多个不同类型的弱学习器,将这些弱学习器的输出作为输入来训练一个次级学习器,以得到最终的预测结果。本文采用的是 Stacking 集成学习策略。

4.4 基于迁移学习的新冠肺炎 CT 图像分类方法

4.4.1 微调方式的迁移学习训练

包括卷积神经网络在内的深度学习算法往往需要大量的数据进行训练,在数据量不足的情况下很容易出现过拟合的问题。由于医学图像经常涉及病人的隐私,而且专家注释数据集的成本非常高,因此数据缺乏问题在用深度学习研究医学图像的领域中更为突出。对此很多研究工作都试图通过迁移学习来克服这个问题。ImageNet 数据集包含了数百万张自然图像,分为 1000 种图像类别,是目前应用最广泛的大规模图像数据集之一。在许多领域中,利用在 ImageNet 数据集上预训练好的模型进行迁移学习训练都取得了不错的成果,ImageNet 数据集也成为了医学领域中热门的迁移学习对象。

SARS-CoV-2 CT 数据集相对来说较大,依照迁移学习方法选择的四种不同情况来说,载入预训练好模型的权重作为初始权重进行微调比较合适。在本实验中,我们利用在 ImageNet 数据集上预训练好的几种卷积神经网络模型,来提取图像的基础特征,并用微调的方式继续学习 SARS-CoV-2 CT 数据集中胸部 CT 图像的复杂特征。由于不同卷积神经网络的深度和复杂度不同,采用微调方式中冻结部分卷积层的方法难以确定最佳冻结层数,且在实验过程中表现并不好,对于微调方式本实验采用载入权重后训练所有参数的方式。

在载入不同卷积神经网络在 ImageNet 大型数据集上预先训练好的模型后,对整个网络的参数进行微调。本实验中我们依然选取了 VGG19、InceptionV3、

ResNet50、DenseNet169 四种神经网络作为迁移训练实验对象。在第三章中我们已经验证了在分类层进行改进方法的有效性，因此我们仍然对迁移的模型进行同样的改进。在载入 ImageNet 大型数据集上预先训练的模型后我们保留除最后分类层之外的其余所有层，删去网络原来的分类层结构。这一步操作只需将载入模型的 `include_top` 属性设置为 `False` 就能完成。再加入全局平均池化层+BN 层+256 个节点的全连接层+BN 层+2 个节点的 `softmax` 层作为新的分类层。本实验将第三章中手动搭建的卷积神经网络的改进模型的实验结果作为迁移学习训练的对照。

对 SARS-CoV-2 CT 数据预处理后，将预训练好的网络在 SARS-CoV-2 CT 数据集上进行训练，设置 SARS-CoV-2 CT 阳性图像的标签为 0，SARS-CoV-2 CT 阴性图像的标签为 1。损失函数选择交叉熵损失函数 `CrossEntropyLoss`，优化器使用 Adam 优化器并设置参数为：`lr=0.003`，`betas=(0.9, 0.999)`。通过 `ReduceLROnPlateau` 来调整学习率，并设置初始参数为 `factor=0.70`，`patience=5`，`verbose=1`，`min_lr=1e-4`。训练的模型用.h5 格式的文件保存。

4.4.2 集成分类模型的构建

在微调方式的迁移学习训练实验中，我们发现 InceptionV3、ResNet50、DenseNet169 三种卷积神经网络都取得了不错的分类效果。如果使用集成学习方法将这三种卷积神经网络训练好的模型作为弱学习器，将这些弱学习器结合获得强学习器，则可能获得更好的分类效果。我们将 InceptionV3、ResNet50、DenseNet169 三种卷积神经网络在 SARS-CoV-2 CT 数据集上以微调方式迁移学习训练好的模型进行集成，采用 Stacking 集成策略，集成模型的结构如图 4-1 所示。

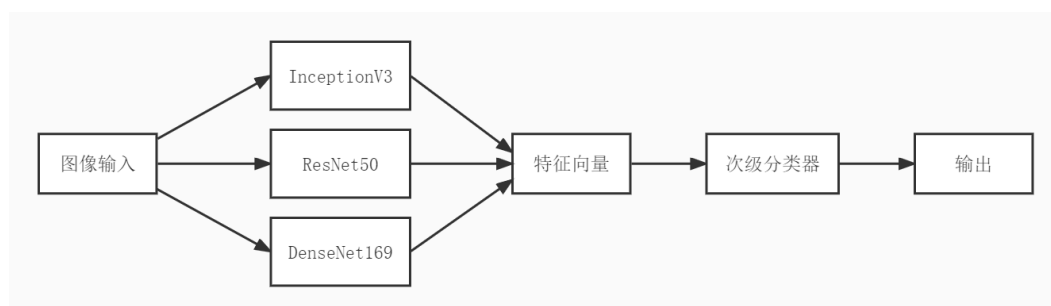


图 4-1 集成模型结构示意图

如上图所示，将输入图像经过 InceptionV3、ResNet50、DenseNet169 三种卷积神经网络在 SARS-CoV-2 CT 数据集上以微调方式迁移学习训练好的模型，把

三个从全局平均池化层输出的特征向量拼接成一个新的特征向量，再送入一个新的次级分类器，完成集成模型最终的分类。

4.4.3 微调与特征提取相结合的迁移学习训练

本实验结合了两种方法对新冠肺炎 CT 图像进行分类实验。由于 COVID-CT 数据集较小，在 COVID-CT 数据集上，本实验采用把预训练网络当作特征提取器的方式，来进行迁移学习训练。考虑到如果源域与目标域图像特征有较高的相似度，迁移起来效果可能更好，所以将 SARS-CoV-2 CT 数据集作为源域进行迁移学习。本实验尝试将 SARS-CoV-2 CT 数据集在 ImageNet 数据集预训练模型进行迁移学习后的网络权重作为特征提取器进行迁移学习训练，即微调与特征提取相结合的迁移训练方法。前后两次迁移的方式不同，前者采用微调的方式，后者采用特征提取的方式。具体实验过程如下：首先下载之前 SARS-CoV-2 CT 数据集在 ImageNet 数据集预训练模型迁移学习后的网络权重，然后冻结所有卷积层，用 COVID-CT 数据集以特征提取的方式仅对剩下的分类层进行训练。

本实验以 COVID-CT 数据集直接在 ImageNet 数据集预训练模型的迁移学习训练，即微调方式的迁移训练，作为微调与特征提取相结合的迁移训练方法的对照。

COVID-CT 数据预处理后，将预训练好的网络模型在 COVID-CT 数据集上进行训练，设置 COVID-CT 阳性图像的标签为 0，COVID-CT 阴性图像的标签为 1。损失函数选择交叉熵损失函数 CrossEntropyLoss，优化器使用 Adam 优化器并设置参数：lr=0.003，betas=(0.9, 0.999)。用 ReduceLROnPlateau 来调整学习率，并设置初始参数 factor=0.70, patience=5, verbose=1, min_lr=1e-4。训练的模型用.h5 格式的文件保存。

4.5 基于迁移学习的新冠肺炎 CT 图像分类结果

4.5.1 微调方式的迁移训练和集成模型的分类结果

按照 4.3.1 中的参数设置，使用 VGG19、InceptionV3、ResNet50、DenseNet169 四种神经网络在 ImageNet 大型数据集上预先训练的模型，使用微调方式对 SARS-CoV-2 CT 数据集迭代 100 次训练，具体训练过程的训练集和验证集的准确率变化以及迭代过程的误差损失如图 4-2 至图 4-5 所示。

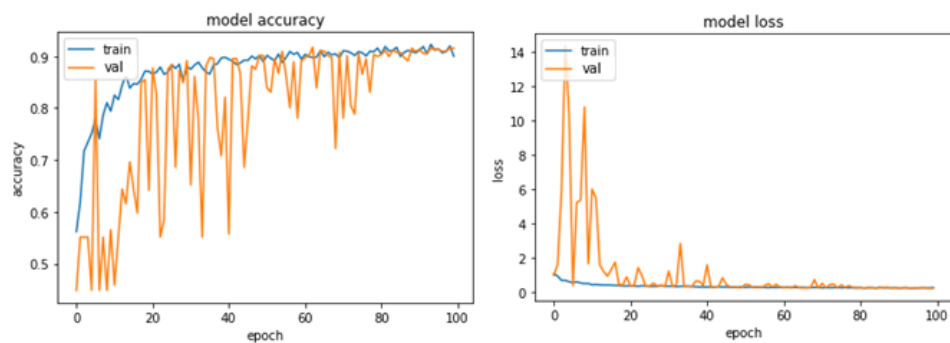


图 4-2 VGG19 迁移学习训练集和验证集准确率和损失

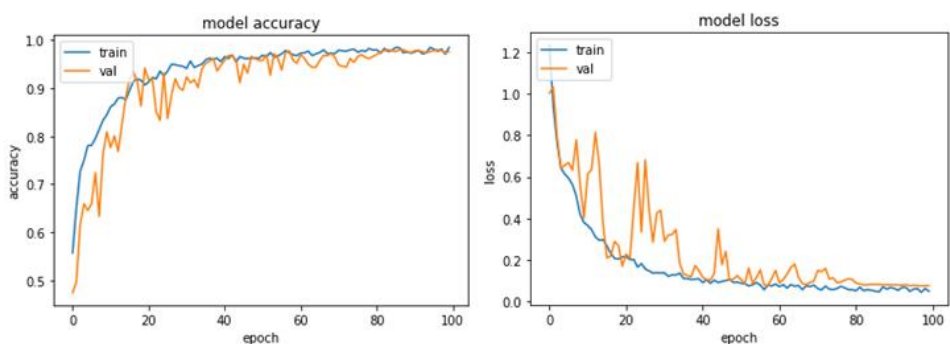


图 4-3 InceptionV3 迁移学习训练集和验证集准确率和损失

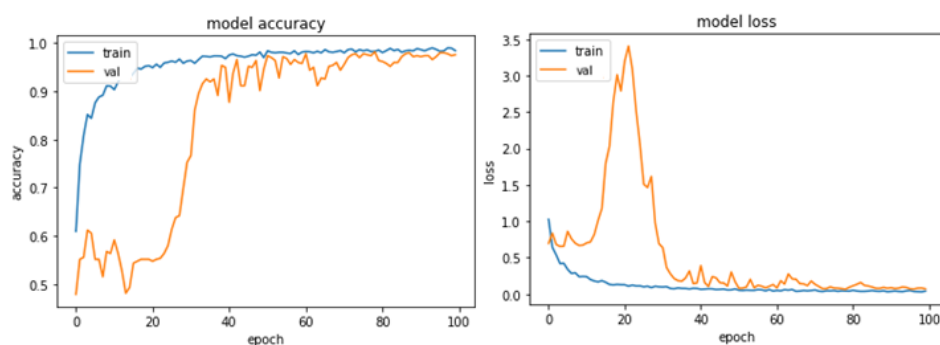


图 4-4 ResNet50 迁移学习训练集和验证集准确率和损失

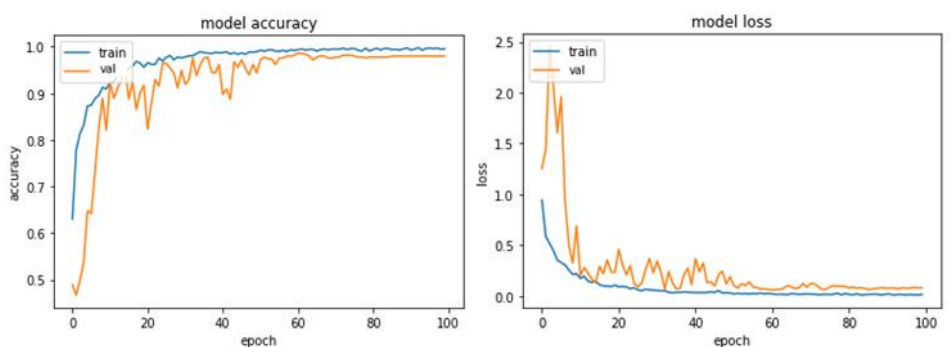


图 4-5 DenseNet169 迁移学习训练集和验证集准确率和损失

本节对 VGG19、InceptionV3、ResNet50、DenseNet169 四种模型进行迁移学

习训练，并用测试集验证结果如表 4-1 所示。

表 4-1 迁移学习的各个模型实验结果

Module	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)
VGG19	91.54	95.31	89.05	92.07
InceptionV3	97.59	97.21	97.99	97.60
ResNet50	97.79	98.37	97.19	97.77
DenseNet169	97.59	98.36	96.79	97.56

从表 4-1 中可以看出 InceptionV3、ResNet50、DenseNet169 三种模型的迁移学习训练在 SARS-CoV-2 CT 数据集都表现出较高的性能，其分类准确率都达到了 97% 以上。VGG19 模型的迁移学习训练在 SARS-CoV-2 CT 数据集上的表现稍差，只有 91.54%。

将表 4-1 对比表 3-4 可以发现，四种模型迁移学习训练后的各个评价指标都比手动搭建模型从头训练要有较大的提升。VGG19、InceptionV3、ResNet50、DenseNet169 四种模型的迁移学习的分类准确率训练相对于手动搭建模型都有 2.2%~3.5% 左右的提升。实验证明使用微调方式的迁移学习方法可以提高模型在新冠肺炎 CT 图像分类任务的准确率。

对 InceptionV3、ResNet50、DenseNet169 三种卷积神经网络在 SARS-CoV-2 CT 数据集上以微调方式迁移学习训练好的模型进行集成，并用测试集验证结果如表 4-2 所示。

表 4-2 集成模型的实验结果

Module	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)
Ensemble Model	98.59	99.59	97.59	98.58

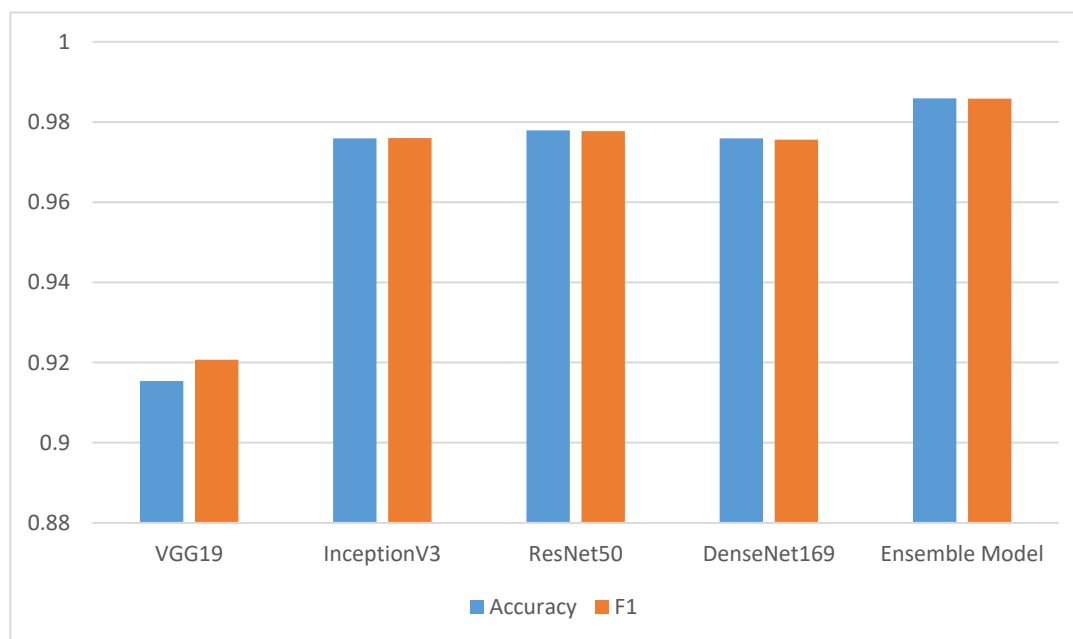


图 4-6 不同分类模型的效果比较

从以上表格和统计图中可以看出，集成分类模型的分类效果优于单个弱学习器的分类效果，在准确率和 F1 分数上都有所提升，并且集成分类模型取得了在本文中最好的分类效果，准确率为 98.59%。

我们将当前已有的研究结果与本文所提出的方法进行对比。如 Wang^[21]等人使用 InceptionNet 检测 COVID-19 CT 图像。他们收集了 1065 张确诊 COVID-19 CT 图像和诊断为病毒性肺炎的 CT 图像，其中 325 张 CT 图像为新冠肺炎确诊，740 张 CT 图像为病毒性肺炎，识别准确率为 89.50%。Xu 等人^[22]使用 3D 深度学习模型从肺部 CT 图像集中分割出候选感染区域，使用 ResNet 结合位置注意力机制提取特征。他们对来自 90 个 CT 样本的共 1710 张 CT 图像进行分类，其中包括 357 张 COVID-19 CT 图像、390 张甲型流感病 CT 图像和 963 张健康病例图像，最后得到的总体分类准确率为 86.70%。Li 等人^[23]开发了 2019 冠状病毒疾病检测系统（CVNET），从胸部 CT 图像中提取出特征，用于检测冠状病毒感染。利用 COVNet 对 3322 名患者的 4356 张胸部 CT 图像进行了训练，包括 1296 张 COVID-19 CT 图像、1735 张社区获得性肺炎（CAP）CT 图像和 1325 张健康病例 CT 图像。从 COVNet 获得的 COVID-19 和 CAP 的 AUC 值分别为 96%和 95%。由于不同文献中的数据集和分类对象均不相同，因此不能简单的说哪篇文章的实验结果更好。Jaiswa 等人^[27]借助卷积神经网络和预训练的 DenseNet201，设计了一种深度迁移学习模型。该模型对胸部 CT 扫描图像进行分类，其训练集、验证集、测试集的分类准确率分别为 99.82%、97.4%、96.25%。其数据集和本文中使用的数据集均为 SARS-CoV-2 CT 数据集，本文实验结果中集成分类模型取

得了最好的分类效果，测试集分类准确率为 98.59%，与之相比高出 2.34%，验证了本文方法的有效性。

4.5.2 微调与特征提取相结合的迁移训练分类结果

由于在微调方式的迁移学习训练实验中，VGG19 模型的表现与其它三种网络模型相比稍有差距，本实验选取 InceptionV3、ResNet50、DenseNet169 三种模型进行微调与特征提取相结合的迁移训练实验。按照 4.3.2 中的参数设置，首先下载之前 SARS-CoV-2 CT 数据集在 ImageNet 数据集预训练模型经过微调方式的迁移学习后的网络权重，然后冻结所有卷积层，用 COVID-CT 数据集以特征提取的方式仅对剩下的分类层进行训练。

实验还设置了 COVID-CT 数据集直接在 ImageNet 数据集预训练模型上以微调方式迁移学习训练作为对照实验。COVID-CT 数据集通过微调与特征提取相结合的迁移学习和仅通过微调方式的迁移学习效果对比如表 4-3。

表 4-3 两种迁移学习方法的性能比较

训练方式	Module	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)
微调	InceptionV3	86.66	82.66	89.85	86.1
	ResNet50	85.33	87.5	82.89	85.13
	DenseNet169	90	86.48	92.75	89.5
微调与特征提取相结合	InceptionV3	89.33	85.33	92.75	88.89
	ResNet50	88.66	91.42	87.67	89.51
	DenseNet169	91.33	85.89	97.10	91.15

表 4-3 列出了 InceptionV3、ResNet50、DenseNet169 三种模型在采用微调方式以及采用微调与特征提取方式相结合两种迁移学习方法下，模型的准确率、查准率、召回率、F1 分数的比较结果。从表 4-3 中可以看出使用微调与特征提取相结合方式的迁移学习对比微调方式的迁移学习在各个评价指标说都有所提升。说明在数据集较小的情况下，使用微调与特征提取方式相结合的迁移学习方法，可以进一步提升迁移学习模型的分类效果。

4.6 可视化研究

在上文中，我们使用微调迁移学习方式以及对 InceptionV3、ResNet50、DenseNet169 三种卷积神经网络迁移学习模型进行集成后得到的集成分类模型，

在对新冠肺炎 CT 图像的分类任务中都取得了不错的效果。但是卷积神经网络模型学习的过程具有不透明性和不可解释性，其结果往往难以信任。

我们知道卷积神经网络可以从数据中提取重要特征，并按照这些特征对图像进行分类。但学习过程中模型具体是按照哪些特征进行分类，以及模型使用的这些分类特征是否与现实诊断中使用的特征相关，这些都是我们难以得知的。如果我们能将卷积神经网络某一层抽象程度高的特征图可视化，并标注其重点关注的特征区域，我们就可以分析网络是否学习到了正确的特征信息。

Zhou 等人^[59]分析了卷积神经网络视觉编码过程，提出了类激活映射方法（CAM，Class Activation Mapping）。该方法利用全局平均池化层（GAP，Global Average Pooling）替换掉了全连接层，好处是减少了全连接层大量的参数，在一定程度上避免过拟合问题，能够获得较强的鲁棒性，更重要的是在最后一个卷积层生成和目标类别数量一致的特征图，使每个特征图赋予的意义更明确。通过计算全局平均池化层之后线性操作的权重，求对应特征图的加权和后覆盖到原始图像上，最终产生对应的热力图进行可视化。然而 CAM 方法的一个比较严重的缺点是需要改变模型的结构，因此对模型需要重新训练，限制了 CAM 方法的使用场景。

为此 Selvaraju 等人^[60]对 CAM 方法进行改进，提出了 Grad-CAM 方法。Grad-CAM 是基于梯度的类激活映射，CAM 通过使用全局平均池化层得到权重，而 Grad-CAM 直接通过梯度的全局平均得到权重。因此 Grad-CAM 解决了 CAM 需要修改模型的问题，使用更加方便。由于网络的最后一个卷积层特征抽象程度最高，语义信息最丰富，因此通常对最后一个卷积层使用 Grad-CAM。具体过程首先通过正向传播得到最后一层卷积层的特征图，然后反向传播计算最后一层卷积层的梯度，用梯度的全局平均来计算权重，然后求对应特征图的加权和并覆盖到原始图像上，就能得到对图像上各个像素显示不同关注程度的热力图。本文使用 Grad-CAM 来对模型分类过程中关注的特征进行可视化。

我们选取了微调方式迁移学习的 ResNet50 网络模型作为可视化模型，提取最后一层卷积层的特征图做基于梯度的类激活映射，通过热力图的方式标注其重点关注的特征区域。热力图中颜色越接近于红色表示该区域在分类时越受到关注。分类正确的新冠肺炎患者 CT 图像和热力图如图 4-7 所示，分类错误的新冠肺炎患者 CT 图像和热力图如图 4-8 所示。

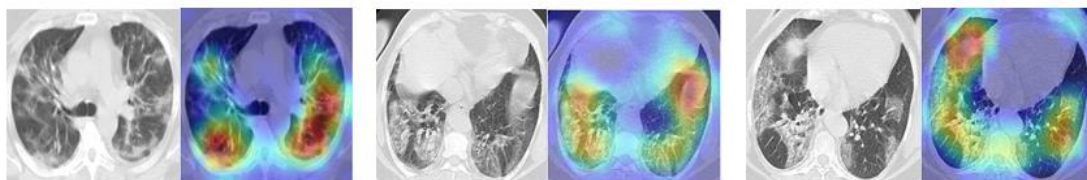


图 4-7 分类正确的新冠肺炎患者 CT 图像和热力图

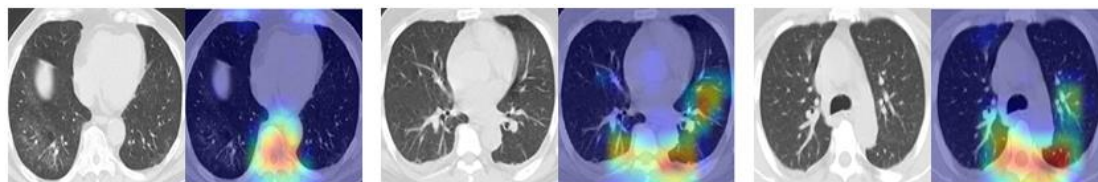


图 4-8 分类错误的新冠肺炎患者 CT 图像和热力图

由图 4-7 和图 4-8 中可以看出，分类正确的新冠肺炎患者 CT 图像其热力图关注的区域更接近新冠肺炎 CT 影像学特征，而当热力图关注其他无关区域时往往会得到错误的分类结果。说明通过梯度的类激活映射方法对模型分类结果的解释具有合理性。

4.7 本章小结

本章介绍了实验中用到的另一个数据集，然后介绍了迁移学习的使用准则。在考虑迁移学习的使用准则后对 SARS-CoV-2 CT 数据集设计了基于微调方式的迁移训练实验，选取了 VGG19、InceptionV3、ResNet50、DenseNet169 四种神经网络作为迁移训练实验模型。对于图像数量较少的数据集，设计了微调和特征提取相结合的方式迁移学习训练，并选取 InceptionV3、ResNet50、DenseNet169 三种神经网络作为迁移训练实验模型。

对比微调方式的迁移训练和第三章中手动搭建的卷积神经网络的改进模型的实验结果，实验结果证明微调方式的迁移学习的方法可以有效的提高模型在新冠肺炎 CT 图像分类的准确率。通过对几种微调方式迁移学习训练好的模型进行集成，集成模型的分分类准确率达到到了 98.59%。此外还在小数据集 COVID-CT 上，对比基于微调方式和基于微调和特征提取相结合的方式，证明了使用微调与特征提取方式相结合的迁移学习方法，可以进一步提升迁移学习模型的分分类效果。最后，通过基于梯度的类激活映射可视化方法，对模型的分分类结果进行解释。

第五章 新冠肺炎检测系统的实现

本文前几章的所做的工作都是为了提升对新冠肺炎 CT 图像的分类准确率，最终获得的分类模型也取得了不错的分类效果。为能更好地利用本文的分类模型，在本章研究了基于 Flask 框架的新冠肺炎检测系统。新冠肺炎检测系统的设计结合了 MVC 三层架构，前端使用 HTML 和 CSS 设计，后端使用 Flask 的 Web 框架进行设计。

5.1 系统需求分析

5.1.1 系统编写目的

对新冠肺炎影像学图像进行分类是当前热门的研究话题，但通过查阅文献发现，许多学者的研究内容都停留在构建和训练模型这一阶段，只是对模型的分类效果做出评价，并没有将优化好的模型进行充分利用。在当今新冠肺炎蔓延的情况下，许多医院都有对新冠肺炎 CT 图像进行识别分类的需求。本文研究的新冠肺炎检测系统可以为用户提供快速检测肺部 CT 图像的方法，让本文研究得到的新冠肺炎 CT 图像分类模型得到充分的应用。

本系统主要以训练好的卷积神经网络模型为基础，结合 Flask Web 开发框架，实现图片上传、图片检测等功能，让用户可以简便快捷的对肺部 CT 图像进行新冠肺炎检测。

5.1.2 功能需求分析

系统主要包括图片上传和 CT 图像检测等功能，用户启动 Flask 后台服务后会自动在浏览器上弹出图片上传界面，只需将图片上传，系统会加载模型对图片进行特征提取，对图片进行预测并返回预测结果。

5.2 Flask 框架

Flask、Django 及 Tornado 是目前主流的 Python Web 编程框架^[61]。Django^[62]是功能最完整的重量级 web 框架，自带如数据库 ORM 组件、缓存、权限管理、用户认证等常用工具，多用于开发功能齐全的企业级网站。Tornado^[63]是注重性能的轻量级 web 框架，它最受瞩目的是异步非阻塞 IO 的处理方式，同时它还拥有较出色的抗负载能力，但相对来说 Tornado 的入门门槛较高。Flask 以微型框

架著称，简洁灵活的同时还容易扩展，可以用于开发 Web 应用程序和服务，但多用于简单需求和小型应用。它只实现核心功能，使开发人员可以灵活地在实现过程中根据需要添加功能。本文开发的新冠肺炎检测系统功能需求较为简单，因此选择使用 Flask 框架。

Flask 有两个核心库^[64]，Werkzeug 和 Jinja2。Werkzeug 负责提供路由、调试和 Web 服务器网关接口(WSGI)，Jinja2 一个功能齐全模板引擎。Flask 本身不支持数据库访问、用户身份验证或任何其他高级实用程序，但它可以通过扩展集成来添加所有这些功能。Flask 应用程序可以放入单个 Python 文件中，也可以将其模块化成应用程序。

5.3 系统结构设计

本实验设计的新冠肺炎检测系统主要包括两个部分，分别是图像上传模块、图像检测模块。其中图像上传模块主要实现 CT 图像的选择、图像的上传以及图像尺寸的调整。图像检测模块是本系统中最主要的部分，用来实现本文训练好的分类模型的调用以及返回最终的检测结果并显示。新冠肺炎检测系统的结构如图 5-1 所示。

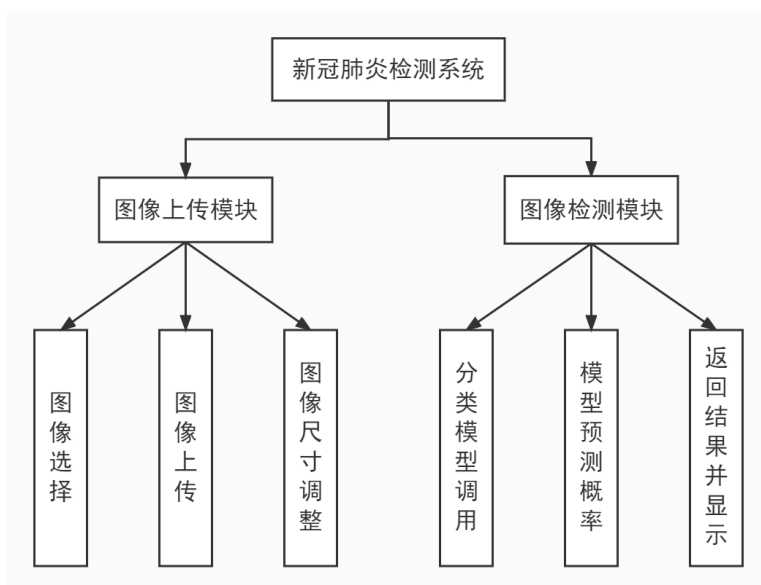


图 5-1 系统结构图

5.4 系统功能实现

5.4.1 图像上传模块



图 5-2 图像上传界面

图 5-2 为图像上传界面，主要实现 CT 图像的选择、图像的上传以及图像尺寸的调整。上传界面由 HTML 和 CSS 完成网页的实现与布局，用 CSS 文件控制 HTML 内页面的布局和样式。通过 JavaScript 编写的 js 文件可以完成前端页面的逻辑处理。例如访问 HTML 元素，JavaScript 可使用 `document.getElementById` 方法，并用 `fileSelect.addEventListener` 监听 HTML 中的事件。

图中虚线框内是图像上传的区域，可以通过单击虚线框内从本地文件中选取图片，也可以通过将图片拖放到虚线框内来选中图片。点击检测按钮后，即可将图片上传并调用图像检测模块进行检测。为防止用户选择图片错误，图像上传界面上设置了清除按钮，点击后可取消图片选中来重新选择图片，也可以直接在虚线框内点击或者将图片拖放到虚线框内来重新选择图片。图像上传后，会自动调整为 224×224 的尺寸。

在调用图像检测模块前，必须确保图片选中以上传图片。因此当未选择图片就点击检测按钮时，应停止上传并提示警告。对此我们使用 JavaScript 中 `alert` 方法弹出提示窗口，来告诉用户必须先选中图片，如图 5-3 所示。

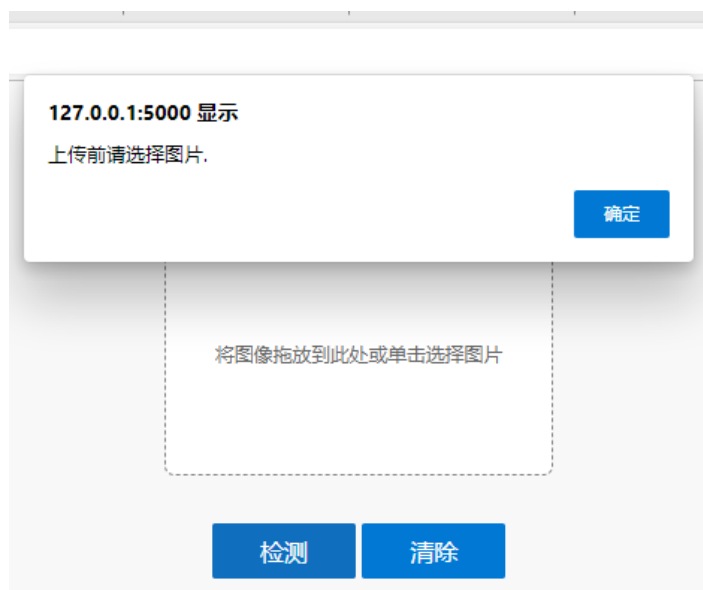


图 5-3 提示弹窗

5.4.2 图像检测模块

图像检测模块是本系统中最主要的部分,用来实现本文训练好的分类模型的调用以及返回最终的检测结果并显示。首先将本文训练好的分类模型放在指定文件夹下,以便在后端加载模型。在用户启动 Flask 后台服务器后,首先显示图像上传界面,在用户按要求选择图片并点击检测按钮后,Web 服务器首先存储上传的图像并加载训练好的分类模型,然后定义一个 `model_predict` 函数,传入图片和模型,对图片预处理后使用 Python 的 `predict()` 函数来进行前向传播的训练,并得到对该图片的预测结果。Flask 调用 `jsonify()` 函数将产生响应的结果返回给前端,发送 JSON 响应,生成预测结果的输出。这里使用 `jsonify()` 函数而不使用 `json()` 函数,其好处是让返回的响应符合 HTTP 协议,同时可以自动去掉不需要的空格。最后将生成的预测结果通过 JavaScript 在浏览器中显示。

图像检测模块的工作流程如图 5-4 所示。

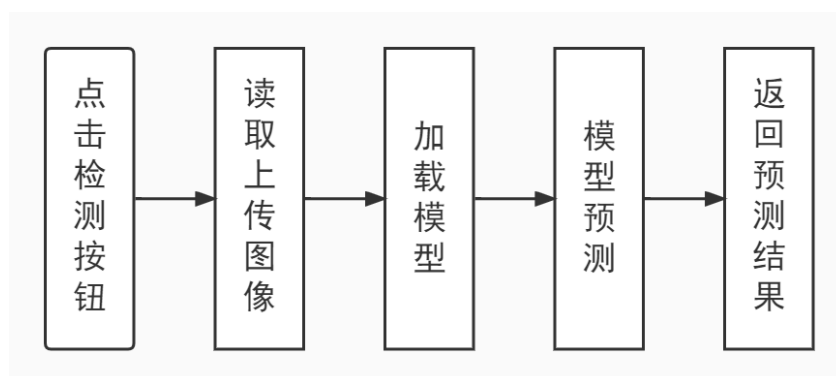


图 5-4 图像检测模块工作流程图

5.4.3 新冠肺炎检测系统演示

本文的新冠肺炎检测系统简单易用。首先启动 Flask 后台服务后会自动在浏览器上弹出图片上传界面，然后按要求选择图片，之后点击检测按钮就可以完成上传图片、加载模型、模型预测以及返回预测结果等过程，能够直接获得模型预测该图片为新冠肺炎的概率。选择图片，点击检测按钮并返回预测结果的演示如图 5-5 所示。

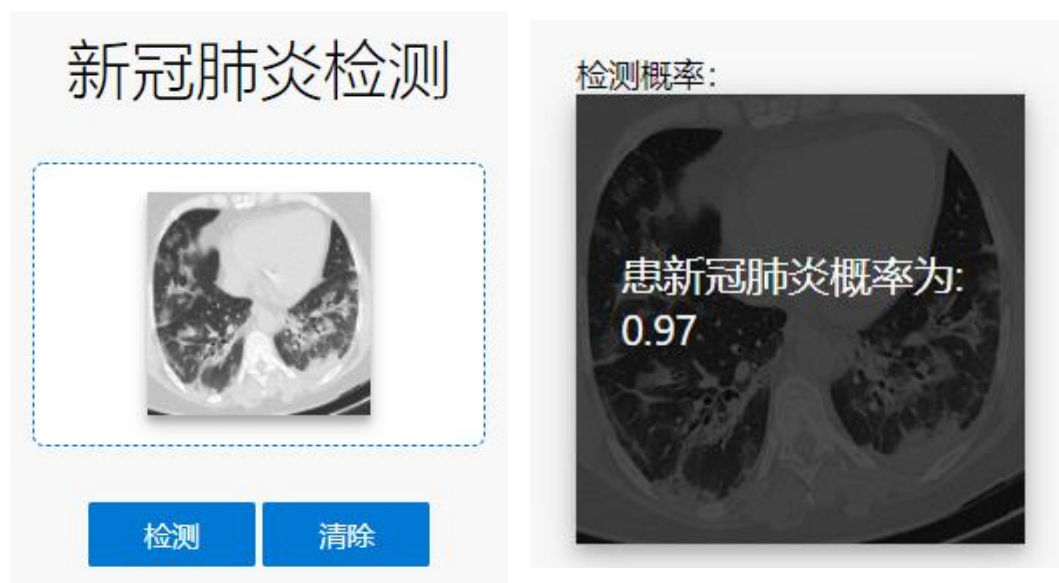


图 5-5 新冠肺炎检测系统演示

5.5 本章小结

本章主要对新冠肺炎检测系统进行设计与实现，首先进行了系统需求分析，其次介绍了 Python Web 编程框架的选择及系统的结构设计，最后分别对图像上传模块和图像检测模块进行系统功能实现，并对新冠肺炎检测系统进行演示。

第六章 总结与展望

6.1 总结

随着新冠肺炎在全世界范围内不断蔓延,以及感染人数的不断增加,配合疫情防控工作需要对新冠肺炎进行检测尤为重要。能够尽早且准确地检测识别新冠肺炎,对防止疫情扩散,提高新冠肺炎患者的治疗效果等方面都具有重大意义。而利用计算机辅助诊断技术进行新冠肺炎检测,可以在一定程度上避免漏诊与误诊,能够帮助医生做出更准确的诊断。本文针对新冠肺炎的 CT 影像,使用深度学习的方法对其进行分类研究,并取得了一定的成果。

本文研究了基于 VGGNet、GoogLeNet、ResNet、DenseNet 等经典卷积神经网络模型对新冠肺炎 CT 图像的分类效果。并在此基础上,通过改进分类层结构、嵌入 SE 和 CBAM 模块等方式来提升网络的分类性能。其中通过改进分类层结构的方法,InceptionV3、ResNet50、DenseNet169 三种网络模型改进分类层结构后相较于原模型准确率提升了 2%左右,VGG19 改进分类层结构后的分类性能提升很大,相对于原模型的总体分类准确率提升了 24.08%,实验说明了改进卷积神经网络分类层结构方法的有效性。通过对改进后的 ResNet50 模型嵌入 SE 模块和 CBAM 模块,其模型分类准确率进一步提升了,说明了在卷积神经网络中嵌入 SE 模块和 CBAM 模块能够提升在新冠肺炎 CT 图像分类任务上的性能。

为进一步提升对新冠肺炎 CT 图像的分类效果,本文使用了迁移学习和集成学习的方法。在分类准确率上,VGG19、InceptionV3、ResNet50、DenseNet169 四种模型的微调方式迁移学习训练相对于手动搭建模型都有 2.2%~3.5%左右的提升,说明了迁移学习的有效性。最后对 InceptionV3、ResNet50、DenseNet169 三种卷积神经网络在 SARS-CoV-2 CT 数据集上以微调方式迁移学习训练好的模型进行集成,集成分类模型取得了本文的最佳分类效果,分类准确率为 98.59%。

对于 COVID-CT 图像数据集数据量过少的问题,本文提出了基于微调和特征提取相结合的迁移学习方法。首先下载 SARS-CoV-2 CT 数据集在 ImageNet 数据集预训练模型以微调方式迁移学习后的网络权重,然后冻结所有卷积层,用 COVID-CT 数据集仅对剩下的分类层进行训练,即特征提取的迁移学习方式。设置对照实验,将 COVID-CT 数据集直接在 ImageNet 数据集预训练好的模型上以微调方式进行迁移学习训练。用测试集对于本文构建的基于微调和特征提取相结合的迁移模型进行验证,并采用多种评价指标对实验结果进行评估。对实验

结果进行对比发现，采用微调和特征提取相结合的迁移模型均优于直接从 ImageNet 数据集上微调的迁移学习模型，实验结果验证了本方法的有效性。本文还通过基于梯度的类激活映射可视化方法，得到对图像上各个像素显示不同关注程度的热力图，对模型的分类结果做出一定的解释。

最后本文用 Flask Web 开发框架搭建了新冠肺炎检测系统，能够用已训练好的模型来预测用户上传的图像，为用户提供了简便快捷的新冠肺炎 CT 图像检测方式。

6.2 展望

本文研究了基于卷积神经网络和迁移学习方法对新冠肺炎 CT 图像分类，但由于个人能力不足，对深度学习的研究和了解有限，研究过程中还存在许多不足之处，未来的研究工作可以从以下几个方面展开：

（1）模型优化。选择更多的分类模型并进行参数调节，以找出更加优秀的分类模型；

（2）结合多种医学影像数据进行分类并比较。本文只使用了新冠肺炎 CT 图像进行分类，今后有机会可以尝其他的医学影像数据，如 X 胸片数据、磁共振成像数据等进行对比研究；

（3）本文研究中的数据量十分有限，希望能有更大规模、更高质量的数据集来支持相关的研究；

（4）将分类模型部署到云平台（比如亚马逊、阿里云等）上，方便以后将本文中的代码复现和迁移应用；

（5）本文仅对新冠肺炎 CT 图像进行分类，虽然可以对新冠肺炎进行检测，但用途仍然十分有限。如果可以获得数量足够的新冠肺炎不同进展阶段的图像数据，还可以对新冠肺炎进行分级。此外还可以采用目标检测算法，识别病灶的具体位置，能够进一步辅助医生进行诊断

致谢

光阴似箭，转眼间研究生三年的生活即将结束，而我也即将步入新的人生旅程。回首这三年的时光，遇到过困难与挑战，也有过收获与喜悦。在此，我谨向读研期间和在论文写作过程中帮助、关心过我的老师、同学表达自己深深的感恩之情。

首先，我要感谢我的导师张会雄老师。从确定论文题目到思路和框架的整理以及论文内容的充实的过程中，张会雄老师都给予了我细致耐心的辅导，使我在写作过程中豁然开朗。张老师严肃认真的治学态度、精益求精的工作作风深深激励着我。在生活和就业上，张老师也给了我很多关怀和指导。在这里对张老师表示诚挚的感谢。

其次我要感谢教研室的师兄师姐，他们在学习和生活中给予我非常多的帮助；感谢教研室的同学和师弟，感谢我的室友们，正是因为有了你们的共同学习和陪伴才使我这三年的研究生生活丰富多彩。

最后我要感谢我的父母，他们的支持和付出为我提供了坚强的后盾，使我一路披荆斩棘、勇往直前。由衷祝愿他们身体健康。

参考文献

- [1]. Zhu N, Zhang D, Wang W, Li, X, et al, A Novel Coronavirus from Patients with Pneumonia in China, 2019 , New England Journal of Medicine, 382(8):727-733, 2020.
- [2]. Perlman, S., Another Decade, Another Coronavirus. (2020) New England Journal of Medicine, 382(8): 760:762, 2020.
- [3]. Hui DS, I Azhar E, Madani TA, Ntoumi F, Kock R, et al, The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health - The latest 2019 novel coronavirus outbreak in Wuhan, China, International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases. 91: 264-266, 2020.
- [4]. Singhal T. A review of coronavirus disease-2019 (COVID-19)[J]. The indian journal of pediatrics, 2020, 87(4): 281-286.
- [5]. Wang W, Xu Y, Gao R, et al. Detection of SARS-CoV-2 in different types of clinical specimens[J]. Jama, 2020, 323(18): 1843-1844.
- [6]. 张瑞,吴国球.SARS-CoV-2 的生物学性状和病毒检测方法以及 COVID-19 的治疗现状[J]. 东南大学学报(医学版),2020,39(04):550-556.
- [7]. Singh D, Kumar V, Kaur M. Classification of COVID-19 patients from chest CT images using multi-objective differential evolution-based convolutional neural networks[J]. European Journal of Clinical Microbiology & Infectious Diseases, 2020, 39(7): 1379-1389.
- [8]. Tingting Y, Junqian W, Lintai W, et al. Three-stage network for age estimation[J]. CAAI Transactions on Intelligence Technology, 2019, 4(2): 122-126.
- [9]. 新型冠状病毒感染的肺炎影像学诊断指南(2020 第一版)[J].医学新知,2020,30(01):22-34.
- [10]. Jin Y H, Cai L, Cheng Z S, et al. A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version)[J]. Military Medical Research, 2020, 7(1): 1-23.
- [11]. Pan Y, Guan H, Zhou S, et al. Initial CT findings and temporal changes in patients with the novel coronavirus pneumonia (2019-nCoV): a study of 63 patients in Wuhan, China[J]. European radiology, 2020, 30(6): 3306-3309.
- [12]. 徐秋贞,李勇刚.江苏省新型冠状病毒感染肺炎医学影像学检查、诊断与防控规范专家共识(2020 年第 1 版)[J].东南大学学报(医学版),2020,39(02):121-139.
- [13]. Xiong Y, Sun D, Liu Y, et al. Clinical and high-resolution CT features of the COVID-19 infection: comparison of the initial and follow-up changes[J]. Investigative radiology, 2020.
- [14]. Chang C C, Lin C J. LIBSVM: a library for support vector machines[J]. ACM transactions on

- intelligent systems and technology (TIST), 2011, 2(3): 1-27.
- [15].Brown M P S, Grundy W N, Lin D, et al. Knowledge-based analysis of microarray gene expression data by using support vector machines[J]. Proceedings of the National Academy of Sciences, 2000, 97(1): 262-267.
- [16].Bartsch Jr G, Mitra A P, Mitra S A, et al. Use of artificial intelligence and machine learning algorithms with gene expression profiling to predict recurrent nonmuscle invasive urothelial carcinoma of the bladder[J]. The Journal of urology, 2016, 195(2): 493-498.
- [17].高瞻,陈文,姜鹏,江家文,何强,赵亮.基于 XGBoost 的新冠肺炎智能检测系统实现[J].中国数字医学,2020,15(11):123-127.
- [18].Chen T, Guestrin C. Xgboost: A scalable tree boosting system[C]//Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. 2016: 785-794.
- [19].Barstugan M, Ozkaya U, Ozturk S. Coronavirus (covid-19) classification using ct images by machine learning methods[J]. arXiv preprint arXiv:2003.09424, 2020.
- [20].Sethy P K, Behera S K, Ratha P K, et al. Detection of coronavirus disease (COVID-19) based on deep features and support vector machine[J]. 2020.
- [21].Wang S, Kang B, Ma J, et al. A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19)[J]. European radiology, 2021: 1-9.
- [22].Xu X, Jiang X, Ma C, et al. A deep learning system to screen novel coronavirus disease 2019 pneumonia[J]. Engineering, 2020, 6(10): 1122-1129.
- [23].Li L, Qin L, Xu Z, et al. Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT[J]. Radiology, 2020.
- [24].Zheng C, Deng X, Fu Q, et al. Deep learning-based detection for COVID-19 from chest CT using weak label[J]. MedRxiv, 2020.
- [25].Das N N, Kumar N, Kaur M, et al. Automated deep transfer learning-based approach for detection of COVID-19 infection in chest X-rays[J]. Irbm, 2020.
- [26].Pathak Y, Shukla P K, Tiwari A, et al. Deep transfer learning based classification model for COVID-19 disease[J]. Irbm, 2020.
- [27].Jaiswal A, Gianchandani N, Singh D, et al. Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning[J]. Journal of Biomolecular Structure and Dynamics, 2021, 39(15): 5682-5689.
- [28].Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. arXiv preprint arXiv:2010.11929, 2020.
- [29].Gao X, Qian Y, Gao A. COVID-VIT: Classification of COVID-19 from CT chest images based on vision transformer models[J]. arXiv preprint arXiv:2107.01682, 2021.

- [30].江晶晶,高春,冯富娟,于晓辉,张久聪.新型冠状病毒肺炎无症状感染者的检测方法研究进展[J].海南医学院学报,2022,28(03):161-165.DOI:10.13210/j.cnki.jhmu.20211221.001.
- [31].Schaar M, Alaa A M, Floto A , et al. How artificial intelligence and machine learning can help healthcare systems respond to COVID-19[J]. Machine Learning, 2021, 110(1):1-14.
- [32].窦瑞欣. 深度学习算法在医学影像学中的应用及研究进展[J]. 中国医学计算机成像杂志, 2018, 024(005): 369-372.
- [33].Huang Y, Gao L, Preuhs A, et al. Field of view extension in computed tomography using deep learning prior[M]//Bildverarbeitung für die Medizin 2020. Springer Vieweg, Wiesbaden, 2020: 186-191.
- [34].Nowak S, Faron A, Luetkens J A, et al. Fully automated segmentation of connective tissue compartments for CT-based body composition analysis: a deep learning approach[J]. Investigative radiology, 2020, 55(6): 357-366.
- [35].Hubel D H, Wiesel T N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex[J]. The Journal of physiology, 1962, 160(1): 106.
- [36].Fukushima K, Miyake S. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition[M]//Competition and cooperation in neural nets. Springer, Berlin, Heidelberg, 1982: 267-285.
- [37].Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors[M].Neurocomputing: foundations of research. MIT Press, 1988:533-536.
- [38].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.
- [39].Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. Advances in neural information processing systems, 2012, 25.
- [40].Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. The journal of machine learning research, 2014, 15(1): 1929-1958.
- [41].K. Simonyan ,A. Zisserman . Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
- [42].Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1-9. Lin M, Chen Q, Yan S. Network in network[J]. arXiv preprint arXiv:1312.4400,2013.
- [43].He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
- [44].Li W, Wang Z, Wang Y, et al. Classification of high-spatial-resolution remote sensing scenes method using transfer learning and deep convolutional neural network[J]. IEEE Journal of

- Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13: 1986-1995.
- [45].Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4700-4708.
- [46].Weiss K, Khoshgoftaar T M, Wang D D. A survey of transfer learning[J]. Journal of Big data, 2016, 3(1): 1-40.
- [47].Pan S J, Yang Q. A survey on transfer learning[J]. IEEE Transactions on knowledge and data engineering, 2009, 22(10): 1345-1359.
- [48].Angelov P, Almeida Soares E. SARS-CoV-2 CT-scan dataset: A large dataset of real patients CT scans for SARS-CoV-2 identification[J]. MedRxiv, 2020.
- [49].王一超, 韦建文. 基于高性能计算平台的 TensorFlow 应用探索与实践[J]. Research & Exploration in Laboratory, 2017, 36(12).
- [50].Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.
- [51].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.
- [52].Valtemir de Alencar a Silva, Ig lbert Bittencourt, Jose carlos Maldonada. Automatic Question Classifiers: a Systematic Review[J]. IEEE Transactions on Learning Technologies. 2018, 10. PP(99): 1-1.
- [53].Ms. Yasin Kabir, Sanjay Madria. A Deep Learning Approach for Tweet Classification and Rescue Scheduling for Effective Disaster Management[J]. 2019, 8. arXiv: 1908.01456.
- [54].Zhao J, Zhang Y, He X, et al. Covid-ct-dataset: a ct scan dataset about covid-19[J]. arXiv preprint arXiv:2003.13865, 2020, 490.
- [55].Sun S, Wang S, Wei Y. A new ensemble deep learning approach for exchange rates forecasting and trading[J]. Advanced Engineering Informatics, 2020, 46: 101160.
- [56].Polikar R. Ensemble learning[M]//Ensemble machine learning. Springer, Boston, MA, 2012: 1-34.
- [57].Breiman L. Bagging predictors[J]. Machine learning, 1996, 24(2): 123-140.
- [58].周志华,陈世福.神经网络集成[J].计算机学报,2002(01):1-8.
- [59].Zhou B, Khosla A, Lapedriza A, et al. Learning deep features for discriminative localization[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 2921-2929.
- [60].Selvaraju R R, Cogswell M, Das A, et al. Grad-cam: Visual explanations from deep networks via gradient-based localization[C]//Proceedings of the IEEE international conference on

computer vision. 2017: 618-626.

- [61].Ghimire D. Comparative study on Python web frameworks: Flask and Django[J]. 2020.
- [62].白相辰. 基于 Django 框架的 Web 在线教育平台的设计与实现[D].北京交通大学,2019.DOI:10.26944/d.cnki.gbfju.2019.000099.
- [63].林国梁. 基于 Web 的移动终端取证管理系统的设计与实现[D].厦门大学,2018.
- [64].牛作东,李捍东.基于 Python 与 flask 工具搭建可高效开发的实用型 MVC 框架[J].计算机应用与软件,2019,36(07):21-25.