

基于深度残差神经网络的肝脏局部病灶分类

黄 林, 孙小淇*

青岛大学数学与统计学院, 山东 青岛

收稿日期: 2024年12月26日; 录用日期: 2025年1月18日; 发布日期: 2025年1月29日

摘 要

本研究提出了一种基于特征提取的深度神经网络(ResNet), 用于对局灶性肝脏病变的超声波图像进行分类。这些优势使其能够有效地从超声图像中提取肝脏损伤的迹象, 并实现准确的分类。基于包含血管瘤、脂肪肝、肝转移瘤、肝囊肿和正常肝脏的超声图像数据集, 该模型在实验验证中提供了出色的分类效果。在测试集上, 该模型的准确率达到了93.99%。本文将该模型与AlexNet和VGGNet模型进行了比较, 结果表明作者获得的模型分类效果更好, 在准确率、参数数量和学习效率方面都有显著优势, 并且具有很强的泛化能力。这项研究在病灶性肝脏病变的超声波图像分类任务中具有潜在的应用价值, 可为临床医生提供准确、快速的辅助诊断工具。

关键词

ResNet, AlexNet, VGGNet, CNN, 特征提取, 超声图像, 肝脏病变, 图像分类

Local Liver Lesion Classification Based on Deep Residual Neural Network

Lin Huang, Xiaoqi Sun*

School of Mathematics and Statistics, Qingdao University, Qingdao Shandong

Received: Dec. 26th, 2024; accepted: Jan. 18th, 2025; published: Jan. 29th, 2025

Abstract

This study proposes a feature extraction-based deep neural network (ResNet) for classifying ultrasound images of focal liver lesions. These advantages enable it to effectively extract signs of liver damage from ultrasound images and achieve accurate classification. Based on an ultrasound image dataset containing hemangioma, fatty liver, liver metastases, liver cysts, and normal liver, the model provided excellent classification results in experimental validation. On the test set, the model

*通讯作者。

文章引用: 黄林, 孙小淇. 基于深度残差神经网络的肝脏局部病灶分类[J]. 应用数学进展, 2025, 14(1): 352-359.
DOI: 10.12677/aam.2025.141035

achieved an accuracy of 93.99%. This article compares this model with AlexNet and VGGNet models. The results show that the model obtained by the author has better classification results, has significant advantages in accuracy, number of parameters, and learning efficiency, and has strong generalization ability. This study has potential application value in the task of ultrasonic image classification of focal liver lesions and can provide clinicians with an accurate and rapid auxiliary diagnostic tool.

Keywords

ResNet, AlexNet, VGGNet, CNN, Feature Extraction, Ultrasound Image, Liver Lesions, Image Classification

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

肝癌一直是全球关注的疾病[1], 根据世界卫生组织国际癌症研究机构(IARC/WHO)的最新统计, 2020年将有 90.57 万人被诊断为肝癌, 83.02 万人死于肝癌, 成为全球 46 个国家癌症死亡的三大原因之一; 预计 2040 年全球肝癌新诊断人数和死亡人数将增加 55% [1]。肝细胞癌死亡率极高, 病情发展较快, 发病率逐年上升。肝细胞癌是全球第四大常见恶性肿瘤, 每年约有 84.1 万新发病例和 78.2 万死亡病例, 在全球大部分地区, 男性的发病率和死亡率是女性的 2 到 3 倍。常见的肝脏病变包括脂肪肝、酒精性肝损伤、急性病毒性肝炎、心源性肝肿大、肝硬化、肝损伤、肝囊肿、肝血管瘤、肝脓肿和原发性肝癌等。因此, 肝脏病变的早期筛查和诊断是非常必要和具有挑战性的, 全世界的研究人员和临床医生都为此付出了努力, 做了大量的研究和实践。

随着超声波换能器、图像处理和计算机技术的不断进步, 超声波图像的分辨率和质量有了显著提高, 使医生能够更准确地诊断和评估病情。肝脏超声图像是一种常用的医学成像技术, 通过超声波在人体组织中的传播和反射特性, 对肝脏结构和病变进行无创检查和评估。这种成像技术可以清晰地显示肝脏的大小、形状、质地和血流情况, 帮助医生诊断和监测各种肝脏疾病, 如脂肪肝、肝囊肿、肝硬化和肝肿瘤等。由于肝脏超声波检查无辐射、操作简便、价格低廉、可实时观察, 因此已被广泛应用于临床。

在人工智能和机器学习技术不断发展的基础上, 人工智能技术越来越多地应用于医学影像的诊断和分析, 并取得了良好的发展。传统机器学习在肝脏超声图像分类中的应用已经取得了显著进展。肝脏疾病, 如肝炎、肝硬化和肝癌, 早期诊断至关重要, 而超声成像作为一种无创且经济的诊断工具, 广泛应用于临床实践。传统机器学习方法通过提取图像特征, 辅助医生进行肝脏病变的识别与分类。但是传统的机器学习方法依赖手工提取特征, 受限于设计, 训练较简单, 计算需求低, 所以精度通常较低, 泛化能力较差, 适用于低资源环境。随着卷积神经网络的发展, 不断超越传统机器学习图像分类, 在超声图像分类领域表现出极大的优越性。卷积神经网络(CNN)是一种广泛应用于图像处理领域的深度学习模型。在医学图像分析领域, 卷积神经网络已成为一个非常有前景的研究方向。通过将医学图像的分类、分割和诊断等任务自动化和标准化, CNN 可以极大地帮助临床医学的发展和完善。许多机器学习方法已被用于诊断各类肝病。

Wang [2]等人新开发的深度学习放射组学(DLRE)前体层摄影术在评估肝纤维化分期方面显著优于其他方法。Yasaka K. [3]等人证明, 使用卷积神经网络(CNN)深度学习方法在动态对比增强计算机断层扫描(CT)中区肝肿块诊断性能优于传统方法。Wu [4]等人将 3 层 DBN 应用于从造影剂增强超声(CEUS)视频序列中提取的时间强度曲线(TIC), 实现了局灶性肝脏病变的良恶性分类, 获得了 86.36%的最高准确率。Sato M. [5]通过对 B 型超声图像以及病历和血液数据中的其他信息进行联合训练, 提出了一种用于良恶性肿瘤识别的人工智能模型, 其准确率为 96.3%, AUROC 为 0.994。Tarek M. [6]等人提出了一种基于深度学习技术的堆叠稀疏自动编码器特征表示, 其整体分类准确率达到 97.2%, 优于(多重支持向量机、K-最近邻和 Naive Bayes)等技术。然而, 这些机器学习方法大多使用浅层架构。鉴于深度神经网络强大的特征和学习能力, 基于特征提取的深度卷积神经网络用于肝脏超声图像分类是一个值得深入研究的课题。本文提出了一种基于特征提取的深度残差神经网络(ResNet), 用于肝脏病变的超声图像分类, 提高了肝脏病灶分类的准确性。

2. 材料和方法

我们工作中使用的数据由俄罗斯圣彼得堡卫生部 N.N. Petrov 肿瘤学 NMRC 的专家提供。这些数据包含该研究中心对患者进行的超声检查的匿名视频记录。对于每个视频记录, 超声诊断医生都会根据以下超声参数创建一个文本文档来描述肝脏的状态: 评估实质的结构和回声, 评估器官的血管结构, 识别弥漫性和局部性变化的迹象, 以及局灶性病变的存在, 并根据普遍认可的标准对变化进行完整描述。由于超声图像的标准方向和注释对于清晰传输静态图像中呈现的信息非常重要。静态图像中呈现的信息包括经腹部图像, 其中换能器沿纵向(矢状面)定向; 以及腹部平面图像, 其中换能器位于横向(轴向)。

所涉病例的视频包含两个方向(矢状面扫描、肋下扫描)。回顾性收集了 2022 年 8 月至 2023 年 12 月期间共 92 个超声视频, 包括(17 个血管瘤视频、17 个脂肪肝视频、16 个肝转移视频、25 个肝囊肿视频和 17 个正常肝脏视频)。

由于采集的肝脏超声是利用高频声波生成的实时视频。为此, 应用图像处理技术对肝脏病变的常规超声图像进行分割, 作为训练和测试分类算法的基础, 以进一步提取特征和病变分类。共获得 3368 张肝脏超声图像, 包括(651 张血管瘤图像、608 张脂肪肝图像、536 张转移瘤图像、1085 张肝囊肿超声图像和 488 张正常肝脏图像)。

2.1. 图像降噪

超声图像的成像机制以及成像设备之间的差异导致超声图像中含有大量的斑点噪声, 且图像对比度和分辨率都较低, 边界不清晰, 因此为了提高直接分割的效果, 需要预先对获取的超声图像进行滤波。本文采用高斯滤波器对切分后的图像进行降噪。高斯滤波器是一种线性平滑滤波器, 适用于消除高斯噪声, 在图像去噪中应用广泛。其原理是对整幅图像的像素值进行加权, 对于每个像素, 其值都是由其自身的值和邻域内的其他像素值加权平均后得到的。二维高斯公式如下:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

本文采用[5 × 5]卷积对图像中每个像素进行扫描, 用卷积确定的邻域内像素的加权平均灰度值代替模板中心像素点的值。图 1 为降噪前后对比, 经高斯滤波处理后的图像与未降噪图像相比, 图像质量和边缘锐度均有所提高, 说明高斯滤波在局部性肝脏病变超声图像处理中具有良好的效果, 有助于提高后续深度卷积神经网络模型的性能和稳定性。

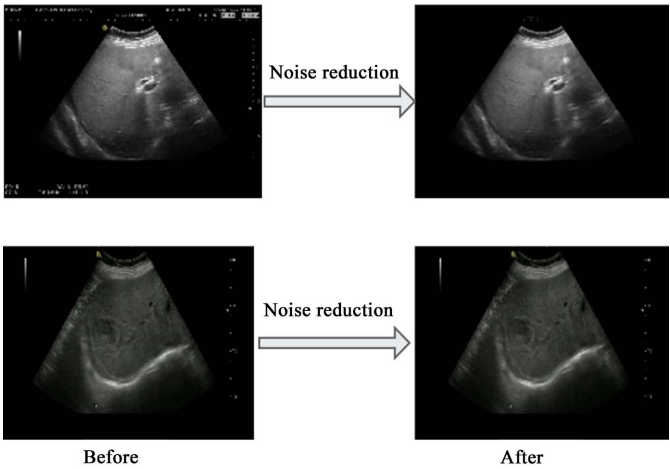


Figure 1. Before and after Gaussian filtering noise reduction comparison diagram
图 1. 高斯滤波降噪前后对比图

2.2. 特征提取

肝脏超声图像中病变的感兴趣区域(RIO)通常呈现以下三种情况(图 2)。

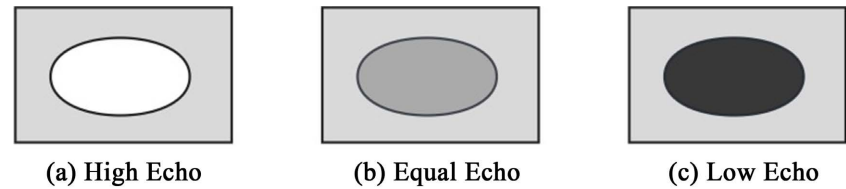


Figure 2. Classification of lesion regions of interest in liver ultrasound images
图 2. 肝脏超声图像中病变感兴趣区域分类

- (a) 低回声 - 肝脏中的白色物质，当存在于实体器官中时，通常代表器官
- (b) 等回声 - 肝脏呈灰色，通常是正常器官或腔体的标志。
- (c) 低回声 - 肝脏呈黑色：这主要见于局部分布的脂肪组织或良性

为此，我们需要标记病变的感兴趣区域，既可以手动确定病变的感兴趣区域(ROI)，也可以通过计算机提取。在本文中，为提高效率和客观性，采用 U-Net 对图像中 ROI 进行标记。U-Net 是一种卷积神经网络(CNN)，最早用于医学图像分割，在肝脏超声图像分割中表现出色。如下将采集到的肝脏病变超声原始图像通过 U-Net 标注，得到病变感兴趣区域如图 3：

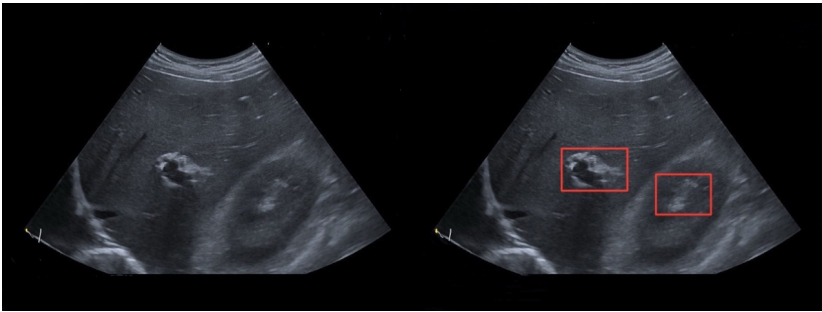


Figure 3. Before and after lesion labelling
图 3. 病灶标记前后对比图

3. 分类方法与结果

卷积神经网络(CNN)能够有效地从图像中提取特征,具备强大的表征学习能力。浅层卷积操作仅能识别图像的基本特征,如轮廓和线条,而随着卷积层数和卷积核尺寸的增加,网络能够捕捉到更加复杂的特征信息。如图 4 所示,为利用卷积神经网络进行肝脏超声病变分类的示意图,即将肝脏病变的超声图像输入到卷积神经网络中,计算出有判别力的深度卷积特征,再输入到分类器中,输出病变的类别属性。

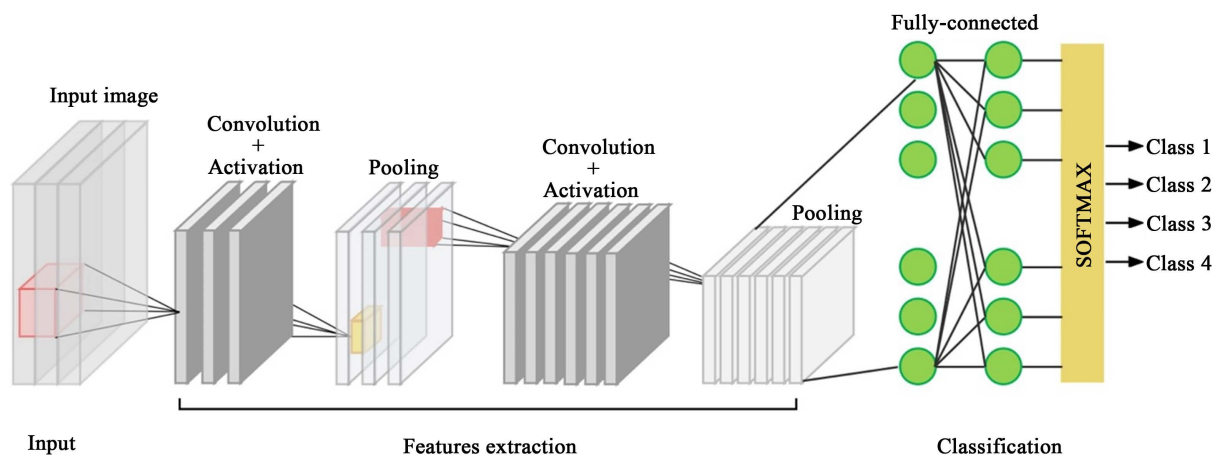


Figure 4. Schematic of convolutional neural network image classification model

图 4. 卷积神经网络图像分类模型示意图

同时,为了提高神经网络的泛化性能,出现了各种不同结构的神经网络,例如 LetNet、AlexNet、VGGNet、GoogLeNet 等[7]。其中, AlexNet 首次引入了 ReLU 激活函数,取代了传统的 Sigmoid 或 Tanh 激活函数,ReLU 的使用加速了训练过程,解决了梯度消失的问题。AlexNet 有 8 层需要训练,不包含局部响应归一化(LRN)层和池化层,前 5 层为卷积层,后 3 层为全连接层,最后的全连接层被 SoftMax 层取代。Simonian 等人[8]设计的卷积神经网络用于视频。较小的卷积核可以增加非线性网络的深度,使网络能够学习更复杂的特征。通常,模型的性能应该随着网络的加深而提高。然而在实际应用中,当网络加深到一定程度时,模型的性能反而开始下降。这是因为在深度网络中,信号需要经过多个非线性变换才能到达后面的层,这会导致前面层的信息丢失或衰减,使模型难以训练。为了解决这个问题,ResNet 中引入了残差块的结构[9]。

如图 5 所示,是普通卷积神经网络和加入了残差连接的卷积神经网络的结构对比。残差块包含一个跨层的 Shortcut Connection,可以让信息直接传递到后面的层,避免信号丢失和衰减。具体来说,其他残差块的输入会添加到输出中形成残差,然后再添加到输出中。这样可以保证深度网络中信息的完整性和稳定性,使网络更易于训练。它在图像分类、目标检测、语义分割等多个领域都取得了不错的效果,成为深度学习领域的重要模型之一。因此,本文采用 ResNet_18 作为分类任务的基础模型,如图 6 所示。ResNet-18 网络由 17 个卷积层和 1 个全连接层组成,每个卷积层采用 $[3 \times 3]$ 卷积核和 ReLU 激活函数来提取图像的局部特征。输入图像首先进入网络,经过多个卷积层和池化层进行特征提取。网络包含 8 个残差块,每个块有两个分支:主路径和残差路径。主路径通过一系列卷积层提取特征,残差路径通过跳跃连接将输入直接加到主路径的输出上。这样的结构有助于缓解梯度消失和爆炸问题,简化网络学习残差函数。在网络的末端,使用全局平均池化层将特征图转化为一维向量,减少参数量并保留空间信息。最终,池化后的输出连接到全连接层。在训练时,ResNet-18 使用基于随机梯度下降(SGD) [10]的反向传播方法,通过最小化交叉熵损失函数来优化网络参数。

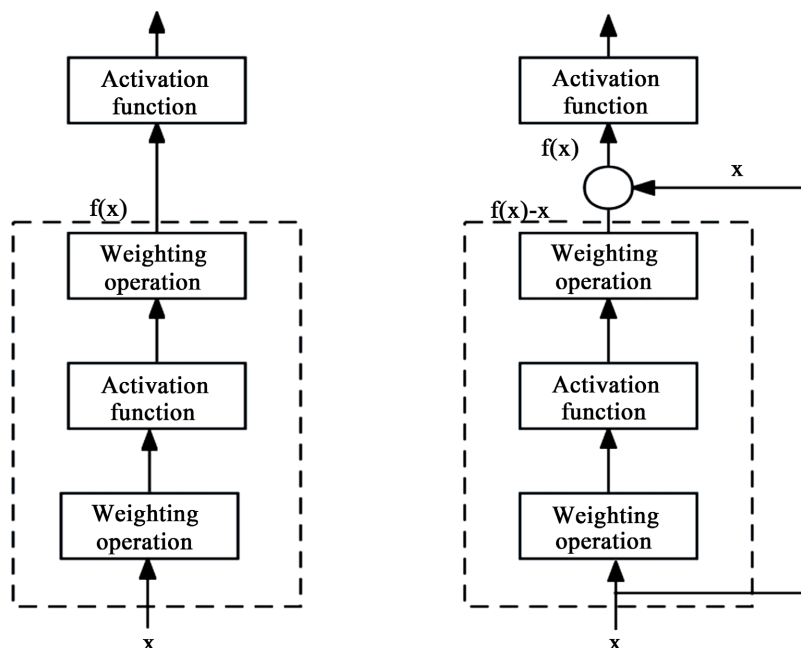


Figure 5. Ordinary convolutional neural network and convolutional neural network with added residual connections
图 5. 普通卷积神经网络与加入残差连接的卷积神经网络

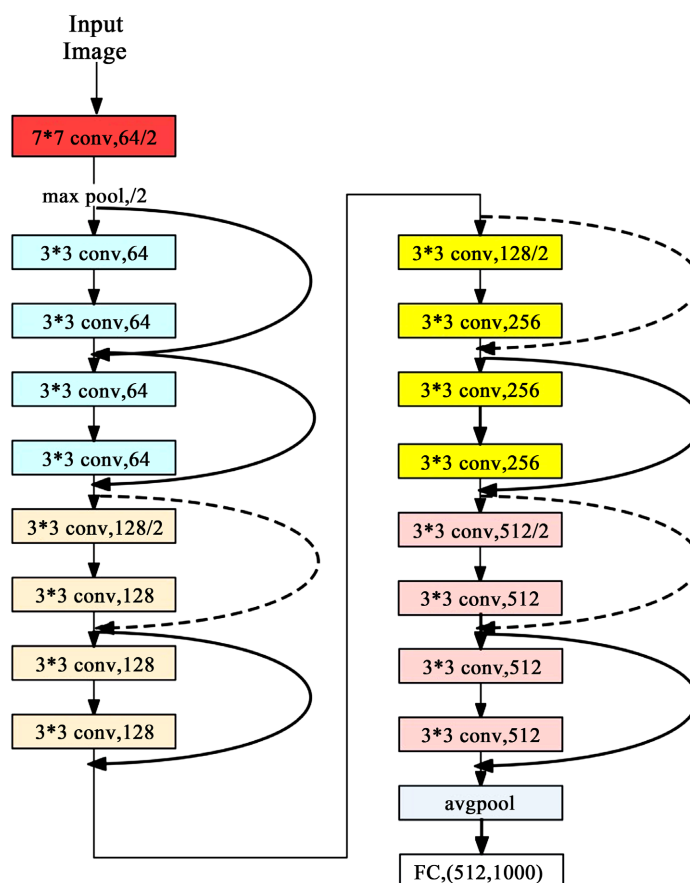


Figure 6. Deep residual neural network ResNet_18
图 6. 深度残差神经网络 ResNet_18

为了保证模型的客观可信度, 将数据随机分为训练集(80%)和测试集(20%)两部分, 其中训练集 2694 个, 测试集 674 个, 训练过程中考虑到计算机的性能, Epoch 设置为 100。训练过程中考虑到计算机的性能, Epoch 设置为 100, 初始学习率设置为 0.0001。训练和测试的准确率和损失率计算如下图 7 所示。

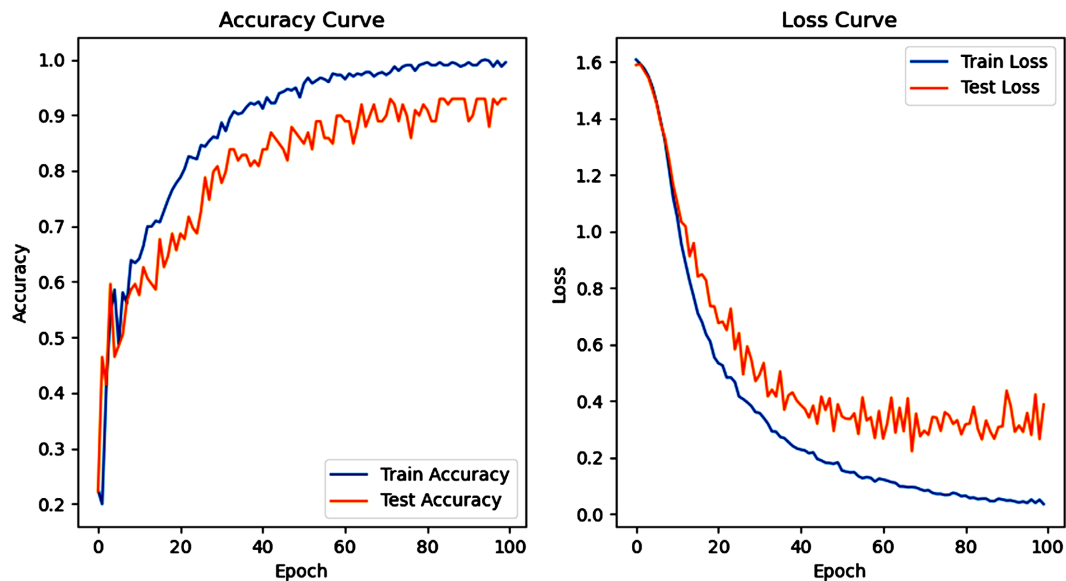


Figure 7. Model training accuracy
图 7. 模型准确率

此外, 我们将分类结果与 AlexNet 和 VGGNet 两个模型进行了比较, 实验结果如表 1 所示。实验结果表明, 我们的模型在测试集上的准确率为 93.99%, 平均准确率为 92.67%, 说明我们的模型能够有效区分不同类型的肝脏病变, 分类准确率较高。进一步分析发现, 与 AlexNet 和 VGGNet 两个模型相比, ResNet_18 模型收敛速度最快, 在训练集上的训练效果最好, 表现出良好的分类性能。通过实验, 基于特征提取的 ResNet_18 训练出了不错的效果, 证明了基于 CNN 对肝脏超声图像进行分类的可行性。

Table 1. Comparison of training effectiveness of different network models
表 1. 不同网络模型的训练效能比较

网络模型	训练集		测试集		
	最高准确率(%)	平均准确率(%)	准确率(%)	灵敏度(%)	特异度(%)
ALexNet	81.96	77.95	74.2	69.37	72.11
VGG	87.08	85.6	83.96	73.85	70.43
ResNet	93.99	92.67	90.01	85.44	88.27

然而, 我们的研究也存在一些局限性。首先, 我们的数据集相对较小, 可能会限制模型的泛化能力。其次, 我们的模型在某些病变类型的分类上可能存在一些错误, 这可能与数据集的不平衡或特定病变的特征提取有关。因此, 未来的研究可以考虑扩大数据集大小, 改进模型结构, 进一步优化训练策略, 以提高模型的性能和鲁棒性。

4. 总结

CNN 作为深度学习网络的代表, 在医学图像分类中得到了广泛应用。本文通过对三种不同深度、结构和功能的 CNN 模型进行分类训练和测试, 结果显示, ResNet-18 模型的分类效果最佳, 准确率较高。研究表明, 深度残差神经网络在医学图像分类任务中, 特别是在肝病诊断方面具有潜在的应用价值。然而, 研究也存在一些局限性, 特征提取和网络设计仍有改进空间, 未来可进一步提升分类效率和模型的泛化能力。在未来的研究中, 将增加数据样本量, 同步微调网络参数, 减少过拟合和欠拟合现象, 提高肝脏病变超声图像分类的准确率和稳定性。同时, 我们也将考虑将该技术应用于临床, 为医生提供更加准确、快速的辅助诊断手段。

基金项目

山东省自然科学基金省、补助/拨款编号: ZR2024MA097。

参考文献

- [1] Rumgay, H., Arnold, M., Ferlay, J., Lesi, O., Cabaasag, C.J., Vignat, J., *et al.* (2022) Global Burden of Primary Liver Cancer in 2020 and Predictions to 2040. *Journal of Hepatology*, **77**, 1598-1606. <https://doi.org/10.1016/j.jhep.2022.08.021>
- [2] Wang, K., Lu, X., Zhou, H., Gao, Y., Zheng, J., Tong, M., *et al.* (2018) Deep Learning Radiomics of Shear Wave Elastography Significantly Improved Diagnostic Performance for Assessing Liver Fibrosis in Chronic Hepatitis B: A Prospective Multicentre Study. *Gut*, **68**, 729-741. <https://doi.org/10.1136/gutjnl-2018-316204>
- [3] Yasaka, K., Akai, H., Abe, O. and Kiryu, S. (2018) Deep Learning with Convolutional Neural Network for Differentiation of Liver Masses at Dynamic Contrast-Enhanced CT: A Preliminary Study. *Radiology*, **286**, 887-896. <https://doi.org/10.1148/radiol.2017170706>
- [4] Wu, K., Chen, X. and Ding, M. (2014) Deep Learning Based Classification of Focal Liver Lesions with Contrast-Enhanced Ultrasound. *Optik*, **125**, 4057-4063. <https://doi.org/10.1016/j.jileo.2014.01.114>
- [5] Sato, M., Kobayashi, T., Soroida, Y., Tanaka, T., Nakatsuka, T., Nakagawa, H., *et al.* (2022) Development of Novel Deep Multimodal Representation Learning-based Model for the Differentiation of Liver Tumors on B-mode Ultrasound Images. *Journal of Gastroenterology and Hepatology*, **37**, 678-684. <https://doi.org/10.1111/jgh.15763>
- [6] Hassan, T.M., Elmogy, M. and Sallam, E. (2017) Diagnosis of Focal Liver Diseases Based on Deep Learning Technique for Ultrasound Images. *Arabian Journal for Science and Engineering*, **42**, 3127-3140. <https://doi.org/10.1007/s13369-016-2387-9>
- [7] Dogo, E.M., Afolabi, O.J., Nwulu, N.I., Twala, B. and Aigbavboa, C.O. (2018) A Comparative Analysis of Gradient Descent-Based Optimization Algorithms on Convolutional Neural Networks. 2018 *International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*, Belgaum, 21-22 December 2018, 92-99. <https://doi.org/10.1109/ctems.2018.8769211>
- [8] Simonyan, K. and Zisserman, A. (2014) Two-Stream Convolutional Networks for Action Recognition in Videos. *NIPS'14: Proceedings of the 28th International Conference on Neural Information Processing Systems*, Volume 1, 568-576.
- [9] Lee, C. and Le, T. (2021) Identifying Faults of Rolling Element Based on Persistence Spectrum and Convolutional Neural Network with Resnet Structure. *IEEE Access*, **9**, 78241-78252. <https://doi.org/10.1109/access.2021.3083646>
- [10] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2017) Imagenet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, **60**, 84-90. <https://doi.org/10.1145/3065386>