

DOI: 10.20271/j.cnki.1673-9833.2026.3008

基于改进 EfficientNetV2 的 急性淋巴细胞白血病分类方法

朱文球, 朱 锐, 邓 立

(湖南工业大学 计算机与人工智能学院, 湖南 株洲 412007)

摘要: 针对急性淋巴细胞白血病图像类别分布不均衡、背景信息复杂等特点, 以及人工诊断耗时较长且易受主观因素影响的挑战, 提出一种 EfficientNet-DSP 白血病分类方法。该方法通过图像增强技术和动态随机失活块以提升模型的泛化能力, 融合残差置换注意力机制以增强模型细节特征的提取能力; 并提出利用 Dy-ODConv 动态卷积学习各个维度的信息, 动态调整卷积核权重, 在降低参数量的同时提升分类准确率。此外, 改进了算法的损失函数, 增强模型处理复杂背景图像时的分类能力。最后, 在 Blood Cells Cancer 数据集上进行实验, 结果显示, EfficientNet-DSP 取得了 98.46% 的图像分类准确率, 相比原始 EfficientNetV2 模型提升了 2.54%, 相较其它算法的最优值提升了 3.61%, 可知所提方法有效提高了对急性淋巴细胞白血病图像的诊断准确率, 可作为医师诊断的参考依据。

关键词: 残差置换注意力; 动态卷积; 动态随机失活块; 损失函数

中图分类号: TP18; R733.71 **文献标志码:** A **文章编号:** 1673-9833(2026)03-0055-08

引文格式: 朱文球, 朱 锐, 邓 立. 基于改进 EfficientNetV2 的急性淋巴细胞白血病分类方法 [J]. 湖南工业大学学报, 2026, 40(3): 55-62.

Classification Method for Acute Lymphoblastic Leukemia Based on Improved EfficientNetV2

ZHU Wenqiu, ZHU Kun, DENG Li

(School of Computer Science and Artificial Intelligence, Hunan University of Technology, Zhuzhou Hunan 412007, China)

Abstract: In view of the uneven distribution of image categories and complex background information in acute lymphoblastic leukemia, as well as the challenges of time-consuming manual diagnosis and susceptibility to subjective factors, an EfficientNet-DSP leukemia classification method has thus been proposed. The generalization ability of the model can be enhanced by the proposed method through image enhancement techniques and dynamic random deactivation blocks, with residual permutation attention mechanism integrated to enhance its ability to extract detailed features. It is proposed to use Dy-ODConv dynamic convolution to learn information from various dimensions, and dynamically adjust the weights of convolution kernels, thus improving classification accuracy while reducing the number of parameters. In addition, the loss function of the algorithm has been improved to enhance the classification ability of the model when processing complex background images. Finally, experiments are conducted on the Blood Cells Cancer dataset, with the results showing that EfficientNet-DSP achieves an image classification accuracy of 98.46%, an improvement of 2.54% compared to the original EfficientNetV2 model, and an improvement of 3.61% compared to the optimal values of other algorithms. It can be concluded that the proposed method effectively improves the diagnostic accuracy of acute lymphoblastic leukemia images, which makes it a reference for physician diagnosis.

收稿日期: 2024-12-10

基金项目: 湖南省教育厅科学研究基金资助项目 (23A0423)

作者简介: 朱文球, 男, 湖南工业大学教授, 主要研究方向为人工智能, 计算机视觉, 大数据技术,

E-mail: wenqiu_zhu@126.com

Keywords: shuffle attention of residual; dynamic convolution (Dy-ODConv); DyDropBlock; loss function

1 研究背景

白血病是一种源于造血干细胞的恶性克隆性疾病，可细分为4种主要类型：慢性髓系白血病、慢性淋巴细胞白血病、急性髓系白血病和急性淋巴细胞白血病^[1]。其中急性淋巴细胞白血病（Acute Lymphoblastic Leukemia, ALL）的主要发病人群是儿童和青少年^[2]，其发病率较高，具有较强的致死性。急性淋巴细胞白血病有多种治疗方案，包括化疗、靶向治疗和干细胞移植等，而这些治疗方案的最终成效依赖于早期的精确诊断。传统诊断 ALL 的方法主要依赖病理医师对血细胞图像的形态学检查，尤其是通过显微镜对外周血或骨髓样本进行观察，这一过程不仅依赖医师经验，还受到图像质量、细胞重叠、背景噪声等因素的影响，导致诊断过程耗时且易产生误差，从而影响诊断的准确性^[3]。

通过训练深度神经网络模型，能够有效识别与白血病相关的细胞形态特征，在大规模样本中快速筛选出异常病例，显著提高诊断的准确率和效率^[4]。A. Hosseini 等^[5]在急性淋巴细胞白血病分类任务过程中，通过聚类和掩码处理后分割背景图像，有效降低了图像背景对训练过程的影响。J. Prellberg 等^[6]采用预训练的 ResNeXt50 模型，结合迁移学习策略，在测试集上取得了 88.91% 的 F_1 加权分数。Pan Y. S. 等^[7]使用预训练的 ResNet50^[8]、ResNet-101 和 ResNet-152 模型，通过微调模型，生成特征映射，并基于最相似邻居的加权多数投票矫正标签，模型最终达到了 91.04% 的 F_1 加权分数。本研究的 4 类细胞间具有较高相似性，更加注重细节部分的区分，但使用预训练对白血病的识别效果有待进一步验证。

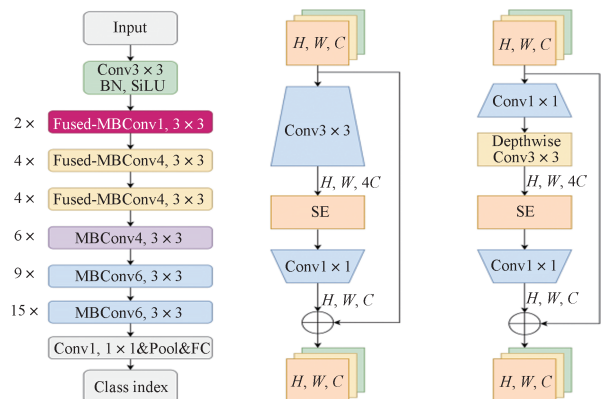
注意力机制用于捕捉图像的特征信息，通过不同的关注机制获取图像不同特征维度的信息。邵宇飞等^[9]基于 ResNeXt50 模型，改进了空洞金字塔特征提取方法，并引入改进的置换注意力（shuffle attention, SA），以提升急性淋巴细胞白血病分类性能。孙天宇等^[10]提出了一种基于稀疏注意力模块的 Vision Transformer 血细胞图像识别方法，该方法能有效筛选出图像中的辨识性区域，提升模型的细粒度特征表达能力。李家成等^[4]设计了一种多尺度空间注意力和通道注意力的卷积模块，以提取不同类别图像中的细颗粒特征信息。这些方法均取得了一定的成果，但在复杂背景下对关键特征的提取能力仍显不

足，在数据集类别不平衡的情况下，如何有效学习少数类别细胞的特征信息仍有待深入研究。

针对以上问题，本研究结合细胞图像的特点，提出了一种基于 EfficientNetV2^[11] 的急性淋巴细胞白血病分类方法 EfficientNet-DSP，通过图像增强和图像分割技术增加数据集的丰富性，降低背景噪声干扰，引入动态随机失活块 DyDropBlock 以降低模型过拟合风险。同时，设计一种残差置换注意力机制以增强模型对图像特征的提取能力；采用动态卷积降低参数的同时提升分类准确率。此外，选用多项式焦点损失函数 PolyFocal Loss 以提升模型在数据集类别不平衡、难分样本情景下的表现，提高模型分类效率。

2 EfficientNetV2 简介

EfficientNetV2 是 EfficientNet 的改进版本，其采用训练感知神经架构搜索和缩放相结合的方法，显著提升了训练速度和参数效率，与传统的 CNN 模型相比，EfficientNetV2 能在较短的时间内完成训练，在保证较高分类精度的同时显著降低计算资源消耗^[11]。本文基于急性淋巴细胞白血病图像分类的需求，在保证高准确率和推理速度的前提下，选用 EfficientNetV2-s 作为基础模型，其结构见图 1。



a) EfficientNetV2 b) Fused-MBConv 模块 c) MBConv 模块
图 1 EfficientNetV2 网络结构

Fig. 1 EfficientNetV2 network structure

本文提出的 EfficientNet-DSP 模型主要由改进后的 Fused-MBConv、MBConv 和 Dy-MBConv 组成，具体见图 2。MBConv 模块是 EfficientNetV2 模型的主要构成模块，采用深度可分离卷积，融合残差置换注意力（shuffle attention of residual, SAR）模块，充分结合通道注意力和空间注意力的优点，提升了模型特征提取能力。此外，模型引入了动态随机失活块，

通过屏蔽连续区域信息, 有效避免了 DropOut 正则化效果不明显的问题, 进而减少模型过拟合风险。Fused-MBConv 是在 MBConv 结构的基础上, 采用普通 3×3 卷积替代原有的 1×1 卷积层和深度可分离 3×3 卷积层, 这提升了通道扩展和空间特征提取能

力, 放在浅层网络中可提升训练速度。Dy-MBConv 模块在 MBConv 结构基础上融合了 Dy-ODconv 模块, 根据输入特征图自适应调整卷积核, 用更少的卷积核获得更优性能, 使模型在训练过程中更高效, 提升模型整体性能和分类准确率。

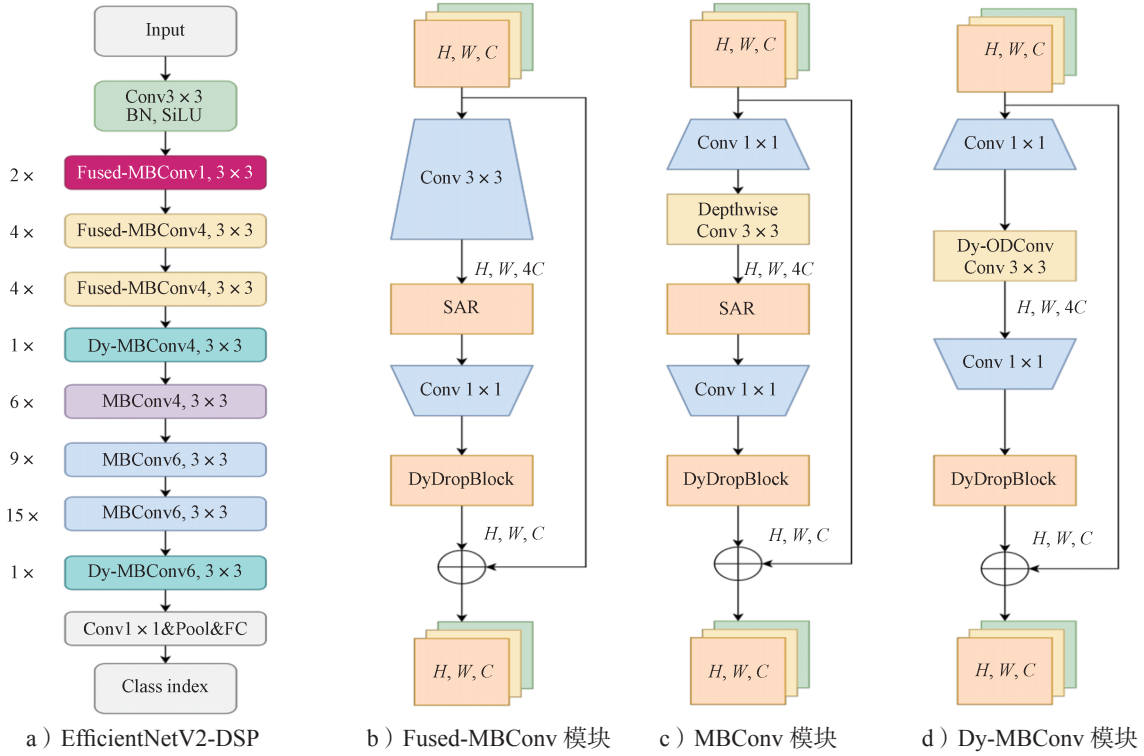


图 2 EfficientNet-DSP 网络结构

Fig. 2 EfficientNet-DSP network structure

3 模型改进

3.1 数据预处理

若训练数据周围太多无关背景, 则会成为干扰学习的噪声, 通过分割背景, 能够将染色明显的淋巴细胞与红细胞、血小板等其它血液细胞分开, 降低噪声干扰。具体步骤如下: 将 RGB 颜色空间转换到 LAB 颜色空间^[8], 如式 (1) (2) 所示, 再将每个像素调整到相应的聚类并分配其最近聚类中心的值, 聚类后的像素值塑形为原始图像尺寸, 并对聚类后的图像进行二值化处理, 将处理后的掩码图像跟原始图像按位与操作, 保留需要学习的部分, 如图 3 所示。

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.231 & 0.070 & 0.072 \\ 0.113 & 0.238 & 0.041 \\ 0.000 & 0.027 & 0.400 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \quad (1)$$

$$\begin{cases} L = 116f(Y/Y_n) - 16, L \in [0, 100]; \\ a = 500[f(X/X_n) - f(Y/Y_n)], a \in [-128, 127]; \\ b = 200[f(Y/Y_n) - f(Z/Z_n)], b \in [-128, 127]. \end{cases} \quad (2)$$

式 (1) (2) 中: X 、 Y 、 Z 为 XYZ 彩色空间的 3

个通道, 由 RGB 颜色空间变换而来; X_n 、 Y_n 、 Z_n 为 CIE 标准定义白色的灰度值, 通常取值 255;

$$f(x) = \begin{cases} x^{1/3}, & x > 0.008 856; \\ (1/3)(29/6)^2 x + 4/29, & \text{otherwise.} \end{cases}$$

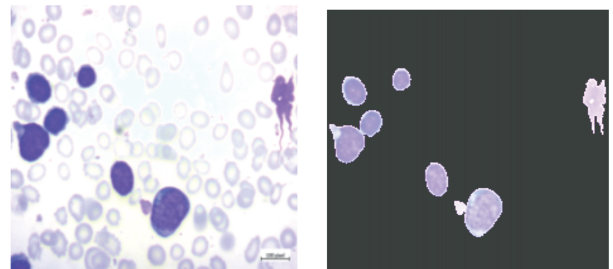


图 3 分割前后的图像对比

Fig. 3 Comparison before and after segmentation

当数据集的数量充足且分布均匀时, 卷积神经网络能充分发挥其特征提取能力。对 Blood Cells Cancer (ALL) 公开数据集进行数据增强, 以获得多样性的数据图像。本实验采用了随机旋转、水平翻转和垂直翻转、添加高斯噪声以及去除背景噪声等方式

进行数据增强机制。

3.2 残差置换注意力机制

注意力机制用于捕捉图像特征信息，通过不同关注机制获取图像不同特征维度上的信息。通道注意力机制 (squeeze-and-excitation, SE)^[12] 将高维全局特征图挤压降维成低维特征向量，然而降维操作削弱了通道间关系的提取能力，带来低效率问题。

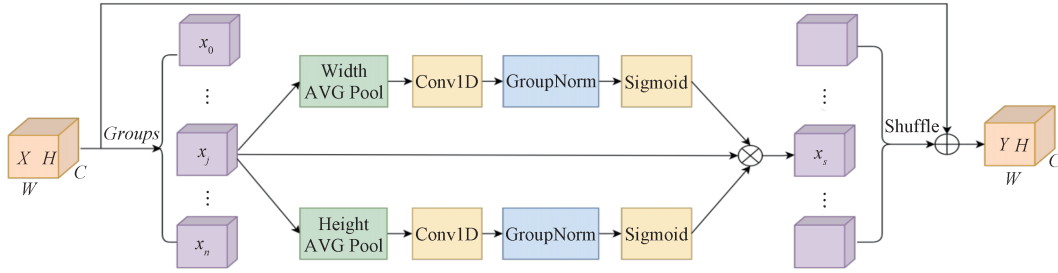


图4 残差置换注意力机制结构示意图

Fig. 4 SAR schematic diagram

首先，将输入特征图 X 在通道维度上重塑为 $Groups$ 个组。每个分组在宽度维度和高度维度进行全局平均池化，得到高度 h 和宽度 w 方向上的特征图 Z_c^h 和 Z_c^w ，如式 (3) (4) 所示：

$$Z_c^h(h) = \frac{1}{H} \sum_{0 \leq i < H} x_c(h, i), \quad (3)$$

$$Z_c^w(w) = \frac{1}{W} \sum_{0 \leq j < W} x_c(j, w). \quad (4)$$

式 (3) (4) 中： c 为通道数； H 和 W 分别为输入特征图的宽度和高度。

相比二维卷积，一维卷积更适合处理顺序信号，且更轻、更快^[14]。用卷积核大小为 7 的一维卷积 F_h 和 F_w 对池化后的两个通道输出特征图进行卷积，使用组归一化 (group normalization, GN) 处理增强的位置信息，采用非线性激活函数 Sigmoid 将它们转换至 $[0, 1]$ 范围内，得到高度和宽度方向上的注意力权重 y^h 和 y^w ，其表达式如下：

$$y^h = \sigma(G_n(F_h(z_h))), \quad (5)$$

$$y^w = \sigma(G_n(F_w(z_w))). \quad (6)$$

式 (5) (6) 中： σ 为 Sigmoid 激活函数； G_n 为组归一化； F_h 、 F_w 为高度维度和宽度维度上的一维卷积。

将当前分组特征图用高度和宽度的空间信息作为注意力权重进行加权编码，输出表达式见式 (7)。

$$y_c = x_c \times y_c^h \times y_c^w. \quad (7)$$

所有组被处理后，会被汇集起来，通过打乱操作对输出结果进行通道排列，使得相互关联的通道合在一起，实现不同组特征之间的信息融合。再将输出特征与原始图像特征融合，减少卷积层堆叠带来的信息

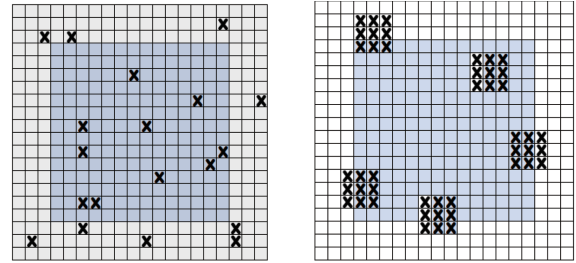
置换注意力 SA^[13] 结合了空间注意力机制和通道注意力机制，将输入特征图划分为多组，并通过随机单元将通道注意力和空间注意力整合到每个组中。

虽然分组之后会带来效率的提升，然而，堆叠多个卷积层后，容易导致长距离信息传递困难，基于上述问题，本文设计了一种融合全局特征的注意力机制，命名为残差置换注意力机制，其结构如图 4 所示。

传递困难的影响，从而提升模型精度。

3.3 动态随机失活块

DropOut 方法随机将部分神经元的输出置零来防止过拟合，然而特征单元还能从连接关系中学习，正则化效果不明显。而随机失活块 DropBlock^[15] 能随机屏蔽一整块连续的空间区域，抑制特征单元之间的相关性，有效去除某些冗余语义信息，减少过拟合风险，DropOut 和 DropBlock 的效果图如图 5 所示。



a) DropOut 效果图

b) DropBlock 效果图

图5 DropOut 和 DropBlock 效果对比图

Fig. 5 DropOut and DropBlock comparison chart

由图 5 可知，随着网络层数增加，模型对数据的拟合能力增强，会导致 DropBlock 保持固定的失活比例，无法动态适应训练过程中的不同阶段，为此，设计动态随机失活块 DyDropBlock 模块，其能够随着网络深度增加逐步提升失活比例，动态调节失活过程。失活比例的计算式如式 (8) 所示。

$$p = (p_f - p_s) / \text{depth} \times n. \quad (8)$$

式中： p 为随机失活比例； p_f 为最终随机失活比例， p_s 为初始随机失活比例； depth 为网络深度数； n 为当前层的网络深度。

DyDropBlock 模块主要有 α 和 β 两个参数，其中 β 为丢弃特征单元的大小， α 为伯努利函数概率，控

制删除特征图单元的数量, 定义为

$$\alpha = \frac{p \times (w \times h)}{\beta^2 \times (w - \beta + 1) \times (h - \beta + 1)} \quad (9)$$

3.4 Dy-ODConv 模块

传统卷积层通过固定大小和方向的卷积核在特征图上提取特征, 但这种方式不适用于复杂的图像特征, 难以适应多样化的场景。故在动态卷积核的思想引入注意力机制, 对动态卷积核进行加权求和, 使特征提取更加灵活全面。

基于此思想, 提出一种基于 ODConv^[16] 并融合 Dy-ReLU^[17] 激活函数的动态卷积模块 Dy-ODConv, 该模块综合考虑空间核尺寸、输入通道数、输出通道数、卷积核数量 4 个维度的注意力机制, 按照位置、通道、滤波器和内核的顺序依次乘以卷积核 W , 捕获更加丰富的上下文信息。其详细结构如图 6 所示。

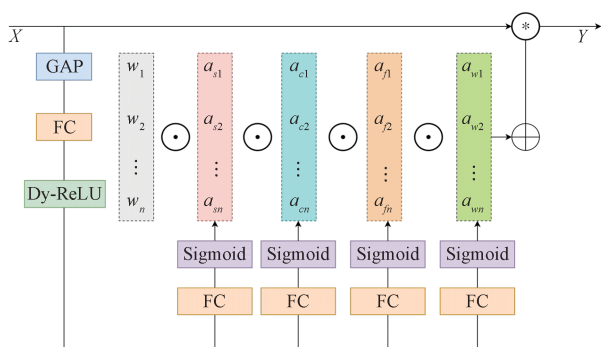


图 6 Dy-ODConv 结构示意图

Fig. 6 Dy-ODConv structural diagram

3.5 多项式焦点损失函数

急性淋巴细胞白血病图像中, 当目标特征仅占据图像中小部分区域时, 会带来模型在学习过程中对目标分类的关注度降低而过多地学习背景特征的问题。Focal Loss 损失函数能够缓解正负样本数量不平衡的问题, 如式 (10) 所示。

$$FL = -(1 - p_t)^\gamma \log(p_t) \quad (10)$$

式中: p_t 为目标类标签的预测概率; γ 为超参数。

然而 Focal Loss 损失函数平等地对待正负样本, 没有突出难分正负样本的贡献。为了动态调整难分正负样本的权重, 基于 Poly Loss^[18] 损失函数框架的 Poly-1 公式, 在 Focal Loss 泰勒多项式的第一个多项式中加入动态调节因子 $(1 + \omega\gamma)$, 其中 ω 为样本数量的权重, 设计一种多项式焦点损失函数 Poly Focal Loss, 该损失函数能够根据正样本比例动态调节 ω 的大小, 有效突出难分正负样本的贡献, 其定义如式 (11) 所示。

$$PFL = -(1 - p_t)^\gamma \log(p_t) + \omega\gamma(1 - p_t)^{\gamma+1} \quad (11)$$

4 实验与结果分析

4.1 实验环境及参数设置

本实验采用了 Blood Cells Cancer (ALL) 公开数据集。该数据集包含 4 个类别, 共计 3 242 张图像, 第一类为良性病例 (benign), 共 512 张图像; 其余 3 类为急性淋巴细胞白血病 (ALL) 组, 分别代表 3 种不同的恶性淋巴细胞亚型: Early Pre-B, Pre-B, Pro-B, 分别为 979, 955, 796 张图像, 图 7 直观展示了各类细胞的代表性图像。

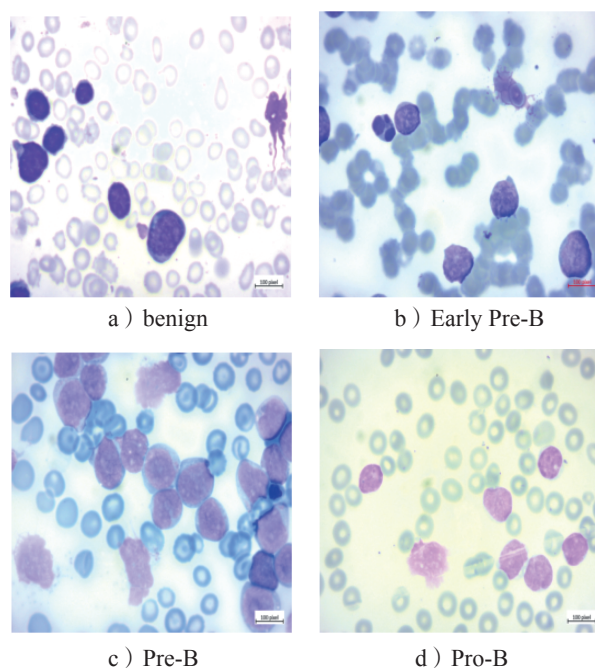


图 7 数据集的 4 类代表性图像

Fig. 7 Four categories of images of the dataset

按 7:2:1 的比例划分训练集、验证集和测试集, 为充分发挥卷积神经网络的特征提取能力, 本实验对训练集数据进行预处理, 包括随机旋转、水平翻转和垂直翻转、高斯噪声及背景去除等, 将训练集扩充到 12 314 张, 并且所有增强后的图像均被统一调整为 224×224 像素, 3 通道。实验环境配置见表 1。

表 1 硬件参数和软件环境

Table 1 Hardware parameters and software environment

名称	相关配置
操作系统	Ubuntu22.04
CPU	Intel Xeon Platinum 8255C
GPU	NVIDIA RTX3080
内存	40 GB
Python 版本	3.12
CUDA	12.1

4.2 评价指标

本实验采用以下指标评估模型性能: 准确率

(accuracy) A 、精确率 (precision) P 、召回率 (recall) R 、 F_1 评价指标, 这些指标的计算方法如下:

$$A = \frac{T_p + T_n}{T_p + F_p + T_n + F_n}, \quad (12)$$

$$P = \frac{T_p}{T_p + F_p}, \quad (13)$$

$$R = \frac{T_p}{T_p + F_n}, \quad (14)$$

$$F_1 = \frac{2 \times P \times R}{P + R}. \quad (15)$$

式(12)~(15)中: T_p 为正样本被预测正确的次数; T_n 为负样本被预测正确的次数; F_p 为负样本被预测为正样本的次数; F_n 为正样本被预测为负样本的次数; A 为正确分类的正样本和负样本数量与所有样本总数的比值; P 为正确分类的正样本数量与所有被分类为正样本总数的比值; R 为正确分类正样本数量与所有实际正样本数量的比值; F_1 为召回率和精确率的调和平均数。

此外, 采用参数量 Params 和浮点运算量 FLOPs 来衡量模型大小以及复杂度。其中, Params 为模型中的参数量, 以百万 (M) 为单位; FLOPs 为模型在一次前向推理过程中所需的浮点运算量, 以 10 亿次 (G) 每秒为单位;

4.3 实验结果及分析

使用 Efficient-DSP 模型对增强后的数据集进行训练, 设置学习率为 0.01, 使用多项式焦点损失函数 Poly Focal Loss, 训练批次为 32, 迭代次数为 100, 得到如图 8 所示的训练集准确率和损失函数变化曲线。从图 8 可以看出, 随着训练次数增加, 训练集的损失值呈现缓慢减少并逐渐稳定的变化趋势。同时, 分类的准确率不断提高, 说明模型分类准确率随着训练次数增加而增加并趋于稳定, 展现出模型良好的分类能力。

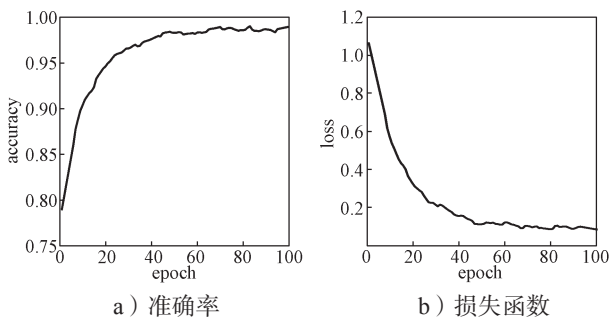


图 8 训练集的准确率和损失函数变化曲线

Fig. 8 Training set accuracy and loss function variation curves

将本文提出模型与经典图像分类模型进行比较,

包括 ShuffleNetV2^[19], GoogleNet^[20], EfficientNet^[21], ResNet34^[8] 和 Vision Transformer^[22] 等。为了获得客观的结果, 所有模型均使用同一数据集, 并采用相同的深度学习框架 PyTorch, 设置相同的学习率、训练批次和迭代次数, 得到的不同网络模型性能对比结果如表 2 所示。

表 2 不同网络模型性能对比结果

Table 2 Performance comparison of different network models

方法	accuracy/ %	precision/ %	recall/ %	F_1 / %	Params/ 10^6	FLOPs/ 10^9 s^{-1}
ResNet34	94.10	94.46	94.00	94.22	21.79	3.67
AlexNet	86.34	86.19	86.35	86.26	61.10	0.71
GoogleNet	90.37	89.34	89.81	89.57	6.62	1.51
Vision Transformer	93.50	93.11	92.24	92.67	57.30	11.28
EfficientNet	93.48	94.40	93.48	93.93	5.28	0.41
EfficientNetV2	96.02	96.09	95.43	95.75	21.45	2.90
ShuffleNetV2	95.03	95.71	95.40	95.55	1.37	0.04
本文方法	98.46	98.52	98.50	98.51	17.63	1.96

根据表 2 中的数据可知, 本文所提模型的图像分类准确率为 98.46%, 精确率为 98.52%, 召回率为 98.50%, F_1 为 98.51%, 这些评价指标值均显著优于对比模型的对应值。与原始模型 (EfficientNetV2) 相比, 改进后的算法在准确率上约提升了 2.54%; 与对比模型最优值相比, 准确率约提升了 3.61%。在模型复杂度方面, 模型参数量为 17.63×10^6 , 仅约为原始模型参数量的 82%; 浮点运算量为 $1.96 \times 10^9 \text{ s}^{-1}$, 仅约为原始模型的 67.6%。尽管 GoogleNet、EfficientNet 和 ShuffleNetV2 等模型具有更少的参数量, 但分类准确率低于本文所提模型的。综上所述, 改进后的模型在降低模型复杂度的同时, 有效提升了急性淋巴细胞白血病图像分类任务中的准确率, 实现了模型性能与效率的平衡优化。

通过分析数据集图像提出的残差置换注意力机制, 与 SE、ECA (efficient channel attention)^[23] 和 CBAM (convolutional block attention module)^[24] 等常规注意力机制进行比较, 所得结果如表 3 所示。

表 3 不同注意力机制性能实验对比结果

Table 3 Performance experiment comparison results of different attention mechanisms

方法	accuracy	precision	recall	F_1
EfficientnetV2+SE	96.02	96.09	95.43	95.75
EfficientnetV2+ECA	96.94	97.14	96.43	96.78
EfficientnetV2+CBAM	97.55	97.44	97.45	97.44
EfficientnetV2+SAR	97.84	98.00	97.85	97.92

由表 3 可知, ECA 注意力机制的图像分类准确率为 96.94%, 相较 SE 注意力机制的对应值约提升

了 0.96%, 但仅通过一维卷积实现局部跨通道高效交互, 缺乏空间信息关注, 提升幅度有限。CBAM 注意力机制的准确率为 97.55%, 相较 SE 注意力机制的图像分类准确率约提升了 1.59%; SAR 注意力机制表现最优, 其图像分类准确率、精确率、召回率、 F_1 分别为 97.84%, 98.00%, 97.85%, 97.92%, 其中准确率相较原模型的 SE 注意力的对应值约提升了 1.89%, 对比其它模型的最优值约提升了 0.30%。这一结果说明, SAR 注意力机制能更有效地增强模型对细节特征的提取能力, 显著提升急性淋巴细胞白血病图像分类准确率。

图 9 为测试集上生成的混淆矩阵, 其中横坐标表示细胞图像的真实类别, 纵坐标表示细胞图像的预测类别。对角线上的元素表示 4 类细胞图像正确分类的样本数, 而非对角线元素则代表错误分类的样本数。

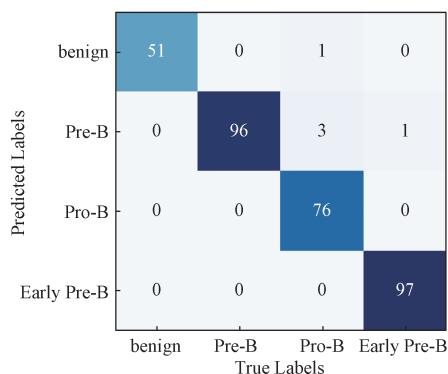


图 9 测试集混淆矩阵

Fig. 9 Test set confusion matrix

通过对比预测结果与实际标签, 可以看出模型较为准确地识别了 4 个类别, 仅有少量样本出现了错误分类。这表明该模型在细胞图像分类任务中表现出较高的准确性和可靠性。

4.4 消融实验

为验证 Efficient-DSP 改进模块的有效性, 进行了消融实验, 实验过程中, 将 EfficientNetV2 作为基准 (base), 得到的实验结果如表 4 所示。

表 4 消融实验对比结果

Table 4 Ablation experiment comparison results %

方法	accuracy	precision	recall	F_1
base	96.02	96.09	95.43	95.75
+DyDropBlock	97.82	97.59	97.78	97.68
+PFL	97.53	97.66	96.72	97.18
+Dy-ODConv	97.55	96.62	97.70	97.15
+SAR	97.84	98.00	97.85	97.92
本文方法	98.46	98.52	98.50	98.51

分析表 4 中数据, 可发现融合动态随机失活块

DyDropBlock、多项式焦点损失函数 PFL、动态卷积核 Dy-ODConv、残差置换注意力 SAR 均提高了模型分类准确率, 且 SAR 注意力机制的效果最明显, 模型的准确率约提升了 1.89%。最终改进算法的图像分类准确率达到 98.46%, 较原模型约提升了 2.54%, 从而验证了本文提出改进方案的有效性。

5 结论

为了更好地辅助医师对急性淋巴细胞白血病进行诊断, 本文提出了一种 EfficientNet-DSP 白血病分类方法。

1) 本文提出的模型与经典图像分类模型比较结果表明, 本文所提模型的图像分类准确率为 98.46%, 精确率为 98.52%, 召回率为 98.50%, F_1 为 98.51%, 这些评价指标值均显著优于对比模型的对应值。

2) 对比实验结果显示, EfficientNet-DSP 分类模型相比原始的 EfficientNetV2 模型在急性淋巴细胞白血病 (ALL) 公开数据集上的图像分类准确率提升了 2.54%, 与对比模型最优值相比, 其最优值提升了 3.61%, 表明了改进算法在急性淋巴细胞白血病分类中的有效性。

3) 消融实验结果验证了改进的模块对整体网络的贡献, 改进算法的图像分类准确率达 98.46%, 较原模型提升了 2.54%;

4) 所提方法能有效提高诊断准确率, 有望成为医师诊断的参考依据。

参考文献:

- [1] 倪敬花, 罗佳, 姜森, 等. 诱导多能干细胞在肿瘤疾病中的应用与前景 [J]. 中国组织工程研究, 2021, 25(13): 2127-2132.
NI Jinghua, LUO Jia, JIANG Sen, et al. Application and Prospect of Induced Pluripotent Stem Cells in Tumor Diseases [J]. Chinese Journal of Tissue Engineering Research, 2021, 25(13): 2127-2132.
- [2] 杨敏, 潘艳莎, 张长玲, 等. 儿童急性淋巴细胞白血病基线数据及早期治疗反应与预后的相关性 [J]. 天津医药, 2024, 52(9): 954-958.
YANG Min, PAN Yansha, ZHANG Changling, et al. Correlation Analysis of Baseline Data, Early Treatment Response and Prognosis in Children with Acute Lymphoblastic Leukemia [J]. Tianjin Medical Journal, 2024, 52(9): 954-958.
- [3] DAS P K, DIYA V A, MEHER S, et al. A Systematic Review on Recent Advancements in Deep and Machine

- Learning Based Detection and Classification of Acute Lymphoblastic Leukemia[J]. IEEE Access, 2022, 10: 81741–81763.
- [4] 李家成, 叶哲江, 张鹏飞. 融合多尺度注意力和累积学习的白血病分类识别[J]. 现代电子技术, 2023, 46(19): 49–54.
LI Jiacheng, YE Zhejiang, ZHANG Pengfei. Leukemia Classification and Recognition Based on Multi-Scale Attention and Cumulative Learning[J]. Modern Electronics Technique, 2023, 46(19): 49–54.
- [5] HOSSEINI A, ESHRAGHI M A, TAAMI T, et al. A Mobile Application Based on Efficient Lightweight CNN Model for Classification of B-ALL Cancer from Non-Cancerous Cells: A Design and Implementation Study[J]. Informatics in Medicine Unlocked, 2023, 39: 101244.
- [6] PRELLBERG J, KRAMER O. Acute Lymphoblastic Leukemia Classification from Microscopic Images Using Convolutional Neural Networks[C]//ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging. Singapore: Springer, 2019: 53–61.
- [7] PAN Y S, LIU M X, XIA Y, et al. Neighborhood-Correction Algorithm for Classification of Normal and Malignant Cells[C]//ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging. Singapore: Springer, 2019: 73–82.
- [8] HE K M, ZHANG X Y, REN S Q, et al. Deep Residual Learning for Image Recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 770–778.
- [9] 邵宇飞, 于潍赫, 刘阳. 基于注意力残差网络的急性淋巴细胞白血病分类[J]. 计算机系统应用, 2024, 33(6): 117–125.
SHAO Yufei, YU Weihe, LIU Yang. Acute Lymphocytic Leukemia Classification Based on Attention Residual Network[J]. Computer Systems and Applications, 2024, 33(6): 117–125.
- [10] 孙天宇, 朱庆涛, 杨健, 等. 基于改进 Vision Transformer 的血细胞图像识别方法研究[J]. 生物医学工程学杂志, 2022, 39(6): 1097–1107.
SUN Tianyu, ZHU Qingtao, YANG Jian, et al. An Improved Vision Transformer Model for the Recognition of Blood Cells[J]. Journal of Biomedical Engineering, 2022, 39(6): 1097–1107.
- [11] TAN M X, LE Q V. EfficientNetV2: Smaller Models and Faster Training[C]//International Conference on Machine Learning. [S. l.]: PMLR, 2021: 10096–10106.
- [12] HU J, SHEN L, SUN G. Squeeze-and-Excitation Networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 7132–7141.
- [13] ZHANG Q L, YANG Y B. SA-Net: Shuffle Attention for Deep Convolutional Neural Networks[C]//2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Toronto: IEEE, 2021: 2235–2239.
- [14] XU W, WAN Y. ELA: Efficient Local Attention for Deep Convolutional Neural Networks[EB/OL]. [2024–10–12]. <https://arxiv.org/abs/2403.01123>.
- [15] GHIASI G, LIN T Y, LE Q V. DropBlock: A Regularization Method for Convolutional Networks[EB/OL]. [2024–10–12]. <https://arxiv.org/abs/1810.12890>.
- [16] QIAN J J, LIN J, BAI D, et al. Omni-Dimensional Dynamic Convolution Meets Bottleneck Transformer: A Novel Improved High Accuracy Forest Fire Smoke Detection Model[J]. Forests, 2023, 14(4): 838.
- [17] CHEN Y P, DAI X Y, LIU M C, et al. Dynamic ReLU[C]//Computer Vision-ECCV 2020. Cham: Springer International Publishing, 2020: 351–367.
- [18] LENG Z, TAN M, LIU C, et al. PolyLoss: A Polynomial Expansion Perspective of Classification Loss Functions[EB/OL]. [2024–10–12]. <https://arxiv.org/abs/2204.12511>.
- [19] YANG H J, LIU J X, MEI G M, et al. Research on Real-Time Detection Method of Rail Corrugation Based on Improved ShuffleNet V2[J]. Engineering Applications of Artificial Intelligence, 2023, 126: 106825.
- [20] SZEGEDY C, LIU W, JIA Y Q, et al. Going Deeper with Convolutions[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston: IEEE, 2015: 1–9.
- [21] TAN M X, LE Q V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks[EB/OL]. [2024–10–12]. <https://arxiv.org/abs/1905.11946>.
- [22] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An Image Is Worth 16 × 16 Words: Transformers for Image Recognition at Scale[EB/OL]. [2024–10–12]. <https://arxiv.org/abs/2010.11929>.
- [23] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020: 11531–11539.
- [24] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional Block Attention Module[C]//Computer Vision-ECCV 2018. Cham: Springer, 2018: 3–19.

(责任编辑: 廖友媛)