

模型创新及相关说明

目录结构说明

```
environment.yml          # 创建测试环境
project/
├─ run.py                # 测试入口
├─ model/                # 测试模型和数据处理配置目录
│   └─ data.py           # 数据加载文件
│   └─ ConvNextModel.py  # MyConvNext模型
│   └─ DenseNetModel.py  # MyDenseNet模型
│   └─ MyConv.pth         # MyConvNext预训练权重
│   └─ MyDense.pth        # MyDenseNet预训练权重
├─ train_cls/            # BIRADS分类训练目录
│   └─ ConvNextModel.py  # MyConvNext模型
│   └─ data_cls.py       # 训练数据加载和处理
│   └─ func.py           # 训练及测试函数
│   └─ train_cls.py       # 训练入口
└─ train_fea/            # 特征识别训练目录
    └─ data_binary.py     # 训练数据加载和处理
    └─ DenseNetModel.py   # MyDenseNet模型
    └─ testloss.py        # F1 loss
    └─ train_binary.py     # 训练入口
```

测试流程

1.使用conda创建 environment.yml 中的新环境

```
conda env create -f environment.yml
```

2.将testB数据集放到与project目录同级

3.激活新环境 `conda activate py39` , 进入project文件夹

3.运行 `run.py` 文件, 等待片刻既可在project目录下得到 `cla_pre.csv` 和 `fea_pre.csv` 文件

一、概述

对于超声乳腺影像的BIRADS分类及特征识别任务, 我们采取了一系列创新措施和技术优化, 训练出了表现优异、推理速度快且占用资源少的模型。具体措施及其创新包括:

- **模型设计**
 - 设计了基于 ConvNeXt 的神经网络模型 **MyConvNeXt**, 用于执行 BIRADS 六分类任务。
 - 设计了基于 DenseNet 的神经网络模型 **MyDenseNet**, 用于执行特征识别二分类任务。
- **数据处理**
 - 应用了 **CutMix** 技术进行图像增广, 以增强模型的泛化能力。
 - 开发了 **CutBlack** 算法, 自动去除乳腺图像中的无用黑色边框, 提高数据质量。
- **训练策略**
 - 在BIRADS分类训练过程中采用了 **过采样** 策略, 以解决数据不平衡的问题。

- 在特征识别任务中采取**加权损失**策略, 以解决不同特征识别难度不同的问题。
- 创新设计并使用了 **F1 Loss** 作为辅助损失函数, 优化模型对不同类别的预测性能。

二、模型介绍及其创新

在乳腺癌超声图像的分类和特征识别任务中, 我们选择了没有Transformer结构的两个基准模型。尽管现如今ViT ([An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale](#)) 在许多视觉任务中取得了显著的性能, 但其成功往往依赖于大规模数据集的训练 (如图1所示)。然而, 我们的任务中训练集仅有 2500 张图像, 这样的小样本数据集难以充分发挥 ViT 的潜力, 容易导致过拟合和性能下降。同时, Transformer 相较于 CNN 来说计算量更大, 资源消耗更多, 计算时间更长。因此, 我们选择了 ConvNeXt 和 DenseNet 作为基准模型, 并在此基础上进行了多项创新, 以更好地适应小样本数据集的挑战。值得一提的是, 我们设计的轻量级模型MyConvNeXt和MyDense, 分别拥有**107MB**和**37MB**的参数规模。尽管体积小巧, 它们在保证快速推理速度和较低资源占用的同时, 仍能提供令人满意的性能表现。

When trained on mid-sized datasets such as ImageNet without strong regularization, these models yield modest accuracies of a few percentage points below ResNets of comparable size. This seemingly discouraging outcome may be expected: Transformers lack some of the inductive biases

inherent to CNNs, such as translation equivariance and locality, and therefore do not generalize well when trained on insufficient amounts of data.

However, the picture changes if the models are trained on larger datasets (14M-300M images). We find that large scale training trumps inductive bias. Our Vision Transformer (ViT) attains excellent results when pre-trained at sufficient scale and transferred to tasks with fewer datapoints. When pre-trained on the public ImageNet-21k dataset or the in-house JFT-300M dataset, ViT approaches or beats state of the art on multiple image recognition benchmarks. In particular, the best model reaches the accuracy of 88.55% on ImageNet, 90.72% on ImageNet-Real, 94.55% on CIFAR-100, and 77.63% on the VTAB suite of 19 tasks.

图1：ViT 论文中关于小样本数据集的效果说明

1. MyConvNeXt——执行 BIRADS 六分类任务

(1) ConvNeXt 模型简介

ConvNeXt ([ConvNeXt: A ConvNet for the 2020s](#)) 是 2022 年提出的一种创新的卷积神经网络架构。其通过采用大卷积核的深度可分离卷积、倒残差块 (Inverted Bottleneck)、全局响应归一化 (GRN)、LayerNorm 替代 BatchNorm, 以及多阶段设计等创新, 显著提升了模型的表达能力和性能, 实现了与 Transformer 相当甚至更优的视觉任务表现。(ConvNeXt结构图如图2)

在乳腺癌超声图像分类任务中, 图像往往伴随着大量的噪声和复杂的背景信息, ConvNeXt 凭借其大核卷积和模块化设计, 能够有效适应超声图像中复杂的纹理特征, 从而在噪声环境中提取出关键信息, 准确捕捉到病灶区域的微小特征。此外, ConvNeXt 的深度可分离卷积设计在计算效率上的优化, 使其能够在保持高精度的同时实现更快的推理速度和更低的资源消耗。

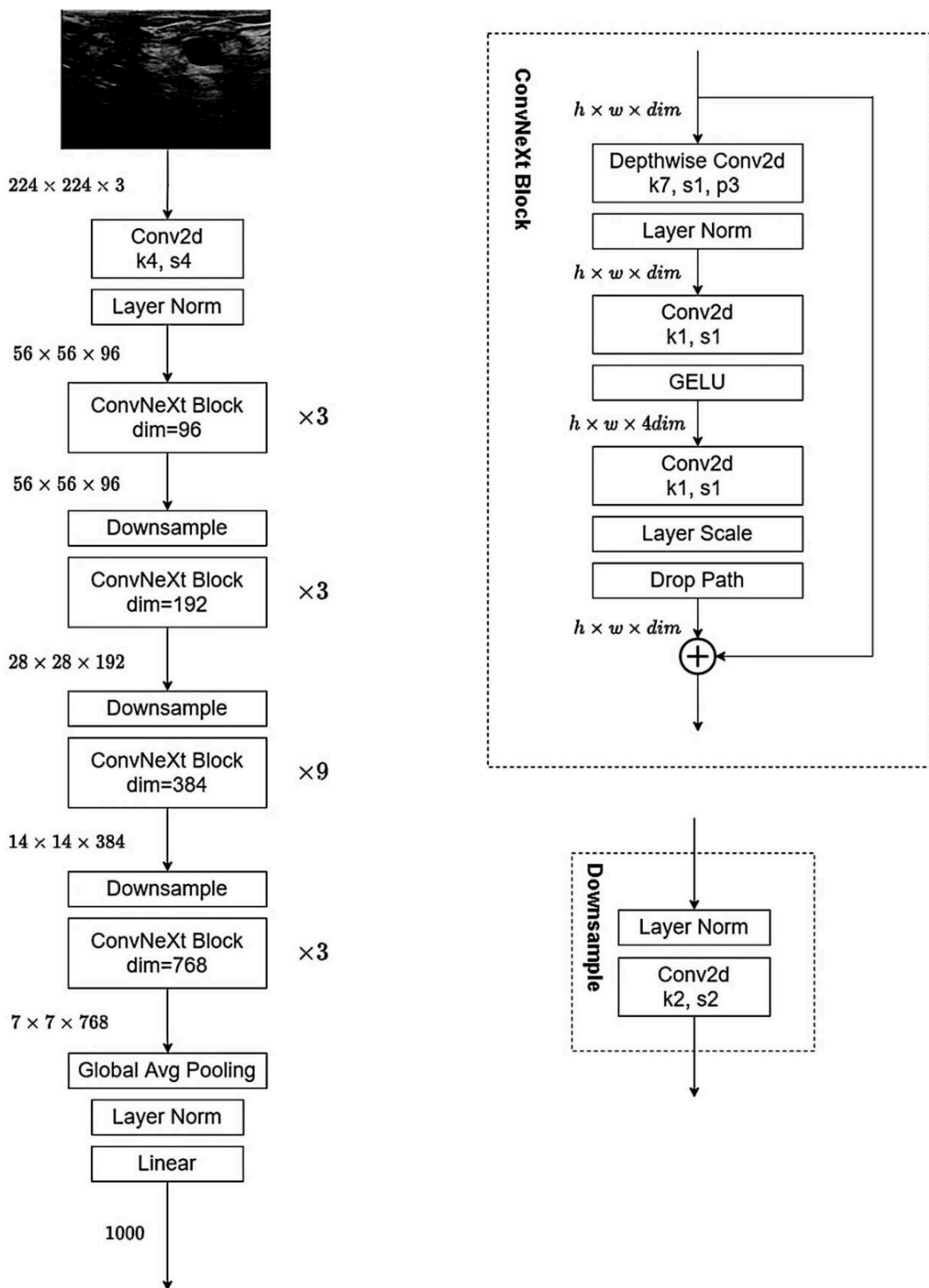


图2：ConvNeXt 结构图

(3) MyConvNeXt 模型的创新之处

基于ConvNeXt架构的优势，我们开发了一个针对乳腺癌超声特征分类任务的自定义版本——**MyConvNeXt**，并进行了多个创新性优化以适应任务需求。

1. 多层次的 Dropout 策略

- 在 CNBlock (ConvNeXt基本块) 和分类器中分别使用不同强度的 Dropout 进行正则化: 对于 CNBlock, 使用较小的 Dropout 率 (如 0.05), 在提供正则化的同时, 不会过度破坏特征的连续性, 保护重要特征。对于 分类器, 使用较大的 Dropout 率 (如 0.5), 有效防止过拟合, 使网络学习更加鲁棒的特征表示。

2. 动态调整随机深度 (Stochastic Depth)

- 动态调整每个 CNBlock 的随机深度参数 (sd_rate)。随机深度是一种正则化技术, 在训练过程中以一定概率随机跳过残差连接中的某些层, 从而简化模型并减少过拟合。

cur_sd_rate 从 0 线性增加到 0.6, 逐渐增加模型的正则化强度, 避免早期层过早受到过强的正则化影响, 确保深层网络能够更好地泛化。

3. 改进分类器层

- 修改了 ConvNeXt 模型的分层器层, 增加了额外的线性层和激活函数, 提高了模型的表达能力。分类器层包括归一化、线性变换、激活函数和最终的分层输出。

通过上述创新, MyConvNeXt 模型能够更好地适应小样本数据集, 减少过拟合的风险, 特别是在乳腺癌超声图像分类任务中, 使得模型在训练过程中更稳定地学习到关键特征, 提高了分类的准确性和泛化能力。

2. MyDenseNet——执行特征识别二分类任务

(1) DenseNet 介绍及选择理由

DenseNet([Densely Connected Convolutional Networks \(DenseNet\)](#)) 是一种先进的卷积神经网络结构, 其核心特点在于密集连接机制, 旨在优化信息流和梯度传播的效率, 能够更高效地利用特征, 减少冗余。这种特征复用策略使得 DenseNet 在参数量较少的情况下也能展现出卓越的性能, 降低了过拟合的风险。

对于乳腺癌超声图像, DenseNet 的密集连接特性有助于提取关键特征, 减弱噪声干扰, 提高分类精度。此外, DenseNet 避免了重复计算, 训练效率高, 推理速度快, 具有实际应用价值。

(2) MyDenseNet 模型的创新之处

基于DenseNet架构的优势, 我们开发了一个针对乳腺癌超声特征分类任务的自定义版本——MyDenseNet, 并进行了多个创新性优化以适应任务需求。MyDenseNet结构图如下所示:

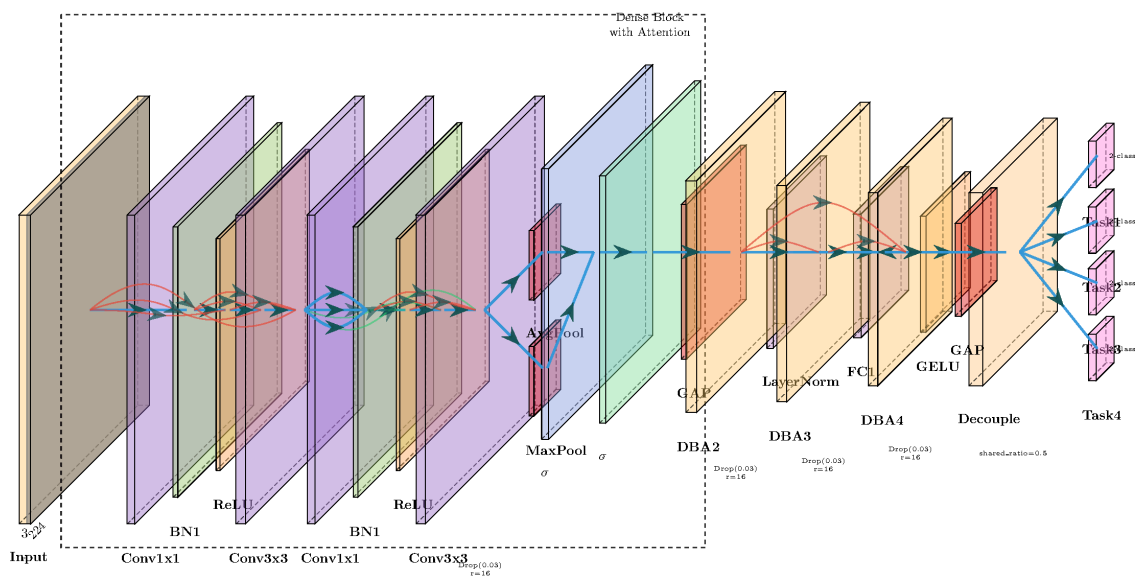


图3：MyDenseNet结构图

1. 多层次注意力机制的融合

引入了通道注意力（Channel Attention）和空间注意力（Spatial Attention）机制。

- **通道注意力**：通过平均池化和最大池化分支捕获通道信息，使用共享的 MLP 网络进行特征转换，最终通过 Sigmoid 函数得到通道权重。我们认为通道注意力机制特别适合 DenseNet 这样的密集连接型模型，因为它能够学习到不同通道的重要性，突出更加关键的特征，从而增强特征的利用率，进一步提升模型的性能

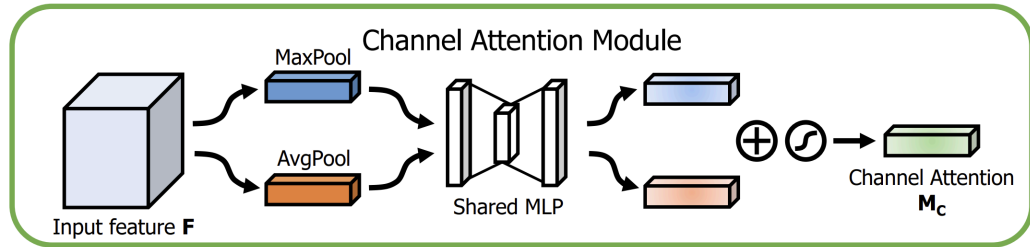


图4：通道注意力示意图

- **空间注意力**：空间注意力模块则更加关注空间维度的信息。通过对特征图在通道维度上的平均值和最大值进行聚合，使用卷积层学习空间位置的重要性权重。这种机制特别适合乳腺超声影像数据，因为能关注图像的关键部分，即乳腺癌结节，从而提高分类的准确性。

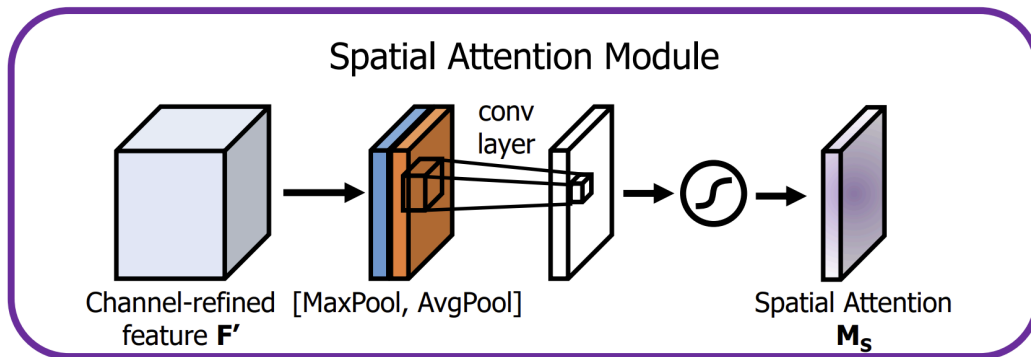


图5：空间注意力示意图

2. 自适应 ratio 设置

相较于传统的一直使用固定的ratio值,受到ConvNeXt中stochastic depth丢弃率逐层递增的启发,我们让 ratio 随着网络深度的增加而增大 ($\text{ratio} = 16 * (i + 1) // 2$), 在浅层使用较小的 ratio 保留更多细节特征, 在深层使用更大的 ratio 进行特征压缩。这样既有利于处理更抽象的语义信息, 也可以控制计算复杂度, 提高计算效率。

3. 多任务学习与任务特定解耦模块

为了同时处理多个二分类任务, 引入了**多任务学习**, 每个任务的输出通过独立的线性层实现, 确保了模型在处理不同任务时能够保持其独特的特征表示. 特别的, 我们设计了**任务特定解耦模块** (TaskSpecificDecoupling)。该模块将特征空间分为共享特征空间和任务特定特征空间, 通过共享投影器提取通用特征, 任务特定投影器捕获独特信息, 最后融合形成完整的任务表示。实现了任务间的知识迁移, 同时保持任务独立性, 减少计算和存储开销。

4. DenseNet 结构优化

- **多层次的 Dropout 策略**：在 denseblock 和分类器中分别使用不同强度的 Dropout 进行正则化: 对于 Denseblock, 使用较小的 Dropout 率 (如 0.03), 在提供正则化的同时, 不会过度破坏特征的连续性, 保护重要特征。对于 分类器, 使用较大的 Dropout 率 (如 0.5), 有效防止过拟合, 使网络学习更加鲁棒的特征表示。
- **规范化和激活函数创新**：使用 LayerNorm 替代 BatchNorm, 帮助稳定梯度流动. 采用双重激活函数 (ReLU + GELU), ReLU 防止梯度消失, GELU 提供平滑的梯度传播, 多层次的非线性变换增强了模型的表达能力。

三、数据处理

针对乳腺癌超声图像数据集，我们设计了多种数据增强方法：CutMix 聚焦癌结点，CutBlack 去除黑色边框，几何变换与模糊提升数据多样性。这些方法有效提高了数据质量和模型的鲁棒性，帮助模型在新数据上更好地泛化。

1. CutMix 增强

CutMix 是该数据处理流程中的核心部分之一。通过裁剪出图像 A 中包含癌结点的边框区域，并将其覆盖在图像 B 的癌结点区域上，同时沿用图像 A 的标签，从而实现癌结点特征的组合增强。针对本次比赛，测试集不提供肿瘤结点的BBox坐标，CutMix处理更加合理的利用了训练集上的BBox数据。

优势：

- **聚焦癌结点：**有效引导模型关注癌结点区域，提升特征学习能力。
- **丰富样本特征：**通过组合不同图像的癌结点特征，生成多样化的训练样本，扩展训练视野。

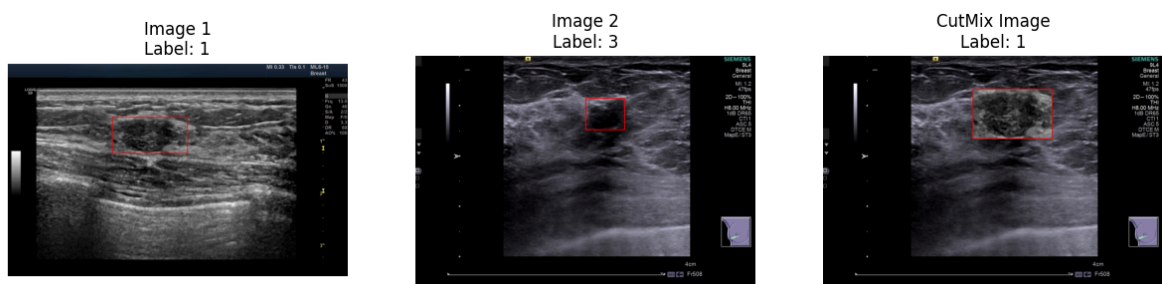


图6：CutMix 处理效果(BBox边框在训练时不会画出)

2. CutBlack 处理

我们独立设计了CutBlack方法，来裁剪精准地裁剪掉乳腺超声图像中的黑色边框，减少无关背景对模型的干扰。该方法的实现为：逐列逐行扫描图像，获取该列/行的所有像素值，计算黑色像素比例，当黑色比例超过阈值时裁剪。CutBlack处理解决了数据集来源不一,图片质量不一的问题,构建出具有针对性的训练数据。

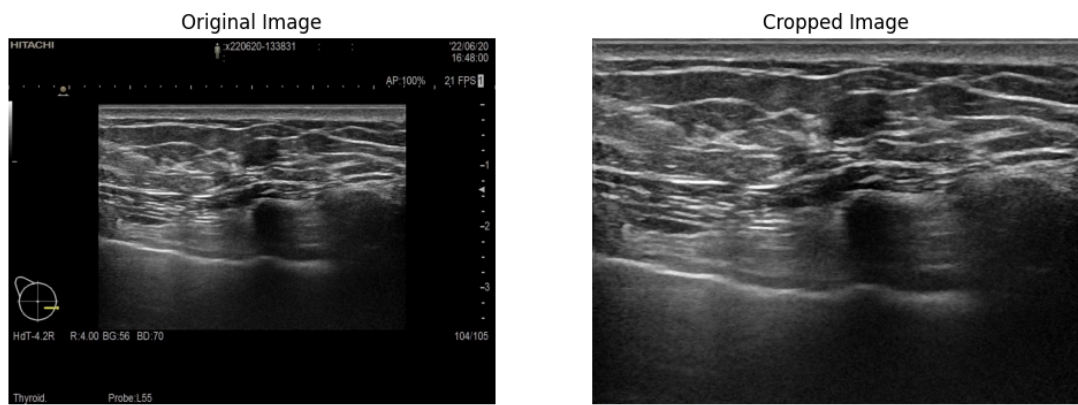


图7：CutBlack 处理效果

3. 其他增强方法

除了 CutMix 和 CutBlack，还使用了以下经典的图像增强方法：

- **几何变换**：包括随机旋转、翻转、缩放与裁剪等，增加了图像在空间维度上的变化，让模型更能适应不同方向、不同尺度的癌结节特征。
- **高斯模糊**：通过适当的模糊处理，提升模型对低清晰度或噪声干扰的鲁棒性，尤其在边缘模糊的癌结节区域表现突出。
- **随机擦除**：通过随机地擦除图像中的某些区域来增加训练数据的多样性，从而提高模型的泛化能力，减少过拟合的产生。

四、训练策略

在模型训练过程中，我们实施了多项创新机制，以提升训练效率、优化资源使用和改善模型性能。以下是具体的改进措施：

1. 混合精度训练与标签平滑

- **混合精度训练**：在六分类和二分类任务中均采用了混合精度训练技术，通过智能结合FP32（单精度）和FP16（半精度）计算，显著降低了显存占用，同时加快了训练速度。
- **标签平滑**：同样应用于六分类和二分类任务，通过将硬标签转换为软标签，减少了模型过拟合的风险，增强了模型的泛化能力。

2. 六分类任务中的过采样策略

- **过采样与差异化权重分配**：为了解决六分类任务中样本分类不平衡的问题，特别是在乳腺超声数据集中患者疾病类型的不平衡，我们采用了过采样的**差异化权重分配**的方法。即我们对少数类样本进行过采样，增加其在训练集中的数量，并为这些样本赋予更高的权重。通过差异化权重的平衡机制，有效提高了少数类别的识别率，提升了模型的整体性能。

3. 二分类任务中的任务权重机制与F1损失函数

- **任务权重机制**：鉴于不同任务的学习难度存在差异，我们为每个任务设置**差异化损失权重**，确保较难任务获得更多的关注，防止简单任务主导整个训练过程。
- **F1损失函数**：为了解决高准确率下F1分数仍有较大提升空间的问题，我们设计了一种**基于概率计算的F1损失函数**作为辅助函数。这种方法克服了传统F1分数不可微的限制，通过计算TP、FP、FN来得出可微的F1分数，从而直接促进F1分数的提升。