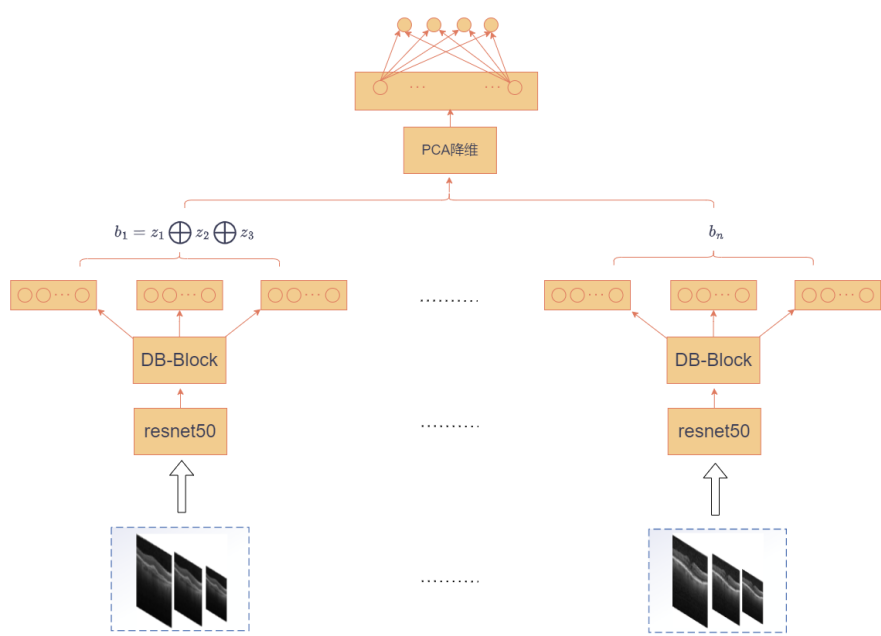


第3章 技术方案

3.1 模型结构

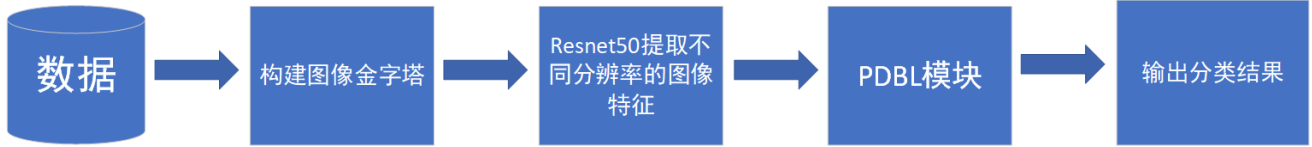
在临床实践中，由于多尺度的图片信息对于组织病理学的图像分析至关重要，病理学家通过观察不同放大倍数的物镜下的组织切片。受此启发，我们采用了了Plug-and-Play Pyramidal Deep-Broad Learning (PDBL) 模型的思想，对数据集中的每一个样本，建立了3种分辨率尺度的图像金字塔 ($112 \times 112, 160 \times 160, 224 \times 224$)，并基于resnet50网络框架，对视网膜OCT图像的病例进行了多分类处理。整体结构如下图所示：



为了充分提取特征图像在不同分辨率下的全局信息，对于每一个数据集中的样本，我们做了两个操作：1. 我们将每个样本缩放为 $112 \times 112, 160 \times 160, 224 \times 224$ 这3个尺度，记为 $X = \{X_1, X_2, X_3\}$ ；2. 在ResNet50的每个阶段的最后一层，对得到的特征图像的每个通道进行全局平均池化操作。使得初始输入和卷积神经网络都能够考虑到多尺度的图像信息。

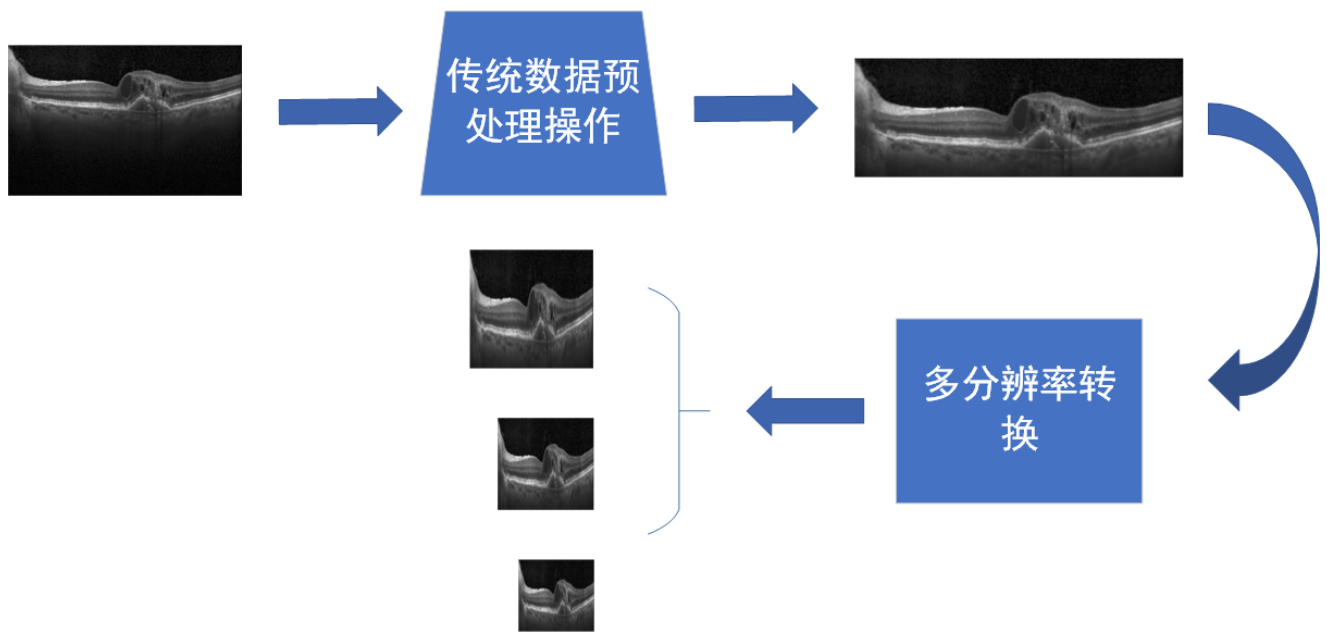
3.2 模型原理

我们将上述复杂的模型结构进一步抽象，可以得到以下流程图：



本部分将根据抽象后的流程图，依次介绍模型的每一个模块。

3.2.1 图像预处理



由于OCT图像的产生、传输和存储过程中会不可避免地出现图像清晰度下降、对比度偏低和包含噪声等降质现象，因此，需要进行图像预处理操作，对图片进行去噪、增强对比度等操作来提高图片的质量。本模型的预处理方法主要分为以下两部分：

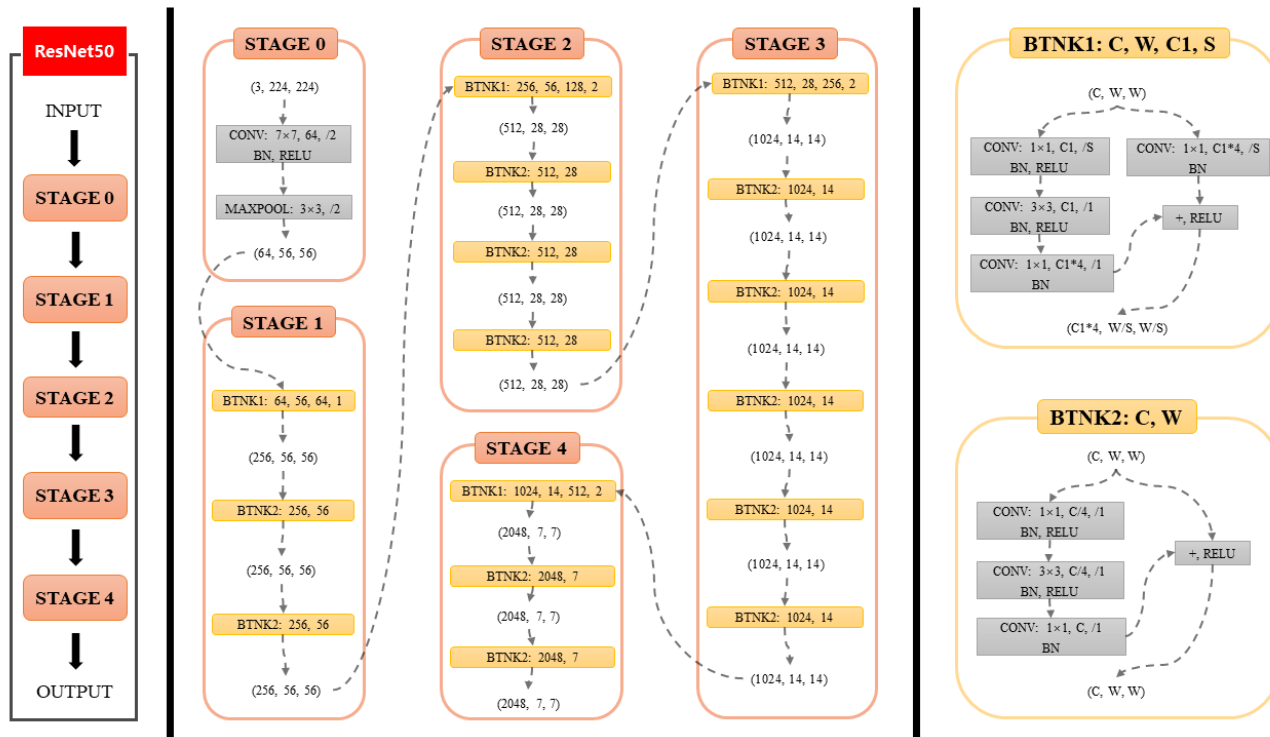
1. 传统图像预处理法：高斯模糊、二值化处理、中值滤波、形态学操作、图像拟合、图像位置调整、图像剪裁、图像归一化等
2. 多分辨率转换：为了获取多分辨率的信息，我们使用了双线性插值法对每个数据样本进行不同尺度的缩放，分别为 224×224 , 160×160 , 112×112 这3个尺度。

3.2.2 构建图像金字塔

受到病理学家通过不同放大倍数的物镜来观察组织样本的启发，对于每一个数据集中的样本，我们采用了OpenCV库的resize函数，将每个样本缩放为 112×112 , 160×160 , 224×224 这3个尺度，记为 $X = \{X_1, X_2, X_3\}$ ，使得卷积神经网络也能够考虑多尺度的图像信息。

3.2.3 ResNet50

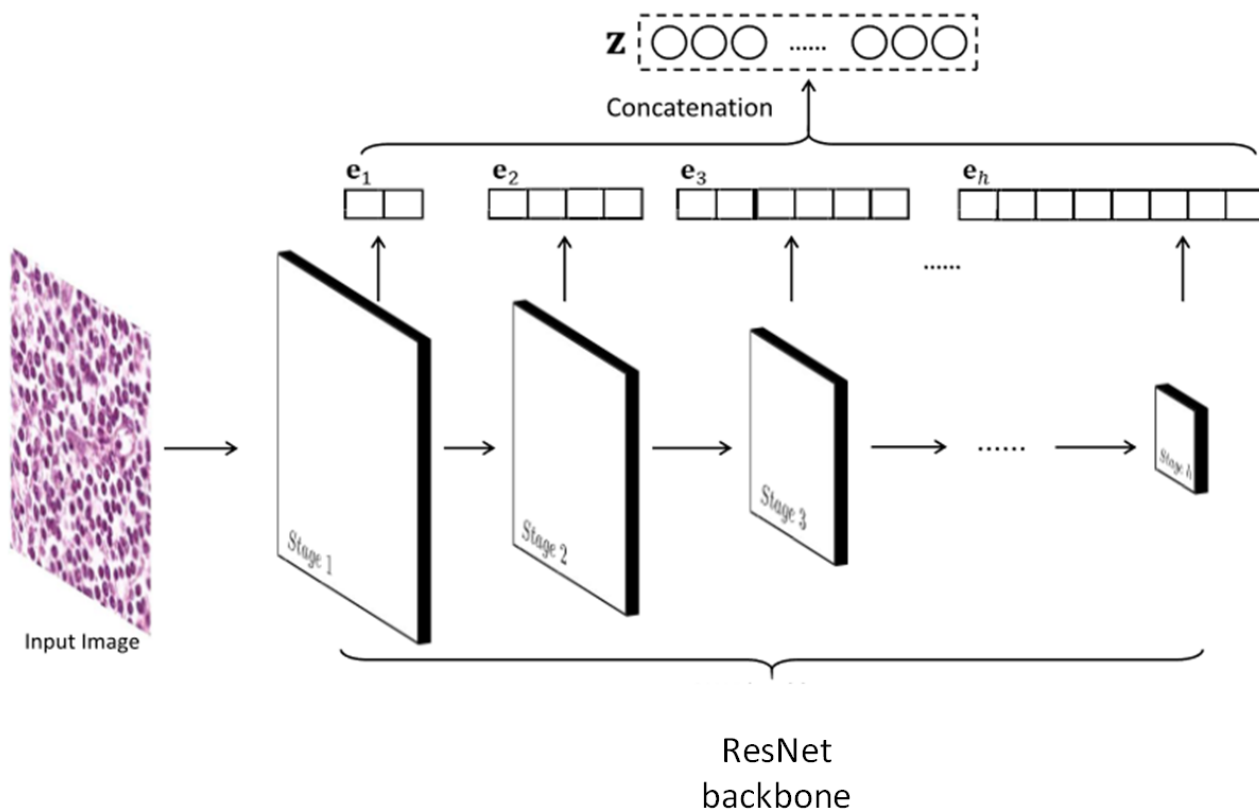
由于ResNet50出色的性能和训练效果，ResNet在计算机视觉领域的各种任务中得到了广泛的应用。它通过引入残差学习的概念，通过构建残差块来构建深层网络；并且引入了残差连接，有助于解决深层神经网络中的梯度消失问题。作为ResNet的一个子版本，ResNet50分为5个阶段，如图所示（图片引用自：<https://zhuanlan.zhihu.com/p/353235794>）：



我们旨在通过ResNet50对不同分辨率的图像 X_1, X_2, X_3 进行特征提取，ResNet50每个阶段的最后一层将会作为DB-Block模块的输入。

3.2.4 Deep-Broad Block (DB-Block) 模块:

DB-Block是用来提取图像金字塔特征的一个模块，对于 $X = \{X_1, X_2, X_3\}$ 中的子图像 X_i ，我们首先从ResNet50主干的每个阶段的最后一层来提取中间深度特征 f_k （ k 表示第几个阶段）， f_k 将通过自适应全局平均池化被压缩为通道特征向量 e_k ，最终将所有的 e_k 拼接，得到结果 z_i ，如 X_1 对应 z_1 。具体过程如下图所示：



对于每个样本 i ，按照上面的过程，我们可以得到一个 b_i ， b_i 为一个 $1 \times c$ 大小矩阵， c 为不同ResNet50不同阶段最后一层图片的通道数之和，即

$$c = \sum_0^4 c_i$$

其中 c_i 为第 i 个阶段最后一层的特征图所具有的通道数。

3.2.5 主成分分析和分类实现

矩阵构建和正则化处理：

经过DB-Block模块，最终可以得到一组结合了多尺度特征图像信息的特征向量 $b_1, b_2, b_3, \dots, b_n$ 。将它们按行排列，并进行正则化处理，就可以得到 B 矩阵。

$$B = \delta([b_1^T, b_2^T, \dots, b_n^T]^T)$$

其中， δ 表示矩阵正则化。

PCA降维

为了减少特征维数和冗余，我们采用主成分分析方法对矩阵 B 进行降维处理。首先求得矩阵 B 的协方差矩阵 C ，再对 C 进行奇异值分解（SVD）得到 U 矩阵。通过公式 $A = BU$ 对矩阵 B 进行降维处理，计算得到的 A 表示降维后得到的最终结果矩阵。最后，将 A 矩阵与通过伪逆算法计算得到PDBL的权重 W_{PDBL} 相乘，即可得到最终每个样本的分类概率。样本将被分入其计算所得概率值最大的类中。具体计算步骤如下：

1. 协方差矩阵 C 的计算

$$C = \frac{1}{n} B^T B$$

协方差矩阵 C 表示表现不同维度间的相关性。矩阵中元素的越接近1，说明二者正相关性越强；越接近-1，表明二者负相关性越强。

2. 奇异值分解（SVD）

$$[U, \Sigma, V^T] = SVD(C)$$

SVD是一种矩阵分解的方法，由于奇异值减低得非常快，很多情况下前1%的奇异值的和就占了全部奇异值之和的99%以上的比例。所以我们可以用最大的 k 个奇异值和对应的左右奇异向量来近似描述矩阵。

3. 右乘奇异矩阵压缩列数

$$A = BU$$

通过矩阵右乘这一列变换操作， B 矩阵被降维至 A 矩阵

概率映射

最终，每个样本的分类结果可以通过以下公式得到：

$$Y = AW_{PDBL}$$

其中 W_{PDBL} 是PDBL的权重，可以通过违逆算法求得。 Y 是该分类问题的概率矩阵，一行表示一个样本，在本问题中， Y 矩阵有4列，表示4个分类标签。矩阵元素 $a_{i,j}$ 表示样本 i 被分到第 j 类的概率。