



EE369 课程大作业

基于3D卷积神经网络的医学图像分类

沈力 517030910243

2020年1月1日



上海交通大學

SHANGHAI JIAO TONG UNIVERSITY

项目完成情况



- 最终Leaderboard上的分数： 0.73313
- Leaderboard上显示的名字： 517030910243_沈力
- 总共提交次数： 51 (34+17)
- 是否使用小号刷分： 是
- 方法简述： 基于3D-DenseNet卷积神经网络的医学图像分类
- 主要使用的代码框架： Keras+TensorFlow
- 模型大小（MB）： 5.28 (3*1.76) MB
- 亮点：
 - 数据增强，对原数据集分别对三个坐标平面做镜像；
 - 训练前模型加载初始权重，提高训练效率和预测精度；
 - 多次预测结果平均，最小化难预测的样本的预测误差。
- 代码链接： <https://github.com/Sshenl/SJTU-M3DV>

问题描述



- Medical 3D Voxel Classification(M3DV):
 - 这是一个肺部结节的分类项目。数据集是一系列的3D肺部结节的CT扫描。对于每个样本，包含：
 - 一个大小为100x100x100、已经经过预处理的3D体素，其中心为肺部结节（病灶）
 - 由医生给出的结节分割mask
 - 有某种方式得出的二分类结果，0 / 1
- 目标：
 - 在一个 *train_val* 数据集上训练并评估自己的模型，然后在 *test* 数据集上预测分类结果，要求预测的二分类结果尽可能的准确。
- 方法：
 - 使用3D卷积神经网络进行深度学习。

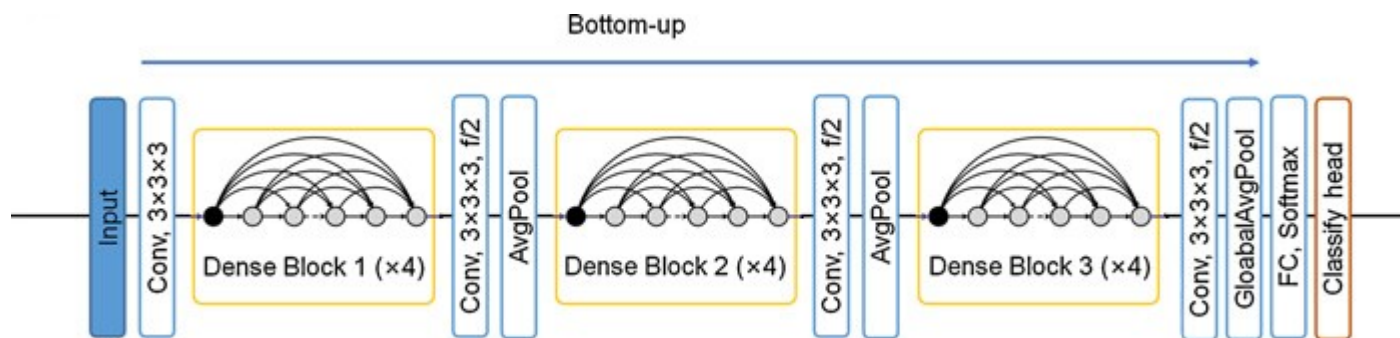
模型设计



■ Dense Convolutional Network (DenseNet)

- DenseNet是一个以前馈地方式将每个层与其它层连接的密集卷积网络，它具有缓解了消失梯度问题，加强了特征传播，鼓励特征重用，并大大减少了参数的数量等优点。
- 本次项目的DenseNet实现代码参考自：

<https://github.com/duducheng/DenseSharp>

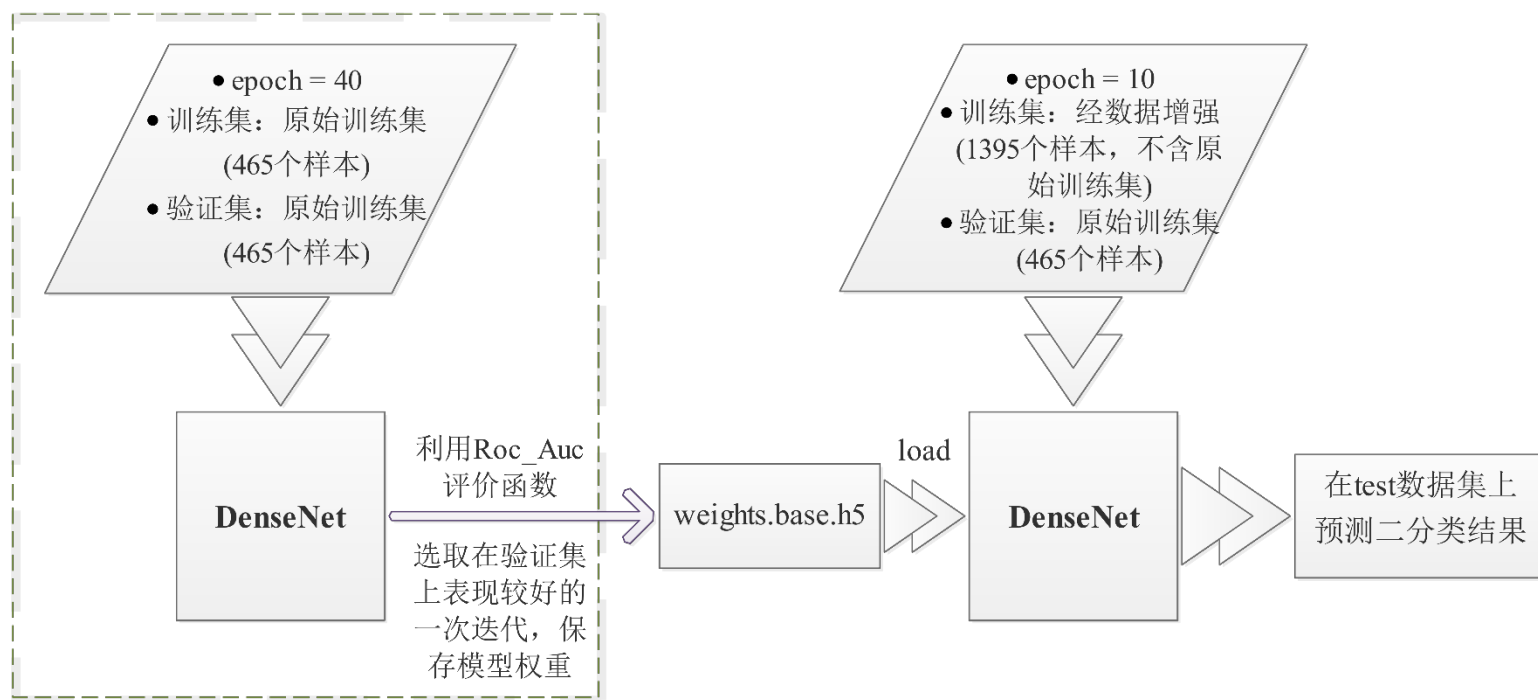


DenseNet网络结构示意图

模型设计



■ 基于DenseNet的3D医学图像分类模型



- 在Keras中的callbacks函数中加入Roc_Auc评价函数的代码参考自:
https://blog.csdn.net/qq_16234613/article/details/79437952

性能分析



■ 预测准确性

- 加载weights.base.h5权重，仅预测，在验证集上Roc_Auc评分大于0.78，在test测试集上Kaggle评分大于0.65；
- 加载weights.base.h5权重，在1395个样本上训练并预测。多数情况下，前两个epoch的模型的预测结果在Kaggle上评分大于0.69；
- 将三个在测试集上表现最好的模型的预测结果（Kaggle上评分分别为0.70557、0.70674和0.70762）取平均，最终得到的结果在Kaggle上评分为0.73313。

■ 运行时间

- 配置：AWS g4dn.2xlarge实例
- 样本量：每个epoch训练1395个样本，每个样本维度为32x32x32
- 单个epoch训练时间约为60 s

■ 模型大小： 5.28 (3*1.76) MB

技巧设计



■ 数据预处理

- 原始数据集内样本中包括肺部结节的CT扫描图voxel和结节分隔图seg，样本数据维度为100x100x100。为了充分地利用数据同时减轻计算负担，我们将voxel和seg矩阵点乘，然后取其中心32x32x32部分。

■ 数据增强

a) Mixup

- Mixup令模型在处理样本和样本之间的区域时表现为线性，这种线性建模减少了在预测训练样本以外的数据时的不适应性。其处理方式为：

$$\tilde{x} = \lambda x_i + (1 - \lambda)x_j$$

$$\tilde{y} = \lambda y_i + (1 - \lambda)y_j$$

- 其中， (x_i, y_i) 和 (x_j, y_j) 是从训练数据中随机抽取的两个样本，且 $\lambda \in [0,1]$ 。

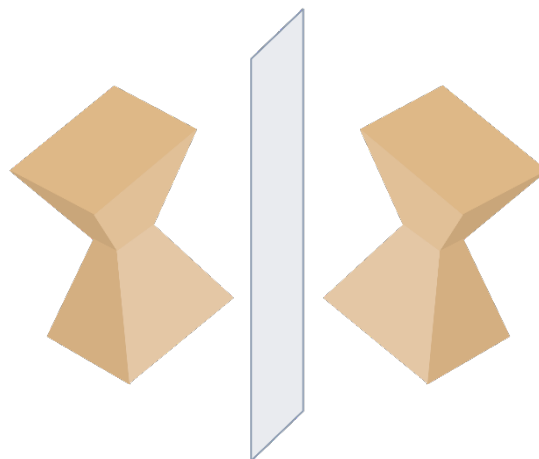
技巧设计



■ 数据增强

b) 镜像变换

- 对于3D图像，镜像变换就是将原图像相对于某一平面做镜像，同时镜像变换不改变样本的标签。对于一个 $32 \times 32 \times 32$ 大小的样本，我们可以分别相对 $x=16$ 、 $y=16$ 和 $z=16$ 平面做镜像，从而得到一个经数据增强后的含1395个样本的训练集。



3D图像镜像变换示意图

技巧设计



■ 模型加载初始权重

- 在神经网络训练中，好的权重初始化会加速训练过程并提升训练精度。
- 初始权重的获取：
 - 使用DenseNet，训练集和验证集均为原始的465个样本，选取在验证集上的Roc_Auc评分在0.8左右，且在测试集上的预测结果的Kaggle评分最高的一次迭代的权重，即该权重在训练集上表现出了一定的过拟合，但不是严重的过拟合以至于在测试集上表现糟糕。

■ 多次预测结果平均

- 对于一些较难预测的样本，用同一模型不同次训练下的预测结果可能会产生很大的差别，将多次预测结果取平均则可以将这个误差最小化。

讨论



■ 数据增强的效果

- 我们首先使用了mixup用作数据增强，发现mixup虽能缓解过拟合的速度，但对提高在测试集上的预测效果却没有显著帮助。
- 我们从旋转、缩放等数据增强方法中受到启发，想到了使用镜像变换用作数据增强。镜像变换结合我们最终的模型，不仅有缓解过拟合的作用，而且将模型在测试集上预测结果的评分上限提高到了0.7以上。

■ 模型预训练和加载初始权重的配合

- 我们首先使用原始数据集在模型中训练得到一个轻度过拟合但在测试集上仍表现较好的权重。然后将此权重作为初始权重，同时将训练集更换为经数据增强后的1395个样本（不含原数据集）再次训练。
- 经过多次训练我们发现在前几次迭代中，受到数据增强的作用，初始权重的过拟合得到一定程度的缓解，同时使得前两次迭代的模型在测试集上的预测准度进一步提高，预测结果在Kaggle上的评分基本保持在0.69~0.71之间。

讨论



■ 可能的改进

1. 尝试使用更多的数据增强方法，包括对mixup进一步调参，以求在缓解过拟合的同时提升模型在测试集上表现。
2. 尝试调节学习率、正则化参数和loss函数等参数，获得更好的抗过拟合效果和预测效果。
3. 可以多次训练获取更多预测结果取平均，当训练次数足够大时，某几次预测错误的影响就可以忽略不计，而对于非常难预测的样本，其预测结果最终会稳定在0.5左右，不至于产生较大误差。

Thank You



上海交通大学

SHANGHAI JIAO TONG UNIVERSITY