

# 基于卷积神经网络的糖尿病视网膜病变分类网络

梅骏逸 2111876

网络空间安全学院 信息安全、法学

2022 年 11 月



1 数据集

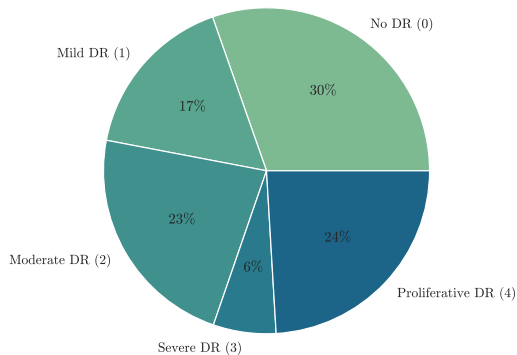
2 模型架构

3 实验

## 数据集

# 数据集

- DDR 数据集 [5]
- 样本不均衡的问题。
- 预训练模型
- Label Smoothing
- 尝试减少样本不均衡问题带来的对模型的影响，并且取得了较好的效果。



## 数据集的增强

在前期的尝试中发现对数据增强能够显著提高模型的性能，实验中使用 PyTorch 内置工具对图片进行了变换：

- ❶ 随机翻转（纵向、横向）
- ❷ 随机旋转一定角度
- ❸ TrivialAugment（随机选择增强操作和增强幅度）
- ❹ 随机改变图像的色调
- ❺ 随机抹除部分像素的内容

## 模型架构

# 基本架构

- 使用预训练模型及 Patch Embedding 作为 Backbone 提取特征;
- CBAM 注意力机制模块 [11]

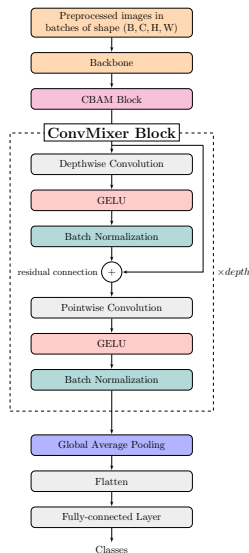
$$w_1 = \sigma(\text{MLP}(\text{AvgPool}(x_0)) + \text{MLP}(\text{MaxPool}(x_0)))$$

$$x_1 = w_1 \otimes x_0$$

$$w_2 = \sigma(\text{Conv}_{2 \rightarrow 1}([\text{AvgPool}(x_1); \text{MaxPool}(x_1)]_{\text{concat}}))$$

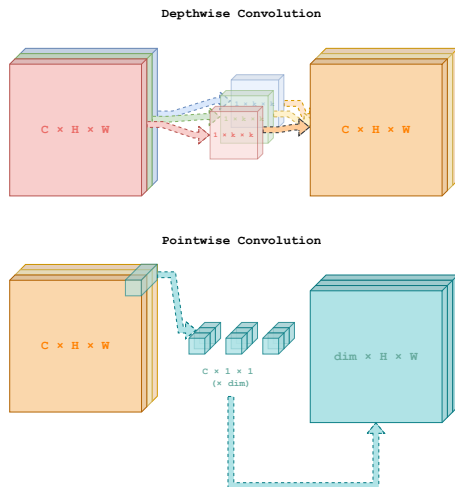
$$x_2 = w_2 \otimes x_1$$

- ConvMixer Block[10] 组成后续网络
- Global Average Pooling + Flatten + 全连接层
- 实现 CABNet GAB+CAB 注意力机制并进行了实验



## Depthwise 及 Pointwise 卷积

- 将一次卷积运算拆分成两次独立的操作；
- Depthwise 卷积对每个通道使用不同的卷积核；
- Pointwise 卷积在每一个点上对通道进行卷积；





# Backbones

- Patch Embedding
  - ConvMixer 原网络
  - 将输入进行分块卷积（步长与卷积核大小均为  $p$ ）
- ConvNeXt
  - 综合了过去各种网络中的优点
  - 具有良好的性能 [7]
  - 实验中使用冻结参数连接 CBAM 和 ConvMixer、不冻结参数连接 CBAM 和 ConvMixer、不冻结参数更换全连接层这三种方式进行训练并进行了对比。
- EfficientNet
  - 较少的参数量和较快的速度 [9]
- DenseNet
  - 将每一层的处理结果连接（Concatenate）[3]

## GAB+CAB

Global Attention 过程可以表示如下

$$w_1 = \sigma(\text{Conv}(\text{AvgPool}(x_0)))$$

$$x_1 = w_1 \otimes x_0$$

$$w_2 = \sigma(\text{AvgPool}_{\text{channel}}(x_1))$$

$$x_2 = w_2 \otimes x_1$$

Category Attention 模块过程大致如下

- ① 将通道数通过卷积增广到  $k \times L$ ，其中  $L$  表示类别数量，通过池化得到每一类的权重；
- ② 在改变维数后对  $k$  所在维度池化
- ③ 将两种操作结果对应元素相乘并在通道维数上池化得到权重
- ④ 权重与 Global Attention 结果相乘得到 CAB 的处理结果。

最后通过池化、Flatten 和全连接层得到分类结果。  
实验中根据 CABNet 论文中的实验结果，取  $k = 5$ 。

## 实验

## 训练设置

- AdamW 优化器
- Label Smoothing + CrossEntropy 作为损失函数
- 学习率下降

# 评估指标

- Accuracy
- Kappa[1]

● 令  $f_{ij}$  表示混淆矩阵中第  $i$  行第  $j$  列, Kappa 系数的定义如下:

$$\text{Kappa} = \frac{p_o - p_e}{1 - p_e}$$

$$p_o = \frac{1}{n} \sum_i^k f_{ii}$$

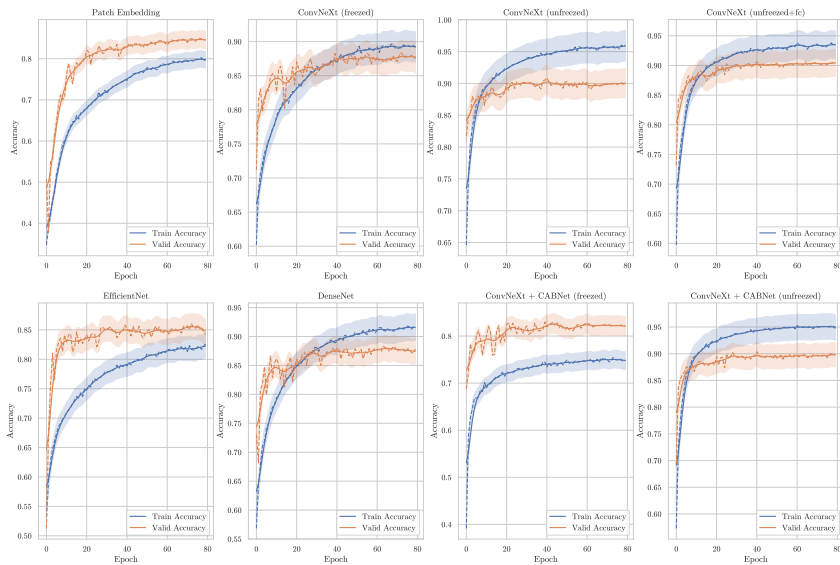
$$p_e = \frac{1}{n^2} \sum_i^k f_{i\cdot} \times f_{\cdot i}$$

- 在样本不均衡的情况下考虑小样本的分类情况
- 更好地评估分类精度

## 实验结果 I

| Backbone        | Parameters     | Classification Net | Depth | Accuracy      | Kappa         | Image   | Batch Size |
|-----------------|----------------|--------------------|-------|---------------|---------------|---------|------------|
| Patch-Embedding | not pretrained | CBAM + ConvMixer   | 8     | 0.6911        | 0.4618        | 480×480 | 128        |
| ConvNeXt        | unfreezed      | FC                 | 8     | <b>0.8399</b> | <b>0.7227</b> | 480×480 | 32         |
| ConvNeXt        | unfreezed      | CBAM + ConvMixer   | 4     | <b>0.8399</b> | 0.7218        | 480×480 | 32         |
| ConvNeXt        | unfreezed      | GAB + CAB          | -     | 0.8284        | 0.7001        | 480×480 | 32         |
| ConvNeXt        | freezed        | CBAM + ConvMixer   | 8     | 0.7526        | 0.5599        | 480×480 | 128        |
| ConvNeXt        | freezed        | GAB + CAB          | -     | 0.7023        | 0.4717        | 480×480 | 128        |
| EfficientNet    | unfreezed      | CBAM + ConvMixer   | 8     | 0.7571        | 0.5725        | 480×480 | 32         |
| DenseNet        | unfreezed      | CBAM + ConvMixer   | 8     | 0.7840        | 0.6190        | 480×480 | 32         |

## 实验结果 II



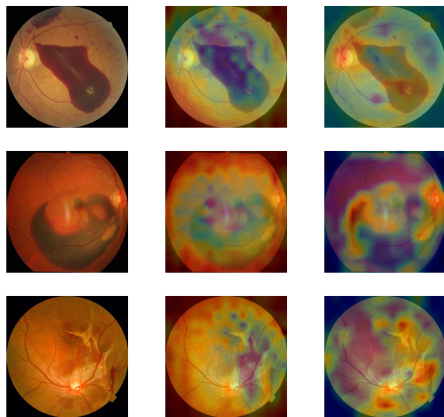
## 分析

- 从零开始训练的使用 Patch Embedding 的网络效果较差
- 冻结了参数的 ConvNeXt 加上后续网络和未冻结参数的 EfficientNet 和 DenseNet 加上后续网络结果相差不大
- 未冻结参数的 ConvNeXt 后接 CBAM+ConvMixer 训练时比纯全连接层略好，但测试结果相差并不大。
- 从 Kappa 结果根据 Landis 和 Koch 等人提出的标准 [4] 来看，ConvNeXt 不冻结参数训练得到的分类结果具有高度的一致性（Substantial）
- ConvNeXt 作为特征提取器后使用 FC, CBAM+ConvMixer 以及 CABNet 都能够取得较好的结果。



## CBAM 与 CABNet 注意力机制对比

图为以 ConvNeXt 为 Backbone 的网络下使用 CBAM + ConvMixer 和 CABNet 中注意力机制得到的结果。可以直观地发现 GAB+CAB 对原始图片中特征的刻画的确是更加准确的。



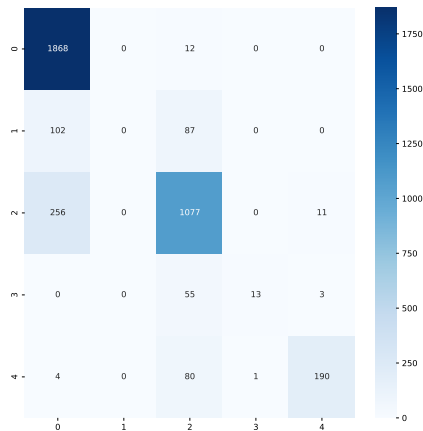
(a) 原始图片

(b) CBAM

(c) GAB+CAB

# 混淆矩阵

使用 ConvNext+CBAM+ConvMixer 模型在 test 集上得到的混淆矩阵如右图。



## 参考文献 I

- [1] Jacob Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46, 1960.
- [2] Along He, Tao Li, Ning Li, Kai Wang, and Huazhu Fu. Cabnet: Category attention block for imbalanced diabetic retinopathy grading. *IEEE Transactions on Medical Imaging*, 40(1):143–153, 2021.
- [3] Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q. Weinberger. Densely connected convolutional networks, 2016.
- [4] J. Richard Landis and Gary G. Koch. The measurement of observer agreement for categorical data. *Biometrics*, 33(1):159–174, 1977.
- [5] Tao Li, Yingqi Gao, Kai Wang, Song Guo, Hanruo Liu, and Hong Kang. Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening. *Information Sciences*, 501:511 – 522, 2019.
- [6] Min Lin, Qiang Chen, and Shuicheng Yan. Network in network, 2013.
- [7] Zhuang Liu, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, and Saining Xie. A convnet for the 2020s, 2022.
- [8] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization, 2017.
- [9] Mingxing Tan and Quoc V. Le. Efficientnet: Rethinking model scaling for convolutional neural networks. 2019.
- [10] Asher Trockman and J. Zico Kolter. Patches are all you need?, 2022.
- [11] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. Cbam: Convolutional block attention module, 2018.

# 谢谢!

---

实验报告及演示文稿使用 Overleaf 及  $\text{\LaTeX}$  编写  
报告及演示文稿中的示意图使用  $\text{\LaTeX}$ , TikZ 以及 draw.io 完成