



南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

本科生毕业设计（论文）

题 目： 面向眼科影像识别的临床像素校准
 网络模型研究

姓 名： 赵冀鲁

学 号： 11912821

系 别： 计算机科学与工程系

专 业： 计算机科学与技术

指导教师： 刘江 教授

2023 年 5 月 8 日

诚信承诺书

1. 本人郑重承诺所呈交的毕业设计(论文),是在导师的指导下,独立进行研究工作所取得的成果,所有数据、图片资料均真实可靠。

2. 除文中已经注明引用的内容外,本论文不包含任何其他人或集体已经发表或撰写过的作品或成果。对本论文的研究作出重要贡献的个人和集体,均已在文中以明确的方式标明。

3. 本人承诺在毕业论文(设计)选题和研究内容过程中没有抄袭他人研究成果和伪造相关数据等行为。

4. 在毕业论文(设计)中对侵犯任何方面知识产权的行为,由本人承担相应的法律责任。

作者签名: 赵冀鲁

2023 年 5 月 8 日

面向眼科影像识别的临床像素校准网络模型研究

赵冀鲁

(计算机科学与工程系 指导教师：刘江教授)

[摘要]：近年来，眼科影像检查已经成为了眼病筛查和诊断的常用手段。研究人员在计算机辅助诊断领域取得了诸多进展，各种基于卷积神经网络的模型被用于眼科影像处理。然而，很少有研究工作将眼科影像中的临床特征纳入到卷积神经网络的特征表示中，从而提升网络在眼疾上的分类性能和模型在决策过程中的可解释性。为此，本文提出了一个简洁高效的临床像素特征重新校准模块 (clinical pixel feature recalibration block, CPF)，旨在利用临床特征的潜力来提升网络在眼病上的分类能力。其利用跨通道池化将特征图中每个空间位置的临床像素特征提取出来，然后通过临床融合对每个空间位置的特征进行重新校准，从而将临床特征间的相对重要性融入到卷积神经网络的特征图中。本文在两个公开数据集 (OCTMNIST 和 RETINAMNIST) 和 CIFAR 数据集上进行了综合实验，实验结果显示：相比于先进的注意力方法，CPF 拥有优越的性能。此外，本文通过深入的临床特征权重可视化分析来增强本文模型在决策过程中的可解释性。

[关键词]： 眼科图像； CPF； 可解释性； 注意力； 分类

[ABSTRACT] : Ophthalmic image examination has become a commonly-acknowledged way for ocular disease screening and diagnosis. Clinical features extracted from ophthalmic images play different roles in affecting clinicians making diagnosis results, but how to incorporate these clinical features into convolutional neural network (CNN) representations has been less studied. In this paper, we propose a simple yet practical module, \textit{Clinical Pixel Feature Recalibration Module (CPF)}, aiming to exploit the potential of clinical features to improve the ocular disease recognition performance of CNNs. CPF first extracts clinical pixel features from each spatial position of all feature maps by clinical cross-channel pooling, then estimates each spatial position recalibration weight in a pixel-independent clinical fusion. By infusing the relative importance of clinical features into feature maps at the pixel level, CPF is supposed to enhance the representational ability of CNNs. Our CPF is easily inserted into existing CNNs with negligible overhead. We conduct comprehensive experiments on two publicly available ophthalmic image datasets and CIFAR datasets, and the results show the superiority and generation ability of CPF over advanced attention methods. Furthermore, this paper presents an in-depth weight visualization analysis to investigate the inherent behavior of CPF, aiming to improve the interpretability of CNNs in the decision-making process.

[Keywords]:Ophthalmic image; CPF; interpretability; attention;
classification

目录

1. 绪论.....	1
2. 相关工作.....	2
2.1 计算机辅助眼病诊断.....	2
2.2 注意力机制.....	2
3. 方法.....	4
3.1 临床跨通道池化.....	4
3.2 临床融合.....	4
3.2 算法部署与诊断系统.....	5
4. 数据集与实验设定.....	5
4.1 数据集.....	5
4.2 实验设定.....	6
5. 数据结果与讨论.....	8
5.1 结果对比.....	8
5.2 可视化与可解释性.....	9
5.3 消融实验.....	10
6. 总结.....	12
参考文献(宋体四号, 加粗).....	13
致谢(宋体四号, 加粗).....	15

1. 绪论

根据世界卫生组织的数据显示[1]，全球至少有 22 亿人视力受损，其中，至少有 10 亿人的视力损伤本来可以有预防或治愈的可能。而作为人口大国，中国是世界上视力损伤患者最多的国家之一。因此，大力发展眼保健服务是非常有必要的。而眼科影像检查是眼部疾病诊断的常用方法，在过去的很多年中，眼科影像检查显著影响了各种眼疾的诊断，并为青光眼、糖尿病视网膜病变和老年性黄斑变性等疾病的治疗提供了指导。

近年来，研究人员在用于眼科影像识别的计算机辅助诊断领域取得了很大进展。Zhang[2]等人提出了一种自适应特征挤压网络来对白内障和视网膜疾病进行分类。Fu[3]等人提出了一个多级深度网络来检测角度闭合。Kurma[4]等人提出了一种用于糖尿病视网膜病变分类的混合深度神经网络。Raghavendra[5]等人提出了一个深度卷积神经网络（CNN）来识别青光眼。然而，这些方法尚未充分利用临床特征的潜力来增强深度 CNN 的特征表征。

为了解决这个问题，研究人员将注意力机制纳入到 CNN 中，进而提升其表征能力。其中，SE[6]模块可以强调重要通道并通过在通道之间构建长期依赖关系来抑制冗余通道。ECA[7]模块可以对通道之间的局部范围依赖关系进行建模。CCA[8]模块则可以自适应地估计每个通道的相对重要性。我们发现这些注意力方法通常应用空间池化方法来聚合空间维度的空间上下文特征，然而这些空间上下文特征是全局统计信息，临床特征则通常来源于局部统计信息。因此，我们提出了这样一个问题：是否可以通过局部统计信息提取出临床像素特征，利用它们来增强 CNN 的表征能力？

基于这个问题，本文提出了一个高效的注意力模块，即临床像素特征重新校准模块（CPF），该模块通过像素特征重新校正的方式将临床像素特征明确地融入到 CNN 的特征图表示中。CPF 自适应地估计每个临床像素特征的相对重要性，然后独立地重新校准空间位置，允许 CNN 突出病变区域并忽略冗余区域。图 1 介绍了 CPF 的一般框架，由两个主要组成部分组成：临床跨通道池化模块和临床融合模块。临床跨通道池化能够提取沿着通道轴的每个空间位置的临床像素特征。随后是临床融合，该融合在通道维度上动态融合临床像素特征，生成不同空间位置的注意力权重，用于重新校准空间位置。为了证明我们的 CPF 的有效性和效率，

我们在两个公开可用的医学图像数据集（OCTNIST 和 RETINAMNIST[9]）和 CIFAR 数据集上进行了广泛的实验。结果表明，与最先进的注意力方法相比，CPF 在性能和模型复杂性之间实现了更好的平衡。例如，在 OCTMIST 数据集上，我们的 CPF 在减少了 10.8% 的参数的前提下，在准确性方面优于 SE 7.6%。

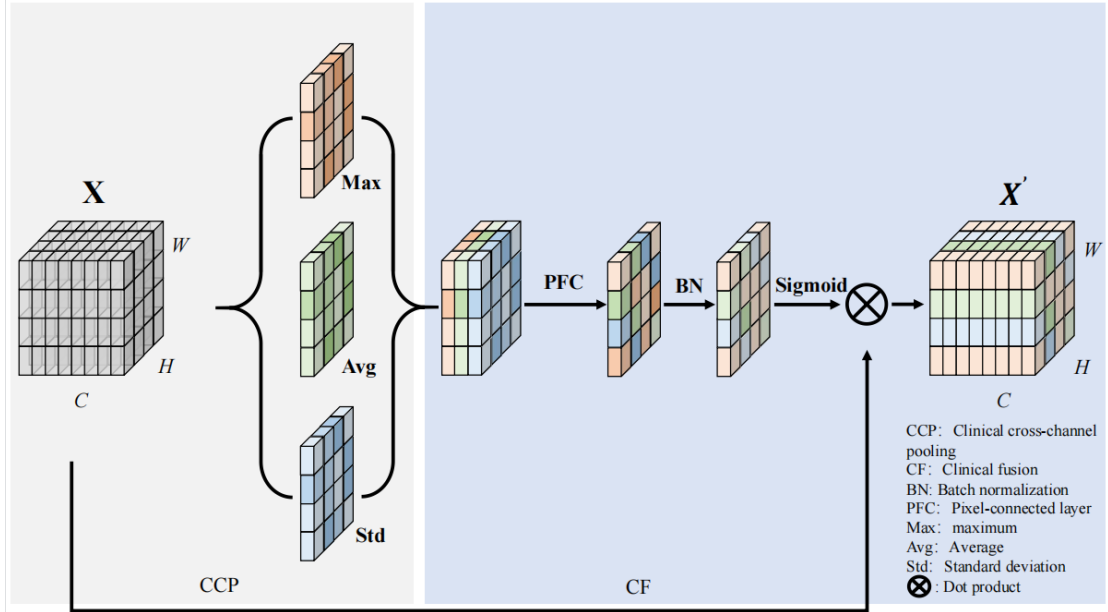


图 1 临床像素特征重新校准模块（CPF）。该模块有两个主要组成部分：临床跨通道池化（CCP）和临床融合（CF）。

2. 相关工作

2.1 计算机辅助眼病诊断

近年来，深度神经网络在计算机辅助诊断方面取得了巨大进步。Perdomo[10]等人提出了一种专门的 CNN 来分类糖尿病相关的视网膜疾病。Hao[11]等人提出一种用于评估前房角的混合变异感知网络。Xu[12]等人提出了一个全局局部 CNN 以识别基于眼底的白内障。Li[13]等人设计了一个用于白内障诊断的无注释恢复网络。Das[14]等人开发了一种多尺度深度特征融合网络对黄斑病变进行分类。Fu[3]等人提出了一种用于自动闭角检测的深度学习系统。Zhang[15]等人提出了一个基于区域的集成和重新校准网络用于白内障分类。然而，这些方法侧重于设计复杂的 CNN 模型，很少融合 CNN 的临床特征以提高性能。

2.2 注意力机制

注意力机制已证明它们显著提高了深度神经网络在各种学习任务中的性能。

目前的注意力机制大致可以分为三种类型：通道注意力、空间注意力以及两者的结合。

Squeeze-and-excitation networks (SE) [6]是最具代表性的通道注意方法之一，它可以强调信息通道并抑制冗余通道。SENet 主要有挤压模块和激活模块两部分组成。挤压模块通过全局平均池化捕获全局空间信息。激活模块通过使用全连接层和非线性层来捕获通道间的关系并输出注意力向量，最后通过注意力向量对输入特征的每一个通道进行重新校准。SENet 在强调重要通道信息忽视冗余通道信息方面表现优越。但是，SENet 也有相应的缺陷。由于在捕获空间信息时 SENet 只使用了简单的全局平均池化，因此 SENet 缺乏捕获高阶统计信息的能力。

Gather-excite (GE) [16]可以捕获远程空间上下文信息，是非常具有代表性的空间注意力方法。GENet 结合了空间汇集和激活操作。首先，它从特征图中的每个空间邻域提取上下文信息，对不同空间位置之间的关系进行建模。然后它通过注意力图对特征图中空间位置上的特征信息进行重新校准。GENet 跟 SENet 一样都是轻量级的网络，它可以在抑制冗余特征的同时强调重要特征。

Non-local (NL-Net) [17]引入了一个非局部 (NL) 块，这是一种基于空间注意力的方法，通过自注意力机制捕获空间位置之间的长距离依赖关系。众所周知，卷积神经网络可以通过卷积层的堆叠来增大感受野，但是特定层的感受野仍旧有限，使得网络缺乏对全局信息的掌握。而 NL-Net 则通过自注意力机制捕获空间位置之间的长距离依赖关系，使网络在掌握局部信息的同时兼顾全局信息。但是长距离依赖在医学影像中并非特别适用，因为在很多情况下，我们需要网络去关注病理区域而忽视冗余区域。

External attention (EA) [18]利用两个线性层来估计特征图空间位置的相对重要性。在此之前，自注意力机制只考虑单个样本中的相关性，忽略不同样本之间的潜在关系。而外部注意力则通过使用双重归一化代替 softmax，从而使网络在捕获不同样本相关性的同时减少计算量。

Convolutional Block Attention Module (CBAM) [19] 和 Bottleneck Attention Module (BAM) [20]结合了通道注意力和空间注意力方法。CBAM 在通道和空间两个独立的维度进行注意力图的推断，同时通过全局池化对全局信息进行聚合。CBAM 由通道注意力模块和空间注意力模块组成，通道注意力模块将特

征图在空间维度上进行聚合，而空间注意力模块则基于通道对特征图进行压缩。通过将空间注意力和通道注意力进行解耦，CBAM 实现了计算效率的提升，但是由于其空间注意力模块仍旧采用卷积来生成注意力图，使 CBAM 的空间子模块会受到感受野的限制。BAM 通过使用扩张卷积有效地扩大了自己的感受野，使其在强调或抑制通道和空间维度的信息时具有良好的表现。但是，BAM 仍旧缺乏捕获长距离依赖和编码跨域关系的能力。

与现有方法相比，本文利用临床特征的潜力来重新制定空间位置重新校准，而无需对空间位置之间的远程依赖性进行建模。我们率先设计了一种临床跨通道池化方法，在通道维度上提取三个临床像素特征。

3. 方法

3.1 临床跨通道池化

现有的方法广泛应用空间池化方法来提取全局上下文特征。然而，很少有研究利用跨通道池化方法来提取像素上下文特征。因此，我们提出了一个设计良好的临床跨通道池化 (CCP) 来提取沿通道轴的三个临床像素特征：平均值 (Avg)、最大值 (Max) 和标准偏差 (Std)。具体来说，每个空间位置 $p(i, j)$ 的三个临床像素特征可以通过以下方式计算：

$$\begin{aligned}\mu(i, j) &= \frac{1}{C} \sum_{k=1}^C x(k, i, j) \\ m(i, j) &= \max_{(i, j)} (x(1, i, j), \dots, x(C, i, j)) \\ \sigma(i, j) &= \sqrt{\frac{1}{C} \sum_{k=1}^C (x(k, i, j) - \mu(i, j))^2} \\ t(i, j) &= [m(i, j), \mu(i, j), \sigma(i, j)]\end{aligned}$$

其中， $\mu(i, j)$, $m(i, j)$, and $\sigma(i, j)$ 分别表示空间位置 $p(i, j)$ 上的平均、最大值、标准差特征。

3.2 临床融合

在临床跨通道池化之后，我们提出了一个临床融合算子，将临床像素特征转换为空间位置注意力权重。临床融合算子可以调整与每个空间位置相关的临床像

素特征的相对重要性，以便关注或忽略它们。为了实现这个目的，我们构建了一个简单而实用的像素连接层 (PFC)、批量归一化层 (BN) 和 sigmoid 激活函数的组合。给定临床像素特征表示 $T \in \mathbb{R}^{3 \times H \times W}$ 作为输入，特征融合算子使用可学习参数 $\mathbb{R}^{3 \times H \times W}$ 执行逐像素编码：

$$Z = W \cdot T$$

$Z \in \mathbb{R}^{N \times HW}$ 是每个像素位置的编码临床像素特征。然后我们应用 BN 来促进训练和 sigmoid 函数作为门控机制来生成空间位置注意力权重：

$$G = \sigma(\text{BN}(Z))$$

其中 G 表示空间位置注意力权重， σ 表示 sigmoid 函数。

最后，将原始输入张量 X 以点积方式乘以注意力权重 G 进行增强，因此输出张量 $X' \in \mathbb{R}^{C \times H \times W}$ 可以通过以下公式得到：

$$X' = X \cdot G$$

3.3 算法部署与诊断系统

我们采用了 ResNet18 和 ResNet50 来体现 CPF 相较于先进的注意力方法的优势，以此来证明临床像素特征在提升 CNN 表征能力上的潜力。我们将 CPF 模块和残差块结合形成残差-CPF 模块。分类器和损失函数我们分别采用了 softmax 函数和交叉熵损失函数。

与此同时，为了能够更加直观地展示我们的网络，我们搭建了一个简单的诊断系统。我们的诊断系统采用了前后端分离的模式进行开发。前端使用基于 python 的 PyQt 框架，后端采用了基于 python 的 django 框架。系统界面如图 2 所示，我们的系统会展示原始的分类图像以及分类结果。与此同时，系统还会以热力图的形式对图像进行分析，从而给临床医生更加精细化的辅助。

4. 数据集和实验设定

4.1 数据集

4.1.1 OCTMNIST

OCTMNIST 是一个公开的视网膜疾病数据集，该数据集包含四种视网膜疾病类型：正常、玻璃膜疣、脉络膜新生血管 (CNV) 和糖尿病性黄斑水肿 (MDE)。它包含 109,309 张光学相干断层扫描 (OCT) 图像，图像大小为 28×28 。为了

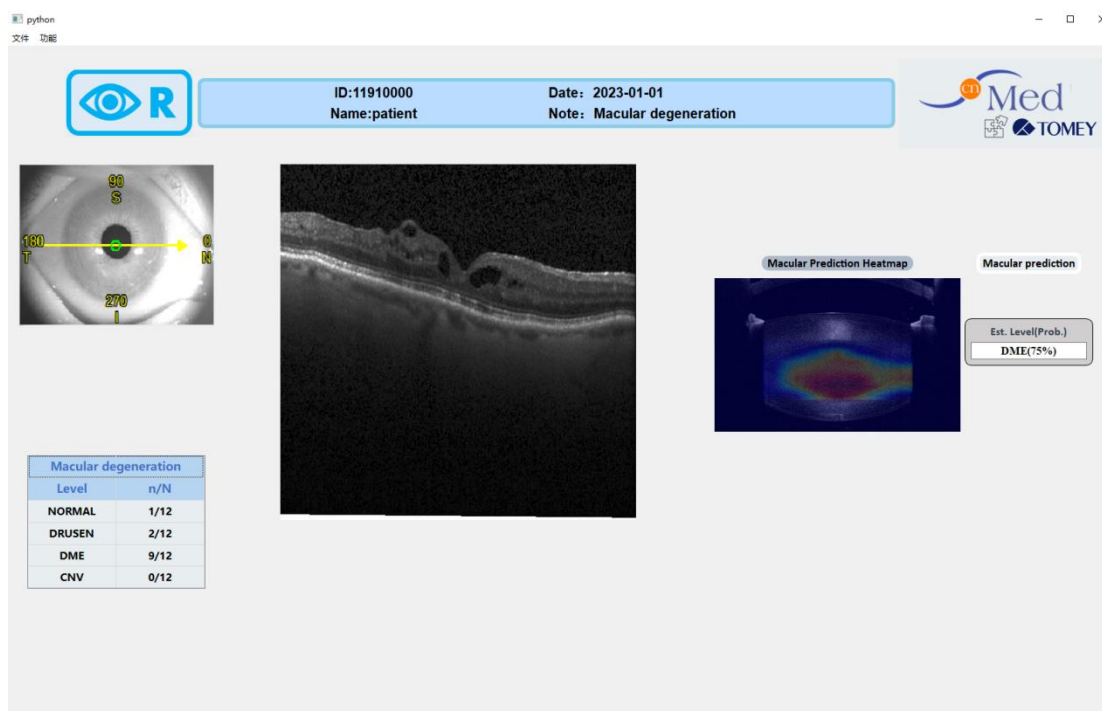


图 2 诊断系统界面

确保对比的公平性，我们遵循相同的数据集拆分方法 [9]，它具有三个子集：训练集、验证集、和测试集。

4.1.2 RETINAMNIST

RETINAMNIST 是一个用于糖尿病视网膜病变分级的眼底图像数据集，包含包含 1600 张图像（图像大小为 28×28 ）。文献[9]将其分为三个子集：训练集（1080）、验证集（120）和测试集（400），本文采用这种数据集拆分方法。

4.2 实验设定

4.2.1 评价指标

在本文中，我们使用了以下指标来验证网络的整体性能：准确率（ACC）、F1 和 Kappa 系数。

表 1 以 ResNet18 和 ResNet50 为主干时, CPF 和最先进的注意力方法在 OCTMNIST 数据集上的性能比较和复杂度比较。

Method	ACC	F1	Kappa	Params
ResNet18	76.20	73.61	68.27	11.18M
+ CBAM	76.60	73.90	68.80	11.26M
+ BAM	77.00	74.15	69.33	11.18M
+ SE	72.50	67.79	63.33	11.27M
+ GE	78.60	76.07	71.47	11.45M
+ NL	75.60	72.26	67.47	11.91M
+ CPF	79.40	76.81	72.53	11.18M
ResNet50	78.00	75.82	70.67	23.50M
+ CBAM	77.80	75.11	70.40	26.04M
+ BAM	78.10	75.10	70.81	23.70M
+ SE	72.50	67.69	63.33	26.05M
+ GE	78.30	75.44	71.07	27.46M
+ NL	65.80	58.78	54.4	44.76M
+ CPF	80.10	77.67	73.47	23.52M

4.2.2 基线

我们使用了 CBAM, BAM, SE, GE 和 NL 去测试我们的 CPF 的有效性。

4.2.3 实现细节

我们使用 Pytorch 实现我们的 CPF 和其他先进的注意力方法。在训练过程中, 我们使用随机梯度下降优化器对模型进行训练, 初始学习率和训练代数分别设定为 0.0025 和 150, 学习率每 20 代会缩减百分之五。此外, 我们遵循标准的数据增强方法来增强训练数据, 例如随机翻转方法和随机裁剪方法。我们的全部实验在配备了 NVIDIA A6000 的工作站上进行。

表 2 以 ResNet18 和 ResNet50 为主干时, CPF 和最先进的注意力方法在 RETINAMNIST 数据集上的性能比较和复杂度比较。

Method	ACC	F1	Kappa	Params
ResNet18	52.00	31.99	29.45	11.17M
+ CBAM	52.25	34.81	31.73	11.27M
+ BAM	53.00	38.27	32.93	11.19M
+ SE	51.00	36.43	31.05	11.27M
+ GE	50.00	33.27	29.46	11.45M
+ NL	51.75	34.03	31.51	11.92M
+ CPF	53.50	38.30	33.85	11.18M
ResNet50	51.50	33.76	31.15	23.51M
+ CBAM	51.50	33.95	32.29	26.04M
+ BAM	50.25	33.08	29.78	23.69M
+ SE	47.75	31.65	26.16	26.05M
+ GE	53.25	35.51	33.04	27.45M
+ NL	52.50	30.21	30.75	44.77M
+ CPF	53.50	32.36	34.34	23.51M

5. 实验结果与讨论

5.1 结果对比

5.1.1 在 OCTMNIST 数据集上的性能对比

我们以 ResNet18 和 ResNet50[21]为主干, 在 OCTMNIST 数据集上将我们的 CPF 和先进的注意力方法进行了性能对比, 表 1 展示了我们的实验结果。从实验结果我们可以看出, 在模型复杂度较低的情况下, CPF 在性能方面比其他先进的注意力方法要优越。我们可以看到我们的 CPF 几乎与 ResNet18 共享所有参数, 但在性能上我们的 CPF 在三个评价指标上均超过 ResNet18 3%。值得注意的是, 在以 ResNet50 为主干的实验中, 尽管 NL 的参数量要比 CPF 大 90%以上, 但是 CPF 的

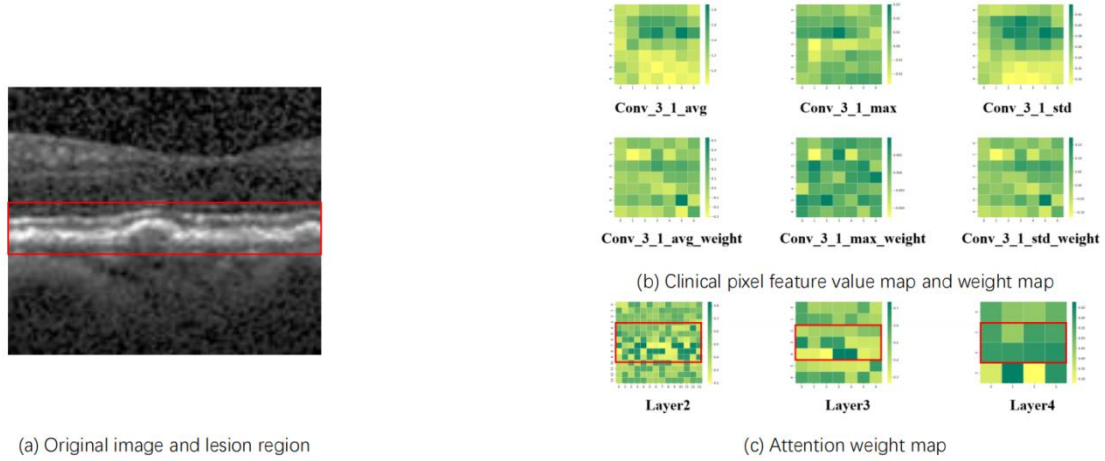


图 3 (a) 原始图像和病理区域（被红框标注出来）；(b) 在 OCTMNIST 上基于 ResNet50 的 CPF 在第三阶段的临床像素特征值分布和权重分布；(c) 在 OCTMNIST 数据集上基于 ResNet50 的 CPF 在三个阶段的注意力权重分布。

性能在准确率上比 NL 高 14%。而在参数量少于 CBAM 和 BAM 的情况下，我们的 CPF 在性能上提升了 2%。这证明了不同临床像素上下文特征的临床融合方式相对于局部融合方式的优越性。

5.1.2 在 RETINAMNIST 数据集上的性能对比

表 2 展示了我们的 CPF 和其他先进的注意力方法在视网膜图像上的分类结果。我们可以看到，在采用 ResNet18 和 ResNet50 作为主干的情况下，CPF 在性能和模型复杂性之间取得了比先进的注意方法更好的平衡。在以 ResNet18 为主干的情况下，CPF 在 F1 分数上优于 GE，在类似参数下，kappa 值优于 4%。值得注意的是，CPF 比基于 ResNet18 和 ResNet50 的 NL 获得了超过 2.34% 的 kappa 增益和 1.75% 的精度 (acc) 增益，尽管 NL 的参数量比 CPF 要大 6.6%。我们的实验结果表明将基于像素的上下文信息合并到特征图中这一方法是非常有效的。

5.2 可视化和可解释性

5.2.1 临床像素特征值图和权重图可视化

图 3(b) 展示了 OCT 图像在 ResNet50 的高层阶段的三个临床像素特征值图（第一行）及其对应的像素特征权重图（第二行）。我们可以观察到 avg 和 std 像素特征的像素特征值分布相似：上部区域的像素特征值大于下部区域的像

表 3 以 ResNet50 为主干时，不同池化方式在 OCTMNIST 数据集上的性能比较。

Method	ACC	F1	Kappa
ResNet50	78.00	75.82	70.67
+ Max	78.40	75.94	71.21
+ Avg	79.10	76.90	72.13
+ Std	78.90	77.06	71.87
+ Max + Avg	79.50	77.34	72.67
+ Max + Std	79.60	77.48	72.81
+ Avg + Std	79.50	77.60	72.67
+ Max + Std + Avg (CPF)	80.10	77.67	73.47

素特征值。相反，边界区域的最大像素特征值大于中心区域。此外，三个临床像素特征的像素特征权重分布与临床像素特征图不同，表明 CPF 可以在每个像素位置自适应地为临床像素特征设置不同的权重，这有利于引导 CNN 对病变区域的关注。

5.2.2 注意力权重图可视化

图 3(c) 展示了 ResNet18 三个阶段的 CPF 注意力权重图。可以看出，当网络深入时，注意力权重值之间的差异变得明显。值得注意的是，CPF 在高阶段的注意力权重分布与病变分布一致，证明我们的 CPF 可以让 CNN 强调重要区域并忽略冗余区域。

5.3 消融实验

5.3.1 临床跨通道池化

我们使用 Resnet50 作为骨干网络，在 OCTMNIST 数据集上将我们提出的临床跨通道池化方式与其他池化方式进行了对比，表 3 展示了不同池化方式完成分类任务的结果。通过观察实验结果我们可以发现，单独的池化方法可以给网络带来性能增益，而两种池化方法的结合可以实现更高的性能提升。我们提出的临床跨通道池化在性能上相比其他池化方法达到了最优，而这证明了提取三种临床像素特征这一方法的优越性。

5.3.2 验证

表 4 以 ResNet18 和 ResNet50 为主干时, CPF 和最先进的注意力方法在 CIFAR 数据集上的性能比较和复杂度比较。

Method	CIFAR-10		CIFAR-100	
	ACC	Params	ACC	Params
ResNet18	93.02	11.17M	74.56	11.22M
+ CBAM	95.19	11.26M	77.82	11.31M
+ BAM	95.20	11.20M	78.09	11.24M
+ SE	94.84	11.27M	75.19	11.32M
+ GE	95.14	11.55M	77.64	11.56M
+ NL	93.38	11.96M	71.97	12.01M
+ CPF	95.26	11.18M	78.51	11.23M
ResNet50	93.62	23.52M	78.51	23.71M
+ CBAM	95.70	26.05M	80.13	26.24M
+ BAM	95.54	23.88M	80.00	24.06M
+ SE	95.35	26.06M	79.28	26.64M
+ GE	95.44	27.87M	79.54	28.05M
+ NL	94.00	46.17M	72.15	46.36M
+ CPF	95.87	23.53M	80.26	23.71M

为了展示我们提出的模块在自然图像上的泛化能力,表 4 展示了我们的 CPF 和其他先进的注意力方法在 CIFAR 数据集上的结果。CIFAR 数据集由 CIFAR-10 (10 种标签) 和 CIFAR-100 (100 种标签) 组成, 它们都包含 50000 张大小为 32x32 像素的训练图像和 10000 张大小为 32x32 像素的测试图像。我们将训练代数和批量大小分别设置为 200 和 128。初始学习率设定为 0.1, 每 40 代学习率缩减百分之十。通过观察表格,我们可以清楚地发现,我们的 CPF 达到了最好的分类性能。与此同时,相较于其他先进的注意力方法,我们的 CPF 在 ResNet18 和 ResNet50 的基础上新引入的参数数量最少。这些结果向我们证明了 CPF 模块的有效性并不仅仅局限于医学图像。

6. 总结

本文提出了一种轻量级但高效的架构单元，即临床像素特征重新校准模块 (CPF)，它根据临床像素特征的相对重要性动态重新校准空间位置响应。通过将临床像素特征注入特征图的空间位置，CPF 显著提高了 CNN 的表征能力。对两个眼科图像数据集和 CIFAR 数据集的广泛实验表明，本文的 CPF 比最先进的注意力方法在有效性和效率之间保持更好的权衡。此外，本文验证了 CPF 引导 CNN 关注重要空间位置并抑制冗余位置的能力。

参考文献

- [1] ORGANIZATION W H, et al. World report on vision[J]., 2019.
- [2] ZHANG X, XIAO Z, HIGASHITA R, et al. Adaptive feature squeeze network for nuclear cataract classification in AS-OCT image[J/OL]. JBI, 2022, 128: 104037. <https://www.sciencedirect.com/science/article/pii/S1532046422000533>. DOI: <https://doi.org/10.1016/j.jbi.2022.104037>.
- [3] FU H, XU Y, LIN S, et al. Angle-closure detection in anterior segment OCT based on multilevel deep network[J]. IEEE transactions on cybernetics, 2019, 50(7): 3358–3366.
- [4] KUMAR G, CHATTERJEE S, CHATTOPADHYAY C. DRISTI: a hybrid deep neural network for diabetic retinopathy diagnosis[J/OL]. Signal, image and video processing, 2021, 15(8): 1679–1686. <https://europepmc.org/articles/PMC8051933>. DOI: 10.1007/s11760-021-01904-7.
- [5] RAGHAVENDRA U, FUJITA H, BHANDARY S V, et al. Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images[J]. Information Sciences, 2018, 441: 41–49.
- [6] HU J, SHEN L, SUN G. Squeeze-and-Excitation Networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018: 7132–7141. DOI: 10.1109/CVPR.2018.00745.
- [7] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks[C]//CVPR. 2020: 11531–11539. DOI: 10.1109/CVPR42600.2020.01155.
- [8] ZHANG X, XIAO Z, HU L, et al. CCA-Net: Clinical-awareness attention network for nuclear cataract classification in AS-OCT[J]. KBS, 2022, 250: 109109.
- [9] YANG J, SHI R, WEI D, et al. MedMNIST v2: A Large-Scale Lightweight Benchmark for 2D and 3D Biomedical Image Classification[J]. arXiv preprint arXiv:2110.14795, 2021.
- [10] PERDOMO O, RIOS H, RODRÍGUEZ F J, et al. Classification of diabetes-related retinal diseases using a deep learning approach in optical coherence tomography[J]. Computer methods and programs in biomedicine, 2019, 178: 181–189.
- [11] HAO J, LI F, HAO H, et al. Hybrid variation-aware network for angle-closure assessment in As-Oct[J]. TMI, 2021, 41(2): 254–265.
- [12] XU X, ZHANG L, LI J, et al. A Hybrid Global-Local Representation CNN Model for Automatic Cataract Grading[J]. IEEE Journal of Biomedical and Health Informatics, 2020, 24(2): 556–567. DOI: 10.1109/JBHI.2019.2914690.
- [13] LI H, LIU H, HU Y, et al. An Annotation-Free Restoration Network for Cataractous Fundus Images[J]. TMI, 2022, 41: 1699–1710.
- [14] DAS V, DANDAPAT S, BORA P K. Multi-scale deep feature fusion for automated classification of macular pathologies from OCT images[J/OL]. Biomedical Signal Processing and Control, 2019, 54: 101605. <https://www.sciencedirect.com/science/article/pii/S1746809419301867>. DOI: <https://doi.org/10.1016/j.bspc.2019.101605>.

- [15] ZHANG X, XIAO Z, FU H, et al. Attention to region: Region-based integration-and-recalibration networks for nuclear cataract classification using AS-OCT images[J]. Medical Image Analysis, 2022: 102499.
- [16] HU J, SHEN L, ALBANIE S, et al. Gather-Excite: Exploiting Feature Context in Convolutional Neural Networks[C]//NIPS' 18: Proceedings of the 32nd International Conference on Neural Information Processing Systems. Montréal, Canada: Curran Associates Inc., 2018: 9423-9433.
- [17] WANG X, GIRSHICK R, GUPTA A, et al. Non-local Neural Networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018: 7794-7803. DOI: 10.1109/CVPR.2018.00813.
- [18] GUO M H, LIU Z N, MU T J, et al. Beyond Self-Attention: External Attention Using Two Linear Layers for Visual Tasks[J]. T-PAMI, 2022: 1-13. DOI: 10.1109/TPAMI.2022.3211006.
- [19] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]//ECCV. 2018: 3-19.
- [20] PARK J, WOO S, LEE J Y, et al. A simple and light-weight attention module for convolutional neural networks[J]. IJCV, 2020, 128(4): 783-798.
- [21] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//CVPR. 2016: 770-778.

致谢

时光飞逝，日月如梭。随着毕业设计的完成，本人的大学生涯也迎来尾声。四年的求学之路，我经历了很多，也成长了很多。一路走来，有太多记忆值得珍藏，有太多人值得感谢。

首先，我要向我的学术导师刘江教授致以真诚的感谢。感谢您在我的大学生涯中给予我的诸多帮助和指导，您为人处世的原则和对待工作的热情让我受益良多，您的“快乐科研”让我对科研这条路有了新的认知，您带领的 iMED 团队让我感受到了人心的温暖。这些知识与情怀，将伴随我一生，成为我人生路上的珍贵财富。

我还要感谢组内的章晓庆博士给予我的帮助。感谢您一直以来的指导与宽容。感谢您在我面临困难时的慷慨帮助，感谢您在我犯错时的极尽包容。正是有您的细心指导，我才能慢慢走上科研的道路。

最后，我要感谢我的家人和朋友们。感谢我的家人一直以来的默默支持，这是我能够坚定前行的动力。也感谢我的朋友们的热心帮助，与你们一起的岁月于我而言弥足珍贵。