arXiv:2303.03817v1 [cs.CV] 7 Mar 2023

面向眼底图像的区域与空间感知异常检测

Jingqi Niu[1], *Shiwen Dong*[1] ,*Qinji Yu*[1],*Kang Dang*[2*], *Xiaowei Ding*[1,2*]

[1]中国上海上海交通大学[2]中国上海Voxelcloud Inc

摘要

近年来，异常检测在眼部疾病诊断中受到广泛关注。现有的大多数眼底图像异常检测研究在显著的视网膜结构（如血管、视杯和视盘）处得到的异常分数相对较高。本文提出一种面向眼底图像的区域与空间感知异常检测（ReSAD）方法，通过获取局部区域信息和长距离空间信息，减少正常结构中的假阳性。ReSAD迁移预训练模型以提取正常眼底图像的特征，并对像素级特征应用区域与空间感知特征组合模块（ReSC），以构建记忆库。在测试阶段，ReSAD利用记忆库将分布外样本判定为异常。在两个公开基准数据集上，我们的方法显著优于现有的眼底图像异常检测方法。

索引词— 异常检测、迁移学习、眼底图像

1. 引言

深度学习推动了计算机辅助疾病诊断的革命性发展 [1]。然而，现有的监督式深度学习模型通常需要大量训练数据，而这类数据很难获取，尤其是对于罕见疾病。因此，我们研究异常检测问题，仅依赖正常数据集构建能够表示健康受试者的模型，并识别患者中的任何异常。异常检测是各种医学图像分析应用中广泛研究的课题 [2, 3]; 然而，现有研究中仅有少数关注眼底图像中的异常检测。本文针对彩色眼底图像中的异常检测问题展开研究，由于病变尺寸较小且具有多样性，该问题极具挑战性。

我们的工作基于最先进的基于代表样本的异常检测方法。这类方法在训练阶段将正常图像映射到特定的特征空间，并在模型测试时将分布外样本检测为异常。与先前的工作类似 [4, 5], ，我们直接利用 ImageNet 分类得到的深度表征，无需进行额外的特征微调。为了使这些特征适用于眼底图像异常检测任务，我们提出了一种新颖的区域与空间感知特征组合（ReSC）模块。区域感知特征模块利用局部图像区域内的距离相关关系，而空间感知特征模块则构建贯穿整个眼底图像的长程空间关系。实验表明，我们的区域与空间感知特征组合模块能够显著减少误报，并提高异常检测质量。

我们提出的区域与空间感知异常检测方法（ReSAD）的整体流程如下。首先，ReSAD对预训练的ImageNet模型进行适配，以提取正常眼底图像的特征。为了减少假阳性，引入区域与空间感知特征组合模块，以融合局部区域信息和长程空间信息，并利用这些特征构建记忆库。在测试阶段，ReSAD计算测试特征与记忆库中特征之间的距离，以确定分布外异常样本。

贡献如下：(1) 我们提出了一种基于表征的区域与空间感知异常检测方法（ReSAD），利用局部区域信息和长程空间信息来减少假阳性并提升异常检测准确率。(2) 在IDRiD和ADAM数据集上，我们的方法在像素级AUC方面分别比最先进模型高出6.2%和8.7%。此外，仅使用10%的正常图像作为训练数据，我们的方法就能超越大多数最先进方法。

2. 相关工作

本节讨论当前针对眼底图像的最先进异常检测工作。

基于重建的方法广泛应用于自然图像和医学图像的异常检测中 [2]。fAno-GAN [6] 利用生成对抗模型通过重建学习正常数据的分布，并通过重建误差衡量异常分数。WDMT- Net [7] 提出了一种用于异常检测的多任务编码器-解码器网络，并采用权重衰减跳跃连接。该

* 通讯作者

基于重建的方法在正常视网膜结构中的重建误差相对较大，从而导致假阳性。

很少有研究探索基于表征的眼底图像异常检测方法。Deep-IF [8] 提出了一种迁移学习方法，并训练孤立森林模型进行异常检测。MDFSC [9]采用自编码器构建特征空间，并将多尺度稀疏编码用于异常检测。现有方法未考虑眼底图像的像素级特征，导致性能下降，也无法对视网膜病变进行可解释性分析。

与现有方法相比，本文提出的 ReSAD 数据效率更高，并能减少正常结构区域中的假阳性。

3. 方法

ReSAD 框架由三部分组成：特征提取（§3.1）、区域与空间感知记忆库构建（§3.2）和推理（§3.3）。完整流程如图 1 所示。

3.1. 特征提取

为了检测眼底图像中的异常，需要在训练阶段构建正常数据的分布。本文采用一个在 ImageNet 上 ϕ 预训练的模型，提取训练集中的正常图像特征。遵循先前的工作 [5], ，我们采用类似 ResNet 的模型作为特征提取器。

我们将输入图像表示为 $I_k \in \mathbb{R}^{h\times w\times 3}$ (k =1，.....，N)，其中 N表示训练图像的数量，h 和 w 分别表示输入图像的高度和宽度I_k。眼底图像异常检测需要细粒度的高级特征，以定位较小且具有语义相关性的视网膜病变，因此我们从预训练模型 ϕ的 i_{th} 阶段块 ϕ_i中提取多级特征图$f_{ik} \in \mathbb{R}^{h_i\times w_i\times c_i}$ (i = 1，2，3，4)，其计算方式如下：

$$\begin{aligned} f_{ik} &= \phi_i(f_{(i-1)k}), \\ f_{1k} &= \phi_1(I_k), \end{aligned} \tag{1}$$

为控制内存消耗并保持全面的特征表示，ReSAD 利用中间特征 f_{2k}、 f_{3k} 构建多级特征图F_k。为保持空间分辨率，ReSAD 将特征图 f_{3k} 上采样至与 f_{2k} 相同的空间形状，并将特征图 f_{2k} 与 f_{3k} 拼接，得到：

$$F_k = \mathrm{concat}(f_{2k}, \mathrm{T}_u(f_{3k})), \tag{2}$$

其中 $F_k \in \mathbb{R}^{h_2\times w_2\times c}$、 $c = c_2 + c_3$ 和 T_u 表示上采样变换。

3.2. 区域与空间感知记忆库构建

眼底图像的某些关系局限于特定区域，例如视盘杯。眼底图像的某些关系则与区域无关，结构相关性并不集中在某个区域，例如眼底血管（上方与下方眼底血管之间的相关性）。

为构建眼底图像中的关系，ReSAD 使用区域与空间感知特征组合模块（ReSC），为像素级特征提取局部区域信息和长程空间信息，并利用这些特征构建记忆库。区域感知特征模块可以提取视网膜结构区域中的距离相关关系。空间感知特征模块则可以构建眼底结构之间的空间关系，而不受其在空间维度上的距离影响。

首先，ReSC 使用固定的区域距离感知滤波器ψ_r 提取局部区域信息，其中 r 表示滤波器 ψ的尺寸。滤波器的权重会随其与中心的距离增加而线性衰减。区域感知特征图 R_k 可按如下方式计算：

$$R_k = \psi_r(F_k), \tag{3}$$

其中， $R_k \in \mathbb{R}^{h_2\times w_2\times c}$ 与 F_k具有相同的形状。

其次，ReSC 对特征图 F_k 执行空间自注意力操作，以获得空间感知特征图P_k。ReSC 首先将 F_k 的空间维度展平，并计算空间注意力图 $G_k = mul(F_k, F_k^{\mathsf{T}})$，其中 mul表示矩阵乘法， $G_k \in \mathbb{R}^{h_2w_2\times h_2w_2}$。在空间注意力图 G_k中，任意两个相似特征都可以相互促进，而不受两者之间空间距离的影响。空间感知特征图 P_k 可按如下方式计算：

$$P_k = F_k \circ softmax(G_k), \tag{4}$$

其中， $\circ$ 表示Hadamard积， $P_k \in \mathbb{R}^{h_2\times w_2\times c}$ 与 F_k具有相同的形状。

最后，我们可以通过以下方式计算 C_k：

$$C_k = P_k + R_k. \tag{5}$$

为了为每个像素级特征构建一个记忆库，Re-SAD 将区域与空间感知特征 $C_k^{(x,y)}$ 整合在一起，其中 $C_k^{(x,y)}$ 表示正常图像 I_k中位置（x， y）处的像素级特征。通过整合训练集中的所有图像在所有位置上的像素级特征，我们可以得到一个区域与空间感知记忆库 B，如下所示：

$$B = \bigcup_{k,x,y} C_k^{(x,y)}, k \in N, x \in h_2, y \in w_2, \tag{6}$$

其中， $C_k^{(x,y)}$ 表示正常图像 I_k在位置（x， y）处的像素级特征。 B 包含 $N \times h_2 \times w_2$ 特征，且

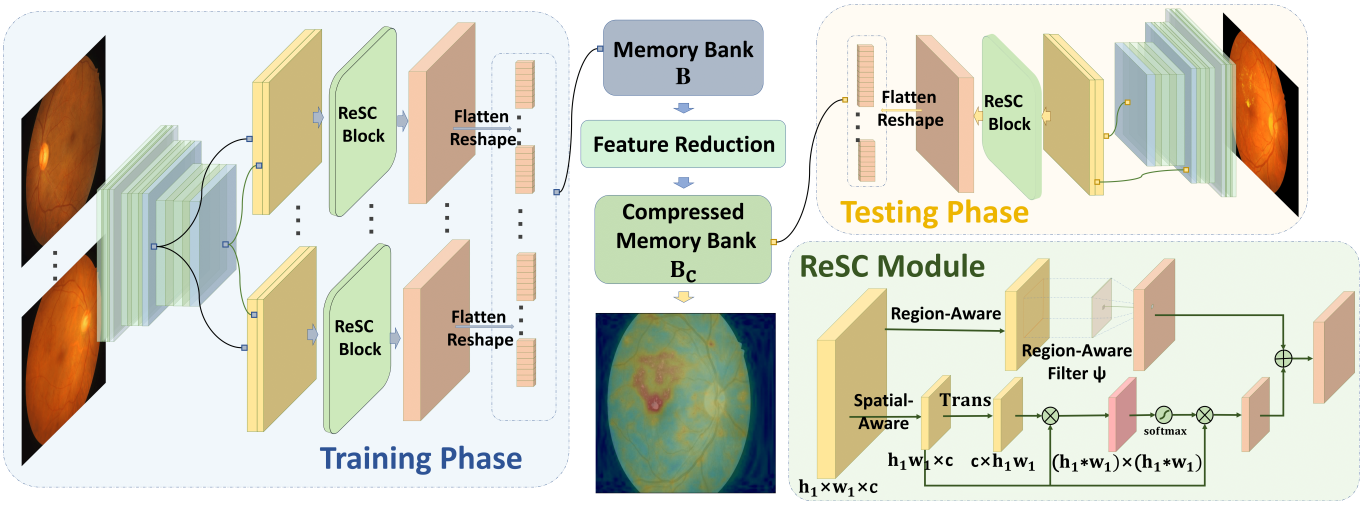

图1。ReSAD训练流程。首先，ReSAD通过预训练模型提取多层级特征图。其次，ReSC模块将局部区域信息和长程信息引入像素级特征，并构建记忆库。在测试阶段，ReSAD通过记忆库将分布外样本判定为异常。

每个特征有 c 个通道，其中包含局部区域和长距离空间信息。

此外，为了减小 B的大小，ReSAD遵循 [5],，采用 [10] 提出的迭代贪心近似方法来获取压缩记忆库 B_c，该记忆库保留了大部分信息，同时内存消耗更低。

3.3. 推理

对于单张图像 $I_{test} \in \mathbb{R}^{h\times w\times 3}$，我们可以根据公式5计算其区域与空间感知特征图 C_{test}，并计算测试图像 $d^{x,y}_{test}$ 中每个像素级特征的距离，如下所示：

$$d^{x,y}_{test} = min_{b\in B_c}\|b - c^{x,y}_{test}\|_2, \tag{7}$$

其中，b，$c^{x,y}_{test} \in \mathbb{R}^c$，$b$ 表示压缩记忆库 B_c中的像素级特征向量；$c^{x,y}_{test}$ 表示图像 I_{test} 在位置（x，y）处的区域与空间感知特征。通过整合空间位置上的 $d^{x,y}_{test}$，我们可以得到完整的距离图 D_{test}。

将 D_{test} 插值至原始图像的空间尺寸 I_{test}，即可得到像素级异常分数图$A_{test} \in \mathbb{R}^{h\times w}$。此外，我们将异常分数图 A_{test} 中的最大分数用作图像级异常分数a_{test}。

4. 实验

4.1. 数据集

我们在两个公开基准数据集 IDRiD 和 ADAM 上评估了所提出的方法。

IDRiD：印度糖尿病视网膜病变图像数据集 [11] (IDRiD)包含 134 张用于训练的正常图像、32 张正常图像和 81 张用于测试的异常图像。每张异常图像都包含四种不同病变的掩码，包括微动脉瘤（MA）、软性渗出物（SE）、硬性渗出物（EX）和出血（HE）。

ADAM：年龄相关性黄斑变性自动检测挑战赛数据集（ADAM） [12] 包含 282 张用于训练的正常图像和 118 张用于测试的异常图像。每张异常图像都包含五种不同病变的掩码，包括玻璃膜疣、渗出物、出血、瘢痕及其他病变。

原始眼底图像尺寸过大，无法用于构建记忆库（IDRiD 为 4288×2848，ADAM 为 2124×2056），因此我们将每张图像缩放至 784×784，以对眼底图像进行异常检测。

4.2. 评估指标

按照以往的研究 [3, 7],，采用曲线下面积（AUC）评估性能。此外，我们通过平衡准确率（准确率）[13], 评估该方法的性能，该指标对真正例和真负例预测进行归一化。

4.3. 实现细节

在我们的实现中，我们采用 WideResnet-50 [14] 作为预训练的 ImageNet模型。训练样本数量的缩减比例为 0.1。稀疏随机投影 ϵ 的权重为 0.90。距离感知滤波器 ψ_r 的半径 r 为 12。所有代码均使用 PyTorch 在配备 24GB 内存的单块 NVIDIA RTX 3090 GPU 上实现。

4.4. 实验结果

我们将所提方法与多种最先进的方法进行比较：fAnoGAN [6], AnoGAN [15], MemAE [16], WDMT- Net [7], 自编码器 [17]。此外，我们还将

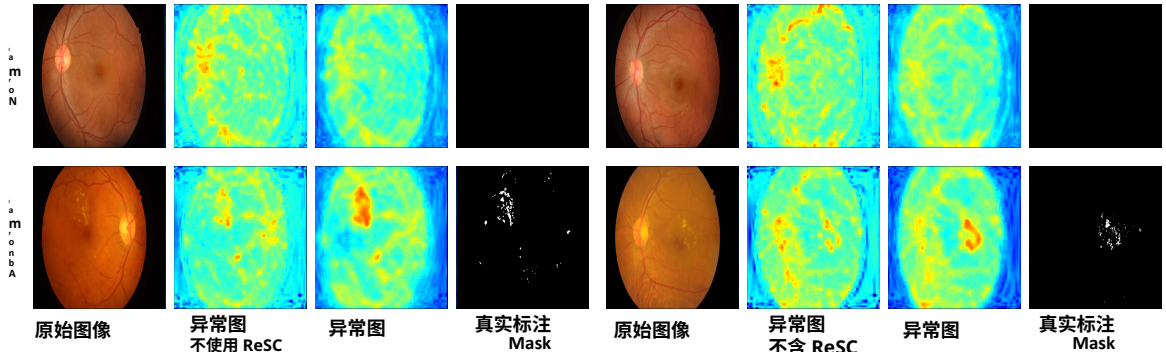

图 2。ReSAD 在 IDRiD数据集上的定性结果。

方法	IDRiD [11]		ADAM [12]	
	AUC	ACC	AUC	ACC
自编码器 [17]	0.687	0.638	0.662	0.633
fAnoGAN [6]	0.729	0.715	0.716	0.693
AnoGAN [15]	0.635	0.659	0.677	0.661
MemAE [16]	0.677	0.630	0.670	0.668
WDMT-Net [7]	0.687	0.638	0.668	0.653
PatchCore [5]	0.845	0.758	0.733	0.678
Ours	**0.907**	**0.819**	**0.820**	**0.732**

表1。IDRiD和ADAM的像素级结果。

方法	AUC	ACC
自编码器 [17]	0.818	0.782
WDMT-Net [7]	0.680	0.653
MemAE [16]	0.722	0.689
PatchCore [5]	0.777	0.738
fAnoGAN [6]	0.780	0.790
Ours	**0.941**	**0.902**

表2。IDRiD 的图像级结果。

自然图像中的基于表征的最先进方法 PatchCore [5] 用作实验基线。如表1和表2所示，在 IDRiD 和 ADAM 数据集上，ReSAD 在像素级异常检测方面均优于所有基线，并且在 IDRiD 数据集的图像级结果上也达到了最先进的性能。

图2展示了异常分数图的可视化结果。可以看出，网络对异常区域有较强的激活。此外，从图中还可以看出，与未使用ReSC模块相比，ReSC能够减少正常视网膜结构区域中的假阳性，并在眼底图像异常检测方面呈现出更好的可视化效果。

4.5. 消融

在消融实验中，我们通过评估在IDRiD数据集上不使用ReSC模块（§3.1）、不使用区域感知特征模块、不使用空间感知特征模块以及使用ReSC模块时的性能，研究ReSC模块的重要性。如表3所示，与未使用ReSC模块相比，使用该模块后图像级AUC提高了17.2%，准确率提高了14.6%，像素级AUC提高了4.3%，准确率提高了4.2%，这证明ReSC模块显著提升了眼底图像异常检测的性能。

此外，为了证明 ReSAD 的数据效率，我们减少了 IDRiD训练集中的正常图像。如表3所示，减少正常图像对性能的影响很小。即使仅使用 10% 的训练数据，ReSAD 也能达到

消融	IDRiD [11]			
	图像级AUC	图像级准确率	像素级AUC	像素级准确率
10%	0.913	0.855	0.898	0.808
40%	0.916	0.840	0.899	0.809
70%	0.925	0.870	0.901	0.812
不使用 ReSC	0.769	0.756	0.864	0.777
不使用区域模块	0.812	0.785	0.878	0.792
不使用空间模块	0.915	0.848	0.897	0.807
使用 ReSC	0.941	0.902	0.907	0.819

表 3. 训练数据数量与 ReSC 的消融研究。I-AUC、I-ACC 分别表示图像级AUC和图像级准确率，P-AUC、P-ACC 分别表示像素级AUC和像素级准确率。

图像级异常检测的 AUC 为 91.3%、准确率为 85.5%，像素级异常检测的 AUC 为 89.8%、准确率为 80.8%，仍优于大多数基线结果。

5. 结论

本文提出了一种用于眼底图像的新型异常检测方法 ReSAD。ReSAD采用预训练的ImageNet模型提取特征，并设计了区域与空间感知特征组合模块（ReSC）来构建记忆库。最后，ReSAD利用记忆库生成图像级和像素级异常分数。通过在ADAM和IDRiD上对我们的方法进行评估，我们的方法显著优于现有最先进的异常检测方法，证明了ReSAD在眼底异常检测中的有效性。

6. 符合伦理标准声明

本研究采用开放获取的**IDRiD** 和**ADAM** 数据集中的人类受试者数据进行回顾性研究。由于开放获取数据所附许可证已确认无需伦理审批，因此未要求获得伦理批准。

7. 参考文献

[1] Dinggang Shen、GuorongWu 和 Heung-Il Suk，《医学图像分析中的深度学习》，《生物医学工程年度评论》，第 19 卷，第 221 页，2017 年。

[2] Yu Cai、Hao Chen、Xin Yang、Yu Zhou 和 Kwang- Ting Cheng，“胸部X光片异常检测的双分布差异”，*arXiv*预印本 *arXiv:2206.03935*，2022 年。

[3] Kang Zhou、Jing Li、Weixin Luo、Zhengxin Li、Jianlong Yang、Huazhu Fu、Jun Cheng、Jiang Liu 和 Shenghua Gao，“用于医学图像异常检测的代理桥接图像重建网络”，*IEEE*医学影像汇刊，第 41 卷，第 3 期，第 582–594 页，2021 年。[4] Mohammadreza Salehi、Niousha Sadjadi、Soroosh Baselizadeh、Mohammad H Rohban 和 Hamid R Rabiee，“用于异常检测的多分辨率知识蒸馏”，载于*IEEE/CVF*计算机视觉与模式识别会议论文集，2021 年，第 14902–14912 页。[5] Karsten Roth、Latha Pemula、Joaquin Zepeda、Bernhard Sch¨olkopf、Thomas Brox 和 Peter Gehler，“迈向工业异常检测中的全面召回”，载于*IEEE/CVF*计算机视觉与模式识别会议论文集，2022 年，第 14318–14328 页。[6] Thomas Schlegl、Philipp Seeb¨ock、Sebastian M Wald- stein、Georg Langs 和 Ursula Schmidt-Erfurth，“f-anogan：基于生成对抗网络的快速无监督异常检测”，医学图像分析，第 54 卷，第 30–44 页，2019 年。[7] Wentian Zhang、Xu Sun、Yuexiang Li、Haozhe Liu、Nanjun He、Feng Liu 和 Yefeng Zheng，“一种采用权重衰减跳跃连接训练的多任务网络，用于视网膜眼底图像异常检测”，载于医学图像计算与计算机辅助干预国际会议。施普林格，2022 年，第 656–666 页。[8]Khalil Ouardini、Huijuan Yang、Balagopal Unnikr- ishnan、Manon Romain、Camille Garcin、Houssam Zenati、J Peter Campbell、Michael F Chiang、Jayashree Kalpathy-Cramer、Vijay Chandrasekhar 等，“迈向视网膜图像实用型无监督异常检测”，载于少标签和不完美数据条件下的领域自适应、表征迁移与医学图像学习，第 225–234 页。施普林格，2019 年。[9] Sourya Dipta Das、Saikat Dutta、Nisarg A Shah、Dwarikanath Mahapatra 和 Zongyuan Ge，“利用多尺度深度特征稀疏编码检测视网膜图像中的异常”，载于*2022 IEEE* 第 *19* 届生物医学成像国际研讨会（*ISBI*）。IEEE，2022 年，第 1–5 页。[10] Ozan Sener 和 Silvio Savarese，“卷积神经网络的主动学习：一种核心集方法”，*arXiv*预印本 *arXiv:1708.00489*，2017 年。[11] Prasanna Porwal、Samiksha Pachade、Manesh Kokare、Girish Deshmukh、Jaemin Son、Woong Bae、Lihong Liu、Jianzong Wang、Xinhui Liu、Liangxin Gao 等，“IDRiD：糖尿病视网膜病变分割与分级挑战赛”，医学图像分析，第 59 卷，第 101561 页，2020 年。[12] Huihui Fang、Fei Li、Huazhu Fu、Xu Sun、Xingxing Cao、Fengbin Lin、Jaemin Son、Sunho Kim、Gwenole Quellec、Sarah Matta 等，“ADAM挑战赛：从眼底图像中检测年龄相关性黄斑变性”，*IEEE*医学影像汇刊，2022 年。[13] Jeffrey P Mower，“PREP套件：用于植物线粒体基因、叶绿体基因和用户定义比对的预测性RNA编辑器”，《核酸研究》，第 37 卷，增刊 2，第 W253–W259 页，2009 年。[14] Sergey Zagoruyko 和 Nikos Komodakis，“宽残差网络”，*arXiv*预印本 *arXiv:1605.07146*，2016 年。[15] Thomas Schlegl、Philipp Seeb¨ock、Sebastian M Wald- stein、Ursula Schmidt-Erfurth 和 Georg Langs，“利用生成对抗网络引导标志物发现的无监督异常检测”，载于医学图像处理国际会议。施普林格，2017 年，第 146–157 页。[16] Dong Gong、Lingqiao Liu、Vuong Le、Budhaditya Saha、Moussa Reda Mansour、Svetha Venkatesh 和 Anton van den Hengel，“记忆正常模式以检测异常：用于无监督异常检测的记忆增强深度自编码器”，载于*IEEE/CVF*计算机视觉国际会议论文集，2019 年，第 1705–1714 页。[17] Christoph Baur、Benedikt Wiestler、Shadi Albarqouni 和 Nassir Navab，“用于脑部MR图像无监督异常分割的深度自编码模型”，载于*MICCAI*脑病变研讨会。施普林格，2018 年，第 161–169 页。