杜 涛, 闫建红. DI-MobileNet:基于轻量化网络的骶髂关节炎识别方法[J]. 智能计算机与应用,2025,15(2):138-143.
DOI:10.20169/j.issn.2095-2163.24102205

DI-MobileNet:基于轻量化网络的骶髂关节炎识别方法

杜 涛, 闫建红
(太原师范学院 计算机科学与技术学院, 山西 晋中 030619)

摘 要: 针对传统疾病识别网络参数量大,部署效率低的特点,本文提出一种基于改进 MobileNetV2 模型的分类模型 DI-MobileNet,该模型以轻量化网络 MobileNetV2 为基模型,基于 Inception 提出多尺度卷积结构;将空洞卷积嵌入模型中,以更高效的方式提取骶髂关节图像不同尺度特征。采用公共数据集 Digital Knee X-ray 与私有骶髂关节数据集进行实验验证,实验结果表明 DI-MobileNet 模型参数量为 2.23 M,远低于常规卷积神经网络,在 Digital Knee X-ray 数据集上准确率达到 93.05%,在骶髂关节数据集上准确率达到了 97.33%,均高于其他模型,较原模型分别提高了 1.4 和 6.31 个百分点。

关键词: 骶髂关节; Inception; 空洞卷积; MobileNetV2
中图分类号: TP391.7 **文献标志码**: A **文章编号**: 2095-2163(2025)02-0138-06

DI-MobileNet: A lightweight network-based method to identify sacroiliac arthritis

DU Tao, YAN Jianhong
(School of Computer Science and Technology, Taiyuan Normal University, Jinzhong 030619, Shanxi, China)

Abstract: Aiming at the characteristics of sacroiliac joint diseases with large intraclass variation and small interclass variation, a classification model DI-MobileNet is proposed to improve the MobileNetV2 model, which takes the lightweight network MobileNetV2 as the base model, proposes a multi-scale convolution structure based on Inception, and subsequently embeds the null convolution into the model to extract sacroiliac iliac joint images with different scale features. The public dataset Digital Knee X-ray with sacroiliac joint dataset is used for validation. Experiments show that the number of parameters of DI-MobileNet model is 2.23 M, which is much lower than that of conventional convolutional neural network. The accuracy reached 93.05% on the Digital Knee X-ray dataset and 97.33% on the sacroiliac joint dataset, which were both higher than the other models, and improved 1.4 and 6.31 percentage points compared with the original model, respectively.

Key words: Sacroiliac joint; Inception; Dilated convolution; MobileNetV2

0 引 言

骶髂关节是由骶骨和髂骨上的凹凸不平的耳状关节面组成,两者密切结合,并被韧带包围,以约束和稳定关节的运动[1]。由于骶髂关节相关疾病在不同类别之间差异不大,但在个体之间存在显著差异,且由于关节位置较深,呈窄条状,韧带和骨质覆盖了大部分关节的后部,只有极小部分可见,因此,目前临床主要依赖影像学来检测骶髂关节疾病。

医学影像通常涵盖 X 射线(X-ray),计算机断层扫描(Computed Tomography, CT)与核磁共振成像(Nuclear Magnetic Resonance Imaging, MRI)。膝关节 X 射线和骶髂关节 CT 数据集涉及不同的成像方式,有助于验证模型在多模态数据上的鲁棒性和准确性,同时可以检验模型对于共同疾病特征的捕捉能力。虽然膝关节与骶髂关节位置不同,但两者在解剖结构上存在相似性[2]。这种相似性表现在骨骼结构、关节轮廓、软组织位置等方面,同时有助于验证模型在不同关节疾病识别上的泛化能力。目前,骶髂炎的 CT 诊断依赖于观看者,放射科医生和医疗机构的结果可能不一致,许多体征并不是骶髂炎所特有的,也可能出现在退行性疾病和其他非骶髂关节炎相关疾病中[3]。因此,在疑似骶髂炎患者的 CT 图像上,需要辅助工具对骶髂疾病类型进行识别。

基金项目: 太原师范学院校级研究生教育创新项目(SYYJSYC-2470); 山西省研究生精品教学案例项目(2024AL27)。
作者简介: 杜 涛(1999—),男,硕士研究生,主要研究方向:深度学习,人工智能。
通信作者: 闫建红(1972—),女,博士,教授,主要研究方向:机器学习,人工智能,计算机视觉。Email:yan_jian_hong@163.com。
收稿日期: 2024-10-22

随着深度学习的发展，越来越多的神经网络被应用于医学图像诊断领域。2020 年，Castro-Zunti R[4] 使用骶髂关节 CT 图像进行二分类的研究，借助了迁移学习方法进行网络训练，获得了 99.0% 的准确率。然而，迁移学习使用的 ImageNet 预训练网络权重和骶髂关节图像之间存在较大差异，且迁移网络权重参数较大，实际应用中难以部署。Proft 等[5] 提出使用一工神经网络，能够在 X 射线片上精确检测出与轴性脊柱炎（axSpA）诊断相关的明确的骶髂炎，接近专家的表现。但 X 线不能显示活动性病变，对 axSpA 的早期诊断价值有限，并且在轻微糜烂时效果较差。2021 年，Bressem K K 等[6] 使用深度学习方法检测 X 平片上是否能准确检测出明确的放射学骶髂关节炎；2023 年，李府天[7] 使用先分割后分类的方法，对骶髂关节 CT 图像进行二、三、五分类，分别获得了 0.95、0.94 和 0.86 的分类准确率。上述两种特征提取方法都侧重于对局部特征的提取，导致网络缺失了对全局特征的关注。同年，Zhang K 等[8] 根据修订的纽约标准将骶髂关节炎分为了 3 级，并使用 3D 卷积神经网络对骶髂关节进行分级诊断研究，3D 卷积虽然可以考虑到 CT 图像的切片间距，但计算量较大，导致网络训练速度较慢。

基于深度学习方法的骶髂关节疾病识别解决了人工检测和传统影像学方法的不足，能达到较高的识别准确率和识别速度，但现有的针对骶髂关节识别的深度学习方法也存在一些不足：为了追求小的计算量，模型通常采用小卷积核提取特征，可能忽略对于全局特征的关注；直接引入全局特征后模型参数量变大，不利于模型的部署应用，实用性差。基于此，本文的主要创新点总结如下：

（1）采用轻量化模型 MobileNetV2 模型，为了弥补参数量下降带来的准确率的下降，引入了 Inception 模块，让网络自适应地选择不同尺寸的卷积核，能在同一卷积层内捕获不同尺寸信息，提升了提取的特征质量；

（2）引入空洞卷积到 MobileNetV2 模型中，弥补卷积神经网络本身只能提取局部特征的缺点，增加卷积核的感受野，同时又不增加计算量，从而提升模型的预测效率。

1　传统 MobileNetV2 模型简介

本文选用 MobileNetV2 作为基模型，其由 Google 团队于 2018 年提出，是一种边缘计算模型，延续了 MobileNetV1 的深度可分离卷积思想，并引入了倒残差结构和线性瓶颈结构。相较传统卷积模型，如 Resnet 和卷积神经网络，MobileNetV2 具有更高的准确率和更低的计算成本，更适合轻量级疾病诊断需求[9]。

如图 1 所示，倒残差结构是一种高效的卷积神经网络模块，广泛应用于轻量化网络设计。其先通过扩展层增加通道数，引入更多特征表达能力；接着使用深度可分离卷积提取特征，减少计算复杂度；最后通过压缩层降低通道数，恢复到输入通道维度。此结构在减少计算量和参数的同时，保持了模型性能。

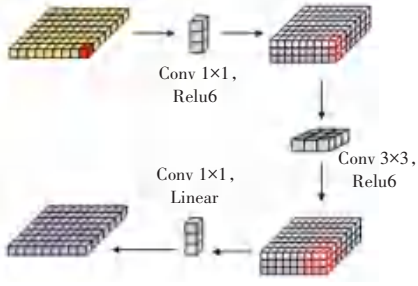

图 1　深度可分离卷积

Fig. 1　Deeply separable convolution

2　DI-MobileNet 模型架构

DI-MobileNet 模型采用多尺度特征融合方法，融入空洞卷积，在保证低参数量的同时拥有较高的准确性，模型整体结构如图 2 所示，由空洞卷积、倒残差结构以及 Inception-x 模块构成，包含特征提取阶段和特征拼接阶段。特征提取阶段，引入空洞卷积来增加卷积核的感受野，提取出包含更多信息的特征；在特征拼接阶段，引入 Inception 结构对前期提取出的特征进行多尺度拼接融合，最后输出模型预测结果。

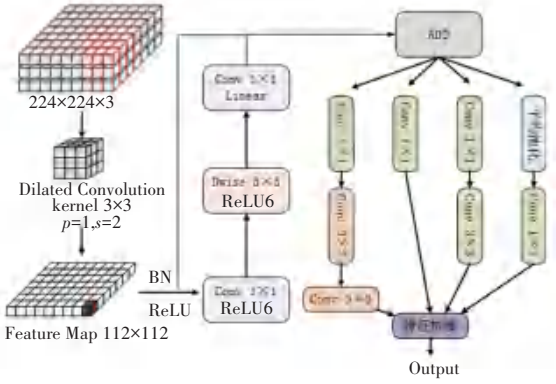

图 2　DI-MobileNet 整体架构

Fig. 2　DI-MobileNet Overall Architecture

2.1 空洞卷积核

在标准的轻量级 MobileNetV2 模型中,卷积核主要包括普通卷积核和深度可分离卷积核,通常采用较小的 3×3 卷积核,以减少模型参数量和计算负担。在靠近输入的初始层,特征图的分辨率较高,而小尺寸的卷积核具有较小的局部感受野,限制了对特征的捕捉。尽管这种卷积核有助于提取细微特征,但由于受到分辨率的限制,难以捕捉到更全面的特征,因此引入了空洞卷积核的概念。

空洞卷积核在卷积核内部引入固定间隔的零元素来扩大感受野,从而增强卷积核的感知范围,而不增加参数数量[10]。这种方式有效地扩展了每层卷积神经网络的感受野,使得网络能够捕获更广泛的上下文信息,从而提升特征提取的能力,如图 3 所示。

图 3 空洞卷积示意图

Fig. 3 Schematic diagram of a hollow convolution

在图 3 中,图 3(a)展示了常规 3×3 卷积核的感受野,而图 3(b)和图 3(c)则显示了分别采用扩张率为 2 和 3 时的卷积核感受野变化情况,分别为 5×5 和 7×7,且均未使用填充。通过调整扩张率来扩大卷积核的感受野,能够在不增加参数数量的情况下提升网络的感知范围。

在 MobileNetV2 模型中,保持每个网络层的步长设置不变,仅在第二层的深度可分离卷积层中将原本使用步长为 1 的深度卷积层中的普通卷积核替换为空洞卷积核,以保持输出特征图的大小不变。这一调整不会对后续网络结构产生影响,具体网络结构如图 4 所示。

设输入特征图为 $I=R^{h\times w\times m}$,其中 h,w,m 代表特征图的高、宽和输入特征图的通道数。用扩张率为 r 的同尺寸大小的卷积核 K 进行无填充的空洞卷积操作,可得到 $R^{(h-s-(s-1)(r-1)+1)\times(w-s-(s-1)(r-1)+1)\times n}$ 的输出特征图 O_d,s 代表卷积核的尺寸,n 代表输出特征图通道数,公式如下:

$$O_d(y,x,j)=\sum_{i=1}^{m}\sum_{u,v=1}^{s}K(u,v,i,j)I(y+u+(u-1)(r-1)-1,x+v+(v-1)(r-1)-1,i) \tag{1}$$

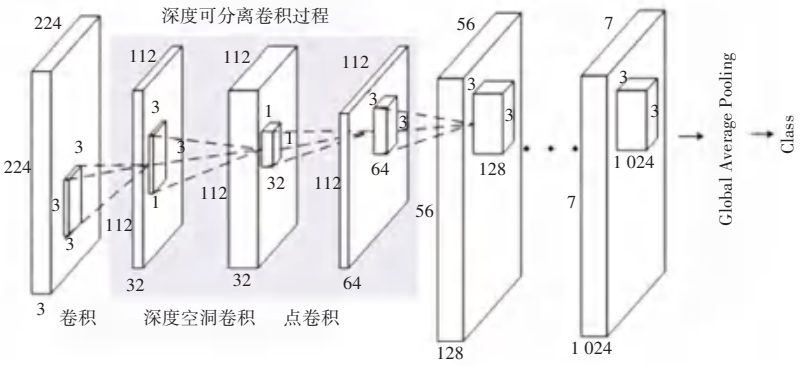

图 4 深度空洞卷积

Fig. 4 Deep void convolution

由公式(1)可知,空洞卷积层的总计算量为 $s\times s\times m\times(h-s-(s-1)(r-1)+1)\times(w-s-(s-1)(r-1)+1)\times n$,参数量为 $s\times s\times m\times n$,与标准卷积的参数量一样,但感受野扩大了,证实空洞卷积核扩张局部感受野可以有效提高分类精度。

2.2 多尺度特征提取模块 Inception-x

Inception 是 GoogleNet 中的核心模块,旨在提升神经网络对多尺度特征的提取能力,其关键理念在于并行应用不同尺寸的卷积核和池化操作,以增加网络的深度和宽度[11]。这种策略使得网络可以自适应地选择最适合的卷积核尺寸,有助于捕捉各种尺度的信息。

Inception-x 架构如图 5 所示。Inception-x 模块包含 4 个并行组件:3×3 卷积用于捕获中等尺度的空间相关性,拆分原始 5×5 卷积为两个单独的 3×3 卷积,用于捕捉更广泛的空间相关性;3×3 最大池化提供了另一种形式的空间抽象;最后,通过拼接操作将各个尺寸卷积核和池化操作的特征图连接在一起,有效地融合了不同尺度的特征信息,但模型在处理多尺度信息时增加了网络的复杂性。Inception-x 模块借鉴了残差网络结构的设计,通过添加残差连接来促进更深的网络训练,有助于解决深层网络中梯度消失和

爆炸问题。将 Inception 模块中的 5×5 卷积分解为 2 个连续的 3×3 卷积，在卷积之前使用 1×1 卷积降维的同时也能捕捉特征图通道之间的局部相关性，在减少参数量和计算量的同时保持了感受野一致[12]，能够有效处理骶髂关节图像的多尺度问题。

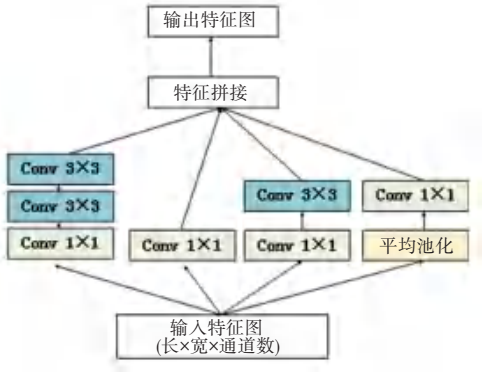

图 5　Inception-x 架构

Fig. 5　Inception-x architecture

3　实验与结果分析

在公共数据集及私有数据集上进行对比实验，测试模型效果及参数量，通过消融实验验证所提模型的有效性。

3.1　实验设置

实验基于 Windows10 操作系统，GPU 型号为 Nvidia GeForce RTX3090，深度学习框架为 Pytorch。采用了能够自适应调整学习率的 Adam W 优化器。Adam W 优化器改进了 Adam 优化器中权重衰减的不足，减少了调参的复杂性。损失函数选用了交叉熵损失函数，并采用等间隔调整学习率的策略，初始学习率为 0.000 1，每经过 10 轮训练学习率就会下降 0.000 01。每个训练批次包含 32 张图片，总训练迭代次数设定为 100 轮。为防止模型过拟合设置了提前停止，当模型连续 10 轮损失函数未下降便停止训练，并保留之前表现最好的模型参数。

3.2　评价指标

使用准确率、灵敏度和特异性 3 种指标评估模型性能。准确率衡量模型正确识别骶髂关节疾病类型样本数量占总样本数的比例，公式如下：

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

灵敏度指标用于衡量被正确辨别的正例个数占所有正类的比例，公式如下：

$$SEN = \frac{TP}{TP + FN} \tag{3}$$

特异性指标表示被正确辨别的负类个数占所有负类的比例，公式如下：

$$SPE = \frac{TN}{TN + FP} \tag{4}$$

其中，TP 表示真阳性，表示被正确划分为正例的数量；TN 表示真阴性，表示被正确划分为负例的数量；FP 表示假阳性，表示被错误划分为正例的数量；FN 表示假阴性，表示被错误划分为负例的数量。

在多分类实验中，计算每个类别的宏平均值作为该类的最终结果。宏平均首先计算每个类别的评价指标，包括准确率、灵敏度和特异性，然后将各类别的结果求和并计算平均值。这种方法能够有效处理多类别情况，确保每个类别对评价指标的贡献都得到了充分考虑。计算公式如下：

$$\overline{ACC} = \frac{1}{n}\sum_{i=1}^{n} ACC_i \tag{5}$$

$$\overline{SEN} = \frac{1}{n}\sum_{i=1}^{n} SEN_i \tag{6}$$

$$\overline{SPE} = \frac{1}{n}\sum_{i=1}^{n} SPE_i \tag{7}$$

其中，ACC_i、SEN_i、SPE_i 表示第 i 个类别的准确率、灵敏度和特异性，n 为类别总数。

使用评价指标和宏平均方法，能够全面评估模型在多类别分类任务中的性能，并得出综合的评估结果。

3.3　数据集

膝关节 X 射线数据集来自公开数据集库，包括从知名医院和诊断中心收集的 1 650 张膝关节电子 X 射线图像，原始图像是 8 位灰度图像，每个 X 射线膝关节图像均由 Kellgren 和 Lawrence 两位医学专家手动注释标记。

骶髂关节疾病四分类数据集由大型三甲医院影像中心提供。数据源自 2023 年至 2024 年期间采用 CT 机收集的 40 例骶髂关节疾病。实验构建了 4 种等级的骶髂关节疾病类型，分别为骶髂关节炎、退行性改变、未见异常和致密性骨炎，将所有 CT 图像切分为单张切片进行分类，共计 3 728 张切片，所有图像均由两位经验丰富的影像科医师进行筛选和标注。

3.4　对比实验

3.4.1　公共数据集对比实验

为评估模型性能及其对医学图像分类的有效性，选择了一些经典的分类网络模型 Resnet-152、

Densenet-121、Regnetx_002[13]，ViT-B/16 及 ViT-Deit[14]进行对比实验，在膝关节数据集上实验结果见表 1。

表 1 公共数据集模型性能比较

Table 1 Comparison of public dataset

模型	*ACC*/%	*SEN*/%	*SPE*/%	参数量/M
Resnet-152	84.59	92.13	84.88	60.21
Densenet-121	84.13	92.31	87.33	6.90
RegNetx_002	82.64	87.22	83.15	6.20
ViT-B/16	85.24	90.25	87.28	86.01
ViT-Deit	86.56	92.65	89.47	77.90
本文	**93.72**	**92.67**	**98.26**	**2.23**

表 1 可见，在膝关节 X 射线数据集中，本文算法相比常规卷积参数量大幅降低，相较于 ViT 模型参数量更是降低了数十倍；同时在准确率、灵敏度及特异度等指标上分别提高了 7.16%、0.02%，8.49%，表明相较于传统模型，本文方法都可以用较小的参数量得到更好的效果。

3.4.2 私有骶髂关节数据集对比实验

将上述模型与本文模型在骶髂关节数据集上进行对比实验，结果见表 2。

表 2 骶髂关节数据集模型性能比较

Table 2 Comparison of model performance for the sacroiliac joint dataset

模型	*ACC*/%	*SEN*/%	*SPE*/%	参数量/M
Resnet-152	90.52	90.15	96.76	60.20
Densenet-121	91.54	90.35	97.02	6.96
RegNetx_002	92.58	92.17	97.44	6.20
ViT-B/16	84.95	84.49	94.65	86.00
ViT-Deit	93.76	93.32	97.51	77.90
本文	**97.33**	**97.23**	**99.14**	**2.23**

在骶髂关节数据集上，本文模型同样在保持较小参数量的前提下，在准确率，敏感度及特异度 3 个评价指标比已知最好的模型 ViT-Deit 分类准确率分别提升了 3.47%、3.91%和 1.63%，相较于其他自注意模型，可以更好地捕捉骶髂关节的细节局部特征，提升了对于骶髂关节疾病的分类效果。

3.5 消融实验

为了验证网络各模块的有效性，在本文的骶髂关节疾病私有图像数据集上进行消融实验。实验以 MobileNetV2 网络为基准方法，分别在此网络上增加 Inception-x、空洞卷积核（Dilated Conv），实验结果见表 3。

表 3 不同模块对模型性能的影响

Table 3 Effect of different modules on the performance of the model

模块	参数量/M	膝关节 *X* 射线数据集			骶髂关节数据集		
		ACC/%	*SEN*/%	*SPE*/%	*ACC*/%	*SEN*/%	*SPE*/%
MobileNets	1.66	92.32	89.91	97.85	91.02	90.91	96.92
+Inception-x	2.23	93.14	91.93	98.01	95.05	94.92	98.36
+Dilated Conv	2.23	**93.72**	**92.67**	**98.26**	**97.33**	**97.23**	**99.14**

由表 3 可见，添加 Inception-x 模块后，参数量有小幅度上升，但同时在膝关节 X 射线数据集上准确率、敏感度和特异度也都有所上升，在骶髂关节上准确率、敏感度和特异度也分别增加了 4.03%，4.01%，1.44%，表明多尺度的特征提取有利于提取到更好的特征，但会带来额外参数量；添加空洞卷积核后，3 个指标在骶髂关节数据集上分别增加了 2.28%、2.31%和 0.78%，参数量并没有明显变化，表明扩大后的感受野可以有效地提取图像中的分类特征，不会带来附加参数。实验证明，加入了两种模块后，模型收敛速度更快且准确率更高。

4 结束语

疾病辅助诊断模型通常参数量较大，实际生活中难以部署。对此，本文提出 DI-MobileNet 模型，将空洞卷积核及 Inception-x 引入 MobileNetV2 中，在骶髂关节 CT 图像分类任务中应用，能够全面挖掘病变区域的深浅特征。实验结果表明，本文模型能够充分挖掘骶髂关节图像的局部和全局特征，参数量较少，便于部署在实际应用中，能够为医生的诊断治疗提供参考。在未来的工作中，会进一步提高模型对于骶髂关节 CT 图像中的细节特征的提取能力，同时扩大数据集规模，以进一步提高模型的性能表现。

参考文献

[1] KIM K, LEE C, KIM K. Sacroiliitis in inflammatory bowel disease on abdominal computed tomography: prevalence, misses,

and associated factors [J]. Scandinavian Journal of Rheumatology，2024，53(4)：1-7.

[2] LAMBERT R，HERMANN K，DIEKHOFF T. Low-Dose computed tomography for axialspondyxo arthritis：Update on use and limitations [J]. Current Opinion in Rheumatology，2021(33)：326-332.

[3] DENIS P，HENNING W. Clinical and imaging characteristics of osteitis condensans ilii as compared with axial spondyloarthritis [J]. Rheumatology，2020，59(12)：3798-3806.

[4] CASTRO-ZUNTI R，PARK E，CHOI Y，et al. Early detection of ankylosing spondylitis using texture features and statistical machine learning，and deep learning，with some patient age analysis[J]. Computerized Medical Imaging and Graphics，2020(82)：101718.

[5] PROFT F，VAHLDIEK J，NICOLAES J，et al. Analysis of the performance of an artificial intelligence algorithm for the detection of radiographic sacroiliitis in an independent cohort of axSpA patients including both Nr-axSpA and r-axSpA [J]. Arthritis Rheumatol，2022，74.

[6] BRESSEM K，VAHLDIEK J，ADAMS L，et al. Deep learning for detection of radiographicsacroiliitis：achieving expert-level performance[J]. Arthritis Research & Therapy，2021，23(1)：1-10.

[7] 李府天. 基于骶髂关节 CT 影像的强直性脊柱炎自动分级研究[D]. 汕头：汕头大学，2023.

[8] ZHANG Ke，LUO Guibo，LI Wenjuan，et al. Automatic image segmentation and grading diagnosis of sacroiliitis associated with AS using a deep convolutional neural network on CT images[J]. Journal of Digital Imaging，2023，36(5)：2025-2034.

[9] MA Yuqian，FANG Zitong，WEN Jiang，et al. Gesture recognition based on time-of-flight sensor and residual neural network[J]. Journal of Computer and Communications，2024，12(6)：103-114.

[10] MONTALBO F J P. S3AR U-Net：A separable squeezed similarity attention-gated residual U-Net for glottis segmentation [J]. Biomedical Signal Processing and Control，2024(92)：106047.

[11] YANG Le，YU Xiaoyun，ZHANG Shaoping，et al. GoogLeNet based on residual network and attention mechanism identification of rice leaf diseases [J]. Computers and Electronics in Agriculture，2023(204)：107543.

[12] ELUL D，LEVIN N. The role of population receptive field sizes in higher-order visual dysfunction [R]. Current Neurology and Neuroscience Reports. 2024：1-10.

[13] KARTHICK S. Deep RegNet-150 architecture for single image super resolution of real-time unpaired image data[J]. Applied Soft Computing，2024，162：111837.

[14] MARINA B，ASKHAT S，SERGEY O，et al. Deep learning approaches to automatic chronic venous disease classification[J]. Mathematics，2022，10(19)：3571.