

融合双重注意力与加权动态卷积的 车辆损伤分类模型

翟永杰, 王紫萱, 张祯琪, 周迅琪, 王乾铭

(华北电力大学自动化系, 河北 保定 071003)

摘要: 针对车险理赔客户上传的车辆损伤图像中存在损伤类型形态相似、分类困难的问题, 提出了一种适用于车辆损伤分类的模型 ResAWDNet。首先, 为有效增强模型对损伤特征的提取能力, 使用加权动态卷积代替原有的下采样操作, 依据输入特征动态调整卷积核权重, 提高模型对不同尺度和方向特征的适应性, 从而更准确地捕捉损伤的细微差异。其次, 为了使模型关注图像中的显著性判别区域和特征通道, 在主干网络的卷积层后嵌入了双重注意力机制, 同时学习空间和通道维度上的重要权重, 提升模型对关键信息的捕捉能力, 进一步提升模型在损伤分类任务中的决策准确性。最后, 基于真实事故案例的车辆损伤图片数据集进行实验验证。实验结果表明, ResAWDNet 模型在车辆损伤分类任务中切实可行且优势显著, 整体分类准确率达到 73.79%。与基线模型相比, ResAWDNet 在多数损伤类型的分类上均展现出更高的准确率, 有力地证明了该模型的有效性。

关键词: 智能定损; 图像分类; 深度学习; 注意力机制; 动态卷积

中图分类号: TP 391.41

DOI: 10.11996/JGj.2095-302X.2026010017

文献标识码: A

文章编号: 2095-302X(2026)01-0017-12

A vehicle damage classification model incorporating dual attention and weighted dynamic convolution

ZHAI Yongjie, WANG Zixuan, ZHANG Zhenqi, ZHOU Xunqi, WANG Qianming

(Department of Automation, North China Electric Power University, Baoding Hebei 071003, China)

Abstract: To address the challenges of morphological similarity and the resulting difficulty in classifying vehicle damage images uploaded by clients for auto insurance claims, a model named ResAWDNet was proposed for vehicle damage classification. Firstly, to effectively augment the model's capacity for extracting damage features, the traditional down sampling operation was replaced with weighted dynamic convolution. This approach dynamically adjusted the weights of convolutional kernels based on the input features, thereby enhancing the model's adaptability to features of varying scales and orientations. As a result, it enabled more precise capture of the subtle differences in vehicle damage. Secondly, to ensure that the model could concentrate on the salient discriminative regions and feature channels within the images, a dual attention mechanism was embedded after the convolutional layers of the backbone network. This mechanism concurrently learned the important weights in both spatial and channel dimensions, significantly enhancing the model's ability to capture crucial information. Consequently, it further enhanced the decision-making accuracy of the model in the task of vehicle damage classification. Finally, experimental validation was conducted based on a dataset of vehicle damage images sourced from real accident cases. The experimental

收稿日期: 2025-02-28; 定稿日期: 2025-06-23; 通信作者: 王乾铭, E-mail: qianmingwang@ncepu.edu.cn

Received: 28 February, 2025; Finalized: 23 June, 2025; Corresponding author: WANG Qianming, E-mail: qianmingwang@ncepu.edu.cn

基金项目: 国家自然科学基金(62373151); 河北省自然科学基金(F2023502010); 国家自然科学基金联合基金项目重点支持项目(U21A20486); 河北省在读研究生创新能力培养资助项目(CXZZSS2025152)

Foundation items: National Natural Science Foundation of China (62373151); Natural Science Foundation of Hebei Province (F2023502010); Joint Funds of the National Natural Science Foundation of China (U21A20486); Project for Cultivating the Innovative Ability of Full-time Postgraduate Students Studying in Hebei Province (CXZZSS2025152)

results demonstrated that the ResAWDNet model was feasible and offered significant advantages for vehicle damage classification tasks, achieving an accuracy rate of 73.79%. Compared with baseline models, ResAWDNet achieved higher accuracy in classifying multiple types of damages, robustly validating the effectiveness of the proposed model.

Keywords: intelligent damage assessment; image classification; deep learning; attention mechanism; dynamic convolution

中华人民共和国国家发展和改革委员会于2020年发布“数字化转型伙伴行动”倡议^[1],旨在推动各行业的数字化转型,提升企业的智能化水平。对于保险公司而言,结合现代智能技术实现车辆智能定损以进行车险智慧化转型是实现业务提升的重要途径,也是提升保险公司运营效率、客户满意度的有效手段。车辆智能定损技术^[2]利用深度学习算法,采用人工拍摄方式获取车辆损伤外观图像,通过分析大量汽车相关数据和损伤图像判定汽车的损坏情况和损伤类型,进而确定受损类型以及评估部件受损程度,给出赔付方案。因此,

基于计算机视觉的车辆智能定损技术能够有效降低保险定损时投入的人力和物力成本,提高车辆定损效率。

智能定损中常见的损伤可分为车身刮蹭、车身划痕、车身变形、玻璃破损、玻璃划痕、错位等多种类别,如图1所示。由于车辆的外观部件形态多样,且损伤发生的情况、角度和场景均不同,导致同一类别的损伤在视觉上表现差异较大,难以分辨。同时损伤图片存在模糊、质量不佳的问题,为损伤辨别增加了难度。因此需要提出针对性的模型能够迅速且精准地对车辆损伤图像进行分类。

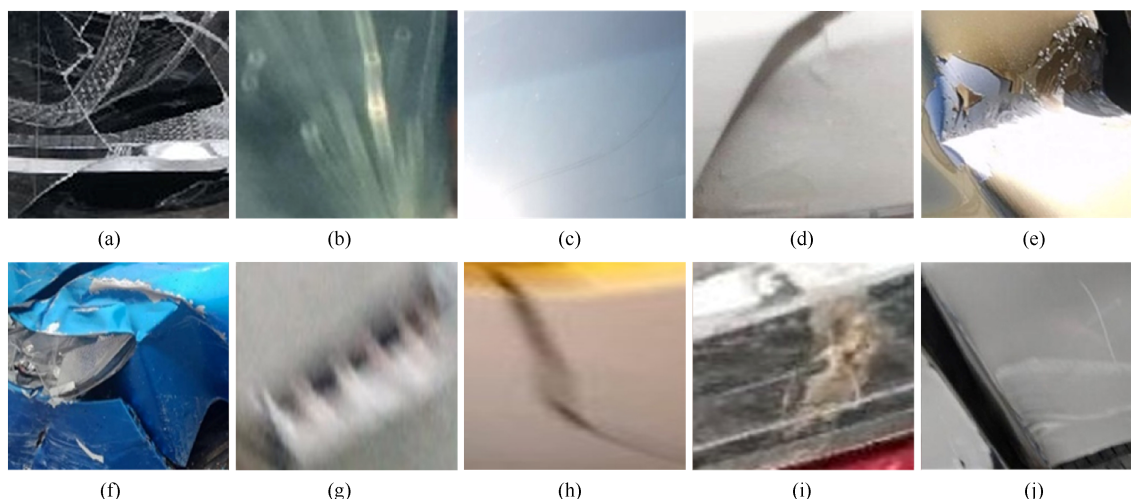


图1 10类损伤类型展示((a) 丢失; (b) 玻璃破损; (c) 玻璃划痕; (d) 轻度变形; (e) 中度变形; (f) 重度变形; (g) 错位; (h) 撕裂; (i) 车身刮擦; (j) 车身划痕)

Fig. 1 Display of the ten injury types ((a) Loss; (b) Glass breakage; (c) Glass scratches; (d) Mild deformation; (e) Moderate deformation; (f) Severe deformation; (g) Misalignment; (h) Tearing; (i) Body scuffing; (j) Body scratches)

近年来,深度学习技术在损伤分类领域已经取得了诸多进展和成就。如武兵和田莹^[3]提出的YOLOv8-RDD模型应用于道路损伤检测;LIU等^[4]开发了基于Transformer结合分块与掩码机制的模型用于工业气缸套缺陷检测;PARK等^[5]提出了一种卷积神经网络成像分析,识别检测零件表面缺陷;WU等^[6]针对小样本下的钢表面缺陷分类问题,优化了特征提取网络,为实际工业生产中的小样本损伤分类提供了新思路;LIU等^[7]提出了一种基于多尺度特征融合的深度学习方法,用于复杂工业零部件的表面损伤分类。然而,上述工作大多针对金

属缺陷、表面划痕等单一损伤类型,车辆损伤类型多样,形态多变,已有的算法在车辆损伤分类任务中难以取得理想表现,针对车辆损伤分类的具体研究仍处于相对初级的探索阶段。

本文结合保险公司实际车险案例图片构建车辆损伤数据集,基于ResNet50模型,针对车辆损伤分类任务进行了深入研究和改进。通过嵌入双重注意力、使用加权动态卷积代替原有的下采样,提出一种适用于车辆损伤分类的模型。本文主要贡献如下:

1) 以ResNet50模型作为基础框架,提出了适

用于车辆损伤的分类模型 ResAWDNet, 实现了车辆损伤特征的高效识别与分类。

2) 采用加权动态卷积实现下采样操作, 增强模型特征提取能力, 提高模型对不同特征的适应性; 提出双重注意力模块并嵌入卷积层后, 提高模型对图像中关键信息的捕捉能力, 提升模型的性能与准确性。

3) 基于真实事故案例采集图像, 验证了本文提出的 ResAWDNet 模型在车辆损伤分类任务中的可行性和优越性。

1 相关工作

1.1 图像分类算法

图像分类是计算机视觉领域的重要研究方向之一, 其旨在利用图像的视觉特征自动将图片划分到某一类别^[8]。常见的图像分类算法包括传统图像分类算法和基于深度学习的图像分类算法^[9]。传统算法需人工设计特征提取器通过 SVM (Support Vector Machine)、随机森林等分类器进行决策^[10], 例如贺敏雪等^[11]融合车标对称性特征通过 CRC (Collaborative Representation Based Classification) 分类器识别; ANANDA 和 PUTRI^[12]利用 K-最近邻算法对车辆损伤特征进行分类匹配, 结合案例推理借鉴过往案例经验, 有效提高了车损检测的准确性与效率。然而, 人工设计的特征在捕捉复杂特征形态方面存在局限性, 难以契合实际业务中车辆损伤多样化的需求。

随着深度学习技术的迅猛发展, 基于深度学习的车辆损伤图像分类算法在近年来逐渐成为研究热点。MISHRA 等^[13]基于 YOLOv8 和 ResNet50 构建的 MD R-CNN (Multi-Detection Region-Convolutional Neural Networks) 模型, 通过多尺度特征融合提升车辆损伤分类性能, 但对不良光照下的损伤识别效果有限; PENG 等^[14]提出了多视角融合与对齐的车辆损伤检测模型 MVA-CDD (Multi-View Car Damage Detection Model), 但对小尺度损伤的敏感度较低; SHUBHAM 和 BANERJEE^[15]运用卷积神经网络(Convolutional Neural Networks, CNN)提取车辆损伤图像特征, 并结合 SVM 进行分类, 有效提高了损伤识别精度, 但对复杂光照、多角度拍摄的损伤图像鲁棒性欠佳。王心旷^[16]提出基于可变形卷积网络(Deformable Convolution Networks, DCN)改进的 DCN+算法, 通过集成多尺度训练策略、采用混合焦点损失函数等手段, 提升

了小尺度车损目标的识别准确率, 但在刮擦、凹陷等复杂类别的识别精度方面仍有待进一步提高。

尽管上述研究通过优化网络架构、改进特征提取机制以及创新损失函数设计等手段, 在模型精度提升、计算效率优化等方面取得一定进展, 但在车险理赔的特殊应用场景中仍面临着形态相似损伤特征区分困难、模型泛化能力弱的问题。为此, 本文提出了一种适用于车辆损伤的分类网络, 旨在为车辆定损提供更准确、高效的解决方案。

1.2 下采样方法

下采样(又名降采样), 是指通过一定的算法或操作, 将数据的空间分辨率降低、减少数据量的过程。在深度学习中, 较低层次的特征通常对应于局部的细节信息, 而较高层次的特征则更加抽象, 反映了全局结构和语义信息。通过下采样, 模型可以逐渐从局部细节中抽象出更具概括性的特征, 从而更好地适应不同的任务需求。如在图像分类任务中, ResNet^[17]通过步幅卷积实现下采样操作, 减少特征图的尺寸, 从而降低计算复杂度; 在目标检测任务中, Faster R-CNN(Faster Region-Convolutional Neural Networks)^[18]通过最大池化和步幅卷积实现下采样, 用于生成不同尺度的特征图, 以便检测不同大小的目标; 在语义分割任务中, DeepLab^[19]系列通过空洞卷积和步幅卷积实现下采样, 以便更好地捕捉图像的上下文信息; 在生成对抗网络(Generative Adversarial Networks, GAN)中, Progressive GAN^[20]通过步幅卷积实现下采样, 生成不同分辨率的图像。

然而, 目前常用池化操作和步幅卷积的下采样方法存在一定局限性。一方面其采样特征难以适配多样化图像内容, 面对复杂多变的图像场景, 缺乏灵活性与针对性, 导致特征提取效果不佳。另一方面, 在处理过程中, 全局性信息利用不足, 模型无法充分捕捉图像全局结构与语义关联, 限制了对图像整体理解, 影响任务精度与性能^[21]。针对这一问题, 本文提出了加权动态卷积下采样方式, 通过自适应调整卷积核权重, 实现对不同图像区域的有针对性采样, 提高模型的泛化性。

1.3 注意力机制

注意力机制能够帮助模型聚焦于图像中与任务相关的区域, 改善模型对于关键信息的处理能力, 从而提升模型的性能^[22]。如 HU 等^[23]关注通道之间的关系提出的 SE(Squeeze-and-Excitation Networks); JADERBERG 等^[24]基于空间关系注意

力提出的 STN(Spatial Transformer Networks)和 WOO 等^[25]提出融合空间注意力机制和通道注意力机制的 CBAM(Convolutional Block Attention Module)等注意力机制,验证了注意力机制增强模型表征能力的有效性。

近年来,尽管注意力机制在深度学习领域快速发展,但在车辆损伤图像分类任务中表现欠佳。SeerAttention^[26]通过可学习的门自适应选择重要区块平衡准确性和速度,但主要针对面向语言任务,难以捕捉车辆损伤的非规则分布。EMA(Efficient Multi-scale Attention)^[27]依赖时序特征加权,难以在静态图像中动态捕捉局部关键区域。EPSA(Efficient Pyramid Squeeze Attention)^[28]多尺度空间聚合但忽略了通道交互,ECA(Efficient Channel Attention)^[29]

纯通道增强缺乏空间敏感性,复杂损伤形态的表征能力受限。综上,由于任务目标、数据特性及模型需求存在差异,现有的注意力机制难以直接应用到车辆损伤分类任务中。

2 本文模型

2.1 整体结构

针对车险理赔客户上传的车辆损伤图像中存在损伤类型形态相似、分类困难的问题,本文提出了一种适用于车辆损伤分类的模型ResAWDNet。以ResNet50作为基础框架,并引入加权动态卷积(Weighted Dynamic Convolution, WDCov)与双重注意力机制(Dual Attention Mechanism, DAM)来优化车辆损伤分类模型。模型结构如图2所示。

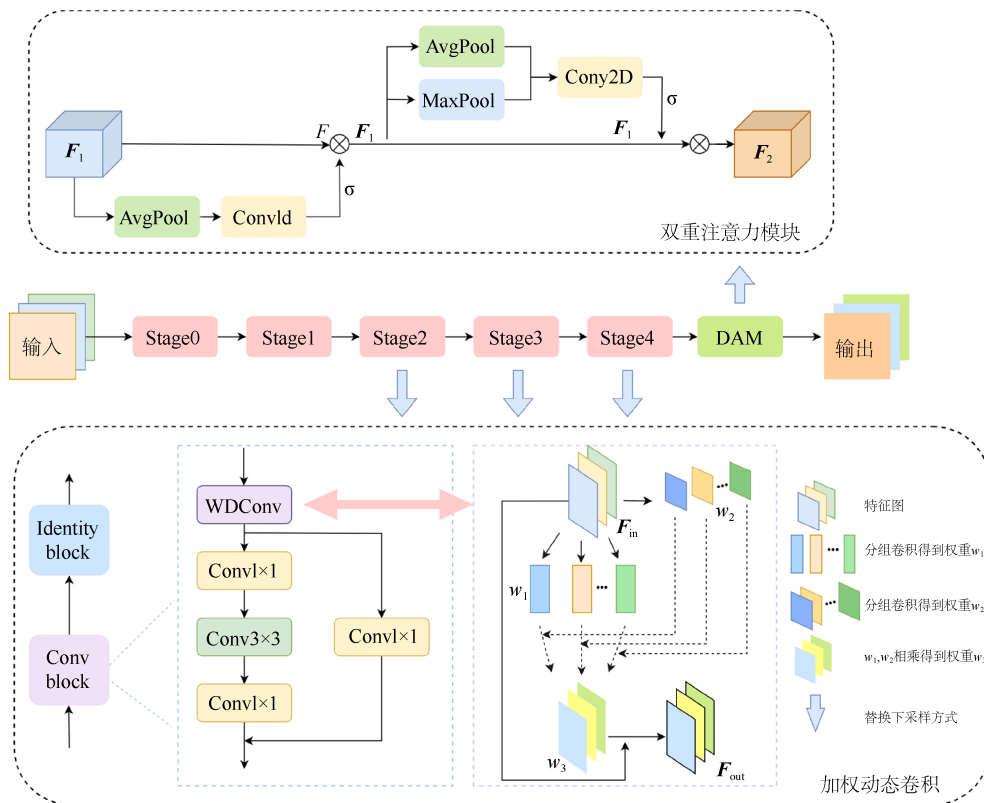


图2 ResAWDNet 模型结构图

Fig. 2 Structure of ResAWDNet model

图像输入模型后,首先会经过Stage0和Stage1的初步处理。从Stage2开始,模型采用加权动态卷积替代ResNet50网络Stage2~Stage4中原有的下采样操作。接着依次经过卷积操作,完成特征提取与变换,最终输出特征图。Stage2~Stage4均采用这样的加权动态卷积处理,各阶段处理后的特征图作为后续阶段的输入,实现了特征的逐步丰富与

细化。

经 Stage4 处理后的特征图,作为双重注意力模块的输入。在此模块中,通过通道注意力与空间注意力依次处理,引导模型重点关注图像中的关键区域以及具有重要信息的特征通道,提高模型对于损伤部位的敏感程度,进一步提升整体分类性能。

2.2 加权动态卷积

加权动态卷积下采样层的核心作用是动态地调整卷积核权重, 利用调整后的权重对经过填充处理的输入特征图进行分组卷积操作。相比于传统的固定卷积核, 动态卷积能够根据输入数据的特征动态调整卷积核的权重, 有助于提取更具有代表性的

特征。在下采样的过程中, 模型需保留更多有用的信息, 以减少特征信息的丢失。本文设计并采用加权动态卷积进行下采样操作, 使得下采样操作更加灵活, 可以根据输入数据的不同, 动态地调整卷积核的权重, 从而更好地适应不同的任务需求。加权动态卷积结构如图3所示。

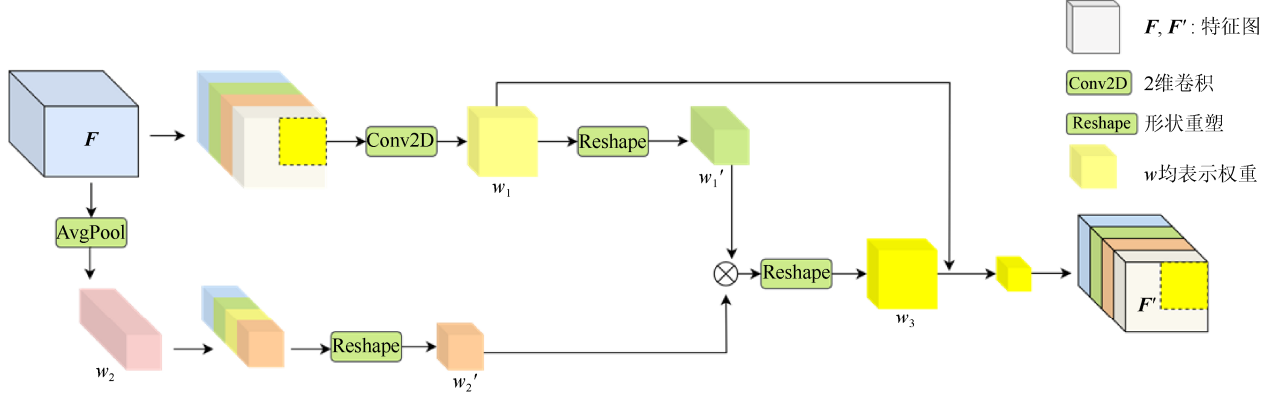


图 3 加权动态卷积结构图

Fig. 3 Structure of weighted dynamic convolution

输入的特征图 F 首先通过计算得卷积核权重 w_1 , 该过程采用分组卷积的形式, 即将输入通道分成若干组, 每组独立进行卷积操作, 即

$$w_1 = \text{Conv}(F), \text{BN}(F) \quad (1)$$

式中: Conv 表示卷积操作; BN 表示批量归一化。

随后将权重 w_1 重新塑形, 并在通道维度上重复, 使其与输入特征图的通道数匹配。塑形后的权重为

$$w_1' \in R^{(B, G, 1, K^2, H, W)} \quad (2)$$

式中: B 表示一次处理的样本大小; G 表示分组数量; 1 为维度; K 表示卷积核大小; H 和 W 分别表示输出特征图的高度和宽度。

再通过计算得到权重 w_2 , 重新塑形后记为 w_2' 。将权重 w_1' 与 w_2' 相乘, 以调整每个通道的权重, 输出结果记为 w_3 。此过程可以表示为

$$w_2 = \text{Conv}[\text{Conv}(\text{AvgPool}(F))] \quad (3)$$

$$w_2' \in R^{(B, G, C/G, 1, 1, 1)} \quad (4)$$

$$w_3 = w_1' \times w_2' \quad (5)$$

式中: AvgPool 表示平均池化; B 表示一次处理的样本大小; G 表示分组数量; C 表示输入张量的通道数。

文献[30]提到, 低通滤波器能够减少图像中噪声对特征的干扰。因此将得到的权重 w_3 经过 Softmax 操作, 约束权重为正并求和为 1, 使得生成的权重是低通的, 从而减弱特征噪声的干扰。随后重新调整 w_3 的形状, 以便在后续的分组卷积操

作中能够正确地进行计算, 即

$$w_3' = \text{reshape}[\text{Softmax}(w_3)] \quad (6)$$

对输入特征图进行反射填充, 并展开生成滑动窗口的特征块。最后, 将得到的权重 w_3' 与展开后的特征块相乘, 并对卷积核维度进行求和, 生成最终的输出特征图, 即

$$\text{pad}_F = \text{pad}(F) \quad (7)$$

$$\text{pad}'_F = \text{reshape}[\text{unfold}(\text{pad}_F)] \quad (8)$$

$$F_{\text{out}} = \text{reshape}[\text{sum}(w_3' \times \text{pad}'_F)] \quad (9)$$

加权动态卷积模块将自适应加权的思想与分组卷积相结合, 充分利用局部信息和全局信息调整卷积核权重, 使模型在轻量化的同时能够加强表达能力, 整体算法流程如下:

算法 1. WDCConv。

输入: 特征图 x 。

输出: 最终的输出特征 x_{out} 。

1. $w_1 \leftarrow \text{reshape}[\text{BN}(\text{Conv}(x))], \text{repeat}[\text{BN}(\text{Conv}(x))]$
2. **if** $\text{use_channel} == \text{true}$ **then**
3. $w_2 \leftarrow \text{Conv}[\text{Conv}(\text{Avg}(x))]. \text{reshape}$
4. **else**
5. $w_2 \leftarrow 1$
6. **end**
7. $w_3 \leftarrow w_1, w_2$
8. **for** i **from** 2 **to** 3 **do**:对高度和宽度维度进

3 实验结果

3.1 数据及实验环境

本文使用车辆损伤分类数据集对本文模型进行评估, 车辆损伤分类数据集由保险公司提供, 该数据集来源于其在车险理赔业务中的实际车辆损伤勘查记录, 包含车身、前窗和侧窗等多类车辆部位, 涵盖错位、玻璃破损、玻璃裂痕、中度变形、轻微变形、丢失、车身划痕、车身刮擦、重度变形和撕裂共 10 类损伤类型。总计 30 463 张图片, 其中训练集与测试集样本图像比例为 3 : 1, 分别为 22 850 和 7 613 张。数据集各类图片数量见表 1。

表 1 数据集构成

Table 1 Data set composition

损伤类型	训练集/张	测试集/张	图片数量/张
错位	2 749	916	3 665
玻璃破损	913	304	1 217
玻璃裂痕	933	311	1 244
中度变形	1 493	497	1 990
轻微变形	1 488	496	1 984
丢失	2 683	894	3 577
车身划痕	4 336	1445	5 781
车身刮擦	4 116	1371	5 487
重度变形	1 455	485	1 940
撕裂	2 684	894	3 578

为进一步评估本模型的性能, 本文还选用公开的车辆损伤数据集 CarDD^[32]进行实验, 该数据集涵盖凹痕、划痕、裂纹、玻璃破碎、轮胎漏气和车灯损坏 6 种常见车辆损伤类型, 适用于多种计算机视觉任务。

本文所用操作系统为 Ubuntu 18.04 LTS, GPU 为 Geforce RTX 3090 显卡, CUDA 版本为 12.5, 以 Python3.8 为程序开发语言。采用随机梯度下降法 (Stochastic Gradient Descent, SGD) 作为优化模型的方法, 输入图像尺寸设为 224×224×3 像素, 批次大小设为 8, 经实验确定学习率设为 0.000 1, 迭代次数设为 100。学习率调整结果见表 2、训练过程准确率与损失曲线如图 5 所示。

表 2 学习率调整结果

Table 2 Outcome of learning rate tuning

学习率	Acc_1%
0.01	51.67
0.005	53.58
0.001	62.45
0.000 5	69.50
0.000 1	73.79
0.000 05	73.19
0.000 01	72.81

注: 加粗数据表示最优值。

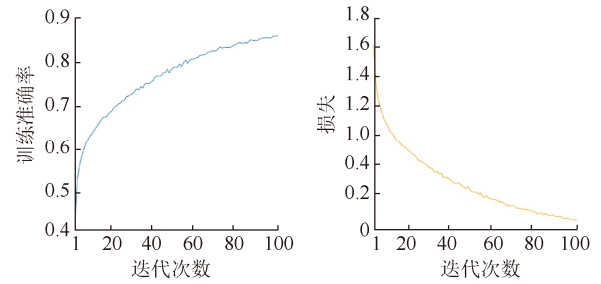


图 5 训练过程准确率与损失曲线

Fig. 5 Training process accuracy and loss curve

3.2 实验结果及分析

以每一类别的准确率 Accuracy、每一类别的精确率 Precision、整体的 Top-1 准确率和 Top-5 准确率作为车辆损伤分类的评价标准。其中 Top-1 准确率是指排名第一的类别与实际结果相符的准确率, 简称为准确率; Top-5 准确率是指排名前 5 的类别包含实际结果的准确率。两者大小均介于 [0,1] 之间, 数值越大, 表示分类效果越好。每一类别的准确率及精确率计算式分别为

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (18)$$

$$Precision = \frac{TP}{TP + FP} \quad (19)$$

式中: TP (True Positive) 表示正样本被正确识别的数量; FP (False Positive) 表示误报的负样本数量; TN (True Negative) 表示负样本被正确识别的数量; FN (False Negative) 表示漏报的正样本数量。

3.2.1 消融实验

本文以 ResNet 模型为基线模型实现了所提出模型 ResAWDNet, 将各模块在保险公司真实应用场景下的车辆损伤分类数据集上进行实验并测试了 2 种定量指标, 实验结果见表 3。Baseline 代表基础模型, Baseline+WDCConv 代表引入加权动态卷积下采样的模型, Baseline+DAM 代表加入双重注意力的模型, ResAWDNet 代表结合双重注意力和加权动态卷积下采样的模型。

表 3 消融实验结果

Table 3 Results of ablation experiments

模型	Acc_1%	Acc_5%
Baseline	71.88	97.24
Baseline+ WDCConv	73.05	97.16
Baseline+DAM	72.97	97.14
ResAWDNet (本文模型)	73.79	97.68

注: 加粗数据表示最优值。

由表 3 可知, 基线模型 Top-1 准确率、Top-5 准确率分别为 71.88% 和 97.24%, 而分别引入

WDCConv 和 DAM 下采样使 Top-1 准确率分别提升了 1.17% 和 1.09%。两者结合构成 ResAWDNet 使 Top-1 准确率达到 73.79%，相较基线模型提升 1.91%，超越单一模块的改进幅度，体现了模块设计的有效性。

表 4 展示了消融过程中各损伤类型的分类准确率。由表中结果可以看出引入 DAM 和 WDCConv 下

采样模块对大多数损伤类型分类准确率均有提升，其中引入 DAM 机制对错位、中度变形 2 类损伤影响较大，采用 WDCConv 下采样对玻璃破损、撕裂等损伤影响较大。尽管消融过程中部分类别的准确率存在一定的波动，但从整体实验结果来看，二者结合使模型在损伤分类中展现出更强的能力，验证 ResAWDNet 的融合策略是有效的。

表 4 消融过程中各损伤类型的分类准确率

Table 4 Accuracy of classification of each injury type during ablation process

损伤类型	Baseline	+DAM	+WDCConv	ResAWDNet
错位	79.26	85.37	82.97	82.10
玻璃破损	76.64	78.62	80.92	77.96
玻璃裂痕	64.95	68.81	67.85	71.06
中度变形	25.75	47.89	23.94	30.99
轻度变形	51.81	46.17	53.23	49.80
丢失	73.60	73.60	78.19	80.09
车身划痕	90.73	89.34	88.86	92.25
车身刮擦	73.89	73.01	73.52	75.13
重度变形	52.99	63.30	70.72	61.44
撕裂	63.87	66.89	70.13	67.90

注：加粗数据表示最优值。

3.2.2 对比试验

为验证 DAM 的有效性，本文进行了一系列对比实验。在保证超参数设置相同的情况下，选取 ResNet50 作为基础模型框架，对比加入 SE^[23]通道注意力、CBAM^[25]卷积注意力等注意力机制与加入本文所提出的 DAM 机制后的性能表现，具体内容见表 5。由表中结果可知，相较于未引入任何注意力机制的基线 ResNet50 模型，应用 SE^[23]、CBAM^[25]和 ECA^[29]等注意力机制均能在一定程度上提升模型在图像分类任务上的准确率。然而，在相同的实验条件下，本文提出的加入 DAM 机制的模型表现尤为突出，相较于基线模型分类准确率提高了 1.09%，与近年提出的 EMA 和 EPSA 等注意力相比也占优势。

表 5 注意力机制效果对比

Table 5 Comparison of the effects of attention mechanisms

注意力机制	Acc_1%	Acc_5%
Baseline	71.88	97.24
Baseline+SE ^[23]	72.32	97.48
Baseline+CBAM ^[25]	72.53	97.74
Baseline+EMA ^[27]	72.85	97.33
Baseline+EPSA ^[28]	72.61	97.36
Baseline+ECA ^[29]	72.93	97.62
Baseline+RGA ^[33]	72.49	97.35
Baseline+CPCA ^[34]	72.61	97.22
Baseline+DAM	72.97	97.14

注：加粗数据表示最优值。

图 6 展示了在基线模型中引入 DAM 机制前后的可视化对比结果，通过热力图^[35]呈现了模型在处理图像数据时对不同区域的关注度。在未加入注意力机制的情况下，基线模型的注意力分布较为分散，未能有效聚焦于图像中的关键损伤部位，加入 DAM 机制后，模型更加专注于图像中的损伤区域，有效减少了无关信息的干扰，显著提升了模型对关键信息的捕捉能力。如在变形损伤中，加入 DAM 机制后热力图的高热度区域与损伤位置贴合更为准确，能够准确覆盖变形区域，表明其能更有效地聚焦关键损伤特征。对比基线模型，加入 DAM 机制后热力图中高热度区域集中在玻璃实际破损或裂痕周围，即使在反光影响下，也能有效聚焦于损伤本身，充分体现出 DAM 机制在反光等复杂光照条件下对损伤检测的优势。

基于车辆损伤数据集，对基线模型和本文提出的 ResAWDNet 模型进行了对比。从表 6 结果可以看出，ResAWDNet 在多数损伤类型上的分类准确率显著高于基线模型，其中，“玻璃裂痕”分类准确率由 64.95% 提升至 71.06%，相较于基线模型提升了 6.11%；“重度变形”分类准确率由 52.99% 提升至 61.44%，提升了 8.45%。尽管部分类别的分类准确率仍有提升空间，但从整体角度看，ResAWDNet 在多数类别上实现了准确率的提升。

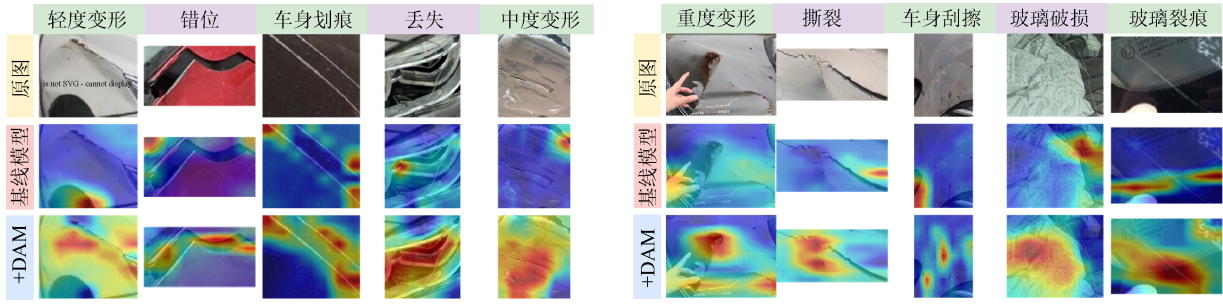


图 6 引入双重注意力机制前后的可视化对比结果

Fig. 6 Visualization comparison results before and after the introduction of the dual attention mechanism

表 6 各损伤类型分类效果对比/%

Table 6 Comparison of classification effects by injury type/%

损伤类型	Baseline		ResAWDNet	
	Acc	Pre	Acc	Pre
错位	79.26	83.75	82.10	81.56
玻璃破损	76.64	75.08	77.96	81.72
玻璃裂痕	64.95	73.49	71.06	74.16
中度变形	25.75	43.86	30.99	49.04
轻度变形	51.81	49.68	49.80	49.60
丢失	73.60	72.78	80.09	74.11
车身划痕	90.73	88.09	92.25	86.33
车身刮擦	73.89	68.72	75.13	71.23
重度变形	52.99	67.95	61.44	62.47
撕裂	63.87	68.63	67.90	77.03

注: 加粗数据表示最优值。

为验证本文所提出模型的优越性, 将数据集分别在 AlexNet 和 GoogleNet 等经典的卷积神经分类

网络模型、RepLKNet、StartNet 等前沿的网络模型和 ViT、SwinT 等 Transformer 模型中进行训练和测试, 结果见表 7。通过对比各模型的实验结果, 可以明显看出, 与其他经典的 CNN 模型相比, 本文提出的结合 DAM 和 WDCConv 下采样的 ResAWDNet 模型在分类任务中展现出了显著的优势。具体来说, ResAWDNet 在 Top-1 和 Top-5 的准确率分别达到了 73.79%和 97.68%, 均优于如 AlexNet(57.22%)、GoogleNet(62.17%)、MobileNet (58.08%)等经典模型, 以及 MobileViT(72.19%)、DenseNet(72.72%)、RepLKNet(72.75%)等前沿模型, 充分体现了 ResAWDNet 模型在车辆损伤分类任务中的有效性和先进性。

表 7 与其他模型对比

Table 7 Comparison with other models

模型	Acc_1/%	Acc_5/%	Flops	Params/M	
AlexNet ^[36]	57.22	92.51	309.16 M	14.60	
GoogleNet ^[37]	62.17	94.33	1.58 G	6.99	
MobileNet ^[38]	58.08	94.02	327.55 M	3.50	
ShuffleNet ^[39]	71.93	97.48	152.71 M	2.28	
DenseNet ^[40]	72.72	97.11	2.90 G	7.98	
EfficientNet ^[41]	69.80	96.97	412.83 M	5.29	
RegNet ^[42]	72.77	97.65	207.35 M	2.32	
EfficientNetv2 ^[43]	71.97	97.01	2.89 G	21.46	
FasterNet ^[44]	73.36	97.74	4.45 G	31.18	
RepLKNet ^[45]	72.75	97.52	-	304.66	
StarNet ^[46]	60.28	94.48	427.33 M	2.87	
ResNet ^[17]	71.88	97.24	4.13 G	25.56	
Vision Transformer ^[47]	VIT-B16	64.59	95.97	16.88 G	103.03
	VIT-B32	68.53	97.02	4.37 G	88.19
	VIT-L16	72.32	97.90	59.69 G	304.12
	VIT-L32	66.08	96.64	15.28 G	328.89
Swin Transformer ^[48]	SwinT-T	72.76	97.60	4.37 G	28.27
	SwinT-S	73.11	97.20	8.55 G	49.56
	SwinT-B	72.90	97.65	23.57 G	109.07
MobileViT ^[49]	72.19	97.29	273.67 M	1.27	
ResAWDNet	73.79	97.68	3.94 G	26.42	

注: 加粗数据表示最优值。

为进一步评估本模型在损伤分类任务中的泛化性, 选用公开的车辆损伤数据集 CarDD 进行对比实验。表 8 展示了本方法与当前主流图像分类算法在 CarDD 数据集上的对比结果, 从中可看出, 本模型在分类表现上更为优异。

3.2.3 预测可视化

为检验 ResAWDNet 模型的分类准确性, 对其分类结果进行展示, 通过输入不同的图片, 预测其损伤类型。图 7 展示了基线模型和本文 ResAWDNet 模型的分类结果。与基线模型相比, ResAWDNet 模型展现出了更为出色的分类能力, 在处理相似或模糊损伤特征时也表现出了较高的准确率, 充分验

证了 ResAWDNet 模型在损伤类型分类任务上的优越性和可靠性。

表 8 在 CarDD 数据集上的对比
Table 8 Comparison on the CarDD dataset

模型	Acc_1/%	Acc_5/%
ShuffleNet ^[39]	58.77	99.60
DenseNet ^[40]	59.09	99.84
FasterNet ^[44]	54.81	99.75
ResNet ^[17]	59.18	99.51
ViT-L16 ^[47]	58.85	99.76
SwinT-S ^[48]	59.82	99.68
MobileViT ^[49]	60.15	99.78
ResAWDNet	60.43	99.68

注: 加粗数据表示最优值。

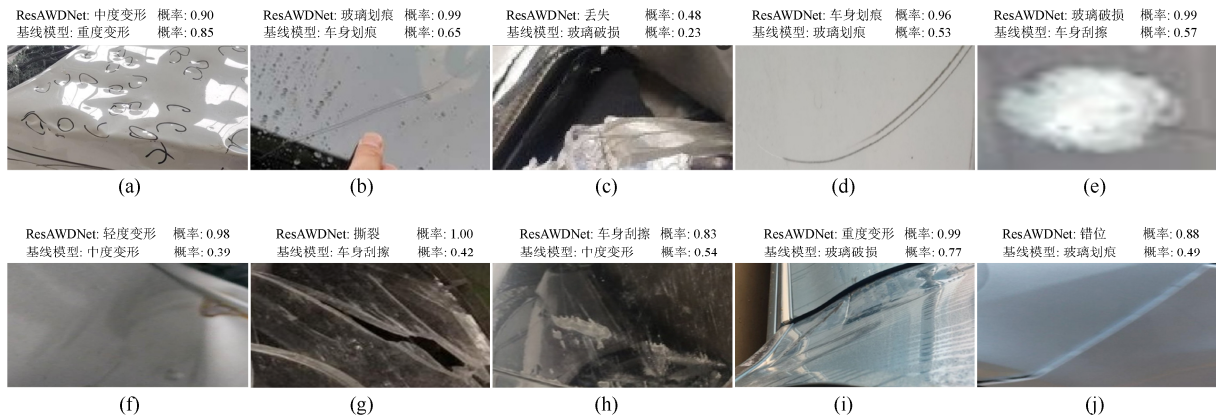


图 7 分类结果展示((a) 中度变形; (b) 玻璃划痕; (c) 丢失; (d) 车身划痕; (e) 玻璃破损; (f) 轻度变形; (g) 撕裂; (h) 车身刮擦; (i) 重度变形; (j) 错位)

Fig. 7 Classification results display ((a) Medium deformation; (b) Glass crack; (c) Missing; (d) Scratch; (e) Glass breakage; (f) Mild deformation; (g) Tearing; (h) Scratches; (i) Severe deformation; (j) Dislocation)

4 结束语

本文研究致力于提升车辆损伤分类算法的精度。本文提出了 DAM 机制与 WDCConv 模块, 并将二者相结合, 以 ResNet50 作为基础框架构建了 ResAWDNet 模型, 实现了在车辆损伤分类任务中的性能提升。WDCConv 模块能够根据不同输入动态调整卷积核, 优化特征提取过程, 有效减轻数据集中低质量图像对分类性能的不利影响。DAM 机制能够同时捕捉通道和空间维度的关键信息, 使模型在处理车辆损伤的相似图像数据时更加高效和准确。在车辆损伤数据集上的实验结果表明, ResAWDNet 模型在车辆损伤分类任务中表现出色。与基线模型相比, ResAWDNet 准确率提升了 1.91%, 达到了 73.79%。在各类损伤类型的分类上, ResAWDNet 也普遍表现出更高的准确率, 特别是在错位、玻璃裂痕和丢失等类型上, 准确率提升尤

为显著。

本文通过提出 DAM 机制与 WDCConv 模块, 并将其有效结合构建 ResAWDNet 模型, 针对车辆损伤分类任务中的图像损伤特征形态相似性高、分类难度大的问题进行了深入研究, 为车辆多形态部件损伤缺陷精确识别提供了有效的技术方案。然而, 该方法也存在一定的局限性, 对相互遮挡的损伤区域特征提取能力有限, 分类准确率有待提升。在未来的工作中, 将探索更鲁棒的特征提取与解耦机制, 与多目标检测算法结合, 部署到车辆理赔现场实际运行环境中, 从客户端上传待定损图像到对车辆部件状态研判, 实现端到端部署和运行。

参考文献 (References)

- [1] 赵子豪, 申颖, 李薇. 基于图像识别的车辆智能定损应用研究[J]. 保险职业学院学报, 2019, 33(3): 73-77.
ZHAO Z H, SHEN Y, LI W. Application and value research about apps of vehicle survey and loss assessment based on

- image recognition[J]. *Journal of Insurance Professional College*, 2019, 33(3): 73-77 (in Chinese).
- [2] 翟永杰, 李佳蔚, 陈年昊, 等. 融合改进 Transformer 的车辆部件检测方法[J]. *图学学报*, 2024, 45(5): 930-940.
ZHAI Y J, LI J W, CHEN N H, et al. The vehicle parts detection method enhanced with Transformer integration[J]. *Journal of Graphics*, 2024, 45(5): 930-940 (in Chinese).
- [3] 武兵, 田莹. 基于注意力机制的多尺度道路损伤检测算法研究[J]. *图学学报*, 2024, 45(4): 770-778.
WU B, TIAN Y. Research on multi-scale road damage detection algorithm based on attention mechanism[J]. *Journal of Graphics*, 2024, 45(4): 770-778 (in Chinese).
- [4] LIU Q, HUANG X H, SHAO X Y, et al. Industrial cylinder liner defect detection using a transformer with a block division and mask mechanism[J]. *Scientific Reports*, 2022, 12(1): 10689.
- [5] PARK J K, KWON B K, PARK J H, et al. Machine learning-based imaging system for surface defect inspection[J]. *International Journal of Precision Engineering and Manufacturing-Green Technology*, 2016, 3(3): 303-310.
- [6] WU S Q, ZHAO S Y, ZHANG Q Q, et al. Steel surface defect classification based on small sample learning[J]. *Applied Sciences*, 2021, 11(23): 11459.
- [7] LIU W, QIU J, WANG Y, et al. Multiscale feature fusion convolutional neural network for surface damage detection in retired steel shafts[J]. *Journal of Computing and Information Science in Engineering*, 2024, 24(4): 041005..
- [8] 王瑞芳. 基于字典学习的图像分类算法研究[D]. 重庆: 重庆邮电大学, 2020.
WANG R F. Research on image classification algorithm based on dictionary learning[D]. Chongqing: Chongqing University of Posts and Telecommunications, 2020 (in Chinese).
- [9] 张鹏飞, 石志良, 李晓垚, 等. 基于深度学习的主轴承盖分类识别算法[J]. *图学学报*, 2021, 42(4): 572-580.
ZHANG P F, SHI Z L, LI X Y, et al. Classification algorithm of main bearing cap based on deep learning[J]. *Journal of Graphics*, 2021, 42(4): 572-580 (in Chinese).
- [10] 董潇. 卷积神经网络的图像分类优化算法研究[D]. 淮南: 安徽理工大学, 2020.
DONG X. Research on image classification optimization algorithm of convolutional neural network[D]. Huainan: Anhui University of Science & Technology, 2020 (in Chinese).
- [11] 贺敏雪, 余焯, 程茹秋. 特征增强策略驱动的车标识别[J]. *中国图象图形学报*, 2021, 26(5): 1030-1040.
HE M X, YU Y, CHENG R Q. Vehicle logo recognition method based on feature enhancement[J]. *Journal of Image and Graphics*, 2021, 26(5): 1030-1040 (in Chinese).
- [12] ANANDA B, PUTRI R A. K-nearest neighbor algorithm and case base reasoning on xenia car damage detection expert system[J]. *Journal of Computer Networks, Architecture and High Performance Computing*, 2024, 6(2): 633-646.
- [13] MISHRA S, KAMAL D, SENTHIL KUMAR K. Vehicle damage identification using deep learning techniques[C]//2024 IEEE International Students' Conference on Electrical, Electronics and Computer Science. New York: IEEE Press, 2024: 1-6.
- [14] PENG J B, DONG S B, YUAN H, et al. Car damage detection based on multi-view fusion and alignment: dataset and method[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2025, 26(4): 4717-4730.
- [15] SHUBHAM, BANERJEE D. Robust car damage identification through CNN and SVM techniques[C]//The 4th International Conference on Technological Advancements in Computational Sciences. New York: IEEE Press, 2024: 101-107.
- [16] 王心旷. 基于深度学习的车辆外观损伤识别及其图像生成方法研究[D]. 合肥: 中国科学技术大学, 2024.
WANG X K. Research on car exterior damage recognition and image generation based on deep learning[D]. Hefei: University of Science and Technology of China, 2024 (in Chinese).
- [17] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2016: 770-778.
- [18] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [19] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs[EB/OL]. [2024-12-28]. <https://arxiv.org/abs/1412.7062>.
- [20] KARRAS T, AILA T, LAINE S, et al. Progressive growing of GANs for improved quality, stability, and variation[EB/OL]. [2024-12-28]. <https://openreview.net/forum?id=Hk99zCeAb>.
- [21] 顾正华, 刘嘎琼, 邵长斌, 等. 深度检测方法中融合大小感受野机制的下采样算法[J]. *计算机科学与探索*, 2024, 18(10): 2727-2737.
GU Z H, LIU G Q, SHAO C B, et al. Downsampling algorithm with fusion of different receptive field sizes in deep detection methods[J]. *Journal of Frontiers of Technology*, 2024, 18(10): 2727-2737 (in Chinese).
- [22] 谢东升. 基于深度学习的车辆智能定损算法研究[D]. 天津: 天津大学, 2019.
XIE D S. Research on vehicle intelligent damage location algorithm based on deep learning[D]. Tianjin: Tianjin University, 2019 (in Chinese).
- [23] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2018: 7132-7141.
- [24] JADERBERG M, SIMONYAN K, ZISSERMAN A. Spatial transformer networks[C]//The 29th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2015: 2017-2025.
- [25] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//The 15th European Conference on Computer Vision. Cham: Springer, 2018: 3-19.
- [26] GAO Y, ZENG Z, DU D, et al. SeerAttention: learning intrinsic sparse attention in your LLM[EB/OL]. [2024-12-28]. <https://arxiv.org/abs/2410.13276>.
- [27] OUYANG D L, HE S, ZHANG G Z, et al. Efficient multi-scale attention module with cross-spatial learning[C]//ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing. New York: IEEE Press, 2023: 1-5.
- [28] ZHANG H, ZU K K, LU J, et al. EPSANet: an efficient pyramid squeeze attention block on convolutional neural network[C]//The 16th Asian Conference on Computer Vision. Cham: Springer, 2022: 541-557.
- [29] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: efficient channel attention for deep convolutional neural networks[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2020: 11531-11539.
- [30] CHEN L W, FU Y, WEI K X, et al. Instance segmentation in

- the dark[J]. *International Journal of Computer Vision*, 2023, 131(8): 2198-2218.
- [31] WANG F, JIANG M Q, QIAN C, et al. Residual attention network for image classification[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2017: 3156-3164.
- [32] WANG X K, LI W J, WU Z C. CarDD: a new dataset for vision-based car damage detection[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2023, 24(7): 7202-7214.
- [33] ZHANG Z Z, LAN C L, ZENG W J, et al. Relation-aware global attention for person re-identification[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2020: 3183-3192.
- [34] HUANG H J, CHEN Z G, ZOU Y, et al. Channel prior convolutional attention for medical image segmentation[J]. *Computers in Biology and Medicine*, 2024, 178: 108784.
- [35] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-cam: visual explanations from deep networks via gradient-based localization[C]//2017 IEEE International Conference on Computer Vision. New York: IEEE Press, 2017: 618-626.
- [36] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]//The 26th International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2012: 1097-1105.
- [37] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2015: 1-9.
- [38] HOWARD A G, ZHU M L, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[EB/OL]. [2024-12-28]. <https://arxiv.org/abs/1704.04861>.
- [39] ZHANG X Y, ZHOU X Y, LIN M X, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2018: 6848-6856.
- [40] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2017: 2261-2269.
- [41] TAN M X, LE Q. EfficientNet: rethinking model scaling for convolutional neural networks[EB/OL]. [2024-12-28]. <https://proceedings.mlr.press/v97/tan19a.html>.
- [42] XU J, PAN Y, PAN X L, et al. RegNet: self-regulated network for image classification[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2023, 34(11): 9562-9567.
- [43] TAN M X, LE Q. EfficientNetV2: smaller models and faster training[EB/OL]. [2024-12-28]. <https://proceedings.mlr.press/v139/tan21a>.
- [44] CHEN J R, KAO S H, HE H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2023: 12021-12031.
- [45] DING X H, ZHANG X Y, HAN J G, et al. Scaling up your kernels to 31×31: revisiting large kernel design in CNNs[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2022: 11953-11965.
- [46] MA X, DAI X Y, BAI Y, et al. Rewrite the stars[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2024: 5694-5703.
- [47] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: transformers for image recognition at scale[EB/OL]. [2024-12-28]. <https://openreview.net/forum?id=YicbFdNTTy>.
- [48] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]//2021 IEEE/CVF International Conference on Computer Vision. New York: IEEE Press, 2021: 9992-10002.
- [49] MEHTA S, RASTEGARI M. MobileViT: light-weight, general-purpose, and mobile-friendly vision transformer[EB/OL]. [2024-12-28]. <https://openreview.net/forum?id=vh-0sUt8HIG>.