

汽车发动机积碳图像细粒度分类方法研究

陈香, 陈友兴*, 张航佳, 刘昱彤

(中北大学信息与通信工程学院, 山西太原 030051)

摘要: 由于汽车发动机积碳图像在传统模型中的识别效果不佳, 结合实际应用需求, 提出了一种针对积碳图像细粒度特征的分类方法。对于喷油嘴积碳图像分类任务中难以聚焦形态差异大的喷油嘴区域的问题, 提出基于CA注意力机制的积碳图像识别方法。通过添加CA注意力机制, 实现不同通道和空间位置权重的调整, 提升模型对关键特征的聚焦效果。针对活塞顶积碳图像分类任务中积碳分布较分散而难以全面捕捉到关键空间特征的问题, 提出基于SG-former的自适应积碳图像识别方法。通过自动分配全局注意力权重, 有效提取重点区域的细粒度特征, 同时通过BlurPool池化最大化积碳特征在下采样过程的平移不变性。最终, 喷油嘴部位和活塞顶部位的积碳程度识别准确率分别达到88.52%和88.20%。

关键词: 积碳检测; 图像分类; RepVGG; 注意力机制; SG-former

中图分类号: TP391 **文献标识码:** A **doi:** 10.62756/csjs.1671-7449.2025024

引用格式: 陈香, 陈友兴, 张航佳, 等. 汽车发动机积碳图像细粒度分类方法研究[J]. 测试技术学报, 2025, 39(2): 121-129.

CHEN Xiang, CHEN Youxing, ZHANG Hangjia, et al. Research on fine-grained classification method of carbon deposit images in automobile engines[J]. Journal of Test and Measurement Technology, 2025, 39(2): 121-129.

Research on Fine-Grained Classification Method of Carbon Deposit Images in Automobile Engines

CHEN Xiang, CHEN Youxing*, ZHANG Hangjia, LIU Yutong

(School of Information and Communication Engineering, North University of China, Taiyuan 030051, China)

Abstract: Because of the poor recognition effect of carbon deposit images of automobile engines in the traditional model and the practical application requirements, a classification method for the fine-grained features of carbon deposit images is proposed. For the problem that it is difficult to focus on the nozzle region with large morphological differences in the nozzle carbon image classification task, a carbon image recognition method based on the CA mechanism is proposed. The weights of different channels and spatial locations are adjusted to improve the focusing effect of the model on key features by adding the CA attention mechanism. To address the problem that it is difficult to fully capture the key spatial features in the task of carbon image classification of the piston top, an adaptive carbon image recognition method based on SG-former is proposed, which effectively extracts the fine-grained features in the key areas by

收稿日期: 2024-09-29

基金项目: 山西省回国留学人员科研资助项目(2022-145); 山西省重点研发计划资助项目(202302020101008); 山西省研究生科研创新资助项目(2024KY607)

作者简介: 陈香(2000-), 女, 硕士生, 主要从事计算机视觉和图像处理研究。E-mail: 18734292079@163.com。

* **通信作者:** 陈友兴(1978-), 男, 教授, 博士, 主要从事人工智能、信息探测与处理研究。E-mail: chenyouxing@nuc.edu.cn。

automatically assigning the global attention weights and maximizes the translational invariance of the carbon features in the downsampling process through BlurPool pooling. Eventually, the accuracy of carbon degree recognition in the injector nozzle area and piston top area reaches 88.52% and 88.20%, respectively.

Key words: carbon deposit detection; image classification; RepVGG; attention mechanism; SG-former

0 引言

积碳是汽车发动机内汽油或柴油与润滑油不完全燃烧的产物,主要集中在进气道、进气阀、喷油嘴、缸壁和活塞顶等部位。它不仅会导致汽车冷启动困难、怠速不稳、发动机过热、油耗增加、动力下降、机件早期损耗等多种故障,还会对发动机整体性能产生严重的负面影响。因此,定期进行汽车的积碳检测并清理不仅有助于维持车辆的正常运行,还能减少对环境的不良影响。

目前,汽车发动机积碳检测的研究重点是基于传感器和内窥镜的视觉积碳检测。Roe等^[1]通过测量发动机油道温度以确定积碳堆积程度。Wang等^[2]通过收集发动机运行条件下的离子和气缸压力信号,利用循环神经网络对发动机积碳进行检测。钱飞等^[3]提出了一种基于内燃机爆震声音的积碳分布测定装置,通过采集与放大音频信号,分析发动机缸内的积碳分布情况进行可视化。可以看出,基于传感器的积碳检测需要发动机在某些特定条件并且通过其他信息间接判断积碳堆积情况,检测较为繁琐,局限性较大。相比之下,基于内窥镜的检测方法更加直观,更易广泛推广。卢曼斯等^[4]通过采集活塞表面积碳图像来推测积碳分布;尚忱等^[5]提出一种车辆积碳监测系统,对存在异常气味的发动机内积碳进行视觉图像检测。然而,以上检测通常依赖专业人员,其判别过程具有较强的主观性。因此,肖剑平等^[6]采用机器学习技术对各部位积碳图像与预设图像进行比较,以实现积碳程度的自动识别,提高了检测效率并降低了人工成本,但识别率仍较低。

近年来,深度学习在图像识别中表现卓越^[7-11],尤其是卷积神经网络,通过自动学习层次化特征,显著提升了分类和可视化效果,在工业应用中展现了良好的部署性、稳定性和可优化性^[12-14]。基于此,本文提出了一种基于特征优化的缸壁积碳程度判别模型,达到了86.9%的准确

率,比传统模型提升了2.8%^[15],进一步验证了深度学习在积碳检测领域的可行性和有效性。

但汽车发动机结构复杂,且积碳产生和聚集受多种因素影响,单一部位的模型难以全面表征发动机整体积碳分布。因此,还需要融合其他部位的积碳程度以准确判别整体情况,从而选择合适的清理液。进气道和进气阀部位的采集图像角度基本固定,特征区域相对恒定统一,极少有重度噪声干扰的图像,处理相对简单;相比之下,喷油嘴部位图像由于采集远近距离交替,导致喷油嘴形态及周围污染程度不同;活塞顶部位积碳分布较为分散,外界条件干扰严重导致采集图像质量参差不齐。这些因素导致传统的深度卷积神经网络(Dynamic Convolution Neural Network, DCNN)模型在处理这些部位时,容易将注意力分散到许多无关特征上,而忽视了感兴趣区域中的积碳特征,从而影响识别精度。

注意力机制能够有效弥补DCNN模型的不足,通过对图像中不同区域加权,实现对重要区域的动态聚焦,从而提高目标检测和图像分类的精度与效率^[16]。基于这一思想,国内外研究者提出了多种变种机制,如通道注意力机制(SENNet^[17]、SKNet^[18]、ECANet^[19])、空间注意力机制(GENet^[20]、RANet^[21])、以及混合注意力机制(CBAM^[22]、DANet^[23]、CA^[24])等,增强了网络对图像重要特征的关注能力。其中,CA注意力机制通过调整通道特征的权重,使得模型能更有效地聚焦于图像中的关键区域,提升目标检测和图像分类的精度。同时,SG-Former^[25]引入自适应注意力机制,能够精准捕捉图像细粒度特征。

因此,本文针对喷油嘴部位积碳图像,提出了基于CA注意力机制的积碳识别模型来提升对喷油嘴部位积碳的聚焦关注性;针对活塞顶部位积碳图像,提出基于SG-Former的自适应积碳图像细粒度识别模型,旨在最大化下采样过程的平移不变性和对全局重要特征的细粒度信息的提取。

1 主要算法描述

1.1 基于 CA 注意力机制的喷油嘴部位积碳图像识别方法

1.1.1 RepVGG

图 1 为 RepVGG^[26] 的主要模型结构, 在训练阶段中, 最开始采用的是简单的一分支残差结构, 这种结构适合于初期网络设计, 能够快速学习积碳图像的基本特征, 如纹理特征。然而, 随着网络层数的增加, 特征图像的分辨率逐渐降低, 空间结构损失严重, 并且一部分特征在传递过程中逐渐消隐, 失去了具有判别性的特征。为了解决这一问题, RepVGG 在网络深度增加时引入了更复杂的多分支并行残差结构。这种结构通过并行学习不同类型的特征, 能更好地适应特征提取的变化, 增强网络的特征提取能力, 从而提高模型的整体性能和鲁棒性。

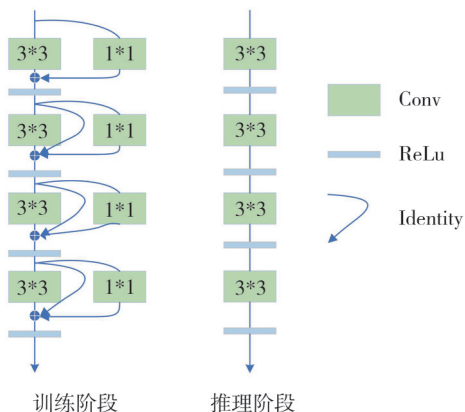


图 1 RepVGG 结构
Fig. 1 RepVGG structure

而在推理阶段, 我们将 1*1 卷积和恒等映射转换为 3*3 卷积, 并将其与另一个 3*3 卷积融合, 从而实现 RepVGG 结构的重参数化。这一过程简化了推理的网络模型, 加快模型部署。其过程如下:

卷积运算表示为 $y = kx + b$, 其中, k 为卷积权重参数, b 为偏置。BN 层包括均值 μ 、方差 σ 、尺度因子 γ 和偏置 β , BN 层的运算公式为

$$bn(x, \mu, \sigma, \gamma, \beta) = \frac{(x - \mu)\gamma}{\sigma} + \beta. \quad (1)$$

因此, Conv-BN 的结果为

$$y' = \frac{((kx + b) - \mu)\gamma}{\sqrt{\sigma^2 + \epsilon}} + \beta. \quad (2)$$

所以, 式(2)等效的卷积权重参数(k')和偏置

参数(b')为

$$\begin{cases} k' = \frac{\gamma k}{\sigma}, \\ b' = -\frac{\mu \gamma}{\sigma} + \beta. \end{cases} \quad (3)$$

因此, 结构重参数的最终结果为

$$y'' = bn(xk^{(3)}, \mu^{(3)}, \sigma^{(3)}, \gamma^{(3)}, \beta^{(3)}) + bn(xk^{(1)}, \mu^{(1)}, \sigma^{(1)}, \gamma^{(1)}, \beta^{(1)}) + bn(x, \mu^{(0)}, \sigma^{(0)}, \gamma^{(0)}, \beta^{(0)}) = k^{(3)'}x + b^{(3)'} + k^{(1)'}x + b^{(1)'} + x + b^{(0)'}. \quad (4)$$

1.1.2 CA 注意力机制

CA 注意力机制的优势在于其可对输入特征沿一个方向捕获通道之间的长期依赖关系, 并沿另一个方向捕获对象的位置信息。目标位置信息被保存到生成的注意力图中, 然后通过乘法运算添加到输入特征图中, 以进行下一次卷积运算, 这有助于模型更好地定位喷油嘴区域。

CA 注意力机制结构如图 2 所示, 具体操作如下: 对输入的特征图($Y \in \mathbb{R}^{C \times H \times W}$)在水平和垂直方向上执行两个独立的一维全局池化操作, 将 Y 分别聚合为两个具有特定方向和位置信息的特征映射(式(5), 式(6))。

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i), \quad (5)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w). \quad (6)$$

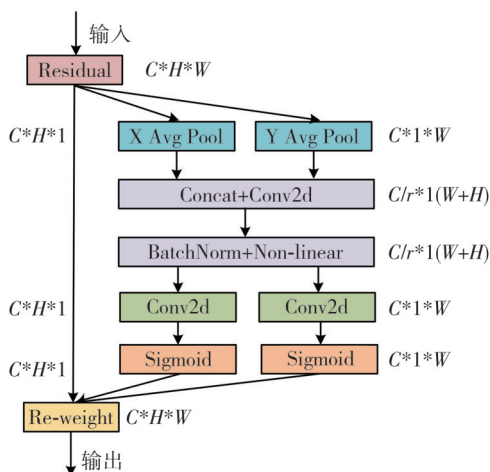


图 2 CA 注意力机制

Fig. 2 Coordinate attention

拼接得到的两个方向特征图, 并采用 1*1 卷积降低维度为原来的 C/r 得到 $1 \times (W \times H) \times C/r$ 的特征图 Z , 接着将特征图 Z 按照原来的高度和宽度进行卷积核为 1*1 的卷积分别得到通道数与原来一样的特征图, 通过 sigmoid 激活函数得到 g^h 和 g^w

(式(7), 式(8))。

$$g^h = \sigma(F_h(Z^h)), \quad (7)$$

$$g^w = \sigma(F_w(Z^w)). \quad (8)$$

最终, 将得到的权重和原始输入图像加权相乘, 得到加权的注意力特征图。

1.1.3 模型结构

喷油嘴部位模型的详细结构如表 1 所示, 以 RepVGG 为主体架构, 强化特征信息的提取, 增加模型推理速度, 并在第一个 stage 和最后一个 stage 添加 CA 注意力机制, 提高对喷油嘴部位积碳的关注。

1.2 基于 SG-former 的自适应活塞顶积碳图像识别方法

1.2.1 下采样模块

卷积在步进大于 2 进行下采样操作时, 可能会忽略采样定理, 导致信号发生混叠, 即输入发生微妙变化可能会引起输出的剧烈变化^[27-29]。在信号处理中, 通常在下采样前加入低通滤波来抗混叠, 但直接加入低通滤波会影响整个网络的性能。

在积碳识别过程中, 传统的卷积操作会导致某些积碳特征的丢失, 从而造成分类准确度下降。因此, 在下采样模块中利用 BlurPool 池化^[29]最大化下采样过程中的平移不变性, 最大程度保证积碳特征的提取。具体应用在卷积下采样层(以原

步长为 2 的卷积为例), 主要为以下 3 个步骤, 如图 3 所示。

表 1 喷油嘴积碳模型结构

Tab. 1 Injector nozzle carbon modelling structure

| 阶段 | 输出特征图大小 | 网络分支结构(训练阶段) |
|---------|---------|--|
| Stage 1 | 112*112 | $\begin{bmatrix} 3*3, 48 \\ \oplus [1*1, 48] \end{bmatrix} *1, \text{stride} = 2$ |
| 注意力机制 | 112*112 | 1*CA |
| Stage 2 | 56*56 | $\begin{bmatrix} [3*3, 48] \\ \oplus [1*1, 48] \\ \oplus \text{恒等映射} \end{bmatrix} *1, \text{stride} = 1$ |
| Stage 3 | 28*28 | $\begin{bmatrix} [3*3, 96] \\ \oplus [1*1, 96] \\ \oplus \text{恒等映射} \end{bmatrix} *1, \text{stride} = 2$ |
| Stage 4 | 14*14 | $\begin{bmatrix} [3*3, 192] \\ \oplus [1*1, 192] \\ \oplus \text{恒等映射} \end{bmatrix} *13, \text{stride} = 1$ |
| Stage 5 | 7*7 | $\begin{bmatrix} 3*3, 1280 \\ \oplus [1*1, 1280] \end{bmatrix} *1, \text{stride} = 2$ |
| 注意力机制 | 7*7 | 1*CA |
| | 1*1 | 自适应平均池化全连接层 |

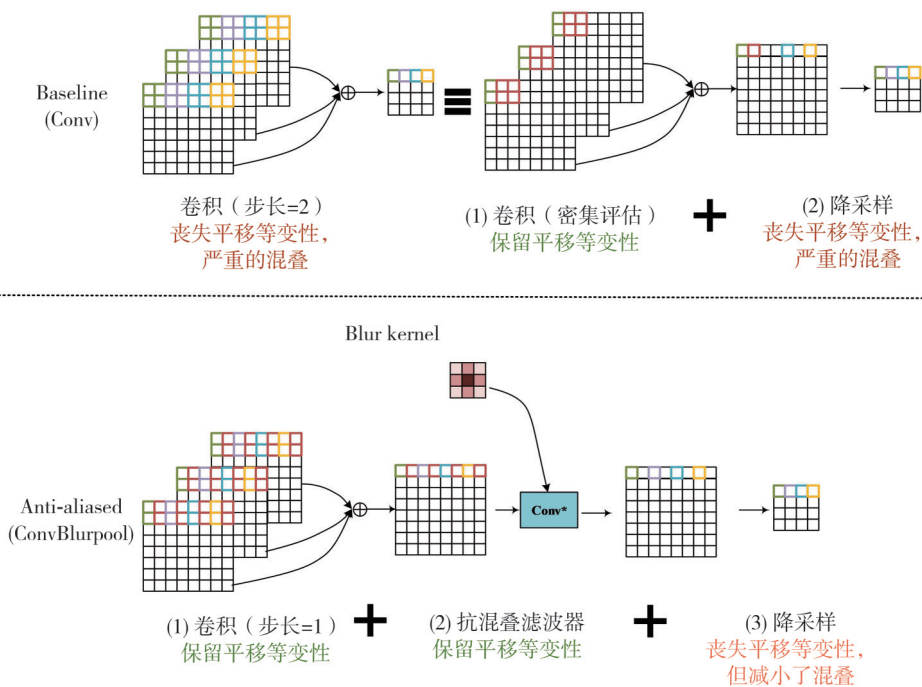


图 3 抗混叠的卷积下采样层

Fig. 3 Anti-aliasing conv downsampling layers

1) 将卷积步长设置为 1, 可以有效地保留输入信号的平移等变性。

2) 应用低通滤波器(模糊核)对相邻像素进行加权平均。这种处理方式有助于在特征图中更好地保留信息, 从而使得即便输入图像发生轻微的平移, 输出特征图也不会出现明显的变化。此外, 这种模糊处理能保留更多的上下文信息, 使得特征图对局部噪声和不精确性具有更强的鲁棒性, 避免了因局部特征的微小变化而引发的显著差异。经过模糊处理后, 再进行下采样, 这一过程最大限度地消除了混叠现象。

3) 最后, 对步骤 2) 的输出进行下采样, 输出结果。

1.2.2 自适应全局显著区域关注模块

如图 4 所示, 活塞顶部积碳的分布较为分散。传统的卷积神经网络在局部特征提取上表现优异, 但难以全面捕捉到一些重要的积碳空间特征。自注意力机制可在提取图像特征的同时关注整个图像的全局信息。然而, 这种机制的计算成本较高。为降低成本, 通常会在全局自注意力中缩短序列, 但这可能导致精细特征提取变得粗粒度。于是, 本文采用了自适应全局显著区域关注模块(SG-former)。该模块利用显著性图, 在训练中动

态调整 token 分配。具体而言, 显著区域获得更多 token, 实现细粒度关注, 而次要区域则分配较少 token, 从而提高效率的同时保持全局感受野。自适应全局显著模块融合策略如图 5 所示。

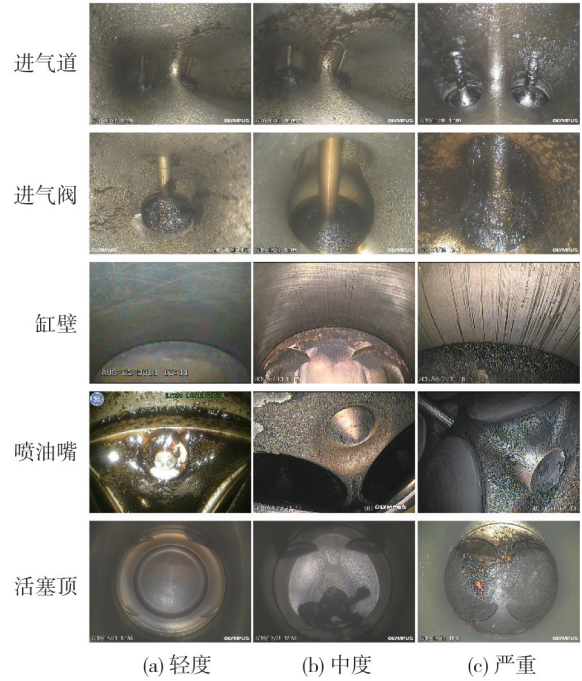


图 4 汽车发动机积碳数据集实例

Fig. 4 Example of a carbon deposit dataset for a vehicle engine

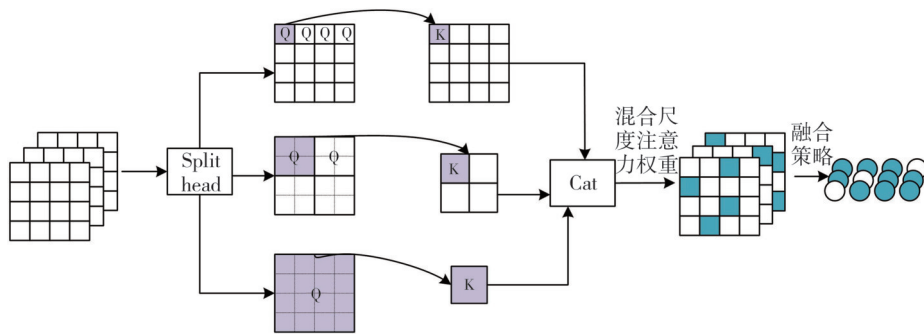


图 5 自适应全局显著信息关注模块融合策略

Fig. 5 Adaptive global salient information concern module fusion strategy

自适应全局显著模块融合策略具体过程如下:

将输入特征 X 映射到 Q, K, V 中, 然后将 N 个头分为 n 组, 对第 j 组的第 i 个头, 按尺度 s_j ($j=1, 2, 3, \dots, n$) 将 (K, V) 中的每个 s_j*s_j 的 token 合并成一个后, 将 (Q, K, V) 窗口分割成多个, 将每个 (K, V) 窗口设置为 M , 对应的 Q 设置 s_jM*s_jM 。

$$Q_i = XW_i^Q, \quad (9)$$

$$K_i = Merge(X, s_j)W_i^K, \quad (10)$$

$$V_i = Merge(X, s_j)W_i^V, \quad (11)$$

式中: $Merge(X, s_j)$ 表示将 X 中每个 s_j*s_j 的 token

通过步长为 s_j 的卷积合并成一个, $s_j=1$ 时, (Q, K, V) 具有相同大小。

$$Atten_i = Softmax\left(\frac{P(Q_i, s_jM)P(K_i^T, M)}{\sqrt{d_h}}\right), \quad (12)$$

$$n_i = Atten_i P(V_i, M), \quad (13)$$

式中: $P(Q_i, s_jM)$ 表示对查询窗口进行 s_jM*s_jM 分区, $s_jM*s_jM = h*w$ 时, 所有的 (K, V) 用来匹配 Q , 确保全局注意力的提取。

$$S_i = \frac{1}{hw} \sum_{m=1}^{hw} Atten_i^{m,n}, \quad (14)$$

$$S = \sum_{i=1}^n S_i, \quad (15)$$

式中： S 为最后输出的注意力去权重引导，包含局部细粒度和全局信息。

将上述输出的 S 按重要程度均匀分为 n 个子区域 s^1, \dots, s^n (n 越小越重要)，对应的输入 X 也划分为 X^1, \dots, X^n 。

$$Q = XW^Q, \quad (16)$$

$$K = IAM(X, S, r)W^K, \quad (17)$$

$$V = IAM(X, S, r)W^V, \quad (18)$$

式中： r 为对子区域的聚合力，子区域越重要， r 越小。 IAM 根据 r 聚合力重新分配输入 X^1, \dots, X^n 的 token。

最后通过聚合函数对划分的 X^1, \dots, X^n 进行重新聚合后重新拼接得到最终的提取特征。

$$\hat{X} = Cat(F(X^1, r_1), \dots, F(X^n, r_n)). \quad (19)$$

2 实验结果与分析

2.1 实验数据集

汽车发动机图像由汽车公司提供，主要包括5个部位：进气道、进气阀、缸壁、喷油嘴、活塞顶；每个部位分为3种程度：轻度、中度、严重，如图4所示。汽车各部位积碳不同程度数据分布如表2所示。

表2 数据集分布

Tab. 2 Distribution of datasets

| | 轻度/张 | 中度/张 | 严重/张 |
|-----|-------|-------|------|
| 活塞顶 | 1 322 | 2 827 | 478 |
| 喷油嘴 | 398 | 2 080 | 439 |
| 缸壁 | 1 450 | 1 450 | 550 |
| 进气道 | 834 | 411 | 871 |
| 进气阀 | 814 | 882 | 550 |

将每个部位积碳数据集按6:2:2划分训练集、验证集、测试集，并在训练时对所有图像做标准化处理，加快模型收敛。

2.2 环境设置

本实验采用Intel(R) Xeon(R) CPU E5-2690 V3@ 2.6 GHz处理器，GPU型号为Geforce RTX 2080 Ti，操作系统为Ubuntu 18.04.5 LTS，软件采用Pytorch深度学习框架，编程语言为Python。实验超参数设置如表3所示。

表3 超参数设置

Tab. 3 Hyperparameter settings

| 超参数 | 大小 |
|-----------|------------------|
| Epoch | 300 |
| Batchsize | 16 |
| 损失函数 | Crossentropyloss |
| 优化器 | Adam |

2.3 评价指标

本文使用的性能评估标准如下。其中在计算这些指标时所用到的参量包括： TP 为模型将正类正确分类为正类的样本数； FP 为模型将负类错误地分类为正类的样本数； FN 为模型将正类错误地分类为负类的样本数； TN 为模型将负类正确分类为负类的样本数。

1) 准确率(Accuracy)

$$Accuracy = \frac{\sum_{i=0}^{N-1} (f(x_i) = y_i)}{N}. \quad (20)$$

2) 精准率(Precision)

$$Precision = \frac{TP}{TP + FP}. \quad (21)$$

3) 召回率(Recall)

$$Recall = \frac{TP}{TP + FN}. \quad (22)$$

4) 特异度(Specificity)

$$Specificity = \frac{TN}{TN + FP}. \quad (23)$$

5) F1分数(F1-score)

$$F1\text{-score} = \frac{2 * Precision * Recall}{Precision + Recall}. \quad (24)$$

2.4 模型性能分析

2.4.1 基于RepVGG的部位识别性能分析

部位识别采用RepVGG-A0，多尺度提取各部位结构特征信息。RepVGG-A0对部位的识别率非常优越，准确率高达99.88%，表4还列出了RepVGG-A0对不同部位的识别精准率、召回率、特异度、F1-score，其中缸壁和活塞顶部位的识别精确度达100%，F1-score均高于99%，证实了其稳健的分类特性。

2.4.2 基于CA注意力机制的喷油嘴部位积碳图像识别性能分析

通过参考与积碳程度识别类似的图像识别任务，选取了性能最优的模型作为本文的主干框架。结果如表5所示，RepVGG-A0的效果最好，识别

率达到了 87.58%。

表 4 RepVGG-A0 部位识别结果

Tab. 4 RepVGG-A0 site identification results

| | Precision | Recall | Specificity | F1-score |
|-----|-----------|---------|-------------|----------|
| 进气道 | 0.998 0 | 1.000 0 | 1.000 0 | 1.000 0 |
| 缸壁 | 1.000 0 | 0.997 0 | 1.000 0 | 0.998 5 |
| 活塞顶 | 1.000 0 | 1.000 0 | 1.000 0 | 1.000 0 |
| 进气阀 | 0.995 0 | 0.998 0 | 0.999 0 | 0.996 5 |
| 喷油嘴 | 0.999 0 | 1.000 0 | 1.000 0 | 0.999 5 |

之后本文对比了在 RepVGG-A0 不同位置添加 CA 注意力机制对喷油嘴积碳程度识别模型性

能的影响(如表 6 所示)。结果显示,基于 CA 注意力机制的模型有效改善了多通道与空间之间的联系,减少了冗余信息的干扰。在不同位置添加 CA 注意力机制后, Precision、Recall、Specificity、F1-score、Accuracy 指标基本都有所提高,尤其在第 1 和第 5 个 stage 添加 CA 注意力机制, Precision、Recall、Specificity、F1-score、Accuracy 相对原始模型分别提升了 0.67、1.50、0.80、1.08、0.94 百分点,均获得了最高分数。

表 5 喷油嘴积碳图像分类模型对比实验

Tab. 5 Comparative experiments of the nozzle carbon models

| 模型 | Precision | Recall | Specificity | F1-score | Accuracy |
|------------------|----------------|----------------|----------------|----------------|----------------|
| VGG11 | 0.834 3 | 0.851 7 | 0.915 0 | 0.842 9 | 0.849 1 |
| Resnet50 | 0.846 3 | 0.856 7 | 0.919 0 | 0.851 5 | 0.854 4 |
| Mobilenet-V2 | 0.844 0 | 0.844 3 | 0.913 7 | 0.844 2 | 0.853 1 |
| Efficientnet | 0.831 3 | 0.867 0 | 0.921 7 | 0.848 8 | 0.850 5 |
| RepVGG-A0 | 0.872 3 | 0.867 7 | 0.925 3 | 0.870 0 | 0.875 8 |
| Convnext | 0.812 3 | 0.818 3 | 0.898 3 | 0.815 3 | 0.825 1 |
| Nfnnet | 0.851 3 | 0.835 7 | 0.911 0 | 0.843 4 | 0.853 1 |
| Ghostnet | 0.828 0 | 0.835 3 | 0.907 3 | 0.831 7 | 0.830 4 |
| ViT | 0.826 0 | 0.752 3 | 0.868 7 | 0.787 4 | 0.802 4 |
| Swin Transformer | 0.808 3 | 0.819 7 | 0.898 7 | 0.814 0 | 0.821 1 |

表 6 CA 注意力不同位置对比

Tab. 6 Comparison of different locations of coordinate attention

| 模型 | Precision | Recall | Specificity | F1-score | Accuracy |
|-------------|----------------|----------------|----------------|----------------|----------------|
| RepVGG-A0 | 0.872 3 | 0.867 7 | 0.925 3 | 0.870 0 | 0.875 8 |
| Stage1-CA | 0.868 3 | 0.883 3 | 0.932 7 | 0.875 8 | 0.879 8 |
| Stage1/2-CA | 0.873 7 | 0.873 3 | 0.928 3 | 0.874 0 | 0.878 5 |
| Stage1/3-CA | 0.876 7 | 0.868 3 | 0.926 0 | 0.872 5 | 0.878 5 |
| Stage1/4-CA | 0.870 0 | 0.882 0 | 0.932 7 | 0.876 0 | 0.881 2 |
| Stage1/5-CA | 0.879 0 | 0.882 7 | 0.933 3 | 0.880 8 | 0.885 2 |

2.4.3 基于 SG-Former 的自适应活塞顶积碳图像识别性能分析

同喷油嘴部位一样,本文对主干网络进行了对

比实验,如表 7 所示,综合实验结果表明 Resnet50 的效果最佳。为了验证所提方法的有效性,还对活塞顶数据集进行了消融实验。

表 7 活塞顶积碳图像分类模型对比实验

Tab. 7 Comparative experiments on piston top carbon models

| 模型 | Precision | Recall | Specificity | F1-score | Accuracy |
|------------------|----------------|----------------|----------------|----------------|----------------|
| VGG11 | 0.831 3 | 0.806 0 | 0.907 0 | 0.818 5 | 0.851 7 |
| Resnet50 | 0.838 0 | 0.806 7 | 0.901 7 | 0.822 0 | 0.855 0 |
| Mobilenet-V2 | 0.802 3 | 0.780 0 | 0.894 0 | 0.791 0 | 0.821 4 |
| Efficientnet | 0.793 7 | 0.821 3 | 0.904 3 | 0.807 3 | 0.842 0 |
| Repvvg-A0 | 0.816 0 | 0.841 7 | 0.916 3 | 0.828 6 | 0.848 5 |
| Convnext | 0.813 3 | 0.777 3 | 0.889 7 | 0.794 9 | 0.837 7 |
| Nfnnet | 0.798 7 | 0.874 3 | 0.922 3 | 0.834 8 | 0.851 7 |
| Ghostnet | 0.783 7 | 0.824 7 | 0.907 7 | 0.803 6 | 0.832 3 |
| ViT | 0.651 0 | 0.673 7 | 0.850 7 | 0.662 1 | 0.738 1 |
| Swin Transformer | 0.839 3 | 0.783 0 | 0.898 0 | 0.810 2 | 0.839 8 |

在消融实验中,分别引入了 Blur 池化和 SG-former 模块,如表 8 所示,可以看出 Precision、Recall、Specificity、F1-score、Accuracy 均有所上升,当将两个模块一起融入到模型中,模型的效果得到了最好的体现。

2.4.4 对比传统模型的进气阀、进气道积碳图像识别性能分析

由于进气阀和和进气道积碳特征相对稳定,因此本文利用传统卷积神经网络对其进行训练,测试结果如表 9 所示。

表8 活塞顶积碳识别消融实验
Tab. 8 Piston top carbon identification and ablation experiment

| 模型 | Precision | Recall | Specificity | F1-score | Accuracy |
|--------------------|----------------|----------------|----------------|----------------|----------------|
| Resnet50 | 0.838 0 | 0.806 7 | 0.901 7 | 0.822 0 | 0.855 0 |
| Resnet50+BlurPool | 0.841 7 | 0.847 0 | 0.920 0 | 0.844 3 | 0.866 9 |
| Resnet50+SG-Former | 0.848 3 | 0.821 3 | 0.915 7 | 0.834 6 | 0.870 1 |
| 所提方法 | 0.868 0 | 0.852 7 | 0.921 0 | 0.860 3 | 0.882 0 |

表9 进气阀、进气道积碳程度识别不同模型对比

Tab. 9 Comparison of different models for identifying the carbon deposition degree of intake valves and intake tracts

| 部位 | 模型 | Precision | Recall | Specificity | F1-score | Accuracy |
|-----|--------------|----------------|----------------|----------------|----------------|----------------|
| 进气阀 | Vgg11 | 0.960 7 | 0.946 7 | 0.974 7 | 0.953 6 | 0.953 0 |
| | Vgg13 | 0.972 0 | 0.959 3 | 0.981 7 | 0.965 6 | 0.966 4 |
| | Vgg16 | 0.956 7 | 0.937 3 | 0.970 7 | 0.946 9 | 0.946 3 |
| | Vgg19 | 0.967 7 | 0.962 0 | 0.980 7 | 0.964 8 | 0.964 2 |
| | Resnet18 | 0.975 0 | 0.968 7 | 0.984 | 0.971 8 | 0.970 9 |
| | Resnet34 | 0.979 0 | 0.979 3 | 0.989 7 | 0.979 1 | 0.979 9 |
| | Resnet50 | 0.987 0 | 0.982 3 | 0.991 3 | 0.984 7 | 0.984 3 |
| | Mobilenet-V2 | 0.967 3 | 0.956 7 | 0.979 7 | 0.962 0 | 0.961 9 |
| | Mobilenet-V3 | 0.984 3 | 0.980 3 | 0.990 3 | 0.982 3 | 0.982 1 |
| | Efficientnet | 0.982 0 | 0.980 0 | 0.989 0 | 0.981 0 | 0.979 9 |
| 进气道 | Vgg11 | 0.931 3 | 0.930 0 | 0.973 7 | 0.930 7 | 0.945 5 |
| | Vgg13 | 0.930 0 | 0.942 3 | 0.976 3 | 0.936 1 | 0.947 9 |
| | Vgg16 | 0.939 3 | 0.950 3 | 0.979 3 | 0.944 8 | 0.955 0 |
| | Vgg19 | 0.919 3 | 0.909 7 | 0.967 0 | 0.914 5 | 0.933 6 |
| | Resnet18 | 0.942 3 | 0.935 7 | 0.976 3 | 0.939 0 | 0.952 6 |
| | Resnet34 | 0.938 3 | 0.952 0 | 0.979 7 | 0.945 1 | 0.955 0 |
| | Resnet50 | 0.893 3 | 0.904 3 | 0.961 3 | 0.898 8 | 0.917 1 |
| | Mobilenet-V2 | 0.898 3 | 0.886 0 | 0.952 7 | 0.892 1 | 0.907 6 |
| | Mobilenet-V3 | 0.927 3 | 0.940 3 | 0.975 3 | 0.933 8 | 0.945 5 |
| | Efficientnet | 0.907 3 | 0.906 0 | 0.964 3 | 0.906 7 | 0.926 5 |

通过对比,可以看出对于进气阀部位的积碳程度识别, Resnet50模型在各个方面都占有绝对优势,分类准确率高达98.43%, F1-score也达到了98.47%;对于进气道部位的积碳程度识别, Resnet34则表现更好,虽然精确度相较Resnet18降低0.4百分比,但其余评价指标都相对较高,分类准确率也达到了95.50%, F1-score也达到94.51%。

3 结论

本文提出了一种针对积碳图像细粒度特征的分类方法,旨在实现对汽车发动机总体积碳程度的判别。针对背景信息干扰严重且形态差异大的喷油嘴部位积碳,提出了一种基于CA注意力机制的RepVGG网络提升对喷油嘴部位积碳的聚焦性,并分析了不同位置的CA注意力机制对模型的影响;针对积碳分布较散的活塞顶部位积碳提出了一种基于SG-Former的自适应细粒度分类网络,同时采用Blurpool池化来最大化下采样过程的平移不变性。此外,还利用传统模型对积碳特征相对稳定的进气道和进气阀部位进行了最优模型选择。最终对多输入图像的部位识别率达到了99.88%,其中喷油嘴部位、活塞顶、进气道、进气阀的积碳程度识别也分别达到了88.52%、88.20%、95.50%、98.43%。另外

在本课题组之前的工作中,针对缸壁部位提出的积碳程度识别方法的准确率已达到86.9%^[15]。基于以上方法,有望在实际应用中实现对汽车发动机总体积碳程度判别。

参考文献:

- [1] ROE A P, THOMPSON S A, PETROU A T, et al. System for determining piston damage based on carbon deposit growth: US20180030890[P]. 2018-02-01.
- [2] WANG J Q, DING W Q, DENG J, et al. Method for detecting and distinguishing early combustion and carbon deposits of engine based on ion current: NL2029619[P]. 2023-06-02.
- [3] 钱飞,王婷,钱宸浩,等. 基于内燃机的爆震声音判定气缸内积碳分布的测定装置: CN116698272A [P]. 2023-09-05.
- [4] 卢曼斯,司徒白雪,赵炎. 燃料发动机内部积碳厚度检测装置: CN207881628U[P]. 2018-09-18.
- [5] 尚忱. 应用物联网传感器监测车辆积炭的装置: CN108240840A[P]. 2018-07-03.
- [6] 肖剑平,季明,兰龙,等. 发动机检测方法、系统、计算机设备和存储介质: CN113284131A[P]. 2021-08-20.
- [7] ZHANG Q, SHU J, CHEN C, et al. Optimization of pneumonia CT classification model using RepVGG and spatial attention features [J]. Frontiers in Medi-

- cine, 2023, 10: 1233724.
- [8] 李昊璇, 苏艳琼. 基于改进 YOLOv8 的道路缺陷检测[J]. 测试技术学报, 2024, 38(5): 506-512.
LI Haoxuan, SU Yanqiong. Road defect detection based on improved YOLOv8 [J]. Journal of Test and Measurement Technology, 2024, 38(5): 506-512. (in Chinese)
- [9] HUANG H, YOU Z, CAI H, et al. Fast detection method for prostate cancer cells based on an integrated ResNet50 and YOLOv5 framework [J]. Computer Methods and Programs in Biomedicine, 2022, 226: 107184.
- [10] ATILA Ü, UÇAR M, AKYOL K, et al. Plant leaf disease classification using EfficientNet deep learning model[J]. Ecological Informatics, 2021, 61: 101182.
- [11] 左昊轩, 黄祺成, 杨佳昊, 等. 基于改进 YOLOv5s 的作物黄化曲叶病检测方法[J]. 农业机械学报, 2023, 54(S1): 230-238.
ZUO Haoxuan, HUANG Qicheng, YANG Jiahao, et al. Improved YOLOv5s-based detection method for crop yellow leaf curl virus disease[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(S1): 230-238. (in Chinese)
- [12] CHEN L, LIN S, LU X, et al. Deep neural network based vehicle and pedestrian detection for autonomous driving: a survey[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(6): 3234-3246.
- [13] LIU Y, ZHANG Z, LIU X, et al. Deep learning-based image classification for online multi-coal and multi-class sorting [J]. Computers & Geosciences, 2021, 157: 104922.
- [14] LI E, ZENG L, ZHOU Z, et al. Edge AI: on-demand accelerating deep neural network inference via edge computing [J]. IEEE Transactions on Wireless Communications, 2019, 19(1): 447-457.
- [15] HUANG Q, CHEN Y, WANG Z, et al. Research on discrimination method of carbon deposit degree of automobile engine based on deep learning [J]. Computing and Informatics, 2024, 43(1): 126-148.
- [16] ZHU X, CHENG D, ZHANG Z, et al. An empirical study of spatial attention mechanisms in deep networks [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019: 6687-6696.
- [17] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [18] LI X, WANG W, HU X, et al. Selective kernel networks [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 510-519.
- [19] MISHRA S K, RAI G, SAHA S, et al. Efficient channel attention based encoder-decoder approach for image captioning in Hindi [J]. ACM Transactions on Asian and Low-Resource Language Information Processing, 2021, 21(3): 49.
- [20] CAO Y, XU J, LIN S, et al. GCNet: non-local networks meet squeeze-excitation networks and beyond[C]//2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), 2019: 1971-1980.
- [21] WANG Z, XU J, LIU L, et al. RANet: ranking attention network for fast video object segmentation [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019: 3977-3986.
- [22] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[M]//Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 3-19.
- [23] ZEGARRA F C, VARGAS-MACHUCA J, CORONADO A M. Comparison of CNN and CNN-LSTM architectures for tool wear estimation [C]//2021 IEEE Engineering International Research Conference (EIRCON), 2021: 1-4.
- [24] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 13708-13717.
- [25] REN S, YANG X, LIU S, et al. SG-former: self-guided transformer with evolving token reallocation [C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV), 2023: 5980-5991.
- [26] DING X, ZHANG X, MA N, et al. RepVGG: making VGG-style ConvNets great again [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 13728-13737.
- [27] ENGSTROM L, TRAN B, TSIPRAS D, et al. A rotation and a translation suffice: fooling CNNs with simple transformations [C]//Proceedings of the International Conference on Learning Representations (ICLR), 2019: 1-21.
- [28] AZULAY A, WEISS Y. Why do deep convolutional networks generalize so poorly to small image transformations [J]. Journal of Machine Learning Research, 2019, 20(184): 1-25.
- [29] ZHANG R. Making convolutional networks shift-invariant again [C]//International Conference on Machine Learning, PMLR, 2019: 7324-7334.