

文章编号: 1671-7449(2024)04-0441-07

## 基于非刚性特征的发动机积碳程度判别模型

黄倩, 王召巴, 陈香, 张航佳, 陈友兴\*

(中北大学信息与通信工程学院, 山西太原 030051)

**摘要:** 汽车发动机积碳的长期累积会引起发动机动力下降、油耗上升、排放性能降低等问题, 对发动机及时的检测和清理可以有效缓解积碳造成的影响。本文提出了一种新的基于非刚性特征的模型, 用于积碳程度的判别。首先, 在模型中使用可变形卷积调整卷积核的偏移位置和偏移幅度, 提高网络的有效感受野, 提取非刚性特征信息; 其次, 根据核区域中像素的相关性, 使用自适应指数度量池化核神经元进行加权, 以此捕获更好的细节特征信息; 最后, 添加基于自注意机制的特征增强模块, 获取特征图的上下文信息。经过实验测试, 本文方法的测试准确率为87.1%, 各项指标相较原始模型平均提高2.5%, 证明本文方法可以提取有效的积碳程度判别特征。

**关键词:** 积碳程度识别; 非刚性特征; 可变形卷积; 下采样; 自注意机制

**中图分类号:** TP391

**文献标识码:** A

**doi:** 10.3969/j.issn.1671-7449.2024057

**引用格式:** 黄倩, 王召巴, 陈香, 等. 基于非刚性特征的发动机积碳程度判别模型[J]. 测试技术学报, 2024, 38(4): 441-447.

HUANG Qian, WANG Zhaoba, CHEN Xiang, et al. The discriminative model of carbon deposit degree of engine based on non-rigid features[J]. Journal of Test and Measurement Technology, 2024, 38(4): 441-447.

## The Discriminative Model of Carbon Deposit Degree of Engine Based on Non-Rigid Features

HUANG Qian, WANG Zhaoba, CHEN Xiang, ZHANG Hangjia, CHEN Youxing\*

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

**Abstract:** Long-term accumulation of carbon deposit in car engines can result in decreased engine power, increased fuel consumption, and diminished emission performance, highlighting the critical importance of timely engine detection and cleaning for effectively mitigating the impact of carbon deposit. A novel model based on non-rigid features has been proposed to discriminate the degree of carbon deposit. Firstly, deformable convolution is employed in the model to adjust the offset position and amplitude of the convolution kernel, enhancing the effective receptive field of the network and extracting non-rigid feature information. Subsequently, neurons are weighted based on the correlation of pixels within the kernel region using an adaptive exponential metric pooling kernel to capture more precise feature information. Finally, a feature improvement module based on a self-attention mechanism is incorporated to extract comprehensive contextual information from feature maps. The model's test accuracy after experimental testing is 87.1%, and the indexes have increased by 2.5% on average compared to the original model. Demonstrat-

**收稿日期:** 2024-01-02

**基金项目:** 山西省回国留学人员科研资助项目(2022-145); 中北大学研究生科技立项资助课题(2022180506)

**作者简介:** 黄倩(2000-), 女, 硕士生, 主要从事计算机视觉和图像处理研究。E-mail: 15891473422@163.com。

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

ing the capability of our proposed method to extract effective features for carbon deposit degree discrimination. The approach has the potential to theoretically justify the widespread promotion of the degree model discrimination for carbon deposit.

**Key words:** discriminate the degree of carbon deposit; non-rigid features; deformable convolution; down-sample; self-attention mechanism

## 0 引言

汽车发动机积碳是燃料、润滑油在高温高压作用下不完全燃烧所产生的氧化物,积碳脱落会影响机油润滑活塞环和缸壁的摩擦,导致发动机拉缸和燃烧汽油,若不及时检测清理则会导致发动机报废。因此,研究一种针对汽车发动机积碳程度检测的方法十分重要。

针对汽车发动机积碳程度检测的方法可以分为基于物理量的检测和基于视觉的检测。前者首先利用传感器设备获取发动机的状态信息量,再结合阈值判断是否存在积碳<sup>[1-3]</sup>。该方法需要综合多种信息且检测结果不够直观,在实际应用中存在很大的局限性;后者主要通过内窥镜成像进行判断,如图1所示为发动机燃烧室内部图像。该方法相较基于物理量的检测方法更直观、易推广<sup>[4]</sup>。但在图像程度分类时却存在如下问题:①由于判别的主要依据划痕的形态和大小不固定,所以不存在刚性的纹理特征和显著的轮廓特征;②由于图像采集时成像的色调并不一致,因此并不具备颜色特征;③划痕不具备固定的角点,所以不存在明显的结构特征。因此基于传统的手工设计分类器无法提取有效的程度判别特征<sup>[5]</sup>。

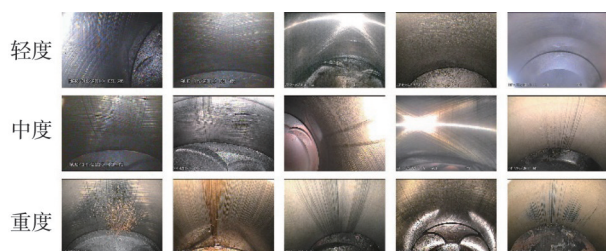


图1 发动机燃烧室内部图像样例

Fig. 1 Images of the inside of the engine's combustion chamber

基于深度学习的方法,通过更深层的网络结构可以学到隐藏在图像深层的特征,增加分类特点,提高图像分类准确度。但标准卷积采样位置受限于传统的矩形形状,无法适应非刚性的积碳程度判别特征。因此,引入可变形卷积(Deformable

Convolution, DConv),推广了尺度、各向异性的纵横比和旋转等各种变换,通过调制采样点的偏移量达到适应物体特征的效果<sup>[6]</sup>。池化可以帮助模型捕获卷积所缺少的全局信息,从而应对积碳图像中特征分布分散的问题。但最大池化、平均池化等常规的池化方式在应对纹理性较强的图像时会产生纹理断层,导致特征信息的丢失。因此,在进行池化操作时应考虑当前核区域内所有的像素贡献。自适应指数池化核融合单个核向量与其均值的通道相似性及softmax加权,通过捕捉最重要的信息和保留结构,最大限度减少信息的损失<sup>[7]</sup>。通过注意力机制对输入特征进行加权以强化特征图中的上下文信息,但基于挤压激励的注意力机制都只关注同一空间不同通道的信息,或者只关注了不同空间同一通道的信息,缺乏特征之间的内在联系<sup>[8-10]</sup>。双端注意力机制则通过使用自注意力关注相邻像素和相邻通道间的相似度,从两个维度增强特征间的依赖关系,从而增强特征表达<sup>[11]</sup>。

本文针对发动机图像中积碳程度判别特征难以获取的问题提出了一种新的深度神经网络架构,通过提高有效感受野的区域大小、精准加权图像的细节特征和加强特征像素邻域间的依赖关系等方法,有效改善积碳程度分类准确率低下的现状,为汽车发动机积碳程度的自动检测提供技术支持。

## 1 模型方法

### 1.1 可变形卷积

关于有效感受野<sup>[12]</sup>和显著区域<sup>[13-15]</sup>的研究表明,理论感受野中只有一小部分像素对最终的网络预测有显著影响,有效感受野由网络权值和采样位置共同控制。为了提高卷积模块的区域适应性,增强模型泛化能力,引入DConv提取积碳程度判别特征。 $\mathfrak{R}=\{(-1,-1),(-1,0),\dots,(0,1),(1,1)\}$ 为一个标准的 $3\times 3$ 卷积, $\mathfrak{R}$ 中的元素是相对于中心位置的偏移量。对于标准卷积采样位置

的输出可以表示为

$$y(p_0) = \sum_{p_n \in \mathcal{R}} w(p_n) \cdot x(p_0 + p_n), \quad (1)$$

式中:  $p_0$  是输出特征图  $y$  上的位置映射;  $p_n$  为  $\mathcal{R}$  中的元素位置;  $w$  为权值。可变形卷积相较普通卷积增加了一个方向参数  $\{\Delta p_n | n = 1, \dots, N\}$  和幅度调制参数  $\{\Delta m_n | n = 1, \dots, N\}$ ,  $N = |\mathcal{R}|$ , 则采样位置的输出可以表示为

$$y(p_0) = \sum_{p_n \in \mathcal{R}} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \cdot \Delta m_n. \quad (2)$$

如图 2 所示, 偏移域  $\Delta p_n$  和调制标量  $\Delta m_n$  通过平行的卷积操作获取, 输出偏移的尺寸与输入特征图的大小一致, 维度为  $3N$ , 前  $2N$  个通道表示  $x$  和  $y$  两个方向上的偏移量, 剩下的  $N$  个通道通过 sigmoid 层形成调制标量。融合后的偏移量不仅可以在感知输入特征时调整偏移方向, 还可以调制来自不同空间位置的输入特征幅度, 从而更好捕获非刚性特征, 同时避免局部曝光过度导致梯度过大对预测造成影响。

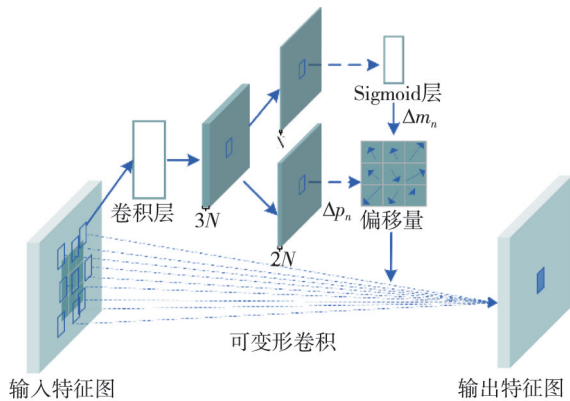


图 2 可变形卷积示意图

Fig. 2 Schematic of deformable convolution

## 1.2 指数自适应池化

为了更好地保留原始图像的细节信息, 采用更加全面的自适应指数加权池化, 包含两种池化核。

第 1 种是使用核区域  $P$  中单个核向量  $a_i$  与其均值  $\bar{a}$  的指数相似性来确定核向量权值, 称为 eDSCPool, 计算公式为式(3)。相似度的计算没有使用余弦相似度或基于距离的相似度方法, 而是根据 Dice-Sørensen 系数(DSC)计算, 如式(4)。

$$\tilde{a}_{eDSC} = \sum_{i \in P} \frac{e^{DSC(\bar{a}, a_i)}}{\sum_{j \in P} e^{DSC(\bar{a}, a_j)}} \cdot a_i, \quad (3)$$

$$DSC(\bar{a}, a_i) = \frac{2 \cdot \|\bar{a} \cdot a_i\|}{\|\bar{a}\|^2 + \|a_i\|^2}. \quad (4)$$

这种考虑矢量长度的方法有效克服了距离异常产生的像素伪影, 同时指数  $e$  避免了距离为 0 的情况, 使池化在反向传播阶段可微, 降低了梯度消失的风险。

第 2 种是基于 soft max 加权来放大更大强度的特征激活<sup>[16]</sup>, 称为 eMPool, 计算公式为

$$\tilde{a}_{eM} = \sum_{i \in P} \frac{e^{a_i}}{\sum_{j \in P} e^{a_j}} \cdot a_i. \quad (5)$$

soft max 加权通过概率加权使较大的激活对最终输出产生更大的影响, 同时也确保将最小的权重值分配给最低的激活。与 Lee 等<sup>[17]</sup>引入的平均和最大池化融合策略一致, 将式(3)和式(5)通过可训练的权重掩模  $\beta$  来创建光滑近似平均和光滑近似最大值的组合, 称为 eAPool, 如下式:

$$\tilde{a}_{eA} = \tilde{a}_{eDSC} \cdot \beta + \tilde{a}_{eM} \cdot (1 - \beta). \quad (6)$$

## 1.3 双端注意力机制

为了增强积碳程度判别特征间的长期依赖关系, 有效利用特征图的上下文关系, 利用自注意力从空间和通道两个维度建立相邻空间和相邻通道之间的依赖关系, 从而自适应地整合局部特征及其全局依赖关系。如图 3 所示, 分别获取图像的位置注意图和通道注意图增强特征表达。输入特征图  $O \in \mathbb{R}^{C \times H \times w}$ , 通过 3 个卷积层生成特征映射并进行形状重塑为特征图  $\{A, B, C\} \in \mathbb{R}^{C \times N}$ , 其中  $N = H \times W$  为像素数, 对特征映射  $A$  的转置与特征映射  $B$  做矩阵乘法, 通过 soft max 层来计算位置注意图  $P \in \mathbb{R}^{N \times N}$ 。其中  $P_{ji}$  测量第  $i$  个位置对第  $j$  个位置的影响, 当两个位置的特征表示越相似, 它们之间的相关性就越大。

$$P_{ji} = \frac{e^{(A_i \cdot B_j)}}{\sum_{i=1}^N e^{(A_i \cdot B_j)}}. \quad (7)$$

之后, 将注意图与特征映射  $C$  做矩阵乘法, 将输出结果重塑为特征图  $X \in \mathbb{R}^{C \times H \times w}$ ,  $X_i$  的计算执行像素级运算, 其结果表达式为

$$X_i = \sum_{i=1}^N (P_{ji} C_i) + O_{i_0}. \quad (8)$$

通道注意图的获取直接对原始特征图进行重塑, 以保持不同通道间的依赖关系不变。为了方便表达, 将重塑后的特征映射记为  $\{D, E, F\} \in \mathbb{R}^{C \times N}$ , 对特征映射  $E$  和  $D$  的转置做矩阵乘法, 通过 soft max 层获取最大相关的通道注意图  $Q \in \mathbb{R}^{C \times C}$ ,  $Q_{ji}$  表示第  $i$  个通道

对第 $j$ 个通道的影响。

通过通道注意力获得的特征图  $Y \in \mathbb{R}^{C \times H \times W}$  表示为

$$Q_{ji} = \frac{e^{(E_i \cdot D_j)}}{\sum_{i=1}^N e^{(E_i \cdot D_j)}}, \quad (9)$$

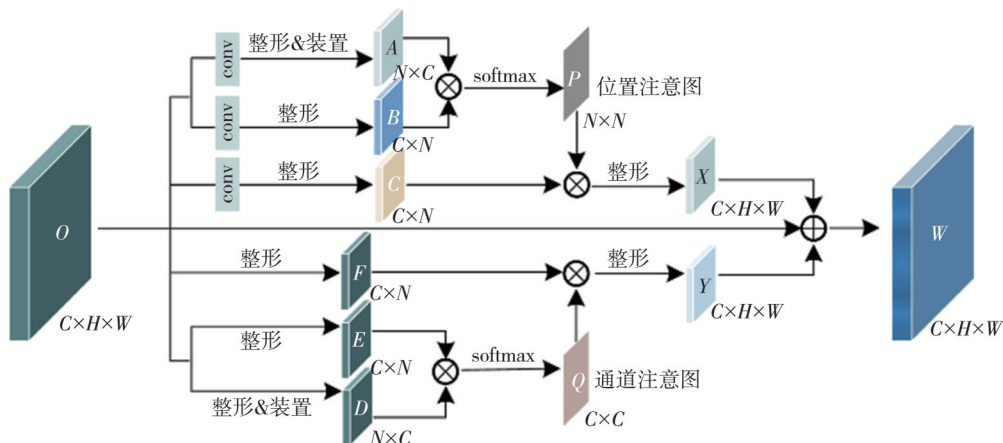


图3 双端注意力机制

Fig. 3 Dual attention mechanism

## 1.4 模型结构

积碳程度判别模型的详细网络结构如表1所示。

表1 积碳程度判别模型的详细网络结构

Tab. 1 Detailed network structure of the carbon deposit degree discrimination model

层名称	输出特征图大小	结构
Conv	112×112	7×7, stride2, 64 3×3 eAPool, stride2, 64
Block1	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$
Block2	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$
Block3	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$
Block4	7×7	$\begin{bmatrix} \text{DConv}3 \times 3, 512 \\ \text{DConv}3 \times 3, 512 \end{bmatrix} \times 2$
Attention module	7×7	$\begin{bmatrix} \text{SAM} \\ \text{CAM} \end{bmatrix} \times 1$
	1×1	Average pool Dropout Fully connected Softmax

骨干网络采用ResNet-18<sup>[18]</sup>, 本文针对积碳程度判别特征对骨干网络进行改进, 采用自适应指数池化捕获全局特征, 将Block4中的卷积替换为可变形

$$Y_i = \sum_{i=1}^N (Q_{ji} F_i) + O_i. \quad (10)$$

每个通道的最终特征是所有通道的特征与原始特征的加权和, 它模拟了特征映射之间的长程语义依赖关系, 有助于提高特征的可辨别性。

卷积, 更好地适应非刚性特征。增加注意力机制模块获取更有价值的特征信息。添加dropout层, 保证网络在学习到有效特征元素的同时避免了过拟合的产生。结构一栏中给出了网络设计的详细参数, 包括卷积核的大小, 步径长度和通道数。

## 2 实验

### 2.1 数据集介绍及预处理

表2中给出了数据分布情况, 原数据集存在类别分布不均及数据量偏少的问题, 因此针对原始数据集做两步处理, 首先, 对数据量做均衡化处理, 避免网络偏向学习。以重度数据量为基准抽取定量数据, 抽取原则上避免相似度较高且类别模糊的图像。对抽样数据按8:2的比例划分训练集和测试集; 其次, 对训练集数据进行扩充, 以此提高模型泛化性能和鲁棒性。

表2 数据集分布

Tab. 2 Distribution of datasets

标签	类别	原始数据		训练集/张 (扩充20倍)	测试集/张
		数量/张	占比/%		
0	轻度	1 450	42%	8 800	900
1	中度	1 450	42%	8 800	900
2	重度	550	16%	8 800	110
	总计	3 450	100%	26 400	1 910

本文采用翻转、裁剪、仿射、模糊、锐化、加噪、dropout、色彩抖动等多种数据增强方式,对每张图片同时进行多次变化,以达到提高差异性的目的,图 4 为单张样本扩充示例,首张为原图像,其余为扩充图像。将剩余数据补充为测试集,模拟实际应用测试。

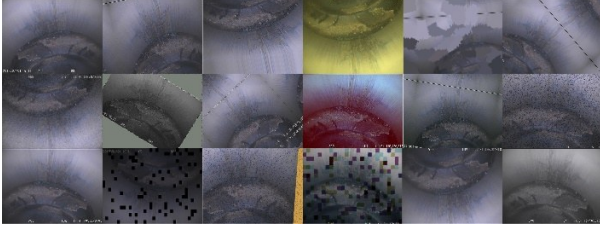


图 4 样本扩充示例图

Fig. 4 Sample Expansion Examples

## 2.2 实验参数设置

本文实验在 PyCharm 集成编译环境下进行, GPU 为 NVIDIA GeForce GTX 1080 Ti\*3, 实验采用并行训练的方式, 批处理大小为 64, 训练轮次为 100, 优化器采用 SGD 优化器, 初始学习率为 0.001, 学习率调整采用余弦衰减策略衰减至 0.000 1, 损失函数采用交叉熵损失函数。为了避免过拟合的发生, 在训练时对图像添加 Cutmix<sup>[19]</sup>, 即对任意两张图片进行裁剪混合,  $x \in \mathbb{R}^{W \times H \times C}$  代表输入图片,  $y$  为图片标签, 将任意两张输入图片  $(x_A, y_A)$ ,  $(x_B, y_B)$  混合生成新的训练样本  $(\tilde{x}, \tilde{y})$ , 混合操作定义为

$$\tilde{x} = M \odot x_A + (1 - M) \odot x_B, \quad (11)$$

$$\tilde{y} = \lambda y_A \oplus (1 - \lambda) y_B, \quad (12)$$

式中:  $M \in \{0, 1\}^{W \times H}$  为二进制掩码, 用来抽取和填充图像,  $\oplus$  是两个标签的融合,  $\lambda$  符合 beta 分布  $\lambda \sim \text{Beta}(\alpha, \alpha)$ , 在实验中将  $\alpha$  设为 1, 即  $\lambda$  服从  $(0, 1)$  的均匀分布, 将每个处理批次的混合概率设为 0.1。

## 2.3 评价指标

实验使用 Accuracy, Precision, Recall, F1-score 作为评价指标。其中 F1-score 利用精确率和召回率综合评价模型的分类性能。评价指标的计算表达式为

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}, \quad (13)$$

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (14)$$

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (15)$$

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (16)$$

由于测试集数量不均衡, 因此在计算 F1-score 时对每个类别添加加权系数, 加权系数由数据量分配占比决定, 则每个指标的得分计算如下式

$$S = \frac{N_{c_0}}{N_{\text{total}}} S_0 + \frac{N_{c_1}}{N_{\text{total}}} S_1 + \frac{N_{c_2}}{N_{\text{total}}} S_2, \quad (17)$$

式中:  $N_{c_i} (i=0, 1, 2)$  为类别数;  $N_{\text{total}}$  为测试集总量,  $S_i (i=0, 1, 2)$  为对应类别 F1-score 分值。

## 2.4 实验结果分析

### 2.4.1 模型对比实验

在骨干网络的选取中对比了几种与积碳程度判别较为相似的任务中应用较为广泛的几种骨干网络模型<sup>[20]</sup>。包括 VGGNet16, DenseNet169 和 EfficientNet-B4。从表 3 的实验结果可以看到, VGGNet16 和 ResNet18 简单的堆叠结构在积碳程度判别任务中有着较好的性能表现, 其中 ResNet18 因为捷径连接操作在各项指标中的表现最佳。但 DenseNet169 的密集连接网络并没有得到更好的结果, 相反, 由于图像噪声因素的干扰, 造成网络过拟合, 识别效果反而不佳。EfficientNet-B4 虽然有着较为高效的模型架构, 但由于移动翻转瓶颈卷积块中采用  $1 \times 1$  的点卷积, 降低了局部特征间的关联性, 导致模型识别准确率略显不足。

表 3 模型对比实验结果

Tab. 3 Results of model comparison experiments

Model	Accuracy/%	Precision/%	Recal/%	F1-score/%
VGGNet16	83.6	85.1	83.5	84.6
DenseNet169	79.9	83.0	80.0	80.7
EfficientNet-B4	80.4	83.7	80.4	82.0
ResNet18	84.6	85.2	84.6	84.9

### 2.4.2 可变形卷积模块的性能表现

在网络结构中探索了可变形卷积模块的版本以及添加的位置对分类准确率的影响, 可变形卷积模块的版本包括只调整卷积方向偏移的 DCNv1<sup>[21]</sup>, 同时调整方向偏移和调制幅度的 DCNv2, 包含权值共享机制的 DCNv3<sup>[22]</sup>。位置对

比添加在不同的Block中,表4和表5给出了可变形卷积模块和位置的对比结果。

表4 可变形卷积模块的对比结果

Tab. 4 Comparison results for deformable convolutional modules

Structure	Accuracy/%	Precision/%	Recall/%	F1-score/%
model-w/-DCNv1	84.7	85.7	84.7	85.3
model-w/-DCNv2	85.0	86.0	85.0	85.5
model-w/-DCNv3	84.4	85.3	84.4	84.9

表5 可变形卷积模块的位置对比结果

Tab. 5 Positional comparison results for deformable convolutional modules

Structure	Accuracy/%	Precision/%	Recall/%	F1-score/%
Block1-w/-DCNv2	85.0	86.0	85.0	85.5
Block2-w/-DCNv2	84.9	85.8	84.9	85.3
Block3-w/-DCNv2	85.4	86.1	85.4	85.7
Block4-w/-DCNv2	85.8	86.5	85.8	86.1

表4中的可变形卷积均添加在Block1中,结果显示,DCNv1和DCNv2模块对积碳程度分类模型的性能均有不同程度的提升,其中DCNv2考虑了卷积核的调制幅度筛选到更有价值的信息,DCNv3的权重共享机制模糊了图像的细节特征导致准确率降低。根据表4的实验结果,分析了DCNv2在不同特征层次中的性能表现,实验结果表明,可变形卷积在更深层次的特征提取中能发挥更大效能,原因在于底层的像素分布较为密集,可变形卷积核的大小限制导致特征提取差异不够明显,而高层特征的稀疏分布使特征差异更加显著。

#### 2.4.3 eAPool中池化核的性能表现

在本节中讨论了eAPool中光滑近似平均池化eDSCPool和光滑近似最大池化eMPool两种池化核各自的性能表现,实验结果表明,单独使用eDSCPool比单独使用eMPool分类效果更好,证明基于DSC系数的指数相似度量在应对细粒度特征时能较好地保留细节特征,从而提高类别的可分离性。融合eDSCPool和eMPool两种池化核的eAPool获取了更加丰富的下采样信息,可以进一步提高实验结果。

表6 池化核的性能对比结果

Tab. 6 Performance comparison results for pooled kernels

Structure	Accuracy/%	Precision/%	Recall/%	F1-score/%
eAPool-w/o-eDSCPool	85.1	85.8	85.2	85.5
eAPool-w/o-eMPool	85.3	86.1	85.3	85.7
eAPool	85.5	87.0	85.5	86.2

#### 2.4.4 消融实验

针对本文使用的多种特征增强方法,做了消融实验来分析各个模块对网络性能的影响,表7对比了原始骨干网络ResNet18和模型改进后的实验结果,针对各个模块的性能效用,分别设计不同的模型结构进行对比,结构A中未使用可变形卷积的方法,仅对图像的细节特征和邻域特征的依赖关系进行强化。结构B中未使用eAPool池化核,图像细节有部分损失,但网络的有效感受野得到有效提高。结构C中未添加基于自注意的特征增强模块,损失了特征间的上下文信息。从实验结果来看,以上3种结构在模型中发挥着不同的效用,经过改进后的总体网络分类准确率有了明显的提升,但缺少任何一种结构都会降低模型的性能。证明有效感受野、细节特征和长程依赖关系在积碳程度判别特征的获取中均发挥着重要作用。

表7 不同特征采样模块对模型训练指标的影响

Tab. 7 Impact of different feature sampling modules on model training metrics

Structure	Accuracy/%	Precision/%	Recall/%	F1-score/%
ResNet18	84.6	85.2	84.6	84.9
A	86.0	86.8	86.0	86.4
B	86.5	87.3	86.5	86.9
C	86.2	87.1	86.2	86.6
Proposed method	87.1	87.8	87.1	87.4

## 3 结论

本文针对积碳程度分类任务中无法提取有效判别特征的问题展开研究,提出了一种新的基于非刚性特征的模型用于积碳程度判别。模型针对特征覆盖范围较广,缺乏明确边界条件的问题采用可变形卷积增大特征捕获有效感受野,避免局部梯度较大形成判别干扰。针对下采样过程中细节特征损失较为严重的现象,采用自适应的指数相似度量池化核,仔细考核每一个像素对分类性能的贡献,从而提升区分精度。针对图像类间分离性差的特点,添加基于自注意机制的特征增强模块,增强特征间的长程依赖关系。从本文的实验结果可以看出,提出的模型在积碳程度分类任务中有着较好的性能表现,可以为汽车发动机积碳程度的自动识别系统搭建提供技术支持。

#### 参考文献:

[1] 穆大芸,吕宪勇,宋国梁,等.一种喷气阀积碳的检

- 测方法、装置及电子设备: CN110005524B [P]. 2019-09-20.
- [ 2 ] 杨万里, 刘岩, 洪安新, 等. 一种发动机气缸盖、发动机、车辆及沉积物处理方法: CN113818970B [P]. 2022-12-23.
- [ 3 ] PRABHAKAR B, CONTI R. Non-contact process for engine deposit layer measurement: US11650043 [P]. 2021-03-25.
- [ 4 ] 晏娟, 邵传刚, 王利锋, 等. 一种燃料发动机内积碳厚度检测装置: CN209131594U [P]. 2019-07-19.
- [ 5 ] 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.
- [ 6 ] ZHU X, HU H, LIN S, et al. Deformable ConvNets v2: more deformable, better results [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 930-938.
- [ 7 ] STERGIOU A, POPPE R. AdaPool: exponential adaptive pooling for information-retaining downsampling [J]. IEEE Transactions on Image Processing, 2022, 32: 251-266.
- [ 8 ] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [ 9 ] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]// Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3-19.
- [ 10 ] LEE H J, KIM H E, NAM H. SRM: a style-based recalibration module for convolutional neural networks [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 1854-1862.
- [ 11 ] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 3146-3154.
- [ 12 ] LUO W, LI Y, URTASUN R, et al. Understanding the effective receptive field in deep convolutional neural networks [C]//Proceedings of the 30th International Conference on Neural Information Processing Systems, 2016: 4905-4913.
- [ 13 ] ZINTGRAF L M, COHEN T S, ADEL T, et al. Visualizing deep neural network decisions: prediction difference analysis [J]. arXiv preprint arXiv: 1702.04595, 2017.
- [ 14 ] ZHOU B, KHOSLA A, LAPEDRIZA A, et al. Learning deep features for discriminative localization [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2921-2929.
- [ 15 ] DABKOWSKI P, GAL Y. Real time image saliency for black box classifiers [C]//Proceeding of the 31st International Conference on Neural Information Processing Systems, 2017: 6970-6979.
- [ 16 ] STERGIOU A, POPPE R, KALLIATAKIS G. Refining activation downsampling with SoftPool [C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 10337-10346.
- [ 17 ] LEE C Y, GALLAGHER P W, TU Z. Generalizing pooling functions in convolutional neural networks: mixed, gated, and tree [EB/OL]. <http://arxiv.org/abs/1509.08985v2>.
- [ 18 ] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [ 19 ] YUN S, HAN D, CHUN S, et al. Cutmix: regularization strategy to train strong classifiers with localizable features [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, , 2019: 6023-6032.
- [ 20 ] 梁美彦, 张倩楠, 任竹云, 等. 基于注意力机制的结肠癌病理学图像识别研究 [J]. 测试技术学报, 2022, 36(2): 93-100.
- LIANG Meiyen, ZHANG Qiannan, REN Zhu Yun, et al. Research on identification of colon pathology image based on attention mechanism [J]. Journal of Test and Measurement Technology, 2022, 36(2): 93-100. (in Chinese)
- [ 21 ] DAI J, QI H, XIONG Y, et al. Deformable convolutional networks [C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 764-773.
- [ 22 ] WANG W, DAI J, CHEN Z, et al. InternImage: exploring large-scale vision foundation models with deformable convolutions [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 14408-14419.