

面向轻量化网络的火焰快速识别^①



薛颂东¹, 曹旺旺¹, 王 斌²

¹(太原科技大学 计算机科学与技术学院, 太原 030024)

²(中北大学 大数据学院, 太原 030051)

通信作者: 薛颂东, E-mail: xuesongdong@tyust.edu.cn

摘 要: 为解决火焰图像识别在边缘设备、移动端设备环境下模型体积大、准确率低、实时性能差的问题。首先选取 ShuffleNetV2 作为轻量化主干神经网络, 保证模型的实时性; 其次, 设计了一种新的注意力模块 SCDAM (space and channel dual attention module) 去同时考虑通道和空间的关联性, 针对不同特征的重要程度去赋予不同权重并有效提高模型精度; 然后, 设计了一种多尺度特征融合模块, 使提取到的特征在空间尺度上更加丰富, 加强网络对不同尺度的适应性; 最后将 SCDAM 模块以及多尺度模块引入到 ShuffleNetV2 中并利用迁移学习方式优化模型参数, 进一步提高模型精度。在参数量和计算量仅有微量增加的情况下, 本算法的精度比 ShuffleNetV2 提升了 3.2%, 且单次推理速度仅耗时 8.7 ms。实验证明, 该算法更加适合应用在计算资源有限情况下, 如火药火焰的识别与监控。

关键词: 火焰图像识别; ShuffleNetV2; 注意力机制; 多尺度特征融合; 迁移学习

引用格式: 薛颂东, 曹旺旺, 王斌. 面向轻量化网络的火焰快速识别. 计算机系统应用, 2023, 32(4): 274-282. <http://www.c-s-a.org.cn/1003-3254/9017.html>

Fast Flame Recognition for Lightweight Network

XUE Song-Dong¹, CAO Wang-Wang¹, WANG Bin²

¹(College of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan 030024, China)

²(School of Data Science and Technology, North University of China, Taiyuan 030051, China)

Abstract: Flame image recognition faces a large model size, low accuracy, and poor real-time performance in edge equipment and mobile terminal equipment environments. In order to solve these problems, firstly, ShuffleNetV2 is selected as a lightweight backbone neural network to ensure the real-time performance of the model. Secondly, a new space and channel dual attention module (SCDAM) is designed to analyze the relevance of channels and spaces, and different weights are given according to the importance of different features, so as to effectively improve the model accuracy. Then, in order to enrich the extracted features on the spatial scale, a multi-scale feature fusion module is designed to enhance the adaptability of the network to different scales. Finally, the SCDAM and multi-scale module are introduced into ShuffleNetV2, and the model parameters are optimized by transfer learning, so as to further improve the model accuracy. With only a slight increase in the amount of parameters and calculations, the accuracy of the proposed algorithm is 3.2% higher than that of ShuffleNetV2, and the single inference takes only 8.7 ms. Experiments show that the proposed algorithm is more suitable for conditions with limited computing resources, such as the identification and monitoring of gunpowder flame.

Key words: flame image recognition; ShuffleNetV2; attention mechanism; multi-scale feature fusion; transfer learning

① 基金项目: 教育部产学合作协同育人项目 (202102076011); 山西省高校教学改革创新项目 (J2021441); 山西省高等学校科技创新项目 (2021L322); 太原科技大学研究生教育改革研究课题 (XJG21019); 山西省基础研究计划 (201801D121126, 20210302124165); 研究生教育创新项目 (2019BY115, 2019BY107)

收稿时间: 2022-09-01; 修改时间: 2022-09-30; 采用时间: 2022-10-08; csa 在线出版时间: 2022-12-09

CNKI 网络首发时间: 2022-12-12

火灾是最危险的灾害类型之一,威胁着人类的生命财产、生态环境和基础设施。减少火灾造成的损失具有重要的理论和现实意义^[1-5]。尤其是当有用到火药的地方,如果火药在被小火星碰到时,就会迅速燃烧,甚至过多的火药可能会引起爆炸。当火药在刚要燃烧的瞬间迅速检测出来,并将其直接扑灭,以防止燃烧和爆炸发生,这一点则极为重要。

火的特征可分为静态特征和动态特征。静态特征包括目标信息和空间结构信息,动态特征包括整体运动特征和随机运动特征^[6]。早期的火灾探测方法通常根据这些特征中的一个或多个来识别火灾。吴茜茵等人^[7]根据火焰形状的纵横比和质心高度系数来判别火焰;Ouyang等人^[8]通过分析火焰的边缘信息,改进了Sobel算子,通过提取火焰颜色的边缘梯度特征来识别火焰。然而,早期这种手工提取的特征是高度冗余的,尽可能多的深度静态特征的智能提取是不可能的^[9,10]。并且这种传统的手工提取特征的方法检测速度很慢,达不到实时性,其次识别精度也不高,不能满足实际应用要求。

随着视频监控设备的日益普及和计算机视觉技术的发展,卷积神经网络得到了一个迅猛的发展,并且在各种视觉领域都取得了非常大的成果。2012年,AlexNet^[11]在ImageNet数据集上达到15.3%的top5错误率,与传统手工特征提取方法相比,在精度方面有了一个质的提升。随之出现了更深和更宽的卷积神经网络VGG^[12]以及GoogLeNet^[13]用于提升网络精度。而随后出现的带有残差结构的ResNet^[14]网络很好地缓解了网络过深而导致的梯度消失以及过拟合等问题。在火焰识别方面,卷积神经网络被用来识别当前环境中是否有火焰产生,从而起到预防作用。并且深度学习火焰识别方法最大的优点便是不再需要手动提取火焰特征^[15-17]。但是这些大型网络参数量和计算量庞大,不易部署到边缘设备上运行。而当前用于实际火焰识别项目的硬件设备为边缘设备或者移动端设备,因此要求网络需要足够轻量化并且能够保持一定的识别准确率。

为了促使网络轻量化,可部署到移动端和边缘设备上运行,2016年提出了SqueezeNet^[18],是最早的轻量级卷积神经网络,主要包括两部分:squeeze和expand。Squeeze中通过 1×1 卷积核对上一层特征图进行卷积,从而降低特征图维数;expand则是经过一组连续的 1×1 卷积和 3×3 卷积之后拼接构成。SqueezeNet的精度与AlexNet相比虽然没有降低,模型参数量也只有

AlexNet的1/50,但是他加深了网络的深度,并丧失了网络的并行能力,不能够很好地去满足实时性。因此,Qiang等人^[19]为了加快检测和识别速度,以SqueezeNet为基础,大幅压缩参数,并引入残差结构来集成高级和低级特征,从而提高精度降低计算成本;Zhang等人^[20]为了能够可靠的实时探测森林火灾,将SqueezeNet中的普通卷积替换为深度卷积,并在其Fire模块中使用通道混洗操作以作为特征通信器。

2017年,谷歌提出了采用多个深度可分离卷积去代替常规卷积进行堆叠形成的轻量级神经网络MobileNet^[21]。深度可分离卷积使用深度卷积去代替常规卷积进行特征提取,之后利用点卷积进行各通道之间的信息交流,有效减少模型参数。后来为了进一步优化MobileNet的网络结构,提出了带有反向残差结构的MobileNetV2^[22],但其小目标检测能力并不出众。为了针对早期火焰多为小目标的问题,He等人^[23]通过移植轻量级算法MobileNet SSD,并结合多特征融合技术,提高了算法的小目标检测能力;文献[24]以MobileNetV2网络为主干网络,并引入通道注意力机制,实现特征增强,在保证算法轻量化的同时有效提升网络的检测性能;Bouguezzi等人^[25]提出了一个增强版的MobileNet,通过对MobileNet中的深度可分离卷积单元进行修改,使模型精度获得了显著的提高。

ShuffleNet^[26]中利用shuffle操作去替换常规组卷积之后的点卷积操作,对组卷积之后的特征图在通道维度上随机均匀打乱,再进行组卷积操作,从而实现不同分组之间的信息交流。这样做可以很大程度的减少网络参数量和计算量,但是ShuffleNet中组卷积使用过多,且元素级操作过多,导致运行速度较慢。Yang等人^[27]通过奇异值分解对卷积进行分解优化,并用h-swish激活函数代替ReLU激活函数来改进ShuffleNet,以达到减少参数量和计算量的目的;Chen等人^[28]对ShuffleNet引入注意力机制以及采用新的激活函数来改善分类准确率低实时性差的问题。

深度可分离卷积是先学习空间之间的相关性,再学习通道之间的相关性,而Xception^[29]的Xception模块则完全相反,他是先学习通道之间的相关性,再学习空间之间的相关性,并且在学习过程中未使用激活函数,却能有效提升网络性能。Shi等人^[30]通过调整网络分支和网络层对Xception网络改进,从而实现了高精度的分类。

这些轻量化网络模型的出现,使模型速度更快,体积更小,同时能保持精度,能够在边缘设备中更好的运行.在本文中则以 ShuffleNetV2^[31]为基础,自主构建火焰图像数据集作为训练样本,从以下几个方面设计快速准确的火焰图像识别方法并用于实际项目中:1)设计了一种通道加空间的双重注意力模块 SCDAM (space and channel dual attention module),赋予重要特征一个更大的权重,同时抑制无用特征,在保证网络轻量化的基础上进一步去提升网络精度;2)设计多尺度特征融合模块,从而提升网络对不同尺度的一个适应性;3)将上述所提出的注意力模块和多尺度模块合理的嵌入到 ShuffleNetV2 中去,从而提高网络的检测能力;4)利用迁移学习对网络进行训练,更好地去提升网络性能.

1 算法设计

1.1 整体结构设计

ShuffleNetV2 是在 2018 年所提出的一种轻量化神经网络,其主要创新点是提出了 4 条轻量化原则,并通过这 4 条原则对 ShuffleNetV1 中的结构进行了改进.

本算法选取 ShuffleNetV2 为主干网络,在保证其轻量化的基础上加入自主设计的双重注意力机制模块 (SCDAM) 以及改进之后的多尺度特征融合模块,利用迁移学习来提升网络的分类准确率,改进之后的网络结构图如图 1 所示.图 1 中在 ShuffleNetV2 第 1 阶段以及第 3 阶段之后分别嵌入 SCDAM 注意力模块,在第 4 阶段之后嵌入多尺度模块,最终输出分类置信度,判断是否存在火焰.

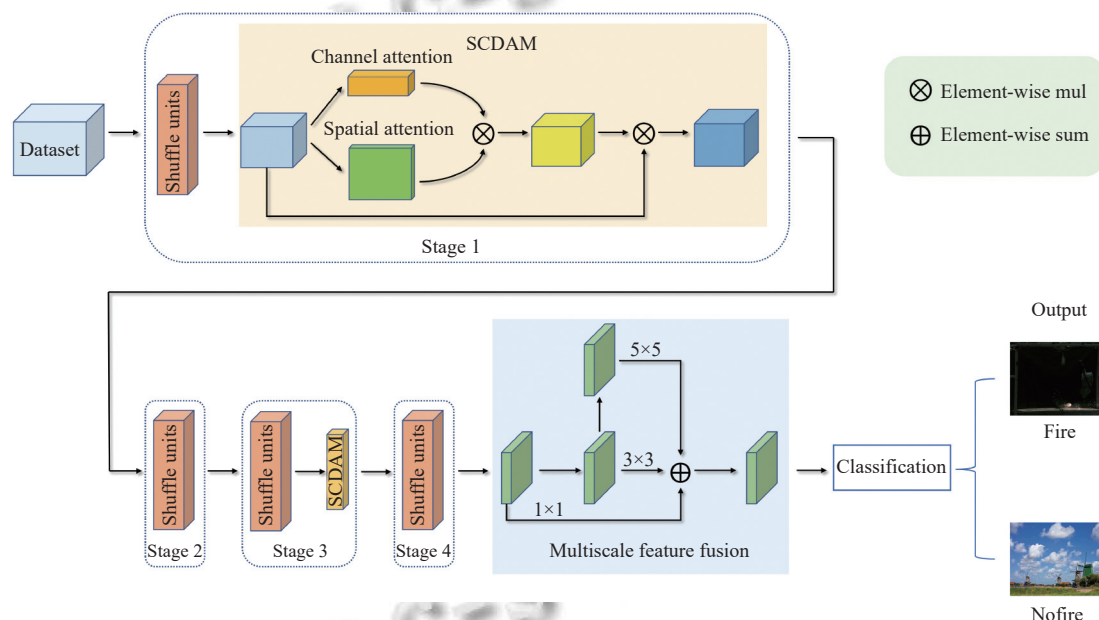


图 1 改进 ShuffleNetV2 整体框架图

1.2 SCDAM 模块设计

注意力机制最早是在自然语言处理中出现,而在视觉领域,注意力机制通过对特征图附加权重,从而去加强重要特征的提取,抑制不重要特征.目前,已有大量针对注意力机制的研究.最早出现的 SENet^[32]借鉴注意力机制思想,设计了一种通道注意力模块,引入 squeeze 和 excitation 两个过程,首先通过全局平均池化层将特征图在空间维度上压缩为 $1 \times 1 \times C$,再经过两个全连接层对特征图赋予权重,提升网络的学习能力.之后提出的 CBAM 模块^[33]在 SENet 上加以改进,不

止关注通道注意力,也对空间上的注意力加以关注,并证明了比单通道注意力效果更好.最近提出的 ECANet^[34]模块在 SENet 模块基础上,采用一维卷积避免通道降维,并实现了局部跨信道交互,使注意力模块变得更加轻量化,且分类准确率得到了很大程度的提升.

本文提出了一种新的空间加通道的双重注意力模块 SCDAM,其整体构成图如图 2 所示.当输入的特征图为 $X_U \in R^{C \times H \times W}$ 时,其中, C 、 H 、 W 分别为特征图的通道数、高度、宽度,经过两条不同分支的计算,上侧分支为通道注意力计算权重,得到特征尺寸为 $C \times 1 \times 1$

的权重 W_C ;下侧分支为空间注意力计算权重,得到特征尺寸为 $1 \times H \times W$ 的权重 W_S .将两个特征权重维度分别扩展为原始输入特征维度大小,然后将通道权重和空间权重相乘,并将相乘后的结果进行归一化操作,这里采用 Sigmoid 非线性激活函数进行归一化,得到最终的权重结果 $X_W \in R^{C \times H \times W}$.将 X_W 与原始输入特征图

X_U 相乘,赋予输入特征权重.表达式为:

$$X_W = \sigma(f_e(W_C) \otimes f_e(W_S)) \quad (1)$$

$$X_V = X_W \otimes X_U \quad (2)$$

其中, σ 为非线性激活函数 Sigmoid, f_e 为扩展函数, \otimes 表示元素间乘法, X_V 为赋予权重的特征图,其中 $X_V \in R^{C \times H \times W}$.

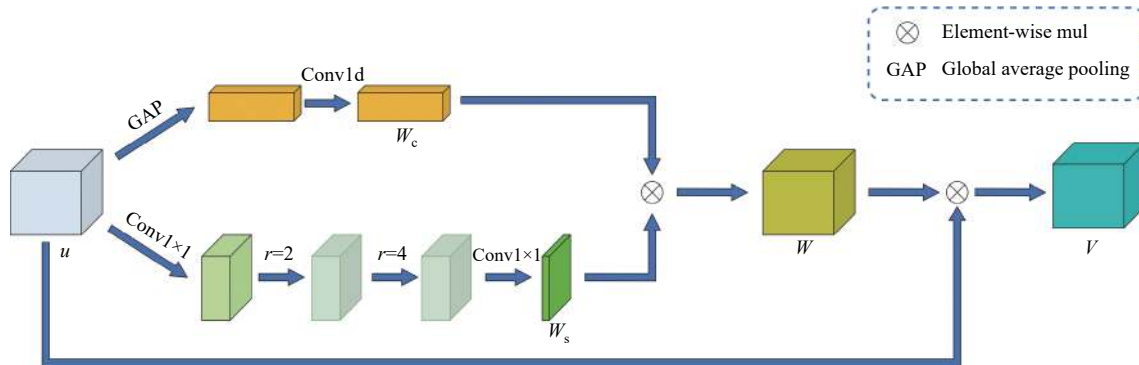


图2 SCDAM 注意力机制整体模块图

通道注意力机制的具体实现过程为,首先输入特征图的全局信息在空间维度上经过全局平均池化层进行压缩,压缩之后的通道特征图尺寸为 $C \times 1 \times 1$,之后再经过一维卷积去实现各信道之间的交流.这里不再采用全连接层去实现各通道之间的信息交流,一方面是因为一维卷积的计算量和参数量与全连接层相比会大幅度减小,使网络更加轻量化;另一方面则是能够避免降维,减少特征信息的丢失,针对通道的学习,不降维是非常重要的.通道注意力的表达式如下:

$$W_C = F_{1d}(F_{GAP}(X_U)) \quad (3)$$

其中, F_{1d} 为卷积核大小为3的一维卷积, F_{GAP} 为全局平均池化操作.

空间注意力是为了突出或抑制特定位置点的特征,在这里我们首先对输入特征图采用 1×1 卷积进行降维处理,然后使用两个深度可分离的 3×3 空洞卷积,空洞率分别为2和4.因为大的感受野可以更好的捕捉语义信息,采用深度可分离的空洞卷积可以扩大感受野,同时保持较低的复杂度,采用空洞率为2和4是为了避免提取的特征过于分散,同时也为了避免特征语义在空间上不连续.最后再用一个 1×1 卷积对输入特征做进一步降维处理,将其压缩为 $1 \times H \times W$ 的特征图.空间注意力的表达式如下:

$$W_S = F(f_4(f_2(F(X_U)))) \quad (4)$$

其中, F 代表普通卷积操作, f_2 和 f_4 分别代表空洞率为2和4的空洞卷积.

1.3 多尺度特征融合模块设计

为了提取到的特征在空间尺度上更加丰富,利于后续的火焰识别,采用多尺度特征融合方式. GoogLeNet 提出的 InceptionV1 模块利用 1×1 , 3×3 , 5×5 这3个不同大小的卷积核和 3×3 的池化操作对输入的特征进行提取,不止加宽了网络,同时增加了网络对不同尺度的适应性,提升了网络的性能.

在本文中我们对 InceptionV1 中模块进行改进,改进后的多尺度特征融合模块如图3所示.首先对输入特征使用 1×1 卷积进行降维处理,同时将处理后的结果作为 1×1 尺度特征提取的分支.然后进行两次 3×3 深度可分离卷积操作,将第1次 3×3 深度可分离卷积之后的结果作为 3×3 尺度特征提取的分支.在VGG中提出两个 3×3 卷积提供的感受野相当于一个 5×5 卷积提供的感受野,因此在这里将第2次 3×3 深度可分离卷积之后的结果作为 5×5 尺度特征提取的分支.将3条分支的结果进行相加进行特征融合,采用相加而非拼接是为了减少模型的参数量,融合之后使用 1×1 卷积做升维处理,恢复到输入特征时的维度大小.

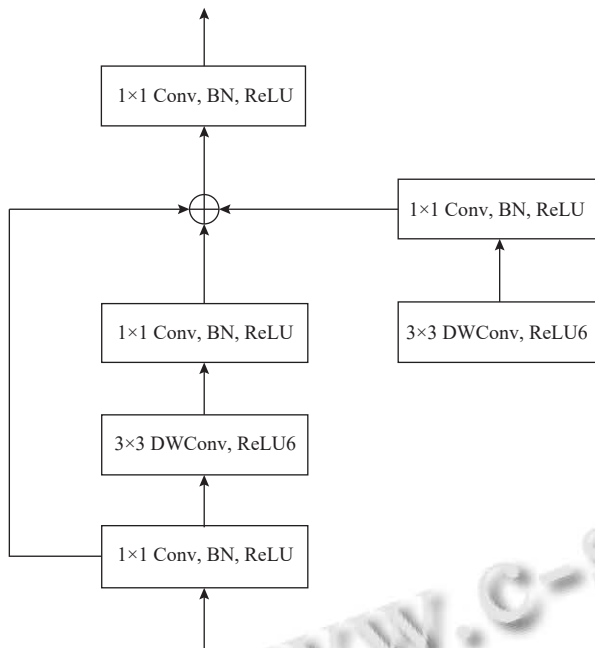


图3 多尺度特征融合模块

1.4 迁移学习

为了使模型获得更好的初始化权重参数, 利用模型在相似任务中的学习权重, 在此基础上采用迁移学习方式, 避免模型从零开始学习。

本研究利用 ShuffleNetV2 在 ImageNet 数据集中的预训练权重作为初始化权重, 由于对模型进行改进导致部分结构发生变化, 因此在导入权重时需要跳过不匹配的操作层。

2 实验与结果

2.1 实验环境与数据集介绍

本实验使用 Windows 10 系统, 深度学习框架采用 PyTorch 框架, CPU 为 Intel Core i7-10870, GPU 为 NVIDIA GeForce RTX2060 (显存 6 GB)。

所采用数据集来源为 3 部分, 第 1 部分来源于为网络收集, 第 2 部分为现场采集, 第 3 部分为公共数据集 (fire image data set for Dunnings 2018 study)。实验所用数据集共 19 182 张图片, 图片尺寸全设为 320×240, 分为两个类别 fire 和 nofire, 其中 fire 包含 9 887 张图片, nofire 包含正常非火图片以及类火干扰性较大的图片 (如夕阳朝阳图片, 灯泡灯光, 不锈钢反射光等), 共计 9 295 张图片, 具体样本数量如表 1 所示。代表性图片如图 4 所示。

在实验过程中将数据集按 8:2 划分为训练集和测

试集, 同时为加强模型泛化能力, 对数据进行随即裁剪, 翻转, 亮度变化等数据增强方式进行扩充。在模型训练过程中, 优化算法采用 Adam 优化器; 训练批次为 160, 批大小为 8; 损失函数使用交叉熵损失函数; 学习率为 0.0001。

表 1 数据集样本分布

Data set category		Dataset size
fire	—	9 887
nofire	Normal	6 727
	Interference	2 568



图4 代表性图片

2.2 注意力模块性能对比实验

为验证所设计注意力模块 SCDAM 的性能, 以 MobileNetV3 为基础网络, 将不同的注意力机制模块嵌入 MobileNetV3 进行对比, 在这里选用 MobileNetV3 为基础网络是因为 MobileNetV3 本身加入了注意力模块 SENet, 并且已经选好了加入注意力模块的位置, 我们只需要将其他注意力模块替换掉 MobileNetV3 本身自带的 SENet 进行对比即可。实验选取了 CBAM, BAM, ECA, SGE 以及 MobileNetV3 本身携带的 SENet 在准确率, 模型大小, 参数量, 计算量这 4 个方面进行对比来验证本文所提注意力模块 SCDAM 的有效性。与不同注意力模块的对比如表 2 所示。

表 2 各注意力模块效果对比

Model	Accuracy	Model size	Params	FLOPs
	(%)	(M)	(M)	(G)
MobileNetV3 (SENet)	91.8	16.2	4.205	0.346
MobileNetV3 (CBAM)	91.1	11.9	3.069	0.347
MobileNetV3 (BAM)	92.1	13.5	3.475	0.403
MobileNetV3 (ECA)	93.1	10.4	2.692	0.346
MobileNetV3 (SGE)	92.5	10.4	2.692	0.346
MobileNetV3 (SCDAM)	93.6	11.4	2.911	0.377

由表 2 可知, 本文所提出的注意力模块在精度方面与其他 5 个注意力模块相比有一个较为明显的提升, 且本文所提出的注意力模块的参数量比 SENet, CBAM, BAM 少, 与 ECA 和 SGE 相比也没有增加太多。整体

在准确率与参数量上达到一个比较适中的状态。

在 160 次迭代中, 嵌入不同注意力模块的 MobileNet-V3 准确率变化如图 5 所示. 绿色实线代表嵌入本文所提出的 SCDAM 模块的 MobileNetV3 在迭代过程中的准确率变化, 并且在 105 轮之后的迭代过程中绿色实线整体优于其他颜色实线。

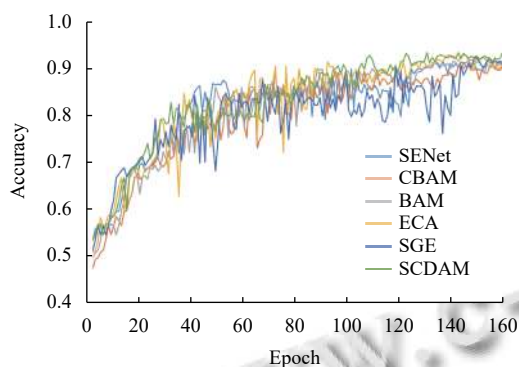


图 5 各模块迭代过程

综上所述, 有效证明了 SCDAM 的有效性, 尤其是在准确率上有较大提升. 主要包括两个方面: 一是在通道注意力方面选取一维卷积实现各通道之间的交流, 避免使用全连接层降维再升维的操作, 不止参数量和计算量有大幅度的减少, 且避免了降维过程中的信息丢失, 使最后的权重分配更加合理; 二是在空间注意力方面, 选取深度可分离的空洞卷积代替普通卷积, 有效的增加了感受野, 更好地去捕捉语义信息, 并且连续采用了两个不同空洞率的空洞卷积, 避免特征语义在空间上的不连续, 使提取到的显著特征图更加合理. 通道注意力和空间注意力两方面的优化效果叠加, 使得 SCDAM 的效果提升明显。

为了进一步验证本文所提出注意力模块的性能与其泛化性, 在公共数据集 CIFAR-100 上, 将该模块分别嵌入不同的网络进行测试. 实验选取了 6 种典型的网络, 分别以代表加深网络深度从而提升效果的 VGG16, 开创残差网络的 ResNet18, 稠密连接网络的 DenseNet, 实现多尺度的 GoogLeNet, 轻量化网络的 MobileNetV2 和 ShuffleNetV2 加以对比. 实验效果如表 3 所示。

由表 3 可知, 在不同的神经网络中加入本文所提出的注意力模块 SCDAM, 与原网络相比均有一定程度的效果提升, 准确率上平均提升幅度为 0.83%, 且在 VGG16 和 ShuffleNetV2 中提升幅度较大, 分别为 1.1% 和 1%。

表 3 不同神经网络效果对比

神经网络	Base	+SCDAM	提升率
VGG16	0.721	0.732	0.011
ResNet18	0.747	0.756	0.009
GoogLeNet	0.672	0.68	0.008
DenseNet	0.768	0.773	0.005
MobileNetV2	0.742	0.749	0.007
ShuffleNetV2	0.746	0.756	0.01

2.3 消融实验

本文以 ShuffleNetV2 为主干网络, 将 SCDAM 模块和改进之后的多尺度特征融合模块分别嵌入 ShuffleNetV2 中以改进 ShuffleNetV2 网络的性能. 通过消融实验证明本文提出的 SCDAM 模块, 多尺度特征融合模块以及迁移学习都对模型性能提升有所贡献, 消融实验结果如表 4 所示。

表 4 消融实验结果

Network	因素			Accuracy	Params (M)
	SCDAM	Multi-scale	Transfer learning		
ShuffleNetV2	—	—	—	0.909	1.256
	√	—	√	0.936	1.260
	—	√	√	0.933	1.286
	—	—	√	0.928	1.256
Ours	√	√	√	0.941	1.290

由表 4 可知, 利用迁移学习训练网络相比直接训练网络在精度方面会有一个巨大的提升, 模型的准确率可以提高 1.9%; 在利用迁移学习训练网络的基础上加入注意力模块 SCDAM, 可以看出在参数量微涨的情况下, 模型准确率可以提高 0.8%, 达到 93.6% 的准确率; 而将注意力模块换成多尺度特征融合模块后, 进一步加强网络对尺度的适应性, 可以发现, 在参数量略微增加的情况下, 准确率也可提升 0.5%; 最终本文在迁移学习的基础下初始化模型权重参数, 通过引入注意力模块 SCDAM 和多尺度特征融合模块去改进 ShuffleNetV2, 改进之后的 ShuffleNetV2 模型与原模型相比在准确率方面可以提升 3.2%, 达到 94.1% 的准确率, 且参数量与原模型相比只增加了 0.034M, 实现了在轻量化的基础上精度的进一步提升。

2.4 算法对比与分析

(1) 效果对比: 本文通过改进 ShuffleNetV2 实现了优良的效果, 在图 6 中我们通过一些可视化的样例展示了我们方法的优越性. 图中红色区域代表模型最后输出时所关注的区域, 颜色越深代表关注程度越高. 其

中左侧图片是普通火焰的样例,可以看出本文方法对火焰区域的关注程度都重于 ShuffleNetV2. 右侧图片是火药点燃时所产生火焰的样例,由于火药点燃产生

的火星微小,不易被检测出,而本文通过对 ShuffleNetV2 添加注意力模块以及多尺度模块,使模型进一步加深对不明显火焰区域特征的提取,从而将微小火星检测出来.

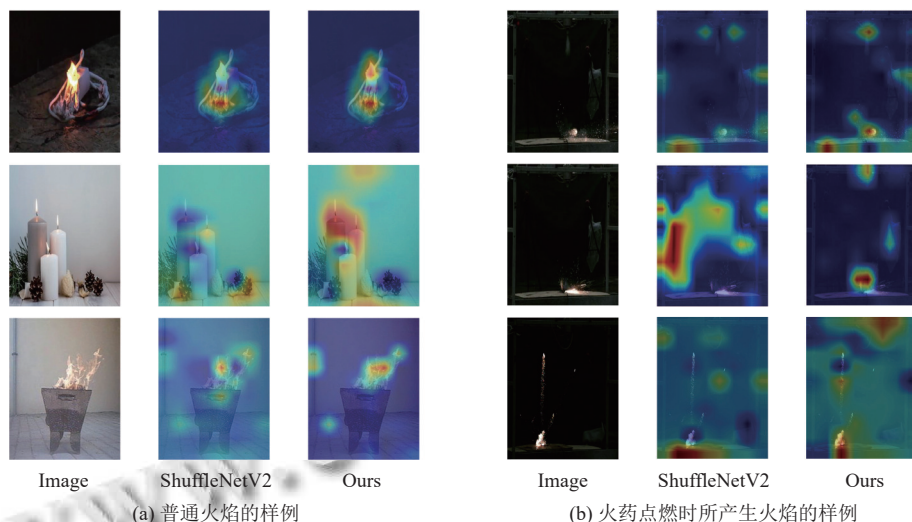


图6 模型最终输出时所关注区域的比较

图7中展示了在部分样例中,本文方法与 ShuffleNetV2 在各个阶段对火焰目标关注程度的对比. 在第2阶段前以及第4阶段前加入注意力模块使模型在不

同区域赋予不同的权重比重,在输出前加入的多尺度模块使模型进一步关注火焰本身区域,便于后续的火

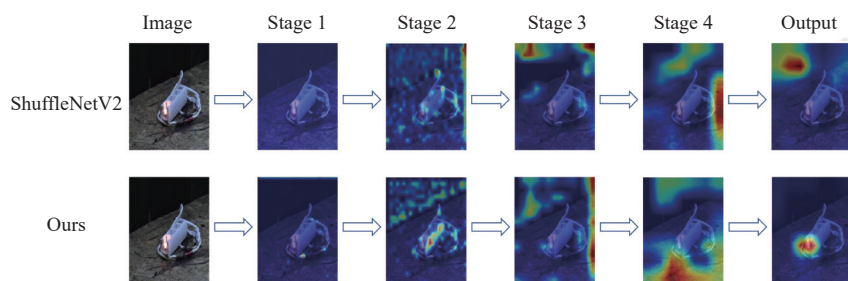


图7 模型在各阶段关注程度的样例

为了进一步验证本文方法在识别微小火焰方面的优越性,在本文所自制数据集上,将本文模型与 VGG, ResNet 等经典网络模型以及具有代表性的 MobileNet 系列和 ShuffleNet 系列的轻量化网络模型进行对比,对比结果如表5所示.

由表5可知,在本文自建数据集中 VGG16 与 MobileNetV2 的准确率最低,只有 91.1%. 本文算法准确率最高,可以达到 94.1%,且参数量与计算量相对都比较少.

(2) 推理速度对比: 为了验证本文在实时性方面也

能满足要求,在自建数据集中随机选取 100 张图片进行模型推理速度测试,结果表明:本文在 ShuffleNetV2 基础上采用迁移学习方式,加入 SCDAM 注意力模块以及多尺度模块,在精度提升 3.2% 的幅度下,平均推理速度也能达到 8.7 ms,完全能够满足本项目在实时性上的要求.

3 结论与展望

本文针对火药在燃烧时产生的小火苗进行及时检测,并将其扑灭避免火灾发生. 为了满足模型实时性要

求且能够在边缘设备中快速运行,选取轻量化网络 ShuffleNetV2 为基础网络,将所设计的空间加通道的双重注意力模块以及多尺度模块嵌入到 ShuffleNetV2 中,显著提高了模型的特征提取能力,以较小的开销代价提升了模型的准确率,且本文所提出的模型平均推理每张图片的速度仅为 8.7 ms,完全能够达到本项目在实时性上的要求。未来研究方向主要为如何在保持网络轻量化的前提下去进一步提升模型的抗干扰能力,且能够更准确地去分辨出各种类火干扰图片。

表5 各模型在本文自建数据集中训练结果对比

Model	Accuracy	Params (M)	FLOPs (G)
VGG16 ^[12]	0.911	178.309	23.674
ResNet18 ^[14]	0.931	11.178	2.826
MobileNetV2 ^[22]	0.911	2.226	0.487
MobileNetV3 ^[35]	0.918	4.205	0.346
ShuffleNetV1 ^[26]	0.917	0.919	0.224
ShuffleNetV2 ^[31]	0.928	1.256	0.231
Ours	0.941	1.290	0.235

参考文献

- Muhammad K, Ahmad J, Mehmood I, *et al.* Convolutional neural networks based fire detection in surveillance videos. *IEEE Access*, 2018, 6: 18174–18183. [doi: [10.1109/ACCESS.2018.2812835](https://doi.org/10.1109/ACCESS.2018.2812835)]
- Luo YM, Zhao L, Liu PZ, *et al.* Fire smoke detection algorithm based on motion characteristic and convolutional neural networks. *Multimedia Tools and Applications*, 2018, 77(12): 15075–15092. [doi: [10.1007/s11042-017-5090-2](https://doi.org/10.1007/s11042-017-5090-2)]
- Sudhakar S, Vijayakumar V, Kumar CS, *et al.* Unmanned aerial vehicle (UAV) based forest fire detection and monitoring for reducing false alarms in forest-fires. *Computer Communications*, 2020, 149: 1–16. [doi: [10.1016/j.comcom.2019.10.007](https://doi.org/10.1016/j.comcom.2019.10.007)]
- Sharma A, Singh PK, Kumar Y. An integrated fire detection system using IoT and image processing technique for smart cities. *Sustainable Cities and Society*, 2020, 61: 102332. [doi: [10.1016/j.scs.2020.102332](https://doi.org/10.1016/j.scs.2020.102332)]
- 王林, 赵红. 改进 YOLOv3 的火灾检测. *计算机系统应用*, 2022, 31(4): 143–153. [doi: [10.15888/j.cnki.csa.008404](https://doi.org/10.15888/j.cnki.csa.008404)]
- Foggia P, Saggese A, Vento M. Real-time fire detection for video-surveillance applications using a combination of experts based on color, shape, and motion. *IEEE Transactions on Circuits and Systems for Video Technology*, 2015, 25(9): 1545–1556. [doi: [10.1109/TCSVT.2015.2392531](https://doi.org/10.1109/TCSVT.2015.2392531)]
- 吴茜茜, 严云洋, 杜静, 等. 多特征融合的火焰检测算法. *智能系统学报*, 2015, 10(2): 240–247.
- Ouyang JN, Bu LP, Yang ZK, *et al.* An early flame identification method based on edge gradient feature. *Proceedings of the 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference*. Xi'an: IEEE Press, 2018. 642–646.
- Xie YK, Zhu J, Cao YG, *et al.* Efficient video fire detection exploiting motion-flicker-based dynamic features and deep static features. *IEEE Access*, 2020, 8: 81904–81917. [doi: [10.1109/ACCESS.2020.2991338](https://doi.org/10.1109/ACCESS.2020.2991338)]
- 罗梦研, 刘雁飞. 基于轻量化卷积神经网络的服装分类方法. *计算机系统应用*, 2019, 28(3): 223–228. [doi: [10.15888/j.cnki.csa.006821](https://doi.org/10.15888/j.cnki.csa.006821)]
- Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 2017, 60(6): 84–90. [doi: [10.1145/3065386](https://doi.org/10.1145/3065386)]
- Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv:1409.1556*, 2015.
- Szegedy C, Liu W, Jia YQ, *et al.* Going deeper with convolutions. *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston: IEEE, 2015: 1–9. [doi: [10.1109/cvpr.2015.7298594](https://doi.org/10.1109/cvpr.2015.7298594)]
- He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE Press, 2016. 770–778. [doi: [10.1109/cvpr.2016.90](https://doi.org/10.1109/cvpr.2016.90)]
- Muhammad K, Ahmad J, Lv ZH, *et al.* Efficient deep CNN-based fire detection and localization in video surveillance applications. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2019, 49(7): 1419–1434. [doi: [10.1109/TSMC.2018.2830099](https://doi.org/10.1109/TSMC.2018.2830099)]
- Huang HY, Kuang P, Fan L, *et al.* An improved multi-scale fire detection method based on convolutional neural network. *Proceedings of the 17th International Computer Conference on Wavelet Active Media Technology and Information Processing*. Chengdu: IEEE, 2020. 109–112.
- 徐小平, 余香佳, 刘广钧, 等. 利用改进 AlexNet 卷积神经网络识别石墨. *计算机系统应用*, 2022, 31(2): 376–383. [doi: [10.15888/j.cnki.csa.008311](https://doi.org/10.15888/j.cnki.csa.008311)]
- Iandola FN, Han S, Moskewicz MW, *et al.* SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size. *arXiv:1602.07360*, 2016.
- Qiang BH, Zhai YJ, Zhou ML, *et al.* SqueezeNet and fusion

- network-based accurate fast fully convolutional network for hand detection and gesture recognition. *IEEE Access*, 2021, 9: 77661–77674. [doi: [10.1109/ACCESS.2021.3079337](https://doi.org/10.1109/ACCESS.2021.3079337)]
- 20 Zhang JM, Zhu HQ, Wang PY, *et al.* ATT Squeeze U-Net: A lightweight network for forest fire detection and recognition. *IEEE Access*, 2021, 9: 10858–10870. [doi: [10.1109/ACCESS.2021.3050628](https://doi.org/10.1109/ACCESS.2021.3050628)]
- 21 Howard AG, Zhu ML, Chen B, *et al.* MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv:1704.04861, 2017.
- 22 Sandler M, Howard A, Zhu ML, *et al.* MobileNetV2: Inverted residuals and linear bottlenecks. *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 4510–4520. [doi: [10.1109/CVPR.2018.00474](https://doi.org/10.1109/CVPR.2018.00474)]
- 23 He HF, Huang PJ, Xiong AM, *et al.* Design of embedded real-time video fire AI recognition system. *Proceedings of 2021 International Conference on Computer Communication and Artificial Intelligence*. Guangzhou: IEEE, 2021. 52–55.
- 24 任坤, 黄泷, 范春奇, 等. 基于多尺度像素特征融合的实时小交通标志检测算法. *信号处理*, 2020, 36(9): 1457–1463. [doi: [10.16798/j.issn.1003-0530.2020.09.010](https://doi.org/10.16798/j.issn.1003-0530.2020.09.010)]
- 25 Bouguezzi S, Faiedh H, Souani C. Slim MobileNet: An enhanced deep convolutional neural network. *Proceedings of the 18th International Multi-Conference on Systems, Signals & Devices*. Monastir: IEEE, 2021. 12–16.
- 26 Zhang XY, Zhou XY, Lin MX, *et al.* ShuffleNet: An extremely efficient convolutional neural network for mobile devices. *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 6848–6856. [doi: [10.1109/cvpr.2018.00716](https://doi.org/10.1109/cvpr.2018.00716)]
- 27 Yang DD, Peng HY, Wan TZ. Attachment fault diagnosis for tidal stream turbine based on improved ShuffleNetV2. *Proceedings of 2021 China Automation Congress*. Beijing: IEEE, 2021. 6844–6849. [doi: [10.1109/CAC53003.2021.9728227](https://doi.org/10.1109/CAC53003.2021.9728227)]
- 28 Chen ZC, Yang J, Chen LF, *et al.* Garbage classification system based on improved ShuffleNetV2. *Resources, Conservation and Recycling*, 2022, 178: 106090. [doi: [10.1016/j.resconrec.2021.106090](https://doi.org/10.1016/j.resconrec.2021.106090)]
- 29 Chollet F. Xception: Deep learning with depthwise separable convolutions. *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 1800–1807. [doi: [10.1109/cvpr.2017.195](https://doi.org/10.1109/cvpr.2017.195)]
- 30 Shi CP, Xia RY, Wang LG. A novel multi-branch channel expansion network for garbage image classification. *IEEE Access*, 2020, 8: 154436–154452. [doi: [10.1109/ACCESS.2020.3016116](https://doi.org/10.1109/ACCESS.2020.3016116)]
- 31 Ma NN, Zhang XY, Zheng HT, *et al.* ShuffleNetV2: Practical guidelines for efficient CNN architecture design. *Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018. 122–138.
- 32 Hu J, Shen L, Albanie S, *et al.* Squeeze-and-excitation networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011–2023. [doi: [10.1109/TPAMI.2019.2913372](https://doi.org/10.1109/TPAMI.2019.2913372)]
- 33 Woo S, Park J, Lee JY, *et al.* CBAM: Convolutional block attention module. *Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018. 3–19. [doi: [10.1007/978-3-030-01234-2_1](https://doi.org/10.1007/978-3-030-01234-2_1)]
- 34 Wang QL, Wu BG, Zhu PF, *et al.* ECA-Net: Efficient channel attention for deep convolutional neural networks. *Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020. 11531–11539. [doi: [10.1109/cvpr42600.2020.01155](https://doi.org/10.1109/cvpr42600.2020.01155)]
- 35 Howard A, Sandler M, Chen B, *et al.* Searching for MobileNetV3. *Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. Seoul: IEEE, 2019. 1314–1324. [doi: [10.1109/ICCV.2019.00140](https://doi.org/10.1109/ICCV.2019.00140)]

(校对责编: 牛欣悦)