

# 基于改进 Resnet18 网络的火灾图像识别

陈跨越<sup>a</sup>, 王保云<sup>a,b</sup>

(云南师范大学 a.数学学院; b.云南省现代分析数学及应用重点实验室, 昆明 650500)

**摘要:** 针对传统卷积神经网络进行火灾图像识别时, 准确率不高、特征难以提取、网络的平移不变性较弱等问题, 对 Resnet18 网络进行改进, 使其具有更高的性能和准确性。首先, 在 Resnet18 网络的卷积层前插入空间变换网络(spatial transform networks, STN)。对于连续多个相同大小的卷积层, 只在第一个卷积层前添加 STN, 共添加 5 个, 并且在全连接层后添加 dropout 层防止过拟合。然后, 使用迁移学习(transfer learning, TL)的方法对火灾进行分类识别。实验结果表明, 改进后的 Resnet18 网络准确率、召回率、 $F_1$  值和 AUC 值等各项指标性能优于 Resnet18 网络和其他深度学习识别算法, 能够对火灾图像进行快速、准确地识别。

**关键词:** 火灾检测; 卷积神经网络; 空间变换网络; Resnet18; HSI 色彩模型; 迁移学习

**中图分类号:** TP391.41; TP183; X928.7 **文献标志码:** A **文章编号:** 1000-2367(2024)04-0101-10

火灾严重威胁了人类的生命和财产安全, 在很大程度上阻碍了社会的发展。因此, 火灾预测的意义在于能够提前发现潜在的火灾风险, 并采取相应的措施来减少火灾的发生以及对人类和财产的损失。通过火灾预测技术, 可以对建筑物、森林、工厂等进行监测和分析, 及时发现潜在的火灾隐患, 提高火灾预警能力, 有效降低火灾事故的发生率和损失。

在早期主要使用感光式、感烟式等各种传感器<sup>[1]</sup>对火灾进行检测。但该检测系统易受周围电子噪声及环境的影响, 误报率极高, 并且这种方式不能掌握现场火灾的情况。随着计算机图像处理和人工智能技术的进步, 基于深度学习的火灾检测方法发展迅速。该方法可以自动从图像中提取特征量, 从而实现对火灾的快速检测和报警。相比传统的火灾检测方法, 基于深度学习的方法具有更高的准确度和更快的响应速度, 同时还能够自动学习和优化模型, 提高检测的鲁棒性和可靠性。因此, 基于深度学习的火灾检测方法在实际应用中具有广泛的应用前景和 market 价值。

近年来, 越来越多的国内外学者开始通过卷积神经网络(convolutional neural network, CNN)<sup>[2]</sup>来对火灾图像进行预测分类并取得了良好的识别效果。比如经典的卷积神经网络模型 VGG16。它包含 16 层卷积层和全连接层, 主要特点是使用了非常小的卷积核( $3 \times 3$ ), 并且将多个卷积层和池化层串联在一起, 形成了非常深的网络结构。这种设计可以增加网络的深度, 从而提高网络的表达能力和分类准确率。同时, 使用小的卷积核可以减少模型的参数数量, 降低过拟合的风险。蒋珍存等<sup>[3]</sup>通过对 VGG16 模型的改进, 采用迁移学习技术, 有效地将火灾样本划分为不同的类别。吴彩朋等<sup>[4]</sup>改进轻量级卷积神经网络模型 SqueezeNet, 采用数据增强技术增加火灾数据样本, 通过应用运动检测技术, 从图像中提取出火焰的特征, 从而实现对火灾图像的

收稿日期: 2023-05-19; 修回日期: 2023-07-06.

基金项目: 国家自然科学基金(61966040).

作者简介: 陈跨越(2000—), 男, 山东济宁人, 云南师范大学硕士研究生, 研究方向为图像处理和深度学习, E-mail: 2590701362@qq.com.

通信作者: 王保云(1977—), 男, 云南玉溪人, 云南师范大学副教授, 博士, 研究方向为机器学习及图像处理, E-mail: wspbmlly@163.com.

引用本文: 陈跨越, 王保云. 基于改进 Resnet18 网络的火灾图像识别[J]. 河南师范大学学报(自然科学版), 2024, 52(4): 101-110. (Chen Kuayue, Wang Baoyun. Fire image recognition based on improved Resnet18 network[J]. Journal of Henan Normal University(Natural Science Edition), 2024, 52(4): 101-110. DOI: 10.16366/j.cnki.1000-2367.2023.05.19.0001.)

识别.JACEK 等<sup>[5]</sup>利用网上提供的森林火灾图像库,采用 AlexNet 网络框架来识别火灾图像中的烟雾,也取得了良好的效果.

尽管改进后的神经网络的性能有了显著提高,但依然存在一些问题.一般的深度学习网络会随着网络层数的增加导致算法的复杂度逐渐递增,同时使调取参数的数量增大.在网络结构中加入残差模块可以在一定程度上解决这一问题,因此文章的基础模型采用具有 18 层网络结构的残差网络 Resnet18.另外,当卷积神经网络进行卷积和池化处理时,其下采样过程可能会削弱网络的平移特征,从而导致其稳定性受到影响.空间变换网络(spatial transformation network,STN)<sup>[6]</sup>引入了一个新的可学习的空间转换模块.它能够把原始的图像纠正成为网络想要的理想图像,可以使模型具有空间不变性.由于所收集到的火灾图像的数量有限,因此采用迁移学习<sup>[7]</sup>技术.它可以有效地解决火灾样本量不足这一问题,同时缩短训练时间.迁移学习的策略是对网络结构进行微调(fine-tune),这可以有效保持模型的特征提取层并且实现对模型更加精确的分类和识别.

针对火灾图像特征不明显、网络模型过拟合和梯度消失以及不变性能力弱等火灾图像识别问题.本研究采用 Resnet18 网络结合空间变换网络进行训练,把原始的图像纠正成为网络想要的理想图像,提高网络的不变性;然后用迁移学习对提取数据集特征进行分类.

## 1 原理和方法

### 1.1 改进的 Resnet18 网络

一般的卷积神经网络当层数少时学习图像特征不够充分,当层数多时容易导致模型过拟合和梯度消失以及训练持续时间过长等问题.残差神经网络(ResNet)于 2015 年被提出<sup>[8]</sup>,其特有的网络结构可以克服上述问题.残差模块(BasicBlock)的结构如图 1(a)所示.

文章采用具有 18 层网络结构的残差网络 Resnet18.其网络结构为:首先,输入样本经过一个 kernel\_size 为  $7 \times 7$ , stride 为 2, padding 为 3 且 out\_channels 为 64 的卷积层.然后,在卷积层后接一个 kernel\_size 为  $3 \times 3$ , stride 为 2, padding 为 1 的最大池化层.接着使用 4 个模块,每个模块均由两个具有同样输出通道数的残差块构成,且均使用  $3 \times 3$  大小的卷积核.其中第一个模块的输入和输出通道数均为 64,在其后的每个模块的通道数都是前一个模块通道数的 2 倍.另外,Resnet18 网络模型中的残差块在卷积层之后添加了 batch normalization(BN)<sup>[9]</sup>对数据进行归一化处理,可以增加网络的稳定性.其残差模块结构如图 1(b)所示.最后,在 4 个模块后接一个平均池化层,并用 Softmax 函数<sup>[10]</sup>对提取到的特征进行分类.具体 Resnet18 网络结构如图 2 所示.图 2 中实线表示残差块中的通道数保持不变,虚线表示通道数进行翻倍.

虽然 Resnet18 网络定义了一个非常健壮模型,但仍然受到限制,缺乏以计算和参数有效的方式对输入数据进行空间不变能力的应对.空间变换原理和迁移学习可以与 ResNet18 网络结合使用以提高模型性能和泛化能力.对于空间变换原理,可以在 ResNet18 网络中引入空间变换器模块,从而使得网络可以对输入进行平移、旋转、缩放等操作,增强网络的鲁棒性和泛化能力.这样可以使得 ResNet18 网络更加适应不同的数据集和任务.对于迁移学习,可以将已经训练好的 ResNet18 网络作为预训练模型,然后新的数据集上微调模型参数,以适应新的任务.这样可以加快模型训练的速度并提高模型的性能.同时,在迁移学习中还可以使

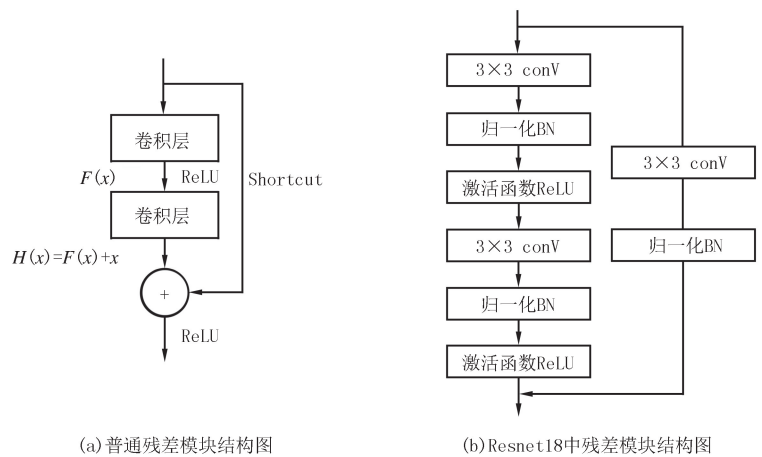


图1 残差模块结构示意图

Fig.1 Residual module structure diagram

用一些技巧,如冻结预训练模型的某些层、调整学习率等,以进一步提高模型性能.

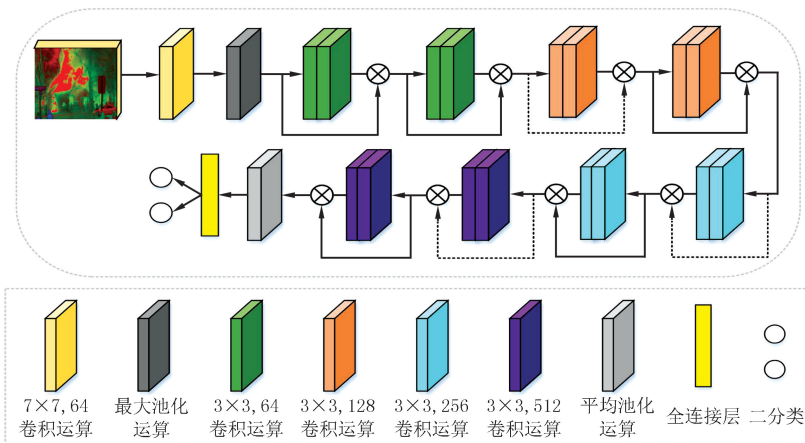


图2 Resnet18网络分类模型

Fig.2 Resnet18 network classification model

本研究主要对 Resnet18 网络做了如下改进.因为火灾识别是一个二分类问题,无需使用复杂的分类层,所以对 Resnet18 网络原有的卷积层、激活函数和池化层进行保留.将 STN 加入 Resnet18 网络卷积层之前,通过对输入图像进行平移、旋转、缩放等变换操作,使得网络可以更好地适应不同的输入图像.对于连续多个相同大小的卷积层,则只在第一个卷积层前面添加 STN,因此共需要添加 5 个 STN.使用 ReLU 激活函数<sup>[11]</sup>,其计算公式如式(1),增加了 dropout 层<sup>[12]</sup>防止过拟合.最后,使用迁移学习对火灾图像进行分类.首先使用在 ImageNet 数据集上训练得到的模型,来初始化 Resnet18 网络的权重.然后将改进后的 Resnet18 网络应用于火灾数据集上,通过对网络进行微调来对火灾图像进行分类,

$$f(x) = \max(0, x). \tag{1}$$

改进后的 Resnet18 网络结构框架如图 3 所示.

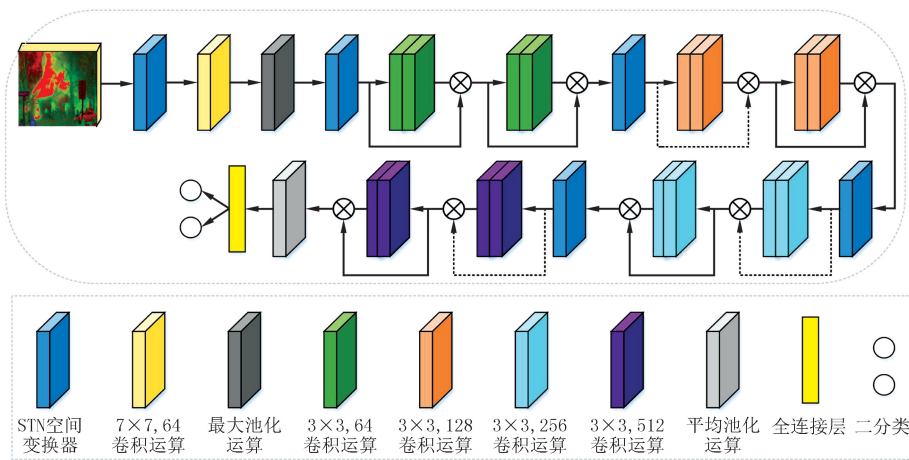


图3 改进的Resnet18网络火灾分类模型

Fig.3 Improved Resnet18 network fire classification model

### 1.2 空间变换网络原理

空间变换网络(STN)采用卷积网络技术,可以有效地减少输入图像中数据的多样性,从而提升分类准确度,而无需改变网络结构.该网络能够根据分类或其他任务的需求,插入现有网络的任意位置.自动调整数据的位置,包括平移、缩放、旋转以及其他几何变换,进而实现空间上的优化.

STN 由定位网络(Localisation net),网络生成器(Grid generator)和采样器(Sampler)3 部分组成<sup>[13]</sup>.其网络结构如图 4 所示.

Localisation net 是一个卷积神经网络,用于学习变换参数矩阵  $\theta$ .  $\theta$  的尺寸大小取决于所采用的变换方式.本研究的火灾图像识别属于 2D 仿射变换,那么  $\theta$  在本研究中是一个 6 维( $2 \times 3$ )向量.

Grid generator 通过变换参数  $\theta$ ,生成一个输入与输出图像之间的坐标关系网格.网格生成器是一种映射关系  $T_\theta$ ,其二维仿射变换关系式如式(2).

$$\begin{pmatrix} x_i^s \\ y_i^s \end{pmatrix} = T_\theta(G_i) = A_\theta \begin{pmatrix} x_i^t \\ y_i^t \\ 1 \end{pmatrix} = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \end{bmatrix} \begin{pmatrix} x_i^t \\ y_i^t \\ 1 \end{pmatrix}, \quad (2)$$

其中,  $(x_i^s, y_i^s)$  与  $(x_i^t, y_i^t)$  分别表示输入图像  $U$  和输出图像  $V$  中每个像素的坐标.

Sampler 是将输入图像按网格的坐标对照关系转换为一个新图像,转发给分类卷积神经网络.在计算得到  $T_\theta$  后,利用双线性插值<sup>[14]</sup>,可以根据式(3)由  $U$  得到  $V$ .

$$V_i^c = \sum_n^H \sum_m^W U_{nm}^c \max(0, 1 - |x_i^s - m|) \max(0, 1 - |y_i^s - n|), \quad (3)$$

其中,  $U_{nm}^c$  是输入图像  $(n, m, c)$  位置的像素值,  $V_i^c$  是输出图像  $(x_i^t, y_i^t, c)$  位置的像素值.

在求得  $V_i^c$  后,对  $U_{nm}^c, x_i^s, y_i^s$  求偏导,以便根据 loss 进行网络的反向传播,具体求偏导公式如式(4)~(7).

$$\frac{\partial V_i^c}{\partial U_{nm}^c} = \sum_n^H \sum_m^W \max(0, 1 - |x_i^s - m|) \max(0, 1 - |y_i^s - n|), \quad (4)$$

$$\frac{\partial V_i^c}{\partial x_i^s} = \sum_n^H \sum_m^W U_{nm}^c \max(0, 1 - |y_i^s - n|) \begin{cases} 0, & |m - x_i^s| \geq 1, \\ 1, & m \geq x_i^s, \\ -1, & m < x_i^s, \end{cases} \quad (5)$$

$$\frac{\partial V_i^c}{\partial y_i^s} = \sum_n^H \sum_m^W U_{nm}^c \max(0, 1 - |x_i^s - m|) \begin{cases} 0, & |n - y_i^s| \geq 1, \\ 1, & n \geq y_i^s, \\ -1, & n < y_i^s, \end{cases} \quad (6)$$

对  $\theta$  求导为:

$$\frac{\partial V_i^c}{\partial \theta} = \begin{pmatrix} \frac{\partial V_i^c}{\partial x_i^s} \cdot \frac{\partial x_i^s}{\partial \theta} \\ \frac{\partial V_i^c}{\partial y_i^s} \cdot \frac{\partial y_i^s}{\partial \theta} \end{pmatrix}, \quad (7)$$

其中,  $\frac{\partial x_i^s}{\partial \theta}, \frac{\partial y_i^s}{\partial \theta}$  根据具体的变换函数便可得到.

通过以上 3 个部分的结合,便形成了完整的 STN.

### 1.3 迁移学习

通常训练一个深度卷积神经网络所需的样本数量超过上百万个.实际上,能够搜集到的火灾图像却极其有限.因此,采用迁移学习技术,可以在有限的数据集上,迅速地将已学习的特征应用于新的分类任务中.迁移学习中,已有的知识称为源领域(source domain,  $D_s$ ),公式如式(8),待解决的知识称为目标域(target domain,  $D_t$ ),公式如式(9).其中,  $X_s$  和  $X_t$  分别是源领域和目标域的特征空间,  $P(X_s)$  和  $P(X_t)$  分别是特征

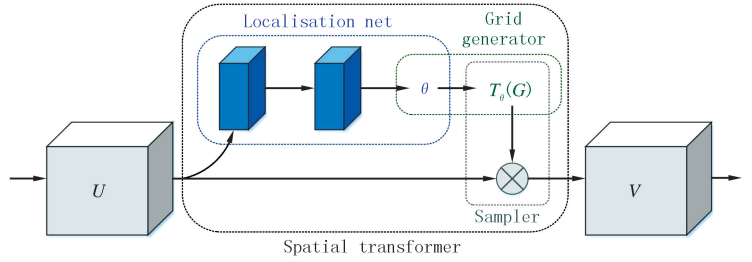


图4 空间变换网络原理图

Fig.4 Spatial transformation network schematic diagram



空间对应的边际概率分布.

$$D_s = \{X_s, P(X_s)\}, \tag{8}$$

$$D_t = \{X_t, P(X_t)\}. \tag{9}$$

迁移学习的目的是在源领域( $D_s$ )和目标领域( $D_t$ )不同或者源任务(source task,  $T_s$ )和目标任务(target task,  $T_t$ )不同的情况下,用  $D_s$  和  $T_s$  的特征来提升  $T_t$  的预测效果.因为神经网络的结构可以直接进行迁移,所以本研究采用基于模型的迁移学习方法(model based transfer Learning)<sup>[15]</sup>.该方法是指从  $D_s$  和  $D_t$  中找到共享的参数信息,以实现迁移学习.微调(fine-tune)就是模型参数迁移的很好体现.

迁移学习可以有效克服深度学习所需要的大量样本训练模型的局限性,从而实现在有限的数据集上训练出更加高效的模型.fine-tune 可以有效地保持模型的特征提取层,并且通过对 ImageNet 数据集上的 Resnet18 网络进行微调,从而实现对模型更加精确的分类和识别.Resnet18 在 ImageNet 数据集上训练输出特征数是 1 000,而对于火灾分类来说,需要分 2 类,因此要改一下全连接层的输出.在学习率的设置上,考虑到用于特征提取的 Resnet18 网络已经经过预训练,所以要设置较低的学习率来更好的拟合模型,设为 0.000 1.重新训练改进后的 Resnet18 网络,对火灾图像进行分类识别.本研究的迁移学习训练过程如图 5 所示.

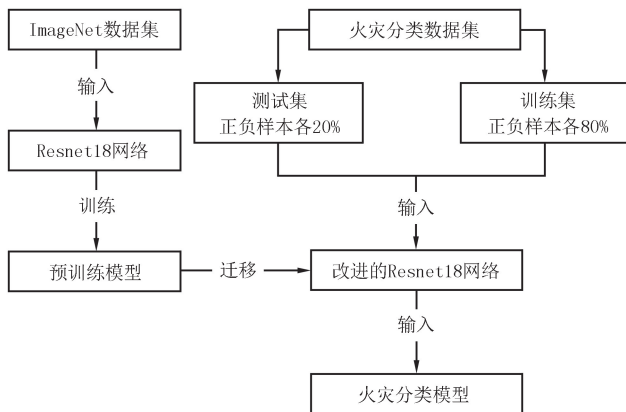


图5 迁移学习训练过程  
Fig.5 Transfer learning training process

## 2 实验数据

### 2.1 数据收集

本研究所采用数据来自 Visifire 数据集和项目组自行收集的火灾图像.共包括 15 000 张图片,其中不同场景下有火灾图片共 6 600 张,无火灾图片共 8 400 张.其部分样本如图 6 所示.其中图 6(a)为不同光照强度、近景远景、不同场景下的火灾数据集部分样本,图 6(b)为亮光、红黄物体等类似火灾的无火灾数据集部分样本.

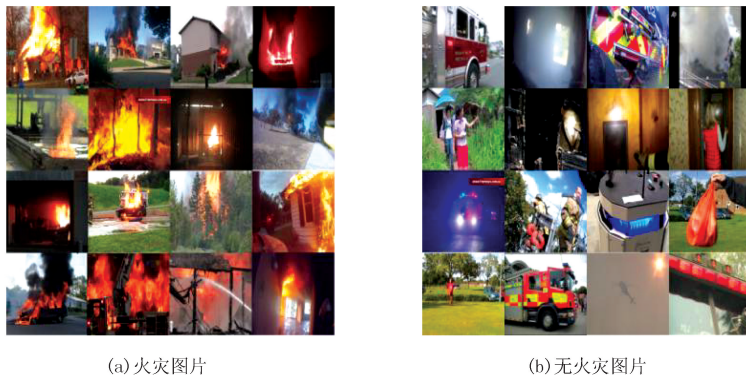


图6 实验图像示例  
Fig.6 Experimental image example

### 2.2 图片预处理

与 RGB(red,green,blue)图像相比,HSI(hue,saturation,intensity)<sup>[16]</sup>图像更能有效地利用颜色信息,因此更适用于目标识别和分割等应用场景.HSI 直接用颜色三要素:色调(hue)、饱和度(saturation)和亮度

(intensity)来准确地表达出颜色的特征。

有多种方法将图片的 RGB 色彩格式转换到 HSI 色彩格式,其中最常用的是几何推导法,具体转换过程如下式:

$$\begin{aligned} R, G, B &\in [0, 255], \\ (R', G', B') &= \frac{(R, G, B)}{255}, \\ \theta &= \arccos \frac{R' - G' + R' - B'}{2\sqrt{(R' - G')^2 + (R' - B')(G' - B')}}}, \\ H &= \begin{cases} \theta, & B' \leq G', \\ 2\pi - \theta, & B' > G' \end{cases} \\ I &= \frac{R' + G' + B'}{3}, \\ S &= 1 - \frac{3\min\{R', G', B'\}}{R' + G' + B'}, \end{aligned}$$

其中,  $R, G, B$  分别表示火灾样本图片 RGB 格式中每一个像素的通道数,  $H, S, I$  分别表示 HSI 色彩格式中的色调、饱和度、亮度的通道数。

通过转换公式,将数据集中的图像由 RGB 格式转换成 HSI 色彩格式.火灾 RGB 格式和 HSI 格式图像及其各自通道图像如图 7 所示,将其  $H, S, I$  三通道作为卷积神经网络的输入。

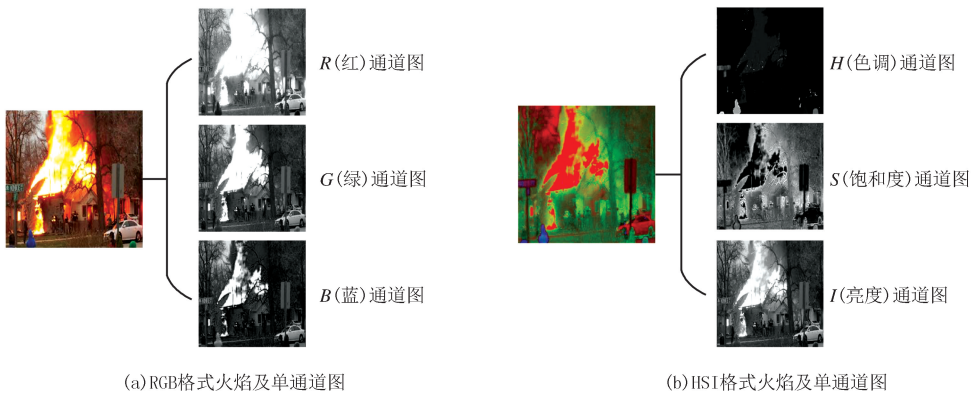


图7 RGB与HSI火焰图

Fig.7 RGB and HSI flame graphs

通过预处理,将图片转换为 HSI 格式.在进入网络之前,先将样本图片统一调整为  $224 \times 224 \times 3$  大小.为了满足模型的训练要求,将有火灾的图片作为正样本(记为 0),将无火灾的图片作为负样本(记为 1).为了避免模型在训练过程中产生过拟合或欠拟合现象,各类别的训练集、测试集按照 4 : 1 的比例进行划分,那么原始数据划分情况如表 1 所示。

表 1 原始数据划分

Tab. 1 Raw data partitioning			
类别	训练集	测试集	总计
有火灾(正样本)	5 280	1 320	6 600
无火灾(负样本)	6 720	1 680	8 400

### 3 实验结果与分析

#### 3.1 实验环境与参数设置

实验设备为 64 位 Microsoft Windows10 操作系统,英特尔 Core i5-8250U@1.6 GHz 四核处理器,威刚 DDR4 2 666 MHz 8 GB 内存,NVIDIA GeForce 930MX 显卡。

对数据集的训练和测试均在 Pytorch 平台上进行,编程语言为 Python.模型训练时采用 14 G 内存

NVIDIA A16 图形处理器 GPU 加速,具体训练设定如下:训练轮数 epochs 设定为 240,迭代次数 batchsize 设定为 64,学习率  $lr$  设定为 0.001,优化方法为随机梯度下降法(SGD),分类函数为 softmax 函数,损失函数为  $\text{nn.CrossEntropyLoss}()$ (交叉熵损失函数)<sup>[17]</sup>,其运算公式如式(10).

$$\text{Loss} = -\frac{1}{n} \sum_{i=1}^n y_i \left( \ln \frac{1}{1 + e^{-x}} \right) + (1 - y) \ln \left( 1 - \frac{1}{1 + e^{-x}} \right). \quad (10)$$

### 3.2 评价指标

在测试集中根据神经网络的结果,以准确率、精确率、召回率、 $F_1$  值、AUC 值等作为评价指标对网络模型进行评估.这些评价指标均是数值越高,模型的分类效果越好.其中准确率(Accuracy,  $A$ )是指数据样本中被正确识别的个数占总样本的比例,其计算公式如式(11),其中  $P$  表示火灾样本被正确识别的个数,  $N$  表示无火灾样本被正确识别的个数,  $F$  表示无火灾样本被识别为火灾样本的个数,  $T$  表示火灾样本被识别为无火灾样本的个数,

$$A = \frac{P + N}{P + N + F + T}. \quad (11)$$

精确率(Precision,  $P$ )是指火灾样本被正确识别的个数占所有被识别为火灾样本个数的比例,其计算公式为:

$$P = \frac{P}{P + F}. \quad (12)$$

召回率(Recall,  $R$ )是指火灾样本被正确识别的个数占所有实际火灾样本个数的比例,其计算公式为:

$$R = \frac{P}{P + T}. \quad (13)$$

$F_1$  值是精确率和召回率的加权平均值,其值在 0~1 之间,且越大越好,其计算公式为:

$$F_1 = \frac{2 \cdot P \cdot R}{P + R} = \frac{2P}{2P + F + T}. \quad (14)$$

AUC(area under ROC curve)<sup>[18]</sup>值是坐标轴与 ROC 曲线下方围成的面积,取值在 0.5~1.0 之间,且越接近于 1 表示模型分类效果越好.ROC 曲线是反映敏感性和特异性连续变量的综合指标,以伪正类率(false positive rate,  $FPR$ )为横坐标,以真正类率(true positive rate,  $TPR$ )为纵坐标.其中伪正类率是指无火灾样本被识别为火灾样本的个数占所有实际无火灾样本个数的比例,其计算公式为:

$$FPR = \frac{F}{F + N}. \quad (15)$$

真正类率是指火灾样本被识别正确的个数占所有实际火灾样本个数的比例,其计算公式如下:

$$TPR = \frac{P}{P + T}. \quad (16)$$

### 3.3 识别结果与分析

为了验证本研究改进网络模型的可行性,以同样的条件对 Resnet18、Resnet18+HSI、Resnet18+HSI+STN、Resnet18+HSI+STN+TL 这几种网络分别进行 20 组平行实验.并与蒋珍存等<sup>[3]</sup>改进的 VGG16 网络模型、吴彩朋等<sup>[4]</sup>改进的轻量化 SqueezeNet 网络模型以及 JACEK 等<sup>[5]</sup>采用的 AlexNet 网络模型等进行比较.由于每组实验结果较为相似,将其汇总后计算平均值.训练后网络模型各项性能的对比如表 2.

从表 2 中可以得到如下结论:

1)当 Resnet18 网络不加 HSI、STN、TL 时,与文献[3-5]提出的算法进行比较.可以看出:3 种算法对火灾识别的准确率均高于 Resnet18 网络.其中,吴彩朋等<sup>[4]</sup>改进的轻量化 SqueezeNet 网络模型最好,准确率为 90.67%,召回率为 83.94%, $F_1$  值为 0.887, AUC 值为 0.961,分别高出 Resnet18 网络 7.54%,1.21%,0.075,0.056.

2)当 Resnet18 网络使用 HSI 色彩格式火灾数据集时,与使用 RGB 色彩格式相比,准确率、召回率、 $F_1$  值、AUC 值均有所提高,分别提高了 1.2%,5%,0.018,0.02.但与蒋珍存等<sup>[3]</sup>3 种算法相比,各项指标还均偏低.

3)当 Resnet18 网络加入 STN,并且使用 HSI 色彩格式火灾数据集时,各项指标明显高于原始 Resnet18 网络.其中准确率达到 87.30%,略低于蒋珍存等<sup>[3]</sup>3 种算法. $F_1$  值与 AUC 值分别为 0.860,0.941,分别高于 JACEK 等<sup>[5]</sup>算法 0.011,0.01.

4)当 Resnet18 网络加入 STN,使用 HSI 色彩格式火焰数据集,并且使用迁移学习的方法对火灾进行分类识别时.准确率、召回率、 $F_1$  值、AUC 值均达到最高,分别高于 Resnet18 网络 7.89%、6.13%、0.08、0.056. AUC 值与吴彩朋等<sup>[4]</sup>算法相当,均为 0.961,准确率、召回率、 $F_1$  值分别高于吴彩朋等<sup>[4]</sup>算法 0.35%、4.92%、0.005.

表 2 各模型对比

Tab.2 Comparison of models

模型	A/%	P/%	R/%	$F_1$	AUC 值
Resnet18	83.13	79.71	82.73	0.812	0.905
Resnet18+HSI	84.33	78.72	87.73	0.830	0.925
Resnet18+HSI+STN	87.30	88.31	84.85	0.860	0.941
Resnet18+HSI+STN+TL(本文算法)	91.02	94.04	88.86	0.892	0.961
蒋珍存等 <sup>[3]</sup> 算法	88.34	97.83	75.08	0.850	0.945
吴彩朋等 <sup>[4]</sup> 算法	90.67	94.14	83.94	0.887	0.961
JACEK 等 <sup>[5]</sup> 算法	88.07	95.46	76.52	0.849	0.931

根据实验结果绘制图像并进行分析:

Resnet18 网络及改进的 Resnet18 网络各模型测试集准确率随迭代次数的变化趋势如附录图 S1 所示,损失值的变化趋势如附录图 S2 所示,ROC 曲线图如附录图 S3 所示.

由图 S1 可明显看出,从训练 40 轮后,Resnet18 网络加入 STN,使用 HSI 色彩格式火焰数据集,并且使用迁移学习的方法对火灾进行分类识别时,网络模型的准确率明显高于 Resnet18 网络模型及对 Resnet18 模型的其他改进.并且从图像中能明显看出,使用迁移学习后的网络较不使用迁移学习的网络准确率达到稳定的训练轮数明显要更少,说明使用迁移学习更有利于提取火焰特征,降低成本.从附录图 S2 可以看出,几种模型损失值的变化趋势基本保持一致,且最终都能收敛到 0,说明学习率设定为 0.001 是妥当的.从附录图 S3 可以看出,Resnet18+HSI+STN+TL 模型具有更高的 AUC 值,计算得 AUC 值为 0.961,说明此模型对火灾具有更精准的认识能力.

为了进一步验证本文提出的算法可行性,根据实验结果绘制 Resnet18+HSI+STN+TL 模型、蒋珍存等<sup>[3]</sup>、吴彩朋等<sup>[4]</sup>以及 JACEK 等<sup>[5]</sup>采用网络模型图像进行比较.其中各模型测试集准确率随迭代次数的变化趋势如附录图 S4 所示,损失值的变化趋势如附录图 S5 所示,ROC 曲线图如附录图 S6 所示.

由附录图 S4 可明显看出,各模型测试集准确率的变化趋势基本保持一致,从训练 50 轮后,Resnet18+HSI+STN+TL 与吴彩朋等<sup>[4]</sup>算法的准确率明显高于蒋珍存等<sup>[3]</sup>算法与 JACEK 等<sup>[5]</sup>算法的准确率,并且二者的准确率相当.另外,Resnet18+HSI+STN+TL 的准确率变化趋势要比吴彩朋等<sup>[4]</sup>算法更稳定,说明 Resnet18+HSI+STN+TL 具有更强的稳定性,对于火灾识别效果更好.从附录图 S5 可以看出,各模型损失值的变化趋势基本保持一致,且最终都能收敛到 0.从附录图 S6 可以看出,Resnet18+HSI+STN+TL 模型与吴彩朋等<sup>[4]</sup>算法具有相同的 AUC 值,为 0.961,高于其他两种算法,说明 Resnet18+HSI+STN+TL 模型对火灾具有精准的认识能力,能够更有效地发现火灾隐患.

### 3.4 识别错误与原因分析

部分样本类别识别错误的 RGB 格式与对应的 HSI 格式图片如附录图 S7 所示.其中附录图 S7(a)展示的是将有火灾识别为无火灾的部分样本,分析其原因是火焰受到建筑物遮挡、红黄颜色物体、烟雾等因素的干扰.附录图 S7(b)展示的是将无火灾识别为有火灾的部分样本,分析原因是图片样本中有浓烟不能准确辨别或有与火焰的颜色、形状都相似的物体.



## 4 结 语

针对传统深度学习技术的火灾图像检测算法准确率不太高、特征难以提取、网络的平移不变性较弱等问题,基于深度迁移学习设计了一种改进 Resnet18 网络的图像型火灾检测方法,将 Resnet18 网络与空间变换网络相结合,使用迁移学习技术进行训练,得到的火灾识别模型分类准确率为 91.02%。通过实验结果对比,相对于原始 Resnet18 网络性能大幅度提高,与蒋珍存等<sup>[3]</sup>改进的 VGG16 网络模型、吴彩朋等<sup>[4]</sup>改进的轻量化 SqueezeNet 网络模型以及 JACEK 等<sup>[5]</sup>采用的 AlexNet 网络模型图像进行比较,改进后的 Resnet18 网络整体性能也优于这 3 种网络。本研究改进的模型能够更好地通过图像检测技术识别不同场景下的火灾,为改进网络提供了一种新的思考方式,对火灾的预防有更大的参考价值。

在本次研究中发现烟雾对火灾的识别有着密切联系,还不能准确辨识出有浓烟但无火灾发生的情形。另外,当红黄物体与火焰同时出现时,特别是红黄物体占据主要部分时,网络会对火焰的辨识度降低,下一步需要研究出更好的方法来改善烟雾及干扰物体在火灾图像识别中的问题,进一步提高模型的准确率。

附录见电子版(DOI:10.16366/j.cnki.1000-2367.2023.05.19.0001)。

## 参 考 文 献

- [1] GUTMACHER D, HOEFER U, WÖLLENSTEIN J. Gas sensor technologies for fire detection[J]. *Sensors and Actuators B: Chemical*, 2012, 175: 40-45.
- [2] LIU M H, XIE T S, CHENG X, et al. Focused Dropout for convolutional neural network[J]. *Applied Sciences*, 2022, 12(15): 7682.
- [3] 蒋珍存, 温晓静, 董正心, 等. 基于深度学习的 VGG16 图像型火灾探测方法研究[J]. *消防科学与技术*, 2021, 40(3): 375-377.  
JIANG Z C, WEN X J, DONG Z X, et al. Research on fire detection of improved VGG16 image recognition based on deep learning[J]. *Fire Science and Technology*, 2021, 40(3): 375-377.
- [4] 吴彩朋, 王兴鹏. 面向视频火灾火焰识别的轻量化卷积神经网络模型[J]. *安全与环境学报*, 2023, 23(10): 3584-3590.  
WU C P, WANG X P. Lightweight convolutional neural network model for flame recognition in fire video[J]. *Journal of Safety and Environment*, 2023, 23(10): 3584-3590.
- [5] JACEK J, MACIEJ S, MICHAL W. Smoke Detection in a Digital Image with the Use of Convolutional Network[C]// *Conference on Reconnaissance and Electronic Warfare Systems*. [s.l.: s.n.], 2019.
- [6] 陈颖, 张祺, 李文举, 等. 参数合成空间变换网络的遥感图像一致性配准[J]. *中国图象图形学报*, 2021, 26(12): 2964-2980.  
CHEN Y, ZHANG Q, LI W J, et al. Consistent registration of remote sensing images in parametric synthesized spatial transformation network[J]. *Journal of Image and Graphics*, 2021, 26(12): 2964-2980.
- [7] 姜兴家, 刘志云, 宋志敏, 等. 基于迁移学习的船舶机舱火灾图像检测方法[J]. *大连海事大学学报*, 2023, 49(1): 103-109.  
JIANG X J, LIU Y Z, SONG Z M, et al. Image detection method of ship engine room fire based on transfer learning[J]. *Journal of Dalian Maritime University*, 2023, 49(1): 103-109.
- [8] 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*. [s.l.: s.n.], 2016.
- [9] BIN TUFAIL A, ULLAH I, REHMAN A U, et al. On disharmony in batch normalization and dropout methods for early categorization of Alzheimer's disease[J]. *Sustainability*, 2022, 14(22): 14695.
- [10] 陈悦, 杨柳, 李帅, 等. 基于 Softmax 函数增强卷积神经网络—双向长短期记忆网络框架的交通拥堵预测算法[J]. *科学技术与工程*, 2022, 22(29): 12917-12926.  
CHEN Y, YANG L, LI S, et al. Traffic congestion prediction algorithm based on CS-BiLSTM framework[J]. *Science Technology and Engineering*, 2022, 22(29): 12917-12926.
- [11] 蒋昂波, 王维维. ReLU 激活函数优化研究[J]. *传感器与微系统*, 2018, 37(2): 50-52.  
JIANG A B, WANG W W. Research on optimization of ReLU activation function[J]. *Transducer and Microsystem Technologies*, 2018, 37(2): 50-52.
- [12] 陈伟, 王复淞, 郭婧, 等. 基于 Dropout-多尺度空洞卷积神经网络的轴承故障诊断[J]. *机电工程*, 2023, 40(5): 644-654.  
CHEN W, WANG F S, GUO J, et al. Rolling bearing fault diagnosis based on Dropout-multi-scale dilated convolution neural network[J]. *Journal of Mechanical & Electrical Engineering*, 2023, 40(5): 644-654.
- [13] 张祺. 基于空间变换网络的遥感图像配准算法研究[D]. 上海: 上海应用技术大学, 2022.
- [14] MASTYŁO M. Bilinear interpolation theorems and applications[J]. *Journal of Functional Analysis*, 2013, 265(2): 185-207.

- [15] 魏甫豫,张振宇,梁桂珍.基于卷积神经网络下昆虫种类图像识别应用研究[J].河南师范大学学报(自然科学版),2022,50(6):96-105.  
WEI F Y,ZHANG Z Y,LIANG G Z.Research on application of insect species image recognition based on convolutional neural network [J].Journal of Henan Normal University(Natural Science Edition),2022,50(6):96-105.
- [16] SANDULA P,KOLANU H R,OKADE M.CNN-based camera motion classification using HSI color model for compressed videos[J].Signal,Image and Video Processing,2022,16(1):103-110.
- [17] 郑力源.基于深度学习的脑肿瘤图像分类研究与实现[D].成都:电子科技大学,2021.
- [18] KOTTAS M,KUSS O,ZAPF A.A modified Wald interval for the area under the ROC curve(AUC) in diagnostic case-control studies[J].BMC Medical Research Methodology,2014,14:1-9.

## Fire image recognition based on improved Resnet18 network

Chen Kuayue<sup>a</sup>, Wang Baoyun<sup>a,b</sup>

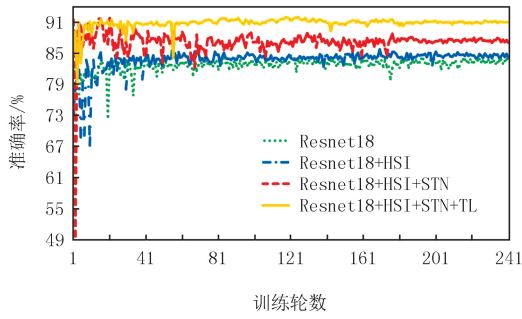
(a. School of Mathematics; b. Yunnan Key Laboratory of Modern Analytical Mathematics and Applications,  
Yunnan Normal University, Kunming 650500, China)

**Abstract:** In view of the problems such as low accuracy, difficult feature extraction and weak translation invariance of the network during fire image recognition by traditional convolutional neural network, this paper improved Resnet18 network to make it have higher performance and accuracy. First, the spatial transformation network(STN) is inserted in front of the convolution layer of the Resnet18 network. For multiple convolution layers of the same size in a row, only the STN is added before the first convolution layer, a total of five are added, and the dropout layer is added after the fully connected layer to prevent overfitting. Then, the transfer learning(TL) method is used to classify and identify fires. Experimental results show that the improved Resnet18 network accuracy rate, recall rate,  $F_1$  value and AUC value are superior to Resnet18 network and other deep learning recognition algorithms, and can quickly and accurately identify fire images.

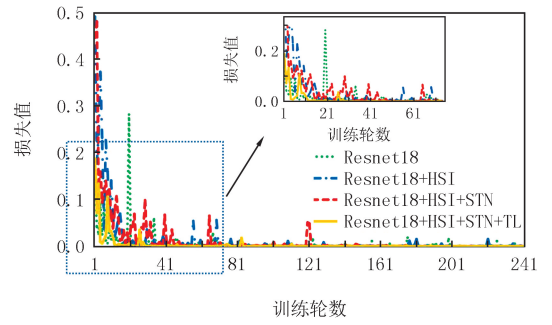
**Keywords:** fire detection; convolutional neural network; spatial transformation network; Resnet18; HSI color model; transfer learning

[责任编辑 陈留院 赵晓华]

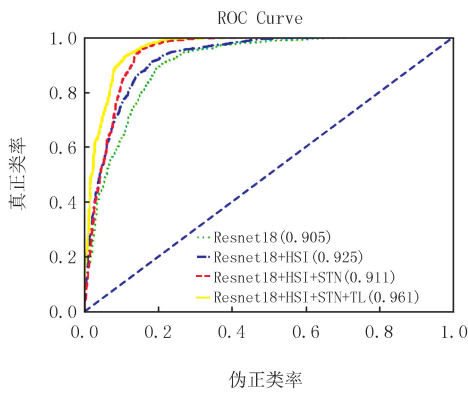
# 附录



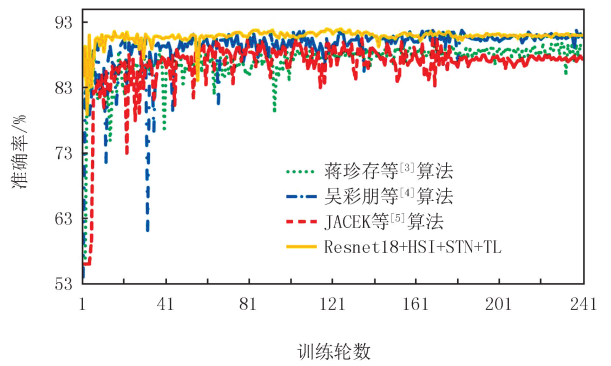
图S1 Resnet18相关网络模型测试集准确率变化曲线  
Fig.S1 Accuracy curve of Resnet18 correlation network model test set



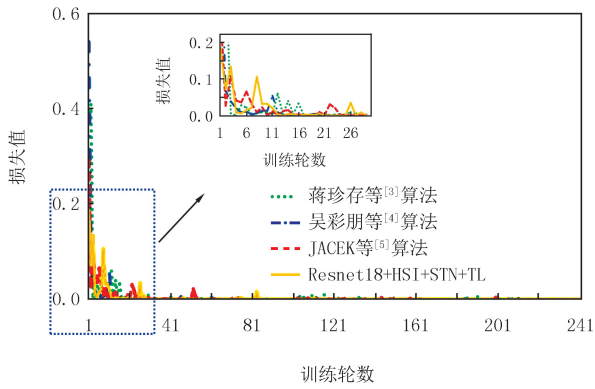
图S2 Resnet18相关网络模型损失值变化曲线  
Fig.S2 Loss value change curve of Resnet18 related network model



图S3 Resnet18相关网络模型ROC曲线图  
Fig.S3 ROC curve of Resnet18 related network model

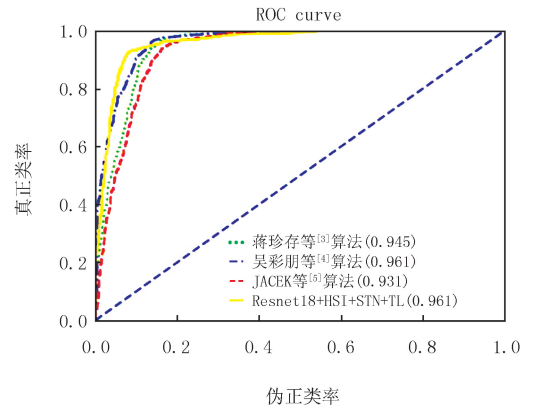


图S4 各模型测试集准确率变化曲线  
Fig.S4 Accuracy curve of each model test set



图S5 各模型损失值变化曲线

Fig.S5 Change curve of loss value of each model



图S6 各模型ROC曲线图

Fig.S6 ROC curve of each model



(a) 将正样本识别为负样本



(b) 将负样本识别为正样本

图S7 模型分类错误部分样本

Fig.S7 Model classification error partial sample