

基于迁移学习的火焰图像识别技术研究^①黎传琛^② 白 勇^③ 陈益民

(海南大学信息科学技术学院 海口 570228)

摘 要 针对火灾视频中复杂背景环境下火焰识别问题,提出了一种基于迁移学习的火焰图像智能识别方法。采用以深度学习预训练的模型为基础并经过迁移学习作为特征提取器提出特征,而后进行特征融合并结合传统机器学习分类器方法进行识别的流程。在所提出的流程中采用逻辑回归和 Xgboost 两种机器学习算法作为最终分类器进行了实验,结果表明识别准确率得到较大的提升。最终识别时只需要输入原始图片,就能够自动得出识别结果,在火灾视频中复杂环境下的火焰识别方面取得了很好的效果。

关键词 火焰图像识别,深度学习,迁移学习,机器学习,特征提取

0 引 言

对火灾的监控,一直是一个被广泛关注的问题。由于火灾会造成很大的人员伤亡与财产损失,而由于所有的地方都由人工监控并不现实,基于视频自动分析的方法成为了一个非常自然的选择。视频自动分析拥有相对于传统人工监控无可比拟的优势。

计算机视觉领域与图像处理领域对火焰识别的研究近年来取得了长足发展。文献[1]通过图像处理的方式把视频图像处理后得到需要的特征(如背景动态变化,R,G,B颜色),再设定一个阈值进行判断。文献[2]利用光流图残差来做纹理分割,同时可以检测动态的背景,改进了传统方法中背景需要固定的要求。利用传统机器学习算法进行判决可带来相比传统方法较大的性能提升。文献[3]利用支持向量机来进行二分类。文献[4]对火焰像素的判别就是先通过传统的计算机视觉对图像处理,而后对比了逻辑(logistic)回归方法与最大投票方法。近年来基于深度学习的卷积神经网络在 Imagenet 数据集上带来了令人印象深刻的多分类识别率,学得

的模型具有很好的表示学习能力与泛化能力。文献[5]提出在输入神经网络前先对火焰图片进行预处理,把原图先转换成 HSV 域图,分割出高能域,然后通过做帧间差分获得动态变化区域。这样进行处理后再输入神经网络能够减少神经网络学习数据。文献[6]在火灾识别中使用卷积神经网络,且在卷积神经网络提出的最后一层的特征图上进行滑动搜索框搜索来识别火和烟,识别速度非常快,实现了仅需输入原图而不需进行预处理。文献[7]利用深度卷积神经网络识别烟雾。文献[8]利用 3D 卷积提取时间特征与 2D 卷积提取空间特征结合进行动态场景分类。文献[9]在网络的池化层后加入了一个归一化层,得到更好的烟雾识别。文献[10]先使用卷积神经网络判断是否有火,然后在第五层池化进行上采样,最终可以得到火焰位置。

火焰识别研究领域没有业界公认的火焰图片数据集,目前研究中火焰识别性能统计使用的图片数偏少。因此,本文自行搭建了火焰识别数据集,并对数据集进行预处理来增强数据集。使用了传统的机器学习方法进行实验,采取了深度学习经迁移学习后的模型。本文首先考虑使用现有迁移学习方法建

① 国家自然科学基金(61561017)资助项目。

② 男,1993年生,硕士生;研究方向:机器学习,图像处理等;E-mail: 625559571@qq.com

③ 通信作者,E-mail: bai@hainu.edu.cn

(收稿日期:2018-07-04)

立火焰识别模型;为进一步提高识别性能,又提出了基于迁移学习提取火焰特征;然后把多个迁移学习模型提取到的特征进行特征融合,最后再结合分类器进行火焰识别。实验结果表明,特征融合后使用分类器可取得更高的识别准确率与受试者工作特征曲线(receiver oprating characteristic cure, ROC)。

1 数据预处理与传统分类方法

本文中的火焰图片数据集来自 Imagenet 与百度、Google 上的火焰图片,以及 Youtube 视频截取的帧。由于视频截取出来的帧,图片的相似度太高,且每个视频间的差异非常大,有不同光照环境与复杂场景,因此对视频帧的识别比对独立的火焰图片识别更加困难。本文把视频拆分成的帧作为主要的测试集。测试集与训练集均自己标注标签。数据集如表 1 所示。标签 flame 为正样本,代表有火焰的图片;而标签 others 则为负样本,主要选取了一些可能出现火灾的背景如房子、车、树木,可能误判的光照、枫叶等。其余的则非常分散随机选取。

表 1 火焰图片数据集组成

训练集	值
训练集火焰图片	1643
训练集数据增强火焰帧图片	3563
训练集负样本	5144
测试集火焰图片	192
测试集火灾视频帧	1437
测试集负样本	1611

由于负样本数量比较大,本文中数据集仅对火焰正样本使用了数据增强。数据增强技术即以保持标签相同的方式改变数组表示方式来改变训练数据的方法。本文采用了旋转、剪裁、缩放、翻转、图片合成等数据增强方法。使用的数据增强技术具体如下:

(1) 图片合成。使用图像处理(PIL)模块把容易出现的背景图如车、房子、树与切割出来的火焰图片进行融合,合成后标注为 flame,而背景图表示为 others,可以更好地让神经网络学习到什么是背景,

什么是需要学习的目标。

(2) 输入归一化。由于缺少足够的训练集,在输入前进行了归一化处理来避免欠拟合。归一化后的值学习起来的速度也更快。

(3) 随机剪裁、缩放、翻转后把其中有火焰的图像保存下来扩大数据集。

(4) 图片后续根据输入模型的不同,变换成相应的尺寸。

(5) 图片根据转换到 HSV 颜色区域,通过设定相应阈值取得高能量区域。

(6) 图片通过主成分分析(principal component analysis, PCA)降维,减少了大量占总方差比例小的成分,以适应于传统机器学习分类方法。

使用传统分类方法进行图像分类时,一般把图片的每个像素当作一个特征,输入进分类器前需要进行展平,展平后数据维度非常高,轻易能达到万维,甚至十万维。减少输入进分类器的数据非常重要。在火焰检测中进行处理,如颜色空间变化后,进行颜色直方图提取,或者通过掩膜操作得到感兴趣区域,用背景残差得出运动区域,并进行纹理特征提取,并进行 PCA 降维,这些方法都可以减少整体数据。进行多种方式进行特征提取后合并输入分类器,让分类器得到最后的分类结果。

2 迁移学习

2.1 迁移学习

迁移学习就是把已训练好的模型参数迁移到新的模型来帮助新模型训练。考虑到大部分数据和任务存在相关性,迁移学习可以将已经学到的模型参数通过某种方式来分享给新模型从而加快并优化模型的学习效率。

在数据集不够且测试集与训练集分布并不是特别相似的情况下,如果使用深层的网络结构去进行训练,由于深层网络的表示学习能力太强,非常容易造成过拟合。而不使用深层的网络结构,识别率又不能满足实际需求。此时使用迁移学习则是一个非常好的思路。把一个预先训练好的模型部分挪移到另外一个任务上使用。对于图像识别问题来说一般

是只需要训练最后的几层,可以大量减少训练时间,且识别率也较高。对于火灾视频帧识别,单独的火灾图像数量比较少,因此选择使用迁移学习来实现火灾视频帧的识别。

本文的迁移学习模型选择了 Xception^[11] 与 InceptionV3^[12] 深度学习模型作为基础模型。这两个模型已有在 Imagenet 上训练的参数可用。Inception 模块组的目的是将空间结构简化,同时把空间信息转化为高阶抽象的特征信息,输出的特征向量维度为 2 048。Xception 则是分别按不同通道进行一次卷积(生成输入通道数张特征图)再将这些特征图

一起进行第二次卷积。输出的特征向量维度也是 2 048。

2.2 基于现有迁移学习模型的火焰识别方法

由于每次训练一个模型需要较长时间,且每个迁移模型提出的特征会有些许区别,本文采用了并行训练多个现有迁移模型的方式。由于 InceptionV3 与 Xception 推荐的输入图片尺寸都是 299,则图片尺寸只要变换为 299 就可以同时输入两个模型。因此选取了 InceptionV3 与 Xception 作为同时并行训练的模型。搭建的火焰识别流程图如图 1 所示。

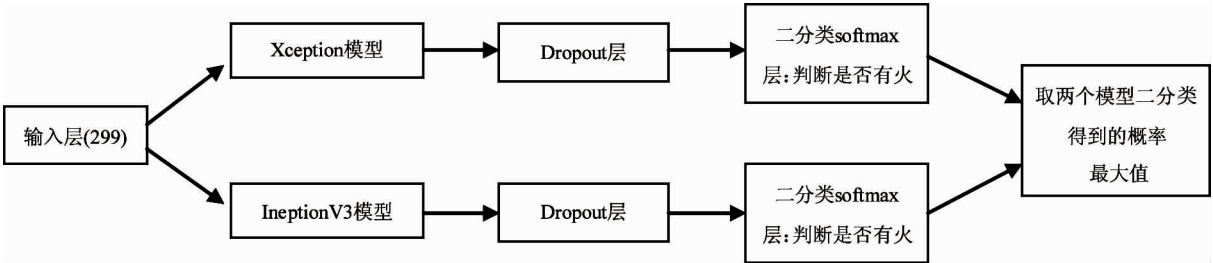


图 1 基于现有迁移学习模型的火焰识别流程图

如图 1 所示,输入层维度为 299,同时输入两个不同模型 InceptionV3 与 Xception 进行并行训练,分别设置全连接层。输出是否有火焰的分类结果概率。综合两个模型判断的结果概率的输出最大值,最终可以看作进行了一种投票来判断是否有火焰,避免了单一模型针对某些图片的盲点。

训练时保留了迁移模型在 Imagenet 上训练得到的所有结构与参数。冻结卷积层,包含 top 层,进行全连接层的训练。训练集使用上节介绍的通过数据增强后的训练集,测试集则不使用数据增强。由于网络结构非常深,使用了学习速率自动衰减,防止测试集损失摆动幅度过大,并设置了早停止方法来防止过拟合;同时使用了 nadam 优化器,并行进行参数更新。

3 基于迁移学习模型特征提取的火焰识别方法

3.1 迁移学习模型提取特征

迁移学习在 Imagenet 使用大量图片训练因而具有强大的特征提取能力,能够提取出高维特征。

相比之下,人为条件下可能只能够提取到几十上百种的特征,且传统方法中把每个像素当作特征点输入,维度非常高。深度模型中有大量的卷积核,可以看作是多种特征提取器。如 ResNet50 单独一个模型可以得到 2 048 维输出,可以看成 2 048 维的特征。

整个迁移学习模型可当成是特征提取器,多模型提取出的特征融合后再使用机器学习方法进行分类判断。在迁移模型进入全连接层之前,选择不包括 top 层而直接输出特征向量,对多个模型进行了测试并最终选择使用 InceptionV3、Xception、ResNet50^[13] 这 3 个模型。由于 InceptionV3 与 Xception 模型比较类似而加入了 ResNet50 以弥补前两种模型提取不到的特征。有时太深层的神经网络在训练集准确率上反而下降而不如浅层神经网络,而人为让神经网络某些层跳过下一层神经元的链接,隔层相连,以弱化每层关系。

数据集直接使用全部测试集与训练集,并不需要做预处理。由于需要特征提取器得出的特征进行对比,则训练集与测试集在开始时就需要先使用多

个深度学习模型提取出特征,提取特征后搭建一层全连接层对每一个模型提出的特征单独进行一次逻辑回归训练,使用 sigmoid 激活函数,可以选择在此测试中比较优秀的模型。InceptionV3 和 Xception

模型抽取的特征单独判断具有比较高的准确率。而 InceptionResNetV2 识别率非常低,另外的如 VGG 虽然识别率不低,但是并不优秀。整体的流程如图 2 所示。

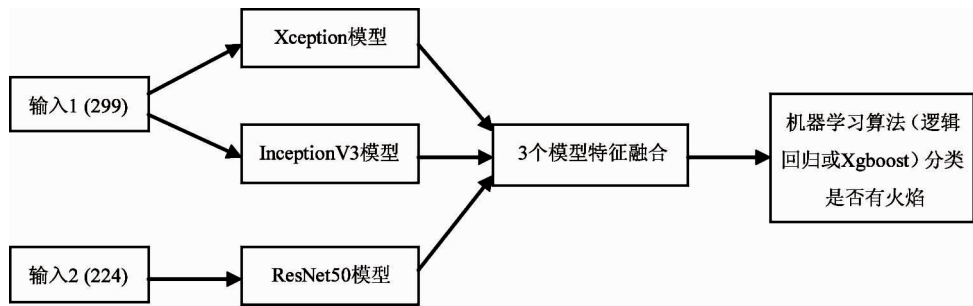


图 2 迁移学习结合机器学习分类器的火焰识别流程图

迁移模型得到的特征是一个 numpy 数据,可使用 numpy 中的 concatenate 函数进行拼接,设置拼接方向为 axis = 1 方向,进行 3 个模型的特征融合。由于 3 个模型提出的特征维度都是 2 048,拼接后得到 6 144 维特征。特征融合扩大了特征数。

3.2 基于特征融合的机器学习分类

选择好 3 个深度模型后,抽取特征进行融合,并使用传统机器学习算法分类器。分类器选择逻辑回归或 Xgboost 算法。后者是一种集成学习 boosting 方法,是一族可将弱学习器提升为强学习器的方法。

3.2.1 逻辑回归模型

搭建一层全连接层:在输入全连接层前需要把抽取出来的特征做融合处理,处理后维度是(, 6 144),维度较高。设置了丢弃层(丢弃率为 0.3)防止过拟合,并且设置好自动学习率衰减防止测试集 loss 的大幅度摆动。逻辑回归的 sigmoid 函数公式如下:

$$y = \frac{1}{1 + e^{-(\omega^T x + b)}} \quad (1)$$

式(1)中,函数的定义域为全体实数,值域在 [0,1] 之间,x 轴在 0 点对应的结果为 0.5。当 x 取值足够大的时候,可以看成 0 或 1 类问题,大于 0.5 可以认为是 1 类问题,反之是 0 类问题,而刚好是 0.5,则可以划分至 0 类或 1 类。通过 sigmoid 函数可以得到每个测试样本的概率值,可与标签比对优化损失函数。

3.2.2 Xgboost 模型

Xgboost 算法实际上是一种集成树。优势则在于显式地将树模型的复杂度作为正则项加在优化目标上。Xgboost 借鉴了随机森林的做法,支持列抽样,不仅能降低过拟合,还能减少计算。对于特征的值有缺失的样本,Xgboost 可以自动学习出它的分裂方向,实现了一种分裂节点寻找的近似算法。Xgboost 的损失函数如下:

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2 \quad (2)$$

$$L(\mathcal{O}) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (3)$$

式(2)中 $\Omega(f)$ 是树复杂度项,由叶子数量与 $L2$ 正则组成; γ 和 λ 是 Xgboost 自己定义的,读者可以设定它们的值,显然, γ 越大,表示越希望获得结构简单的树,因为此时对较多叶子节点的树的惩罚越大。 λ 越大也是越希望获得结构简单的树, ω 则为与真实值的残差。式(3)中 i 表示样本编号, k 表示树的棵树,由于 loss 和复杂度项都是凸函数,所以有最小值。通过上式得到最优化解后,代表指定了一个树结构,对每一次尝试对已有叶子加入一个分割。树节点分裂公式如下:

$$Gain = \frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right] - \gamma \quad (4)$$

式(4)中,括号中第 1 项为左子树分数,第 2 项为右子树分数,第 3 项为不分割可以得到的分数。 γ 为加入新叶子节点引入的复杂度。此参数可以在建

树过程中动态选择是否要加入一个节点。*Gain* 如果是正的,并且值越大,就越值得切分。*Xgboost* 的输入维度为 $(m, 6144)$, m 为样本数,把抽取的特征融合后直接输入分类器。

Xgboost 参数选择如表 2 所示。

表 3 *Xgboost* 参数设置

参数	值
最大深度	10
学习率	0.85
损失函数	binary_logistic
列采样	0.7
样本随机采样	0.85
Min_child_weight	1
Early_stopping	30

集成树模型由于输出的是得分,分类时评价指标一般是选择 ROC 曲线,即受试者工作特征曲线。纵坐标为真正例率 (*TPR*),横坐标为假正例率 (*FPR*),其计算公式如下:

$$TPR = \frac{TP}{TP + FN} \tag{5}$$

$$FPR = \frac{FP}{TN + FP} \tag{6}$$

分类结果混淆矩阵如表 3 所示。

表 3 分类结果混淆矩阵

真实情况	预测结果	
	正例	反例
正例	<i>TP</i>	<i>FN</i>
反例	<i>FP</i>	<i>TN</i>

4 实验与结果分析

实验使用 *Keras* 平台,采用基于 *Imagenet* 训练得到的迁移学习模型,使用本文作者自己搭建的数据集进行训练与测试,调用 *GPU* 进行训练。

首先进行传统机器学习分类方法的实验,分类器采用了支持向量机 (support vector machine, *SVM*) 与 *Lightgbm* 两种。*SVM* 使用了高斯核, *C* 值设置为 200,最终结果能够收敛。*Lightgbm* 使用了 *gdbt* 树

作为弱分类器,设置调整好参数。而对图像的处理方法则分为:(1)训练集、测试集归一化;(2)把原图转换到 *HSV* 颜色空间。通过实验设定不同的取值范围,使用 *opencv* 的掩膜操作把高能量区域作为感兴趣区域,掩盖掉非高能区域,实际上也减少了大量的数据;(3)对图像进行 *PCA* 降维,选取降维后各主成分的方差值占总体方差 0.99 的主成分。实验结果的 ROC 曲线如图 3 ~ 图 5 所示。

如图所示,*SVM* 标注的虚线代表了使用 *SVM* 分类器,*Lgb* 标注的实线代表了使用 *Lightgbm* 分类器,*luck* 标注的圆形点连线表示随机猜测。实验对比采用了 ROC 曲线。其中由于 *SVM* 输出的是 0,1 标签判断值,而 *Lightgbm* 输出的结果是得分,所以 *SVM* 的 ROC 曲线呈折线状,而 *Lightgbm* 则是不

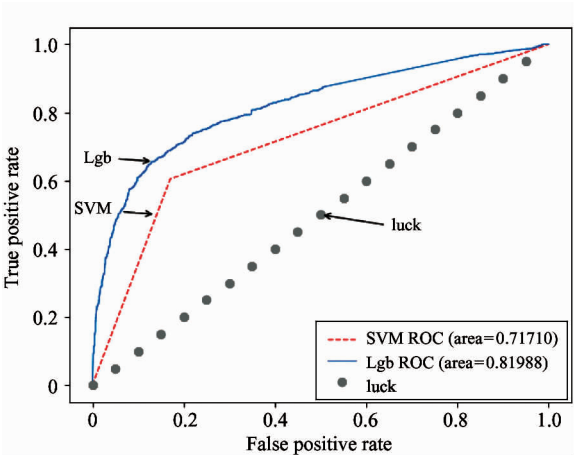


图 3 感兴趣区域提取后输入分类器

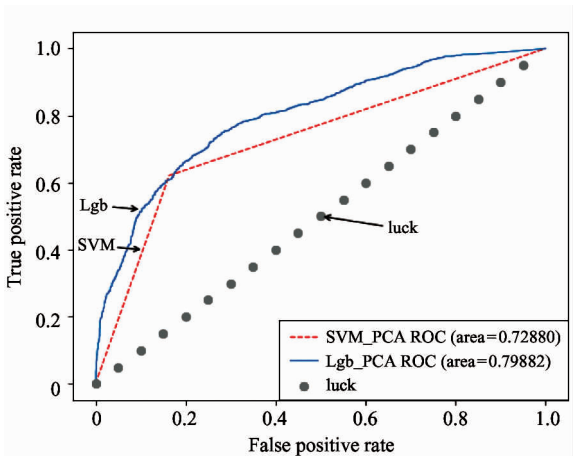


图 4 感兴趣区域提取后 PCA 降维输入分类器

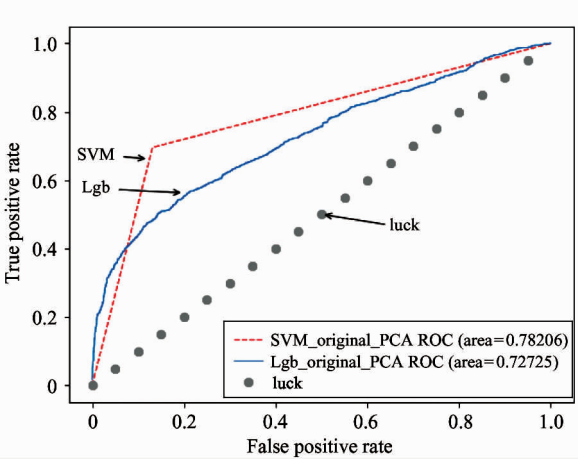


图5 原图 PCA 降维后输入分类器

规则的曲线状。对比图 3 与图 4,可以看出在同样经过感兴趣区域提取后,进行了 PCA 降维,对于 Lightgbm 分类器来说 ROC 曲线下方与坐标轴围成的面积大小(AUC)值下降幅度不大,仅为 0.02 左右,而对于 SVM 分类器, AUC 值反而上升了 0.01 左右。对比图 4 与图 5,可以看出通过对原图进行 PCA 降维,与对提出感兴趣区域后进行 PCA 降维,对 SVM 和 Lightgbm 的影响非常不同。SVM 在使用原图 PCA 作为输入时明显比使用感兴趣区域 PCA 作为输入时得到的 AUC 更接近 1,高了近 0.06。而 Lightgbm 的 AUC 下降了近 0.07。对比图 3 与图 5,可以看出两种减少输入分类器数据的方法对两种分类器影响不同,对 SVM 来说感兴趣区域提取方法并不是很好的方法,通过图 5 可以看出原始图片使用了 PCA 降维后输入 SVM 比原始图片直接输入 SVM,

得到了更大的 AUC 值;原始图片使用 PCA 降维后输入 Lightgbm 比原始图片直接输入 Lightgbm, AUC 值反而下降了。

实验的结果显示出对于 Lightgbm 这种分类器来说,特征的维数非常重要,对分类结果影响很大。原图特征维度为 49 152,使用 PCA 降维后,维度降低到了 4 839;而对原图进行感兴趣区域提取后,把原图的特征维度降低到 5 215。通过对比图 3 与图 5,可以看出输入的维度越高,Lightgbm 就能够得到越好的识别效果。

对于 SVM 来说,使用 PCA 降维的效果更好,在太高维度情况下,SVM 得到的结果并不好,则更加适合维度不太高的情况。而减少了原图数据的感兴趣区域提取方式对于 SVM 性能也影响较大。

综上所述,原图由于维度太高,几乎不能直接输入分类器,而进行降维或者感兴趣区域提取,人为减少了输入分类器的数据量,会带来数据损失。总体的 AUC 值,最好也仅仅超过 0.81,不能满足真正的实际需求,且训练速度较慢。

而后对 VGG16 神经网络结构进行实验,将原图转换到 HSV 颜色空间后,通过实验设定不同的取值范围,使用 opencv 的掩膜操作把高能量区域作为感兴趣区域,掩盖非高能区域,使得模型学习到的主要是感兴趣的火焰区域。实验设置了早停止与学习率衰减,通过实验选取结果最好的阈值后得到的结果如图 6 所示。

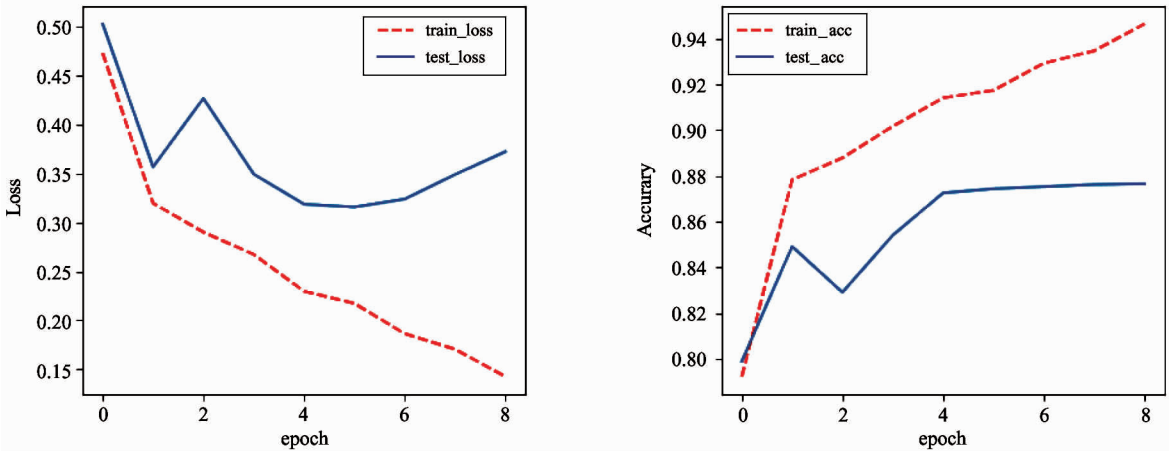


图6 基于 VGG16 结构的火焰识别的损失率与准确率

如图 6 所示,虚线代表的是训练集,实线是测试集。由于提出了感兴趣区域,数据量比原图更少,训练速度得到加快。在 3 到 4 个 epoch 就已经得到很好的结果。整体结果图也比较平稳,最终测试集准确率收敛到了 87% 左右,在 4 到 6 个 epoch 时已经达到了损失最低点,处于 0.33 附近。在 6 个 epoch 后测试集损失已经开始上升,而训练集损失依然在下降,已经开始有过拟合趋势,由于设置了早停止,在 8 epoch 时候整个训练停止。最终 87% 左右的测试集准确率并不能很好地满足实际的需求。

上图显示出,虽然选取阈值来提出高能量区域

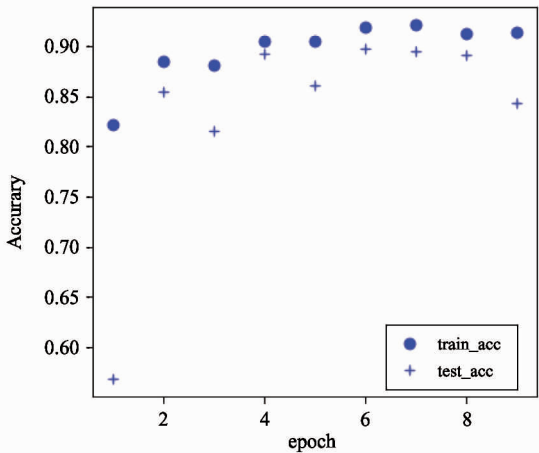
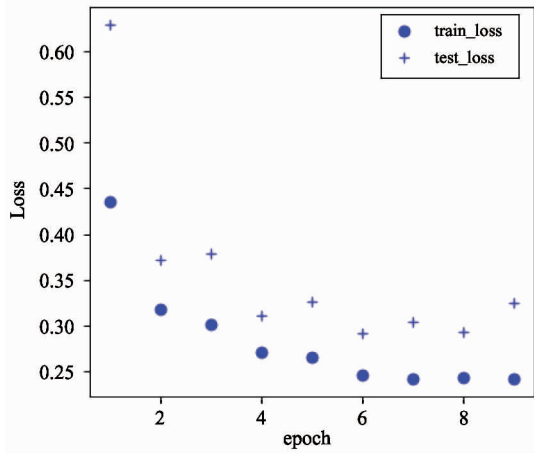


图 7 基于现有迁移学习模型的火焰识别的损失率与准确率

图 7 中圆点表示的是训练集,十字点表示测试集。横坐标是迭代周期 epoch,左图纵坐标是损失率,右图纵坐标为准确率。可知使用迁移学习得到测试集不到 90% 的识别率,损失也在 0.3 以上。且图中显示出损失在第 8 次迭代后就开始上升,设置了早停止,如果不尽早停止,继续再训练损失会继续上升。识别率下降,而不到 90% 的识别率显然还不能够满足实际要求。

图 7 反映出虽然迁移学习模型在 Imagenet 上基于大量图片进行了训练,得到了一个对各种图像较好的识别模型,然而对某种具体任务,如本文中火灾图像的识别,模型中的参数不能够完美拟合。

然后对基于迁移学习结合机器学习分类器的火焰识别分类方法进行实验。这里使用迁移模型 In-

ceptionV3, Xception, ResNet50 提取特征后进行特征融合采用逻辑回归来分类结果,如图 8 所示。

图 8 中虚线表示训练集,实线表示测试集,横坐标为迭代周期 epoch。左图纵坐标为损失率,右图纵坐标为准确率。可知使用了多迁移学习模型提取特征,进行特征融合后使用逻辑回归判断的准确收敛到了 96.3% 左右,损失率也降低到 0.10 左右。因此,所提出的方法具有很好的火焰识别性能。

特征融合后使用 Xgboost 作为分类器,ROC 曲线结果如图 9 所示。

图 9 中横坐标为假正例率 (FPR),纵坐标为真正例率 (TPR),得到的 ROC 曲线围成的面积为 0.9845。可知使用 Xgboost 得到了比较接近于 1 的 AUC 值,显示了很好的火焰识别性能。

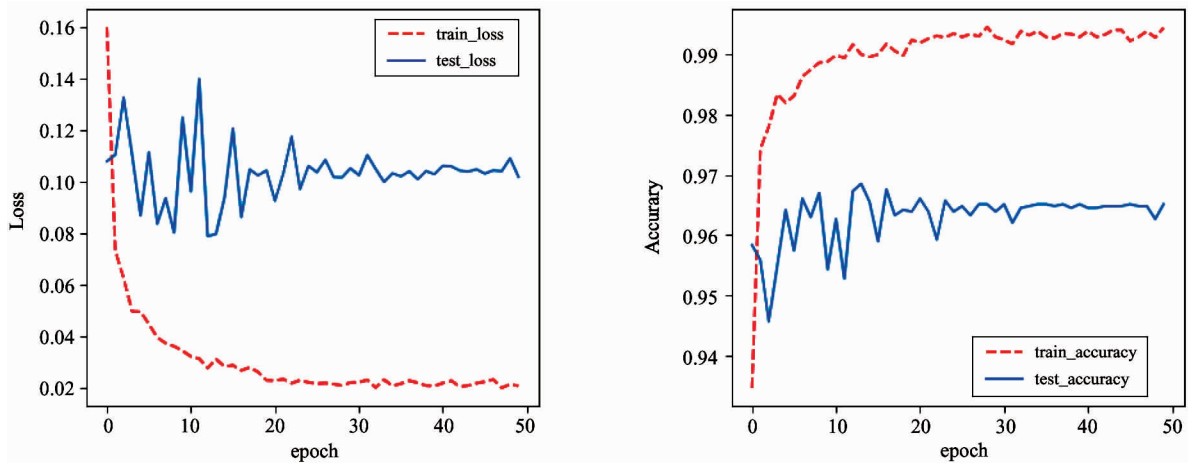


图 8 特征融合后使用逻辑回归算法

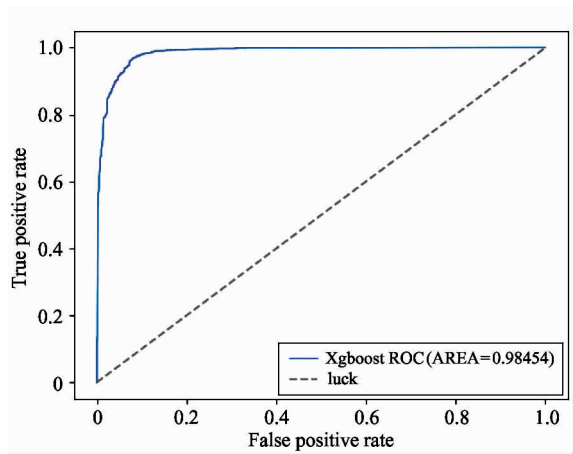


图 9 Xgboost ROC 曲线图

5 结 论

本文使用了图像处理、数据增强与迁移学习 3 种模型结合的方法,采用了把 3 种深度学习预训练并经过迁移学习后的模型作为特征提取器提取火焰特征并进行特征融合,然后再使用传统机器学习分类算法(逻辑回归或 Xgboost)进行分类判断。实验表明:基于现有迁移学习框架火焰识别方法获得了低于 90% 的分类准确率,还不足以使用在火焰分类的任务上。而通过迁移学习进行特征提取并进行融合后,再使用逻辑回归分类的方式得到了 96% 以上的识别率,并且使用 Xgboost 算法也达到了接近 0.985 的 AUC。实验表明本文提出的方法对视频帧中复杂场景下火焰识别有非常好的识别率与鲁棒性。

参考文献

- [1] ÇetinA E, Dimitropoulos K, Gouverneur B, et al. Video fire detection-review [J]. *Digital Signal Processing*, 2013,23(6):1827-1843
- [2] Chetverikov D, Fazekas S, Haindl M. Dynamic texture as foreground and background[J]. *Machine Vision and Applications*, 2011, 22(5):741-750
- [3] 张楠. 基于视频图像的火灾检测与识别方法研究 [D]. 广州:华南理工大学,2013
- [4] Toulouse T, Rossi L, Celik T, et al. Automatic fire pixel detection using image processing: a comparative analysis of rule-based and machine learning-based methods[J]. *Signal Image & Video Processing*, 2016, 10(4):647-654
- [5] Maksymiv O, Rak T, Menshikova O. Deep convolutional network for detecting probable emergency situations[C]. In: Proceedings of the IEEE 1st International Conference on Data Stream Mining & Processing, Lviv, Ukraine, 2016. 23-27
- [6] Frizzi1 S, Kaabi R, Bouchouicha M, et al. Convolutional neural network for video fire and smoke detection[C]. In: Proceedings of the 42nd Annual Conference of the IEEE Industrial Electronics Society, Florence, Italy, 2016. 877-882
- [7] Tao C, Zhang J, Wang P. Smoke detection based on deep convolutional neural networks[C]. In: Proceedings of the 2016 International Conference on Industrial Informatics-Computing Technology, Intelligent Technology, Industrial Information Integration, Wuhan, China, 2017. 150-153

[8] Zheng J, Cao X, Zhang B, et al. Bi-heterogeneous convolutional neural network for UAV-based dynamic scene classification[C]. In: Proceedings of the Integrated Communications, Navigation & Surveillance Conference, Herndon, USA, 2017. 1-28

[9] Yin Z, Wan B, Yuan F, et al. A deep normalization and convolutional neural network for image smoke detection [J]. *IEEE Access*, 2017. 5(99) :18429-18438

[10] Zhang Q, Xu J, Xu L, et al. Deep convolutional neural networks for forest fire detection[C]. In: Proceedings of the International Forum on Management, Education and Information Technology Application, Guangzhou, China, 2016. 568-575

[11] Chollet F. Xception: deep learning with depthwise separable convolutions[C]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017. 1800-1807

[12] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the Inception Architecture for Computer Vision[J]. *Computer Science*, 2015. 2818-2826

[13] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016. 770-778

Research on flame image identification based on transfer learning

Li Chuanchen, Bai Yong, Chen Yimin
(College of Information Science and Technology, Hainan University, Haikou 570228)

Abstract

Aiming at the problem of flame identification under the complex background environment of fire video, a method that employs transfer learning on pre-trained deep learning models to identify the flame in images is proposed. Transfer learning models are used as feature extractors to extract the features in image, and the extracted image features are fused as input to traditional machine learning classifier to identify flame. In the proposed procedure, two machine learning classifiers, logistic regression and Xgboost, are used for experiments. The experimental results show that the flame identification accuracy is improved significantly. In the proposed method, the original images are inputed only, and then it can automatically generate identification results. The proposed method can achieve good results in flame identification under the complex background environment of fire video.

Key works: flame image identification, deep learning, transfer learning, machine learning, feature extract