

基于对比学习和伪异常合成的无监督火灾检测^①



叶伟华¹, 吴云涛², 李佐勇^{3,4}

¹(福州软件园科技创新发展有限公司, 福州 350101)

²(福建理工大学 计算机科学与数学学院, 福州 350118)

³(闽江学院 计算机与控制工程学院, 福州 350121)

⁴(福建省信息处理与智能控制重点实验室, 福州 350121)

通信作者: 叶伟华, E-mail: ywh@fzrjy.com

摘要: 传统的火灾检测方法大多基于目标检测技术, 存在火灾样本获取难度高、人工标注成本高的问题. 为解决该问题, 本研究提出了一种基于对比学习和伪异常合成的无监督火灾检测模型. 为了实现无监督图像特征学习, 提出了交叉输入对比学习模块. 然后, 引入了一个记忆原型学习正常场景图像的特征分布, 通过特征重建实现对火灾场景的判别. 并且, 提出了伪异常火灾场景合成方法和基于欧氏距离的异常特征区分损失, 使模型对于火灾场景具有针对性. 根据实验表明, 我们的方法在 Fire-Flame-Dataset 和 Fire-Detection-Image-Dataset 两个公开火灾检测数据集上的图像级 AUC 分别达到 89.86% 和 89.56%, 优于 PatchCore、PANDA、Mean-Shift 等主流图像异常检测算法.

关键词: 火灾检测; 异常检测; 对比学习; 记忆机制; 无监督学习

引用格式: 叶伟华, 吴云涛, 李佐勇. 基于对比学习和伪异常合成的无监督火灾检测. 计算机系统应用, 2024, 33(6): 28–36. <http://www.c-s-a.org.cn/1003-3254/9529.html>

Unsupervised Fire Detection Based on Contrastive Learning and Synthetic Pseudo Anomalies

YE Wei-Hua¹, WU Yun-Tao², LI Zuo-Yong^{3,4}

¹(Fuzhou Software Park Technology Innovation Development Co. Ltd., Fuzhou 350101, China)

²(College of Computer Science and Mathematics, Fujian University of Technology, Fuzhou 350118, China)

³(College of Computer and Control Engineering, Minjiang University, Fuzhou 350121, China)

⁴(Fujian Provincial Key Laboratory of Information Processing and Intelligent Control, Fuzhou 350121, China)

Abstract: Traditional fire detection methods are mostly based on object detection techniques, which suffer from difficulties in acquiring fire samples and high manual annotation costs. To address this issue, this study proposes an unsupervised fire detection model based on contrastive learning and synthetic pseudo anomalies. A cross-input contrastive learning module is proposed for achieving unsupervised image feature learning. Then, a memory prototype that learns the feature distribution of normal scene images to discriminate fire scenes through feature reconstruction is introduced. Moreover, a method for synthesizing pseudo anomaly fire scenes and an anomaly feature discrimination loss based on Euclidean distance are proposed, making the model more targeted toward fire scenes. Experimental results demonstrate that the proposed method achieves an image-level AUC of 89.86% and 89.56% on the publicly available Fire-Flame-Dataset and Fire-Detection-Image-Dataset, respectively, surpassing mainstream image anomaly detection algorithms such as PatchCore, PANDA, and Mean-Shift.

Key words: fire detection; anomaly detection; comparative learning; memory mechanism; unsupervised learning

① 基金项目: 国家自然科学基金 (61972187)

收稿时间: 2023-12-14; 修改时间: 2024-01-17; 采用时间: 2024-02-01; csa 在线出版时间: 2024-04-30

CNKI 网络首发时间: 2024-05-07

1 引言

火灾是威胁公众安全和社会发展的最经常、最普遍的灾害之一。据不完全统计,2023年中国发生的火灾数量约为15.2万起,造成直接经济损失约500亿。防范火灾一直是消防工作的难点和重点,越早地发现火灾则能越早采取行动控制火势,减少人员伤亡和财产损失。因此,及时地发现火情十分重要。近年来,随着深度学习和计算机视觉技术的发展,利用深度学习技术进行火灾检测已经成为一大研究热点^[1]。

当前,大多数研究使用目标检测算法进行火灾检测^[2-9],它们通常基于YOLO^[10]或Faster R-CNN^[11]。这类方法对数据的要求较高,需要大量的火灾图像数据,并且需要规范的人工标注信息。但是,在真实工业场景和社区场景中,由于火灾的高危险性,普遍存在着真实火灾数据少、获取数据困难的问题。因此,还有少数研究采用无监督的方式训练火灾检测模型^[12-16],以追求较低的数据获取成本,但由于缺乏监督信息,检测精度往往较低。

为了应对当前存在的问题,我们提出了基于对比学习和伪异常合成的无监督火灾检测方法,实现了低数据成本且高精度的火灾检测。首先,我们的方法将火灾检测问题视为一项无监督图像异常检测问题,具有无监督方法训练成本低的优点。为了充分利用无监督信息的图像数据训练火灾检测模型,我们利用对比学习技术训练编码器学习正常图像的实例级判别特征。并且,受到Ning等人^[17]的启发,我们引入了可学习的记忆原型来重建图像特征,从而实现火灾场景和正常场景的判别。其次,为了使模型对于火灾场景具有针对性,同时缓解仅使用正常数据训练模型导致部分样本检测置信度较低的问题,我们提出了伪异常火灾场景合成方法,通过正常样本合成火灾场景样本,并提出了基于欧氏距离的异常特征区分损失,在训练过程中充分利用合成的火灾场景样本提升模型性能。

本研究的主要贡献如下:(1)我们将火灾检测视为一项无监督异常检测任务,有效避免了在真实工业场景和社区场景中,由于火灾的高危险性,真实火灾数据少、获取数据困难的问题。(2)提出了一个交叉输入对比学习模块,利用交叉输入的方式提升对比学习的性能,同时设计了一个记忆原型和特征重建模块来学习正常样本的特征分布,并通过重建的方式判别火灾场景图像。(3)提出了一种伪异常火灾场景合成方法,同

时提出了基于欧氏距离的异常特征区分损失来充分利用伪异常图像,使模型对于火灾场景具有针对性,并缓解了仅使用正常数据训练模型导致部分样本检测置信度较低的问题。

我们在两个公共火灾检测数据集 Fire-Flame-Dataset 和 Fire-Detection-Image-Dataset 上开展了广泛的实验,通过与多个相关模型进行对比,本文所提出的模型取得了最优的检测性能。

2 相关工作

基于深度学习的火灾检测技术主要通过监控摄像头设备获取的视觉信息来检测一定区域内是否存在火情。目前,主流的火灾检测方法分为两类:一类为基于目标检测技术的火灾检测方法,另一类为无监督火灾检测方法。

2.1 基于目标检测的火灾检测方法

基于目标检测技术的火灾检测方法的思路是通过YOLO^[10]、Faster R-CNN^[11]等目标检测方法检测输入图片中是否存在包含火焰的区域,若检测到火焰则判定为出现火灾。例如,Barmoutis等人^[2]提出了一种基于图像的火灾检测方法,该方法将Faster R-CNN与基于高阶线性动力系统的多维纹理分析相结合,实现了高性能火灾检测。Zhao等人^[3]提出了一种改进的Fire-YOLO深度学习算法,对特征提取网络进行三维扩展,增强了火灾小目标识别的特征传播能力,在提高网络性能的同时降低了参数量。Zhang等人^[4]提出在YOLOv5网络中引入Swin Transformer,在不改变模型深度的情况下增强了模型的感受野和特征提取能力。张旭彤等人^[5]则在YOLOv5模型的基础上引入了CIoU损失函数以提高检测精度。Talaat等人^[6]则基于YOLOv8目标检测算法提出了一种智能城市火灾探测方法。在火灾初期,小火焰目标检测是一个挑战,针对该问题,Lu等人^[7]提出了随机源对角交换的数据增强策略。而汪子健等人^[8]提出一种新的损失函数eIoU来替代传统的IoU损失函数,解决了在检测火焰小目标时预测框和真实框出现无交集的情况。针对传统火灾监测算法参数量大的问题,Huang等人^[9]利用GhostNet、深度可分离卷积和SENet对YOLOX-L进行轻量级改进,在实现高精度火灾检测的同时减少了92.6%的参数量。

2.2 无监督火灾检测方法

基于目标检测技术的火灾检测方法的局限性在于

需要大量符合目标检测规范的人工标注数据,但火灾数据在现实中往往较为稀缺,且标注成本较高.无监督火灾检测方法则不需要对数据进行人工标注,其使用大量无火灾的图像进行训练.例如,宋焕生等人^[12]提出一种基于改进教师学生网络的隧道火灾检测方法.通过无监督学习对无火灾的样本训练从而检测火灾,同时在用于知识蒸馏的残差块中加入注意力机制以减少重要信息损失,提高火灾检测精度. Buza 等人^[13]提出了一种基于 RGB 色彩空间的无监督野火早期检测方法,通过使用 RGB 色彩空间,从而消除耗时的色彩空间转换任务. Ajith 等人^[14]利用空间、时间和运动信息从视频帧中提取所需的区域,并对光流、散度和强度值进行特征融合,最后利用这些特征以无监督的方式将像素划分为不同的类别.李权威等人^[15]则基于 Dignet ANN 无监督聚类算法实现火灾检测模型,并提出了环境模式阈值自适应方法进一步提升检测精度.为了解决森林火灾背景复杂多样的问题, Yang 等人^[16]通过学习一个单分类模型来解决火灾检测问题,并提出了一个批量决策策略以提升模型性能.

2.3 本文方法

由于没有可用的监督信息,实现高准确率的无监督火灾检测十分具有挑战性.为了应对当前的挑战,本研究提出了一种无监督火灾检测方法.我们的方法与已有火灾检测方法的主要差别在于:(1)我们将火灾检测视为一项无监督异常检测任务,有效避免了在真实工业场景和社区场景中,由于火灾的高危险性,真实火灾数据少、获取数据困难的问题.(2)我们提出了交叉输入对比学习模块,有效学习图像的实例级特征,同时,设计了一个记忆原型和特征重建模块,通过重建的方式判别火灾场景图像.(3)我们提出了一种伪异常火灾场景合成方法,利用正常图像合成具备火灾场景特征的伪异常图像,同时提出了基于欧氏距离的异常特征区分损失来充分利用伪异常图像,使模型对于火灾场景具有针对性,并缓解了仅使用正常数据训练模型导致部分样本检测置信度较低的问题.

3 方法

本节将详细描述我们的方法.第 3.1 节对研究内容和模型框架进行介绍.第 3.2 介绍了判别特征的对比学习模块.第 3.3 节描述了记忆原型和特征重建模块.第 3.4 节阐述了伪异常火灾场景合成方法.

3.1 模型整体框架

我们的方法将火灾检测任务视为一项无监督图像异常检测任务.对于一个场景图像,如果其中包含火焰,则将其视为异常图像,反之视为正常图像.我们的目的是仅使用正常图像数据来训练网络模型,并在推理时实现对正常图像和异常图像的判别.我们的模型由交叉输入对比学习模块、记忆原型和特征重建模块和伪异常火灾场景合成方法 3 个部分组成.首先,由交叉输入对比学习模块以无监督的方式学习图像的特征.然后,通过记忆原型和特征重建模块利用这些图像特征生成记忆原型,并根据记忆原型进行特征重建.最后,通过伪异常火灾场景合成方法合成伪异常火灾样本提升模型对于火灾场景的针对性.模型整体框架如图 1 所示,图 1 中箭头表示数据流动的方向,这些箭头被分为黑、蓝和红 3 种颜色,分别对应当前数据流所属的部分为交叉输入对比学习模块、记忆原型和特征重建模块和伪异常火灾场景合成方法.模型的具体细节将在下面的小节中详细介绍.

3.2 交叉输入对比学习模块

我们遵循 He 等人^[18]的思想,利用自监督对比学习来学习图像的实例级特征.通常,对比学习框架有两个编码器,它们通过定义正、负样本对进行对比优化.在 He 等人^[18]提出的方法中,每个图像形成的样本对被分别输入两个结构相同的编码器 E_q 和 E_k 进行对比学习,但由于样本对中的两个图像总是由不同增强方式得到,因此这种方式在学习过程中无法充分利用样本对之间的关联信息,这将导致最终学到的特征表示不够充分,对比学习的效果受到影响.受到 Caron 等人^[19]的启发,我们不仅将每个图像形成的样本对分别输入编码器 E_q 和 E_k ,而且交叉地将样本对分别输入编码器 E_k 和 E_q .在调换输入编码器 E_k 和编码器 E_q 的过程中,两个编码器可以相互匹配和学习,使得学到的特征表示更具鲁棒性和泛化能力.接下来将详细描述该方法.

首先,给定一个全部由正常图像构成的数据集 $X = [X_1, X_2, \dots, X_n]$.为了进行对比学习,我们使用图像增强的方式生成样本对,图像增强的方式包括随机翻转,随机裁剪,随机对比度变化和高斯模糊.对于图像 $X_i \in X$,它被增强两次后形成正样本对 (X_i^a, X_i^b) ,而对于不同的图像 X_i 和 X_j 则形成负样本对 (X_i^a, X_j^b) ,其中 $i \neq j$.

然后,将来自同一图片的正样本对 (X_i^a, X_i^b) 中的 X_i^a 和 X_i^b 分别输入两个 ResNet50 编码器 E_q 和 E_k .经由

编码器输出的样本对的特征 f_i^{aq} 和 f_i^{bk} 的表达式为:

$$f_i^{aq} = E_q(X_i^a), f_i^{bk} = E_k(X_i^b) \quad (1)$$

为了方便后续计算记忆原型, 我们通过创建一个队列 Q_f 来存放每个 mini-batch 的特征, 此时将 f_i^{bk} 存入

队列 Q_f 中. 同时, 将 X_i^a 和 X_i^b 对称的输入编码器 E_k 和 E_q , 经由编码器输出的样本对的特征 f_i^{bq} 和 f_i^{ak} 的表达式为:

$$f_i^{bq} = E_q(X_i^b), f_i^{ak} = E_k(X_i^a) \quad (2)$$

同样, 将 f_i^{bk} 存入队列 Q_f 中.

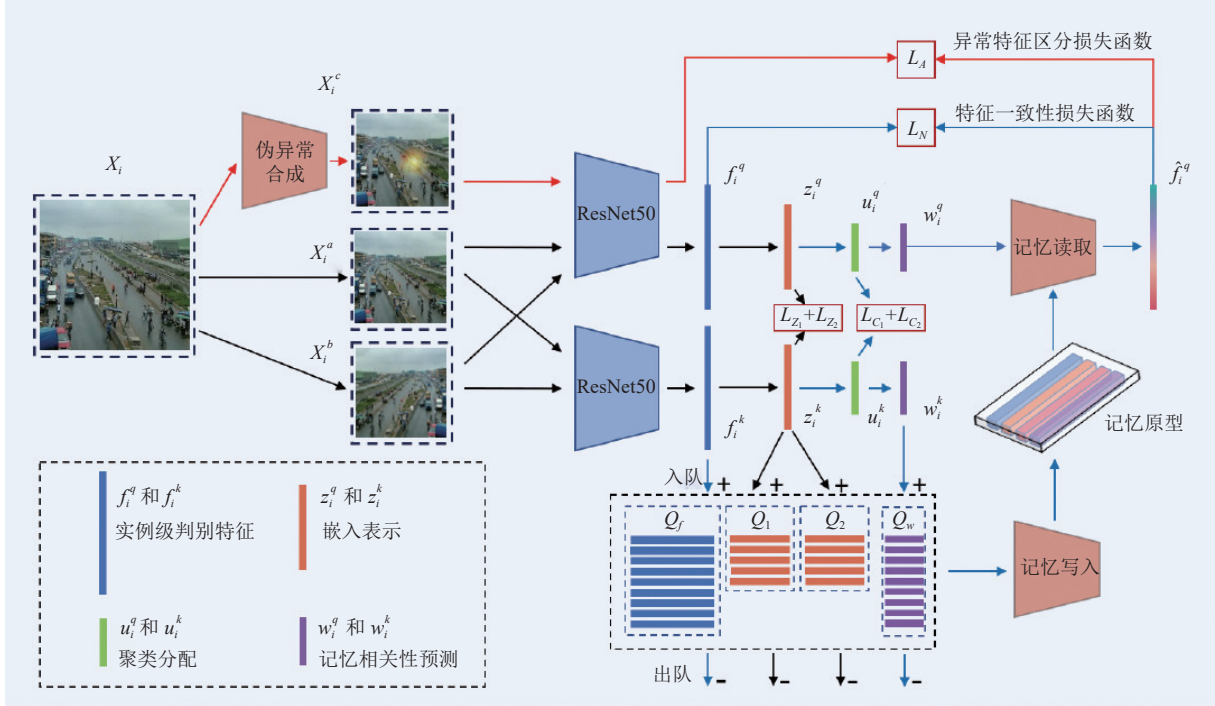


图1 模型整体框架

接着, 在编码器 E_q 、 E_k 之后建立一组非线性嵌入层 G_q 和 G_k , 得到 f_i^{aq} 和 f_i^{bk} 的嵌入表示 z_i^{aq} 和 z_i^{bk} , 表达式为:

$$z_i^{aq} = G_q(f_i^{aq}), z_i^{bk} = G_k(f_i^{bk}) \quad (3)$$

同时, 将 f_i^{bq} 和 f_i^{ak} 也经过非线性嵌入层 G_q 、 G_k , 得到嵌入表示 z_i^{bq} 和 z_i^{ak} , 表达式为:

$$z_i^{bq} = G_q(f_i^{bq}), z_i^{ak} = G_k(f_i^{ak}) \quad (4)$$

为了更好地优化对比学习的嵌入表示, 我们通过创建两个队列 Q_1 和 Q_2 来扩展前面 mini-batch 的负样本集, 其中 Q_1 存放每次迭代产生的 z_i^{bq} , Q_2 存放每次迭代产生的 z_i^{ak} . 队列的容量为 $N_{Q_1} = N_{Q_2} = 4096$, 在最初的几次迭代之后将被填满, 之后 mini-batch 中的新嵌入表示将被添加到队列中, 并且删除最早的嵌入表示. 优化对比学习的嵌入表示的损失函数 L_{Z_1} 定义为:

$$L_{Z_1} = -\frac{1}{N} \sum_{i=1}^N \log \frac{\exp(z_i^{aq} \cdot z_i^{bk} / T_z)}{\sum_{z^{bk} \in Z_1} \exp(z_i^{aq} \cdot z^{bk} / T_z)} \quad (5)$$

其中, N 为 mini-batch 的大小, T_z 为温度参数. Z_1 定义为:

$$Z_1 = \{z_i^{bq}\}_{i=1}^N \cup \{z_{i_{Q_1}}^{bk}\}_{i_{Q_1}=1}^{N_{Q_1}} \quad (6)$$

其中, $z_{i_{Q_1}}^{bk}$ 表示存储在队列 Q_1 中的特征嵌入表示. 优化对比学习的嵌入表示的损失函数 L_{Z_2} 定义为:

$$L_{Z_2} = -\frac{1}{N} \sum_{i=1}^N \log \frac{\exp(z_i^{bq} \cdot z_i^{ak} / T_z)}{\sum_{z^{ak} \in Z_2} \exp(z_i^{bq} \cdot z^{ak} / T_z)} \quad (7)$$

其中, N 为 mini-batch 的大小, T_z 为温度参数. Z_2 定义为:

$$Z_2 = \{z_i^{ak}\}_{i=1}^N \cup \{z_{i_{Q_2}}^{ak}\}_{i_{Q_2}=1}^{N_{Q_2}} \quad (8)$$

其中, $z_{i_{Q_2}}^{ak}$ 表示存储在队列 Q_2 中的特征嵌入表示.

3.3 记忆原型和特征重建模块

由于火灾图像的特征与正常图像的特征偏差较大,

需要根据特征空间中的距离对其进行分离. 因此, 受到 Ning 等人^[17]的启发, 我们引入了一个记忆模块, 该模块产生一个可更新的记忆原型. 但与之不同的是, 由于在第 3.2 节中我们进行了对称输入的对比学习, 因此记忆原型的计算也由 $f_i^{b_k}$ 和 $f_i^{a_k}$ 共同参与.

为了计算记忆原型并学习正常样本和记忆原型之间的关系, 我们使用深度聚类方法输出每个样本的聚类预测. 在嵌入层之后, 定义一个聚类层 $H_q(\cdot)$ 和 $H_k(\cdot)$. 嵌入表示 $z_i^{a_q}$ 和 $z_i^{b_k}$ 经过聚类层后输出的聚类分配向量定义为:

$$u_i^{a_q} = H_q(z_i^{a_q}), u_i^{b_k} = H_k(z_i^{b_k}) \quad (9)$$

同样地, 嵌入表示 $z_i^{b_q}$ 和 $z_i^{a_k}$ 经过聚类层后的输出定义为:

$$u_i^{b_q} = H_q(z_i^{b_q}), u_i^{a_k} = H_k(z_i^{a_k}) \quad (10)$$

其中, $u_i^{a_q}$ 和 $u_i^{b_k}$ 以及 $u_i^{b_q}$ 和 $u_i^{a_k}$ 都是大小为 $1 \times C$ 的聚类分配向量, C 为簇号. 考虑一个 mini-batch 的聚类分配向量, 它们分别由矩阵 U^{a_q} 和 U^{b_k} 以及 U^{b_q} 和 U^{a_k} 构建, 定义为:

$$\begin{cases} U^{a_q} = \begin{bmatrix} u_1^{a_q} \\ \dots \\ u_N^{a_q} \end{bmatrix}_{N \times C}, U^{b_k} = \begin{bmatrix} u_1^{b_k} \\ \dots \\ u_N^{b_k} \end{bmatrix}_{N \times C} \\ U^{b_q} = \begin{bmatrix} u_1^{b_q} \\ \dots \\ u_N^{b_q} \end{bmatrix}_{N \times C}, U^{a_k} = \begin{bmatrix} u_1^{a_k} \\ \dots \\ u_N^{a_k} \end{bmatrix}_{N \times C} \end{cases} \quad (11)$$

其中, 元素 $u_{ic}^{a_q}$ 和 $u_{ic}^{b_k}$ 以及 $u_{ic}^{b_q}$ 和 $u_{ic}^{a_k}$ 是样本 X_i 增强后属于簇 c 的簇对数, $c = 1, \dots, C$. 将矩阵写成列分块矩阵形式:

$$\begin{cases} U^{a_q} = [v_1^{a_q}, \dots, v_C^{a_q}]_{N \times C}, U^{b_k} = [v_1^{b_k}, \dots, v_C^{b_k}]_{N \times C} \\ U^{b_q} = [v_1^{b_q}, \dots, v_C^{b_q}]_{N \times C}, U^{a_k} = [v_1^{a_k}, \dots, v_C^{a_k}]_{N \times C} \end{cases} \quad (12)$$

实现聚类的损失函数 L_{C_1} 定义为:

$$L_{C_1} = -\frac{1}{C} \sum_{c=1}^C \log \frac{\exp(v_c^{a_q} \cdot v_c^{b_k} / T_C)}{\sum_{c'=1}^C \exp(v_c^{a_q} \cdot v_{c'}^{b_k} / T_C)} \quad (13)$$

其中, $v_c^{a_q} \in U^{a_q}$ 和 $v_c^{b_k} \in U^{b_k}$, T_C 是温度参数.

实现聚类的损失函数 L_{C_2} 定义为:

$$L_{C_2} = -\frac{1}{C} \sum_{c=1}^C \log \frac{\exp(v_c^{b_q} \cdot v_c^{a_k} / T_C)}{\sum_{c'=1}^C \exp(v_c^{b_q} \cdot v_{c'}^{a_k} / T_C)} \quad (14)$$

其中, $v_c^{b_q} \in U^{b_q}$ 和 $v_c^{a_k} \in U^{a_k}$, T_C 是温度参数.

记忆相关性预测反映了样本和每个聚类中心之间

的关系, 定义为:

$$\begin{cases} w_i^{a_q} = \phi_q(u_i^{a_q}), w_i^{b_k} = \phi_k(u_i^{b_k}) \\ w_i^{b_q} = \phi_q(u_i^{b_q}), w_i^{a_k} = \phi_k(u_i^{a_k}) \end{cases} \quad (15)$$

其中, $\phi(\cdot)$ 为 Softmax 函数. 为了计算记忆原型, 我们通过创建一个队列 Q_w 来存放每个 mini-batch 的记忆相关性预测将 $w_i^{b_k}$ 和 $w_i^{a_k}$ 存入队列 Q_w 中.

记忆写入: 为了找到更好的一致性特征, 记忆原型应该在训练过程中更新. 因此, 模型采用一个记忆写操作, 根据存储在队列 Q_f 中的特征 $f_{i_Q}^k$ 及存在队列 Q_w 中的记忆相关性预测 $w_{i_Q}^k$ 生成新的记忆原型. 写入操作为:

$$m_c = \sum_{i_Q=1}^{N_Q} w_{i_Q}^k f_{i_Q}^k \quad (16)$$

其中, $f_{i_Q}^k$ 和 $w_{i_Q}^k$ 是存储在队列 Q_f 和 Q_w 中的前一个 mini-batch 的图像特征和记忆相关性预测. 利用队列中样本的输出来更新记忆原型, 以保持记忆原型的稳定性.

记忆读取: 为了保证特征的一致性, 本项目需要最小化样本特征与其对应的记忆原型之间的距离. 给定记忆矩阵 $M = [m_1, m_2, \dots, m_C]_{d \times C}$, 其中每一列为一个记忆原型 m_c , d 为每个记忆原型的特征维数. 这些记忆原型是每个集群的聚类中心. 考虑到样本通常在几个相关的集群中共享某些特征, 我们采用软赋值记忆读取:

$$\hat{f}_i^q = \sum_{c=1}^C w_{ic}^q m_c \quad (17)$$

其中, w_{ic}^q 表示由 ϕ_q 输出的记忆相关性预测, 包括了 $w_i^{a_q}$ 和 $w_i^{b_q}$.

在训练过程中, 模型优化的目标是在当前的记忆原型和记忆相关性预测下, 使 f_i^q 和 \hat{f}_i^q 之间的距离最小化, 基于欧氏距离的特征一致性损失函数定义为:

$$L_N = \frac{1}{N} \sum_{i=1}^N \|\hat{f}_i^q - f_i^q\|^2 \quad (18)$$

其中, f_i^q 表示由 E_q 输出的实例级判别特征, 包括 $f_i^{a_q}$ 和 $f_i^{b_q}$, N 为批量大小.

动量更新: 在模型优化过程中, E_q 充当一般网络, E_k 充当动量网络. 它们的参数逐渐更新为:

$$\theta^k = \alpha \theta^k + (1 - \alpha) \theta^q \quad (19)$$

其中, θ^q 和 θ^k 分别是 E_q , G_q , H_q 和 E_k , G_k , H_k 的参数集, $\alpha \in [0, 1)$ 是动量系数.

3.4 伪异常火灾场景合成

我们的方法将火灾检测问题视为无监督图像异常

检测问题,该方法当前存在一个难题,即由于仅使用正常数据训练模型,所以模型对于“火灾”的概念是模糊的,在实际检测的过程中,往往存在“很像火灾的正常”图像和“很像正常的火灾”图像,这会导致模型的检测结果置信度较低.为了解决这个问题,我们提出了伪异常火灾场景合成方法.我们的方法不同于大多数伪异常合成方法,大多数方法使用正常样本和合成的伪异常样本进行监督学习,训练一个能够判断正常和异常的判别器^[20].我们考虑到火灾异常是无界的,即火灾场景有无数种情况,而使用伪异常合成方法合成的火灾情景是有限的.因此,我们将合成的伪异常火灾情景作为重建的负面例子对模型优化过程产生影响,从而在训练过程中促使模型更加关注正常样本与火灾样本之间的区别,而不是图像中的其他无关信息.通过这种方式,能够使模型对于火灾场景具有针对性,并缓解仅使用正常数据训练模型导致部分样本检测置信度较低的问题.

首先,我们用红、橙、黄和黑等颜色的像素点人工合成一张 50×50 像素的火焰素材.然后,对于每个图像 $X_i \in X$,在每次迭代中,我们将人工合成的火焰素材按照一定范围内随机的透明度、旋转角度、对比度、饱和度覆盖在图像 X_i 上以获得人工合成的火灾场景 X_i^c ,表示为:

$$X_i^c = F(X_i) \quad (20)$$

其中, $F(\cdot)$ 表示合成火灾场景的操作. X_i^c 具备火灾场景的显著特征,如图2所示.

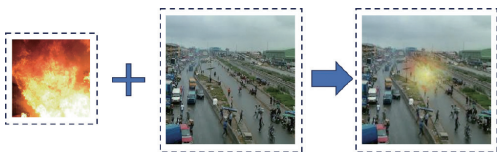


图2 伪异常火灾场景合成示例

然后,将 X_i^c 输入编码器 E_q 以得到特征 f_i^c ,表示为:

$$f_i^c = E_q(X_i^c) \quad (21)$$

由于 f_i^c 代表的是火灾场景图像的特征,因此在训练过程中,模型优化的目标是在当前的记忆原型和聚类预测下,使 f_i^c 和 \hat{f}_i^q 之间的距离最大化,基于欧氏距离的异常特征区分损失函数定义为:

$$L_A = -\frac{1}{N} \sum_{i=1}^N \hat{f}_i^q - f_i^c^2 \quad (22)$$

其中, N 为 mini-batch 的大小.

3.5 训练和推理

模型训练时,由对比损失函数 L_{Z_1} 和 L_{Z_2} 、聚类损失函数 L_{C_1} 和 L_{C_2} 、基于欧氏距离的特征一致性损失函数 L_N 以及基于欧氏距离的异常特征区分损失函数 L_A 对模型整体进行联合优化,模型总体损失函数表示为:

$$L_{\text{all}} = L_{Z_1} + L_{Z_2} + L_{C_1} + L_{C_2} + \varepsilon(L_N + L_A) \quad (23)$$

其中, ε 为权重超参数.

在模型推理过程中,只需将待推理的图像输入编码器 E_q 中,模型即可输出图像对应的异常分数 S ,异常分数 S 表达为:

$$S = \|\hat{f}_i^q - f_i^q\|^2 \quad (24)$$

异常分数 S 代表的含义是模型对于当前图像的重建误差值,重建误差越小说明该图像越可能属于正常图像,反之重建误差越大则该图像越有可能属于火灾场景图像.因此,最终模型根据异常分数的大小来判别正常图像与火灾场景图像.

4 实验分析

4.1 数据集和评估方法

我们在 Fire-Flame-Dataset 和 Fire-Detection-Image-Dataset 两个公共数据集上对该方法进行了评估.

Fire-Flame-Dataset (<https://github.com/DeepQuestAI/Fire-Smoke-Dataset>) 包含 1 000 个包含火焰的图像和 1 000 个正常图像,其中训练集包含带火图像和正常图像各 900 个,测试集包含带火图像和正常图像各 100 个.这些图像中涵盖了室外街区场景和房屋室内场景.我们的方法使用全部 200 个测试集图像进行测试,并使用训练集中 900 个正常图像进行训练.

Fire-Detection-Image-Dataset (<https://github.com/cair/Fire-Detection-Image-Dataset>) 包含带火图像和正常图像,其中包括了各种场景和不同的火灾情况(强度、亮度、大小、环境等),共有 110 个带火图像和 541 个正常图像.我们的方法将数据集中“Fire Images”目录下的 110 个带火图像和“Normal Images 1”目录下 157 个正常图像划分为测试集进行测试,其余目录中共 384 个正常图像划分为训练集进行训练.

评估方法如下:我们使用 ROC 曲线下的面积 ROC-

AUC 和准确率 ACC 作为所有数据集的评价度量。

4.2 实验细节

实验均运行于 1 台显卡为 NVIDIA GeForce RTX4090 (24 GB), CPU 为 Intel Core i7-13700KF 的台式计算机上, 该方法使用 PyTorch 实现. 实验中相关参数设置如表 1 所示.

表 1 实验中相关参数设置

参数名称	数值设置
输入图像尺寸	224×224
批量大小	32
学习率	0.0001
优化器	AdamW
权重衰减	0.0005
训练轮数	1000
记忆原型数量C	10
队列 Q_1 和 Q_2 大小	4096
队列 Q_f 和 Q_w 大小	8192
T_z	0.1
T_C	0.1
ε	0.00005

4.3 检测结果

为了直观地展现模型的结果, 我们从 Fire-Detection-Image-Dataset 数据集的测试集中随机抽取了两个火灾场景图像和两个正常场景图像, 并分别展现了在 PANDA^[21]、MCOD^[17]和我们的模型上的推理结果以及真实标签, 如图 3 所示. 其中红色边框表示异常, 绿色边框表示正常. 从图 3 中可以看出, 我们的模型不仅能够准确检测出火灾场景, 并且对于容易产生误判的正常场景也能够正确进行判断, 具有较高的鲁棒性.



图 3 模型检测结果

4.4 对比实验

为了保证实验的公平性, 我们在数据集和标签条件都相同的情况下进行对比实验. 我们选择了近 3 年

内的 4 种图像异常检测 SOTA (state-of-the-art) 模型 PatchCore^[22]、PANDA^[21]、Mean-Shift^[23]和 MCOD^[17] 与我们的方法进行比较.

实验的 AUC 指标结果如表 2 所示, 其中 FFD 代表 Fire-Flame-Dataset 数据集, 而 FDID 代表 Fire-Detection-Image-Dataset 数据集. 从表 2 中可以看出, 在 FFD 数据集上我们的方法取得了 89.86% 的 AUC, 而在 FDID 数据集上我们的方法取得了 89.56% 的 AUC. 模型在 FFD 和 FDID 两个数据集上获得的 ROC 曲线可视化结果如图 4 所示.

表 2 AUC 结果对比 (%)

方法	年份	FFD	FDID
PatchCore ^[22]	2022	48.19	57.18
PANDA ^[21]	2021	81.46	84.29
Mean-Shift ^[23]	2023	76.57	85.87
MCOD ^[17]	2022	80.21	75.73
Ours	—	89.86	89.56

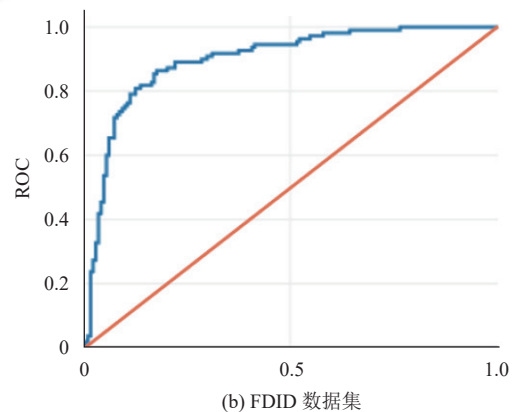
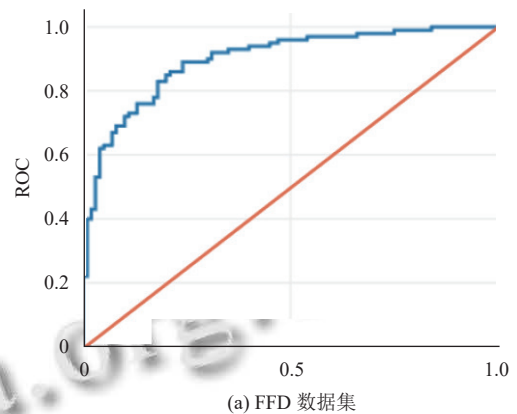


图 4 模型在 FFD 数据集和 FDID 数据集上的 ROC 曲线

可以发现, 与其他检测方法相比, 我们的模型在两个火灾数据集上都获得了最佳的检测性能. 说明其在

火灾检测任务上具有更高的准确性和可靠性,能够更加精准地判别火灾场景与非火灾场景,也说明其具有更好的鲁棒性。

为了更直观地反映模型对火灾场景和正常场景的判别,我们还对模型的 ACC 指标进行了比较。每个模型的 ACC 结果如表 3 所示,其中 FFD 代表 Fire-Flame-Dataset 数据集,FDID 代表 Fire-Detection-Image-Dataset 数据集。从表 3 中可以看出,在 FFD 数据集上我们的方法取得了 82.50% 的 ACC,而在 FDID 数据集上我们的方法取得了 84.26% 的 ACC。与次优方法相比,我们的模型在 FFD 和 FDID 两个火灾数据集上的 ACC 性能分别高出了 6.50% 和 4.49%。这些结果进一步验证了该方法在火灾检测任务中的优越性。

表 3 ACC 结果对比 (%)

方法	年份	FFD	FDID
PatchCore ^[22]	2022	52.50	59.92
PANDA ^[21]	2021	74.50	79.77
Mean-Shift ^[23]	2023	70.50	78.27
MCOD ^[17]	2022	76.00	70.78
Ours	—	82.50	84.26

4.5 消融实验

为了验证我们提出的伪异常火灾场景合成和对称对比学习模块对模型性能的贡献,我们进行了消融实验。首先,我们将 MCOD 模型作为基线模型,然后我们加入伪异常火灾场景合成(SFS)和对称对比学习(SCL)进行实验,消融实验结果如表 4 所示。从表 4 中可知,加入伪异常火灾场景合成在 FFD 和 FDID 两个数据集上模型 AUC 分别提升了 7.69% 和 12.10%,这证明伪异常火灾场景合成能够显著提升模型对火灾的判别能力。在此基础上进一步加入对称对比学习,在 FFD 和 FDID 两个数据集上模型 AUC 分别进一步提升了 1.96% 和 1.73%,说明对称对比学习能够进一步提升模型对火灾的判别能力。

表 4 消融实验结果 (AUC (%))

方法	FFD	FDID
Baseline	80.21	75.73
Baseline+SFS	87.90	87.83
Baseline+SFS+SCL	89.86	89.56

5 结论与展望

在真实工业场景和社区场景中,由于火灾的高危险性,普遍存在着真实火灾数据少、获取数据困难的问题。为了应对当前存在的问题,本研究提出了一种基

于对比学习和伪异常合成的无监督火灾检测方法,利用正常数据训练模型以实现高精度火灾检测。首先,提出了交叉输入对比学习模块,利用交叉输入的方式提升对比学习的性能。然后,引入记忆原型学习正常场景图像的特征分布,并通过重建的方法实现对火灾场景的判别。最后,通过所提出的伪异常火灾场景合成方法和基于欧氏距离的异常特征区分损失,使模型对于火灾场景具有针对性,并缓解了仅使用正常数据训练模型导致部分样本检测置信度较低的问题。在两个公共火灾检测数据集 Fire-Flame-Dataset 和 Fire-Detection-Image-Dataset 上的大量实验证明了我们的模型的优越性。

目前,基于对比学习和伪异常合成的无监督火灾检测的不足在于检测粒度不够精细,只能在图像级别进行检测,接下来的研究工作是设计更精细的方法,在进行图像级检测的同时,实现在图像中对火焰的像素级定位。

参考文献

- 祝玉华, 司艺艺, 李智慧. 基于深度学习的烟雾与火灾检测算法综述. 计算机工程与应用, 2022, 58(23): 1-11.
- Barmpoutis P, Dimitropoulos K, Kaza K, et al. Fire detection from images using Faster R-CNN and multidimensional texture analysis. Proceedings of the 2019 IEEE International Conference on Acoustics, Speech and Signal Processing. Brighton: IEEE, 2019. 8301-8305.
- Zhao L, Zhi LQ, Zhao C, et al. Fire-YOLO: A small target object detection method for fire inspection. Sustainability, 2022, 14(9): 4930. [doi: 10.3390/su14094930]
- Zhang SJG, Zhang FX, Ding YY, et al. Swin-YOLOv5: Research and application of fire and smoke detection algorithm based on YOLOv5. Computational Intelligence and Neuroscience, 2022, 2022: 6081680.
- 张旭彤, 胡鹏, 赵鑫, 等. 火灾预警中基于 YOLO v5 的火源智能检测定位方法. 微电子学与计算机, 2023, 40(3): 67-74.
- Talaat FM, ZainEldin H. An improved fire detection approach based on YOLO-v8 for smart cities. Neural Computing and Applications, 2023, 35(28): 20939-20954. [doi: 10.1007/s00521-023-08809-1]
- Lu KJ, Huang JW, Li JH, et al. MTL-FFDET: A multi-task learning-based model for forest fire detection. Forests, 2022, 13(9): 1448. [doi: 10.3390/f13091448]
- 汪子健, 高焕兵, 侯宇翔, 等. 改进 YOLOX-Nano 的火灾火

- 焰烟雾检测. 计算机系统应用, 2023, 32(3): 265–274. [doi: 10.15888/j.cnki.csa.009000]
- 9 Huang JR, He ZL, Guan YW, *et al.* Real-time forest fire detection by ensemble lightweight YOLOX-L and defogging method. *Sensors*, 2023, 23(4): 1894. [doi: 10.3390/s23041894]
- 10 Redmon J, Divvala S, Girshick R, *et al.* You only look once: Unified, real-time object detection. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016. 779–788.
- 11 Ren SQ, He KM, Girshick R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks. *Proceedings of the 28th International Conference on Neural Information Processing Systems*. Montreal: ACM, 2015. 91–99.
- 12 宋焕生, 文雅, 孙士杰, 等. 基于改进教师学生网络的隧道火灾检测. *图学学报*, 2023, 44(5): 978–987.
- 13 Buza E, Akagic A. Unsupervised method for wildfire flame segmentation and detection. *IEEE Access*, 2022, 10: 55213–55225. [doi: 10.1109/ACCESS.2022.3176724]
- 14 Ajith M, Martínez-Ramón M. Unsupervised segmentation of fire and smoke from infra-red videos. *IEEE Access*, 2019, 7: 182381–182394. [doi: 10.1109/ACCESS.2019.2960209]
- 15 李权威, 宛田宾, 秦俊, 等. 基于 Dignet 无监督学习聚类算法的智能火灾探测. *中国科学技术大学学报*, 2009, 39(7): 769–776, 782.
- 16 Yang XB, Wang Y, Liu XD, *et al.* High-precision real-time forest fire video detection using one-class model. *Forests*, 2022, 13(11): 1826. [doi: 10.3390/f13111826]
- 17 Ning HY, Quan D, Zhang XR, *et al.* Unsupervised outlier detection using memory and contrastive learning. *IEEE Transactions on Image Processing*, 2022, 31: 6440–6454. [doi: 10.1109/TIP.2022.3211476]
- 18 He KM, Fan HQ, Wu YX, *et al.* Momentum contrast for unsupervised visual representation learning. *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020. 9726–9735.
- 19 Caron M, Misra I, Mairal J, *et al.* Unsupervised learning of visual features by contrasting cluster assignments. *Proceedings of the 34th International Conference on Neural Information Processing Systems*. Vancouver: ACM, 2020. 831.
- 20 Astrid M, Zaheer MZ, Lee SI. Synthetic temporal anomaly guided end-to-end video anomaly detection. *Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision*. Montreal: IEEE, 2021. 207–214.
- 21 Reiss T, Cohen N, Bergman L, *et al.* PANDA: Adapting pretrained features for anomaly detection and segmentation. *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021. 2805–2813.
- 22 Roth K, Pemula L, Zepeda J, *et al.* Towards total recall in industrial anomaly detection. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans: IEEE, 2022. 14298–14308.
- 23 Reiss T, Hoshen Y. Mean-shifted contrastive loss for anomaly detection. *Proceedings of the 2023 AAAI Conference on Artificial Intelligence*. Washington: AAAI, 2023. 2155–2162.

(校对责编: 孙君艳)