

基于 SE-U-Net 预测网络的视频异常事件检测方法

王伟胜 王来花* 贾 晴 赵 月

(曲阜师范大学网络空间安全学院 山东 曲阜 273165)

摘 要 针对视频异常检测中存在的數據不平衡问题,提出一种基于 SE-U-Net 预测网络的视频异常检测方法。该方法提取视频帧的显著性图,并将其制作成掩膜对数据进行预处理;利用预处理后的数据对预测模型进行训练,为了使预测模型更关注前景区域的优化,结合注意力机制设计一组新的损失函数用于约束模型的训练。在测试阶段设计一个新的异常评价分数计算方法,通过仅计算视频中显著性区域的预测误差来进行异常检测,缓解数据不平衡问题。利用公共数据集进行相关对比实验以及消融实验验证该方法的有效性。

关键词 视频异常检测 显著性提取 未来帧预测 U-Net 异常评价分数

中图分类号 TP391

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2024.12.022

VIDEO ABNORMAL EVENT DETECTION BASED ON SE-U-NET PREDICTIVE NETWORK

Wang Weisheng Wang Laihua* Jia Qing Zhao Yue

(School of Cyber Science and Engineering, Qufu Normal University, Qufu 273165, Shandong, China)

Abstract Aimed at the problem of data imbalance in video anomaly detection, an anomaly video detection method based on SE-U-Net predictive network is proposed. We extracted the saliency map of the video frame and made it into a mask to preprocess the data. We used the preprocessed data to train the prediction model. In order to make the prediction model pay more attention to the optimization of the foreground area, this paper combined the attention mechanism, and a new set of loss functions were designed to constrain the training of the model. In addition, in the testing phase, this paper designed a new anomaly evaluation score calculation method, and anomaly detection was performed only by calculating the prediction error of the saliency region in the video, which alleviated the problem of data imbalance. Public datasets for comparative experiments and ablation experiments were used to verify the effectiveness of the proposed method for abnormal event detection.

Keywords Video anomaly detection Saliency extraction Prediction of future frames U-Net Anomaly evaluation score

0 引言

随着公共安全问题的日益突出,及时地发现公共场所发生的异常事件是建设平安城市的一个非常重要的环节。在视频数据激增的今天,仅仅依靠传统的人力资源来监测异常事件已经不可行,因此,近年来,随着视频分析技术和模式识别技术的进步,智能的高精度的视频异常检测技术取得了长足的进步,已成为图

像处理、机器视觉、机器学习等领域的热门研究课题之一。

目前常见的异常检测方法一般可分为两个阶段:训练阶段和测试阶段。在训练阶段,通过学习训练数据的外观和运动特征,建立与正常数据相关联的模型;在测试阶段,根据视频数据与模型是否匹配判定视频数据中是否出现异常。在早期的研究方法中,大多利用手工提取的特征来进行视频异常检测^[1-2],通过提取视频的代表性特征,训练模型来表示正常事件的规

律模式,与模型不匹配的事件检测为异常。然而手工提取特征不能很好地表征复杂行为,且提取的特征单一,对复杂的异常场景不具有鲁棒性。随着机器学习、深度学习技术的发展,3D 卷积神经网络^[3]、长短时记忆网络^[4-5]、生成对抗网络^[6]被很好地应用于异常检测任务当中。然而,大部分基于深度学习的异常检测算法在训练网络模型的过程中都忽视了前景与背景的数据不平衡这一问题。为了解决这个问题,本文提出一种基于 SE-U-Net 预测网络的异常视频检测方法,以此来提高异常检测的准确性。本文的具体贡献如下:

1) 由于异常事件常常发生在人眼感兴趣的前景区域,本文提取视频中的显著性图,并将其制作成掩膜对数据进行预处理,从而使得预测网络忽略对背景信息的学习。

2) 为了更好地监督模型训练,本文将数据预处理制作的掩膜添加到损失函数中来进一步约束模型的训练,使得网络更关注前景区域的优化。

3) 在异常评价的过程中,本文仅对显著的前景区域计算预测误差,进一步提高异常检测的准确率。

1 相关工作

1.1 视频异常检测

近些年,深度学习在异常检测方面取得了巨大成功,为了更好地学习场景表示,已经提出了各种基于深度学习的异常检测方法。深度学习框架下的视频异常检测方法主要采用帧生成的方法,利用生成帧的真实性来检测异常事件。帧生成的异常检测方法进一步又可以分为帧重构和帧预测。帧重构方法是通过重构当前帧并通过重构误差来检测异常事件。Hasan 等^[7]提出一种端到端的深度全卷积自编码网络,通过学习正常事件行为的变化规律来重构输入帧进行异常检测。考虑到长短时记忆网络在处理序列问题上的优势,Chong 等^[8]提出利用融合了卷积神经网络和卷积长短时记忆网络的自编码网络来重构当前帧进行异常事件检测。Ravanbakhsh 等^[9]引入生成对抗网络来学习正常视频帧中的正常模式,保证生成帧的真实性,进一步提升了异常检测的准确性。然而,这些异常检测方法都是基于规则训练数据的重建,认为所有异常事件对应于更大的重建误差,但是由于深度神经网络的良好泛化能力,这种假设不一定成立,正常事件和异常事件的重建误差可能会很相似,从而导致无法辨别是否有异常事件的发生。

与重建策略不同的是,有研究者提出利用预测未

来帧网络框架来解决视频异常事件检测问题。帧预测网络是基于视频的先前几帧而不是其本身来模拟当前帧。Liu 等^[10]提出使用 U-Net 网络作为预测网络,并引入生成对抗网络来保证预测帧的真实性,最后通过预测误差来判别异常事件。该模型整体结构简单,性能突出,最终得到了不错的结果。考虑到 U-Net 缺少对时间信息的建模,Song 等^[11]提出利用长短时记忆网络作为预测网络来进行异常检测,该方法能够更好地学习视频之间时间信息。Dong 等^[12]提出了一种基于双重鉴别器的生成对抗网络的半监督方法,通过评估预测帧的质量获得规则分数,并以此来检测异常帧,该方法考虑了视频剪辑中更多的运动信息。基于帧预测模型的视频异常事件检测方法通常具有结构简单、泛化能力强的特点,并且它可以很好地利用时空信息,这对于视频异常检测更有优势。

1.2 数据不平衡分析

数据不平衡问题普遍存在于各种机器学习任务中。如果训练数据不平衡,大多数学习算法会对频繁(即多数)类别产生归纳学习偏差,导致少数类别识别性能差。在视频异常事件检测任务中,数据不平衡的问题主要指的是图像的前景与背景不平衡问题。通常,异常事件只发生在拍摄影像的一小部分前景区域,如图 1 所示,在校园人行道出现了一辆汽车,被认为是一个异常事件,汽车只占拍摄区域的一小部分,背景信息占据了更大一部分区域,我们希望设计的网络可以更多地关注异常区域的建模。然而目前视频异常检测方法通常对视频影像的所有区域同等对待,使得设计的网络结构更偏向于捕捉大量的背景信息^[13],设计的网络失去“焦点”。因此缓解数据不平衡问题是视频异常检测任务重要的一环。



图 1 数据不平衡示例图

1.3 SGSP

SGSP 是一种基于超像素图和时空传播的视频显著性检测算法,用来提取视频中感兴趣区域。该算法分为特征提取、构建超像素图度量运动显著性、帧间时间传播、帧内空间传播四部分。首先,提取超像素级运

动和颜色直方图作为显著性度量的特征。然后构建超像素并根据分割结果建立一个无向加权图,利用图中最短路径算法来生成运动显著性图。利用帧间相似性矩阵来执行显著性在前向和后向的时间传播,以获得具有更好一致性的时域显著性图。最后,通过利用帧内相似性矩阵来执行局部和全局显著性的空间特征传播,获得质量更高的时空显著性图。因此,本文选用 SGSP 算法来获取视频帧的显著性区域,以缓解图像前景与背景不平衡的问题。

2 SE-U-Net

考虑到利用预测模型解决异常检测任务结构简单并且可以充分地利用时空信息,本文搭建预测模型来解决视频异常检测问题。整体网络框架 SE-U-Net 如图 2 所示,主要分为两部分:数据的预处理和预测模型的搭建。其中数据预处理是对视频帧做显著性提取 SE(Saliency Extraction)操作,获取视频帧的感兴趣区域,并选用 U-Net 网络作为该方法的预测模型。详细介绍如下。

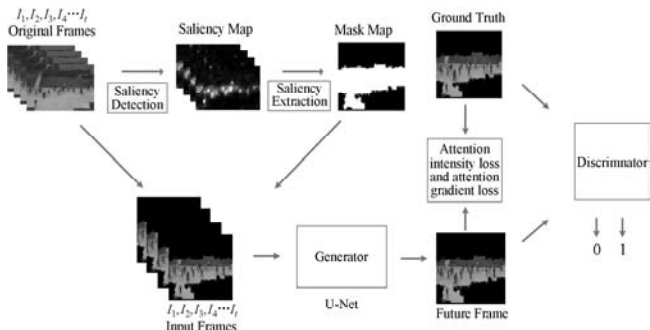
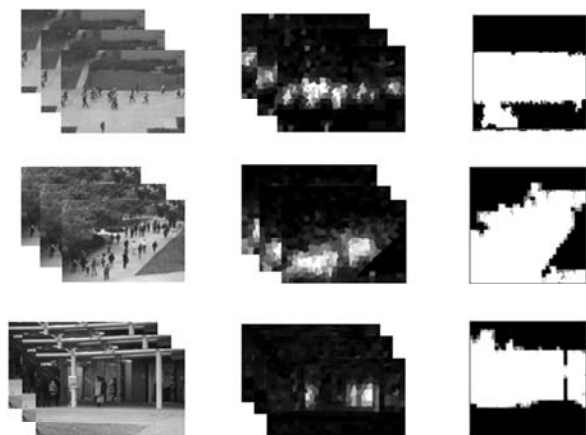


图2 预测网络整体架构

2.1 数据预处理

显著性检测研究中一个重要的理论是注意心理学中的视觉关注机制^[14]。该机制让人们看到视觉场景时,能够自动滤除场景中冗余的视觉信息同时保留大部分重要的视觉信息,从而让视觉系统专注于处理场景中的显著区域。考虑到视频中的异常区域可能仅仅发生图像的前景区域,然而目前的异常检测算法通常对视频帧的所有区域同等对待,会造成数据不平衡的问题,使得神经网络失去焦点,不再关注前景区域的优化。因此,本文提出一种数据预处理的方法,该方法利用显著性检测算法提取视频中显著性区域并将其制作成掩膜对数据进行预处理。详细步骤如下:

步骤1 首先利用 SGSP 显著性检测算法^[15]提取时空显著性图 S ,得到的显著图如图 3(b)所示。



(a) 原图 (b) 显著性图 (c) 掩膜图

图3 Ped1、Ped2 和 Avenue 数据集对应的显著性图和掩膜图

步骤2 为了进一步屏蔽图像中背景信息的噪声,我们将上一步得到的时空显著性图 S 进行二值化处理:

$$S^b(i, j) = \begin{cases} 1 & S(i, j) > g \\ 0 & S(i, j) \leq g \end{cases} \quad (1)$$

式中: i, j 表示图像的空间索引, g 表示最佳分割阈值。

步骤3 利用二值化后的显著性图制作一个只包含数据集中所有感兴趣区域的掩膜 M :

$$M = S_1^b \cup S_2^b \cup \dots \cup S_t^b \quad t = 1, 2, \dots, T \quad (2)$$

式中: S_t^b 表示二值化后的视频帧中显著性区域, $\cup$ 代表对应像素相位或操作。

步骤4 将只包含感兴趣区域的掩膜利用图像处理算法去除小的连通区域,得到最终的掩膜如图 3(c)所示,然后将其与原始数据进行点乘,即可得到最终用于模型训练的训练集。

$$T_{\text{rainData}} = O_t \cdot bwareaopen(M) \quad t = 1, 2, \dots, T \quad (3)$$

式中: O_t 表示原始的数据集, T_{rainData} 表示数据预处理后的数据集, $\cdot$ 表示矩阵对应的像素乘积操作, $bwareaopen(\cdot)$ 函数表示去除小的连通区域。

2.2 未来帧预测

本文的预测网络由生成器和鉴别器两部分组成,与传统的自动编码器结构相比, Ronneberger 等^[16]提出的 U-Net 网络包括具有相同分辨率的高层和低层之间的跳跃连接,实现了高层语义信息和低层外观信息的结合, U-Net 的跳跃连接结构有助于还原池化层导致部分细节信息的丢失。因此,本文选用 U-Net 作为预测网络框架的生成器,该网络由编码和解码两部分组成,编码部分中的每一层都包含两个 3×3 卷积层和一个 2×2 池化层,用于下采样,解码部分的每一层都与对应的编码部分进行连接,用于上采样,卷积过后都有一个用于校正的 ReLU 激活函数。对于鉴别器我们遵循先前研究的框架,用 PatchGAN 来实现鉴别器部分。

2.3 损失函数

基于正常事件是可以预测的,异常事件是不可被预测的基本思想,本文的优化目标是训练一个能够预测未来帧的模型,通过该模型可以准确地预测正常数据的未来帧,并通过真实帧与预测帧之间的差异判断视频中是否存在异常。考虑到视频异常检测任务中空间信息和运动信息的重要性,在模型训练的过程中分别引入了图像注意力强度损失函数、注意力梯度损失函数、光流损失函数和对抗训练损失来保证预测帧和真实帧在外观以及运动信息上的一致性。

为了最小化预测帧与真实帧之间的差距,本文将掩膜 M 引入到图像注意力强度损失函数 ℓ_{mit} 中,以此来约束模型的训练:

$$\ell_{\text{mit}}(\hat{I}_t, I_t) = \|M \cdot (\hat{I}_t - I_t)\|_2^2 \quad (4)$$

式中: I_t 表示预测帧, $\hat{I}_t$ 表示真实帧, “ $\cdot$ ” 表示对应位置的乘法操作。

强度损失捕获了主要的内容信息,但是预测帧的模糊还是不可避免,为了解决这个问题,在模型训练的过程中引入了图像注意力梯度损失函数 ℓ_{mgd} :

$$\ell_{\text{mgd}}(\hat{I}_t, I_t) = \sum_{i,j} \|M \cdot (|\hat{I}_t^{i,j} - \hat{I}_t^{i-1,j}| - |I_t^{i,j} - I_t^{i-1,j}|) + M \cdot (|\hat{I}_t^{i,j} - \hat{I}_t^{i,j-1}| - |I_t^{i,j} - I_t^{i,j-1}|)\|_1 \quad (5)$$

式中: i, j 表示视频帧的空间索引。

2.4 对抗训练

为了使预测帧更加真实,本文引入生成对抗网络。它包含一个生成网络 G 和一个判别网络 D , G 的目的学习生成难以被 D 判别的帧,而 D 的目的是能够判别出 G 生成的帧。在实践中,对抗训练是以一种交替更新的方式来实现的。

2.4.1 生成器的训练

当训练生成器 G 时,固定鉴别器 D 的参数,生成器的对抗损失函数 ℓ_{adv}^G 定义如下:

$$\ell_{\text{adv}}^G = \sum_{i,j} \frac{1}{2} (D(\hat{I}_t^{i,j}) - 1)^2 \quad (6)$$

式中: i 和 j 表示视频帧的空间索引。

基于上述约束函数,我们生成器的最后目标函数设计如下:

$$\ell_G = \lambda_{\text{mit}} \ell_{\text{mit}}(\hat{I}_{t+1}, I_{t+1}) + \lambda_{\text{mgd}} \ell_{\text{mgd}}(\hat{I}_{t+1}, I_{t+1}) + \lambda_f \ell_f + \lambda_{\text{adv}} \ell_{\text{adv}}^G \quad (7)$$

式中: $\lambda_{\text{mit}}, \lambda_{\text{mgd}}, \lambda_f, \lambda_{\text{adv}}$ 表示各个损失函数的权重; ℓ_f 代表光流损失,光流信息是利用 FlowNet^[17] 网络来估计的,FlowNet 是预先训练好的,它的所有参数都是固定的,它可以用来约束动作信息的学习,作为基于 RGB 内容损失的补充。

2.4.2 鉴别器的训练

当训练鉴别器时,固定生成器的参数,鉴别器的对抗损失函数 ℓ_D 设计如下:

$$\ell_D(\hat{I}_t, I_t) = \sum_{i,j} \frac{1}{2} \{ (D(\hat{I}_t^{i,j}) - 0)^2 + (D(I_t^{i,j}) - 1)^2 \} \quad (8)$$

式中: $\hat{I}_t^{i,j} = G(I_1, I_2, \dots, I_{t-1})$, i, j 表示空间索引。

2.5 异常评价分数

在测试阶段,利用训练后的生成器来预测未来帧。受文献[10]中的异常评价方法启发,本文通过计算预测帧与真实帧之间的 PSNR 来衡量两者之间的相似程度,达到异常检测的目的。考虑到异常仅发生在感兴趣区域,本文仅计算显著性区域的 PSNR 值,结合 PSNR 计算公式本文设计的异常评价分数如下:

$$RS(\hat{I}_t) = \frac{PSNR(M_{\text{psnr}} \cdot \hat{I}_t, M_{\text{psnr}} \cdot I_t) - P_{\text{SNR}_{\text{min}}}}{P_{\text{SNR}_{\text{max}}} - P_{\text{SNR}_{\text{min}}}} \quad (9)$$

式中: $P_{\text{SNR}_{\text{max}}}, P_{\text{SNR}_{\text{min}}}$ 分别表示测试视频中 PSNR 值最大值和最小值; M_{psnr} 表示视频剪辑中小数据集对应的显著性区域的掩膜; $PSNR(\cdot)$ 为计算 PSNR 的函数。 $PSNR(\cdot)$ 计算如下:

$$PSNR(\hat{I}, I) = 10 \log_{10} \frac{[\max_i]_i^2}{\frac{1}{N} \sum_{i=0}^N (I^i - \hat{I}^i)^2} \quad (10)$$

在测试阶段,根据视频帧的得分 $RS(\hat{I}_t)$ 来预测它是正常还是异常,如图 4 所示,当视频中出现异常(汽车,自行车闯入),对应的 $RS(\hat{I}_t)$ 降低。因此,可以设置一个阈值来区分规则帧和不规则帧。

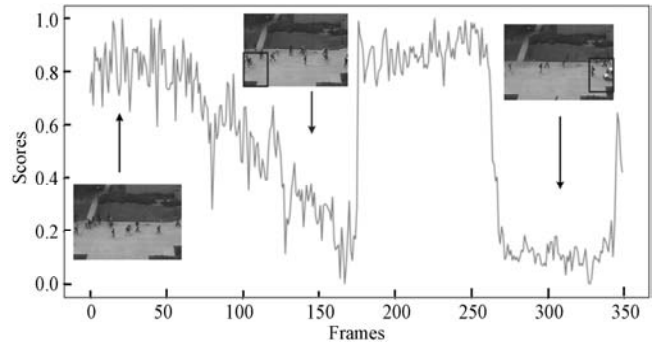


图 4 Ped2 数据集中一段视频序列对应的 RS

3 实验

本节使用公开的异常检测数据集来评估提出的异常检测方法,包括 UCSD^[18] 行人数据集和 CUHK^[19] 大道数据集,UCSD 又分为 Ped1 和 Ped2 两个子数据集。其中:训练数据集中只包含正常事件,测试数据集包含异常事件和正常事件。图 5 给出了含有异常事件的数

据集的部分示例。

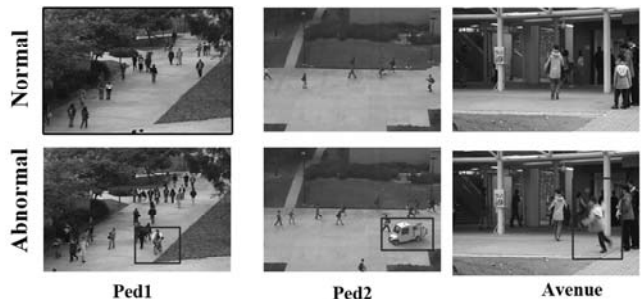


图5 带有异常事件的数据集示例

3.1 参数设置

为了更好地训练网络,在实验的过程中,本文将视频中所有帧的像素强度归一化到 $[-1,1]$ 区间,调整视频帧的分辨率为 256×256 。对于数据集 Ped1 和 Ped2,生成器和鉴别器的学习率分别设置为 0.000 1 和 0.000 01,而对于 avenue 数据集,生成器和鉴别器的学习率分别设置为 0.000 2 和 0.000 02。对于不同的数据集, λ_{mit} 、 λ_{mgd} 、 λ_f 、 λ_{adv} 的系数也略有不同。通过实验发现,当 $\lambda_{mit}=1$ 、 $\lambda_{mgd}=1$ 、 $\lambda_f=0.01$ 、 $\lambda_{adv}=0.05$ 时,在数据集 ped1 能够取得最好的实验效果;而对于数据集 ped2 和 avenue,当 $\lambda_{mit}=1$ 、 $\lambda_{mgd}=1$ 、 $\lambda_f=2$ 、 $\lambda_{adv}=0.05$ 时能达到最佳的实验效果。在测试阶段利用 ROC 曲线获取的最佳阈值作为判别规则帧和不规则帧的阈值,经实验验证,本文在 Ped1、Ped2、Avenue 数据集上的阈值分别设为 0.700 1、0.651 9 和 0.692 7。

3.2 与相关方法的对比

本文利用误报率与召回率绘制接收者操作曲线 (ROC 曲线) 和曲线下面积 (AUC) 值作为模型评价的指标。ROC 曲线下的面积代表 AUC 值,值越大表示异常样本越有可能排在负样本之前,即模型的判别效果越好。

为了验证本文方法的有效性,将本文方法与算法 MPPCA^[1]、MDT^[2]、3DConv-AE^[7]、ConvLSTM-AE^[20]、Stacked RNN^[5]、Pred_Unet^[10]、Driven_loss^[21]、Dual-Discriminator^[12]、Predict-reconstruction^[22] 进行实验对比。表 1 给出了在 Ped1、Ped2 和 Avenue 三个数据集上的对比结果。通过表 1 中给出的 AUC 值可以发现基于手工制作特征的视频异常检测方法 AUC 值较低;而基于重构误差及预测未来帧进行异常检测的方法, AUC 值有了明显的提高。此外,在 Dual-Discriminator 方法中, Ped1 数据集上的异常检测的准确率要高于本文方法,主要原因是该方法在鉴别器中输入了视频帧的光流图,相比于本文方法增加了模型的复杂度以及训练时间,并且在其他数据集上本文方法优于该方法。综上,可以说明本文方法的性能要优于其他方法,证明

了本文方法对异常检测的有效性。

表 1 不同方法在三个常用数据集上的 AUC(%)

方法	Ped1 AUC	Ped2 AUC	Avenue AUC
MPPCA ^[1]	59.0	69.3	N/A
MDT ^[2]	81.8	82.9	N/A
3DConv-AE ^[18]	75.0	85.0	80.0
ConvLSTM-AE ^[19]	75.5	88.1	77.0
Stacked RNN ^[5]	N/A	92.2	81.7
Pred_Unet ^[10]	83.1	95.4	85.1
Driven_loss ^[21]	83.9	96.0	86.0
Dual-Discriminator ^[12]	84.9	95.5	N/A
Prediction reconstruction ^[22]	82.6	96.2	83.7
本文方法	84.3	96.2	86.7

图 6、图 7 给出了本文方法与目前主流方法 Pred_Unet^[10]、Driven_loss^[21] 在 Avenue、Ped2 数据集上的 ROC 曲线。可以发现本文方法的 ROC 曲线在大多数阈值下都优于其他两种方法。

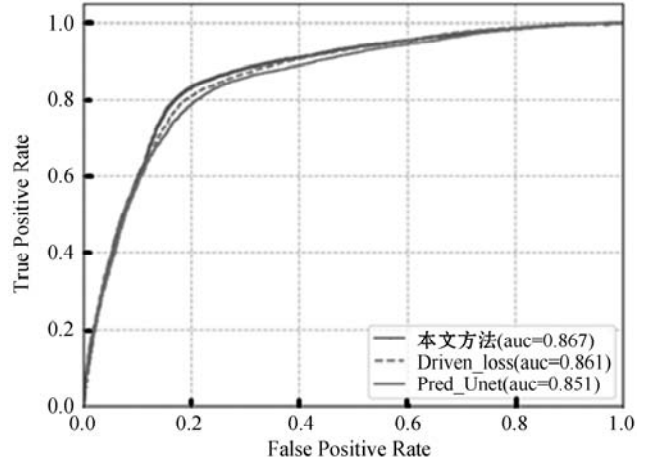


图 6 Avenue 数据集上的 ROC 曲线对比

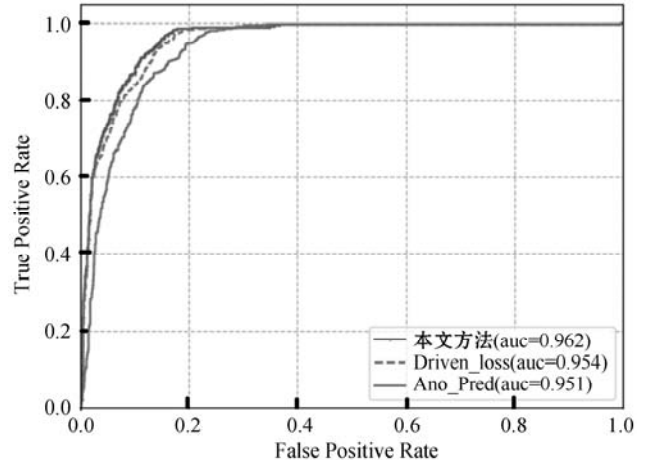


图 7 Ped2 数据集上的 ROC 曲线对比

为了进一步验证本文方法的有效性,本文计算正常帧和异常帧 RS 的平均值,然后将两者之间的差值

表示为 ΔS , ΔS 值越大意味着网络区分正常帧和异常帧能力越强。通过图 8 可以看出,本文方法的 ΔS 比 Pred_Unet 更大,说明本文方法在辨别正常帧和异常帧的能力上更强。

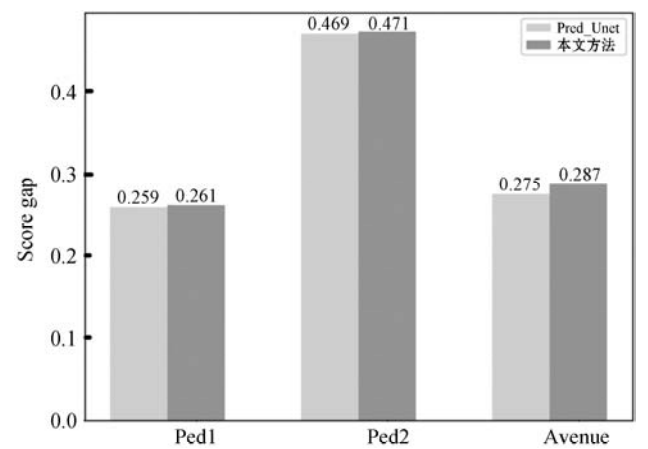


图 8 正常帧和异常帧 RS 的平均值之间的差值

3.3 消融实验

为了证明本文提出的数据预处理方法对异常事件检测的重要性,设计两组消融实验,具体如下:

1) 掩膜对模型性能的影响分析。通常,监控视频中的异常目标只出现在视频帧的局部区域,然而现有的视频异常检测算法大多选择整幅图像的总误差作为异常评价的标准,抑制局部异常的同时,也增强了背景噪声的影响。本文提出仅计算显著性区域的预测误差来进行异常检测,为此,设计一组消融实验来验证将掩膜引入到异常评价分数对模型性能的影响,通过表 2 可以发现模型的 AUC 值提高大约 1.4%,验证了将掩膜引入到异常评价分数设计的必要性。

表 2 掩膜对模型性能的影响(%)

模型	Ped1	Ped2	Avenue
With mask PSNR	84.3	96.2	86.7
Without mask PSNR	83.0	94.6	85.7

2) 图像注意力损失函数对模型性能的影响分析。引入图像注意力损失函数的目的是缓解图像前景与背景不平衡的问题,使得设计的预测网络更关注前景区域的建模,忽略背景信息产生的噪声,影响预测帧的生成质量。为了验证注意力损失函数对模型性能的影响,在 Ped1、Ped2、Avenue 数据集上设计一组消融实验,实验结果如表 3 所示(为了节省时间,这里的结果是迭代一万次得到的),可以看出,本文设计的注意力损失函数会使得预测模型异常检测准确率提高,验证了注意力损失函数对异常检测的有效性,这也说明了本文方法可以更好地关注视频中的前景区域,缓解了数据中图像的前景与背景不平衡的问题。

表 3 注意力损失函数对模型性能影响(%)

模型	Ped1	Ped2	Avenue
With attention loss	83.6	94.3	85.2
Without attention loss	82.9	93.9	84.4

4 结 语

由于目前视频异常检测算法对图像所有区域同等对待,这会造成一个数据不平衡的问题,本文针对此问题,提出一种基于 SE-U-Net 预测网络的视频异常检测方法。本文方法首先提取视频帧的显著性图,并将其制作成掩膜对数据进行预处理;然后利用预处理后的数据对预测模型进行训练,为了使预测模型更关注前景区域的优化,本文结合注意力机制设计一组新的损失函数用于约束模型的训练;此外,在测试阶段,本文仅计算显著性区域的预测误差,进一步提高异常检测的准确性,并通过相关实验结果和消融研究验证本文方法的有效性。

参 考 文 献

[1] Kim J, Grauman K. Observe locally, infer globally: A space-time MRF for detecting abnormal activities with incremental updates [C]//IEEE International Conference on Computer Vision and Pattern Recognition, 2009: 2921 - 2928.

[2] Mahadevan V, Li W X, Bhalodia V, et al. Anomaly detection in crowded scenes[C]//IEEE International Conference on Computer Vision and Pattern Recognition, 2010: 1975 - 1981.

[3] Sabokrou M, Fayyaz M, Fathy M, et al. Deep-cascade: Cascading 3D deep neural networks for fast anomaly detection and localization in crowded scenes[J]. IEEE Transactions on Image Processing, 2017, 26(4): 1992 - 2004.

[4] Sabokrou M, Pourreza M, Fayyaz M, et al. AVID: Adversarial visual irregularity detection[C]//Asian Conference on Computer Vision, 2018: 488 - 505.

[5] Luo W X, Liu W, Gao S H. A revisit of sparse coding based anomaly detection in stacked RNN framework[C]//IEEE International Conference on Computer Vision, 2017: 341 - 349.

[6] Schlegl T, Seebock P, Waldstein S M, et al. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery[C]//International Conference on Information Processing in Medical Imaging, 2017: 146 - 157.

[7] Hasan M, Choi J, Neumann J, et al. Learning temporal regularity in video sequences: Supplementary material [C]//IEEE International Conference on Computer Vision and Pat-

- tern Recognition, 2016:733 – 742.
- [8] Chong Y S, Tay Y H. Abnormal event detection in videos using spatiotemporal autoencoder[C]//International Symposium on Neural Networks, 2017:189 – 196.
- [9] Ravanbakhsh M, Nabi M, Sangineto E, et al. Abnormal event detection in videos using generative adversarial nets [C]//IEEE on International Conference on Image Processing, 2018.
- [10] Liu W, Luo W X, Lian D Z, et al. Future frame prediction for anomaly detection—A new baseline[C]//IEEE International Conference on Computer Vision and Pattern Recognition, 2018:6536 – 6545.
- [11] Song H, Sun C, Wu X, et al. Learning normal patterns via adversarial attention-based autoencoder for abnormal event detection in videos[J]. IEEE Transactions on Multimedia, 2020, 22(8) :2138 – 2148.
- [12] Dong F, Zhang Y, Nie X S. Dual discriminator generative adversarial network for video anomaly detection[J]. IEEE Access, 2020, 8:88170 – 88176.
- [13] Zhao B, Wu X, Feng J S, et al. Diversified visual attention networks for fine-grained object classification [J]. IEEE Transactions on Multimedia, 2017, 19(6) :1245 – 1256.
- [14] Liu K, Ma H. Exploring Background-bias for anomaly detection in surveillance videos [C]//27th ACM International Conference on Multimedia, 2019:1490 – 1499.
- [15] Liu Z, Li J H, Ye L W, et al. Saliency detection for unconstrained videos using superpixel-level graph and spatiotemporal propagation[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2017, 27(12) :2527 – 2542.
- [16] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015:234 – 241.
- [17] Dosovitskiy A, Fischer P, Ilg E, et al. Flownet: Learning optical flow with convolutional networks[C]//IEEE International Conference on Computer Vision, 2015:2758 – 2766.
- [18] Lu C, Shi J P, Jia J. Abnormal event detection at 150 FPS in MATLAB [C]//IEEE International Conference on Computer Vision, 2013:2720 – 2727.
- [19] Mahadevan V, Li W X, Bhalodia V, et al. Anomaly detection in crowded scenes[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010:1975 – 1981.
- [20] Luo W X, Liu W, Gao S H. Remembering history with convolutional LSTM for anomaly detection[C]//IEEE International Conference on Multimedia and Expo, 2017:439 – 444.
- [21] Zhou J T, Zhang L, Fang Z W, et al. Attention-driven loss for anomaly detection in video surveillance[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 30(12) :4639 – 4647.
- [22] Tang Y, Zhao L, Zhang S, et al. Integrating prediction and resconstruction for anomaly detection[J]. Pattern Recognition Letters, 2020, 129:123 – 130.
- ~~~~~
- (上接第 85 页)
- [9] Chen L, Li X, Ji H, et al. Computation offloading balance in small cell networks with mobile edge computing[J]. Wireless Networks, 2019, 25(7) :4133 – 4145.
- [10] Ma X, Lin C, Zhang H, et al. Energy-aware computation offloading of IoT sensors in cloudlet-based mobile edge computing[J]. Sensors, 2018, 18(6) :1945.
- [11] You C S, Huang K B, Chae H, et al. Energy-efficient resource allocation for mobile-edge computation offloading[J]. IEEE Transactions on Wireless Communications, 2016, 16(3) :1397 – 1411.
- [12] Tran T X, Pompili D. Joint task offloading and resource allocation for multi-server mobile-edge computing networks[J]. IEEE Transactions on Vehicular Technology, 2018, 68(1) :856 – 868.
- [13] Park L, Lee C, Na W, et al. Two-stage computation offloading scheduling algorithm for energy-harvesting mobile edge computing[J]. Energies, 2019, 12(22) :4367.
- [14] Min M H, Xiao L, Chen Y, et al. Learning-based computation offloading for IoT devices with energy harvesting[J]. IEEE Transactions on Vehicular Technology, 2019, 68(2) :1930 – 1941.
- [15] Mao Y, Zhang J, Letaief K B. Dynamic computation offloading for mobile edge computing with energy harvesting devices [J]. IEEE Journal on Selected Areas in Communications, 2016, 34(12) :3590 – 3605.
- [16] Neely M J. Stochastic network optimization with application to communication and queueing systems[J]. Synthesis Lectures on Communication Networks, 2010, 3(1) :201 – 211.
- [17] Wang F, Xu J, Wang X, et al. Joint offloading and computing optimization in wireless powered mobile-edge computing systems[J]. IEEE Transactions on Wireless Communications, 2017, 17(3) :1784 – 1797.
- [18] Zhang Z K, Wu J G, Chen L, et al. Collaborative task offloading with computation result reusing for mobile edge computing[J]. The Computer Journal, 2019, 62(10) :10.
- [19] Zhao T C, Zhou S, Song L Q, et al. Energy-optimal and delay-bounded computation offloading in mobile edge computing with heterogeneous clouds[J]. China Communications, 2020, 17(5) :191 – 210.
- [20] Bai T, Pan C H, Deng Y S, et al. Latency minimization for intelligent reflecting surface aided mobile edge computing [J]. IEEE Journal on Selected Areas in Communications, 2020, 38(11) :2666 – 2682.