

Research on Pedestrian Re-recognition in Public Video Surveillance Images

Siyuan Li

Southeast University, Nanjing, Jiangsu, 210018, China

Abstract

The research on pedestrian re-recognition in public video surveillance images aims to solve the problem of identifying and tracking pedestrians in public video surveillance images. This study mainly involves extracting pedestrian features from surveillance videos and achieving cross camera and cross scene pedestrian re-recognition by comparing feature information. The core issue of this study is how to effectively extract and compare pedestrian features. Traditional pedestrian re-recognition methods are usually based on manually extracted features such as shape, texture, and behavioral features. However, these methods often struggle to cope with issues such as changes in lighting, perspective, pedestrian posture and clothing, and occlusion. To address these issues, deep learning technology has been widely applied in pedestrian re-recognition tasks.

Keywords

public video; pedestrian re-recognition; deep learning technology

公共视频监控图像中的行人重识别研究

李思源

东南大学, 中国·江苏 南京 210018

摘要

公共视频监控图像中的行人重识别研究旨在解决在公共视频监控图像中识别和追踪行人的问题。该研究主要涉及从监控视频中提取行人特征,通过比对特征信息,实现跨摄像头、跨场景的行人重识别。该研究的核心问题是如何有效地提取和比较行人的特征。传统的行人重识别方法通常基于手工提取的特征,如形状、纹理和行为特征等。然而,这些方法往往难以应对光照变化、视角变化、行人姿态和衣着变化以及遮挡等问题。为了解决这些问题,深度学习技术被广泛应用于行人重识别任务中。

关键词

公共监控; 行人重识别; 深度学习技术

1 引言

行人重识别的研究涉及如何处理遮挡问题。一种可能的方法是使用光流估计和运动补偿来推断被遮挡的行人特征,还可以通过分析行人的运动模式来解决视角变化和光照变化问题。在实际应用中,行人重识别技术可以用于安全监控、智能交通系统等领域。例如,通过行人重识别技术,可以在公共场所自动识别和追踪特定行人的行动轨迹,提高安全性和效率。总的来说,公共视频监控图像中的行人重识别研究是一个具有挑战性和实用性的课题。随着深度学习技术的发展和应用,行人重识别的准确性和可靠性将得到进一步提升。

2 深度学习技术的概念

深度学习是机器学习领域中的一个新的研究方向,它

被引入机器学习使其更接近于最初的目标——人工智能^[1]。深度学习是学习样本数据的内在规律和表示层次,这些学习过程中获得的信息对诸如文字、图像和声音等数据的解释有很大的帮助。它的最终目标是让机器能够像人一样具有分析学习能力,能够识别文字、图像和声音等数据。深度学习是一个复杂的机器学习算法,在语言和图像识别方面取得的效果,远远超过先前相关技术。它在搜索技术、数据挖掘、机器学习、机器翻译、自然语言处理、多媒体学习、语音、推荐和个性化技术以及其他相关领域都取得了很多成果。

3 深度学习在行人重识别中的最新发展和应用

3.1 卷积神经网络 (CNN)

卷积神经网络是一类包含卷积计算且具有深度结构的前馈神经网络,是深度学习的代表算法之一^[2]。它被广泛应用于图像识别、语音识别、自然语言处理等领域。CNN模型的特点包括局部感知、权重共享和分层次特征采集。局部感知是指网络中的每个节点只与输入图像的局部区域相连,

【作者简介】李思源(1995-),男,中国山东五莲人,在读博士,从事图像处理研究。

这使得网络可以更好地捕捉局部特征。权值共享是指卷积神经网络中所有神经元的权值都是共享的，这大大减少了模型的参数数量。分层次特征采集是指 CNN 模型分为多个层次，每个层次都负责提取不同的特征，从低级的像素特征到高级的语义特征。CNN 模型的基本结构包括卷积层、池化层和全连接层。卷积层负责提取输入图像的特征，池化层负责降低特征的维度，全连接层则负责将前面两层提取到的特征进行整合，输出最终的分类结果。CNN 模型在很多领域都有广泛的应用，例如在图像识别领域，CNN 可以用于分类、分割和目标检测等任务；在语音识别领域，CNN 可以用于语音特征提取和语音建模等任务；在自然语言处理领域，CNN 可以用于文本分类、情感分析等任务。在行人重识别任务中，CNN 可以用于提取行人的外观特征，包括形状、纹理和行为特征等^[3]。通过训练 CNN 模型，可以使得其自动学习到更加鲁棒的特征表达，从而提高行人重识别的准确性。

3.2 循环神经网络 (RNN)

循环神经网络是一种深度学习模型，它以序列数据为输入，并在序列的演进方向进行递归。RNN 模型在序列数据的处理上有很强的优势，可以处理时间序列数据、语音数据、自然语言数据等。RNN 模型有两种主要的变体：一种是时间递归神经网络，另一种是结构递归神经网络。时间递归神经网络是指序列的当前输出与前面的输出有关，而结构递归神经网络则是根据人的认知是基于过往的经验和记忆这一观点提出的。在 RNN 模型中，所有的节点（循环单元）都是按照链式连接的，形成一个环状结构。RNN 模型的主要特点是它能够记忆之前的输入信息，并将其作为当前输入的一部分，这使得 RNN 模型能够处理具有时序依赖性的数据。RNN 模型在自然语言处理、语音识别、机器翻译等领域都有广泛的应用。例如，在机器翻译中，RNN 可以用于编码源语言句子并将其转换为目标语言句子；在语音识别中，RNN 可以用于将语音信号转换为文本^[4]。总之，循环神经网络是一种强大的深度学习模型，适用于处理序列数据，具有记忆性和时序依赖性等特点。

3.3 生成对抗网络 (GAN)

生成对抗网络是一种深度学习模型，由一个生成器和一个判别器组成。生成器的任务是生成逼真的图像数据，而判别器的任务是判断输入的图像数据是否真实。在训练过程中，生成器和判别器会进行对抗性的训练，最终使得生成器能够生成出足以欺骗判别器的图像数据。GAN 的应用非常广泛，如在图像生成、图像修复、风格迁移等方面都有应用。GAN 还可以用于数据增强，通过生成新的数据来增加训练集的数量，从而提高模型的泛化能力。此外，GAN 还可以用于图像分割、目标检测等任务^[5]。生成对抗网络是一种非常强大的深度学习模型，具有广泛的应用前景。它通过生成器和判别器的对抗性训练，能够生成出高质量的

图像数据，并且在很多计算机视觉任务中都取得了很好的效果。

总之，深度学习在行人重识别任务中具有重要的作用。通过使用深度学习技术，可以自动学习行人的特征表达，处理复杂的行人姿态、衣着和遮挡等问题，从而提高行人重识别的准确性。

4 行人重识别的挑战

4.1 图像分辨率低

公共视频监控系统中的摄像头通常具有固定的分辨率，而高分辨率的摄像头可能会导致存储空间和传输带宽的需求增加。在将监控视频传输到存储设备或监控中心的过程中，通常会使用压缩技术来减少传输的数据量。然而，这种压缩可能会导致图像质量的损失，包括降低图像的分辨率^[6]。在进行行人重识别等视频处理任务时，为了减少计算量和存储空间的需求，通常会对视频进行降采样处理，即减少每个像素点的数量，从而降低图像的分辨率。

4.2 视角变化

不同的摄像头有不同的拍摄视角，因此从不同的摄像头拍摄到的行人图像可能会有不同的视角。例如，有的摄像头可能只能拍摄行人的侧面或背面，而无法拍摄到行人的正面或完整的图像。行人的姿态和衣着也会随着时间的变化而发生变化，例如行人的行走姿势、服装搭配等。这些变化可能会导致在相邻的两帧图像中，行人的姿态和衣着出现较大的差异，从而影响重识别的准确性。

4.3 光照变化

在行人重识别中，光照变化是一个常见的问题。光照变化会导致图像中的亮度、阴影和反射等信息发生变化，从而改变行人的外观特征。这使得行人重识别变得更加困难，因为模型需要更好地处理光照变化带来的挑战。光照变化的原因可能包括自然光线的变化、不同时间段的照明条件差异、摄像头视角和位置的变化等^[7]。例如，在跨越摄像头的行人重识别任务中，由于不同摄像头放置的位置和角度不同，拍摄到的行人图像的光照条件可能会有所不同，这可能导致模型难以准确匹配不同图像中的行人。为了减轻光照变化对行人重识别的影响，可以尝试采用一些方法，如使用合适的光照归一化方法、增加数据集的多样性等。此外，还可以通过改进模型的结构和训练方法，提高模型对光照变化的适应性。例如，可以使用注意力机制来增强模型对行人关键部位的关注度，从而提高模型在光照变化情况下的准确性。

为了解决这些挑战，研究者们正在不断探索新的深度学习模型和方法。例如，使用注意力机制来增强模型对行人关键部位的关注度；使用数据增强技术来增加训练数据量，提高模型的泛化能力；使用 GAN 来生成与目标行人相似的图像，从而增加训练数据的多样性等。这些方法为解决行人重识别中的挑战提供了新的思路和方向。

5 公共视频监控图像中的行人重识别方法

公共视频监控图像中的行人重识别方法主要包括以下步骤:

①采集数据:从监控摄像头获取原始视频数据。这些摄像机通常位于不同环境下的不同地方,最有可能的是,这些原始数据包含了大量复杂且有噪声的背景杂波。

②生成边界框:从原始视频数据中提取包含人物图像的包围盒。边界框通常由行人检测或跟踪算法获得。

③标注训练数据:跨相机进行标注。由于摄像机之间的差异很大,训练数据标注对于有区别的重复识别模型学习通常是必不可少的。

④训练模型:用先前标注的人物图像/视频训练一个有区别的、健壮的重识别模型。

⑤行人检索:给定一个感兴趣的人(查询)和一个图库集,使用前一阶段学习的重识别模型提取特征表示,计算与图库的相似度并由高到低进行排序,即可获取检索到的排名列表。

以上方法主要基于深度学习技术,如卷积神经网络(CNN)、循环神经网络(RNN)、生成对抗网络(GAN)等,可以自动学习图像中的特征表达,处理复杂的行人姿态、衣着和遮挡等问题,从而提高行人重识别的准确性。同时,注意力机制可以增强模型对行人关键部位的关注度,数据增强技术可以增加训练数据的多样性,提高模型的泛化能力。

6 结论

公共视频监控图像中的行人重识别研究是一个具有挑战性和实用性的课题。尽管深度学习技术已经取得了很大的进展,并在行人重识别任务中展现出了优越的性能,但仍然存在一些挑战和问题需要进一步解决。首先,跨视角行人重识别是一个重要的挑战。由于不同摄像头拍摄到的行人图像

视角不同,行人姿态、衣着等特征可能会发生变化,这给准确识别行人带来了困难。因此,需要研究更好的方法来提取行人的特征表示,并提高模型对不同视角下行人图像的适应性。其次,遮挡也是行人重识别中的一个重要问题。在实际视频监控场景中,行人可能会出现不同程度的遮挡,这使得提取行人的外观特征变得更加困难。为了解决这个问题,可以尝试使用光流估计和运动补偿等方法来推断被遮挡的行人特征。此外,算法的鲁棒性也是一个需要考虑的问题。由于视频监控图像通常会受到光照变化、噪声等因素的干扰,这会对模型的性能产生影响。因此,需要研究更好的方法来提高模型的鲁棒性,以应对这些干扰因素。最后,未来的研究需要进一步关注这些问题,并探索更有效的解决方法。同时,还需要进一步改进行人重识别的技术,提高其准确性和稳定性,以便更好地应用于安全监控、智能交通系统等领域。

参考文献

- [1] 王俊丽,郭福荣.基于深度学习的行人重识别研究[J].计算机应用研究,2021,38(11):1-5.
- [2] 张明,赵鹏飞,王坤.基于卷积神经网络的行人重识别研究[J].计算机工程与应用,2020,56(16):1-7.
- [3] 负俊豪,李文杰.基于跨摄像头跨场景的行人重识别研究[J].计算机工程与应用,2020,56(12):1-7.
- [4] Liao S, Hu Y, Zhu X ,et al.Person re-identification by Local Maximal Occurrence representation and metric learning[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),IEEE, 2015.
- [5] MA B, SU Y, JURIEF. Local descriptors encoded by fisher vectors for person re-identification[C]// ICCV, 2015.
- [6] 宋婉茹,赵晴晴,陈昌红,等.行人重识别研究综述[J].智能系统学报,2017,12(6):770-780.
- [7] 张化祥,刘丽.行人重识别研究综述[J].山东师范大学学报:自然科学版,2018,33(4): 379-387.