**北京工业大学研究生开题报告**

**学位级别： □博士 ■硕士**

**学 号： S202274129**

**研究生姓名： 张一鸣**

**指导教师姓名： 贾熹滨**

**专业名称： 计算机技术**

**所在学院： 信息学部**

**开题报告时间： 2023/12/19**

**北京工业大学研究生院制表**

**注意**： 本表基本情况及报告正文由研究生本人填写，硕士不少于3000字，博士不少于5000字。格式要求：正文文字部分为5号宋体、单倍行间距排版，A4纸双面打印装订。

开题报告评价部分分别由指导教师及专家组书写。开题报告会结束后一周之内将报告原件交院（所）研究生教学秘书处。

**一、基本情况**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **研究生姓名** | | 张一鸣 | | **学 号** | S202274129 |
| **院、系** | | 信息学部 | | **指导教师姓名及职称** | 贾熹滨（教授） |
| **学科、专业** | | 计算机技术 | | **入学年月** | 2022年9月 |
| **校外指导老师姓名及职称** | | 崔国勤（教授） | | **填表日期** | 2023年12月 |
| 1. **研究方向、论文选题范围：**   研究方向：医疗影像处理  论文选题范围：深度学习、计算机视觉、域泛化   1. **拟定论文题目：**   医疗影像域泛化分类算法研究   1. **论文科研课题属于哪一级科研项目，经费来源及金额（**课题来源选项分为国家计委、科委项目、国家经贸委项目、国家自然科学基金项目、国务院其他部门项目、主管部门（部委级）项目、省、市、自治区项目、国际合作项目、学校级项目、自选项目、其它）**：**   经费支持：基于视频的目标跟踪和行为认知技术研究与开发 ，企业横向课题，项目编号：46007013201802，经费：10万   1. **论文类型（基础研究、应用研究、开发研究、其它）**   应用基础研究 | | | | | |
| **摘**  **要** | **选题研究内容和意义简介（限400字）：**  随着越来越来多的摄像头被部署到公共场所中，有大量的视频数据可供研究者研究和利用。行人再识别(Person re-identification，re-id)是利用计算机视觉技术判断视频或者图像数据中是否存在特定行人的技术，也就是在给定的图库（gallery）中去查找特定的(query)行人图像。  在行人再识别的研究过程中，背景噪声严重影响到模型的识别精度，这种误差来源于原始的检测算法，因此在现有检测框架的基础上改进后适用于行人再识别数据集以得到检测效果更好的数据集。同时，针对不同特征的贡献度是不同的，使用特定的特征加权算法赋予每个部分特征权值，放大具有良好判别性的局部特征；最后，针对于具有区分度的细节信息往往存在于小物体中，引入多尺度特征融合机制，加强网络对小物体特征的提取，达到进一步提升准确率的效果。 | | | | |
| **关键词（用分号隔开、最多5个）** | | | **多尺度特征融合；特征加权；行人再识别；细节特征** | | |

**报 告 正 文**

**（一）选题依据与研究内容**

**1、选题依据（专业学位类别或领域的研究意义、国内外研究现状等）**

**1.1 研究意义**

随着科技的发展和社会的进步，数以万记的大型摄像头网络被部署在了火车站、机场、高校校园、和大型娱乐场所等公共场所。这些场所的具有地理空间跨度大、人流量多、并且相互摄像头之间可能并没有重叠区域这样的特点。大量摄像头的部署，给人们提供了海量的视频数据。在智能分析算法没有被应用到这些数据之前，只能靠人工来监控这些视频，既费时又费钱。随后通过将视频分析算法应用到大量的数据上能够确定人们在一个场景中的长期活动和行为特征，使得监控变得更加主动。例如：通过分析一个人的行为轨迹来判断此人的活动是否正常同时也可以预测危害公共安全的事件并及时向安保部门报警。

行人再识别（Person re-identification）也称行人重识别，是利用计算机视觉技术判断图像或者视频序列中是否存在特定行人的技术[1]。广义意义上来说，可以认为行人再识别是一个图像检索的子问题，即给定一个监控摄像头拍摄的行人图像，检索该行人图像是否在其他摄像头拍摄的行人图像中出现过，并且要求给定行人图像的摄像头与检索图像的摄像头没有重叠的监控区域。这样就致使行人再识别的研究面临了很多挑战。首先，不同区域的摄像头存在摄像角度，光照的差异，导致同一个人在不同摄像头的拍摄下所产生的图像差别较大，为后续的图像预处理工作带来很大的挑战[2]；其次，由于行人并没有在提前告知的情况下就进行行人再识别，导致所拍摄到的图像可能会受到人体姿态的影响较大，为最后的识别带来巨大的挑战；最后，行人在场所进行活动时会受到遮挡物的遮挡，导致所拍摄的图像并没有记录下完整的行人，在后期利用卷积神经网络进行特征提取时面临巨大的困难。

综上所述，不同区域的摄像头拍摄的视角、光照、以及人体的姿态的影响，使得提取一个具有鲁棒性很好的特征变得很困难。同时，体态相同的不同行人在穿着相同颜色的衣服时，很容易被识别为同一个人，这样的情况也为行人再识别带来的巨大的挑战[3]。行人再识别本质上是基于人的外观形体的一个相似性比较，主要是人体外观的颜色特征、行走时的步态特征等，所以行人再识别是一个极具挑战性的课题。

**1.2 国内外研究现状**

行人再识别任务应该包括两个任务，首先是从图像或者视频中检测行人，其次是从检测好的行人图像中提取特征，进行相似性度量。但是随着研究方向的不断细化，第一个问题已经作为一个独立课题即行人检测来研究。因此行人再识别技术的研究方向主要分为两类：基于特征的描述方法和基于度量学习法。在近两年，利用改进后的生成对抗网络（GAN）对行人再识别的数据集进行背景的风格迁移以及生成新的行人再识别数据成为一个新的研究方向，也为行人再识别应用到现实生活中提供了大量可靠的数据集。

**1.2.1 基于特征的描述方法**

20世纪90年代开始，人们开始尝试利用贝叶斯的方法来解决多摄像头追踪的问题[4]。到了2006年，行人再识别的概念第一次被提出，同时成为了一个独立的计算机视觉的研究方向[5]。到了2014年，当深度学习的方法加入到行人再识别领域的时候，使得效果得到了大大的提升[6]。传统的行人再识别可以根据数据的格式划分为基于图像的行人再识别的研究和基于视频的行人再识别的研究。由于图像数据和视频数据的格式上的差异，研究方法也出现了很大的差异。基于图像的行人再识别数据为静态的图像，因此需要根据手工设计的算法或者利用深度学习中的卷积神经网络来提取鲁棒的特征。而视频数据则存在着连续的图像帧，可以利用的信息包含了时间信息。因此，研究者们常使用CNN+RNN的组合或者是3DCNN来提取视频数据中的时间信息，以此来加强图像帧之间的联系，得到更加鲁棒的特征，使得识别的效果更加接近于现实生活[7]。

在深度学习没有应用到行人再识别领域之前，研究者们通常利用手工提取的特征来进行行人再识别。D. Gray等人首先提取图像的颜色和纹理特征，将RGB空间的图像转化成HSL和YCbCr颜色空间，观察对数颜色空间中目标像素值的分布，并且颜色特征在不同光照或者角度等行人识别的不同环境中具有一定的不变性[8]。A.Krizhevsky等人通过提取行人的形状特征也取得了不错的效果[9]。由于形状特征如方向梯度直方图（HOG）特征以及布局不变特征SIFT、SURF和Covariance描述子具有一定的鲁棒性，也被很多研究者所使用。

后来的出现的语义属性特征也被多位研究者所使用。研究者发现，可以通过语义信息来判断两张图像中是否属于同一行人，比如颜色、衣服以及携带的包等信息。相同的行人在不同的摄像头的拍摄下，语义属性这个特征很少变化。Layne等人采用15种语义来描述行人，包括鞋子、头发颜色长短、是否携带物品等，分类器用 SVM 定义每幅行人图像的以上语义

属性。结合语义属性重要性加权以及与底层特征融合，最终描述行人图像[10]。Shi 等人对图像超像素划分，最近分割算法对图像块定义多种特征属性，颜色、位置和 SIFT 特征，效果也有明显地提高[11]。

随着深度学习应用的热度不断上升，也被应用到行人再识别领域。由于卷积神经网络具有局部连接和权值共享两大特征，很快就代替手工特征成为研究者提取特征的主要方式。所谓局部连接，就是卷积层的节点仅仅和其前一层的部分节点相连接，只用来学习局部特征。局部感知结构的构思理念来源于动物视觉的皮层结构，其指的是动物视觉的神经元在感知外界物体的过程中起作用的只有一部分神经元。在计算机视觉中，图像中的某一块区域中，像素与像素之间的相关性与像素与像素之间的距离同样相关，距离较近的像素间相关性较强，距离较远则相关性就比较弱，由此可以局部相关理论也适用于计算机视觉的图像处理领域。因此，局部感知采用部分神经元接受图像信息，再通过综合全部的图像信息达到增强图像信息的目的。并且这种局部连接的方式大幅减少了参数数量，加快了学习速率，同时也在一定程度上减少了过拟合的可能。卷积神经网络的另一大特征是权值共享，意味着每一个滤波器在遍历整个图像的时候，过滤器的参数是固定不变的，这样做大大的减少了网络的参数，使得训练更深的卷积神经网络变得可能。利用卷积神经网络提取出来的特征具有平移不变性，从而使得特征的鲁棒性更高。

2012年，McLaughlin 等人利用简单的在ImageNet数据集上预训练的基于AlexNet结构的卷积神经网络，并用这个数据集进行微调，对图像提取颜色和光流特征，用卷积神经网络处理得到高层特征[12]。2016年，T. Xiao等人对来自各个领域的数据训练出同一个卷积神经网络（CNN），有些神经元学习各个领域共享的表征，而其他的神经元对特定的某个区域有效，得到鲁棒的 CNN 特征表示[13]。早期的行人再识别研究的重点还在于如何提取更加鲁棒的全局特征，就是用整张图得到一个特征向量进行图像检索。在2017年，Xuan Zhang等人在论文中提出，同时提取global feature和local feature已经成为行人再识别领域一种常见且比较有效果的一种手法,采用了动态规划来解决行人再识别中常见的misalignment问题，并声称检测精确度已经超过了人类的标注水平[14]。2017年，Yifan Sun等人提出的PCB模型，首先将整张图像输入到基础网络里得到base feature，在基础特征上在height的维度上将特征向量平均分为六份，这一想法假设人体在图像空间上的分布是均匀的[15]。但是这样做的缺点是没有考虑到misalignment，同时也有背景噪声添加到特征提取的过程中，一定程度的影响了最终的精度指标。2018年，Guanshuo Wang等人设计的MGN网络，一举成为各个数据集的得分榜首，其巧妙地设计了三个网络分支同时提取了global和part feature,其中两个part branch 分别将人体划分为两份和三份[16]。同时，网络的损失函数除了常见的分类损失，还加入了triplet loss 约束，将同类距离变得更小，而拉大不同类之间的距离。除了常见的分割条带的方法，还有很多其他比较复杂的body part的划分方式。2018年，Mahdi M. Kalayeh等人则利用人体语义分割模型，将人体划分为头部、上半身、下半身和鞋子四部分，产生这四部分的激活图和提取出的特征向量进行矩阵点乘操作从而得到身体各个部分的特征向量[17]。但是这样做使得行人再识别的精度很大的程度上依赖于人体语义分割的效果，于是作者添加了一条全局分支，用来提取全局特征以此来保证最终的精度。2018年，Maoqing Tian同样利用分割模型将人体分割出来，形成三个输入源，主要探讨了背景信息在行人再识别到底产生的作用[18]。结论表示，带背景训练的模型要比不带背景训练的模型更加鲁棒，说明背景信息在行人再识别的模型中并不全是噪音。人体骨骼关键点对于描述人体姿态，预测人体行为至关重要。因此人体骨骼关键点检测是诸多计算机视觉任务的基础，例如动作分类，异常行为检测，以及自动驾驶等等。近年来，随着深度学习技术的发展，人体骨骼关键点检测效果不断提升，已经开始广泛应用于计算机视觉的相关领域。2018年，M. Saquib Sarfraz等人利用人体关键点检测模型预测人体在图像中相对于摄像机的位置，将视觉角度划分为front、back、以及side三个视角，想解决由不同摄像头摄像角度带来的差异，从而提取出更鲁棒的特征[19]。

**1.2.2 基于度量学习的方法**

由于摄像机的视角、尺度、光照、服饰与姿态变化、分辨率不同以及存在遮挡等问题，不同摄像头间可能会失去连续的位置和运动信息，使用欧氏距离、巴氏距离等标准的距离度量来度量行人外观特征的相似度不能获得很好的重识别效果，因此，研究者们提出通过度量学习的方法。在行人再识别问题上，具体为同一行人的不同图片形似度大于不同行人的不同图片。该方法获得一个新的距离度量空间，使得同一行人不同图像的特征距离小于不同人的距离。距离度量学习方法一般是基于马氏距离而进行。2002 年，Xing 等人提出以马氏距离为基础的度量学习算法，根据样本的类别标签，将具有相同标签的样本组成正样本对，反之组成负样本对，并以此作为约束训练得到一个马氏矩阵，通过这样学习到的距离尺度变换，使得相同的人的特征距离减小，而不同的人特征距离增大，以此开创了行人重识别中距离度量学习的先河[20]。

行人再识别中常见的损失包括分类损失、判别损失和其他像Triplet loss一样的距离损失函数。其中，Triplet loss是非常常用的一种deep metric learning方法，在图像检索领域有着非常广泛的应用，比如人脸识别、行人再识别、商品检索等。传统的Triplet loss训练需要一个三元组，包括三张图片：anchor,positive和negative。传统的Triplet loss的缺点在于随机从训练集中挑选出三张图片，因此可能会出现很简单的样本，导致训练没有意义。因此，在2017年，Alexander Hermans等人在论文中提出，通过在mini batch中挑选难样本的Triplet hard loss是最适合于行人再识别的。由于行人再识别存在图片数据集小、每个行人的图片数量少的特点。Triplet hard loss通过构造三元组可以生成远多于图片数量的样本组，使得模型的效果更好，同时也可以有效地缓解过拟合问题。Zheng等人提出 PRDC 算法，身份相同的人的图像组成同类样本对，不同身份的行人图像之间组成异类样本对，获得度量函数对应的系数矩阵，优化目标函数使得同类样本对之间的匹配距离小于异类样本对之间的距离，对每一个样本，选择一个同类样本和异类样本与其形成三元组，在训练过程通过最小化异类样本距离减去同类样本距离的和，得到满足约束的距离度量矩阵[21]。 Pedagadi 等人提出 LFDA 算法进行度量学习，该方法在进行特征提取的时候，首先提取不同特征的主要成分，然后拼接成特征向量。在距离度量学习上，该方法考虑不是对所有样本点都给予相同的权重，考虑到了局部样本点，应用局部 Fisher 判别分析方法为降维的特征提供有识别能力的空间，提高度量学习的识别率[22]。

**1.2.3 基于生成对抗网络的研究方法**

生成对抗网络是一种深度学习模型，是近年来无监督学习最具前景的方法之一。模型通过框架中的生成模型和判别模型的相互博弈学习产生近似与输入的输出。因此，研究者通常用生成对抗网络做数据增强，主要的解决的问题就是小数据集、数据量不足。因此，生成对抗网络的这个特点被应用到行人再识别的数据集上。2017年，Zheng Z等人首次将生成对抗网络用于行人再识别领域，论文提出一个标签平滑的方法一次来解决随机生成的图像的标签标注问题[23]。2017年，Zhong 等人在利用生成对抗网络的同时考虑到了摄像头之间的差异问题，如角度、光线等，论文利用生成对抗网络将一个摄像头所拍摄的图像风格迁移到另一个摄像头所拍摄的图像中，使得将多个行人再识别的数据集的图像融合到一起变成可能，训练数据量得到大大的提升[24]。行人再识别中一个难点就是每个人的姿态不同，为了克服这个问题，2017年，Qian X等人首先定义了人体的八个姿态，这八个姿态基本涵盖了各个角度。随后利用对抗生成网络对每一张图片都生成这样标准的八个姿态的图片，不但增加了训练样本量，同时也解决了由于人体姿态的差异带来的识别精度上的差异[25]。

综上所述，从行人再识别的三个不同的研究方向描述了国内外目前研究的现状，提出了目前行人再识别领域中如样本量少、摄像头拍摄的图片的风格差异等问题。基于以上文献的经验，本课题主要从基于特征的描述方法以及基于度量学习的行人再识别两个方向展开研究，提出自己的想法。

**2、选题的研究内容、研究目标以及拟解决的关键问题等**

**2.1研究内容**

本课题拟通过利用现有的性能较好的行人检测算法对原有的数据集进行再检测，得到检测效果更加精确的行人再识别数据集，同时采用Se-net模型的特征向量加权的思想使得区分度更明显的特征向量贡献度更大[26]。本课题拟从3个方面进行：

（1）行人再识别背景噪声和misalignment问题的研究

传统的检测算法带来的背景噪声和misalignment为随后的识别算法带来了极大影响。所以，利用现有的效果较好的行人检测算法对数据集进行重新检测，可以极大程度的减小背景噪音的影响，同时可以避免misalignment带来不匹配的问题。

（2）基于part feature的特征加权的研究

将特征分成不同部分后，不同特征在最终识别精确度的贡献度也是不同的。因此，本课题拟采用Se-net网络的思想，对每一个部分特征进行特征加权，使得权重大的特征在最终的识别精度上贡献度更大，从而达到提升识别精度的效果。

（3）多尺度特征融合算法的研究

通过对旧模型分类错误的样本的研究发现，颜色特征已经成为行人再识别系统最主要的识别特征。但是最具有区分度的信息往往存在于小物体中，如行人的背包、行人的鞋子等等。所以对小物体的关注可以更好的区分两个人的身份。实验表明，浅层的网络更关注于细节信息，高层的网络更关注于语义信息。本课题拟采用多尺度特征融合的方式，将高层特征和低层特征进行融合，得到更加具有区分度的特征向量。

**2.2研究目标**

本课题旨在设计一个能够提取具有良好区分度的part feature的框架，采用最新的行人检测算法对数据集进行重新检测，再结合Se-net网络的特征加权思想，进一步放大具有良好判别行的特征；而在基础网络的设计上，采取多尺度特征的融合，加强网络对小物体等细节的关注，得到区分度高的特征向量，从而实现整个行人再识别系统的识别精度的提升。

**2.3拟解决的关键问题**

为了构建一个完整高效的行人再识别系统，实现提取区分度高的特征向量，本课题拟解决的关键问题主要基于以下3个方面：

1. 如何解决数据集本身带来的背景噪声和misalignment

传统的market1501数据集在DPM检测算法的基础上进行预处理，但是DPM算法已经是近十年前的检测算法[27]。通过查看原有数据集可以发现，由于检测效果不准确所带来的背景噪声和misalignment的影响已经为随后的识别算法带来了很大的影响。因此，利用目前最好的行人检测算法对行人再识别的数据集进行重新的检测，进一步的减小由于检测算法带来的误差，从而达到提升识别算法的精度。

1. 如何利用Se-net网络实现部分特征的加权

针对不同part feature对于最终分类结果的贡献度是不同的，因此引入Se-net的思想，通过权值的大小来描述对不同特征的关注程度。权值作为重要组成部分，能够反映不同特征的分类贡献率，需要通过学习的方法获得。所以在整个网络训练的过程中，权值的获取是一个待解决的关键问题。

1. 如何用多尺度特征融合实现对小物体的关注

在行人再识别的识别过程中，具有很高的区分度的信息往往存在于小物体中，因此引入多尺度特征融合实现对细节特征的挖掘。不同阶段的特征存在不同的信息，所以在融合的过程中，不同阶段的特征融合是一个待解决的关键问题。

**3、拟采取的研究方案（研究思路、技术路线或研究方法）及可行性分析**

**3.1研究思路**

如图1所示，本课题的研究将依据数据集的制作与预处理、特征提取、特征加权以及模型测试等步骤进行。

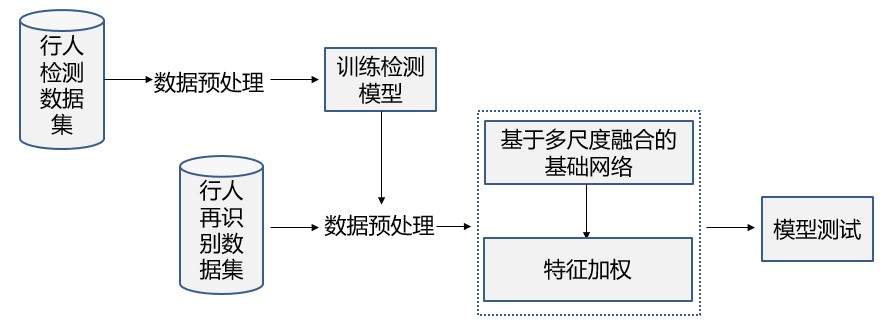


图1研究框架

**3.2技术路线**

**3.2.1数据集的获取与预处理**

**1) 数据集的处理**

Market-1501数据集每张图片的分辨率为64\*128，因此在选择行人检测数据集的时候，应该选择和行人再识别数据集相似的数据集进行模型训练，如在图像分辨率上的相似，行人在图像上的空间分布等等。同时，基于现在效果较好的检查算法如faster-rcnn[28]（二步法）和SSD[29]（一步法）的基础上进行行人检测框架的训练。

**2) 行人再识别数据集的处理**

本课题在行人再识别的数据集的选择上拟采用market-1501数据集，该数据集在清华大学中采集，包括1501个行人、总共44905张图像。每个行人至少由两个摄像头捕获到，并且在一个摄像头中可能具有多张图像。利用训练好的行人检测模型重新检测行人再识别数据集，同时采用插值算法将分割后的图像统一到同一尺寸方便后续处理，最后将整个数据集做归一化处理。

**3.2.2 part feature的提取**

part feature提取的过程如图2和图3所示，卷积神经网络不同层次的特征旨在编码不同层次的信息。高层特征更加关注语义信息，所以分类常采用高层特征。而低层特征包含更多的细节信息。而将高层特征和低层特征进行融合操作，使得更多的细节信息能够得到进一步的分析，对提升行人再识别的精度也有很大的好处。

本课题拟采用的融合方式类似于FPN网络，将高层特征进行上采样操作后与低层特征融合[30]。这里的融合方式与其他融合方式相比，将各层特征充分利用，并没有单独的使用某一层特征进行预测，使得网络可以更多的关注于图像的细节信息。

鉴于此种融合方式与其他融合方式相比具有以上优势，故多尺度特征融合的任务拟采用高低层特征融合的方式。

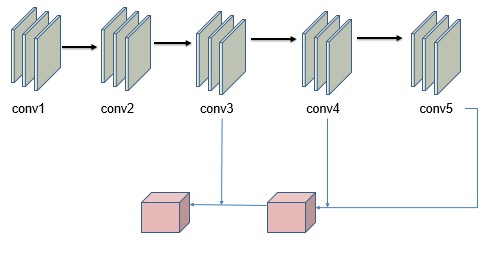
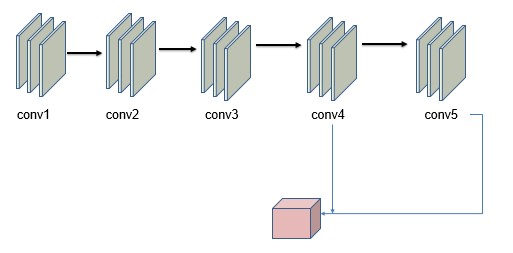


图2 conv3、conv4和conv5特征融合

****图3 conv4和conv5特征融合

在融合高层特征和低层特征的过程中，可以采用不同阶段的特征进行融合。如图2所示，将conv3、conv4和conv5三个阶段的特征融合，以期望得到更多的细节信息。而图3则采用conv4和conv5两个阶段的特征进行融合。本课题拟采取对比实验的手法，找出识别精度最好的特征融合组合。

**3.2.3 特征加权及分类**

特征加权的流程如图4所示，该过程由两部分组成，特征加权模块和分类模块。在特征加权模块中采用Se-net网络的思想，将其中的通道加权的做法变换行人部分特征加权的思想，通过特征加权模块，最终得到每一个部分特征的权值，将权值与特征相乘得到权重特征向量。在保证识别精度的前提下，使得区分度明显的特征向量贡献度更大，达到提升整个系统识别精度的效果。

当得到已经加权的特征向量后，分别再经过全连接层进行分类，得到每一个身份的预测概率值。利用交叉熵损失函数计算每一个特征向量的损失，进而计算出梯度值，利用随机梯度下降算法优化模型，更新整个网络的参数的值，达到训练网络的效果。

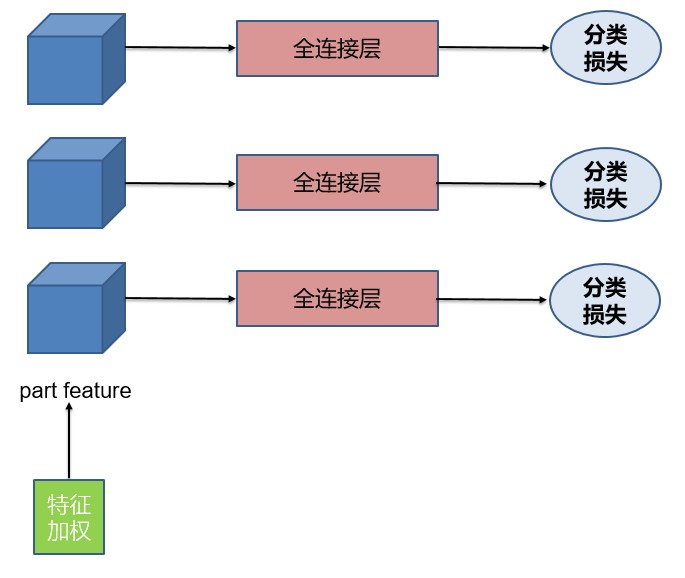


图4 特征加权模块和分类模块

**3.2.4实验验证**

对提取特征的算法采取对比实验的方法来验证本课题提出的算法的有效性，选择几种具有代表性的特征提取算法来进行对比，并将这些算法产生的分类结果和本课题提出的算法产生的分类结果进行对比，以此更加全面的评价该特征生成模型的有效性。除了本课题采用的多尺度特征融合算法，同时可以对比其他常见的融合特征算法。除此之外，对本课题提出的算法采用剥离实验的方式验证每一种算法的有效性。

**3.3可行性分析**

本课题基于part feature的行人再识别的研究，依靠前人的理论基础及实现工具，结合现有的深度学习、模式识别、统计学等方法的研究成果保证了课题的可实现性。采用Se-net和多尺度特征融合提取更有区分度的特征，已有人验证了其有效性。因此，我们提出的引入特征加权和多尺度特征融合模型具有一定的可行性。

**4、论文研究进度计划（含研究进度计划安排、研究经费预算及解决办法）、预期研究结果**

课题研究进度计划安排如下：

|  |  |
| --- | --- |
| 2019年1月～2月 | 收集阅读相关文献及资料，整理检测数据集和re-id数据集。 |
| 2019年3月～6月 | 完成检测模型的测试以及行人再识别网络模型的搭建。 |
| 2019年6月～9月 | 深度学习模型的性能优化和改进，并且预计在mAP的指标上提升2%左右。 |
| 2019年10月～2020年1月 | 通过统计和分析当前分类结果，进一步优化模型，提升分类结果准确率。总结工作，发出小论文一篇。 |
| 2020年1月～4月 | 撰写毕业论文，总结手中工作，组合出效果最好的模型结构。 |
| 2020年5月 | 毕业答辩。 |

预期研究结果：

构建基于part feature的行人再识别模型，同时还有特征加权和多尺度特征融合，得到具有良好判别性的特征，结合行人再识别常见的损失函数，最终实现一个对小物体敏感的行人再识别模型，并发表相关的软件著作、专利和论文。

**（二）论文研究工作基础及条件保障**

1. **工作基础（含入学以来取得研究成果、参与或承担的科研项目情况等）**
2. 入学以来已经完成了模式识别、机器学习、统计分析方法、人工智能等课程的学习；
3. 在此基础上阅读了大量的相关文献，拥有进行此课题研究的理论基础；
4. 在中星微公司实习9个月，具备了一定的行人再识别基础知识，并且搭建模型的能力。
5. **工作条件（设备条件、实验场地条件、可能遇到的困难及应对措施）**
6. 设备条件：笔记本以及GPU服务器
7. 实验场地条件：北京工业大学信西304b实验室；
8. 可能遇到的问题：
9. 知识储备不足：需要不断地阅读领域相关文献，积累研究所需知识；
10. 代码编写困难：需要查阅各项资料，并向导师进行咨询；
11. 实验结果不好：通过调参和优化算法来提升实验结果；
12. 论文撰写问题：可参考优质论文，并向导师咨询请求协助。

**（三）参考文献（参考文献目录用标准规范格式，在正文引用处须标注）**

[1] LI Y, WU Z, KARANAM S, et al. Real-world re-identification in an airport camera network[C]//International Conference on Distributed Smart Cameras. Venice, Italy, 2014: 35.

[2] Zhong Z , Zheng L , Zheng Z , et al. Camera Style Adaptation for Person Re-identification[J]. IEEE Transactions on Image Processing, 2017.

[3] Wang J , Zhu X , Gong S , et al. Transferable Joint Attribute-Identity Deep Learning for Unsupervised Person Re-Identification[J]. 2018.

[4] CAI Q, AGGARWAL J K. Tracking human motion using multiple cameras[C]//International Conference on Pattern Recognition. Vienna, Austria, 1996: 68–72.

[5] GHEISSARI N, SEBASTIAN T B, HARTLEY R. Person re-identification using spatiotemporal appearance[C]// IEEE Conference on Computer Vision and Pattern Recognition. New York, USA, 2006: 1528–1535.

[6] JUNGLING K, BODENSTEINER C, ARENS M. Person re-identification in multi-camera networks[C]//Computer Vision and Pattern Recognition Workshops. Colorado,USA, 2010: 709–716.

[7] WANG T, GONG S G, ZHU X, et al. Person re-identification by video ranking[C]//European Conference on Computer Vision. Zurich, Switzerland, 2014: 688–703.

[8] GRAY D, TAO H. Viewpoint invariant pedestrian recognition with an ensemble of localized features[C]//European Conference on Computer Vision. Marseill, France,2008: 262–275

[9] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]//International Conference on Neural InformationProcessing Systems. Doha, Qatar, 2012: 1097–1105.

[10] LAYNE R, HOSPEDALES T M, GONG S G. Person Re-identification by Attributes[C]//The British Machine Vision Conference. Nottingham, Park, 2014, 2(3): 8.

[11] SHI Z, HOSPEDALSE T M, XIANG T. Transferring a semantic representation for person re-identification and search[C]//Computer Vision and Pattern Recognition. Boston, USA, 2015: 4184–4193.

[12] MCLAUGHLIN N, RINCON J M, MILLER P. Recurrent Convolutional Network for Video-based Person Re-Identification[C]//IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2012: 51–58

[13] XIAO T, LI H, OUYANG W, et al. Learning deep feature representations with domain guided dropout for person re-identification[C]//IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2016: 1249–1258

[14] Zhang X , Luo H , Fan X , et al. AlignedReID: Surpassing Human-Level Performance in Person Re-Identification[J]. 2017.

[15] Sun Y , Zheng L , Yang Y , et al. Beyond Part Models: Person Retrieval with Refined Part Pooling (and a Strong Convolutional Baseline)[J]. 2017.

[16] Wang G , Yuan Y , Chen X , et al. Learning Discriminative Features with Multiple Granularities for Person Re-Identification[J]. 2018.

[17] Kalayeh M M , Basaran E , Gokmen M , et al. Human Semantic Parsing for Person Re-identification[J]. 2018.

[18] Mahdi M. Kalayeh，Emrah Basaran，et al. Eliminating Background-bias for Robust Person Re-identification[J].2018.

[19] Sarfraz M S , Schumann A , Eberle A , et al. A Pose-Sensitive Embedding for Person Re-Identification with Expanded Cross Neighborhood Re-Ranking[J]. 2017.

[20] XING E P, NG A Y, JORDAN M I, et al. Distance metric learning, with application to clustering with side-information[C]//International Conference on Neural Information Processing Systems. Vancouver: MIT Press, 2002:521–528.

[21] ZHENG W, GONG S, XIANG T. Person re-identification by probabilistic relative distance comparison[C]//IEEE conference on Computer Vision and Pattern Recognition. Colorado Springs, USA, 2011: 649–656.

[22] PEDAGADI S, ORWELL J, VELASTIN S, et al. Local fisher discriminant analysis for pedestrian re-identification [C]//IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA, 2013: 3318–3325

[23] Zheng Z, Zheng L, Yang Y. Unlabeled samples generated by gan improve the person re-identification baseline in vitro[J]. arXiv preprint arXiv:1701.07717, 2017.

[24] Zhong Z, Zheng L, Zheng Z, et al. Camera Style Adaptation for Person Re-identification[J]. arXiv preprint arXiv:1711.10295, 2017.

[25] Qian X, Fu Y, Wang W, et al. Pose-Normalized Image Generation for Person Re-identification[J]. arXiv preprint arXiv:1712.02225, 2017.

[26] Hu J , Shen L , Albanie S , et al. Squeeze-and-Excitation Networks[J]. 2017.

[27] Felzenszwalb, P., Girshick, R., McAllester, D., Ramanan, D.: Object Detection with Discriminatively Trained Part-Based Models. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 9 (2010).

[28] Ren S , He K , Girshick R , et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 39(6):1137-1149.

[29] Liu W , Anguelov D , Erhan D , et al. SSD: Single Shot MultiBox Detector[J]. 2015.

[30] Lin T Y , Dollár, Piotr, Girshick R , et al. Feature Pyramid Networks for Object Detection[J]. 2016.

**三、开题报告评价（本项分别由指导教师及专家组填写）**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **指导老师对该生选题报告的简要评语（**本栏由指导教师在开题报告会之前填写完毕）**：**  本课题是基于part feature的行人再识别技术的研究与分析，面对较少的数据集能提出很好的数据  增强的方法。同时针对于局部特征贡献度大小的问题提出特征加权的方法，提出自己的拟解决方案，并且方案具有相关实验的支持，选题具有很好的应用意义。同时，充分的分析了国内外现有的方案，并针对于相应的优缺点做出分析和改进，阐述清楚，具有可行性。  鉴于以上情况，同意其参加开题。    **指导教师签名： 年 月 日** | | | | | | |
| **开**  **题**  **报**  **告**  **会** | **开题报告会时间： 2019年 01月 16日 上午 09时— 12时 地点：信西305A** | | | | | |
| **评**  **审**  **专**  **家**  **组**  **成** | 姓名 | 职称 | 所在单位及学科专长 | 博导**/**硕导 | （出席者）签名 |
| 组长：胡长建 | 高级研究员 | 自然语言处理、人工智能 | 硕导 |  |
| 马伟 | 副教授 | 北京工业大学 计算机科学与技术 | 硕导 |  |
| 才智 | 副教授 | 北京工业大学 计算机科学与技术 | 硕导 |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| **评审意见：**（由评审专家组填写）  **（正常）通过 √；**  **不合格 ，**  **前重做开题报告。**  **专家评审评语：**  **评审组长签名： 年 月 日** | | | | | |