

DOI: 10.7652/xjtuxb202001007

采用类内迁移学习的红外/可见光异源图像匹配

毛远宏, 贺占庄, 马钟, 毕瑞星, 王竹平

(西安微电子技术研究所, 710065, 西安)

摘要: 为解决异源图像匹配中样本量过少和成像原理不同导致成像差异的问题,提出了一种采用类内迁移学习的异源图像匹配网络(PairsNet)。该网络由特征提取子网络和匹配度量子网络两部分组成。特征提取子网络中存在4条卷积神经网络分支,其通过卷积神经网络分支提取出红外图像和可见光图像的特征。将可见光图像作为源域、红外图像作为目标域进行迁移学习,通过减小两个域中样本特征的类内最大均值差异距离,实现了源域和目标域对应图像类别上精准的样本特征分布对齐。匹配度量子网络使用2个全连接层和1个softmax层进行串联,评估出异源图像特征的匹配度。构建了红外和可见光图像数据集,进行端到端的训练和测试。结果表明:与当前使用预训练模型微调的方法相比,PairsNet的准确率提升了10.54%,可见光图像匹配网络的能力可以有效迁移到异源图像匹配网络。

关键词: 异源图像;图像匹配;迁移学习

中图分类号: TP391 **文献标志码:** A **文章编号:** 0253-987X(2020)01-0049-07

Infrared-Visible Heterogenous Image Matching Based on Intra-Class Transfer Learning

MAO Yuanhong, HE Zhanzhuang, MA Zhong, BI Ruixing, WANG Zhuping

(Xi'an Microelectronics Technology Institute, Xi'an 710065, China)

Abstract: To solve the problem of imaging difference and too few training samples in infrared-visible images matching, a matching network based on intra-class transfer learning is proposed. The network consists of feature extraction subnetwork and metric subnetwork. Owing to four convolutional neural network (CNN) branches in feature extraction subnetwork, the proposed network is referred to as PairsNet for short. The CNN branches extract infrared and visible image features. The visible image is used as the source domain and the infrared image as the target domain. By reducing the maximum mean discrepancy distance of intra-class samples in the two domains, more accurate sample distribution alignment is achieved in the intra-class feature space. The metric subnetwork uses two full connection layers and one softmax layer in series to evaluate infrared-visible image matching performance. Infrared and visible image data sets are built for end-to-end training and testing. The experimental results show that the accuracy of PairsNet is improved by 10.54%, compared with the network fine-tuning method based on pre-training model. The ability of visible image matching network can be effectively transferred to the infrared-visible image matching network.

Keywords: infrared-visible images; image matching; transfer learning

收稿日期: 2019-06-14。 作者简介: 毛远宏(1980—),男,博士生;贺占庄(通信作者),男,研究员,博士生导师。 基金项目: 国家自然科学基金青年科学基金资助项目(61702413)。

网络出版时间: 2019-09-26

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1069.T.20190926.0906.002.html>

从不同类型的传感器获得的图像信息被称为异源图像。异源图像匹配是计算机视觉中的一项基本任务,它通过比较不同传感器图像在结构、关系、纹理、边缘和灰度的相关性,评估异源图像的相似性匹配。异源图像匹配有着重要的应用价值,例如:在基于视觉的导航和制导领域,异源图像匹配可以支持在夜间和复杂气候条件下无人飞行器的导航和制导^[1-2];在目标识别领域,能够实现待识别目标在多模态影像之间的匹配,从而有效提高识别性能^[3-4]。

异源图像匹配相对于传统可见光图像匹配来说,面临更多的困难。首先,由于各种传感器成像原理不同,异源图像之间的差异远远大于单一可见光图像。其次,异源图像是不同的传感器在不同时间、不同角度拍摄,即使同一场景,也可能存在时间、视角和形变等因素的影响。最后,在可见光图像领域,用于深度学习的数据集中图像数量通常数以百万计。例如,VGG16网络^[5]进行预训练使用的ImageNet 2012数据集,仅训练集图像就达到了128万张,而异源图像的样本采集成本高,在实际操作过程中很难获取大量带有标记的数据,因此基于数据驱动的深度学习方法难以直接运用于异源图像匹配。

异源图像匹配一直是计算机视觉领域的研究热点,各种方法不断涌现。传统异源图像匹配的方法主要包括基于灰度的匹配和基于手工特征的匹配。基于灰度的匹配方法通过计算图像中像素点灰度的统计指标,比较图像之间的相似性^[6-8]。这种方法的缺点主要是计算量相对大,并且对异源图像中常出现的灰度失真和目标形变比较敏感。基于手工特征的匹配方法通过SIFT^[9]、SURF^[10]、ORB^[11]等描述子,从红外和可见光图像中分别提取图像的特征。由于传感器的不同成像原理,使用手工特征的方法,在可见光图像上很明显的特征点在相应的红外图像上并不显著,反之亦然。这导致基于手工特征的异源图像匹配准确率不高。

深度学习的方法兴起之后,在计算机视觉领域中分类、检测、语义分割等方面都取得了前所未有的性能。在可见光的图像块匹配方面,目前基于深度学习的方法也取得了优异的性能,主流的方法有MatchNet^[12]和Deep Compare Network^[13],它们都采用了孪生神经网络的架构,由两个CNN分支分别来提取待匹配图像的特征,通过比较两张图像的相似性来评估是否匹配。通过神经网络来选择更合适的特征表达,MatchNet和Deep Compare Network在可见光图像块的匹配性能上超过了目前主

流的手工特征方法。但是,这些深度学习网络需要大量可见光图像样本进行有监督学习,同时要求样本之间服从独立同分布。对于异源图像来说,这两个条件都难以满足。在异源图像匹配过程中,首先获取异源图像对代价过高,缺乏大量有标注的异源图像数据来训练网络。另一方面,由于成像差异,可见光和红外图像之间不是同分布的关系,因此直接使用可见光匹配网络进行训练,或者基于可见光的预处理模型进行网络微调,在异源图像匹配上效果都不理想。

由于上述原因,本文希望采用迁移学习来解决此问题,将可见光图像作为源域,将红外图像作为目标域。迁移学习可以解决目标域中训练样本过少的问题,避免网络出现过拟合。更重要的是,迁移学习通过域适配,能够使得源域和目标域在高维特征空间的分布距离减小,可以解决源域和目标域样本分布差异的问题。

迁移学习是机器学习的一个重要分支。近年来,随着深度学习的快速发展,两者之间的结合也更加紧密^[14-16]。通过迁移学习方法,希望将可见光图像(源域)的深度网络学习能力迁移给红外图像域(目标域)。本文提出基于源域和目标域的内迁移学习,对于红外图像特征学习过程进行引导,让两个域高层特征分布接近,以提高异源图像匹配效果。通过上述改进,本文提出的网络相比之前方法,性能有较大提升。

1 PairsNet 总体架构

PairsNet主要由特征提取子网络和匹配度量子网络两大部分组成,网络结构如图1所示。特征提取子网络是整个PairsNet的核心,主要用于异源图像的特征提取。在训练阶段特征时,网络分为4个分支,每个分支都采用了VGG16网络结构。每个VGG16网络中卷积层包含5个卷积Block结构。每个Block结构由两个或者3个卷积层、一个激活层和一个池化层组成。为了减少分支的参数数量,在本文中,VGG16最后的两个全连接层被全局平均池化层(GAP)替代。每个VGG16分支最终从一幅输入图像提取出 $1 \times 4\,096$ 维的向量。

在PairsNet训练过程中,4张图像组成一个训练样本。如图1所示:分支1和分支2输入来自同一场景的异源图像,其中分支1输入可见光图像,分支2输入红外图像;分支3和分支4输入来自同一场景的异源图像,其中分支3输入可见光图像,分支

4 输入红外图像。在特征提取子网络中采用此种网络设计,希望通过神经网络的训练,使得红外图像和可见光图像的特征分布能够相互接近,实现源域和目标域之间的迁移学习。

在 PairsNet 中,分支 2 和分支 3 之间主要聚焦于异源图像匹配工作,通过匹配度量量子网络给出匹配的概率。匹配度量量子网络包含两个全连接层(FC)和 softmax 层。分支 2 和分支 3 输出异源图像特征,两个全连接层主要用来学习特征匹配的度量准则。softmax 层用于评估分支 2 和分支 3 中异源图像的匹配度。

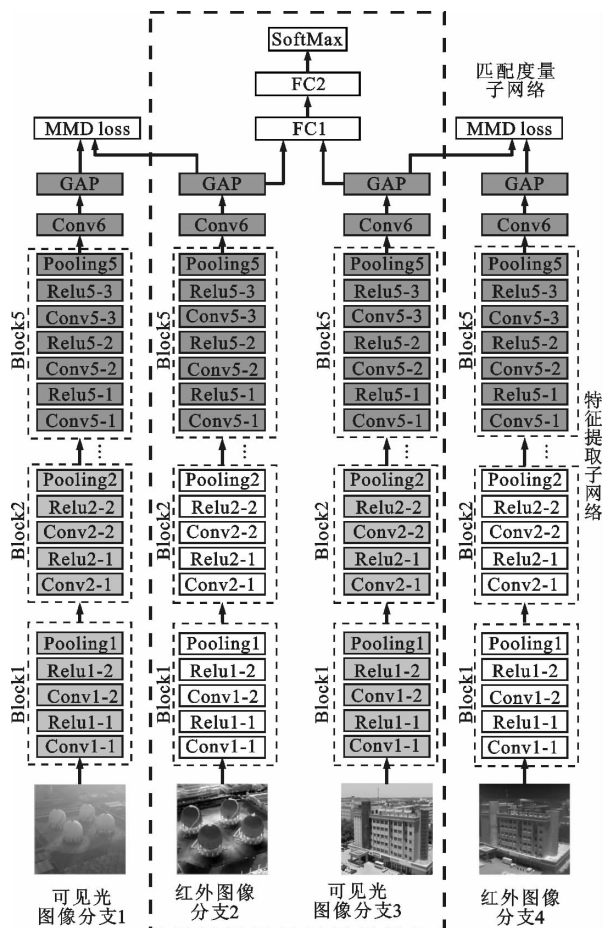


图 1 PairsNet 结构图

在 PairsNet 训练的过程中,待优化的损失函数如下

$$\lambda_1 d_1(D_s, D_t) + \lambda_2 d_2(D_s, D_t) + \sum_{i=1}^n \text{softmax}(x_i, y_i) \quad (1)$$

式中:函数 d_1 为分支 1 可见光图像特征和分支 2 红外图像特征的 MMD(maximum mean discrepancy)距离;函数 d_2 为分支 3 可见光图像特征和分支 4 红外图像特征的 MMD 距离; D_s 表示源域特征在高维

空间的数据分布; D_t 表示目标域特征在高维空间的数据分布; n 为训练样本总数; λ_1 和 λ_2 为超参数。通过多次实验, λ_1 和 λ_2 均设置为 0.10,则网络能够达到最好的性能。通过优化两个 MMD 距离,实现了源域和目标域在特征空间分布相似。式(1)第 3 项为 softmax 损失函数,用它来衡量分支 2 和分支 3 输入的异源图像是否匹配,0 表示不匹配,1 表示匹配。

需要说明的是,迁移学习相关的网络结构和 MMD 距离仅在 PairsNet 训练过程中工作。在测试过程中,PairsNet 仅保留分支 2 和分支 3,网络输入为待匹配的异源图像对。PairsNet 测试阶段的网络结构为图 1 中虚线框所包含的部分。

2 异源图像匹配的内迁移学习

虽然红外和可见光图像在底层成像原理上存在差异,但是对于同一场景的异源图像仍然有相似之处,因此异源图像中使用迁移学习具有可行性。

同时,由于异源图像样本的获取比较困难,样本数量少,从零开始训练一个全新的异源图像匹配网络容易引起过拟合。对于此类问题,目前常用的方法是在可见光图像的预训练模型上,使用异源图像样本进行网络参数微调。广义上讲,网络参数微调也是属于基于模型参数的一种迁移学习。此种方法虽然在一定程度上可以降低过拟合的风险,但是严格上讲,红外和可见光的图像样本之间不是独立同分布的关系,仅仅采用网络微调的方法,不同类别的样本特征存在着重叠和交错,无法保证两个域中相同类别的特征在高维空间中的分布能够对齐,因此在异源图像匹配任务中需要更加精准的迁移学习方法。

为了说明网络微调中的问题,使用 t-SNE^[17] 的方法可视化了使用网络微调方法后的高层特征分布。网络微调采用了可见光图像匹配中 MatchNet 架构,使用 ImageNet 数据集预训练的权值来初始化两个 VGG16 分支,再使用异源图像样本对于网络进行重训练。从高空无人机拍摄的图片类别中选取了 5 种类别,分别为天然气储气罐、房屋、大楼、立交桥和工厂厂区。图 2 为基于网络微调的异源图像特征分布,其中可见光的 5 种类别各自聚集,圆点代表可见光图像域(源域)中样本的高层特征,三角形代表红外图像域(目标域)中样本的高层特征。由图 2 可以看出,在高维空间中源域和目标域的样本特征不能保证在类别上的分布完全对齐。

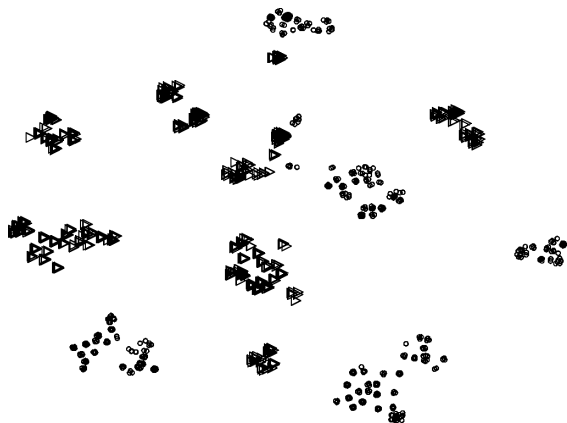


图2 基于网络微调的异源图像特征分布

传统的迁移学习通常是关注全部样本的域适应,而忽略了两个域中类内特征的相似性。在异源图像收集阶段,在不同场景下分别采集了红外和可见光图像样本,因此目标域中的图像有明确的分类标签。相比于JDA方法^[16]中通过迭代的方式对目标域的样本类别设置伪标签,异源图像匹配中的目标域的条件分布信息更加清晰。因此,本文希望通过条件分布自适应方法来适配两个域的条件分布 $P(y_s|x_s)$ 和 $P(y_t|x_t)$,使得它们相似。其中: x_s 为源域样本, y_s 为源域图像中的类别; x_t 为目标域样本, y_t 为目标域图像中的类别。

在图1中,PairsNet在每个VGG16分支的全局平均池化层都输出了 $1 \times 4\,096$ 维的向量,其分支1和分支2之间的MMD距离、分支3和分支4之间的MMD距离,都是为了比较源域和目标域中同类别的两个分布之间的距离,如下式所示

$$D(D_s, D_t)^2 = \sum_{c=1}^C \left\| \frac{1}{m_c} \sum_{x_s \in D_s^{(c)}} \phi(x_s) - \frac{1}{n_c} \sum_{x_t \in D_t^{(c)}} \phi(x_t) \right\|_H^2 = \sum_{c=1}^C \left(\frac{1}{m_c^2} \sum_{i,j=1}^{m_c} k(x_s^i, x_s^j) - \frac{2}{m_c n_c} \sum_{i,j=1}^{m_c, n_c} k(x_s^i, x_t^j) + \frac{1}{n_c^2} \sum_{i,j=1}^{n_c} k(x_t^i, x_t^j) \right) \quad (2)$$

式中: C 为异源图像的分类数量; m_c 是源域中一个分类的样本数量; n_c 是目标域中对应相同分类的样本数量; $\phi(x)$ 代表了样本特征到再生希尔伯特空间的映射。MMD距离平方展开后,再生希尔伯特空间的内积将转化为核函数 $k(x_i, x_j)$,以方便计算。

为了使核函数表示更加有效,在MMD距离计算中,先使用一个Minibatch中的样本计算出方差 σ^2 ,再分别以 $(0.5\sigma^2, \sigma^2, 2\sigma^2)$ 构造3个高斯核。使用这3个核的MMD距离之和来评估源域和目标域的

分布相似性。

为了实现样本能够按照类别计算MMD距离,当构建训练数据集时,在一个Minibatch内部,人为地将分支1和分支2的输入设置为同类异源图像,这样可以保证一个Minibatch中在同一分类计算MMD距离。分支3和分支4的输入图像也遵循相同的原则来计算MMD距离。在PairsNet的整体训练过程中,不要连续输入包含相同类别的样本批次,需要交替优化不同类别的样本,从而取得更好的训练效果。

3 实验

3.1 测试数据

目前还没有公开的可见光和红外图像匹配的数据集,本文自行构建了数据集进行实验。数据的采集是通过在无人机上搭载红外摄像头和可见光摄像头,获取了不同类型物体在不同角度、光照、尺度、旋转的图像。红外相机的拍摄波长为 $7.5 \sim 13.5 \mu\text{m}$ 。由于红外摄像头对于颜色和细致的纹理不敏感,样本主要选择不同形状和轮廓的目标。在数据收集的过程中,无人机在不同的高度对目标进行拍摄。在后期的数据处理过程中,对于目标区域进行相应的裁剪,人工整理成对应的图像对。

在数据采集后,一共选择了2 016幅图像(包含可见光和红外图像),共25个场景类别,包含立交桥、道路、天然气罐、各类型建筑物等。在数据集中,可见光和红外图像的比例为1:1。数据集中80%的样本用来进行训练,剩余样本用来进行测试。如果两幅图像来自同一个场景,样本标签为1,为正样本;如果两幅图像来自不同的场景,样本标签为0,为负样本。在数据集中,正负样本的比例为1:1。在一个正样本中,要求输入图像之间的拍摄角度偏差不能超过 30° 。要匹配的关键信息点的物体的大小占整个图像块的尺寸不能小于50%。如果为同一场景,其样本标签为1;如果为不同样本,样本标签为0。图3为PairsNet的训练样本,其中奇数列为负样本,偶数列为正样本。

需要说明的是,PairsNet的测试样本相对于训练样本,仅保留了图3中的分支2和分支3的图像输入,训练样本标签由分支2和分支3输入的图像决定。测试集标注样本标签的方法和训练集保持一致。

3.2 测试方法

在训练PairsNet之前,需要一个高精度的可见

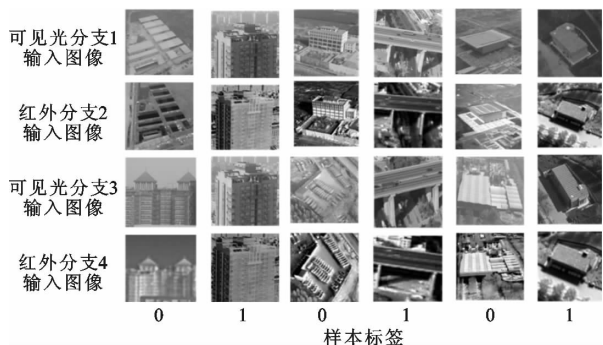


图 3 PairsNet 的训练样本

光匹配网络作为后续迁移学习中源域的基线。相对于红外图像,可见光图像的获取相对容易。通过拍摄和网络收集了大量图像,并进行了数据集整理和标注,样本数量为异源图像样本的 10 倍以上。可见光样本在拍摄角度、形变、尺度等方面相比于异源图像样本来说,也更加丰富多样。使用 MatchNet 架构,将 VGG16 作为特征提取分支,使用 ImageNet 预训练的权值作为初始值进行再训练。通过大量数据训练,获得了一个准确率为 92.31% 的可见光图像匹配网络。

PairsNet 训练过程中,首先使用上面预训练好的可见光网络权重来初始化其特征提取子网络的 4 个 VGG16 分支。匹配度量网络中的两个全连接层使用 Xavier 方法进行初始化,Minibatch 大小设置为 16。梯度下降算法采用 Momentum 算法,参数为 0.9。网络基础学习率为 $10^{-3} \sim 10^{-5}$,按照迭代次数均匀下降,迭代次数设置为 50 000。网络每层的学习率等于网络基础学习率乘以各层的学习率乘子,全连接网络的学习率乘子设置为 1,其余层的学习率乘子设置为 0.01。

实验使用的硬件平台为 Nvidia TITAN XP GPU,软件平台使用 Caffe 进行训练和测试。

3.3 测试结果

为了证明 PairsNet 的有效性,在实验过程中增加了不同的模型和方法的对比实验,参与实验的方法和网络如下。

(1)由于红外和可见光图像在物体边缘轮廓有相似之处,在红外和可见光图像的传统匹配方法中,有基于边缘增强和特征提取相结合的传统异源图像匹配方法^[18]。在对比实验中,先采用传统匹配方法进行边缘增强,再人工提取特征点,最后通过 RANSAC 迭代选取最相似的特征点来进行匹配。

(2)基于 MatchNet,随机初始化开始训练。特征提取子网络使用 VGG16 网络作为分支,直接使

用异源图像样本来进行训练。由于样本过少,网络很快过拟合。

(3)基于 MatchNet,使用 VGG16 在 ImageNet 数据集上预训练模型来初始化特征提取子网络,然后使用异源图像样本进行再训练,对网络参数进行微调。由于小样本场景下业界通常采用此种训练方法,因此在实验结果的评估中,将其设置为比较基线。

(4)PairsNet 在底层卷积采用独立参数,高层卷积共享参数。在特征提取子网络中红外图像和可见光图像分支的底层卷积采用独立权值。对于异源图像来说,由于不同的成像原理,底层特征存在比较大的差异,如果在两个域的底层都采用相同的滤波器来提取特征,可能无法针对各自域的成像特点来选择最适应的滤波器。因此,在特征提取子网络中可见光分支和红外分支采用独立的底层卷积参数,通过反向传播算法,让可见光图像域和红外图像域中 Block1 和 Block2 中的卷积层去自主地选择最适合本域的滤波器。在 PairsNet 中,高层卷积层 Block3 ~ Block5 中卷积层参数设计为所有分支共享。因为在高层特征中,希望能够将两个域中有共性的特征信息提取出来。在图 1 所示的特征提取子网络中,4 个 VGG16 网络分支中使用相同灰度的卷积层表示卷积层参数共享。

(5)PairsNet 在特征提取子网络中的红外图像和可见光图像分支,无论是底层还是高层卷积,各个分支完全共享权值。

为了避免正负样本不均衡的影响,业界通常采用 ROC 曲线来衡量分类问题的性能。通常采用在 95% 召回率(真阳性率)条件下,测试结果中假阳性的百分比(错误率)作为衡量的标准,其错误率越低说明越好。

基于实验结果,上述方法的 ROC 曲线如图 4 所示。由图可以看出,本文方法相比于 MatchNet(微调)方法,错误率下降了 10.54%。通过多项对比实验,得出以下结论。

(1)由于可见光和红外的成像差异大,使用传统的手工描述子,无法保证提取出异源图像的共同特征,匹配效果不理想。

(2)在红外图像和可见光图像的底层卷积使用独立的权值,可以更加有针对性地提取网络的底层特征,提高特征的表达能力。

(3)在 PairsNet 中,采用基于类别的迁移学习,可使迁移的效果更加精准,减少同一分类内部可见光和红外图像的数据分布差异。

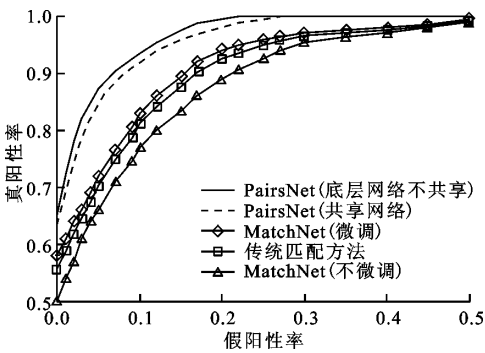


图 4 不同方法的 ROC 曲线

为了描述更加具体,本文可视化了 PairsNet 迁移学习后特征提取网络的输出特征,如图 5 所示。相比于图 2,可以明显看出,通过使用基于类别对齐的迁移学习之后,类内可见光和红外图像特征分布更加紧密。



图 5 PairsNet 异源图像特征样本分布

图 6 展示了 PairsNet 在异源图像中的测试结果,网络推断异源图像对为正样本,则匹配预测值为 1,否则预测值为 0。相比之前的匹配方法,PairsNet 在这些样本上实现了准确的匹配判断。对于正样本来说,异源图像存在着比较大的角度和尺度变化,PairsNet 具有更强的鲁棒性;对于负样本来说,虽然异源图像存在一定的相似性,PairsNet 依然能够有效地区分。

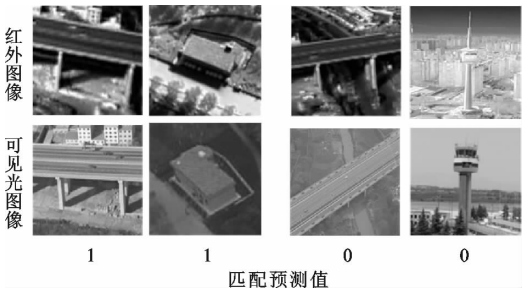


图 6 PairsNet 异源图像测试结果

为了更加客观地比较不同算法和网络架构的效率,在测试过程中对相应的运行速度进行了比较,结果如表 1 所示。在测试过程中,神经网络仅进行前向推理,不进行反向传播。在 CPU 上的运行速度可以反映出不同算法的运算复杂度。通过在 CPU 上的运行速度可以看出,传统算法在计算量上明显小于基于深度学习的匹配算法。在 GPU 平台上,由于目前硬件和软件平台对于端到端深度学习架构的良好支持,基于深度学习方法的处理速度和传统的异源图像匹配方法相差不大。

PairNet 在特征提取子网络中使用全局平均池化层替换了 VGG16 先前的全连接层,因此计算速度相比 MatchNet 有所提高。

表 1 不同算法运行速度对比 s^{-1}

硬件平台	运行速度		
	传统方法	MatchNet	PairsNet
Intel CPU 2.40 GHz	2.857	0.278	0.283
GPU Nvidia TITAN XP	63.217	60.063	61.541

4 结 论

本文将深度学习和迁移学习方法相结合,提出了 PairsNet 来解决异源图像匹配存在的样本过少和成像差异的问题。将红外图像域作为目标域,可见光图像域作为源域,为了便于两个域之间的样本进行特征分布比较,PairsNet 使用了多路 VGG16 分支来分别提取红外和可见光图像特征。与迁移学习中传统的域适配方法相比,PairsNet 优化了两个域中样本特征在相同类别的 MMD 距离。在高层特征空间中,两个域样本在类内实现更加准确的分布对齐。在实验中,自行构建了用于红外和可见光图像迁移学习的数据集,对于 PairsNet 进行了端到端训练。通过上述改进,PairsNet 的图像匹配错误率相比于参数微调的方法下降了 10.54%,并且没有影响在 GPU 设备上的前向推理速度。

参考文献:

[1] 李露, 蒋宏, 任章. 红外与可见光复合寻的制导中的快速图像配准方法 [J]. 哈尔滨工程大学学报, 2009, 30(2): 228-232.
LI Lu, JIANG Hong, REN Zhang. A fast image registration algorithm for compound IR/visible light homing guidance [J]. Journal of Harbin Engineering University, 2009, 30(2): 228-232.

- [2] LI H, DING W, CAO X, et al. Image registration and fusion of visible and infrared integrated camera for medium-altitude unmanned aerial vehicle remote sensing [J]. *Remote Sensing*, 2017, 9(5): 441.
- [3] 叶华, 朱明早, 王日兴. 红外和可见光图像互补融合的运动目标检测方法 [J]. *红外技术*, 2015, 37(8): 648-654.
- YE Hua, ZHU Minghan, WANG Rixing. Fusion of complementary information from infrared and visual image for moving object detection [J]. *Infrared Technology*, 2015, 37(8): 648-654.
- [4] CHENG K-S, LIN H-Y. Automatic target recognition by infrared and visible image matching [C]//14th IAPR International Conference on Machine Vision Applications. Piscataway, NJ, USA: IEEE, 2015: 312-315.
- [5] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. [2019-08-15]. <https://arxiv.org/pdf/1409.1556.pdf>.
- [6] ZHAO F, HUANG Q, GAO W. Image matching by normalized cross-correlation [C]// IEEE International Conference on Acoustics, Speech, and Signal Processing. Piscataway, NJ, USA: IEEE, 2006: 729-732.
- [7] MAES F, COLLIGNON A, VANDERMEULEN D, et al. Multi-modality image registration by maximization of mutual information [C]// Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis. Piscataway, NJ, USA: IEEE, 1996: 14-22.
- [8] SIM D-G, KWON O-K, PARK R-H. Object matching algorithms using robust Hausdorff distance measures [J]. *IEEE Transactions on Image Processing*, 1999, 8(3): 425-429.
- [9] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision and Image Understanding*, 2004, 60(2): 91-110.
- [10] BAY H, TUYTELAARS T, GOOL L J V. SURF: speeded up robust features [J]. *Computer Vision and Image Understanding*, 2006, 110(3): 404-417.
- [11] RUBLEE E, RABAUD V, KONOLIGE K, et al. ORB: an efficient alternative to SIFT or SURF [C]// IEEE International Conference on Computer Vision. Piscataway, NJ, USA: IEEE, 2011: 2564-2571.
- [12] HAN X, LEUNG T, JIA Y, et al. MatchNet: unifying feature and metric learning for patch-based matching [C]// IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2015: 3279-3286.
- [13] ZAGORUYKO S, KOMODAKIS N. Learning to compare image patches via convolutional neural networks [C]// IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2015: 4353-4361.
- [14] PAN S J, YANG Q. A survey on transfer learning [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2010, 22(10): 1345-1359.
- [15] LONG M, CAO Y, WANG J, et al. Learning transferable features with deep adaptation networks [J]. *Journal of Machine Learning Research*, 2015, 37: 97-105.
- [16] LONG M, WANG J, DING G, et al. Transfer feature learning with joint distribution adaptation [C]// IEEE International Conference on Computer Vision. Washington, DC, USA: IEEE, 2013: 2200-2207.
- [17] VAN DER MAATEN L, HINTON G E. Visualizing data using t-SNE [J]. *Journal of Machine Learning Research*, 2008, 9: 2579-2605.
- [18] 纪利娥, 杨风暴, 王志社, 等. 基于边缘图像和 SURF 特征的可见光与红外图像的匹配算法 [J]. *红外技术*, 2012, 34(11): 629-635.
- JI Lie, YANG Fengbao, WANG Zhishe, et al. Visible and infrared image matching algorithm based on edge image and SURF features [J]. *Infrared Technology*, 2012, 34(11): 629-635.

(编辑 荆树蓉)