

基于筛选深度特征的红外图像目标识别方法

廖莎莎

(巢湖学院 信息工程学院, 安徽 合肥 238000)

摘要: 红外成像是现代战场侦察的重要手段, 基于红外图像的目标识别技术可为情报解译提供重要支撑。针对红外图像目标识别, 提出基于筛选深度特征的方法。设计适当结构的 ResNet 对红外图像进行特征学习, 对于每个卷积层的输出特征图进行矢量化处理, 获得相应的特征矢量。针对各个特征图的深度特征矢量, 基于斯皮尔曼等级相关系数评价它们与原始图像的相关性。然后, 通过门限判决算法选取若干具有高相关性的深度特征。经过筛选得到的深度特征可剔除了不必要的冗余成分, 从而提升后续分类的精度和稳健性。采用联合稀疏表示模型对筛选得到的若干深度特征进行表征和分类, 最终获取待识别样本的所属类别。因此, 方法可有效结合 ResNet 多层次深度特征的鉴别力, 从而提高最终的识别性能。实验在公开的中波红外目标图像数据集 (MWIR) 开展, 利用原始测试样本、模拟噪声样本和模拟遮挡样本对方法性能进行测试和分析。实验结果表明: 相比现有的部分红外目标识别方法, 提出方法可取得更强的有效性和稳健性。

关键词: 红外图像; 目标识别; ResNet; 斯皮尔曼等级相关系数; 联合稀疏表示
中图分类号: TP391 **文献标志码:** A **DOI:** 10.3788/IRLA20210372

Infrared image target recognition method based on selected deep features

Liao Shasha

(College of Information Engineering, Chaohu University, Hefei 238000, China)

Abstract: Infrared imaging is an important means of modern battlefield reconnaissance, and target recognition technology based on infrared images can provide important support for intelligence interpretation. Aiming at target recognition in infrared images, a method based on selected deep features was proposed. A ResNet with a proper structure was designed to perform feature learning on infrared images, and the output feature maps from each convolutional layer was vectorized to obtain a corresponding feature vector. For the deep feature vectors of different feature maps, their correlations with the original image were evaluated based on the Spearman rank correlation coefficient. Afterwards, several deep features with high correlations were selected through the threshold decision algorithm. The deep features obtained after selection can eliminate unnecessary redundant components, thereby improving the accuracy and robustness of subsequent classification. The joint sparse representation model was used to characterize and classify the selected deep features, and finally the category of the sample can be identified. Therefore, the proposed method can effectively combine the discrimination of the multi-level deep features learned from ResNet, thereby improving the final recognition performance. The experiments were carried out in the public mid-wave infrared target image dataset (MWIR), using the original test

收稿日期: 2021-12-01; 修订日期: 2022-01-20

基金项目: 巢湖学院教改项目 (ch19jxyj37); 巢湖学院质量工程 (ch19kfk19); 国家级课题 (ZGD-Y-2018-24)

作者简介: 廖莎莎, 女, 实验师, 硕士, 主要研究方向为计算机应用、图像处理等。

samples, simulated noisy samples and simulated occluded samples to test and analyze the performance of the method. The experimental results show that the proposed method can achieve stronger effectiveness and robustness compared with some existing infrared target recognition ones.

Key words: infrared image; target recognition; ResNet; Spearman rank correlation coefficients; joint sparse representation

0 引言

随着红外传感元器件和红外成像技术的发展,高分辨率红外图像在军事和民用领域得到广泛运用^[1-5]。在民用领域,红外夜视仪、夜间交通管控系统等设施设备发挥了重要作用。在军事领域,基于红外成像的导弹末制导系统具有精度高、针对性强等特点。红外图像目标识别技术可依托观测图像获取感兴趣目标的所属类别,为情报解译提供支撑。红外目标识别方法一般区分特征提取和分类决策两个阶段。根据现有文献,红外目标识别方法中涉及的特征包括几何外形、局部纹理、像素分布特性等^[6-9]。分类过程中,根据提取得到的特征进行分类器的选择或设计,典型代表包括支持向量机(SVM)^[10]、稀疏表示分类(SRC)^[11]等。深度学习理论的出现和发展为信号处理、图像分析提供了重要手段。特别地,在图像识别领域,以卷积神经网络(CNN)为代表的深度学习模型成为炙手可热的工具,并验证了其有效性。参考文献[12-14]就针对红外图像目标识别问题设计了深度网络模型并进行了应用。特别地,参考文献[14]以某卷积神经网络为基础对红外图像进行特征提取,将各个卷积层输出的特征图进行处理,获取相应的深度特征矢量。最后,采用联合稀疏表示模型对多层次深度特征进行联合决策。相比传统直接应用深度模型进行端到端分类的方法相比,多层次深度特征的联合运用进一步提高了识别性能,反映了方法的有效性。

文中提出基于筛选深度特征的红外图像目标识别方法,主要是对参考文献[14]中联合深度特征方法进行两个方面的改进优化。一方面,采用 ResNet 替换参考文献[14]中的卷积神经网络。根据现有文献报道,ResNet 通过引入残差学习算法可显著提高深度网络的稳健性^[15-16]。相比一般的卷积神经网络,ResNet 输出的多层次深度特征具有更强的鉴别力。另一方面,对于 ResNet 各个卷积层输出的深度特征矢量,基于斯皮尔曼等级相关系数^[17-18]评价它们与原

始图像的相关性,进而合适的门限选取若干相关性较高的深度特征矢量。最终,基于联合稀疏表示模型^[14,19]对选取的多层次深度特征矢量进行表征分类,获取决策结果。与参考文献[14]中利用全部深度特征矢量的算法相比,文中通过有效的特征筛选可以提出部分冗余甚至负面信息,从而提高分类的精度和效率。因此,通过充分运用 ResNet 学习得到的深度特征,文中方法可以进一步提高红外图像识别性能。实验中,采用公开的中波红外目标图像数据集(MWIR)对设计方法进行测试和分析,分别在原始样本、噪声样本和遮挡样本三类情形下与部分现有方法进行对比。实验结果反映了提出方法的有效性和稳健性。

1 基于 ResNet 的深度特征学习

ResNet 作为一种代表性的卷积神经网络,在图像分割、分类识别等领域得到了广泛应用和验证^[15-16]。与一般的卷积神经网络相比,ResNet 能够随着网络层数的增加,获得更佳丰富的深度特征,同时可避免梯度消失问题。假设 $H(x)$ 为最佳映射,利用堆叠的非线性层得到新映射 $F(x) = H(x) - x$,从而计算新的最佳映射 $F(x) = H(x) + x$,提高网络的整体学习性能和分类能力。

文中将 ResNet 引入红外图像目标识别,主要用于学习多层次的深度特征。图 1 显示了文中设计的针对 10 类目标分类问题的 ResNet 整体结构,共包含 12 层,在每个卷积层均可以输出特征图,通过矢量化操作获得相应的深度特征矢量。与一般的卷积神经网络结构相比,ResNet 能够将输入与后续非相邻层直接相连,从而最大程度保持了图像中的有效信息。此外,ResNet 降低了网络学习的整体难度,有利于提高网络学习效率。基于图 1 可获得红外目标图像多层次深度特征。这些深度特征可从不同方面描述目标特性,联合运用它们可为后续的分类识别提供有效支撑。

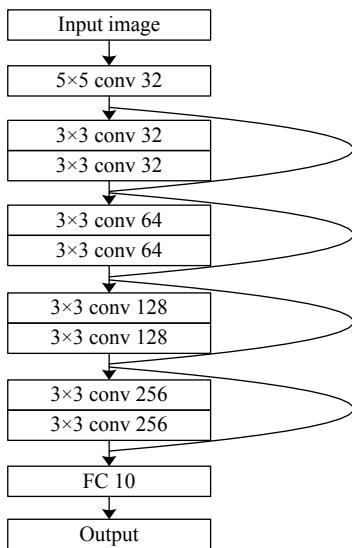


图 1 ResNet 结构

Fig.1 Architecture of ResNet

2 斯皮尔曼等级相关系数

基于 ResNet 学习得到的多层次深度特征对于描述红外图像中的目标特征具有互补性,联合运用这些信息有利于提高识别性能。然而,实际过程中并不是所有层次的深度特征对于正确识别都有贡献,其中不乏冗余分量,不仅会降低识别效率还可以降低识别精度。为此,有必要对多层次深度特征矢量进行可靠筛选,筛选保留的深度特征用于后续的分类过程。

文中采用斯皮尔曼等级相关系数作为深度特征筛选的基础相似度度量准则。斯皮尔曼等级相关稀疏用于评价两个变量之间的统计相关性^[17-18],有不依赖样本分布的优点。设定 P 和 Q 为两个 n 维的变量, P_i 、 Q_i 分别为两者的第 $i(1 \leq i \leq n)$ 个元素。按照升序或降序对 P 和 Q 的进行重新排列,得到新的变量序列 p 、 q 。其中, p_i 表示 P_i 在 P 中的排序, q_i 表示 Q_i 在 Q 中的排序。然后,计算差分集合 $d = p_i - q_i, 1 \leq i \leq n$ 。在此基础上, P 和 Q 的斯皮尔曼等级相关系数按照下式计算:

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \quad (1)$$

根据上述过程,斯皮尔曼等级相关系数可用于快速计算不同深度特征矢量与原始红外图像之间的相关性。在此基础上,设置适当门限 T ,对于相关性高于门限的深度特征矢量予以保留。反之,相关性小于门

限的深度特征矢量则予以剔除,不再进入后续的分类过程。

3 识别方法描述

3.1 联合稀疏表示

联合稀疏表示模型是处理多个关联稀疏表示问题的有效工具,通过利用内在关联提高求解精度,为信号重构,图像识别等应用提供支撑^[14, 19]。对于测试样本 y ,假设其经过筛选后得到的 K 个 ResNet 深度特征矢量为: $[y^{(1)} y^{(2)} \dots y^{(K)}]$ 。它们的联合稀疏表示模型构建如下:

$$\min_{\beta} \left\{ g(\beta) = \sum_{k=1}^K \|y^{(k)} - A^{(k)} \alpha^{(k)}\| \right\} \quad (2)$$

式中: $A^{(k)}$ 表示对应第 k 个深度特征的全局字典,记为所有训练样本获取的第 k 个深度特征的矩阵合集; $\alpha^{(k)}$ 为稀疏系数矢量; $\beta = [\alpha^{(1)} \alpha^{(2)} \dots \alpha^{(K)}]$ 则为系数矩阵。

公式 (2) 给出了联合稀疏表示的最基本形式,但并未考虑不同成分之间的联系。为此,实际应用的联合稀疏表示模型增加了相关性约束,新的目标函数如下式所示:

$$\min_{\beta} g(\beta) + \lambda \|\beta\|_{2,1} \quad (3)$$

式中: λ 正则化系数。 ℓ_1/ℓ_2 范数对稀疏矩阵的约束使得不同深度特征对应的系数矢量倾向于相近的分布特点,即它们之间的相关性。

对公式 (3) 进行求解,可获取稀疏系数矩阵,即不同深度特征对应的稀疏系数矢量。在此基础上,按照类别对测试样本的深度特征进行联合重构,计算重构误差之和,用于最终的类别决策,如下式所示:

$$\text{identity}(y) = \min_i \sum_{k=1}^K \|y^{(k)} - A_i^{(k)} \alpha_i^{(k)}\| \quad (4)$$

式中: $A_i^{(k)}$ 为第 i 个训练类别对应第 k 个深度特征的局部字典; $\alpha_i^{(k)}$ 为对应的系数矢量。

3.2 识别流程

根据上述分析,文中设计如图 2 所示的识别流程。首先,根据所有训练样本对图 1 所示的 ResNet 进行训练,获得可靠的网络参数。对于测试样本,在 ResNet 的各个卷积层构建相应的深度特征矢量。在该计算书,利用斯皮尔曼等级相关系数进行深度特征

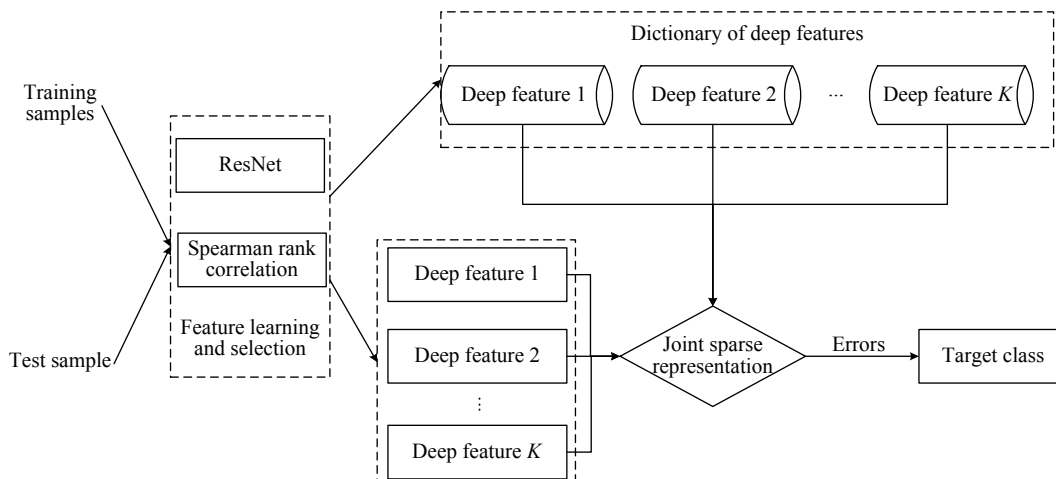


图 2 基于筛选深度特征的红外图像目标识别流程

Fig.2 Procedure of infrared imagery target recognition based on selected deep features

的筛选, 获取若干个深度特征矢量。与此同时, 所有训练样本按照测试样本选取的深度特征进行对应抽取, 并相应构建全局字典。最终, 测试样本选取的深度特征通过联合稀疏表示模型进行分析和重构, 根据最小重构误差的准则判定其所属的目标类别。

4 实验及讨论

4.1 数据集介绍

实验基于中波红外目标数据集 (MWIR) 实施。该数据集在合作条件下对 10 类地面车辆目标 (Pickup、SUV、BTR70、BRDM2、BMP2、T72、ZSU23-4、2S3、MTLB、D20) 进行了全方位高分辨率红外图像采集, 相关数据集的介绍可参见参考文献 [13]。红外观测图像可以反映出不同目标的外形特点与热源分布特征, 从而达到区分不同目标的目标。根据参考文献 [14], 文中对任一类目标选用 120 幅图像作为训练样本, 另外 100 幅不同的图像作为测试样本。

从现有文献选取若干方法进行对比分析, 包含参考文献 [10] 中基于 SVM 的方法, 参考文献 [11] 中基于 SRC 的方法; 参考文献 [13] 中采用 CNN 的方法以及参考文献 [14] 中联合深度特征的方法 (记为 JSR-Deep)。其中 CNN 方法仅仅采用一般的 CNN 网络实现端到端的特征学习与分类。为全面测试提出方法的性能, 后续实验分别在原始测试样本、模拟噪声测试样本以及模拟遮挡测试样本上开展, 通过上述三类场景的测试与对比, 可对提出方法的有效性和稳健性进行较为综合的评估。

4.2 结果与讨论

4.2.1 原始测试样本

首先基于原始测试样本进行实验。如前文所述, 原始测试样本与训练样本一样采集自合作条件, 两者之间存在较高的相似度, 因此识别难度相对较低。采用提出方法对各类目标的 100 个测试样本进行识别, 正确识别样本数目分别为 98、98、98、100、100、100、99、97、98、98 (对应 4.1 节描述的目标类别顺序), 据此可得各类目标的正确识别率。为综合评估方法性能, 文中定义平均识别率为 $P_{av} = N_c / N_T$, 其中 N_c 为正确识别样本数, N_T 为测试样本总数。在此条件下, 分别计算文中方法和四类对比方法的平均识别率, 如表 1 所示。对比可知, 文中方法在当前条件下的平均识别率为 98.4%, 高于对比方法, 体现其更强的有效性。与 CNN 方法相比, 文中通过引入鉴别力更强的 ResNet 进一步提升了分类能力。与 JSR-Deep 方法相比, 文中通过优化网络结构以及进行必要的深

表 1 各方法对原始测试样本的结果对比

Tab.1 Comparison of results from different methods on the original test samples

Method	Average recognition rate
Proposed	98.4%
SVM	94.5%
SRC	95.1%
CNN	96.6%
JSR-Deep	97.8%

度特征筛选保证了特征的可靠性,进一步促进了识别性能的提升。

4.2.2 模拟噪声样本

噪声发生在红外图像采集和后续处理的各个环节,因此实际过程中非合作条件下的红外图像往往受到较为严重的噪声干扰。相比合作条件下的高信噪比 (SNR) 红外图像,噪声干扰下的样本中,目标特性变得模糊,从而加剧了识别难度。文中在原始样本的基础上,按照设定信噪比的原则对其进行加性噪声添加,构造信噪比分别为 10 dB、5 dB、0 dB、-5 dB 和 -10 dB 条件下的噪声测试集。在此基础上,分别采用提出方法和对比方法对五个不同噪声测试集进行分类,统计结果如图 3 所示。文中方法在各个噪声水平下均取得最高的平均识别率,体现其最优的噪声稳健性。CNN 方法在噪声干扰条件下性能下降最为剧烈,反映了深度学习模型对于噪声等异常条件存在一定的不适应性。JSR-Deep 通过对多层次深度特征的联合表征,一定程度降低了噪声的影响,提升了识别稳健性。文中方法则是通过结合 ResNet 优势,同时进行剔除不必要的特征,确保了最终对于噪声样本的识别性能。

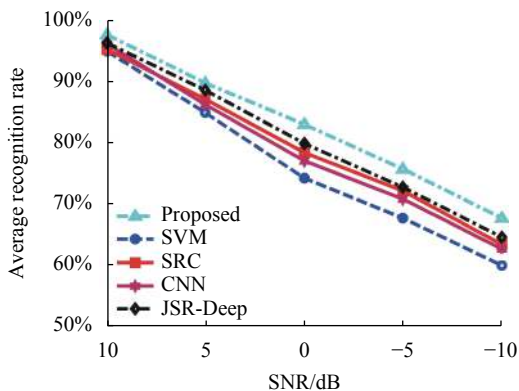


图 3 各方法对模拟噪声样本的结果对比

Fig.3 Comparison of results from different methods on the simulated noisy test samples

4.2.3 模拟遮挡样本

与噪声干扰的情形类似,在非合作条件下获取的红外目标图像也可以存在遮挡的情况,导致目标局部特性不能体现在获取的图像内。为测试提出方法在遮挡条件下的性能,文中首选对原始测试样本中的目标部分区域进行模拟遮挡。遮挡水平采用被遮挡目

标区域占目标全部区域的比例进行衡量。据此,该实验构造了遮挡比例分别为 5%、10%、15%、20% 和 25% 条件下的测试集,进而对提出方法和对比方法进行测试,获得如图 4 所示的统计结果。由此可以看出,提出方法在遮挡条件下相对可保持更强的稳健性。CNN 方法的结果与噪声干扰类似,性能下降最为剧烈。其余方法的性能趋势也与噪声干扰基本相同。文中方法综合了 ResNet 多层次深度特征的优势,通过进行了必要的特征筛选,确保影响正确识别的不良因素得以剔除,因此可进一步提高方法对于遮挡样本的识别稳健性。

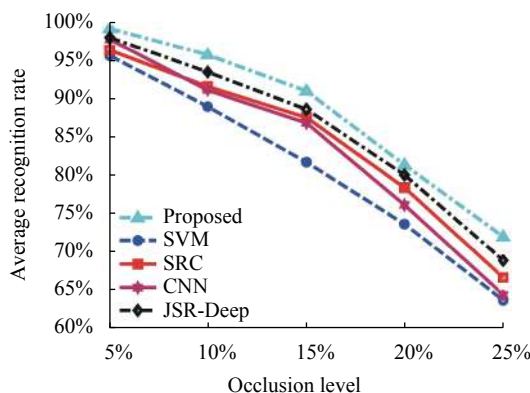


图 4 各方法对模拟遮挡样本的结果对比

Fig.4 Comparison of results from different methods on the simulated occluded test samples

5 结论

文中提出基于筛选深度特征的红外图像目标识别方法。在特征提取阶段,采用 ResNet 学习红外图像的多层次深度特征。同时,基于斯皮尔曼等级相关系数筛选其中的部分深度特征,剔除不必要的冗余成分。在分类阶段,采用联合稀疏表示对筛选得到的深度特征进行分析,获取最终的决策结果。因此,方法可结合 ResNet 多层次深度特征的优势,提高最终识别性能。实验在中波红外目标数据集上开展,分别对原始测试样本、模拟噪声样本和模拟遮挡样本进行分类,结果验证了方法的有效性和稳健性。

参考文献:

[1] Cao H J, Liu N, Xu J, et al. Infrared image adaptive inverse histogram enhancement technology [J]. *Infrared and Laser*

- Engineering*, 2020, 49(4): 0426003. (in Chinese)
- [2] Gu Y, Liu J, Shen H H, et al. Infrared dim-small target detection based on an improved multiscale fractal feature [J]. *Optics and Precision Engineering*, 2020, 28(6): 1375-1386. (in Chinese)
- [3] Wang Y Y, Zhao Y H, Luo H B, et al. Dynamic range compression and detail enhancement of sea-surface infrared image [J]. *Infrared and Laser Engineering*, 2019, 48(1): 0126003. (in Chinese)
- [4] Wang H Q, Xu T F, Sun X L, et al. Infrared-visible video registration with matching motion trajectories of targets [J]. *Optics and Precision Engineering*, 2018, 26(6): 1533-1541. (in Chinese)
- [5] Qiao T Y, Cai L H, Li N, et al. Opposite target measurement based on infrared radiation characteristic system [J]. *Chinese Optics*, 2018, 11(5): 804-811. (in Chinese)
- [6] Wu C L, Hao Y P, Zhang L, et al. Infrared target recognition algorithm based on multi-feature fusion [J]. *Journal of Projectiles, Rockets, Missiles and Guidance*, 2019, 39(3): 39-44. (in Chinese)
- [7] Su J, Yang L, Zhang Y Y. Infrared target recognition algorithm based on fragment matching and graph searching [J]. *Acta Armamentarii*, 2015, 36(5): 854-860. (in Chinese)
- [8] Xie Z H, Liu G D. Infrared face recognition based on co-occurrence histogram of multi-scale local binary patterns [J]. *Infrared and Laser Engineering*, 2015, 44(1): 391-397. (in Chinese)
- [9] Zhao A G, Wang H L, Yang X G, et al. An affine invariant method of forward looking infra-red target recognition [J]. *Laser & Optoelectronics Progress*, 2015, 52: 071501. (in Chinese)
- [10] Zhang D, Zhang J S, Yao K M, et al. Infrared ship-target recognition based on SVM classification [J]. *Infrared and Laser Engineering*, 2016, 45(1): 0104004. (in Chinese)
- [11] Su J, Yang L, Zhang Y Y. An infrared target recognition method based on kernel sparse coding [J]. *Infrared Technology*, 2016, 38(3): 230-235. (in Chinese)
- [12] Huang L H, Cao L H, Li N, et al. A state perception method for infrared dim and small targets with deep learning [J]. *Chinese Optics*, 2020, 13(3): 527-536. (in Chinese)
- [13] D'acremont A, Fablet R, Baussard A, et al. CNN-based target recognition and identification for infrared imaging in defense systems [J]. *Sensors*, 2019, 19: 2040.
- [14] Shi G J. Target recognition method of infrared imagery via joint representation of deep features [J]. *Infrared and Laser Engineering*, 2021, 50(3): 20200399. (in Chinese)
- [15] Zhu C P, Yang Y. Face detection and recognition in monitoring video based on YOLO2 and ResNet algorithm [J]. *Journal of Chongqing University of Technology (Natural Science)*, 2018(8): 170-175. (in Chinese)
- [16] Shi G Z, Chen M, Zhang C Y. Accurate traceability system of crab based on improved deep residual network [J]. *Chinese Journal of Liquid Crystals and Displays*, 2019, 34(12): 1202-1209. (in Chinese)
- [17] Jia X F, Guo Y C, Huang Y R, et al. Salt & pepper noise detection algorithm for color images based on Spearman rank correlation [J]. *Journal of University of Science and Technology of China*, 2019, 49(1): 63-70. (in Chinese)
- [18] Wang T. Infrared small target detection based on Spearman rank correlation coefficient [J]. *Journal of Science Technology and Engineering*, 2017, 17(2): 234-238. (in Chinese)
- [19] Ding B Y, Wen G J. Exploiting multi-view SAR images for robust target recognition [J]. *Remote Sensing*, 2017, 9(11): 1150.