

基于迁移学习 SAE 的无人机目标识别算法研究

谢冰¹, 段哲民¹, 郑宾², 殷云华²

- (1. 西北工业大学 电子信息学院, 陕西 西安 710072;
2. 中北大学 电气与工程学院, 山西 太原 030051)

摘要: 无人机在复杂战场环境下, 因敌我双方无人机外形、颜色等特征较为相似, 如何准确地对敌方无人机识别是实现其自主导航及作战任务执行的关键。由于受敌方无人机飞行速度、形状、尺寸、姿态等的改变及气象环境因素的影响, 无法准确地对其进行识别与分类。针对这一问题, 提出基于迁移学习卷积稀疏自动编码器(Sparse Auto-Encoder, SAE)实现对航拍多帧图像中敌方目标对象的识别与分类。算法首先借助 SAE 对源领域数据集中大量无标记样本进行无监督学习, 获取其局部特征; 然后, 采用池化层卷积神经网络(CNN)算法提取目标图像全局特征; 最后, 送入 Softmax 回归模型实现目标对象的识别与分类。实验结果表明: 与传统非迁移学习的 SAE 算法及基于底层视觉特征学习的识别算法相比, 该算法具有更高的准确性。

关键词: 无人机自主导航; 目标识别分类; 稀疏自动编码器; 卷积神经网络; 迁移学习中图分类号: TP391.4 文献标志码: A DOI: 10.3788/IRLA201847.0626001

Research on UAV target recognition algorithm based on transfer learning SAE

Xie Bing¹, Duan Zhemin¹, Zheng Bin², Yin Yunhua²

- (1. School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China;
2. School of Electrical and Control Engineering, North University of China, Taiyuan 030051, China)

Abstract: UAV in complex battlefield environment, because the two sides of the UAV shape, color and other characteristics are more similar, how to identify enemy UAV accurately is the key to realize the autonomous navigation and combat mission execution. However, due to changes in the speed, shape, size, attitude of enemy UAV and the impact of meteorological and environmental factors, they can not be accurately identified and classified. Aiming at this problem, a kind of sparse auto-encoder(SAE) based on the transfer learning was proposed, and the target objects in the multi-frame aerial images were identified. The algorithm firstly used SAE to study the unsupervised learning of a large number of unmarked samples in the data concentration of source domain, and obtained its local characteristics. Then, the global feature response of the aerial image in the target domain was extracted by the convolution neural network (CNN) algorithm. Finally, the different categories of target objects were identified and classified by the

收稿日期: 2018-01-02; 修订日期: 2018-02-18

基金项目: 瞬态冲击技术重点实验室基金(61426060103162606007)

作者简介: 谢冰(1989-), 男, 博士生, 主要从事红外目标识别及红外超分辨率重建方面的研究。Email: 664897555@qq.com

导师简介: 段哲民(1953-), 男, 教授, 博士生导师, 主要从事目标识别与图像处理方面的研究。Email: zhemin@nwpu.edu.cn

Softmax regression model. The experimental results show that new algorithm proposed in this paper for multiple target objects in aerial multi-frame images is better than more traditional non-transfer learning SAE algorithm, and underlying visual feature recognition transfer learning algorithm, which has higher recognition rate.

Key words: autonomous navigation of UAV; classification and identification of target objects; sparse auto-encoder; convolutional neural network; transfer learning

0 引言

复杂战场环境下, 如何实现对敌方空中目标进行有效及准确识别, 是实现无人机定位、导航、避障及执行攻击任务的关键, 对敌方空中多目标对象的识别能够保障我方无人机在复杂空天战场完成侦察及作战任务。目前, 基于视觉的目标对象识别与分类技术已逐步成为国内外研究的焦点。目标对象识别的关键在于对目标对象特征的提取, 目前, 在目标识别领域以统计学特征提取方法主要包括: 基于时域的 HOG 方法、LBP 方法、GIST 方法、彩色直方图方法以及基于频域的 FFT、小波变换、Gabor 变换、哈达玛变换以及 K-L(PCA)变换等特征提取方法。这些特征提取的方法主要依靠人工经验以及先验知识来获取目标图像中的灰度、颜色、纹理等底层视觉特征, 该类方法很难获取有效表达的目标图像本质特征信息^[1], 对目标图像的识别^[2]和分类精度不高。

近年来, 面向海量无标记数据的无监督特征学习技术已经成为世界各国专家学者研究的热点, 其通过模拟人眼^[3]对图像进行扫描感知, 进而获取数据内部最本质的特征信息, 通过对原始信号进行逐层特征变换, 可将样本从原空间的特征表示变换到新的特征空间, 自动地学习目标图像^[4]深层次化的特征表示, 从而更有利于目标对象的识别和分类。卷积稀疏自动编码器 (Sparse Auto-Encoder, SAE) 方法^[5-7]能够对大量无标记的样本数据进行自我复原训练, 并将特征提取技术推广到很多有标记且样本数有限的应用领域, 该模型不需事先定义特征, 并且能够自动学习隐含在数据中的关系, 从而使特征学习更具表达力。

针对复杂的无人机作战环境下, 由于航拍到的敌方无人机样本数量极少, 无法为无监督特征学习提供所需要的大量训练样本, 因此, 文中提出一种基

于迁移学习的卷积稀疏自动编码器算法对我方无人机飞行途中航拍到的多帧敌方无人机图像进行识别。算法具体设计如下: 首先, 对源领域中大量无标记图像进行图像子块的随机采样, 利用 SAE 学习各局部特征^[8]; 然后, 采用池化层卷积神经网络, 将 SAE 学习到的局部特征向量在三类航拍敌方无人机图像上进行卷积, 得到其全局特征响应; 最后, 将全局特征响应送入 softmax 回归模型进行分类。实验验证了文中提出迁移学习 SAE 算法^[9]能够对三类航拍多帧敌方无人机图像进行有效、准确地识别和分类^[10]。

1 算法设计流程

文中提出算法主要包含 4 个部分: 采样和预处理模块、源领域特征学习模块、目标领域全局特征提取模块以及目标识别分类模块等。算法整体框架如图 1 所示。

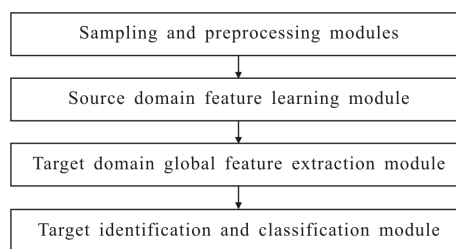


图 1 算法设计框架

Fig.1 Algorithm design frame

如图 1 所示, 采样和预处理模块从源领域无标记数据集 STL-10^[11]中随机采集多个图像子块, 并对采集的图像子块进行白化处理, 从而加强各图像块的边缘特性; 无监督特征学习模块采用反向传播方法对 SAE 神经网络^[12]进行训练, 在隐藏层中加入稀疏性约束学习源领域各图像子块的局部特征; 基于池化层的卷积神经网络 (Convolutional Neural Network, CNN) 技术对目标领域的全局特征进行提

取,从而获取目标领域航拍图像各不同位置的全局特征响应,通过池化操作得到具有旋转及缩放不变性的图像特征;目标识别与分类模块将目标领域图像学习到的全局特征向量送入 softmax 回归模型,从而实现目标对象的识别与分类。

2 算法设计

2.1 图像局部特征提取

2.1.1 自动编码器

自动编码器是无监督学习算法中的一种^[13-14],其根据目标输入与输出值相等,即 $y^{(i)}=x^{(i)}$ 来进行数据的自我复原训练。传统自动编码器主要包括输入层、隐藏层和输出层。其模型如图 2 所示。

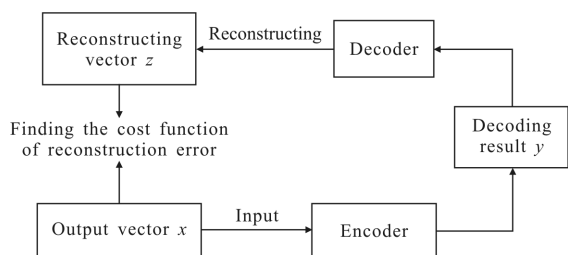


图 2 稀疏编码器模型图

Fig.2 Model diagram of automatic encoder

如图 2 所示,给定无标记数据集 $\{x^{(i)}\}_{i=1}^m$ 后,对输入层中的训练数据 $x^{(i)}$ 进行编码,得到隐藏层特征表达公式(1):

$$y^{(i)}=f_{\theta}(x^{(i)})=s(Wx^{(i)}+b) \quad (1)$$

式中: $\theta=(W, b)$, W 为输入层权值矩阵, b 为输入层偏置向量; $s(x)$ 为激活函数,选用 sigmoid 激活函数,并且令 $s(x)=1/(1+e^{-x})$; 然后,对隐藏层特征表达进行解码操作映射到输出层 z , 由此可得重构向量 z' , 如公式(2)所示:

$$z'=g\theta'(y^{(i)})=s(W'y^{(i)}+b') \quad (2)$$

式中: b' 为隐层的偏置向量; $\theta'=(W', b')$, W' 为隐层到输出层的权值矩阵, 并取 $W'=W^T$; 将参数合并为 $\theta=\{\theta, \theta'\}=\{W, b, b'\}$; 最后,通过训练调节 θ 值,使输入与输出的重构误差代价函数为最小值,传统自动编码器采用代价函数 MSE 如公式(3)所示:

$$j_{\text{cost}}(\theta)=j_{\text{MSE}}(\theta)=\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n \left(\frac{1}{2} \|z_j^{(i)} - x_j^{(i)}\|^2 \right) \quad (3)$$

式中: n 为样本维数。

2.1.2 SAE

SAE 由 Bengio 在 2007 年提出,其通过对自动编码器隐层单元的响应添加约束条件,使得隐藏层单元的大多神经元处于“抑制状态”,而只有少量的神经元处于“兴奋状态”。从而提高了算法对数据特征学习的能力,文中采用一种零相位成分分析(Zero-phase Component Analysis, ZCA)方法,对 SAE 学习到的源领域图像子块进行白化处理,从而强化图像子块的边缘信息。假设源领域采集到的第 i 个图像块的尺寸为 $n \times n$, 按照 R、G、B 分量对其排序,可得到 $m=n \times n \times 3$ 维的向量 $x(i)$, 白化处理后可得输入向量 $x'(i)=W_{\text{white}}x^{(i)}$, W_{white} 表示 $m \times m$ 维白化变换系数矩阵。则 SAE 的第 i 维隐层响应向量为^[14]:

$$a^{(i)}=\sigma(Wx^{(i)}+b_1)=\sigma(W_{\text{SAE}}W_{\text{white}}x^{(i)}+b_1) \quad (4)$$

式中: W_{SAE} 为连接 SAE 隐层和白化处理后各个图像块的输入权重; b_1 代表输入偏置; $\sigma(\cdot)$ 为激活函数。而 $W=W_{\text{SAE}}W_{\text{white}}$ 是白化处理后的整体权重系数,代表隐藏层与原始数据间的映射关系。经白化处理后,输入数据将超出 $[0, 1]$ 的范围,因此,在数据重建时不再需要激活函数 $\sigma(\cdot)$ 对 SAE 的输出进行映射:

$$\hat{x}^{(i)}=W_{\text{SAE}}^T a^{(i)}+b_2 \quad (5)$$

式中: $\hat{x}^{(i)}$ 是第 i 个复原样本; W_{SAE}^T 是输出权重; b_2 是输出偏置。SAE 神经网络通过反向传播训练寻找最小代价函数。为防止过拟合和保持隐层响应的稀疏性,需要对代价函数加入权重衰减项和稀疏性惩罚项,整体代价函数为:

$$J_{\text{SAE}}(W_{\text{SAE}}, b_2)=\frac{1}{N} \sum_{i=1}^N \frac{1}{2} \|\hat{x}^{(i)} - W_{\text{white}}x^{(i)}\|^2 + \lambda \|W_{\text{SAE}}\|^2 + \beta \sum_{j=1}^s \left(\rho \log \frac{\rho}{\hat{\rho}_j} + (1-\rho) \log \frac{1-\rho}{1-\hat{\rho}_j} \right) \quad (6)$$

式中: N 是无标记训练样本数; λ 是权重衰减系数; β 是稀疏惩罚权重; s 是隐层单元数; ρ 是目标稀疏值; $\hat{\rho}_j$ 是所有训练样本在第 j 个隐层单元上的平均响应。

对 SAE 进行训练后得到的输入权重 W_{SAE} 是对应图像子块不同位置的权重系数,根据该权重系数在某图像块上得到的隐层响应就是该图像块的特征。假设从源领域学习包含白化处理的整体权重为 W_s , 输入偏置为 b_{1s} , 目标领域下某图像块 x_T , 对该图像块的特征响应如公式(7)所示:

$$a_i = \sigma(W_s x_i + b_{is}) \quad (7)$$

2.2 基于 CNN 目标图像全局特征提取与分类

为了得到敌方三种不同类型无人机多帧图像中的全局特征,首先,将利用 SAE 算法从源领域数据库中学习到的 $n \times n$ 维图像子块局部特征在每帧航拍敌方无人机图像上进行卷积操作,进而得到每帧航拍敌方无人机的特征响应;然后,将这些特征响应组合为各帧图像的全局特征,并用于后续的分类训练。全局特征提取与分类的整体结构框图如图 3 所示。

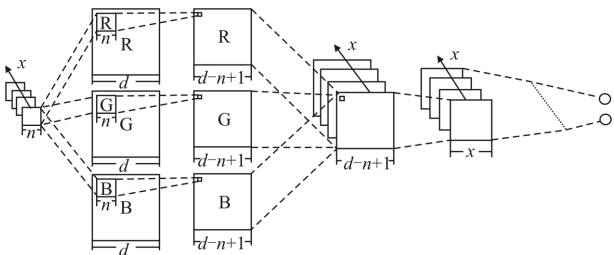


图 3 目标图像全局特征提取和分类结构框图

Fig.3 Global feature extraction and classification structure diagram of target image

2.2.1 全局特征提取

如图 3 所示,假设 SAE 在 $n \times n$ 维图像子块中学习得到 K 个局部特征(K 为 SAE 中隐层单元数),将这些局部特征在尺寸为 $l \times l$ 的每帧航拍无人机图像上进行步长为 1 的逐像素点卷积^[15],得到 K 个 $(l-n+1) \times (l-n+1)$ 维全局特征图案。

文中利用 SAE,在 RGB 色彩空间上进行局部特征值提取,为了提高运算效率,文中在对每帧图像卷积提取全局特征之前,通过 SAE 学习到每个局部特征权重,并按照 3 个色彩通道分别进行拆分;然后,将其与 $l \times l$ 维各帧航拍敌方无人机图像的 R、G、B 分量分贝进行逐点卷积,从而得到 3 个 $(l-n+1) \times (l-n+1)$ 维全局特征向量;最后,分别对 3 个分量上的全局特征向量求和,以得到每帧航拍敌方无人机图像的最终全局特征向量。

$$a = f(WW_{ZW}^{RGB} x' + b) \quad (8)$$

式中: W 和 b 分别代表 SAE 学习到的局部特征系数; x' 代表训练图像样本中每个待卷积子区域的值; W_{ZW}^{RGB} 代表白化预处理阶段中的 ZCA 白化系数。

2.2.2 池化层

CNN^[15]中的池化(pooling)层是对前一层不同位

置的特征进行聚合,例如:将前一层特征图案中的某个区域的平均值(或最大值)作为该区域统计表示。该操作可降低维度过拟合,并且使聚合特征在空间上具有缩放及旋转不变。如图 3 所示, $(l-n+1) \times (l-n+1)$ 维全局特征图案经过池化操作将变为 $p \times p$ 维,得到 $p \times p \times k$ 维特征。

目前,最常用的池化方法是对池化区域取均值或者最大值作为聚合进行结果的统计。参考文献[5]得出结论:在白化预处理情况下,卷积自动编码器更适宜采用平均池化方法。因此,文中采取该原则,选择平均池化方式。在最后输出分类层采用 Softmax 回归模型对三类航拍多帧敌方无人机进行训练,从而得出网络最后一层的全连接参数。

3 实验验证

分别采用文中提出的分类算法、非迁移学习 SAE 分类算法以及基于底层视觉特征描述子的分类算法对三类航拍敌方无人机图像进行识别并分类。为保证实验结果的一般性,每种算法均采用 5 次实验取平均值得出,借助识别准确性(accuracy)指标对算法进行性能比较验证。实验中的部分目标样本集为我方无人机在飞行过程中,机载传感器在同一时刻不同焦距拍摄的 300 帧三类不同姿态、不同形状、不同飞行速度的 RGB 敌方无人机图像,每帧图像的整体几何尺寸为 64×64 ,其中包括场景和目标,需要识别的三类无人机目标图像在整幅图像中占据的比例很小,部分样本如图 4 所示。

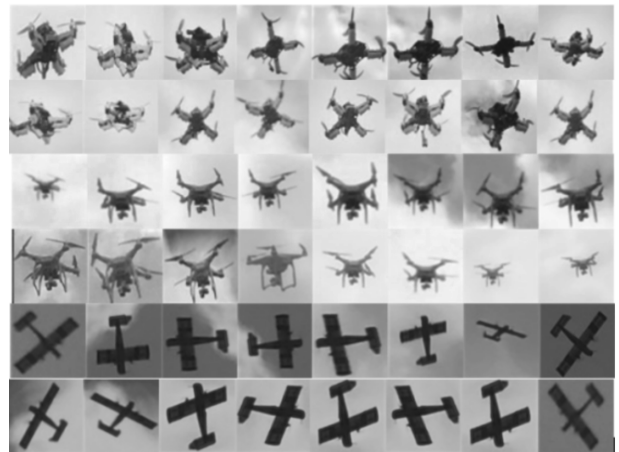


图 4 三类敌方无人机图像

Fig.4 Three kinds of UAV images

迁移学习所使用的源数据集为 STL-10, 该数据集来源于世界最大的图像识别数据库 ImageNet, 其是专门面向无监督学习算法研究的数据库。该数据集中的标记数据包含了 10 类 100 000 张 96×96 大小的 RGB 彩色图片, 内容分布广泛。该数据库中的部分图片如图 5 所示。



图 5 STL-10 数据集上部分图像

Fig.5 Part of the image on the STL-10 data set

3.1 文中算法与非迁移学习 SAE 分类算法比较

非迁移学习 SAE 分类算法直接对航拍 300 帧不同类型的敌方无人机进行随机采样 10 000 个 8×8 图像块进行局部特征学习; 文中基于迁移学习 SAE 算法首先对 STL-10 数据库采集与非迁移学习相同数量的图像块进行局部特征学习, 文中提出算法的实验流程如图 6 所示。

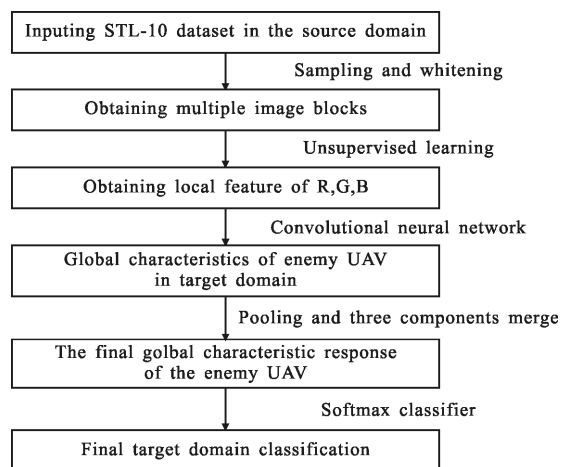


图 6 文中算法实验流程图

Fig.6 Experimental flow chart of the algorithm

白化预处理阶段的正则化常数设置为 0.1, SAE 隐藏层单元数 $s=400$ (即: 对应的 400 个自学习特征), 权重衰减系数 $\lambda=3e-3$, 稀疏惩罚权重 $\beta=5$, 目标稀疏值=0.035; 全局特征提取时, 卷积网络的池化区域尺寸设置为 19×19; 无人机分类时, 设置 softmax 回归模型迭代次数依次为 10、30、50、70 和 100, 来测试文中算法的稳定性。最后, 分别选取三类敌方无人机中的 60 幅图片作为测试集进行分类实验, 在不同迭代条件下进行 5 次交叉验证, 以 accuracy 指标评价迁移学习和非迁移学习算法的分类性能。传统非迁移学习 SAE 算法对三类敌方无人机学习到的局部特征如图 7 所示。

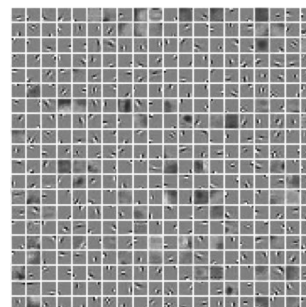


图 7 SAE 在三类航拍敌方无人机所学的局部特征

Fig.7 Local features of three types of aerial enemy UAV learned by SAE

文中提出的迁移学习 SAE 算法首先在 STL-10 数据集上进行局部特征学习, 学习到的局部特征向量可视化表示分别如图 8 所示。

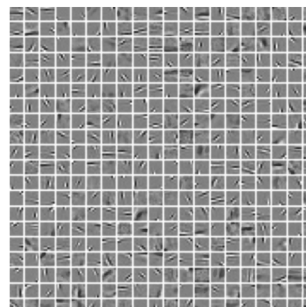


图 8 SAE 在 STL-10 上所学的局部特征

Fig.8 Local feature on STL-10 data set learned by SAE

图 8 中, 从 STL-10 数据库上学习到的局部特征边缘较为清晰; 图 7 所示在三类敌方无人机数据集上学习到的局部特征权重比较模糊, 且每个局部特征向量的边缘性能较弱。

由于局部特征向量边缘性能强弱反应了局部特

征之间的相似性,相似性越大,则表明利用 SAE 进行局部特征提取的效果越好。

不同迭代次数下,文中提出的迁移 SAE 分类算法与非迁移学习 SAE 分类算法的 Accuracy 指标实验结果如表 1 所示。

表 1 文中提出的迁移 SAE 分类算法与非迁移学习算法的分类 Accuracy 结果

Tab.1 Classification Accuracy based on the proposed transfer SAE algorithm in this paper and non-transfer learning algorithm

Iterations	10	30	50	70	100
Transfer learning algorithm	90.01%	88.11%	91.12%	91.31%	91.54%
Non-transfer learning algorithm	86.32%	82.42%	87.43%	88.54%	88.78%

从表 1 可以看出:相同参数下,随着迭代次数的增加,非迁移学习 SAE 分类算法与文中提出的迁移学习 SAE 分类算法的整体性能趋于稳定。但是,从分类结果的 Accuracy 指标上可以看出:基于迁移学习的分类算法比非迁移学习的分类算法效果要好。这是因为有标记的三类敌方无人机样本数目只有 300 个,在利用 SAE 对其进行局部特征提取后,各局部特征向量间的相似度很小,容易出现过拟合现象。因此,利用相似性小的局部特征向量卷积到每帧目标样本图像上得到的目标样本图像的全局特征向量并不能很好地表示出三类敌方无人机全部特征,造成识别效果不佳。实验结果进一步验证了在缺乏有标记的小样本下,利用迁移学习 SAE 算法能够取得比非迁移学习更好的识别和分类效果。

3.2 文中算法与底层视觉特征描述子分类算法比较

选取 RGB 色彩直方图、GIST(Generalized Search Trees) 描述子以及适用于纹理检测的 LBP(Local Binary Pattern)描述子分别作为分类特征,对 300 帧三类敌方无人机图像进行全局特征值提取,送入 softmax 回归模型对其进行分类。不同迭代次数下本文提出的迁移学习 SAE 分类算法与基于底层视觉特征描述子分类算法在 Accuracy 指标的实验结果如

表 2 所示。

如表 2 所示,不同迭代次数下,文中提出的基于迁移学习 SAE 算法的分类 Accuracy 指标能够得比传统基于 RGB 色彩直方图、GIST 描述子分类算法以及纹理检测的 LBP 描述子分类算法更好的分类性能。实验结果进一步验证了文中基于迁移学习的 SAE 算法能够较好地模拟人眼对图像进行扫描感知来获取目标图像的高层视觉特征,比传统人工经验和先验知识来获取目标图像中的灰度、颜色、纹理等底层视觉特征的方法具有更好的图像表达,对目标图像的识别及分类效果较好。

表 2 文中提出 SAE 分类算法与底层视觉特征描述子分类 Accuracy 结果

Tab.2 Classification Accuracy based on the proposed SAE algorithm in this paper and underlying visual feature descriptor

Iterations	10	30	50	70	100
Proposed algorithm in this paper	90.01%	88.11%	91.12%	91.31%	91.54%
RGB histogram	74.32%	75.66%	76.43%	76.89%	75.79%
GIST descriptor	80.12%	79.87%	81.32%	79.68%	80.78%
LBP descriptor	76.64%	77.65%	78.34%	78.85%	79.66%

4 结 论

为了提高复杂战场环境下对敌方无人机识别与分类准确性困难的问题,文中提出一种基于迁移学习的 SAE 目标图像识别与分类算法,通过从源领域大量无标记图像中学习局部特征,结合 CNN 将其应用于三类多帧敌方无人机图像的识别与分类中。算法首先利用 SAE 对源领域数据集中的大量无标记样本进行无监督学习,学习其局部特征;然后,采用池化层的 CNN 提取目标领域有标记图像的全局特征响应;最后,送入 softmax 回归模型^[6]中对其进行分类。实验结果表明:文中提出的基于迁移学习的 SAE 算法能够应用于小样本多帧敌方无人机图像,较传统 SAE 算法及基于底层特征描述子分类算法取得更好的识别与分类效果。

参考文献:

- [1] Jia Ping, Xu Ning, Zhang Ye. Automatic target recognition based on local feature extraction [J]. *Optics and Precision Engineering*, 2013, 21(7): 1898–1905. (in Chinese)
- [2] Li Min, Zhou Zhenhua, Zhang Guilin. Image metric in performance evaluation of automatic target recognition algorithm[J]. *Infrared and Laser Engineering*, 2007, 36(3): 412–416. (in Chinese)
- [3] Xu Fang, Liu Jinghong, Zeng Dongdong, et al. Unsupervised ship detection and recognition based on Visual salience [J]. *Optics and Precision Engineering*, 2017, 25(5): 1300–1311. (in Chinese)
- [4] Jiang Shan, Zhang Rui, Han Guangliang, et al. Multi-feature fusion moving target tracking in gray image with complex background[J]. *Chinese Optics*, 2016, 9(3): 320–328. (in Chinese)
- [5] Lecun Y, Bengio Y, Hinton G. Deep learning [J]. *Nature*, 2015, 521(7553): 436–444.
- [6] Chen T, Borth D, Darrell T, et al. DeepSentibank: visual sentiment concept classification with deep convolutional neural networks[J]. *Computer Science*, 2014, 65(5): 1087–1095.
- [7] You Q, Luo J, Jin H, et al. Robust image sentiment analysis using progressively trained and domain transferred deep networks [C]//29th AAAI Conference on Artificial Intelligence (AAAI), 2015: 381–388.
- [8] Zhang F, Du B, and Zhang L. Saliency-guided unsupervised feature learning for scene classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, 53(4): 2175–2184.
- [9] Zhuang Fuzhen, Luo Ping, He Qing, et al. Progress in transfer learning[J]. *Journal of Software*, 2015, 261: 26–39. (in Chinese)
- [10] Song Jianzhong. The development trend of intelligent image processing[J]. *Chinese Optics*, 2011, 4(5): 431–440. (in Chinese)
- [11] Coates A, Lee H, Ng A Y. An analysis of single-layer networks in unsupervised feature learning [C]//14th International Conference on Artificial Intelligence and Statistics, 2011: 215–223.
- [12] Eng J, Zhang Z, Eyben F, et al. Autoencoder-based unsupervised domain adaptation for speech emotion recognition [J]. *IEEE Signal Processing Letters*, 2014, 21(9): 1068–1072.
- [13] Zhou J T, Pan S J, Tsang I W, et al. Hybrid heterogeneous transfer learning through deep learning [C]//28th AAAI Conference on Artificial Intelligence(AAAI), 2014: 2213–2219.
- [14] Guo Yingkai, Yang Jie, Lu Zhenggang. Unsupervised texture segmentation based on gaussian markov random field and neural network [J]. *Infrared and Laser Engineering*, 2000, 29: 225–229. (in Chinese)
- [15] Vedaldi A, Lenc K. MatConvNet: convolutional neural networks for matlab [C]//23rd ACM International Conference on Multimedia, 2015: 689–692.
- [16] Luo Haibo, Xu Lingyun, Hui Bin, et al. Present situation and prospect of target tracking methods based on deep learning [J]. *Infrared and Laser Engineering*, 2017, 46(5): 0502002. (in Chinese)