基于深度神经网络的低空弱小无人机目标检测研究

王靖宇[1,2], 王霰禹[1,2], 张科[1,2], 蔡宜伦[1,2], 刘越[1,2]

(1.航天飞行动力学技术重点实验室, 陕西 西安 710072; 2.西北工业大学 航天学院, 陕西 西安 710072)

摘 要:针对低空无人机目标视觉特征较弱,传统识别模型在目标尺度较小时易受干扰导致识别精度下降等问题,提出了一种基于多隐含层深度神经网络的弱小无人机目标检测模型。根据低空监视图像输入特性和弱小无人机目标视觉表征特点,设计了包含多个隐含层的多通道深度神经网络模型结构,并通过建立多尺度、多角度、多背景条件下的无人机目标图像数据库,完成了对深度网络模型参数的训练及优化。仿真结果表明,所设计的深度模型对低空无人机目标具有较好的变尺度检测能力和抗干扰效果,体现出良好的鲁棒性和潜在工程应用前景。

关 键 词:低空无人机;目标识别;深度神经网络;多隐含层

中图分类号:TP391.4 **文献标志码**:A **文章编号**:1000-2758(2018)02-0258-06

随着无人机使用门槛的不断降低,由于缺乏统一的行业标准和规范,导致无人机的“黑飞”问题日益突出,使得无人机被滥用的可能性在显著增加,低空空域内无人机的非合作入侵飞行事件在国内外已屡见不鲜,不仅对公民隐私和生命财产安全造成了危害,严重制约了无人机产业化的发展,更是对公共安全和国家安全带来了十分巨大的威胁[1-2]。因此,需要采用视频图像信息对低空无人机目标进行有效检测,以实现后续的防护压制。

近年来,随着以深度学习为代表的人工智能方法在图像处理领域的不断发展,大量目标检测方法都采用深度神经网络来完成任务。文献[3]针对利用图像数据与深度信息进行目标识别时,存在信息融合不一致、特征设计不统一等问题,研究了基于卷积神经网络的多模信息融合识别方法,使得特征学习可以根据目标特性自动进行。然而,无人机在远距离上的视觉特性较弱、尺寸较小,是较为典型的弱小视觉目标。目前,基于深度神经网络的弱小目标检测方法研究主要针对交通标志、空中飞鸟等弱小目标。其中,文献[4]提出了一种基于深度神经网络的小尺寸道路标示识别框架,通过采用全卷积网络来提供道路标识的位置线索,并采用深度卷积神经网络来实现对标识的检测。文献[5]针对图像中鸟类弱小目标的检测问题,提出了一种基于卷积神经网络和全卷积网络的检测识别方法,通过提取多尺度特征来有效区分鸟和背景区域,从而实现对鸟类弱小目标的检测与识别。

此外,蒋兆军等人[6]针对基于深度学习的无人机检测方法进行了研究,提出通过训练卷积神经网络来实现对无人机的识别。文献[7]则将无人机目标的检测任务简化为检测问题,通过对背景图像和前景运动区域进行建模,利用帧间差分原理实现了对弱小无人机目标的检测和识别。然而,上述研究工作只采用了不包含其他运动目标的视频和图像数据进行测试实验,并且训练样本所覆盖的目标尺度较为固定,基本没有考虑无人机的尺度变化问题,使得上述方法在用于实际弱小无人机目标检测时具有较大局限性。与此同时,由于采用深度学习方法进行目标检测时,训练数据的数量与质量对提高检测模型的精度和泛化能力来说较为关键。但由于无人

收稿日期:2017-06-12 **基金项目**:国家自然科学基金(61174204,61101191,61502391)、航天支撑基金(N2015KC0121)、航天飞行动力学技术重点实验室开放基金、陕西省自然科学基础研究计划(2017JM6043)与中央高校基本科研业务费资助

作者简介:王靖宇(1985—),西北工业大学讲师,主要从事智能图像处理及探测导引技术研究。

机目标较为特殊,当前还尚未有较为成熟的大型图像数据库可以用来进行模型训练,使得现有研究工作受到了一定限制。针对上述问题,本文首先建立了由 15 000 张无人机图像组成的低空无人机目标图像数据库,其中包含了不同大小、角度、颜色、纹理和背景的无人机目标。其次,建立了由多个隐含层组成的深度神经网络模型,通过对输入无人机图像进行卷积、池化等操作,分析目标图像各像素间的空间关系,并有效提取目标的图像特征信息,同时,通过将较大尺寸的二维的图像信息转换为较小尺寸下的多通道特征图,以克服目标变尺度检测问题。最后,利用云台式光电探测系统在室外采集的无人机目标飞行视频进行了实验验证。结果表明,本文提出的深度神经网络模型能够有效检测空中弱小无人机目标,并在视场内有其他干扰物存在时具有较好的鲁棒性。

1　低空无人机目标数据库建立

由于现有研究工作中还尚未有较为成熟的无人机目标图像数据库,会对面向无人机目标检测的深度神经网络模型训练带来较大困难。文献[6]在深度神经网络的训练过程中虽然建立并使用了一个具有 3 560 张图像的无人机目标图像数据库,但其训练样本的数量仍较小,难以保证深度神经网络模型在训练过程中能够快速收敛。为了提高深度神经网络模型的检测精度和泛化能力,本文首先建立了一个由 15 000 张无人机目标图像组成的训练数据库。

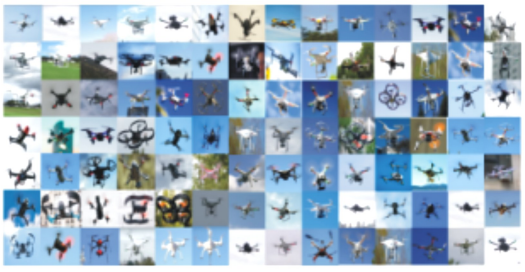


图 1　无人机目标图像数据库中的样本示例

具体而言,先是通过在大型网络图像数据库中进行搜索和在包含无人机目标的视频中截取图像帧,采集了约 20 000 张、包含约 30 种无人机目标的图像样本。随后,采用图像去噪、尺寸归一化等手段对图像进行预处理,得到约 16 000 张质量较好的无人机目标图像。最后,利用人工在图像中对目标进行标注,并去除目标特性较差的图像,得到 15 000 张含有无人机目标的图像样本来组成训练数据库。本文所建立的无人机目标图像数据库中的样本示例如图 1 所示,可以看出数据库中的样本图像包含了蓝天、草地、阴天、树林等不同环境背景,并且无人机目标的形状、大小和颜色均各不相同,从而提高训练数据库的样本质量。

2　基于深度神经网络的无人机目标检测模型

2.1　多隐含层深度神经网络结构设计

深度神经网络在图像目标检测领域已有广泛应用,并有着非常优异的效果[8]。但现有研究较少考虑待检测目标尺寸,现有用于目标检测的深度神经网络的输入尺寸都大于 100×100 像素,而待检测的原图像尺寸更是大于神经网络的输入尺寸,使得神经网络有充足的数据进行检测。在无人机目标检测过程中,捕捉到的无人机图像通常只有不到 40×40 像素,可用于检测的数据非常有限,因此现有研究提出的深度神经网络并不适用于这类问题。

本文提出并建立的深度神经网络模型结构如图 2 所示,神经网络包括共 23 层,其中 21 层为隐含层,包含 5 层卷积层,3 层全连接层,7 层 ReLU 层,3 层归一化层和 3 个池化层;深度神经网络的输入层接受的数据为 65×65 像素的 RGB 彩色图像,第一层卷积层的卷积核尺寸为 5×5×3,共 64 个,进行步长为 2 的卷积操作,随后经过步长为 2、尺寸为 3×3 像素的池化层和一个正则化层后,输出一组尺寸为 17×17×64 的数据。利用小尺度卷积核对数据进行细致的滤波处理,充分提取输入图像中弱小无人机目标的多层次视觉特征。从而使得所建立的深度神经网络能够较好地适应弱小尺度视觉目标,

第二层卷积层由 128 个尺寸为 3×3×64 的卷积核构成,对第一层输出的数据进行步长为 1 的卷积运算,通过提高卷积核个数来增加图像通道数量,从而为后续的特征提取和检测提供更多数据支持,在经过一次正则化操作后输入第三层卷积层,包含 64 个尺寸为 3×3×128 的卷积核,仅进行一次步长为 1 的卷积操作,输出尺寸为 17×17×64 的数据。在这 2 次卷积后未进行最大池化操作,从而使得本文所

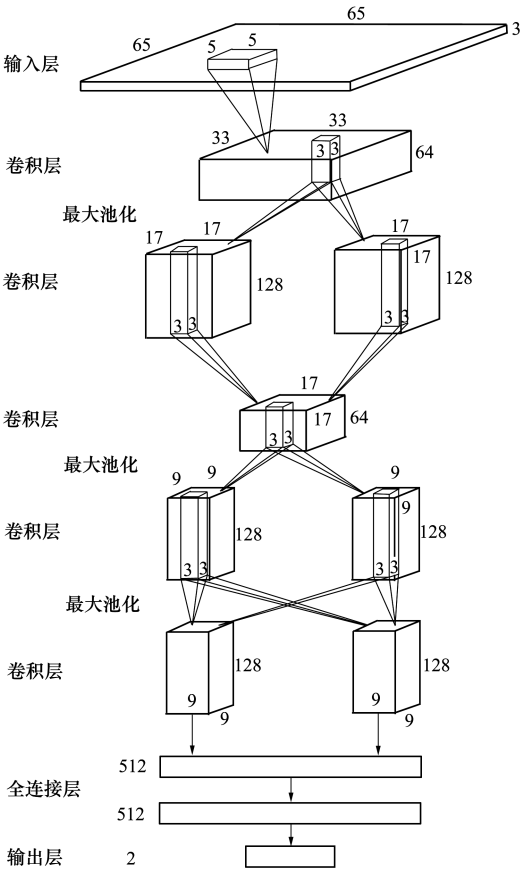


图 2 多隐含层深度神经网络结构

提出的深度神经网络能够充分地保留输入目标的图像细节信息。

第四层卷积层包含 128 个尺寸为 3×3×64 个卷积核,对第三层的输出数据进行步长为 2 的卷积操作,随后进行尺寸为 3×3 像素的池化操作并再一次对所得数据进行一次卷积操作,卷积核尺寸为 3×3×128,共 128 个,最后对数据进行一次步长为 2、尺寸为 3×3 的池化操作,将数据尺寸缩减为 5×5×128。通过对图像进行 5 次卷积操作,将其视觉特征抽象成多通道的小尺度特征图,包含了原图像中不同尺度下的信息,从而能实现对不同尺度目标的准确检测。将这组数据输入由 2 组 512 个神经元构成的全连接层,输出层为 2 个神经元,分别代表“是”和“不是”2 个判断结果。

2.2 基于深度神经网络的目标检测方法

深度神经网络模型的训练采用的是传统的反向传播算法,模型的参数通过随机数进行初始化。利用建立的无人机图像数据库作为训练数据,对图像进行判断,通过将判断结果与标准结果的差反馈到各参数,对参数数值进行微调使其逐渐逼近最佳值。

首先将训练数据通过网络输入层输入深度神经网络,经过每一层的前向传播最终得到判断结果,对于 l 层的前向传播公式为:

$$\begin{cases} z^{(l)} = w^{(l)} a^{(l-1)} + b^{(l)} \\ a^{(l)} = f(z^{(l)}) \end{cases}$$

式中, $a^{(l)}$ 为第 l 层的输出, 对于卷积层,$z^{(l)} = \sum \mathrm{conv}(a_j^{(l-1)}, w_j^{(l)}) + b_j^{(l)}$,随后计算整个网络的代价值 $J(w,b)$

$$\begin{cases} J(w,b) = \dfrac{1}{m}\left(\sum\limits_{i=1}^{m} J(w,b;x^{(i)},y^{(i)}) \right) \\ J(w,b;x^{(i)},y^{(i)}) = \dfrac{1}{2}(y^{(i)} - h_{w,b}(x^{(i)}))^2 \end{cases}$$

式中, $h_{w,b}(x^{(i)})$ 为第 i 个分类的计算值;$y^{(i)}$ 表示第 i 个分类的真实值;m 表示类别数。此处我们要得到最佳的 w 和 b,以使代价值 $J(w,b)$ 最小,即达到分类的目的。

通过反向传播算法计算误差相对各参数的梯度,并通过梯度下降法对各参数进行微调

在本文所建立的深度神经网络训练过程中,网络各参数经过反复微调,逐渐逼近提取和描述无人机目标图像的最佳模型参数,利用上节中所建立的大样本目标图像数据库进行训练,采用较小的学习速度可以使参数更精准拟合多尺度无人机目标的视觉表征模型,使所建立的深度神经网络模型能够更加准确地适用于对不同无人机目标的检测。

但深度神经网络仍存在算法时间复杂度较高的不足,直接利用深度神经网络在视频中进行无人机目标的检测无法满足实时性的需求。深度神经网络时间复杂度的计算公式为:

$$Time \sim O\left(\sum_{l=1}^{D} M_l^2 \cdot K_l^2 \cdot C_{l-1} \cdot C_l \right)$$

式中, D 为神经网络中卷积层数,M_l 为第 l 层输出特征图的边长,K_l 为第 l 层卷积核的边长,C_l 为卷积层的输出通道数。经计算,本文提出的神经网络完成一次识别需要进行超过 6.5×10^7 次运算。

因此本文利用一种时间复杂度较低的运动目标检测算法提取可能存在的无人机目标,本方法通过对传统 Haar 特征进行随机投影降维,描述图像中的局部信息,可以快速准确地表征图像中的弱小目标,通过对比临近帧间的局部信息相似性,提取图像中产生明显变化的区域,实现快速的运动目标检测。深度神经网络仅对图中的运动目标进行检测,从而

减少深度神经网络的计算量,提高算法的整体检测速度。

3　仿真验证与分析

为了验证本文所提出无人机目标检测模型在实际场景中的有效性和准确率,本文建立了一套基于云台式可见光摄像头的低空光电探测系统,其系统组成如图 3 所示。



图 3　云台式可见光摄像头(左)及视频画面(右)

其中,图 3 左侧为低空光电探测系统所采用的云台式可见光摄像头,图 3 右侧图像为该系统在实验环境下所获取的原始视频画面。同时,在测试实验中采用了 2 款国内消费级无人机市场上的主流机型作为无人机目标,其外观如图 4 所示。



图 4　测试飞行无人机示例

如图 4 所示,实验采用了 2 款国内消费级无人机市场上的主流机型作为低空无人机目标,其实际尺寸分别为 438 mm×451 mm×301 mm(左)和 221 mm×205 mm×117mm(右)。在实验测试过程中,选择摄像头的视场场景(见图 3 右)为飞行测试环境并保持当前视场不变,当摄像头实际成像焦距设置为近焦端(f=100 mm)且视频图像分辨率为 1 920×1 080 时,图 4 中的 2 种无人机目标在水平方向 600 m 距离处的成像尺寸约为 15×15 像素。

同时,根据低空光电探测系统的可能部署区域,选择了城区和郊区 2 种环境进行低空无人机目标的探测检测实验。其中,城区环境位于城市的中心区域,高层建筑较多,且摄像头视场中时常出现飞鸟、杂物等其他干扰物体。郊区环境则位于城区和山区相邻交界处的野外区域,无高层建筑但存在较多飞鸟、蚊虫等飞行干扰物体。为了避免无人机飞行对周围环境产生负面影响,实验中无人机目标的飞行高度不低于 50 m 但未超过 150 m,城区环境下水平最远飞行距离约为 800 m,郊区环境下水平最远飞行距离约为 2 200 m,2 架无人机的最大横向机动飞行速度约为 10 m/s。

在城区环境实验中,为了使飞行中的无人机目标能够位于摄像头的视场中央,调整云台在垂直方向上抬高了约 15°视角并带有约 2 倍左右的光学变焦。在实验中,我们采用了移动计算终端对光电摄像头的视频流进行获取和处理,其处理器为 i7-7700HQ,内存为 16G,图形计算卡为 GeForce GTX1050,对视频中每一帧输入图像的处理时间约为 250 ms。需要注意的是,受拍摄时间、太阳位置、空气质量等实验条件变化的影响,图 3 右侧中视频图像的画质与实验结果(见图 5~图 9)中所采集到图像的画质存在视觉上的差异。

图 5 至图 9 为采用本文所建立的深度神经网络对低空无人机目标进行检测的结果。其中,图 5 和图 6 分别为视频画面的第 50 帧和第 150 帧,图中左侧无人机目标为悬停状态,右侧无人机为飞行运动状态。从图中可以看出,绿色方框显示本文算法对 2 个无人机目标都能进行准确检测,并且当右侧无人机进行机动飞行时仍然能够准确地对运动中的无人机目标进行正确检测。

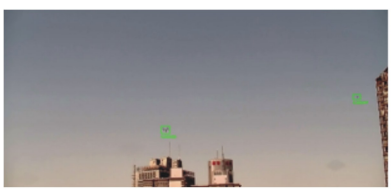


图 5　实验视频图像第 50 帧结果(2 个目标悬停)

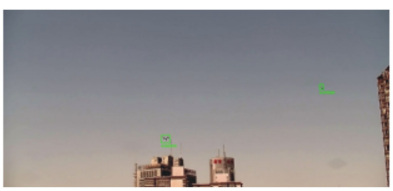


图 6　实验视频图像第 150 帧结果(左侧无人机目标悬停,右侧无人机目标机动)

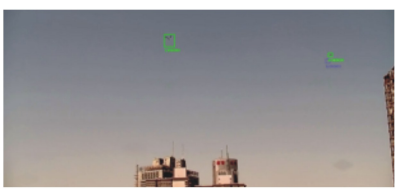


图 7　实验视频图像第 300 帧结果(2 个目标同时机动)

图 7 给出了 2 个无人机目标同时进行机动时的检测结果,为原视频图像的第 300 帧。可以看出,本文算法所产生的绿色滑动窗口具有 99% 的检测概率,能够准确地定位无人机目标的图像所在区域。在图 5 至图 7 中,虽然视频中无人机目标所在区域的背景图像是仅有天空和云层的简单背景,但由于摄像头成像质量较差,实际采集到的视频图像在上述区域仍有大量的图像噪点和毛刺,是较为明显的干扰。

图 7 中右侧无人机目标处的蓝色窗口即为算法产生的其他滑动窗口,但通过采用本文所建立的深度神经网络进行目标检测,实现了对右侧飞行无人机的正确检测(绿色框)。



图 8 郊区实验视频图像第 50 帧结果

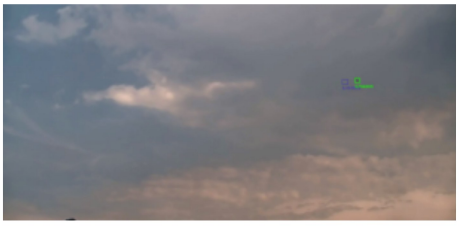


图 9 郊区实验视频图像第 200 帧结果

图 8 和图 9 显示了在郊区环境下,利用本文所建立深度神经网络检测模型得到实验结果的第 50 帧和第 200 帧。其中,无人机在视频中正进行由左向右的快速飞行机动。在该实验环境条件下,无人机目标距离摄像头最远约为 2 200 m,摄像头为 25 倍光学变焦,蓝色和绿色框均为本文算法所产生的检测滑动窗口。可以看出,虽然摄像头已经为 25 倍光学变焦,但由于无人机目标距离较远,其视觉尺寸仅有约 20×20 像素,利用传统算法已经无法将其与其他干扰目标区分。受实验采用的移动计算终端计算能力限制,本文算法的检测结果由绿色窗口表示,但其检测概率为 99%,说明本文所建立的深度神经网络检测模型能够有效实现对远距离弱小无人机目标的检测。

此外,为了充分验证本文所建立深度神经网络模型的检测精度,还采用基于特征和分类器组合的经典检测方法对静态图片中的无人机目标检测进行了对比实验。其中,采用了基于 Dense-SIFT 的强特征对局部图像块进行描述,通过结合 3 种不同的编码策略:词袋(bag of visual words,BoVW),费舍尔向量(Fisher vector,FV)和局部特征聚合描述符(vector of locally aggregated descriptors,VLAD),并利用支持向量机(support vector machine,SVM)作为分类器来得到最终的结果。同时,为了有效验证深度神经网络在处理含噪声数据时的鲁棒性,实验数据不仅含有无噪声的样本,还含有添加了 10 dB 和 20 dB 高斯噪声的图像样本,其实验结果如表 1 所示。

表 1 对比实验的检测结果

指标		DNN	BoVW	FV	VLAD
无噪声	ACC	0.992 7	0.921 4	0.977 2	0.955 5
	mAP	0.991 2	0.917 6	0.977 5	0.954 8
10 dB 噪声	ACC	0.962 7	0.810 8	0.856 9	0.874 0
	mAP	0.9621	0.787 4	0.853 4	0.867 2
20 dB 噪声	ACC	0.958 1	0.901 2	0.946 7	0.945 7
	mAP	0.962 5	0.902 3	0.944 8	0.945 1

其中,ACC 表示精度(accuracy),而 mAP 则表示平均正确率均值(mean average accuracy)。可以看出,采用本文提出的深度神经网络进行分类检测时,其 ACC 和 mAP 指标都要优于采用特征和分类器组合的传统方法。同时,在处理含有不同强度噪声的图像数据时,本文所建立的深度神经网络模型仍能取得较高的总体精度和平均精度,显示出本文方法具有较好的鲁棒性。

3 结 论

本文针对低空弱小无人机目标存在机动飞行和其他视觉干扰物体情况下的检测问题,提出了一种基于多隐含层深度神经网络的目标检测模型。在建立无人机目标大样本数据库的基础上,提出并设计了一个具有多隐含层的无人机深度神经网络检测模型,通过对无人机目标多尺度视觉特性的深度表征与抽取,有效提高了检测模型的泛化能力,实现了对远距离弱小无人机目标的准确检测。实验结果表明,本文所设计的深度检测模型具有较好的变尺度检测能力和抗干扰效果,能够完成对多个无人机目标的实时检测,在工程应用上具有一定参考价值。

参考文献:

[1] 蔡亚梅,姜宇航,赵霜. 国外反无人机系统发展动态与趋势分析[J]. 航天电子对抗,2017,33(2):59-64
Cai Yamei, Jiang Yuhang, Zhao Shuang. Development Status and Trend Analysis of Counter UAV Systems[J]. Aerospace Electronic Warfare, 2017,33(2): 59-64 (in Chinese)

[2] 李国军. 民用无人机恐怖袭击风险评估与防御策略研究[J]. 中国人民公安大学学报:社会科学版,2017,33(3):9-14
Li Guojun. Study on Risk Assessment and Defensive Strategy of Terrorist Attacks on Civil UAV [J]. Journal of People's Public Security University of China, Social Sciences Edition, 2017, 33(3):9-14 (in Chinese)

[3] Wang A, Lu J, Cai J, et al, Large-Margin Multi-Modal Deep Learning for RGB-D Object Recognition[J]. IEEE Trans on Multimedia, 2015, 17(11): 1887-1898

[4] Zhu Yingying, Zhang Chengquan, Zhou Duoyou, et al. Traffic Sign Detection and Recognition Using Fully Convolutional Network Guided Proposals[J]. Neurocomputing, 2016, 214 (suppl C): 758-766

[5] Takeki, Akito, Tu Tuan Trinh, Ryota Yoshihashi, et al. Combining Deep Features for Object Detection at Various Scales: Finding Small Birds in Landscape Images[J]. IPSJ Trans on Computer Vision and Applications 2016 8(1):5

[6] 蒋兆军,成孝刚,彭雅琴,等. 基于深度学习的无人机识别算法研究[J]. 电子技术应用,2017,43(7):84-87
Jiang Zhaojun, Cheng Xiaogang, Peng Yaqin, et al. A Novel UAV Recognition Algorithm Based on Deep Learning Approach [J]. Application of Electronic Technique, 2017,43(7): 84-87 (in Chinese)

[7] Rozantsev, Artem, Vincent Lepetit, Pascal Fua. Detecting Flying Objects Using a Single Moving Camera.[J]. IEEE Trans on Pattern Analysis and Machine Intelligence 2017, 39(5): 879-92

[8] Rawat W, Wang Z. Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review[J]. Neural Computation, 2017, 29(9): 2352-2449

Small UAV Target Detection Model Based on Deep Neural Network

Wang Jingyu[1,2], Wang Xianyu[1,2], Zhang Ke[1,2], Cai Yilun[1,2], Liu Yue[1,2]

(1.National Key Laboratory of Aerospace Flight Dynamics, Xi'an 710072, China;
2.School of Astronautics, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: Unmanned aerial vehicle (UAV) has relatively small size and weak visual characteristics. The recognition accuracy of traditional object detection methods can decrease sharply when complex background and distraction objects exist. In this paper, we proposed a novel deep neural network (DNN) model for small UAV target recognition task. Based on the visual characteristics of surveillance image and UAV target, a multi-channel DNN is designed. Training and optimization of the DNN are completed with self-constructed UAV image database. Simulation results show that the proposed DNN model can achieve good results in recognizing the variable-scale UAV target and have compatible performance in distinguishing the interference and that the proposed model is robust and has a great potential prospect for engineering application.

Keywords: unmanned aerial vehicle(UAV); object recognition; deep neural network(DNN); multi-hidden layer; neural networks; optimization