

无人机目标分类的深度卷积网络设计与优化^①



皮 骏¹, 张志力², 李 想², 张春泽³

¹(中国民航大学 通航学院, 天津 300300)

²(中国民航大学 航空工程学院, 天津 300300)

³(天津讯联科技有限公司, 天津 310600)

通讯作者: 张志力, E-mail: 1397548395@qq.com

摘 要: 针对传统无人机目标分类方法效率低、特征提取能力不足和适应性差等问题, 通过对无人机自身特点和现有分类方法的分析, 提出了引入注意力机制优化深度卷积神经网络的无人机分类方法. 设计多组对比实验, 根据实验效果设计出模型结构为 3 层卷积层、3 层池化层、2 层全连接层的卷积神经网络进行训练, 得到最优的无人机目标分类模型, 再引入卷积注意力模块对特征图元素进行加强和抑制, 引入批归一化层加速模型收敛, 提升泛化能力. 实验结果表明: 引入卷积注意力模块和批归一化层优化后的无人机目标分类模型的识别率达到 92.44%, 较优化前提升 1.5%, 相比于其它神经网络模型具有识别率高、收敛速度快的优点, 可以基本满足实际场景中无人机目标分类的要求.

关键词: 无人机分类; 卷积神经网络; 注意力机制; 批归一化

引用格式: 皮骏, 张志力, 李想, 张春泽. 无人机目标分类的深度卷积网络设计与优化. 计算机系统应用, 2021, 30(5):290-297. <http://www.c-s-a.org.cn/1003-3254/7950.html>

Design and Optimization of Deep Convolutional Neural Network for UAV Target Classification

PI Jun¹, ZHANG Zhi-Li², LI Xiang², ZHANG Chun-Ze³

¹(School of General Aviation, Civil Aviation University of China, Tianjin 300300, China)

²(School of Aeronautical Engineering, Civil Aviation University of China, Tianjin 300300, China)

³(Tianjin Xunlian Technology Co. Ltd., Tianjin 310600, China)

Abstract: Aiming at the problems such as low efficiency, limited ability of feature extraction, and poor adaptability of traditional classification methods for UAV targets, this study proposes a UAV classification method that introduces attention modules to optimize deep convolutional neural networks by analyzing the characteristics of UAVs and existing classification methods. Multiple sets of comparative experiments are designed for a model structure of a convolutional neural network with three convolutional layers, three pooling layers, and two fully connected layers according to the experimental results for training to obtain the optimal classification model for UAV targets. Then, the convolutional block attention module is introduced to strengthen and suppress feature map elements, and the batch normalization layer is introduced to accelerate convergence and improve generalization capabilities of the model. Experimental results show that after introduction of convolution block attention modules and batch normalization layers, the recognition rate of the classification model for UAV targets rises by 1.5% to 92.44%. Its advantages of high recognition rate and fast convergence over other neural network models can basically meet the requirements of UAV target classification in actual scenes.

Key words: UAV classification; Convolutional Neural Network (CNN); attention module; batch normalization

^① 收稿时间: 2020-09-22; 修改时间: 2020-10-21, 2020-11-12; 采用时间: 2020-11-17; csa 在线出版时间: 2021-04-28

随着科技水平的提高,无人机(Unmanned Aerial Vehicle, UAV)在军事领域发展迅猛,其在多次战争中立下显赫战功,成为战场利器.此外,无人机更是凭借其体积小、灵活、用途广泛等优势,逐渐进入民用市场,在灾害检测救援、农业、地质勘探、测绘、治安反恐等领域都发挥着不可替代的作用,更加贴近生活的自拍、航拍都可以消费级无人机作为载体来完成.民用无人机火热的同时,也给空域管理带来新的挑战.目前无人机的实际监管还存在很大的难点,如何保障空域安全和有效利用有限的空间资源、防止无人机违规飞行、降低违法行为的概率成为现阶段亟待解决的问题.

无人机目标识别技术是反无人机系统的重要组成部分,包括无人机的分类和目标检测,主要是将监测系统所拍摄得到的图像进行分析并加以处理,从中正确识别出无人机目标^[1].近年来,随着以深度学习为代表的人工智能方法在图像识别领域的应用愈发成熟,大量的目标识别任务都采用卷积神经网络(Convolutional Neural Network, CNN)来完成,与传统机器学习方法相比其具有更强大的特征学习和特征表达能力^[2].文献[3]针对低空快速无人机,提出一种优化YOLOv3无人机识别模型,利用残差网络和多尺度融合的方式进行优化,结果表明该方法的平均准确率高于原始方法,但在昏暗条件下的准确率还有待验证.文献[4]针对无人机目标较小时易受干扰的问题,提出一种基于多隐含层深度神经网络的弱小无人机检测模型,完成对深度神经网络模型的训练和优化,结果表明该模型具有良好的变尺度检测能力和抗干扰效果.文献[5]提出一种基于双通道的快速低空无人机检测算法,利用红外和可见光无人机图像同时输入到残差网络中进行特征提取,利用多尺度预测网络对无人机目标进行分类和位置回归.文献[6]通过搭建深度残差网络和预测网络,提出一种基于红外图像的低空无人机检测识别方法.

以上方法基本都是通过设计多通道、多尺度融合的神经网络模型来增强网络的特征提取和抗干扰能力,而将注意力机制用于强化无人机特征提取的研究却稍显空白.传统的CNN模型在对图像进行特征提取时,往往难以有效地将图像主体特征与噪声特征区分开来,而引入注意力机制,可以捕获场景中目标对象的显著性信息,增强其特征提取能力.2014年,Google Deep Mind团队率先在循环神经网络模型上引入注意力机制完成

多个物体的高效准确识别,验证了注意力机制在图像识别领域的有效性^[7].褚晶辉等^[8]针对表情特征易受干扰的问题,通过通道和空间注意力模型对特征元素进行加强或抑制,并证明了该方法的优越性.如何将注意力机制与神经网络更为有效的结合,已经成为深度学习各领域共同关注的问题.

基于以上背景,设计了一种UAV-CNN模型对三类具有代表性的大疆无人机进行自动分类.该模型在AlexNet网络基础上进行改进,通过调整模型层数、卷积核尺寸等多项参数,以期获得最适应无人机数据集的模型,并引入批归一化和通道/空间注意力机制对模型进行优化,抑制噪声干扰,以提升模型的泛化性和特征提取能力,最终建立专门适用于无人机分类的神经网络模型.结果表明该模型在无人机数据集上具有较高的准确率和收敛速度,可实现针对大疆无人机的有效分类.

1 基本原理

1.1 卷积神经网络

卷积神经网络是深度学习中最具代表性和应用最广泛的结构之一,如图1所示.经过训练的模型通过交替连接的卷积层和池化层,将输入图像转换成一系列的特征图,最后通过全连接层完成对图像特征的分类表达.CNN强大的特征表达能力也使其在中药药材识别^[9]、医疗诊断^[10]、青铜器铭文识别^[11]、虫害识别^[12]等多方面的应用倍受到人们的关注.

1.1.1 卷积层

卷积层作用是对输入图像进行特征提取,其内部包含多个卷积核,分别与输入数据进行卷积运算,产生多个特征图,并作为下一层网络的输入.卷积核和卷积层数的选择一定程度上会影响网络的性能,需要适当的设计才能有效的避免过拟合和欠拟合.卷积运算的具体公式如式(1):

$$y_{conv} = \sum_{j=0}^{J-1} \sum_{i=0}^{I-1} x_{m+i,n+j} \omega_{ij} + b, 0 \leq m \leq M, 0 \leq n \leq N \quad (1)$$

式(1)中, x 表示接受区域为 (M, N) 的二维向量; ω 表示长和宽分别为 j, i 的卷积核; b 表示每个输出特征映射所加的偏置项; y_{conv} 表示输出的卷积结果; M 表示二维向量的长, N 为二维向量的宽; f 表示激活函数.

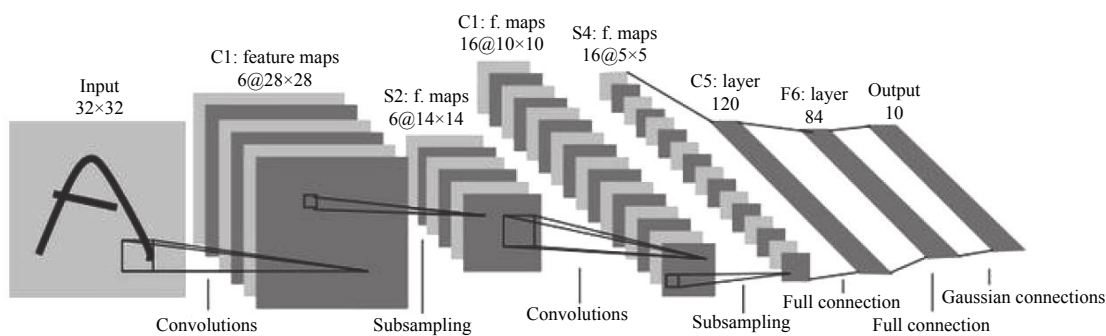


图1 卷积神经网络基本结构

1.1.2 池化层

在卷积层对特征进行提取后,输出的特征图会进入池化层进行特征选择和信息过滤,池化层不会引起额外的参数,可以在保持特征尺度不变性的情况下,大大减少网络的参数.常见的池化方式有最大池化(max pooling)和均值池化(average pooling).以最大池化为例,其公式如式(2)所示:

$$\begin{cases} f_{\text{pool}} = \max(x_{m,n}, x_{m+1,n}, x_{m,n+1}, x_{m+1,n+1}) \\ 0 \leq m \leq M, 0 \leq n \leq N \end{cases} \quad (2)$$

式(2)中, f_{pool} 表示特征图像最大池化结果.

1.2 批归一化

批归一化(Batch Normalization, BN)^[13]是一种由Ioffe和Szegedy提出的训练优化方法,其注意到中间协变量迁移,使得网络训练过程对参数初始值敏感,收敛速度变慢,把BN作为神经网络的一部分,可以使每层的输入有相同的、稳定的分布. BN简单有效,合理配置可以有效提升网络的训练速度,并且可以在一定程度上提升网络的准确率.

1.3 注意力机制算法及结构

注意力机制(Attention Module, AM)是一种模拟人脑注意力机制而衍生出的模型,通过计算注意力的概率分布,来突出某些关键部分对输出的影响. 2017年, Hu等^[14]提出基于通道间权值的分配来实现通道间重要性选择的算法 SeNet,引起了人们对于通道注意力(Channel Attention Module, CAM)的关注. 2018年, Woo等^[15]基于此提出卷积注意力模块(Convolution Block Attention Module, CBAM),其在通道注意力基础上,又加入空间注意力(Spatial Attention Module). CBAM结构兼顾通道和空间的权值选择,是一个较为轻量级的模块,可以简单的嵌入卷积神经网络模型中,

优化模型提取特征的能力.其结构如图2所示.

1.3.1 通道注意力模型

通道注意力的作用是筛选出包含重要语义信息的特征通道.输入的特征图 $H \times W \times C$ (H 表示特征图的高度, W 表示宽度, C 表示通道数)分别经过最大池化和均值池化,得到两个 $1 \times 1 \times C$ 的特征向量,将其输入一个含有隐藏层的多层感知机中,最后将分别输出的特征向量的元素对应求和,经过 Sigmoid 函数激励后得到最后的通道注意力特征向量 M_c .它可以理解为通道的权重,包含重要信息的通道会被赋予更大的权值.将通道注意力特征向量 M_c 与输入特征图对应相乘即可得到经通道注意力模块作用后的特征图.其结构如图3所示.

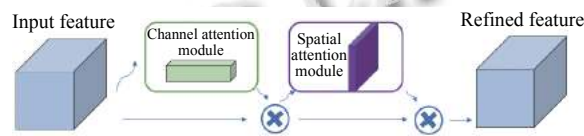


图2 卷积注意力模块结构

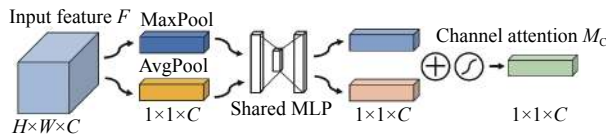


图3 通道注意力模型

1.3.2 空间注意力机制

空间注意力的作用筛选出包含重要特征信息的空间位置.其在通道维度上分别做最大池化和均值池化,然后得到两个 $H \times W \times 1$ 的特征图,并将其串联得到 $H \times W \times 2$ 的特征图.再使用一个包含 $7 \times 7 \times 2 \times 1$ 卷积核的卷积层进一步提取特征,经 Sigmoid 函数激励得到空

间注意力特征图(如图4)。空间注意力特征图可以理解为一个通道中每个像素区域的权值,包含重要信息的像素区域会被赋予更大的权重。将空间注意力特征图 M_s 与输入特征图对应相乘,即可得到经空间注意力模块作用后的特征图。

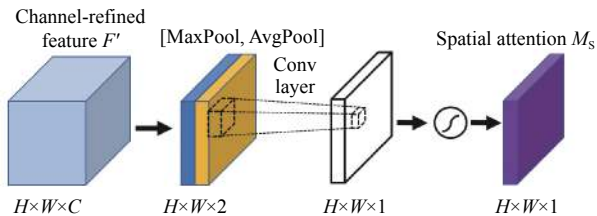


图4 空间注意力模型

2 实验与分析

2.1 数据集与实验环境

数据集采用自建的大疆无人机数据集,选取3款

典型的大疆无人机作为研究对象,分别为晓 Spark 系列,精灵 Phantom 系列,悟 Inspire 系列,如图5所示。图像数量分别为543张、550张、495张,总计1588张。选取处于复杂背景下的无人机图像,例如湖泊、丛林、岩石、荒漠、火山等,以增加了图像的多样性,更符合实际应用场景。将原图像压缩到 $227 \times 227 \times 3$ 像素,并将像素值归一化到 $[0, 1]$ 之间,以便于模型的收敛。每类无人机图片按照4:1的比例随机划分训练集和测试集,其中训练集共计1271张,用于网络训练,测试集共317张,用于测试网络性能。部分无人机数据集如图6所示。



图5 无人机型号



图6 部分无人机数据集

实验环境配置如下,硬件:CPU为Intel Core(TM) i5-8300H,内存为16 GB, GPU为NVIDIA GTX 1050Ti,软件:操作系统为Windows10 64 bit版本,编程框架为基于Python的Tensorflow框架,编译软件为Spyder。

2.2 无人机分类模型设计

由于CNN本身的数学模型还没有得到完善的数学验证和解释,学术界普遍采用以实验效果为导向的研究方式。例如VGG的深层网络结构、GoogLeNet中Inception模块以及ResNet的short connection等方法对于网络改善的有效性都是通过实验验证的。因此本文设计多组对比实验,分析不同模型结构对无人机目标分类效果的影响,并选取最优结构。

模型设计思路源于AlexNet模型,其针对大型目

标分类任务,网络结构深,模型复杂计算量大,考虑到无人机数据集样本较少和实验设备的限制,所以本文将在AlexNet模型基础上改进,探索适用于无人机数据集的模型。

实验参数设置如下,学习率设为0.0001, Batch Size为100,优化器选择Adam,损失函数选择交叉熵(cross entropy),随机失活(Dropout)率设为0.7,迭代250次,取最后10次迭代的平均值作为该模型的准确率,每个模型进行3组实验,取其平均值。实验中将综合考量模型的测试集准确率(test accuracy),训练时间(times),准确率首次到达80%的迭代数(Epoch-80%),准确率首次到达90%的迭代数(Epoch-90%)以及准确率90%以上的迭代数占总迭代数的比率(rate)的平均值,来作为衡量网络性能的参照标准。

实验首先对卷积层数进行选取,在卷积层、池化层、全连接层堆叠的模型中分别搭建卷积核为3×3的不同层数(2层、3层、4层、5层)的网络结构,并基于不同卷积层数设计的网络分别进行3组实验,各项参数如表1所示。

表1 不同卷积层层数模型的各项参数

Number of Conv layers	Test accuracy(%)	Times (min)	Epoch		Rate
			80%	90%	
Two	85.90	42	33	-	-
Three	91.17	83	28	89	45.4
Four	90.82	95	27	99	43.2
Five	87.28	324	149	-	-

分析可知设计2层卷积层时,由于模型所需训练的参数较少,训练速度快,但模型容量小,拟合能力不足,平均准确率仅85.90%。设计3层卷积层时,平均准确率达到91.17%,且Epoch-80%和Epoch-90%的值较高,说明模型收敛速度较快,准确率90%以上迭代数的占比最高,说明模型稳健性较好。设计4层和5层卷积层时,模型容量较大,参数较多,随着训练时间的增加,准确率反而呈现下降的趋势。综合考虑,后续的结构设计和优化均建立在3层卷积层的基础上。

卷积核尺寸选取是另一个重要的关乎网络性能的参数。第2组实验在3层卷积层的基础上,设立3种不同尺寸卷积核的模型。第1种按照经典AlexNet网络中卷积核由大到小的惯例,依次选择9×9、5×5、3×3的卷积核。第2种均选取5×5的卷积核,第3种均选取3×3的卷积核,分别进行3组实验,各项参数如表2所示。

表2 不同卷积核尺寸模型的各项参数

Size of Conv kernel	Test accuracy(%)	Times (min)	Epoch		Rate
			80%	90%	
9×9, 5×5, 3×3	88.16	120	92	-	-
5×5, 5×5, 5×5	89.00	85	33	79	23.6
3×3, 3×3, 3×3	91.17	83	28	89	45.4

分析可知卷积核尺寸均为3×3的模型具有最高的准确率,并且尺寸越大,模型的准确率越低,模型参数越复杂,所需的训练时间也越长。第一个卷积层选用9×9卷积核的模型,平均需要经过92次迭代才能达到80%的准确率,远低于选择小尺寸的卷积核。该结果也证实了采用较小卷积核的网络可提取更加细致、深层次的无人机图像特征,进而提高识别准确率,同时减少参数。

第3组实验对每层卷积核数量进行选取,设置5组对照实验,结果表明卷积核数量对影响网络的较为明显,卷积核数量不足将导致网络准确率低,而数量过多亦会导致准确率下降,适当的卷积核数量显得尤为重要,卷积核数量分别选择48, 96, 128时,准确率最高。各项参数如表3所示。

表3 不同卷积核数量模型的各项参数

Number of Conv kernels	Test accuracy(%)	Times (min)	Epoch		Rate
			80%	90%	
16, 32, 64	88.37	49	22	-	-
32, 64, 96	89.95	64	23	113	23.3
48, 96, 128	91.17	83	28	89	45.4
48, 96, 160	90.84	85	37	102	36.2
48, 96, 192	90.60	88	28	111	26

全连接层将池化层下采样获得的特征图映射到样本空间,将具有类别区分的局部信息组合分类,在这一层后,神经元被压平,最终形成3200个神经元。第4组实验探究全连接层层数对网络性能影响,分别设置层数为1层、2层或3层的全连接层。实验发现全连接层数设置为2层且神经元个数为1200、100的网络结构的准确率最高为91.17%,各项参数如表4所示。

第5组实验探究池化核网络的影响,典型的LeNet-5和VGG-16网络均采用k=2, s=2大小的池化核,而AlexNet网络采用k=3, s=2大小的池化核,其中k表示池化核或滑动窗口的大小,s表示滑动步长。实验的各项参数如表5所示。

表4 不同全连接层层数和神经元数目模型的各项参数

Number of FC layers and neurons	Test accuracy(%)	Times (min)	Epoch		Rate
			80%	90%	
One(300)	90.74	80	16	119	27.8
One(1200)	90.14	81	33	107	18.4
Two(500+100)	89.65	82	32	169	3.8
Two(1200+100)	91.17	83	28	89	45.4
Two(1600+100)	89.53	83	33	136	17.7
Three(2000+1200+100)	88.01	94	116	-	-

表5 不同池化方式模型的各项参数

Methods of pooling	Test accuracy(%)	Times (min)	Epoch		Rate
			80%	90%	
Max pooling (f=3, s=2)	91.17	83	27.6	89	45.4
Max pooling (f=2, s=2)	86.74	83	33	-	-
Avg pooling (f=3, s=2)	89.98	81	21	107	8.8
Avg pooling (f=2, s=2)	85.37	81	12	-	-

分析可知选用最大池化的网络的准确率比使用均值池化的网络的准确率高1.2%左右,最大池化选取该

区域内最大的值,可以更好的保留图像的纹理特征,均值池化是对区域内所有不为零的像素值取平均,以此得到数据特征,该方法更好地保留了对图像背景信息的提取.而不同池化核的影响更大,经实验证明不论是采用最大池化还是均值池化,采用 3×3 池化核的网络比 2×2 卷积核的网络准确率高4~5个百分点,因此选用滑动窗口大小大于滑动步长产生的重叠池化,可以有效的提升网络的分类能力.

综合上述的实验结果,设计一个10层的CNN模型,其中8层为隐含层,包括3层卷积层,3层池化层,2层全连接层.CNN模型结构和参数如表6所示.

表6 CNN模型结构和参数

Layer	Kernel size	Strides	Output size	Output number
Input			227×227	3
Convolution1	3×3	2	113×113	48
Max pooling1	3×3	2	56×56	48
Convolution2	3×3	2	27×27	96
Max pooling2	3×3	2	13×13	96
Convolution3	3×3	1	11×11	128
Max Pooling3	3×3	2	5×5	128
Fully connected1			1×1	1200
Fully connected2			1×1	100
Output				3

2.3 无人机分类模型优化

2.3.1 BN对模型优化的影响

在AlexNet网络中使用了局部响应归一化层(LRN)来增加网络的泛化能力,但通过数据实际测试后发现对网络性能并没有太大的改善,反而增加了训练时间.而BN作为一种简单有效的网络优化方法,可以对每一层的输入数据做归一化处理,稳定其数据分布的同时,加速网络的训练,并允许梯度迭代的过程中使用更大的学习率,且不会引起梯度更新的发散.BN通过引入两个可以和神经网络训练参数一起训练的参数,很好的保留了神经网络的原始表达能力.因此引入BN层对模型进行优化,加入BN层后模型如图7所示.

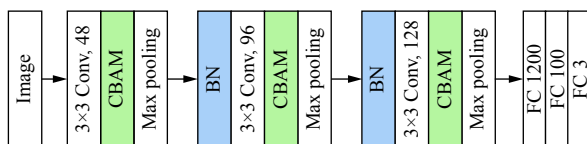


图7 优化后的模型

2.3.2 注意力机制对网络优化的影响

由于无人机本身置于复杂环境中,与背景之间区

分度较小以及部分图像从视频中截取而来像素较低,存在较大的噪声干扰,二者都会影响网络分类的效果.同时在网络学习的过程中,随着网络层数的增加,这些噪声像素在特征图中的权值也会逐步增加,这将会对最终的分类结果有较大的影响,因此如何弱化噪声干扰给网络训练带来负面影响,是网络优化任务重点解决的问题.

对于众多特征,并不是每一个特征层都能提取到有用的信息,而通道注意力的引入可以通过再学习的方式允许网络进行特征信息的重新校准,使模型对包含重要特征信息的特征层赋予更大的权值,同时减小无关信息特征层的权值,帮助模型获取表达能力更强的语义信息.空间注意力可以捕捉图像中对于某一特定任务贡献显著的某些区域,忽略带来负面效应的区域,有效的增强模型的拟合能力,提升分类效果.

因此在CNN模型中引入卷积注意力模型,经初步实验分析将通道和空间注意力模型设置在卷积层和池化层中间可以取得最优的效果,引入通道和空间注意力后的模型如图7所示.

2.3.3 优化结果分析

首先设置CNN模型与引入CBAM的CNN模型的对比实验.为获得更为合理、准确的比较结果,将迭代次数由250增加至1000次,使得模型可以充分训练,通道注意力模型的压缩比设置为8,其余实验参数与前文所述一致.由于模型的准确率存在一定范围内的波动,因此取后100次迭代的准确率的平均值作为模型最终的准确率.

本部分实验首先分析引入CBAM对模型的影响.分别对先前设计的普通的CNN模型与基于CBAM的CNN模型,应用无人机数据集进行训练,然后将两者的识别准确率进行比较.每次10次迭代记录其测试样本的准确率,然后将记录好的数据绘制成图,即图8所示的测试准确率(test accuracy)与迭代次数(epoch)关系图.

曲线A表示未经CBAM优化的CNN模型的识别准确率曲线,曲线B表示经CBAM优化后模型的识别准确率曲线.由图8可以看出本文设计的模型收敛较快,在100次迭代前后就可以达到90%的识别准确率,200次迭代已经可以基本收敛.说明本文设计的模型训练效果较好,能够快速达到较为理想的训练状态.经CBAM优化后的模型具有更好的平均准确率,可以达

到 92.13%, 较优化前提升 1.19%。本文认为注意力机制的引入较好的增强了特征的判别性, 起到了特征选择的作用, 提升了分类效果。

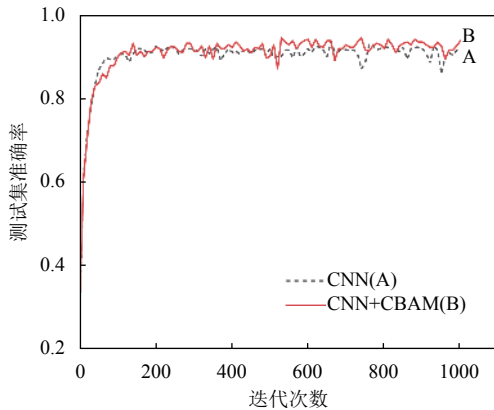


图8 测试准确率与迭代次数关系图

实验的第二部分在引入 CBAM 的基础上再引入批归一化层 (BN) 对模型进行改进。实验结果如图 9 所示。曲线 B 表示仅 CBAM 优化的模型的识别准确率曲线, 曲线 C 表示经 CBAM 和 BN 共同优化后的模型的识别准确率曲线。

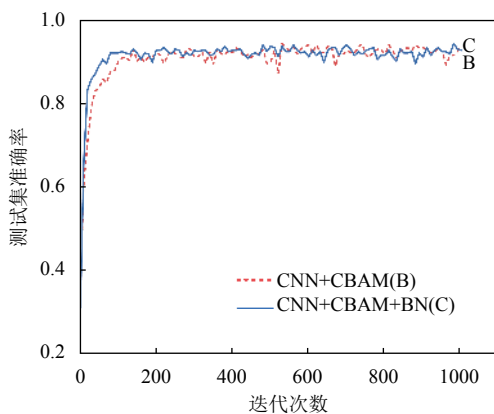


图9 测试准确率与迭代次数关系图

由图 9 可以看出引入 BN 层后模型的收敛速度明显加快, 在获得较高的准确率的基础上, 模型震荡水平更小, 说明模型更加稳定。由表 7 可以看出经 CBAM 和 BN 共同优化后的模型的准确率可以达到 92.44%, 较仅引入 CBAM 的模型又提升 0.31%。

2.3.4 与其他卷积神经网络模型对比分析

采用不同的卷积神经网络模型对无人机数据集进行识别, 实验结果如表 7 所示, 识别准确率曲线对比如图 10 所示, 可以看出本文提出的基于 CBAM 和 BN

的模型的最终准确率可以达到 92.44%, 较 LeNet-5 和 AlexNet 模型分别提升了 11.82% 和 20.78%, 且收敛速度更快, 参数量较经典的 AlexNet 模型大大减少。证明本文设计的模型更加适用于无人机目标的识别, 可以满足基本的大疆无人机图像的识别需求。

表7 不同模型在无人机数据集上各项参数对比

模型	测试准确率(%)	时长(h)	参数(MB)
CNN	90.94	6.5	≈4.1
CNN+CBAM	92.13	7	≈4.1
CNN+CBAM+BN	92.44	6.7	≈4.1
LeNet-5	80.62	0.2	≈0.1
AlexNet	71.66	12.5	≈62

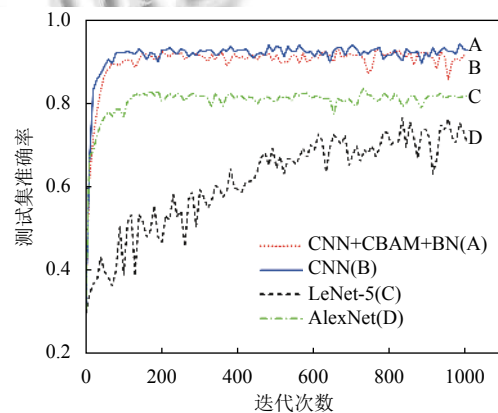


图10 不同模型在无人机数据集上识别准确率对比图

3 结论

本文设计了一个基于深度卷积神经网络的无人机目标分类模型, 可以完成多款大疆无人机的分类任务。通过多组对比实验选择最适用于无人机数据的模型参数, 然后引入卷积注意力机制对网络进行优化, 加强模型的特征提取能力, 同时引入批归一化层, 加速网络的训练, 提升模型稳健性。实验结果表明本文提出的方法适用于无人机数据集的目标分类, 较 LeNet-5 和 AlexNet 模型分别提升了 11.82% 和 20.78%, 可以满足实际场景中无人机目标分类的要求。但本文的数据集仅选择了 3 种型号的无人机, 数量较少。后续将研究更深层次的模型上的多类型目标分类任务。

参考文献

- 于鹏. 基于无人机图像的识别算法研究 [硕士学位论文]. 哈尔滨: 哈尔滨工程大学, 2018.
- 卢宏涛, 张秦川. 深度卷积神经网络在计算机视觉中的应用研究综述. 数据采集与处理, 2016, 31(1): 1-17.

- 3 马旗, 朱斌, 张宏伟, 等. 基于优化 YOLOv3 的低空无人机检测识别方法. 激光与光电子学进展, 2019, 56(20): 201006.
- 4 王靖宇, 王霁禹, 张科, 等. 基于深度神经网络的低空弱小无人机目标检测研究. 西北工业大学学报, 2018, 36(2): 258–263. [doi: 10.3969/j.issn.1000-2758.2018.02.008]
- 5 马旗, 朱斌, 程正东, 等. 基于双通道的快速低空无人机检测识别方法. 光学学报, 2019, 39(12): 1210002.
- 6 马旗, 孙晓军, 张杨, 等. 基于红外图像的低空无人机检测识别方法. 弹箭与制导学报, 2020, 40(3): 150–154.
- 7 朱张莉, 饶元, 吴渊, 等. 注意力机制在深度学习中的研究进展. 中文信息学报, 2019, 33(6): 1–11. [doi: 10.3969/j.issn.1003-0077.2019.06.001]
- 8 褚晶辉, 汤文豪, 张姗, 等. 一种基于注意力模型的面部表情识别算法. 激光与光电子学进展, 2020, 57(12): 121015.
- 9 王一丁, 郝晨宇, 李耀利, 等. 基于深度学习的小样本中药材粉末显微图像识别. 计算机应用, 2020, 40(5): 1301–1308.
- 10 车翔玖, 董有政. 基于多尺度信息融合的图像识别改进算法. 吉林大学学报(工学版), 2020, 50(5): 1747–1754.
- 11 李文英, 曹斌, 曹春水, 等. 一种基于深度学习的青铜器铭文识别方法. 自动化学报, 2018, 44(11): 2023–2030.
- 12 赵立新, 侯发东, 吕正超, 等. 基于迁移学习的棉花叶部病虫害图像识别. 农业工程学报, 2020, 36(7): 184–191. [doi: 10.11975/j.issn.1002-6819.2020.07.021]
- 13 Ioffe S, Szegedy C. Batch Normalization: Accelerating deep network training by reducing internal covariate shift. Proceedings of the 32nd International Conference on Machine Learning. Lille, France. 2015. 448–456.
- 14 Hu J, Shen L, Sun G. Squeeze-and-excitation networks. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. 2018. 7132–7141.
- 15 Woo S, Park J, Lee JY, *et al.* CBAM: Convolutional block attention module. Proceeding of the 15th European Conference on Computer Vision. Munich, Germany. 2018. 3–19.