第47卷 第7期 计 算 机 工 程 2021年7月
Vol.47 No.7 Computer Engineering July 2021

·图形图像处理· 文章编号：1000-3428(2021)07-0273-08 文献标志码：A 中图分类号：TP391.41

基于轻量级卷积神经网络的人脸检测算法

朱灵灵[1]，高 超[2]，陈福才[2]

（1.郑州大学 中原网络安全研究院，郑州 450000；2.中国人民解放军战略支援部队信息工程大学，郑州 450000）

摘 要：针对人脸检测在移动端应用时面临的移动设备计算能力及存储资源受限等问题，设计一种基于轻量级卷积神经网络的改进人脸检测算法Lightweight-SSH。基于单点无头人脸检测器(SSH)人脸检测算法，采用基于MobileNet的轻量级卷积神经网络对样本数据进行特征提取，减少模型的参数量和计算量，通过在SSH网络的检测模块中引入可变形卷积层，提升卷积神经网络对人脸形变的建模能力。在Wider Face数据集上的实验结果表明，与常用人脸检测算法相比，Lightweight-SSH算法在保证检测精度的前提下，明显降低模型复杂度，并提高了模型检测速度。

关键词：人脸检测；可变形卷积；MobileNet网络；单点无头人脸检测器；Lightweight-SSH算法

开放科学（资源服务）标志码（OSID）：

中文引用格式：朱灵灵，高超，陈福才.基于轻量级卷积神经网络的人脸检测算法[J].计算机工程，2021，47(7)：273-280.

英文引用格式：ZHU L L，GAO C，CHEN F C.Face detection algorithm based on lightweight convolutional neural network[J].Computer Engineering，2021，47(7)：273-280.

Face Detection Algorithm Based on Lightweight Convolutional Neural Network

ZHU Lingling[1]，GAO Chao[2]，CHEN Fucai[2]

（1.Zhongyuan Network Security Research Institute，Zhengzhou University，Zhengzhou 450000，China；
2.PLA Strategic Support Force Information Engineering University，Zhengzhou 450000，China）

【Abstract】When deployed to mobile terminals，face detection applications are usually limited by the computing power and storage resources of the devices.To address the problem，an improved face detection algorithm called Lightweight-SSH using lightweight convolutional neural network is proposed. This face detection algorithm is designed based on the Single Stage Headless Face Detector（SSH） algorithm，and employs MobileNet-based lightweight convolutional neural network to extract the features of sample data，reducing the number of parameters and the amount of computation.A deformable convolution layer is introduced into the detection module of the SSH network to improve the ability of the convolutional neural network to model human face deformation.Experimental results on the Wider Face dataset show that，compared with commonly used face detection algorithms，the Lightweight-SSH algorithm significantly reduces model complexity and improves model detection speed while ensuring detection accuracy.

【Key words】face detection；deformable convolution；MobileNet；Single Stage Headless Face Detector（SSH）；Lightweight-SSH algorithm

DOI：10. 19678/j. issn. 1000-3428. 0057249

0 概述

人脸检测算法是计算机视觉信息处理的经典问题，经过多年的研究已经比较成熟，但随着移动设备的快速普及，人脸检测算法需要应用在手持设备、车载设备等移动终端上。然而，这些终端设备在计算能力、存储能力等方面相对服务器和个人计算机设备局限性很强。为使基于深度网络的人脸检测算法能够较好地运行在移动终端设备上，基于轻量级卷积神经网络的算法得到人们越来越多的关注。

国内外学者已经对人脸检测算法进行了深入研究，提出了诸多经典方法。文献[1]提出用于人脸检测任务的多任务神经网络模型，其主要采用3个级联网络，运用候选框加分类器的思想进行快速高效的人脸检测；文献[2]提出的Faster RCNN将特征抽取、候选框提取、边界框回归、分类整合在一个网络

基金项目：国家自然科学基金（61601513）。

作者简介：朱灵灵（1995—），女，硕士研究生，主研方向为计算机视觉、图像处理；高 超，助理研究员；陈福才，研究员。

收稿日期：2020-01-16 **修回日期：**2020-05-29 **E-mail：**chaosndsc@163.com

计算机工程 www.ecice06.com

中,使得综合性能有较大提高;文献[3]提出的SSD是一种基于单步候选区回归的目标检测算法。

在2012年举行的ImageNet大赛中,文献[4]将卷积神经网络AlexNet应用在图像分类中并取得冠军,之后,基于深度学习的检测方法得到快速发展;文献[5]提出的GoogLeNet网络通过增大网络规模,包括深度和宽度,提高网络性能;2014年,文献[6]提出的VGG主要采用小尺寸的卷积核,被广泛应用于视觉领域的各类任务中;文献[7]提出的ResNet网络将网络的层数推广到了前所未有的规模。这些卷积网络主要是通过加深网络层数提高模型准确率,但增加了网络的整体复杂度,使得网络难以优化,容易过拟合。同时,现有的多数深度学习模型常会有较大的参数量,可以达到数百甚至上千兆,而移动端的计算能力、存储资源较弱,很难将较大的深度学习模型部署在移动端上。因此,研究一种轻量级人脸检测技术具有很好的工程应用价值和研究意义。2016年,文献[8]提出Inception V3,该方法将1个二维卷积拆分为2个相对较小的卷积,如将7×7卷积拆分成1×7和7×1卷积,从而可以降低参数量;同年,文献[9]提出的SqueezeNet是采用1×1卷积核代替3×3卷积核,有效减少了输出特征图的通道数。文献[10]提出轻量级卷积网络MobileNet,在尽量保证检测精度相当的基础上,减少该网络的参数量和计算量。

传统的卷积神经网络在采样过程中由于采用方块型的卷积核,其几何变换建模的能力受到限制,模型的检测性能受到影响。本文将文献[11]中可变形卷积层应用到卷积神经网络中,提出一种基于轻量级卷积神经网络的改进人脸检测算法Lightweight-SSH。采用MobileNet轻量级卷积神经网络对样本数据进行特征提取,并通过改进后的单点无头人脸检测器(Single Stage Headless Face Detector,SSH)[12]对样本数据中的人脸进行检测。

1 相关工作

基于深度学习的人脸检测算法主要由卷积神经网络特征提取器和目标检测器2个部分组成,如图1所示。卷积神经网络特征提取器对加载后的数据集进行特征提取,得到特征图后采用目标检测器对其进行分类和回归[13]。本文主要从上述2个方面对SSH检测算法进行分析并改进:1)采用MobileNet轻量级卷积神经网络代替VGG-16网络来提取特征,有效降低模型复杂度;2)在SSH检测算法的检测模块部分加入可变形卷积的方法,有效提升模型对提取到的特征的处理能力,提高模型检测准确率。

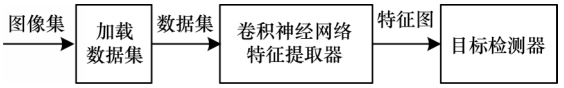


图1 人脸检测算法系统结构

Fig.1 System structure of face detection algorithm

1.1 SSH人脸检测算法

检测小尺寸人脸仍然是一项具有挑战性的任务,在通常情况下检测性能较好的人脸检测器常常速度较慢,参数量较大。文献[11]等提出的SSH检测器是单阶段检测,与两阶段最大的不同是直接同时进行分类和anchor的回归;SSH骨干网络采用的是去掉3个全连接层的VGG-16网络,这些属性使SSH快速且轻量级。但需要注意的是,SSH网络的检测速度较快是相对的,且依赖于GPU,VGG-16网络含有的参数量太大。另外,与其他常用的检测算法不同的是,SSH网络不是依靠输入图像金字塔来检测具有不同尺寸人脸的图像,而是在不同特征层上进行检测来保持尺度的不变。SSH的目的是设计一个减少推理时间、占用内存低和尺度不变性的网络结构,并可以有效地合并上下文,增加感受野,同时对卷积层提取的信息进行分类定位。

1.2 MobileNet轻量级网络

MobileNet的核心是使用一种深度可分离卷积作为其基本单元来代替原来的标准卷积,经过这样的操作可以达到近似传统卷积的效果,同时可以显著降低网络的参数量和计算量,减少卷积核的冗余表达。在MobileNet的网络结构中,所有层后都加上BN和非线性激活函数ReLU,除最后的连接层没有非线性激活函数,其直接将结果传送到softmax层中进行分类。

SSH算法采用VGG-16网络对样本进行特征提取,即包含16个卷积层和全连接层。采用VGG-16简化网络结构参数量较大,在ImageNet数据集上,共包含了1.38亿个参数,导致模型特征参数量太大,影响检测速度。针对这一现象,本文采用MobileNet神经网络作为基础网络对样本数据进行特征提取,有效降低模型复杂度。

1.3 可变形卷积层

可变形卷积与可变形部件模型[14](Deformable Part Model,DPM)不同,传统卷积层采用规则形状的卷积核,一般采用方块形的卷积核对样本进行采样,若采用二维网格图表示,则如图2(a)所示,几何变换能力受到限制。而可变形卷积针对这一限制,在原有卷积核上加入一个偏移的向量,如图2(b)~图2(d)所示,该偏移量是通过网络结构的一个平行标准卷积单元计算得到,进而可以通过梯度反向传播进行端到端的学习[14]。加上偏移量的学习后,可以根据当前需要检测的图像内容进行动态调整卷积核的大小和位置,其直观效果就是不同位置的卷积

www.ecaeb.com

核采样点位置会根据图像内容发生变化从而适应不同物体的形状、大小等几何形变。增强模型对于非刚性物体变换的建模能力，并提高对物体的检测效果。

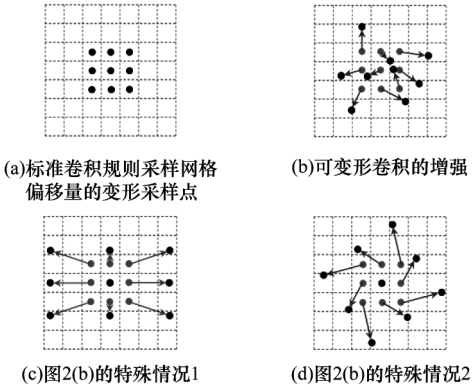


图2　卷积核采样点的位置分布

Fig.2　Location distribution of convolution kernel sampling points

图3所示为可变形卷积网络结构。偏移量通过一个卷积层学习得到，该卷积层的卷积核与普通卷积核一样，输出的偏移尺寸和输入的特征图尺寸一致。偏置区域是变形卷积外加的待训练参数，大小和输入层图片一致，卷积窗口在偏置区域上滑动呈现卷积像素偏移的效果。传统卷积网络少了图3中虚线框的部分。

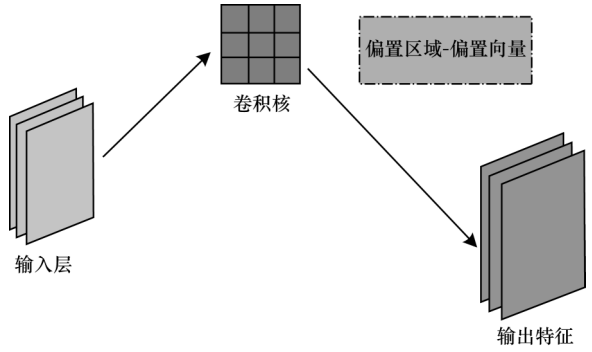


图3　可变形卷积网络结构

Fig.3　Structure of deformable convolutional network

2　Lightweight-SSH人脸检测算法

人脸检测算法SSH最大的特色就是尺度不相关性。MTCNN[1]网络在预测时，主要是通过多次输入不同尺寸大小的图片并对其进行预测再通过非极大值抑制（NMS）；而SSH只需要1遍传播就可以完成预测。SSH的实现途径主要就是对VGG不同的卷积输出层做了3个分支（M1，M2，M3），每个分支都是用类似的流程预测检测和分类结果。SSH的速度较快、占用内存低的特性是相对的，并不能很好地应用在移动端上。本文主要从2个方面对SSH进行改进并设计Lightweight-SSH网络结构，该模型结构如图4所示。

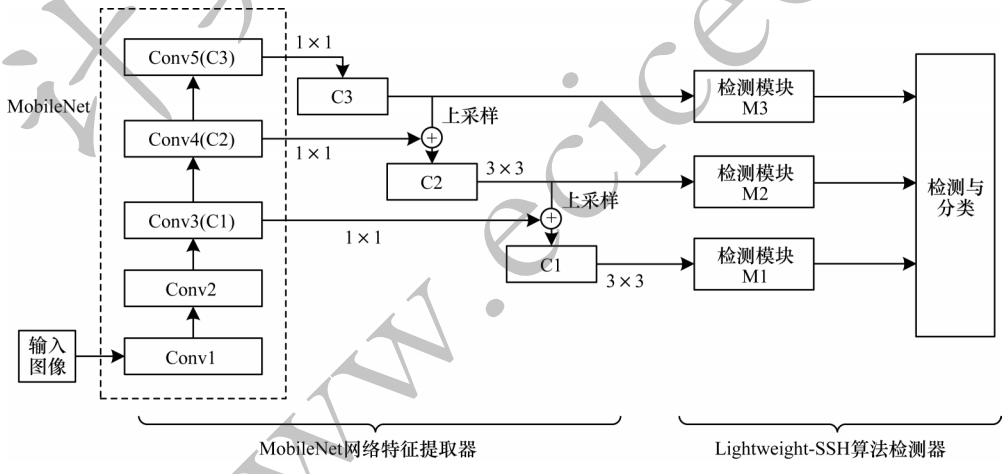


图4　Lightweight-SSH模型结构

Fig.4　Structure of Lightweight-SSH model

2.1　MobileNet网络结构的改进

传统卷积神经网络虽然有良好的特征提取能力，但是模型复杂度较高、计算量较大，无法很好地适用于硬件性能有限的设备[13]。SSH算法采用VGG-16对样本数据进行特征提取，但VGG-16的参数量较大，影响检测速度。在ImageNet数据集上，VGG-16包含1.38亿个参数，1.0MobileNet-224包含4.2×10^6个参数，而两者准确率相当。采用MobileNet轻量级网络替代VGG-16来对样本数据进行特征提取，可以在保证检测精度相当的同时显著降低模型的计算复杂度。

2.1.1　深度可分离卷积与通道缩减

MobileNet的核心是使用深度可分离卷积作为其基本单元来代替原来的标准卷积[15-16]。图5所示为标准卷积、深度卷积和1×1逐点卷积示例，使用深度卷积核与逐点卷积核代替标准卷积核，可以达到与标准卷积相同的卷积效果。假定输入特征图F大小是$D_F\times D_F\times M$，而输出特征图G大小是$D_F\times D_F\times N$，

假定输入、输出特征图的宽高相同皆为D_F，M、N分别表示输入、输出的通道数。设定$D_K \times D_K$为卷积层中卷积核K的大小[17]。

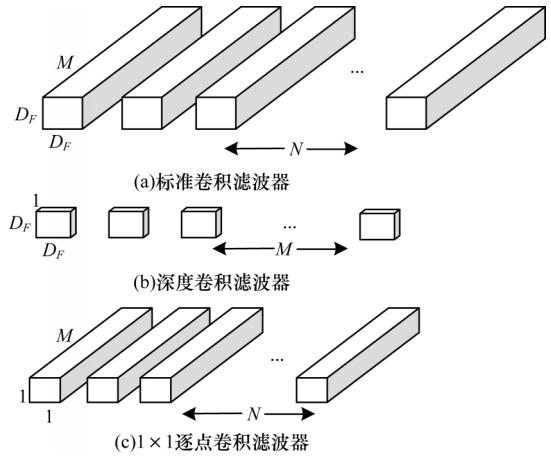


图5 标准卷积与深度可分离卷积

Fig.5 Standard convolution and depth separate convolution

对于标准卷积而言，假设卷积核的步长为1，其卷积过程如式(1)所示，经过相应计算可以得到输出特征图：

$$G_{k,l,n}=\sum_{i,j,m}K_{i,j,m,n}\times F_{k+i-1,l+j-1,m} \tag{1}$$

计算量为：$D_K \times D_K \times M \times N \times D_F \times D_F$。

对于深度可分离卷积而言，深度卷积操作如式(2)所示：

$$G_{k,l,m}=\sum_{i,j}K_{i,j,m}\times F_{k+i-1,l+j-1,m} \tag{2}$$

其中：K的尺寸为$D_K \times D_K \times M$，计算量为$D_K \times D_K \times M \times D_F \times D_F$；1×1的逐点卷积的计算量大小为$M \times N \times D_F \times D_F$。总的计算量为两者相加，即$D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F$。

通过式(3)可以得到将标准卷积操作分解为2个步骤后减少的计算量，即两者计算总量的比值：

$$\frac{D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F}{D_K \times D_K \times M \times N \times D_F \times D_F}=\frac{1}{N}+\frac{1}{D_K^2} \tag{3}$$

MobileNet主要使用尺寸大小为3×3的深度可分离卷积，即D_K=3，通过式(3)的结果可以看出，采用该操作可以将计算量减少8~9倍，提高了计算的效率。

引入宽度因子α可以控制模型大小，缩减通道，降低模型参数量，如式(4)所示：

$$\frac{D_K^* D_K \times \alpha M \times D_F \times D_F + \alpha M \times \alpha N \times D_F \times D_F}{D_K \times D_K \times M \times N \times D_F \times D_F}=\frac{\alpha}{N}+\frac{\alpha}{D_K^2} \tag{4}$$

令α=0.25，引入宽度因子α后，计算量约降为原来的1/16。

2.1.2 多级特征融合

在卷积神经网络提取图像特征的过程中，网络层数越深，特征图包含的语义信息越多，更有利于检测大尺寸人脸；网络层数越低，特征图包含的空间信息越多，更有利于检测小尺寸人脸。Lightweight-SSH是以MobileNet作为特征提取的骨干网络，为更好地检测不同尺寸的人脸，利用SSH算法的尺度不变性，即通过不同的检测层对不同尺寸的人脸进行分类和预测检测，在该网络的不同卷积输出层做3个分支(M1、M2、M3)，分别对小、中、大尺寸的人脸进行检测。如图4所示，首先采用自底向上的线路选取Conv3、Conv4、Conv5 3个层的特征图作为检测层，记为C1、C2、C3，然后使用自顶向下和横向连接分别计算3个层的输出。自底向上即为网络的前向过程，在此过程中，经过特定卷积层后特征图的大小会改变，而经过其他一些层时不会改变，将同等特征图大小的层归为一个阶段，每次抽取每个阶段的最后一层输出。自顶向下的过程进行上采样从而使其结果和自底向上生成的特征图大小相同，而横向连接则是将两者进行融合，采用1×1卷积层输入进行降维，从而可以降低本模型的大小。为消除上采样的混叠效应，在融合之后通过3×3的卷积核对上一步的融合结果进行卷积。将计算得到的输出特征图分别送入检测模块M1、M2、M3。

2.2 Lightweight-SSH算法检测

在Lightweight-SSH算法中主要采用检测模块(M1、M2、M3)对特征提取部分得到的特征图进行处理，检测模块结构如图6所示，2个1×1卷积核分别对应输出分类和回归结果，该模块中包含1个简单的上下文模块来增加感受野的有效性。

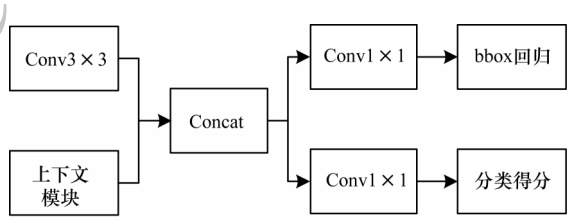


图6 检测模块结构

Fig.6 Structure of detection module

在卷积神经网络中，感受野是指卷积神经网络中每一层输出特征图上的单个像素点所对应的输入图像的映射区域大小[18]。而SSH检测算法中的上下文模块主要用来增大感受野，提高准确率，即通过使用串联的卷积层代替较大尺寸的卷积层从而增加卷积层的感受野同时可以减少参数量，并将2个3×3卷积层和3个3×3卷积层并联的结构作为检测模块的目标尺寸。在此基础上，将原来SSH结构的上下文模块中的3×3卷积层替换为可变形卷积层，即可以

计算机工程 www.ecice06.com

针对不同形状的物体采用不同的几何形状进行采样，提高了对不规则物体的形变建模能力。

可变形卷积核相较于传统卷积核，其主要是引入一个偏移量$\{\Delta p_n | n=1,2,\cdots,N\}$，其中，$N=|R|$。每一个位置$p_0$的输出结果为：

$$y(p_0)=\sum_{p_n} w(p_n)x(p_0+p_n+\Delta p_n) \tag{5}$$

其中：p_n表示在卷积窗口中任意一个像素点；$w(p_n)$表示像素点p_n的权重；x表示输入层像素点的集合；Δp_n表示像素点p_n的偏移量。

具体的上下文模块结构如图7所示，DCN表示可变形卷积层。采用传统卷积层的模型主要是由数据本身所具有的多样性来提高对物体几何形变的适应能力，即该模型不能很好地适应目标物体的形变，从而不能较好地对样本进行采样。而可变形卷积层通过一个平行的标准卷积单元计算得到偏移量后，可以根据当前需要检测的目标物体进行动态调整卷积核的大小和位置，其直观效果即卷积核采样点位置会根据目标物体发生自适应变化，从而适应不同物体的形状、大小等变化。

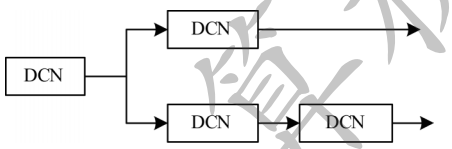


图7　改进的上下文模块结构

Fig.7　Improved context module structure

通过对SSH检测算法进行改进，本文提出的Lightweight-SSH检测模型在检测速度、检测精度、模型大小等性能上都有所提升，从而可以更好地应用在人脸检测的移动端设备。

2.3　检测模型目标函数

对于任何训练的锚点框i，通过训练最小化多任务损失函数，其包含分类任务和边框回归任务，如式(6)所示：

$$L=L_{\text{cls}}(p_i,p_i^*)+p_i^*L_{\text{box}}(t_i,t_i^*) \tag{6}$$

其中：$L_{\text{cls}}(p_i,p_i^*)$为人脸分类的损失函数；$p_i$是预测锚点$i$为人脸的概率；$p_i^*$为1表示正锚点，为0表示负锚点；$L_{\text{box}}(t_i,t_i^*)$为人脸框的回归损失函数，$t_i=\{t_x,t_y,t_w,t_h\}$和$t_i^*=\{t_x^*,t_y^*,t_w^*,t_h^*\}$分别表示正锚点的预测框和真实框的坐标。

人脸分类的损失函数L_{cls}是二分类的softmax损失，具体表示如式(7)所示：

$$L_{\text{cls}}(p_i,p_i^*)=-\log_a p_i \tag{7}$$

人脸框的回归损失函数L_{box}采用L_1损失函数，如式(8)所示：

$$L_{\text{box}}(t_i,t_i^*)=\sum_{i\in\{x,y,w,h\}}\text{smooth}_{L_1}(t_i^*-t_i) \tag{8}$$

其中：

$$\text{smooth}_{L_1}(x)=\begin{cases}0.5x^2, & |x|<1\\ |x|-0.5, & \text{其他}\end{cases} \tag{9}$$

3　实验结果与分析

3.1　实验数据集

本文采用Wider Face[19]数据集对人脸检测定位任务和人脸分类任务进行训练，该数据集共有32 203张人脸图片，并标记393 703个具有高度变化性的人脸，其中包含遮挡、化妆、角度、不同尺度、姿势等情况下变化的图片，如图8所示。在本文实验中，将该数据集的人脸图片随机划分，40%的数据作为训练集，10%的数据作为验证集，50%的数据作为测试集，分别包含约12 881、3 220、16 102张人脸图片。同时，训练集、验证集、测试集标注的人脸数量分别约为84 600、18 500、290 600个。



图8　Wider Face中不同尺寸、姿势、遮挡、化妆等情况下变换的示例

Fig.8　Examples of transformtions in Wider Face different Scale, Pose, Occlusion, Makeup, etc.

3.2　评价指标

本文实验结果评价指标采用精确率(P)与召回率(R)。

精确率是针对预测结果而言，表示被正确预测的人脸数量占样本中所有预测为人脸的比例。精确率计算公式如下：

$$P=\frac{T_{\text{TP}}}{T_{\text{TP}}+F_{\text{FP}}} \tag{10}$$

其中：TP表示被正确预测的人脸数量；FP表示被错误预测为人脸的数量。

召回率是针对原本的样本而言，表示样本中被正确预测的人脸数量占样本中总人脸数的比例。召回率计算公式如下：

$$R=\frac{T_{\text{TP}}}{T_{\text{TP}}+F_{\text{FN}}} \tag{11}$$

其中：FN表示人脸被错误预测为背景的数量。

3.3　实验设置

交并比(IoU)一般指目标预测框和真实框的交集和并集的比例。在训练过程中，通过设置IoU的阈值划分正负样本。在训练时，当设置IoU大于阈

计算机工程 www.ecice.com

值0.5时，锚点框匹配到正样本中；当IoU小于0.3时，则匹配到负样本中；未匹配的锚点框在训练中被忽略。锚点框总数约为102 300个，而由于大多数锚点框（约大于99%，约有101 277个）被匹配为负样本，本文采用标准的困难样本挖掘（OHEM）来缓解正、负训练样本之间的显著不平衡。根据损失值对负锚框进行排序，并选择损失最大的anchors，这样负样本和正样本之间的比例至少为3∶1，最终被忽略掉的锚点框约为98 208个。

对于Wider Face数据集的测试，采用Box voting即边框投票机制，并设置IoU阈值为0.4。在检测模型检测目标对象时，每个目标附近会产生多个边界框，将这些边界框的得分进行排序，选中最高分及其对应的框；遍历其余边界框，选取和当前最高分边界框的重叠面积大于设定的阈值（0.4）的边界框；最后取这些边界框的加权平均值作为最终预测的边界框位置，即为预测的人脸框位置。

实验采用分段训练的策略，在模型训练过程中，按照设定的训练节点调整学习率的大小，减少模型训练过程中的损失。其中，初始学习率设置为0.001，在5个轮次后上升为0.01，之后在第55个和第68个轮次时，学习率分别除以10。学习率的设置采用了warm up预热学习率的方式，由于在刚开始训练时，模型的权重是随机初始化的，此时若选择一个较大的学习率，可能带来模型的不稳定，选择warm up的方式，可以使得开始训练的5个批次内学习率较小，在预热的小学习率下，模型可以趋于稳定，等模型相对稳定后可以再选择预先设置的学习率（如0.01）进行训练，使得模型收敛速度变得更快，模型效果更佳。动量参数设置为0.9，每个训练批次随机读取32张图片。在训练过程中采用的是随机梯度下降（SGD）算法优化整个网络模型。

预训练模型采用的是Gluon Model Zoo的标准版MobileNet0.25预训练模型。本文的所有实验都是在Ubuntu 16.04.5 LTS系统下利用MXNet框架完成的，系统的硬件环境为TITAN XP处理器。

3.4 实验结果

为进一步验证模型的性能，将本文算法Lightweight-SSH与近年来的主流算法SSH、多任务级联卷积神经网络（Multitask Cascade Convolutional Neural Network，MTCNN）和多任务深度卷积网络（Multi-Scale Deep Convolutional Neural Network，MSCNN[20]）在相同的实验环境和相同的数据集上进行对比，在Wider Face人脸验证集的Easy、Medium、Hard 3个子集上得到的P-R曲线如图9所示。其中3个子集根据边缘框的检测情况划分为简单、中等、困难3个难度等级。

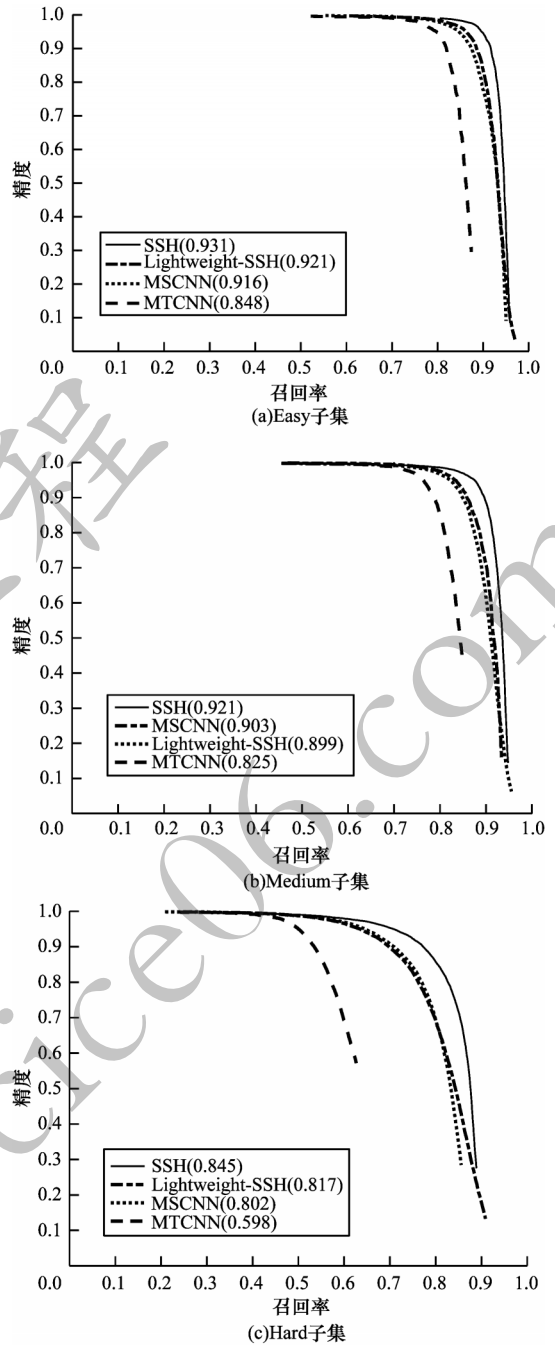


图9 Wider Face人脸验证集测试结果

Fig.9 Test results of the Wider Face verification set

从图9可以看出，Lightweight-SSH算法较MSCNN、MTCNN算法在检测精度上占有优势；而该算法在3个子集上与SSH算法的性能相当，但Lightweight-SSH算法能更好地部署在移动端。在Hard子集上能够达到0.82的检测精度，表明该算法检测小尺寸人脸具有良好的效果。

将Lightweight-SSH算法与当前常用的人脸检测算法在Wider Face上进行性能比较。具体的比较结果如表1所示，其中粗体为最优值。

www.ecice06.com

表1　不同算法在Wider Face数据集上的性能对比

Table 1　Performance comparison of different algorithms on Wider Face datasets

算法	测试精度	浮点运算量/(Operations·s^{-1})	模型大小/MB	检测时间/ms
MSCNN	0.802	326.400	28.45	60
MTCNN	0.598	267.200	25.70	45
VGG16-SSH	**0.845**	312.300	15.65	74
MobileNet-SSH	0.817	**18.645**	**1.46**	**23**
Lightweight-SSH	0.825	28.736	1.69	25

从表1可以看出，相比MSCNN、MTCNN人脸检测算法，Lightweight-SSH人脸检测算法在检测精度、浮点运算数、模型大小、检测时间各个性能中都有大幅优化，占有明显优势。同时，将SSH算法中的骨架网络VGG16网络替换为MobileNet网络后，检测精度稍有差距，但浮点运算数约降为原来算法的1/15，大大减小了模型大小，并提高了单幅图片检测速度。Lightweight-SSH算法在MobileNet-SSH算法基础上模型大小虽稍有增加，但提高了检测精度，且满足移动设备的条件。

通过上述比较实验结果可知，采用MobileNet轻量级卷积网络作为基础网络来提取特征，有效降低了模型参数量和计算量，减小了模型大小；将卷积神经网络中的相应传统卷积层替换为可变形卷积层，有效地提高了人脸检测的准确率，表明了Lightweight-SSH算法的有效性。

3.5　检测效果

图10所示为本文算法与MSCNN、MTCNN人脸检测算法的检测结果对比，矩形框表示检测出的人脸位置。由图10(a)、图10(b)、图10(c)可见，在检测密集的小人脸的情况下，Lightweight-SSH算法在检测小尺寸人脸上有较好的优势。



(a)Lightweight-SSH人脸检测效果　(b)MSCNN人脸检测效果

(c)MTCNN人脸检测效果

图10　不同算法的检测结果对比

Fig.10　Comparison of different algorithms detection results

4 结束语

针对移动端计算能力及存储资源受限的问题，本文提出一种基于轻量级卷积神经网络的改进人脸检测算法Lightweight-SSH，并从2个方面对现有的人脸检测算法SSH进行改进。使用MobileNet轻量级卷积神经网络代替VGG-16作为基础网络对样本数据进行特征提取，降低了模型的参数量和计算量，减少了内存占用性能。通过在SSH网络的检测模块中引入可变形卷积层，即可对不同形状的物体采用不同的几何形状进行采样，提升了卷积神经网络的形变建模能力，从而提高非约束环境下的人脸检测性能。实验结果表明，本文算法在保证准确性的前提下，能够提高模型处理速度，降低模型复杂程度，有效提升基于移动端的人脸检测算法性能。下一步将针对移动端的人脸检测率低于PC端的问题，在保证模型检测的时间性能基础上优化网络结构，从而更好地进行人脸检测。

参考文献

[1] ZHANG K P, ZHANG Z P, LI Z F, et al. Joint face detection and alignment using multi-task cascaded convolutional networks[J]. IEEE Signal Processing Letters, 2016, 23(10): 1499-1503.

[2] REN S, HE K, GIRSHICK R, et al. Faster r-CNN: towards real-time object detection with region proposal networks[EB/OL]. [2019-12-10]. https://arxiv.org/pdf/1506.01497.pdf.

[3] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detecto[C]//Proceedings of European Conference on Computer Vision. Berlin, Germany: Springer, 2016: 21-37.

[4] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]//Proceedings of IEEE NIPS'12. Washington D. C., USA: IEEE Press, 2012: 1097-1105.

[5] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2015: 1-9.

[6] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. [2019-12-10]. https://arxiv.org/pdf/1409.1556.

[7] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2016: 770-778.

[8] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2016: 2818-2826.

[9] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size[EB/OL]. [2019-12-10]. https://arxiv.org/pdf/1602.07360.

[10] HOWARD A G, ZHU M L, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[EB/OL]. [2019-12-10]. https://arxiv.org/pdf/1704.04861.

[11] DAI J, QI H, XIONG Y, et al. Deformable convolutional networks[C]//Proceedings of IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE Press, 2017: 764-773.

[12] NAJIBI M, SAMANGOUEI P, CHELLAPPA R, et al. SSH: single stage headless face detector[C]//Proceedings of IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE Press, 2017: 4875-4884.

[13] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[EB/OL]. [2019-12-10]. https://arxiv.org/pdf/1612.03144.

[14] FELZENSZWALB P F, GIRSHICK R B, MCALLESTER D, et al. Object detection with discriminatively trained part-based models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9): 1627-1645.

[15] 宋超，许道云，秦永彬. 分离多路卷积神经网络研究[J]. 计算机工程，2017，43(6)：145-149，157.
SONG C, XU D Y, QIN Y B. Research on detached multiple convolutional neural network[J]. Computer Engineering, 2017, 43(6): 145-149, 157. (in Chinese)

[16] 黄良辉，康祖超，张昌凡，等. 基于轻量级卷积神经网络的人脸识别方法[J]. 湖南工业大学学报，2019，32(2)：43-47.
HUANG L H, KANG Z C, ZHANG C F, et al. Research on face recognition technology based on lightweight convolutional neural networks[J]. Journal of Hunan University of Technology, 2019, 32(2): 43-47. (in Chinese)

[17] 蔡汉明，随玉腾，张镇，等. 基于深度可分离卷积神经网络的农作物病害识别方法[J]. 安徽农业科学，2019，47(11)：244-246，252.
CAI H M, SUI Y T, ZHANG Z, et al. Crop disease recognition based on depthwise separable convolutional neural network[J]. Journal of Anhui Agricultural Sciences 2019, 47(11): 244-246, 252. (in Chinese)

[18] 孙贵华，陈淑荣. 一种改进的RefineDet多尺度人脸检测方法[J]. 电子技术应用，2019，45(8)：34-39.
SUN G H, CHEN S R. An improved RefineDet multi-scale face detection method[J]. Application of electronic technology, 2019, 45(8): 34-39. (in Chinese)

[19] YANG S, LUO P, LOY C C, et al. Wider Face: a face detection benchmark[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2016: 5525-5533.

[20] CAI Z W, FAN Q F, FERIS R S, et al. A unified multi-scale deep convolutional neural network for fast object detection[C]//Proceedings of European Conference on Computer Vision. Berlin, Germany: Springer, 2016: 354-370.

编辑 索书志

计算机工程 www.ecice06.com