

doi:10.3969/j.issn.1673-9833.2019.02.008

基于轻量级卷积神经网络的人脸识别方法

黄良辉¹, 康祖超², 张昌凡², 程涛²

(1. 广东南海鹰视通达科技有限公司, 广东 佛山 528200; 2. 湖南工业大学 电气与信息工程学院, 湖南 株洲 412007)

摘要: 针对传统卷积神经网络在人脸识别中模型复杂程度高、处理数据较慢的问题, 提出一种轻量级卷积神经网络算法。首先, 通过对数据集采用剪裁、旋转等方式增强样本数据; 然后, 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, 并采用 SSD 目标检测器对样本数据中的人脸进行识别; 最后, 利用 Python 编程实现上述算法, 并与传统的人脸识别算法进行比较。实验结果表明, 采用的轻量级卷积神经网络算法在不失精度的前提下, 处理速度更快, 模型复杂程度更低。

关键词: 轻量级卷积神经网络; MobileNet; 目标检测; 人脸识别

中图分类号: TP391

文献标志码: A

文章编号: 1673-9833(2019)02-0043-05

引文格式: 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, 33(2): 43-47.

Research on Face Recognition Technology Based on Lightweight Convolutional Neural Networks

HUANG Lianghui¹, KANG Zuchao², ZHANG Changfan², CHENG Tao²

(1. Guangdong Nanhai Eagle View Technology Co., Ltd., Foshan Guangdong 528200, China;
2. College of Electrical and Information Engineering, Hunan University of Technology, Zhuzhou Hunan 412007, China)

Abstract: In view of the flaws of high complexity and slow data processing of face recognition found in traditional convolution neural networks, a lightweight convolutional neural network algorithm has thus been proposed. Firstly, the sample data is to be enhanced by clipping and rotating the data set. Then, a lightweight convolutional neural network based on MobileNet is used to extract the features from sample data, and an SSD target detector is used for face recognition in sample data. Finally, the above algorithm is implemented by Python programming, followed by a comparison with the outcome of the traditional face recognition algorithm. The experimental results show that the proposed lightweight convolution neural network algorithm is faster in the processing speed with a lower model complexity and retained accuracy.

Keywords: lightweight convolutional neural network; MobileNet; target detection; face recognition

1 研究背景

随着人机交互技术的多样化发展, 人脸识别被

越来越多地应用在可穿戴电子设备上。但是现阶段的可穿戴电子设备硬件性能较低, 传统的人脸识别算法在硬件性能受限时, 无法快速或根本无法实现

收稿日期: 2018-11-11

基金项目: 国家自然科学基金资助项目 (61773159)

作者简介: 黄良辉 (1980-), 男, 广东广州人, 广东南海鹰视通达科技有限公司董事长, 博士, 主要研究方向为便携式电子设备交互技术, E-mail: raphael.huang@eagle-view.cn

预设功能^[1]。因此,研究一种网络复杂程度较低,处理速度较快的人脸识别算法显得尤为重要。

国内外学者已经对人脸识别算法进行了深入研究,提出了特征脸^[2]、弹性图匹配和LBP(local binary patterns)^[3]等诸多经典方法。如李倩玉等^[4]构建了一个能够自动提取深层特征的特征提取器,并对数据进行ZCA(zero-mean component analysis)白化处理,提出一种改进深层网络的人脸识别算法,该算法中所使用的卷积核主要通过无监督学习获得,该方法相较于传统方法在准确度和性能上都有所提高。Zhang K.等^[5]针对面部检测时各种姿势、照明和遮挡问题,提出一个深度级联任务框架,采用由粗略到精细的方式预测面部和地标位置,并保证了实时性能。杨洁等^[6]研究了一种改进的SSD(single shot multibox detector)卷积网络视频目标检测模型,其能够有效地识别视频中较远场景下的目标,且减少了计算量,降低了硬件的内存资源消耗。但是上述算法都存在计算复杂程度高的特点,在可穿戴电子设备的应用上具有一定的局限性。

随着对卷积神经网络算法研究的不断深入,在保证算法准确度的前提下,对其网络复杂程度和可实现性提出了更高要求。何长婷^[7]通过迁移学习构建了基于MobileNet模型的图像质量分类器,对不同质量的图像进行分类,得到了较高的准确率,有效地减少了模型的复杂程度,保证了模型的准确性。孙旭等^[8]针对人脸在动态视频中的检测,提出单网络多目标的人脸检测方法,结合半自动化数据标注策略和多任务学习思想,满足了监控视频对特征分析的准确性和实时性需求。MobileNet是一种轻量级的卷积神经网络,能在硬件性能有限的前提下实现数据特征提取;SSD是一种准确度高且检测快速的目标检测器^[9]。研究表明,结合SSD和MobileNet的MobileNet_SSD网络能够快速精确地实现人脸识别。基于上述研究工作,本文提出一种基于MobileNet的轻量级卷积神经网络对样本数据进行特征提取,并通过SSD目标检测器对样本数据中的人脸进行识别的算法。该算法可实现人脸识别,并对数据集采用剪裁、旋转等方式做样本数据增强;最后与传统卷积神经网络人脸识别算法比较,验证了算法的优越性和有效性。

2 系统描述

基于轻量级卷积神经网络的人脸识别算法由3部分组成:预处理模块、轻量级卷积神经网络特征提取器和目标检测器,如图3所示。预处理模块通过人工采集的图像样本,经过数据增强处理得到更大的图像

集,然后对图像进行标注得到合适的数据集。轻量级卷积神经网络特征提取器主要完成数据集的高层次特征提取。通过目标检测器直接对提取的特征图进行多尺度分类和检测,最终实现人脸识别。

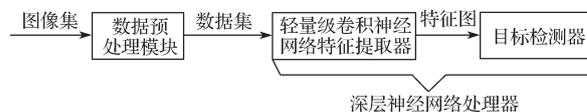


图1 算法系统结构

Fig. 1 Algorithm system architecture

3 算法实现

3.1 数据预处理模块

数据预处理模块主要由数据增强和数据标定构成。数据增强是对采集的图片样本进行翻转、裁剪及亮度调节等操作,能够在样本较少的前提下增加数据集数量,同时提高数据集复杂度,有效提高模型的识别效果。数据标定主要利用labelImg工具对经过数据增强处理的数据进行数据标定,从而得到需要的数据集。

课题组标定的人脸数据主要为人脸所在宽高、人脸检测框坐标点及人脸标签信息,见表1。

表1 数据制作

Table 1 Data production

文件	标签	人脸大小		人脸检测框位置			
		宽	高	x_{min}	y_{min}	x_{max}	y_{max}
Picture1.jpg	ZXC	556	300	157	1	416	265
Picture2.jpg	ZXC	500	321	76	1	418	265
Picture3.jpg	ZXC	500	354	236	24	422	184
Picture4.jpg	ZXC	499	346	59	22	324	251
Picture5.jpg	ZXC	500	350	184	1	410	258

注:数据标定完成后的文件为csv格式,而tensorflow平台下需要制作成record格式文件,才能生成tensorflow平台能够识别的标准训练集。因此,需将csv格式文件转化为record格式文件。

3.2 轻量级神经网络处理器

传统卷积神经网络虽然在数据处理中具有良好的特征提取效果,但是模型的计算复杂度较高,无法适用硬件性能有限的设备。因此,在轻量级神经网络处理器中采用基于MobileNet的轻量级卷积神经网络对样本数据进行特征提取;最后,采用SSD目标检测器对样本数据中的人脸进行识别。

3.2.1 MobileNet 特征提取器设计

课题组采用的特征提取模型主要采用了轻量级的卷积神经网络,MobileNet网络架构是一种结构简化、小尺寸的适用于移动设备的视觉模型。其具有功耗低、结构小的特点,但依然可被用于大规模的图像检测和分类中,并最大程度地保证了模型的准确度^[10]。MobileNet网络主要包括卷积层、BatchNormal层、

ReLU层以及最大池化层。

1) 卷积层。MobileNet 卷积网络在卷积层将传统的卷积神经网络分为 Depthwise 卷积和 Pointwise 卷积。Depthwise 卷积为 3×3 的卷积核, Pointwise 卷积为 1×1 的卷积核, 通过批量标准化 (batch normalization, BN) 和线性修正单元 (rectified linear unit, ReLU) 激活函数来构建网络的卷积层。MobileNet 卷积网络在保证网络准确性的前提下, 减少了模型的网络参数, 传统的卷积神经网络卷积核 K 的参数量^[10]如式(1)所示:

$$K = D_K \cdot D_K \cdot M \cdot N \cdot D_F \cdot D_F \quad (1)$$

式中: D_K 为传统卷积核的大小; M 为输入数据的通道数; N 为所采用核的个数; D_F 为输入数据的大小。

MobileNet 的卷积核参数量 K' 如式(2)^[11]所示:

$$K' = D_K \cdot D_K \cdot M \cdot D_F \cdot D_F + M \cdot N \cdot D_F \cdot D_F \quad (2)$$

通过卷积分解压缩, 计算量比如式(3)所示:

$$\frac{D_K \cdot D_K \cdot M \cdot D_F \cdot D_F + M \cdot N \cdot D_F \cdot D_F}{D_K \cdot D_K \cdot M \cdot N \cdot D_F \cdot D_F} = \frac{1}{N} + \frac{1}{D_K \cdot D_K} < 1 \quad (3)$$

通过比较发现, MobileNet 网络模型参数计算量是传统卷积神经网络计算量的 $1/N + 1/(D_K \cdot D_K)$, 其网络的计算量大幅度减少, 提高了系统性能且精确率下降程度较低。

2) BatchNormal 层。该层对输入数据的各维度进行批量规范化, 在系统中加入该层可有效提升网络学习速率。

3) ReLU 层。这是一个非线性激活单元, 通常作为激活函数用在全连接层之后。该层的表达式如式(4)所示^[12]。

$$L(X) = \max(X, 0) \quad (4)$$

式中: $L(X)$ 为网络层的函数; X 为该层输入的数据张量。此处是取 X 与 0 中的较大值。

ReLU 层没有需要迭代的参数, 其更多地被用在现代神经网络中的激活函数中, 因为和 sigmoid 和 tanh 函数比较, 其计算更加简单, 极大地减少了训练中的计算开销; 同时 ReLU 是单端饱和的激活函数, 可有效减少 sigmoid 这种双端饱和函数在梯度反向传播时造成的衰减现象, 能够加快训练的收敛速度。

4) 最大池化层。该层针对特征图上每个 2×2 区域的像素值取出一个最大值来减小特征图的尺寸, 即取出每个 2×2 区域中 4 个像素值中的最大值, 最终输出图片的宽和高是输入图片的一半。该层使网络在前向传播的过程中抽象到更高层级, 使卷积层可以使用不同的尺度对图像进行特征提取。

3.2.2 目标检测器

目前基于深度学习的目标检测模型很多, 像

RCNN (regions with convolutional neural network features) 系列及 YOLO (you only look once)^[13] 等都是通过预选框进行目标检测。如 RCNN 系列的目标检测框架主要通过两步过程完成检测工作: 第一步, 对输入样本进行特征图提取; 第二步, 将提取的特征图再次通过卷积神经网络进行处理, 最后进行分类及检测。基于 RCNN 系列的检测模型的最大不足之处在于对数据的处理速度太慢, 同时通过采用 Selective Search 算法得出候选框的处理中存在多个重复区域的计算^[11]。本文采用的 SSD 方法直接对一张图片进行处理, 对最后不同尺度的特征图进行分类检测, 检测样本中的人脸位置。SSD 目标检测网络是通过 RCNN 系列检测模型衍生出来的、检测更快的模型^[14], 继承了 Faster RCNN 的良好性能, 又在其基础上改进优化, 最终演变为精确度较高、检测速率更快的网络模型。SSD 物体检测模型如图 2 所示。

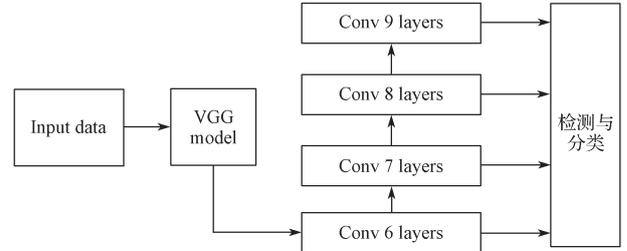


图 2 SSD 分类器

Fig. 2 SSD classifier

SSD 模型结构包含 VGG (visual geometry group) 卷积结构、卷积层以及多尺度分类器。本研究将 MobileNet 网络结构改进的新网络模型和目标检测 SSD 卷积神经网络模型结合, 然后进行人脸检测, 其改进之处在于将 SSD 网络中用于特征提取的 VGG Model 替换为 MobileNet 网络, 从而在保证模型识别精度的前提下最大化地减少了整体网络的计算复杂度。网络目标损失函数包括定位损失函数与回归损失函数^[13], 这 2 个函数可在训练过程中对网络结构的准确度进行判定, 利用损失函数更新网络参数, 目标损失函数 $L(x, c, a, b)$ 如式(5)所示^[15]:

$$L(x, c, a, b) = \frac{1}{N} (L_{\text{conf}}(x, c) + \alpha L_{\text{loc}}(x, a, b)) \quad (5)$$

式中: N 为单元盒数目; α 为权值且其初始值为 1; x 为中心坐标; c 为类别置信度预测值; a 为先验框所对应边界框的位置预测值; b 为 ground truth 的位置参数。

其定位损失函数 $L_{\text{loc}}(x, a, b)$ ^[15] 为

$$L_{\text{loc}}(x, a, b) = \sum_{i \in \text{Pos}} \sum_{m \in \{c_x, c_y, w, h\}} x_{ij}^k \text{smooth}_{L1}(l_i^m - \hat{g}_j^m) \quad (6)$$

式中: Pos 表示正样本边框的位置; c_x, c_y 分别为边

框左上角顶点的坐标; w 、 h 分别为边框的宽和高; x_{ij}^k 为第 i 个预测框与第 j 个真实框关于第 k 个类别是否一致, 匹配为 1, 否则为 0; $smooth_{L1}$ 是对位置误差采用 Smooth L1 Loss; l_i^m 为预测框; \hat{g}_i^m 为真实框。 x 、 a 、 b 参数为目标函数传递目标位置信息^[16]; 回归损失函数为识别种类属性进行函数评估, 作为人脸识别判断依据。

其置信度误差损失函数为^[15]

$$L_{\text{conf}}(x, c) = - \sum_{i \in \text{Pos}} x_{ij}^p \log(\hat{c}_i^p) - \sum_{i \in \text{Neg}} \log(\hat{c}_i^0) \quad (7)$$

式中: $x_{ij}^p \log(\hat{c}_i^p)$ 为预测框 i 与真实框 j 关于类别 p 匹配, 该概率预测值越高, 损失越小; $x_{ij}^p \in \{0, 1\}$ 为一个指数参数, 当其值等于 1 时, 表示第 i 个先验框与第 j 个 ground truth 匹配; \hat{c}_i^p 为类别置信度预测值; Neg 为负样本的边框位置; \hat{c}_i^0 表示预测框没有物体, 该概率越高则预测为背景的概率越高, 该参数概率通过 *softmax* 产生。

4 实验及结果分析

4.1 实验环境以及实验数据

本研究中的软硬件实验平台为 PC 端, 操作系统为 Ubuntu16.04, GPU 显卡为 NVIDIA GeForce GTX 970, 系统型号为 x64_84, 内存为 8 GB, 框架采用 tensorflow 深度学习框架, 实验数据为人工采集数据并进行标注。

4.2 PC 端人脸识别系统实验结果分析

用 Python 语言搭建上述神经网络, 在 Linux 系统中运行、训练网络, 卷积神经网络识别损失率和检测损失率的变化曲线如图 3 所示。

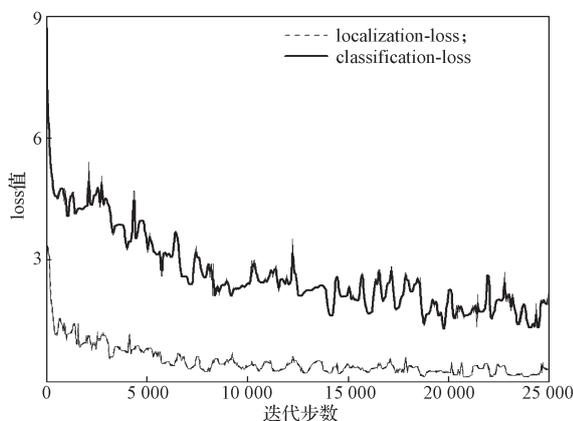


图 3 分类与定位损失 loss 值

Fig. 3 Loss value of classification and location

由图 3 可知, 接近 25 000 步时, 该模型的收敛效果趋于稳定, 可能由于手工标注的原因, loss 值存在一定的跳变, 但是其人脸检测损失值最低降到 0.2

左右, 而人脸识别的损失值也基本稳定于 2.0 左右。通过一系列人脸识别任务之后, 对人脸识别性能进行评估。表 2 为不同模型的性能对比。由表 2 可得到基于 Face_data 自制的数据集在不同网络模型上进行检测任务所产生的平均精确度。该试验由于设备不同等因素影响, 实验结果存在一定的误差, 与理想状态下的目标检测效果有一定的差距。

表 2 不同模型性能对比

Table 2 Performance comparison of different models				
模型	数据库	平均精度 / %	模型大小 (计算参数)	检测时间 / s
SSD_ResNet	Face_data	75.0	42 605 504	2.312
SSD_Inception	Face_data	73.2	10 173 112	2.103
SSD_MobileNet	Face_data	71.5	3 191 072	1.908

由表 2 可得, 本研究采用方法与传统的特征提取方法 ResNet 与 Inception 结合 SSD 进行人脸检测做对比。ResNet 是一种深度残差网络, 能够在加深网络深度的同时保证特征的有效提取; Inception 网络则是把 1×1 、 3×3 、 5×5 的卷积核和 3×3 的 pooling 层堆叠在一起, 不仅增加了网络的宽度, 同时增加了网络对尺度的适应性。对比发现, SSD_MobileNet 相对于其他模型, 其准确度相差不大, 并且网络的模型参数最少, 数据处理速度更快。采用 SSD_MobileNet 网络能够减少网络的计算参数, 同时具有更快的处理速度, 其网络能够更好地应用到硬件设备当中。

4.3 Android 人脸识别系统实现

由于基于 tensorflow 平台开发的轻量级卷积神经网络模型只适合在 ARM 架构的系统中运行, 智能眼镜需装载 ARM 架构的主控芯片, 并且主控芯片本身提供完整 Android 系统环境, 才能将模型移植到智能眼镜之中。

本系统开发使用的环境为 Android Studio, 对上述训练好的文件通过 tensorflow 平台提供的 tf.mobile 进行进一步提取和优化后集成到 AndroidStudio 的 Assets 文件中, 完成整个系统的集成工作。通过智能眼镜摄像头采集图片并进行识别, 识别效果相对实验结果准确度相近, 但是略低于实验结果, 整个系统处理速度也相对较快, 达到了整合到智能眼镜上实现的预期效果。

5 结语

针对传统卷积神经网络在人脸识别中, 模型复杂程度高、处理数据较慢、其网络无法在硬件性能受限快速实现或根本无法实现其预设功能的问题, 本文提出一种 MobileNet_SSD 改进网络模型人脸识别方法。通过对数据集采用剪裁、旋转等方式增强样本数

据; 然后采用基于 MobileNet 的轻量级卷积神经网络对样本数据进行特征提取; 最后采用 SSD 目标检测器对数据中人脸进行识别, 系统网络模型在训练过程中得到很好的收敛效果, 模型训练计算复杂度降低, 不占用内存性能。在不失准确度的前提下, 模型处理速度更快, 复杂程度更低。但是训练网络结果仍然有很大的提升空间, 网络的精确度提高与性能关系并未达到高度一致性, 因此在今后实验中将进一步对网络进行改进, 完善网络结构, 实现更高鲁棒性。

参考文献:

- [1] 胡骏飞, 文志强, 谭海湖. 基于二值化卷积神经网络的手势分类方法研究 [J]. 湖南工业大学学报, 2017, 31(1): 76-80.
HU Junfei, WEN Zhiqiang, TAN Haihu. Research on Gesture Classification Method Based on Binarized Convolutional Neural Network[J]. Journal of Hunan University of Technology, 2017, 31(1): 76-80.
- [2] CAI Z W, FAN Q F, FERIS R S, et al. A Unified Multi-Scale Deep Convolutional Neural Network for Fast Object Detection[C]//European Conference on Computer Vision 2016. Amsterdam: Springer, 2016: 354-370.
- [3] 马国峻, 周海东. 轻量级智能终端人脸识别系统研究与实现 [J]. 通讯学报, 2015, 36 (增刊 1): 149-156.
MA Guojun, ZHOU Haidong. Research and Implementation of Lightweight Intelligent Terminal Face Recognition System[J]. Journal of Communications, 2015, 36(S1): 149-156.
- [4] 李倩玉, 蒋建国, 齐美彬. 基于改进深层网络的人脸识别算法 [J]. 电子学报, 2017, 45(3): 619-625
LI Qianyu, JIANG Jianguo, QI Meibin. Face Recognition Algorithm Based on Improved Deep Networks[J]. Acta Electronica Sinica, 2017, 45(3): 619-625.
- [5] ZHANG K, ZHANG Z, LI Z, et al. Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks[J]. IEEE Signal Processing Letters, 2016, 23(10): 1499-1503.
- [6] 杨洁, 陈明志, 吴智秦, 等. 基于 SSD 卷积网络的视频目标检测研究 [J]. 南华大学学报(自然科学版), 2018, 32(1): 79-86.
YANG Jie, CHENG Mingzhi, WU Zhiqing, et al. Research on Video Target Detection Based on SSD Convolution Network[J]. Journal of University of South China(Science & Technology), 2018, 32(1): 79-86.
- [7] 何长婷. 课堂签到系统中的人脸识别方法研究与实现 [D]. 合肥: 中国科学技术大学, 2018.
HE Changting. Research and Implementation of Face Recognition Method in Classroom Sign-in System[D]. Hefei: University of Science and Technology of China, 2018.
- [8] 孙旭, 胡伟, 李瑞瑞. 基于 Lighten CNN 的自动人脸分类模型的优化方法 [J]. 计算机应用, 2018, 38(增刊 1): 32-35.
SUN Xu, HU Wei, LI Ruirui. Optimization Method of Automatic Face Classification Model Based on Lighten CNN[J]. Journal of Computer Applications, 2018, 38(S1): 32-35.
- [9] AGRAWAL N, PRABHAKARAN V, WOBBER T, et al. Design Tradeoffs for SSD Performance[J]. Security & Privacy IEEE, 2008, 7(2): 57-70.
- [10] QIN Z, ZHANG Z N, CHEN X T, et al. FD-MobileNet: Improved MobileNet with a Fast Downsampling Strategy[J/OL]. [2018-02-13]. <https://arxiv.org/abs/1802.03750>.
- [11] 陈方. MobileNet 压缩模型的研究与优化 [D]. 武汉: 华中师范大学, 2018.
CHEN Fang. Research and Optimization of MobileNet Compression Model[D]. Wuhan: Central China Normal University, 2018.
- [12] HOSSAIN A K M, JIN Yunye, SOH Weeseng, et al. SSD: A Robust RF Location Fingerprint Addressing Mobile Devices' Heterogeneity[J]. IEEE Transactions on Mobile Computing, 2013, 12(1): 65-77.
- [13] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection [C]//CVPR 2016: Proceedings of the 2016 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE, 2016: 779-788.
- [14] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal Loss for Dense Object Detection[C]//IEEE Transactions on Pattern Analysis & Machine Intelligence 2017. Venice: IEEE, 2017: 2999-3007.
- [15] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot MultiBox Detector[C]// European Conference on Computer Vision 2016. Munich: Springer, 2016: 21-37.
- [16] 叶远征, 李小霞, 李旻择. 基于 MS-KCF 模型的图像序列中人脸快速稳定检测 [J]. 计算机应用, 2018, 38(8): 2192-2197.
YE Yuanzheng, LI Xiaoxia, LI Minze. Rapid Stable Detection of Human Faces in Image Sequence Based on MS-KCF Model[J]. Journal of Computer Applications, 2018, 38(8): 2192-2197.

(责任编辑: 申剑)