

# 深度学习识别光网络单元故障的设计与应用

汤斯鹏<sup>1</sup>, 池鸿源<sup>1</sup>, 张培炜<sup>1</sup>, 张炳华<sup>2</sup>, 蔡毅<sup>3</sup>

(1. 中国移动通信集团广东有限公司 AI 能力支撑中心, 广东 汕头 515000;

2. 中国移动通信集团广东有限公司 AI 能力支撑中心, 广东 广州 510000;

3. 华南理工大学 软件学院, 广东 广州 510000)

**摘要:**为解决依赖装维上门鉴别光网络单元故障带来的不便,可以从机器视觉入手实现自动化故障识别。近年,ImageNet 挑战赛的成功推动了物体识别技术的跨越式发展,特别是基于卷积的深度学习技术在视觉识别方面已经达到人类水平,为光网络单元故障的自动识别提供了技术基础。文章对识别光网络单元的工作状态进行了研究,将设备工作状态分为7个场景,提出了利用手机 APP 采集图片识别故障的解决方案并投入了实际生产;重点阐述了深度学习模块的设计与实现,提出一种通过算法整合的方式综合运用物体检测和图像分类算法,分3阶段逐步求精,解决了图片过滤,光网络单元型号和状态识别等问题,实现了基于计算机视觉自动识别光网络单元故障。从数据上看产品的端到端准确率超过84%,识别速度达到10 FPS,月均提供服务超过1万人次,在减少用户等待的同时节约了人力资源。

**关键词:**深度学习;物体检测;图片分类;客户服务;光网络单元

中图分类号:TP391.4

文献标识码:A

文章编号:1673-629X(2020)05-0211-05

doi:10.3969/j.issn.1673-629X.2020.05.040

## Design and Application of Identifying Malfunctions in Optical Network Units Based on Deep Learning

TANG Si-peng<sup>1</sup>, CHI Hong-yuan<sup>1</sup>, ZHANG Pei-wei<sup>1</sup>, ZHANG Bing-hua<sup>2</sup>, CAI Yi<sup>3</sup>

(1. AI Capability Support Center of China Mobile Communications Group Guangdong Co., Ltd., Shantou 515000, China;

2. AI Capability Support Center of China Mobile Communications Group Guangdong Co., Ltd., Guangzhou 510000, China;

3. School of Software Engineering, South China University of Technology, Guangzhou 510000, China)

**Abstract:** To solve the inconvenience caused by the failure of the manual identification optical network unit, automatic fault identification can be realized from machine vision. The success of the ImageNet Challenge has promoted the leap-forward development of object recognition technology, especially the convolution-based deep learning technology has reached the human level in visual recognition, providing a technical basis for automatic identification of optical network unit malfunctions. We study the working status of optical network units, divide the working status of equipment into seven scenarios, and put forward a solution of using the mobile phone APP to collect pictures to identify faults and put it into actual production. We focus on the design and implementation of the deep learning module and propose a comprehensive method of object detection and image classification by algorithm integration, which is in three stages, solving the problems of image filtering, optical network unit model and status recognition, and finally achieving the automatic identification of optical network unit faults based on computer vision. Saving human resources while reducing user waiting, the end-to-end accuracy rate of the product exceeds 84%, the recognition speed reaches 10 FPS, and the monthly service provides more than 10 000 times.

**Key words:** deep learning; object detection; image classification; customer service; optical network unit

## 0 引言

旨在完成百万级规模数据分类和检测的

ImageNet 挑战赛 2009 年开始在 CVPR 举办,自 2012 至 2017 年挑战赛结束冠军队伍均采用了深度学习算

收稿日期:2019-06-04

修回日期:2019-10-09

网络出版时间:2020-01-10

基金项目:广东省特支计划青年拔尖项目(2015TQ01X633)

作者简介:汤斯鹏(1987-),男,硕士,研究方向为深度学习算法应用。

网络出版地址:<http://kns.cnki.net/kcms/detail/61.1450.TP.20200110.1118.008.html>

法,物体分类错误率从 0.28 降到了 0.03;物体识别的平均准确率从 0.23 上升到了 0.66。ImageNet 的成功推动了基于卷积神经网络的深度学习技术在视觉领域的快速发展,利用卷积、池化操作具备位移不变性、能够合并相似语意特征的特点,在自然场景下图像物体的分类和检测的精度上已经接近并超过人类的水平,为解决视觉分类和物体检测问题提供了强大的支撑。

光网络单元(以下简称 ONU)属于接入网的用户侧设备<sup>[1]</sup>,是电信运营商宽带业务实现“光电转换”的重要设备。不同厂商生产的 ONU 外形不尽相同(见图 1),普通用户受专业限制,一旦 ONU 异常或故障,

往往难以描述并只能等待装维人员上门检修。对用户而言浪费时间等待,对运营商而言消耗大量人力成本检修。如果可以通过计算机视觉技术识别故障甚至引导客户自助排除简单故障,有利于提升用户体验同时降低企业人力成本。

文中以物体检测实现故障识别为切入点,综合应用了 SSD, Inception-V4 和 Inception-Resnet-V2,提出了一种识别 ONU 故障的设计,并在“广东移动手机营业厅”APP 中投入了实际应用。从用户数据上看,该系统的端到端准确率超过 84%。

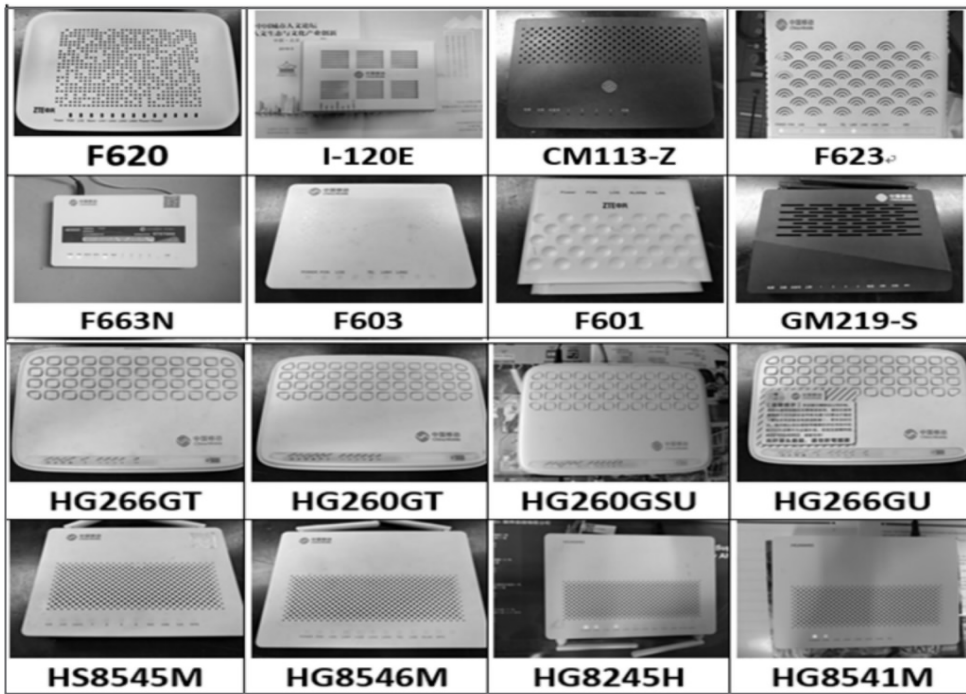


图 1 几款不同型号的 ONU 设备示例

## 1 相关技术

### 1.1 SSD

SSD(single shot multibox detector<sup>[2]</sup>)是一种单步骤的目标检测方法:比同为单步骤的 YOLO V1<sup>[3]</sup>速度快而精度达到两步骤方法 Faster-RCNN<sup>[4]</sup>的水平;SSD 的思想是用小尺寸卷积核对预设候选框预测坐标偏移量和分类概率;SSD 产生固定比例的候选框集合并对这些候选框中存在的目标类的概率,使用非最大化抑制步骤以产生最终检测,利用多个尺寸的特征图预测不同长宽比的候选框综合起来提升检测精度;是一种同时权衡了精度和速度的端到端的目标检测算法。

SSD 在主干网络的末尾添加了卷积层。卷积层的尺寸从大到小,得到不同尺度检测的概率。每个添加的特征层(或可选的基础网络中已存在的特征层)

可以使用一组卷积滤波器产生固定的预测集合。对于具有  $p$  个通道的大小为  $m \times n$  的特征层,使用  $3 \times 3 \times p$  卷积核进行卷积操作,使每个类别产生一个预测概率,以及相对于候选框的坐标偏移。

### 1.2 Inception 网络及其演进

Inception<sup>[5]</sup>网络是卷积神经网络分类器发展史上一个重要的里程碑。在 Inception 出现之前,大部分流行卷积神经网络仅仅是把卷积层堆叠得越来越多,使网络越来越深,以此希望能够得到更好的性能。但这种简单粗暴的方法存在一些问题:会导致更大的参数空间,更容易过拟合、需要更多的计算资源、网络越深,梯度容易消失,优化困难。Inception 的目标是,提高网络计算资源的利用率,在计算量不变的情况下,提高网络的宽度和深度。Inception 把全连接改成稀疏连接,卷积层也是稀疏连接,卷积网络可以近似最优局部稀疏结构,并且该结构下可以用现有的密度矩阵计算硬

件实现,产生的结果就是 Inception,其作用是替代了人工确定卷积层中过滤器的类型或者是否创建卷积层和池化层,让网络自动学习参数。Inception-V2<sup>[6]</sup>在 Inception 的基础上增加了批数据归一化。Inception-V3<sup>[7]</sup>在 Inception 的基础上对卷积核进行了分解,将 5 \* 5 的卷积核替换成 2 个 3 \* 3 的卷积核;将 n \* n 的卷积核替换成 1 \* n 和 n \* 1 的卷积核堆叠。这样加速了计算,还增加了网络的非线性,减小了过拟合的概率。

### 1.3 Inception-Resnet-V2 和 Inception-V4

在 Inception-V3 的基础上,受 ResNet 的优越性能启发引入残差结构,提出了 Inception-Resnet-V2<sup>[8]</sup>,并修改 Inception 模块提出了 Inception-V4 结构,基于 Inception-V4 的网络实验发现在不引入残差结构的基础上也能达到和 Inception-Resnet-V2 结构相似的结果。ILSVR2012 比赛中 Inception-V3、Inception-V4、Inception-Resnet-V2 的 Top-5 Error 分别为 4.3%、

3.8% 和 3.7%。

## 2 系统设计和实现

### 2.1 ONU 常见状态

不同生产厂商制造的 ONU 设备的型号和外形不尽相同。据广东移动统计,截至 2017 年 12 月,广东移动投放至市场的 ONU 类型有 34 种。各种型号 ONU 市场占有率如表 1 所示。虽然 ONU 型号众多,但 ONU 设备上具有 Power、Pon、Los、Lan,4 种类型指示灯,通过 ONU 设备上指示灯的状态组合,可以识别 ONU 所处的工作状态。如表 1 所示,ONU 设备的指示灯状态可以分为 7 个场景,包含了 ONU 常见的 7 种状态。从表中可以知道场景 1、4、5、6 代表设备故障,场景 2、3、7 代表设备正常。如图 2 所示,不同型号的 ONU 外形不同,指示灯的排列和显示不尽相同,因此识别 ONU 的型号是必要的。

表 1 ONU 指示灯对应场景

| 场景 | Power 灯 | Pon 灯 | Los 灯 | Lan 灯 | 场景对应状态描述      |
|----|---------|-------|-------|-------|---------------|
| 1  | 灭       | 灭     | 灭     | 灭     | 设备电源故障        |
| 2  | 亮       | 灭     | 灭     | 灭     | 设备启动中         |
| 3  | 亮       | 灭     | 灭     | 亮     | 设备启动中         |
| 4  | 亮       | 灭     | 亮     | 灭     | 光信号弱、无局域网信号   |
| 5  | 亮       | 灭     | 亮     | 亮     | 光信号弱、有局域网信号   |
| 6  | 亮       | 亮     | 灭     | 灭     | 设备工作正常、无局域网信号 |
| 7  | 亮       | 亮     | 灭     | 亮     | 设备工作正常、有局域网信号 |

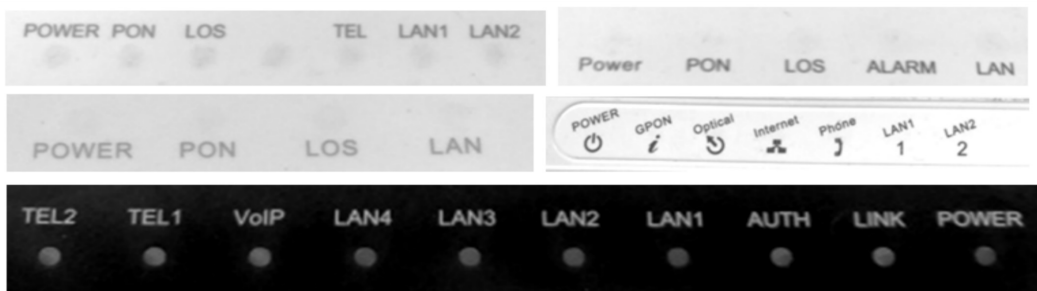


图 2 不同型号 ONU 指示灯

### 2.2 整体结构设计

针对以上陈述,文中提出的系统从视觉领域着手主要实现了以下核心功能:

- 识别自然场景图像中是否包含 ONU 设备;
- 识别图像中 ONU 设备型号;
- 识别图像中 ONU 设备指示灯对应场景。

核心功能的开发采用了谷歌开源的 tensorflow<sup>[9]</sup>深度学习框架。

为了完成整体系统应用,系统采用了 CS 模式<sup>[10]</sup>。用户打开 APP 调用摄像头采集图像,利用 webserver<sup>[11]</sup>方式上传图像到后台 GPU 服务器,GPU 服务器

调用经过训练的模型识别 ONU 型号和场景并返回手机 APP 前端,APP 前端使用虚拟现实<sup>[12]</sup>方式与用户交互。数据流程如图 3 所示。下面章节将详细介绍识别相关的核心功能设计与实现,其余部分将不再赘述。

### 2.3 深度学习模块设计与实现

深度学习模块实现分为以下 3 个阶段:

阶段 1(检测 ONU):自然场景下的用户图片往往采集随意且包含许多噪音,为了聚焦主体任务,有必要将图片中无关信息进行裁剪。使用 SSD 算法进行物体检测,检测的目标是 ONU 和 ONU 指示灯在图片中的位置,如图 4 所示。值得注意的是此时并未进行分

类,而是将所有型号的 ONU 共同训练,检测的物体只有 2 种类型(类型 1:ONU,类型 2:ONU 指示灯),目的是为了将不同型号的特征充分融合,提升鲁棒性,同时有利于减轻数据不平衡所带来的不利影响。用户的输入可能不是 ONU 图片,但是这个阶段并不进行数据过滤,而是将所裁剪的图片送入阶段 2 的分类模型中。

算法训练分类模型可以预测出 ONU 型号,为了能够判定用户是否按要求输入,训练 ONU 型号分类器时除已知的 ONU 型号类型之外增加了一个“未知”分类。巧妙地将“未知”分类的图片在型号分类时融入分类器的训练,不仅可以加快模型收敛速度,而且有利于提升 ONU 型号的分类准确率、提升判定该物体是否为 ONU 的准确率。即,经过阶段 1 的检测会将图片中最像 ONU 的部分裁剪后再进行分类,能够有效降低分类的难度。

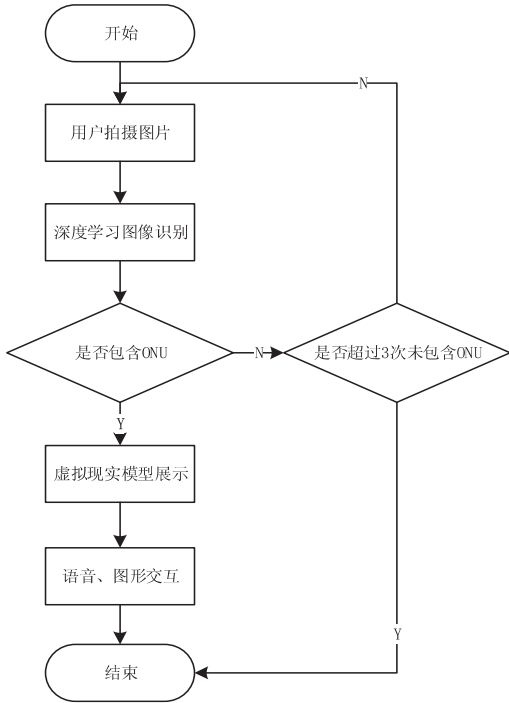


图 3 应用数据流程



图 4 ONU 检测

阶段 3(分类 ONU 指示灯):如图 2 所示,各种 ONU 的指示灯不尽相同,为了识别 ONU 的各种场景,可以将每个场景作为一个类来对待。同时为了平衡识别时长和识别准确率,使用 Iception-V4 算法分类 ONU 的 7 个场景;为了保证准确率,每个型号的 ONU 训练一个场景分类模型,输入数据使用经过 SSD 裁剪后的 ONU 指示灯图片。3 个阶段的数据流程如图 5 所示。

阶段 2(分类 ONU):完成位置检测将裁剪后类型 1(ONU 分类)的图片送入 Inception-Resnet-V2 分类

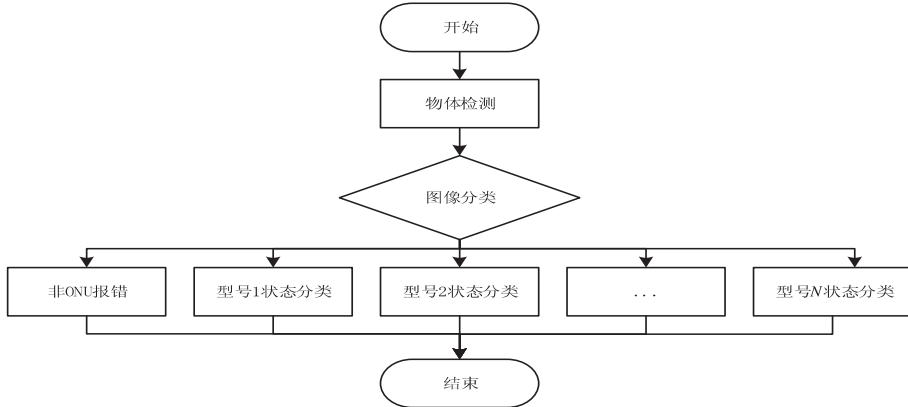


图 5 深度学习模块数据流程

### 3 模型训练

#### 3.1 损失函数

阶段 1 检测 ONU 的损失函数沿用了 SSD 的损失函数,SSD 损失函数分为两个部分,即对应搜索框的位置  $L_{loc}$  和类别置信度  $L_{conf}$  :

$$L(x, c, l, g) = \frac{1}{N} [L_{conf}(x, c) + \alpha L_{loc}(x, l, g)]$$

其中,  $N$  是匹配的包围框的个数,  $x$  表示匹配的框是否

属于类别  $p$ , 取值  $(0, 1)$ ;  $l$  是预测框,  $g$  是真实值;  $c$  是指所框选目标属于类别  $p$  的置信度。

阶段 2 和阶段 3 的损失函数均使用了交叉熵,  $y_i$  就表示期望输出,  $y_{i-}$  表示原始的实际输出。

$$loss = - \sum_{i=1}^n y_i \log(y_{i-})$$

#### 3.2 训练细节

为了实现冷启动,在未得到用户数据之前,模型的训练采用了人工拍摄的数据进行,在应用上线 1 个月

后收集用户数据进行标注。在原有模型的基础上进行训练,加入用户数据显著提升了模型的准确率,同时在训练的过程中进行了角度、亮度、对比度、饱和度的数据增强<sup>[13]</sup>,提升模型鲁棒性,利用数据过采样<sup>[14]</sup>的方式解决型号 ONU 数据不平衡问题。SSD 中输入图片的分辨率为  $300 \times 300$ ,L2 正则化系数为  $4 \times 10^{-5}$ ,使用了 RMSProp 优化器,初始学习率为 0.004,衰减系数为 0.95;在 Inception-V4 和 Inception-Resnet-V2 中输入图片统一缩放到  $299 \times 299$ ,使用了 Adam<sup>[15]</sup>优化器,初始学习率为 0.001,衰减率为 0.93。图 6 是分类指示

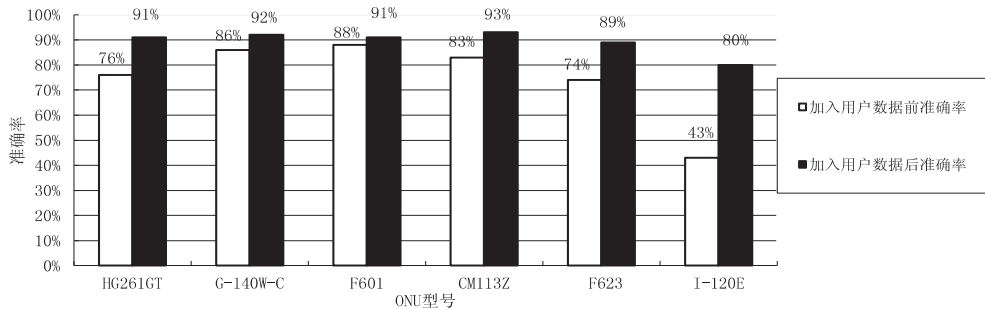


图 6 准确率对比

## 4 结束语

提出了使用深度学习技术识别 ONU 故障的设计和实现方法,巧妙地综合运用了检测和分类的思想,将核心模块分为 3 个阶段逐步求精实现,有效减少了环境的干扰,运用分类的方法对非 ONU 图片可以有效过滤,最终识别出不同 ONU 型号的不同场景。经分析,端到端准确率超过 84%,产品上线“广东移动手机营业厅”,据统计该功能的月均使用次数超过 1 万人次,不仅有效节约了客户等待时间,而且减少了装维上门成本,是电信行业提升客户服务品质的一次成功探索。

## 参考文献:

- [1] 迟延龙. 基于家庭用户的宽带网络接入技术研究[J]. 计算机光盘软件与应用, 2014, 17(2): 292-293.
- [2] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: single shot multibox detector [C]//European conference on computer vision. Amsterdam: Springer, 2016: 21-37.
- [3] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas, NV: IEEE, 2016: 779-788.
- [4] GIRSHICK R. Fast r-cnn [C]//Proceedings of the IEEE international conference on computer vision. [s. l.]: IEEE, 2015: 1440-1448.
- [5] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Boston, MA: IEEE,

灯在加入用户数据之前和加入用户数据之后的准确率对比,最终分类 ONU 指示灯的平均准确率是 87%。运用同样的方法,检测 ONU 准确率达到 98% (50% IOU)、分类 ONU 的准确率达到 99%,整体端到端准确率为三者乘积,达到 84.4%;硬件方面 GPU 使用了一块 NVIDIA Tesla P100 显卡,从大规模压力测试情况来看使用 Inception-Resnet-V2 与 Inception-V4 组合达到了精度和速度的平衡,实际识别速度可达 10 FPS。

2015:1-9.

- [6] SANTURKAR S, TSIPRAS D, ILYAS A, et al. How does batch normalization help optimization? [C]//Advances in neural information processing systems. CAN: NIPS, 2018: 2483-2493.
- [7] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. [s. l.]: IEEE, 2016: 2818-2826.
- [8] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning [C]//Thirty-first AAAI conference on artificial intelligence. [s. l.]: AAAI, 2017: 4278-4284.
- [9] ABADI M, BARHAM P, CHEN J, et al. Tensorflow: a system for large-scale machine learning [C]//12th USENIX symposium on operating systems design and implementation (OSDI 16). [s. l.]: USENIX, 2016: 265-283.
- [10] 侯淑英. B/S 模式和 C/S 模式优势比较 [J]. 沈阳教育学院学报, 2007, 9(2): 98-100.
- [11] 吕 曦, 王化文. Web Service 的架构与协议 [J]. 计算机应用, 2002, 22(12): 62-65.
- [12] 宗兴寰. 浅议计算机虚拟现实技术发展及应用 [J]. 现代信息技术, 2019, 3(4): 91-93.
- [13] MASI I, TRÂN A T, HASSNER T, et al. Face-specific data augmentation for unconstrained face recognition [J]. International Journal of Computer Vision, 2019, 127(6-7): 642-667.
- [14] 古 平, 杨 炀. 面向不均衡数据集少数类细分的过采样算法 [J]. 计算机工程, 2017, 43(2): 241-247.
- [15] 张 慧. 深度学习中优化算法的研究与改进 [D]. 北京: 北京邮电大学, 2018.