

基于 YOLOv3 网络的电能表示数识别方法^①



龚安, 张洋, 唐永红

(中国石油大学(华东), 青岛 266580)

通讯作者: 龚安, E-mail: 414625329@qq.com

摘要: 随着智能电网的不断发展, 基于数字图像处理方法的电能表自动抄表系统被广泛应用, 为提升传统电能表示数自动识别的准确率, 提出了一种基于 YOLOv3 (You Only Look Once) 网络的电能表示数识别新方法. 对于电能表图像, 构建基于 YOLOv3-Tiny 网络的计数器定位模型并训练, 使用训练完毕的模型定位计数器目标区域, 裁剪计数器区域生成计数器图像; 对于计数器图像, 构建基于 YOLOv3 网络的计数器识别模型并训练, 使用训练完毕的模型识别计数器目标区域的数字. 选择巴西巴拉那联邦大学公开的电能表数据集作为研究对象, 通过与 YOLOv2-Tiny 定位模型、CR-NET 识别模型的对比实验, 表明了本方法具有更高的定位准确率和识别准确率.

关键词: 电能表; YOLOv3 网络; 目标检测; 图像识别

引用格式: 龚安, 张洋, 唐永红. 基于 YOLOv3 网络的电能表示数识别方法. 计算机系统应用, 2020, 29(1): 196-202. <http://www.c-s-a.org.cn/1003-3254/7215.html>

Automatic Reading Method of Electric Energy Meter Based on YOLOv3

GONG An, ZHANG Yang, TANG Yong-Hong

(China University of Petroleum, Qingdao 266580, China)

Abstract: With the continuous development of smart grid, the automatic reading system of electric energy meter based on digital image processing method is widely used. To improve the accuracy of automatic reading of traditional electric energy meter, a new method of automatic reading of electric energy meter based on YOLOv3 network is proposed. For the electric energy meter image, a counter positioning model based on the YOLOv3-Tiny network is constructed and trained, the trained target model is used to locate the counter target area, and the counter area is generated to achieve a counter image. For the counter image, a counter recognition model based on the YOLOv3 network is constructed and trained, and the trained model is used to identify the number of the counter target area. The electric energy meter data set published by the Federal University of Paraná Brazil was selected as the research object. The comparison experiment with YOLOv2-Tiny positioning model and CR-NET recognition model shows that the proposed method has higher positioning accuracy and recognition accuracy.

Key words: electric energy meter; YOLOv3 network; target detection; image recognition

1 引言

在电力行业中, 传统电能表的示数信息采集大都采用人工抄表, 但人工抄表工作繁琐, 效率较低, 因此智能自动抄表系统的出现迎合了时代的需要. 近年来,

许多数字图像处理方法已经被应用于电能表示数自动识别^[1-5], 一类是根据人工设计的特征进行定位和识别的方法^[1-4], 如张宗健使用边缘检测方法定位计数器目标区域, 然后根据数字的笔画和形状特征进行数字识

① 基金项目: 国家科技重大专项 (2017ZX05013-001)

Foundation item: National Science and Technology Major Program (2017ZX05013-001)

收稿时间: 2019-05-30; 修改时间: 2019-06-28; 采用时间: 2019-07-05; csa 在线出版时间: 2019-12-27

别^[1]; 贝澄洁使用图像阈值处理方法定位计数器目标区域, 然后根据投影法和数字的笔画特征进行数字识别^[2]. 另一类是基于深度学习自动提取特征的定位和识别方法^[5], 如 Rayson Laroca 等使用 YOLOv2-Tiny^[6]网络定位计数器目标区域, 并进行了 CR-NET^[7]、CRNN^[8]等网络在数字识别部分的对比实验^[5]. 基于人工设计的特征进行目标检测存在检测准确率低、训练耗时长等缺点, 而基于深度学习的目标检测能够自动提取特征, 有效提升检测精度, 因此目前基于深度学习的目标检测得到了广泛应用^[9,10].

因为 YOLO 网络检测速度快、准确率较高, 所以为提升电能表示数识别的准确率, 本文提出一种基于 YOLOv3 网络的电能表示数识别新方法. 首先构建基于 YOLOv3-Tiny 网络的计数器定位模型定位计数器

区域, 然后构建基于 YOLOv3^[11]网络的计数器识别模型进行数字识别, 实验结果表明本方法对电能表示数的识别具有更高的准确率.

2 YOLO 网络的相关理论

本章主要介绍了 YOLO 网络检测的基本流程、网络训练过程和网络的结构单元.

2.1 YOLO 网络检测的基本流程

基于 YOLO 网络的检测方法将特征提取、目标定位、目标分类统一于一个卷积网络中, 将物体检测问题作为一个回归问题来解决, 将测试图像输入完成训练的网络便能得到目标物体的位置及其所属的类别及其相应的置信度. YOLO 网络检测的基本流程如图 1 所示.

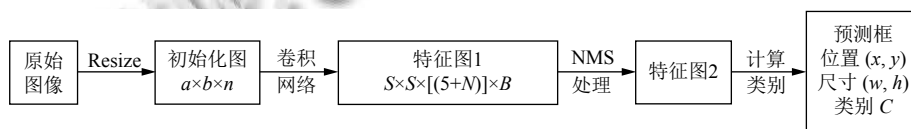


图 1 YOLO 网络检测流程

具体流程如下:

(1) 分辨率初始化: 将原始图像的分辨率初始化为 $a \times b$, 图像的通道数为 n .

(2) 卷积网络提取特征: 将 resize 后的图像数据输入卷积网络, 输出一个 $S \times S \times [(5+N) \times B]$ 的特征图, 该特征图将图像划分为 $S \times S$ 个网格单元, 每个网格单元输出 B 个预测框, 每个预测框将输出 4 种信息, 包括预测框的位置 (x, y) 、尺寸 (w, h) 、置信度 c 以及预测不同类别的条件概率 p_1, \dots, p_N , 其中类别的数量为 N .

预测框的置信度 c 用公式表示为:

$$c = p(\text{Object}) * IOU_{\text{pred}}^{\text{truth}} \quad (1)$$

$$IOU_{\text{pred}}^{\text{truth}} = \frac{\text{area}(t) \cap \text{area}(p)}{\text{area}(t) \cup \text{area}(p)} \quad (2)$$

其中, $p(\text{Object})$ 表示应的网格单元是否存在目标 (Object) 的中心点, 若存在, 那么 $p(\text{Object}) = 1$, 否则 $p(\text{Object}) = 0$; $IOU_{\text{pred}}^{\text{truth}}$ 表示预测框和真实框面积的交并比; $\text{area}(t)$ 表示真实框的面积, $\text{area}(p)$ 表示预测框的面积.

(3) NMS 处理删除冗余框: NMS (Non-Maximum Suppression) 即非极大值抑制算法, 对于预测框组成的集合 A , 采用下面的计算方式: 选择具有最大置信度的

预测框 M , 将其从 A 集合中移除并加入到最终的检测结果 D 中, 将 A 中剩余预测框中与 M 的 IOU 大于阈值 N_t 的框从 A 中移除, 重复这个过程, 直到 A 集合为空. 使用 NMS 算法将特征图 1 中的冗余框的置信度设置为 0, 生成特征图 2.

(4) 计算预测类别: 预测框对于第 i 类的置信度分数 s_i , 用公式表示为:

$$s_i = p_i * c \quad (3)$$

其中, p_i 表示当前预测框对应第 i 类的条件概率, c 表示当前预测框的置信度.

预测类别取 s_i 最大时对应的类别 i , 预测框对应的预测类别用公式表示为:

$$C = \arg \max (s_i) \quad (4)$$

2.2 YOLO 网络的训练过程

YOLO 网络的训练过程可分为 4 个步骤:

(1) 由 2.1 节中得到特征图 2 中每个网格单元每个预测框的 $x, y, w, h, c, p_1, \dots, p_N$.

(2) 根据损失函数计算误差.

YOLO 网络的损失函数=预测框和真实框的坐标尺寸误差+预测框和真实框的置信度误差+分类误差^[6,11,12].

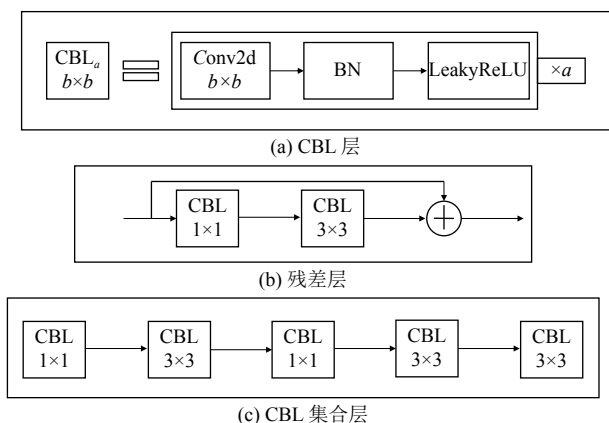


图2 网络基本结构单元

预测框与真实框的坐标尺寸误差 $loss1$ 用公式表示为:

$$loss1 = \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B l_{ij}^{obj} (2 - w_i \times h_i) \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + (w_i - \hat{w}_i)^2 + (h_i - \hat{h}_i)^2 \right] \quad (5)$$

其中, λ_{coord} 表示预测位置的系数, 该值设置为5; l_{ij}^{obj} 判断第*i*个网格单元中第*j*个预测框是否负责目标obj, 与目标obj的真实框IOU最大的预测框负责目标obj, 若负责预测, 那么该值为1, 否则为0; $(w_i, y_i), (w_i, h_i)$ 表示预测框位置和尺寸; $(\hat{x}_i, \hat{y}_i), (\hat{w}_i, \hat{h}_i)$ 表示真实框位置和尺寸.

预测框和真实框的置信度误差 $loss2$ 用公式表示为:

$$loss2 = \sum_{i=0}^{S^2} \sum_{j=0}^B l_{ij}^{obj} (c_i - \hat{c}_i)^2 + \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^B l_{ij}^{noobj} (c_i - \hat{c}_i)^2 \quad (6)$$

其中, λ_{noobj} 表示网格内没有目标时的系数, 该值设置为0.5; l_{ij}^{noobj} 判断第*i*个网格单元中第*j*个预测框是否不负责目标obj, 若不负责, 那么该值为1, 否则为0; c_i 表示预测框置信度; \hat{c}_i 表示真实框置信度.

类别条件概率误差 $loss3$ 用公式表示为:

$$loss3 = \sum_{i=0}^{S^2} \sum_{C \in classes} l_{ij}^{obj} (p_i(C) - \hat{p}_i(C))^2 \quad (7)$$

其中 $p_i(C)$ 表示第*i*个网格单元预测C类的条件概率, $\hat{p}_i(C)$ 第*i*个网格单元是C类的真实概率.

- (3) 根据误差计算卷积核中每个权重的梯度.
- (4) 应用基于梯度的优化算法更新权重.

2.3 网络基本结构单元

(1) CBL层: 指集合了卷积操作、BatchNormal^[13]

操作和 LeakyReLU 激活操作的网络层. 结构如图2(a)所示, Conv2d表示卷积操作, 卷积核数量为*a*, 大小为 $b \times b$, 该操作使用权重组卷积核 $\begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1b} \\ w_{21} & w_{22} & \cdots & w_{2b} \\ \cdots & \cdots & \cdots & \cdots \\ w_{b1} & w_{b2} & \cdots & w_{bb} \end{bmatrix}$ 在数字图像上滑动, 将图像点上的像素灰度值与对应的卷积核上的数值相乘, 然后将相乘后的值相加作为卷积核中间像素对应的图像上像素的灰度值, 具有稀疏交互、参数共享、等变表示等优点.

BN表示批标准化操作, 该操作将特征规整为均值为0、方差为1的正态分布. 可以加快收敛速度, 缓解训练过拟合问题, 增强模型泛化能力.

LeakyReLU是激活函数, 公式如下:

$$y_i = \begin{cases} x_i, & \text{if } x_i \geq 0 \\ \frac{x_i}{a_i}, & \text{if } x_i < 0 \end{cases} \quad (8)$$

其中, a_i 是 $(1, +\infty)$ 区间内的固定参数.

(2) 最大池化层 (Maxpool): 用某一位置的相邻矩形区域内的最大值来代替网络在该位置的输出, 该操作能提高计算速度和所提取特征的鲁棒性.

(3) 上采样层 (Upsample): 通过双线性插值方法实现上采样操作.

(4) 残差层^[14]: 将浅层网络特征和深层网络特征融合, 有效解决了深度网络训练过程的退化问题, 结构如图2(b)所示.

(5) CBL集合层: 指集合了5个CBL层的集合. 结构如图2(c)所示.

2.4 YOLOv3-Tiny 和 YOLOv3 网络

2.4.1 机制和性能

YOLOv3-Tiny网络以Darknet-Tiny网络为特征提取网络, 使用Darknet-Tiny网络前7个CBL层和6个池化层, 设计了两个不同尺度的特征图输出, 并且通过上采样将两个不同分辨率的特征图进行了融合; 通过聚类先验框尺寸, 对每个尺度的输出给定3个先验框.

YOLOv3网络以DarkNet53网络为特征提取网络, 使用DarkNet53网络前52个CBL层, 存在有23个残差层结构, 设计了3个不同尺度的特征图输出, 并且通过上采样将3个不同分辨率的特征图进行了融合; 通过聚类先验框尺寸, 对每个尺度的输出给定3个先验框.

在COCO图像数据集的目标检测任务中, 当真实

框和预测框的 IOU 阈值为 0.5 时, YOLOv3-Tiny 网络的 mAP 为 33.1, FPS 为 220 f/s; YOLOv3 网络的 mAP 为 55.3, FPS 为 35 f/s.

2.4.2 适用场景

根据 YOLOv3-Tiny 和 YOLOv3 网络的结构和在 COCO 数据集上的性能表现, 可以得出 YOLOv3-Tiny 速度较快, mAP 较低, 适用于图像分辨率高, 目标种类少的简单场景; YOLOv3 速度较慢, mAP 较高, 适用于图像分辨率低, 目标种类多的复杂场景.

3 定位和识别模型

为排除电能表图像中无关信息的干扰, 首先需要对计数器目标定位, 然后裁剪计数器目标区域, 最后进行计数器数字识别. 根据相关理论和电能表计数器图像的特点, 给出计数器定位和识别模型.

电能表图像较大, 且目标区域单一, 因此基于 YOLOv3-Tiny 网络构建计数器定位模型. 为了提高定位的准确率, 测试时将预测框的尺寸扩大为原来的 1.1 倍. 计数器图像较小, 目标种类相对较多, 因此基于 YOLOv3 网络构建计数器识别模型. 为了避免同一个数字位置出现多个不同类别的预测框, 测试时对 NMS 算法进行改进, 在 NMS 算法中添加规则: 如果存在任意两个预测框的 $IOU > 0.6$, 那么将保留置信度最高的预测框, 将其余预测框全部删除.

计数器定位和识别的流程如图 3 所示.

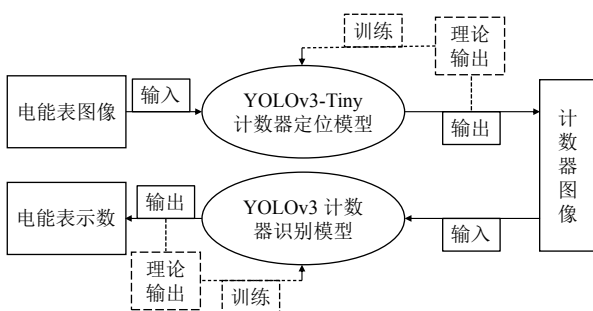


图 3 定位和识别的流程

针对电能表图像宽高比特征, 将输入的分辨率初始化为 416×416 ; 网络结构如图 4 所示.

针对计数器图像宽高比特征, 将输入的分辨率初始化为 416×128 ; 网络结构如图 5 所示.

综上, 计数器定位和识别的整体设计思路即使用基于 YOLOv3-Tiny 网络的单目标检测方法实现计数

器定位, 使用基于 YOLOv3 网络的多目标检测方法实现计数器识别. 定位和识别模型的速度取决于输入图片的大小, 网络的参数数量和结构; 准确率取决于输入图片的质量, 网络的训练是否充分等.

Type	Filters	Size	Input	Output
CBL	16	$3 \times 3 / 1$	$416 \times 416 \times 3$	$416 \times 416 \times 16$
Max		$2 \times 2 / 2$	$416 \times 416 \times 16$	$208 \times 208 \times 16$
CBL	32	$3 \times 3 / 1$	$208 \times 208 \times 16$	$208 \times 208 \times 32$
Max		$2 \times 2 / 2$	$208 \times 208 \times 32$	$104 \times 104 \times 32$
CBL	64	$3 \times 3 / 1$	$104 \times 104 \times 32$	$104 \times 104 \times 64$
Max		$2 \times 2 / 2$	$104 \times 104 \times 64$	$52 \times 52 \times 64$
CBL	128	$3 \times 3 / 1$	$52 \times 52 \times 64$	$52 \times 52 \times 128$
Max		$2 \times 2 / 2$	$52 \times 52 \times 128$	$26 \times 26 \times 128$
CBL	256	$3 \times 3 / 1$	$26 \times 26 \times 128$	$26 \times 26 \times 256$
Max		$2 \times 2 / 2$	$26 \times 26 \times 256$	$13 \times 13 \times 256$
CBL	512	$3 \times 3 / 1$	$13 \times 13 \times 256$	$13 \times 13 \times 512$
Max		$2 \times 2 / 2$	$13 \times 13 \times 512$	$13 \times 13 \times 512$
CBL	1024	$3 \times 3 / 1$	$13 \times 13 \times 512$	$13 \times 13 \times 1024$

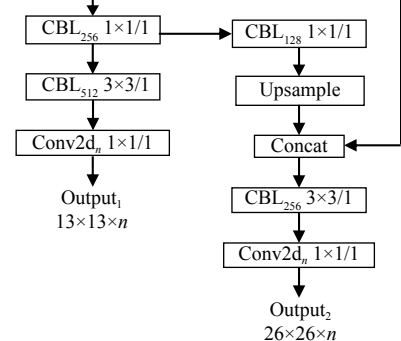


图 4 YOLOv3-Tiny 网络结构

4 实验分析

4.1 实验环境

实验的计算机配置和环境如下: CPU 处理器为 Intel(R) Core(TM) i7-6700K CPU @ 4.00 GHz \times 8; GPU 处理器为 NVIDIA GeForce GTX TITAN X; 内存 (RAM) 为 16.00 GB; 操作系统为 64 bit Windows10 专业版; 编程语言为 C++; 集成开发环境为 Visual Studio2017; 网络模型通过基于 DarkNet 的深度学习框架实现.

4.2 实验数据

计数器定位数据: 本文所用数据集共含有 2000 张图片和对应的标签文件, 选取数据集中的 800 张电能表图片作为训练集; 从剩余 1200 张中选取 100 张图片作为验证集; 从剩余的 1100 张中选取 100 张图片作为测试集.

计数器识别数据: 将数据集中的原始图片按计数器位置标签裁剪生成计数器图片, 选取裁剪生成 800 张计数器图片作为训练集. 从剩余 1200 张中选取

400 张计数器图片作为验证集. 将剩余的 800 张计数器图片作为测试集.

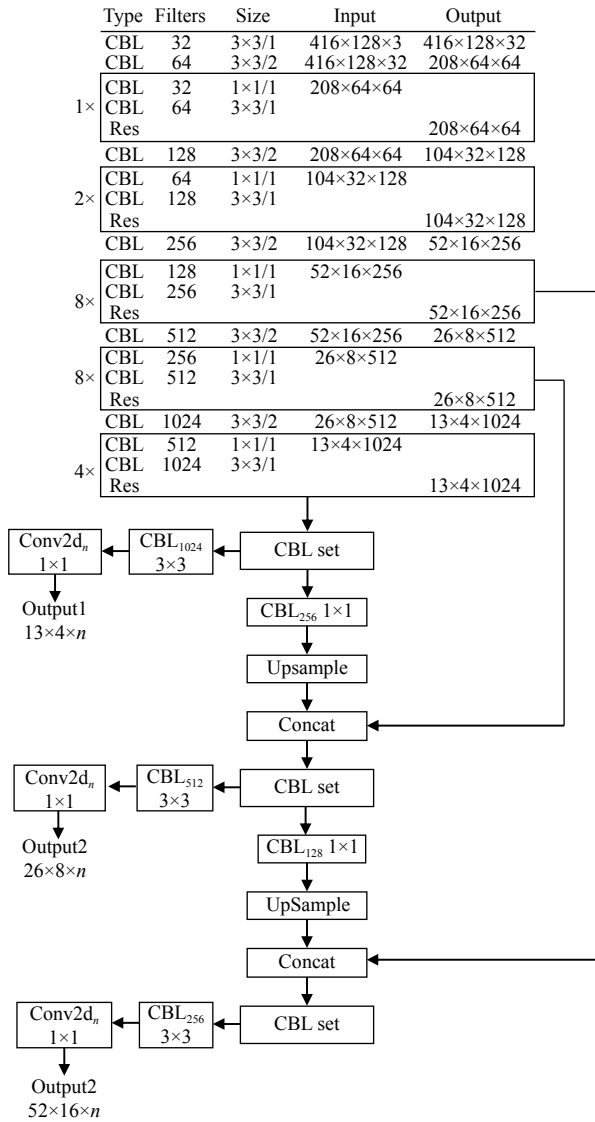


图5 YOLOv3 网络结构

4.3 评价指标

4.3.1 定位模型的的评价指标

评分 S 用公式表示为:

$$S = S_1 + S_2 \times 100 \quad (9)$$

$$S_1 = \text{mean} \left(\sum IOU_{\text{Pred}}^{\text{Truth}} \right) \quad (10)$$

$$S_2 = \frac{N(\text{True})}{N(\text{Total})} \quad (11)$$

其中, S_1 表示预测框和真实框的交并比的平均值, S_2 表示当 IOU 阈值=0.5 时每张图像的定位计数器目标的准

确率, $N(\text{True})$ 表示定位准确的数量, $N(\text{Total})$ 表示在总目标数量, 计数器检测中的验证集和测试集中的 $N(\text{Total})$ 均为 100, 评分 S 用于选择最优权重.

4.3.2 识别模型的评价指标

单个数字识别的准确率公式如下:

$$F_1 = \frac{N_1(\text{True})}{N_1(\text{Total})} \times 100\% \quad (12)$$

其中, $N_1(\text{True})$ 表示预测识别的数字数量, $N_1(\text{Total})$ 表示总数字数量, 验证集的 $N_1(\text{Total}) = 400 \times 5 = 2000$, 测试集的 $N_1(\text{Total}) = 800 \times 5 = 4000$.

整个计数器识别的准确率公式如下:

$$F_2 = \frac{N_2(\text{True})}{N_2(\text{Total})} \times 100\% \quad (13)$$

其中, $N_2(\text{True})$ 表示正确识别的计数器数量, $N_2(\text{Total})$ 表示计数器的数量, 验证集的 $N_2(\text{Total}) = 400$, 测试集的 $N_2(\text{Total}) = 800$.

4.4 实验过程

4.4.1 定位模型的实验过程

(1) 配置训练参数: 图像尺寸初始化为 $416 \times 416 \times 3$; 采用随机梯度下降法 (SGD), 每次迭代使用 1 个样本对参数进行更新; 优化方案采用动量法, 参数值=0.9; 权重衰减正则项值=0.0005; 学习率=0.001; 迭代次数共 50 000 次; 网络需要识别的物体种类数为 1, 因此最后一层卷积核数量 $n = (1 + 5) \times 3 = 18$; 开启多尺度训练, 每经 10 批次训练强制改变输入图像大小进行训练; 使用 K-means 聚类方法对训练集真实框的尺寸聚类获得先验框尺寸, 配置先验框尺寸参数.

(2) 训练网络: 根据 2.2 节的训练过程训练网络, 每迭代 1000 次保存权重文件.

(3) 验证网络: 设置 IOU 阈值=0.5, 计算不同权重文件对验证集验证的评分 S , 选取最优权重 W .

(4) 测试网络: 选择验证步骤所得的最优权重 W 对测试集进行测试, 获取电能表图像中计数器的位置 $[x, y, w, h]$.

4.4.2 识别模型的实验过程

(1) 配置训练参数: 图像尺寸初始化为 $416 \times 128 \times 3$; 采用小批量梯度下降法 (MBGD), 其中 $\text{batch}=64$, $\text{sub}=4$, 每次迭代使用 64 个样本对参数进行更新; 优化方案采用动量法, 参数值=0.9; 权重衰减正则项值=0.0005; 学习率=0.001; 迭代次数共 20 000 次; 网络需要识别的物体种类数为 10, 因此最后一层卷积核数量

$n = (10 + 5) \times 3 = 45$; 使用 K-means 聚类方法对训练集真实框的尺寸聚类获得先验框尺寸, 配置先验框尺寸参数。

(2) 训练网络: 根据 2.2 节的训练过程训练网络, 每迭代 1000 次保存权重文件。

(3) 验证网络: 设置 IOU 阈值=0.6, 计算不同权重文件对验证集验证所得的单个数字识别准确率 F_1 和整个计数器示数识别准确率 F_2 , 用于选择最优权重 W 。

(4) 测试网络: 选择验证步骤所得的最优权重 W 对测试集进行测试, 获取计数器图像中的示数 $[a_1, a_2, a_3, a_4, a_5]$ 以及每个数字 a_i 的位置 $[x_i, y_i, w_i, h_i]$ 。

4.5 实验结果

4.5.1 计数器定位的实验结果

将基于 YOLOv3-Tiny 网络的定位模型与基于 YOLOv2-Tiny 网络的定位模型进行对比实验。

(1) 验证结果: 选取不同迭代次数对应的权重文件, 对验证集进行验证的结果如图 6 所示。

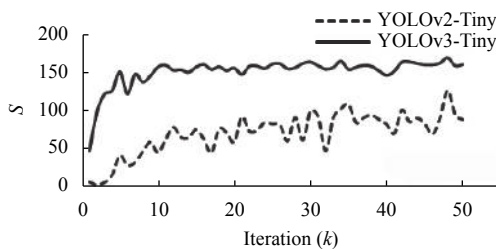


图 6 YOLOv2-Tiny 与 YOLOv3-Tiny 验证集对比结果

(2) 测试结果: 根据验证集结果, YOLOv3-Tiny 和 YOLOv2-Tiny 网络均在迭代次数 48 000 处对应权重验证评分较高, 因此分别选择迭代次数 48 000 对应的权重进行测试, 对测试集进行测试的结果如表 1 所示。

由表 1 可得, 相比 YOLOv2-Tiny 网络, 基于 YOLOv3-Tiny 网络的计数器定位模型的精度更高, 当 IOU 阈值为 0.5 时, 预测准确率达到 95%。将预测框的尺寸扩大为原来的 1.1 倍, 定位准确率达到 100%。

4.5.2 计数器识别的实验结果

YOLO 网络存在许多变体, 其中较为成功的是

CR-NET (Character Recognition NETWORK) 字符识别网络, 该网络将 YOLOv2-Tiny 网络中的部分池化层改为卷积层, 结构简单, 参数量少, 本文将基于 YOLOv3 网络的识别模型与基于 CR-NET 网络的识别模型进行对比实验。

表 1 YOLOv2-Tiny 与 YOLOv3-Tiny 测试集对比结果

网络类型	迭代次数对应的最优权重	计数器定位准确率 (%) (IOU 阈值=0.5)	FPS (f/s)
YOLOv2-Tiny	48k_weight	78	14.3
YOLOv3-Tiny	48k_weight	95	14.3

(1) 验证结果: 选取不同迭代次数对应的权重文件, 对验证集进行验证的结果如图 7、图 8 所示。

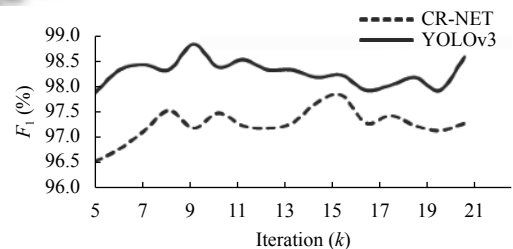


图 7 对单个数字识别的准确率验证集对比结果

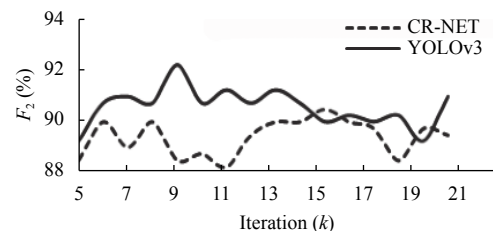


图 8 对整个计数器示数识别准确率验证集对比结果

(2) 测试结果: 根据验证集结果, CR-NET、YOLOv3 网络分别在迭代次数 15 000、9000 时对应的权重验证准确率较高, 因此选择上述两个权重进行测试, IOU 阈值设置为 0.6, 对测试集进行测试的结果如表 2 所示。由表 2 可得, 修改识别模型的 NMS 算法后提高了计数器整体识别的准确率; 基于 YOLOv3 网络的计数器识别模型的准确率达到 92.13%, 相比 CR-NET 网络, 准确率提升 2.63%。

表 2 两个网络的测试集对比结果

网络类型	迭代次数对应的最优权重	正确识别数字的数量	单个数字识别准确率 (%)	未修改 NMS 计数器识别准确率 (%)	修改 NMS 后计数器识别准确率 (%)	FPS (f/s)
CR-NET	15k_weight	3899	97.48	88.25	89.50	200
YOLOv3	9k_weight	3938	98.45	91.25	92.13	67

5 效果展示

5.1 计数器定位

定位效果展示如图9所示。

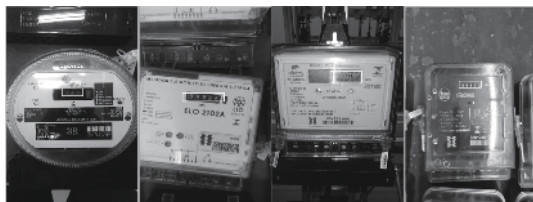


图9 计数器定位图片展示

5.2 计数器识别

(1) 正确识别的示例如图10所示。



图10 正确识别图片

(2) 错误识别的示例如图11所示。主要原因包括：像素点缺失，电表处于跳转状态，图片模糊等。

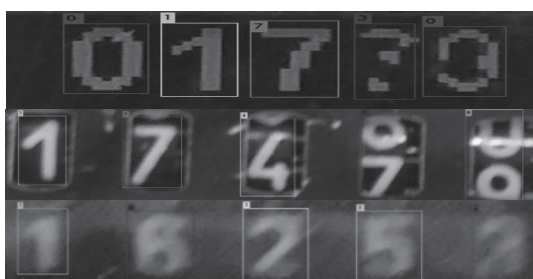


图11 错误识别图片

6 结束语

本文为提高电能表自动识别准确率，构建了基于YOLOv3-Tiny网络的计数器定位模型和基于YOLOv3网络的计数器识别模型。实验表明基于YOLO网络的模型对数字式电表的自动识别适用性较高，后续将研究基于YOLO网络的模型对指针式电表示数识别的问题。

参考文献

1 张宗健. 基于数字图像识别的字轮电能表读数系统研究

[硕士学位论文]. 北京: 北京交通大学, 2010.

2 贝澄洁. 单相电表图像识别技术研究[硕士学位论文]. 南京: 南京理工大学, 2014.

3 张晓颖. 电能表数字图像识别技术与实现[硕士学位论文]. 上海: 上海交通大学, 2007.

4 凌璐. 基于图像处理的电能表示数识别技术. 中国电机工程学会电力系统自动化专业委员会2012年学术交流会议论文集. 厦门, 中国. 2012. 1-5.

5 Laroca R, Barroso V, Diniz MA, *et al.* Convolutional neural networks for automatic meter reading. *Journal of Electronic Imaging*, 2019, 28(1): 013023.

6 Redmon J, Farhadi A. YOLO9000: Better, faster, stronger. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI, USA. 2017. 6517-6525.

7 Silva SM, Jung CR. Real-time brazilian license plate detection and recognition using deep convolutional neural networks. *Proceedings of 2017 30th SIBGRAPI Conference on Graphics, Patterns and Images*. Niteroi, Brazil. 2017. 55-62.

8 Shi BG, Bai X, Yao C. An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(11): 2298-2304. [doi: 10.1109/TPAMI.2016.2646371]

9 王潇天. 基于深度学习的目标检测研究与应用. *电子制作*, 2018, (22): 40-41, 86. [doi: 10.3969/j.issn.1006-5059.2018.22.018]

10 吴帅, 徐勇, 赵东宁. 基于深度卷积神经网络的目标检测综述. *模式识别与人工智能*, 2018, 31(4): 335-346.

11 Redmon J, Farhadi A. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018.

12 Redmon J, Divvala S, Girshick R, *et al.* You only look once: Unified, real-time object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA. 2016. 779-788.

13 Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015.

14 He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA. 2016. 770-778.