

DOI:10.19651/j.cnki.emt.2107837

基于 CNN 的水表指针读数识别及 STM32 实现方案设计

张鹏飞 叶哲江 杨嘉林 李家成

(昆明理工大学 信息工程与自动化学院 昆明 650500)

摘要: 为了提高卷积神经网络对于水表指针读数识别的准确率,同时实现将卷积神经网络移植到 STM32 单片机中运行,使用了包含 2 913 张水表指针图片的数据集对 GoogLeNet 和 ResNet-18 进行迁移学习和测试,其中 GoogLeNet 的测试集准确率为 89.37%,ResNet-18 的测试集准确率为 93.24%。借鉴于 ResNet-18 模型的跳跃连接思想,使用了高低层特征融合的方法,在保证感受野大小不变的前提下将 7×7 大卷积核替换为 3 个 3×3 小卷积核的串接以减少网络的参数量,同时减低网络的深度,加快了训练时网络的收敛,之后设计了一个对于水表指针读数识别准确率更高和收敛更快的卷积神经网络模型,此模型的测试集准确率为 95.11%。为克服 STM32 单片机存储资源极其有限的困难,在保证较高准确率的前提下进一步减小网络规模从而降低网络参数量,设计出模型的测试集准确率为 91.51%,训练过程在 PC 端使用 MATLAB 深度学习工具箱完成,生成的 onnx 模型仅有 948 KB 大小,运行占用 RAM 大小为 437.14 KB。

关键词: 卷积神经网络;STM32 单片机;GoogLeNet;ResNet-18;特征融合;感受野

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 520.2060

Recognition of water meter pointer reading based on CNN and design of STM32 implementation scheme

Zhang Pengfei Ye Zhejiang Yang Jialin Li Jiacheng

(Department of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: In order to improve the accuracy of convolutional neural network for water meter pointer reading recognition and realize the operation of convolutional neural network transplanted into STM32 microcontroller, A data set containing 2 913 water meter pointer pictures was used for transfer learning and testing of GoogLeNet and ResNet-18. The accuracy of GoogLeNet test set was 89.37%, and that of ResNet-18 test set was 93.24%. Based on the jumping connection idea of ResNet-18 model, the method of feature fusion of high and low levels is used. On the premise that the size of receptive field remains unchanged, the 7×7 large convolution kernel is replaced by three 3×3 small convolution kernels in series to reduce the number of network parameters, reduce the depth of the network, and speed up the convergence of the network during training. Then, a convolutional neural network model with higher accuracy and faster convergence for water meter pointer reading is proposed. The test set accuracy of this model is 95.11%. In order to overcome the difficulty of extremely limited storage resources of STM32 microcontroller and further reduce the network size and the number of network parameters on the condition of ensuring high accuracy, the test set accuracy of the designed model is 91.51%. The training process is completed using MATLAB deep learning toolbox on PC, and the generated onnx model is only 948 KB in size. The running footprint of RAM is 437.14 KB.

Keywords: convolutional neural network;STM32 SCM;GoogLeNet;ResNet-18;feature fusion;receptive field

0 引言

近几年来,智能水表开始普及,极大地方便了自来水供

应商的抄表和用水量统计工作,节省了大量的人力和财力,但是,一些老旧小区、偏远地区仍然使用着旧式的水表,所以这些地区的水表抄表工作依旧存在着不便,如果更换智能

水表,将花费高昂费用,且工作量大,因此,本文的目的是仅在旧式水表的基础上设计一种小型自动抄表系统。

卷积神经网络(convolutional neural network,CNN)是一种对于图像识别和分类有着良好效果的深度学习算法,广泛应用于人脸识别^[1]、机器故障检测^[2]、噪声检测^[3]、医学图像识别^[4]、指针式仪表读数识别^[5],将 CNN 应用于仪器仪表盘的读数识别具有较高的准确率。常见水表的读数方式通常有两种,一是直接读取水表界面的数字,二是通过水表界面的指针进行读数,常用的机器识别水表读数的算法可分为两类,一类是基于图像分析处理的算法^[6-8],另一类则是 CNN 算法,类似于车牌识别^[9],CNN 对水表的数字识别有着很好的效果^[10]。随着 CNN 算法的快速发展,AlexNet^[11]、VGGNet^[12]、GoogLeNet 和 ResNet 等经典模型相继被提出,基于改进的 VGG-16 迁移学习在水表的数字读数识别有着较高的准确率^[13]。

旧式水表一般具有多个表盘,读取多个表盘指针读数乘上相应倍率再相加即可得到当前用水量。考虑到水表可能安置在潮湿的环境,采集表盘界面的图像受水雾和污渍等噪声的干扰严重,使用传统图像分析处理的方法识别水表读数需要使用复杂的去噪算法,对于 CNN,在训练网络时,在数据集中引入添加各种噪声的图片可增强网络对噪声干扰图片的识别能力。

STM32 单片机具有功耗低、性能好、体积小、成本低的众多优点,将 CNN 算法移植到 STM32 单片机用于水表指针读数的识别有一定的实际应用意义。由于 STM32 单片机内存资源极其有限,为解决神经网络运行时内存不足的问题,人为扩展外部存储资源是其中一种解决方法^[14],本文在缩减 CNN 参数、调整神经网络结构上做出改进,使其仅靠片内存储资源就可运行,降低了成本,同时也保证了水表指针读数识别的较高准确率。

1 方法与实验

1.1 数据集

通过摄像头采集了无噪声干扰的图片 100 张,轻度水雾和污渍干扰的图片 208 张,重度水雾和污渍干扰的图片 195 张,经过人为添加高斯噪声、椒盐噪声、斑点噪声、泊松噪声及中值滤波、旋转操作后,将数据集扩充到了 3 884 张图片,图片像素为 120×160 。原始图片及各类预处理后的图片如图 1 所示。

针对于单片机内部存储资源有限的情况,将数据集图片统一转化为灰度图用于训练 STM32 单片机可运行的卷积神经网络模型,可减少运行时计算量,节省更多的 RAM 空间。

表盘圆弧按照刻度分为 10 等份,根据指针针尖所在的区域进行分类,以指针逆时针方向最近的一个读数为标签,数据集分类情况如表 1 所示。将数据集按照 7:1:2 的比例划分为训练集、验证集和测试集。

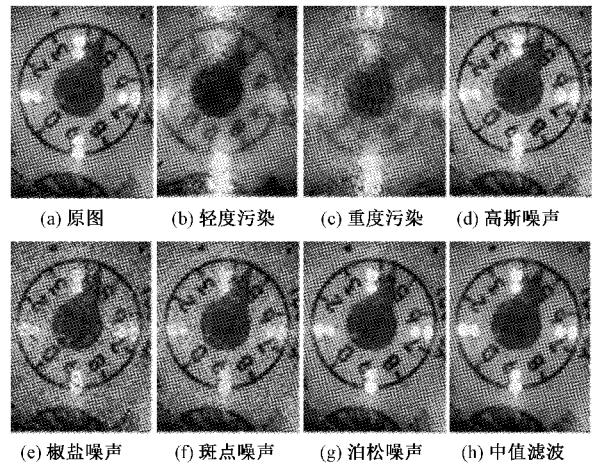


图 1 数据集图片

表 1 数据集分类情况

类别	0	1	2	3	4	5	6	7	8	9
数量/张	368	368	396	340	424	384	412	424	393	375

1.2 GoogLeNet 模型

GoogLeNet 是一种经典的卷积神经网络模型^[15],凭借其在 2014 年 ILSVRC (ImageNet large scale visual recognition challenge) 竞赛中取得的优异成绩,引起了计算机视觉领域众多学者的关注。GoogLeNet 在结构上添加了具有创新意义的 Inception 模块,这是一种尝试增加卷积神经网络宽度来提高网络性能的一种方法,Inception v1 模块如图 2 所示,其内部使用了 3 种不同的大小的卷积核,意味着可以获得不同大小的感受野,其中 1×1 卷积核可以起到降维和减少网络参数数量的作用,可实现跨通道信息交互,其参数量只有先前 AlexNet 参数量的 1/12。GoogLeNet 结构中使用了 9 个 Inception v1 模块串联,其中的两个 Inception v1 模块引入了辅助分类器防止梯度消失,有效地利用了中间层的识别能力。本文在实验中使用了 GoogLeNet 模型对水表的表盘读数进行了识别,取得了较好的效果。

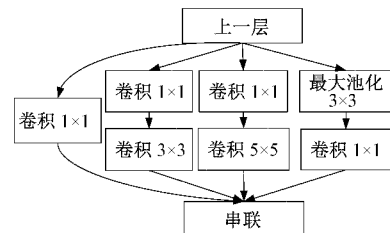


图 2 Inception v1 模块

1.3 ResNet-18 模型

ResNet 是 2015 年 ILSVRC 竞赛的第 1 名^[16],ResNet 提出了残差学习单元,这是一种通过跳跃连接再相加的特征融合方式,可以有效地利用高层特征和低层特征的各自

信息。同时残差学习单元很好地缓解深度网络的梯度消失现象,加快了网络的收敛^[17]。残差学习单元如图3所示。

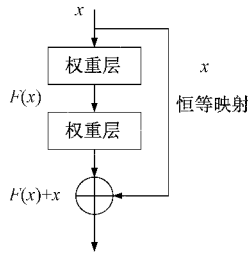


图3 残差学习单元结构

残差学习单元的运算过程如下:

$$y_l = h(x_l) + F(x_l, w_l) \quad (1)$$

式中: y_l 为第 l 个残差块的输出特征, x_l 为第 l 个残差块的输入特征, w_l 为第 l 个残差块的权重及偏置参数, h 为恒等映射, F 为残差函数。

$$x_{l+1} = f(y_l) \quad (2)$$

式中: f 为ReLU激活函数。由于 h 为恒等映射, $h(x_l) = x_l$,由式(1)和(2)可得出式(3):

$$x_{l+1} = x_l + F(x_l, w_l) \quad (3)$$

通过迭代,第 L 个残差块输入特征 x_L 等于第 l 个残差块输入特征加上第 l 个到第 $L-1$ 个残差的累加。

$$x_L = x_l + \sum_{i=l}^{L-1} F(x_i, w_i) \quad (4)$$

根据链式求导法则,残差网络的反向求导过程为:

$$\frac{\partial \epsilon}{\partial x_l} = \frac{\partial \epsilon}{\partial x_L} \frac{\partial x_L}{\partial x_l} = \frac{\partial \epsilon}{\partial x_L} \left[1 + \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, w_i) \right] \quad (5)$$

式中: ϵ 为损失函数,可以看出其中小括号中的1是前层残差块短路机制的体现,可以保证一部分梯度不会经过任何的衰减,另一部分的梯度也由链式求导法则的累乘变成累加,这样就有了更好的稳定性,可以有效地缓解梯度消失和梯度爆炸的问题,在一定程度上为更深层网络的搭建提供了保证。

1.4 本文模型

借鉴于ResNet-18的结构,在本文模型中引入支路连接,实现高低层网络特征实现特征融合。由于本文数据集中图片的语义特征很明显,更容易通过卷积运算提取到图片特征,同时由于本文的数据集样本数量有限,对于过深或者更宽的网络可能会造成过拟合的结果,从而导致网络的泛化能力下降,可能会严重影响网络的训练,对于本文数据集,可以尝试适当的减少网络的层数,随后的实验结果表明确实有效。为扩展数据集将图片旋转后生成的图片尺寸不一,在训练网络时,将图片的像素统一调整为 165×165 ,所设计的卷积神经网络模型的输入图片尺寸为 $165 \times 165 \times 3$ 大小,本文模型结构如图4所示。

输入为 $165 \times 165 \times 3$ 的RGB图像,所有卷积核及池化层的Padding设置为Same,其中批归一化(batch

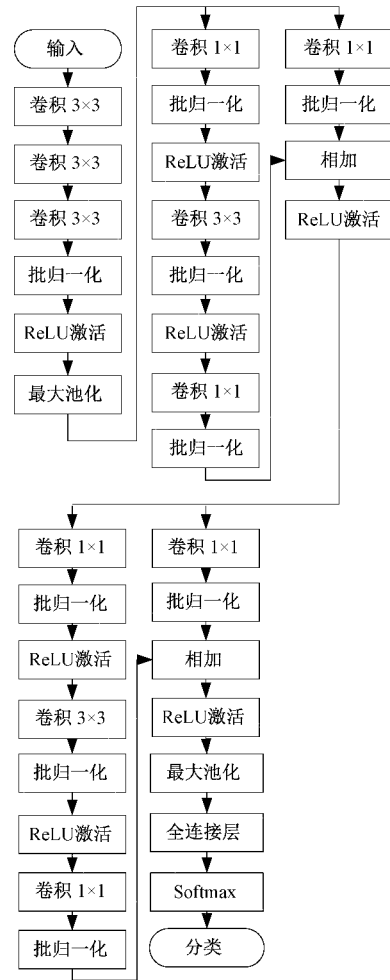


图4 本文模型结构

normalization, BN)操作可以有效减少内部协变量偏移(internal covariate shift, ICS)从而加速网络的训练^[18]。本文模型中,使用了3层串联的 3×3 小卷积核代替了ResNet-18中 7×7 的大卷积核,根据感受野计算公式,3层串联的 3×3 卷积核与1个 7×7 卷积核的感受野是相同的,但在输入为RGB图像并且每一个卷积层的卷积核个数相同时,3层串联的 3×3 卷积核参数数量为: $3 \times 3 \times 3 \times 3 \times N = 81N$,1层 7×7 卷积核的参数数量为: $7 \times 7 \times 3 \times N = 147N$, N 为每一层的卷积核的个数,前者的参数量相比于后者可减少45%,可以节省大量的计算资源。根据感受野计算公式:

$$r_l = r_{l-1} + [(k_l - 1) \times \prod_{i=1}^{l-1} s_i] \quad (6)$$

式中: r_l 为第 l 层的感受野大小, k_l 为第 l 层卷积核尺寸, s 为对应卷积核的步长。在本文模型中,第1层卷积核的步长为2,第2层和第3层的卷积核步长为1,根据式(6),3层 3×3 卷积核串联后的感受野大小为13;在ResNet-18中, 7×7 卷积核的步长为2,其感受野大小也为: $1 + [(7-1) \times 2] = 13$ 。同时,在ResNet-18结构上出相

应改动,使用了 $1 \times 1, 3 \times 3$ 和 1×1 卷积核代替了 $3 \times 3, 3 \times 3$ 卷积核,替换如图 5 所示。

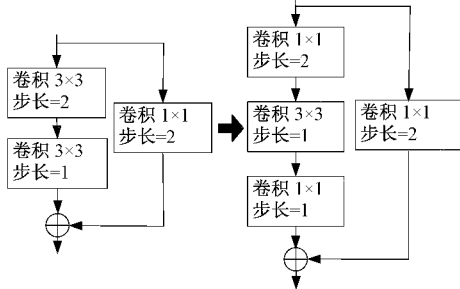


图 5 卷积核替换

图 5 中省略了批归一化层和 ReLU 非线性函数激活层,事实上这样替换可以减少参数量,节省计算资源,对于后面训练网络加快收敛有较好的效果,对于测试集的分类准确率也有所提高。

在最后的全连接层之前使用了最大池化层替换平均池化层。最大池化能够更好地提取具有强烈语义信息的特征,平均池化有利于提取全局特征,对于指针和表盘这类图像,仅指针位置和表盘数字的像素值才具有强烈的语义特征,最大池化更有利于提取到这些纹理特征。所以在网络模型中更多的使用最大池化更有利于提高网络的识别准确率。在实验过程中发现,最后一个池化层使用最大池化的效果的确优于平均池化。

在损失函数上,本文模型采用了分类问题最具效果的交叉熵,其表达式如式(7)所示。

$$\epsilon = -\frac{1}{X} \sum_{i=1}^Y \sum_{c=1}^Y y_{ic} \log(p_{ic}) \quad (7)$$

式中: X 为样本数量, Y 为分类类数,本文中 X 取 10, y_{ic} 为符号函数,取 0 或 1,如果样本 i 的真实分类等于 c ,则取 1,反之取 0。 p_{ic} 为样本 i 属于类别 c 的预测概率。在实际训练过程中,将输出映射到 $0 \sim 1$ 之间,对于每一个样本的 Softmax 层输出的是一个数值位于 $0 \sim 1$ 之间的向量,代表对应预测出每一个标签的概率。交叉熵的值越小,表示预测结果与实际标签越接近。

1.5 STM32 可运行模型及硬件设计

STM32 系列单片机作为一种价格低廉、体积小、功耗低的微控制器,被广泛应用于各种电子产品。本文所采用的单片机型号为 STM32F767ZI,该型号单片机具有 512 KB 的片内 RAM 和 2 048 KB 的 Flash。

使用 MATLAB 深度学习工具箱将训练好的卷积神经网络模型生成 .onnx 文件,使用 STM32 CubeMX 中的 Cube AI 软件包将 onnx 模型转化成 C 语言代码^[19],STM32 单片机实现卷积神经网络算法的流程如图 6 所示。

由于所采用的单片机仅有 512 KB 的 RAM,必须考虑 CNN 模型运算时产生的临时数据量不可超过单片机的 RAM 大小,在保证较高识别准确率的前提下,必须减少卷

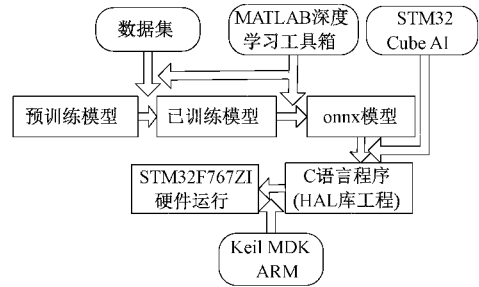


图 6 STM32 实现 CNN 流程

积核的通道数从而减少临时特征数据所占 RAM 的大小,同时减少卷积神经网络模型的层数从而减少整个网络的参数量,整个模型的大小不可超过单片机的 ROM,将所有数据集转化成灰度图,将网络输入层的尺寸改为 $165 \times 165 \times 1$,可以降低模型在单片机上的运算复杂度,减少单片机的运算负担,针对于单片机,将本文模型进行简化,设计出的卷积神经网络模型如图 7 所示。图 7 中 s 为卷积核滑动的步长, n 为特征图通道数,为减少计算量,故将 n 进行缩减。在结构上依旧保留了本文模型中 $1 \times 1, 3 \times 3$ 和 1×1 三个卷积核串联的方式,虚线框的卷积层或池化层的 Padding 为 Same,其余卷积层和池化层的 Padding 为 0。

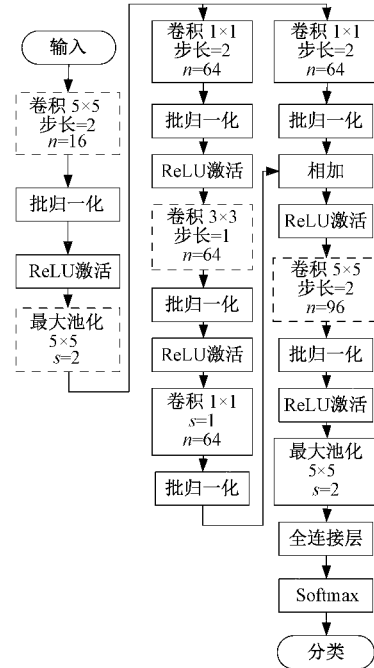


图 7 STM32 运行模型

在 PC 端将上述模型训练完成并生成的 onnx 模型仅有 948 KB 大小,使用 STM32 Cube AI 工具分析网络模型是否可运行,此模型占用 Flash 资源 937.39 KB,占用 RAM 资源 437.14 KB,可以运行在 STM32F767ZI 单片机上,分析结果如图 8 所示。

硬件系统如图 9 所示,整个硬件装置可直接固定在水表界面上方,将摄像头置于表盘指针上方,调整好距离和角

复杂度: 41329566 MACC
Flash 占用: 937.93KiB (2.00MiB)
RAM 占用: 437.14KiB (512KiB)
实现压缩率: -

图8 Cube AI 分析结果

度。STM32 单片机可控制 GC0308 摄像头模块定时采集一张水表指针图片,可将 GC0308 摄像头设置为 YUV 格式,调整输出图片尺寸为 160×120 ,直接可提取 Y 分量获得水表指针图的灰度图像,非采集图像时间段摄像头模块可处于休眠状态以降低功耗,闪光灯仅为摄像头采集图像时提供光源,摄像头模块所需时钟信号可由单片机内部时钟分频得到。

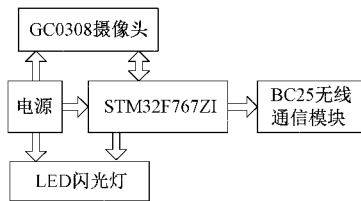


图9 硬件系统结构

图像采集完成后通过直接存储器访问(direct memory access, DMA)缓存到单片机内部,将所得到的灰度图片矩阵扩充到 165×165 大小,随后调用卷积神经网络程序对图片进行读数识别,再通过 BC25 无线通信模块即可向外界发送用水量信息,BC25 是一款高性能、低功耗的 NB-IoT (narrow band internet of things) 无线通信模块,支持国内运营商物联网平台。电源采用电池供电即可保证系统的长时间低功耗状态运行,此硬件设计方案可完成水表远程抄表系统的全部功能^[19]。

1.6 模型训练过程

使用 MATLAB 深度学习工具箱进行卷积神经网络的训练,针对本文问题依次对 GoogLeNet、ResNet-18 及本文模型进行测试。使用附加动量的随机梯度下降法(stochastic gradient descent with momentum, SGDM)更新网络参数,设置小批量训练,批量大小为 15,训练 20 轮,训练集共迭代 2 720 次,每 100 次迭代进行一次验证集验证,GoogLeNet、ResNet-18、本文模型和 STM32 可运行模型训练过程如图 10~13 所示,训练完成后,进行测试集的测试。

由图中各模型的训练过程可知,本文模型相比较于 GoogLeNet 和 ResNet-18 其收敛速度更快,在接近 1 500 次迭代后,其验证集的准确率稳定在 97% 左右,验证集的损失趋近于 0。

2 结果

混淆矩阵列出了测试集各类别在模型上的分类情况,其中对角线的元素值越大表明此分类模型的性能越好,由

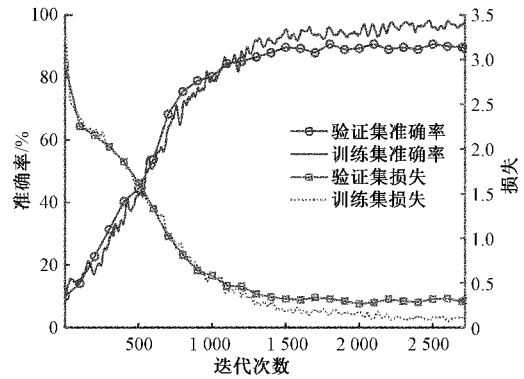


图10 GoogLeNet 训练过程

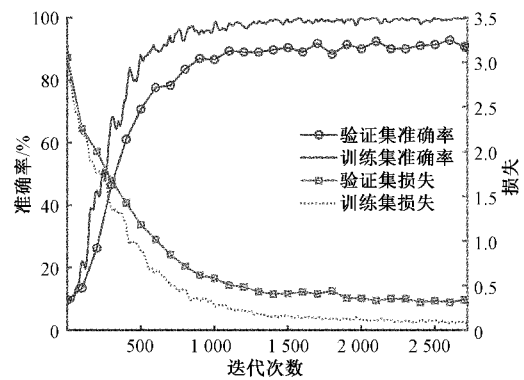


图11 ResNet-18 训练过程

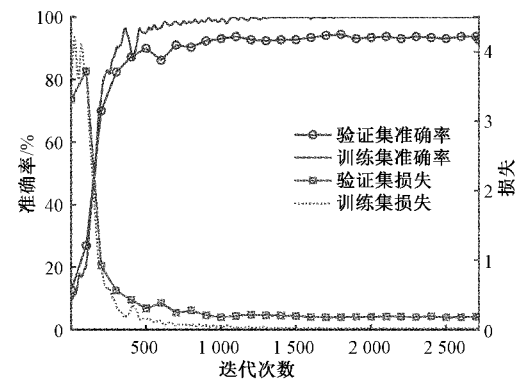


图12 本文模型训练过程

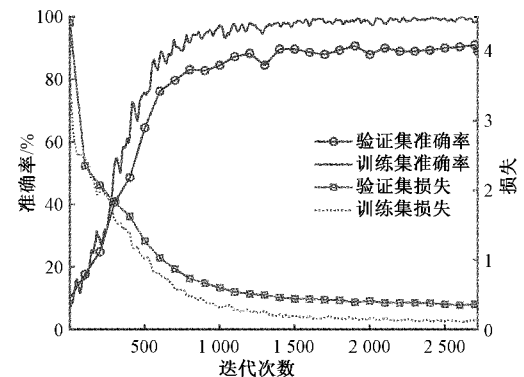


图13 STM32 可运行模型训练过程

各模型的混淆矩阵可以看出,多数分类错误集中在对角线附近,表明各模型对于指针位于刻度值边界的样本分类存在较大误差,这其中原因更大程度是人为给数据集添加标签时,对于指针处于刻度值边界的样本分类所属存在较大争议,这在一定程度上影响了模型训练后的分类性能。各模型混淆矩阵如图 14~17 所示。对比于各模型的混淆矩阵,本文模型混淆矩阵的元素值大小更集中于对角线,所以本文模型的识别效果优于 GoogLeNet 和 ResNet-18,各模型的测试集识别准确率如表 2 所示。

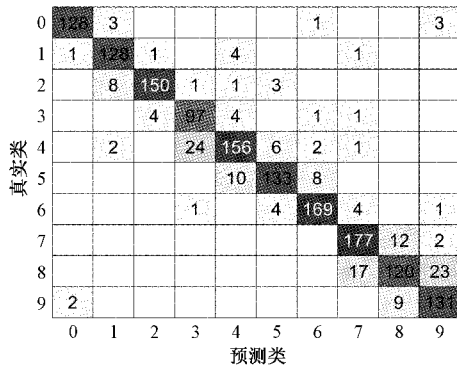


图 14 GoogLeNet 分类混淆矩阵

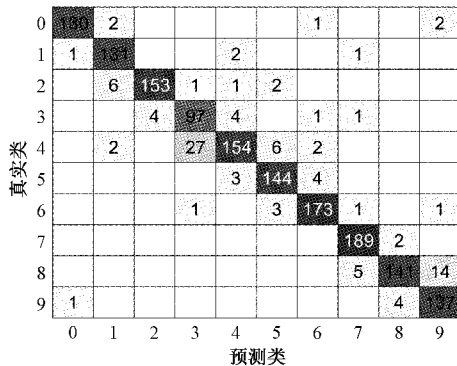


图 15 ResNet-18 分类混淆矩阵

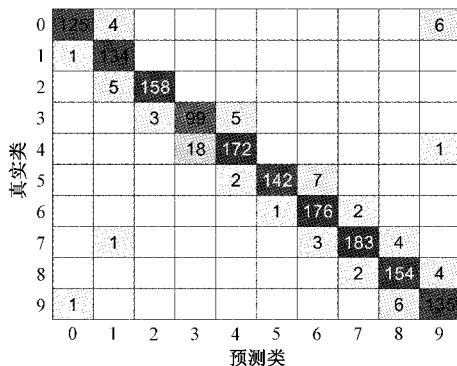


图 16 本文模型分类混淆矩阵

AUC(area under curve)值是衡量一个二分类器性能的重要指标,AUC 表示的是 ROC(receiver operating characteristic)曲线与坐标轴所围成的右下方面积。本文

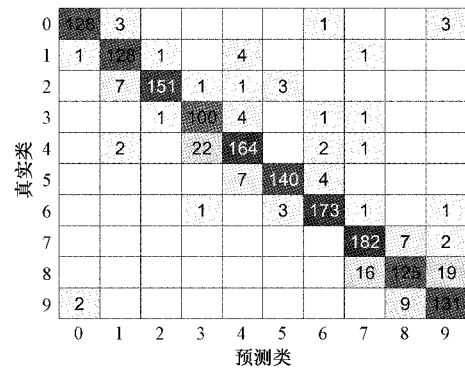


图 17 STM32 可运行模型分类混淆矩阵

表 2 各模型准确率

模型	GoogLeNet	ResNet-18	本文模型	STM32 运行模型
验证集准确率/%	89.64	93.44	95.11	91.51

在 10 分类上将其中一类作为正样本,其余 9 类作为负样本,分 10 次计算出模型对于每一类分类结果的 AUC 值。4 种模型对 10 个分类结果的 AUC 值如图 18 所示,本文模型的 AUC 值在各类别的分类结果上基本高于 GoogLeNet 和 ResNet-18。

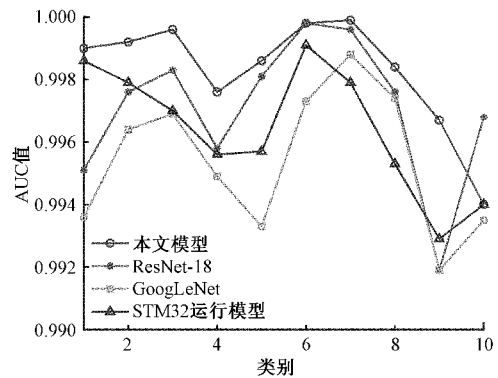


图 18 各模型 AUC 值

3 结 论

本文先后测试 GoogLeNet 和 ResNet-18 两种经典卷积神经网络模型对水表仪表盘读数的识别性能,ResNet-18 的准确率略高于 GoogLeNet,针对于本文问题,根据 ResNet-18 的思想搭建了新的网络模型,降低了网络的深度,使得网络收敛速度加快,使用多个小卷积核代替一个大卷积核的方法,既保证了感受野的大小又减少了网络的参数,节省了计算资源。实验结果表明,相比于 GoogLeNet,本文模型准确率提高了 5.47%,对比于 ResNet-18 准确率提高了 1.67%,本文模型性能在水表仪表盘读数分类问题上优于二者。

将本文模型进行简化,使其可以移植到 STM32 单片机上运行,同时也保证了较高的识别准确率,为嵌入式硬件实现 CNN 算法提供了一种途径,基于 STM32 单片机和卷积神经网络算法设计了一个自动抄表的硬件系统方案,为基于 CNN 算法的小型水表自动抄表装置的开发提供了经验。

参考文献

- [1] 赵梓涵,李东新.基于改进 SRCNN 算法的人脸识别技术研究[J].国外电子测量技术,2020,39(12):74-79.
- [2] 唐波,陈慎慎.基于深度卷积神经网络的轴承故障诊断方法[J].电子测量与仪器学报,2020,34(3):88-93.
- [3] SHANTHAKUMAR S, SHAKILA S, SUNETH P, et al. Environmental sound classification using deep learning[J]. Instrumentation, 2020,7(3):15-22.
- [4] 林楨哲,王桂棠,陈建强,等.基于残差网络深度学习的肺部 CT 图像结节良恶性分类模型[J].仪器仪表学报,2020,41(3):248-256.
- [5] 董云龙,刘行,袁烨,等.基于细节注意力卷积神经网络的仪表自动化识别方法[J].中国科学:技术科学,2020,50(11):1437-1448.
- [6] 余坤,刘明娟,王凡彬,等.基于图像的旋翼式水表指针读数自动识别研究[J].湖北工业大学学报,2016,31(4):34-37.
- [7] 张志锋,王凤琴,田二林,等.基于机器视觉的指针式仪表读数识别[J].控制工程,2020,27(3):581-586.
- [8] 祁建,张燕.基于多特征信息的水表指针自动识别方法[J].计算机与数字工程,2015,43(4):676-679.
- [9] 周娟,易爽.基于卷积神经网络的模糊车牌字符识别[J].中国刑警学院学报,2021(4):124-128.
- [10] YANG F, JIN L, LAI S, et al. Fully convolutional sequence recognition network for water meter number reading[J]. IEEE Access, 2019, 7: 11679-11687.
- [11] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [J]. Advances in Neural Information Processing Systems, 2012, 25: 1097-1105.
- [12] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. ArXiv Preprint,2014, ArXiv:1409.1556.
- [13] 吉训生,谭凯凯.基于卷积神经网络的水表读数识别方法[J].传感器与微系统,2020,39(12):130-133.
- [14] 刘明,曹银杰,耿相珍,等.在微控制器芯片实现神经网络的方法[J].现代电子技术,2020,43(22):1-5,9.
- [15] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1-9.
- [16] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [17] HE K, ZHANG X, REN S, et al. Identity mappings in deep residual networks[C]. European Conference on Computer Vision, Springer, Cham, 2016: 630-645.
- [18] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [C]. International Conference on Machine Learning, PMLR, 2015: 448-456.
- [19] 江涛.基于 MSP430 的智能远传水表设计[J].电子测量技术,2019,42(2):67-72.

作者简介

张鹏飞,硕士研究生,主要研究方向为计算机视觉。

E-mail:zhangpf0819@163.com

叶哲江(通信作者),副教授,硕士生导师,主要研究方向为信号处理、嵌入式系统。

E-mail:546819613@qq.com

杨嘉林,副教授,硕士生导师,主要研究方向为信息技术。

E-mail:455168215@qq.com

李家成,硕士研究生,主要研究方向为生物医学图像处理。

E-mail:1071709595@qq.com