

基于轻量级网络的 PCB 芯片文字识别

蒋子敏, 刘宁钟, 沈家全

(南京航空航天大学 计算机科学与技术学院, 江苏 南京 211106)

摘要:随着人工智能的兴起,深度学习技术在多个领域有了广泛的应用与发展,将深度学习技术与 PCB 芯片文字识别相结合,实现具体的场景应用,具有重要意义。基于深度学习进行文字识别,完成图像中文字信息的自动获取,进一步提高了准确率,极大地节约了人工及时间成本。但传统深度学习模型巨大的参数量以及内存消耗等限制了其在小型设备如移动终端上的应用与发展,难以满足人们日益增长的需求。基于此,该文提出了一种基于轻量级网络的文字识别算法 LWTR。该算法框架主要包括卷积神经网络(进行特征提取)、循环神经网络(进行标签预测)以及 CTC(实现转录),最终得到预测序列。为减小模型参数量,进行通道数的统一并采用多路小卷积及堆叠 Dense Layer 充分提取特征。同时,为加速网络的收敛,提高模型的泛化能力,引入 BN 归一化。结果表明,该算法在 PCB 芯片数据集中文字识别准确率达到 89.58%,与现有文字识别算法相比,在准确率几乎没有下降的情况下具有更小的模型与更快的速度。

关键词:PCB 芯片;文字识别;轻量级神经网络;卷积神经网络;循环神经网络

中图分类号:TP391.4

文献标识码:A

文章编号:1673-629X(2021)12-0055-06

doi:10.3969/j.issn.1673-629X.2021.12.010

PCB Chip Text Recognition Based on Lightweight Network

JIANG Zi-min, LIU Ning-zhong, SHEN Jia-quan

(School of Computer Science and Technology, Nanjing University of
Aeronautics and Astronautics, Nanjing 211106, China)

Abstract: With the rise of artificial intelligence, deep learning technology has been widely used and developed in many fields. It is of great significance to combine deep learning technology with PCB chip text recognition and realize specific domain applications. Finishing text recognition based on deep learning to realize automatic acquisition of text information in images has further improved recognition accuracy and greatly saved labor and time costs. However, the huge parameter amount and memory consumption of the traditional deep learning models limit its application and development on small devices such as mobile terminals, and it is difficult to meet the increasing needs of people nowadays. Therefore, a text recognition algorithm LWTR is proposed based on lightweight network. The algorithm framework mainly includes convolutional neural network for feature extraction, recurrent neural network for label prediction and CTC to achieve transcription, and finally gets the predicted sequence. In order to reduce the amount of model parameters, unified number of channels, multi-path small convolution and stacked Dense Layer are used to fully extract features. At the same time, in order to accelerate the convergence of the network and improve the generalization ability of the model, Batch Normalization (BN) is introduced. The results show that the proposed algorithm has achieved 89.58% of the text recognition accuracy in our PCB chip dataset. And compared with the existing text recognition algorithms, it has a smaller model and faster speed with almost no decrease in accuracy.

Key words: PCB chip; text recognition; lightweight neural network; convolutional neural network; recurrent neural network

0 引言

作为工业电子领域的核心基础组件之一,PCB (printed circuit board) 电路板被广泛应用于计算机、汽车电子、通信等多个领域,在电子元器件电气连接中具有重要的支撑与互连作用^[1]。作为 PCB 电路板中重要的元器件之一,芯片在文字识别中可以检测在封装过程中是否出现印刷错误等缺陷,为 PCB 提供基础质

量保障,是至关重要的环节。

人工进行芯片文字识别将会消耗大量的时间,且容易出现失误,准确率不高。目前,一般与计算机技术相结合加以实现。传统的文字识别技术主要包含预处理、版面分析、特征提取以及后处理等操作。该流程对于背景简单、字体大小一致的版面文字具有较好的处理效果,且已经趋于成熟,但面对一些复杂的场景图片

收稿日期:2021-01-07

修回日期:2021-05-11

基金项目:国家自然科学基金(61375021)

作者简介:蒋子敏(1996-),女,硕士,研究方向为计算机视觉和模式识别;刘宁钟,教授,博导,研究方向为计算机视觉和模式识别。

时很难将文字从图片中完整提取。与场景文字类似,芯片文字具有背景噪声多、字符粘连、字体大小不一致等特点,使用传统的图像处理技术对 PCB 芯片文字进行识别时无法取得较好效果。

近年来,随着科技的不断发展与进步,深度学习被广泛应用于目标检测、语音识别等,文字识别作为基础及热点问题之一^[2-4],在计算机视觉领域引起了广泛的关注与研究。使用深度学习的方法对芯片文字进行识别,准确率更高,且更具鲁棒性。而随着社会不断进步,人们对于人工智能的需求日益加深,在良好需求实现的基础上对模型的速度以及大小有着更进一步的要求,即在保证一定的精度下,尽可能地减少网络模型的计算量与参数量,提高模型运行速度,使其能够运行在移动端、嵌入式等设备中,减少大模型在实际操作中存在的限制。但传统深度学习网络模型较大,难以满足如今移动终端、便携式设备的需求。为了解决上述问题,减小模型大小,提升 PCB 芯片文字识别的效率与速度,该文提出了一种基于轻量级^[5-7]PCB 芯片文字识别算法 LWTR。

1 相关工作

文字识别作为计算机视觉领域的基础及热点问题之一,已经引起了广泛的讨论与关注。早期传统的文字识别主要通过计算机图像处理方法实现,包括图像预处理、版面分析、图像行列切分、单个字符模型训练以及后处理等步骤^[8]。该类传统方法对每一个流程模块的性能都有着较高的标准要求,因此在文字识别的应用范围上具有一定的局限性,在很多场景下,该方法无法将文字从背景复杂、噪声较多、光照不一的图像中分割、训练以及识别,同时,检测与识别流程较为繁琐,难以泛化出较好的模型。

近年来,深度学习技术在不断发展与进步,并在多个领域都有了广泛而普遍的应用^[9],如目标检测、图像分类、语音识别等等,极大地提高了生产生活效率。作为计算机视觉领域的基础及热点研究问题之一,文字识别技术亦得到了显著的发展。其逐渐舍弃传统的图像处理方法,不断向深度学习的框架靠近。将深度学习算法应用于文字检测与识别之中,与传统算法相比,其过程更为简单,具有更好的性能。不同于传统方法的手工特征提取,基于深度学习的方法将图片作为输入,通过自动特征提取等操作,即可得到输出结果,整体而言更加鲁棒,且模型的泛化能力更强。目前,基于深度学习实现端到端文字识别主要分为通过连接主义分类^[10](connectionist temporal classification, CTC)或注意力机制(attention mechanism, AM)的识别方法,分别以 CNN+RNN+CTC^[11]与 CNN+RNN+Attention^[12]

框架为基础。这两种框架的主要区别为将循环神经网络^[13](recurrent neural network, RNN)得到的序列标签转录为最终识别结果的对齐方式差异。目前,CRNN^[14]作为场景文字识别算法被广泛应用,在该算法中,作者选用 VGGNet^[15]作为输入图像的 CNN 特征提取网络,将提取到的特征图转换为特征序列,并将特征序列送入基于双向长短期记忆网络^[16](Bidirectional long short-term memory, Bi-LSTM)实现的循环神经网络层进行序列标签预测,最终通过 CTC 完成转录。该类方法在场景文字识别中得到了良好的效果,但传统的深度学习算法如 VGG、ResNet^[17]等模型参数量较大,计算量较大,受小型设备内存限制较大,难以满足人们日益增长的轻量化应用需求。

基于以上问题,该文提出了一种轻量级文字识别算法 LWTR。该算法网络框架以 CNN+RNN+CTC 为基准实现,共包含特征序列提取层、循环神经网络层以及转录层。针对传统深度模型参数量多、受设备内存限制较大的问题,该文引入轻量级网络 PeleeNet^[18],以该网络为基础并借鉴 ShuffleNet V2^[19]相关设计原则对其进行改进,在很大程度上减少了计算量以及模型参数量,减少小型设备中因模型太大对于内存的限制,提升文字识别速度,使其更加适用于轻量级场景。同时,在循环神经网络层中加入归一化^[20](batch normalization, BN),从而加速网络的收敛,并进一步提高了模型的泛化能力。

2 LWTR 模型

该文提出了一种轻量级 PCB 芯片文字识别算法 LWTR,能够实现快速而准确的高性能 PCB 芯片文字识别。受文字识别网络架构以及轻量级网络的启发,该模型网络由特征序列提取层、循环神经网络层和转录层三部分组成。LWTR 整体结构框架如图 1 所示。

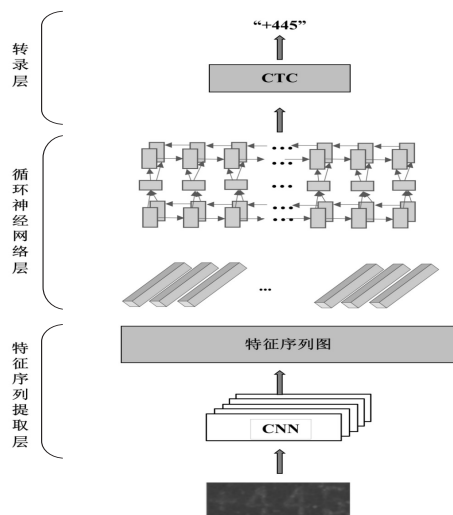


图 1 LWTR 框架

2.1 特征序列提取层

图像输入到网络,将经 CNN 卷积神经网络提取特征图以及特征图到特征序列的转换部分称为特征序列提取层。关于特征提取网络,传统方法通常采用一些较为基础的算法模型实现丰富的特征提取,如 VGG 模型,并取得了良好的效果。但该类网络具有较多的参数量以及计算量,所得到的模型较大,且内存消耗大,在小型设备的应用上受限较多,难以实现快速、准确的高性能任务。同时,较深的网络在训练时收敛较慢,需要大量的训练数据。

为实现轻量级文字识别算法,减少计算成本、内存限制的影响,在特征序列提取层中,该文设计了一种轻量级卷积神经网络用以实现特征提取,在很大程度上提升了 LWTR 整体识别网络的效率及性能。

特征序列提取层的网络如表 1 所示。该层网络共

包含 4 个特征提取的阶段以及特征图到特征序列转换的过程。通过摒弃原 PeleeNet 的阶段 4,选用阶段 0 至阶段 3 共四个阶段进行特征提取。同时,为了满足文中数据集文字长序列的特点,将原网络中每个阶段最后 2×2 的平均池化层改为 2×1 的平均池化层,并且在特征提取的最后一层添加了一个 1×1 的卷积层。受 ShuffleNet V2 相关设计原则的启发,在每个阶段 Dense Layer 的设计中,统一卷积核的通道数,保持其输入输出通道数的一致,从而加快识别速度,且阶段 1,2,3 输出特征图的通道数分别为 128,256 以及 512,从而保证了特征信息的充分提取。同时,为了加速网络的收敛,在各阶段最后使用 BN 进行归一化,进一步提升模型的泛化能力。最终,将所得到的特征图映射为特征序列,并送入循环神经网络层中进行更进一步的操作。

表 1 特征序列提取层

阶段	详细	输出
输入	文字图片	$100 \times 32 \times 1$
阶段 0	Stem block	$25 \times 8 \times 16$
	Dense block	Dense Layer $\times 7$
阶段 1	Transition layer	1×1 conv, stride 1 Batch Normalization
	1×2 average pool, stride 1×2	$25 \times 4 \times 128$
	Dense block	Dense Layer $\times 8$
阶段 2	Transition layer	1×1 conv, stride 1 Batch Normalization
	1×2 average pool, stride 1×2	$25 \times 2 \times 256$
	Dense block	Dense Layer $\times 16$
阶段 3	Transition layer	1×1 conv, stride 1 Batch Normalization
	1×2 average pool, stride 1×2	$25 \times 1 \times 512$
	1×1 conv, stride 1	$25 \times 1 \times 512$
	Batch Normalization	
	Map-to-Sequence	

2.2 循环神经网络层

循环神经网络层主要由 Bi-LSTM 构成。图片经特征序列提取层得到的特征序列被送入 RNN 中得到预测标签,最终采用 CTC 转录实现文字识别。即对于特征序列层提取而得的特征序列 $X = x_1, x_2, \dots, x_L$ (L 为序列长度)输入到循环网络层中,生成一系列预测序列标签 $H = h_1, h_2, \dots, h_L$,其中每个 x_i 对应一个 h_i 。对于 RNN 在训练中存在的梯度消失导致难以完成长序列的预测问题,常通过改变其单元结构加以优化。作为一种特殊的 RNN,长短期记忆网络^[21](long short-term memory, LSTM)通过引入门机制及细胞状态,有

效解决了长期依赖的瓶颈问题,使得网络可以处理任意长度的序列问题。对于 LSTM 而言,其看到、记住的信息均来自过去,无法获得之后的相关信息。而在文字识别的算法应用中,文字序列过去与未来的信息间充满关联,上下文间处于互补、一致状态,如主语、时态的确定时前后信息都至关重要。因此,该文选择将两个 LSTM 进行前后组合为 Bi-LSTM,从而充分获得文字的前后语义信息,对于文字的整体识别具有重要意义。

同时,该文采用多层 Bi-LSTM 构建循环神经网络层。与单层 Bi-LSTM 相较,堆叠的多层 Bi-LSTM

可以获得更高层次的抽象,从而提取更为充分的信息,在文字识别过程中起到了至关重要的作用。

2.3 CTC 转录层

作为 LWTR 网络模型的最后一部分,转录层负责将循环神经网络层中预测得到的序列标签转化为最终文字识别结果加以输出,是文字识别框架中至关重要的一步。对于循环神经网络而言,若采用常规 SoftMax 损失函数,则每一列的输出均对应一个字符元素。该过程需要在训练时对每张图片进行字符及其位置的标记,再将其对应到特征序列获取对应标签方能进行,耗费了大量的时间及精力。CTC 作为一种时序分类算法,多用于处理序列任务中标注问题中的输入以及输出标签的难以自动对齐问题。通过引入 CTC 算法至损失函数的计算,将其与 RNN 相结合,无需对训练数据进行预分割以及后续处理操作,只需输入与输出序列。整体网络的输入为未进行分割、对齐的序列标签,输出则为经过规则映射得到的序列标签中概率最大的序列。该类应用避免了字符或音素级别的标注,在单一网络架构中完成对序列的建模,呈现出一种用循环神经网络标记序列数据的新方法,真正实现端到端的网络模型,进一步简化了任务的流程,极大地提升了文字识别的效率,在语音识别、文字识别等场景中被广泛应用。

在文中文字识别的处理中,CTC 主要在训练时进行标签的对齐并参与损失函数的计算过程。同时,在预测过程中,CTC 通过计算条件概率将预测的标签文字转换为最终的输出结果,从而实现文字识别的整体过程。因此,该网络可以使用序列标签作为监督进行训练。对于给定的特征输入序列 x ,输入图片文字识别为序列 l 的概率如公式(1)所示:

$$p(l|x) = \sum_{\pi \in \mathcal{B}^{-1}(l)} p(\pi|x) \quad (1)$$

其中, \mathcal{B} 变换为序列到序列的映射函数,即对 RNN 层的输出进行变换, $\pi \in \mathcal{B}^{-1}(l)$ 代表所有经过 \mathcal{B} 变换后输出为 l 的路径。在该过程中,首先会去除重复标签,随后将空格进行移除,如 \mathcal{B} 将“--hh-e-l-ll-oo--”(其中“-”代表 blank 空格)映射为“hello”。选择概率最大的序列 l 作为最终结果,从而完成输入与输出序列的自动对齐,避免了传统方法中对输入序列的手动分割,实现了端到端文字识别任务,进一步提高了模型的效率与性能。

2.4 损失函数

对于训练数据集 $X = \{I_i, l_i\}_i$,其中 I_i 代表训练图像, l_i 代表对应图像文字的真实值标签序列。则本实验中算法的训练目标为最小化真实条件概率的最大负对数似然函数。该损失函数的定义如式(2)所示:

$$\mathcal{O} = - \sum_{I_i, l_i \in X} \log p(l_i | y_i) \quad (2)$$

其中, y_i 代表图像 I_i 经卷积神经网络层以及循环神经网络层产生的序列。损失函数通过图像以及它的真实标签序列计算损失值。基于此,本算法研究中的网络可以基于训练图像及其真实标签序列值进行端到端的训练,从而避免了人工标记单个字符的过程,进一步提高了效率。

3 实验结果与分析

3.1 评价标准

关于文字识别的评价,通常存在两种标准,分别为编辑距离以及准确率,其中准确率可以细分为字符准确率、字段准确率以及图片文字识别准确率。

编辑距离又称莱文斯坦距离 (Levenshtein distance, LD),在信息学、计算机科学等领域被广泛应用。其作为一种特殊的量化算法常常被用来衡量两个字符串序列之间的相似度,即对于两个单词或者是两个字符串,从其中一个转换为另一个所需的最少编辑操作次数。如对于单词 *cafe* 与 *coffee*,这两者的编辑距离为 3,对于单词 *set* 与 *sit*,其编辑距离为 1。对于两者间的操作转换,仅限于插入、删除以及替换共计三种单字符的操作。通过编辑距离,当预测标签与真实值之间的差异编辑距离越小,相似度越大,则识别准确率越高。

对于文字识别准确率的计算,针对其基准不同,可分为字符准确率、字段准确率以及图片文字识别准确率。字符识别准确率为识别正确的字符数占总字符的比率,字段识别准确率为图像中完全识别正确的字段占总字段数的比率,而图像文字识别准确率则用识别正确的图片数量除以总图片数。

其中,在字段识别以及图片文字识别准确率中,识别正确的字段以及图片要求字段或图片中文字每一个字符识别均完全正确。在现实应用中,由于其直观与便捷性,准确率常常被选用为文字识别的评价指标。在本实验中,由于数据集中图片为分割完成的单独字段行,采用字段准确率(图片准确率)作为文中文字识别的评价指标。

3.2 实验设置

实验在 Tensorflow 深度学习框架上进行,操作系统为 Ubuntu 16.04, CPU 为 i7-8700k,显卡为 GeForce GTX 1080 Ti,内存为 15.6 GiB, 11 GB 显存。

采用基于批处理的随机梯度下降方法进行模型的优化。基本学习率设为 0.01,学习衰减率设置为 0.1,重量衰减值设为 0.000 5,采用 Kaiming^[22] 初始化进行参数的初始化过程。在训练阶段,通过最小化 CTC 损

失函数共同训练 LWTR 中的卷积神经网络与循环神经网络。由于 PCB 文字数据集数据量较小,对该部分数据进行增加噪声、调整对比度等操作实现数据增广,增广后的数据集作为该文的训练集进行训练。同时,对于输入图像,将其缩放至 100×32 的大小进行网络的训练以及测试。

3.3 实验结果与分析

将 PCB 芯片数据集运行在不同的网络架构之中,实验结果如表 2 所示。实验表明,该文提出算法的识别准确率为 89.58%,比以 VGG 16 为主干网络的 CRNN 算法识别率低 0.42%。可以清晰地发现,采用所提出的 LWTR 算法模型,相比以 VGG 16 为主干的提取网络,在准确率几乎没有下降的情况下,模型的参数减少了一半多,同时,与 PeleeNet 为主干架构的网络模型相比,所提出方法的识别准确率高了 2.5%,模型大小减小了 36%。实验验证了所提出的算法的有效性,即在 PCB 芯片数据集上进行文字识别,与其他网络相比,具有更好的性能。

表 2 实验对比

方法名称	准确率/%	模型大小/M
VGG 16+CRNN	90.00	69.9
PeleeNet+CRNN	87.08	49.8
LWTR(Our method)	89.58	31.8

在循环神经网络层,采用堆叠的 Bi-LSTM 进行特征序列到标签的预测。为深入探讨 Bi-LSTM 的堆叠深度以及卷积层中通道数的大小对于文字识别模型整体性能的影响,本实验在 PCB 芯片数据集上通过 RNN 深度不同以及通道数不同的设置分别进行消融实验。消融实验结果如表 3 所示。其中,本模型中通道数的选择为 8,Scale 代表所选取的通道数为模型通道数的倍数,当 Scale=0.5 时,通道数为 4,当 Scale=1 时,通道数为 8。通过消融实验可以发现,循环神经网络层中 Bi-LSTM 的叠加深度以及卷积神经网络中通道数对实验结果有不同的影响。由于不同的 RNN 层数以及通道数直接影响着模型的大小、计算量以及运行速度,当选用一层 Bi-LSTM 时,虽然其参数量与叠加的 RNN 相比较小,且速度较快,但网络收敛较慢,无法充分学习,在同等情况下(通道数相同),相较层次较深的网络无法得到可观的识别结果。而当 Bi-LSTM 深度较深、通道数较大时,会在一定程度上造成信息的冗余,从而带来信息的丢失,不利于识别准确率的提升。同时,由于深度较深,模型参数量较大,其整体的识别速度大幅度下降。实验表明,在使用 2 层 Bi-LSTM 且通道数为 8 时,即采用该文提出的算法时,在识别准确率几乎等同于最好值 90.00% 时,具有更

小的模型以及更快的速度。其得到的结果可以在速度、模型大小以及识别准确率间谋取平衡,达到较高的性能,具有较好的综合实验效果。

表 3 消融实验对比

Scale	网络结构	PCB 芯片数据集/%	模型大小/M	速度/FPS
0.5	1 层 Bi-LSTM	68.75	17.6	25
	2 层 Bi-LSTM	86.67	30.2	17
	3 层 Bi-LSTM	86.67	42.8	13
1	1 层 Bi-LSTM	81.67	19.2	25
	2 层 Bi-LSTM	89.58	31.8	18
	3 层 Bi-LSTM	87.50	44.4	14
2	1 层 Bi-LSTM	87.50	23.7	24
	2 层 Bi-LSTM	90.00	36.3	17
	3 层 Bi-LSTM	84.58	48.9	13

3.4 结果可视化

将 LWTR 网络模型应用于 PCB 数据集中进行芯片文字的识别测试,测试结果图如图 2 所示。图 2 所展示的为部分测试图像以及其对应的模型识别结果。从图 2 可以得出,该文提出的算法应用于 PCB 芯片文字识别时取得了较好的效果,在背景各异、字体及大小不同的图片文字识别中具有较强的鲁棒性。



图 2 结果可视化

4 结束语

该文主要提出了一种轻量级文字识别算法 LWTR 并将其应用于 PCB 芯片文字识别中。通过引入改进的轻量级网络进行特征提取,大大减少了网络的参数量,提高了识别速度,同时,引入 BN 归一化加速网络收敛。实验证明,该文提出方法在保证识别准确率的前提下获得了更小的模型与更快的识别速度,相比传统深度模型更具轻量级特点。

参考文献:

- [1] 蔡茂蓉. PCB 缺陷检测系统的研究与实现[J]. 信息系统工程, 2009(6): 98-100.
- [2] 黄攀. 基于深度学习的自然场景文字识别[D]. 杭州: 浙江大学, 2016.
- [3] 蒋瓌. 基于深度学习的发票识别系统[D]. 南京: 南京邮

- 电大学,2020.
- [4] 陈桂安.端到端的自然场景文字检测与识别神经网络的研究与实现[D].上海:上海师范大学,2019.
- [5] HONG S, ROH B, KIM K, et al. PVANet: lightweight deep neural networks for real-time object detection[J]. arXiv: 1611.08588, 2016.
- [6] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: efficient convolutional neural networks for mobile vision applications[J]. arXiv:1704.04861, 2017.
- [7] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size[J]. arXiv:1602.07360, 2016.
- [8] 孙 华, 张 航. 汉字识别方法综述[J]. 计算机工程, 2010, 36(20): 194-197.
- [9] LONG S, HE X, YAO C. Scene text detection and recognition; the deep learning era[J]. International Journal of Computer Vision, 2021, 129(1): 161-184.
- [10] GRAVES A. Connectionist temporal classification[M]//Supervised sequence labelling with recurrent neural networks. Berlin: Springer, 2012: 61-93.
- [11] HE P, HUANG W, QIAO Y, et al. Reading scene text in deep convolutional sequences [C]//Proceedings of the AAAI conference on artificial intelligence. Phoenix, Arizona, USA: AAAI Press, 2016: 3501-3508.
- [12] SUTSKEVER I, VINYALS O, LE Q V. Sequence to sequence learning with neural networks[C]//Neural information processing systems. Montreal, QC, Canada: MIT Press, 2014: 3104-3112.
- [13] GRAVES A, MOHAMED A, HINTON G. Speech recognition with deep recurrent neural networks [C]//2013 IEEE international conference on acoustics, speech and signal processing. Vancouver, Canada: IEEE, 2013: 6645-6649.
- [14] SHI B, BAI X, YAO C. An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(11): 2298-2304.
- [15] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv:1409.1556, 2014.
- [16] GRAVES A, SCHMIDHUBER J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures[J]. Neural Networks, 2005, 18(5-6): 602-610.
- [17] 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. Las Vegas, NV, USA: IEEE, 2016: 770-778.
- [18] WANG R J, LI X, LING C X. Pelee: a real-time object detection system on mobile devices [C]//Advances in neural information processing systems, Montréal, Canada: [s. n.], 2018: 1963-1972.
- [19] MA N, ZHANG X, ZHENG H T, et al. Shufflenet v2: practical guidelines for efficient cnn architecture design [C]//Computer vision - ECCV 2018. Munich, Germany: Springer, 2018: 122-138.
- [20] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [J]. arXiv:1502.03167, 2015.
- [21] MALHOTRA P, VIG L, SHROFF G, et al. Long short term memory networks for anomaly detection in time series [C]//23rd European symposium on artificial neural networks, computational intelligence and machine learning. Bruges, Belgium: [s. n.], 2015: 89-94.
- [22] HE K, ZHANG X, REN S, et al. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification [C]//Proceedings of the IEEE international conference on computer vision. Santiago, Chile: IEEE, 2015: 1026-1034.