

基于单字符注意力的全品类鲁棒车牌识别

穆世义^{1,2} 徐树公^{1,2}

摘要 复杂场景下的高精度车牌识别仍然存在着许多挑战,除了光照、分辨率不可控和运动模糊等因素导致的车牌图像质量低之外,还包括车牌品类多样产生的行数不一和字数不一等困难,以及因拍摄角度多样出现的大倾角等问题.针对这些挑战,提出了一种基于单字符注意力的场景鲁棒的高精度车牌识别算法,在无单字符位置标签信息的情况下,使用注意力机制对车牌全局特征图进行单字符级特征分割,以处理多品类车牌和倾斜车牌中的二维字符布局问题.另外,该算法通过使用共享参数的多分支结构代替现有算法的串行解码结构,降低了分类头参数数量并实现了并行化推理.实验结果表明,该算法在公开车牌数据集上实现了超越现有算法的精度,同时具有较快的识别速度.

关键词 车牌识别,注意力机制,字符分割,字符分类

引用格式 穆世义,徐树公.基于单字符注意力的全品类鲁棒车牌识别.自动化学报,2023,49(1):122-134

DOI 10.16383/j.aas.c211210

Full-category Robust License Plate Recognition Based on Character Attention

MU Shi-Yi^{1,2} XU Shu-Gong^{1,2}

Abstract There are still many challenges for high-precision vehicle license plate recognition in complex scenarios. In addition to the low quality of license plate images caused by factors such as poor illumination, low resolution, and motion blur, challenges also include different variant numbers of characters and lines for different license plate categories, as well as large inclination caused by the various camera locations. In response to these challenges, this paper proposes a scene-robust high-precision license plate recognition algorithm based on character attention, which performs character level segmentation on the global feature map of the license plate images without character position label information. Such character level segmentation can deal with the 2D character layout problems in multi-category license plates and inclined license plates. In addition, this algorithm uses a shared weight classification header structure to replace the serial decoding structure used in existing algorithms, which reduces the number of classification header parameters and realizes parallel inference. The experimental results show that the algorithm achieves high accuracy which surpasses the existing algorithms on the public-domain data sets, and meanwhile has a faster recognition speed.

Key words License plate recognition, attention mechanism, character segmentation, character classification

Citation Mu Shi-Yi, Xu Shu-Gong. Full-category robust license plate recognition based on character attention. *Acta Automatica Sinica*, 2023, 49(1): 122-134

机动车车牌作为机动车重要的身份标识,对车牌的精准识别具有较高的应用价值.车牌识别技术已经广泛应用在交通监控、门禁管理、智慧交通等场景中.然而大部分现有算法仅在光照、拍摄距离和拍摄角度等相对固定的受限场景中达到实用的准确度,甚至有些识别系统只能识别单品类的车牌.

本文主要研究车牌识别中多品类车牌兼容和复

杂场景下的大角度倾斜的两大挑战.

多品类兼容的难点在于不同品类的车牌的字符布局不同.除字符个数不同外,字符空间布局差异较大,尤其是单行文本车牌的一维布局和双行文本车牌的二维布局之间的差异难以兼容.

角度倾斜的车牌和角度水平的车牌之间不同之处也在于其空间布局方式的不同.水平的车牌图像中字符序列沿水平方向从左到右依次排列.倾斜车牌图像中的字符布局存在高度差,呈现出一种二维的对角线布局状态.

针对不同长度的单行文本车牌,通常可采用基于连接时序分类(Connectionist temporal classification, CTC)的算法进行识别. Li 等^[1]提出了基于卷积神经网络与 CTC 结合的算法. Wu 等^[2]提出了 DenseNet^[3]结合 CTC 的算法. Yang 等^[4]提出了

收稿日期 2021-12-20 录用日期 2022-04-07

Manuscript received December 20, 2021; accepted April 7, 2022

国家自然科学基金(61871262)资助

Supported by National Natural Science Foundation of China (61871262)

本文责任编辑 金连文

Recommended by Associate Editor JIN Lian-Wen

1. 上海大学通信与信息工程学院 上海 200444 2. 上海先进通信与数据科学研究院 上海 200444

1. School of Communication and Information Engineering, Shanghai University, Shanghai 200444 2. Shanghai Institute for Advanced Communication and Data Science, Shanghai 200444

基于卷积、全连接分类头和 CTC 的 HomoNet 算法. He 等^[5] 采用了场景文本识别领域中广泛应用的卷积神经网络结合循环神经网络 (Recurrent neural network, RNN) 的组合方式, 利用循环神经网络对前后语义信息进行建模. Lee 等^[6]、Cheng 等^[7] 和 Zou 等^[8] 提出了一些基于一维注意力的算法用于提高识别准确率. 此类针对一维水平布局的单行字符而设计的算法无法很好地解决二维布局的倾斜车牌和双行车牌.

可兼容一维和二维布局的方法有四种: 基于多个分类头的算法、基于单字符检测分割的方法、基于空间矫正的方法和基于二维注意力的方法.

基于多个分类头的算法对单双行车牌设计不同的分类头以实现分而治之. 依据车牌品类分类结果选择对应品类的分类头进行字符分类. Qin 等^[9] 在识别器中设计分支结构预测车牌品类后选择对应的字符分类头. 此类方法解决了单双行的兼容问题, 但是无法应对大角度倾斜导致的二维空间布局问题.

基于单字符检测分割的算法^[10-14] 对车牌图像中的单字符进行逐一检测定位, 并逐一进行分割分类, 根据空间位置关系对多个分类结果进行排序重组即组成所预测的车牌字符序列. 该种算法可有效地解决倾斜导致的二维布局问题, 同时可兼容单双行的车牌识别. 但是存在不足之处, 训练阶段需要每个单字符的位置和类别标签进行监督训练.

基于空间矫正的算法将倾斜的车牌图像整体矫正为水平状态再进行车牌识别. Luo 等^[15] 提出了基于空间变换网络^[16] 的车牌识别算法, 周晓君等^[17] 提出了相似的算法用于车牌图像的空间矫正. 此类算法可将倾斜单行文本的二维布局矫正为一维布局, 但是无法将双行文本矫正为单行文本.

基于二维注意力的算法则无需对输入图像进行空间矫正, 而是利用二维注意力在特征提取过程中更加关注单字符区域特征从而依次对单字符进行识别. Li 等^[18] 提出基于长短时记忆网络的二维注意力机制用于场景文本识别, Xu 等^[19] 和 Zhang 等^[20] 将该注意力机制应用在复杂场景的车牌识别, 提高了倾斜车牌的识别鲁棒性. Hu 等^[21] 提出了采用图卷积网络 (Graph convolutional neural network, GCN) 代替基于长短时记忆网络的二维注意力结构应用于文本识别, 加快了网络推理速度. Yan 等^[22] 引入 GCN 增强了单字符特征的表达能力, 有效地提升了场景文本识别的鲁棒性.

上述算法中基于单字符检测分割的算法和基于二维注意力的算法可用于解决双行字符和倾斜车牌的二维布局识别问题. 前者需要大量的单字符位置

标签, 后者无需单字符位置标签. 然而后者结构中的串行解码器耗时较大, 并且无法实现并行化. 在实际应用中, 部署车牌识别算法的嵌入式设备计算能力较弱, 因此需要较小的计算复杂度和较少的参数量才能够满足车牌识别应用部署阶段的实时性需求.

本文提出基于字符注意力的识别网络 (Character attention based recognition network, CAR-Net), 采用二维注意力的机制对全局特征图进行单字符特征分割, 提出可并行化的多分支结构代替现有的串行解码结构, 提高了算法推理速度. 并在多个分类头之间进行参数共享, 有效降低网络参数量. 本文的主要贡献包括:

1) 采用了单字符注意力解决单双行布局不同的兼容难题.

2) 设计了参数共享的多字符分类头网络, 实现多字符的并行化预测.

3) 在中国城市停车场数据集 (Chinese city parking dataset, CCPD) 和中国车牌数据集 (China license plate dataset, CLPD) 上实现了超越现有公开算法的车牌识别准确率

1 车牌识别算法

1.1 算法结构分析与改进

基于循环神经网络的二维注意力算法结构见图 1 中算法 1. 该种算法在循环解码器的解码过程中依次进行二维注意力图预测, 并采用全连接 (Fully connected, FC) 网络进行单字符分类. 该种算法的串行结构无法实现并行化推理, 并且单步耗时较大. 由于循环神经网络在预测当前一步的字符时需要前一步的字符预测结果的编码作为输入, 当解码过程中的某一步发生识别错误时, 会影响到后一步的注意力图预测, 直接导致识别结果错位.

针对注意力漂移现象, Wang 等^[23] 提出解耦合的单字符注意力机制应用于场景文本识别和手写字识别, 解决了注意力漂移问题. 其结构见图 1 中算法 2. 该算法将注意力预测过程与循环解码过程解耦, 先对多个字符的注意力图进行并行化预测, 再进行循环解码的分类过程. 该算法对自然场景中的字符串识别性能提升显著, 但是该种结构保留了循环神经网络作为串行解码器. 自然场景中的字符串多为单词或有规律的序列, 字符间存在前后语义关联. 然而在车牌识别的任务中这种规律并不存在, 除第 1 位字符和第 2 位字符表示地区以外, 其余字符可看作为随机排列的字符序列, 这种序列并不存在语义关联. 因此, 本文在解耦合注意力的基础上

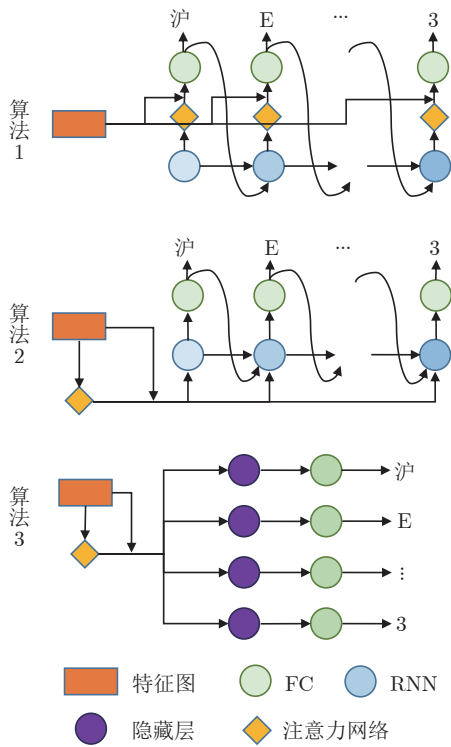


图 1 注意力机制改进

Fig. 1 Evolution of attention mechanism

优化了串行解码器结构, 采用了可并行化推理的多分支分类头进行替代. 此并行分支结构见图 1 中算法 3. 与基于解耦合注意力的算法 B 类似, 首先并行化地预测多个单字符注意力图, 利用注意力图对多个字符特征进行解耦分离. 不同之处在于本文直接采用多分支全连接分类头进行单字符预测, 而不是采用之前的串行解码结构. 这种多分支结构实现了并行化推理, 相对于循环神经网络组成的串行结构, 并行化结构优势为可节约推理阶段的耗时, 并且全连接网络的简单结构相较于循环神经网络能够更轻易地在嵌入式设备上实现部署.

本文提出多分支结构对多个位置的字符进行独立预测, 将车牌字符序列识别任务转化为多个并行

的单字符识别任务. 事实上多个独立字符识别任务并不是完全不相关的任务. 不同位置的分类头所需预测的字符类别是高度重叠的, 故本文将多个独立的字符分类任务融合为单个字符分类任务. 实现方式是在多个分类头之间进行参数共享, 既降低了网络参数量又扩大了单个分类头的训练样本. 多分支的分类头结构被简化成为单支分类头的孪生结构, 模型结构及模型参数量得以简化.

1.2 CARNet 车牌识别网络结构

本文提出基于单字符注意力的 CARNet 车牌识别算法的整体框架见图 2, 主要包含以下 3 个部分:

- 1) 轻量化特征提取, 由多层卷积网络组成. 主干网络从图像中提取出多种尺度的全局特征图, 代表不同的视野域特征. 并将多个尺度不同的全局特征图融合形成一个全局特征图.
- 2) 单字符注意力网络, 使用卷积下采样结构和反卷积上采样结构结合, 预测每个字符的注意力图, 并进行单字符特征分割.
- 3) 并行化单字符分类, 使用共享参数的并行多分支结构对分割出的单字符特征进行并行化分类.

1.3 轻量化特征提取

本文在 Xception^[24] 基础上改进得到 19 层卷积结构作为特征提取网络. 基础单元由深度可分离卷积组成. 本文提出的轻量化特征提取网络相对于现有算法^[23, 25] 中常用的 Resnet45 特征提取网络结构的优点在于获得相同准确率表现的情况下使用更少的网络参数, 具有更快的推理速度. 轻量化特征提取网络包括浅层特征提取模块、中层特征提取模块和深层特征提取模块 3 个模块. 随着网络加深, 网络可获得更大视野域的特征. 19 层卷积结构由多个块组成, 每个块之间由卷积核大小为 1 的卷积层进行跳跃连接. 其细节设置如图 3 所示, 其中虚线区域结构重复 2 次.

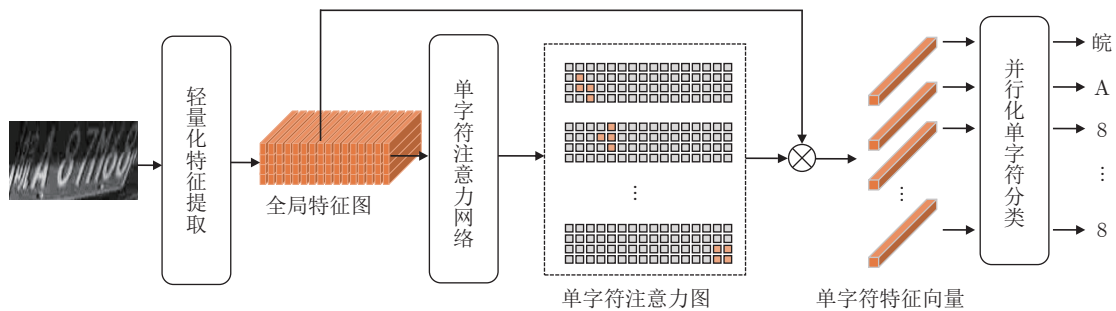


图 2 CARNet 算法结构图

Fig. 2 Framework of the proposed algorithm CARNet

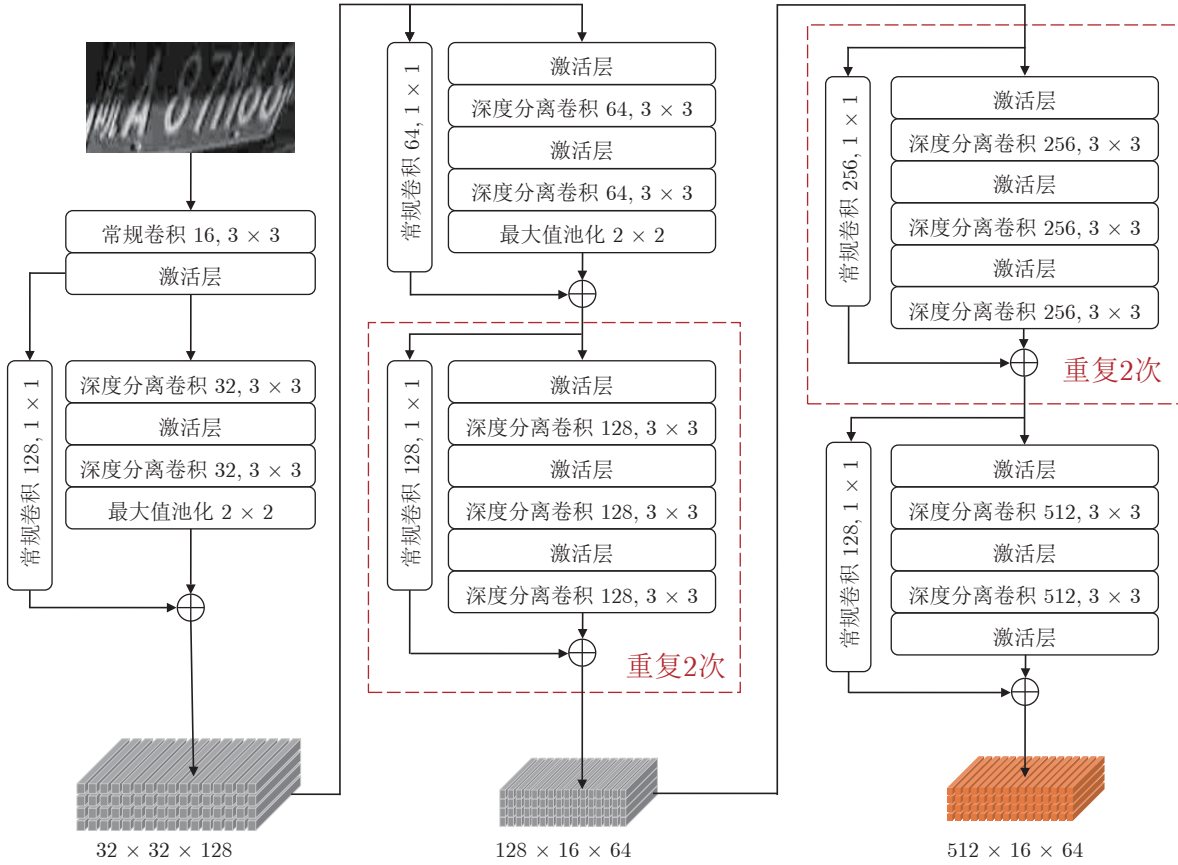


图3 轻量化特征提取

Fig.3 Lightweight feature extraction

输入的车牌图像为 $X \in \mathbf{R}^{C \times H \times W}$, 在本文的实验中网络输入车牌图像的宽 W 为 256 像素, 车牌图像的高 H 为 64 像素, 并且为单通道的灰度图. 本文轻量化特征提取网络所输出的多尺度特征图包括: 浅层特征图 $f_L \in \mathbf{R}^{32 \times (H/2) \times (W/2)}$ 、中层特征图 $f_M \in \mathbf{R}^{128 \times (H/4) \times (W/4)}$ 和深层特征图 $f_S \in \mathbf{R}^{512 \times (H/4) \times (W/4)}$.

使用式 (1) 的方式对多个尺度的特征图进行特征融合. 首先对浅层特征 f_L 进行卷积下采样操作以调整其尺寸, 再与中层特征 f_M 相加, 再次经过卷积操作将其扩大通道数与 f_S 尺寸相同后与深层特征 f_S 相加. 融合后的特征图为 $f \in \mathbf{R}^{512 \times (H/4) \times (W/4)}$, 所获得的全局特征图 f 包含三种尺度视野域的特征信息, 有利于解决车牌检测误差导致的字符尺度不一的问题.

$$f = \text{Conv}(\text{Conv}(f_L) + f_M) + f_S \quad (1)$$

1.4 单字符注意力网络

在现有的基于单字符检测分割的车牌识别算法中, 通过字符分割算法将一张多字符车牌图像分割为多张单字符图像, 从而进一步对单字符图像进行

单字符分类. 在一些基于注意力机制的文本识别算法中利用二维注意力对字符特征进行加权. 可以观察到二维注意力图和原图中的单字符位置存在一一对应的空间分布. 受此启发, 本文利用该种空间对应关系, 利用二维注意力图对二维的全局特征图进行单字符特征分割, 实现单字符特征间的解耦合.

如图 4 所示, 单字符注意力网络的输入为融合后的全局特征图 f 输出为多个字符特征注意力图 $A \in \mathbf{R}^{T \times (H/4) \times (W/4)}$. A 由 T 张单字符注意力图 $\alpha \in \mathbf{R}^{1 \times (H/4) \times (W/4)}$ 组成, 其中 T 为最大可识别字符长度. 单字符注意力网络结构由 4 个卷积层和 4 个反卷积层结构组成. 上下采样结构的对称位置之间存在跳跃连接, 这种连接的方式和图像分割任务图像分割任务中的“U”形网络^[26]方式相同.

与图像分割任务不同的是, 本文分割过程用于特征空间的特征图裁剪, 而非对原始输入图像进行的单字符图像裁剪. 单字符注意力网络的参数在训练过程中仅使用最终的字符识别损失进行监督训练, 并未使用像素级别的单字符位置标签进行监督训练. 与 Zhang 等^[20]的基于循环神经网络的二维注意力的串行结构不同, 本文使用的单字符注意力结

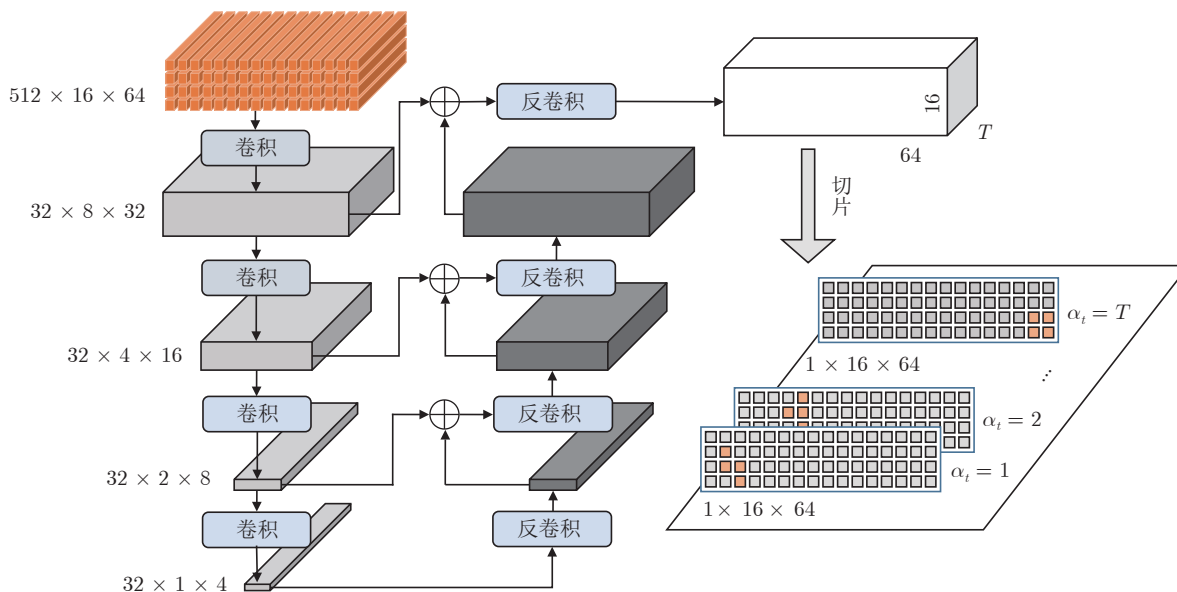


图 4 单字符注意力网络

Fig.4 Single character attention network

构为并行结构,可并行化地预测出全部单字符的注意力图.实现了各字符注意力之间的解耦合,后一个字符的注意力预测不依赖于前一个字符的位置信息,可有效地避免注意力漂移现象的发生.同时并行结构相对于串行结构能够节约网络推理的耗时,实现更快的推理速度.

本文以注意力图为权重对全局特征图中的各像素位置的特征进行加权,保留单字符特征,并抑制背景纹理和其他字符的特征.可理解为利用注意力图对全局特征图进行像素级别的软分割,分割出多个单字符各自独有的字符特征图.单字符特征图的分割过程如式(2)所示,分割后的第 t 个字符的字符特征为 $f_t \in \mathbf{R}^{512 \times (H/4) \times (W/4)}$ 由该字符的注意力图 α_t 和全局特征 f 进行哈达玛积计算而来.多次进行乘法操作对 f 中的多个单字符特征进行逐个增强,同时抑制了背景特征,实现对多个单字符特征的分割分离.其分割过程如图5所示,注意力图 α_t 的数值较大位置所对应 f 中的相同位置的特征得以保留,其余位置的特征被抑制.

$$f_t = \alpha_t f \quad (2)$$

$$c_t = \sum_{i=1}^{\frac{H}{4}} \sum_{j=1}^{\frac{W}{4}} f_{t,i,j} \quad (3)$$

基于式(3)对分割分离后的单字符特征图 f_t 进行求和,获得所需的第 t 个字符特有的单字符特征向量 $c_t \in \mathbf{R}^{512 \times 1}$,用于后续并行化单字符分类.这种求和操作并行化地进行 T 次即可分离出全部字符特有的单字符特征向量.

1.5 并行化单字符分类

上述字符特征分割过程将单字符特征从全局特征中分离并映射为 T 个长度为512的单字符特征向量 c_t .为提高车牌识别算法的运行速度,避免使用基于长短时记忆网络的串行解码器结构.本文在解码阶段设计了多个共享参数的多分支并联解码结构,实现多个字符的并行化解码.每个分支结构由两层全连接网络组成分类头网络.第1层全连接层作为隐藏层输出特征向量为 $h_t \in \mathbf{R}^{256 \times 1}$,第2层全连接层输出为单字符分类结果 $y_t \in \mathbf{R}^{N \times 1}$,其中 N 为车牌字符类别数.为降低网络参数量,多个分支之间进行网络参数共享.

$$h_t = W_1 c_t + b_1 \quad (4)$$

$$y_t = W_2 h_t + b_2 \quad (5)$$

式中, W_1 和 W_2 表示全连接层的可学习权重参数, b_1 和 b_2 表示偏置参数.本文采用的多分支分类头输出结果为固定长度的字符序列,而不同品类车牌的字符长度不同,例如蓝色车牌为7位字符,新能源车牌为8位字符.因此实验中对短的字符序列使用“&”符号进行占位补齐.

1.6 损失函数

$$Loss = -\frac{1}{T} \sum_{t=1}^T \sum_{i=1}^N y'_{t,i} \log_2(y_{t,i}) \quad (6)$$

单个分支的分类损失函数为交叉熵损失函数,字符序列预测总损失函数为多个分支分类损失的平

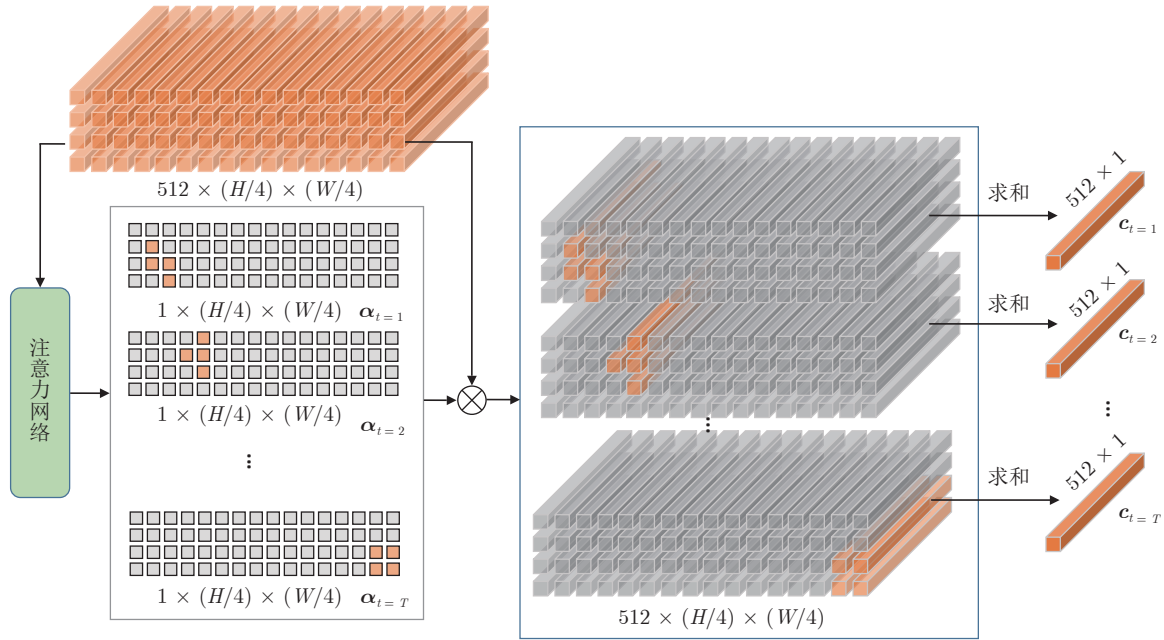


图 5 单字符特征分割

Fig. 5 Single character feature segmentation

均值. 如式 (6) 所示, 其中 $\mathbf{y}'_i \in \mathbf{R}^{N \times 1}$ 为字符分类标签. 本文所提出的 CARNet 车牌识别算法在训练过程中仅使用式 (6) 所表示的平均交叉熵损失进行监督训练.

2 车牌识别数据集

2.1 CCPD 车牌数据集

中国车牌数据集 CCPD 由 Xu 等^[27] 提出, 该车牌数据集包含 7 个子集共计 28 万张蓝色车牌. 其中基础集图像数量最大, 并且为识别难度较小的常规车牌图像, 明暗集为光照条件极端亮或暗的车牌图像, 远近集为拍摄较远和较近的车牌图像, 旋转集为水平角度旋转的车牌图像, 倾斜集为倾斜角度较大的车牌图像, 天气集为不同天气条件下采集的车牌图像, 挑战集为图像质量较差的高难度车牌图像. 该数据集的标签数据中包含车牌位置标签和车牌字符标签, 每张图像中仅包含一个车牌.

2.2 CLPD 车牌测试数据集

CLPD 由 Zhang 等^[20] 提出, 该数据集包含 1200 张中国大陆各地区的车牌图像, 省份类别字符种类分布相对于 CCPD 更均衡. 除少量黄色单行车牌和新能源车牌以外, 其余均为小型汽车蓝色车牌. 标签数据与 CCPD 数据集标签相似, CLPD 中的标签数据包含车牌位置标签和车牌字符标签, 且每张测试图像中仅有 1 个车牌.

2.3 混合品类车牌测试数据集

SYSU-ITS^[28] 车牌数据集为混合品类车牌识别测试集, 所包含的车牌品类丰富, 包括小型车蓝牌、新能源绿牌、大型车后牌、教练车牌、港澳车牌和大型车前牌. 该数据集的标签数据中仅包括车牌字符类别标签, 无车牌位置标签, 且每张测试图像中仅有 1 个车牌. 由于 SYSU-ITS 中子集样本数量较小, 本文在该数据集基础上进行了样本扩充. 添加了新的测试样本, 并对原有的样本进行随机图像增广, 增广方式包括高斯模糊、运动模糊、对比度调整和图像质量压缩. 扩充后的样本量为原有测试集的 10 倍.

2.4 合成车牌图像训练数据集

为了更好地验证本文算法在混合品类车牌上的性能表现. 本文使用图像合成算法对 CCPD 训练样本进行了大量样本扩充. 主要针对多省份车牌和多品类车牌两个方面进行了扩充.

Wu 等^[2]、Sun 等^[29]、Han 等^[30] 和 Sun 等^[31] 提出基于生成对抗网络 (Generative adversarial networks, GAN)^[32] 的车牌生成算法用于扩充真实样本中稀缺部分的车牌样本. 本文采用了类似于 Wu 等^[2] 的生成方法, 结合图像生成与图像质量退化的两种方式进行训练样本扩充. 合成的品类包括各省份的蓝色车牌、新能源车牌、黄色单双行车牌及其他稀有品类车牌. 图像质量退化方式包括运动模糊、高斯模糊、图像压缩和随机光影. 合成图像的可视化

效果如图 6 所示. 合成的混合品类车牌训练集样本共计 10 万张, 不同省份和不同品类的车牌样本数量分布均衡.



图 6 脚本生成的车牌样本

Fig. 6 License plate samples generated by script

3 实验测试与结果

3.1 实验实施细节

为公平地比较本文算法对车牌识别速度和识别准确率, 以 Zhang 等^[20] 为基线算法, 并在本文实验的软硬件环境中参照该论文细节对其进行了复现实验. 复现的准确率和速度与原文献 [20] 实验结果接近. 为了控制车牌检测算法对车牌识别准确率造成的影响, 本文算法和 GCN^[22] 对比算法均采用与基线算法 Zhang 等^[20] 相同的车牌检测算法进行车牌定位和图像裁剪.

本文所有实验均在相同软硬件环境进行: GPU 为显存 6 GB 的英伟达 1660 s 显卡, CPU 为 AMD3600 处理器. 实验中使用的神经网络框架为 PyTorch, 训练数据的批大小为 32, 学习率为 0.001, 优化器为 Adam, 训练轮数为 150. 输入图像的宽和高分别为 256 像素和 64 像素.

3.2 算法评估指标

车牌识别准确率 R_{LP} 定义为所有测试样本中预测结果正确的样本数 N_C 占测试样本总数 N_{LP} 的比例. 其中车牌中所有字符全部预测正确则判定为该车牌识别正确.

$$R_{LP} = \frac{N_C}{N_{LP}} \quad (7)$$

字符准确率 R_{Char} 定义为所有测试样本中预测正确的单字符个数 M_C 占总样本中所有字符个数 M_{Char} 的比例.

$$R_{Char} = \frac{M_C}{M_{Char}} \quad (8)$$

中文字符准确率 R_{C_Char} 定义为所有测试样本中的中文字符正确个数 M_{C_C} 占所有测试样本中中文字符总数 M_{C_Char} 的比例. 其中中文字符包括各地区简称及“警”、“挂”、“领”、“使”等.

$$R_{C_Char} = \frac{M_{C_C}}{M_{C_Char}} \quad (9)$$

西文字符准确率 R_{W_Char} 定义为所有测试样本中的西文字符预测正确的个数 M_{W_C} 占全部西文字符总数 M_{W_Char} 的比例. 其中西文字符包括数字和英文字符.

$$R_{W_Char} = \frac{M_{W_C}}{M_{W_Char}} \quad (10)$$

3.3 算法识别准确率分析

3.3.1 CCPD 数据集测试评估

与 Zhang 等^[20] 和 Xu 等^[27] 的样本划分方式相同, 表 1 实验过程使用 CCPD 基础集的二分之一样本作为训练数据集, 用于对车牌检测网络和车牌识别网络进行训练. 剩余的基础集二分之一样本和其他各子集样本用于性能测试. 此阶段实验不使用额外合成数据集.

为公平对比各识别算法的有效性, 表 1 实验中 Zhang 等^[20]、GCN^[22] 与本文 CARNet 算法均使用同样的车牌检测算法, 其中 GCN 算法是对 Yan 等^[22] 提出的场景文本识别算法的复现, 训练设置和本文算法保持一致. 从表 1 中可以看出, 本文算法在 CCPD 各测试子集上均超过现有算法的最高准确率. 各子集的平均准确率达到 99.5%, 超越其他算法的最高准确率 0.7%, 并在最难的挑战子集上的识别准确率提升 2%.

表 2 给出了本文算法在 CCPD 测试集上与 GCN 算法^[22] 的各项评估指标比较结果. 括号内为多次实验的准确率的标准差. 由表 2 可以看出, 本文算法在各测试子集上的车牌准确率、字符准确率、中文字符准确率和西文字符准确率, 与对比算法中的较优算法 GCN 算法相比, 均具有显著优越性.

3.3.2 CLPD 数据集测试评估

表 3 实验结果显示, 在仅使用 CCPD 基础集作为训练数据的条件下, CARNet 算法在 CLPD 测试集上的准确率显著高于现有算法. 并且在引入混合数据集训练后识别准确率得到进一步提升. 其中混合数据集由 CCPD 和合成图像组成, 合成图像包含蓝色、绿色和黄色车牌数量相等共计 10 万张.

表 1 在 CCPD 上的车牌识别准确率 (%)
Table 1 License plate recognition accuracy on CCPD (%)

算法	平均	基础集	明暗集	远近集	旋转集	倾斜集	天气集	挑战集
Li 等 ^[1]	94.4	97.8	94.8	94.5	87.9	92.1	86.8	81.2
Xu 等 ^[27]	95.5	98.5	96.9	94.3	90.8	92.5	87.9	85.1
Wang 等 ^[28]	96.6	98.9	96.1	96.4	91.9	93.7	95.4	83.1
Zou 等 ^[8]	97.8	99.3	98.5	98.6	92.5	94.4	<u>99.3</u>	86.6
Yang 等 ^[4]	97.5	99.1	96.9	95.9	97.1	98.0	97.5	85.9
Qin 等 ^[33]	97.5	99.5	93.3	93.7	98.2	95.9	98.9	<u>92.9</u>
Qiao 等 ^[34]	96.9	99.0	97.1	95.5	95.0	96.5	95.9	83.1
Zhang 等 ^[20]	98.5	99.6	98.8	98.8	96.4	97.6	98.5	88.9
Liu 等 ^[35]	98.74	<u>99.73</u>	99.05	<u>99.23</u>	97.62	98.40	98.89	88.51
GCN	<u>98.79</u>	99.70	<u>99.07</u>	98.96	<u>98.33</u>	<u>98.82</u>	98.66	89.42
CARNet	99.50 (0.02)	99.89 (0.01)	99.57 (0.08)	99.56 (0.04)	99.68 (0.04)	99.80 (0.01)	99.38 (0.06)	94.92 (0.09)

表 2 本文算法有效性评估 (%)
Table 2 Evaluation of the effectiveness of the algorithm of this paper (%)

评估指标	算法	平均	基础集	明暗集	远近集	旋转集	倾斜集	天气集	挑战集
R_{LP}	GCN	98.79 (0.10)	99.70 (0.03)	99.07 (0.11)	98.96 (0.11)	98.33 (0.25)	98.82 (0.19)	98.66 (0.13)	89.42 (0.54)
	CARNet	99.50 (0.02)	99.89 (0.01)	99.57 (0.08)	99.56 (0.04)	99.68 (0.04)	99.80 (0.01)	99.38 (0.06)	94.92 (0.09)
R_{Char}	GCN	99.74 (0.02)	99.95 (0.01)	99.83 (0.02)	99.79 (0.01)	99.68 (0.05)	99.78 (0.03)	99.77 (0.02)	97.28 (0.14)
	CARNet	99.90 (0.01)	99.98 (0.01)	99.94 (0.01)	99.93 (0.01)	99.95 (0.01)	99.97 (0.01)	99.90 (0.01)	98.89 (0.01)
R_{C_Char}	GCN	99.72 (0.02)	99.87 (0.01)	99.78 (0.02)	99.78 (0.01)	99.56 (0.06)	99.71 (0.03)	99.70 (0.06)	98.18 (0.07)
	CARNet	99.92 (0.01)	99.99 (0.01)	99.93 (0.02)	99.89 (0.02)	99.95 (0.01)	99.98 (0.01)	99.95 (0.02)	99.13 (0.01)
R_{W_Char}	GCN	99.74 (0.02)	99.97 (0.01)	99.83 (0.02)	99.80 (0.01)	99.70 (0.05)	99.80 (0.03)	99.78 (0.01)	97.13 (0.16)
	CARNet	99.90 (0.01)	99.98 (0.01)	99.94 (0.01)	99.93 (0.01)	99.95 (0.01)	99.97 (0.01)	99.89 (0.01)	98.85 (0.01)

表 3 在 CLPD 上的车牌识别准确率
Table 3 License plate recognition accuracy on CLPD

算法	训练数据	R_{LP} (%)
Zhang 等 ^[20]	CCPD 基础集	70.8
GCN ^[22]	CCPD 基础集	74.4
Zou 等 ^[8]	CCPD 基础集	80.7
Liu 等 ^[35]	CCPD 基础集	80.3
CARNet	CCPD 基础集	82.2
CARNet	混合数据集	83.9

3.3.3 混合品类数据集评估

表 4 实验为本文和对比算法在混合品类车牌数

据集上的识别准确率. 此实验使用 CCPD 蓝色车牌样本和算法合成的多品类车牌样本进行混合训练, 测试集是在 SYSU-ITS 基础上扩充的混合品类测试集. 本文 CARNet 算法和 GCN 对比算法均采用相同的实验设置和车牌检测器. 表 4 结果表明, 本文算法在各品类车牌上的识别准确率均显著高于对比算法, 尤其是在双行大型车后牌上性能优势明显.

3.4 算法时间复杂度分析

3.4.1 GPU 平台算法速度评估

表 5 给出了本文和对比算法在 GPU 平台的识别速度比较结果. 得益于本文采用的并行化解码结

表 4 在混合品类车牌上的识别准确率
Table 4 Recognition accuracy on mixed types of license plates

车牌类别	数量 (张)	GCN ^[22] (%)	CARNet (%)
蓝牌车牌	1050	96.3	99.0
新能源绿牌	1010	41.9	78.5
大型车后牌	660	35.8	62.1
教练车牌	860	44.9	74.8
港澳车牌	1580	53.7	74.9
大型车前牌	980	63.2	75.8

表 5 各算法速度比较
Table 5 Comparison of algorithm speed

车牌识别算法	车牌识别耗时 (ms)
Zhang 等 ^[20]	7.9
GCN ^[22]	18.7
CARNet	4.9

构, 本文算法对单张车牌图像的识别耗时减少至 4.9 ms, 与基线算法 Zhang 等^[20] 串行结构识别耗时 7.9 ms 相比, 本文识别算法的识别速度提升了 38%. 与对比算法中准确率较高的 GCN 算法相比, 单张图像识别速度提升了 74%.

3.4.2 嵌入式平台算法速度评估

本文将 CARNet 车牌识别算法运行在 3 款低功耗设备进行速度测试, 分别为英伟达的 Jetson Nano 和 Jetson TX2, 以及海思的 Hi3516DV300. 前 2 种英伟达平台使用 Pytorch 神经网络框架进行网络推理. Hi3516DV300 采用 NNIE 神经网络框架进行开发部署. 速度对比结果见表 6, 本文的车牌识别算法在 3 种低功耗嵌入式设备速度达到 22FPS 至 33FPS, 可很好地满足低功耗场景下的实时性需求.

表 6 低功耗嵌入式硬件测试
Table 6 Low-power embedded device test

算法	硬件平台	推理引擎	耗时 (ms)
Qin 等 ^[9]	Jetson Nano	TensorFlow	68
CARNet	Jetson Nano	Pytorch	41
CARNet	Jetson TX2	Pytorch	30
CARNet	Hi3516DV300	NNIE	46

3.5 算法消融实验与分析

本节讨论本文算法涉及的改进部分对识别准确率和算法复杂度的影响. 实验中的训练数据集均为 CCPD 基础集的二分之一, 不使用合成数据集. 测试数据集是 CCPD 中基础集的余下二分之一和其

他子集. 表 7 ~ 9 中准确率为全部测试子集的车牌识别准确率的平均值, 参数量为车牌识别网络所需的全部网络参数量, 计算复杂度为车牌识别网络在推理过程中所需的复杂度, 不包含车牌检测网络所需的复杂度.

表 7 特征提取网络消融实验
Table 7 Feature extraction ablation experiment

特征提取	R_{LP} (%)	参数量	计算复杂度 (GMacs)
Resnet45 ^[25]	99.5	13.94 M	14.66
Xception19	99.5	1.87 M	1.71

表 8 分类头参数共享消融实验
Table 8 Classification head weight sharing ablation experiment

参数共享	R_{LP} (%)	参数量	计算复杂度 (GMacs)
是	99.5	1.87 M	1.71
否	99.4	3.82 M	1.71

表 9 单字符注意力消融实验
Table 9 Ablation experiments for single-character attention

单字符注意力	R_{LP} (%)	参数量	计算复杂度 (GMacs)
有	99.5	1.87 M	1.71
无	99.1	3.04 M	1.02

3.5.1 特征提取网络消融实验

表 7 实验结果表明, 轻量化的 19 层类 Xception 结构在保持准确率相近的情况下, 由于深度可分离卷积的设计优势, 可有效地降低参数量和计算复杂度. Xception19 结构比 Resnet45 结构参数量降低 87%, 计算复杂度下降 88%.

3.5.2 分类头参数共享消融实验

如表 8 所示, 参数共享的识别网络整体参数量降低约 51%, 同时车牌识别准确率提升了 0.1%. 这是由于当使用非共享参数的多个分类头时, 多个分类头是独立的单字符分类任务. CCPD 数据集中大部分的样本是在合肥市采集, 第 2 位字符的类别分布存在严重的长尾效应. 当采用共享参数的分类头时, 第 2 位字符和后几位字符的分支共用一组网络参数, 长尾效应得以轻微改善.

3.5.3 字符注意力消融实验

表 9 为单字符注意力消融实验. 由图 5 可以看出, 若取消本文结构中的单字符注意力机制, 则会导致式 (3) 计算出的多个单字符特征向量完全相等, 若用共享参数的分类头则会导致多个分支预测的字符类别相同, 网络模型则无法有效地进行监督

训练. 因此表 9 无注意力对比实验仅可使用参数不共享的多个并行化分类头. 表 9 实验结果表明, 若取消单字符注意力机制可降低网络整体计算复杂度, 但是准确率也会降低 0.4%.

3.6 单字符注意力可视化

为展示本文提出的单字符注意力机制在不同品类车牌的可视化效果, 本文对各品类车牌的字符注

意力图进行了可视化, 如图 7 和图 8 所示. 由于注意力的尺寸为输入车牌图像尺寸的四分之一, 在可视化的过程中将注意力图插值到原始尺寸的大小并转化为热力图与原图叠加. 其中高亮区域代表注意力所关注区域, 亮度越大代表该位置的注意力值接近 1, 反之接近于 0. 图 7 和图 8 中第 1 行图像为原始车牌图像, 第 2 行开始依次为各字符注意力图效果. 可以看出, 每个字符的注意力图与原图中字



图 7 常见 7 字符车牌注意力图

Fig. 7 Attention maps of seven-character license plates



图 8 双行黄牌、新能源车牌及黑色车牌注意力图

Fig. 8 Attention maps of double-line and new energy and black plate licenses

符所在位置一一对应. 图 7 为常见 7 字符车牌图像, 图 7(a) 为低分辨率的车牌, 图 7(b) 为倾斜的蓝色牌照, 图 7(c) 为白色警车车牌, 图 7(d) 为模糊的教练车牌, 图 7(e) 为光线条件较差的黄色车牌.

图 8(a) 为大型汽车双行字符后牌, 且存在角度倾斜, 本文的单字符注意力能够自适应地换行并准确定位第 2 行的首位字符. 图 8(b) 和图 8(c) 分别为新能源小型汽车和新能源大型汽车车牌. 由于这种类型的车牌比其他车牌多出 1 个字符. 相对于前 7 个字符的注意力图的亮度, 第 8 个字符的注意力图亮度偏暗, 这是由于训练数据集中多数的车牌为 7 个字符车牌, 较少部分车牌样本为 8 个字符绿色车牌. 故第 8 个字符多数情况下为“&”占位符, 类别分布非常不均衡. 在图 8(b) 和图 8(c) 中, 最后 1 行可视化效果中新能源车牌的第 8 个字符的注意力仍能够准确聚集于最后 1 个字符, 而在 7 个字符车牌中第 8 个字符注意力亮点无明显聚集. 本文提出的字符注意力机制可以有效地判断第 8 个字符为常规字符或为占位符“&”, 能够实现对不同字符个数的自适应. 综合观察图 7 和图 8 各种品类车牌的注意力图效果, 本文提出的单字符注意力能够实现对不同行数、字数、颜色的车牌进行精准地单字符特征定位.

3.7 复杂场景测试样例

考虑到第 3.3 节算法识别准确率分析实验中测试集的车牌场景固定, 本文收集各种更加复杂场景下的车牌图像进行车牌检测和车牌识别的端到端测试. 测试场景包含各类角度、光照和分辨率. 如图 9 所示, 第 1 行左数第 1 张图的车牌倾斜角度接近 45 度, 具有很大识别难度. 得益于单字符注意力机制, 本文算法能够对倾斜车牌图像中的单字符做到精准的特征分割和分类. 本文提出的识别算法在图 9 各种真实的复杂场景中对各品类车牌能够保持一定的鲁棒性.

3.8 识别失败样例分析

尽管本文算法在公开数据集上获得了比较高的识别准确率, 但在极端场景下仍然存在不足. 图 10 展示了本文实验中的一些失败案例, 可以看出在低分辨率、车牌形变、极端光照条件和污渍遮挡的情况下仍会出现识别错误的情况. 这种错误通常发生在相似字符之间, 如“E”和“F”、“D”和“0”、“7”和“Z”、“A”和“4”等.

4 结束语

针对多品类车牌和复杂场景下的倾斜车牌识别



图 9 真实复杂场景下的检测识别测试
Fig. 9 Detection and recognition test in complex scene



图 10 识别错误示例
Fig. 10 Recognition error cases

难题, 本文提出一种基于单字符注意力的二维特征分割方案, 将车牌识别这个序列识别任务简化为多分支单字符分类任务. 在不需要单字符位置标签监督训练的情况下, 充分发挥二维注意力的优势实现并行化的单字符特征分割. 基于深度可分离卷积设计出更加轻量化的特征提取主干网络, 并且设计了共享参数的多分支结构代替了基于循环神经网络的串行解码器结构, 提升了车牌识别算法的速度, 降低了模型参数量.

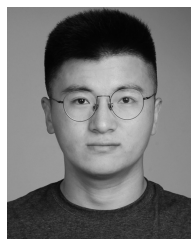
本文提出的 CARNet 算法同时兼容国内绝大多数品类常规和特种车牌识别, 实现对不同字数和行数车牌的兼容识别. 但是稀有品类的车牌在实验中仅使用合成数据集进行监督训练, 所合成的训练样本与真实车牌图像样本仍然存在一定的域差异,

识别准确率低于常规蓝色车牌准确率。未来可通过采集更多训练样本或通过域迁移的方法来提高稀有品类车牌识别效果。

References

- Li H, Wang P, You M Y, Shen C H. Reading car license plates using deep neural networks. *Image and Vision Computing*, 2018, **72**: 14–23
- Wu C H, Xu S G, Song G C, Zhang S Q. How many labeled license plates are needed. In: Proceedings of the Chinese Conference on Pattern Recognition and Computer Vision. Guangzhou, China: 2018. 334–346
- Huang G, Liu Z, Maaten L V, Weinberger K Q. Densely connected convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017. 2261–2269
- Yang Y X, Xi W, Zhu C K, Zhao Y H. HomoNet: Unified license plate detection and recognition in complex scenes. In: Proceedings of the International Conference on Collaborative Computing: Networking, Applications and Worksharing. Shanghai, China: 2020. 268–282
- He M X, Hao P. Robust automatic recognition of Chinese license plates in natural scenes. *IEEE Access*, 2020, **8**: 173804–173814
- Lee C Y, Osindero S. Recursive recurrent nets with attention modeling for ocr in the wild. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 2231–2239
- Cheng Z Z, Bai F, Xu Y L, Zheng G, Pu S L, Zhou S G. Focusing attention: Towards accurate text recognition in natural images. In: Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017. 5086–5094
- Zou Y J, Zhang Y J, Yan J, Jiang X X, Huang T J, Fan H S, et al. A robust license plate recognition model based on bi-LSTM. *IEEE Access*, 2020, **8**: 211630–211641
- Qin S X, Liu S J. Efficient and unified license plate recognition via lightweight deep neural network. *IET Image Processing*, 2020, **14**(16): 4102–4109
- Guo J M, Liu Y F. License plate localization and character segmentation with feedback self-learning and hybrid binarization techniques. *IEEE Transactions on Vehicular Technology*, 2008, **57**(3): 1417–1424
- Silva S M, Jung C R. License plate detection and recognition in unconstrained scenarios. In: Proceedings of the European Conference on Computer Vision. Munich, Germany: 2018. 580–596
- Gou C, Wang K F, Yao Y J, Li Z X. Vehicle license plate recognition based on extremal regions and restricted boltzmann machines. *IEEE Transactions on Intelligent Transportation Systems*, 2016, **17**(4): 1096–1107
- Park S H, Yu S B, Kim J, Yoon H. An all-in-one vehicle type and license plate recognition system using yolov4. *Sensors*, 2022, **22**(3): 921–939
- Wang D, Tian Y M, Geng W H, Zhao L, Gong C. LPR-Net: Recognizing Chinese license plate in complex environments. *Pattern Recognition Letters*, 2020, **130**: 148–156
- Luo C J, Jin L W, Sun Z H. Moran: A multi-object rectified attention network for scene text recognition. *Pattern Recognition*, 2019, **90**: 109–118
- Jaderberg M, Simonyan K, Zisserman A, Kavukcuoglu K. Spatial transformer networks. In: Proceedings of the Annual Conference on Neural Information Processing Systems. Montreal, Canada: 2015. 2017–2025
- Zhou Xiao-Jun, Gao Yuan, Li Chao-Jie, Yang Chun-Hua. Multi-objective optimization based multi-task learning for end-to-end car license plates recognition. *Control Theory and Applications*. 2021, **38**(5): 676–688
(周晓君, 高媛, 李超杰, 阳春华. 基于多目标优化多任务学习的端到端车牌识别方法. 控制理论与应用, 2021, **38**(5): 676–688)
- Li H, Wang P, Shen C H, Zhang G Y. Show, attend and read: A simple and strong baseline for irregular text recognition. In: Proceedings of the 33rd AAAI Conference on Artificial Intelligence. Honolulu, USA: 2019. 8610–8617
- Xu H, Guo Z H, Wang D H, Zhou X D, Shi Y. 2D license plate recognition based on automatic perspective rectification. In: Proceedings of the 25th International Conference on Pattern Recognition. Milan, Italy: IEEE, 2021. 202–208
- Zhang L J, Wang P, Li H, Li Z, Shen C H. A robust attentional framework for license plate recognition in the wild. *IEEE Transactions on Intelligent Transportation Systems*, 2021, **22**(11): 6967–6976
- Hu W Y, Cai X C, Hou J, Yi S, Lin Z P. GTC: Guided training of ctc towards efficient and accurate scene text recognition. In: Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York, USA: 2020. 11005–11012
- Yan R J, Peng L R, Xiao S Y, Yao G. Primitive representation learning for scene text recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE, 2021. 284–293
- Wang T W, Zhu Y Z, Jin L W, Luo C J, Chen X X, Wu Y Q, et al. Decoupled attention network for text recognition. In: Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York, USA: AAAI, 2020. 12216–12224
- Chollet F. Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017. 1251–1258
- Shi B G, Yang M K, Wang X G, Lyu P Y, Yao C, Bai X. Aster: An attentional scene text recognizer with flexible rectification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, **41**(9): 2035–2048
- Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Proceedings of the International Conference on Medical Image Computing and Computer-assisted Intervention. Munich, Germany: 2015. 234–241
- Xu Z B, Yang W, Meng A J, Lu N X, Huang H, Ying C C, et al. Towards end-to-end license plate detection and recognition: A large dataset and baseline. In: Proceedings of the European Conference on Computer Vision. Munich, Germany: 2018. 255–271
- Zhao Y T, Yu Z, Li X Y. Evaluation methodology for license plate recognition systems and experimental results. *IET Intelligent Transport Systems*, 2018, **12**(5): 375–385
- Sun M, Zhou F, Yang C, Yin X C. Image generation framework for unbalanced license plate data Set. In: Proceedings of the International Conference on Data Mining Workshops. Beijing, China: IEEE, 2019. 883–889

- 30 Han B G, Lee J T, Lim K T, Choi D H. License plate image generation using generative adversarial networks for end-to-end license plate character recognition from a small set of real images. *Applied Sciences*, 2020, **10**(8): 2780–2796
- 31 Sun Y F, Liu Q, Chen S L, Zhou F, Yin X C. Robust Chinese license plate generation via foreground text and background separation. In: Proceedings of the 11th International Conference on Image and Graphics. Haikou, China: 2021. 290–302
- 32 Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B. Generative adversarial nets. In: Proceedings of the Advances in Neural Information Processing Systems. Montreal, Canada: 2014. 2672–1680
- 33 Qin S X, Liu S J. Towards end-to-end car license plate location and recognition in unconstrained scenarios. *Neural Computing and Applications*, 2022, **34**: 21551–21566
- 34 Qiao L, Chen Y, Cheng Z Z, Xu Y L, Niu Y, Pu S L, et al. Mango: A mask attention guided one-stage scene text spotter. In: Proceedings of the 35th AAAI Conference on Artificial Intelligence. Virtual Event: 2021. 2467–2476
- 35 Liu Q, Chen S L, Li Z J, Yang C, Chen F, Yin X C. Fast recognition for multidirectional and multi-type license plates with 2D spatial attention. In: Proceedings of the 16th International Conference on Document Analysis and Recognition. Lausanne, Switzerland: 2021. 125–139

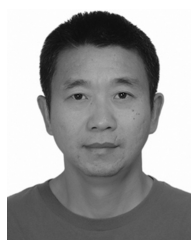


穆世义 上海大学通信与信息工程学院硕士研究生. 主要研究方向为深度学习, 计算机视觉和光学字符识别.

E-mail: mushiyishu@shu.edu.cn

(MU Shi-Yi Master student at the School of Communication and Information Engineering, Shanghai

University. His research interest covers deep learning, computer vision and optical character recognition.)



徐树公 上海大学通信与信息工程学院教授. 主要研究方向为无线通信和模式识别. 本文通信作者.

E-mail: shugong@shu.edu.cn

(XU Shu-Gong Professor at the School of Communication and Information Engineering, Shanghai Uni-

versity. His research interest covers wireless communication and pattern recognition. Corresponding author of this paper.)