

Abstract

由于无监督字符标注方法没有实现, 所以这里先占个位置, 后续再补充。

Introduction

随着车辆数量的增加, 车牌识别技术在交通监控管理、停车场管理、被盗车辆识别、嫌疑车辆追踪等领域中得到了广泛应用。车牌数据集作为车牌识别技术的重要基础资源, 数据集的质量直接影响到整个车牌识别系统的性能。现有的数据集如 KETI-ALPR [1] 和 KarPlate [2] 在一定程度上取得了成功, 但这些数据集规模较小, 存在缺乏多样性, 缺乏字符级别标注等问题。Han et al. [3] 通过基于生成对抗网络(GAN)的 LP-GAN 方法, 生成了大量的车牌数据, 但仍没有从根本上解决数据集的多样性问题。Wang et al. [4] 提出了一个包含真实和合成车牌的数据集, 通过 2D 贴图的方式生成合成数据, 但在合成数据中没有考虑到凸起的车牌字符边缘的光照、阴影等因素。随着车牌规则的更新和新能源车牌的发行, 车牌字符数量和种类的增加。Park et al. [5] 的数据集丰富了韩国车牌的多样性, 并且提供了更多地区的前缀, 仍存在部分地区车牌数据缺失的问题。

Table 1 是上述数据集的对比信息, 包含了数据的类型, 数据集的规模, 是否包含多行车牌, 是否包含字符级别标注。从对比结果可以看出, 现有的车牌数据集大多数缺乏字符级别标注, 或者缺乏数据多样性。

Dataset	Type	Scale	Multi-line	Char-level Annotation
Han et al. [3]	Real + Synthetic	18,000	No	No
KETI-ALPR [1]	Real	3,000	Yes	No
KarPlate [2]	Real	4,267	Yes	No
Wang et al. [4]	Real + Synthetic	10,500+500,000	Yes	No
Park et al. [5]	Real	6,878	Yes	Yes
Ours	Real + Synthetic	36,245+100,000	Yes	Yes

Table 1: Comparison of the existing datasets

因此, 本文提出了一个由真实和合成数据构成的新数据集 SKLP136, 包含单双行车牌以及字符级别的标注信息, 并对新能源车牌和缺少的部分地区的车牌数据进行了补充。最后, 本文针对嵌入式设备上的车牌识别任务, 提出了一个轻量级的车牌识别模型 TinyLPR, 用于提高车牌识别的速度和准确率。

SKLP136 Dataset

我们收集了室内外不同场景, 白天夜晚不同光照条件下的 36,245 张真实的韩国车牌数据, Figure 1 展示了部分具有代表性的单双行, 不同类型的韩国真实的车牌数据。



Figure 1: Different license plate types including single-line and multi-line

为了进一步丰富不同地区的车牌数据, 我们借助 Blender 软件开发了一个车牌合成器, 合成了韩国不同类型的共计 100,000 张车牌合成数据 (如图 x, 这里缺个图)。相较于 Wang et al. [4] 的 2D 贴图的传统合成方法, 我们的合成器使用 3D 模型, 考虑了不同光照条件下的车牌凸起字符边缘的反光、阴影等因素, 并且针对夜间车牌灯对车牌的光照影响, 进行了相对应的调整, 使合成数据更加真实。(这里补一个局部特写图用于与 2D 贴图作对照)。



Figure 2: Comparison of 2D texture mapping and 3D model synthesis

SKLP136 包含了不同类型的单行和双行车牌, 包含了韩国所有地区的前缀, 以及字符级的标注信息, 能够更好地适应韩国车牌的识别任务。Table 2 展示了 SKLP136 中韩文字符的种类, 包括数字、韩文字符和地区前缀。

Type	Characters	Total
Numerical	0,1,2,3,4,5,6,7,8,9	10
Korean alphabetic	가 (Ga), 나 (Na), 다 (Da), 라 (Ra), 마 (Ma), 거 (Geo), 너 (Neo), 더 (Deo), 러 (Reo), 머 (Meo), 버 (Beo), 서 (Seo), 어 (Eo), 저 (Jeo), 고 (Go), 노 (No), 도 (Do), 로 (Ro), 모 (Mo), 보 (Bo), 소 (So), 오 (O), 조 (Jo), 구 (Gu), 누 (Nu), 두 (Du), 루 (Ru), 무 (Mu), 부 (Bu), 수 (Su), 우 (U), 주 (Ju), 하 (Ha), 허 (Heo), 호 (Ho), 바 (Ba), 사 (Sa), 아 (A), 자 (Ja), 배 (Bae)	40
Area prefixes	서울 (Seoul), 부산 (Busan), 대구 (Daegu), 인천 (Incheon), 광주 (Gwangju), 대전 (Daejeon), 울산 (Ulsan), 세종 (Sejong), 경기 (Gyeonggi), 강원 (Gangwon), 충북 (Chungbuk), 충남 (Chungnam), 전북	17

	(Jeonbuk),전남 (Jeonnam),경북 (Gyeongbuk),경남 (Gyeongnam),제주 (Jeju)	
--	--	--

Table 2: South Korean license plate corresponding characters

如 Table 3 所示, SKLP136 数据集提供了共计 136,245 张车牌数据, 其中 122,620 张用于训练, 13,625 用于验证。数据集中单行车牌和双行车牌的比例约为 7:3, 其中单行车牌占 93,222 张, 双行车牌占 43,023 张。数据集中包含了 10 个数字字符, 40 个韩文字符和 17 个地区前缀, 共计 67 个字符种类, Table 4 展示了 SKLP136 数据集中数字、韩文字符和地区前缀的分布情况。

Type	Training	Validation	Total
Single-line	83,855	9,367	93,222
Multi-line	38,765	4,258	43,023
Total	122,620	13,625	136,245

Table 3: Collected data for different types of license plates

Class	Character	Training	Validation	Total
1	1	64800	7141	71941
2	2	78340	8737	87077
3	3	86043	9719	95762
4	4	75401	8413	83814
5	5	76880	8623	85503
6	6	77770	8657	86427
7	7	82295	9080	91375
8	8	75489	8316	83805
9	9	71975	7861	79836
10	0	59892	6612	66504
11	가 (Ga)	2420	267	2687
12	나 (Na)	2405	289	2694
13	다 (Da)	2425	274	2699
14	라 (Ra)	2390	293	2683
15	마 (Ma)	2489	262	2751
16	거 (Geo)	2511	287	2798
17	너 (Neo)	2542	282	2824
18	더 (Deo)	2418	275	2693
19	러 (Reo)	2380	249	2629
20	머 (Meo)	2340	272	2612
21	버 (Beo)	2571	253	2824
22	서 (Seo)	2443	268	2711
23	어 (Eo)	2515	305	2820
24	저 (Jeo)	2456	288	2744
25	고 (Go)	2598	301	2899
26	노 (No)	2555	296	2851

Class	Character	Training	Validation	Total
27	ㄷ (Do)	2704	297	3001
28	ㄹ (Ro)	2473	291	2764
29	ㅁ (Mo)	2675	321	2996
30	ㅂ (Bo)	2379	271	2650
31	ㅅ (So)	2545	282	2827
32	ㅇ (O)	2491	239	2730
33	ㅈ (Jo)	2743	304	3047
34	ㅊ (Gu)	2333	261	2594
35	ㅊ (Nu)	2584	284	2868
36	ㄷ (Du)	2505	285	2790
37	ㄷ (Ru)	2302	267	2569
38	ㅁ (Mu)	2343	268	2611
39	ㅂ (Bu)	2668	286	2954
40	ㅅ (Su)	2491	291	2782
41	ㅇ (U)	2573	274	2847
42	ㅈ (Ju)	2400	262	2662
43	ㅊ (Ha)	1975	238	2213
44	ㅊ (Heo)	1482	150	1632
45	ㅊ (Ho)	2293	242	2535
46	ㅂ (Ba)	12786	1391	14177
47	ㅅ (Sa)	5414	571	5985
48	ㅅ (A)	8233	907	9140
49	ㅈ (Ja)	5382	589	5971
50	ㅂ (Bae)	5388	593	5981
51	서울 (Seoul)	2764	326	3090
52	부산 (Busan)	2016	212	2228
53	대구 (Daegu)	2109	210	2319
54	인천 (Incheon)	3140	355	3495
55	광주 (Gwangju)	1932	206	2138
56	대전 (Daejeon)	1941	220	2161
57	울산 (Ulsan)	2023	215	2238
58	세종 (Sejong)	1735	182	1917
59	경기 (Gyeonggi)	12270	1389	13659
60	강원 (Gangwon)	2022	225	2247
61	충북 (Chungbuk)	1965	195	2160
62	충남 (Chungnam)	2049	212	2261
63	전북 (Jeonbuk)	2026	213	2239
64	전남 (Jeonnam)	1997	186	2183
65	경북 (Gyeongbuk)	1953	218	2171

Class	Character	Training	Validation	Total
66	경 남 (Gyeongnam)	1924	236	2160
67	제 주 (Jeju)	1926	215	2141
Total		917,297	101,799	1,019,096

Table 4: Korean license plate dataset characters distribution

XXXXXXXX

接下来是表述如何进行字符级别的标注的流程，但由于无监督的方法还没有实现，所以这里先占个位置，后续再补充。

下面是一些字符级别标注的效果图片，先拿别人的占个位置，后续再补充。



Figure 3: Character-level annotation of license plate characters

以下是关于车牌识别的方法的 **related works**，后续再补充。

针对多行车牌的识别，主流的方法分为两类：基于检测器的两阶段方法和端到端的方法。

Two-stage Methods

基于检测器的两阶段方法通常先使用检测器检测车牌的字符，根据返回的每个字符的预测结果和边界框的位置来对车牌识别的最终结果进行归纳[2], [5], [6]。此类方法需要对车牌的每个字符进行分别标注，数据的标注工作量大，并且字符检测阶段的错误会直接影响到识别阶段，产生误差累积，从而导致整体车牌识别性能的下降。

End-to-end Methods

端到端的方法通常将车牌识别任务作为一个整体识别问题，直接输入车牌图片即可识别出车牌字符。这类方法大多依赖于 RNN, LSTM 等序列模型，通常不需要字符级别的标注，但是需要更多的计算资源，因此速度较慢。对于计算资源极度有限的嵌入式设备，基于序列模型的方法通常不适用。

References

- [1] J.-Y. Sung and S.-B. Yu, "Real-time automatic license plate recognition system using YOLOv4," in *2020 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)*, 2020, pp. 1–3.
- [2] C. Henry, S. Y. Ahn, and S.-W. Lee, "Multinational license plate recognition using generalized character sequence detection," *Ieee Access*, vol. 8, pp. 35185–35199, 2020.
- [3] B.-G. Han, J. T. Lee, K.-T. Lim, and D.-H. Choi, "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*, vol. 10, no. 8, p. 2780–2781, 2020.
- [4] H. Wang, Y. Li, L.-M. Dang, and H. Moon, "Robust Korean license plate recognition based on deep neural networks," *Sensors*, vol. 21, no. 12, p. 4140–4141, 2021.

- [5] S.-H. Park, S.-B. Yu, J.-A. Kim, and H. Yoon, "An all-in-one vehicle type and license plate recognition system using YOLOv4," *Sensors*, vol. 22, no. 3, p. 921–922, 2022.
- [6] B. Wang, H. Xiao, J. Zheng, D. Yu, and C. L. P. Chen, "Character Segmentation and Recognition of Variable-Length License Plates Using ROI Detection and Broad Learning System," *Remote Sensing*, vol. 14, no. 7, p. 1560–1561, 2022.