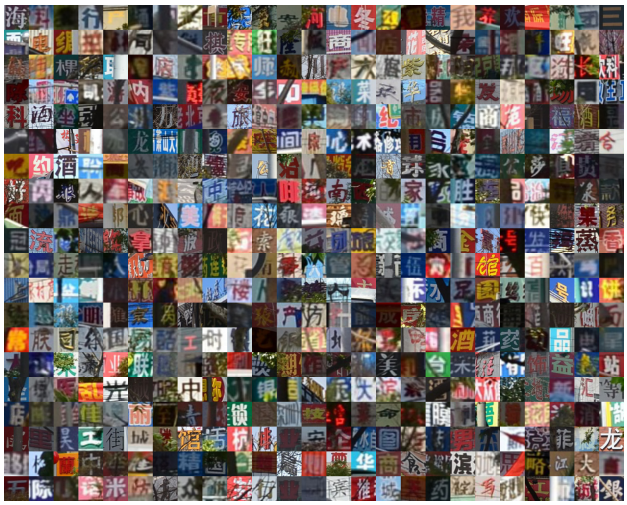
Supplementary Material for the Paper Entitled "A Large Chinese Text Dataset in the Wild"

This supplementary material provides more details of the dataset, recognition results, end-to-end detection and recognition results in the paper entitled "A Large Chinese Text Dataset in the Wild".

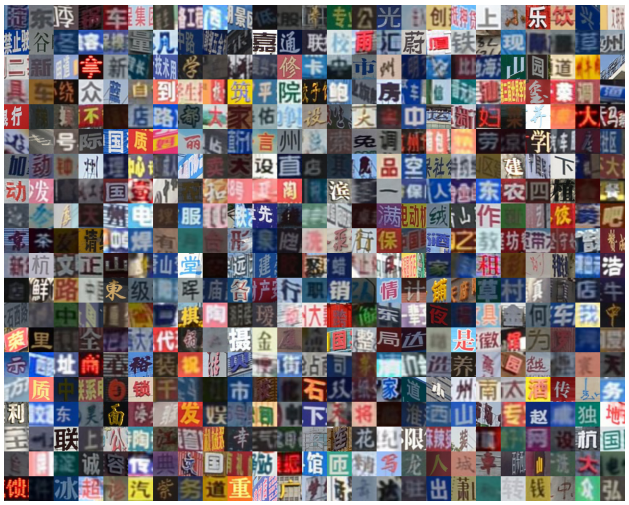
In Figure 1, we provide more Chinese character instances in our dataset, with different attributes. In Figure 2, we provide more images in our dataset with corresponding annotation. In Table 1, we provide the number of character instances with different attribute combination and different sizes in the training and testing set, respectively. All above information demonstrate that our dataset is a challenging dataset due to its large size, diversity and complexity.

For the task of character recognition from cropped regions, in Table 2, we show the recognition results of more randomly chosen character instances using different methods. In Table 3, we give the top-1 accuracy of 50 randomly chosen character categories using different methods. In Table 4, we also give the top-1 accuracy using Google Inception for character instances with different attribute combination and different sizes, respectively.

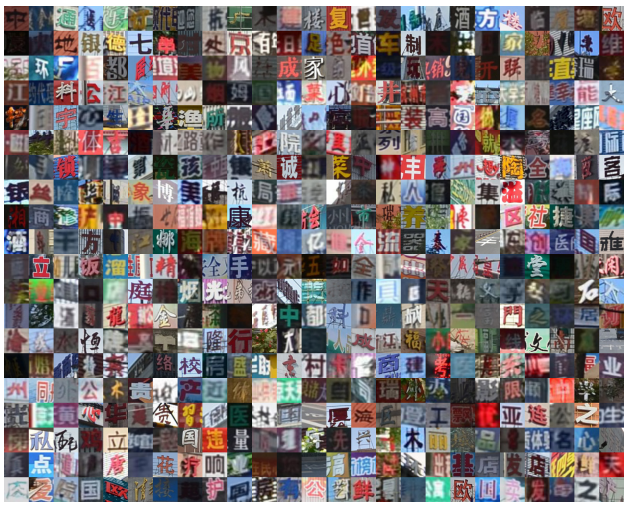
For the task of character detection from images using YOLOv2 and SSD512, we provide the AP of 50 randomly chosen character categories in Table 5. We also provide the recall rates of character instances with different attribute combinations and different sizes, respectively, in Table 6. More detection results of YOLOv2 and SSD512 are given in Figure 3 and Figure 4, respectively.



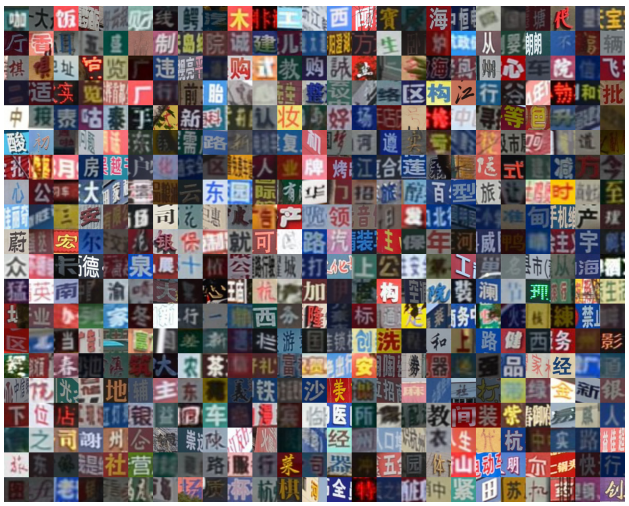
(a) occluded



(b) not occluded

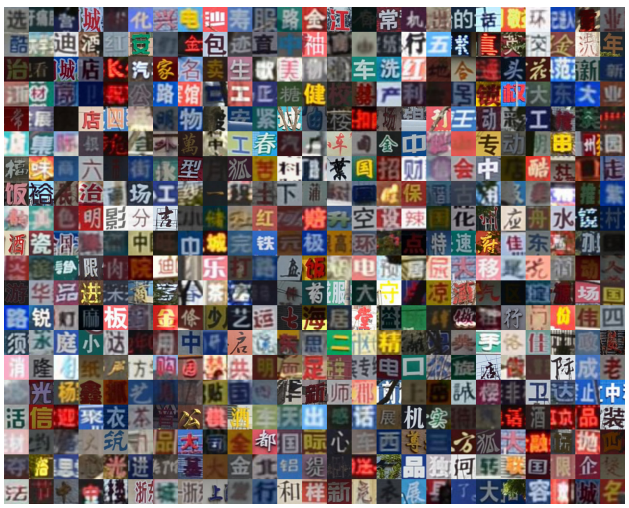


(c) complex background

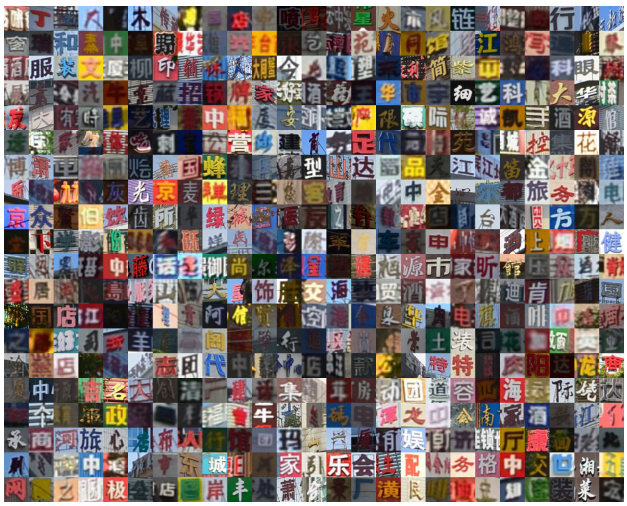


(d) clean background

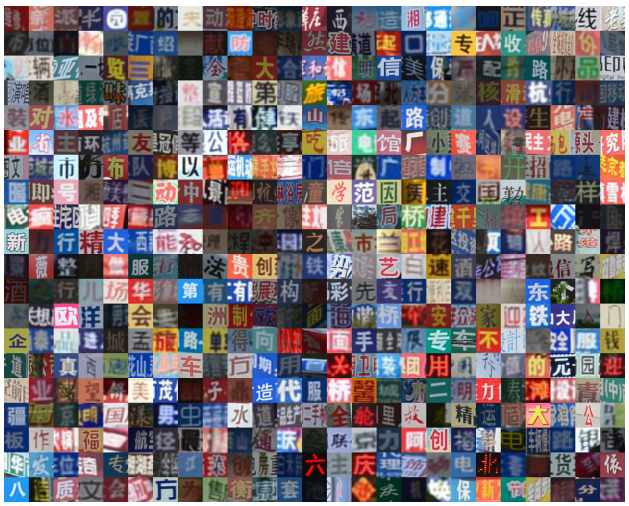
(e) distorted



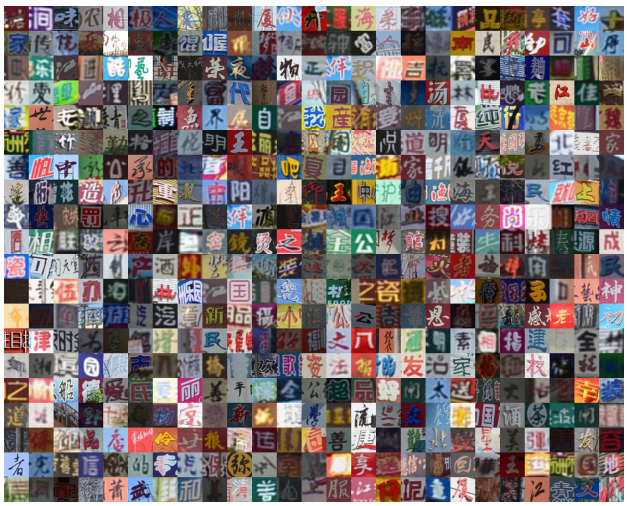
(f) frontal



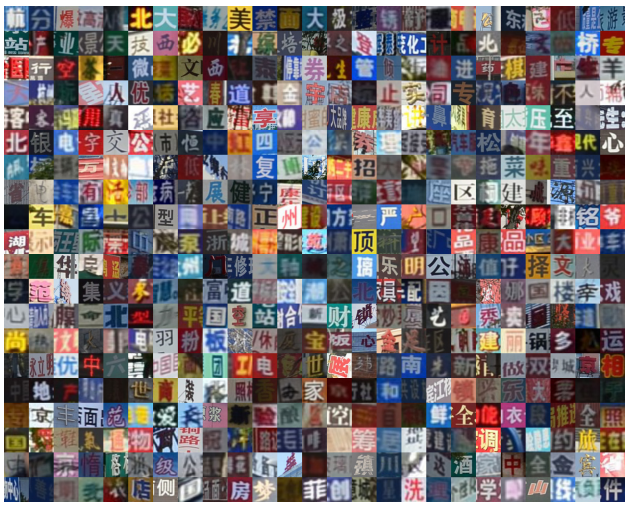
(g) 3D raised



(h) planar



(i) wordart



(j) not wordart

(k) handwritten

(l) printed

Fig. 1: More Chinese character instances with different attributes in our dataset.

打造东山最大的服饰折扣广场
爱尚城1-10

KTV

Fig. 2: More Images in our dataset with corresponding annotation. Character instances are shown in green while ‘ignore’ regions are shown in yellow.

TABLE 1: The number of character instances with different attribute combination and different sizes in the training and testing set, respectively. From left to right, we provide the attribute combination, the number of character instances in training/testing set, the number of character instances with large size in training/testing set, the number of character instances with medium size in training/testing set, the number of character instances with small size in training/testing set. Small, medium, and large refer to character size < 16, $\in [16, 32)$ and ≥ 32, respectively. $\sim$ denotes without a specific attribute, & denotes logical and. e.g., occluded & $\sim$bgcomplex denotes character instances which have the occluded attribute but do not have the complex background attribute.

attribute combination	all	large	medium	small
occluded & bgcomplex	36367 / 7480	18308 / 3190	14471 / 3387	3588 / 903
occluded & distorted	25184 / 6646	10283 / 2594	11879 / 3155	3022 / 897
occluded & raised	44377 / 11290	21564 / 4468	19000 / 5482	3813 / 1340
occluded & wordart	10112 / 1888	5023 / 946	4285 / 776	804 / 166
occluded & handwritten	615 / 200	277 / 78	283 / 103	55 / 19
occluded & $\sim$bgcomplex	73199 / 17757	24392 / 5305	36141 / 9003	12666 / 3449
occluded & $\sim$distorted	84382 / 18591	32417 / 5901	38733 / 9235	13232 / 3455
occluded & $\sim$raised	65189 / 13947	21136 / 4027	31612 / 6908	12441 / 3012
occluded & $\sim$wordart	99454 / 23349	37677 / 7549	46327 / 11614	15450 / 4186
occluded & $\sim$handwritten	108951 / 25037	42423 / 8417	50329 / 12287	16199 / 4333
bgcomplex & distorted	58917 / 15843	23030 / 6151	26780 / 7290	9107 / 2402
bgcomplex & raised	111061 / 27240	55990 / 11788	44716 / 12475	10355 / 2977
bgcomplex & wordart	28670 / 5296	12485 / 2264	12038 / 2328	4147 / 704
bgcomplex & handwritten	2595 / 713	1147 / 312	1208 / 332	240 / 69
bgcomplex & $\sim$occluded	194486 / 47115	70431 / 15309	83504 / 21561	40551 / 10245
bgcomplex & $\sim$distorted	171936 / 38752	65709 / 12348	71195 / 17658	35032 / 8746
bgcomplex & $\sim$raised	119792 / 27355	32749 / 6711	53259 / 12473	33784 / 8171
bgcomplex & $\sim$wordart	202183 / 49299	76254 / 16235	85937 / 22620	39992 / 10444
bgcomplex & $\sim$handwritten	228258 / 53882	87592 / 18187	96767 / 24616	43899 / 11079
distorted & raised	57825 / 16838	22910 / 6731	28312 / 8070	6603 / 2037
distorted & wordart	17229 / 3787	6839 / 1579	7587 / 1523	2803 / 685
distorted & handwritten	938 / 322	408 / 92	435 / 174	95 / 56
distorted & $\sim$occluded	183066 / 50030	52800 / 13645	88469 / 23987	41797 / 12398
distorted & $\sim$bgcomplex	149333 / 40833	40053 / 10088	73568 / 19852	35712 / 10893
distorted & $\sim$raised	150425 / 39838	40173 / 9508	72036 / 19072	38216 / 11258
distorted & $\sim$wordart	191021 / 52889	56244 / 14660	92761 / 25619	42016 / 12610
distorted & $\sim$handwritten	207312 / 56354	62675 / 16147	99913 / 26968	44724 / 13239
raised & wordart	19331 / 4524	9508 / 2182	8393 / 1909	1430 / 433
raised & handwritten	62 / 36	26 / 32	36 / 4	0 / 0
raised & $\sim$occluded	170310 / 47516	67994 / 16725	76868 / 23205	25448 / 7586
raised & $\sim$bgcomplex	103626 / 31566	33568 / 9405	51152 / 16212	18906 / 5949
raised & $\sim$distorted	156862 / 41968	66648 / 14462	67556 / 20617	22658 / 6889

Continued on next page

TABLE 1 – continued from previous page

attribute combination	all	large	medium	small
raised & ∼wordart	195356 / 54282	80050 / 19011	87475 / 26778	27831 / 8493
raised & ∼handwritten	214625 / 58770	89532 / 21161	95832 / 28683	29261 / 8926
wordart & handwritten	203 / 35	91 / 21	73 / 10	39 / 4
wordart & ∼occluded	60416 / 12164	22154 / 4451	27813 / 5389	10449 / 2324
wordart & ∼bgcomplex	41858 / 8756	14692 / 3133	20060 / 3837	7106 / 1786
wordart & ∼distorted	53299 / 10265	20338 / 3818	24511 / 4642	8450 / 1805
wordart & ∼raised	51197 / 9528	17669 / 3215	23705 / 4256	9823 / 2057
wordart & ∼handwritten	70325 / 14017	27086 / 5376	32025 / 6155	11214 / 2486
handwritten & ∼occluded	6238 / 1520	2181 / 502	3010 / 700	1047 / 318
handwritten & ∼bgcomplex	4258 / 1007	1311 / 268	2085 / 471	862 / 268
handwritten & ∼distorted	5915 / 1398	2050 / 488	2858 / 629	1007 / 281
handwritten & ∼raised	6791 / 1684	2432 / 548	3257 / 799	1102 / 337
handwritten & ∼wordart	6650 / 1685	2367 / 559	3220 / 793	1063 / 333
∼occluded & ∼bgcomplex	508820 / 133178	104588 / 24867	231293 / 59406	172939 / 48905
∼occluded & ∼distorted	520240 / 130263	122219 / 26531	226328 / 56980	171693 / 46752
∼occluded & ∼raised	532996 / 132777	107025 / 23451	237929 / 57762	188042 / 51564
∼occluded & ∼wordart	642890 / 168129	152865 / 35725	286984 / 75578	203041 / 56826
∼occluded & ∼handwritten	697068 / 178773	172838 / 39674	311787 / 80267	212443 / 58832
∼bgcomplex & ∼distorted	432686 / 110102	88927 / 20084	193866 / 48557	149893 / 41461
∼bgcomplex & ∼raised	478393 / 119369	95412 / 20767	216282 / 52197	166699 / 46405
∼bgcomplex & ∼wordart	540161 / 142179	114288 / 27039	247374 / 64572	178499 / 50568
∼bgcomplex & ∼handwritten	577761 / 149928	127669 / 29904	265349 / 67938	184743 / 52086
∼distorted & ∼raised	447760 / 106886	87988 / 17970	197505 / 45598	162267 / 43318
∼distorted & ∼wordart	551323 / 138589	134298 / 28614	240550 / 61573	176475 / 48402
∼distorted & ∼handwritten	598707 / 147456	152586 / 31944	262203 / 65586	183918 / 49926
∼raised & ∼wordart	546988 / 137196	110492 / 24263	245836 / 60414	190660 / 52519
∼raised & ∼handwritten	591394 / 145040	125729 / 26930	266284 / 63871	199381 / 54239
∼wordart & ∼handwritten	735694 / 189793	188175 / 42715	330091 / 86399	217428 / 60679

TABLE 2: More examples of the recognition task. The character instances are randomly chosen from the testing set. In each row, from left to right, we give: the cropped region of a character instance, the ground truth character category, and the recognition results of different methods. Correct recognitions are painted with green. The number shows the confidence (%) of the results.

line No.	instance	category	AlexNet	OverFeat	Inception	ResNet50	ResNet152
1		场	场 98.6	场 100.0	场 100.0	场 99.6	场 100.0
2		车	车 100.0	车 100.0	车 99.7	车 98.8	车 99.4
3		色	色 100.0	色 100.0	色 100.0	色 100.0	色 100.0
4		银	银 99.8	银 99.9	银 96.4	银 98.7	银 96.1
5		路	路 78.4	路 84.8	路 95.2	路 98.5	路 96.0
6		司	司 93.1	司 99.3	司 99.7	司 99.8	司 99.9
7		宇	字 8.1	字 55.5	字 74.3	容 42.1	字 92.7
8		限	限 97.2	限 100.0	限 99.9	限 100.0	限 99.9
9		饰	饰 100.0	饰 100.0	饰 100.0	饰 100.0	饰 100.0
10		向	奥 57.1	奥 38.7	向 85.0	向 91.6	向 98.3
11		埠	维 43.3	峰 22.9	维 51.3	啡 21.3	维 44.3
12		全	公 41.3	全 55.4	全 71.5	全 47.1	全 58.6
13		道	持 1.6	梦 7.9	辆 26.4	号 18.8	警 14.1
14		州	州 100.0	州 100.0	州 100.0	州 100.0	州 100.0
15		持	持 100.0	持 100.0	持 98.9	持 99.8	持 99.5
16		娱	程 91.0	程 77.0	程 42.8	媒 42.5	媒 49.4
17		宏	宏 100.0	宏 100.0	宏 99.6	宏 100.0	宏 100.0
18		培	培 99.6	培 100.0	培 99.7	培 99.4	培 100.0
19		关	关 100.0	关 100.0	关 99.2	关 96.9	关 97.9
20		搜	州 6.9	州 10.2	捷 35.9	推 86.9	推 65.2
21		区	区 100.0	区 100.0	区 100.0	区 99.9	区 99.9
22		京	京 98.7	京 99.9	京 99.1	京 99.7	京 99.6
23		式	式 100.0	式 100.0	式 100.0	式 100.0	式 100.0
24		产	严 57.8	产 72.8	产 80.1	产 46.0	产 73.7
25		农	农 95.3	农 100.0	农 99.9	农 99.9	农 99.8
26		文	幼 10.6	公 26.6	公 26.5	公 37.0	公 33.5
27		食	食 96.1	食 91.9	食 99.4	食 79.1	食 91.2
28		虫	中 18.4	少 57.9	电 33.5	串 20.4	安 55.8
29		东	东 12.2	东 92.9	东 59.3	车 90.8	车 75.7
30		行	行 100.0	行 100.0	行 100.0	行 100.0	行 100.0
31		前	前 99.9	前 99.9	前 99.7	前 92.7	前 94.1
32		行	行 100.0	行 100.0	行 100.0	行 100.0	行 100.0

Continued on next page

TABLE 2 – continued from previous page

line No.	instance	category	AlexNet	OverFeat	Inception	ResNet50	ResNet152
33		合	合 99.9	合 100.0	合 99.0	合 98.2	合 98.9
34		车	车 100.0	车 100.0	车 99.8	车 99.8	车 99.5
35		开	开 100.0	开 100.0	开 100.0	开 99.7	开 99.9
36		宝	库 5.1	圣 12.3	车 23.0	季 7.5	作 24.1
37		支	支 23.3	支 46.5	支 32.5	支 79.1	支 93.1
38		定	定 100.0	定 100.0	定 99.9	定 100.0	定 100.0
39		时	村 25.6	时 37.7	时 53.5	时 77.9	村 48.2
40		林	林 100.0	林 100.0	林 100.0	林 99.1	林 98.1
41		腾	腾 100.0	腾 100.0	腾 100.0	腾 100.0	腾 100.0
42		第	第 100.0	第 100.0	第 99.9	第 99.7	第 99.8
43		联	联 100.0	联 100.0	联 99.9	联 99.9	联 99.9
44		楼	楼 98.2	楼 100.0	楼 100.0	楼 97.4	楼 94.7
45		公	公 100.0	公 100.0	公 100.0	公 100.0	公 100.0
46		印	印 100.0	印 100.0	印 100.0	印 99.9	印 100.0
47		葵	类 47.5	委 40.0	烫 12.6	烫 66.5	卖 24.8
48		华	华 99.1	华 99.7	华 100.0	华 100.0	华 100.0
49		子	子 98.8	子 100.0	子 99.7	子 99.1	子 99.4
50		培	培 99.0	培 100.0	培 99.4	培 98.5	培 99.8

TABLE 3: The top-1 accuracy (%) of 50 character categories randomly chosen from the testing set. From left to right, we give: the character category, the top-1 accuracy in percentage achieved by AlexNet, OverFeat, Google Inception, ResNet50 and ResNet152, respectively, and the number of instances of each character category in the testing and in the training set, respectively. For each character category, among different methods, the highest accuracy is highlighted in bold.

line No.	category	AlexNet	Overfeat	Inception	ResNet50	ResNet152	testing	training
1	新	78.3	82.2	**87.9**	84.9	87.3	535	4810
2	市	75.1	77.3	**86.3**	85.5	83.9	578	4450
3	南	79.4	82.4	**88.8**	85.5	87.4	427	3399
4	区	90.7	91.6	**94.2**	92.2	94.0	450	3225
5	话	76.2	80.1	**86.6**	85.4	85.4	357	3189
6	西	84.8	86.4	89.1	88.7	**89.4**	433	3102
7	招	83.3	87.9	**90.9**	87.1	84.8	132	2034
8	精	68.0	68.8	**80.8**	69.6	76.8	125	1619
9	成	79.9	81.5	**92.4**	87.0	87.5	184	1559
10	元	82.7	86.4	88.0	86.9	**89.5**	191	1403
11	售	81.1	83.3	**88.6**	85.6	84.1	132	1252
12	体	77.2	85.1	85.1	85.1	**86.8**	114	888

Continued on next page

TABLE 3 – continued from previous page

line No.	category	AlexNet	Overfeat	Inception	ResNet50	ResNet152	testing	training
13	米	91.2	80.7	89.5	**93.0**	91.2	57	868
14	请	79.2	82.3	**89.2**	86.2	**89.2**	130	822
15	间	83.7	88.2	**93.5**	87.6	92.8	153	748
16	云	83.1	85.5	**89.2**	86.7	86.7	83	727
17	贸	71.1	71.9	**82.2**	77.0	77.0	135	608
18	险	85.9	**87.5**	**87.5**	**87.5**	**87.5**	64	497
19	國	62.1	71.8	**76.7**	64.1	70.9	103	494
20	岛	80.0	80.0	**89.4**	83.5	87.1	85	482
21	视	68.1	72.3	**80.9**	74.5	76.6	47	435
22	敬	88.2	**94.1**	**94.1**	**94.1**	**94.1**	17	430
23	武	60.0	55.0	**65.0**	60.0	60.0	20	415
24	由	82.4	88.2	76.5	**94.1**	88.2	17	412
25	先	79.3	81.0	81.0	**82.8**	77.6	58	383
26	苑	80.0	85.5	**89.1**	85.5	85.5	55	367
27	智	89.1	94.5	**98.2**	94.5	**98.2**	55	356
28	勤	73.7	84.2	**89.5**	73.7	84.2	19	318
29	干	63.8	61.7	66.0	**70.2**	63.8	47	300
30	恩	42.9	61.9	66.7	**71.4**	**71.4**	21	280
31	寓	79.3	**82.8**	**82.8**	79.3	**82.8**	29	266
32	要	57.5	57.5	**65.0**	**65.0**	62.5	40	242
33	责	54.5	**90.9**	**90.9**	**90.9**	**90.9**	11	241
34	任	**87.5**	**87.5**	**87.5**	**87.5**	**87.5**	8	222
35	藏	85.7	85.7	**90.5**	**90.5**	**90.5**	21	217
36	委	81.8	81.8	**86.4**	**86.4**	**86.4**	22	203
37	饺	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	4	203
38	膜	83.3	83.3	**100.0**	**100.0**	**100.0**	6	190
39	莱	14.3	**28.6**	21.4	14.3	14.3	14	190
40	套	78.3	**95.7**	**95.7**	**95.7**	**95.7**	23	189
41	致	75.0	75.0	**93.8**	81.2	**93.8**	16	188
42	燕	81.6	**86.8**	**86.8**	81.6	78.9	38	180
43	玻	92.3	92.3	**97.4**	94.9	**97.4**	39	176
44	边	50.0	60.0	**70.0**	60.0	60.0	10	164
45	扬	82.4	82.4	**88.2**	**88.2**	82.4	17	160
46	蒸	**92.9**	78.6	85.7	85.7	78.6	14	149
47	溪	72.7	76.4	**85.5**	83.6	**85.5**	55	146
48	仙	57.9	63.2	**78.9**	73.7	**78.9**	19	143
49	泽	68.8	68.8	**75.0**	**75.0**	**75.0**	16	133
50	韩	**83.3**	77.8	77.8	77.8	77.8	18	131

TABLE 4: The top-1 accuracy (%) using Google Inception for character instances with different attribute combination and different sizes. From left to right, we provide the attribute combination, the top-1 accuracy of all character instances, top-1 accuracy in percentage of character instances with large, medium, and small sizes, respectively. Small, medium, and large refer to character size < 16, $\in [16, 32)$ and ≥ 32, respectively. $\sim$ denotes without a specific attribute, & denotes logical and. e.g., occluded & $\sim$bgcomplex denotes character instances which have the occluded attribute but do not have the complex background attribute.

attribute combination			all	large	medium	small
occluded	&	bgcomplex	61.5	64.5	60.6	55.8
occluded	&	distorted	59.6	61.3	60.2	52.7
occluded	&	raised	64.0	67.8	63.9	51.6
occluded	&	wordart	54.1	53.7	54.8	52.9
occluded	&	handwritten	41.9	52.6	48.1	18.8
occluded	&	∼bgcomplex	65.6	68.1	66.7	58.6
occluded	&	∼distorted	66.1	69.8	66.6	59.3
occluded	&	∼raised	64.7	65.9	66.0	60.2
occluded	&	∼wordart	65.2	68.6	65.8	58.1
occluded	&	∼handwritten	64.5	66.9	65.2	58.2
bgcomplex	&	distorted	73.4	73.6	72.3	75.8
bgcomplex	&	raised	75.2	78.7	73.1	70.0
bgcomplex	&	wordart	67.4	63.7	68.6	75.8
bgcomplex	&	handwritten	55.6	62.1	54.1	40.5
bgcomplex	&	∼occluded	78.0	80.5	77.7	75.0
bgcomplex	&	∼distorted	76.8	80.4	76.5	72.7
bgcomplex	&	∼raised	76.3	76.9	77.4	74.4
bgcomplex	&	∼wordart	76.8	80.3	76.0	73.2
bgcomplex	&	∼handwritten	76.0	78.1	75.5	73.6
distorted	&	raised	71.5	75.7	69.3	62.9
distorted	&	wordart	61.7	63.2	60.9	59.5
distorted	&	handwritten	48.0	52.9	51.0	38.8
distorted	&	∼occluded	76.5	80.2	77.0	71.5
distorted	&	∼bgcomplex	75.1	79.3	76.1	69.0
distorted	&	∼raised	75.7	78.1	77.3	71.2
distorted	&	∼wordart	75.8	79.0	76.2	71.0
distorted	&	∼handwritten	74.8	77.2	75.4	70.5
raised	&	wordart	63.0	61.6	64.8	62.6
raised	&	handwritten	83.3	83.3	-	-
raised	&	∼occluded	79.0	82.6	78.7	72.1
raised	&	∼bgcomplex	77.1	80.8	77.7	69.0
raised	&	∼distorted	78.0	81.7	78.0	70.8
raised	&	∼wordart	77.4	81.9	76.6	69.7
raised	&	∼handwritten	76.2	79.7	75.8	69.3

Continued on next page

TABLE 4 – continued from previous page

attribute combination			all	large	medium	small
wordart	&	handwritten	20.0	-	20.0	-
wordart	&	∼occluded	71.1	71.6	71.3	69.8
wordart	&	∼bgcomplex	69.9	72.8	69.6	66.6
wordart	&	∼distorted	72.2	71.7	72.4	72.8
wordart	&	∼raised	71.5	72.9	71.0	70.2
wordart	&	∼handwritten	69.0	68.6	69.4	69.0
handwritten	&	∼occluded	57.0	68.2	56.8	47.4
handwritten	&	∼bgcomplex	56.1	71.4	57.2	46.5
handwritten	&	∼distorted	58.0	69.2	57.9	47.2
handwritten	&	∼raised	55.8	66.5	56.2	45.7
handwritten	&	∼wordart	56.4	66.9	57.2	45.7
∼occluded	&	∼bgcomplex	81.8	85.7	83.9	77.4
∼occluded	&	∼distorted	82.5	85.6	84.6	78.4
∼occluded	&	∼raised	81.5	84.5	83.7	77.6
∼occluded	&	∼wordart	81.6	85.3	83.2	77.3
∼occluded	&	∼handwritten	81.1	83.9	82.6	77.2
∼bgcomplex	&	∼distorted	81.9	84.7	84.0	78.1
∼bgcomplex	&	∼raised	80.8	83.8	82.9	77.0
∼bgcomplex	&	∼wordart	80.7	83.9	82.5	76.6
∼bgcomplex	&	∼handwritten	80.3	83.0	81.9	76.5
∼distorted	&	∼raised	81.5	84.1	83.8	78.1
∼distorted	&	∼wordart	81.2	84.6	82.8	77.3
∼distorted	&	∼handwritten	80.8	83.4	82.3	77.4
∼raised	&	∼wordart	80.6	83.2	82.7	76.9
∼raised	&	∼handwritten	80.2	82.3	82.2	76.9
∼wordart	&	∼handwritten	80.0	82.8	81.2	76.3

TABLE 5: AP (%) of 50 character categories randomly chosen from the testing set using YOLOv2 and SSD512. From left to right, we provide the character category, AP of all character instances, AP of character instances with large, medium, and small sizes, respectively, and the number of character instances in the testing set and training set, respectively.

		YOLOv2				SSD512					
line No.	**category**	**all**	**large**	**medium**	**small**	**all**	**large**	**medium**	**small**	**testing**	**training**
1	新	83.4	91.0	85.8	69.1	78.8	84.4	82.1	65.2	447	4810
2	市	82.6	94.6	87.3	61.4	79.0	90.5	79.3	69.8	521	4450
3	南	80.6	92.1	87.9	69.7	74.3	78.3	78.0	74.0	373	3399
4	区	80.4	78.4	88.0	71.8	76.5	78.0	83.4	68.2	460	3225
5	话	81.5	88.9	94.9	69.5	69.8	79.3	86.3	57.1	357	3189
6	西	81.8	89.8	88.7	64.9	79.9	87.9	84.0	68.9	363	3102
7	招	85.1	90.2	95.3	60.4	78.3	89.8	86.7	52.4	184	2034
8	成	73.0	73.5	84.8	66.0	69.5	74.4	76.7	67.1	229	1559
9	精	79.8	100.0	82.6	70.7	73.4	98.2	74.9	63.3	173	1619
10	元	76.3	80.4	87.2	64.5	72.0	84.6	78.5	61.7	192	1403
11	售	76.6	92.8	84.9	59.4	75.2	88.3	83.6	58.5	183	1252
12	体	79.8	89.6	83.4	63.3	79.7	90.4	81.3	69.1	157	888
13	请	75.3	86.7	91.1	57.5	68.2	79.0	78.9	58.0	119	822
14	米	71.8	70.2	76.8	69.9	69.0	58.6	76.9	67.9	76	868
15	间	71.2	86.0	80.9	58.0	72.5	83.7	85.8	63.6	65	748
16	云	70.3	69.7	74.4	54.2	62.6	62.4	65.4	50.2	81	727
17	贸	74.2	98.0	78.7	44.1	65.7	88.9	73.4	37.1	58	608
18	國	48.2	71.0	29.6	0.4	32.1	51.4	14.1	0.5	60	494
19	险	65.1	88.9	68.8	49.0	63.4	72.7	67.4	63.7	46	497
20	岛	49.5	94.9	51.0	11.0	54.3	97.6	45.4	32.7	39	482
21	先	79.8	75.7	84.6	70.3	74.6	61.0	81.4	64.3	105	383
22	视	65.3	86.5	64.0	61.2	63.4	85.7	57.4	78.5	57	435
23	苑	74.5	81.4	80.3	33.9	64.0	73.3	71.3	21.8	55	367
24	敬	63.5	62.5	78.5	12.5	37.7	42.9	48.6	5.6	15	430
25	智	76.8	87.2	85.1	15.7	73.1	83.6	79.7	32.2	50	356
26	武	82.5	91.7	83.0	61.1	75.5	81.2	87.5	61.1	26	415
27	由	65.9	100.0	70.2	58.5	70.6	100.0	76.1	70.3	31	412
28	干	48.0	86.3	40.8	44.3	34.9	73.4	35.8	8.3	22	300
29	勤	37.5	38.5	62.4	31.1	39.4	46.7	46.9	50.0	18	318
30	恩	82.4	97.8	80.9	80.0	67.9	82.4	72.1	60.1	52	280
31	寓	70.4	84.2	78.1	65.9	15.8	46.4	14.9	21.4	26	266
32	要	69.4	82.7	87.2	27.7	67.7	76.8	90.9	28.4	32	242
33	责	85.9	95.8	94.1	61.2	64.5	80.5	69.1	45.0	60	241
34	任	90.9	92.8	95.3	83.7	72.2	67.7	79.3	71.2	46	222
35	套	67.8	100.0	87.7	37.9	52.3	77.8	64.6	35.8	39	189
36	藏	53.7	-	51.5	66.4	37.8	-	37.1	58.0	13	217
37	委	55.5	-	75.6	54.4	48.1	-	91.7	42.9	17	203
38	燕	52.2	100.0	54.1	25.4	22.1	100.0	21.0	1.2	19	180

Continued on next page

TABLE 5 – continued from previous page

		YOLOv2				SSD512					
line No.	**category**	**all**	**large**	**medium**	**small**	**all**	**large**	**medium**	**small**	**testing**	**training**
39	玻	61.9	69.0	64.7	43.3	33.1	40.8	39.2	4.8	19	176
40	溪	74.8	62.9	86.1	71.0	52.8	40.8	56.4	75.3	28	146
41	致	58.1	81.0	72.2	59.6	35.6	79.2	54.4	47.8	20	188
42	饺	79.3	67.7	83.1	100.0	38.3	34.9	41.4	100.0	13	203
43	膜	59.9	100.0	62.3	0.0	37.0	50.0	39.5	0.0	10	190
44	莱	52.1	100.0	7.1	-	1.1	1.0	1.6	-	2	190
45	扬	79.0	74.4	100.0	100.0	60.4	66.7	100.0	7.7	5	160
46	边	88.1	100.0	80.0	100.0	77.3	100.0	76.2	100.0	8	164
47	泽	67.0	85.0	76.7	36.9	63.4	100.0	73.0	27.8	24	133
48	蒸	35.4	-	42.6	-	25.0	-	25.0	-	4	149
49	仙	50.5	100.0	1.4	-	26.0	50.0	4.2	-	2	143
50	韩	89.7	100.0	100.0	62.5	52.5	75.0	71.4	9.6	13	131

TABLE 6: The recall rates (%) of character instances with different attribute combinations and different sizes using YOLOv2 and SSD512. From left to right, we provide the attribute combination, the recall rate of all character instances, the recall rates of character instances with large, medium, and small sizes, respectively. Small, medium, and large refer to character size < 16, $\in [16, 32)$ and ≥ 32, respectively. $\sim$ denotes without a specific attribute, & denotes logical and. e.g., occluded & $\sim$bgcomplex denotes character instances which have the occluded attribute but do not have the complex background attribute.

			YOLOv2				SSD512			
attribute combination			**all**	**large**	**medium**	**small**	**all**	**large**	**medium**	**small**
occluded	&	bgcomplex	55.7	60.4	53.3	44.1	53.4	60.3	47.9	40.2
occluded	&	distorted	55.3	57.3	56.9	44.6	46.7	50.2	46.7	38.6
occluded	&	raised	57.8	60.8	57.2	52.4	53.8	57.8	51.2	48.7
occluded	&	wordart	49.9	50.2	46.8	44.8	44.6	46.7	41.0	40.6
occluded	&	handwritten	45.7	50.8	42.1	0.0	44.2	49.2	44.7	0.0
occluded	&	$\sim$bgcomplex	57.5	62.2	58.9	48.6	51.4	55.6	53.0	43.8
occluded	&	$\sim$distorted	57.5	63.0	57.4	48.6	53.8	59.9	53.5	44.3
occluded	&	$\sim$raised	56.2	62.4	57.4	45.3	50.4	56.8	51.9	40.1
occluded	&	$\sim$wordart	57.5	62.8	58.0	47.9	52.5	58.6	52.2	43.3
occluded	&	$\sim$handwritten	57.1	61.6	57.5	47.8	52.0	57.5	51.7	43.2
bgcomplex	&	distorted	62.1	65.7	63.0	49.4	53.8	58.7	53.7	45.7
bgcomplex	&	raised	70.1	76.1	67.6	56.2	65.7	73.4	61.8	52.1
bgcomplex	&	wordart	60.2	64.1	60.8	46.6	56.1	59.8	54.7	42.6
bgcomplex	&	handwritten	47.3	51.0	41.9	43.8	42.9	46.4	39.0	46.9
bgcomplex	&	$\sim$occluded	69.7	76.8	71.5	56.0	64.6	72.6	65.6	53.2
bgcomplex	&	$\sim$distorted	70.0	77.3	71.7	56.6	66.8	75.5	67.5	53.9
bgcomplex	&	$\sim$raised	65.1	68.6	70.6	54.7	60.0	63.5	64.8	52.2

Continued on next page

TABLE 6 – continued from previous page

attribute	combination		YOLOv2 all	YOLOv2 large	YOLOv2 medium	YOLOv2 small	SSD512 all	SSD512 large	SSD512 medium	SSD512 small
bgcomplex	&	∼wordart	68.5	74.8	69.9	55.6	63.7	71.5	64.1	52.8
bgcomplex	&	∼handwritten	68.1	74.2	69.5	55.2	63.3	70.8	63.6	52.2
distorted	&	raised	64.6	70.8	64.1	49.8	55.8	62.3	55.3	43.4
distorted	&	wordart	51.3	56.5	52.2	34.8	44.5	49.9	44.5	31.1
distorted	&	handwritten	44.1	46.6	42.3	42.9	43.4	46.6	38.5	57.1
distorted	&	∼occluded	64.9	72.0	67.7	53.8	55.3	60.7	57.4	49.9
distorted	&	∼bgcomplex	64.4	72.1	67.7	54.0	54.4	59.3	57.0	49.8
distorted	&	∼raised	63.3	68.9	67.4	54.0	53.5	56.6	56.4	50.5
distorted	&	∼wordart	64.5	70.7	67.0	54.0	54.8	59.8	56.6	49.9
distorted	&	∼handwritten	63.9	69.9	66.4	53.2	54.3	59.2	56.2	49.1
raised	&	wordart	60.2	64.2	55.0	60.7	55.3	59.8	49.2	52.3
raised	&	handwritten	46.7	34.6	100.0	-	63.3	50.0	100.0	-
raised	&	∼occluded	73.7	80.4	73.9	59.5	68.1	76.0	67.9	54.4
raised	&	∼bgcomplex	71.0	75.9	73.3	59.4	64.9	70.0	67.3	54.2
raised	&	∼distorted	73.1	78.2	73.6	61.5	69.3	76.0	69.0	57.2
raised	&	∼wordart	71.4	77.3	71.8	58.2	66.1	73.3	65.9	53.6
raised	&	∼handwritten	70.6	76.1	70.7	58.4	65.3	72.0	64.8	53.5
wordart	&	handwritten	44.0	33.3	-	75.0	60.0	47.6	-	75.0
wordart	&	∼occluded	62.1	67.9	62.2	51.4	56.7	62.6	55.2	46.4
wordart	&	∼bgcomplex	60.5	65.0	59.9	52.6	54.3	59.6	52.5	47.4
wordart	&	∼distorted	63.0	66.9	62.5	56.3	58.0	62.3	55.9	51.0
wordart	&	∼raised	60.5	65.1	62.9	48.2	54.8	59.6	55.4	44.2
wordart	&	∼handwritten	60.4	64.9	60.2	50.7	54.9	59.8	53.4	45.8
handwritten	&	∼occluded	55.9	59.1	56.0	41.8	44.5	48.4	42.2	41.8
handwritten	&	∼bgcomplex	62.0	66.9	65.2	36.8	46.3	51.4	46.4	34.2
handwritten	&	∼distorted	56.3	60.0	56.0	39.7	44.7	48.9	43.7	38.1
handwritten	&	∼raised	54.5	59.6	53.0	40.0	43.8	48.4	42.1	40.0
handwritten	&	∼wordart	54.5	59.3	53.5	37.9	44.0	48.6	42.7	37.9
∼occluded	&	∼bgcomplex	72.7	81.4	78.4	63.1	67.0	74.3	72.3	60.4
∼occluded	&	∼distorted	74.6	83.4	80.2	64.0	70.7	80.2	76.0	61.7
∼occluded	&	∼raised	71.1	78.9	77.6	62.2	65.6	71.7	71.4	60.0
∼occluded	&	∼wordart	72.5	81.0	77.4	62.2	67.0	75.0	71.4	59.6
∼occluded	&	∼handwritten	72.0	79.8	76.6	61.9	66.5	74.0	70.6	59.2
∼bgcomplex	&	∼distorted	73.1	80.4	79.1	64.3	69.0	76.3	74.9	61.8
∼bgcomplex	&	∼raised	70.7	78.5	76.5	62.4	65.1	71.1	70.4	60.0
∼bgcomplex	&	∼wordart	71.4	79.3	76.6	62.3	65.7	72.1	70.6	59.6
∼bgcomplex	&	∼handwritten	70.8	77.8	75.8	62.1	65.1	70.9	69.8	59.3
∼distorted	&	∼raised	71.9	80.0	78.6	63.2	68.0	75.9	74.6	61.0
∼distorted	&	∼wordart	73.0	80.9	78.1	63.1	69.2	78.0	74.1	60.7
∼distorted	&	∼handwritten	72.4	79.5	77.2	63.0	68.6	76.4	73.1	60.5
∼raised	&	∼wordart	70.2	77.7	76.1	61.7	64.7	70.6	70.2	59.3
∼raised	&	∼handwritten	69.8	76.6	75.6	61.3	64.3	69.8	69.6	58.9
∼wordart	&	∼handwritten	70.7	77.8	74.9	61.2	65.3	72.2	69.0	58.5

Fig. 3: More detection results by YOLOv2. For each image, we give the detected character instances and their bounding boxes. Correct detections are shown in green while wrong detections are shown in yellow.

Fig. 4: More detection results by SSD512. For each image, we give the detected character instances and their bounding boxes. Correct detections are shown in green while wrong detections are shown in yellow.