摘要: 在测试集上评估图像分类算法精度

1. 实验准备

编写评估代码、挑选训练完的模型文件

2. 实验步骤

2.1. 运行准备

导入所需工具包、图像预处理、载入测试集、导入模型,编写评估代码。

2.2. 评估代码

主要分为以下指标:

1. 图像分类预测结果

D:/SanYeC云南省	1 云南省	0.97279 广西省	0.01956 贵州省	0.00703	0.97279	0.01956	0	0.00052	0.00703	0.0001
):/SanYeC广西省	1 广西省	0.98155 陕西省	0.01817 贵州省	0.00016	8.00E-05	0.98155	0	4.00E-05	0.00016	0.01817
):/SanYeC广西省	1 广西省	0.50931 陕西省	0.4885 浙江省	0.0008	0.00069	0.50931	0.00017	0.0008	0.00053	0.4885
):/SanYeC广西省	1 广西省	0.79198 陕西省	0.20777 云南省	0.00012	0.00012	0.79198	0	0.00012	1.00E-05	0.20777
:/SanYeC广西省	1 广西省	0.67181 陕西省	0.32766 云南省	0.00041	0.00041	0.67181	4.00E-05	4.00E-05	4.00E-05	0.32766
):/SanYeC广西省	1 广西省	0.9725 陕西省	0.02603 云南省	0.00127	0.00127	0.9725	3.00E-05	5.00E-05	0.00011	0.02603
:/SanYeC广西省	1 广西省	0.94027 陕西省	0.0394 云南省	0.01939	0.01939	0.94027	0.00011	0.00038	0.00045	0.0394
:/SanYeC广西省	1 广西省	0.95716 贵州省	0.04134 陕西省	0.00099	6.00E-05	0.95716	0	0.00044	0.04134	0.00099
):/SanYeC广西省	1 广西省	0.98196 陕西省	0.01529 云南省	0.00143	0.00143	0.98196	0	0.00071	0.00061	0.01529
:/SanYeC广西省	1 广西省	0.79888 陕西省	0.18859 贵州省	0.01182	0.00023	0.79888	0	0.00048	0.01182	0.18859
:/SanYeC广西省	1 广西省	0.91726 陕西省	0.07383 贵州省	0.00806	0.00034	0.91726	2.00E-05	0.0005	0.00806	0.07383
):/SanYeC广西省	1 广西省	0.98655 陕西省	0.00802 云南省	0.00394	0.00394	0.98655	2.00E-05	0.0001	0.00136	0.00802
):/SanYeC广西省	1 广西省	0.92929 陕西省	0.06963 云南省	0.00088	0.00088	0.92929	0.00012	1.00E-05	7.00E-05	0.06963
):/SanYeC广西省	1 广西省	0.97868 陕西省	0.01832 云南省	0.00271	0.00271	0.97868	1.00E-05	0.00012	0.00016	0.01832
):/SanYeC广西省	1 广西省	0.6597 陕西省	0.33575 云南省	0.00227	0.00227	0.6597	6.00E-05	0.00209	0.00013	0.33575
:/SanYeC广西省	1 陕西省	0.79111 广西省	0.20302 贵州省	0.00435	4.00E-05	0.20302	2.00E-05	0.00147	0.00435	0.79111
):/SanYeC广西省	1 广西省	0.69866 陕西省	0.30117 云南省	8.00E-05	8.00E-05	0.69866	2.00E-05	6.00E-05	0	0.30117
):/SanYeC广西省	1 广西省	0.96759 陕西省	0.03225 贵州省	0.0001	1.00E-05	0.96759	0	4.00E-05	0.0001	0.03225
):/SanYeC广西省	1 广西省	0.87385 陕西省	0.12602 云南省	0.0001	0.0001	0.87385	1.00E-05	2.00E-05	2.00E-05	0.12602
):/SanYeC广西省	1 广西省	0.99718 陕西省	0.0027 未知	5.00E-05	4.00E-05	0.99718	5.00E-05	2.00E-05	2.00E-05	0.0027
):/SanYeC广西省	1 广西省	0.99917 陕西省	0.00075 贵州省	6.00E-05	1.00E-05	0.99917	0	0	6.00E-05	0.00075
:/SanYeC云南省	1 云南省	0.98748 广西省	0.01184 贵州省	0.00052	0.98748	0.01184	0	0.00014	0.00052	2.00E-05
):/SanYeC广西省	1 广西省	0.99457 贵州省	0.00426 陕西省	0.00088	0.00026	0.99457	1.00E-05	2.00E-05	0.00426	0.00088
):/SanYeC广西省	1 广西省	0.99667 陕西省	0.00283 云南省	0.00038	0.00038	0.99667	4.00E-05	2.00E-05	6.00E-05	0.00283
):/SanYeC广西省	1 广西省	0.99151 陕西省	0.00832 贵州省	0.00015	0	0.99151	0	1.00E-05	0.00015	0.00832
):/SanYeC广西省	1 广西省	0.97732 陕西省	0.02009 贵州省	0.00247	8.00E-05	0.97732	2.00E-05	1.00E-05	0.00247	0.02009
):/SanYeC广西省	1 广西省	0.98838 陕西省	0.01096 贵州省	0.0006	2.00E-05	0.98838	0	3.00E-05	0.0006	0.01096
):/SanYeC广西省	1 广西省	0.7421 陕西省	0.25787 云南省	2.00E-05	2.00E-05	0.7421	0	1.00E-05	0	0. 25787
):/SanYeC广西省	1 广西省	0.98236 陕西省	0.01761 云南省	1.00E-05	1.00E-05	0.98236	0	1.00E-05	1.00E-05	0.01761
):/SanYeC广西省	1 广西省	0.94964 陕西省	0.05035 云南省	1.00E-05	1.00E-05	0.94964	0	0	0	0.05035
):/SanYeC广西省	1 广西省	0.68774 陕西省	0.31223 云南省	1.00E-05	1.00E-05	0.68774	1.00E-05	0	0	0.31223
:/SanYeC广西省	1 广西省	0.99339 陕西省	0.00651 云南省	9.00E-05	9.00E-05	0.99339	1.00E-05	0	1.00E-05	0.00651
):/SanYeC广西省	1 陕西省	0.53501 广西省	0.46499 云南省	0	0	0.46499	0	0	0	0.53501
):/SanYeC广西省	1 广西省	0.99674 陕西省	0.00288 未知	0.00027	6.00E-05	0.99674	0.00027	4.00E-05	1.00E-05	0.00288
):/SanYeC广西省	1 广西省	0.70904 陕西省	0.2818 云南省	0.00708	0.00708	0.70904	0.00073	0.00124	0.0001	0.2818
D:/SanYeC广西省	1 广西省	0.93107 陕西省	0.06873 云南省	0.00016	0.00016	0.93107	1.00E-05	1.00E-05	4.00E-05	0.06873
D:/SanYeC广西省	1 广西省	0.9945 陕西省	0.0033 贵州省	0.00201	0.00019	0.9945	0	0	0.00201	0.0033

图 1

- 2. 测试集整体准确率
- 3. top-n 准确率

图 2

- 4. 分类报告
- 5. 各类别准确率 (recall)

	pr	ecision	recall fi	-score	support	
	云南省		0.90	0.93	220	
	广西省	0.89	0.96	0.93	298	
	未知	1.00	1.00	1.00	164	
	折江省	0.95	0.92	0.94	197	
	贵州省	0.93	0.93	0.93	200	
	陕西省	0.91	0.85	0.88		
accur	асу			0.94	1174	
macro	avg	0.94	0.93	0.93	1174	
weighted	avg	0.94	0.94	0.94	1174	
100%		6/6 [00:0	0<00:00, 10	9.32it/s	l	
		ecision	recall fi	l-score	support ac	curacy
云南省		0.947619	0.904545	0.925581	220.0	0.90454
广西省		0.893750	0.959732	0.925566	298.0	0.95973
未知		1.000000	1.000000	1.000000	164.0	1.000000
浙江省		0.947917	0.923858	0.935733	197.0	0.92385
贵州省		0.934673	0.930000	0.932331	200.0	0.93000
陕西省		0.910112	0.852632	0.880435		0.85263
macro avg	0	.939012 0	.928461 0.	933274	1174.0 0.	928461

图 3

6. 混淆矩阵

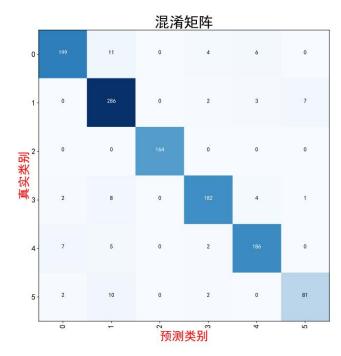
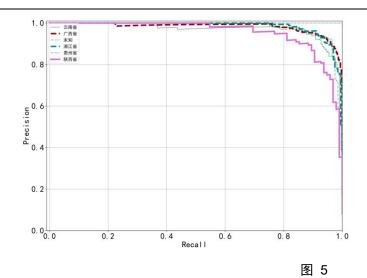
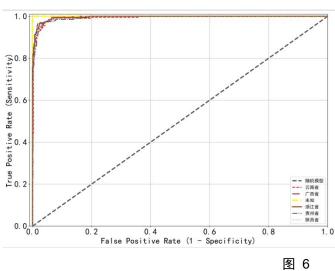


图 4

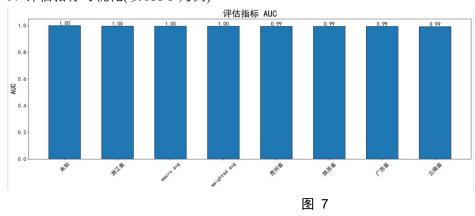
7. PR 曲线







9. 评估指标可视化(以 AUC 为例)



3. 实验结果

将各项指标整理到表格文件、保存图像等

class_2.json
class_5,json
class_10.json
■ kaggle_test_predict.ipynb
S 测试集预测结果.csv
₽ 各类别PR曲线,pdf
₽ 各类别ROC曲线.pdf
₽ 各类别评估指标柱状图-AUC.pdf
S 各类别准确率评估指标.csv
♂ 各类别准确率评估指标柱状图-recall.pdf
₽ 混淆矩阵.pdf

图 8