

街景 patch 分类研究过程梳理

1、对图像进行缩放及裁剪

 <code>_01_resize.py</code>	
 <code>_00_sv</code>	 <code>_01_sv_resize_crop_result</code>
原始街景文件夹	1、resize 到 800*400，尺寸可自由调整 2、居中保留图片中心 800*400 部分
	<div><div> location_120 .151699,30.2 0601532_he ad0_pitch1...</div><div> location_120 .151699,30.2 0601532_he ad90_pitch...</div><div> location_120 .151699,30.2 0601532_he ad180_pitch...</div><div> location_120 .151699,30.2 0601532_he ad270_pitch...</div><div> location_120 .156949,30.2 1349094_he ad0_pitch1...</div><div> location_120 .156949,30.2 1349094_he ad90_pitch...</div><div> location_120 .156949,30.2 1349094_he ad180_pitch...</div><div> location_120 .156949,30.2 1349094_he ad270_pitch...</div></div> <div><div> location_120 .157262,30.2 1372579_he ad0_pitch1...</div><div> location_120 .157262,30.2 1372579_he ad90_pitch...</div><div> location_120 .157262,30.2 1372579_he ad180_pitch...</div><div> location_120 .157262,30.2 1372579_he ad270_pitch...</div><div> location_120 .160266,30.2 3934918_he ad0_pitch1...</div><div> location_120 .160266,30.2 3934918_he ad90_pitch...</div><div> location_120 .160266,30.2 3934918_he ad180_pitch...</div><div> location_120 .160266,30.2 3934918_he ad270_pitch...</div></div> <div><div> location_120 .163274,30.2 135979_he d90_pitch15...</div><div> location_120 .163274,30.2 135979_he d180_pitch...</div><div> location_120 .163274,30.2 135979_he d270_pitch...</div><div> location_120 .164091,30.2 3866929_he ad0_pitch1...</div><div> location_120 .164091,30.2 3866929_he ad90_pitch...</div><div> location_120 .164091,30.2 3866929_he ad180_pitch...</div><div> location_120 .164091,30.2 3866929_he ad270_pitch...</div></div> <div><div> location_120 .151699,30.2 0601532_he ad0_pitch1...</div><div> location_120 .151699,30.2 0601532_he ad90_pitch...</div><div> location_120 .151699,30.2 0601532_he ad180_pitch...</div><div> location_120 .151699,30.2 0601532_he ad270_pitch...</div><div> location_120 .156949,30.2 1349094_he ad0_pitch1...</div><div> location_120 .156949,30.2 1349094_he ad90_pitch...</div><div> location_120 .156949,30.2 1349094_he ad180_pitch...</div><div> location_120 .156949,30.2 1349094_he ad270_pitch...</div></div>


2、对街景分 patch

可自由设置把一张街景分割多少行、多少列。

<


3、语义分析


为了保证语义分析的准确性，这里对街景源文件和 patch 后的文件，分别进行 ade20k（支持 150 类别）语义分析。可用于优化后面街景 patch 的聚类分析，也可用于其他分析。

**_03_imgs_seg.py**

语义标签色标，见 ade20k_color 文件。


**_03_sv_ss**

**resize_crop_ss**

**split_patch_ss**



**ss_result_HZSVI.csv**


**ss_PATCH_result_HZSVI.csv**


该语义分析共包含 150 中分类，


id	wall	building:edifice	sky	floor:flooring	tree	ceiling	road:route	bed	window:pane:window	grass	cabinet	sidewalk:pavement
location_120.1499625.30	0.03041875	0.157121875	0.050521875	0	0.372246875	0	0.098853125	0	0	0.046353125	0	0.003659375
location_120.1499625.30	0.012975	0.000175	0.02931875	0	0.557159375	0	0.187925	0	0	0.13948125	0	0.0690625
location_120.1499625.30	0.02920625	0.080659375	0.21208125	0	0.253228125	0	0.264265625	0	0.037215625	0	0.009621875	0
location_120.1499625.30	0.001784375	0.000709375	0.157075	0	0.471640625	0	0.236465625	0	0.094503125	0	0.032440625	0
location_120.1500626.30	0.344929125	0.055809375	0.001875	0	0.479975	0	0.06228125	0	0	0	0	0.029384375
location_120.1500626.30	0.2008875	0.020146875	0.037940625	0	0.523853125	0	0.085046875	0	0	0	0	0
location_120.1500626.30	0.3509375	0.002053125	0.02493125	0	0.556240625	0	0.05955625	0	0	0	0	0
location_120.1500626.30	0.060140625	0.070628125	0.018240625	0	0.54684375	0	0.0776125	0	0	0	0	0.003828125
location_120.1503159.30	0.3193875	0	0.1551875	0	0.383303125	0	0.09588125	0	0	0	0	0.000103125
location_120.1503159.30	0.048809375	0.06828125	0.036440625	0	0.5248875	0	0.0777375	0	0	0.001190625	0	0
location_120.1503159.30	0.346165625	0	0.21160625	0	0.34945625	0	0.07620625	0	0	0	0	0
location_120.1503159.30	0.113928125	0.0025	0.044684375	0	0.527	0	0.099825	0	0	0.000775	0	0.000134375
location_120.1503504.30	0.02886875	0.21778125	0.07000625	0	0.292465625	0	0.1918625	0	0.037834375	0	0.00125625	0
location_120.1503504.30	0.009128125	0.0043625	0.031340625	0	0.52739375	0	0.19011875	0	0	0.1659625	0	0.06754375
location_120.1503504.30	0.0083	0.004396875	0.096046875	0	0.44085	0	0.33	0	0.078078125	0	0.022628125	0
location_120.1503504.30	0.01230625	0.085965625	0.16559375	0	0.37715	0	0.214584375	0	0	0.056875	0	0.026165625
location_120.1506485.30	0.34671875	0	0	0	0.511046875	0	0.028975	0	0	0	0	0.027871875
location_120.1506485.30	0.001828125	0	0.04245625	0	0.70154375	0	0.064146875	0	0	0	0	0
location_120.1506485.30	0.35360625	0	0.052903125	0	0.489446875	0	0.064778125	0	0	0	0	0
location_120.1506485.30	0.068396875	0	0	0	0.697125	0	0.084028125	0	0	0	0	0.0016625
location_120.1507568.30	0.0320125	0.206934375	0.05721875	0.02418125	0.307125	0	0.108709375	0	0	0.138165625	0	0
location_120.1507568.30	0	0.230730625	0.211428125	0	0.21213125	0	0.111134375	0	0	0.14106875	0	0
location_120.1507568.30	0	0.35853125	0.11149375	0	0.257628125	0	0.059773125	0	0	0.0765	0	0.069578125
location_120.1507568.30	0.000365625	0.075059375	0.048178125	0	0.462903125	0	0.10635625	0	0	0.196296875	0	0
location_120.1511427.30	0.00785625	3.75e-05	0.0097625	0.029846875	0.585421875	0	0.0127	0	0	0.007421875	0	0
location_120.1511427.30	5.94e-05	0.02434375	0.036734375	0	0.56125625	0	0.17319375	0	0	0.026815625	0	0.000103125
location_120.1511427.30	0.000471875	0.05396875	0.116	0	0.46336875	0	0.166353125	0	0	0.015253125	0	0


4、街景 patch 筛选


可自由设定目标类别的占比，基于该占比数据（如 2-建筑；大厦 9-窗玻璃；窗户 39-栏杆；扶手，占据图像大于 50%）筛选目标街景 patch。


**_04_imgs_filter.py**

**_04_imgs_filter**

**ss_PATCH_result_HZSVI**


**ss_result_HZSVI**

**ss_PATCH_result_HZSVI_filter.csv**

**ss_result_HZSVI_filter.csv**


文件夹下面含有筛选后的街景、街景 patch 文件，以及对应的语义标签 csv 数据。

5、基于 LLaVA: Large Language and Vision Assistant，对街景 patch 进行语义描述，增加对应的词向量，用于优化后面街景 patch 的聚类分析。




what is it?

The image shows a tall building with many windows, likely an apartment building or a hotel. The windows are arranged in a vertical manner, and the building is situated in an area with some greenery, as there are trees nearby.



what is it?


The image shows a large apartment building with many windows, where some of the windows are open.

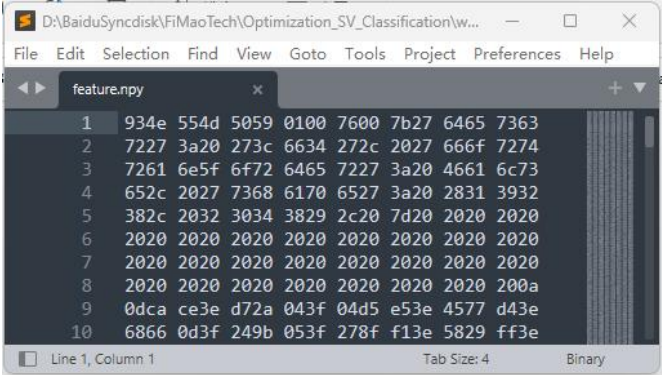


what is it

The image features a large apartment building with many windows and air conditioners attached to the outside. The building is made of concrete and has a gray color tone. The air conditioners are situated on various floors and sections of the building, providing cooling and comfort for the residents. The presence of numerous windows and air conditioners indicates that the building is designed to accommodate multiple residents and offer them individual temperature control within their living spaces.


6、采用 HorNet 框架的预训练模型，提取图片特征向量。


 `_05_imgs_class_results.py`





```
feature.npy
1 934e 554d 5059 0100 7600 7b27 6465 7363
2 7227 3a20 273c 6634 272c 2027 666f 7274
3 7261 6e5f 6f72 6465 7227 3a20 4661 6c73
4 652c 2027 7368 6170 6527 3a20 2831 3932
5 382c 2032 3034 3829 2c20 7d20 2020 2020
6 2020 2020 2020 2020 2020 2020 2020 2020
7 2020 2020 2020 2020 2020 2020 2020 2020
8 2020 2020 2020 2020 2020 2020 2020 200a
9 0dca ce3e d72a 043f 04d5 e53e 4577 d43e
10 6866 0d3f 249b 053f 278f f13e 5829 ff3e
```



7、结合步骤 3 的语义分析数据、步骤 5 的图像描述词向量、步骤 6HorNet 模型提取图片特征向量，进行聚类分析。


_05_imgs_class_results



0



1



2



3



4



5



6



7



8



9



10



11



12



13



14



15



16



17



18


19



20



21



22



23


51



location_120
.1556674,30.
24050064_h
ead90_pitc...



location_120
.1556674,30.
24050064_h
ead90_pitc...



location_120
.1560622,30.
23920848_h
ead180_pit...



location_120
.1560622,30.
23920848_h
ead270_pit...



location_120
.1564884,30.
24055808_h
ead0_pitch...



location_120
.1564884,30.
24055808_h
ead270_pit...



location_120
.1568697,30.
23810132_h
ead0_pitch...



location_120
.1568697,30.
23810132_h
ead270_pit...

























45



location_12
0.1556674,
30.2405006
4_head0_...



location_12
0.1560622,
30.2392084
8_head0_...



location_12
0.1560622,
30.2392084
8_head0_...


location_12
0.1560622,
30.2392084
8_head27...


location_12
0.1560652,
30.2395690
3_head0_...


location_12
0.1560652,
30.2395690
3_head0_...


location_12
0.1560652,
30.2395690
3_head18...


location_12
0.1560652,
30.2395690
3_head27...

