街景patch分类研究过程梳理

1、对图像进行缩放及裁剪

|  |  |
| --- | --- |
|  | |
|  |  |
| 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 |
|  |  |

1. 对街景分patch

可自由设置把一张街景分割多少行、多少列。

|  |
| --- |
|  |
|  |
|  |

1. 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。

|  |  |
| --- | --- |
| **语义标签色标，见ade20k\_color文件。** | |
|  | |
|  |  |
|  |  |
|  |  |
| 该语义分析共包含150中分类， | |

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。

|  |  |
| --- | --- |
|  | |
|  | |
|  | 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。

|  |
| --- |
|  |
|  |
|  |

6、采用HorNet框架的预训练模型，提取图片特征向量。

|  |  |
| --- | --- |
|  |  |
|  |  |

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。

|  |
| --- |
|  |
|  |
| 51 |
| 45 |