**关于初赛晋级审核的通知（重要）**

各位选手，大赛初赛即将结束，进入资格审核阶段，请务必于10月19日晚10点（北京时间）前填写完成以下问卷表格，以便审核，若未及时填写，可能会被取消晋级资格：

【填写链接】https://jinshuju.net/f/ktwNEC

【以下信息请务必仔细阅读】  
1、请初赛线上最终成绩排名top110的团队填写（分B榜的赛道以B榜为准，不分的赛道以A榜为准），经审核后如无特殊情况，线上最终成绩top100团队晋级复赛；  
2、各团队最晚需于10月19日晚上10点（北京时间）前完成提交，建议尽量提前提交，避免晚提交使得官方无法及时和你的团队进行质询沟通，造成误判；  
3、组委会将根据提交的材料进行晋级审核及详细抽查，排名top 20的团队将至少有10支队伍被详细抽查审核，其他团队将随机详细抽查审核；  
4、若被抽查团队（将不公布被抽查名单）没有填写本问卷，视为自动放弃比赛，将取消成绩，若审核不合格亦将取消成绩，晋级资格由排名后续团队补足；  
5、在提供的gitl链接项目中，选手需要保留所有的 commit，以便展现项目迭代过程；如有版本迭代，保留所有 tag 与 release；应当包含方案说明文件，所有的模型和工程文件，保证模型和成绩可复现；最终代码以 master 分支为准，可适当合并分支（本阶段为主观评审）；  
6、经审核后，晋级名单将不晚于10月24日发布并公示，届时可在赛题数据下载页查看晋级名单，在发布名单的7天内，若有异议或审核均可向官方邮箱contact@datafountain.cn详细说明。

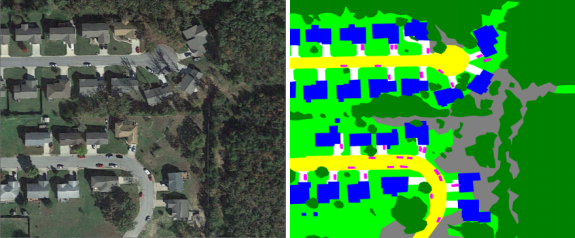
请大家务必保证填写信息真实有效，审核期间可能会有工作人员与你沟通，请积极配合，谢谢！

**赛道背景**

地物要素分类是将地表相对固定的物体分类的系统体系，是地表地物要素观测与测绘的重要手段之一。基于遥感影像开展地理国情监测、“耕地红线”、“生态红线”等实际应用，受“同物异谱”、“同谱异物”现象影响，分析与处理的难度极大，目前主要采用人工方式提取地物要素，效率低，耗资巨大，迫切需要自动化、高精度的地物要素提取方法。AI+遥感影像赛道，旨在充分利用大数据、人工智能等先进技术，突破面向大规模多源异构高分数据信息提取与分析的技术瓶颈，形成一批高效、可用、实用的先进算法，提升高分对地观测信息服务能力。

**赛道任务**

对高分辨率光学遥感图像中各类地物光谱信息和空间信息进行分析，将图像中具有语义信息的各个像元分别赋予语义类别标签；以包含典型土地利用分类的光学遥感图像为处理对象，选手使用主办方提供的遥感图像进行土地利用类型语义分割处理，主办方依据评分标准对结果进行综合评价。



依据现有的遥感地物分类要求，结合现有的地物分类实际需求，参照地理国情监测、“三调”等既有地物分类标准，依据遥感地物“所见即所得”原则，设计地物要素分类体系，共涉及一级大类8种，二级子类17种，详见表：

| **一级大类** | **二级分类** | **分类示例** | **参考说明** |
| --- | --- | --- | --- |
| 一级大类 | 二级分类 | 分类示例 | 参考说明 |
| 水体 | 水体 | 河流、湖泊、水库、坑塘、沟渠、海洋等 | 视觉效果为纯水面，如蓝色、浅蓝色等，由于水体污染等情况，水体可能出现浅绿到深绿色，并且由于水深的不同，水体可能出现颜色变幻。甚至由于传感器角度与太阳角度，出现镜面反射的情况，水体也可能出现高亮的情况，另外在大风天，高分辨率影像的水体也可能出现鱼鳞纹。 |
| 交通运输 | 道路 | 快速路、主干路、次干路和支路等 | 深灰色或灰白色，带状，连续，有明显机动车道或高架匝道，内有行驶车辆。从影像角度而言，只考虑明显可以将地物分割为不同地块的道路，例如小区内部细小道路不予考虑，只标注周边的公路 |
| 交通运输 | 机场 | 飞行区、航站楼等 | 位于城市中心区以外，与高速公路联系，占地面积大，有明显白色飞机跑道，大体量的候机大厅，并配有大规模仓储库房。勾画时，仅勾画像元级别精度的跑道与配属的航站楼，以及其他的拥有飞机的配属建筑。 |
| 交通运输 | 火车站 | 火车站站台及候车厅 | 火车站因其背景特殊（建于铁道之上）且形制特殊，可以单独作为一类地物进行提取。 |
| 建筑 | 建筑物 | 住宅、工厂等人工建筑 | 村庄、城市各类建筑，如住宅、厂房，排列较为整齐，住宅通常为红顶黑顶，厂房多为蓝顶、白顶或灰顶等。 |
| 耕地 | 普通耕地 | 普通耕地 | 耕地多为规则矩形组成，成片分布，纹理细致。此处的耕地仅为在该地块上无任何附加建筑与覆盖的耕地，例如农业大棚另做一类。耕地在不同时相上、不同气候背景下有不同的光谱特征，例如休耕期与作物生长期的光谱特征截然不同，在南北方的相同时期，因气候背景不同，其光谱特征也会由于作物类型不同而不同。 |
| 耕地 | 农业大棚 | 耕地另一种形式，如温室 | 农业大棚为具有一定规模、排列紧密整齐，条状浅蓝绿色或白色窄形。 |
| 耕地 | 自然草地 | 自然草地 | 指以生长草本植物为主，覆盖度在5%以上的各类草地，包括以牧为主的灌丛草地和郁闭度在10%以下的疏林草地。自然草地颜色以浅绿色至深绿色为主，间杂矮小灌木，图像纹理杂乱。 |
|  | 绿地绿化 | 包括公园绿地、生态景观绿地、防护绿地等人工草地 | 城市公园，内部有大面积洁净水面或观赏塔、亭，绿地率不小于65%；大型郊野公园影像特征与城市公园相似，占地面积更大，多位于城市边缘地区或绿隔地区；防护绿地一般位于高速路、高压线、铁路两侧，或燃气、污水处理厂等危险源四周，带状分布、均匀宽度，起隔离、卫生和安全的作用。 |
| 耕地 | 自然林 | 乔木林、灌木林 | 多位于丘陵地区，成片分布，深绿色，具有明显的沟脊走向，山坡阳坡颜色以浅绿色至深绿色为主，背阴坡以深绿色至黑色阴影覆盖为主。从影像判读中，其植被郁闭度极高，有明显的垂直分布结构。 |
|  | 人工林 | 果园林、苗圃等 | 边界规则，行株距清晰，树冠小，深色点状纹理，内部道路系统清晰连贯。树冠阴影表现为黑色的晕斑等明显有人为修正痕迹的林地。 |
| 裸土 | 自然裸土 | 自然裸土、戈壁、沙漠 | 主要以自然原因造成的土壤裸露为主，例如因土壤贫瘠、土层较薄造成的土壤裸露，颜色以褐色或者灰色为主，或者因气候原因造成土壤沙化进而形成的戈壁滩、沙漠等，颜色以黄色为主。 |
| 裸土 | 人为裸土 | 指因建筑需求等原因破坏地表覆盖而造成裸土裸露的地表 | 因建设需要，在建设之前，需要将地表开挖，然后运走土方，以便填充构筑物，例如打桩作为地基稳定建筑物，所以该地物存在时间较短，且纹理复杂，颜色以亮黄色为主，但是有时候因环境需求，须铺设防尘网等防止扬尘的设施，所以部分情况下，在遥感影像上呈浅绿色。 |
| 其它 | 光伏 | 太阳能光伏发电板 | 太阳能光伏板是光谱和纹理相对于背景很特殊的一种地物，其颜色以深蓝色为主，纹理固定 |
| 其它 | 停车场 | 社会停车场用地/公共交通场站/ | 位于主干道、次干道两侧或车站前甚至房顶，或大型购物中心附近。有数量较多、排列整齐或者间杂排列的汽车。 |
| 其它 | 操场 | 篮球场、排球场、羽毛球场、网球场等 | 椭圆形红色或者蓝色跑道，内有绿色或者蓝色等其他醒目色的绒状足球场，亦或者有明显线条的室外篮球场、羽毛球场灯体育设施 |
| 其它 | 其它无法确定归属地物 | / | / |

依据所述地物要素分类标准与星载遥感影像处理实际需求，以多种地物要素的高分光学遥感图像的像素级分类为目标，设计赛题如下：  
a. 初赛：算法在地物要素按一级大类（8类）分类的能力。进阶考察算法地物要素分类的准确性；  
b. 复赛：算法在地物要素按二级子类（17类）分类的能力。进阶考察算法分类的准确性，测试数据尺寸变化的适应性，支持统一的接口调用；  
c. 决赛：算法在地物要素按二级子类（17类）分类的能力。进阶考察现场答辩评议，强调模型复杂度和效率考察，算法分类的准确性，测试数据尺寸变化的适应性，支持统一的接口调用；

**数据简介**

（1）来源：数据为0.1米-4米分辨率的高分一、二、六号，高景二号，北京二号，以及部分航空等数据源的可见光、多光谱载荷图像，由鹏城实验室和协办单位合作采集、标注、构建；  
（2）规模：100万+张遥感影像语义分割样本数据；  
（3）用途：土地利用动态监测，矿产资源开发状况和地质灾害的调查与监测；生态环境监管调查与评价，水环境监测与评估，空气环境监测与评价；耕地数量与质量调查。  
（4）方法  
初赛：10万高分光学影像和标注文件（一级分类），20万测试图片数据；  
复赛：10万高分光学影像和标注文件（二级分类），30万测试图片数据；  
决赛：10万高分光学影像和标注文件（二级分类），30万测试图片数据，强调模型复杂度和效率考察：

**数据说明**

（1）整体数据说明  
a.原始影像  
影像格式为tif，包含R、G、B三个波段，训练集影像尺寸为256 \* 256像素，初赛测试集影像尺寸为256 \* 256像素，复赛、决赛初步确定测试集尺寸可变。  
b.标签数据  
标签格式为单通道的png，每个像素的标签值由一个三位数表示，使用‘uint16’数据类型存储，该三位数包含了一级和二级两个类别信息，百位上的一个数字表示一级类别，十位和个位上的两个数字一起表示二级类别。  
一级类别共分为8类，一级类别及对应的百位上的数字如下表所示：

| **一级类别** | **对应标签百位上的数字** |
| --- | --- |
| 水体 | 1 |
| 交通运输 | 2 |
| 建筑 | 3 |
| 耕地 | 4 |
| 草地 | 5 |
| 林地 | 6 |
| 裸土 | 7 |
| 其它 | 8 |

二级类别共分为17类，二级类别及对应的十位、个位上的数字如下表所示：

| **二级类别** | **对应标签十位及个位上的数字** |
| --- | --- |
| 水体 | 01 |
| 道路 | 02 |
| 建筑物 | 03 |
| 机场 | 04 |
| 火车站 | 05 |
| 光伏 | 06 |
| 停车场 | 07 |
| 操场 | 08 |
| 普通耕地 | 09 |
| 农业大棚 | 10 |
| 自然草地 | 11 |
| 绿地绿化 | 12 |
| 自然林 | 13 |
| 人工林 | 14 |
| 自然裸土 | 15 |
| 人为裸土 | 16 |
| 其它 | 17 |

（2）初赛标签数据  
初赛中，要求对一级类别分类，分类的结果为表3中的8个类别，提供的标签数据中，三位数标签的十位及个位数赋值为0，用百位上的一个数字来表示一级类别，例如：对于标签值为100的像元，对照表3，表示该像元的一级类别为水体。  
（4）复赛/决赛标签数据  
复赛和决赛中，要求对二级类别分类，分类的结果为表4中的17个类别，三位数标签的十位和个位一起表示二级类别，例如，对于标签值为204的像元，该像元的二级类别编码为‘04’，对照表4，表示该像元的二级类别为机场。

**提交要求**

（1）初赛  
预测结果压缩包提交，参赛者将测试集的预测结果放入results文件夹并压缩为zip格式，将压缩包上传到大数据竞赛平台，平台进行在线评分，实时排名。  
（2）复赛及决赛  
复赛及决赛作品提交要求将于初赛结束前2周公布。

**提交示例**

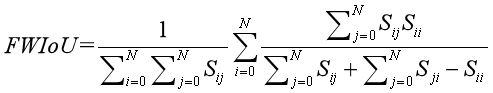
（1）初赛  
a.向大赛平台提交参与评测的results.zip内容为：  
results:  
1.png  
2.png

注意事项: 预测结果中的单个文件名需和预测图片命名方式一致，如预测图片中包含1.jpg，则预测结果中必须有1.png，其中1.png的数据格式和提供的训练集的标注文件保持一致。  
b.主观评测的提交材料：

* 应当包含必要的代码级样例展示；
* 应当包含详细的解题思路说明、项目运行环境和运行办法等信息，方便大赛专家评委进行成绩有效性核实；  
  c.参赛项目模型和工程文件：
* 建议保留从工程创建至截止日期所有的 commit，以便展现项目迭代过程；
* 如有版本迭代，建议保留所有 tag 与 release；
* 应当包含所有的模型和工程文件，保证模型和成绩可复现；
* 最终代码以 master 分支为准，请适当合并分支；
* 成绩需要可以成功复现，选手有义务及时响应技术委员会的复现要求；
* 其他技术委员会视情况指定的审核要求；

（2）复赛  
复赛提交示例细则将在初赛结束后提供。

7、评测标准  
（1）初赛/复赛  
初赛、复赛使用国际通用指标加权交并比FWIoU， 根据每个类出现的频率为其设置权重，具体计算公式为：



（2）决赛  
决赛在FWIoU的技术上引入模型效率及模型复杂度评价指标：  
模型效率：模型效率以6小时为基础归一化为百分制，超过6小时得分为0。  
模型复杂度：模型复杂度以模型大小50MB为基础归一化为百分制，超过50MB得分为0。  
决赛客观得分 = 0.5*FWIoU*100 + 0.3*E + 0.2*C。

注意：

1、评审说明

* 初赛排行榜采用 A/B 榜机制，其中A 榜计算提交结果文件中一定比例数据的成绩，B 榜计算剩余数据的成绩，A、B榜阶段前将发布相应测试集；
* 在初赛A榜阶段，每个队伍每天最多可提交3次结果文件参与评测，平台实时评测出分，在A榜阶段内各团队最高分参与排名，排行榜实时更新；
* 在初赛B榜阶段，每个队伍每天可多次提交结果文件，但平台仅对当日最后1次提交进行评测，于当晚12:00:00评测出分，在B榜阶段内各团队最高分参与排名，排行榜每晚12:00:00定时更新；
* 主观评审规则将由组委会统一发布；
* 决赛加分项：基于昇腾的算法模型，决赛成绩可加分（细则待复赛结束后公布）。

2、本次比赛允许选手自主使用外部数据集和预训练模型辅助训练比赛模型，但应满足如下要求：

* 选用的预训练模型必须已经在学术会议或期刊上公开发表（不含 arxiv 发表）；是论文作者公布的原始模型、各大深度学习平台（如 Pytorch，Tensorflow 等）官方提供的预训练模型或者是在2020年10月23日之前发布的预训练模型；
* 不允许选手对大赛数据集进行额外标注并用于训练，一经发现将取消参赛资格；
* 比赛阶段，选手应配合工作人员进行随机核查。若使用了外部数据和预训练模型，选手应将用到的外部数据集和预训练模型一并提供给工作人员，并提供说明文档、详实阐述训练过程；
* 复赛有望晋级队伍将需要进行结果复现。届时选手需向工作人员提供其用到的外部数据集和预训练模型（没有用到则不需提供）和说明文档，并在组委会提供的平台上从零开始复现；
* 参赛人员需保证使用的数据集和模型不侵犯第三方权益，任何侵权行为与主办方及其相关方、平台方无涉。