



数据智能中台-高质量数据集

用户操作手册

天翼云科技有限公司

目录

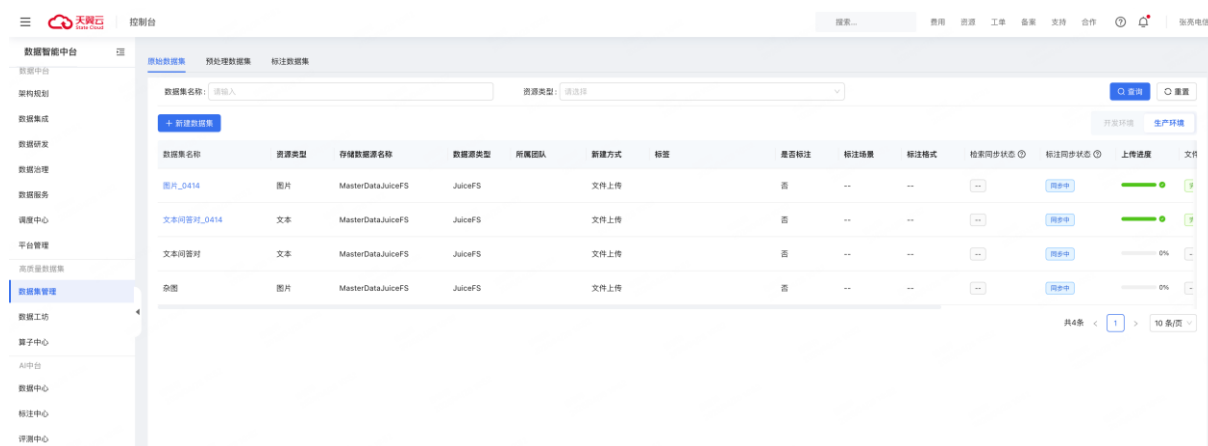
1.	高质量数据集平台功能介绍	3
1.1.	数据集管理.....	3
1.2.	数据加工	4
1.3.	数据评测	4
1.4.	算子广场	6
1.5.	算子管理列表.....	7
2.	高质量数据集平台操作说明	7
2.1.	创建数据源（以 Ceph 数据源为例）	7
1.	创建 Ceph 数据源	7
2.2.	数据集管理.....	10
2.2.1.	新建原始数据集&数据集管理	10
2.2.2.	预处理数据集管理	14
2.2.3.	标注数据集管理	18
2.3.	创建数据加工任务.....	19
2.4.	数据评测	28
2.4.1.	规则定义	28
1.	创建评测任务	31
2.	查看评测报告	33
2.5.	新增算子	34
2.5.1.	新增引用内置算法的算子	34

2.5.2. 新增引用内置算法模型的算子	37
2.5.3. 新增引用第三方 Http 算法的分析类算子	42
2.5.4. 新增引用第三方 Http 算法的生成类算子	46
2.5.5. 算子前端组件配置	50

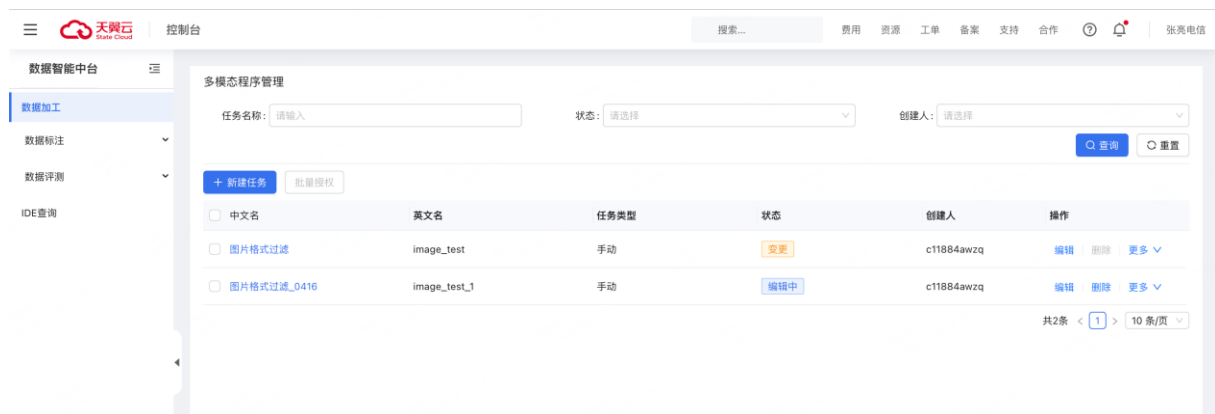
1. 高质量数据集平台功能介绍

1.1. 数据集管理

数据集管理用于对生产出来的数据集进行统一的管理。



1.2. 数据加工

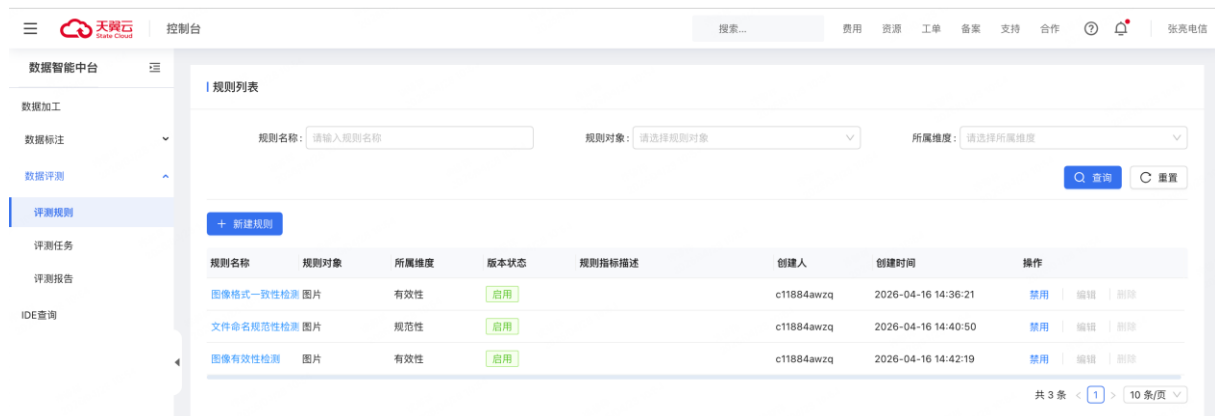


数据加工提供对多模态数据进行采集、处理的任务的管理功能，包括任务新建、任务编辑、任务删除、任务授权及批量授权等操作。

任务新增具体操作见 1.3 高质量数据集平台操作案例的 1.3.2 创建任务部分。

1.3. 数据评测

数据评测基于数据的元信息及信通院对于数据集质量的评测方法理论，进行规则定义与创建，在评测任务中通过规则的组合，设置不同的规则权重来从多个维度进行数据集数据质量的评价，最终输出质量评测报告。



任务评测管理

任务名称:

评测对象:

任务状态:

创建人:

[+ 新建评测任务](#)

任务中文名	任务英文名	任务状态	评测对象	创建人	创建时间	操作
电信问答对评测	teleqa	成功	电信问答对_文本质量统...		2026-04-09 10:20:41	执行 编辑
问答对抽取数据集评测	QA_report	成功	问答对抽取加工_智能文...		2026-04-02 10:36:33	执行 编辑
车辆破损demo	damaged_vehicle_demo	成功	车辆破损demo_图片增...		2026-03-26 22:09:09	执行 编辑
交通路口车辆灰度数据集质量评测	Vehicle_Gayscale_Dataset_a...	成功	交通路口车辆灰度数据集...		2025-12-24 15:44:44	执行 编辑
电动车入梯数据集质量评测	Evaluation_of_Electric_Vehicle_	成功	交通路口车辆灰度数据集...		2026-01-23 16:44:05	执行 编辑
文件大小_demo_cq	document_demo_cq	成功	文档解析演示demo_智...		2026-01-20 11:07:06	执行 编辑
检测测试	testdoc114	成功	交通路口数据集视频处理...		2026-01-14 15:32:06	执行 编辑
文档评测演示demo	docstest	成功	文档解析演示demo_智...		2025-12-24 16:39:48	执行 编辑

报告查询

报告名称:

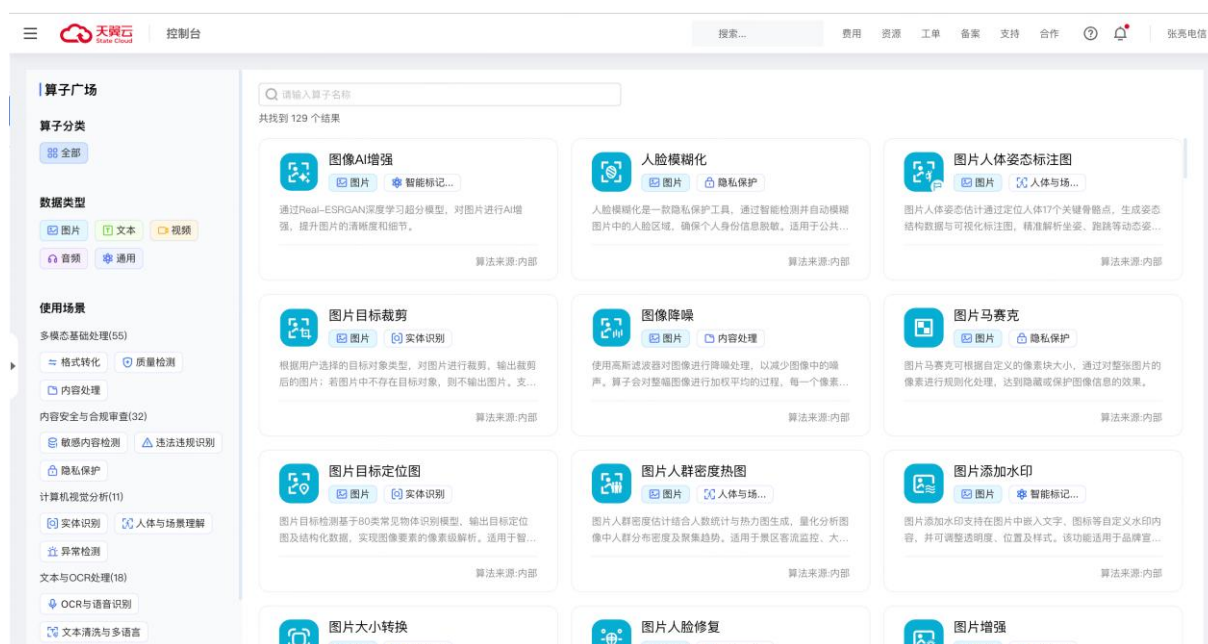
所属任务:

评测对象:

创建人:

报告名称	所属任务	版本号	来源	评测对象	创建人	最近执行时间	操作
电信问答对评测_报告	电信问答对评测	v2	原始评测	电信问答对_文本质量统计_20260409		2026-04-09 10:21:23	查看报告 人工复核
电信问答对评测_报告	电信问答对评测	v1	原始评测	电信问答对_文本质量统计_20260409		2026-04-09 10:20:46	查看报告 人工复核
问答对抽取数据集评测_报告	问答对抽取数据集评测	v3	原始评测	问答对抽取加工_智能文档解析_20260402		2026-04-02 11:53:05	查看报告 人工复核
问答对抽取数据集评测_报告	问答对抽取数据集评测	v2	原始评测	问答对抽取加工_智能文档解析_20260402		2026-04-02 11:51:11	查看报告 人工复核
问答对抽取数据集评测_报告	问答对抽取数据集评测	v1	原始评测	问答对抽取加工_智能文档解析_20260402		2026-04-02 10:39:27	查看报告 人工复核
车辆破损demo_报告	车辆破损demo	v4	原始评测	车辆破损demo_图片增强_20260327		2026-03-27 08:11:51	查看报告 人工复核
车辆破损demo_报告	车辆破损demo	v3	原始评测	车辆破损demo_图片增强_20260327		2026-03-27 08:11:15	查看报告 人工复核
车辆破损demo_报告	车辆破损demo	v2	原始评测	车辆破损demo_图片增强_20260327		2026-03-27 08:07:14	查看报告 人工复核
车辆破损demo_报告	车辆破损demo	v1	原始评测	车辆破损demo_图片增强_20260326		2026-03-26 22:09:12	查看报告 人工复核
交通路口车辆灰度数据集质量评测_报告	交通路口车辆灰度数据集质量评测	v4	原始评测	交通路口车辆灰度数据集生产流程_图像灰度...		2026-03-18 20:30:35	查看报告 人工复核

1.4. 算子广场



算子广场提供高质量数据集平台所有算子（包括内置算子及用户自定义算子）的全局视图，用户可以通过算子广场查看算子的功能介绍、所用算法支持的输入格式、所需配置项及算子实现的功能实例等信息，对算子有初步的了解，有助于后续使用对应的算子进行数据的处理。

算子广场左侧为算子分类及算子标签，可对算子根据标签进行更细致的筛选。

点击算子卡片后会进入算子详情页面，如下所示：



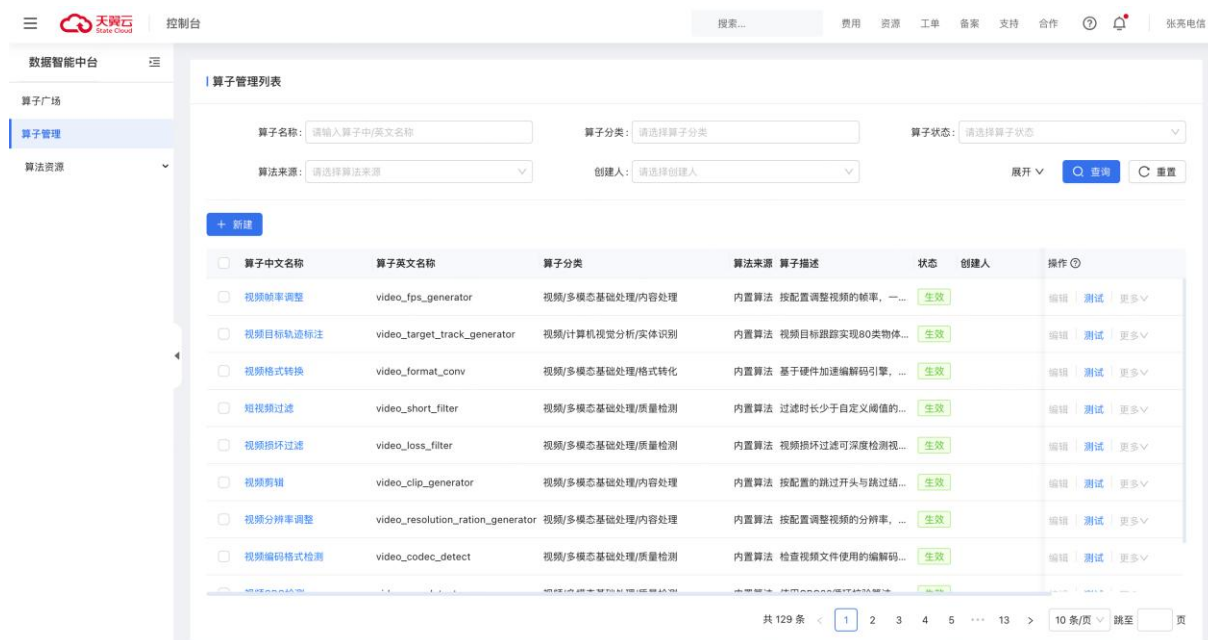
算子详情页面可查看算子的详细功能介绍以及算法信息等；

算法信息包括支持处理的文件格式：image（图片）、video（视频）、audio（音频）、text（文

本), 以及每个算法需要用户配置的参数信息等。

同时, 在算子详情页还会以功能示例的形式, 以处理前的数据形态和处理后的数据形态对比展示, 让用户实际感受到算子的具体功能, 对算子有进一步的了解。

1.5. 算子管理列表



算子管理列表提供对整个高质量数据集平台所有数据处理算子的管理能力, 包括算子新建、算子编辑、算子删除、算子上下线等操作;

算子新增具体操作见 1.3 高质量数据集平台操作案例的 1.3.1 算子新增部分。

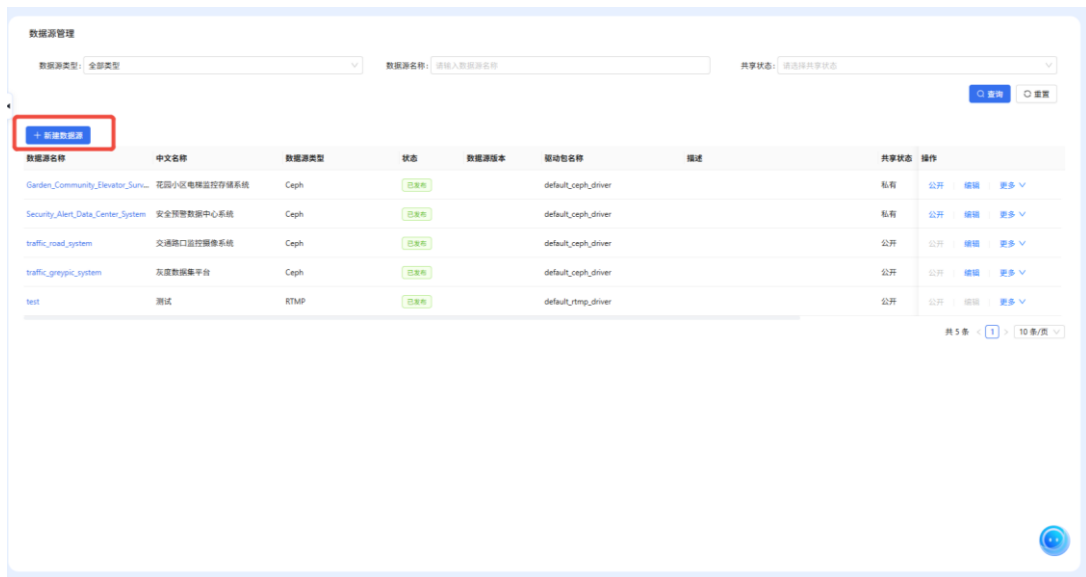
2. 高质量数据集平台操作说明

2.1. 创建数据源 (以 Ceph 数据源为例)

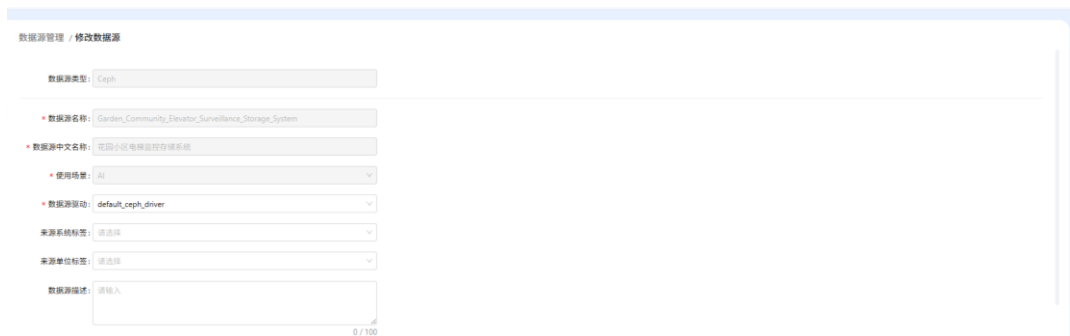
1. 创建 Ceph 数据源

第一步: 新建数据源

在数据源管理-数据源接入中，点击新建数据源按钮。

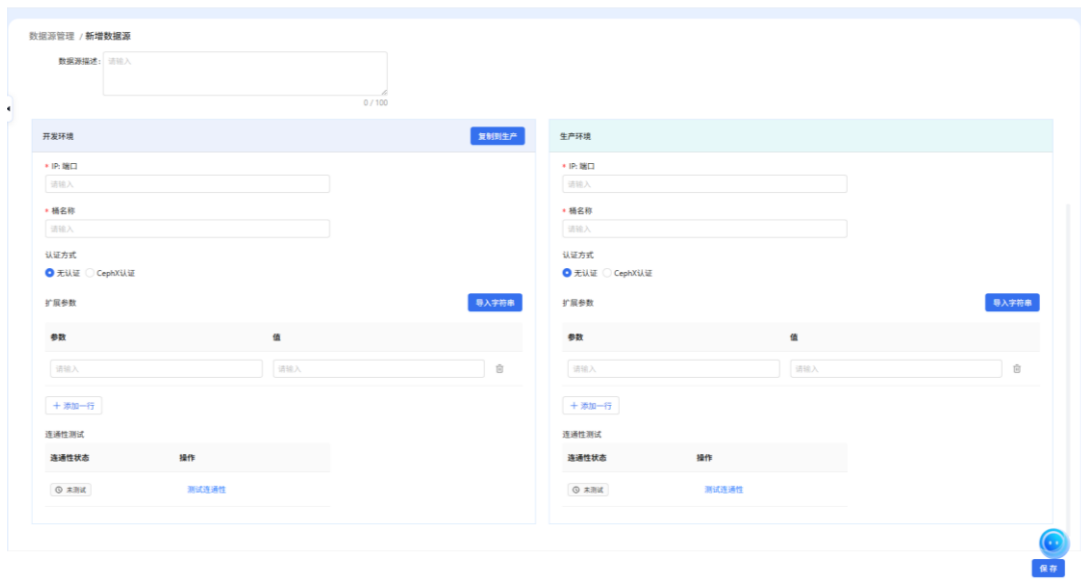


在弹窗中，选择对象存储分类，选择 Ceph，进入 Ceph 连接信息配置页面；
配置页面分为上下两个部分，上半部分主要维护数据源的基础信息；



- 数据源类型：在弹窗中已选择，不可更改；
- 数据源名称：数据源的英文名称，作为数据源的唯一标识，支持大小写英文字母、数字、下划线；
- 数据源中文名称：数据源的中文名称，无唯一性要求；
- 使用场景：
- 数据源驱动：用于连接数据库的驱动，系统已内置常用驱动，如需特殊驱动文件，则可在驱动管理中进行上传，一般情况下使用内置驱动即可；
- 来源系统标签：用于标识该数据源的来源系统，标签值在标签管理中进行维护，非必填；
- 来源单位标签：用于标识该数据源的来源单位，标签值在标签管理中进行维护，非必填；
- 数据源描述：数据源的描述信息；

下半部分是数据源的连接信息，包括 IP、认证、连通测试等；



关于开发环境和生产环境：基于生产隔离的理念设计的功能，将测试与生产的数据隔离，防止测试行为影响生产业务；如本身不存在此类顾虑，可使用复制到生产，使用同一个 ip 和桶。

IP: 端口: Ceph 数据源所在的服务器的 IP 地址及其端口号，以英文冒号分隔；

桶名称: Ceph 存储中的分桶的名称，用于定位桶；

认证方式: 分为无认证和 CephX 认证两种；

Access Key: 用于 Ceph 身份认证的 Access Key 信息；

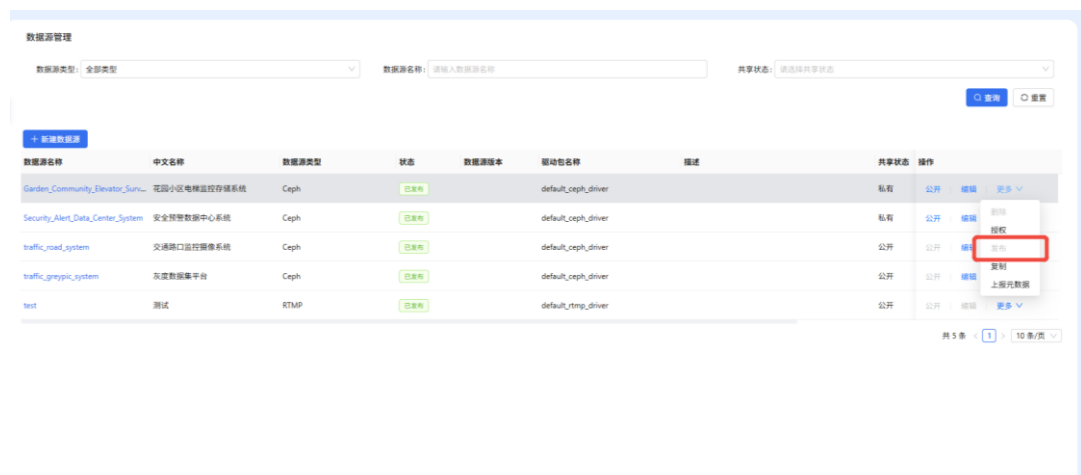
Secret Key: 用于 Ceph 身份认证的 Secret Key 信息；

扩展参数: 如连接 Ceph 需要其他额外参数，可以 k-v 键值对的形式维护在扩展参数中；

连通性测试: 当所有信息填充完成后，点击连通性测试，可测试信息填写是否正确，系统是否能够成功访问到数据库中。

第二步：发布数据源

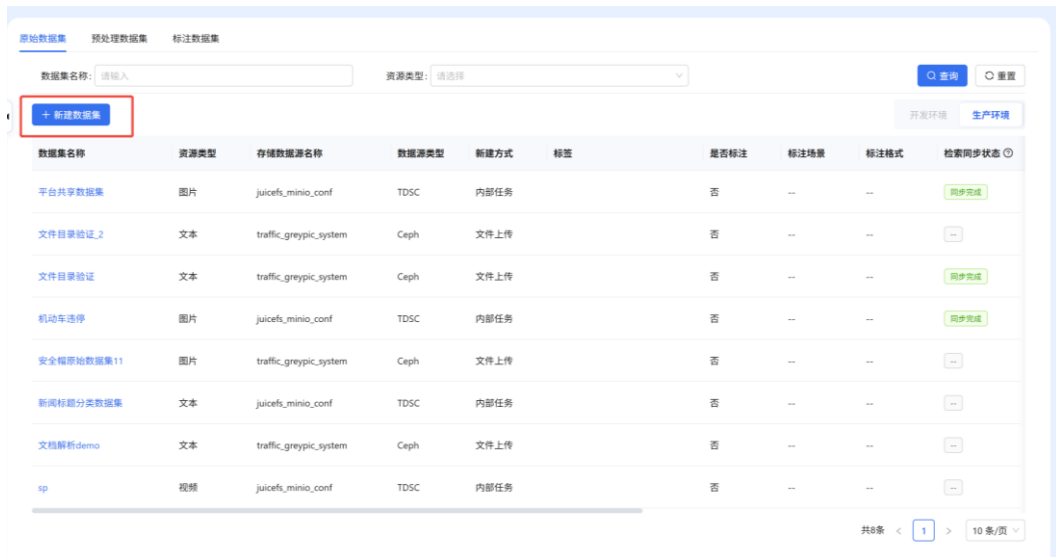
回到列表后，点击对应数据源的更多按钮，点击发布，可将数据源发布，此时数据源可用。



2.2. 数据集管理

2.2.1. 新建原始数据集&数据集管理

原始数据集管理页面支持用户进行已有数据集的纳管及上传；
点击【新建数据集】，可进行原始数据集的纳管和上传。

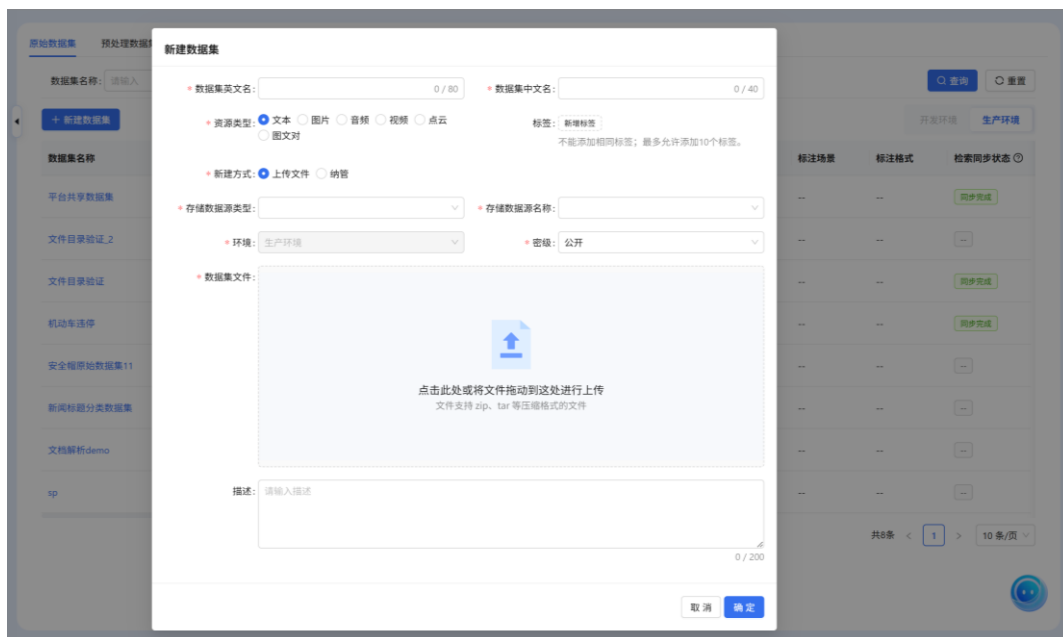


在弹窗中，可以选中新建原始数据集的方式，分为上传文件和纳管两种；

上传文件：支持用户将本地的数据集文件通过拖拽方式上传到平台；

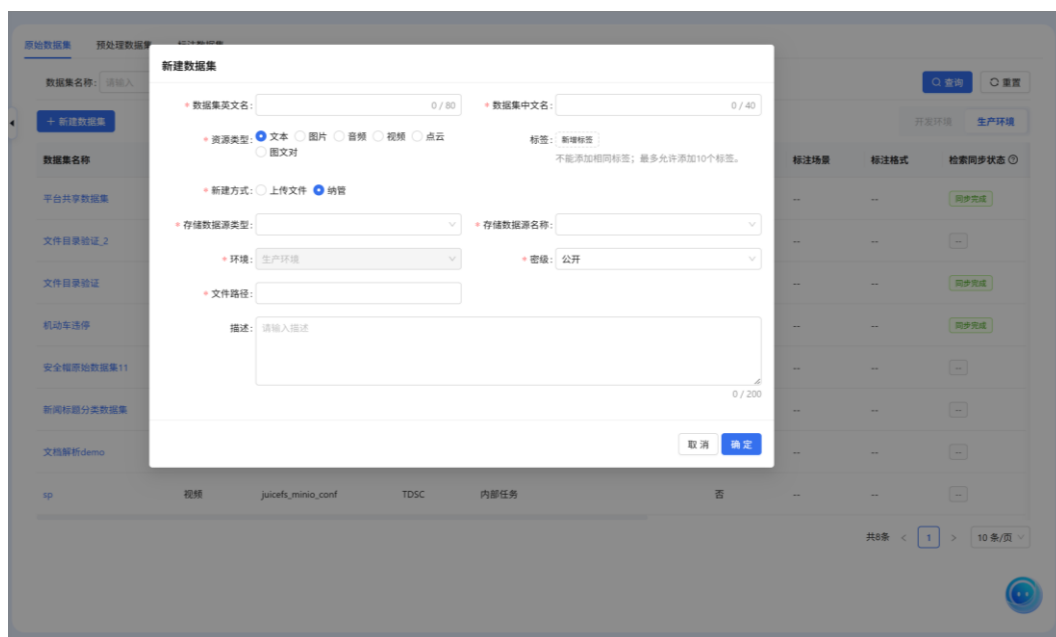
纳管：支持用户选择某个数据源的某个路径，将该路径下的文件形成一个原始数据集；

当选择上传文件方式，需维护数据集的中英文名称、选择数据集的资源类型，选择数据集标签等；并选择需要将该文件上传至哪个类型的哪个数据源，上传完成后，系统会自动生成存放路径。



当选择纳管方式，需维护数据集的中英文名称、选择数据集的资源类型、选择数据集标签等，

同时，选择需纳管的数据集所在的数据源及其路径，此时系统会将所填路径下的文件打包为一个原始数据集。在数据集管理中进行管理。



当数据集上传完成后，可在数据集管理列表查看到该原始数据集：



列表有五个状态，其代表的含义如下：

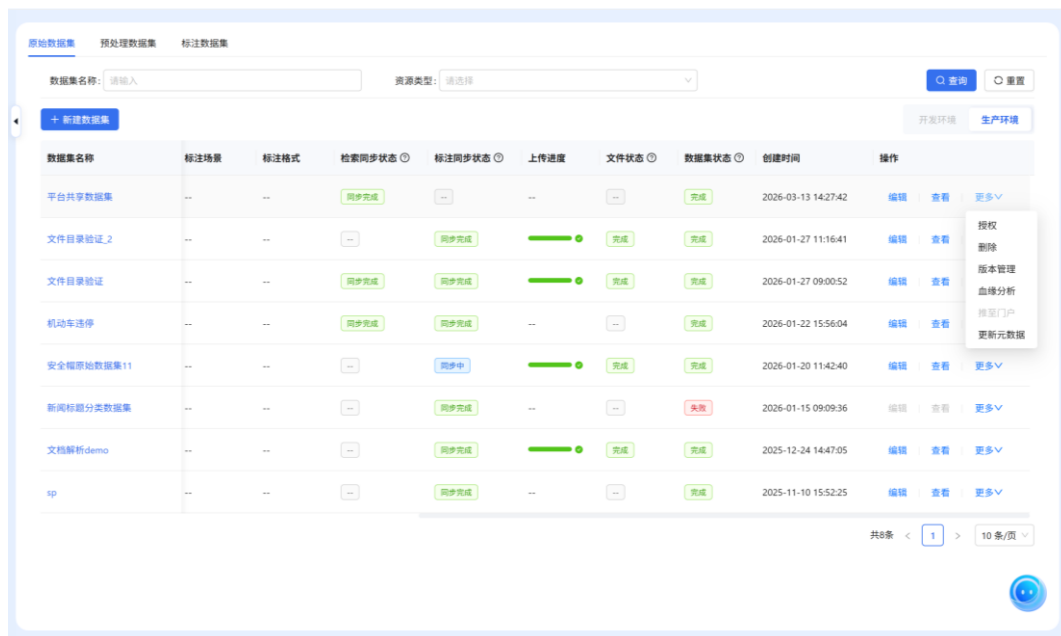
【1】检索同步状态：表示该数据集的信息是否已同步给数据检索，同步完成后，该数据集即可在数据检索页面进行检索；

【2】标注同步状态：表示该数据集的信息是否已同步至标注平台，同步完成后，该数据集即可在创建标注任务时进行选择；

【3】上传进度：选择上传文件方式时该状态有意义，表示上传的进度；纳管方式下无意义，以 [-] 标注；

【4】文件状态：选择上传文件方式时该状态有意义，表示文件解析的结果；纳管方式下无意义，以 [-] 标注；

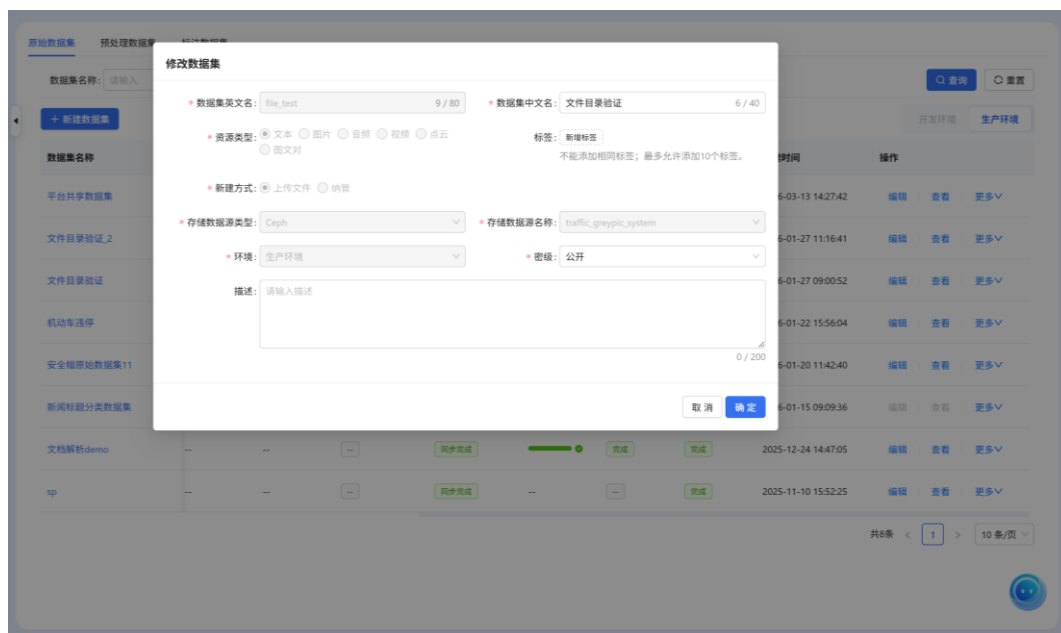
【5】数据集状态：表示该数据集是否生成完成；



数据集名称	标注场景	标注格式	检索同步状态	标注同步状态	上传进度	文件状态	数据集状态	创建时间	操作
平台共享数据集	--	--	同步完成	--	--	--	完成	2026-03-13 14:27:42	编辑 查看 更多
文件目录验证_2	--	--	-	同步完成	<div style="width: 100%;"></div>	完成	完成	2026-01-27 11:16:41	编辑 查看 更多
文件目录验证	--	--	同步完成	同步完成	<div style="width: 100%;"></div>	完成	完成	2026-01-27 09:00:52	编辑 查看 更多
机动车违停	--	--	同步完成	同步完成	--	--	完成	2026-01-22 15:56:04	编辑 查看 更多
安全帽原始数据集11	--	--	-	同步中	<div style="width: 100%;"></div>	完成	完成	2026-01-20 11:42:40	编辑 查看 更多
新闻标题分类数据集	--	--	-	同步完成	--	--	失败	2026-01-15 09:09:36	编辑 查看 更多
文档解析demo	--	--	-	同步完成	<div style="width: 100%;"></div>	完成	完成	2025-12-24 14:47:05	编辑 查看 更多
sp	--	--	-	同步完成	--	--	完成	2025-11-10 15:52:25	编辑 查看 更多

针对原始数据集的操作，分为以下几类：

【编辑】，仅能够编辑数据集的中文名称、数据集的标签、密级信息；



修改数据集

数据集英文名: file_test 9 / 80

数据集中文名: 文件目录验证 6 / 40

资源类型: 文本 图片 音频 视频 点云 文档对

新建方式: 上传文件 纳管

存储数据类型: Ceph

存储数据源名称: traffic_greypic_system

环境: 生产环境

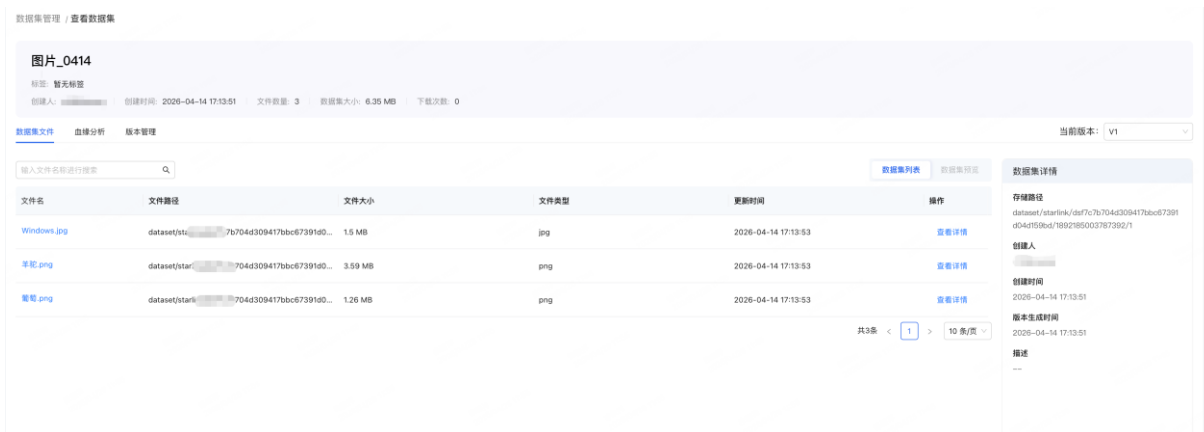
密级: 公开

标签: 新建标签
不能添加相同标签; 最多允许添加10个标签。

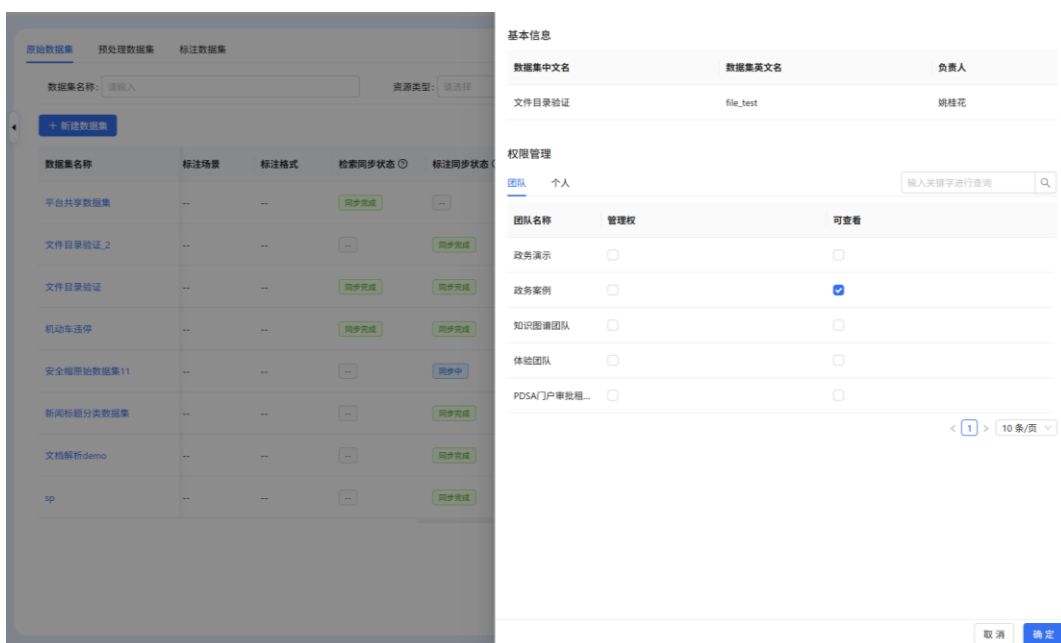
描述: 请输入描述 0 / 200

取消 确定

【查看】指查看数据集的内容；



【授权】支持将该数据集授权给个人或团队；



【删除】指能够删除该数据集；

【版本管理】指能够查看数据集的版本情况，可进行版本管理【注意：原始数据集仅有一个版本】



【血缘分析】支持查看该数据集的血缘情况；

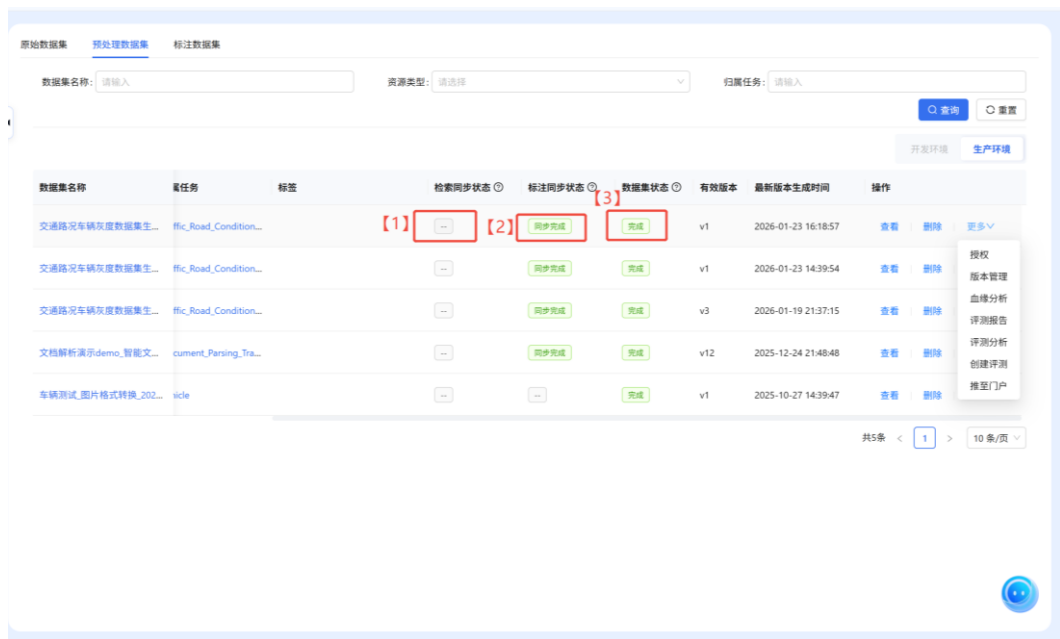


【推至门户】当具备数据集门户时操作有效，可以将该数据集推送到门户的管理端。

【更新元数据】点击后会将该数据集元数据信息进行更新。

2.2.2. 预处理数据集管理

预处理数据集是由数据加工任务生成的数据集：

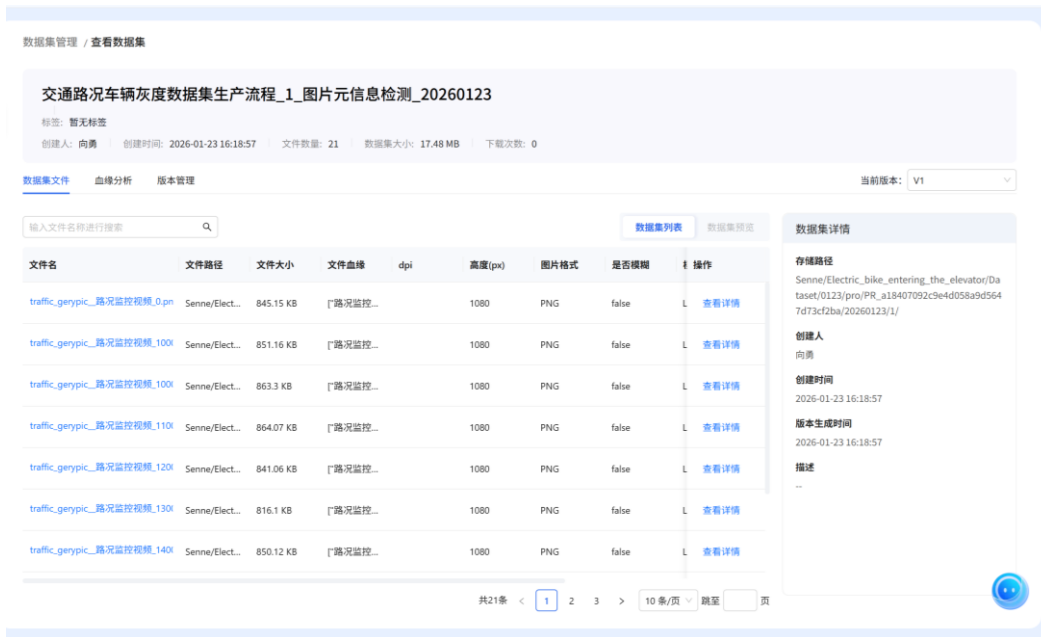


【1】检索同步状态：表示该数据集的信息是否已同步给数据检索，同步完成后，该数据集即可在数据检索页面进行检索；

【2】标注同步状态：表示该数据集的信息是否已同步至标注平台，同步完成后，该数据集即可在创建标注任务时进行选择；

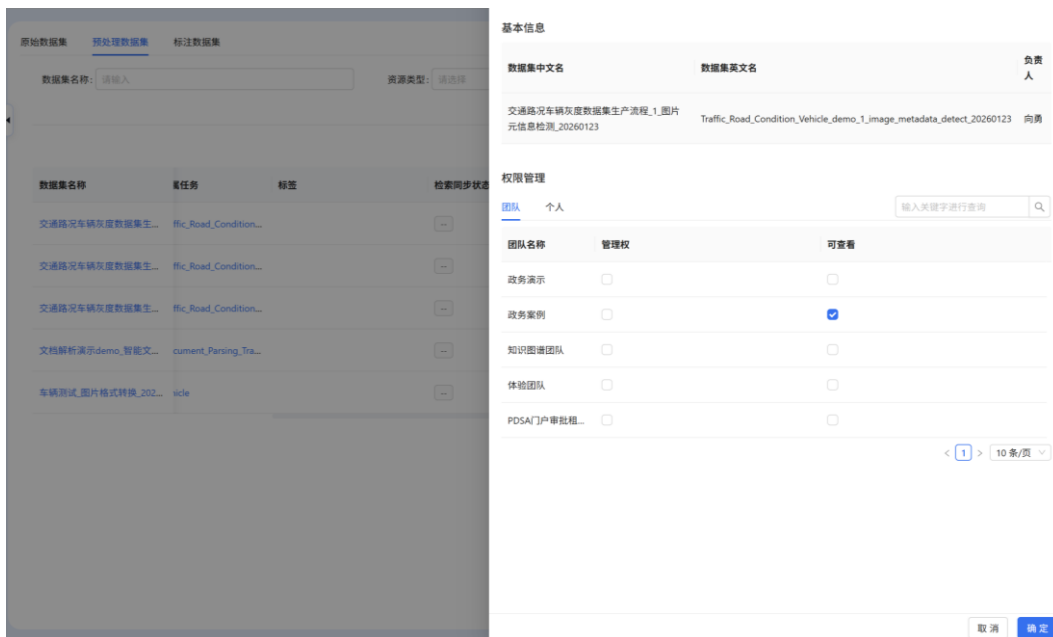
【3】数据集状态：表示该数据集是否生成完成；

【查看】支持查看数据集的详情；

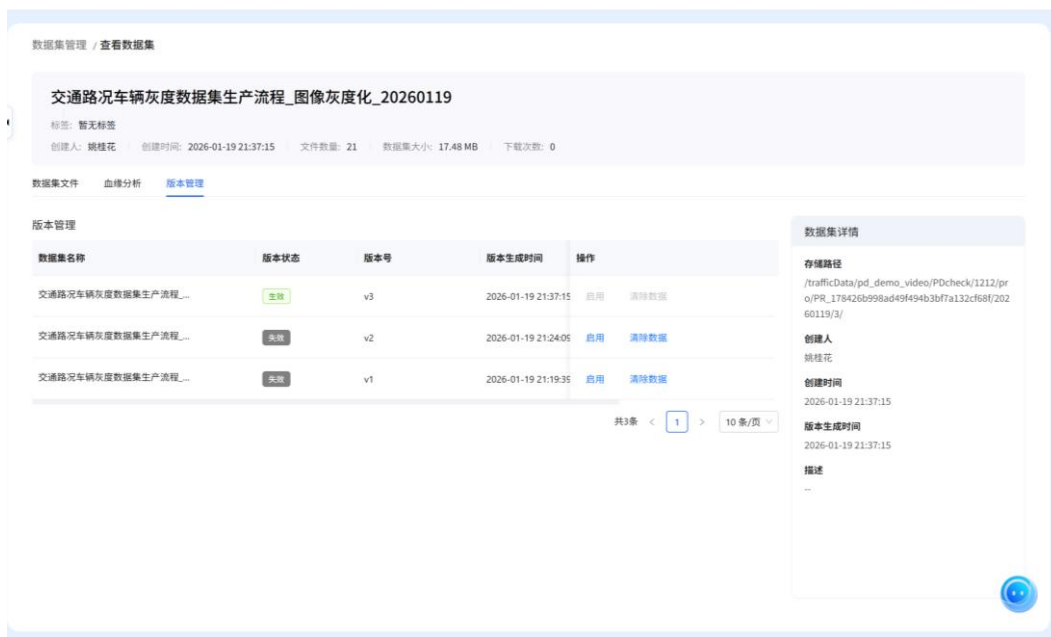


【删除】支持删除该数据集；

【授权】支持将数据集授权给个人或团队；



【版本管理】指能够查看数据集的版本情况，可进行版本管理；同一时刻只能拥有一个有效版本。



【血缘分析】支持查看该数据集的血缘情况；



【评测报告】支持查看该数据集已有的评测报告，点击查看报告可跳转至报告页面；

历史评测报告

评测报告名	报告版本	评测执行时间	归属评测任务	评测对象	评测对象版本	评测结果	操作
Evaluation_of_Electric_Vehicle_report	v4	2026-02-09 16:39:09	Evaluation_of_Electric_Vehicle_	交通路况车辆灰度数据集 集生产流程_1_图片元信息检测_20260123	v1	L4应用级	查看报告
Evaluation_of_Electric_Vehicle_report	v3	2026-02-09 16:05:51	Evaluation_of_Electric_Vehicle_	交通路况车辆灰度数据集 集生产流程_1_图片元信息检测_20260123	v1	L4应用级	查看报告
Evaluation_of_Electric_Vehicle_report	v2	2026-02-03 17:10:05	Evaluation_of_Electric_Vehicle_	交通路况车辆灰度数据集 集生产流程_1_图片元信息检测_20260123	v1	L4应用级	查看报告
Evaluation_of_Electric_Vehicle_report	v1	2026-01-23 16:44:08	Evaluation_of_Electric_Vehicle_	交通路况车辆灰度数据集 集生产流程_1_图片元信息检测_20260123	v1	L4应用级	查看报告

共4条 < 1 > 10条/页

关闭

【评测分析】指用户可选择该数据集的所有评测报告，按照整体及分维度两个方面进行数据集评测得分的趋势分析。

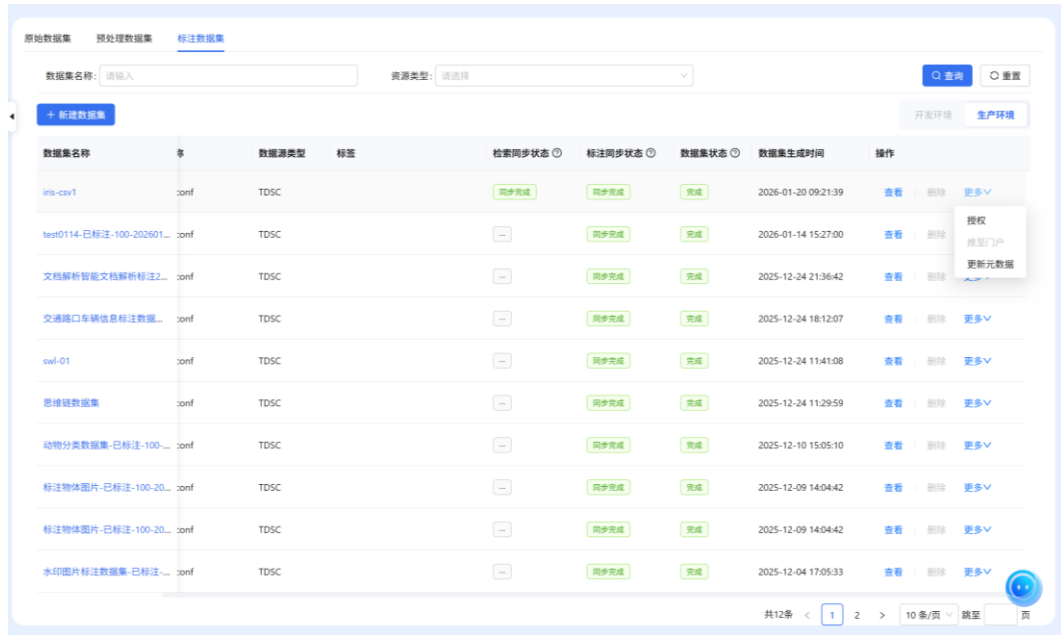


【创建评测】支持快速创建该数据集的评测任务，跳转至评测任务创建页面，详见 1.3.9.2.

【推至门户】当具备数据集门户时操作有效，可以将该数据集推送到门户的管理端。

2.2.3. 标注数据集管理

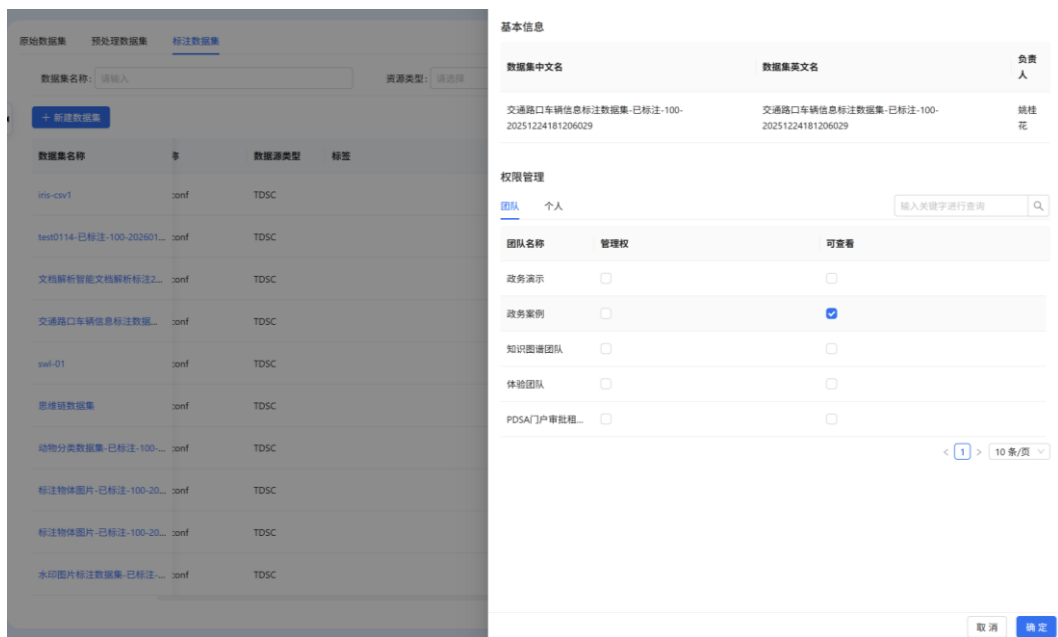
标注数据集是由标注模块产生的数据集。



【查看】支持查看标注数据集详情；



【授权】支持将数据集授权给个人或团队；



【推至门户】当具备数据集门户时操作有效，可以将该数据集推送到门户的管理端。

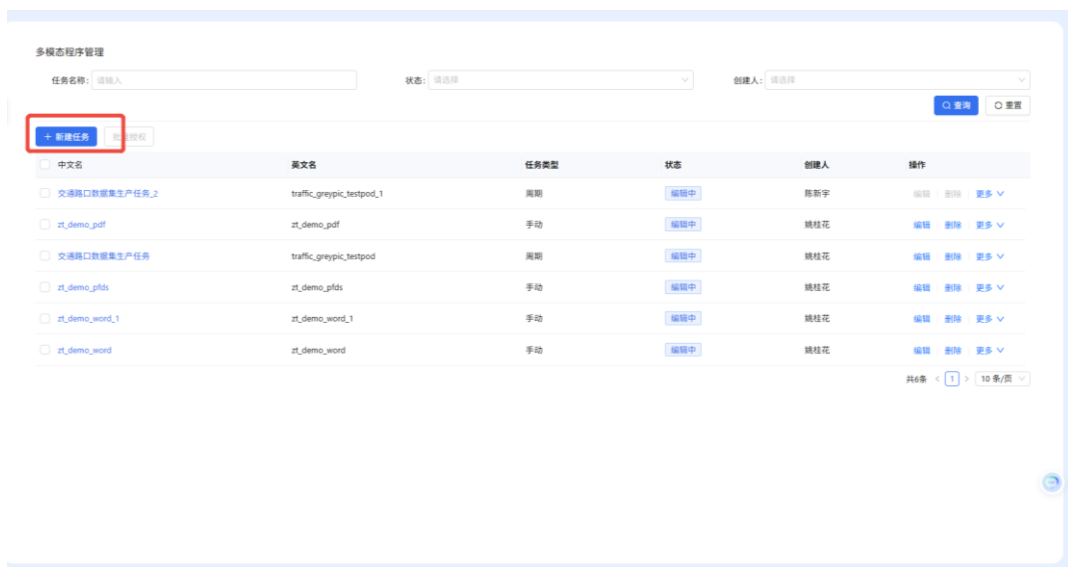
【更新元数据】点击后会将该数据集元数据信息进行更新。

2.3. 创建数据加工任务

本案例以采集监控视频，并对视频进行抽帧，将抽帧图片进行图片灰度化生成数据集为例进行配置指导；

第一步：新建任务画布

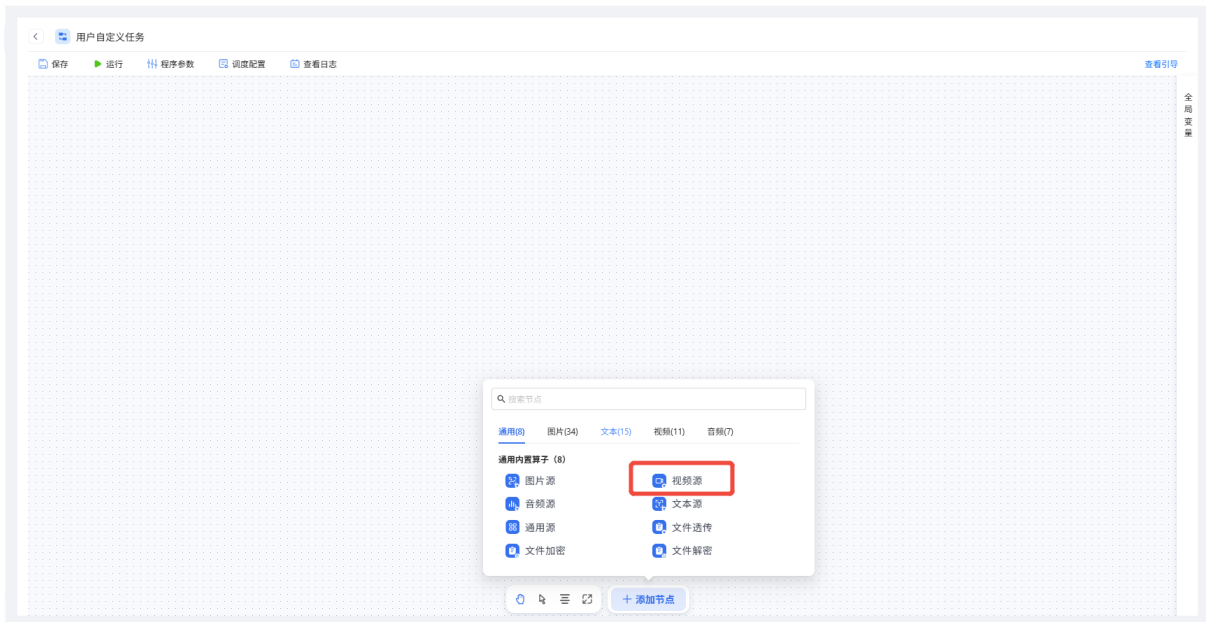
在数据工坊-数据加工中，点击新建任务按钮。



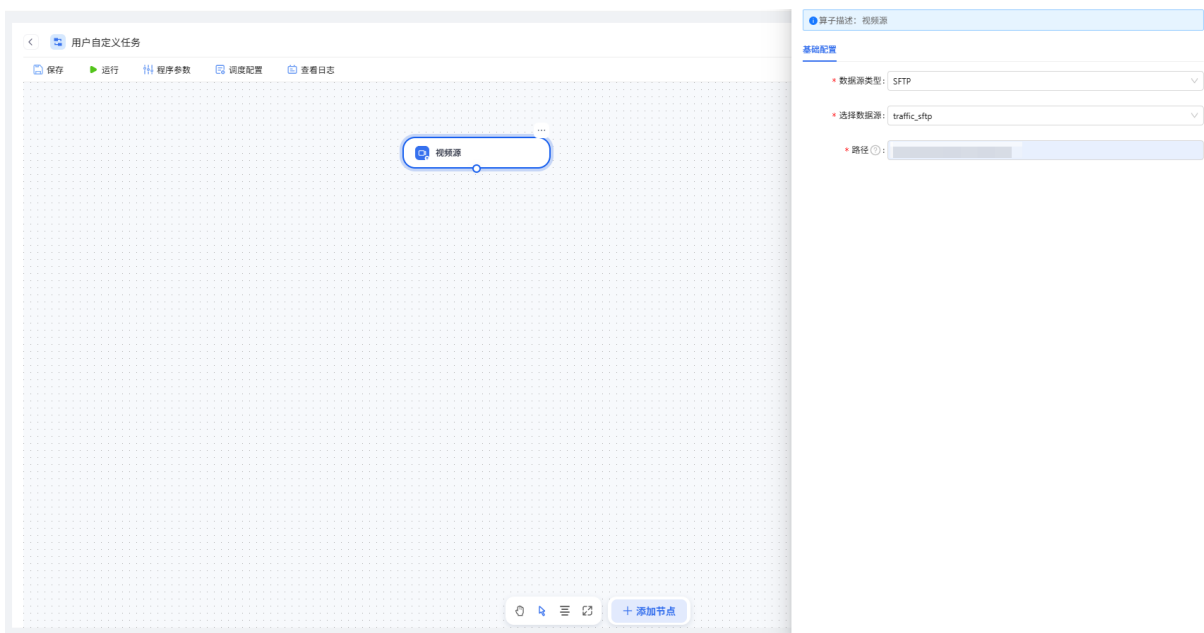
此时会弹出输入任务中英文名的弹窗，选择任务类型为周期或手动，输入后进入任务画布；

周期类型：周期任务提供各类源算子进行周期数据处理，不提供文件上传与流媒体算子；

手动类型：手动任务提供各类源算子进行单次数据处理。同时提供文件上传算子与流媒体源算

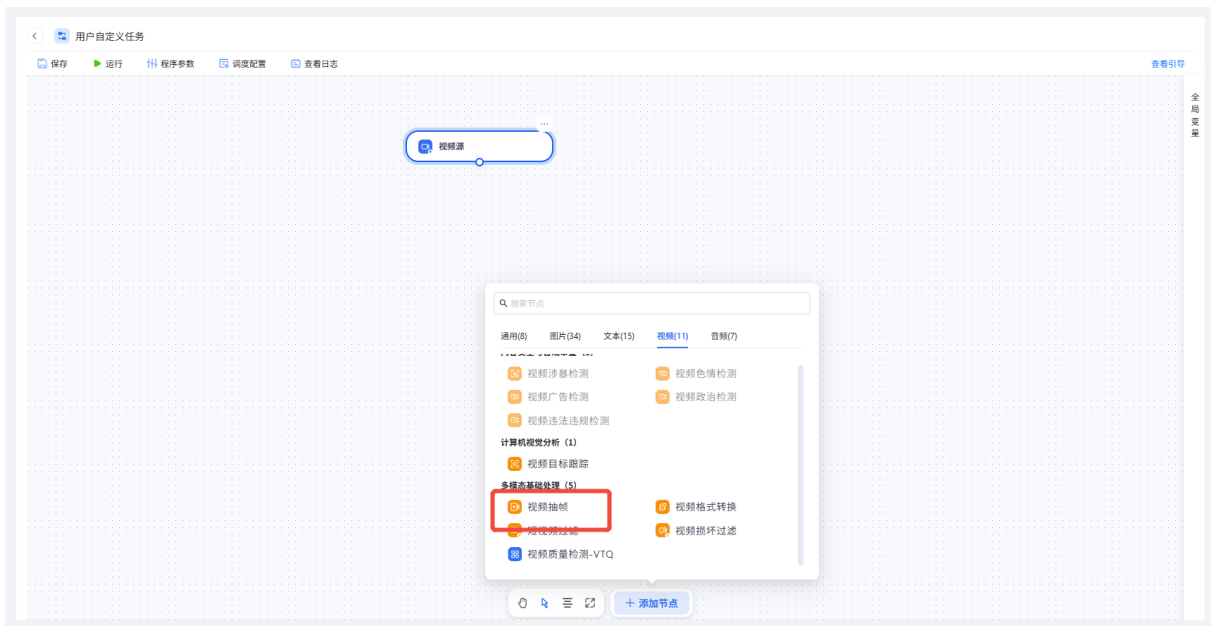


点击【视频源】算子，拖拽到画布中，双击视频源算子进行算子配置；

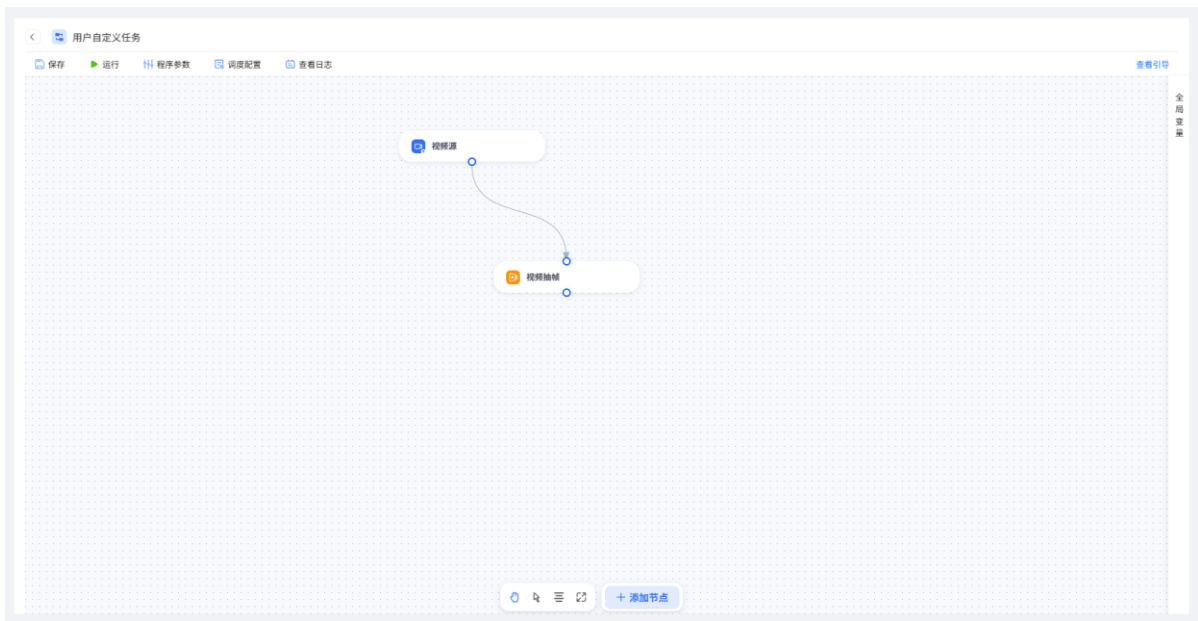


选择需要读取视频的数据源类型及数据源，以及视频存放的文件路径；（注：此处路径可使用全局变量 $\{batchNo\}$ 进行时间替换）点击画布即可保存配置信息；

第三步：点击添加节点，在视频-多模态基础处理分类中，点击【视频抽帧】算子

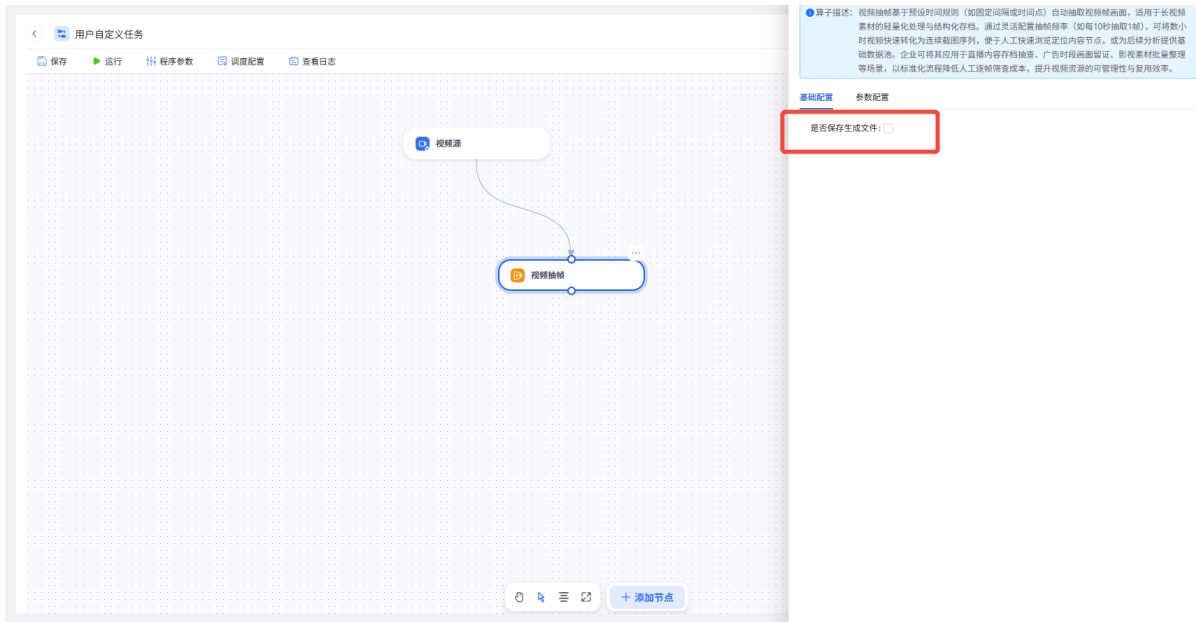


点击算子并将其拖拽到画布中，将【视频源】算子连接到【视频抽帧】算子；

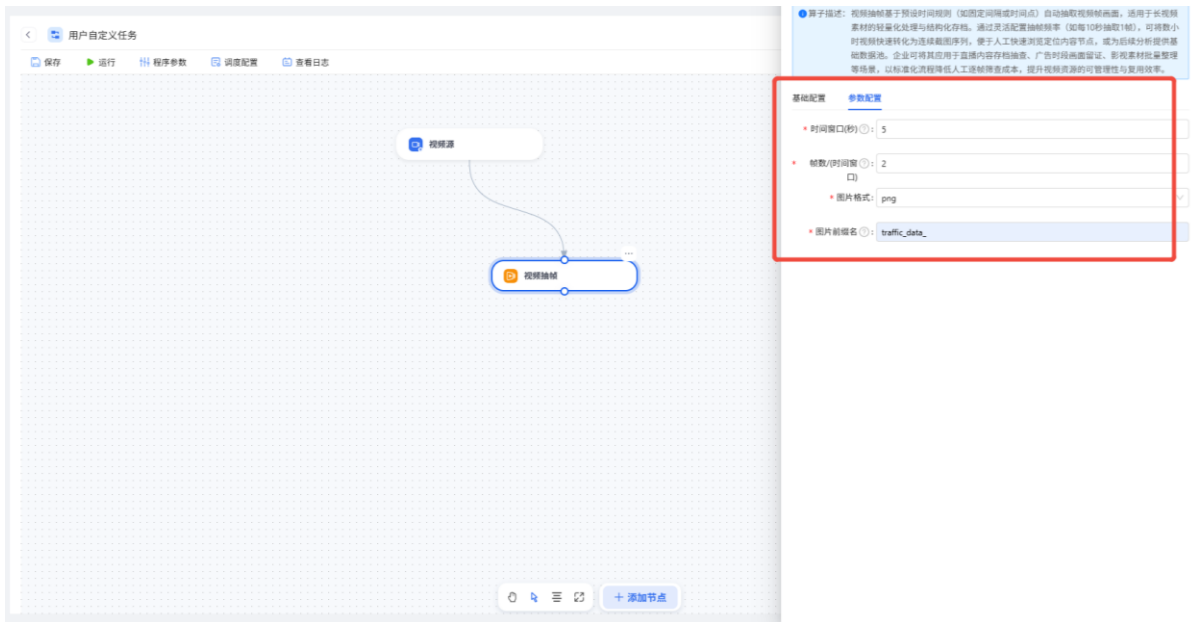


双击【视频抽帧】算子，进行抽帧配置；

算子配置-基础配置中，有是否保存生成文件选项，此配置项表示会将抽帧得到的图片存储至指定的存储路径，若不勾选此项，不会额外存储抽帧结果，也不影响后续算子接收文件。



点击参数配置，此处可配置抽帧规则，根据需要进行抽帧规则的配置。



时间窗口：抽取自定义帧数的单位时间；

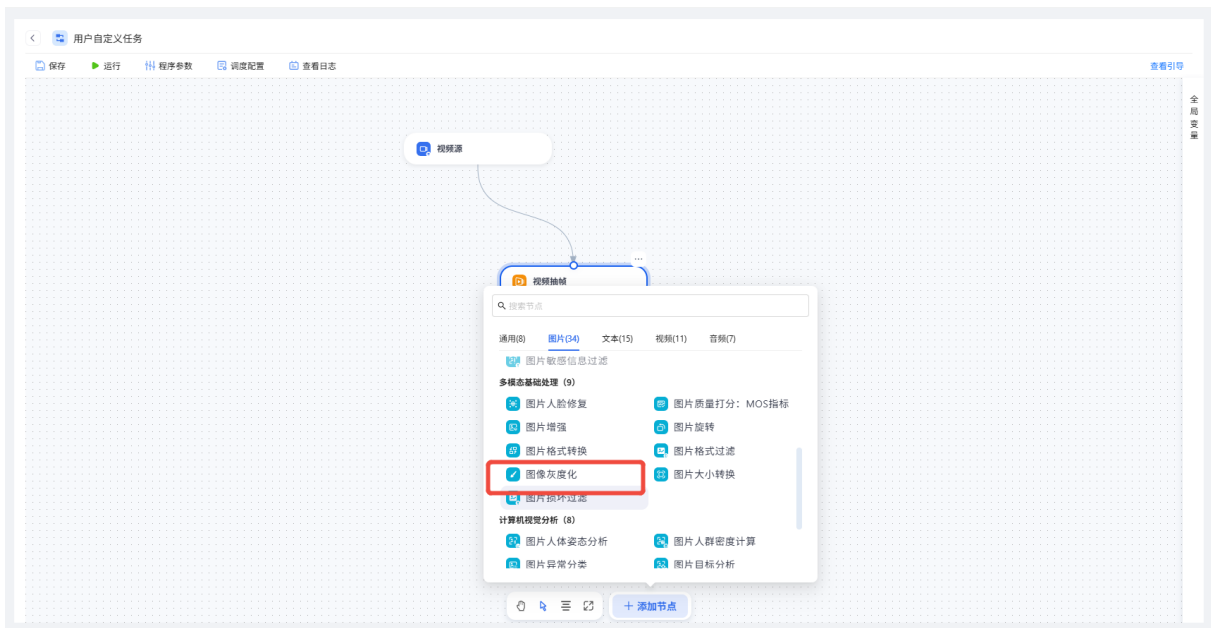
帧数/（时间窗口）：每单位时间窗口内采集的帧数，比如，时间窗口为 2 秒，帧数为 4，即每 2 秒时间内均匀抽取 4 帧；

图片格式：抽取帧图片保存为何种格式；

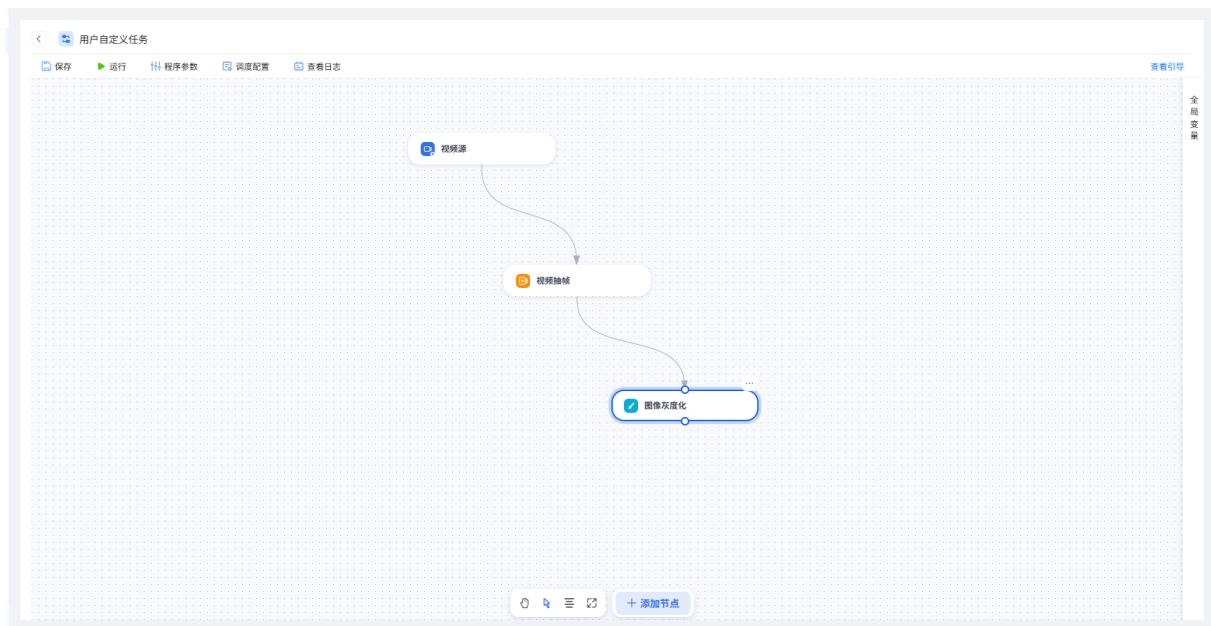
图片前缀名：表示保存后的图片命名前缀；

配置完成后，点击画布保存配置信息；

第四步：点击添加节点，在图片-多模态基础处理分类中，点击【图像灰度化】算子



点击算子并将其拖拽到画布中，将【视频抽帧】算子连接到【图像灰度化】算子；



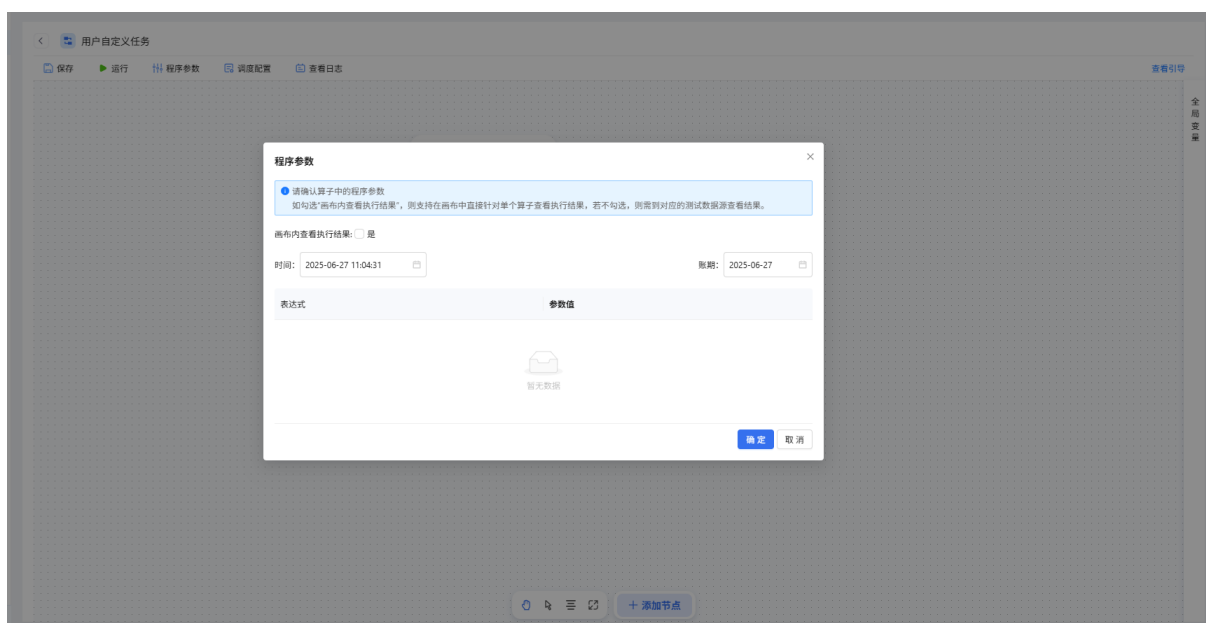
双击【图像灰度化】算子，进行算子配置；

由于该算子无需用户额外配置参数，所以算子配置里仅有基础配置，用户可选择是否另外保存算子输出的文件至指定路径；



所有算子配置完毕后，将整个画布进行保存，可以进行测试；

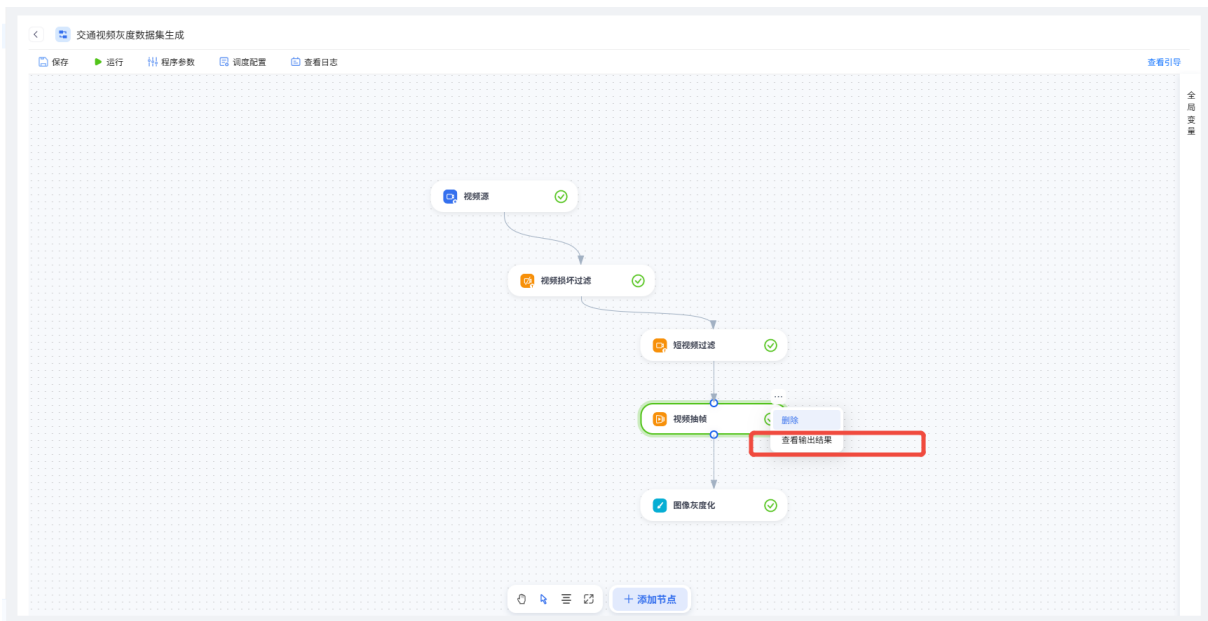
点击运行，会弹出程序参数确认框，若您使用了全局变量，将会在此处展示；



弹窗中配置项【画布内查看执行结果】，勾选后才能够看到在画布中看到算子运行后的结果；本案例勾选为是；

点击确定后，任务开始运行测试；

测试完成后，针对处理过程的算子，可以点击算子右上角的更多选项，查看算子输出结果：



【视频抽帧】算子结果查看:

The screenshot shows the '视频抽帧' (Video Frame Extraction) operator result viewer. It displays a grid of traffic video frames and a table with columns: file_path, sample_file_path, file_name, file_length, and 操作 (Action).

file_path	sample_file_path	file_name	file_length	操作
/tmp/tmp8ufd5mce/traffic_data_33.png	dmodel/resource/sample_data/PR_867833694e3d40b095d608e7966...	traffic_data_33.png	2769611	预览
/tmp/tmp8ufd5mce/traffic_data_5005.png	dmodel/resource/sample_data/PR_867833694e3d40b095d608e7966...	traffic_data_5005.png	2944775	预览
/tmp/tmp8ufd5mce/traffic_data_10010.png	dmodel/resource/sample_data/PR_867833694e3d40b095d608e7966...	traffic_data_10010.png	2926827	预览
/tmp/tmp8ufd5mce/traffic_data_15015.png	dmodel/resource/sample_data/PR_867833694e3d40b095d608e7966...	traffic_data_15015.png	2907133	预览
/tmp/tmp8ufd5mce/traffic_data_20020.png	dmodel/resource/sample_data/PR_867833694e3d40b095d608e7966...	traffic_data_20020.png	2782195	预览
			7764851	预览

【图像灰度化】算子结果查看:



确认无误后，可进行调度配置，可直接采用默认值，或根据业务需要调整优先级、执行周期、生效和失效时间等；



保存调度信息后，返回任务列表，进行任务上线，当上线申请审批通过后，任务生效；

中文名	英文名	状态	创建人	操作
<input type="checkbox"/> 用户自定义任务	user_sun0	编辑中	admin	编辑 删除 更多
<input type="checkbox"/> test_hj	test_hj	生效	admin	编辑 删除 上线 下线
<input type="checkbox"/> 交通规则及度数据生成	traffic_data	变更	admin	编辑 删除 更多
<input type="checkbox"/> test_http	test_http	编辑中	admin	编辑 删除 更多
<input type="checkbox"/> test	test	编辑中	admin	编辑 删除 更多
<input type="checkbox"/> test0617	test0617	编辑中	admin	编辑 删除 更多
<input type="checkbox"/> update_file_20250612	update_file_20250612	编辑中	admin	编辑 删除 更多
<input type="checkbox"/> 长链接文本	test_long_text	变更	admin	编辑 删除 更多
<input type="checkbox"/> test_003	test_003	编辑中	admin	编辑 删除 更多
<input type="checkbox"/> 建站方案任务	test_sghd	编辑中	admin	编辑 删除 更多

2.4. 数据评测

2.4.1. 规则定义

点击规则列表的【新建规则】按钮，跳转至规则定义页面；

规则名称	规则对象	所属维度	规则状态	规则描述	创建人	创建时间	操作
图片格式规范	图片	规范性	禁用		姚桂花	2025-10-31 09:42:54	启用 编辑 删除
文件大小	图片	规范性	启用	文件大小	姚桂花	2025-10-31 10:03:22	禁用 编辑 删除

本案例以配置“图片宽高校验”来进行说明；

第一步：维护规则的基础信息，包括规则名称、规则对象（指该规则针对的数据类型）、所属维度（定义该规则所属的维度）、规则描述等；

规则基础信息

规则名称: 图片宽高校验 规则对象: 图片 所属维度: 有效性

规则描述: 校验图片宽度是否满足1920及以上且高度满足1080及以上

第二步：定义该规则的模板及检测规则；

规则模板：分为指标匹配和指标统计两种；

指标匹配：选定一个或多个指标，设定匹配值，根据每个数据样本的该项元数据的值与条件进行判定，判断是否满足来计算分值；

指标统计：选定多个指标，设定包含数量值，根据每个数据样本具备 x 个已选定指标来计算分值。

计算指标：下拉选择，可多选，选择需要进行条件判断的指标项；

命中条件：配置指标值的判定条件的区域；可配置多个判断条件；

满足以下【所有】条件视为样本命中：指配置的所有条件，需要全部满足才会判定为样本满足规则，否则视为不满足条件，即【&】关系；

满足以下【任意】条件视为样本命中：指配置的所有条件，只需满足任一个就会判定为样本满足规则，即【||】关系；

计算公式：指标匹配固定项，用于计算该规则的命中比率，公式为：

$$\text{命中样本比率 (计算结果)} = \text{命中样本数} / \text{总样本数}$$

分值定义：用于定义命中比率满足某个区间时，规则得分情况；

如：命中比率 90%以上，该条规则得分为 100，即命中比率大于等于 0.9 且小于等于 1 时，该条规则得 100 分；

如图，本次案例配置的分值定义含义为：

当计算的命中比率大于等于 0.9 且小于等于 1 时，该规则得 100 分；

当计算的命中比率大于等于 0.6 且小于 0.9 时，该规则得 60 分；

当计算的命中比率大于等于 0.4 且小于 0.6 时，该规则得 40 分；

当计算的命中比率小于 0.4 时，该规则得 0 分；

如：当一个数据集总样本数为 100，该规则的命中样本数为 76，则命中比率为 0.76，此时该规则得分为 60 分。

此处需注意：判定值为计算的命中比率，一定为 0~1 之间，配置时注意选择的判定条件，所有判定区间的并集需要满足 [0,1]。

规则定义

规则模板: 指标匹配 计算指标: 宽度 × 高度

命中条件: 满足以下 所有 条件视为样本命中

宽度: 数值 大于等于 1920

高度: 数值 大于等于 1080

+ 添加条件

计算公式: $X=A/B$ 请按照既定参数进行公式编写。X:计算结果; A:命中样本数; B:总样本数

分值定义	规则得分	判定指标	判定条件	判定值	操作
100	计算结果	左右闭区间(a,b)	0.9	- 1	删除
60	计算结果	左闭右开区间(a,b)	0.6	- 0.9	删除
40	计算结果	左闭右开区间(a,b)	0.4	- 0.6	删除
0	计算结果	小于	0.4		删除

+ 添加一行

点击确定后, 规则创建完成, 此时为禁用状态, 表示该规则还不能用于评测任务。
在规则管理列表, 点击启用, 即可将规则转化为启用状态, 此时可用于评测任务。

规则列表

规则名称: 请输入规则名称 规则对象: 请选择规则对象 所属维度: 请选择所属维度

查询 重置

+ 新建规则

规则名称	规则对象	所属维度	版本状态	规则指标描述	创建人	创建时间	操作
图片宽高校验	图片	有效性	禁用	校验图片宽度是否满足1920及以上且高度满足1080及以上	姚桂花	2025-11-10 18:36:08	启用 编辑 删除
图片格式规范	图片	规范性	禁用		姚桂花	2025-10-31 09:42:54	启用 编辑 删除
文件大小	图片	规范性	启用	文件大小	姚桂花	2025-10-31 10:03:22	禁用 编辑 删除

共 3 条 < 1 > 10 条/页

规则列表

规则名称: 请输入规则名称 规则对象: 请选择规则对象 所属维度: 请选择所属维度

查询 重置

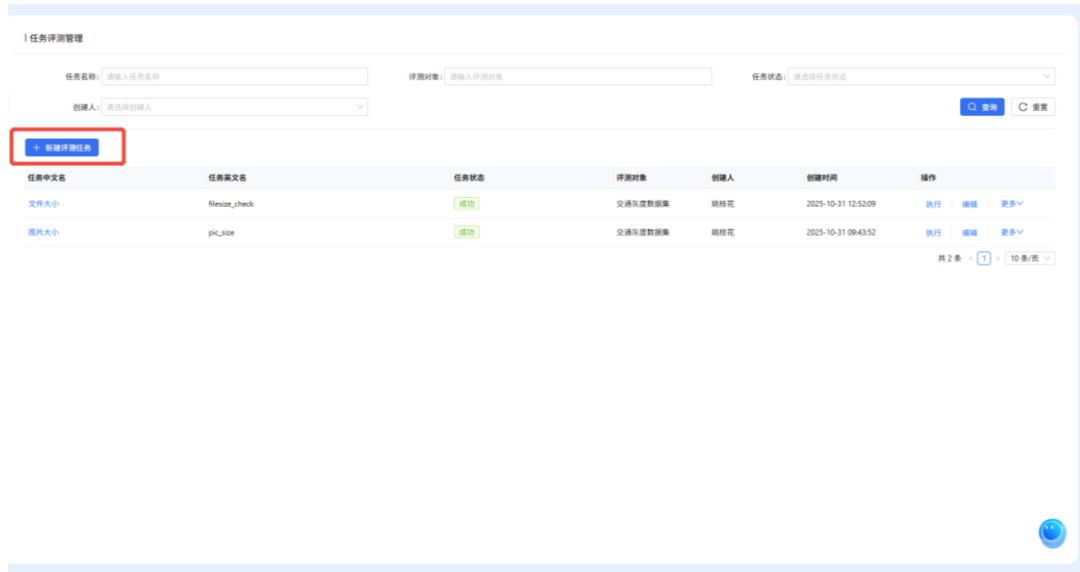
+ 新建规则

规则名称	规则对象	所属维度	版本状态	规则指标描述	创建人	创建时间	操作
图片宽高校验	图片	有效性	启用	校验图片宽度是否满足1920及以上且高度满足1080及以上	姚桂花	2025-11-10 18:36:08	禁用 编辑 删除
图片格式规范	图片	规范性	禁用		姚桂花	2025-10-31 09:42:54	启用 编辑 删除
文件大小	图片	规范性	启用	文件大小	姚桂花	2025-10-31 10:03:22	禁用 编辑 删除

共 3 条 < 1 > 10 条/页

1. 创建评测任务

在评测任务管理列表，点击【新建评测任务】按钮，跳转至评测任务新建页面。



本案例以“交通灰度数据集评测”来进行说明；

第一步：维护任务基础信息，包括任务中文名及任务英文名；

注意：任务英文名具有唯一性校验，不要重复；

任务基础信息

* 任务中文名: 交通数据集评测

* 任务英文名: traffic_grey_data_Evaluation

第二步：维护评测配置

评测对象类型： 主要指需要评测的数据集的数据类型，分为图片、音频、视频、文本四种；

评测对象： 选择在已选的类型中，需要进行评测的数据集（如类型选择图片，则此处只可选择图片类型的数据集）

评测对象版本： 由于数据集有版本概念，所以此处需要选择针对哪个版本进行评测；

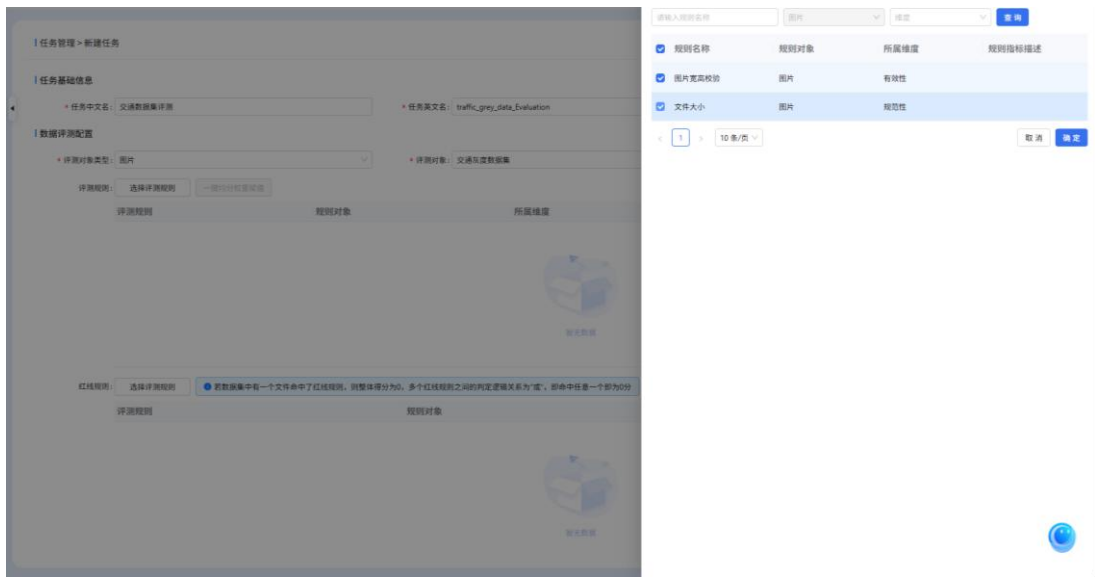
数据评测配置

* 评测对象类型: 图片

* 评测对象: 交通灰度数据集

* 评测对象版本: V1

评测规则：点击【选择评测规则】按钮，右侧弹出规则列表，勾选需要的规则；



点击确定后，选择的规则就会回显到任务配置页面，此时可针对每个规则设置权重。

评测规则:

评测规则	规则对象	所属维度	分值权重(%)	操作
图片宽高校验	图片	有效性	<input type="text" value="50"/>	删除
文件大小	图片	规范性	<input type="text" value="50"/>	删除

如图，分值权重为 50%，表示该条规则的得分在评测总体得分中占比 50%；如图中两条规则权重各为 50%，规则得分分别为 100 分和 76 分，结合权重后该评测总体得分为 $100*50\%+76*50\%=50+38=88$ 分。

红线规则：指用于判定合法性的规则，通常选择合法合规检测类规则，用于拦截含有违法违规样本的数据集，当存在红线规则时，优先判定红线规则是否命中，若红线规则命中，则数据集判定违规，分数为 0。



数据集定级配置：根据信通院标准，将数据集评测等级分为 L1~L5 五个等级，根据规则+权重的综合计算方式，按照分值落位的区间判定该数据集所属的等级。

此处需注意：分值范围所有区间的并集需满足 $[0, 100]$ 。

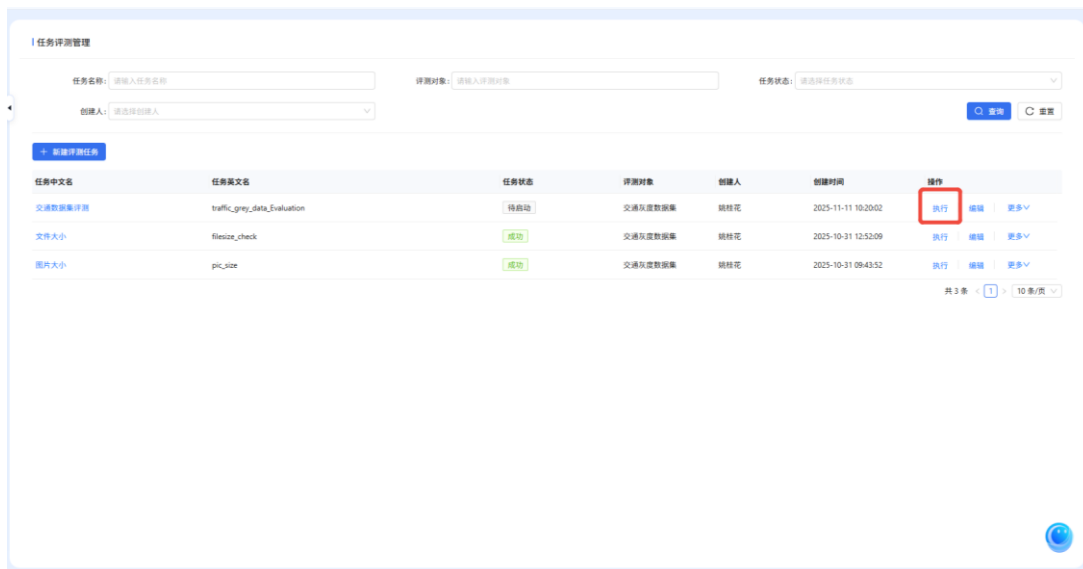
数据集定级配置

数据集定级: 分值范围

L1 1级	等于	<input type="text" value="100"/>
L2 2级	左闭右开区间(a,b)	<input type="text" value="90"/> <input type="text" value="100"/>
L3 3级	左闭右开区间(a,b)	<input type="text" value="80"/> <input type="text" value="90"/>
L4 4级	左闭右开区间(a,b)	<input type="text" value="70"/> <input type="text" value="80"/>
L5 5级	小于	<input type="text" value="70"/>

[返回](#)

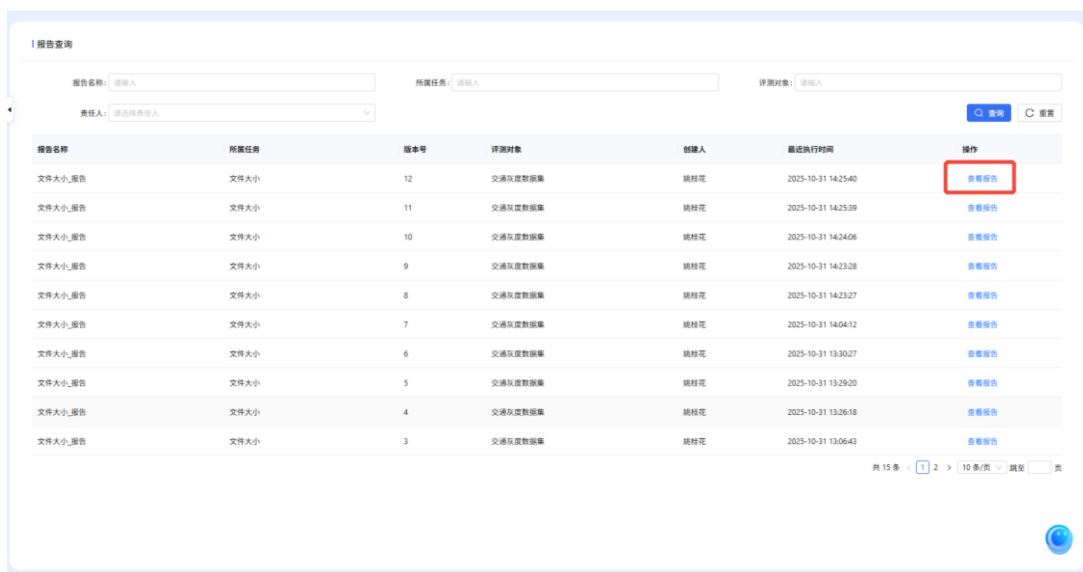
配置完成后，点击确定，会返回到任务管理列表，点击列表中的执行按钮，开始进行评测；



执行成功后，可在评测报告页面查看评测报告。

2. 查看评测报告

进入评测报告管理列表，可看到所有的评测报告，包括多次执行同一个评测任务的不同版本的报告，点击查看报告，可进入报告详细页面。



在报告页面，可以看到具体的报告情况。



至此本案例结束。

2.5. 新增算子

2.5.1. 新增引用内置算法的算子

本案例以引用内置图片灰度化算法，新建**图片灰度化算子**为例进行配置操作指导：

第一步，新建算子

在算子管理列表，点击【新建】按钮，进入算子新增页面：

算子管理列表

算子名称: 算子分类: 算子状态:

算法来源: 创建人:

+ 新建

<input type="checkbox"/>	算子中文名称	算子英文名称	算子分类	算法来源	算子描述	状态	创建人	创建时间	操作
<input type="checkbox"/>	视频帧率调整	video_fps_generator	视频/多模态基础处理/内容处理	内置算法	按配置调整视频的帧率。...	生效		2026-03-13 14:12:48	编辑 测试 更多
<input type="checkbox"/>	视频目标轨迹标注	video_target_track_generator	视频/计算机视觉分析/实体识别	内置算法	视频目标跟踪实现80类物体...	生效		2026-03-13 14:12:48	编辑 测试 更多
<input type="checkbox"/>	视频格式转换	video_format_conv	视频/多模态基础处理/格式转化	内置算法	基于硬件加速编解码引擎。...	生效		2026-03-13 14:12:48	编辑 测试 更多
<input type="checkbox"/>	短视频过滤	video_short_filter	视频/多模态基础处理/质量检测	内置算法	过滤时长少于自定义阈值的...	生效		2026-03-13 14:12:48	编辑 测试 更多
<input type="checkbox"/>	视频帧丢失过滤	video_loss_filter	视频/多模态基础处理/质量检测	内置算法	视频帧丢失过滤可深度检测视...	生效		2026-03-13 14:12:48	编辑 测试 更多
<input type="checkbox"/>	视频剪辑	video_clip_generator	视频/多模态基础处理/内容处理	内置算法	按配置的帧头与帧结尾...	生效		2026-03-13 14:12:48	编辑 测试 更多
<input type="checkbox"/>	视频分辨率调整	video_resolution_ratio_generator	视频/多模态基础处理/内容处理	内置算法	按配置调整视频的分辨率。...	生效		2026-03-13 14:12:48	编辑 测试 更多
<input type="checkbox"/>	视频编码格式检测	video_codec_detect	视频/多模态基础处理/质量检测	内置算法	检查视频文件的编解码...	生效		2026-03-13 14:12:48	编辑 测试 更多
<input type="checkbox"/>	视频CRC检测	video_crc_detect	视频/多模态基础处理/质量检测	内置算法	使用CRC32循环校验算法。...	生效		2026-03-13 14:12:48	编辑 测试 更多
<input type="checkbox"/>	视频AI增强	video_ai_enhance_generator	视频/媒体增强与智能创作/智能标记与生成	内置算法	通过Real-ESRGAN深度学习...	生效		2026-03-13 14:12:48	编辑 测试 更多

共 129 条 < 1 2 3 4 5 ... 13 > 10 条/页 跳至 页

第二步，填写算子基础信息

在跳转的算子新建页面，填写算子英文名、中文名、选择算子标签、填写算子描述、选择算子输入、输出类型、选择算法来源及算法类型；

算子管理列表 > 新建算子

算子基础信息

- 算子英文名称: 英文、数字、下划线, 以英文开头, 最低4字符, 最多40字符
- 算子中文名称: 中文、英文、数字、下划线, 最低4字符, 最多40字符
- 算子标签: 请选择标签
- 算子描述: 建议不低于100字, 介绍算子作用、所用的算法及对应的参数如何使用
- 算子输入类型: 图片 视频 文本 音频 通用
- 算子输出类型: 图片 视频 文本 音频 通用
- 算法来源: 内置算法 内置模型服务 第三方Http
- 算法类型: 分析类 生成类

算子英文名称：支持英文、数字、下划线，必须以英文开头，最少 4 字符，最多 40 字符，具备唯一性；

算子中文名称：支持中文、英文、数字、下划线，最少 4 字符、最多 40 字符；

算子标签：下拉选择，表示该算子归属于哪一场景；

算子描述：用于详细描述算子的功能，建议填写 100 字以上，详细描述算子的作用及算子对应的业务价值等信息；

算子输入类型：表示该算子是处理何种类型的数据，本案例选择图片；

算子输出类型：表示该算子输出何种类型的数据，本案例选择图片；

算法来源：表示算法的来源，本案例选择内置算法；

算法类型：表示算法是生成类算法还是分析类算法，本案例选择生成类；表示该算法是模态到模态的算法；

第三步，定义算法信息

由于本案例选择的算法来源是内置算法，此处可在所有内置算法中进行选择；

算子管理列表 > 新建算子

算子基础信息

算子英文名称: 英文、数字、下划线, 以英文开头, 最低4字符, 最多40字符

算子中文名称: 中文、英文、数字、下划线, 最低4字符, 最多40字符

算子标签: 请选择标签

算子描述: 建议不低于100字, 介绍算子作用、所用的算法及对应的参数如何使用

算子输入类型: 图片 视频 文本 音频 通用

算子输出类型: 图片 视频 文本 音频 通用

算法来源: 内置算法 内置模型服务 第三方Http

算法类型: 分析类 生成类

算法信息

选择算法: [下拉菜单]

前端组件定义

控制类型

用枚举	是否必填	默认值	配置说明	操作
<input type="checkbox"/>	<input checked="" type="checkbox"/>			删除

算子基础配置信息

展示生成文件配置项: 是 否

注意: 此处若勾选“是”, 将会在算子基础配置部分展示保存生成文件的相关配置项(数据来源和路径)。建议生成类的算子此处选择“是”。

展示分析结果配置项: 是 否

注意: 此处若勾选“是”, 将会在算子基础配置部分展示保存分析结果的相关配置项。算子的分析结果会保存到数据表中, 建议分析类的算子此处选择“是”。

本案例选择【图像灰度化】算法；

第四步，查看前端组件定义

由于内置算法的前端组件是固定且不可更改的，根据所选算法不同，前端组件定义也不同，此处选择的【图像灰度化】算法无需用户填写参数，故前端组件定义为空。



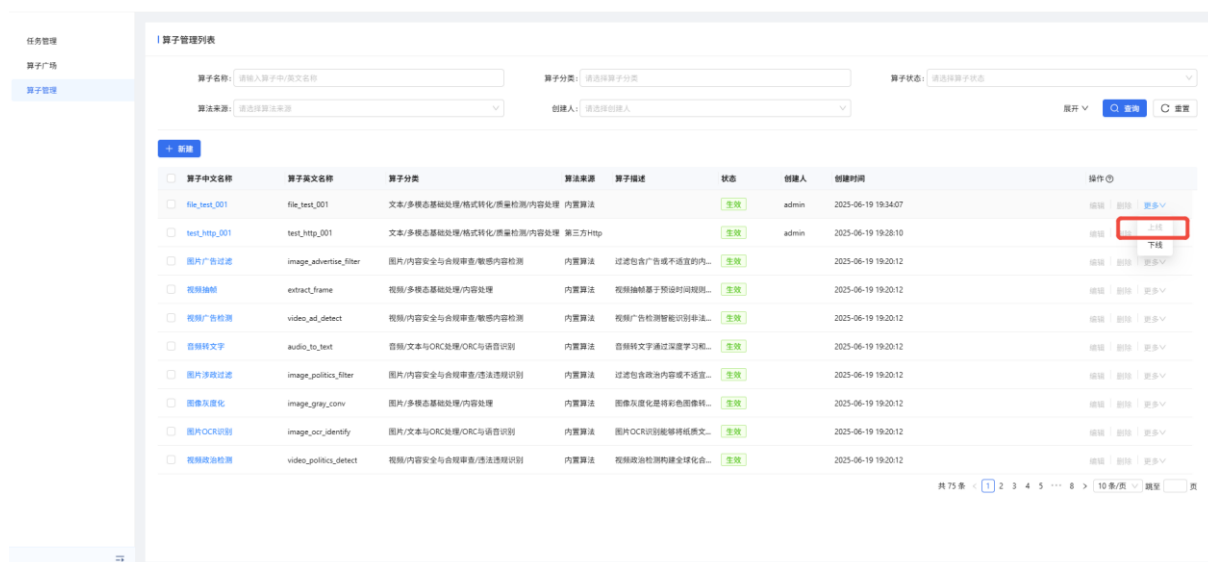
第五步，设置算子基础配置信息

最后设置该算子是否保存生成的文件或是生成的分析结果。一般来说，生成类算法会生成新的文件，此时建议勾选展示生成文件配置项，用于配置生成文件的存储路径；分析类算法会生成可解析的半结构化分析结果，此时建议勾选展示分析结果配置项，用于配置分析结果的存储表及映射关系。



第六步，保存算子

点击右下角【确定】按钮，即可保存算子，此时算子状态为【编辑中】。可在算子管理列表进行算子上线操作，上线后，算子会在算子广场及任务画布中展示。



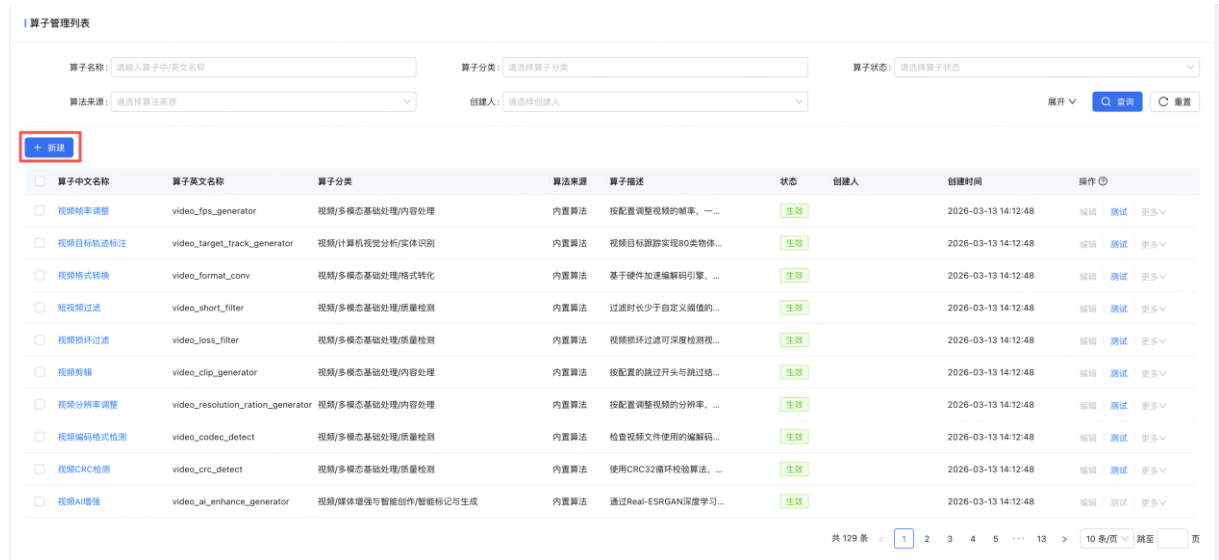
自此本案例结束。

2.5.2. 新增引用内置算法模型的算子

本案例以引用图片涉黄检测算法，新建**图片涉黄检测**算子为例进行配置指导。

第一步，新建算子

在算子管理列表，点击【新建】按钮，进入算子新增页面；



第二步，填写算子基础信息

在跳转的算子新建页面，填写算子英文名、中文名、选择算子标签、填写算子描述、选择算子输入、输出类型、选择算法来源及算法类型；



算子英文名称：支持英文、数字、下划线，必须以英文开头，最少 4 字符，最多 40 字符，具备唯一性；

算子中文名称：支持中文、英文、数字、下划线，最少 4 字符、最多 40 字符；

算子标签：下拉选择，表示该算子归属于哪一场景；

算子描述：用于详细描述算子的功能，建议填写 100 字以上，详细描述算子的作用及算子对应的业务价值等信息；

算子输入类型：表示该算子是处理何种类型的数据，本案例选择图片；

算子输出类型：表示该算子输出何种类型的数据，本案例选择图片；

算法来源：表示算法的来源，本案例选择内置模型服务；

算法类型：表示算法是生成类算法还是分析类算法，本案例选择分析类；表示该算法是模态到数据的算法；

第三步，定义算法信息

由于本案例选择的算法来源是内置模型服务，此处可在所有内置模型服务中进行选择；



本案例选择【图片涉黄检测】算法

第四步，查看输入参数定义

由于内置模型服务的算法输入参数与返回参数已经固定，所以配置页面提供 body 参数与返回结果的示例，帮助您进行前端组件的定义与输出字段的定义。



第五步，配置输出参数

所选算法的返回示例如下：

```
{
  "status":0,
  "result":[
    {
      "id":"0"
      "size":"680,510",
      "objects":[
        {
          "cls_name":"裸露",
          "confidence":"0.575",
          "xywh":"0.08392,0.16195,0.87576,0.99976"
        }
      ]
    }
  ]
  "message":"success"
}
```

此时，在配置页面的输出参数部分，可以定义算子输出字段与算法输出字段的映射关系，按照如下方式进行配置：

某分析类算法可以识别图片中的物体并返回中英文结果，某次执行的返回结果如下：

```
{
  "status":0,
  "result":[
    {
      "tag":"image_recognition",
      "en_tags":[
        "car",
        "cat"
      ],
      "cn_tags":[
        "汽车",
        "猫"
      ]
    }
  ],
  "message":"success"
}
```

我们所需的结果为中文名和英文名，则在输出参数中配置为：

算子输出字段	算法输出字段
en_tags	<code>\$.result[*].en_tags</code>
cn_tags	<code>\$.result[*].cn_tags</code>

此时算子输出字段 en_tags 取值为：`["car", "cat"]`；算子输出字段 cn_tags 取值为：`["汽车"、"猫"]`

所以最终该算子的输出参数配置如下：

返参取值映射: 字段类型	算子输出字段	算子输出字段中文名	算法输出字段	操作
String	result_pic		<code>\$.result[*].object[*].cls_name</code>	删除
String	confidence		<code>\$.result[*].object[*].confidence</code>	删除
String	xywh		<code>\$.result[*].object[*].xywh</code>	删除

[配置案例说明](#) + 添加

字段类型	算子输出字段	算法输出字段
string	result_pic	<code>\$.result[*].object[*].cls_name</code>
string	confidence	<code>\$.result[*].object[*].confiden</code>

		ce
string	xywh	\$.result[*].object[*].xywh

字段类型与算子输出字段根据实际情况及需要进行修改，算法输出字段的根据算法返回示例中的 json 填写对应的取值表达式。

第六步，配置前端组件

定义前端组件时，需要查看算法的输入参数，识别是否需要用户进行参数输入。如本案例所需那的算法输入的 body 参数如下：

```
[
  {
    "data_type": "image",
    "id": "",
    "data": "${data_base64}"
  }
]
```

这里需要注意的是，若算法需传入的参数为文件，可以使用内置变量 `data_base64`，在 body 参数中将该变量填写到对应的存放位置中（如上面输入框的 data 参数），系统会将文件转换为 Base64 形式，并传输给算法。

分析算法的输入参数可知，本案例中所选算法无需用户输入额外参数，所以前端组件定义为空即可。



第七步，设置算子基础配置信息

最后设置该算子是否保存生成的文件或是生成的分析结果。一般来说，生成类算法会生成新的文件，此时建议勾选展示生成文件配置项，用于配置生成文件的存储路径；分析类算法会生成可解析的半结构化分析结果，此时建议勾选展示分析结果配置项，用于配置分析结果的存储表及映射关系。

算子基础配置信息

* 显示保存数据集配置项: 是 否

注意:

此处若勾选“是”，将会在算子基础配置部分展示出保存数据集的相关配置项（数据源和路径），建议生成类的算子此处选择“是”。

* 显示分析结果配置项: 是 否

注意:

此处若勾选“是”，将会在算子基础配置部分展示出分析结果的相关配置项，算子的分析结果会保存到数据表中，建议分析类的算子此处选择“是”。

第八步，保存算子

点击右上角【确定】按钮，即可保存算子，此时算子状态为【编辑中】。可在算子管理列表进行算子上线操作，上线后，算子会在算子广场及任务画布中展示。

自此本案例结束。

2.5.3. 新增引用第三方Http算法的分析类算子

本案例通过调用第三方算法的Http接口进行暴力检测算子的新增

第一步，新建算子

在算子管理列表，点击【新建】按钮，进入算子新增页面；

算子中文名称	算子英文名称	算子分类	算法来源	算子描述	状态	创建人	创建时间	操作
视频帧率调整	video_fps_generator	视频/多模态基础处理/内容处理	内置算法	按配置调整视频的帧率。...	生效		2026-03-13 14:12:48	编辑 测试 更多
视频目标轨迹标注	video_target_track_generator	视频/计算机视觉分析/物体识别	内置算法	视频目标跟踪实现80类物体...	生效		2026-03-13 14:12:48	编辑 测试 更多
视频格式转换	video_format_conv	视频/多模态基础处理/格式转化	内置算法	基于硬件加速编解码引擎。...	生效		2026-03-13 14:12:48	编辑 测试 更多
短视频过滤	video_short_filter	视频/多模态基础处理/质量检测	内置算法	过滤时长少于自定义阈值的...	生效		2026-03-13 14:12:48	编辑 测试 更多
视频循环过滤	video_loss_filter	视频/多模态基础处理/质量检测	内置算法	视频循环过滤可深度检测视...	生效		2026-03-13 14:12:48	编辑 测试 更多
视频剪辑	video_clip_generator	视频/多模态基础处理/内容处理	内置算法	按配置的跳过开头与跳过结...	生效		2026-03-13 14:12:48	编辑 测试 更多
视频分辨率调整	video_resolution_ratio_generator	视频/多模态基础处理/内容处理	内置算法	按配置调整视频的分辨率。...	生效		2026-03-13 14:12:48	编辑 测试 更多
视频编解码检测	video_codec_detect	视频/多模态基础处理/质量检测	内置算法	检查视频文件使用的编解码...	生效		2026-03-13 14:12:48	编辑 测试 更多
视频CRC检测	video_crc_detect	视频/多模态基础处理/质量检测	内置算法	使用CRC32循环校验算法。...	生效		2026-03-13 14:12:48	编辑 测试 更多
视频AI增强	video_ai_enhance_generator	视频/媒体增强与智能创作/智能标记与生成	内置算法	通过Real-ESRGAN深度学习...	生效		2026-03-13 14:12:48	编辑 测试 更多

第二步，填写算子基础信息

在跳转的算子新建页面，填写算子英文名、中文名、选择算子标签、填写算子描述、选择算子输入、输出类型、选择算法来源及算法类型；

算子英文名称: [输入框] * 算子中文名称: [输入框] * 算子标签: [选择标签]

算子描述: [输入框]

算子输入类型: 图片 视频 文本 音频 通用

算子输出类型: 图片 视频 文本 音频 通用

算法来源: 内置算法 内置模型服务 第三方Http

算法类型: 分析类 生成类

算子英文名称: 支持英文、数字、下划线，必须以英文开头，最少 4 字符，最多 40 字符，具

备唯一性；

算子中文名称：支持中文、英文、数字、下划线，最少 4 字符、最多 40 字符；

算子标签：下拉选择，表示该算子归属于哪一场景；

算子描述：用于详细描述算子的功能，建议填写 100 字以上，详细描述算子的作用及算子对应的业务价值等信息；

算子输入类型：表示该算子是处理何种类型的数据，本案例选择图片；

算子输出类型：表示该算子输出何种类型的数据，本案例选择图片；

算法来源：表示算法的来源，本案例选择第三方 Http；

算法类型：表示算法是生成类算法还是分析类算法，本案例选择分析类；

第三步，定义前端组件

第三方 http 接入都会基于接口文档，所以在配置前，您会从接口文档中知晓算法所需的输入参数，以及是否需要用户输入额外参数，如本案例中，需要调用的算法输入参数如下：

```
[
  {
    "data_type": "image",
    "id": "",
    "data": "${data base64}"
  }
]
```

从输入参数分析，可知本算法无需用户填写额外参数，故前端组件为空



第四步，定义算法信息

定义算法信息时，需要输入调用的算法的 url 地址；

```
http://1.0.0.127:8080/analysis/predict 【仅示例】
```



配置页面中的选择算法配置项，为系统默认配置项，用户不可更改，用于执行算法的调用逻辑。

第五步，定义算法输入参数

将所调用算法的接口文档提供的输入参数填写至 head 参数或 body 参数填写框中；（需要用户填写额外参数的前端组件配置及输入参数配置见 2.9.3.1.5 前端组件配置部分）

输入参数定义

head参数	参数名	参数值	操作

body参数: 输入框 表单 +添加

```
{
  "data_type": "image",
  "id": "",
  "data": "${data_base64}"
}
```

注意：
定义的前端组件需要在Body中通过引用变量的方式与对应的参数进行绑定。如定义的前端组件参数为id，算法所需请求参数为info_id，那么在body中需配置info_id.\$id。
若算法需传入的参数为文件，可以使用内置变量\${data_base64}，在body参数中将该变量填写到对应的存放位置中（如上面输入框的data参数），系统会将文件转换为Base64形式，并传输给算法。
若算法需传入的参数为文本内容，可以使用内置变量\${data_string}，在body参数中将该变量填写到对应的存放位置中，系统会读取文档的文本内容，并传输给算法。

这里需要注意的是，若算法需传入的参数为文件，可以使用内置变量`${data_base64}`，在 body 参数中将该变量填写到对应的存放位置中（如上面输入框的 data 参数），系统会将文件转换为 Base64 形式，并传输给算法。

第六步，定义算子输出参数

根据所调用算法的接口文档提供的返回参数，配置算子的输出字段，并将算子输出字段与算法输出字段进行映射。如本案例中调用算法的返回参数如下：

```
{
  "status": 0,
  "result": [
    {
      "is_violence": 1,
      "desc": "violence in office"
    }
  ],
  "message": "success"
}
```

此时，在配置页面的输出参数部分，可以定义算子输出字段与算法输出字段的映射关系，按照如下方式进行配置：

某分析类算法可以识别图片中的物体并返回中英文结果，某次执行的返回结果如下：

```
{
  "status":0,
  "result":[
    {
      "tag":"image_recognition",
      "en_tags":[
        "car",
        "cat"
      ],
      "cn_tags":[
        "汽车",
        "猫"
      ]
    }
  ],
  "message":"success"
}
```

我们所需的结果为中文名和英文名，则在输出参数中配置为：

算子输出字段	算法输出字段
en_tags	<code>\$.result[*].en_tags</code>
cn_tags	<code>\$.result[*].cn_tags</code>

此时算子输出字段 en_tags 取值为：`["car", "cat"]`；算子输出字段 cn_tags 取值为：`["汽车"、"猫"]`

所以最终该算子的输出参数配置如下：

返回取值映射	字段类型	算子输出字段	算法输出字段	操作
	String	is_violence	<code>\$.result[*].is_violence</code>	删除
	String	desc	<code>\$.result[*].desc</code>	删除

[配置案例说明](#) +添加

字段类型	算子输出字段	算法输出字段
string	is_violence	<code>\$.result[*].is_violence</code>
string	desc	<code>\$.result[*].desc</code>

第七步，设置算子基础配置信息

最后设置该算子是否保存生成的文件或是生成的分析结果。一般来说，生成类算法会生成新的文件，此时建议勾选展示生成文件配置项，用于配置生成文件的存储路径；分析类算法会生成可解析的半结构化分析结果，此时建议勾选展示分析结果配置项，用于配置分析结果的存储表及映射关系。



第八步，保存算子

点击右下角【确定】按钮，即可保存算子，此时算子状态为【编辑中】。可在算子管理列表进行算子上线操作，上线后，算子会在算子广场及任务画布中展示。

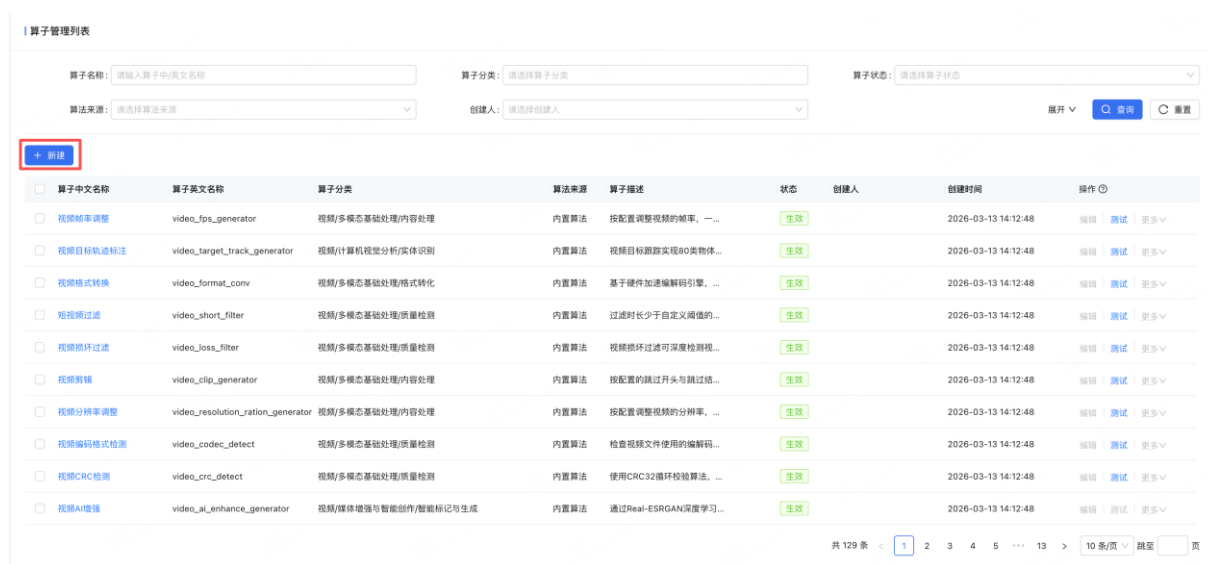
自此本案例结束。

2.5.4. 新增引用第三方 Http 算法的生成类算子

本案例通过调用第三方算法的 Http 接口进行图片清晰化算子的新增

第一步，新建算子

在算子管理列表，点击【新建】按钮，进入算子新增页面；



第二步，填写算子基础信息

在跳转的算子新建页面，填写算子英文名、中文名、选择算子标签、填写算子描述、选择算子输入、输出类型、选择算法来源及算法类型；

算子英文名称：支持英文、数字、下划线，必须以英文开头，最少 4 字符，最多 40 字符，具备唯一性；

算子中文名称：支持中文、英文、数字、下划线，最少 4 字符、最多 40 字符；

算子标签：下拉选择，表示该算子归属于哪一场景；

算子描述：用于详细描述算子的功能，建议填写 100 字以上，详细描述算子的作用及算子对应的业务价值等信息；

算子输入类型：表示该算子是处理何种类型的数据，本案例选择图片；

算子输出类型：表示该算子输出何种类型的数据，本案例选择图片；

算法来源：表示算法的来源，本案例选择第三方 Http；

算法类型：表示算法是生成类算法还是分析类算法，本案例选择生成类；

第三步，定义前端组件

第三方 http 接入都会基于接口文档，所以在配置前，您会从接口文档中知晓算法所需的输入参数，以及是否需要用户输入额外参数，如本案例中，需要调用的算法输入参数如下：

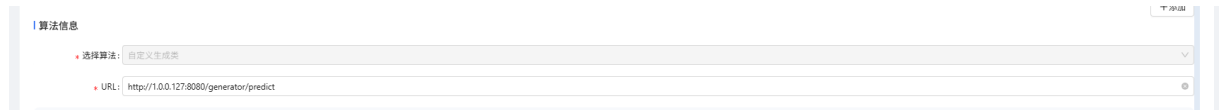
```
[
  {
    "data_type": "image",
    "id": "",
    "data": "${data base64}"
  }
]
```

从输入参数分析，可知本算法无需用户填写额外参数，故前端组件为空

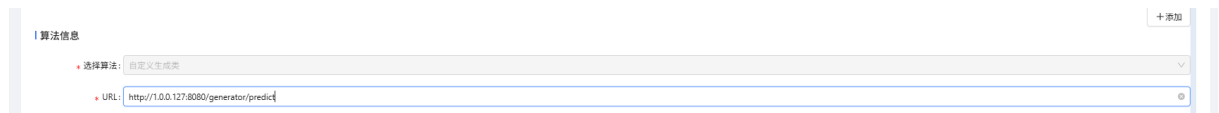
第四步，定义算法信息

定义算法信息时，需要输入调用的算法的 url 地址；

```
http://1.0.0.127:8080/generator/predict 【仅示例】
```



配置页面中的选择算法配置项，为系统默认配置项，用户不可更改，用于执行算法的调用逻辑。



第五步，定义算法输入参数

将所调用算法的接口文档提供的输入参数填写至 head 参数或 body 参数填写框中；（需要用户填写额外参数的前端组件配置及输入参数配置见 2.9.3.1.5 前端组件配置部分）



这里需要注意的是，若算法需传入的参数为文件，可以使用内置变量`${data_base64}`，在 body 参数中将该变量填写到对应的存放位置中（如上面输入框的 data 参数），系统会将文件转换为 Base64 形式，并传输给算法。

第六步，定义算子输出参数

与分析类算法不同的时，生成类算法返回的是 base64 编码后的文件，所以，在算子输出参数配置处有一些不同。如本次案例中的算法返回如下：

```
{
  "status": 0,
  "base64": "/9j/4AAQSkZJRg.....",
  "message": "success"
}
```

算子输出参数按如下逻辑配置：

某生成类算法返回为 base64 编码，某次执行的返回结果如下：

```
{
  "status":0,
  "result":[
    {
      "data":"base64ksdjsldkhfs.....",
    }
  ],
  "message":"success"
}
```

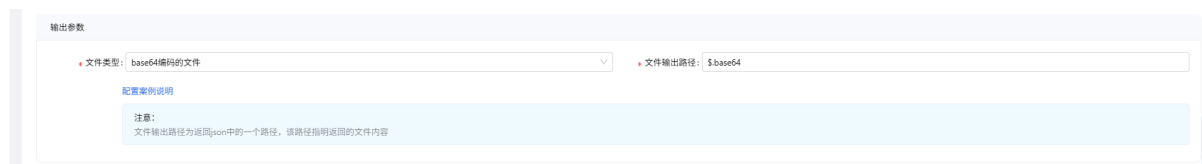
我们需要取到该文件，则在文件输出路径输入`$.result[*].data`，系统会将 base64 编码转换为文件进行输出；

同样的，某生成类算法返回为 text，某次执行的返回结果如下：

```
{
  "status":0,
  "result":[
    {
      "translate":"你好！世界！",
    }
  ],
  "message":"success"
}
```

我们需要取到这段文本，则在文件输出路径输入`$.result[*].translate`，系统会将取到的文本存储为 txt 文件进行输出。

此时需要取到算法返回的文件，则在输出参数中进行如下配置：



文件类型	文件输出路径
base64 编码的文件	\$. base64

第七步，设置算子基础配置信息

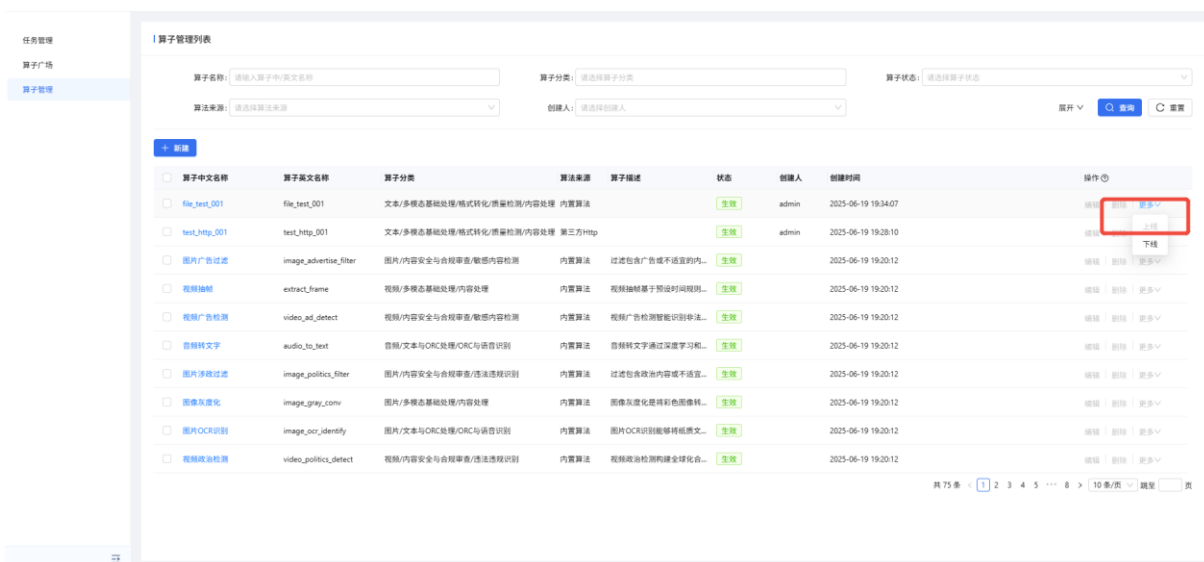
最后设置该算子是否保存生成的文件或是生成的分析结果。一般来说，生成类算法会生成新的

文件，此时建议勾选展示生成文件配置项，用于配置生成文件的存储路径；分析类算法会生成可解析的半结构化分析结果，此时建议勾选展示分析结果配置项，用于配置分析结果的存储表及映射关系。



第八步，保存算子

点击右下角【确定】按钮，即可保存算子，此时算子状态为【编辑中】。可在算子管理列表进行算子上线操作，上线后，算子会在算子广场及任务画布中展示。



自此本案例结束。

2.5.5. 算子前端组件配置

上面的案例中，所选算法都是无需用户配置额外参数的算法，本案例专门进行需要用户输入参数的配置指导；

我们假设某 http 算法的输入参数示例如下：

```

[
  {
    "data_type": "image",
    "id": "",
    "data": "${data base64}",
    "length": "20px",
    "width": "10px"
  }
]

```

此时，算法需要用户填写的参数为长度 length 和宽度 width，于是在前端组件定义处，定义两个前端组件如下：

控件类型	参数名	参数类型	配置项名称	可用枚举	是否必填	默认值	配置说明	操作
输入框	user_length	INTEGER	长度		<input checked="" type="checkbox"/> 是			删除
输入框	user_width	INTEGER	宽度		<input checked="" type="checkbox"/> 是			删除

+ 添加

- 控件类型选择输入框，表示该算子需要用户填写内容；其备选项还包括单选下拉框、多选下拉框、复选框、文本域等；
- 参数名：唯一标识控件的英文名，也是用于在输入参数中配置的变量名；
- 参数类型：表示该组件所填写的数据类型，针对输入框分为 string 和 integer 两种；
- 配置项名称：表示该组件在前端展示的中文配置项名称；
- 可用枚举：当控件类型选择下拉框时，可在此处配置下拉选项；
- 是否必填：标识该组件是否为必填项；
- 默认值：若填写了默认值，则会默认填充默认值；
- 配置说明：针对该组件填写的一些说明，若填写则会在配置项名称后显示”？“，鼠标 hover 显示说明内容；

前端组件定义完成后，我们需要将前端组件的参数名以变量的方式绑定到算法是输入参数中，如上输入参数中，在配置页面的 body 参数应如下配置：

```
[
  {
    "data_type": "image",
    "id": "",
    "data": "${data base64}",
    "length": "${user_length}",
    "width": "${user_width}"
  }
]
```

body参数:

输入框 表单

十进制

```
{
  "data_type": "image",
  "id": "",
  "data": "${data base64}",
  "length": "${user_length}",
  "width": "${user_width}"
}
```

注意:

定义的前端组件需要在Body中通过引用变量的方式与对应的参数进行绑定。如定义的前端组件参数为id，算法所需请求参数为info_id，那么在body中需配置=fo_id\$id。
若算法传入的参数为文件，可以使用内置变量\${data_base64}，在body参数中将变量填写到对应的存放位置中（如上面输入框的data参数），系统会将文件转换为Base64形式，并传输给算法。
若算法传入的参数为文本内容，可以使用内置变量\${data_string}，在body参数中将变量填写到对应的存放位置中，系统会读取文档的文本内容，并传输给算法。