

ModelArts

数据标注

文档版本 01

发布日期 2023-12-09



版权所有 © 华为云计算技术有限公司 2023。保留一切权利。

未经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为云计算技术有限公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为云计算技术有限公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

目 录

1 数据标注简介.....	1
2 人工标注.....	3
2.1 创建标注作业.....	3
2.2 图片标注.....	9
2.2.1 图像分类.....	9
2.2.2 物体检测.....	17
2.2.3 图像分割.....	28
2.3 文本标注.....	35
2.3.1 文本分类.....	35
2.3.2 命名实体.....	40
2.3.3 文本三元组.....	45
2.4 音频标注.....	50
2.4.1 声音分类.....	50
2.4.2 语音内容.....	54
2.4.3 语音分割.....	57
2.5 视频标注.....	60
2.6 查看标注作业.....	64
2.6.1 查看创建的作业.....	64
2.6.2 查看参与标注的作业.....	65
3 智能标注.....	66
3.1 创建智能标注作业.....	66
3.2 难例确认.....	70
4 团队标注.....	73
4.1 团队标注简介.....	73
4.2 创建和管理团队.....	73
4.2.1 管理团队.....	73
4.2.2 管理成员.....	75
4.3 创建团队标注任务.....	76
4.4 登录 ModelArts-Console.....	79
4.5 启动团队标注任务.....	81
4.6 审核团队标注任务结果.....	82

4.7 验收团队标注任务结果.....	83
---------------------	----

1

数据标注简介

说明

数据管理模块即将下线，对未使用过数据管理的用户不可见。

模型训练过程中需要大量已标注的数据，因此在模型训练之前需要进行数据标注作业。ModelArts为用户提供了标注数据的能力：

- **人工标注**：用户创建单人标注作业，对数据进行手工标注。
- **智能标注**：在标注一定量的数据情况下，用户可以通过启动智能标注任务对数据进行自动标注，提高标注的效率。
- **团队标注**：对于大批量的数据，用户可以通过创建团队标注作业，进行多人协同标注。

人工标注

对于不同类型的数据，用户可以选择不同的标注类型。当前ModelArts支持如下类型的标注作业：

- 图片
 - 图像分类：识别一张图片中是否包含某种物体。
 - 物体检测：识别出图片中每个物体的位置及类别。
 - 图像分割：根据图片中的物体划分出不同区域。
- 音频
 - 声音分类：对声音进行分类。
 - 语音内容：对语音内容进行标注。
 - 语音分割：对语音进行分段标注。
- 文本
 - 文本分类：对文本的内容按照标签进行分类处理。
 - 命名实体：针对文本中的实体片段进行标注，如“时间”、“地点”等。
 - 文本三元组：针对文本中的实体片段和实体之间的关系进行标注。
- 视频
视频标注：识别出视频中每个物体的位置及分类。目前仅支持mp4格式。

智能标注

除了人工标注外，ModelArts还提供了智能标注功能，快速完成数据标注，为您节省70%以上的标注时间。智能标注是指基于当前标注阶段的标签及图片学习训练，选中系统中已有的模型进行智能标注，快速完成剩余图片的标注操作。

目前只有“图像分类”和“物体检测”类型的数据集支持智能标注功能。

团队标注

数据标注任务中，一般由一个人完成，但是针对数据集较大时，需要多人协助完成。ModelArts提供了团队标注功能，可以由多人组成一个标注团队，针对同一个数据集进行标注管理。

团队标注功能当前仅支持“图像分类”、“物体检测”、“文本分类”、“命名实体”、“文本三元组”、“语音分割”类型的数据集。

不同类型数据集支持的功能列表

其中，不同类型的 data 集，支持不同的功能，详细信息请参见[表1-1](#)。

表 1-1 不同类型数据集支持的功能

数据集类型	标注类型	人工标注	智能标注	团队标注
图片	图像分类	支持	支持	支持
	物体检测	支持	支持	支持
	图像分割	支持	-	-
音频	声音分类	支持	-	-
	语音内容	支持	-	-
	语音分割	支持	-	支持
文本	文本分类	支持	-	支持
	命名实体	支持	-	支持
	文本三元组	支持	-	支持
视频	视频标注	支持	-	-
自由格式	-	-	-	-
表格	-	-	-	-

2 人工标注

2.1 创建标注作业

由于模型训练过程需要大量有标签的数据，因此在模型训练之前需对没有标签的数据添加标签。您可以通过创建单人标注作业或团队标注作业对数据进行手工标注，或对任务启动智能标注添加标签，快速完成对图片的标注操作，也可以对已标注图片修改或删除标签进行重新标注。

标注作业支持的数据类型

对于不同类型的数据集，用户可以选择不同的标注任务，当前ModelArts支持如下类型的标注任务。

- 图片
 - 图像分类：识别一张图片中是否包含某种物体。
 - 物体检测：识别出图片中每个物体的位置及类别。
 - 图像分割：根据图片中的物体划分出不同区域。
- 音频
 - 声音分类：对声音进行分类。
 - 语音内容：对语音内容进行标注。
 - 语音分割：对语音进行分段标注。
- 文本
 - 文本分类：对文本的内容按照标签进行分类处理。
 - 命名实体：针对文本中的实体片段进行标注，如“时间”、“地点”等。
 - 文本三元组：针对文本中的实体片段和实体之间的关系进行标注。
- 视频
 - 视频标注：识别出视频中每个物体的位置及分类。目前仅支持mp4格式。

前提条件

在进行数据标注前，需要创建相应类型的数据集。

说明

数据管理模块即将下线，对未使用过数据管理的用户不可见。

操作步骤

1. 登录ModelArts管理控制台，在左侧菜单栏中选择“数据管理> 数据标注”，进入“数据标注”管理页面。
2. 在数据标注管理页面，单击页面右上角“创建标注作业”，进入“创建标注作业”页面，根据需求创建不同类型的标注作业。
 - a. 填写标注作业基本信息，标注作业的“名称”和“描述”。

图 2-1 标注作业基本信息

The screenshot shows a form for creating an annotation job. It includes fields for 'Name' (dataset-name) and 'Description'. The 'Description' field has a character limit of 0/256.

- b. 根据您的需求，选择“标注场景”和“标注类型”。

图 2-2 选择标注场景和标注类型



- c. 针对不同类型的标注作业，需填写参数不同，请参考如下类型标注作业对应的参数介绍。

- **图片 (图像分类、物体检测、图像分割)**
- **音频 (声音分类、语音内容、语音分割)**
- **文本 (文本分类、命名实体、文本三元组)**
- **视频**

- d. 参数填写无误后，单击页面右下角“创建”。

标注作业创建完成后，系统自动跳转至数据标注管理页面，针对创建好的标注作业，您可以执行智能标注、发布、修改和删除等操作。

图片（图像分类、物体检测、图像分割）

图 2-3 图像分类和物体检测类型的参数

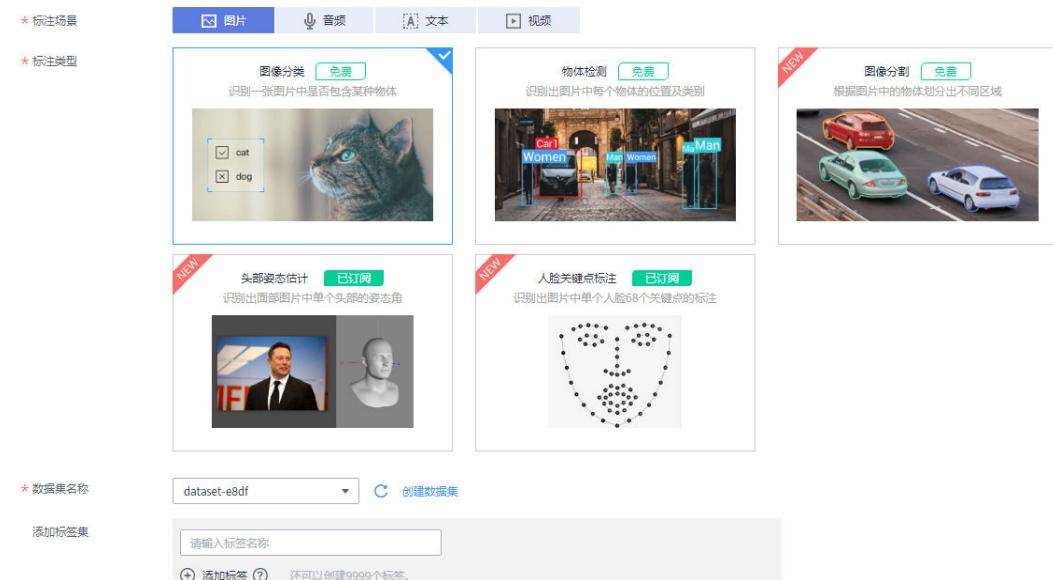


表 2-1 图片类型标注作业的详细参数

参数名称	说明
数据集名称	选择支持当前标注类型的数据集。
添加标签集	<ul style="list-style-type: none">设置标签名称：在标签名称文本框中，输入标签名称。长度为1~1024字符。添加标签：单击“添加标签”可增加多个标签。设置标签颜色：“物体检测”和“图像分割”类型标注作业需设置此参数。在每个标签右侧的标签颜色区域下，可在色板中选择颜色，或者直接输入十六进制颜色码进行设置。设置标签属性：针对“物体检测”类型标注作业，在设置完标签颜色后，可在右侧单击加号，增加对应的标签属性。标签属性用于区分同一标签物体的不同属性。例如，黄色小猫、黑色小猫。标签为cat，颜色为不同的标签属性。
启用团队标注	<p>选择是否启用团队标注。图像分割暂不支持团队标注，当选择图像分割类型时，界面不显示此参数。</p> <p>启用团队标注功能，需填写对应的团队标注任务“类型”，同时选择对应的“标注团队”及参与标注的“团队成员”。参数详细介绍请参见创建团队标注任务。</p> <p>在启用“团队标注”前，需确保您已经在“标注团队”管理页面，添加相应的团队以及成员。如果没有标注团队，可直接从界面链接跳转至“标注团队”页面，添加您的团队并为其添加成员。详细指导请参见添加团队。</p> <p>启用团队标注功能的数据集，在创建完成后，可以在“标注类型”中看到“团队标注”的标识。</p>

音频（声音分类、语音内容、语音分割）

图 2-4 声音分类、语音内容、语音分割类型的参数

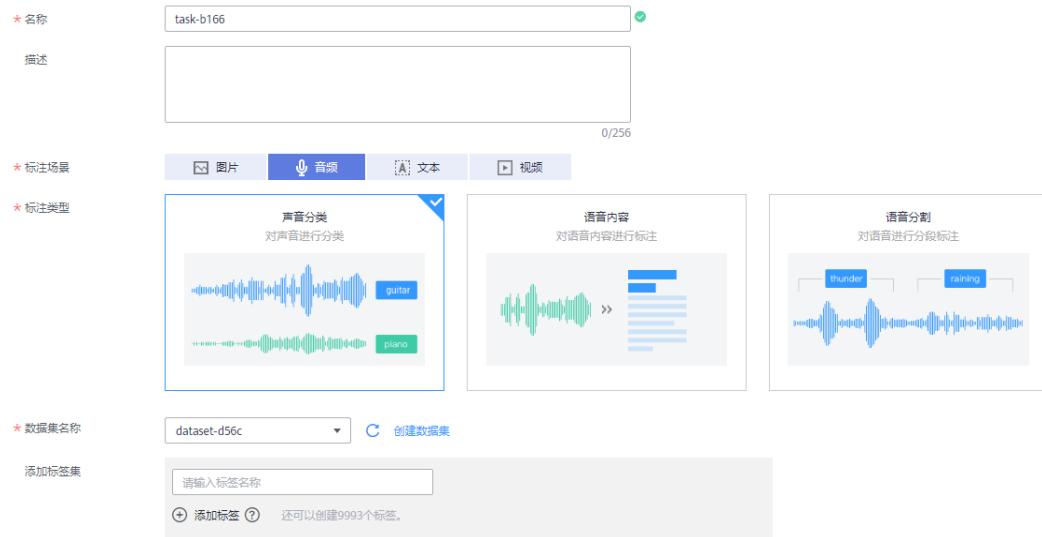


表 2-2 音频类型标注作业的详细参数

参数名称	说明
数据集名称	选择支持当前标注类型的数据集。
添加标签集（声音分类）	<p>“声音分类”类型的标注作业可以添加标签集。</p> <ul style="list-style-type: none">设置标签名称：在标签名称文本框中，长度为1~1024字符。添加标签：单击“添加标签”可增加多个标签。
标签管理（语音分割）	<p>“语音分割”类型的标注作业，支持标签管理。</p> <ul style="list-style-type: none">单标签 单标签适用于一段音频标注只有一种类别的音频，通常标注一个标签。<ul style="list-style-type: none">设置标签名称：在“标签名”列输入标签名称。长度为1~1024字符。设置标签颜色：在“标签颜色”列设置标签颜色。可在色板中选择颜色，或者直接输入十六进制颜色码进行设置。多标签 多标签适用于多维度标注，例如在一段音频标注噪音与人说话的声音两种类别，其中说话的声音还可以标注为不同人的声音。单击“新建标签类别”可添加多个标签类别，一个标签类别可以包含多个标签。“标签类别”和“标签名”只能是中文、字母、数字、英文句号、下划线或中划线组成的合法字符串。长度为1~256字符。<ul style="list-style-type: none">设置标签类别：在“标签类别”输入标签类别的名称。设置标签名称：在“标签名”输入标签名称。添加标签：单击“添加标签”可增加多个标签。

参数名称	说明
启用语音内容标注（语音分割）	仅“语音分割”类型数据集支持设置，默认关闭。如果启用此功能，支持针对语音内容进行标注。
启用团队标注（语音分割）	仅“语音分割”类型支持团队标注，因此选择创建语音分割类型时，支持设置是否启用团队标注。 启用团队标注功能，需填写对应的团队标注任务“类型”，同时选择对应的“标注团队”及参与标注的“团队成员”。参数详细介绍请参见 创建团队标注任务 。 在启用“团队标注”前，需确保您已经在“标注团队”管理页面，添加相应的团队以及成员。如果没有标注团队，可直接从界面链接跳转至“标注团队”页面，添加您的团队并为其添加成员。详细指导请参见 添加团队 。 启用团队标注功能的数据集，在创建完成后，可以在“标注类型”中看到“团队标注”的标识。

文本（文本分类、命名实体、文本三元组）

图 2-5 文本分类、命名实体、文本三元组类型的参数

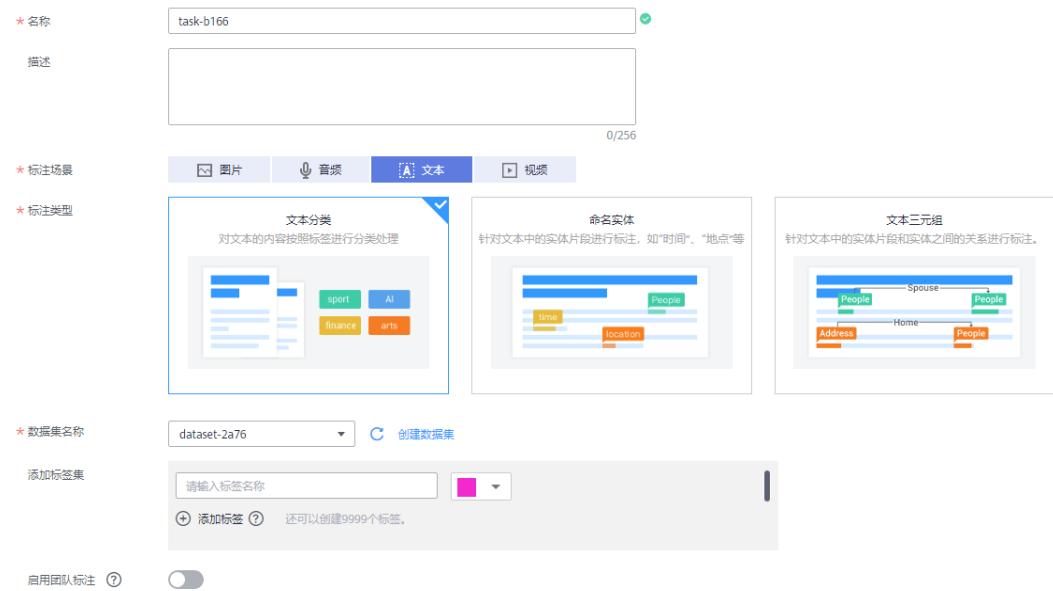
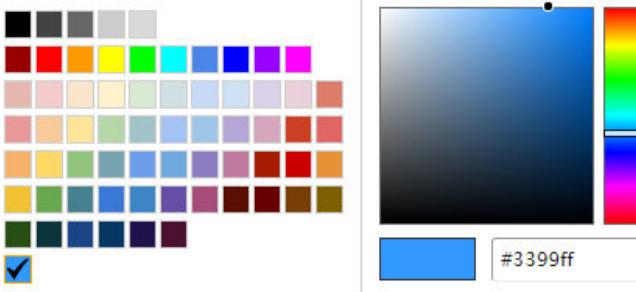


表 2-3 文本类型标注作业的详细参数

参数名称	说明
数据集名称	选择支持当前标注类型的数据集。

参数名称	说明
添加标签集（文本分类、命名实体）	<ul style="list-style-type: none">设置标签名称：在标签名称文本框中，输入标签名称。长度为1~1024字符。添加标签：单击“添加标签”可增加多个标签。设置标签颜色：在每个标签右侧的标签颜色区域下，可在色板中选择颜色，或者直接输入十六进制颜色码进行设置。 
添加标签集（文本三元组）	<ul style="list-style-type: none">实体标签：需设置标签名以及标签颜色。可在颜色区域右侧单击加号增加多个标签。关系标签：关系标签为两个实体之间的关系。需设置起始实体和终止实体，您需要先添加至少2个实体标签后，再添加关系标签。 
启用团队标注	<p>选择是否启用团队标注。</p> <p>启用团队标注功能，需填写对应的团队标注任务“类型”，同时选择对应的“标注团队”及参与标注的“团队成员”。参数详细介绍请参见创建团队标注任务。</p> <p>在启用“团队标注”前，需确保您已经在“标注团队”管理页面，添加相应的团队以及成员。如果没有标注团队，可直接从界面链接跳转至“标注团队”页面，添加您的团队并为其添加成员。详细指导请参见添加团队。</p> <p>启用团队标注功能的数据集，在创建完成后，可以在“标注类型”中看到“团队标注”的标识。</p>

视频

图 2-6 视频类型的参数

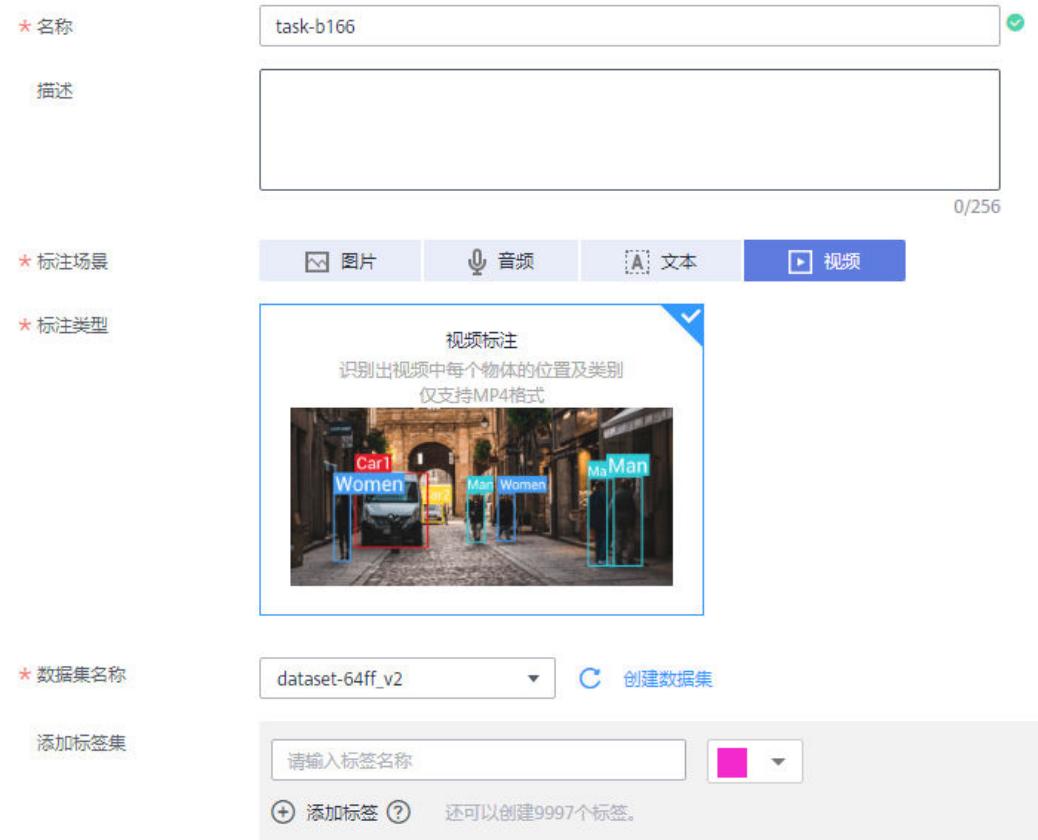


表 2-4 视频类型标注作业的详细参数

参数名称	说明
数据集名称	选择支持当前标注类型的数据集。
添加标签集	<ul style="list-style-type: none">设置标签名称: 在标签名称文本框中，输入标签名称。长度为1~1024字符。添加标签: 单击“添加标签”可增加多个标签。设置标签颜色: 在每个标签右侧的标签颜色区域下，可在色板中选择颜色，或者直接输入十六进制颜色码进行设置。

2.2 图片标注

2.2.1 图像分类

由于模型训练过程需要大量有标签的图片数据，因此在模型训练之前需对没有标签的图片添加标签。您可以通过手工标注或智能一键标注的方式添加标签，快速完成对图片的标注操作，也可以对已标注图片修改或删除标签进行重新标注。

针对图像分类场景，开始标注前，您需要了解：

- 图片标注支持多标签，即一张图片可添加多个标签。
- 标签名是由中文、大小写字母、数字、中划线或下划线组成，且不超过1024位的字符串。

开始标注

1. 登录ModelArts管理控制台，在左侧菜单栏中选择“数据管理> 数据标注”，进入“数据标注”管理页面。
2. 在标注作业列表右侧“所有类型”页签下拉选择标注类型。基于“标注类型”选择需要进行标注的标注作业，单击标注作业名称进入标注作业标注详情页。

图 2-7 下拉选择标注类型



3. 在标注作业标注详情中，展示此标注作业下全部数据。

同步新数据

ModelArts会自动将数据集中新增的数据同步至标注作业，包含数据及当前标注作业支持的标注信息。

为了快速获取数据集中最新数据，可在标注作业详情页的“全部”、“未标注”或“已标注”页签中，单击“同步新数据”，快速将数据集中的数据添加到标注作业中。

说明

问题现象：

将已标注好的数据上传至OBS，同步数据后，显示为未标注。

原因分析：

可能是OBS桶设置了自动加密导致此问题。

解决方法：

需要新建OBS桶重新上传数据，或者取消桶加密后，重新上传数据。

筛选数据

在标注作业详情页面，默认展示作业中全部未标注的数据，您可以在“全部”、“未标注”或“已标注”页签下，在筛选条件区域，单击¹，添加筛选条件，快速过滤出您想要查看的数据。

支持的筛选条件如下所示，您可以设置一个或多个选项进行筛选。

- 难例集：难例或非难例。
- 标签：您可以选择全部标签，或者基于您指定的标签，选中其中一个或多个。
- 文件名或目录：根据文件名称或者文件存储目录筛选。
- 标注人：选择执行标注操作的帐号名称。
- 样本属性：表示自动分组生成的属性。只有启用了自动分组任务后才可使用此筛选条件。
- 数据属性：筛选数据的来源，选择“全部”或“推理”。

图 2-8 筛选条件



标注图片（手工标注）

在标注作业详情页中，展示了此数据集中“全部”、“未标注”和“已标注”的图片，默认显示“未标注”的图片列表。单击图片，即可进行图片的预览，对于已标注图片，预览页面下方会显示该图片的标签信息。

1. 在“未标注”页签，勾选需进行标注的图片。
 - 手工点选：在图片列表中，单击勾选图片左上角的选择框，进入选择模式，表示图片已勾选。可勾选同类别的多个图片，一起添加标签。
 - 批量选中：如果图片列表的当前页，所有图片属于一种类型，可以在图片列表的右上角单击“选择当前页”，则当前页面所有的图片将选中。
2. 为选中图片添加标签。
 - a. 在右侧的“添加标签”区域中，单击“标签名”右侧的文本框中设置标签。单击“标签名”右侧的文本框，然后从下拉列表中选择已有的标签。如果有标签无法满足要求时，直接在文本框中添加新标签。
 - b. 单击“确定”。此时，选中的图片将被自动移动至“已标注”页签，且在“未标注”和“全部”页签中，标签的信息也将随着标注步骤进行更新，如增加的标签名称、各标签对应的图片数量。

图 2-9 添加标签



说明

若您还不太清楚如何进行标注，可参考数据集详情页面的“标注样例说明”完成标注。

1. 登录ModelArts管理控制台，选择数据管理>数据标注进入数据标注页。
2. 在“我创建的”或“我参与的”页签下，找到您需要标注的数据集。
3. 单击数据集名称，进入标注详情页。（默认直接进入“未标注”页签）。
4. 在标注详情页右上角单击“标注样例说明”。

图 2-10 标注样例说明

② 标注样例说明

查看已标注图片

在标注任务详情页，单击“已标注”页签，您可以查看已完成标注的图片列表。图片缩略图下方默认呈现其对应的标签，您也可以勾选图片，在右侧的“选中文件标签”中了解当前图片的标签信息。

修改标注

当数据完成标注后，您还可以进入已标注页签，对已标注的数据进行修改。

- **基于图片修改**

在标注作业详情页面，单击“已标注”页签，然后在图片列表中选中待修改的图片（选择一个或多个）。在右侧标签信息区域中对图片信息进行修改。

修改标签：在“选中文件标签”区域中，单击操作列的编辑图标，然后在文本框中输入正确的标签名，然后单击确定图标完成修改。

删除标签：在“选中文件标签”区域中，单击操作列的删除图标删除该标签。此操作仅删除选中图片中的标签。

图 2-11 编辑标签

选中文件标签		
标签名称	标签数量	操作
dog_image	1	
image100	1	

- **基于标签修改**

- 在标注作业详情页，单击右侧区域的“标签管理”，显示全部标签列表。

- 修改标签：单击操作列的“修改”，然后在弹出的对话框中输入修改后的标签名，然后单击“确定”完成修改。修改后，之前添加了此标签的图片，都将被标注为新的标签名称。
- 删除标签：单击操作列“删除”，之前添加了此标签的图片，都将删除此标签。

图 2-12 标签管理

全部标签 12		标签管理 C
标签名称	标签数量	操作
animal	168	
cake	50	
cat_image	19	
data	293	

图 2-13 全部标签的信息

操作列	属性	操作
<input type="checkbox"/> 标签名称	--	修改 删除
<input type="checkbox"/> animal	--	修改 删除
<input type="checkbox"/> cake	--	修改 删除
<input type="checkbox"/> cat_image	--	修改 删除
<input type="checkbox"/> data	--	修改 删除
<input type="checkbox"/> Data	--	修改 删除
<input type="checkbox"/> dd	--	修改 删除
<input type="checkbox"/> dog_image	--	修改 删除

- 单击标注作业操作列的“标签”，可跳转至标签管理页。
 - 单击操作列的“修改”，即可完成标签的修改。
 - 单击操作列的“删除”，即可删除该标签。

添加数据

除了同步数据集中的新数据外，您还可以在标注作业中，直接添加图片，用于数据标注。添加的数据将先导入至标注任务关联的数据集中，然后标注任务会自动同步数据集中最新的数据。

1. 在标注作业详情页面，单击“全部”、“已标注”或“未标注”页签，然后单击左上角“添加数据”，选择添加数据。

图 2-14 添加数据



2. 在弹出的导入对话框中，选择数据来源和导入方式，选择导入的数据路径和数据标注状态。

图 2-15 添加图片



3. 在导入对话框中，单击“确定”，完成添加数据的操作。

您添加的图片将自动呈现在“全部”的图片列表中，也可单击“添加数据>查看历史记录”，进入“任务历史”界面，可查看相应的导入历史。

图 2-16 查看历史数据

任务历史						
创建时间	导入方式	导入路径	样本总数	导入样本总数	导入已标注样本数	导入状态
2021/11/08 09:40:4...	对象存储服务 (OBS...)	obs://tupiantest1/s...	16	16	0	成功

删除图片

通过数据删除操作，可将需要丢弃的图片数据快速删除。

在“全部”、“未标注”或“已标注”页面中，依次选中需要删除的图片，或者选择“选择当前页”选中该页面所有图片，然后单击“删除”。在弹出的对话框中，根据实际情况选择是否勾选“同时删除OBS源文件”，确认信息无误后，单击“确定”完成图片删除操作。

图 2-17 删除图片



其中，被选中的图片，其左上角将显示为勾选状态。如果当前页面无选中图片时，“删除”按钮为灰色，无法执行删除操作。

说明

如果勾选了“同时删除OBS源文件”，删除图片操作将删除对应OBS目录下存储的图片，此操作可能会影响已使用此源文件的其他数据集或数据集版本，有可能导致展示异常或训练/推理异常。删除后，数据将无法恢复，请谨慎操作。

标注人员管理

若您创建的标注作业，开启了团队标注，“标注人员管理”页面中可查看团队标注作业的标注详情。添加、修改或删除标注成员。

1. 登录“数据管理>数据标注”，在“我创建的”页签下可查看所有的标注作业列表。
2. 在作业列表的“名称”列，根据标注作业名称找到对应的团队标注作业。（团队标注作业的名称后带有🔍标识。）
3. 单击作业操作列的“更多>标注人员管理”。或单击作业名称进入作业详情，继续单击右上角“团队标注>标注人员管理”，进入成员管理页面。

图 2-18 进入标注人员管理页（1）

名称	状态	进度	时间	操作
task-aa55 🔍	dataset-c42d	33% (20/60)	0 2021/05/13 18:33:15 GMT+08:00	智能标注 标签 发布 删...
dataset-7416 🔍	dataset-7416	0% (0/20)	.. 2021/05/13 16:47:48 GMT+08:00	标签 发布 任务统计
dataset-c42d 🔍	dataset-c42d	33% (20/60)	0 2021/05/13 16:23:04 GMT+08:00	智能标注 标签 发布 更多

图 2-19 进入标注人员管理页（2）



- 添加成员：
单击页面“添加成员”，选择成员名称，单击确定。
在操作列，选择“发送邮件”，可将该标注任务以邮件的方式发送至该标注成员。
- 修改成员信息：
单击操作列的“修改”，可修改该成员的角色。
- 删除标注成员：
单击操作列的“删除”可删除该标注成员的所有信息。

2.2.2 物体检测

由于模型训练过程需要大量有标签的图片数据，因此在模型训练之前需对没有标签的图片添加标签。您可以通过手工标注或智能一键标注的方式添加标签，快速完成对图片的标注操作，也可以对已标注图片修改或删除标签进行重新标注。

针对物体检测场景，开始标注前，您需要了解：

- 图片中所有目标物体都要标注。
- 目标物体清晰无遮挡的，必须画框。
- 画框仅包含整个物体。框内包含整个物体的全部，画框边缘不可与待标注的物体的边缘轮廓相交，在此基础之上确保边缘和待标注物体间不要留着空隙，避免背景对模型训练造成干扰。

开始标注

1. 登录ModelArts管理控制台，在左侧菜单栏中选择“数据管理> 数据标注”，进入“数据标注”管理页面。
2. 在标注作业列表右侧“所有类型”页签下拉选择标注类型，基于“标注类型”选择需要进行标注的标注作业，单击标注作业名称进入标注作业标注详情页。

图 2-20 下拉选择标注类型



3. 在标注作业标注详情中，展示此标注作业下全部数据。

同步新数据

ModelArts会自动将数据集中新增的数据同步至标注作业，包含数据及当前标注作业支持的标注信息。

为了快速获取数据集中最新数据，可在标注作业详情页的“全部”、“未标注”或“已标注”页签中，单击“同步新数据”，快速将数据集中的数据添加到标注作业中。

说明书

问题现象：

将已标注好的数据上传至OBS，同步数据后，显示为未标注。

原因分析：

可能是OBS桶设置了自动加密导致此问题。

解决方法：

需要新建OBS桶重新上传数据，或者取消桶加密后，重新上传数据。

筛选数据

在标注作业详情页面，默认展示作业中“未标注”的数据，您可以在“全部”、“未标注”或页签下，在筛选条件区域，单击“”，添加筛选条件，快速过滤出您想要查看的数据。

支持的筛选条件如下所示，您可以设置一个或多个选项进行筛选。

- 难例集：难例或非难例。
- 标签：您可以选择全部标签，或者基于您指定的标签，选中其中一个或多个。
- 文件名或目录：根据文件名称或者文件存储目录筛选。
- 标注人：选择执行标注操作的帐号名称。
- 样本属性：表示自动分组生成的属性。只有启用了自动分组任务后才可使用此筛选条件。
- 数据属性：筛选数据的来源，选择“全部”或“推理”。

图 2-21 筛选条件



标注图片（手工标注）

标注作业详情页中，展示了此数据集中“全部”“未标注”和“已标注”的图片，默认显示“未标注”的图片列表。

1. 在“未标注”页签图片列表中，单击图片，自动跳转到标注页面。在标注页面，常用按钮的使用可参见表2-6。
2. 在页面上方工具栏选择合适的标注图形，系统默认的标注图形为矩形。本示例使用矩形工具进行标注。

说明

页面左侧可以选择多种形状对图片进行标注。标注第一张图片时，一旦选择其中一种，其他图片默认使用此形状进行标注，用户可以根据自己需求再进行切换。

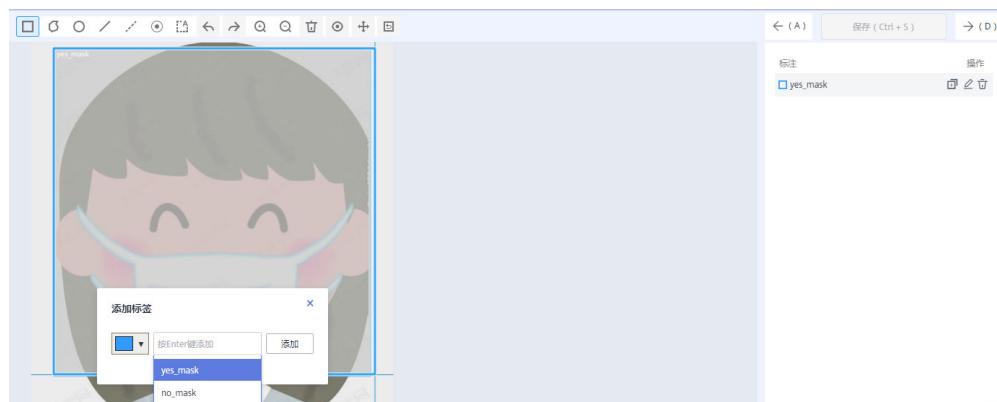
表 2-5 支持的标注框

图标	使用说明
	矩形。也可使用快捷键【1】。鼠标单击标注对象左上角边缘位置，界面将出现矩形框，移动鼠标使得矩形框覆盖标注对象，然后单击完成标注。
	多边形。也可使用快捷键【2】。在标注对象所在范围内，鼠标左键单击完成一个点的标注，沿着物体的形状边缘，通过鼠标指定多个点，最终单击到第一个点的位置，由所有的点组成一个多边形形状。使得需标注的对象在此标注框内。
	圆形。也可使用快捷键【3】。在标注对象中，选择物体的中心点位置，单击鼠标确定圆心，然后移动鼠标，使得圆形框覆盖标注对象，然后再单击鼠标完成标注。
	直线。也可使用快捷键【4】。在标注对象中，选择物体的起始点，单击鼠标确定直线的起始点，然后使得直线覆盖标注对象，然后再单击鼠标完成标注。
	虚线。也可使用快捷键【5】。在标注对象中，选择物体的起始点，单击鼠标确定虚线的起始点，然后使得虚线覆盖标注对象，然后再单击鼠标完成标注。
	点。也可使用快捷键【6】。单击图片中的物体所在位置，即可完成点的标注。

3. 在弹出的添加标签文本框中，直接输入新的标签名，在文本框前面选中标签颜色，然后单击“添加”。如果已存在标签，从下拉列表中选择已有的标签，单击“添加”。

逐步标注图片中所有物体所在位置，一张图片可添加多个标签。完成一张图片标注后，可单击图片右上角来切换下一张（也可使用快捷键【D】直接切换），快速选中其他未标注的图片，然后在标注页面中执行标注操作。

图 2-22 添加物体检测标签



4. 单击页面上方“返回数据标注预览”查看标注信息，在弹框中单击“确定”保存当前标注并离开标注页面。

选中的图片被自动移动至“已标注”页签，且在“未标注”和“全部”页签中，标签的信息也将随着标注步骤进行更新，如增加的标签名称、标签对应的图片数量。

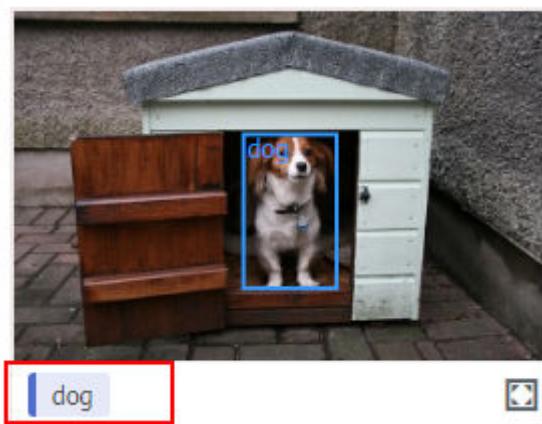
表 2-6 标注界面的常用按钮

按钮图标	功能说明
↶	撤销上一个操作。也可使用快捷键【 Ctrl+Z 】
↷	重做上一个操作。也可使用快捷键【 Ctrl+Shift+Z 】
+	放大图片。也可以使用滚轮进行放大。
-	缩小图片。也可以使用滚轮进行缩小。
☒	删除当前图片中的所有标注框。也可使用快捷键【 Shift+Delete 】
☒	显示或隐藏标注框。只有在已标注图片中可使用此操作。也可使用快捷键【 Shift+H 】
+	拖动，可将标注好的框拖动至其他位置，也可以选择框的边缘，更改框的大小。也可使用【 X+鼠标左键 】
☒	复位，与上方拖动为同组操作，当执行了拖动后，可以单击复位按钮快速将标注框恢复为拖动前的形状和位置。也可使用快捷键【 Esc 】

查看已标注图片

在标注作业详情页，单击“已标注”页签，您可以查看已完成标注的图片列表。可在每张图片的下方显示当前图片的标签信息。

图 2-23 标签信息



快速复核

当前的标注作业无法实现批量复核，如果有某一样本的标签修改或者删除，只能进入到标注页面详情进行，操作繁琐。为了简化用户操作，实现此功能，用户可以批量进行标注信息的审核或者修改，提升用户效率。

1. 登录ModelArts管理控制台，在总览页选择“数据管理>数据标注”，进入“我创建的”页签，在右上方的作业类型中下拉选择对应类型的标注作业。（仅物体检测与图像分割支持快速复核功能）
2. 在物体检测类型的标注作业列表，单击标注作业名称，进入标注详情页。
3. 单击“已标注”页签的“快速复核”，进入复核页面，对标注结果进行确认。

图 2-24 进入快速复核



4. 快速复核，支持您按照标签批量复核。
 - a. 在复核页面，单击“按照标签过滤”，选择需要复核的标签类型图片。
 - b. 在当前页面，您可以选择对当前的标签类型的图片，按照标注面积排序，或按照宽高比排序。
 - c. 依次单击需要复核的图片，在标注页面拖动图片的标注框，即可重新完成标注。（修改后的图片会带有“已修改”的信息。）
 - d. 您也可以选中需要删除标签的图片，单击右上方的 ，删除原始的标注信息。（删除后的图片会带有“已删除”的信息）

图 2-25 已修改



图 2-26 已删除

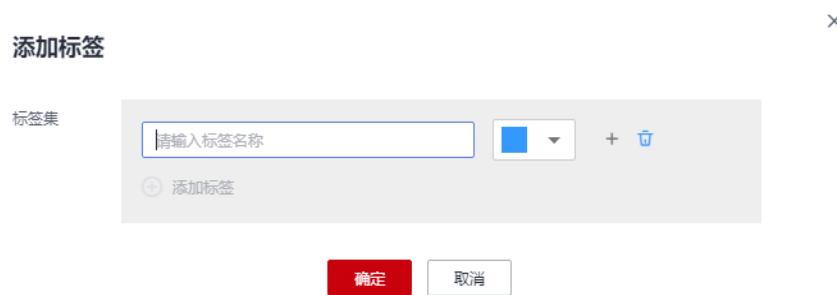


- e. 您也可以对当前已标注的图片标签信息进行修改。
- 选中待复核的图片，单击右侧的“全部标签”区域的 \oplus 按钮。
 - 输入新的标签，单击“确定”。

图 2-27 全部标签



图 2-28 添加标签



- 标注页面和标签都修改完成后，单击“应用所有修改”，在弹出的对话框单击“确定”，自动返回至标注概览页，同时会覆盖原始的标注数据。

图 2-29 应用所有修改



6. 若您对修改后的数据不满意，也可以单击“放弃修改”选择放弃本次修改，保持原有的标注数据。

图 2-30 放弃修改



表 2-7 快速复核界面的常用按钮

按钮图标	功能说明
	删除原有的标注数据，删除后可重新标注。
	还原本页所有操作至未复核页面。
	撤销上一步操作。
	重做上一步操作。

修改标注

当数据完成标注后，您还可以进入已标注页签，对已标注的数据进行修改。

- **基于图片修改**

在标注作业详情页面，单击选择“已标注”页签，然后在图片列表中选中待修改的图片，单击图片跳转到标注页面，在右侧标签信息区域中对图片信息进行修改。

- 修改标签：“标注”区域中，单击编辑图标，在文本框中输入正确的标签名，然后单击确定图标完成修改。也可以单击标签，在图片标注区域，调整标注框的位置和大小，完成调整后，右键选择“修改”，填入修改后的标签名称，单击“修改”即可保存修改。
- 删除标签：在“标注”区域中，单击删除图标即可删除此图片中的标签。标签删除后，单击页面左上角的“返回数据标注预览”离开标注页面，在弹出对话框中保存标注。图标的标签全部删除后，该图片会重新回到“未标注”页签。

图 2-31 编辑物体检测标签

标注	操作
sofa	编辑
bicycle	编辑

- **基于标签修改**

- 在标注作业详情页中，单击右侧的“标签管理”页签，在标签管理详情页，显示全部标签的信息。
 - 修改标签：单击操作列的“修改”，然后在弹出的对话框中输入修改后的标签和颜色，然后单击“确定”完成修改。修改后，之前添加了此标签的图片，都将被标注为新的标签名称。
 - 删除标签：单击操作列的“删除”，或选中需要删除的标签名称，单击上方的“删除”可删除此标签。

图 2-32 进入标签管理

全部标签 5		标签管理
标签名称		标签数量
black Cat	5	
Cat	7	
person	0	
werewr	1	
white Cat	5	

图 2-33 全部标签

操作	标签名称	属性	标签颜色	操作
修改	black Cat	矩形框	■	修改 删掉
修改	Cat	矩形框	■	修改 删掉
修改	person	矩形框	■	修改 删掉
修改	werear	圆形	■	修改 删掉
修改	white Cat	矩形框	■	修改 删掉

- 单击标注作业操作列的“标签”，也可跳转至标签管理页。

图 2-34 从标注作业列表进入标签管理



- 单击操作列的“修改”，即可完成标签的修改。
- 单击操作列的“删除”，即可删除该标签。

添加数据

除了同步数据集中的新数据外，您还可以在标注作业中，直接添加图片，用于数据标注。添加的数据将先导入至标注任务关联的数据集中，然后标注任务会自动同步数据集中最新的数据。

1. 在标注作业详情页面，单击“全部”、“已标注”或“未标注”页签，然后单击左上角“添加数据”，选择添加数据。

图 2-35 添加数据



2. 在弹出的导入对话框中，选择数据来源和导入方式，选择导入的数据路径和数据标注状态。

图 2-36 添加图片



3. 在导入对话框中，单击“确定”，完成添加数据的操作。

您添加的图片将自动呈现在“全部”的图片列表中，也可单击“添加数据>查看历史记录”，进入“任务历史”界面，可查看相应的导入历史。

图 2-37 查看历史数据

任务历史						
创建时间	导入方式	导入路径	样本总数	导入样本总数	导入已标注样本数	导入状态
2021/11/08 09:40:4...	对象存储服务 (OBS...)	obs://tupiantest1/s...	16	16	0	成功

删除图片

通过数据删除操作，可将需要丢弃的图片数据快速删除。

在“全部”、“未标注”或“已标注”页面中，依次选中需要删除的图片，或者选择“选择当前页”选中该页面所有图片，然后单击“删除”。在弹出的对话框中，根据实际情况选择是否勾选“同时删除OBS源文件”，确认信息无误后，单击“确定”完成图片删除操作。

图 2-38 删除图片



其中，被选中的图片，其左上角将显示为勾选状态。如果当前页面无选中图片时，“删除”按钮为灰色，无法执行删除操作。

说明

如果勾选了“同时删除OBS源文件”，删除图片操作将删除对应OBS目录下存储的图片，此操作可能会影响已使用此源文件的其他数据集或数据集版本，有可能导致展示异常或训练/推理异常。删除后，数据将无法恢复，请谨慎操作。

标注人员管理

若您创建的标注作业，开启了团队标注，“标注人员管理”页面中可查看团队标注作业的标注详情。添加、修改或删除标注成员。

1. 登录“数据管理>数据标注”，在“我创建的”页签下可查看所有的标注作业列表。
2. 在作业列表的“名称”列，根据标注作业名称找到对应的团队标注作业。（团队标注作业的名称后带有🔍标识。）
3. 单击作业操作列的“更多>标注人员管理”。或单击作业名称进入作业详情，继续单击右上角“团队标注>标注人员管理”，进入成员管理页面。

图 2-39 进入标注标注人员管理页（1）



图 2-40 进入标注人员管理页（2）



- 添加成员：

单击页面“添加成员”，选择成员名称，单击确定。

在操作列，选择“发送邮件”，可将该标注任务以邮件的方式发送至该标注成员。

- 修改成员信息：

单击操作列的“修改”，可修改该成员的角色。

- 删 除标注成员：

单击操作列的“删除”可删除该标注成员的所有信息。

2.2.3 图像分割

由于模型训练过程需要大量有标签的图片数据，因此在模型训练之前需对没有标签的图片添加标签。您可以通过在ModelArts控制台进行标注，也可以对已标注图片修改或删除标签进行重新标注。

针对图像分割场景，开始标注前，您需要了解：

- 图片中需要提取轮廓的物体都要标注。
- 支持使用多边形标注。
 - 多边形标注，根据目标物体的轮廓绘制多边形。
- 多边形标注或极点标注时，标注框必须在图片范围内，超出图片将导致后续作业异常。

开始标注

1. 登录ModelArts管理控制台，在左侧菜单栏中选择“数据管理> 数据标注”，进入“数据标注”管理页面。
2. 在标注作业列表右侧“所有类型”页签下拉选择标注类型。基于“标注类型”选择需要进行标注的标注作业，单击标注作业名称进入标注作业标注详情页。

图 2-41 下拉选择标注类型



3. 在标注作业标注详情中，展示此标注作业下全部数据。

同步新数据

ModelArts会自动将数据集中新增的数据同步至标注作业，包含数据及当前标注作业支持的标注信息。

为了快速获取数据集中最新数据，可在标注作业详情页的“全部”、“未标注”或“已标注”页签中，单击“同步新数据”，快速将数据集中的数据添加到标注作业中。

说明书

问题现象：

将已标注好的数据上传至OBS，同步数据后，显示为未标注。

原因分析：

可能是OBS桶设置了自动加密导致此问题。

解决方法：

需要新建OBS桶重新上传数据，或者取消桶加密后，重新上传数据。

筛选数据

在标注作业详情页面，默认展示作业中“未标注”数据，您可以在“全部”、“未标注”或页签下，在筛选条件区域，单击 \downarrow ，添加筛选条件，快速过滤出您想要查看的数据。

支持的筛选条件如下所示，您可以设置一个或多个选项进行筛选。

- 难例集：难例或非难例。
- 标签：您可以选择全部标签，或者基于您指定的标签，选中其中一个或多个。
- 文件名或目录：根据文件名称或者文件存储目录筛选。
- 标注人：选择执行标注操作的帐号名称。
- 样本属性：表示自动分组生成的属性。只有启用了自动分组任务后才可使用此筛选条件。
- 数据属性：筛选数据的来源，选择“全部”或“推理”。

图 2-42 筛选条件



标注图片（手工标注）

标注作业详情页中，展示了此标注作业中“全部”、“未标注”和“已标注”的图片，默认显示“未标注”的图片列表。

1. 在“未标注”页签图片列表中，单击图片，自动跳转到标注页面。在标注页面，常用按钮的使用可参见[表2-9](#)。
2. 选择标注方式。

在标注页面，上方工具栏提供了常用的[标注方式及常用按钮](#)，系统默认的标注方式为多边形标注。选择多边形标注或极点标注。

□ 说明

标注第一张图片时，一旦选择其中一种，其他所有图片都需要使用此方式进行标注。

表 2-8 标注方式

图标	使用说明
	多边形。在标注对象所在范围内，鼠标左键单击完成一个点的标注，沿着物体的形状边缘，通过鼠标指定多个点，最终单击到第一个点的位置，由所有的点组成一个多边形形状。使得需标注的对象在此标注框内。

表 2-9 工具栏常用按钮

按钮图标	功能说明
	撤销上一个操作。
	重做上一个操作。
	放大图片。
	缩小图片。
	删除当前图片中的所有标注框。
	显示或隐藏标注框。只有在已标注图片中可使用此操作。
	拖动，可将标注好的框拖动至其他位置，也可以选择框的边缘，更改框的大小。
	复位，与上方拖动为同组操作，当执行了拖动后，可以单击复位按钮快速将标注框恢复为拖动前的形状和位置。
	全屏显示标注的图片。

3. 标注物体。

识别图片中的物体，单击左键分别定位物体的最上、最左、最下、最右的位置点。确定位置后，单击标注区域，将弹出对话框，填入标签名称，单击“确定”。完成一张图片标注后，可单击图片下方  展开缩略图，查看图片列表，快速选中其他未标注的图片，然后在标注页面中执行标注操作。

4. 单击页面上方“返回数据标注预览”查看标注信息，在弹框中单击“确定”保存当前标注并离开标注页面。

选中的图片被自动移动至“已标注”页签，且在“未标注”和“全部”页签中，标签的信息也将随着标注步骤进行更新，如增加的标签名称、标签对应的图片数量。

查看已标注图片

在标注作业详情页中，单击“已标注”页签，您可以查看已完成标注的图片列表。单击图片进入图片标注详情，可在右侧的“当前文件标签”中了解当前图片的标签信息。

快速复核

当前的标注作业无法实现批量复核，如果有某一样本的标签修改或者删除，只能进入到标注页面详情进行，操作繁琐。为了简化用户操作，实现此功能，用户可以批量进行标注信息的审核或者修改，提升用户效率。

1. 登录ModelArts管理控制台，在总览页选择“数据管理>数据标注”，进入“我创建的”页签，在右上方的作业类型中下拉选择对应类型的标注作业。（仅物体检测与图像分割支持快速复核功能）
2. 在物体检测类型的标注作业列表，单击标注作业名称，进入标注详情页。
3. 单击“已标注”页签的“快速复核”，进入复核页面，对标注结果进行确认。

图 2-43 进入快速复核



4. 快速复核，支持您按照标签批量复核。
 - a. 在复核页面，单击“按照标签过滤”，选择需要复核的标签类型图片。
 - b. 在当前页面，您可以选择对当前的标签类型的图片，按照标注面积排序，或按照宽高比排序。
 - c. 依次单击需要复核的图片，在标注页面拖动图片的标注框，即可重新完成标注。（修改后的图片会带有“已修改”的信息。）
 - d. 您也可以选中需要删除标签的图片，单击右上方的 ，删除原始的标注信息。（删除后的图片会带有“已删除”的信息）

图 2-44 已修改



图 2-45 已删除

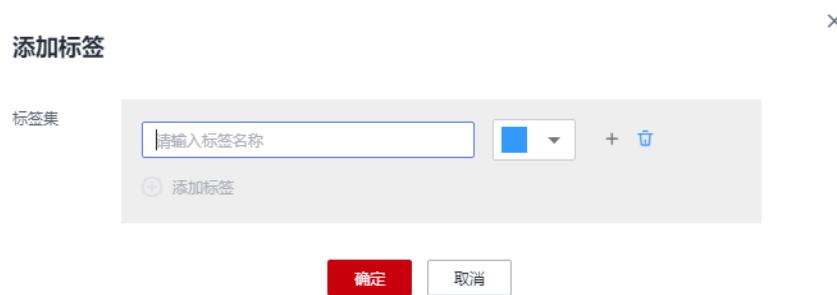


- e. 您也可以对当前已标注的图片标签信息进行修改。
- 选中待复核的图片，单击右侧的“全部标签”区域的 \oplus 按钮。
 - 输入新的标签，单击“确定”。

图 2-46 全部标签



图 2-47 添加标签



5. 标注页面和标签都修改完成后，单击“应用所有修改”，在弹出的对话框单击“确定”，自动返回至标注概览页，同时会覆盖原始的标注数据。

图 2-48 应用所有修改



- 若您对修改后的数据不满意，也可以单击“放弃修改”选择放弃本次修改，保持原有的标注数据。

图 2-49 放弃修改



表 2-10 快速复核界面的常用按钮

按钮图标	功能说明
	删除原有的标注数据，删除后可重新标注。
	还原本页所有操作至未复核页面。
	撤销上一步操作。
	重做上一步操作。

修改标注信息

当数据完成标注后，您还可以进入已标注页签，对已标注的数据进行修改。

在数据标注详情页面，单击“已标注”页签，然后在图片列表中选中待修改的图片，单击图片跳转到标注页面，在右侧标签信息区域中单击此图片已添加的标注信息。

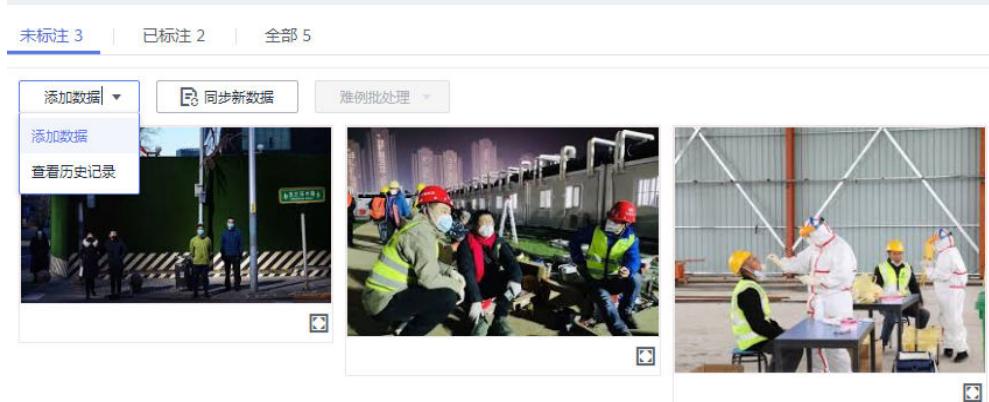
- 修改标签：在“标注”区域中，单击编辑图标，在弹出框中输入正确的标签名或标签颜色，然后单击完成修改。也可以单击标签，在图片标注区域，调整标注框的位置和大小，完成调整后，单击其他标签即可保存修改。
 - 删除标签：在“标注”区域中，单击删除图标即可删除此图片中的标签。图片的标签全部删除后，该图片会重新回到“未标注”页签。
- 标注信息修改后，单击页面左上角的“返回数据标注预览”离开标注页面，在弹出对话框中单击“确定”保存修改。

添加数据

除了同步数据集中的新数据外，您还可以在标注作业中，直接添加图片，用于数据标注。添加的数据将先导入至标注任务关联的数据集中，然后标注任务会自动同步数据集中最新的数据。

- 在标注作业详情页面，单击“全部”、“已标注”或“未标注”页签，然后单击左上角“添加数据”，选择添加数据。

图 2-50 添加数据



- 在弹出的导入对话框中，选择数据来源和导入方式，选择导入的数据路径和数据标注状态。

图 2-51 添加图片





3. 在导入对话框中，单击“确定”，完成添加数据的操作。

您添加的图片将自动呈现在“全部”的图片列表中，也可单击“添加数据>查看历史记录”，进入“任务历史”界面，可查看相应的导入历史。

图 2-52 查看历史数据

任务历史						
创建时间	导入方式	导入路径	样本总数	导入样本总数	导入已标注样本数	导入状态
2021/11/08 09:40:4...	对象存储服务 (OBS...)	obs://tupiantest1/s...	16	16	0	成功

删除图片

通过数据删除操作，可将需要丢弃的图片数据快速删除。

在“全部”、“未标注”或“已标注”页面中，依次选中需要删除的图片，或者选择“选择当前页”选中该页面所有图片，然后单击左上角“删除图片”。在弹出的对话框中，根据实际情况选择是否勾选“同时删除OBS源文件”，确认信息无误后，单击“确定”完成图片删除操作。

其中，被选中的图片，其左上角将显示为勾选状态。如果当前页面无选中图片时，“删除图片”按钮为灰色，无法执行删除操作。

说明

如果勾选了“同时删除OBS源文件”，删除图片操作将删除对应OBS目录下存储的图片，此操作可能会影响已使用此源文件的其他数据集或数据集版本，有可能导致展示异常或训练/推理异常。删除后，数据将无法恢复，请谨慎操作。

2.3 文本标注

2.3.1 文本分类

由于模型训练过程需要大量有标签的数据，因此在模型训练之前需对没有标签的文本添加标签。您也可以对已标注文本进行修改、删除和重新标注。

针对文本分类场景，是对文本的内容按照标签进行分类处理，开始标注前，您需要了解：

- 文本标注支持多标签，即一个标注对象可添加多个标签。
- 标签名是由中文、大小写字母、数字、中划线、下划线或特殊符号组成，且不超过1024位的字符串。

开始标注

1. 登录ModelArts管理控制台，在左侧菜单栏中选择“数据管理> 数据标注”，进入“数据标注”管理页面。
2. 在标注作业列表右侧“所有类型”页签下拉选择标注类型，基于“标注类型”选择需要进行标注的标注作业，单击标注作业名称进入标注作业标注详情页。

图 2-53 下拉选择标注类型



3. 在标注作业标注详情中，展示此标注作业下全部数据。

同步新数据

ModelArts会自动将数据集中新增的数据同步至标注作业，包含数据及当前标注作业支持的标注信息。

为了快速获取数据集中最新数据，可在标注作业详情页的“未标注”页签中，单击“同步新数据”，快速将数据集中的数据添加到标注作业中。

说明

问题现象：

将已标注好的数据上传至OBS，同步数据后，显示为未标注。

原因分析：

可能是OBS桶设置了自动加密导致此问题。

解决方法：

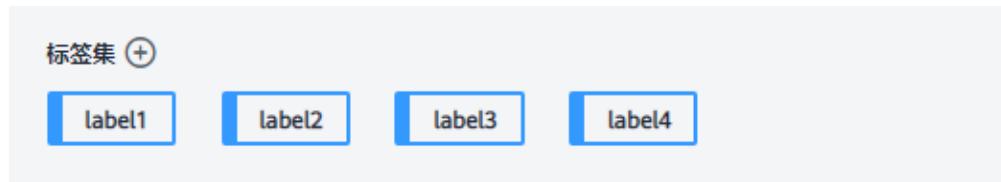
需要新建OBS桶重新上传数据，或者取消桶加密后，重新上传数据。

标注文本

标注作业详情页中，展示了此标注作业中“未标注”和“已标注”的文本，默认显示“未标注”的文本列表。

1. 在“未标注”页签文本列表中，页面左侧罗列“标注对象列表”。在列表中单击需标注的文本对象，选择右侧“标签集”中的标签进行标注。一个标注对象可添加多个标签。
以此类推，不断选中标注对象，并为其添加标签。

图 2-54 文本分类标注

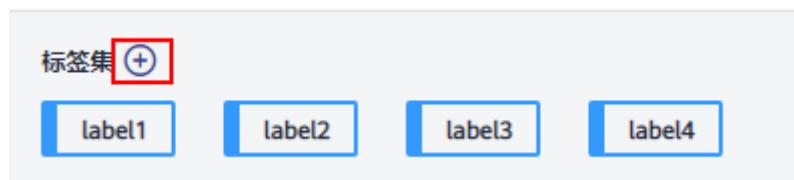


2. 当所有的标注对象都已完成标注，单击页面下方“保存当前页”，完成“未标注”列表的文本标注。

添加标签

- 在“未标注”页签添加：单击页面中标签集右侧的加号，然后在弹出的“新增标签”页中，添加标签名称，选择标签颜色，单击“确定”完成标签的新增。

图 2-55 添加标签 (1)



- 在“已标注”页签添加：单击页面中标签集右侧的加号，然后在弹出的“新增标签”页中，添加标签名称，选择标签颜色，单击“确定”完成标签的新增。

图 2-56 添加标签 (2)



图 2-57 新增标签

新增标签



查看已标注文本

在标注作业详情页，单击“已标注”页签，您可以查看已完成标注的文本列表。您也可以在右侧的“全部标签”中了解当前标注作业支持的所有标签信息。

修改标注

当数据完成标注后，您还可以进入已标注页签，对已标注的数据进行修改。

- **基于文本修改**

在标注作业详情页，单击“已标注”页签，然后在文本列表中选中待修改的文本。

在文本列表中，单击文本，当文本背景变为蓝色时，表示已选择。当文本有多个标签时，可以单击文本标签上方的~~x~~删除单个标签。

- **基于标签修改**

在标注作业详情页，单击“已标注”页签，在图片列表右侧，显示全部标签的信息。

- 批量修改：在“全部标签”区域中，单击操作列的编辑图标，然后在文本框中修改标签名称，选择标签颜色，单击“确定”完成修改。

- 批量删除：在“全部标签”区域中，单击操作列的删除图标，在弹出对话框中，可选择“仅删除标签”或“删除标签及仅包含此标签的标注对象”，然后单击“确定”。

添加文件

除了同步新数据外，您还可以在标注详情页面中，直接添加数据，用于数据标注。

1. 在标注作业详情页面，单击“未标注”页签，然后单击左上角“添加数据”。
2. 在弹出的导入对话框中，选择数据来源、导入方式、导入路径等参数，导入数据。单击确定。

导入数据的详细操作介绍请参见[数据接入简介](#)。

图 2-58 导入数据



删除文件

通过数据删除操作，可将需要丢弃的数据快速删除。

- 在“未标注”页面中，单击选中需要删除的文本对象，然后单击左上角“删除”，即可完成文本的删除操作。
- 在“已标注”页面中，选中待删除的文本对象，然后单击“删除”，删除单个文本。或者选择“选择当前页”选中该页面所有文本，然后单击左上角“删除”，即可完成当前页所有文本的删除操作。

其中，被选中的文本，其背景将显示为蓝色。

标注人员管理

若您创建的标注作业，开启了团队标注，“标注人员管理”页面中可查看团队标注作业的标注详情。添加、修改或删除标注成员。

- 登录“数据管理>数据标注”，在“我创建的”页签下可查看所有的标注作业列表。
- 在作业列表的“名称”列，根据标注作业名称找到对应的团队标注作业。（团队标注作业的名称后带有标识。）
- 单击作业操作列的“更多>标注人员管理”。或单击作业名称进入作业详情，继续单击右上角“团队标注>标注人员管理”，进入成员管理页面。

图 2-59 进入标注人员管理页 (1)

操作	任务名称	状态	进度	最后更新时间	操作
更多	task-aa55	进行中	33% (20/60)	0 2021/05/13 18:33:15 GMT+08:00	智能标注 标签 发布 删除
更多	dataset-7416	待标注	0% (0/20)	.. 2021/05/13 16:47:48 GMT+08:00	标签 发布 任务统计
更多	dataset-c42d	进行中	33% (20/60)	0 2021/05/13 16:23:04 GMT+08:00	智能标注 标签 发布

图 2-60 进入标注人员管理页 (2)



- 添加成员：
单击页面“添加成员”，选择成员名称，单击确定。
在操作列，选择“发送邮件”，可将该标注任务以邮件的方式发送至该标注成员。
- 修改成员信息：
单击操作列的“修改”，可修改该成员的角色。
- 删除标注成员：
单击操作列的“删除”可删除该标注成员的所有信息。

2.3.2 命名实体

命名实体场景，是针对文本中的实体片段进行标注，如“时间”、“地点”等。开始标注前，您需要了解：

实体命名标签名是由中文、大小写字母、数字、中划线、下划线或特殊符号组成，且不超过1024位的字符串。

开始标注

1. 登录ModelArts管理控制台，在左侧菜单栏中选择“数据管理> 数据标注”，进入“数据标注”管理页面。
2. 在标注作业列表右侧“所有类型”页签下拉选择标注类型，基于“标注类型”选择需要进行标注的标注作业，单击标注作业名称进入标注作业标注详情页。

图 2-61 下拉选择标注类型



3. 在标注作业标注详情中，展示此标注作业下全部数据。

同步新数据

ModelArts会自动将数据集中新增的数据同步至标注作业，包含数据及当前标注作业支持的标注信息。

为了快速获取数据集中最新数据，可在标注作业详情页的“未标注”页签中，单击“同步新数据”，快速将数据集中的数据添加到标注作业中。

说明

问题现象：

将已标注好的数据上传至OBS，同步数据后，显示为未标注。

原因分析：

可能是OBS桶设置了自动加密导致此问题。

解决方法：

需要新建OBS桶重新上传数据，或者取消桶加密后，重新上传数据。

标注文本

标注作业详情页中，展示了此标注作业中“未标注”和“已标注”的文本，默认显示“未标注”的文本列表。

1. 在“未标注”页签文本列表中，页面左侧罗列“标注对象列表”。在列表中单击需标注的文本对象，在右侧标签集下显示的文本内容中选中需要标注的部分，然后选择右侧“标签集”中的标签进行标注。
以此类推，不断选中标注对象，并为其添加标签。

图 2-62 命名实体标注

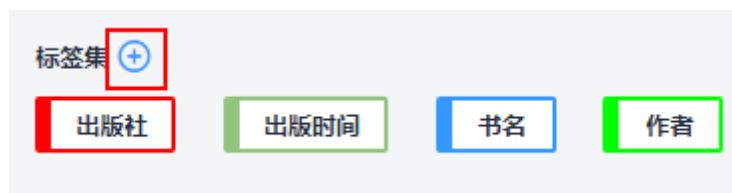


2. 单击页面下方“保存当前页”完成文本标注。

添加标签

- 在“未标注”页签添加：单击页面中标签集右侧的加号，然后在弹出的“新增标签”页中，添加标签名称，选择标签颜色，单击“确定”完成标签的新增。

图 2-63 添加命名实体标签 (1)



- 在“已标注”页签添加：单击页面中标签集右侧的加号，然后在弹出的“新增标签”页中，添加标签名称，选择标签颜色，单击“确定”完成标签的新增。

图 2-64 添加命名实体标签 (2)



图 2-65 新增命名实体标签

新增标签



查看已标注文本

在数据集详情页，单击“已标注”页签，您可以查看已完成标注的文本列表。您也可以在右侧的“全部标签”中了解当前数据集支持的所有标签信息。

修改标注

当数据完成标注后，您还可以进入“已标注”页签，对已标注的数据进行修改。

在数据集详情页，单击“已标注”页签，在右侧标签信息区域中对文本信息进行修改。

- **基于文本修改**

在数据集详情页，单击“已标注”页签，然后在文本列表中选中待修改的文本。

手工点选删除：在文本列表中，单击文本，当文本背景变为蓝色时，表示已选择。在页面右侧，单击文本标签上方的~~x~~删除单个标签。

- **基于标签修改**

在数据集详情页，单击“已标注”页签，在图片列表右侧，显示全部标签的信息。

- 批量修改：在“全部标签”区域中，单击操作列的编辑按钮，然后在文本框中添加标签名称，选择标签颜色，单击“确定”完成修改。
- 批量删除：在“全部标签”区域中，单击操作列的删除按钮，在弹出对话框中，可选择“仅删除标签”或“删除标签及仅包含此标签的标注对象”，然后单击“确定”。

添加文件

除了同步新数据外，您还可以在标注详情页面中，直接添加数据，用于数据标注。

1. 在标注作业详情页面，单击“未标注”页签，然后单击左上角“添加数据”。
2. 在弹出的导入对话框中，选择数据来源、导入方式、导入路径等参数，导入数据。单击确定。

导入数据的详细操作介绍请参见[数据接入简介](#)。

图 2-66 导入数据



删除文件

通过数据删除操作，可将需要丢弃的文件数据快速删除。

- 在“未标注”页面中，单击选中需要删除的文本对象，然后单击左上角“删除”，即可完成文本的删除操作。
- 在“已标注”页面中，选中待删除的文本对象，然后单击“删除”，删除单个文本。或者选择“选择当前页”选中该页面所有文本，然后单击左上角“删除”，即可完成当前页所有文本的删除操作。

其中，被选中的文本，其背景将显示为蓝色。

标注人员管理

若您创建的标注作业，开启了团队标注，“标注人员管理”页面中可查看团队标注作业的标注详情。添加、修改或删除标注成员。

- 登录“数据管理>数据标注”，在“我创建的”页签下可查看所有的标注作业列表。
- 在作业列表的“名称”列，根据标注作业名称找到对应的团队标注作业。（团队标注作业的名称后带有^人标识。）
- 单击作业操作列的“更多>标注人员管理”。或单击作业名称进入作业详情，继续单击右上角“团队标注>标注人员管理”，进入成员管理页面。

图 2-67 进入标注人员管理页 (1)

	task-aa55	dataset-c42d	33% (20/60)	0 2021/05/13 18:33:15 GMT+08:00	..	智能标注 标签 发布	删除
▼		dataset-7416	0% (0/20)	-- 2021/05/13 16:47:48 GMT+08:00	..	标签 发布 任务统计	任务统计
▼		dataset-c42d	33% (20/60)	0 2021/05/13 16:23:04 GMT+08:00	..	智能标注 标签 发布	标注人员管理 修改

图 2-68 进入标注人员管理页 (2)



- 添加成员：
单击页面“添加成员”，选择成员名称，单击确定。
在操作列，选择“发送邮件”，可将该标注任务以邮件的方式发送至该标注成员。
- 修改成员信息：
单击操作列的“修改”，可修改该成员的角色。
- 删除标注成员：
单击操作列的“删除”可删除该标注成员的所有信息。

2.3.3 文本三元组

三元组标注适用于标注出语句当中形如（主语/Subject，谓词/Predicate，宾语/Object）结构化知识的场景，标注时不但可以标注出语句当中的实体，还可以标注出实体之间的关系，其在依存句法分析、信息抽取等自然语言处理任务中经常用到。

文本三元组类型的数据标注，需要关注两种标签，“实体标签”和“关系标签”。
“关系标签”需设置对应的“起始实体”和“终止实体”。

- 支持设置多个“实体标签”和“关系标签”。一个文本数据中，也可以标注多个“实体标签”和“关系标签”
- 创建数据集时定义的“实体标签”，不支持删除。

注意事项

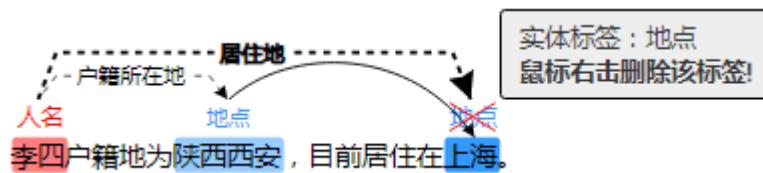
在开始标注之前，需确保标注作业对应的“实体标签”和“关系标签”已定义好。
“关系标签”需设置对应的“起始实体”和“终止实体”。“关系标签”只能添加至其设置好的“起始实体”和“终止实体”之间。

例如，如图2-69所示，当两个文本都被标注为“地点”，那么针对这两个实体，无法添加本示例中的任意一个关系标签。当无法添加某个关系标签时，界面将显示一个红色的叉号，如图2-70所示。

图 2-69 实体标签和关系标签的示例



图 2-70 无法添加关系标签



开始标注

1. 登录ModelArts管理控制台，在左侧菜单栏中选择“数据管理> 数据标注”，进入“数据标注”管理页面。
2. 在标注作业列表右侧“所有类型”页签下拉选择标注类型，基于“标注类型”选择需要进行标注的标注作业，单击标注作业名称进入标注作业标注详情页。

图 2-71 下拉选择标注类型



3. 在标注作业标注详情中，展示此标注作业下全部数据。

同步新数据

ModelArts会自动将数据集中新增的数据同步至标注作业，包含数据及当前标注作业支持的标注信息。

为了快速获取数据集中最新数据，可在标注作业详情页的“未标注”页签中，单击“同步新数据”，快速将数据集中的数据添加到标注作业中。

说明

问题现象：

将已标注好的数据上传至OBS，同步数据后，显示为未标注。

原因分析：

可能是OBS桶设置了自动加密导致此问题。

解决方法：

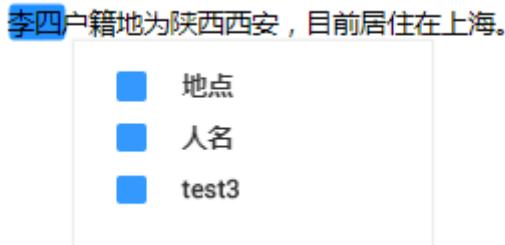
需要新建OBS桶重新上传数据，或者取消桶加密后，重新上传数据。

标注文本

标注作业详情页中，展示了此标注作业中“未标注”和“已标注”的文本，默认显示“未标注”的文本列表。

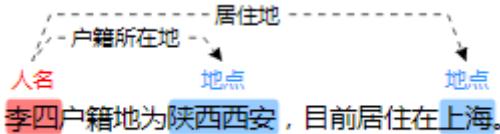
- 在“未标注”页签文本列表中，页面左侧罗列“标注对象列表”。在列表中单击需标注的文本对象，选中相应文本内容，在页面呈现的实体类型列表中选择实体名称，完成实体标注。

图 2-72 实体标注



- 在完成多个实体标注后，鼠标左键依次单击起始实体和终止实体，在呈现的关系类型列表中选择一个对应的关系类型，完成关系标注。

图 2-73 关系标注



- 当所有的标注对象都已完成标注，单击页面下方“保存当前页”完成“未标注”列表的文本标注。

说明

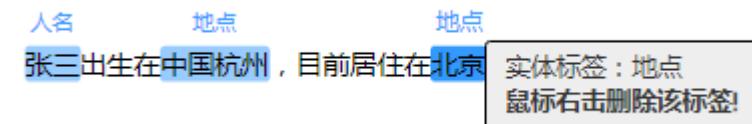
“文本三元组”类型的数据集，不支持在标注页面修改标签，需要进入“标签管理”页面，修改“实体标签”和“关系标签”。

修改标注

当数据完成标注后，您还可以进入已标注页签，对已标注的数据进行修改。

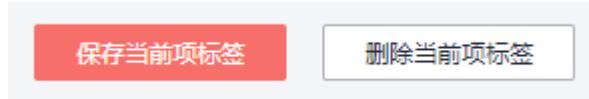
在标注作业详情页，单击“已标注”页签，在左侧文本列表中选中一行文本，右侧区域显示具体的标注信息。将鼠标移动至对应的实体标签或关系类型，单击鼠标右键，可删除此标注。单击鼠标左键，依次单击连接起始实体和终止实体，可增加关系类型，增加关系标注。

图 2-74 在文本中修改标签



您也可以在单击页面下方的“删除当前项标签”按钮，删除选中文本对象中的所有标签。

图 2-75 删除当前项标签



添加文件

除了同步新数据外，您还可以在标注详情页面中，直接添加数据，用于数据标注。

1. 在标注作业详情页面，单击“未标注”页签，然后单击左上角“添加数据”。
2. 在弹出的导入对话框中，选择数据来源、导入方式、导入路径等参数，导入数据。单击确定。

导入数据的详细操作介绍请参见[数据接入简介](#)。

图 2-76 导入数据



删除文件

通过数据删除操作，可将需要丢弃的文件数据快速删除。

- 在“未标注”页面中，单击选中需要删除的文本，然后单击左上角“删除”，即可完成文本的删除操作。
- 在“已标注”页面中，选中待删除的文本，然后单击“删除”，删除单个文本。或者勾选“选择当前页”选中该页面所有文本，然后单击左上角“删除”，即可完成当前页所有文本的删除操作。

其中，被选中的文本，其背景将显示为蓝色。如果当前页面无选中文本时，“删除”按钮为灰色，无法执行删除操作。

标注人员管理

若您创建的标注作业，开启了团队标注，“标注人员管理”页面中可查看团队标注作业的标注详情。添加、修改或删除标注成员。

- 登录“数据管理>数据标注”，在“我创建的”页签下可查看所有的标注作业列表。
- 在作业列表的“名称”列，根据标注作业名称找到对应的团队标注作业。（团队标注作业的名称后带有?标识。）
- 单击作业操作列的“更多>标注人员管理”。或单击作业名称进入作业详情，继续单击右上角“团队标注>标注人员管理”，进入成员管理页面。

图 2-77 进入标注人员管理页 (1)

操作	名称	状态	进度	时间	操作	更多
...	task-aa55 ?	dataset-c42d	33% (20/60)	0 2021/05/13 18:33:15 GMT+08:00	...	智能标注 标签 发布 删除 更多
...	dataset-7416	dataset-7416	0% (0/20)	.. 2021/05/13 16:47:48 GMT+08:00	...	标签 发布 任务统计
...	dataset-c42d	dataset-c42d	33% (20/60)	0 2021/05/13 16:23:04 GMT+08:00	...	智能标注 标签 发布 更多

图 2-78 进入标注人员管理页 (2)



- 添加成员：
单击页面“添加成员”，选择成员名称，单击确定。
在操作列，选择“发送邮件”，可将该标注任务以邮件的方式发送至该标注成员。
- 修改成员信息：
单击操作列的“修改”，可修改该成员的角色。
- 删除标注成员：
单击操作列的“删除”可删除该标注成员的所有信息。

2.4 音频标注

2.4.1 声音分类

由于模型训练过程需要大量有标签的音频数据，因此在模型训练之前需对没有标签的音频添加标签。通过ModelArts您可对音频进行一键式批量添加标签，快速完成对音频的标注操作，也可以对已标注音频修改或删除标签进行重新标注。

开始标注

1. 登录ModelArts管理控制台，在左侧菜单栏中选择“数据管理> 数据标注”，进入“数据标注”管理页面。
2. 在标注作业列表右侧“所有类型”页签下拉选择标注类型，基于“标注类型”选择需要进行标注的标注作业，单击标注作业名称进入标注作业标注详情页。

图 2-79 下拉选择标注类型



3. 在标注作业标注详情中，展示此标注作业下全部数据。

同步新数据

ModelArts会自动将数据集中新增的数据同步至标注作业，包含数据及当前标注作业支持的标注信息。

为了快速获取数据集中最新数据，可在标注作业详情页的“未标注”和“已标注”页签中，单击“同步新数据”，快速将数据集中的数据添加到标注作业中。

说明

问题现象：

将已标注好的数据上传至OBS，同步数据后，显示为未标注。

原因分析：

可能是OBS桶设置了自动加密导致此问题。

解决方法：

需要新建OBS桶重新上传数据，或者取消桶加密后，重新上传数据。

标注音频

标注作业详情页中，展示了此标注作业中“未标注”和“已标注”的音频，默认显示“未标注”的音频列表。单击音频左侧 ，即可进行音频的试听。

1. 在“未标注”页签，勾选需进行标注的音频。
 - 手工点选：在音频列表中，单击音频，当右上角出现蓝色勾选框时，表示已勾选。可勾选同类别的多个音频，一起添加标签。
 - 批量选中：如果音频列表的当前页，所有音频属于一种类型，可以在列表的右上角单击“选择当前页”，则当前页面所有的音频将选中。
2. 添加标签。
 - a. 在右侧的“添加标签”区域中，单击“标签”下侧的文本框设置标签。

方式一（已存在标签）：单击“标签”下方的文本框，在快捷键下拉列表中选择快捷键，然后在标签文本输入框中选择已有的标签名称，然后单击“确定”。

方式二（新增标签）：在“标签”下方的文本框中，在快捷键下拉列表中选择快捷键，然后在标签文本输入框中输入新的标签名称，然后单击“确定”。

- b. 选中的音频将被自动移动至“已标注”页签，且在“未标注”页签中，标签的信息也将随着标注步骤进行更新，如增加的标签名称、各标签对应的音频数量。

□ 说明

快捷键的使用说明：为标签指定快捷键后，当您选择一段音频后，在键盘中按一下快捷键，即可为此音频增加为此快捷键对应的标签。例如“aa”标签对应的快捷键是“1”，在数据标注过程中，选中1个或多个文件，按“1”，界面将提示是否需要将此文件标注为“aa”标签，单击确认即可完成标注。

快捷键对应的是标签，1个标签对应1个快捷键。不同的标签，不能指定为同一个快捷键。快捷键的使用，可以大大提升标注效率。

图 2-80 添加音频标签



查看已标注音频

在标注作业详情页，单击“已标注”页签，您可以查看已完成标注的音频列表。单击音频，可在右侧的“选中文件标签”中了解当前音频的标签信息。

修改标注

当数据完成标注后，您还可以进入“已标注”页签，对已标注的数据进行修改。

• 基于音频修改

在标注作业详情页面，单击“已标注”页签，然后在音频列表中选中待修改的音频（选择一个或多个）。在右侧标签信息区域中对标签进行修改。

- 修改标签：在“选中文件标签”区域中，单击操作列的编辑图标，然后在文本框中输入正确的标签名，然后单击确定图标完成修改。
- 删除标签：在“选中文件标签”区域中，单击操作列的删除图标删除该标签。

• 基于标签修改

在标注作业详情页面，单击“已标注”页签，在音频列表右侧，显示全部标签的信息。

图 2-81 全部标签信息

标签	数量	快捷键	操作
8	1	6	
aaaa	7	1	
qqqq	1	2	
test	1	--	

- 修改标签：单击操作列的编辑图标，然后在弹出的对话框中输入修改后的标签名，然后单击“确定”完成修改。修改后，之前添加了此标签的音频，都将被标注为新的标签名称。
- 删除标签：单击操作列的删除图标，在弹出的对话框中，根据提示框选择需要删除的对象，然后单击“确定”完成删除。

添加音频

除了同步新数据外，您还可以在标注详情页面中，直接添加数据，用于数据标注。

1. 在标注作业详情页面，单击“未标注”或“已标注”页签，然后单击左上角“添加数据”。
2. 在弹出的导入对话框中，选择数据来源、导入方式、导入路径等参数，导入数据。单击确定。

导入数据的详细操作介绍请参见[数据接入简介](#)。

图 2-82 导入数据



删除音频

通过数据删除操作，可将需要丢弃的音频数据快速删除。

在“未标注”或“已标注”页面中，选中需要删除的音频，或者选择“选择当前页”选中该页面所有音频，然后单击左上角“删除音频”，在弹出的对话框中，根据实际情况选择是否勾选“同时删除OBS源文件”，确认信息无误后，单击“确定”完成音频删除操作。

其中，被选中的音频，其右上角将显示为勾选状态。如果当前页面无选中音频时，“删除音频”按钮为灰色，无法执行删除操作。

说明

如果勾选了“同时删除OBS源文件”，删除音频操作是将删除对应OBS目录下存储的音频。此操作可能会影响已使用此源文件的其他数据集或数据集版本，有可能导致展示异常或训练/推理异常。删除后，数据将无法恢复，请谨慎操作。

2.4.2 语音内容

由于模型训练过程需要大量有标签的音频数据，因此在模型训练之前需对没有标签的音频添加标签。通过ModelArts您可对音频进行一键式批量添加标签，快速完成对音频的标注操作，也可以对已标注音频修改或删除标签进行重新标注。

开始标注

1. 登录ModelArts管理控制台，在左侧菜单栏中选择“数据管理> 数据标注”，进入“数据标注”管理页面。
2. 在标注作业列表右侧“所有类型”页签下拉选择标注类型，基于“标注类型”选择需要进行标注的标注作业，单击标注作业名称进入标注作业标注详情页。

图 2-83 下拉选择标注类型



3. 在标注作业标注详情中，展示此标注作业下全部数据。

同步新数据

ModelArts会自动将数据集中新增的数据同步至标注作业，包含数据及当前标注作业支持的标注信息。

为了快速获取数据集中最新数据，可在标注作业详情页的“未标注”页签中，单击“同步新数据”，快速将数据集中的数据添加到标注作业中。

说明

问题现象：

将已标注好的数据上传至OBS，同步数据后，显示为未标注。

原因分析：

可能是OBS桶设置了自动加密导致此问题。

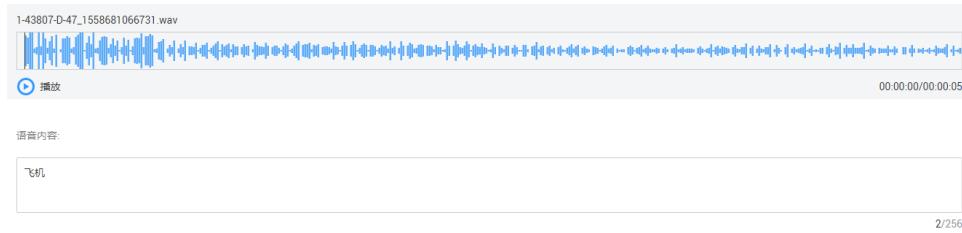
解决方法：

需要新建OBS桶重新上传数据，或者取消桶加密后，重新上传数据。

标注音频

标注作业详情页中，展示了此数据集中“未标注”和“已标注”的音频，默认显示“未标注”的音频列表。

1. 在“未标注”页签左侧音频列表中，单击目标音频文件，在右侧的区域中出现音频，单击音频下方 ，即可进行音频播放。
2. 根据播放内容，在下方“语音内容”文本框中填写音频内容。
3. 输入内容后单击下方的“确认标注”按钮完成标注。音频将被自动移动至“已标注”页签。

图 2-84 语音内容音频标注

查看已标注音频

在标注作业详情页，单击“已标注”页签，您可以查看已完成标注的音频列表。单击音频，可在右侧的“语音内容”文本框中了解当前音频的内容信息。

修改标注

当数据完成标注后，您还可以进入“已标注”页签，对已标注的数据进行修改。

在标注作业详情页，单击“已标注”页签，然后在音频列表中选中待修改的音频。在右侧标签信息区域中修改“语音内容”文本框中的内容，单击下方的“确认标注”按钮完成修改。

添加音频

除了同步新数据外，您还可以在标注详情页面中，直接添加数据，用于数据标注。

1. 在标注作业详情页面，单击“未标注”页签，然后单击左上角“添加数据”。
2. 在弹出的导入对话框中，选择数据来源、导入方式、导入路径等参数，导入数据。单击确定。

导入数据的详细操作介绍请参见[数据接入简介](#)。

图 2-85 导入数据

删除音频

通过数据删除操作，可将需要丢弃的音频数据快速删除。

在“未标注”或“已标注”页面中，选中需要删除的音频，然后单击左上角“删除音频”，在弹出的对话框中，根据实际情况选择是否勾选“同时删除OBS源文件”，确认信息无误后，单击“确定”完成音频删除操作。

说明

如果勾选了“同时删除源OBS文件”，删除音频操作是将删除对应OBS目录下存储的音频。此操作可能会影响已使用此源文件的其他数据集或数据集版本，有可能导致展示异常或训练/推理异常。删除后，数据将无法恢复，请谨慎操作。

2.4.3 语音分割

由于模型训练过程需要大量有标签的音频数据，因此在模型训练之前需对没有标签的音频添加标签。通过ModelArts您可对音频添加标签，快速完成对音频的标注操作，也可以对已标注音频修改或删除标签进行重新标注。

开始标注

1. 登录ModelArts管理控制台，在左侧菜单栏中选择“数据管理> 数据标注”，进入“数据标注”管理页面。
2. 在标注作业列表右侧“所有类型”页签下拉选择标注类型，基于“标注类型”选择需要进行标注的标注作业，单击标注作业名称进入标注作业标注详情页。

图 2-86 下拉选择标注类型



3. 在标注作业标注详情中，展示此标注作业下全部数据。

同步数据源

ModelArts会自动从数据集输入位置同步数据至数据集详情页，包含数据及标注信息。

为了快速获取OBS桶中最新数据，可在数据集详情页的“未标注”页签中，单击“同步数据源”，快速将通过OBS上传的数据添加到数据集中。

说明书

问题现象：

将已标注好的数据上传至OBS，同步数据后，显示为未标注。

原因分析：

可能是OBS桶设置了自动加密导致此问题。

解决方法：

需要新建OBS桶重新上传数据，或者取消桶加密后，重新上传数据。

标注音频

标注作业详情页中，展示了此标注作业中“未标注”和“已标注”的音频，默认显示“未标注”的音频列表。

- 在“未标注”页签左侧音频列表中，单击目标音频文件，在右侧的区域中出现音频，单击音频下方，即可进行音频播放。
- 根据播放内容，选取合适的音频段，在下方“语音内容”文本框中填写音频标签和内容。

图 2-87 语音标签音频标注



- 输入内容后单击下方的“确认标注”按钮完成标注。音频将被自动移动至“已标注”页签。

查看已标注音频

在标注作业详情页，单击“已标注”页签，您可以查看已完成标注的音频列表。单击音频，可在右侧的对应的文本框中了解当前音频的内容信息。

修改标注

当数据完成标注后，您还可以进入“已标注”页签，对已标注的数据进行修改。

- 修改标签：在数据标注详情页，单击“已标注”页签，然后在音频列表中选中待修改的音频。在右侧下方的标签信息区域中修改语音内容中的“标签”和，单击下方的“确认标注”按钮完成修改。
- 删除标签：单击目标编号操作列的“删除”，删除该段音频的标注。您也可以单击标注音频文件上方的删除标注，然后单击“确认标注”。

添加音频

除了同步新数据外，您还可以在标注详情页面中，直接添加数据，用于数据标注。

1. 在标注作业详情页面，单击“未标注”页签，然后单击左上角“添加数据”。
2. 在弹出的导入对话框中，选择数据来源、导入方式、导入路径等参数，导入数据。单击确定。

图 2-88 导入数据



删除音频

通过数据删除操作，可将需要丢弃的音频数据快速删除。

在“未标注”或“已标注”页面中，选中需要删除的音频，然后单击左上角“删除音频”，在弹出的对话框中，根据实际情况选择是否勾选“同时删除OBS源文件”，确认信息无误后，单击“确定”完成音频删除操作。

说明

如果勾选了“同时删除OBS源文件”，删除音频操作是将删除对应OBS目录下存储的音频。此操作可能会影响已使用此源文件的其他数据集或数据集版本，有可能导致展示异常或训练/推理异常。删除后，数据将无法恢复，请谨慎操作。

标注人员管理

若您创建的标注作业，开启了团队标注，“标注人员管理”页面中可查看团队标注作业的标注详情。添加、修改或删除标注成员。

1. 登录“数据管理>数据标注”，在“我创建的”页签下可查看所有的标注作业列表。
2. 在作业列表的“名称”列，根据标注作业名称找到对应的团队标注作业。（团队标注作业的名称后带有标识。）
3. 单击作业操作列的“更多>标注人员管理”。或单击作业名称进入作业详情，继续单击右上角“团队标注>标注人员管理”，进入成员管理页面。

图 2-89 进入标注人员管理页 (1)

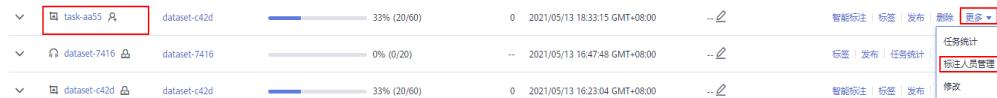


图 2-90 进入标注人员管理页 (2)



- **添加成员:**
单击页面“添加成员”，选择成员名称，单击确定。
在操作列，选择“发送邮件”，可将该标注任务以邮件的方式发送至该标注成员。
- **修改成员信息:**
单击操作列的“修改”，可修改该成员的角色。
- **删除标注成员:**
单击操作列的“删除”可删除该标注成员的所有信息。

2.5 视频标注

由于模型训练过程需要大量有标签的视频数据，因此在模型训练之前需对没有标签的视频添加标签。通过ModelArts您可对视频添加标签，快速完成对视频的标注操作，也可以对已标注视频修改或删除标签进行重新标注。

说明

视频标注仅针对视频帧进行标注。

开始标注

1. 登录ModelArts管理控制台，在左侧菜单栏中选择“数据管理> 数据标注”，进入“数据标注”管理页面。
2. 在标注作业列表右侧“所有类型”页签下拉选择标注类型，基于“标注类型”选择需要进行标注的标注作业，单击标注作业名称进入标注作业标注详情页。

图 2-91 下拉选择标注类型



3. 在标注作业标注详情中，展示此标注作业下全部数据。

同步数据源

ModelArts会自动从数据集输入位置同步数据至数据集详情页，包含数据及标注信息。

为了快速获取OBS桶中最新数据，可在数据集详情页的“已标注”或“未标注”页签中，单击“同步数据源”，快速将通过OBS上传的数据添加到数据集中。

说明

问题现象：

将已标注好的数据上传至OBS，同步数据后，显示为未标注。

原因分析：

可能是OBS桶设置了自动加密导致此问题。

解决方法：

需要新建OBS桶重新上传数据，或者取消桶加密后，重新上传数据。

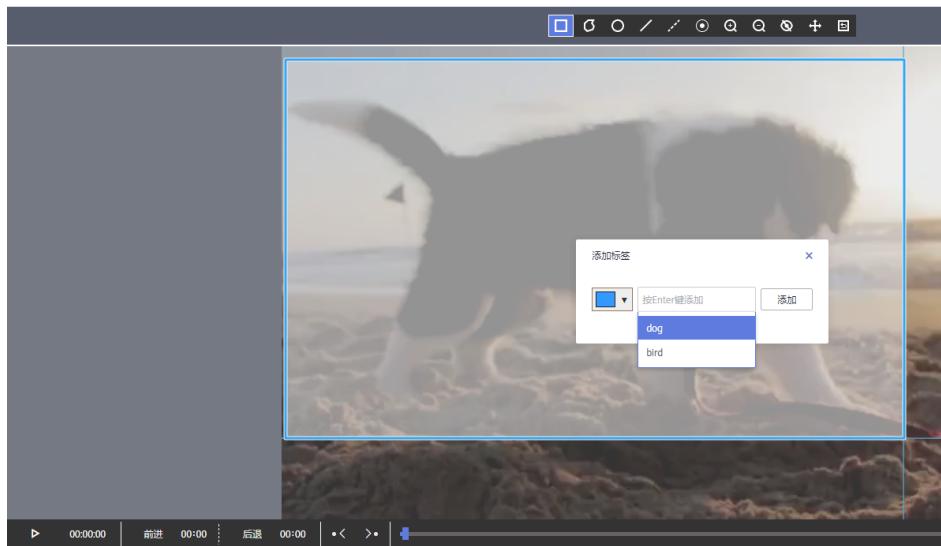
视频标注

标注作业详情页中，展示了此数据集中“未标注”、“已标注”和“全部”的视频。

1. 在“未标注”页签左侧视频列表中，单击目标视频文件，打开标注页面。
2. 在标注页面中，播放视频，当视频播放至待标注时间时，单击进度条左侧的暂停按钮，将视频暂停至某一帧对应的画面。
3. 在上方区域选择标注框，默认为矩形框。使用鼠标在视频画面中框出目标，然后在弹出的添加标签文本框中，直接输入新的标签名，在文本框前面选中标签颜色，单击“添加”完成1个物体的标注。如果已存在标签，从下拉列表中选择已有的标签，然后单击“添加”完成标注。逐步此画面中所有物体所在位置，一帧对应的画面可添加多个标签。

支持的标注框与“物体检测”类型一致，详细描述请参见物体检测章节的[表2 标注界面的常用按钮](#)。

图 2-92 视频标注



4. 上一帧对应的画面标注完成后，在进度条处单击播放按钮继续播放，在需要标注处暂停，然后重复执行步骤3完成整个视频的标注。
单击界面右上角的“标注列表”，在“当前文件标签”的详情页将呈现当前视频带标注的时间点。

图 2-93 当前文件标签信息

当前文件标签

00:00:01 删除 下拉菜单

5. 单击页面左上角“返回数据标注预览”，页面将自动返回标注作业详情页面，同时，标注好的视频将呈现在“已标注”页签下。

常见问题

Q：视频数据集无法显示或者无法播放视频？

A：若无法显示和播放视频，请检查视频格式类型，目前只支持MP4格式。

修改标注

当数据完成标注后，您还可以进入“已标注”页签，修改标注数据。

- 在“已标注”页签下，单击目标视频文件，在标注页面的右上方选择“标注列表”进入“当前文件标签”详情页面，可单击时间点右侧小三角展开详情，您可以修改或删除标签。
- 修改标签：在“当前文件标签”详情页，单击标签右侧的编辑按钮，标签名称可进行修改。

- 删除标签：在“当前文件标签”详情页，单击标签右侧的删除按钮，将直接删除此标签。如果单击画面时间右侧的删除按钮，将删除此画面下的所有标签。

图 2-94 修改标注

当前文件标签

00:00:07			
dog			
00:00:02			
dog			
00:00:09			
dog			

添加视频

除了同步新数据外，您还可以在标注详情页面中，直接添加数据，用于数据标注。

- 在标注作业详情页面，单击“未标注”或“已标注”页签，然后单击左上角“添加数据”。
- 在弹出的导入对话框中，选择数据来源、导入方式、导入路径等参数，导入数据。单击确定。

图 2-95 导入数据



删除视频

通过数据删除操作，可将需要丢弃的视频数据快速删除。

在“全部”、“未标注”或“已标注”页面中，依次选中需要删除的视频，或者选择“选择当前页”选中该页面所有视频，然后单击上边的“删除”。在弹出的对话框中，根据实际情况选择是否勾选“同时删除OBS源文件”，确认信息无误后，单击“确定”完成视频删除操作。

其中，被选中的视频，其左上角将显示为勾选状态。如果当前页面无选中视频时，“删除视频”按钮为灰色，无法执行删除操作。

说明

如果勾选了“同时删除OBS源文件”，删除视频操作将删除对应OBS目录下存储的视频，此操作可能会影响已使用此源文件的其他数据集或数据集版本，有可能导致展示异常或训练/推理异常。删除后，数据将无法恢复，请谨慎操作。

2.6 查看标注作业

2.6.1 查看创建的作业

在ModelArts数据标注页面可查看用户自己创建的标注作业。

操作步骤

1. 登录ModelArts管理控制台，在左侧菜单栏选择“数据管理>数据标注”，进入数据标注页面。
2. 在数据标注详情页选择“我创建的”，即可展示所有自己创建的标注作业。用户可查看自己创建的标注作业的相关信息，

图 2-96 我创建的

名称	数据集	标注进度 (已标注个数/总数)	待确认个数	创建时间	描述	操作
test-obj	person-object-detection	100% (334/334)	--	2023/02/15 16:37:35 GMT+08:00	..	智能标注 标签 发布 删除 更多 ▾
ExeML_9db7	dataset-647f	62% (1110/1790)	--	2023/02/07 19:31:22 GMT+08:00	..	标签 发布 删除 任务统计 修改
ExeML_Secc	dataset-5ffe	0% (0/20)	--	2023/02/07 17:32:19 GMT+08:00	..	智能标注 标签 发布 删除 更多 ▾
ExeML_4338	dataset-99c5	0% (0/0)	--	2023/02/07 17:29:09 GMT+08:00	..	标签 发布 删除 任务统计 修改

标注作业复制

1. 登录ModelArts管理控制台，在左侧菜单栏选择“数据管理>数据标注”，进入数据标注页面。
2. 在数据标注列表页，“我创建的”页签下，选择需要复制的标注任务。
3. 单击作业操作列的“更多>复制”。
4. 在标注任务复制的弹窗中，填写作业描述，作业名称task-xxxx-copy-xxxx，其中xxxx为系统生成的随机码，用来区分新作业与被复制作业。也可以修改新生成的作业名称。单击“确定”。

复制标注任务

您正在进行标注任务 dataset-a4d9 的复制操作，点击“确定”后会基于此任务生成新的标注任务。

名称

dataset-a4d9-copy-4a7f

描述

0/256

⚠ 复制操作基于作业关联的数据集进行任务重建，如数据集样本发生变化，新生成的作业会同步变化，同时不支持标注结果的复制。

确定取消

5. 复制完成后，在标注作业列表页即可查询新的标注任务，拷贝标注作业信息包含，拷贝标注任务的样本、标签、团队标注信息。

2.6.2 查看参与标注的作业

在ModelArts数据标注页面可查看用户参与的标注作业。

前提条件

创建标注作业时，启用了团队标注。

操作步骤

1. 登录ModelArts管理控制台。在左侧菜单栏选择“数据管理>数据标注”，进入数据标注详情页面。
2. 在数据标注详情页选择“我参与的”，即可展示所有参与过标注的标注作业，用户可查看自己参与过的标注作业的详情。包括标注团队的成员、标注进展等情况。

3 智能标注

3.1 创建智能标注作业

除了人工标注外，ModelArts还提供了智能标注功能，快速完成数据标注，为您节省70%以上的标注时间。智能标注是指基于当前标注阶段的标签及图片学习训练，选中系统中已有的模型进行智能标注，快速完成剩余图片的标注操作。

背景信息

- 目前只有“图像分类”和“物体检测”类型的标注作业支持智能标注功能。
- 启动智能标注时，需标注作业存在至少2种标签，且每种标签已标注的图片不少于5张。
- 启动智能标注时，必须存在未标注图片。
- 启动智能标注前，保证当前系统中不存在正在进行中的智能标注任务。
- 检查用于标注的图片数据，确保您的图片数据中，不存在RGBA四通道图片。如果存在四通道图片，智能标注任务将运行失败，因此，请从数据集中删除四通道图片后，再启动智能标注。

说明

数据管理模块即将下线，对未使用过数据管理的用户不可见。

启动智能标注作业

1. 登录ModelArts管理控制台，在左侧菜单栏中选择“数据管理 > 数据标注”，进入“数据标注”管理页面。
2. 在标注作业列表中，选择“物体检测”或“图像分类”类型的标注作业，单击操作列的“智能标注”启动智能标注作业。
3. 在弹出的“启动智能标注”对话框中，选择智能标注类型，可选“主动学习”或者“预标注”，详见[表3-1](#)和[表3-2](#)。

表 3-1 主动学习

参数	说明
智能标注类型	“主动学习”。“主动学习”表示系统将自动使用半监督学习、难例筛选等多种手段进行智能标注，降低人工标注量，帮助用户找到难例。
算法类型	针对“图像分类”类型的数据集，您需要选择以下参数。 “快速型”：仅使用已标注的样本进行训练。 “精准型”：会额外使用未标注的样本做半监督训练，使得模型精度更高。
计算节点规格	即智能标注任务使用的资源规格，仅支持GPU规格，且免费公测。 说明 智能标注创建时免费，但OBS存储会按需收费，请参考 计费详情 。为保证您的资源不浪费，标注作业与后续任务完成后，请及时清理您的OBS桶。
计算节点个数	默认为1，表示单机模式。目前仅支持此参数值。

表 3-2 预标注

参数	说明
智能标注类型	“预标注”。“预标注”表示选择用户模型管理里面的模型，选择模型时需要注意模型类型和数据集的标注类型相匹配。预标注结束后，如果标注结果符合平台定义的标准标注格式，系统将进行难例筛选，该步骤不影响预标注结果。
选择模型及版本	<ul style="list-style-type: none">“我的AI应用”。您可以根据实际需求选择您的模型。您需要在目标AI应用的左侧单击下拉三角标，选择合适的版本。您的模型导入参见创建AI应用。
计算节点规格	在下拉框中，您可以选择目前ModelArts支持的节点规格选项。
计算节点个数	默认为1。您可以根据您的实际情况选择，最大为5。

□□ 说明

针对“物体检测”类型的标注作业，选择“主动学习”时，只支持识别和标注矩形框。

图 3-1 启动智能标注（图像分类）

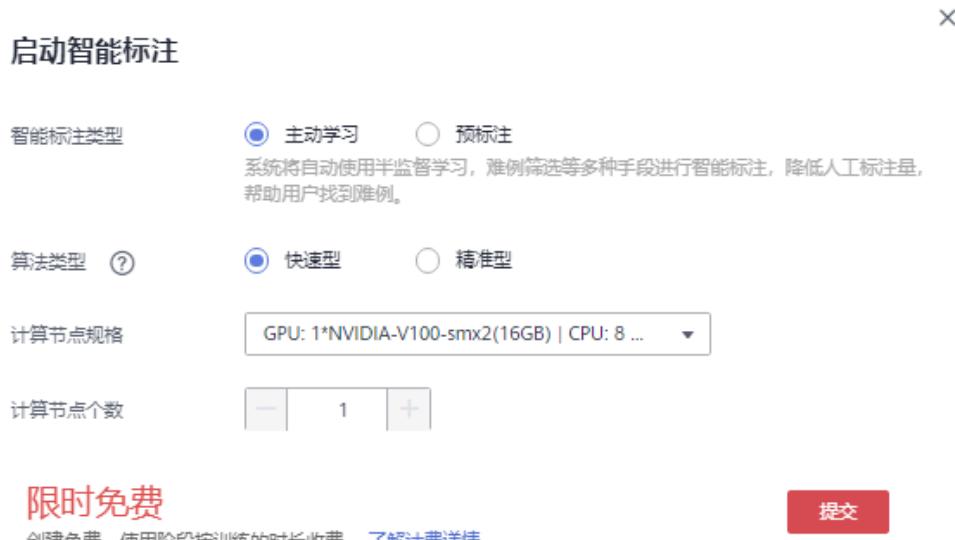


图 3-2 启动智能标注（物体检测）



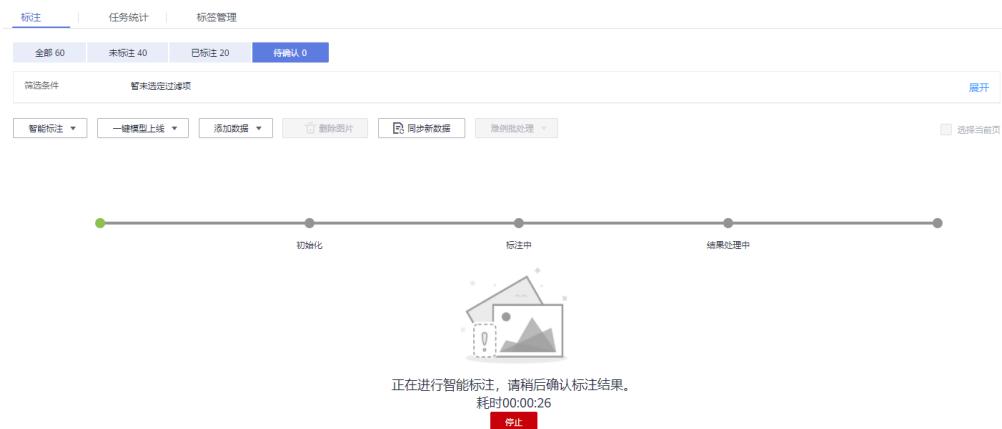
图 3-3 启动智能标注（预标注）



4. 完成参数设置后，单击“提交”，即可启动智能标注。
5. 在标注作业列表中，单击标注作业名称进入“标注作业详情”页。
6. 在“数据集概览页标注作业详情页”，选择“标注”页签，单击“待确认”页签，即可查看智能标注进度。

您也可以在该页签，“启动智能标注”或者查看“智能标注历史”

图 3-4 标注进度



说明

当系统中智能标注任务过多时，因免费资源有限，可能会出现排队的情况，导致作业一直处于“标注中”的状态。请您耐心等待，为确保您的标注作业能顺利进行，建议您避开高峰期使用。

7. 智能标注完成后，“待确认”页面将呈现所有标注后的图片列表。

- 图像分类标注作业

在“待确认”页面查看标签是否准确，勾选标注准确的图片，然后单击“确认”完成智能标注结果的确认。确认完成后的图片将被归类至“已标注”页面下。

针对标为“难例”的图片，您可以根据实际情况判断，手工修正标签。详细操作及示例请参见[•针对“图像分类”数据集](#)。

- 物体检测标注作业

在“待确认”页面，单击图片查看标注详情，查看标签及目标框是否准确，针对标注准确的图片单击“确认标注”完成智能标注结果的确认。确认完成后的图片将被归类至“已标注”页面下。

针对标为“难例”的图片，您可以根据实际情况判断，手工修正标签或目标框。详细操作及示例请参见[•针对“物体检测”数据集](#)。

相关问题

- **智能标注失败，如何处理？**

当前智能标注为免费使用阶段，当系统的标注任务过多时，因免费资源有限，导致任务失败，请您重新创建智能标注任务或建议您避开高峰期使用。

- **智能标注时间过长，如何处理？**

当前智能标注为免费使用阶段，当系统的标注任务过多时，因免费资源有限，需要排队，您的标注任务会长时间处于“标注中”状态，请您耐心等待。或建议您避开高峰期使用。

3.2 难例确认

在数据量很大的标注任务中，标注初期由于已标注图片不足，智能标注的结果无法直接用于训练。若对所有的未标注数据——进行调整确认仍然需要较大的人力和时间成本。为了更快地完成标注任务，在对未标注数据进行智能标注的任务中，ModelArts嵌入了自动难例发现功能。该功能会对剩余未标注图片的标注优先级给出建议。因为标注优先级高的图片的智能标注结果未达到预期，所以称之为难例。

ModelArts平台提供的自动难例发现功能，在智能标注以及数据采集筛选过程中，将自动标注出难例，建议对难例数据进一步确认标注，然后将其加入训练数据集中，使用此数据集训练模型，可得到精度更高的模型。首先，针对智能标注和采集筛选任务，难例的发现操作是系统自动执行的，无需人工介入，仅需针对标注后的数据进行确认和修改即可，提升数据管理和标注效率。其次，您可以基于难例的情况，补充类似数据，提升数据集的丰富性，进一步提升模型训练的精度。

在数据集管理中，对难例的管理有如下场景。

- [智能标注后，确认难例](#)
- [将数据集中的数据标注为难例](#)

说明

目前只有“图像分类”和“物体检测”类型的数据集支持难例发现功能。

智能标注后，确认难例

“智能标注”任务执行过程中，ModelArts将自动识别难例，并完成标注。当智能标注结束后，难例标注结果将呈现在“待确认”页签，建议您对难例数据进行人工修正，然后确认标注。

1. 登录ModelArts管理控制台，在左侧菜单栏中选择“数据管理>数据标注”，单击“我创建的”页签可查看所有作业列表。

2. 在标注作业列表中，选择“物体检测”或“图像分类”类型的标注作业，单击标注作业名称进入“标注作业详情”。
3. 在“标注作业详情页”，选择“待确认”页签，查看并确认难例。

说明

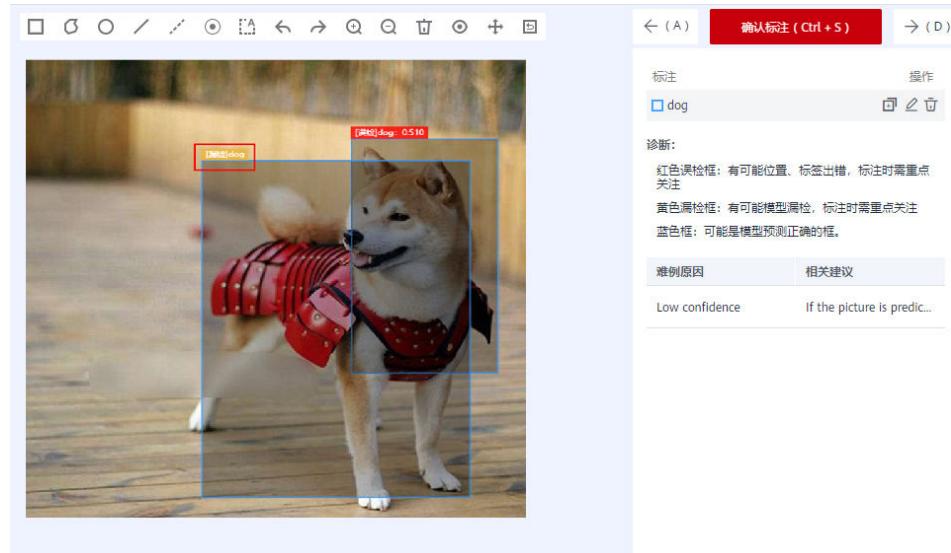
只有当智能标注任务完成后，待确认页签才会显示标注数据。否则，此页签内容为空。智能标注操作请参见[创建智能标注作业](#)。

- 针对“物体检测”标注作业

在“待确认”页签中，单击图片展开标注详情，查看图片数据的标注情况，如标签是否准确、目标框位置添加是否准确。如果智能标注结果不准确，建议手工调整标签或目标框，然后单击“确认标注”。完成确认后，重新标注的数据将呈现在“已标注”页签下。

如图3-5所示的难例，dog标签的目标框位置不准确，使用标注框重新标注，如图中的“漏检”目标框，然后需要将原先标注错误的目标框删除，即“误检”标签框。手工调整后，单击“确认标注”完成难例确认。

图 3-5 物体检测的难例确认



- 针对“图像分类”标注作业

在“待确认”页签中，查看标注难例的图片，其添加的标签是否准确。勾选标注不准确的图片，删除错误标签，然后在右侧“标签名”处添加准确标签。单击“确认”，勾选的图片及其标注情况，将呈现在“已标注”页签下。

选中的图片为标注错误图片，在右侧删除错误标签，然后在标签名处添加“狗”的标签，然后单击“确认”，完成难例确认。

将数据集中的数据标注为难例

针对标注作业中，已标注或未标注数据，也可以将图片数据标注为难例。标注为难例的数据，对后续模型训练中，通过内置规则提升模型精度。

1. 登录ModelArts管理控制台，在左侧菜单栏中选择“数据管理>数据标注”，单击“我创建的”页签可查看所有作业列表。
2. 在标注作业列表中，选择“物体检测”或“图像分类”类型的标注作业，单击数据集名称进入“标注作业详情页”。

3. 在标注作业详情页，选择“标注”页签，单击“已标注”、“未标注”或“全部”页签，勾选需标注为难例的图片，然后单击“难例批处理 > 确认为难例”。完成标注后，图片预览时，其右上角将显示为“难例”。

图 3-6 确认为难例



4 团队标注

4.1 团队标注简介

数据标注任务中，一般由一个人完成，但是针对数据集较大时，需要多人协助完成。ModelArts提供了团队标注功能，可以由多人组成一个标注团队，针对同一个数据集进行标注管理。

说明

团队标注功能当前仅支持“图像分类”、“物体检测”、“文本分类”、“命名实体”、“文本三元组”、“语音分割”类型的数据集。

针对启用团队标注功能的数据标注任务，支持创建团队标注任务，将标注任务指派给不同的团队，由多人完成标注任务。同时，在成员进行数据标注过程中，支持发起验收、继续验收以及查看验收报告等功能。

团队标注功能是以团队为单位进行管理，数据集启用团队标注功能时，必须指定一个团队。一个团队可以添加多个成员。

- 一个帐号最多可添加10个团队。
- 如果数据集需要启用团队标注功能，当前帐号至少拥有一个团队。如果没有，请执行[添加团队](#)操作添加。

4.2 创建和管理团队

4.2.1 管理团队

团队标注功能是以团队为单位进行管理，数据集启用团队标注功能时，必须指定一个团队。一个团队可以添加多个成员。

背景说明

- 一个帐号最多可添加10个团队。
- 如果数据集需要启用团队标注功能，当前帐号至少拥有一个团队。如果没有，请执行[添加团队](#)操作添加。

添加团队

1. 在ModelArts管理控制台左侧导航栏中，选择“数据管理>标注团队”，进入“标注团队”管理页面。
2. 在“标注团队”管理页面，单击“添加团队”。
3. 在弹出的“添加团队”对话框中，填写团队“名称”和“描述”，然后单击“确定”。完成标注团队的添加。

图 4-1 添加团队

添加团队

名称

描述

0/256

确定 取消

团队添加完成后，“标注团队”管理页面呈现新添加的团队，在页面右侧区域，可以查看团队详情。新添加的团队，其成员列表为空，请参考[添加成员](#)操作，为您的团队添加成员。

删除团队

当已有的团队不再使用，您可以执行删除操作。

在“标注团队”管理页面中，选中需删除的团队，然后单击“删除”。在弹出的对话框中，确认信息无误后，单击“确定”完成团队删除。

图 4-2 删除团队



4.2.2 管理成员

新添加的团队，其成员列表为空。您需要根据实际情况添加即将参与标注任务的成员信息。

一个团队最多支持添加100个成员，当超过100时，建议分为多个团队进行管理。

添加成员

1. 在ModelArts管理控制台左侧导航栏中，选择“数据管理>标注团队”，进入“标注团队”管理页面。
2. 在“标注团队”管理页面，从左侧团队列表中选择一个团队，单击团队，其右侧区域将呈现“团队详情”。
3. 在“团队详情”区域，单击“添加成员”。
4. 邮箱作为团队管理中的唯一标识，不同成员不能使用同一个邮箱。您填写的邮箱地址将被记录并保存在ModelArts中，仅用于ModelArts团队标注功能，当成员删除后，其填写的邮箱信息也将被一并删除。

其中，“角色”支持“Labeler”、“Reviewer”和“Team Manager”，“Team Manager”只能设置为一个人。

需要注意的是：目前不支持从标注任务中删除labeler。labeler的标注必须通过审核后，才能同步到最终结果，不支持单独分离操作。

成员添加完成后，团队详情区域中将呈现此成员的信息。

修改成员信息

团队中的成员，当其信息发生变化时，可以编辑其基本情况。

1. 在“团队详情”区域，选择需修改的成员。
2. 在成员所在行的“操作”列，单击“修改”。在弹出的对话框中，修改其“描述”或“角色”。

成员的“邮箱”无法修改，如果需要修改邮箱地址，建议先删除此成员，然后再基于新的邮箱地址添加新成员。

“角色”支持“Labeler”、“Reviewer”和“Team Manager”，“Team Manager”只能设置为一个人。

删除成员

- **删除单个成员**

在“团队详情”区域，选择需要删除的成员，单击“操作”列的“删除”。在弹出的对话框中，确认信息无误后，单击“确定”完成删除操作。

- **批量删除**

在“团队详情”区域，勾选需删除的成员，然后单击“删除”。在弹出的对话框中，确认信息无误后，单击“确定”完成多个成员的删除操作。

图 4-3 批量删除



4.3 创建团队标注任务

如果您在创建标注作业时，即启用团队标注，且指派了某一团队负责标注，系统将默认基于此团队创建一个标注任务。您可以在创建数据标注任务后，在“我创建的”页面查看此任务。

您还可以重新创建一个团队标注任务，指派给同一团队的不同成员，或者指派给其他标注团队。

团队标注作业的创建方式

- 从控制台的“数据管理 > 数据标注”页面进入，创建标注作业时，打开“启用团队标注”开关，同时指定一个标注团队，或者指定标注管理员。

图 4-4 创建团队标注作业

启用团队标注

* 名称: task-8461

类型: 指定标注团队 指定标注管理员

选择标注团队: team-7b84 ! 请至少选中一个labeler

将数据集的未标注文件立即分配给指定的人力进行标注和审核。团队成员收到系统发送的邮件后，按邮件提示进行标注和审核。

成员名称	角色	创建时间
member03@xxx.com	Labeler	--
member02@xxx.com	Labeler	--
member01@xxx.com	Labeler	--
member2@huawei.com	Labeler	--
member01@huawei.com	Labeler	--

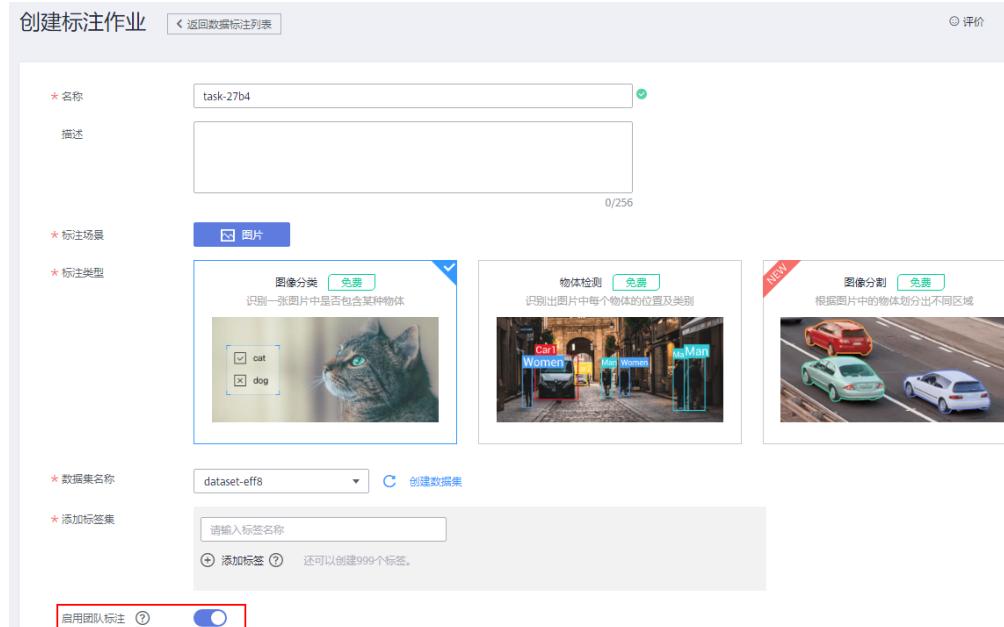
5 总条数: 8 < 1 2 >

自动将新增文件同步给标注团队。
勾选后，数据集中新增的文件会自动动态同步到已启动的团队标注任务中。

团队标注的文件自动加载智能标注结果。
勾选后，本次团队标注任务中的文件将含有智能标注的标注结果，标注员可直接确认或修改。

- 从控制台的“数据管理 > 数据集”进入数据集页面，在需要进行团队标注的数据集的操作列，单击“标注”，进入创建标注作业页面，打开“启用团队标注”开关。对于同一个数据集，可以创建多个团队标注任务。

图 4-5 打开启用团队标注



说明

- 只有当创建团队标注任务时，标注人员才会收到邮件。创建标注团队及添加标注团队的成员并不会发送邮件。此外，当所有样本都是已标注状态时，创建团队标注任务也不会收到邮件。
- 标注任务创建完成后，会将所有未标注状态的样本分配给标注人员。分配采用随机均分的策略，不支持重复分配。

创建团队标注任务

同一个数据集，支持创建多个团队标注作业，指派给同一团队的不同成员，或者指派给其他标注团队。

- 登录ModelArts管理控制台，在左侧菜单栏中选择“数据管理 > 数据集”，打开数据集列表。
- 在数据集列表中，选择支持团队标注的数据集，单击数据集名称进入数据集概览页。
- 在数据集概览页页面，右侧的“标注任务”区域，可查看此数据集已有的标注任务。单击“新建标注任务”开始创建新任务。

图 4-6 标注任务

任务名称	标注进度	样本数量	描述
task-cdf0	0/41	41	--
task-3752	3/41	41	--

或者也可以从“数据管理 > 数据标注”页面进入，单击“创建标注作业”进入创建标注作业页面。

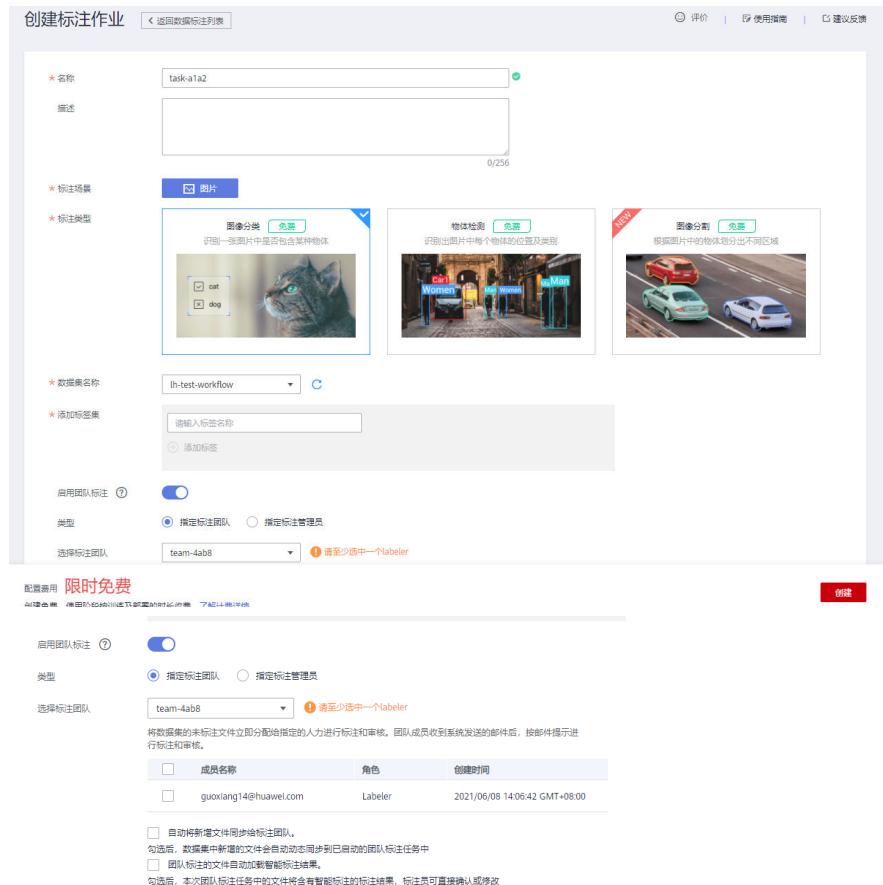
4. 在弹出的“创建标注作业”页面中，填写相关参数，然后单击“确定”，完成任务创建。
 - “名称”：设置此任务的名称。
 - “标注场景”：选择标注作业的任务类型。
 - “标签集”：展示当前数据集已有的标签及标签属性。
 - “启用团队标注”：选择打开，并配置如下团队标注相关参数。
 - “类型”：设置任务类型，支持“指定标注团队”或“指定标注管理员”。
 - “选择标注团队”：任务类型设置为“指定标注团队”，需在此参数中指定一个团队，同时勾选此团队中某几个成员负责标注。下拉框中将罗列当前帐号下创建的标注团队及其成员。
 - “选择标注接口人”：任务类型设置为“指定标注管理员”，需在所有团队的“Team Manager”中选择一人作为管理员。
 - “自动将新增图片同步给标注团队”：根据需要选择是否将任务中新增的数据自动同步给标注人员。
 - “团队标注的图片自动加载智能标注结果”：根据需要选择是否将任务中智能标注待确认的结果自动同步给标注人员。

说明

团队标注加载智能标注结果的处理步骤：

- 如果类型选择“指定标注团队”，需要先创建团队标注任务，然后执行智能标注任务。
- 如果类型选择“指定标注管理员”，在“我参与的”页签下选择团队标注任务，单击“分配任务”。

图 4-7 创建团队标注任务



任务创建完成后，您可以在“我创建的”页签下看到新建的任务。

4.4 登录 ModelArts-Console

在ModelArts中，一般用户使用数据标注功能，直接是在“数据管理”模块操作，此模块包含数据管理所有能力，数据集管理、数据标注、数据导入导出、智能标注、团队标注和管理等。团队标注任务创建成功后，团队成员登录ModelArts-Console查看相关任务。

1. 团队标注任务创建成功后，团队成员收到标注任务的邮件。

图 4-8 任务邮件



- 单击任务邮件中的标注任务地址, 跳转至ModelArts数据管理>数据标注页面的“我参与的”页签。

图 4-9 登录 ModelArts-Console



- 在“我参与的”页签下, 可查看您的标注任务。

图 4-10 标注任务

数据标注							评价	使用指南
我创建的		我参与的		操作			筛选器	
任务名称	标注类型	样本数	未标注/待确认/已驳回/待审核/审核通过	创建时间	创建人	角色	操作	
task-5d9e	图像分类	4	4/0/0/0/0	2023/01/05 17:36:27 GMT+08:00	模型管理员	Labeler	<button>标注</button>	

数据标注访问地址可以查看如下表格获取，另外，如果团队成员绑定了邮箱，可以收到任务通知邮件，成员也可以通过邮件中给出的地址访问ModelArts-Console标注地址。

表 4-1 ModelArts 标注地址

局点	ModelArts-Console访问地址
中国-香港	https://console-intl.huaweicloud.com/modelarts/?region=ap-southeast-1&locale=zh-cn#/dataLabel?tabActive=labelConsole
亚太-新加坡	https://console-intl.huaweicloud.com/modelarts/?region=ap-southeast-3&locale=zh-cn#/dataLabel?tabActive=labelConsole
亚太-曼谷	https://console-intl.huaweicloud.com/modelarts/?region=ap-southeast-2&locale=zh-cn#/dataLabel?tabActive=labelConsole

登录后，仅显示当前用户（此邮箱用户）相关的团队标注任务及其相关数据。

4.5 启动团队标注任务

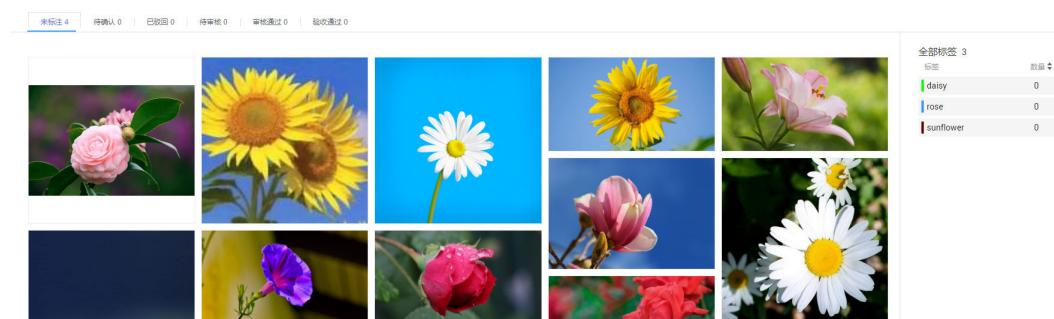
登录到console标注页面后在“我参与的”页签下，可查看到分配的标注任务，单击任务名称，可进入标注页面。不同类型的标注作业，标注方式不同，详细请参见：

- [图像分类](#)
- [物体检测](#)
- [文本分类](#)
- [命名实体](#)
- [文本三元组](#)
- [语音分割](#)

在标注页面中，每个成员可查看“未标注”、“待确认”、“已驳回”、“待审核”、“审核通过”、“验收通过”的图片信息。请及时关注管理员驳回以及待修正的图片。

当团队标注任务中，分配了Reviewer角色，则需要对标注结果进行审核，审核完成后，再提交给管理员验收。

图 4-11 成员标注平台



4.6 审核团队标注任务结果

团队标注成员完成后，团队审核者可以对标注结果进行审核。

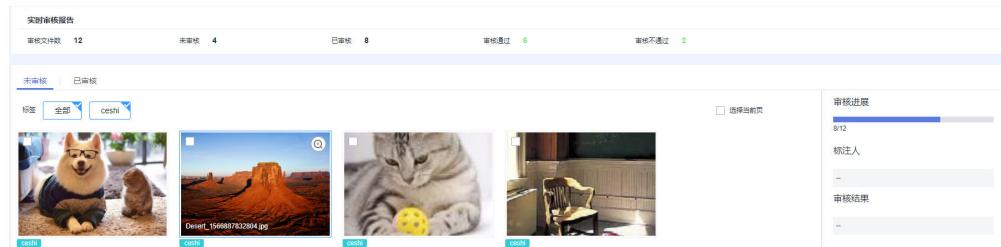
1. 登录ModelArts管理控制台，左侧菜单栏选择“数据管理>数据标注”，在数据标注页面选择“我参与的”，在任务列表“操作”列单击“审核”，发起审核。

图 4-12 发起审核



2. 在审核页面中，审核人员可以查看“未审核”、“已审核”、“审核通过”、“审核不通过”的样本。

图 4-13 标注结果审核



3. 审核人员可以在审核页面的右侧选择“审核结果”（“通过”或“不通过”）。当选择审核结果为“通过”时，需设置“验收评分”（分“A”、“B”、“C”、“D”四个选项，“A”表示最高分），如图4-14所示。当选择审核结果为“不通过”时，可以在文本框中写明驳回原因，如图4-15所示。

图 4-14 审核通过



图 4-15 审核不通过



4.7 验收团队标注任务结果

任务验收（管理员）

- **发起验收**

当团队的成员已完成数据标注，标注作业的创建者可发起验收，对标注结果进行抽验。只有当标注成员存在标注完成的数据时，才可以发起验收，否则发起验收按钮为灰色。

- a. 登录ModelArts管理控制台，在左侧菜单栏中选择“数据管理 > 数据标注”，打开数据标注管理页。
- b. 选择“我参与的”，选择团队标注作业，单击作业名称进入“标注作业详情页”，单击右上角“团队标注>验收”，发起验收。

图 4-16 发起验收



- c. 在弹出的对话框中，设置“抽样策略”，可设置为“按百分比”，也可以设置为“按数量”。设置好参数值后，单击“确定”启动验收。
- “按百分比”：按待验收图片总数的一定比例进行抽样验收。
- “按数量”：按一定数量进行抽样验收。

图 4-17 发起验收



- d. 验收启动后，界面将展示实时验收报告，您可以在右侧选择“验收结果”（“通过”或“不通过”）。当选择验收结果为“通过”时，需设置“验收评分”（分“A”、“B”、“C”、“D”四个选项，“A”表示最高分），如图4-19所示。当选择验收结果为“不通过”时，可以在文本框中写明驳回原因，如图4-20所示。

图 4-18 查看实时验收报告



图 4-19 设置验收结果为“通过”

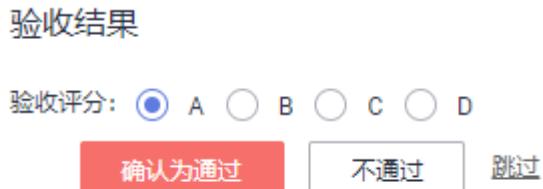


图 4-20 设置验收结果为“不通过”

验收结果



- 继续验收

针对未完成验收的任务，可以继续验收。针对未发起过验收流程的任务，不支持“继续验收”，按钮为灰色。

在“任务统计>标注进展”页签中，针对需继续验收的任务，单击“继续验收”。系统直接进入“实时验收报告”页面，您可以继续验收未验收的图片，设置其“验收结果”。

- 完成验收

继续验收完成后，单击右上角“完成验收”在完成验收窗口，您可以查看本标注作业的验收情况，如抽样文件数等，同时设置如下参数，然后进行验收。只有完成验收，标注信息才会同步到标注作业的已标注页面中。

一旦标注数据完成验收，团队成员无法再修改标注信息，只有数据集创建者可修改。

表 4-2 完成验收的参数设置

参数	说明
对已标注数据修改	<ul style="list-style-type: none">不覆盖：针对同一个数据，不使用当前团队标注的结果覆盖已有数据。覆盖：针对同一个数据，使用当前团队标注的结果覆盖已有数据。覆盖后无法恢复，请谨慎操作。

参数	说明
验收范围	<ul style="list-style-type: none">全部通过：被驳回的样本，也会通过。全部驳回：已经通过的样本，需要重新标注，下次验收时重新进行审核。剩余全部通过：已经驳回的会驳回，其余会自动验收通过。剩余全部驳回：样本抽中的通过的，不需要标注了，未通过和样本未抽中的需要重新标注验收。

图 4-21 完成验收**完成验收**

验收通过后，标注结果才会同步到数据集已标注

**查看验收报告**

针对进行中或已完成的标注任务，都可以查看其验收报告。登录管理控制台，选择“数据管理>数据标注”，在数据标注页选择“我创建的”，并单击某条团队标注的任务名称，进入标注详情页。在右上角单击“验收报告”，即可在弹出的“验收报告”对话框中查看详情。

图 4-22 查看验收报告

验收报告

进行中验收统计信息

验收通过率	-%	抽样文件数	--	未验收	--
已验收	--	验收通过	--	验收不通过	--

已完成验收统计信息

验收通过率	-%	抽样文件数	--	未验收	--
已验收	--	验收通过	--	验收不通过	--

确定

删除标注任务

验收结束后，针对不再使用的标注任务，您可单击任务所在行的删除。任务删除后，未验收的标注详情将丢失，请谨慎操作。但是数据集中的原始数据以及完成验收的标注数据仍然存储在对应的OBS桶中。