ModelArts

Data Labeling

Issue   01
Date   2025-01-06

HUAWEI CLOUD COMPUTING TECHNOLOGIES CO., LTD.

Huawei Cloud Computing Technologies Co., Ltd.

Address: Huawei Cloud Data Center Jiaoxinggong Road
             Qianzhong Avenue
             Gui'an New District
             Gui Zhou 550029
             People's Republic of China

Website: https://www.huaweicloud.com/intl/en-us/

Contents

1 Introduction to Data Labeling

⬡ **NOTE**

Data labeling is supported only in the following regions: CN North-Beijing4, CN North-Beijing1, CN East-Shanghai1, CN South-Guangzhou, CN Southwest-Guiyang1, CN-Hong Kong, AP-Singapore, AP-Bangkok, AP-Jakarta, LA-Santiago, LA-Sao Paulo1, and LA-Mexico City2.

Model training requires a large amount of labeled data. Therefore, before training a model, label data. ModelArts provides you with the following labeling functions:

- **Manual Labeling**: allows you to manually label data.
- **Auto Labeling**: allows you to automatically label remaining data after a small amount of data is manually labeled.
- **Team Labeling**: allows you to perform collaborative labeling for a large amount of data.

Manual Labeling

Create a labeling job based on the dataset type. ModelArts supports the following types of labeling jobs:

- Images
  - Image classification: identifies a class of objects in images.
  - Object detection: identifies the position and class of each object in an image.
  - Image segmentation: segments an image into different areas based on objects in the image.
- Audio
  - Sound classification: classifies and identifies different sounds.
  - Speech labeling: labels speech content.
  - Speech paragraph labeling: segments and labels speech content.
- Text
  - Text classification: assigns labels to text according to its content.
  - Named entity recognition: assigns labels to named entities in text, such as time and locations.

- – Text triplet: assigns labels to entity segments and entity relationships in the text.
- Video

  Video labeling: identifies the position and class of each object in a video. Only the MP4 format is supported.

Auto Labeling

In addition to manual labeling, ModelArts also provides the auto labeling function to quickly label data, reducing the labeling time by more than 70%. Auto labeling means learning and training are performed based on the labeled images and an existing model is used to quickly label the remaining images.

Only datasets of image classification and object detection types support the auto labeling function.

Team Labeling

Generally, a small data labeling task can be completed by an individual. However, team work is required to label a large dataset. ModelArts provides the team labeling function. A labeling team can be formed to manage labeling for the same dataset.

The team labeling function supports only datasets for image classification, object detection, text classification, named entity recognition, text triplet, and speech paragraph labeling.

Dataset Functions

Dataset functions vary depending on dataset types. For details, see **Table 1-1**.

Table 1-1 Functions supported by different types of datasets

Dataset Type	Labeling Type	Manual Labeling	Auto Labeling	Team Labeling
Images	Image classification	Supported	Supported	Supported
	Object detection	Supported	Supported	Supported
	Image segmentation	Supported	N/A	N/A
Audio	Sound classification	Supported	N/A	N/A
	Speech labeling	Supported	N/A	N/A

Dataset Type	Labeling Type	Manual Labeling	Auto Labeling	Team Labeling
	Speech paragraph labeling	Supported	N/A	Supported
Text	Text classification	Supported	N/A	Supported
	Named entity recognition	Supported	N/A	Supported
	Text triplet	Supported	N/A	Supported
Video	Video labeling	Supported	N/A	N/A
Free format	N/A	N/A	N/A	N/A
Table	N/A	N/A	N/A	N/A

2 Manual Labeling

2.1 Creating a Labeling Job

Model training requires a large amount of labeled data. Therefore, before training a model, label data. You can create a manual labeling job labeled by one person or by a group of persons (team labeling), or enable auto labeling to quickly label images. You can also modify existing labels, or delete them and re-label.

Labeling Job Types

Create a labeling job based on the dataset type. ModelArts supports the following types of labeling jobs:

- Images
  - Image classification: identifies a class of objects in images.
  - Object detection: identifies the position and class of each object in an image.
  - Image segmentation: segments an image into different areas based on objects in the image.
- Audio
  - Sound classification: classifies and identifies different sounds.
  - Speech labeling: labels speech content.
  - Speech paragraph labeling: segments and labels speech content.
- Text
  - Text classification: assigns labels to text according to its content.
  - Named entity recognition: assigns labels to named entities in text, such as time and locations.
  - Text triplet: assigns labels to entity segments and entity relationships in the text.
- Videos

  Video labeling: identifies the position and class of each object in a video. Only the MP4 format is supported.

Prerequisites

Before labeling data, create a dataset.

Procedure

1. Log in to the ModelArts management console. In the navigation pane on the left, chooseData Preparation > **Label Data**.

   📖 **NOTE**

   Data labeling is supported only in the following regions: CN North-Beijing4, CN North-Beijing1, CN East-Shanghai1, CN South-Guangzhou, CN Southwest-Guiyang1, CN-Hong Kong, AP-Singapore, AP-Bangkok, AP-Jakarta, LA-Santiago, LA-Sao Paulo1, and LA-Mexico City2.
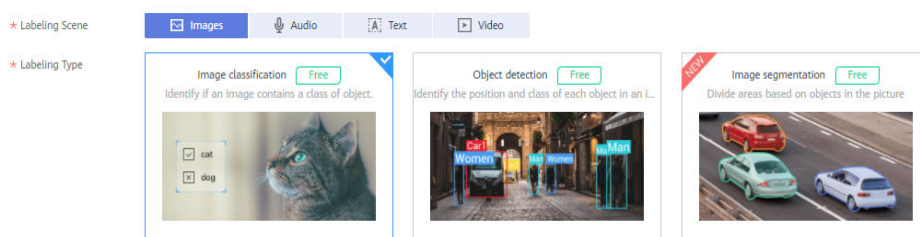
2. On the **Data Labeling** page, click **Create Labeling Job** in the upper right corner. On the page that is displayed, create a labeling job.

   a. Enter basic information about the labeling job, including **Name** and **Description**.

   Figure 2-1 Basic information about a labeling job

   b. Select a labeling scene and type as required.

   Figure 2-2 Selecting a labeling scene and type

   c. Set the parameters based on the labeling job type. For details, see the parameters of the following labeling job types:

   - **Images (Image Classification, Image Segmentation, and Object Detection)**

   - **Audio (Sound Classification, Speech Labeling, and Speech Paragraph Labeling)**

   - **Text (Text Classification, Named Entity Recognition, and Text Triplet)**

   - **Videos**

   d. Click **Create** in the lower right corner of the page.

After the labeling job is created, the data labeling management page is displayed. You can perform the following operations on the labeling job: start auto labeling, publish new versions, modify the labeling job, and delete the labeling job.

Images (Image Classification, Image Segmentation, and Object Detection)

Figure 2-3 Parameters of labeling jobs for image classification and object detection

Table 2-1 Parameters of an image labeling job

Parameter	Description
Dataset Name	Select a dataset that supports the labeling type.
Label Set	- Label name: Enter a label name with 1 to 1024 characters. **Add Label**: Click **Add Label** to add one or more labels. - Label color: Set label colors for object detection and image segmentation labeling jobs. Select a color from the color palette on the right of a label, or enter the hexadecimal color code to set the color. **Add Label Attribute**: For an object detection labeling job, you can click the plus sign (+) on the right to add label attributes after setting a label color. Label attributes are used to distinguish different attributes of the objects with the same label. For example, yellow kittens and black kittens have the same label **cat** and their label attribute is **color**.

Parameter	Description
Team Labeling	Enable or disable team labeling. Image segmentation does not support team labeling. Therefore, this parameter is unavailable when you use image segmentation.
	After enabling team labeling, enter the type of the team labeling job, and select the labeling team and team members. For details about the parameter settings, see **Creating a Team Labeling Job**.
	Before enabling team labeling, ensure that you have added a team and members on the **Labeling Teams** page. If no labeling team is available, click the link on the page to go to the **Labeling Teams** page, and add your team and members. For details, see **Adding a Team**.
	After a dataset is created with team labeling enabled, you can view the **Team Labeling** mark in **Labeling Type**.

Audio (Sound Classification, Speech Labeling, and Speech Paragraph Labeling)

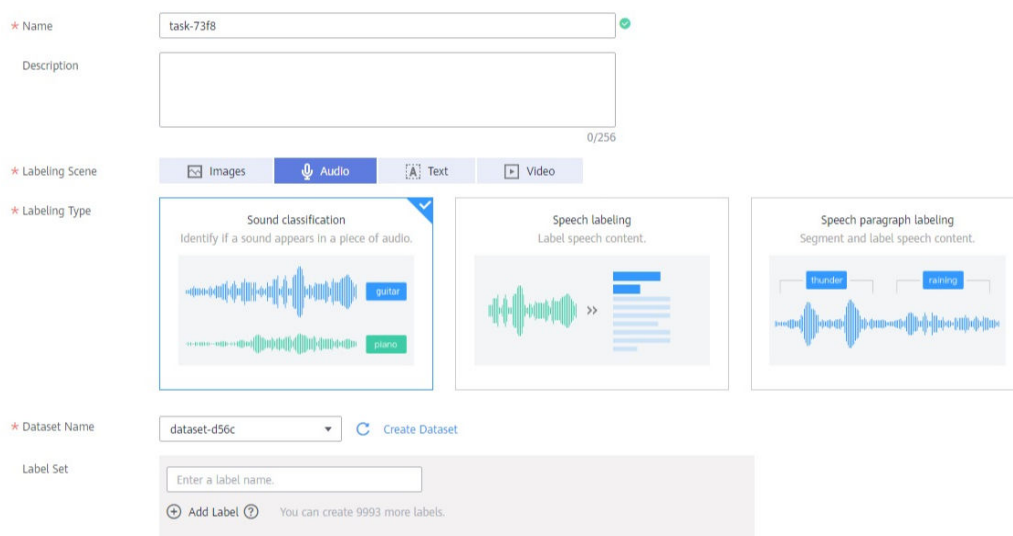**Figure 2-4** Parameters of labeling jobs for sound classification, speech labeling, and speech paragraph labeling

Table 2-2 Parameters of an audio labeling job

Parameter	Description
Dataset Name	Select a dataset that supports the labeling type.

Parameter	Description
Label Set (for sound classification)	You can add a label set for labeling jobs of sound classification. ● Label name: Enter 1 to 1024 characters in the **Label Set** text box. ● **Add Label**: Click **Add Label** to add one or more labels.

Label Management (for speech paragraph labeling)	Label management is available for speech paragraph labeling.
	● **Single Label**
	A single label is used to label a piece of audio that has only one class.
	  – **Label**: Enter a label name, with 1 to 1024 characters.
	  – **Label Color**: Set the label color in the **Label Color** column. You can select a color from the color palette or enter a hexadecimal color code to set the color.
	● **Multiple Labels**
	Multiple labels are suitable for multi-dimensional labeling. For example, you can label a piece of audio as both noise and speech. For speech, you can label the audio with different speakers. You can click **Add Label Class** to add multiple label classes. A label class can contain multiple labels. The label class or name contains 1 to 256 characters. Only letters, digits, periods (.), underscores (_), and hyphens (-) are allowed.
	  – **Add Label Class**: Enter a label class.
	  – **Label**: Enter a label name.
	  – **Add Label**: Click **Add Label** to add one or more labels.

Speech Labeling (for speech paragraph labeling)	Only datasets for speech paragraph labeling support speech labeling. By default, speech labeling is disabled. If this function is enabled, you can label speech content.
Team Labeling (for speech paragraph labeling)	Only datasets of speech paragraph labeling support team labeling. After enabling team labeling, enter the type of the team labeling job, and select the labeling team and team members. For details about the parameter settings, see **Creating a Team Labeling Job**. Before enabling team labeling, ensure that you have added a team and members on the **Labeling Teams** page. If no labeling team is available, click the link on the page to go to the **Labeling Teams** page, and add your team and members. For details, see **Adding a Team**. After a dataset is created with team labeling enabled, you can view the **Team Labeling** mark in **Labeling Type**.

Text (Text Classification, Named Entity Recognition, and Text Triplet)

Figure 2-5 Parameters of labeling jobs for text classification, named entity recognition, and text triplet
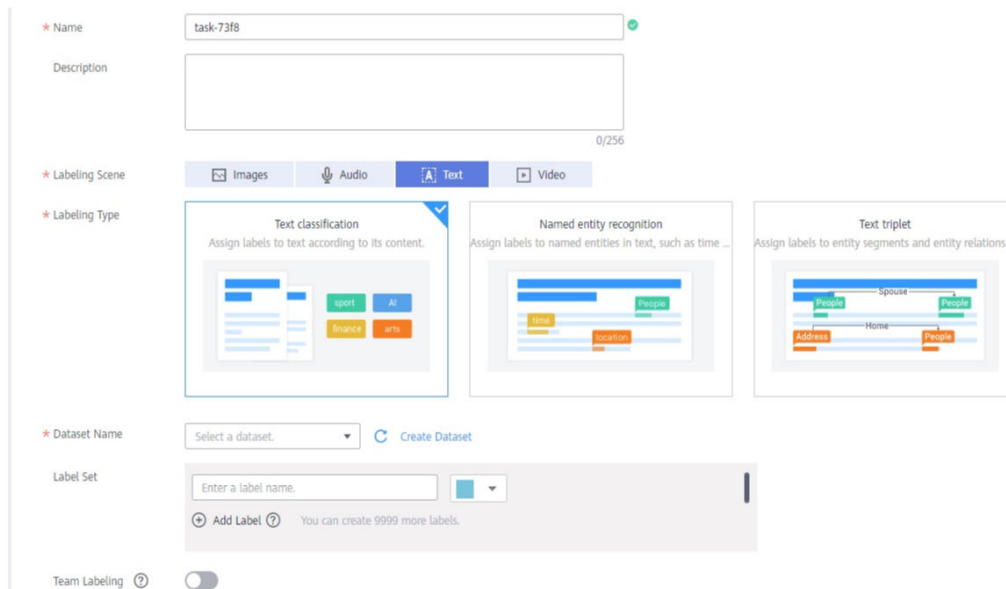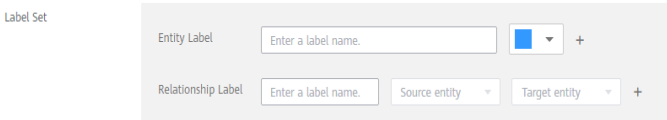

Table 2-3 Parameters of a text labeling job

Parameter	Description
Dataset Name	Select a dataset that supports the labeling type.
Label Set (for text classification and named entity recognition)	- Label name: Enter a label name, with 1 to 1024 characters. **Add Label**: Click **Add Label** to add one or more labels. - Label color: Select a color from the color palette or enter the hexadecimal color code to set the color. **Figure 2-6** Setting the label color

Parameter	Description
Label Set (for text triplet)	For datasets of the text triplet type, set entity labels and relationship labels. ● **Entity Label**: Set the label name and label color. You can click the plus sign (+) on the right of the color area to add multiple labels. ● **Relationship Label**: a relationship between two entities. Set the source entity and target entity. Therefore, add at least two entity labels before adding a relationship label. **Figure 2-7** Adding a label
Team Labeling	Enable or disable team labeling. After enabling team labeling, enter the type of the team labeling job, and select the labeling team and team members. For details about the parameter settings, see **Creating a Team Labeling Job**. Before enabling team labeling, ensure that you have added a team and members on the **Labeling Teams** page. If no labeling team is available, click the link on the page to go to the **Labeling Teams** page, and add your team and members. For details, see **Adding a Team**. After a dataset is created with team labeling enabled, you can view the **Team Labeling** mark in **Labeling Type**.

Videos

Figure 2-8 Parameters of a video labeling job

Table 2-4 Parameters of a video labeling job

Parameter	Description
Dataset Name	Select a dataset that supports the labeling type.
Label Set	• Label name: Enter a label name, with 1 to 1024 characters. • **Add Label**: Click **Add Label** to add one or more labels. • Label color: Select a color from the color palette or enter the hexadecimal color code to set the color.

2.2 Image Labeling

2.2.1 Image Classification

Training a model uses a large number of labeled images. Therefore, label images before the model training. You can add labels to images by manual labeling or auto labeling. In addition, you can modify the labels of images, or remove their labels and label the images again.

Before labeling an image in image classification scenarios, pay attention to the following:

- You can add multiple labels to an image.
- A label name can contain a maximum of 1024 characters, including letters, digits, hyphens (-), and underscores (_).

Starting Labeling

1. Log in to the ModelArts management console. In the navigation pane on the left, choose **Data Preparation** > **Label Data**.

2. On the right of the labeling job list, select a labeling type from the job type drop-down list. Click the job to be performed based on the labeling type. The details page of the job is displayed.

Figure 2-9 Selecting a labeling type
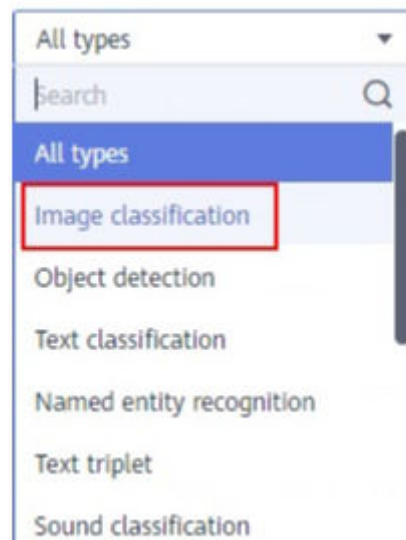
3. The job details page displays all data of the labeling job.

Synchronizing New Data

ModelArts automatically synchronizes data and labeling information from datasets to labeling jobs.

To quickly obtain the latest data in a dataset, on the **All statuses**, **Unlabeled**, or **Labeled** tab page of the labeling job details page, click **Synchronize New Data**.

📖 **NOTE**

Symptom:

After the labeled data is uploaded to OBS and synchronized, the data is displayed as unlabeled.

Possible causes:

Automatic encryption is enabled in the OBS bucket.

Solution:

Create an OBS bucket and upload data again, or disable bucket encryption and upload data again.

Filtering Data

On the **All statuses**, **Unlabeled**, or tab page, click ⌄ in the filter criteria area and add filter criteria to quickly filter the data you want to view.

The following filter criteria are available. You can set one or more filter criteria.

- **Example Type**: Select **Hard example** or **Non-hard example**.
- **Label**: Select **All** or one or more labels you specified.
- **File Name** or **Path**: Filter files by file name or file storage path.
- **Labeled By**: Select the name of the user who labeled the image.
- **Sample Attribute**: Select the attribute generated by auto grouping. This filter criterion can be used only after auto grouping is enabled.
- **Data Attribute**: Select **All** or **Inference** to filter the data source.

**Figure 2-10** Filter criteria
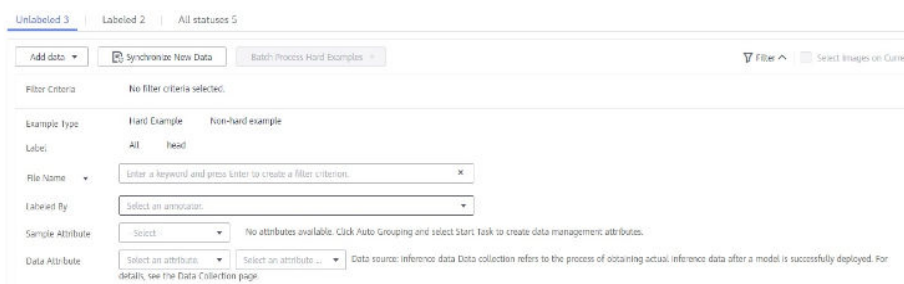
Manually Labeling Images

The labeling job details page displays the **All statuses**, **Unlabeled**, and **Labeled** tab pages. The **Unlabeled** tab page is displayed by default. Click an image to preview it. For the images that have been labeled, the label information is displayed at the bottom of the preview page.

1. On the **Unlabeled** tab page, select the images to be labeled.
   - Manual selection: In the image list, click the selection box in the upper left corner of an image to enter the selection mode, indicating that the image is selected. You can select multiple images of the same type and add labels to them together.

– Batch selection: If all the images on the current page of the image list belong to the same type, you can click **Select Images on Current Page** in the upper right corner to select all the images on the current page.

2. Add labels to the selected images.

a. In the label adding area on the right, set a label in the **Label** text box.

Click the **Label** text box and select an existing label from the drop-down list. If the existing labels cannot meet the requirements, input a label in the text box.

b. Click **OK**. The selected images are automatically moved to the **Labeled** tab page. On the **Unlabeled** and **All statuses** tab pages, the labeling information is updated along with the labeling process, including the added label names and the number of images for each label.

Figure 2-11 Adding a label
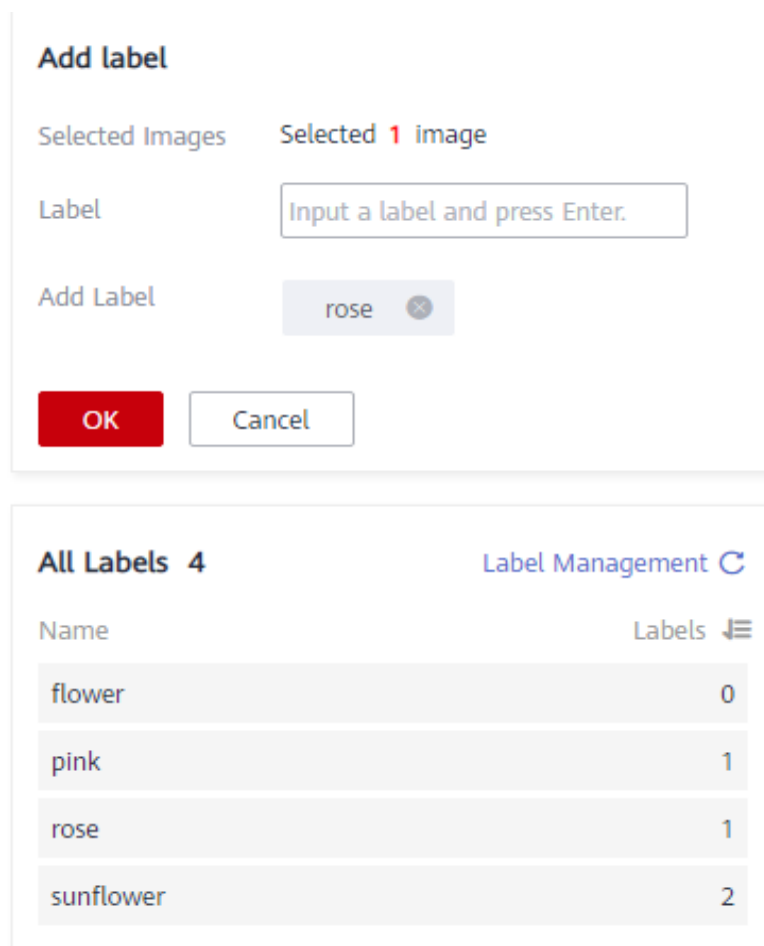
📖 **NOTE**

For details about how to label data, see **Labeling Description** on the dataset details page.

1. Log in to the ModelArts management console. In the navigation pane on the left, choose **Data Preparation** > **Label Data**. The **Data Labeling** page is displayed.

2. On the **My Creations** or **My Participations** tab page, find the dataset to be labeled.

3. Click the dataset name. The labeling details page is displayed. (By default, the **Unlabeled** tab page is displayed.)

4. In the upper right corner of the labeling details page, click **Labeling Description**.

Figure 2-12 Labeling Description

Viewing Labeled Images

On the labeling job details page, click the **Labeled** tab to view the list of labeled images. By default, the corresponding labels are displayed under the image thumbnails. You can also select an image and view the label information of the image in the **Labels of Selected Images** area on the right.

Modifying Labeled Data

After labeling data, you can modify labeled data on the **Labeled** tab page.

- **Modifying based on images**

  On the labeling job details page, click the **Labeled** tab, and select one or more images to be modified from the image list. Modify the image information in the label information area on the right.

  Modifying a label: In the **Labels of Selected Images** area, click the edit icon in the **Operation** column, enter the correct label name in the text box, and click the check mark to complete the modification.

  Deleting a label: In the **Labels of Selected Images** area, click the delete icon in the **Operation** column to delete the label. This operation deletes only the labels added to the selected image.

Figure 2-13 Modifying a label

- **Modifying based on labels**

  – On the labeling job details page, click **Label Management**. All labels are displayed on the list.

- Modifying a label: Click **Modify** in the **Operation** column. In the dialog box that is displayed, enter a new label name and click **OK**. After the modification, the images with the label added will use the new label name.

- Deleting a label: Click **Delete** in the **Operation** column to delete the label from all images with the label added.

Figure 2-14 Label management

Figure 2-15 All labels

- Click **Label** in the **Operation** column of the target labeling job to go to the label management page.

    - Click **Modify** in the **Operation** column of the target label to modify it.

    - Click **Delete** in the **Operation** column of the target label to delete it.

Adding Data

In addition to the data automatically synchronized from datasets, you can directly add images to labeling jobs for labeling. The added data is first imported to the dataset associated with the labeling job. Then, the labeling job automatically synchronizes the latest data from the dataset.
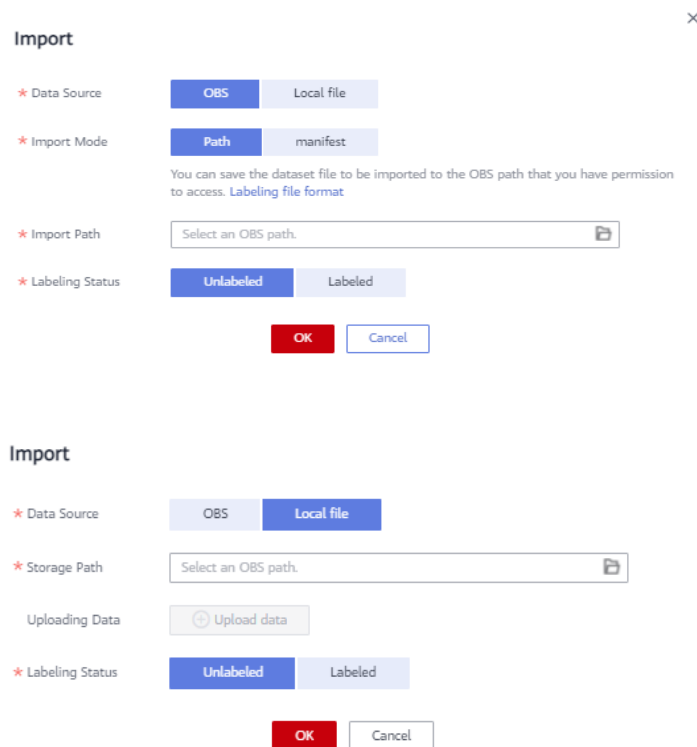
1. On the labeling job details page, click **All statuses**, **Labeled**, or **Unlabeled** tab, click **Add data** in the upper left corner.

Figure 2-16 Adding data

2. Configure the data source, import mode, import path, and labeling status.

    Figure 2-17 Adding images

3. Click **OK**.

    The images you have added will be automatically displayed in the image list on the **All statuses** tab page. You can choose **Add data** > **View historical records** to view task history.

    Figure 2-18 Viewing historical data

Deleting Images

You can quickly delete the images you want to discard.

On the **All statuses**, **Unlabeled**, or **Labeled** tab page, select the images to be deleted or click **Select Images on Current Page**, and click **Delete**. In the displayed

dialog box, select or deselect **Delete the source files from OBS** as required. After confirmation, click **Yes** to delete the images.

Figure 2-19 Deleting Images

If a tick is displayed in the upper left corner of an image, the image is selected. If no image is selected on the page, the **Delete** button is unavailable.

☐ NOTE

If you select **Delete the source files from OBS**, images stored in the OBS directory will be deleted accordingly. This operation may affect other dataset versions or datasets using those files, for example, leading to an error in page display, training, or inference. Deleted data cannot be recovered. Exercise caution when performing this operation.

Managing Annotators

If team labeling is enabled for a labeling job, view its labeling details on the **Annotator Management** tab page. Additionally, you can add, modify, or delete annotators.
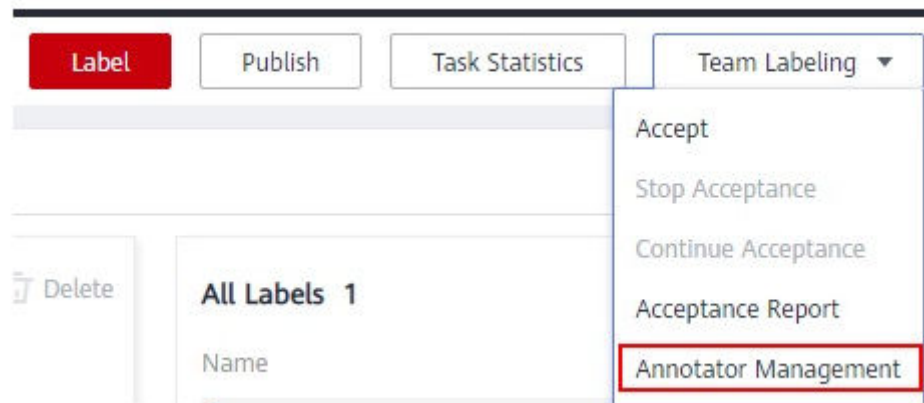
1. Choose **Data Preparation** > **Label Data**. On the **My Creations** tab page, view the list of all labeling jobs.

2. Locate the row that contains the target team labeling job. (The name of a team labeling job is followed by ⚲.)

3. Choose **More** > **Annotator Management** in the **Operation** column. Alternatively, click the job name to go to the job details page, and choose **Team Labeling** > **Annotator Management** in the upper right corner.

Figure 2-20 Annotator management (1)

Figure 2-21 Annotator management (2)

- Adding an annotator

  Click **Add Member**, select a member name, and click **OK**.

  Click **Send Email** in the **Operation** column to send the labeling job to the annotator by email.

- Modifying annotator information

  Click **Modify** in the **Operation** column to modify the role of the annotator.

- Deleting an annotator

  Click **Delete** in the **Operation** column to delete the annotator.

2.2.2 Object Detection

Training a model uses a large number of labeled images. Therefore, label images before the model training. You can add labels to images by manual labeling or auto labeling. In addition, you can modify the labels of images, or remove their labels and label the images again.
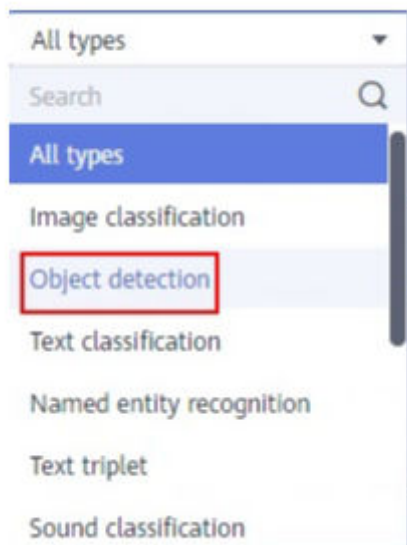
Before labeling an image in object detection scenarios, pay attention to the following:

- All target objects in the image must be labeled.
- Target objects are clear without any blocking and contained within bounding boxes.
- Only the entire object must be contained within a bounding box. The bounding box contains the entire object. The edge of the bounding box cannot intersect the edge outline of the object to be labeled. Ensure that there is no gap between the edge and the object to be labeled to prevent the background from interfering with the model training.

Starting Labeling

1. Log in to the ModelArts management console. In the navigation pane on the left, choose **Data Preparation** > **Label Data**.

2. In the labeling job list, select a labeling type from the **All type** drop-down list, click the job to be performed based on the labeling type. The details page of the job is displayed.

Figure 2-22 Selecting a labeling type

3. The job details page displays all data of the labeling job.

Synchronizing New Data

ModelArts automatically synchronizes data and labeling information from datasets to the labeling job.

To quickly obtain the latest data in a dataset, on the **All statuses**, **Unlabeled**, or **Labeled** tab page of the labeling job details page, click **Synchronize New Data**.

📖 **NOTE**

Symptom:

After the labeled data is uploaded to OBS and synchronized, the data is displayed as unlabeled.

Possible causes:

Automatic encryption is enabled in the OBS bucket.

Solution:

Create an OBS bucket and upload data again, or disable bucket encryption and upload data again.

Filtering Data

On the **All statuses**, **Unlabeled**, or tab page, click ⌄ in the filter criteria area and add filter criteria to quickly filter the data you want to view.

The following filter criteria are available. You can set one or more filter criteria.

● **Example Type**: Select **Hard example** or **Non-hard example**.

● **Label**: Select **All** or one or more labels you specified.

● **File Name** or **Path**: Filter files by file name or file storage path.

● **Labeled By**: Select the name of the user who labeled the image.

- **Sample Attribute**: Select the attribute generated by auto grouping. This filter criterion can be used only after auto grouping is enabled.

- **Data Attribute**: Select **All** or **Inference** to filter the data source.

Figure 2-23 Filter criteria
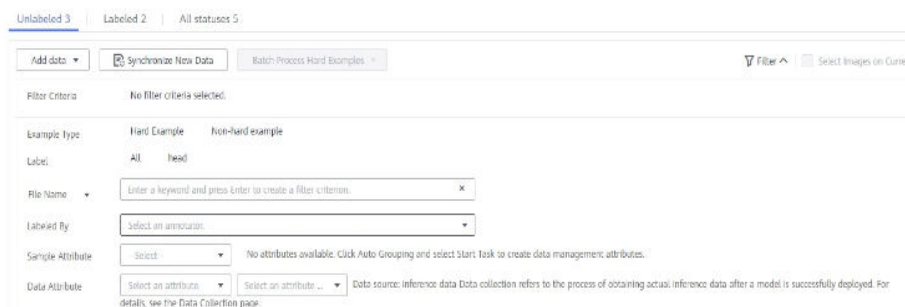
Manually Labeling Images

The labeling job details page displays the **All statuses**, **Unlabeled**, and **Labeled** tab pages. The **Unlabeled** tab page is displayed by default.

1. On the **Unlabeled** tab page, click an image. The system automatically directs you to the page for labeling the image. For details about how to use common buttons on this page, see **Table 2-6**.

2. In the tool bar, select a proper labeling shape. The default labeling shape is a rectangle. In this example, the rectangle is used for labeling.

   ◻ **NOTE**

   In the tool bar, multiple tools are provided for you to label images. After you select a shape to label the first image, the shape automatically applies to subsequent images. You can switch the shape as required.

Table 2-5 Supported bounding box

Icon	Description
◻	Rectangle. You can also press **1**. Click the edge of the upper left corner of the object to be labeled. A rectangle will be displayed. Drag the rectangle to cover the object and click to label the object.
⬠	Polygon. You can also press **2**. In the area where the object to be labeled is located, click to label a point, move the mouse and click multiple points along the edge of the object, and then click the first point again. All the points form a polygon. In this way, the object to be labeled is within the bounding box.
○	Round. You can also press **3**. Click the center point of an object, and move the mouse to draw a circle to cover the object and click to label the object.

Icon	Description
/	Straight. You can also press **4**. Click to specify the start and end points of an object, and move the mouse to draw a straight line to cover the object and click to label the object.
⁄⁄⁄	Dashed line. You can also press **5**. Click to specify the start and end points of an object, and move the mouse to draw a dashed line to cover the object and click to label the object.
⊙	Dot. You can also press **6**. Click the object in an image to label a point.

3. In the **Add Label** text box, enter a new label name, select the label color, and click **Add**. Alternatively, select an existing label from the drop-down list.

   Label all objects in an image. Multiple labels can be added to an image. After labeling an image, click the right arrow (or press D) in the upper right corner of the image to switch to the next image and label the image.

   Figure 2-24 Adding an object detection label
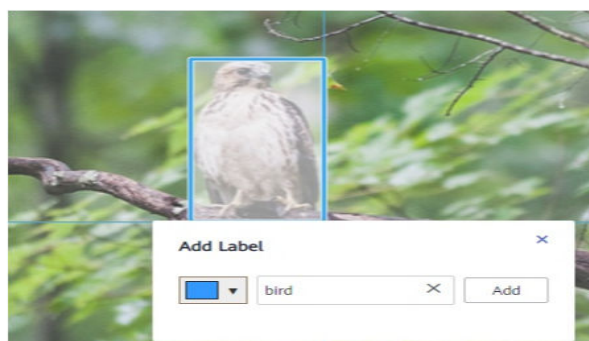
4. Click **Back to Data Labeling Preview** in the upper left part of the page to view the labeling information. In the dialog box that is displayed, click **Yes** to save the labeling settings.

   The selected images are automatically moved to the **Labeled** tab page. On the **Unlabeled** and **All statuses** tab pages, the labeling information is updated along with the labeling process, including the added label names and the number of images for each label.

   **Table 2-6** Common icons on the labeling page

Button	Features
↰	Cancel the previous operation. You can also press **Ctrl+Z**.
↱	Redo the previous operation. You can also press **Ctrl+Shift+Z**.

Button	Features
⊕	Zoom in an image. You can also use the mouse wheel to zoom in.
⊖	Zoom out an image. You can also use the mouse wheel to zoom out.
🗑	Delete all bounding boxes on the current image. You can also press **Shift+Delete**.
⊘	Show or hide a bounding box. This operation can be performed only on a labeled image. You can also press **Shift +H**.
✛	Drag a bounding box to another position or drag the edge of the bounding box to resize it. You can also use **X** + left mouse button.
⟲	Reset a bounding box. After dragging a bounding box, you can click this button to quickly restore the bounding box to its original shape and position. You can also press **Esc**.

Viewing Labeled Images

On the labeling job details page, click the **Labeled** tab to view the list of labeled images. The labels of each image are displayed below the image.

**Figure 2-25** Labels
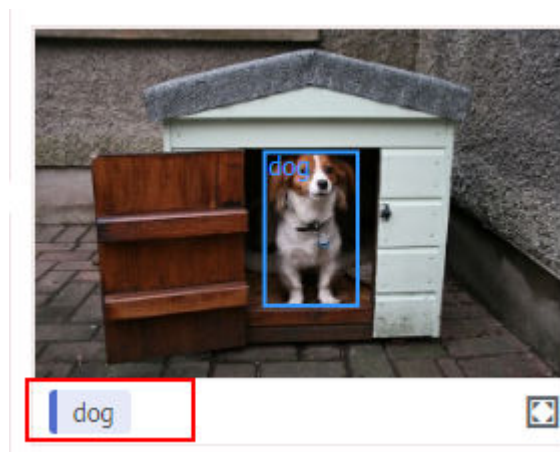
Quick Review

To simplify operations, ModelArts provides quick review so that you can batch review and modify labeled data.

1. Log in to the ModelArts management console. In the navigation pane, choose **Data Preparation** > **Label Data**. On the **My Creations** tab page, select the

target labeling job type from the **All types** drop-down list in the upper right corner. (Only object detection and image segmentation support quick review.)

2. In the labeling job list, click the target labeling job. The labeling details page is displayed.

3. Click **Quick Review** on the **Labeled** tab. On the displayed page, confirm the labeling results.

Figure 2-26 Quick Review
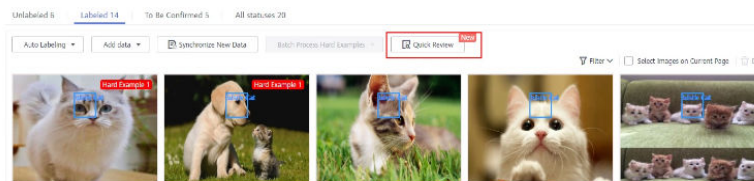
4. Batch review images of the same label.

   a. On the review page, select the label type from the drop-down list next to **Filter by Label**.

   b. Sort images of the selected label type by bounding box area or aspect ratio.

   c. Click an incorrectly labeled image, and then drag the labeling box to relabel the image. (**Modified** is displayed on the modified images.)

   d. You can select the incorrectly labeled images, and then click 🗑 in the upper right corner to delete the label. (**Deleted** is displayed on the images whose label has been deleted.)

   Figure 2-27 Modified

   Figure 2-28 Deleted

   e. You can also modify the label of a labeled image.

      i. Select the target images and click ⊕ in the **All Labels** area on the right.

      ii. Type a new label and click **OK**.
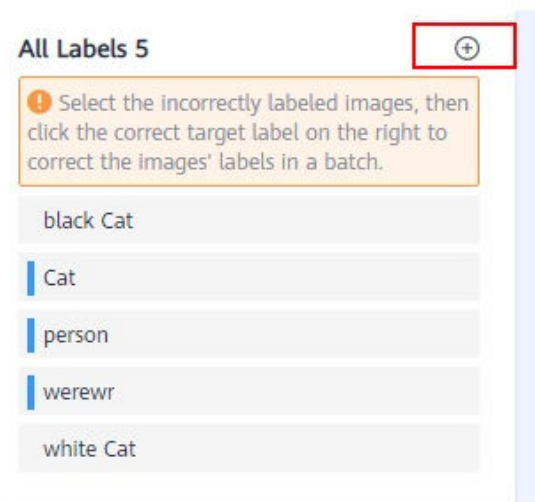
Figure 2-29 All Labels

Figure 2-30 Adding a label
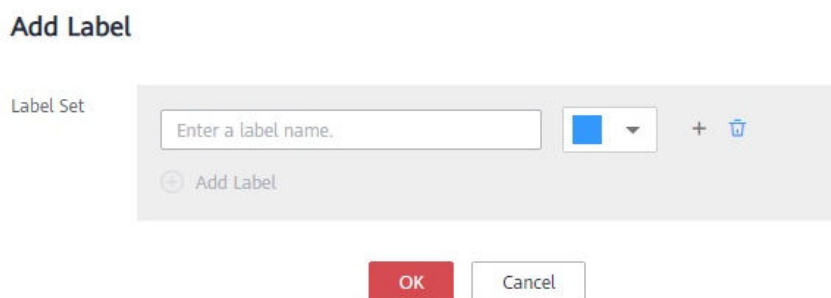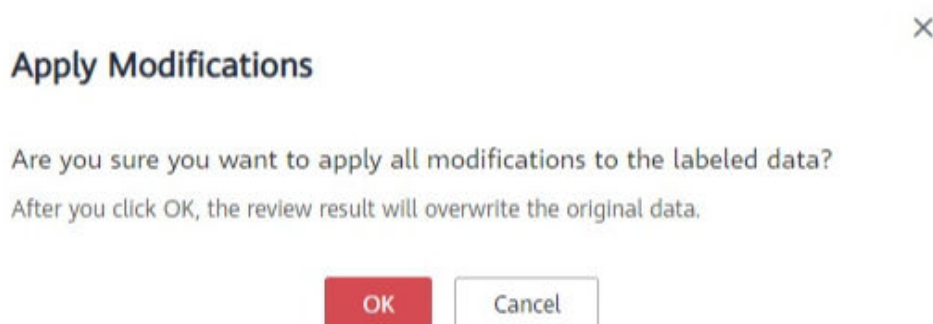

5.   After the modification, click **Apply Modifications**. In the displayed dialog box, click **OK**. The system automatically returns to the labeling overview page and overwrites the original labeling data.

Figure 2-31 Apply Modifications

6.   If you are not satisfied with the modified data, you can click **Cancel Modifications** to retain the original labeling data.

Figure 2-32 Cancel Modifications

Cancel Modifications

Are you sure you want to cancel all modifications in this review?

After you click OK, all the modifications in this review will be cancelled.

| OK | Cancel |

Table 2-7 Buttons on the quick review page

Button	Features
🗑	Delete the label.
⏮	Undo all operations on the current page.
←	Undo the previous operation.
→	Redo the previous operation.

Modifying Labeled Data

After labeling data, you can modify labeled data on the **Labeled** tab page.
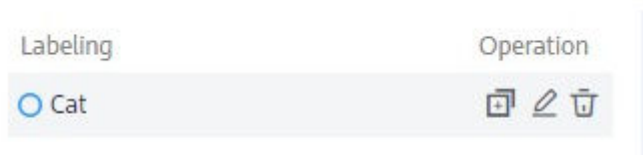
- **Modifying based on images**

  On the labeling job details page, click the **Labeled** tab and then the image to be modified. The labeling page is displayed. Modify the image information in the label information area on the right.

  – Modifying a label: In the **Labeling** area, click the edit icon, enter the correct label name in the text box, and click the check mark to complete the modification. Alternatively, click a label. In the image labeling area, adjust the position and size of the labeling box. After the adjustment, right-click the labeling box and choose **Modify** from the shortcut menu. Enter the new label and click **Modify** to save the modification.

  – Deleting a label: In the **Labeling** area, click the deletion icon to delete a label from the image.

    After deleting the label, click **Back to Data Labeling Preview** in the upper left corner of the page to exit the labeling page. In the dialog box that is displayed, save the modification. After all labels of an image are deleted, the image is displayed on the **Unlabeled** tab page.

Figure 2-33 Editing an object detection label

- **Modifying based on labels**
  - On the labeling job details page, click **Label Management** on the right. All label information is displayed.

    ■ Modifying a label: Click **Modify** in the **Operation** column. In the dialog box that is displayed, enter a new label name, select a new label color, and click **OK**. After the modification, the images with the label added will use the new label name.

    ■ Deleting a label: Click **Delete** in the **Operation** column, or select the label to be deleted and click **Delete Label** above the label list.

Figure 2-34 Label Management

Figure 2-35 All labels

  - Alternatively, click **Label** in the **Operation** column of the target labeling job to go to the label management page.

Figure 2-36 Accessing the label management page from the labeling job list

- Click **Modify** in the **Operation** column of the target label to modify it.
- Click **Delete** in the **Operation** column of the target label to delete it.

Adding Data

In addition to the data automatically synchronized from datasets, you can directly add images to labeling jobs for labeling. The added data is first imported to the dataset associated with the labeling job. Then, the labeling job automatically synchronizes the latest data from the dataset.

1. On the labeling job details page, click **All statuses**, **Labeled**, or **Unlabeled** tab, click **Add data** in the upper left corner.

Figure 2-37 Adding data
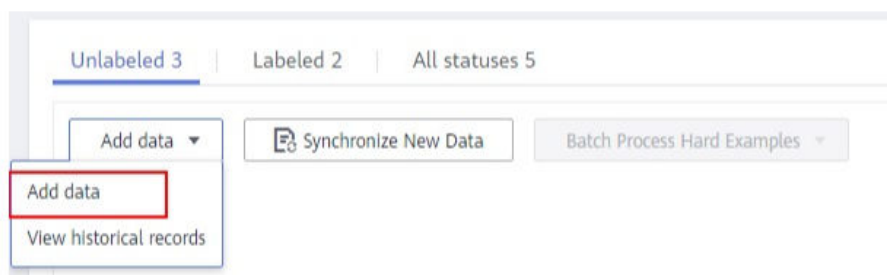
2. Configure the data source, import mode, import path, and labeling status.

Figure 2-38 Adding images

3. Click **OK**.

   The images you have added will be automatically displayed in the image list on the **All statuses** tab page. You can choose **Add data** > **View historical records** to view task history.
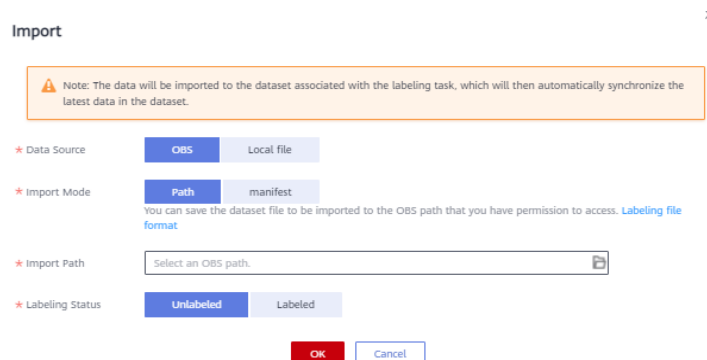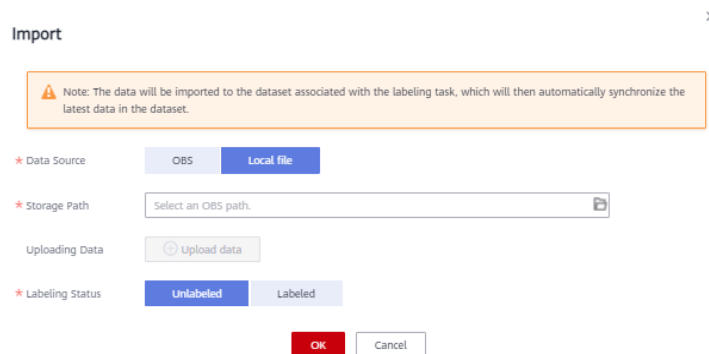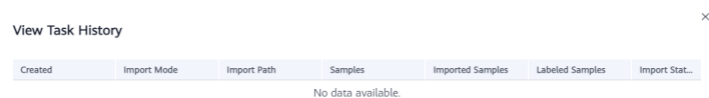
   Figure 2-39 Viewing historical data

Deleting Images

You can quickly delete the images you want to discard.

On the **All statuses**, **Unlabeled**, or **Labeled** tab page, select the images to be deleted or click **Select Images on Current Page**, and click **Delete**. In the displayed dialog box, select or deselect **Delete the source files from OBS** as required. After confirmation, click **Yes** to delete the images.

Figure 2-40 Deleting images

If a tick is displayed in the upper left corner of an image, the image is selected. If no image is selected on the page, the **Delete** button is unavailable.

📖 **NOTE**

If you select **Delete the source files from OBS**, images stored in the OBS directory will be deleted accordingly. This operation may affect other dataset versions or datasets using those files, for example, leading to an error in page display, training, or inference. Deleted data cannot be recovered. Exercise caution when performing this operation.

Managing Annotators

If team labeling is enabled for a labeling job, view its labeling details on the **Annotator Management** tab page. Additionally, you can add, modify, or delete annotators.

1. Choose **Data Preparation** > **Label Data**. On the **My Creations** tab page, view the list of all labeling jobs.

2. Locate the row that contains the target team labeling job. (The name of a team labeling job is followed by $\overset{\mathcal{R}}{}$.)

3. Choose **More** > **Annotator Management** in the **Operation** column. Alternatively, click the job name to go to the job details page, and choose **Team Labeling** > **Annotator Management** in the upper right corner.

Figure 2-41 Annotator management (1)

Figure 2-42 Annotator management (2)
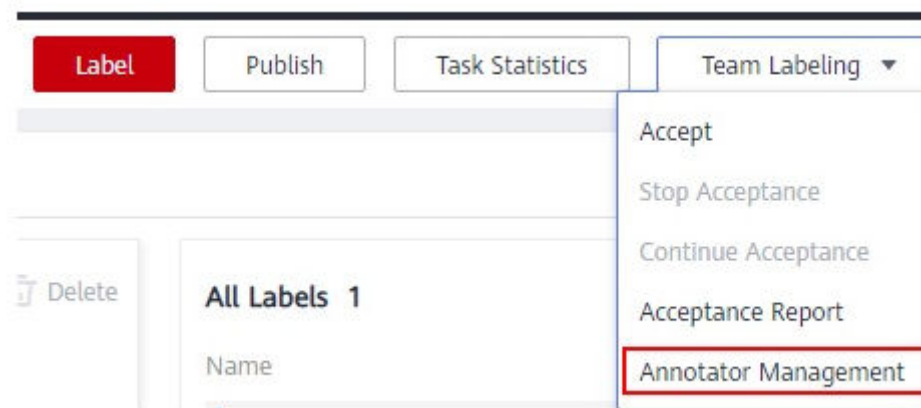
- Adding an annotator

  Click **Add Member**, select a member name, and click **OK**.

  Click **Send Email** in the **Operation** column to send the labeling job to the annotator by email.

- Modifying annotator information

  Click **Modify** in the **Operation** column to modify the role of the annotator.

- Deleting an annotator

  Click **Delete** in the **Operation** column to delete the annotator.

2.2.3 Image Segmentation

Training a model uses a large number of labeled images. Therefore, label images before the model training. You can label images on the ModelArts management console. Alternatively, modify labels, or delete them and label them again.

Before labeling an image in image segmentation scenarios, pay attention to the following:

- All objects whose contours need to be extracted from the image must be labeled.

- Polygons can be used for labeling.
  - In polygon labeling, draw a polygon based on the outline of the target object.
- When labeling an image, ensure that the polygons are within the image. Otherwise, an error will occur in subsequent operations.

Starting Labeling

1. Log in to the ModelArts management console. In the navigation pane on the left, choose **Data Preparation** > **Label Data**.

2. On the right of the labeling job list, select a labeling type from the job type drop-down list. Click the job to be performed based on the labeling type. The details page of the job is displayed.

   Figure 2-43 Selecting a labeling type
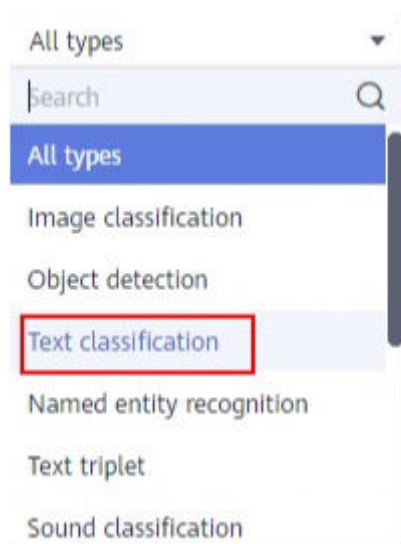
3. The job details page displays all data of the labeling job.

Synchronizing New Data

ModelArts automatically synchronizes data and labeling information from datasets to the labeling job.

To quickly obtain the latest data in a dataset, in the **All statuses**, **Unlabeled**, or **Labeled** tab of the labeling job details page, click **Synchronize New Data**.

📖 NOTE

Symptom:

After the labeled data is uploaded to OBS and synchronized, the data is displayed as unlabeled.

Possible causes:

Automatic encryption is enabled in the OBS bucket.

Solution:

Create an OBS bucket and upload data again, or disable bucket encryption and upload data again.

Filtering Data

In the **All statuses** or **Unlabeled** tab, click ∨ in the filter criteria area and add filter criteria to quickly filter the data you want to view.

The following filter criteria are available. You can set one or more filter criteria.

- **Example Type**: Select **Hard example** or **Non-hard example**.
- **Label**: Select **All** or one or more labels you specified.
- **File Name** or **Path**: Filter files by file name or file storage path.
- **Labeled By**: Select the name of the user who labeled the image.
- **Sample Attribute**: Select the attribute generated by auto grouping. This filter criterion can be used only after auto grouping is enabled.
- **Data Attribute**: Select **All** or **Inference** to filter the data source.

Figure 2-44 Filter criteria
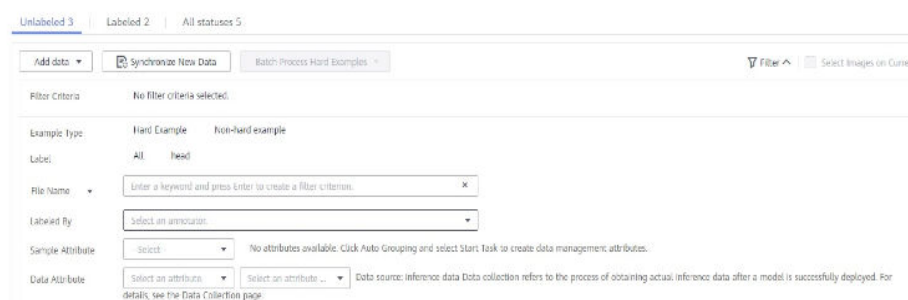
Manually Labeling Images

The labeling job details page displays the **All statuses**, **Unlabeled**, and **Labeled** tabs. The **Unlabeled** tab is displayed by default.

1. In the **Unlabeled** tab, click an image. The system automatically directs you to the page for labeling the image. For details about how to use common buttons on this page, see **Table 2-9**.
2. Select a labeling method.

   On the labeling page, common **labeling methods** and **buttons** are provided in the toolbar. By default, polygon labeling is selected. Use polygon or pole labeling as needed.

 📖 NOTE

> After you select a method to label the first image, the labeling method automatically applies to subsequent images.

Table 2-8 Labeling methods

Icon	Description
⬡	Polygon. In the area where the object to be labeled is located, click to label a point, move the mouse and click multiple points along the edge of the object, and then click the first point again. All the points form a polygon. In this way, the object to be labeled is within the bounding box.

Table 2-9 Toolbar buttons

Button	Features
↩	Cancel the previous operation.
↪	Redo the previous operation.
⊕	Zoom in an image.
⊖	Zoom out an image.
🗑	Delete all bounding boxes on the current image.
⊘	Show or hide a bounding box. This operation can be performed only on a labeled image.
✛	Drag a bounding box to another position or drag the edge of the bounding box to resize it.
⮌	Reset a bounding box. After dragging a bounding box, you can click this button to quickly restore the bounding box to its original shape and position.
⛶	Display the labeled image in full screen.

3. Label an object.

   Identify an object in an image. Click the top, bottom, leftmost, and rightmost points on the object contour. In the displayed dialog box, set the label name and click **Add**. After labeling an image, click ⌃ below the image to view in the image list and click an unlabeled image to label the new image.

4. Click **Back to Data Labeling Preview** in the upper left part of the page to view the labeling information. In the displayed dialog box, click **Yes** to save the labeling settings.

The selected images are automatically moved to the **Labeled** tab. In the **Unlabeled** and **All statuses** tabs, the labeling information is updated along with the labeling process, including the added label names and the number of images for each label.

Viewing Labeled Images

On the labeling job details page, click the **Labeled** tab to view the list of labeled images. Click an image to view its labeling information in the **File Labels** area on the right.

Quick Review

To simplify operations, ModelArts provides quick review so that you can batch review and modify labeled data.

1. Log in to the ModelArts management console. In the navigation pane, choose **Data Preparation** > **Label Data**. On the **My Creations** tab page, select the target labeling job type from the **All types** drop-down list in the upper right corner. (Only object detection and image segmentation support quick review.)

2. In the labeling job list, click the target labeling job. The labeling details page is displayed.

3. Click **Quick Review** on the **Labeled** tab. On the displayed page, confirm the labeling results.

Figure 2-45 Quick Review
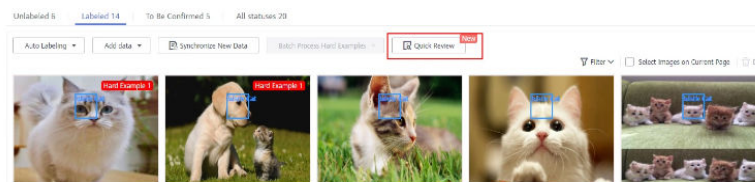
4. Batch review images of the same label.

   a. On the review page, select the label type from the drop-down list next to **Filter by Label**.

   b. Sort images of the selected label type by bounding box area or aspect ratio.

   c. Click an incorrectly labeled image, and then drag the labeling box to relabel the image. (**Modified** is displayed on the modified images.)

   d. You can select the incorrectly labeled images, and then click 🗑 in the upper right corner to delete the label. (**Deleted** is displayed on the images whose label has been deleted.)

Figure 2-46 Modified

Figure 2-47 Deleted

e. You can also modify the label of a labeled image.

   i.   Select the target images and click ⊕ in the **All Labels** area on the right.

   ii. Type a new label and click **OK**.

Figure 2-48 All Labels
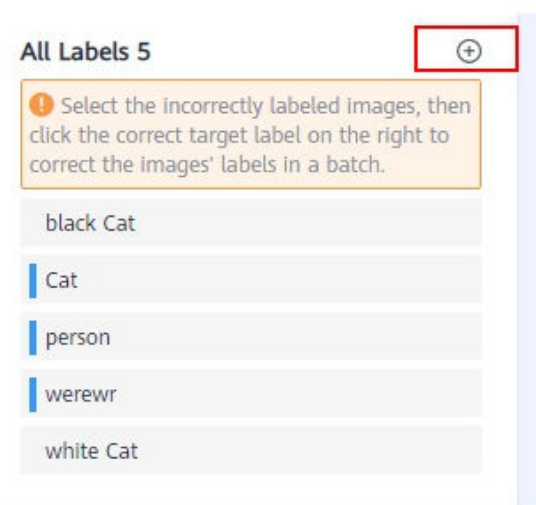
Figure 2-49 Adding a label
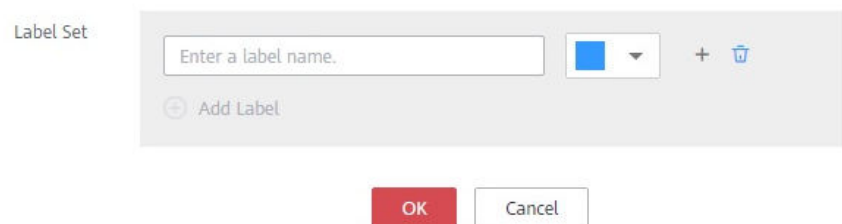
5. After the modification, click **Apply Modifications**. In the displayed dialog box, click **OK**. The system automatically returns to the labeling overview page and overwrites the original labeling data.

Figure 2-50 Apply Modifications
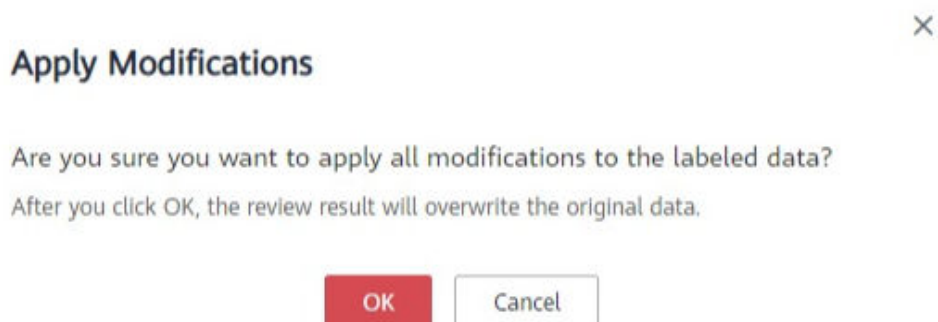
Apply Modifications   ✕

Are you sure you want to apply all modifications to the labeled data?

After you click OK, the review result will overwrite the original data.

[OK]   [Cancel]

6. If you are not satisfied with the modified data, you can click **Cancel Modifications** to retain the original labeling data.

Figure 2-51 Cancel Modifications

Cancel Modifications

Are you sure you want to cancel all modifications in this review?

After you click OK, all the modifications in this review will be cancelled.

[OK]   [Cancel]
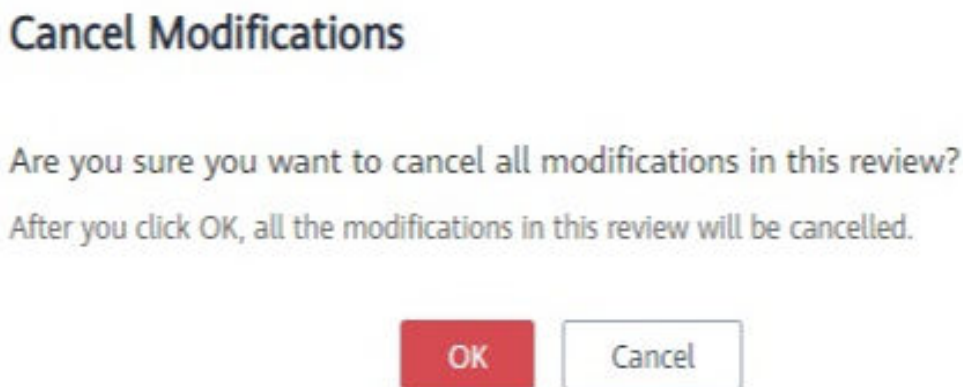
Table 2-10 Buttons on the quick review page

Button	Features
🗑	Delete the label.
↩	Undo all operations on the current page.
←	Undo the previous operation.
→	Redo the previous operation.

Modifying a Label

After labeling data, you can modify labeled data in the **Labeled** tab.

On the labeling details page, click the **Labeled** tab and then the image to be modified. On the displayed labeling page, modify the labeling information in the **File Labels** area on the right.

- Modifying a label: In the **Labeling** area, click the edit icon, set the target label name or color in the displayed dialog box, and click [✓] to save the modification. Alternatively, click a label to be modified. In the image labeling area, adjust the position and size of the bounding box. After the adjustment is complete, click another label to save the modification.

- Deleting a label: In the **Labeling** area, click the deletion icon to delete a label from the image. After all labels of an image are deleted, the image is displayed in the **Unlabeled** tab.

After the labeling information is modified, click **Back to Data Labeling Preview** in the upper left part of the page to exit the labeling page. In the displayed dialog box, click **Yes** to save the modification.

Adding Data

In addition to the data automatically synchronized from datasets, you can directly add images to labeling jobs for labeling. The added data is first imported to the dataset associated with the labeling job. Then, the labeling job automatically synchronizes the latest data from the dataset.

1. On the labeling job details page, click **All statuses**, **Labeled**, or **Unlabeled** tab, click **Add data** in the upper left corner.

Figure 2-52 Adding Data
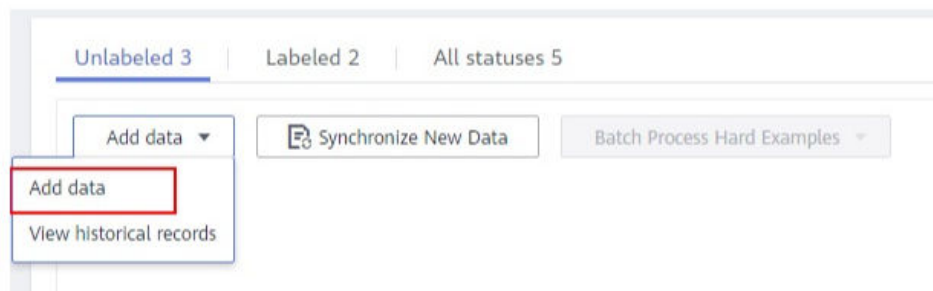
2. Configure the data source, import mode, import path, and labeling status.

Figure 2-53 Adding images
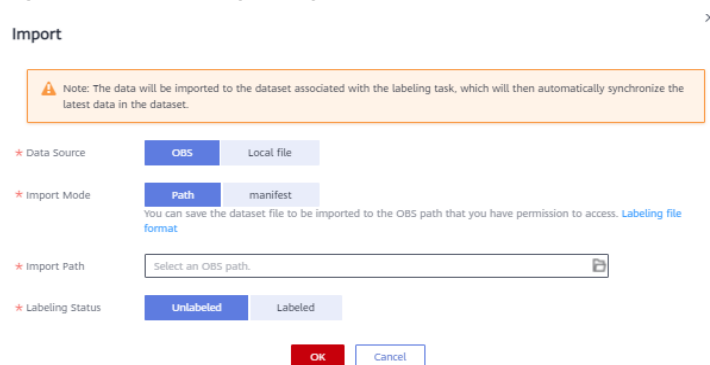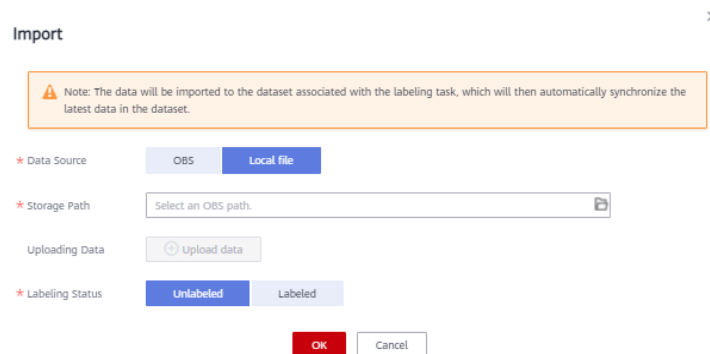
3. Click **OK**.

   The images you have added will be automatically displayed in the image list in the **All statuses** tab. You can choose **Add data** > **View historical records** to view task history.

   Figure 2-54 Viewing historical data
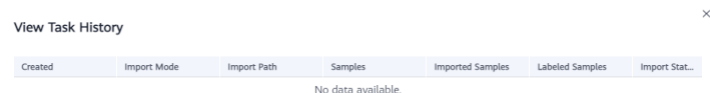
Deleting Images

You can quickly delete the images you want to discard.

In the **All statuses**, **Unlabeled**, or **Labeled** tab, select the images to be deleted or click **Select Images on Current Page**, and click **Delete** in the upper left corner to delete them. In the displayed dialog box, select or deselect **Delete the source files from OBS** as required. After confirmation, click **Yes** to delete the images.

If a tick is displayed in the upper left corner of an image, the image is selected. If no image is selected on the page, the **Delete** button is unavailable.

◯ NOTE

If you select **Delete the source files from OBS**, images stored in the OBS directory will be deleted accordingly. This operation may affect other dataset versions or datasets using those files, for example, leading to an error in page display, training, or inference. Deleted data cannot be recovered. Exercise caution when performing this operation.

2.3 Text Labeling

2.3.1 Text Classification

Model training requires a large amount of labeled data. Therefore, before the model training, add labels to the files that are not labeled. In addition, you can modify, delete, and re-label the labeled text.

Text classification classifies text content based on labels. Before labeling text content, pay attention to the following:

- Text labeling supports multiple labels. That is, you can add multiple labels to a labeling object.

- A label name can contain a maximum of 1024 characters, including letters, digits, hyphens (-), underscores (_), and special characters.

Starting Labeling

1. Log in to the ModelArts management console. In the navigation pane on the left, choose **Data Preparation** > **Label Data**.

2. In the labeling job list, select a labeling type from the **All type** drop-down list, click the job to be performed based on the labeling type. The details page of the job is displayed.

Figure 2-55 Selecting a labeling type

3. The job details page displays all data of the labeling job.

Synchronizing New Data

ModelArts automatically synchronizes data and labeling information from datasets to the labeling job.

To quickly obtain the latest data in the datasets, in the **Unlabeled** tab of the labeling job details page, click **Synchronize New Data**.

☐ **NOTE**

Symptom:

After the labeled data is uploaded to OBS and synchronized, the data is displayed as unlabeled.

Possible causes:

Automatic encryption is enabled in the OBS bucket.

Solution:

Create an OBS bucket and upload data again, or disable bucket encryption and upload data again.

Labeling Text Files

The labeling job details page displays the **Unlabeled** and **Labeled** tabs. The **Unlabeled** tab is displayed by default.

1. In the **Unlabeled** tab, the objects to be labeled are listed in the pane on the left. In the list, click the text object to be labeled, and select a label in the **Label Set** area in the right pane. Multiple labels can be added to a labeling object.

   You can repeat this operation to select objects and add labels to the objects.

   Figure 2-56 Labeling for text classification
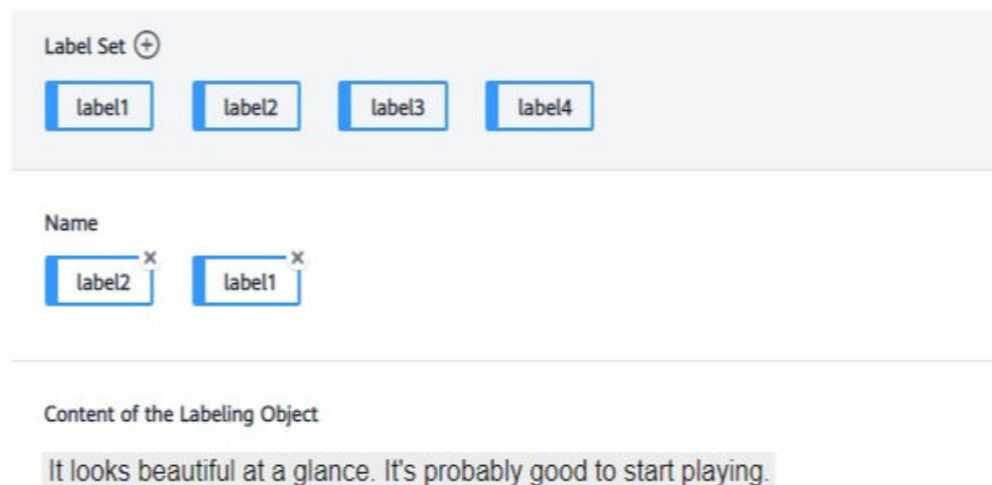
2. After all objects are labeled, click **Save Current Page** at the bottom of the page.

Adding a Label

- Adding labels in the **Unlabeled** tab: Click the plus sign (+) next to **Label Set**. On the displayed **Add Label** page, add a label name, select a label color, and click **OK**.

  Figure 2-57 Adding a label (1)

- Adding labels in the **Labeled** tab: Click the plus sign (+) next to **Label Set**. On the displayed **Add Label** page, add a label name, select a label color, and click **OK**.

Figure 2-58 Adding a label (2)
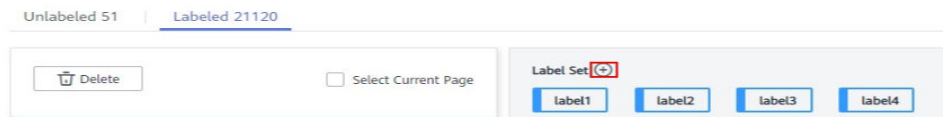
Figure 2-59 Adding a label
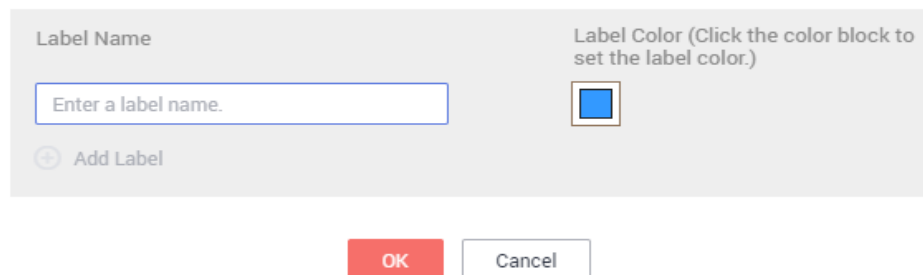
Viewing the Labeled Text

On the labeling job details page, click the **Labeled** tab to view the list of labeled texts. You can also view all labels supported by the labeling job in the **All Labels** area on the right.

Modifying Labeled Data

After labeling data, you can modify labeled data in the **Labeled** tab.

- **Modifying based on texts**

  On the labeling job details page, click the **Labeled** tab and select the text to be modified from the text list.

  In the text list, click the text. When the text background turns blue, the text is selected. If a text file has multiple labels, you can click ✖ above a label to delete the label.

- **Modifying based on labels**

  On the labeling job details page, click the **Labeled** tab. The information about all labels is displayed on the right.

  - Batch modification: In the **All Labels** area, click the edit icon in the **Operation** column, modify the label name in the text box, select a label color, and click **OK**.

  - Batch deletion: In the **All Labels** area, click the deletion icon in the **Operation** column to delete the label. In the displayed dialog box, select **Delete the label** or **Delete the label and objects with only the label**, and click **OK**.

Adding a File

In addition to the data synchronized, you can directly add data on labeling job details page for labeling.

1. On the labeling job details page, click the **Unlabeled** tab, click **Add data** in the upper left corner.

2. Configure input data and click **OK**.

Figure 2-60 Importing data

Deleting a File

You can quickly delete the files you want to discard.

- In the **Unlabeled** tab, select the text to be deleted, and click **Delete** in the upper left corner.

- In the **Labeled** tab, select the text to be deleted, and click **Delete**. Alternatively, tick **Select Current Page** to select all text objects on the current page and click **Delete** in the upper left corner.

The background of the selected text is blue.

Managing Annotators

If team labeling is enabled for a labeling job, view its labeling details on the **Annotator Management** tab page. Additionally, you can add, modify, or delete annotators.

1. Choose **Data Preparation** > **Label Data**. On the **My Creations** tab page, view the list of all labeling jobs.

2. Locate the row that contains the target team labeling job. (The name of a team labeling job is followed by .)

3. Choose **More** > **Annotator Management** in the **Operation** column. Alternatively, click the job name to go to the job details page, and choose **Team Labeling** > **Annotator Management** in the upper right corner.

Figure 2-61 Annotator management (1)

Figure 2-62 Annotator management (2)
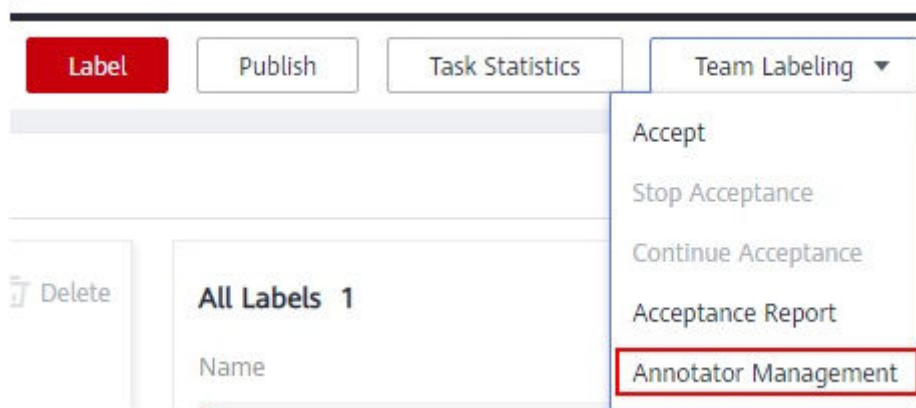
- Adding an annotator

  Click **Add Member**, select a member name, and click **OK**.

  Click **Send Email** in the **Operation** column to send the labeling job to the annotator by email.

- Modifying annotator information

  Click **Modify** in the **Operation** column to modify the role of the annotator.

- Deleting an annotator

  Click **Delete** in the **Operation** column to delete the annotator.

2.3.2 Named Entity Recognition

Named entity recognition assigns labels to named entities in text, such as time and locations. Before labeling, pay attention to the following:

A label name of a named entity can contain a maximum of 1024 characters, including letters, digits, hyphens (-), underscores (_), and special characters.

Starting Labeling

1. Log in to the ModelArts management console. In the navigation pane on the left, choose **Data Preparation** > **Label Data**.
2. In the labeling job list, select a labeling type from the **All type** drop-down list, click the job to be performed based on the labeling type. The details page of the job is displayed.

Figure 2-63 Selecting a labeling type

3. The job details page displays all data of the labeling job.

Synchronizing New Data

ModelArts automatically synchronizes data and labeling information from datasets to the labeling job.

To quickly obtain the latest data in the datasets, in the **Unlabeled** tab of the labeling job details page, click **Synchronize New Data**.

◻ NOTE

Symptom:

After the labeled data is uploaded to OBS and synchronized, the data is displayed as unlabeled.

Possible causes:

Automatic encryption is enabled in the OBS bucket.

Solution:

Create an OBS bucket and upload data again, or disable bucket encryption and upload data again.

Labeling Text Files

The labeling job details page displays the **Unlabeled** and **Labeled** tabs. The **Unlabeled** tab is displayed by default.

1. In the **Unlabeled** tab, the objects to be labeled are listed in the pane on the left. In the list, click the text object to be labeled, select a part of text displayed under **Label Set** for labeling, and select a label in the **Label Set** area in the right pane.

    You can repeat this operation to select objects and add labels to the objects.

Figure 2-64 Labeling for named entity recognition
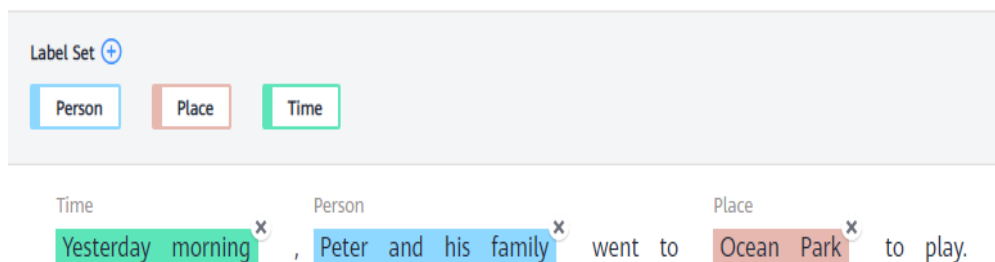
2. Click **Save Current Page** in the lower part of the page to complete the labeling.

Adding a Label

- Adding labels in the **Unlabeled** tab: Click the plus sign (+) next to **Label Set**. On the displayed **Add Label** page, add a label name, select a label color, and click **OK**.

  Figure 2-65 Adding a named entity label (1)
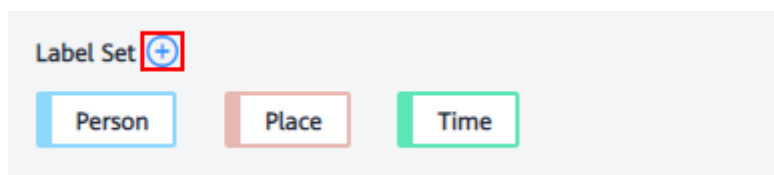
- Adding labels in the **Labeled** tab: Click the plus sign (+) next to **Label Set**. On the displayed **Add Label** page, add a label name, select a label color, and click **OK**.

  Figure 2-66 Adding a named entity label (2)
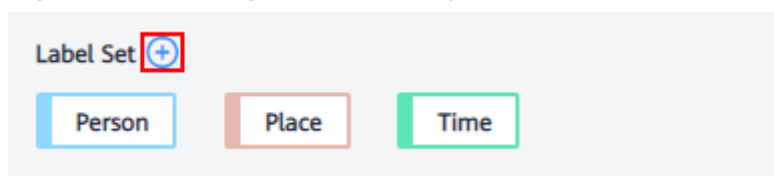
  Figure 2-67 Adding a named entity label

Viewing the Labeled Text

On the dataset details page, click the **Labeled** tab to view the list of the labeled text. You can also view all labels supported by the dataset in the **All Labels** area on the right.

Modifying Labeled Data

After labeling data, you can modify labeled data in the **Labeled** tab.

On the labeling job details page, click the **Labeled** tab, and modify the text information in the label information area on the right.

- **Modifying based on texts**

  On the labeling job details page, click the **Labeled** tab, and select the text to be modified from the text list.

  Manual deletion: In the text list, click the text. When the text background turns blue, the text is selected. On the right of the page, click ✕ above a text label to delete the label.

- **Modifying based on labels**

  On the labeling job details page, click the **Labeled** tab. The information about all labels is displayed on the right.

  - Batch modification: In the **All Labels** area, click the edit icon in the **Operation** column, add a label name in the text box, select a label color, and click **OK**.

  - Batch deletion: In the **All Labels** area, click the deletion icon in the **Operation** column to delete the label. In the displayed dialog box, select **Delete the label** or **Delete the label and objects with only the label**, and click **OK**.

Adding a File

In addition to the data synchronized, you can directly add data on labeling job details page for labeling.

1. On the labeling job details page, click the **Unlabeled** tab, click **Add data** in the upper left corner.

2. Configure input data and click **OK**.

Figure 2-68 Importing data

Deleting a File

You can quickly delete the files you want to discard.

● In the **Unlabeled** tab, select the text to be deleted, and click **Delete** in the upper left corner.

● In the **Labeled** tab, select the text to be deleted, and click **Delete**. Alternatively, tick **Select Current Page** to select all text objects on the current page and click **Delete** in the upper left corner.

The background of the selected text is blue.

Managing Annotators

If team labeling is enabled for a labeling job, view its labeling details on the **Annotator Management** tab page. Additionally, you can add, modify, or delete annotators.

1. Choose **Data Preparation** > **Label Data**. On the **My Creations** tab page, view the list of all labeling jobs.

2. Locate the row that contains the target team labeling job. (The name of a team labeling job is followed by 👤.)

3. Choose **More** > **Annotator Management** in the **Operation** column. Alternatively, click the job name to go to the job details page, and choose **Team Labeling** > **Annotator Management** in the upper right corner.

Figure 2-69 Annotator management (1)

Figure 2-70 Annotator management (2)
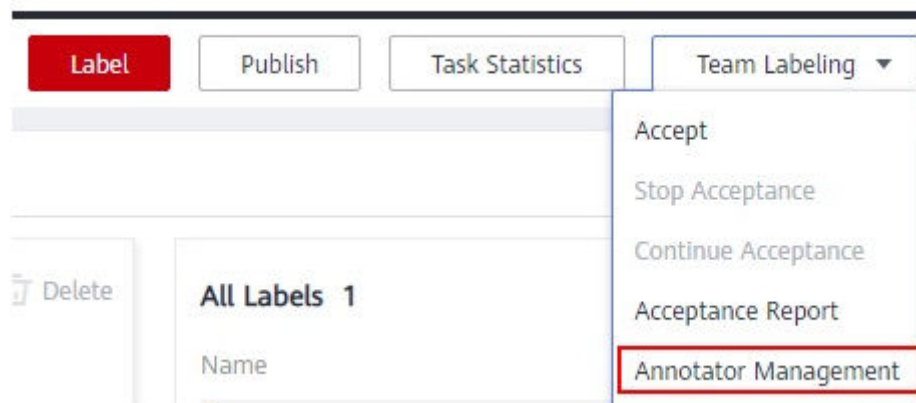
- Adding an annotator

  Click **Add Member**, select a member name, and click **OK**.

  Click **Send Email** in the **Operation** column to send the labeling job to the annotator by email.

- Modifying annotator information

  Click **Modify** in the **Operation** column to modify the role of the annotator.

- Deleting an annotator

  Click **Delete** in the **Operation** column to delete the annotator.

2.3.3 Text Triplet

Triplet labeling is suitable for scenarios where structured information, such as subjects, predicates, and objects, needs to be labeled in statements. With this function, not only entities in statements, but also relationships between entities can be labeled. Triplet labeling is often used in natural language processing tasks such as dependency syntax analysis and information extraction.

Text triplet labeling involves two classes of important labels: **Entity Label** and **Relationship Label**. For **Relationship Label**, set its **Source entity** and **Target entity**.

- You can define multiple entity and relationship labels for a text object.
- The **Entity Label** defined during dataset creation cannot be deleted.

Precautions

Before labeling, ensure that the **Entity Label** and **Relationship Label** of a labeling job have been defined. For **Relationship Label**, set its **Source entity** and **Target entity**. **Relationship Label** must be between the defined **Source entity** and **Target entity**.

For example, if two entities are labeled as **Place**, you cannot add any relationship label between them, as shown in **Figure 2-71**. If a relationship label cannot be added, a red cross is displayed, as shown in **Figure 2-72**.

Figure 2-71 Example of entity and relationship labels
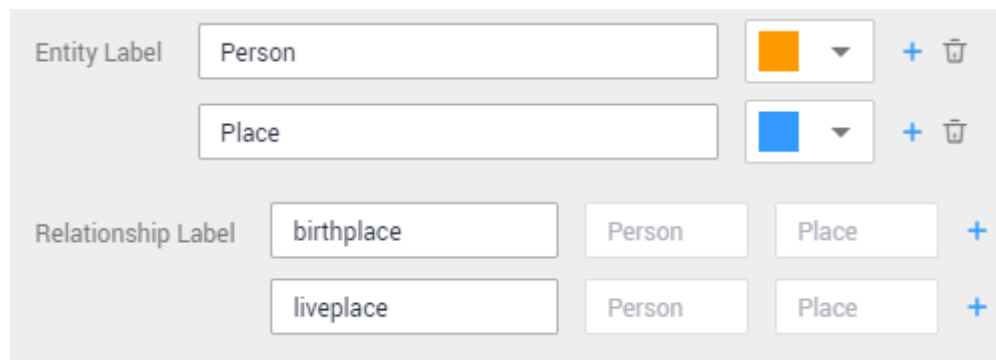
Figure 2-72 Failure of adding a relationship label
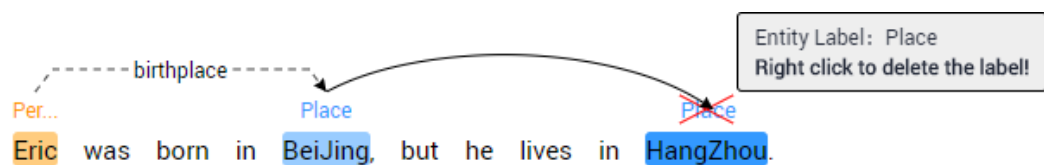
Starting Labeling

1. Log in to the ModelArts management console. In the navigation pane on the left, choose **Data Preparation** > **Label Data**.

2. In the labeling job list, select a labeling type from the **All type** drop-down list, click the job to be performed based on the labeling type. The details page of the job is displayed.

   Figure 2-73 Selecting a labeling type

3. The job details page displays all data of the labeling job.

Synchronizing New Data

ModelArts automatically synchronizes data and labeling information from datasets to the labeling job.

To quickly obtain the latest data in the datasets, in the **Unlabeled** tab of the labeling job details page, click **Synchronize New Data**.

◻ NOTE

Symptom:

After the labeled data is uploaded to OBS and synchronized, the data is displayed as unlabeled.

Possible causes:

Automatic encryption is enabled in the OBS bucket.

Solution:

Create an OBS bucket and upload data again, or disable bucket encryption and upload data again.

Labeling Text Files

The labeling job details page displays the **Unlabeled** and **Labeled** tabs. The **Unlabeled** tab is displayed by default.

1. In the **Unlabeled** tab, the objects to be labeled are listed in the pane on the left. In the list, click a text object, select the corresponding text content on the right pane, and select an entity name from the displayed entity list to label the content.

    Figure 2-74 Labeling an entity
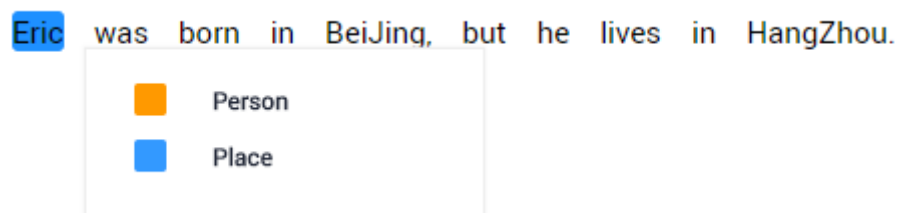
2. After labeling multiple entities, click the source entity and target entity in sequence and select a relationship type from the displayed relationship list to label the relationship.

    Figure 2-75 Labeling a relationship
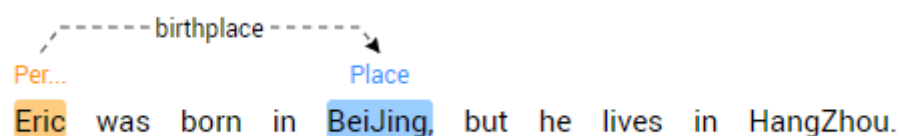
3. After all objects are labeled, click **Save Current Page** at the bottom of the page.

📖 **NOTE**

> You cannot modify the labels of a dataset in the text triplet type on the labeling page. Instead, click **Label Management** and modify the **Entity Label** and **Relationship Label**.

Modifying Labeled Data

After labeling data, you can modify labeled data in the **Labeled** tab.

On the labeling job details page, click the **Labeled** tab. Select a text object in the left pane and the right pane displays the detailed label information. You can move your cursor to the entity or relationship label, and right-click to delete it. You can also click the source entity and target entity in sequence to add a relationship label.

Figure 2-76 Modifying a label in the text
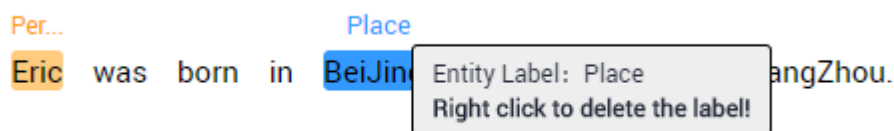
You can click **Delete Labels on Current Item** at the bottom of the page to delete all labels in the selected text object.

Figure 2-77 Deleting current labels
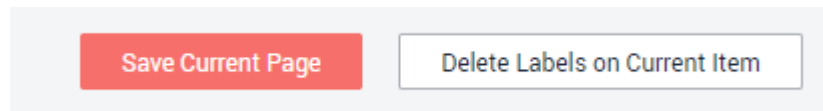
Adding a File

In addition to the data synchronized, you can directly add data on labeling job details page for labeling.

1. On the labeling job details page, click the **Unlabeled** tab, click **Add data** in the upper left corner.
2. Configure input data and click **OK**.

Figure 2-78 Importing data
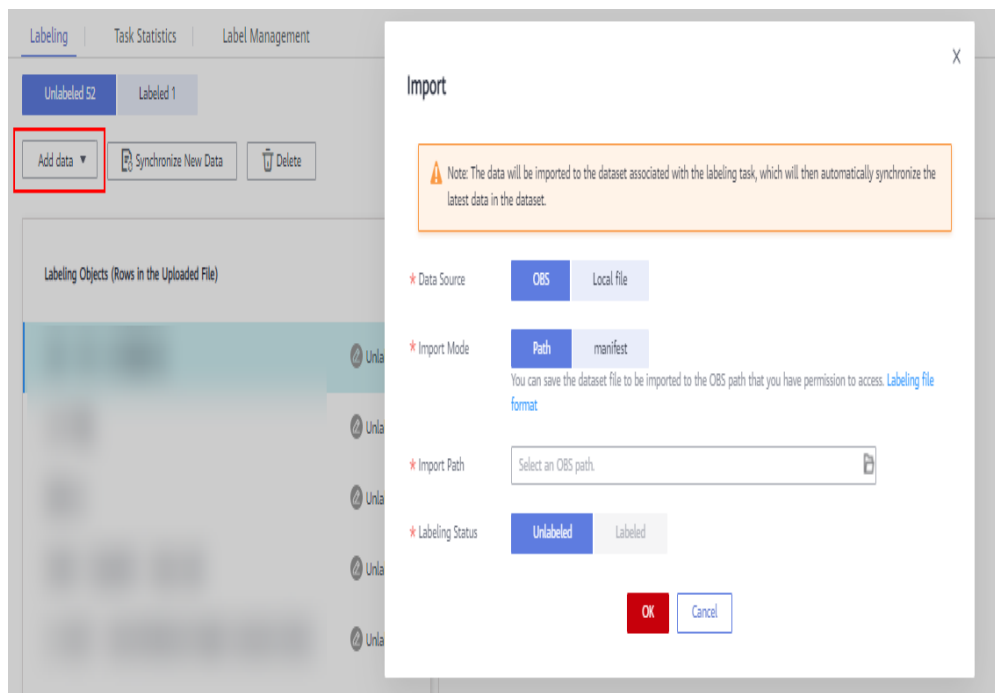
Deleting a File

You can quickly delete the files you want to discard.

- In the **Unlabeled** tab, select the text to be deleted, and click **Delete** in the upper left corner.
- In the **Labeled** tab, select the text to be deleted, and click **Delete**. Alternatively, tick **Select Current Page** to select all text objects on the current page and click **Delete** in the upper left corner.

The background of the selected text is blue. If no text is selected on the page, the **Delete** button is unavailable.

Managing Annotators

If team labeling is enabled for a labeling job, view its labeling details on the **Annotator Management** tab page. Additionally, you can add, modify, or delete annotators.

1. Choose **Data Preparation** > **Label Data**. On the **My Creations** tab page, view the list of all labeling jobs.
2. Locate the row that contains the target team labeling job. (The name of a team labeling job is followed by 👤.)
3. Choose **More** > **Annotator Management** in the **Operation** column. Alternatively, click the job name to go to the job details page, and choose **Team Labeling** > **Annotator Management** in the upper right corner.

Figure 2-79 Annotator management (1)

Figure 2-80 Annotator management (2)
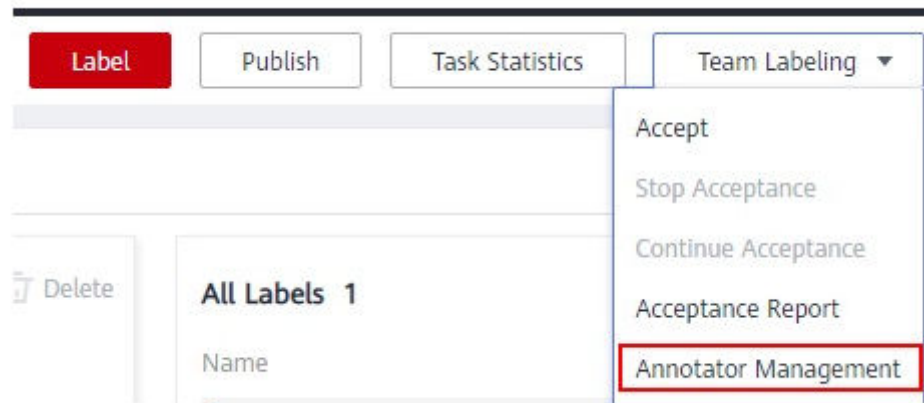
- Adding an annotator

  Click **Add Member**, select a member name, and click **OK**.

  Click **Send Email** in the **Operation** column to send the labeling job to the annotator by email.

- Modifying annotator information

  Click **Modify** in the **Operation** column to modify the role of the annotator.

- Deleting an annotator

  Click **Delete** in the **Operation** column to delete the annotator.

2.4 Audio Labeling

2.4.1 Sound Classification

Sound classification is to classify different sounds.

Model training requires a large amount of labeled data. Therefore, before the model training, label the unlabeled audio files. ModelArts enables you to label audio files in batches by one click. In addition, you can modify the labels of audio files, or remove their labels and label the audio files again.

Only Chinese and English are supported for audio labeling.

Starting Labeling

1. Log in to the ModelArts management console. In the navigation pane on the left, choose **Data Preparation** > **Label Data**.

2. In the labeling job list, select a labeling type from the **All type** drop-down list, click the job to be performed based on the labeling type. The details page of the job is displayed.

Figure 2-81 Selecting a labeling type

3. The job details page displays all data of the labeling job.

Synchronizing New Data

ModelArts automatically synchronizes data and labeling information from datasets to the labeling job.

To quickly obtain the latest data in the datasets, in the **Unlabeled** or **Labeled** tab of the labeling job details page, click **Synchronize New Data**.

📖 **NOTE**

Symptom:

After the labeled data is uploaded to OBS and synchronized, the data is displayed as unlabeled.

Possible causes:

Automatic encryption is enabled in the OBS bucket.

Solution:

Create an OBS bucket and upload data again, or disable bucket encryption and upload data again.

Labeling Audio Files

The labeling job details page displays the **Unlabeled** and **Labeled** tabs. The

Unlabeled tab is displayed by default. Click on the left of the audio to preview the audio.

1. In the **Unlabeled** tab, select the audio files to be labeled.
   - Manual selection: In the audio list, click the target audio. If the blue check box is displayed in the upper right corner, the audio is selected. You can select multiple audio files of the same type and label them together.
   - Batch selection: If all audio files of the current page belong to one type, you can click **Select Current Page** in the upper right corner of the list to select all the audio files on the page.

2. Add labels.

   a. In the label adding area on the right, set a label in the **Label** text box.

      Method 1 (the required label already exists): In the right pane, select a shortcut from the **Shortcut** drop-down list, select an existing label name from the **Label** text box, and click **OK**.

      Method 2 (adding a label): In the right pane, select a shortcut from the **Shortcut** drop-down list, and enter a new label name in the **Label** text box.

   b. The selected audio files are automatically moved to the **Labeled** tab. In the **Unlabeled** tab, the labeling information is updated along with the labeling process, including the added label names and the number of audio files corresponding to each label.

   📖 **NOTE**

   Shortcut key description: After specifying a shortcut key for a label, you can select an audio file and press the shortcut key to add a label for the audio file. Example: Specify **1** as the shortcut key for the **aa** label. Select one or more files and press **1**. A message is displayed, asking you whether to label the files with **aa**. Click **OK**.

   Each label has a shortcut key. A shortcut key cannot be specified for different labels. Shortcut keys can greatly improve the labeling efficiency.

Figure 2-82 Adding an audio label
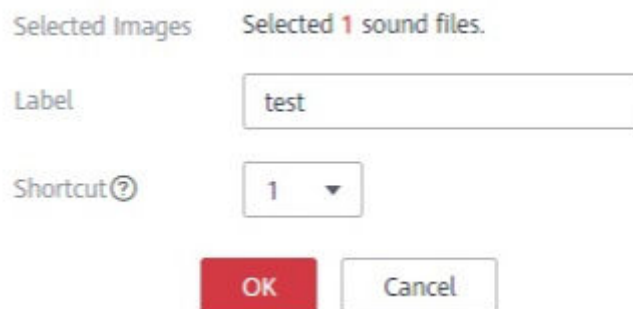
Viewing the Labeled Audio Files

On the labeling job details page, click the **Labeled** tab to view the list of labeled audio files. Click an audio file. You can view the label information about the audio file in the **File Labels** area on the right.

Modifying Labeled Data

After labeling data, you can modify labeled data in the **Labeled** tab.

- **Modifying based on audio**

  On the labeling job details page, click the **Labeled** tab. Select one or more audio files to be modified from the audio list. Modify the label in the label details area on the right.

  – Modifying a label: In the **File Labels** area, click the edit icon in the **Operation** column, enter the correct label name in the text box, and click the check mark to complete the modification.

– Deleting a label: In the **File Labels** area, click the delete icon in the **Operation** column to delete the label.

- **Modifying based on labels**

  On the labeling job details page, click the **Labeled** tab. The information about all labels is displayed on the right.

  Figure 2-83 Information about all labels

All Labels 4		Label Management C	
Label	Count ⬇≡	Shortcut	Operation
8	1	6	✎ 🗑
aaaa	7	1	✎ 🗑
qqqq	1	2	✎ 🗑
test	1	--	✎ 🗑

  – Modifying a label: Click the edit icon in the **Operation** column. In the displayed dialog box, enter the new label name and click **OK**. After the modification, the new label applies to the audio files that contain the original label.

  – Deleting a label: Click the deletion icon in the **Operation** column. In the displayed dialog box, select the object to be deleted as prompted and click **OK**.

Adding Audio Files

In addition to the data synchronized, you can directly add data on labeling job details page for labeling.

1. On the labeling job details page, click the **Unlabeled** or **Labeled** tab, click **Add data** in the upper left corner.
2. Configure input data and click **OK**.

Figure 2-84 Importing data
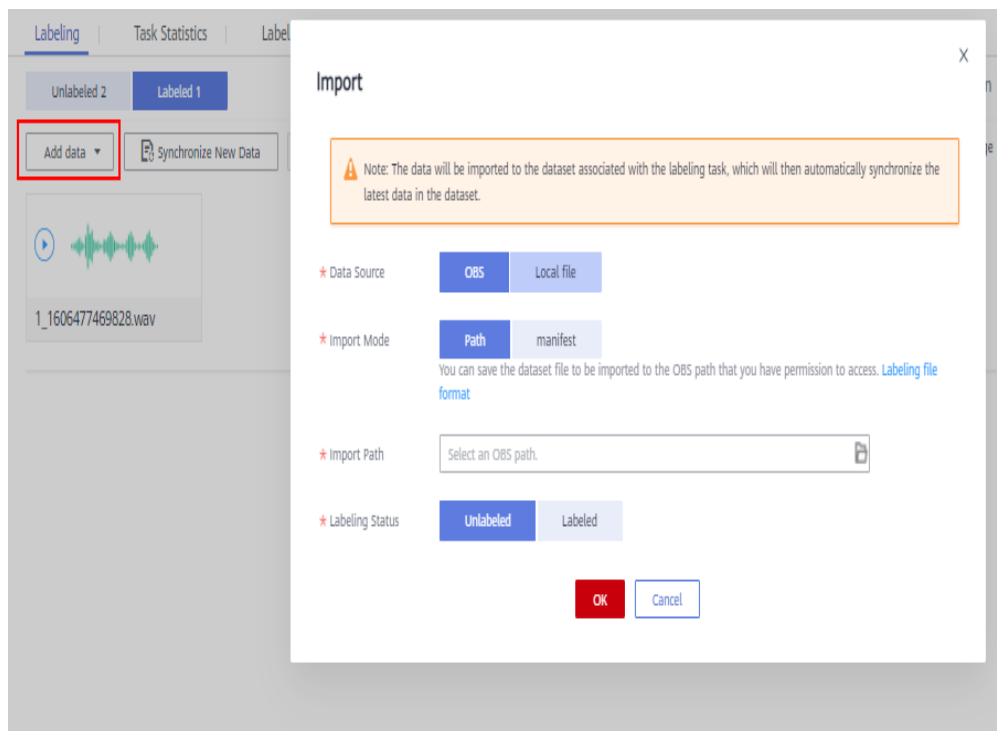
Deleting Audio Files

You can quickly delete the audio files you want to discard.

In the **Unlabeled** or **Labeled** tab, select the audio files to be deleted one by one or tick **Select Current Page** to select all audio files on the page, and then click **Delete File** in the upper left corner. In the displayed dialog box, select or deselect **Delete the source files from OBS** as required. After confirmation, click **OK** to delete the audio files.

If a tick is displayed in the upper right corner of an audio file, the audio file is selected. If no audio file is selected on the page, the **Delete File** button is unavailable.

📖 **NOTE**

If you select **Delete the source files from OBS**, audio files stored in the corresponding OBS directory will be deleted when you delete the selected audio files. Deleting source files may affect other dataset versions or datasets using those files. As a result, the page display, training, or inference is abnormal. Deleted data cannot be recovered. Exercise caution when performing this operation.

2.4.2 Speech Labeling

Speech labeling is to label speech content.

Model training requires a large amount of labeled data. Therefore, before the model training, label the unlabeled audio files. ModelArts enables you to label audio files in batches by one click. In addition, you can modify the labels of audio files, or remove their labels and label the audio files again.

Only Chinese and English are supported for audio labeling.

Starting Labeling

1. Log in to the ModelArts management console. In the navigation pane on the left, choose **Data Preparation** > **Label Data**.

2. In the labeling job list, select a labeling type from the **All type** drop-down list, click the job to be performed based on the labeling type. The details page of the job is displayed.

Figure 2-85 Selecting a labeling type

3. The job details page displays all data of the labeling job.

Synchronizing New Data

ModelArts automatically synchronizes data and labeling information from datasets to the labeling job.

To quickly obtain the latest data in the datasets, in the **Unlabeled** tab of the labeling job details page, click **Synchronize New Data**.

📖 **NOTE**

Symptom:

After the labeled data is uploaded to OBS and synchronized, the data is displayed as unlabeled.

Possible causes:

Automatic encryption is enabled in the OBS bucket.

Solution:

Create an OBS bucket and upload data again, or disable bucket encryption and upload data again.

Labeling Audio Files

The labeling job details page displays the labeled and unlabeled audio files. The **Unlabeled** tab is displayed by default.

1. In the audio file list in the **Unlabeled** tab, click the target audio file. In the
    area on the right, the audio file is displayed. Click ▶ below the audio file to
    play the audio.
2. In **Speech Content**, enter the speech content.
3. After entering the content, click **Label** to complete the labeling. The audio file
    is automatically moved to the **Labeled** tab.

Figure 2-86 Labeling speech content

Viewing the Labeled Audio Files

On the labeling job details page, click the **Labeled** tab to view the list of labeled
audio files. Click the audio file to view the audio content in the **Speech Content**
text box on the right.

Modifying Labeled Data

After labeling data, you can modify labeled data in the **Labeled** tab.

On the labeling job details page, click the **Labeled** tab and select the audio file to
be modified from the audio file list. In the label information area on the right,
modify the content of the **Speech Content** text box, and click **Label** to complete
the modification.

Adding an Audio File

In addition to the data synchronized, you can directly add data on labeling job
details page for labeling.

1. On the labeling job details page, click the **Unlabeled** tab, click **Add data** in
    the upper left corner.
2. Configure input data and click **OK**.
    .

Figure 2-87 Importing data

Deleting Audio Files

You can quickly delete the audio files you want to discard.

In the **Unlabeled** or **Labeled** tab, select the audio files to be deleted, and then click **Delete File** in the upper left corner. In the displayed dialog box, select or deselect **Delete the source files from OBS** as required. After confirmation, click **OK** to delete the audio files.

NOTE

If you select **Delete the source files from OBS**, audio files stored in the corresponding OBS directory will be deleted when you delete the selected audio files. Deleting source files may affect other dataset versions or datasets using those files. As a result, the page display, training, or inference is abnormal. Deleted data cannot be recovered. Exercise caution when performing this operation.

2.4.3 Speech Paragraph Labeling

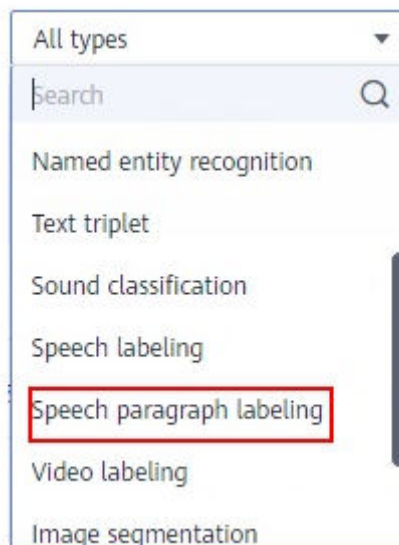Speech paragraph labeling is to segment and label speech content.

Model training requires a large amount of labeled data. Therefore, before the model training, label the unlabeled audio files. ModelArts enables you to label audio files. In addition, you can modify the labels of audio files, or remove their labels and label the audio files again.

Only Chinese and English are supported for audio labeling.

Starting Labeling

1. Log in to the ModelArts management console. In the navigation pane on the left, choose **Data Preparation** > **Label Data**.

2. In the labeling job list, select a labeling type from the **All type** drop-down list, click the job to be performed based on the labeling type. The details page of the job is displayed.

Figure 2-88 Selecting a labeling type
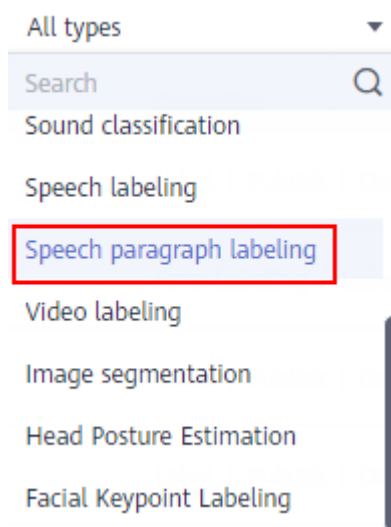
3. The job details page displays all data of the labeling job.

Synchronizing Data Sources

ModelArts automatically synchronizes data and labeling information from datasets to the labeling job.

To quickly obtain the latest data in the OBS bucket, click **Synchronize Data Source** in the **Unlabeled** tab of the labeling job details page to add the data uploaded using OBS to the dataset.

> **NOTE**
>
> Symptom:
>
> After the labeled data is uploaded to OBS and synchronized, the data is displayed as unlabeled.
>
> Possible causes:
>
> Automatic encryption is enabled in the OBS bucket.
>
> Solution:
>
> Create an OBS bucket and upload data again, or disable bucket encryption and upload data again.

Labeling Audio Files

The labeling job details page displays the **Unlabeled** and **Labeled** tabs. The **Unlabeled** tab is displayed by default.
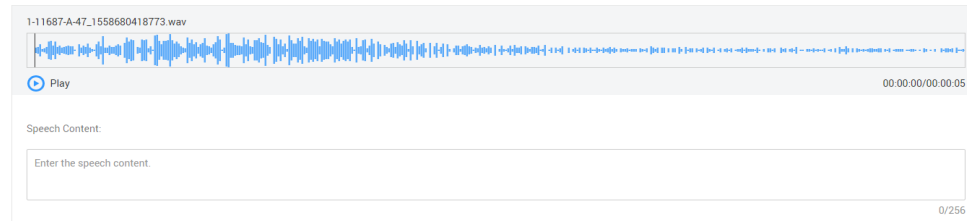
1. In the audio file list in the **Unlabeled** tab, click the target audio file. In the area on the right, the audio file is displayed. Click below the audio file to play the audio.

2. Select an audio segment based on the content being played, and enter the audio file label and content in the **Speech Content** text box.

Figure 2-89 Speech paragraph labeling
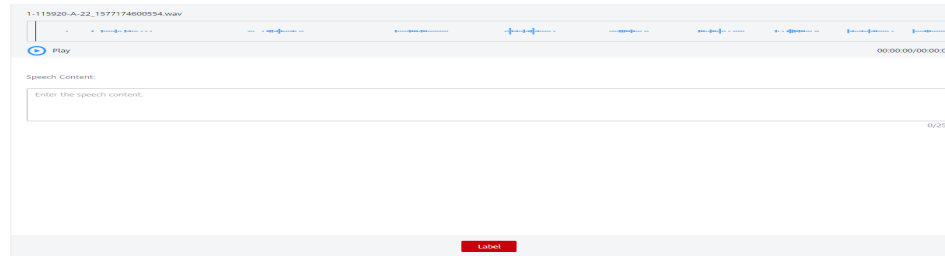
3. After entering the content, click **Label** to complete the labeling. The audio file is automatically moved to the **Labeled** tab.

Viewing the Labeled Audio Files

On the labeling job details page, click the **Labeled** tab to view the list of labeled audio files. Click the audio file to view the labeling information on the right.

Modifying Labeled Data

After labeling data, you can modify labeled data in the **Labeled** tab.

- Modifying a label: On the labeling details page, click the **Labeled** tab, and select the audio file to be modified from the audio file list. In the right area, modify labeling information and click **Label** to complete the modification.
- Deleting a label: Click **Delete** in the **Operation** column of the target number to delete the label of the audio segment. Alternatively, you can click 🗑 above the labeled audio file to delete the label. Then click **Label**.

Adding an Audio File

In addition to the data synchronized, you can directly add data on labeling job details page for labeling.

1. On the labeling job details page, click the **Unlabeled** tab, click **Add data** in the upper left corner.
2. Configure input data and click **OK**.
   .

Figure 2-90 Importing data

Deleting Audio Files

You can quickly delete the audio files you want to discard.

In the **Unlabeled** or **Labeled** tab, select the audio files to be deleted, and then click **Delete File** in the upper left corner. In the displayed dialog box, select or deselect **Delete the source files from OBS** as required. After confirmation, click **OK** to delete the audio files.

☐ **NOTE**

If you select **Delete the source files from OBS**, audio files stored in the corresponding OBS directory will be deleted when you delete the selected audio files. Deleting source files may affect other dataset versions or datasets using those files. As a result, the page display, training, or inference is abnormal. Deleted data cannot be recovered. Exercise caution when performing this operation.

Managing Annotators

If team labeling is enabled for a labeling job, view its labeling details on the **Annotator Management** tab page. Additionally, you can add, modify, or delete annotators.

1. Choose **Data Preparation** > **Label Data**. On the **My Creations** tab page, view the list of all labeling jobs.

2. Locate the row that contains the target team labeling job. (The name of a team labeling job is followed by 🐾.)

3. Choose **More** > **Annotator Management** in the **Operation** column. Alternatively, click the job name to go to the job details page, and choose **Team Labeling** > **Annotator Management** in the upper right corner.

Figure 2-91 Annotator management (1)

Figure 2-92 Annotator management (2)
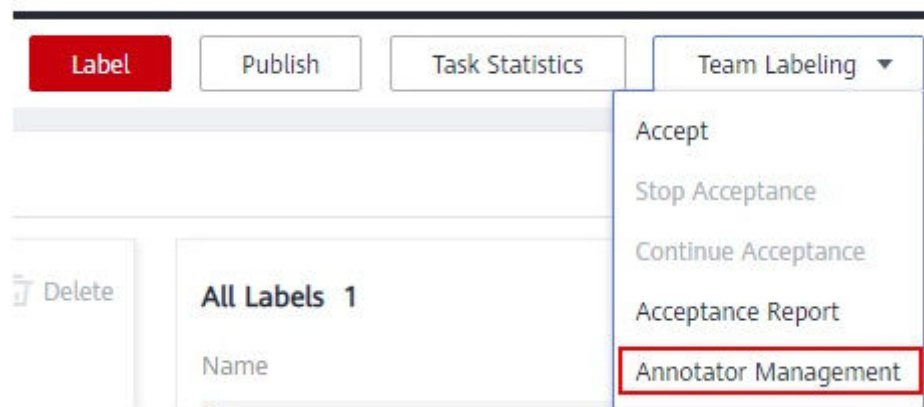
- Adding an annotator

  Click **Add Member**, select a member name, and click **OK**.

  Click **Send Email** in the **Operation** column to send the labeling job to the annotator by email.

- Modifying annotator information

  Click **Modify** in the **Operation** column to modify the role of the annotator.

- Deleting an annotator

  Click **Delete** in the **Operation** column to delete the annotator.

2.5 Video Labeling

Model training requires a large amount of labeled video data. Therefore, before the model training, label the unlabeled video files. ModelArts enables you to label video files. In addition, you can modify the labels of video files, or remove their labels and label the video files again.

📖 **NOTE**

> Video labeling applies only to video frames.

Starting Labeling

1. Log in to the ModelArts management console. In the navigation pane on the left, choose **Data Preparation** > **Label Data**.

2. In the labeling job list, select a labeling type from the **All type** drop-down list, click the job to be performed based on the labeling type. The details page of the job is displayed.

Figure 2-93 Selecting a labeling type
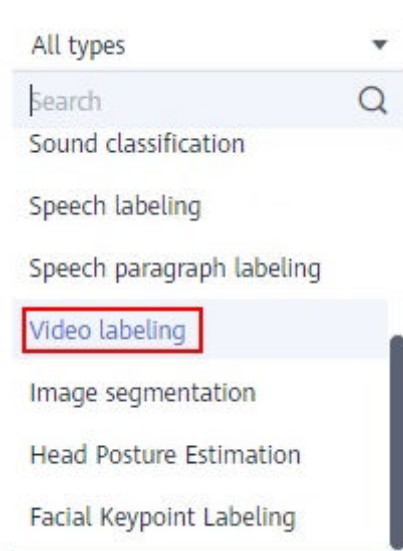
3. The job details page displays all data of the labeling job.

Synchronizing Data Sources

ModelArts automatically synchronizes data and labeling information from **Input Dataset Path** to the dataset details page.

To quickly obtain the latest data in the OBS bucket, in the **Labeled** or **Unlabeled** tab of the labeling job details page, click **Synchronize Data Source**.

☐ **NOTE**

Symptom:

After the labeled data is uploaded to OBS and synchronized, the data is displayed as unlabeled.

Possible causes:

Automatic encryption is enabled in the OBS bucket.

Solution:

Create an OBS bucket and upload data again, or disable bucket encryption and upload data again.

Video Labeling

The labeling job details page displays the **Unlabeled**, **Labeled**, and **All statuses** tabs.

1. In the **Unlabeled** tab, click the target video file in the video list on the left. The labeling page is displayed.

2. Play the video. When the video is played to the time point to be labeled, click the pause button in the progress bar to pause the video to a specific image.

3. In the upper pane, select a bounding box. By default, a rectangular box is selected. Drag the mouse to select an object in the video image, enter a new label name in the displayed **Add Label** text box, select a label color, and click **Add** to label the object. Alternatively, select an existing label from the drop-

down list and click **Add** to label the object. Label all objects in the image. Multiple labels can be added to an image.

The supported labeling boxes are the same as those for object detection. For details, see **Common icons on the labeling page**.

4. After the previous image is labeled, click the play button on the progress bar to resume the playback. Then, repeat **3** to complete labeling on the entire video.

Click **Label List** in the upper right corner of the page. The time points when the video is labeled are displayed.

Figure 2-94 File labels

File Labels

00:00:03   🗑   ⌄
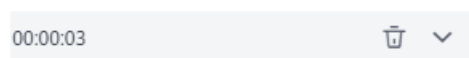
5. Click **Back to Data Labeling Preview** in the upper left corner of the page. The labeling job details page is displayed, and the labeled video file is displayed in the **Labeled** tab.

FAQs

Q: What can I do if the video dataset cannot be displayed or videos cannot be played?

A: If this issue occurs, check the video format. Only MP4 videos can be displayed and played.

Modifying Labeled Data

After labeling data, you can modify labeled data in the **Labeled** tab.

● In the **Labeled** tab, click the target video file. In the upper right corner of the labeling page, click **Label List** to go to the **File Labels** page. You can click the triangle icon on the right of the time point to view details, modify labels, and delete labels.

● Modifying a label: On the **File Labels** area, click the edit button on the right of a label to modify it.

● Deleting a label: On the **File Labels** area, click the delete button on the right of a label to delete it. If you click the delete icon on the right of the image time, all labels on the image are deleted.

Figure 2-95 Modifying Labeled Data
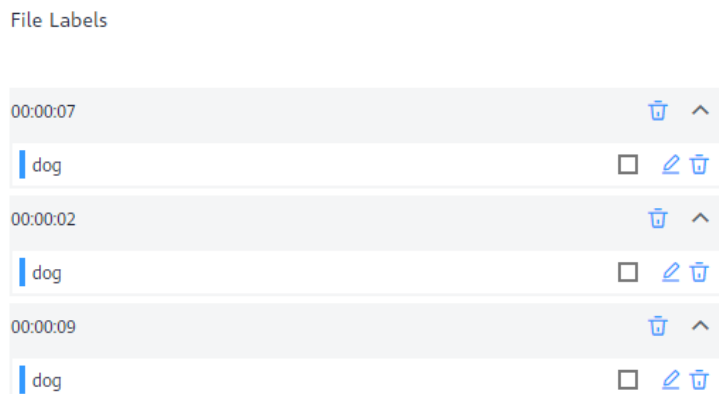
Adding Video Files

In addition to the data synchronized, you can directly add data on labeling job details page for labeling.

1. On the labeling job details page, click the **Unlabeled** or **Labeled** tab, click **Add data** in the upper left corner.
2. Configure the data source, import mode, and other parameters, and click **OK**.

Figure 2-96 Importing data
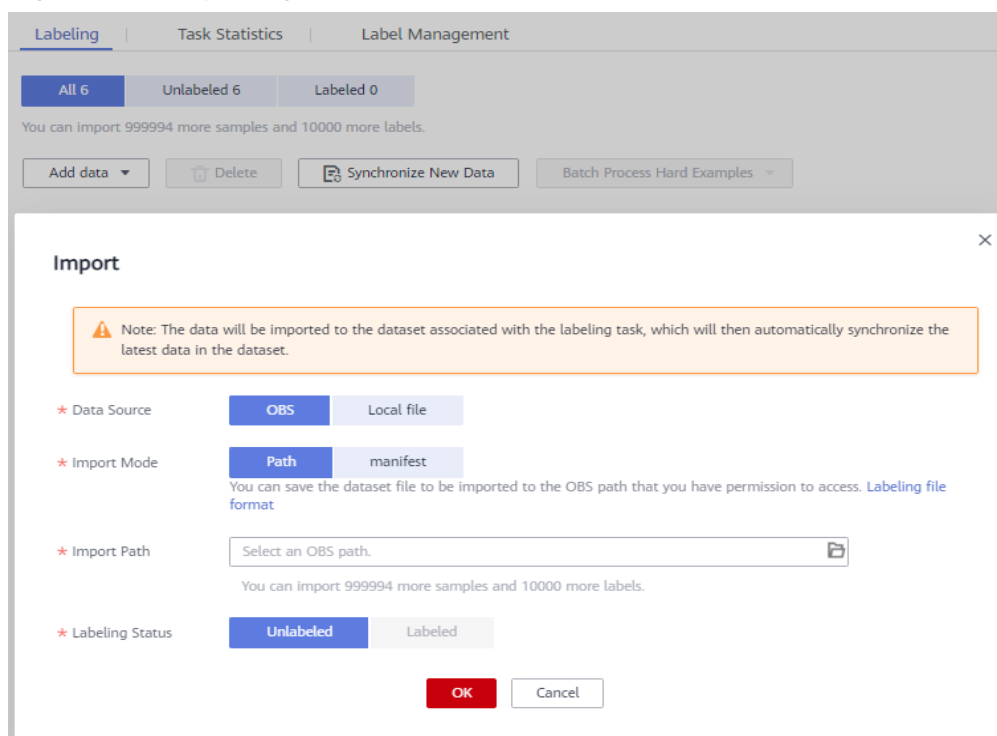
Deleting a Video File

You can quickly delete the video files you want to discard.

In the **All statuses**, **Unlabeled**, or **Labeled** tab, select the video files to be deleted or click **Select Images on Current Page** to select all video files on the page, and

click **Delete** in the upper part to delete the video files. In the displayed dialog box, select or deselect **Delete the source files from OBS** as required. After confirmation, click **OK** to delete the videos.

If a tick is displayed in the upper left corner of a video file, the video file is selected. If no video file is selected on the page, the **Delete** button is unavailable.

📖 **NOTE**

If you select **Delete the source files from OBS**, video files stored in the corresponding OBS directory will be deleted when you delete the selected video files. Deleting source files may affect other dataset versions or datasets using those files. As a result, the page display, training, or inference is abnormal. Deleted data cannot be recovered. Exercise caution when performing this operation.

2.6 Viewing Labeling Jobs

2.6.1 Viewing My Created Labeling Jobs

On the ModelArts **Data Labeling** page, view your created labeling jobs in the **My Creations** tab.

Procedure

1. Log in to the ModelArts management console. In the navigation pane on the left, choose **Data Preparation** > **Label Data**. The **Data Labeling** page is displayed.

2. On the **My Creations** tab, view all labeling jobs created by you. You can view information about these labeling jobs.

    Figure 2-97 My Creations

Copying a Labeling Job

1. Log in to the ModelArts management console. In the navigation pane on the left, choose **Data Preparation** > **Label Data**. The **Data Labeling** page is displayed.

2. On the **My Creations** tab, locate the labeling job you want to copy.

3. Choose **More** > **Copy** in the **Operation** column of the job.

4. In the **Copy Task** dialog box, enter the job description and job name *Task name*-**copy**-*xxxx*, where *xxxx* is a randomly generated code to distinguish the new job from the copied job. You can also change the name of the new job. Click **Yes**.

Copy Task

You are copying the labeling task **dataset-a4d9** . After you click OK, a new labeling task is created based on the task.

Name   dataset-a4d9-copy-c78f

Description

                                           0/256

⚠ The copy operation rebuilds the job based on the dataset associated with the job. If the dataset sample changes, the new job will change accordingly. In addition, the labeling result cannot be copied.

Yes   No

5. After the labeling job is copied, you can obtain the new labeling job on the labeling job list page. The new labeling job information includes the samples, labels, and team labeling information.

2.6.2 Viewing My Participated Labeling Jobs

On the ModelArts **Data Labeling** page, view your participated labeling jobs in the **My Participations** tab.

Prerequisites

Team labeling is enabled when a labeling job is created.

Procedure

1. Log in to the ModelArts management console. In the navigation pane on the left, choose **Data Preparation** > **Label Data**. The **Data Labeling** page is displayed.

2. Click the **My Participations** tab to view the labeling jobs you have participated in, including the members in the labeling team and the labeling progress.

3 Auto Labeling

3.1 Creating an Auto Labeling Job

In addition to manual labeling, ModelArts also provides the auto labeling function to quickly label data, reducing the labeling time by more than 70%. Auto labeling means learning and training are performed based on the labeled images and an existing model is used to quickly label the remaining images.

Context

- Only labeling jobs of image classification and object detection types support auto labeling.
- There are at least two types of labels in the labeling job for auto labeling, and each label has been added to at least five images.
- At least one unlabeled image must exist when you enable auto labeling.
- Before starting an auto labeling job, ensure that no auto labeling job is in progress.
- Before starting an auto labeling job, ensure that the image data does not contain any RGBA four-channel images. These images will cause the job to fail. Delete them from the dataset if you find any.

Starting an Auto Labeling Job

1. Log in to the ModelArts management console. In the navigation pane on the left, choose **Data Preparation** > **Label Data**. The **Data Labeling** page is displayed.

   📖 **NOTE**

   Data labeling is supported only in the following regions: CN North-Beijing4, CN North-Beijing1, CN East-Shanghai1, CN South-Guangzhou, CN Southwest-Guiyang1, CN-Hong Kong, AP-Singapore, AP-Bangkok, AP-Jakarta, LA-Santiago, LA-Sao Paulo1, and LA-Mexico City2.

2. In the labeling job list, locate the row containing a labeling job of the object detection or image classification type and click **Auto Labeling** in the **Operation** column.

3. On the **Enable Auto Labeling** page, select **Active learning** or **Pre-labeling**. For details, see **Table 3-1** and **Table 3-2**.

Table 3-1 Active learning

Parameter	Description
Auto Labeling Type	**Active learning**: The system uses semi-supervised learning and hard example filtering to perform auto labeling, reducing manual labeling workload and helping you find hard examples.
Algorithm Type	For a dataset of the image classification type, set the following parameters: **Fast**: Use the labeled samples for training. **Precise**: Use labeled and unlabeled samples for semi-supervised training, which improves the model precision.
Specifications	Resource specifications used by an auto labeling job. **NOTE** Creating an auto labeling job is free, but you will be billed for OBS storage based on usage. For details, see **Product Pricing Details**. To avoid wasting resources, clear your OBS bucket after labeling jobs and their subsequent tasks are complete.
Compute Nodes	The default value is **1**, indicating the single-node system mode. Only this parameter value is supported.

Table 3-2 Pre-labeling

Parameter	Description
Auto Labeling Type	**Pre-labeling**: Select a model in the **My AI Applications** tab. Ensure that the model type matches the dataset labeling type. After the pre-labeling is complete, if the labeling result complies with the standard labeling format defined by the platform, the system filters hard examples. This step does not affect the pre-labeling result.
Model and Version	● **My AI Applications**: Select a model as required. Click the drop-down arrow on the left of the target AI application and select a proper version. For details about how to import a model, see **Creating a Model**
Specifications	In the drop-down list, you can select the node specifications supported by ModelArts.
Compute Nodes	The default value is **1**. You can select a value based on site requirements. The maximum value is **5**.

 NOTE

For labeling jobs of the object detection type, only rectangular boxes can be recognized and labeled when **Active learning** is selected.

Figure 3-1 Enabling auto labeling (image classification)

Figure 3-2 Enabling auto labeling (object detection)

Figure 3-3 Enabling auto labeling (pre-labeling)

4. After setting the parameters, click **Submit** to enable auto labeling.

5. In the labeling job list, click a labeling job name to go to the labeling job details page.

6. Click the **To Be Confirmed** tab to view the auto labeling progress.

   You can also enable auto labeling or view the auto labeling history in this tab.

Figure 3-4 Labeling progress

📖 **NOTE**

If there are too many auto labeling jobs, they may have to wait in a queue due to limited free resources. This means that they will stay in the labeling state until their turn comes. To ensure that your labeling job can run properly, you are advised to avoid peak hours.

7. After auto labeling is complete, all the labeled images are displayed on the **To Be Confirmed** page.

   – Image classification labeling job

On the **To Be Confirmed** page, check whether labels are correct, select the correctly labeled images, and click **OK** to confirm the auto labeling results. The confirmed image will be categorized to the **Labeled** page.

You can modify the labels of the images that are marked as hard examples according to your needs. For details, see **For labeling jobs of the image classification type**.

– Object detection labeling job

On the **To Be Confirmed** page, click images to view their labeling details and check whether labels and target bounding boxes are correct. For the correctly labeled images, click **Labeled** to confirm the auto labeling results. The confirmed image will be categorized to the **Labeled** page.

You can modify the labels or target bounding boxes of the images marked as hard examples according to your needs. For details, see **For labeling jobs of the object detection type**.

FAQs

- **What can I do if auto labeling fails?**

  Auto labeling is free of charge. If there are too many auto labeling jobs, they may have to wait in a queue due to limited free resources. Create an auto labeling job again or avoid peak hours.

- **What can I do if auto labeling takes a long time?**

  Auto labeling is free of charge. If there are too many auto labeling jobs, they may have to wait in a queue due to limited free resources. You are advised to avoid peak hours.

3.2 Confirming Hard Examples

In a labeling task that processes a large amount of data, auto labeling results cannot be directly used for training because the labeled images are insufficient at the initial stage of labeling. It takes a lot of time and manpower to adjust and confirm all unlabeled data one by one. To accelerate labeling progress, ModelArts embeds an auto hard example detection function for labeling unlabeled data in an auto labeling task. This function provides suggestions on labeling priorities for remaining unlabeled images. The auto labeling result of an image with high labeling priority is not as expected. Therefore, this case is called a hard example.

The auto hard example detection function is used to automatically label hard examples during auto labeling and data collection and filtering. Further confirm and label hard example data, and add labeling results to the training dataset to obtain a trained model with higher precision. No manual intervention is required for hard example detection, and you only need to confirm and modify the labeled data, improving data management and labeling efficiency. In addition, you can supplement data similar to hard examples to improve the variety of the dataset and further improve the model training precision.

Hard example management involves three scenarios.

- **Confirming Hard Examples After Auto Labeling**

- **Labeling Data in a Dataset as Hard Examples**

> **NOTE**
>
> Only datasets of image classification and object detection types support the auto hard example detection function.

Confirming Hard Examples After Auto Labeling

During the execution of an auto labeling task, ModelArts automatically detects and labels hard examples. After the auto labeling task is complete, the labeling results of hard examples are displayed in the **To Be Confirmed** tab. Modify hard example data and confirm the labeling result.

1. Log in to the ModelArts management console. In the navigation pane on the left, choose **Data Preparation** > **Label Data**. On the **Data Labeling** page, click **My Creations**.

2. In the labeling job list, select a labeling job of the object detection or image classification type and click the labeling job name to go to the labeling job details page.

3. In the **Labeling** tab, click **To Be Confirmed** to check and confirm hard examples.

   > **NOTE**
   >
   > Labeling data is displayed in the **To Be Confirmed** tab only after the auto labeling task is complete. Otherwise, no data is available in the tab. For details about auto labeling, see **Creating an Auto Labeling Job**.

   – For labeling jobs of the object detection type

     In the **To Be Confirmed** tab, click an image to expand its labeling details. Check whether labeling information is correct, for example, whether the label is correct and whether the target bounding box is correctly added to the right position. If the auto labeling result is inaccurate, manually adjust the label or target bounding box and click **Labeled**. Then, the re-labeled data is displayed in the **Labeled** tab.

   – For labeling jobs of the image classification type

     In the **To Be Confirmed** tab, check whether labels added to images with the **Hard example** mark are correct. Select the images that are incorrectly labeled, delete the incorrect labels, and add correct labels in **Label** on the right. Click **OK**. The selected images and its labeling details are displayed in the **Labeled** tab.

     The selected images are incorrectly labeled. Delete the incorrect labels on the right, add the **dog** label in **Label**, and click **OK** to confirm the hard examples.

Labeling Data in a Dataset as Hard Examples

In a labeling job, labeled or unlabeled image data can be labeled as hard examples. Data labeled as hard examples can be used to improve model precision through built-in rules during subsequent model training.

1. Log in to the ModelArts management console. In the navigation pane on the left, choose **Data Preparation** > **Label Data**. On the **Data Labeling** page, click **My Creations**.

2. In the labeling job list, select a labeling job of the object detection or image classification type and click the labeling job name to go to the labeling job details page.

3. On the labeling job details page, click the **Labeled**, **Unlabeled**, or **All** tab, select the images to be labeled as hard examples, and choose **Batch Process Hard Examples** > **Confirm Hard Example**. After the labeling is complete, a **Hard example** mark will be displayed in the upper right corner of a preview image.

<h1 style="text-align: right;">4 Team Labeling</h1>

4.1 Team Labeling Overview

Generally, a small data labeling job can be completed by an individual. However, team work is required to label a large dataset. ModelArts provides the team labeling function. A labeling team can be formed to manage labeling for the same dataset.

📖 **NOTE**

> Team labeling is supported only in the following regions: CN North-Beijing4, CN Southwest-Guiyang1, CN-Hong Kong, AP-Singapore, AP-Bangkok.
>
> Team labeling is available only to datasets for image classification, object detection, text classification, named entity recognition, text triplet, and speech paragraph labeling.

For labeling jobs with team labeling enabled, you can create team labeling jobs and assign them to different teams so that team members can complete the labeling jobs together. During data labeling, members can initiate acceptance, continue acceptance, and view acceptance reports.

Team labeling is managed in a unit of teams. To enable team labeling for a dataset, a team must be specified. Multiple members can be added to a team.

- An account can have a maximum of 10 teams.
- An account must have at least one team to enable team labeling for datasets. If the account has no team, add a team by referring to **Adding a Team**.

4.2 Creating and Managing Teams

4.2.1 Managing Teams

Team labeling is managed in a unit of teams. To enable team labeling for a dataset, a team must be specified. Multiple members can be added to a team.

Context

- An account can have a maximum of 10 teams.

- An account must have at least one team to enable team labeling for datasets. If the account has no team, add a team by referring to **Adding a Team**.

Adding a Team

1. In the left navigation pane of the ModelArts management console, choose **Data Preparation > Labeling Teams**. The **Labeling Teams** page is displayed.

2. On the **Labeling Teams** page, click **Add Team**.

3. In the displayed **Add Team** dialog box, enter a team name and description and click **OK**. The labeling team is added.

   The new team is displayed on the **Labeling Teams** page. You can view team details in the right pane. There is no member in the new team. Add members to the new team by referring to **Adding a Member**.

Deleting a Team

You can delete a team that is no longer used.

On the **Labeling Teams** page, select the target team and click **Delete**. In the dialog box that is displayed, click **OK**.

Figure 4-1 Deleting a team
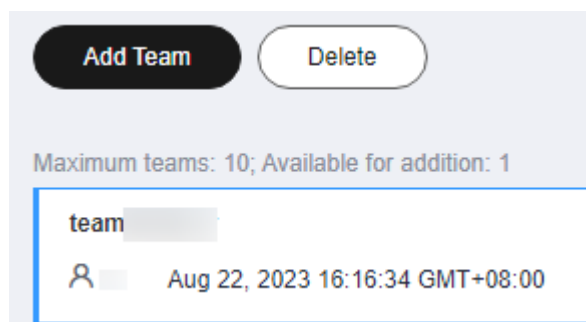
4.2.2 Managing Team Members

There is no member in a new team. You need to add members who will participate in a team labeling job.

A maximum of 100 members can be added to a team. If there are more than 100 members, add them to different teams for better management.

Adding a Member

1. In the left navigation pane of the ModelArts management console, choose **Data Preparation > Labeling Teams**. The **Labeling Teams** page is displayed.

2. On the **Labeling Teams** page, select a team from the team list on the left and click a team name. The team details are displayed in the right pane.

3. In the **Team Details** area, click **Add Member**.

4. An email address uniquely identifies a team member. Different members cannot use the same email address. The email address you enter will be recorded and saved in ModelArts. It is used only for ModelArts team labeling. After a member is deleted, the email address will also be deleted.

Possible values of **Role** are **Labeler**, **Reviewer**, and **Team Manager**. Only one **Team Manager** can be set.

No annotator cannot be deleted from a labeling team with labeling tasks assigned. The labeling result of an annotator can be synchronized to the overall labeling result only after the annotator's labeling is approved, and the labeling result cannot be filtered.

Information about the added member is displayed in the **Team Details** area.

Modifying Member Information

You can modify member information if it is changed.

1. In the **Team Details** area, select the desired member.
2. In the row containing the desired member, click **Modify** in the **Operation** column. In the displayed dialog box, modify the description or role.

   The email address of a member cannot be changed. To change the email address of a member, delete the member, and set a new email address when adding a member.

   Possible values of **Role** are **Labeler**, **Reviewer**, and **Team Manager**. Only one **Team Manager** can be set.

Deleting Members

- **Deleting a single member**

  In the **Team Details** area, select the desired member, and click **Delete** in the **Operation** column. In the dialog box that is displayed, click **OK**.

- **Batch Deletion**

  In the **Team Details** area, select members to be deleted and click **Delete**. In the dialog box that is displayed, click **OK**.

Figure 4-2 Batch deletion

4.3 Creating a Team Labeling Job

If you enable team labeling when creating a labeling job and assign a team to label the dataset, the system creates a labeling job based on the team by default. After creating the labeling job, you can view the job on the **My Creations** tab page of the dataset.

You can also create a team labeling job and assign it to different members in the same team or to other labeling teams.

Methods

- Choose **Data Preparation** > **Label Data** on the console. When creating a labeling job, enable **Team Labeling** and select a team or task manager.

Figure 4-3 Enabling team labeling

- Choose **Asset Management** > **Datasets** on the console. In the **Operation** column of the target dataset, click **Labeling**. On the **Create Labeling Job** page that is displayed, enable **Team Labeling**. You can create multiple team labeling jobs for the same dataset.

**Figure 4-4** Enabling team labeling

📖 NOTE

- Team members receive emails for team labeling jobs. No email will be sent when you create a labeling team or add members to a labeling team. Additionally, after all samples are labeled, no email will be sent when you create a team labeling job.
- After a team labeling job is created, all unlabeled samples are assigned to annotators randomly and evenly.

Procedure

You can create multiple team labeling jobs for the same dataset and assign them to different members in the same team or to other labeling teams.

1. Log in to the ModelArts management console. In the left navigation pane, choose **Asset Management** > **Datasets**.

2. In the dataset list, select a dataset that supports team labeling, and click the dataset name to go to the **Dashboard** tab page of the dataset.

3. In the **Labeling Job** area, view existing labeling jobs of the dataset. Click **Create** to create a job.

Figure 4-5 Labeling Job
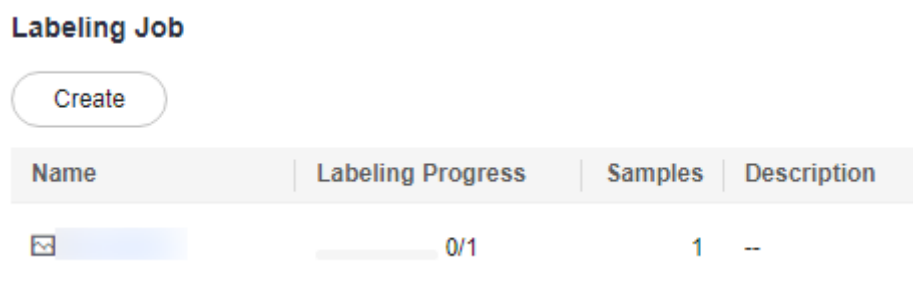
Alternatively, you can choose **Data Preparation** > **Label Data** and click **Create Labeling Job**.

4. In the displayed **Create Labeling Job** page, set parameters and click **Create**.

   - **Name**: Enter a job name.

   - **Labeling Scene**: Select the type of the labeling job.

   - **Label Set**: All existing labels and label attributes of the dataset are displayed.

   - **Team Labeling**: Click the button on the right and set the following parameters:

     - **Type**: Select a job type, **Team** or **Task Manager**.

     - **Select Team**: If **Type** is set to **Team**, select a team and members for labeling. The drop-down list displays the labeling teams and their members created by the current account.

     - **Select Task Manager**: If **Type** is set to **Task Manager**, select one **Team Manager** member from all teams as the task manager.

■ **Automatically synchronize new files to the team labeling task**:
New files in the dataset will be automatically synchronized to the
labeling job that has been started.

■ **Automatically load the intelligent labeling results to files that
need to be labeled**: Files are automatically labeled. Annotators can
then accept or modify the labels.

☐ NOTE

The process of loading auto labeling results to a team labeling job is as
follows:

● If you set **Type** to **Team**, you are required to create a team labeling task
before executing the job.

● If you set **Type** to **Task Manager**, select a team labeling job on the **My
Participations** tab page and click **Assign Task**.

Figure 4-6 Creating a team labeling job
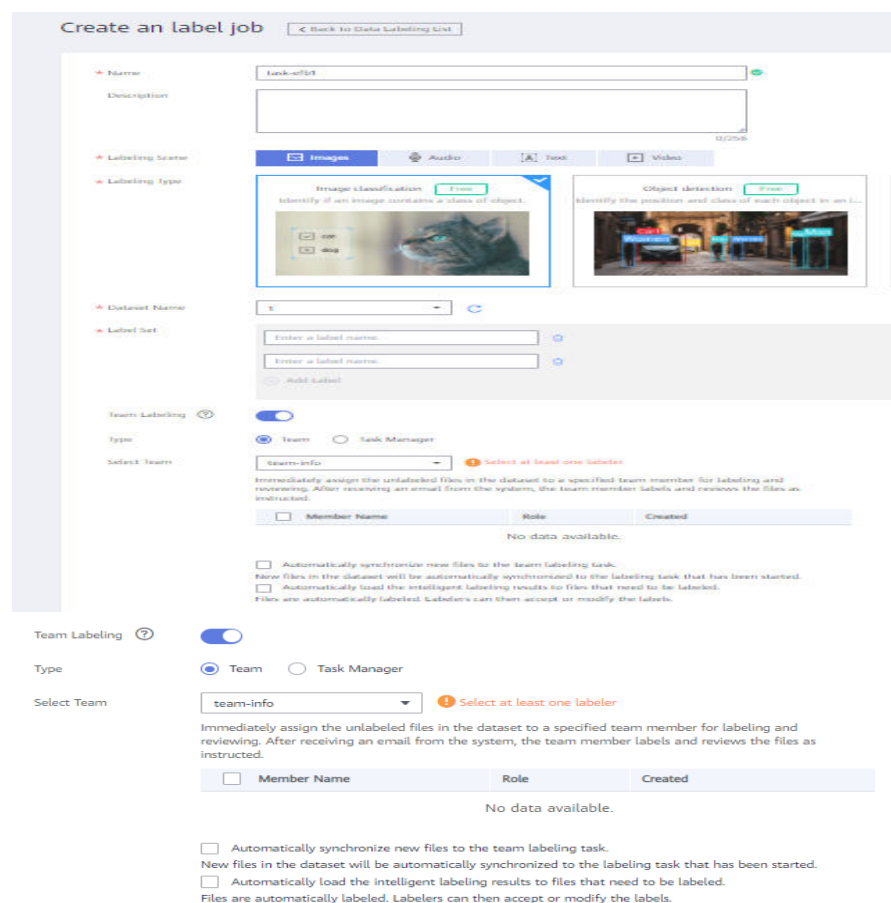
After the job is created, you can view the new job on the **My Creations**
tab page.

4.4 Logging In to ModelArts

After a team labeling job is created, team members can log in to the ModelArts
console to view the job.

1. After a labeling job is created, receive a labeling notification email as a team member to which the job is assigned.

2. Click the labeling job link in the email. The **Data Preparation** > **Data Labeling** > **My Participations** tab page on the ModelArts console is displayed. If you have not logged in to the console, log in to it first.

3. On the **My Participations** tab page, you can view your labeling jobs.

Figure 4-7 My Participations
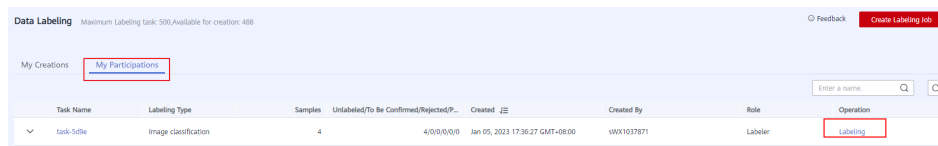
The following table lists the data labeling access addresses. If a team member has bound an email address, the team member can receive a job notification email and access the data labeling console using the address provided in the email.

Table 4-1 ModelArts labeling address

Region	URL of the data labeling console
CN-Hong Kong	https://console-intl.huaweicloud.com/modelarts/?region=ap-southeast-1&locale=en-us#/dataLabel?tabActive=labelConsole
AP-Singapore	https://console-intl.huaweicloud.com/modelarts/?region=ap-southeast-3&locale=en-us#/dataLabel?tabActive=labelConsole
AP-Bangkok	https://console-intl.huaweicloud.com/modelarts/?region=ap-southeast-2&locale=en-us#/dataLabel?tabActive=labelConsole

Upon your login, only the team labeling jobs and related data of the current user (the mailbox user) are displayed.

4.5 Starting a Team Labeling Job

After logging in to the data labeling page on the management console, you can click the **My Participations** tab page to view the assigned labeling job and click the job name to go to the labeling page. The labeling method varies depending on the labeling job type. For details, see the following:

- **Image Classification**
- **Object Detection**
- **Text Classification**
- **Named Entity Recognition**
- **Text Triplet**
- **Speech Paragraph Labeling**

On the labeling page, each member can view the images that are not labeled, to be confirmed, rejected, to be reviewed, approved, and accepted. Pay attention to the images rejected by the administrator and the images to be corrected.

If the Reviewer role is assigned for a team labeling job, the labeling result needs to be reviewed. After the labeling result is reviewed, it is submitted to the administrator for acceptance.

Figure 4-8 Labeling platform
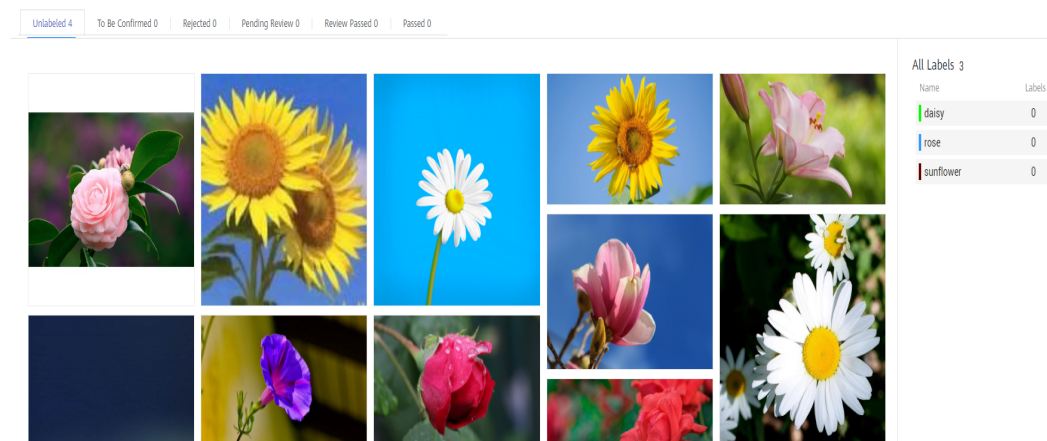
4.6 Reviewing Team Labeling Results

After team labeling is complete, the reviewer can review the labeling result.

1. Log in to the ModelArts management console. In the navigation pane, choose **Data Preparation** > **Label Data**. On the **Data Labeling** page, click **My Participations**. Locate the row containing the target labeling job and click **Review** in the **Operation** column to initiate the review.

   Figure 4-9 Initiating review
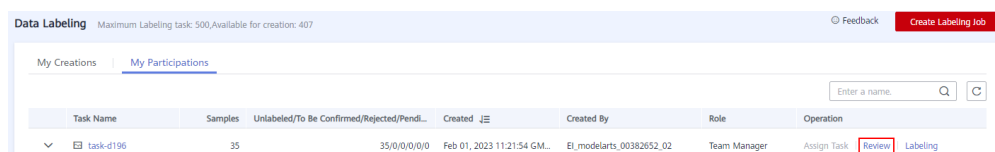
2. On the review page, check the samples that are not reviewed, reviewed, approved, or rejected.

   Figure 4-10 Labeling result review
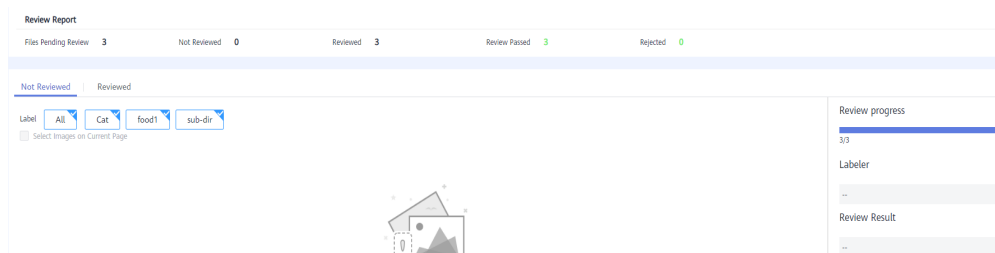
3. Choose **Confirm** or **Reject** on the right of the review page.

   If you choose **Confirm**, set **Rating** to **A**, **B**, **C**, or **D**. Option **A** indicates the highest score. If you choose **Reject**, enter the rejection reason in the text box.

4.7 Accepting Team Labeling Results

Task Acceptance (Administrator)

- **Initiating acceptance**

  After team members complete data labeling, the labeling job creator can initiate acceptance to check labeling results. The acceptance can be initiated only when a labeling member has labeled data. Otherwise, the acceptance initiation button is unavailable.

  a. Log in to the ModelArts management console. In the left navigation pane, choose **Data Preparation** > **Label Data**.

  b. On the **My Participations** tab page, click a team labeling job to go to its details page. Choose **Team Labeling** > **Accept** in the upper right corner.

  Figure 4-11 Initiating acceptance
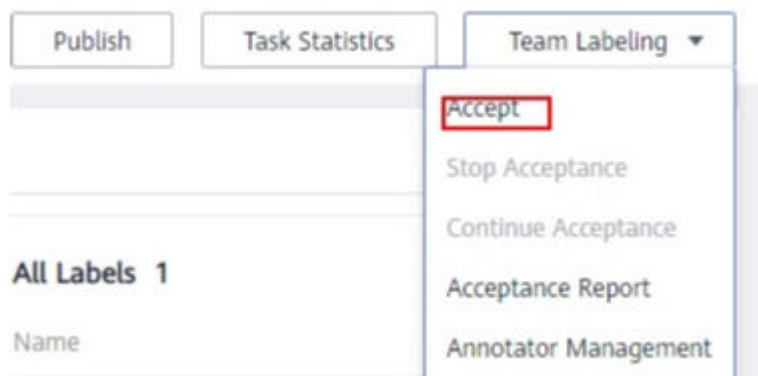
  c. In the displayed dialog box, set **Sample Policy** to **By percentage** or **By quantity**. Click **OK** to start the acceptance.

     By percentage: Sampling is performed based on a percentage for acceptance.

     By quantity: Sampling is performed based on quantity for acceptance.

Figure 4-12 Initiating acceptance

d. After the acceptance is initiated, an acceptance report is displayed on the console. In the **Acceptance Result** area on the right, click **Pass** or **Reject**.

If you click **Pass**, set **Rating** to **A**, **B**, **C**, or **D**. Option **A** indicates the highest score. See **Figure 4-14**. If you click **Reject**, enter your rejection reasons in the text box. See **Figure 4-15**.
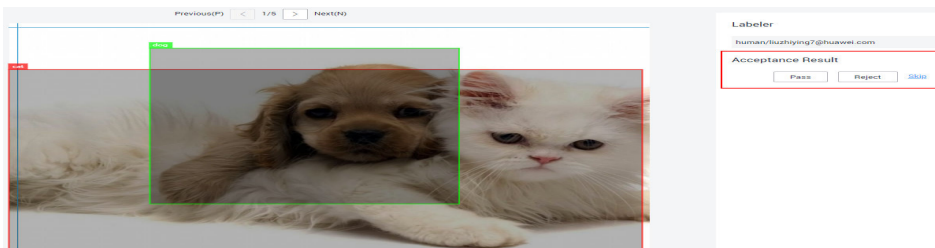
Figure 4-13 Viewing a real-time acceptance report
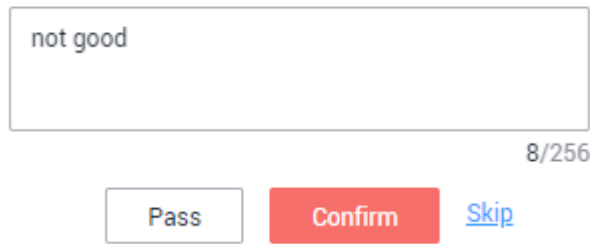

Figure 4-14 Pass

Figure 4-15 Reject

Acceptance Result

not good

8/256

Pass   Confirm   Skip
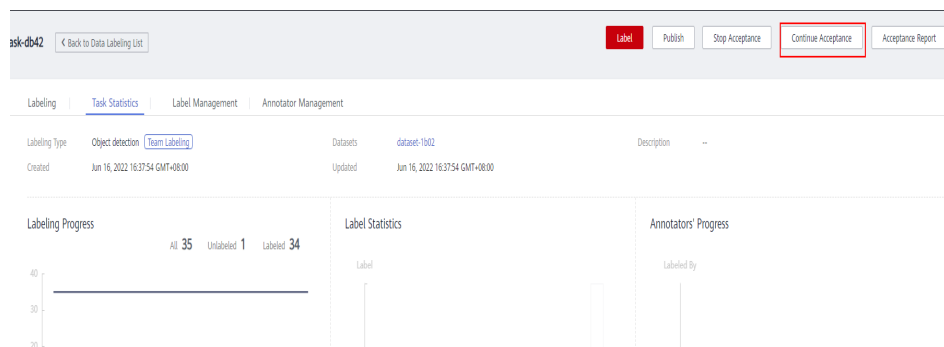
- **Continuing acceptance**

  You can continue accepting tasks whose acceptance is not completed. For tasks for which an acceptance process is not initiated, the **Continue Acceptance** button is unavailable.

  In the **Labeling Progress** pane on the **Task Statistics** tab page, click **Continue Acceptance** to continue accepting jobs. The **Real-Time Acceptance Report** page is displayed. You can continue to accept the images that are not accepted.

- **Finishing acceptance**

  After the continue acceptance is complete, click **Stop Acceptance** in the upper right corner. On the page that is displayed, view the acceptance status of the labeling job, such as the number of sampled files, configure parameters, and perform the acceptance. The labeling information is synchronized to the **Labeled** tab page of the labeling job only after the acceptance is complete.

  Once the labeled data is accepted, team members cannot modify the labeling information. Only the dataset creator can modify the labeling information.

Table 4-2 Parameters for finishing acceptance

Parameter	Description
Modifying Labeled Data	• **Not overwrite**: For the same data, do not overwrite the existing data with the labeling result of the current team. • **Overlays**: For the same data, overwrite the existing data with the labeling result of the current team. Overwritten data cannot be recovered.
Acceptance Scope	• **All passed**: All items, including the rejected ones will pass the review. • **All rejects**: All items, including the ones that have passed the review will be rejected. In this case, the passed items must be labeled and reviewed again in the next acceptance. • **All remaining items pass**: The rejected items are still rejected, and the remaining items will automatically pass the review. • **All remaining items rejects**: The selected items that have passed the review do not need to be labeled. All the selected items that have been rejected and the items that have not been selected must be labeled again for acceptance.

Figure 4-16 Finishing acceptance

Viewing an Acceptance Report

You can view the acceptance report of an ongoing or finished labeling job. Log in to the management console and choose **Data Preparation** > **Label Data**. On the **Data Labeling** page, select **My Creations** and click the name of a team labeling job. The job details page is displayed. In the upper right corner of the page, click **Acceptance Report**. In the displayed dialog box, view report details.

Figure 4-17 Viewing an acceptance report

Deleting a Labeling Job

After a job is accepted, delete it if the labeling job is no longer used. After a job is deleted, the labeling details that are not accepted will be lost. However, the original data in the dataset and the labeled data that has been accepted are still stored in the corresponding OBS bucket.