oduct is in a pre-release state and might change or have limited support. For more information, see the <u>product laun</u>
<u>_</u> (/products/#product-launch-stages).

This page describes how to prepare your training and test data so that AutoML Video Intelligence
Classification can create a custom video annotation model for you.

- AutoML Video Intelligence Classification supports the following video formats (shown below) for training your model or requesting a prediction (annotating a video). Maximum file size is 50GB (up to 3 hours in duration). Individual video files with malformed or empty timestamps in the container aren't supported.

    - .MOV

    - .MPEG4

    - .MP4

    - .AVI

- The training data should be as close as possible to the data on which predictions are to be made. For example, if your use case involves blurry and low-resolution videos (such as from a security camera), your training data should be composed of blurry, low-resolution videos. In general, you should also consider providing multiple angles, resolutions, and backgrounds for your training videos.

- AutoML Video Classification models can't generally predict labels that humans can't assign. So, if a human can't be trained to assign labels by looking at the video for 1-2 seconds, the model likely can't be trained to do it either.

- We recommend about 1000 training videos per label. The minimum per label is 10, or 50 for advanced models. In general it takes more examples per label to train models with multiple labels per video, and resulting scores are harder to interpret.

- The model works best when there are at most 100x more videos for the most common label than for the least common label. We recommend removing very low frequency labels.

- Consider including a `None_of_the_above` label and videos that don't match any of your defined labels. For example, for an animal dataset, include videos of animals outside of your labeled varieties, and label them as `None_of_the_above`. This can improve the accuracy of your model. Note that, while any label name will work, `None_of_the_above` is treated specially by the system.

The data in a dataset is divided into three datasets when training a model: a training dataset, a validation dataset, and a test dataset.

A training dataset is used to build a model. The model tries multiple algorithms and parameters while searching for patterns in the training data. As the model identifies patterns, it uses the validation dataset to test the algorithms and patterns. The best performing algorithms and patterns are chosen from those identified during the training stage.

After the best performing algorithms and patterns have been identified, they are tested for error rate, quality, and accuracy using the test dataset.

Both a validation and a test dataset are used in order to avoid bias in the model. During the validation stage, optimal model parameters are used, which can result in biased metrics. Using the test dataset to assess the quality of the model after the validation stage provides an unbiased assessment of the quality of the model.

To identify your training, testing, and validation data, use CSV files as described in the next section.

Once your files have been uploaded to Google Cloud Storage, you can create CSV files that list all of your training data and the category labels for that data. The CSV files can have any filenames, must be in the same bucket as your video files, must be UTF-8 encoded, and must end with a `.csv` extension.

There are four files that you can use for training and verifying your model:

File	Description
Model training file list	Contains paths to the train, test, and validate CSV files. This file is used to identify the locations of up to three separate CSV files that describe your training and testing data.

Here are some examples of the contents of the file list CSV file:

Example 1:

Example 2:

Training data	Used to train the model. Contains paths to video files, start and end times for video segments, and labels identifying the subject of the video segment. If you specify a training data CSV file, you must also specify a testing data CSV file.
Test data	Used for testing the model during the training phase. Contains paths to video files, start and end times for video segments, and labels identifying the subject of the video segment. If you specify a testing data CSV file, you must also specify a training data CSV file.
Unassigned data	Used for both training and testing the model. Contains paths to video files, start and end times for video segments, and labels identifying the subject of the video segment. Rows in the unassigned file are automatically divided into train and test data. 80% for training and 20% for testing. You can specify only an unassigned data CSV file without training and testing data CSV files. You can also specify only the training and testing data CSV files without an unassigned data CSV file.

The training, test, and unassigned files have one row for each video in the set you are uploading, with these columns in each row:

1. **The content to be categorized or annotated.** This field contains Google Cloud Storage URI for the video. Google Cloud Storage URIs are case-sensitive.

2. **A label that identifies how the video is categorized.** . Labels must start with a letter and only contain letters, numbers, and underscores. You can specify multiple labels for a video by adding multiple rows in the CSV file that each identify the same video segment, with a different label for each row.

3. **Start and end time of the video segment.** These two, comma-separated fields identify the start and end time of the video segment to analyze, in seconds. The start time must be less than the end time. Both values must be non-negative and within the time range of the video. For

example, `0.09845,1.3600555`. To use the entire content of the video, specify a start time of 0 and an end time of the full length of the video or "inf". For example, `0,inf`.

Here are some example rows for the CSV data files:

Single label:

Multi-label on the same video segment:

Using `inf` to indicate the end of a video:

For best results, you should include at least several hundred training video segments per label to create an accurate model. This number can vary depending on the complexity of your data.

You can also provide videos in the CSV data file without specifying any labels. You must then use the AutoML Video Classification UI (https://console.cloud.google.com/video-intelligence/introduction) to apply labels to your data before you train your model. To do so, you only need to provide the Cloud Storage URI for the video followed by three commas, as shown in the following example.

You do not need to specify validation data to verify the results of your trained model. AutoML Video Classification automatically divides the rows identified for training into training and validation data. 70% for training and 30% for validation.

Save the contents as a CSV file in your Google Cloud Storage bucket.

on errors with CSV

Using unicode characters in labels. E.g. Japanese characters are not supported.

Using spaces and non-alphanumeric characters in labels.

Empty lines.

Empty columns (lines with two successive commas).

Incorrect capitalization of Cloud Storage video paths.

Incorrect access control configured for your video files. Your service account should have read or greater access, or must be publicly-readable.

References to non-video files (such as PDF or PSD files). Likewise, files that are not video files but that have been ren with a video extension will cause an error.

URI of video points to a different bucket than the current project. Only videos in the project bucket can be accessed.

Non-CSV-formatted files.