

# Quotas & Limits

This document contains current restrictions and usage quotas on use of AutoML Vision. This page will be updated to reflect any changes to these restrictions and usage quotas.

These quotas apply to each AutoML Vision project and are shared across all applications and IP addresses using a given project.

You can edit your quotas up to their maximum values by selecting **Edit Quotas** from the **Quotas** page of the [Google Cloud Platform Dashboard](#)

([https://console.cloud.google.com/project/\\_/quotas?service=automl.googleapis.com](https://console.cloud.google.com/project/_/quotas?service=automl.googleapis.com)). Edit the quotas for the **Cloud AutoML API**. To request an increase above the maximum quota, edit your quota with your requested increase and justification and submit your update. You will be notified when your request has been received. You might be contacted for more information regarding your request. After your request has been reviewed, you will be notified whether it has been approved or denied.

AutoML Vision has quotas associated with creating models, and making requests to the models.

## Modeling quotas

Activity	Default quota
Model creation requests per minute	10
Concurrent model training requests	5

## Request quotas

The current usage quotas for AutoML Vision for non-prediction requests is 600 per minute per project. It applies collectively to all deployed models.

### Online prediction requests

For Models deployed after the [Beta Refresh](#)

([https://cloud.google.com/vision/automl/docs/release-notes#October\\_09\\_2019](https://cloud.google.com/vision/automl/docs/release-notes#October_09_2019)), the number of online prediction per minute for AutoML Vision are based on how many nodes are deployed. You can adjust the number of nodes when you [deploy your model](#)

([https://cloud.google.com/vision/automl/docs/deploy#automl\\_vision\\_classification\\_deploy\\_model\\_node\\_count-web-integrated](https://cloud.google.com/vision/automl/docs/deploy#automl_vision_classification_deploy_model_node_count-web-integrated))

. However, the total number of nodes across all the deployed models in the project is subject to the quota below.

Activity	Default quota
The number of deployed model nodes	10

### Offline batch prediction requests

The current usage quotas for AutoML Vision are 5 *concurrent* batch prediction requests per minute per project. These quotas apply collectively to all models.

## Limits

Type of limit	Value
Image file size	Maximum: 30MB
Images per dataset	Maximum: 1,000,000
Labels per dataset	Minimum: 2 Maximum: 5000

Type of limit	Value
Images per label	Minimum: 10 Recommended: 1000
Batch input CSV file size	Maximum: 100MB
Number of images in batch input	Maximum: 1,000,000

*Last updated December 4, 2019.*