Image Reader

Updated: July 8, 2015

Loads images from Azure BLOB Storage into a dataset

Category: OpenCV Library Modules (https://msdn.microsoft.com/en-us/library/azure/dn905946.aspx)

Module Overview

You can use the **Image Reader** module to get images from Azure Blob storage and create an image dataset from them.

When you read an image from blob storage using **Image Reader**, the image is represented as a series of numeric values for the red, green, and blue channels, along with the image file name. You can then pass this dataset to the Score Model (https://msdn.microsoft.com/en-us/library/azure/dn905995.aspx) module, and connect a pre-trained image classification model to predict the image type.

The Pretrained Cascade Image Classification (https://msdn.microsoft.com/en-us/library/azure/dn913079.aspx) provided in Azure Machine Learning currently supports recognition of faces in frontal view; other image classifiers are not yet available.

Note
See the Technical Notes section for a list of the requirements that apply to images. Note also that you cannot use image datasets with Machine Learning / Train (https://msdn.microsoft.com/en-us/library/azure/dn905846.aspx) or Cross-Validate Model (https://msdn.microsoft.com/en-us/library/azure/dn905852.aspx).

How Images are Processed

Image Reader determines the type of an image by reading the first few bytes of the content, not by the file extension.

The following image formats are supported:

- Windows bitmap files: *.bmp, *.dib
- JPEG files: *.jpeg, *.jpg, *.jpe

- JPEG 2000 files: *.jp2
- Portable Network Graphics: *.png
- Portable image format: *.pbm, *.pgm, *.ppm
- Sun Raster: *.sr, *.ras
- TIFF files: *.tiff, *.tif

Each row of the output dataset contains data from one image, ordered as follows:

- The first column contains image names.
- All other columns contain flattened data from the red, green, and blue color channels, in that order.

  Therefore, depending on the color depth of the image and the image format, there could be many thousands of columns for a single image.
- The transparency channel is ignored.
- In the output dataset, the rows are sorted alphabetically by image name.

How to Configure Image Reader

This example assumes that you have uploaded multiple images to an Azure blob storage, in a container designated for that purpose only.

Each image must be fairly small and have the same dimensions and color channels.

1. Add the to your experiment.
2. Add the Pretrained Cascade Image Classification (https://msdn.microsoft.com/en-us/library/azure/dn913079.aspx) and the Score Model (https://msdn.microsoft.com/en-us/library/azure/dn905995.aspx) module.
3. Configure the **Image Reader** as follows:
   - Specify the location of the images by specifying the authentication method, private or public.

     If the image set is in a blob that has been configured for public access through Shared Access Signatures (https://azure.microsoft.com/documentation/articles/storage-dotnet-shared-access-signature-part-1/)(SAS), type the URL to the container that holds the images.

     If the images are stored in a private account in Azure storage, select **Account**, and then type the account name as it appears in the management portal.

Paste in the primary or secondary account key, which is provided in the Azure portal.

For **Path to container**, type just the container name.

4. Connect the output of **Image Reader** to the Score Model (https://msdn.microsoft.com/en-us/library/azure/dn905995.aspx) module.

5. Run the experiment.

6. To view the results of the experiment, we recommend that you add the Project Columns (https://msdn.microsoft.com/en-us/library/azure/dn905883.aspx) module, and select only these columns:

    - Image Name
    - Scored Labels
    - Scored Probabilities

Technical Notes

The following requirements apply to images processed by the **Image Reader** module:

- All images must be the same shape.
- All images must have the same color channels. For example, you cannot mix grayscale images with RBG images.
- There is a limit of 65536 pixels per image. However, the number of images is not limited.
- If you specify a blob container as the source, the container must not contain blobs. Ensure that the container contains only images before running the module.

Module Parameters

Name	Range	Type	Default	Description
Please specify authentication type	List	AuthenticationType	Account	Public or Shared Access Signature (SAS) URI or user credentials

URI	Any	String		Uniform Resource Identifier with SAS or public access
Account name	Any	String		Name of the Azure Storage account
Account key	Any	SecureString		Key associated with the Azure Storage account
Path to container, directory or blob	Any	String		Path to blob or name of table

Output

Name	Type	Description
Results dataset	Data Table (https://msdn.microsoft.com/en-us/library/azure/dn905851.aspx)	Dataset with downloaded images

Exceptions

For a list of all error codes, see Machine Learning Module Error Codes (https://msdn.microsoft.com/en-us/library/azure/dn905910.aspx).

Exception	Description
Error 0003 (https://msdn.microsoft.com/en-us/library/azure/dn906003.aspx)	Exception occurs if one or more inputs are null or empty.
Error 0029 (https://msdn.microsoft.com/en-us/library/azure/dn905890.aspx)	Exception occurs when invalid URI is passed.
Error 0009 (https://msdn.microsoft.com/en-us/library/azure/dn906034.aspx)	Exception occurs if the Azure storage account name or container name is specified incorrectly.
Error 0015 (https://msdn.microsoft.com/en-us/library/azure/dn906021.aspx)	Exception occurs if the database connection has failed.

Error 0030 (https://msdn.microsoft.com/en-us/library/azure/dn906053.aspx)	Exception occurs when it is not possible to download a file.
Error 0049 (https://msdn.microsoft.com/en-us/library/azure/dn906049.aspx)	Exception occurs when it is not possible to parse a file.
Error 0048 (https://msdn.microsoft.com/en-us/library/azure/dn905937.aspx)	Exception occurs when it is not possible to open a file.

See Also

Pretrained Cascade Image Classification (https://msdn.microsoft.com/en-us/library/azure/dn913079.aspx)
A-Z List of Machine Learning Studio Modules (https://msdn.microsoft.com/en-us/library/azure/dn906033.aspx)