IBM PowerAI Vision Tutorial

NCSA HAL Tutorial Series
Ke Xu
What is PowerAI Vision?

- PowerAI Vision (also called IBM Visual Insights) is a GUI platform that empowers subject matter experts to label, train, and deploy deep learning vision models, without coding or deep learning expertise.
- It is a streamline processes to label, train, monitor and deploy models using a unified interface.
Supported vision tasks

- The following three computer vision tasks are supported:

Select type of training

<table>
<thead>
<tr>
<th>Image classification</th>
<th>Object detection</th>
<th>Action detection</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.jpg" alt="Image of a lion" /></td>
<td><img src="image2.jpg" alt="Image of a kitchen" /></td>
<td><img src="image3.jpg" alt="Image of baseball players" /></td>
</tr>
<tr>
<td>Classify an entire image according to its visual content based on the relationship of nearby pixels.</td>
<td>Detect instances of particular classes within images or videos using rectangular or polygonal bounding boxes.</td>
<td>Classify brief, temporal actions in video. Actions of interest will be located in both space and time.</td>
</tr>
</tbody>
</table>
Supported vision models

Only the following models are supported by PowerAI Vision:

- **Image classification:**
  - GoogLeNet

- **Object detection:**
  - Faster R-CNN
  - Tiny Yolov2
  - Detectron
  - Single Shot Detector (SSD)

- **Action detection:**
  - Structured Segment Network (SSN)
Account login

- Contact HAL admin to create a PowerAI Vision account
- Open a terminal and type in:
  - `ssh -L 8443:hal16:443 <hal_urs_name>@hal.ncsa.illinois.edu`
  - You will enter the HAL login node, leave it there
- Open a browser and enter the following url:
  - `https://localhost:8443/visual-insights`
  - You may need to manually enter the url again since sometimes your browser will correct your url and delete the 8843 port number
  - You may be prompt for insecure connection, just click “advanced” and accept the insecurity
- You need to enter your PowerAI Vision username and password at the login page the 1st time you login
Account login

● After login, you will see the homepage:

Welcome to IBM Visual Insights

Create Dataset
Start by adding images and video files to a data set.

Prepare Data
Label objects in images and video frames, then use auto-labeling to finish adding labels to the whole data set; assign categories to images or videos, or label actions in videos.

Train Model
Select a few custom options to create your model.

Deploy Model
Deploy the trained model and receive an API link for an inference device.
Object detection in videos

PowerAI Vision is a highly encapsulated pipeline to train and deploy models. The same pipeline is used for different vision tasks. In this tutorial, we pick object detection as an example.

For image classification and action detection, the pipeline is the same except for different input data preparation, which we will introduce after the object detection example.
Create dataset

1. Download sample video from: https://ibm.ent.box.com/s/gf52cgp455n19oyn71dhwqd1unvimlzj
2. Click Data Sets in the navigation bar to open the Data Sets page.
3. From the Data set page, click the icon and name the data set “object-detect-video”.
4. To add a video to the data set, click the data set and click Import file or drag the video to the + area.

* You cannot navigate away from the IBM Visual Insights page or refresh until the upload completes. You can navigate to different pages within IBM Visual Insights during the upload.
Create dataset

IBM Visual Insights

Data Set / object-detect-video

- Total files: 1
- Matching files: 1
- Selected files: 0

- Train model
- Augment data
- Auto label
- Export data set

- Assign category
- Label objects
- Label actions

- Select

Drop files here

No objects added

Uncategorized

Import files

0 frames
Data augmentation

1. You can apply data augmentation by clicking on the “Augment data” button on top of your dataset homepage

2. The following data augmentation techniques are supported:
   a. Image blurring
   b. Change brightness and contrast
   c. Image sharpening
   d. Image rotation

3. A new data set will be created that contains the original data as well as additional images created by using these filters.
Data augmentation

Augment data

Select the filters to use to augment your data. A new data set will be created that contains the original data as well as additional images created by using these filters. Visit the Knowledge Center for more information about data augmentation.

Currently selected: 3
New creations: 15
New total: 19

Blur
Color
Sharpen

Rotate sample
Rotate

Cancel
Continue
Labeling objects in a video

1. Select the video from your data set and select Label Objects.
2. Click Auto capture frames and specify a value for Capture Interval (Seconds) that will result in at least five frames. We will select this option and specify 10 seconds.
3. If you used Auto capture frames, verify that there are enough of each object type in the video frames.
4. Create new object labels for the data set by clicking Add new by the Objects list. Enter Car, click Add.
5. Label the objects in the frames:
   a. Select frame by frame
   b. Select the correct object label, for example, "Car".
   c. If an image has more than one object, you must label all objects.
   d. Try to draw bboxes as close to the objects as possible
Labeling objects in a video
Training a model

1. From the Data set page, click Train.
2. Fill out the fields on the Train Data set page, ensuring that you select Object Detection. We will choose Tiny Yolov2.
3. Click Train.

* you can only select from the models that PowerAI Vision supports

* since PowerAI Vision is a browser based API, you must leave your browser, connection, and computer on during training, otherwise the training will be interrupted.
Training a model

Object detection
Tiny YOLO v2

2.2 HOURS LEFT

Initialized
04/11/2020, 10:47 PM

Stop training

Training options

Core ML

Max iteration
40000

Ratio
0.8
Training a model

Loss VS Iteration

- Blue line: Test Loss
- Yellow line: Train Loss
Finished training

- After the model has been trained, you can see the final test accuracy, hyperparameters, and learning curve.
- You can also download the CoreML file for deployment on mobile/embedded devices; CoreML format is only supported for a few models.
- After the model has been trained, you can choose to deploy the model or export in PowerAI Vision format.
Finished training

Trained Model / Traffic camera_model

Trained models are created from prepared data sets.
This model can be validated and deployed prior to exporting for PowerAI Inference Engine (PIE).

Object detection
Data set: Traffic camera

Created
04/12/2020, 3:26 AM
By: kexu6

Model hyperparameters

Model Assets

Core ML files
60 MB
Deploying a model

1. Click Models from the menu.
2. Select the model you created in the previous section and click Deploy.
3. Specify a name for the model, and click Deploy.
4. Double-click the deployed model to get the API endpoint and test other videos or images against the model.

- You can choose to upload an image or provide an url to the image to test the model
- You can adjust the confidence threshold for the model
Deploying a model
Data preparation for image classification

- There must be at least two categories.
- Each category must have at least five images.
- If an image has more than one object, you must label all objects.

1. Create new object labels for the data set by clicking Add new by the Objects list.
2. Open an image.
3. Choose Box or Polygon from the bottom left and draw it around the object you want to classify
4. Assign the corresponding object label to the box you just draw
Data preparation for image classification

Label objects / image-classification
Data preparation for action detection

- Each action label must be in the range of 5 - 1000 frames.
- At least 10 instances of each action tag in the data set are recommended.
- Transcode videos with FPS greater than 30 down to 30 FPS.
- Crop the video so that actions should take up a large part of the frame.

1. Open the data set that contains the video you want to label.
2. Create an action tag in the data set by expanding Actions on the left and clicking Add action.
3. Select the video and click Label actions.
4. Enter the start and end time stamps for an action.
5. Specify an action name, either by selecting an existing tag or by entering the name for a new tag, then click Create action.
Data preparation for action detection
How to train using a public dataset?

1. Click Data Sets in the navigation bar to open the Data Sets page. Create a new data set and give it a name like before.

2. From the COCO download site, click 2017 Train images to download the train2017.zip file.

3. Create a new file that contains just the images that you want from train2017 by running a command such as the following:

   `ls train2017 | grep jpg | head -20000 >/tmp/flist`

4. From the COCO download site, click 2017 Train/Val annotations to download the annotations_trainval2017.zip file.

5. From annotations_trainval2017.zip, extract the annotations/instances_train2017.json file, which is the COCO annotation file for object detection.
How to train using a public dataset?

6. Add annotations/instances_train2017.json to the file of images that you created and compress them into a zip file.

7. From your new data set, click Import file and select the zip file you just created.
Thank you!