Practice-Object Detection

Label Formats for Bounding Box

	file type
KITTI	txt
Pascal VOC	XML
Imagenet	XML
YOLO	XML
MS COCO	json

PASCAL VOC dataset

```
<annotation>
      <filename>2012_001003.jpg</filename>
      <folder>VOC2012</folder>
      <object>
            <name>person</name>
            <actions>
                  <jumping>0</jumping>
                  <other>0</other>
                  <phoning>0</phoning>
                  <playinginstrument>1</playinginstr
                  <reading>0</reading>
                  <ridingbike>0</ridingbike>
                  <ridinghorse>0</ridinghorse>
                  <running>0</running>
                  <takingphoto>0</takingphoto>
                  <usingcomputer>0</usingcomputer>
                  <walking>0</walking>
            </actions>
            <bndbox>
                  <xmax>420</xmax>
                  <xmin>214</xmin>
                  <ymax>320</ymax>
                  <ymin>40</ymin>
            </bndbox>
            <difficult>0</difficult>
            <pose>Unspecified</pose>
            <point>
                  <x>281</x>
                  <y>163</y>
            </point>
      </object>
```

KITTI Dataset

http://www.cvlibs.net/datasets/kitti/eval_object.php

The KITTI Vision Benchmark Suite

A project of Karlsruhe Institute of Technology and Toyota Technological Institute at Chicago

Karlsruhe Institute of Technology

home | setup | stereo | flow | scene flow | odometry | object | tracking | road | semantics | raw data | submit results | jobs

Andreas Geiger (MPI Tübingen) | Philip Lenz (KIT) | Christoph Stiller (KIT) | Raquel Urtasun (University of Toronto)

Object Detection Evaluation 2012

- Download left color images of object data set (12 GB)
- Download right color images, if you want to use stereo information (12 GB)
- Download the 3 temporally preceding frames (left color) (36 GB)
- Download the 3 temporally preceding frames (right color) (36 GB)
- Download Velodyne point clouds, if you want to use laser information (29 GB)
- Download camera calibration matrices of object data set (16 MB)
- Download training labels of object data set (5 MB)

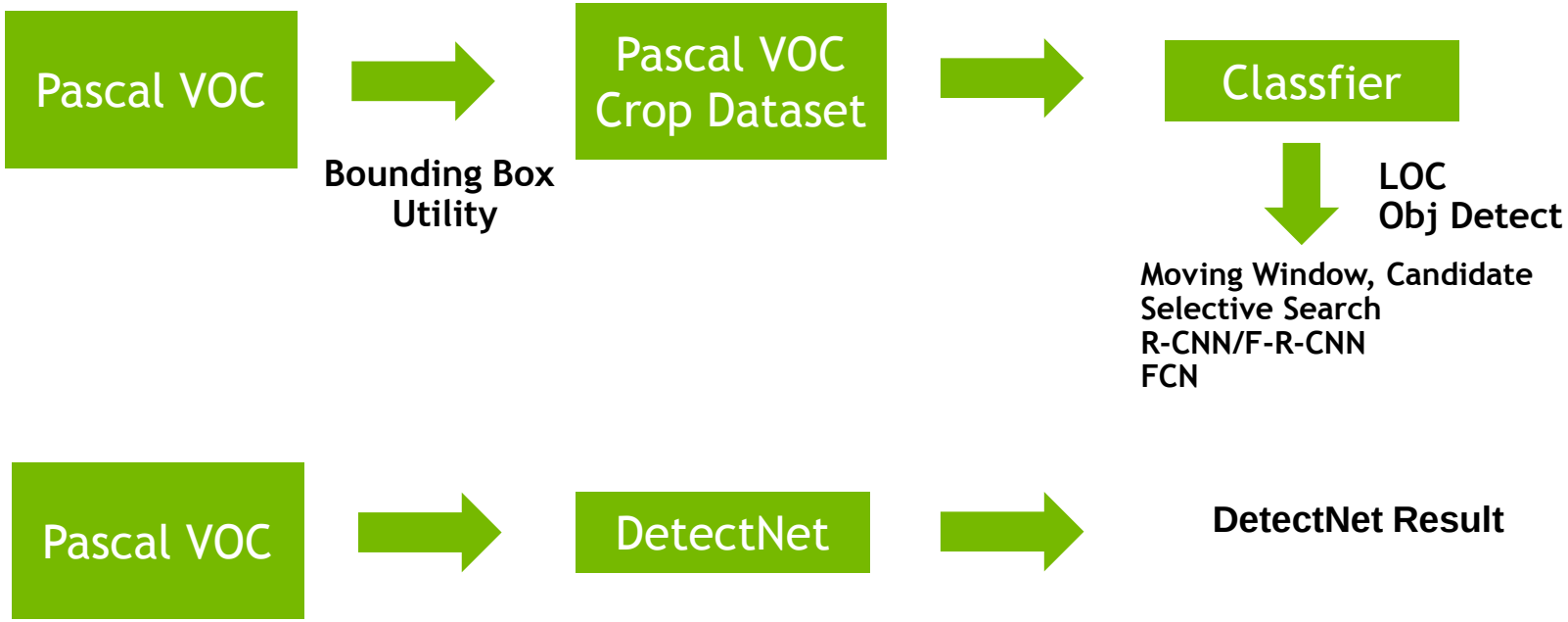
KITTI dataset

002152.Png 1242x375pixel

```
002152.txt
Car 0.00 0 -1.78 701.27 177.10 845.43 278.19 1.51 1.59 4.28 2.70 1.61 13.06 -1.58
Car 0.00 0 0.57 60.00 186.16 277.52 263.57 1.50 1.52 3.65 -9.18 1.80 15.10 0.04
Car 0.49 1 -2.06 845.05 187.60 1241.00 374.00 1.43 1.51 3.33 3.02 1.57 5.38 -1.57
Car 0.00 2 -1.72 681.90 173.99 763.48 239.51 1.56 1.57 3.78 2.86 1.62 19.25 -1.57
Car 0.00 0 1.75 397.79 184.95 488.42 246.25 1.47 1.71 3.85 -4.48 1.82 19.70 1.53
Car 0.00 1 1.74 447.34 175.36 513.59 229.68 1.73 1.65 3.70 -4.45 1.84 25.03 1.56
Car 0.00 1 -1.67 656.66 175.59 710.36 218.98 1.54 1.66 3.79 2.73 1.67 27.78 -1.57
Car 1.00 0 -2.71 1161.40 213.98 1241.00 374.00 1.45 1.61 3.51 2.93 1.59 1.06 -1.57
Car 0.00 2 1.69 487.45 177.08 537.11 219.72 1.68 1.63 3.94 -4.13 1.88 30.75 1.56
Van 0.00 1 -1.64 638.73 148.40 689.27 201.78 2.85 2.46 5.85 3.01 1.58 41.78 -1.57
Car 0.00 0 -1.70 535.28 178.56 559.18 193.94 1.34 1.69 3.60 -5.76 1.88 65.74 -1.79
DontCare -1 -1 -10 560.02 172.77 632.76 193.12 -1 -1 -1 -1000 -1000 -1000 -10
```

```
Type : Van
Truncated :
Occluded : 1 partial
Alpha : angle
BB : Xmin,Ymin,Xmax,Ymax
Dimensions :
Location :
Rotation :
Score :
```

Pipeline for Object Detection



Moving Windows (simple deployment)

```
MODEL_FILE = JOB_DIR + MODEL_JOB_NUM + '/deploy.prototxt'
PRETRAINED = JOB_DIR+ MODEL_JOB_NUM + '/snapshot_iter_270.caffemodel'

caffe.set_mode_gpu()
# Initialize the Caffe model using the model trained in DIGITS
net = caffe.Classifier(MODEL_FILE, PRETRAINED,
                       channel_swap=(2,1,0),
                       raw_scale=255,
                       image_dims=(256, 256))


start = time.time()
for i in range(0,rows):
    for j in range(0,cols):
        grid_square = input_image[i*256:(i+1)*256,j*256:(j+1)*256]
        # subtract the mean image
        grid_square -= mean_image
        # make prediction
        prediction = net.predict([grid_square])
        detections[i,j] = prediction[0].argmax()
end = time.time()
```

FCN (segmentation) to detect object

```
net = caffe.Net(MODEL_FILE,PRETRAINED,caffe.TEST)
net.blobs['data'].reshape(1, 3, input_image.shape[0], input_image.shape[1])
net.reshape()
transformer = caffe.io.Transformer({'data': net.blobs['data'].data.shape})
transformer.set_transpose('data', (2,0,1))
transformer.set_channel_swap('data', (2,1,0))
transformer.set_raw_scale('data', 255.0)
```

```
my_cmap = copy.copy(plt.cm.get_cmap('jet')) # get a copy of the jet color map
my_cmap.set_bad(alpha=0) # set how the colormap handles 'bad' values

# Feed the whole input image into the model for classification
out = net.forward(data=np.asarray([transformer.preprocess('data', input_image)]))

# Create an overlay visualization of the classification result
im = transformer.deprocess('data', net.blobs['data'].data[0])
classifications = out['softmax'][0]
classifications = imresize(classifications.argmax(axis=0),input_image.shape,interp='bilinear').astype('float')
classifications[classifications==0] = np.nan
```

KITTI Dataset

http://www.cvlibs.net/datasets/kitti/eval_object.php

The KITTI Vision Benchmark Suite

A project of Karlsruhe Institute of Technology and Toyota Technological Institute at Chicago

Karlsruhe Institute of Technology

home | setup | stereo | flow | scene flow | odometry | object | tracking | road | semantics | raw data | submit results | jobs

Andreas Geiger (MPI Tübingen) | Philip Lenz (KIT) | Christoph Stiller (KIT) | Raquel Urtasun (University of Toronto)

Object Detection Evaluation 2012

- Download left color images of object data set (12 GB)
- Download right color images, if you want to use stereo information (12 GB)
- Download the 3 temporally preceding frames (left color) (36 GB)
- Download the 3 temporally preceding frames (right color) (36 GB)
- Download Velodyne point clouds, if you want to use laser information (29 GB)
- Download camera calibration matrices of object data set (16 MB)
- Download training labels of object data set (5 MB)

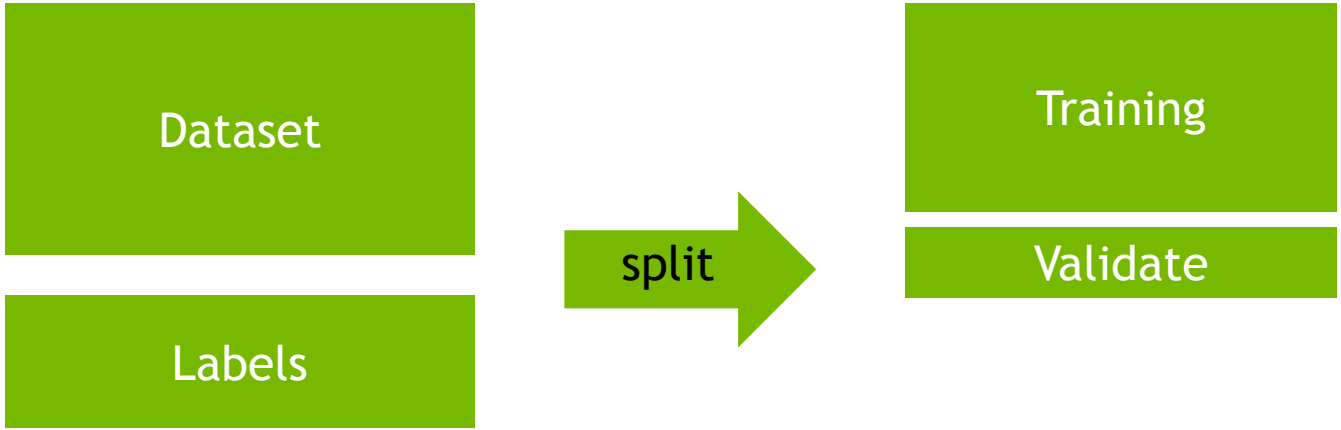
KITTI dataset

002152.Png 1242x375pixel

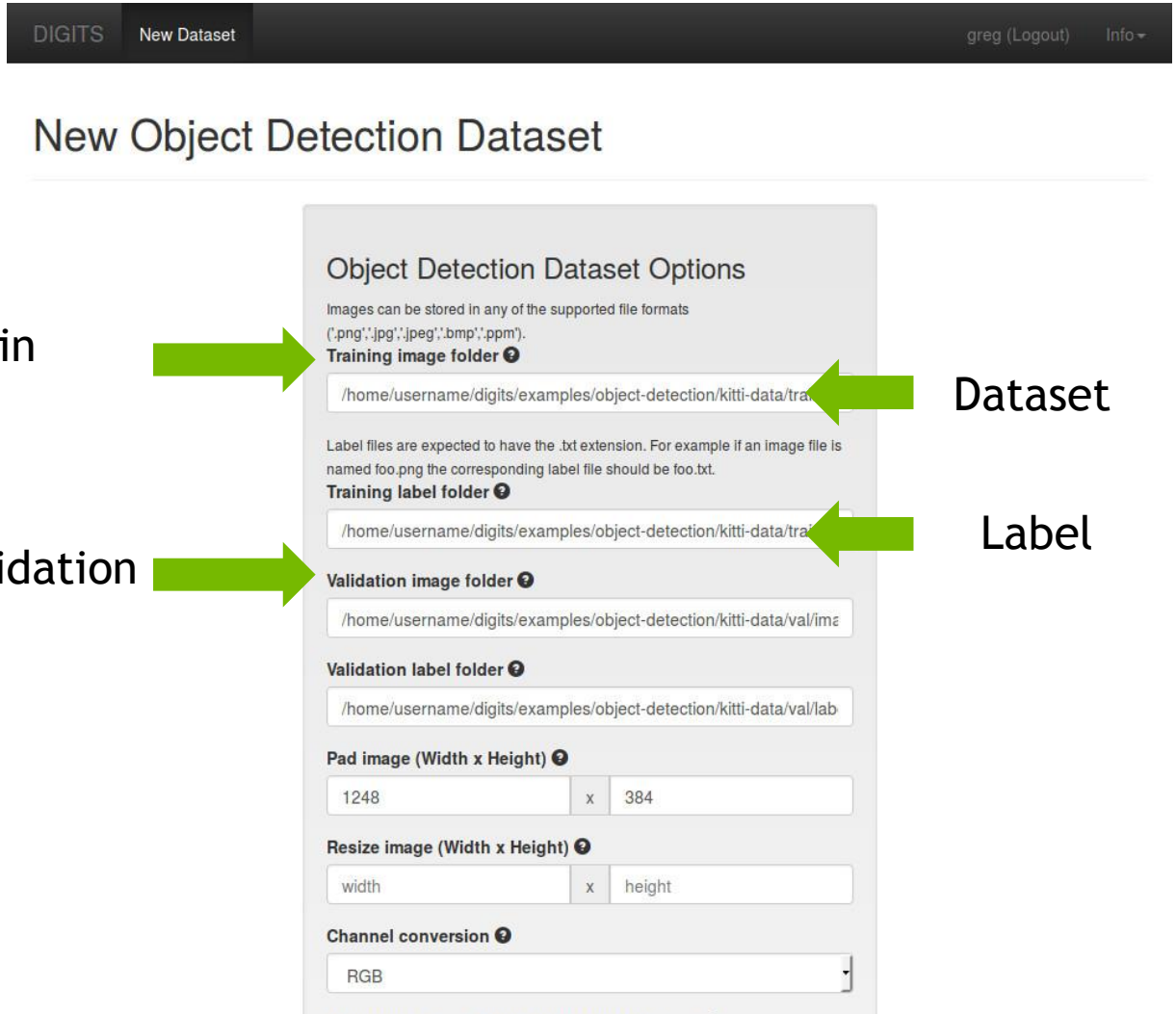
```
002152.txt
Car 0.00 0 -1.78 701.27 177.10 845.43 278.19 1.51 1.59 4.28 2.70 1.61 13.06 -1.58
Car 0.00 0 0.57 60.00 186.16 277.52 263.57 1.50 1.52 3.65 -9.18 1.80 15.10 0.04
Car 0.49 1 -2.06 845.05 187.60 1241.00 374.00 1.43 1.51 3.33 3.02 1.57 5.38 -1.57
Car 0.00 2 -1.72 681.90 173.99 763.48 239.51 1.56 1.57 3.78 2.86 1.62 19.25 -1.57
Car 0.00 0 1.75 397.79 184.95 488.42 246.25 1.47 1.71 3.85 -4.48 1.82 19.70 1.53
Car 0.00 1 1.74 447.34 175.36 513.59 229.68 1.73 1.65 3.70 -4.45 1.84 25.03 1.56
Car 0.00 1 -1.67 656.66 175.59 710.36 218.98 1.54 1.66 3.79 2.73 1.67 27.78 -1.57
Car 1.00 0 -2.71 1161.40 213.98 1241.00 374.00 1.45 1.61 3.51 2.93 1.59 1.06 -1.57
Car 0.00 2 1.69 487.45 177.08 537.11 219.72 1.68 1.63 3.94 -4.13 1.88 30.75 1.56
Van 0.00 1 -1.64 638.73 148.40 689.27 201.78 2.85 2.46 5.85 3.01 1.58 41.78 -1.57
Car 0.00 0 -1.70 535.28 178.56 559.18 193.94 1.34 1.69 3.60 -5.76 1.88 65.74 -1.79
DontCare -1 -1 -10 560.02 172.77 632.76 193.12 -1 -1 -1 -1000 -1000 -1000 -10
```

```
Type : Van
Truncated :
Occluded : 1 partial
Alpha : angle
BB : Xmin,Ymin,Xmax,Ymax
Dimensions :
Location :
Rotation :
Score :
```

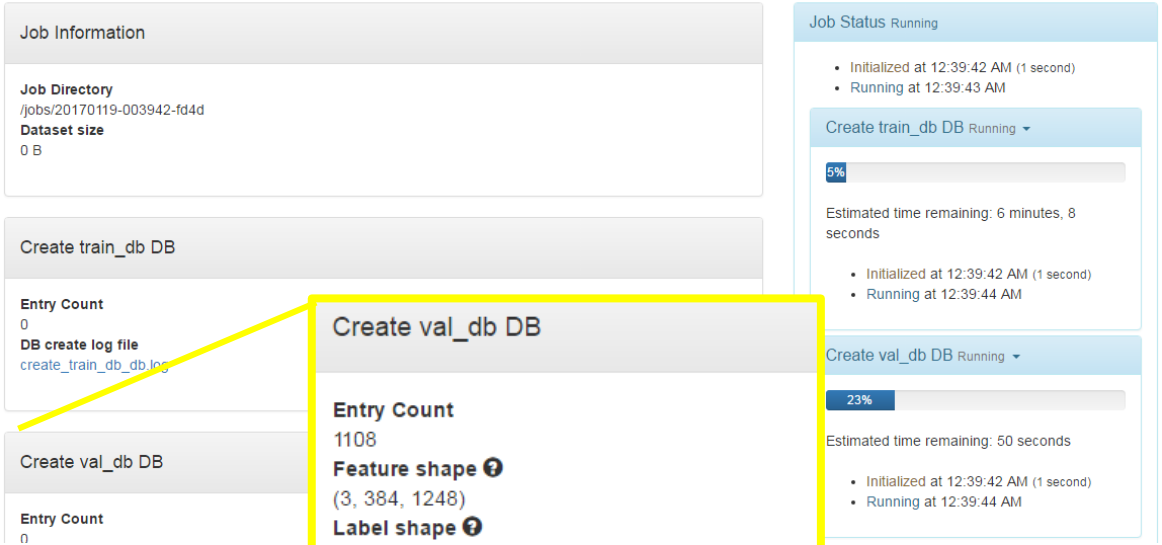
Prepare dataset

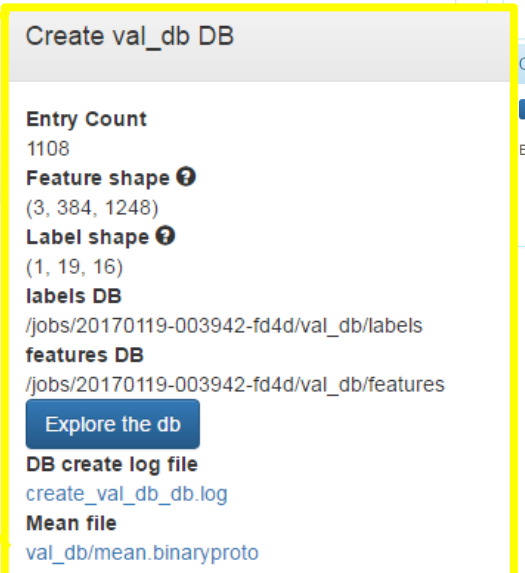



```
$ ./prepare_kitti_data.py
Extracting zipfiles ...
Unzipping data_object_label_2.zip ...
Unzipping data_object_image_2.zip ...
Unzipping devkit_object.zip ...
Calculating image to video mapping ...
Splitting images by video ...
Creating train/val split ...
Done.
```


DIGITS
New Dataset
greg (Logout)
Info
New Object Detection Dataset
Object Detection Dataset Options
Images can be stored in any of the supported file formats ('.png','.jpg','.jpeg','.bmp','.ppm').
Training image folder
/home/username/digits/examples/object-detection/kitti-data/tra
Label files are expected to have the .txt extension. For example if an image file is named foo.png the corresponding label file should be foo.txt.
Training label folder
/home/username/digits/examples/object-detection/kitti-data/tra
Validation image folder
/home/username/digits/examples/object-detection/kitti-data/val/ima
Validation label folder
/home/username/digits/examples/object-detection/kitti-data/val/lab
Pad image (Width x Height)
1248
x
384
Resize image (Width x Height)
width
x
height
Channel conversion
RGB
Train
Validation
Dataset
Label

Processing to generate LMDB





Exploring kitti-data (/jobs/20170119-003942-fd4d/train_db/features) images

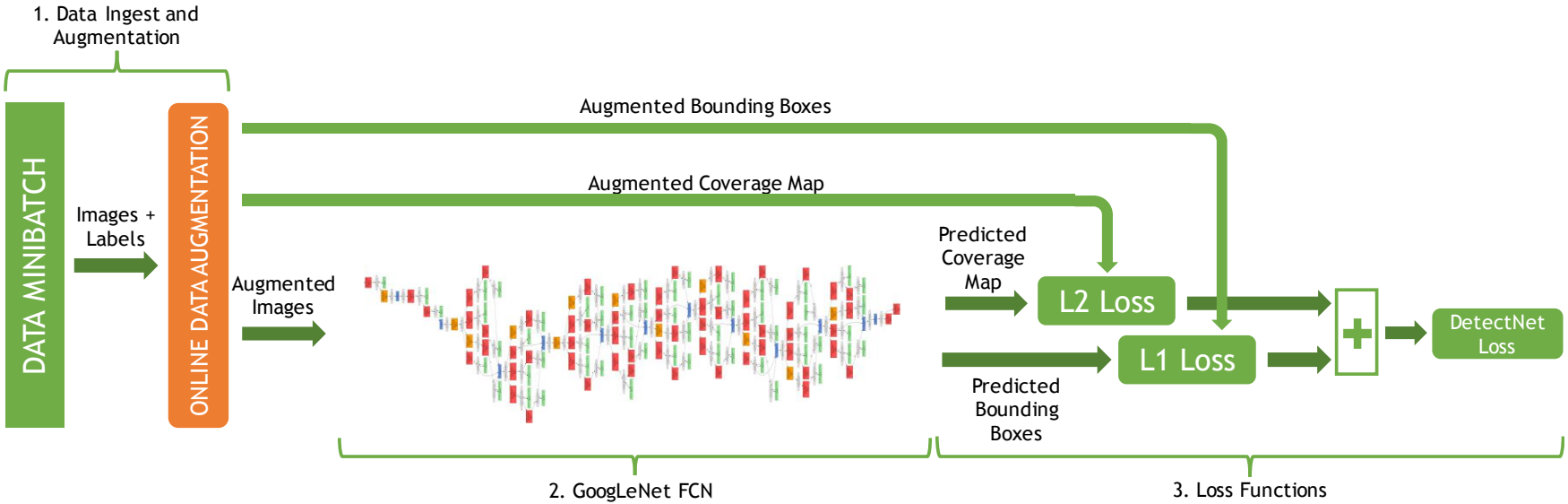
Show all images

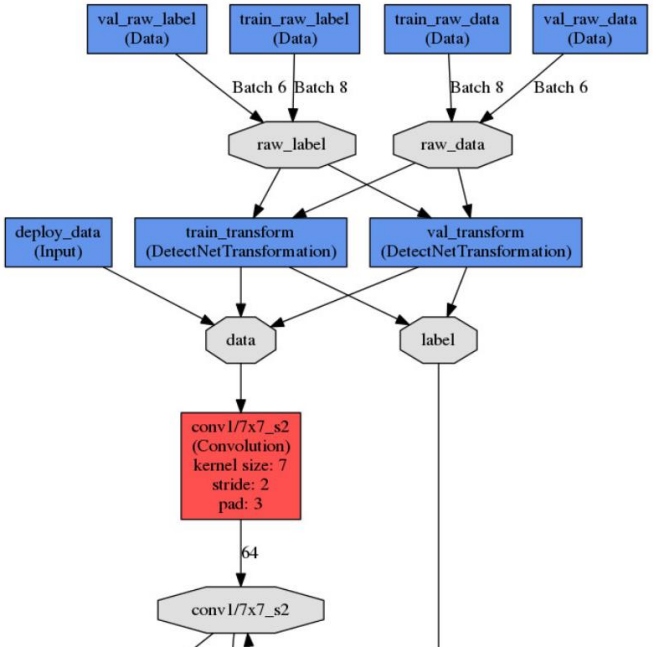
Items per page: 10 - **25** - 50 - 100

« 0 1 2 3 4 5 ... 254 »

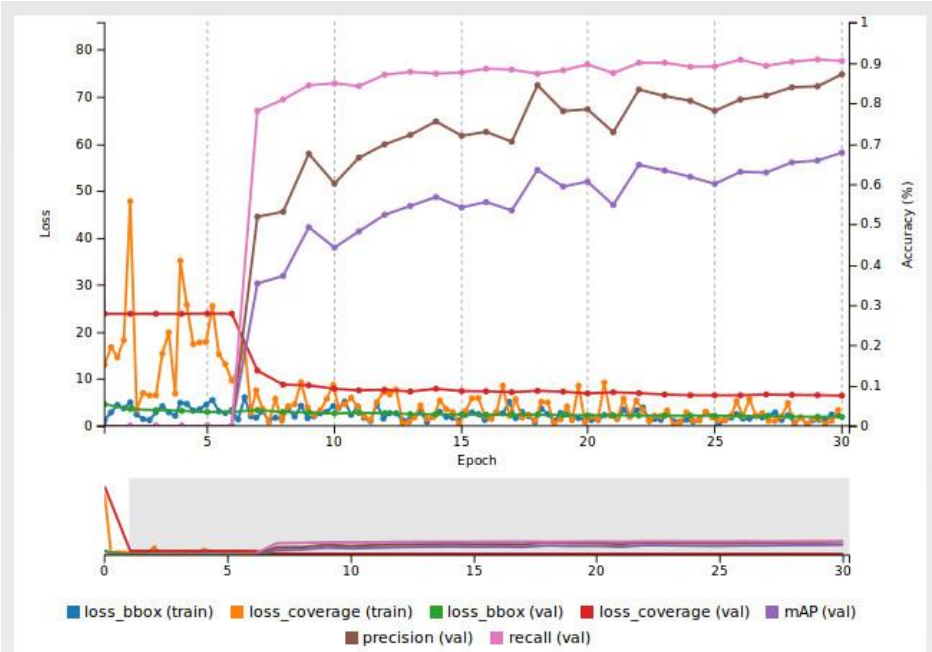
Model Setting DetectNet

$CAFFE_ROOT/examples/kitti/detectnet_network.prototxt




val_raw_label
(Data)
train_raw_label
(Data)
train_raw_data
(Data)
val_raw_data
(Data)
Batch 6
Batch 8
Batch 8
Batch 6
raw_label
raw_data
deploy_data
(Input)
train_transform
(DetectNetTransformation)
val_transform
(DetectNetTransformation)
data
label
conv1/7x7_s2
(Convolution)
kernel size: 7
stride: 2
pad: 3
64
conv1/7x7_s2

Training DetectNet



Inference

Source image

Inference visualization