Bootstrapping object detection from motion

DAVID GILLSJÖ, CENTRE FOR MATHEMATICAL SCIENCES, LUND UNIVERSITY
Collaborators

PhD. Saurabh Gupta, FAIR

Prof. Jitendra Malik, UC Berkeley

Prof. Kalle Åström, Lund University
Motivation

Chair!
Motivation
Overview

Policy → Simulation

Simulation → Detector

Detector → SfM

SfM → Retraining

Retraining → Weights

Weights → B.boxes

B.boxes → Pose

Pose → SfM

Policy → Loss

Loss → Policy
Simulation - SUNCG

- 45k 3D scenes
- OpenGL rendering
SUNCG examples
Object detection

Ross Girshick, 2015, Fast R-CNN slides
Faster R-CNN

- NIPS 2015 - Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun
- PyTorch Detectron
  [https://github.com/roytseng-tw/Detectron.pytorch]
- ResNet50, Feature Pyramid Network

Simple Neural Net

\[ y_i = \sigma \left( \sum_{j=1}^{\infty} w_{ij} x_j + b_i \right) \]

\[ \sigma(x) = \max(0, x) \]
Convolutional layer

CS231n Stanford Course
Conv. operation

\[
\begin{array}{ccccc}
\times(i,j,1) & & \times(i,j,2) & & \\
1 & -2 & 0 & 5 & \\
7 & 0 & 1 & -3 & \\
2 & -2 & 2 & -2 & \\
-5 & -4 & 3 & 1 & \\
\end{array}
\]

\[
\begin{array}{ccccc}
w(i,j,1) & & w(i,j,2) & & \\
1 & -2 & & & \\
3 & 1 & & & \\
\end{array}
\]

\[
\begin{array}{ccc}
y(:, :, 1) & & \\
25 & & \\
\end{array}
\]
Downsampling - max pooling

\[
\begin{array}{cccc}
0 & 2 & -1 & -2 \\
4 & 8 & 9 & 1 \\
1 & -3 & 5 & 3 \\
5 & 2 & 6 & 7 \\
\end{array}
\]

\[
\begin{array}{cccc}
8 & 9 \\
5 & 7 \\
\end{array}
\]

LTH, Mathematics, FMAN45
Conv. operation
ConvNet - Backbone

- Trained for classification
- ImageNet:
  - 1000 classes
  - 1,431,167 images
- VGG-16,
  ResNet-50/101/152
ResNet

Deep Residual Learning for Image Recognition, Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun
ResNet

Deep Residual Learning for Image Recognition, Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun
ResNet

$F(x)$

$H(x) = F(x) + x$

Deep Residual Learning for Image Recognition, Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun
ResNet

\[ F(x) \]

\[ H(x) = F(x) + x \]

Deep Residual Learning for Image Recognition, Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun
RPN - Anchors

RPN - Anchors

RPN - Objectiveness and bbox regression

Faster R-CNN, Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun
RPN - Post-processing

- Non-Max Supression - IoU 0.7
- Keep the 2000 with best Objectiveness
Region Proposal Pooling

Region Proposal Pooling

Hi-res input image: 3 x 800 x 600 with region proposal

Hi-res conv features: C x H x W with region proposal

Rol conv features: C x h x w for region proposal

Max-pool within each grid cell

Convolution and Pooling

Fei-Fei Li, Andrej Karpathy, & Justin Johnson
R-CNN


bicycle
p=0.96
R-CNN post-processing

- Remove background objects
- NMS per class
- Score threshold
Feature Pyramid Network

https://medium.com/@jonathan.hui/understanding-feature-pyramid-networks-for-object-detection-fpn-45b227b9106c
Mask R-CNN

https://github.com/facebookresearch/Detectron
Overview
Experiment #1 setup

• Given bounding box correspondence, can we improve?
• 500/200 houses – 150k/60k images
Experiment #1 setup

- Given bounding box correspondence, can we improve?
- 500/200 houses – 150k/60k images
Result

AR [browkin_results/oracle_gen2]

Dataset

iou0.5_s0_a0_d0_dr0 - 0.30
iou0.25_s0_a0_d0_dr0 - 0.28
 coco - 0.23
 oracle - 0.53

mAR
Problems along the way

- Label mapping COCO to SUNCG
- Ambiguity in SUNCG models
- Forgetting/overfitting during re-training
- SfM with OpenGL rendered images
- Detectron loss on validation set
Problems along the way

- Label mapping COCO to SUNCG
- Ambiguity in SUNCG models
- Forgetting/overfitting during re-training
- SfM with OpenGL rendered images
- Detectron loss on validation set
Problems along the way

- Label mapping COCO to SUNCG
- Ambiguity in SUNCG models
- Forgetting/overfitting during re-training
- SfM with OpenGL rendered images
- Detectron loss on validation set
Problems along the way

- Label mapping COCO to SUNCG
- Ambiguity in SUNCG models
- Forgetting/overfitting during re-training
- SfM with OpenGL rendered images
- Detectron loss on validation set
Problems along the way

- Label mapping COCO to SUNCG
- Ambiguity in SUNCG models
- Forgetting/overfitting during re-training
- SfM with OpenGL rendered images
- Detectron loss on validation set
Development environment
Development environment

FROM nvidia/cuda:8.0-cudnn7-devel

# Arguments
ARG user=pytorch
ARG uid=1000
ARG gid=1000
ARG gpu_arch=sm_01

# Install some dependencies
RUN apt-get update && apt-get install -y \\
    wget \\
    git \\
    python3.5-tk \\
    python3-pip \\
    sudo \\
    vim \\
    unzip \\
    libqt5-qtbase5 \\
    && \\
    apt-get clean && \\
    apt-get autoremove && \\
    rm -rf /var/lib/apt/lists/*

# gdrive dependency
RUN pip3 install requests

# Install repo dependencies
RUN pip3 install \\
    opencv-python-headless \\
    cffi