CSC2548: Machine Learning in Computer Vision
Introduction

Sanja Fidler

January 10, 2018

UNIVERSITY OF TORONTO

Instructor Info

- **Instructor**:

Sanja Fidler (`fidler@cs.toronto.edu`)

- **Office**: 386 in Pratt
- **Office hours**: Send email for appointment

This course has no TAs, so please bare with me!

Course Information

- **Class time**: Wed at 12-2pm
- **Location**: SS 1070
- **Class Website**:

  http://www.cs.toronto.edu/~fidler/teaching/2018/CSC2548.html
- The class will use Piazza for **announcements** and **discussions**:

  piazza.com/utoronto.ca/winter2018/csc2548/home
- Your grade will **not depend on your participation on Piazza**

Course Prerequisites

Good to know:

- Basics of Machine Learning, Neural Networks

Otherwise you'll need some reading

Requirements and Grading

- This course is a seminar course. We'll be reading papers on computer vision, covering various ML techniques. Thus, how much you learn greatly depends on how prepared everyone comes to class.
- Each student expected to write short reviews of two papers per week, present a paper/topic, and do a project
- **Grading**
  - Participation (attendance, participation in discussions, reviews): 15%
  - Presentation (presentation of papers in class): 25%
  - Project (proposal, final report): 60%

Project

Logistics:

- Need to hand in a **report** and do a **presentation**
- Can work **individually** or in **pairs**

Types of projects:

- Great project (A+): nice new research. Does not need to be fully tested by time of presentation
- Good result on a popular benchmark
- Can also implement an existing paper (max grade A, depending how challenging the method is)
- Simply running existing code is not sufficient

Term Work Dates

Term Work	**Due Date**
Reviews	one day before class (Tue)
Project Proposal	Feb 20
Project Report	end of April
Project Presentation	end of April

- All dates are for 2018

Lateness

Deadline Reviews / project should be submitted **by 11.59pm on the date they are due**. Anything from 1 minute late to 24 hours will count as **one late day**.

Lateness Each student will be given a total of **3 free late days**. **After you have used the 3 day budget, each late day will have a 10% penalty.**

Discount You have a budget of 1 missing review without penalty. You do not need to do reviews for the week you present.

Machine Learning

Focus on Deep Learning

- Convolutional Neural Networks
- Recurrent Neural Networks
- Graph Neural Networks
- Reinforcement Learning
- Variational autoencoders, GANs
- Graphical models

Computer Vision

Topics:

- Object detection
- Semantic and instance segmentation
- Stereo, flow
- Action recognition
- Tracking
- 3D scene understanding
- Captioning, VQA, retrieval
- Image/video generation, style transfer

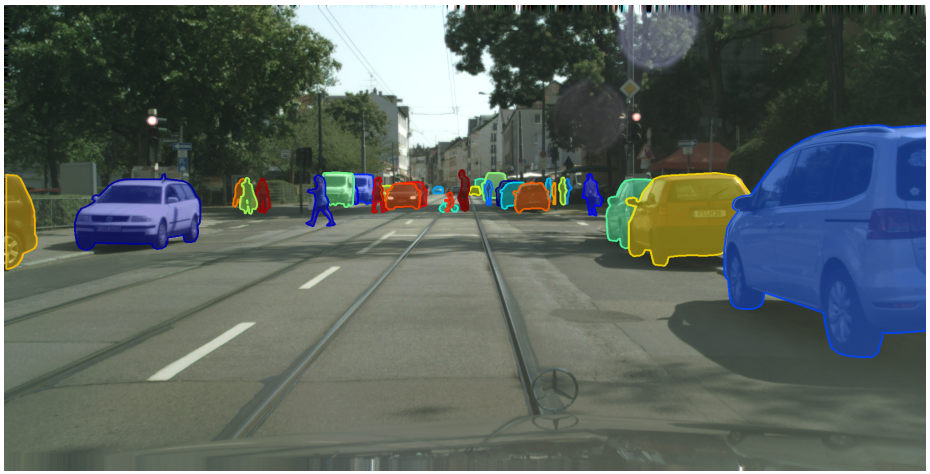
How:

- Overview of topic
- We'll try to cover some old techniques (even if no learning)
- And some of the latest ones

Benchmarks, Resources

Cityscapes: Semantic and instance segmentation

https://www.cityscapes-dataset.com

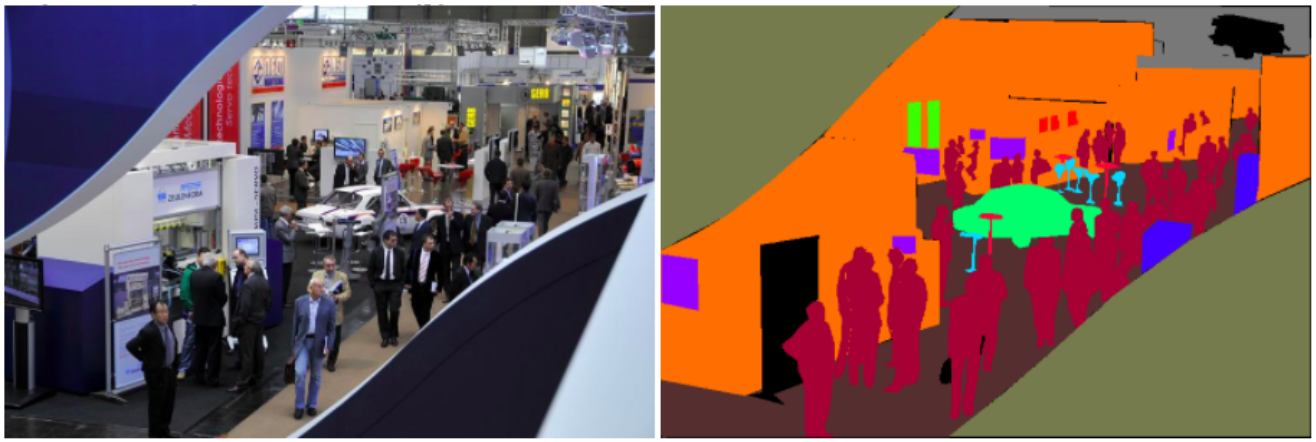


	name	fine	coarse	16-bit	depth	video	sub	AP	AP 50%	AP 100m	AP 50m	Runtime [s]
	Mask R-CNN [COCO]	yes	no	no	no	no	no	32.0	58.1	45.8	49.5	n/a
	SegNet	yes	yes	no	no	no	no	29.5	55.6	43.2	45.8	0.5
	Mask R-CNN [fine-only]	yes	no	no	no	no	no	26.2	49.9	37.6	40.1	n/a
	SGN	yes	yes	no	no	no	no	25.0	44.9	38.9	44.5	n/a
	Pixelwise Instance Segmentation with a Dynamically Instantiated Network	yes	yes	no	no	no	no	20.0	38.8	32.6	37.6	n/a
	Deep Watershed Transformation	yes	no	no	no	no	2	19.4	35.3	31.4	36.8	n/a
	Semantic Instance Segmentation with a Discriminative Loss Function	yes	no	no	no	no	2	17.5	35.9	27.8	31.0	n/a
	Boundary-aware Instance Segmentation	yes	no	no	no	no	2	17.4	36.7	29.3	34.0	n/a
	LCIS	yes	no	no	no	no	no	15.1	30.8	24.2	25.8	n/a

Benchmarks, Resources

PASCAL: Semantic segmentation, detection; 10K images, 20 object classes
http://host.robots.ox.ac.uk/pascal/VOC/voc2012/index.html

ADE20k: Semantic segmentation; 20K images, 150 classes, open voc
http://sceneparsing.csail.mit.edu/



Benchmarks, Resources

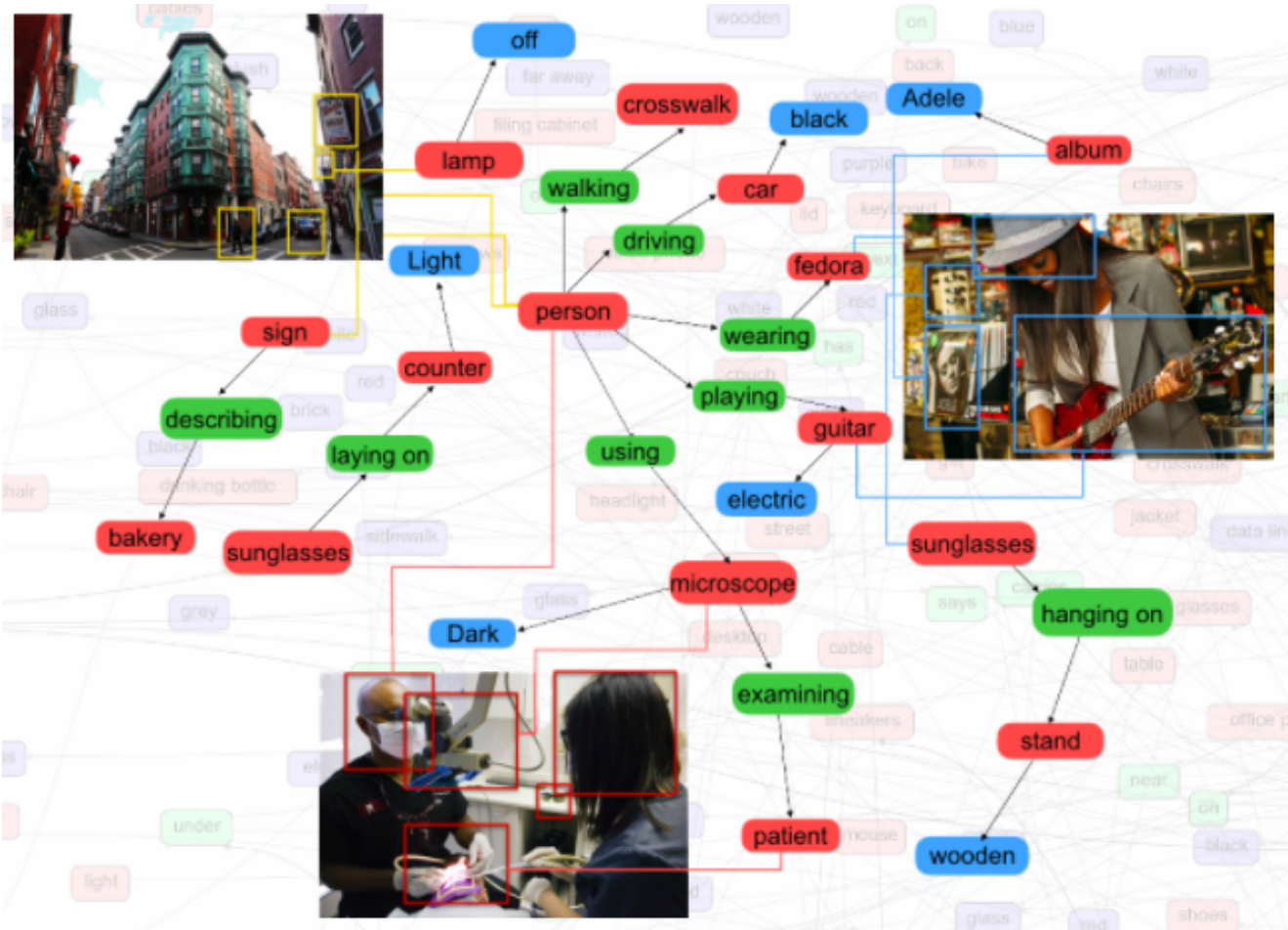
MS-COCO: Detection, segmentation, keypoints, captioning, VQA; 200K images, 80 object classes http://cocodataset.org/

The man at bat readies to swing at the pitch while the umpire looks on.

A large bus sitting next to a very tall building.

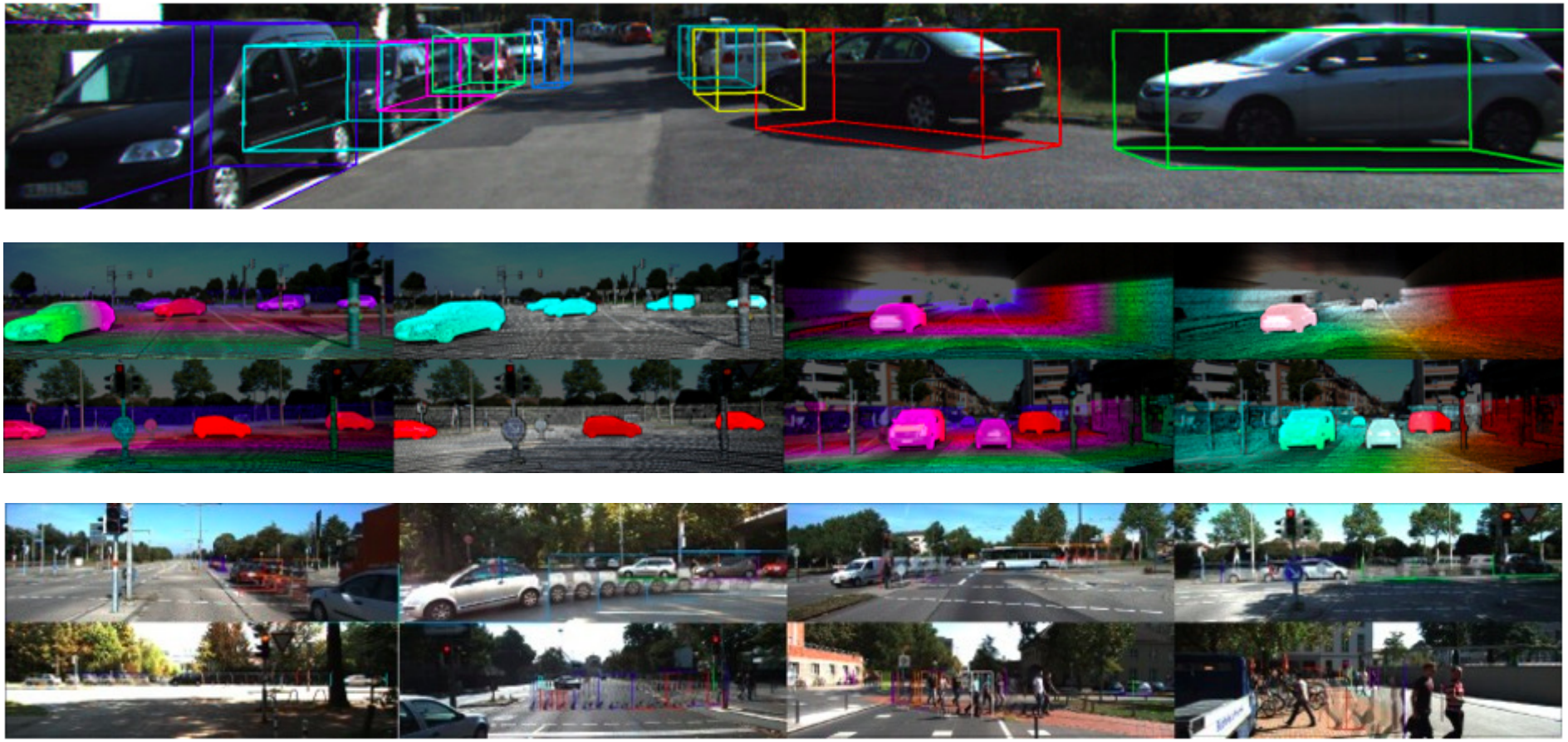
Benchmarks, Resources

Visual Genome: VQA, relationship prediction, attributes, detection...
http://visualgenome.org/



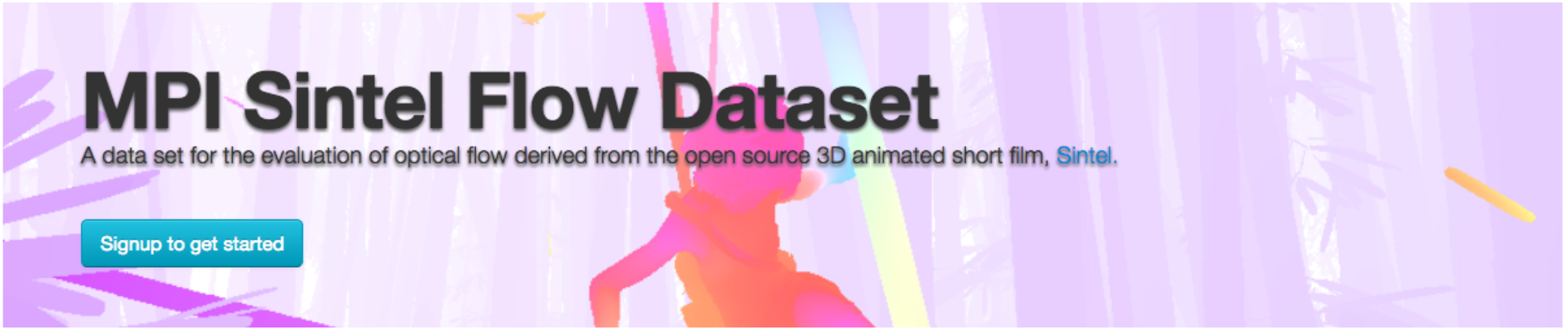
Benchmarks, Resources

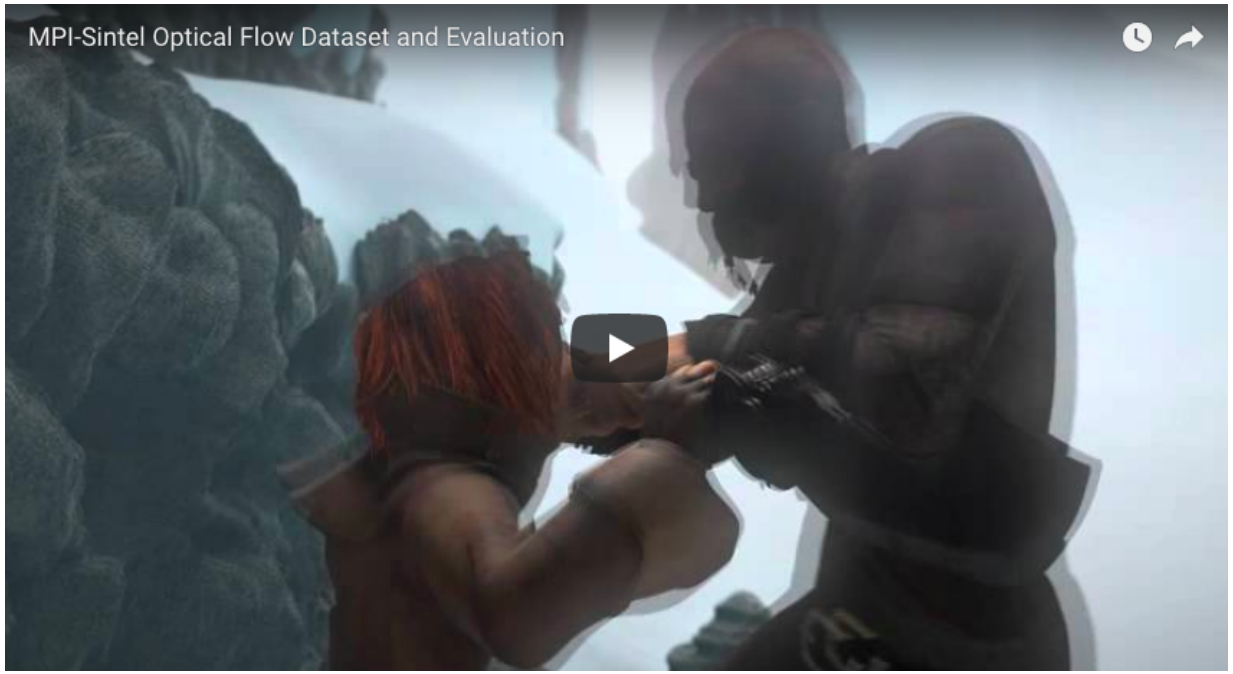
KITTI: Detection (2D, 3D), stereo, flow, tracking, road, odometry
http://www.cvlibs.net/datasets/kitti/index.php



Benchmarks, Resources

Sintel: Flow, http://sintel.is.tue.mpg.de/





Benchmarks, Resources

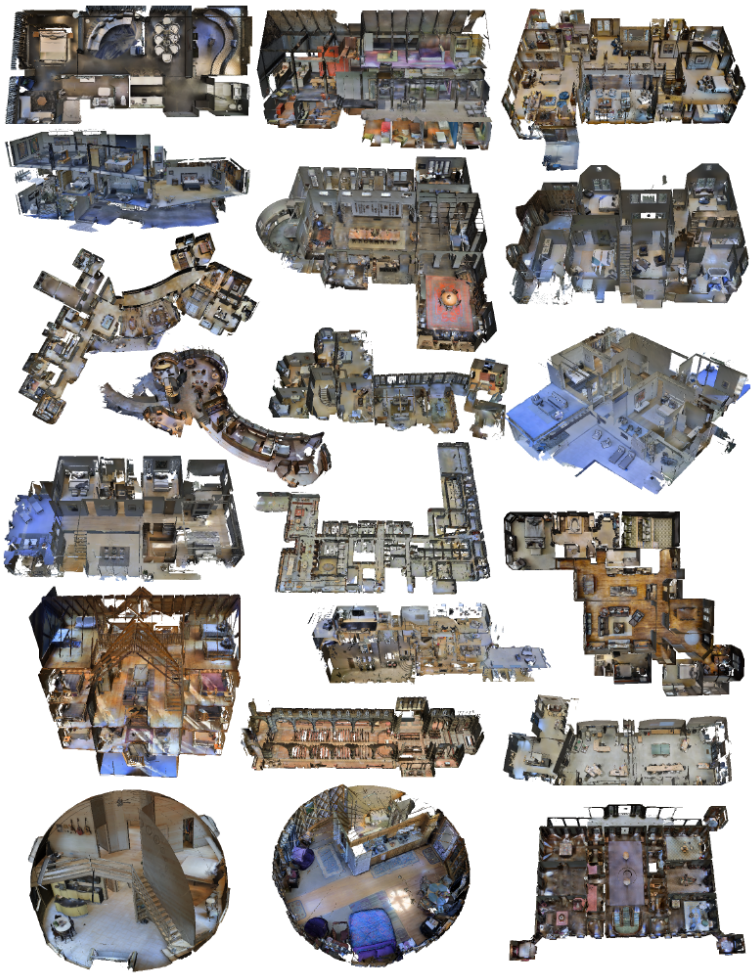
SceneNN: RGB-D segmentation

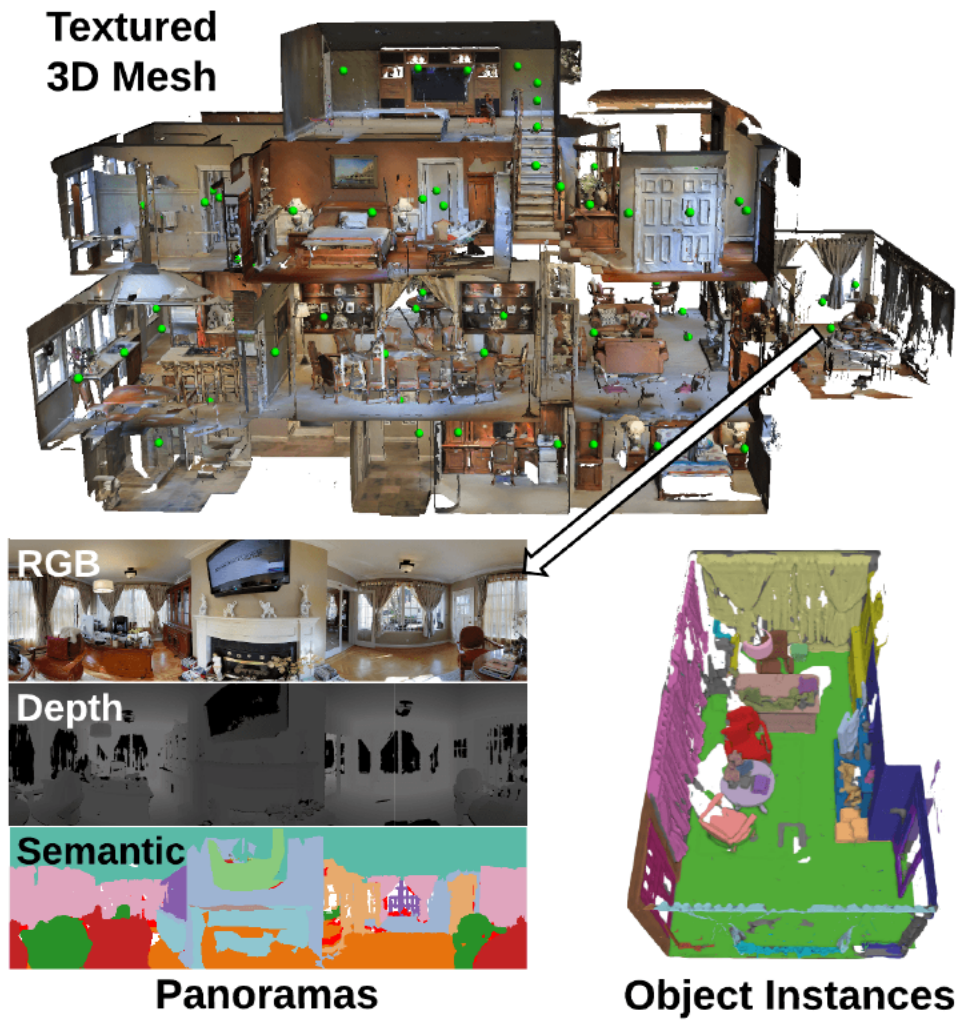
http://people.sutd.edu.sg/~saikit/projects/sceneNN/

Benchmarks, Resources

Matterport3D: RGB-D segmentation, depth estimation

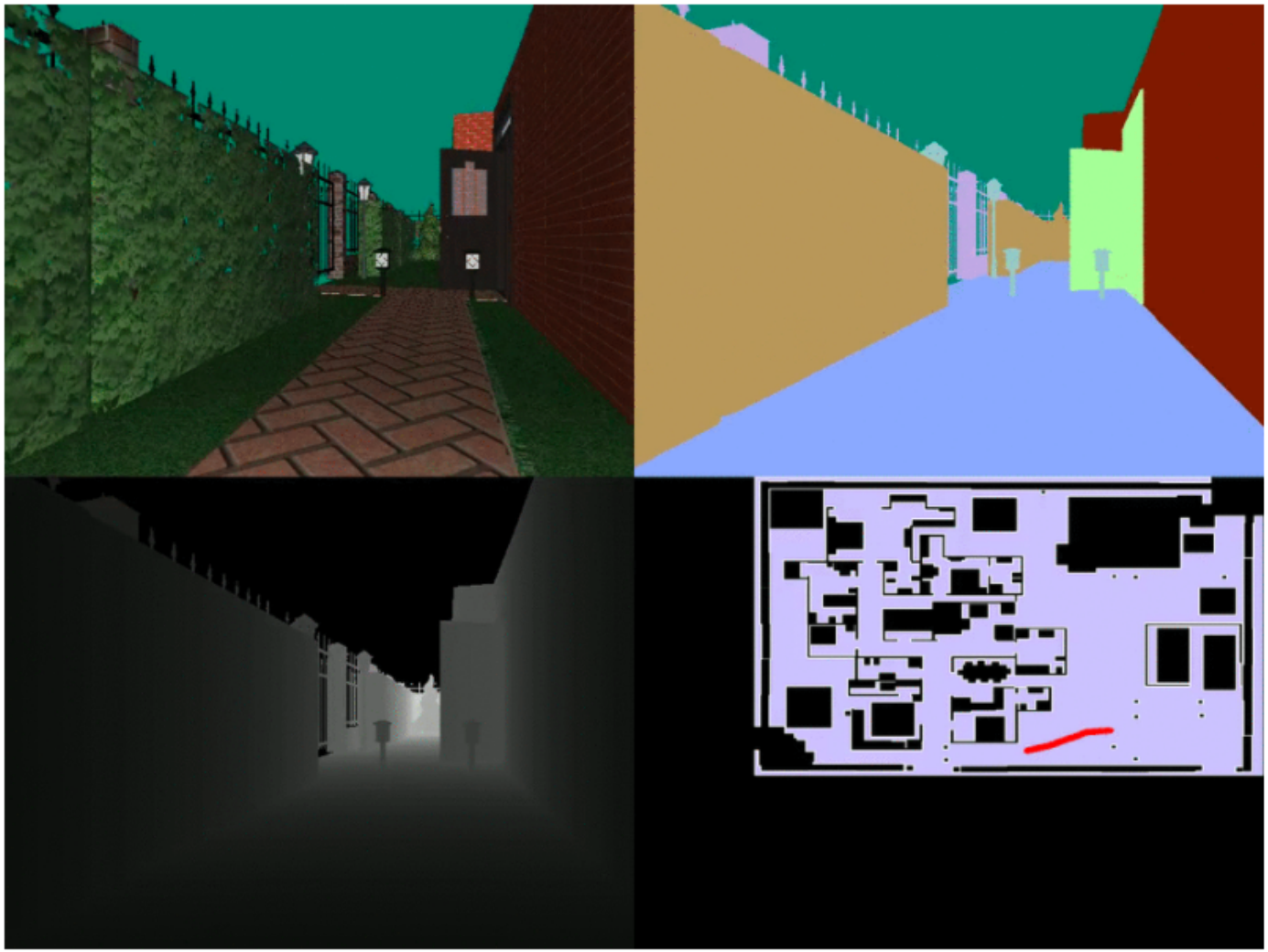
https://matterport.com/blog/2017/09/20/announcing-matterport3d-research-dataset/





Benchmarks, Resources

House3D: Room navigation, grounded VQA
https://github.com/facebookresearch/House3D



Benchmarks, Resources

Something Something: Action classification

https://www.twentybn.com/datasets/something-something

Poking a stack of cans so the stack collapses

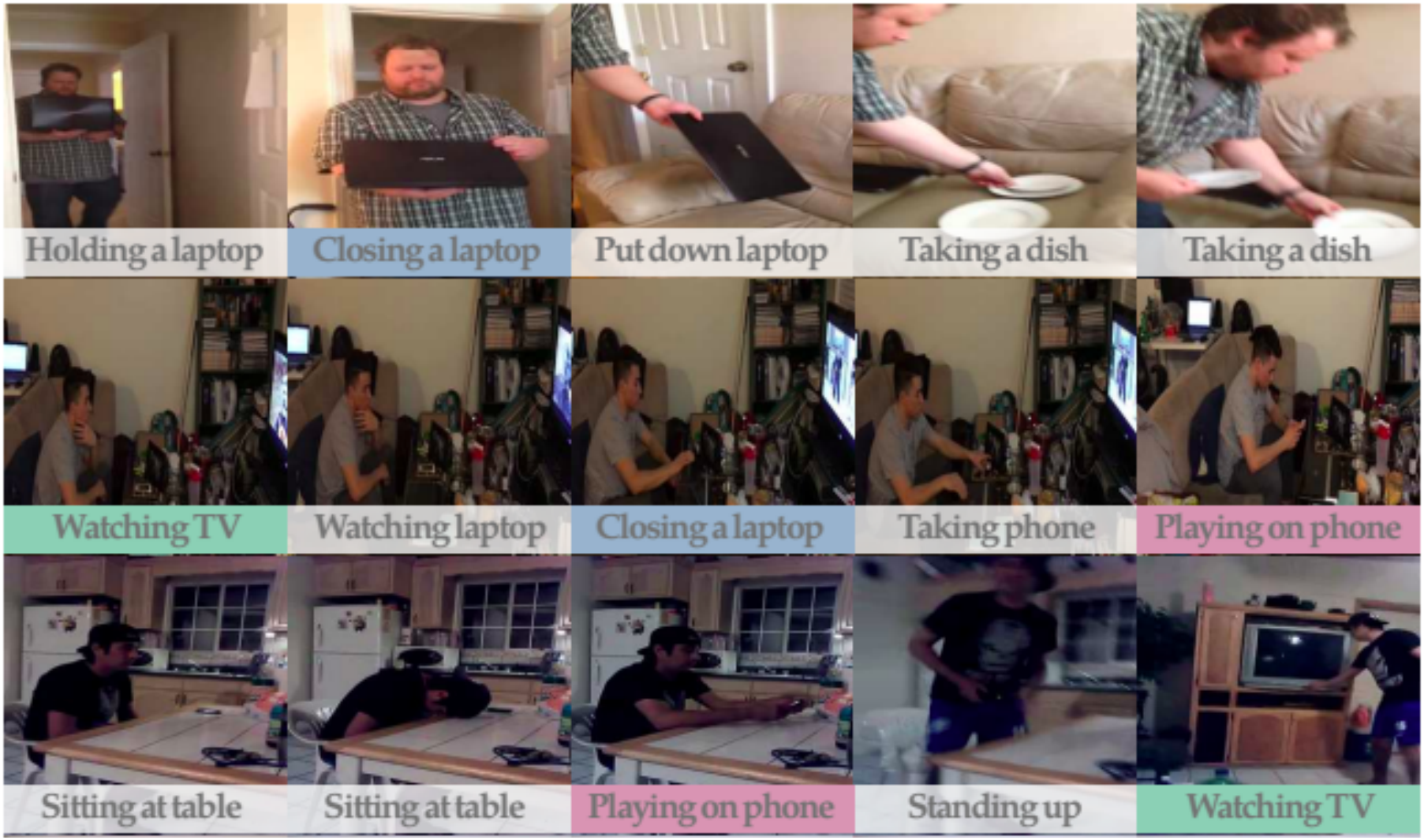
Plugging cable into charger

Closing dishwasher

20BN-SOMETHING-SOMETHING-DATASET	
Total number of videos	108,499
Training Set	86,017
Validation Set	11,522
Test Set (w/o labels)	10,960
Labels	174

Benchmarks, Resources

Charades: Activity parsing; 10k videos
http://allenai.org/plato/charades/



Benchmarks, Resources

MovieQA: Video-based QA

http://movieqa.cs.toronto.edu/

Movie	The Adjustment Bureau	Snatch.	Revolutionary Road
Question	Why does David abandon Elise at the hospital after she sprains her ankle?	Why does a robber tell Franky to buy a gun from Boris?	Why does April die?
Story		- When you get to London... - if you want a gun... - call this number. - Boris?	*April dies in the hospital due to complications following the abortion.*
Correct answer	To protect both Elise and himself from Thompson's threats	Because the robber and Boris want to steal the diamond from Franky	She performs an abortion on her own
Wrong answer 1	Because he wants to be with someone that can walk	He wants to hook him up	Due to injuries from an accident
Wrong answer 2	He wants to run the Bureau and he cannot do it with a limping wife	He plans on robbing and killing him	She kills herself
Wrong answer 3	He does not abandon her, he stays with her	Because otherwise Boris would kill him	Due to a drug overdose
Wrong answer 4	He wants to save her from World War I and the Great Depression	The robber plans to steal a painting from Franky	She is shot