VisDA-17
Visual Domain Adaptation Challenge

Xingchao Peng[1] Ben Usman[1] Neela Kaushik[1] Judy Hoffman[2] Dequan Wang[2] Kate Saenko[1]

[1] Department of Computer Science, Boston University
[2] EECS, University of California Berkeley

10/29/017

Domain Adaptation

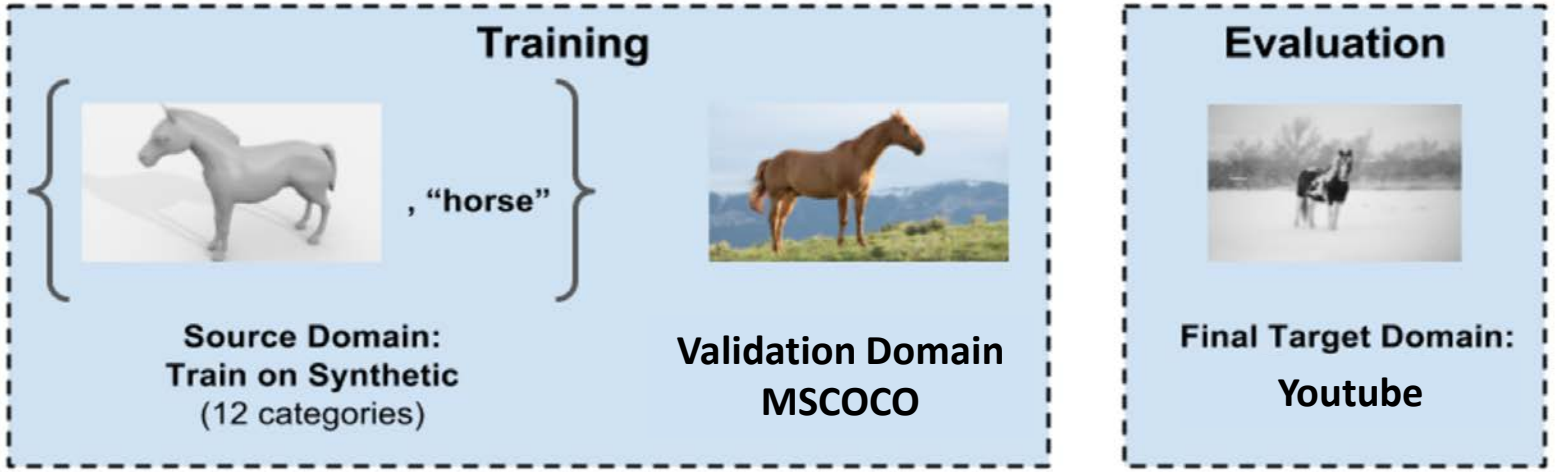
ICCV2017 Workshop Challenge

Evaluation servers go live
June 23rd

Final submission
September 29th

Winners notified
October 13th

Classification Track

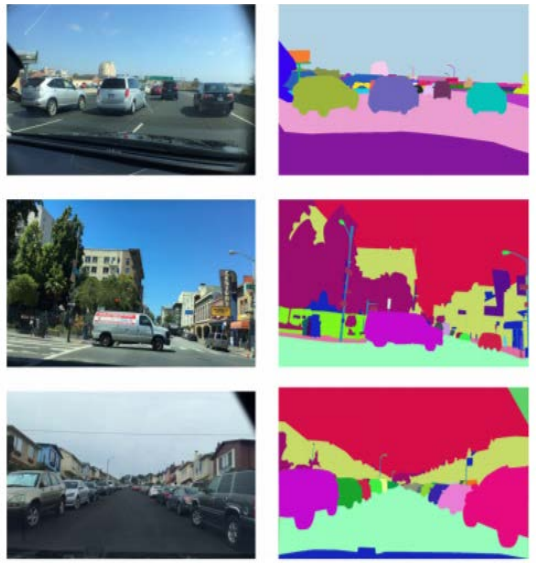


Semantic Segmentation Track

Source: Grand Theft Auto

Validation: CityScapes Real Dashcam

Test: NEXAR real dashcam



Data Generation – Classification Training

2017 VisDA Challenge sponsored by nexar NSF

CAD Model Libraries

ShapeNet

SHREC 3D

SketchUp

NTU 3D

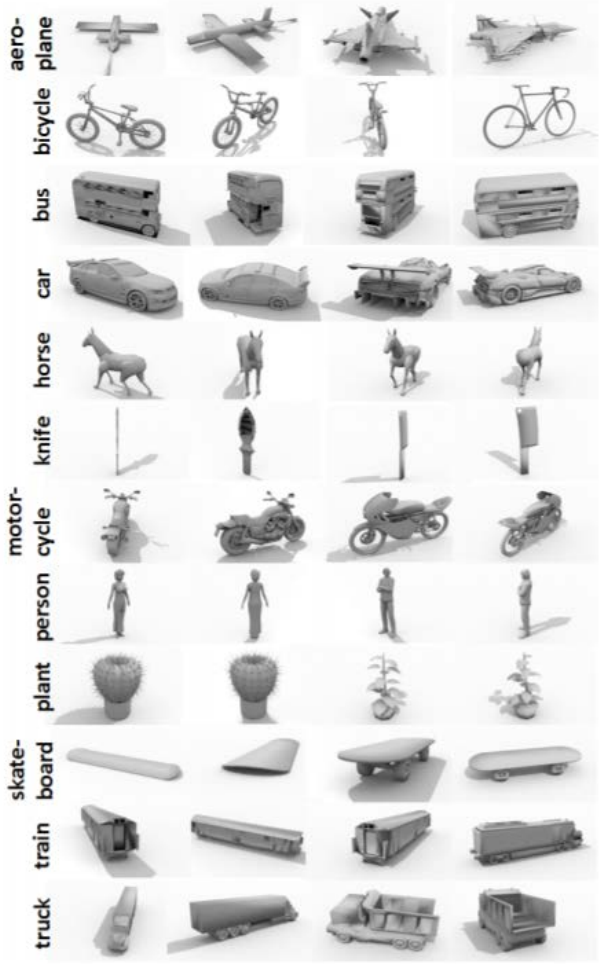
CAD Models

3,466 CAD models
12 Categories

Blender Rendering

Per model, we have
20 different viewpoints
4 lighting conditions

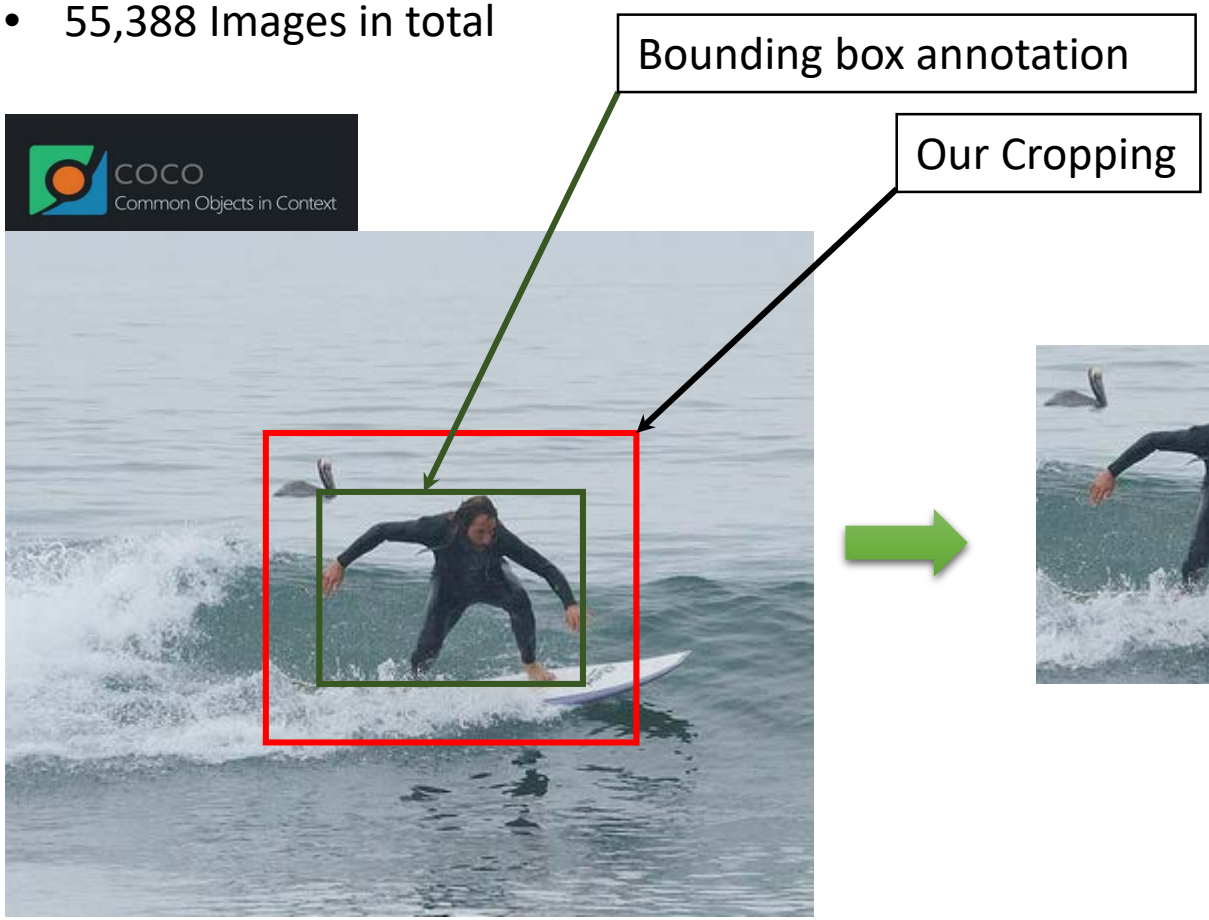
Synthetic Images

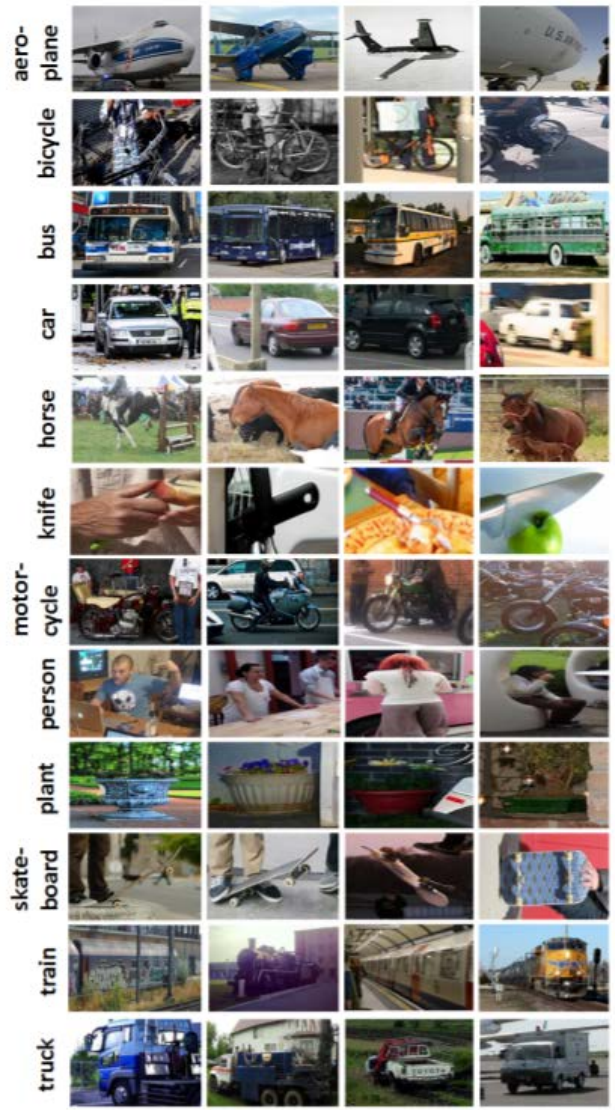


152,397 Images

Data Generation – Classification Validation

- Cropped from COCO dataset
- Padded by retaining an additional 50% of its cropped height and weight
- Padded images under 70x70 pixels were excluded
- 55,388 Images in total





55,388 Images

Data Generation – Classification Testing

2017 VisDA Challenge sponsored by nexar NSF

- Cropped from YouTuBe BBox dataset
- Padded by retaining an additional 50% of its cropped height and weight
- Padded images under 70x70 pixels were excluded
- 72,372 frame Images in total

72,372 Images

Dataset Overview – Classification

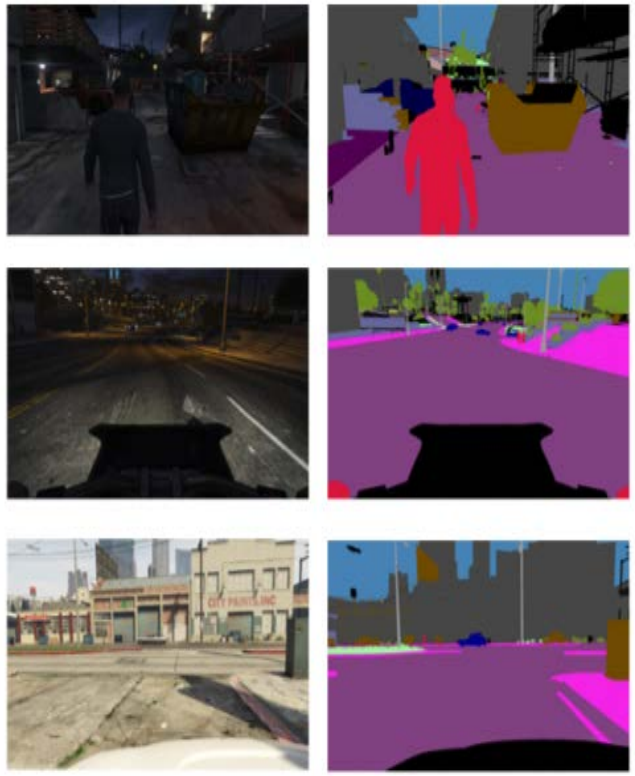
2017 VisDA Challenge
sponsored by nexar NSF

Category	Training		Validation	Testing
	Models	Images	Images	Images
aeroplane	179	14,309	3,646	5,196
bicycle	93	7,365	3,475	4,272
bus	208	16,640	4,690	6,970
car	160	12,800	10,401	7,280
horse	119	9,512	4,691	6,331
knife	178	14,240	2,075	5,491
motorcycle	217	17,360	5,796	8,079
person	152	12,160	4,000	7,673
plant	135	10,371	4,549	4,287
skateboard	146	11,680	2,281	2,762
train	200	16,000	4,236	7,264
truck	120	9,600	5,548	6,767
total	1,907	152,397	55,388	72,372

Data Generation – Segmentation

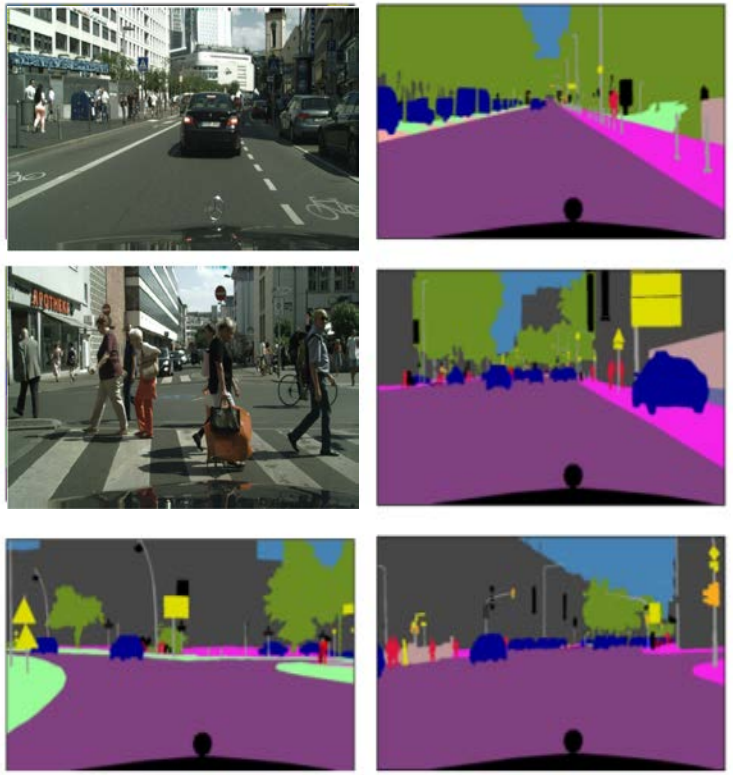
2017 VisDA Challenge sponsored by nexar NSF

Training – Synthetic GTA5



24,966 high quality labeled frames from open-world computer game, GTA 5

Validation - CityScapes



5,000 images

NEXAR DashCam Images

1,500 images

Challenge Statistics

	Classification	Segmentation
#Participated Team	42 teams	45 teams
Submissions	242	167
Train/val data release	June 19, 2017	
Test data release	Sept. 8, 2017	
Challenge ended	September 29th, 11:59pm ET	

Special Thanks

2017 VisDA Challenge sponsored by nexar NSF

Xingchao Peng[1] Ben Usman[1] Neela Kaushik[1] Judy Hoffman[2] Dequan Wang[2] Kate Saenko[1]

BOSTON UNIVERSITY

[1] Department of Computer Science, Boston University
[2] EECS, University of California Berkeley

Competition Website: http://ai.bu.edu/visda-2017/
Evaluation Server will remain open

VisDA17 Tech Report:

```
@article{peng2017visda,
  title={VisDA: The Visual Domain Adaptation Challenge},
  author={Peng, Xingchao and Usman, Ben and Kaushik, Neela and Hoffman, Judy and Wang, Dequan and Saenko, Kate},
  journal={arXiv preprint arXiv:1710.06924},
  year={2017}
}
```

Leaderboard TOP 5 results

2017 VisDA Challenge sponsored by nexar NSF

Leaderboard: ImageNet Pretraining — CAD->Real Classification Track

#	User	Team Name	Per Category Accuracy												MeanAcc		
			plane	bcycl	bus	car	horse	horse	mcycl	person	plant	sktbd	train	truck	s	a	
1	GF_ColourLab_UEA		96.9	92.4	92.0	97.2	95.2	98.8	86.3	75.3	97.7	93.3	94.5	93.3	45.3	92.8	47.5
2	NLE_DA		94.3	86.5	86.9	95.1	91.1	90.0	82.1	77.9	96.4	77.2	86.6	88.0	64.3	87.7	23.4
3	BUPT_OVERFIT	BUPT_OVERFIT	95.7	67.0	93.4	97.2	90.6	86.9	92.0	74.2	96.3	66.9	95.2	69.2	63.2	85.4	22.2
4	MIL	UTokyo_MIL	90.2	84.1	85.4	96.7	89.2	59.5	79.4	75.9	95.8	61.2	91.4	80.3	41.2	82.4	41.2
5	Yifei	ETH_CVL	89.0	65.7	86.9	92.3	78.3	67.1	97.1	64.8	88.4	70.6	91.4	85.4	81.4	81.4	

Leaderboard: ImageNet Pretraining — GTA->Real Segmentation Track

#	User	Team Name	Per Category IoU																		MeanIoU	
			road	sdwlk	bldng	wall	fence	pole	light	sign	vgttn	trrn	sky	person	rider	car	truck	bus	train	mcycl	bcycl	s	a
1	RTZH	MSRA	87.0	38.5	74.7	23.7	30.5	41.1	45.2	36.9	72.1	32.6	90.4	55.9	26.8	80.0	23.4	25.1	28.7	44.6	46.0	35.6	47.5
2	_piotr_		85.3	42.4	53.4	17.3	31.5	39.0	45.6	29.1	77.1	27.3	61.7	57.8	46.1	75.5	27.5	36.4	26.3	51.1	18.1	0.1	44.7
3	whung	VLLAB	87.2	33.3	70.2	13.6	27.8	29.3	32.9	27.9	77.2	28.6	90.3	47.0	35.7	78.0	24.8	18.0	9.1	37.4	38.1	30.3	42.4
4	yuhchen		85.7	36.0	75.4	22.6	25.5	23.9	28.8	18.0	75.6	34.6	90.0	50.8	26.0	77.0	27.1	25.3	6.9	47.4	12.1	34.5	41.5
5	yzou2	CMU_Tartan	87.8	37.4	72.0	24.2	32.4	31.9	32.9	28.7	75.6	33.8	91.8	51.1	6.2	74.1	23.5	1.0	6.1	34.8	0.4	29.8	39.3

Classification Winners

2017 VisDA Challenge sponsored by nexar NSF

Third place

Beijing University of Posts and Telecommunications

*Zhiqun He***

Source 63.2 Adapted 85.4

Second place

Naver Labs Europe

Stéphane Clinchant, Gabriela Csurka, Boris Chidlovskii

Source 64.3 Adapted 87.7

First place

Colour Lab at University of East Anglia

Geoffry French

Source 45.3 Adapted 92.8

Segmentation Winners

2017 VisDA Challenge sponsored by nexar NSF

Third place

Vision and Learning Lab at U.C. Merced

Wei-Chih Hung, Yi-Hsuan Tsai, Ming-Hsuan Yang

Source 30.3 Adapted 42.4

Second place

Active Vision Lab at Oxford University

Piotr Bilinski, Victor Prisacariu

Source – Adapted 44.7

First place

Multimedia Search and Mining group at MSRA

Yiheng Zhang, Zhaofan Qiu, Ting Yao, Tao Mei **

Source 35.6 Adapted 47.5