# **Deep Learning** and Computer Vision

Ryan Fox Milwaukee Machine Learning 12/5/17 Ryan Fox Email: <u>ryan@foxrow.com</u> Web: <u>https://foxrow.com</u> Twitter: @ryan\_fox





**Deep Learning and Computer Vision** 

What is computer vision? How can CNNs help? How can I use it?





**Deep Learning and Computer Vision** 

# What is computer vision? How can CNNs help? How can I use it?



What is computer vision?

- Imagery or video
- Gain information at a high level of abstraction
- Often emulate human vision at computer speed and scale



### Drones - 3D modeling





Image classification
grass, outdoor, people,
large, field, park,
group, sitting, table,
man, standing, grassy,
cake, crowd, display,
ball, riding, horse, air,
umbrella



### Object recognition





### OCR (AKA reading)





**Deep Learning and Computer Vision** 

# What is computer vision? **How can CNNs help?** How can I use it?



### What's a convolution?







https://commons.wikimedia.org/wiki/File:Comparison\_convolution\_correlation.svg

0	2	1	5							B
	0	<b>C</b>	0	-	1	2	1		4	
4	8	6	8	X	2	Л	2	/ 16 =	4	
1	2	2	g		<b>∠</b>	4	2	7 10 -		
-				-	1	2	1			
2	1	1	7		•					
<b>—</b>										

(0\*1 + 2\*2 + 1\*1 + 4\*2 + 8\*4 + 6\*2 + 1\*1 + 2\*2 + 2\*1)/16 = 4

0	2	1	5			•	-					K
Λ	Q	6	o		L	2	L			Λ	5	
4	0	0	0	x	2	Δ	2	/ 16 =	4	4		
1	2	2	9					7 10				
					1	2	1					
2	1	1	7									

(2\*1 + 1\*2 + 5\*1 + 8\*2 + 6\*4 + 8\*2 + 2\*1 + 2\*2 + 9\*1)/16 = 5



http://cs231n.github.io/assets/cnnvis/filt1.jpeg



<u>https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf</u> (17k+ citations!)

### 2011 ImageNet - 26% top-5 error 2012 ImageNet - 15% top-5 error



http://image-net.org/challenges/talks/2016/ILSVRC2016\_10\_09\_clsloc.pdf

### CNNs are good at:

- Single data points
- Recognizing a broad number of classes
- Known knowns

### CNNs are bad at:

- Combining multiple images
- Time series data
- Small training corpora/novel items



Image classification grass, outdoor, people, large, field, park, group, sitting, table, man, standing, grassy, cake, crowd, display, ball, riding, horse, air, umbrella



## Object recognition





### Face/person detection



Finding Tiny Faces, Hu, Ramanan, https://arxiv.org/abs/1612.04402v1

### Segmentation

Classification

#### Classification + Localization

#### **Object Detection**

#### Instance Segmentation



CAT

Single object





CAT, DOG, DUCK

Multiple objects

http://cs231n.stanford.edu/slides/2016/winter1516\_lecture8.pdf

### NSFW classification



OCR

**太阳阳众**院



Artificial neural networks based vehicle license plate recognition, Kocer, Cevik, https://www.sciencedirect.com/science/article/pii/S1877050910005442

### Superresolution



Original / PSNR



SC / 25.58 dB  $\,$ 



Bicubic / 24.04 dB



SRCNN / 27.95 dB

Image Super-Resolution Using Deep Convolutional Networks, Dong et al., https://arxiv.org/pdf/1501.00092.pdf

### **Generative Adversarial Networks**



Intriguing properties of neural networks, Szegedy et al., https://arxiv.org/pdf/1312.6199v4.pdf

### Classification:

- VGG16/VGG19
- ResNet



### Inception



https://research.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html

### Localization:

R-CNN/Fast R-CNN/Faster R-CNNYOLO/YOLOv2/YOLO9000



- Xception
- MobileNet/SqueezeNet

Model	Size	Top-1 Accuracy	Top-5 Accuracy	Parameters	Depth
Xception	88 MB	0.790	0.945	22,910,480	126
VGG16	528 MB	0.715	0.901	138,357,544	23
VGG19	549 MB	0.727	0.910	143,667,240	26
ResNet50	99 MB	0.759	0.929	25,636,712	168
InceptionV3	92 MB	0.788	0.944	23,851,784	159
InceptionResNetV2	215 MB	0.804	0.953	55,873,736	572
MobileNet	17 MB	0.665	0.871	4,253,864	88

**Deep Learning and Computer Vision** 

# What is computer vision? How can CNNs help? **How can I use it?**



### Keras - <u>https://keras.io</u>

model = Sequential() model.add(Conv2D(32, (3, 3), input shape=input shape)) model.add(Activation('relu')) model.add(MaxPooling2D(pool\_size=(2, 2))) 65 model.add(Conv2D(32, (3, 3))) model.add(Activation('relu')) 67 model.add(MaxPooling2D(pool size=(2, 2))) model.add(Conv2D(64, (3, 3))) model.add(Activation('relu')) model.add(MaxPooling2D(pool size=(2, 2))) model.add(Flatten()) 74 model.add(Dense(64)) model.add(Activation('relu')) model.add(Dropout(0.5)) 78 model.add(Dense(1)) model.add(Activation('sigmoid')) 80 model.compile(loss='binary crossentropy', 81 optimizer='rmsprop', 82 metrics=['accuracy']) 83

https://blog.keras.io/building-powerful-image-classification-models-using-very-little-data.html

OpenCV - <u>https://opencv.org</u>



### DIL - <u>https://github.com/foxrow/dil</u>

load DJI\_001.jpg
highlight person car





### AWS, GCP, Azure vision APIs



Resources: <u>https://keras.io/</u> <u>https://opencv.org/</u> <u>http://www.image-net.org/</u> <u>https://github.com/tesseract-ocr/tesseract</u>



**Resources:** 

http://host.robots.ox.ac.uk/pascal/VOC/

https://vision.cornell.edu/se3/projects/microsoft-coco

https://adeshpande3.github.io/

https://foxrow.com/assets/cnns.pdf



Ryan Fox Email: <u>ryan@foxrow.com</u> Web: <u>https://foxrow.com</u> Twitter: @ryan\_fox

> Male Black hair Sunglasses No facial hair Happy

