Learning to see: How machines learn to understand images?

Muhammad Asad

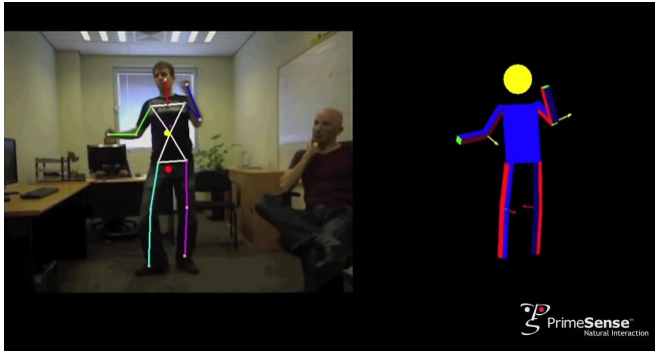
City, University of London, UK

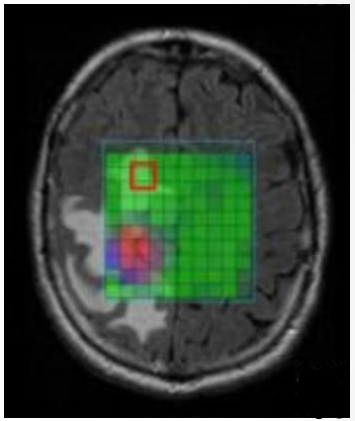
This talk

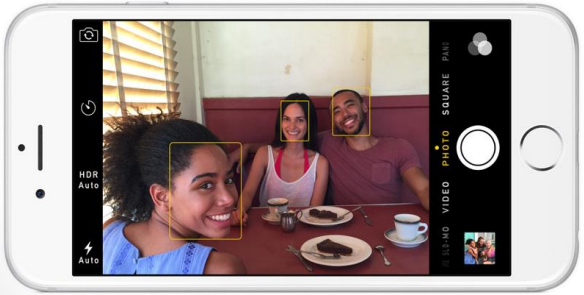
- Introduction
- How does a computer see?
- What is an image?
- Extracting useful information
- How does a computer learn?
- Applications
- Demos (if time allows)


COMPUTER VISION
IS EVERYWHERE
memegenerator.net

Some applications…



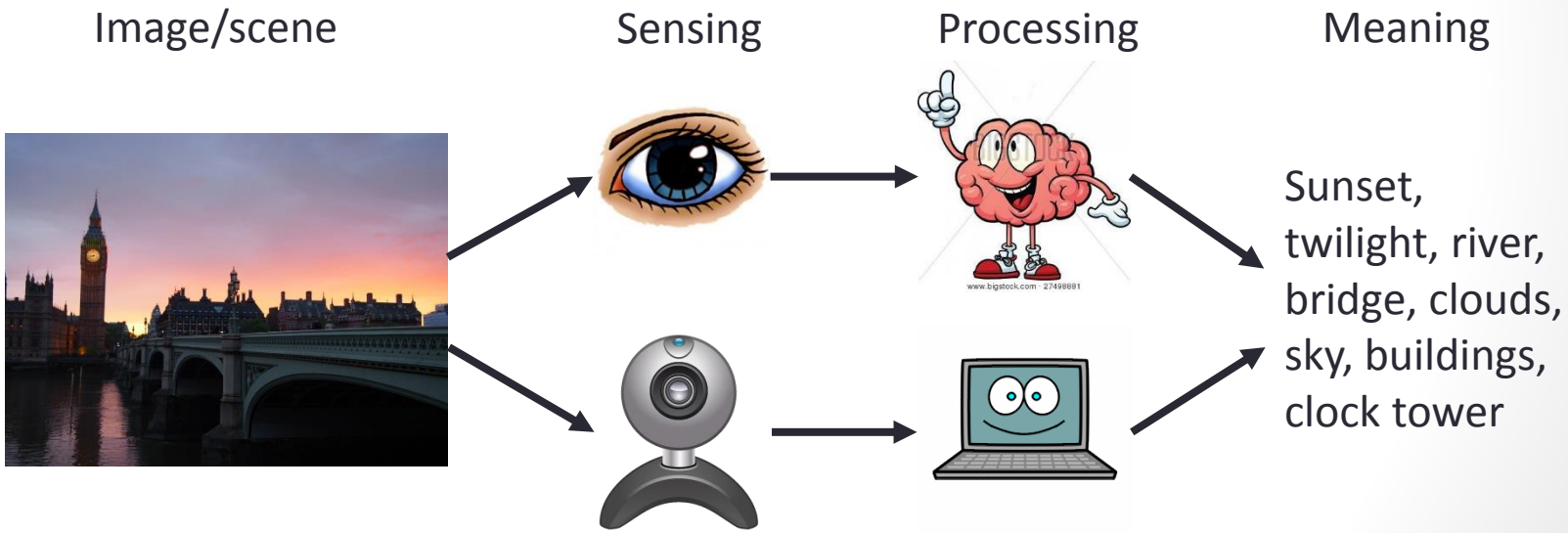




A person riding a motorcycle on a dirt road.

How does a computer see?

- How do we “see” something? What processes does it involve?
- How does a computer achieve the same task?

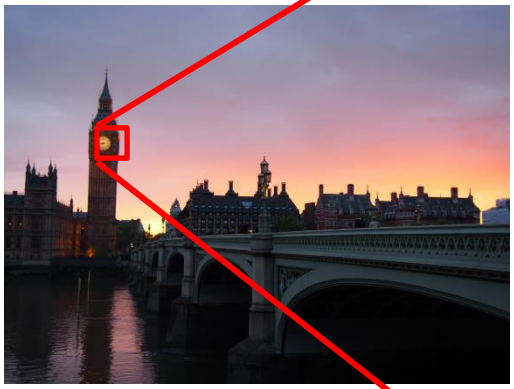


https://www.cs.princeton.edu/courses/archive/spring07/cos424/lectures/li-guest-lecture.pdf

What is an image?

What is an image?

What is an image?



R: 49 G: 62 B: 21	R: 44 G: 52 B: 23	R: 40 G: 44 B: 28	R: 40 G: 44 B: 31	R: 47 G: 46 B: 30	R: 70 G: 60 B: 39	R: 62 G: 51 B: 49	R:145 G:109 B:127	R:197 G:172 B:189	R:197 G:176 B:192	R:198 G:176 B:191	R:197 G:175 B:190	R:199 G:172 B:189	R:199 G:172 B:190
R: 65 G: 67 B: 26	R: 65 G: 53 B: 30	R: 63 G: 47 B: 36	R: 64 G: 55 B: 37	R: 71 G: 56 B: 42	R: 60 G: 55 B: 37	R: 55 G: 45 B: 46	R:115 G: 83 B: 97	R:192 G:157 B:176	R:202 G:174 B:192	R:200 G:174 B:190	R:200 G:173 B:189	R:200 G:171 B:188	R:200 G:170 B:188
R: 91 G: 44 B: 16	R: 82 G: 41 B: 20	R: 75 G: 42 B: 26	R: 78 G: 45 B: 26	R: 68 G: 42 B: 30	R: 50 G: 38 B: 35	R: 43 G: 35 B: 40	R: 90 G: 39 B: 33	R:205 G:131 B:130	R:202 G:174 B:192	R:202 G:172 B:188	R:202 G:172 B:188	R:203 G:172 B:188	R:204 G:171 B:187
R:163 G: 86 B: 37	R:148 G: 84 B: 37	R: 96 G: 55 B: 27	R: 69 G: 36 B: 24	R: 63 G: 38 B: 30	R: 53 G: 33 B: 33	R: 50 G: 35 B: 39	R:106 G: 40 B: 18	R:228 G:138 B:116	R:204 G:175 B:194	R:206 G:173 B:189	R:206 G:173 B:188	R:206 G:173 B:188	R:206 G:172 B:188
R:202 G:143 B: 78	R:201 G:144 B: 82	R:199 G:143 B: 70	R:119 G: 80 B: 38	R: 52 G: 36 B: 34	R: 63 G: 42 B: 42	R: 67 G: 47 B: 47	R:106 G: 40 B: 23	R:222 G:127 B:112	R:208 G:175 B:194	R:207 G:174 B:188	R:208 G:174 B:189	R:209 G:174 B:189	R:209 G:174 B:188
R:234 G:177 B: 97	R:229 G:173 B: 96	R:224 G:167 B: 85	R:190 G:134 B: 58	R: 73 G: 49 B: 33	R: 52 G: 34 B: 39	R: 64 G: 42 B: 44	R: 79 G: 32 B: 35	R:192 G:123 B:139	R:213 G:178 B:192	R:210 G:175 B:188	R:213 G:174 B:189	R:213 G:174 B:188	R:212 G:174 B:188
R:166 G:118 B: 58	R:184 G:133 B: 79	R:228 G:170 B: 88	R:204 G:145 B: 63	R: 93 G: 62 B: 31	R: 58 G: 35 B: 37	R: 66 G: 43 B: 43	R: 65 G: 30 B: 37	R:184 G:127 B:149	R:217 G:179 B:192	R:215 G:174 B:188	R:216 G:174 B:188	R:216 G:175 B:188	R:216 G:175 B:189
R:215 G:155 B: 87	R:224 G:167 B: 94	R:212 G:155 B: 77	R:196 G:139 B: 61	R: 84 G: 58 B: 30	R: 56 G: 37 B: 40	R: 63 G: 44 B: 45	R: 68 G: 29 B: 37	R:186 G:132 B:153	R:220 G:180 B:192	R:218 G:175 B:188	R:218 G:175 B:187	R:218 G:175 B:188	R:219 G:175 B:189
R:194 G:139 B: 81	R:199 G:143 B: 78	R:196 G:141 B: 69	R:137 G: 92 B: 38	R: 52 G: 35 B: 29	R: 54 G: 35 B: 36	R: 58 G: 40 B: 41	R: 66 G: 35 B: 44	R:188 G:138 B:157	R:223 G:180 B:193	R:221 G:176 B:188	R:220 G:177 B:187	R:221 G:176 B:188	R:222 G:176 B:187
R:138 G: 95 B: 61	R:140 G: 96 B: 62	R:106 G: 73 B: 39	R: 71 G: 43 B: 27	R: 52 G: 34 B: 32	R: 66 G: 42 B: 41	R: 67 G: 45 B: 44	R: 71 G: 38 B: 47	R:199 G:145 B:164	R:225 G:180 B:191	R:223 G:176 B:187	R:222 G:177 B:187	R:224 G:177 B:186	R:222 G:177 B:186

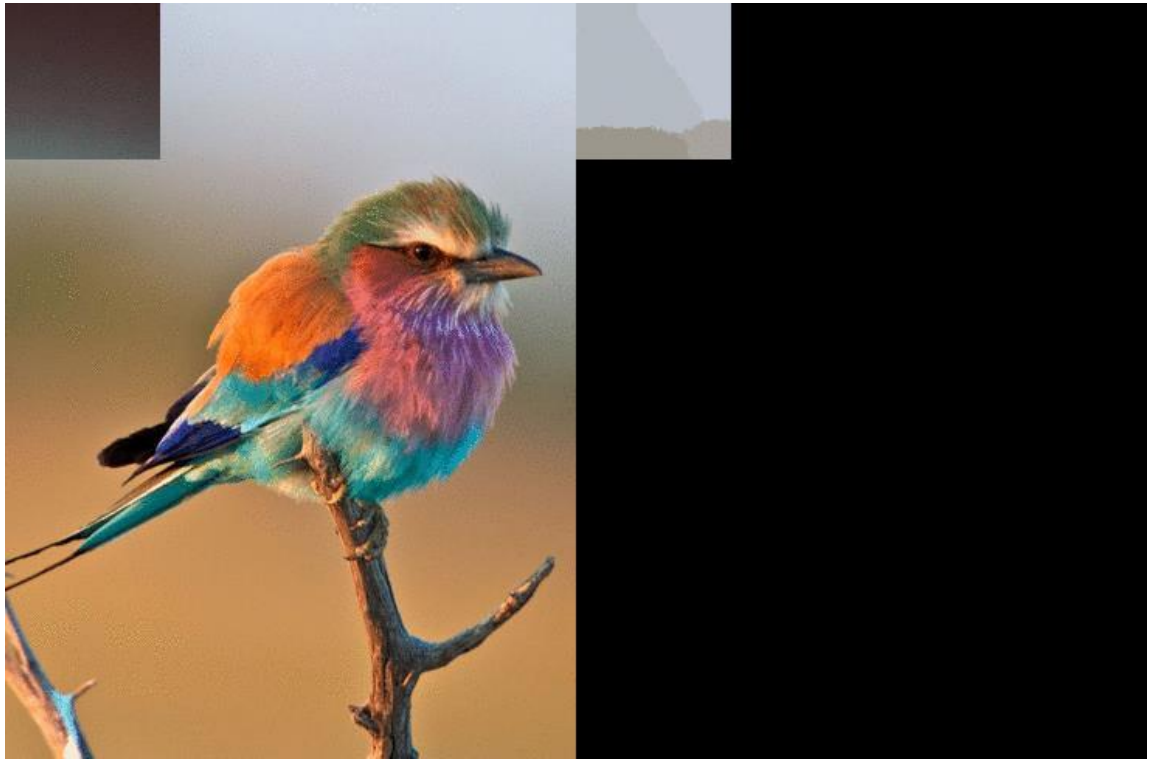
Useful information

- Problems with images
  - Too much information

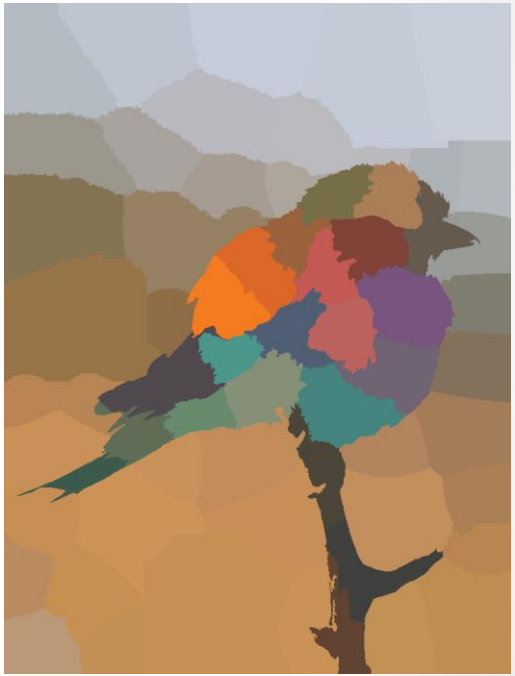
- Features of interest are extracted using a set of mathematical operations that represent:
  - Shape Information
  - Color(s) Information
  - Texture Information

https://uk.pinterest.com/explore/hummingbird-photos/

How are the features extracted?



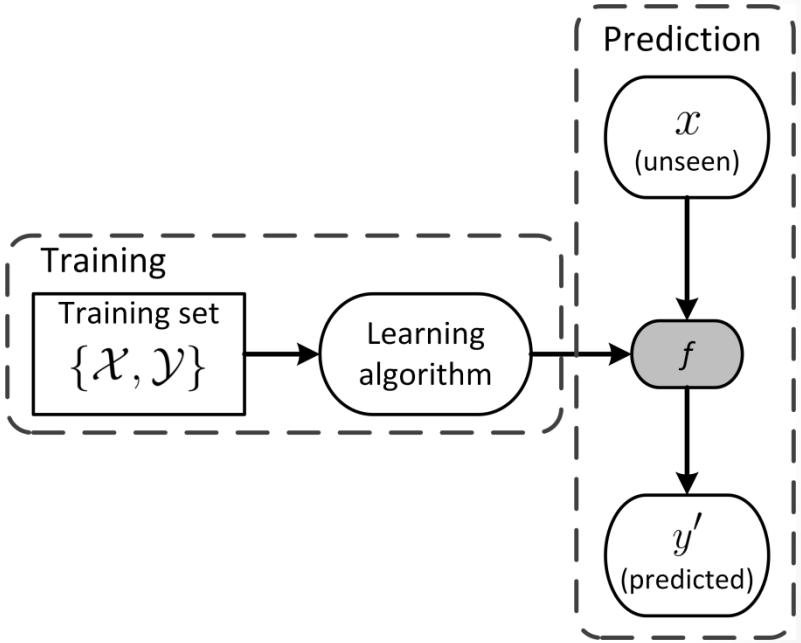
Feature extraction



https://uk.pinterest.com/explore/colourful-birds/

How does a computer learn?

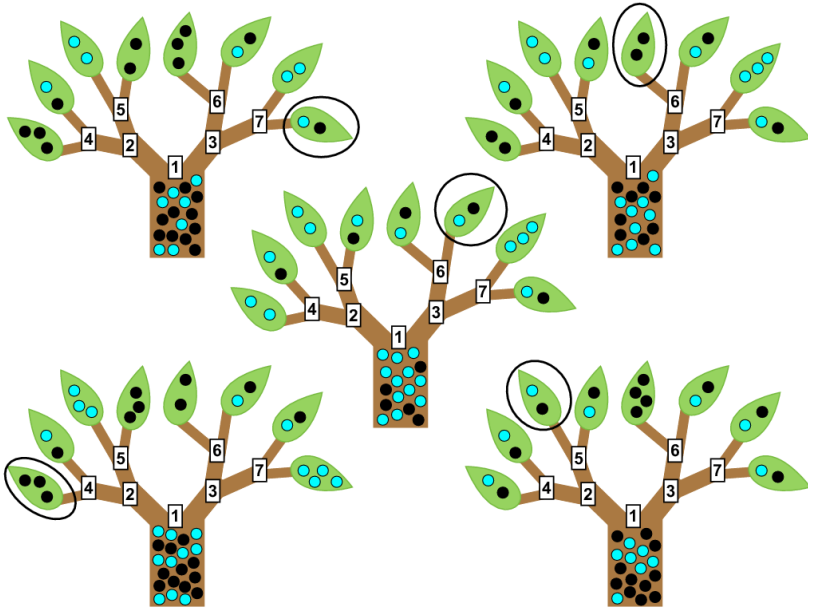
- Given a training set $\{\mathcal{X}, \mathcal{Y}\}$
  - x - Input features
  - y - Target labels
- Learn function $f : \mathcal{X} \mapsto \mathcal{Y}$
  - f could be a regression or classification based on y

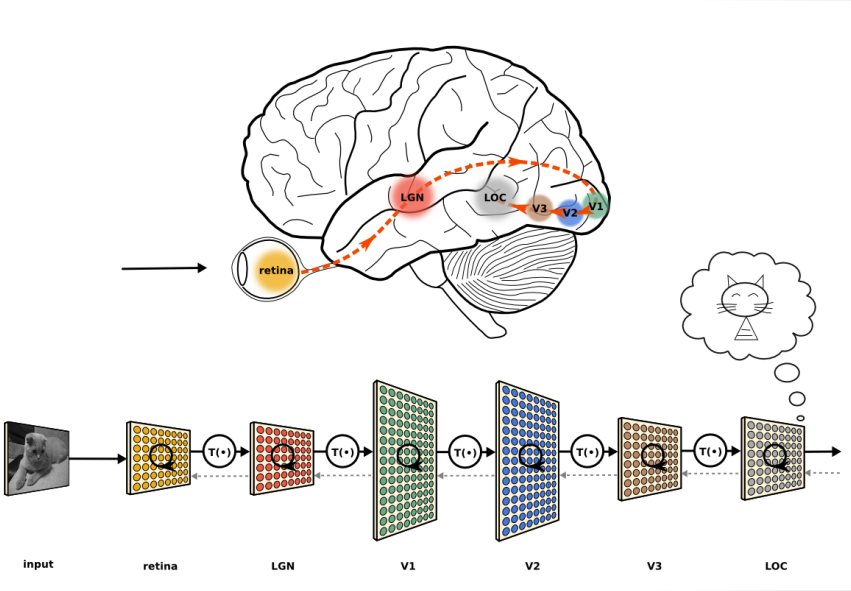


Inspiration from nature…

http://machinelearningmastery.com/tune-machine-learning-algorithms-in-r/
http://scicurious.scientopia.org/2011/05/04/science-101-the-neuron/

Inspiration from nature…

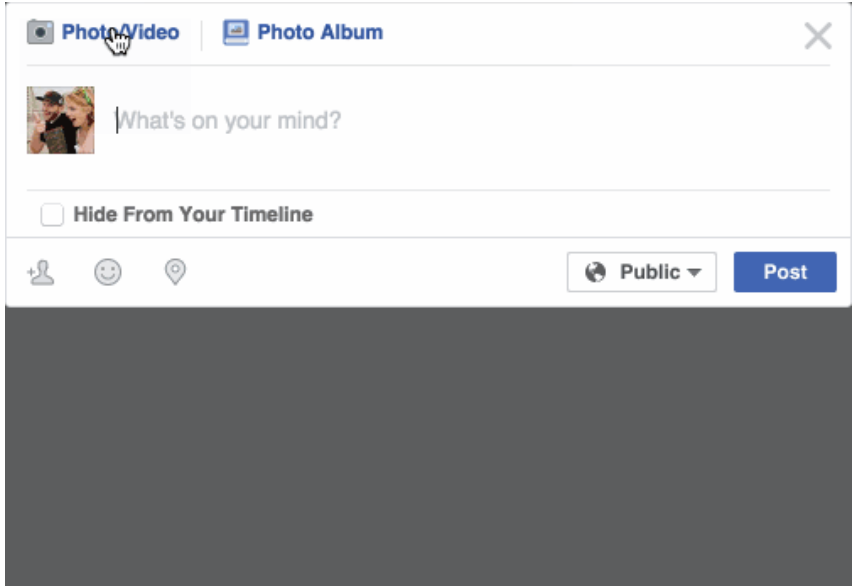


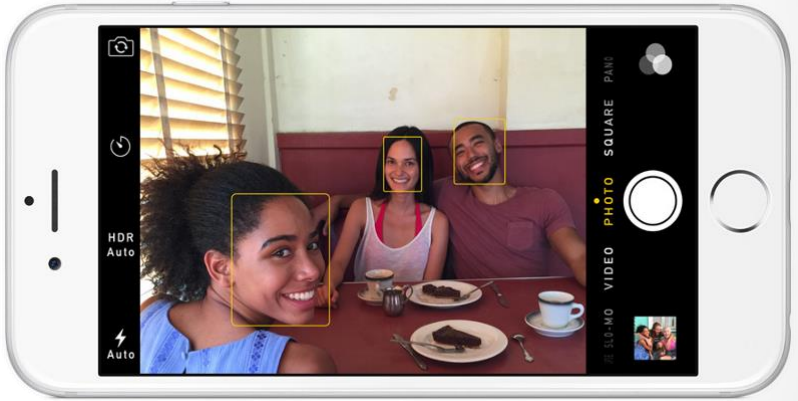


https://inspirehep.net/record/1335130/plots
https://neuwritesd.org/2015/10/22/deep-neural-networks-help-us-read-your-mind/

Applications

Face recognition…





Machine Learning is Fun! Part 4: Modern Face Recognition with Deep Learning
https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78

Recognizing objects…

https://aiexperiments.withgoogle.com/

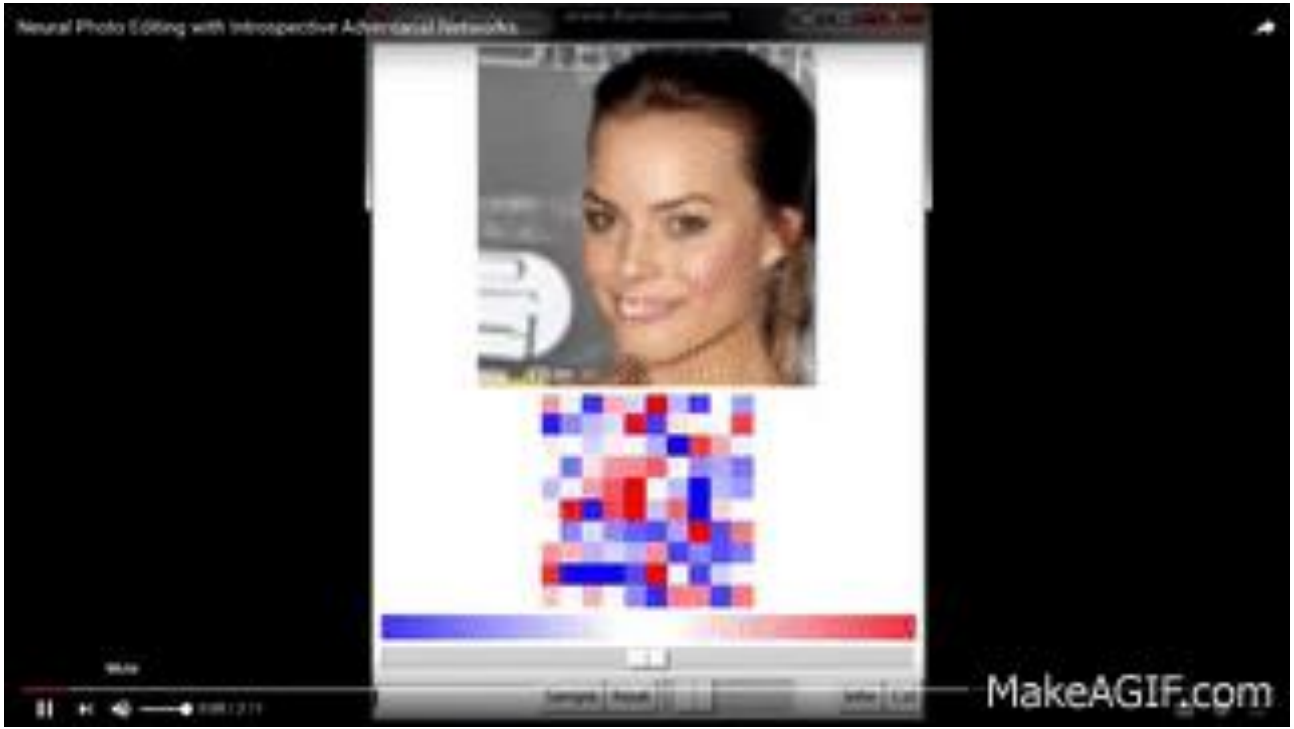
Recognizing drawings…

https://aiexperiments.withgoogle.com/

Photo editing…

NIPS 2016 – Generative Adversarial Networks – Ian Goodfellow
https://www.youtube.com/watch?v=AJVyzd0rqdc

Photo editing…



NIPS 2016 – Generative Adversarial Networks – Ian Goodfellow
https://www.youtube.com/watch?v=AJVyzd0rqdc

Style transfer from art…

+

https://deepart.io/

Style transfer from art…

https://deepart.io/

Image captioning…

A person riding a motorcycle on a dirt road.

A picture is worth a thousand (coherent) words: building a natural description of images
https://research.googleblog.com/2014/11/a-picture-is-worth-thousand-coherent.html

Image captioning…

A herd of elephants walking across a dry grass field.

A picture is worth a thousand (coherent) words: building a natural description of images
https://research.googleblog.com/2014/11/a-picture-is-worth-thousand-coherent.html

Image captioning…

A refrigerator filled with lots of food and drinks.

A picture is worth a thousand (coherent) words: building a natural description of images
https://research.googleblog.com/2014/11/a-picture-is-worth-thousand-coherent.html

Self driving cars…

Nick's Tesla –Self Driving Car Demo Video Analyzed
https://www.youtube.com/watch?v=YntM3bVVago

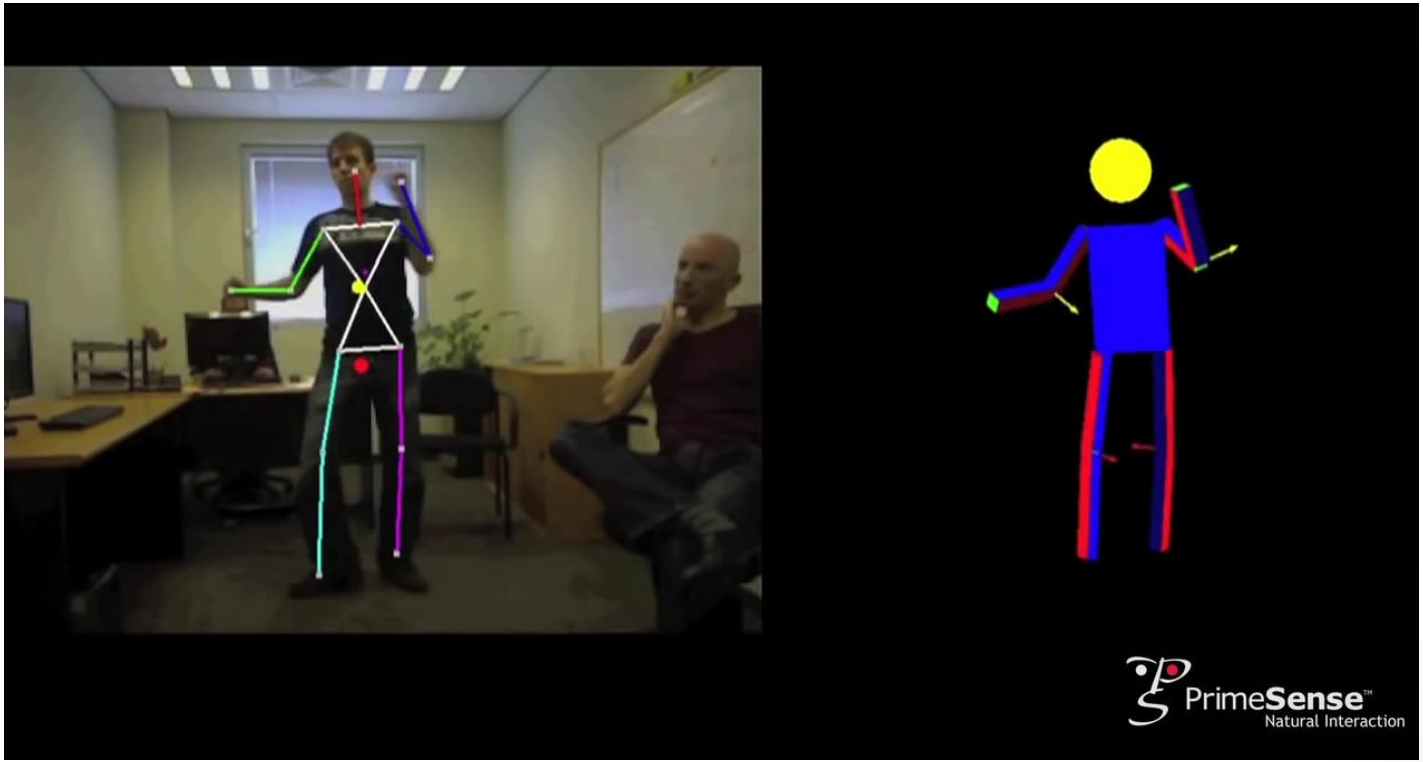

MEDIUM RANGE VEHICLE CAMERA
MakeAGIF.com

Augmented Reality…

Revealing invisible changes…

Revealing invisible changes in the world
https://www.youtube.com/watch?v=e9ASH8IBJ2U

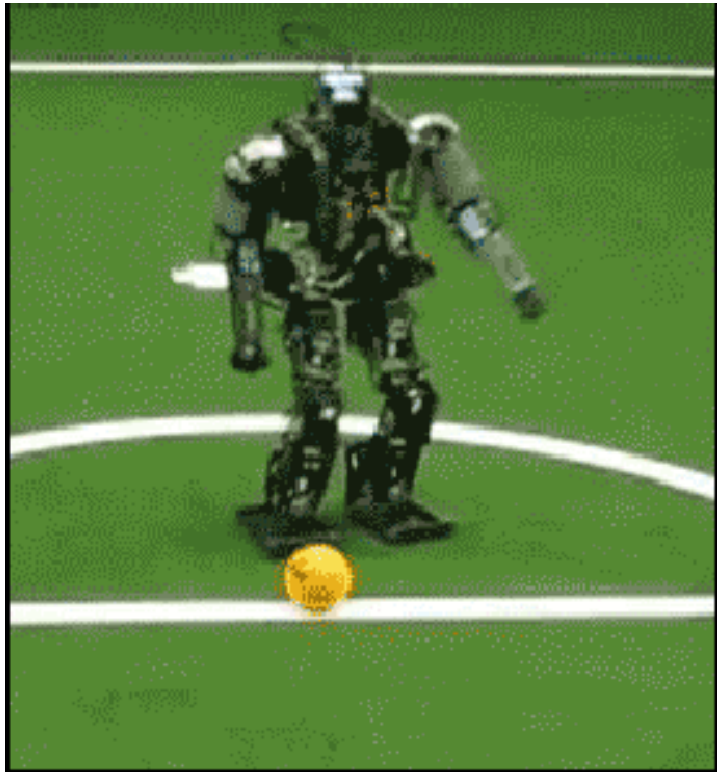
Understanding human motion…



Demos

- https://quickdraw.withgoogle.com/
- http://cs.stanford.edu/people/karpathy/convnetjs/demo/image_regression.html
- https://aiexperiments.withgoogle.com/

Thank you!



Muhammad Asad
City, University of London, UK
Muhammad.Asad.2@city.ac.uk

http://masadcv.github.io/