



Deep Learning for Graphics

Data sets

Niloy Mitra

UCL

Iasonas Kokkinos

UCL/Facebook

Paul Guerrero

UCL

Vladimir Kim

Adobe Research

Kostas Rematas

U Washington

Tobias Ritschel

UCL



Timetable

	Niloy	Iasonas	Paul	Vova	Kostas	Tobias
Introduction	X	X	X			X
Theory	X					
NN Basics		X				X
Supervised Applications						
Data						X
Unsupervised Applications			X			
Beyond 2D	X		X	X		
Outlook	X	X	X	X	X	X

Modalities

- 2D images
- 3D datasets
- Data augmentation
- Synthetic image training data

2D image data sets

Image Datasets

- MNIST
 - Handwritten digits
 - 28x28 images
 - 10 classes
 - 60k train/10k test



Image Datasets

- CIFAR10
 - Object images
 - 10 classes
 - 32 x 32 pixels
 - 50k train/10k test

airplane



automobile



bird



cat



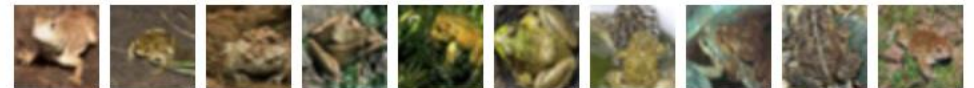
deer



dog



frog



horse



ship



truck



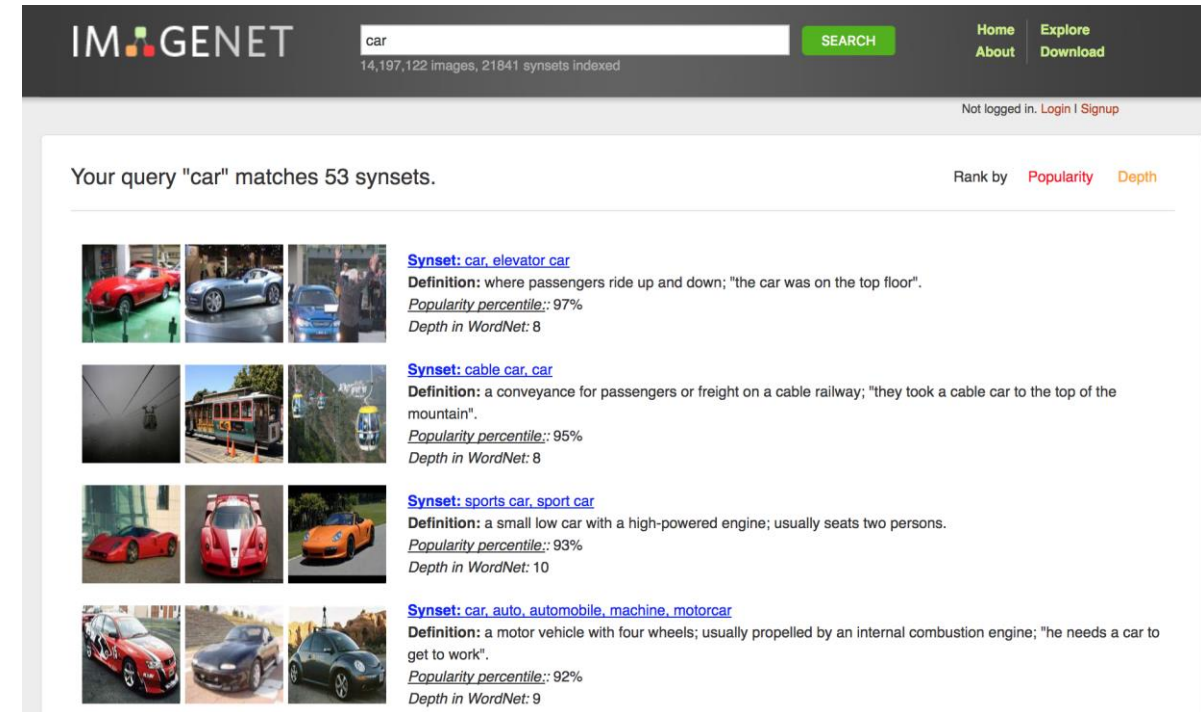
Image Datasets

- PASCAL VOC
 - Multiple objects per image
 - 20 classes
 - Labels for classification, segmentation, detection



Image Datasets

- ImageNet
 - The main "fuel" for deep learning
 - 1000 classes
 - Classification/Detection (200 classes)
 - Structure from WordNet



The screenshot shows the ImageNet website interface. At the top, the logo "IMAGENET" is displayed next to a search bar containing the word "car" and a green "SEARCH" button. Below the search bar, it indicates "14,197,122 images, 21841 synsets indexed". Navigation links for "Home", "About", "Explore", and "Download" are visible. A user status bar shows "Not logged in. Login | Signup".

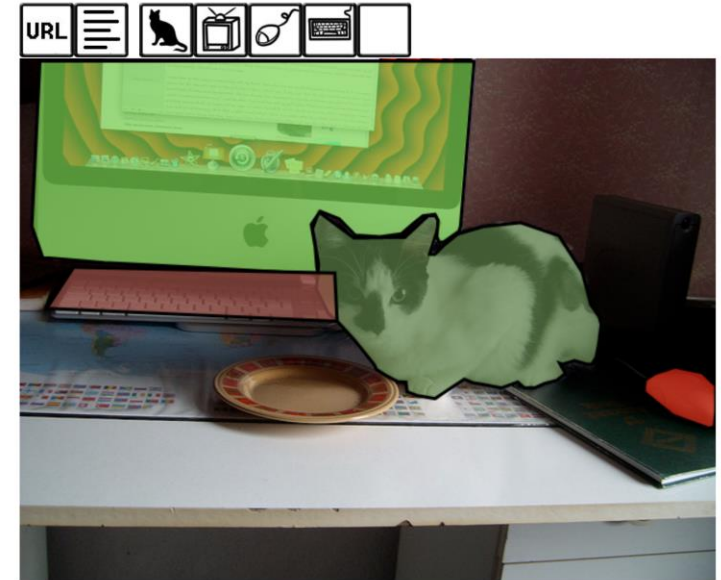
The main content area displays the search results for the query "car", stating "Your query 'car' matches 53 synsets." and a "Rank by" dropdown menu set to "Popularity".

The results are organized into four rows, each representing a different synset:

- Synset: car, elevator car**
Definition: where passengers ride up and down; "the car was on the top floor".
Popularity percentile: 97%
Depth in WordNet: 8
- Synset: cable car, car**
Definition: a conveyance for passengers or freight on a cable railway; "they took a cable car to the top of the mountain".
Popularity percentile: 95%
Depth in WordNet: 8
- Synset: sports car, sport car**
Definition: a small low car with a high-powered engine; usually seats two persons.
Popularity percentile: 93%
Depth in WordNet: 10
- Synset: car, auto, automobile, machine, motorcar**
Definition: a motor vehicle with four wheels; usually propelled by an internal combustion engine; "he needs a car to get to work".
Popularity percentile: 92%
Depth in WordNet: 9

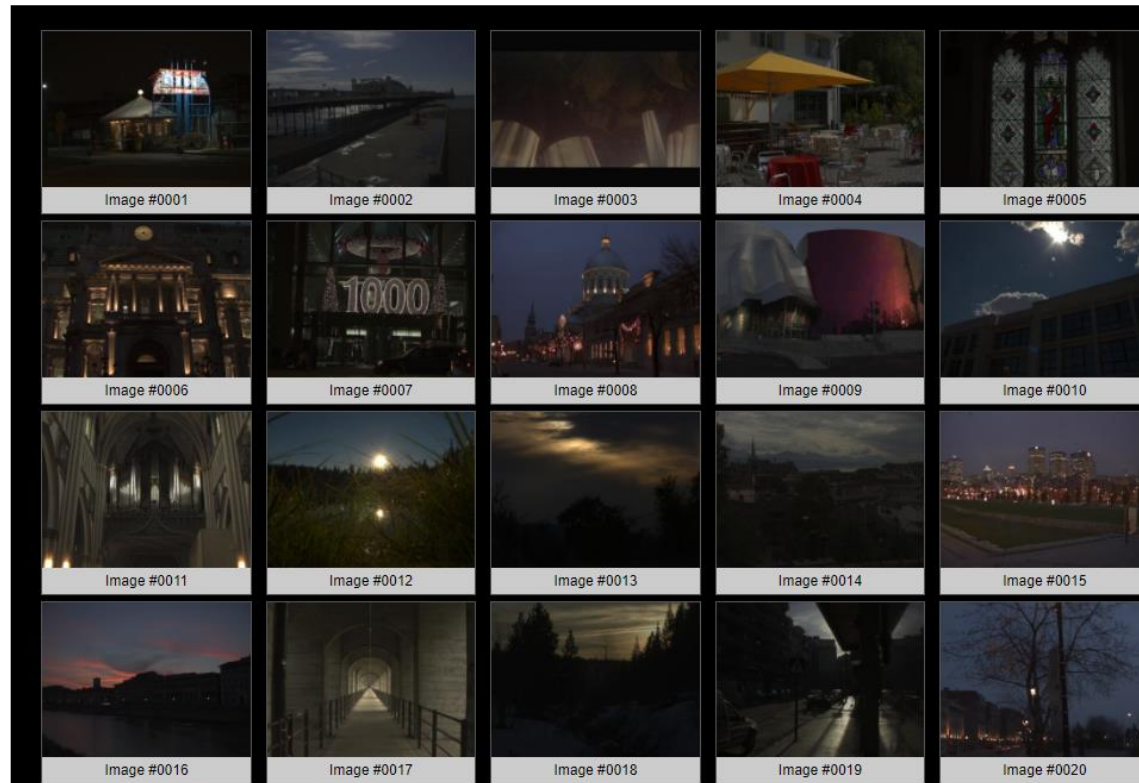
Image Datasets

- MS COCO
 - Boost to DL class/instance segmentation and keypoint detection
 - 80 classes
 - 200k images
 - Instance segmentation masks (>1 mil)
 - Human keypoints (250k)



LDR 2 HDR

- <http://hdrv.org/hdrcnn/material/testset/index.html>

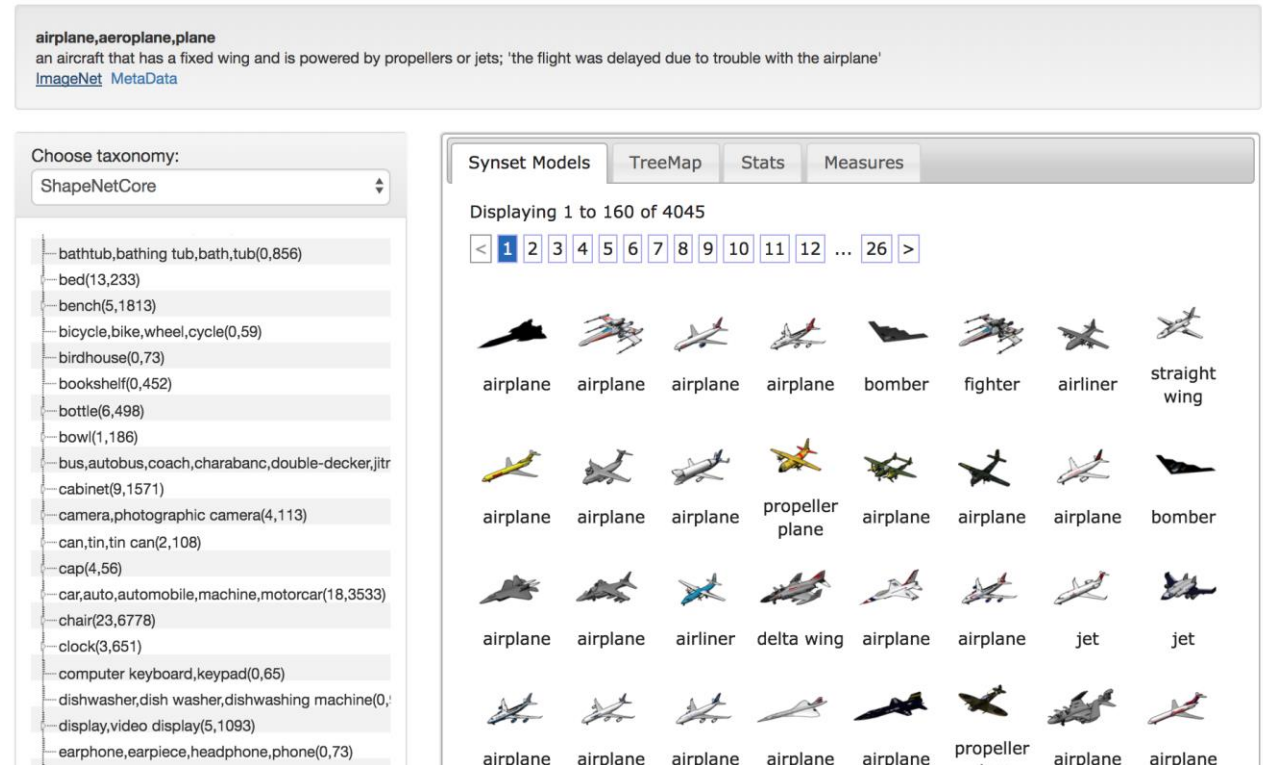


3D data sets

3D Datasets

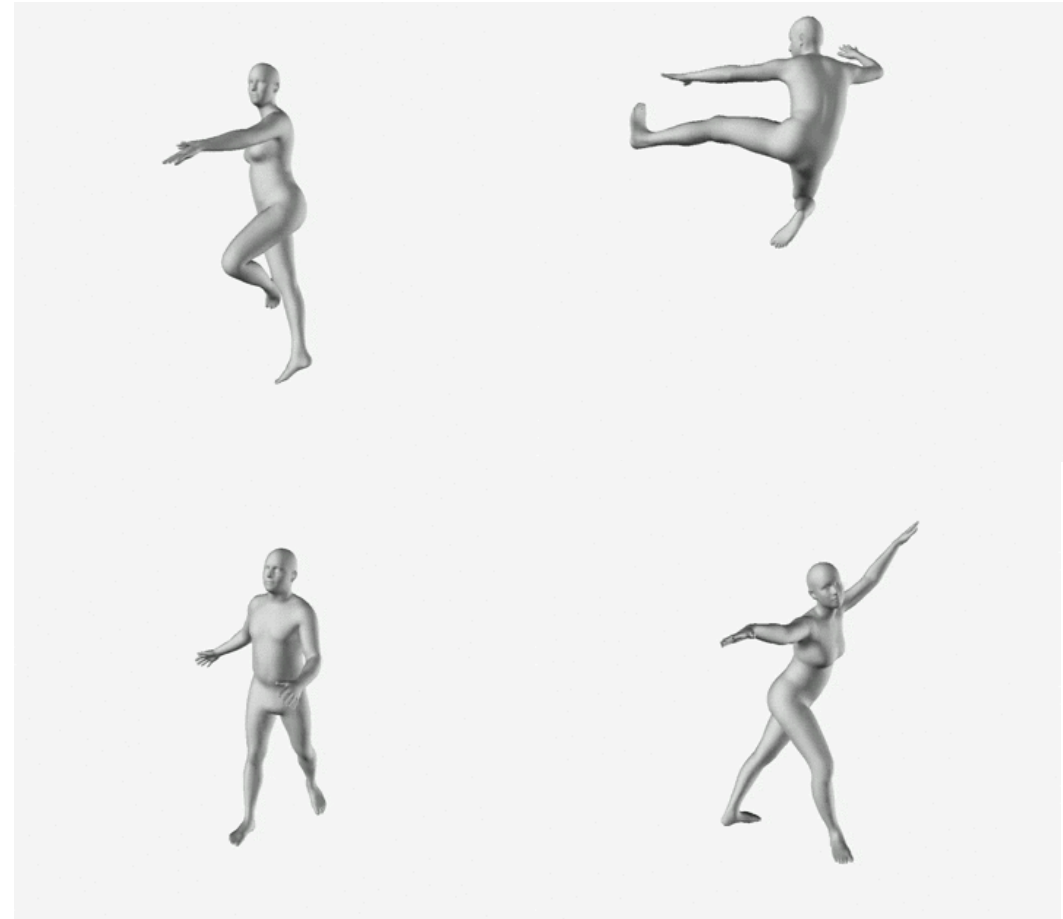
- ShapeNet

- Similar to ImageNet but for CAD models
- 55 common categories
- 10k+ models
- ShapeNetCore
 - 12k models
 - Additional annotations (real world dimensions, materials,...)



3D Datasets

- SMPL
 - Parametric human shape model
 - 72 parameters control pose and human shape
 - Fully differentiable
 - Useful for human shape estimation, motion capture etc



Data Augmentation

Data Augmentation

- Augment existing data with image operations to reduce overfitting
- Much larger dataset
- Learn expected transformations



Data Augmentation

- Mirror



Data Augmentation

- Mirror
- Rotation



Data Augmentation

- Mirror
- Rotation
- Translation



Data Augmentation

- Mirror
- Rotation
- Translation
- Zoom



Data Augmentation

- Mirror
- Rotation
- Translation
- Zoom
- Blur



Data Augmentation

- Mirror
- Rotation
- Translation
- Zoom
- Blur
- Noise



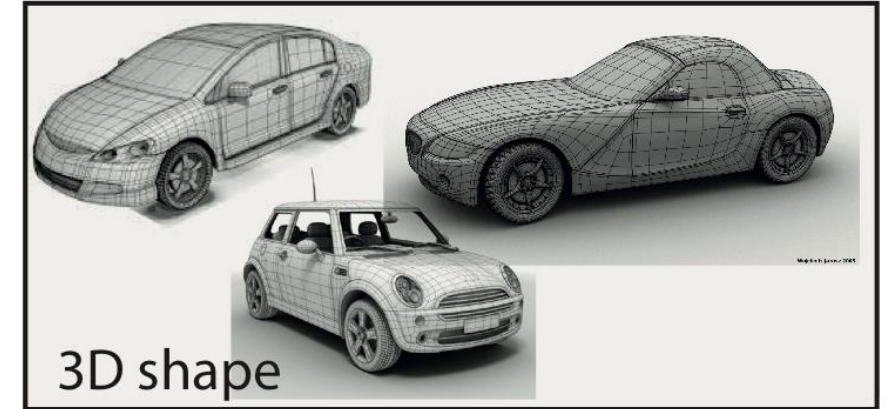
Data Augmentation

- Mirror
- Rotation
- Translation
- Zoom
- Blur
- Noise
- Color transforms



Creating Your Own

Material/Illumination



Material/Illumination



Decomposition

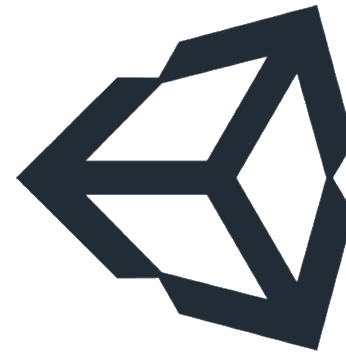
$$\text{Image} = \text{Occlusion} \times (\text{Albedo} \times \text{Irradiance} + \text{Specular})$$

$$\text{Image} = \text{Bottom}^* + \text{Top}^* + \dots + \text{Left}^* + \text{Right}^*$$

$$\text{Top}^* = \text{Occlusion} \times (\text{Albedo} \times \text{Irradiance Top} + \text{Specular Top})$$

Synthetic data

Synthetic Data

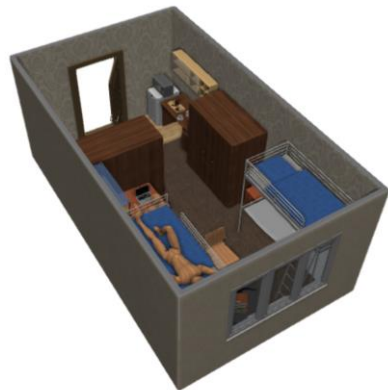


Synthetic Data for DL

- 3D models + renderer = unlimited data
- Suitable for data hungry approaches such as deep networks
- Higher fidelity -> smaller discrepancy between synthetic and real

How To Generate Synthetic Data

- What you need
 - 3D models with task-specific annotation
 - Renderer
- Example: Indoor depth estimation (McCormac et al)



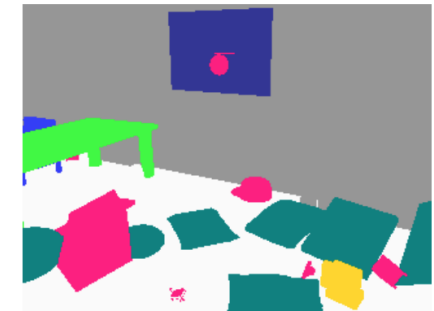
3D Room with furniture

+



NVIDIA OptiX renderer

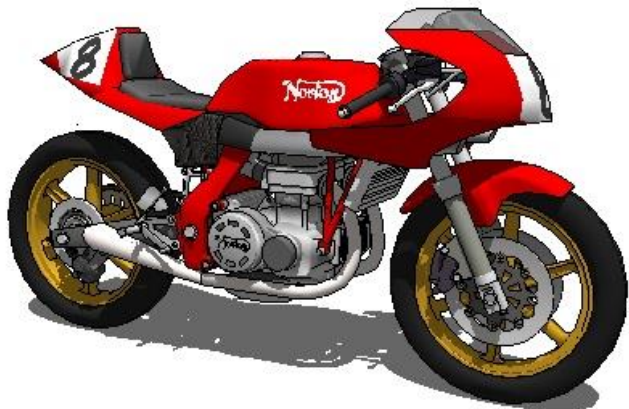
=



Data with annotation

How To Generate Synthetic Data

- How much fidelity?



OpenGL



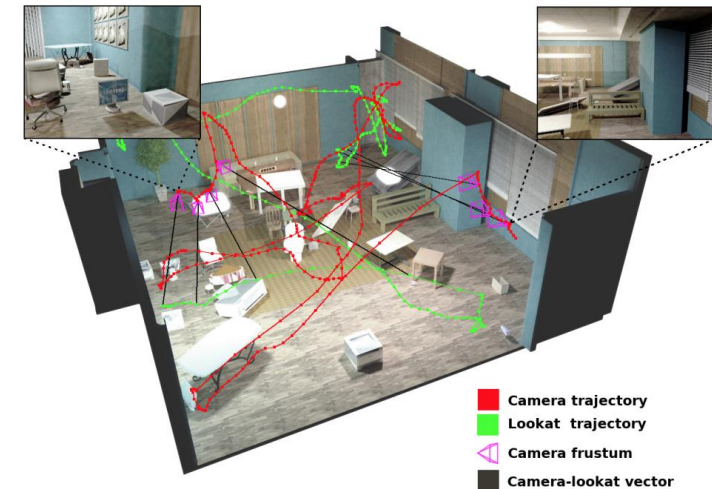
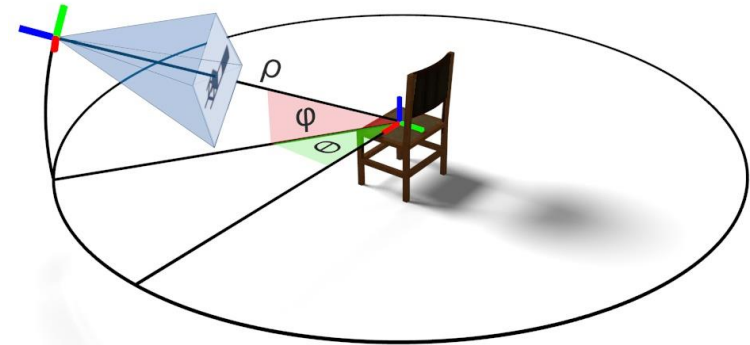
Natural Illumination + AO



Path tracer, global Illumination (VRay)

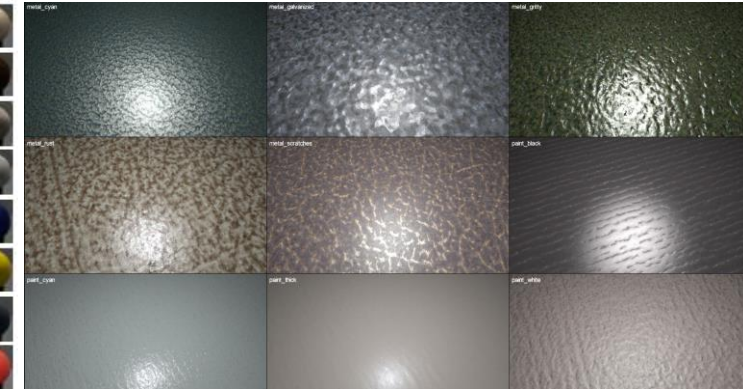
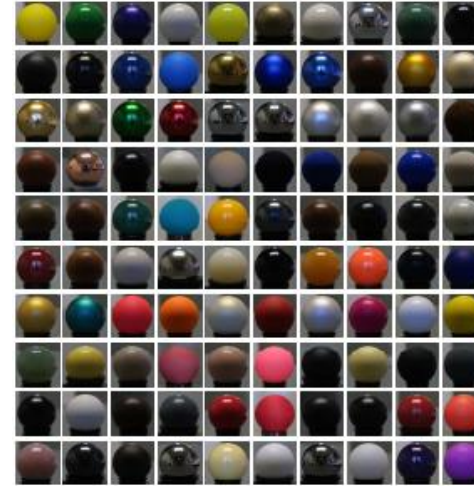
Selecting viewpoint

- Object in the center
 - Sample from the hemisphere
 - Multiple FOV, Target, Up, etc
- Scene
 - Simulate human camera path
 - Optimize camera position for a particular objective (eg segmentation)

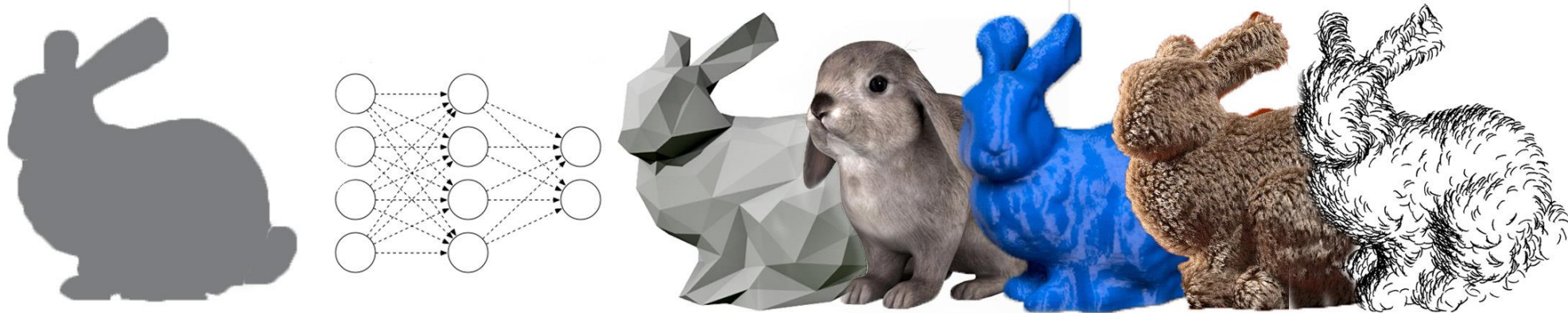


Selecting Appearance

- Materials
 - Capture the variability of real world objects
 - BRDF, textures (MERL DB)
- Illumination
 - Capture the effect of environment
 - Increase realism
 - Laval indoor DB
<http://indoor.hdrdb.com/>



Thank you!



<http://geometry.cs.ucl.ac.uk/dl4g/>