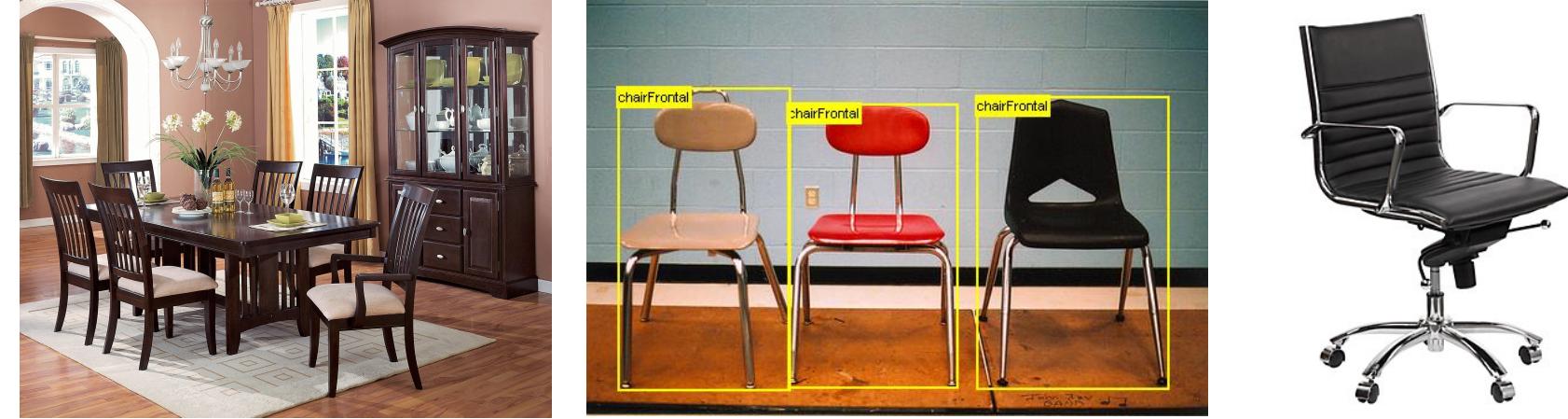


Estimating Chair Pose Using ConvNets

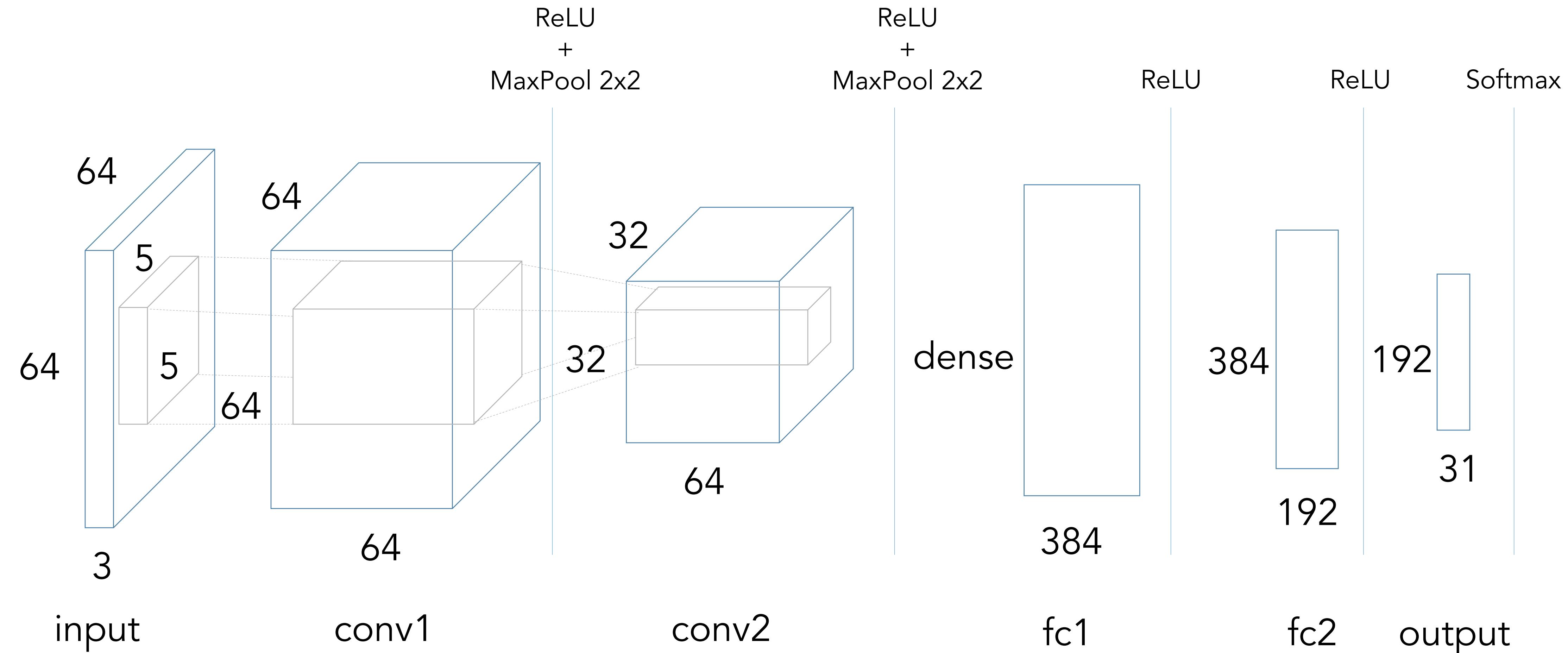
Keunhong Park and Aditya Sankar

Motivation



- Estimating the 3D pose of objects in 2D images is a challenging task
- Accurately labeled 3D pose data is difficult to acquire at scale
- Can enable a variety interesting applications such as interior design, remodeling, robot vision, virtual reality
- We initially focus on a commonly occurring class of objects: Chairs

Network Architecture

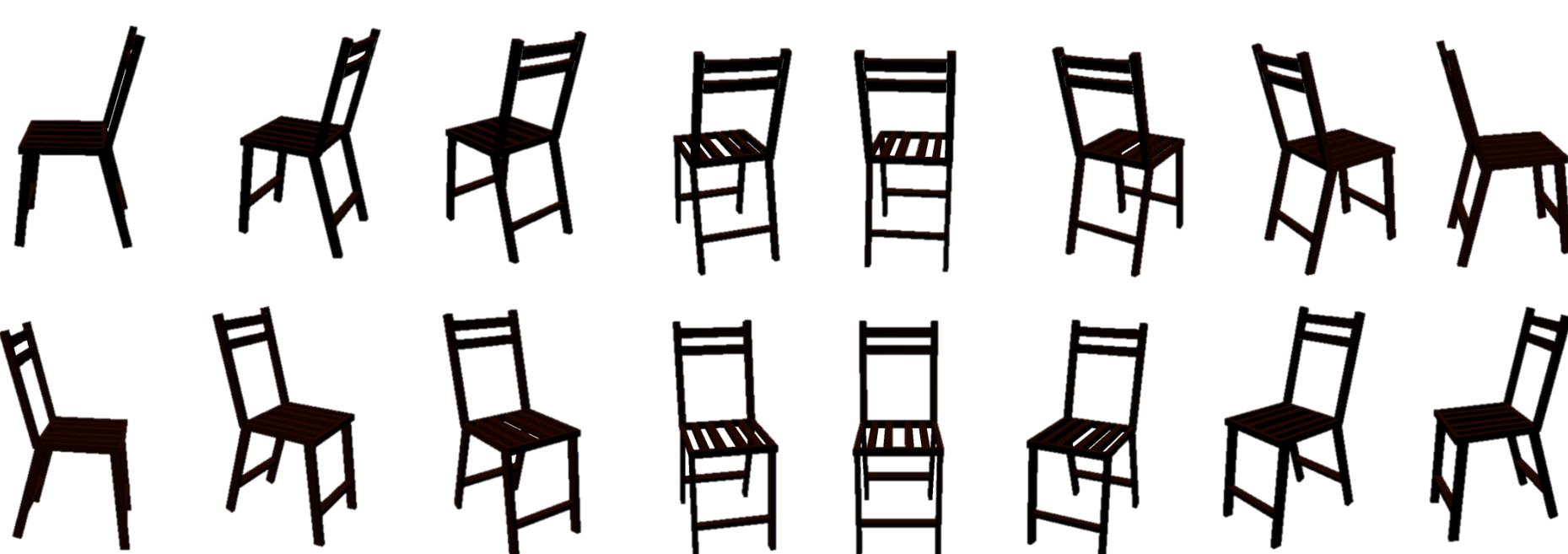


Training Data

While it is difficult to acquire labeled training data, we observe that large repositories of 3D models exist online. We leverage these to generate synthetic training examples by rendering chairs in various poses. A method also proposed by [1].



Chair 3D Models



Example Rendered Synthetic Views

We have 1393 distinct models with 31 discretized azimuth & 2 elevation angles. Totaling 172732 training images.

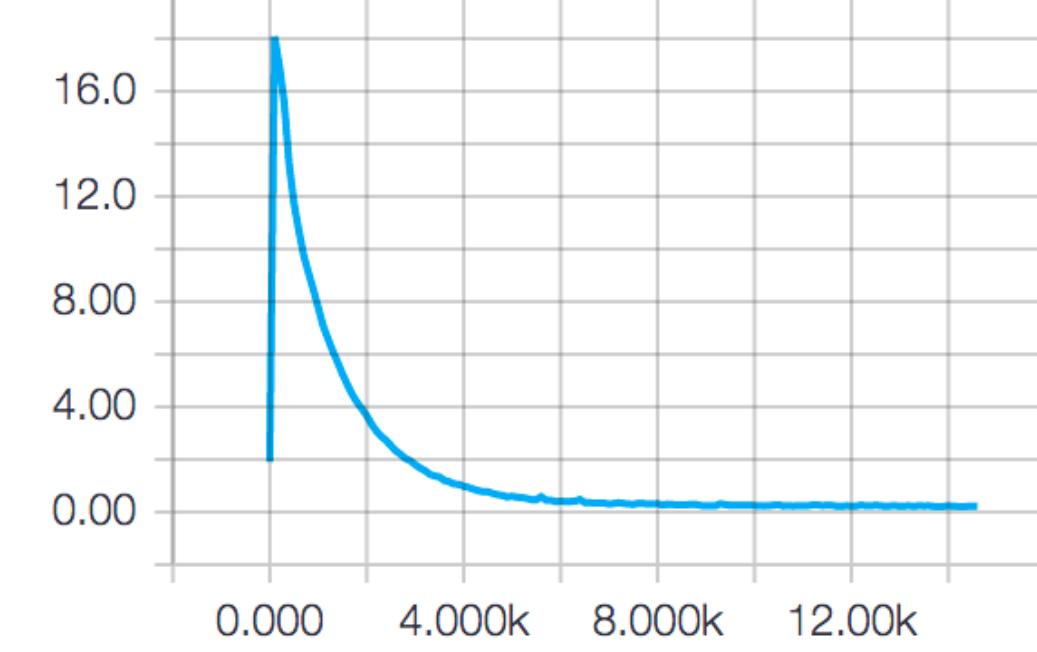
[1] Seeing 3D Chairs, Aubry et al., CVPR 2014

Results

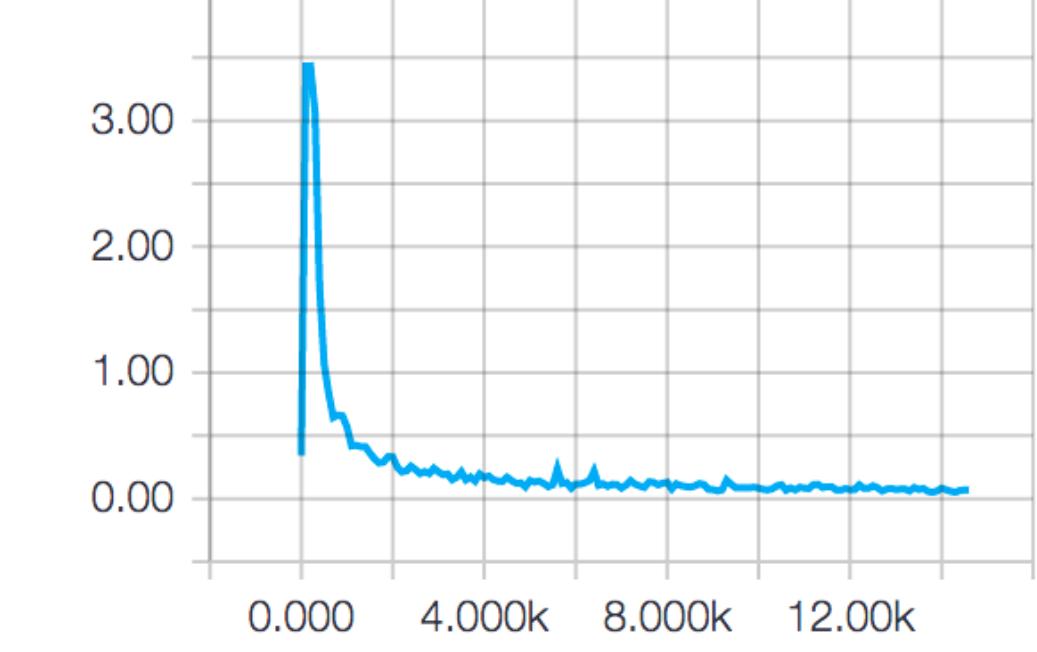
Quantitative

Synthetic Images		PASCAL VOC	
Precision (Training)	1.000	Precision	0.1415
Precision (Testing)	0.924	Precision (top k)	0.3184
		Precision (within 1 class)	0.2712

Loss vs Iterations



Cross Entropy vs Iterations



Failure Cases

- Cluttered backgrounds
- Noise, Not centered
- Occlusion



Interactive Demo: ~3.33fps



Qualitative (product images from Bing)

Image	Confidence/Azimuth	4.6356 / 325	4.4407 / 336	3.6498 / 313
0264.jpg				
0173.jpg				
0218.jpg				
0057.jpg				
0436.jpg				