Deep Learning for Graphics

Supervised Applications

Niloy Mitra
UCL

Iasonas Kokkinos
UCL/Facebook

Paul Guerrero
UCL

Vladimir Kim
Adobe Research

Kostas Rematas
U Washington

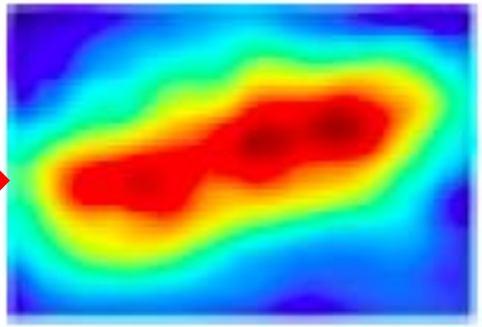
Tobias Ritschel
UCL

Timetable

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

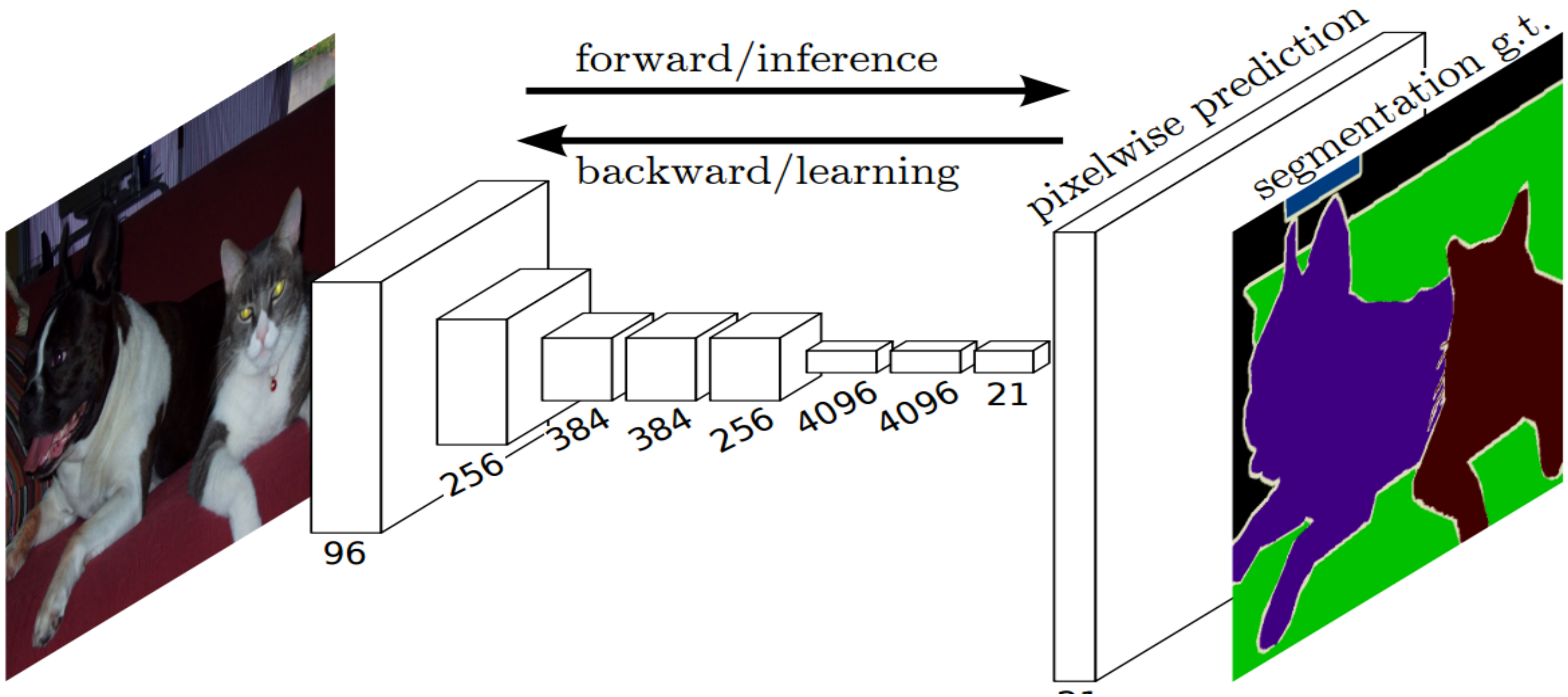
Fully-Convolutional Network (FCN)

FCN



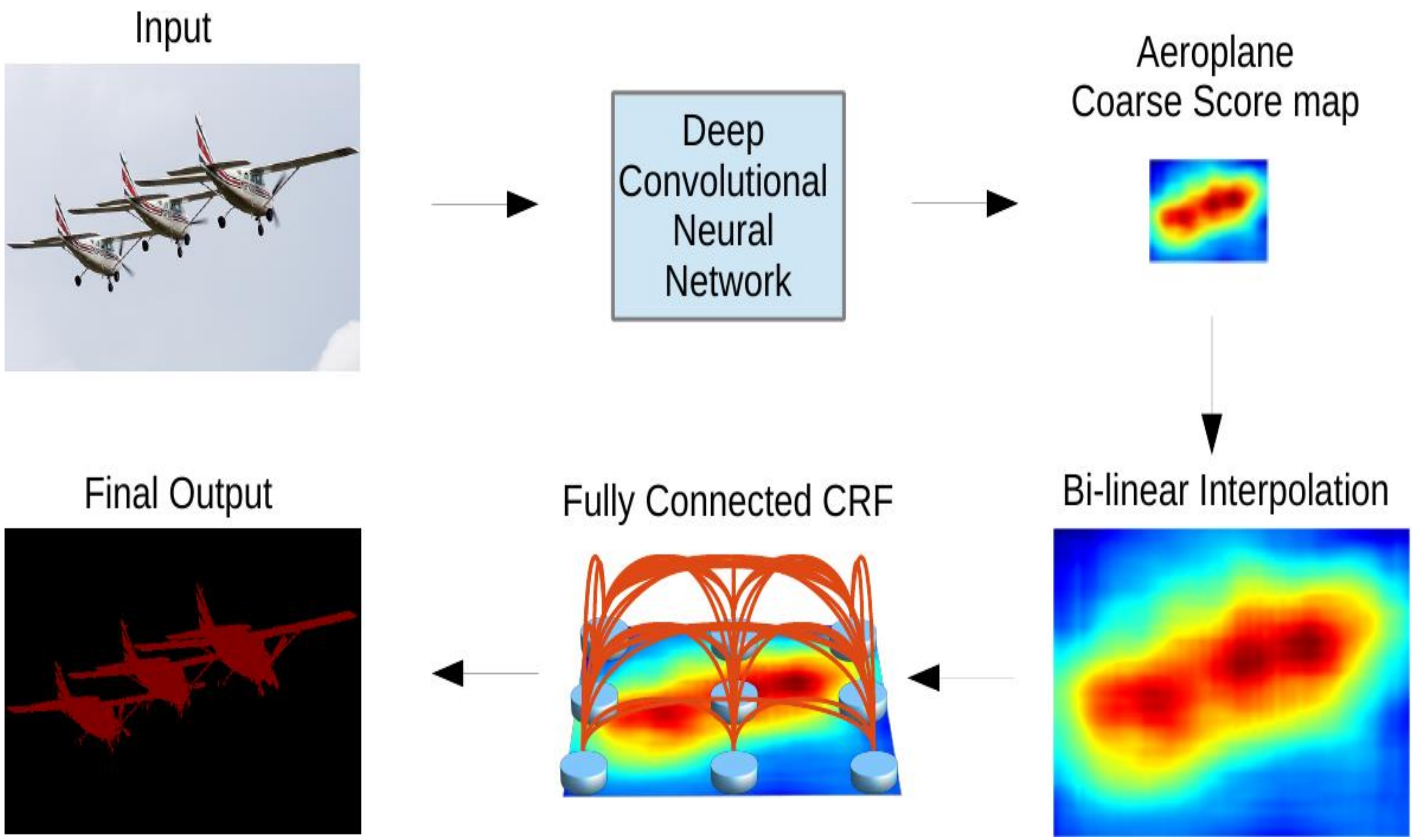
Fast (shared convolutions)
Simple (dense)

FCN-based semantic segmentation



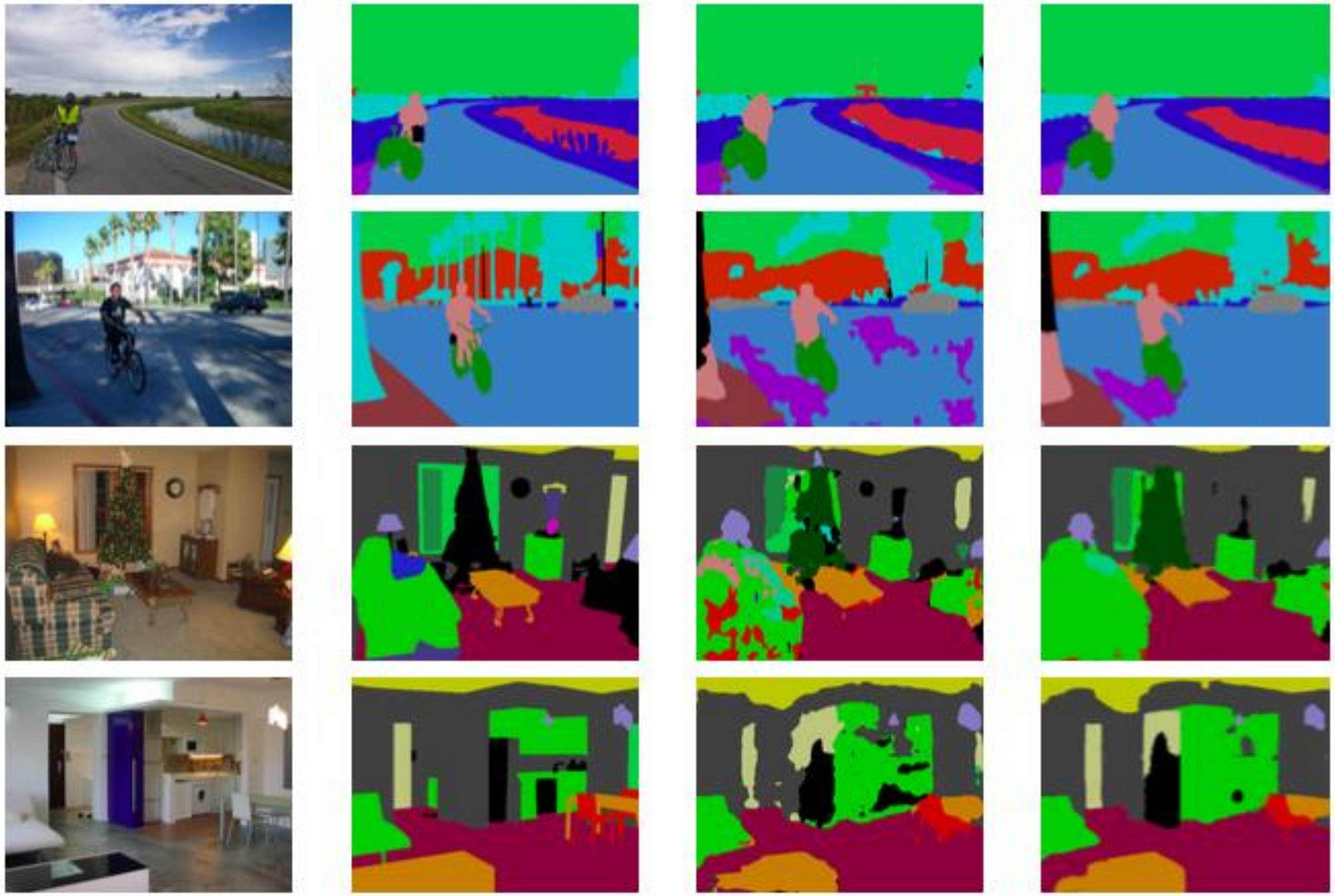
J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. *CVPR*, 2015

FCN-CRFs: Deeplab



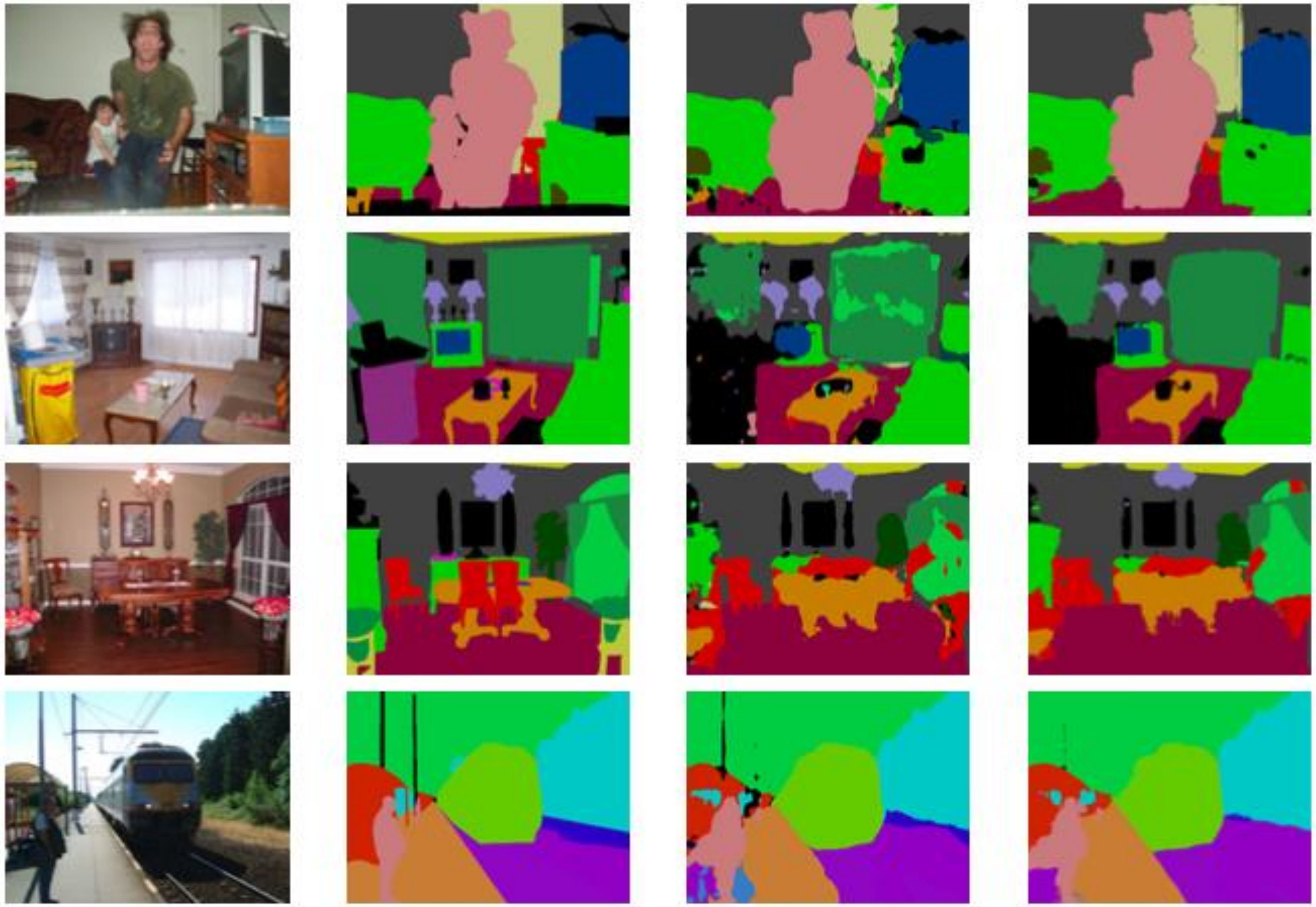
L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy and A. Yuille,
Deeplab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs, PAMI 2016

Deeplab v2 results



Ground truth FCN FCN-DCRF

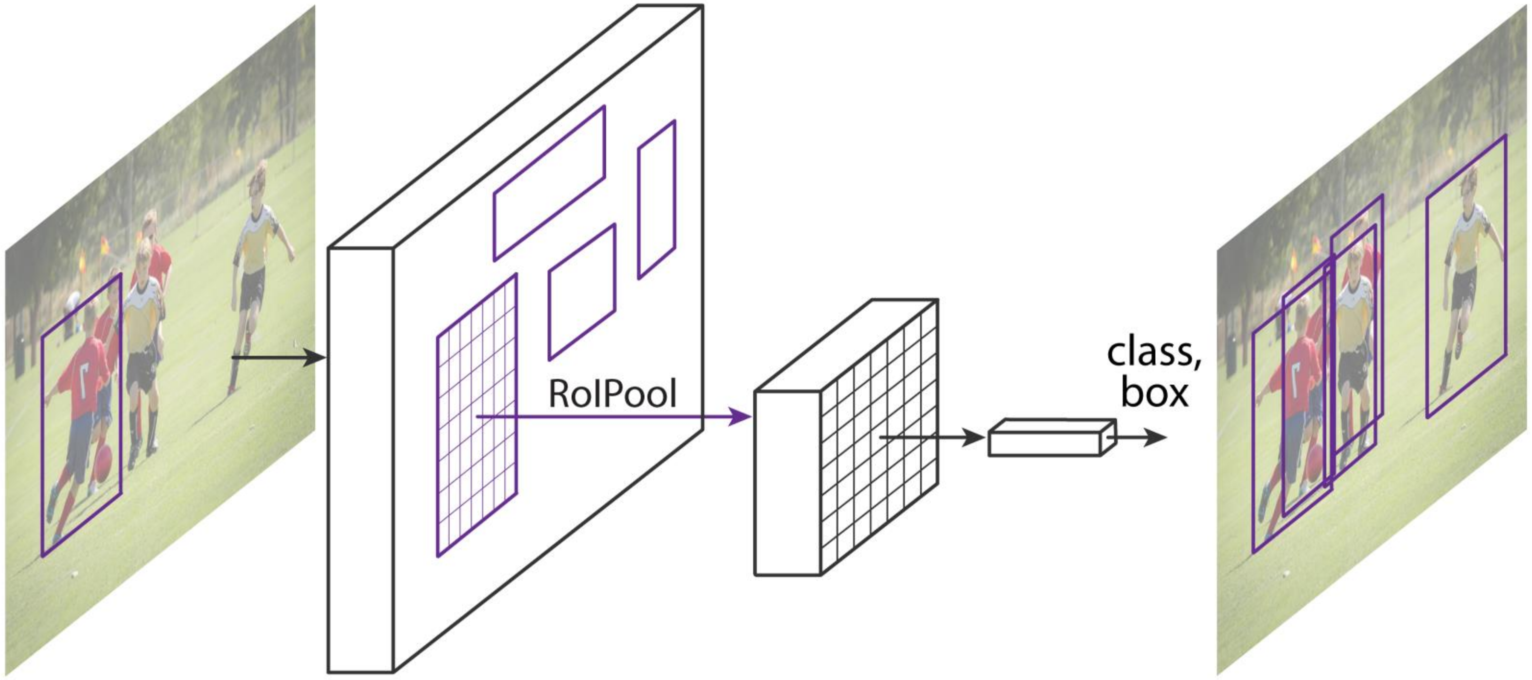
Deeplab v2 results



Ground truth FCN FCN-DCRF

Object Detection: Fast(er)-RCNN

- Fast/Faster R-CNN
  - ✓Good speed
  - ✓Good accuracy
  - ✓Intuitive
  - ✓Easy to use

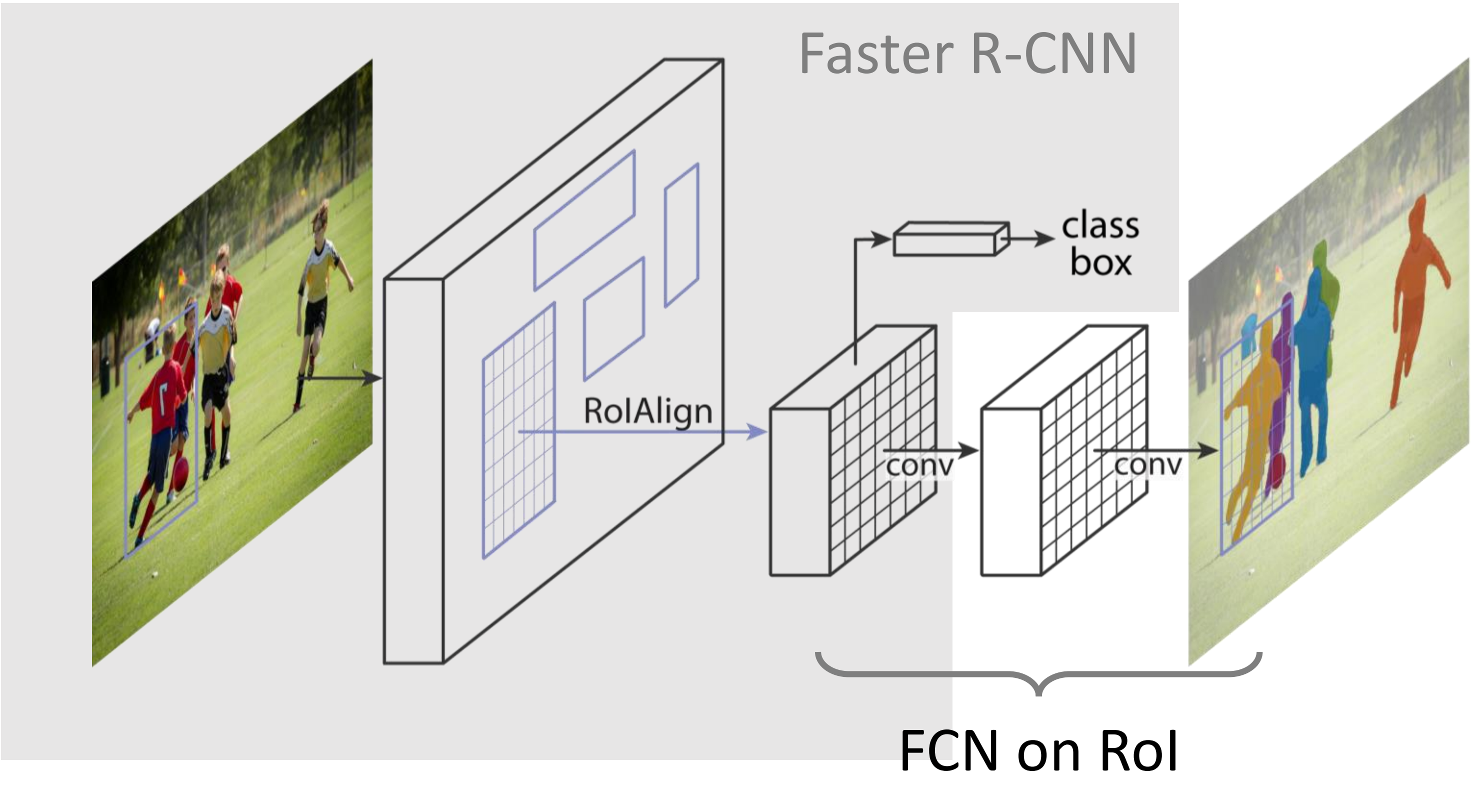


Ross Girshick. “Fast R-CNN”. ICCV 2015.

Shaoqing Ren, Kaiming He, Ross Girshick, & Jian Sun. “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks”. NIPS 2015.

Mask R-CNN

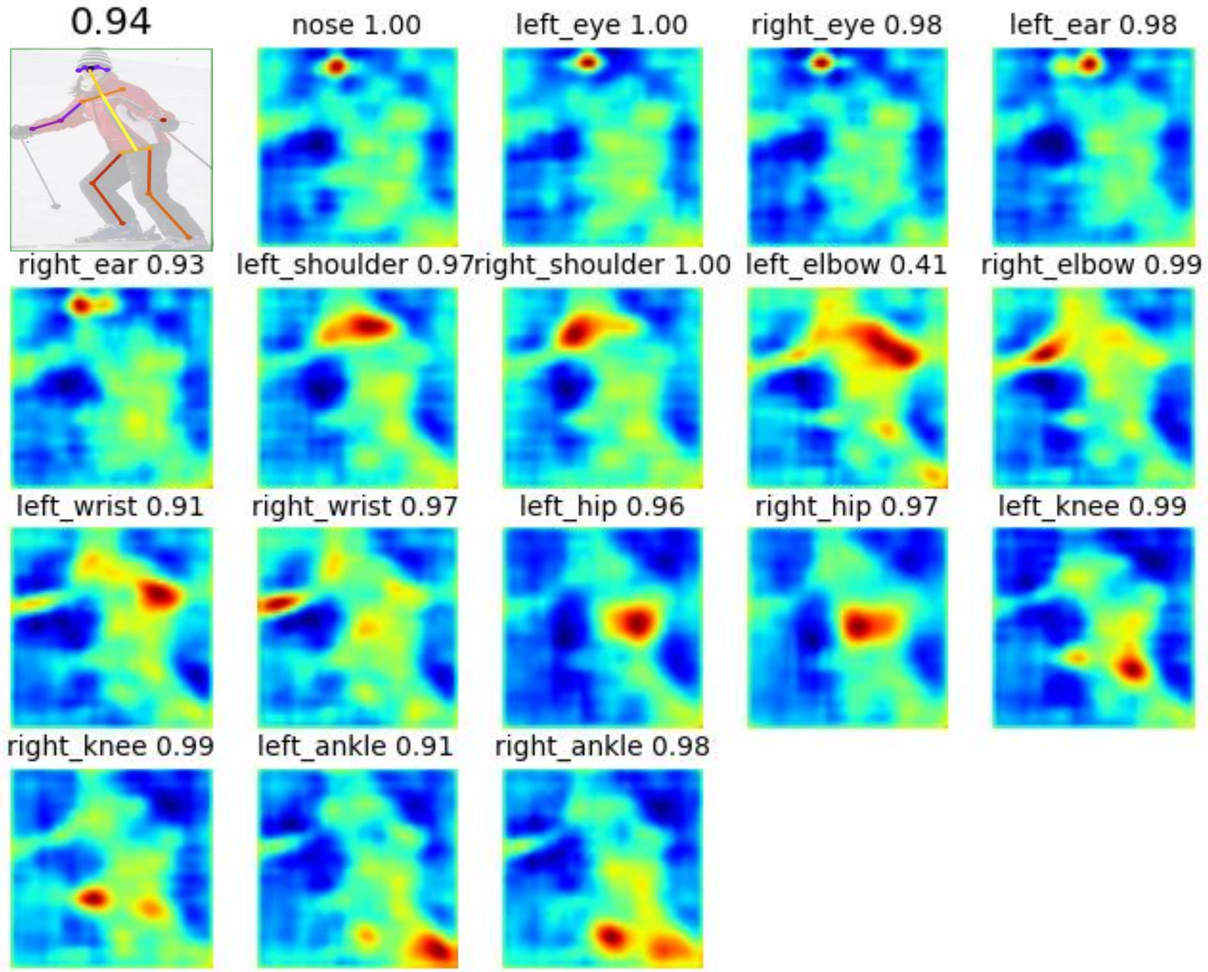
- Mask R-CNN = **Faster R-CNN** with **FCN** on RoIs

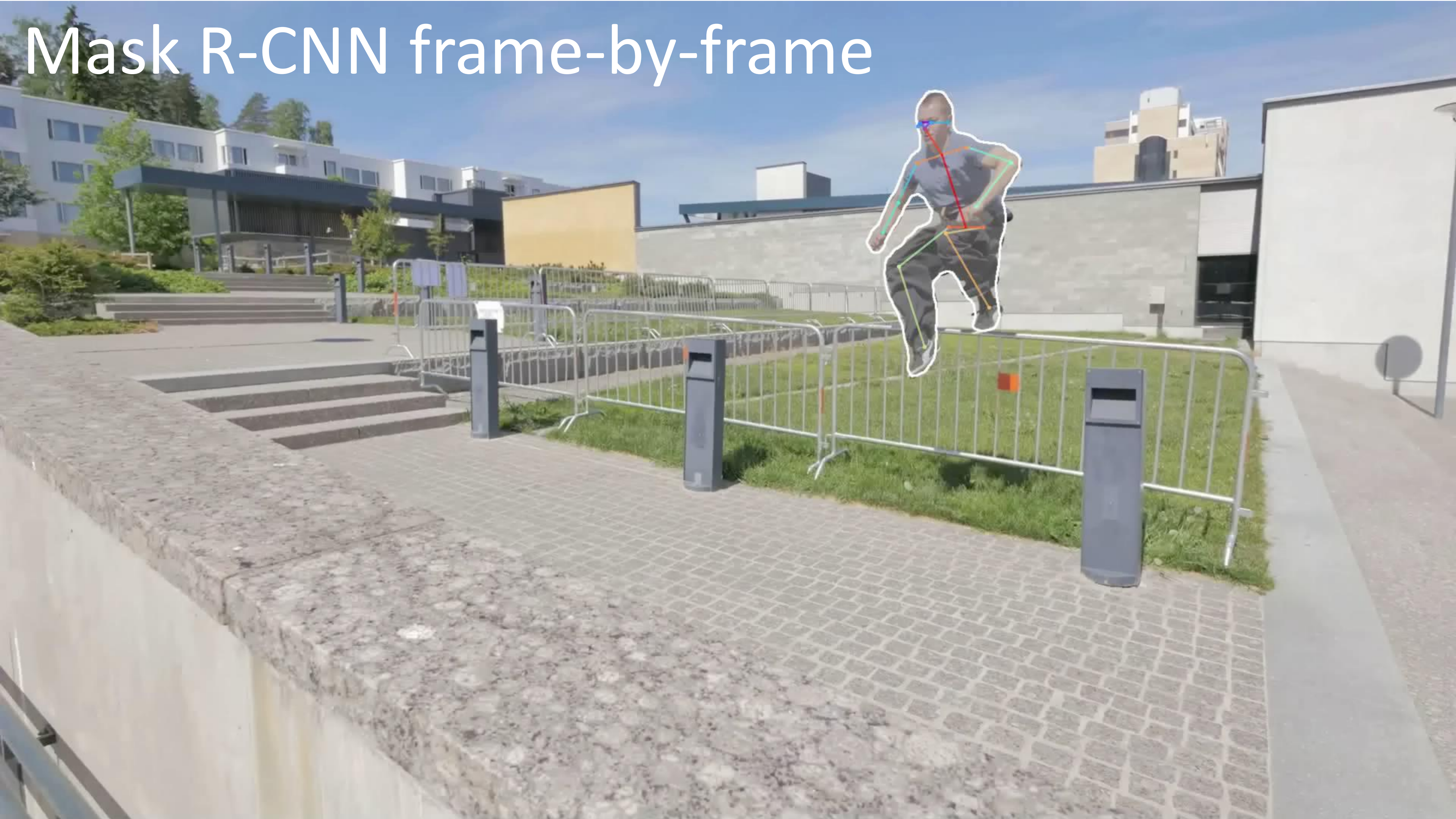


Mask R-CNN results on COCO

Mask R-CNN for Human Keypoint Detection

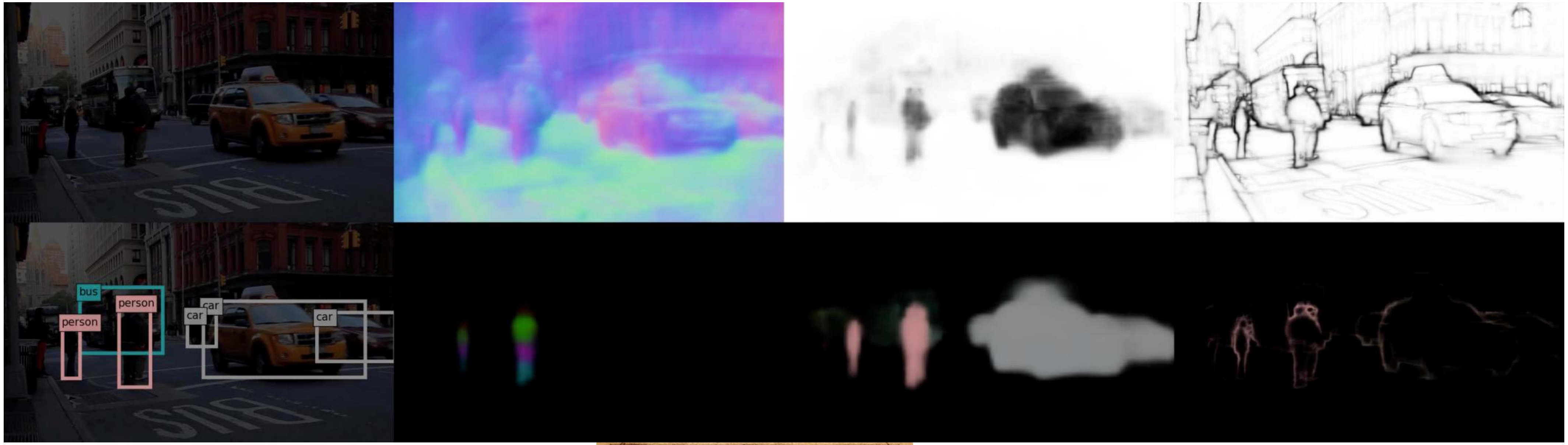
- 1 keypoint = 1-hot "mask"
- Human pose = 17 masks
- Softmax over spatial locations
  - e.g. 56^2-way softmax on 56x56

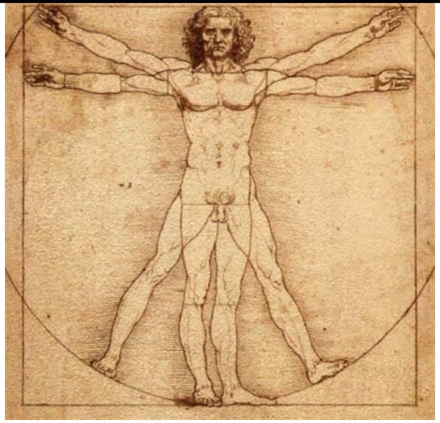



Mask R-CNN frame-by-frame


Mask R-CNN frame-by-frame

UberNet: a "universal" network for all tasks





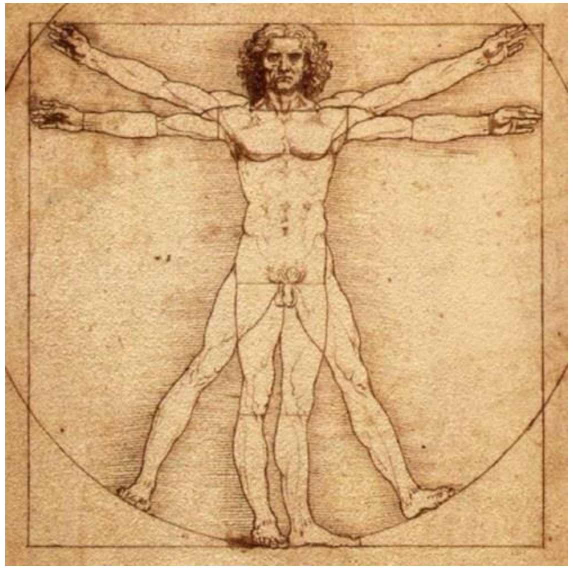
https://github.com/jkokkin/UberNet

I. Kokkinos, UberNet: Training a Universal CNN for *Low- Mid- and High-Level* Vision, CVPR 2017

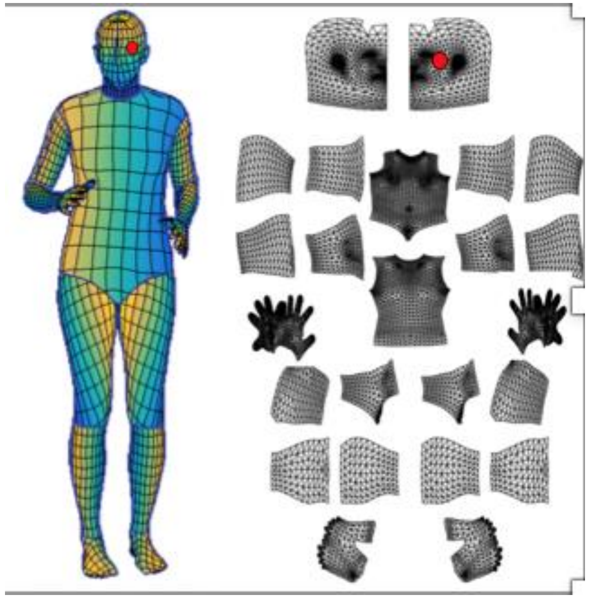
What is the ultimate vision task?

"Inverse graphics": understand how an image was generated from a scene

If we focus on a single object category: surface-based models

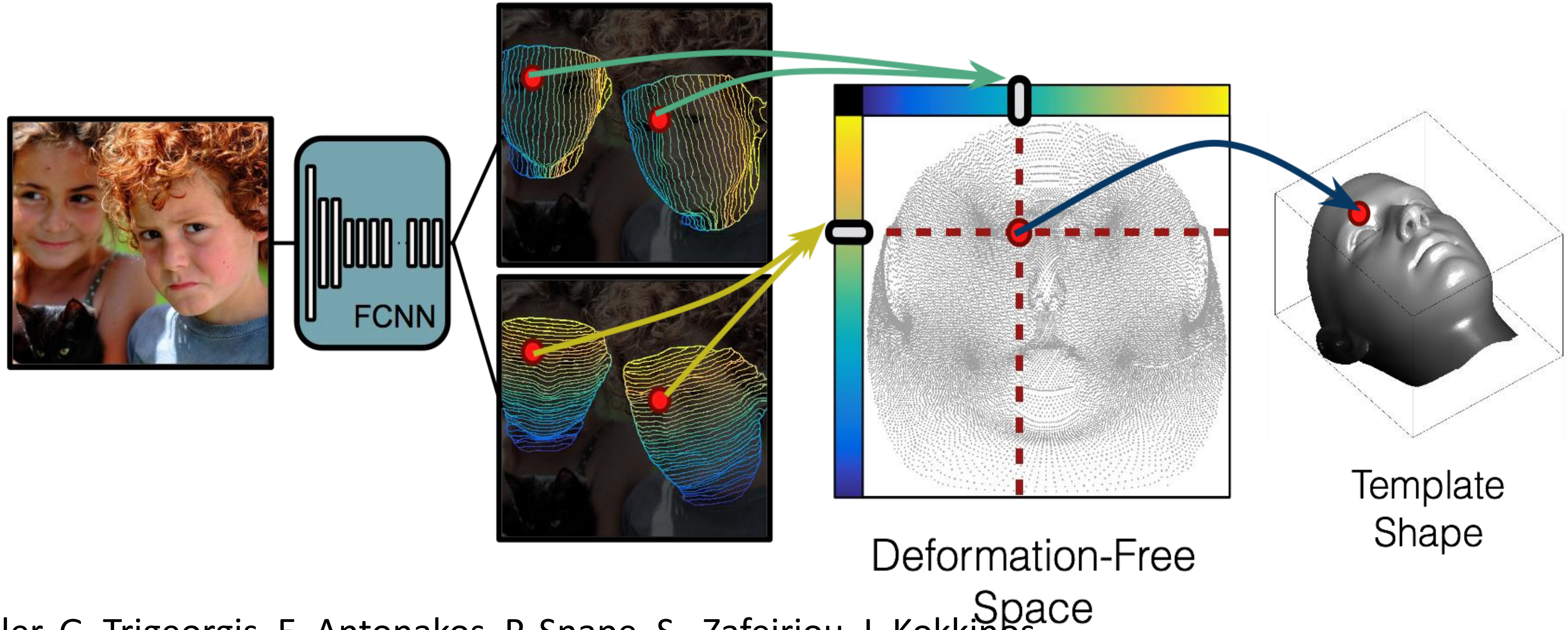


UberNet:
Universal Network



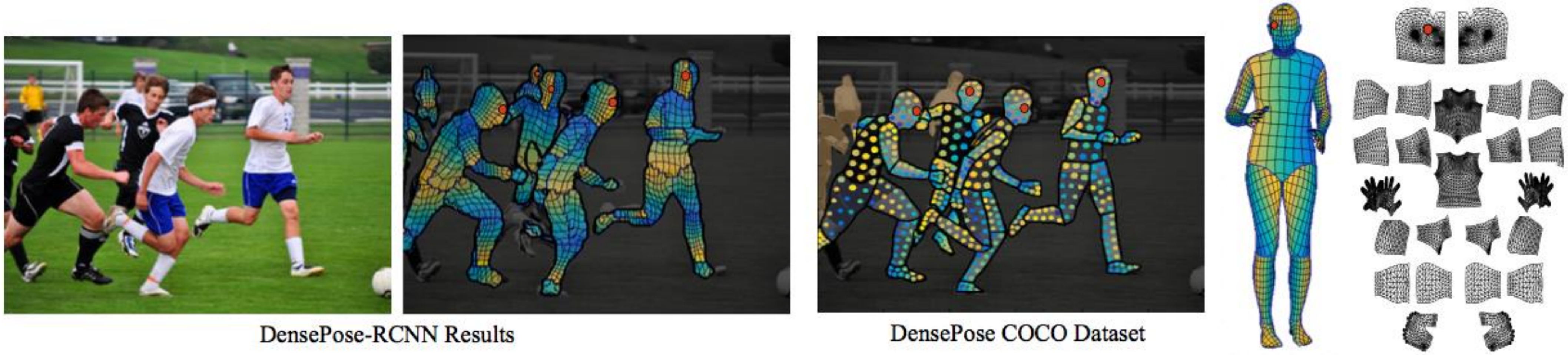
DensePose:
Unified model

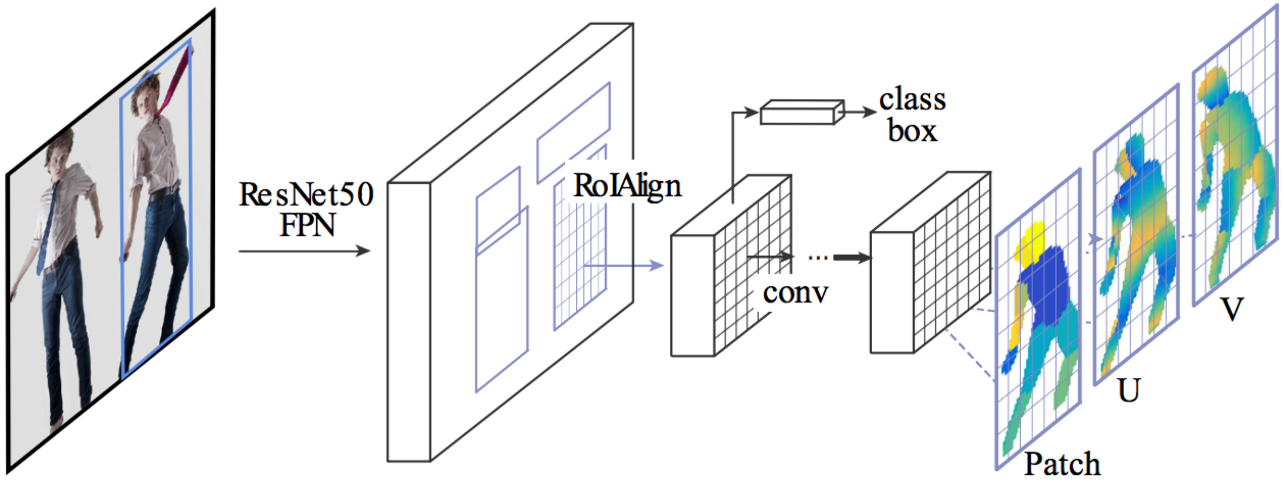
DenseReg: dense image-to-face regression



R. A. Guler, G. Trigeorgis, E. Antonakos, P. Snape, S. Zafeiriou, I. Kokkinos,
DenseReg: Fully Convolutional Dense Shape Regression In-the-Wild, CVPR 2017

DensePose: dense image-to-body correspondence

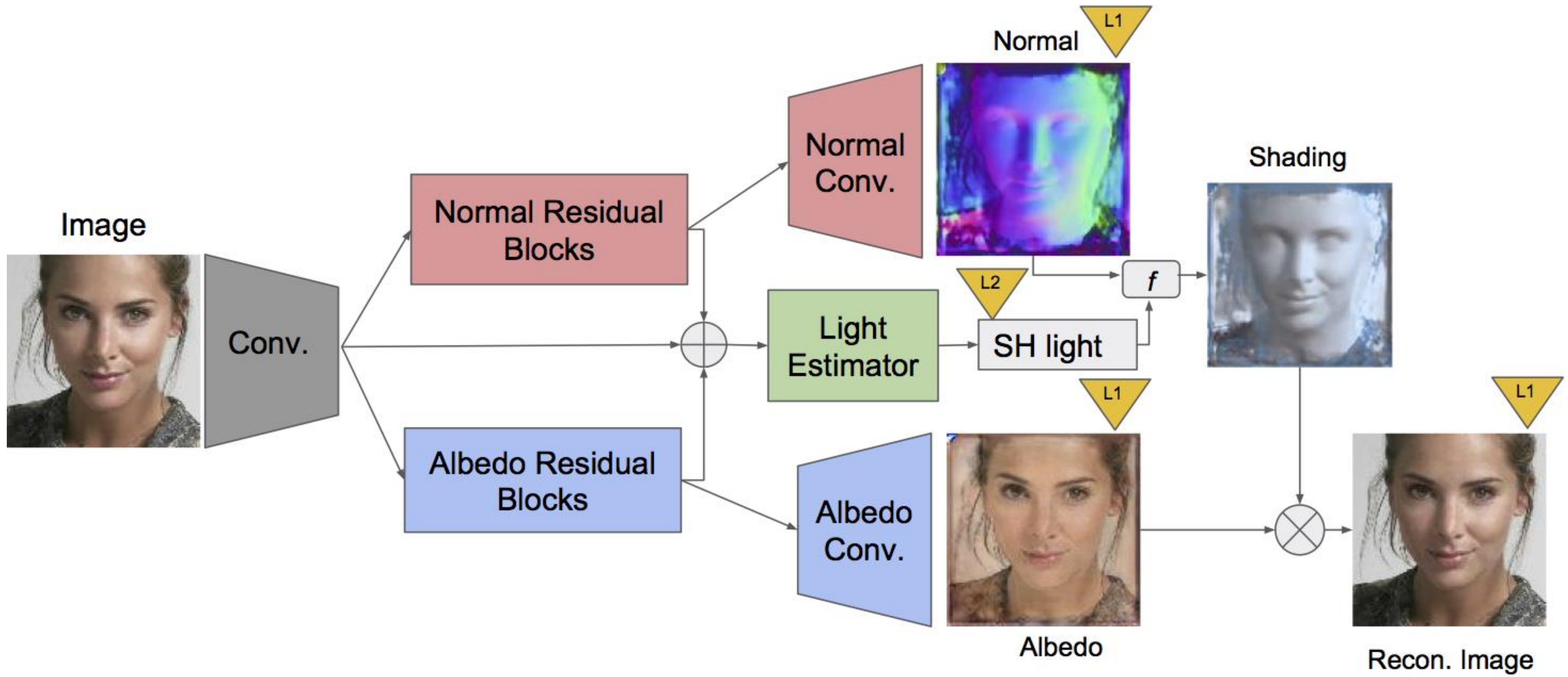




DensePose-RCNN: ~25 FPS

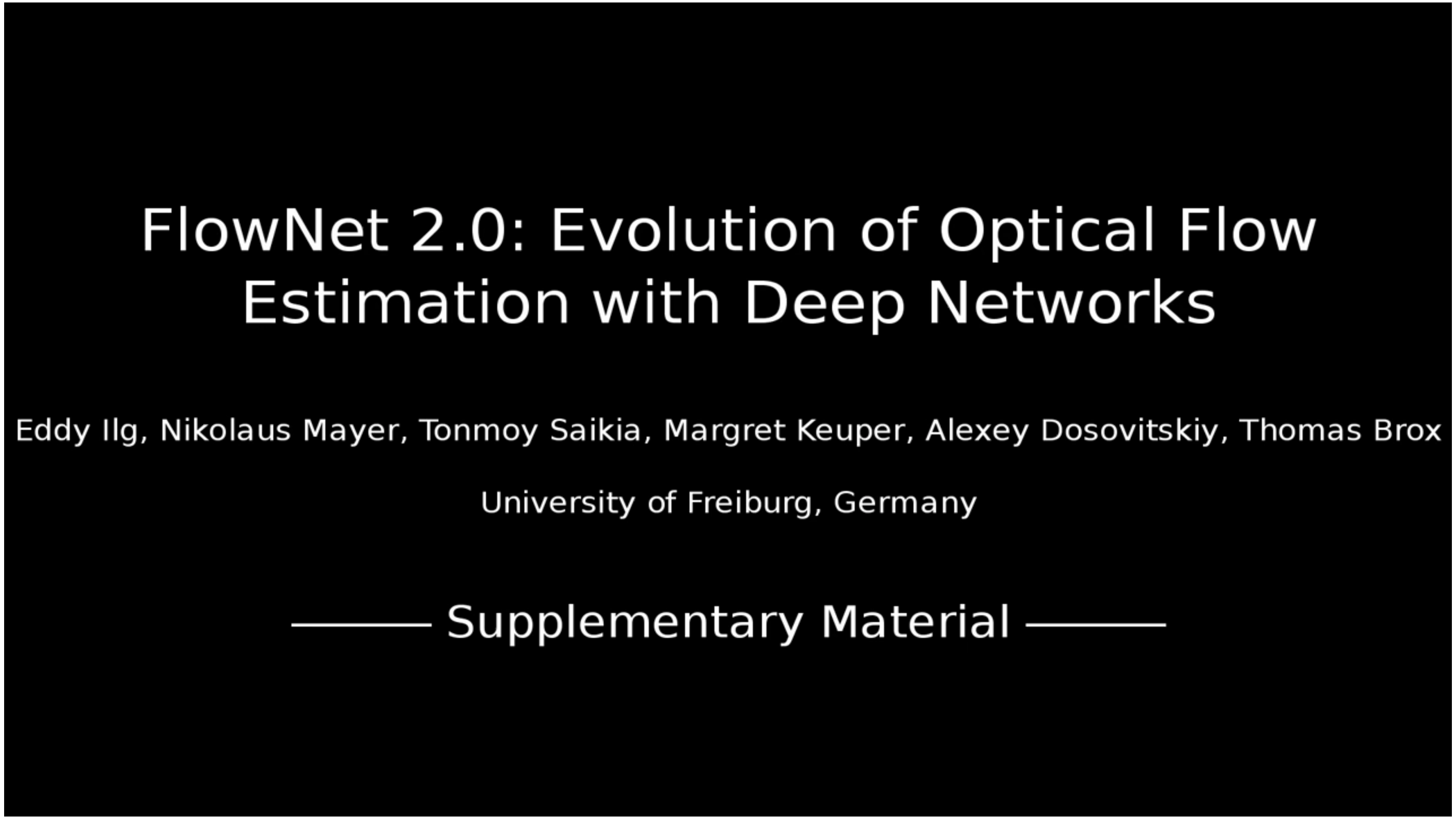
R. A. Guler, N. Neverova, I. Kokkinos “DensePose: Dense Human Pose Estimation In The Wild”, CVPR’18

SFSNet: incorporating image formation in model

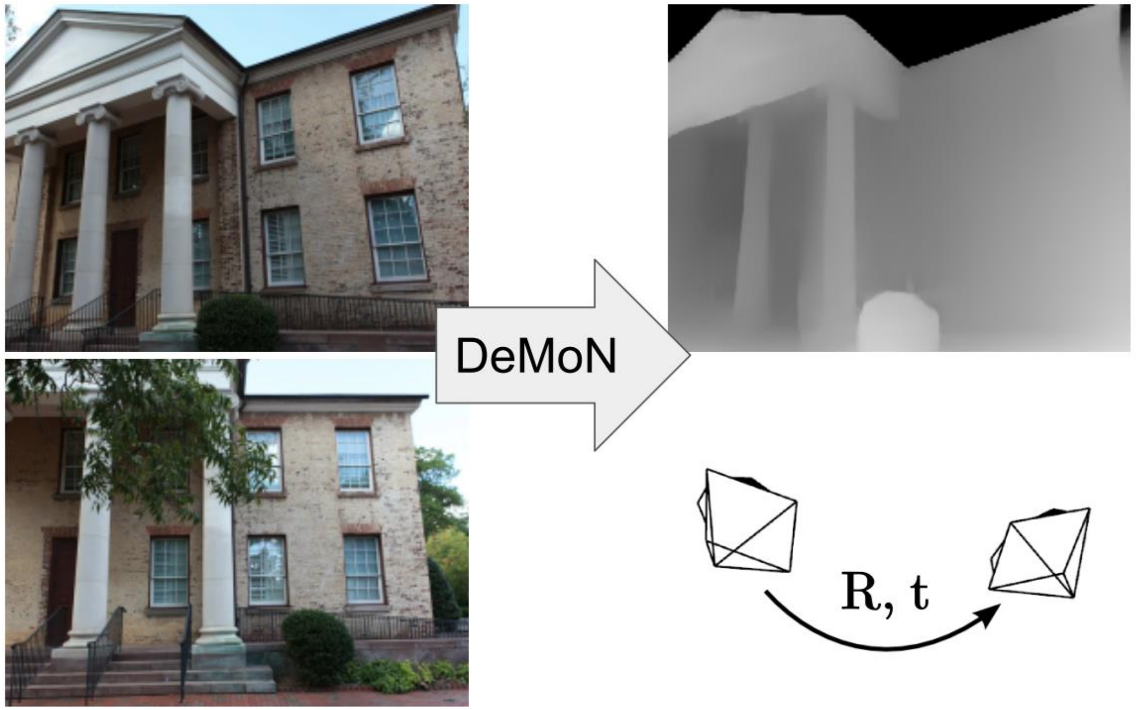


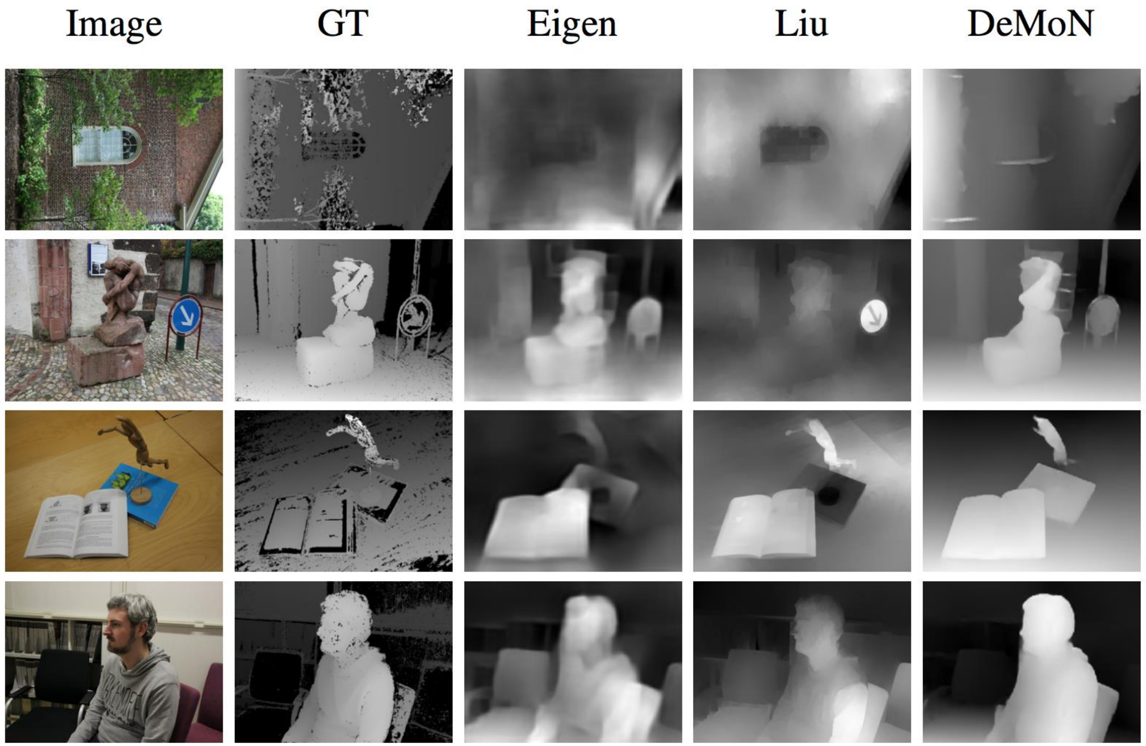
SfSNet: Learning Shape, Reflectance and Illuminance of Faces ‘in the wild' Soumyadip Sengupta Angjoo Kanazawa Carlos D. Castillo David W. Jacobs, CVPR 2018

Beyond single frames: end-to-end optical flow



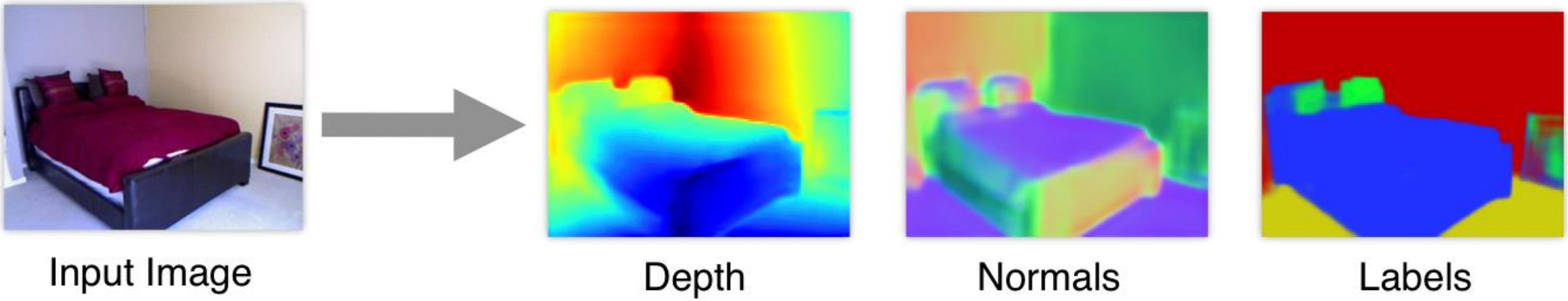
End-to-end Structure From Motion





- DeMoN: Depth and Motion Network for Learning Monocular Stereo, B. Ummenhofer, et al, CVPR 2017
- Unsupervised learning of depth and ego-motion from video, T Zhou, M Brown, N Snavely, DG Lowe, CVPR 2017

Monocular depth & normal estimation

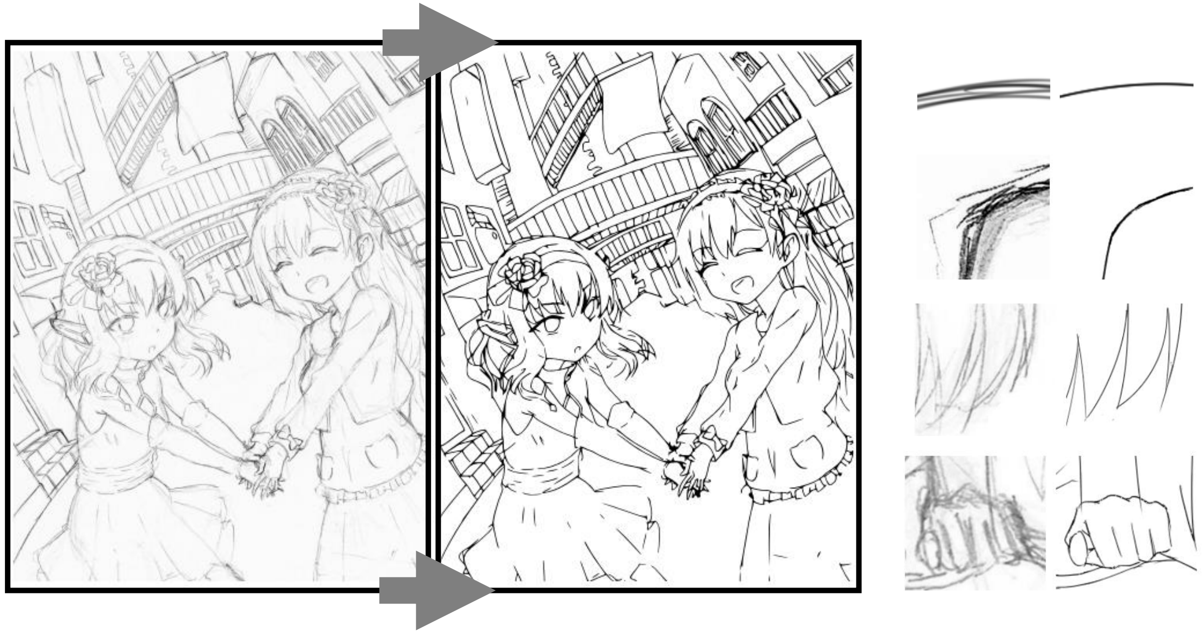


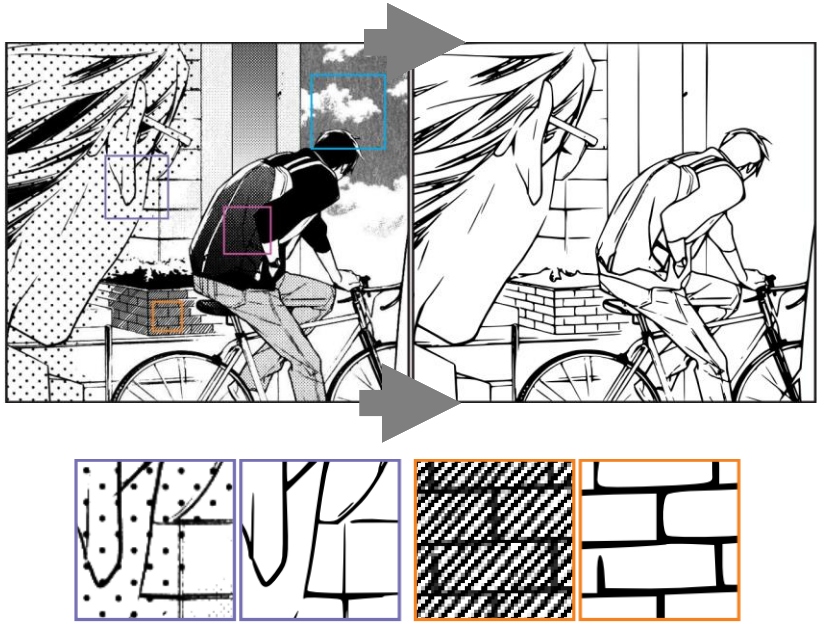
- D. Eigen and R. Fergus, Predicting Depth, Surface Normals and Semantic Labels with a Common Multi-Scale Convolutional Architecture, ICCV 2015

Graphics applications

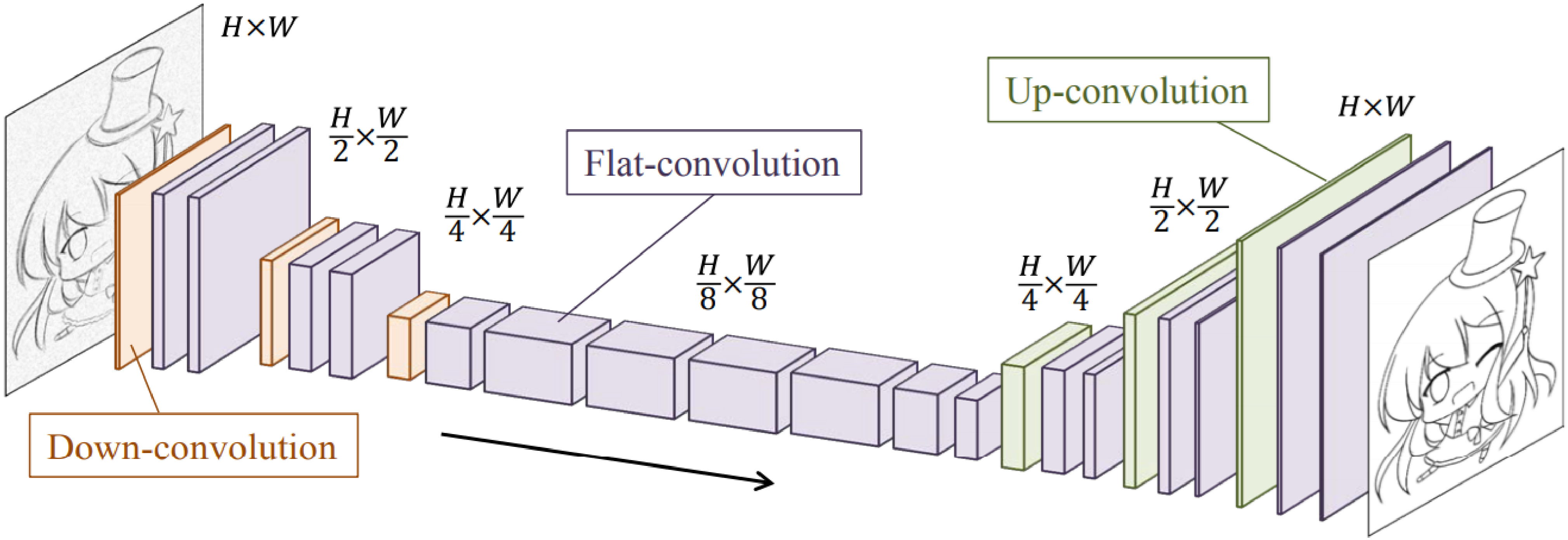
Sketch Simplification

- *Learning to Simplify: Fully Convolutional Networks for Rough Sketch Cleanup*, Simon-Serra et al., 2016
- *Deep Extraction of Manga Structural Lines*, Li et al., 2017





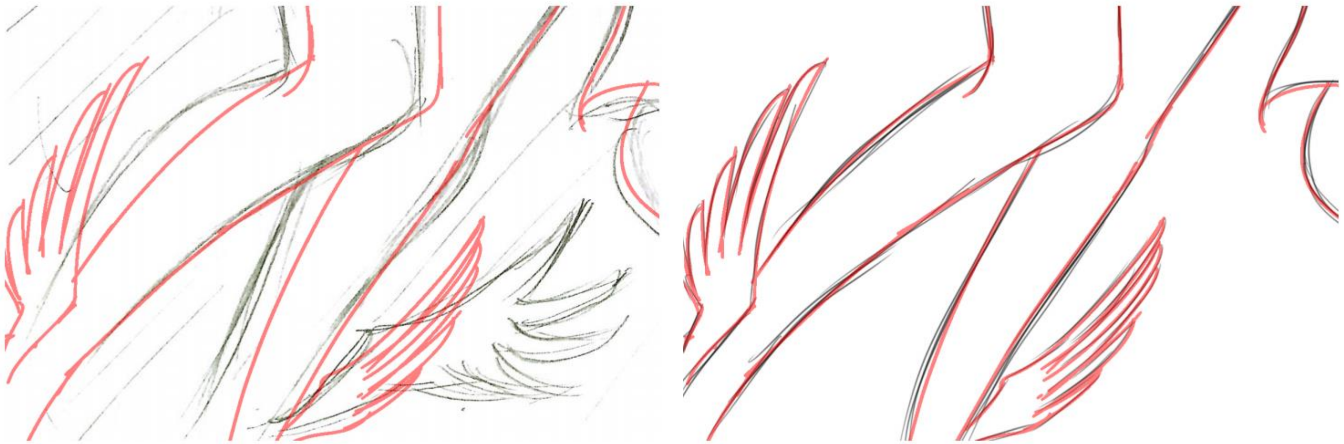
Sketch Simplification: *Learning to Simplify*



Learning to Simplify: Fully Convolutional Networks for Rough Sketch Cleanup, Simo-Serra et al.

Sketch Simplification: *Learning to Simplify*

- Loss for thin edges saturates easily
- Authors take extra steps to align input and ground truth edges



Pencil: input
Red: ground truth

Learning to Simplify: Fully Convolutional Networks for Rough Sketch Cleanup, **Simo-Serra et al.**

Image Decomposition

- A selection of methods:
- *Direct Instrinsics*, Narihira et al., 2015
- *Learning Data-driven Reflectance Priors for Intrinsic Image Decomposition,* Zhou et al., 2015
- *Decomposing Single Images for Layered Photo Retouching*, Innamorati et al. 2017

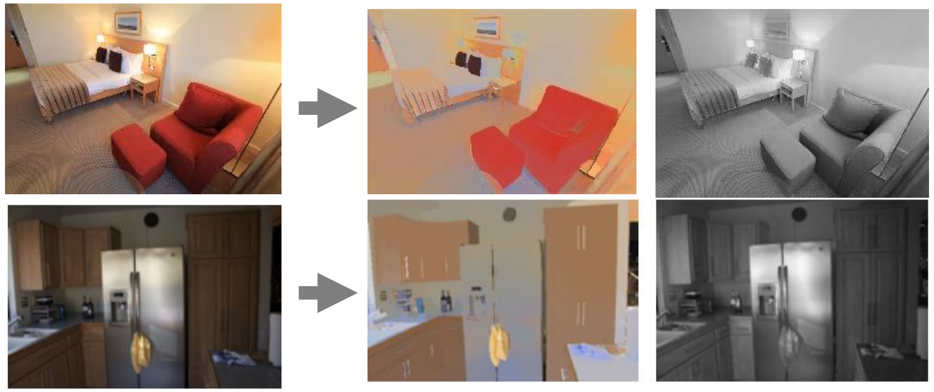
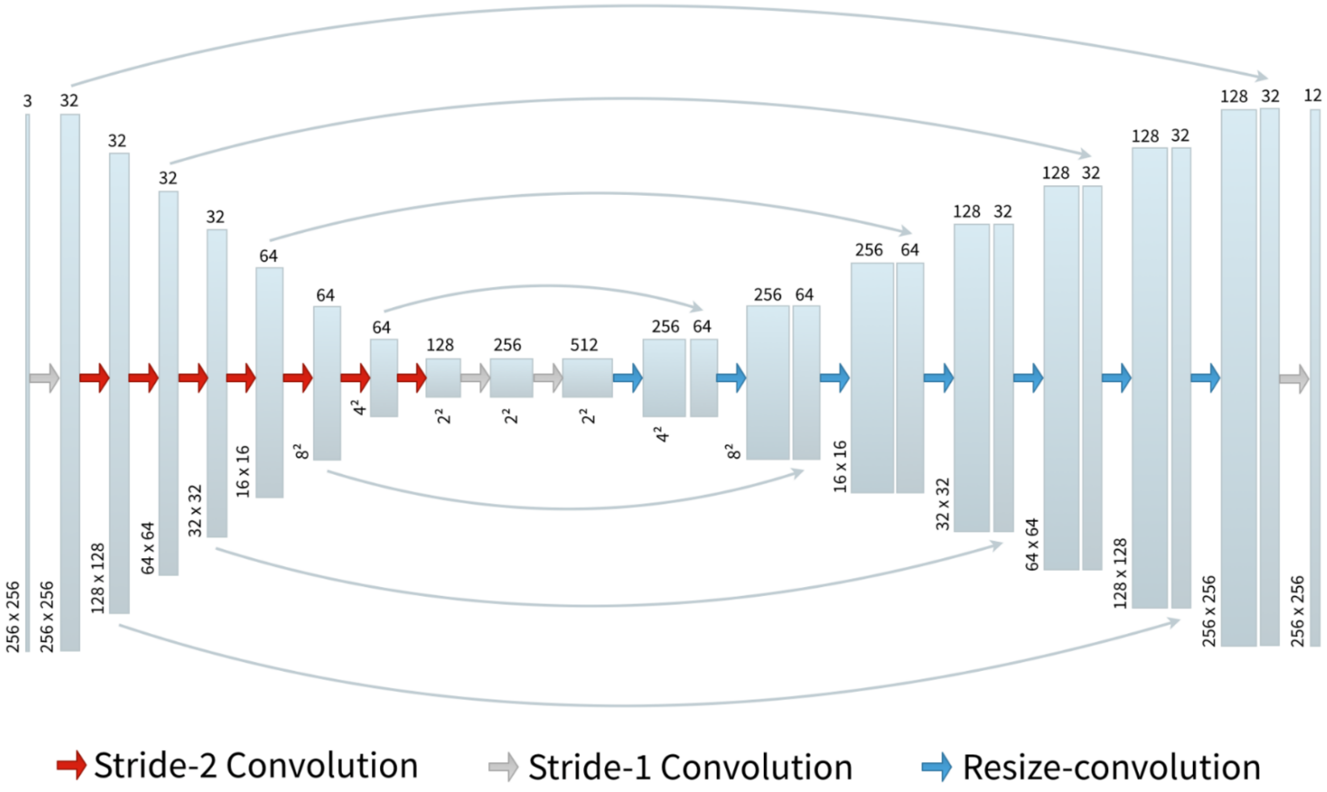
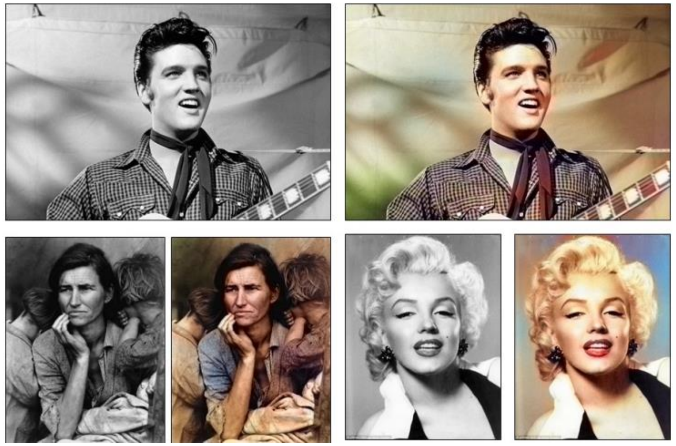


Image Decomposition: Decomposing *Single Images for Layered Photo Retouching*

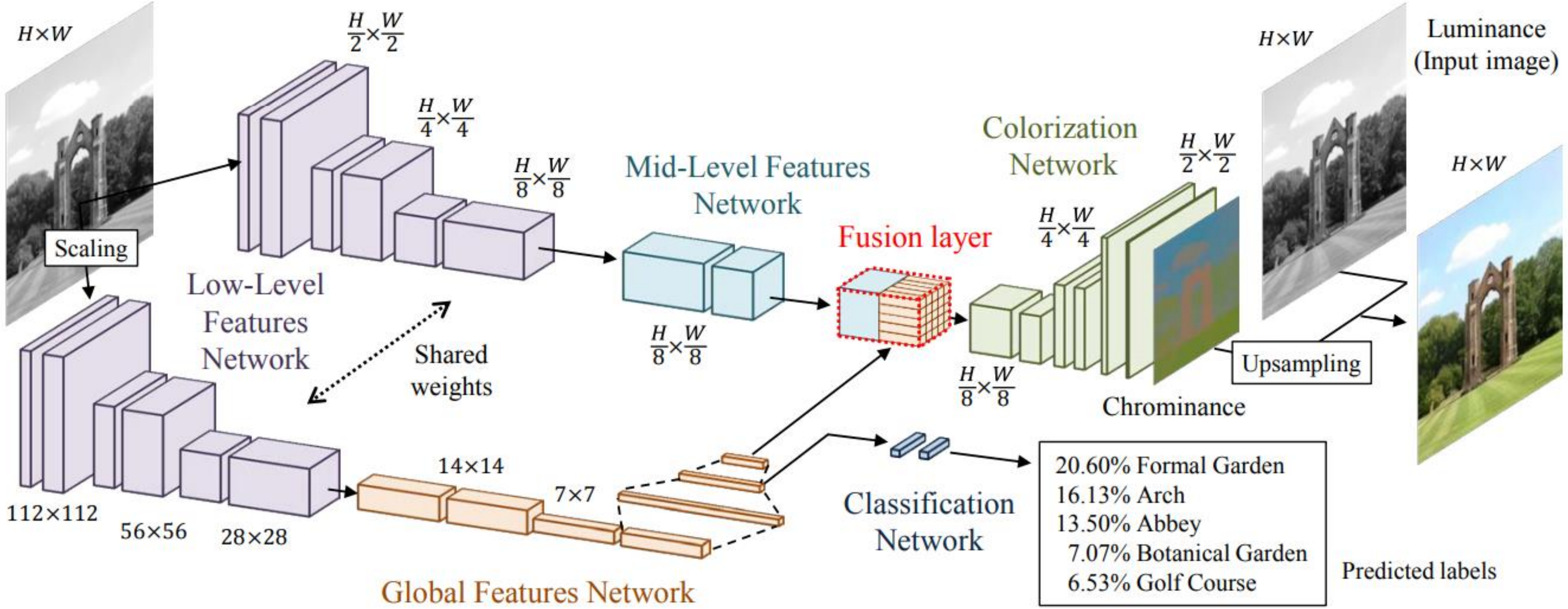


Colorization

- Concurrent methods:
  - *Let there be Color!*, Iizuka et al., 2016
  - *Colorful Image Colorization*, Zhang et al. 2016
  - *Learning Representations for Automatic Colorization,* Larsson et al., 2016
  - *Real-Time User-Guided Image Colorization with Learned Deep Priors*, Zhang et al. 2017

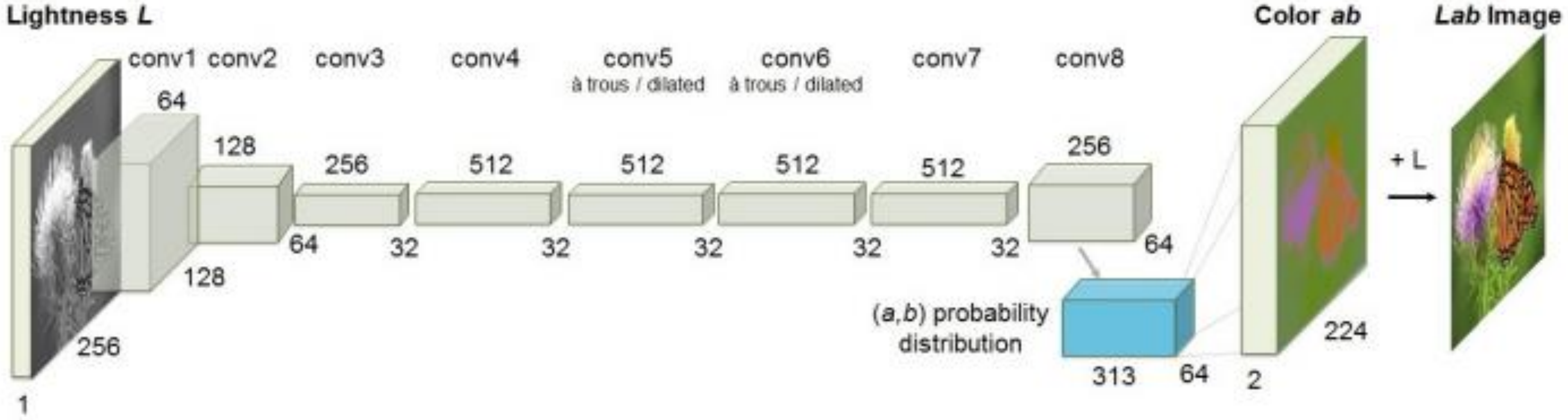


Colorization: *Let There Be Color!*



Let there be Color!: Iizuka et al.

Colorization: *Colorful Image Colorization*



input

output

direct regression probability distr.

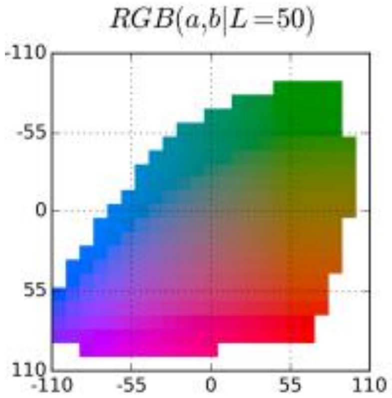
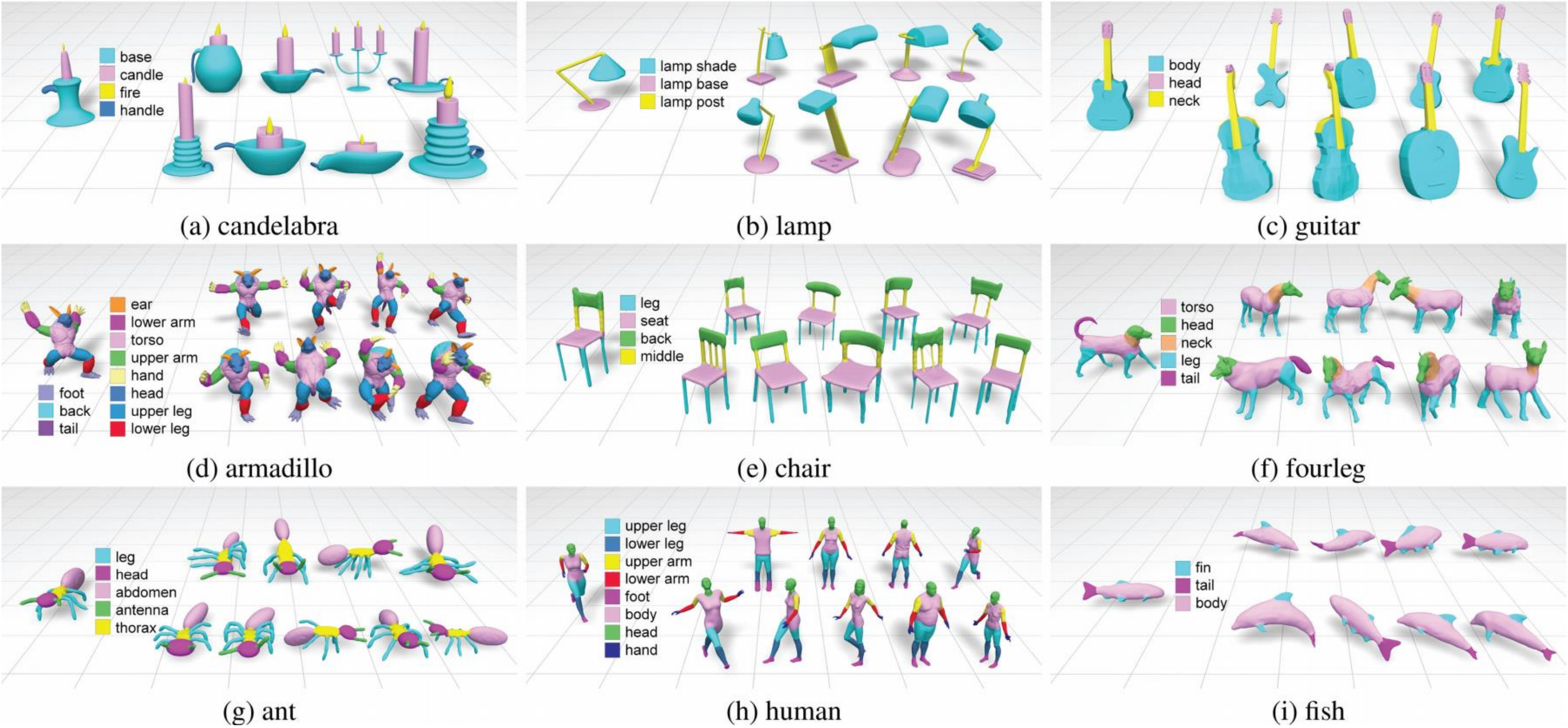


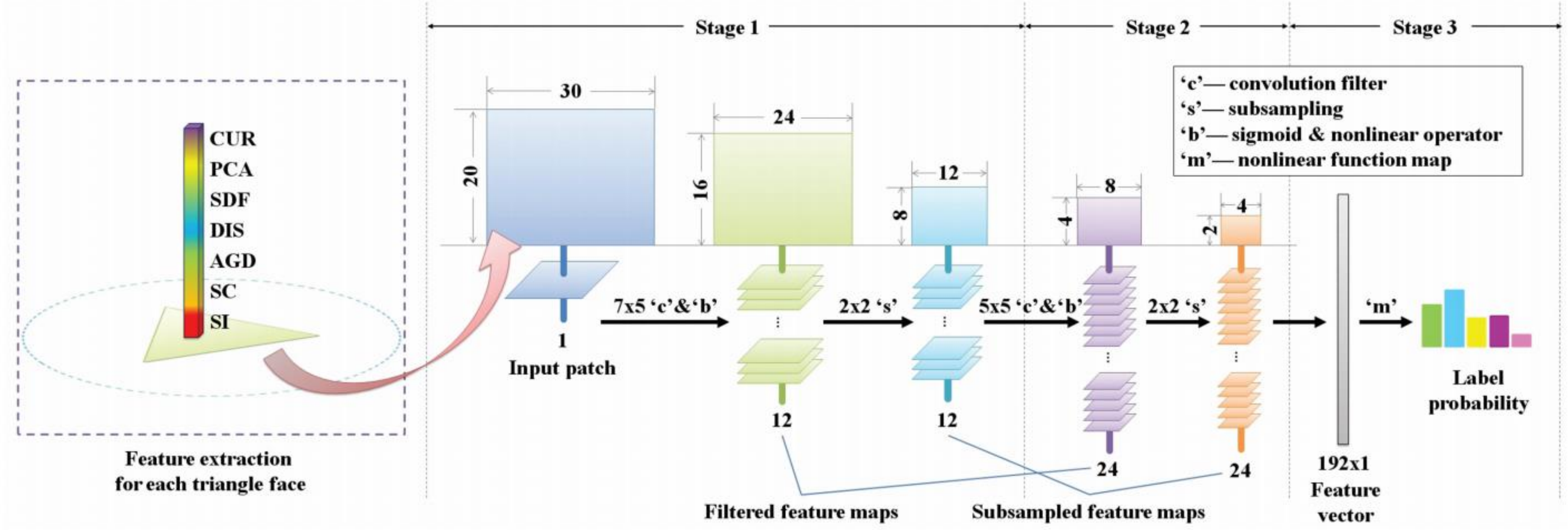
Image Credit: *Colorful Image Colorization*, Zhang et al.

Mesh Labeling / Segmentation

3D Mesh Labeling via Deep Convolutional Neural Networks, Guo et al. 2016



Mesh Labeling / Segmentation



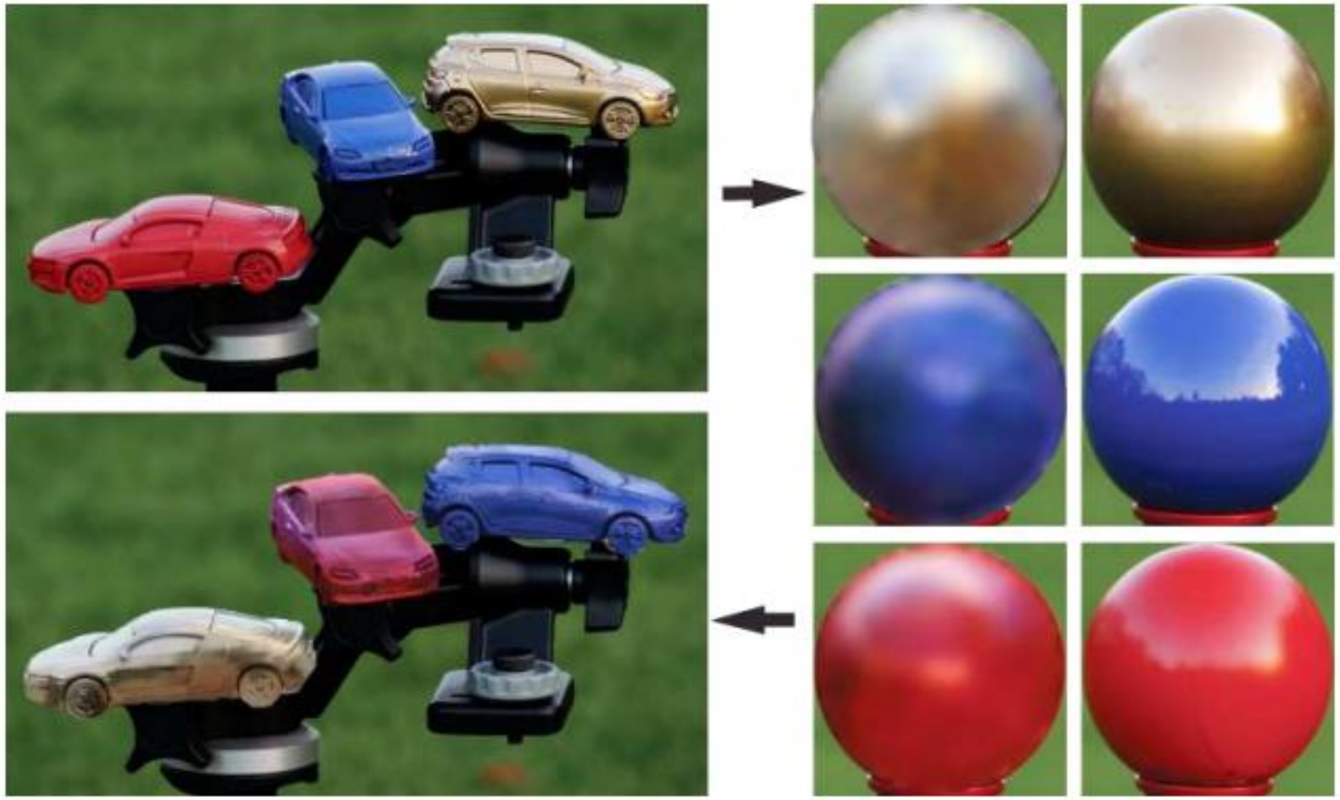
3D Mesh Labeling via Deep Convolutional Neural Networks, Guo et al.

LDR to HDR Image Reconstruction:

- Concurrently:
- *Deep Reverse Tone Mapping*, Endo et al. 2017
- *HDR image reconstruction from a single exposure using deep CNNs*, Eilertsen et al. 2017

Reflectance Maps

- Paint a sphere as if it is made of a material under a certain illumination of another object in a photo



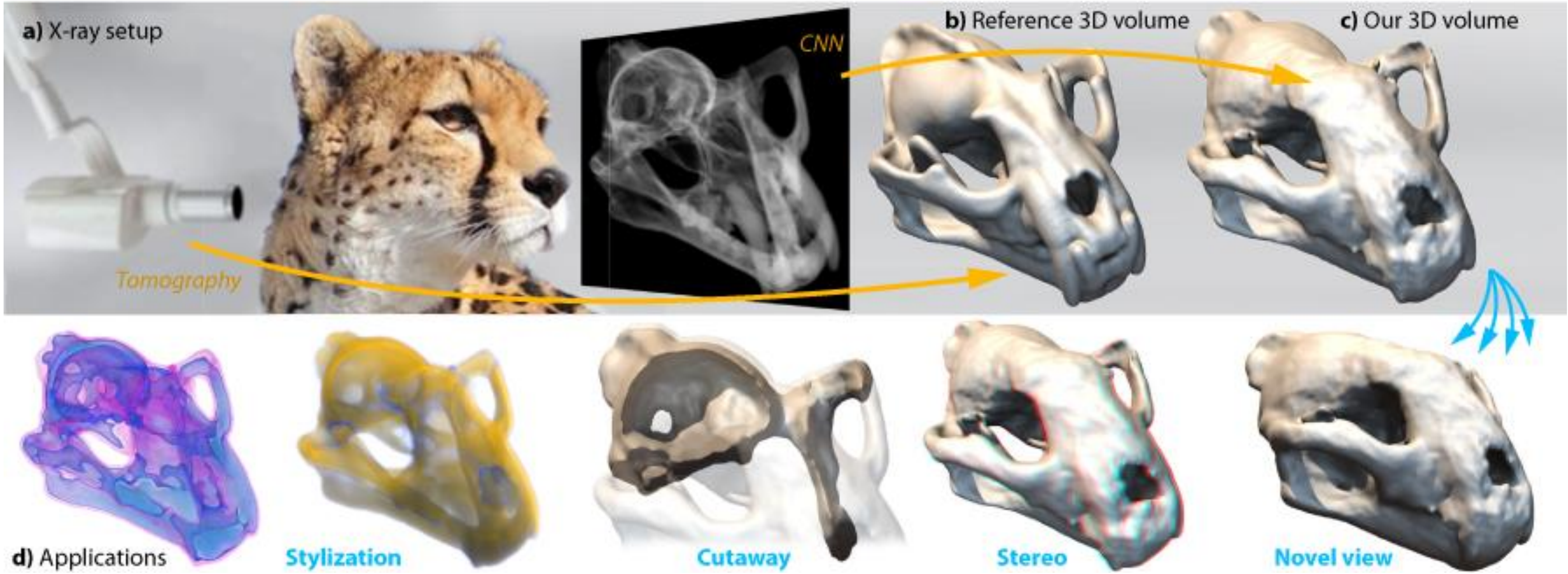
Deep Reflectance Maps. Rematas et al. **CVPR 2015**

DeLight

- Factor BRDF and (HDR) Illumination

Reflectance and Natural Illumination from Single-Material Specular Objects Using Deep Learning. Georgoulis et al. **PAMI 2017**

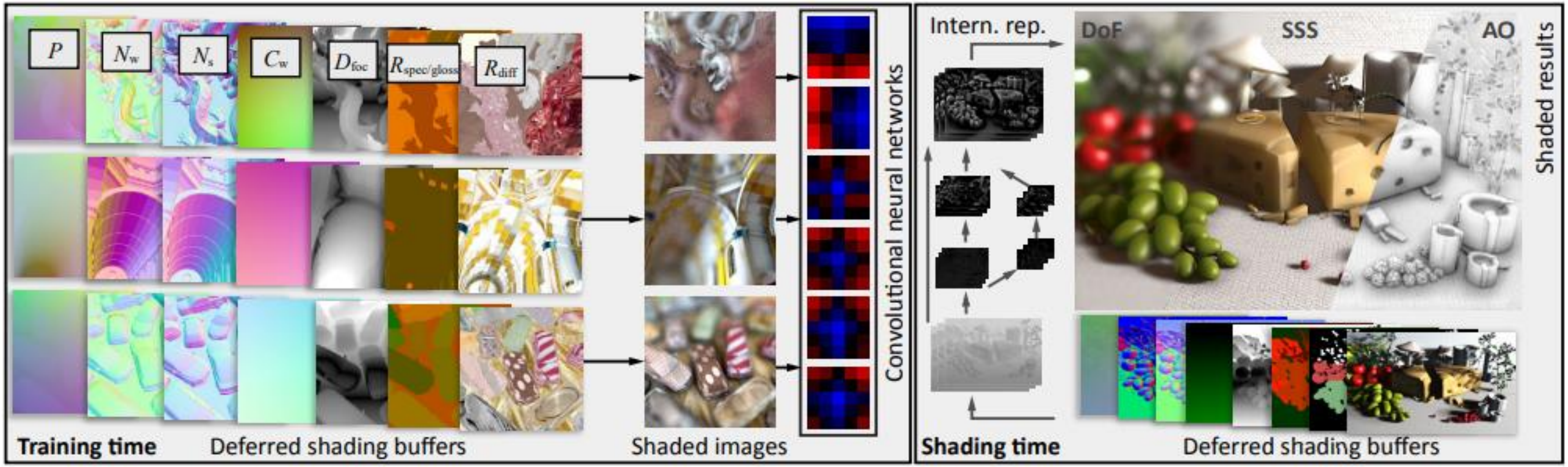
3D volumes form Xrays



Single-Image Tomography: 3D Volumes from 2D Cranial X-Rays. Henzler et al. **EG 2018**

Deep Shading

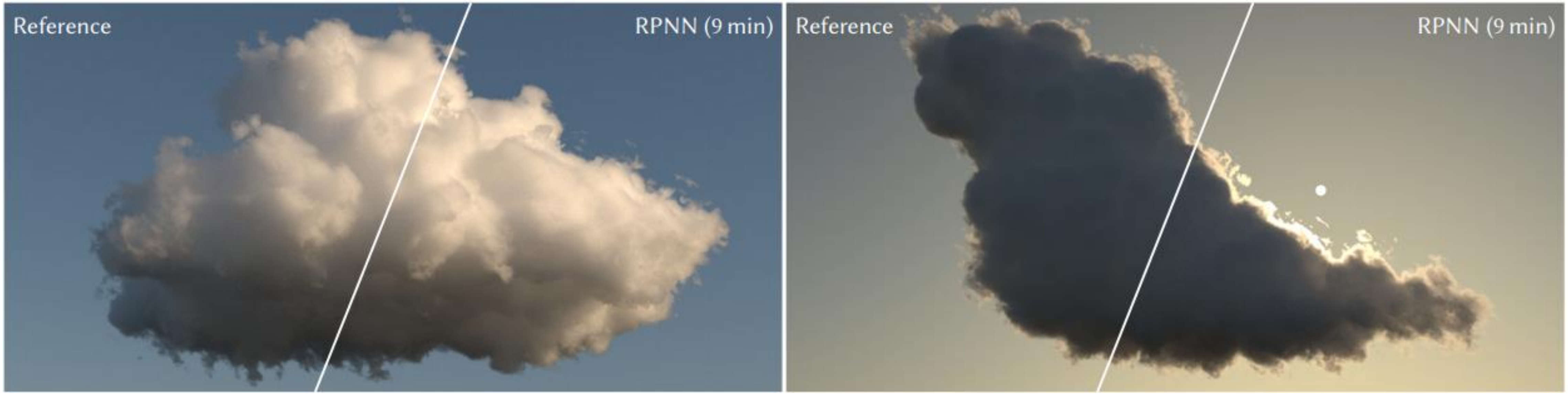
- Paint a z-buffer like a path tracer (AO, DOF, MB)



Deep Shading, Nalbach et al. EGSR 2017

Rendering Atmospherics

Deep Scattering: Rendering Atmospheric Clouds with Radiance-Predicting Neural Networks, Kallweit et al. **SIGGRAPH Asia 2017**

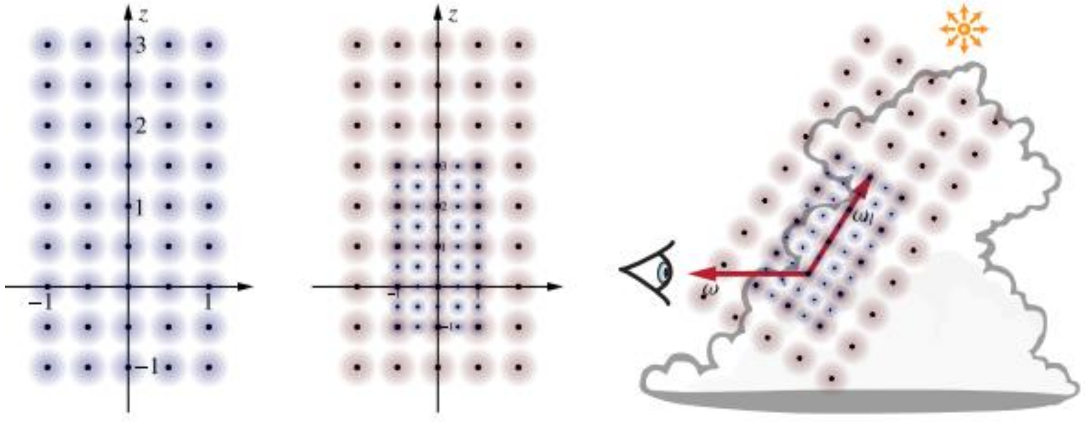


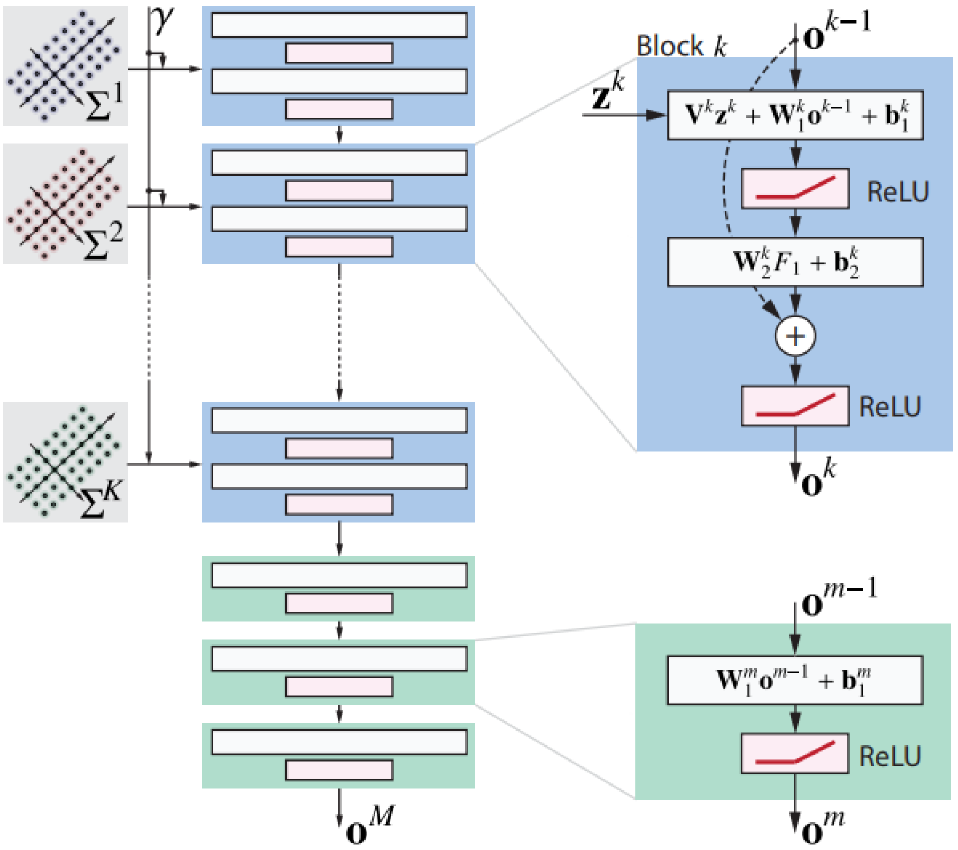
Speed up approx. 24 x

Speed up approx. 24 x

Rendering Atmospherics: *RPNN*

In: Hierarchical representation of a cloud patch
Out: incoming indirect radiance at patch center
(incoming direct radiance is computed directly)

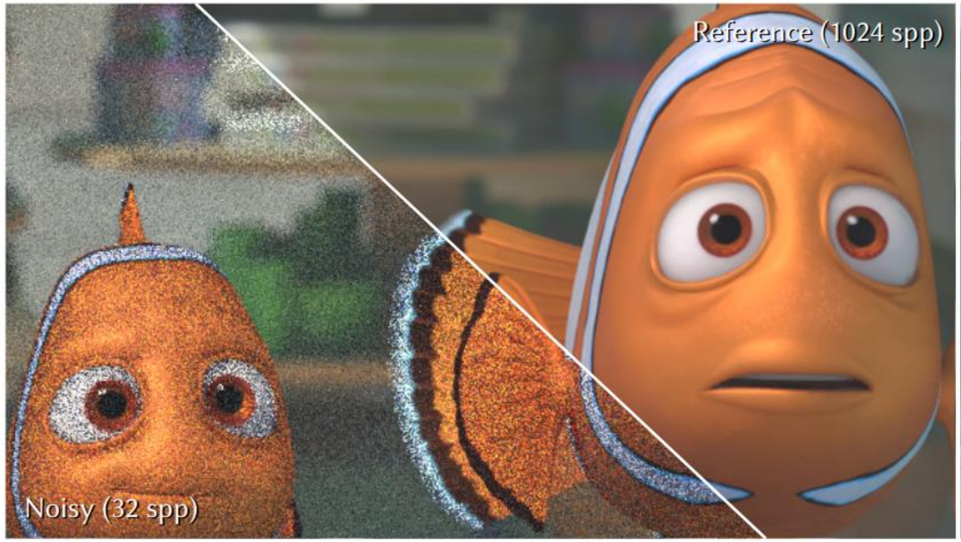




Deep Scattering: Rendering Atmospheric Clouds with Radiance-Predicting Neural Networks, Kallweit et al. **SIGGRAPH Asia 2017**

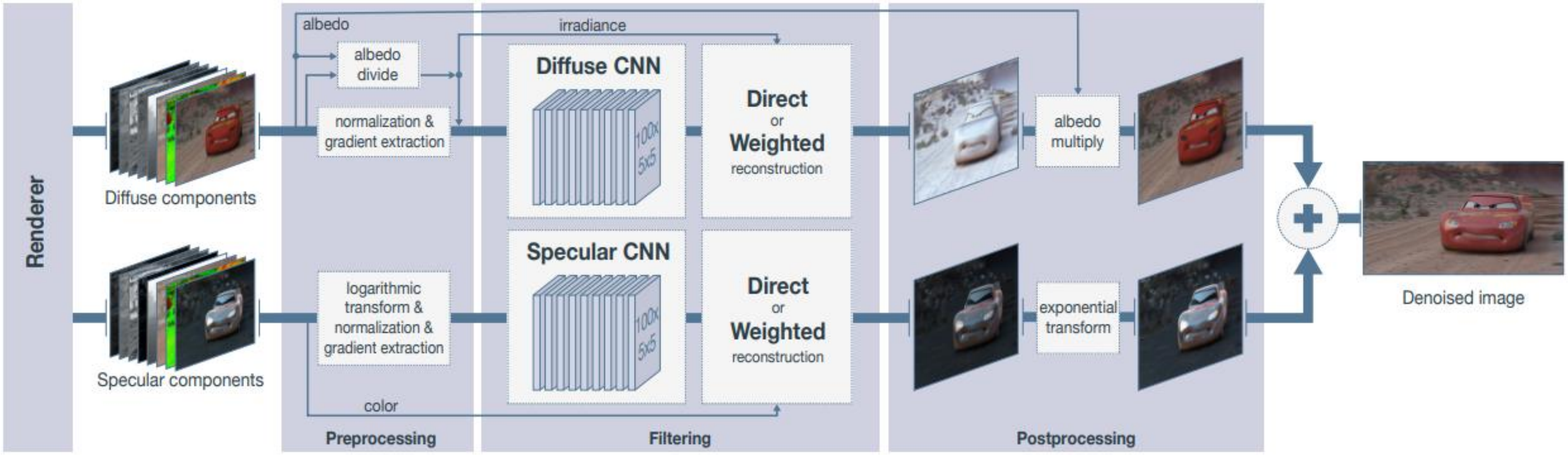
Denoising Renderings

- Concurrent:
- *Kernel-Predicting Convolutional Networks for Denoising Monte Carlo Renderings,* Bako et al. 2017
- *Interactive Reconstruction of Monte Carlo Image Sequences using a Recurrent Denoising Autoencoder,* Chaitanya et al. 2017 (more on Autoencoders later)



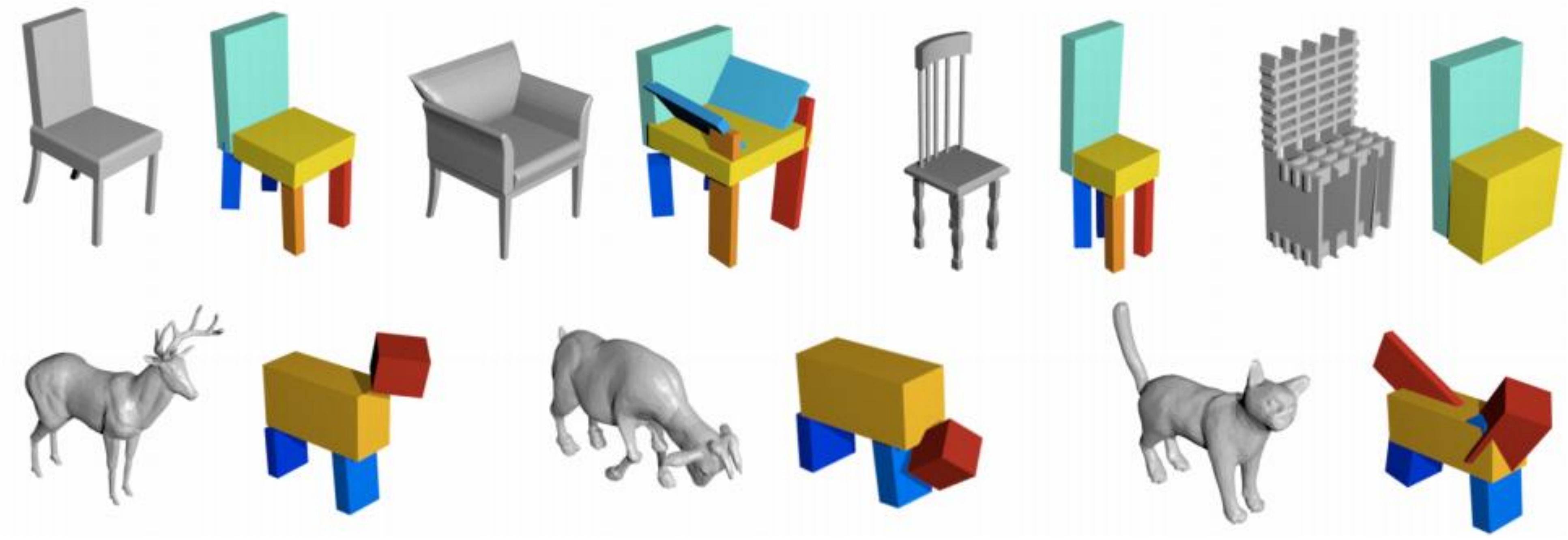
Kernel-Predicting Convolutional Networks for Denoising Monte Carlo Renderings, **Bako et al.**

Denoising Renderings:



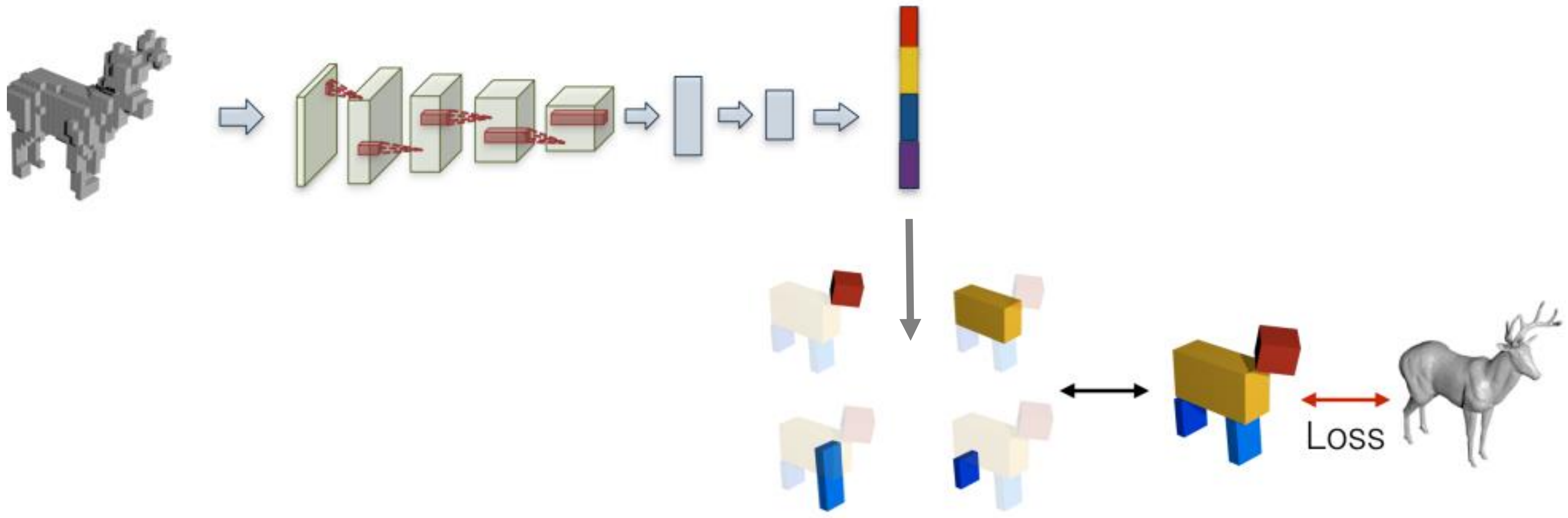
Kernel-Predicting Convolutional Networks for Denoising Monte Carlo Renderings, Bako et al. **SIGGRAPH 2017**

Geometry Abstraction / Simplification



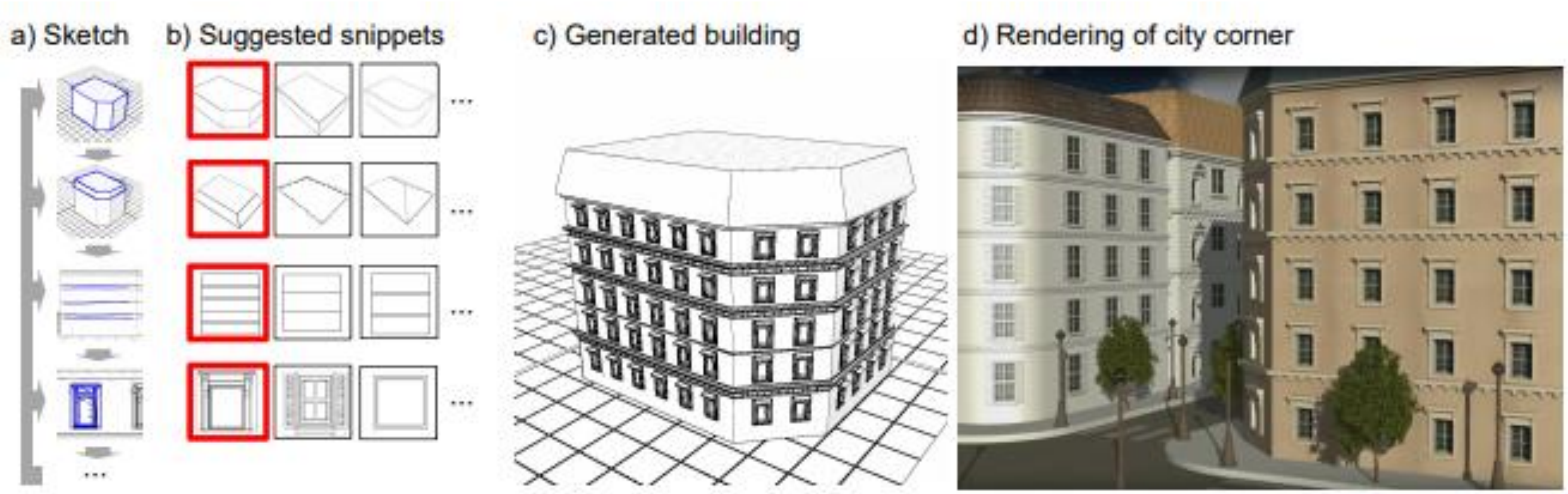
Learning Shape Abstractions by Assembling Volumetric Primitives, Tulsiani et al. 2016

Geometry Abstraction / Simplification:



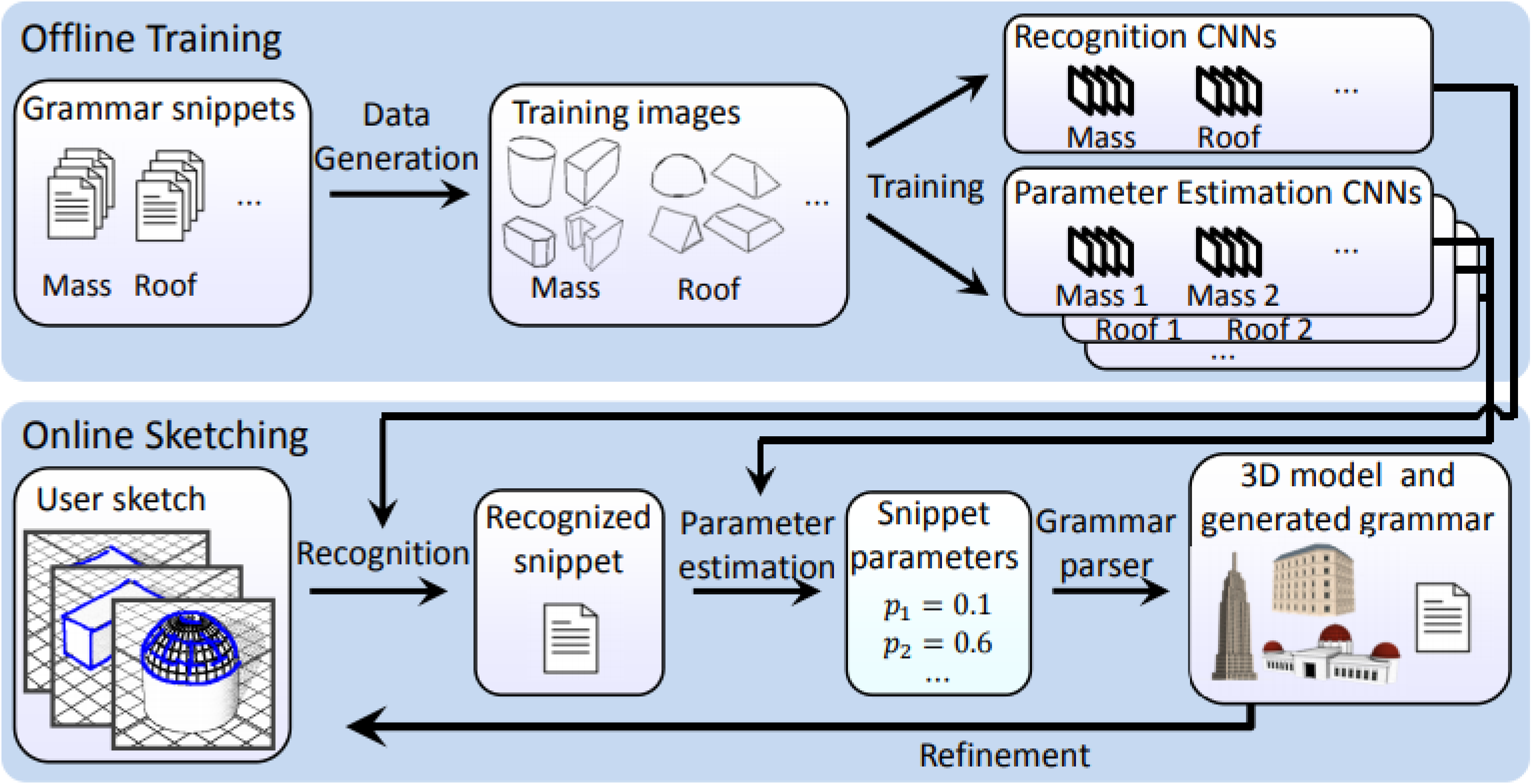
Learning Shape Abstractions by Assembling Volumetric Primitives, Tulsiani et al. 2016

Procedural Parameter Estimation



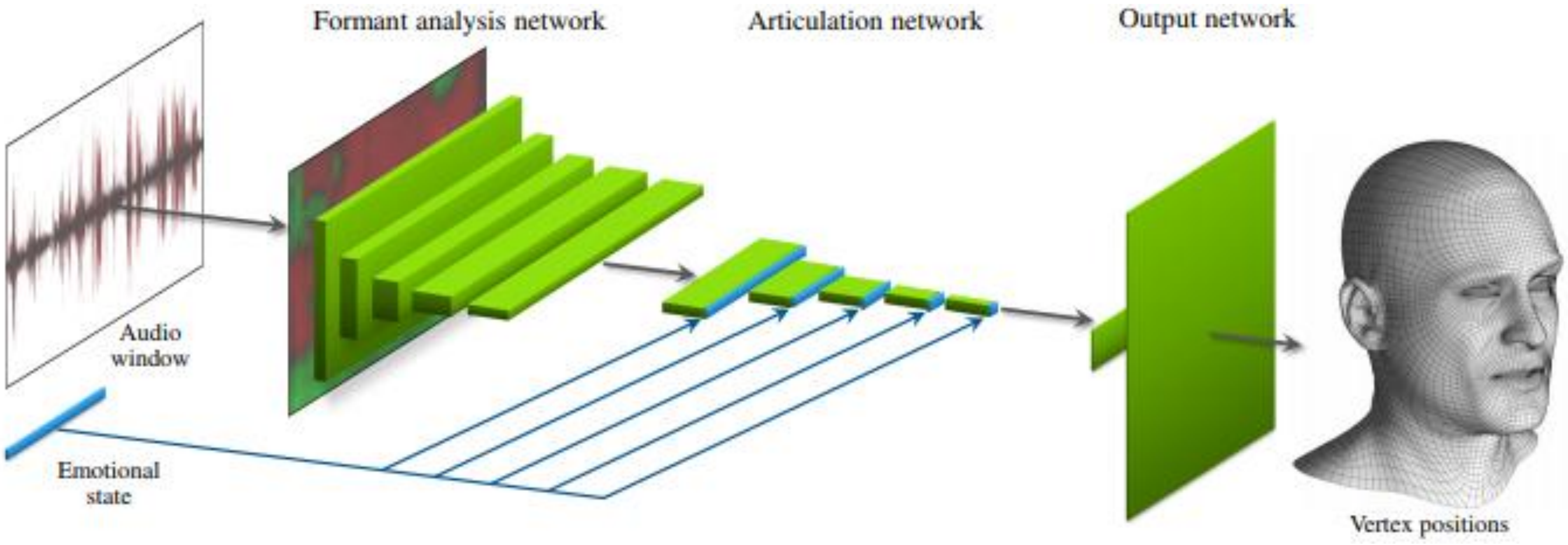
Interactive Sketching of Urban Procedural Models, Nishida et al. 2016

Procedural Parameter Estimation: *Interactive Sketching of Urban Procedural Models*



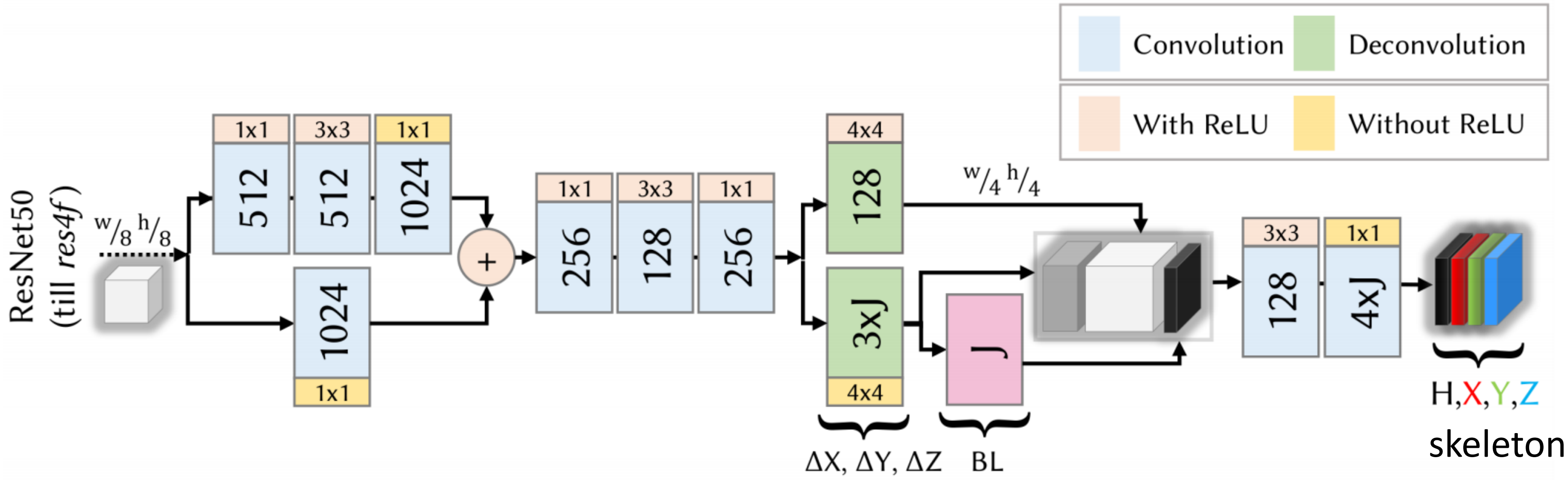
Interactive Sketching of Urban Procedural Models, **Nishida et al.**

Audio-driven facial animation



Audio-Driven Facial Animation by Joint End-to-End Learning of Pose and Emotion, Karras et al. 2017

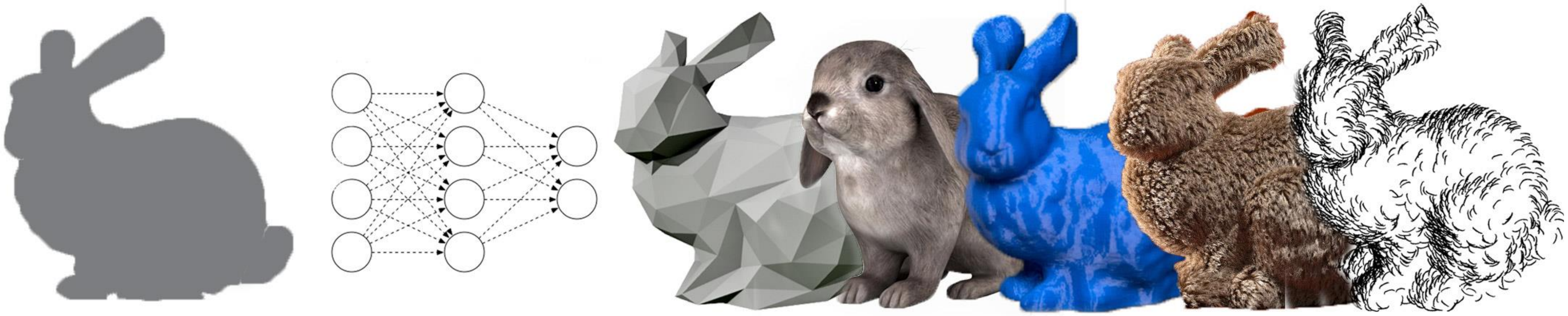
3D Pose Estimation: *VNECT*



skeleton joint heatmap
and 3d positions

VNect: Real-time 3D Human Pose Estimation with a Single RGB Camera, Mehta et al., **SIGGRAPH 2017**

Thank you!



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