

# Robust Vehicle Parsing Using Synthetic Data

## September Update

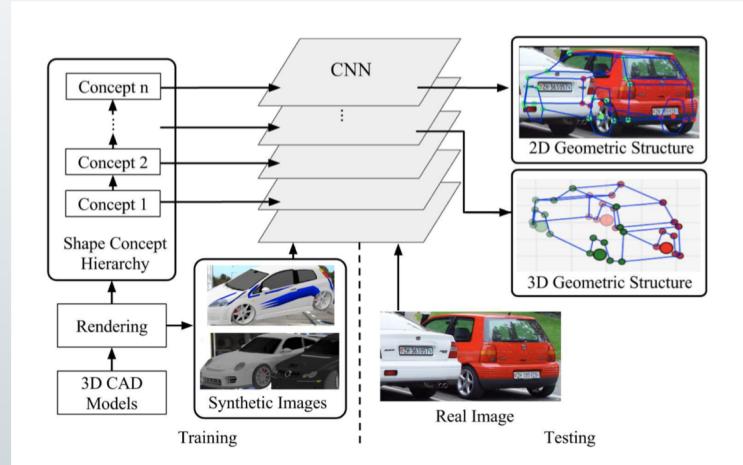
Thomas Liao

09/05/2018

## Motivation for predicting car key points using synthetic data

- Car- human activities: provide relative geometry between vehicles and humans for complex interactions for car-human activities.
- Car tracking: provide tools for similarity metric and trajectory prediction for car tracking.

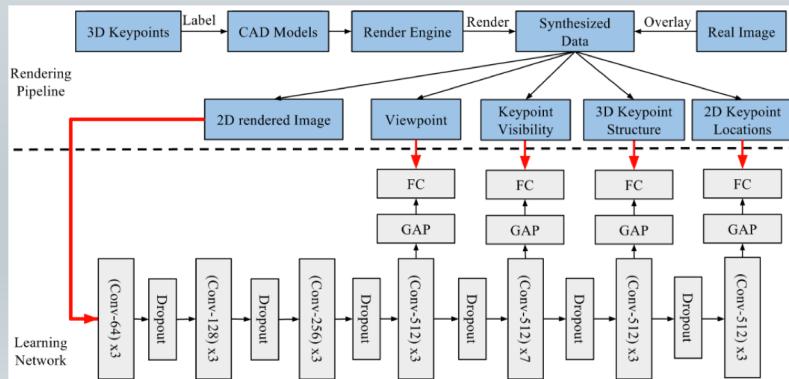
## Previous Model: Overview and Failure Cases



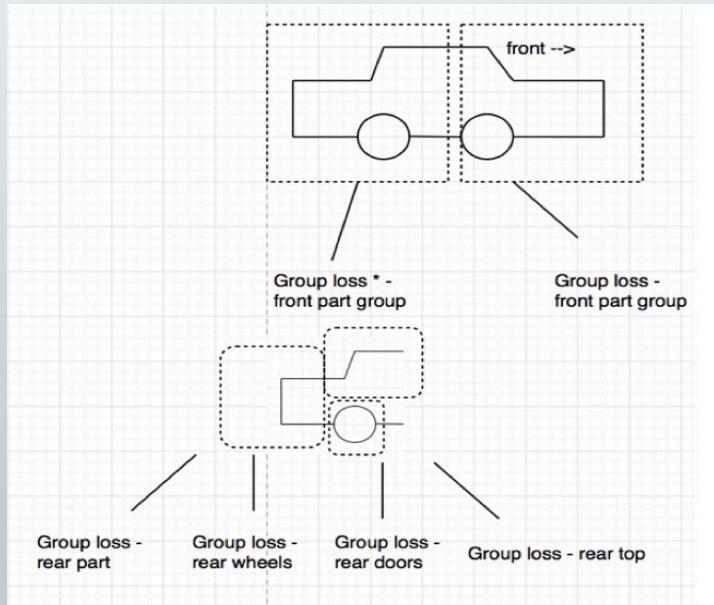
Failure case 1: occlusion/truncation



Failure case 2: deformable parts



# Improvement – hierarchical structure-aware loss function



Insights:

1. local context info for key point inference (especially in heavily-occluded situation).
2. robust deformable parts (doors, trunks) key point inference.

Hierarchical structure-aware loss function:

$$L_{hs} = \frac{1}{N} \sum_{i=1}^N \|P_i - G_i\|_2 + \sum_{k=1}^L \lambda_k \sum_{i=1}^{N_G^k} \|P_{S_i^k} - G_{S_i^k}\|_2$$

$P_i, G_i$ : prediction heat map and ground truth heat map of  $i^{th}$  key point ( $i = 1, 2, \dots, N$ )  
 $P_{S_i^k}, G_{S_i^k}$ : group prediction heat map, ground ground truth heat map of group  $S_i^k$   
where  $S_i^k$  is  $i^{th}$  group within  $k^{th}$  hierarchical level.  
 $N_G^k$  is number of groups in  $k^{th}$  hierarchical level

# Improved Car Model: Deformable Cars

● Ground truth

● Prediction

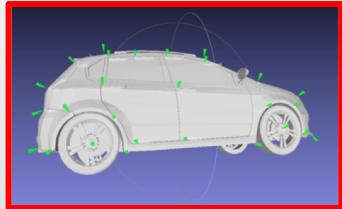
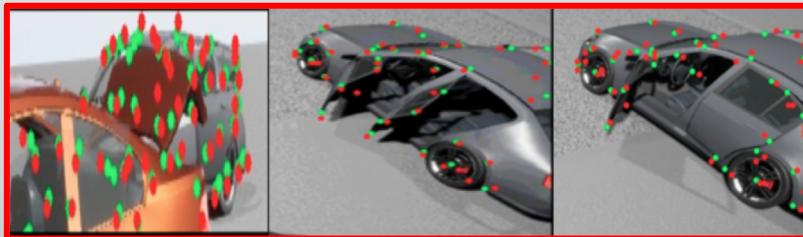


Fig. 3D object model for rendering engine (UE4.) data synthesis.

Full view (no occlusion / truncation)



Moderate occlusion/truncation



Heavy occlusion/truncation



Fig. Visualization on deformable cars w.r.t. occlusion level.

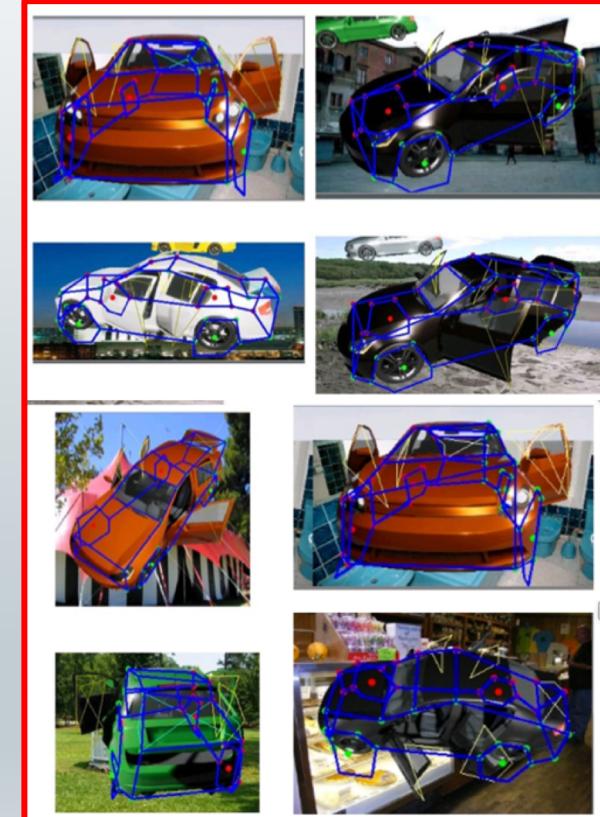


Fig. Deformable car visualization.

# Deep supervised multi-resolution with coarse-to-fine stepwise probabilistic framework for occlusion-aware / occlusion robust neural network.

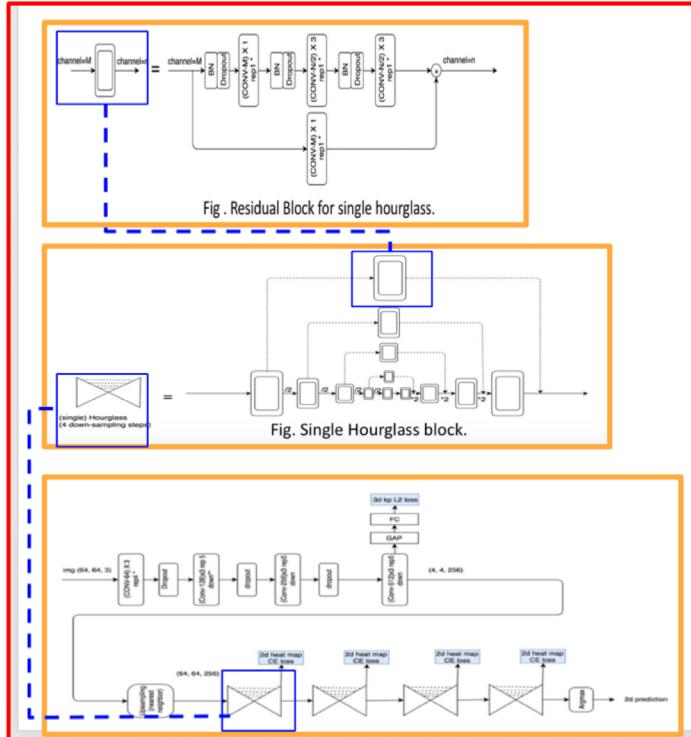


Fig 1. hg\_23d model

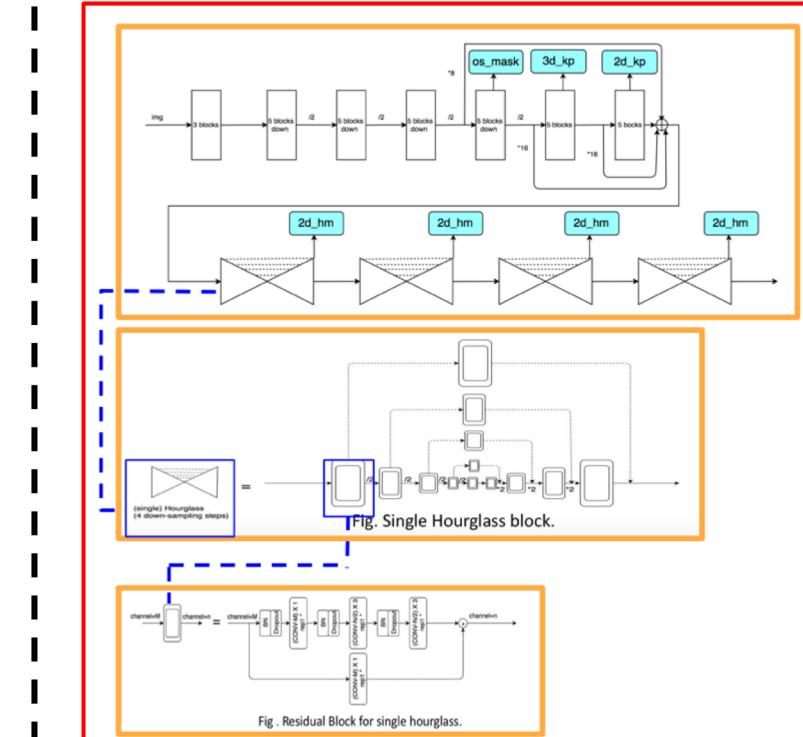


Fig 2. hg\_fusion\_os23d model

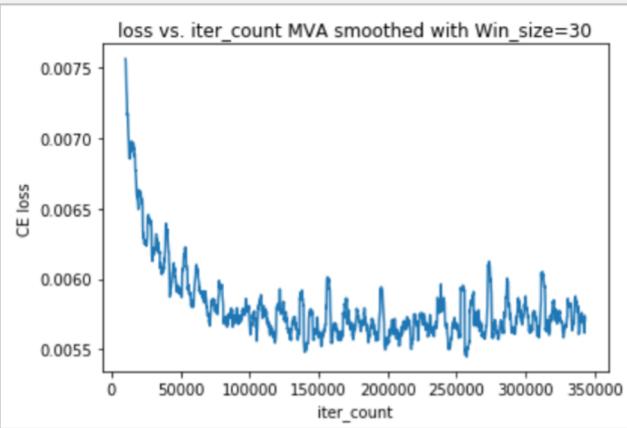


Fig. Combined training losses vs training step.

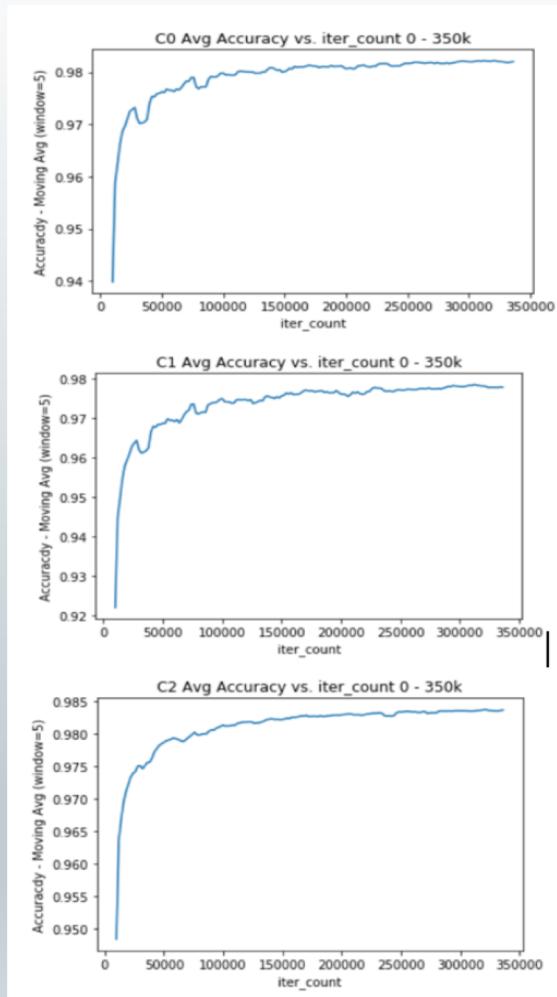


Fig. Full view / truncation / multi car occlusion accuracies vs. training step.

C0: full view

C1: cropped image

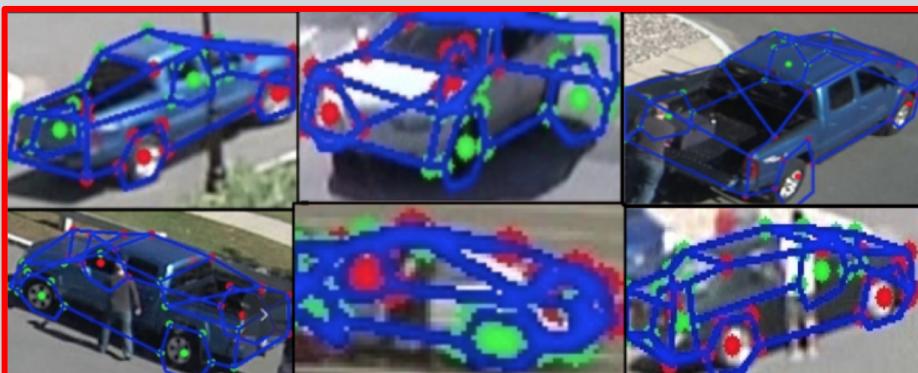
C2: multi car

# Improved performance on occlusion

Before



After



Model	full *	truncate *	multi-car *	occ *
L2_23d	90.6	44.4	71.9	70.7
L2_os23d	89.0	39.6	70.2	74.6
hg_23d	88.3	<b>48.7</b>	72.2	<b>78.0</b>
hg_os23d	89.4	<b>49.6</b>	72.8	<b>80.2</b>

Table 1. PCK[ $\alpha = 0.1$ ] accuracies (%) for 2D key points localization on KITTI-3D dataset.

\* full: car fully visible, 788 images.

\* truncate: car with truncation, 436 images.

\* multi-car: multiple car, 696 images.

\* occ: car occluded by other objects, 120 images.

Fig. Qualitative comparison for hg\_os23d performance with origin model.