FIRE: THE FIRST-YEAR INNOVATION & RESEARCH EXPERIENCE

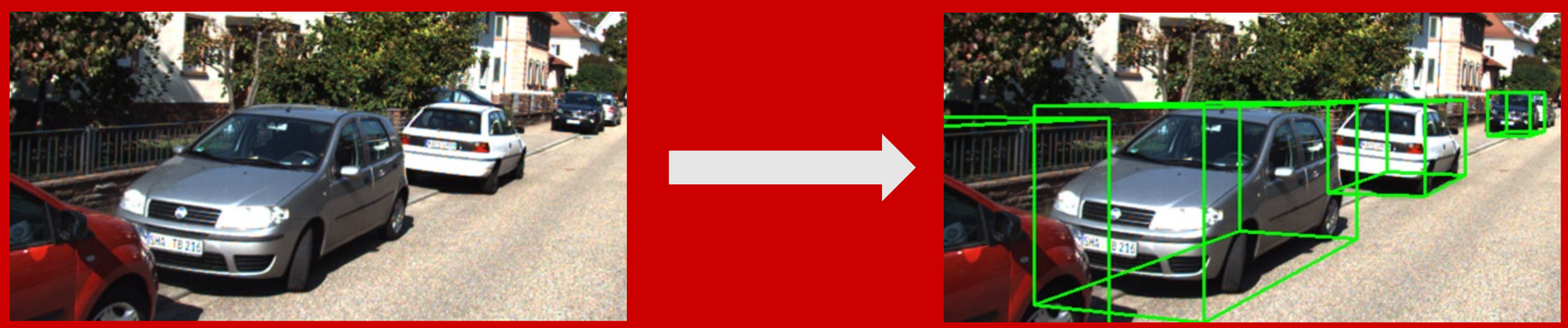
CAPITAL ONE MACHINE LEARNING

Monocular 3D Object Detection

Jerry Lan, Richard Gao, Siyuan Peng, Vladimir Leung

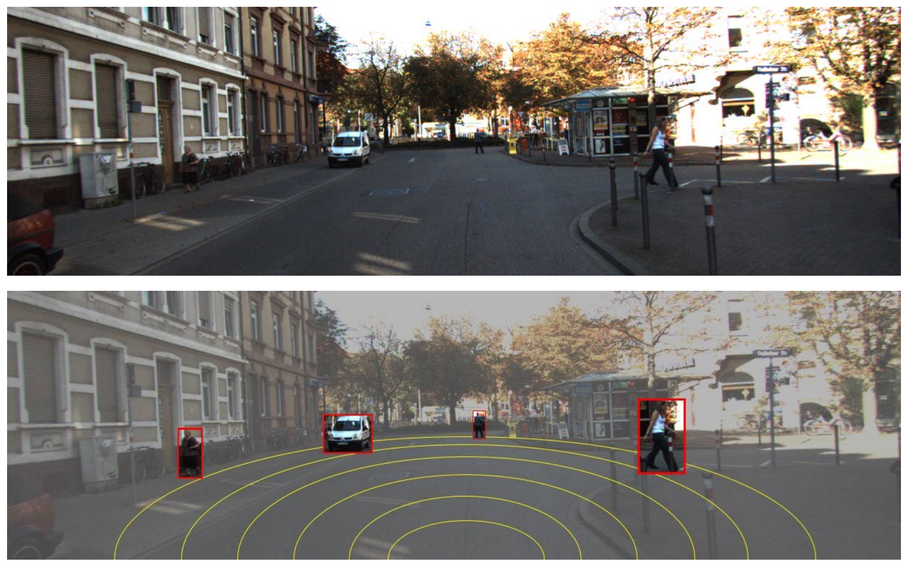
3D Object Detection From A Single Camera

Training computers to understand 3D space



Future Application

Method Logic

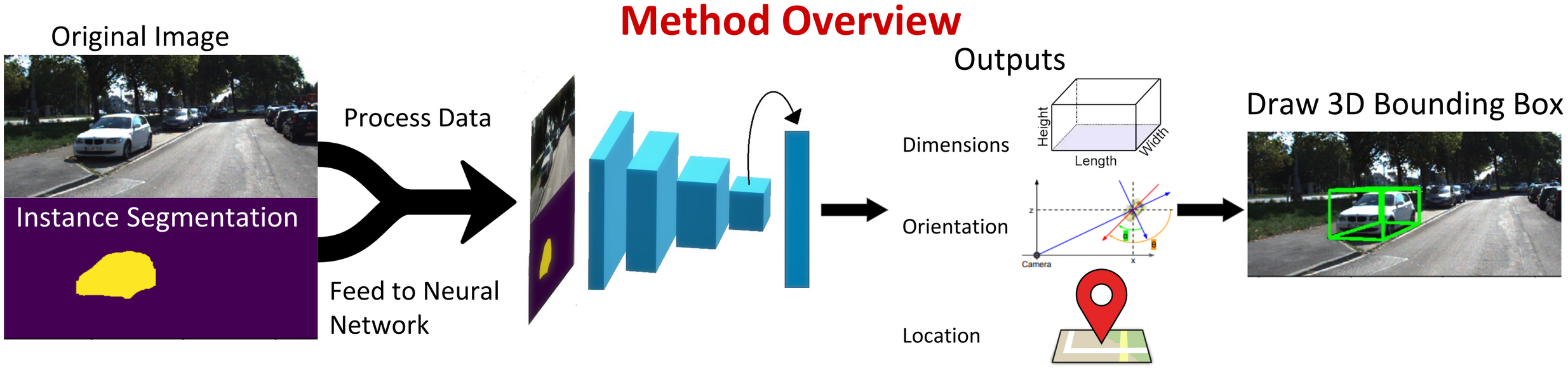


Step 1:
Look at the image!

Step 2:
Identify objects

Step 3:
Determine distance & orientation

Method Overview



Dataset

- KITTI provides over 7500 full-size images.
- 40,000 labeled instances of objects.
- Ground truths are measured by laser scanner and GPS system.

Challenges

- Limited data.
- Long period of training time.
- Performance balance between difference outputs (dimension, orientation, location).

Contributions

- Classification Formulation: predict orientation of the object based on the sector rather than the actual angle, making it easier to train.
- Modular Network: each segment of our network can be trained and used independently, so that poorly performing segments can be replaced without issue.

Next Steps

- Data augmentation to simulate additional training data
- Refine the loss function