Lecture 11: Detection, Segmentation, and Feature Extraction in Computer Vision

Areej Alasiry

CIS 6217 – Computer Vision for Data Representation

College of Computer Science, King Khalid University

Outline

1. Semantic Segmentation
2. Object Detection
3. Instance Segmentation
4. HoG/SIFT Features

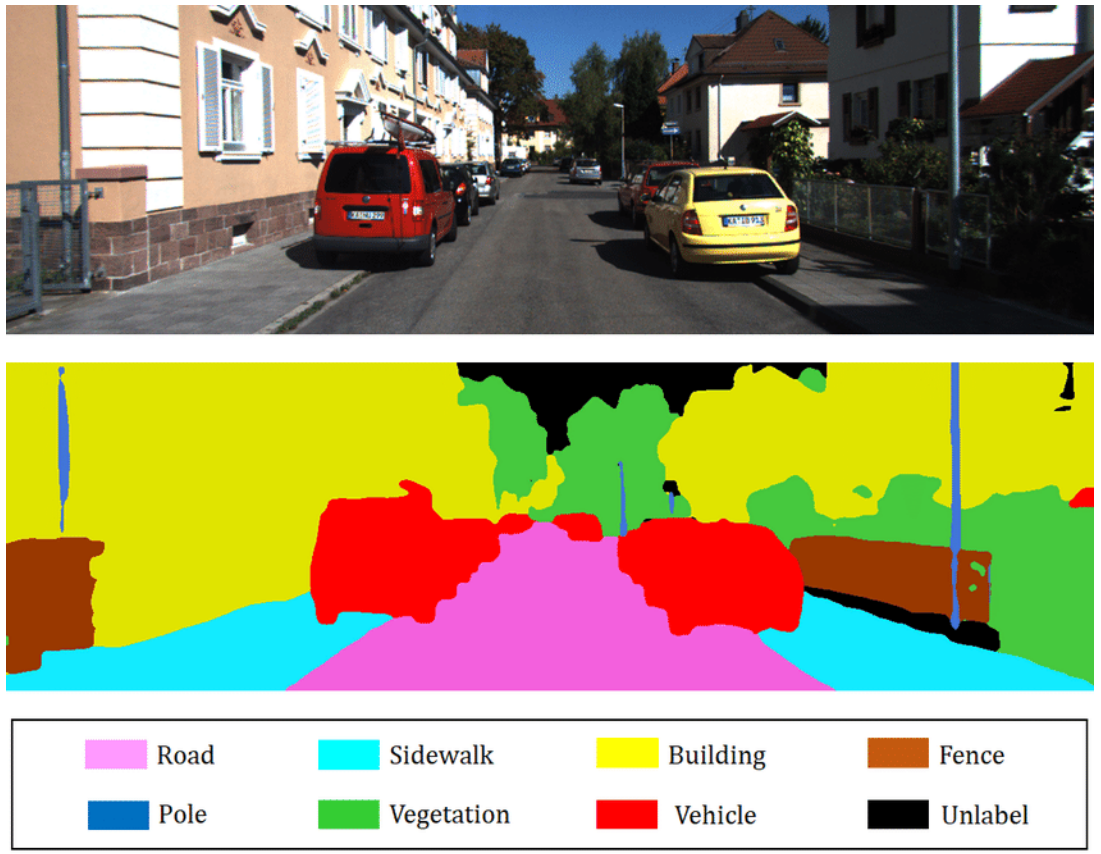
Why Detection & Segmentation?

- Essential for:
  - Autonomous driving
  - Medical imaging
  - Robotics, tracking
  - Surveillance
  - Scene understanding

Semantic Segmentation

- Assigns a **class label to every pixel** in the image.
  All objects of the same class share the same label.

- Common Architecture:
  - **Fully Convolutional Networks (FCN)**
  - **U-Net** (skip connections)
  - **SegNet**
  - **DeepLab (v1–v3+)**

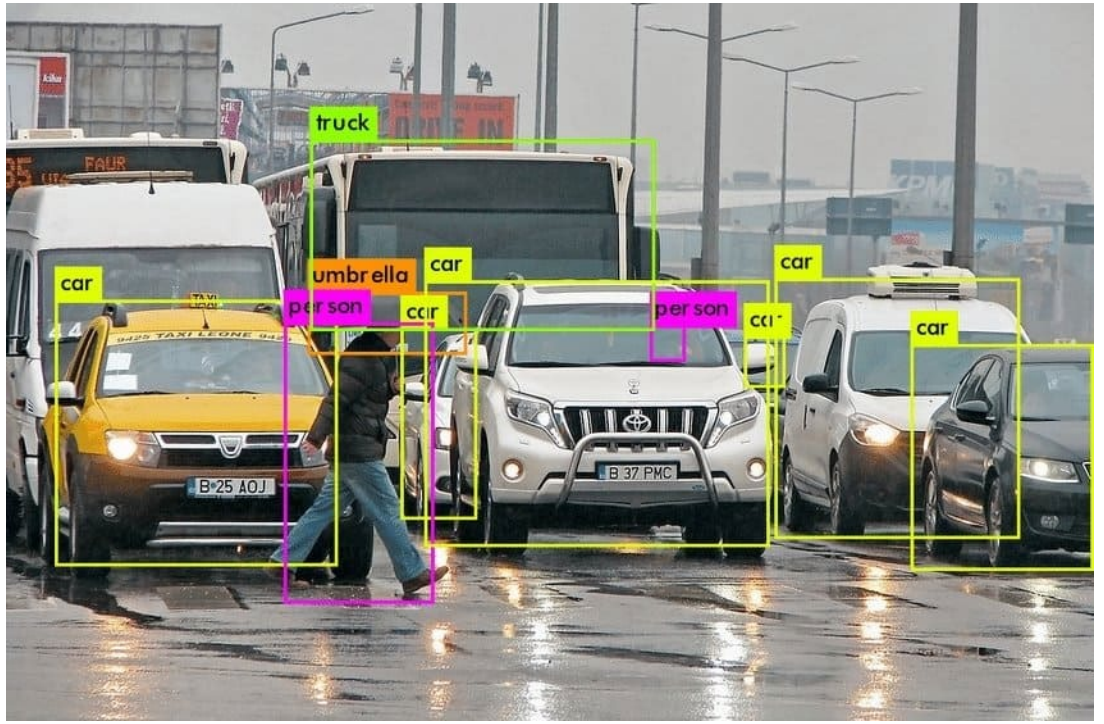


Team, T. A. (2021, October 6). *Semantic segmentation: A complete guide*. https://towardsai.net/p/l/machine-learning-7

Object Detection

- Locate and classify **each object instance** with a bounding box.

- Two components:
  - **Classification** (what is the object?)
  - **Localization** (where is it?)



Klingler, N. (2025, April 4). *Explore Yolov4: Speedy object detection mastery*. viso.ai. https://viso.ai/deep-learning/yolov4/

Object Detection Pipline

1. Two-Stage Detectors

- R-CNN, Fast R-CNN, Faster R-CNN
- Region proposal → classification + bounding box regression
- High accuracy

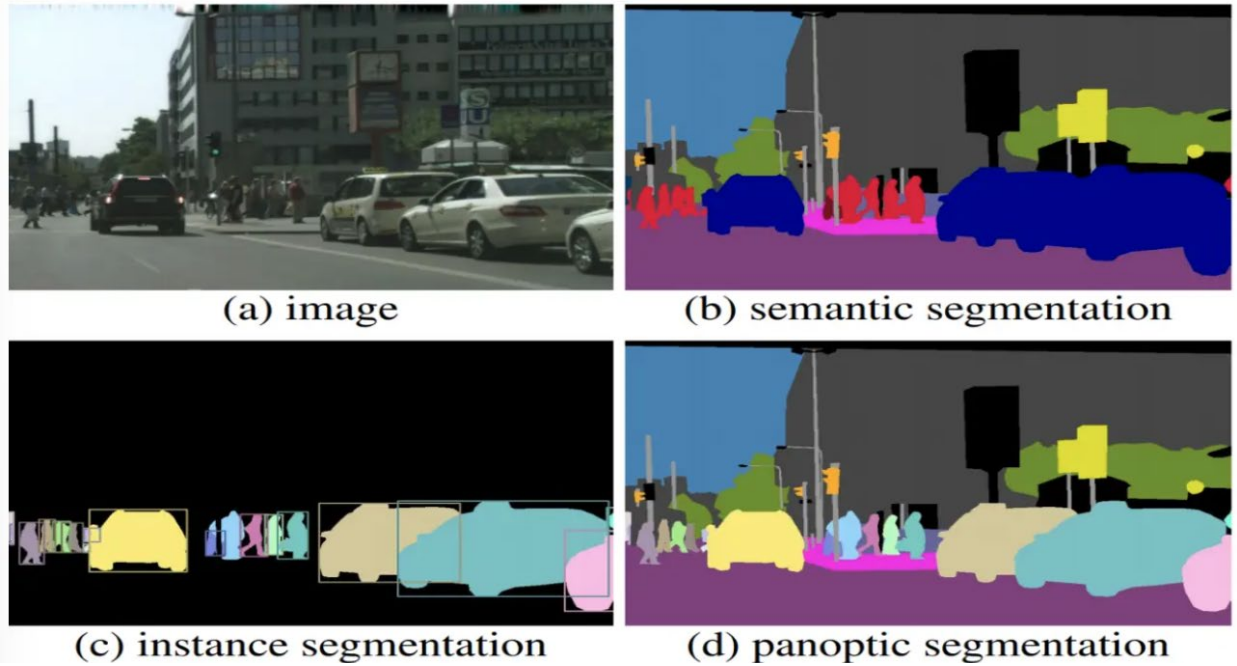
2. One-Stage Detectors

- YOLO (You Only Look Once)
- SSD (Single Shot Detector)
- RetinaNet
- No proposal stage → fast, real-time detection

- Output: For each object:
  - Bounding box (x, y, width, height)
  - Class label
  - Confidence score

Instance Segmentation

- Pixel-level segmentation for each object instance separately
- Combines object detection and semantic segmentation
- It first detects each object using a bounding box, and then generates a pixel-wise mask for each detected object within that box.



Singh, S. (2025, August 28). *Instance segmentation vs Semantic vs Panoptic: Labellerr*. Labellerr AI. https://www.labellerr.com/blog/semantic-vs-instance-vs-panoptic-which-image-segmentation-technique-to-choose/

Feature Extraction in Classical CV

- **Histogram of Oriented Gradients (HoG)**
- Extracts **edge orientation** and **shape** information
- Steps:
  - Compute image gradients
  - Form orientation histograms in small cells
  - Normalize over blocks
  - Concatenate into a feature vector
- **Strengths:**
  - Good for detecting shapes (e.g., pedestrians)
  - Robust to illumination & contrast variations

- **Scale-Invariant Feature Transform (SIFT)**
- Detects **stable keypoints** & describes them
- Steps:
  - Detect scale-space extrema
  - Localize keypoints
  - Assign

Classical vs Deep Features

Aspect	Classical (HoG/SIFT)	Deep Features (CNN)
Feature Design	Manual	Learned from data
Robustness	Moderate	Very high
Performance	Good	State-of-the-art
Data Need	Low	High

References

- Guide to CNNs for CV – Khan et al. (2018)
- Deep Learning with Python – Chollet (2018)
- Deep Learning in Computer Vision – Awad & Hassaballah (2020)

- Deep Learning for Vision Systems by Mohamed Elgendy (2020)