Sensor Fusion and Artificial Intelligence in Autonomous Vehicles (AVs)

Rohollah (Roham) Moghadam, PhD

Masoud Ghodrat Abadi, PhD

SACRAMENTO STATE

WAVE

WORKFORCE FOR AUTONOMOUS VEHICLE ENGINEERING

MTI MINETA TRANSPORTATION INSTITUTE

Learning Objectives

- **Describe** sensors and perception systems used in autonomous vehicles
- **Explain** why no single sensing modality is sufficient in real-world driving scenarios
- **Describe** the concept and benefits of sensor fusion for perception in AVs
- **Understand** how artificial intelligence and machine learning support perception systems in AVs
- **Differentiate** neural networks, deep neural networks, and CNNs used in AV perception
- **Recognize** how sensor fusion improves robustness and safety under challenging conditions

Sensor

- A **sensor** is a device that detects, measures, and responds to physical, chemical, or environmental changes (inputs) such as light, heat, motion, or pressure. It converts these physical phenomena into signals usually electrical that can be read by a user or transmitted to a controller.

  Examples: Camera (light), radar(radio waves), LiDAR(laser), IMU (acceleration and rotation)

- Sensors act as "**eyes and ears**" for technology, bridging the gap between the physical and digital worlds.

Sensors in Autonomous Transportation

- Vehicles must:
  - Perceive the environment
  - Understand road users and infrastructure
  - Make safe decisions in real time
- Humans rely on vision and hearing → AVs rely on **multiple sensors working together**

Categories of Sensors in Autonomous Transportation

- **Exteroceptive sensors** (environment)
  - Cameras, LiDAR, Radar, Ultrasonic
- **Proprioceptive sensors** (vehicle state)
  - IMU, wheel encoders, odometry
- **Positioning sensors**
  - GPS/GNSS

Perception System in Autonomous Vehicle

- Autonomous vehicles rely on a **Perception System** to:
  - Sense the environment
  - Interpret traffic participants
  - Build a consistent world representation
- Perception is the **first and most safety-critical stage** in the AV pipeline.

Definition of Perception System

- **Perception System** is a system that processes multi-modal sensor data to achieve accurate, robust, and real-time understanding of the driving environment.
- Key characteristics:
  - **Understand**
  - **Combine**
  - **Interpret**

Why Multi-Modal Perception is Necessary

- No single sensor satisfies all perception requirements
- Sensors have complementary properties
  - Cameras → semantics
  - LiDAR → geometry
  - Radar → motion & robustness
- **Sensor fusion** is **fundamental**, not optional

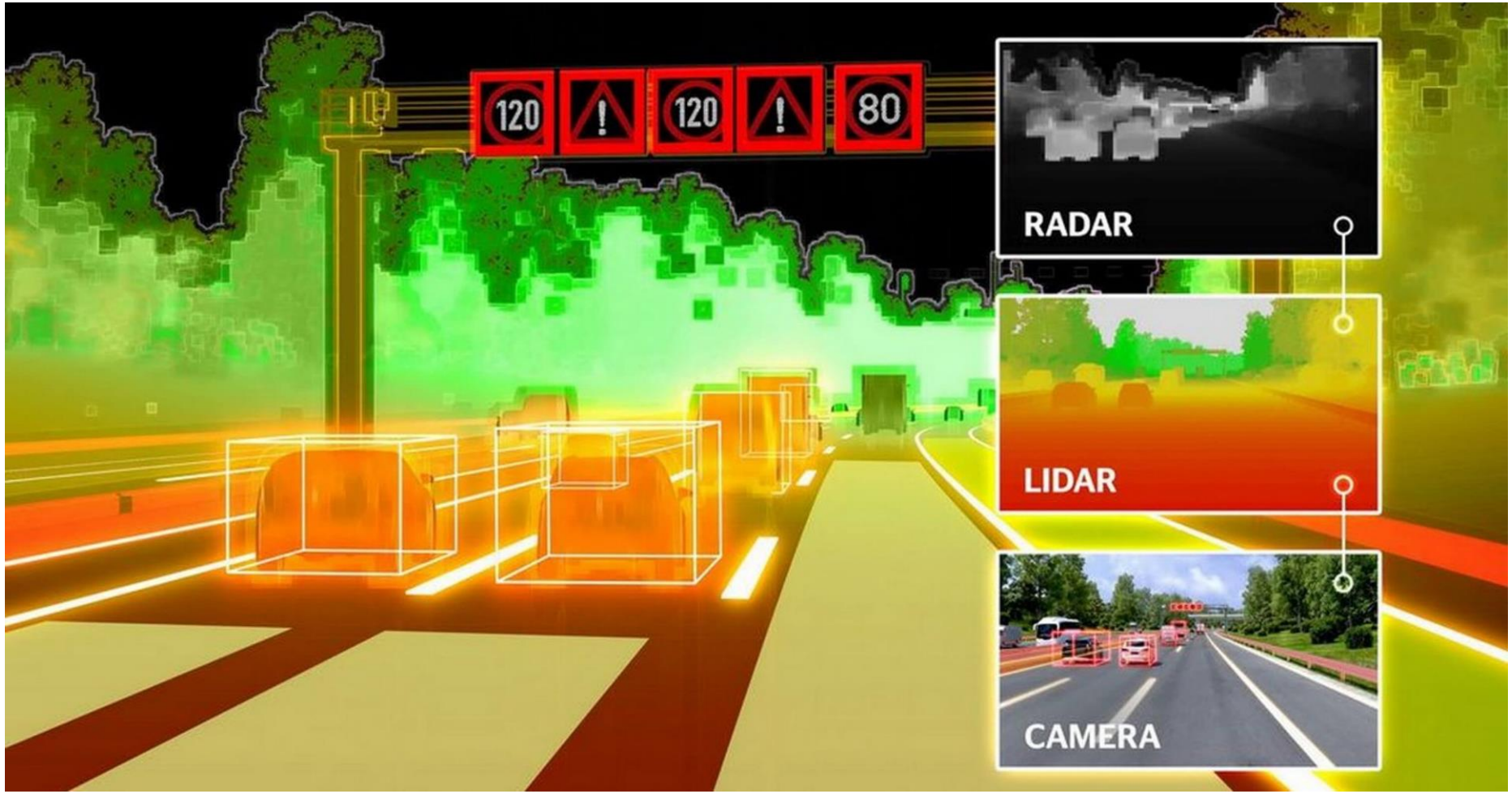
Sensor Fusion

- Sensor fusion is the process of **combining data from multiple sensors** to form a more accurate and reliable understanding of the environment.
- Instead of relying on one sensor, the fusion uses multiple sensors to cover each other's weaknesses

Role of Sensor Fusion in Perception

- **Sensor fusion:**
  - Combines heterogeneous sensor data
  - Reduces ambiguity
  - Improves reliability under sensor degradation
- **Fusion directly supports:**
  - Object detection
  - Semantic segmentation
  - Localization

Combination of Different Sensors



https://www.image-engineering.de/library/blog/articles/1157-challenges-for-cameras-in-automotive-applications

Sensor Fusion in Autonomous Driving Tasks

- Object detection and tracking
- Lane recognition
- Localization and mapping
- Collision avoidance and emergency braking

Sensor Fusion & Popular algorithms

- **Kalman Filter:**
  - Common in linear systems
  - Used in vehicle motion models
- **Extended Kalman Filter:**
  - Common in nonlinear systems
  - Used in vehicle localization, and pose estimation

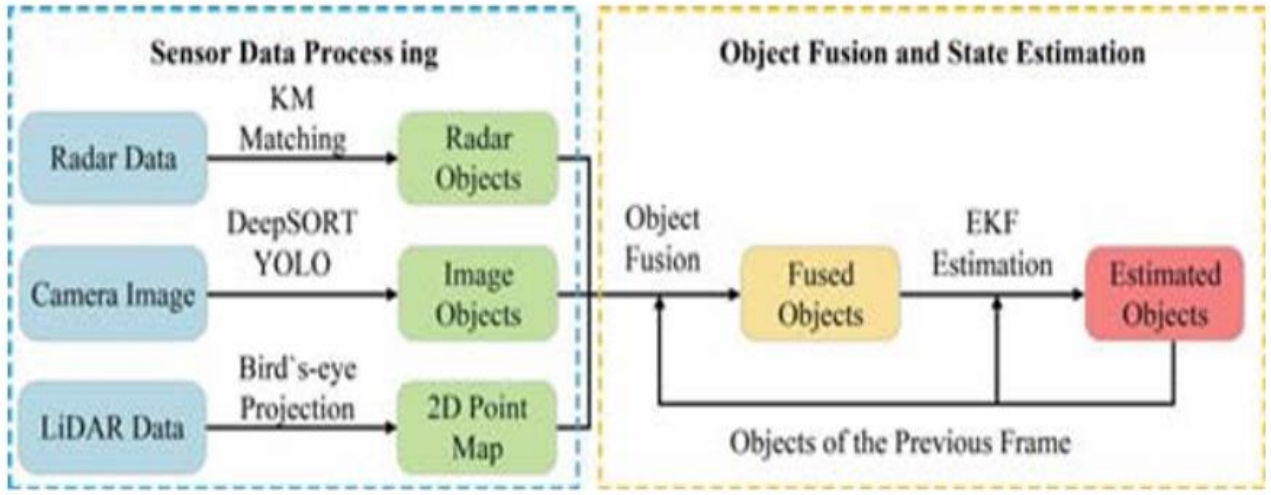
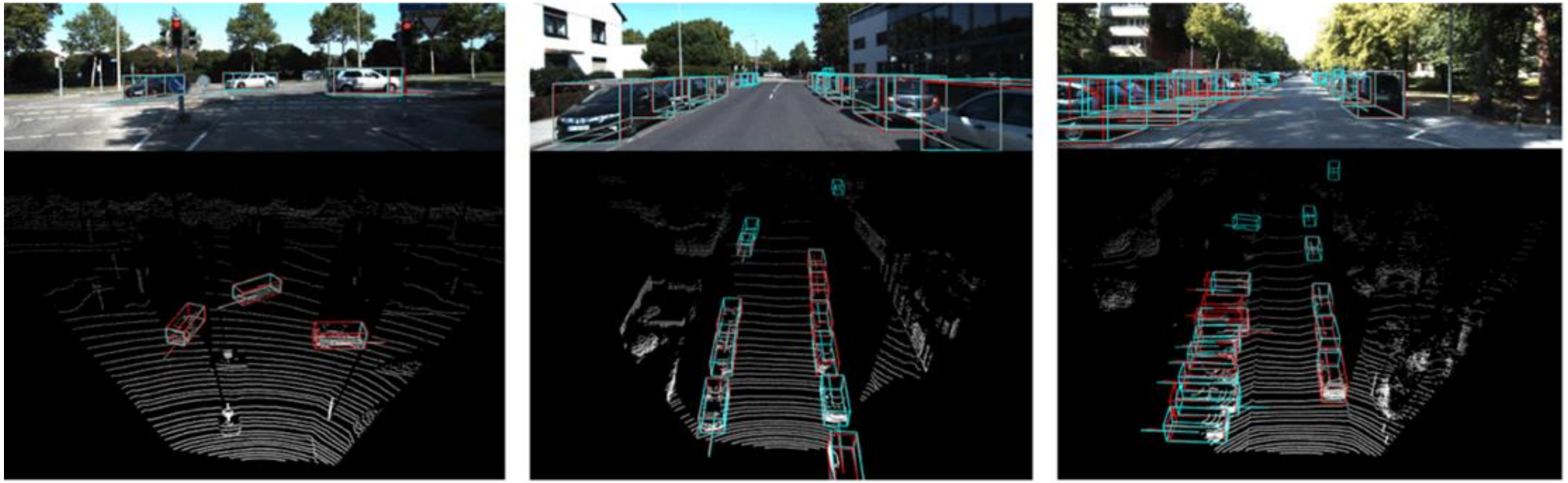


Figure adapted from Li et al., "Object Detection and State Estimation of Autonomous Vehicles with Multi-Sensor Information Fusion," IEEE Conference Publication, 2021.

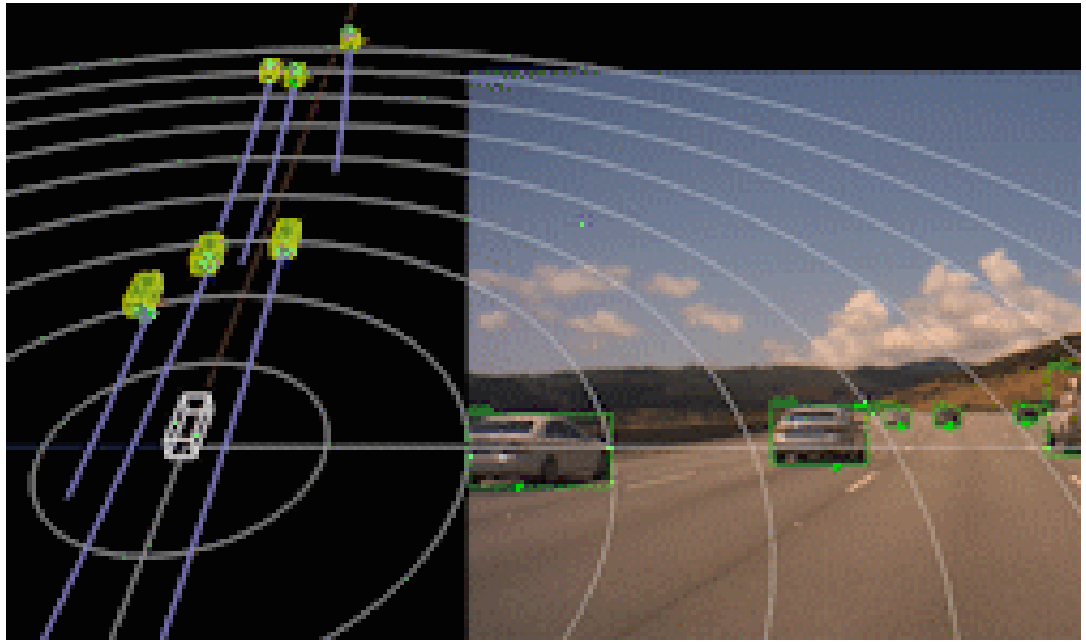
Camera, and LiDAR fusion



LiDAR and Camera Sensor Fusion in Self-Driving Cars

Applications of Sensor fusion in AVs

- **Perception & Object Detection:**
  - **Camera + LiDAR + RADAR**
  - Fusion allows for 3D mapping, object detection, and classification
- **Lane Detection**
  - **Camera + IMU + GPS**
  - Popularly used to keep AV in lane and proper lane changing
  - Improved redundancy from simple camera algorithms



DRIVE Labs: Covering Every Angle with Surround Camera-Radar Fusion | NVIDIA Technical Blog

Sensor Fusion and Safety

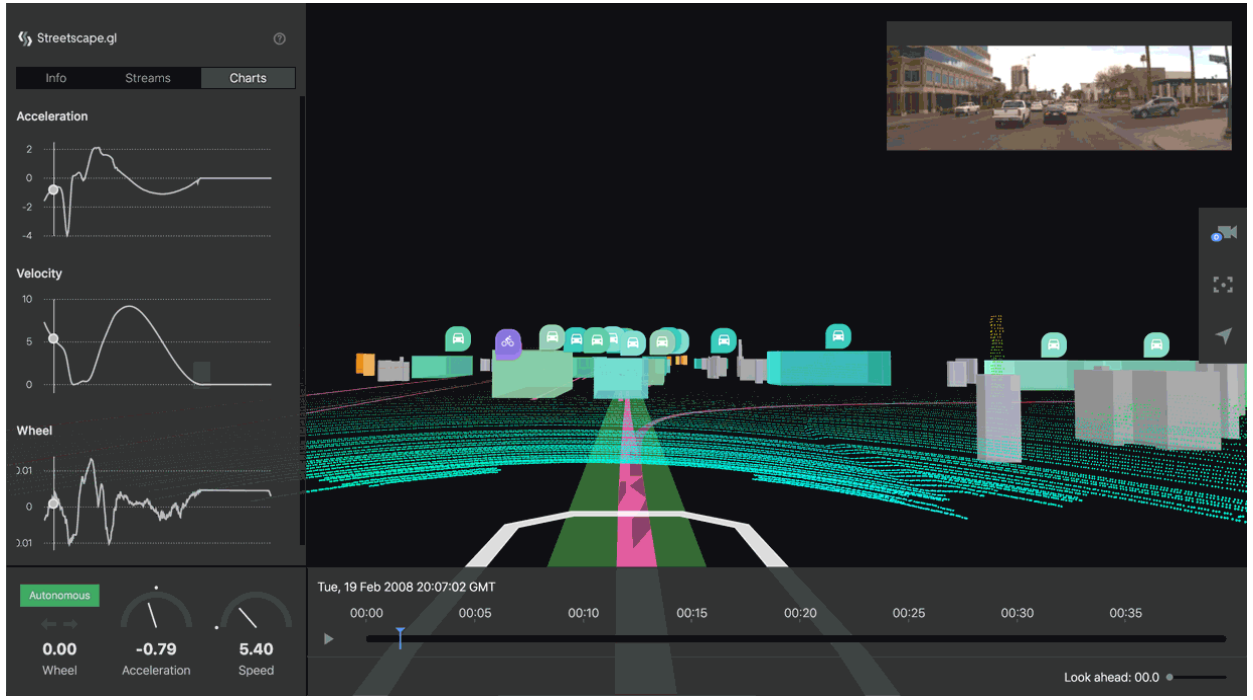
- Fusion Improves Performance In:
  - Poor lighting (night, glare, shadows)
  - Adverse weather (rain, fog, snow)
  - Partial sensor failure or degradation
- Why It Matters
  - Sensors back each other up when one becomes unreliable
  - Reduces single-point failures
  - Increases system robustness, reliability, and trustworthiness

What is Artificial Intelligence (AI)

- AI: Systems that perform tasks requiring human-like intelligence
- In AVs, AI is mainly used for:
  - Seeing
  - Recognizing
  - Interpreting the environment
- Think of AI as **software that learns from experience**

Artificial Intelligence (AI) in AVs

- AI in AVs refers to **data-driven algorithms** that learn patterns from data which can then make decisions and adapt in a dynamic environment
- **Perception stage:** Data from sensors used to understand environment
- **Prediction stage:** Predicting changes of other vehicles, obstacles, and pedestrians
- **Assist in Decision making:** Road management strategies (Turns, yielding, overtaking)



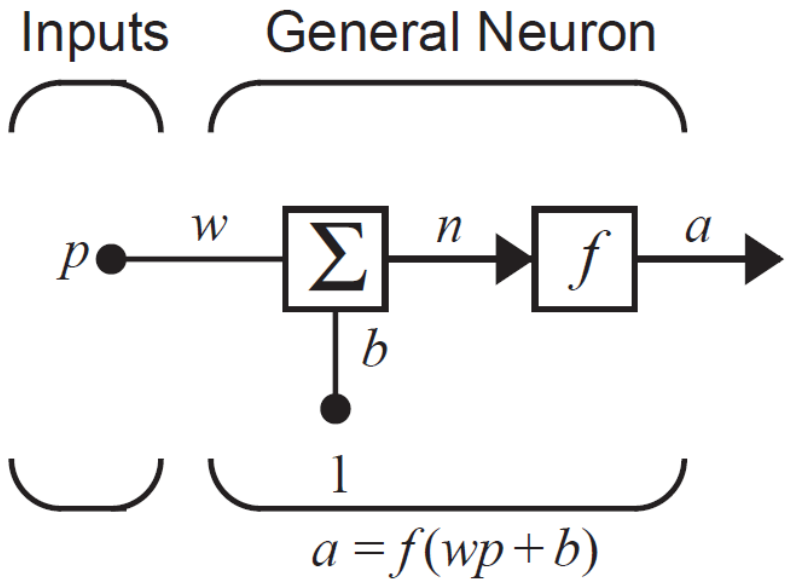
Introducing AVS, an Open Standard for Autonomous Vehicle Visualization from Uber | Uber Blog

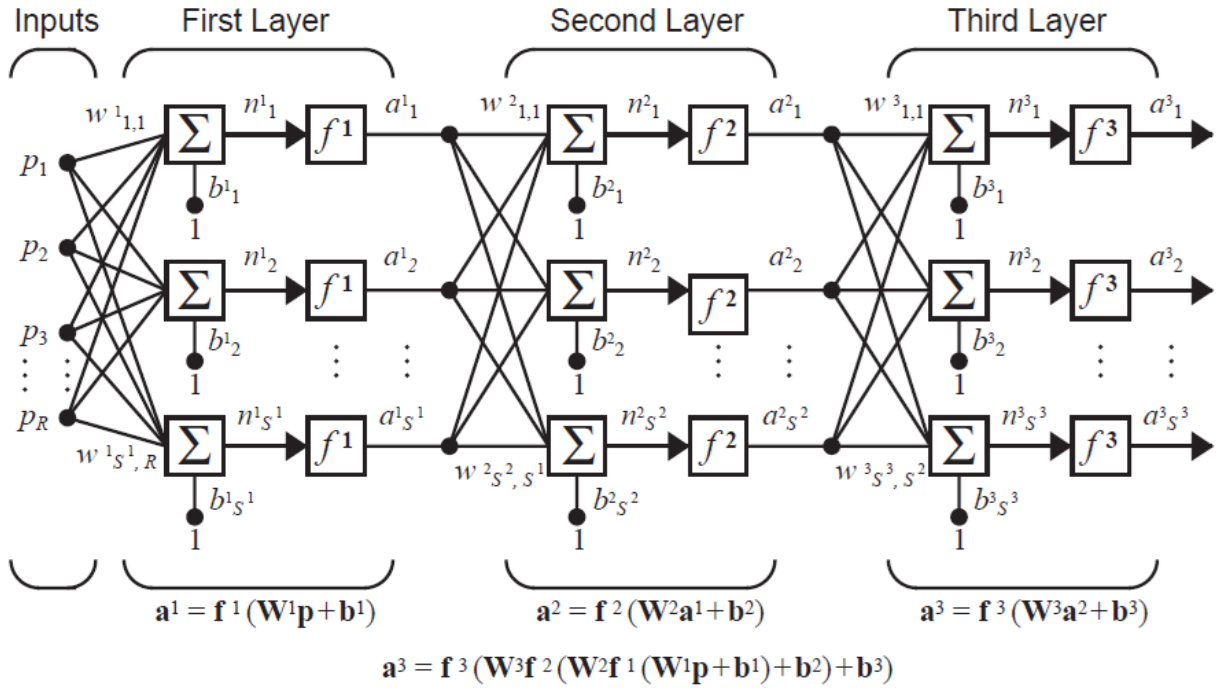
What is Machine Learning

- A subset of Artificial Intelligence
- Systems **learn patterns from data** instead of using fixed rules
- Performance **improves with more data and experience**
- Example in AVs:
  - Show many labeled images of cars, pedestrians, and lanes, the model learns to **recognize them automatically** in new situations

What is Neural Network

- Inspired by how the human brain processes information
- Made of layers of simple processing units (**neurons**)
- Each layer extracts **increasingly complex features**

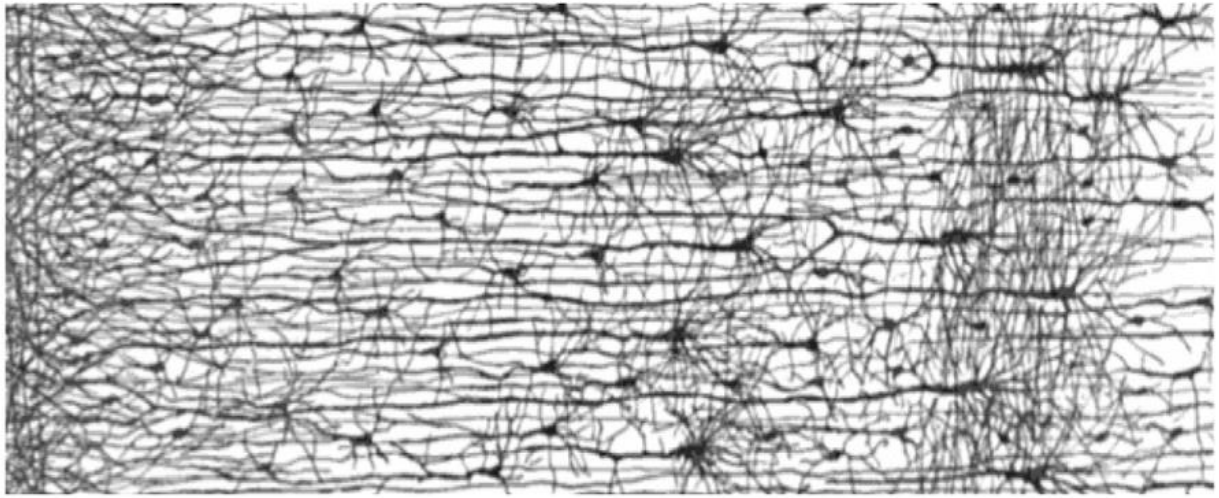


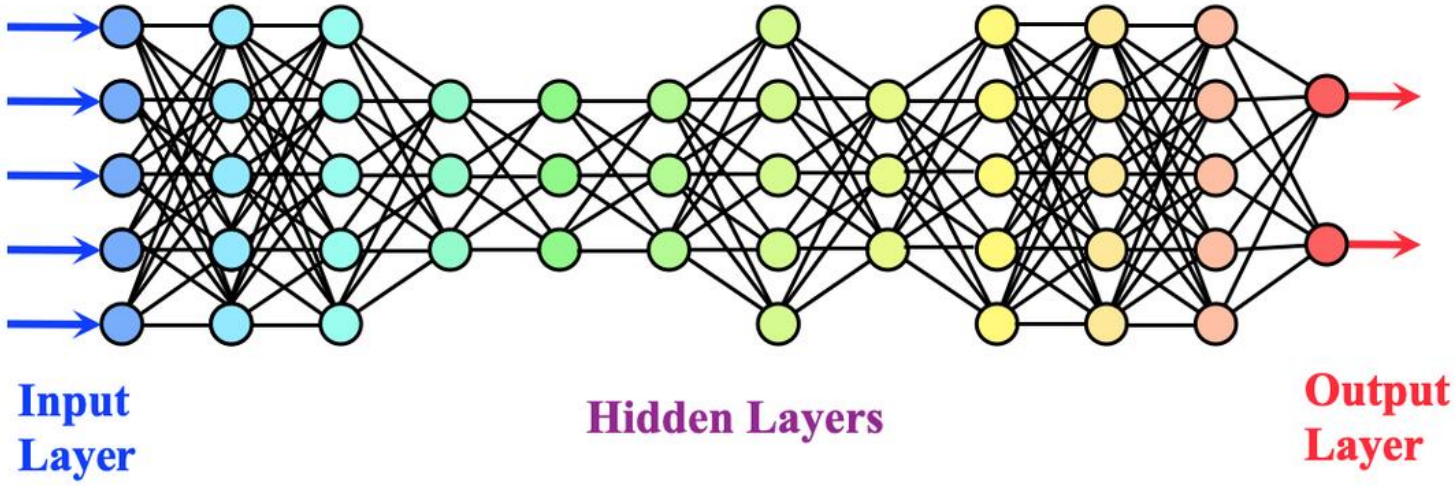


Deep Neural Network

- A neural network with many hidden layers that learns complex and hierarchical patterns from large datasets *(simple features → meaningful representations)*
- DNN can model nonlinear, real-world relationships, scale effectively with large sensor datasets
- In AVs, DNN used for object detection, lane recognition, and traffic sign recognition and forms the backbone of modern perception systems

Deep Neural Network



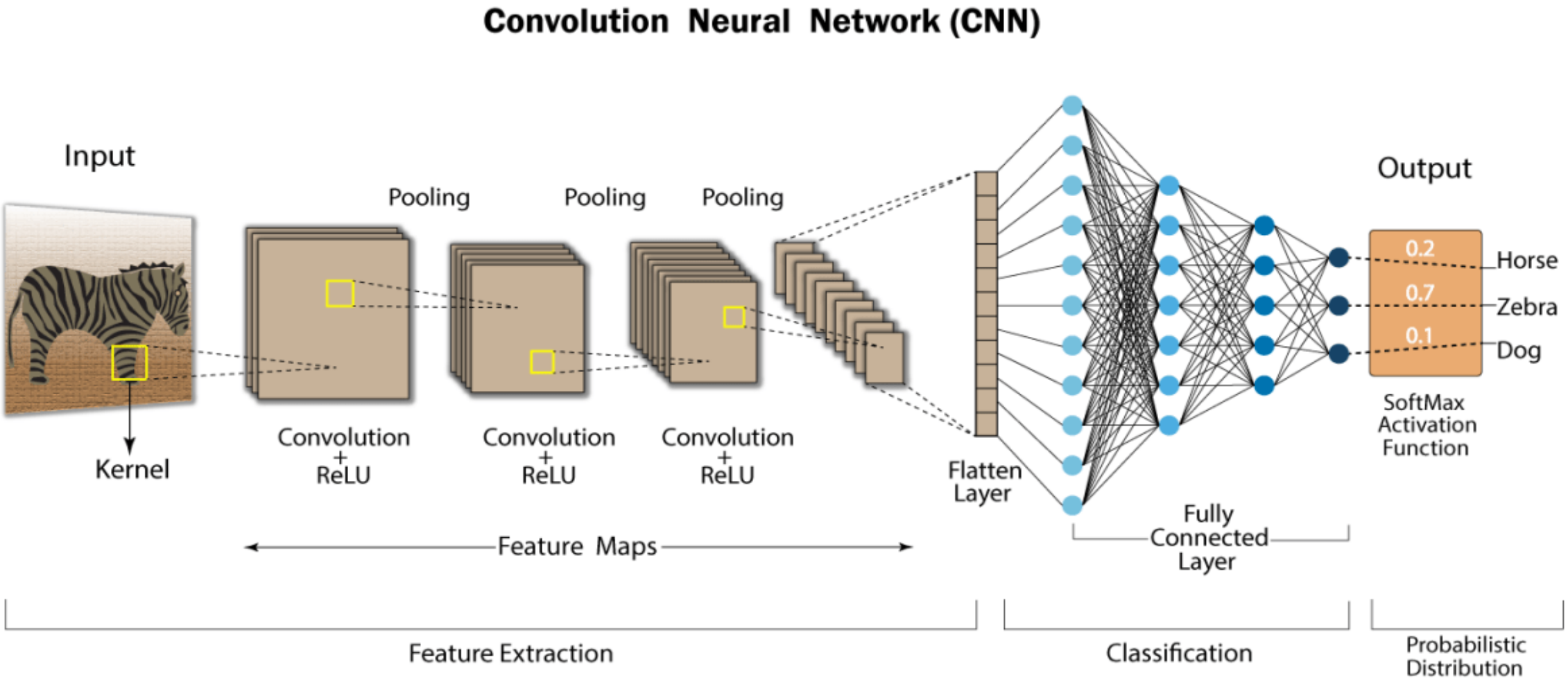


https://mriquestions.com/what-is-a-neural-network.html

Convolutional Neural Network (CNN)

- A specialized neural network for images and spatial data
- Automatically learns visual features through layers (edges→shapes→objects)
- Widely used in AVs:
  - Object detection (detect pedestrians, recognize vehicles)
  - Identify road features (lane detection, traffic sign recognition)
- CNN reduces the need for manual feature design
- CNN enables robust perception from camera data

Convolutional Neural Network (CNN)

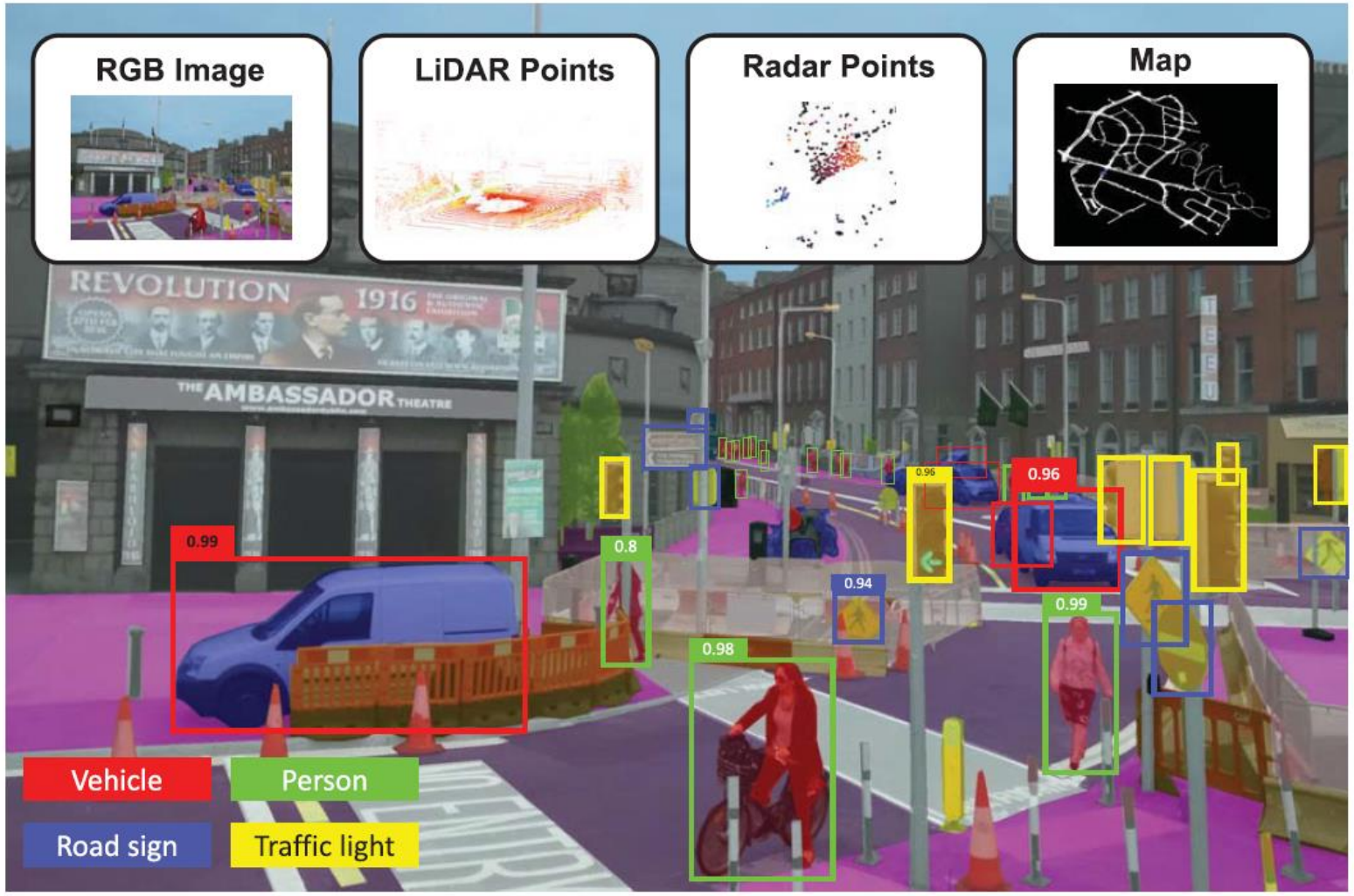


https://medium.com/@prathammodi001/convolutional-neural-networks-for-dummies-a-step-by-step-cnn-tutorial-e68f464d608f

Sensor Fusion + CNN

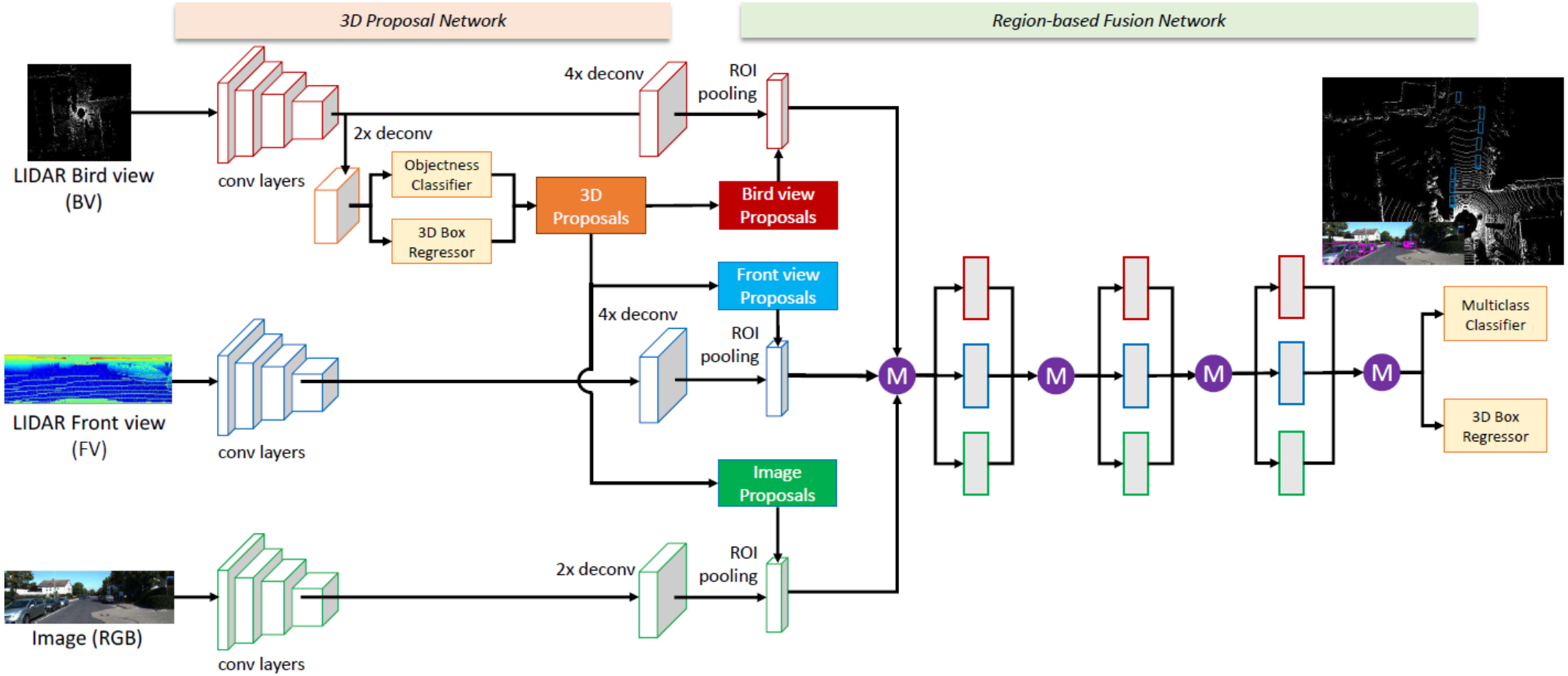
- Cameras → CNN extracts rich visual and semantic features
- LiDAR → Provides accurate 3D geometry and depth structure
- Radar → Measures motion and relative speed, robust in bad weather
- AI learns how to fuse complementary information from all sensors
- This results in a unified and robust understanding of the environment and is called **multi-modal perception**

Multi-View 3D Object Detection Network



A complex urban scenario for autonomous driving [2] .

Multi-View 3D Object Detection Network



Multi-View 3D Object Detection Network for Autonomous Driving [1]

Sensor Fusion Challenges

- **High Data Volume**: Multiple sensors generate large amounts of data that must be processed in real time.
- **Computational Demands**: Sensor fusion and AI algorithms require significant onboard computing power.
- **Sensor Calibration and Alignment**: Accurate spatial and temporal alignment between sensors is critical for reliable fusion.
- **AI Explainability**: Understanding and interpreting AI-based fusion decisions remains difficult.

Sensor Fusion Connection to Civil Engineering

- Infrastructure must support:
  - Clear markings
  - Consistent signage
  - Predictable layouts
- Poor infrastructure → poor perception
- CE design directly affects AV perception quality

References

1. X. Chen, H. Ma, J. Wan, B. Li and T. Xia, "Multi-view 3D Object Detection Network for Autonomous Driving," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017, pp. 6526-6534, doi: 10.1109/CVPR.2017.691.
2. D. Feng et al., "Deep Multi-Modal Object Detection and Semantic Segmentation for Autonomous Driving: Datasets, Methods, and Challenges," in IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 3, pp. 1341-1360, March 2021, doi: 10.1109/TITS.2020.2972974.
3. Waymo LLC, *Waymo Safety Report*, Waymo, Mountain View, CA, USA, 2017.