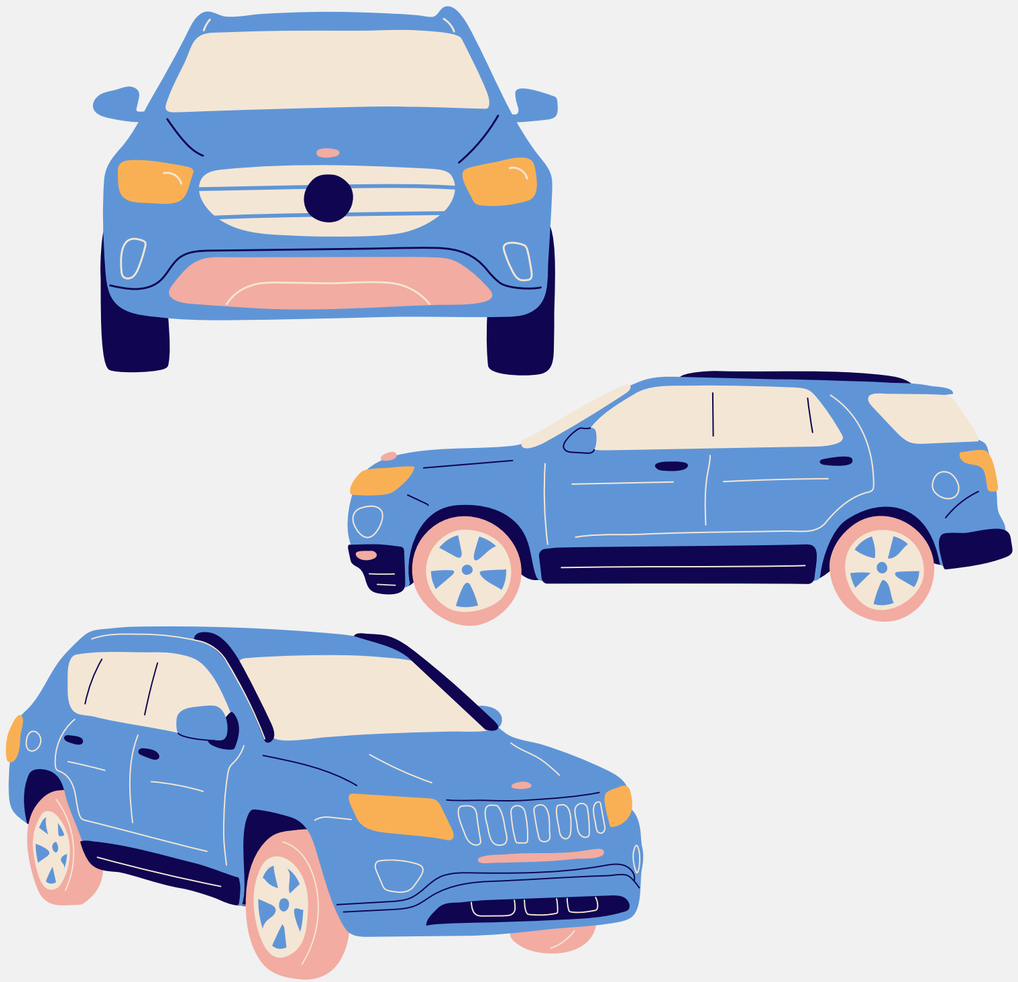
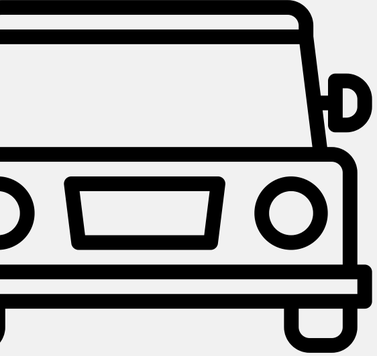
ASME EFEST

Computer Vision for Autonomous Vehicles

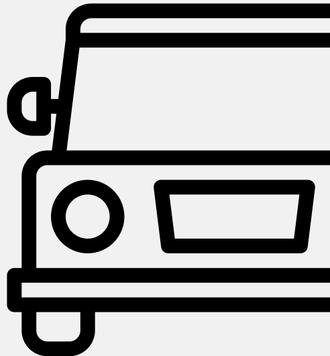
Jana Kabrit



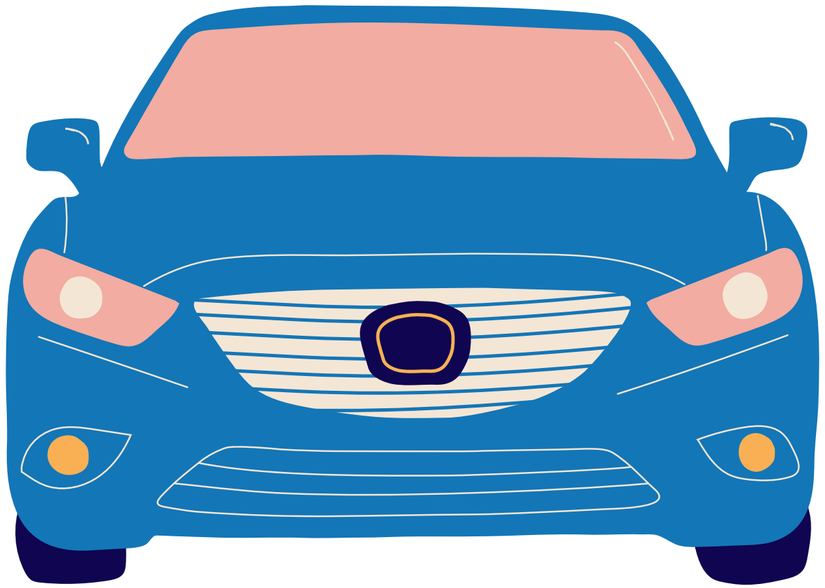


Who am I?

- Applied Research Engineer @ AEMS AI
- Artificial Intelligence Instructor @ Zaka AI



Today's Topics

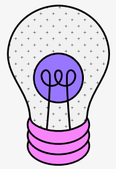


- What actually is AI?
- Levels of AI
- Types of AI
- Fields of AI
- Computer Vision for Autonomous Vehicles
- CV-AV: Pipelines and State of The Art
- Hands-on Example

What Is AI?

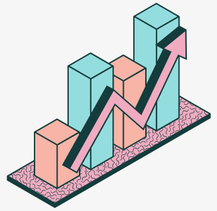
The term AI, coined in the 1950s, refers to the simulation of human intelligence by machines.

Levels of AI



Artificial Intelligence

Hard-coding



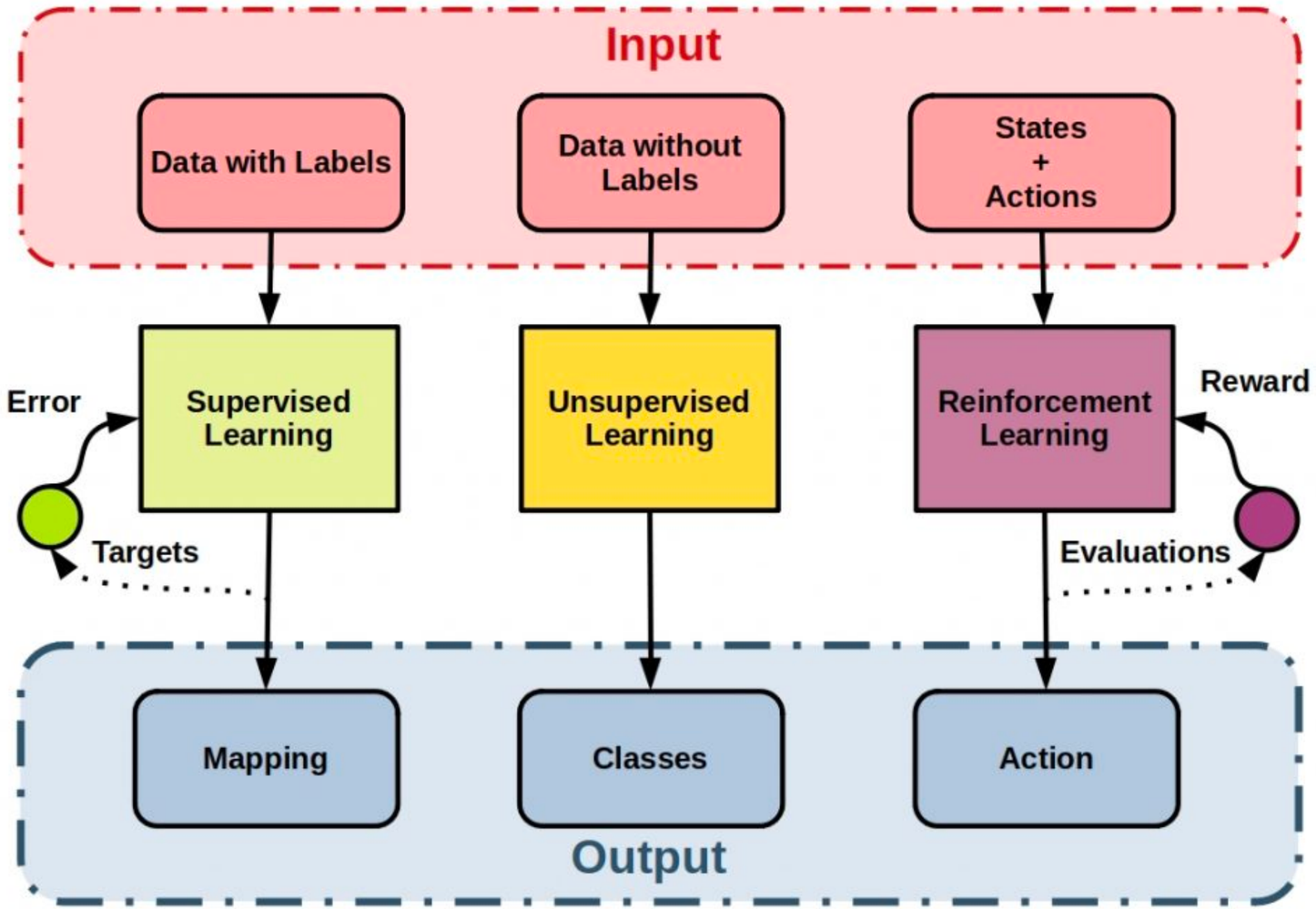
Machine Learning

Using statistical models

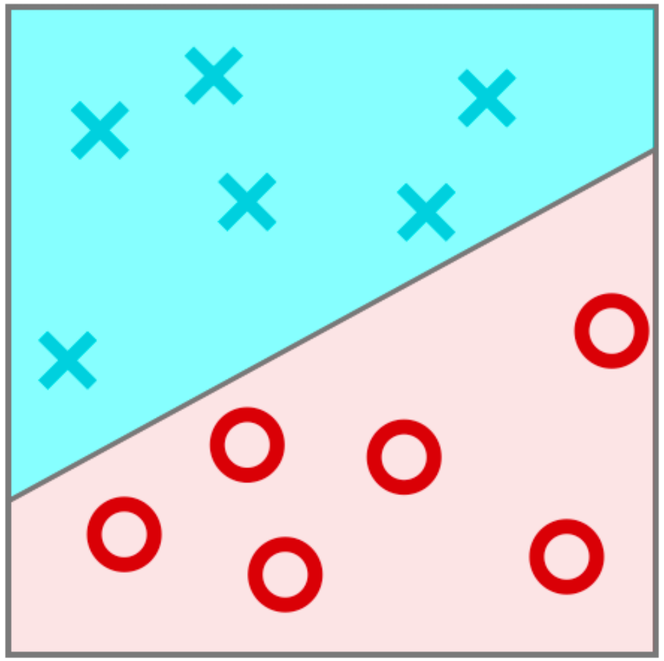
Deep Learning

Using Neural Networks

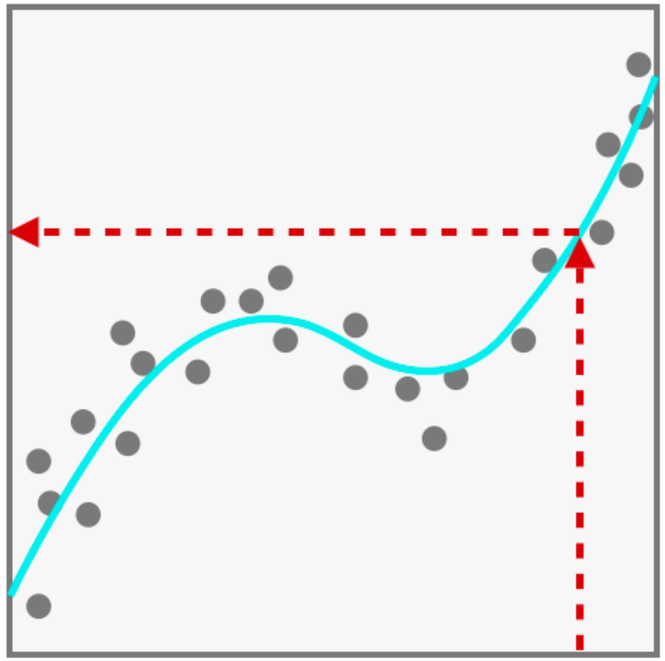
Types of AI


Input
Data with Labels
Data without Labels
States + Actions
Error
Supervised Learning
Unsupervised Learning
Reinforcement Learning
Reward
Targets
Evaluations
Mapping
Classes
Action
Output

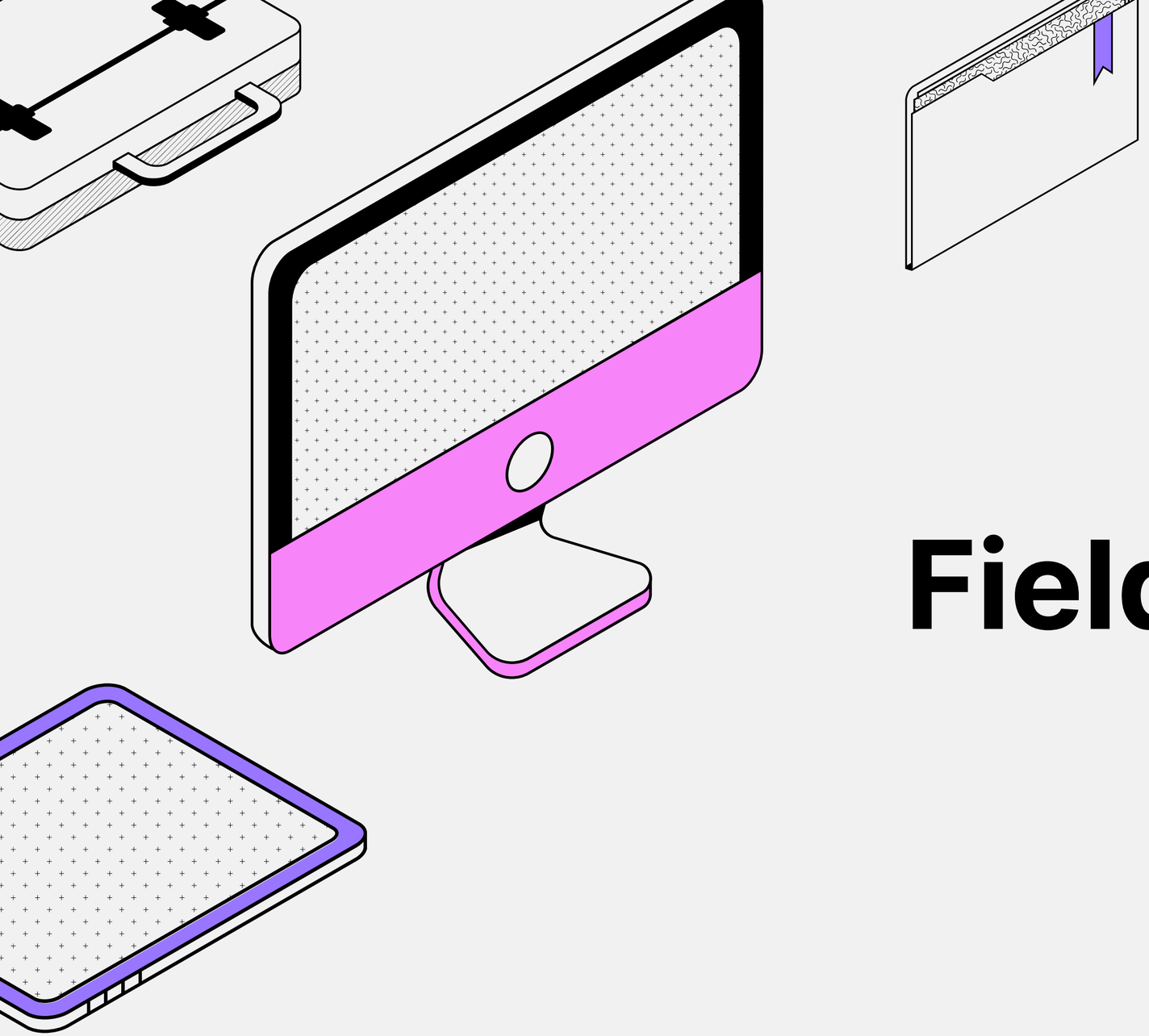
Diving into Supervised Learning



Classification



Regression

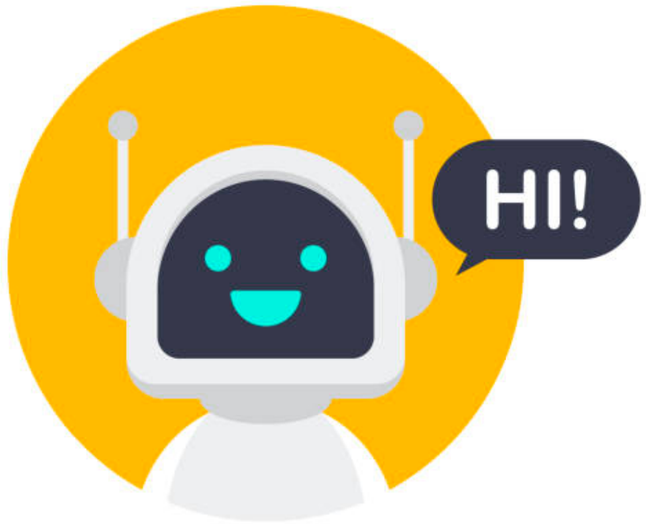


Fields of AI

Natural Language Processing

What is NLP?

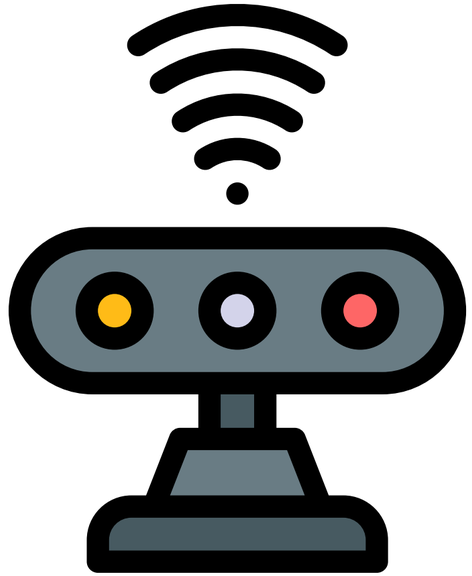
The application of computational techniques to the analysis and synthesis of natural language and speech.



Time Series Analysis

What is TSA?

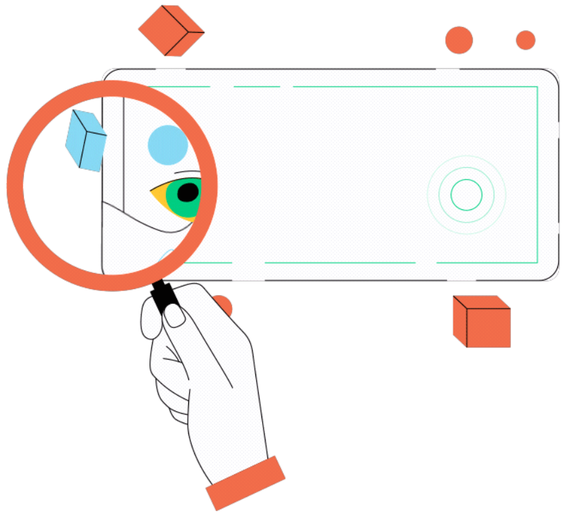
Time series analysis is a specific way of analyzing a sequence of data points collected over an interval of time.



Computer Vision

What is CV?

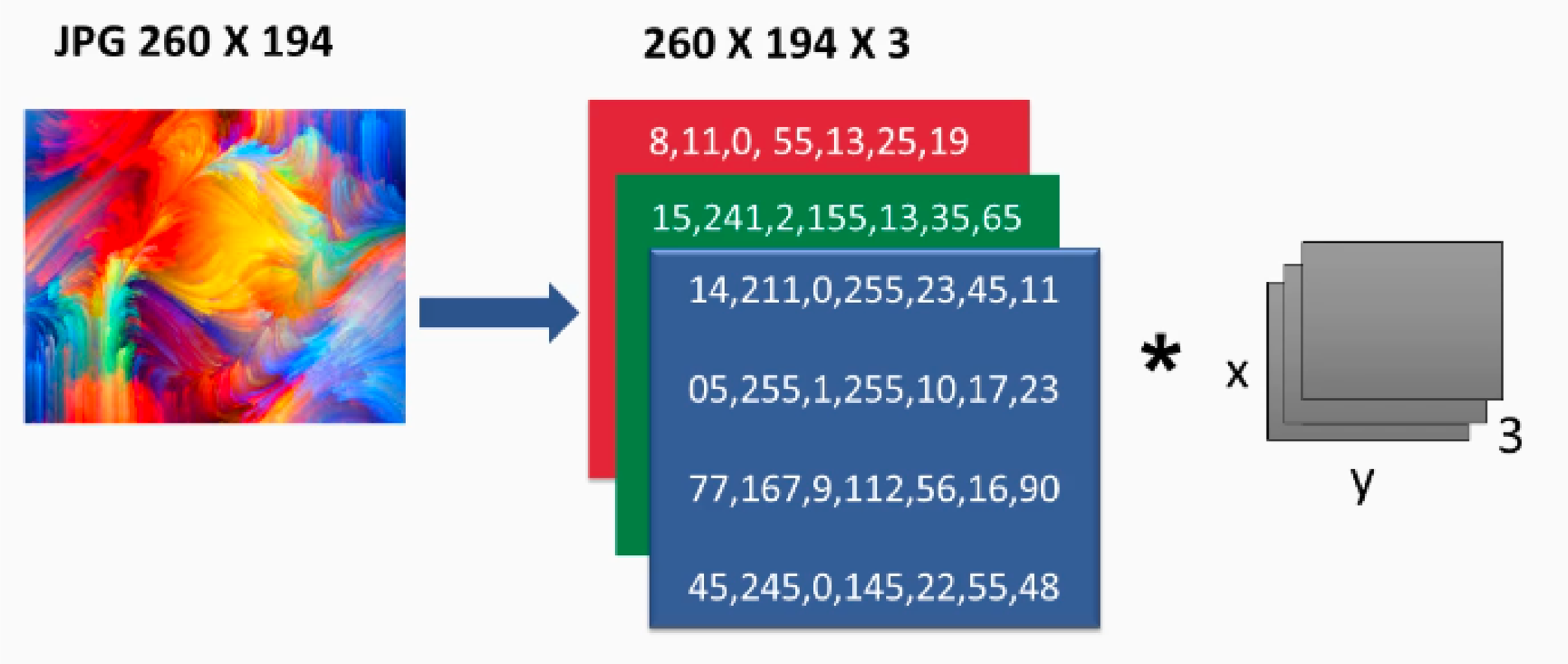
Computer vision is a field of artificial intelligence (AI) that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs



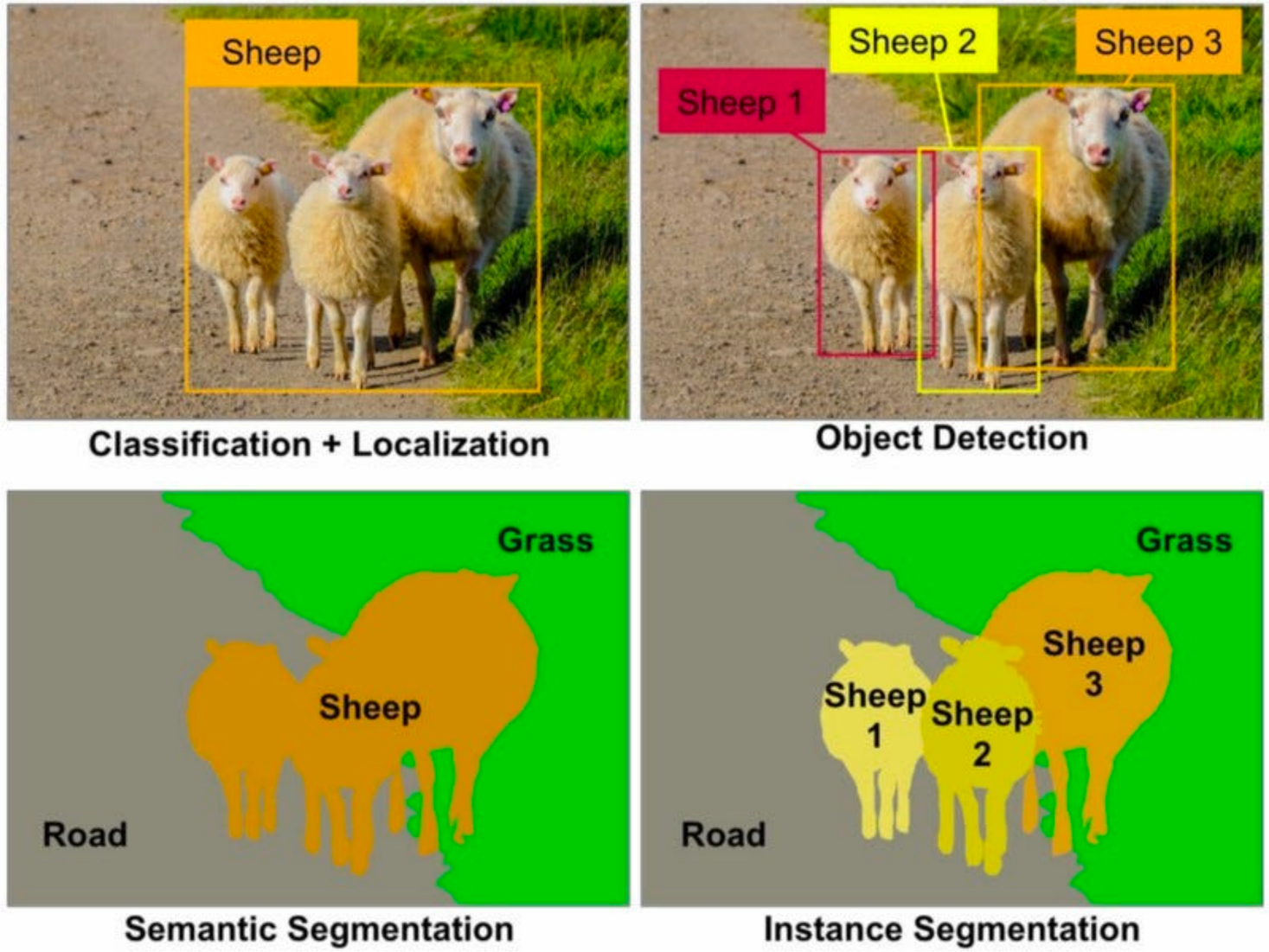
Computer Vision for Autonomous Vehicles

How do we process images?

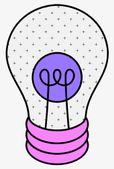
Videos are made of frames (ie images), which are made of pixels.



Types of Classification in CV



Computer Vision Use Cases in Vehicle Companies

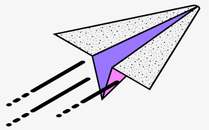


Manufacturing Automation

Most of the mass-produced cars in the industry are made in automated manufacturing facilities using CV.

Driving Assessment

Computer vision-enabled cameras inside the car are used to analyze the behavior of the driver. If the systems detect anomalies or unusual activities.

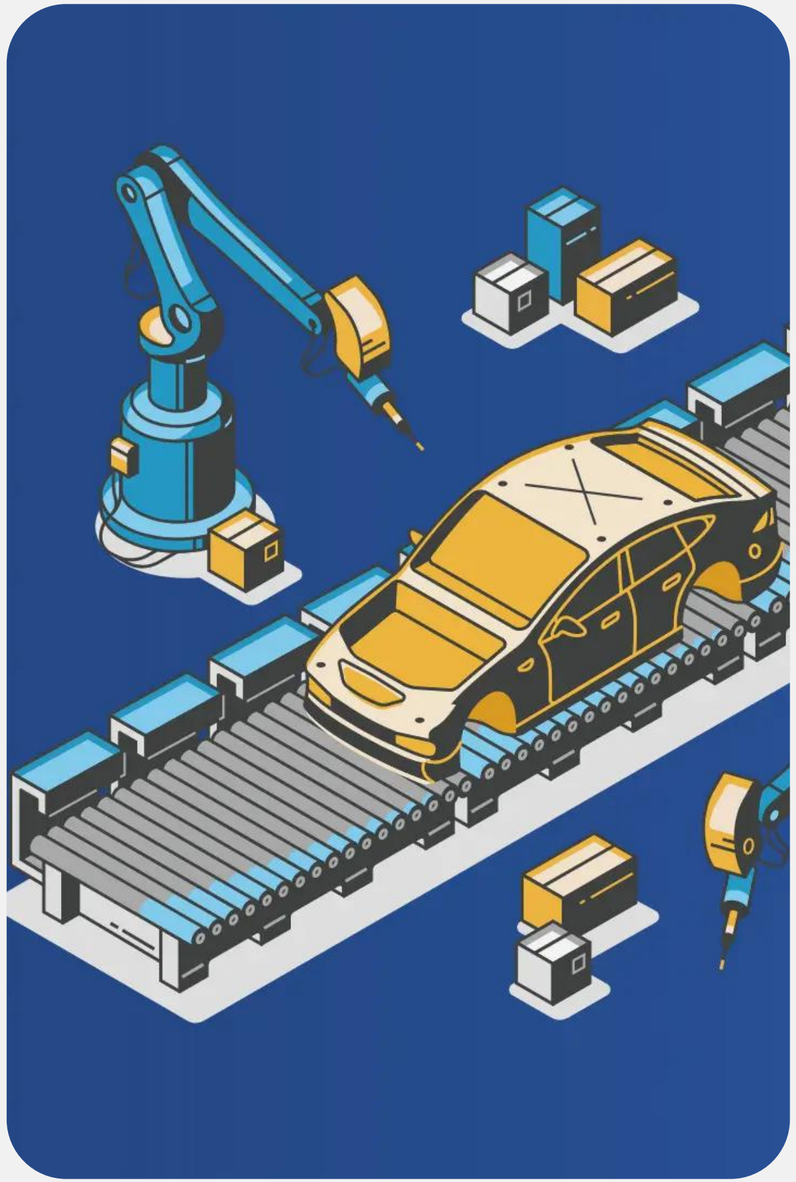


Dashcam Footage

Ranging from self-driving non-commercial cars to heavy-duty logistics trucks, computer vision and AI-enabled self-driving systems are revolutionizing the automotive sector.

Manufacturing Automation

In the automotive sector, vehicle assembly lines are automated through bots and computer vision systems to eliminate human errors in repetitive assembly tasks.

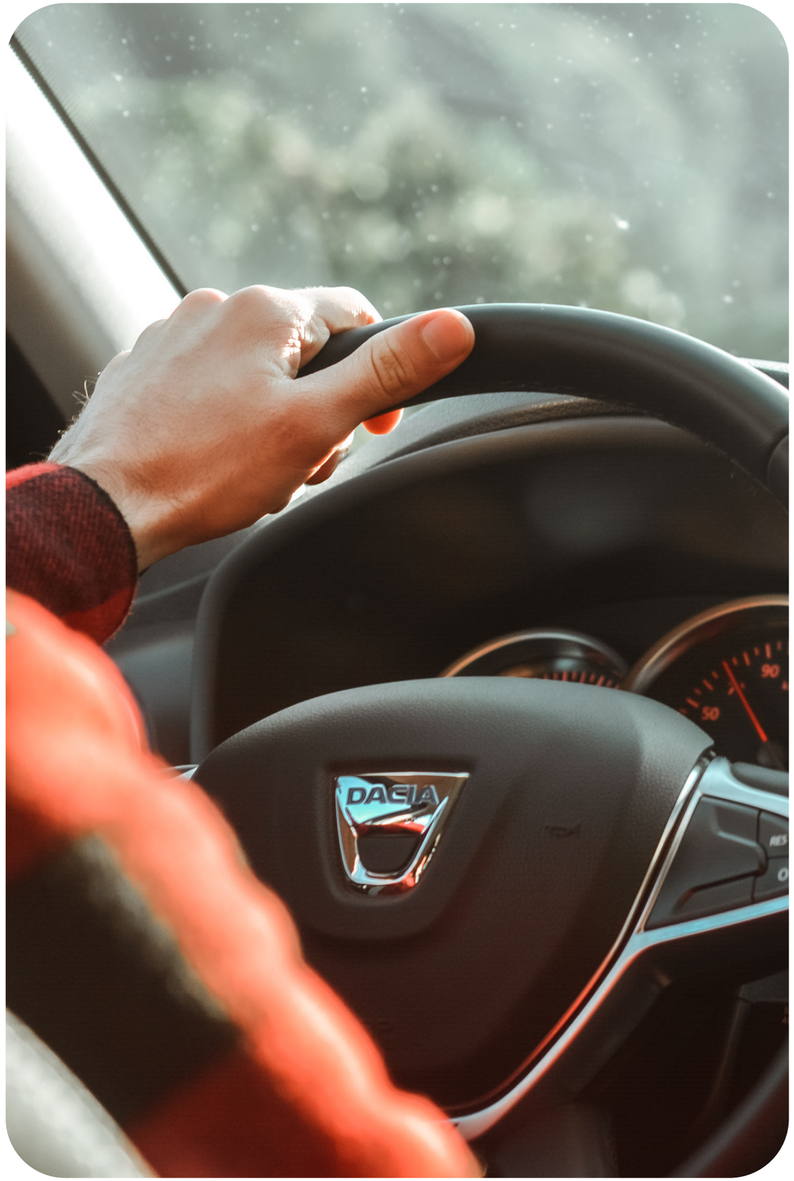


Manufacturing Automation

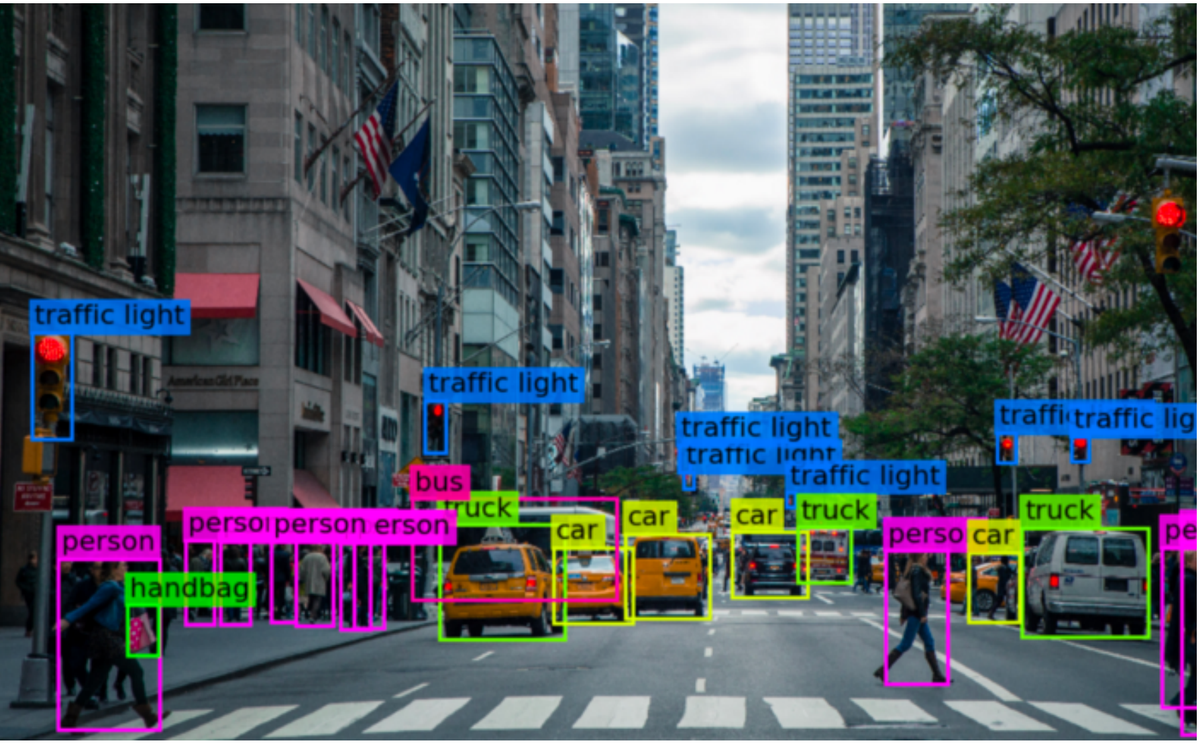
See how BMW uses computer vision and AI to inspect vehicle model tags.

Driving Assessment

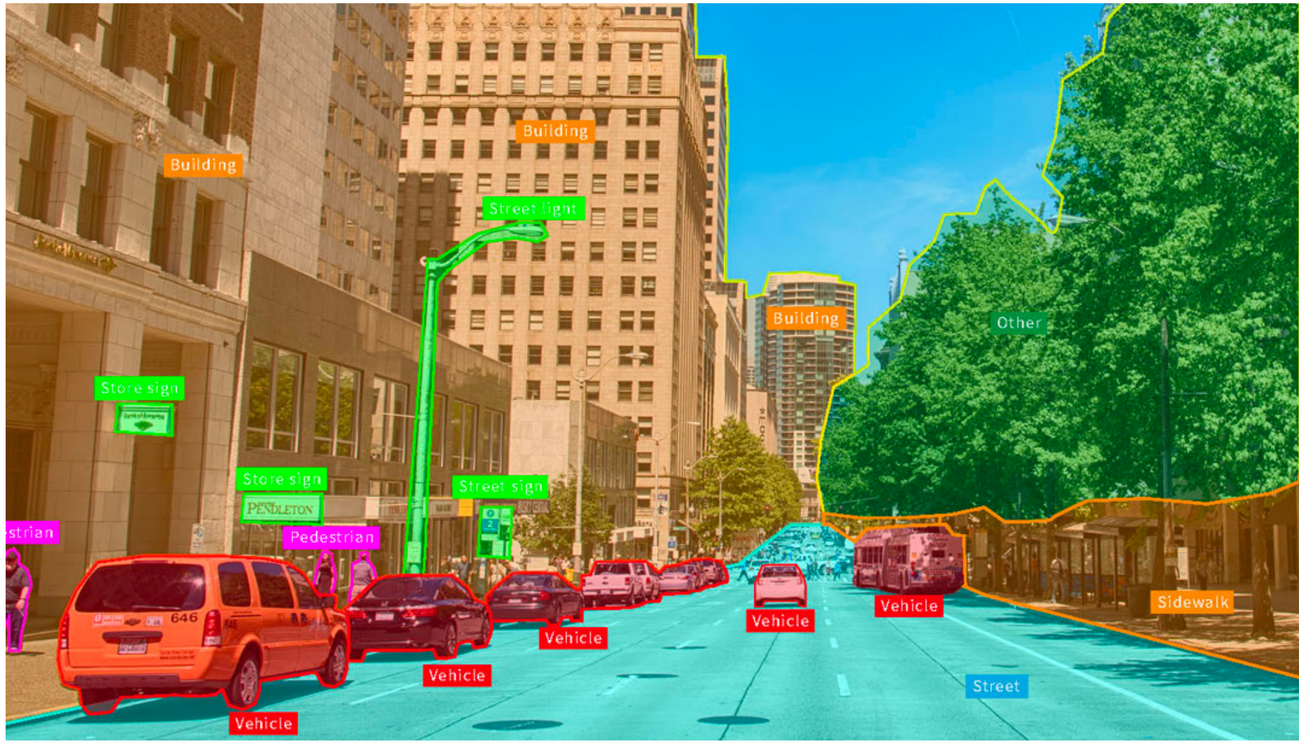
- Driver distraction
- Driver fatigue or tiredness
- Usage of a mobile phone
- Falling asleep
- Low blinking rate



Dashcam Footage



Object Detection



Instance Segmentation

How can we use dashcam footage?

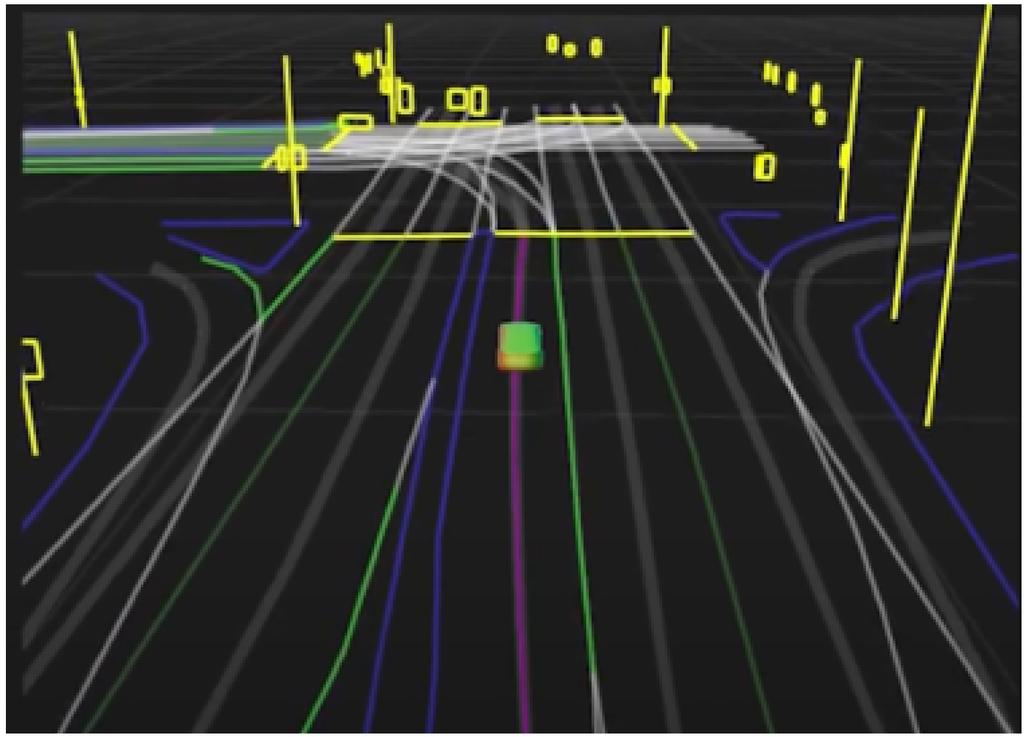
1. Perception

The car needs to see and classify traffic lights, pedestrians, road signs, walkways, parking spots, lanes, and much more. Not only that, it also needs to know the exact distance between itself and the objects around it. Perception is more than just seeing and classifying, it enables the system to evaluate the distance and decide to either slow down or brake.

This was usually done by a combination of cameras and sensors, but now cameras are enough!

How can we use dashcam footage?

2. Localization

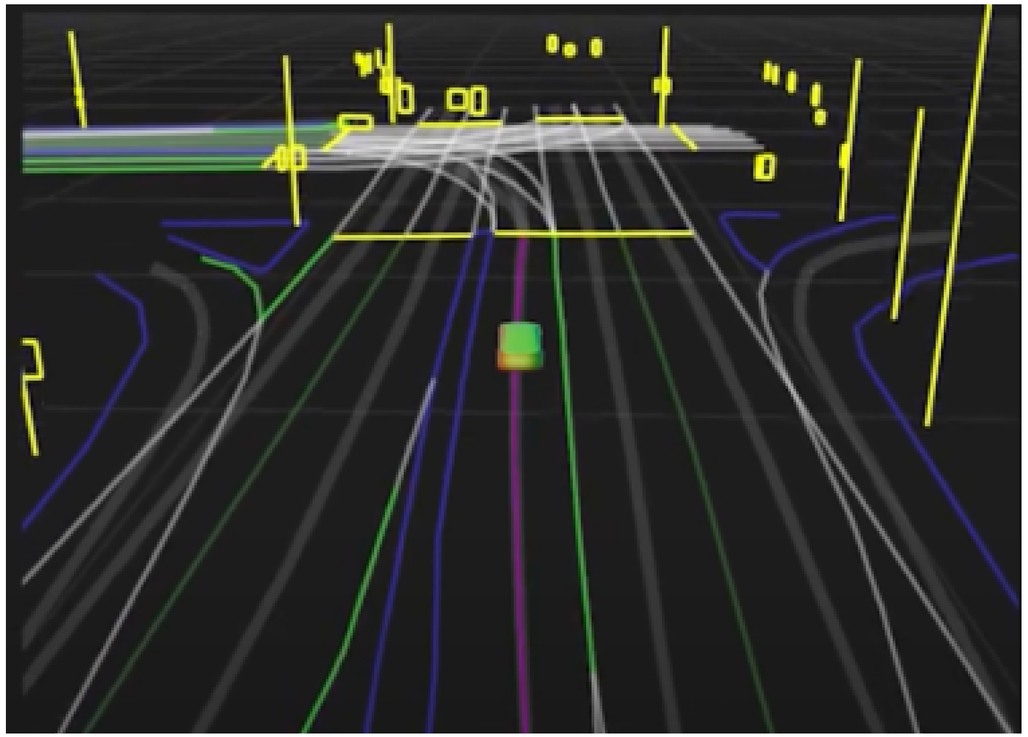


Localization algorithms in self-driving cars calculate the position and orientation of the vehicle as it navigates – a science known as Visual Odometry (VO).

VO works by matching key points in consecutive video frames. The mapping algorithm, such as Simultaneous localization and mapping (SLAM), computes the position and orientation of each object nearby with respect to the previous frame and helps to classify roads, pedestrians, and other objects around.

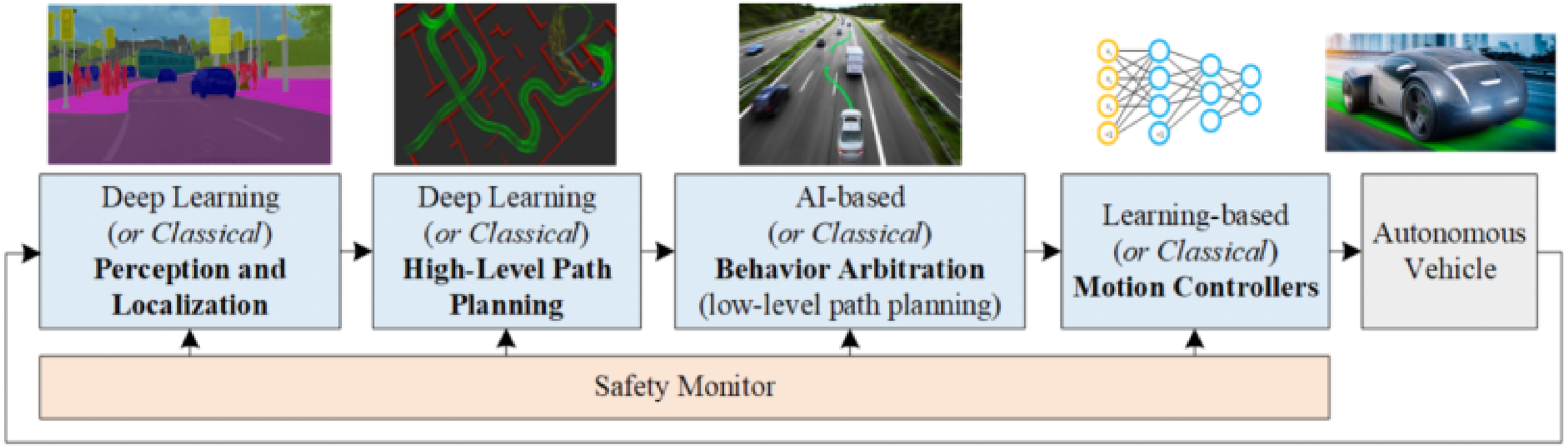
How can we use dashcam footage?

2. Prediction

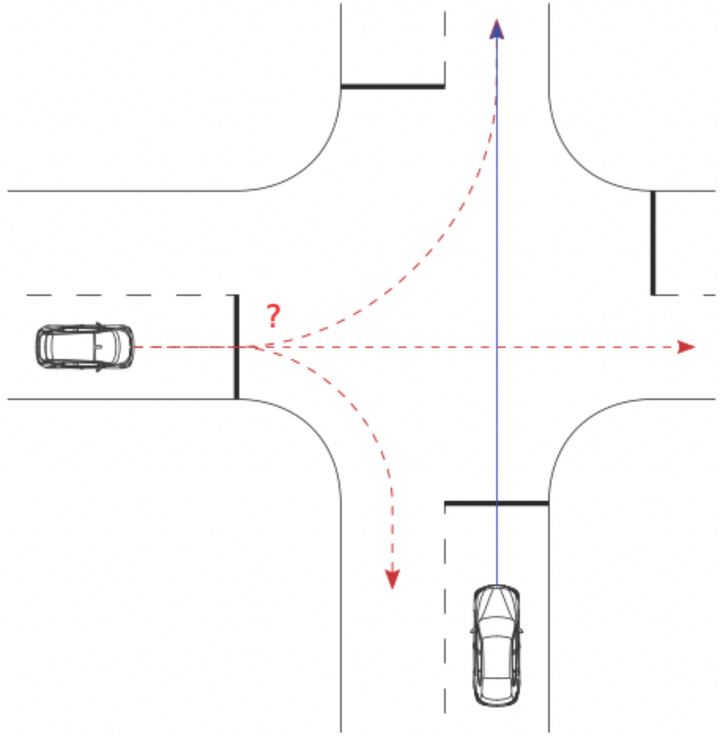


The car has a 360-degree view of its environment that enables it to perceive and capture all the information and process it. Once fed into the deep learning algorithm, it can come up with all the possible moves that other road users might make. It’s like a game where the player has a finite number of moves and tries to find the best move to defeat the opponent.

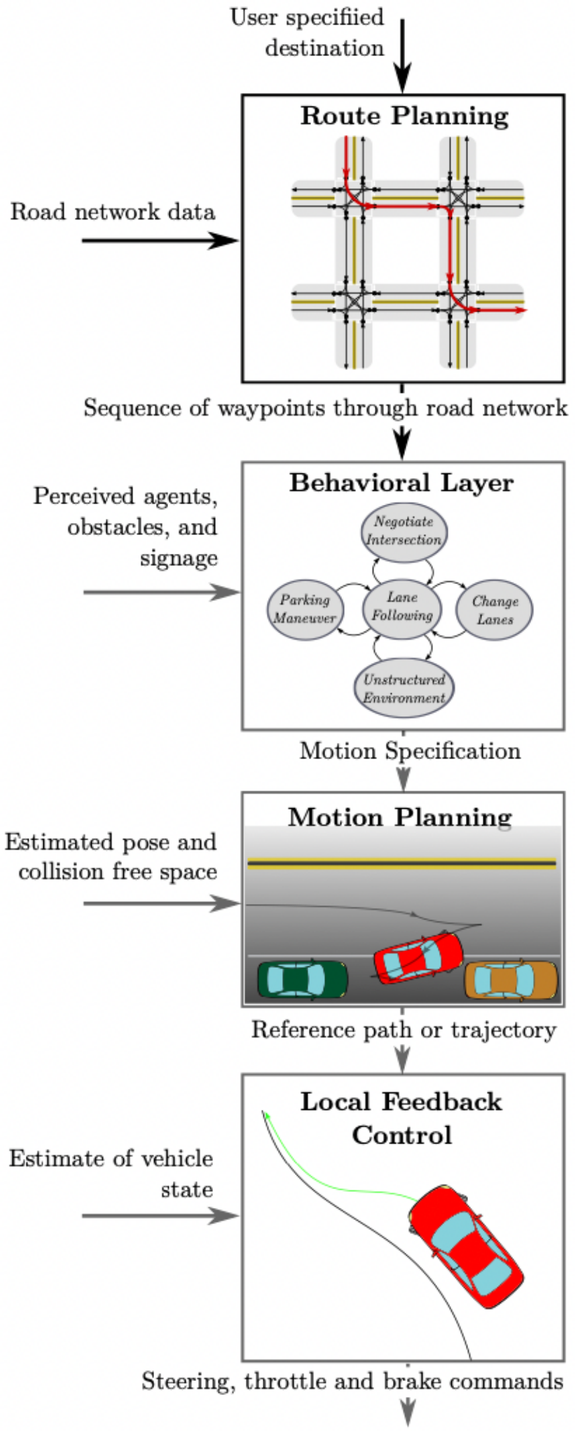
CV-AV: Pipelines and State of The Art



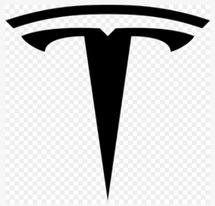
Self-driving cars: the full pipeline



Self-driving cars: the full pipeline



State of the Art Models



Tesla

Hydranet

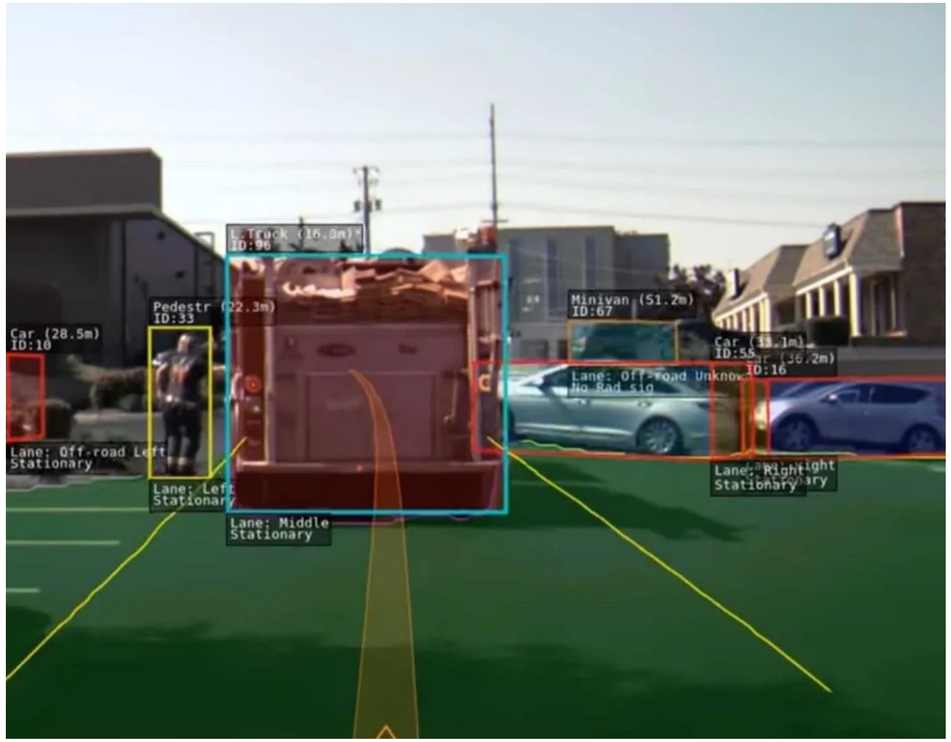
Google Waymo

ChauffeurNet

Nvidia

Self driving car: Drive OS

Tesla HydraNet



- HydraNet was introduced by Ravi et al. in 2018. It was developed for semantic segmentation, for improving computational efficiency during inference time.
- HydraNets are a type of neural network architecture that can be used for multiple tasks.
- Tesla uses HydraNets in their self-driving cars. They have modified the HydraNet architecture slightly to use a shared backbone. This means that all of the branches share the same early layers of the network.
- Tesla's HydraNet is trained on all of the object's data. It has task-specific heads that allow the model to predict task-specific outputs, such as steering angle and throttle.
- Tesla's HydraNet can also project a birds-eye view of the environment, which gives the car much more dimensionality to navigate properly.

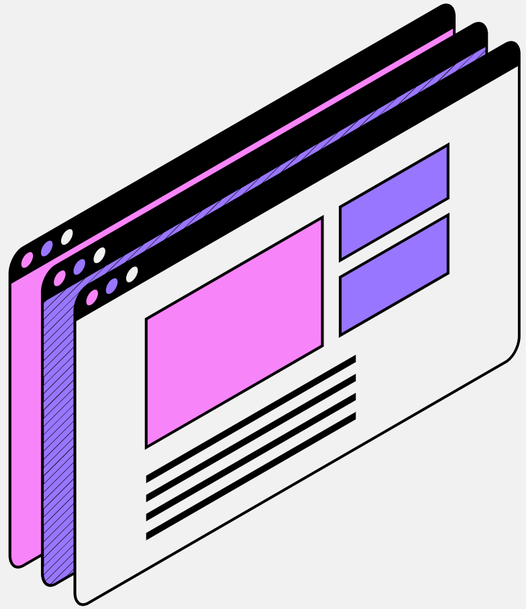
Google Waymo

- ChauffeurNet is a neural network used by Google Waymo to train self-driving cars.
- It uses a convolutional neural network (CNN) to extract features from the perception system's data.
- The CNN features are then fed to a recurrent agent network (AgentRNN), which predicts the car's driving trajectory.
- ChauffeurNet is trained using imitation learning, which means that it learns to drive by imitating expert human drivers.
- ChauffeurNet is trained on both real and synthetic data. The synthetic data includes deviations from normal driving conditions, such as obstacles, unnatural scenes, and perturbations to the trajectory path.
- ChauffeurNet is a mid-level system, which means that it is not part of the end-to-end deep learning algorithm that is used to control the car. Instead, it provides the planning and control system with a driving trajectory.

Google Waymo

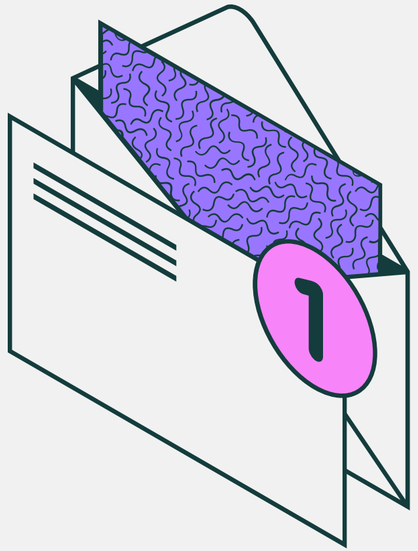
- Nvidia's self-driving car uses a convolutional neural network (CNN) as its primary algorithm.
- Unlike Tesla, Nvidia uses 3 cameras, one on each side and one at the front.
- The CNN is capable of operating on roads that don't have lane markings, including parking lots.
- The CNN learns features and representations that are necessary for detecting useful road features.
- The CNN is an end-to-end system, which means that it optimizes all processing steps at the same time.
- This leads to better performance than systems that decompose the problem into separate steps, such as lane marking detection, path planning, and control.
- The CNN is also smaller than systems that decompose the problem, because it learns to solve the problem with a minimal number of processing steps.

Hands-On Section



Vehicle Detection for Cars, Trucks, and Busses

Thank you!



Linkedin.com/jana-kabrit