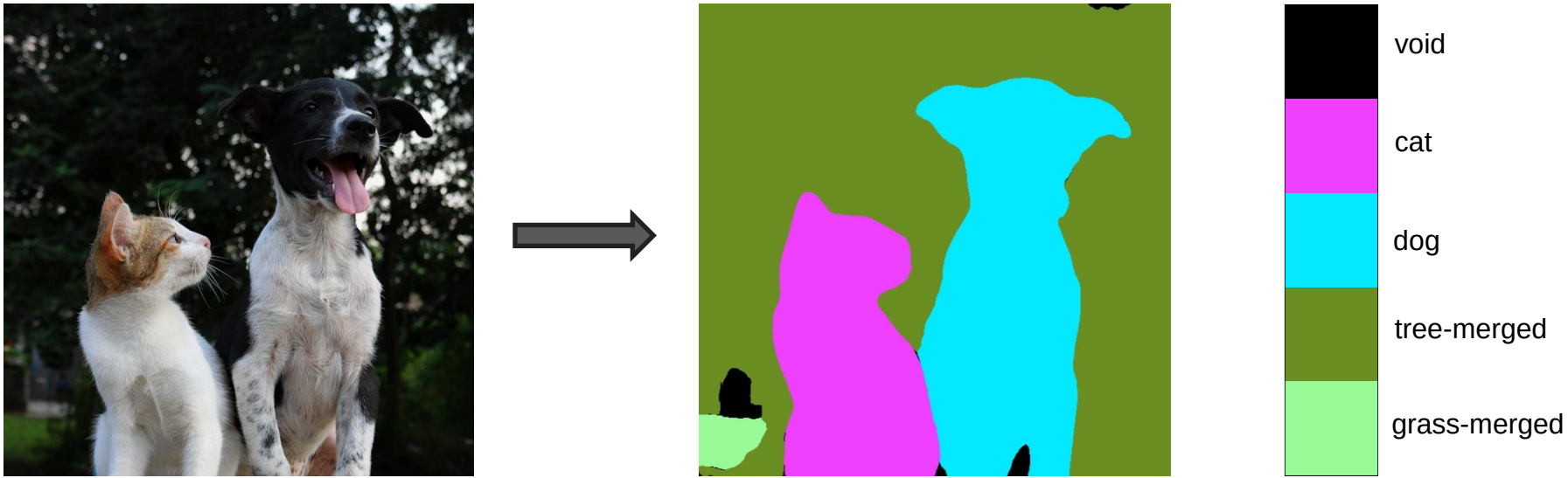
# Introductory seminar on Computer Vision

Sara Concas, Cagliari Digital Lab 2024 - Day 3


sara.concas90c@unica.it

# Semantic segmentation

Task that involves classifying each pixel in an image into a predefined category or class.

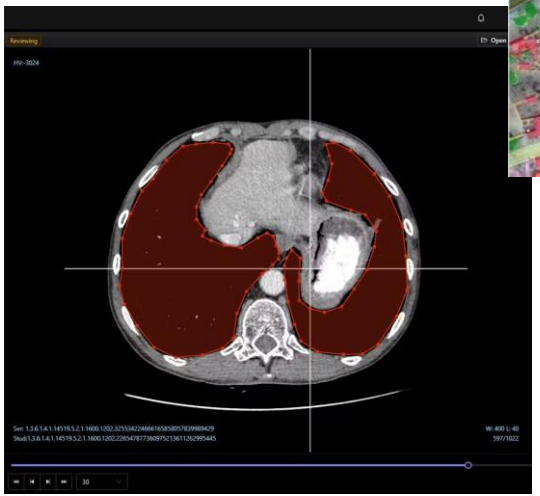



If we have 2 objects of the same category (e.g. 2 cats) the segmentation map does not distinguish these as separate objects.

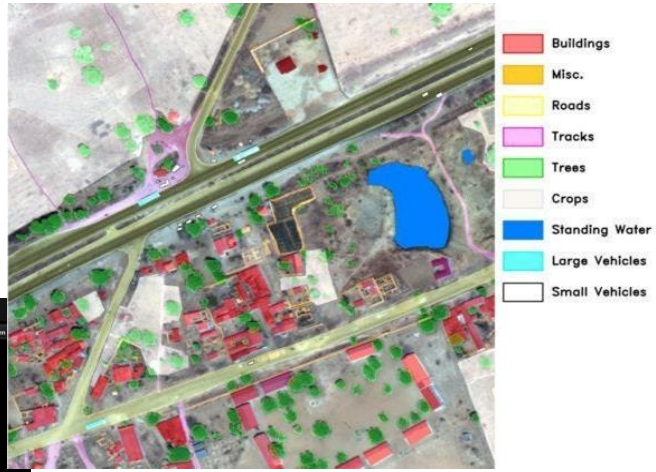
# Semantic segmentation - Applications

Task that involves classifying each pixel in an image into a predefined category or class.

**Autonomous Driving**: identifying road elements like cars, pedestrians, lanes, and traffic signs.



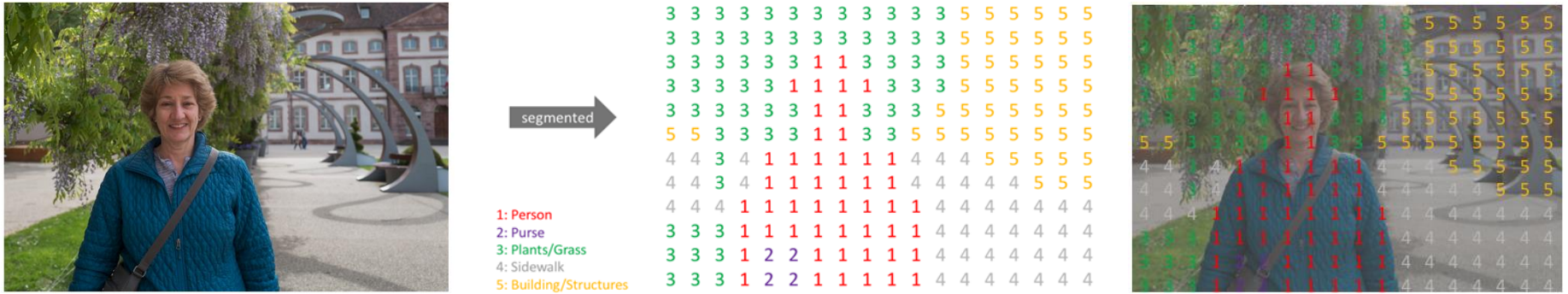
**Medical Imaging**: segmenting different types of tissues, organs, or tumors in medical scans.
https://encord.com/blog/medical-image-segmentation/




**Satellite Imagery**:land cover classification, where different terrains like water, vegetation, and urban areas are segmented.
https://gsitechnology.com/beginners-guide-to-segmentation-in-satellite-images/

# Semantic segmentation - Task

**GOAL**: take a RGB color image or a grayscale image and output a segmentation map where each pixel contains a class label represented as an integer.




https://www.jeremyjordan.me/semantic-segmentation/

# Semantic segmentation - SVM

SVM analyzes the extracted visual features and learns to classify each pixel into different semantic categories

It constructs a hyperplane that maximally separates different classes, enabling accurate pixel labeling.

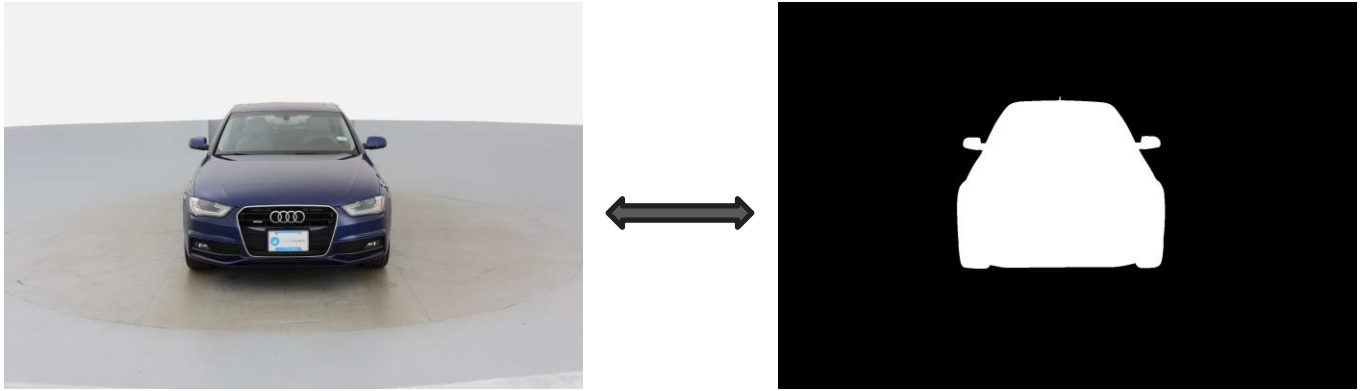
**Good**
**when data is linearly separable**

**But**
**It depends on the quality of the features and on the selection of correct parameters**

# Semantic segmentation – Car Example

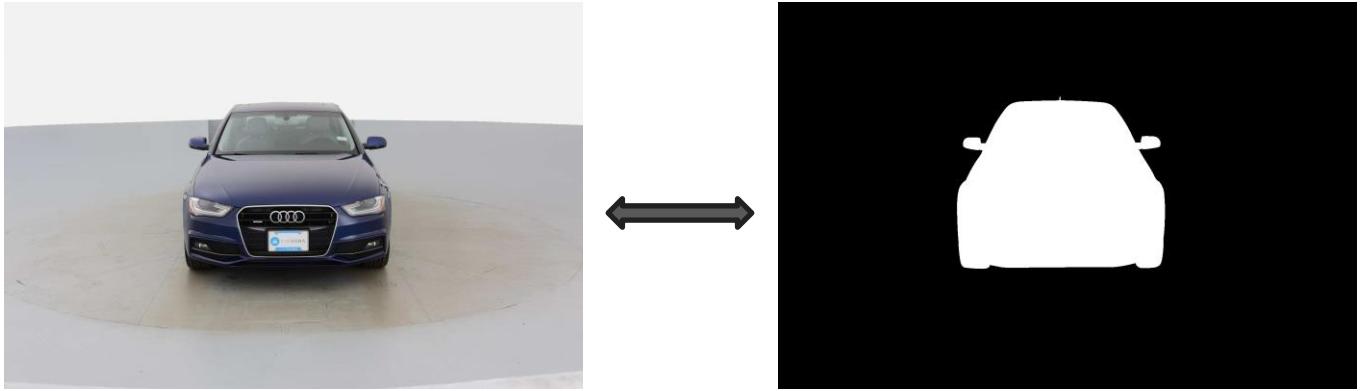
**GOAL**: segment the images from the Carvana Image Masking Dataset using two classes: foreground and background.

The dataset contains a large number of car images, along with a .gif file that contains the manually cutout mask for each image.



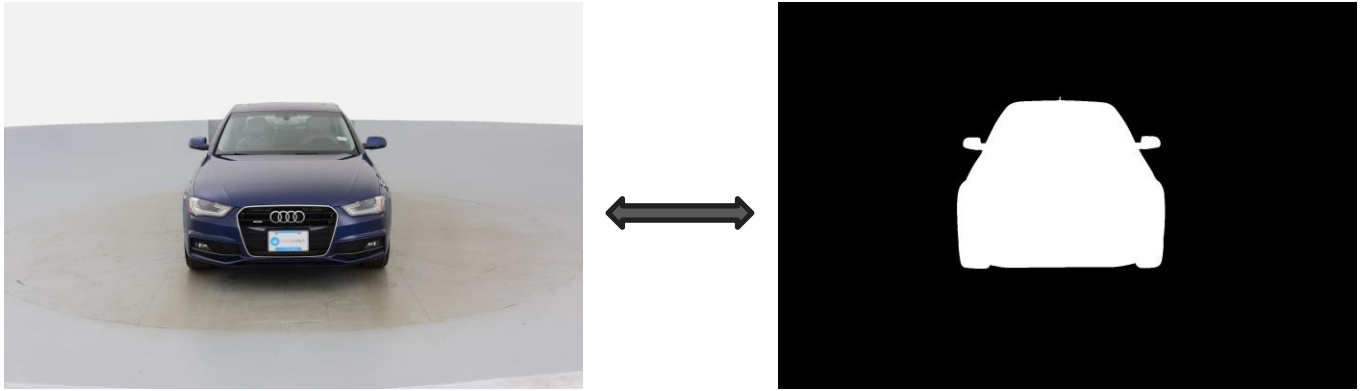
# Semantic segmentation – Car Example

- **Images**: each pixel represents a feature.
- **Masks**: each pixel represents the label of the correspoding pixel in the image.
  - 1 (white) for the foreground (the car in our case)
  - 0 (black) for the background



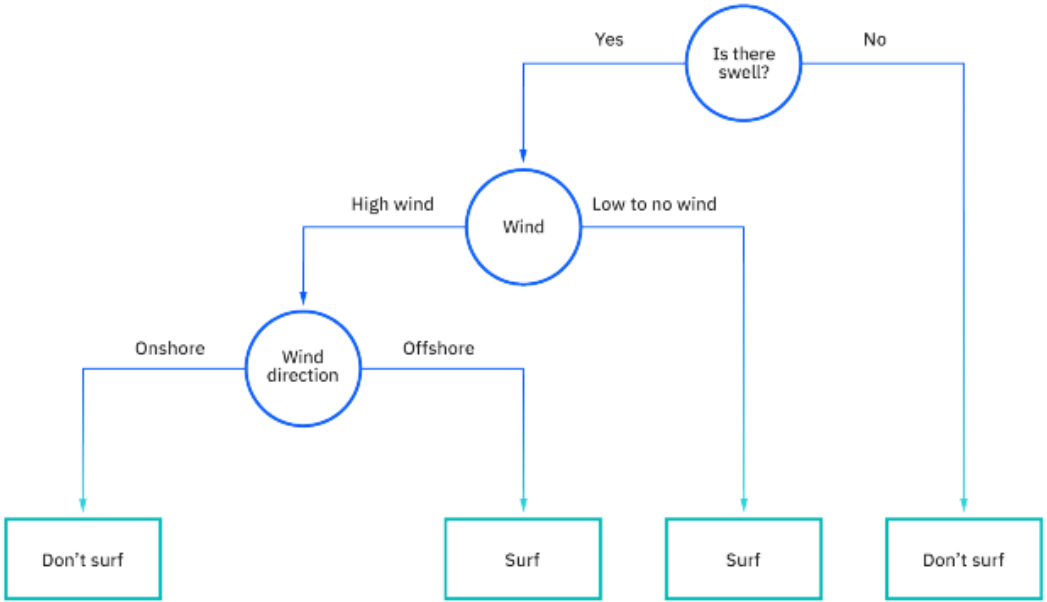
# Semantic segmentation – SVM Car Example

https://bit.ly/3yNquPF



# Semantic segmentation – Random Forests

RF uses a collection of decision trees to classify pixels based on their extracted visual features.




It is a tree-like structure where each internal node represents a decision based on the value of an input feature, each branch represents the outcome of the decision, and each leaf node represents a final output or prediction (such as a class label in classification or a value in regression).

https://www.ibm.com/it-it/topics/decision-trees

# Semantic segmentation – Random Forests

RF uses a collection of decision trees to classify pixels based on their extracted visual features.

Random Forest's added randomness minimizes overfitting and enhances the model's capacity for generalization.

**Good**
**Robust against overfitting**

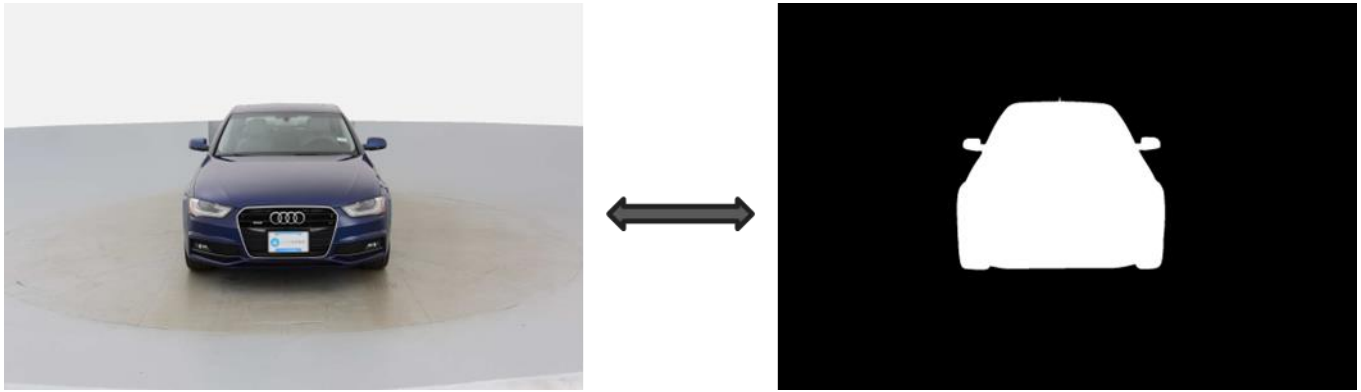
**But**
**Difficulty handling complex structures**
**May lose fine-grained details**

https://www.ibm.com/topics/random-forest

# Semantic segmentation – Random Forests Car Example

GOAL: segment the images from the Carvana Image Masking Dataset using two classes: foreground and background.
We will use the same dataset and compare the result.

https://bit.ly/3YYOABX

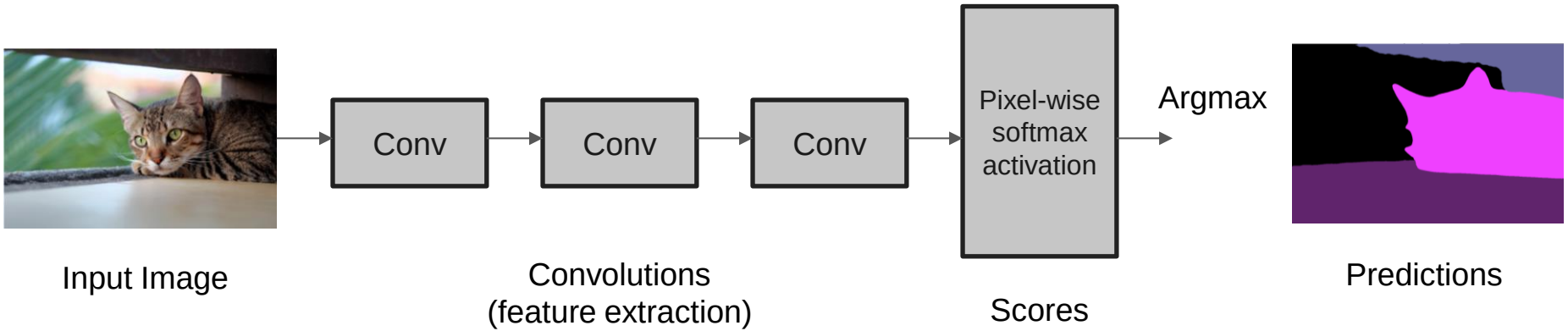


# Semantic segmentation – Neural Networks

Neural networks, particularly convolutional neural networks (CNNs), have been at the forefront of advancements in semantic segmentation.

DL methods have gradually replaced traditional methods like SVMs and Random Forest. Although they require more data, time and computational resources, they are able to outperform other methods.

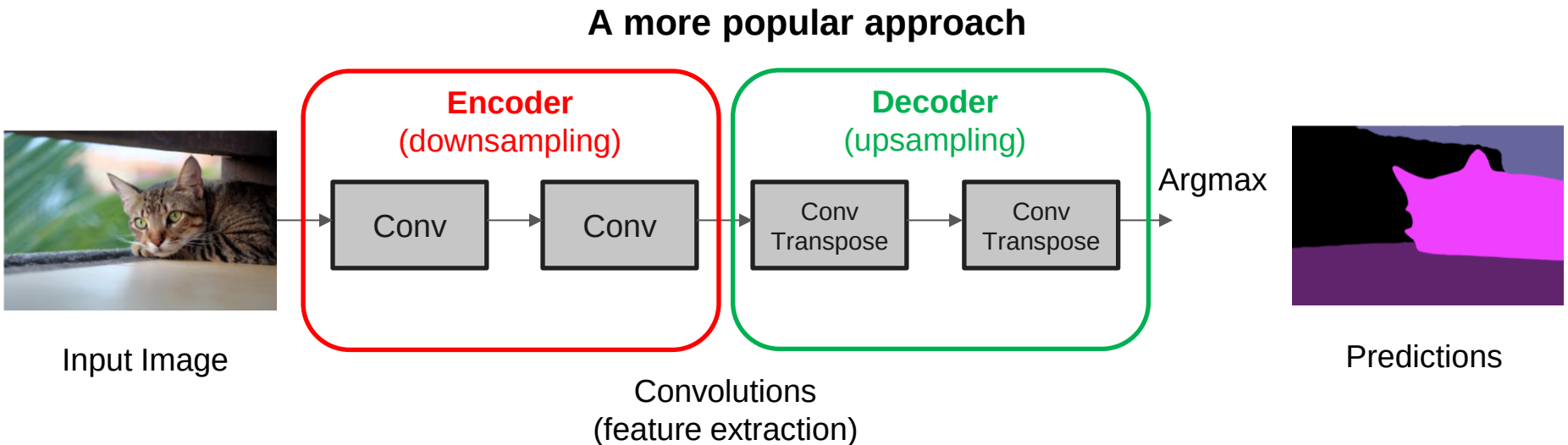
**A very simple example of Architecture**



# Semantic segmentation – Neural Networks

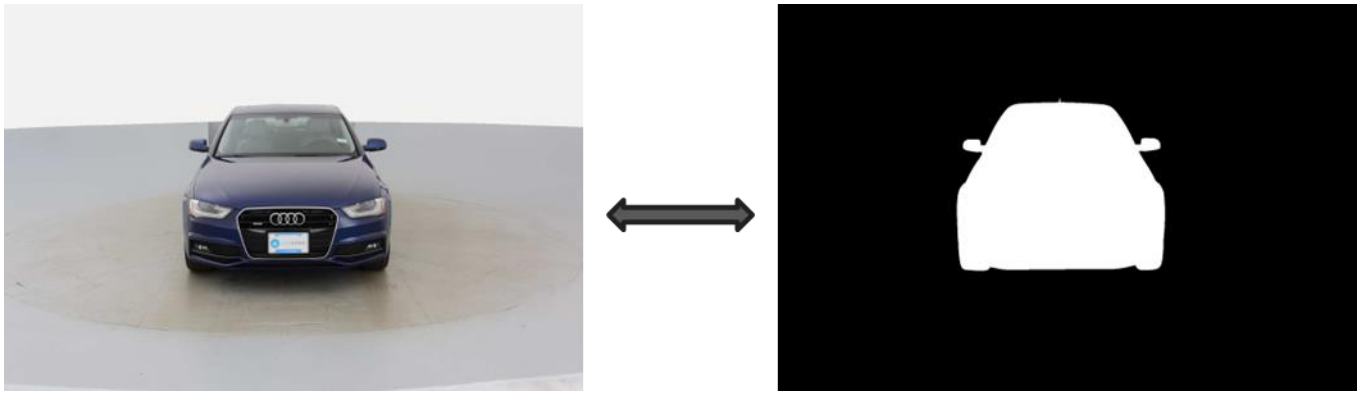
Neural networks, particularly convolutional neural networks (CNNs), have been at the forefront of advancements in semantic segmentation.

DL methods have gradually replaced traditional methods like SVMs and Random Forest. Although they require more data, time and computational resources, they are able to outperform other methods.



# Semantic segmentation – Neural Networks  Car Example

https://bit.ly/4fUW0fu



# Semantic Segmentation – Car Example

Let’s compare the results!

Original Image

Ground Truth mask

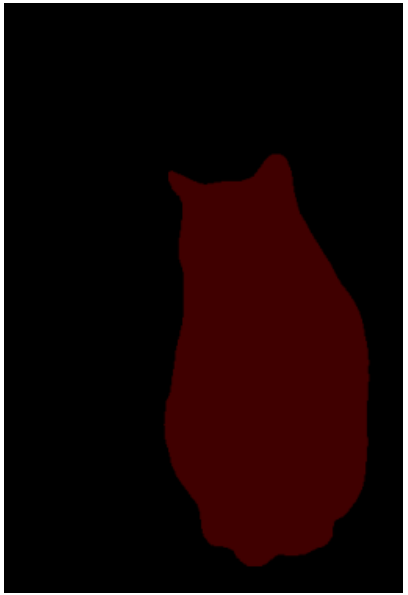
SVM Prediction

RF Prediction

NN Prediction

# Semantic Segmentation – DeepLab

DeepLab is a state-of-art deep learning model for semantic image segmentation, where the goal is to assign semantic labels (e.g., person, dog, cat and so on) to every pixel in the input image.



We will use **DeepLabv3** implementation, pretrained on the PASCAL Visual Object Classes Dataset (PASCAL VOC) . This is a well-known object detection, segmentation, and classification dataset.

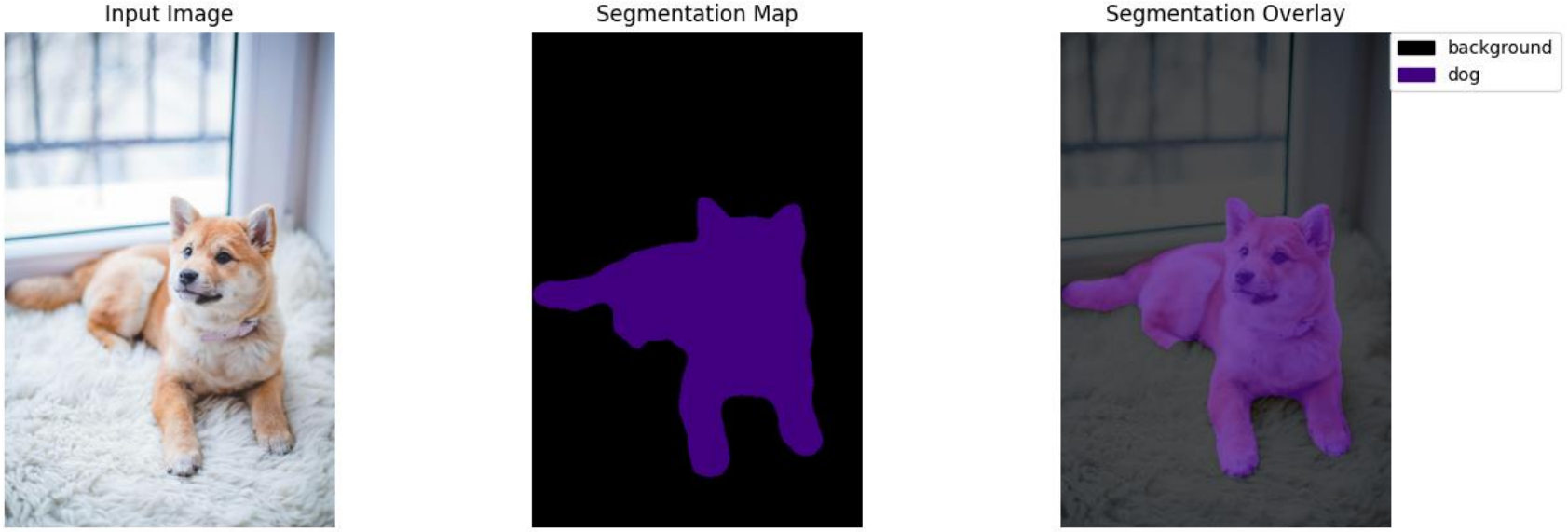
It includes 20 object categories, including common objects like cars, bicycles, and animals, as well as more specific categories such as boats, sofas, and dining tables.

Annotations include object bounding boxes and class labels for object detection and classification tasks, and segmentation masks for the segmentation tasks.

https://github.com/tensorflow/models/tree/master/research/deeplab

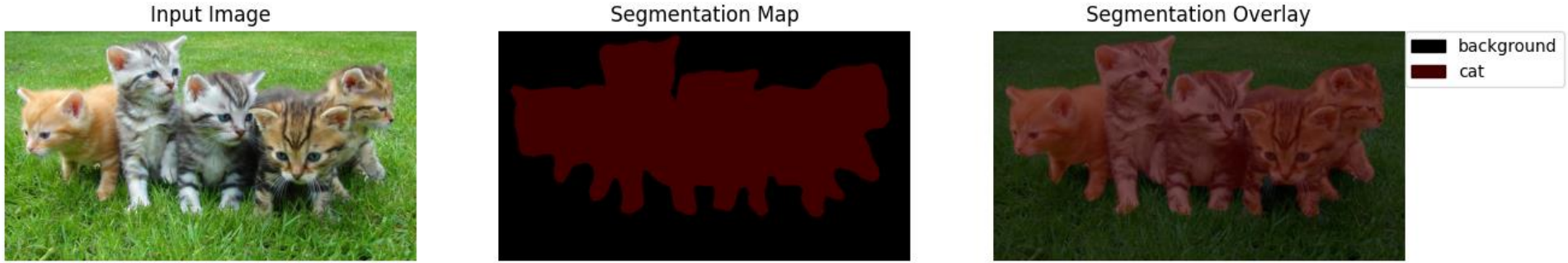
# Semantic Segmentation – DeepLab Example

https://bit.ly/3MeW243



# Semantic, Instance and Panoptic Segmentation

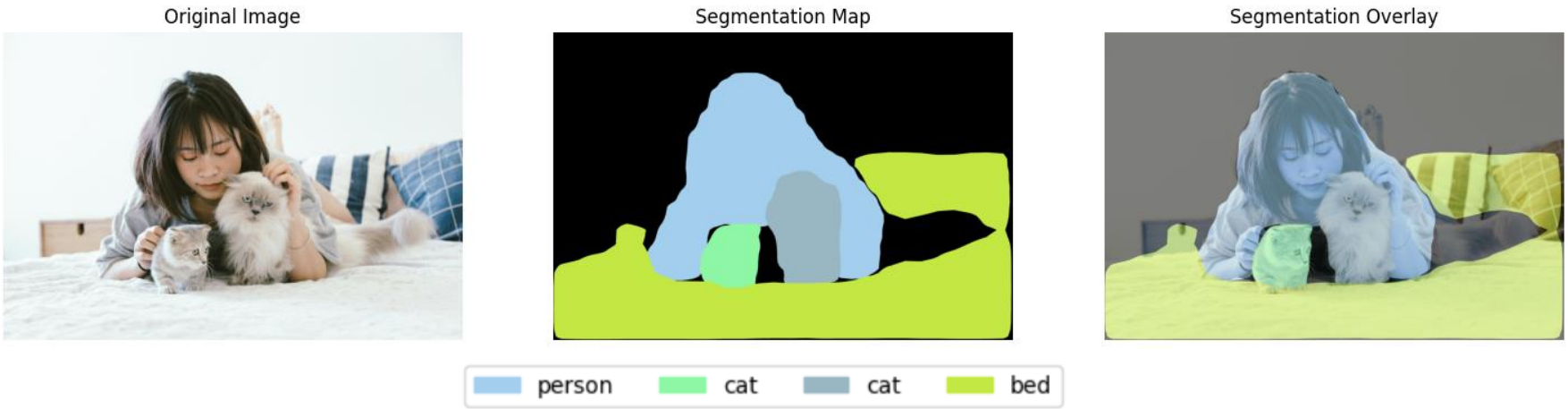
- **Semantic Segmentation**: involves classifying each pixel in an image into a predefined set of categories or classes. The goal is to label every pixel with a class, but not to distinguish between different instances of the same class.




It groups together all pixels that belong to the same class

# Semantic, Instance and Panoptic Segmentation

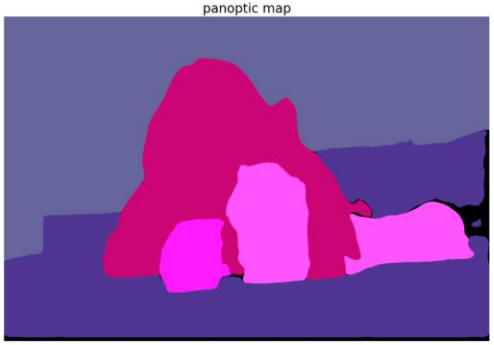
- **Instance Segmentation:** not only classifies each pixel but also distinguishes between different instances of the same class. This means that it identifies and separates individual objects within the same category.



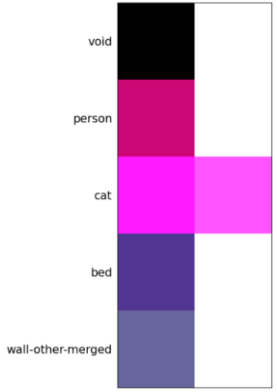

Each instance of an object gets a unique label (in this example, in fact, we see a different color for the two cats)

# Semantic, Instance and Panoptic Segmentation

- **Panoptic Segmentation:** combines both semantic and instance segmentation into a unified framework. It provides pixel-level classification for all pixels (like semantic segmentation) and distinguishes between different instances of objects (like instance segmentation).



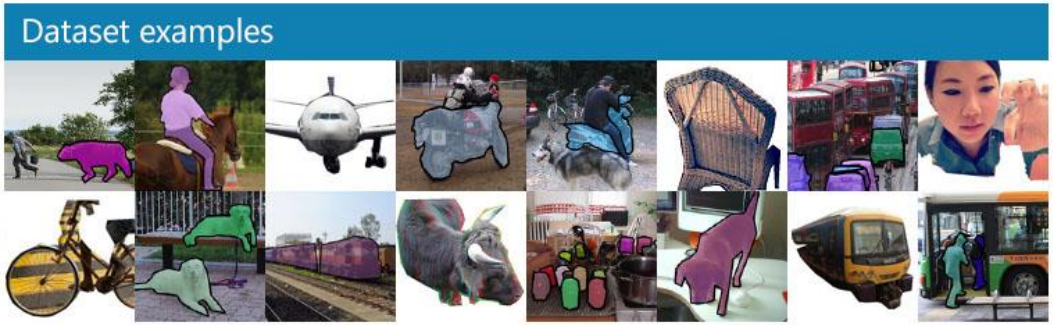



It assigns every pixel in the image a unique label but also considers its class.

# Instance and Panoptic Segmentation - Example

We will use the open-source library Detectron2, developed by Facebook AI Research (FAIR) for object detection and instance segmentation tasks.
It is a high-level framework built on top of PyTorch, designed to make it easy for developers and researchers to create, train, and deploy computer vision models.

We will use a model trained on the Common Objects in Context (COCO) dataset, a large-scale dataset designed for various computer vision tasks such as object detection, segmentation, and captioning. Developed by Microsoft, it is widely used in research and competitions due to its challenging and diverse set of images, rich annotations, and large size.

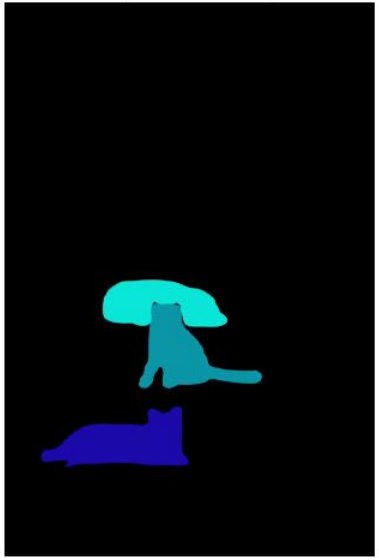


# Instance Segmentation - Example

https://bit.ly/3XduRwN

Original Image

Segmentation Map



Segmentation Overlay

# Panoptic Segmentation - Example

https://bit.ly/3SZ9L2X

Original Image

Panoptic Map