CV / VLM

Unit 2: Introduction to Object Detection (OD)

2.2.3

Region-Based Object Detection

Introduction to R-CNN Families

TIL-AI
TODAY I LEARNED AI

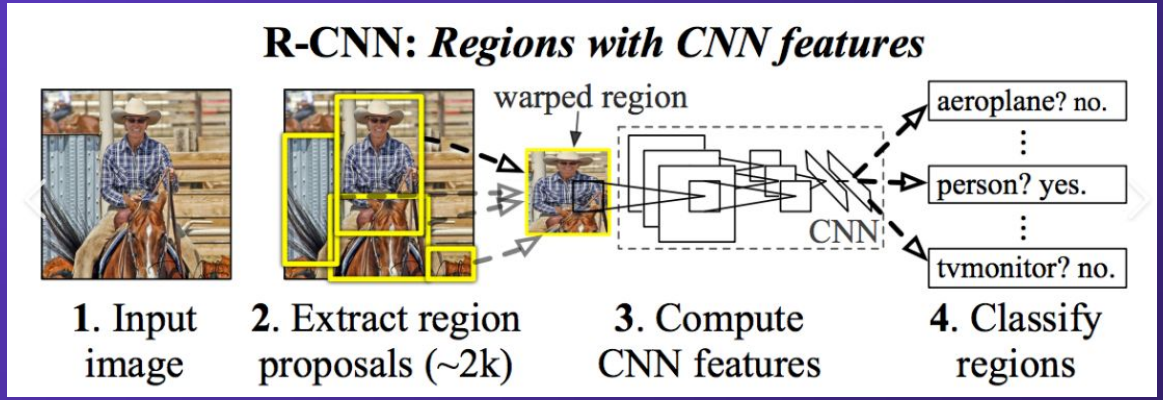
Intro to Region-CNN (R-CNN)

The process of R-CNN detection is as follows:

1. Generate **bottom-up** region proposals (selective search)

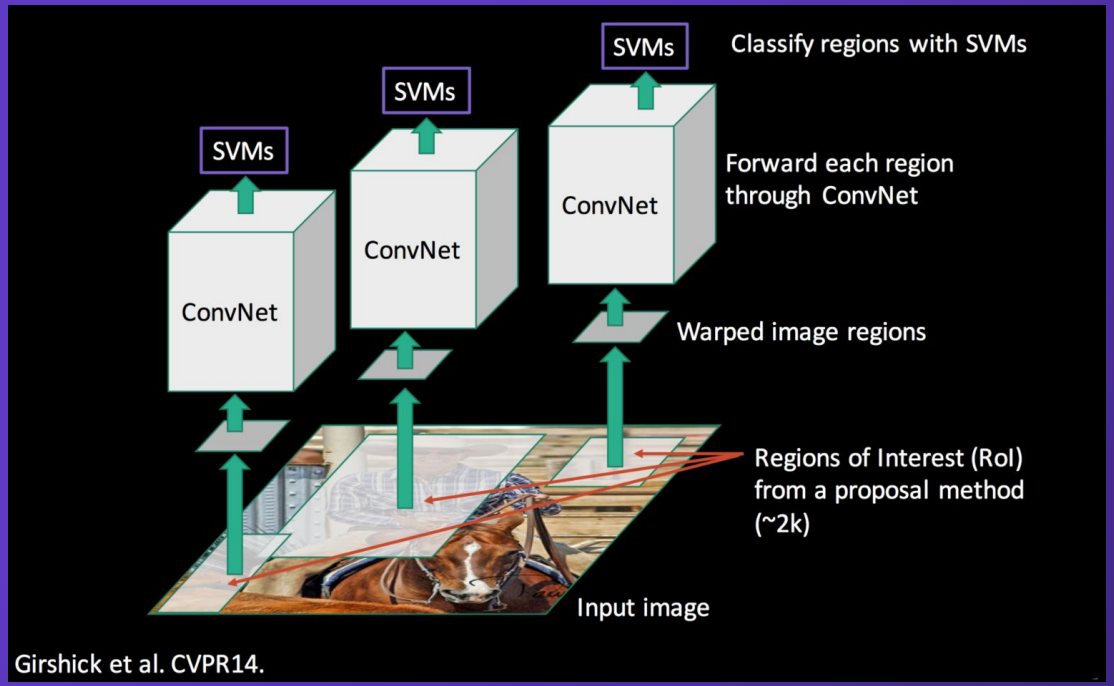
2. **Warping & Feature Extraction:** Extract features for each proposal using CNNs

3. **Classification:** SVM to segregating the objects into classes



A Survey on Deep Learning Based Approaches for Scene Understanding in Autonomous Driving Zhiyang Guo- 2021/4

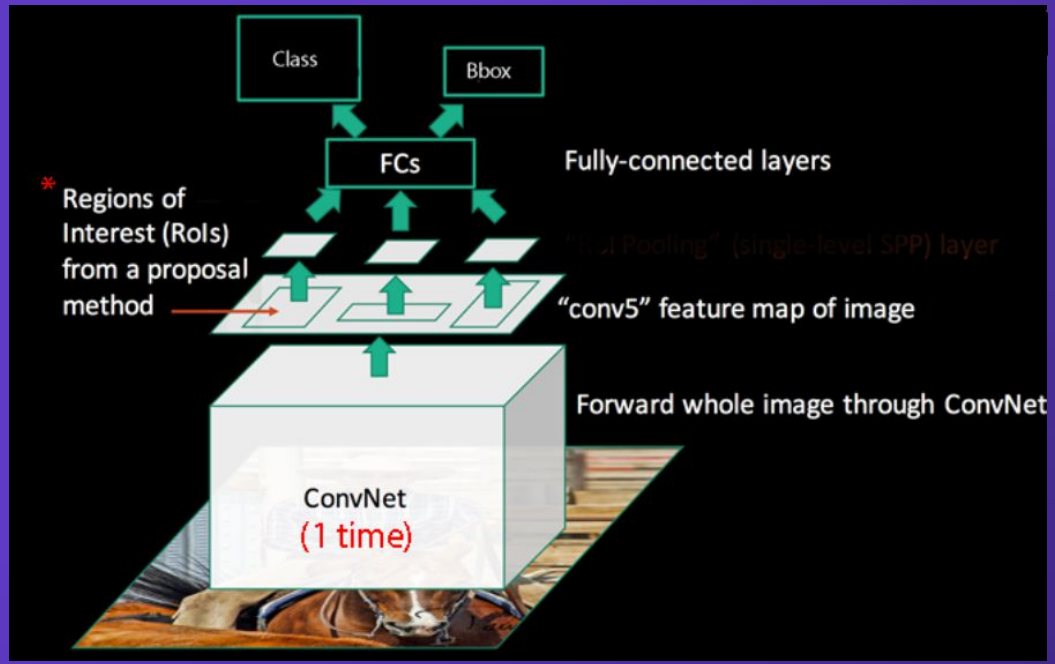
R-CNN - Technique Overview



Steps

1. Region of interest proposal
2. Warped Image Regions
3. Individual CNN per region
4. Classifying regions with SVM

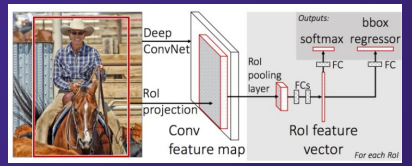
Fast R-CNN - Technique Overview



Steps

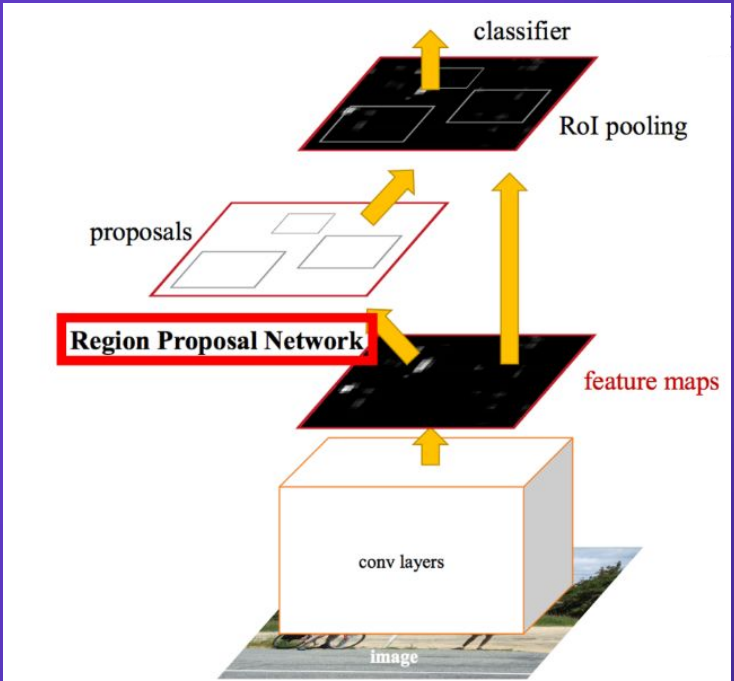
1. CNN to extract feature maps

2. Regions of interest (RoI) on feature maps with max pooling
3. Fully connected layers



4. Softmax classification and bounding box regression

Faster R-CNN - Technique Overview



Steps

1. CNN to extract feature maps.

2. **RPN (RoI Pooling)**: Determine regions of interest from the feature maps with a seperate CNN.

3. Fully connected layers

4. Classification with softmax and bounding box regression

Summary of R-CNN Families

Complexity →

Feature	R-CNN	Fast R-CNN	Faster R-CNN
Region Proposal	External Algorithm (e.g. Selective Search)	Region of Interest Pooling (Maximum Pooling)	Region Proposal Network (RoI Pooling)
CNN Feature Extraction	Independent for image & ROIs	Shared for entire image	Shared for entire image
Localization	Bounding Box	Bounding Box	Bounding Box
Accuracy	Moderate	Similar to R-CNN	Better
Advantages	Pretrained-CNN features for object detection.	Faster than R-CNN with Pooling	Fastest, potentially more accurate ROI proposals
Disadvantages	Slow due to multiple CNN passes	Relies on external algorithm for ROI proposals	More complex architecture

TIL-AI
TODAY I LEARNED AI