Application of deep neural networks to

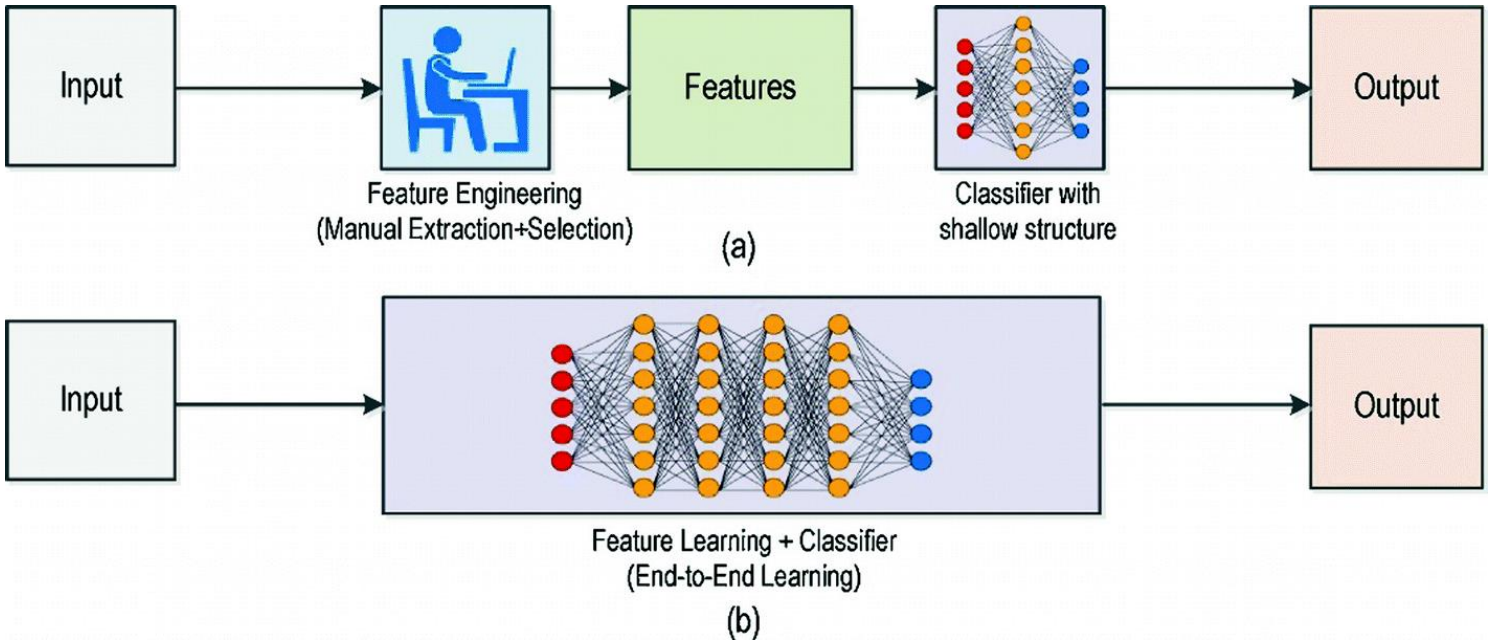
Computer Vision

dayche.com . گروه دایکه

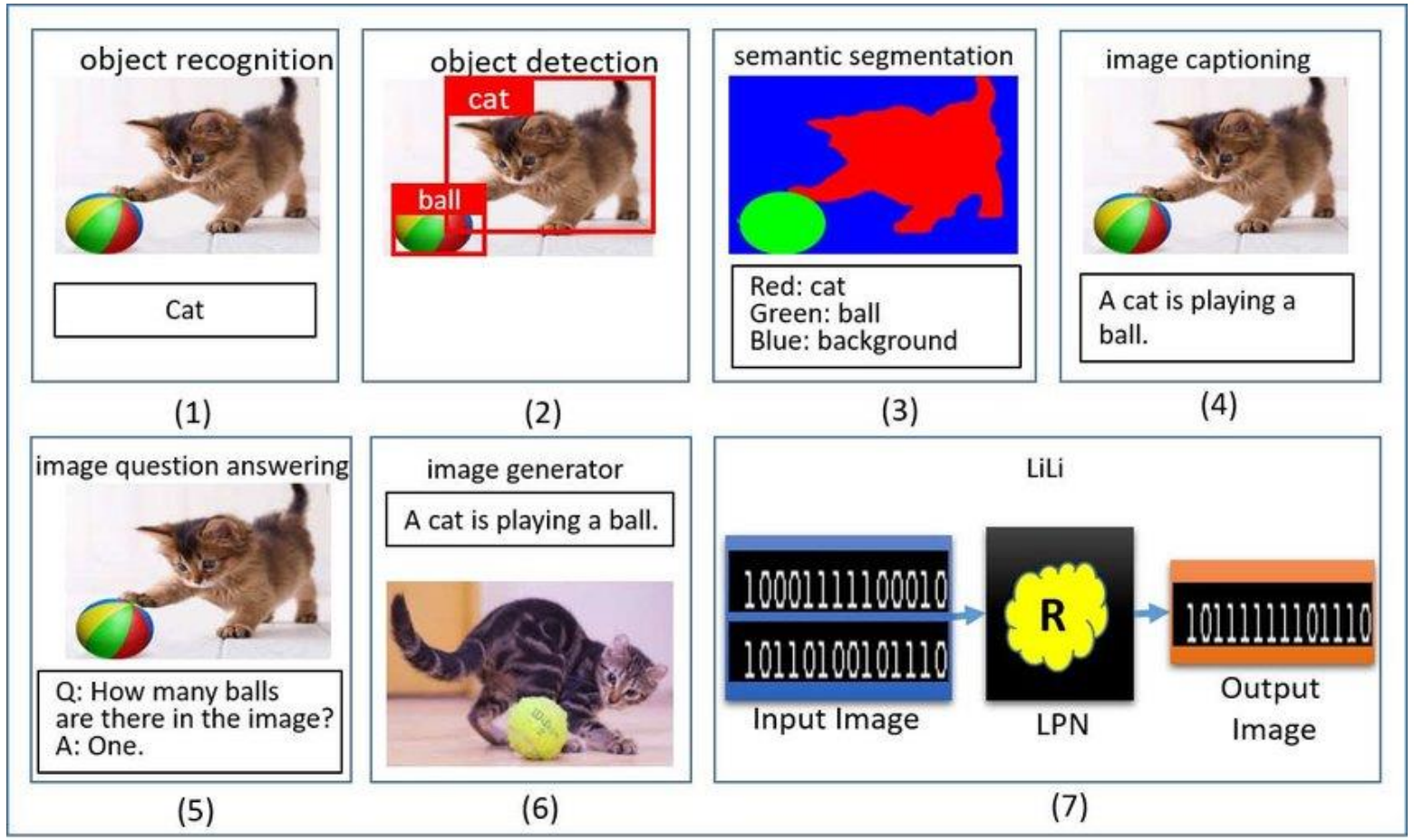
دای‌کِه DAYCHE

Computer vision

- We human beings get information from different sources among which eyes are so important
  - In some applications the machinery tools need to be equipped to eye.



Tasks in computer vision

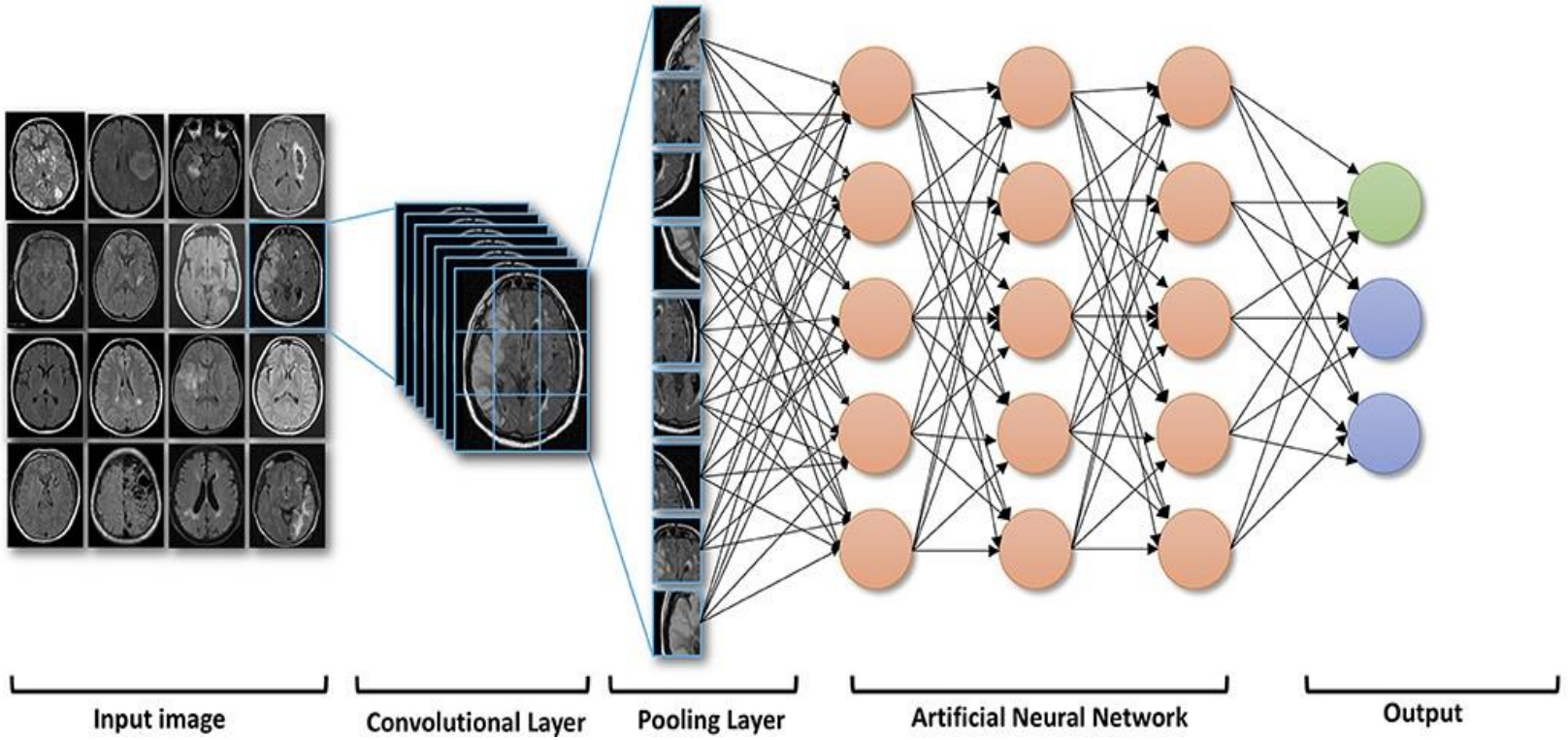


تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
گروه دایکه | dayche.com

OCR

تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

Medical diagnosis



تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

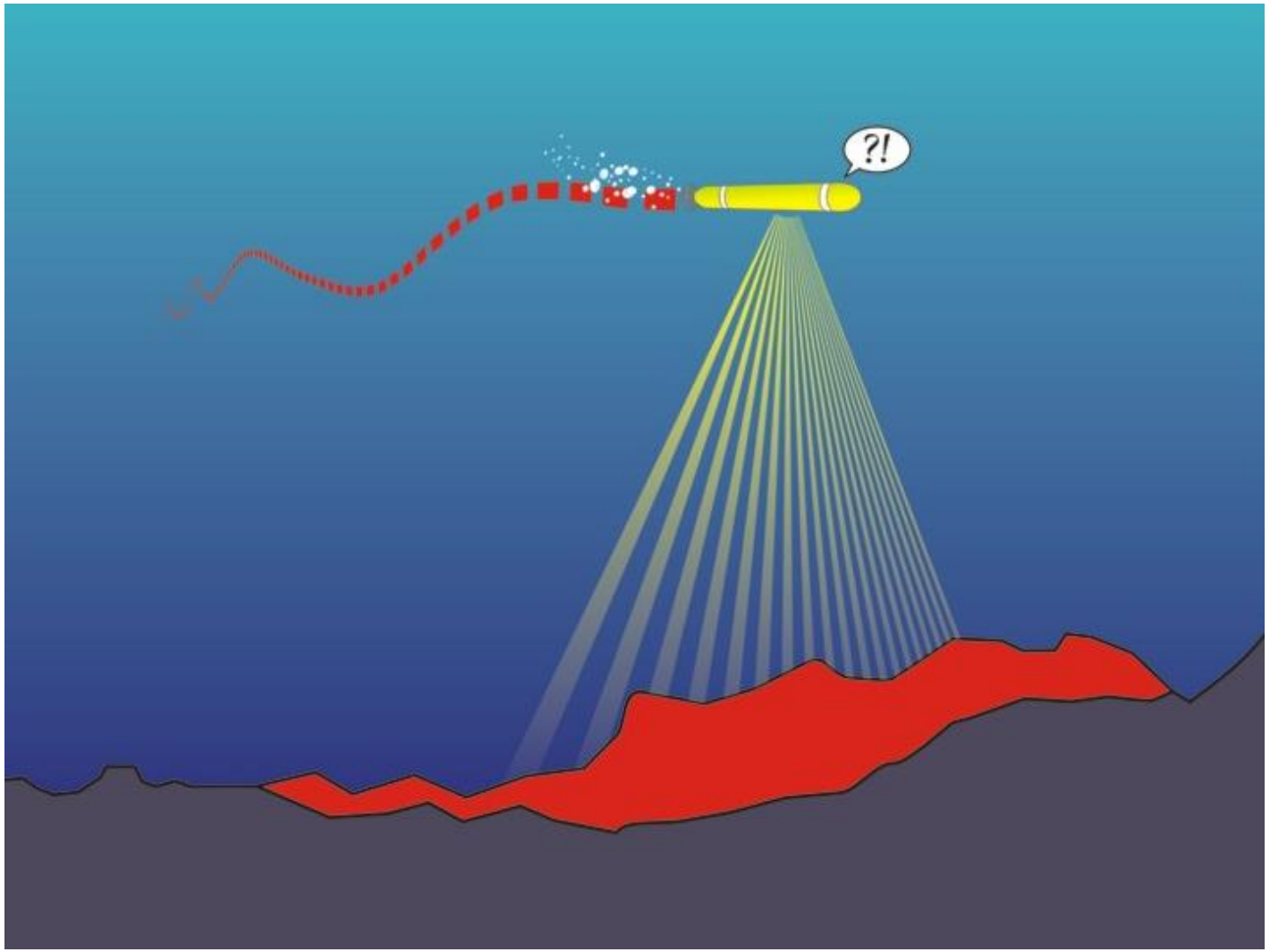
Advanced Driver-assistance Systems

تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

Image based surveillance systems

تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

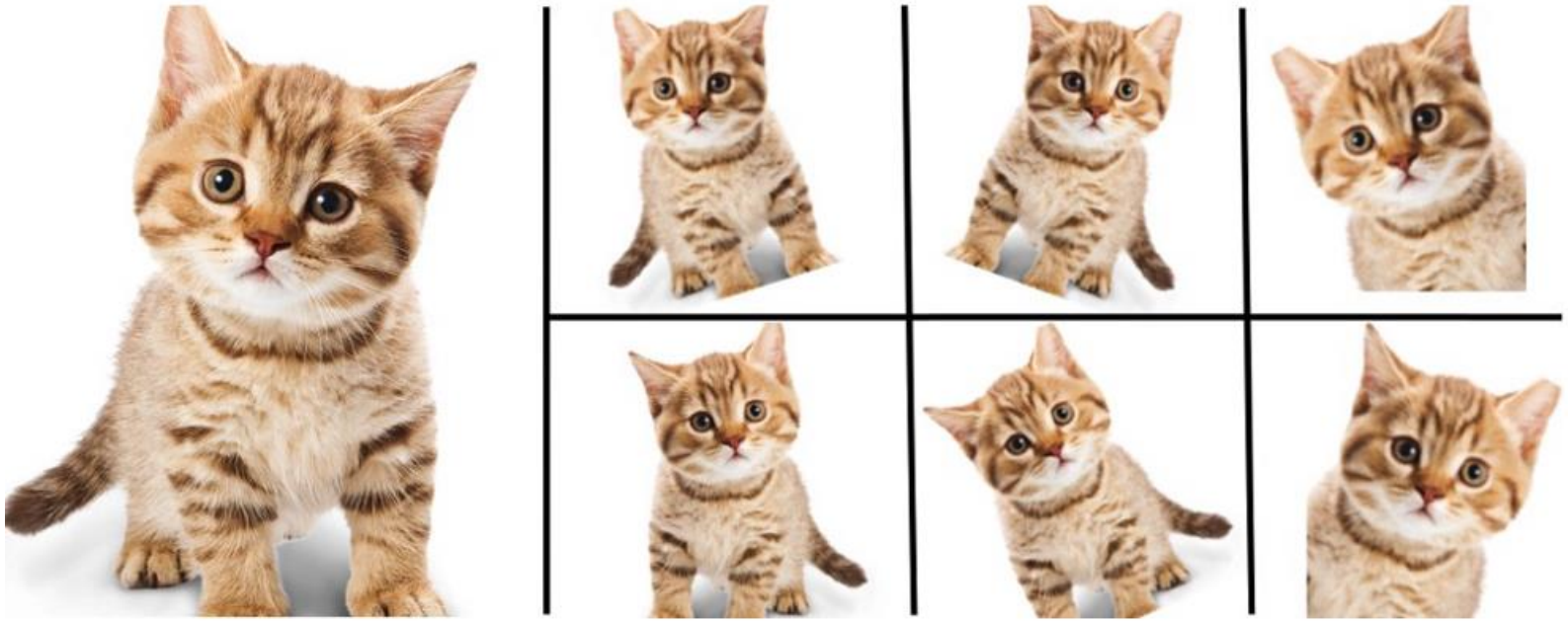
Navigation system (DSMAC)



تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

Image augmentation

- Deep models are DATA HUNGRY! How to acquire a large dataset?
  - Image augmentation – random transformation on images make them to be enlarged in size.
  - AlexNet augment 3 million images to obtain 14 million images

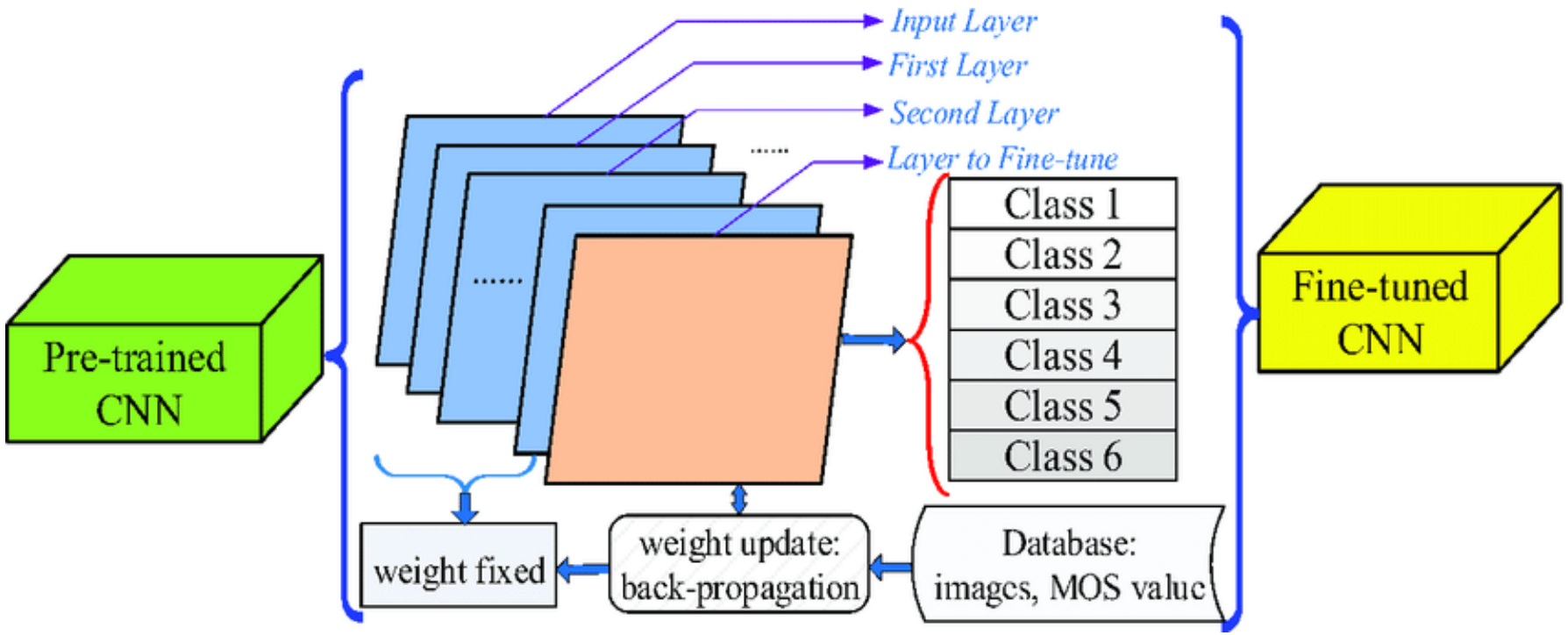


It will improve the generalization of model, how?

- Rotation may help the model to produce features which are not location dependent
- Color transformation mat improve the sensitivity to light.
- Can we always use this approach?

تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
گروه دایکه | dayche.com

Fine tuning



تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

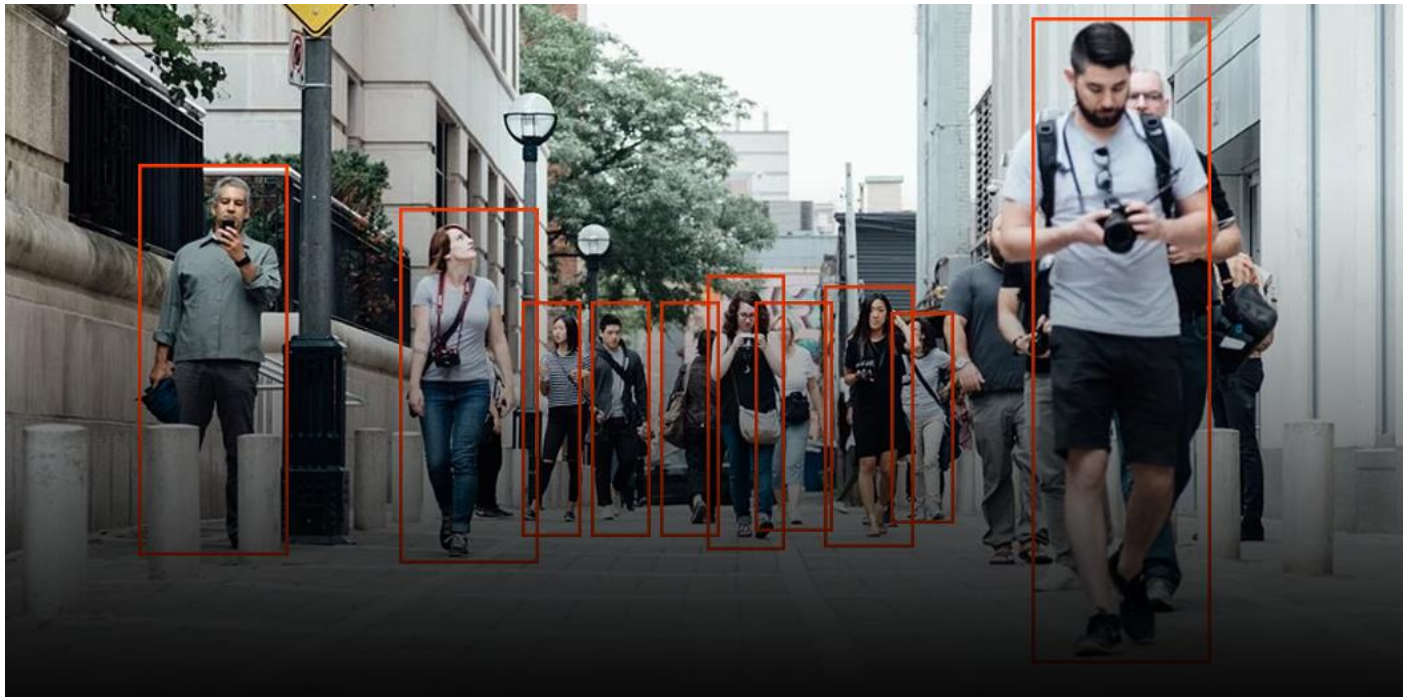
Object recognition

- Object recognition
  - A computer vision task at which we are seeking to answer the question of whether there is an object of specific type in an image.
  - We develop a model whose input is an image and output is a vector of probability measure.

ImageNet dataset

Object detection

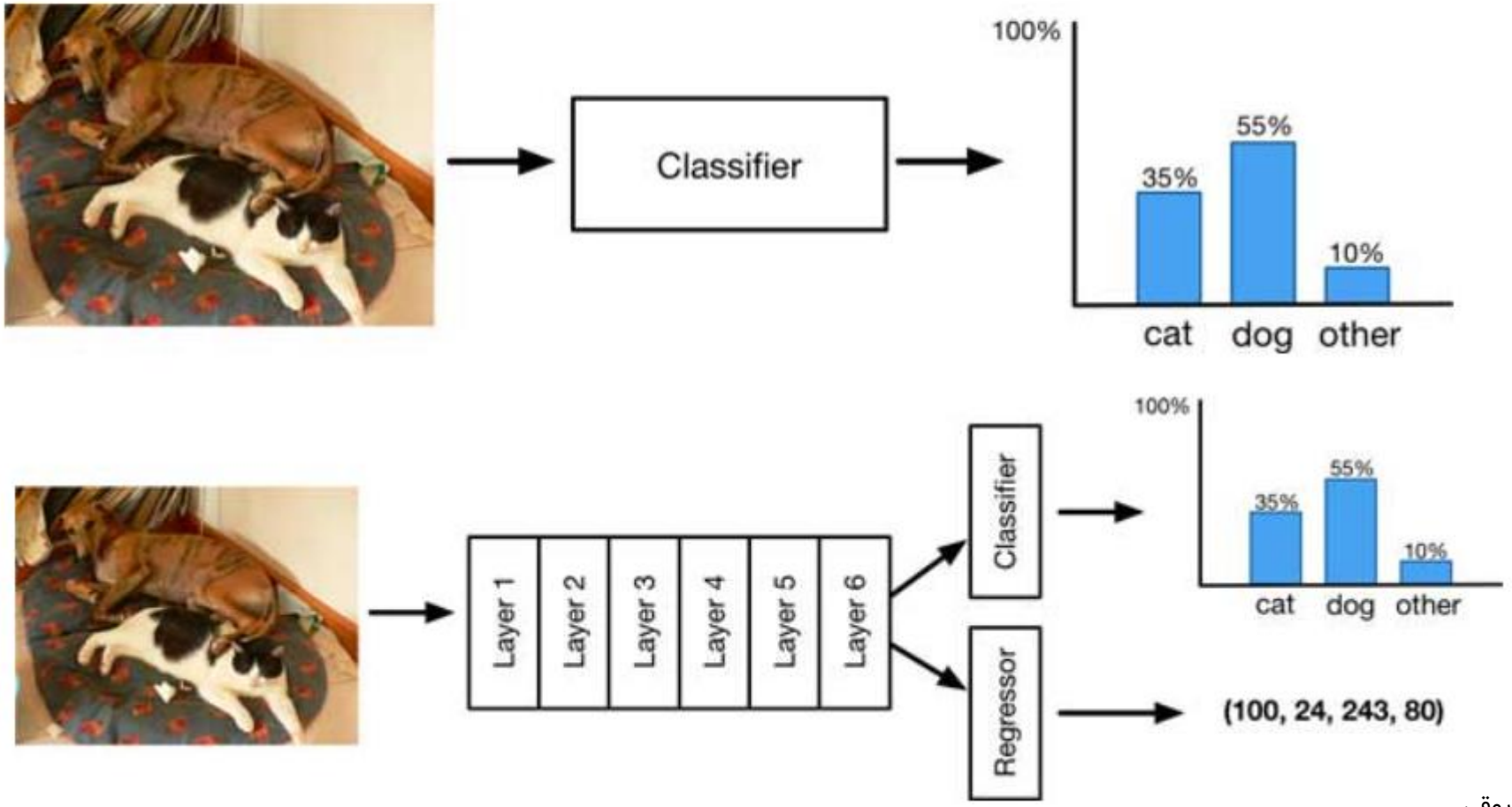
- Object detection
  - An object recognition task which requires to localize objects as well.



How to develop such a model?

- The idea is simply object recognition, yet how to localize the objects?
- We should feed the region which is likely to be host of object and feed it as network input, but how?
  - Is it the best solution? Think of inference time.

Object recognition vs object detection



تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

Intuition

What is available in real-world? How to find the proposal region?

What model need to classify

تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
گروه دایکه | dayche.com

Exhaustive search

تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

Exhaustive search

تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

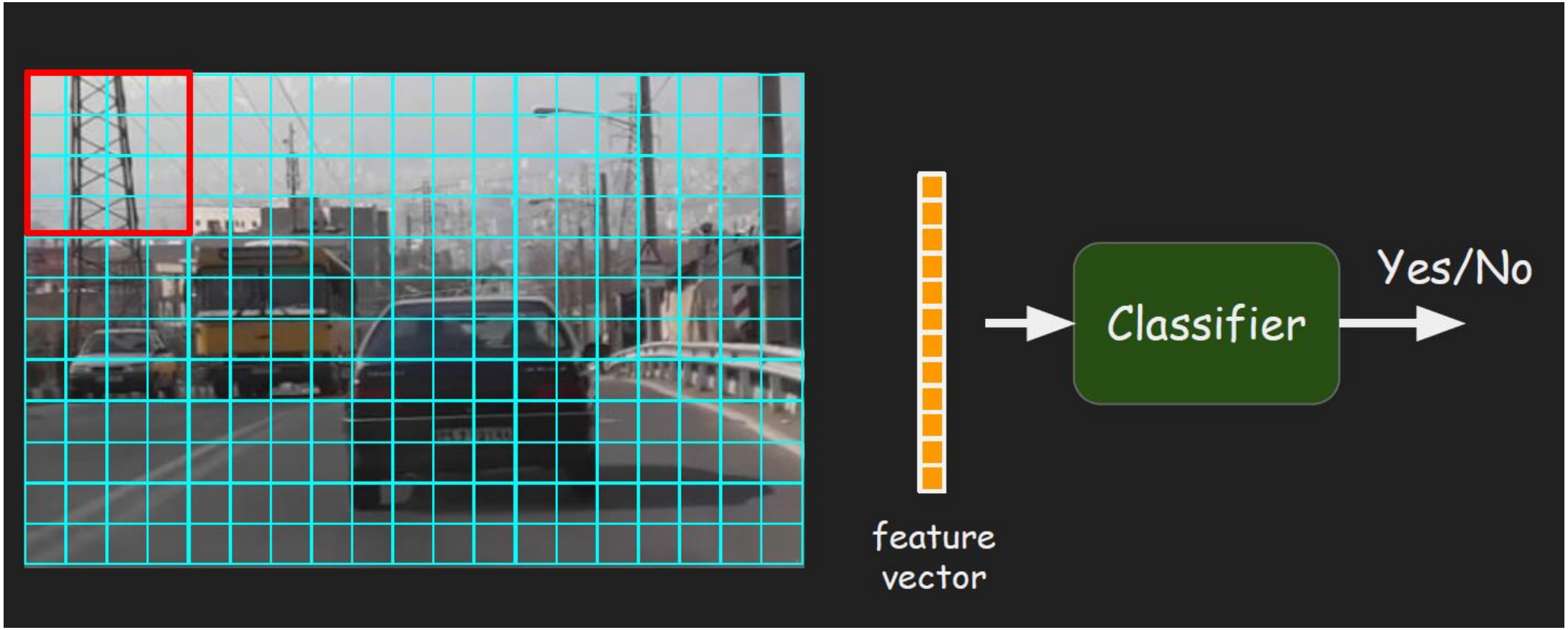
Exhaustive search

تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

Exhaustive search

- This was one trial! We need to scan the image with different size windows since the objects are of different size
- Does it make sense? How to can do the same in a smart way? (Anchor box)

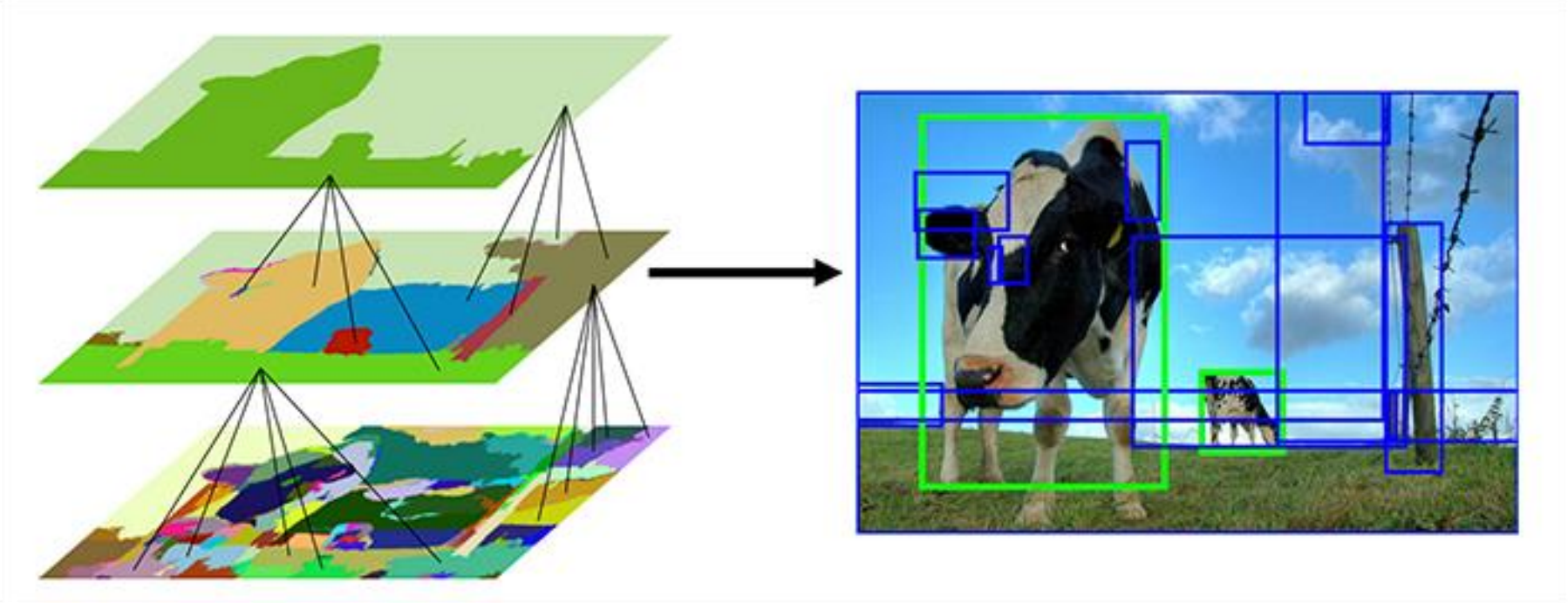
Sliding on feature maps



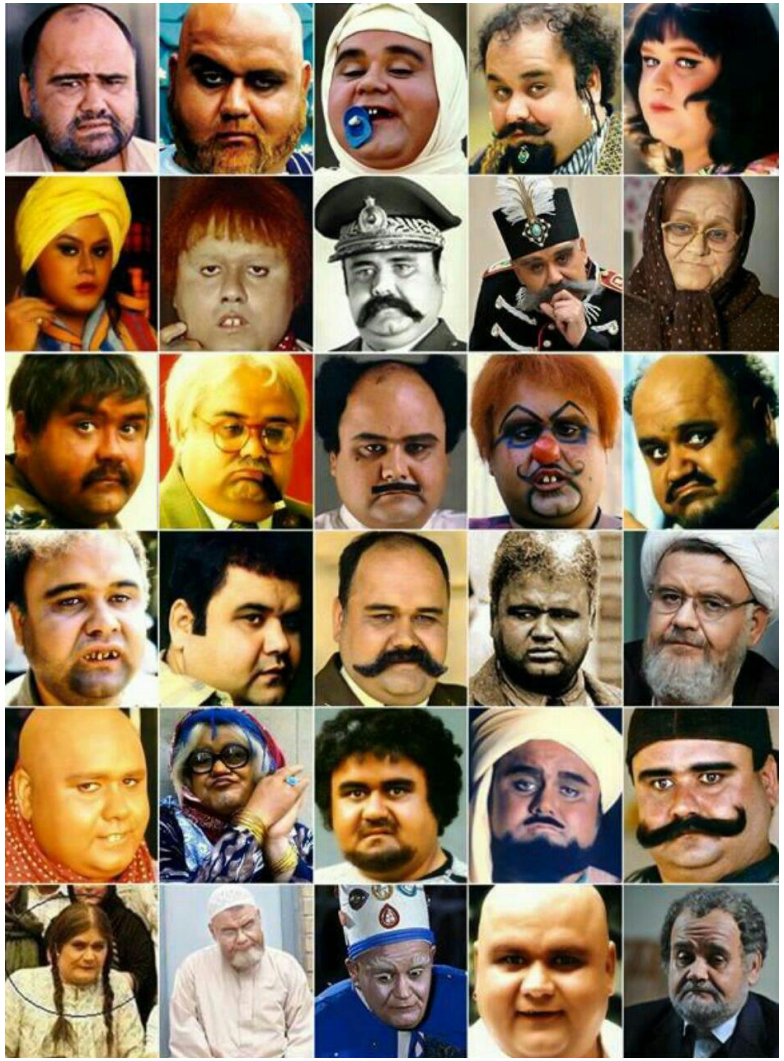
تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
گروه دایکه | dayche.com

Region proposal – Selective search

Combination of graph-based segmentation and exhaustive search



Is the image itself the input of classifier?



- Suppose the region proposal algorithm output this frames.
- All of this images belong to the same person.
- As machine learning algorithms assign weights to given inputs and process it to construct the output, they will be failed if such raw frames fed as their inputs? Why? There is no stationarity.
- We need to capture features which are common in all of frames

- Classical computer vision algorithms (edge detectors, HOG, …)
- Convolutional neural networks

تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
گروه دایکه | dayche.com

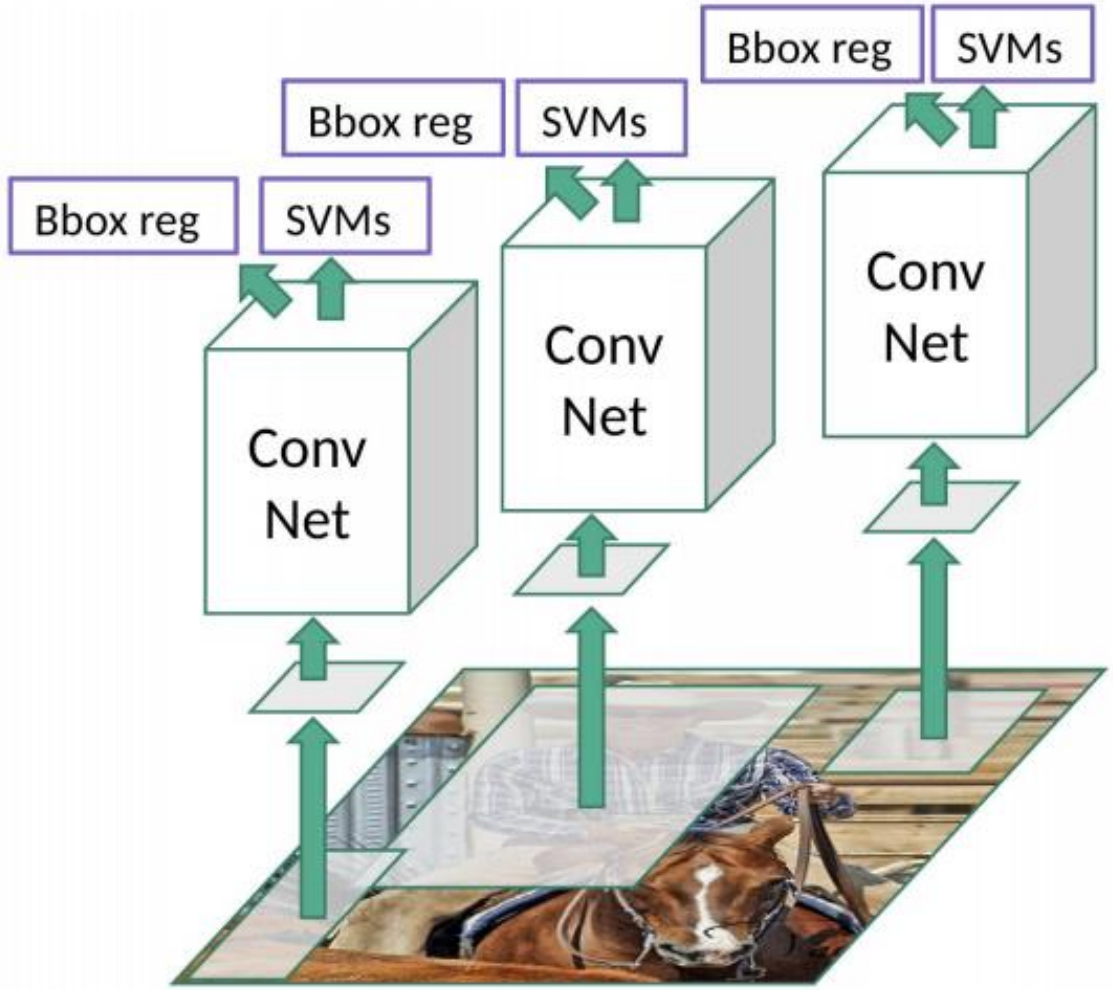
Region-based Convolutional neural network (R-CNN)

- Collect data as more as you can
- Labeling manually

- Selective search to locate objects
- Feed each RoI to a detector network (whichever pre-trained networks after fine-tuning on custom dataset)
- Train a SVM classifier
- Train a regression model for object coordinate prediction

- They are not efficient since their inference phase take 47 seconds.
- Multi-stage training
- Require storage in test phase

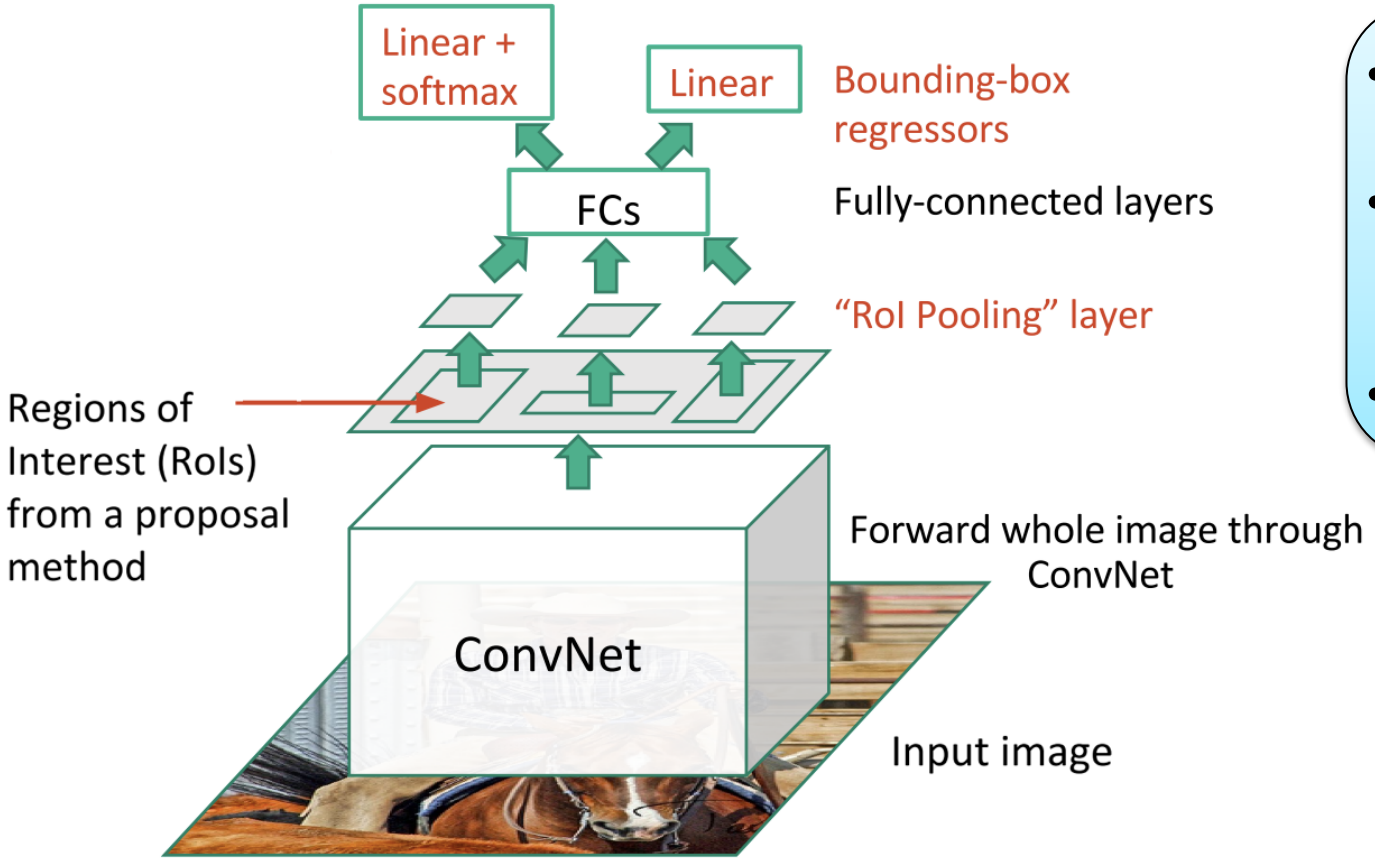
R-CNN architecture



Training pipelines:

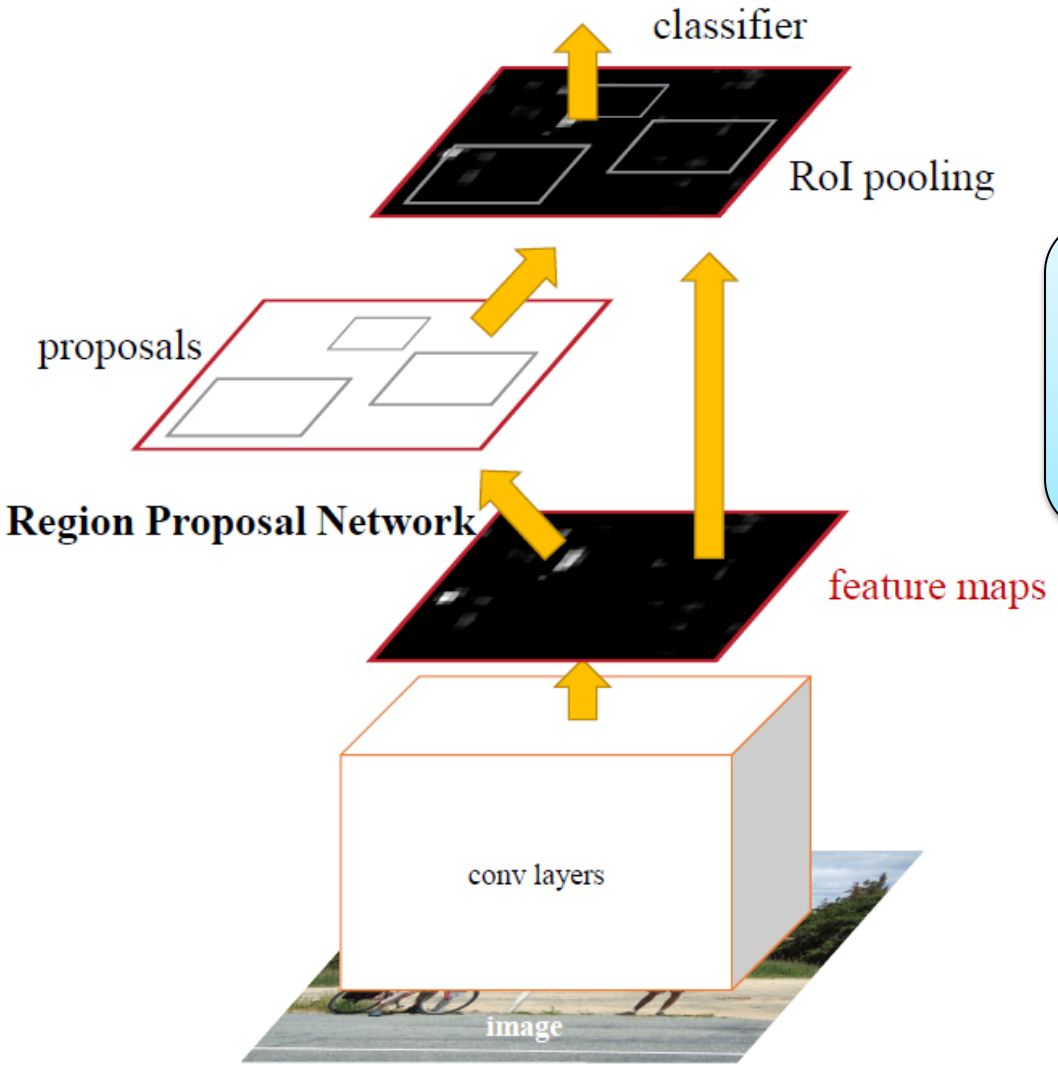
- Fine-tuning the detection network
- Training SVM classifiers
- Training regression network

Fast R-CNN



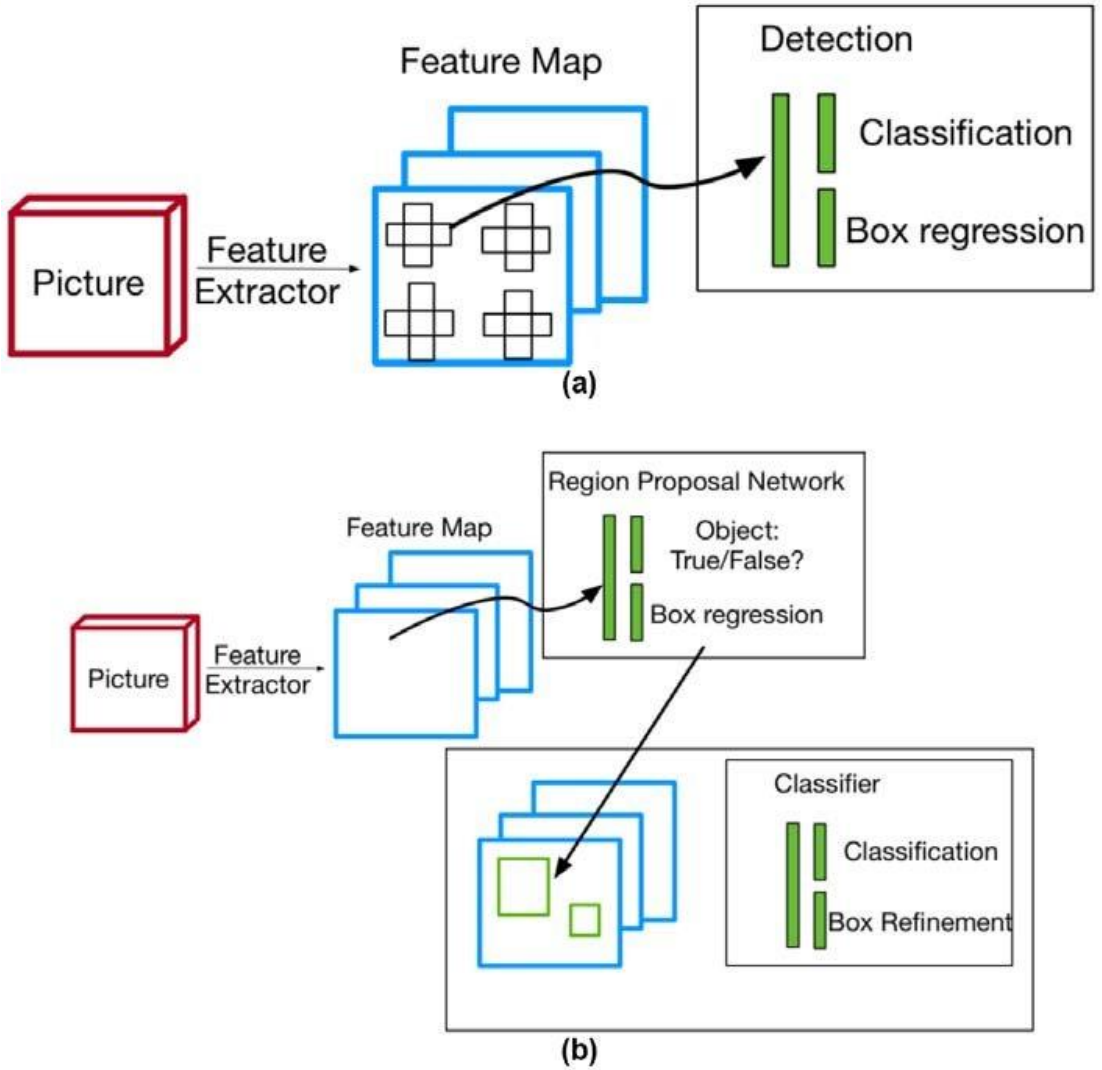
- Inspired by SPPnet with single level RoI pooling layer
- Use selective search and then find the corresponding coordinates on obtained feature maps
- The accuracy is 0.5 frame per second

Faster R-CNN

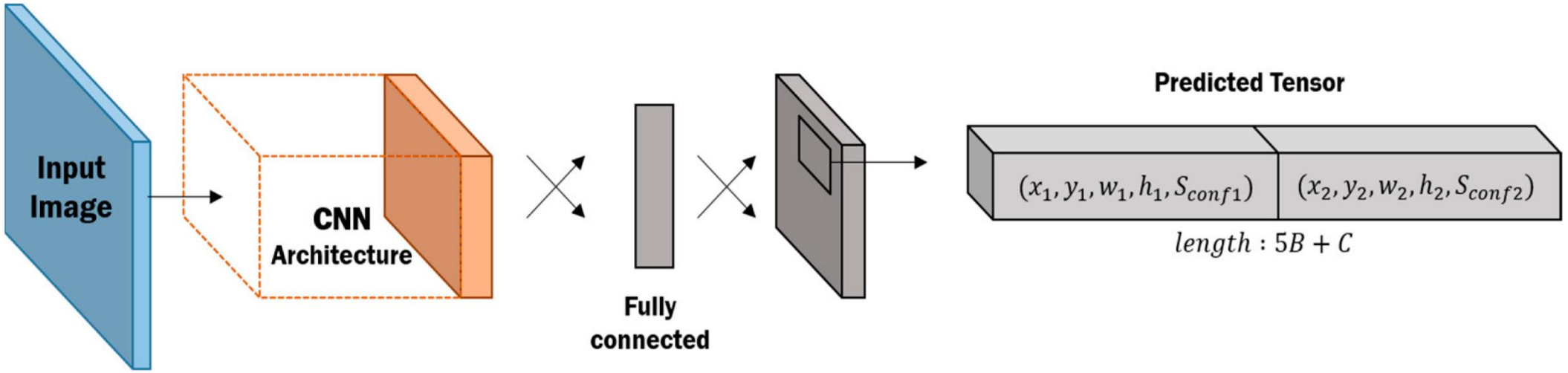


- Introduce region proposal network
- Share the detection network and proposal network
- The accuracy is 5 frame per second

Single stage detectors

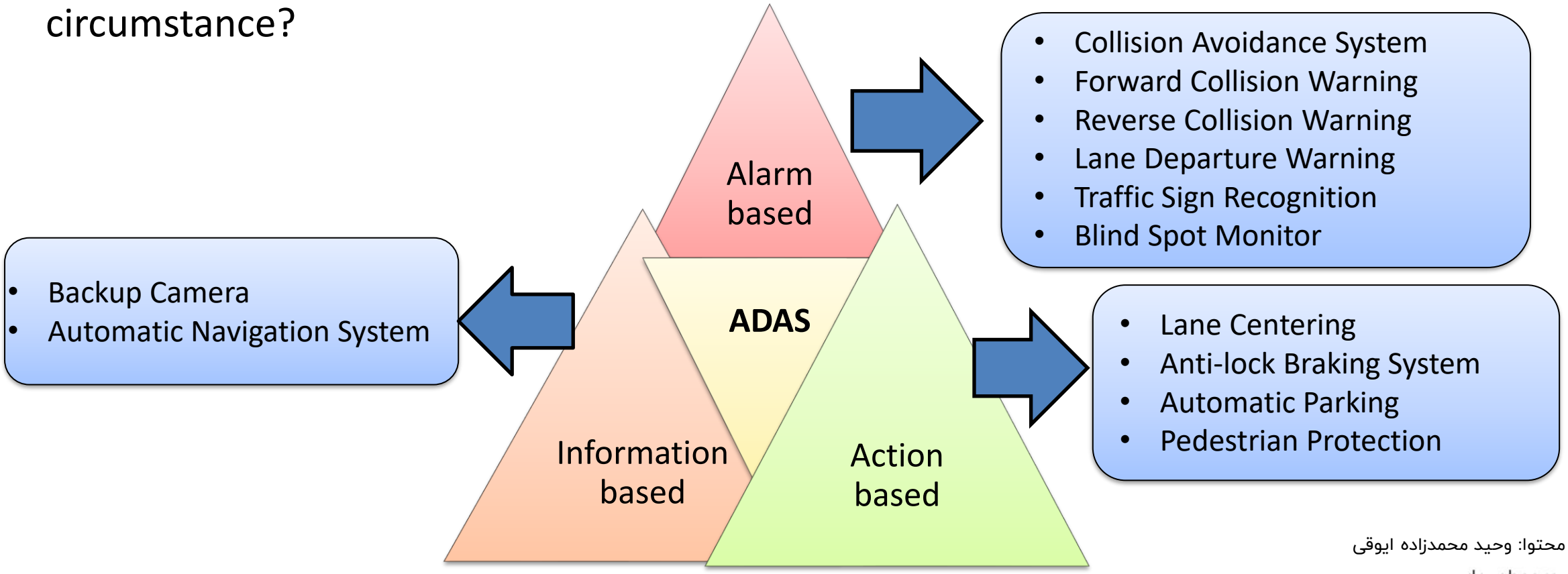


YOLO



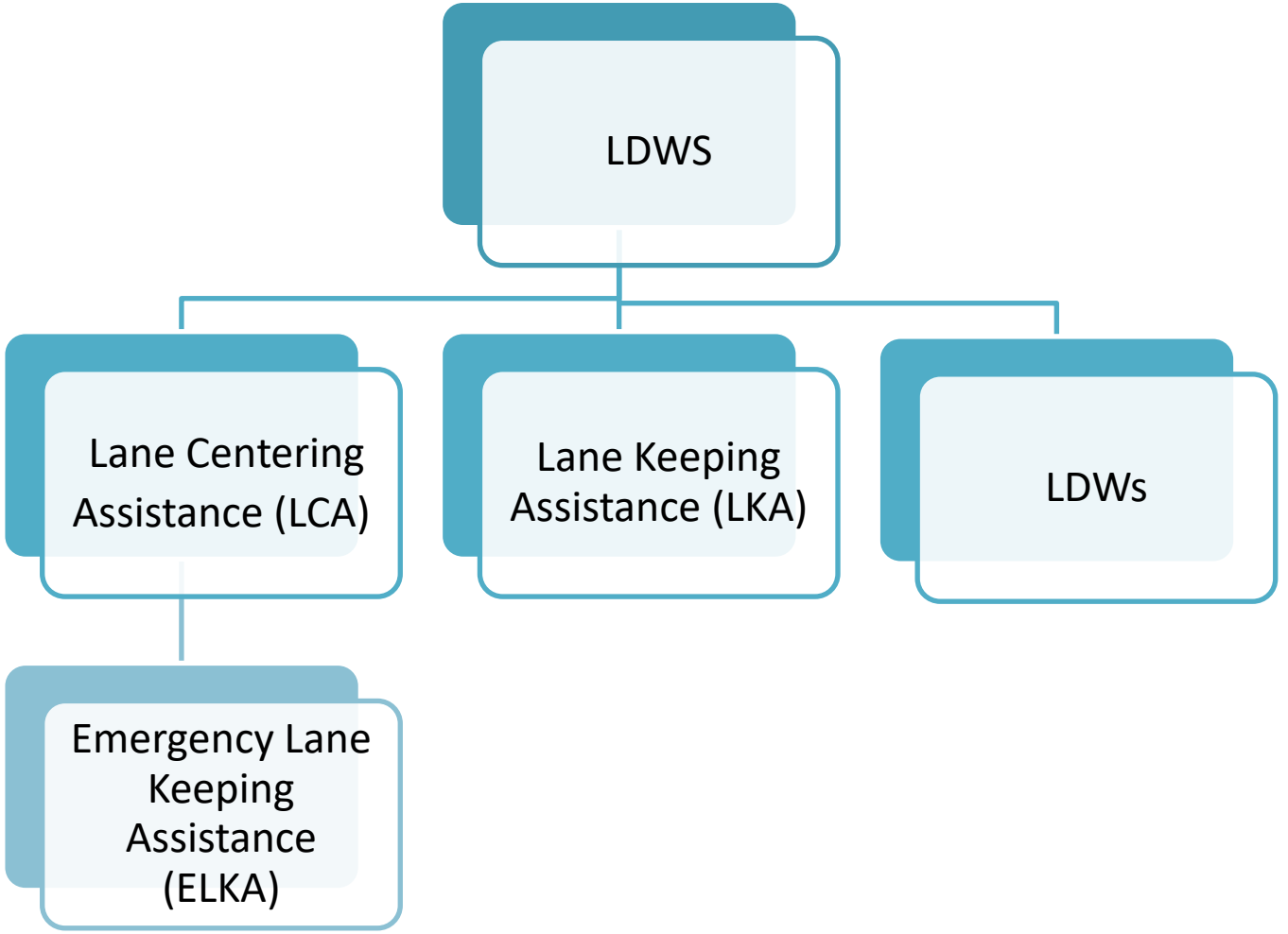
Advanced Driver-assistance Systems

- How to help the drivers with making a suitable decision in the dangerous circumstance?



Lane Departure Warning Systems (LDWs)

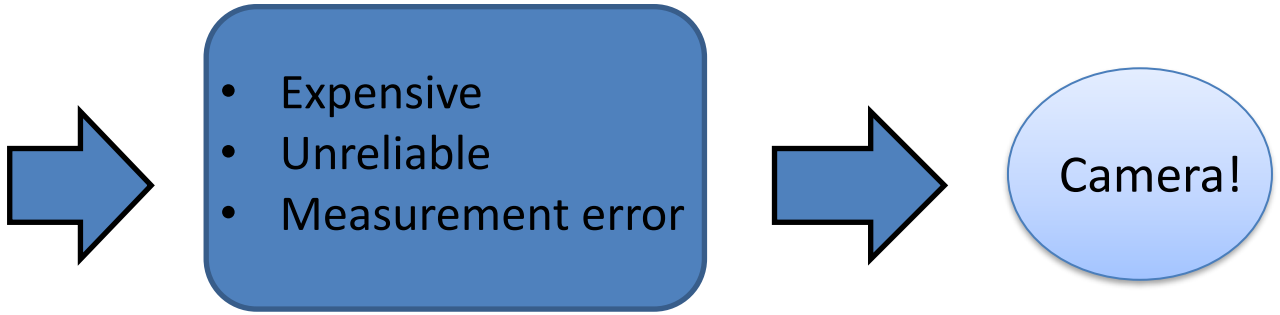
- Different type of services



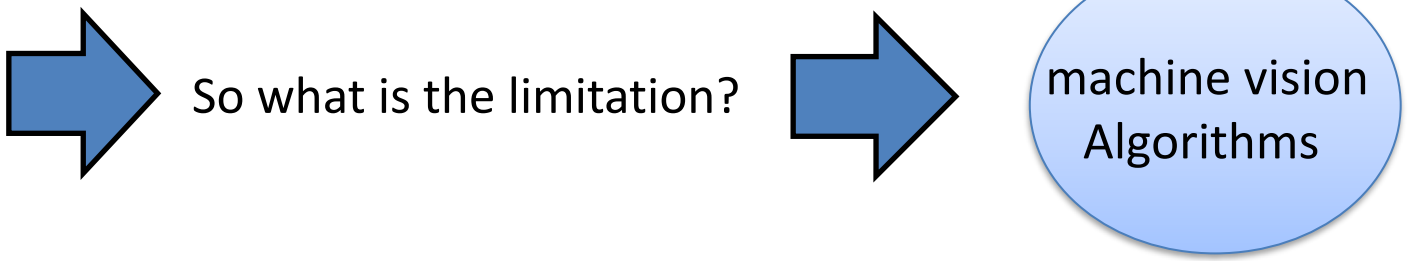
- Unwanted departure
- Cognitive biases
- Drowsiness effects

Requirements

- Different type of technology
  - Laser
  - Global positioning system (GPS)
  - Infrared sensors



- Vision-based lane detection systems
  - Inexpensive
  - Reliable
  - A fruitful source of information



Traditional machine vision – pipelines

- Calibration
- Color analysis
- Perspective correction
- Edge detection
- Hough line detection

- Mapping from 3D scene to 2D images
- Calibration remedies this problem

- We are seeking to the white and yellow color exist in images
- Which color space would be suitable for analysis?

- Change the car perspective to birds’ eye perspective
- Not essential but will make the process easier

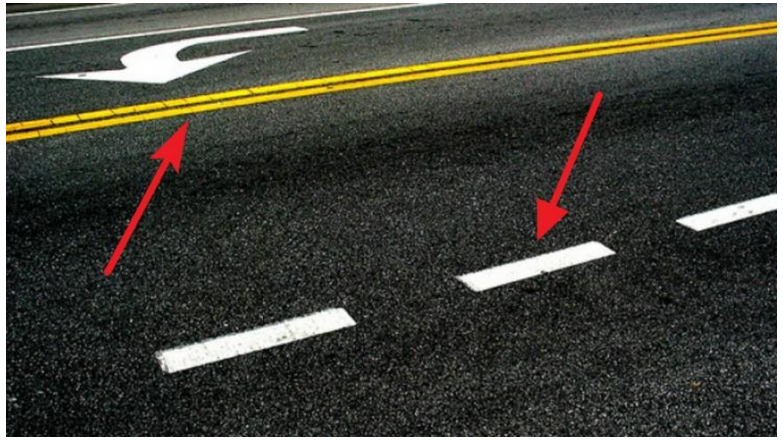
- Highlight the lane borders
- Give the additional information up

- Find the lane coordinates
- Sliding windows for tracking the lanes

Drawbacks

What is the problem with traditional machine vision algorithms?

The lanes are clearly visible

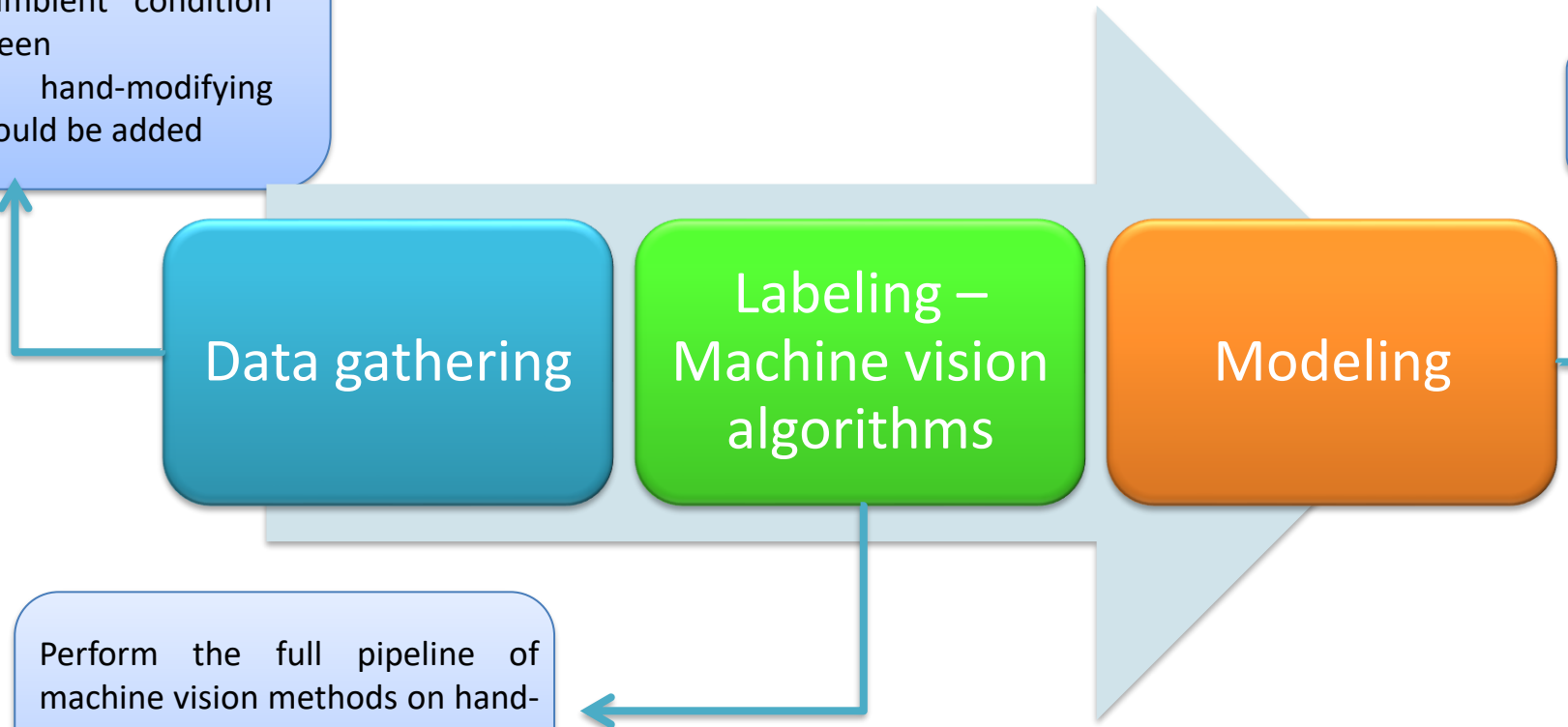


Where are the lanes?

The sunlight is nearly of white pixels!

Deep learning usage

- The most important step in every data-driven problem
- Different ambient condition should be seen
- Sometimes hand-modifying features should be added



- CNN architecture
- Regression model

Perform the full pipeline of machine vision methods on hand-modified images to compute the lane coordinates

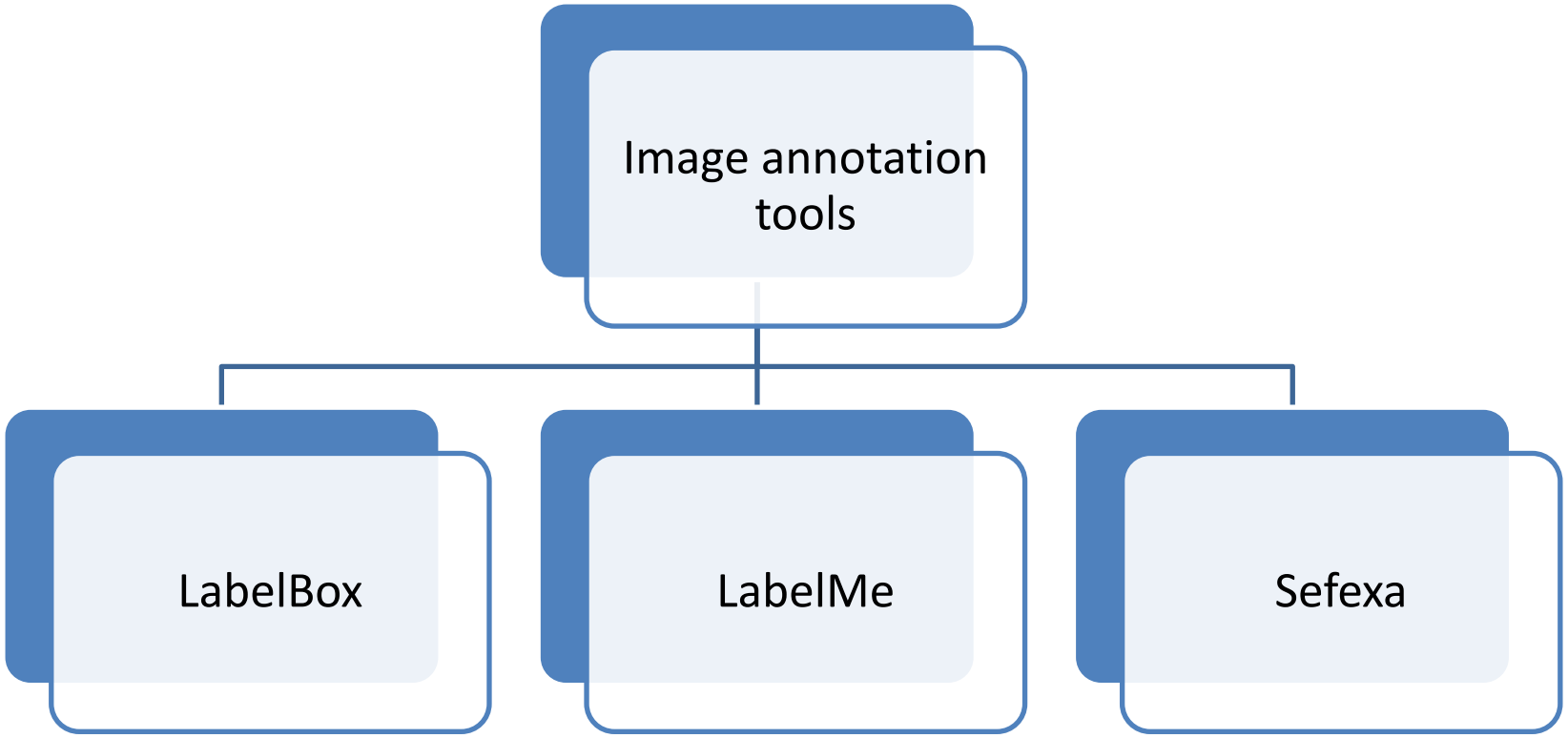
Data gathering

- Data is collected using my smart phone – 48 mega pixel, 30 frame per second
- Belongs to the Yadegar Imam, Chamran, and Hashemi Rafsanjani Highway

Frame for Yadegar Imam Highway

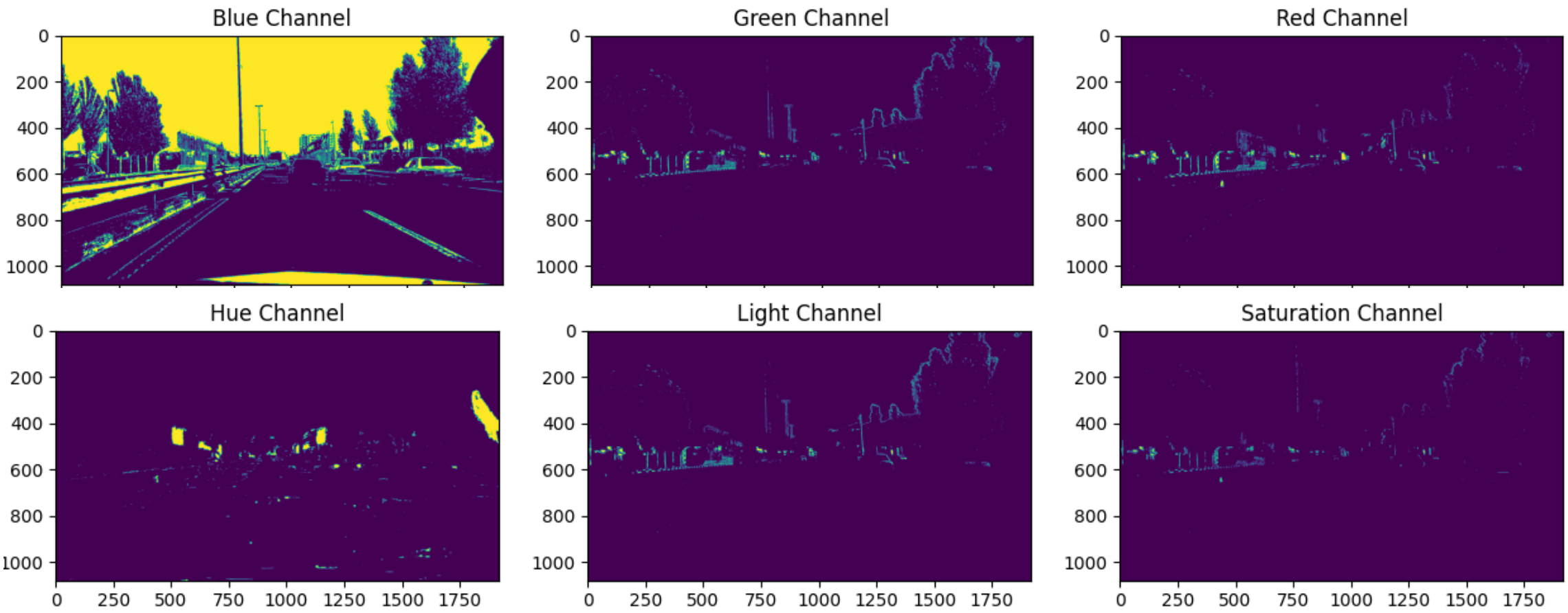
Frame for Hashemi Rafsanjani Highway

Image annotation tools



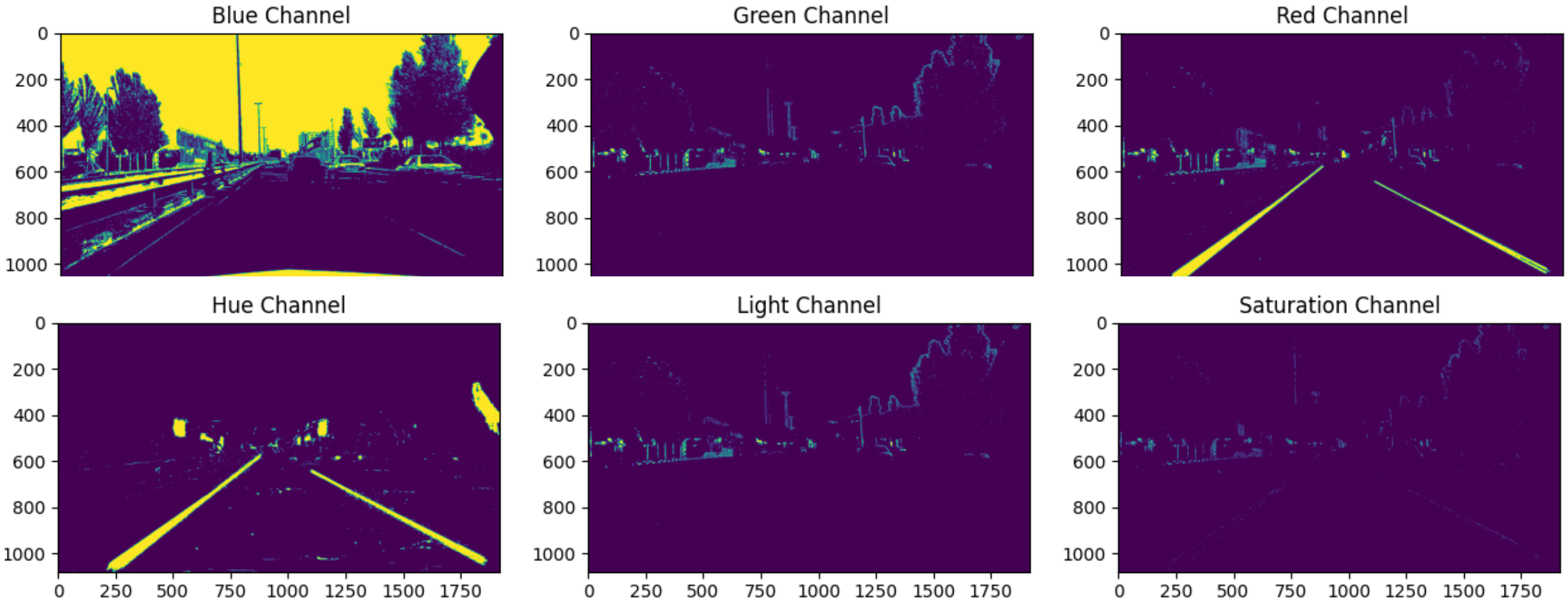
تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

Color space analysis



تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
گروه دایکه | dayche.com

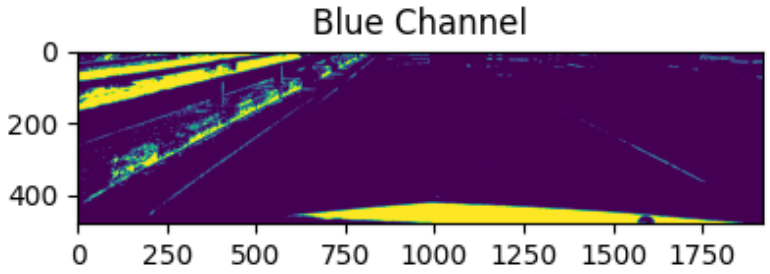
Handy-modified images

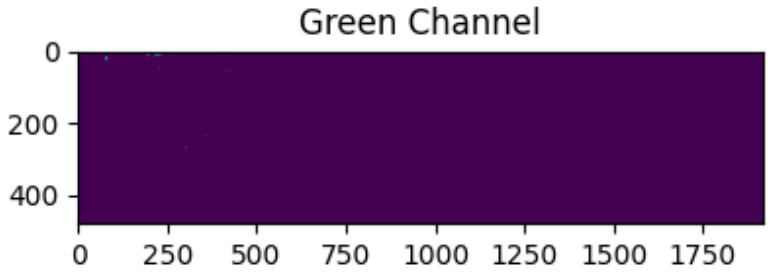


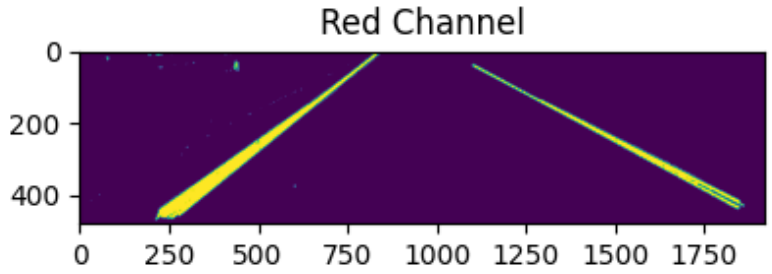
تولید محتوا: وحید محمدزاده ایوقی

Annotation – Hough line transform

- We are seeking to lanes – so remove additional information

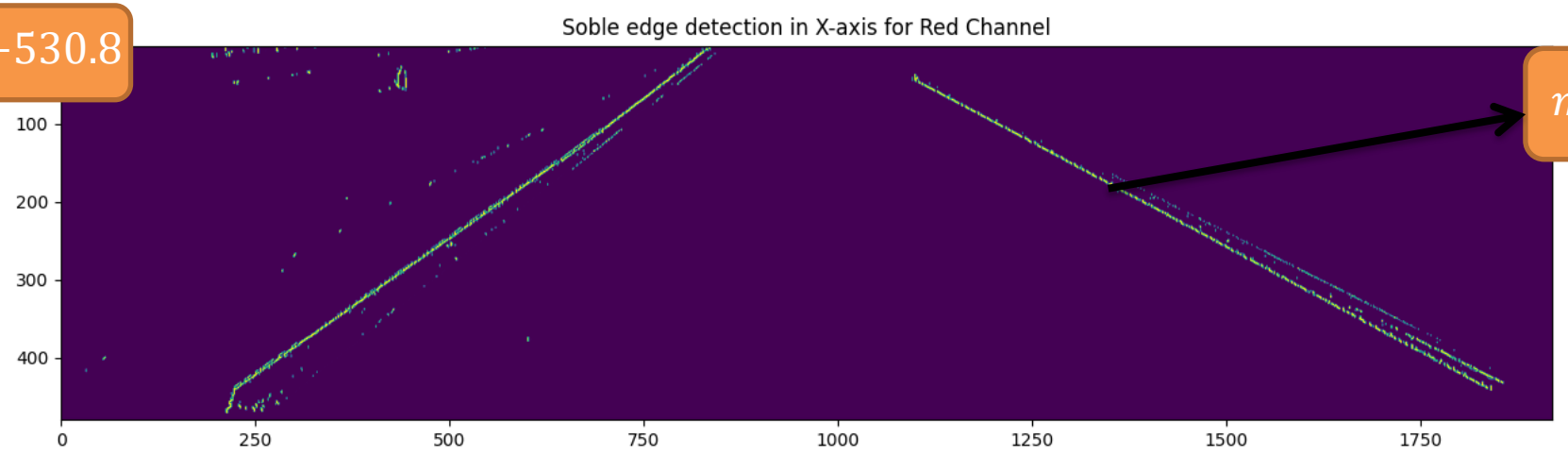






- Hough-line transform

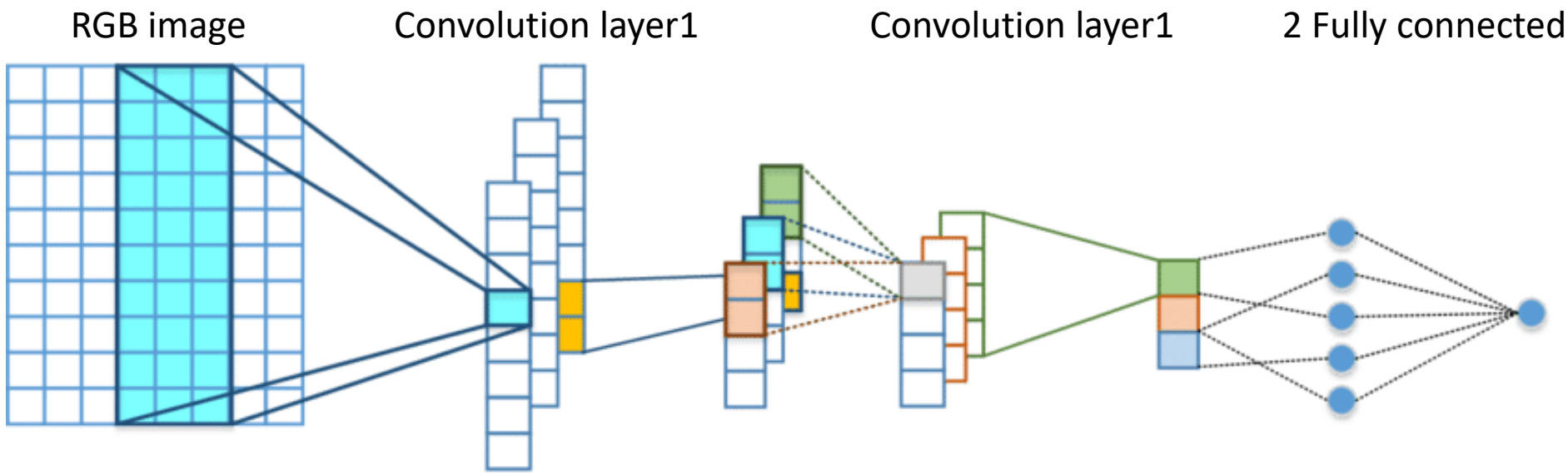
$m = 0.52, b = -530.8$



$m = -0.73, b = 623.9$

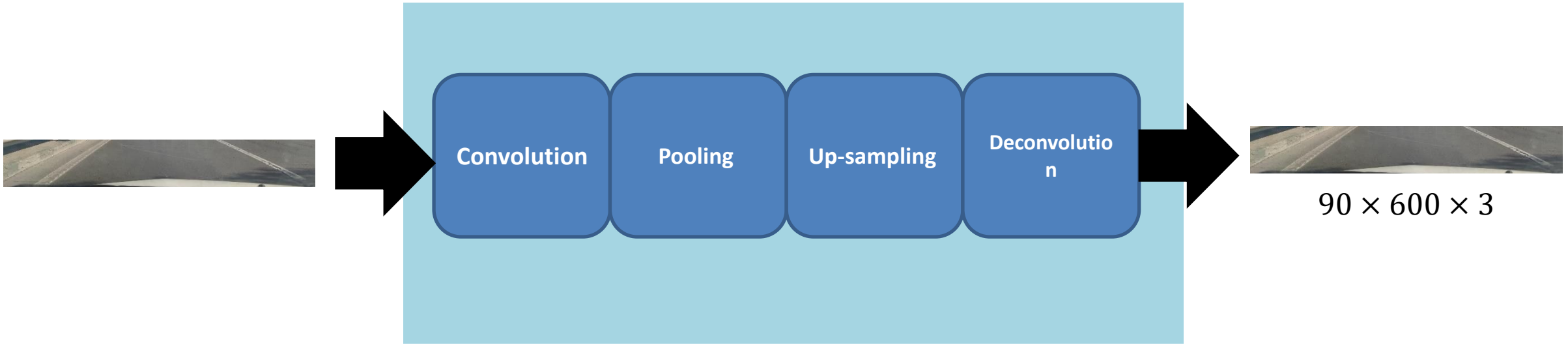
تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
گروه دایکه | dayche.com

Network architecture



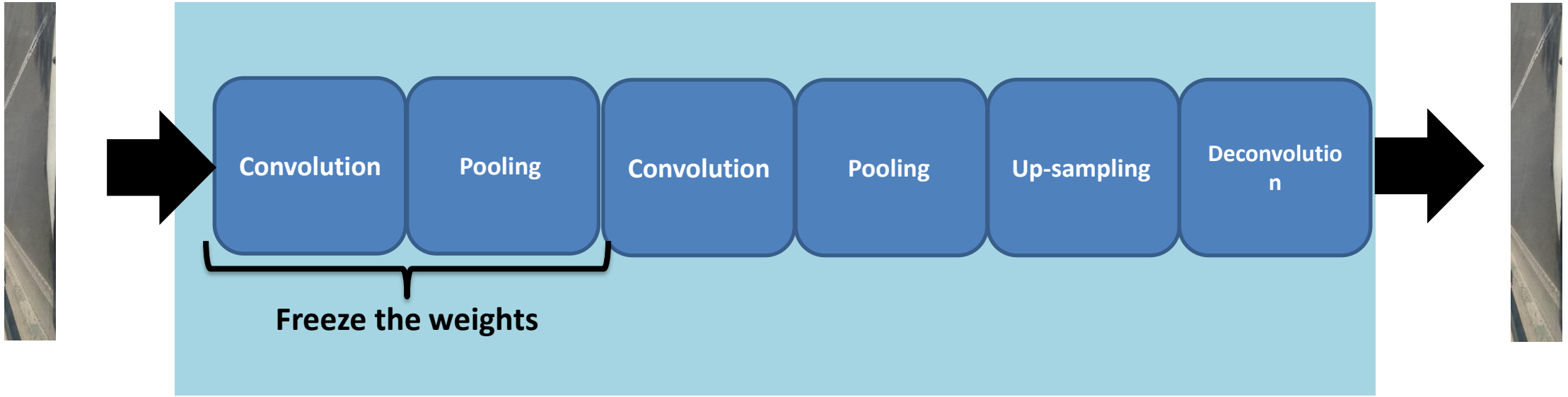
- There are nearly 90000 parameters
- The network aims to be trained based on 1435 selected images
- Some images are not suitable, so they should be removed. Finally 892 images are selected .

Layer-wised training – first convolution layer



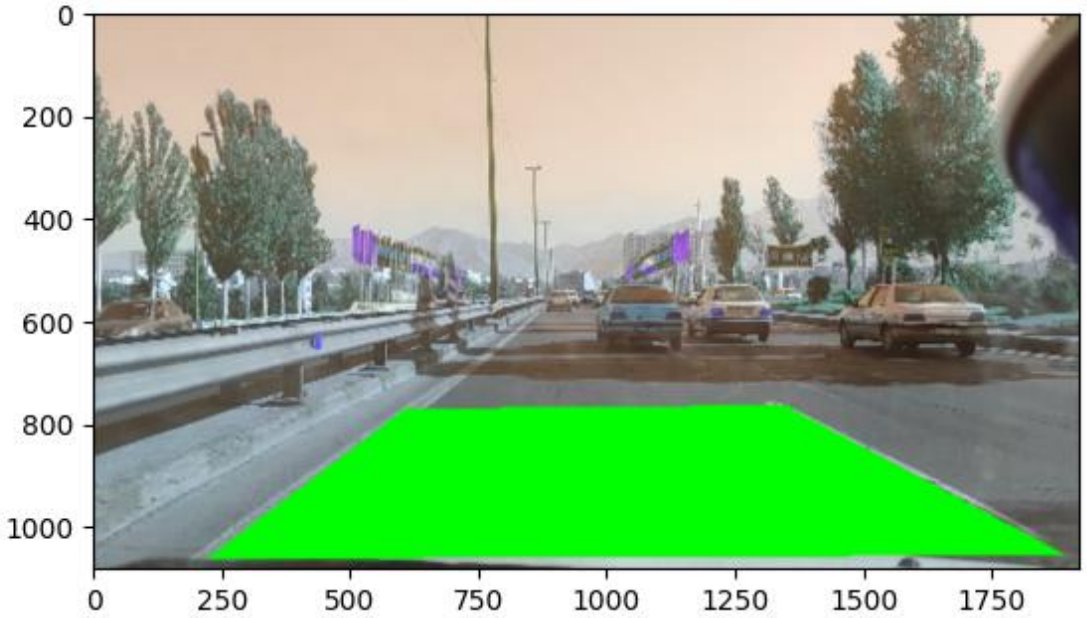
The network is trained for 12000 images in an unsupervised manner!

Layer-wised training – second convolution layer



The network is trained for 12000 images in an unsupervised manner!

Output result



تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه