Paper review

Study on State-of-Art Object Detection Based on Deep Learning

Fabio Cecchinato
100720

João Cordeiro
76548

Marco Santos
83192

deti universidade de aveiro
departamento de eletrónica, telecomunicações e informática

universidade de aveiro

deti universidade de aveiro departamento de eletrónica, telecomunicações e informática
universidade de aveiro

Computer vision → Object detection → Convolutional Neural Networks

- ★ Neural Networks
- ★ Convolutional layers
- ★ mAP (Mean Average Precision) and IoU (Intersection over Union)
- ★ Outperforming results

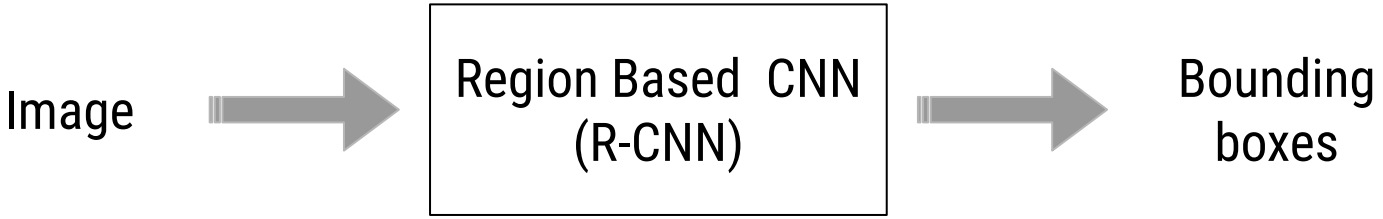
R-CNN, Fast R-CNN and Faster R-CNN

Paper review

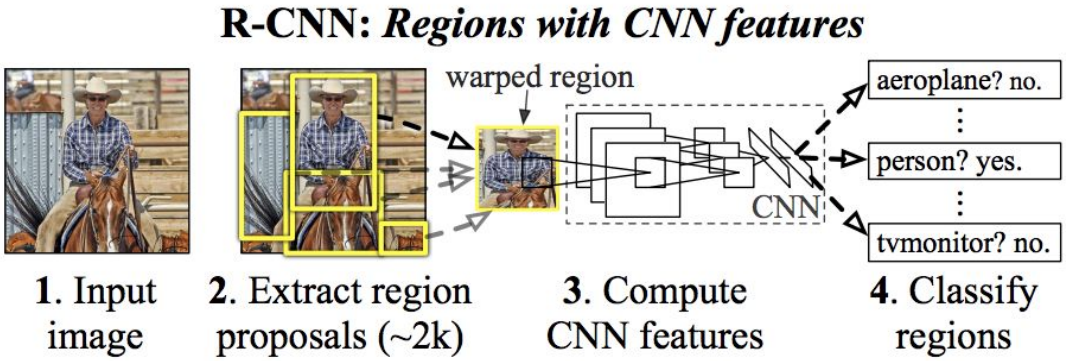
Study on State-of-Art Object Detection Based on Deep Learning

deti universidade de aveiro
departamento de eletrónica, telecomunicações e informática

universidade de aveiro

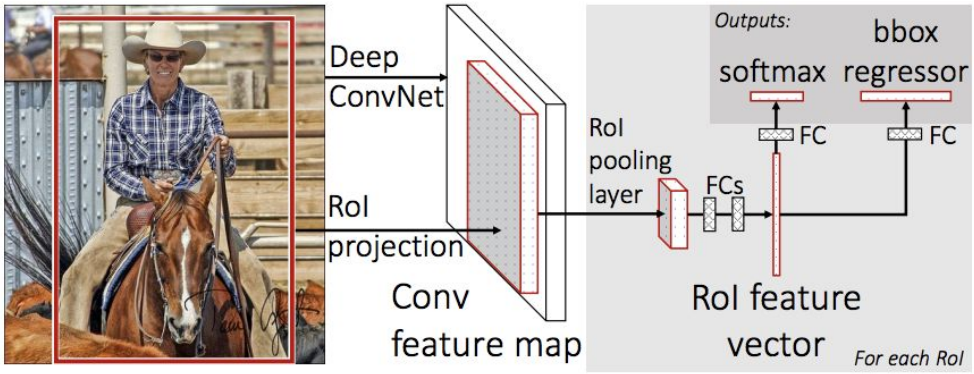


- ★ Pre-training to extract ~2000 category independent region proposals (Regions of Interest - RoI)
- ★ Warping of the RoI to feed the CNN
- ★ The CNN extracts the features from the RoI, than a SVM classifies them



Problems:

- ★ Computationally very expensive (analyze 2000 RoI per image)
- ★ Both a CNN and a SVM are required



Fast R-CNN:

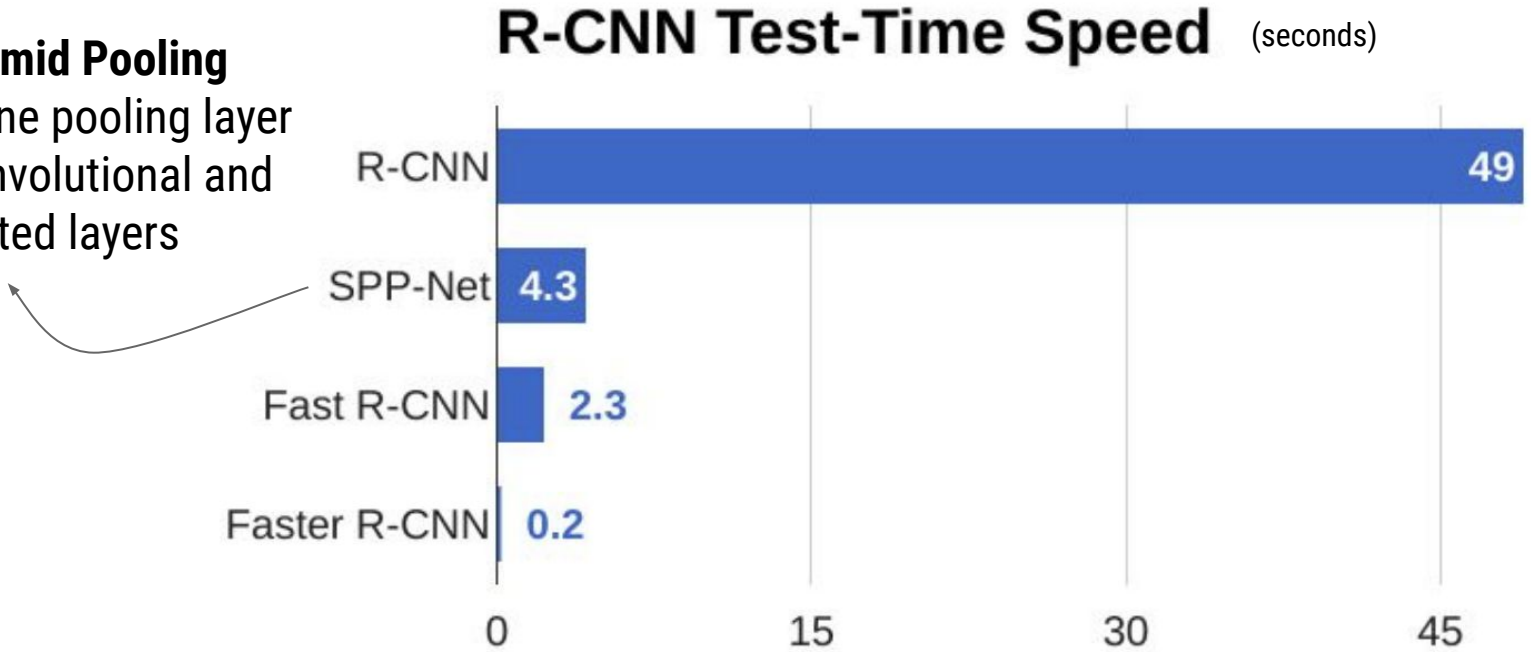
- ★ No initial RoI: they are computed from the features extracted by the CNN
- ★ Substantial improvement on the speed, since for each image just one passing through the network is required

Faster R-CNN:

- ★ The bottleneck of R-CNN and Fast R-CNN is the RoI selection, which is very time consuming
  - It uses a *selective search* algorithm
- ★ To speed-up the process:
  - Feed the input image in a Region Proposal Network (to extract the RoI) and then to a CNN for the classification of each RoI

deti universidade de aveiro departamento de eletrónica, telecomunicações e informática
universidade de aveiro

Spatial Pyramid Pooling
more than one pooling layer between convolutional and fully connected layers



SSD: Single Shot MultiBox Detector

Paper review

Study on State-of-Art Object Detection Based on Deep Learning

deti universidade de aveiro
departamento de eletrónica, telecomunicações e informática

universidade de aveiro

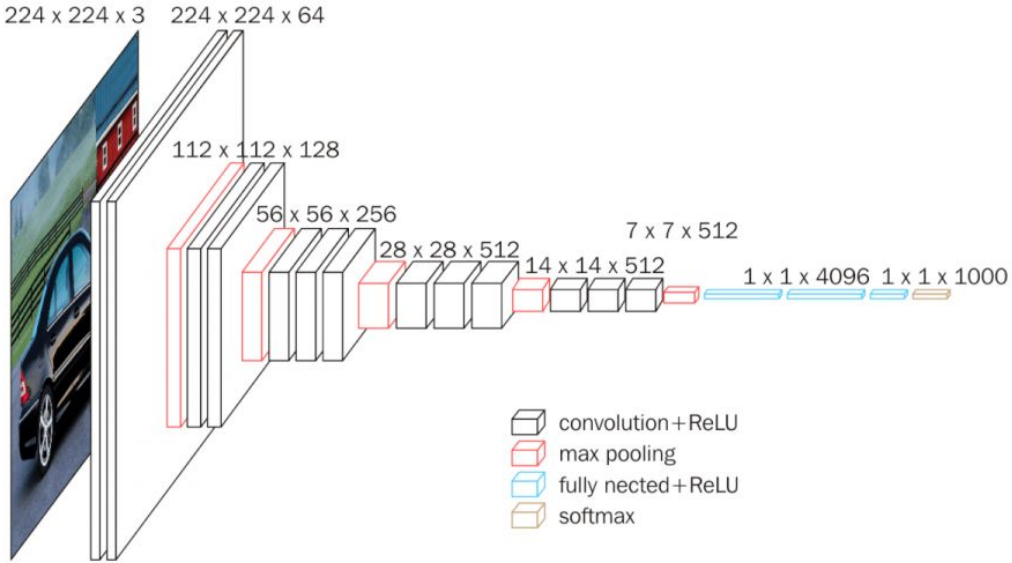
Is based on a feed-forward convolutional network that creates a fixed-size collection of bounding boxes and detects object classes within those boxes.

SSD only takes a single look at the image in order to analyze it.
Does it using the VGG-16 architecture due to its high performance when it comes to image classification.

SSD- Multi-scale feature maps for detection

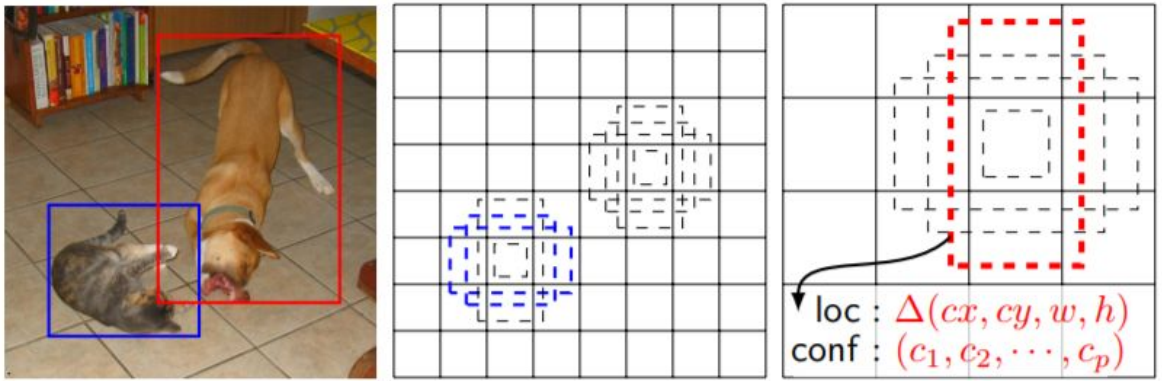
Using a convolutional feature layer attached to the end of the network, its possible to decrease the size of the image and detect classes in multiple scales.

Each feature can produce a fixed amount of predictions.
A 3∗3∗nchannels kernel is used



deti universidade de aveiro
departamento de eletrónica, telecomunicações e informática
universidade de aveiro

A set of default bounding boxes are added to each feature map cell, predicting the offset relative to the default bounding box and category.



Confidence Loss:

$L = (L_{conf} + \alpha * L_{loc})/N$

where N is the number of matched default bounding boxes, L conf the confidence loss, α balancing contributor and L loc the location loss.

The Confidence Loss permit to reduce the number of bounding boxes.

YOLO: You Only Look Once

Paper review

Study on State-of-Art Object Detection Based on Deep Learning

deti universidade de aveiro
departamento de eletrónica, telecomunicações e informática

universidade de aveiro

Is a single shot multibox detector. Differ from SDD on detection and classification strategies.

- ★ **Dimension Cluster:** automatically find a good default bounding box.
- ★ **Direct location prediction:** easier to learn parametrization.
- ★ **Batch Normalization:** improved convergence.

Darknet-19 architecture:

★ Faster than VGG-16.

★ Slightly less accurate than VGG-16.

Two more interactions, YOLO9000 and YOLOv3

Analysis of the paper

Paper review

Study on State-of-Art Object Detection Based on Deep Learning

deti universidade de aveiro
departamento de eletrónica, telecomunicações e informática

universidade de aveiro

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	1
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	1
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	1
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	-
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	-
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	1-2
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	1-2
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	1-2
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	-
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	-
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	-
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	-
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	-

- ★ Most of the aspects of the list are present
- ★ The ones that are not present are related to things that are relevant in a paper only if this is about a new work proposal, while the paper under analysis is just a review on the state of the art

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	-
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	-
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	2
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	2-4
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	-
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	2-4
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	2-4
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	-
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	-
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	4
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	4
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	4
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	4-5

Other considerations:

- ★ Sometimes makes use of colloquial language
  - “Why’s that?”
  - Use of “we” instead of a more impersonal approach
- ★ Some sources are not properly cited
  - References assigned to the wrong sources
  - Missing links to references

used in 1998 to detect digits [1]. But what was the reason for this recent popularity then? AlexNet. Back in 2010 the *ImageNet Large Scale Visual Recognition Challenge (ILSVRC)* was launched and in just 2 years a major revolution in the field of computer vision was on. Why’s that? In 2012 researchers Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton

In terms of limitations related to our work, we didn’t feel like there was any major obstacle during our research for the state of the art in this field, although if we knew about Arxiv

a given image. This method [5] is slow and a time-consuming process that has a big impact in the performance of the network. In order to solve these problems, another variant of *R-CNN* called *Faster R-CNN* was developed. This method

more information about them, to then find some articles on Medium (Towards Data Science) explaining their evolution as