A Case Study: Deep Learning Methods for Vehicle Damage Detection

Wai Po Kevin Teng

Content

Motivation

- Vehicle is a common means of transportation susceptible to damages
- External condition of vehicles is crucial for insurance claim, repair purpose and reselling purpose
- Manual visual inspection is tedious and time consuming
- The need of an automated system without compromising speed and efficiency
- Deep learning based computer vision as an optimum for object detection
- Leveraging deep learning methods for vehicle damages detection

Scope of Work

- Vehicle type: car
- Deep learning algorithm: Convolutional Neural Networks (CNNs)
- Deep learning architecture:
  - ✘ Backbone (feature extractor) architectures
  - ✓ SOTA Object detection frameworks
- Performance based on PASCAL VOC 2012 challenge
- No actual implementation

Background (Object Detection)

- Object recognition + Object localization:
  - What is the object? => Object Recognition
  - Where is the object? => Object Localization
- **Two stage detector:**
  - First stage relies on region proposal module for object proposals
  - Second stage attempts to classify and localize the objects
  - Example: R-CNN family (R-CNN, Fast R-CNN, Faster R-CNN)
- **Single stage detector:** Classify and localize objects in a single module
  - Example: YOLO

Background (Backbone)

- Network architectures that extract features
- SOTA: VGGNet, GoogleNet, ResNet, DenseNet, etc.

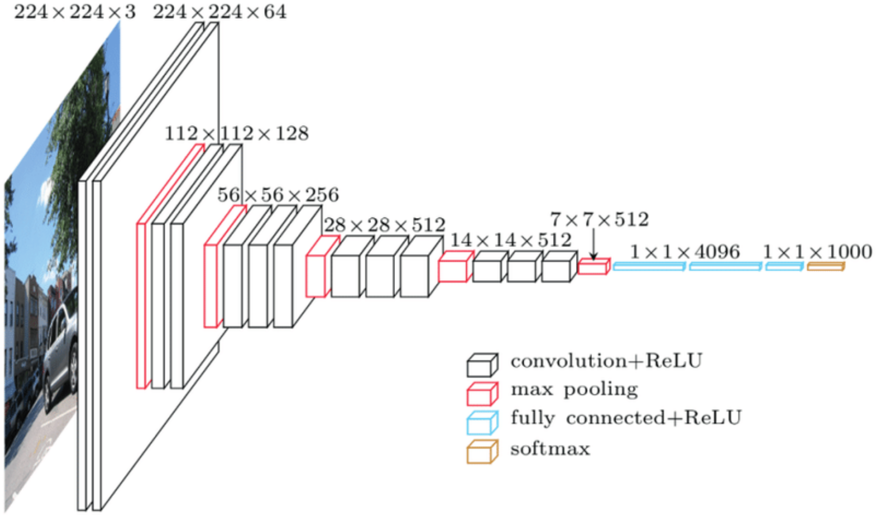

Figure 1: VGGNet architecture. Image source by Bacanin and Bezdan (2019).

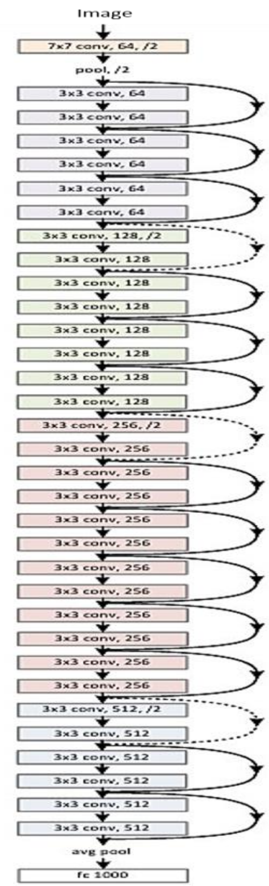

Figure 2: ResNet architecture. Image source by Kaiming He et al. (2015).

Background (R-CNN)

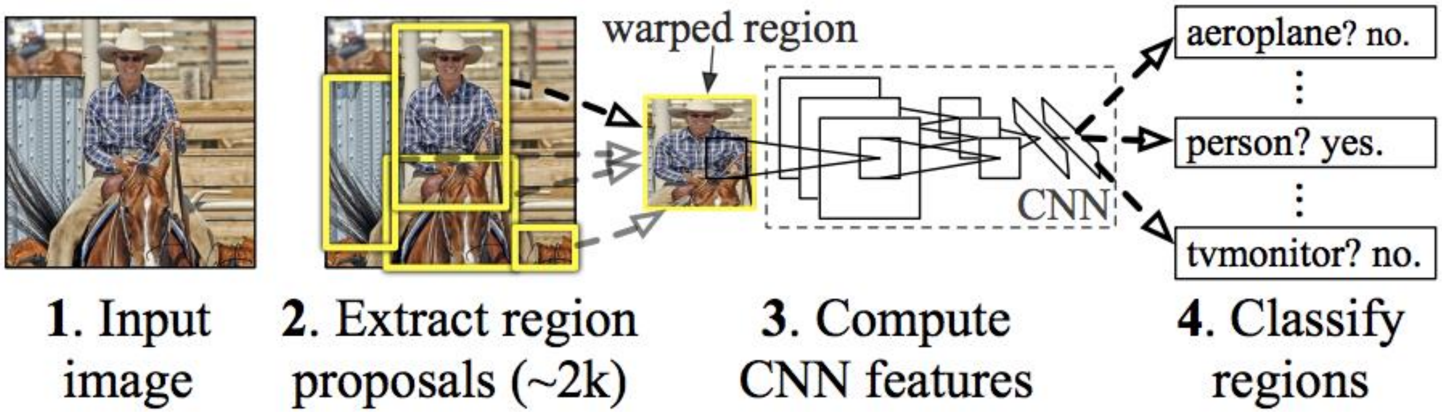

Figure 3: Procedures of R-CNN while detecting an object. Image source by Girshick et al. (2013).

- Proposed by Girshick et al. (2013)
- Selective search for region proposals extraction
- CNN for region proposals feature extraction
- Support vector machine (SVM) as a classifier

Background (Fast R-CNN)

- Drawbacks of RCNN:

  - Feature extraction on selective search candidates is computational expensive

- Proposed by Girshick (2015)
- Region proposal extractor (selective search) on CNN features
- ROI pooling on selective search candidate in one layer

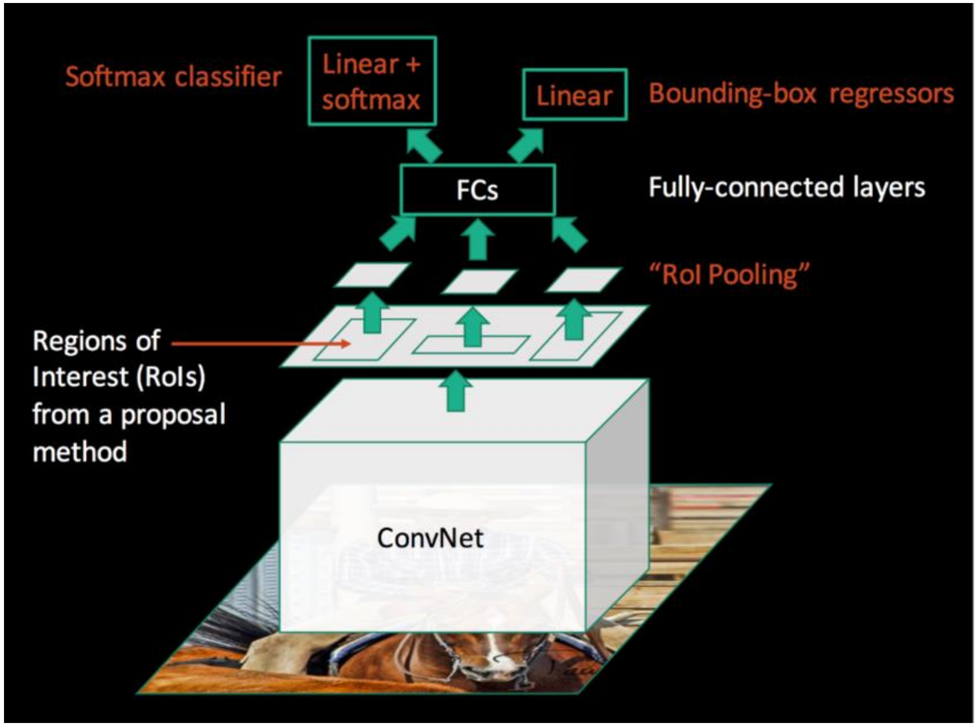

Figure 4: Procedures of Fast R-CNN while detecting an object. Image source by Jonathan Hui (2017).

Background (Faster R-CNN & Mask R-CNN)

- Drawback of Fast R-CNN:

    - Selective search is a greedy approach

- Proposed by Ren et al. (2016).
- Fully convolutional network (FCN) [Long et al. (2014)] as region proposal network (RPN) for feature maps
- Additional instance segmentation branch -> Mask R-CNN (He et al. 2017)

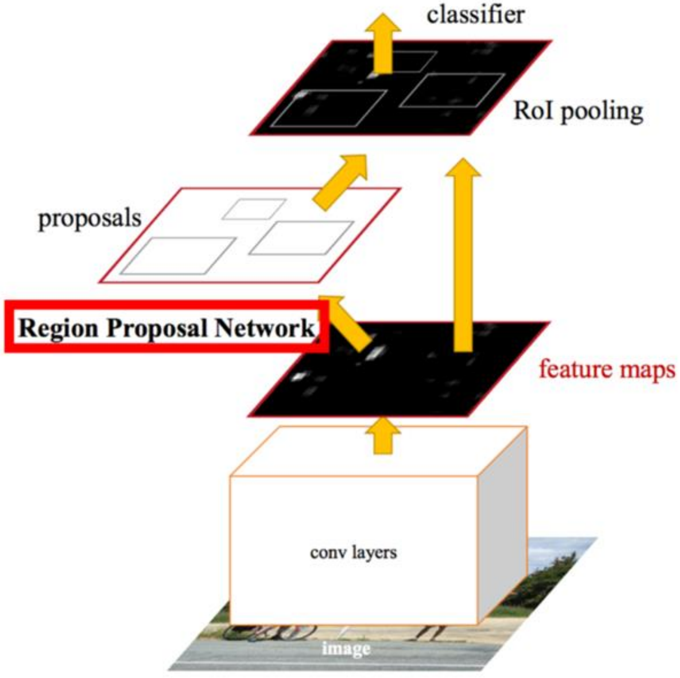

Figure 5: Procedures of Faster R-CNN. Image source by Jonathan Hui (2017).

Background (YOLO)

- You Only Look Once (YOLO) proposed by Redmon et al. (2015)
- Readdress object detection as a regression task
- Split image into SxS grid, with each grid cells predicts multiple bounding boxes and confidence score
- Each bounding boxes contains class probability

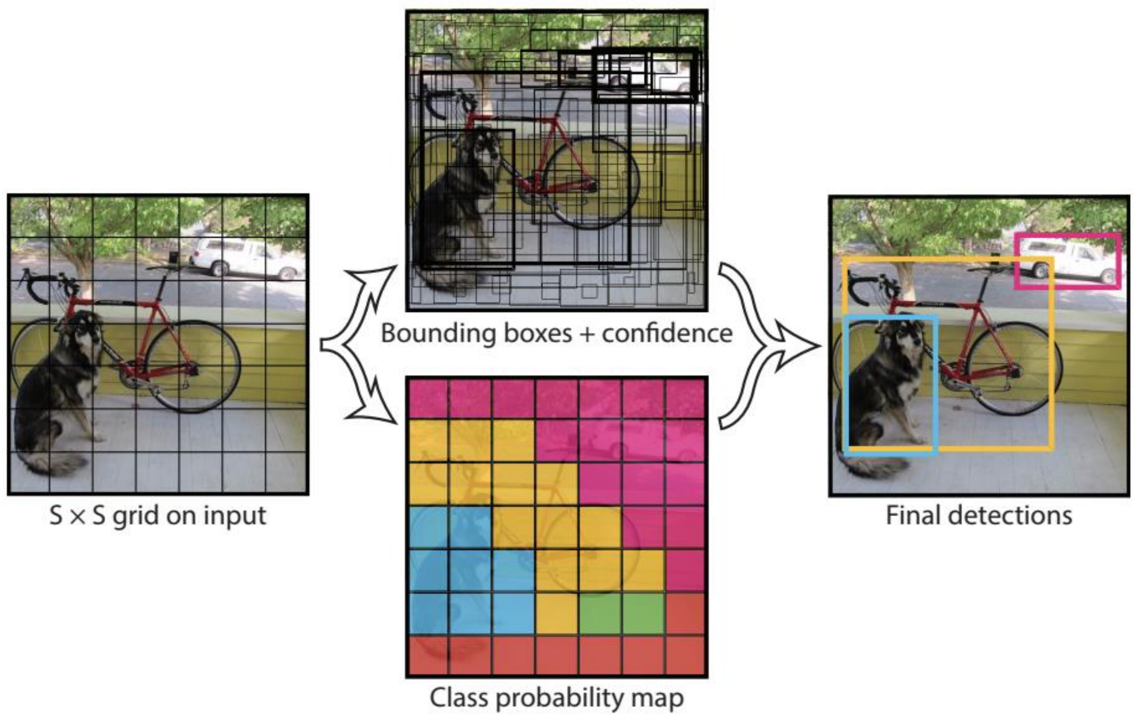

Figure 6: Procedures of Faster R-CNN. Image source by Redmon et al. (2015).

Background (Comparison Results)

Model	Year	Backbone	Size	$AP_{0.5}$	FPS
R-CNN	2014	AlexNet	224	58.50%	0.02
Fast R-CNN	2015	VGG-16	Variable	65.70%	0.43
Faster R-CNN	2016	VGG-16	600	**67.00%**	5
YOLO	2015	GoogleNet	448	57.90%	**45**

- Two stage detectors (R-CNN, Fast R-CNN, Faster R-CNN) has higher $AP_{0.5}$ scores but slow in real time object detection
- Single stage detector (YOLO) has object detection speed but does not fare well on small or clustered objects

Experiment Setup (Data Description)

- **Data set preparation:**
  - ➢ Scarcity of publicly available data set for car damage classification
  - ➢ Web scraping through keywords
  - ➢ Formulate types of car damages

Figure 7: Image gallery of different damage-types, starting from (left column to right column) bumper dent, scratches, door dent, glass shatter, head-lamp broken, tail-lamp broken and smashed. Image adapted from Malik et al. (2020).

Experiment Setup (Data Description)

- **Data preprocessing:**
  - Resize image (Homogeneity)
  - Normalize image pixel intensity
  - Bounding boxes annotations, class label annotations
  - Segmentation annotations (optional)
- **Data Augmentation:** Rotation, zooming, shearing, etc.
- **Transfer learning:** transfer knowledge from pre-trained model will share similar features on car damages detection.

Experiment Setup (Car Damaged Area Detector)

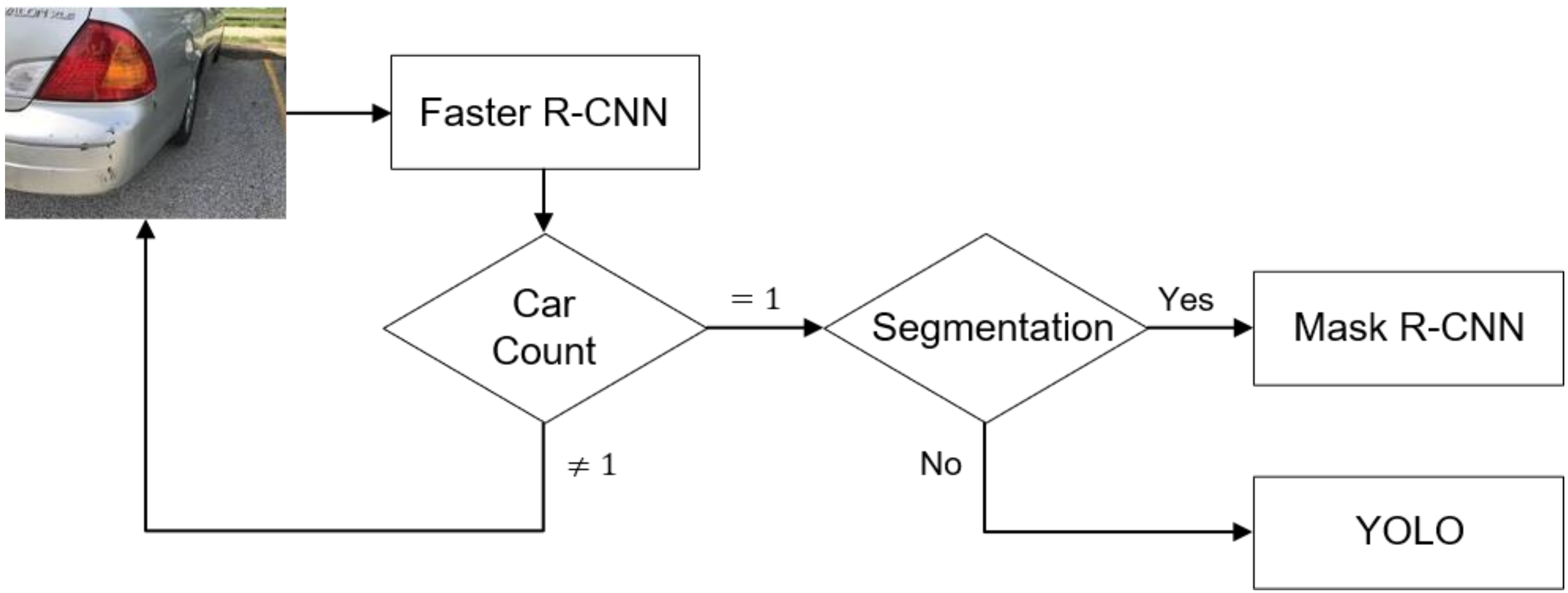

Figure 8: Flow chart on car damaged area detection pipeline.

- Faster R-CNN to calculate the number of cars in an image
- Car damaged area detection with YOLO
- Car damaged area detection + segmentation with Mask R-CNN

References

1. Nebojsa Bacanin and Timea Bezdan "Convolutional Neural Network Layers and Architectures ", International Scientific Conference on Information Technology and Data Related Research, 2019.
2. Kaiming He et al. "Deep Residual Learning for Image Recognition". In: vol. abs/1512.03385. 2015.
3. Ross B. Girshick et al. "Rich feature hierarchies for accurate object detection and semantic segmentation". In: vol. abs/1311.2524. 2013.
4. Jonathan Hui "Fast R-CNN and Faster R-CNN". Jonathan Hui Blog, https://jhui.github.io/2017/03/15/Fast-R-CNN-and-Faster-R-CNN/, 2017. [Accessed 27.02.2022].
5. Ross B. Girshick. "Fast R-CNN". In: vol. abs/1504.08083. 2015.
6. Jonathan Long, Evan Shelhamer, and Trevor Darrell. "Fully Convolutional Networks for Semantic Segmentation". In: vol. abs/1411.4038. 2014.
7. Joseph Redmon et al. "You Only Look Once: Unified, Real-Time Object Detection". In: vol. abs/1506.02640. 2015.
8. Hashmat Shadab Malik et al. Deep Learning Based Car Damage Classification and Detection. EasyChair Preprint no. 3008. EasyChair, 2020.
9. Shaoqing Ren et al. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks". In: vol. abs/1506.01497. 2016.
10. Kaiming He et al. "Mask R-CNN". In: vol. abs/1703.06870. 2017.

Thank You for Your Kind Attention!