# 阅读纲要

## 1 自己的总结、评价以及应用

该论文是关于object detection based deep-learning的概述以及相关技术介绍（object detection architectures），其中还介绍了的一些应用。

. In this paper, we provide a review of deep learning-based

object detection frameworks.

## 2 文章的主要问题（abstract、疑问句中）

The problem definition of object detection is to determine

where objects are located （localization）in a given image (object localization)

and which category（classification） each object belongs to.

关于object detection based deep-learning的概述以及相关技术介绍（object detection architectures），其中还介绍了的一些应用。

Therefore, the pipeline of traditional object detection models can be mainly divided into three stages: informative region selection, feature extraction, and classification

## 3 结论（abstract以及conclusion中）

## This paper provides a detailed review on deep learning-based object detection frameworks that handle different subproblems, such as occlusion, clutter,and low resolution, with different degrees of modifications on R-CNN.

## 4 思路脉络（小标题中的关键句）

1. INTRODUCTION

Object detection的主要任务

classifying different images　estimate the concepts and locations

of objects

Object detection的应用

object detection　consists of different subtasks such as face detection [2], [S2], pedestrian detection [3], [S2], and skeleton detection [4], [S3] and is related to many applications, including image classification [5], [6], human behavior

analysis [7], [S4], face recognition [8], [S5], and autonomous driving [9], [10].

Object detectiond的主要问题

The problem definition of object detection is to determine

where objects are located （localization）in a given image (object localization)

and which category（classification） each object belongs to.

Object detection的pipelines

Therefore, the pipeline of traditional object detection models can be mainly divided into three stages: informative region selection, feature extraction, and classification

1. Informative Region Selection
2. Feature Extraction
3. Classification

II. BRIEF OVERVIEW OF DEEP LEARNING

A. History: Birth, Decline, and Prosperity

B. Architecture and Advantages of CNN

The advantages of CNN against traditional methods can be

summarized as follows.

1) Hierarchical feature representation, which is the multilevel representations from pixel to high-level semantic features learned by a hierarchical multistage

Fig. 2. Two types of frameworks: region proposal based and regression/classification based. SPP: spatial pyramid pooling [64], FRCN: faster R-CNN [16], RPN: region proposal network [17], FCN: fully convolutional network [65], BN: batch normalization [43], and Deconv layers: deconvolution layers [54] .structure [15], [53], can be learned from data automatically and hidden factors of input data can be disentangled through multilevel nonlinear mappings.

2) Compared with traditional shallow models, a deeper architecture provides an exponentially increased expressive capability.

3) The architecture of CNN provides an opportunity to jointly optimize several related tasks together (e.g., fast R-CNN combines classification and bounding box regression into a multitask learning manner).

4) Benefitting from the large learning capacity of deep CNNs, some classical computer vision challenges can be recast as high-dimensional data transform problems and solved from a different viewpoint

III.GENERIC OBJECT DETECTION

Generic object detection aims at locating and classifying existing objects in any one image and labeling them with rectangular BBs to show the confidences of existence.

The frameworks of generic object detection methods can mainly

be categorized into two types (see Fig. 2).

1. Region Proposal-Based Framework

①R-CNN

②SPP-net

③Fast R-CNN

④Faster R-CNN

1. Regression/Classification-Based Framework

One-step frameworks based on global regression/classification, mapping straightly from image pixels to bounding box coordinates and class probabilities, can reduce time expense.

①YOLO

②SSD

1. Experimental Evaluation

IV. SALIENT OBJECT DETECTION

Visual saliency detection, one of the most important and challenging tasks in computer vision, aims to highlight the most dominant object regions in an image.

Broadly, there are two branches of approaches in salient object detection, namely, BU [127] and TD [128].

1. Deep Learning in Salient Object Detection
2. Experimental Evaluation

V. FACE DETECTION

1. Deep Learning in Face Detection
2. Experimental Evaluation

VI. PEDESTRIAN DETECTION

Recently, pedestrian detection has been intensively studied, which has a close relationship to pedestrian tracking [189], [190], person reidentification [191], [192],and robot navigation [193], [194].

PEDESTRIAN DETECTION存在的问题

to explicitly modelthe deformation and occlusion, part-based models [200] and

explicit occlusion handling （遮挡处理）[201], [202] are of concern.

The confusion of hard background instances, which is in contrast to the interference from multiple categories in generic object detection.

1. Deep Learning in Pedestrian Detection

目前的解决方案

① Zhang et al

adapt generic Faster R-CNN [17] to pedestrian detection.They modified the

downstream classifier by adding boosted forests to shared, high-resolution conv feature maps and taking an RPN to handle small instances and hard negative examples.

② Tian et al

proposed a deep learning framework called DeepParts, which makes decisions based on an ensemble of extensive part detectors.

1. Experimental Evaluation

VII. PROMISING FUTURE DIRECTIONS AND TASKS

The first one is small object detection such as occurring in COCO data set and in face detection task.

The second one is to release the burden on manual labor and accomplish real-time object detection, with the emergence of the large-scale image and video data.

The third one is to extend typical methods for 2-D object detection to adapt 3-D object detection and video object detection, with the requirements from autonomous driving, intelligent transportation, and intelligent surveillance.

VIII. CONCLUSION

This paper provides a detailed review on deep learning-based object detection frameworks that handle different subproblems.

Then, three other common tasks, namely, salient object

detection, face detection, and pedestrian detection, are also

briefly reviewed.

Finally, we propose several promising future directions to gain a thorough understanding of the object detection landscape.

NOTES:

①第一遍阅读时不查生词；

②整理专业术语、概念表。

附表：专业术语、概念