Faculty of Mathematics
and Computer Science

Machine learning course (ML)

Object detection using deep neural networks

Comănac Dragoș-Mihail

Department of Computer Science, Babeş-Bolyai University
1, M. Kogalniceanu Street, 400084, Cluj-Napoca, Romania
E-mail: dragos.comanac@stud.ubbcluj.ro

Abstract

Keywords:

1. Introduction

Computer vision is a scientific field that deals with the extraction of meaningful information from visual data such as images or videos. As a consequence, over time, computer vision has emerged as a key field in the domain of artificial intelligence because it intends to replicate the functions of the human visual cortex. This is possible due to the continuous development of optical hardware, which nowadays can exceed the capabilities of the human eye, but probably more important are the huge quantities of data that are available to more and more people.

There are various methods that can be used in the field of computer vision, such as hand crafted features, but in this context of big data, the most relevant methods are related to machine learning, and more recently to deep learning which is better suited to visual data. Through supervised learning, deeper and deeper neural network models are now able to learn the complex patterns found in large amounts of data, thus they are a good fit for solving computer vision problems.

One such problem is object detection, which consists of locating and classifying objects in an image. This problem is relevant in many domains such as automotive, logistics or even assistive technologies. Also, it has more potential than simple classification, because the objects are localized and this forces the learning algorithm to look for the very specific patterns that describe the objects, as opposed to classification where other patterns might be learned if the data distribution is unbalanced.

Given the potential of deep neural networks in solving the problem of object detection, the purpose of this paper is to study the recent architectural trends in building object detectors.

I wish to present in week 9.

[4] [2] [3] [1] [5]

2. Other sections to be added

3. Discussion

4. Conclusions and future work

References

[1] Feng, C., Zhong, Y., Gao, Y., Scott, M.R., Huang, W., 2021. TOOD: Task-aligned One-stage Object Detection, in: IEEE/CVF International Conference on Computer Vision (ICCV), IEEE Computer Society. pp. 3490–3499.

[2] Girshick, R., Donahue, J., Darrell, T., Malik, J., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 580–587.

[3] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., Berg, A.C., 2016. SSD: Single Shot MultiBox Detector, in: European conference on computer vision, Springer. pp. 21–37.

[4] Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You Only Look Once: Unified, Real-Time Object Detection, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779–788.

[5] Tian, Z., Shen, C., Chen, H., He, T., 2019. FCOS: Fully Convolutional One-Stage Object Detection, in: Proceedings of the IEEE/CVF international conference on computer vision, pp. 9627–9636.