COMP 576 Final Project

Dynamic Vehicle Detection

Shanli Ding, Mingzhao Liu, Yikun Li

Nov 2021

Abstract

Our team is aiming to train a vehicle detection model with Region-based Convolutional Neural Network(R-CNN). Firstly, we will train the model with static vehicle images, including 10 classes total 1142 cases, and train-test split in around 8:2 proportion. The data set is found in Kaggle website, you can find the reference at the bottom of this proposal. We plan to train and refine our model firstly on this smaller dataset. And then use the refined model to run on RGB CIFAR-10 dataset, 60000 cases with train-test 5:1, and improve the accuracy of the classification one step further.

Background

Computer vision is defined as a field of study that seeks to develop techniques to help computers acquire, process, analyze and understand digital images or videos, which is an interdisciplinary field that has attracted huge attention in recent years, and self-driving vehicles have come to the main stage [1]. Another integral part of computer vision is object detection, which consists of two separate tasks that are classification and localization [2] and aiding in pose estimation, vehicle detection, surveillance, etc. The purpose of our project is to do vehicle detection, which includes 10 classes of different types of vehicles based on a Region-based Convolutional Neural Network (R-CNN) model. The key concept behind the R-CNN series is region proposals. Region proposals are used to localize objects within an image.

State of art method

- most common method: R-CNN
- Potential alternative method: Faster R-CNN and VGG16

Broader scope and Goal

One of many so-called artificial intelligence or machine learning, which describes a scene as the same as a human's thinking logic. We have already done the beginning step in CNN and RNN to recognize the image with input-labels by CIFAR-10 grayscale dataset. Therefore, Our goal is to go further and do a large-scale RGB 3-channel image classification with higher accuracy and tolerance in a chaotic scene, for example detecting human beings and vehicles during rush hour in the morning. We are expecting to train an R-CNN model at first to do the still object detection, and next we will try to implement a faster R-CNN[3] model to boost our accuracy. Lastly, we will iterate the process several times to enhance our model for image classification.

Proposed solution and major contributions

- We will introduce R-CNN and Faster R-CNN to large scale image classification. Furthermore, we will contribute also on video tracking and in ~30FPS.
- Our model will support high resolution images with noisy background classification, instead of CIFAR-10 low resolution and clear background.

Proposed experiments-datasets and Model

Datasets

We will use the “Vehicle Dataset”[4] and RGB CIFAR-10 as our dataset. “Vehicle Dataset” contains 10 classes in total 1142 cases, and the train-test split in around 8:2 proportion. We plan to train and test our model firstly on this smaller dataset. And then use the polished model to run on RGB CIFAR-10 dataset, 60000 cases with train-test 5:1, and improve the accuracy of the classification one step further.

Model

Since the nature of the problem is image classification, there are lots of efficient methods to accomplish it. And most common and effective methods are various CNN-based models. In our experiment setup, we will implement the basic R-CNN model similar to the structure shown in Figure 1 and then test the model with bigger datasets, such as CIFAR-10 to evaluate the accuracy of the model. If enough time is provided, we will try to implement faster R-CNN and make the comparison with these two models.

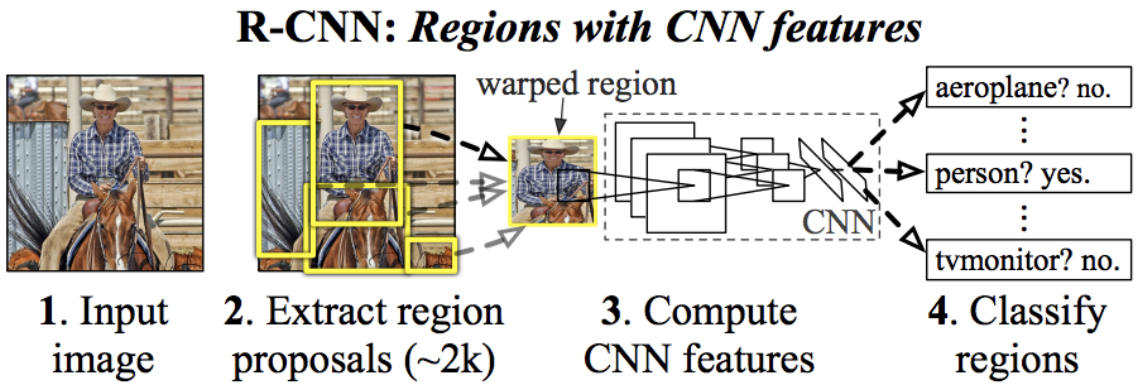

Figure 1. principle of R-CNN implementation[5]

Project execution plan

The main goal and targets for this project are:

- Survey the novel R-CNN implementation and understand their designs and implementation methods
- Implement the basic R-CNN model and faster R-CNN if time provided
- Evaluate the model with the two datasets mentioned above
- Reconstruct and model with better accuracy
- Formulate the final report

Feasibility and limitations of approach

Due to the limitation of the computation power, we plan to mainly work on the Google collabtory using Tensorflow and Keras to train and test the model.

Potential impact

Vehicle detection technology can be used in many different fields. For instance, the police can use this technique to distinguish the vehicle type in order to lock the criminal's position really quickly. Another application of this technology is that transportation surveillance is able to recognize the illegal vehicle on the specific road.

References

[1] Rohith Gandhi. R-CNN, Fast R-CNN, Faster R-CNN, YOLO - Object Detection Algorithms. Jul 9. 2018

[2] Chingis Oinar. Object Detection Explained: R-CNN. Mar. 20, 2021.

[3] Espinosa J.E., Velastin S.A., Branch J.W. (2017) Vehicle Detection Using Alex Net and Faster R-CNN Deep Learning Models: A Comparative Study. In: Badioze Zaman H. et al. (eds) Advances in Visual Informatics. IVIC 2017. Lecture Notes in Computer Science, vol 10645. Springer, Cham. https://doi.org/10.1007/978-3-319-70010-6_1
[4] https://www.kaggle.com/rishabkoul1/vechicle-dataset
[5]https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms-36d53571365e