


Chapter 19

Rifle Detection and Performance Evaluation Using Deep Learning Frameworks

Adeyemi Abel Ajibesin

 <https://orcid.org/0000-0001-6518-0231>
American University of Nigeria, Nigeria

Doken Edgar

American University of Nigeria, Nigeria

ABSTRACT

Deep learning models being used to improve human life has been an ongoing domain of research. Violence, especially with the proliferation of arms, has been on the increase worldwide. Many tragedies have occurred right across the globe, leading to people losing their lives as a result of being shot at with guns. This research sought to use deep learning frameworks to detect rifles in images and assess their performance based on the metrics of accuracy and F1 score. The study used a combination of images from Google open images and other sources to form a dataset of 2105 images; 1857 of those was used to train YOLOv3 and RetinaNet models to detect rifles, using Darknet-53 and ResNet50 respectively as the backbone networks. The models were evaluated after training using a test dataset containing 248 images, both the training and evaluation of the models were carried out using scripts written in Python. The results obtained showed that YOLOv3 had better output in terms of accuracy, precision, recall, and consequently, the F1 scored better than RetinaNet

INTRODUCTION

Overview

Computer vision is a subarea of artificial intelligence that is focused on the study and automation of visual perception tasks (Rodriguez, 2020). For humans, it is quite a simple and straightforward task to

DOI: 10.4018/978-1-6684-6937-8.ch019

identify and recognize familiar objects, even in scenarios that may be unfriendly due to variations in light or distortions due to movements (Sharma & Thakur, 2017). However, these simple tasks of recognizing objects are very difficult for computers (Sharma & Thakur, 2017). As advances are being made in the field of artificial intelligence, being able to have computers recognize objects in videos and images is quite important. This process is known as object recognition (Zhao et al., 2017). Though sometimes used interchangeably, detecting, and classifying objects from an image are not the same. Classification is the ability to say what object is present in an image, that is, labeling the content of an image. Detection goes a little beyond that: it is not enough to say what object is in an image, but where the said object is located within the image (Rodriguez, 2020). For example, an algorithm can determine that an image contains a car, but with detection, it would take it further and provide a bounded box of where the car is situated within the image. Hence a combination of object classification and detection provides a more complete understanding of an image (Zhao et al., 2017).

Traditionally in machine learning, image classification was based on feature extraction of specific features from the training set. This makes it unable to extract other types of features from the training data (Krishna et al., 2018). This shortfall is addressed via the use of deep learning (Lee et al., 2009). With a deep learning model, the model isn't given a specific set of features, but it's allowed to learn through its method of computing. The deep learning model mimics the structure of the human brain and utilizes several algorithms expressed as Artificial Neural Networks (Krishna et al., 2018).

In the field of computer vision, object classification and detection have been studied and applied in many fields like autonomous vehicles, facial recognition (Bashbaghi et al., 2019), agriculture (Geffen et al., 2020), medicine and health and so many others. Object detection is one of the important features needed in the navigation of autonomous vehicles, without which it will be almost impossible to have a vehicle that is safe for passengers in the vehicle as well as other road users (Lewis, 2016). In virtually all of the fields that computer vision is applied, one important attribute to have is accuracy (Lewis, 2016).

Another important application of object detection is surveillance and security (Kanehisa & Neto, 2019). Due to a lot of areas making use of closed-circuit television and drones for surveillance, manning the output can be overwhelming and tedious for the operators. This is where object detection models can be used to assist security agencies to anticipate dangerous situations before they occur. The models can be used to detect firearms (Kanehisa & Neto, 2019) and other dangerous items from an image or video input, and then alert security agencies to respond before the situation escalates. Having such a system in place can sometimes be the difference between life and death situations, because, although human visual perception is usually quick and precise, it is prone to error if someone keeps watching the same thing for a long period (Narejo et al., 2021). To curb this limitation, the availability of huge datasets, faster processing, and computing capabilities using GPUs and machine learning algorithms are being considered to effectively develop surveillance systems that will function for long periods with high accuracy. Advances in machine learning and also image processing algorithms such as the usage of convolutional neural networks (CNNs) have made it possible to develop smart and effective surveillance systems. Convolutional neural networks can automatically extract features from data (An et al., 2012), unlike traditional means of extraction where the features are manually selected (Searle et al., 1997).

The state-of-the-art models used in object detection can be divided into two categories: one-stage and two-stage detectors. The two-stage detectors generally have higher accuracy than the one-stage detectors, but they use more computational power and take more time to execute than the one-stage detectors. But this does depend on what convolutional network is used as the backbone network as well as other configurations (Garcia et al., 2021). Two-stage detectors are so-called because the frameworks'

29 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/rifle-detection-and-performance-evaluation-using-deep-learning-frameworks/318075

Related Content

Neuro-Immune Model Based on Bio-Inspired Methods for Medical Diagnosis

Fatiha Djahafiand Abdelkader Gafour (2022). *International Journal of Ambient Computing and Intelligence* (pp. 1-18).

www.irma-international.org/article/neuro-immune-model-based-on-bio-inspired-methods-for-medical-diagnosis/293176

Different Approaches for Cooperation with Metaheuristics

José M. Cadenas, Ma Carmen Garrido, Enrique Muñoz, Carlos Cruz-Corona, David A. Peltaand José L. Verdegay (2009). *Encyclopedia of Artificial Intelligence* (pp. 480-487).

www.irma-international.org/chapter/different-approaches-cooperation-metaheuristics/10290

An Efficient Kinetic Range Query for One Dimensional Axis Parallel Segments

T. Hemaand K. S. Easwarakumar (2018). *International Journal of Intelligent Information Technologies* (pp. 48-62).

www.irma-international.org/article/efficient-kinetic-range-query-one/190654

Behavioral Analytics of Consumer Complaints

Md Shamim Hossain (2023). *AI-Driven Intelligent Models for Business Excellence* (pp. 42-67).

www.irma-international.org/chapter/behavioral-analytics-of-consumer-complaints/315393

3D Gesture Recognition Based on Handheld Smart Terminals

Yunhe Li, Yi Xieand Qinyu Zhang (2018). *International Journal of Ambient Computing and Intelligence* (pp. 96-111).

www.irma-international.org/article/3d-gesture-recognition-based-on-handheld-smart-terminals/211174