ADVANCED MACHINE LEARNING – FINAL PROJECT REPORT

TOPIC: OBJECT DETECTION

NAME OF THE PROFESSOR: DR. CHAOJIANG WU
BA-64061-001

NAME OF THE STUDENT: ARRABACHALA SISIRA
STUDENT ID: 811244504

DEEP LEARNING
OBJECT DETECTION in COMPUTER VISION

ABSTRACT:

Object detection is the main task in computer vision, and it identifies and locates objects in images or videos. In this Research Paper, we give a detailed analysis of recent improvements in deep learning-based object detection. We will discuss the overview of current state of the art object detection models, along with their modifications. We also will discuss the most recent developments in object detection datasets, training techniques, and evaluation measures. And we also cover the limitations and restrictions of deep learning-based object detection and potential solutions to these issues.

INTRODUCTION:

Object detection algorithms locate and classify items in an image or video using machine learning and deep learning methods. The purpose is to detect and accurately find all instances of specified items, such as cars, pedestrians, or animals, and then draw bounding boxes around them. Deep learning models in object detection are known to perform better than classic computer vision algorithms in the tasks. Object Detection is a key task in computer vision, and it has a wide range of applications including self-driving cars, tracking, surveillance systems and picture and video analysis and many more. Deep learning-based algorithms have made great progress in object detection tasks in recent years, overcoming classical methods. Convolutional neural networks (CNNs) have become the preferred method for object detection applications from the development of deep learning.

The state-of-the-art deep learning object detection models:

- Faster R-CNN (Region-based Convolutional Neural Network): This model combines a region proposal network (RPN) with a Fast R-CNN network to achieve high accuracy and speed in object detection. Faster R-CNN has become the standard model for object detection.
- Mask R-CNN (Mask Region-based Convolutional Neural Network): This model improves Faster R-CNN by combining a branch for predicting object masks with the existing branch for bounding box identification. Mask R-CNN proved great performance in instance segmentation tasks.

- YOLO (You Only Look Once): This model is a real-time object detection system that predicts bounding boxes and class probabilities directly from full images in one evaluation. It has achieved a good balance between accuracy and speed.

In the object detection, the Faster R-CNN, Mask R-CNN, and YOLO models performed excellently. CNNs can learn features automatically from raw data and resolve a variety of problems such as scale variation, occlusion, and background clutter. Faster R-CNN combines a region proposal network (RPN) with a Fast R-CNN network to create and classify object proposals. Object detection in a single feedforward pass is performed by single-stage models such as YOLO and SSD (Single Shot MultiBox Detector), making them fast and suitable for real-time applications. Faster R-CNN and Mask R-CNN are two-stage models that detect objects in two stages and those are region proposal and classification. Even though Two stage models are more accurate they are slower than single stage models.

LITERATURE REVIEW:

Deep learning-based object detection models require a huge amount of labeled data for training. Some data augmentation strategies, such as random cropping, scaling, and flipping, have been proposed to solve the scarcity of labeled data. Transfer learning has also been widely used to train object detection models, in which a pre-trained CNN model is employed as an extractor of features and only the classification head is modified on the target dataset. Some advanced training procedures, such as multi-task learning, where the model is trained to execute multiple related tasks simultaneously, and curriculum learning, where the model is trained on progressively harder instances, have been developed in recent years.

Evaluation metrics play an important role in determining the performance of object detection models. Precision, recall, and average precision (AP) are the most often used measures for object detection. Precision is the percentage of correctly detected objects among all detected objects, whereas recall is the percentage of correctly detected objects among all ground truth objects. At various confidence levels, AP is a common statistic that combines precision and recall. Some advanced evaluation measures, such as mean average precision, which computes the AP over many object categories, and Intersection over Union thresholds, have been presented in recent years.

The development of deep learning-based object detection has been greatly helped by the availability of large-scale object detection datasets. Popular datasets for object detection include ImageNet, COCO (Common Objects in Context), and PASCAL VOC (Visual Object Classes). Millions of annotated photos make up these datasets, which have been used to develop and test deep learning-based object detection models. Open Images and Visual Genome are two recent datasets that include more varied object categories and intricate object interactions than previous ones.

Industry Applications of Object Detection based Deep Learning:

a. Healthcare: Deep learning-based object detection can help in a variety of medical imaging applications, including tumor diagnosis, abnormality identification, and organ segmentation. Deep learning models, for example, may analyze MRI data to detect brain cancers and segment tumor regions for radiation therapy planning. They can also examine X-rays for fractures, lung nodules, and other irregularities. Deep learning-based object detection can help in disease detection and treatment, improving patient outcomes.

b. Transportation: Applications for object detection based on deep learning in transportation include object tracking, traffic management, and self-driving cars. In self-driving cars, deep learning models can analyze sensor data from cameras and radars to find and classify roadside items including other cars, pedestrians, and traffic signals. This might boost the security and effectiveness of transportation infrastructure. Deep learning models in the transportation industry can improve safety, reduce traffic congestion, and even improve the traffic flow.

c. Security: Deep learning-based object detection is very useful for many security applications like surveillance systems, threat detection, facial recognition and many more. Deep learning models in surveillance systems can discover and track persons and items of interest by analyzing streamed footage from CCTV cameras. This can help detect possible attacks and prevent security breaches. Deep learning models in threat detection can scan data from numerous sources, such as social media, CCTV cameras, and sensor networks, to identify potential security concerns. Deep learning-based object detection can also help with facial identification, such as detecting people in a crowd or validating people's identities in a safe environment.

Future Developments and Current Limitations in Object Detection:

Despite significant developments in deep learning-based object detection, there are still specific challenges and limitations. The imbalance between positive and negative instances is one of the primary issues, which might lead to biased models. Another difficulty is that some object detection algorithms have significant computational and memory needs, limiting their adoption in resource-constrained devices. Furthermore, some items, such as small and strongly obstructed objects, remain difficult to recognize using deep learning.

Another deep learning model's current problems in object detection is their inability to handle occlusion and scale variation well. Another main thing is that the requirement for huge amounts of labeled data, which can be costly and requires a lot of time to collect. Approaches like transfer learning and semi-supervised learning may aid in minimizing the amount of labeled data required. Deep learning models have the ability to become even more

effective and accurate in object detection tasks in the future. The use of attention mechanisms in deep learning models for object detection is one area of study.

The development of more effective and lightweight object detection models is another field of research that may be helpful in environments with limited resources, such as mobile devices and drones. Additionally, for tasks like object detection, deep learning models have the capacity to incorporate many modalities, like auditory and visual input. This could allow models to recognize and detect objects utilizing information from numerous sources, leading to more accurate and trustworthy object detection.

Deep learning models have performed well in object detection tasks, but there is still an opportunity for improvement. The field's future advancements will probably focus on improving deep learning models that can handle a variety of object detection issues by making them more accurate and efficient.

Potential Solutions:

Some potential solutions have been presented to overcome these issues and constraints. Some approaches, such as focal loss and Online Hard Example Mining also called OHEM, have been proposed to address the imbalance problem by adjusting the loss function and focusing on hard instances. In order to help with occlusion and size shift, models can utilize attention techniques to concentrate on particular areas of a picture. Some strategies, such as model compression and quantization, have been developed to minimize model size and complexity in order to reduce computational costs. Also, certain strategies for improving the detection of small and strongly occluded objects, such as objectness priors and attention algorithms, have been presented.

Conclusion:

We offered a detailed analysis of current improvements in deep learning-based object detection in this research paper. We discussed the most recent object identification models, datasets, training methodologies, and evaluation measures. We even looked at the restrictions and limits of deep learning-based object detection as well as possible solutions for these issues. In recent years many object detection deep learning model methods have come into the market and have improved a lot with many advancements. To improve the accuracy, effectiveness, and adaptability of deep learning-based object detection models, we still have to do more research.

REFERENCES AND CITATIONS:

- https://machinelearningmastery.com/object-recognition-with-deep-learning/
- https://www.mathworks.com/discovery/object-detection.html#:~:text=Object%20detection%20is%20a%20computer,learning%20to%20produce%20meaningful%20results.
- https://en.wikipedia.org/wiki/Object_detection
- https://www.fritz.ai/object-detection/
- https://www.upgrad.com/blog/ultimate-guide-to-object-detection-using-deep-learning/#:~:text=Deep%20learning%20algorithms%20like%20YOLO,learning%20algorithms%20for%20object%20detection.
- https://www.sciencedirect.com/science/article/pii/S1877050918308767
- https://viso.ai/deep-learning/object-detection/
- https://analyticsindiamag.com/top-8-algorithms-for-object-detection/

- https://www.analyticsvidhya.com/blog/2022/03/a-basic-introduction-to-object-detection/
- https://www.g2.com/articles/object-detection
- https://paperswithcode.com/task/object-detection
- https://arxiv.org/pdf/1807.05511.pdf
- https://www.researchgate.net/publication/338253407_A_STUDY_ON_OBJECT_DETECTION
- https://towardsdatascience.com/12-papers-you-should-read-to-understand-object-detection-in-the-deep-learning-era-3390d4a28891
- https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4042549
- https://www.youtube.com/watch?v=TB-fdISzpHQ&ab_channel=MichiganOnline
- https://www.youtube.com/watch?v=ATw1Dy4p1GU&ab_channel=RanjanSharma