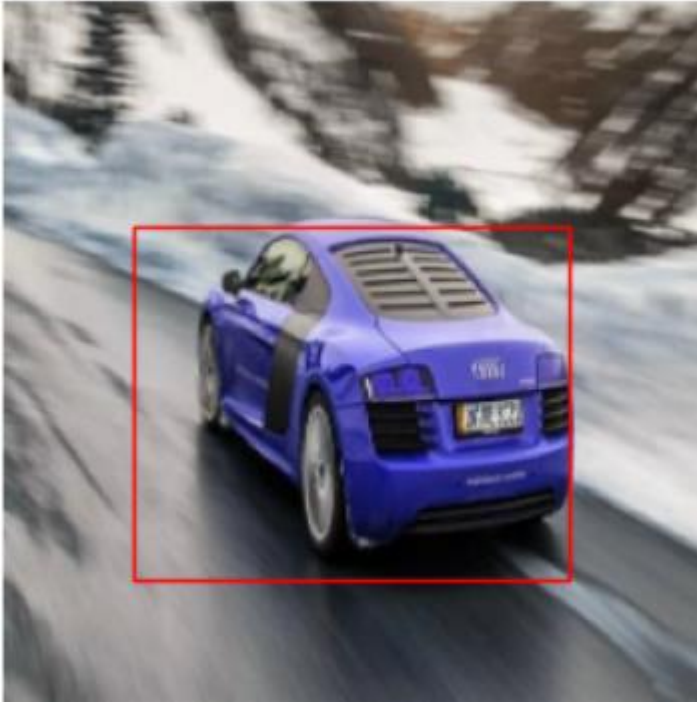


Rapport car detection



Gemaakt door Joshua Veens

Inhoudsopgave

Introductie:	3
Background:	3
Methodebeschrijving:	3
Experimentenbeschrijving:	4
Experiment 1: Meerdere auto's in één afbeelding	4
Experiment 2: Auto's van bovenaf	4
Conclusie/discussie:	4
Literatuurlijst.....	5

Introductie:

In mijn eindopdracht wordt een algoritme ontwikkeld om voertuigen te detecteren in afbeeldingen. Het detecteren van voertuigen is een belangrijke taak in verschillende toepassingen, zoals verkeersmonitoring, beveiliging en autonome voertuigen. In dit project wordt gebruik gemaakt van de VGG16-architectuur, een populaire deep learning-technologie voor objectdetectie.

Background:

Het detecteren van voertuigen in afbeeldingen is een uitdagende taak vanwege de variabiliteit in de grootte, oriëntatie en positie van voertuigen in verschillende omgevingen. Een veelgebruikte methode voor het detecteren van objecten in afbeeldingen is het gebruik van Convolutional Neural Networks (CNN's). Een CNN is een type neuraal netwerk dat geïnspireerd is op de visuele cortex van dieren. Het maakt gebruik van convolutielagen om automatisch kenmerken uit de afbeelding te extraheren en classificatielagen om de objecten te detecteren.

Een voorbeeld van een paper dat gebruik maakt van een CNN voor het detecteren van auto's is "You Only Look Once: Unified, Real-Time Object Detection" van Joseph Redmon et al. (2016). Het voorgestelde netwerk, genaamd YOLO, is in staat om objecten in real-time te detecteren en te classificeren met behulp van slechts één voorwaartse doorloop van het netwerk. Het netwerk maakt gebruik van een grid-celstructuur om de afbeelding te partitioneren en gebruikt vervolgens een convolutionele laag om voor elke cel een set van bounding boxes en classificatie scores te voorspellen. De voorspellingen van alle cellen worden vervolgens gecombineerd om de uiteindelijke detectie te verkrijgen.

Een ander voorbeeld van een paper dat gebruik maakt van CNN's voor het detecteren van auto's is "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks" van Shaoqing Ren et al. (2016). Het voorgestelde netwerk, genaamd Faster R-CNN, maakt gebruik van een Region Proposal Network (RPN) om kandidaat-regio's te genereren waarin objecten kunnen worden gedetecteerd. Vervolgens worden deze kandidaat-regio's gebruikt om objecten te detecteren met behulp van een CNN. Het netwerk maakt gebruik van een zogenaamde "region of interest pooling layer" om het beeld van elke regio om te zetten in een vast formaat, onafhankelijk van de grootte en aspectverhouding van de regio.

Op pagina 6 van "You Only Look Once: Unified, Real-Time Object Detection" van Joseph Redmon et al. (2016). Worden deze 2 algoritmes met elkaar vergeleken.

Methodebeschrijving:

Het algoritme maakt gebruik van de VGG16-architectuur voor objectdetectie. Het is getraind op een dataset met afbeeldingen van voertuigen in verschillende omgevingen, zoals steden en snelwegen. Het algoritme is in staat om auto's te detecteren en hier een vierkant omheen te tekenen. (Bounding box) Het maakt gebruik van een neuraal netwerk om de afbeeldingen te analyseren en te detecteren waar de auto's zich bevinden.

Experimentenbeschrijving:

Experiment 1: Meerdere auto's in één afbeelding

Er is een set afbeeldingen getest waarin meerdere auto's aanwezig zijn. Meestal weet het algoritme de auto's in de afbeelding te vinden, maar het heeft moeite wanneer er meer dan twee auto's in de afbeelding staan. Het VGG16-algoritme presteert hier beter dan het eenvoudige Sequential model. De bounding boxes worden nauwkeuriger geplaatst, en de gemiddelde afstand tot de ground truth is kleiner.

Experiment 2: Auto's van bovenaf

In dit experiment zijn afbeeldingen gebruikt waarin voertuigen van bovenaf zichtbaar zijn, bijvoorbeeld vanuit een droneperspectief. Beide modellen hadden moeite met deze taak, vooral wanneer de voertuigen verder weg zijn of er door ruis van bijvoorbeeld bomen. Het Sequential model presteerde in deze situatie verrassend iets beter, hoewel beide modellen ver ondermaats presteerden.

Conclusie/discussie:

De resultaten tonen aan dat het algoritme over het algemeen goed presteert bij het detecteren van voertuigen in standaardomgevingen. Het VGG16-model laat betere prestaties zien dan het Sequential model, met name wanneer er meerdere voertuigen in beeld zijn. In complexere situaties zoals luchtbeelden of afbeeldingen met veel verschillende auto's, nemen de prestaties sterk af.

Hoewel de nauwkeurigheid bij eenvoudige beelden redelijk is, zijn er nog aanzienlijke verbeterpunten nodig om het algoritme betrouwbaar in de praktijk toe te passen. Verder onderzoek zou zich kunnen richten op het uitbreiden van de dataset, het toevoegen van moeilijkere trainingsbeelden en het experimenteren met geavanceerdere netwerken zoals YOLOv5 of EfficientDet. Daarnaast blijft rekenkracht een belangrijk aandachtspunt; VGG16 is relatief zwaar, wat het moeilijk maakt om real-time prestaties te behalen zonder GPU.

Literatuurlijst

- 1): ["You Only Look Once: Unified, Real-Time Object Detection"](#)
- 2): ["Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks"](#)