

ONDERZOEKSVOORSTEL

Een proof of concept van object detection waarbij verborgen relaties kunnen worden vastgesteld in een GIS georiënteerde casestudy adhv Deep Learning.

Bachelorproef, 2023-2024

Jona Fouquaert

E-mail: jona.fouquaert@student.hogent.be

Co-promotor: S. Beekman (Synalco, sigrid.beekman@synalco.be)

Samenvatting

Geografische informatiesystemen (GIS) zijn digitale systemen die geografische gegevens analyseren, visualiseren en beheren om ruimtelijke patronen en relaties te begrijpen is geworteld in de geografie

Keuzerichting: AI & Data Engineering

Sleutelwoorden: object detection, GIS, deep learning

Inhoudsopgave

1	Introductie	1
2	State-of-the-art	1
3	Methodologie	1
4	Verwacht resultaat, conclusie	2

1. Introductie

Geografische informatiesystemen (GIS) zijn digitale systemen die geografische gegevens analyseren, visualiseren en beheren om ruimtelijke patronen en relaties te begrijpen. Het wordt vandaag de dag door bedrijven gebruikt in verschillende sectoren. Nu er vraag is naar GIS-applicaties, zijn er ook ondernemingen die zich specialiseren in deze trend, zoals bijvoorbeeld het bedrijf geoAI. Zij combineren geo IT met het experimentele veld van machine learning. Dit kan ingezet worden om verschillende doeleinden te bereiken, zoals bijvoorbeeld automatische kaartgeneratie, voorspellende modellering, Het thema waar wij ons op gaan focussen is geospatiale patroonherkenning. Dit gaat over het ontdekken van verborgen relaties tussen verschillende objecten. Hier komt de sterkte van AI naar boven. Het is in staat patronen te ontdekken die wij als mensen moeilijk begrijpen. Meer specifiek ga ik remote sensing classificatie toepassen op lidar data om zo vooraf bepaalde klassen te visualiseren. Het model dat ik ga gebruiken is het Convolutional Neural Network (CNN), een deep learning-algoritme dat in staat is hiërarchische kenmerken in afbeeldingen te visualiseren. CNN's delen de afbeeldingen op in kleine delen a.d.h.v. convolutionele lagen. Hiermee kunnen ze een afbeelding volledig apart analyseren. Voor dit onderzoek maak ik gebruik van het opensource

QGIS-programma en de Semi-Automatic Classification en clusteranalyse-plugin. Concreet is de onderzoeksvraag: Hoe kunnen verborgen relaties tussen verschillende geospatiale objecten worden blootgelegd met behulp van de "Deepness: Deep Neural Remote Sensing" "clusteranalyse" plug-ins in QGIS, gebruikmakend van Convolutional Neural Networks (CNN) en clusteralgoritmes? Deelvragen: - Wat zijn de optimale architectuur en hyperparameters voor het CNN-algoritme bij het identificeren van relaties tussen objecten? - Zijn er factoren zoals terreinomstandigheden, seizoenen of landgebruik die de prestaties van het CNN-algoritme beïnvloeden bij de identificatie van bepaalde objecten? - Welke parameters en instellingen zijn van invloed op de resultaten van de clusteringalgoritmes in QGIS? - In hoeverre verschillen de resultaten van de Agglomerative Hierarchical en K-Means clusteringalgoritmes op onze dataset?

2. State-of-the-art

Net zoals bij andere competenties speelt machine learning een alsmaar grotere rol in geografische applicaties. Convolutional Neural Networks (CNN's)

3. Methodologie

Dit onderzoek mond uit op het uitvoeren van verschillende experimenten om zo te achterhalen of er verborgen relaties tussen objecten kunnen worden waargenomen. Een experiment wordt opgedeeld in drie grote lijnen: preprocessing, object detection en post-analyse. Preprocessing wordt uitgevoerd door de pixels en resolutie van de input layer te specificeren. Vervolgens wordt er

een Open Neural Network Exchange (ONNX) model gekozen dat getraind is voor een gelijkaardig project. De parameters van het model zijn dan meestal reeds gefinetuned. In de plugin-interface kun je ze nog aanpassen. De output van het gekozen model is een vectorlaag met de verschillende objecten gemarkeerd. Deze vectorlaag wordt de input van de tweede plugin die wordt gebruikt: clusteranalyse. Deze plugin voert een unsupervised cluster algoritme (Agglomerative Hierarchical of K-Means) uit op een gekozen vectorlaag. Ook hier wordt nog wat preprocessing uitgevoerd. Door feature cleaning toe te passen vallen nutteloze features weg. Vervolgens kun je op de data ook nog scaling toepassen (Normalized of Standardized). Nadien worden het algoritme en het aantal clusters gekozen. De output is een nieuwe vectorlaag met de clusters afgebeeld. Ten slotte maak ik gebruik van de analysetools ingebouwd in QGIS, door vectorpunten te verbinden en zo een overzicht te krijgen van mogelijke relaties tussen verschillende objecten.

4. Verwacht resultaat, conclusie