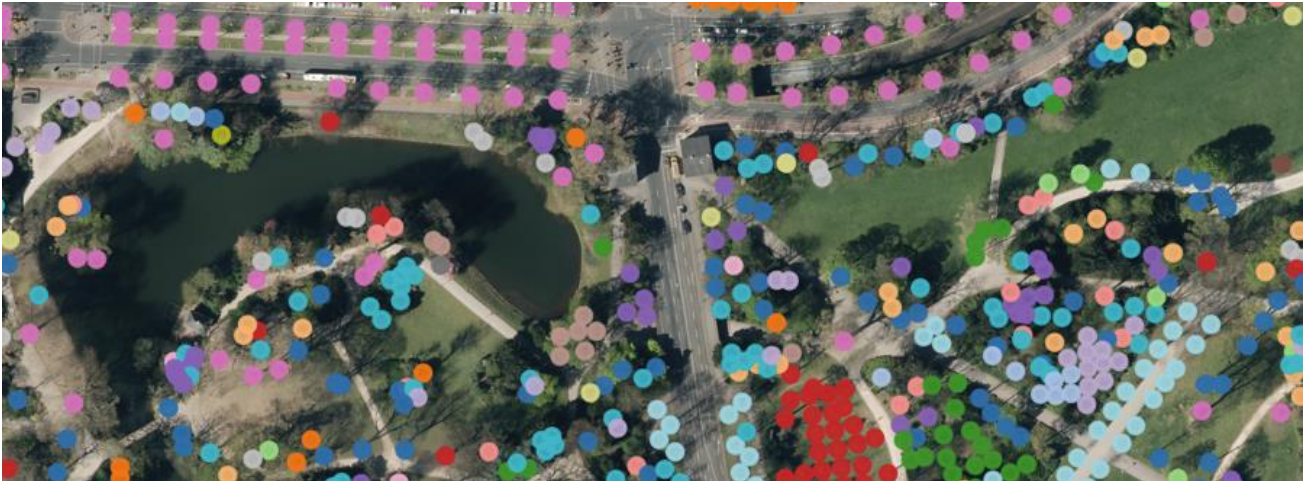


Ausschreibung Masterarbeit

Pretraining von Dual-Stream-Architekturen zur Verarbeitung hochauflösender Luftbilder



Ausgangslage

Moderne Fernerkundungssysteme – darunter Drohnen, Flugzeuge und Satelliten – liefern hochauflösende Bilddaten, die in der Stadt- und Raumplanung, der Katastrophenschutzvorbereitung, der Landnutzungsklassifikation sowie in der Umweltüberwachung eingesetzt werden. Da die manuelle Annotation solcher Daten mit erheblichem Zeit- und Kostenaufwand verbunden ist, liegt ein Großteil der verfügbaren Luftbilddaten unannotiert vor. Selbstüberwachte Pretraining-Verfahren – etwa Masked Autoencoder (MAE) oder kontrastive Methoden wie MoCo und DINO – bieten die Möglichkeit, leistungsfähige Bildrepräsentationen allein aus unannotiertem Bildmaterial zu erlernen und diese anschließend mit vergleichsweise wenigen Labels auf spezifische Aufgaben zu übertragen. Hochauflösende Luftbilder stellen dabei besondere Anforderungen an die Modellarchitektur: Relevante Objekte – etwa einzelne Bäume, Fahrzeuge oder Gebäudekanten – sind im Verhältnis zur Gesamtszene sehr klein, sodass Standard-Transformerarchitekturen mit festen Patch-Größen feine räumliche Details verlieren können. Dual-Stream-Architekturen begegnen diesem Problem, indem sie zwei parallele Verarbeitungspfade kombinieren: einen hochauflösenden Pfad zur Erhaltung feiner Strukturen und einen niedrigauflösenden Pfad zur Erfassung globaler Kontextinformation. Solche Architekturen haben sich in der Bildverarbeitung natürlicher Szenen bereits bewährt, ihr Potenzial im Kontext des selbstüberwachten Pretrainings auf Fernerkundungsdaten ist jedoch bislang kaum untersucht.

Problemstellung

Bestehende Pretraining-Ansätze für Fernerkundungsbilder basieren überwiegend auf Standard-Encoder-Architekturen wie Vision Transformern (ViT) oder ResNet-Varianten, die für eine einheitliche, globale Merkmalsextraktion ausgelegt sind. Für hochauflösende Luftbilder mit stark variierenden Objektgrößen und feinen räumlichen Details sind diese Architekturen jedoch nur bedingt geeignet.

Dual-Stream-Architekturen versprechen hier Abhilfe, werfen jedoch eine Reihe offener Fragen auf: Wie lassen sich selbstüberwachte Pretraining-Ziele – insbesondere rekonstruktionsbasierte Verfahren wie MAE – sinnvoll auf eine Zwei-Pfad-Architektur übertragen? Wie sollten die beiden Ströme während des Pretrainings miteinander interagieren, etwa durch Stream-übergreifende

Ansprechpartner

Moritz Weiß | E-Mail: weiss@uni-wuppertal.de

Aufmerksamkeitsmechanismen oder gemeinsame Rekonstruktionsziele? Und lassen sich die so erlernten Repräsentationen effektiv auf nachgelagerte Aufgaben wie semantische Segmentierung oder Objektdetektion übertragen?

Vorgehensweise und erwartete Ergebnisse

Die Arbeit soll den Einsatz selbstüberwachter Pretraining-Strategien für Dual-Stream-Architekturen auf hochauflösenden Luftbilddaten systematisch untersuchen.

Contact Person

Moritz Weiß | **E-Mail:** weiss@uni-wuppertal.de