

Zusammenfassung

Haze and cloud-removal in multi-spectral optical satellite images for landcover classification

Masterarbeit von Ziqi Gu

Motivation & Relevanz:

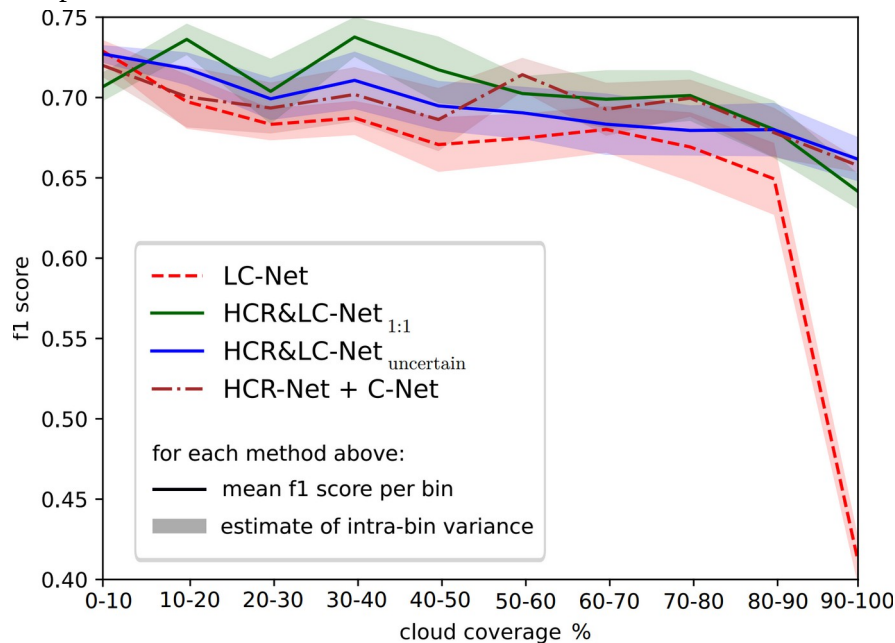
Optische Satelliten leisten einen wichtigen Beitrag zum stetigen Beobachten unseres Planeten. Allerdings sind über die Hälfte aller Satellitenaufnahmen von Wolken und anderen atmosphärischen Störungen betroffen. Der aktuelle Forschungsstand demonstriert, dass neuronale Netze trotz dessen etablierte Fernerkundungsanwendungen durchzuführen vermögen, und bewölkte Pixel schlichtweg ignorieren lernen. Dies wirft die Frage auf, ob es von zusätzlichem Vorteil wäre wolkenbefangene Pixel und deren verdeckte Information explizit zu rekonstruieren um dann mit den dazugewonnenen Informationen weiterzuarbeiten. Das Ziel dieser Masterarbeit ist darauf Antworten zu finden. Während vorige Studien entweder bloß bewölkte Satellitenbilder rekonstruierten oder ungeachtet dessen eine Remote Sensing Anwendung lösten, kombiniert die vorliegende Arbeit beides explizit und untersucht die sich daraus ergebenden Vorteile. Als beispielhafte Anwendung wird die Klassifizierung von teils bewölkten Satellitenaufnahmen in Landbedeckungskategorien adressiert. Das Training moderner neuronaler Netze zwecks solcher Anwendungen involviert üblicherweise große Datensätze zahlreicher beschrifteter Zielwerte, deren manuelle Annotation sehr kostspielig ist. Daher adressiert die vorliegende Arbeit abschließend das Problem, die benötigten Annotationen zu reduzieren, um die vorgeschlagenen Methoden auch in der Praxis anwendbar zu machen.

Ansatz & Methodik:

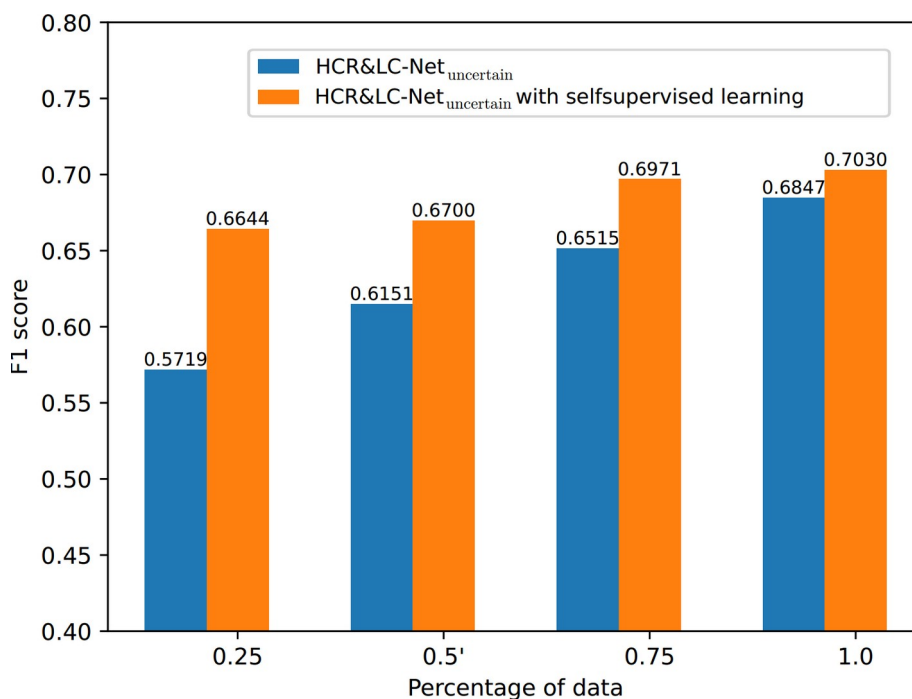
Zur Adressierung des beschriebenen Problems wird ein globaler Datensatz an Landflächen-aufnahmen verwendet und mit einem komplementären Datensatz bewölkter Beobachtungen kombiniert. Um den Vorteil der Bildrekonstruktion zu evaluieren, wird ein neues multi-task Netzwerk entworfen, das gleichzeitig sowohl Wolken entfernt als auch Landflächen klassifiziert. Beide Aufgaben werden dank eines stochastischen Ansatzes systematisch gemeinsam optimiert und das resultierende Netzwerk schließlich mit vergleichbaren Lösungen experimentell evaluiert. In der Analyse wird auf Fairness geachtet: Vergleichbare Methoden ähneln sich in ihrer Komplexität, der Anzahl lernbarer Parameter und ihren Bausteinen---entscheidender Unterschied ist „ob“ und „wie“ mittels einer expliziten Rekonstruktion bewölkter Pixel Einfluss auf die Klassifizierung von Landbedeckungen genommen wird. Zwecks weiterführender Analysen werden Experimente parametrisiert und so zum Beispiel untersucht, wie sich die Klassifizierungsgüte je Ansatz als Funktion des Wolkenbedeckungsanteils ändert. Abschließend wird das Problem adressiert, dass die bislang berücksichtigten neuronalen Netzwerke zwar leistungsfähig aber auch enorm datenhungrig sind: Um die Anzahl erforderlicher Landschafts-Annotationen in den Trainingsdaten zu reduzieren, wird ein self-supervised learning Ansatz erwogen. Konkret werden sogenannte pre-text tasks entwickelt, welche ein Netzwerk zunächst lösen lernt, bevor es sich der eigentlichen Aufgabe der Landschaftsklassifizierung widmet. Die pre-text Aufgaben entstehen aus einer intelligenten Augmentierung unannotierter Satellitenbeobachtungen, die in künstlich generierte Annotationen resultieren, welche die Netzwerke bereits vorab vorherzusagen lernen. Dadurch könnten neuronale Netzwerke nützliche Repräsentationen erwerben, welche für die eigentliche Problemstellung hilfreich sein mögen und zu einer Reduktion an letztlich erforderlichen Annotationen führen könnten.

Resultate & Fazit:

Die Auswertung auf dem globalen Datensatz von Landbedeckungskategorien zeigt, dass das vorgeschlagene Netzwerk Landschaften besser kategorisiert als vergleichbare Lösungen, einschließlich eines zweiteiligen Ansatzes der zunächst entwirrt und anschließend klassifiziert. Alle Netzwerke, die explizit bewölkte Pixel rekonstruieren, funktionieren besser als Ansätze welche bloß implizit lernen sie zu ignorieren. Figur 1 zeigt, dass letztere insbesondere bei starker Bewölkung scheitern während eine explizite Rekonstruktion Robustheit verleiht, selbst bei flächendeckender Bewölkung.



Wie mit neuronalen Netzwerken üblich, werden alle berücksichtigten Methoden auf einem sehr großen Datensatz mit über 100.000 Annotierungen trainiert. Figur 2 berichtet die Klassifizierungsleistung des bislang besten Netzwerks ohne und mit dem vorgeschlagenen self-supervised learning Ansatz. Beachtlicherweise können mit bloß 25% der Annotationen bereits rund 95% der ursprünglichen Leistung erzielt werden. Auch unter Berücksichtigung aller Annotationen ermöglicht die Verwendung des entwickelten pretext tasks eine weitere Steigerung der erbrachten Leistung.



Die Resultate der vorliegenden Masterarbeit wurden bereits Ende Juni im Workshop „[Multimodal Learning for Earth and Environment](#)“ der *IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR)* präsentiert. Eine erweiterte Fassung der veröffentlichten Studie befindet sich aktuell in Begutachtung in der wissenschaftlichen Fachzeitschrift *IEEE Transactions on Geoscience and Remote Sensing*. Die vorliegenden Ergebnisse sind vielversprechend und von wissenschaftlicher Relevanz, sodass mit vertiefender Forschung in die beschriebenen Themengebiete geplant wird.