

Monitoring der Vereisung von Schweizer Seen mit Webcams

Dieser Bericht basiert auf einem nationalen Projekt (initiiert von MeteoSchweiz) über See-Vereisung, auch im Rahmen internationaler Aktivitäten, die mit Klima und seinen Veränderungen zu tun haben. Klimaveränderungen werden oft durch Indikatoren beobachtet. See-Eis ist einer dieser (Teil-)Indikatoren. Für das See-Eis-Monitoring werden oft satelliten-basierte (optische) Fernerkundungsdaten mit ca. einem Bild pro Tag herangezogen. Limitierende Faktoren der automatischen Auswertung dieser Daten sind dabei die geometrische Auflösung (normalerweise ein Pixel auf dem Boden ist 250 m plus) und Wolken. Alternativ soll in diesem Beitrag das Potenzial des Webcam-basierten See-Eis-Monitorings aufgezeigt werden. Ziel ist dabei eine pixelweise Klassifikation mittels moderner Methoden des maschinellen Lernens, um die vereiste See-Oberfläche in einzelnen Webcam-Bildern zu bestimmen.

Cet article base sur un projet national (initié par MétéoSuisse) au sujet de la glaciation de lacs, ceci aussi dans le cadre d'activités internationales en rapport avec le climat et ses changements. Les changements climatiques sont souvent observés par des indicateurs. La glace des lacs est une de ces indicateurs (partiels). Pour le monitoring des glaces lacustres on se sert souvent des dates télédéteectées satellitaires (optiques) d'approximativement une image par jour. Les facteurs limitatifs de l'analyse automatique de ces dates sont la résolution géométrique (normalement 1 pixel au sol correspondant à 250 m plus) et les nuages. Comme alternative cet article veut démontrer le potentiel du monitoring eau-glace basé sur webcam. Le but recherché est la classification par pixel à l'aide de méthodes modernes de l'apprentissage automatique afin de déterminer la surface verglacée des lacs dans des images de webcam individuelles.

Quest'articolo si basa su un progetto nazionale (avviato da MeteoSvizzera) sul tema del congelamento dei laghi, che rientra nelle attività internazionali correlate al clima e ai cambiamenti climatici. I cambiamenti climatici sono sovente osservati in base a indicatori, dei quali il ghiaccio lacustre ne fa parte. Per il monitoraggio del ghiaccio lacustre si ricorre spesso a dati (ottici) di telerilevamento satellitare, con un'immagine al giorno. Purtroppo le nuvole e la risoluzione (di norma un pixel a terra è di oltre 250 m) sono dei fattori limitativi per la valutazione automatica di questi dati. Come alternativa in quest'articolo si analizza il potenziale racchiuso dal monitoraggio del ghiaccio lacustre tramite webcam. Lo scopo consiste nell'effettuare una classificazione per pixel attraverso i metodi più all'avanguardia di machine learning per misurare la superficie ghiacciata di un lago nelle singole immagini delle webcam.

M. Rothermel, M. Xiao, M. Tom,
E. Baltasvias, K. Schindler

1. Einführung

Die kontinuierliche Beobachtung von Klimaindikatoren ist von immenser Bedeutung, um Trends und die Dynamik komplexer Klimasysteme zu verstehen.

Solche Informationen sind essentiell, um bedenkliche Entwicklungen zu erkennen und gegebenenfalls geeignete Gegenmassnahmen ergreifen zu können. In einer Initiative der Vereinigten Nationen (UN) in 1992, unterstützt durch mehrere internationale Organisationen (besonders die World Meteorological Organization), wurde das *Global Climate Observing System* (GCOS) gegründet (in der Schweiz

vertreten durch GCOS CH, bei der MeteoSchweiz). Im Rahmen dessen wurden die wichtigsten Klima-Indikatoren der Atmosphäre, Ozeane und der Landmassen identifiziert, genannt *Essential Climate Variables* (ECVs). Dabei wurden Seen und deren Vereisung als ein ECV aufgenommen. Um Zuverlässigkeit, Genauigkeit, Kostenaufwand und die allgemeine Eignung von verschiedenen Sensoren und Auswertungsmethoden für das Monitoring von See-Eis zu evaluieren, wurde vom MeteoSchweiz (Schweizer Bundesamt für Meteorologie und Klimatologie) das Projekt «Integrated Monitoring of Ice in Selected Swiss Lakes» initiiert (für eine Übersicht siehe www.prs.igp.ethz.ch/research/current_projects/integrated-monitoring-of-ice-swiss-lakes.html). Dabei umfassen die evaluierten Methoden optische Satellitendaten, Webcam-Bilder und in-situ Messungen (hauptsächlich von Temperatur im Wasser). Der Einsatz von Webcam-Bildern wurde als eine alternative und «untypische» Datenquelle untersucht. Solche Bilder werden normalerweise in kurzen Zeitabständen (10 min – 1 h) aufgenommen und weisen eine vergleichsweise (zu Satelliten) gute geometrische Auflösung auf und sind viel weniger von Wolken beeinflusst. Sie haben aber auch gewisse Probleme, auch bei der Interpretation (automatischen oder manuellen), wie im Kapitel 2 näher erläutert wird. Webcam-Bilder sind sehr oft öffentlich auf Internet-Plattformen gratis zugänglich. In diesem Beitrag soll das Potenzial des Eis-Monitorings basierend auf Webcam-Bildern aufgezeigt werden. Die Prädiktion der prozentual vereisten Seeoberfläche mittels moderner Techniken des maschinellen Lernens steht in dieser Arbeit im Vordergrund.

2. Zugrundeliegende Daten

Neben speziell für die Phänologie zusammengestellten Webcam Kollektionen existieren zahlreiche Internet-Plattformen, auf welchen Sammlungen von Webcam-Streams bestimmter Regionen gesammelt und zum Download bereitgestellt werden. So konnten für den Winter



Abb. 1: Beispielbild Kamera 0.



Abb. 2: Beispielbild Kamera 1.

2016/2017 Sequenzen von insgesamt 27 Kameras, welche Bilder von sechs ausgewählten Schweizer Seen (Silsersee, Silvaplana-see, St. Moritzsee, Sihlsee, Greifen-see und Aegerisee) automatisiert gesammelt und gespeichert werden.

Für erste Tests wurden Kameras im Bereich des St. Moritz Sees herangezogen (Abb. 1 und Abb. 2). Um implementierte Algorithmen zu evaluieren, wurden zwei Kameras mit unterschiedlicher Bodenauf- lösung ausgewählt. Kamera 0 liefert Bildsequenzen mit geringerer Bodenauf- lösung, Bilder der Kamera 1 weisen eine höhere Auflösung auf. Die Herausforde- rungen für die Prädiktion der Eisbede- ckung stellt vor allem die Ähnlichkeit des Aussehens von Eis, Wasser und Schnee dar, was selbst die manuelle Interpretati- on ohne temporalen Kontext erschwert (Abb. 3). Dies ist zum Teil auch auf die Problematik limitierter Bildqualität zu- rückzuführen, welche durch Kompressi- onsartefakte noch verstärkt wird. Des Weiteren stellen verschiedenste Beleuch- tungsverhältnisse und grosse Varianzen der Bodenauf- lösung innerhalb der einzel- nen Bilder zusätzliche Herausforderungen für die semantische Klassifizierung dar.

3. Methodik – Fully Convolutional DenseNets

Ziel der implementierten Prozesskette ist eine semantische Segmentierung von Webcam-Bildern, welche eine pixelweise Prädiktion der Klassen *Schnee*, *Eis* und *Wasser* liefern soll. Des Weiteren wurde eine Klasse *Clutter* eingeführt, um den im Rahmen von Sportveranstaltungen auf- gebauten Objekten (hauptsächlich Zelte und Absperrungen) Rechnung zu tragen. Die semantische Segmentierung wurde mit einem Fully Convolutional Network [1] realisiert. Die verwendete Netzwerkarchi-

tektur [2] stellt eine klassische Encod- er-Decoder Architektur dar (Abb. 4) und nutzt DenseBlocks (DB) [3] als grund- legendes Modul. Im Encoder wird ein Eingangsbild mit einer gewissen Anzahl von Filterkernen gefaltet (DB) und die Auflösung reduziert (TD). Letzteres ver- grössert die Umgebung («Sichtfeld»), welche in den folgenden Faltungen ein- bezogen wird. Der Decoder ist verant- wortlich für das erneute Upsampling (TU), um letztendlich semantische Karten der ursprünglichen Auflösung zu generieren. Das Vermindern der Auflösung im En- coder hat zur Folge, dass Detailinformati-

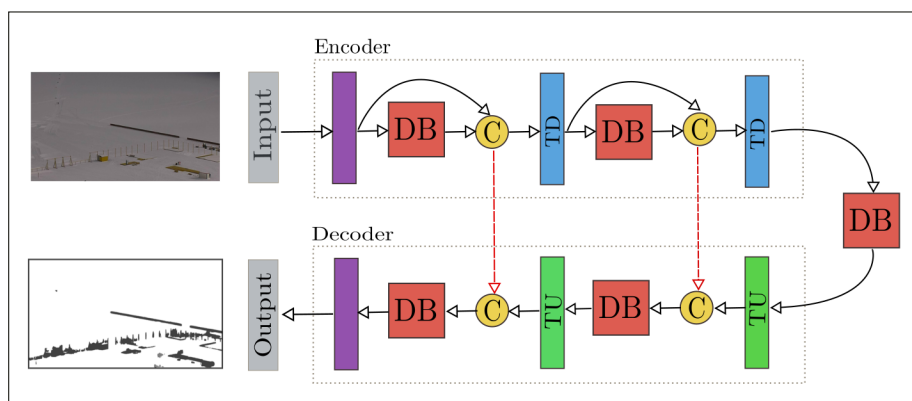


Abb. 4: Encoder-Decoder Struktur des verwendeten Netzwerks.

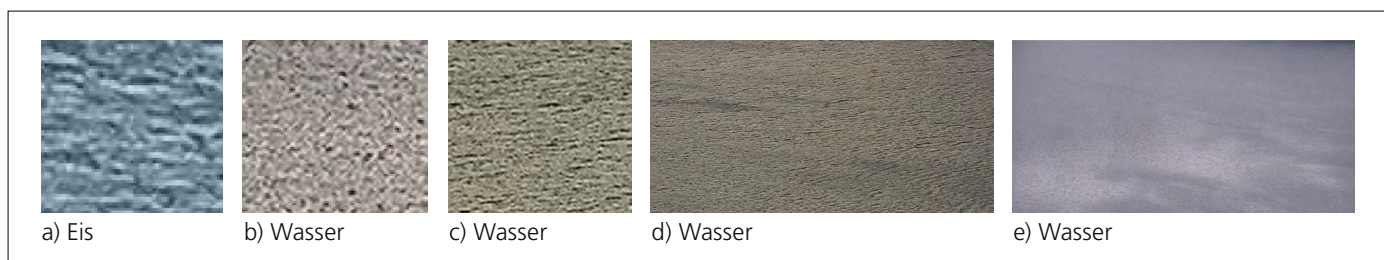


Abb. 3: Beispiele von schwer zu interpretierender Textur.

on verloren geht. Dieses Problem wird gelindert, indem Informationen höherer Auflösung vom Encoder zum Decoder übergeben werden. Dies ist durch so genannte Skip-Connections (rote Pfeile in Abb. 4) realisiert. Im letzten Schritt des Decoders ist eine Softmax-Schicht implementiert, welche Wahrscheinlichkeiten der einzelnen Klassen für jeden Pixel generiert. Eine detailliertere Beschreibung der verwendeten Architektur ist in [4] zu finden.

4. Experimente

Für die Evaluation wurden in einem ersten Schritt die Bilder der zwei Kameras manuell annotiert. Das Resultat sind semantische Karten mit pixelweiser Annotation der Klassen *Wasser*, *Eis*, *Schnee* und *Clutter* und dienen als Groundtruth. Für jede der zwei Kameras wurde ein Netzwerk trainiert. Bilder der einzelnen Kameras wurden dafür zufällig in zwei Teile (75 % und 25 %) aufgeteilt. Dabei wurden 80 % des grösseren Teils (60 % des gesamten Datensatzes) für das Training und 20 % des grösseren Teils (15 % des gesamten Datensatzes) für die Validierung verwendet. Der kleinere Teil (25 % des gesamten Datensatzes) dient als Testdatensatz.

4.1 Resultate für die semantische Segmentierung

Tabellen 1 und 2 zeigen die Konfusionsmatrizen der erhaltenen Segmentierungen. Für Kamera 0 wurde dabei für die drei wichtigsten Klassen *Wasser*, *Eis* und *Schnee* ein Recall (s. Tabelle 1) im Bereich von 88.3 % bis 98.0 % erreicht. Für Precision wurden Werte im Bereich 90.3 % bis 96.9 % erzielt. Analog dazu wurden für die hochauflösende Kamera 1 Recall und Precision im Bereich 86.4 %–97.9 % bzw. 83.5 %–98.5 % für die drei Hauptklassen erreicht. Die «Overall Accuracy» ist für beide Kameras mit 95.3 % bzw. 95.7 % vergleichbar. Die Klasse *Clutter* weist für beide Kameras niedrigere Recall- und Precision-Werte auf. Dies ist auf die ungenaue manuelle Annotation der kleinen Objektstrukturen

| Klassen | Wasser | Eis | Schnee | Clutter | Recall |
|------------------|--------------|-------------|-------------|-------------|--------|
| Wasser | 28.57 | 0.29 | 0.31 | 0.00 | 98.0% |
| Eis | 0.40 | 4.56 | 0.20 | 0.00 | 88.3% |
| Schnee | 0.51 | 0.20 | 8.29 | 0.07 | 91.4% |
| Clutter | 0.00 | 0.00 | 0.08 | 0.29 | 76.7% |
| Precision | 96.9% | 90.3% | 93.3% | 80.8% | 95.3% |

Tab. 1: Konfusionsmatrix [in Mio. Pixel], Recall und Precision (in %) für Kamera 0.

| Klassen | Wasser | Eis | Schnee | Clutter | Recall |
|------------------|---------------|--------------|--------------|-------------|--------|
| Wasser | 211.15 | 1.83 | 2.69 | 0.00 | 97.9% |
| Eis | 2.10 | 27.58 | 2.21 | 0.02 | 86.4% |
| Schnee | 1.02 | 3.60 | 94.01 | 0.54 | 94.8% |
| Clutter | 0.00 | 0.02 | 1.22 | 2.82 | 69.4% |
| Precision | 98.5% | 83.5% | 93.9% | 83.4% | 95.7% |

Tab. 2: Konfusionsmatrix [in Mio. Pixel], Recall und Precision (in %) für Kamera 1.

zurückzuführen. Abbildung 5 zeigt Beispiele der semantischen Segmentierung von Bildern der hochauflösenden Kamera 1. Die oberen Zeilen zeigen Beispiele erfolgreicher semantischer Segmentierungen, die letzte Zeile zeigt ein Beispiel für teilweise Fehlklassifikationen. Erwartungsgemäss treten Fehlklassifikationen vermehrt im Bereich mehrdeutiger Textur und in Bereichen, für welche nur eine limitierte Anzahl an Trainingsdaten zur Verfügung steht (im Prozess der Vereisung), auf.

4.2 Resultate für die Prädiktion der prozentual vereisten Oberfläche

Basierend auf den erzeugten semantischen Karten wurde die prozentual vereiste Wasseroberfläche geschätzt. Dafür wurde für jedes Bild die Pixelanzahl der Klassen *Schnee*, *Eis* und *Clutter* (die Klasse *Clutter* repräsentiert vorwiegend Strukturen wie Zäune und Zelte welche im Rahmen von Sportevents aufgebaut werden) addiert und durch die Anzahl aller Seepixel geteilt. Der tägliche Median der so berechneten Werte wurde dann

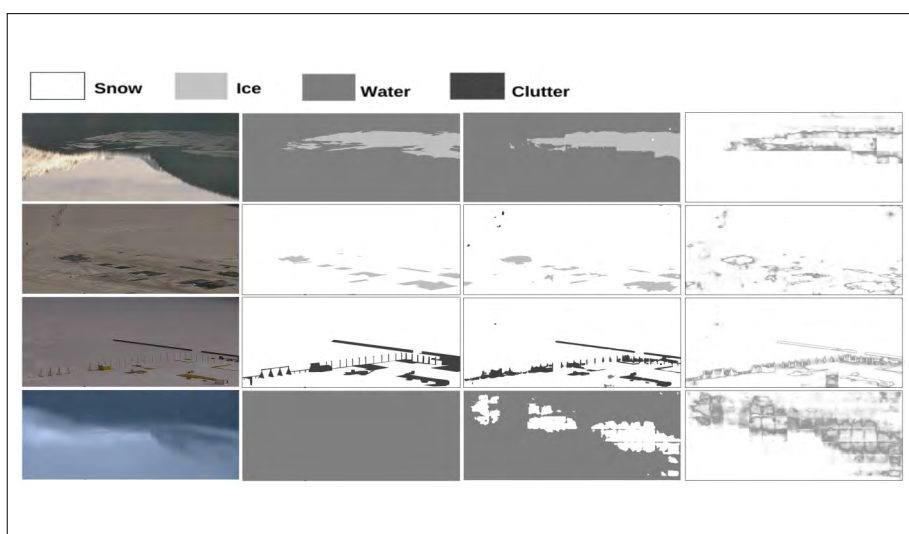


Abb. 5: Beispiele für semantische Segmentierungen für Bilder der hochauflösenden Kamera 1. Links: Bildausschnitte; Mitte: manueller Groundtruth; rechts: Ergebnis der automatischen Klassifikation.

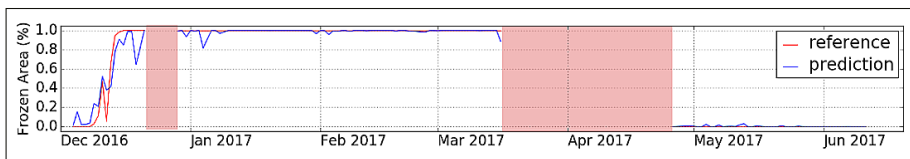


Abb. 6: Prozentual vereiste Oberflächen, Prädiktion und Groundtruth (reference) für niederauflösende Kamera 0, rosa Bereiche markieren Datenlücken.

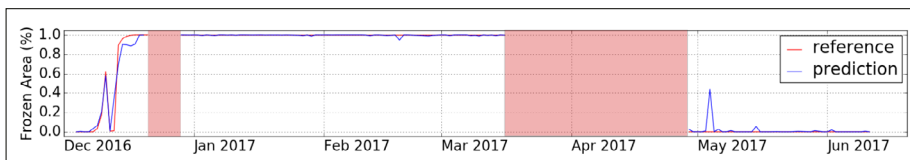


Abb. 7: Prozentual vereiste Oberfläche, Prädiktion und Groundtruth (reference) für hoch auflösende Kamera 1, rosa Bereiche markieren Datenlücken.

als finale Prädiktion der Eisbedeckung herangezogen. Ergebnisse in Abb. 5 und Abb. 6 zeigen eine gute Übereinstimmung mit der Groundtruth. Fehlklassifikationen treten in der Vereisungsphase auf, wobei bei der Auftau-Phase, die auch schwierig ist, keine Daten existierten.

5. Fazit

Webcams stellen eine interessante Alternative für das Monitoring von See-Vereisung dar. So wurde in den durchgeführten Tests eine «Overall Accuracy» für die semantische Segmentierung von Bildern zweier Testkameras von über 90% erreicht. Die daraus abgeleitete Eis-an-Eis-Klassifizierung lieferte weniger als zwei fehlklassifizierte Tage für die niedrig aufgelöste Kamera 0 und 100% richtig

klassifizierte Tage für die höher aufgelöste Kamera 1. Anzumerken ist, dass sich diese Prädiktionen nur auf den beobachteten See-Ausschnitt beziehen. Eine Aussage bezüglich der kompletten Seefläche setzt die Fusion der Ergebnisse mehrerer Kameras und damit eine exakte Georeferenzierung der Kameras voraus, plus einen Einsatz von zusätzlichen Webcams, um die ganze Seefläche abzudecken. Dies und die Ausnutzung temporaler Information von aufeinanderfolgenden Bildern sind Ziele unserer zukünftigen Arbeit. Zusammenfassend kann gesagt werden, dass ein grosses Potenzial der automatischen Interpretation von Webcam-Bildern mittels modernen Ansätzen des maschinellen Lernens, im Speziellen für ökologische Anwendungen, sehr wohl möglich und sinnvoll ist.

Referenzen:

- [1] Jegou, S., Drozdal, M., Vazquez, D., Romero, A. and Bengio, Y., 2016. The one hundred layers tiramisu: Fully convolutional dense-nets for semantic segmentation. <https://arxiv.org/abs/1611.09326>.
- [2] Long, J., Shelhamer, E. and Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In: Proc. CVPR 2015, pp. 3431–3440.
- [3] Huang, G., Liu, Z. and Weinberger, K. Q., 2016. Densely connected convolutional networks. <http://arxiv.org/abs/1608.06993>.
- [4] Xiao, M., Rothermel, M., Tom, M., Galliani, S., Baltsavias, E., Schindler, K., 2018. Lake Ice Monitoring with Webcams. ISPRS Annals of the Photogrammetry, pp. 311–317.

Mathias Rothermel
 Muyan Xiao
 Manu Tom
 Emmanuel Baltsavias
 Konrad Schindler
 Institut für Geodäsie und
 Photogrammetrie, ETH Zürich
 Stefano-Francini-Platz 5
 CH-8093 Zürich
mathias.rothermel@geod.baug.ethz.ch
muxiao@student.ethz.ch
manu.tom@geod.baug.ethz.ch
manos@geod.baug.ethz.ch
schindler@geod.baug.ethz.ch



HEXAGON
SAFETY & INFRASTRUCTURE

M.APP ENTERPRISE

AUS DATEN EINE ENTSCHEIDUNGSRUNDLAGE SCHAFFEN

-  Change Detection in der Amtlichen Vermessung
-  Energie - und Wasserverbrauch in der Stadt visualisieren
-  Infrastrukturschäden - und ausfälle sowie Reaktionszeiten für deren Reparatur überwachen

Unsere Fachleute geben gerne Auskunft:
info-switzerland@hexagonsi.com

