

KATEGORISIEREN VON BILDERN DER MARSOBERFLÄCHE DURCH UNSUPERVISED LEARNING BY BACKPROPAGATION

Merlin Scholz
merlin.scholz@tu-dortmund.de

Zusammenfassung

Um neuronale Netze zur Bildsegmentierung verwenden zu können, benötigt es einen Trainingsdatensatz, welcher zuvor von Hand kategorisiert werden muss. Da dies relativ zeitaufwändig ist, wird hier eine andere Methode vorgestellt, Neuronale Netze ohne vorherige Ground Truth zur Segmentierung zu verwenden. Dies geschieht dadurch, dass die jeweils generierte Segmentierung unter Anderem durch einen Superpixel-Algorithmus wie SLIC dazu lernt, und die so entstehenden Bilder als Zieldatensatz für die nächste Generation des Netzwerkes dient. Angewandt wird dieses Verfahren auf Bilder der Marsoberfläche. Das Hauptziel dieser Arbeit ist eine Methode zur Segmentierung der CTX-Streifen der MRO Context Camera.

I. MOTIVATION

Neuronale Netzwerke werden in den letzten Jahren häufig zur Bildsegmentierung genutzt. Ein weit bekanntes Anwendungsbeispiel findet sich in der Entwicklung des autonomen Fahrens, bei welcher das Bild in verschiedenen Klassen (z.B. Fahrbahn, Ampel) segmentiert wird. Dies geschieht unter der Voraussetzung einer Ground Truth, aus der das Netzwerk lernen kann, wie es die jeweiligen Klassen erkennen kann.

In einigen Anwendungsgebieten dieser Netzwerke existiert allerdings keine Ground Truth, zum Beispiel bei der Analyse der Marsoberfläche. Aufgrund der schlichten Größe des Datensatzes, und der Notwendigkeit von Experten in diesem Fachgebiet, ist es hier nicht leicht oder zeiteffizient möglich, manuell eine Ground Truth als Trainingsdatensatz zu erstellen.

Dieses Problem lässt sich dadurch lösen, dass die anfängliche Klassifizierung durch einen Segmentierungsalgorithmus optimiert wird. Dies wird in Abschnitt III genauer beschrieben.

Marsoberfläche als Trainingsdatensatz, während der hier vorgestellte Ansatz seine Trainingsdaten selbst zur Laufzeit generiert.

Eine Alternative zu grundlegenden CNNs und FCNs besteht aus Deep Neural Networks (DNN), wie AlexNet[7], welches in der Vergangenheit bereits zur Klassifizierung der Marsoberfläche genutzt wurde [14]. Es existiert zudem Vorarbeit zu unbeaufsichtigtem Lernen über Segmentierungsalgorithmen durch DNNs: In [15] werden DNNs dazu genutzt, die Vorgehensweise von Clusteringalgorithmen zu erlernen und gegebenenfalls zu verbessern.

Des Weiteren könnte das Ersetzen des SLIC Algorithmus aus [6] durch einen alternativen Clusteringalgorithmus wie k-Means Clustering oder Mean-Shift Clustering zu besseren Ergebnissen führen. K-Means Clustering wurde bereits in Verbindung mit CNNs ohne annotierte Daten genutzt[2]. Außerdem besteht auch die Möglichkeit, Segmentierungsalgorithmen wie Normalized Cuts[13] oder das Gaussian Mixture Model[3] zu nutzen.

II. VERWANDTE ARBEITEN

Zu dem hier vorgestellten Ansatz wurden bereits einige verwandte Arbeiten veröffentlicht. Die Idee und der grundlegende Ansatz für diese Arbeit basiert auf [6].

Es besteht auch verwandte Arbeit in der Analyse der Marsoberfläche durch neuronale Netze[4]. Dies geschah allerdings mit Bildern der Erdoberfläche oder bereits klassifizierten Bildern der

III. VORGEHENSWEISE

Der erste Schritt dieser Arbeit besteht aus der Programmierung eines Algorithmus ähnlich zu [6]. Dieser funktioniert, indem er wiederholt folgenden Prozess ausführt:

1. Eingabebild durchläuft ein Neuronales Netz mit Argmax Klassifizierung zur Feature Extraktion

2. Eingabebild durchläuft einen Clusteringalgorithmus wie SLIC[1]
3. In 2. bestimmte Cluster werden jeweils der Klasse aus 1. zugewiesen, welche als häufigstes in diesem Cluster auftritt
4. Der Verlust des Neuronalen Netzwerk berechnet sich als Softmax-Loss aus den Ergebnissen aus Schritt 1. und Schritt 3.
5. Durch Backpropagation wird das Neuronale Netz optimiert um den in der nächsten Generation entstehenden Verlust zu minimieren

Somit wird in einem Vorwärts-Durchlauf das Bild segmentiert, während im Rückwärts-Durchlauf die Parameter des Netzwerkes optimiert werden.

Dieser Algorithmus wird dementsprechend so angepasst, dass er vergleichsweise große Eingabedateien verarbeiten kann: Dies geschieht über das Sliding-Window-Verfahren.

Diese Grundlage wird durch einige Veränderungen erweitert; eine dieser Veränderungen ist, wie auch in [6] erwähnt, das Benutzen von größeren und spezialisierteren neuronalen Netzen, wie Fully Convolutional Networks (FCN)[12] für die Segmentierung.

Als Metrik zur Auswertung des Algorithmus wird dieser auf bereits segmentierte Datensätze wie *Common Objects In Context*[8] oder das *Cityscapes Dataset*[5] angewandt. Das Resultat der entwickelten Vorgehensweise wird anschließend mit den jeweils zugehörigen, manuell erstellten Segmentierungen aus den Datensätzen verglichen, also mit deren Ground Truths. Mögliche Metriken bestehen hierbei aus Objektbasierten Metriken wie F-Measure oder Intersection-over-Union oder Partitionsbasierten Metriken wie unter anderem den Rand Index[11] oder Variation of Information[10].

Um die so ermittelten Metriken zu optimieren, werden einzelne Teile des Algorithmus ersetzt: So kann das CNN wie bereits erwähnt durch ein FCN ersetzt werden, während der Segmentierungsalgorithmus SLIC durch k-Means Clustering oder den Mean-Shift Algorithmus ersetzt werden kann.[2] Alternativen hierzu bestehen aus Normalized Cuts[13] oder dem Gaussian Mixture Model[3].

Bei der Abänderung des neuronalen Netzwerkes gilt zu beachten, dass die Initialisierung unter einzelnen Netzwerken immer identisch ist (manuelle Seed-Vorgabe), da die Ergebnisse sonst durch die Zufallswerte verfälscht werden können.

Ultimativ dient diese Vorgehensweise dem Ziel, die Bilder der MRO Context Camera der NASA[9], auch genannt CRX-Streifen, segmentieren und somit analysieren zu können.

REFERENZEN

- [1] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. Slic superpixels. page 15, 2010.
- [2] E. Ahn, A. Kumar, D. Feng, M. Fulham, and J. Kim. Unsupervised Feature Learning with K-means and An Ensemble of Deep Convolutional Neural Networks for Medical Image Classification. *arXiv e-prints*, page arXiv:1906.03359, Jun 2019.
- [3] Z. Ban, J. Liu, and L. Cao. Superpixel segmentation using gaussian mixture model. *IEEE Transactions on Image Processing*, 27(8):4105–4117, Aug 2018.
- [4] J. P. Cohen, H. Z. Lo, T. Lu, and W. Ding. Crater Detection via Convolutional Neural Networks. *arXiv e-prints*, page arXiv:1601.00978, Jan 2016.
- [5] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele. The cityscapes dataset for semantic urban scene understanding. In *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [6] A. Kanezaki. Unsupervised image segmentation by backpropagation. In *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2018.
- [7] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger,

editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012.

pages 478–487, New York, New York, USA, 20–22 Jun 2016. PMLR.

- [8] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick. Microsoft COCO: Common objects in context. In *European Conference on Computer Vision (ECCV)*, 2014.
- [9] M. C. Malin, J. F. Bell III, B. A. Cantor, M. A. Caplinger, W. M. Calvin, R. T. Clancy, K. S. Edgett, L. Edwards, R. M. Haberle, P. B. James, S. W. Lee, M. A. Ravine, P. C. Thomas, and M. J. Wolff. Context camera investigation on board the mars reconnaissance orbiter. *Journal of Geophysical Research: Planets*, 112(E5), 2007.
- [10] M. Meila. Comparing clusterings by the variation of information. In *Learning Theory and Kernel Machines: 16th Annual Conference on Computational Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777, page 173. Springer Science & Business Media, 2003.
- [11] J. Santos and M. Embrechts. On the use of the adjusted rand index as a metric for evaluating supervised classification. volume 2009, pages 175–184, 09 2009.
- [12] E. Shelhamer, J. Long, and T. Darrell. Fully convolutional networks for semantic segmentation. *PAMI*, 2016.
- [13] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 22(8), 2000.
- [14] K. Wagstaff, Y. Lu, A. Stanboli, K. Grimes, T. Gowda, and J. Padams. Deep mars: Cnn classification of mars imagery for the pds imaging atlas. In *AAAI*, 2018.
- [15] J. Xie, R. Girshick, and A. Farhadi. Unsupervised deep embedding for clustering analysis. In M. F. Balcan and K. Q. Weinberger, editors, *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*,