

Hauptseminar

Thema: Effiziente Transformer in der Computer Vision: Werden CNNs abgelöst?

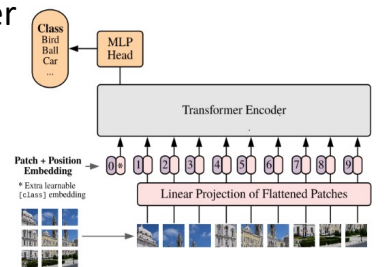
Beschreibung:

Transformer konnten Rekurrente Netzwerke in vielen Bereichen der Sprachverarbeitung bereits ablösen. Getrieben von diesem Erfolg werden Transformer zunehmend auch in der Computer Vision verwendet.

In [1] wurde ein Verfahren gezeigt, welches auf dem ImageNet Datensatz bessere Ergebnisse als Convolutional Neural Networks erzielen kann. Aufbauend auf diesem Erfolg wurden in [2-4] die Genauigkeit und der Berechnungsaufwand optimiert. Aufgrund der Komplexität der Self-Attention wurden in [5-6] weitere Anpassungen an Transformern vorgenommen, um den Speicherbedarf zu reduzieren, und den Durchsatz zu erhöhen.

Detaillierte Aufgabenstellung:

- Recherche zu effizienten Transformern für die Computer Vision ausgehend von der bereitgestellten Literatur
- Erstellung einer Übersicht des State of the Art
- Herausstellen und Aufarbeiten von mindestens einem effizienten und leistungsfähigen Verfahren
- Vorstellung des Themas im Rahmen einer Abschlusspräsentation



Geeignet für:

- Masterstudiengänge

Themengebiet / Schwerpunkte:

- Deep Learning, Machine Learning

Notwendige Voraussetzungen:

- Guter Abschluss der Vorlesung „Neuroinformatik“
- Guter Abschluss der Vorlesung „Deep Learning for Computer Vision“

Literatur:

- [1] Dosovitskiy et al.: [An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale](#), ICLR, 2021
 - [2] d'Ascoli et al.: [ConViT: Vision Transformers with Convolutional Inductive Biases](#), arXiv, 2021
 - [3] Liu et al.: [Swin Transformer: Hierarchical Vision Transformer using Shifted Windows](#), ICCV, 2021
 - [4] Liu et al.: [Swin Transformer V2: Scaling Up Capacity and Resolution](#), arXiv, 2021
 - [5] Lu, et al.: [SOFT: Softmax-free Transformer with Linear Complexity](#), NIPS 2021
 - [6] Wang, et al.: [Linformer: Self-Attention with Linear Complexity](#), arXiv 2020
- Elektronische Literaturdatenbank des FG NI&KR mit Recherchemöglichkeiten
 - Elektronische Konferenzproceedings Datenbank des FG NI&KR
 - Google Scholar scholar.google.com
 - Proceedings der rel. Konferenzen (IROS, ICRA, NIPS, ICML, ICLR, IJCNN, WCCI, ICANN, CVPR, ICCV, ECCV, BMVC, ICPR, ICIP, ...)

Betreuer:

Söhnke B. Fishedick, M.Sc. (soehnke.fishedick@tu-ilmenau.de)
Daniel Seichter, M.Sc. (daniel.seichter@tu-ilmenau.de)

Betr. Hochschullehrer:

Prof. Dr. H.M. Groß

Bearbeiter:

Henning Franke