

# Hauptseminar

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

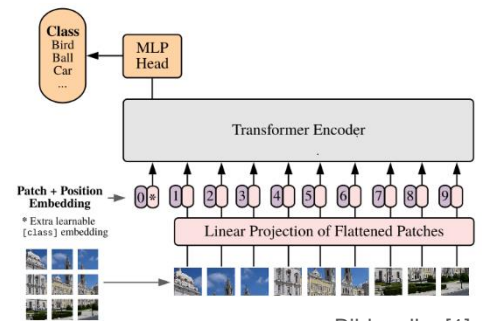
## Beschreibung:

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 (CNNs) erzielen kann. Aufbauend auf diesem Erfolg wurden in [2-5] die Genauigkeit und der Berechnungsaufwand optimiert.

## Detaillierte Aufgabenstellung:

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



Bildquelle: [1]

## Geeignet für:

- Bachelor- / Masterstudiengänge

## Themengebiet / Schwerpunkte:

- Deep Learning, Machine Learning

## Notwendige Voraussetzungen:

- Guter Abschluss der Vorlesung „Neuroinformatik“
- Aktive Teilnahme oder erfolgreicher 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] Mehta et al.: [MobileViT: Light-weight, General-purpose, and Mobile-friendly Vision Transformer](#), arXiv, 2021
  - [4] Liu et al.: [Swin Transformer: Hierarchical Vision Transformer using Shifted Windows](#), ICCV, 2021
  - [5] Liu et al.: [Swin Transformer V2: Scaling Up Capacity and Resolution](#), arXiv, 2021
- Elektronische Literaturdatenbank des FG NI&KR mit Recherchemöglichkeiten
  - Elektronische Konferenzproceedings Datenbank des FG NI&KR
  - Google Scholar [scholar.google.com](https://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-benedikt.fishedick@tu-ilmenau.de](mailto:soehnke-benedikt.fishedick@tu-ilmenau.de))

Daniel Seichter, M.Sc. ([daniel.seichter@tu-ilmenau.de](mailto:daniel.seichter@tu-ilmenau.de))

**Betr. Hochschullehrer:** Prof. Dr. H.M. Groß

**Bearbeiter:** offen