Hauptseminar

Thema: Woher weiß ein Neuronales Netzwerk was es nicht weiß?
Out-of-Distribution Detection und Out-of-Distribution Generalization

Tiefe Neuronale Netzwerke sind dafür bekannt, gut auf unbekannte Daten zu generalisieren. Grundvoraussetzung ist jedoch, dass die unbekannten Daten der gleichen Verteilung entsprechen wie die Trainingsdaten. Ist dies nicht der Fall, dann werden sie als Out-of-Distribution-(OoD)-Daten bezeichnet. Ohne spezielle Maßnahmen ist ein Neuronales Netzwerk nicht in der Lage solche Daten zu erkennen und gibt für OoD-Daten ein Klassifikationsergebnis mit hoher Konfidenz aus. Dies kann jedoch in sicherheitskritischen Anwendungen, wie der Mensch-Roboter-Kollaboration zu gefährlichen Fehlentscheidungen führen. Daher beschäftigen sich aktuelle Forschungsarbeiten mit der Detektion von OoD-Daten (OoD Detection) und mit der besseren Generalisierung wenn OoD-Daten vorliegen (OoD Generalization). Das Ziel dieses Hauptseminars ist die Aufbereitung aktueller Ansätze zur Lösung dieser beiden Problemstellungen.

Aufgabenstellung:

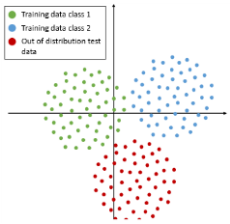
- Aufbereitung des State of the Art zu OoD Detection und OoD Generalization, z.B. anhand der Surveys [6-8]
- Aufarbeiten aktueller Ansätze [1-5] und detaillierte Vorstellung ausgewählter Verfahren
- Abschätzung, wo sich diese Verfahren für die Robotik einsetzen lassen
- Vortrag im Rahmen des Hauptseminars

Geeignet für:

Bachelor- / Masterstudiengänge

Themengebiet / Schwerpunkte:

Deep Learning



Schematische Darstellung von OoD-Daten bei der Klassifikation. Die OoD-Daten (rot) liegen außerhalb des Bereichs der Daten, die das Neuronale Netzwerk während des Trainings zu sehen bekommen hat (grün, blau). Bildquelle: [6]

Erforderliche Vorkenntnisse:

Guter Abschluss der Vorlesung „Neuroinformatik und Maschinelles Lernen“ und Erfahrungen im Bereich Deep Learning
oder erfolgreicher Abschluss der Vorlesung „Deep Learning for Computer Vision“

Zu verwendende Literatur:

[1] **Thulasidasan** et al.: An Effective Baseline for Robustness to Distributional Shift. ICMLA, **2021**.
[2] **Du** et al.: VOS: Learning What You Don't Know by Virtual Outlier Synthesis. ICLR, **2022**.
[3] **Hebbalaguppe** et al.: A Novel Data Augmentation Technique for Out-of-Distribution Sample Detection using Compounded Corruptions. arXiv, **2022**.
[4] **Andreassen** et al.: The Evolution of Out-of-Distribution Robustness Throughout Fine-Tuning. arXiv, **2021**.
[5] **Rame** et al.: Diverse Weight Averaging for Out-of-Distribution Generalization. arXiv, **2022**.
[6] **Gawlikowski** et al.: A Survey of Uncertainty in Deep Neural Networks. arXiv, **2021**.
[7] **Salehi** et al.: A Unified Survey on Anomaly, Novelty, Open-Set, and Out-of-Distribution Detection: Solutions and Future Challenges. preprint TMLR, **2021**.
[8] **Yang** et al.: Generalized Out-of-Distribution Detection: A Survey. preprint TPAMI, **2021**.

- Elektronische Literaturdatenbank des FG NI&KR mit Recherchemöglichkeiten
- Elektronische Konferenzproceedings-Datenbank des FG NI&KR
- IEEE Recherchesystem www.ieeexplore.ieee.org (nur aus dem Uni-Netz bzw. via VPN)
- Google Scholar scholar.google.com
- Suche nach ähnlichen Publikationen connectedpapers.com, arxiv-sanity-lite.com
- Proceedings der relevanten Konferenzen (NeurIPS, ICML, ICLR, IJCNN, WCCI, ICANN, CVPR, ICCV, ECCV, BMVC, AVSS, ICPR, ICIP, …)

Betreuer: Dr. Markus Eisenbach (Markus.Eisenbach@tu-ilmenau.de)
Betr. Hochschullehrer: Prof. Dr. H.-M. Groß
Bearbeiter: **Lena Kellermann**