

# B-UP9 - RoNNi – Verbesserung der Robustheit und Energieeffizienz von Neuronalen Netzwerken durch Äquivalenzprüfung

**Partner:** Marko Kleine Büning, Philipp Kern  
Institut für Theoretische Informatik, KIT

**Projektlaufzeit:** 15.06.2021 – 14.03.2022



## Abstract

Maschinelles Lernen wird heute sowohl in der Produktion als auch in der Steuerung von Fahrzeugen verwendet. Neuronale Netze spielen im Bereich des autonomen Fahrens eine zentrale Rolle. Ihr Einsatz wird dabei durch einen hohen Speicher- und Energiebedarf, sowie fehlende Sicherheitsgarantien erschwert. Äquivalenzprüfung erlaubt es, (unerwünschte) Abweichungen bei Updates von Neuronalen Netzen sichtbar zu machen und gezielt Trainingsdaten zur Minimierung der Abweichung zwischen zwei Netzen zu generieren. Mit formalen Methoden kann man insbesondere verifizieren, ob verkleinerte und damit energieeffizientere Netze immer noch den ursprünglichen Qualitätsanforderungen genügen.

## Zielsetzung und Lösungsansatz

- Die Definition erster abgeschwächter Äquivalenzbegriffe für Neuronale Netzwerke ist bereits erfolgt.
- Beschreibung des relevanten Eingaberaumes von Neuronalen Netzen, wie zum Beispiel gültige Parameter einer Motorsteuerung
- Entwicklung eines Verfahrens zur Äquivalenzprüfung basierend auf Constraint Programming und Abstrakter Interpretation
- Implementierung eines Software-Prototypen für das entwickelte Verfahren
- Optimierung des Verfahrens auf relevanten Neuronalen Netzen

