

## SafeSign

Das sichere Verkehrszeichen 4.1

|                                 |                                                                 |                        |               |
|---------------------------------|-----------------------------------------------------------------|------------------------|---------------|
| <b>Programm / Ausschreibung</b> | Ideen Lab 4.0, Ideen Lab 4.0, Ideen Lab4.0 - Ausschreibung 2019 | <b>Status</b>          | abgeschlossen |
| <b>Projektstart</b>             | 01.03.2020                                                      | <b>Projektende</b>     | 31.08.2021    |
| <b>Zeitraum</b>                 | 2020 - 2021                                                     | <b>Projektlaufzeit</b> | 18 Monate     |
| <b>Keywords</b>                 | Deep Learning; Explainable AI; Ethics in autonomous driving     |                        |               |

### Projektbeschreibung

Mit der zunehmenden Nutzung von automatischen Systemen im Straßenverkehr (Fahrassistenzsysteme, autonome Fahrzeuge) sollen die Verkehrszeichen auch von diesen Systemen sicher "gesehen" werden. Insbesondere in der Übergangsphase (gemischte Nutzung der Straße durch Mensch und Maschine) stellt die potentiell unterschiedliche Wahrnehmung der Verkehrszeichen ein großes Gefahrenpotential dar.

Wir wollen in diesem Projekt untersuchen, inwieweit Störungen bei aktuellen Deep Learning basierten Kennzeichenklassifikationssystemen zu Fehlklassifikationen beitragen. Vom Menschen kann oftmals nicht beurteilt werden wie gut ein Verkehrszeichen für eine künstliche Intelligenz lesbar ist.

Zum Beispiel: ein Mensch fährt mit seinem Fahrzeug auf einen Tunnel zu. Die Anzeige (Wechselverkehrszeichen mit LED Elementen) zeigt aufgrund der Höhenkontrolle "STOP" an. Der Mensch hält. Aufgrund einer Störung (Witterung, Ausfall einzelner LEDs) erkennt ein autonomes Fahrzeug das "STOP" nicht oder interpretiert es falsch. Ein folgenschwerer Unfall kann die Folge sein.

Um in Hinblick auf die in die naher Zukunft liegende gemischte Nutzung der Straße Vertrauen zu schaffen, wollen wir neue Methoden zur Steigerung von Robustheit sowie zur Interpretierbarkeit von KI-Modellen kombinieren und entwickeln. Auf Basis von realen Verkehrszeichenbildern (mit/ohne Störungen) sowie synthetisch erzeugten Störungsbildern soll das Gesamtsystem evaluiert werden. Die technischen Methoden sollen auch im Hinblick auf ethische Grundsätze hinterfragt und evaluiert werden.

Die Projektergebnisse (prototypische Deep Learning Methoden) und insbesondere die Datenbank mit Störungsbildern wird öffentlich verfügbar gemacht, um für österreichische Unternehmen im Bereich Mobilität, Straßeninfrastruktur und autonomes Fahren eine Basis für weitere Entwicklungen zu legen. Insbesondere kann dadurch die Entwicklung zukünftiger Verkehrszeichen angestoßen werden, die für Mensch und Maschine gleichermaßen lesbar sind.

Eine nach ethischen Leitlinien geführte Forschung zu ethisch relevanten Fragestellungen ermöglicht, Vertrauen in künstliche

Intelligenz zu schaffen, womit vertrauensvolle künstliche Intelligenz gesellschaftlich akzeptiert wird. Das ist insbesondere in sicherheitskritischen Bereichen wie Verkehr besonders relevant.

## **Abstract**

The increasing use of automated systems on the road (advanced driver assistance systems, autonomous vehicles) requires that traffic signs can be perceived with high quality by these systems. Especially during the transition from human to fully automated systems (mixed usage of road infrastructure) a different interpretation of signs exposes high risks for road safety.

In this project we will investigate the influence of disturbance on deep learning based traffic sign classification systems. Humans typically can not judge if a sign can be perceived correctly by an artificial intelligence. For example: a human driver approaches a tunnel. The variable traffic sign (based on LEDs) clearly shows "STOP" due to a height incident in the tunnel. The human driver stops the car. Due to disturbances (e.g. weather condition such as snow, ice, fog; defect of some LEDs) the automated system of the following car does not (or not correctly) perceive the sign. A crash happens.

In order to provide trust in future mixed usage situations, we will combine and develop methods to improve deep learning robustness and interpretability. We will evaluate the methods on real images of traffic signs (with / without disturbances) as well as synthetically generated disturbance images. The technical methods and concepts will be constantly reviewed with respect to ethical principles.

The main project results (e.g. prototypical deep learning models, a database of disturbed traffic sign images) will be made available for the public. Austrian companies in the field of mobility, road infrastructure and autonomous driving can start further research and development project based on this results. In particular we expect guidance for the design of future (human and machine readable) traffic signs.

Research that is based on ethical principles and guidelines will increase trust in artificial intelligence systems and boost social acceptance of such systems. In particular for safety relevant areas, such as traffic, this is of utmost importance.

## **Projektkoordinator**

- RISC Software GmbH

## **Projektpartner**

- Autobahnen- und Schnellstraßen- Finanzierungs-Aktiengesellschaft
- Universität Linz