

## SAiEX

Safe Artificial Intelligence with integrated explainable integrity level

|                                 |                                       |                        |               |
|---------------------------------|---------------------------------------|------------------------|---------------|
| <b>Programm / Ausschreibung</b> | BASIS, Basisprogramm, Budgetjahr 2021 | <b>Status</b>          | abgeschlossen |
| <b>Projektstart</b>             | 01.12.2020                            | <b>Projektende</b>     | 30.11.2021    |
| <b>Zeitraum</b>                 | 2020 - 2021                           | <b>Projektlaufzeit</b> | 12 Monate     |
| <b>Keywords</b>                 |                                       |                        |               |

### Projektbeschreibung

Die Fehlaussagen von Maschinellen Lernalgorithmen haben bereits zu katastrophalen Verkehrsunfällen und Toten sowie Schwerverletzten geführt. Die neuesten Erkenntnisse aus dem Bereich von Deep-Learning zeigen, dass man sich nicht einfach auf die Aussage von trainierten Netzen verlassen kann. Noch schlimmer kann bei geringsten Änderungen der Daten das Ergebnis z.B. einer Klassifizierung einer Ampel (rot, orange, grün) falsch sein.

Artificial Intelligence (im Text auch AI, Deutsch: Künstliche Intelligenz – im Text auch KI) ist auf jeden Fall die Zukunft unter anderem für "Autonomes Fahren" und Automatisierungsaufgaben in der Robotik. Artificial Intelligence vereinfacht uns das Leben und gestaltet es sicherer. AVI Systems ist spezialisiert auf die Entwicklung von intelligenten Lösungen für den Öffentlichen Verkehr, die Automotive Branche, die erzeugende Industrie und den Straßenverkehr.

SAiEX hat es sich zum Ziel gesetzt die vorhandenen DNN Algorithmen von AVI, die in den Intelligenten Rückspiegeln eingesetzt werden "sicher" im Sinne der Ergebnisse zu machen.

Konkret werden die folgenden Ziele im Rahmen von SAiEX verfolgt:

- 1) die vorhandenen DNN für den intelligenten Rückspiegel von AVI werden im Zuge von SAiEX weiter optimiert um sie auf die AVI Hardwareplattform Real-Time-Interface synthetisieren zu können
- 2) Verfahren erforschen anhand derer eine Online Aussage über die aktuelle Detektions- und Klassifizierungsergebnisse, um weitere Steuerungseinheiten z.B. für Autonomes Fahren mit dieser wesentlichen Qualitätsinformation zu versorgen
- 3) Erforschen in wie weit die Objektlisten und Distanzwerte die von dem intelligenten Rückspiegel geliefert werden können, sich für Safety Anwendungen bei den zukünftigen ADAS Systemen, welche von der EU gefordert werden als ISO26262 und SOTIF tauglich eingesetzt werden können.

Mit den Forschungsergebnissen von SAiEX können somit neue Maßstäbe in der Aussage über die Zuverlässigkeit von DNN möglich sein. Wichtig ist dabei, dass als konkrete Anwendung der Intelligente Rückspiegel von AVI ausgewählt wurde, prinzipiell jedoch auch noch bewertet werden soll, ob LiDAR und RADAR Daten ebenfalls damit "sicherer" gemacht werden

können.

Im ersten Forschungsjahr wurden wesentliche Erkenntnisse rund um die offline XAI Erfassung und auch die Bedrohungen und Chancen mit adversarial Techniken in Bezug auf synthetische Trainingsdaten als auch als Grundlage für die Explainability geschaffen. Im zweiten Forschungsjahr wird der Schwerpunkt die Online Analyse in einer realen Anwendung wie einem ADAS sein.

## **Projektpartner**

- EYYES GmbH