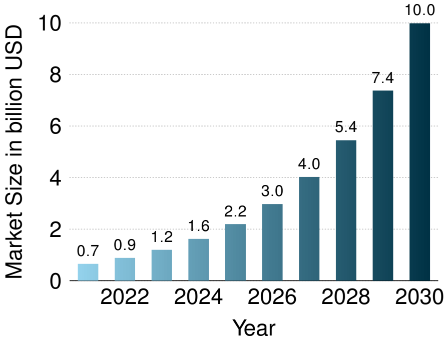
TOWARDS CERTIFIABLE AI IN AVIATION

A Framework for Neural Network Assurance Using Advanced Visualization and Safety Nets

DLR

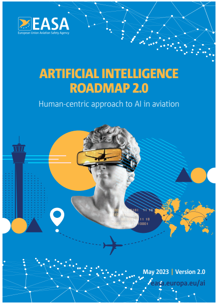
Motivation

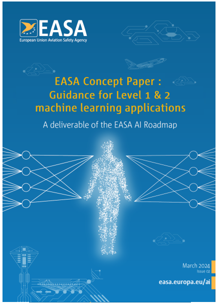
- AI already vital in many domains
- AI in aviation to reach $10 billion by 2030, CAGR >35 %
- Safety in other domains often tread lightly
- Manual inspection still essential, requires tooling



- AI will severely impact future aviation
- Safety is paramount

EASA Roadmap for Safe AI in Aviation





Learning Assurance

"All [. . .] actions [. . .] that error[s] [. . .] have been identified and corrected such that the AI/ML constituent [. . .] provides sufficient generalisation and robustness capabilities."

- EASA AI Roadmap and Concept Papers
- Way towards safe Artificial Intelligence in aviation
- Emphasize a clear and transparent approach

Operational Design Domain

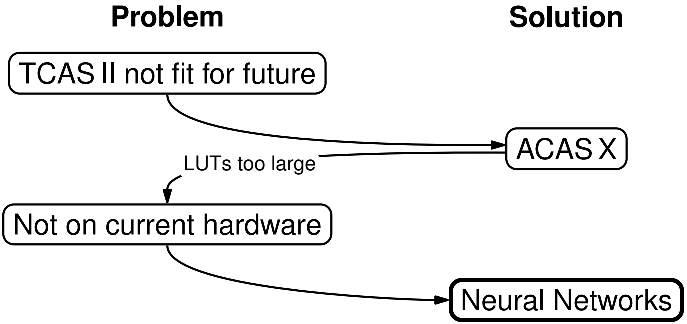
- Developed by SAE International
- Designed for autonomous systems
- Clearly defines environmental conditions
- Enforces boundaries of operation
- Required by EASA for all AI applications

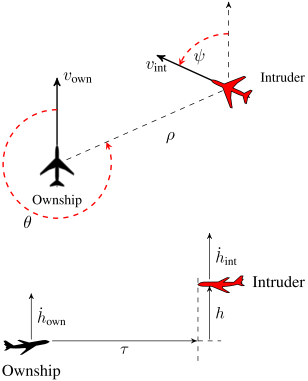
OPERATIONAL DESIGN DOMAIN
- Scenery
  - Geography = Above land
- Dynamic Elements
  - Intruder
    - ...
- Environmental Conditions
  - Wind = 0 kn to 40 kn

"Operating conditions under which a given driving automation system [. . .] is specifically designed to function, including [. . .] **environmental**, **geographical**, and time-of-day restrictions, and [. . .] **traffic** or **roadway** characteristics."

Collision Avoidance

- Collision Avoidance is crucial for safety
- TCAS II is the current standard





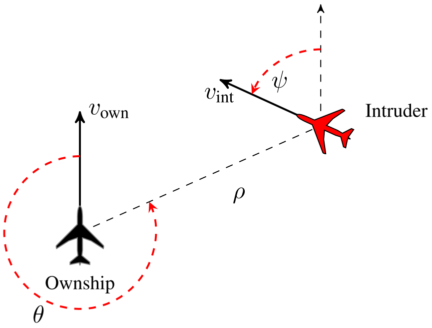
Images based on [8]

[8] Kyle D. Julian and Mykel J. Kochenderfer. “Guaranteeing Safety for Neural Network-Based Aircraft Collision Avoidance Systems”. In: *2019 IEEE/AIAA 38th Digital Avionics Systems Conference (DASC)*. IEEE, Sept. 2019

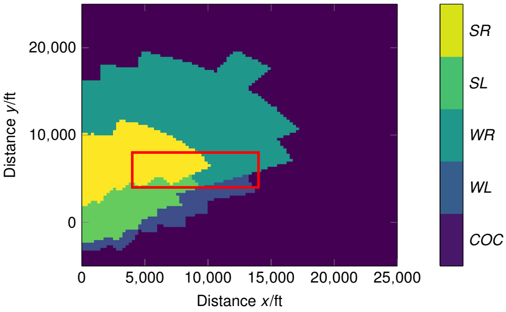
HCAS: A Horizontal Collision Avoidance System

Advisory	Description
COC	clear of conflict
WL	weak left
WR	weak right
SL	strong left
SR	strong right

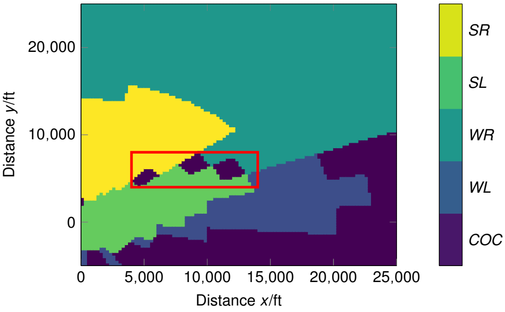
Variable	Unit	Description
ρ	ft	Distance to intruder
θ	°	Bearing angle to intruder
ψ	°	Relative heading angle
v_{own}	$\mathrm{ft\,s^{-1}}$	Ownship's true airspeed
v_{int}	$\mathrm{ft\,s^{-1}}$	Intruder's true airspeed
τ	s	Time to closest point of approach
s_{adv}	–	Previous advisory



Ground truth data[6]



Neural network[8]



HCAS with $\psi = -1^\circ$, $\tau = 5\,\text{s}$, and $s_{\text{adv}} = COC$.

[6] RTCA, Inc. *DO-386 Volume 1 & 2*. Tech. rep. Washington, DC, USA: GlobalSpec, Dec. 17, 2020

[8] Kyle D. Julian and Mykel J. Kochenderfer. "Guaranteeing Safety for Neural Network-Based Aircraft Collision Avoidance Systems". In: *2019 IEEE/AIAA 38th Digital Avionics Systems Conference (DASC)*. IEEE, Sept. 2019

SafetyNet

- Combination of Neural Networks and Lookup Tables
- Lookup Tables save correction data

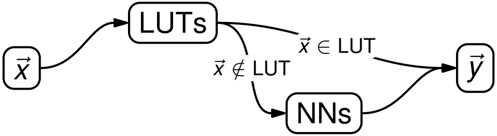
Training

1. Train Neural Network
2. Save correction data in Lookup Table

Inference

1. Query Lookup Tables
2. If input vector not in Lookup Table, infer Neural Networks
3. Return correct output

Datatype	Size	Precision
Lookup Table	large	100 %
Neural Network	small	$\leq$ 100 %
SafetyNet	small	100 %



Feasibility of Calculation

- 100 % assurance only via brute-force calculation
- Using `double`/`f64` precision:

$$n_{\text{input, HCAS}} \approx (2^{64})^4 \cdot 5 \approx 6 \cdot 10^{77}$$
$$n_{\text{input, VCAS}} \approx (2^{64})^4 \cdot 9 \approx 10^{78}$$

✗ Not feasible

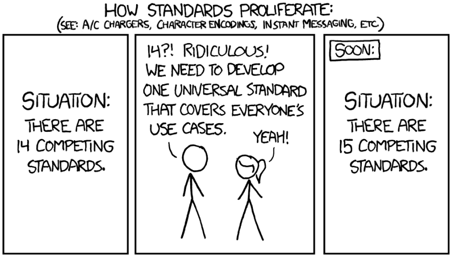
- Using discretization:

$$n_{\text{input, HCAS, disc}} \approx 1.3 \cdot 10^{12}$$
$$n_{\text{input, VCAS, disc}} \approx 8.5 \cdot 10^{11}$$

✓ Feasible in $< 24\,\text{h}$

SafetyNet Manifest

- No open-source implementation available
- Reference implementation
- Use-case independent (with limitations)
- Human-readable and machine-parsable
  $\Rightarrow$ JSON Schema
- JSON Schema allows for automatic validation
- Business logic to be implemented separately



xkcd: Standards (`https://xkcd.com/927/`)

Connection between SafetyNet and ODD

```json
{
 "version": "1.0.0",
 "datatype": "float32",
 "inputs": [
  {
   "index": 0,
   "id": "alt",
   "minimum": 0,
   "maximum": 60000,
   "unit": "ft",
   "ranges": [
    {
     "minimum": 0,
     "maximum": 1000,
     "stride": 10
    },
```

OPERATIONAL DESIGN DOMAIN

- Scenery
  - Geography = Above land
  - Airspace
    - Altitude = {0 ft to 60 000 ft}
- Dynamic Elements
  - Intruder
    - ...
- Environmental Conditions
  - Wind = 0 kn to 40 kn

Business Logic

- Not all necessary validations and implementations can be covered by the JSON schema
- Alignment between files (e.g., data types, inputs, and outputs)
- Four categories of guidelines:

  G General

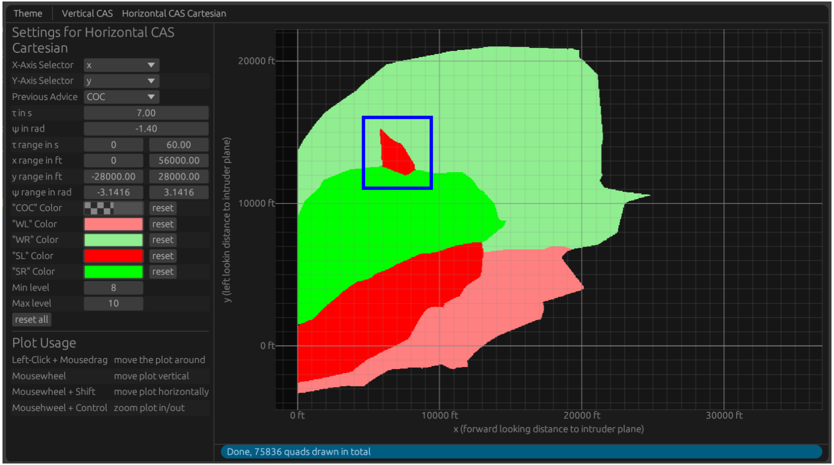
  M Manifest

  L Lookup Table

  N Neural Networks

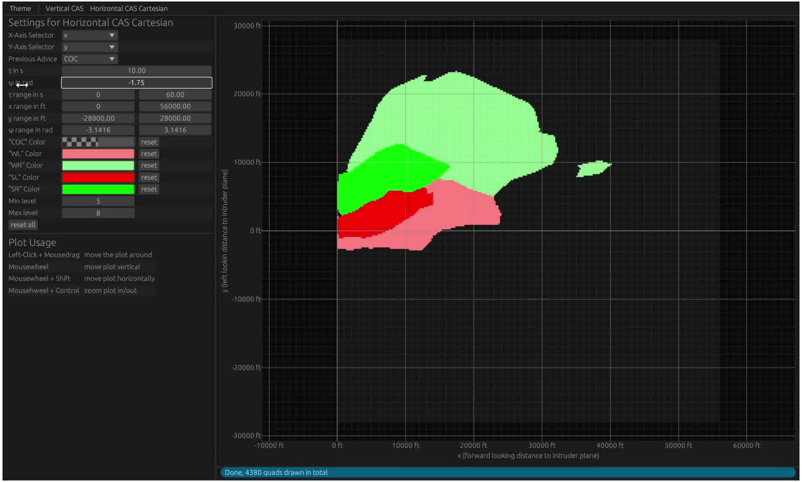
ID	Name
G-001	Available Files
G-002	Datatype Coherence
G-003	Input Coherence
G-004	Input Coverage
G-005	Output Number
G-006	Output Type
G-007	Ensured Responsibility
G-008	Single Responsibility
G-009	Condition Limits
G-010	Wildcard Conditional
M-001	Versioning
M-002	Compatible Versioning
L-001	Correct Output
L-002	Known Format
L-003	Relayed Responsibility
N-001	Correct Output
N-002	Known Format

openCAS



Theme
Vertical CAS
Horizontal CAS Cartesian
Settings for Horizontal CAS Cartesian
X-Axis Selector x
Y-Axis Selector y
Previous Advice COC
τ in s 7.00
ψ in rad -1.40
τ range in s 0 60.00
x range in ft 0 56000.00
y range in ft -28000.00 28000.00
ψ range in rad -3.1416 3.1416
"COC" Color reset
"WL" Color reset
"WR" Color reset
"SL" Color reset
"SR" Color reset
Min level 8
Max level 10
reset all
Plot Usage
Left-Click + Mousedrag move the plot around
Mousewheel move plot vertical
Mousewheel + Shift move plot horizontally
Mousewheel + Control zoom plot in/out
20000 ft
10000 ft
0 ft
y (left lookin distance to intruder plane)
0 ft
10000 ft
20000 ft
30000 ft
x (forward looking distance to intruder plane)
Done, 75836 quads drawn in total

openCAS Demo



Summary and Outlook

- SafetyNets used to detect and prevent incorrect AI behavior
- Common SafetyNet format ensures interoperability
- SafetyNet JSON Schema and business logic for verification and validation
- openCAS reduces input dimensionality, helps understanding AI behavior
- Investigate SafetyNets for other domains
- Provide a 100 % correct SafetyNet for ACAS X_U

- Manual inspection is still essential but requires proper tooling
- SafetyNets can detect and prevent incorrect AI behavior

Contact

Johann Christensen
German Aerospace Center (DLR)
Institute for AI Safety & Security
Rathausallee 12 | 53757 Sankt Augustin | Germany
johann.christensen@dlr.de

References

[1] Precedence Research. *Artificial Intelligence in Aviation Market Size, Share, and Trends 2024 to 2034*. Research rep. 1748. Precedence Research, May 2022

[2] European Union Aviation Safety Agency (EASA). *Artificial Intelligence Roadmap 2.0*. Tech. rep. Version 2.0. Postfach 10 12 53, 50452 Cologne, Germany: European Union Aviation Safety Agency (EASA), May 2023

[3] European Union Aviation Safety Agency (EASA). *EASA Concept Paper: Guidance for Level 1 & 2 Machine Learning Applications*. Tech. rep. Version Issue 02. Postfach 10 12 53, 50452 Cologne, Germany: European Union Aviation Safety Agency (EASA), Apr. 19, 2024

[4] RTCA, Inc. *DO-185B*. Tech. rep. Washington, DC, USA: GlobalSpec, June 19, 2008

[5] RTCA, Inc. *DO-385*. Tech. rep. Washington, DC, USA: GlobalSpec, Oct. 2, 2018

[6] RTCA, Inc. *DO-386 Volume 1 & 2*. Tech. rep. Washington, DC, USA: GlobalSpec, Dec. 17, 2020

[7] Kyle D. Julian et al. “Policy compression for aircraft collision avoidance systems”. In: *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*. IEEE, Sept. 2016

[8] Kyle D. Julian and Mykel J. Kochenderfer. “Guaranteeing Safety for Neural Network-Based Aircraft Collision Avoidance Systems”. In: *2019 IEEE/AIAA 38th Digital Avionics Systems Conference (DASC)*. IEEE, Sept. 2019

Imprint

Topic: **Towards Certifiable AI in Aviation**
A Framework for Neural Network Assurance Using Advanced Visualization and Safety Nets

Date: September 29, 2024 to October 3, 2024

Author: Johann Christensen

Institute: Institute for AI Safety and Security