Uncertainty-Centric Safety Assurance of ML-Based Perception for Automated Driving

Krzysztof Czarnecki
Waterloo Intelligent Systems Engineering (WISE) Lab
University of Waterloo
Uncertainty-Centric Assurance of ML-Based Perception

Uncertainty Influence factors (domain coverage, sensor noise, etc.)

Perceptual Uncertainty

Misclassifications, under-classifications, quantitative errors

Perceptual Uncertainty Aware Responsibility Sensitive Safety (PURSS)

Uncertainty Management

Safety requirements on perception
Uncertainty-Centric Assurance of ML-Based Perception

Perceptual Uncertainty

Influence factors (domain coverage, sensor noise, etc.)


PURSS: Towards a Perceptual Uncertainty-Aware Responsibility Sensitive Safety.
Under submission.

Misclassifications, under-classifications, quantitative errors

Perceptual Uncertainty Aware Responsibility Sensitive Safety (PURSS)

Safety requirements on perception
Responsible Sensitive Safety (RSS)

- Defines responsible behavior to address **behavioral uncertainty**
  - Safe actions when safe and proper response when not safe
- Guarantees no collision when everyone follows the rules
Responsible Sensitive Safety (RSS)

**RULE 1.** Do not hit the car in front (longitudinal distance)

**RULE 2.** Do not cut in recklessly (lateral distance)

**RULE 3.** Right of way is given, not taken

**RULE 4.** Be cautious in areas with limited visibility

**RULE 5.** If you can avoid a crash without causing another one, you must

https://arxiv.org/abs/1708.06374
RULE 1. Safe Following Distance in RSS

\[ d_{\text{min}} = \left[ v_r \rho + \frac{1}{2} a_{\text{max, accel}} \rho^2 + \frac{(v_r + \rho a_{\text{max, accel}})^2}{2 a_{\text{min, brake}}} - \frac{v_f^2}{2 a_{\text{max, brake}}} \right] + \]

Distance traveled due to reaction time \hspace{2cm} Braking distance \hspace{2cm} Distance traveled by front vehicle
RULE 1. Safe Following Distance in RSS

\[ d_{\text{min}} = \left[ v_r \rho + \frac{1}{2} a_{\text{max, accel}} \rho^2 + \frac{(v_r + \rho a_{\text{max, accel}})^2}{2 a_{\text{min, brake}}} - \frac{v_f^2}{2 a_{\text{max, brake}}} \right] + \]

Problem: Assumes perfect perception
Perception Triangle

Real-world situation

True state (unknowable)

Pedestrian speed = 0 activity = standing

Perception

Pedestrian speed = 0.1 activity = walking

Accuracy
Safety Argument Decomposition

ADS

Sensing → Perception → World model → Planning & control → Actuation
RSS as a Constraint on ADS

RSS

Sensing  →  Perception  →  World model  →  Planning & control  →  Actuation

ADS

Sensing  →  Perception  →  World model  →  Planning & control  →  Actuation
Sample RSS-Compliant World Model Schema

Safe following distance

Safe action set Safe(s)
Perception Cases ($s \rightarrow s'$)

Correct Perception

$s \rightarrow s'$ where $s = s'$

Misperception

$s \rightarrow s'$ where $s \neq s'$

Real-world situation

True state (unknowable)

$s$

Pedestrian speed = 0
activity = standing

Perception

$s' = s'$

Pedestrian speed = 0
activity = standing

Real-world situation

True state (unknowable)

$s$

Pedestrian speed = 0
activity = standing

Perception

$s \neq s'$

Pedestrian speed = 0.1
activity = walking
Safety of Perception

Misperception $s \rightarrow s'$ potentially causes safety risk iff

$$\text{Safe}(s') \not\supseteq \text{Safe}(s).$$
Safety-Irrelevant Misperceptions

Misperception $s \rightarrow s'$ where $\text{Safe}(s) = \text{Safe}(s')$
Precise World Model

Real-world situation

True state (unknowable)

Pedestrian speed = 0
activity = standing

Perception

Pedestrian speed = 0.1
activity = walking

Accuracy
Perceptual Uncertainty Handling via Imprecise World Models

Real-world situation

True state (unknowable)

Pedestrian speed = 0
activity = standing

Perception

Imprecise World Model

Pedestrian speed = 0
activity = standing

Pedestrian speed = 0.1
activity = walking

Set of credible states at conf. level $\alpha$
Perceptual Uncertainty Aware RSS (PURSS)

RSS – rules lifted to imprecise world model

Imprecise Perception \rightarrow \text{Imprecise world model} \rightarrow \text{Planning & control} \rightarrow \text{Safe Actions}

ADS

Situation

Imperfect Perception \rightarrow \text{World model} \rightarrow \text{Planning & control} \rightarrow \text{Action}
Lifting World Model Schema to Imprecise World Model Schema

Elementwise lifting:
• Class entity to superclass
• Continuous value to interval
• Discrete value to enumerated set
• Derived attributes via set operations and interval arithmetic
Using Imprecise World Models to Mitigate Misperception

Given an under-perception case, where $S$ is an imprecise model of confidence $\alpha$ perceived when the correct model:

$$s \rightarrow_\alpha S$$

A safe action in an imprecise model must be safe for every precise model covered by the imprecise model.

$$\text{Safe}(S) = \bigcap_{s_i \in S} \text{Safe}(s_i)$$
Different Risk Levels

\[ \alpha = 10^{-4} \quad \alpha = 10^{-9} \]

\[ \alpha = 10^{-4} \quad \alpha = 10^{-9} \]
Imprecise Classification when High Integrity Required

\[ \alpha = 10^{-4} \]

\[ \alpha = 10^{-4} \]

\[ \alpha = 10^{-9} \]
Conservative Action for High Integrity

\[ \alpha = 10^{-4} \]

\[ \alpha = 10^{-4} \]

\[ \alpha = 10^{-9} \]

\[ \text{Safe}(S) = \bigcap_{s_i \in S} \text{Safe}(s_i) \]
Example of Mitigation

Any

No Lane Obstruction in Front

Lane Obstruction in Front (LOF)

- Static LOF
- Front Vehicle

Actions: continue or stop or follow
Safety Requirements on Perception Performance from PURSS

Any

No Lane Obstruction in Front

Lane Obstruction in Front (LOF)

Static LOF

Front Vehicle

Correct LOF/NLOF classification and distance ±5 cm at $\alpha_{LOF} = 10^{-9}$ for 100% of time duration within ODD conditions

Correct FV/SLOF classification and distance ±25 cm and velocity ±0.5 m/s at $\alpha_{FV} = 10^{-4}$ for 90% of time duration within ODD condition
Uncertainty-Centric Assurance of ML-Based Perception

Perceptual Uncertainty

Uncertainty Influence factors (domain coverage, sensor noise, etc.)

Uncertainty Management

Safety requirements on perception

K. Czarnecki and R. Salay. Towards a Framework to Manage Perceptual Uncertainty for Safe Automated Driving. WAISE’18

Misclassifications, under-classifications, quantitative errors

Uncertainty Aware Responsibility Sensitive Safety (PURSS)
Guide to the Expression of Uncertainty in Measurement (GUM)

- True accuracy unknowable
  - Accuracy in ML wrt. test set only
- Must estimate uncertainty
Perception Triangle (Instance-Level)

Real-world situation

True state (unknowable)

Set of credible states (uncertain)

Pedestrian speed = 0
activity = standing

Pedestrian speed = 0.1
activity = walking

Perception algorithm

Perception

Sensory channel

Camera image, radar data

Accuracy
Perceptual Triangle

Real-world situation

True state (unknowable)

Pedestrian speed = 0 activity = standing

Pedestrian speed = 0.1 activity = walking

Set of credible states (uncertain)

Perception

Accuracy

Sensory channel

Perception algorithm

Camera image, radar data

Instance-level

Domain-level (generic)
Perceptual Triangle When Using Supervised ML

**Development**
- Concept
- Partial semantics (examples)
- Sensory data
- Sensory channel
- Data labeling
- Model class selection, training & testing
- Trained Model
- Training & testing
- Development situations and scenarios

**Operation**
- Concept
- Sensory data
- Resulting perception
- Inferred state
- Operational situations and scenarios
- Sensory channel
- Inference
- Operation situations and scenarios
- Sensory channel
Factors Influencing Uncertainty (F1-7)

**Development**
- **Concept**
  - Data labeling (F5)
  - Model class selection, training & testing (F6)
- **Training & testing**
  - Partial semantics (examples) (F2)
  - Sensory channel (F3)
- **Development situations and scenarios**
- Domain shift (F7)

**Operation**
- **Operational situations and scenarios**
  - Sensory channel (F2, F3, F4)
  - Resulting perception
- **Inference**
  - Inferred state
  - Concept
- **Trained Model**

K. Czarnecki and R. Salay. Towards a Framework to Manage Perceptual Uncertainty for Safe Automated Driving. WAISE’18
Factors Influencing Uncertainty (F1-7)

Development

- Concept
  - Data labeling (F5)
  - Model class selection, training & testing (F6)
- Development situations and scenarios (F2)
  - Partial semantics (examples) (F1)
- Sensory data (F3)
  - Sensory channel (F4)
- Inference
  - Resulting perception (F2)
  - Inferred state (F4)

Operation

- Operational situations and scenarios (F7)
- Concept
- Sensory data
- Inferred state

Trained Model

K. Czarnecki and R. Salay. Towards a Framework to Manage Perceptual Uncertainty for Safe Automated Driving. WAISE’18
F3: Scene Uncertainty
F3: Scene Uncertainty

• Surrogate measures
  – range, scale, occlusion level, atmospheric visibility, illumination, clutter and crowding level

• Also part of development data set coverage

• To determine sufficient coverage, compare these measures with
  1. Test set accuracy
  2. Estimated uncertainty by the network
Synthetic Dataset to Study Scene Influence Factors

Scene Influence Factors -> Accuracy
Aleatoric and Epistemic Uncertainty

Predictive Entropy (PE) = H(E(p))

Aleatoric Entropy (AE) = E(H(p))

Mutual Information (MI) = PE - AE  (Epistemic Uncertainty)

Scene Influence Factors -> Uncertainty Estimates

(a) Input Image  (b) Ground Truth  (c) Prediction

(d) Aleatoric Uncertainty  (e) Epistemic Uncertainty  (e) Predictive Uncertainty
Occlusion and Depth -> Uncertainty Estimates

Buu Phan, Samin Khan, and Rick Salay, and Krzysztof Czarnecki. Bayesian Uncertainty Quantification with Synthetic Data. In Proceedings of International Workshop on Artificial Intelligence Safety Engineering (WAISE), SAFECOMP, Turku, Finland, 2019
Occlusion and Depth -> Uncertainty Estimates

|D| = 500

|D| = 3000

|D| = 8000

|D| = 13100
Rain, Clouds, Puddles -> Uncertainty Estimates

(a) Aleatoric Uncertainty Estimates

(b) Epistemic Uncertainty Estimates

(c) Predictive Uncertainty Estimates

(d) mIoU
Coming Soon: Canadian Adverse Driving Conditions Dataset
Summary: Uncertainty-Centric Assurance of ML-Based Perception

- Uncertainty Influence factors (domain coverage, sensor noise, etc.)
- Uncertainty Management
- Perceptual Uncertainty
- Misclassifications, under-classifications, quantitative errors
- Perceptual Uncertainty Aware Responsibility Sensitive Safety (PURSS)
- Safety requirements on perception
Insights and Challenges

- ML currently cannot be assured to certainty levels required for collision avoidance
  - ML is useful for longer-term, anticipatory risk reduction
- Perceptual uncertainty must be considered for the complete, fused perception and over time
  - E.g., different information becomes certain with different delays
- Out-of-distribution detection is still far from being useful in practice
- RSS leads to more conservative automated driving than human driving
  - E.g., negotiation in merging