12th UIC WORLD CONGRESS ON HIGH-SPEED RAIL
8-11 July 2025 - Beijing, China

New Methods for Safety Demonstrations

Machine Learning in Railway Vital Systems

Frédéric HENON

Director Freight, Operations & Safety – UIC - France

4.4 - Rolling Stock & New System Performance

Using New Technologies within Vital Systems

New technologies shall not be banned for safety critical systems

- Even if it is easier
  - Known technologies already used in known vital systems
- Otherwise: limited functions, outdated or unsupported systems

AI and machine learning, cloud, IoT, sensor fusion, distributed & microservices

- Safety: risk analysis, safety demonstration, safety case
  - Applicable for any new vital technologies

  - **But** difficult (impossible ?) to apply for **AI and Machine Learning**
    - *Because new and unknown design after training (ML technology learns from data)*

Machine Learning vs Safety & Standards

Machine learning (ML):

- Training Dataset → neural network → trained neural network → system including trained neural networks

Safety & Safety Standards:

- EN50126: mastering the system's design process, safety based on "how the system is built"
- EN50129: safety demonstration: mitigating all logged systematic & random errors (as much exhaustive as possible)
- EN50128 (EN50716): each software part shall match its detailed specifications

ML is "black box" (internal decisions not understandable) vs Safety is "white box" (transparent system where every decision can be verified)

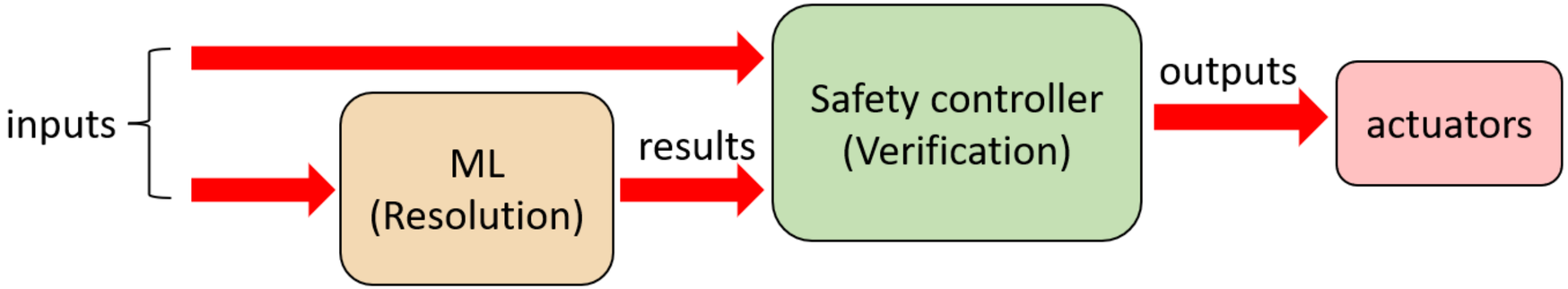
- algorithms used by trained ML are not known by design => systematic errors keep being possible

Machine learning trained to provide verifiable proofs

Possibilities

- Turn the trained ML into a verifiable algorithm: *very difficult even impossible …*
- Train the ML to provide “verifiable results” + “safety verification”
  - Proposed in MIT article “Certified Control: An Architecture for Verifiable Safety of Autonomous Vehicles” (Daniel Jackson and Al. – 2021).

→ ML with verifiable results + Safety Controller (SIL objective)

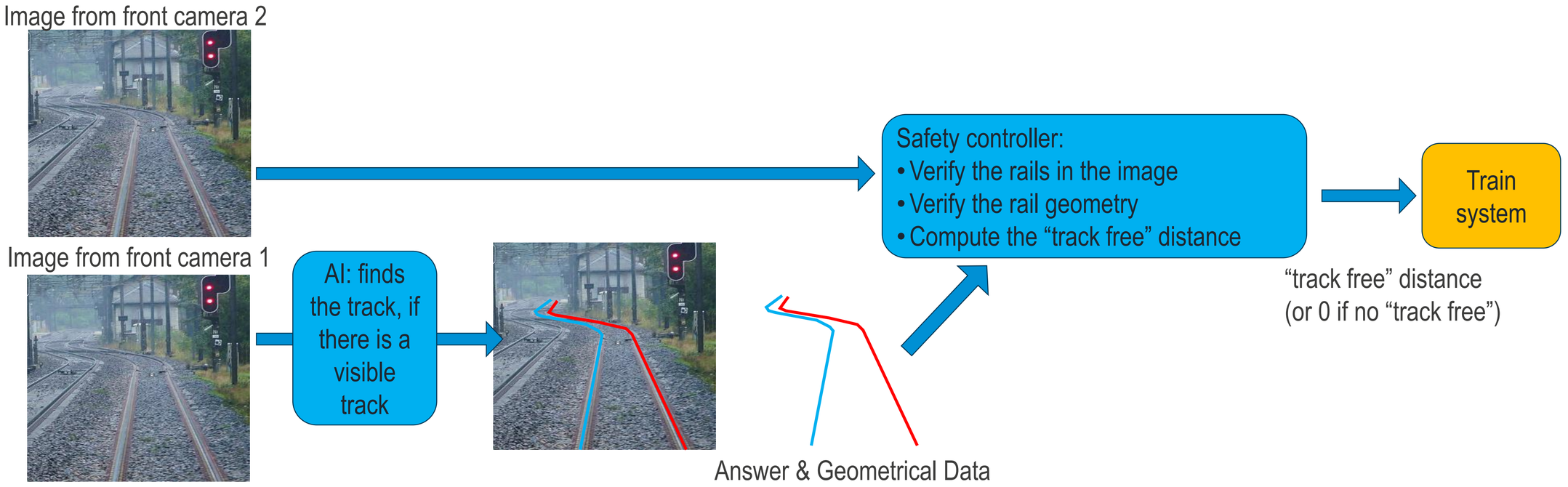


- The Safety Controller shall see the inputs (for verification)
- And the ML results must be chosen to be “verifiable”

Proof Of Concept: “track free” detection

ML with verifiable results + Safety Controller: deserves a POC- proof-of concept

Selected Scenario : **“track free” detection**



“Track Free” detection: Safety Concept

Machine learning: finds the track free path

Safety Controller:

- Verifies that rails are visible at left / right sid
  - Using the 2nd camera
- Verifies the geometry of the path: shall match 1435mm spaced rails, parallel (even if curved), with compatible turns & slopes
- Computes the free track distance from the path

Nothing else than an actual free track can look like 2 parallel lines separated by 1435mm in front of the train

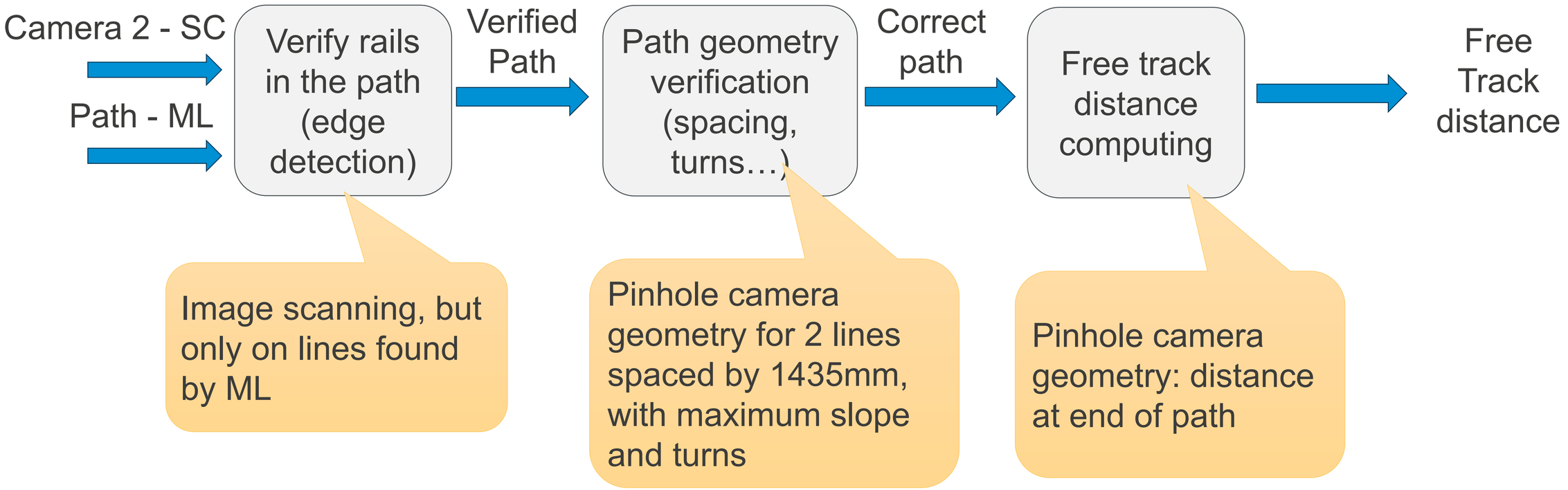
- Correct whatever the ML part does
  - As long as the safety controller is correct (safety part)
- Paths from each camera 1 & camera2, match only if both cameras are working correctly then consistently (including orientation)

ML processes camera image
→ suggests path

Safety controller
→ verifies geometry

Decision taken
→ only if path is safe

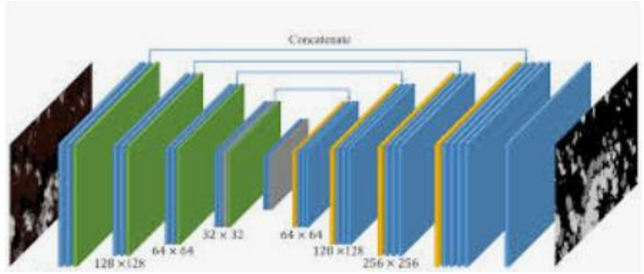
Track Free detection: safety controller



- Safety Controller must be implemented in safety
  - Without ML=> safety by process / proof to specifications

Free track detection: the ML part

Chosen ML model: **U-NET**



- Known for image segmentation
- Trained from scratch using images generated with a train game (MS Train Simulator)
  - Generation of image + path
- Training set 650 images + paths
  - 10 epochs training
- Accuracy obtained ~ 90%

- *Could be improved: training set a bit small and unbalanced…*

Training set

ML found path

Free track detection: safety properties & demonstration

Safety property P1: the output distance shall always be lower than the actual free track distance

- So that the train will stop before reaching the first obstacle
  - *(With some details in case of multiple tracks)*

System safety demonstration method: property oriented

- Based on the controller only: no safety constraint based on ML

1 step example: P1 is ensured by:

- P2: a path validated by the safety controller actually denotes a free track
- P2th: the geometrical formulas used to get the end of path distance are correct
- P2code: the controller correctly implements these formulas (no bugs)
- P2hard: the controller is implemented on a vital platform (no memory upsets, safe compilation, …)

And so on, until reaching leaf properties (ensured by design, or exported safety conditions)

Conclusions: method & proof-of-concept

Machine Learning + Safety Controller: **works** in this case

- Thanks to ML, the controller is quite simple (no need to find the path)
- → **This method works best for problems that are easy to be verified rather than to be solved**

Track Free detection = PoC here, no product

- Using 2D cameras only (real system: sensor fusion)
- No zoom → poor accuracy in the distance
- *But the goal was to ensure the safety*

How the ML + Safety Controller System Works

This system works like a co-pilot / safety inspector:

1. ML suggests a path using camera images
2. The safety controller checks if the path is safe
3. Only validated paths are used to make decisions

This separation ensures that even if ML is uncertain, safety remains guaranteed.

Conclusions

Among new technologies, **AI and machine learning are the most difficult regarding the safety**

- Black box: we never know the decision criteria created by the training…
- Safety is white box oriented

But it is possible to use ML by training it to produce verifiable results

- Results **+ their justifications** as much as possible
- And to develop **safety controllers** to verify those justifications
  - Opening the possibility for the highest safety levels

Thank you
for your attention

CONTACT
Frédéric HENON
Director Operations& Safety, Freight
Tel +33 (0) 6 16 05 05 59
henon@uic.org

HIGHSPEED
中国 · 北京 · 2025