tinyML® Summit

Miniature dreams can come true...

March 28-30, 2022 | San Francisco Bay Area

www.tinyML.org

Plumerai

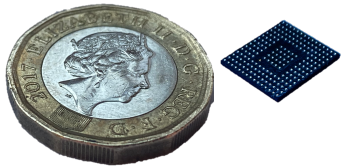
Tiny models with big appetites: Cultivating the perfect data diet

Jelmer Neeven - jelmer@plumerai.com

TinyML Summit - March 29th, 2022

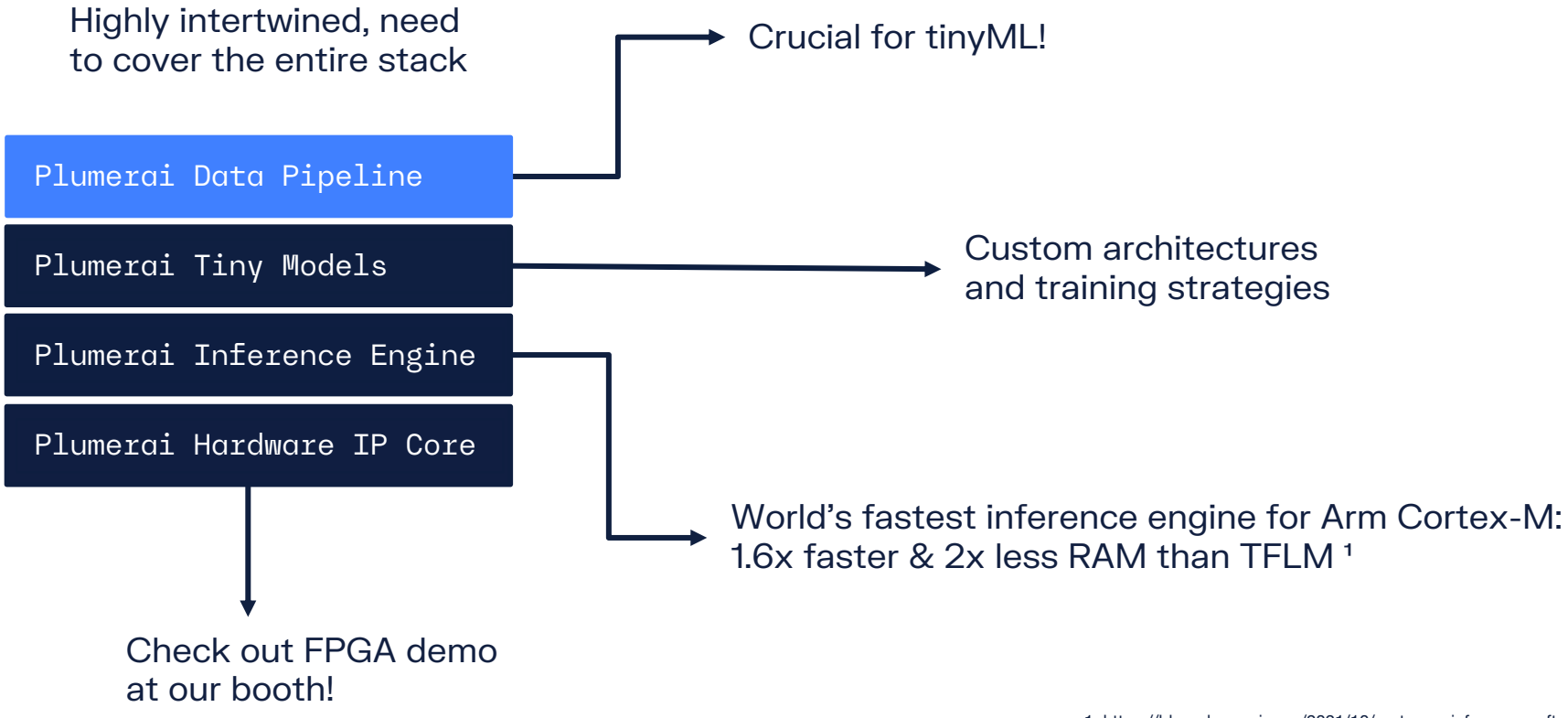
Plumerai people detection

Production-worthy computer vision on Arm Cortex-A SoCs and even $1 Cortex-M microcontrollers.



Come check these demos live at our booth!

How did we do it?

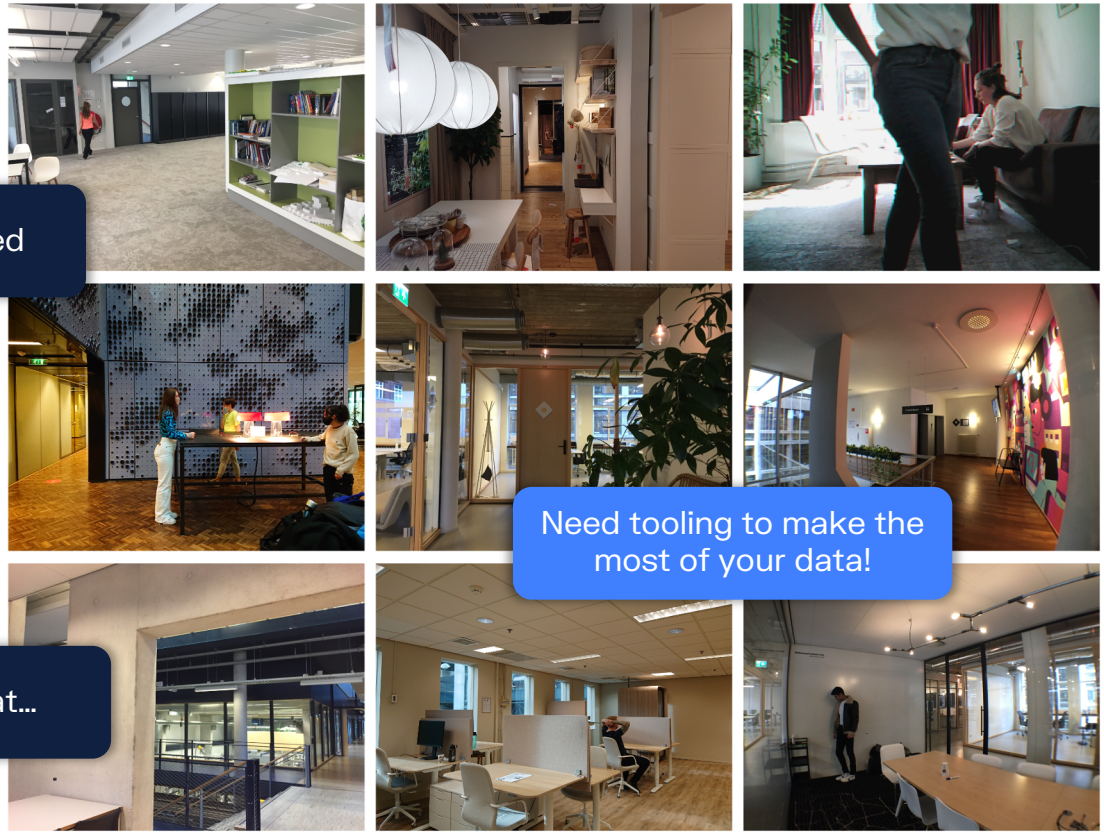


1. https://blog.plumerai.com/2021/10/cortex-m-inference-software/

Why not just use public data?

Plumerai dataset

Not person-centered



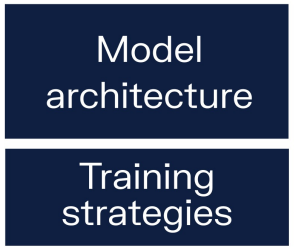
Need tooling to make the most of your data!

Nothing to look at...

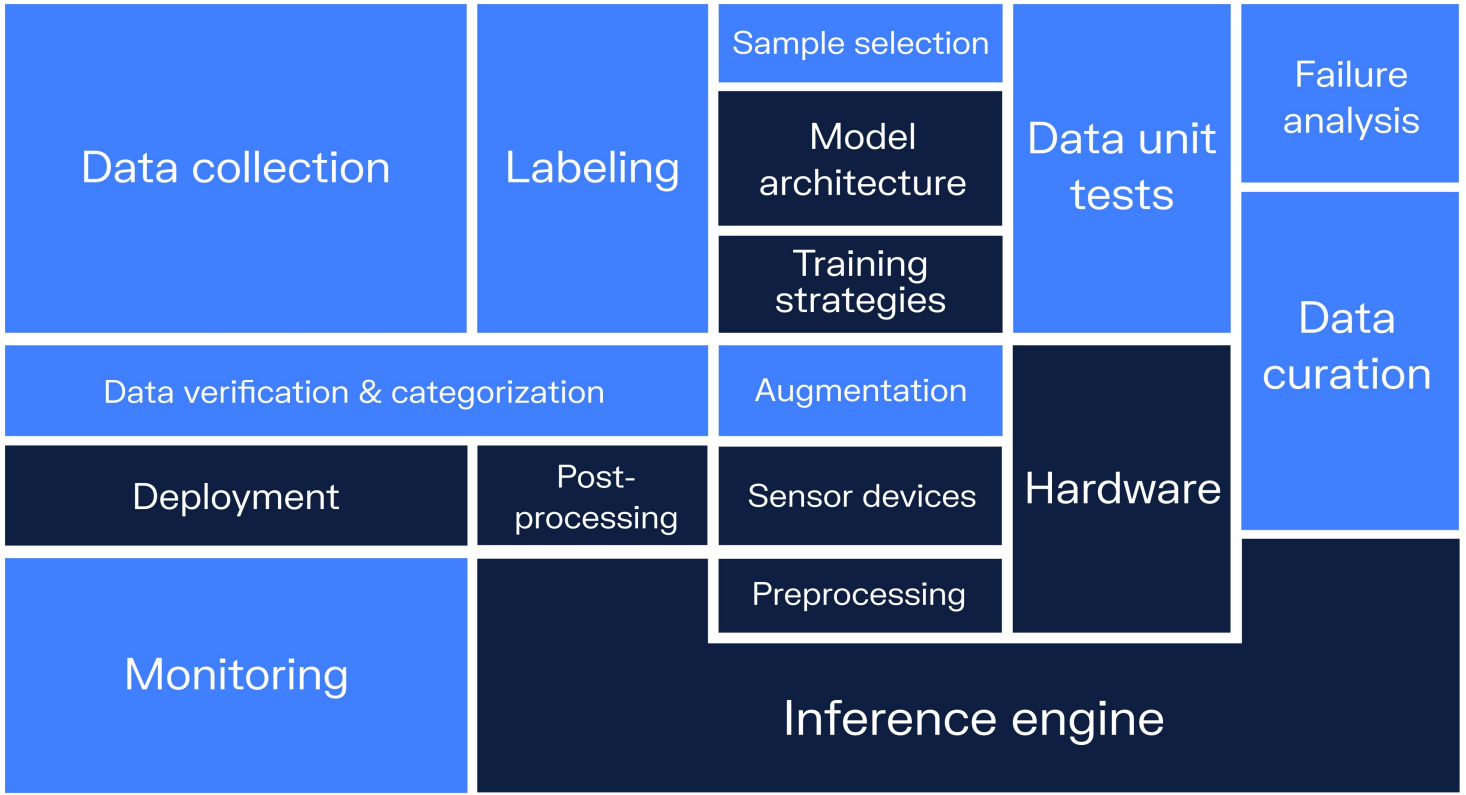
Plumerai

Plumerai infrastructure

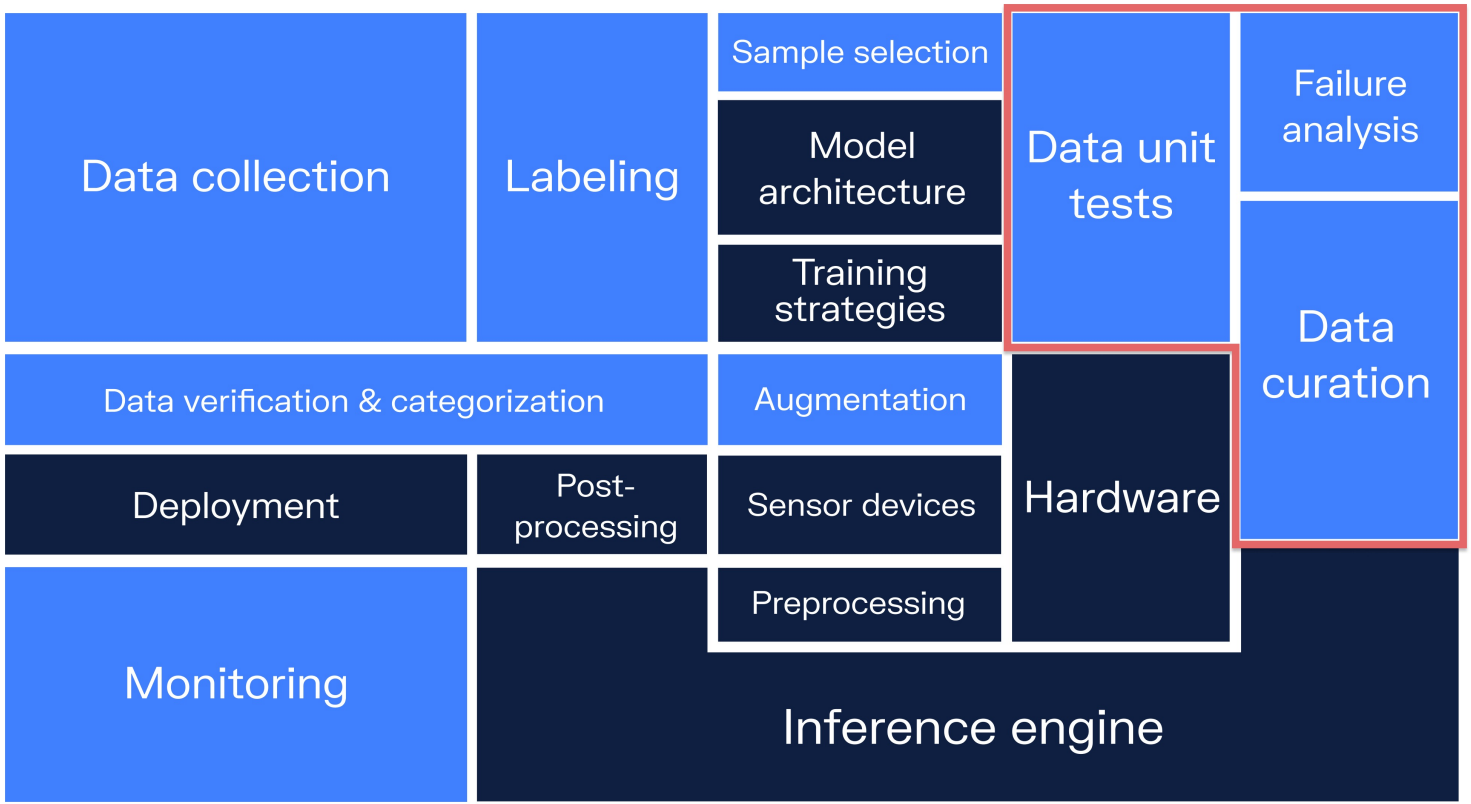
Infrastructure overview



Infrastructure overview



Infrastructure overview



Data unit tests

Cannot capture model behavior in a single metric!

Unit Tests Passed: 73 / 84

✓	empty_desk
✓	empty_hallway
✓	empty_office
✓	person_desk_face
✓	person_desk_back
✓	person_
✓	person_
✓	person_hallway_within_4m
✓	person_hallway_4_7m
✗	person_hallway_beyond_7m
✓	skin_tone_1
✓	skin_tone_2
✓	skin_tone_3

...

Real-life performance

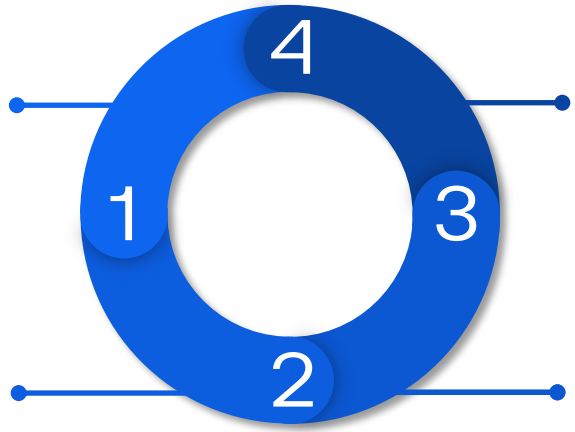
Edge cases

Data curation cycle

Need **good data tooling** to:

Test-driven development

This is an iterative process!



Identify failure cases

Re-train model

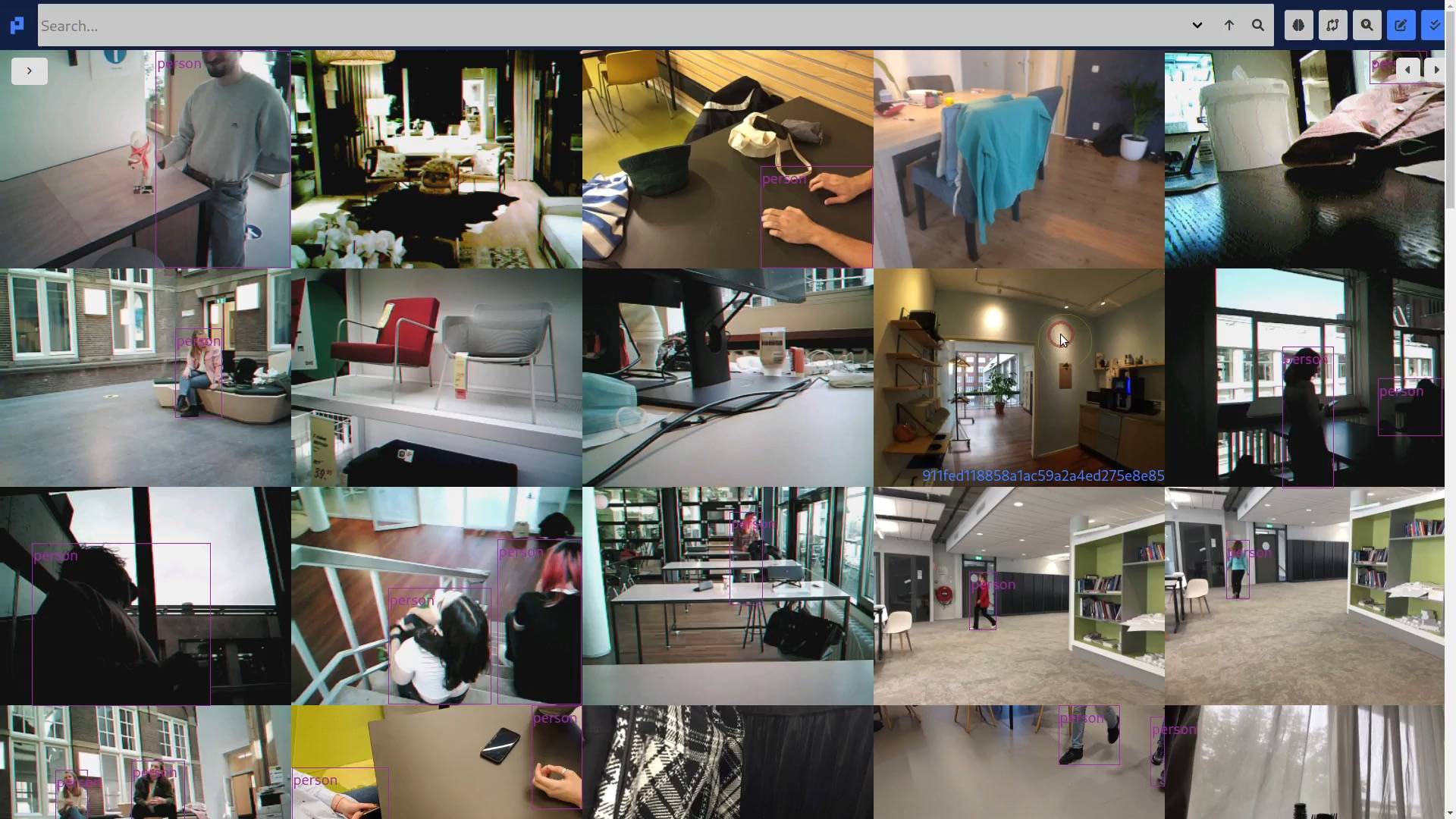
Link problem to training data:

- Visual similarity
- Interactive classification
- Training influence
- …

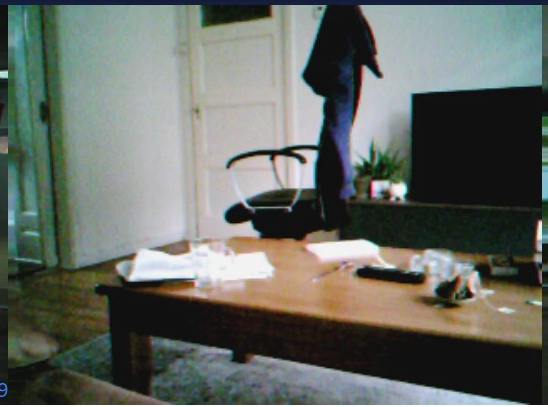
Address problems

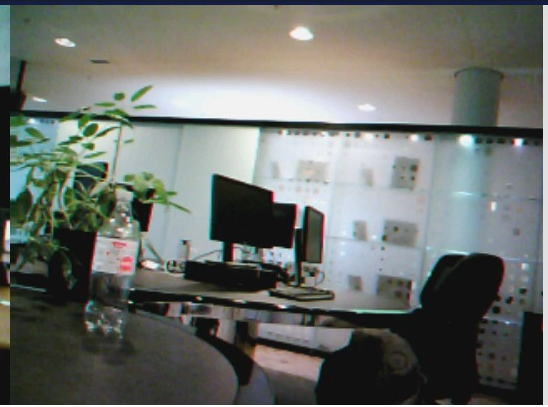
- Correct labels
- Oversampling
- Targeted augmentation
- Supervised Contrastive Learning
- Add / **remove** data
- …

Quality > quantity


Search...
person
911fed118858a1ac59a2a4ed275e8e85

'tinyML_IF_example' in unnest(tags)





Demo recap

- Tooling allows us to identify problematic images:
  - Similar appearance
  - Automatic classification
  - Influence on training process

Demo recap

- Tooling allows us to identify problematic images:
  - **Similar appearance**
  - Automatic classification
  - Influence on training process

Demo recap

- Tooling allows us to identify problematic images:
  - Similar appearance
  - **Automatic classification**
  - Influence on training process

Demo recap

- Tooling allows us to identify problematic images:
  - Similar appearance
  - Automatic classification
  - **Influence on training process**

Demo recap

- Tooling allows us to identify problematic images:
  - Similar appearance
  - Automatic classification
  - Influence on training process

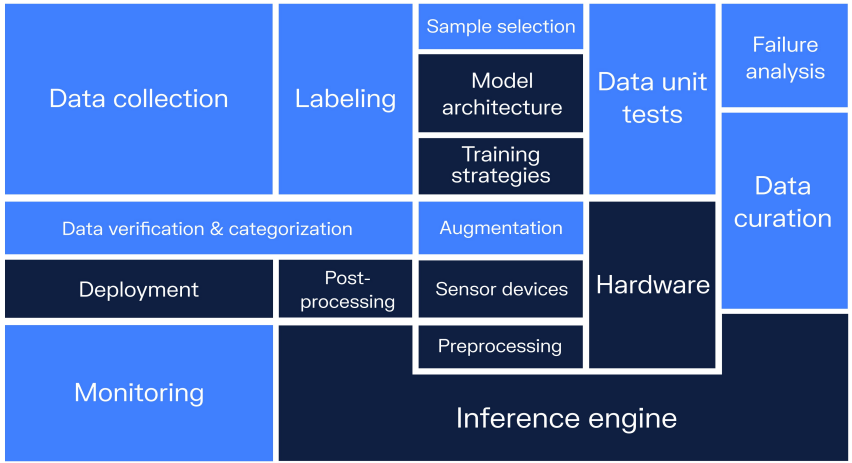
- **Act on them in real-time!**
- **Custom tooling: quickly add / explore new AI-assisted approaches**
  - Needed for scalable data iteration
  - 30M images in total
  - Terabytes of data!

Conclusion

Production-worthy tinyML requires:

- **Vertical integration!**
- Rapid iteration
- High-quality data

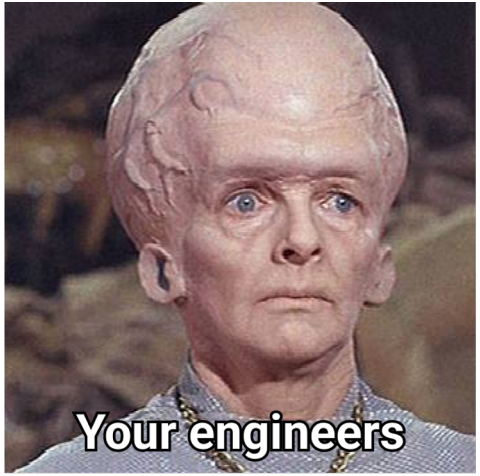
High-quality model!

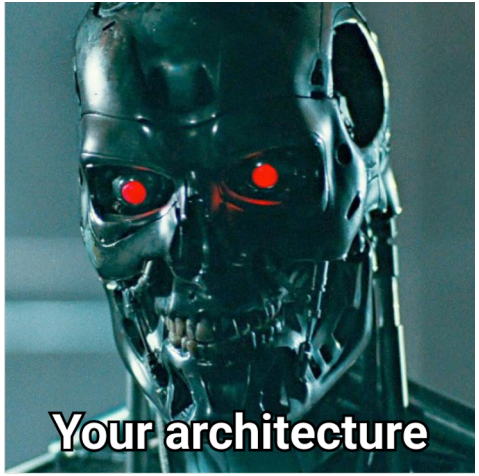


Plumerai

Questions

Jelmer Neeven - jelmer@plumerai.com





Not part of a healthy diet!

Plumerai

Thank you!

Jelmer Neeven - jelmer@plumerai.com

TinyML Summit - March 29th, 2022

tinyML Summit 2022 Sponsors

Copyright Notice

www.tinyml.org