Improve Worksite Safety with AI-Powered Perception

Sabri BAYOUDH

VP Technology

ARCURE BLAXTAIR

Context

- Our mission: Revolutionize worksites by improving both safety and productivity
- Challenges at a worksite
  - Highly complex and dynamic environments
  - Intense interaction between vehicles and pedestrian workers
  - Hazards are omnipresent and can't always be anticipated or managed through logic rules
  - Safety measures should not reduce productivity, they should enhance it

What CNN Can Deliver

EMBEDDED VISION SUMMIT

- Value of an embedded AI vision system
  - Detect and localize
  - Warn the driver
  - Warn the pedestrian
  - Slow down the vehicle

ARCURE | BLAXTAIR

How?

- Adapt the algorithm to meet industrial environment contraints and standards
  - Train using real-world field data that mirrors the actual deployment context
  - Augment the dataset with filters simulating real conditions: weather, dust, visibility, scratch, etc.

How?

- Adapt the algorithm to meet industrial environment contraints and standards
  - Optimize for deployment within an embedded and cost-effective hardware
    - Technical: quantization, pruning, hardware-friendly activations, etc.
    - Design: Select a model architecture that is accelerator-friendly from the start

Luxembourg's Minister of Health stated :

"While health has no price, it does have a cost."

How?

- Aim for a best-in-class product, not a best-in-class algorithm
  - No false alarm, it kills the product

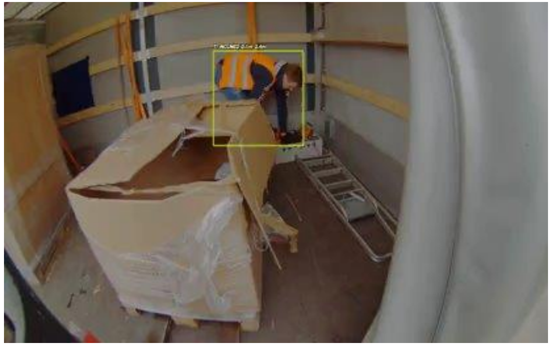
How?

- Aim for a best-in-class product, not a best-in-class algorithm
  - No false alarm, it kills the product
  - Avoid warning the driver when pedestrians are outside the danger zone
    - Requires AI-based localization.

      Enabled through a regression layer added alongside object detection in the model head and supported by a large distance-annotated dataset

How?

- Aim for a best-in-class product, not a best-in-class algorithm
  - Forget average performance metrics, what matters is behavior in all environments in the target market, especially in edge cases
    - Accidents don't happen during clean demos, they happen when unexpected



How?

- Aim for a best-in-class product, not a best-in-class algorithm
  - Forget average performance metrics, what matters is behavior in all environments, especially edge cases
    - Accidents don’t happen during clean demos, they happen when unexpected

What CNNs Can Deliver

- Pro
  - Reduce useless alarms by focusing on pedestrian detection, unlike other techologies (LIDAR, RADAR) which detect all obstacles
- Con
  - Even with pedestrian-only detection, many alerts may still be irrelevant, since workers often move close to vehicles. To improve adoption, the system must better distinguish real danger from normal activity.

Fewer false alarms = more driver trust and better adoption

Using Transformers & Gen AI to Build Efficient and Cost-Effective Algorithms

Context

- From pedestrian detection to risk perception

Approach 1/3

- Ideal solution
  - VLM approach
  - Ask the VLM if the pedestrian is in danger
- Cons
  - Latency: unacceptable to the use case
  - Power consumption: high for an embedded product
  - Cost-efficiency: very expensive

Prompt: Is the pedestrian in danger

Answer: No, he is behind a safety barrier

Approach 2/3

- Intermediate solution
  - Detect pedestrian with bounding box and segment
  - Detect barrier with segmentation
  - Logic rule applied: pedestrian is ignored if fully occluded by a barrier
- Cons
  - Latency: at the edge of acceptable performance
  - Power consumption: just within limits
  - Cost-efficiency: not economically viable

Approach 3/3

- Light solution
  - Add an attribute to the classifier of pedestrians
- Pros
  - Latency: low
  - Power consumption: low
  - Cost effective
- Cons

  Task related and does not generalize. Each attribute requires separate training, but this can be handled efficiently thanks to our extensive field-collected dataset.

Object: Pedestrian
Attribute: Behind Barrier

Get the Best of All Approaches

- Use Approach 1 to identify useless alarms where the person was clearly not at risk: driver, barrier, etc.
- Convert each set of cases into an attribute
- Use Approach 1 to enable semi-automated annotation of the dataset with these new attributes
- Train a lightweight algorithm focused on a single class with a limited set of relevant attributes

Approach Comparison

- Complexity (Cost & Latency Reference)
  - Approach 3 (CNN)
    - Baseline complexity (×1)
  - Approach 2 (Detect & Segment)
    - Approximately 2× more complex
  - Approach 1 (VLM)
    - Up to 100× more complex, with significant cost and latency impact

Approach Comparison

- Performance
  - Performance is use-case dependent
  - In most scenarios relevant to our market, performance is comparable across approaches
  - However, specific edge cases highlight the added value of the more advanced and complete Approach 1 (VLM)

    VLM is overkill but cool in a demo

Approach Comparison

- Performance
  - Approach 1 (VLM)
  - No alert triggered - contextual understanding
  - Approach 2 (Detect and Segment)
  - No alert triggered - pedestrian behind the barrier
  - Approach 3 (CNN)
  - False alert - bounding box does not include the barrier

Pedestrian is safe

Approach Comparison

- Performance
  - Approach 1 (VLM)
  - 🙂 No alert triggered - contextual understanding
  - Approach 2 (Detect and Segment)
  - 😐 No alert triggered - pedestrian behind the barrier
  - Approach 3 (CNN)
  - 😐 Possible alert - if the barrier is not in the database

Approach Comparison

- Performance
  - Approach 1 (VLM)
  - 🙂 No alert triggered - contextual understanding
  - Approach 2 (Detect and Segment)
  - ☹ False alert - pedestrian not behind the barrier
  - Approach 3 (CNN)
  - ☹ False alert - bounding box does not include a barrier

Pedestrian is safe

What About the Hardware?

Hardware Performance and Price Evolution

- Performance improves 12x when switching from FP32 to tensor-INT8 (link)
- Leading ML hardware becomes 40% more energy-efficient each year (link)
- Performance per dollar improves around 30% each year (link)
- The computational performance of machine learning hardware has doubled every 2 years (link)

Source "Epoch AI" focused on high end processing platforms

Wrap up and Key Takeaways

- General AI acceleration fuels progress in embedded AI
- Task-specific models are economically viable, but limited in scope
- Hardware and AI evolve fast, anticipate to avoid early obsolescence
- Adopt an iterative approach: deliver value early and build on it
  - A staged rollout must meet professional quality and reliability requirements
  - Quality and performance are non-negotiable, even in early iterations

Resources

Blaxtair product
https://blaxtair.com/en

Use cases of AI powered perception
https://www.youtube.com/@arcureblaxtair8716

Epoch AI, Trends https://epoch.ai/data-insights/

COCO Database
COCO - Common Objects in Context