IRIM

CAICS
Artificial Intelligence
and Cybersecurity Center

BUILDING AND REGULATING THE METAVERSE

Machine Learning Technologies

Piotr Skrzypczyński

Poznan University of Technology
Institute of Robotics and Machine Intelligence

EMRN
The European Metaverse Research Network

IEEE | IEEE COMPUTER SOCIETY | IEEE COMPUTER SOCIETY VGTC | Inria

Introduction – ML in AR/VR

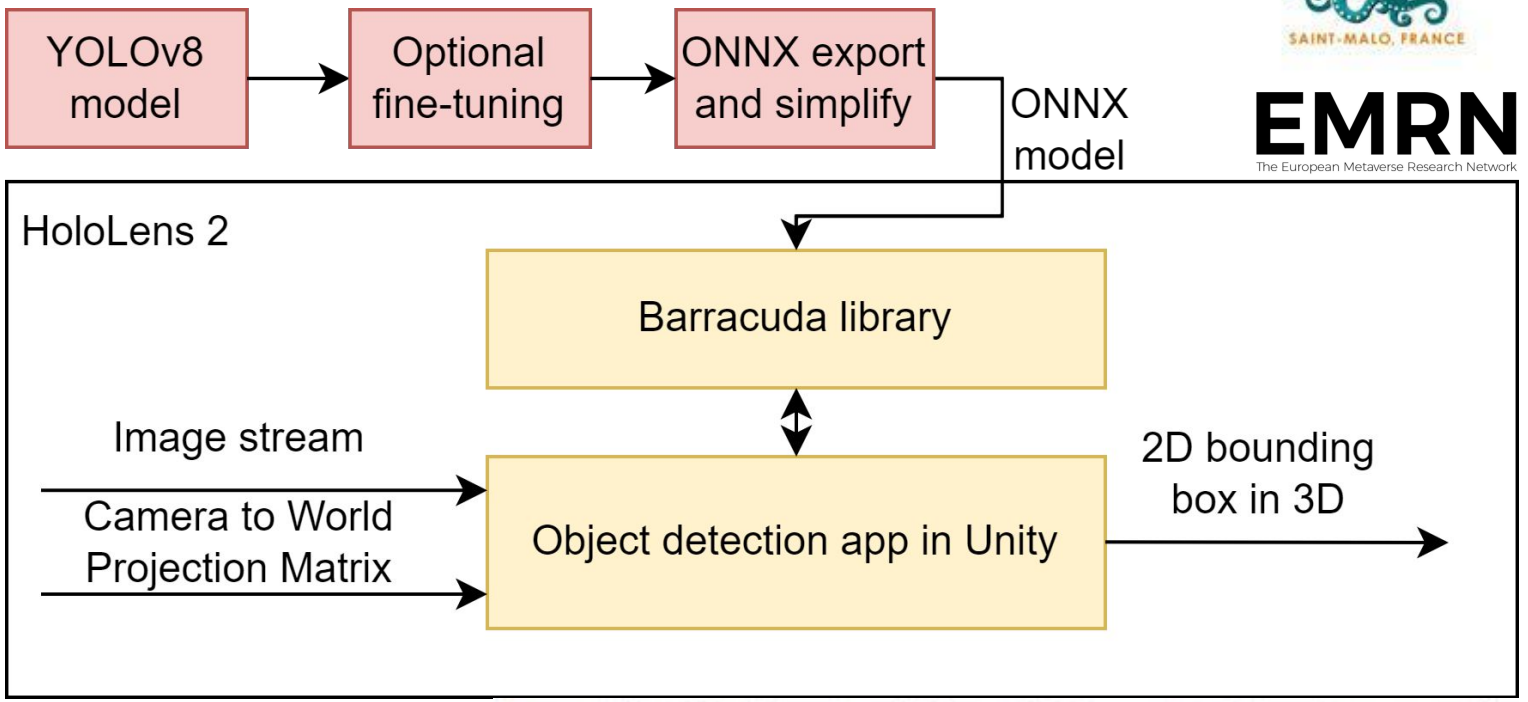
Machine Learning (ML) is revolutionizing AR/VR/XR by enabling:

- **Enhanced perception:** real-time object detection, tracking, and scene understanding.
- **Adaptive decision-making:** intelligent feedback, personalized recommendations, and real-time adjustments.
- **Improved interaction:** more intuitive and immersive user experiences.
- AR/VR is becoming a **core part of the metaverse**, but for real-world applications, it must be **adaptive and efficient**.
- **ML bridges the gap** between static virtual overlays and dynamic, responsive AR environments.
- Use cases span **industry, education, and healthcare**, each requiring different ML techniques

ML-Powered Object Detection in XR

Real-time ML inference on AR devices enables independence from cloud processing, reducing latency.

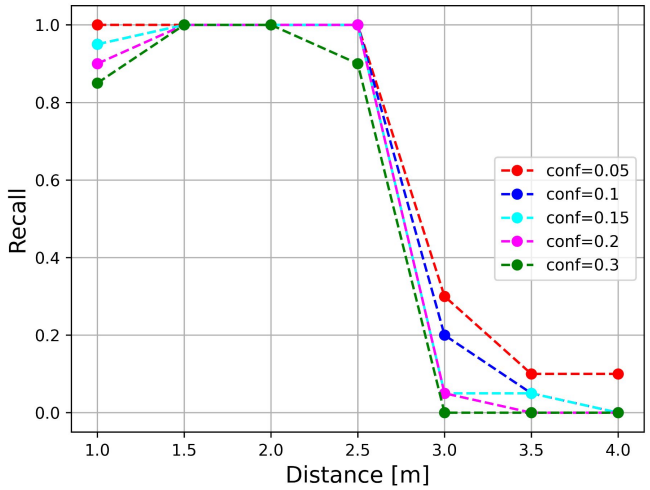
- **Platform:** Microsoft HoloLens 2 with YOLOv8 for real-time object detection.
  - On-device inference vs. server-side computing.
  - Balancing **model size, inference time, and accuracy** (YOLOv8n vs. YOLOv8s).
  - **YOLOv8n (smallest model) achieved real-time performance** (latency < 100ms for 160×160 images).
  - Other models offered better accuracy but required optimizations to run efficiently on resource-limited AR hardware.

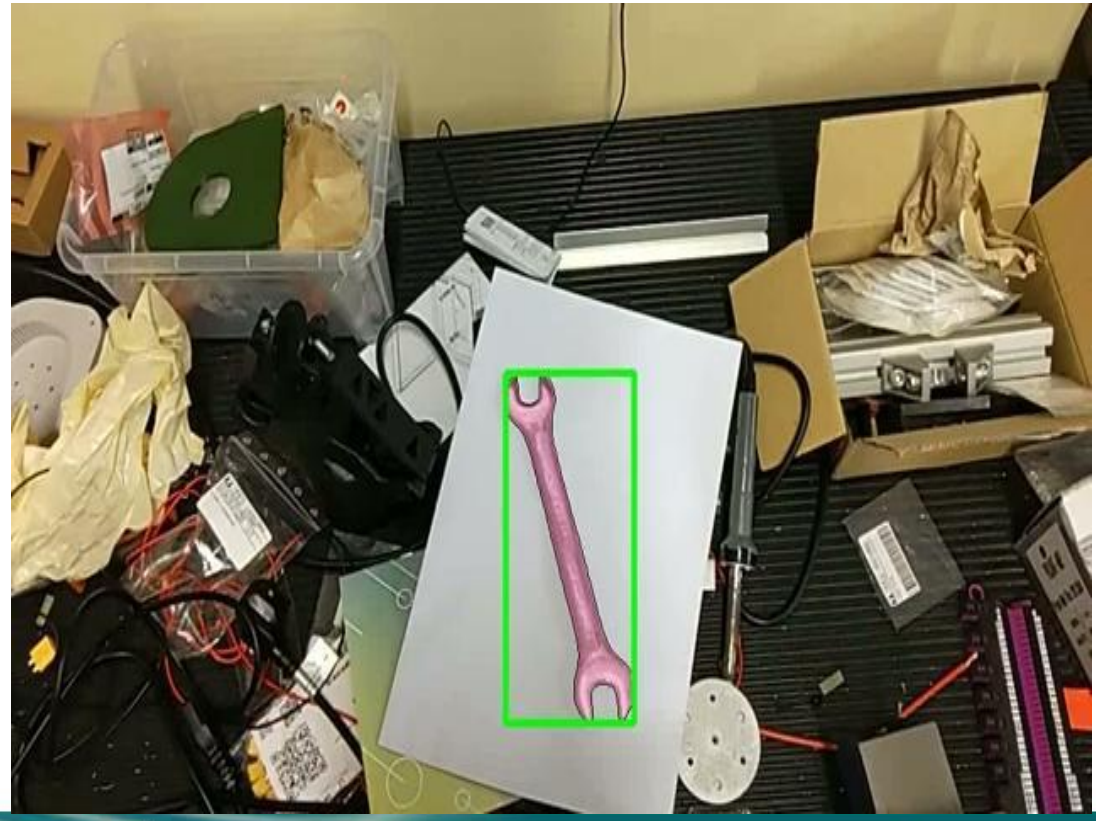


Industrial Application – Visual Search

AR-based ML can assist operators in locating and identifying objects in complex environments.

- **Scenario:** HoloLens 2-based **visual search task** in an industrial workshop.
- **Method:** YOLOv8 object detection, tracking, and augmented search cues.
- **User Study Results:**
  - **No significant reduction in search time**.
  - Need for **more intuitive interaction methods** (e.g. confidence-based overlays).

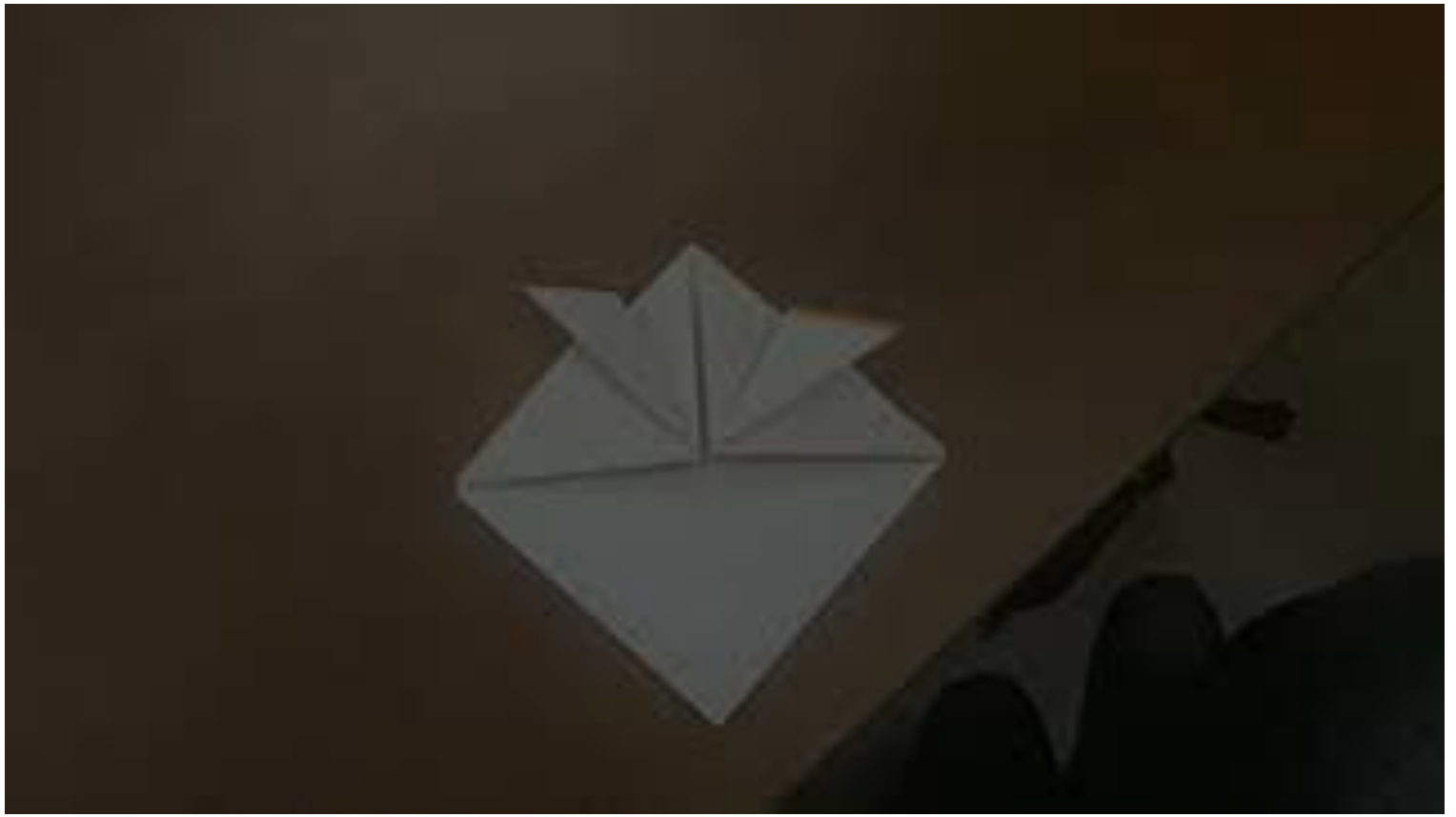




ML-Enhanced Education in XR

ML-powered XR can guide users step by step in manual task completion (origami folding)

- **Approach:**
  - YOLOv8 detects origami folding steps and provides real-time feedback.
  - AR overlays guide users through the folding process.
- **User study results:**
  - **Improved confidence**, but **task completion times increased** due to additional cognitive load.
  - **Users preferred optional feedback** rather than constant validation.

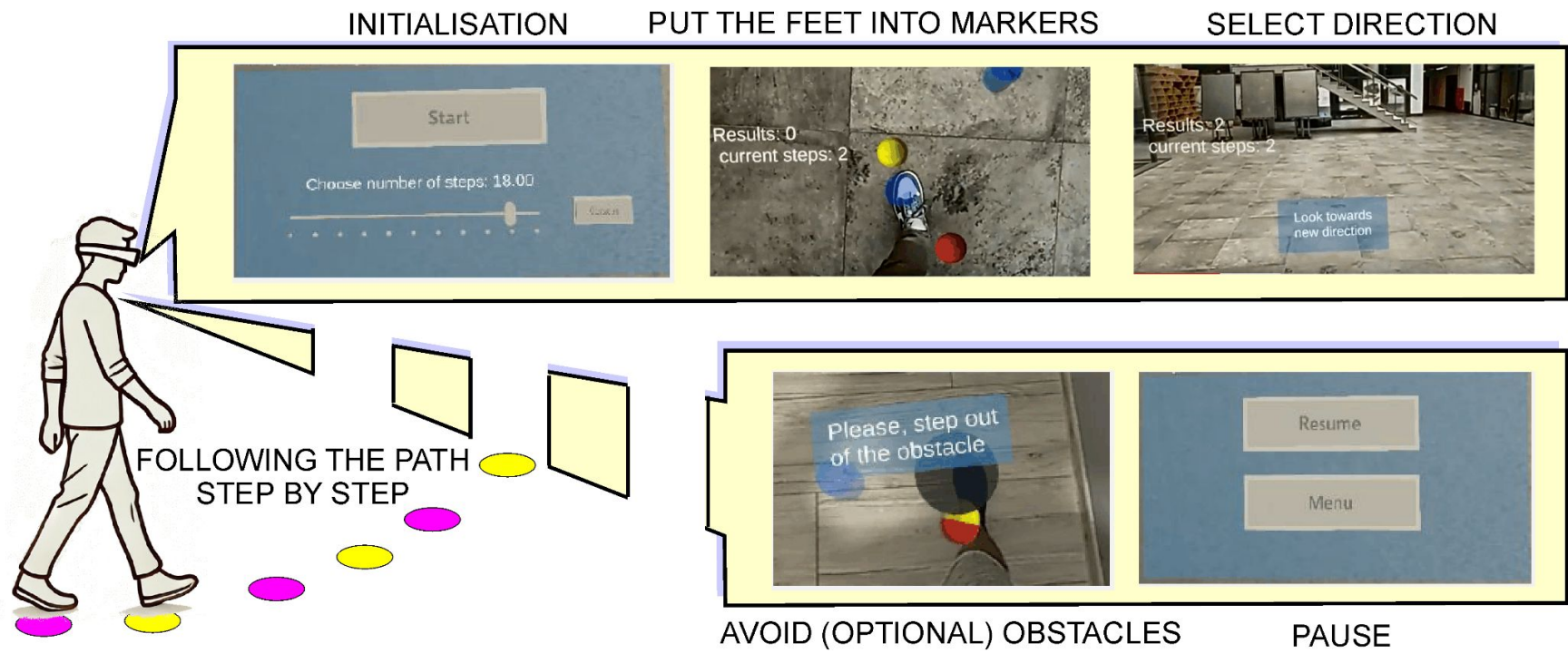


ML for XR-Based Rehabilitation

ML-powered XR enables autonomous rehabilitation for neurological patients by guiding lower limb exercises with real-time feedback using only an RGB camera in AR glasses.

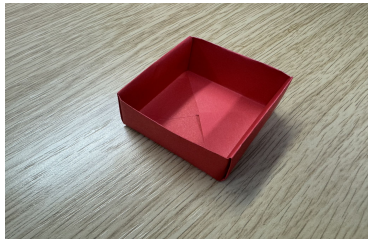
Platform: HoloLens 2 with YOLOv8-Pose for lower limb tracking.

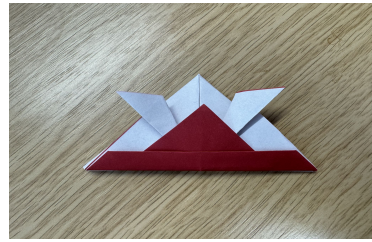
- Detects foot placement & movement patterns.
- Provides real-time step validation & feedback using virtual markers.
- Allows adaptive training paths, including virtual obstacles for dynamic rehabilitation.
- No external sensors or markers required → Cost-effective & scalable for home use.



Challenges

1. **Computational constraints:**
   - Need for **lightweight ML models** optimized for **edge AI**.
   - Balancing **power consumption and performance** for real-time inference.
2. **User trust and interaction:**
   - **False positives** and confidence levels affect usability.
   - Adaptive feedback vs. constant validation—how to **balance guidance** without disrupting user flow?
3. **Scalability and generalization:**
   - ML models should **adapt to diverse environments and user behaviors**.
   - **Transfer learning and federated learning** to improve adaptability.
4. **Security and privacy in ML-driven AR/VR:**
   - **On-device ML vs. cloud processing: edge AI can reduce security and privacy risks** by minimizing data transmission.
   - **Preventing adversarial attacks:** ML models in AR can be **vulnerable to manipulated inputs**, leading to misinformation or incorrect feedback.





Conclusion & Discussion

Final takeaways:

- ML transforms AR/VR by enabling real-time perception, intelligent decision-making, and personalized interactions.
- Applications in industry, education, and healthcare demonstrate its potential, but optimization is needed to improve usability and efficiency.
- Security and privacy must be prioritized in the design of ML-driven AR/VR, especially in sensitive applications like healthcare and education.
- A major challenge for the metaverse: balancing user immersion with trust, security, and cognitive load in ML-powered environments.

Open discussion:

- What are the biggest hurdles in making ML-driven AR/VR truly secure and privacy-aware?
- How can we enhance user trust in AI-powered virtual environments?
- What technological breakthroughs are needed for seamless and ethical ML integration in AR/V?