Current Trends in Machine Vision

General Advances and Edge Deployment Challenges

Deutsches Forschungszentrum für Künstliche Intelligenz
German Research Center for Artificial Intelligence

Alain Pagani | DFKI

DFKI – German Research Center on Artificial Intelligence

dfki ai
Deutsches Forschungszentrum für Künstliche Intelligenz
German Research Center for Artificial Intelligence

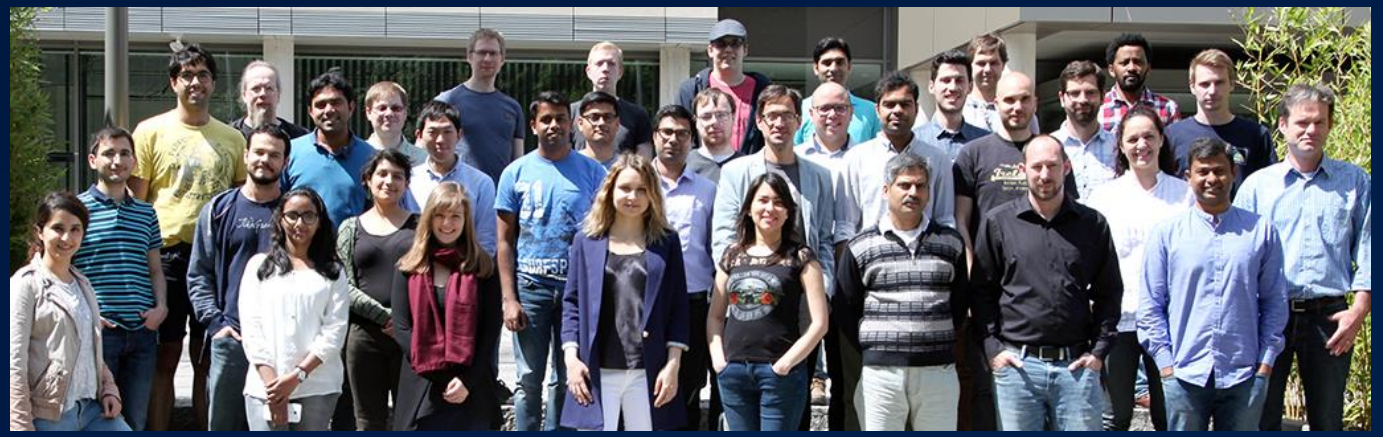
- Since 1988
- 9 locations
- >1000 employees
- All aspects of Artificial Intelligence

DFKI – Department Augmented VIsion

dfki Deutsches Forschungszentrum für Künstliche Intelligenz
German Research Center for Artificial Intelligence
augmented VISION

Augmented Vision department

- Founded in 2008
  Prof. Didier Stricker
- 50 employees and PhD students

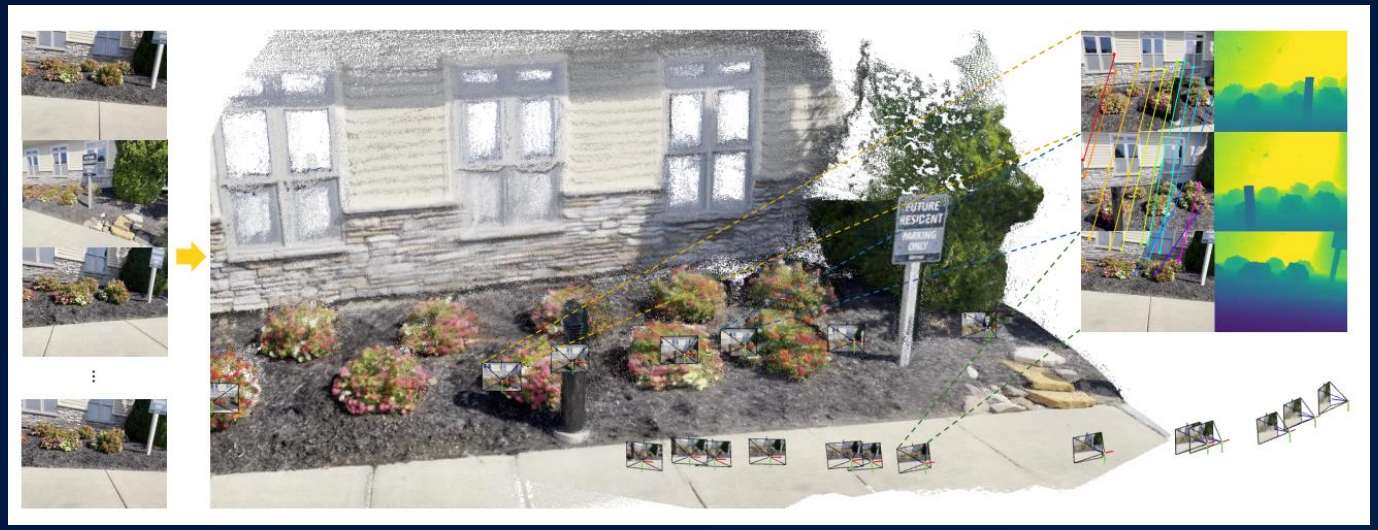


Computer Vision, artificial perception, intelligent sensors, cognitive systems, extended reality

Why Does Machine Vision Matter?

- Key enabler with increasing demand across industries: robotics, AR/VR, surveillance, autonomous systems
- Model complexity vs. compute constraints
- Need for real-time, privacy-aware, energy-efficient vision

Trend #1: 3D Vision and Scene Reconstruction

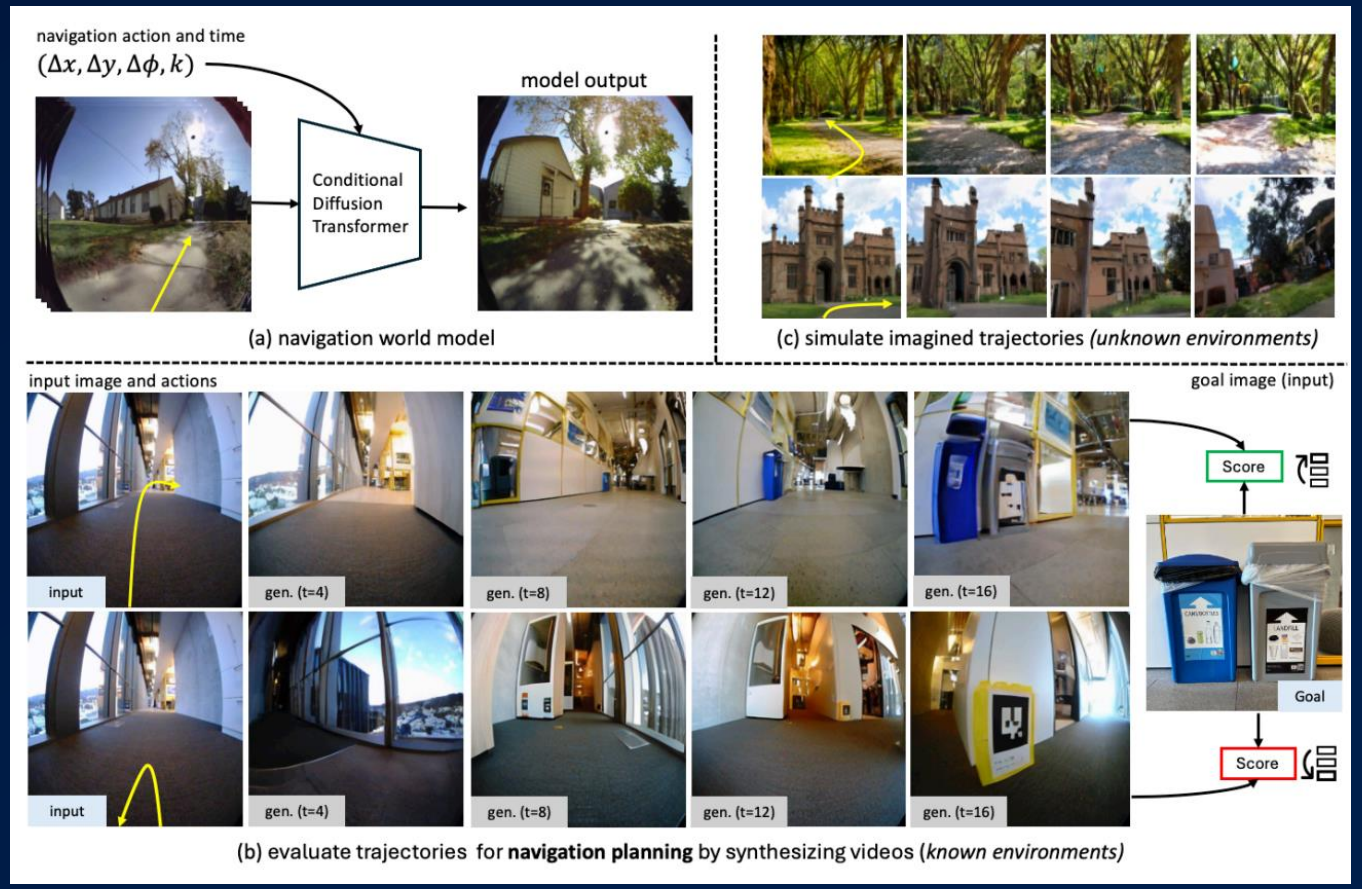


Source: Visual Geometry Grounded Transformer

- Explosion of Gaussian Splatting / Neural rendering variants, monocular depth, SLAM
- CVPR 2025: VGGT - Visual Geometry Grounded Transformer
- Used in AR, autonomous navigation, digital twins

Trend #2: Generative Vision Models

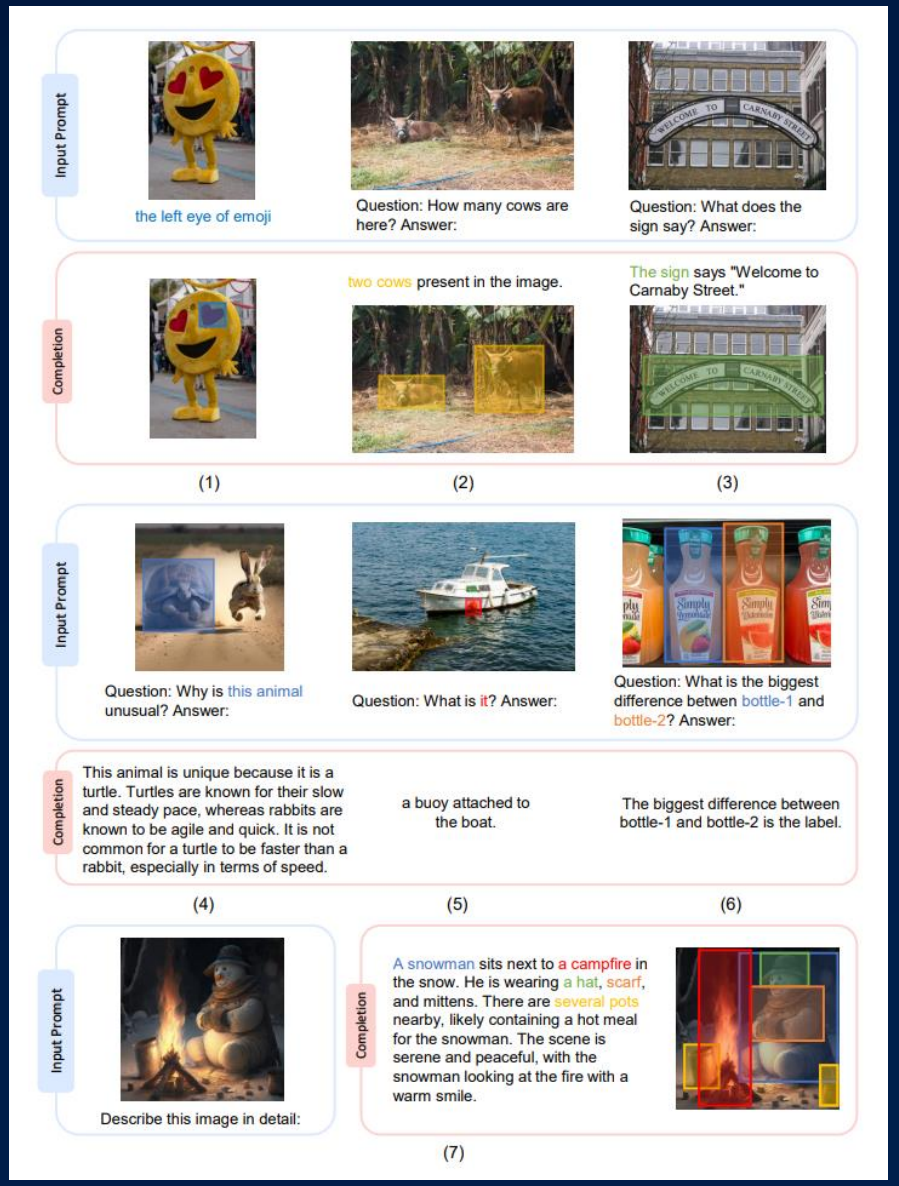
- Diffusion models, GANs for image/video generation
- Video synthesis, inpainting, motion transfer
- Used for synthetic data generation and scene simluation
- Tools for simulation, training data, and creativity



Source: Bar et al, „Navigation World Models“, CVPR 2025

Trend #3: Multi-Modal Vision-Language Models

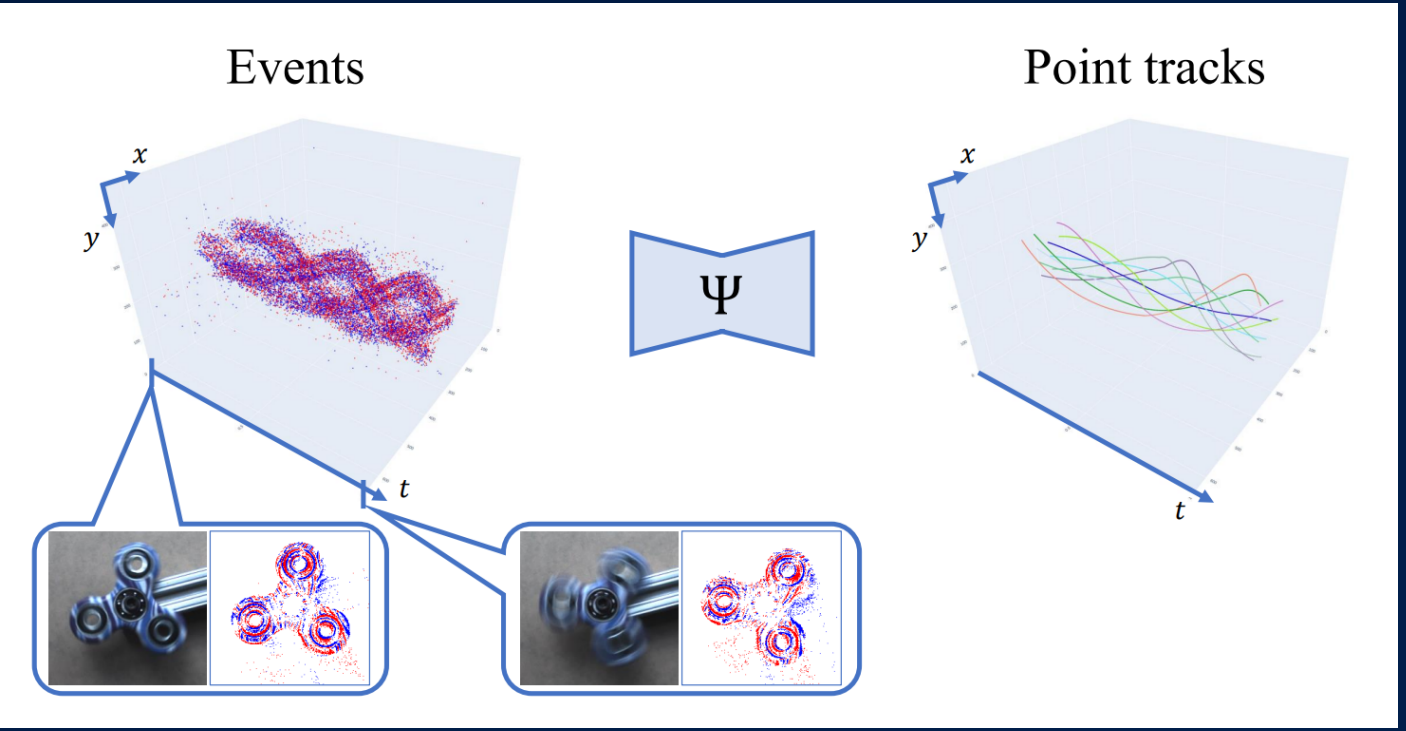
- Combining image/video with text understanding
- Examples: GPT-4V, Flamingo, Kosmos-2, Google VEO
- Applications: captioning, VQA, grounding, robotics



Source: NVISIA Kosmos 2

Trend #4: Neuromorphic and Event-Based Vision

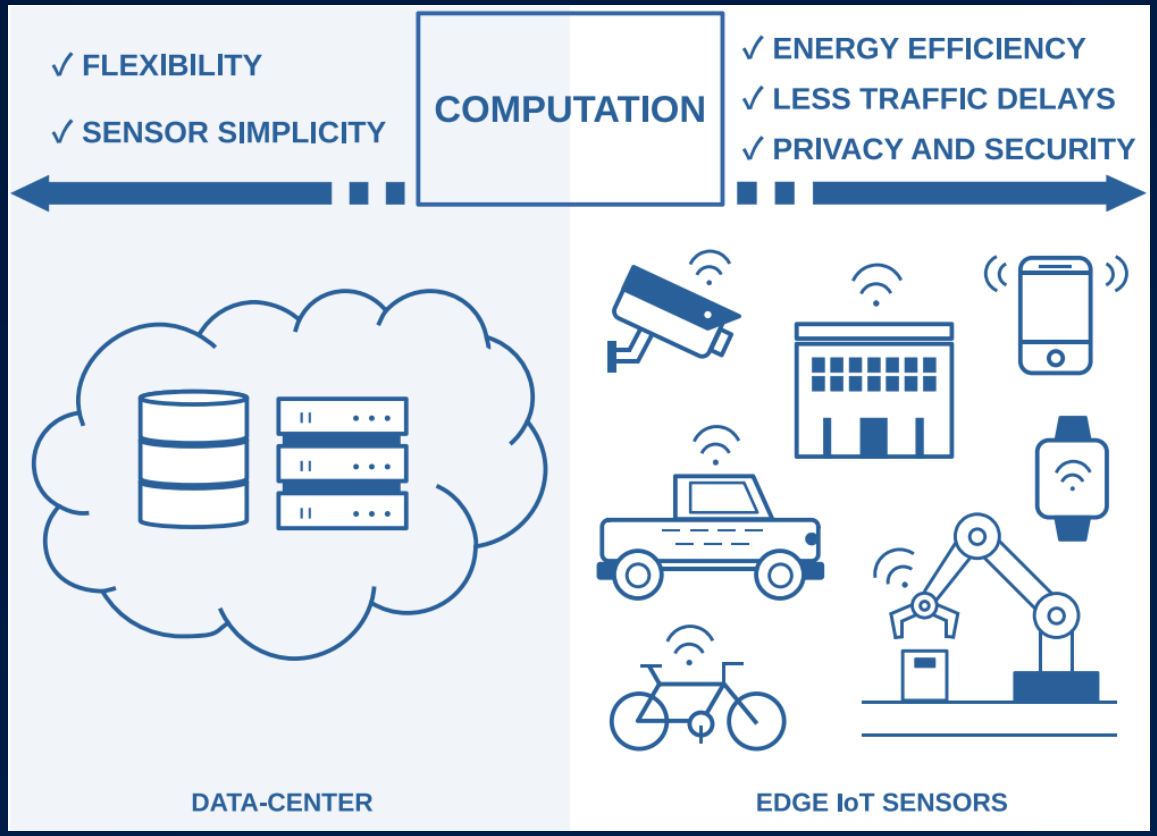
- Sensors encode only changes—low latency, low power
- Compatible with spiking networks, FPGA deployment
- Key use-cases: drones, AR, robotics



Source: Harman et al, „ETAP: Event-based Tracking of Any Point “

From Cloud to Edge: Why It Matters

- Vision models need to run on ASICs, SoCs, FPGAs
- Limited power, memory, and latency
- Trends above must adapt: pruning, quantization, co-design



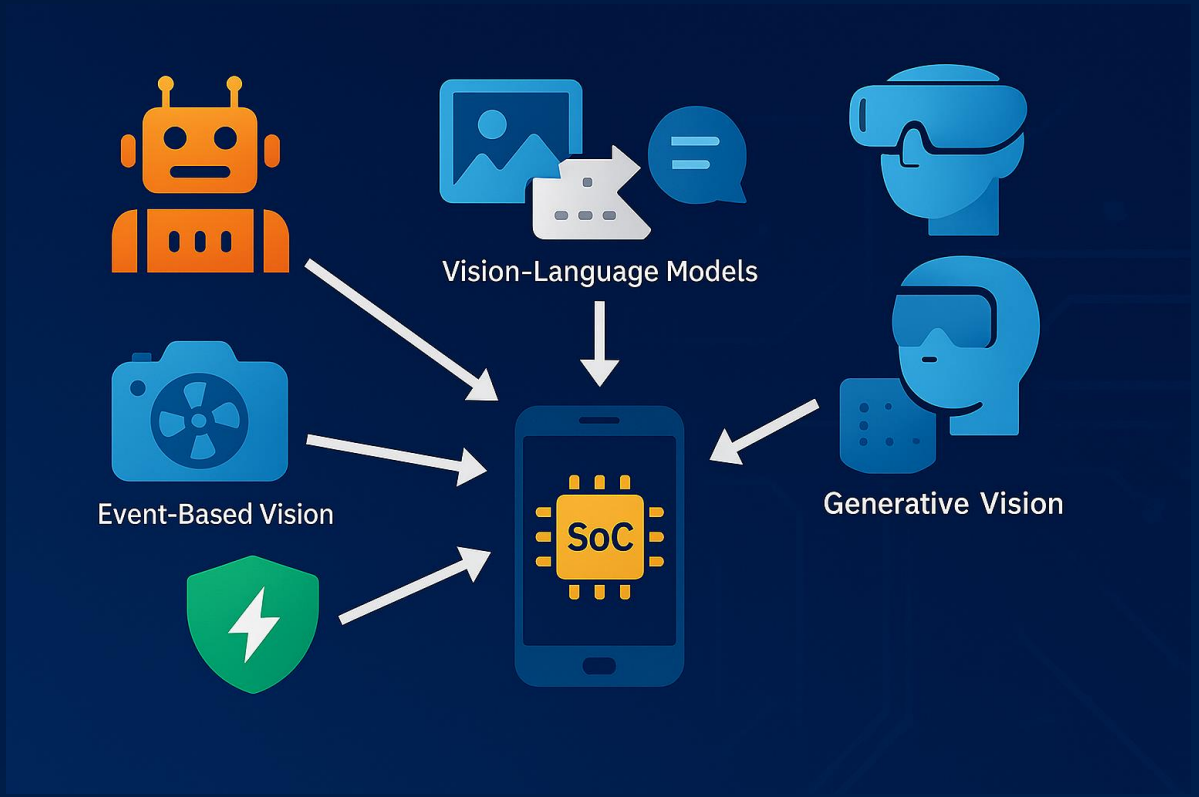
Source: Al Kountayni et al, „DeepEdgeSoC: End-to-end deep learning framework for edge IoT devices“, Internet of Things

Trend #5: Vision on Edge Devices

- Vision models on FPGAs, NPUs, and ASICs
  - FPGA resurgence for real-time, low-power CV (e.g., Xilinx Versal AI Edge, AMD)
  - Example: HOG+SVM pedestrian detection at 4K/60 fps on Zynq UltraScale+
- Smart Cameras on FPGAs:
  - Real-time SLAM on FPGA
  - Tiny YOLO on Raspberry Pi + Google Coral
  - Gesture recognition via low-power CNNs
- Neuromorphic sensors + FPGA = low-latency

Linking General CV Trends to Edge AI

- Vision-Language → mini-transformers on NPUs
- 3D scene understanding → optimized NeRFs for AR glasses
- Event-based CV → direct pipeline on neuromorphic chips
- Generative models → real-time compression or simulation at the edge



What's Next?

- Hardware-model co-design is key
- Emergence of tiny-VLMs, fast-NeRF, and vision transformers on SoCs

- **Edge-native vision is the frontier for robotics, autonomous systems, XR, and beyond**

Thank you!

Deutsches Forschungszentrum
für Künstliche Intelligenz
German Research Center for
Artificial Intelligence

Alain Pagani | DFKI