

Computer Vision

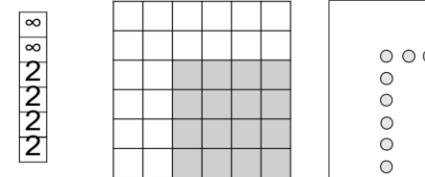


Lecture 11: Trends in Computer Vision

Oliver Bimber

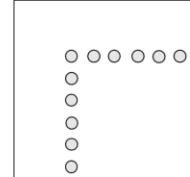
Last Week: 3D Vision

How to represent Depth

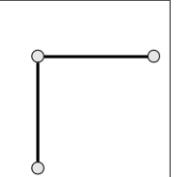


Depth Map

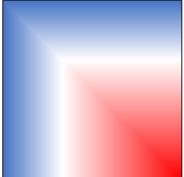
Voxel Grid



Point Cloud



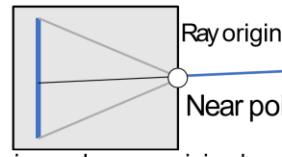
Mesh



Implicit Surface

JYU JOHANNES KEPLER
UNIVERSITY LINZ

Volume Rendering



Parameterize each ray as origin plus direction: $\mathbf{r}(t) = \mathbf{o} + t\mathbf{d}$

Volume Density is $\sigma(\mathbf{p}) \in [0,1]$

Color that a point \mathbf{p} emits in direction \mathbf{d} is $c(\mathbf{p}, \mathbf{d}) \in [0,1]^3$

JYU JOHANNES KEPLER
UNIVERSITY LINZ

Color observed by the camera given by volume rendering equation:

$$C(\mathbf{r}) = \int_{t_n}^{t_f} T(\mathbf{r}(t))\sigma(\mathbf{r}(t))c(\mathbf{r}(t), \mathbf{d})dt$$

$$T(\mathbf{r}(t)) = \exp\left(-\int_{t_n}^t \sigma(\mathbf{r}(s))ds\right)$$

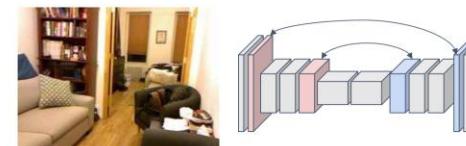
Transmittance: How much light from the current point will reach the camera?

Compute transmittance by accumulating volume density up to current point

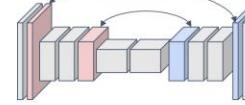
Predicting Depth Maps

Estimate log Depth instead of Depth. Defining y_i as the Ground Truth Depth on Pixel i , and y_i^* its estimated Depth:

$$D_{L2}(y, y^*) = \frac{1}{n} \sum_{i=1}^n (\log y_i - \log y_i^*)^2$$

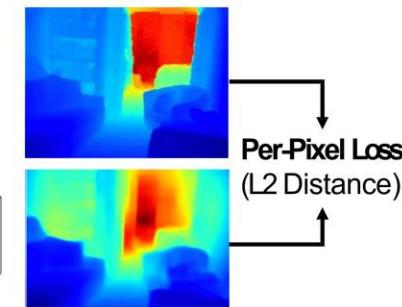


RGB Input Image:
 $3 \times H \times W$



CNN

Measured Depth Image:
 $1 \times H \times W$

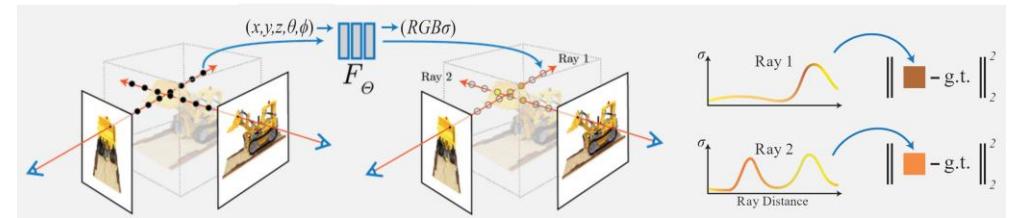


Predicted Depth Image:
 $1 \times H \times W$

JYU JOHANNES KEPLER
UNIVERSITY LINZ

Neural Radiance Fields (NeRFs)

Fully-connected Network: Input Position $\mathbf{p} = x, y, z$ and Direction $\mathbf{d} = \theta, \phi$, and output Volume Density (σ) and RGB color



Mildenhall et al, "Representing Scenes as Neural Radiance Fields for View Synthesis", ECCV/2020

JYU JOHANNES KEPLER
UNIVERSITY LINZ

Course Overview

| CW | Topic | Date | Place | Lab |
|-----|----------------------------------|------------|--|--------------|
| 41 | Introduction and Course Overview | 07.10.2025 | Zoom | Lab 1 |
| 42 | Capturing Digital Images | 14.10.2025 | Zoom | Lab 2 |
| 43 | Digital Image Processing | 21.10.2025 | Zoom | Assignment 1 |
| 44 | Machine Learning | 28.10.2025 | Zoom | |
| 45 | Feature Extraction | 04.11.2025 | Zoom | Open Lab 1 |
| 46 | Segmentation | 11.11.2025 | Zoom | Assignment 2 |
| 47 | Optical Flow | 18.11.2025 | Zoom | Open Lab 2 |
| 48 | Object Detection | 25.11.2025 | Zoom | Assignment 3 |
| 49 | Multi-View Geometry | 02.12.2025 | Zoom | Open Lab 3 |
| 50 | 3D Vision | 10.12.2025 | Zoom | Assignment 4 |
| → 3 | Trends in Computer Vision | 13.01.2026 | Zoom | |
| 4 | Q&A | 20.01.2026 | Zoom | Open Lab 4 |
| 5 | Exam | 27.01.2026 | HS1 (Linz), S1/S3 (Vienna), S5 (Bregenz) | |
| 9 | Retry Exam | 24.02.2026 | tba | |

Not covered: Generative Models for Image Synthesis

- Variational Autoencoders (VAEs)
- Generative Adversarial Networks (GANs)
- Diffusion Models
- etc.
- What is this good for in CV?
 - Training Data Augmentation
 - Solving Inverse Problems
 - etc.



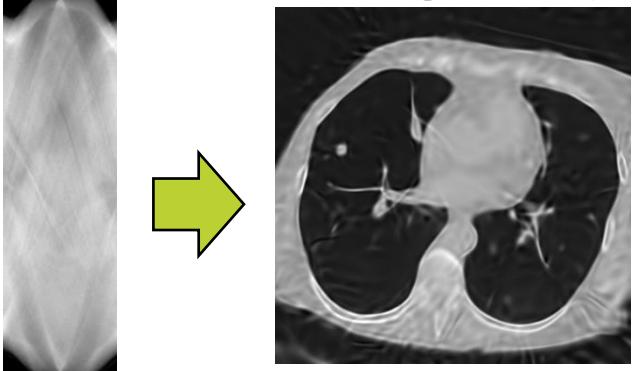
a shiba inu wearing a beret and black turtleneck



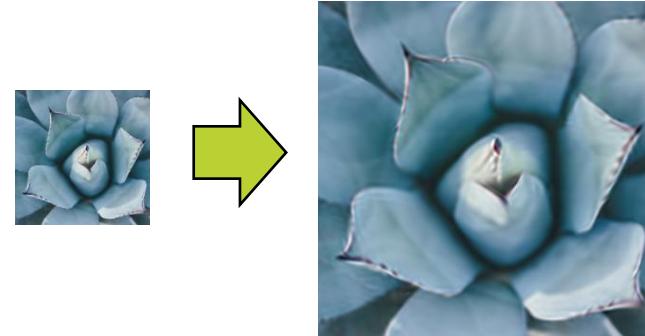
a close up of a handpalm with leaves growing from it

Example: Inverse Problems

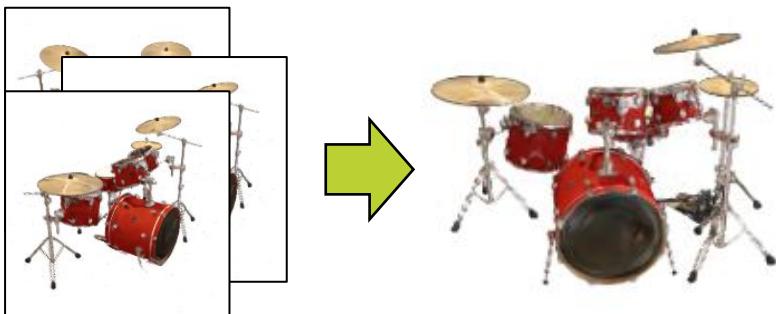
Inverse Tomography



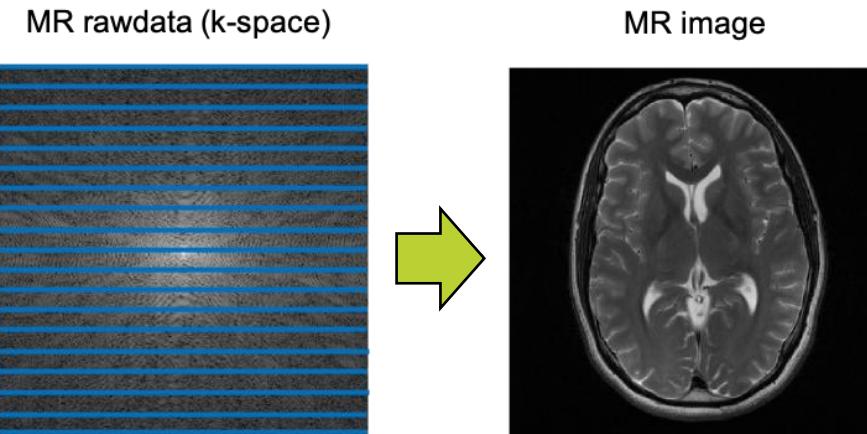
Super Resolution



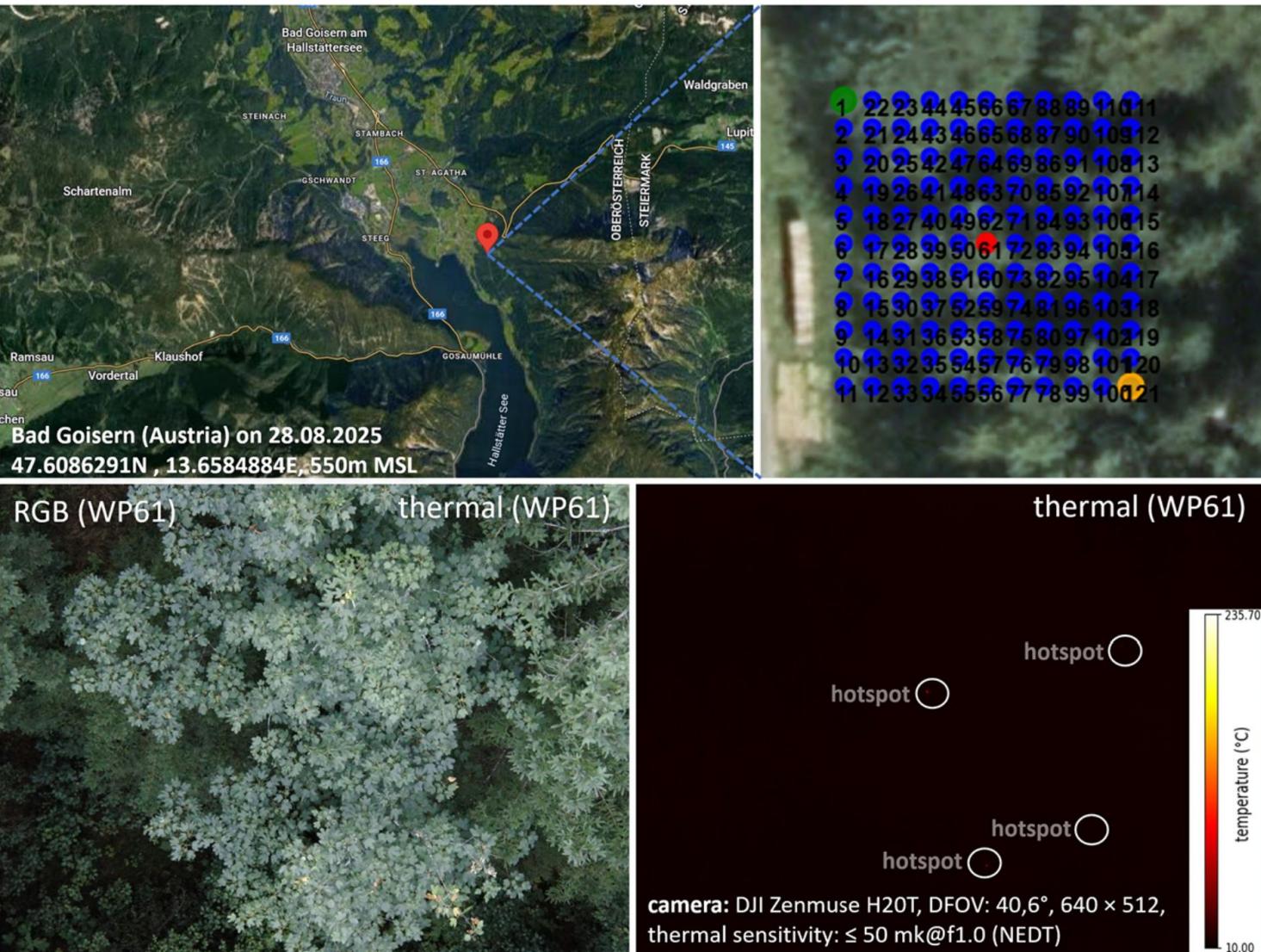
Novel View Synthesis



Sparse MRI



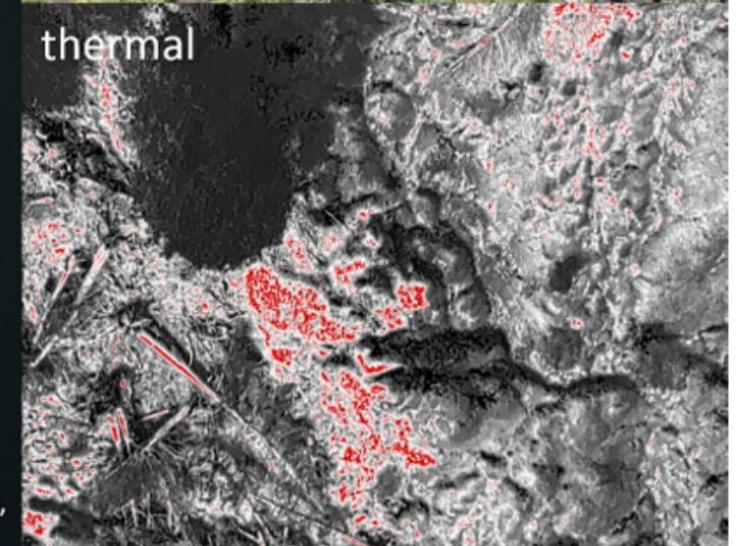
Example from our Research



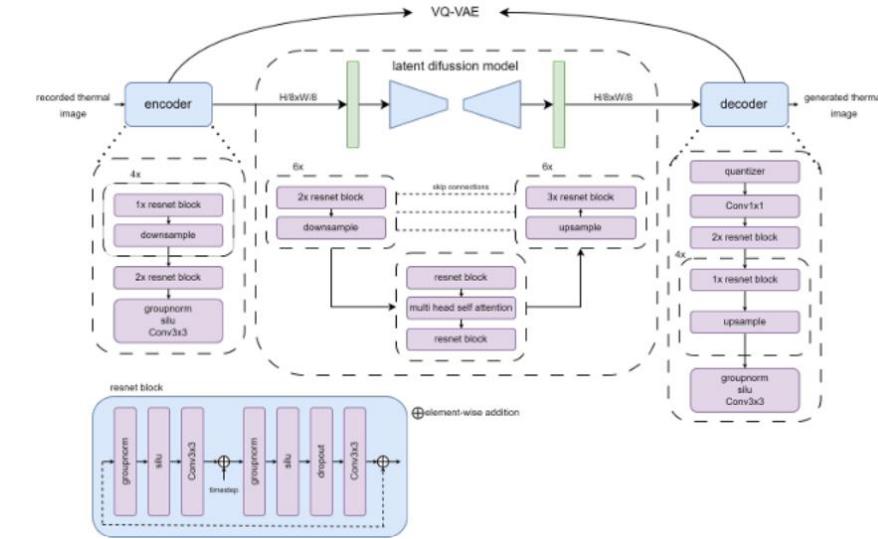
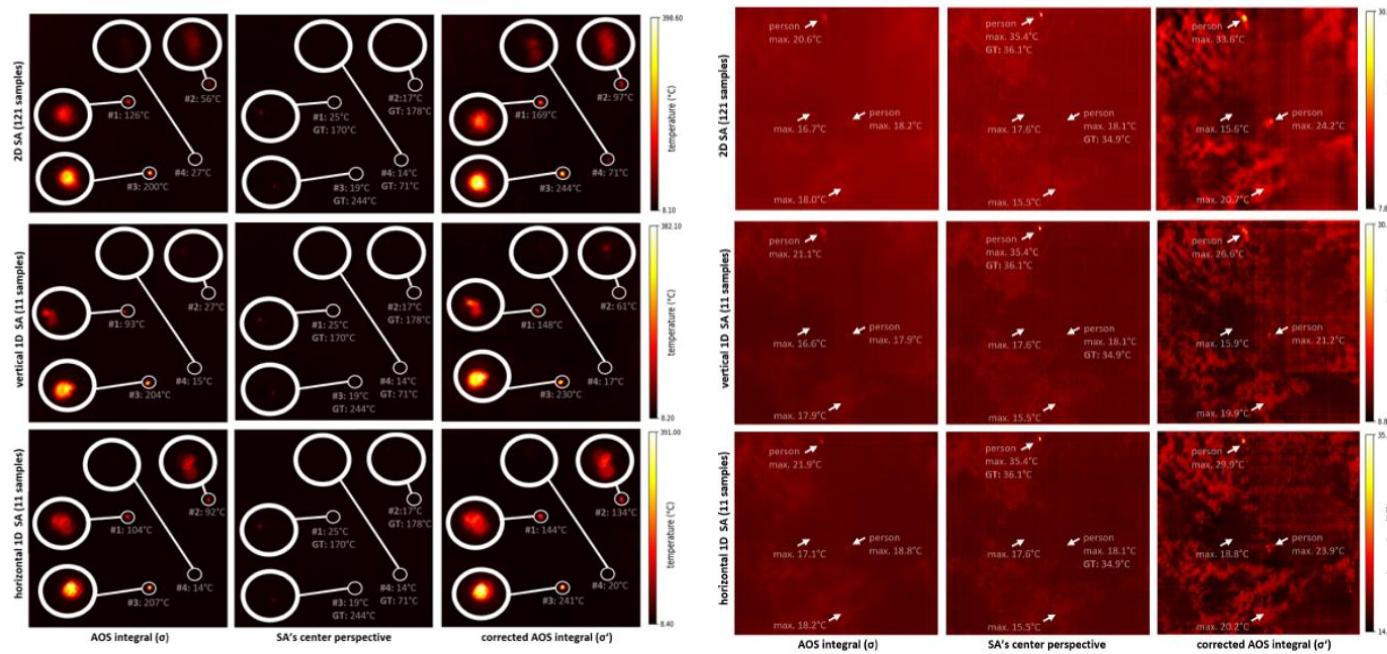
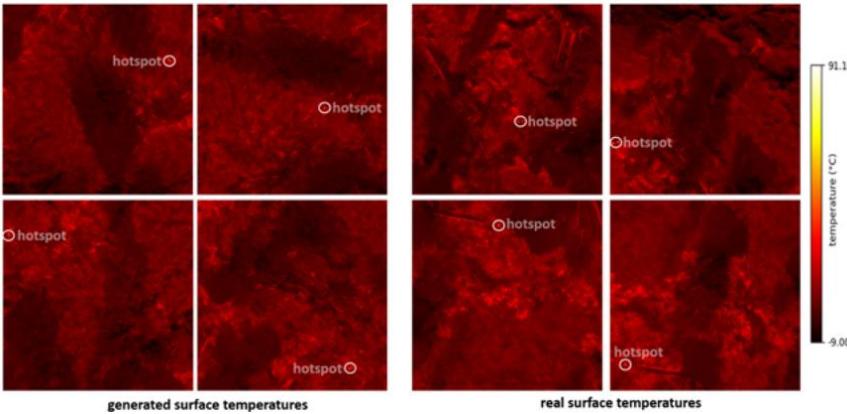
Example from our Research

Wildfire at Grafenberglalm (Austria) on 17.10.2023

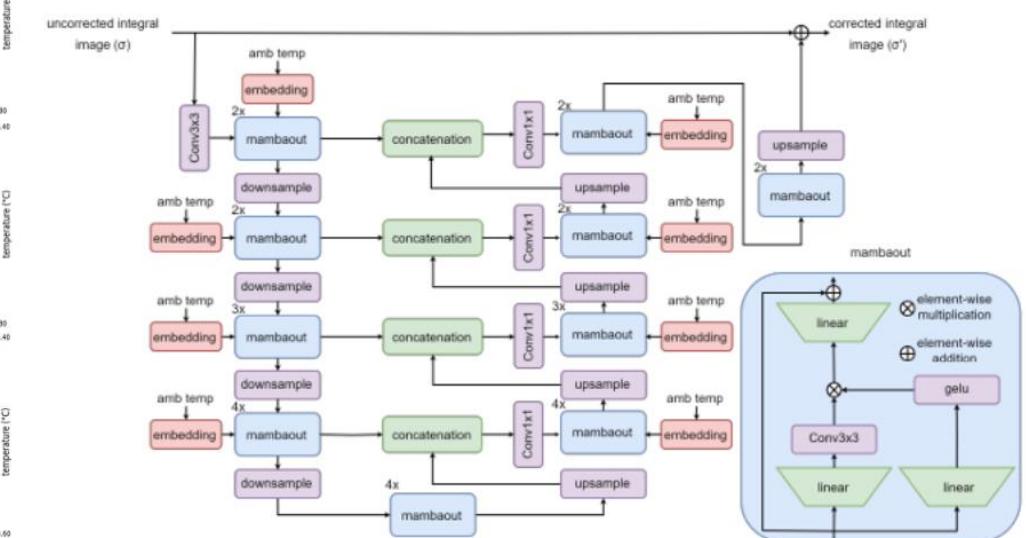
47.46754N , 13.75525E, 1780m MSL



Example from our Research



vector quantized variational autoencoder

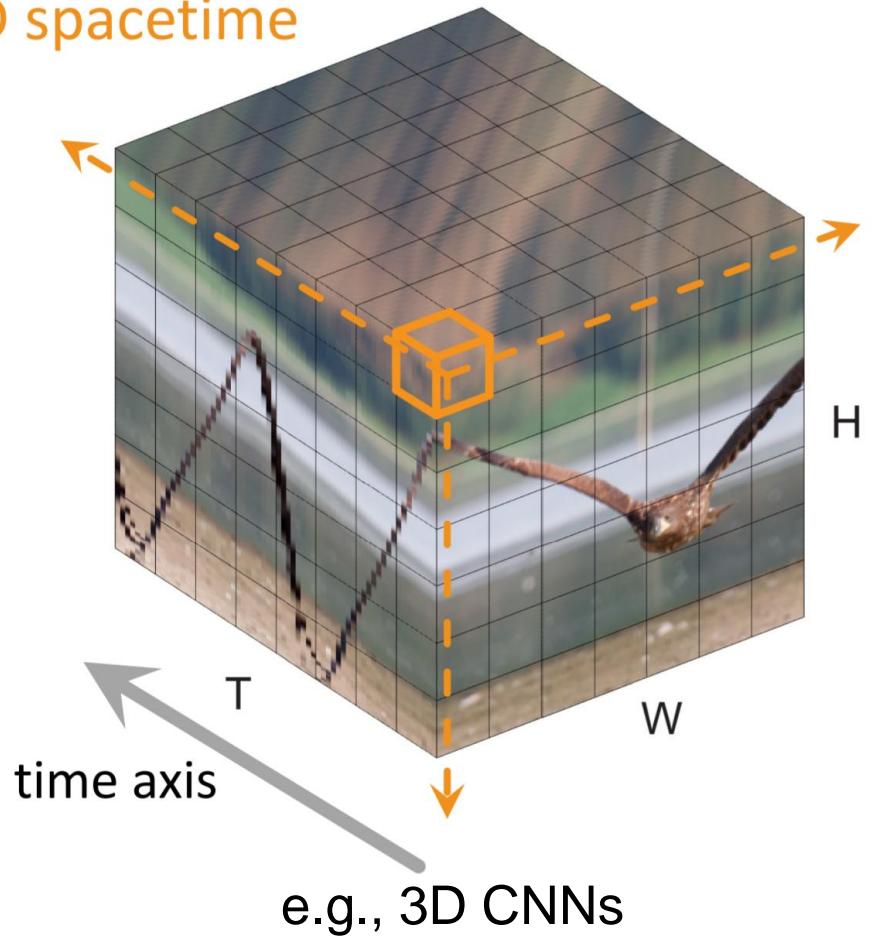


visual state space model (VmambaIR)

Not covered: Sequence Problems

- Sequences in Computer Vision
= Videos
- ConvNets for Sequence Problems
- Recurrent Neural Networks
- Transformers
- etc.

sliding a 3D window
in 3D spacetime



Not covered: Self-Supervised Vision

Common Trick:

- Convert “unsupervised” Problem into “supervised” Empirical Risk Minimization
- Do so by cooking up “Labels” (prediction Targets) from the raw Data itself

Supervised vision

Hand-curated training data
+ Informative
- Expensive
- Limited to teacher's knowledge



Vision in nature

Raw unlabeled training data
+ Cheap
- Noisy
- Harder to interpret



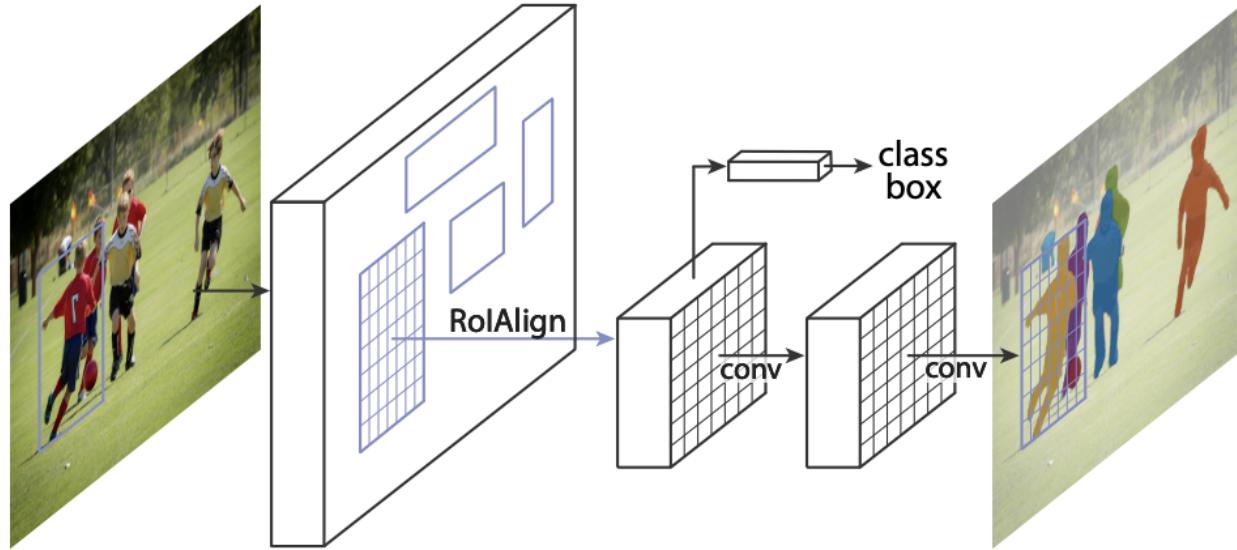
Limits of Machine Learning

“The central challenge in machine learning is that our algorithm must perform well on new, previously unseen inputs—not just those on which our model was trained. The ability to perform well on previously unobserved inputs is called **generalization**.

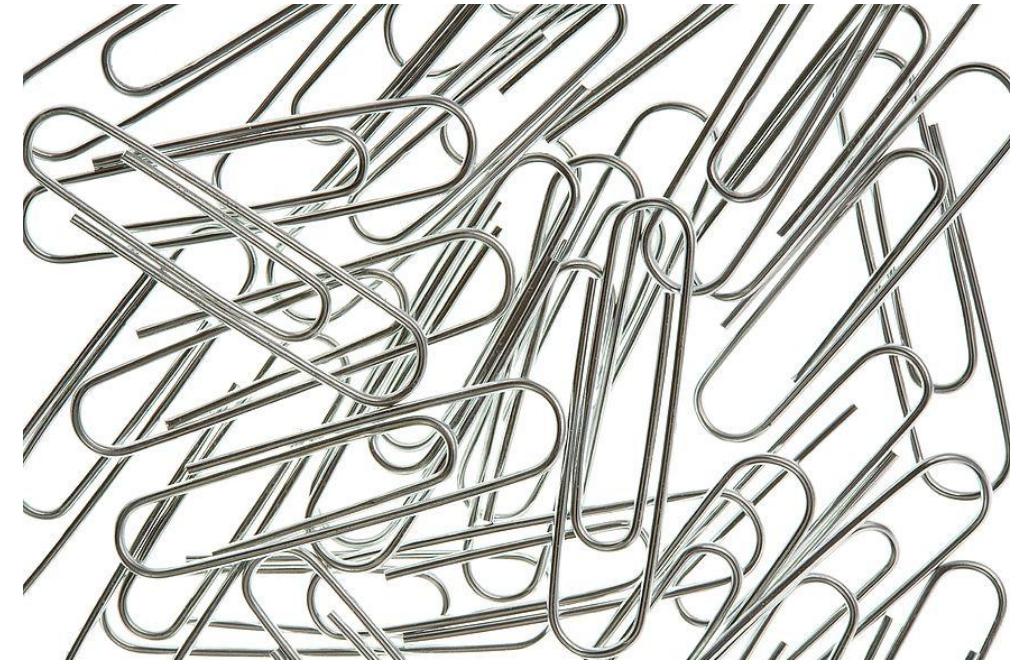
... [this is what] separates machine learning from optimization.”

— Deep Learning Textbook (Goodfellow et al.)

Example: Object Detection

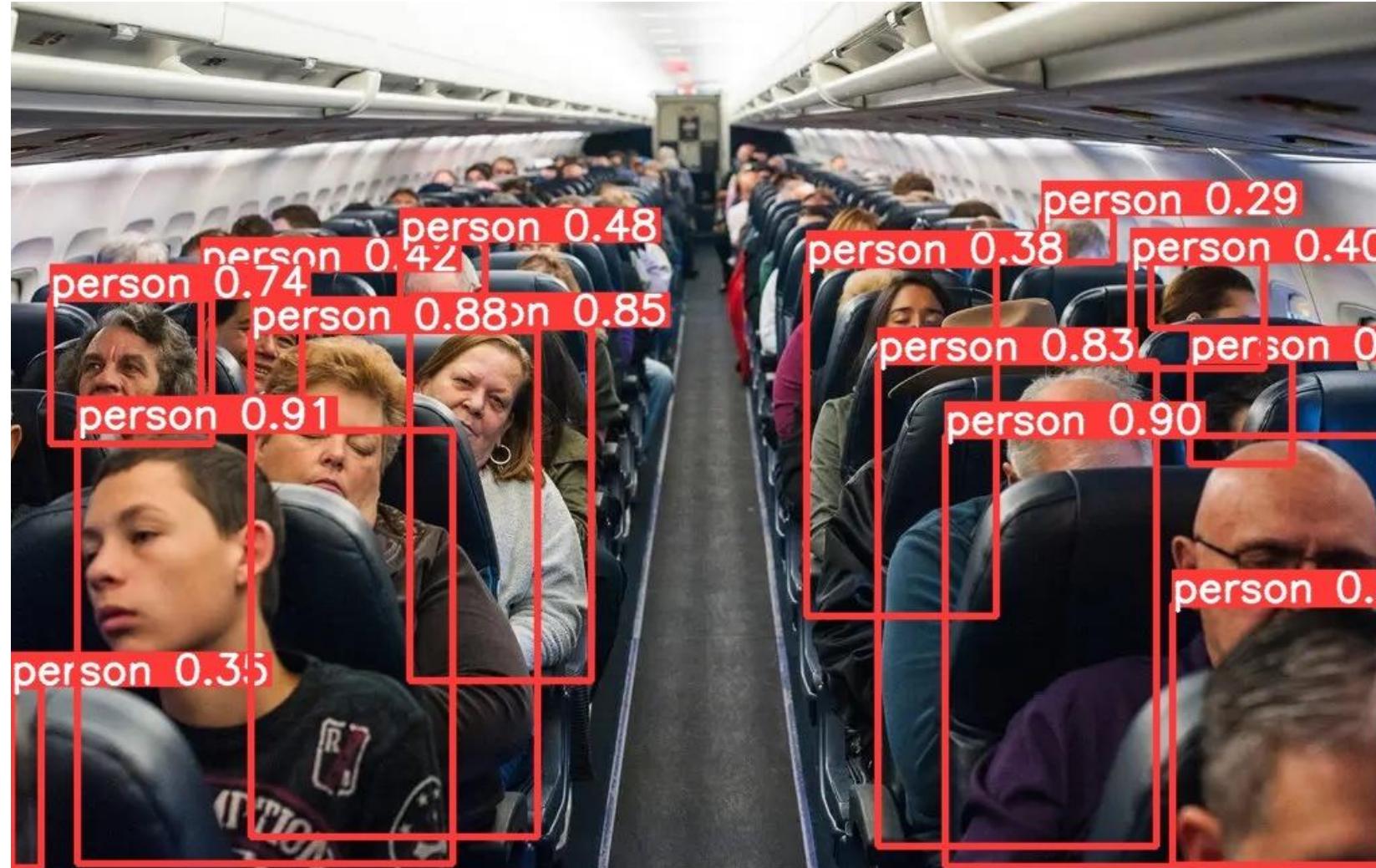


CNN can detect Humans, Cars, etc.



But it can't count these Binder Clips because it's never seen them before! It has not learned what an 'object' is in general.

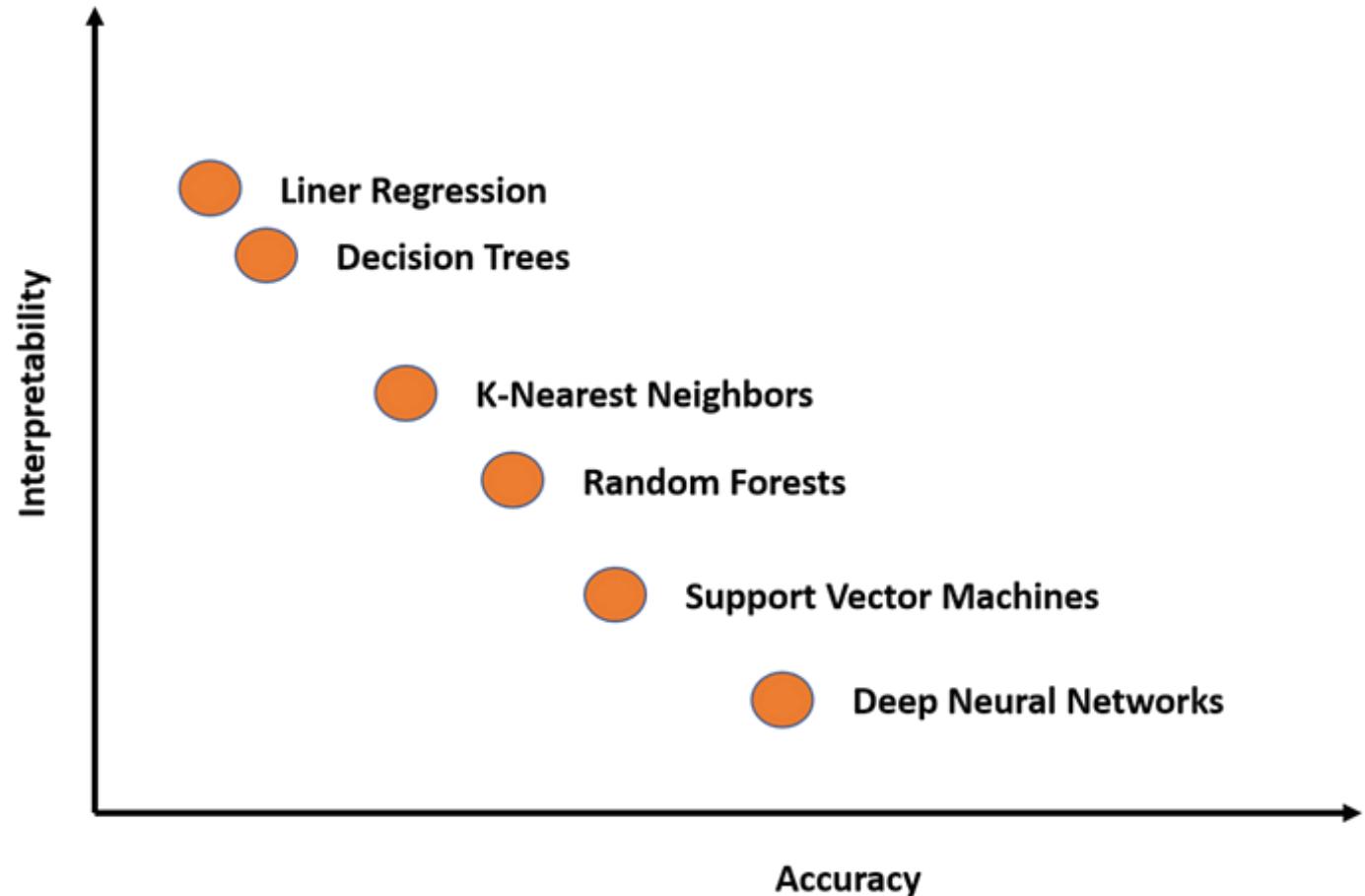
Example: Person Detection



Yolo-v8 can't detect the people in the back.
But we can!

Problem: Interpretability vs. Accuracy

- What Features of an Image is the Model using?
- How does the Model make its Decisions?
- Accuracy-based Training (how about a Bias towards a Majority in the Training Data)?



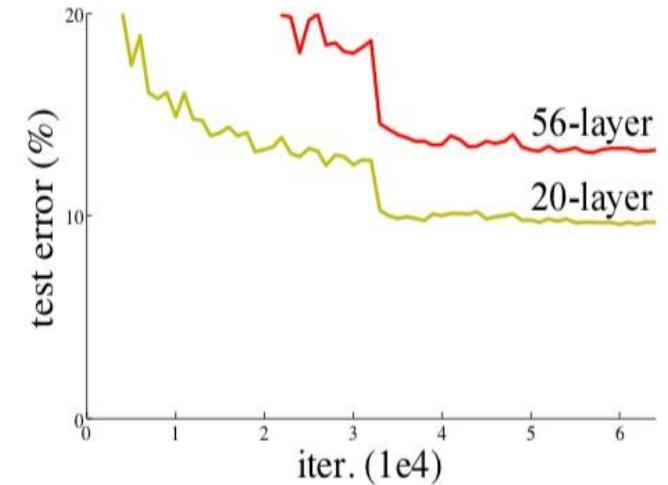
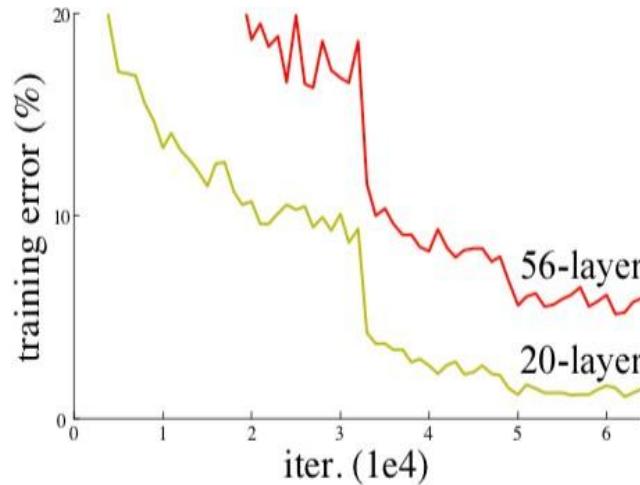
Lessons Learned

Data is not the only Answer

"If you torture the Data long enough,
it will confess to anything"

- *How to Lie With Statistics*
by Darrell Huff

More Parameters != better



Human Knowledge can help (e.g. as done with Regularization in Loss-Functions)

Large Multi-Model Models (LMMs)

- A large multi-modal model **handles multiple data types** (text, images, audio, video) in one system.
- It **learns connections across modalities**, not each one in isolation.
- Examples:
 - GPT (OpenAI): text, image, audio, and video understanding and generation
 - Gemini (Google DeepMind): strong multimodal reasoning with very long context
 - Claude (Anthropic): advanced text + image understanding
 - etc.



Example: stereopair of a dog.

Object Detection with LMMs



what is in the image?

This image features a very cute, small dog, likely a puppy or a toy breed, standing on a stone path.

The Dog

- **Breed Appearance:** It looks like a small terrier or poodle mix (often called a "designer breed" like a Morkie, Maltipoo, or Shichon).
- **Coat:** The fur is white with light tan or cream patches, particularly around the ears and eyes. The texture appears soft and slightly shaggy.
- **Features:** It has dark, expressive eyes, a black nose, and a small tail poking up in the back.

The Setting

- **Background:** The dog is standing in front of a **rustic stone retaining wall** and a set of **stone steps** to the left.
- **Ground:** It is standing on a light-colored concrete or stone walkway with some fallen leaves scattered nearby, suggesting an outdoor garden or patio setting.

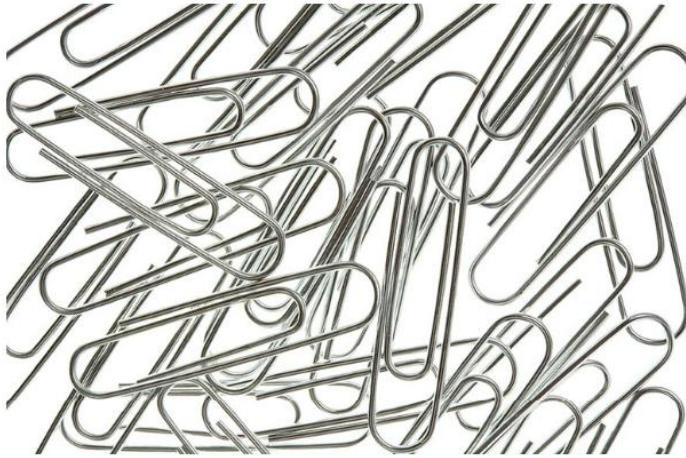
Would you like me to help you identify the specific breed or suggest some names for this little pup?

Segmentation with LMMs



generate an image with the dog segmented

More difficult Example with LMMs



Count how many binder clips are in the image.

There are **zero** binder clips in this image.

The items shown here are actually **paperclips** (specifically the "Gem" style). While they both hold paper together, binder clips are the ones with the black triangular bodies and silver folding arms.

In this pile, there are approximately **30 paperclips**.

Would you like me to help you identify the difference between other types of office supplies, or perhaps help you find where to buy actual binder clips?

Depth from Single Images with LMMs



Here is an image. Generate an image that visualizes reconstructed depth.

Depth from Stereo with LMMs



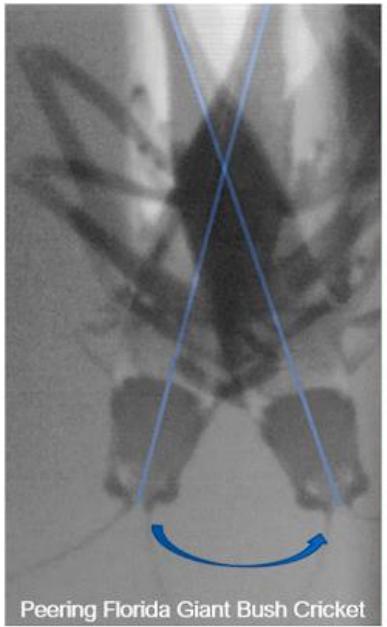
Here is a stereopair. Generate an image that visualizes reconstructed depth.

Novel View Generation with LMMs



generate an image that shows the dog from the side perspective.

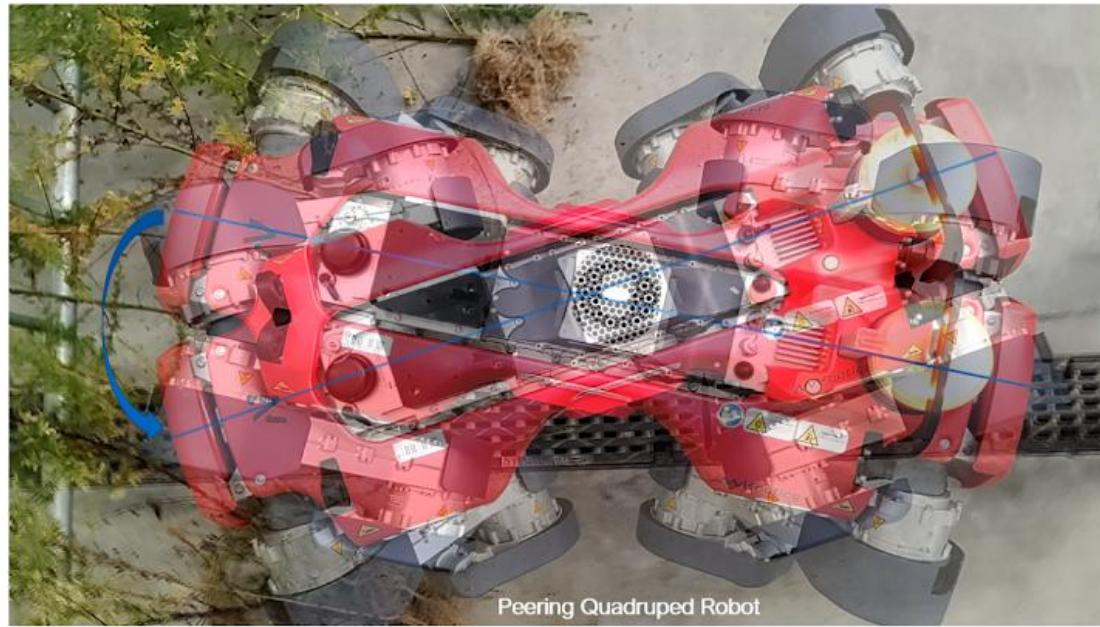
Example from our Research



Peering Florida Giant Bush Cricket



Peering Locust



Peering Quadruped Robot



Camera Equipped Head-Unit



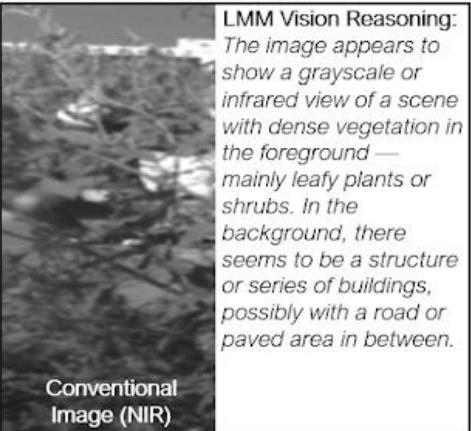
Conventional Image (RGB)

LMM Vision Reasoning:
The image shows some green plants with thin stems and fine leaves growing in front of a metallic fence or railing. Behind the plants, there is an industrial or commercial-looking structure and a concrete floor. The setting appears to be an industrial facility or a technical building exterior.



SA Integral Image (RGB)

LMM Vision Reasoning:
The image appears to be quite blurry, but some features can still be made out. It seems to show an industrial or abandoned facility, possibly a warehouse or a factory. In the center, there's a large structure with circular openings or pipes — maybe ventilation ducts, tubes, or barrels stacked in a frame. The surroundings include a metal framework.



Conventional Image (NIR)

LMM Vision Reasoning:
The image appears to show a grayscale or infrared view of a scene with dense vegetation in the foreground — mainly leafy plants or shrubs. In the background, there seems to be a structure or series of buildings, possibly with a road or paved area in between.



SA Integral Image (NIR)

LMM Vision Reasoning:
The image appears to be a blurry grayscale or low-resolution photo showing an outdoor scene. It looks like there is some vegetation or bushes in the background, and a person or human-like figure is sitting or crouching on the ground near the center of the image. The surroundings seem natural — possibly a garden, park, or wooded area.

Example from our Research



LMM Vision Reasoning:

Based on the sequence of images provided, the horizontal camera movement creates a parallax effect that helps reveal several structures behind the dense foreground vegetation. Through the gaps in the leaves, the following background elements can be identified:

- Metal Staircase/Walkway: On the far left, there is a prominent industrial-style metal staircase or gangway with open treads and railings.
- Industrial Equipment: In the center background, behind the main stems of the plants, there is a piece of blue industrial machinery or a cabinet. It appears to be part of a utility system.
- Horizontal Infrastructure: A large metal pipe or structural railing runs horizontally across the top portion of the frame, spanning the width of the area.
- Building Facade: The primary backdrop is a light-colored (white or light grey) building wall.
- Vertical Support Structures: Several vertical metal poles or supports are visible, which seem to hold up the horizontal railing or form part of a fence/cage structure.
- Ventilation/Grating: To the right of the blue equipment, there are vertical segments that look like metal gratings or ventilation panels.

Example from our Research



LMM Vision Reasoning:

Based on the focal stack provided, as the focus shifts from the foreground vegetation (most prominent in layer 1) to the background, several distinct industrial or architectural features become clear.

Behind the occluding greenery, there is a large, multi-story structure that appears to be an industrial facility or a modern architectural complex. Here is what can be identified:

- Large Circular Apertures: Centered in the mid-background is a prominent array of seven large circular openings arranged in a hexagonal pattern (six around one). These appear to be industrial fans, ventilation ducts, or large-scale pipe outlets integrated into the wall.
- External Staircases and Walkways: On the left side of the frame, there is a clear set of metal industrial stairs and elevated platforms with railings.
- Scaffolding or Support Pillars: Vertical, lattice-like structures (resembling scaffolding or slender support columns) are visible flanking the central circular array.
- Wall Texture: The background wall has a light-colored, possibly concrete or tiled texture, with horizontal lines that become more defined in the later layers of the stack (e.g., layer 7 through layer 10).

Example from our Research



LMM Image Generation (Input: Single Conventional Image):
Generate an image without occlusion in the foreground.

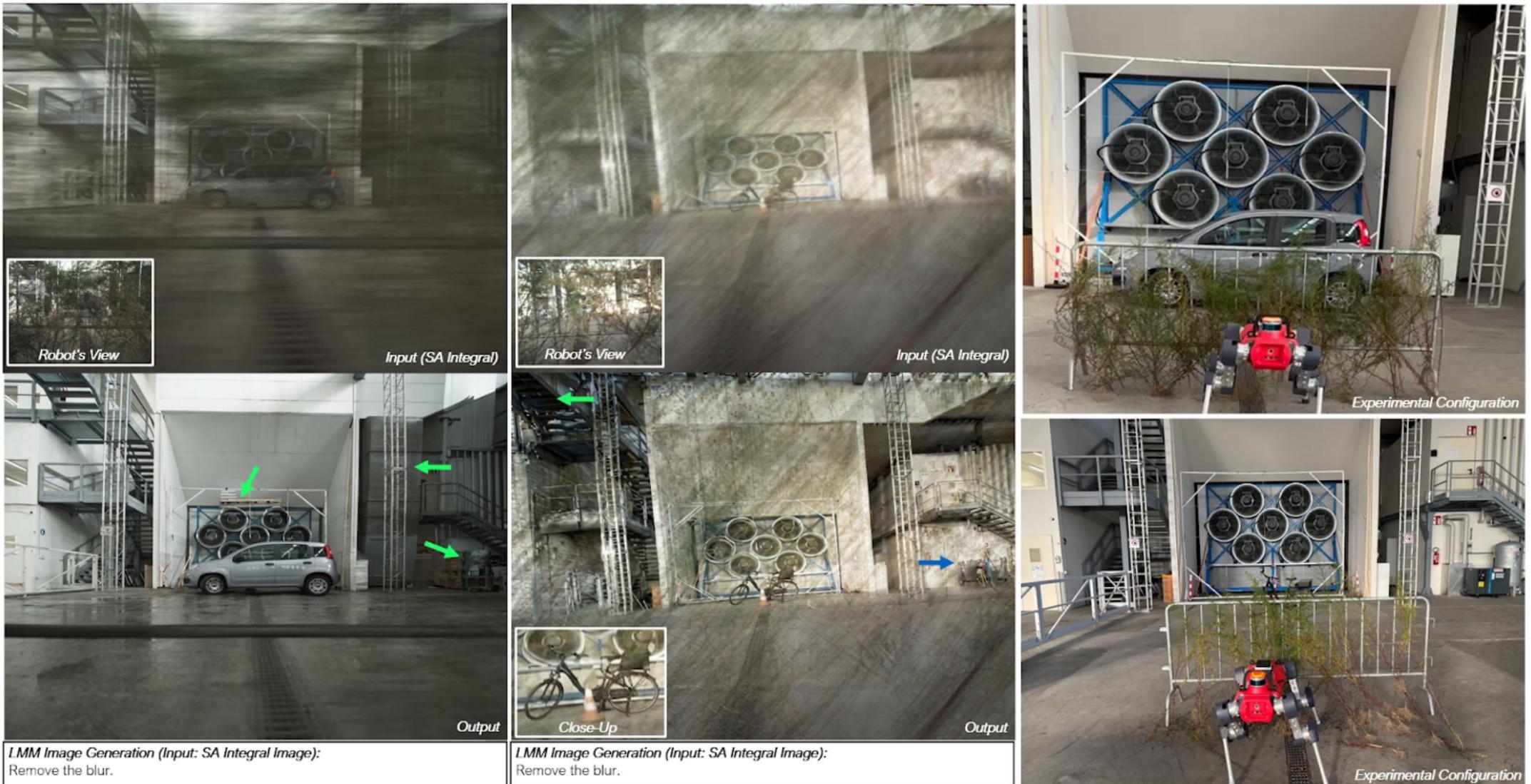


LMM Image Generation (Input: Series of Conventional Images):
Generate an image without occlusion in the foreground.



LMM Image Generation (Input: SA Integral Image):
Remove the blur.

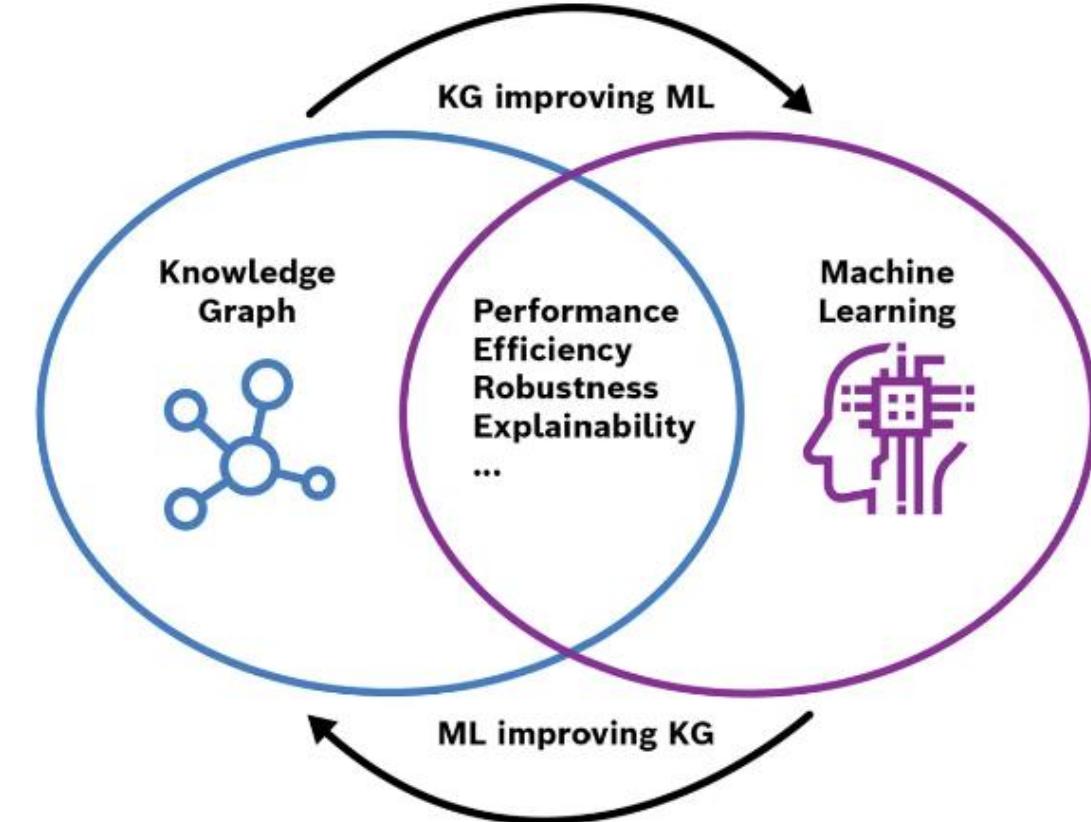
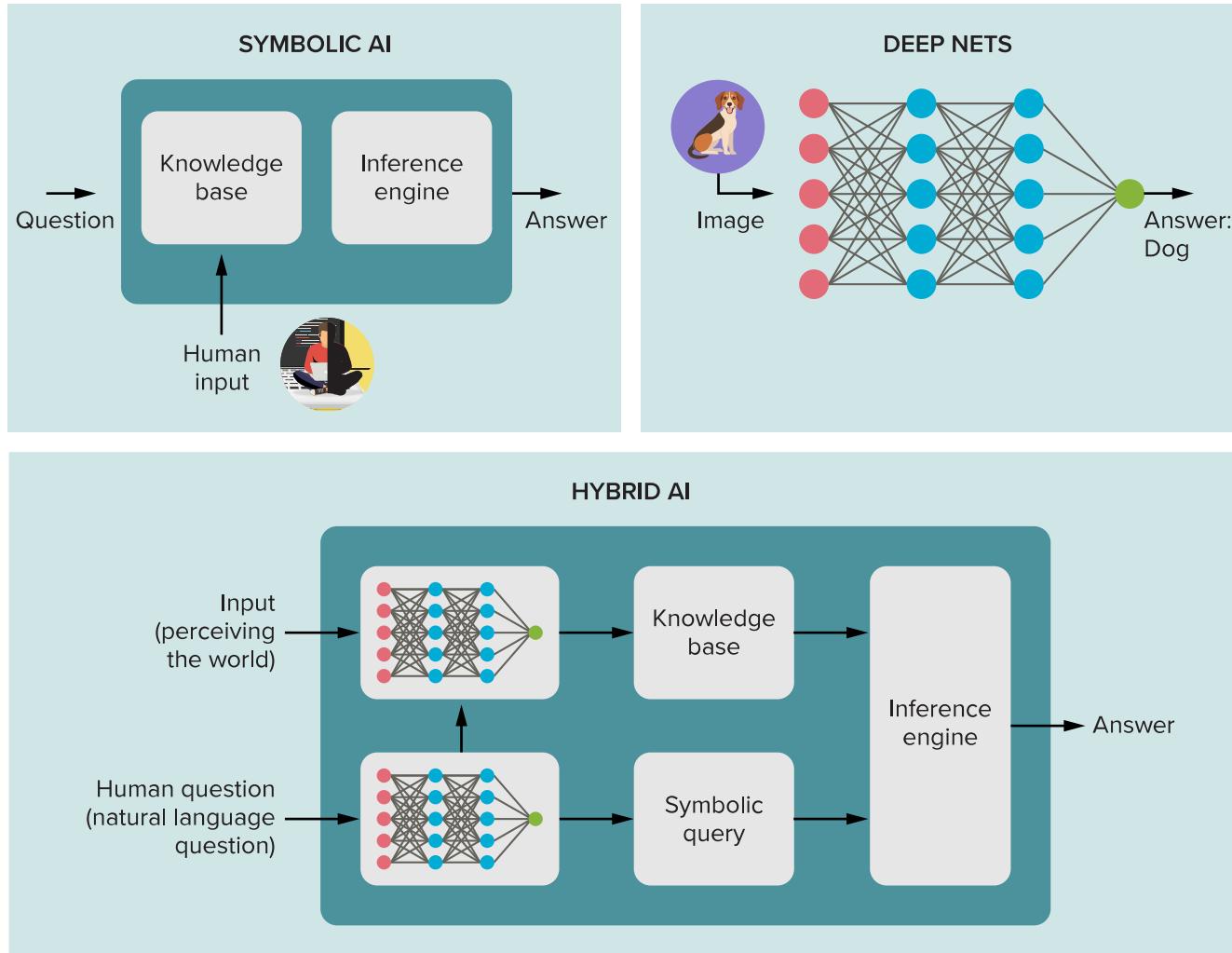
Example from our Research



LLM Limitations

- Todays LMMs are **huge** (e.g., **transformer**) **models**, trained on a **huge amount of (general) data**.
- They can be tweaked with **in-context learning** (adding background information to huge context windows).
- They represent a **general solution to computer vision** (ignoring performance, cost, energy consumption, etc. for now).
- Nevertheless, they are not always right! **They hallucinate** if they can't generalize.
- So, do they **really reason**? Probably not! They only compile and apply the statistical essence of data – but they don't understand.
- And they are still **mainly based on natural language**.
- They don't know much of the **rules** on how our world works.

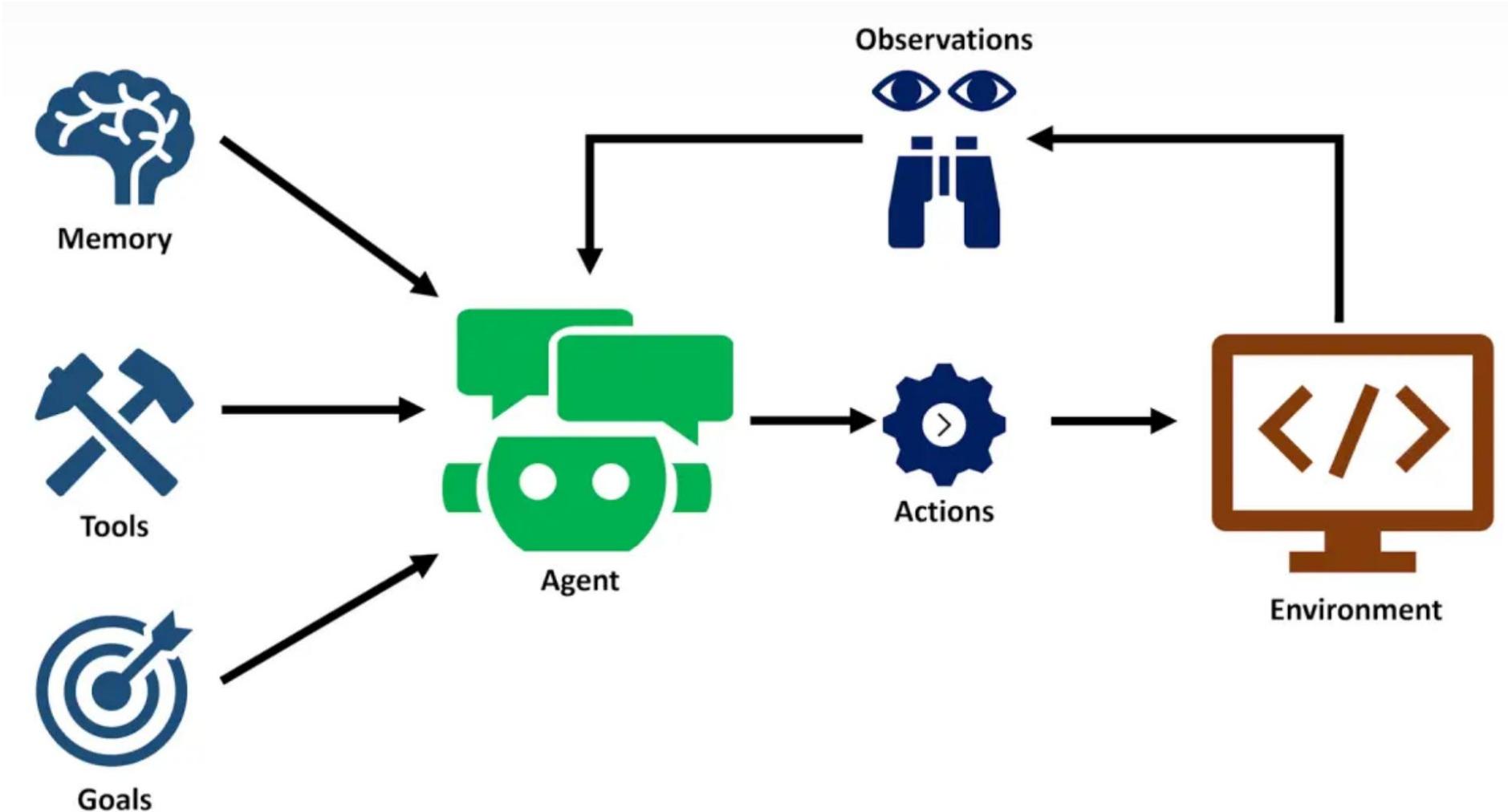
Neurosymbolic AI



SOURCE: REPORTING BY A. ANANTHASWAMY

KNOWABLE MAGAZINE

Agentic AI



Example from our Research

(possibly complex) tasks
(natural language text)

there are two drones, take off both to an altitude of 50m, fly them in a circle with 5 waypoints, then fly drone 1 to Berlin and drone 2 to Rome

Simple Control Loop
In Python

that provides initial task prompts
(e.g. to solve our sub-problems)

In-Context Learning
(reasoning possible because of huge training data (whole internet) and large context windows (whole book) in state of the art LLMs)

GoodAI/charlie-mnemonic

Charlie Mnemonic: The First Personal Assistant with Long-Term Memory

Long-Term Memory

ChatGPT-4o

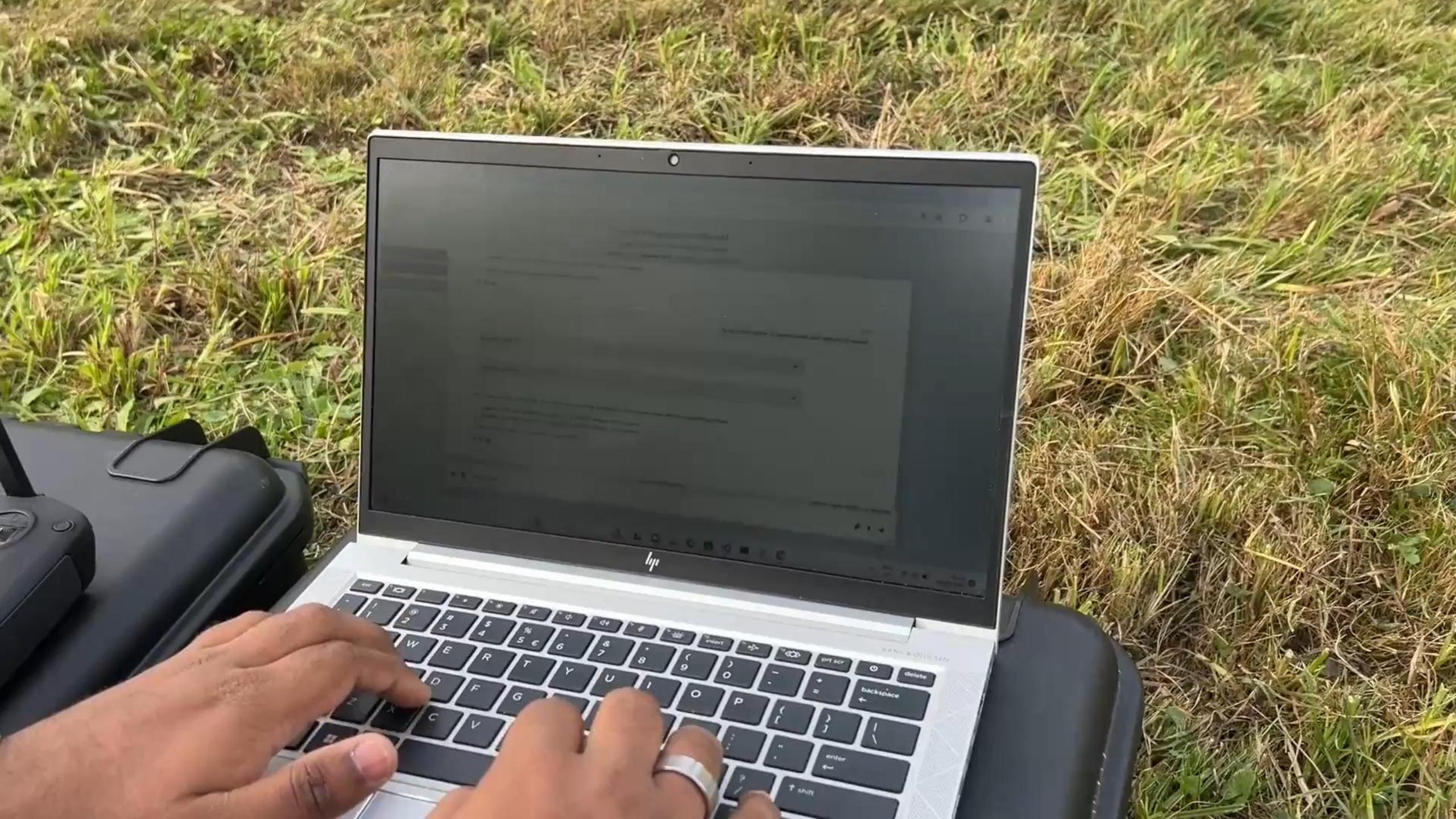
(later possibly o1 with even better reasoning)

API description (natural language text)

```
description = "Control drones using: move(distance, heading, altitude, speed) e.g., move(5, 120, 25, 3). move_by(distance, heading_offset, altitude_offset, speed). move_to(lat, lon, altitude, speed). rotate_drone_to(heading) e.g., rotate_drone_to(120). rotate_drone_by(heading_offset). rotate_gimbal_pitch_by(pitch_offset). rotate_gimbal_yaw_by(yaw_offset). rotate_gimbal_roll_by(roll_offset). Use get_position() for latitude, longitude, altitude. get_altitude() returns altitude. get_heading() returns heading. get_image(camera_mode) captures image e.g., get_image('rgb'). All distances and altitudes in meters, speeds in m/s, offsets (±°), headings (0-359°)."
```

lists of drone instructions
for each drone with properly
chosen parameters
(based on our API
description) → these are
directly parsed and
executed by our CS for
flying real drones OR by
our simulator to simulate
drone flights

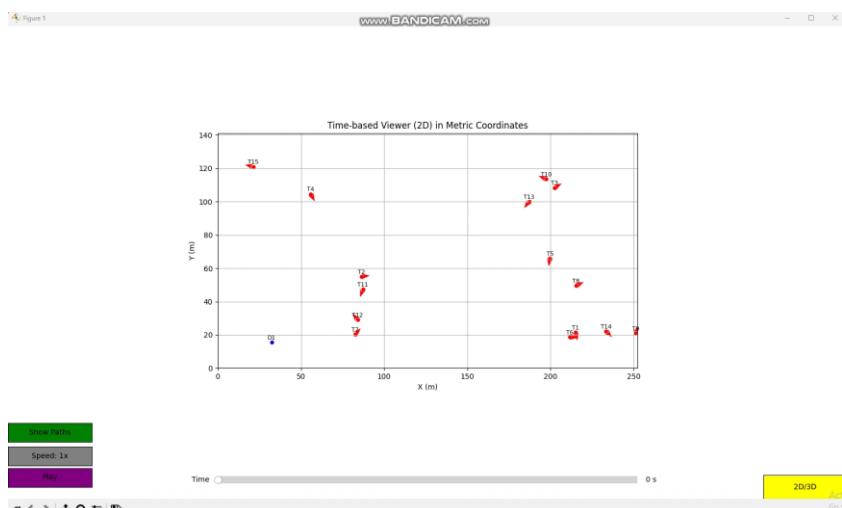
```
{
  "drones": [
    1,
    2
  ],
  "instructions": [
    {
      "drone_id": 1,
      "instructions": [
        {
          "instruction": "move_to",
          "parameters": "0,0,50,5"
        },
        {
          "instruction": "move_to",
          "parameters": "10,0,50,5"
        },
        {
          "instruction": "move_to",
          "parameters": "10,10,50,5"
        },
        {
          "instruction": "move_to",
          "parameters": "0,10,50,5"
        },
        {
          "instruction": "move_to",
          "parameters": "0,0,50,5"
        },
        {
          "instruction": "move_to",
          "parameters": "52.5200,13.4050,50,10"
        }
      ]
    },
    {
      "drone_id": 2,
      "instructions": [
        {
          "instruction": "move_to",
          "parameters": "0,0,50,5"
        },
        {
          "instruction": "move_to",
          "parameters": "10,0,50,5"
        },
        {
          "instruction": "move_to",
          "parameters": "10,10,50,5"
        },
        {
          "instruction": "move_to",
          "parameters": "0,10,50,5"
        },
        {
          "instruction": "move_to",
          "parameters": "0,0,50,5"
        },
        {
          "instruction": "move_to",
          "parameters": "41.9028,12.4964,50,10"
        }
      ]
    }
  ],
  "username": "admin"
}
```



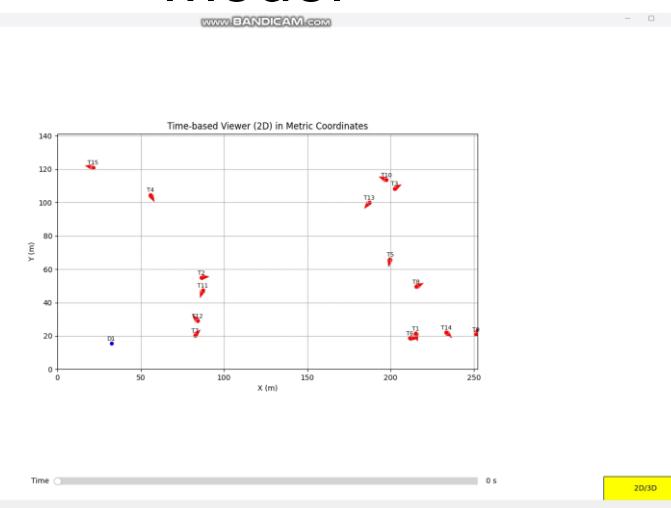
Example from our Research

1 drone, 15 targets

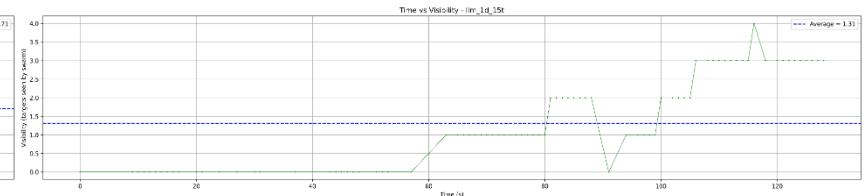
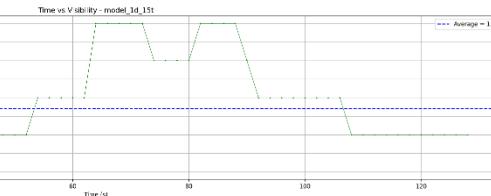
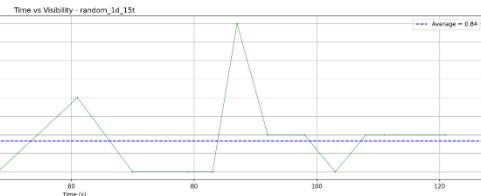
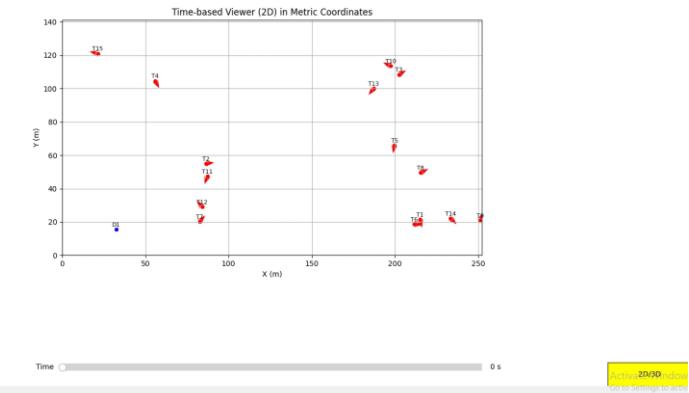
random



model



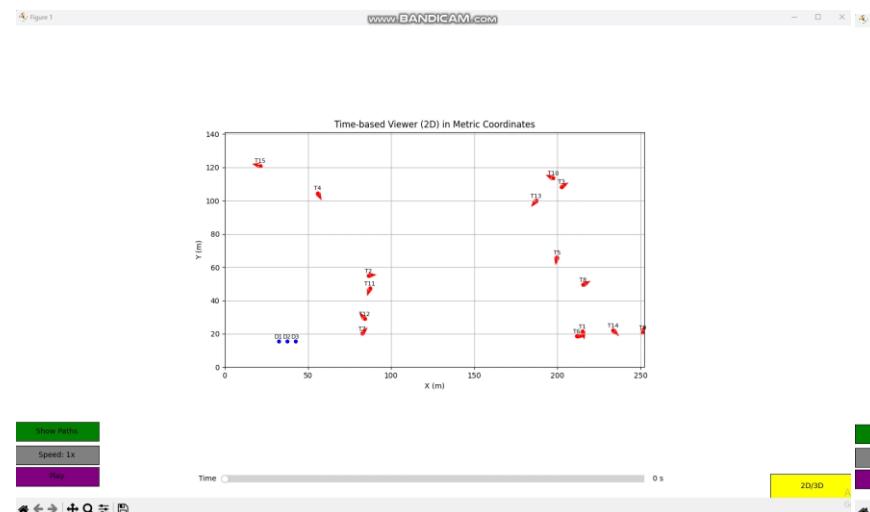
LLM



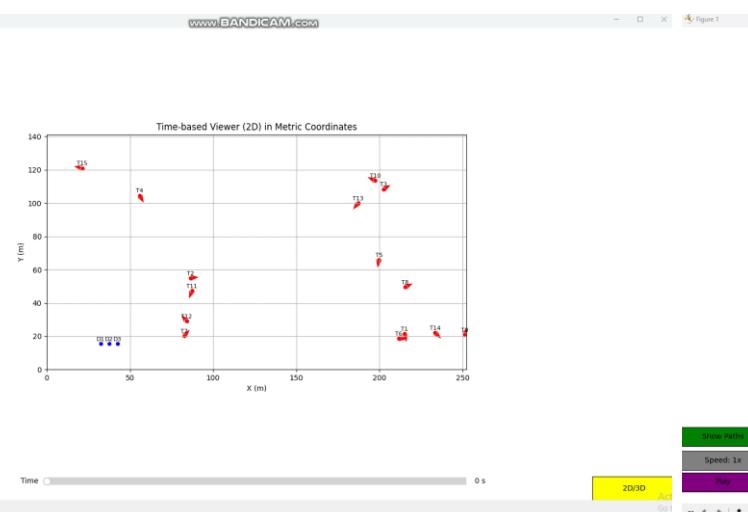
Example from our Research

3 drones, 15 targets

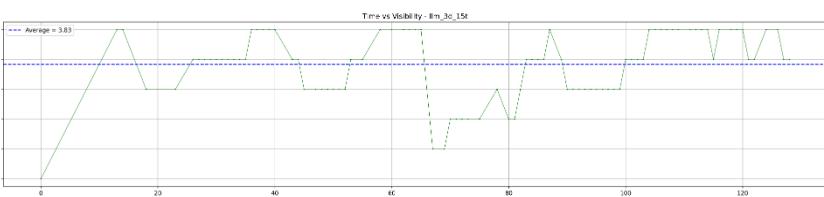
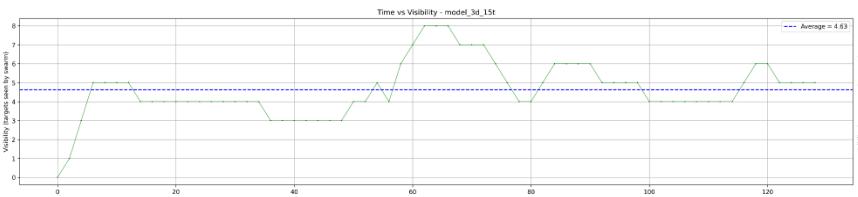
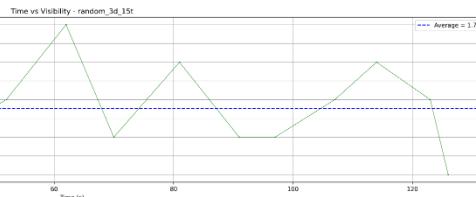
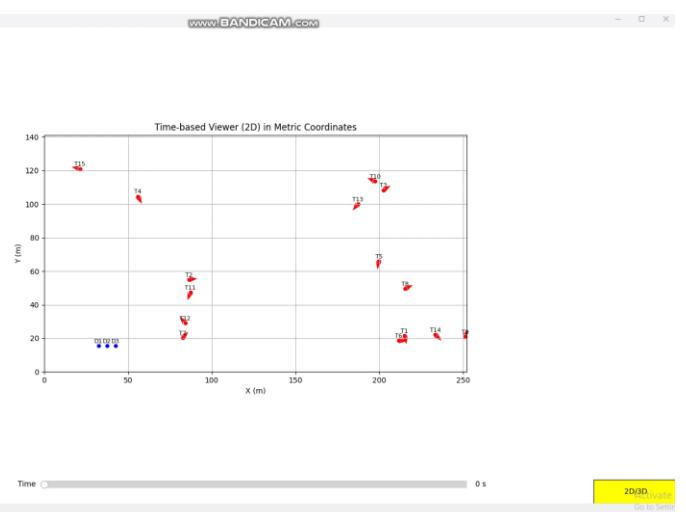
random



model



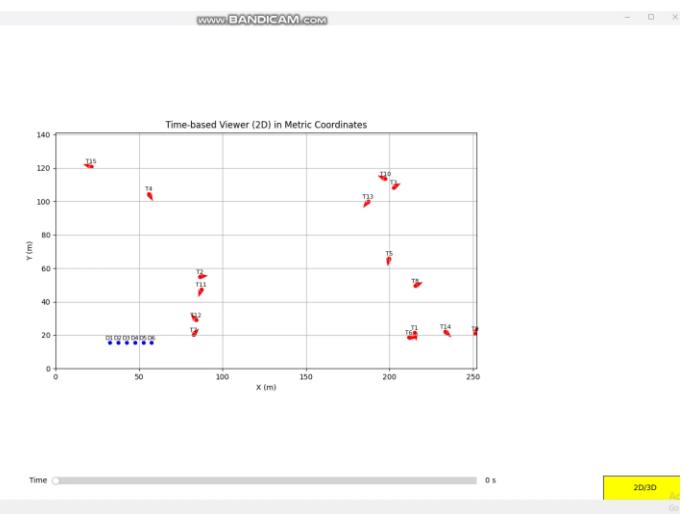
LLM



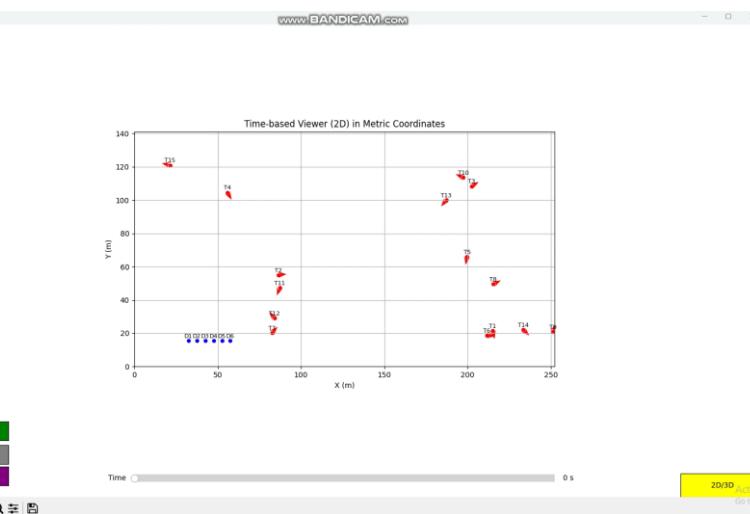
Example from our Research

6 drones, 15 targets

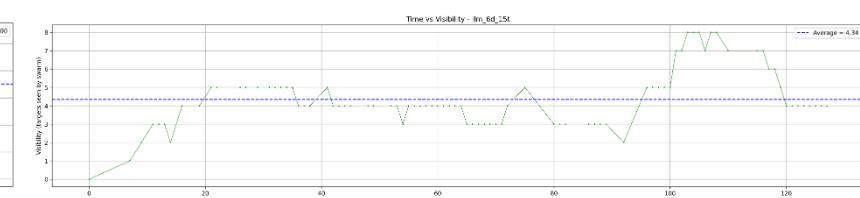
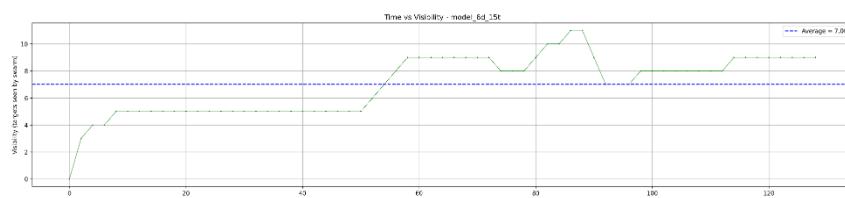
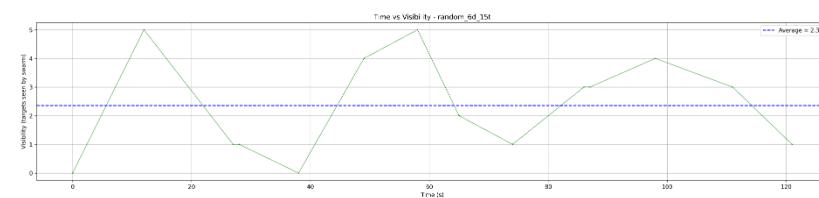
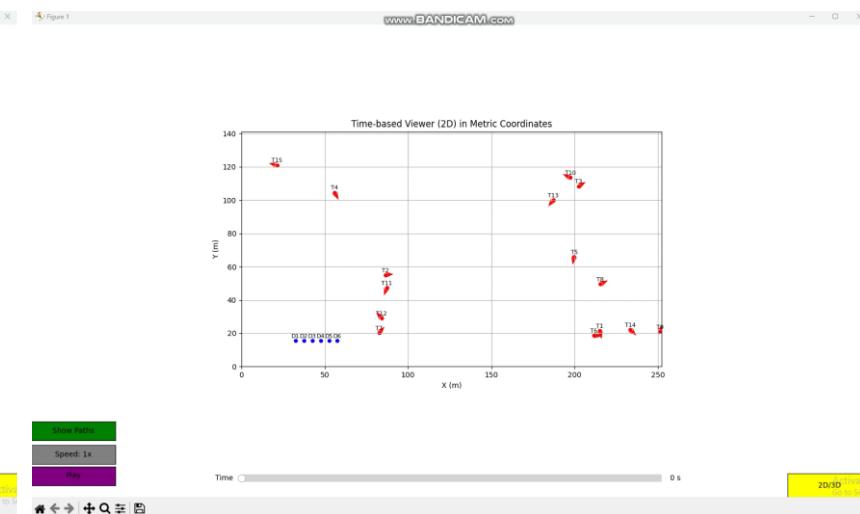
random



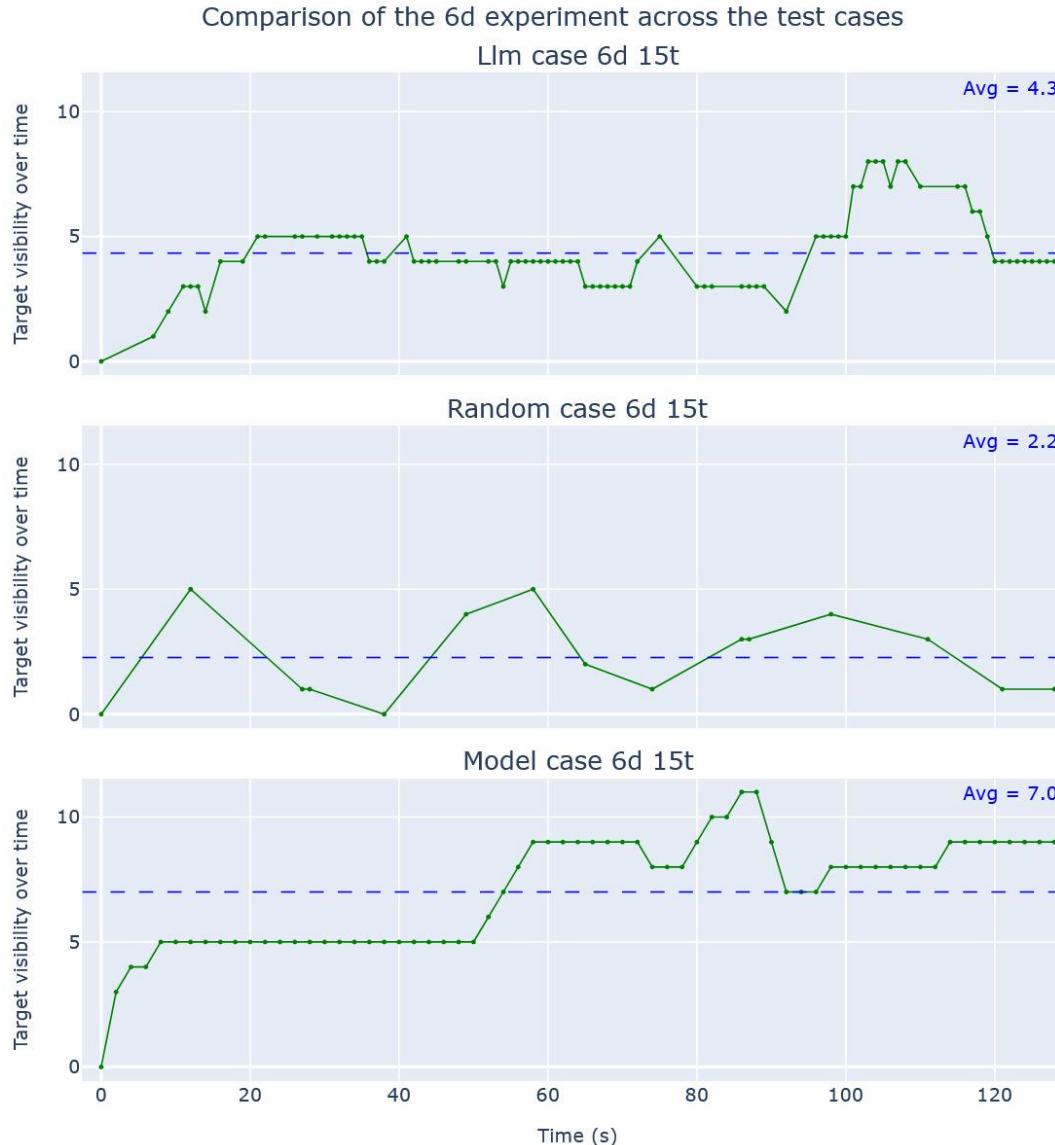
model



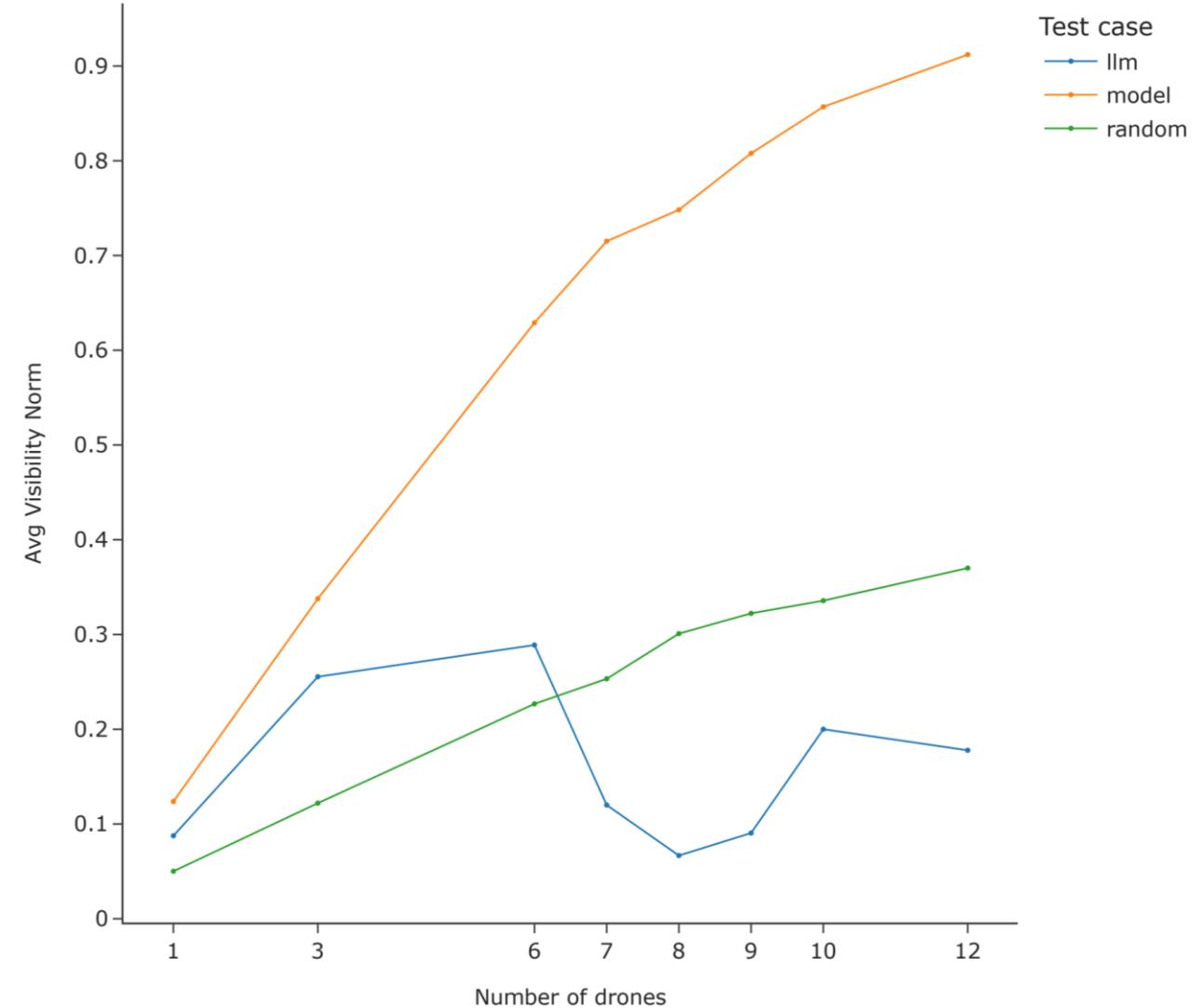
LLM



Example from our Research



Normalized average visibility across all experiments



Conventional Imaging

Has not changed since 350BC (first mentioning of Camera Obscura = Pinhole Camera, by Aristotle)

Became analog in 1826 (first permanent Photograph taken in France)

Became digital in 1975 (Kodak invented the first digital Camera Sensor)

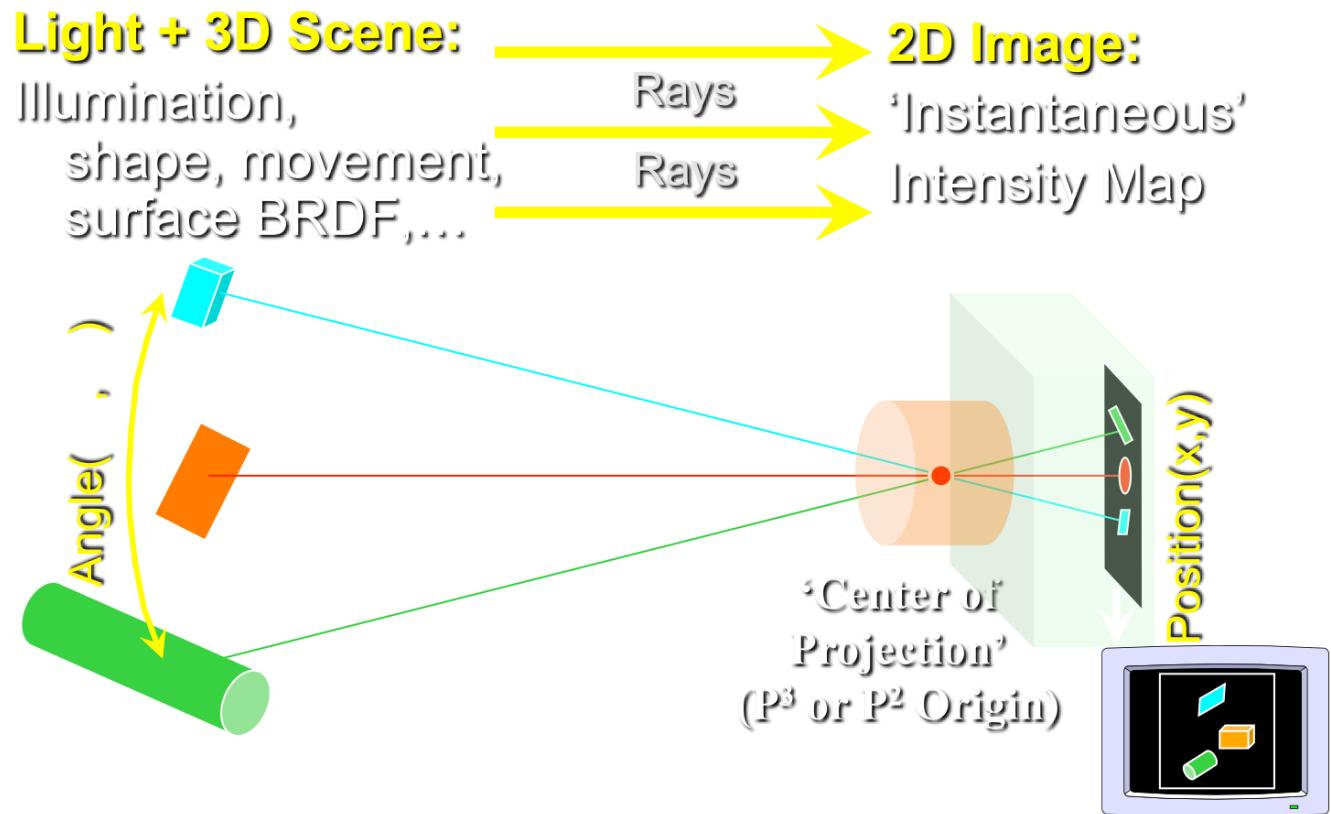
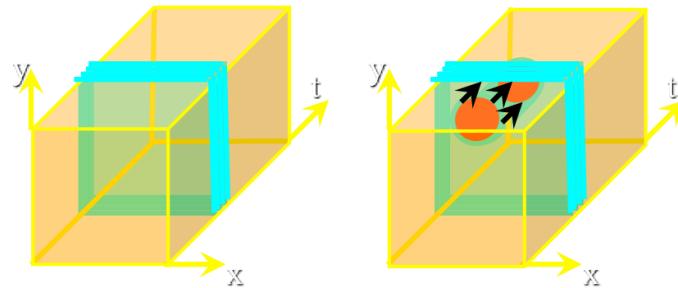
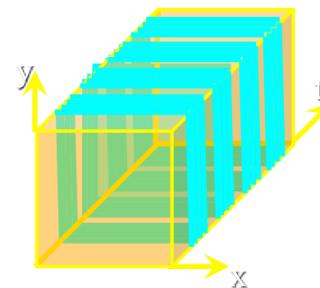


Image Aggregation

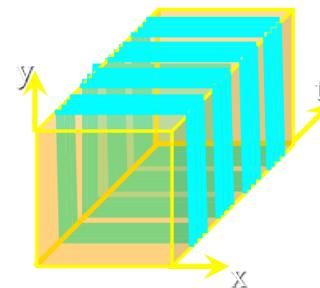
Ordinary
Snapshot



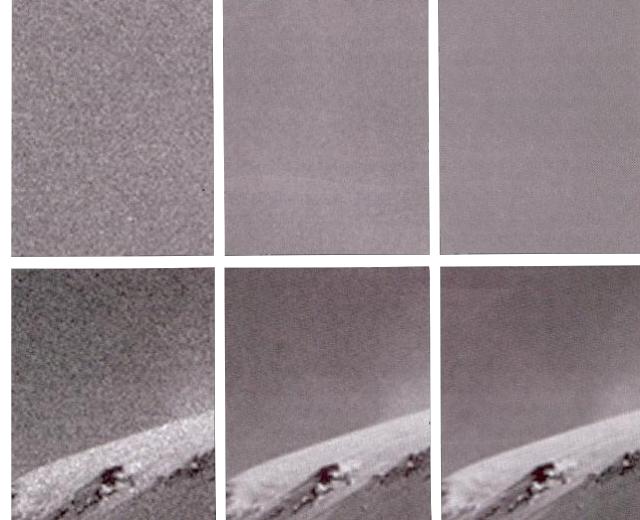
Snapshot with
Motion-Blur



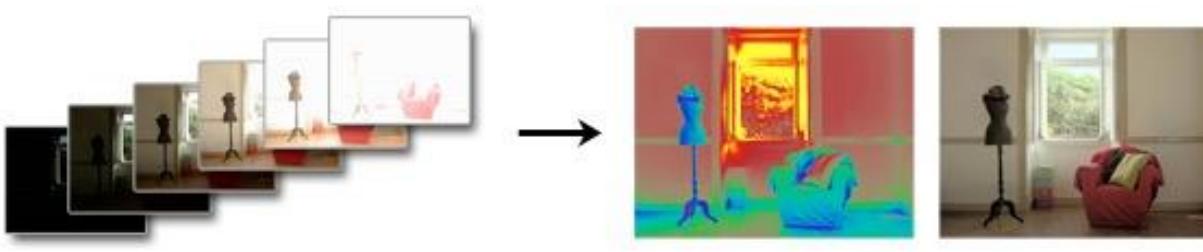
'Motion Picture'
(missing time!)



High-Speed Imaging



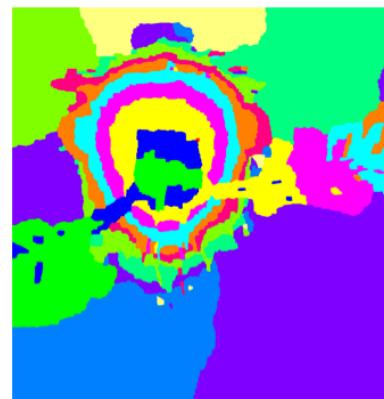
Noise Suppression



High-dynamic Range imaging



Depth from Defocus



Combine multiple conventional
Images computationally

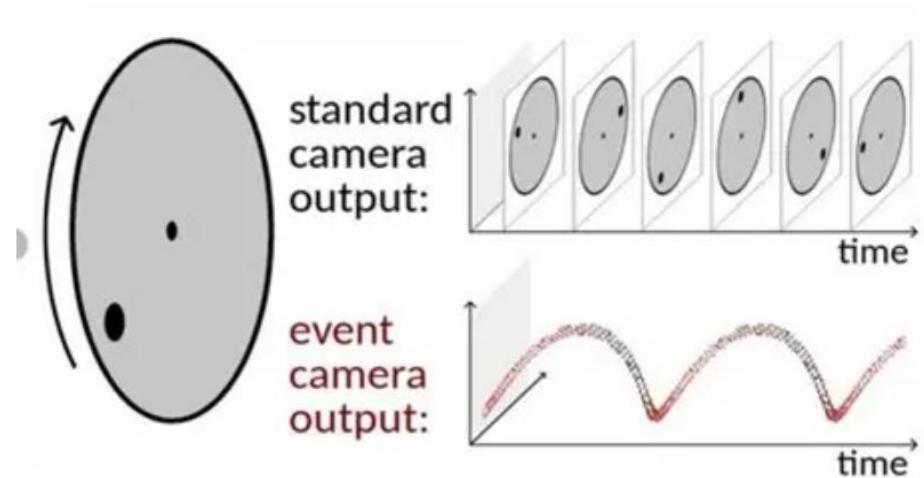


Super-Resolution



Panorama Imaging

Event Cameras



Computational Imaging

2026 Trends in Computer Vision

Gemini

Top Computer Vision Trends in 2026

- **Edge-Native Inference:** A massive shift from cloud-based processing to local **Vision NPUs** (Neural Processing Units). Devices now perform complex real-time analysis on-device to ensure data privacy and near-zero latency.
- **Multimodal "World Models":** Computer vision is no longer isolated. Models now integrate visual data with audio and text simultaneously, allowing AI to "reason" about a scene (e.g., a security camera explaining why a situation looks suspicious in plain English).
- **4D Imaging & Spatial Intelligence:** The integration of **4D imaging radar** and LiDAR has become standard in autonomous systems. This allows for precise depth and velocity sensing in "all-weather" conditions where traditional RGB cameras fail.
- **Event-Based (Neuromorphic) Vision:** Instead of traditional frame-by-frame recording, sensors now only capture pixel changes. This "event-based" approach allows for ultra-high-speed motion tracking with a fraction of the power consumption.
- **Synthetic Data as the Gold Standard:** To solve the "data gap" for rare edge cases (like car accidents or rare medical pathologies), companies are using generative AI to create photorealistic, perfectly labeled synthetic training environments.
- **Foundation Models for Video (SAM 2 & Beyond):** The "Segment Anything" trend has evolved. Systems can now track and segment any object across video frames with perfect consistency, revolutionizing automated video editing and medical surgery assistance.
- **Human-Centric Behavior Analytics:** Beyond facial recognition, 2026 focuses on **intent detection**. Computer vision in retail and automotive now analyzes micro-gestures and eye dynamics to predict a person's next move or state of impairment.

how

where

what

Chat-GPT

Core Technology & Models

- **Foundation & multimodal vision models** — large vision models that combine image, text, and language understanding for richer interpretation and reasoning. apifai
- **Vision-language-action systems** powering embodied AI & robotics (integrating vision with motion and control). Wikipedia
- **Vision Transformers & advanced architectures** replacing older CNNs for better performance. apifai

Deployment & Architecture

- **Real-time edge vision** — on-device inference for low latency and better privacy. Softweb Solutions +1
- **Simulation, digital twins & validation tools** — virtual sensor/data environments for safer training/testing (esp. autonomous systems). SKY ENGI...

Advanced Sensing & Data

- **3D vision & LiDAR integration** — richer spatial understanding for AVs, robotics, AR/VR. Softweb Solutions
- **Event-based and neuromorphic sensors** — highly efficient, low-power vision for dynamic scenes. Softweb Solutions

Ethics & Privacy

- **Privacy-first vision systems** (anonymization, regulatory compliance) as visual AI spreads into society. apifai

Applications & Impact

- **Autonomous mobility & robotics** — vision as a core enabler of safe autonomy. Folio3...
- **Healthcare imaging & diagnostics** — improved clinical insights from visual analysis. Folio3 AI
- **Smart infrastructure & surveillance** — real-time monitoring for cities and industry. Folio3 AI

Claude

Based on the latest information, here are the key computer vision trends for 2026:

- **Physical AI and edge deployment** - Real-time inference running on local AI chips with many award-winning products at CES 2026 BasicAI
- **Edge AI for real-time processing** - Processing data locally on devices for autonomous driving, surveillance, and industrial automation, reducing cloud dependency Softweb Solutions
- **Vision Transformers (ViTs)** - Neural networks using self-attention mechanisms outperforming CNNs in image recognition and object detection Viso.ai
- **Generative AI integration** - Using GANs and diffusion models to create synthetic data and expand training datasets Viso.ai
- **3D computer vision** - Depth sensing using structured light, time-of-flight sensors, and stereo vision for VR, AR, and robotics Viso.ai
- **4D imaging radar fusion** - Combining multiple radar sensors with deep learning for cost-effective autonomous vehicle perception BasicAI
- **Ethical AI and bias mitigation** - Increasing focus on addressing biases in image recognition and analysis systems XenonStack
- **Multimodal AI** - Combining computer vision with natural language processing for richer understanding

Interested in Research Projects?



Basic Research: Light Fields

Disciplines:

Computational Imaging & Optics

Computer Vision / Image Processing

Areas:

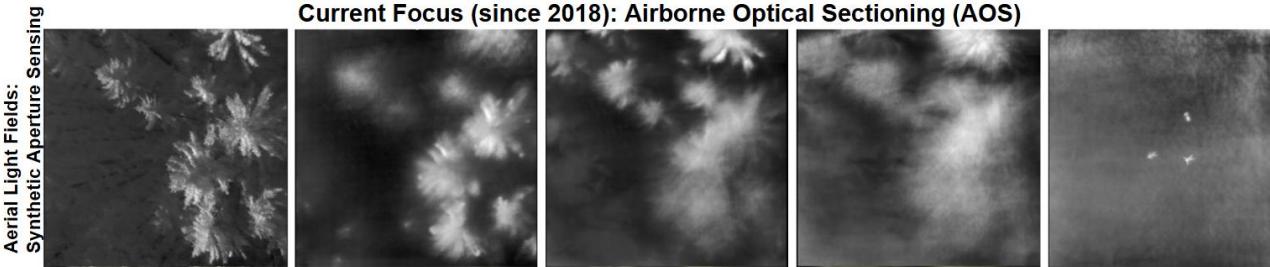
Aerial Imaging, Microscopy, Thin-Film Sensing, 3D Displays, Autonomous Drone and Swarms

Applications:

Search and Rescue, Wildlife Observation, Wildfire Detection, Forest Ecology, Optogenetics, User Interfaces



Current Focus (since 2018): Airborne Optical Sectioning (AOS)



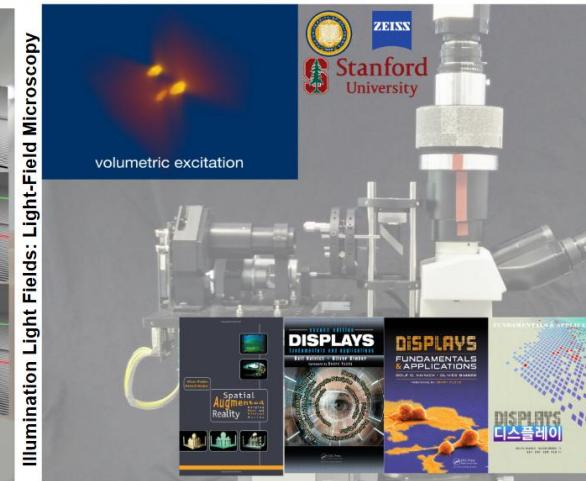
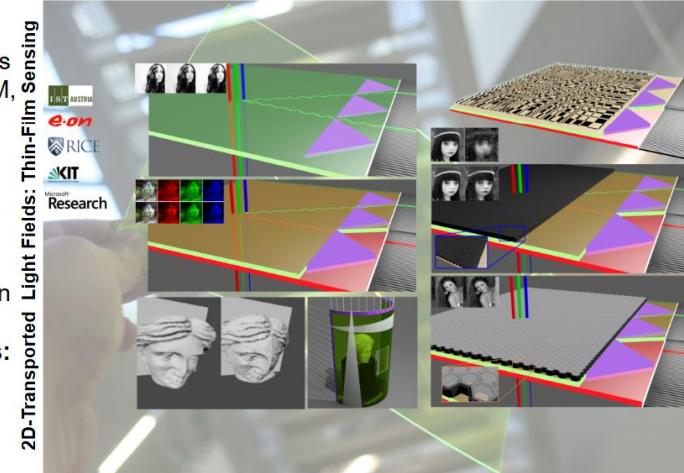
oliver.bimber@jku.at



Publications: Nature Machine Learning, Science Robotics, Nature Communications Engineering, IEEE, ACM, ...

Funding: FWF, DFG, Upper Austria, LIT, EU, Microsoft Research, FFG, Helmholtz Association, Nokia, e.on

Current Collaborators: Cambridge U., EPFL, DLR, BOKU, OOELFV, U. Magdeburg, Helmholtz UFZ



Illumination Light Fields: Light-Field Microscopy

Exam

Places: HS1 (Linz), S1+S3 (Vienna), R4/R5 (Bregenz).

Process: We will check your student IDs at 12:00-12:20. The exam will be open from 12:20-13:20. Results are available after 14:00. On average you need 3 minutes per question (60 minutes in total for 20 questions). However, some answers need more time, while others need less. Please read the questions and answer options very carefully! If you are asked for numbers, please type them in without units (only numbers).

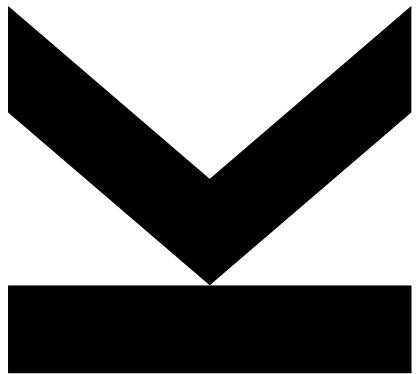
Regulations: This is a closed-book exam. You are allowed to use your laptop (only the test page open - nothing else), clean paper, and pencil only (no additional mobiles, no headsets). Moodle logfiles will be checked for consistent user-IDs and IP-addresses. Any suspected chicanery will lead to a failed attempt of the entire course (without option to attend the retry exam in 2026). All students are responsible to ensure a laptop, browser, internet connection, and access to the Moodle exam from the lecture hall/seminar rooms. Every student who is not present for checking the student ID will not be able to participate. On your desktop only the Moodle exam is allowed to be open.

Rules: Questions and answer options might contain typos. If you have difficulties in understanding (or are unsure about) a question or an answer option, inform the supervisor immediately during the exam. We will neither discard questions with typos from the grading later on, nor accept complains that questions could not be answered correctly due to misunderstandings. Any results that you will see in Moodle are preliminary results until you receive your final grade in KUSSS. Preliminary results might be subject to changes (e.g., when questions have to be removed after the exam) that might have a (positive or negative) influence to your final grade. You have no right to insist on preliminary grading results. Your grades in KUSSS are final.

Course Overview

| CW | Topic | Date | Place | Lab |
|-----|----------------------------------|------------|--|--------------|
| 41 | Introduction and Course Overview | 07.10.2025 | Zoom | Lab 1 |
| 42 | Capturing Digital Images | 14.10.2025 | Zoom | Lab 2 |
| 43 | Digital Image Processing | 21.10.2025 | Zoom | Assignment 1 |
| 44 | Machine Learning | 28.10.2025 | Zoom | |
| 45 | Feature Extraction | 04.11.2025 | Zoom | Open Lab 1 |
| 46 | Segmentation | 11.11.2025 | Zoom | Assignment 2 |
| 47 | Optical Flow | 18.11.2025 | Zoom | Open Lab 2 |
| 48 | Object Detection | 25.11.2025 | Zoom | Assignment 3 |
| 49 | Multi-View Geometry | 02.12.2025 | Zoom | Open Lab 3 |
| 50 | 3D Vision | 10.12.2025 | Zoom | Assignment 4 |
| 3 | Trends in Computer Vision | 13.01.2026 | Zoom | |
| → 4 | Q&A | 20.01.2026 | Zoom | Open Lab 4 |
| 5 | Exam | 27.01.2026 | HS1 (Linz), S1/S3 (Vienna), S5 (Bregenz) | |
| 9 | Retry Exam | 24.02.2026 | tba | |

Thank You



Please fill out evaluation forms.

Stay in touch: 