

Minimalist Vision with Freeform Pixels

Jeremy Klotz and Shree K. Nayar

Columbia University

ECCV 2024 Best Paper

PRESENTER: LILANG LIN

2024/10/13

Lightweight Vision

Workspace Monitoring



shutterstock.com

How many people are in the room?

Which desks are occupied?

Traffic Monitoring



youtu.be/PNCJQkvALVc

How fast is traffic moving?

Is there an accident?

High-Level Inference

Statistics of Objects

Aggregate Motion of Objects

Lightweight Vision

Scene



Camera



Captured Video

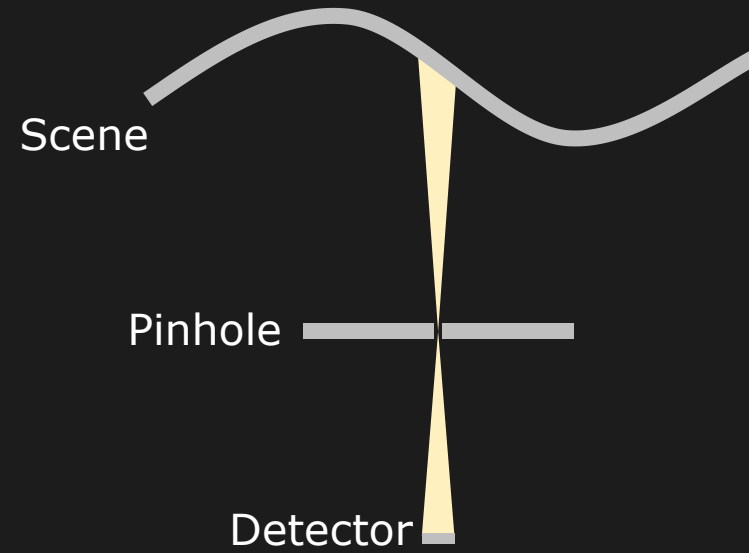
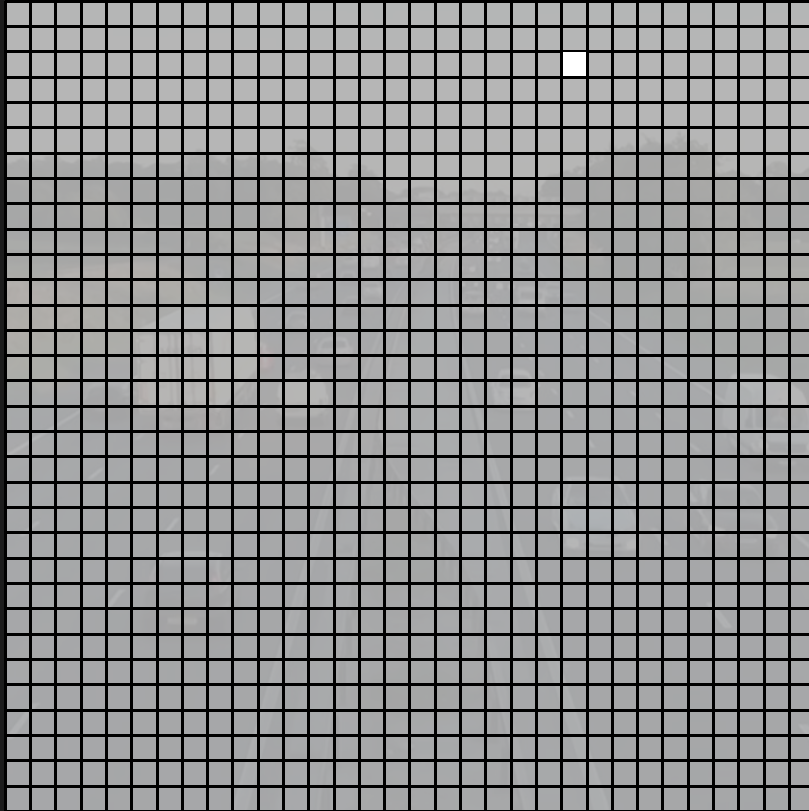


Benefits

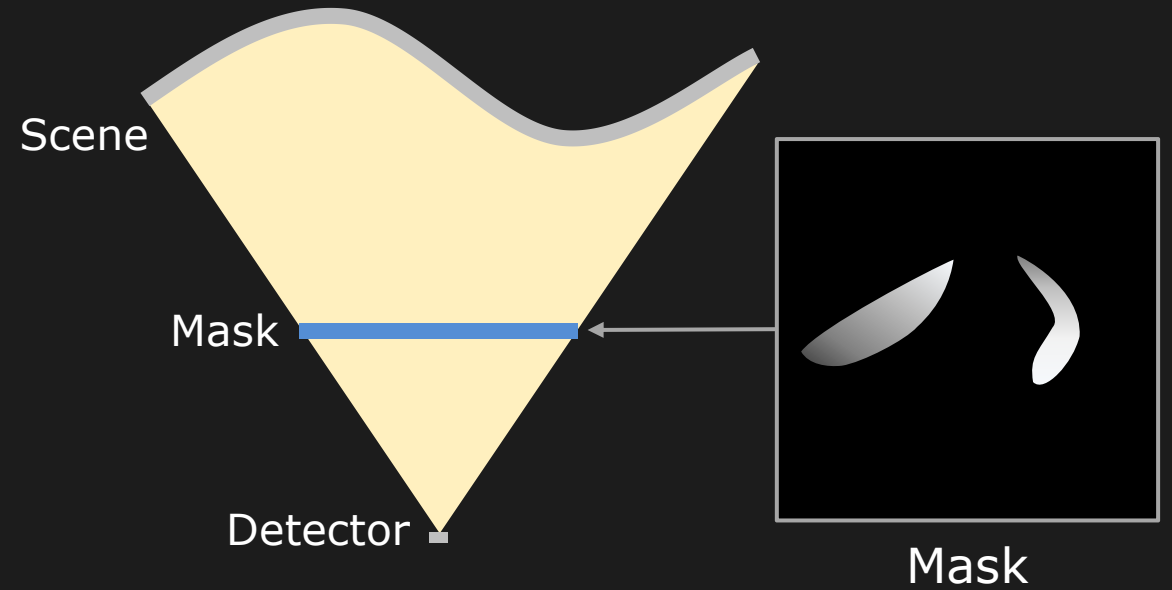
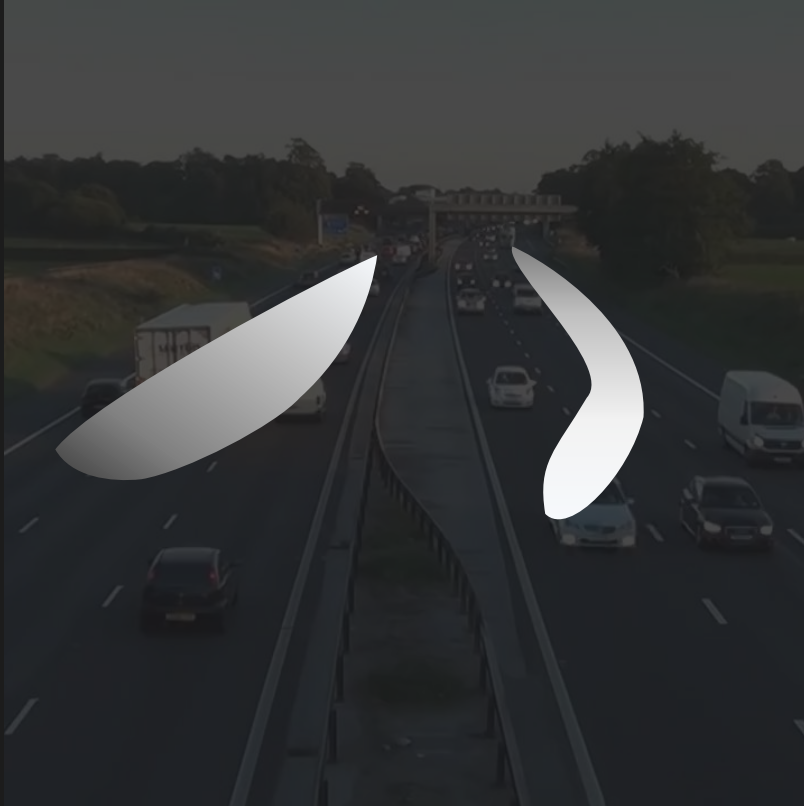
Privacy Preservation

Self-Sustainability

Square Pixels



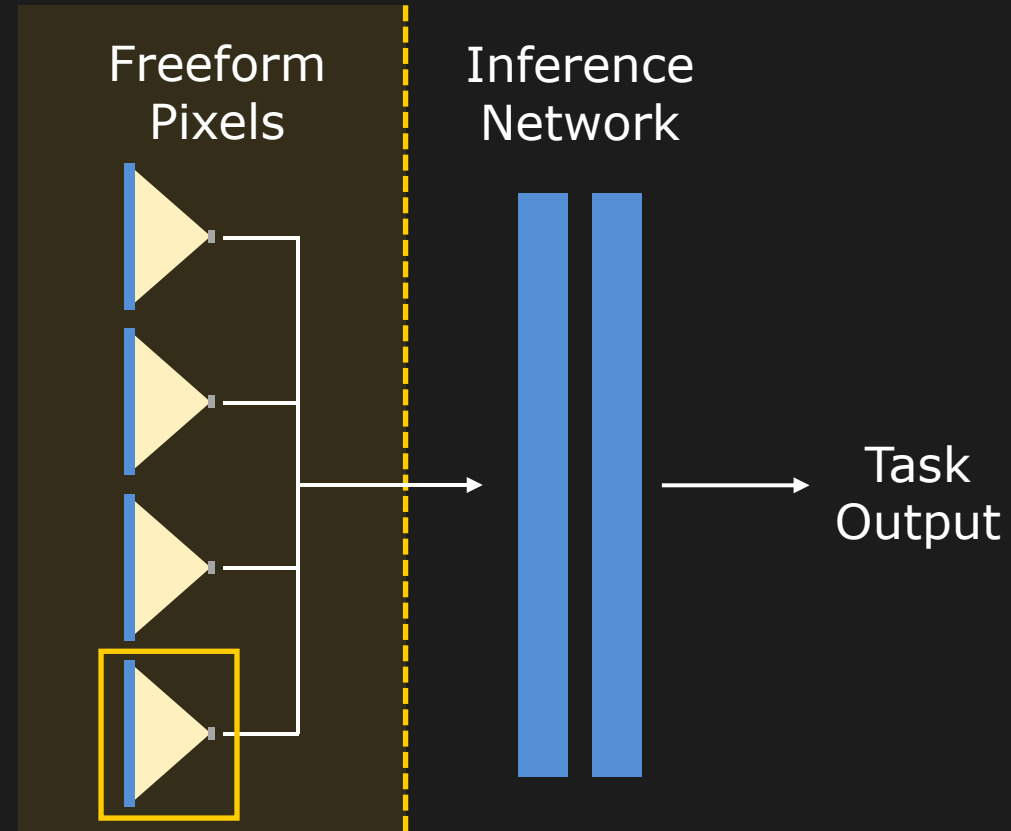
Freeform Pixels



Goal: Design the Fewest Freeform Pixels Needed to Solve a Task

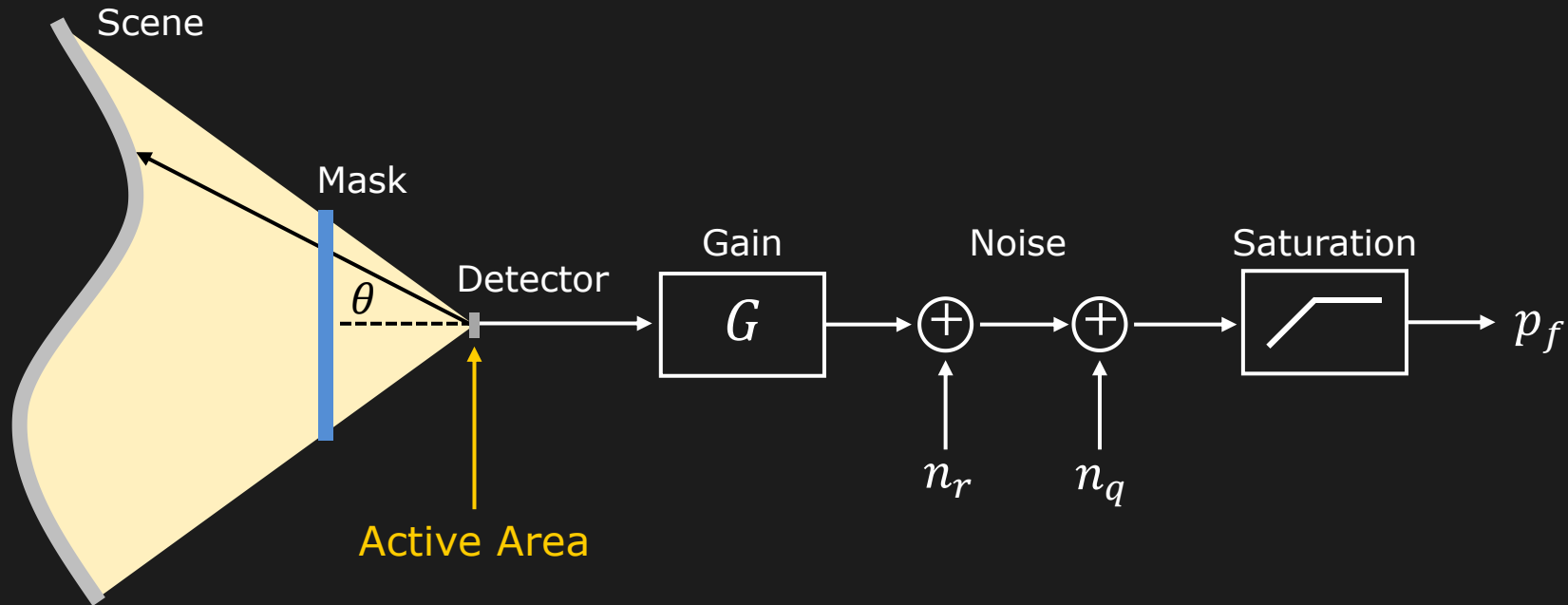
Minimalist Vision

Scene



Fabricated in Hardware

Sensor Model



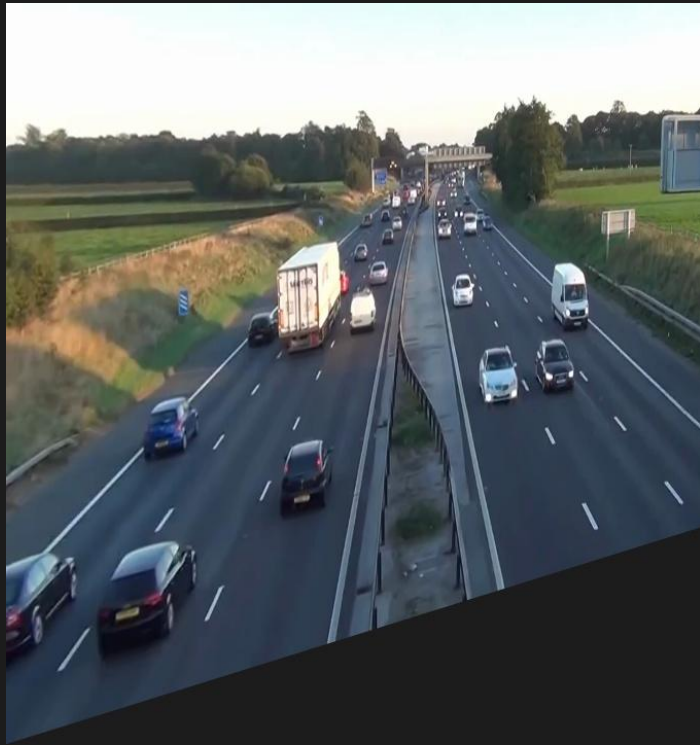
$$\iint_{x,y} I(x,y) \left(\underbrace{M(x,y)}_{\text{Mask}} \cdot \underbrace{d(\theta)}_{\text{Directivity}} \cdot \underbrace{dx dy}_{\text{Active Area}} \right) \cdot \underbrace{G}_{\text{Gain}} + \underbrace{n_r + n_q}_{\text{Noise}} \longrightarrow \underbrace{S(\cdot)}_{\text{Saturation}} \longrightarrow p_f$$

↑
Transmittance in [0, 1]

Measurement is Differentiable w.r.t. Mask Parameters

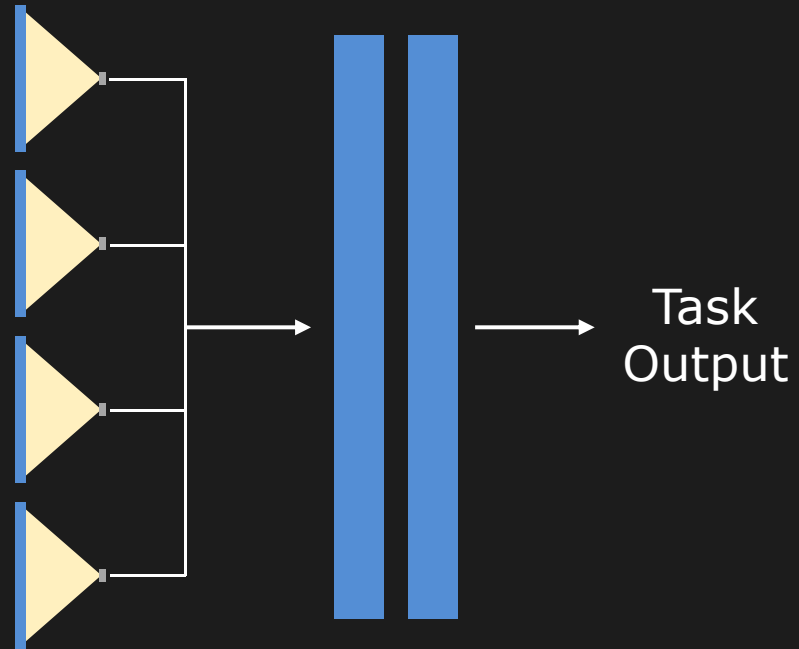
Minimalist Vision with Sensor Model

Scene



Freeform
Pixels

Inference
Network

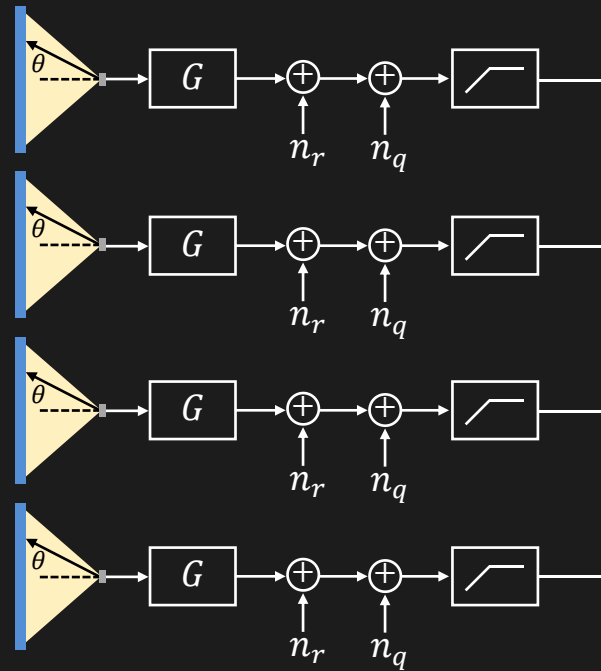


Minimalist Vision with Sensor Model

Scene



Freeform
Pixels



Inference
Network



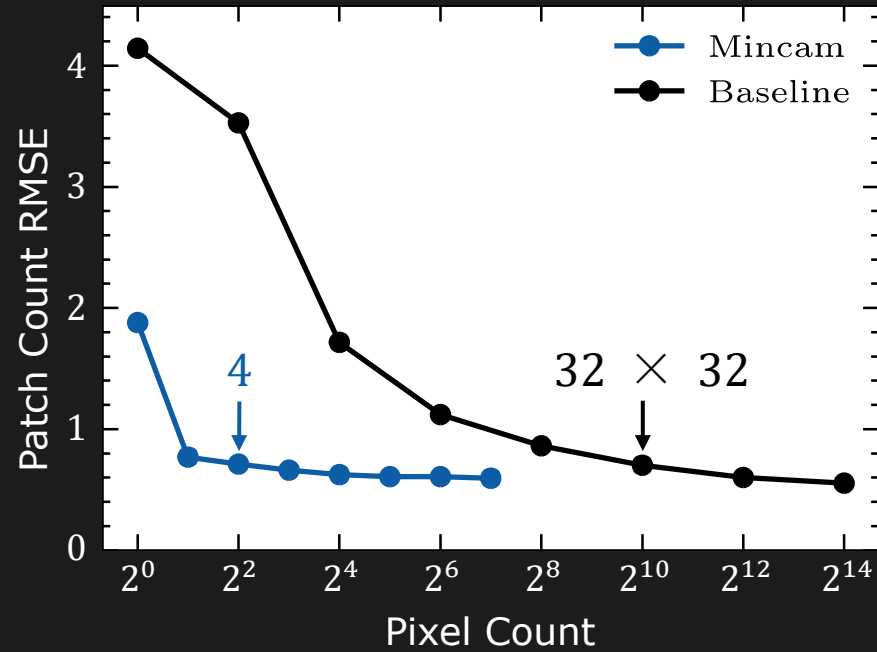
Task
Output

Toy Problem

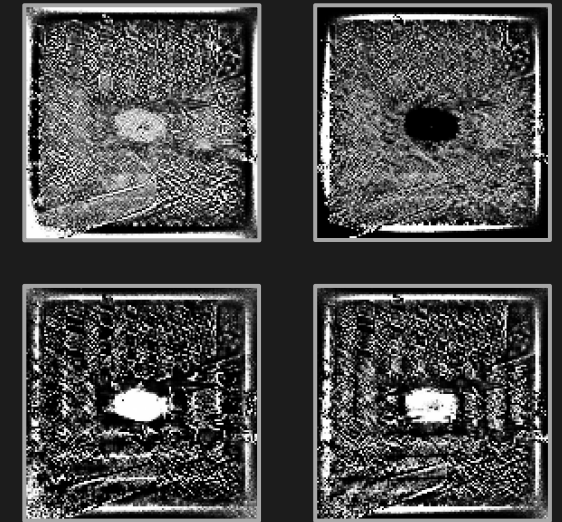
Counting Patches



Counting Performance

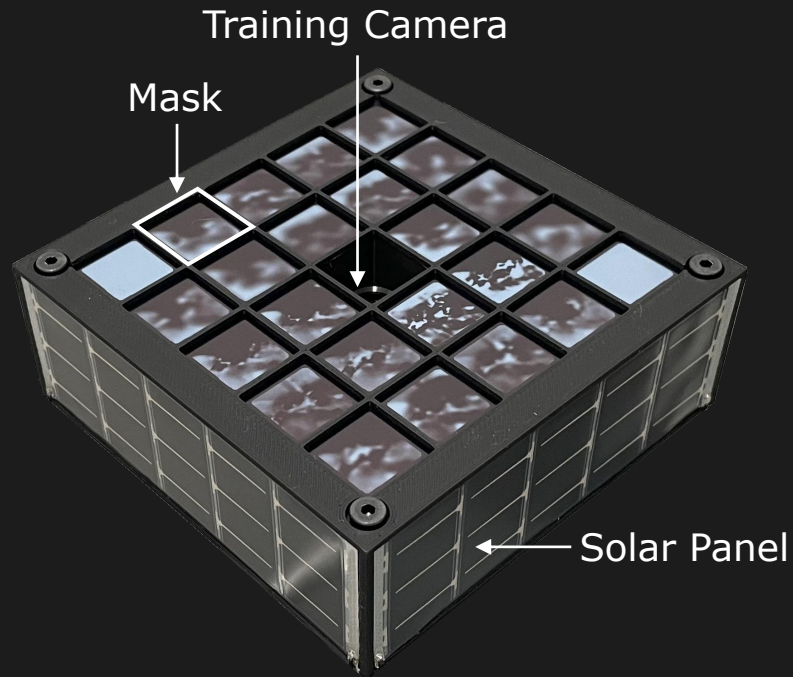


Freeform Pixels



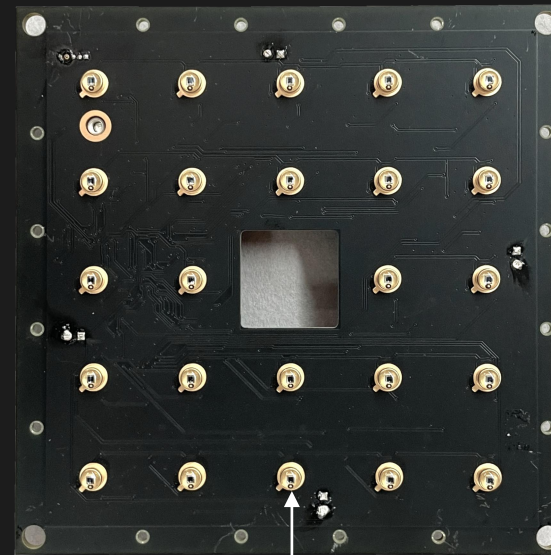
256x Reduction in Pixel Count using Freeform Pixels

Minimalist Camera Prototype



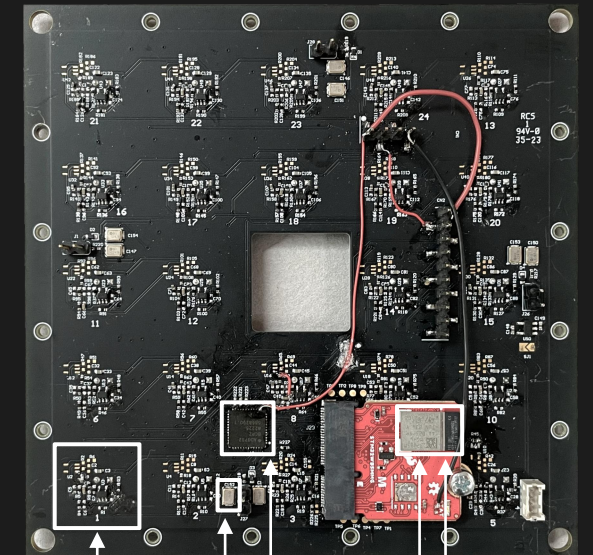
Masks on Transparency

Imaging Board (Front)



Photodiode

Imaging Board (Back)



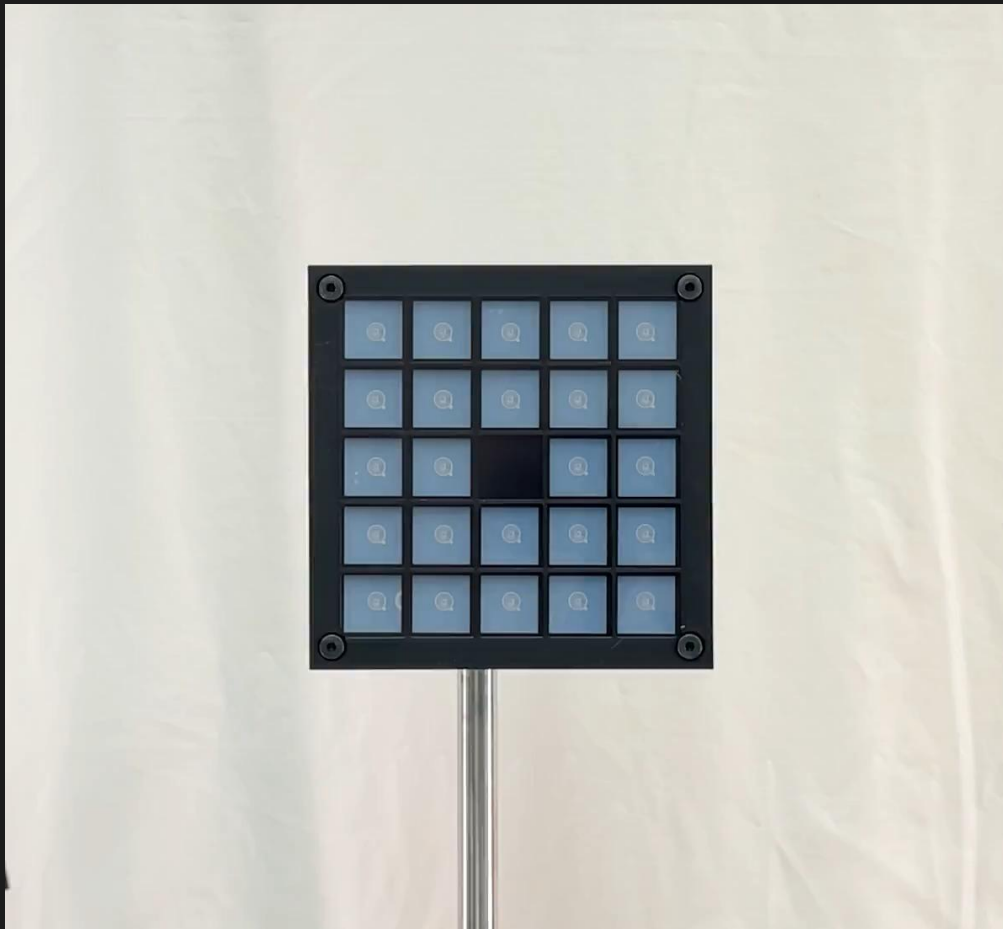
Amplifier

Mux
Supercap

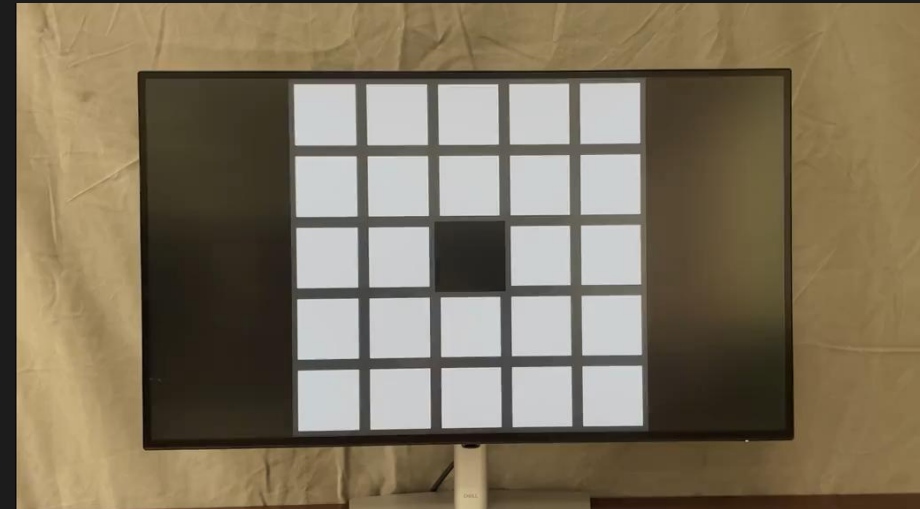
Bluetooth
Microcontroller

Self-Powered Minimalist Camera

Minimalist Camera



Remote Host



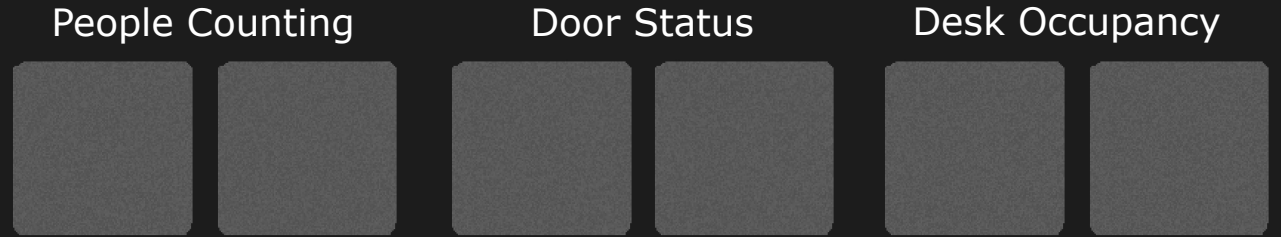
No Battery, Power Supply, or Tether

Workspace Monitoring with 6 Pixels

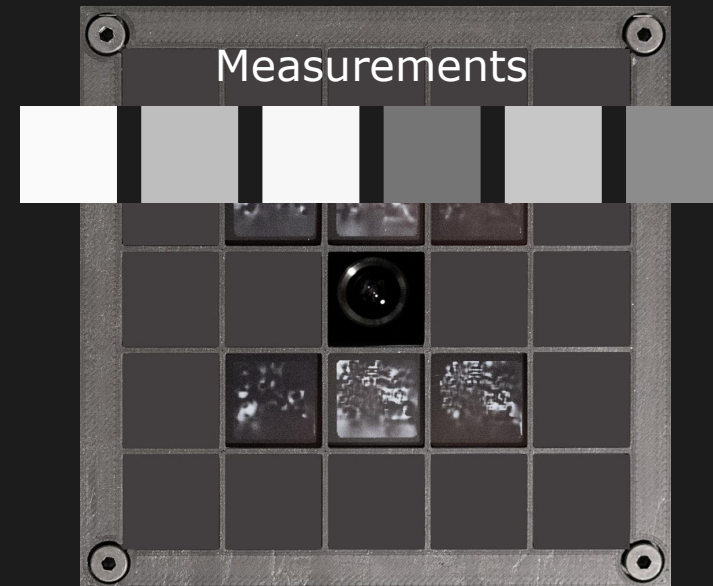
Traditional Camera



Minimalist Camera



Minimalist Camera Prototype



Workspace Monitoring with 6 Pixels

Traditional Camera



Minimalist Camera



Measurements

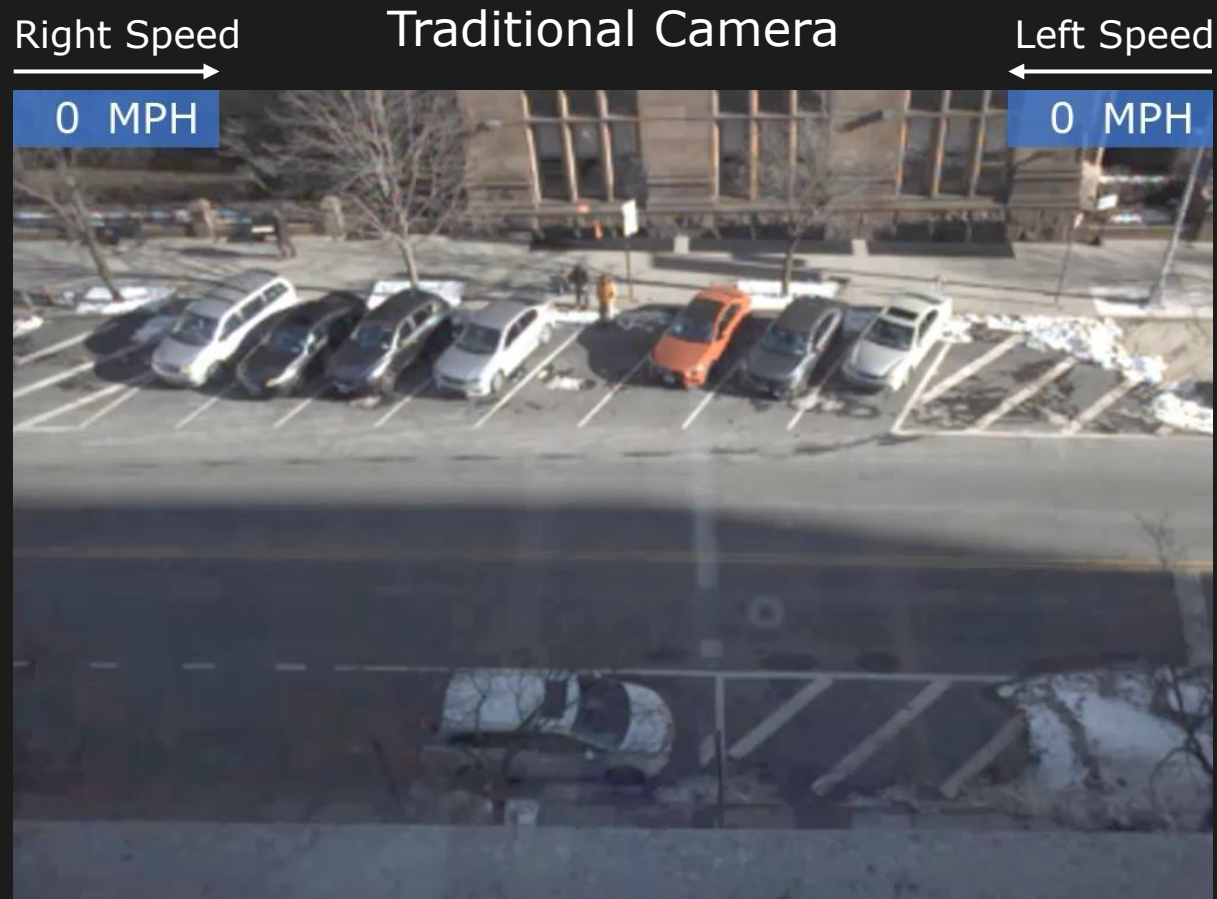


People Count: 1.16 RMS error

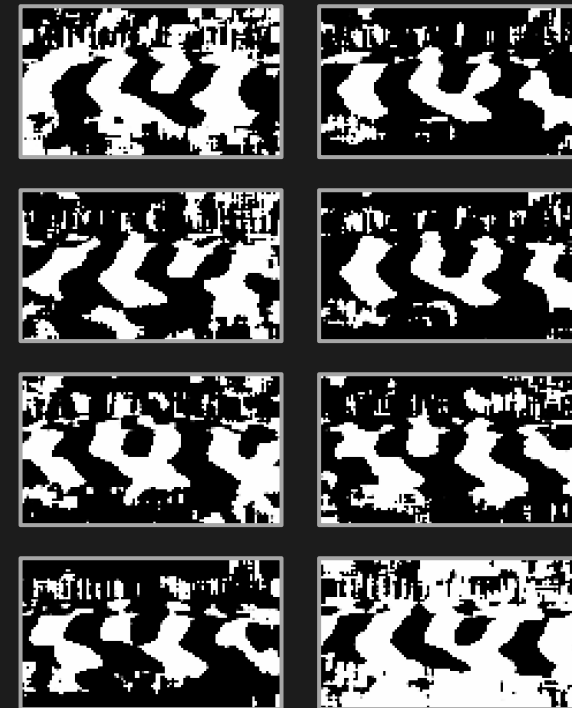
Door Status: 95.4%

Desk Occupancy: 99.2%

Traffic Speed Estimation with 8 Pixels



Minimalist Camera



Right Speed Error (RMS): 2.22 MPH

Left Speed Error (RMS): 2.38 MPH

Minimalist Vision with Freeform Pixels

Cameras on Infrastructure



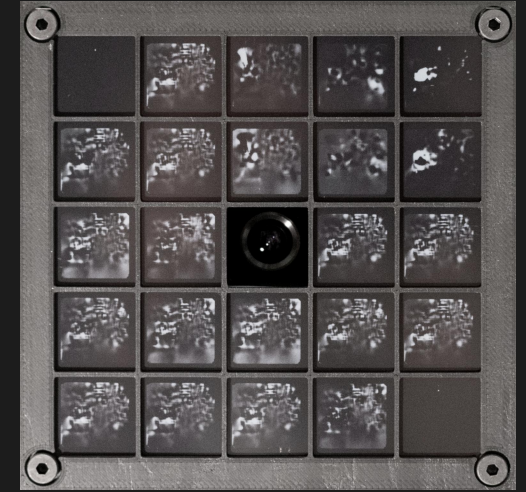
deepsentinel.com

Cameras Carried by People



macworld.com

Minimalist Camera



Acknowledgements

Mikhail Fridberg, Carl Vondrick, Behzad Kamgar-Parsi

Minimalist Vision with Freeform Pixels

Jeremy Klotz and Shree K. Nayar

Columbia University

ECCV 2024

Supported by ONR, NDSEG Fellowship

Why Best Paper

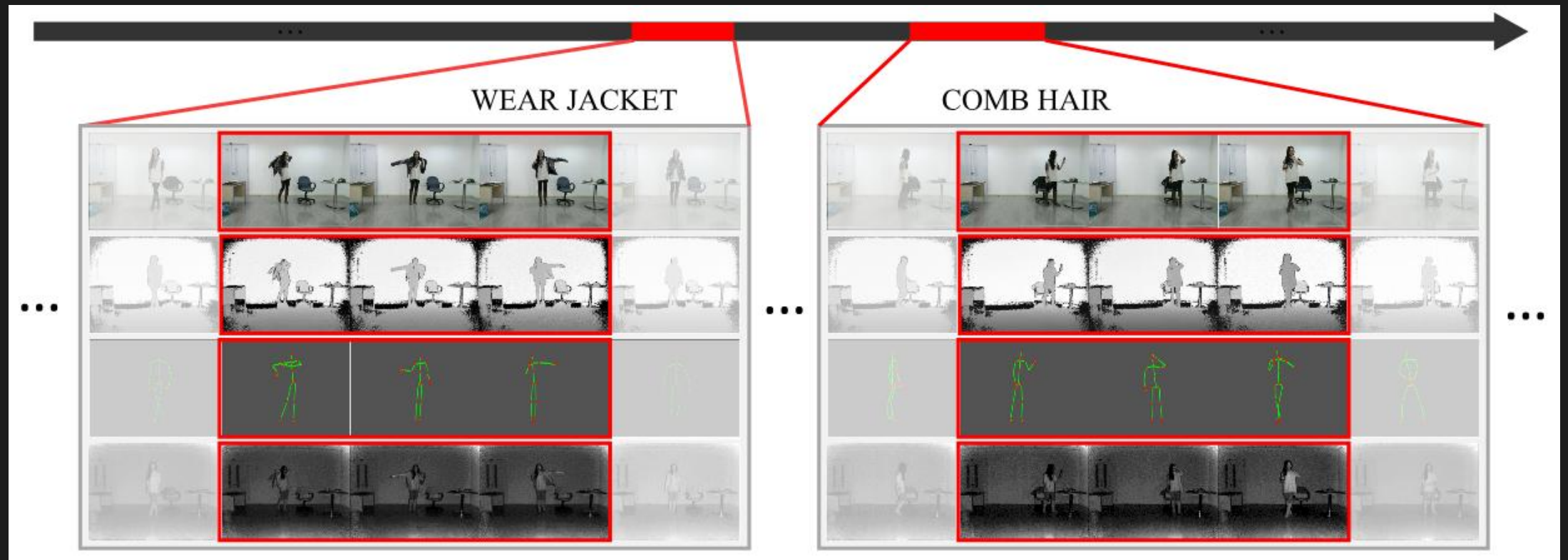
- Minimal information needed to solve a task
 - Compressing
- Algorithm vs Data
 - Freeform pixels as new data modality
 - Fast
 - Privacy Protection

Freeform Pixels

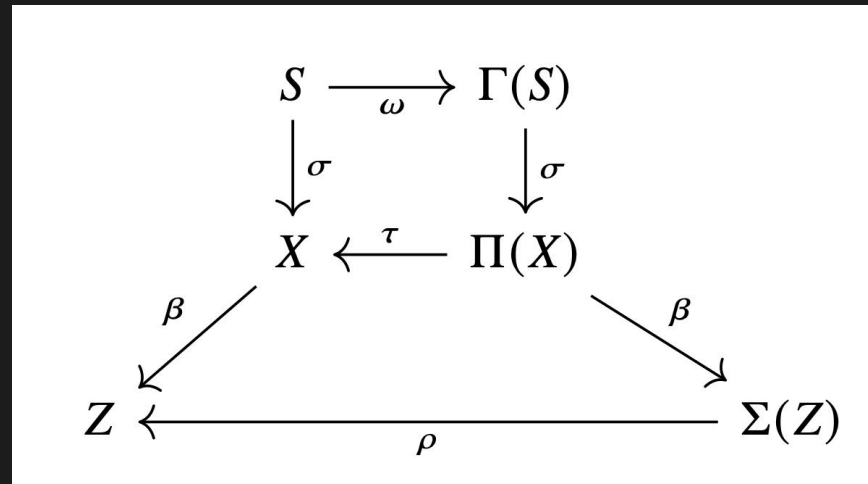
- Data Modality

- Structured Data → Location - Center
 - RGB
 - Event Camera
 - Point Cloud & Mesh
- Semantic Data → Object - Center
 - Text
 - Keypoint & Skeleton
 - Freeform Pixels

Multimodality and Taxonomy



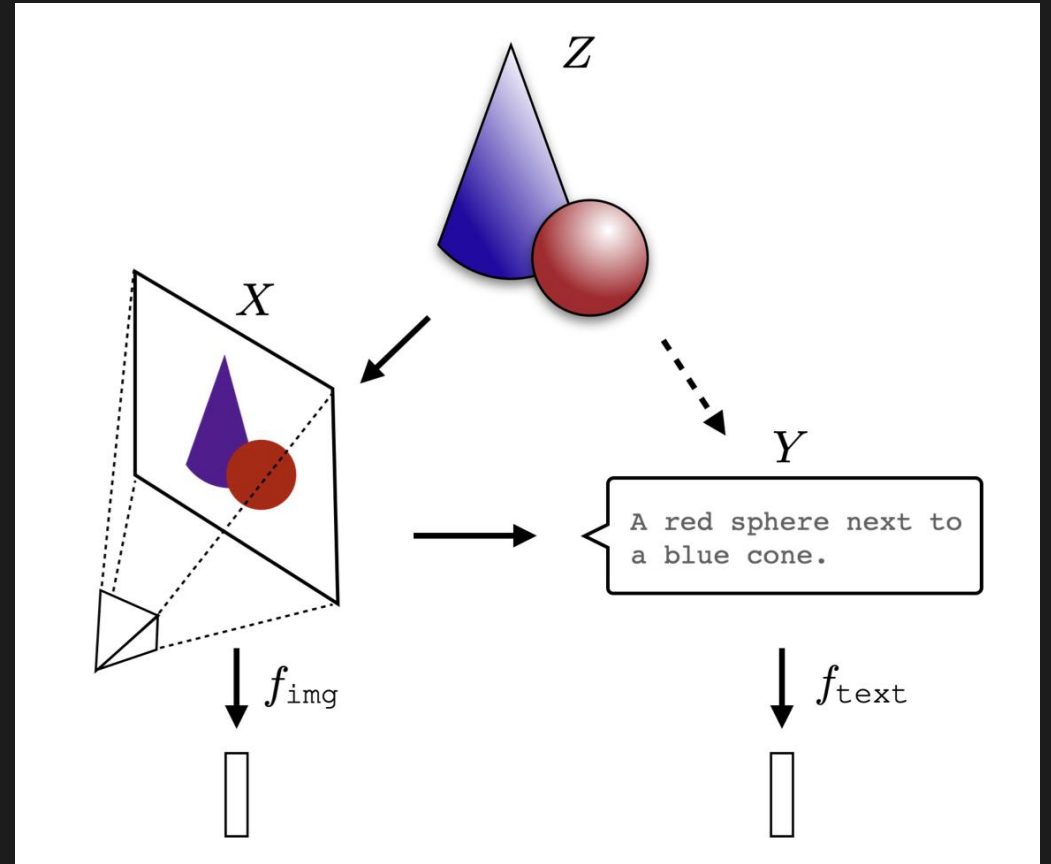
Multimodality and Taxonomy



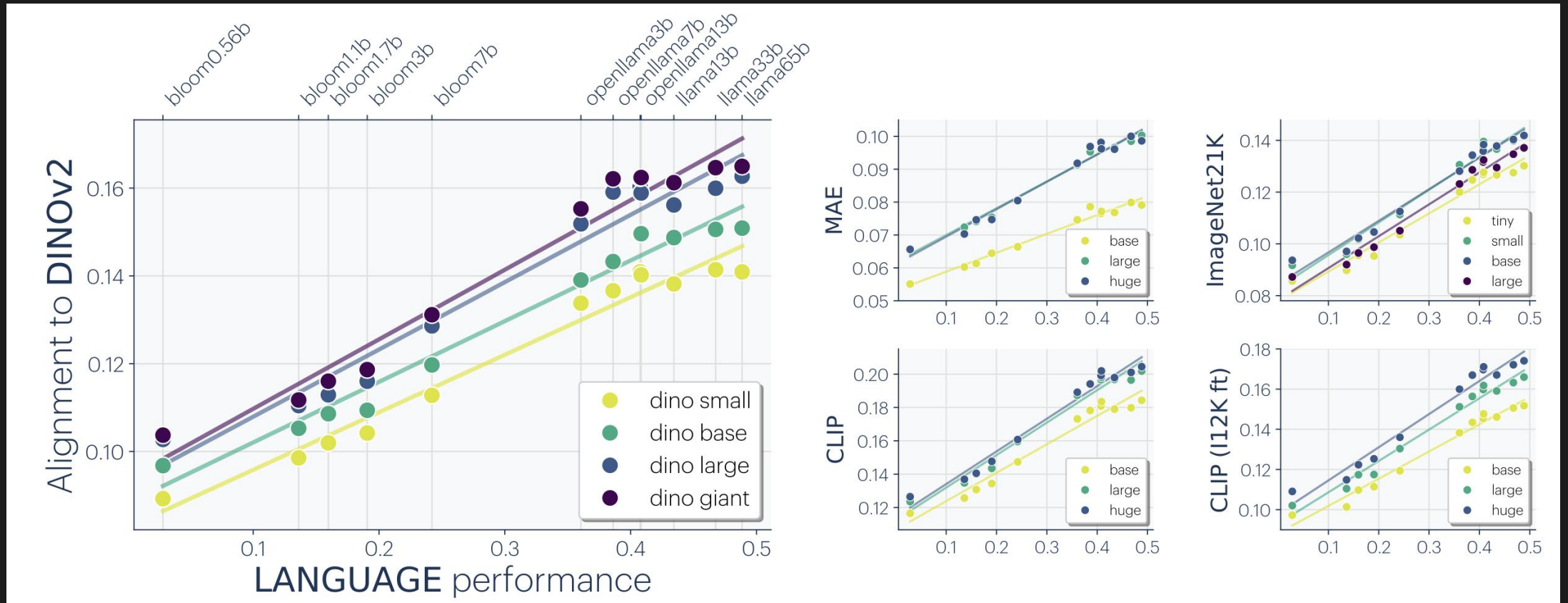
Platonic Representation Hypothesis

The Platonic Representation Hypothesis

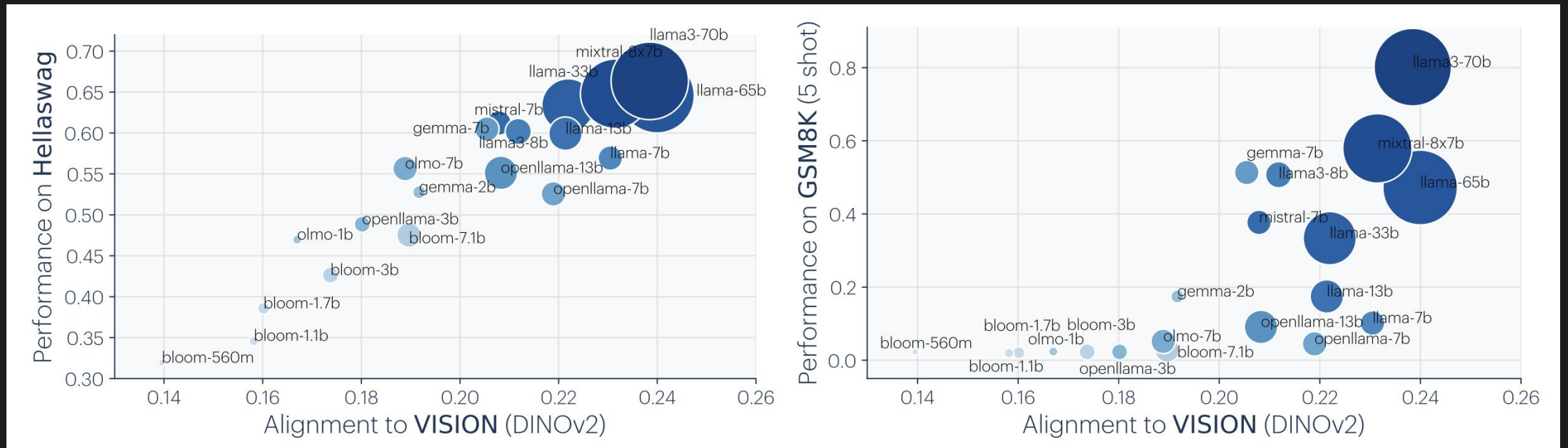
Neural networks, trained with different objectives on different data and modalities, are converging to a shared statistical model of reality in their representation spaces.



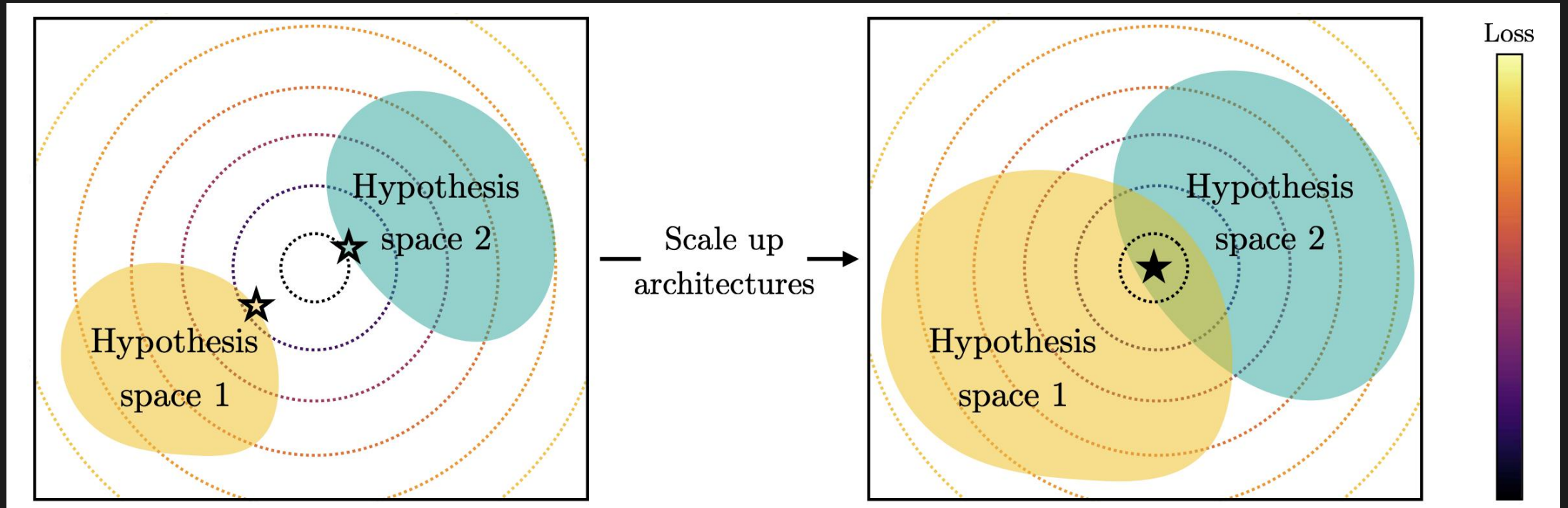
Platonic Representation Hypothesis



Platonic Representation Hypothesis



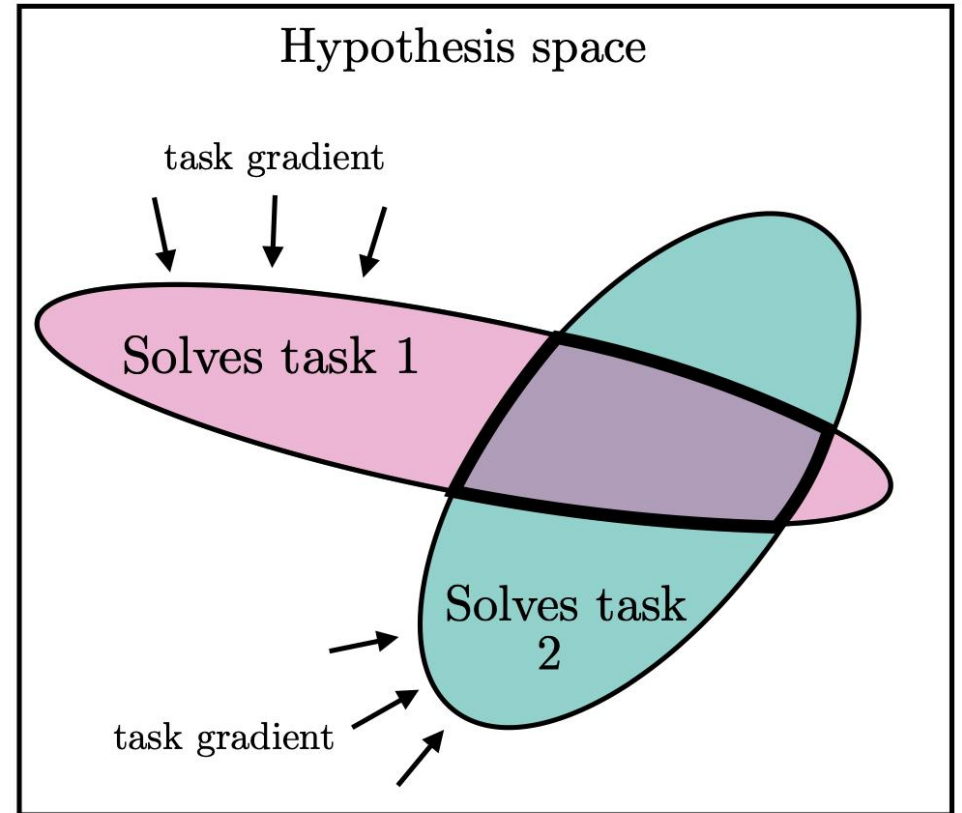
Platonic Representation Hypothesis



Platonic Representation Hypothesis

The Multitask Scaling Hypothesis

There are fewer representations that are competent for N tasks than there are for $M < N$ tasks. As we train more general models that solve more tasks at once, we should expect fewer possible solutions.



Limitations and Discussion

- Scene Changes
- Convert to CNN