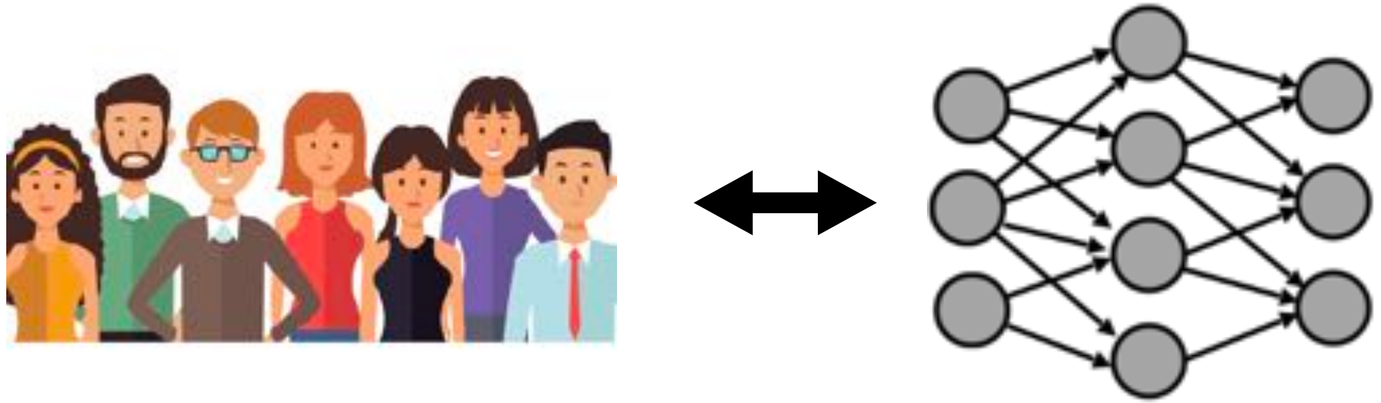
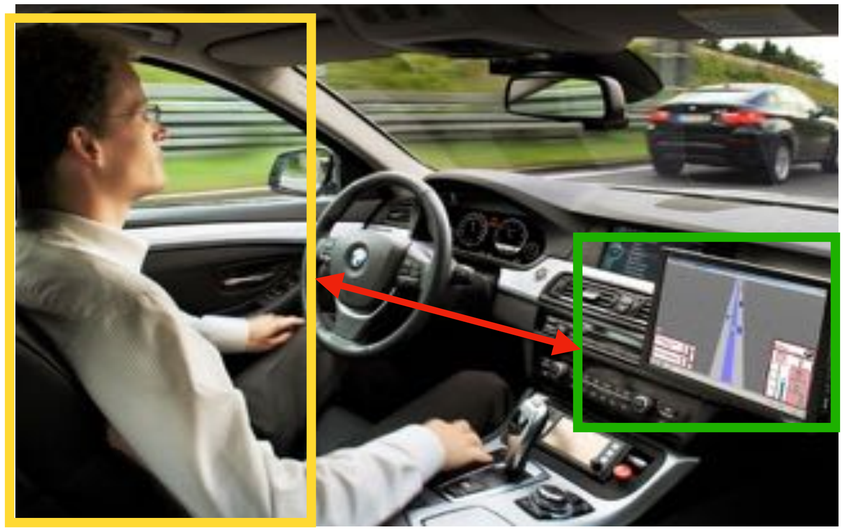


Human-AI Interaction in Computer Vision

Bolei Zhou
Computer Science Department, UCLA

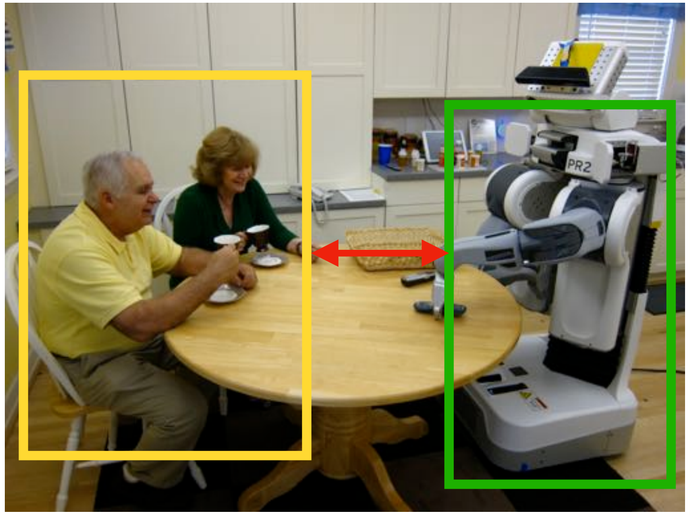
Many real-world Applications of AI and Computer Vision

Self-Driving



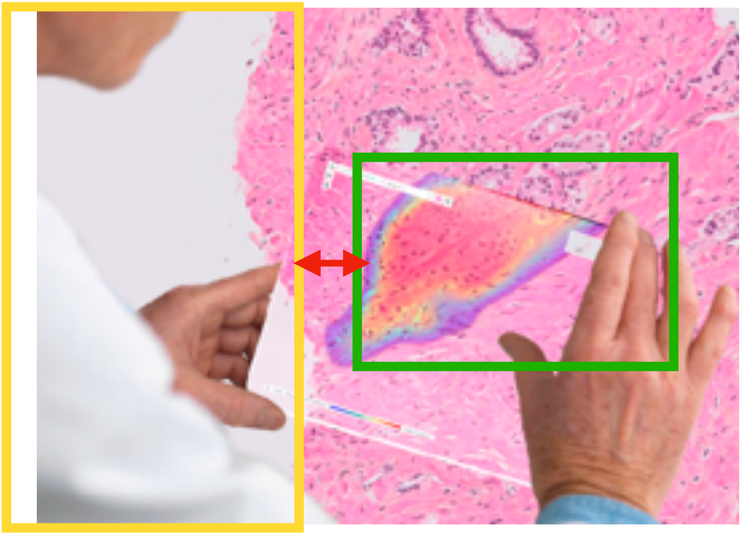
Driver and Autopilot

Home Robot



Elders and Robot

Healthcare

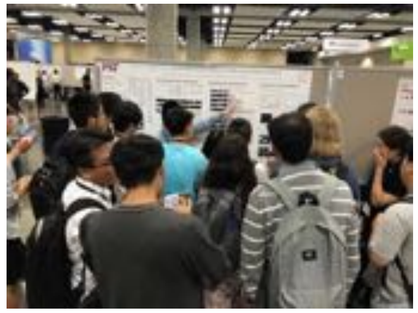


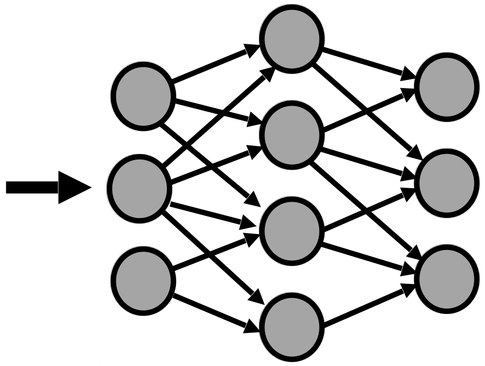
Physician and Medical AI

build trustworthy interactions between humans and AI

AI is achieving human-level and superhuman performance

Classifying image

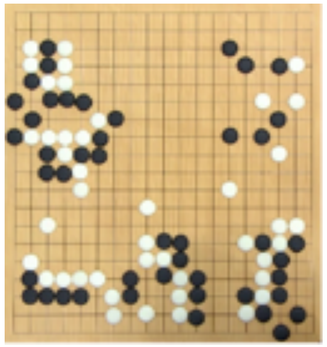


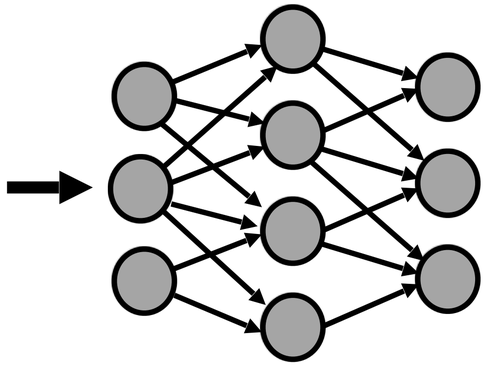


Convention Center

ResNet achieves human-level recognition

Playing strategic game

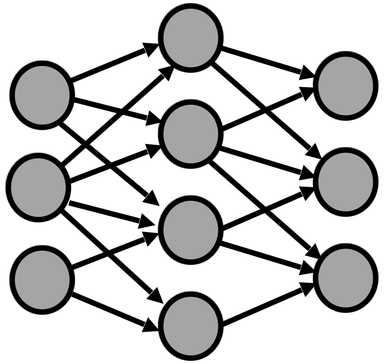




Position 34

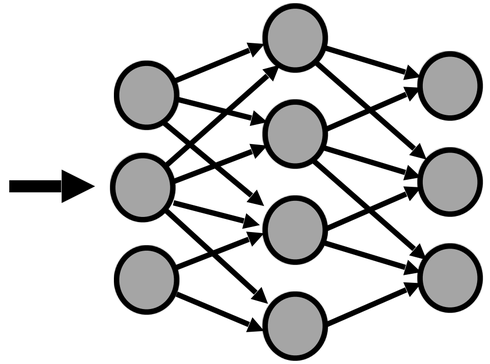
AlphaGo beat Lee Sedal 4-1

Generating image



StyleGAN generates realistic fake images

Steering vehicles



control

Tesla's Autopilot

Narrative of superhuman AI overtaking humans

Machines take over jobs.

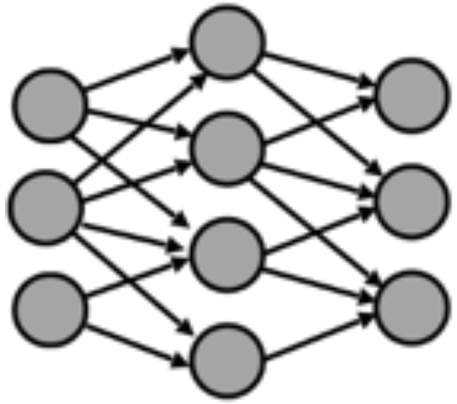
Sci-fi movies: AI threatens mankind

Skynet

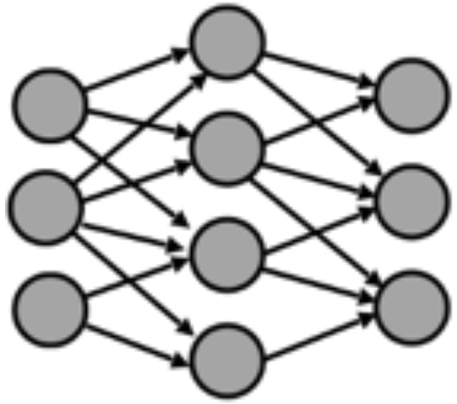
The Terminator

Precogs

Minority Report



Humans and AI shouldn’t be adversarial or competitors



AI should augment human capabilities and offer more possibilities

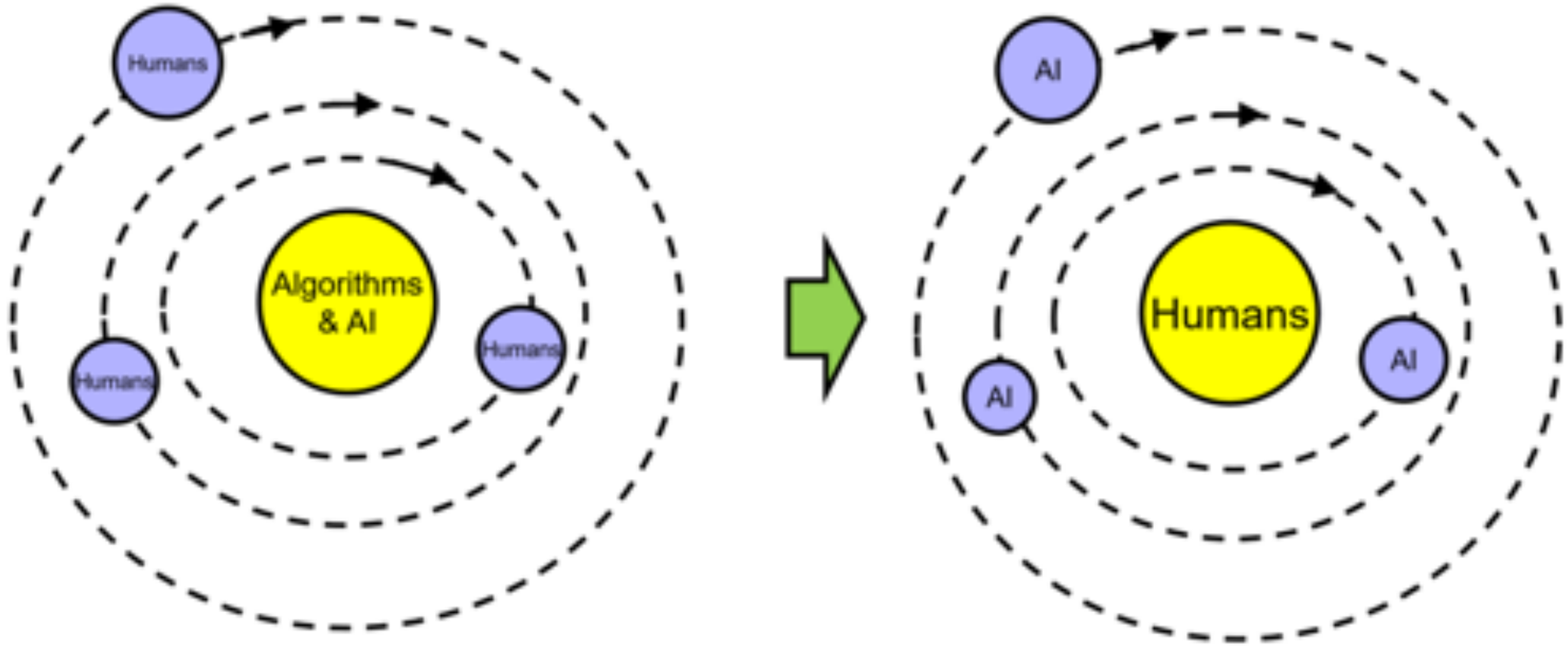
Airplane

Telescope

Reframing the relation between AI and humans

How we develop better AI algorithms to crash benchmarks?

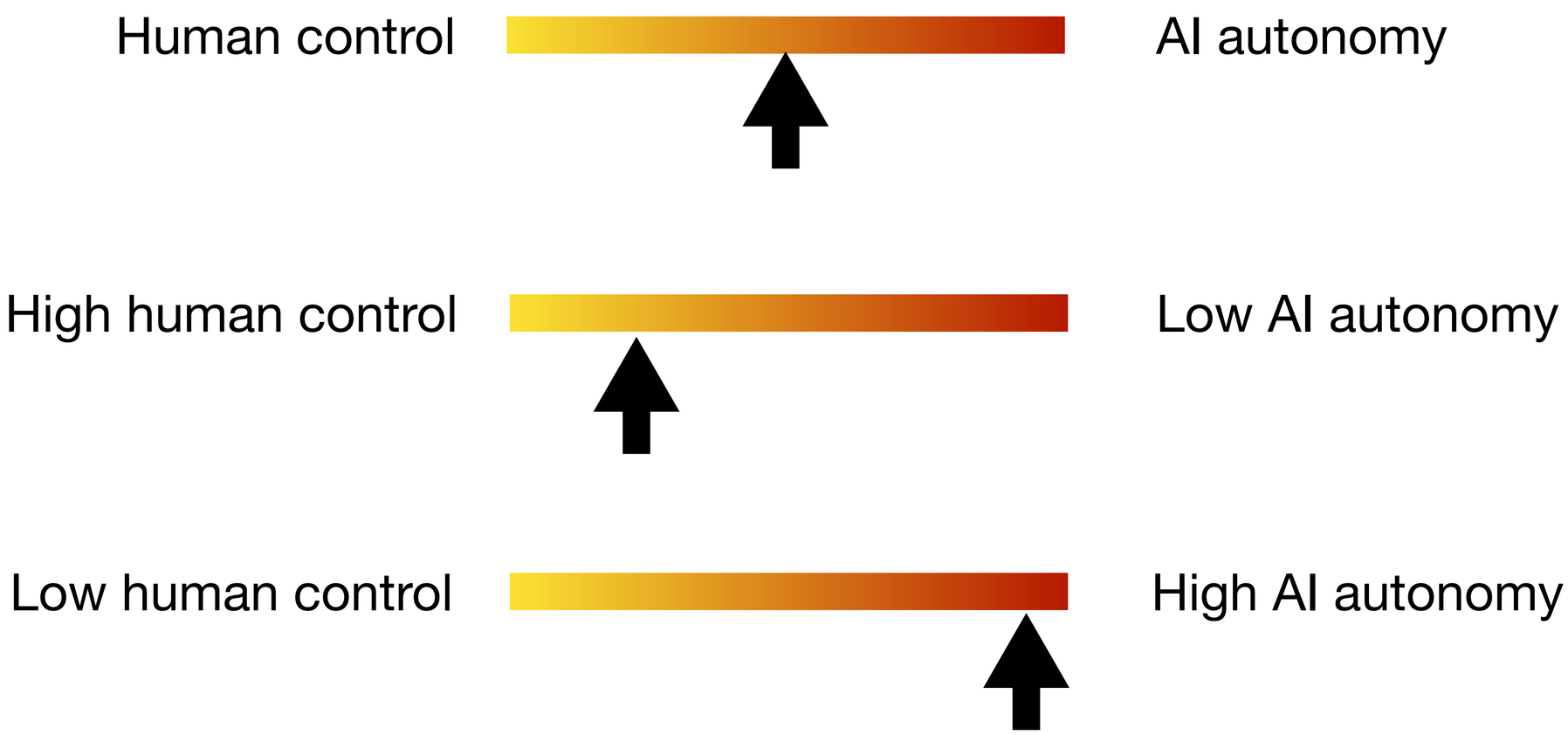
How AI algorithms better serve humans and make our life easier?



Shneiderman, B. (2020). Human-Centered Artificial Intelligence: Three Fresh Ideas. AIS Transactions on Human-Computer Interaction, 12(3), 109-124.

Human control versus AI autonomy

Is it binary?



Adapted from Ben Shneiderman's talk: https://www.youtube.com/watch?v=-VYi_YSNDas

Ensuring human control while increasing autonomy

Reliable, safe & trustworthy systems

Human mastery

High

Human Control

Low

Bike

Piano

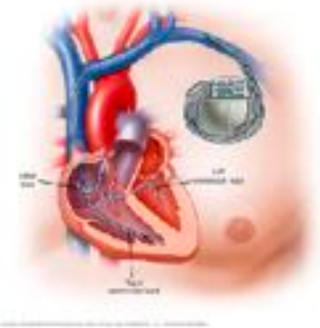
Elevator

Camera

Music box

Landmine

Airbag



Pacemaker

Machine control

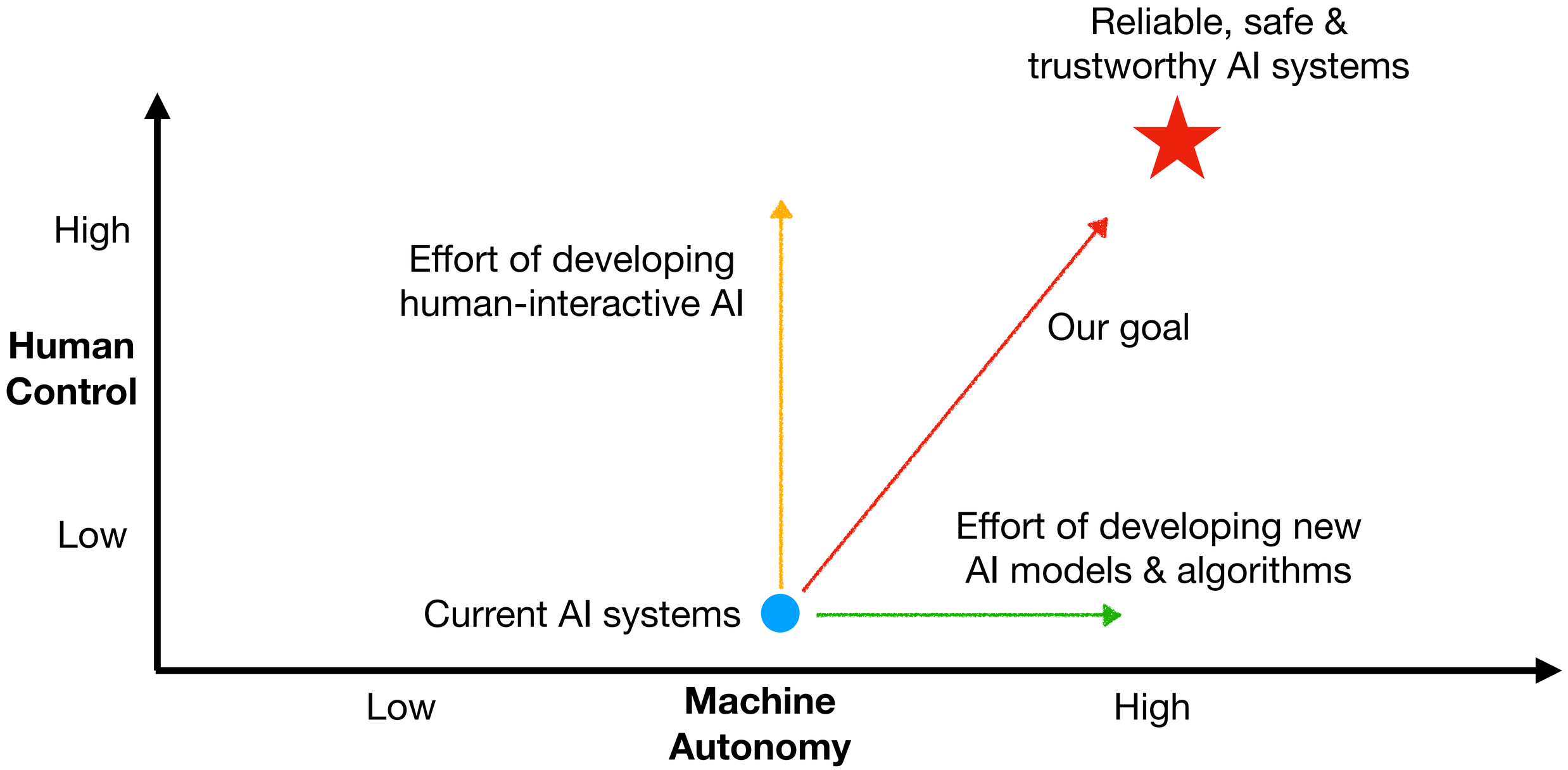
Low

Machine Autonomy

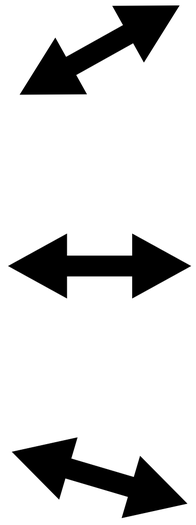
High

Adapted from Ben Shneiderman's talk: https://www.youtube.com/watch?v=-VYi_YSNDas

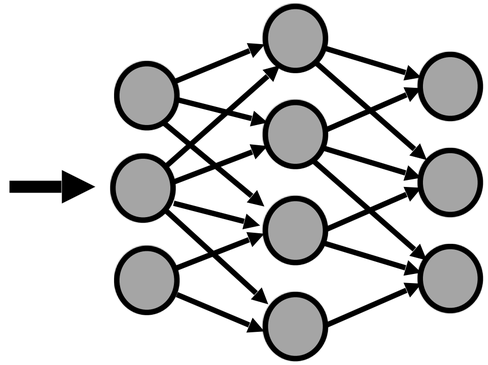
Ensuring human control while increasing autonomy



Human-AI Interaction in Computer Vision

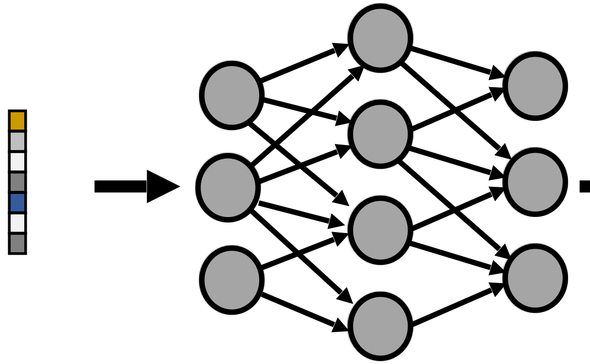


Classifying image

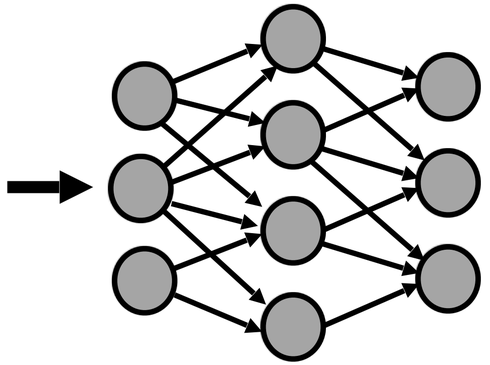


Convention Center

Generating image

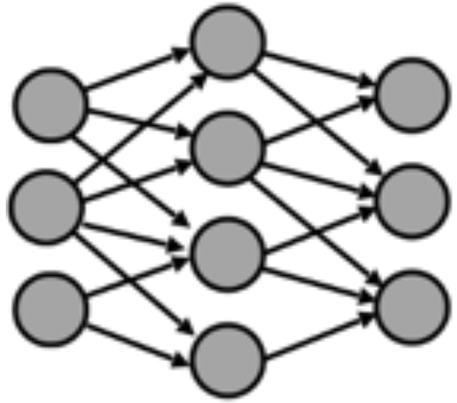


Steering vehicle



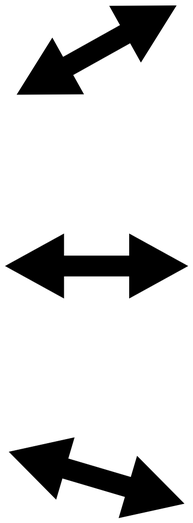
control

Human-AI Interaction

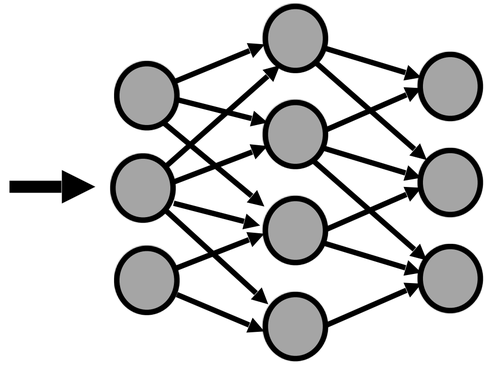


Human-Dog Interaction

Human-AI Interaction in Computer Vision

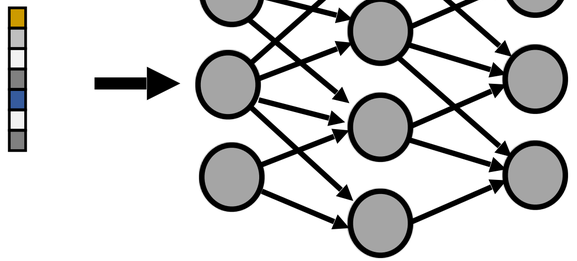


Classifying image

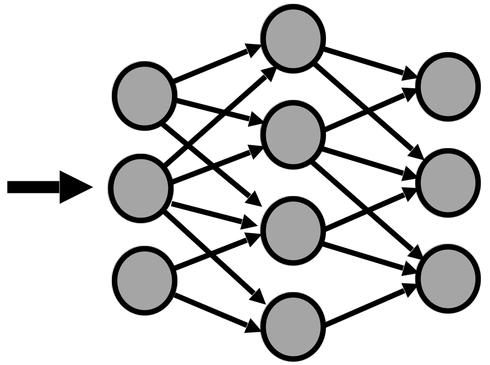


Convention Center

Generating image



Steering vehicle

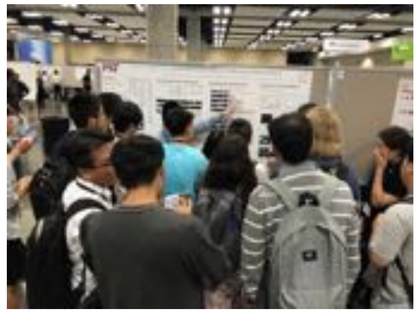


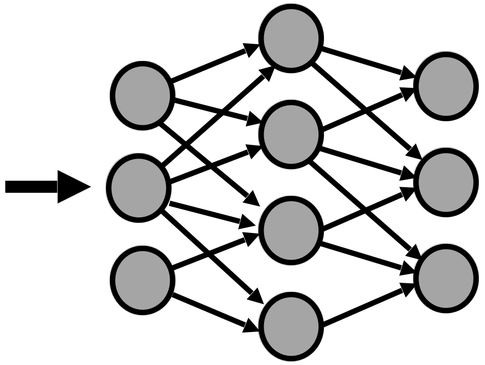
control

Human-AI Interaction in Image Classification

Model interpretability and explainability (to be given by Ruth)

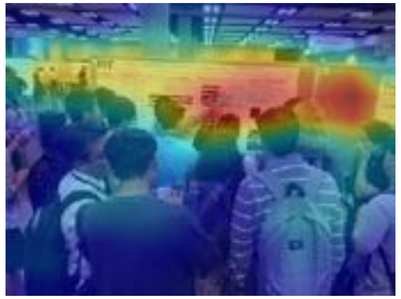
Classifying image

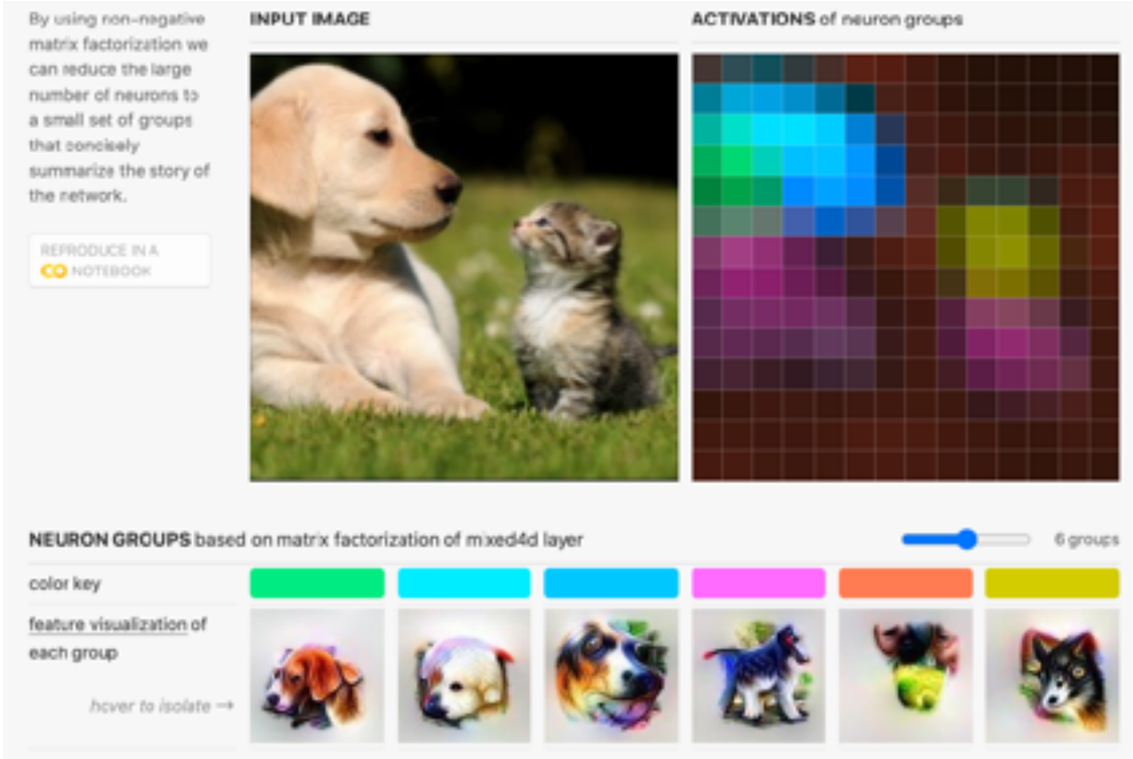




Convention Center

How we humans understand the inner workings and the prediction?

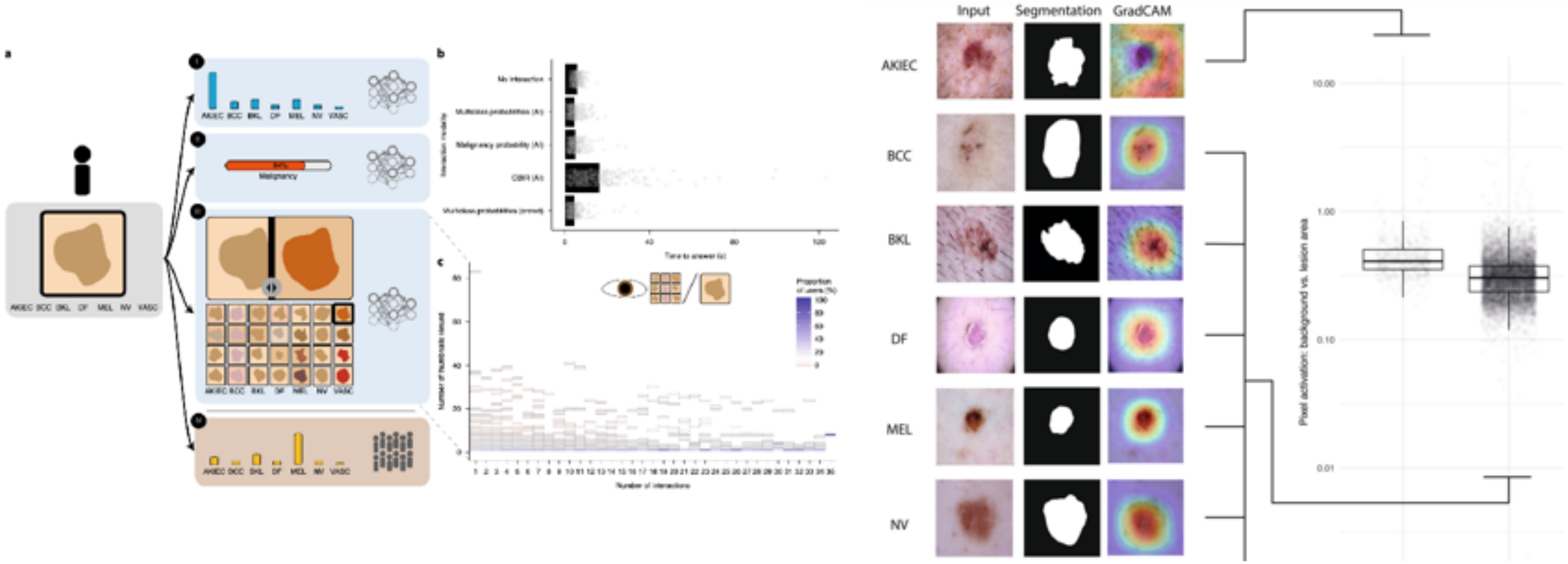




https://distill.pub/2018/building-blocks/

Human-AI Interaction in Image Classification

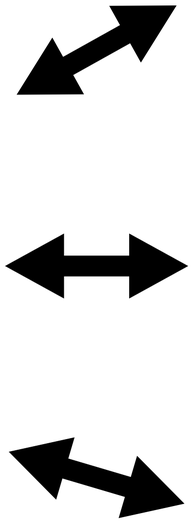
AI-assisted decision making



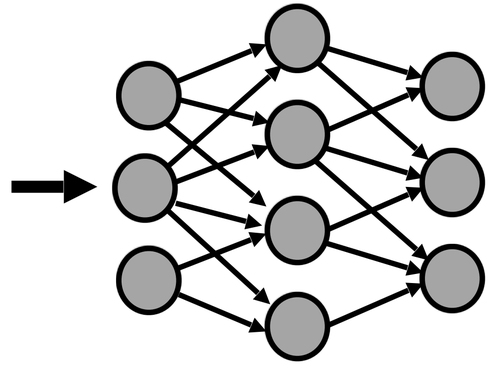
Human–AI collaboration for skin cancer recognition, Nature Medicine, 2020

https://www.nature.com/articles/s41591-020-0942-0.pdf

Human-AI Interaction in Computer Vision

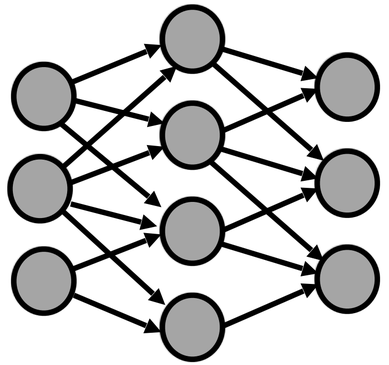


Classifying image

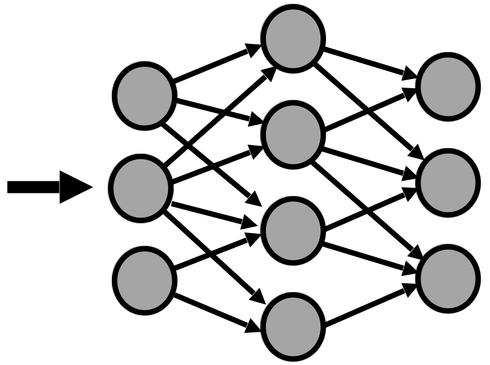


Convention Center

Generating image

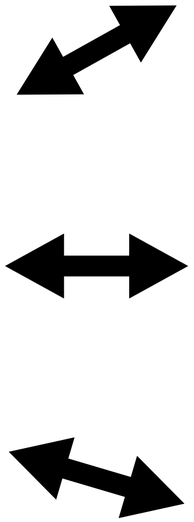


Steering machine

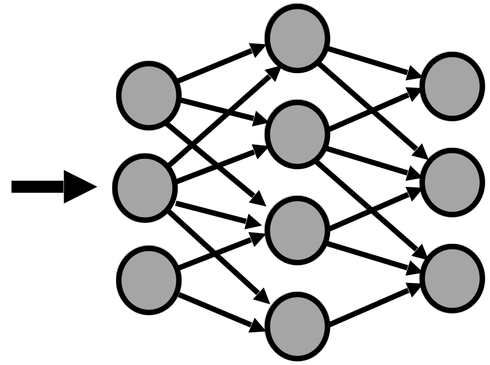


control

Human-AI Interaction in Computer Vision

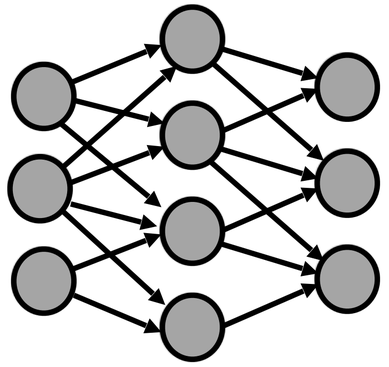


Classifying image

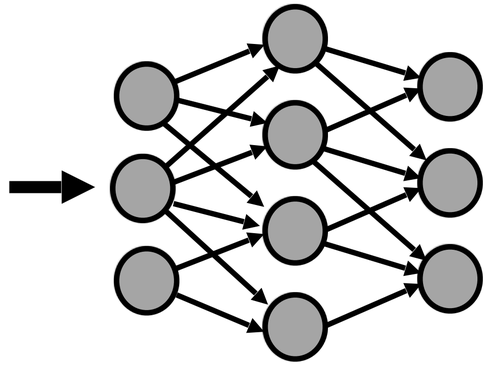


Convention Center

Generating image



Steering machine



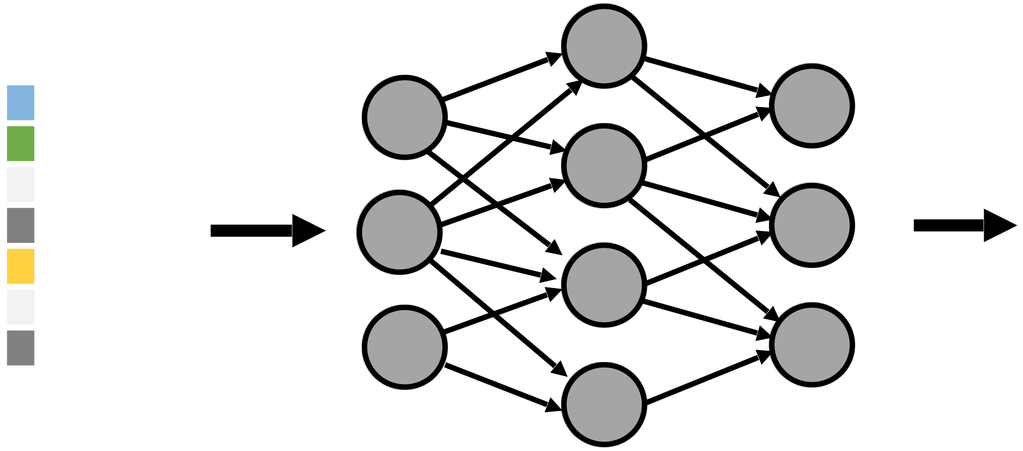
control

Human-AI Interaction in Image Generation

StyleGANv3 from Nvidia

Karras et al. Alias-free Generative Adversarial Networks. NeurIPS'21

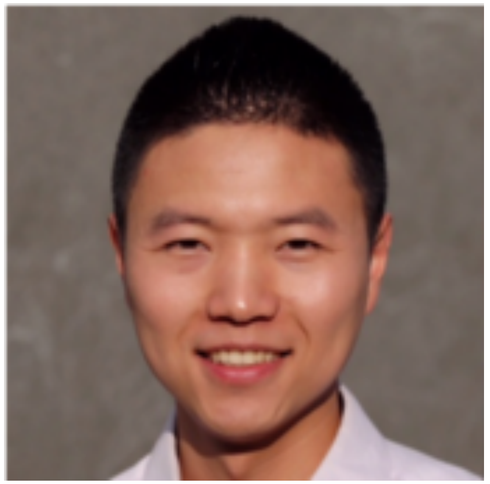
How humans use generative AI to edit and create?



Deep generative model

AI-assisted content editing

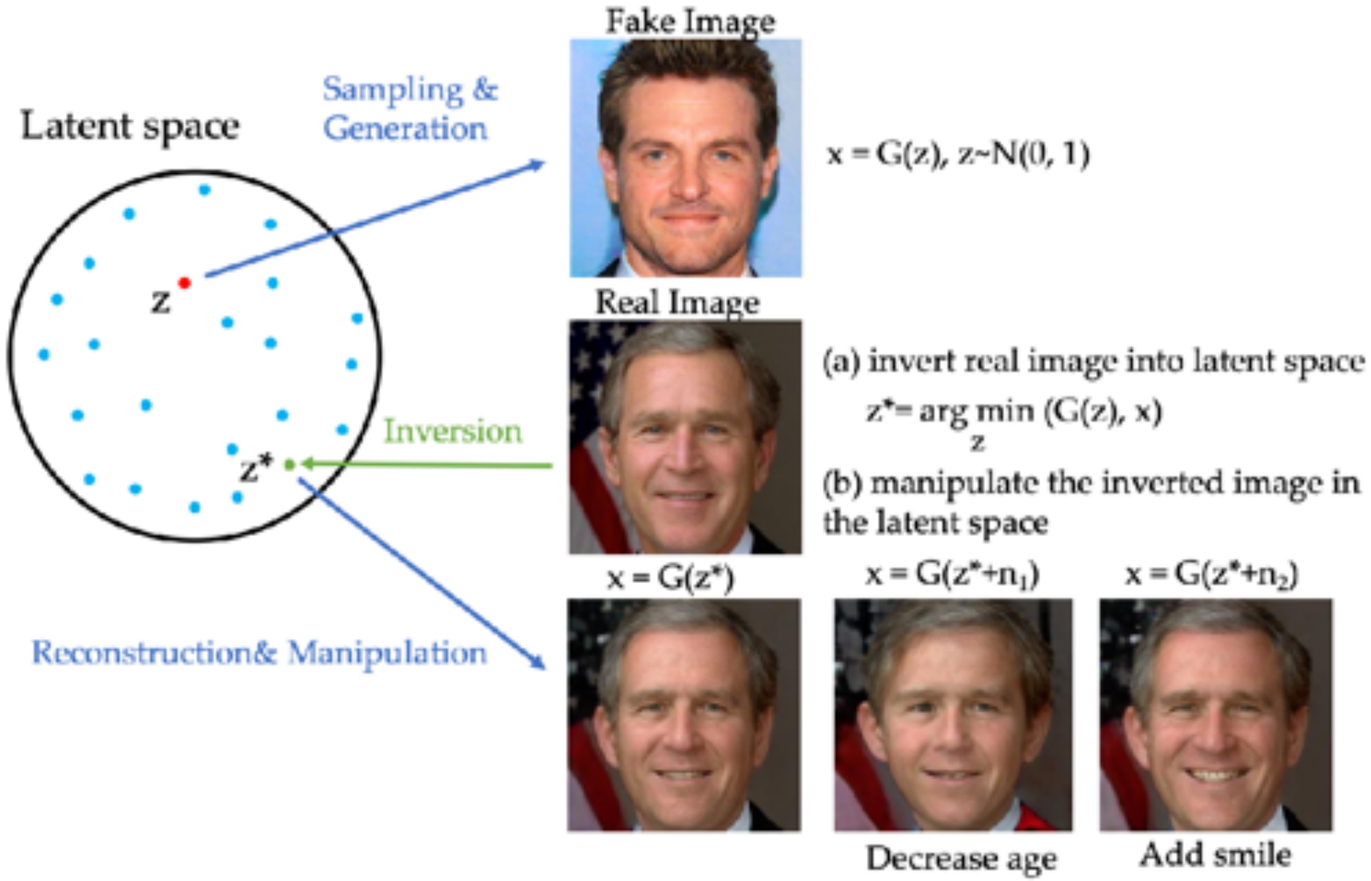
Edit my own photo?



Zhu, Shen, Zhao, Zhou. In-Domain GAN Inversion for Real Image Editing. ECCV 2020

AI-assisted content editing

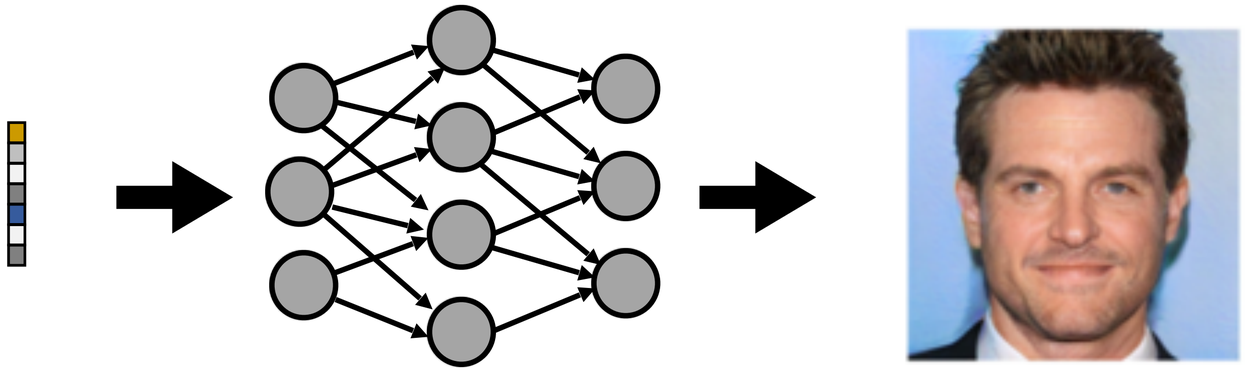
GAN inversion: Inverting real image into latent space



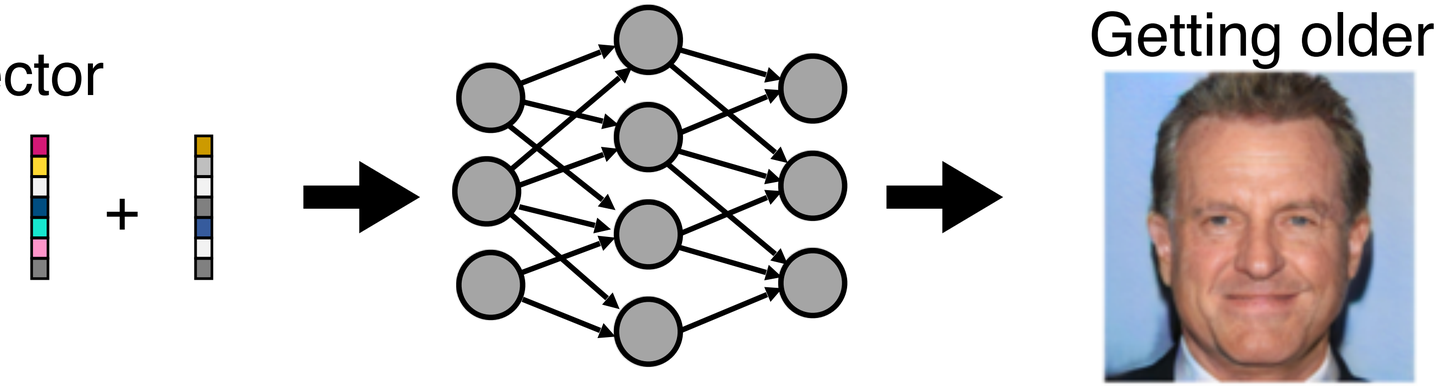
Xia, et al. GAN Inversion: a survey, TPAMI. https://arxiv.org/pdf/2101.05278.pdf

AI-assisted content editing

InterFaceGAN, GAN steerability, HiGAN, SeFa, etc.



Age vector



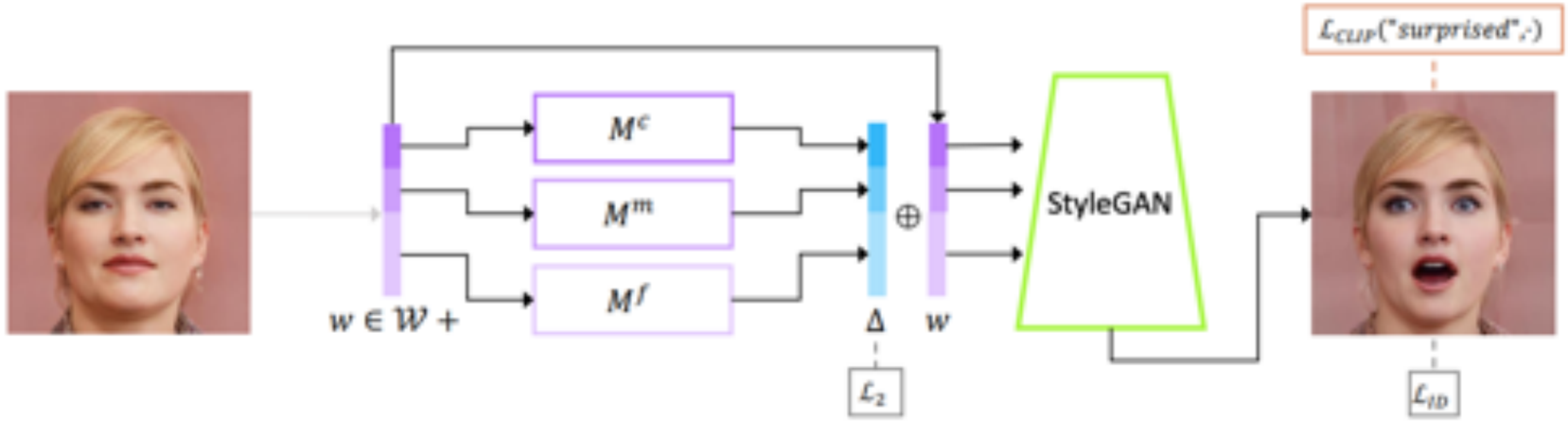
Getting older

Xia, et al. GAN Inversion: a survey, TPAMI. https://arxiv.org/pdf/2101.05278.pdf

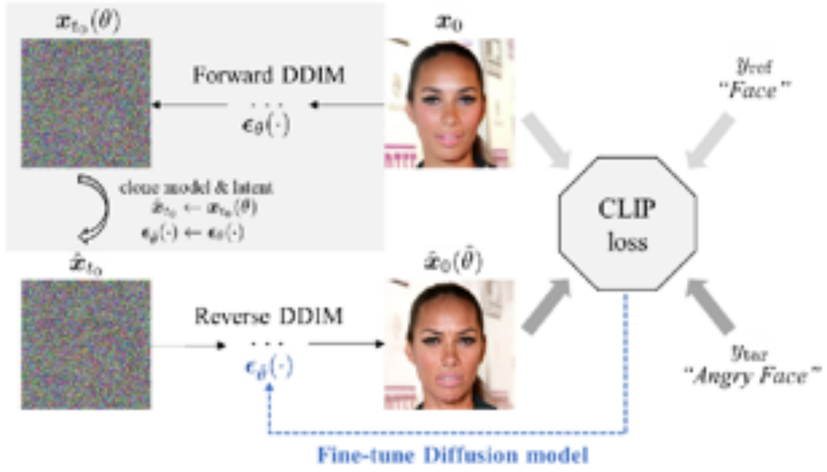
AI-assisted content editing

Text-guided image generation with CLIP model

StyleCLIP (Patashnick et al. ICCV 2021)



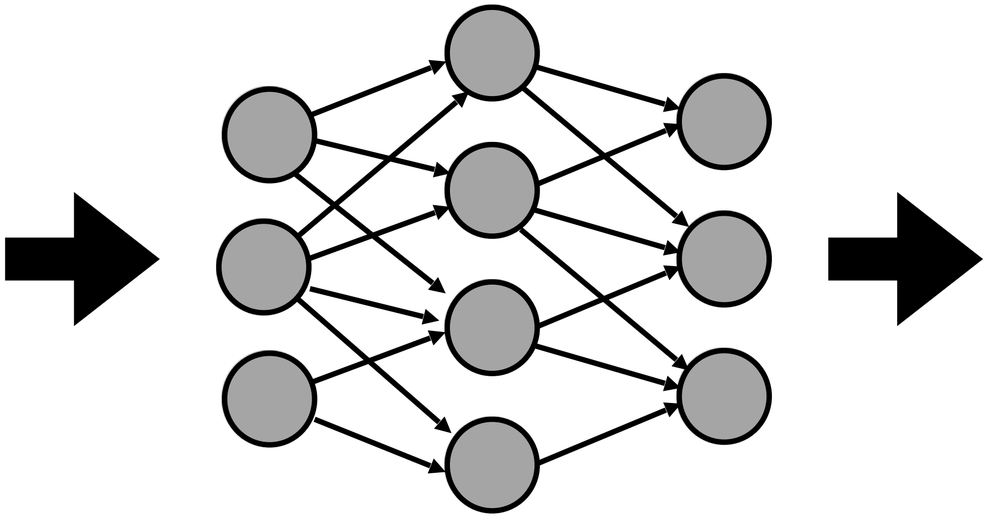
DiffusionCLIP (Kim et al. CVPR 2022)



OpenAI DALL.E 2: Text2Img

Human intent is conveyed through text

An astronaut riding a horse in a photorealistic style

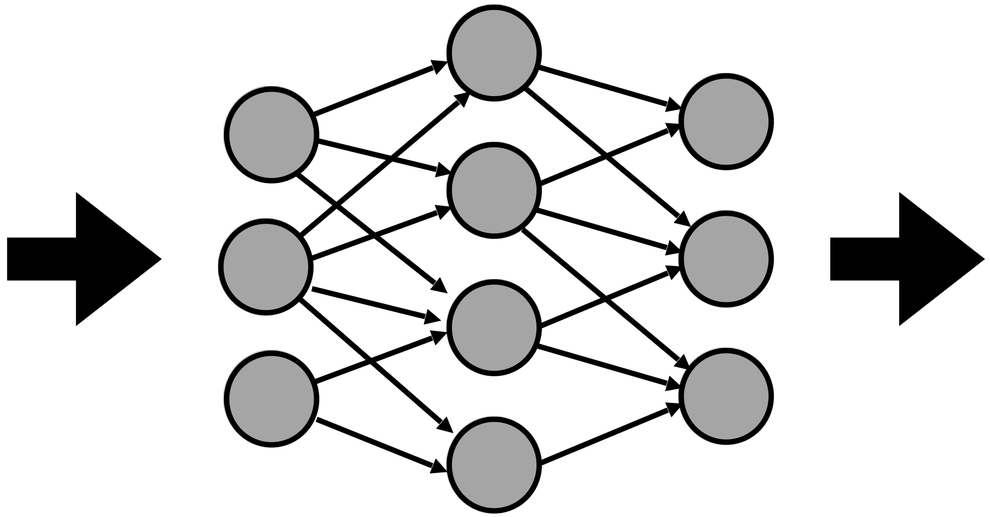


https://openai.com/dall-e-2/

OpenAI DALL.E 2: Text2Img

Human intent is conveyed through text

An astronaut riding a horse as a pencil drawing

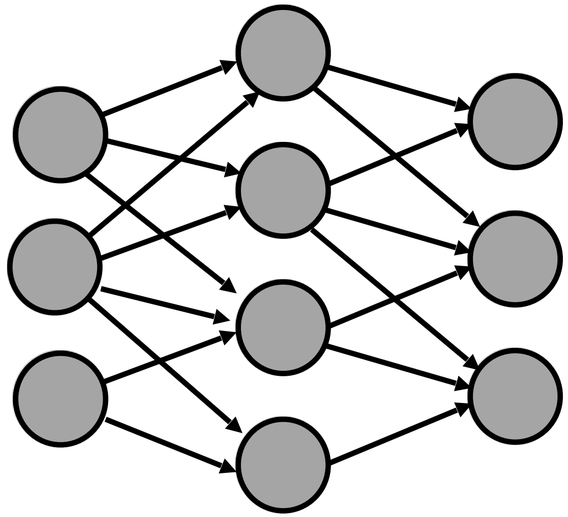


https://openai.com/dall-e-2/

OpenAI DALL.E 2: Text2Img

Human intent is conveyed through text

An astronaut
playing basketball
with cats in space
in a watercolor style

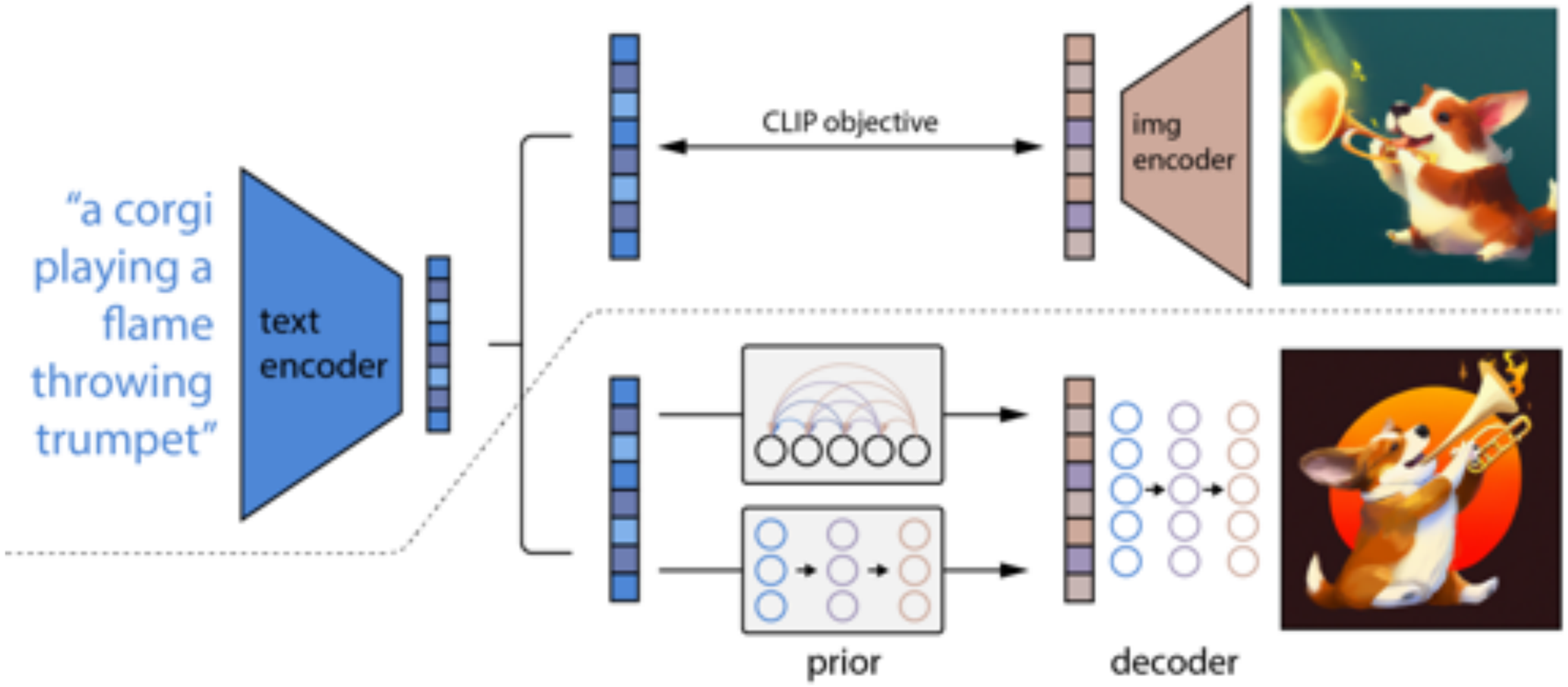


https://openai.com/dall-e-2/

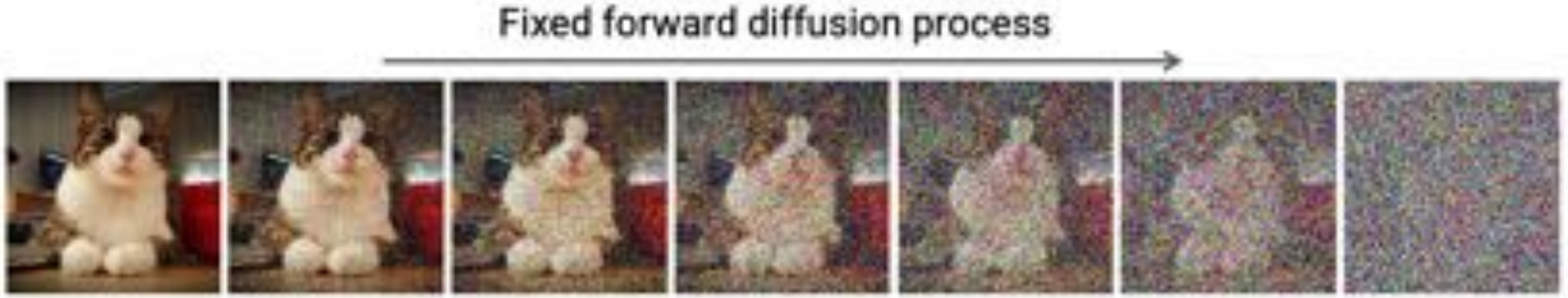
Text-guided Image Generation

Human intent is conveyed through text

CLIP + Diffusion model + Huge amount of text-image pairs (250 million pairs)



Limited Human Controllability for Diffusion Models



$$p_\theta^{(t)}(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t) = \begin{cases} \mathcal{N}(f_\theta^{(1)}(\boldsymbol{x}_1), \sigma_1^2 \boldsymbol{I}) & \text{if } t = 1 \\ q_\sigma(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t, f_\theta^{(t)}(\boldsymbol{x}_t)) & \text{otherwise,} \end{cases}$$

Classifier-guided diffusion sampling

Input: class label y, gradient scale s
$x_T \leftarrow$ sample from $\mathcal{N}(0, \mathbf{I})$
for all t from T to 1 **do**
  $\mu, \Sigma \leftarrow \mu_\theta(x_t), \Sigma_\theta(x_t)$
  $x_{t-1} \leftarrow$ sample from $\mathcal{N}(\mu + s\Sigma \nabla_{x_t} \log p_\phi(y|x_t), \Sigma)$
end for
return x_0

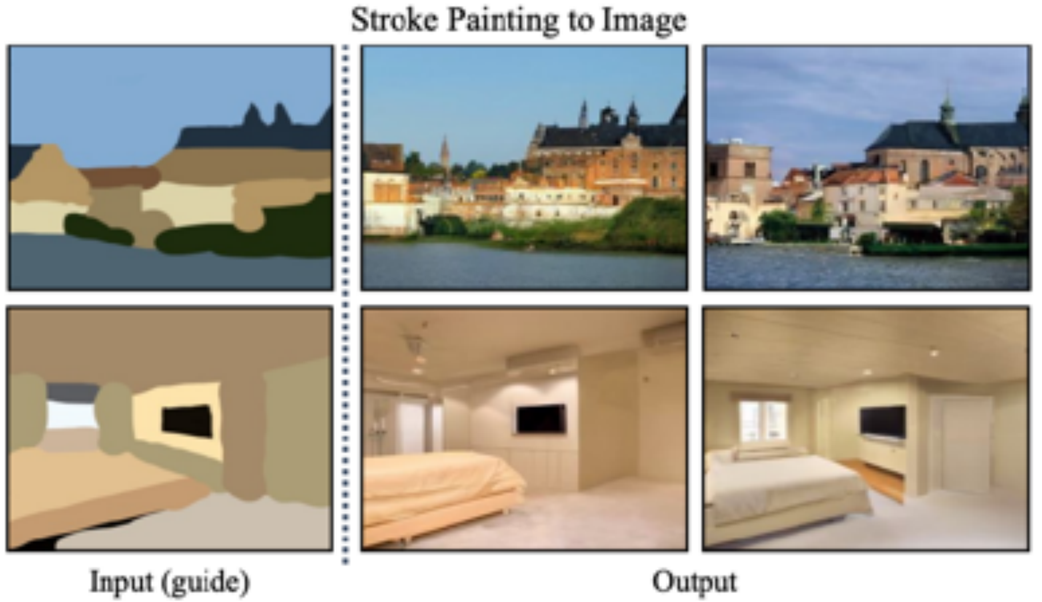
classifier's gradient

Result on an unlabeled diffusion model

https://arxiv.org/pdf/2105.05233.pdf

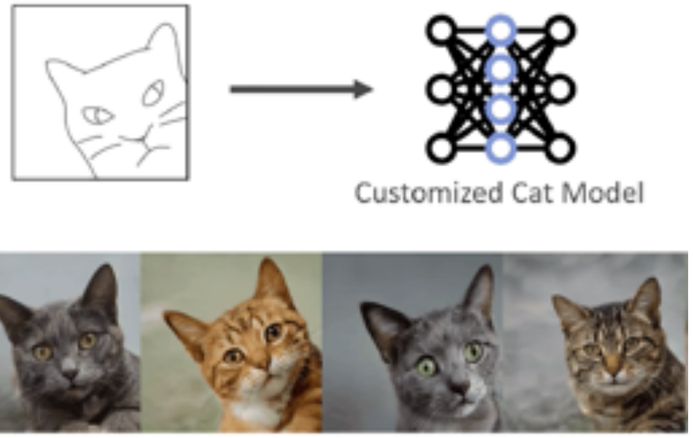
Other ways to convey human intent to generative models

GauGAN by Nvidia



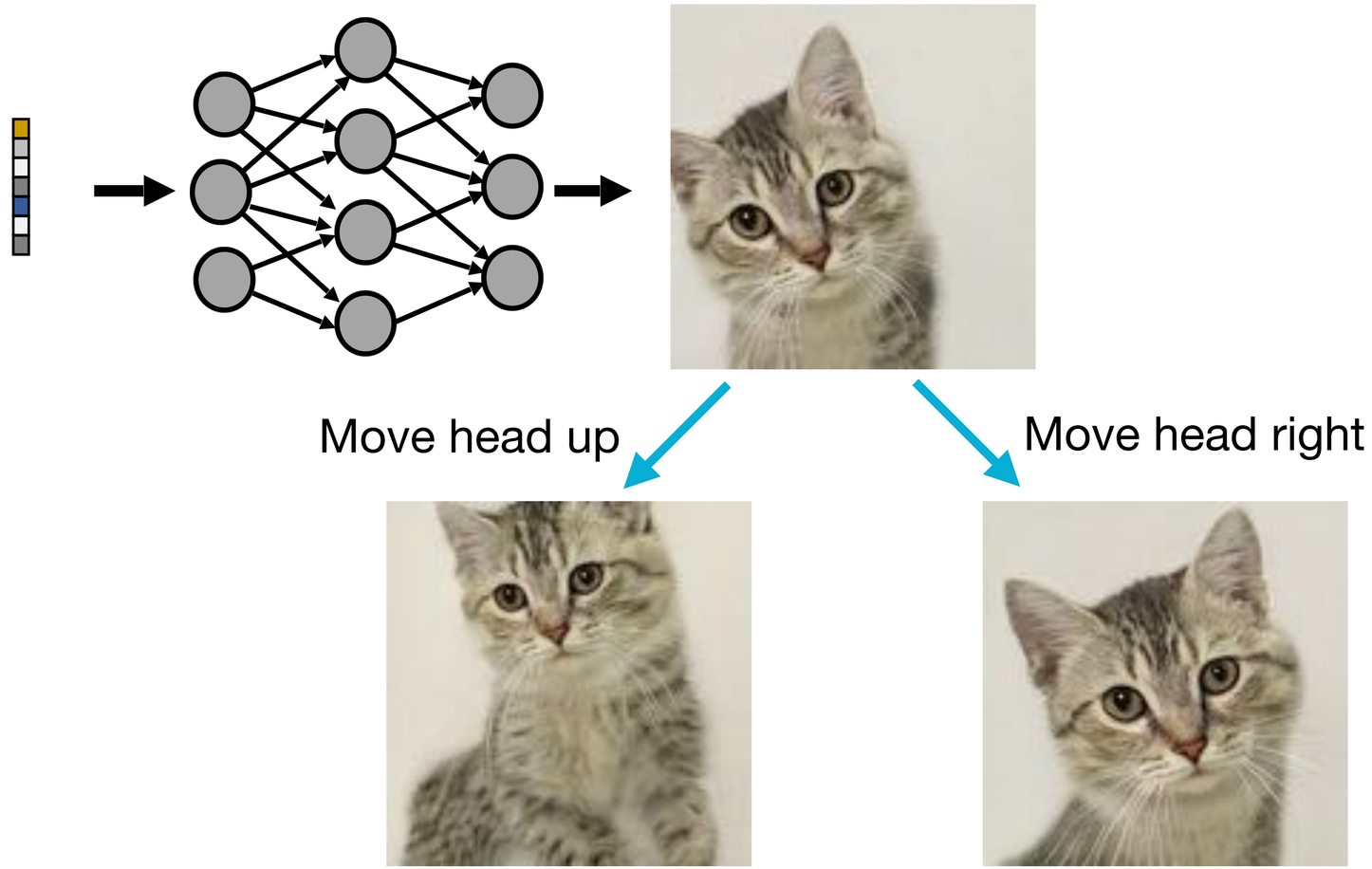
SDEdit. Meng et al. ICLR'22

Scene-Based Text-to-Image Generation with Human Priors
https://arxiv.org/pdf/2203.13131.pdf



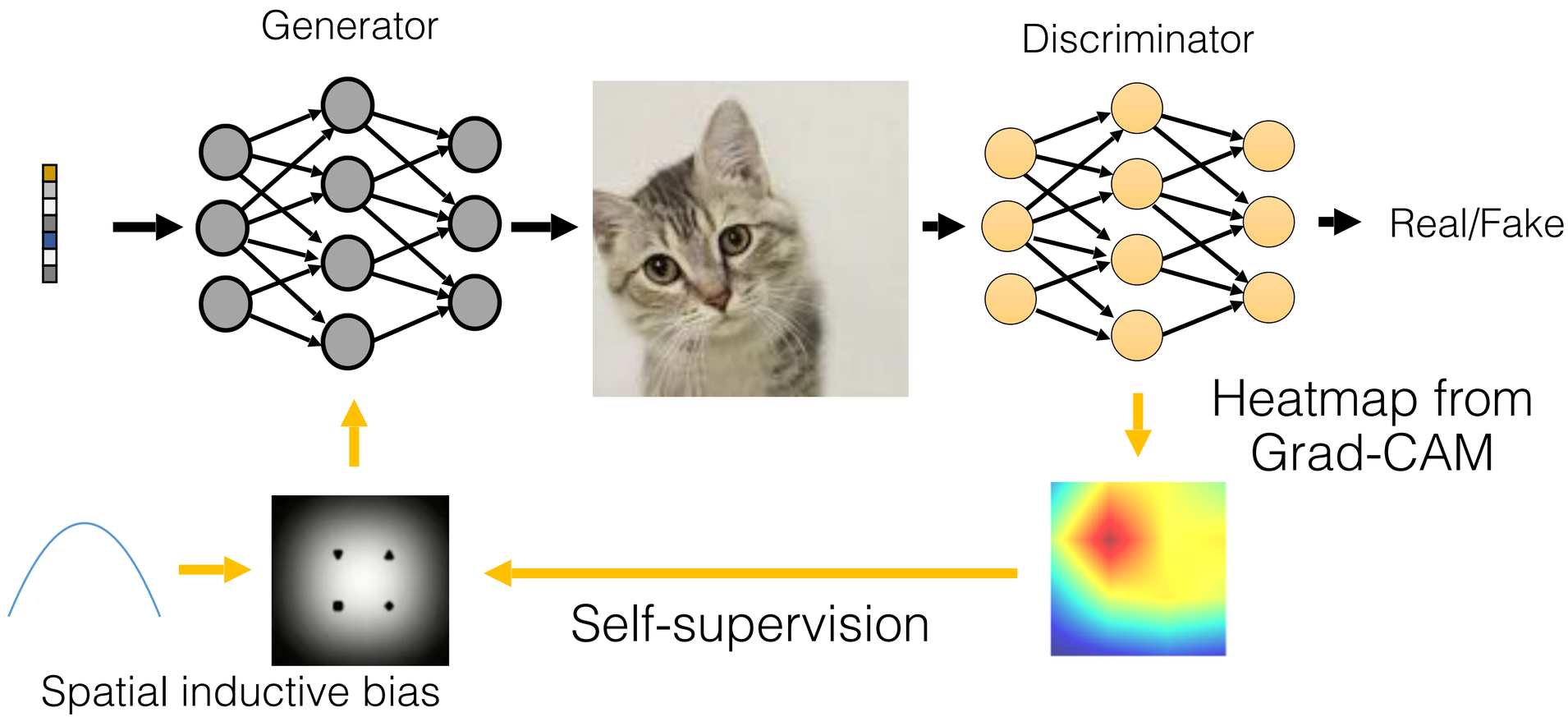
Wang, Bau, Zhu. ICCV'21

Spatial Control of Image Generation

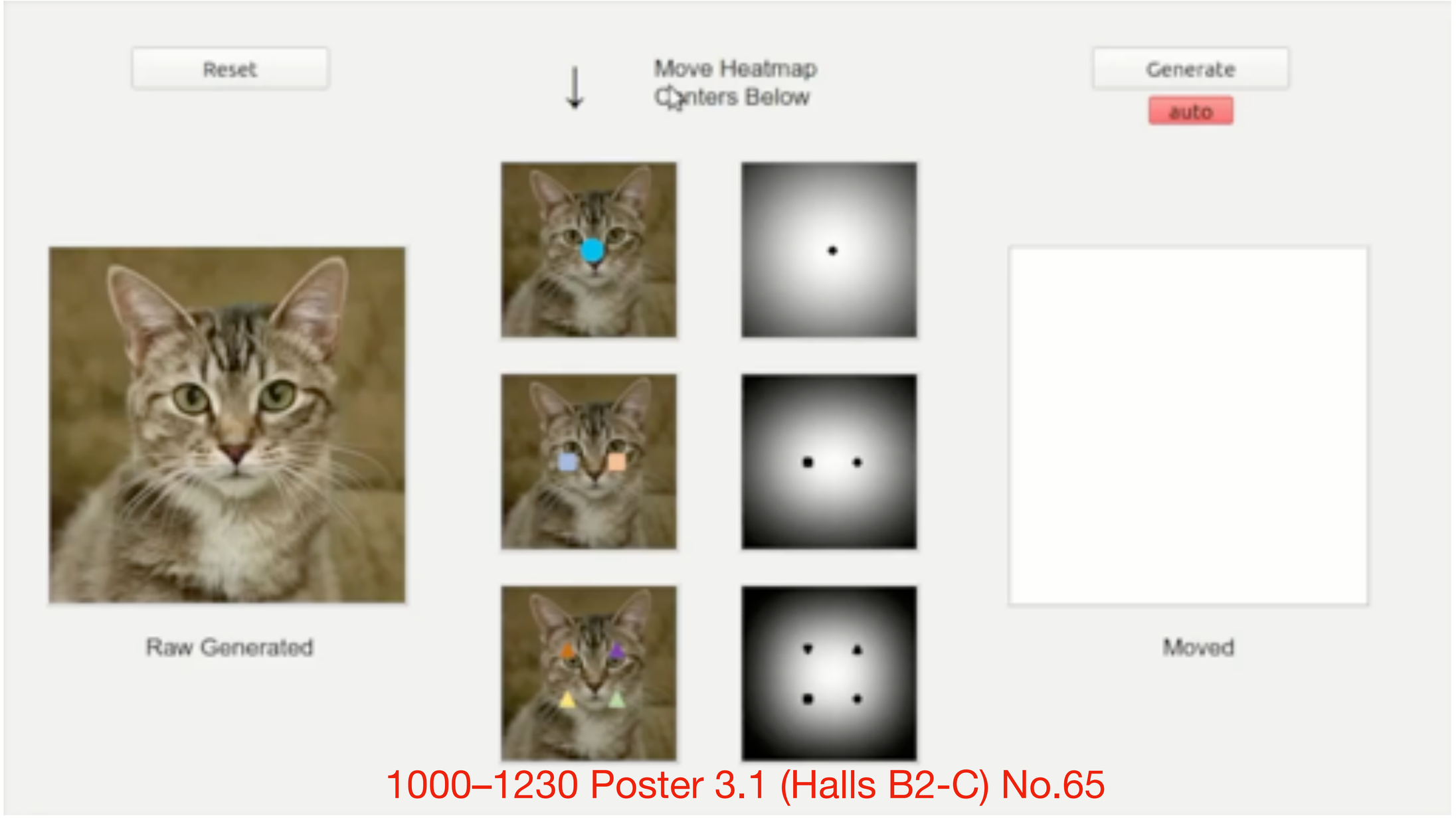


Spatial Control of Image Generation

Self supervision of spatial keypoints in the generator

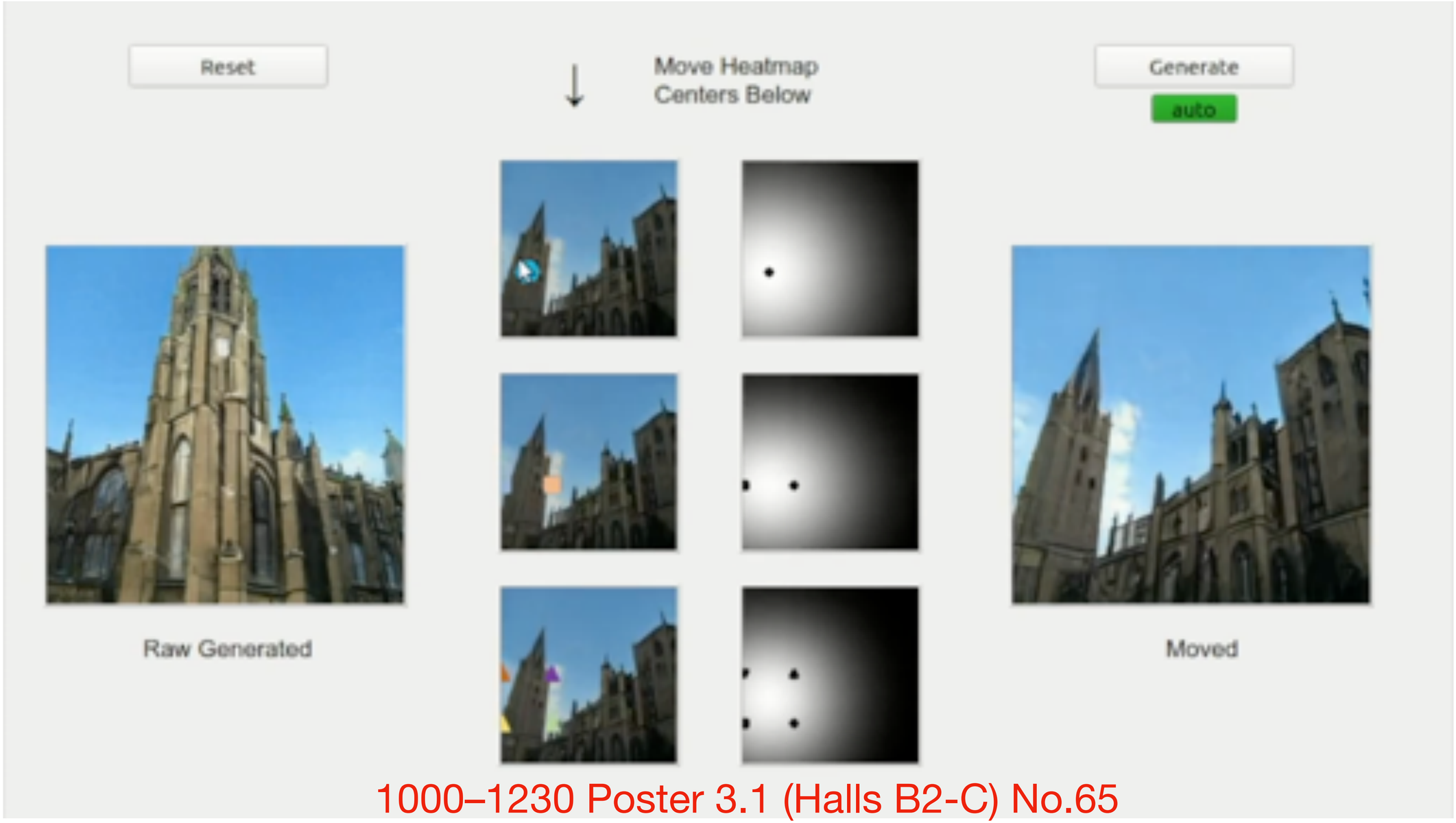


Wang et al. Improving GAN Equilibrium by Raising Spatial Awareness. CVPR'22



1000–1230 Poster 3.1 (Halls B2-C) No.65

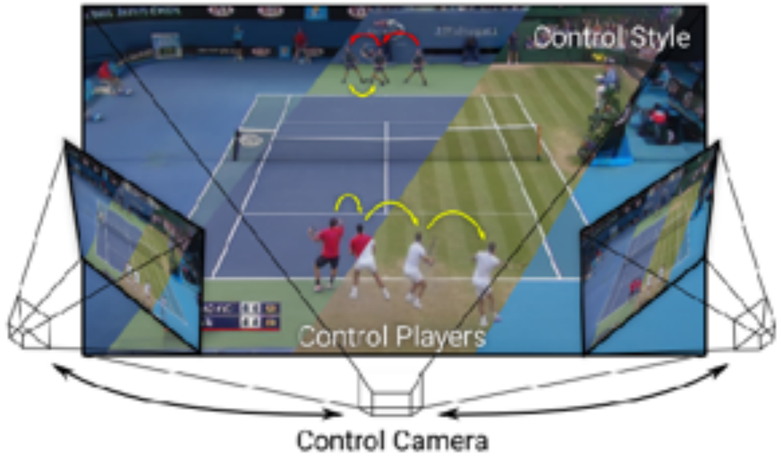
Wang et al. Improving GAN Equilibrium by Raising Spatial Awareness. CVPR'22

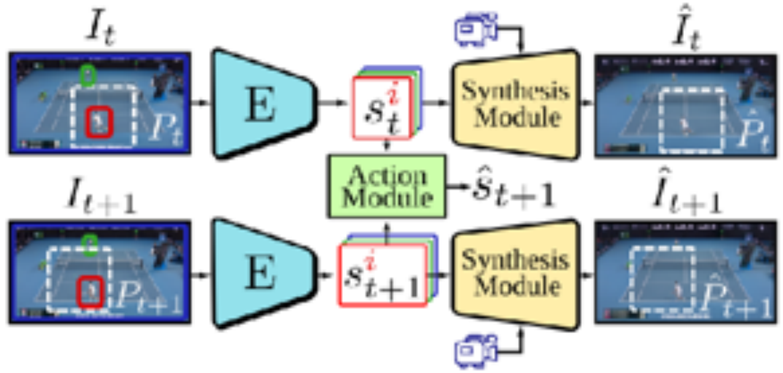


1000–1230 Poster 3.1 (Halls B2-C) No.65

Wang et al. Improving GAN Equilibrium by Raising Spatial Awareness. CVPR'22

Playable Environment on NeRFs

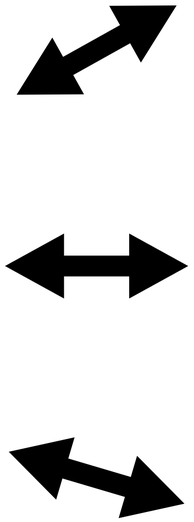




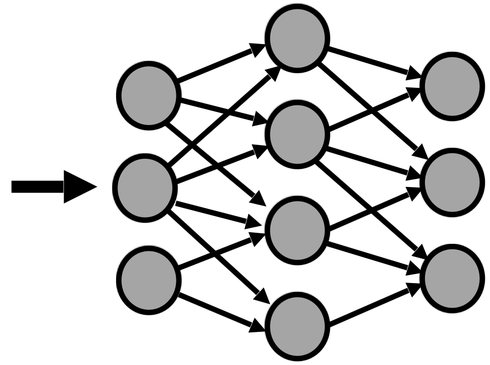
https://willi-menapace.github.io/playable-environments-website/

Menapace, et al. CVPR 2022

Human-AI Interaction in Computer Vision

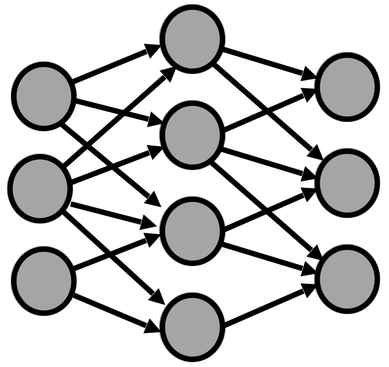


Classifying image

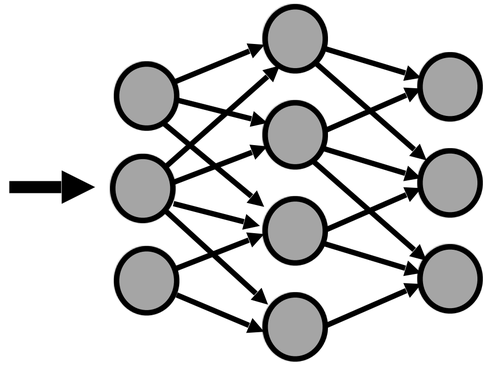


Convention Center

Generating image

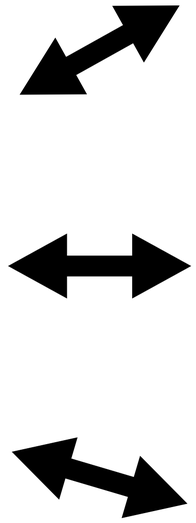


Steering machine

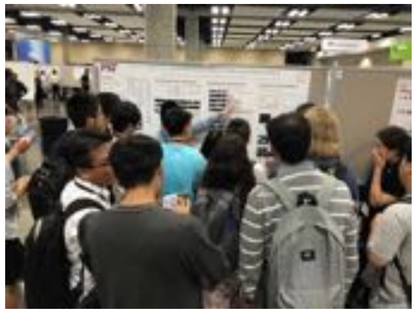


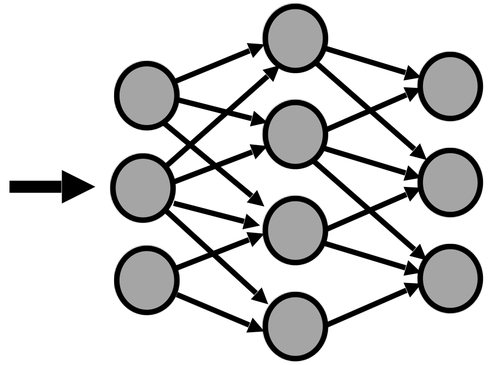
control

Human-AI Interaction in Computer Vision



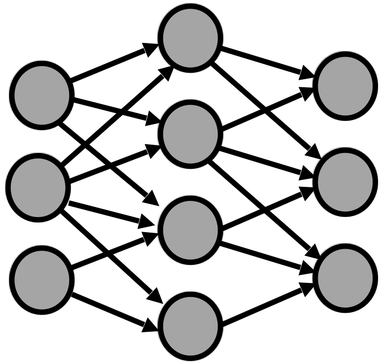
Classifying image



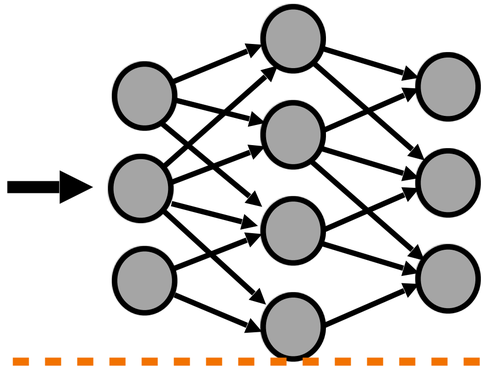


Convention Center

Generating image



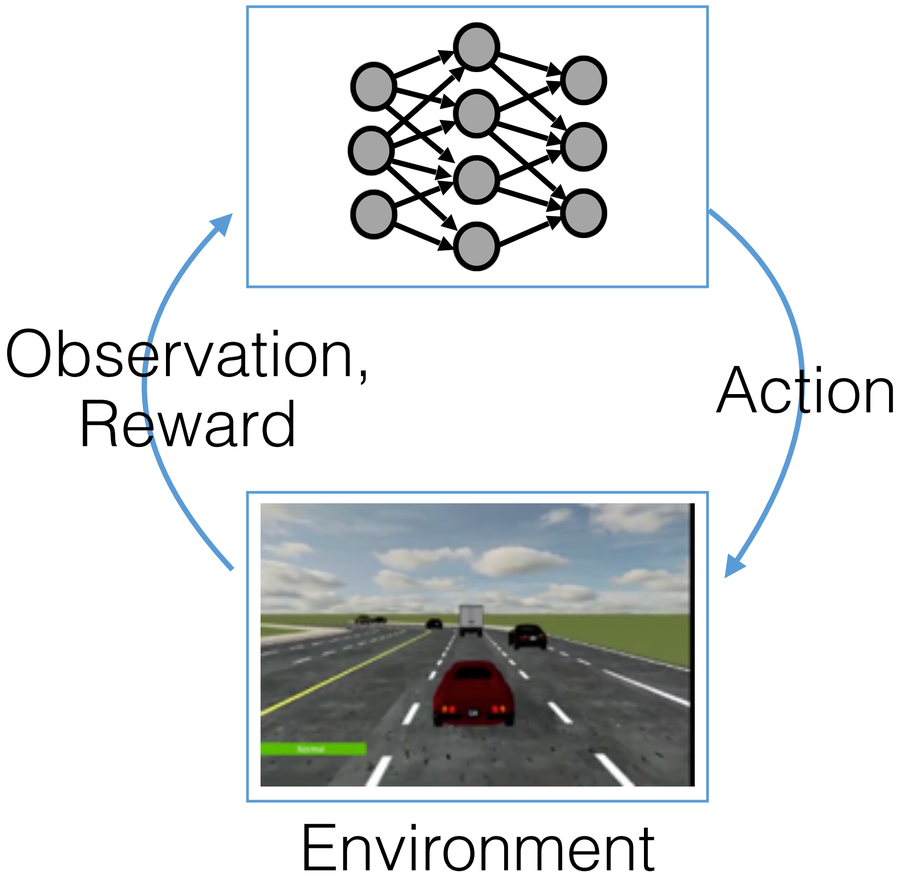
Steering machine



control

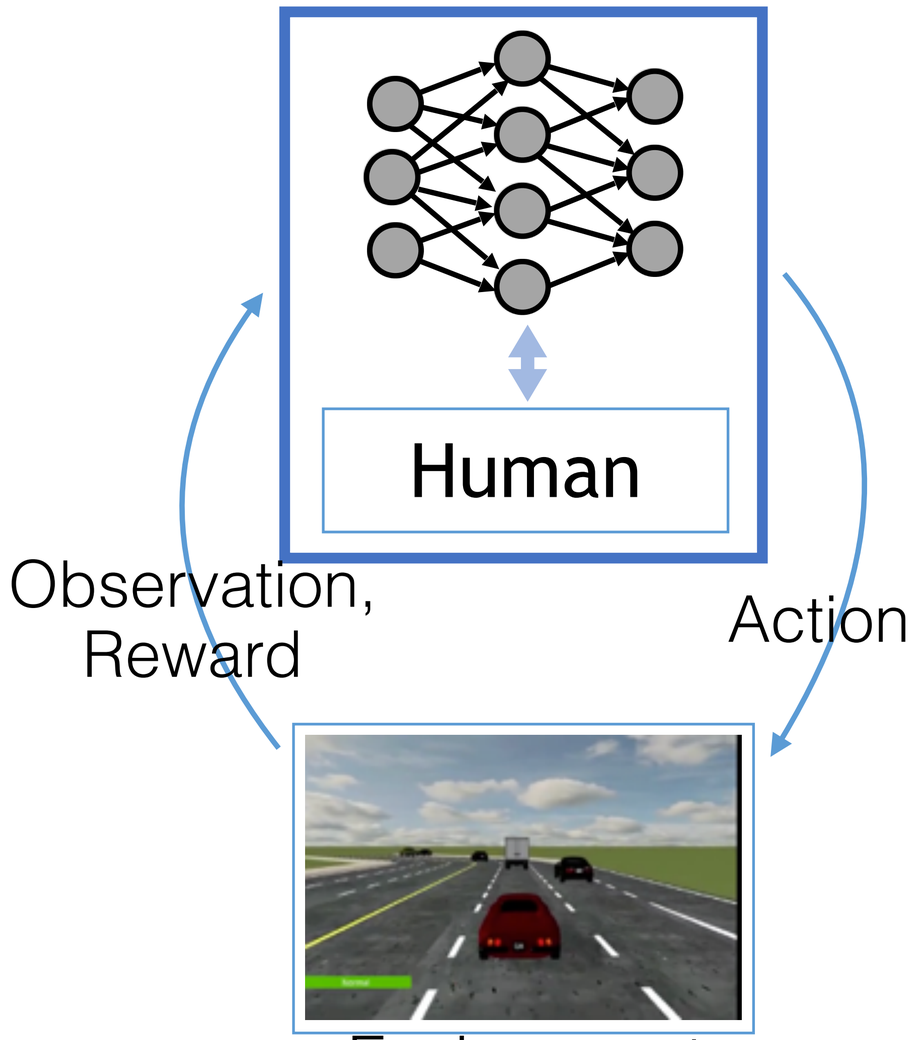
Learning a visuomotor driving policy

Reinforcement Learning



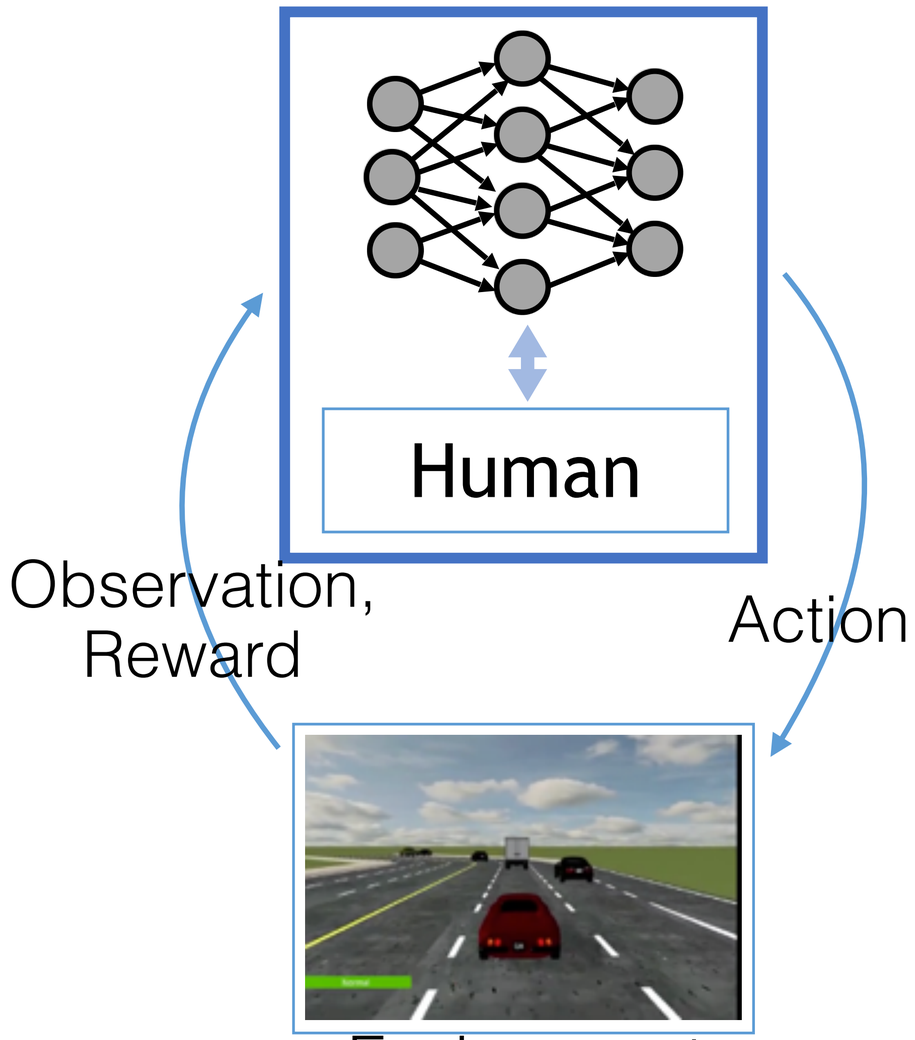
The trained agent is poorly generalizable

Human-in-the-loop learning with safety



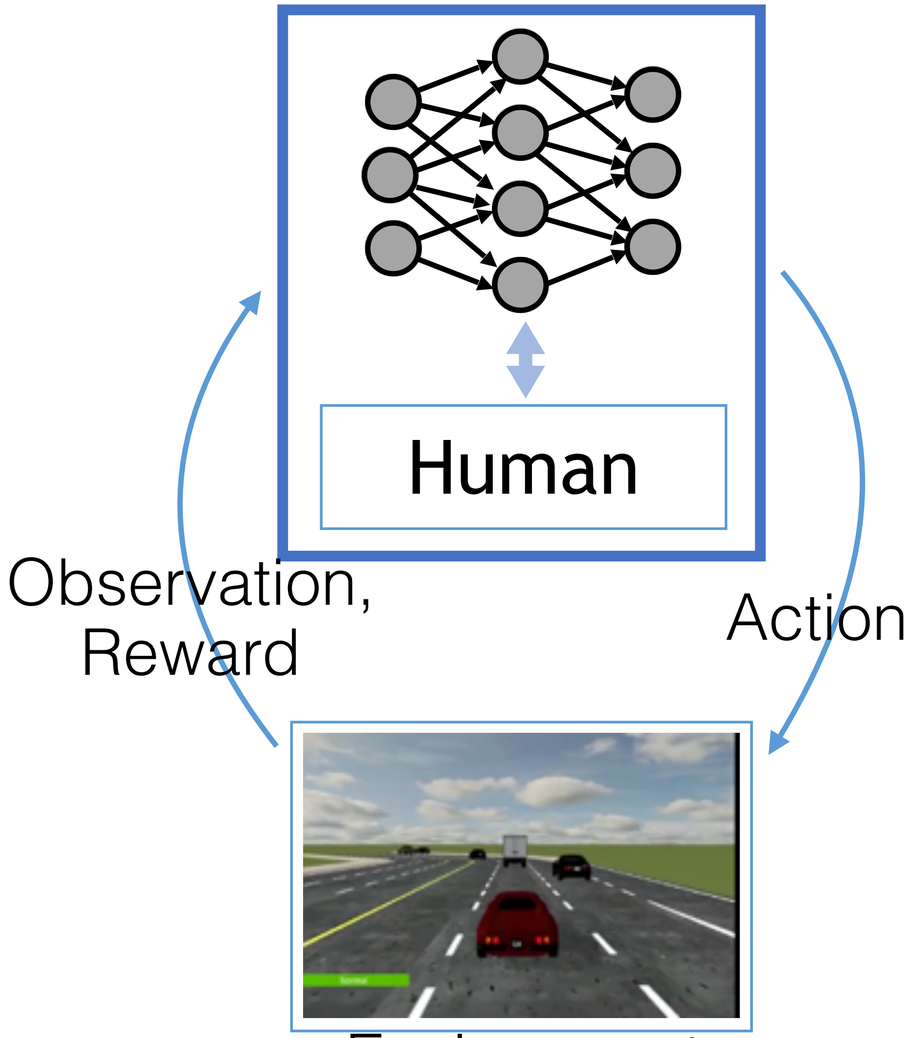
Human subject:
- To safeguard the active exploration of the learner
- To demonstrate correct actions when in danger

Human-in-the-loop learning with safety



Human subject:
- To safeguard the active exploration of the learner
- To demonstrate correct actions when in danger

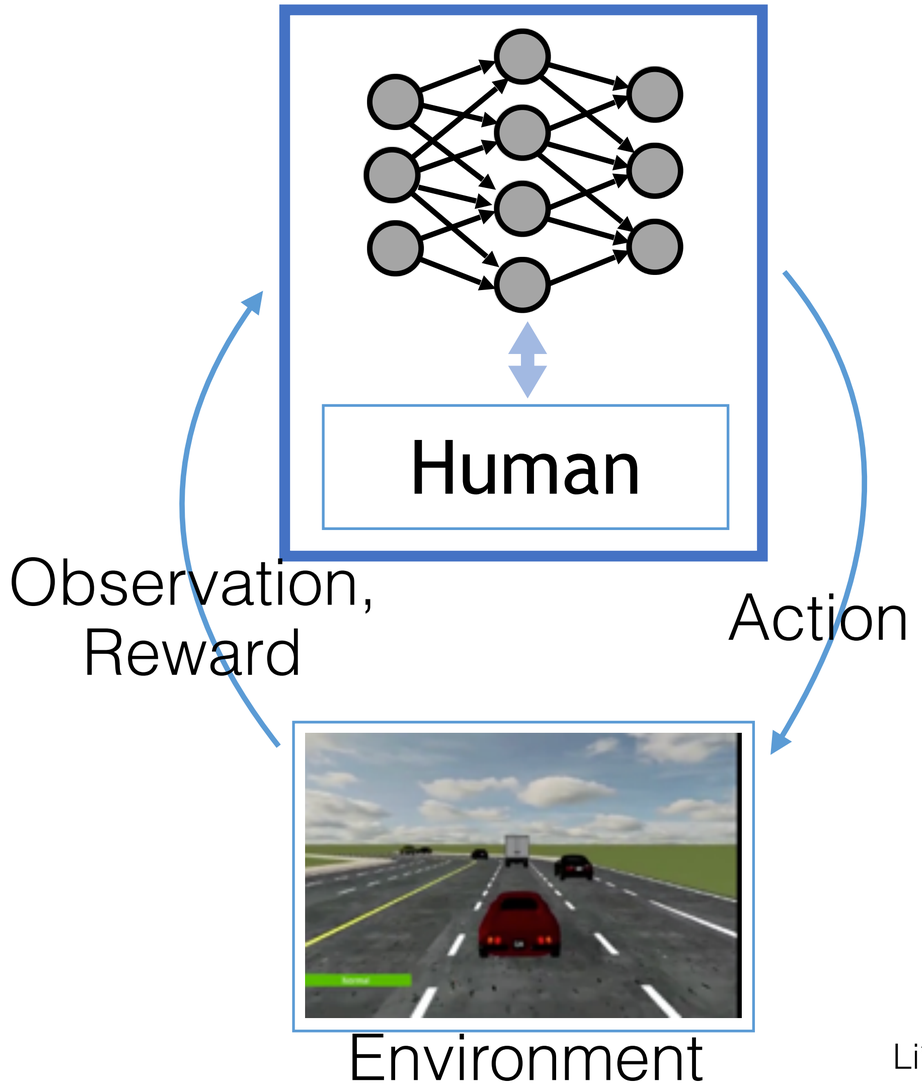
Human-in-the-loop learning



Driving in MetaDrive environment

Li*, Peng*, Zhou. Efficient Learning of Safe Driving Policy via Human-AI Copilot Optimization. ICLR'22

Human-in-the-loop learning



Driving in CARLA environment

Li*, Peng*, Zhou. Efficient Learning of Safe Driving Policy via Human-AI Copilot Optimization. ICLR'22

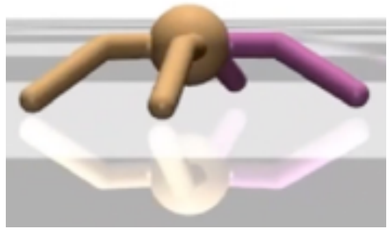
Human-AI Collaboration in Embodied AI Environment

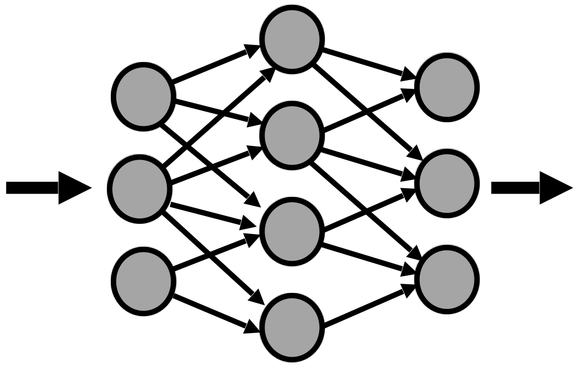
X. Puig, T. Shu, S. Li, Z. Wang, Y. Liao, J. Tenenbaum, S. Fidler, A. Torralba. Watch-And-Help: A Challenge For Social Perception and Human-AI Collaboration. In ICLR 2021

http://virtual-home.org/watch_and_help/

Human-AI Shared Autonomy

Reinforcement Learning agent

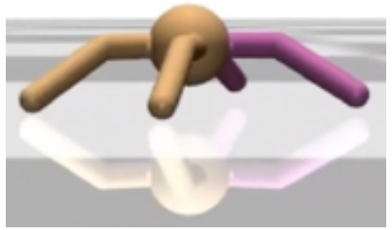


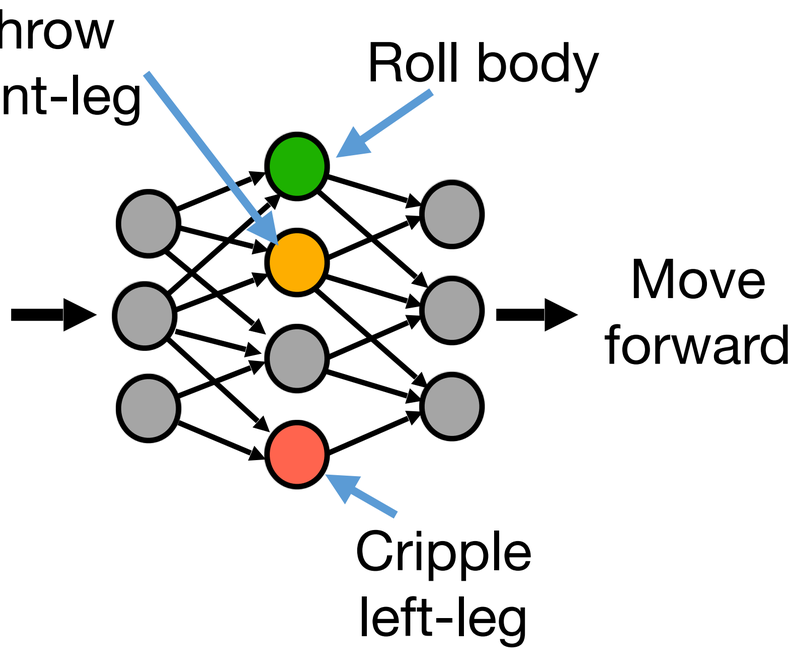


Move
forward

Majoco-Ant trained by SAC

Human-AI Shared Autonomy





Policy Dissection

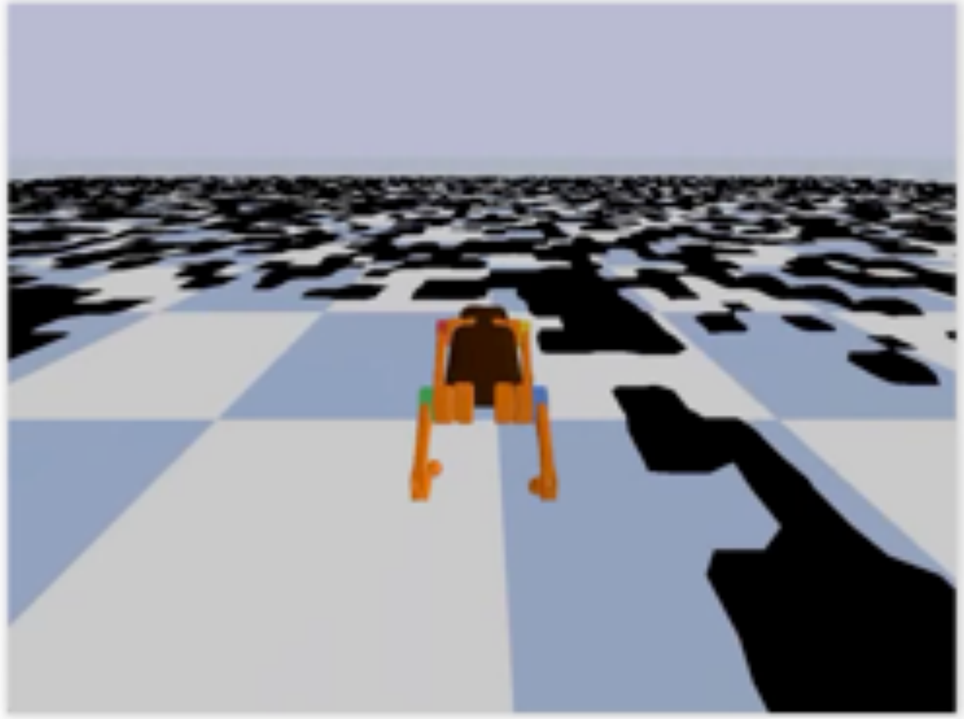
Human-Machine Teaming (HMT) on shared autonomy

Forward

Li, Peng, Wu, Feng, Zhou. Human-AI Shared Control via Frequency-based Policy Dissection. arXiv:2206.00152

Human-AI Shared Autonomy through Policy Dissection

The legged robot is trained to move forward by PPO

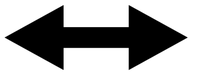


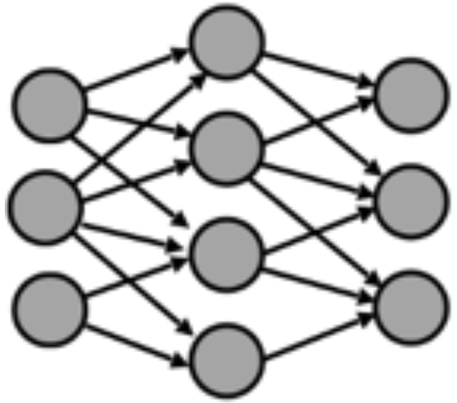
Before Training

Li, Peng, Wu, Feng, Zhou. Human-AI Shared Control via Frequency-based Policy Dissection. arXiv:2206.00152

Human-AI Interaction in Computer Vision

Synergy between machine and human intelligences





Build trustworthy relation with your AI!

A reading list of HAI Interaction: https://github.com/human-centeredAI/awesomeHAI