Defense of Master‘s Thesis

Exploring Predator-Prey Dynamics from Videos using Generative Adversarial Imitation Learning

Jannik Wirtheim

Konstanz, 23.02.2026

Motivation

Reference: „The hunt from above“ – Angela Albi : https://www.campus.uni-konstanz.de/uni-leben/die-kunst-der-haie

Modelling of Multi-Agent Systems

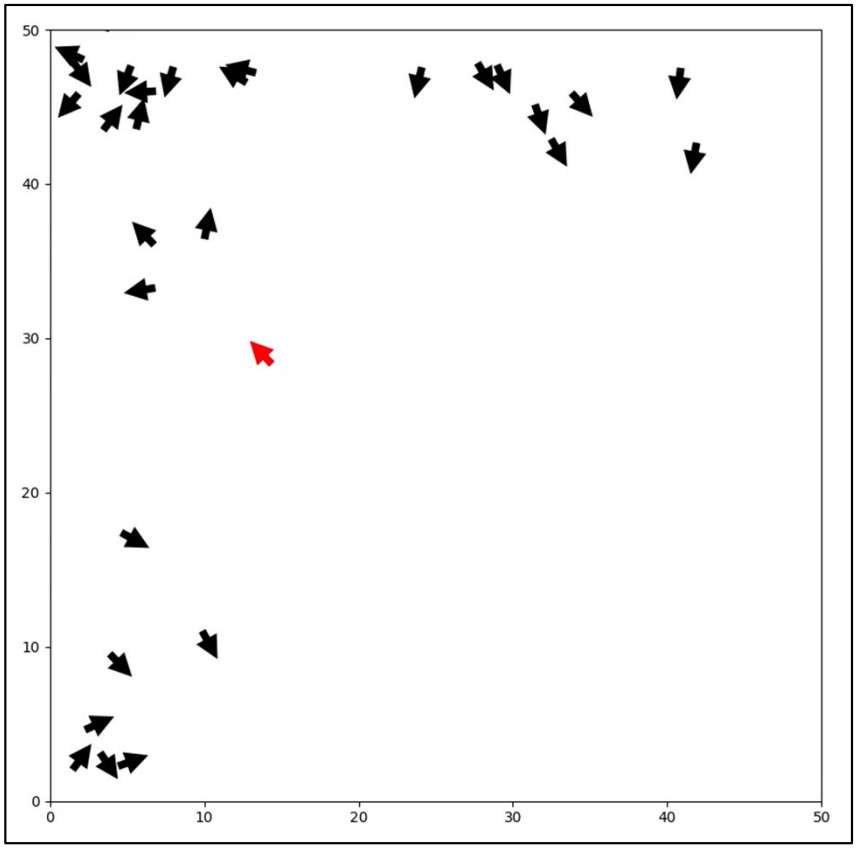
Behavior is modeled using hand-crafted interaction rules.

→ requires domain knowledge from the designer.

Hand-crafted rules

Pros/Cons:
+ Simple rules
+ Reasonable behavior
- Expert knowledge
- Oversimplified dynamics

Examples:
- Reynolds Boids (1987)
- Vicsek Model (1995)
- Couzin Model (2002)
- (Reinforcement Learning (RL))



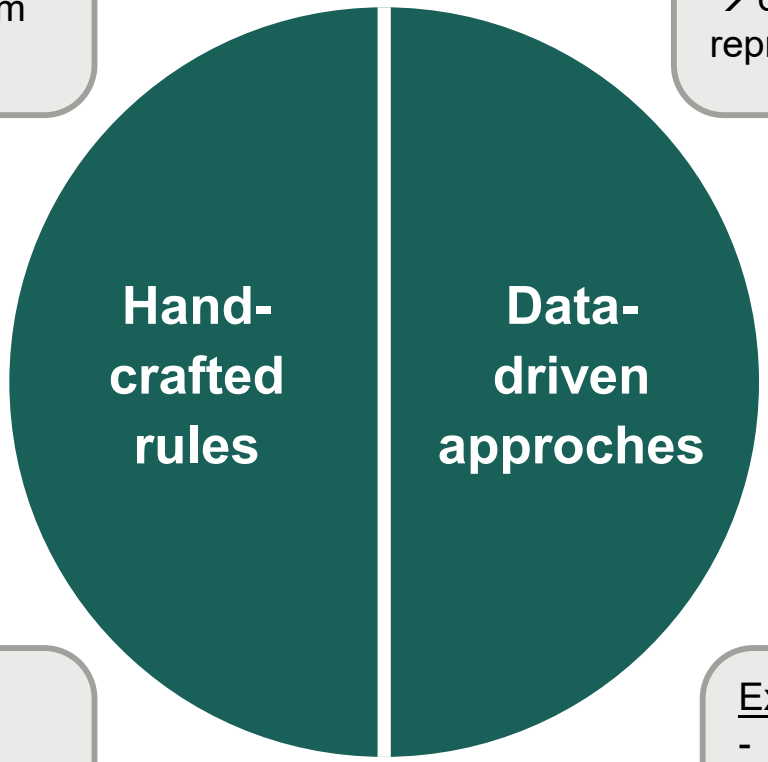
Modelling of Multi-Agent Systems

Behavior is modeled using hand-crafted interaction rules.

→ requires domain knowledge from the designer.

Behavior is learned by imitation from expert demonstrations.

→ desired behavior is implicitly represented in the data.



Pros/Cons:
+ Simple rules
+ Reasonable behavior
- Expert knowledge
- Oversimplified dynamics

Pros/Cons:
+ No manual rule design
+ Complex dynamics
- Data dependency
- Computational problems

Examples:
- Reynolds Boids (1987)
- Vicsek Model (1995)
- Couzin Model (2002)
- (Reinforcement Learning (RL))

Examples:
- Behavioral Cloning (BC)
- Inverse RL
- Generative Adversarial Imitation Learning (GAIL)

Modelling of Predator-Prey Systems

Predator–prey systems represent a specific form of multi-agent systems in which heterogeneous groups pursue contrasting objectives.

Prey strategies:

- Cooperative behavior
- Coordinated motion
- Synchronized directional changes

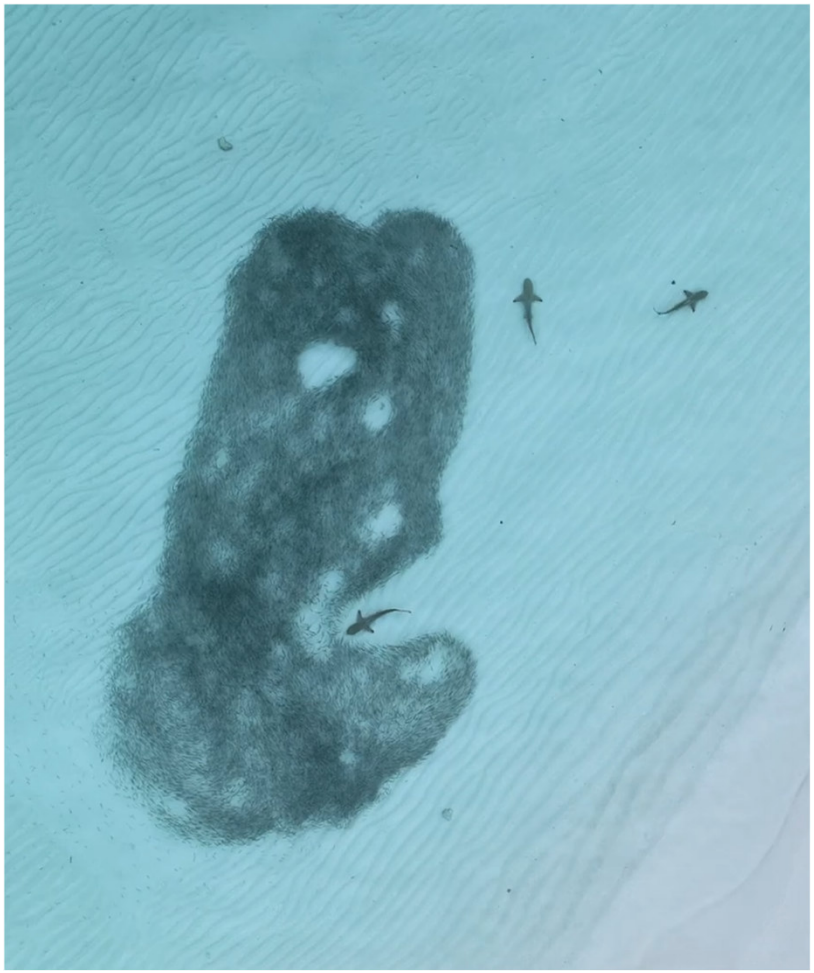
→ Confuse predator & increase survival chance

Predator strategies:

- Dispersion tactics
- Isolated prey

→ Separate individuals from swarm

Challenge: Imitation is driven by a survival-based interplay between cooperative prey and a attack-oriented predator.

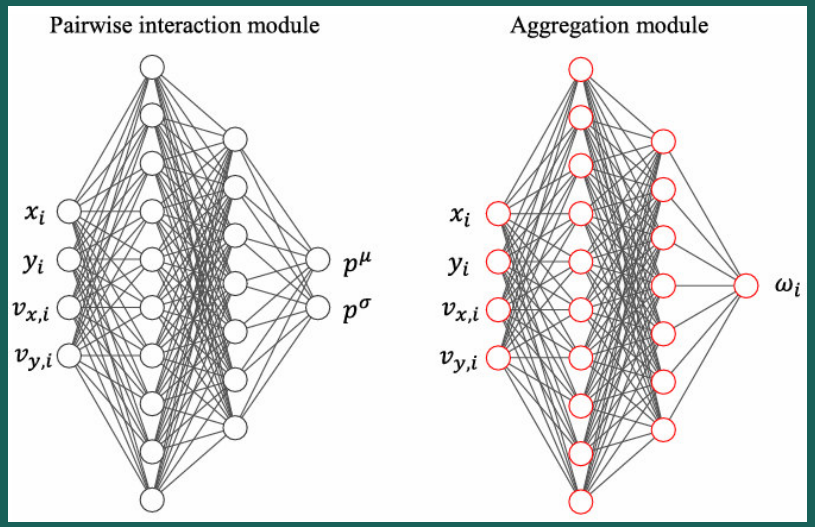


Reference: „The hunt from above“ – Angela Albi : https://www.campus.uni-konstanz.de/uni-leben/die-kunst-der-haie

Related Research

Adversarial imitation learning with deep attention network for swarm systems (Yapei Wu et al., 2025)

- GAIL with shared individual policy
- Couzin-based swarm demonstrations
- Same policy & tensor structure
- Limited to single-species imitation

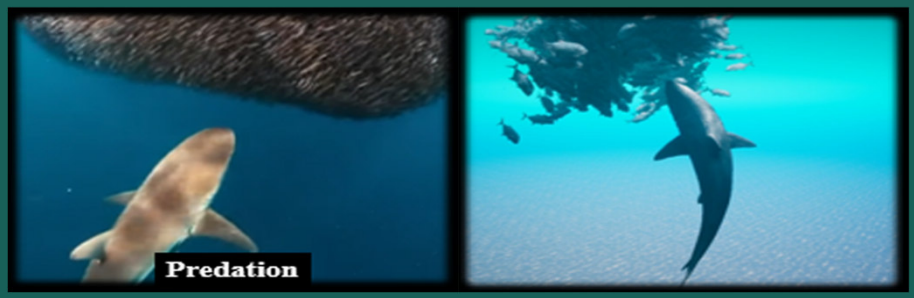


(Yapei Wu et al., 2025, p.4)

→ „Wu Paper“

CBIL: Collective Behavior Imitation Learning for Fish from Real Videos (Yifan Wu et al., 2025)

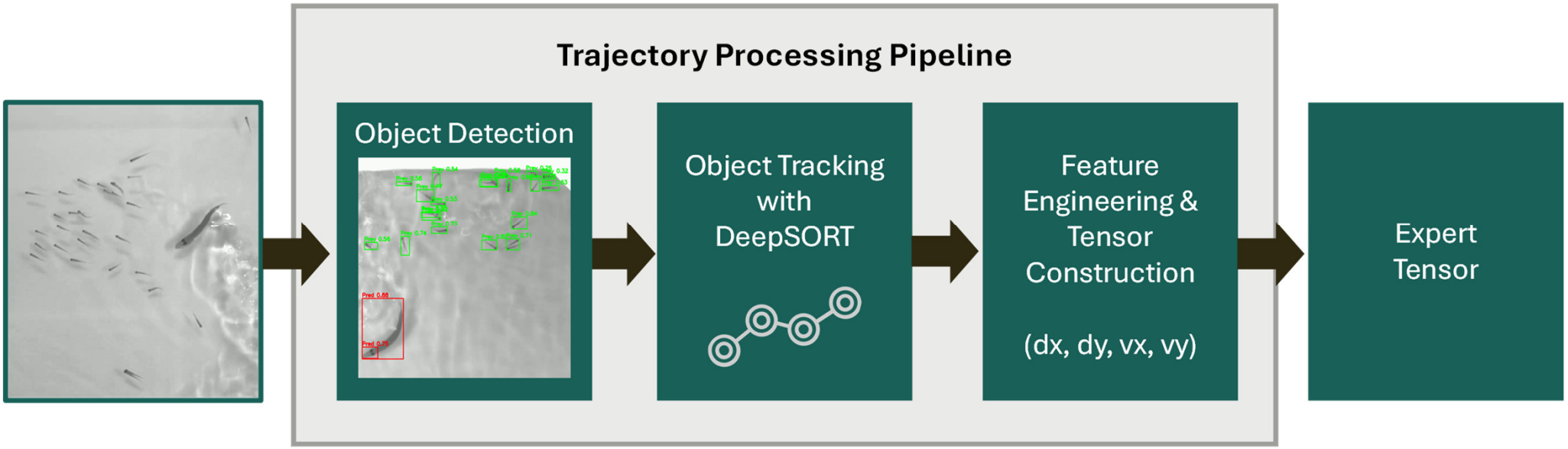
- GAIL on latent video representations
- Transition Encoder
- Feature clustering & reward shaping
- Imitated predation with multi-instance single policy



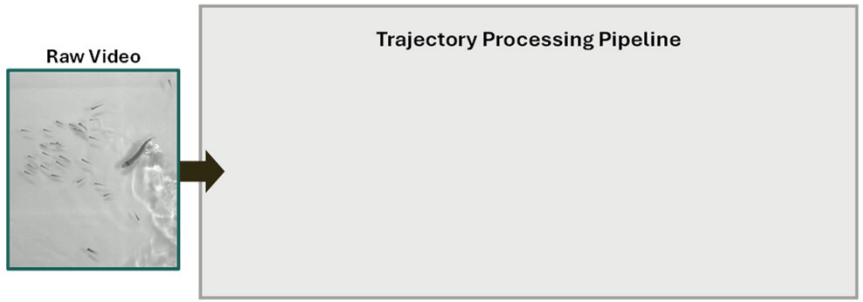
(Yifan Wu et al., 2025, p.9)

→ „CBIL Paper“

Data Collection & Processing



Data Collection & Processing



Recording of predator-prey aquarium:

- Setup & recordings by MPI
- Sunbleak (*Leucaspius delineatus*)
- Northern pike (*Esox lucius*)

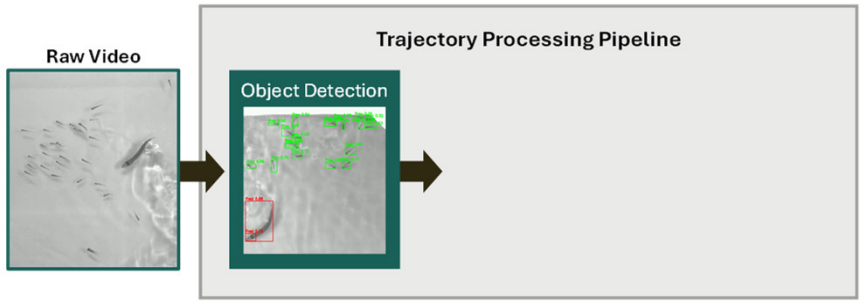
Video-related overview:

- 35 recordings (1 predator, 32 prey)
- total duration 16:42:49 h
- 151.695 frames

Predator-related overview:

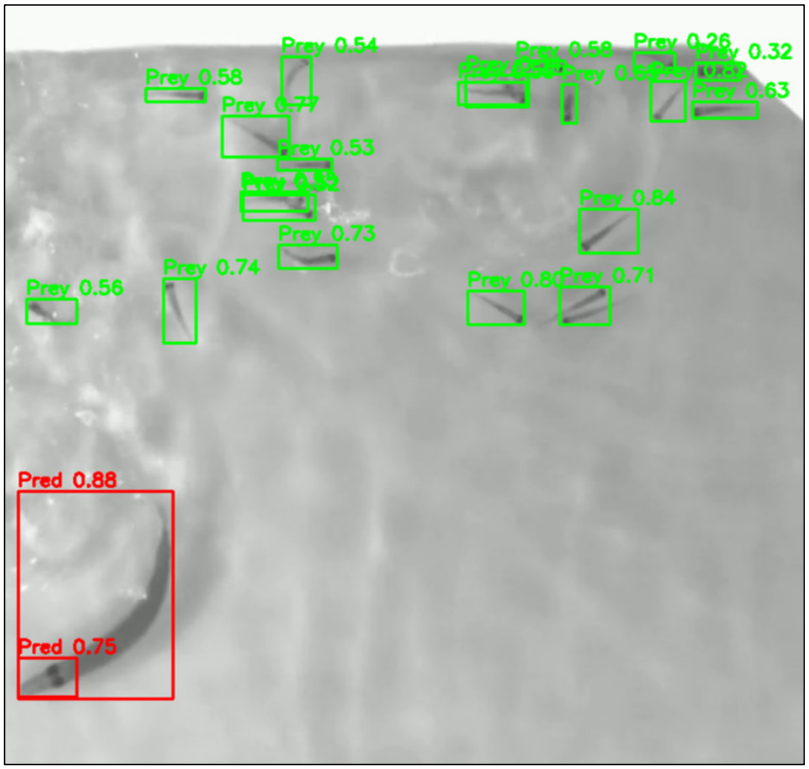
- 32 attacks
- 18.9 seconds
- 0.12% of total recordings

Data Collection & Processing



Object-detection:

- Custom YOLOv11
- fine-tuned on 100 hand-labeled frames
- extract positions



Detection error:

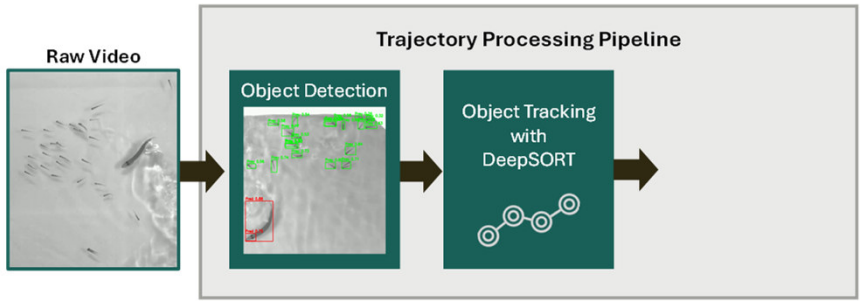
- Predator: MAE ± 0.03
- Prey: MAE ± 3.11

Error causes:

- Dense groups
- Occlusions
- Predation success

→ Attack scenes manually labeled

Data Collection & Processing



Object tracking (DeepSORT):

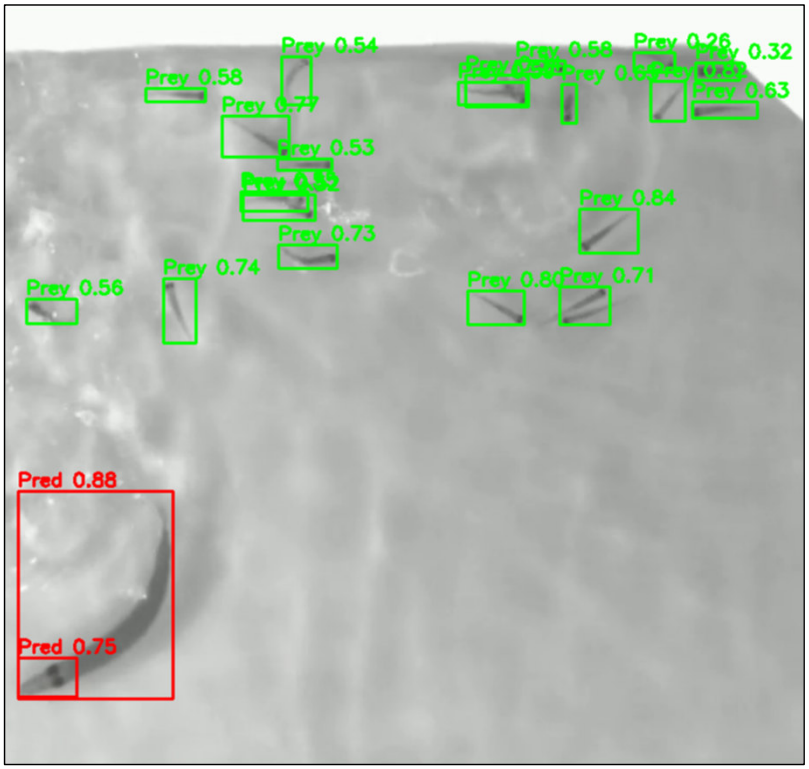
- Data association via Hungarian algorithm
- Kalman filter for state estimation and missed detections
- Produces consistent track IDs over time

Results:

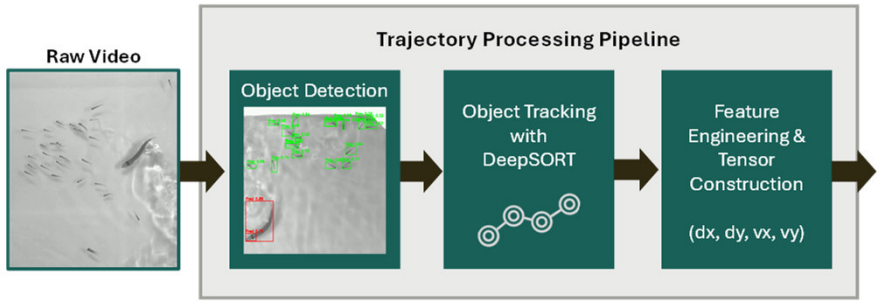
- 793 valid 10-frame windows
- 7.79% of all frames retained

Comparison:

- Wu Paper trained on 1.000 steps
- CBIL Paper on windows of 10



Data Collection & Processing



State / action representation:

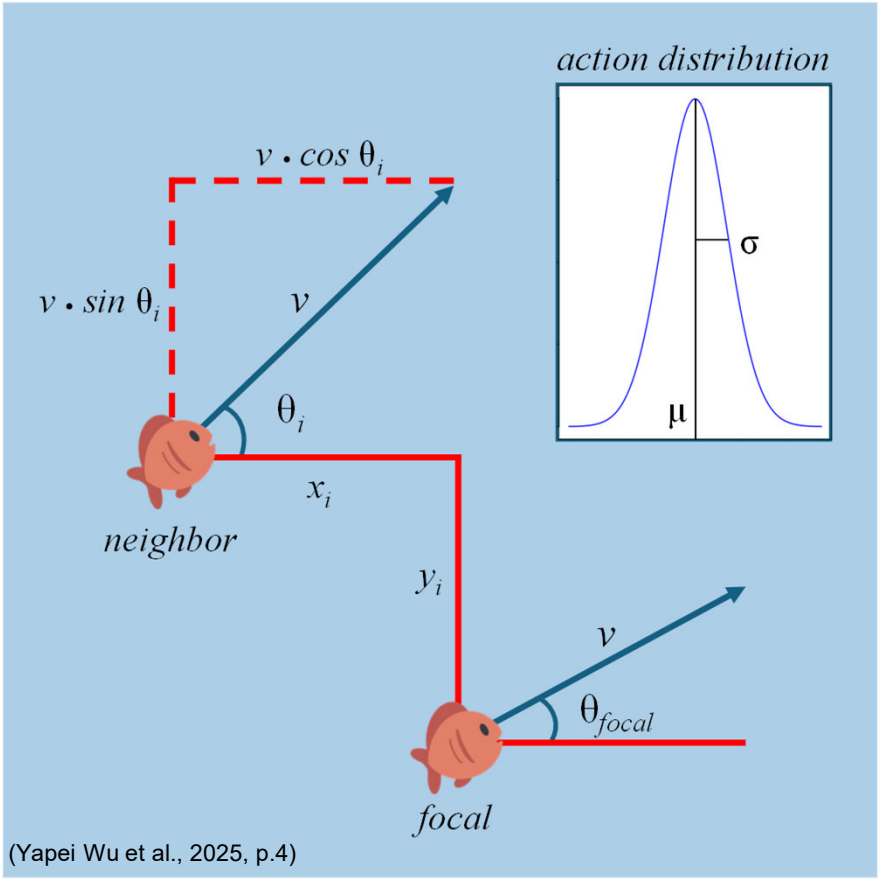
- Pairwise distances (dx, dy)
- Relative velocities (vx, vy)
- Action $\Delta\theta$

→ Prey: [flag, dx, dy, vx, vy, θ]
→ Predator: [dx, dy, vx, vy, θ]

Tensor: [batch, windows, agents, neigh, features]

$$\text{Prey Tensor} \in \mathbb{R}^{793\times10\times32\times32\times6}$$

$$\text{Predator Tensor} \in \mathbb{R}^{793\times10\times1\times32\times5}$$



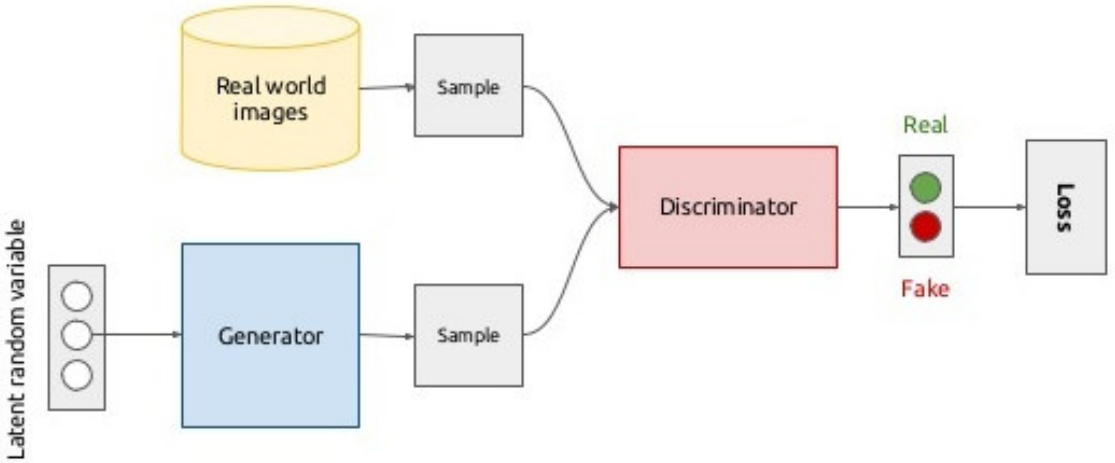
(Yapei Wu et al., 2025, p.4)

Generative Adversarial Networks

GAN:
- Adversarial learning
- Policy vs. discriminator
- Distribution matching

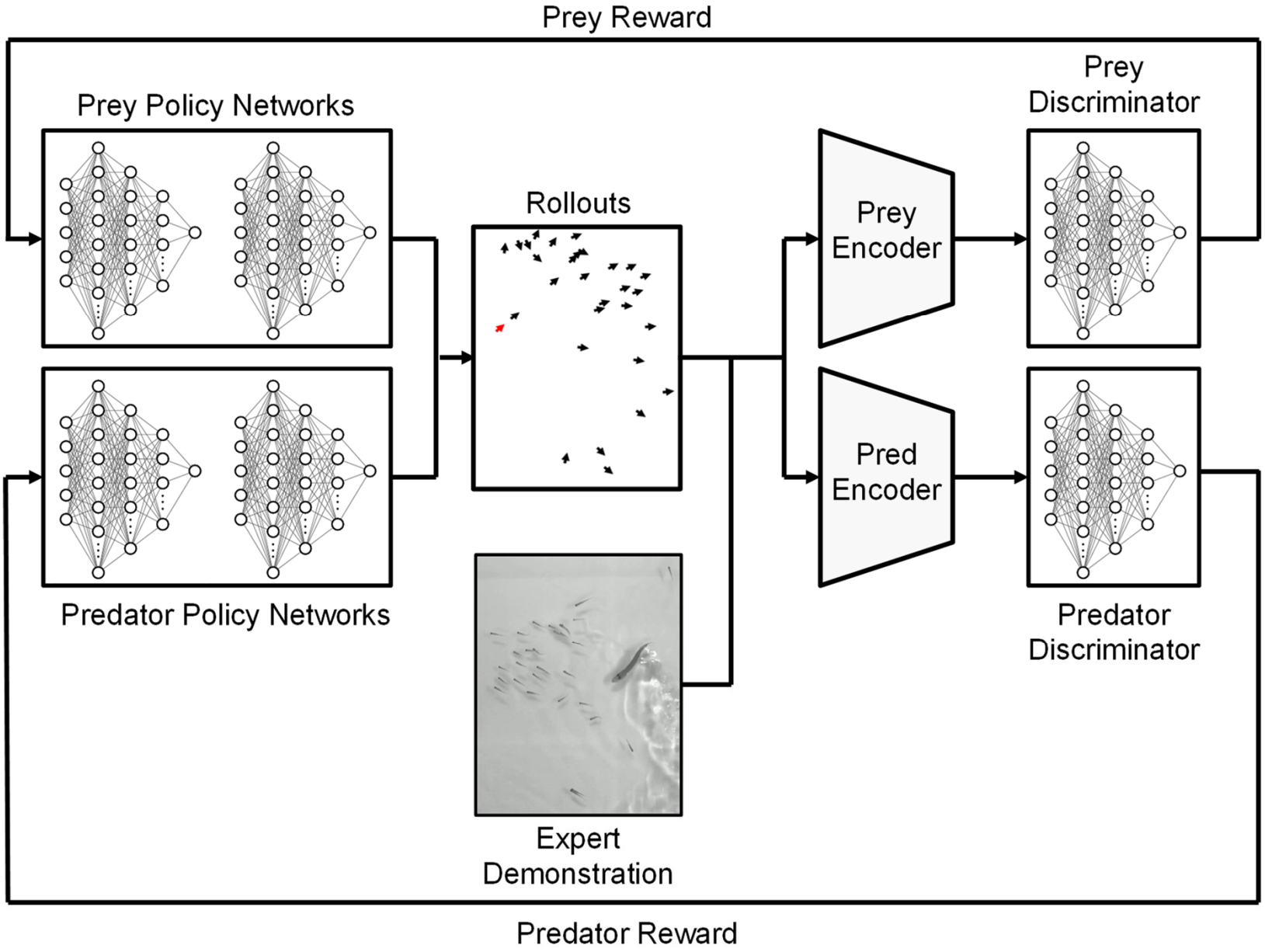
Problems:
- Balance issues
- Mode collapse

→ GAIL naturally scales to multi-agent systems

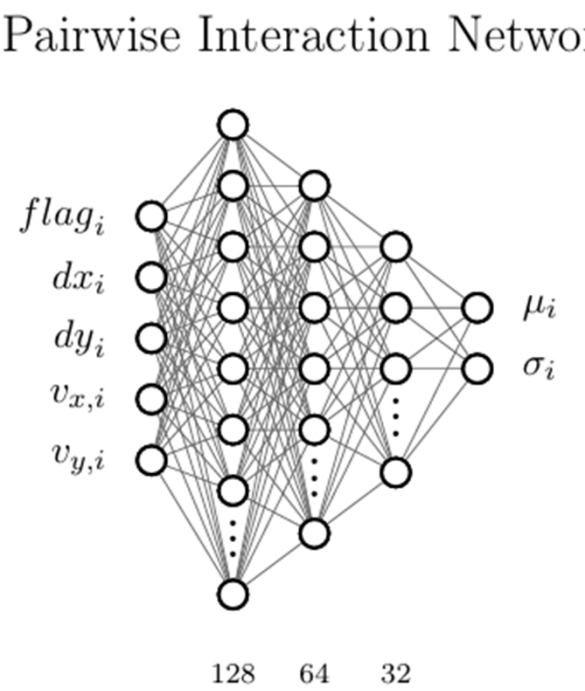


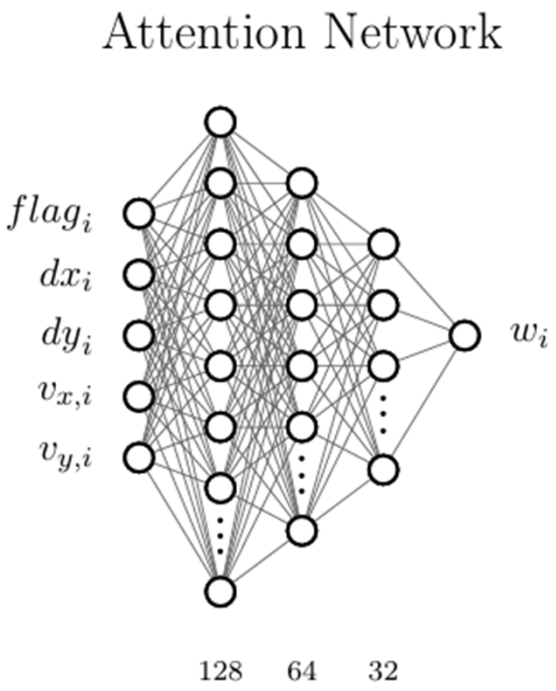
https://collab.dvb.bayern/spaces/TUMIfdv/pages/69119933/Generative+Adversarial+Networks+GANs

Methodology: GAIL



Modular Policy





Pairwise Interaction Network:

- Input: [dx, dy, vx, vy]
- Output: μ, σ
- Parameters of a Gaussian action distribution

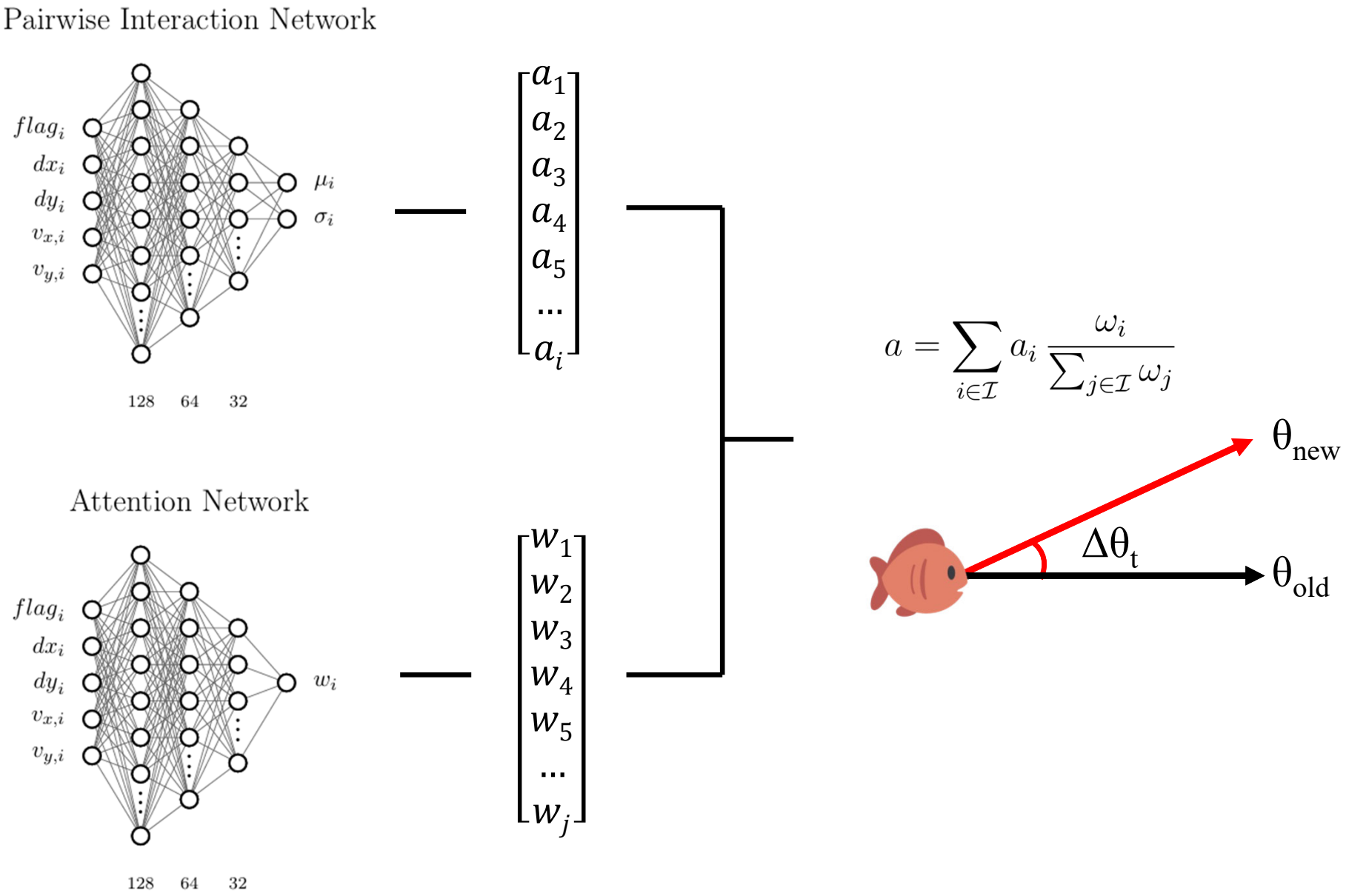
→ **stochastic individual response**

Attention Network:

- Input: [dx, dy, vx, vy]
- Output: w
- Relative influence of each neighbor

→ **interpretable interaction structure**

Modular Policy

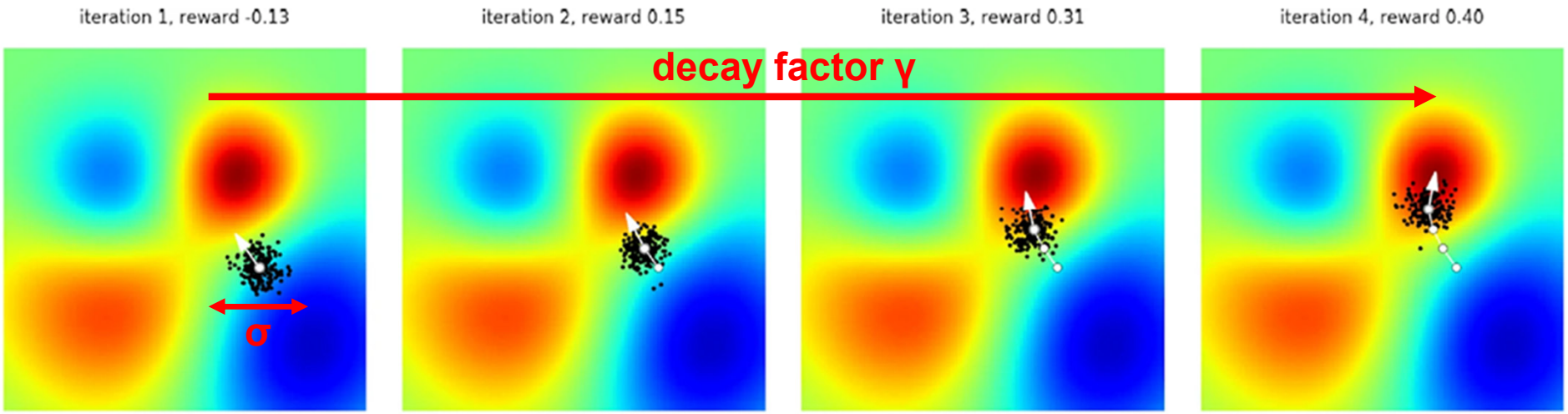


Modular Policy – Evolutionary Strategy

"At every iteration (“generation”), a population of parameter vectors (“genotypes”) is perturbed (“mutated”) and their objective function value (“fitness”) is evaluated“ (Salimans et al., 2017, p.2).

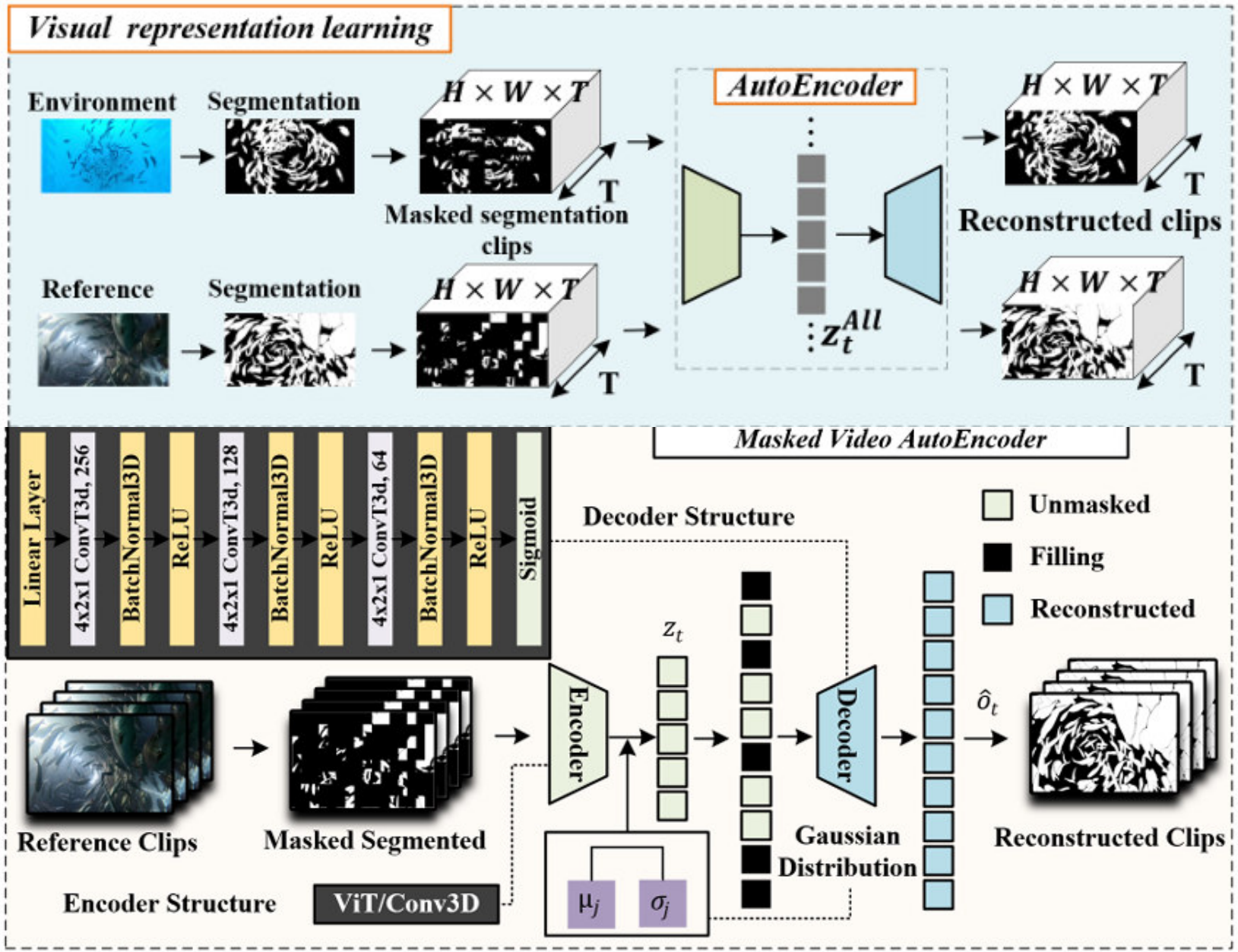
1. Stage:
- Sample perturbation noise
- Apply perturbations to parameters
- Collect rollout rewards

2. Stage:
- Weight perturbations by obtained rewards
- Higher-reward = higher update influence



https://images.ctfassets.net/kftzwdyauwt9/d5acb8a0-a1a1-4772-09f9e1a6550b/e7bccd4dd7532331595032ac7b9e3f14/evo.png

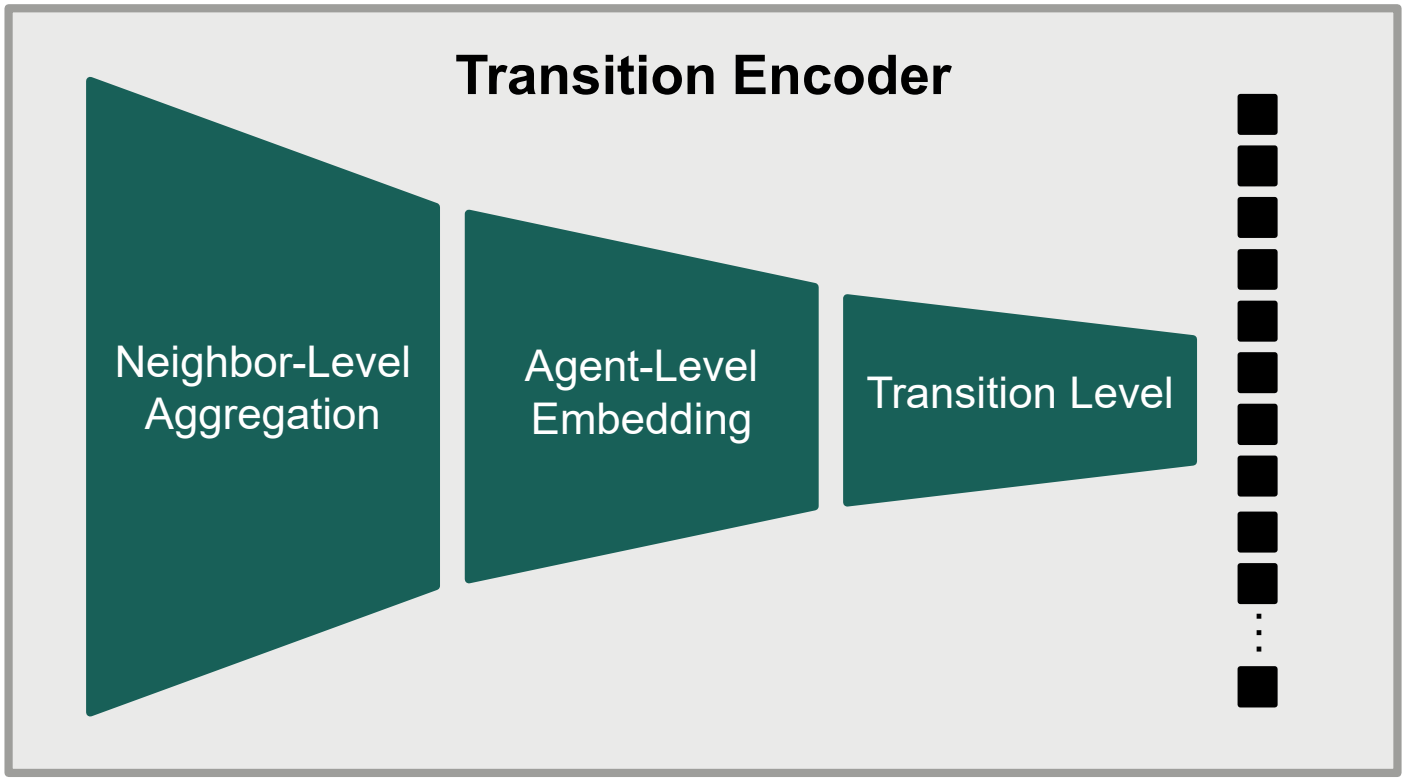
CBIL Transition Encoder



(Yifan Wu et al., 2025, p.4)

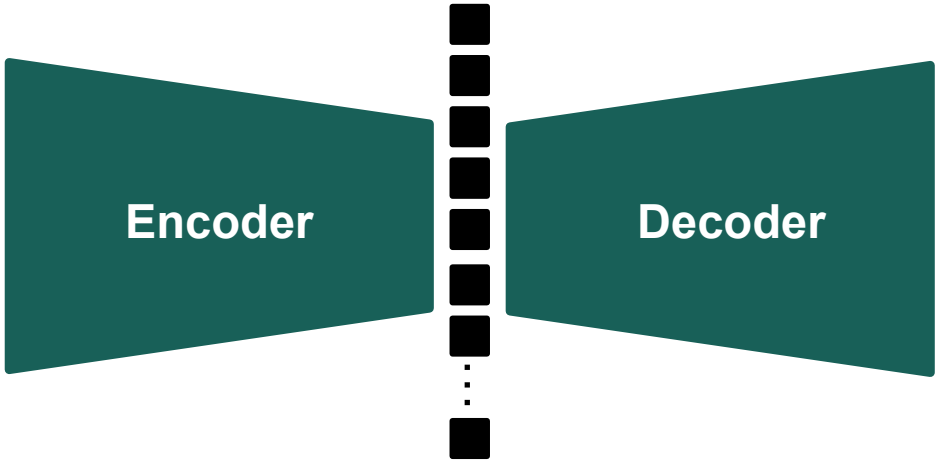
Transition Encoder

Input tensor:
[batch = 10, windows = 10, agents = 33, neigh = 32, feat = 4]



Output tensor: [batch, windows-1, agents, 2z]
Transition Feature (2z) = $[z_t, \Delta z_t]$

Transition Encoder - Training

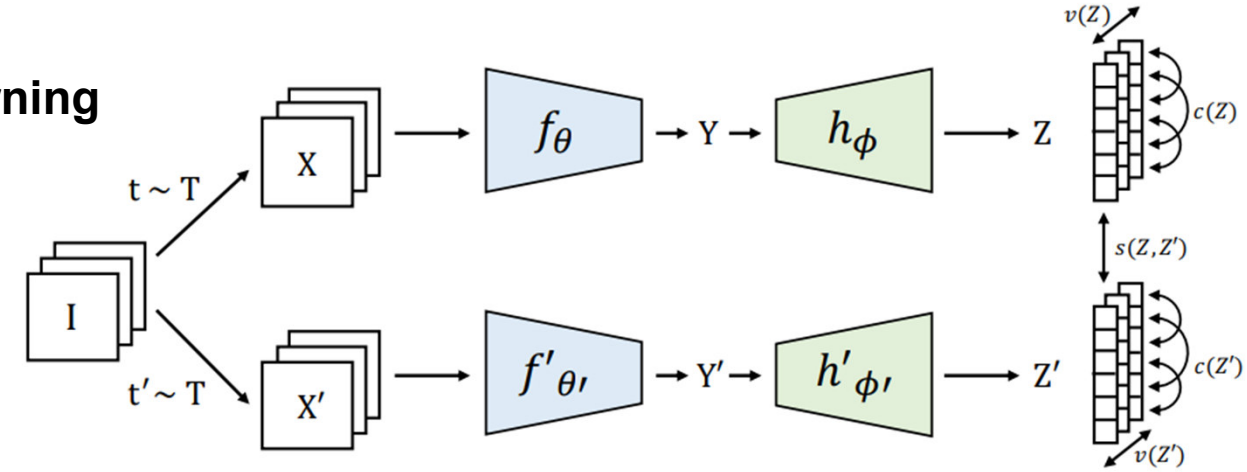


Problems with reconstruction-based training

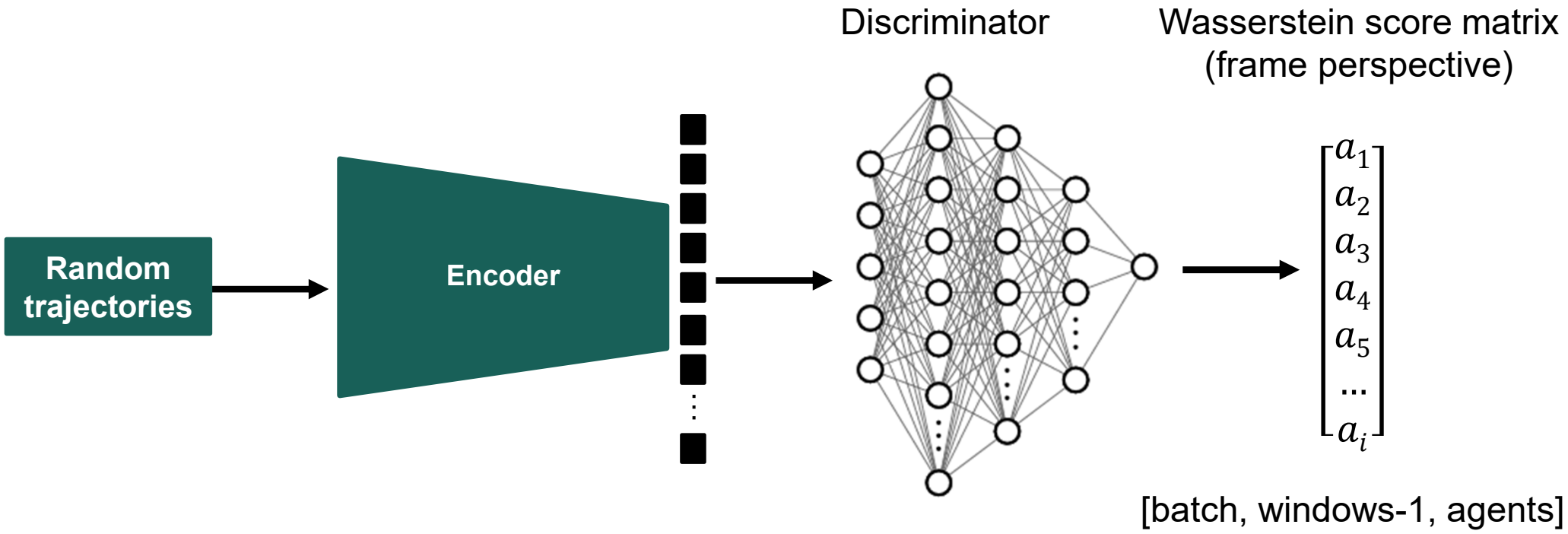
- Poor reconstruction loss → weak latent?
- Decoder not required for GAIL
- Question: Training possible without reconstruction? → **VICReg**

Self-supervised representation learning

- No decoder, direct latent training
- Augmentation:
  - State noise
  - Neighbor dropout
  - Feature dropout



Discriminator

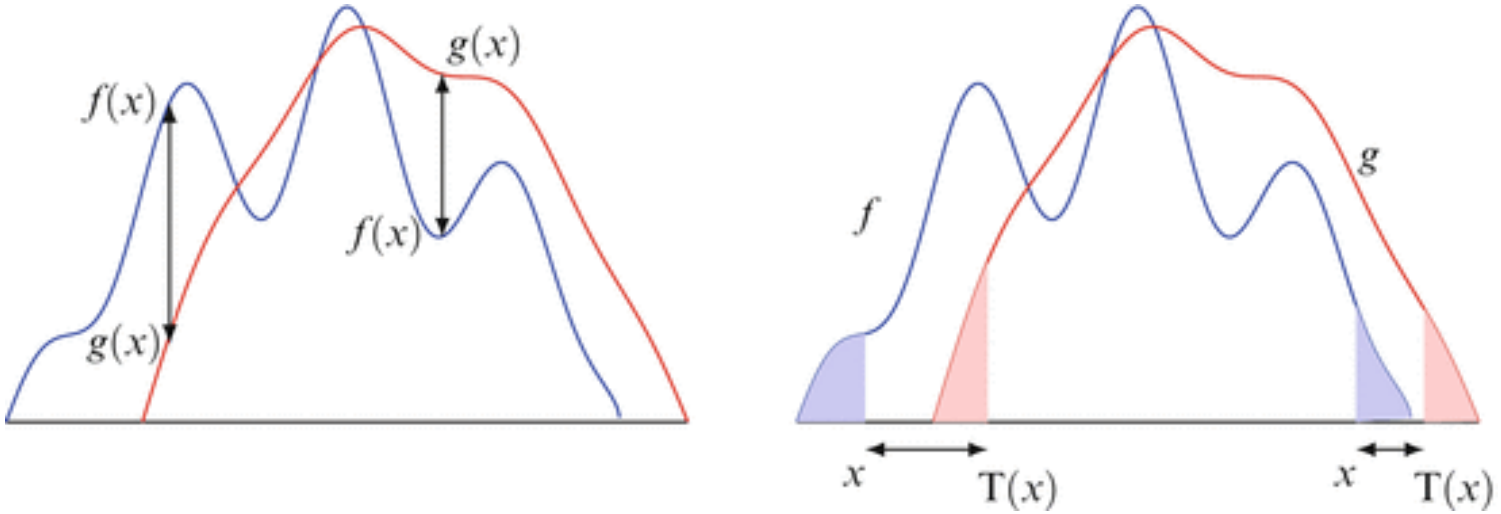


→ **Optimizer: RMSProp**

Wasserstein Loss & Earth-Mover Distance

$$W(\mathbb{P}_E, \mathbb{P}_\pi) = \inf_{\gamma \in \Pi(\mathbb{P}_E, \mathbb{P}_\pi)} \mathbb{E}_{(\tau_E, \tau_\pi) \sim \gamma} \left[\|\tau_E - \tau_\pi\| \right]$$

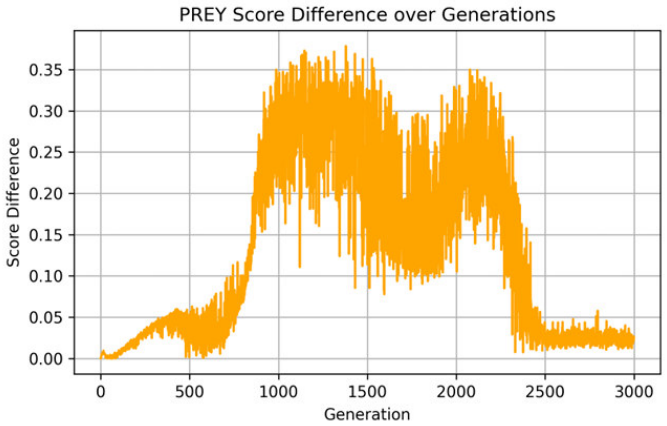
"Intuitively, $\gamma(\tau_E, \tau_\pi)$ indicates how much “mass” must be transported from τ_E to τ_π in order to transform the distributions P_E into the distribution P_π (Arjovsky et al., 2017, p. 4)”.



https://link.springer.com/chapter/10.1007/978-3-319-20828-2_5

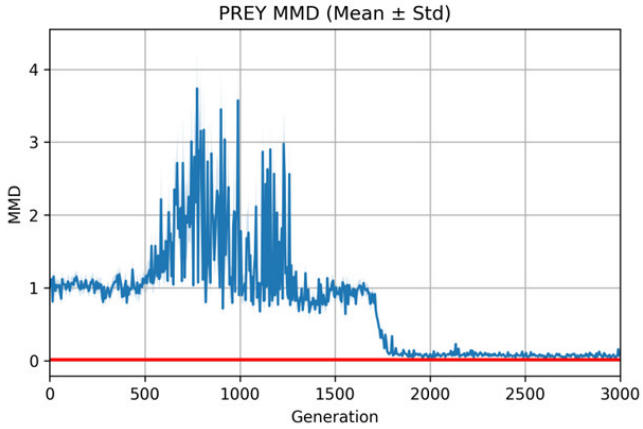
Performance Evaluation

Wasserstein proxy



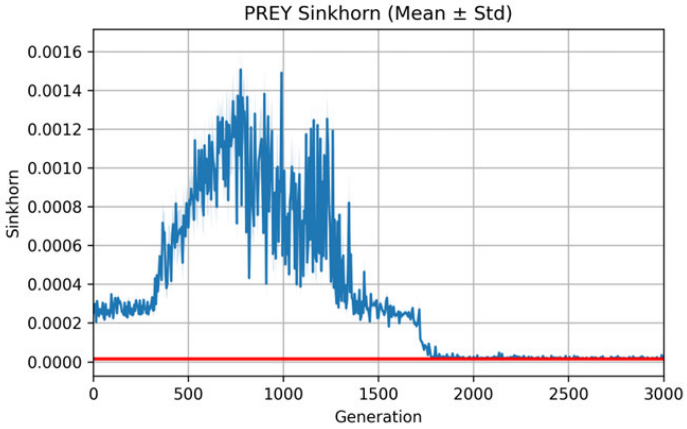
Discriminator score differences

Maximum Mean Discrepancy (MMD)



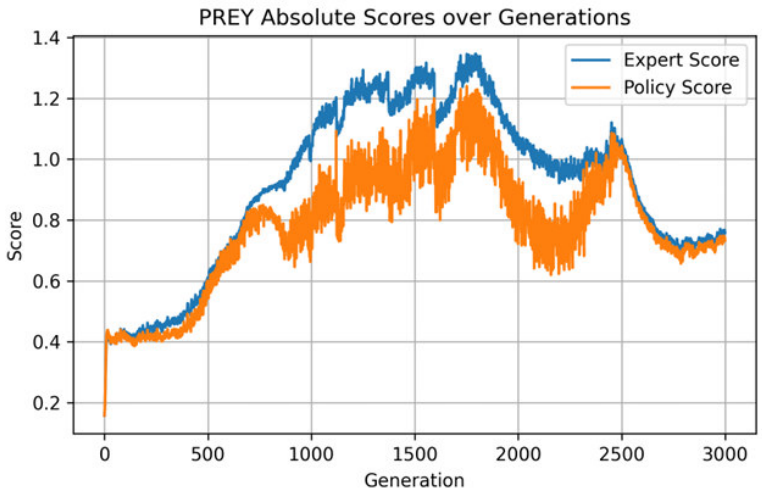
Distance between sampled batches

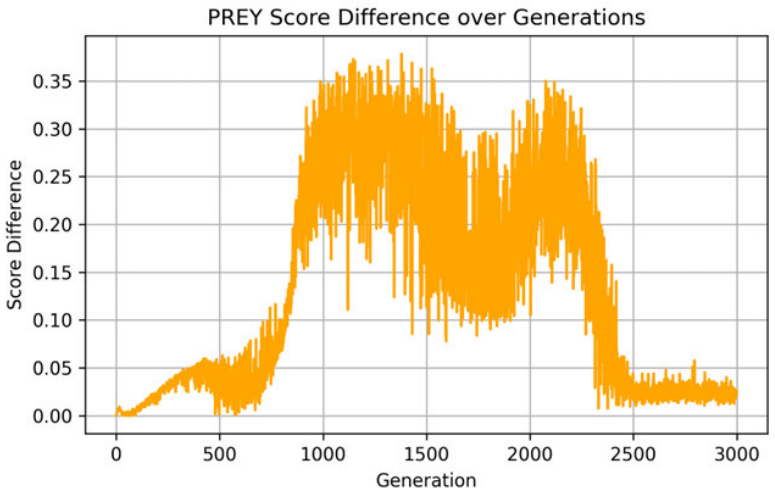
Sinkhorn distance

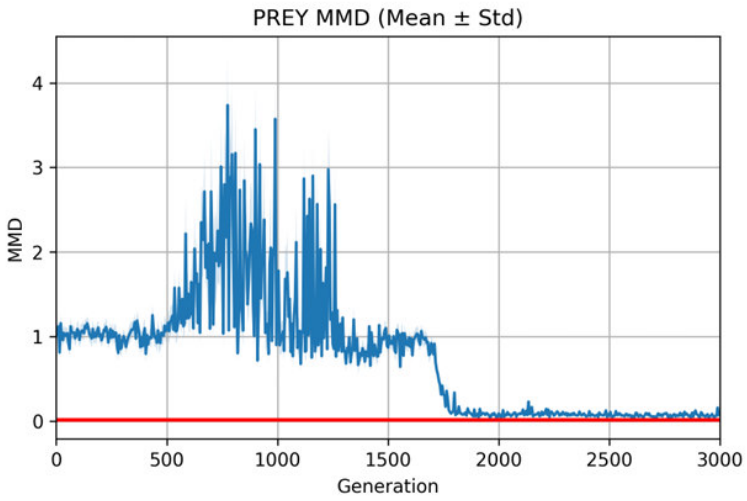


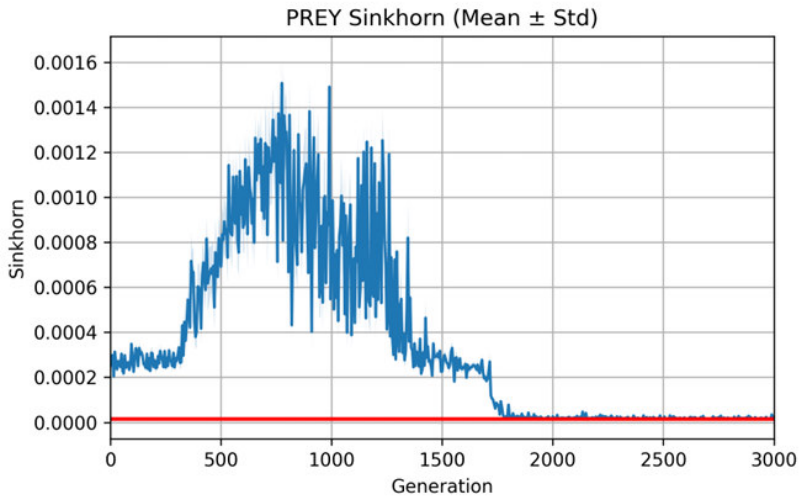
Distance between transition features

Prey-Only Model



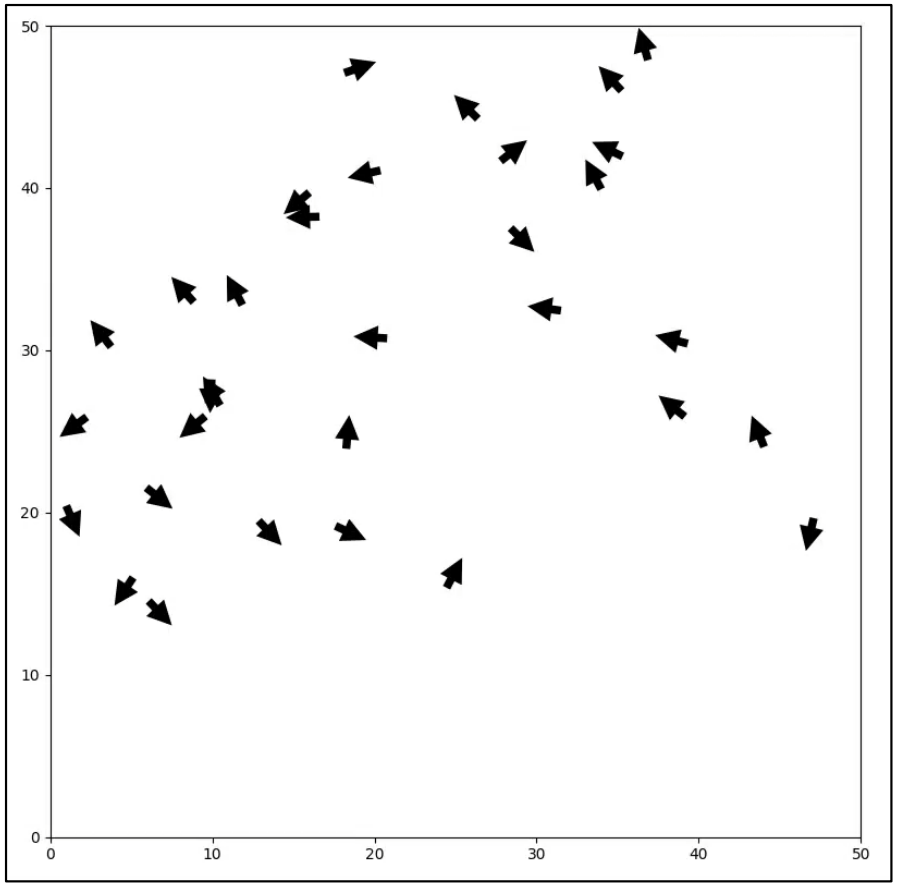




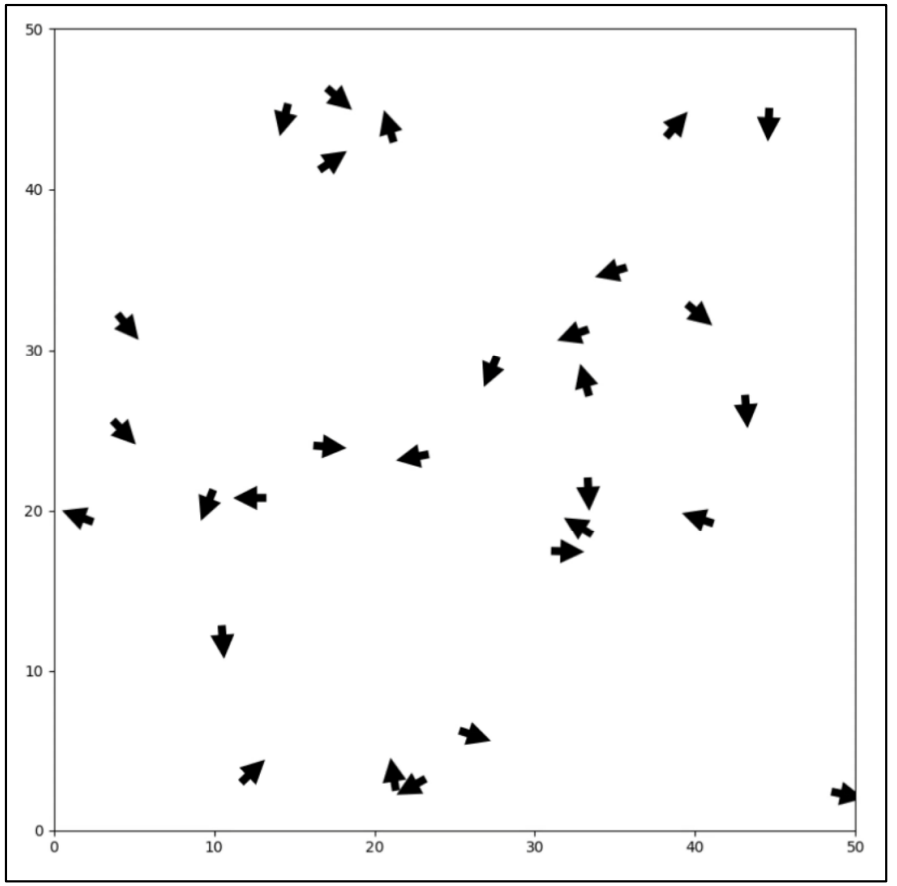


Prey-Only Model

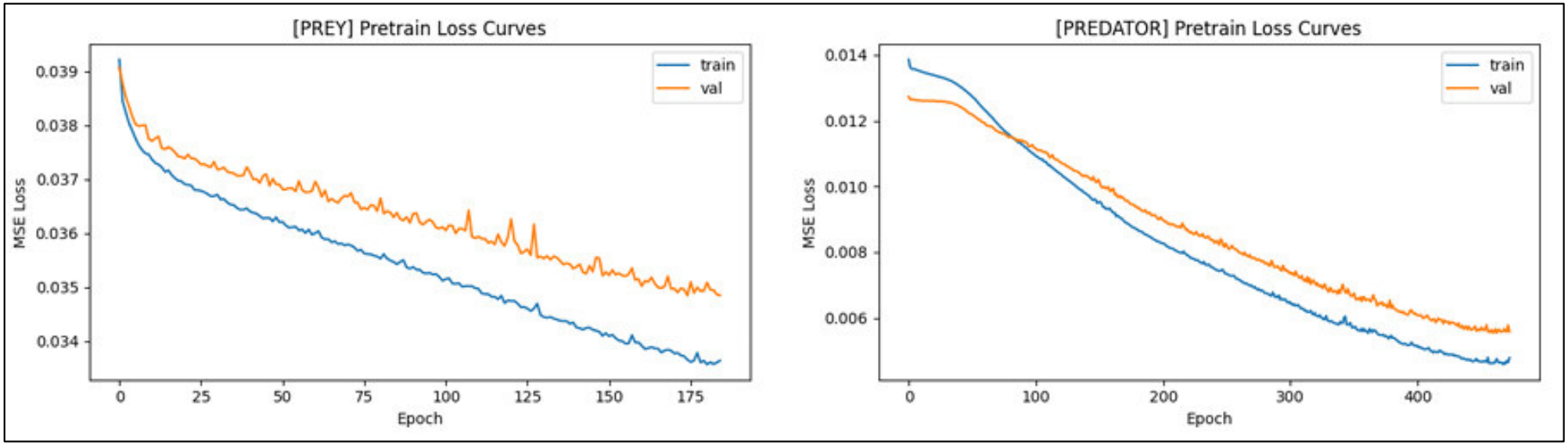
Expert demonstrations
Couzin model



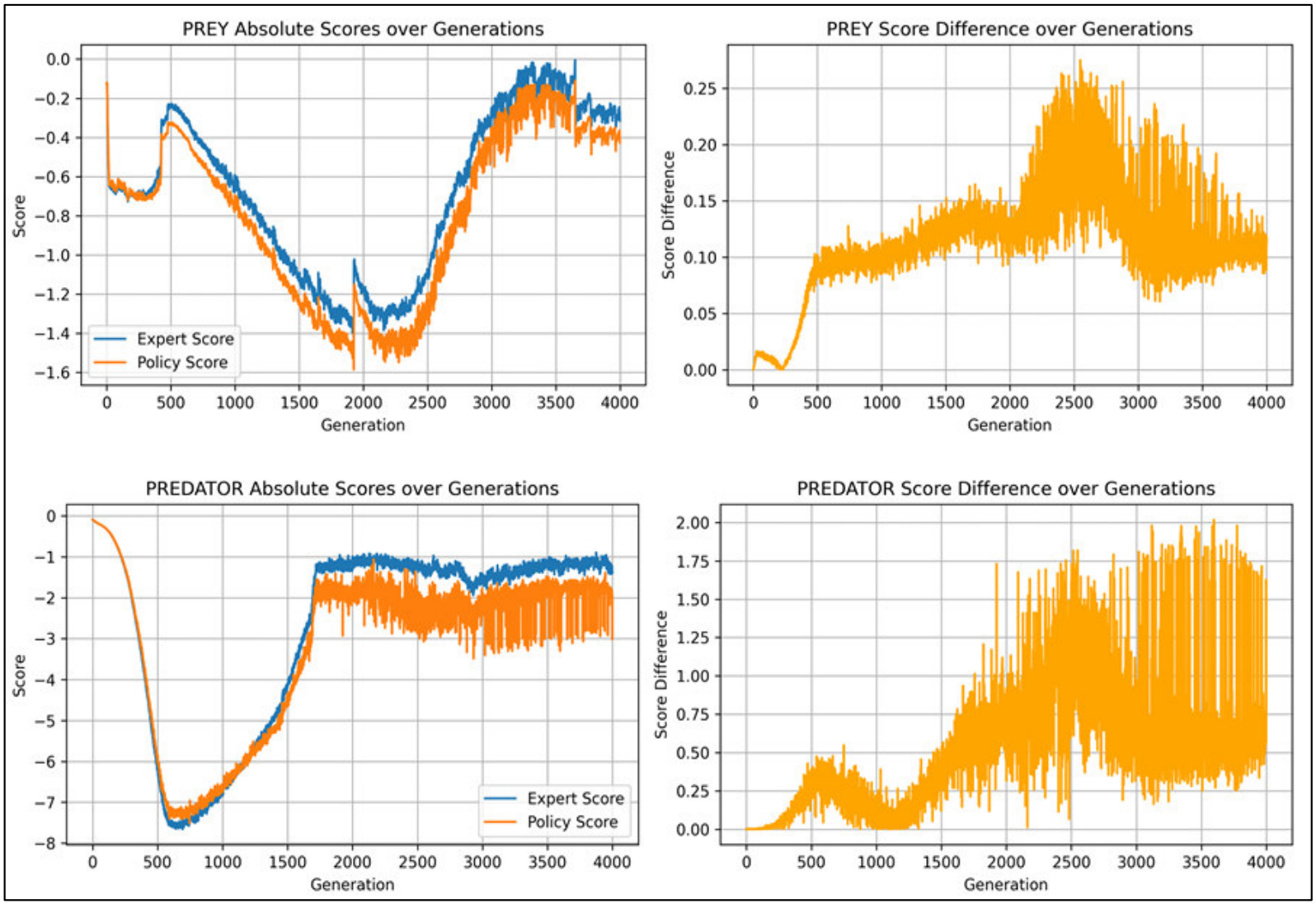
Policy-generated
Prey-only GAIL model



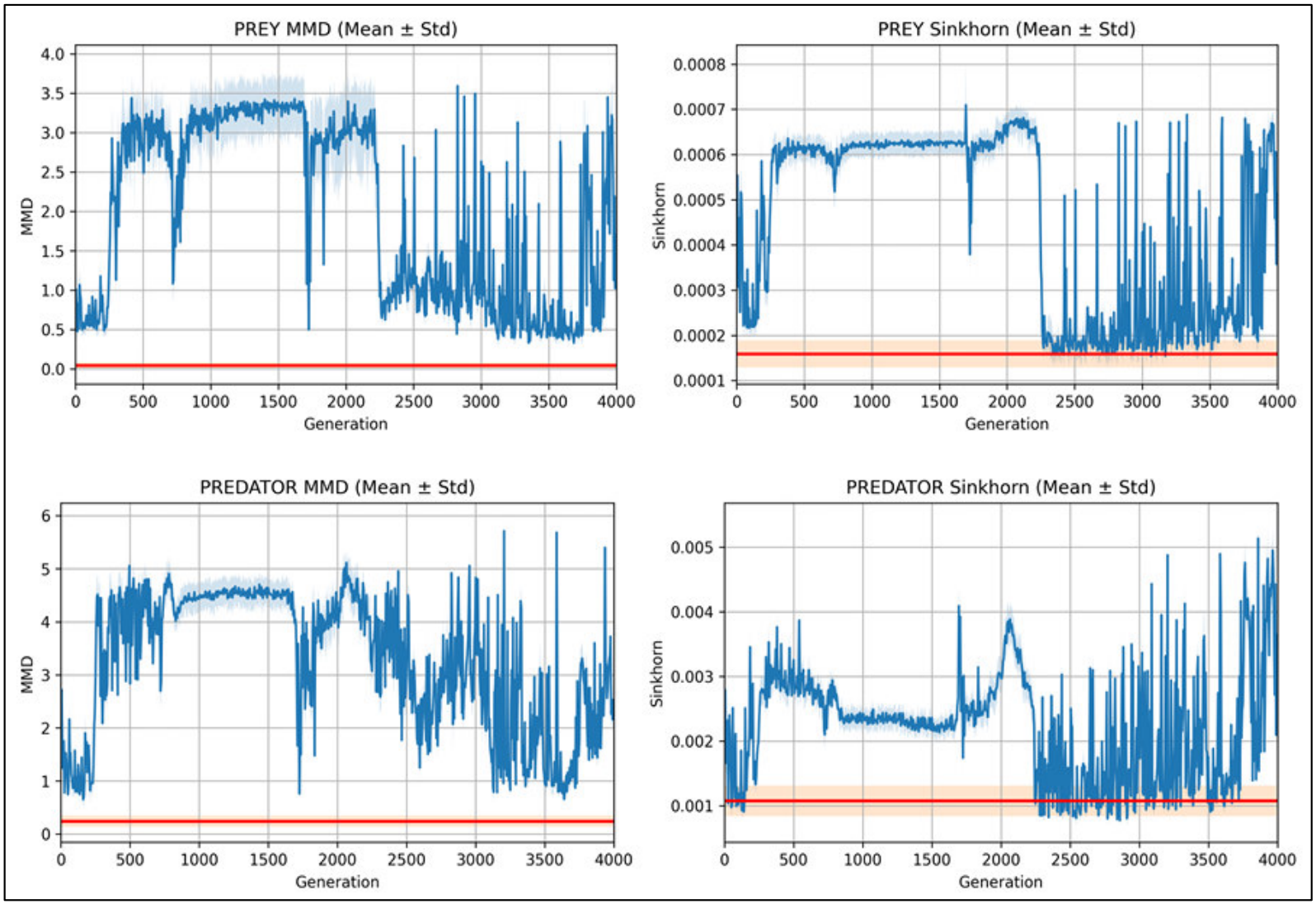
Video Predator-Prey Model



Video Predator-Prey Model

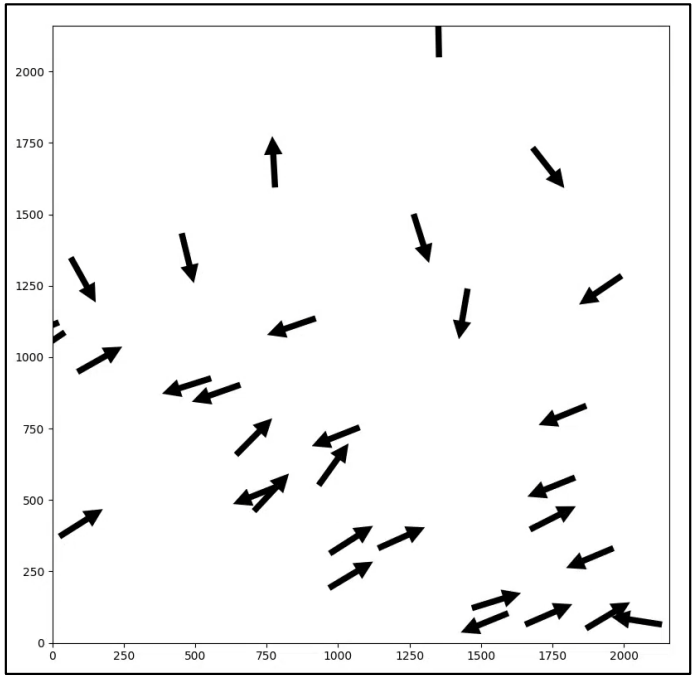


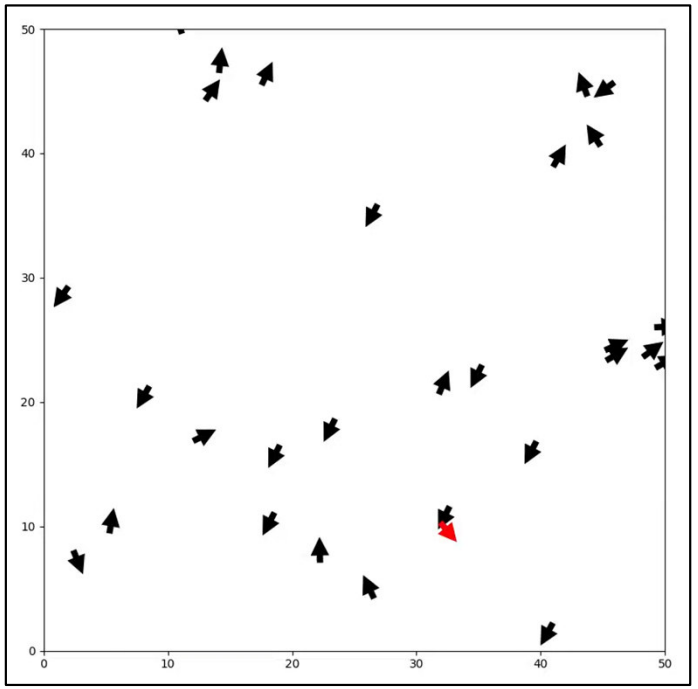
Video Predator-Prey Model



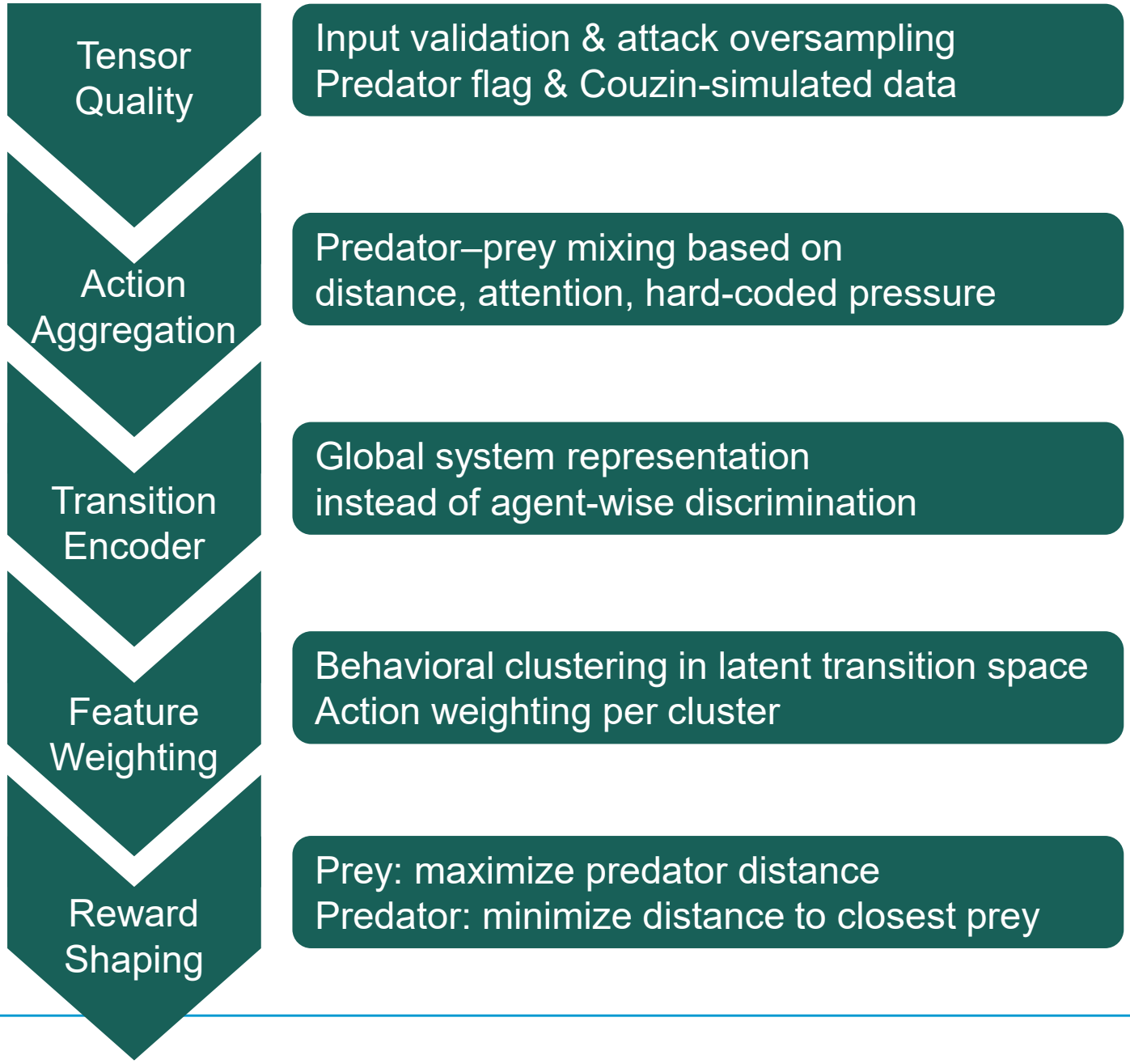
Video Predator-Prey Model

Expert demonstrations

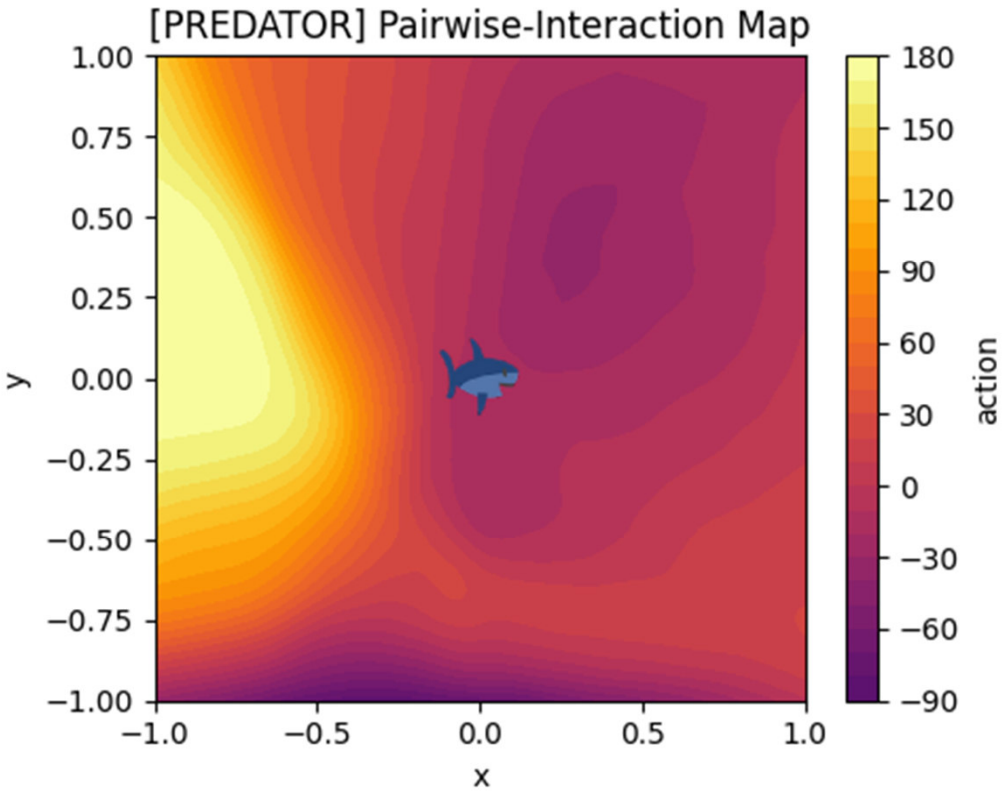


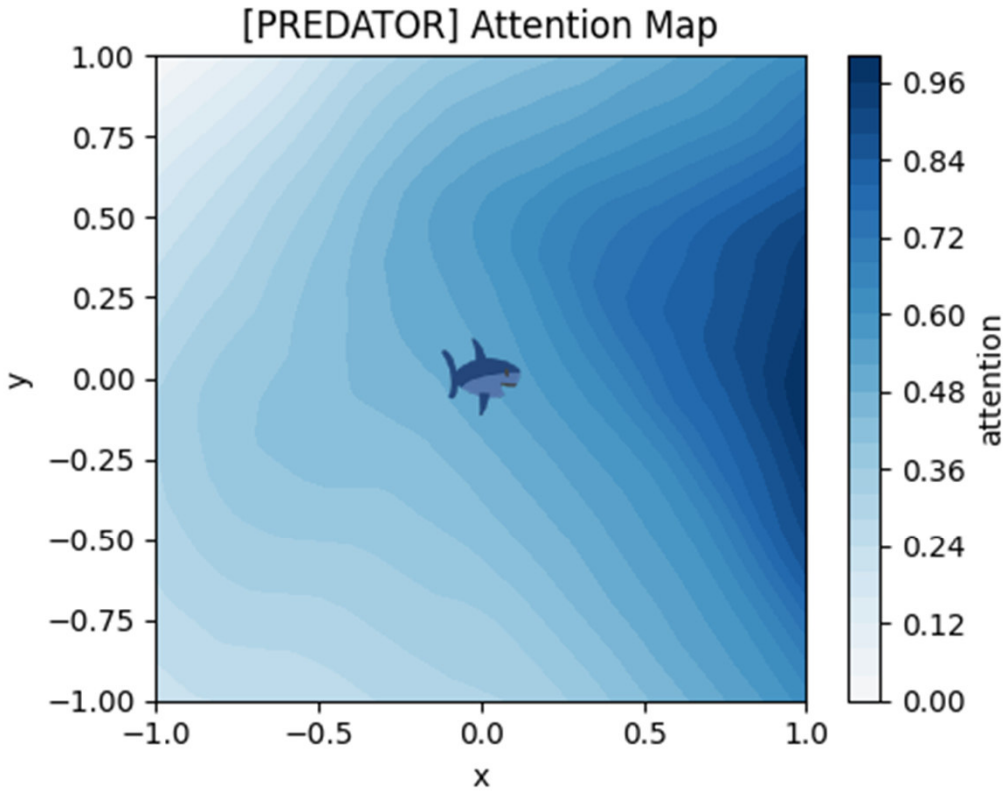


Recovery of missing inter-group dynamics

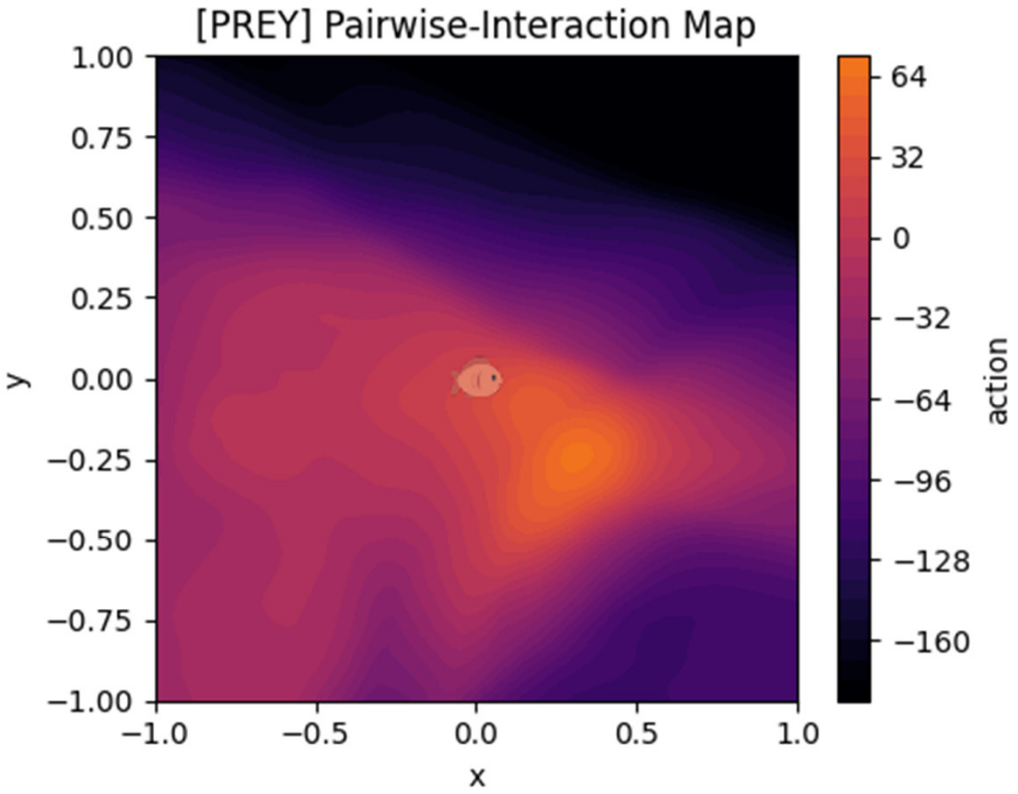


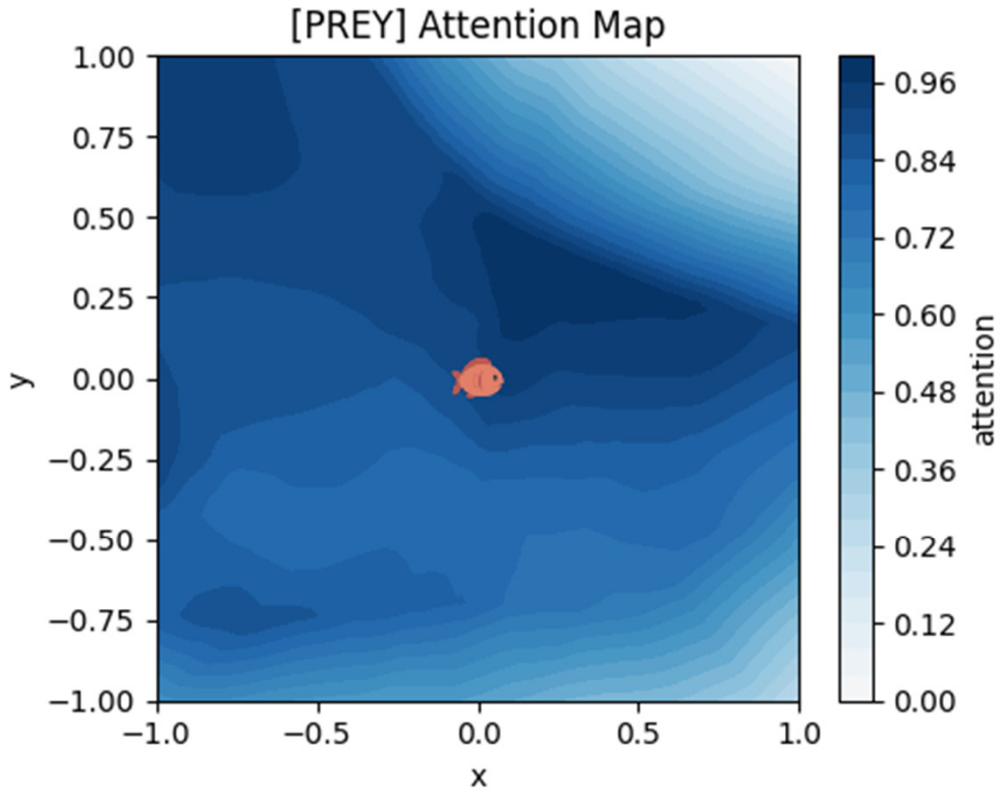
Expertiment: Policy Maps



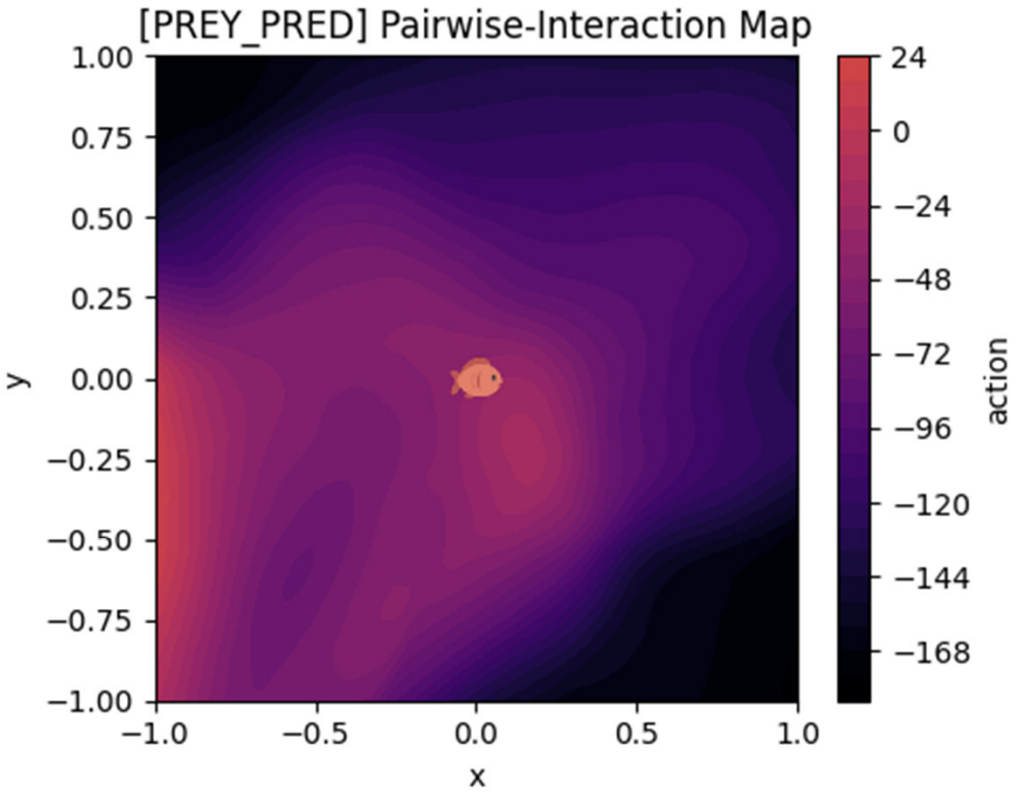


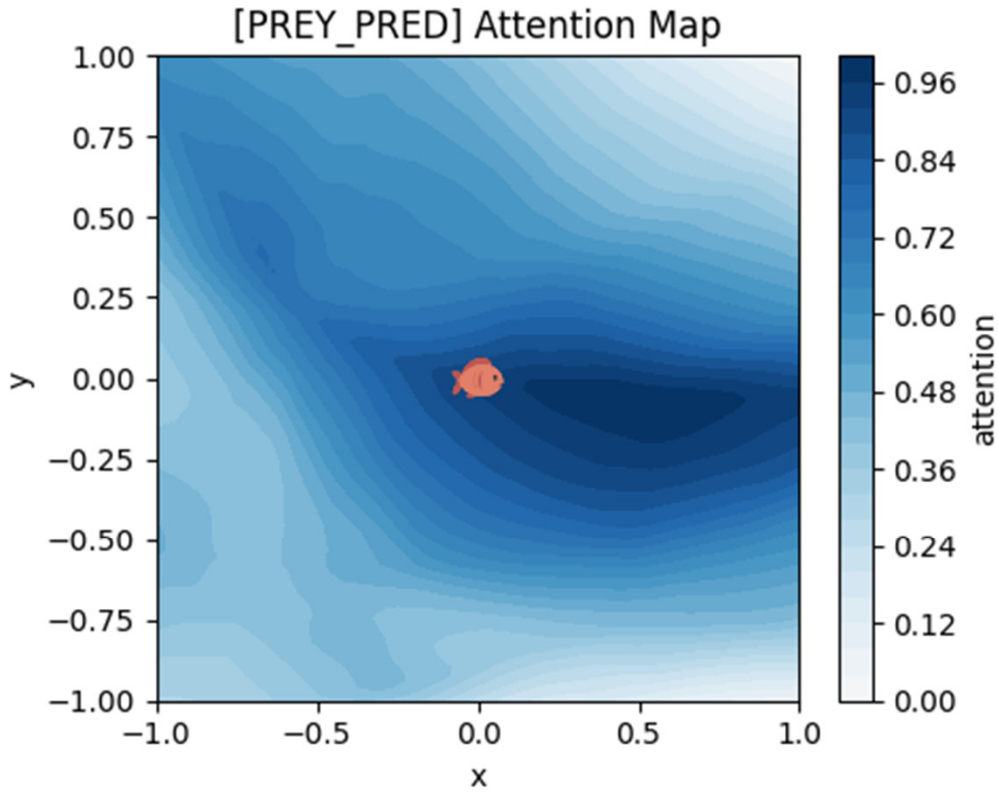
Expertiment: Policy Maps





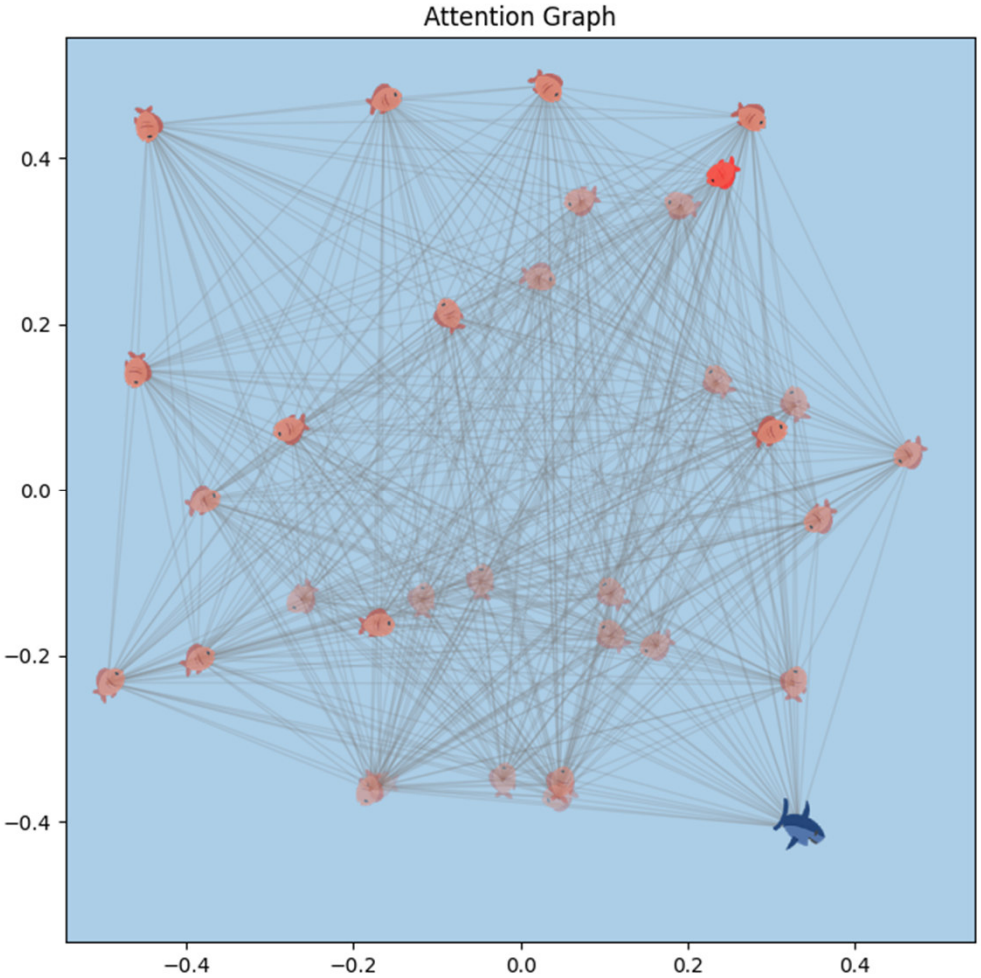
Expertiment: Policy Maps

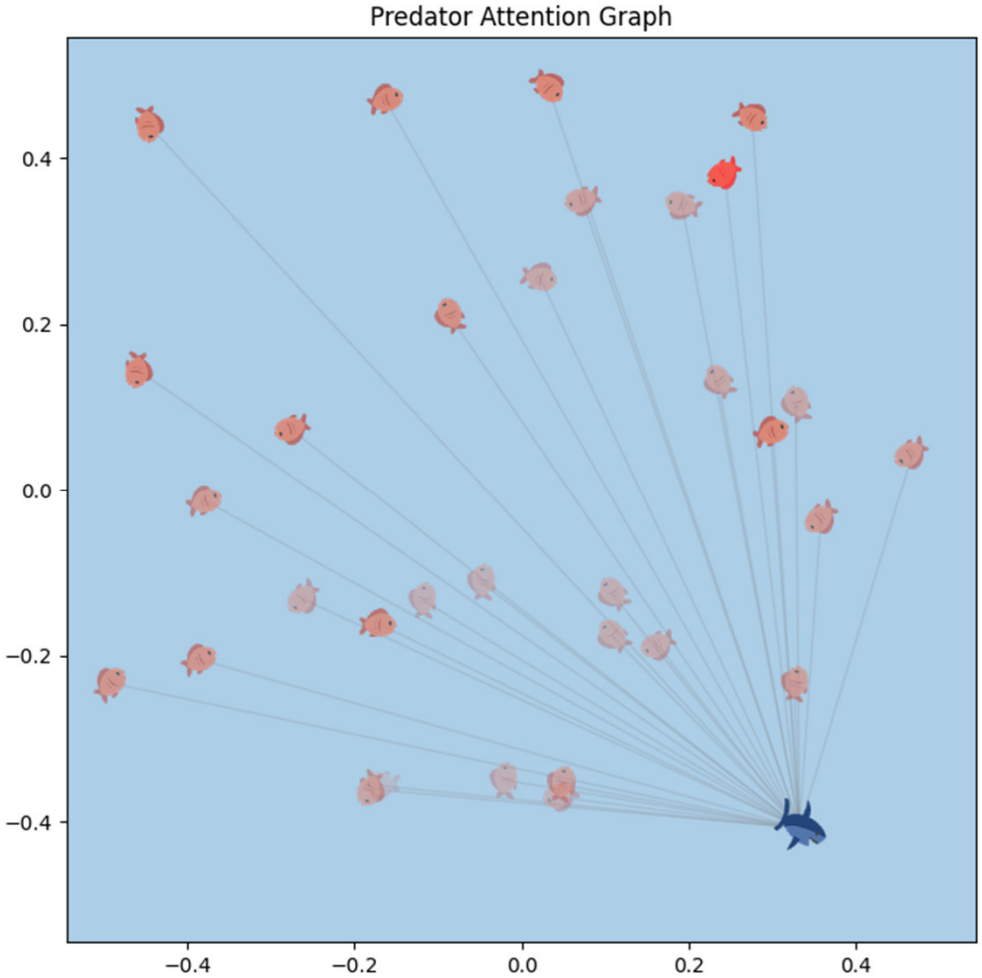




Experiment: Leadership Analysis

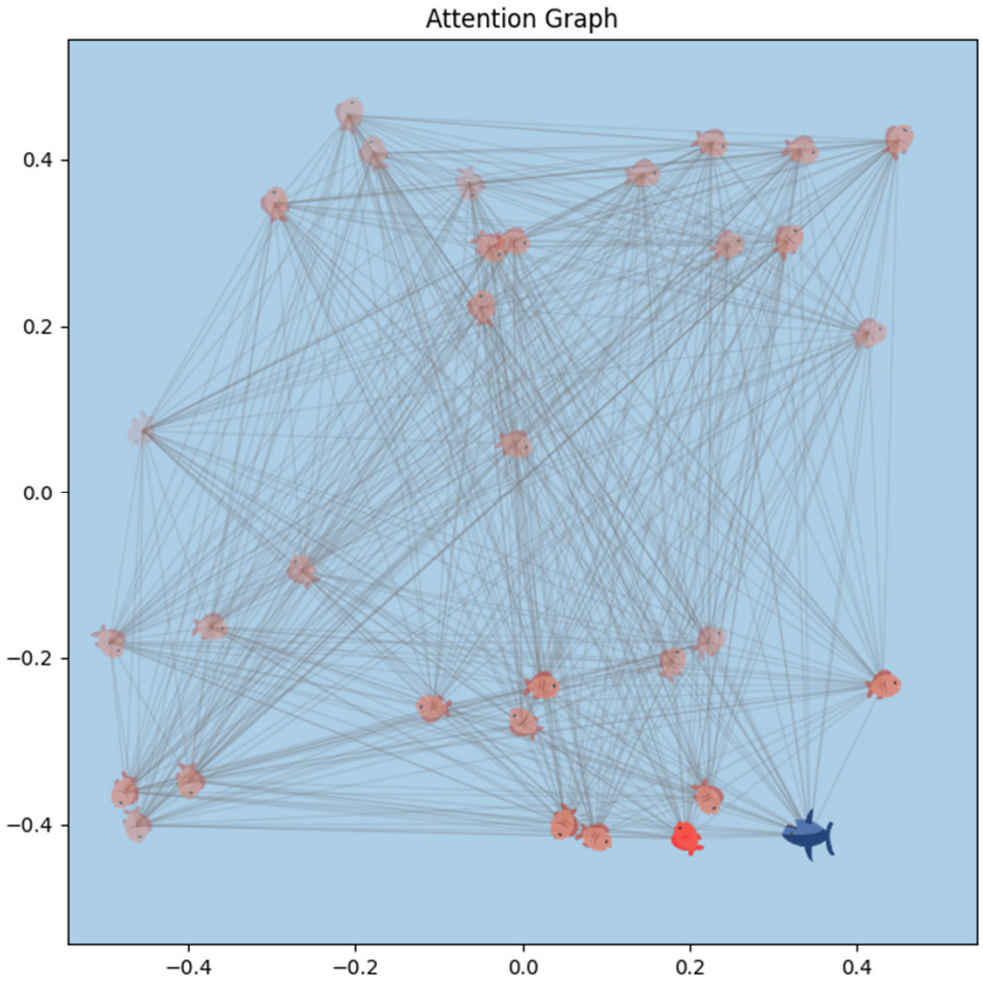
GAIL policies:

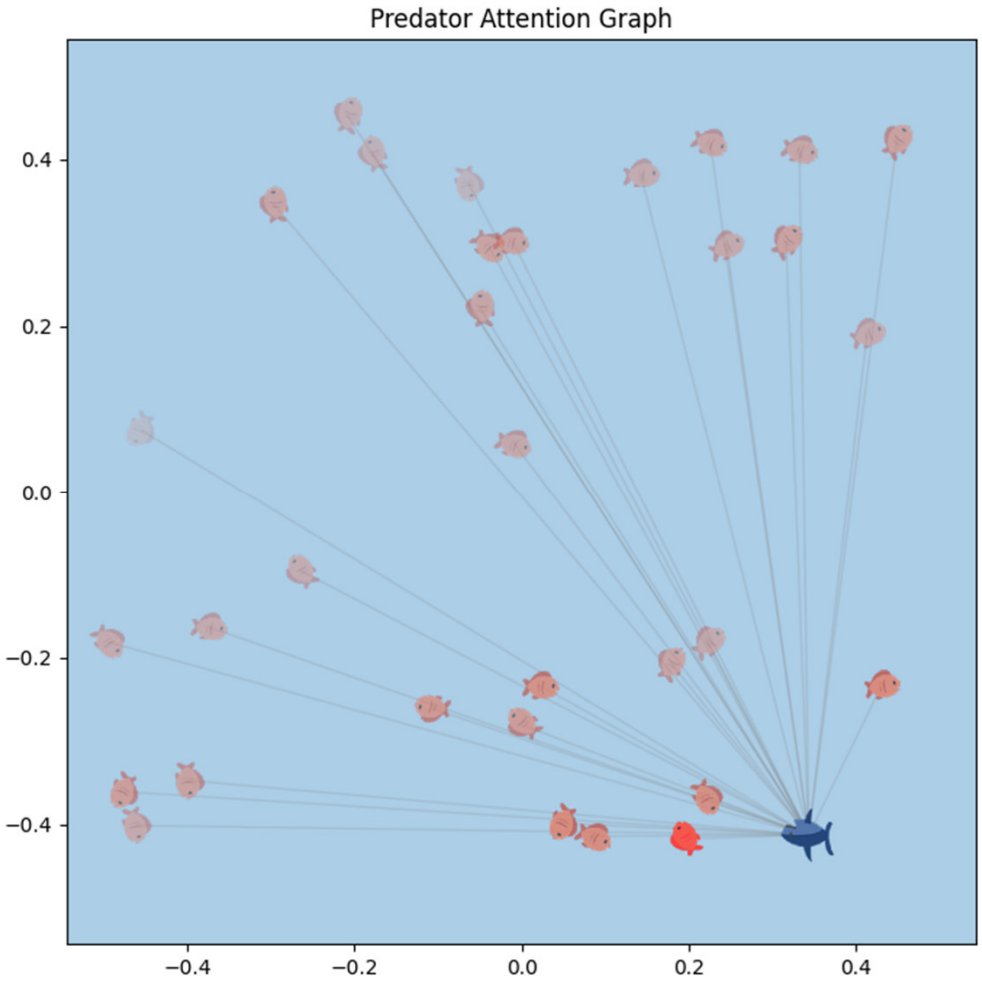




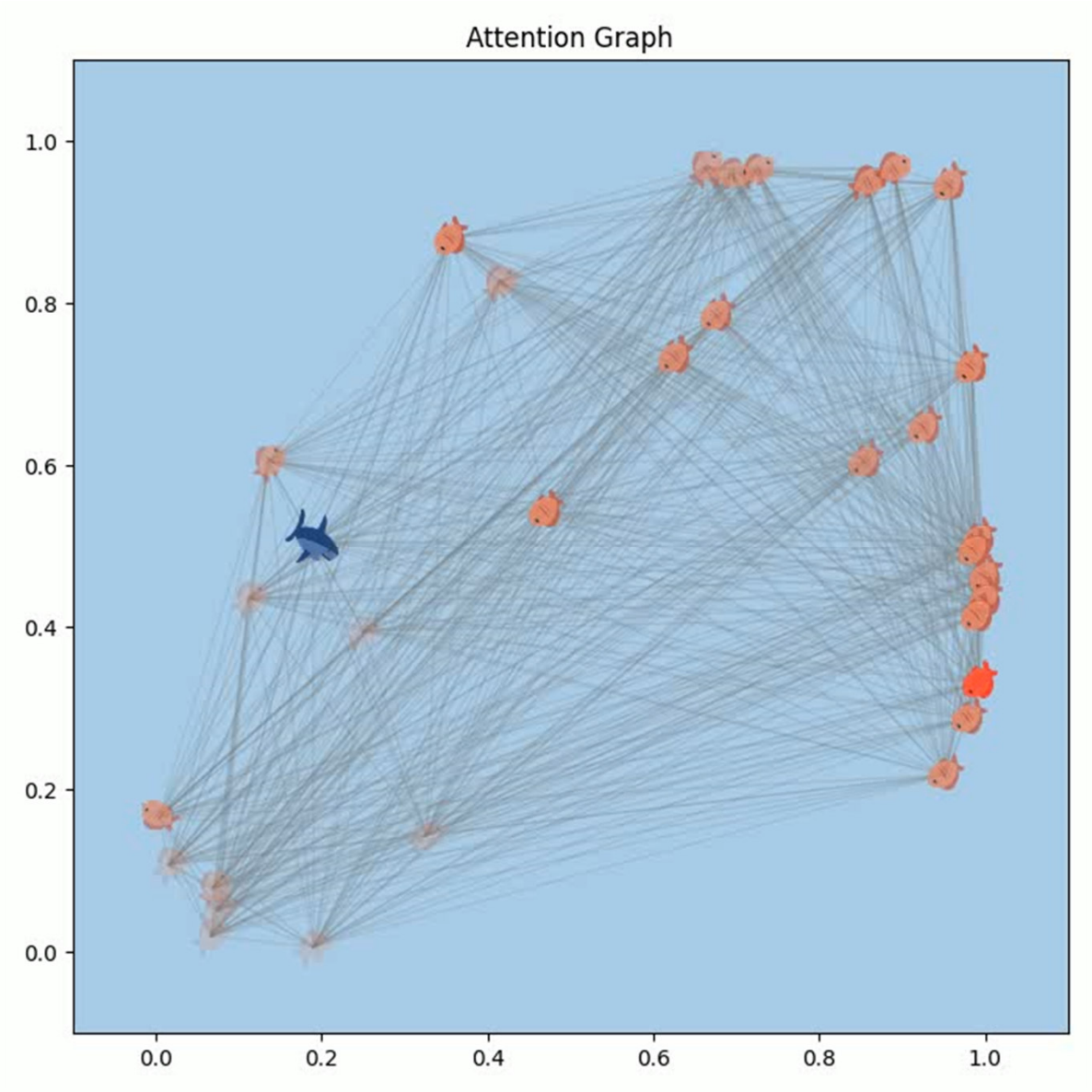
Expertiment: Leadership Analysis

BC policies:

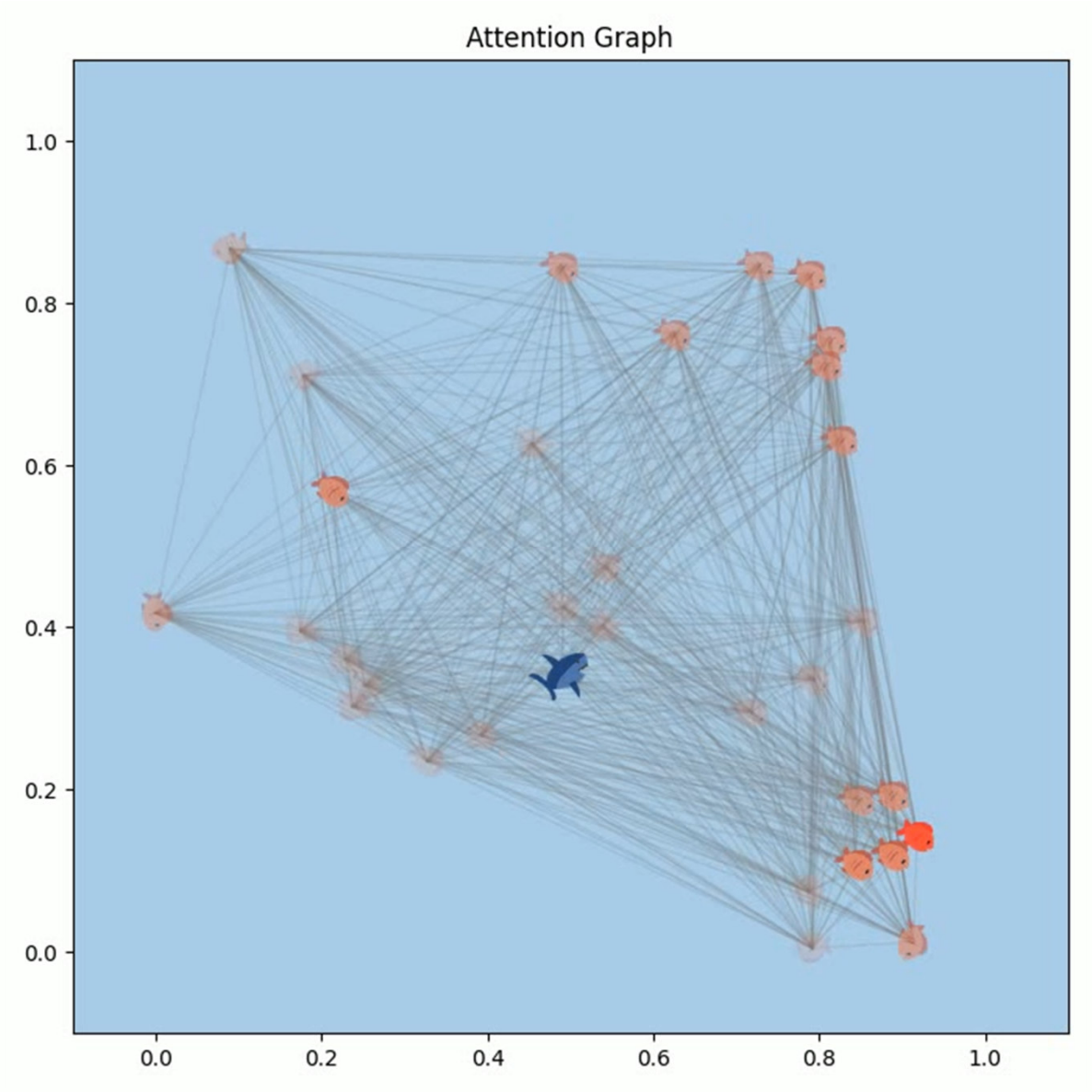




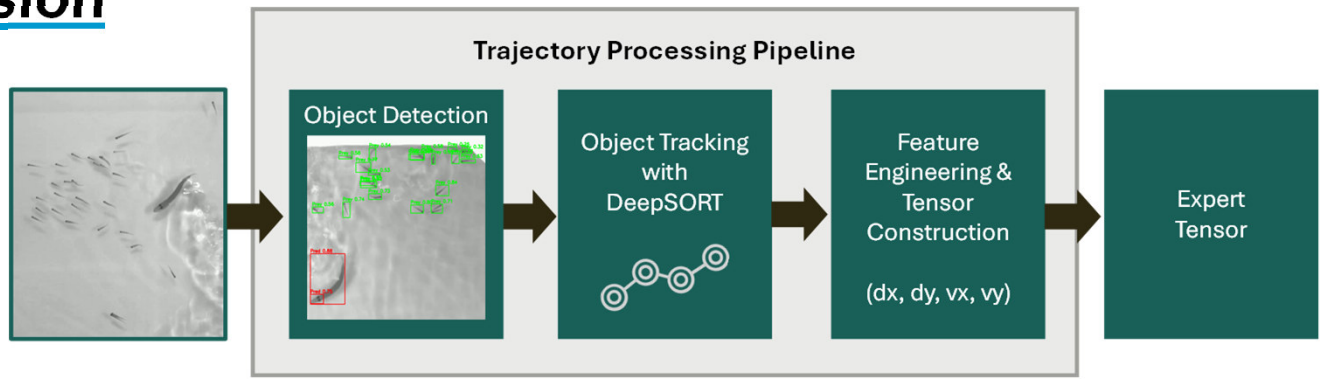
Expertiment : Leadership Analysis



Expertiment : Leadership Analysis



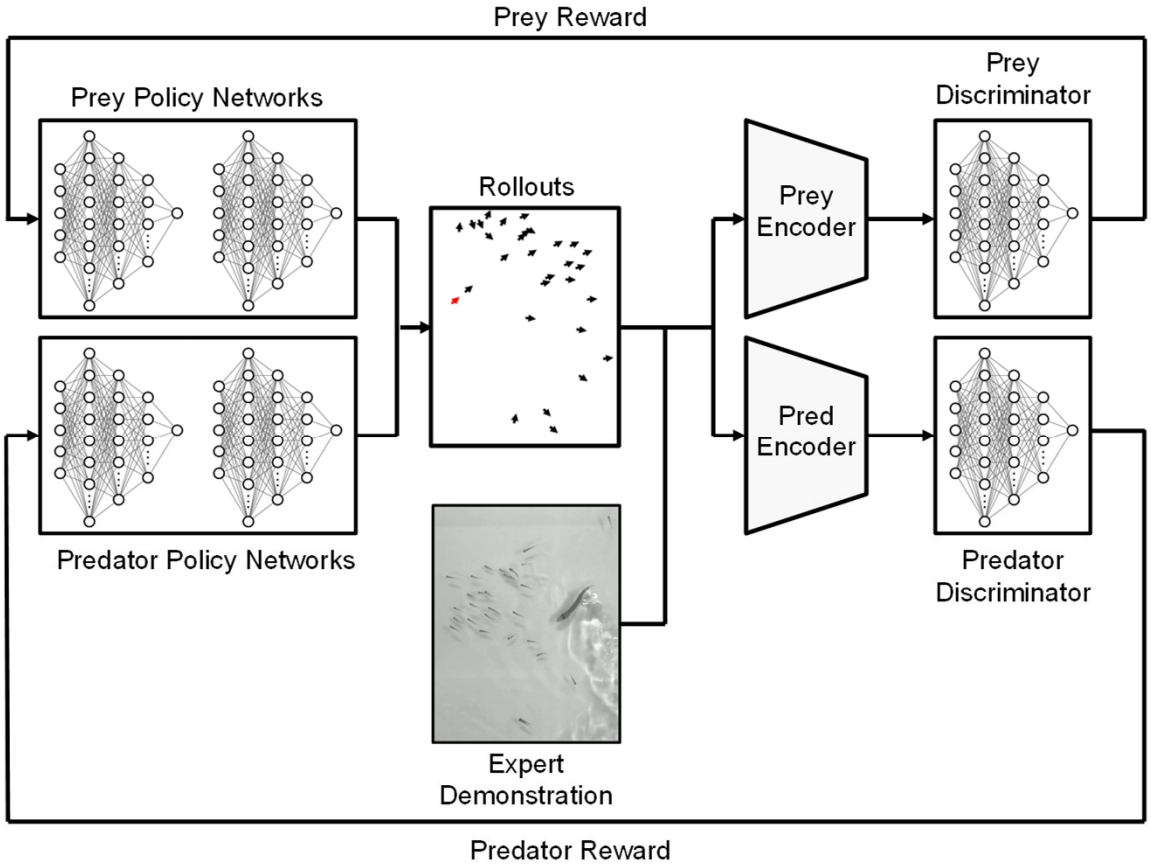
Summary & Discussion



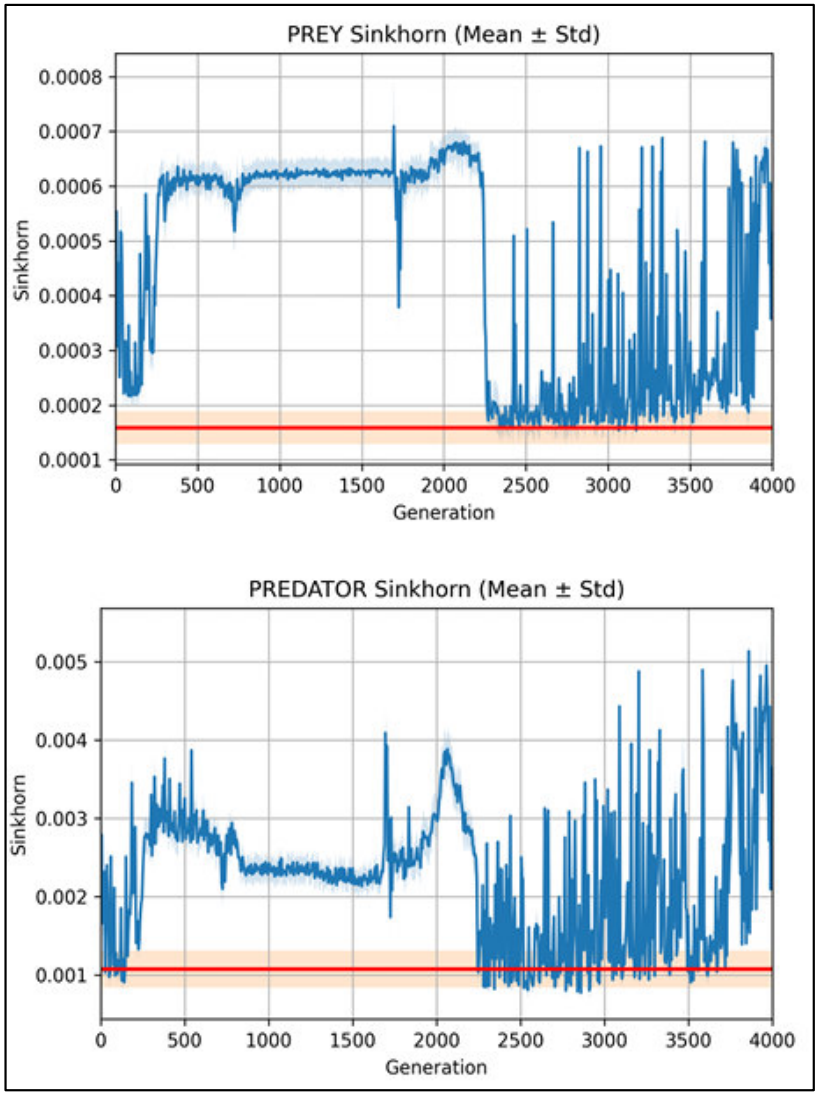
General GAIL pipeline works!
→ Prey-only model

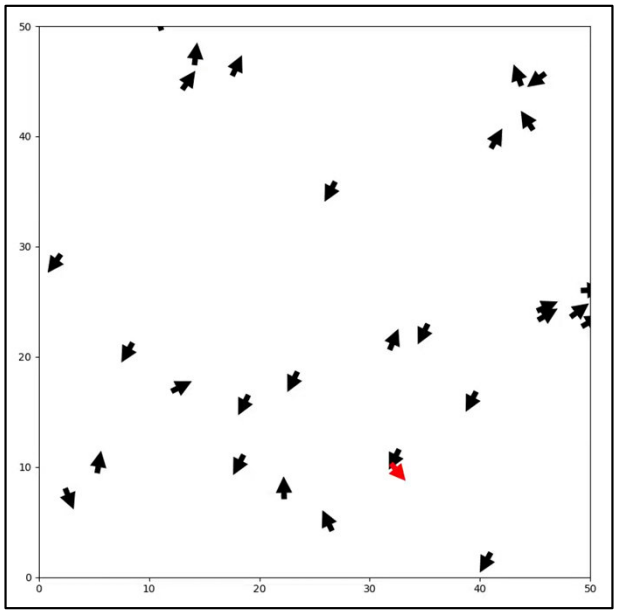
Partial imitation of predator-prey models

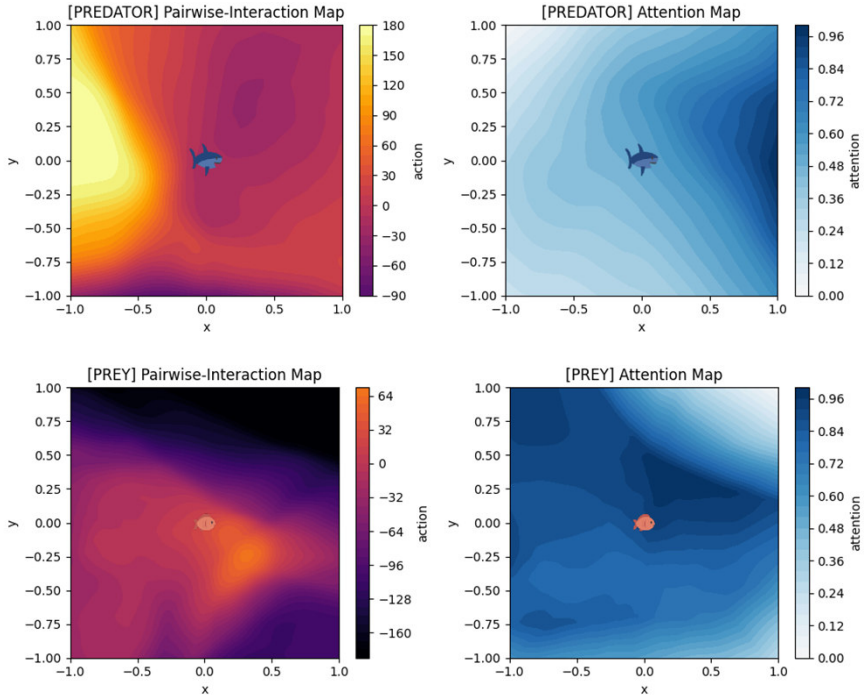
But: missing inter-group dynamics



Summary & Discussion







Limitations & Future Work

Data processing:

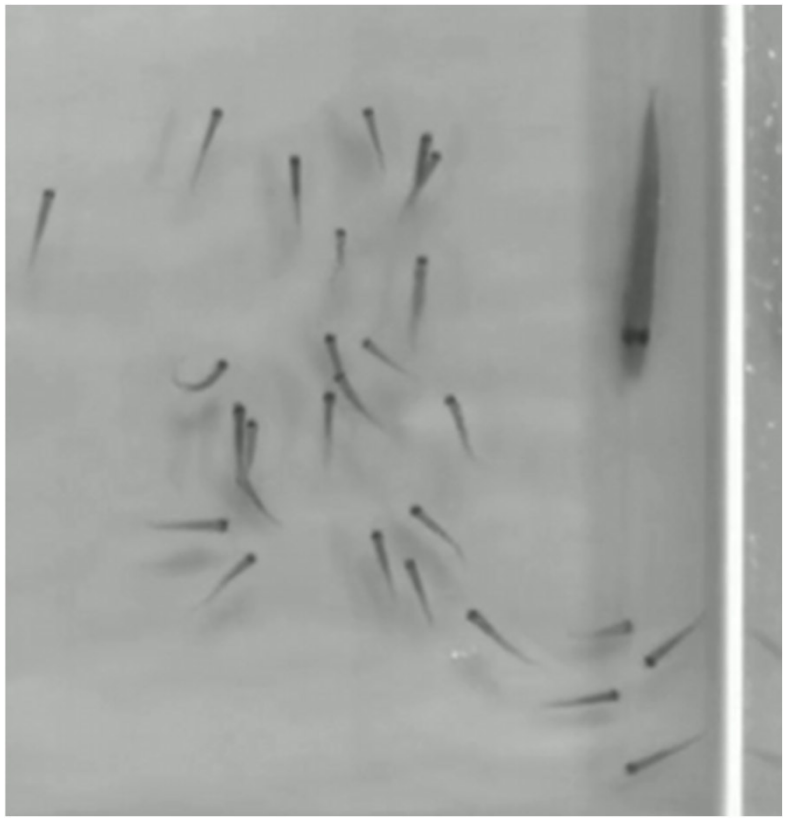
- Large fraction of data remains unused (92.81%)
- Improve tracking continuity and missing detections
- Train on longer temporal contexts and varying group sizes

Predator attacks:

- Definition of attack behavior (hectic movements, …)

Input features:

- Extend policy features with acceleration

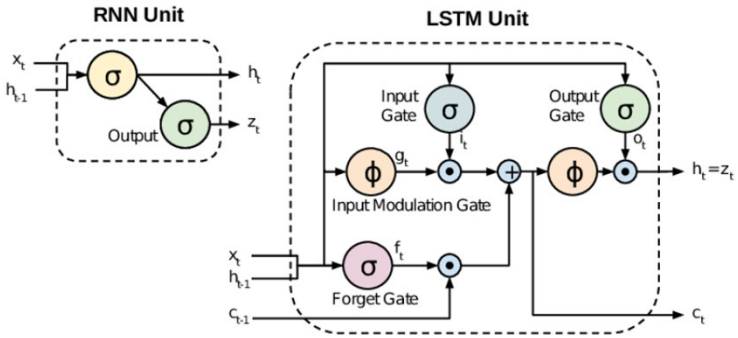


Limitations & Future Work

- Tensor Quality
- Action Aggregation
- Transition Encoder
- Feature Weighting
- Reward Shaping

Sequence processing discriminator

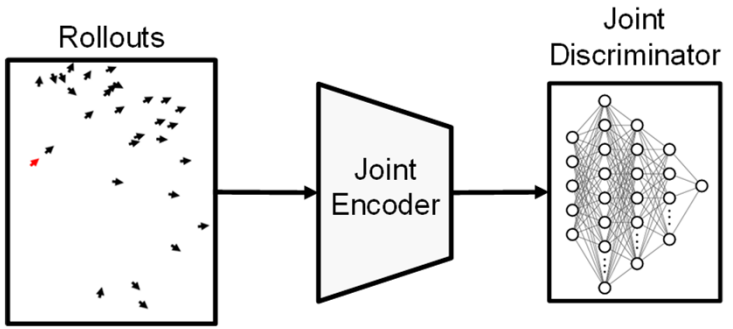
→ Modelling temporal dependencies
→ Train on longer temporal contexts
→ RRN or LSTM



https://ashutoshtripathi.com/2021/07/02/what-is-the-main-difference-between-rnn-and-lstm-nlp-rnn-vs-lstm/

Missing inter-group dynamics

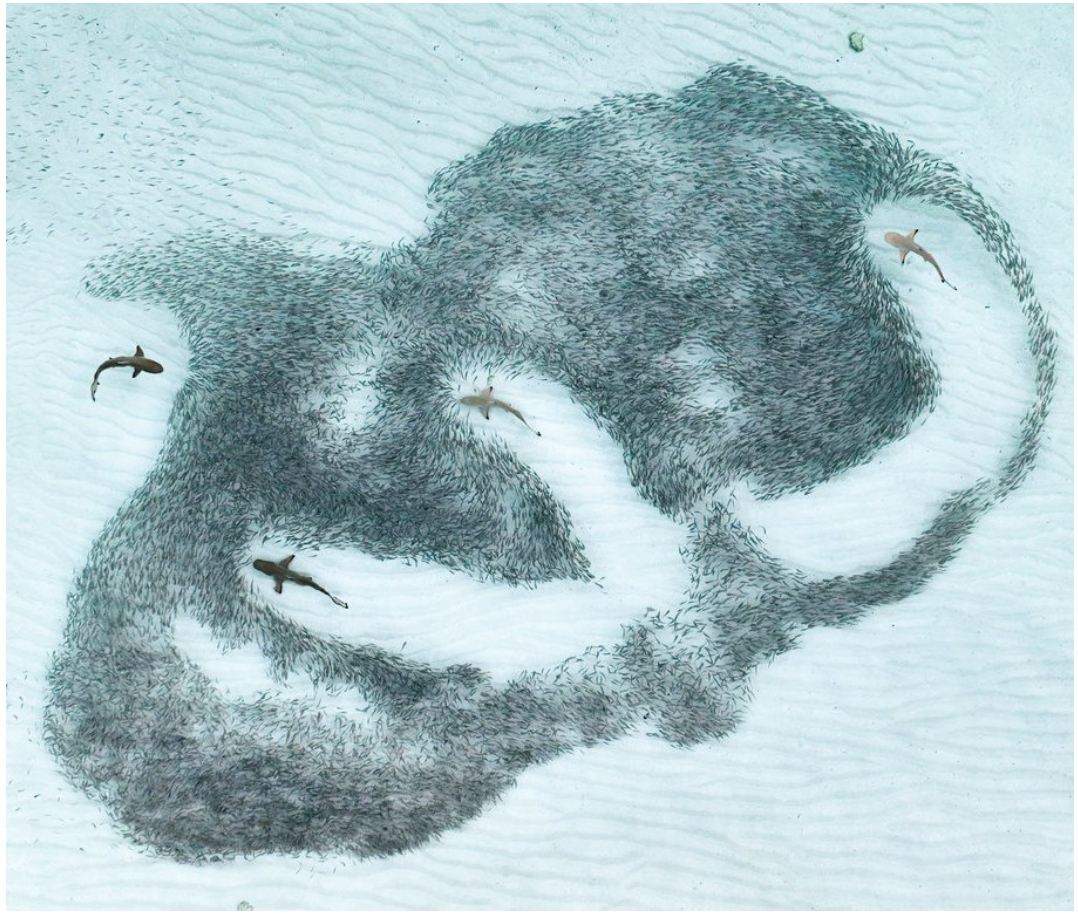
→ CBIL's multi-instance single policy
→ Joint encoder and discriminator
→ 2-stage-setup



Universität Konstanz

Thank you for your Attention!

Jannik Wirtheim
Konstanz



Reference: „The hunt from above“ – Angela Albi: https://www.campus.uni-konstanz.de/uni-leben/die-kunst-der-haie