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

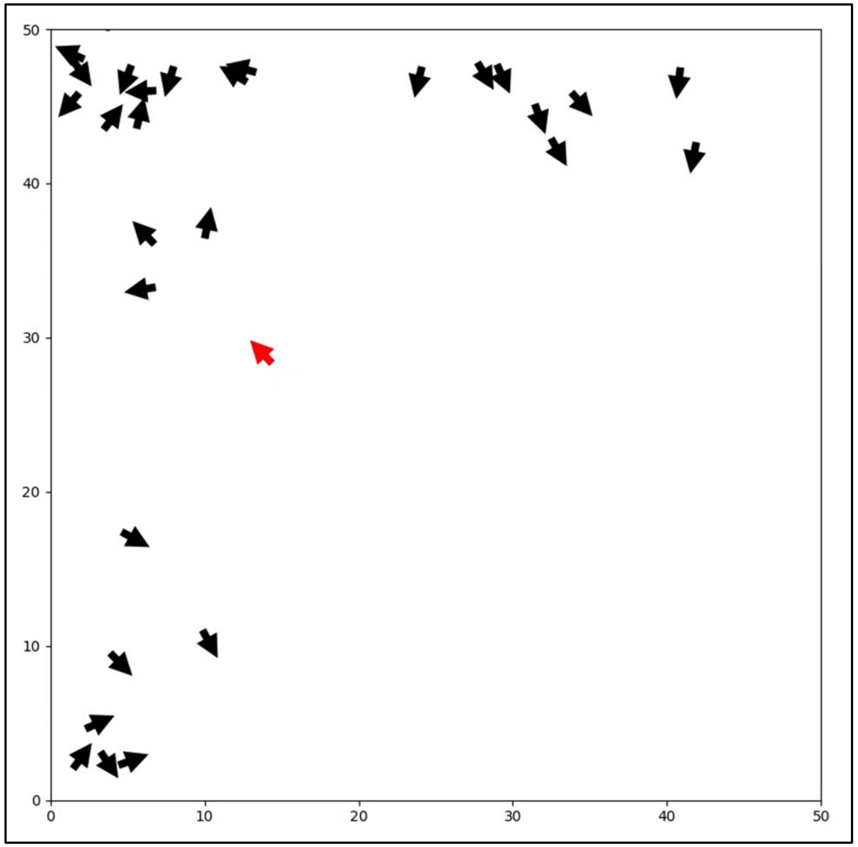
Behavior is modeled using hand-crafted interaction rules.

→ requires domain knowledge from the designer.

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

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

Hand-crafted rules



Modelling of Multi-Agent Systems

Behavior is modeled using hand-crafted interaction rules.

→ requires domain knowledge from the designer.

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

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

Hand-crafted rules

Data-driven approches

Behavior is learned by imitation from expert demonstrations.

→ desired behavior is implicitly represented in the data.

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

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

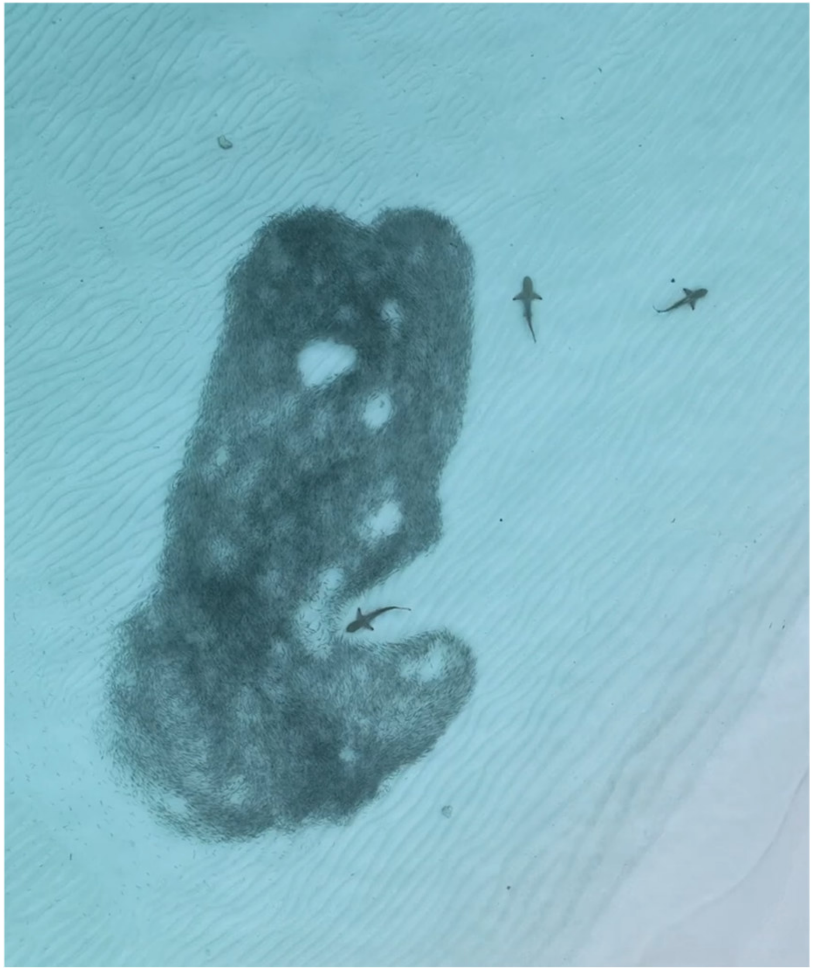
→ Confuse predator & increase survival chance

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

Data-driven Methods

https://skandavaidyanath.github.io/post/inverse-rl-paper/the_problem.png

Behavioral Cloning:
- Supervised learning
- Training on states *s*
- Learning of actions *a*

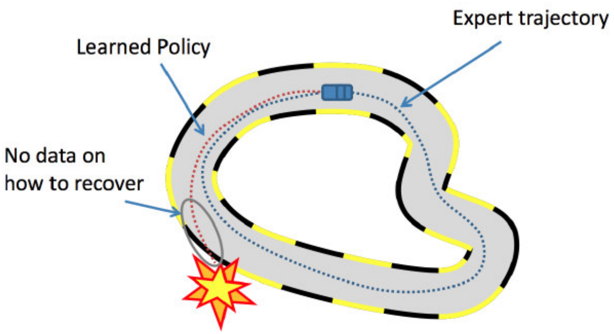
→ *Large amount of data*
→ *Compounding errors*

Inverse RL:
- Recovers reward from expert behavior
- Learns policy from reward
- Avoids manual reward design

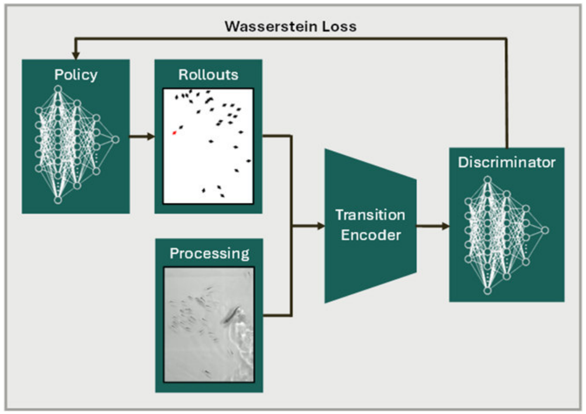
→ *Reward ambiguity*
→ *Computational cost*

GAIL:
- Adversarial learning (policy vs. discriminator)
- Distribution matching
- Scales to multi-agent imitation

→ *Data quality*
→ *Mode collapse*



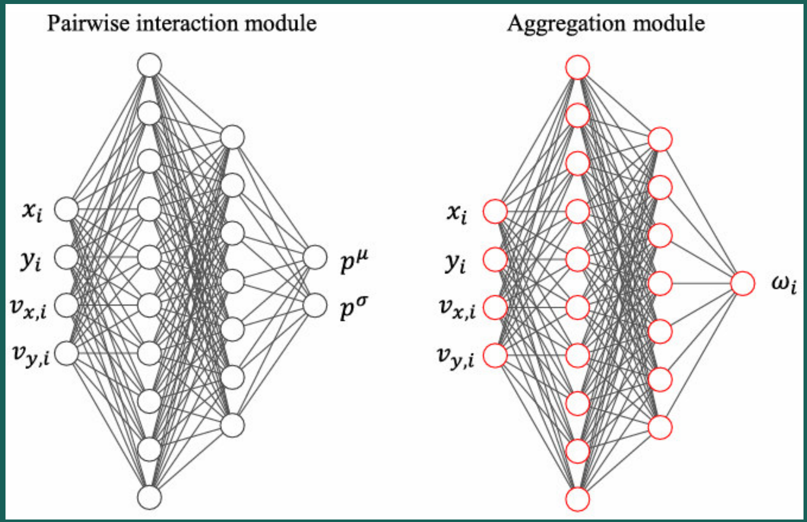
https://smartlabai.medium.com/a-brief-overview-of-imitation-learning-8a8a75c44a9c



Related Research

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

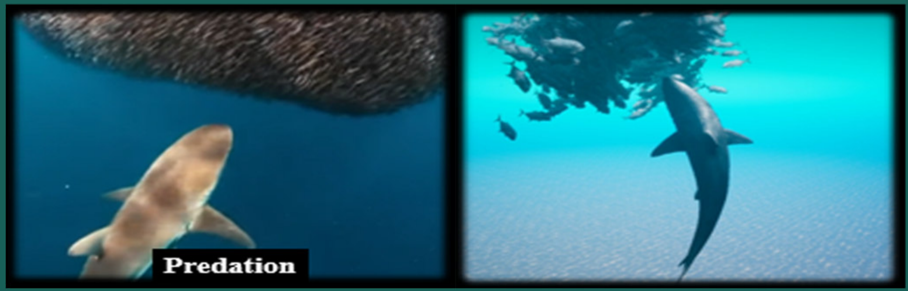
- GAIL with shared individual policy
- Couzin-based swarm demonstrations
- Same policy & tensor structure
- Limited to single-species imitation



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

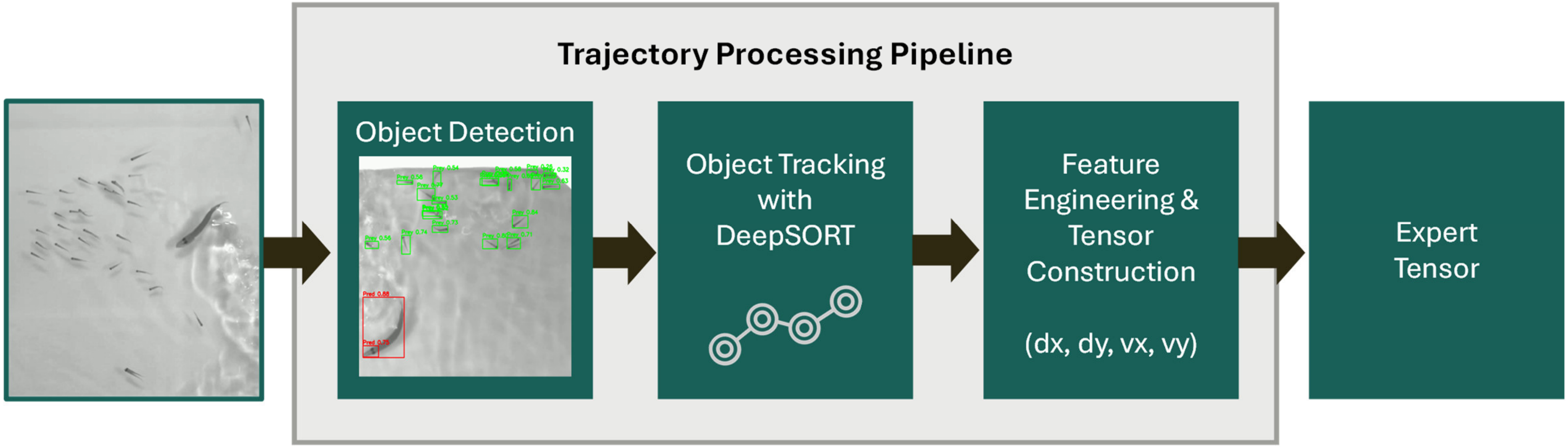
CBIL: Collective Behavior Imitation Learning for Fish from Real Videos (Yifan Wu et al., 2025)

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

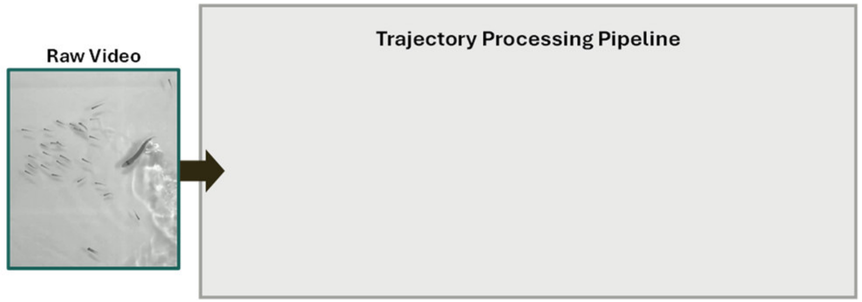


(Yifan Wu et al., 2025, p.9)

Data Collection & Processing



Data Collection & Processing



Recording of predator-prey aquarium:

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

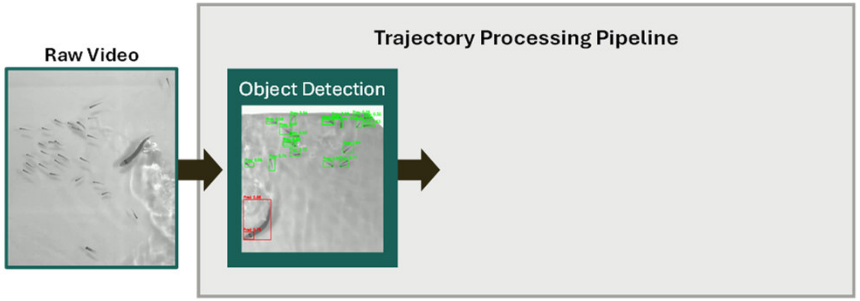
Video-related overview:

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

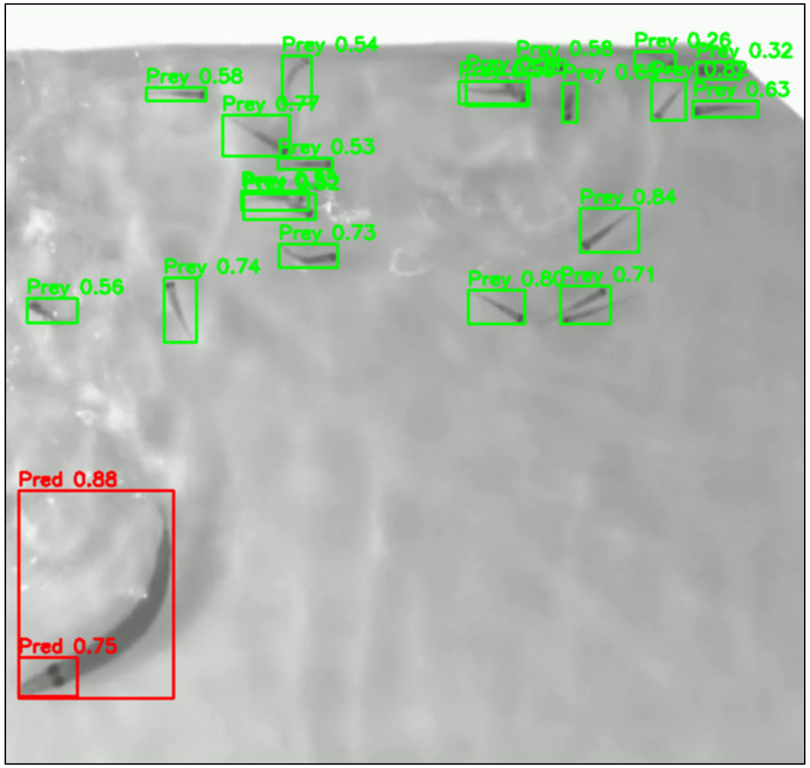
Predator-related overview:

- 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

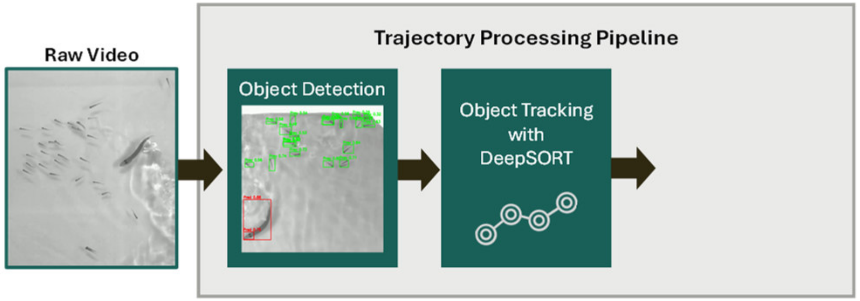


Detection error:
- Predator: MAE ± 0.03
- Prey: MAE ± 3.11

Error causes:
- Dense groups
- Occlusions
- Predation success

→ Attack scenes manually labeled

Data Collection & Processing

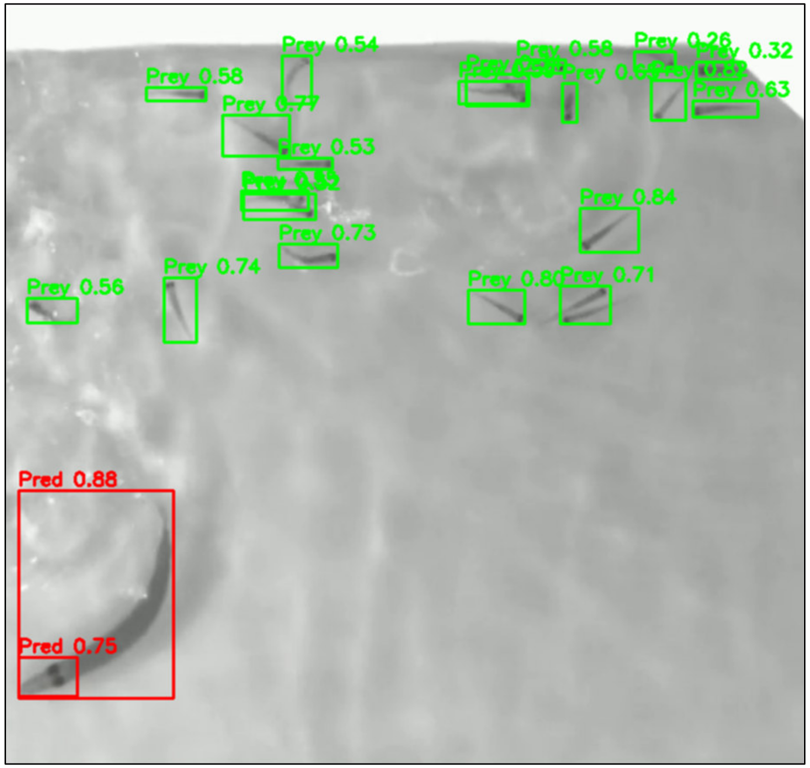


Object tracking (DeepSORT):

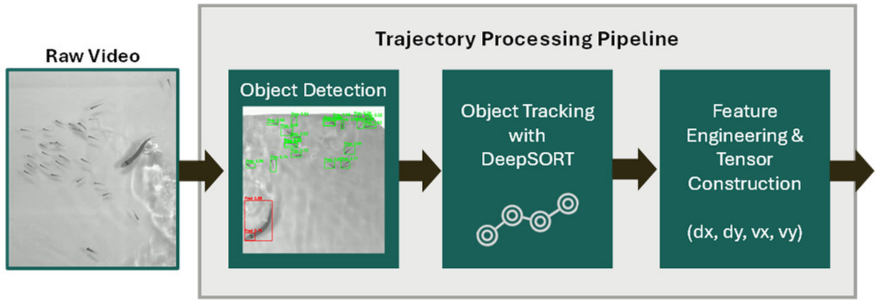
- Data association via Hungarian algorithm
- Kalman filter for state estimation and missed detections
- Produces consistent track IDs over time

Results:

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



Data Collection & Processing



State / action representation:

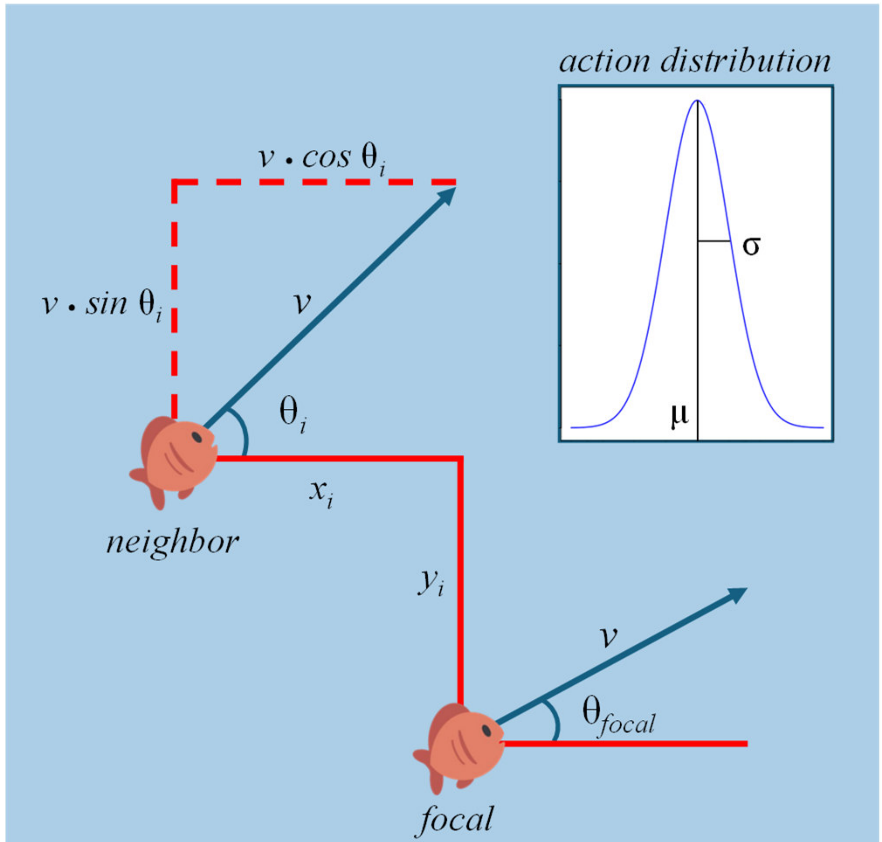
- Pairwise distances (dx, dy)
- Relative velocities (vx, vy)
- Action $\Delta\theta$

→ [dx, dy, vx, vy, θ]

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

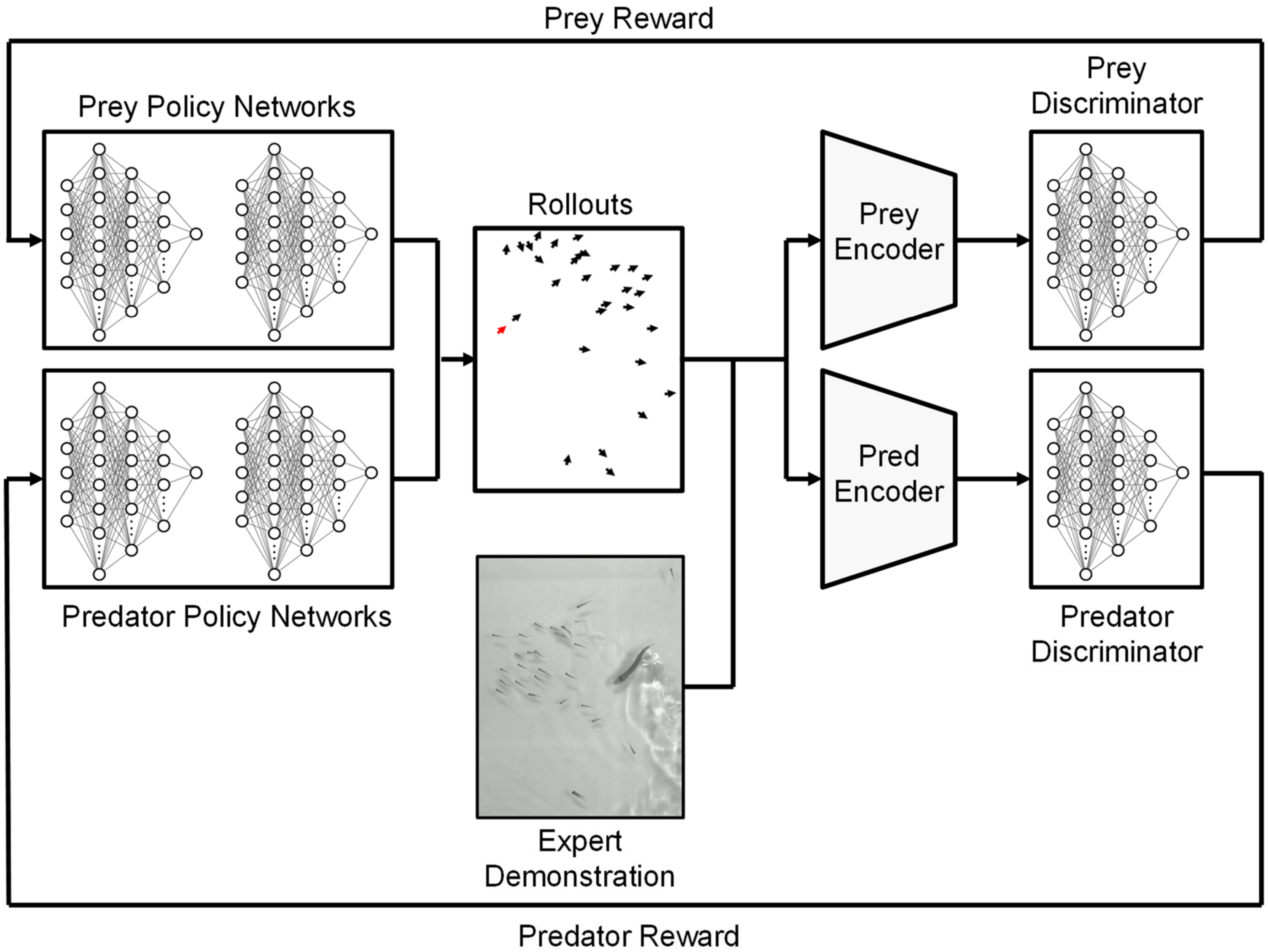
$$\text{Prey Tensor} \in \mathbb{R}^{793\times10\times32\times32\times6}$$

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

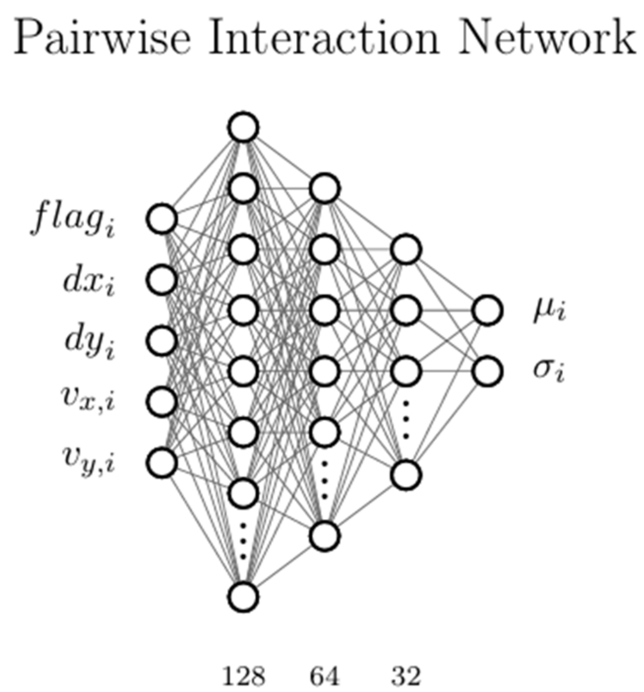


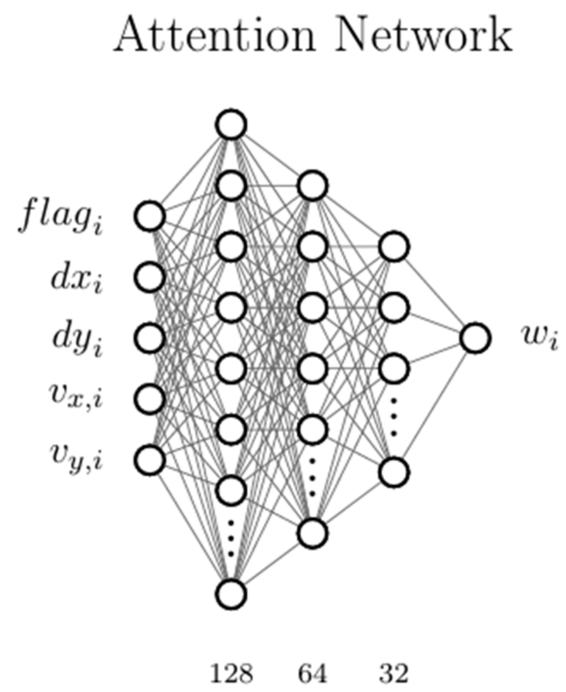
(Yapei Wu et al., 2025, p.4)

Methodology: GAIL



Modular Policy





Pairwise Interaction Network:

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

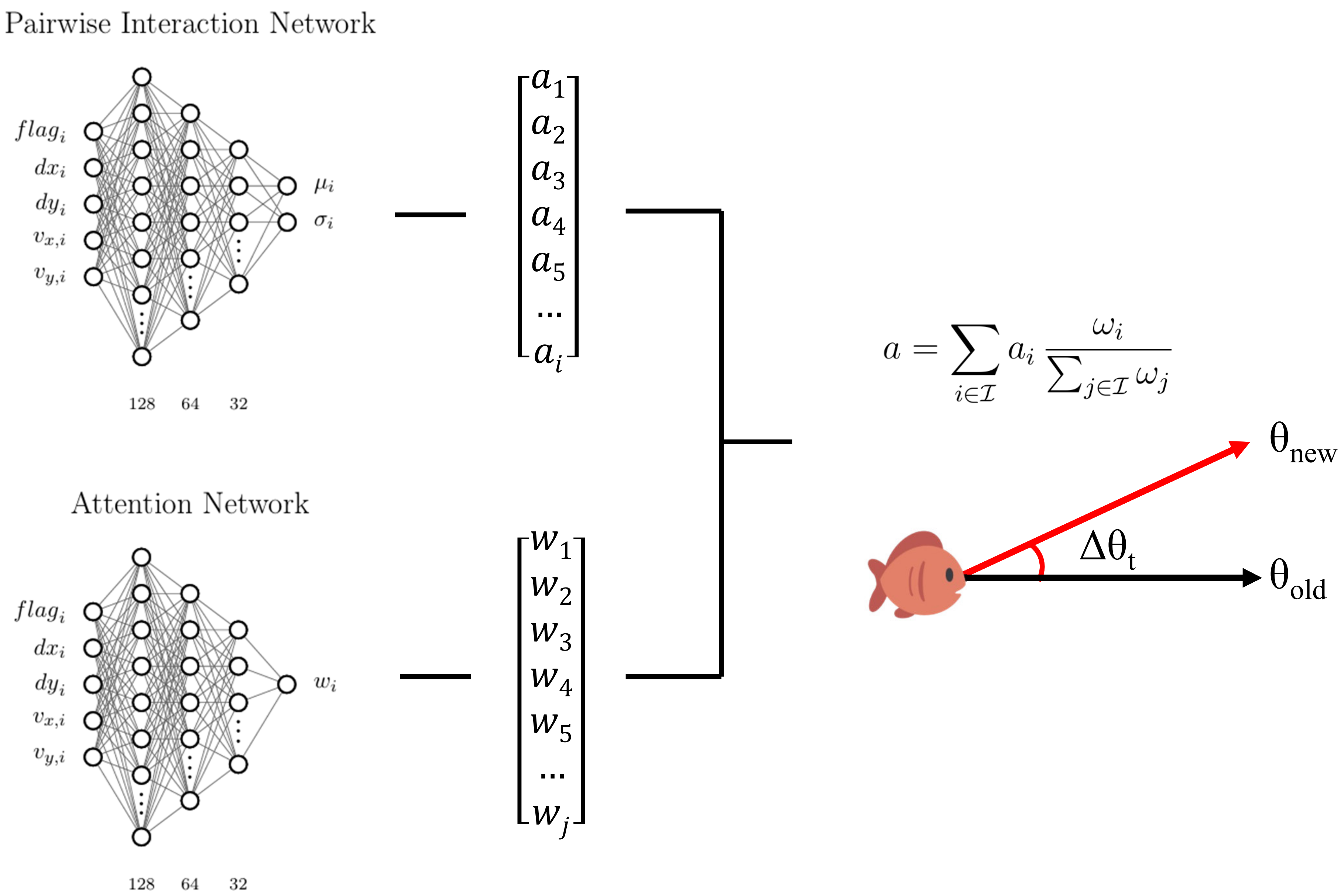
→ **stochastic individual response**

Attention Network:

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

→ **interpretable interaction structure**

Modular Policy

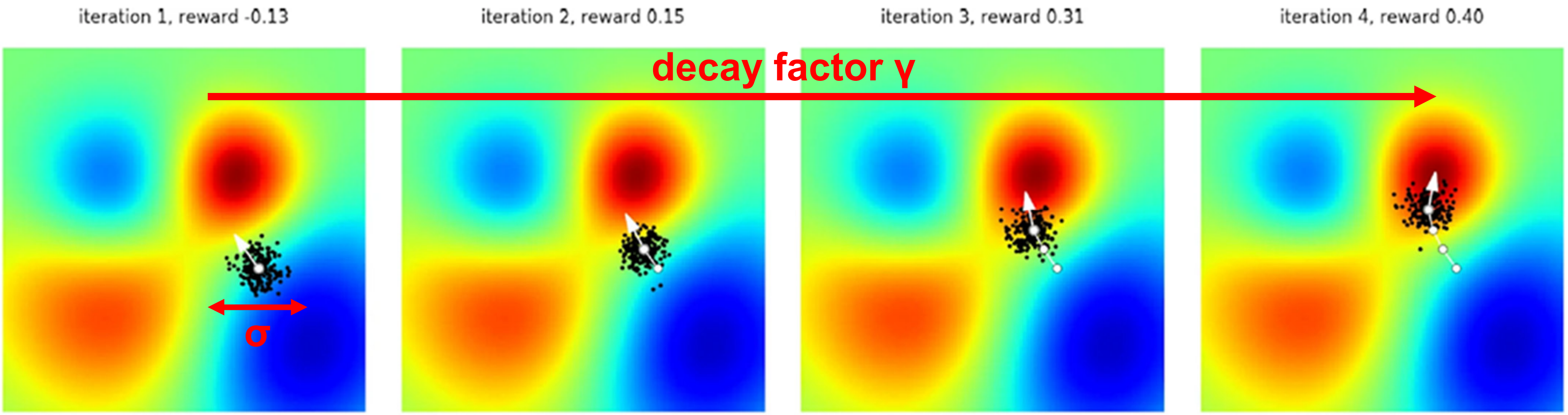


Modular Policy – Evolutionary Strategy

"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).

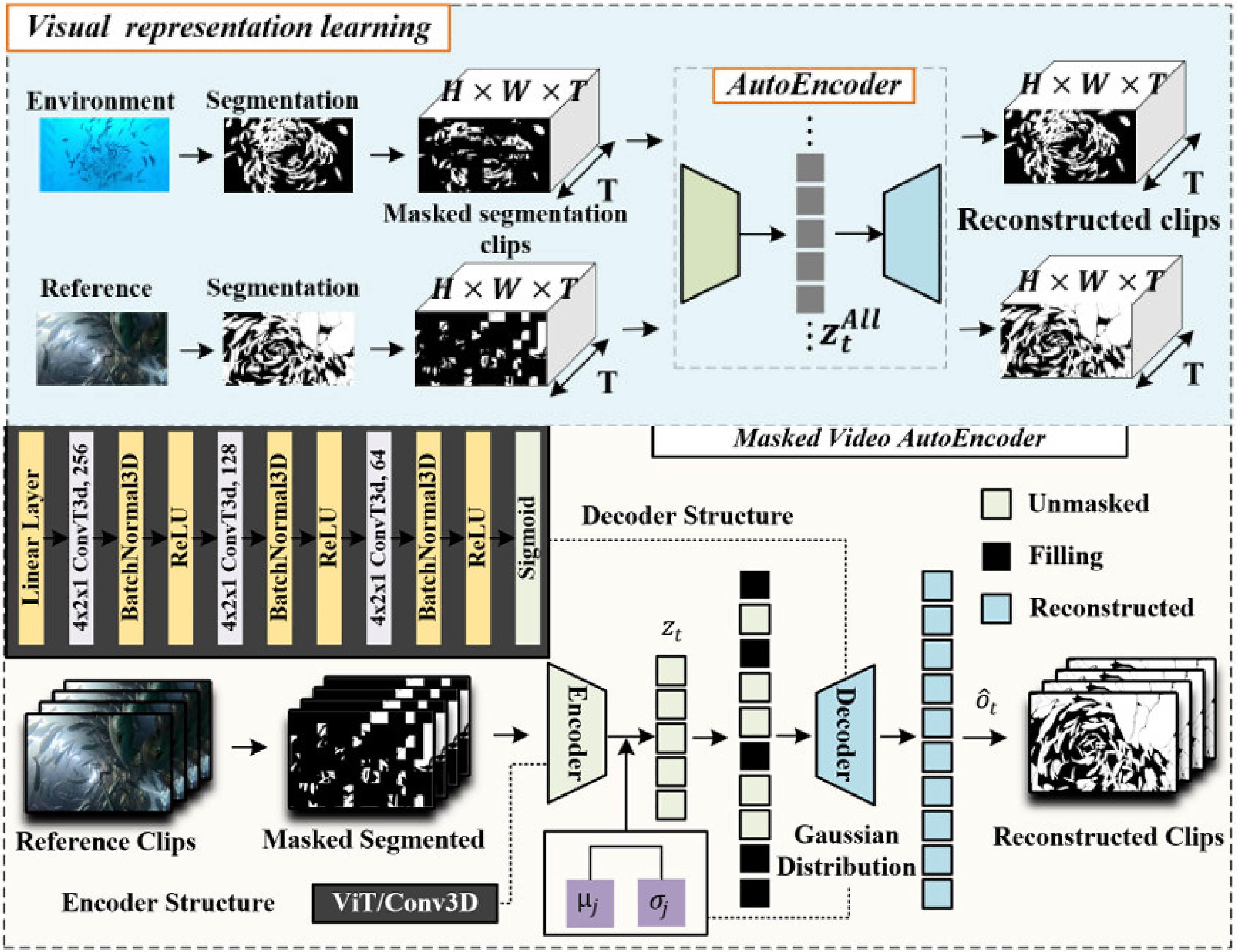
1. Stage:
- Sample perturbation noise
- Apply perturbations to parameters
- Collect rollout rewards

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



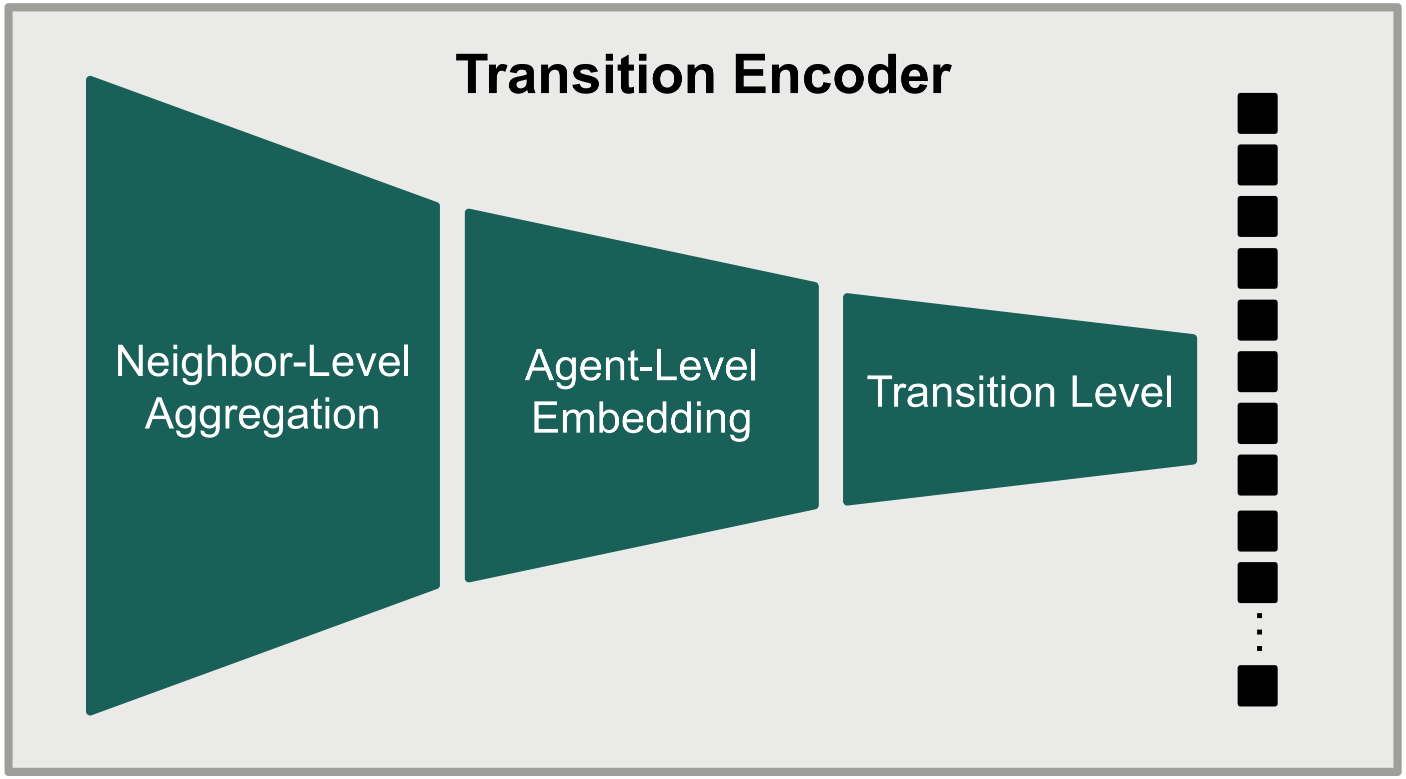
https://images.ctfassets.net/kftzwdyauwt9/d5acb8a0-a1a1-4772-09f9e1a6550b/e7bccd4dd7532331595032ac7b9e3f14/evo.png

CBIL Transition Encoder



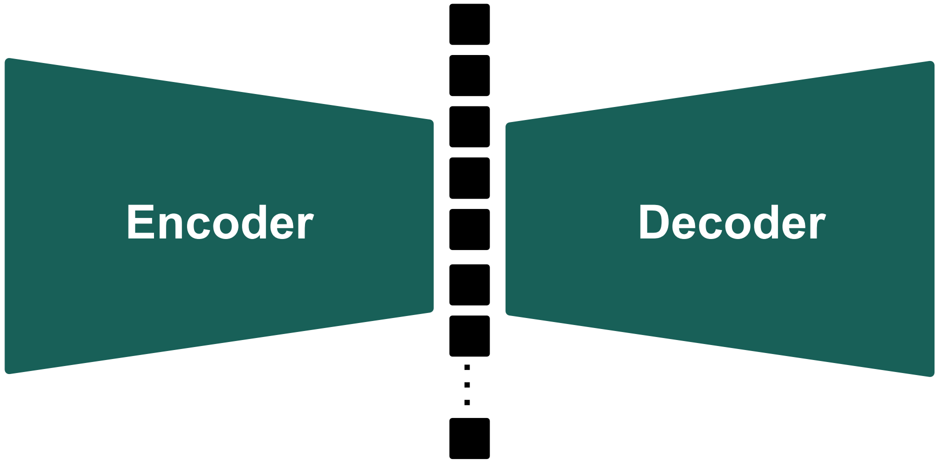
Transition Encoder

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



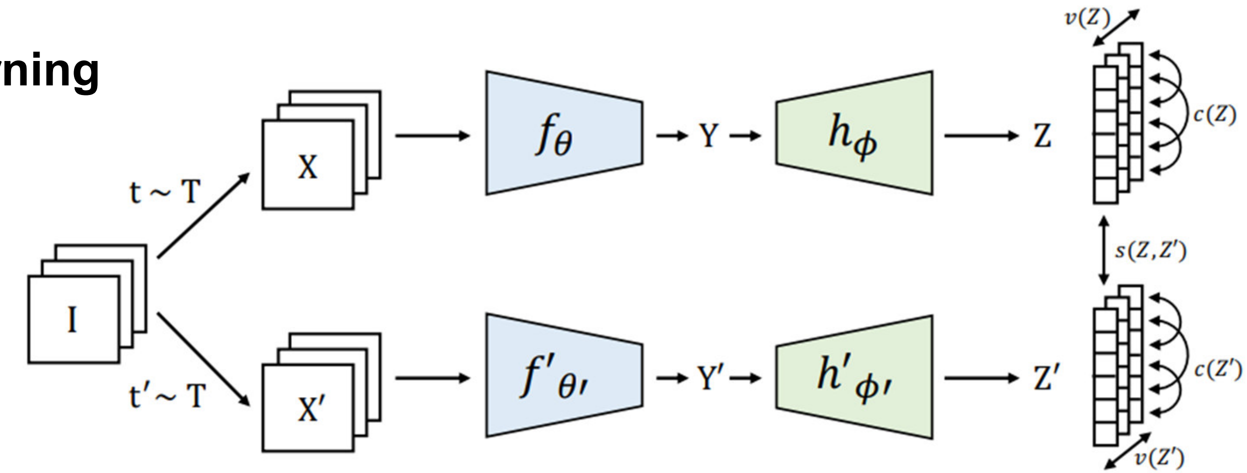
Output tensor: [batch, windows-1, agents, 2z]
Transition Feature (2z) = $[z_t, \Delta z_t]$

Transition Encoder - Training



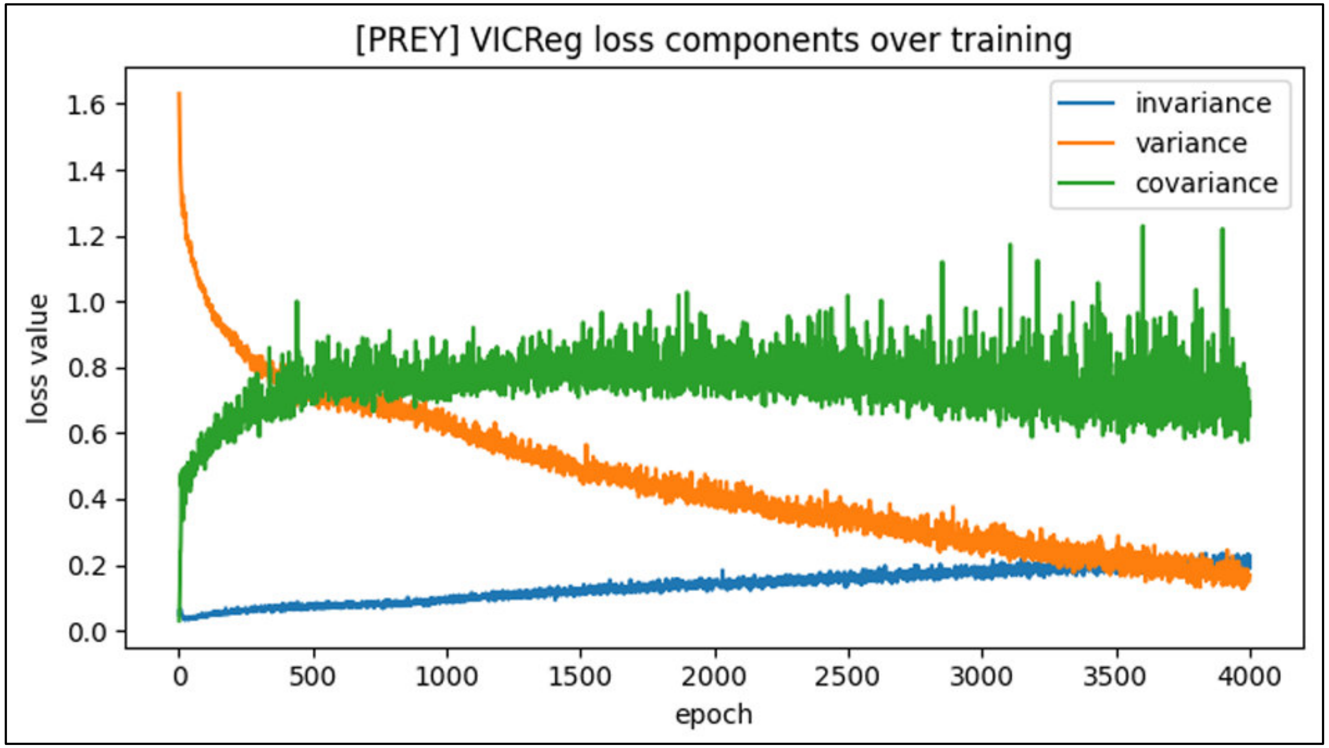
Problems with reconstruction-based training
- Poor reconstruction loss → weak latent?
- Decoder not required for GAIL
- Question: Training possible without reconstruction? → **VICReg**

Self-supervised representation learning
- No decoder, direct latent training
- Prevents representation collapse
- Augmentation:
  - State noise
  - Neighbor dropout
  - Feature dropout

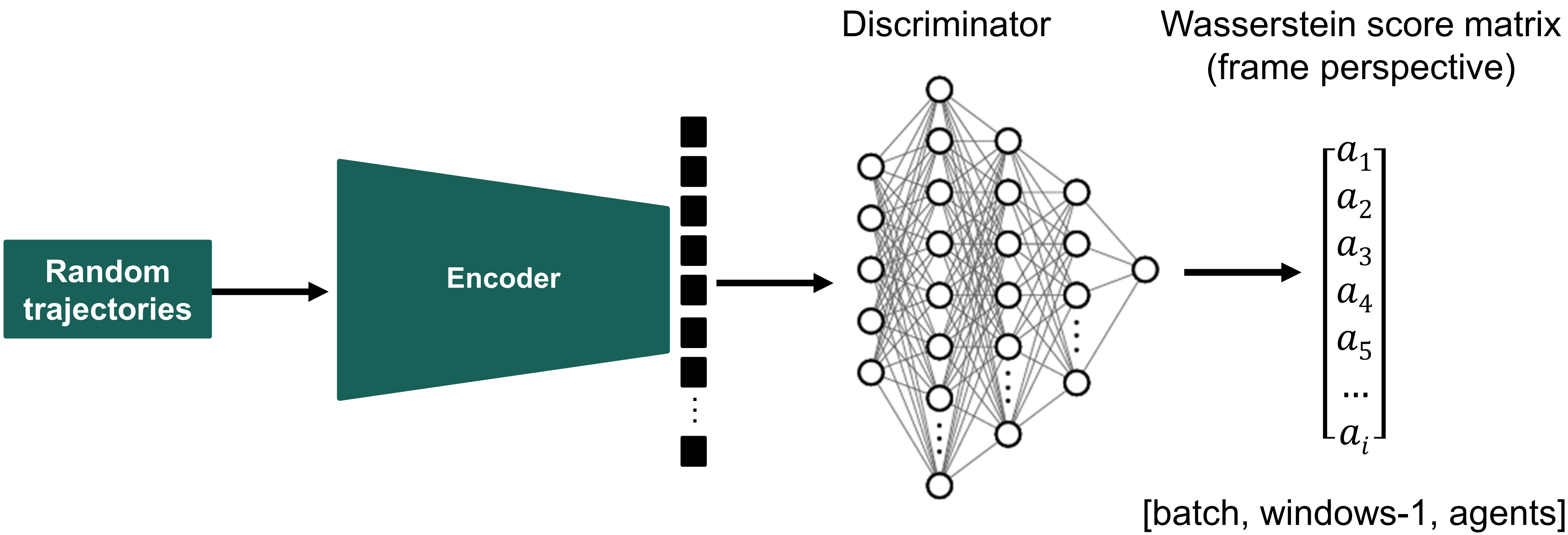


Variance-Invariance-Covariance Regularization

Invariance	MSE between embeddings of augmented views → enforces representation consistency
Variance	Preserve diversity across samples. → Prevents representation collapse
Covariance	Decorrelates embedding dimensions. → Reduces redundancy



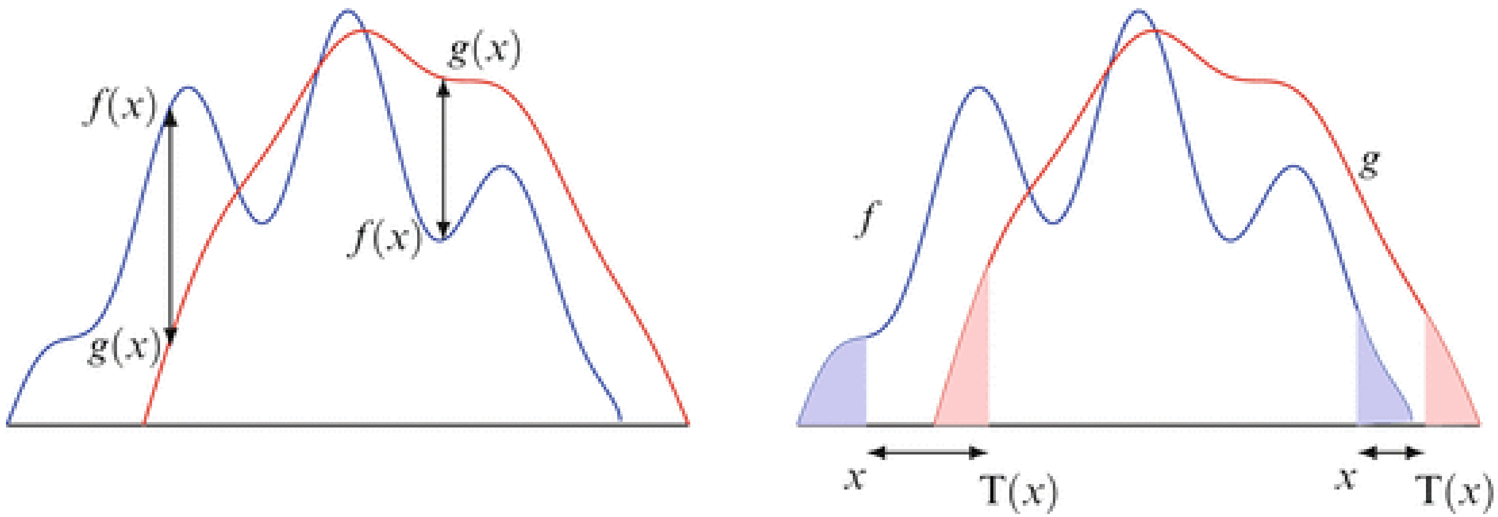
Discriminator



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_π to τ_E in order to transform the distributions P_π into the distribution P_E (Arjovsky et al., 2017, p. 4)”.



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

Gradient Penalty & 1-Lipschitz Constraint

1-Lipschitz Constraint

- Required for a meaningful Wasserstein estimate
- Prevents arbitrarily large score changes for small input differences

$$|D(\tau_E) - D(\tau_\pi)| \leq ||\tau_E - \tau_\pi||$$

Gradient Penalty:

- Approximates the 1-Lipschitz constraint on the discriminator
- Penalizes deviations of $\|\nabla D(\tau)\|$ from 1
- Ensures stable Wasserstein estimation and training

$$\mathcal{L}_{GP} = \mathbb{E}_{\tau\sim\zeta}\left[(\|\nabla_\tau D_\beta(\tau)\| - 1)^2\right]$$

Environment

Base Environment

Analysis and visualization of trained policy networks.

Properties:
- Non-vectorized
- Slower

Used for:
- Policy evaluation
- Metric computation
- Visualization

Vectorized Environment

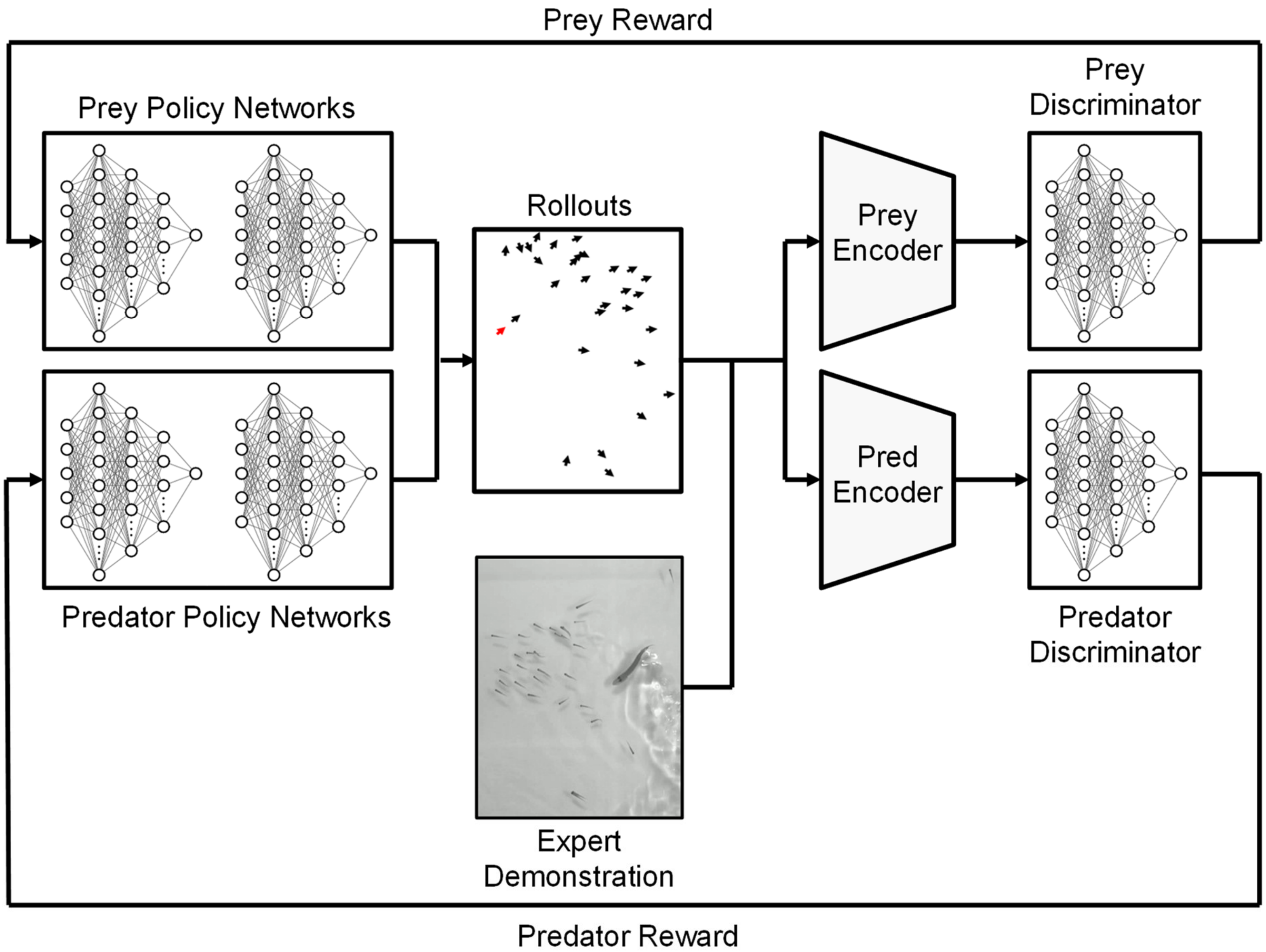
Efficient and fast ES training

Properties:
- Fully vectorized batch simulation
- Tightly coupled to ES rollouts
- GPU-accelerated execution

Used for:
- Parallel evaluation of policy perturbations

Both:
- Initialized through expert states
- Bouncing at walls

Training Stabilization

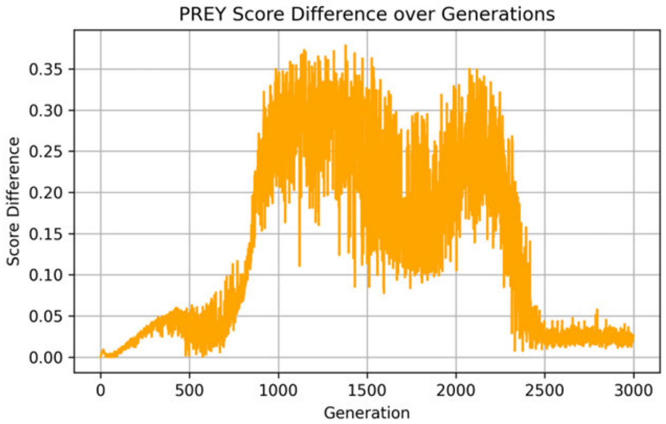


Training Stabilization

ES-Clipping	Limits update magnitude, avoids exploding gradients
EMA	Smooths noisy policy updates over generations
Discriminator-Update-Ratio	Balances discriminator–policy learning
Discriminator Noise	Reduces overfitting to expert vs. generated data
BC-pretraining	Provides informed initialization of policies
Decay factor	Gradually reduces exploration and update scale

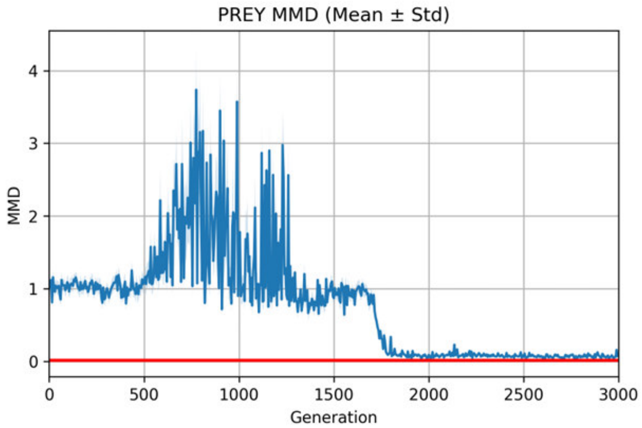
Performance Evaluation

Wasserstein proxy



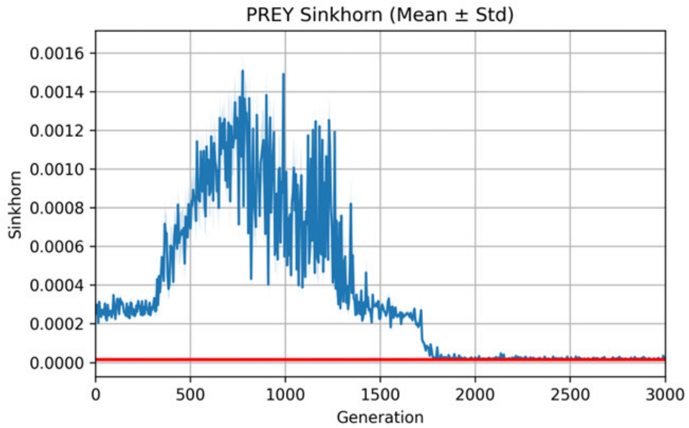
Discriminator score differences

Maximum Mean Discrepancy (MMD)



Distance between sampled batches

Sinkhorn distance



Distance between transition features

Training Procedure

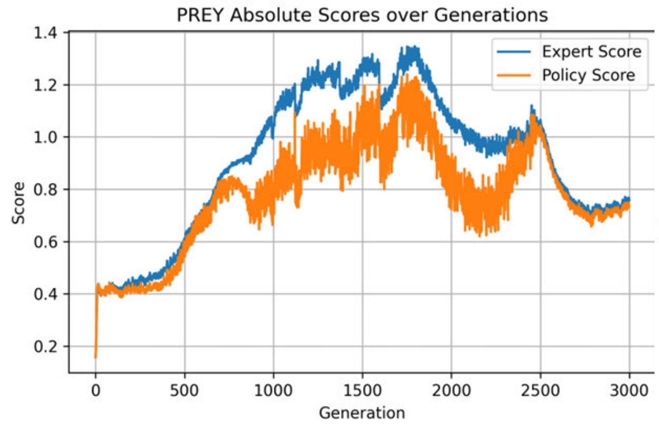
Algorithm 1 GAIL procedure to imitate predator-prey dynamics

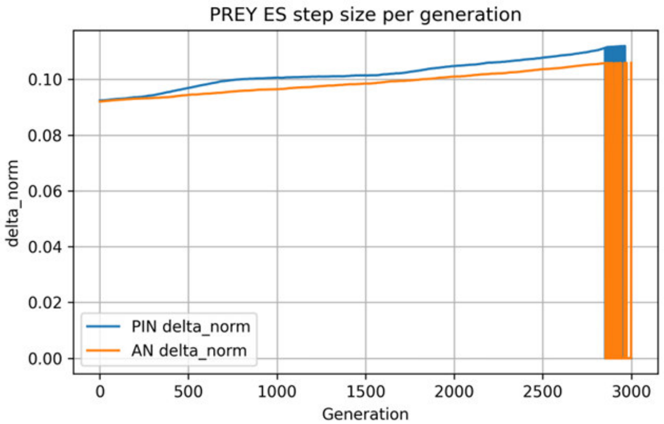
Input: Policies $\pi_{P_{\phi,\varphi}}$ and $\pi_{Y_{\phi,\varphi}}$, Discriminators $D_{P,\beta}$ and $D_{Y,\beta}$,
Encoders $E_{P,\psi}$ and $E_{Y,\psi}$, Expert trajectories $\tau_{\mathrm{E,P}}$ and $\tau_{\mathrm{E,Y}}$

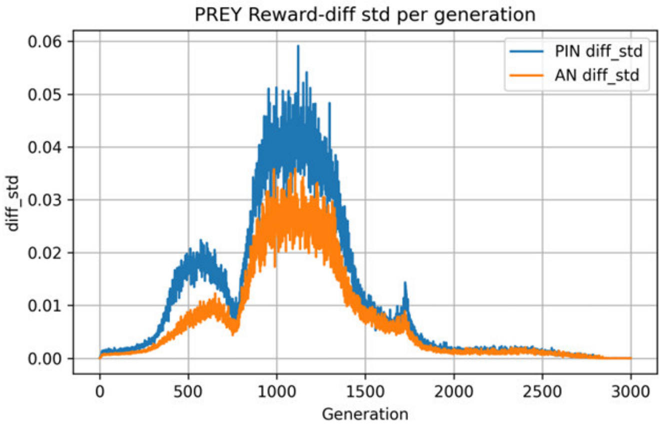
Output: Best predator $\pi^{\star}_{P_{\phi,\varphi}}$ and prey policy $\pi^{\star}_{Y_{\phi,\varphi}}$

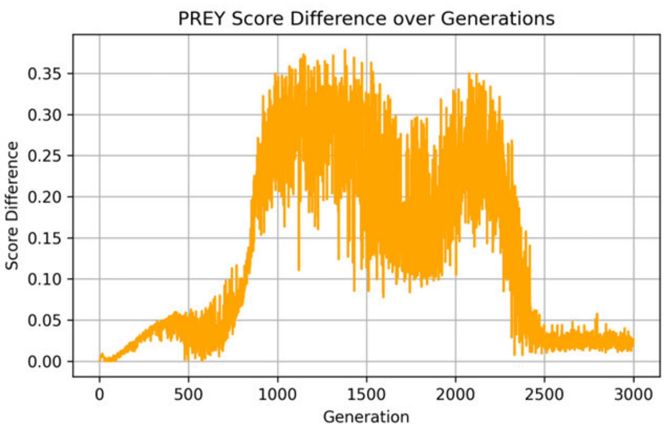
```
Pretrain π_{P_{φ,φ}} and π_{Y_{φ,φ}} with BC
Train and freeze E_{P,ψ} and E_{Y,ψ}
Initialize EMA policies π̄_{P_{φ,φ}} and π̄_{Y_{φ,φ}}
for i = 0, 1, 2, ..., M generations do
    Execute π̄_{P_{φ,φ}} and π̄_{Y_{φ,φ}} to collect rollouts
    for j = 1, ..., J D_{P,β}-update-ratio do
        update predator discriminator D_{P,β}
    end for
    for j = 1, ..., J D_{Y,β}-update-ratio do
        update prey discriminator D_{Y,β}
    end for
    optimize predator PIN π_{P_φ} and AN π_{P_φ} with ES
    optimize prey PIN π_{Y_φ} and AN π_{Y_φ} with ES
    update EMA: π̄_{P_{φ,φ}} ← EMA(π_{P_{φ,φ}}), π̄_{Y_{φ,φ}} ← EMA(π_{Y_{φ,φ}})
    apply decay γ on learning rates λ_P, λ_Y and exploration factor σ_P, σ_Y
end for
```

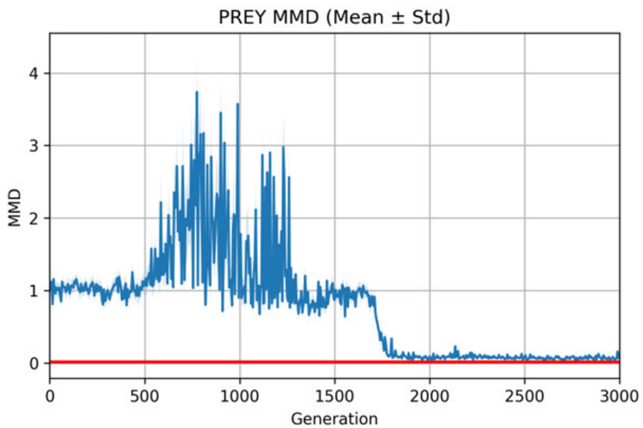
Prey-Only Model

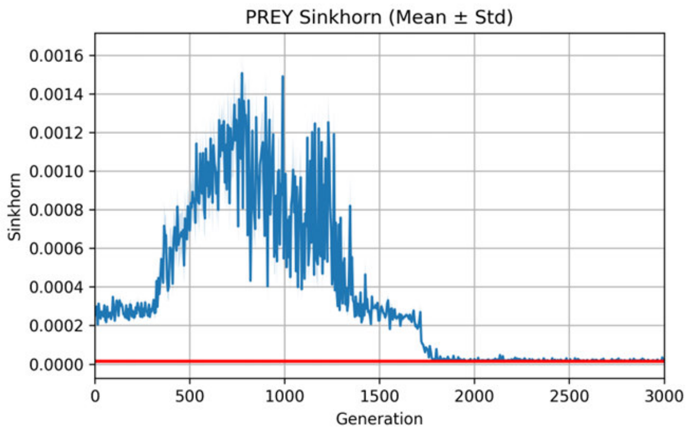






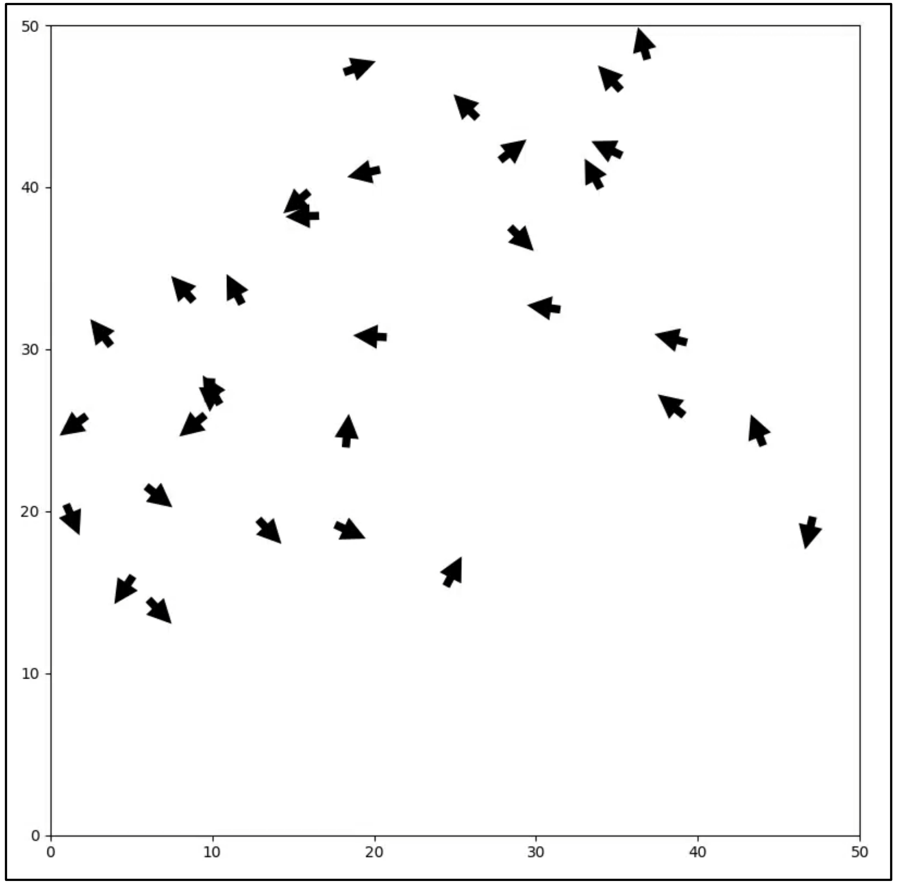




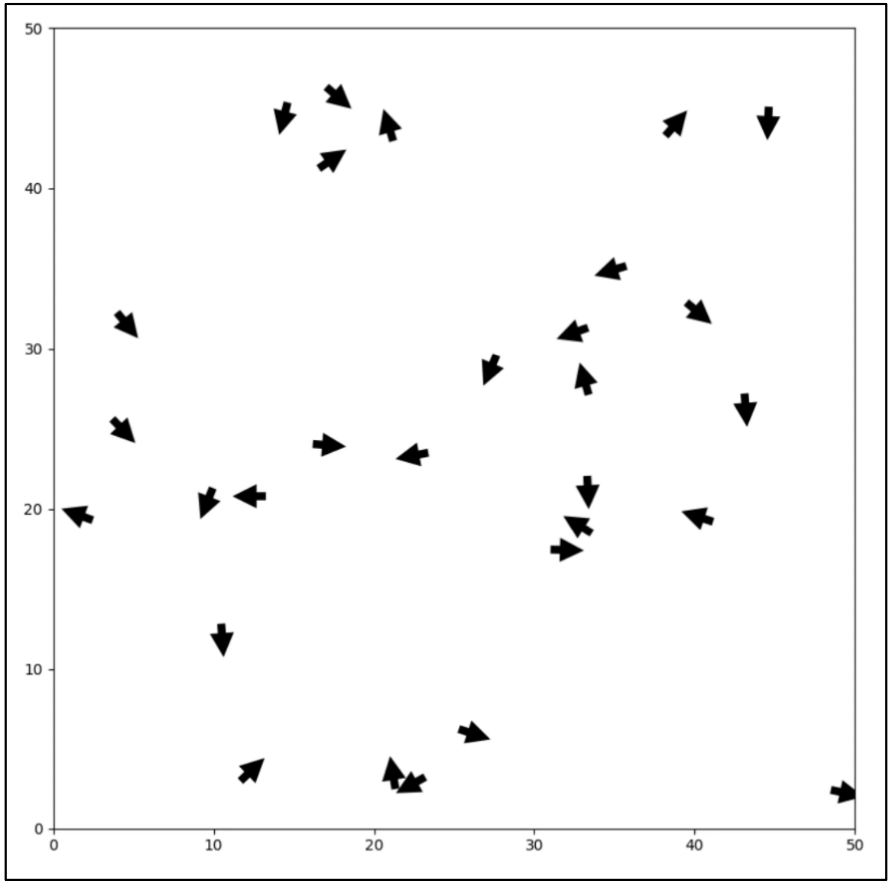


Prey-Only Model

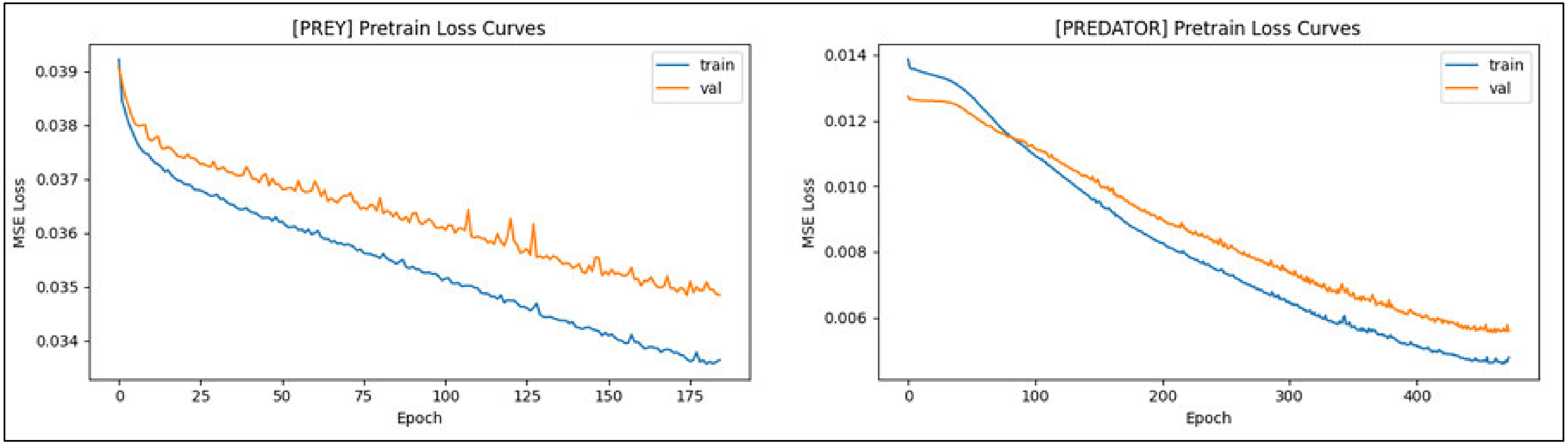
Expert demonstrations
Couzin model



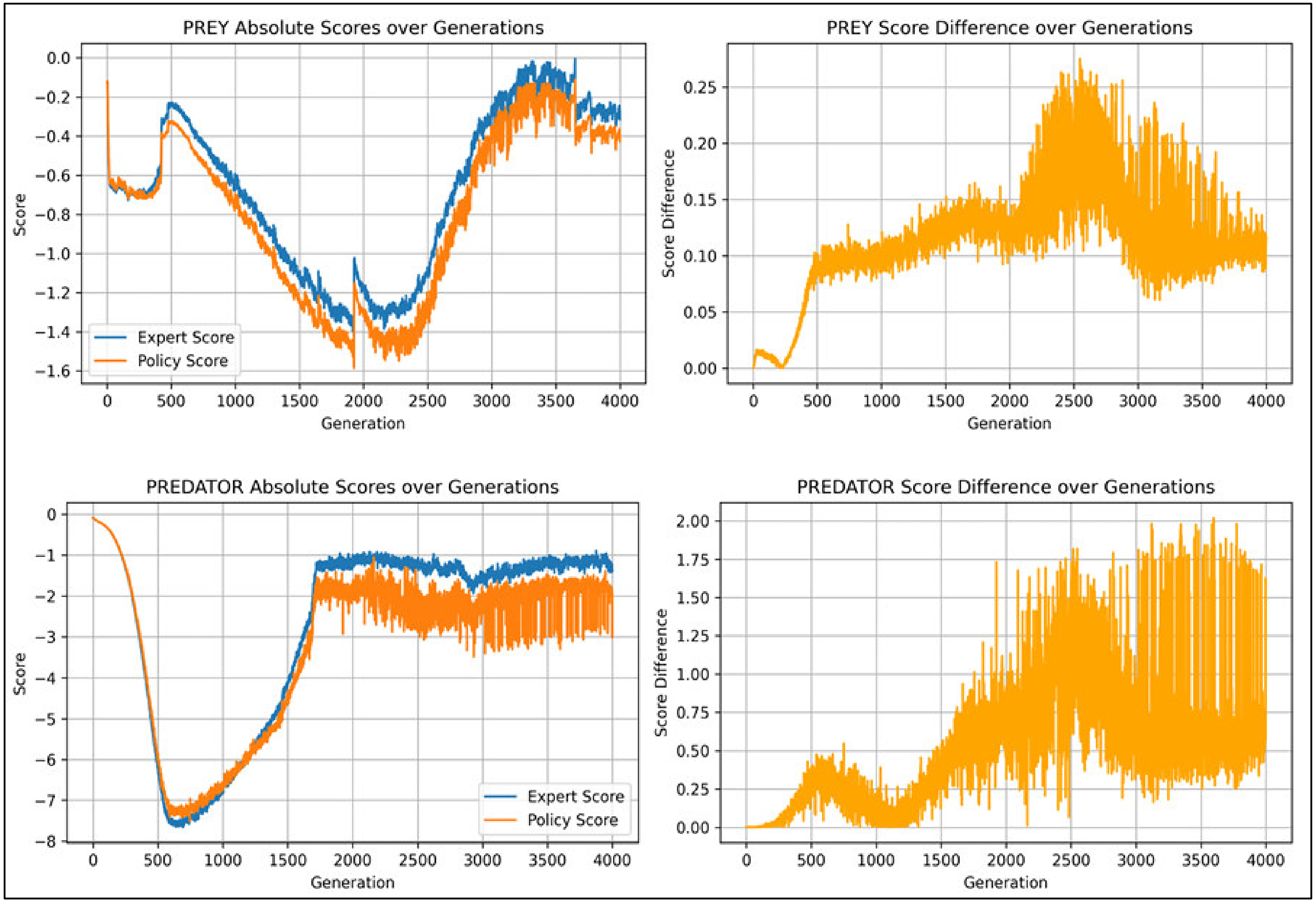
Policy-generated
Prey-only GAIL model



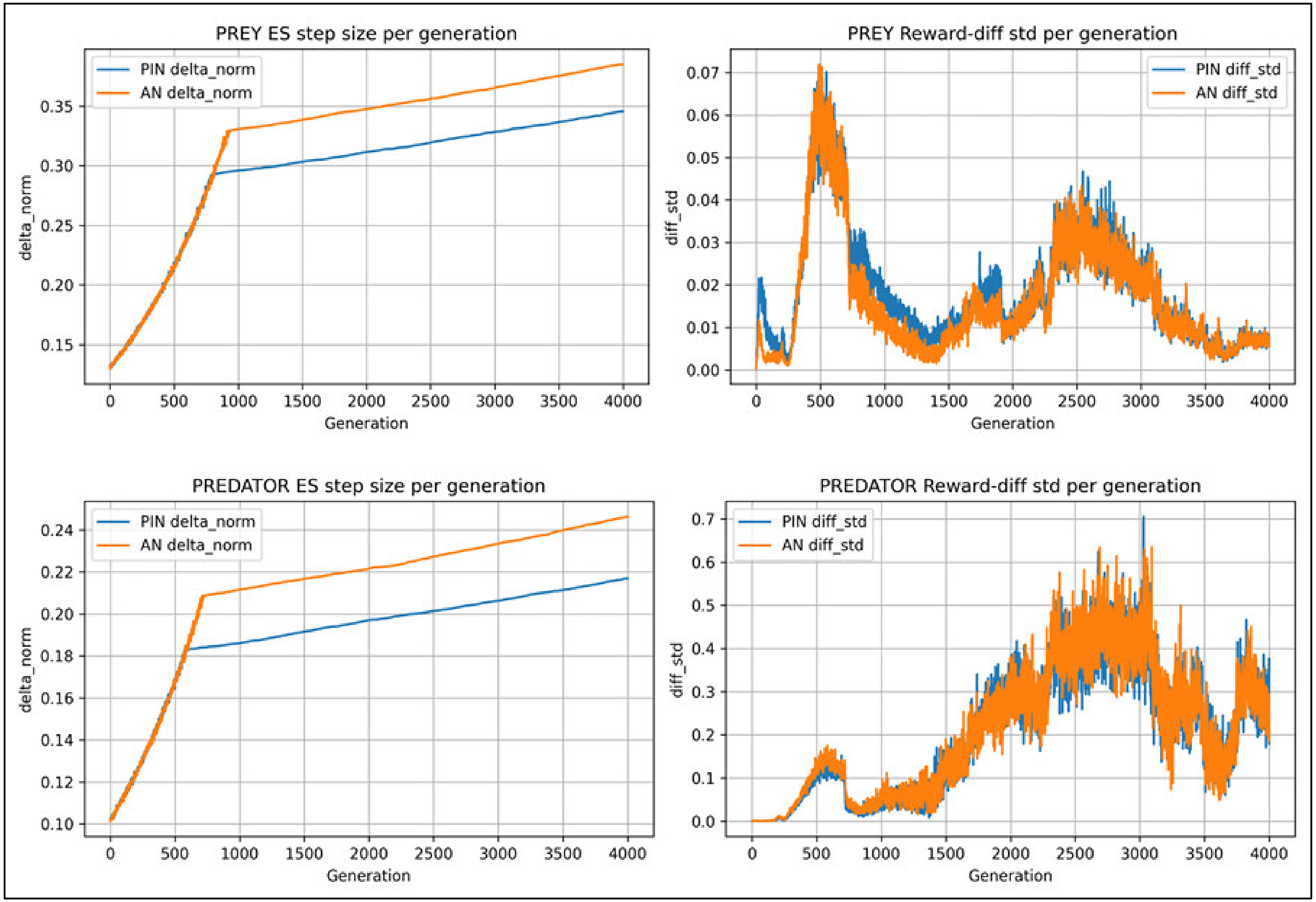
Video Predator-Prey Model



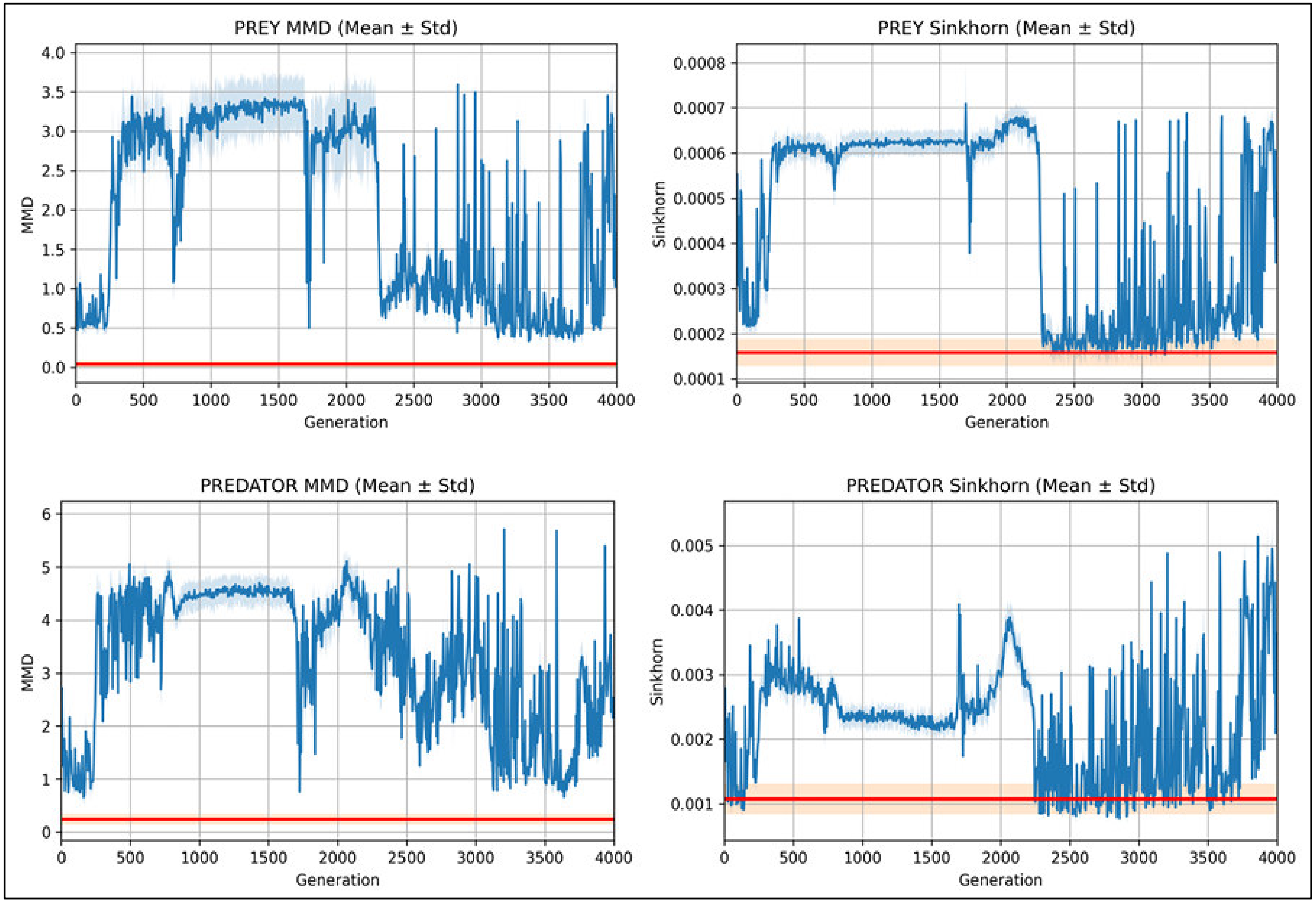
Video Predator-Prey Model



Video Predator-Prey Model

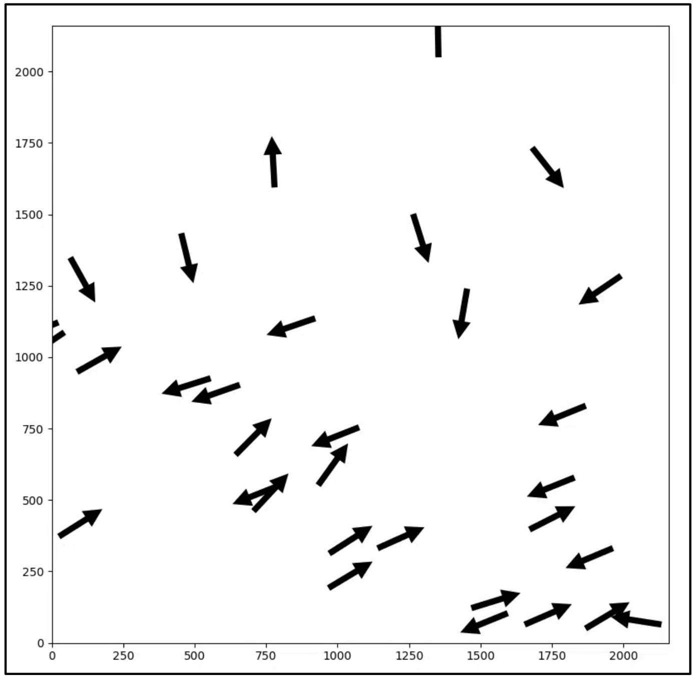


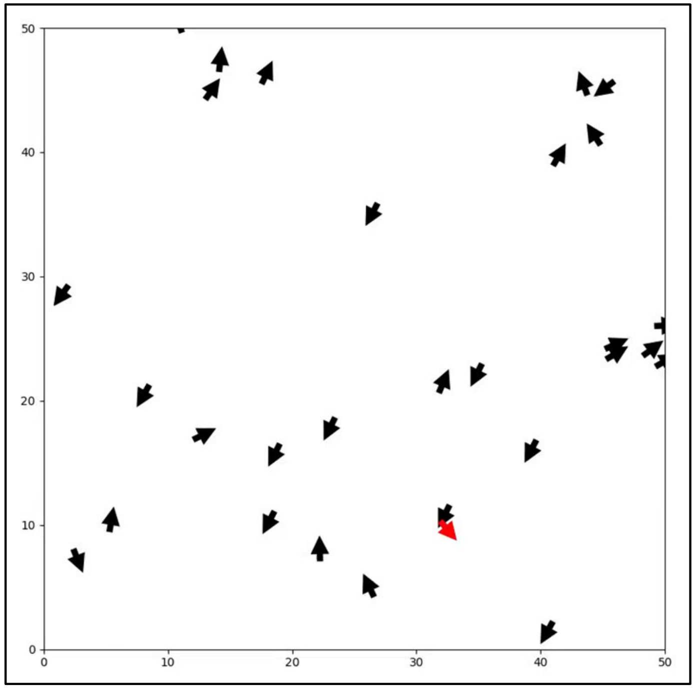
Video Predator-Prey Model



Video Predator-Prey Model

Expert demonstrations





Recovery of missing inter-group dynamics

Tensor Quality
Input validation & attack oversampling
Predator flag & Couzin-simulated data

Action Aggregation
Predator–prey mixing based on distance, attention, hard-coded pressure

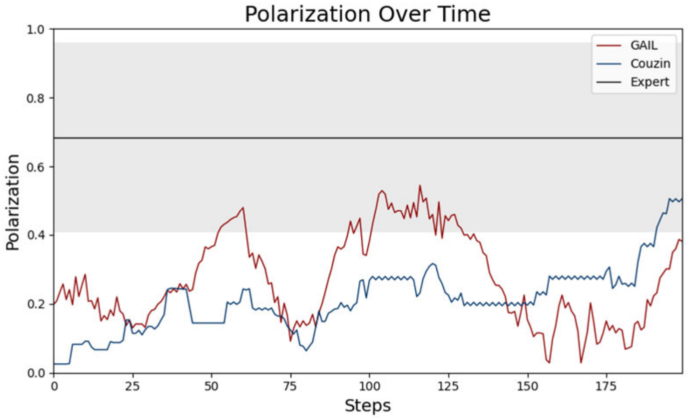
Transition Encoder
Global system representation instead of agent-wise discrimination

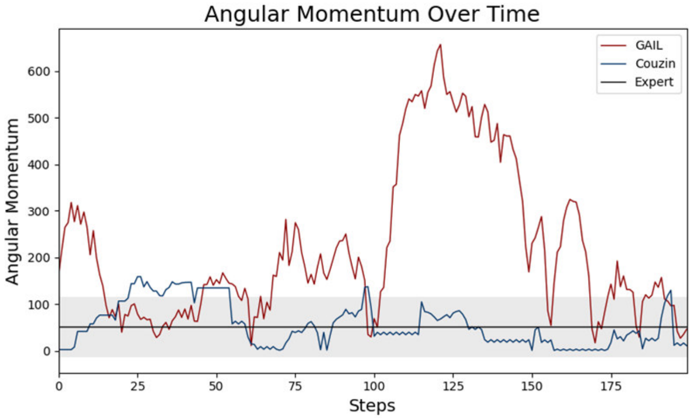
Feature Weighting
Behavioral clustering in latent transition space
Action weighting per cluster

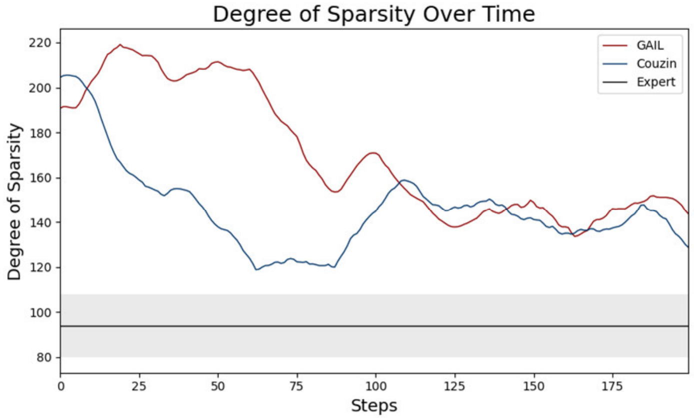
Reward Shaping
Prey: maximize predator distance
Predator: minimize distance to closest prey

Expertiment: Model Comparison

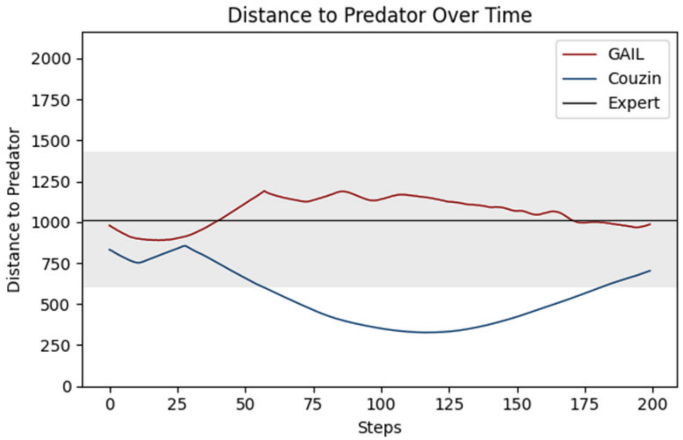
Swarm-related metrics:

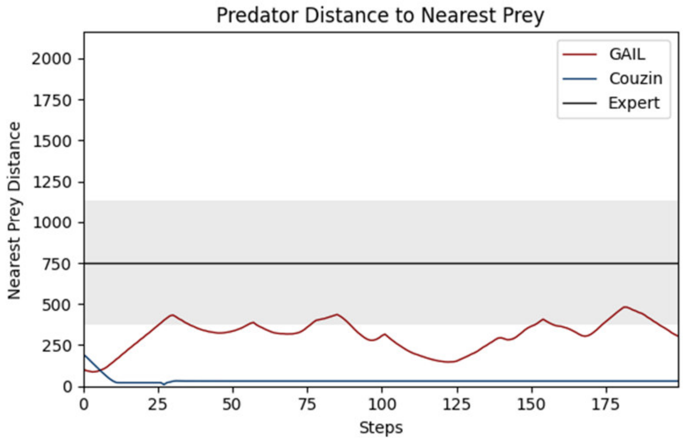


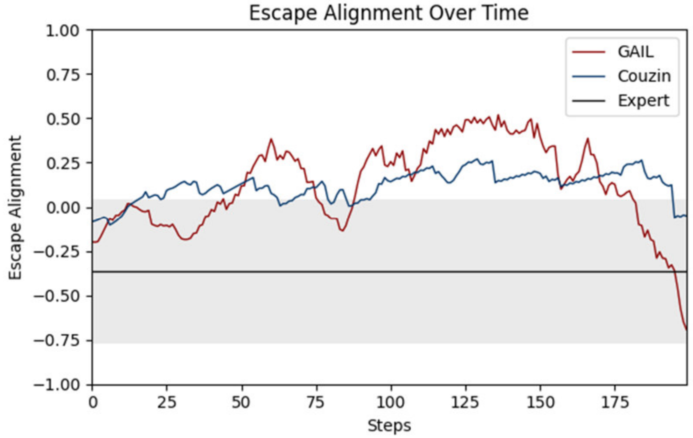




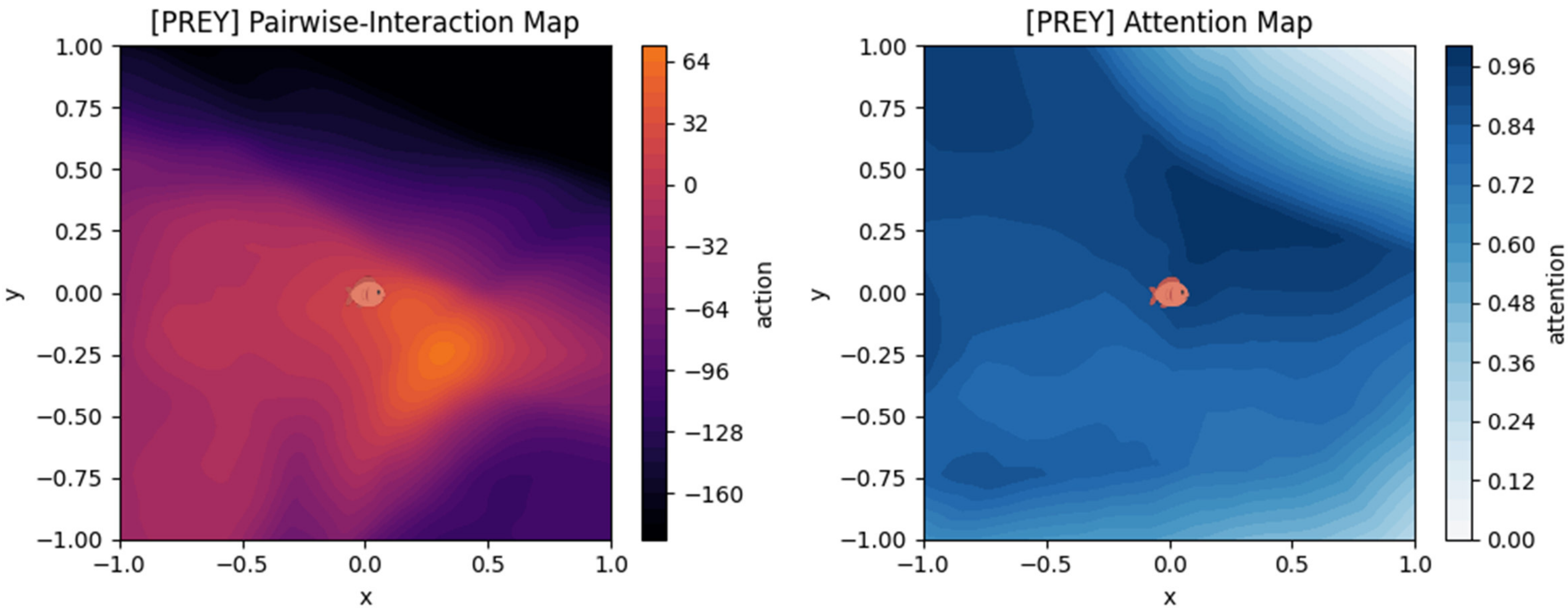
Survival-related metrics:



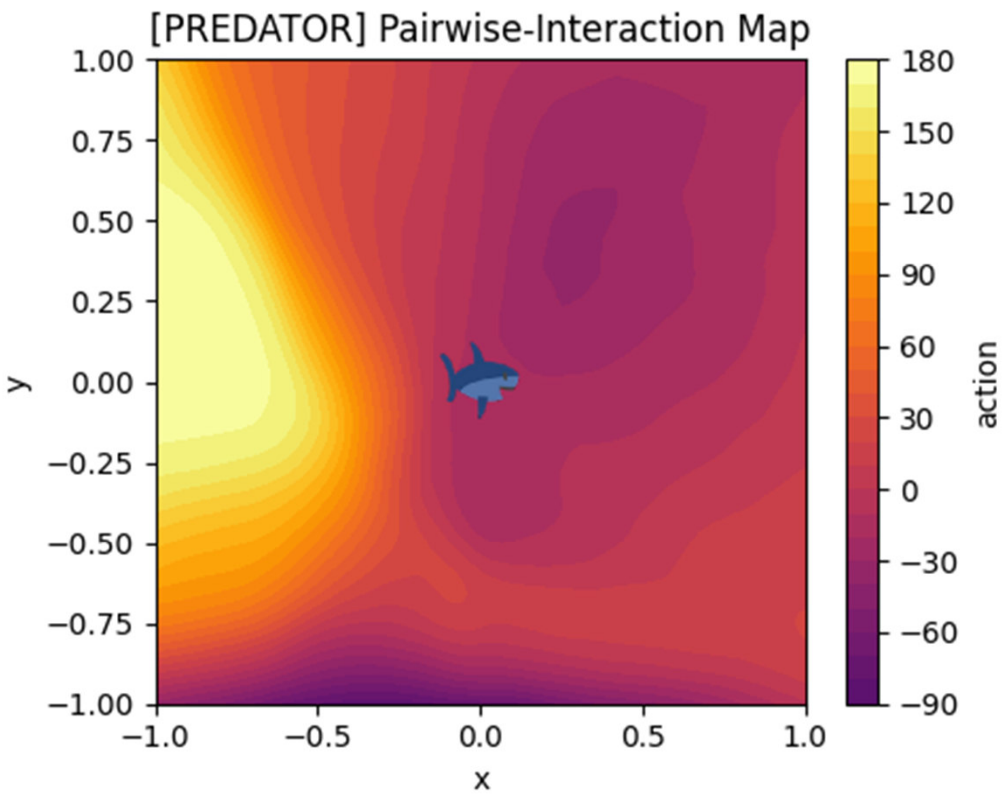


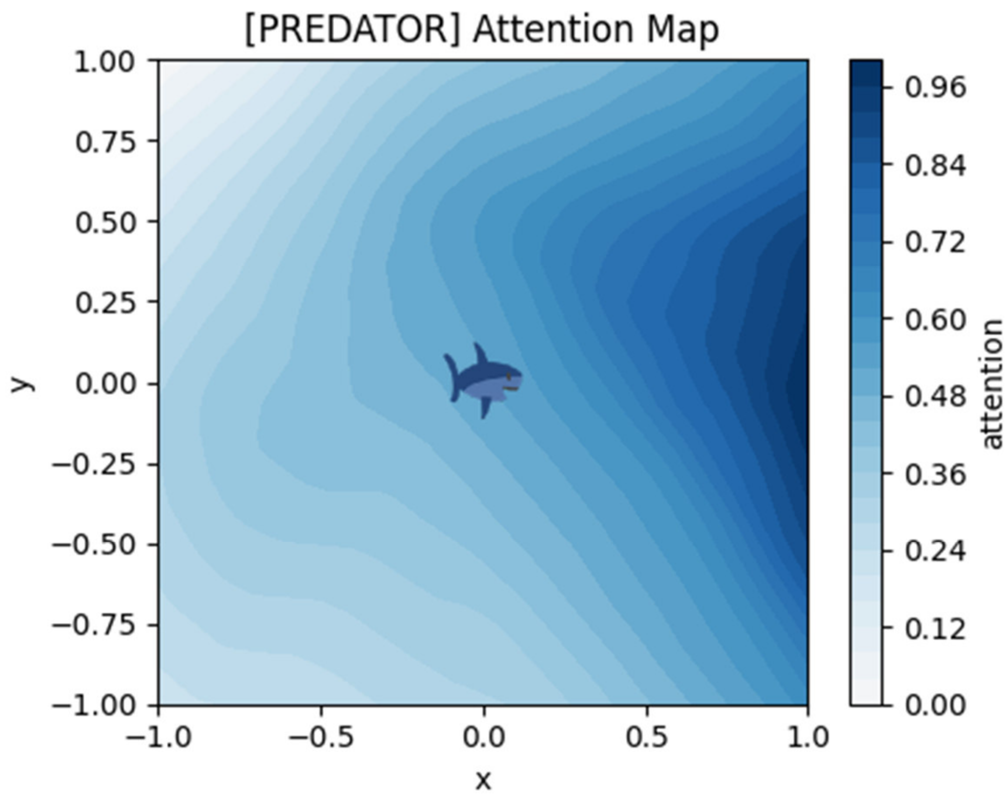


Expertiment: Policy Maps

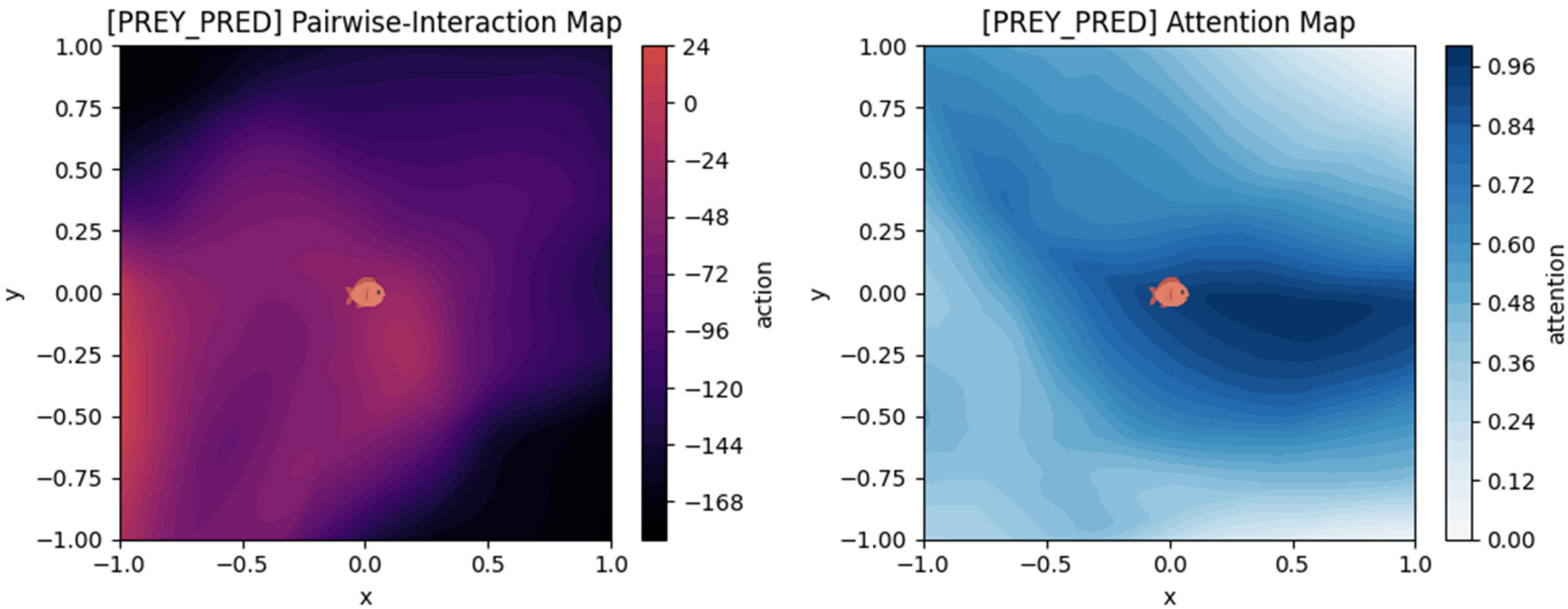


Expertiment: Policy Maps



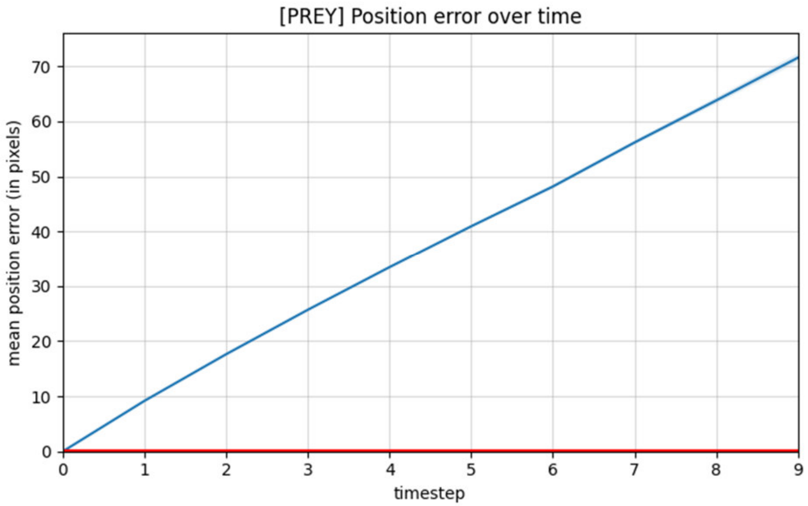


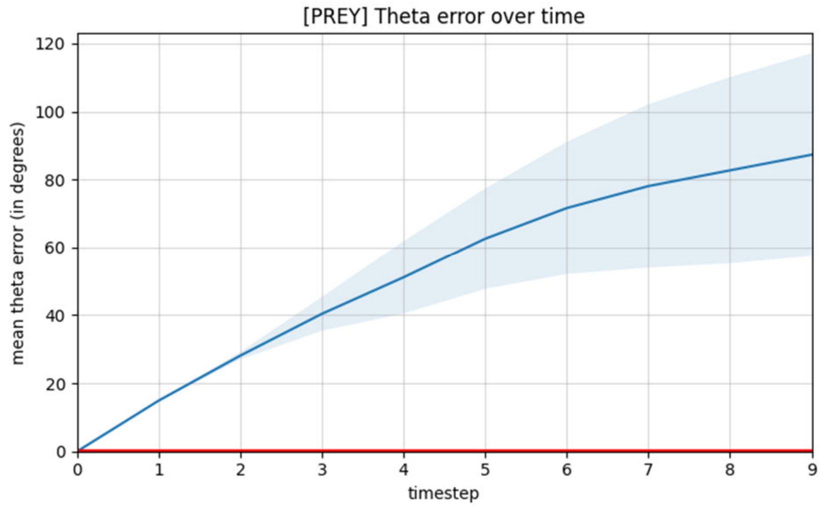
Expertiment: Policy Maps



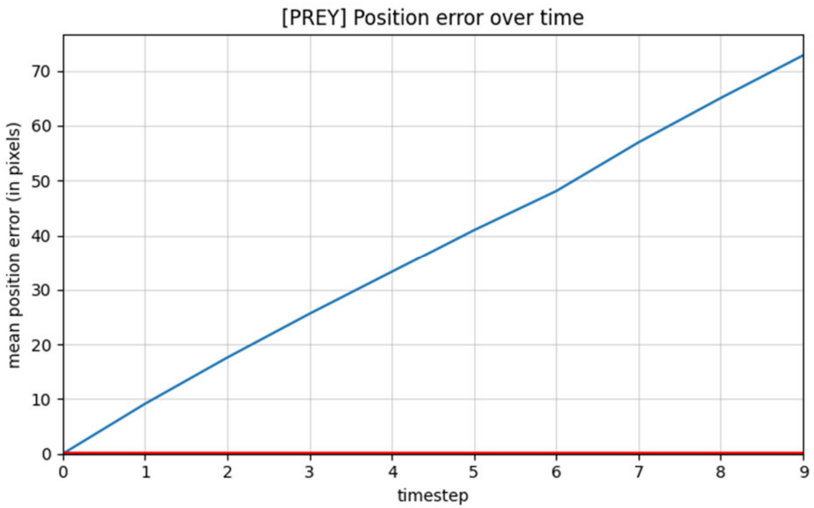
Expertiment: Trajectory Prediction

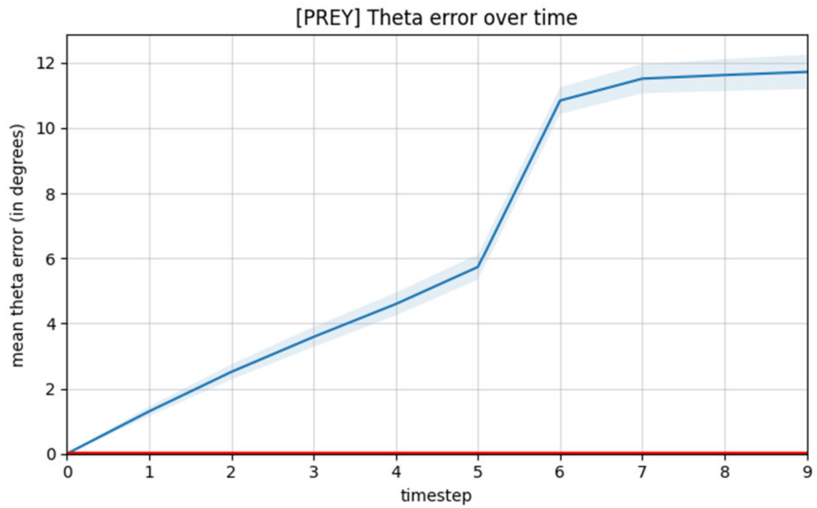
GAIL prey policies:





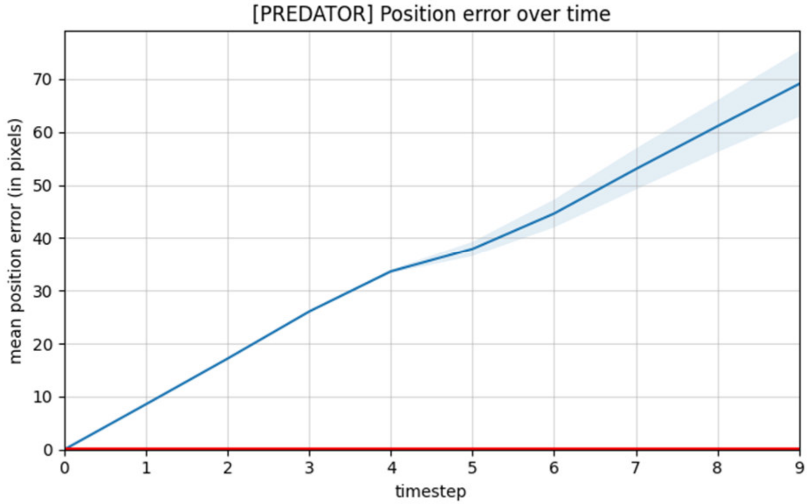
Random prey policies:

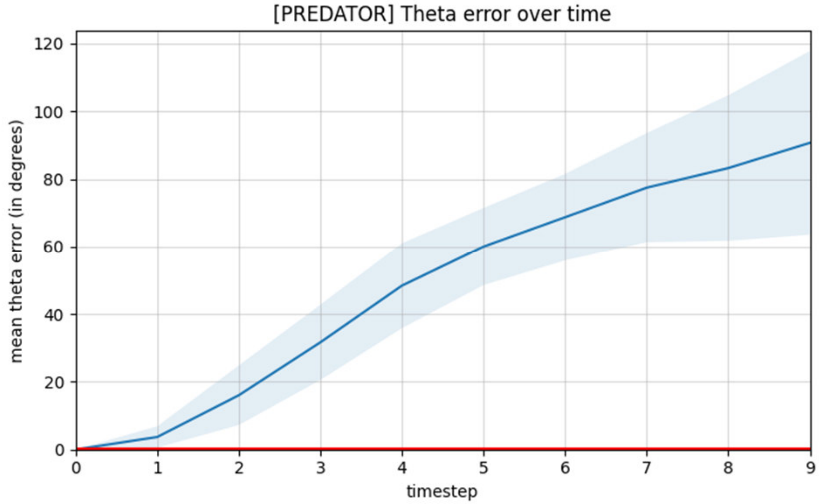




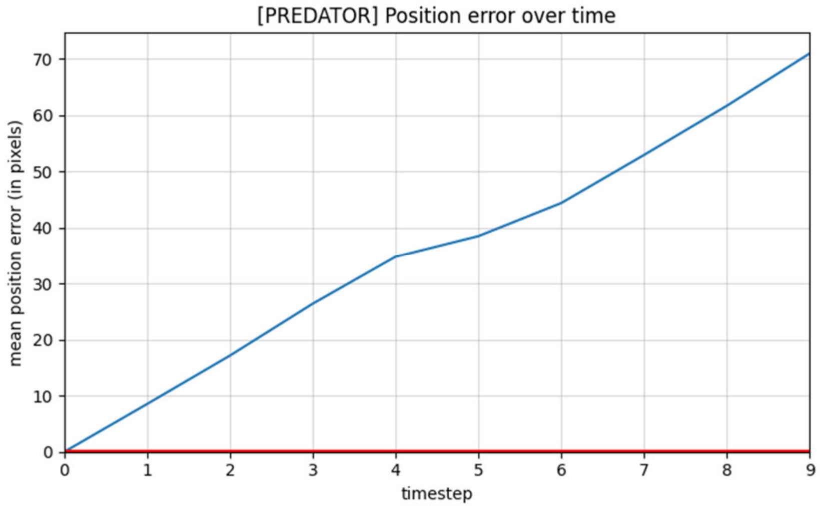
Expertiment: Trajectory Prediction

GAIL predator policies:



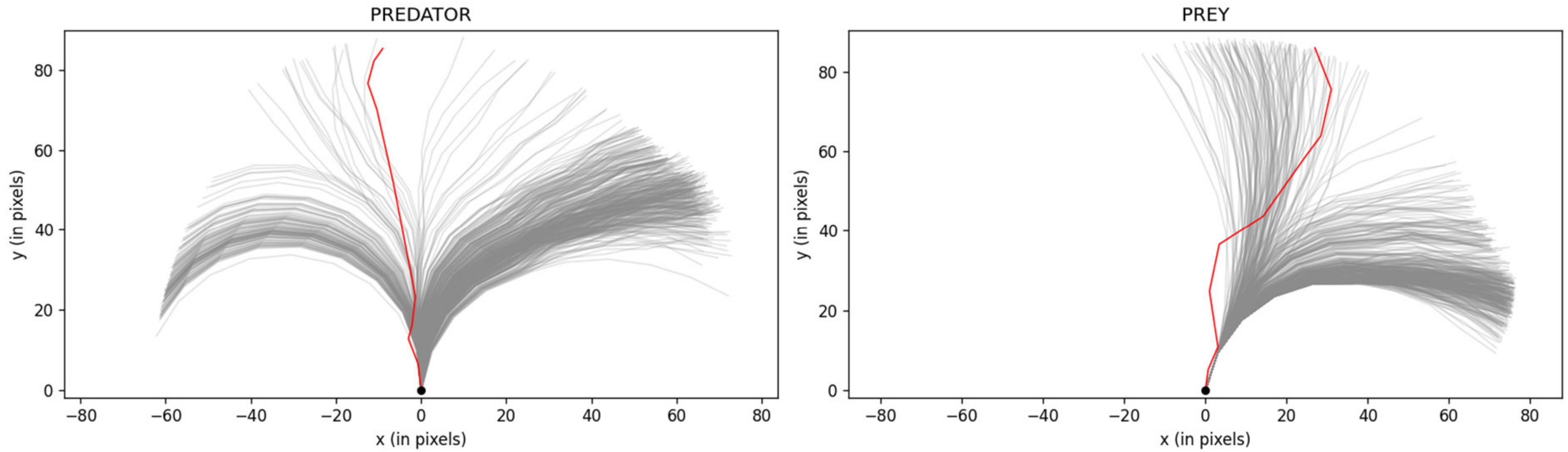


Random predator policies:

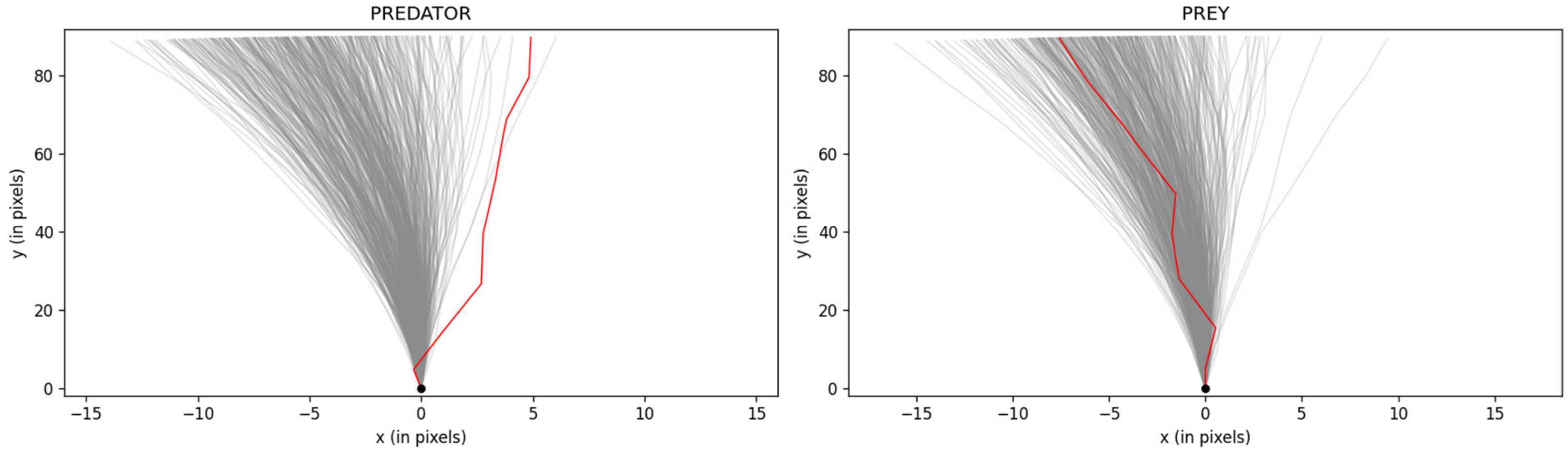


Expertiment: Trajectory Prediction

GAIL predator policies:

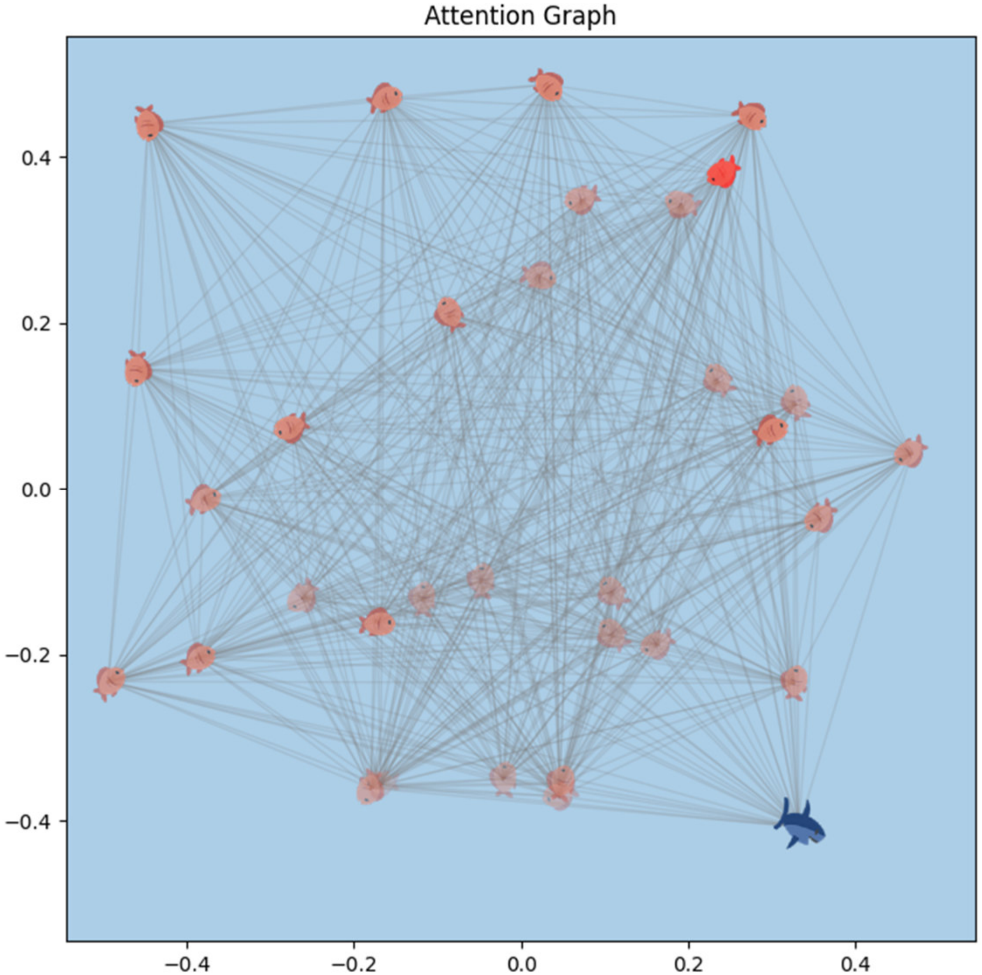


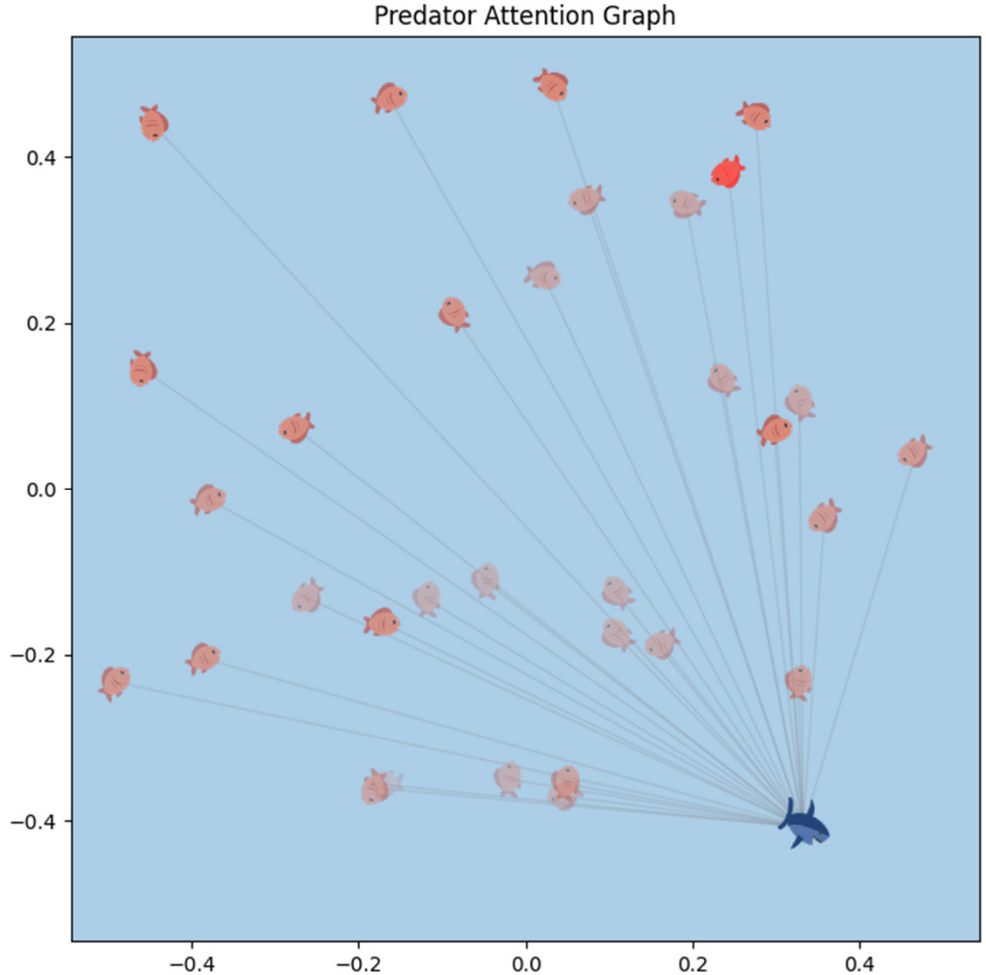
Random predator policies:



Experiment: Leadership Analysis

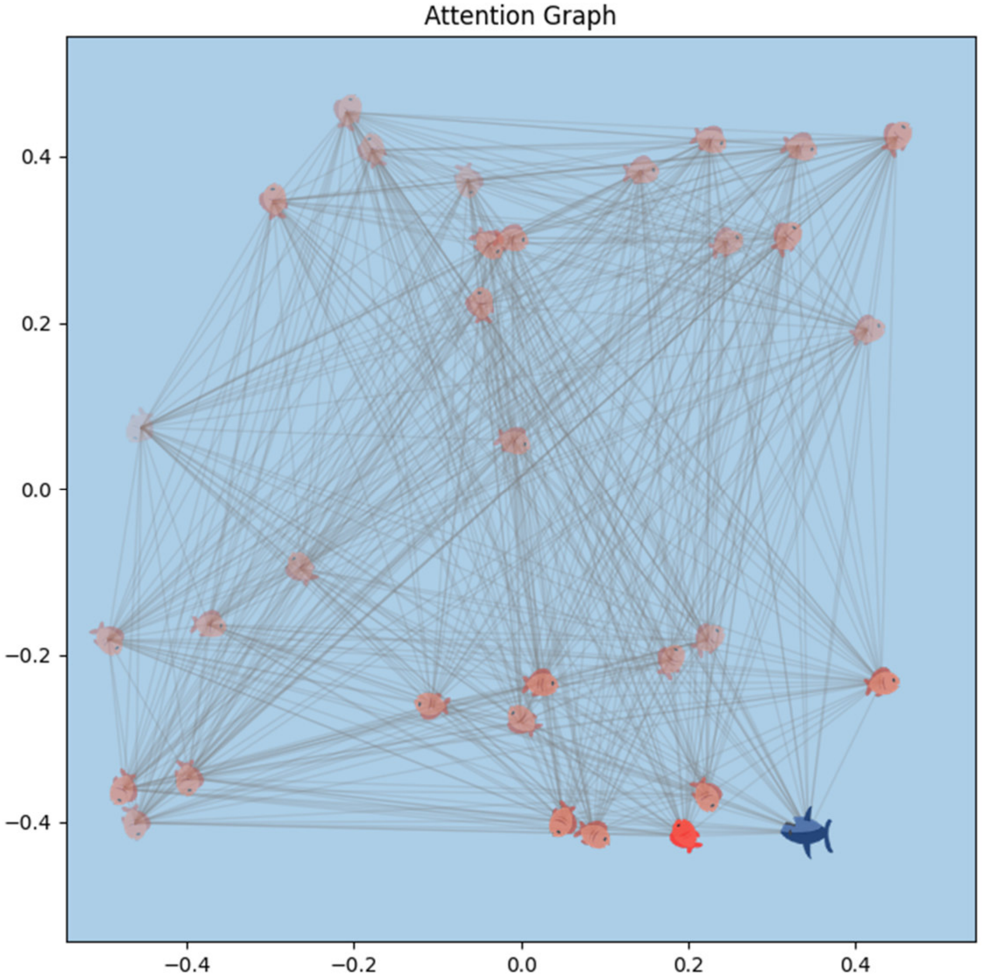
GAIL policies:

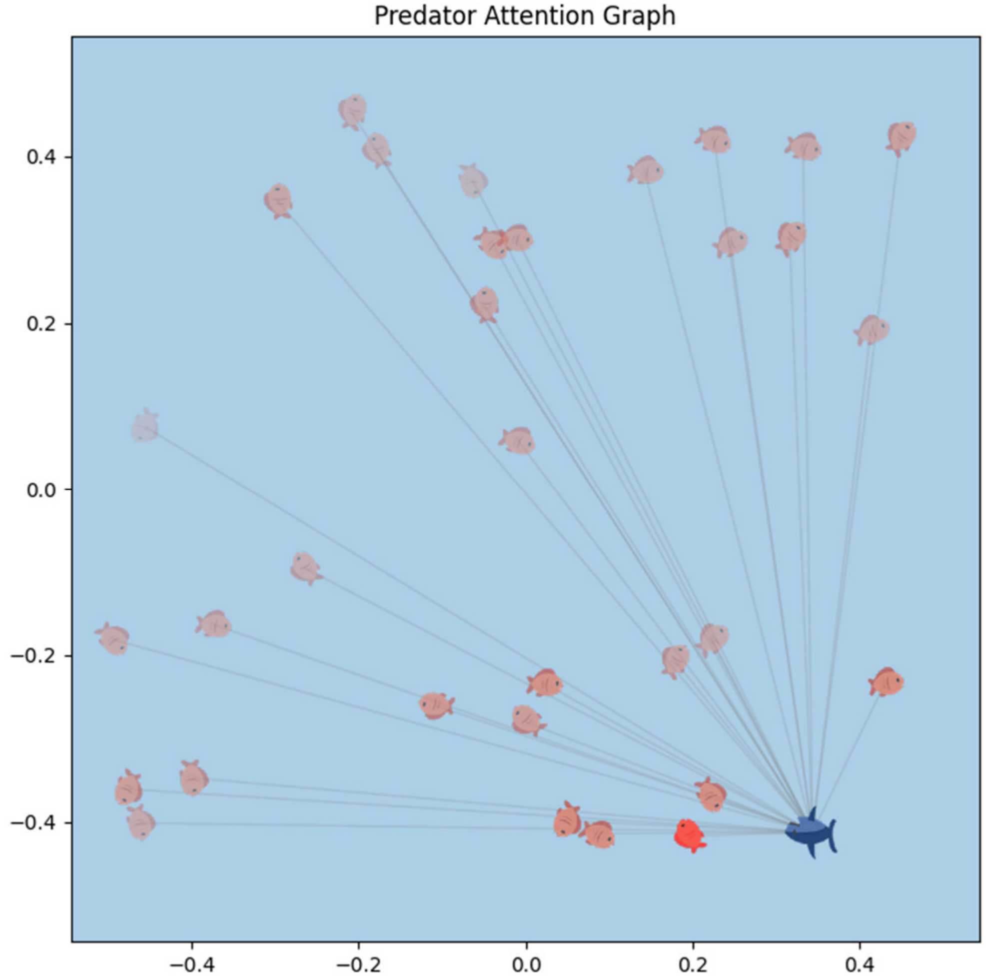




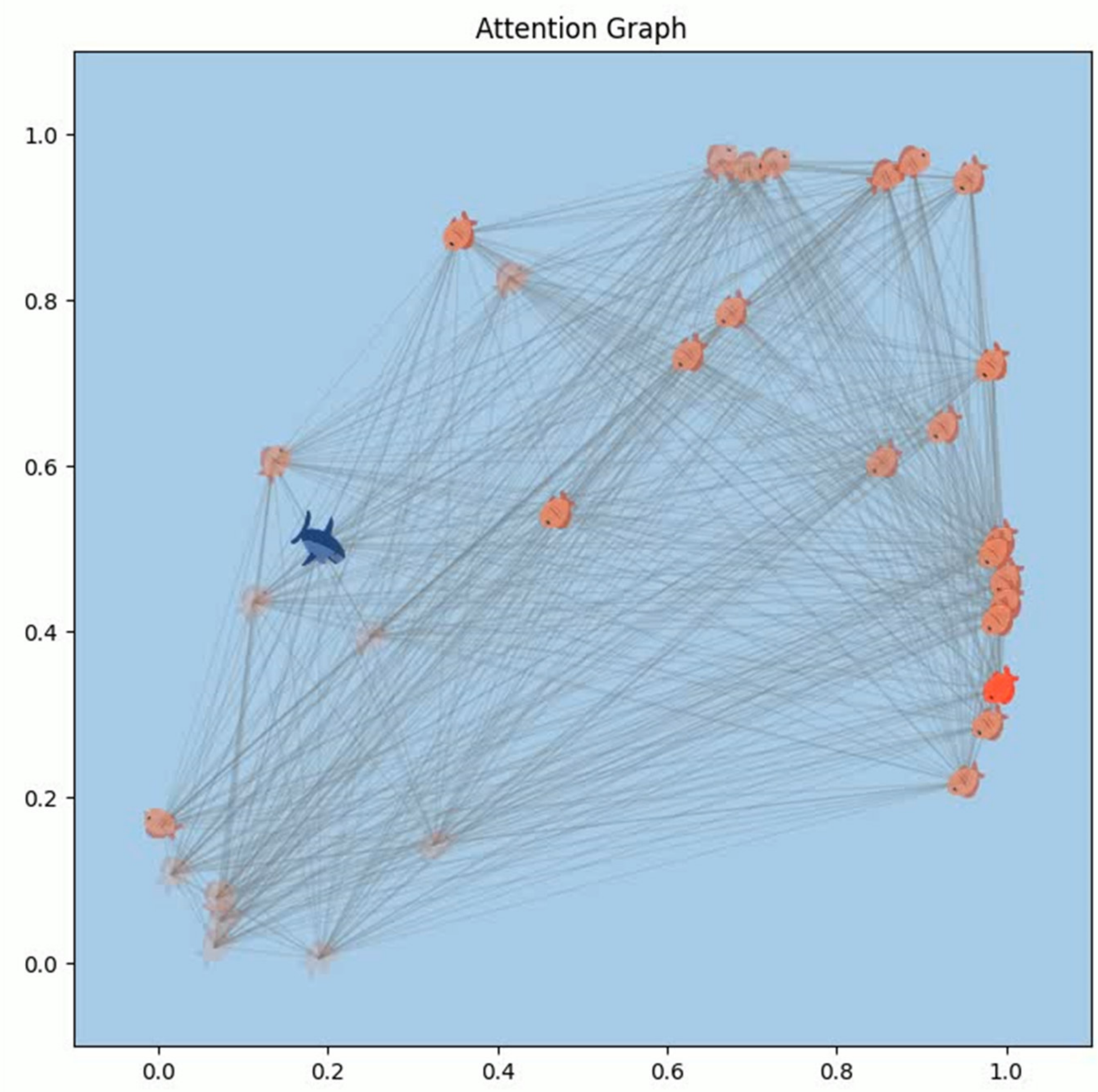
Expertiment: Leadership Analysis

BC policies:

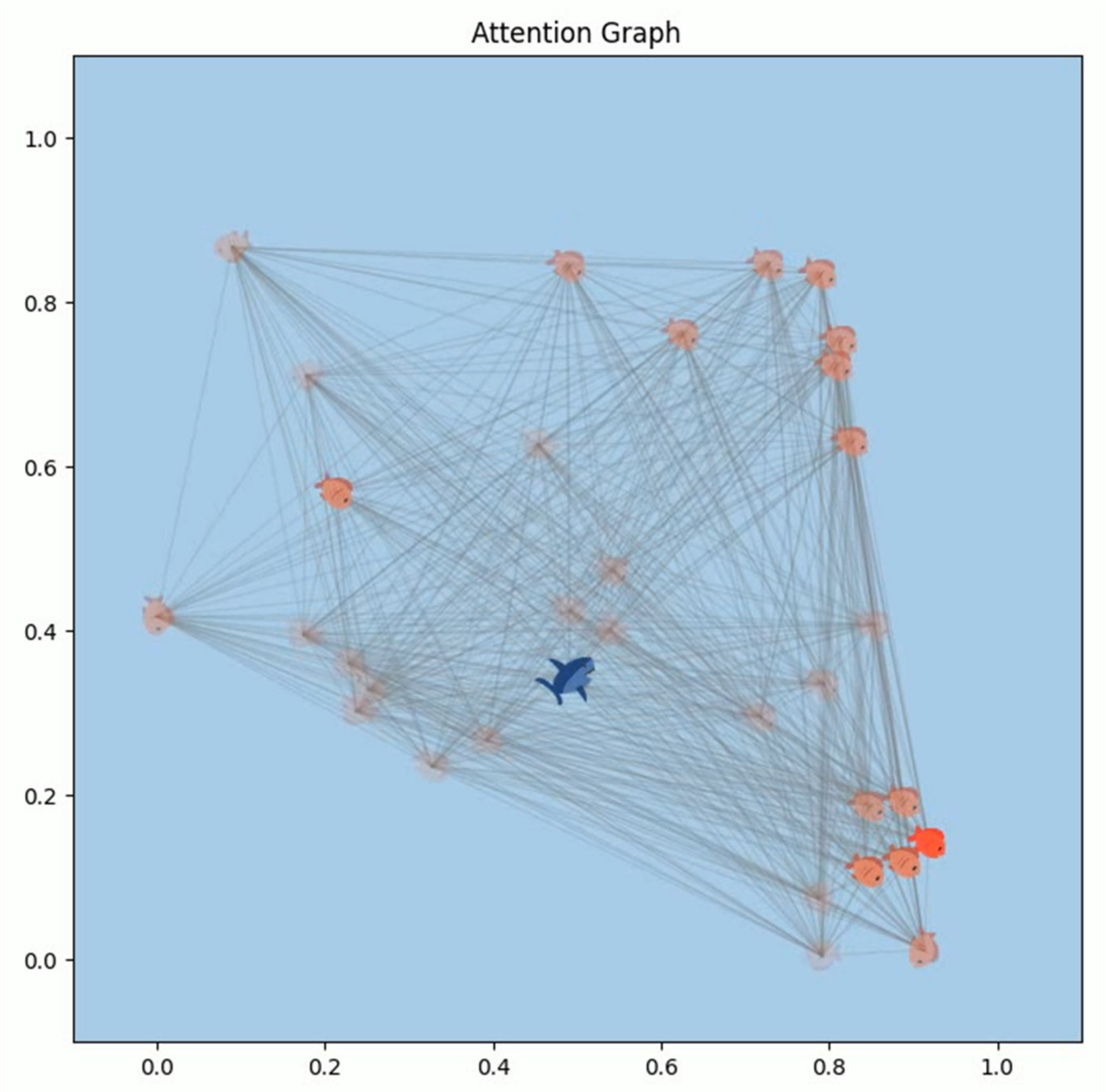




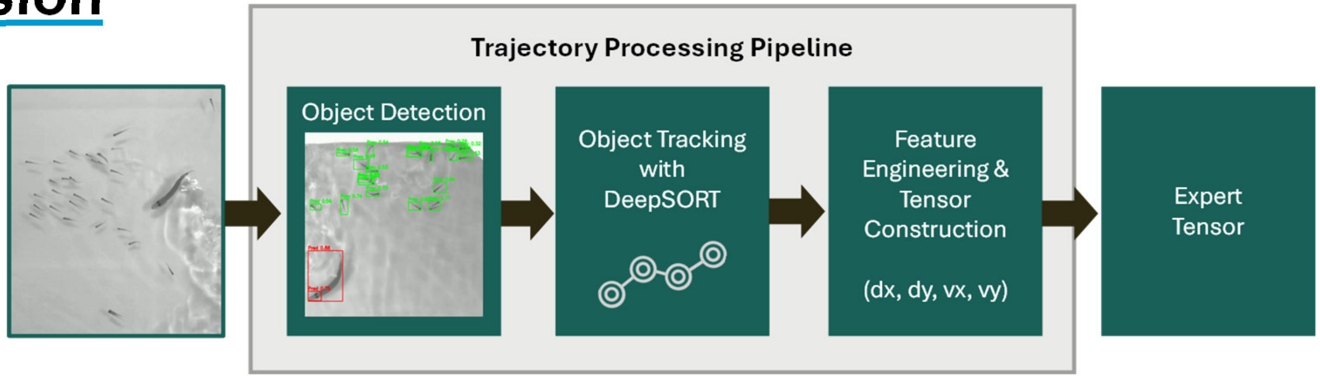
Expertiment: Trajectory Prediction



Expertiment: Trajectory Prediction



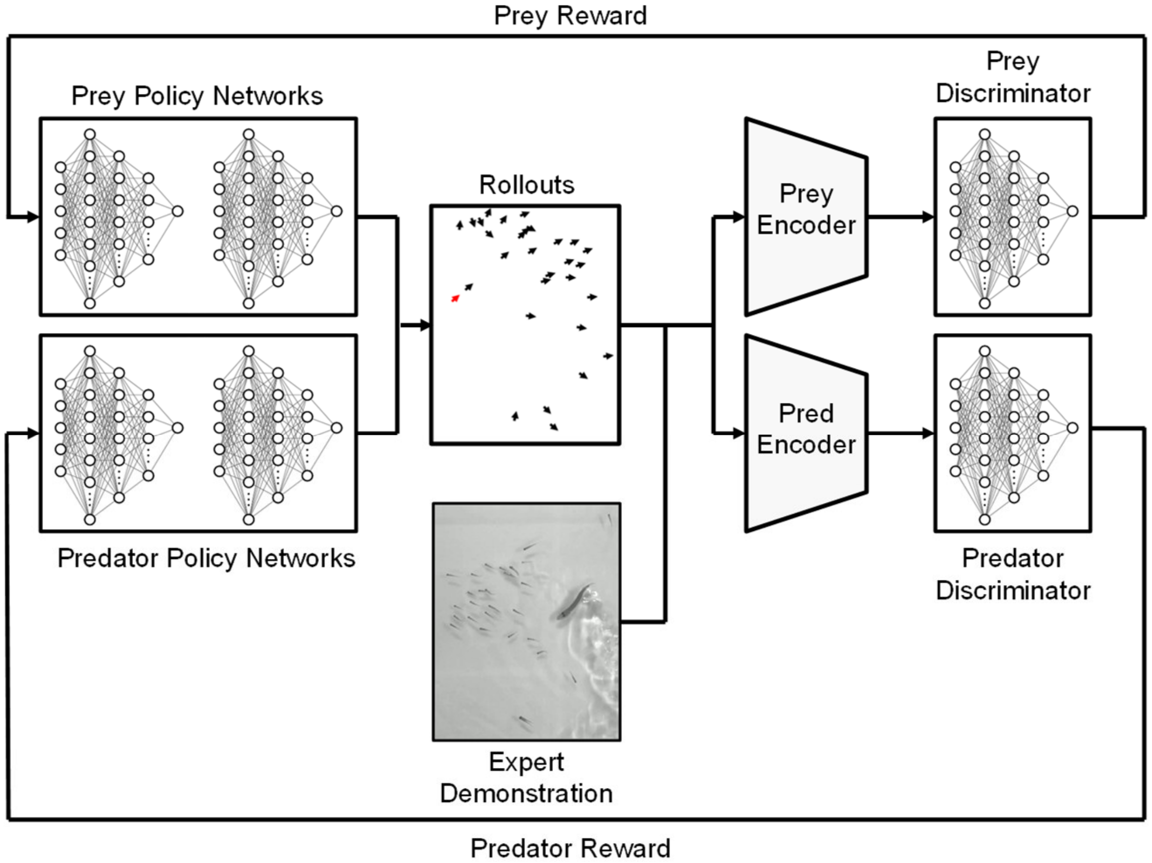
Summary & Discussion



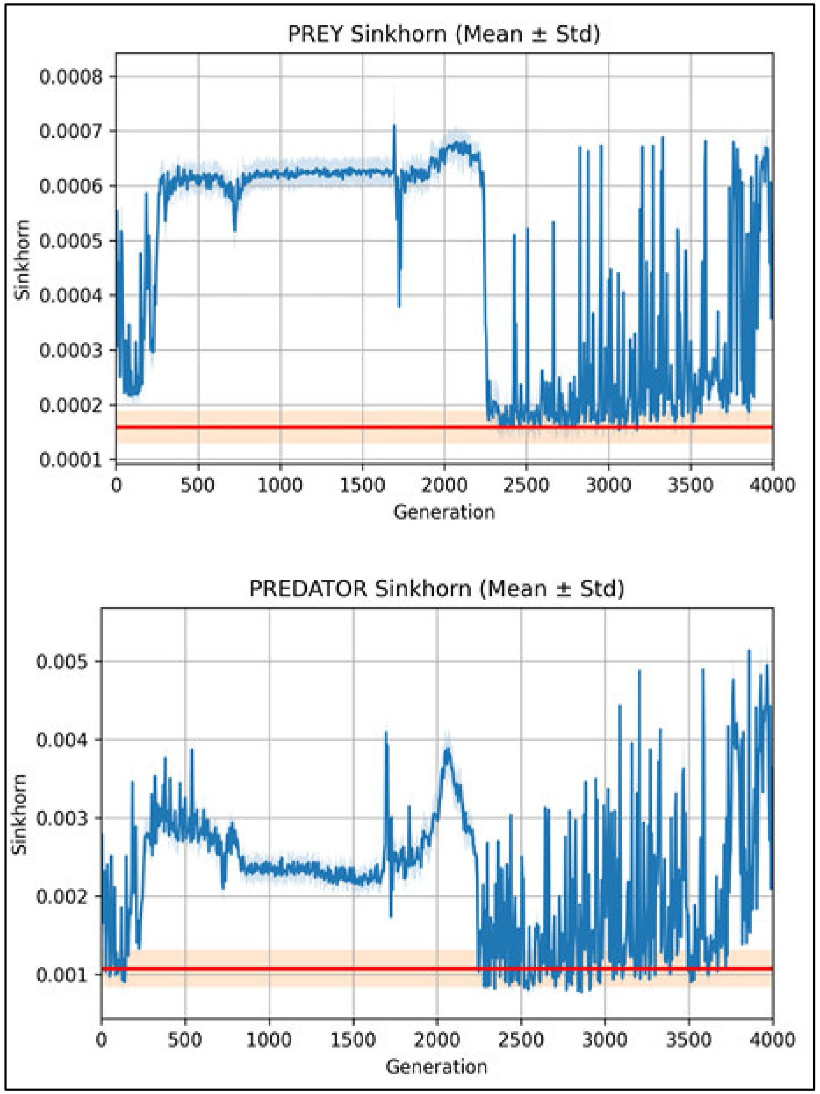
General GAIL pipeline works!
→ Prey-only model

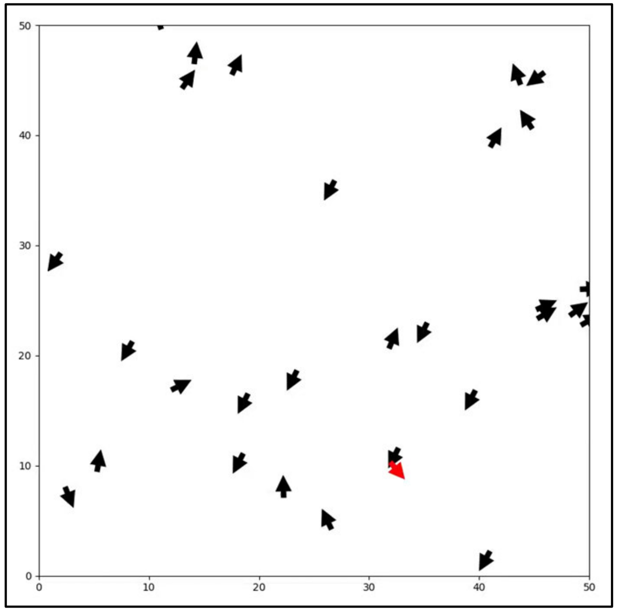
Partial imitation of predator-prey models

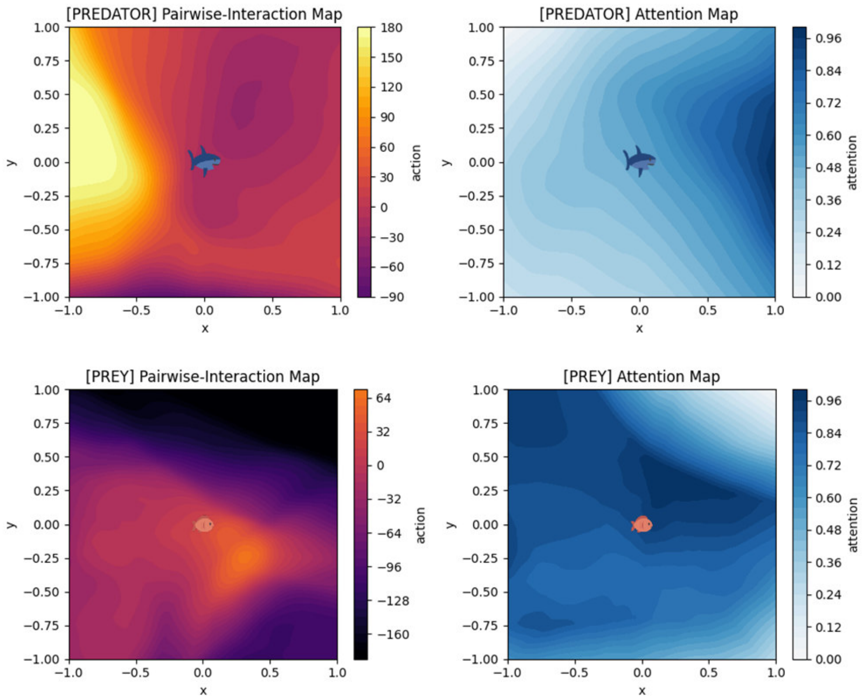
But: missing inter-group dynamics



Summary & Discussion



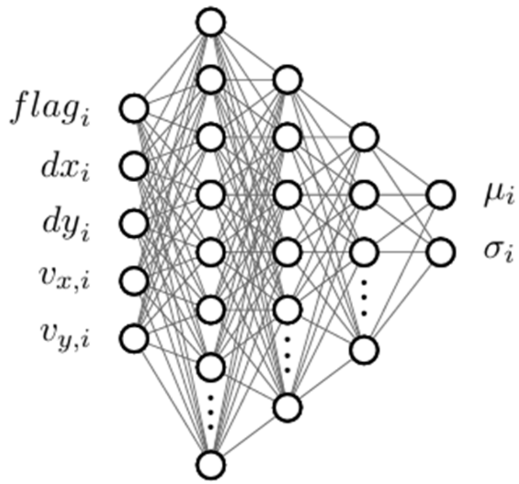




Limitations & Future Work

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

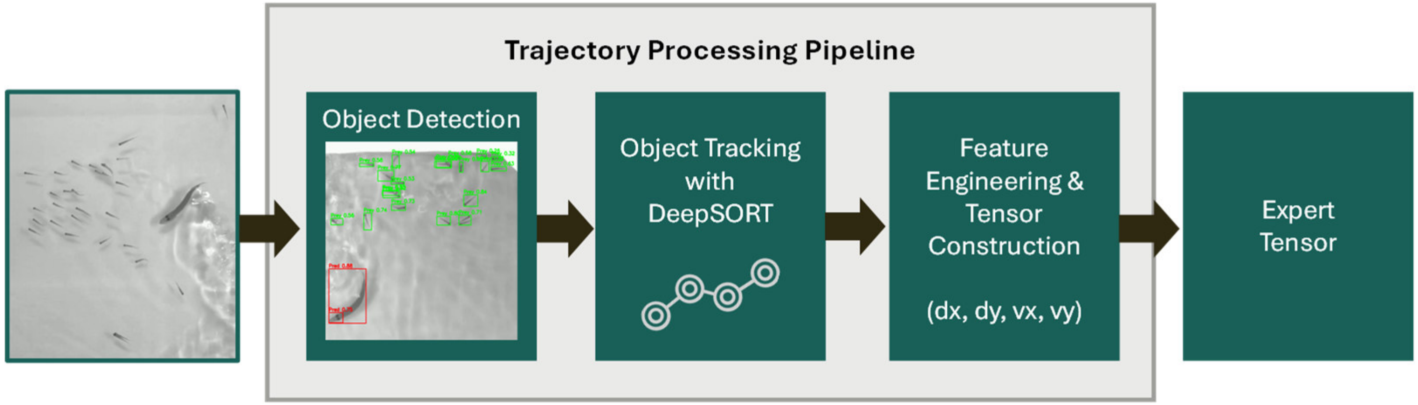


Predator attacks:

- Increase number of attack sequences
- Definition of attack behavior (hectic movements, …)

Input features:

- Extend policy features with acceleration



Limitations & Future Work

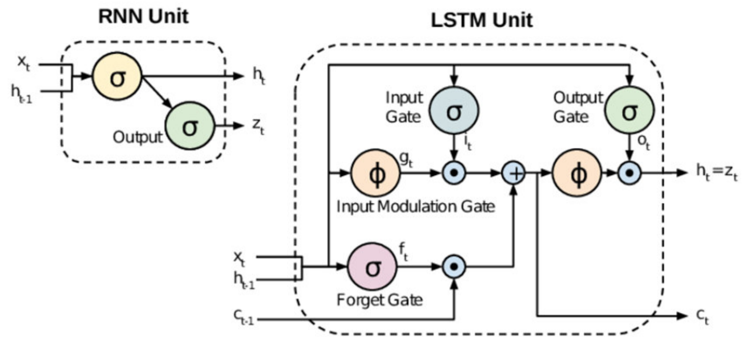
- Tensor Quality
- Action Aggregation
- Transition Encoder
- Feature Weighting
- Reward Shaping

<u>Sequence processing discriminator</u>

→ Modelling temporal dependencies

→ Train on longer temporal contexts

→ RRN or LSTM



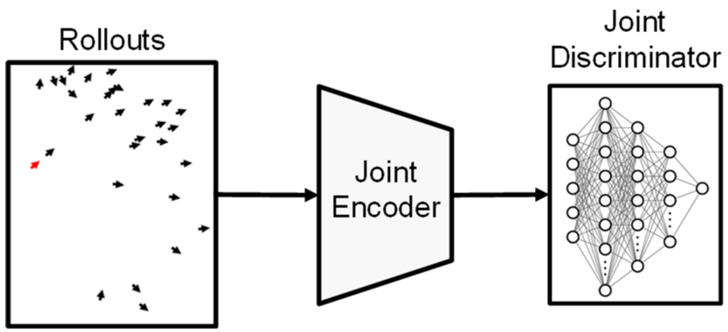
https://ashutoshtripathi.com/2021/07/02/what-is-the-main-difference-between-rnn-and-lstm-nlp-rnn-vs-lstm/

<u>Missing inter-group dynamics</u>

→ CBIL's multi-instance single policy

→ Joint encoder and discriminator

→ 2-stage-setup



Thank you for your Attention!

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

Jannik Wirtheim
Konstanz

Environment Installation

Instruction shows a step by step installation using Anaconda Prompt.

1. Create new conda environment.

```
conda create -n GAIL python=3.10 -y
conda activate GAIL
```

2. Install CUDA enabled pytorch version (Windows + CUDA 12.6).

```
pip install "torch==2.6.0+cu124" --index-url https://download.pytorch.org/whl/cu124
pip install "torchvision==0.21.0+cu124" --index-url https://download.pytorch.org/whl/cu124
```

3. Clone repository.

```
git clone https://github.com/wirthy21/Master-Thesis.git
```

4. Install requirements.

```
cd Master-Thesis
pip install -r requirements.txt
```

5. For data access use the following OneDrive-link: **Password = „predator"**

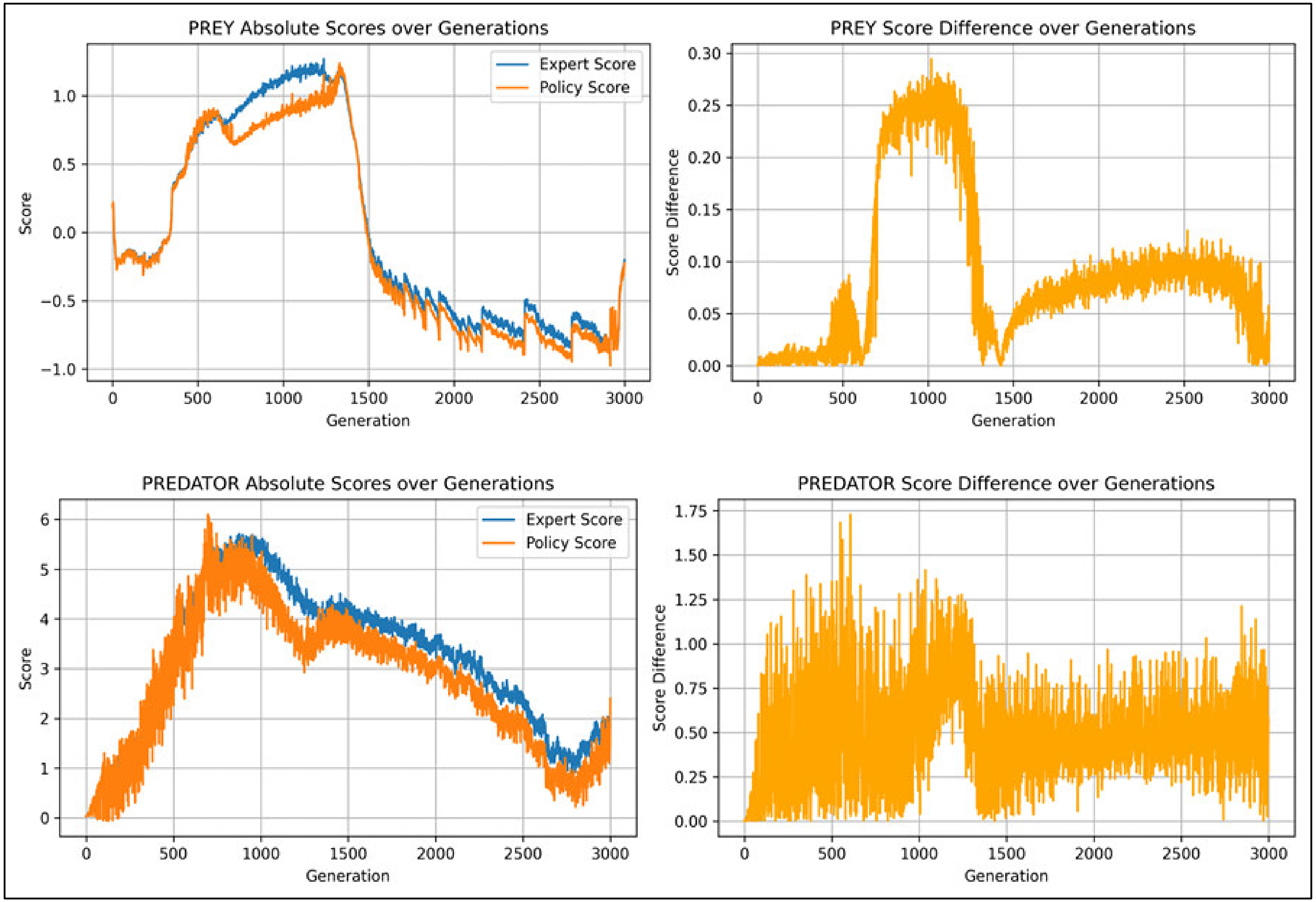
```
https://1drv.ms/f/c/e86ec4f701d097f4/IgCQUyQfwzltSLD6Ojpz2aFxAa7X_cVoBMCKVs6FVY3137c?e=XdhcqM
```

My brain during master’s thesis

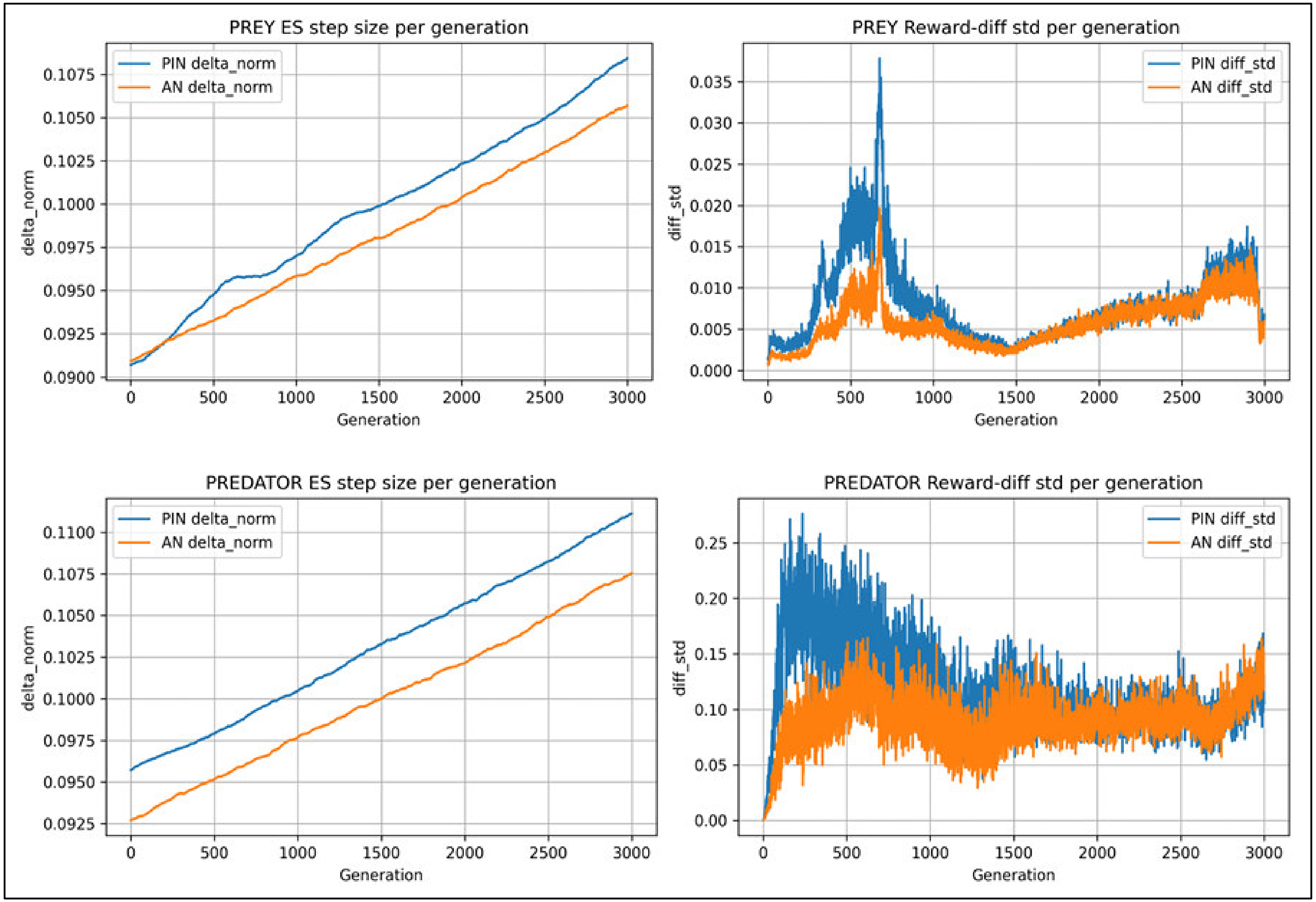
After defense

Fish GIF: https://www.reddit.com/r/Fish/comments/115h1v1/yall_mind_watching_my_spinning_fish_iii_be_back/
Beer GIF: https://tenor.com/de/view/beer-chug-drinking-drink-thirsty-gif-4362970331505802377

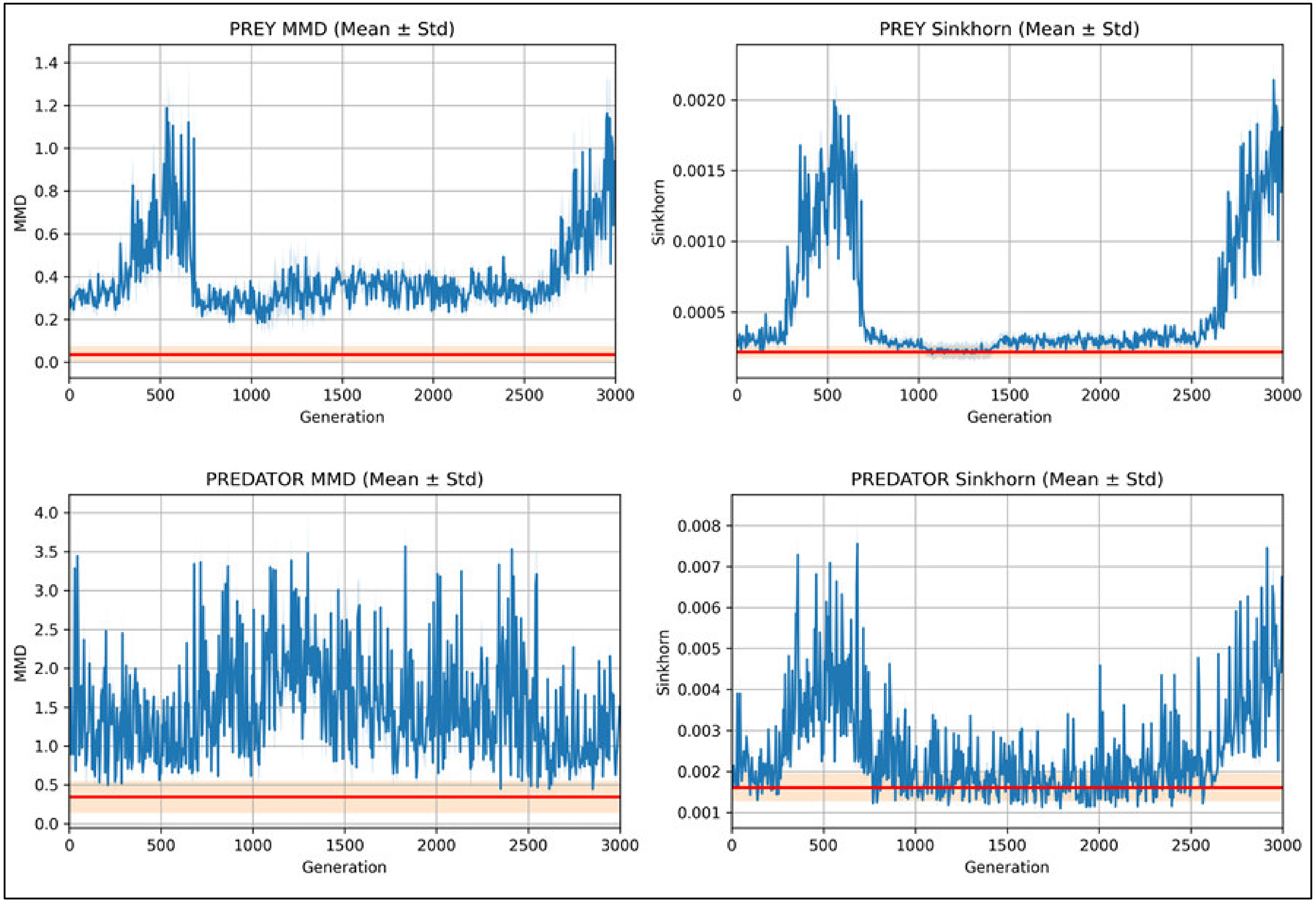
Couzin Predator-Prey Model



Couzin Predator-Prey Model

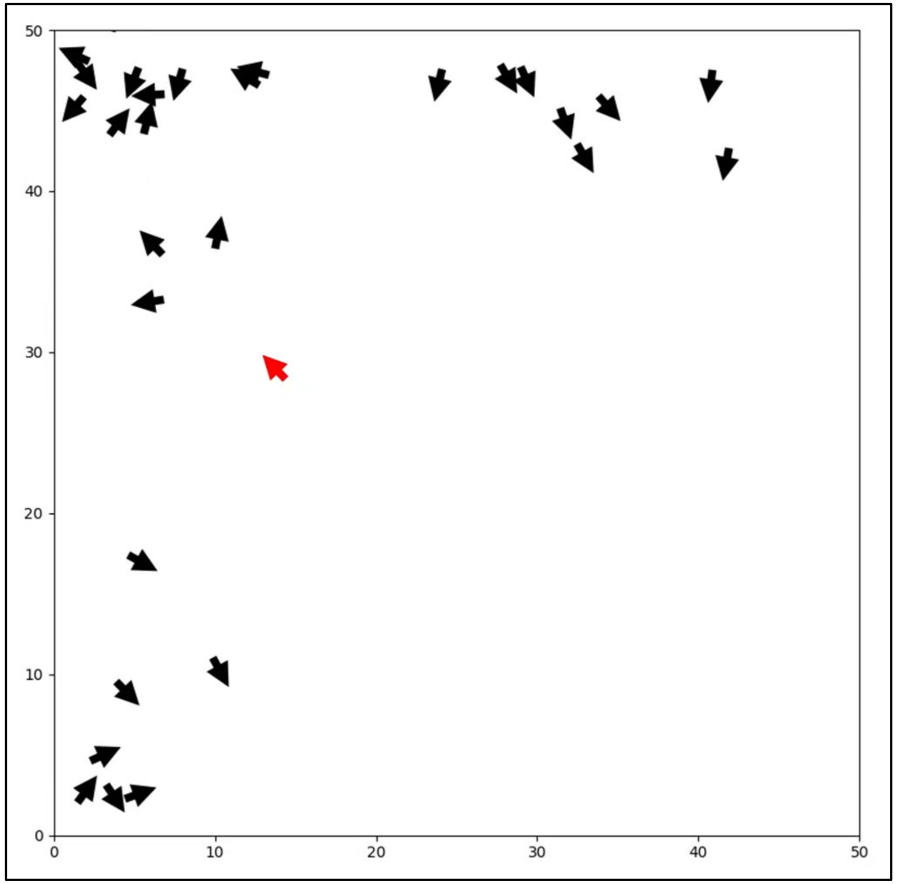


Couzin Predator-Prey Model



Couzin Predator-Prey Model

Expert demonstrations
Couzin model



Policy-generated
Predator-Prey GAIL model