

GNES: Neural-Guided Evolutionary Program Search for Interpretable UAV Swarm Control

Chen Wang^{1,2†} Minfang Lu^{3,4†} Cheng Zhu^{1,2} Xiangke Wang^{3,4} Fei Ming⁵
Yaochu Jin^{5*}

¹College of Systems Engineering & ²National Key Laboratory of Information Systems Engineering, NUDT, Changsha 410073, Hunan, China

³College of Intelligence Science and Technology & ⁴National Key Laboratory of Equipment State Sensing and Smart Support, NUDT, Changsha 410073, Hunan, China

⁵School of Engineering, Westlake University, Hangzhou 310030, Zhejiang, China [†]Equal contribution ^{*}Corresponding

author

Core Takeaway

Neural networks guide *where to search* in discrete program space.

Every UAV deploys a **closed-form symbolic velocity law** (≈ 17.3 nodes) with **zero neural inference overhead** at runtime.

1. Why Symbolic UAV Swarm Control?

- **Deep MARL:** Learns strong policies, but opaque neural weights are hard to audit, verify, or tune before flight.
- **Standard GP / GEP:** Preserves symbolic equations, but wastes simulator rollouts on blind random mutations.
- **Hand-Designed Rules:** Transparent (e.g., Boids/APF), but require manual derivation for each task.

Goal: Combine learned search efficiency with closed-form symbolic interpretability.

2. What GNES Introduces

- **Gene Growth Tree (GGT):** Grow–decode–repair representation over 36 typed primitives ($N_{\max} = 24$).
- **Program-Edit MDP:** States = GGT programs; Actions = 24 validity-preserving genetic operators.
- **GNN + PUCT MCTS:** Policy–value GNNs (Q_ϕ, V_ψ) guide look-ahead edit planning before rollout validation.

Search vs. Execution: GNNs assist offline evolution; deployed controllers are pure equations.

3. Key Empirical Findings

- **3 Swarm Benchmarks:** Flocking, Entrapping, Line Formation + 3D AirSim flight validation.
- **3-Seed Line Formation:** GNES reaches 1.751 ± 0.022 in **841** rollouts, beating DANCE (1.571, 2160) and MAPPO (1.467, 2400) on all seeds.
- **Handover Gain:** At eval 240, median fitness jumps $0.72 \rightarrow 1.05$; collapsed proposals drop $36.7\% \rightarrow 0\%$.

Surrogate Quality: Held-out V_ψ correlates with true rollout fitness ($r = 0.407, \rho = 0.541$).

1. Motivation: Decoupling Neural Search from Symbolic Execution

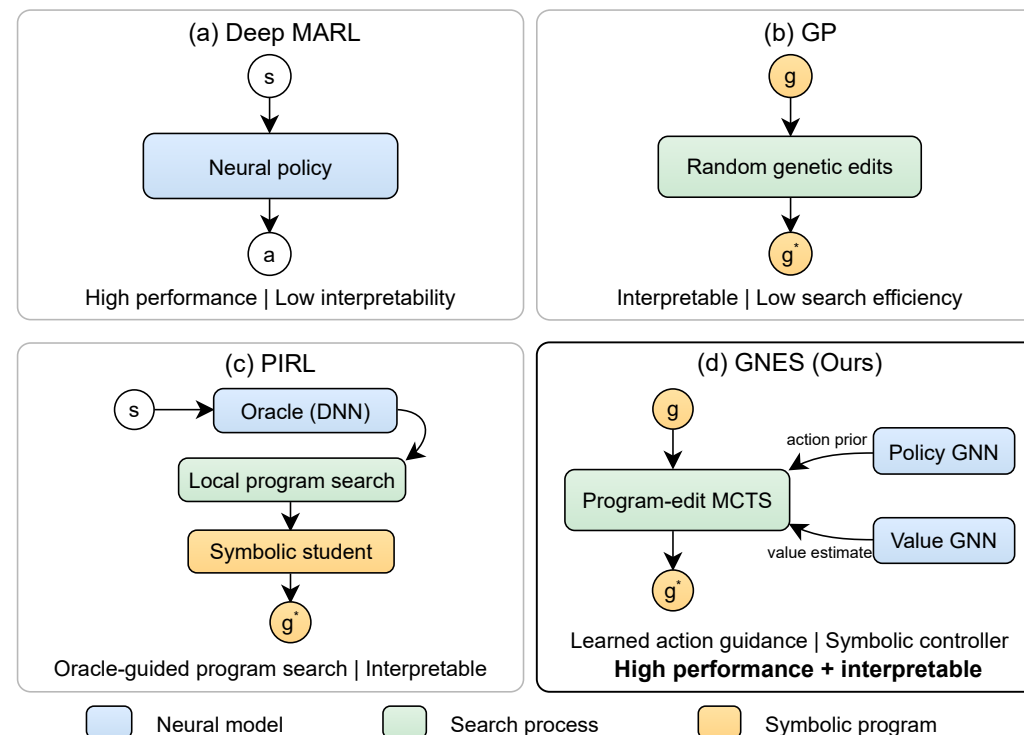
NeurIPS 2026 · GNES · Slide 2 / 8

Existing Paradigms vs. GNES

- **(a) Deep MARL (Black-Box Policy):** Maps local observations directly to actions via neural networks. High expressivity, but no explicit control law for safety inspection or field gain tuning.
- **(b) Genetic Programming (Blind Symbolic Search):** Mutates expression trees randomly. Outputs readable equations, but suffers from severe sample inefficiency and invalid/bloated trees under sparse long-horizon rewards.
- **(c) Oracle Distillation (e.g., PIRL):** Trains a neural teacher first and imitates it symbolically, inheriting teacher approximation bottlenecks.
- **(d) GNES (Ours — Neural-Guided Program Edits):** Learns a **policy–value GNN over program edits** $Q_\phi(g, a), V_\psi(g)$ and plans multi-step rewrites with MCTS before querying the simulator.

Design Automation Perspective: Instead of hand-deriving decentralized interaction rules, GNES discovers closed-form velocity laws automatically while keeping the deployed controller 100% symbolic, auditable, and editable.

Paradigm Comparison (Figure 1)



Key Distinction: Neural networks never sit in the airborne control loop; they only prioritize which symbolic edit to test next.

Dual Role: Every GGT g is simultaneously (1) an executable distributed control law shared by all UAVs and (2) a graph state in the program-edit MDP.

3. GNES Closed-Loop Program-Edit Architecture

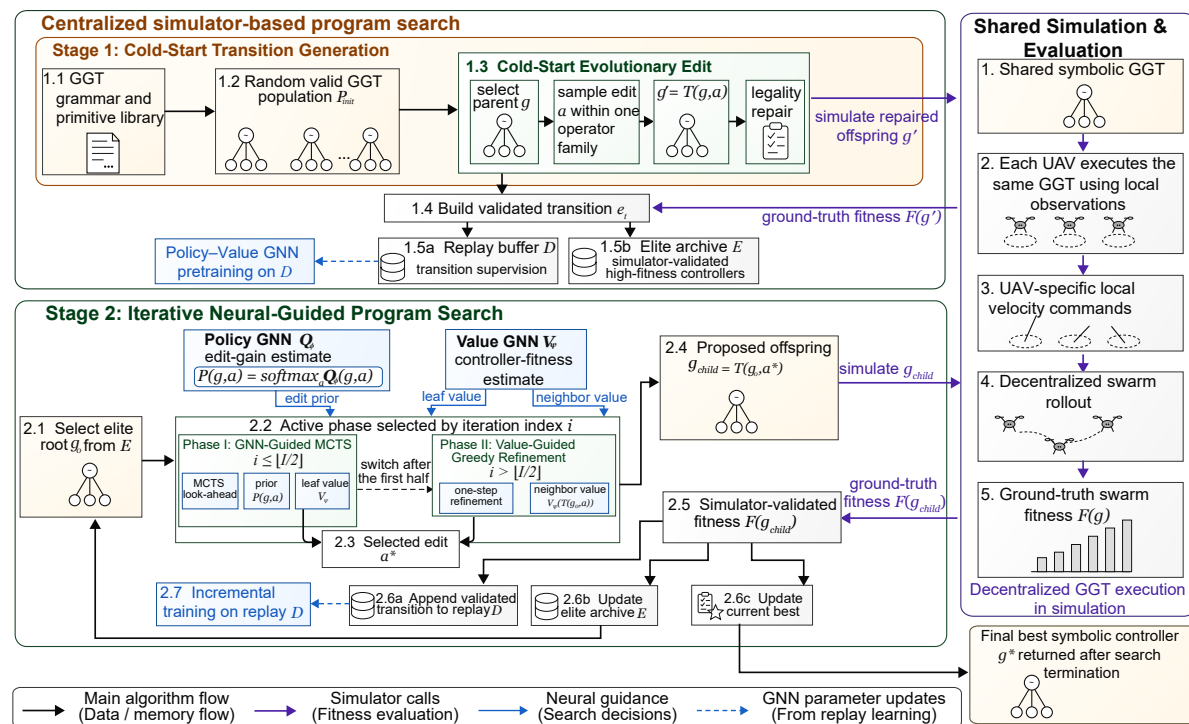
NeurIPS 2026 · GNES · Slide 4 / 8

Two-Stage Neural-Guided Evolutionary Loop

- **Program-Edit MDP:** State $g \in \mathcal{G}$ (valid GGT); Action $a \in \mathcal{A}(g)$ (24 typed operators across substitution, growth, pruning, transformation, crossover); Return $F(g) \in [0, 2]$ from simulator rollout.
- **Stage 1 — Cold-Start Exploration (40 Gens):** Stochastic GGT evolution populates replay buffer $\mathcal{D}$ and elite archive $\mathcal{E}$, then pretrains Policy GNN $Q_\phi(g, a)$ and Value GNN $V_\psi(g)$.
- **Stage 2 — Guided Search (20 Outer Cycles):**
 - **Phase I (PUCT MCTS):** Plans multi-step edits (35 $\rightarrow$ 90 sims) with prior $\text{softmax}_a Q_\phi(g, a)$ and leaf value V_ψ .
 - **Phase II (Value Refinement):** 1-step neighborhood screening via $V_\psi(g')$ for rapid local exploitation.
- **Closed-Loop Grounding:** Verified rollouts continually update $\mathcal{D}$, $\mathcal{E}$, Q_ϕ , and V_ψ .

Surrogate Validation: On $n = 200$ held-out programs, V_ψ achieves Pearson $r = 0.407$ and Spearman $\rho = 0.541$ with true fitness.

GNES System Overview (Figure 2)



Amortized Search Cost: GNN + MCTS filters weak mutations before expensive 300-step rollouts ($\approx +0.8$ s per candidate during offline search; 0 s at deployment).

4. Three Decentralized Swarm Benchmarks & 3D AirSim Validation

NeurIPS 2026 · GNES · Slide 5 / 8

Two-Tier Evaluation Protocol

Tier 1: Quantitative Fitness Engine (Sec. 3.1)

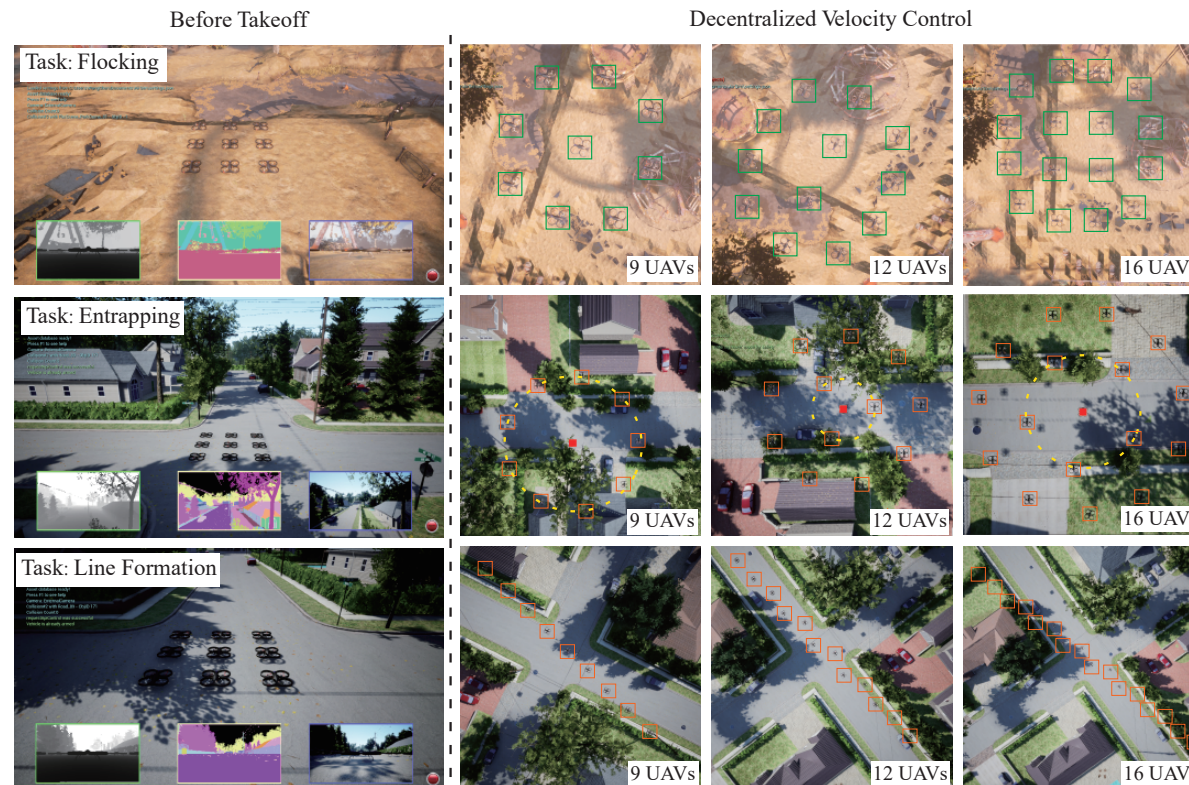
- Evolutionary fitness queries (300 steps/rollout, scale $[0, 2]$) use the planar simulator with drag and saturation.
- **Flocking (9 UAVs)**: Cohesive goal tracking with collision avoidance and alignment.
- **Entrapping (12 UAVs)**: Surround a moving target on a perimeter of radius d_0 .
- **Line Formation (16 UAVs)**: Form a collinear structure matching a moving target's velocity.

Tier 2: Qualitative 3D AirSim Deployment

- Champion GGT controllers discovered in Tier 1 are executed closed-loop on 9/12/16 3D multirotors in **Microsoft AirSim**.
- Validates transfer to 3D six-DOF flight loops without retraining.

Clarity Note: AirSim computes no evolutionary fitness; it provides 3D closed-loop verification.

Closed-Loop Trajectories & 3D AirSim Multirotor Swarms (Figure 4)

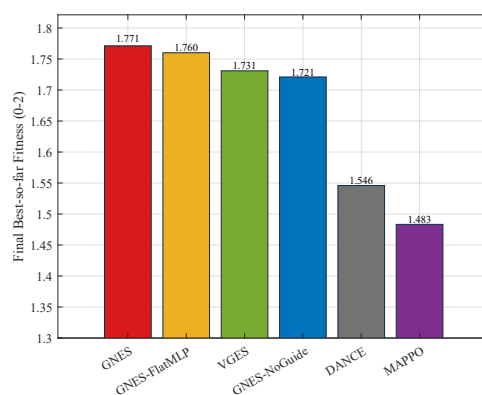
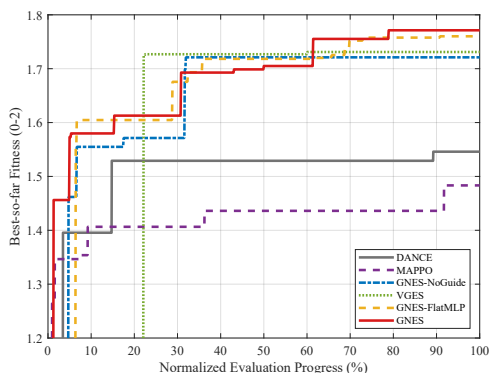
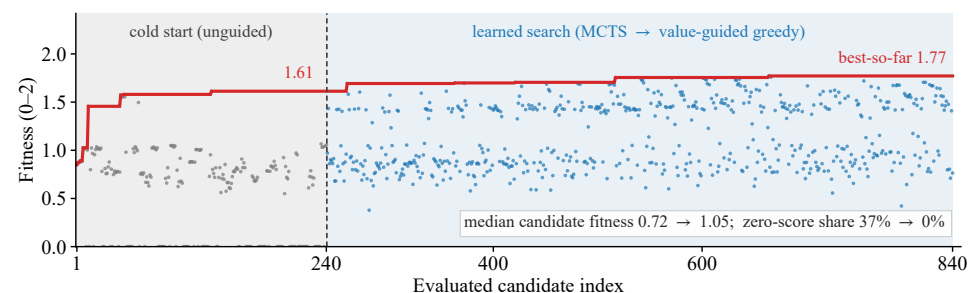


Across All Three Tasks: Evolved symbolic controllers achieve smooth trajectories, balanced inter-UAV spacing, and stable target tracking in both 2D rollouts and 3D AirSim flights.

5. Search Dynamics & Three-Seed Line-Formation Benchmark

NeurIPS 2026 · GNES · Slide 6 / 8

Cold-Start-to-Guided Handover & Trajectories



Handover at Eval 240 (Fig. 6): Switching from unguided GGT edits to GNN + MCTS lifts median fitness 0.72 → 1.05 and eliminates collapsed proposals (36.7% → 0%).

6-Method Comparison over Seeds 20260506–08

Method	Budget	Mean $\pm$ SD	2nd-Half	Key Observation
GNES (Full)	841	1.751 ± 0.022	+0.029	Fastest to 1.72; lowest mid SD (± 0.018)
GNES-FlatMLP	840	1.759 ± 0.002	+0.113	6.7 $\times$ higher mid SD (± 0.121)
GNES-NoGuide	840	1.737 ± 0.014	+0.004	Stagnates after eval 420
VGES (1-Step)	840	1.722 ± 0.027	+0.002	1-step ranking stalls early
DANCE (GP)	2160	1.571 ± 0.024	—	Trails GNES by ≥ 0.13 on all seeds
MAPPO (RL)	2400	1.467 ± 0.017	—	Plateaus under sparse team reward

- **Baseline Separation:** Every program-search variant beats DANCE and MAPPO on **all three seeds** with $\sim 1/3$ of their evaluations (840 vs. 2160/2400).
- **Graph Encoding vs. FlatMLP:** Cuts mid-budget standard deviation by 6.7 $\times$ (± 0.018 vs. ± 0.121).
- **MCTS Look-Ahead vs. VGES / NoGuide:** Sustains second-half improvement (+0.029 vs. +0.002/ + 0.004) near the task ceiling (≈ 1.77).

Wall-Clock Cost: 1313 s (GNES) vs. 633 s (NoGuide), i.e., $\approx +0.8$ s/eval in offline search; **0 s at deployment.**

Closed-Form Controllers & Coordination Motifs

1. Flocking — Motion-Compensated Goal & Neighbor Tracking:

$$\mathbf{v}_{\text{cmd}} = \underbrace{[\mathcal{V}(a, A) + \mathcal{V}(b, B)]}_{\text{target + neighbor attraction}} + \underbrace{\mathcal{V}(d, A) + \mathcal{V}(c, C \cdot A)}_{\text{radial/angular rate feedforward}}$$

2. Entrapping — Density-Normalized Radial Gain:

$$\mathbf{v}_{\text{cmd}} = \underbrace{\mathcal{V}\left(\frac{a - d_0}{b + d_0}, A\right)}_{\text{self-scaling perimeter error}} + \underbrace{\mathcal{V}(b, B)}_{\text{repulsive spacing}}$$

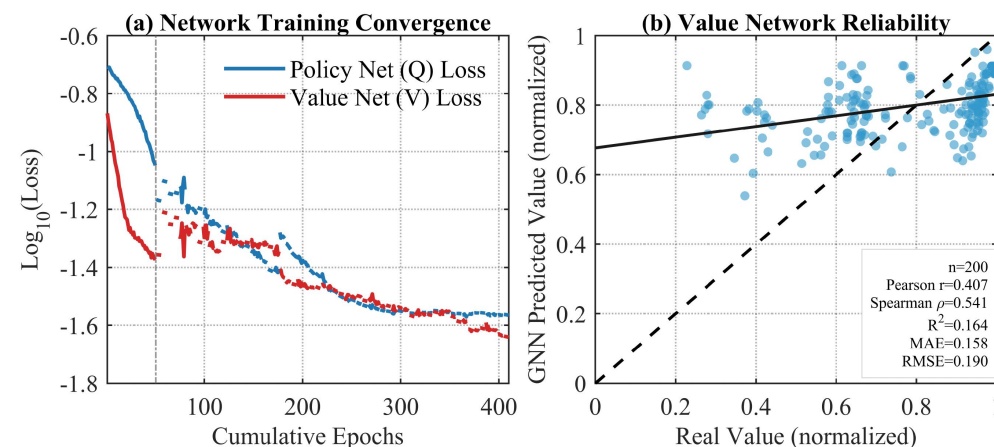
3. Line Formation (Policy C-1) — Damped Origin Anchor:

$$\mathbf{v}_{\text{cmd}} = [\mathcal{V}(a, A) + \mathcal{V}(b, B)] + [\mathcal{V}(d, A) + \mathcal{V}(a, O)]$$

- **Physical Meaning:** Each term carries explicit units and geometry (e.g., dividing radial error $a - d_0$ by local spacing $b + d_0$ automatically damps aggression in dense clusters).
- **Direct Editability:** Engineers can inspect, bound, or tune perimeter radius d_0 and damping terms without retraining.

Zero Runtime Overhead: Deployed controllers execute in microseconds on onboard microcontrollers without GPU/NPU hardware.

Learned GNN Edit Preferences & Diagnostics (Figure 5)



Diagnostic Insight: $Q_\phi(g, a)$ shifts probability mass from growth mutations early in search to parameter/subtree refinement near convergence, stabilizing tree size around 17.3 nodes.

Scope & Honest Limitations

- **Domain Scope (UAV Swarms Only):** The GGT grammar and primitives are tailored to decentralized planar UAV velocity control. Extending to other robotic platforms requires defining domain-appropriate primitives.
- **Simulation-Only Evaluation:** Quantitative optimization uses the planar kinematic simulator; 3D AirSim tests validate closed-loop multirotor execution. Physical outdoor flight requires additional validation under wind gusts, sensor noise, communication delay, and safety envelopes.
- **Statistical & Budget Scope:** The six-method comparison uses three seeds (20260506–08) on a saturating task under unmatched rollout budgets (840–841 vs. 2160/2400), supporting clear baseline separation and search-efficiency ablation findings rather than asymptotic significance claims.

Responsible Research Use: Presented strictly as simulation-based controller synthesis research with human-auditable symbolic outputs.

Summary & Poster Information

Main Contributions:

1. **Program-Edit Formulation:** Casts symbolic UAV swarm controller synthesis as a discrete MDP over validity-preserving Gene Growth Trees (GGTs).
2. **Neural-Guided Evolutionary Search:** Couples topology-aware policy–value GNNs with PUCT MCTS look-ahead to filter weak edits before costly simulator rollouts.
3. **Auditable Symbolic Laws:** Discovers compact velocity controllers (≈ 17.3 nodes) that outperform GP (DANCE) and MARL (MAPPO) baselines with $\sim 1/3$ of the evaluations.

Meet Us at NeurIPS 2026 (Sydney)

Paper: GNES: Neural-Guided Evolutionary Program Search for Interpretable UAV Swarm Control

Authors: Chen Wang[†], Minfang Lu[†], Cheng Zhu, Xiangke Wang, Fei Ming, Yaochu Jin^{*}

Poster ID: #152436 · **Location:** Sydney Exhibition Centre, Hall 1–4