SCIENCE CHINA
Information Sciences

• LETTER •

March 2026, Vol. 69, Iss. 3, 139203:1–139203:2
https://doi.org/10.1007/s11432-025-4733-9

Non-cooperative game-based formation control for underactuated ASVs with prescribed-time performance: an online deep learning method

Wenlong LU[1], Ying ZHAO[1*] & Jun FU[2]

[1]*College of Marine Electrical Engineering, Dalian Maritime University, Dalian 116026, China*
[2]*State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China*

Received 11 March 2025/Revised 5 September 2025/Accepted 23 September 2025/Published online 14 January 2026

Citation Lu W L, Zhao Y, Fu J. Non-cooperative game-based formation control for underactuated ASVs with prescribed-time performance: an online deep learning method. Sci China Inf Sci, 2026, 69(3): 139203, https://doi.org/10.1007/s11432-025-4733-9

Underactuated autonomous surface vehicles (ASVs), originating from the need for efficient maritime operations with reduced actuators, embody profound significance in modern ocean engineering by revealing the evolution of marine robotics and autonomous navigation [1]. However, their coordination in multi-agent formations is hindered by severe challenges caused by time-varying ocean dynamics, communication limitations, and conflicting objectives between individual vessels and collective formation goals [2].

Traditional control methods, reliant on static modeling and constrained to single-objective optimization, lack generalizability and scalability [3], while modern neural network-based approaches, particularly those utilizing recurrent architectures with attention mechanisms [4], suffer from quadratic computational complexity $\mathcal{O}(n^2)$ for n hidden units and strict initial condition requirements $|e(0)| < \rho(0)$, rendering them impractical for real-time maritime applications where vessels cannot be precisely positioned before deployment. Given the trade-offs between computational efficiency and control performance in existing methods, it is imperative to develop an end-to-end, training-efficient, low-cost framework specifically designed for ASV formation control, capable of efficient multi-objective coordination without sacrificing prescribed performance.

Therefore, we propose a comprehensive control framework based on online deep learning and game theory, designed to efficiently coordinate underactuated ASV formations. The framework leverages the temporal learning capabilities of LSTM networks, which capture time-varying dynamics through attention-weighted memory cells, achieving high-quality, adaptive, and generalizable control. Central to its design is the integration of Nash equilibrium seeking strategy with prescribed-time performance functions (PTPF), eliminating traditional initial error constraints while ensuring user-defined convergence.

Method. Consider N underactuated ASVs with dynamics:

$$\dot{\boldsymbol{\eta}}_i = \mathcal{R}(\psi_i)\boldsymbol{\nu}_i, \quad \mathcal{M}_i\dot{\boldsymbol{\nu}}_i = -\mathcal{C}(\boldsymbol{\nu}_i)\boldsymbol{\nu}_i - \mathcal{D}(\boldsymbol{\nu}_i)\boldsymbol{\nu}_i + \boldsymbol{\tau}_i + \boldsymbol{d}_i, \tag{1}$$

where $\boldsymbol{\eta}_i = [x_i, y_i, \psi_i]^{\mathrm{T}} \in \mathbb{R}^3$, $\boldsymbol{\nu}_i = [u_i, v_i, r_i]^{\mathrm{T}} \in \mathbb{R}^3$, and $\boldsymbol{\tau}_i = [\tau_{ui}, 0, \tau_{ri}]^{\mathrm{T}}$ represents underactuated control with zero lateral force.

The framework operates within a non-cooperative game structure $\Gamma = (\mathcal{P}, \mathcal{A}, \mathcal{J})$ over distributed communication topology $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The dual-objective optimization balances global formation and individual tasks:

$$\mathcal{J}_{i,1}(\boldsymbol{\eta}, \boldsymbol{\eta}_{\mathcal{N}_i}(\theta_i)) = \underbrace{\|\boldsymbol{\eta}_i - \boldsymbol{\eta}_{r,i}(\theta_i)\|^2_{\mathcal{Q}_1}}_{\text{individual tracking}} - \underbrace{\gamma \sum_{j\in\mathcal{N}_i} \|\boldsymbol{\eta}_j - \boldsymbol{\eta}_{r,j}(\theta_j)\|^2_{\mathcal{Q}_2}}_{\text{neighbor interaction}},$$

$$\mathcal{J}_{i,2}(\boldsymbol{\theta}) = \underbrace{(\theta_i - \theta_{r,i})^2}_{\text{self progress}} - \underbrace{\delta \sum_{j=1}^{N} (\theta_j - \theta_{r,j})^2}_{\text{team synchronization}}, \tag{2}$$

where θ_i parameterizes individual trajectories evolving as $\dot{\theta}_i = \varrho_{\theta,i}$, with interaction weights $\gamma, \delta < 1/(N-1)$ ensuring convexity.

To incorporate the online learning strategy and LSTM-based time-varying dynamic model (LSTM-TDM) [5], the unknown dynamics is approximated through

$$\dot{\boldsymbol{\nu}}_i = \underbrace{\hat{\mathcal{H}}_t(\boldsymbol{\nu}_i, \boldsymbol{\tau}_i, \boldsymbol{d}_i)}_{\text{LSTM-learned dynamics}} + \mathcal{M}_i^{*-1}\boldsymbol{\tau}_i + \underbrace{\boldsymbol{H}_{0i}}_{\text{modeling error}}, \tag{3}$$

where $\hat{\mathcal{H}}_t = \mathcal{M}_i^{-1}[-\mathcal{C}(\boldsymbol{\nu}_i)\boldsymbol{\nu}_i - \mathcal{D}(\boldsymbol{\nu}_i)\boldsymbol{\nu}_i + \boldsymbol{d}_i] + (\mathcal{M}_i^{-1} - \mathcal{M}_i^{*-1})\boldsymbol{\tau}_i$ represents time-varying uncertainties including unmodeled dynamics, parameter variations, and environmental disturbances, learned via LSTM with attention

$$\alpha_t = \mathrm{softmax}\left(\frac{\mathcal{W}_q\boldsymbol{h}_t \cdot (\mathcal{W}_k\boldsymbol{h}_t)^{\mathrm{T}}}{\sqrt{d_k}}\right), \quad \hat{\mathcal{H}}_t = \sum_{j=1}^{T} \alpha_j \boldsymbol{h}_j. \tag{4}$$

The prescribed-time performance function eliminates initial error constraints through sigmoid modulation:

$$\xi_k = \frac{1}{2}\ln\left(\frac{\delta_{1k}\rho_k(t) + \mathcal{S}(t)e_k}{\delta_{2k}\rho_k(t) - \mathcal{S}(t)e_k}\right), \tag{5}$$

where $\mathcal{S}(t) = \min\{1, \frac{1}{2}[1 + \tanh(\lambda(t - T_s))]\}$ ensures smooth acti-

* Corresponding author (email: zhaoying198916@163.com)

vation at $t = T_s$, and $\rho_k(t)$ evolves according to

$$\dot{\rho}_k(t) = \alpha_k(t) - \beta_k[\rho_k(t) - \rho_{k,f}] \cdot \frac{T_f^{\gamma}}{(T_f + t_0 - t)^{\gamma}}. \quad (6)$$

The singularity at $t = T_f$ created by $(T_f + t_0 - t)^{-\gamma}$ forces exact convergence regardless of initial conditions, eliminating the requirement $|e(0)| < \rho(0)$ of traditional methods.

The Nash equilibrium parallel tracking controller synthesizes formation and individual objectives:

$$\begin{aligned} \gamma_{ui} =& \cos\zeta_{\phi i}\left[\frac{\kappa_{\theta_{di}}\theta_{di} - \varpi_{di}}{\pi_{di}} + \Sigma_{di} + \hat{\bar{\xi}}_i \tanh\left(\frac{\hat{\bar{\xi}}_i\theta_{di}\pi_{di}}{\wp_{\hat{\bar{\xi}}}}\right)\right], \\ \gamma_{vi} =& \sin\zeta_{\phi i}\left[\frac{\kappa_{\theta_{di}}\theta_{di} - \varpi_{di}}{\pi_{di}} + \Sigma_{di} + \hat{\bar{\xi}}_i \tanh\left(\frac{\hat{\bar{\xi}}_i\theta_{di}\pi_{di}}{\wp_{\hat{\bar{\xi}}}}\right)\right], \\ \gamma_{ri} =& \frac{\kappa_{\theta_{\phi i}}\theta_{\phi i} - \varpi_{\phi i}}{\chi_{\phi i}} - \frac{v\cos\zeta_{\phi i} - u\sin\zeta_{\phi i}}{\zeta_{di}} \\ &+ \frac{\hat{\bar{\xi}}_i}{\zeta_{di}}\tanh\left(\frac{\hat{\bar{\xi}}_i\theta_{\phi i}\pi_{\phi i}}{\zeta_{di}\wp_{\bar{\xi}}}\right) + \mu_{ri}\tanh\left(\frac{\mu_{ri}\theta_{\phi i}\pi_{\phi i}}{\wp_{\beta_{ri}}}\right), \end{aligned} \quad (7)$$

where $\pi_{ki} = \mathcal{S}(t)(\delta_{1ki}\rho_{ki}+\delta_{2ki}\rho_{ki})/[2(\delta_{1ki}\rho_{ki}+e_{ki})(\delta_{2ki}\rho_{ki}-e_{ki})]$ and the underactuation compensation through auxiliary dynamics

$$\boldsymbol{\mu}_i = [\alpha_{ui}\tanh(\beta_{ui}), \alpha_{vi}\tanh(\beta_{vi}), \alpha_{ri}\tanh(\beta_{ri})]^{\mathrm{T}} \quad (8)$$

with adaptive laws $\dot{\beta}_{pi} = \cosh^{-2}(\beta_{pi})[\mu_{pi}\beta_{pi} - \omega_{pi}]/\alpha_{pi}$ ensuring bounded evolution.

The composite control law integrates all components:

$$\tau_{pi} = [-\kappa_{pi}z_{pi} - \hat{\mathcal{H}}_{tpi} + \dot{\gamma}_{pi} + \mu_{pi}\beta_{pi}]/m^*_{pi}, \quad p \in \{u, r\}, \quad (9)$$

where $z_{pi} = \nu_{pi} - \gamma_{pi} - \mu_{pi}$ represents tracking errors in the transformed coordinates.

The online LSTM training employs Adam optimizer with learning rate $\alpha = 10^{-6}$, momentum parameters $\beta_1 = 0.8$, $\beta_2 = 0.9$, minimizing the loss function

$$\mathcal{L}_t = \mathbb{E}[\|\mathcal{H}_t - \hat{\mathcal{H}}_t\|^2] + \lambda_r\|\mathcal{W}\|^2. \quad (10)$$

Experiments and results. The experiments were conducted to evaluate formation performance, learning efficiency, and prescribed-time convergence. First, a comparative experiment assessed coordination quality under various disturbance conditions ($\boldsymbol{d}_i = 0.5\sin(0.1t) + 0.3\mathrm{rand}(t)$). Compared to state-of-the-art methods [4], the proposed framework demonstrated superior Nash equilibrium convergence ($\|\nabla\mathcal{J}_i\| < 10^{-3}$) and formation maintenance, particularly under parameter uncertainty.

Second, an ablation study validated the LSTM-TDM effectiveness. For temporal dynamics learning, the framework exhibited rapid convergence (loss $< 10^{-4}$ within 50 iterations) with computational complexity $\mathcal{O}(T/p^2)$ where T is sequence length and p is parallelization factor, significantly outperforming traditional RNN approaches requiring $\mathcal{O}(T \cdot n^2)$. The attention mechanism successfully identified critical temporal patterns, improving approximation accuracy to $\|\hat{\mathcal{H}}_t - \mathcal{H}_t\| < 10^{-3}$.

Lastly, PTPF validation demonstrated convergence at user-defined $T_f = 60$ s regardless of initial conditions: $\boldsymbol{\eta}_1(0) = [40, 20, \pi/2]^{\mathrm{T}}$, $\boldsymbol{\eta}_2(0) = [50, 20, \pi/2]^{\mathrm{T}}$, $\boldsymbol{\eta}_3(0) = [60, 20, \pi/2]^{\mathrm{T}}$ with initial errors exceeding traditional bounds by 300%. The sigmoid activation at $t = 30$ s ensured a smooth transition without performance degradation. Detailed experimental configurations: Adam optimizer ($\alpha = 10^{-6}$, $\beta_1 = 0.8$, $\beta_2 = 0.9$), LSTM with 10 hidden neurons, attention dimension $d_a = 8$. Figure 1 shows the formation trajectories, validating the effectiveness of the proposed algorithm.

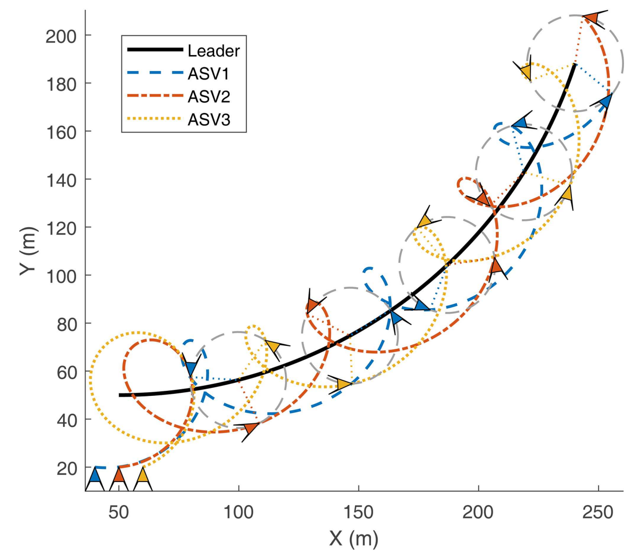

Figure 1 (Color online) Formation trajectories of ASVs executing target encirclement with radius $R = 10$ m and phase separation $2\pi/3$. Despite large initial errors, ASVs converge to the desired formation at the prescribed time.

Discussion and conclusion. We present a comprehensive control framework for underactuated ASV formation, built on online LSTM learning enhanced by Nash equilibrium seeking and PTPF. The framework demonstrates exceptional adaptability to complex ocean dynamics through attention-weighted temporal learning, superior multi-objective coordination achieving proven Pareto optimality, and guaranteed prescribed-time performance without initial constraints. Key innovations include (1) temporal dynamics capture through attention-weighted LSTM eliminating manual modeling with complexity $\mathcal{O}(T/p^2)$, (2) Nash equilibrium balancing individual and collective objectives with convergence guarantee under convexity conditions $\gamma, \delta < 1/(N-1)$, and (3) PTPF with sigmoid modulation ensuring user-defined convergence regardless of initial errors.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 62533004, 62373072, 62073054) and Postdoctoral Natural Science Foundation (Grant No. 2024M764265).

References

1 Naderolasli A, Shojaei K, Chatraei A. Leader-follower formation control of Euler-Lagrange systems with limited field-of-view and saturating actuators: a case study for tractor-trailer wheeled mobile robots. Eur J Control, 2024, 75: 100903
2 Xie F, Liang G, Chien Y R. Robust leader-follower formation control using neural adaptive prescribed performance strategies. Mathematics, 2024, 12: 3259
3 Peng Z H, Wang D, Li T S. Predictor-based neural dynamic surface control for distributed formation tracking of multiple marine surface vehicles with improved transient performance. Sci China Inf Sci, 2016, 59: 092210
4 Xu J, Xia F, Gu N, et al. Parallel path following of underactuated cyber-physical autonomous surface vehicles based on meta-learning extended state observer. IEEE Trans Veh Technol, 2024, 73: 8912–8925
5 Peng Z, Xia F, Liu L, et al. Online deep learning control of an autonomous surface vehicle using learned dynamics. IEEE Trans Intell Veh, 2024, 9: 3283–3292