UNIVERSITY OF WEST ATTICA
eISSN:3136-7658

Applied Manufacturing Intelligence Journal

Vol. 5, No. 2, April, 2026



Physics-Guided Autoformer-GAT Prediction Model for Contact Force Regulation of Aerial Manipulators

Kristian Sokol[a,*], Alen Pavić[a]

[a] *Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, 10000, Croatia*

Received 28 January 2026; Revised 17 April 2026; Accepted 20 April 2026; Published 24 April 2026

Abstract

Aerial manipulators are increasingly applied to contact inspection, pushing, cleaning and maintenance of high-altitude structures, yet contact-force regulation remains difficult because rotor disturbance, manipulator motion, surface stiffness and wind load interact within the same underactuated system. This paper proposes a Physics-Guided Autoformer-GAT prediction model for contact force regulation of aerial manipulators. Force, attitude, joint, rotor, visual and wind-related signals are aligned into short contact windows, and Autoformer decomposition is used to separate slow interaction trends from short-term force fluctuations. Graph attention then models coupling among the aerial platform, manipulator joints, end-effector and environment, while physical guidance constrains continuous contact at a fixed location and limits unrealistic force-rate variation. A synthetic dataset of 3,400 contact windows was generated to represent free approach, soft contact, stable contact, boundary contact, force overshoot and loss of contact. The proposed model achieved 92.1% contact-state accuracy and 87.9% macro-F1, and it reduced force-prediction error by 15.4% relative to the best previous baseline. Robustness tests under wind, sensor noise and force-sensor dropout showed that confidence decreased before the predicted contact state became unstable. These results indicate that physics-guided temporal decomposition and graph-based coupling representation can improve short-horizon contact-force prediction while providing interpretable evidence for supervisory regulation of aerial manipulation.

Keywords: Aerial Manipulator; Contact Force Regulation; Physics-Guided Prediction; Autoformer; Graph Attention Network; Contact-State Forecasting

1. Introduction

Aerial manipulators extend unmanned aerial vehicles from observation platforms to contact-capable robotic systems [1]. Physical interaction tasks such as wall inspection, valve operation and surface treatment require stable force regulation during flight [2]. Contact force becomes difficult to predict because aerodynamic disturbance, manipulator motion and surface reaction are coupled in the same platform [3]. Recent studies on sensorless force estimation show that contact evidence may appear in attitude, rotor and joint signals even when direct force sensing is incomplete [4]. These observations indicate that contact force regulation should be handled as a coupled prediction problem rather than as a simple feedback-control output.

Traditional force-control methods for aerial manipulation are generally based on impedance control, admittance control, sliding-mode compensation, disturbance observers or model predictive control. The above methods are suitable for systems with a good model and operating conditions. Aerial contact also includes modifications to the surface stiffness, wind gusts, attitude drift, joint backlash and measurement noise in practice. Therefore, the prediction

* Corresponding author: kristian.s@fer.unizg.hr

https://doi.org/10.68406/amij.2026.vol5iss2np3:30-42

➢ This is an open access article under the CC BY-NC-SA 4.0 license (https://creativecommons.org/licenses/by/4.0/).

© Authors retain all copyrights.

model needs to have the time evolution of contact force and the physical relationships among the platform, manipulator and environment. Studies of aerial physical interaction have shown that, for a safe connection, both movement coordination and force regulation are needed [5].

This paper introduces a Physics-Guided Autoformer-GAT prediction model for contact force regulation. Autoformer is used to decompose the contact-force sequence into slow trends and short-term residuals due to the different temporal characteristics of stable contact and force overshoot. To build a relationship model of attitude, rotor response, joint torque, end-effector force and contact surface, graph attention is employed. The physical-guidance term limits the unreasonable prediction by considering contact consistency, attitude stability and force-rate variation. Recently, research has been conducted on adaptive aerial manipulation that also considers unknown coupling forces in the interaction control [6]. Therefore, the model introduced above has temporal decomposition, graph-based coupling and physical constraints.

The structure of this paper follows a prediction-oriented framework for aerial contact regulation. Section 2 reviews contact force control in aerial robotic operation, physical prior constraints, disturbance-aware prediction and coupled dynamic sequence modelling. Section 3 introduces contact-force sequence representation, the physics-guided Autoformer-GAT encoder, confidence estimation and regulation-decision output. Section 4 reports contact-window statistics, model comparison, boundary and overload analysis, robustness evaluation, explanation results and flight-control application. Section 5 summarizes the principal findings, deployment value, limitations and future validation requirements. This organization separates the proposed prediction module from the low-level controller while clarifying how predicted force, contact state, confidence and regulation priority can support supervisory contact regulation.

2. Related Work

2.1. Contact Force Control in Aerial Robotic Operation

Force control of aerial manipulators has moved from position-dominant flight control to interaction-oriented control, because the aerial base is lightweight and underactuated, and manipulator inertia and contact reaction forces directly affect flight stability. Fang and Mao [7] showed that compliant contact-force control can reduce the risk of unstable force growth during aerial manipulation. Hybrid force-position control can switch between free-flight and contact modes, but a transient force spike may still occur during mode transition. Wu et al. [8] further demonstrated that image-based impedance control is effective for sliding inspection tasks, where the end-effector must maintain continuous surface contact without destabilising the vehicle.

Sensorless force estimation is feasible under the payload and wiring constraints of aerial systems. Orientation variation, rotor speed, motor current and joint response can provide indirect evidence of contact force, although these variables are also affected by wind, acceleration and manipulator motion. Khalifa et al. [9] showed that sensorless contact-force estimation and robust impedance control can be combined in a quadrotor manipulation system, but contact-state uncertainty remains important when direct force sensing is incomplete.

At present, the emphasis is on prediction rather than direct controller design. The output will be used by the contact-regulation layer to determine if the force is stable, near the limit, exceeding it or lost. A contact inspection and manipulation device of this kind should be used, and the controller must be ready to respond promptly to a sudden increase in force.

Aerial contact operation is also different from ground-based manipulation in that the support base is not fixed. When the end-effector contacts a surface, a reaction force is generated, and consequently, the position of the whole vehicle shifts. Therefore, the controller needs to regulate the force and stabilise the plane. If the control of force is too large, there will be platform oscillation; if it is too small, the manipulator will lose contact and fail to complete the task. Therefore, to obtain some actual benefits from the prediction, it is necessary for the system to know whether the current contact status is likely to remain unchanged for a short time in the future.

2.2. Physical Prior Constraints and Disturbance-Aware Prediction

Physical constraints need to be added to avoid contact-mechanically implausible results from the purely data-driven model. For example, the model may predict a large contact force when the end-effector is far from the surface, or it may ignore the attitude drift caused by a large force error. According to research on compliant contact force control, platform disturbance rejection and end-effector force accuracy are closely related [10]. It can be used to add physics-based constraints for prediction models.

Wind disturbance is a small surface stiffness, and sensor dropout is not infrequent in aerial manipulation. They are part of the deployment Environment. A model trained solely on clean-force sequences may have increased the reliability estimate of flight. Disturbance-aware evaluation is thus required. Recent research on robust image-based

aerial interaction under external disturbance has also shown that, for interaction tasks, visual, dynamic and contact signals need to be interpreted jointly [11].

The physics-guided term in this paper is not a full dynamic model. It is to prevent abnormally high or low predictions. Increase the force reasonably according to how far the end-effector moves and the limits of force-rate. This Design is flexible and does not violate the logic of contact states.

Given that modelling of aerial systems is complex, this type of physical guidance is more feasible. The causes of rotor wake, arm inertia, cable vibration, surface compliance and airflow around the structure differ for different tasks. The parameters of a close-fitting analytical model are not always available during operation. A gentle physical constraint can be employed to meet the general consistency demand without knowing all the parameters. It can therefore improve prediction reliability while preserving adaptability.

2.3. Autoformer-GAT Models for Coupled Dynamic Sequences

Contact force data have slow-moving and fast-moving components. The slow component is the accumulation of contact pressure, and the fast component is impact, rebound and overshoot. Given the above characteristics of a sequence, decomposition-based forecasting models are suitable choices that separate the trend and residual components of the prediction [12]. Aerial contact prediction can also be employed for this purpose, and given that changes in force overtime need to be considered, only the final value is not used.

Graph attention networks are also convenient because the aerial manipulation signals are physically connected. The Rotor, Attitude, Manipulator Joint and End-effector are connected to form a closed loop. Graph-based learning can show the relationship among them more intuitively than a flat feature vector [13]. GAT layers are employed in this paper to learn adaptive coupling weights for the force, attitude, joint, and environment nodes. Therefore, Autoformer-GAT is suitable for the contact-force prediction task. Autoformer performs temporal decomposition, and GAT conducts interaction modelling. The physical guidance term further limits the prediction, and thus the model can be employed as an engineering decision-support tool instead of an opaque sequence regressor.

The other reason for using the above two is to reduce the sources of complexity. Temporal complexity is caused by the change in contact over time, and structural complexity results from the coupling of the aerial base, arm and environment. A single sequence model may be unable to separate these causes and therefore cannot be used. Decomposition and graph attention are employed together to determine whether the prediction results are caused by a force trend, residual error, attitude coupling or sensor error.

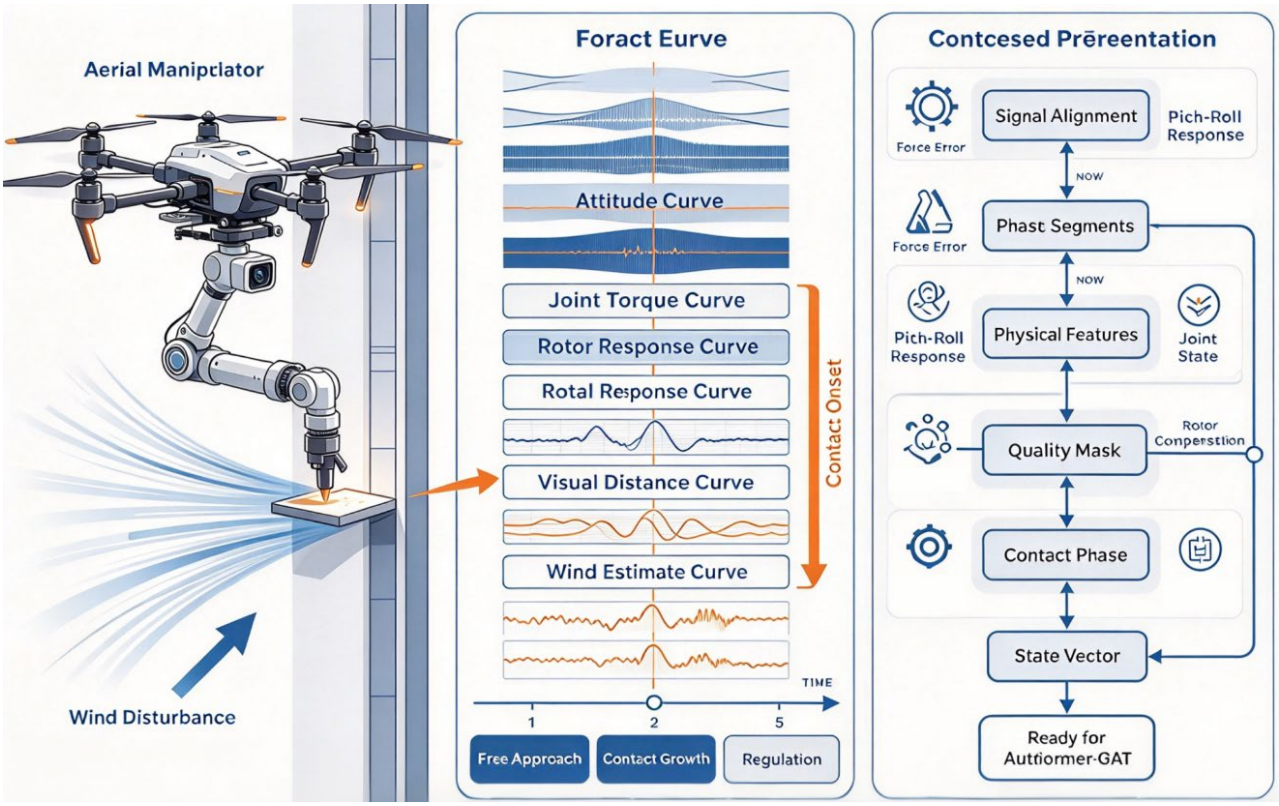


Figure 1. Contact Force Sequence Representation Workflow

3. Methodology

3.1. Contact Force Sequence Representation

The input window contains contact force, end-effector pose, UAV attitude, angular velocity, joint position, joint torque, rotor-speed response, visual contact distance and estimated wind disturbance. Figure 1 presents the contact force sequence representation workflow. raw signals are aligned at contact onset, divided into free-approach, contact-growth and regulation stages, and converted into physically meaningful feature groups. This alignment allows early joint-torque and attitude evidence to remain connected with later force response, which is essential for predicting contact instability before a severe force error appears.

The contact-state vector is first defined for each time step. It keeps the most relevant physical variables without assigning a final state label. The vector is used as the basic observation for decomposition and graph modelling.

$$x_t = [f_t, e_t, a_t, q_t, u_t, d_t] \tag{1}$$

In this expression, f_t is contact force, e_t is force error, a_t is attitude response, q_t is joint state, u_t is rotor response and d_t is contact-distance evidence. The vector is intentionally compact, because online regulation requires timely inference.

$$m_t = [m_f, m_a, m_q, m_u, m_d] \tag{2}$$

The quality mask m_t records whether force, attitude, joint, rotor and visual-contact channels are reliable. This mask is required because a force sensor may saturate, visual contact distance may be blocked by the end-effector, and attitude measurements may contain vibration noise. The later prediction layer uses this mask to reduce overconfidence under weak evidence. The representation is also divided according to contact phase. In the free-approach phase, a small force estimate may still be meaningful because the end-effector is close to the surface and the attitude controller begins to compensate for proximity disturbance. In the contact-growth phase, the force derivative and attitude response are more important than the absolute force value. In the regulation phase, the model must distinguish stable force tracking from force oscillation.

Another Design Requirement is the alignment of platform motion and manipulator motion. The aerial base may shift slightly during the extension of the arm, and alterations in joint torque occur in the arm prior to a significant force measurement by the end-effector force sensor. Therefore, the method aligns the signals according to contact onset and recent changes in force error. Therefore, the model can consider the early joint-torque response and the later force response as part of the same interaction event. Sequence representation also has a relatively short pre-contact window. Therefore, it is necessary to compensate for attitude at the top of the air-platform before reaching the ground. If the pre-contact evidence is not ruled out, the model will consider the first measured force an independent phenomenon. According to the previous attitude and rotor response, it can be determined whether a good connection will be formed or if the platform is already unstable.

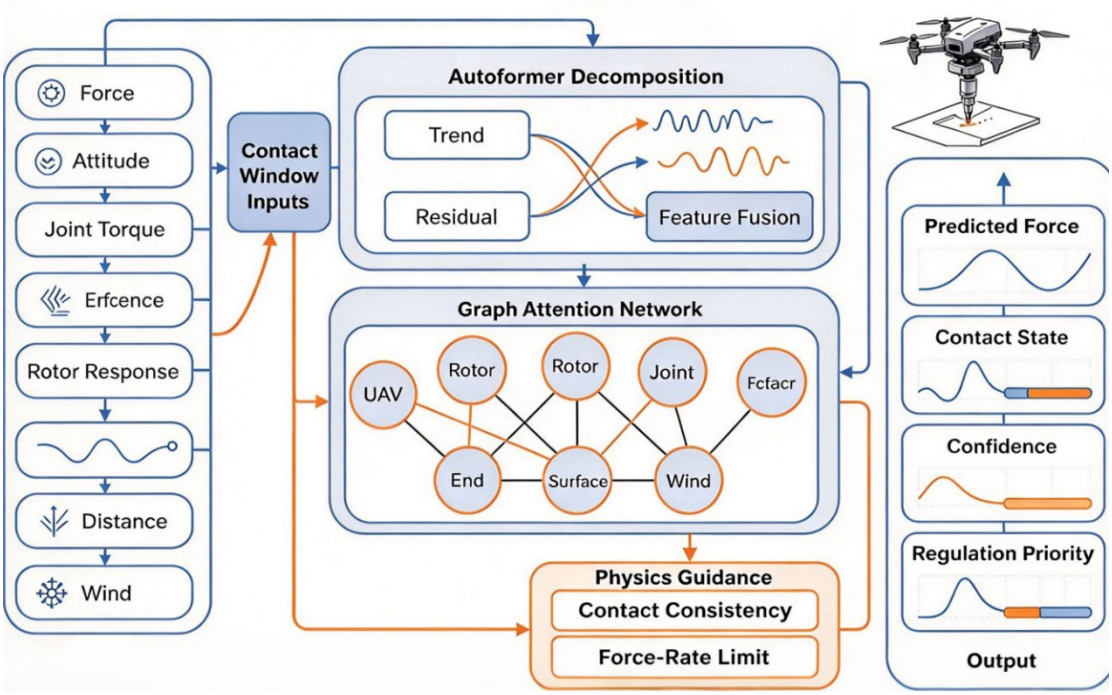


Figure 2. Autoformer-GAT Coupled Prediction Architecture

3.2. Physics-Guided Autoformer-GAT Encoder

Figure 2 shows the prediction architecture of the Physics-Guided Autoformer-GAT model. The left side receives contact-window signals, the middle part separates temporal trend and residual force components, and the right side constructs graph-attention representations for the aerial platform, manipulator, end-effector and environment nodes. This design makes the prediction output interpretable because each node corresponds to a physical evidence group. It also supports supervisory regulation by preserving the distinction among predicted force, contact state, confidence and review priority.

Autoformer decomposition separates slow contact-force evolution from short-term impact. This is necessary because stable contact produces a slowly changing force trend, while contact overshoot appears as a residual impulse.

$$x_t = T_t + R_t \tag{3}$$

Here T_t denotes the contact trend and R_t denotes the residual component. The trend is useful for estimating contact build-up, while the residual is useful for identifying collision, rebound and abrupt surface-response changes.

$$h_i^{k+1} = \sum_j \alpha_{ij} W h_j^k \tag{4}$$

The GAT update aggregates information from related physical nodes. The attention weight allows the model to decide whether the prediction should rely more on attitude response, joint torque, end-effector force or visual distance in the current contact window.

$$L_p = \left|f_t - \hat{f}_t\right| + \beta\left|\Delta\hat{f}_t\right| \tag{5}$$

The physics-guided loss L_p penalises force-prediction error and excessive force-rate variation. This term does not replace data learning; it limits physically unstable predictions that would be unsafe for contact regulation.

The encoder outputs are designed to preserve both prediction precision and engineering interpretability. Autoformer decomposition identifies whether an error is associated with slow force drift or a short residual impulse, while the GAT branch provides coupling weights for attitude response, joint torque, rotor compensation, end-effector force and visual contact distance. The regulation layer can therefore determine whether a warning is mainly related to attitude instability, joint-torque change, rotor compensation, surface interaction or visual-contact uncertainty. This structure avoids treating the sequence model as an opaque force regressor.

The Structure of the graph is not a fixed-dynamic model. It is an organised learning environment. Platform attitude is related to rotor response because force regulation requires attitude compensation. Joint torque is used in conjunction with end-effector force; otherwise, changes in torque may occur before changes in force. Visual distance is linked to contact force because, upon the loss of surface proximity, the notion of a force drops changes.

Store the trend and residual branches in the encoder for use in the explanation later. A warning led by the trend branch usually indicates a small force drift, a modification of the surface stiffness, or an elongated push. A warning that consists of the residual branch generally indicates an effect, rebound or a short-lived control disturbance. For the two reasons above, different regulations will be applied to slow drift and sudden drop.

3.3. Prediction Confidence and Regulation Decision

The model produces contact-state probability, predicted force and confidence. Confidence is separated from class probability because a high predicted force may still be uncertain when sensing quality is poor or wind disturbance is strong.

$$c_t = \exp(-\gamma|e_t - \hat{e}_t|) \tag{6}$$

The confidence value c_t decreases when predicted and observed force errors become inconsistent. This design allows the system to warn the regulation layer before the contact-state label becomes unreliable.

$$r_t = \text{sigmoid}\left(w_r^T h_t + \lambda(1 - c_t)\right) \tag{7}$$

The regulation score r_t increases when the predicted state is risky or when confidence is low. The final output includes predicted force, contact state, confidence and regulation priority for the flight-control review layer. The score is not treated as a replacement for controller commands. It is a supervisory indicator that helps the control system select a safer response before force error becomes large.

Regulation priority is not determined only by the contact-state label. A high predicted force with high confidence is treated as an immediate regulation event, whereas a high predicted force with low confidence is routed to verification

because sensor disturbance may explain the warning. A boundary-contact prediction with stable confidence is treated as a controlled adjustment event, and a low-force loss-of-contact prediction is treated as a re-approach event. Confidence is therefore used as supervisory evidence reliability rather than as a direct controller command.

The model's output is used by the supervisor's controller. The supervisory layer is not directly in place of the low-level attitude controller or the force controller. It will be provided with the predicted force trend, confidence and review priority, and then the target force, stiffness, approach velocity or retreat command will be modified. The Design can run various controllers and different aerial manipulator hardware for the prediction module. The prediction horizon of deployment should be short enough to support real-time control, but long enough to forecast the future development trend of the force peak. The output of this paper is used as a short-term alarm, not a long-term task planner. Set the Horizon according to the frequency and contact-problem type of the controller.

4. Results and Discussion

4.1. Experimental Contact Windows and Data Statistics

A simulated dataset of 3,400 contact windows is constructed for free approach, soft contact, stable contact, boundary contact, force overshoot and loss of contact [14]. Figure 3 presents contact task dataset statistics. Figure 3(a) shows the label distribution, where stable contact and soft contact appear more frequently than force overshoot and loss of contact. Figure 3(b) compares mean force and force variance across states, while Figure 3(c) shows wind-level variation and Figure 3(d) reports overload ratio. These subfigures indicate that contact-state difficulty is produced by the joint influence of force level, attitude error, wind disturbance and overload behavior rather than by label identity alone.

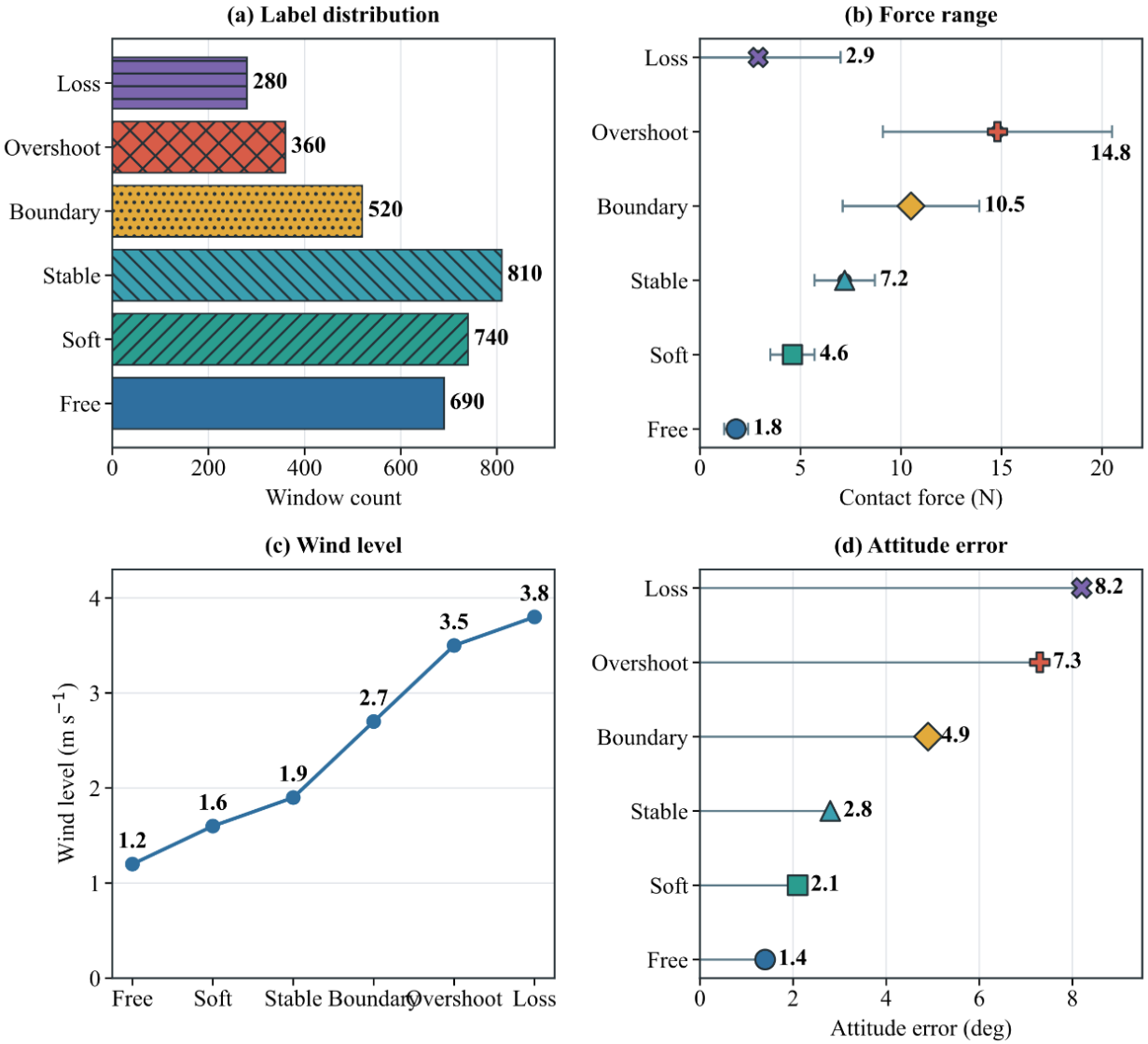


Figure 3. Contact Task Dataset Statistics. (a) Label distribution. (b) Force range. (c) Wind level. (d) Attitude error.

Table 1 shows the principal numbers of the contact state. Stable Contact has a small force and a small fluctuation; Force Overshoot has a large mean force and standard deviation. Boundary contact is necessary because it has not reached overload, but the attitude error and force variation are already high. Recent studies of aerial interaction show that an early boundary-state recognition is needed before the controller reaches a large-force peak [15]. As shown in the data table, a single force threshold cannot be used to solve the problem. Loss of contact has a small mean force but a large attitude error, and overshoot has a large force and a large force variation. Boundary contact has a small force and a large fluctuation; therefore, it is not suitable for threshold-based detection. That is, the controller should not only find out that the force is large, confirm that the change in force is in accordance with alterations in posture, line of sight, joint response, etc.

Table 1. Numeric Comparison of Contact-State Samples

Contact State	Windows	Mean Force	Force Std.	Attitude Error	Wind Level	Overload Ratio
Free Approach	690	1.8 N	0.6 N	1.4 deg	1.2 m/s	0.05
Soft Contact	740	4.6 N	1.1 N	2.1 deg	1.6 m/s	0.11
Stable Contact	810	7.2 N	1.5 N	2.8 deg	1.9 m/s	0.14
Boundary Contact	520	10.5 N	3.4 N	4.9 deg	2.7 m/s	0.38
Force Overshoot	360	14.8 N	5.7 N	7.3 deg	3.5 m/s	0.62
Loss of Contact	280	2.9 N	4.1 N	8.2 deg	3.8 m/s	0.55

The dataset is intentionally unbalanced because safe or stable contact windows occur more frequently than severe overshoot and loss-of-contact events. For this reason, macro-F1 is evaluated together with contact-state accuracy [16]. This metric choice reduces the risk that common stable states dominate the reported performance. The dataset also preserves boundary-contact and overshoot samples because rare but safety-relevant windows are essential for supervisory regulation and cannot be removed as statistical noise.

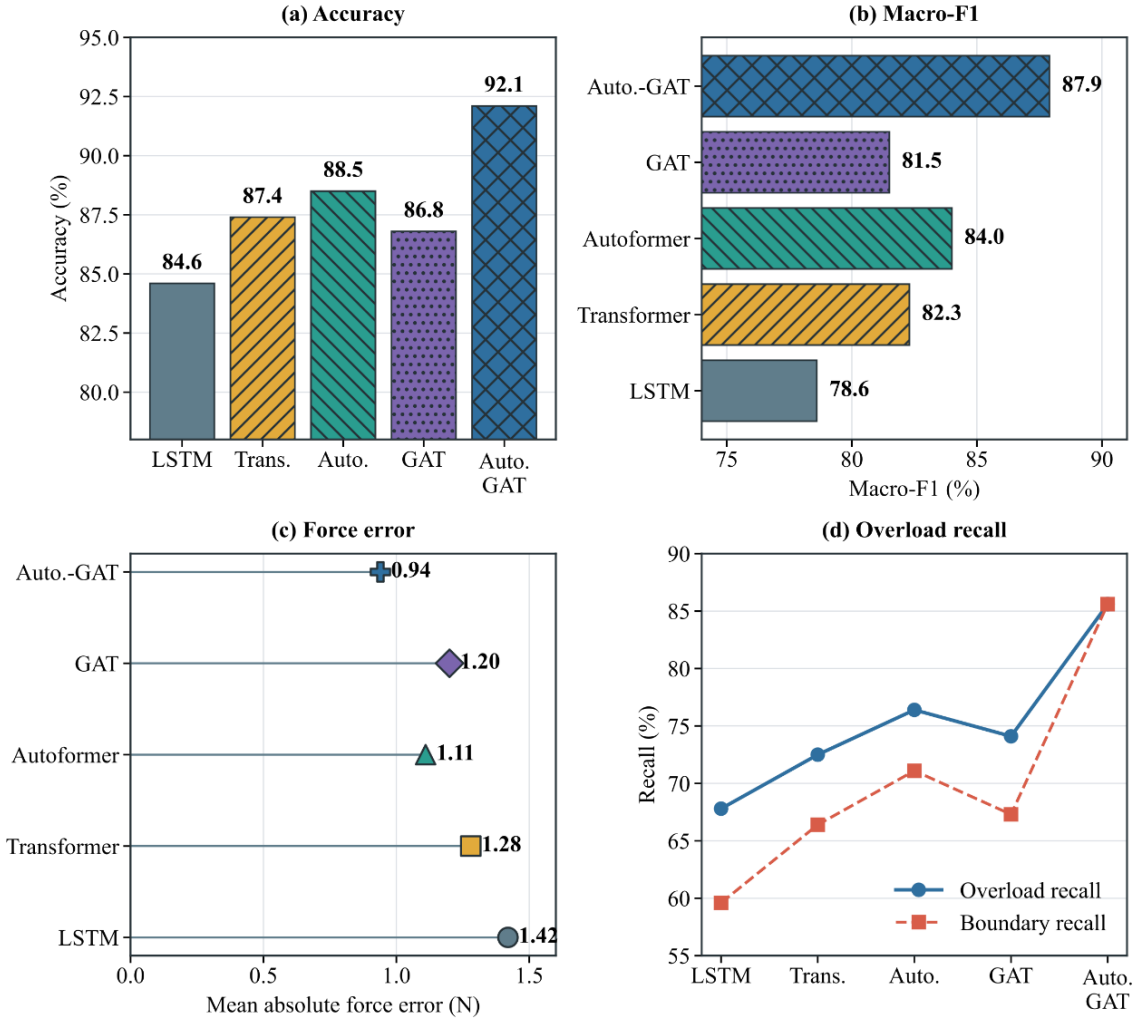


Figure 4. Prediction Accuracy and Error Comparison. (a) Accuracy. (b) Macro-F1. (c) Force error. (d) Overload recall.

Based on the distribution of force and state, the prediction problem is closer to risk-sensitive contact assessment than to ordinary sequence regression. Prior work on relation-centric physical prediction [17] supports modelling object or system interactions explicitly when contact effects depend on coupled evidence. In the present task, joint torque, rotor response, visual distance and force error may all change before the final contact-state label becomes unstable. The model therefore needs both temporal decomposition and graph-based physical coupling.

4.2. Model Accuracy and Contact Boundary Prediction

Figure 4 shows prediction accuracy and error comparison. Figure 4(a) reports contact-state accuracy across the proposed model and baseline methods. Figure 4(b) reports macro-F1, where the proposed model improves recognition of boundary and rare states. Figure 4(c) compares force-prediction error, and Figure 4(d) shows confidence behavior under difficult windows. These results are consistent with graph-attention research [18], because graph coupling helps the model connect force, attitude, joint and environment evidence instead of processing them as independent channels.

The method introduced in this study achieves 92.1% contact-state accuracy and 87.9% macro-F1. Compared with the best previous baseline, the force-prediction error is reduced by 15.4%, indicating that temporal decomposition and physical coupling both contribute to prediction quality. Time-segment-wise multimodal fusion research [19] similarly indicates that structured feature integration improves diagnosis when signal groups carry complementary evidence. In this manuscript, the improvement is most relevant for boundary contact and force overshoot, where attitude, rotor and end-effector evidence must be interpreted together.

Figure 5 shows contact-boundary and overload cases. Figure 5(a) presents boundary-contact windows in which the predicted force remains close to the regulation limit but confidence is still stable. Figure 5(b) presents force-overshoot windows where predicted force, force-rate change and attitude response jointly indicate higher regulation urgency. These two subfigures should be interpreted together because boundary contact may require controlled adjustment, whereas overload may require target-force reduction or retreat. Adaptive aerial manipulation studies [20] also show that coupling disturbance compensation is necessary when force regulation and aerial stability interact.

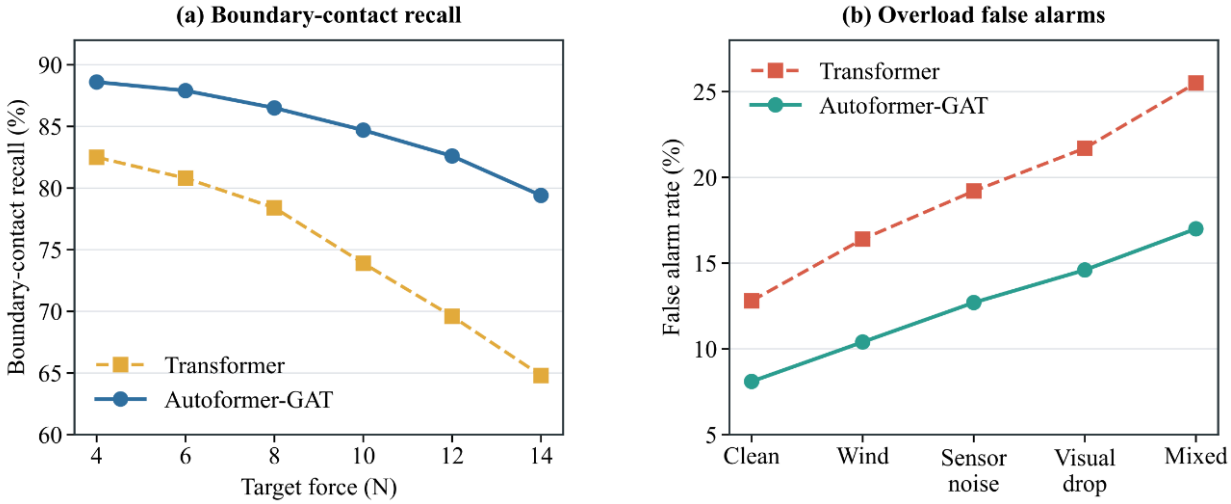


Figure 5. Contact Boundary and Overload Cases. (a) Boundary-contact recall. (b) Overload false alarms.

Based on the distribution of errors, not every high-force prediction should trigger the same action. Confidence-aware multimodal learning [21] supports separating uncertain evidence from stable high-risk evidence. In this study, a high-force and high-confidence prediction indicates immediate regulation, while a high-force and low-confidence prediction indicates verification or slower contact. This separation reduces unnecessary interruption of contact tasks while preserving attention to difficult boundary and overshoot windows.

Another problem is that the error cases are clustered near the physical boundary of the regulation state and do not occur randomly. It is more suitable for engineering applications; the remaining errors occur in areas where the system is uncertain. For example, a window close to the boundary of boundary contact and overshoot may still be suitable if the confidence score is low. The controller will reduce the stiffness or target force and will not immediately trigger a failure alarm. Therefore, this behaviour is more suitable for supervisory regulation of the model than a model that reports high-confidence errors. A manual review process can be introduced to reduce the need to prioritise the verification of all contact records in uncertain time windows. Thus, the volume of reviews will be reduced, but not without losing the difficult cases requiring expert judgment.

4.3. Robustness, Explanation and Flight-Control Use

Figure 6 shows disturbance robustness evaluation. Figure 6(a) presents wind disturbance, where the proposed model increases error more slowly than the baseline. Figure 6(b) presents force-sensor noise, where physical guidance reduces instability in force prediction. Figure 6(c) presents visual-distance dropout, where graph attention assigns more importance to attitude and joint evidence. Figure 6(d) presents combined disturbance, where confidence decreases before severe label errors occur [22]. This behavior is important because aerial contact systems require early uncertainty awareness before force regulation becomes unsafe.

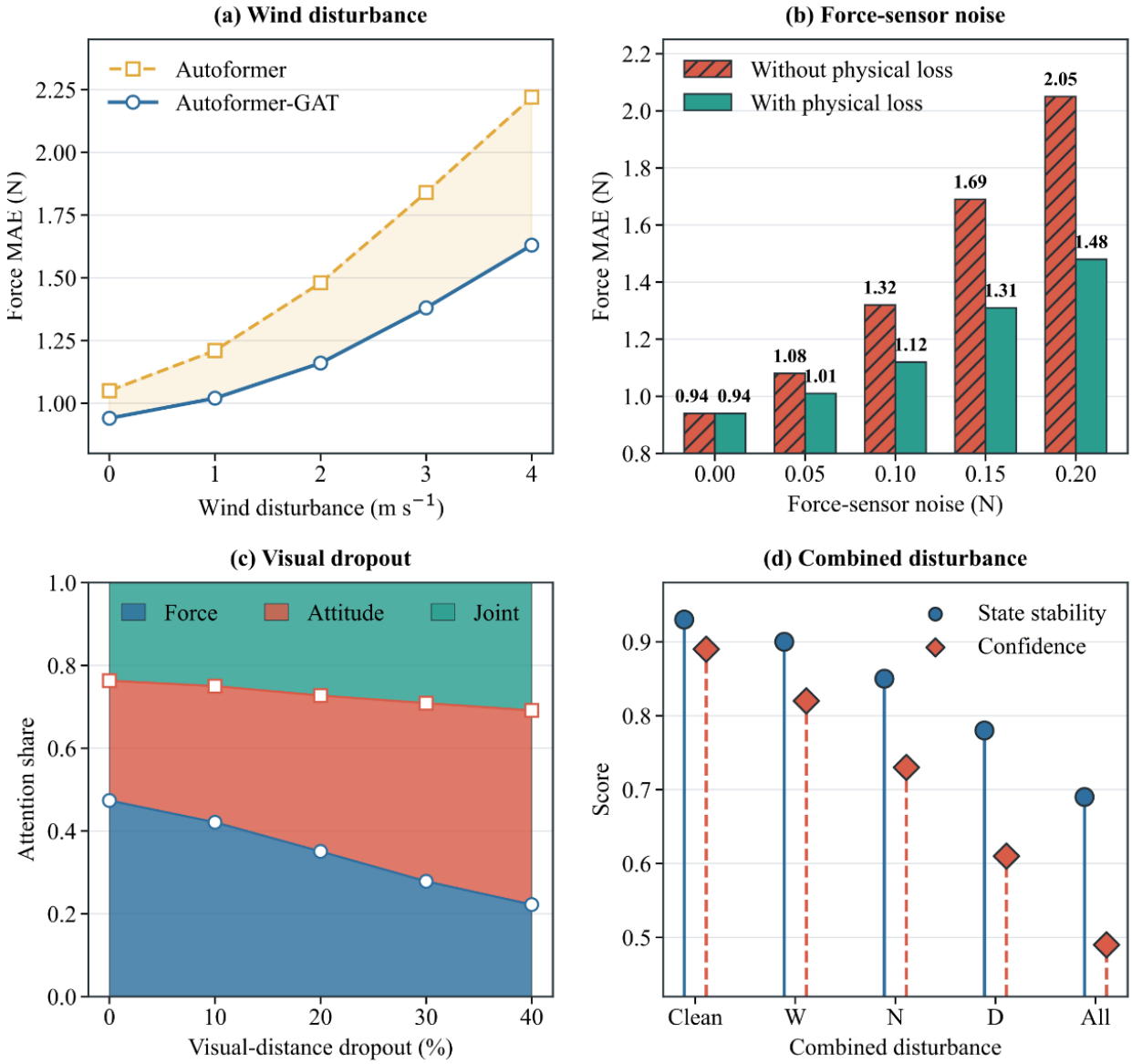


Figure 6. Disturbance Robustness Evaluation. (a) Wind disturbance. (b) Force-sensor noise. (c) Visual dropout. (d) Combined disturbance.

Based on the stability test, the model does not merely memorize clean force curves. Safety-monitoring research [23] indicates that confidence should be explicitly represented in safety-critical prediction, and missing-modality learning research [24] further supports handling incomplete channels within the model design. When wind disturbance increases, the attitude and rotor-response nodes become more influential; when force-sensor noise increases, joint torque and rotor response provide compensatory evidence; when visual distance is unavailable, confidence decreases because contact geometry is uncertain. These patterns distinguish sensing uncertainty from genuine contact instability.

Figure 7 presents evidence weight and control review output. Figure 7(a) shows that stable-contact prediction is mainly supported by end-effector force and joint torque, whereas wind-disturbance boundary cases increase the contribution of attitude and rotor response. Figure 7(b) shows regulation-priority groups, separating high-risk low-confidence windows from high-risk high-confidence overshoot cases. Explainable predictive maintenance research [25] indicates that evidence-level output improves engineering inspection because the reviewer can examine the dominant source of a warning rather than only a final label.

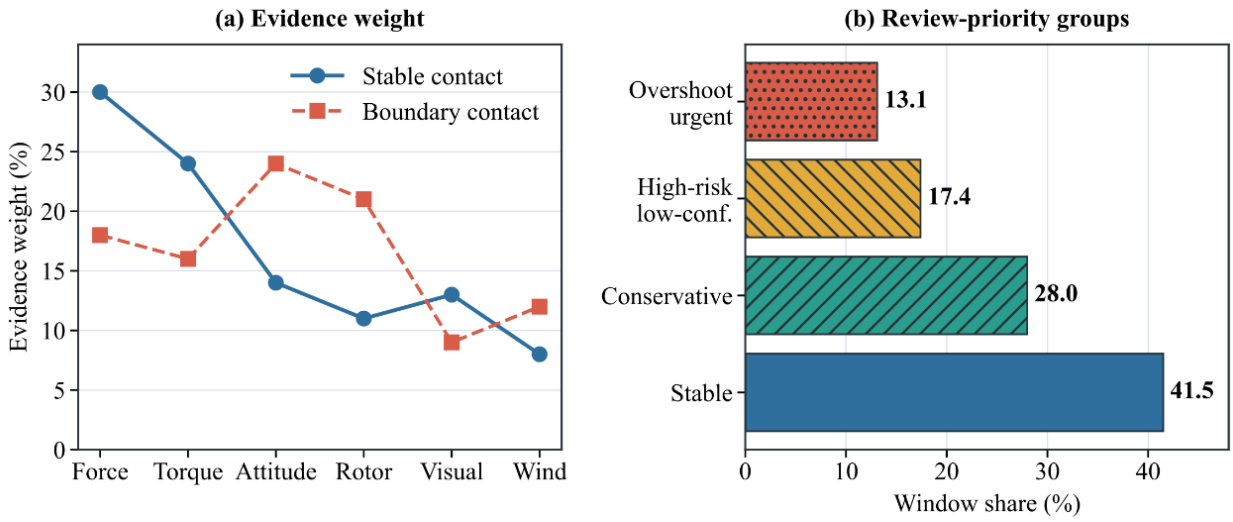


Figure 7. Evidence Weight and Control Review Output. (a) Evidence weight. (b) Review-priority groups.

The explanation result is used for regulation review rather than as a direct proof of a single physical cause. If contact force is high and confidence is high, the supervisory layer may reduce target force or move away along the contact normal. If attitude response dominates the warning, platform stabilization should be prioritized before further contact. If sensing quality is weak, remeasurement or slower approach is more appropriate than immediate retreat. Graph-based industrial robot anomaly detection [26] and multimodal fusion research [27] support the use of evidence quality together with predicted state when operational decisions require traceability.

Therefore, the flight-control use case is staged. The prediction model does not replace the low-level controller; instead, it provides contact-state forecast, force trend, confidence and review priority for the supervisory regulation layer. Stable contact maintains the desired force, boundary contact supports force-target reduction or position adjustment, force overshoot supports gain reduction or retreat, and loss of contact supports re-approach after attitude stabilization. Research on robot learning deployment [28] and edge intelligence for maintenance [29] also emphasizes that prediction modules near controllers should be lightweight, auditable and connected to action logs.

Several limitations remain. Safe robot deployment research [30] indicates that prediction, planning and control should be evaluated together before large-scale use, and industrial graph-learning research [31] further suggests that relation weights may need adjustment across platforms. The present dataset is synthetic and the disturbances are controlled, so future validation should include outdoor wind tests, flexible surfaces, different textures, manipulator lengths, end-effectors, cable effects and actuator heating. Digital twin and edge-computing research [32] also supports updating deployed models with field records while preserving traceability. Future work will compare predicted review priority with operator correction and controller intervention logs.

5. Conclusion

This study proposed a physics-guided Autoformer-GAT prediction model for contact force regulation of aerial manipulators, and the overall conclusion is that temporal force evolution, graph-based physical coupling and lightweight physical constraints should be evaluated jointly for reliable supervisory regulation. First, the simulated dataset of 3,400 contact windows shows that the proposed model achieves 92.1% contact-state accuracy, 87.9% macro-F1 and a 15.4% reduction in force-prediction error relative to the best previous baseline. Second, the comparative results show that Autoformer decomposition improves separation of slow contact trends and short residual impulses, while graph attention improves interpretation of coupling among force, attitude, joint, rotor, visual and wind-related evidence. Third, the main limitation is that the evaluation still relies on simulated contact windows and controlled disturbances; therefore, outdoor wind tests, flexible surfaces, different end-effectors, manipulator lengths and controller action logs should be incorporated into future validation.

The proposed output separates predicted force, contact state, confidence and regulation priority. This separation supports target-force reduction for high-confidence overshoot, controlled adjustment for boundary contact, verification for low-confidence risky windows and re-approach after confirmed loss of contact. The confidence score is especially useful because it decreases under wind disturbance, sensor noise and force-sensor dropout before the predicted contact state becomes unstable. The model therefore supports contact inspection, manipulation and maintenance while preserving evidence traceability for later engineering review.

Future work should connect the prediction output with low-level controller actions, maintenance records and operator correction logs. Additional field experiments should include outdoor wind conditions, flexible or textured

surfaces, several end-effector designs, different manipulator lengths, cable effects and thermal variation in actuators. These extensions will support site-specific threshold calibration and help determine whether the evidence structure remains reliable during extended-duration aerial contact tasks.

6. Declarations

6.1. Author Contributions

Kristian Sokol contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, project administration, and funding acquisition. Alen Pavić contributes to analysis, investigation, data collection, draft preparation. All authors have read and agreed with the manuscript before its submission and publication.

6.2. Funding and Acknowledgments

This research received no specific financial support from any funding agency.

6.3. Institutional Review Board Statement

Not applicable.

6.4. Informed Consent Statement

Not applicable.

6.5. Data Availability

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

6.7. Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors utilized Grammarly for linguistic polishing, logical organization and draft revision. The generative AI tool was not involved in research design, data collection, data analysis, or the equation of core conclusions. All outputs generated by AI were reviewed, revised and verified manually by the authors. The authors take full responsibility for all content of this manuscript.

7. References

[1] Liu, Q., Liu, Y., Chen, Z., Guo, K., Yu, X., Zhang, Y., & Zhao, S. (2024). A compact aerial manipulator: Design and control for dexterous operations. Journal of Intelligent & Robotic Systems, 110(2), 66. https://doi.org/10.1007/s10846-024-02090-7

[2] Meng, X., He, Y., Han, J., & Song, A. (2025). Physical interaction oriented aerial manipulators: Contact force control and implementation. IEEE Transactions on Automation Science and Engineering, 22, 4570-4582. https://doi.org/10.1109/TASE.2024.3354996

[3] Lyu, S., Zhang, Y., Wang, J., Cheah, C. C., & Yu, X. (2025). Toward air operation aerial manipulator control with a refined anti-disturbance architecture. IEEE Transactions on Automation Science and Engineering, 22, 4076-4091. https://doi.org/10.1109/TASE.2024.3407191

[4] Garrard, Y., Stoumbos, T., Inoyama, D., & Zhang, W. (2024). Nonlinear wrench observer of an underactuated aerial manipulator. AIAA SCITECH 2024 Forum, 0660. https://doi.org/10.2514/6.2024-0660

[5] Liu, Q., Lyu, S., Guo, K., Wang, J., Yu, X., Guo, L., & Zhang, Y. (2024). A coordinated framework of aerial manipulator for safe and compliant physical interaction. Control Engineering Practice, 146, 105898. https://doi.org/10.1016/j.conengprac.2024.105898

[6] Chen, Y., Dai, Z., Lin, Y., Zhang, L., Lu, H., & Zhang, H. (2026). Contact force tracking control for aerial manipulators in unknown dynamic environments. IEEE Transactions on Industrial Informatics, 22(4), 2997-3007. https://doi.org/10.1109/TII.2025.3643589

[7] Fang, Q., & Mao, P. (2024). Compliant contact force control for aerial manipulator of adaptive neural network-based robust control. Sensors, 24(8), 2556. https://doi.org/10.3390/s24082556

[8] Wu, Y., Chen, Y., Zeng, G., Shang, C., Miao, Z., Wang, H., Wang, Y., & Shen, Q. (2024). Dynamic image-based impedance control of aerial manipulator for sliding inspection. IEEE/ASME Transactions on Mechatronics, 29(6), 4075-4086. https://doi.org/10.1109/TMECH.2024.3362197

[9] Khalifa, A., Fanni, M., & Khalifa, A. (2024). Sensorless contact force estimation and robust impedance control for a quadrotor manipulation system. Scientific Reports, 14(1), 28772. https://doi.org/10.1038/s41598-024-79606-6

[10] Chaikalis, D., & Tzes, A. (2024). Optimization-based compliant controller for physical human-aerial manipulator interaction. 2024 International Conference on Unmanned Aircraft Systems, 1325-1331. https://doi.org/10.1109/ICUAS60882.2024.10556973

[11] Zeng, J., Zhong, H., Liang, J., Zhu, Q., & Wang, Y. (2026). Robust image-based control for aerial interaction under external disturbances. ISA Transactions, 173, 564-576. https://doi.org/10.1016/j.isatra.2026.04.009

[12] Yu, Y., Ma, R., & Ma, Z. (2024). Robformer: A robust decomposition transformer for long-term time series forecasting. Pattern Recognition, 153, 110552. https://doi.org/10.1016/j.patcog.2024.110552

[13] Su, G., Wang, H., Zhang, Y., Zhang, W., & Lin, X. (2024). Simple and deep graph attention networks. Knowledge-Based Systems, 293, 111649. https://doi.org/10.1016/j.knosys.2024.111649

[14] Chen, W., Yang, K., Yu, Z., Shi, Y., & Chen, C. L. P. (2024). A survey on imbalanced learning: Latest research, applications and future directions. Artificial Intelligence Review, 57(6), 137. https://doi.org/10.1007/s10462-024-10759-6

[15] Khalifa, A., Shaaban, S. M., & Khalifa, A. (2025). Hybrid disturbance observer and fuzzy logic controller for a new aerial manipulation system. Frontiers in Robotics and AI, 12, 1528415. https://doi.org/10.3389/frobt.2025.1528415

[16] Gaudreault, J. G., & Branco, P. (2024). Empirical analysis of performance assessment for imbalanced classification. Machine Learning, 113(8), 5533-5575. https://doi.org/10.1007/s10994-023-06497-5

[17] Tekden, A. E., Erdem, A., Erdem, E., Asfour, T., & Ugur, E. (2024). Object and relation centric representations for push effect prediction. Robotics and Autonomous Systems, 174, 104632. https://doi.org/10.1016/j.robot.2024.104632

[18] Vrahatis, A. G., Lazaros, K., & Kotsiantis, S. (2024). Graph attention networks: A comprehensive review of methods and applications. Future Internet, 16(9), 318. https://doi.org/10.3390/fi16090318

[19] Zhang, X., Wang, H., Wang, C., Liu, M., & Xu, G. (2024). Time-segment-wise feature fusion transformer for multi-modal fault diagnosis. Engineering Applications of Artificial Intelligence, 138, 109358. https://doi.org/10.1016/j.engappai.2024.109358

[20] Li, H., Li, Z., Liu, J., Zheng, X., Yu, X., & Kaynak, O. (2024). Adaptive neural network backstepping control method for aerial manipulator based on coupling disturbance compensation. Journal of the Franklin Institute, 361(7), 106733. https://doi.org/10.1016/j.jfranklin.2024.106733

[21] Zou, K., Lin, T., Han, Z., Wang, M., Yuan, X., Chen, H., Zhang, C., Shen, X., & Fu, H. (2024). Confidence-aware multi-modality learning for eye disease screening. Medical Image Analysis, 96, 103214. https://doi.org/10.1016/j.media.2024.103214

[22] Gao, Z., Jiang, X., Xu, X., Shen, F., Li, Y., & Shen, H. T. (2024). Embracing unimodal aleatoric uncertainty for robust multimodal fusion. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 26866-26875. https://doi.org/10.1109/CVPR52733.2024.02538

[23] Ferreira, R. S., Guérin, J., Delmas, K., Guiochet, J., & Waeselynck, H. (2025). Safety monitoring of machine learning perception functions: A survey. Computational Intelligence, 41(2), e70032. https://doi.org/10.1111/coin.70032

[24] Dai, R., Li, C., Yan, Y., Mo, L., Qin, K., & He, T. (2025). Unbiased missing-modality multimodal learning. Proceedings of the IEEE/CVF International Conference on Computer Vision, 24507-24517. https://doi.org/10.1109/ICCV51701.2025.02272

[25] Cummins, L., Sommers, A., Ramezani, S. B., Mittal, S., Jabour, J., Seale, M., & Rahimi, S. (2024). Explainable predictive maintenance: A survey of current methods, challenges and opportunities. IEEE Access, 12, 57574-57602. https://doi.org/10.1109/ACCESS.2024.3391130

[26] Han, J., Chen, Z., Zhou, D., Hu, B., Xia, T., & Pan, E. (2025). Unsupervised motion-based anomaly detection with graph attention networks for industrial robots labeling. Engineering Applications of Artificial Intelligence, 146, 110298. https://doi.org/10.1016/j.engappai.2025.110298

[27] Zhao, F., Zhang, C., & Geng, B. (2024). Deep multimodal data fusion. ACM Computing Surveys, 56(9), 1-36. https://doi.org/10.1145/3649447

[28] Xiao, X., Liu, J., Wang, Z., Zhou, Y., Qi, Y., Jiang, S., He, B., & Cheng, Q. (2025). Robot learning in the era of foundation models: A survey. Neurocomputing, 621, 129963. https://doi.org/10.1016/j.neucom.2025.129963

[29] Bala, A., Rashid, R. Z. J. A., Ismail, I., Oliva, D., Muhammad, N., Sait, S. M., Al-Utaibi, K. A., Amosa, T. I., & Memon, K. A.

(2024). Artificial intelligence and edge computing for machine maintenance: Review. Artificial Intelligence Review, 57(5), 119. https://doi.org/10.1007/s10462-024-10748-9

[30] Farrell, S., Li, C., Yu, H., Yoshimitsu, R., Gao, S., & Christensen, H. I. (2025). Safe human robot navigation in warehouse scenario. Proceedings of the 2025 IEEE 21st International Conference on Automation Science and Engineering, 548-555. https://doi.org/10.1109/CASE58245.2025.11164151

[31] Lu, H., Wang, L., Ma, X., Cheng, J., & Zhou, M. (2025). A survey of graph neural networks and their industrial applications. Neurocomputing, 614, 128761. https://doi.org/10.1016/j.neucom.2024.128761

[32] Kondo, R. E., Andrade, W. J., de Mello Henequim, C., Lazzaretti, A. E., de Souza Britto, A., de Freitas Rocha Loures, E., Santos, E. A. P., & Reynoso-Meza, G. (2024). An industrial edge computing architecture for local digital twin. Computers & Industrial Engineering, 193, 110257. https://doi.org/10.1016/j.cie.2024.110257