

Article

# Deep Learning-Assisted ES-EKF for Surface AUV Navigation with SINS/GPS/DVL Integration

Yuanbo Yang <sup>1</sup>, Bo Xu <sup>1,\*</sup>, Baodong Ye <sup>1</sup> and Feimo Li <sup>2</sup>

<sup>1</sup> College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin 150001, China; yangyuanbo2024@hrbeu.edu.cn (Y.Y.); yebd@hrbeu.edu.cn (B.Y.)

<sup>2</sup> Institute of Automation, Chinese Academy of Sciences Beijing, Beijing 100190, China; lifeimo2012@ia.ac.cn

\* Correspondence: xubocarter@hrbeu.edu.cn

## Abstract

This study presents a deep learning–assisted integrated navigation scheme implemented on an autonomous underwater vehicle carrying a Chinese domestically developed strapdown inertial navigation system, designed for operation in surface and littoral environments. The system integrates measurements from SINS, the global positioning system, and a Doppler velocity log, while integrating a Decoder-based covariance estimator into the error state-extended Kalman filter. This hybrid architecture adaptively models time-varying processes and measurement noise from raw sensor inputs, greatly improving robustness for surface navigation in dynamic marine environments. To improve learning efficiency, we design a compact and informative feature representation that can be adapted to navigation error dynamics. The novel structure captures temporal dependencies and the evolution of nonlinear error more effectively than typical sequence models, achieving faster convergence and superior accuracy compared to GRU and Transformer baselines. The experimental results based on real sea trial data show that our method significantly outperforms model-based and learning-based methods in terms of navigation solution accuracy and stability, and the adaptive estimation of noise covariance. Specifically, it achieves the lowest RMSE of 0.0274, reducing errors by 94.6–34.6%, compared to conventional ES-EKF-integrated navigation, Transformer, GRU, and a DCE variant. These findings underscore the practical significance of integrating domain-informed filtering methodologies with deep noise modeling frameworks to achieve robust and accurate AUV surface navigation.



Academic Editor: Weicheng Cui

Received: 7 October 2025

Revised: 21 October 2025

Accepted: 22 October 2025

Published: 23 October 2025

**Citation:** Yang, Y.; Xu, B.; Ye, B.; Li, F.

Deep Learning-Assisted ES-EKF for

Surface AUV Navigation with

SINS/GPS/DVL Integration. *J. Mar.*

*Sci. Eng.* **2025**, *13*, 2035. <https://doi.org/10.3390/jmse13112035>

**Copyright:** © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license

(<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** error state-extended Kalman filter; deep learning; autonomous underwater vehicle; integrated navigation

## 1. Introduction

With the continuous development and progress of the machinery, materials, and electronic information industries, underwater robot technology has also been developing rapidly. Compared with remotely operated vehicles (ROVs) that can be controlled in real time by humans, autonomous underwater vehicles (AUVs), as a new generation of unmanned submersibles, are free from cables and deck units. They are characterized by high maneuverability, strong autonomy, and a wide operational range, making them essential tools for submarine cable inspection, marine environmental monitoring, and subsea pipeline inspection [1–5].

AUVs are typically equipped with a suite of pressure-resistant sensors and navigation devices, including strapdown inertial navigation system (SINSs), a magnetometer, a

Doppler velocity log (DVL), a global positioning system (GPS), a depth gauge, an altimeter, and underwater cameras, among others. These navigation and positioning systems each have inherent limitations in terms of accuracy, reliability, and operational robustness. In highly dynamic and complex marine environments, relying on any single sensor modality often fails to meet the stringent navigation requirements of AUVs [6]. Therefore, integrating multiple navigation sensors into a cooperative system to enhance overall performance has become a mainstream approach in modern navigation technology [7–9].

Considering the typical operational workflow of AUVs, which includes deployment and descent, autonomous underwater operation, and recovery [10,11], there exists a clear demand for navigation solutions across both surface and submerged environments. This study focuses on improving the accuracy and robustness of a loosely coupled integrated navigation scheme specifically designed for surface navigation. In this scheme, GPS is utilized to correct the long-term drift in the SINS, thereby mitigating accumulated errors. The DVL improves the short-term accuracy and stability of the system by providing relative velocity measurements. When GPS or DVL signals are unavailable, SINS can still independently provide continuous position, velocity and attitude solutions. Moreover, by employing an optimal estimation approach based on the Kalman filter, the multi-source data from these sensors can be effectively fused to produce reliable and robust navigation outputs [12,13].

The error state-extended Kalman filter (ES-EKF) is the core algorithm typically employed in integrated navigation schemes [14,15]. It fuses heterogeneous sensor measurements to compute integrated navigation solutions. Unlike conventional Kalman filters, the ES-EKF estimates inertial measurement unit (IMU) errors as part of the system state, rather than estimating the full navigation states directly. Since these error states remain small and near zero, the ES-EKF is less susceptible to parameter singularities and global divergence issues [16]. However, the effectiveness of the ES-EKF heavily depends on the manual design of the IMU error model, which requires significant domain expertise and prior knowledge of sensor characteristics [17,18]. In this context, the main source of error in AUV navigation arises from static errors in the IMU, whereas the bias error in the static error of the IMU is usually highly coupled with attitude error and velocity error. Improper modeling can make it difficult to converge the error state estimation of the IMU, thereby affecting the overall accuracy and reliability of the entire navigation system.

In recent years, research on the application of deep learning (DL) in navigation and positioning has advanced rapidly, with particular focus on addressing sensor outages, compensating for dynamic errors, and enhancing positioning performance [19–22]. In research on using DVL or a global navigation satellite system (GNSS) to assist in SINS navigation, studies such as [23–25] have demonstrated how the nonlinear relationship between the IMU and GNSS during periods of normal GPS availability is learnt. When GPS signals are lost, these methods generate pseudo-GPS signals to maintain navigation accuracy. Refs. [26,27] estimate the current vehicle's velocity under complete DVL outages by leveraging inertial data and historical DVL velocity measurements. Refs. [20,21] enhance inertial navigation by using deep learning to model and correct IMU errors without GNSS signals. They integrate neural networks with traditional navigation to improve position and velocity accuracy. One uses an Echo State Network for error modeling; the other applies CNN-BiLSTM to map IMU data to GNSS references. Both show that data-driven correction effectively reduces inertial navigation errors.

These advances in DL indicate that they can directly capture the nonlinearity and dependencies of complex systems from the data, thereby reducing the need for explicit system modeling and feature engineering. This has led to a feasible research trend, namely a hybrid approach that integrates DL with real-time state estimation techniques.

Recent research in real-time state estimation has increasingly adopted hybrid approaches that combine physical modeling with DL [28]. These methods typically concentrate on two main challenges: (1) dynamically modeling time-varying noise, and (2) refining the Kalman gain update mechanism through learning. Methods such as NNAKF [29] and EKFNet [30] focus on learning process and measurement noise covariances directly from data. NNAKF leverages RNNs to dynamically adjust process noise, improving adaptability and tracking accuracy in non-stationary environments. Similarly, EKFNet employs offline training and custom loss functions to refine noise statistics. While effective, these approaches generally do not explore model architecture choices in depth, nor do they systematically investigate feature selection strategies adapted to the complex spatiotemporal characteristics of inertial and navigation data. Ref. [31] explores the accuracy and feasibility of several time series models in estimating the errors of inertial measurement units in the context of two-dimensional simulation data. The other line of research focuses on replacing or enhancing the computation of the Kalman gain using deep learning techniques [32,33]. Among them, KalmanNet proposed in [34], is based on the KF paradigm and innovatively replaces the traditional Kalman gain calculation with an RNN-based encapsulation mechanism. Although it enhances the modeling ability of the temporal characteristics of noise, it also deviates from the theoretical basis of the optimal estimation of the minimum mean square error (MMSE) and lacks a rigorous characterization of the posterior covariance. Nowadays, many research projects involve addressing both challenges simultaneously. However, replacing the Kalman gain with purely data-driven methods may alter the standard ES-EKF framework, which relies on linearized error state propagation and covariance updates. Such a substitution can affect the stability, consistency, and interpretability of the resulting estimator, even though strict optimality guaranteed by the classical Kalman filter does not apply to nonlinear systems like the ES-EKF.

In the study of surface-integrated navigation for AUVs, both the measurement and motion models are generally well defined and can be constructed through physical modeling, requiring minimal reliance on data-driven learning. While both model-based and learning-based algorithms are capable of correcting navigation errors, their performance is often sensitive to the selection of initial parameters and convergence criteria. Furthermore, the error accumulated in inertial navigation is inherently unbounded, meaning that the positioning error from SINS can grow indefinitely over time. This makes dynamic noise modeling within the Kalman filter particularly critical for maintaining long-term navigation accuracy and system robustness.

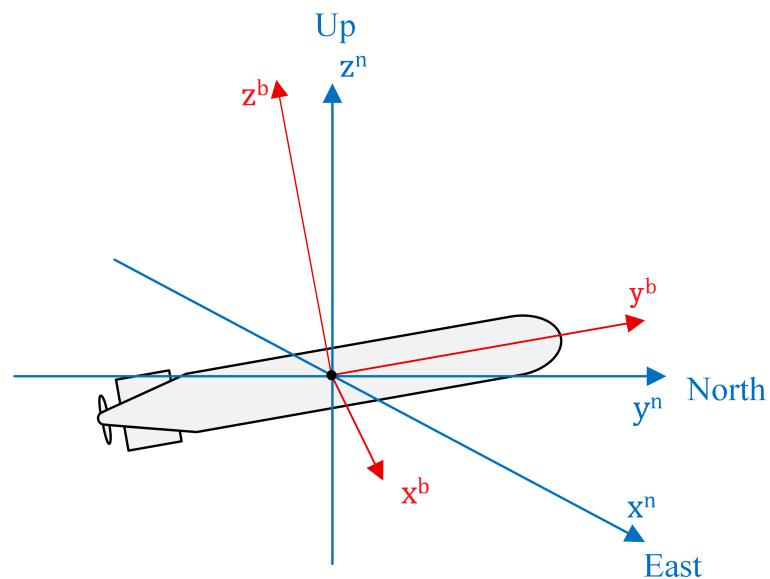
To address this issue, we propose a novel hybrid navigation scheme that integrates a learning-based covariance estimator into the ES-EKF. Our method adaptively optimizes both the process and measurement noise covariance matrices in real time, while preserving the conventional ES-EKF architecture. It can better capture long-distance dependencies and richer cross-temporal difference information. In addition, we design a compact and efficient input feature representation that captures key dynamic patterns while reducing computational overhead. This not only improves the real-time applicability of the model but also accelerates network convergence. Extensive experiments on real-world marine datasets demonstrate that our proposed method outperforms the conventional ES-EKF and typical learning-based models such as a gated recurrent unit (GRU) [35] and Transformer [36], achieving higher accuracy, better stability, and faster convergence in AUV surface navigation scenarios.

The rest of this article is organized as follows: Section 2 introduces the SINS/GPS/DVL integration system model formulation, Section 3 presents the deep learning-assisted error state Kalman filter structure, Section 4 presents the performance analysis of the proposed

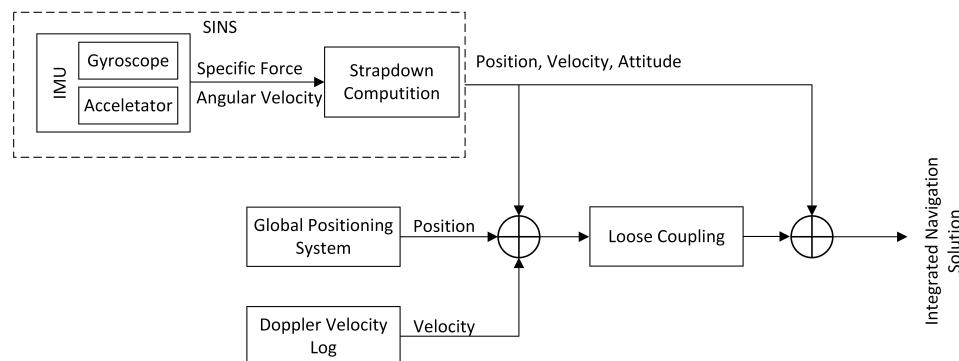
method, and finally Section 5 concludes this article. The code of this work can be found at <https://github.com/Yybcbjy/DL-assisted-ESEKF.git> (accessed on 6 October 2025).

## 2. System Model Formulation

The position and velocity in this study are expressed in the local navigation frame  $n$  (East-North-Up), while the measurements from the IMU are provided in the body frame  $b$ . The body frame is fixed to our AUV, with the  $x$ -axis aligned with the forward direction of the vehicle, the  $y$ -axis pointing to the right side of the vehicle, and the  $z$ -axis pointing directly upward from the AUV. Figure 1 shows the reference frame mentioned above. Depending on the specific application requirements, SINS, GPS, and DVL can be integrated using loosely coupled, tightly coupled, or deeply coupled configurations [37]. Loosely coupled navigation fuses GPS-derived positions, DVL-estimated velocities, and IMU measurements to provide accurate navigation solutions. The INS inherently accumulates drift errors over time, which are periodically corrected by an ES-EKF using GPS and DVL updates. This approach is favored for its simplicity and versatility, as it can accommodate a wide range of INS and GNSS devices, making it ideal for retrofit and enhancement applications. Besides the fused navigation output, the system also provides standalone GNSS and independent INS solutions, enabling effective integrity monitoring during operation. The loosely coupled SINS/GPS/DVL-integrated navigation scheme employed in this study is shown in Figure 2.



**Figure 1.** The navigation frame and body frame of AUV.



**Figure 2.** Loosely coupled integrated navigation scheme based on SINS/GPS/DVL.

The loosely coupled system operates recursively in two main steps: first, the inertial solution continuously integrates accelerometer and gyroscope measurements to estimate the nominal state  $x$ , which comprises the position, velocity, attitude, and bias. Second, when GPS and DVL data are available, the filter algorithm estimates and corrects the error state  $\delta x$  of the inertial solution. The error-state is defined as the small perturbation between the true state and the nominal state:

$$\delta x = x_{true} \ominus x, \quad (1)$$

where  $\ominus$  denotes the appropriate composition operator: for translational components, it reduces to subtraction, while attitude errors are expressed in a small-angle vector form. Concretely, the nominal state  $x$  and the error state  $\delta x$  are defined as follows:

$$\begin{aligned} x &= [\theta_{yaw}, \theta_{pitch}, \theta_{roll}, v_x^n, v_y^n, v_z^n, p_x^n, p_y^n, p_z^n, \zeta_x^b, \zeta_y^b, \zeta_z^b, \\ &\quad \epsilon_x^b, \epsilon_y^b, \epsilon_z^b]^T \in \mathbb{R}^{15 \times 1}, \\ \delta x &= [\delta\theta_{yaw}, \delta\theta_{pitch}, \delta\theta_{roll}, \delta v_x^n, \delta v_y^n, \delta v_z^n, \delta p_x^n, \delta p_y^n, \delta p_z^n, \\ &\quad \delta \zeta_x^b, \delta \zeta_y^b, \delta \zeta_z^b, \delta \epsilon_x^b, \delta \epsilon_y^b, \delta \epsilon_z^b]^T \in \mathbb{R}^{15 \times 1}, \end{aligned} \quad (2)$$

where  $\delta p$ ,  $\delta v$ , and  $\delta\theta$  are the estimated position, velocity, and attitude errors, respectively;  $\delta\epsilon$  and  $\delta\zeta$  are the estimated accelerometer and gyroscope bias errors, respectively. Although our AUV mainly operates at the sea surface, the full navigation state is maintained for consistency with the inertial frame definition. The DVL velocity  $mv$  and GPS position  $mp$  are used as measurements  $y$ :

$$y = [mv_x^n, mv_y^n, mv_z^n, mp_x^n, mp_y^n, mp_z^n]^T \in \mathbb{R}^{6 \times 1}. \quad (3)$$

Their superscripts and subscripts denote the corresponding reference frame and axis.

## 2.1. IMU Measurement Errors Modeling

The IMU measurement noise comprises the combined effects of gyroscope and accelerometer noise, as given below:

$$\tilde{w} = w + \epsilon + S_w + N_w + n_w, \quad (4)$$

$$\tilde{f} = f + \zeta + S_f + N_f + n_f + \delta g, \quad (5)$$

where  $\hat{w}$  and  $w$  denote the true and measured angular velocities obtained from the gyroscope, respectively.  $\hat{f}$  and  $f$  represent the measured and true specific forces of the accelerometer.  $S_{(.)}$  and  $N_{(.)}$  are the scale factor error matrix and the cross-axis coupling error matrix, while  $n_{(.)}$  denotes the sensor noise vector, and  $\delta g$  represents the gravity anomaly measured by the accelerometer. Such decomposition of IMU measurement errors is commonly adopted in navigation studies [38].

## 2.2. Error State Kalman Filter

Before conducting the experiments, we calibrate the inertial sensors to compensate for systematic errors. All sensors are time-synchronized and geometrically aligned, and lever arm compensation is applied to maintain consistency in spatial reference. The underlying nonlinear system of our method can be expressed as follows:

$$\begin{aligned} \dot{x}_t &= f(x_t, u_t) + w_t, \\ y_t &= h(x_t) + v_t. \end{aligned} \quad (6)$$

where  $u_t$  represents IMU inputs,  $w_t$  and  $v_t$  denote process and measurement noise, respectively. The nonlinear dynamics in (6) are linearized to obtain the discrete-time error state model:

$$\begin{aligned}\delta x_{t+1} &= F_t \delta x_t + G_t w_t, \\ y_t &= H_t \delta x_t + v_t,\end{aligned}\quad (7)$$

where  $G_t$  is the noise input matrix. The state transition matrix  $F_t$  and measurement matrix  $H_t$  represent the linearized system and measurement models, respectively. The system propagation is performed by applying the state transition matrix  $F_t$ , which is formulated by linearizing the nonlinear error propagation equations, as discussed in previous studies [39,40]:

$$F_t = \begin{bmatrix} M_1 & M_2 & M_3 & -C_b^n & 0_{3 \times 3} \\ M_4 & M_5 & M_6 & 0_{3 \times 3} & C_b^n \\ 0_{3 \times 3} & M_7 & M_8 & 0_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} \end{bmatrix}, \quad (8)$$

$$M_1 = \begin{bmatrix} 0 & \frac{v_E^n \tan \varphi}{R_N + h} + \omega_{ie} \sin \varphi & -\frac{v_E^n}{R_N + h} - \omega_{ie} \cos \varphi \\ \frac{v_E^n \tan \varphi}{R_N + h} - \omega_{ie} \sin \varphi & 0 & -\frac{v_N^n}{R_M + h} \\ \frac{v_E^n \tan \varphi}{R_N + h} + \omega_{ie} \cos \varphi & \frac{v_N^n \tan \varphi}{R_M + h} & 0 \end{bmatrix}, \quad (9)$$

$$M_2 = \begin{bmatrix} 0 & \frac{1}{R_M + h} & 0 \\ \frac{1}{R_N + h} & 0 & 0 \\ \frac{\tan \varphi}{R_N + h} & 0 & 0 \end{bmatrix}, \quad (10)$$

$$M_3 = \begin{bmatrix} 0 & 0 & \frac{v_N^n}{(R_M + h)^2} \\ -\omega_{ie} \sin \varphi & 0 & -\frac{v_E^n}{(R_N + h)^2} \\ \omega_{ie} \cos \varphi + \frac{v_E^n}{(R_N + h) \cos^2 \varphi} & 0 & -\frac{v_E^n \tan \varphi}{(R_N + h)^2} \end{bmatrix}, \quad (11)$$

$$M_4 = \begin{bmatrix} 0 & -f_U^n & f_N^n \\ f_U^n & 0 & -f_E^n \\ -f_U^n & f_E^n & 0 \end{bmatrix}, \quad (12)$$

$$M_5 = \begin{bmatrix} \frac{v_N^n \tan \varphi - v_U^n}{R_N + h} & \frac{v_E^n \tan \varphi}{R_N + h} + 2\omega_{ie} \sin \varphi & -\frac{v_E^n}{R_N + h} - 2\omega_{ie} \cos \varphi \\ -\frac{v_E^n \tan \varphi}{R_N + h} - 2\omega_{ie} \sin \varphi & -\frac{v_U^n}{R_M + h} & -\frac{v_N^n}{R_M + h} \\ \frac{2v_E^n}{R_N + h} + 2\omega_{ie} \cos \varphi & \frac{v_N^n}{R_M + h} & 0 \end{bmatrix}, \quad (13)$$

$$M_6 = \begin{bmatrix} \frac{v_E^n v_N^n}{(R_N + h) \cos^2 \varphi} + 2\omega_{ie} (v_U^n \sin \varphi + v_N^n \cos \varphi) & 0 & \frac{v_E^n v_U^n - v_E^n v_N^n \tan \varphi}{(R_N + h)^2} \\ -2\omega_{ie} \cos \varphi - \frac{(v_E^n)^2}{(R_N + h) \cos^2 \varphi} & 0 & \frac{v_N^n v_U^n}{(R_M + h)^2} + \frac{(v_E^n)^2 \tan \varphi}{(R_N + h)^2} \\ \frac{2v_E^n}{R_N + h} + 2\omega_{ie} v_E^n \sin \varphi & 0 & -\frac{(v_N^n)^2}{(R_M + h)^2} + \frac{(v_E^n)^2}{(R_N + h)^2} \end{bmatrix}, \quad (14)$$

$$M_7 = \begin{bmatrix} 0 & \frac{1}{(R_M+h)^2} & 0 \\ \frac{1}{(R_N+h)\cos\varphi} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad (15)$$

$$M_8 = \begin{bmatrix} 0 & 0 & -\frac{v_N^n}{(R_M+h)^2} \\ \frac{v_E^n \tan\varphi}{(R_N+h)\cos\varphi} & 0 & -\frac{v_E^n}{(R_N+h)^2 \cos\varphi} \\ 0 & 0 & 0 \end{bmatrix}, \quad (16)$$

where  $R_M$  and  $R_N$  denote the meridian radius of curvature and the prime vertical radius of curvature, respectively.  $\varphi$  denotes the geodetic latitude and  $h$  is the altitude of the AUV.  $C_b^n$  denotes the rotation matrix from the body frame to the navigation frame used for coordinate transformation,  $0_{3 \times 3}$  represents a  $3 \times 3$  zero matrix commonly used as a zero block in state-space matrices, and  $\omega_{ie}$  denotes the Earth's rotation rate vector. The state transition matrix  $F_t$  represents the linearized error-state dynamics around the nominal state  $x$ . Certain values in  $F_t$  are evaluated using the nominal state integrated from IMU measurements. This  $F_t$  allows the filter to propagate the error covariance without requiring additional derivatives or online linearization. For discrete implementation with sampling interval  $\Delta t$ , we use the first-order approximation,

$$\Phi_t \approx I + F_t \Delta t, \quad (17)$$

to propagate the error covariance  $P$ :

$$P_{t|t-\Delta t} = \Phi_t P_{t-\Delta t|t-\Delta t} \Phi_t^T + Q \quad (18)$$

where  $Q$  is the discrete process noise covariance matrix.  $I$  denotes the identity matrix. The measurement matrix can be formulated as follows:

$$H = \begin{bmatrix} 0_{3 \times 3} & I_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & I_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} \end{bmatrix}. \quad (19)$$

The measurement matrix  $H$  remains constant because the GPS and DVL measurement models are already linear in the error-state formulation. Thus, although the analytical structure of  $F_t$  and  $H$  is predetermined, the filter still operates within a linearized, time-varying framework consistent with the ES-EKF methodology. The predicted measurements  $y_{t|t-\Delta t}$  can be computed by applying the measurement matrix  $H$  to the nominal state  $x_t$ :

$$y_{t|t-\Delta t} = H x_{t|t-\Delta t}. \quad (20)$$

The error state and its covariance matrix are estimated during the measurement update phase by incorporating the latest noisy GPS and DVL measurements. The measurement noise covariance matrix  $R$  characterizes the statistical uncertainty of these sensor measurements and directly affects the computation of the Kalman gain  $K_t$ , which determines the relative weighting between the measurements and the current inertial navigation solutions.

$$K_t = P_{t|t-\Delta t} H^T \left[ H P_{t|t-\Delta t} H^T + R \right]^{-1} \quad (21)$$

where  $R$  denotes the measurement noise covariance matrix. Similar to  $Q$ , it is empirically designed based on the known measurement uncertainties. Specifically,  $\sigma_p$  represents the

standard deviation of the GPS position measurements, while  $\sigma_v$  denotes the uncertainty associated with DVL velocity measurements. The measurement noise covariance matrix is defined as follows:

$$R = \begin{bmatrix} \sigma_p^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}. \quad (22)$$

The posterior error state is computed by multiplying  $K_t$  with the innovation, which is defined as the difference between  $y_t$  and  $y_{t|t-\Delta t}$ , both expressed in the local navigation frame:

$$\delta x_t = K_t(y_t - y_{t|t-\Delta t}). \quad (23)$$

The posterior error state covariance matrix is updated according to the Joseph form:

$$P_t = (I - K_t H)P_{t|t-\Delta t}(I - K_t H)^T + K_t R K_t^T. \quad (24)$$

In this study,  $P_t$  represents the covariance of the error-state, not the full navigation state. Although this update form is mathematically identical to that of the classical Kalman filter, it is applied here within the ES-EKF framework to maintain the consistency of the linearized error dynamics.

Finally, the corrected true state can be calculated as follows:

$$\hat{x}_t = x_t \oplus \delta x_t \quad (25)$$

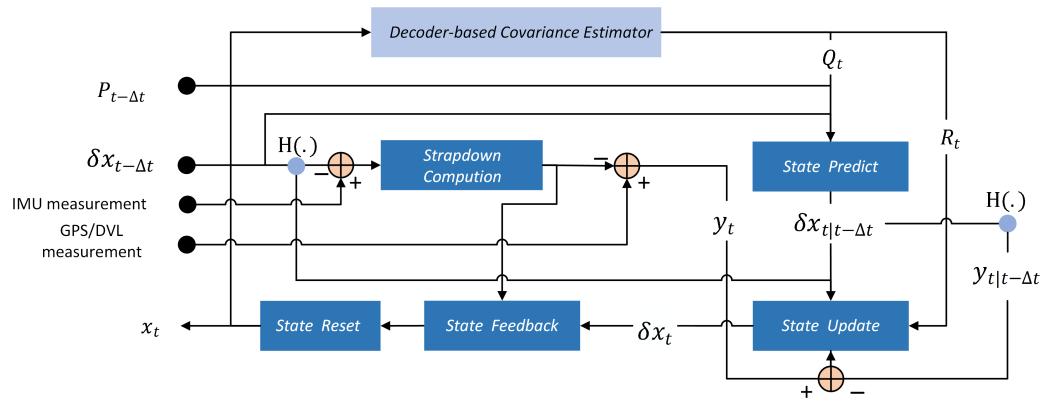
where  $\oplus$  denotes the state composition operator, which incorporates the estimated errors in the attitude, velocity, position, and sensor biases into the nominal state  $x_t$ . Attitude correction is performed using the exponential map on the special orthogonal group,  $SO(3)$ , which maps small error rotation vectors to valid rotation matrices. After correction, the updated nominal state  $\hat{x}_t$  serves as the prior for the next iteration, ensuring a consistent closed-loop estimation process.

### 3. Deep Learning-Assisted Error State Kalman Filter

The performance of ES-EKF largely depends on the degree of mastery of prior knowledge during the model construction process. Although the high-frequency sampling interval adopted in this work allows the linearized state transition matrix to approximate the error dynamics of the nonlinear system, inaccurate specification of the process noise covariance  $Q$  or the measurement noise covariance  $R$  can still cause significant degradation. Misconfiguration can induce numerical instabilities, such as the explosion of the error covariance matrix  $P$  and divergence of the state transition matrix  $e^{F\Delta t}$ , eventually causing filter divergence and estimation failure. Therefore, we propose a DL-assisted ES-EKF scheme that incorporates carefully designed filter-based features and a lightweight neural network architecture to dynamically adjust the noise covariance matrices for improved positioning accuracy and robustness.

#### 3.1. High-Level Architecture

Figure 3 illustrates the proposed architecture of the DL-assisted ES-EKF algorithm. The Decoder-based covariance estimator (DCE) is deeply integrated into the ES-EKF. During the system propagation phase, the error-state is used to extract features required for DCE construction before it is reset. After the ES-EKF completes one full iteration over the navigation dataset, the training phase of the DCE begins. Once trained, the DCE can be embedded into the ES-EKF to dynamically estimate the noise covariance during online navigation.



**Figure 3.** The proposed DL-assisted ES-EKF scheme for integrated navigation.

### 3.2. Input Features

To enable the DCE to effectively learn the implicit characteristics of the accelerometer and gyroscope biases in the IMU, informative input features must be carefully designed during the training stage. Building upon previous work [31,34], we refine and extend the feature extraction process as follows:

- At time step  $t$ , the difference between the estimated error-state and the predicted error-state is as follows:  $f_1 = \delta x_t - \delta x_{t|\bar{t}-\Delta t}$ .
- The difference between the estimated error-state at time step  $t$  and the predicted error-state at time step  $t + \Delta t$  is as follows:  $f_2 = \delta x_t - \delta x_{t+\Delta t|\bar{t}}$ .
- The difference between the measured measurement at time step  $t$  and  $t - \Delta t$  is as follows:  $f_3 = y_t - y_{t-\Delta t}$ .
- The innovation difference at time step  $t$  is as follows:  $f_4 = y_t - y_{t|\bar{t}-\Delta t}$ .
- The measurement difference of the IMU between time step  $t$  and  $t - \Delta t$  is as follows:  $f_5 = [\alpha_t, \omega_t] - [\alpha_{t-\Delta t}, \omega_{t-\Delta t}]$ .
- The evolution of the error state difference at the time step  $t$  is as follows:  $(f_1, f_2) = U \Sigma \underbrace{V^T}_{f_6}$ .
- The evolution of the measurement difference at time step  $t$  is as follows:  $(f_3, f_4) = U \Sigma \underbrace{V^T}_{f_7}$ .

The differences from  $f_1$  to  $f_5$  can eliminate the inherent characteristics of the state and measurement during the propagation process.  $f_6$  and  $f_7$  are decomposed via singular-value decomposition using the matrix  $V$ , capturing the temporal evolution patterns of the error states and measurements, respectively.

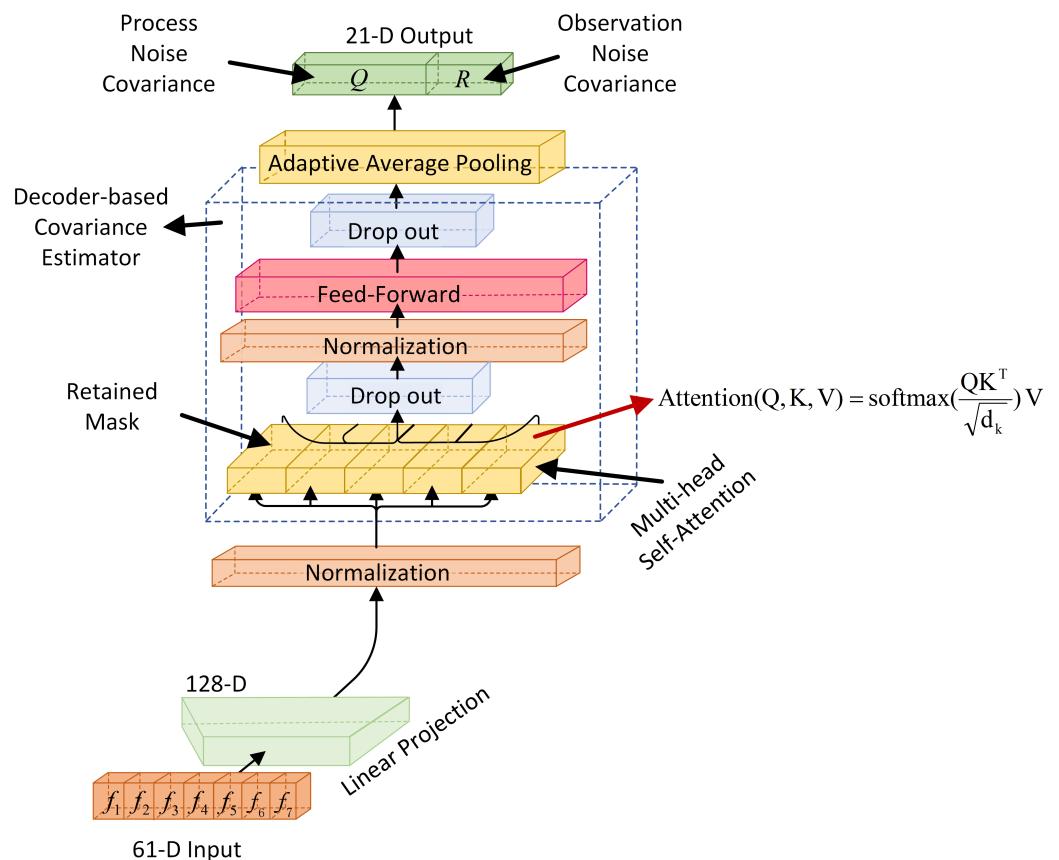
### 3.3. Decoder-Based Covariance Estimator Architecture

To design a neural network architecture suitable for surface-integrated navigation tasks of our AUV, we considered several state-of-the-art time series prediction models and conducted comparative experiments. The final architecture is illustrated in Figure 4.

During the forward propagation process of the DCE, the input sequence  $\delta x$  first passes through a linear projection layer, which maps each 15-dimensional input feature at each time step to a higher-dimensional representation. The representations of these matrix are as follows:

$$\begin{cases} Q = \delta x W_Q \\ K = \delta x W_K \\ V = \delta x W_V \end{cases} \quad (26)$$

where  $W_Q$ ,  $W_K$ , and  $W_V$  are trainable weight matrices. This representation then undergoes normalization through layer normalization and is simultaneously used as the query  $Q$ , key  $K$ , and value  $V$  in the multi-head self-attention mechanism. Owing to the sequential computation nature of the ES-EKF, we employ self-attention to compute the attention scores, which are subsequently passed through a residual connection and a dropout layer to mitigate overfitting. In the attention mechanism,  $d_k$  is the dimension of the key vectors. Next, the output undergoes a second round of layer normalization before being fed into a feed-forward network, which consists of two linear layers separated by a ReLU activation function, followed by another dropout layer. The result of the feed-forward network is added back to its input via another residual connection. Then, a transposition operation is carried out, and the projection is performed through an adaptive average pooling layer and a final linear layer to generate a 21-dimensional output vector. The first 15 dimensions correspond to the diagonal elements of the process noise covariance matrix, while the remaining 6 dimensions correspond to the elements of the measurement noise covariance matrix. To ensure that all values are strictly positive, both parts are passed through a softplus function, followed by the addition of a small positive constant.



**Figure 4.** Decoder-based covariance estimator architecture.

Furthermore, since the sequence length of all inputs is 1, the attention mechanism simplifies to performing scalar self-attention operations on a single vector without any interaction across time steps. The masking operation has no practical effect in this case; however, we chose to retain its form to ensure compatibility with potential future extensions.

### 3.4. Loss Function

The DCE conducts end-to-end training using supervised learning. We use the highly accurate position information obtained by the precision high-performance inertial navi-

gation system (PHINS) as the ground truth, and define the loss function  $\ell$  by comparing it with the position obtained through our DL-assisted method. The specific process is as follows:

$$\ell(t) = \left\| (p_{PHINS} - p)_{E,N} \right\|_2 \quad (27)$$

where  $\ell(t)$  denotes the value of the loss function at time step  $t$ , which measures the deviation between the estimated position  $p$ , based on the predicted  $Q$  and  $R$ , and the ground truth position  $p_{PHINS}$  in the east and north directions of the local horizontal frame. The DCE aims to optimize the model parameters through the gradient descent, achieving this by determining the set of parameters that can efficiently minimize  $\ell(\Theta_t)$ . The parameter update rule using gradient descent is given as follows:

$$\Theta_{(t+\Delta t)} = \Theta_{(t)} - \eta \frac{\partial \ell(\Theta)}{\partial \Theta} \quad (28)$$

where  $\Theta_t$  represents all the parameters of the DCE at time step  $t$ .

## 4. Experiments

### 4.1. Experimental Setting

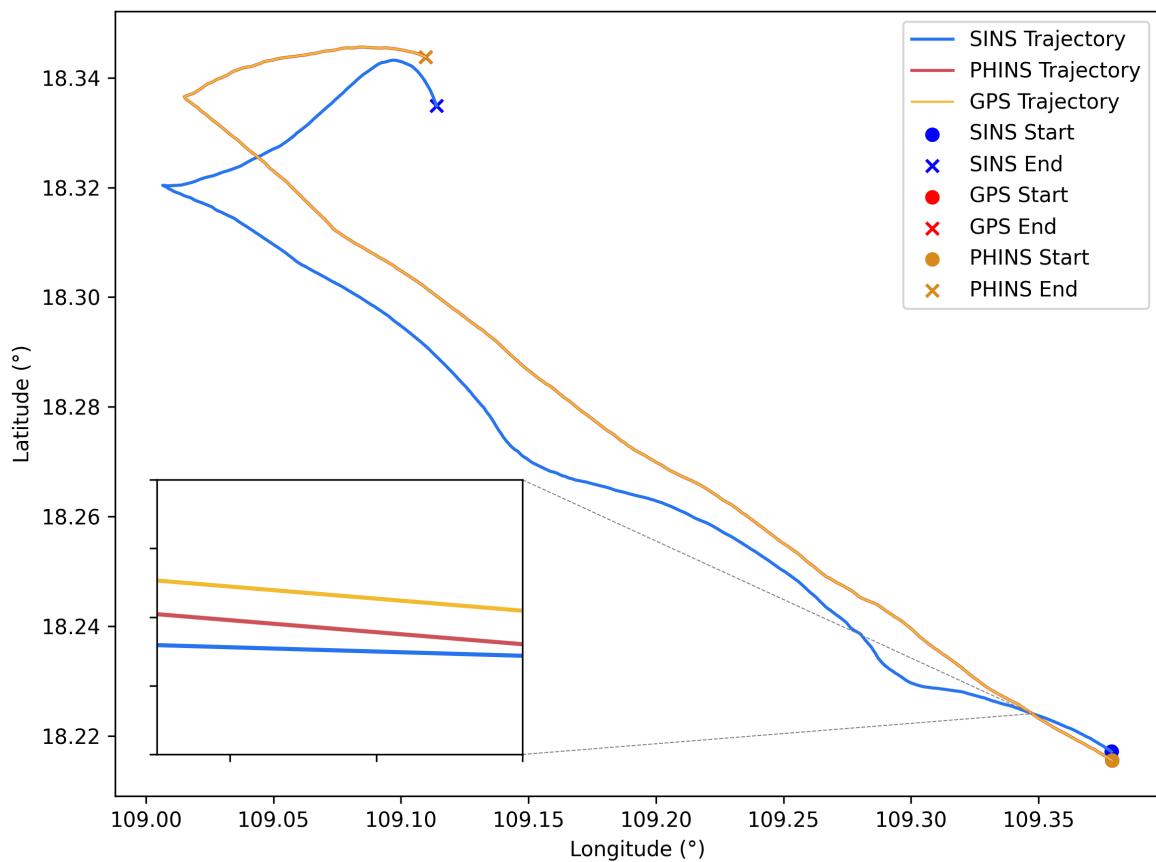
All experiments in this study were conducted on a system equipped with an NVIDIA GeForce GTX 1060 5 GB GPU and an Intel Core i9-10900X CPU, running Python 3.13 and PyTorch 2.7. We conducted data collection in the coastal waters near Sanya, Hainan Province, China, using an AUV equipped with a PHINS (iXblue, Saint-Germain-en-Laye, France), a domestically developed SINS, a GPS receiver (Septentrio, Leuven, Belgium), and a DVL (Teledyne RDI, San Diego, CA, USA) to collect surface navigation data. When all sensors were operational, data was recorded at a frequency of 1 Hz, resulting in a total of 17,130 trajectory points. To preserve local temporal consistency, the dataset was split into groups of 10 consecutive data points each. Within each group, the first seven points were used for training and the remaining three for testing, maintaining a 70:30 split that preserves short-term temporal dependencies essential for training models on sequential navigation data. Although the experiments in this study focus on surface navigation, the Z-axis is retained in the definitions of both the nominal and error states to maintain completeness and consistency with standard AUV navigation formulations. This design also facilitates future extensions of the proposed method to scenarios involving vertical motion or depth variations.

Table 1 presents a comparison of the horizontal positioning errors observed under SINS, GPS, and the integrated navigation system.

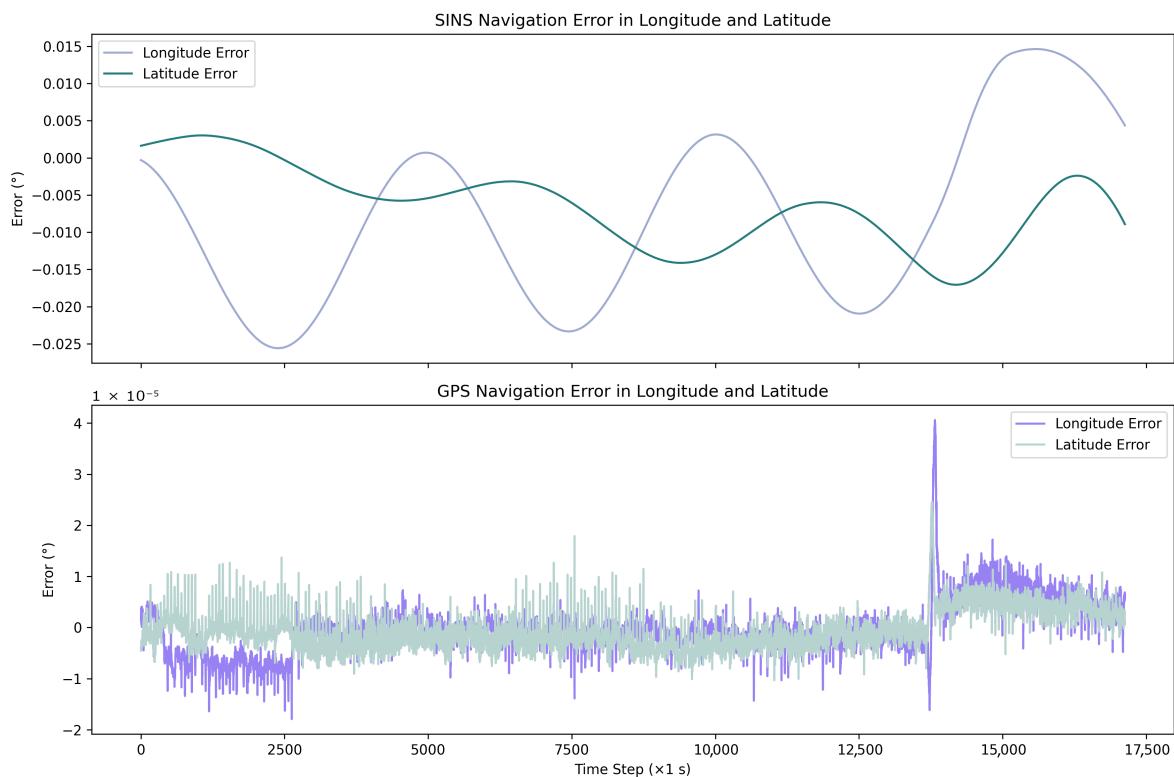
**Table 1.** Position estimation accuracy of our method and other navigation solutions.

Navigation Solutions	Horizontal Error (m)	RMSE
SINS	1715.8108	1209.5987
GPS	0.6714	0.5114
SINS/DVL	2.3478	0.5086
SINS/DVL/GPS	0.6287	0.5016

Additionally, Figure 5 visualizes the aforementioned trajectories, providing an intuitive comparison of the navigation paths. Figure 6 further illustrates the error variations of GPS and SINS along both longitude and latitude directions over time, highlighting the differences in their positioning accuracy and drift characteristics.



**Figure 5.** Comparison of navigation trajectories for different methods.



**Figure 6.** Error variations of GPS and SINS in longitude and latitude directions over time.

The measurement noise covariance matrix  $R$  and the process noise covariance matrix  $Q$  are initialized as follows:

$$Q \times \Delta t = \begin{bmatrix} 0.1 \cdot I_3 & 0 & 0 & 0 & 0 \\ 0 & 0.05 \cdot I_3 & 0 & 0 & 0 \\ 0 & 0 & 0.05 \cdot I_3 & 0 & 0 \\ 0 & 0 & 0 & 0.25 \cdot I_3 & 0 \\ 0 & 0 & 0 & 0 & 0.05 \cdot I_3 \end{bmatrix} \quad (29)$$

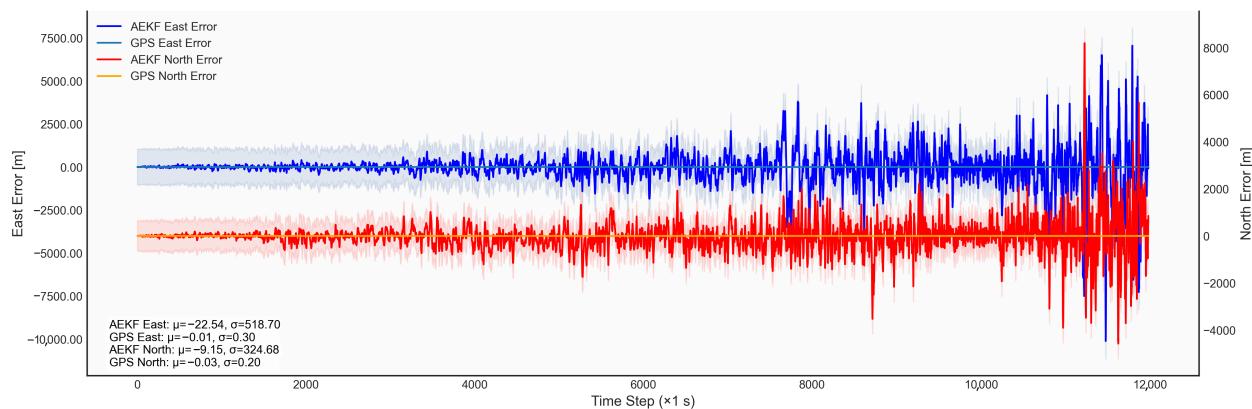
$$R = \begin{bmatrix} 0.8^2 & 0 & 0 & 0 & 0 \\ 0 & 0.6^2 & 0 & 0 & 0 \\ 0 & 0 & 0.3^2 & 0 & 0 \\ 0 & 0 & 0 & 0.6^2 & 0 \\ 0 & 0 & 0 & 0 & 0.4^2 \\ 0 & 0 & 0 & 0 & 0.3^2 \end{bmatrix}$$

where  $I_3$  denotes the  $3 \times 3$  identity matrix.

To ensure efficient model convergence, we employ the AdamW optimizer in combination with the OneCycleLR scheduler. The initial learning rate is set to  $1 \times 10^{-3}$ , with a maximum learning rate of  $1 \times 10^{-1}$ . The scheduler uses a 30% warm-up phase followed by cosine annealing decay, which enables the model to escape sharp local minima early and converge toward flatter, and more generalizable solutions. Empirically, this dynamic learning rate adjustment leads to faster convergence and lower validation errors compared to fixed learning rate baselines.

#### 4.2. Results

Firstly, the performance of the Adaptive Extended Kalman Filter (AEKF), which dynamically adjusts the process and measurement noise covariances, is compared with that of the proposed method in terms of the eastward and northward positional errors [41], as illustrated in Figure 7.



**Figure 7.** Comparison of positioning errors in the East–North frame. The left axis shows the East-direction error and the right axis shows the North-direction error over time. Solid lines represent the error time series for each method, while shaded areas indicate the  $\pm 1$  standard deviation range. Mean ( $\mu$ ) and standard deviation ( $\sigma$ ) values are computed for both directions and methods to quantitatively evaluate performance.

The AEKF employed in our comparative experiments, with a forgetting factor of  $\lambda = 0.995$ , updates the  $Q$  and  $R$  according to the short-term statistical characteristics of the innovation sequence. However, its adaptive capability is inherently constrained by local linearization and Gaussian statistical assumptions, rendering the AEKF vulnerable to transient measurement anomalies and dynamic maneuvers. Such limitations often lead to oscillatory or inconsistent covariance adaptation. In contrast, the proposed method leverages data-driven global mapping to learn the nonlinear relationships and long-term dependencies

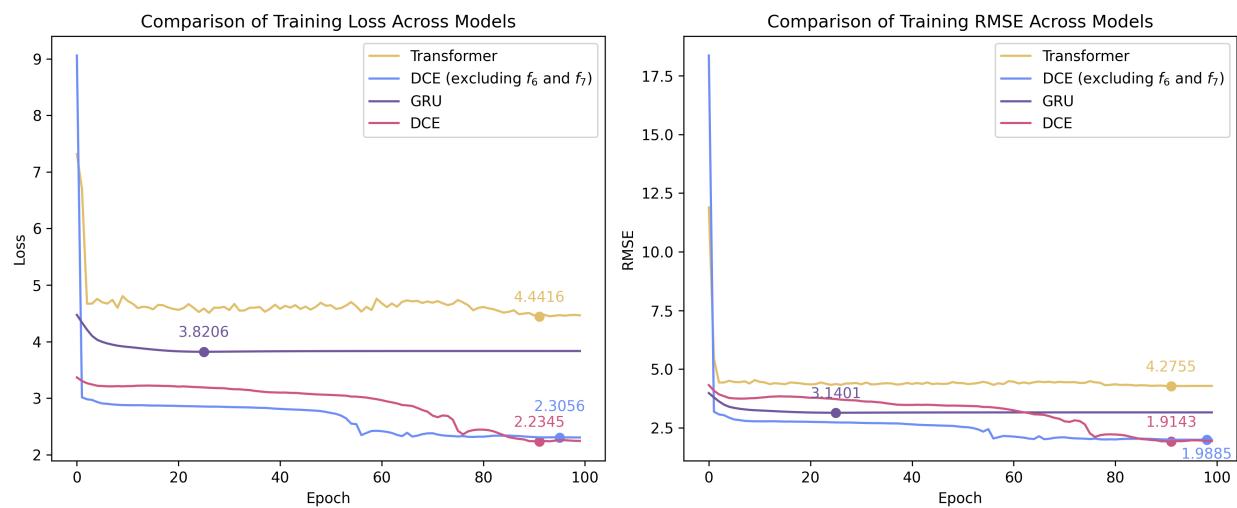
between the optimal noise parameters and system states from observational data. As a result, it effectively captures complex, non-Gaussian, and context-dependent noise patterns, thereby providing smoother and more robust covariance estimation alongside improved noise modeling accuracy.

To further verify the comprehensive improvements in implicit learning efficiency brought about by our designed features and model architecture, we compare the proposed method with two representative models, Transformer and GRU, in terms of positioning accuracy, convergence speed, the duration of a single epoch, and model parameter size. In addition, we include a variant of our method that excludes features  $f_6$  and  $f_7$  to further evaluate the contribution of these inputs. The training parameters of our DCE are listed in Table 2.

**Table 2.** Parameter settings used for training the proposed DCE.

Settings	Value
Epoch	100
Batch	1
Sequence length	1
Multi-heads	4
Linear dimension	128
Input dimension	61
Output dimension	21
The shape of $Q$	(15, 15)
The shape of $R$	(6, 6)
Initial learning rate	$1 \times 10^{-3}$

Figure 8 presents the training curves of the RMSE and loss value for the four models above. The minimum values attained by each model are visually highlighted in the figure to facilitate performance comparison.

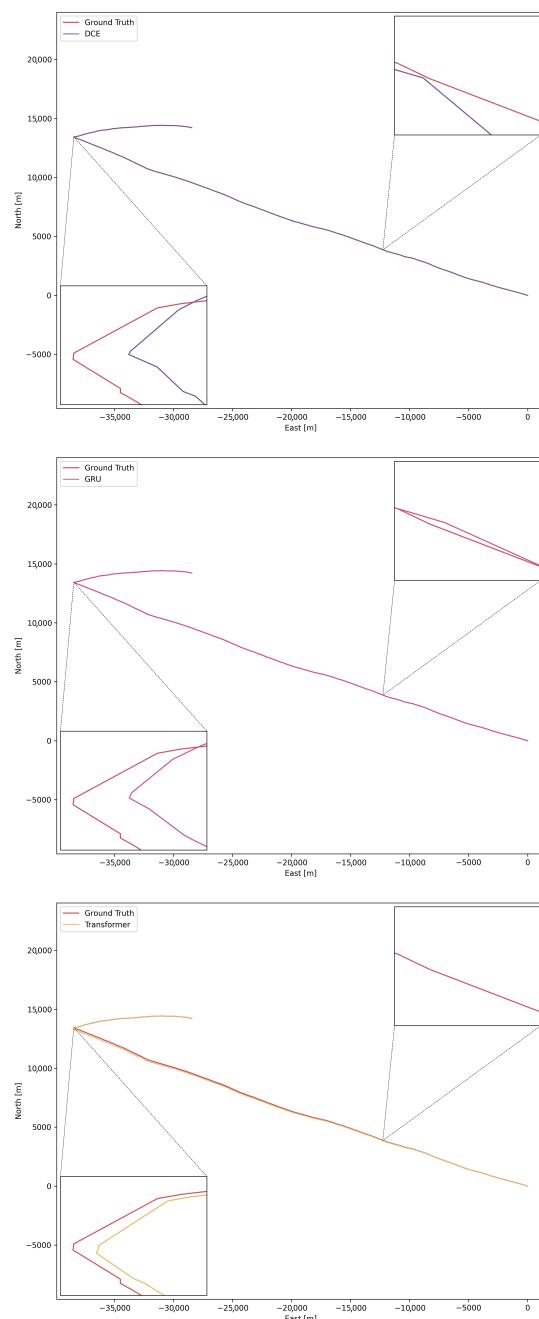


**Figure 8.** The progression of RMSE and loss values throughout training for the four models.

All metrics reported in Table 3 are averaged over multiple runs. Our proposed method demonstrates comprehensive advantages over the compared methods. It achieves the highest positioning accuracy, with the lowest RMSE of 0.0274. Compared to conventional loosely coupled integrated navigation, Transformer, GRU, and the variant of DCE without using  $f_6$  and  $f_7$ , our method reduces the RMSE by 94.6%, 68.7%, 55.9%, and 34.6%, respectively. The horizontal error of our model is minimized to 0.0003, representing reductions of

99.99%, 99.94%, 57.14%, and 40.00% compared to the above approaches, respectively. In addition, it requires fewer trainable parameters, resulting in reduced computational complexity. The testing phase is also more efficient, with the shortest inference time among all methods. This provides a feature design perspective in which appropriate feature extraction techniques trade off some computational efficiency for faster convergence and enhanced accuracy. These results indicate that our DL-assisted ES-EKF scheme has higher accuracy, a lighter structure, and stronger real-time performance in integrated navigation tasks.

As the trajectories of the four models largely overlap, making them difficult to distinguish when plotted together, they are displayed separately in Figure 9. In each plot, two identical segments of the trajectory are enlarged to better highlight the differences in local behavior.



**Figure 9.** Trajectory comparison among the Transformer, DCE, and GRU.

**Table 3.** Evaluation of the proposed method against other covariance estimators.

Covariance Estimator	Loss (RMSE)	Trainable Parameters	Convergence Epoch	Test Duration (s)
Transformer	0.4776 (0.0877)	1,196,693	92	17.44
GRU	0.0007 (0.0622)	175,125	23	14.62
Our model (excluding $f_6$ and $f_7$ )	0.0005 (0.0419)	108,565	95	12.70
Our method	0.0003 (0.0274)	110,229	91	13.72

## 5. Conclusions

In summary, the DL-assisted integrated navigation scheme proposed in this study leverages the powerful representational capabilities of deep neural networks to effectively address the challenges inherent in complex maritime navigation and the intricate error characteristics of inertial measurement units. Specifically, the method is designed for our AUV operating on the water surface, where conventional integrated navigation systems often experience degraded accuracy due to the dynamic and uncertain marine environment. By incorporating learned features and model structures suited to these conditions, the proposed approach enhances both positioning precision and robustness. Compared with model-based and learning-based methods, our proposed approach achieves significantly higher navigation solution accuracy while maintaining acceptable computational complexity and improved interpretability. Nevertheless, the current methodology exhibits a pronounced reliance on the continuous availability of GPS signals, thereby constraining its applicability in scenarios subject to frequent GPS signal outages. In response to this challenge, future research will be directed towards extending the proposed approach to submerged navigation environments and exploring robust mechanisms to ensure consistent and accurate navigation in the presence of GPS signal outages.

**Author Contributions:** Conceptualization, methodology, visualization, writing—original draft and editing, Y.Y.; supervision, project administration, and funding acquisition, B.X.; investigation, B.Y.; supervision, F.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the National Natural Science Foundation of China under Grant No. 62473106, titled “Research on Cross-Medium Navigation and Seamless Positioning Methods for Submersible-Aerial Vehicles” (January 2025–December 2028), the Hainan Provincial Key Research and Development Program under Grant No. ZDYF2024GXJS009.

**Data Availability Statement:** Data are not available for privacy reasons.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Li, Z.; Liu, W.; Li, L.; Guo, L.; Li, L. Modelling of a cable-drogue docking system for AUV. In Proceedings of the Global Oceans 2020: Singapore–U.S. Gulf Coast, Biloxi, MS, USA, 5–30 October 2020; pp. 1–5. [[CrossRef](#)]
2. Shao, H.; Wei, Y.; Guo, T.; Niu, J. An OCSVM Aided Integrated Navigation Fault Tolerance Strategy for Submarine-Pipeline-Detection AUV. In Proceedings of the OCEANS 2021: San Diego–Porto, San Diego, CA, USA, 20–23 September 2021; pp. 1–6. [[CrossRef](#)]
3. Lin, C.; Han, G.; Du, J.; Bi, Y.; Shu, L.; Fan, K. A Path Planning Scheme for AUV Flock-Based Internet-of-Underwater-Things Systems to Enable Transparent and Smart Ocean. *IEEE Internet Things J.* **2020**, *7*, 9760–9772. [[CrossRef](#)]
4. Bobkov, V.; Morozov, M.; Inzartsev, A.; Panin, M. AUV Inspection of Subsea Pipelines Using Information from an Onboard Stereo Camera. In Proceedings of the 2024 International Conference on Ocean Studies (ICOS), Vladivostok, Russia, 8–11 October 2024; pp. 71–75. [[CrossRef](#)]
5. Liu, Z.; Hou, J.; Ning, D.; Zhou, C.; Liang, G.; Zhang, F. Improving Deep Q Network Based on Marketing Psychology for AUV Path Planning in Unknown Marine Environments. *IEEE Internet Things J.* **2025**, *12*, 5476–5487. [[CrossRef](#)]
6. Xu, B.; Guo, Y.; Wang, X. A DVL Measurement Correction Method in High Dynamic Environment and its Application for SINS/DVL Integrated Navigation. In Proceedings of the OCEANS 2024, Singapore, 14–18 April 2024; pp. 1–7. [[CrossRef](#)]

7. Yao, Y.; Xu, X.; Xu, X.; Klein, I. Virtual Beam Aided SINS/DVL Tightly Coupled Integration Method with Partial DVL Measurements. *IEEE Trans. Veh. Technol.* **2023**, *72*, 418–427. [\[CrossRef\]](#)
8. Luo, L.; Huang, Y.; Zhang, Z.; Zhang, Y. A New Kalman Filter-Based In-Motion Initial Alignment Method for DVL-Aided Low-Cost SINS. *IEEE Trans. Veh. Technol.* **2021**, *70*, 331–343. [\[CrossRef\]](#)
9. Feng, K.; Li, J.; Zhang, D.; Wei, X.; Yin, J. Robust Cubature Kalman Filter for SINS/GPS Integrated Navigation Systems with Unknown Noise Statistics. *IEEE Access* **2021**, *9*, 9101–9116. [\[CrossRef\]](#)
10. Sarda, E.; Dhanak, M. Launch and Recovery of an Autonomous Underwater Vehicle from a Station-Keeping Unmanned Surface Vehicle. *IEEE J. Ocean. Eng.* **2018**, *44*, 290–299. [\[CrossRef\]](#)
11. AUV Recovery Technology Development Based on Unmanned Surface Platform. *J. Unmanned Undersea Syst.* **2023**, *31*, 501. [\[CrossRef\]](#)
12. Gao, L.; Fan, Z.; He, T.; Lv, J.; Zhang, X. A Loosely Coupled INS/BDS Integrated Navigation System. In Proceedings of the 2024 5th International Conference on Computer Vision, Image and Deep Learning (CVIDL), Zhuhai, China, 19–21 April 2024; pp. 1054–1058. [\[CrossRef\]](#)
13. Zhao, L.; Qiu, H.; Feng, Y. Analysis of a robust Kalman filter in loosely coupled GPS/INS navigation system. *Measurement* **2016**, *80*, 138–147. [\[CrossRef\]](#)
14. H T, M.; Akram, V.; Reddy, S.; KN, M.G.; LD, U.K. An Error-State Extended Kalman Filter Based State Estimation and Localization Algorithm for Autonomous Systems. In Proceedings of the 2023 International Conference on Smart Systems for Applications in Electrical Sciences (ICSSES), Tumakuru, India, 7–8 July 2023; pp. 1–6. [\[CrossRef\]](#)
15. Kim, C.; Bae, G.; Shin, W.; Wang, S.; Oh, H. EKF-Based Radar-Inertial Odometry with Online Temporal Calibration. *IEEE Robot. Autom. Lett.* **2025**, *10*, 7230–7237. [\[CrossRef\]](#)
16. Gao, P.; Fang, J.; He, J.; Ma, S.; Wen, G.; Li, Z. GRU–Transformer Hybrid Model for GNSS/INS Integration in Orchard Environments. *Agriculture* **2025**, *15*, 1135. [\[CrossRef\]](#)
17. Niu, Z.; Cong, L.; Qin, H. Research on High Precision Attitude Estimation Based on Gait Cycle Modeling with IMU for PDR. In Proceedings of the 2023 3rd International Conference on Electronic Information Engineering and Computer Science (EIECS), Changchun, China, 22–24 September 2023; pp. 574–579. [\[CrossRef\]](#)
18. Brigadnov, I.; Lutonin, A.; Bogdanova, K. Error State Extended Kalman Filter Localization for Underground Mining Environments. *Symmetry* **2023**, *15*, 344. [\[CrossRef\]](#)
19. Gogliettino, G.; Pisoni, F.; Di Grazia, D. Use of Reinforcement Learning to Improve GNSS Satellites Signal Acquisition Search Strategy. In Proceedings of the 2024 IEEE International Workshop on Metrology for Automotive (MetroAutomotive), Bologna, Italy, 26–28 June 2024; pp. 58–63. [\[CrossRef\]](#)
20. Bai, Y.; Jia, W.; Jin, X.; Su, T.; Kong, J. Data-Driven Integrated Inertial Navigation Based on ESN Model. In Proceedings of the 2023 China Automation Congress (CAC), Chongqing, China, 17–19 November 2023; pp. 357–361. [\[CrossRef\]](#)
21. Shao, Y.h.; Han, B.; Luo, Y. A method based on CNN-BiLSTM for UAV navigation error compensation in GNSS denied environment. In Proceedings of the 2023 9th International Conference on Computer and Communications (ICCC), Chengdu, China, 8–11 December 2023; pp. 689–694. [\[CrossRef\]](#)
22. Kanhere, A.V.; Gupta, S.; Shetty, A.; Gao, G. Improving GNSS Positioning using Neural Network-based Corrections. *arXiv* **2022**, arXiv:2110.09581v3.. [\[CrossRef\]](#)
23. Zhang, H.; Xiong, H.; Hao, S.; Yang, G.; Wang, M.; Chen, Q. A Novel Multidimensional Hybrid Position Compensation Method for INS/GPS Integrated Navigation Systems During GPS Outages. *IEEE Sens. J.* **2024**, *24*, 962–974. [\[CrossRef\]](#)
24. Wang, G.; Xu, X.; Yao, Y.; Tong, J. A Novel BPNN-Based Method to Overcome the GPS Outages for INS/GPS System. *IEEE Access* **2019**, *7*, 82134–82143. [\[CrossRef\]](#)
25. Viswanath, A.; Sameer, S.M. A Novel Elman Network Based INS/GPS Fusion Filter to Enhance Tracking Accuracy in UAVs. In Proceedings of the 2021 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT), Online, 17–18 July 2021; pp. 16–20. [\[CrossRef\]](#)
26. Cohen, N.; Yampolsky, Z.; Klein, I. Set-Transformer BeamsNet for AUV Velocity Forecasting in Complete DVL Outage Scenarios. In Proceedings of the 2023 IEEE Underwater Technology (UT), Tokyo, Japan, 6–9 March 2023; pp. 1–6. [\[CrossRef\]](#)
27. Yona, M.; Klein, I. Compensating for Partial Doppler Velocity Log Outages by Using Deep- Learning Approaches. In Proceedings of the 2021 IEEE International Symposium on Robotic and Sensors Environments (ROSE), Virtual Conference, 28–29 October 2021; pp. 1–5. [\[CrossRef\]](#)
28. Feng, S.; Li, X.; Zhang, S.; Jian, Z.; Duan, H.; Wang, Z. A review: State estimation based on hybrid models of Kalman filter and neural network. *Syst. Sci. Control Eng.* **2023**, *11*, 2173682. [\[CrossRef\]](#)
29. Jouaber, S.; Bonnabel, S.; Velasco-Forero, S.; Pilté, M. NNAKF: A Neural Network Adapted Kalman Filter for Target Tracking. In Proceedings of the ICASSP 2021—2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 6–11 June 2021; pp. 4075–4079. [\[CrossRef\]](#)

30. Xu, L.; Niu, R. EKFNet: Learning System Noise Covariance Parameters for Nonlinear Tracking. *IEEE Trans. Signal Process.* **2024**, *72*, 3139–3152. [[CrossRef](#)]
31. Li, S.; Mikhaylov, M.; Pany, T.; Mikhaylov, N. Exploring the Potential of the Deep-Learning-Aided Kalman Filter for GNSS/INS Integration: A Study on 2-D Simulation Datasets. *IEEE Trans. Aerosp. Electron. Syst.* **2024**, *60*, 2683–2691. [[CrossRef](#)]
32. Choi, G.; Park, J.; Shlezinger, N.; Eldar, Y.C.; Lee, N. Split-KalmanNet: A Robust Model-Based Deep Learning Approach for State Estimation. *IEEE Trans. Veh. Technol.* **2023**, *72*, 12326–12331. [[CrossRef](#)]
33. Buchnik, I.; Revach, G.; Steger, D.; van Sloun, R.J.G.; Routtenberg, T.; Shlezinger, N. Latent-KalmanNet: Learned Kalman Filtering for Tracking from High-Dimensional Signals. *IEEE Trans. Signal Process.* **2024**, *72*, 352–367. [[CrossRef](#)]
34. Revach, G.; Shlezinger, N.; Ni, X.; Escoriza, A.L.; van Sloun, R.J.G.; Eldar, Y.C. KalmanNet: Neural Network Aided Kalman Filtering for Partially Known Dynamics. *IEEE Trans. Signal Process.* **2022**, *70*, 1532–1547. [[CrossRef](#)]
35. Chung, J.; Gulcehre, C.; Cho, K.; Bengio, Y. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *arXiv* **2014**, arXiv:1412.3555v1. [[CrossRef](#)]
36. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. *arXiv* **2023**, arXiv:1706.03762. [[CrossRef](#)] [[PubMed](#)]
37. Grewal, M.S.; Weill, L.R.; Andrews, A.P. *Global Positioning Systems, Inertial Navigation, and Integration*; Wiley-Interscience: New York, NY, USA, 2007.
38. Zhang, L.; Zhou, H.; Gao, Y. An in-motion alignment method of AUV SINS/DVL navigation system based on FGO. *Measurement* **2023**, *222*, 113578. [[CrossRef](#)]
39. Bian, Y.; Li, Z.; Wang, G.; Qin, H.; Hu, M.; Qin, X.; Ding, R. A Unidirectional Trend IMM Method for SINS/DVL/USBL Navigation System Under the Time-Varying Noise Environment. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 1–14. [[CrossRef](#)]
40. Sheng, G.; Liu, X.; Zhang, Y.; Shao, Q.; Xu, H.; Cheng, X.; Yuan, X. The EKF-based SINS/DVL integrated navigation for AUV on lie group under hovering condition. *Ocean Eng.* **2025**, *325*, 120742. [[CrossRef](#)]
41. Mehra, R. On the identification of variances and adaptive Kalman filtering. *IEEE Trans. Autom. Control* **1970**, *15*, 175–184. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.