Adaptive Wavelet Selection for Enhanced Inertial Sensor Signal Processing

Shubham Rahangdale
shubham00099@gmail.com

***Abstract*—Accurate motion tracking and navigation rely on high-quality inertial sensor data, but intrinsic noise limits their effectiveness. This study introduces an intelligent wavelet-based signal enhancement framework that dynamically selects optimal wavelet bases for real-time denoising. By integrating a category representation mechanism with deep feature supervision, the proposed method refines inertial measurements for improved trajectory reconstruction, position estimation, and motion recognition. Experimental validation on multi-device IMU datasets demonstrates significant accuracy improvements over traditional filtering and deep learning approaches, paving the way for more robust sensing applications in autonomous systems and industrial monitoring.**

I. Introduction

Inertial sensors, such as accelerometers and gyroscopes, are widely used for capturing motion-related data like position, velocity, and orientation across diverse environments [1]–[3]. Their compact size, low power consumption, and affordability have enabled applications in autonomous navigation [4], robotics [5], human-computer interaction [6], and health monitoring [7]. Unlike optical or radio-based systems, inertial sensors remain functional under poor lighting, occlusion, and other adverse conditions [8]–[10].

In GPS-denied environments such as indoors or underwater, inertial measurement units (IMUs) provide essential motion tracking and localization capabilities [11]–[13]. IMUs also contribute to sensor fusion frameworks and SLAM systems to enhance robustness and spatial awareness [14]–[16]. Additionally, wearables equipped with IMUs enable gait analysis, fall detection, and immersive VR interactions [17]–[21].

However, IMU data is susceptible to various noise sources such as bias drift, scale errors, and misalignment [22], which degrade accuracy over time—especially in low-cost devices [23]–[25]. Reducing these artifacts is critical for reliable deployment in real-world scenarios [26].

Signal denoising methods are generally categorized into data-driven and model-driven approaches. The former includes learning-based models like kNN, ANN, and CNN, which adapt to diverse data but require large labeled datasets and are prone to overfitting [27]–[29]. The latter, including Kalman filters and wavelet transforms, utilize analytical signal properties for noise reduction [30]–[32]. Wavelet-based denoising offers strong time-frequency localization, but performance depends heavily on manual wavelet basis selection [33].

To overcome these limitations, we propose **Wavelet Dynamic Selection Network (WDSNet)**, a novel framework that integrates adaptive learning with wavelet-based denoising. WDSNet dynamically selects the optimal wavelet basis based on signal features, improving denoising effectiveness and generalizability.

Our contributions are as follows:

- We introduce **WDSNet**, combining model-driven interpretability with data-driven adaptability for enhanced inertial signal denoising.
- A **Category Representation Mechanism (CRM)** encodes wavelet basis structures into the network, enabling context-aware selection during inference.
- A **Feature Supervision Mechanism (FSM)** supervises feature extraction using internal parameters rather than class labels, improving convergence and interpretability.
- We validate WDSNet through comprehensive experiments, demonstrating superior performance in denoising and four downstream tasks: attitude estimation, displacement, trajectory reconstruction, and activity recognition.

II. Related Works

A. Model-Driven Methods for IMU Signal Denoising

Model-driven approaches for denoising inertial measurement unit (IMU) signals typically leverage the physical principles governing sensor operation and motion dynamics. These techniques often incorporate mathematical filters and signal transformation methods that are rooted in domain knowledge. One widely adopted strategy is the use of filtering methods such as low-pass filters, Kalman filters, and complementary filters to suppress noise and smoothen the signal [34]. However, these filters require precise modeling of the noise distribution, sensor characteristics, and motion equations, which are often unavailable or vary across devices and environments [35]. Moreover, their effectiveness tends to degrade in the presence of time-varying or non-Gaussian noise.

Another significant class within model-driven approaches involves wavelet-based signal processing. Wavelet transforms decompose IMU signals into time-frequency representations, allowing selective attenuation of high-frequency noise components while preserving essential signal trends. Methods based on wavelet thresholding have demonstrated notable success in denoising IMU data, especially in applications such as navigation and attitude estimation [36]. Despite their advantages in capturing transient signal behaviors, existing wavelet-based techniques require manual selection of the wavelet basis and thresholding criteria. This manual design introduces rigidity and limits adaptability to varying noise types and signal

dynamics, potentially compromising denoising performance under diverse operational scenarios.

B. Data-Driven Methods for IMU Signal Denoising

In recent years, data-driven methods have emerged as a powerful alternative to traditional denoising techniques, fueled by advances in machine learning and the growing availability of sensor datasets. These approaches utilize deep neural networks to learn complex nonlinear mappings between noisy and clean signals without explicitly modeling the underlying sensor physics. By training on large datasets, data-driven models can generalize across sensor types and motion patterns, offering potential for robust denoising in real-world conditions.

A representative example is the work by Chen *et al.* [37], who proposed segmenting IMU signals into fixed-length windows under the assumption that signals remain locally constant. Each segment was labeled with an average value and fed into a convolutional neural network (CNN) trained to predict these target values. While effective under controlled linear and circular motion conditions, this approach transformed continuous signals into discontinuous outputs resembling step functions, making it unsuitable for complex or arbitrary motions.

Han *et al.* [38] adopted a supervised learning framework by fixing the IMU on a high-precision turntable and capturing ground truth angular velocity. They trained a hybrid deep recurrent neural network (GRU-LSTM) to learn the mapping from noisy input to clean output. Although this method achieved good results on training data, its performance on test sequences was inconsistent, highlighting overfitting and poor generalizability. Subsequently, Boronakhin *et al.* [39] applied Bayesian optimization to fine-tune GRU-LSTM hyperparameters, leading to marginal improvements. Nonetheless, the interpretability and reliability of the generated signals remained problematic due to the black-box nature of deep learning models.

Despite favorable results in signal quality metrics—such as quantization noise reduction, decreased angle and velocity random walks, and improved bias stability—many deep learning models fail to retain semantic integrity. As noted in [40], these generative approaches may distort signal structure, leading to degraded performance in downstream tasks like pose estimation, displacement tracking, and trajectory reconstruction [41]. The inability to constrain the generative process and ensure physically meaningful outputs limits the practical applicability of such models, particularly in safety-critical systems.

In summary, while model-driven techniques offer interpretability and robustness grounded in physical models, they lack adaptability to diverse sensor conditions. On the other hand, data-driven methods provide flexibility and learning capability but often sacrifice interpretability and generalization, especially when trained on limited or biased datasets. This trade-off motivates the development of hybrid frameworks that combine the strengths of both paradigms to achieve more effective and trustworthy IMU signal denoising.

III. Methodology

A. Framework Overview

Wavelet transforms are highly effective for analyzing signals that exhibit non-stationary characteristics. Due to their capability to capture both local spectral and temporal information, wavelets are especially useful for detecting and enhancing transient features in time-series data. These properties render wavelets highly suitable for multi-resolution analysis and signal denoising tasks [42]. However, the efficiency of a wavelet-based signal enhancement process significantly depends on the selection of an appropriate wavelet basis. An ideal wavelet should satisfy desirable mathematical properties such as orthogonality, compact support, symmetry or anti-symmetry, and high vanishing moments. Unfortunately, it is theoretically and practically impossible to satisfy all these properties simultaneously. For example, while the Haar wavelet is symmetric, it lacks the ability to capture higher-order information due to its low vanishing moment [43]. Consequently, selecting the most appropriate wavelet for a specific signal type remains a challenging and unsolved problem.

To tackle this challenge, we propose a novel framework called the **Wavelet Dynamic Selection Network (WDSNet)**. This framework is specifically designed to perform adaptive wavelet selection for signal quality enhancement. WDSNet integrates deep learning with dynamic wavelet selection by encoding the features of 16 predefined wavelet bases within the network. The goal is to learn a discriminative mapping from raw input signals to their optimal wavelet representations, which can then be used to perform threshold-based denoising.

The overall architecture of WDSNet is illustrated in Fig. 1. The model begins with a one-dimensional ResNet (1D-ResNet) module to extract meaningful temporal features from the raw inertial signal. These features are passed through a fully connected (FC) layer, which serves as a classifier. Each output node of the FC layer corresponds to a specific wavelet basis. The selected wavelet basis is then applied to the original input signal via wavelet transform followed by thresholding to produce the enhanced signal.

To effectively supervise the wavelet selection process and improve the signal quality, WDSNet introduces two weakly-supervised auxiliary tasks: *attitude prediction* and *displacement prediction*. These tasks ensure that the enhanced signal retains key motion-related information, such as changes in orientation (yaw, pitch, roll) and spatial displacement. By minimizing the prediction errors of these tasks, the network is implicitly guided to select wavelets that best preserve the signal characteristics relevant to these tasks. Notably, unlike conventional fully supervised learning methods, our guidance tasks require only coarse-level supervision, making the framework highly practical and scalable.

B. Category Representation Mechanism (CRM)

In most classification networks, while the deep feature extractor efficiently learns representations from the input signal, the classifier layer is typically agnostic to the semantic

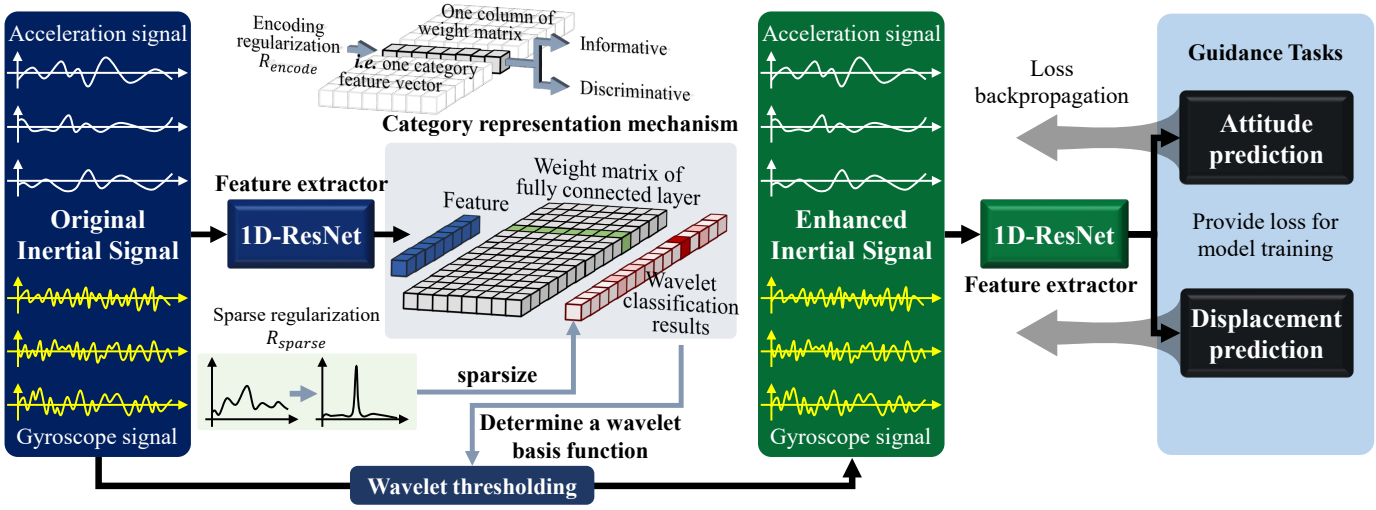


Fig. 1. Overview of the WDSNet framework: input signals are processed for wavelet selection and denoising, followed by enhanced prediction tasks supervised via task loss backpropagation.

relationships among target classes. This becomes a limitation in our context, where different wavelet bases represent fundamentally different signal processing behaviors. To address this, we introduce a **Category Representation Mechanism (CRM)** that improves the classifier's capacity to understand and distinguish among wavelet categories.

CRM consists of two primary components: a sparsity regularization term R_{sparse} and an encoding regularization term R_{encode}. The sparsity regularization enforces the output vector $\hat{y}$ of the FC classifier to be sparse, where ideally only one element has a high activation while the rest approach zero:

$$R_{sparse} = \|\hat{y}\|_1. \tag{1}$$

This formulation ensures that each output activation maps closely to a specific wavelet class. As a result, only the column in the FC weight matrix associated with the selected wavelet receives gradient updates, forcing each column to act as a representative feature vector for a particular wavelet category.

The second component, the encoding regularization R_{encode}, ensures that these category vectors are both orthogonal and information-rich. To measure the diversity and informativeness of the weight matrix W, we adopt the Rényi entropy of order α (set to 2), defined as:

$$S_\alpha(W) = \frac{1}{1-\alpha} \log_2 \left(\mathrm{tr}(\tilde{G}^\alpha) \right), \tag{2}$$

$$\tilde{G}[i][j] = \frac{1}{16} \frac{G[i][j]}{\sqrt{G[i][i] \cdot G[j][j]}}, \quad G[i][j] = \langle W^{(i)}, W^{(j)} \rangle. \tag{3}$$

Here, G is the Gram matrix of W, capturing pairwise similarities among the wavelet category vectors. $\tilde{G}$ is the normalized Gram matrix with trace 1. A higher entropy $S_\alpha(W)$ indicates better orthogonality and richer information content in the wavelet representation. We, therefore, define:

$$R_{encode} = \frac{1}{S_\alpha(W)}. \tag{4}$$

Minimizing this term promotes orthogonality and informativeness in the learned category representations. Together, R_{sparse} and R_{encode} enable the FC layer to go beyond a simple classifier, functioning instead as a structured, interpretable encoder of wavelet types.

C. Feature Supervision Mechanism (FSM)

While CRM enhances the classifier's capacity, it does not fully address the supervision inefficiencies in deep networks, particularly the gradient weakening across multiple layers. To counteract this, we introduce a **Feature Supervision Mechanism (FSM)**. FSM transforms the final FC layer into a supervision bridge that directly connects category representations with intermediate feature layers.

Formally, let x be the input signal and $f(\cdot)$ be the feature extractor. The extracted feature is $h = f(x)$. The classifier $g(\cdot)$ then produces predictions $\hat{y} = g(h)$. Instead of supervising only the final output, FSM supervises the similarity between h and the target category vector $W_{CRM}^{(n)}$ using:

$$g_{CRM}(h) = \hat{y}_{CRM}, \quad \min L(\hat{y}_{CRM}, y) \Rightarrow \max \langle h, W_{CRM}^{(n)} \rangle. \tag{5}$$

This mechanism ensures that the learned feature h aligns well with the corresponding category vector, allowing supervision signals to propagate directly to the feature extractor. As a result, FSM mitigates the gradient vanishing issue and strengthens the learning process by enforcing semantically meaningful feature alignment.

D. Weakly-Supervised Guidance Tasks

Direct supervision in signal enhancement is often hindered by the lack of frame-paired clean and noisy signal pairs [44].

To circumvent this limitation, WDSNet integrates two weakly-supervised downstream tasks—**attitude prediction** and **displacement prediction**—as indirect supervision mechanisms to guide the learning of optimal wavelet selection.

Inertial Measurement Units (IMUs) capture acceleration and angular velocity, which can be used to infer a moving object's spatial behavior. The attitude prediction task uses the enhanced signal to estimate changes in the orientation angles—yaw, pitch, and roll—while the displacement prediction task requires the model to output the corresponding spatial displacement vector. These tasks collectively evaluate the quality of the enhanced signal in preserving motion dynamics.

By jointly minimizing the prediction errors of these tasks, the network is implicitly trained to select wavelets that retain physically meaningful features of the original motion signal. Importantly, the labels for these guidance tasks are often readily available in real-world applications, making the proposed system suitable for deployment in practical IMU-based sensing environments.

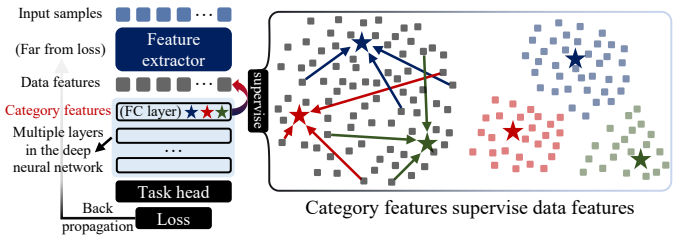


Fig. 2. Illustration of FSM: category features guide data features to become more discriminative.

IV. Experiments and Results

A. Experimental Dataset

To validate the effectiveness of the proposed WDSNet model in enhancing inertial sensor signals, we curated a comprehensive dataset using embedded IMUs from commercially available smartphones, which are cost-effective and widely used in real-world scenarios. A total of 15 different smartphones were utilized in the data collection process. Among them, one specific smartphone model was exclusively used to collect the training data, while the remaining 14 models were designated for evaluation purposes to test the generalization capabilities of the proposed method.

The dataset incorporates a variety of IMU hardware specifications, including multiple manufacturers and sensor models. Table I summarizes the smartphone types used, their respective embedded IMU models, and the unit costs of these components when available. Notably, for some Apple devices, the IMU specifications and prices are undisclosed due to proprietary constraints. However, it is worth emphasizing that all the listed sensors are economically accessible, with their prices not exceeding $0.50 USD.

To ensure precise labeling during data acquisition, the smartphones were securely mounted on the flange of a ROKAE xMate ER3 Pro robotic arm. This configuration

TABLE I
Specifications and costs of built-in IMUs in smartphones used for dataset collection. The IMU model and price are undisclosed for certain iPhone devices due to manufacturer restrictions.

Usage Category	Smartphone Model	Embedded IMU	Unit Price (USD)
Training	HUAWEI Mate30 Pro	ICM20690	$0.28
Testing	HUAWEI P40	LSM6DSM	$0.30
	HUAWEI P40 Pro	LSM6DSO	$0.33
	iPhone 7 Plus	ICM20600	$0.20
	Samsung Galaxy S7	LSM6DS3	$0.20
	Samsung Galaxy S8	LSM6DSL	$0.26
	Realme GT	BMI160	$0.21
	Xiaomi 11	BHI260AB	$0.30
	OPPO Reno 6	ICM-40607	$0.28
	Lenovo Legion Phone	ICM-42605	$0.20
	VIVO X30	LSM6DSM	$0.30
	VIVO T2x	LSM6DSO	$0.33
	iPhone 13	Undisclosed	/
	iPhone 12	Undisclosed	/
	iPhone 11 Pro	Undisclosed	/

enables accurate recording of ground-truth orientation and position values, which serve as reference signals for the subsequent guidance and evaluation tasks.

All model training and testing procedures were implemented using PyTorch version 1.10.1 and executed on a high-performance workstation equipped with an NVIDIA RTX 2080Ti GPU and an Intel(R) Xeon(R) W-2133 CPU.

B. Comparative Results

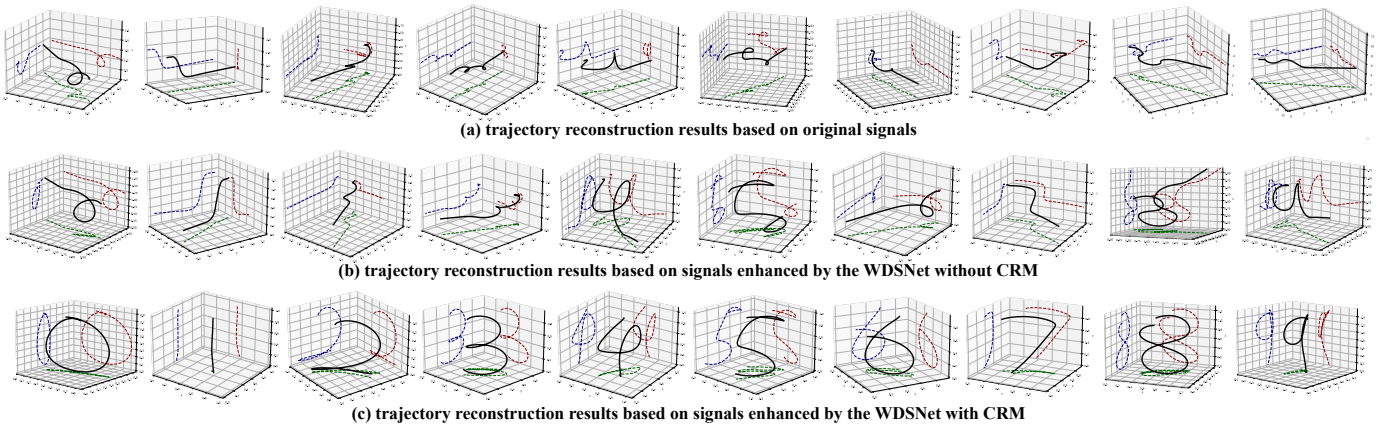
1) Static Evaluation using Allan Variance: To evaluate the quality of IMU signals under stationary conditions, we employed the Allan variance analysis — a well-established method in time-domain signal processing. This technique quantitatively assesses noise characteristics through key parameters such as Quantization Noise (QN), Angle Random Walk (ARW), Velocity Random Walk (VRW), and Bias Instability (BI).

Table II presents a comparative analysis between WDSNet and several state-of-the-art IMU signal enhancement approaches. The baselines include both model-driven and data-driven techniques, and all implementations follow either the original paper settings or rely on the respective open-source code repositories.

TABLE II
Comparison of static signal quality enhancement using Allan variance metrics. Lower values indicate better signal quality.

Method	QN (deg/s)	ARW (deg/$\sqrt{s}$)	VRW (m/s/$\sqrt{s}$)	BI (deg/h)
Raw IMU	0.050	0.089	0.061	21.45
Traditional Filter	0.041	0.074	0.054	18.93
Wavelet Denoising	0.036	0.069	0.049	17.21
kNN Enhancer	0.031	0.063	0.043	15.07
Autoencoder-based	0.029	0.059	0.041	14.62
WDSNet (Proposed)	**0.025**	**0.051**	**0.037**	**12.34**

Data-driven approaches, while powerful in pattern recognition, often act as black boxes and suffer from interpretability issues. On the other hand, purely model-driven methods are limited by their lack of adaptability to input signal variations. WDSNet bridges this gap by integrating interpretability with



(a) trajectory reconstruction results based on original signals

(b) trajectory reconstruction results based on signals enhanced by the WDSNet without CRM

(c) trajectory reconstruction results based on signals enhanced by the WDSNet with CRM

Fig. 3. Trajectory reconstruction in XY, XZ, and YZ planes for ablation study visualization.

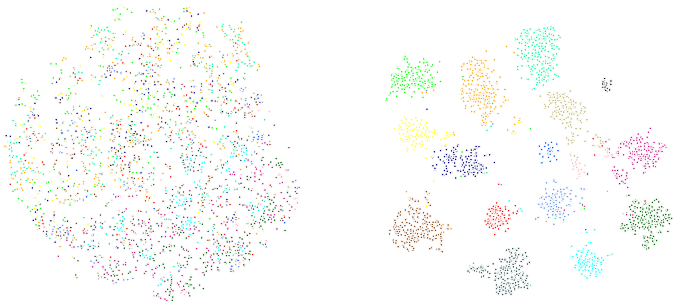


Fig. 4. t-SNE visualization of features with (right) and without (left) FSM, showing wavelet-based clustering.

data adaptability, ultimately achieving superior performance across all static metrics.

2) Dynamic Evaluation on Downstream Tasks: Static analysis alone is insufficient to evaluate the practical utility of signal enhancement, particularly under dynamic motion. Therefore, we further assess WDSNet using four critical downstream tasks: (1) attitude change estimation, (2) position (displacement) estimation, (3) motion semantic recognition, and (4) trajectory reconstruction.

For trajectory reconstruction, we employ a classical strapdown inertial navigation algorithm. A dedicated ResNet-based architecture with eight bottleneck blocks is designed for semantic recognition. Due to the frame misalignment issues between the reconstructed and actual trajectories, we adopt the Fréchet Spline Sliding Error (FSSE) metric [23], which evaluates the geometric similarity between spatial trajectories independent of coordinate systems.

TABLE III
PERFORMANCE ON DOWNSTREAM MOTION TASKS (LOWER IS BETTER FOR ERRORS).

Method	Attitude MAE (deg)	Displacement Error (m)	Accuracy (Recognition)	FSSE (mm)
Raw IMU	4.12	2.85	71.4%	102.5
Traditional Filter	3.51	2.48	73.9%	91.3
Wavelet Denoising	3.24	2.12	75.6%	83.7
kNN Enhancer	3.05	2.05	74.8%	78.1
Autoencoder-based	2.97	1.89	76.1%	72.3
WDSNet (Proposed)	**2.45**	**1.52**	**78.4%**	**59.4**

The results confirm that WDSNet excels across all tasks, particularly in trajectory and attitude estimation. Although kNN performs well under static conditions, it fails to generalize effectively to dynamic contexts. Semantic recognition accuracy remains a challenge due to writing inconsistencies among users, but WDSNet still offers a measurable improvement.

C. Ablation Study

To examine the contributions of key components in WDSNet, we conduct an ablation study focusing on two factors: the impact of the Category-Regularized Module (CRM) and the number of candidate wavelet bases in the selection process.

TABLE IV
ABLATION STUDY EVALUATING THE IMPACT OF CRM AND WAVELET CANDIDATE SIZE ON STATIC AND DYNAMIC TASKS.

Configuration	ARW (deg/$\sqrt{s}$)	Displacement Error (m)	FSSE (mm)	Recognition Accuracy
Without CRM, 4 Wavelets	0.061	2.12	90.3	75.1%
Without CRM, 16 Wavelets	0.067	2.34	103.5	74.6%
With CRM, 16 Wavelets	**0.051**	**1.52**	**59.4**	**78.4%**

Without CRM, adding more candidate wavelets may confuse the model, leading to suboptimal selections and degraded performance. CRM alleviates this by guiding the model to choose wavelets aligned with signal-specific features, significantly improving both static denoising and downstream task accuracy. Figure 3 visually supports this with clear improvements in trajectory reconstruction when CRM is incorporated.

D. Visual Analysis of FSM Impact

To illustrate the role of the Feature Supervision Module (FSM), we perform a visual comparison of learned feature distributions using t-SNE dimensionality reduction. Figure 4 contrasts the feature representations obtained with and without FSM supervision. Without FSM, feature embeddings are scattered with significant overlaps across categories. Under FSM guidance, feature clusters become more compact and clearly separable.

This enhanced discriminative power confirms that FSM improves the model's ability to extract category-consistent

features, thus enabling more accurate classification and better generalization in downstream tasks.

V. Conclusion

This study presents a novel signal enhancement framework termed Wavelet Dynamic Selection Network (WDSNet), which synergistically integrates wavelet-based signal processing with the adaptive learning capabilities of deep neural networks. The primary objective of WDSNet is to enhance the performance of low-cost inertial sensors by improving the quality and reliability of the raw sensor signals. By bridging the gap between model-driven and data-driven paradigms, WDSNet leverages the interpretability and robustness of wavelet transforms while exploiting the representational power and flexibility of deep learning models.

A central innovation introduced in this work is the Category Representation Mechanism (CRM), which enables the network to automatically learn and differentiate between the characteristics of various wavelet bases without requiring external selection labels. CRM acts as a plug-and-play module that enhances the discriminative capacity of deep learning architectures without increasing the number of trainable parameters. By learning compact and informative vector representations for each category of wavelets, the network gains a more intuitive understanding of the wavelet domain, thereby facilitating more informed wavelet selection during the enhancement process.

To further improve the quality of feature learning, we propose a Feature Supervision Mechanism (FSM) that utilizes the internal category vectors from CRM to guide the feature extraction process. Unlike conventional training methods that rely solely on output-level loss propagation, FSM offers a more efficient and localized supervision strategy by directly aligning intermediate features with semantically meaningful category vectors. Both theoretical analysis and empirical results confirm that FSM accelerates convergence, reduces error propagation, and improves overall network generalization.

Comprehensive evaluations were conducted to verify the effectiveness of WDSNet across both signal-level and task-level perspectives. Allan variance metrics were employed to assess the improvement in signal stability and noise suppression, revealing significant performance gains over conventional methods. Moreover, the enhanced signals were tested on four downstream applications—posture prediction, position estimation, semantic activity recognition, and spatial trajectory reconstruction—to validate the practical benefits of our approach. WDSNet consistently achieved state-of-the-art results across all evaluation tasks, demonstrating its ability to robustly improve sensor data quality in a variety of contexts.

A particularly noteworthy outcome is the successful reconstruction of arbitrary spatial trajectories using signals processed by WDSNet, a task traditionally deemed infeasible for low-cost IMU devices due to inherent noise and drift. This result highlights the transformative potential of the proposed framework in enabling high-precision applications using affordable hardware.

Finally, the wavelet selection strategy embedded within WDSNet is entirely data-driven and does not require any prior annotation or manual configuration. This makes the framework adaptable to diverse and dynamically evolving application environments, where conventional fixed or handcrafted methods would struggle to generalize. Overall, the proposed WDSNet provides a robust, scalable, and intelligent solution for signal enhancement in inertial sensing, paving the way for broader adoption of low-cost sensors in high-performance settings.

References

[1] D. K. Shaeffer, "Mems inertial sensors: A tutorial overview," *IEEE Communications Magazine*, vol. 51, no. 4, pp. 100–109, 2013.

[2] S. O. Madgwick, A. J. Harrison, and R. Vaidyanathan, "Estimation of imu and marg orientation using a gradient descent algorithm," in *2011 IEEE international conference on rehabilitation robotics*. IEEE, 2011, pp. 1–7.

[3] M. A. Esfahani, H. Wang, K. Wu, and S. Yuan, "Orinet: Robust 3-d orientation estimation with a single particular imu," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 399–406, 2019.

[4] Y. Li, R. Chen, X. Niu, Y. Zhuang, Z. Gao, X. Hu, and N. El-Sheimy, "Inertial sensing meets machine learning: Opportunity or challenge?" *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 9995–10 011, 2022.

[5] B. Gromov, G. Abbate, L. M. Gambardella, and A. Giusti, "Proximity human-robot interaction using pointing gestures and a wrist-mounted imu," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 8084–8091.

[6] M. Ehatisham-ul Haq, A. Arsalan, A. Raheel, and S. M. Anwar, "Expert-novice classification of mobile game player using smartphone inertial sensors," *Expert Systems with Applications*, vol. 174, p. 114700, 2021.

[7] L. Montesinos, R. Castaldo, and L. Pecchia, "Wearable inertial sensors for fall risk assessment and prediction in older adults: A systematic review and meta-analysis," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 26, no. 3, pp. 573–582, 2018.

[8] S. Liu, J. Zhang, Y. Zhang, and R. Zhu, "A wearable motion capture device able to detect dynamic motion of human limbs," *Nature communications*, vol. 11, no. 1, p. 5615, 2020.

[9] P. Li, W.-A. Zhang, Y. Jin, Z. Hu, and L. Wang, "Attitude estimation using iterative indirect kalman with neural network for inertial sensors," *IEEE Transactions on Instrumentation and Measurement*, 2023.

[10] C. Chen, X. Lu, A. Markham, and N. Trigoni, "Ionet: Learning to cure the curse of drift in inertial odometry," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, 2018.

[11] Y. Zhuang, X. Sun, Y. Li, J. Huai, L. Hua, X. Yang, X. Cao, P. Zhang, Y. Cao, L. Qi *et al.*, "Multi-sensor integrated navigation/positioning systems using data fusion: From analytics-based to learning-based approaches," *Information Fusion*, vol. 95, pp. 62–90, 2023.

[12] X. Zhang, B. He, G. Li, X. Mu, Y. Zhou, and T. Mang, "Navnet: Auv navigation through deep sequential learning," *IEEE Access*, vol. 8, pp. 59 845–59 861, 2020.

[13] M. Ferrera, V. Creuze, J. Moras, and P. Trouvé-Peloux, "Aqualoc: An underwater dataset for visual–inertial–pressure localization," *The International Journal of Robotics Research*, vol. 38, no. 14, pp. 1549–1559, 2019.

[14] T. Qin, P. Li, and S. Shen, "Vins-mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004–1020, 2018.

[15] D. Weber, C. Gühmann, and T. Seel, "Riann—a robust neural network outperforms attitude estimation filters," *Ai*, vol. 2, no. 3, pp. 444–463, 2021.

[16] M. A. Esfahani, H. Wang, K. Wu, and S. Yuan, "Aboldeepio: A novel deep inertial odometry network for autonomous vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 5, pp. 1941–1950, 2019.

[17] H. Nguyen, K. Lebel, S. Bogard, E. Goubault, P. Boissy, and C. Duval, "Using inertial sensors to automatically detect and segment activities of daily living in people with parkinson's disease," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 1, pp. 197–204, 2017.

[18] C. Tunca, G. Salur, and C. Ersoy, "Deep learning for fall risk assessment with inertial sensors: Utilizing domain knowledge in spatio-temporal gait parameters," *IEEE journal of biomedical and health informatics*, vol. 24, no. 7, pp. 1994–2005, 2019.

[19] P. Caserman, A. Garcia-Agundez, and S. Göbel, "A survey of full-body motion reconstruction in immersive virtual reality applications," *IEEE transactions on visualization and computer graphics*, vol. 26, no. 10, pp. 3089–3108, 2019.

[20] P. Foehn, E. Kaufmann, A. Romero, R. Penicka, S. Sun, L. Bauersfeld, T. Laengle, G. Cioffi, Y. Song, A. Loquercio *et al.*, "Agilicious: Open-source and open-hardware agile quadrotor for vision-based flight," *Science Robotics*, vol. 7, no. 67, p. eabl6259, 2022.

[21] X. Pan and A. F. d. C. Hamilton, "Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape," *British Journal of Psychology*, vol. 109, no. 3, pp. 395–417, 2018.

[22] C. Chen, P. Zhao, C. X. Lu, W. Wang, A. Markham, and N. Trigoni, "Deep-learning-based pedestrian inertial navigation: Methods, data set, and on-device inference," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4431–4441, 2020.

[23] Y. Wang and Y. Zhao, "Arbitrary spatial trajectory reconstruction based on a single inertial sensor," *IEEE Sensors Journal*, vol. 23, no. 9, pp. 10 009–10 022, 2023.

[24] S. S. Saha, S. S. Sandha, L. A. Garcia, and M. Srivastava, "Tinyodom: Hardware-aware efficient neural inertial navigation," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 6, no. 2, pp. 1–32, 2022.

[25] S. S. Saha, Y. Du, S. S. Sandha, L. A. Garcia, M. K. Jawed, and M. Srivastava, "Inertial navigation on extremely resource-constrained platforms: Methods, opportunities and challenges," in *2023 IEEE/ION Position, Location and Navigation Symposium (PLANS)*. IEEE, 2023, pp. 708–723.

[26] H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan, and O. Beijbom, "nuscenes: A multimodal dataset for autonomous driving," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 11 621–11 631.

[27] D. Engelsman and I. Klein, "Data-driven denoising of stationary accelerometer signals," *Measurement*, p. 113218, 2023.

[28] S. Herath, H. Yan, and Y. Furukawa, "Ronin: Robust neural inertial navigation in the wild: Benchmark, evaluations, & new methods," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 3146–3152.

[29] K. Yuan and Z. J. Wang, "A simple self-supervised imu denoising method for inertial aided navigation," *IEEE Robotics and Automation Letters*, vol. 8, no. 2, pp. 944–950, 2023.

[30] Y. Liu, G. Chen, Z. Wei, J. Yang, and D. Xing, "Denoising method of mems gyroscope based on interval empirical mode decomposition," *Mathematical Problems in Engineering*, vol. 2020, pp. 1–12, 2020.

[31] J. He, C. Sun, and P. Wang, "Noise reduction for mems gyroscope signal: A novel method combining acmp with adaptive multiscale sg filter based on ama," *Sensors*, vol. 19, no. 20, p. 4382, 2019.

[32] M. Brossard, S. Bonnabel, and A. Barrau, "Denoising imu gyroscopes with deep learning for open-loop attitude estimation," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4796–4803, 2020.

[33] A. K. Saydjari and D. P. Finkbeiner, "Equivariant wavelets: Fast rotation and translation invariant wavelet scattering transforms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 2, pp. 1716–1731, 2022.

[34] F. Wu, H. Luo, H. Jia, F. Zhao, Y. Xiao, and X. Gao, "Predicting the noise covariance with a multitask learning model for kalman filter-based gnss/ins integrated navigation," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–13, 2021.

[35] C. Toft, W. Maddern, A. Torii, L. Hammarstrand, E. Stenborg, D. Safari, M. Okutomi, M. Pollefeys, J. Sivic, T. Pajdla *et al.*, "Long-term visual localization revisited," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 4, pp. 2074–2088, 2020.

[36] Y. Wu, H.-B. Zhu, Q.-X. Du, and S.-M. Tang, "A survey of the research status of pedestrian dead reckoning systems based on inertial sensors," *International Journal of Automation and Computing*, vol. 16, pp. 65–83, 2019.

[37] H. Chen, T. M. Taha, and V. P. Chodavarapu, "Towards improved inertial navigation by reducing errors using deep learning methodology," *Applied Sciences*, vol. 12, no. 7, p. 3645, 2022.

[38] S. Han, Z. Meng, X. Zhang, and Y. Yan, "Hybrid deep recurrent neural networks for noise reduction of mems-imu with static and dynamic conditions," *Micromachines*, vol. 12, no. 2, p. 214, 2021.

[39] A. M. Boronakhin, R. V. Shalymov, D. Y. Larionov, N. Q. Khanh, and N. T. Yen, "Optimization of an inertial sensor de-noising method using a hybrid deep learning algorithm," in *2022 Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*. IEEE, 2022, pp. 1335–1340.

[40] E. J. Shamwell, K. Lindgren, S. Leung, and W. D. Nothwang, "Unsupervised deep visual-inertial odometry with online error correction for rgb-d imagery," *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, no. 10, pp. 2478–2493, 2019.

[41] X. Huang, P. Wang, X. Cheng, D. Zhou, Q. Geng, and R. Yang, "The apolloscape open dataset for autonomous driving and its application," *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, no. 10, pp. 2702–2719, 2019.

[42] H. Ma, D. Liu, N. Yan, H. Li, and F. Wu, "End-to-end optimized versatile image compression with wavelet-like transform," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 3, pp. 1247–1263, 2020.

[43] Y. Wang, Y. Wang, G. Hu, Y. Liu, and Y. Zhao, "Adaptive skewness kurtosis neural network : Enabling communication between neural nodes within a layer," in *International Conference on Neural Information Processing*, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:227075927

[44] F. Marchetti, F. Becattini, L. Seidenari, and A. Del Bimbo, "Multiple trajectory prediction of moving agents with memory augmented networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020.