HYPERSPECTRAL SELF-NAVIGATION DRONE

Chia-Hsiang Lin[1,2], *Qian-Hao Wang*[1,2], *and Jian-Kai Huang*[1,2]

[1]Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan
[2]Institute of Communications Engineering, National Cheng Kung University, Tainan, Taiwan
Email: chiahsiang.steven.lin@gmail.com; n20030325@gmail.com; q38141509@gs.ncku.edu.tw

ABSTRACT

Accurate localization is the fundamental prerequisite for UAV self-navigation. Although GPS is the primary solution, it may become unreliable in GPS-denied environments, for which localization typically relies on IMUs. However, IMU-based localization inevitably suffers from accumulated drift over time, leading to significant positioning errors. To alleviate this issue, RGB vision widely serves as complementary localization information. Nevertheless, RGB-based visual localization is subject to various limitations, including sensitivity to illumination variations and low-texture environments. By contrast, hyperspectral imagery (HSI) signals, remotely sensed/acquired by UAVs, offers rich spectral information that is inherently more robust. Counterintuitively, HSI-based *spectral* localization is computationally more efficient and requires significantly fewer network parameters than RGB-based *spatial* localization, better echoing the requirement of edge AI deployment on board UAVs. Motivated by these advantages, this paper investigates the feasibility of hyperspectral localization for UAV self-navigation (HyperNavi) and demonstrates its potential as a reliable alternative for localization in challenging environments.

***Index Terms*—** Unmanned Aerial Vehicle (UAV), localization, autonomous navigation, Global Positioning System (GPS), RGB vision localization, hyperspectral remote sensing, inertial measurement units (IMUs).

This study was supported by the Emerging Young Scholar Program (namely, the 2030 Cross-Generation Young Scholars Program) of National Science and Technology Council (NSTC), Taiwan, under Grant NSTC 115-2628-E-006-001. This study was also supported by the University Academic Alliance in Taiwan (UAAT), funded by the Ministry of Education (MoE). We thank the National Center for Theoretical Sciences (NCTS) and the National Center for High-performance Computing (NCHC) for providing the computing resources. *(Corresponding Author: Chia-Hsiang Lin)*

1. INTRODUCTION

UAVs have been applied across many fields due to their flexibility [1, 2], where a reliable self-navigation drone technique is essential, with accurate localization serving as a fundamental prerequisite. In general, localization is achieved using the GPS, which may become unreliable in GPS-denied scenarios [3]. Thus, IMUs are employed to provide continuous motion estimation [4, 5].

However, IMU-based localization inevitably suffers from accumulated drift over time [5]. To mitigate this issue, RGB vision is adopted to provide complementary localization information [6–8]. For example, AnyLoc [6] and Sample4Geo [7] serve as auxiliary modules in an IMU-based localization pipeline by intermittently matching RGB vision against a geo-referenced database to correct accumulated drift. Nevertheless, RGB vision localization is sensitive to the illumination variations and low-texture environments [9, 10]. Moreover, these methods [6, 7] primarily rely on global spatial information, thereby significantly increasing computational cost. On the other hand, quantum sensing-based navigation has recently demonstrated the potential in GPS-denied scenarios [11–13]. Particularly, quantum magnetometers enable magnetic-anomaly navigation by measuring subtle geomagnetic variations and matching real-time measurements against pre-built magnetic anomaly maps. Despite the high localization accuracy, it requires high-precision magnetic maps and specialized sensing equipment, hindering the near-term deployment. Meanwhile, quantum techniques have also shown significant potential in hyperspectral remote sensing [14–18].

In contrast, hyperspectral images (HSIs) acquired by UAVs, originally intended for remote object identification [19, 20], are inherently well-suited for UAV self-navigation. Specifically, a hyperspectral geo-referenced database can be constructed from existing airborne remote sensing data or acquired through a preliminary UAV mapping mission, without relying on anomaly

maps that have yet to be widely established. In addition, HSIs provide abundant spectral information and discriminative information in low-texture or visually similar scenes. Moreover, by exploiting the spectral discriminability, localization no longer requires large-scale spatial feature extraction networks and can be achieved with fewer parameters and lower computational cost, matching the real-time and lightweight requirements of onboard edge-AI systems [21].

Thus, this paper investigates the feasibility of HyperNavi for the first time. As shown in Table 1, HyperNavi achieves performance comparable to the state-of-the-art AnyLoc [6] with four orders of magnitude fewer parameters and three orders of magnitude fewer FLOPs.

2. THE PROPOSED METHOD

2.1. Problem Description

As shown in Fig. 1, given the pre-established hyperspectral geo-referenced database $\boldsymbol{Y}_G \in \mathbb{R}^{H\times W\times B}$ with spatial dimensions of $H \times W$ and B spectral bands, we first partition it into non-overlapping patches, each with a spatial size of $p \times p$. Subsequently, we construct the reference patch set $\mathcal{T}_R = \{\boldsymbol{Y}_{R_1}, \ldots, \boldsymbol{Y}_{R_{k^2}}\}$, where $\boldsymbol{Y}_{R_i} \in \mathbb{R}^{p\times p\times B}$ denotes the i-th reference patch, $\forall i = 1 \ldots k^2$. Specifically, $\mathcal{T}_R$ is established by collecting the $k \times k$ neighboring patches centered at the geo-referenced database patch nearest to either the query patch $\boldsymbol{Y}_Q \in \mathbb{R}^{p\times p\times B}$ captured by the UAV or the previous localization result, where $k = 5$ in this work. The spectral features of the query and reference patches are then extracted using the spectral encoding block (SEB). Finally, the symmetric Chamfer matching (SCM) [22] is used to compute the matching scores for localization. The proposed HyperNavi is mathematically defined as

$$\boldsymbol{Y}_E = \arg\max_{\boldsymbol{Y}_R\in\mathcal{T}_R} \mathrm{SCM}\left(\mathrm{SEB}(\boldsymbol{Y}_Q), \mathrm{SEB}(\boldsymbol{Y}_R)\right), \quad (1)$$

where $\boldsymbol{Y}_E \in \mathbb{R}^{p\times p\times B}$ represents the localization result. In the following part, we present the details of SEB and the localization using SCM.

2.2. Spectral Encoding Block

Given the reference patch set $\mathcal{T}_R$ and the query patch $\boldsymbol{Y}_Q$, the SEB extracts the spectral features using a shared Siamese network, which has been adopted in various applications [23]. By leveraging two identical subnetworks with shared parameters, the Siamese network effectively measures the similarity between a pair of inputs. Thus,

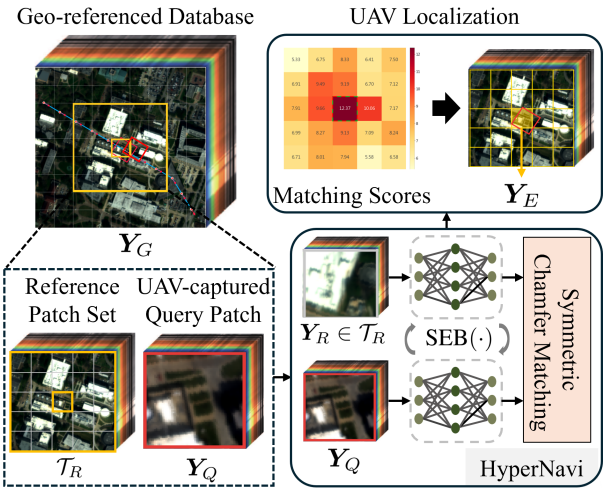


Fig. 1: Overview of HyperNavi architecture. Given the geo-referenced database $\boldsymbol{Y}_G$ and the UAV-captured query patch $\boldsymbol{Y}_Q$, we can establish the reference patch set $\mathcal{T}_R$. Then, $\boldsymbol{Y}_Q$ and $\boldsymbol{Y}_R \in \mathcal{T}_R$ are fed into HyperNavi, which is a Siamese network with two identical SEBs, for spectral feature extraction and matching score computation. Finally, the UAV localization result $\boldsymbol{Y}_E$ is determined as the reference patch with the highest matching score.

the Siamese architecture is well-suited for UAV localization, which aims to identify the reference patch that best matches the query patch captured by the UAV. As shown in Fig. 2, SEB$(\cdot)$ function is defined as

$$\mathrm{SEB}(\boldsymbol{Y}_Q) = \mathrm{Norm}\left(\mathrm{AvgPool}\left(f(\boldsymbol{Y}_Q)\right)\right), \quad (2)$$

where Norm, AvgPool, and $f(\cdot)$ represent the normalization layer, 2-D average pooling layer, and the learnable feature extraction function, respectively. In detail, the function $f(\cdot)$ comprises three consecutive blocks of identical structure, where each block integrates a 1×1 convolutional layer, LayerNorm, and a GELU activation function. For the spectral vector at each spatial location, the function $f(\cdot)$ maps the B-dimensional spectral input into feature representations of dimensions $d = 128$. To incorporate local spatial context while reducing dimensionality, a 4×4 average pooling layer aggregates the pixel-wise embeddings, followed by an ℓ_2 normalization layer to obtain $\mathrm{SEB}(\boldsymbol{Y}_Q) \in \mathbb{R}^{m\times m\times d}$, where $m = p/4$. Due to the nature of the Siamese network, $\mathrm{SEB}(\boldsymbol{Y}_R) \in \mathbb{R}^{m\times m\times d}$ can be derived in the same way.

2.3. Localization

Given $f_Q = \mathrm{SEB}(\boldsymbol{Y}_Q)$ and $f_R = \mathrm{SEB}(\boldsymbol{Y}_R)$, we employ the SCM [22] to measure the similarity between the SEB

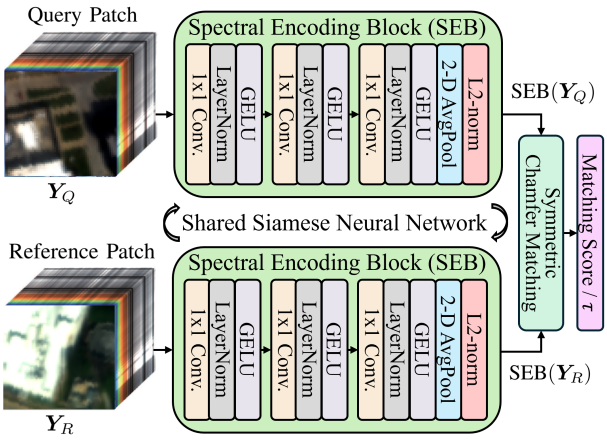


Fig. 2: Detailed architecture of the proposed HyperNavi.

extracted spectral tokens, which is defined as

$$\mathrm{SCM}\left(f_Q, f_R\right) = \frac{1}{2\tau}\left[\mathrm{CS}(\mathcal{S}_Q, \mathcal{S}_R) + \mathrm{CS}(\mathcal{S}_R, \mathcal{S}_Q)\right], \tag{3}$$

where τ denotes the temperature parameter that scales the similarity scores into logits, and $\mathrm{CS}(\cdot)$ denotes the one-directional Chamfer similarity [22]. Here, $\mathcal{S}_Q$ and $\mathcal{S}_R$ denote the spectral token sets derived from f_Q and f_R, respectively, with each set consisting of m^2 spectral tokens of dimension d. Then, $\mathrm{CS}(\cdot)$ is defined as

$$\mathrm{CS}(\mathcal{X}, \mathcal{Y}) = \frac{1}{|\mathcal{X}|}\sum_{\boldsymbol{t}\in\mathcal{X}} \max_{\boldsymbol{u}\in\mathcal{Y}} \langle \boldsymbol{t}, \boldsymbol{u}\rangle, \tag{4}$$

where $\mathcal{X}$ and $\mathcal{Y}$ represent the input spectral token sets, $|\mathcal{X}|$ denotes the cardinality of $\mathcal{X}$, and $\langle \boldsymbol{t}, \boldsymbol{u}\rangle$ denotes the cosine similarity between vectors $\boldsymbol{t}$ and $\boldsymbol{u}$. Finally, the localization result $\boldsymbol{Y}_E$ is identified as the reference patch corresponding to the highest SCM value, as shown in (1).

3. EXPERIMENTS

3.1. Datasets

In this work, we adopt a cross-sensor evaluation setup, where HSIs for training and testing are acquired from AVIRIS-NG [24] and CASI [25, 26] sensors, respectively. Specifically, the training data are derived from the NASA AVIRIS-NG Level-2 surface reflectance product [24], consisting of 27 flight strips in February 2022. Each AVIRIS-NG HSI contains 425 spectral bands spanning about 380–2510 nm, with a ground sampling distance of about 1.8 m. The testing data are obtained from the grss_dfc_2018 dataset [25, 26], acquired in February 2017 over the University of Houston campus. Each HSI in the dataset contains 48 spectral bands spanning 380–1050 nm with a spatial resolution of 1 m, whereas the corresponding RGB imagery has a spatial resolution of 5 cm. HyperNavi takes the CASI HSIs as input, while the RGB-based benchmarks use the corresponding RGB imagery. Due to differences between AVIRIS-NG and CASI configurations, the original AVIRIS-NG bands are spectrally resampled to match CASI's 48 bands.

Table 1: Quantitative comparisons on the grss_dfc_2018 Dataset [25, 26], where EP denotes error propagation.

EP	Metric	AnyLoc	Sample4Geo	HyperNavi
✓	R@1 (↑)	43.6	38.9	**77.9**
	R@5 (↑)	56.4	51.0	**100.0**
✗	R@1 (↑)	75.2	69.8	**77.9**
	R@5 (↑)	98.7	95.3	**100.0**
Time (ms)		132.3	62.8	**6.5**
Params (M)		1136	87.6	**0.11**
GFLOPs		598.81	90.24	**0.25**

3.2. Experimental Settings

In this work, we compare HyperNavi with RGB vision localization methods, including AnyLoc [6] and Sample4Geo [7]. The patch-wise and sequential localization protocol is adopted, where each query patch is first rotated according to the flight heading and then matched against reference patches. The R@1 and R@5 metrics [6] are used for evaluation. Since the overlap ratios between the query patch and the reference patches vary continuously, soft labels are constructed from these overlap ratios. Thus, HyperNavi is trained end-to-end as a k^2-way classification task using a soft cross-entropy loss function [27]. Both baselines follow their original configurations and publicly released weights. HyperNavi is optimized using AdamW with a learning rate of 5×10^{-4}, a weight decay of 5×10^{-4}, and a batch size of 16 for 60 epochs. The checkpoint with the best validation accuracy is used for evaluation. All experiments utilize Python 3.11.15 and PyTorch 2.5.1, conducted on a desktop featuring an Intel Core i9-10900X CPU and an NVIDIA GeForce RTX 2080 Ti GPU.

3.3. Qualitative and Quantitative Analysis

To ensure a fair comparison, a sinusoidal sweep trajectory with 149 query positions is generated. Given the effect of sequential error accumulation, the following two methods for initializing the search window are compared. The first is the setting *with* error propagation, where the search window is centered on the previously

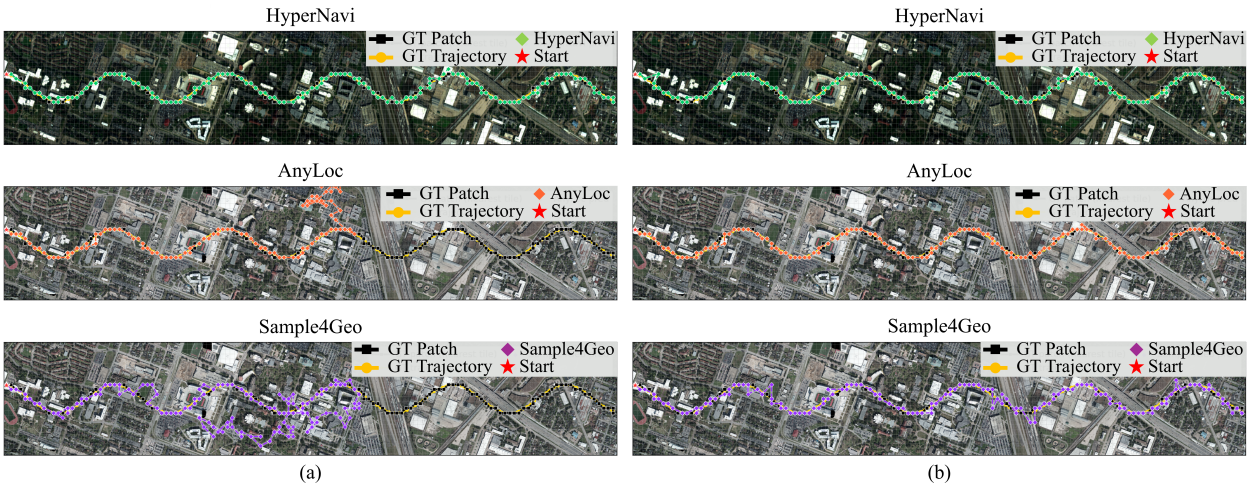


Fig. 3: UAV localization results under (a) with error propagation and (b) without error propagation conditions.

predicted position such that any localized error accumulates along the trajectory. The second is the setting *without* error propagation, in which the search window is centered on the current ground-truth patch, thereby providing an ideal coarse-position prior for evaluating single-step matching capability. As shown in Table 1, HyperNavi attains 77.9% R@1 and 100% R@5 in the error propagation setting, whereas the R@1 of AnyLoc and Sample4Geo drops to 43.6% and 38.9%, and their R@5 reaches only 56.4% and 51.0%, respectively. This shows that once an RGB method misses the target, the local search window fails to recover the ground truth, leading to persistent drift. In the setting without error propagation, despite the performance improvement across all methods, HyperNavi still outperforms AnyLoc and Sample4Geo on R@1 and R@5. The quantitative results indicate that the RGB methods still possess some single-step image-matching ability but are more susceptible to error accumulation during sequential localization. In addition, HyperNavi achieves the remarkable performance in milliseconds using four orders of magnitude fewer parameters and three orders of magnitude fewer FLOPs, demonstrating its potential as a reliable alternative for UAV localization in challenging environments. As shown in Fig. 3, the localization results of HyperNavi are closer to the reference trajectory than AnyLoc and Sample4Geo under both conditions, with and without error propagation.

4. CONCLUSION

For the first time, this paper investigates the feasibility of HyperNavi and demonstrates its potential in GPS-denied environments. Conventionally, RGB vision is used to compensate for the cumulative drift in IMU-based localization, but its performance degrades under low-texture conditions. On the other hand, quantum sensing-based localization has recently shown considerable promise in GPS-denied environments [12]. However, it requires a pre-established database of magnetic maps, which remains unavailable in the near term. In contrast, HSIs are well-suited for UAV self-navigation. Due to the rich spectral information, HSIs can reliably distinguish different materials even under low-illumination or low-texture conditions. Furthermore, unlike the database of magnetic maps, a hyperspectral geo-referenced database can be acquired from existing remote sensing data. Moreover, by prioritizing spectral feature extraction over complex spatial modeling, HyperNavi reduces model parameters by four orders of magnitude and GFLOPs by three orders of magnitude compared to the state-of-the-art RGB vision localization method (i.e., AnyLoc [6]), making it highly suitable for resource-constrained UAV edge platforms. In the future, to bridge current HyperNavi with future quantum sensing-based navigation, we will incorporate quantum neural networks (QUEENs) [14–18,28–30] into HyperNavi to learn quantum representations.

5. ACKNOWLEDGMENT

The authors would like to thank the National Center for Airborne Laser Mapping and the Hyperspectral Image Analysis Laboratory at the University of Houston for acquiring and providing the data used in this study, and the IEEE GRSS Image Analysis and Data Fusion Technical Committee.

6. REFERENCES

[1] Hazim Shakhatreh et al., "Unmanned aerial vehicles (UAVs): A survey on civil applications and key research challenges," *IEEE Access*, vol. 7, pp. 1–63, Apr. 2019.

[2] Mingyang Lyu et al., "Unmanned aerial vehicles for search and rescue: A survey," *Remote Sens.*, vol. 15, no. 13, pp. 1–35, Jun. 2023.

[3] Ni Zhu et al., "GNSS position integrity in urban environments: A review of literature," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 9, pp. 2762–2778, Sep. 2018.

[4] Yusen Wu et al., "UAV localization in GPS-denied environments: A review," *IEEE Trans. Syst., Man, Cybern., Syst.*, pp. 1–17, Jul. 2026.

[5] Manon Kok et al., "Using inertial sensors for position and orientation estimation," *Found. Trends Signal Process.*, vol. 11, no. 1, pp. 1–153, Nov. 2017.

[6] Nikhil Keetha et al., "AnyLoc: Towards universal visual place recognition," *IEEE Robot. Autom. Lett.*, vol. 9, no. 2, pp. 1286–1293, Feb. 2024.

[7] Fabian Deuser et al., "Sample4Geo: Hard negative sampling for cross-view geo-localisation," in *Proc. IEEE ICCV*, Paris, France, 1–6 Oct. 2023, pp. 16847–16856.

[8] Tong Qin et al., "Vins-mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Trans. Robot.*, vol. 34, pp. 1004–1020, Aug. 2018.

[9] Zhen Wang et al., "Sequence matching for image-based UAV-to-satellite geolocalization," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, pp. 1–15, Jan. 2024.

[10] Hu Zhao et al., "TransFG: A cross-view geo-localization of satellite and UAVs imagery pipeline using transformer-based feature aggregation and gradient guidance," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, pp. 1–12, Jan. 2024.

[11] Shiva Raj Pokhrel et al., "Quantum position navigation timing for autonomous vehicles: A Federated learning framework," *IEEE Trans. Veh. Technol.*, vol. 75, pp. 12697–12706, Feb. 2026.

[12] Olga Sambataro et al., "Current trends and advances in quantum navigation for maritime applications: A comprehensive review," *IEEE J. Ocean. Eng.*, vol. 50, pp. 2101–2134, Jul. 2025.

[13] Mia Jukić et al., "Quantum sensing for magnetic-aided navigation in GPS-denied environments," *Quantum Technol. Def. Secur.*, vol. 13202, pp. 23–37, Nov. 2024.

[14] Chia-Hsiang Lin et al., "HyperQUEEN: Hyperspectral quantum deep network for image restoration," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1–20, May 2023.

[15] Chia-Hsiang Lin et al., "HyperKING: Quantum-classical generative adversarial networks for hyperspectral image restoration," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, pp. 1–19, Apr. 2025.

[16] Po-Wei Tang et al., "A quantum-empowered SPEI drought forecasting algorithm using spatially-aware mamba network," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, pp. 1–18, Jun. 2025.

[17] Chia-Hsiang Lin et al., "Quantum feature-empowered deep classification for fast mangrove mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, pp. 1–13, 2025.

[18] Chia-Hsiang Lin et al., "Quantum information-empowered graph neural network for hyperspectral change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, pp. 1–15, Nov. 2024.

[19] Chia-Hsiang Lin et al., "Signal subspace identification for incomplete hyperspectral image with applications to various inverse problems," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, pp. 1–16, Mar. 2024.

[20] Chia-Hsiang Lin et al., "A fast hyperplane-based minimum-volume enclosing simplex algorithm for blind hyperspectral unmixing," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 1946–1961, Dec. 2015.

[21] Sifat Rezwan et al., "Artificial intelligence approaches for UAV navigation: Recent advances and future challenges," *IEEE Access*, vol. 10, pp. 26320–26339, Mar. 2022.

[22] Giorgos Kordopatis-Zilos et al., "Visil: Fine-grained spatio-temporal video similarity learning," in *Proc. IEEE ICCV*, Seoul, Korea, 27 Oct.–2 Nov. 2019, pp. 1–10.

[23] Chia-Hsiang Lin et al., "CODE-MM: Convex deep mangrove mapping algorithm based on optical satellite images," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1–19, Sep. 2023.

[24] R. O. Green et al., "AVIRIS-NG L2 surface reflectance, facility instrument collection, V1," 2023.

[25] "2018 IEEE GRSS Data Fusion Contest," [Online]. Available: `https://www.grss-ieee.org/community/technical-committees/data-fusion`.

[26] Yonghao Xu et al., "Advanced multi-sensor optical remote sensing for urban land use and land cover classification: Outcome of the 2018 IEEE GRSS data fusion contest," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 6, pp. 1709–1724, 2019.

[27] Xiang Li et al., "Generalized focal loss: Learning qualified and distributed bounding boxes for dense object detection," in *Proc. NeurIPS*, Virtual Event, 6–12 Dec. 2020, pp. 21002–21012.

[28] Chia-Hsiang Lin et al., "Quantum-driven multihead inland waterbody detection with transformer-encoded cygnss delay-doppler map data," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, pp. 1–21, Nov. 2025.

[29] Chia-Hsiang Lin et al., "Underdetermined blind source separation via weighted simplex shrinkage regularization and quantum deep image prior," *IEEE Trans. Image Process.*, vol. 35, pp. 3069–3084, Mar. 2026.

[30] Chia-Hsiang Lin et al., "Hyperspectral anomaly detection using einstein fuzzy computing and quantum neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 64, pp. 1–20, May 2026.