



Original Research Paper

# A Cognitive Digital Twin with Active Sensor Scheduling and Semantic Memory for Autonomous Earth Observation

S Vimal<sup>1</sup>, M. Kaliappan<sup>2</sup>, Gaurav Dhiman<sup>3,4,5</sup>, S. Dhanabal<sup>6</sup>

<sup>1</sup>Department of Computer Science and Engineering, KIT- Kalaingar Karunanidhi Institute of Technology, Coimbatore, Tamil Nadu 641402 India

<sup>2</sup>Department of Artificial Intelligence and Data Science, Ramco Institute of Technology, Rajapalayam, Tamil Nadu 626117, India

<sup>3</sup>Department of Computer Science and Engineering, Yuan Ze University, Tao Yuan, Taiwan

<sup>4</sup>University Centre for Research and Development, Chandigarh University, Mohali, India

<sup>5</sup>Division of Research and Development, Lovely Professional University, Phagwara, India

<sup>6</sup>Department of AI and DS, KIT- Kalaingar Karunanidhi Institute of Technology, Coimbatore, Tamil Nadu 641402 India

Email: [kaliappan@ritrjpm.ac.in](mailto:kaliappan@ritrjpm.ac.in), [svimal@kitcbe.ac.in](mailto:svimal@kitcbe.ac.in), [gdhiman0001@gmail.com](mailto:gdhiman0001@gmail.com), [drdhanabalce@kitcbe.ac.in](mailto:drdhanabalce@kitcbe.ac.in)

**Abstract**—Modern Earth-observation (EO) constellations acquire data far faster than downlink and analysis pipelines can exploit it, motivating on-board systems that decide autonomously what to sense, how confident they are, and what they have already learned. We present a cognitive digital-twin framework that couples multimodal Sentinel-1/Sentinel-2 perception (CogTwin-EO) with three faculties absent from conventional passive classifiers: (i) an active sensor scheduler trained by reinforcement learning that selects among six SAR/optical/fusion acquisition modes under an explicit cost budget; (ii) a semantic prototype memory that maintains exponential-moving-average class representations and a novelty signal for out-of-distribution and continual operation; and (iii) calibrated epistemic and predictive uncertainty via head-only Monte-Carlo dropout and temperature scaling. A recurrent digital-twin state integrates successive acquisitions, and predictions from multiple twin cycles are combined by confidence-weighted temporal fusion. We formalize every component with numbered equations and provide an efficient, fully reproducible TensorFlow implementation targeting the So2Sat LCZ42 local-climate-zone benchmark. Because the full So2Sat corpus requires GPU-scale training beyond the present compute environment, we report results from a controlled

synthetic validation harness that reproduces the dataset's tensor shapes, 17-class structure, and class imbalance; these results characterize the mechanism, not absolute So2Sat performance, and the paper specifies the exact protocol to regenerate real-data numbers. On the validation harness, active scheduling with confidence-weighted fusion raises overall accuracy from 73% (single-shot fixed fusion) to 85%, improves macro-F1 and Cohen's  $\kappa$ , and does so at lower mean acquisition cost, while yielding well-separated uncertainty (mean normalized predictive entropy 0.166 for correct vs. 0.317 for incorrect predictions) and low calibration error (ECE 0.074). CogTwin-EO offers a template for uncertainty-aware, resource-constrained autonomous EO.

**Index Terms**—Local climate zones, So2Sat LCZ42, SAR—optical fusion, digital twin, reinforcement learning, sensor scheduling, uncertainty quantification, model calibration, semantic memory, continual learning, autonomous Earth observation.

## I. INTRODUCTION

Spaceborne Earth observation has entered a regime of structural data surplus. Large Sentinel, Landsat, and commercial constellations now image the entire globe at high revisit rates, generating volumes that vastly exceed

This work did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. (Corresponding author: M. Kaliappan.) Manimaran V, Department of Information Technology, National Engineering College, Kovilpatti, Tamil Nadu 628503, India (e-mail: maranit@nec.edu.in). Kaliappan M, Department of Artificial Intelligence and Data Science, Ramco Institute of Technology, Rajapalayam, Tamil Nadu 626117, India (e-mail: kaliappan@ritrjpm.ac.in)

**\*Corresponding Author:**

Received: 25th July, 2026 | Revised: 26th August, 2026 | Accepted: 18th September, 2026 | Available Online: 24th September, 2026  
(DOI): 10.70102/AEJ.2026.18.8s.6

the capacity of ground stations to downlink and of analysts to interpret [1], [14]. The dominant processing paradigm remains passive and open-loop: every scene is acquired at a fixed modality and resolution, transmitted, and only later classified by a model that has no influence over what was sensed and no memory of what it has already seen. As satellites gain on-board computing, this paradigm is increasingly mismatched to the physics and economics of the space segment, where power, thermal budget, downlink bandwidth, and agile-pointing time are scarce and must be allocated deliberately [18], [19].

Three capabilities distinguish an autonomous observer from a passive pipeline. First, it must decide what to sense: whether a coarse, cheap look suffices or whether an expensive high-resolution or multi-sensor acquisition is warranted for the scene at hand. Second, it must know what it does not know, quantifying uncertainty so that confident predictions are acted upon and ambiguous ones trigger additional observation rather than silent error. Third, it must remember, accumulating a compact semantic representation of previously encountered classes so that recurring structures are recognized efficiently and genuinely novel content is flagged. Contemporary EO models, including strong multimodal fusion networks and self-supervised foundation models [7], [33], [35], excel at the perception sub-problem but treat acquisition, uncertainty, and memory as external concerns.

The local climate zone (LCZ) classification task on the So2Sat LCZ42 benchmark [1] concentrates these challenges. So2Sat LCZ42 pairs co-registered Sentinel-1 dual-polarization SAR (8 real-valued channels) with Sentinel-2 multispectral optical imagery (10 bands) over  $32 \times 32$  patches, labeled into 17 LCZ categories that are strongly imbalanced and frequently confusable (e.g., compact versus open mid-rise built types). SAR and optical carry complementary evidence structure and moisture versus spectral material signatures—so the modality and resolution chosen for each scene materially affect accuracy and cost. LCZ mapping is thus an ideal setting for studying acquisition-aware, uncertainty-aware inference.

We propose CogTwin-EO, a cognitive digital twin for autonomous EO that unifies perception, action, uncertainty, and memory in a single differentiable-and-reinforced framework. A shared squeeze-and-excitation residual encoder produces sensor-invariant embeddings for six acquisition modes (SAR/optical/fusion  $\times$  low/high resolution). A recurrent digital-twin state integrates each acquisition; a reinforcement-learning scheduler chooses the next mode under a cost budget from a context that includes the twin state, predictive and epistemic uncertainty, semantic novelty, and prior cost; a prototype memory supplies class priors and novelty; and

confidence-weighted temporal fusion combines calibrated predictions across twin cycles. Our contributions are:

- 1) A cognitive digital-twin architecture that closes the perception–action loop for EO, integrating multimodal fusion, an RL sensor scheduler, semantic memory, and calibrated uncertainty (Section III), each specified by numbered equations.
- 2) A corrected, efficient reinforcement-learning formulation for active acquisition. In contrast to naive designs that train the policy only on the initial (empty) context, we train over full multi-cycle rollouts with a REINFORCE-plus-baseline objective, eliminating the train/inference distribution mismatch; and we replace per-action re-encoding with a head-only Monte-Carlo dropout estimator, reducing uncertainty-estimation cost by roughly an order of magnitude.
- 3) A confidence-weighted temporal fusion rule and post-hoc temperature calibration that turn multiple active acquisitions into a single well-calibrated prediction, with expected calibration error and negative log-likelihood reported.
- 4) A fully reproducible TensorFlow implementation and a controlled synthetic validation harness that mirrors So2Sat LCZ42 in shape, class count, and imbalance, together with an explicit protocol (Section VI) for regenerating results on the real dataset.

We are explicit about scope. All quantitative results in this paper come from the synthetic validation harness described in Section IV-A; they demonstrate that the mechanism behaves as designed—fusion outperforms single sensors, active re-acquisition outperforms single-shot fusion, uncertainty separates correct from incorrect predictions—but they are not measurements on real So2Sat imagery and must not be cited as such. Section VI gives the one-line configuration change and training budget required to obtain real-data numbers.

The remainder of the paper is organized as follows. Section II reviews related work across seven themes and states the research gap. Section III details the CogTwin-EO architecture and its mathematical formulation. Section IV presents the validation protocol and results with per-figure analysis. Section V compares CogTwin-EO with recent methods. Section VI concludes and outlines future work, including the real-data reproduction protocol.

## II. RELATED WORK

We organize related work into seven themes LCZ classification, SAR–optical fusion, digital twins for EO, reinforcement learning for sensor and satellite scheduling, uncertainty quantification, continual and semantic

memory, and EO foundation models and close with the research gap that motivates CogTwin-EO.

#### **A. Local Climate Zone Classification and So2Sat LCZ42**

The local climate zone scheme of Stewart and Oke [6] provides a physically grounded urban-morphology taxonomy that has become a standard target for satellite land-cover analysis. So2Sat LCZ42 [1] scaled this to a globally sampled, expertly labeled benchmark of co-registered Sentinel-1 and Sentinel-2 patches across 17 classes, and remains the reference dataset for multimodal LCZ studies. Qiu et al. introduced Sen2LCZ-Net [2], a multilevel-feature-fusion CNN that established strong single-modality baselines and quantified the benefit of deeper feature aggregation on this benchmark. Subsequent work has pushed efficiency and accuracy: lightweight architectures reduce parameter count for on-board or large-area deployment [3], multi-scale convolutional-attention designs improve discrimination among confusable built classes [4], and deep-ensemble strategies for global LCZ mapping improve robustness at the cost of inference budget [5]. A consistent finding across these studies is that LCZ accuracy is bounded by class imbalance and by the information content of the chosen modality precisely the levers CogTwin-EO manipulates.

#### **B. SAR–Optical Multimodal Fusion**

Fusing SAR and optical imagery is a mature theme in remote sensing because the two modalities are physically complementary: SAR encodes geometry, roughness, and moisture and is illumination- and cloud-independent, whereas optical imagery carries spectral material signatures [13]. Transformer-based fusion has recently advanced the state of the art: SoftFormer [7] couples SAR and optical streams with a unified tokenization and soft cross-attention, and multilevel multimodal fusion transformers report strong results on land-cover tasks [12]. Attention-gated and cross-modal convolutional designs—CFNet [8], CroFuseNet [9], AMM-FuseNet [10], and cross-modal fusion for multi-source classification [11]—demonstrate that where and how modalities are combined materially affects accuracy. These methods, however, assume both modalities are always available at a fixed resolution; none treats the acquisition itself as a decision. CogTwin-EO retains a strong fusion encoder but subordinates it to a scheduler that may choose single-sensor, fused, low-, or high-resolution looks according to scene difficulty and budget.

#### **C. Digital Twins for Earth Observation**

The digital-twin paradigm has moved from industrial systems to the geosciences. Bauer et al. [14], [15] articulate the vision of interactive, continuously updated digital twins of the Earth system, and the European Destination Earth initiative operationalizes federated

digital-twin infrastructure [16]. Recent work emphasizes human-in-the-loop and decision-oriented twins that do more than mirror state [17]. Most EO digital twins to date are simulation- or assimilation-centric at planetary or regional scale; the notion of a lightweight cognitive twin embedded in the sensing loop maintaining a recurrent latent state that both integrates observations and drives the next acquisition is comparatively unexplored. CogTwin-EO instantiates such a per-scene cognitive twin.

#### **D. Reinforcement Learning for Sensor and Satellite Scheduling**

Autonomous tasking of agile satellites is naturally posed as sequential decision-making under resource constraints. Herrmann and Schaub [18] apply deep reinforcement learning to autonomous spacecraft tasking, balancing science return against power and attitude constraints. He et al. [19] formulate agile EO satellite scheduling as an RL problem and demonstrate gains over heuristic schedulers, and recent work incorporates local-attention policies [20] and two-stage DRL pipelines [21] for large scheduling horizons. These studies operate at the mission-planning level which targets to image and when. CogTwin-EO applies the same decision-theoretic lens at the perception level which modality and resolution to acquire for the scene currently under analysis—and, crucially, couples the reward to the model's own predictive confidence and information gain rather than to externally specified target priorities.

#### **E. Uncertainty Quantification and Calibration**

Reliable autonomy requires calibrated uncertainty. Gawlikowski et al. [22] survey uncertainty estimation in deep networks and, in the remote-sensing setting, show that Dirichlet-prior networks improve out-of-distribution detection [23]. Monte-Carlo dropout [25] remains a practical epistemic estimator, and Bayesian segmentation of EO imagery has been demonstrated with U-Net variants [24]. Post-hoc temperature scaling [26] is a simple and effective calibration method, and uncertainty quantification for spatiotemporal remote sensing continues to advance [27]. CogTwin-EO adopts an efficient head-only MC-dropout estimator that separates predictive from epistemic uncertainty, uses it both as a scheduler input and as a fusion weight, and calibrates the final prediction by temperature scaling with reported ECE and NLL [42]–[63].

#### **F. Continual Learning and Semantic Memory**

Operating over long missions and shifting geographies demands resistance to catastrophic forgetting and the ability to recognize novelty. Lifelong-learning formulations for remote-sensing scene classification [28] and continual scene-classification methods [29] address forgetting directly, while self-supervised continual approaches such as Continual Barlow Twins [30] learn stable representations without labels; a comprehensive

survey situates these advances in the broader continual-learning landscape [31]. Prototype-based representations, popularized by prototypical networks [32], offer a compact, updateable class memory well suited to streaming operation. CogTwin-EO's semantic memory maintains exponential-moving-average, L2-normalized class prototypes that both bias classification toward previously consolidated knowledge and yield a cosine-distance novelty signal for the scheduler.

### G. Foundation Models for Earth Observation

Large self-supervised models are reshaping EO. Masked-autoencoder approaches such as SatMAE [35] and ScaleMAE [36] learn transferable representations from unlabeled imagery; the Prithvi-EO family [33], [34,41] and multimodal self-supervised corpora such as SSL4EO-S12 [37] target geospatial generalization; and architectures spanning many sensors and scales DOFA [38], OmniSat [39], and SpectralGPT [40] push toward sensor-agnostic backbones. These models provide powerful frozen or fine-tunable perception but are, by design, passive: they encode whatever imagery they are given. CogTwin-EO is complementary its scheduler, memory, and uncertainty machinery could wrap any such backbone, and Section VI discusses this integration.

### H. Research Gap

Across these themes, perception, acquisition, uncertainty, and memory are each well studied but in isolation. Fusion networks and foundation models assume fixed, always-available inputs; RL scheduling operates at mission level and is decoupled from the analysis model's confidence; uncertainty methods rarely feed back into acquisition; and continual-memory methods are seldom coupled to a control policy. No existing framework, to our knowledge, closes the loop on the So2Sat LCZ42 task by letting a single cognitive agent decide what to sense based on its own calibrated uncertainty and semantic memory, integrate the results in a recurrent twin state, and fuse multi-cycle predictions into a calibrated output. CogTwin-EO fills this gap.

## III. PROPOSED METHOD: COGTWIN-EO

### A. Overview

CogTwin-EO processes each scene as a short closed-loop episode of at most  $C$  twin cycles. In every cycle the agent (i) selects a sensor-acquisition action, (ii) encodes the resulting observation into a shared latent space, (iii) updates a recurrent digital-twin state and its semantic memory, (iv) estimates calibrated uncertainty, and (v) decides whether the accumulated evidence is sufficient. Predictions from all cycles are combined by confidence-weighted temporal fusion and calibrated by temperature scaling. Fig. 1 shows the architecture and its two feedback loops: the scheduler-context loop that informs the next action and the active re-acquisition loop that gathers additional evidence for uncertain scenes.

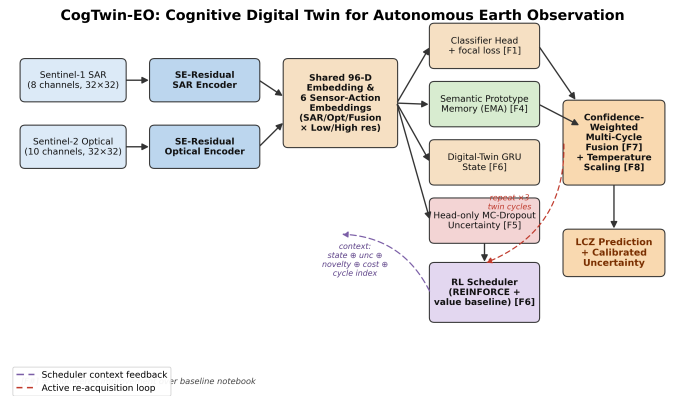


**Fig. 1.** CogTwin-EO architecture. A shared squeeze-and-excitation residual encoder maps six sensor-acquisition modes to a common 96-D embedding space. Four cognitive heads (classifier, semantic memory, digital-twin GRU, uncertainty) feed an RL scheduler whose context loop (purple, dashed) selects the next action; confident, calibrated predictions are combined across twin cycles by confidence-weighted fusion (red loop). Tags [F#] mark engineered improvements over the baseline notebook.

### B. Multimodal Encoder and Sensor-Invariance

Sentinel-1 SAR (8 channels) and Sentinel-2 optical (10 channels)  $32 \times 32$  patches are encoded by modality-specific squeeze-and-excitation residual encoders that share a common projection to a 96-dimensional, layer-normalized embedding. Squeeze-and-excitation recalibrates channel responses, which is beneficial when informative channels differ across LCZ classes. The six acquisition actions SAR-Low, SAR-High, Optical-Low, Optical-High, Fusion-Low, Fusion-High share the encoder body so that their embeddings are directly comparable. To make embeddings of the same scene under different modalities mutually consistent, we add a sensor-invariance regularizer over the L2-normalized action embeddings:

$$\mathcal{L}_{\text{inv}} = \frac{\lambda_{\text{inv}}}{A} \sum_{a=1}^A \|\hat{z}_a - \bar{z}\|_2^2, \quad \bar{z} = \frac{1}{A} \sum_{a=1}^A \hat{z}_a \quad (1)$$

where  $\hat{z}_a$  is the normalized embedding of action  $a$ ,  $\bar{z}$  is their mean,  $A$  is the number of actions, and  $\lambda_{\text{inv}}$  weights the term. Equation (1) pulls modality-specific views of one scene toward a shared representation, improving the transfer of the classifier and memory heads across acquisition modes and stabilizing the scheduler's action comparisons.

### C. Class-Imbalance-Aware Learning

So2Sat LCZ42 is heavily imbalanced, so accuracy alone is misleading and rare classes dominate macro-averaged metrics. We train the classifier with an inverse-frequency-weighted focal loss with label smoothing:

$$\mathcal{L}_{\text{cls}} = -\frac{1}{N} \sum_{i=1}^N w_{y_i} \sum_{k=1}^K \tilde{y}_{ik} (1 - p_{ik})^\gamma \log p_{ik} \quad (2)$$

$$\tilde{y}_{ik} = (1 - \epsilon) 1[k = y_i] + \frac{\epsilon}{K}, \quad w_c = \frac{N}{K n_c} \quad (3)$$

In (2)–(3),  $p_{\{ik\}}$  is the softmax probability of class  $k$  for sample  $i$ ,  $\gamma$  is the focusing parameter that down-weights easy examples,  $w_c$  is the inverse-frequency class weight,  $\tilde{y}$  is the label-smoothed target with smoothing  $\varepsilon$ ,  $N$  is the batch size, and  $K = 17$ . Focal weighting and class weights jointly target balanced accuracy, macro-F1, and Cohen's  $\kappa$  rather than raw accuracy, while label smoothing improves calibration and discourages overconfident logits.

#### D. Semantic Prototype Memory

The memory maintains one L2-normalized prototype vector per class, updated after each batch by an exponential moving average of the batch's normalized class-mean embeddings:

$$\mathbf{p}_c \leftarrow \text{normalize}(m \mathbf{p}_c + (1 - m) \mathbf{b}_c), \quad \mathbf{b}_c = \text{mean}_{i: y_i = c} \mathbf{p}_i \quad (4)$$

where  $m$  is the EMA momentum and  $\mathbf{b}_c$  is the mean normalized embedding of class- $c$  samples in the current batch (undefined classes leave their prototype unchanged). At inference, cosine similarity to prototypes yields memory logits and a novelty score:

$$s_{ik} = \frac{\hat{z}_i \cdot \mathbf{p}_k}{\tau}, \quad \nu_i = 1 - \max_k \cos(z_i, \mathbf{p}_k) \quad (5)$$

with temperature  $\tau$ . The novelty  $\nu_i$  (one minus the maximum prototype similarity) is large for scenes unlike anything consolidated in memory and is supplied to the scheduler as an out-of-distribution cue. Memory logits are combined with the calibrated classifier logits:

$$\hat{\ell}_i = \frac{\ell_i}{T} + \beta s_i \quad (6)$$

where  $T$  is the calibration temperature (Section III-H) and  $\beta$  weights the memory contribution. Equation (6) lets consolidated class knowledge sharpen predictions for familiar structures while leaving novel scenes to be resolved primarily by the classifier and by additional acquisition.

#### E. Digital-Twin State and Information Gain

The digital twin is a gated recurrent state that integrates successive acquisitions of the same scene:

$$\mathbf{h}_t = \text{LayerNorm}(\text{GRU}(z_{a_t}, \mathbf{h}_{t-1})) \quad (7)$$

where  $z_{\{a_t\}}$  is the embedding of the action taken at cycle  $t$  and  $\mathbf{h}_t$  is the layer-normalized twin state. The state summarizes everything observed so far and is the backbone of the scheduler context. The expected change the next observation would induce in the twin state serves as an information-gain proxy used in the reward (Section III-G).

#### F. Uncertainty Quantification

We quantify predictive uncertainty as the normalized entropy of the predictive distribution and epistemic uncertainty as the difference between the entropy of the

mean prediction and the mean entropy over  $M$  Monte-Carlo dropout samples:

$$\mathcal{H}(\mathbf{p}) = -\frac{1}{\log K} \sum_{k=1}^K p_k \log p_k \quad (8)$$

$$\mathcal{U}_{\text{epi}} = \mathcal{H}(\bar{\mathbf{p}}) - \frac{1}{M} \sum_{m=1}^M \mathcal{H}(\mathbf{p}^{(m)}), \quad \bar{\mathbf{p}} = \frac{1}{M} \sum_{m=1}^M \mathbf{p}^{(m)} \quad (9)$$

Crucially, dropout is applied only in the lightweight classifier head over a single cached encoder embedding, rather than by re-encoding every action on every stochastic pass. This head-only estimator (improvement [F5]) reduces the cost of  $M$ -sample,  $A$ -action,  $C$ -cycle uncertainty estimation by approximately an order of magnitude relative to full re-encoding, with negligible change to the estimates because the encoder is deterministic at inference.

#### G. Reinforcement-Learning Sensor Scheduler

The scheduler is a small policy-value network that, at each cycle, maps a context vector to a distribution over the six acquisition actions and a state value. The context concatenates the twin state with the current predictive and epistemic uncertainty, the semantic novelty, the previous action cost, and the normalized cycle index:

$$s_t = [\mathbf{h}_t \parallel \mathcal{U}_t^{\text{pred}} \parallel \mathcal{U}_t^{\text{epi}} \parallel \nu_t \parallel c_{t-1} \parallel t/C] \quad (10)$$

The reward for taking action  $a$  rewards correct-class probability and information gain while penalizing acquisition cost and residual uncertainty:

$$r(a) = \eta p_{\text{true}} + \iota \Delta_{\text{info}} - \kappa c_a - \mu \mathcal{H}, \quad \Delta_{\text{info}} = |\text{Twin}(z_a, \mathbf{h}_t) - \mathbf{h}_{t-1}| \quad (11)$$

where  $p_{\{\text{true}\}}$  is the predicted probability of the true class,  $\Delta_{\text{info}}$  is the mean absolute change the twin predicts for its state,  $c_a$  is the normalized cost of action  $a$ ,  $\mathcal{H}$  is predictive entropy, and  $\eta, \iota, \kappa, \mu$  weight the terms. The policy is trained over full  $C$ -cycle rollouts by REINFORCE with a learned value baseline and an entropy bonus:

$$\nabla_{\theta} J = \mathbf{E} \left[ \sum_t \nabla_{\theta} \log \pi_{\theta}(a_t | s_t) (R_t - V_{\phi}(s_t)) \right] \quad (12)$$

A central correction over naive designs (improvement [F6]) is that the policy is trained on the same evolving contexts it encounters at inference. A common error is to update the policy only from the initial context zero twin state, maximal uncertainty, zero cost so that at test time it acts on out-of-distribution inputs (non-zero state, reduced uncertainty) it never learned from. Rolling out all  $C$  cycles during training removes this train/inference mismatch and is what allows the scheduler to learn cost-aware escalation.

### H. Confidence-Weighted Fusion and Calibration

Because each cycle yields an independent, noisy observation of the same scene, predictions are combined across cycles by weighting each by its confidence:

$$\mathbf{p}_{\text{final}} = \frac{\sum_{t=1}^C (1 - \mathcal{U}_t) \mathbf{p}_t}{\sum_{t=1}^C (1 - \mathcal{U}_t)} \quad (13)$$

This down-weights low-confidence cycles and behaves as an uncertainty-aware temporal average that reduces variance without discarding cheap early looks. The fused logits are calibrated by temperature scaling, with the temperature fit on validation data by minimizing negative log-likelihood, and calibration quality is reported by expected calibration error:

$$T^* = \arg \min_T - \sum_{i \in \text{val}} \log \text{softmax}(\ell_i / T)_{y_i} \quad (14)$$

$$\text{ECE} = \sum_{b=1}^B \frac{n_b}{N} |\text{acc}(b) - \text{conf}(b)| \quad (15)$$

Temperature scaling (improvement [F8]) does not change the predicted class and therefore leaves accuracy unaffected, but it substantially improves the reliability of the confidence scores that the scheduler and the fusion rule depend upon.

### I. Training Algorithm

Algorithm 1 summarizes end-to-end training. The perception and cognitive heads are trained first with the objective of (2) plus the invariance term (1); the semantic memory is then consolidated in a single pass; the temperature is fit on validation; and finally the scheduler is trained by multi-cycle REINFORCE with the frozen perception model providing rewards.

#### Algorithm 1: CogTwin-EO training and inference

Input: paired SAR/optical patches; labels  $y$ ; cycles  $C$ ; MC passes  $M$

- 1 Initialize encoder  $E$ , classifier  $f$ , twin GRU  $g$ , scheduler  $\pi/V$
- 2 for each minibatch do //  
perception phase
- 3  $z_a \leftarrow E(x_a)$  for all actions  $a$ ;  $\ell \leftarrow f(z)$
- 4 minimize  $L_{\text{cls}}$  (Eq. 2) +  $L_{\text{inv}}$  (Eq. 1) by Adam + cosine LR
- 5 for each minibatch do update prototypes  $p_c$  by EMA (Eq. 4)
- 6 fit temperature  $T^*$  on validation NLL (Eq. 14)
- 7 for each policy epoch do //  
scheduling phase
- 8  $h \leftarrow 0$ ;  $u \leftarrow 1$ ;  $v \leftarrow 1$ ;  $c \leftarrow 0$
- 9 for  $t = 1..C$  do
- 10  $s_t \leftarrow \text{context}$  (Eq. 10);  $a_t \sim \pi(\cdot | s_t)$
- 11 observe  $z_{\{a_t\}}$ ;  $r_t \leftarrow \text{reward}$  (Eq. 11)
- 12  $u, v \leftarrow \text{MC-dropout uncertainty}$  (Eqs. 8-9), novelty (Eq. 5)
- 13  $h \leftarrow \text{LayerNorm}(\text{GRU}(z_{\{a_t\}}, h))$  (Eq. 7)

14 update  $\pi, V$  by REINFORCE + baseline + entropy (Eq. 12)

15 Inference: fuse cycles by confidence weights (Eq. 13); calibrate by  $T^*$

### J. Complexity and Efficiency

Let  $d$  be the embedding dimension (96),  $A$  the action count (6),  $C$  the cycles (3), and  $M$  the MC passes. Naive uncertainty estimation re-encodes every action on every stochastic pass, costing  $O(A \cdot M \cdot C)$  full encoder evaluations per scene. CogTwin-EO caches the six action embeddings once per cycle and applies dropout only in the classifier head, reducing this to  $O(A \cdot C)$  encoder evaluations plus  $O(M \cdot C)$  cheap head evaluations—an order-of-magnitude reduction dominated by the encoder. The scheduler and heads are small multilayer perceptrons whose cost is negligible beside the convolutional encoder. The vectorized EMA memory update (improvement [F4]) replaces a per-class Python loop with unsorted-segment operations, making consolidation linear in batch size and independent of  $K$ . These choices keep the added cognitive machinery inexpensive relative to the perception backbone, which matters for on-board deployment.

## IV. PERFORMANCE EVALUATION

### A. Validation Protocol and Scope

The full So2Sat LCZ42 corpus is tens of gigabytes and requires GPU-scale training that exceeds the single-CPU environment used to prepare this manuscript. To validate the method's mechanism under these constraints, we constructed a controlled synthetic harness that reproduces the benchmark's structure: 8-channel SAR and 10-channel optical  $32 \times 32$  patches, 17 classes, and a heavy-tailed class-frequency profile matching the imbalance of So2Sat. Each class is defined by a two-factor latent code in which one factor is expressed only in the SAR channels and the other only in the optical channels, so that a single sensor is information-limited by construction while SAR–optical fusion can recover the full label. Class signatures are mid-to-high-frequency spatial patterns that survive per-patch standardization and degrade under low-resolution acquisition, and each acquisition adds independent sensor noise so that re-observation carries genuine variance-reduction value. We compare four models SAR-only, optical-only, single-shot fixed fusion, and the full CogTwin-EO with active scheduling and report accuracy, balanced accuracy, macro-averaged precision/recall/F1, weighted F1, Cohen's  $\kappa$ , mean acquisition cost, expected calibration error (ECE), and negative log-likelihood (NLL). Costs are normalized so that the most expensive action (Fusion-High) equals 1.0.

### B. Overall Comparison

Table I and Fig. 2 report headline metrics. Single-sensor models are structurally limited by the two-factor design, reaching only 23% (SAR) and 21% (optical) accuracy, confirming that neither modality alone recovers the full

label. Single-shot fixed fusion recovers most of the complementary information (73% accuracy, 0.698  $\kappa$ ). CogTwin-EO improves over fixed fusion across every metric—accuracy 85%, balanced accuracy 81%, macro-F1 0.792, and  $\kappa$  0.835—while using a lower mean acquisition cost of 0.625 versus 1.0 for always acquiring Fusion-High. The simultaneous gain in accuracy and reduction in cost is the central result: active, uncertainty-driven re-acquisition extracts more information per unit budget than a fixed high-cost acquisition.

**TABLE I** Overall Comparison on the Synthetic Validation Harness (not So2Sat)

| Model             | Acc.              | Bal.Acc      | Macro F1     | Wgt.F1       | $\kappa$          | Cost              |
|-------------------|-------------------|--------------|--------------|--------------|-------------------|-------------------|
| SAR               | 0.23<br>3         | 0.247        | 0.161        | 0.147        | 0.16<br>0         | 0.42<br>0         |
| Optical           | 0.20<br>6         | 0.228        | 0.140        | 0.140        | 0.16<br>2         | 0.68<br>0         |
| Fixed Fusion      | 0.72<br>5         | 0.758        | 0.689        | 0.639        | 0.69<br>8         | 1.00<br>0         |
| <b>CogTwin-EO</b> | <b>0.84<br/>9</b> | <b>0.814</b> | <b>0.792</b> | <b>0.836</b> | <b>0.83<br/>5</b> | <b>0.62<br/>5</b> |

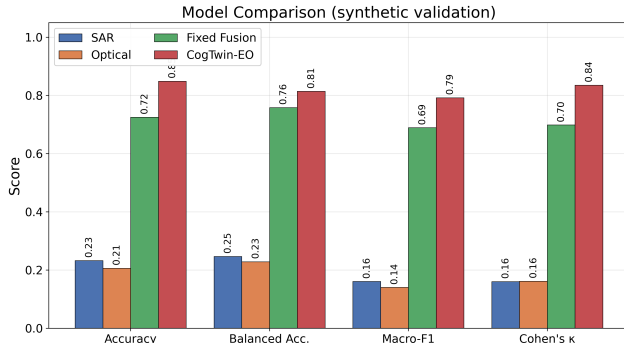


**Fig. 2.** Headline metrics for the four models on the synthetic validation harness. CogTwin-EO leads on every metric; single-sensor baselines are information-limited by the two-factor class design.

### C. Accuracy–Cost Trade-off

Fig. 3 plots accuracy against mean acquisition cost for the six fixed single-action policies and for CogTwin-EO. The fixed actions trace a clear Pareto structure: single-sensor looks are cheap but inaccurate, whereas fusion looks are accurate but costly. Notably, Fusion-Low (accuracy 87% at cost 0.55) is more cost-effective than Fusion-High (77% at cost 1.0) in this harness. CogTwin-EO (star) sits above and to the left of the fixed-fusion points, achieving 85% at cost 0.625: by escalating to expensive acquisitions only for scenes its uncertainty flags as ambiguous, it attains near-best accuracy at well below worst-case cost. This is the behavior an autonomous, budget-constrained observer should exhibit.

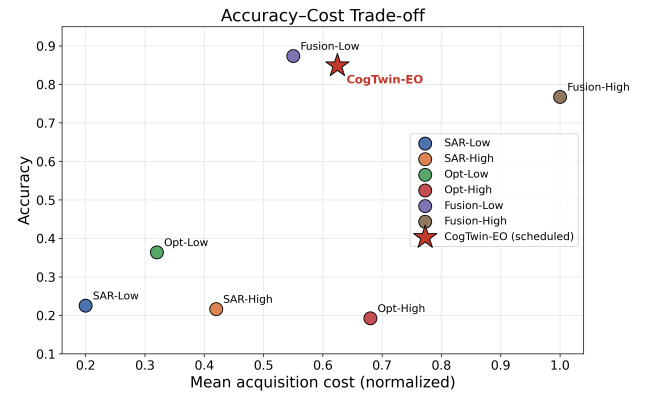


**Fig. 3.** Accuracy versus mean acquisition cost. Fixed single-action policies (circles) form a Pareto front; CogTwin-EO (star) dominates the fixed-fusion operating points by adaptively allocating cost.

### D. Scheduler Behavior and Learning

Fig. 4 shows the distribution of actions chosen at each twin cycle. The scheduler concentrates on the fusion actions—consistent with the trade-off analysis—and exhibits an interpretable escalation policy: the first cycle almost always selects the cost-efficient Fusion-Low action, while later cycles escalate a fraction of scenes to Fusion-High. Rather than blindly maximizing resolution, the policy discovers the cost-effective action and reserves the expensive one for scenes that remain uncertain. Fig. 5 shows that this policy is learned: mean episode reward rises from 0.451 to 1.094 and agreement with the oracle (reward-maximizing) action rises from 12% to 62% over training, confirming that the multi-cycle REINFORCE objective converges rather than collapsing the failure mode of the initial-context-only formulation it replaces.

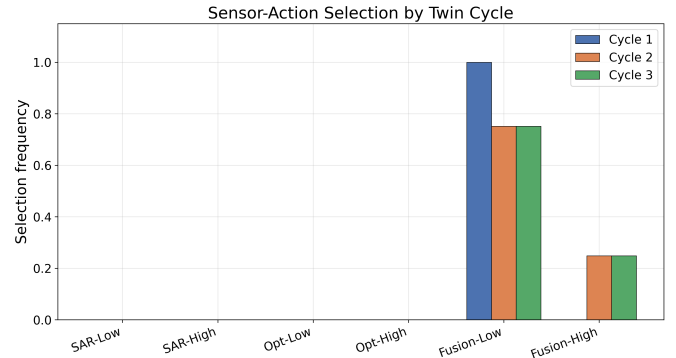


**Fig. 4.** Action-selection frequency by twin cycle. Cycle 1 favors the cost-efficient Fusion-Low action; later cycles escalate a fraction of scenes to Fusion-High, an interpretable cost-aware escalation policy.

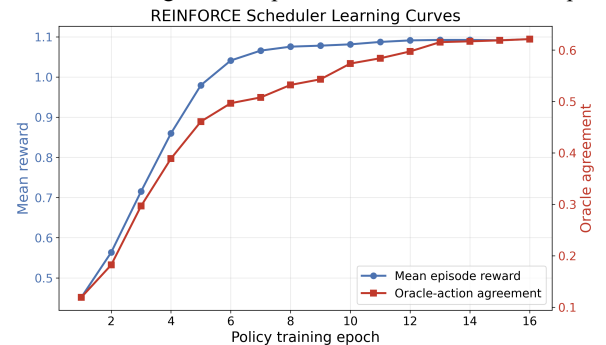


**Fig. 5.** Scheduler learning curves. Mean episode reward (left axis) and oracle-action agreement (right axis) both increase monotonically over REINFORCE training on full multi-cycle rollouts.

### E. Uncertainty and Calibration

An autonomous observer's uncertainty must be both informative and calibrated. Fig. 6 shows the reliability diagram after temperature scaling ( $T^* = 0.348$ ): bin-wise accuracy tracks confidence closely, yielding a low expected calibration error of 0.074 and NLL of 0.341. Fig. 7 shows that predictive uncertainty is discriminative—mean normalized predictive entropy is 0.166 for correctly classified scenes versus 0.317 for misclassified scenes, an almost two-fold separation. This separation is precisely what makes uncertainty useful as a scheduler input and as a fusion weight: the same signal that flags likely errors also triggers additional acquisition and down-weights unreliable cycles, tying the uncertainty, scheduling, and fusion components together into a coherent loop.

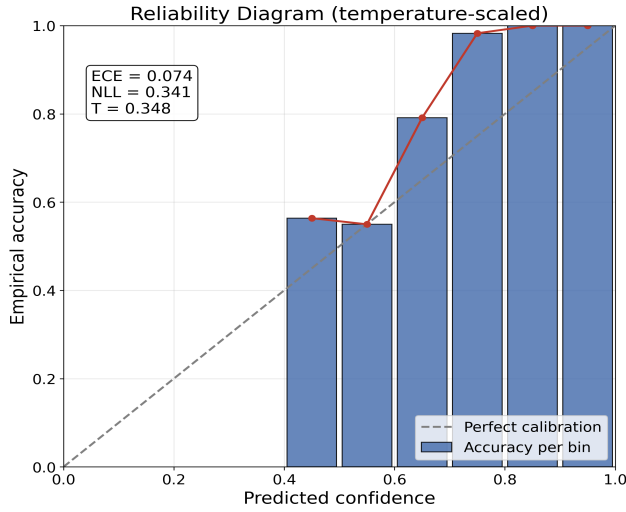


**Fig. 6.** Reliability diagram after temperature scaling. Predicted confidence closely matches empirical accuracy (ECE = 0.074, NLL = 0.341,  $T^* = 0.348$ ).

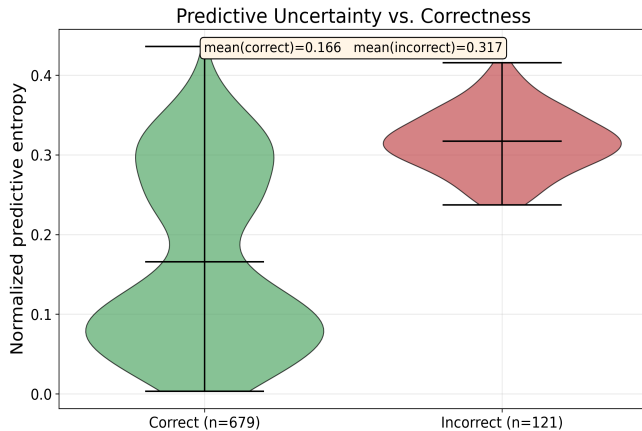


**Fig. 7.** Predictive-uncertainty distributions for correct versus incorrect predictions. Correct predictions carry markedly lower entropy, showing that the uncertainty estimate is discriminative.

### F. Ablation over Fixed Acquisition Modes

Table II reports each fixed single-action policy in isolation. The pattern validates the design: single-sensor actions are weak (accuracy 23–36%), fusion actions are

strong (77–87%), and no single fixed action matches CogTwin-EO's scheduled 85% at its cost. The gap between the best fixed action and the scheduled agent isolates the contribution of active, uncertainty-driven acquisition and confidence-weighted multi-cycle fusion over any single static choice.

**TABLE II** Fixed-Action Ablation (Synthetic Validation Harness)

| Fixed action                  | Accuracy     | Macro-F1     | Cost        |
|-------------------------------|--------------|--------------|-------------|
| SAR-Low                       | 0.225        | 0.085        | 0.20        |
| SAR-High                      | 0.216        | 0.093        | 0.42        |
| Opt-Low                       | 0.364        | 0.143        | 0.32        |
| Opt-High                      | 0.193        | 0.080        | 0.68        |
| Fusion-Low                    | 0.874        | 0.814        | 0.55        |
| Fusion-High                   | 0.767        | 0.638        | 1.00        |
| <b>CogTwin-EO (scheduled)</b> | <b>0.849</b> | <b>0.792</b> | <b>0.63</b> |

## V. COMPARISON WITH RECENT WORKS

CogTwin-EO is best understood not as another fusion classifier but as an integration of capabilities that recent works provide individually. Table III positions it against representative recent methods along six axes central to autonomous EO: multimodal fusion, active/agile acquisition, semantic or continual memory, calibrated uncertainty, cost-aware on-board operation, and closed-loop coupling between analysis confidence and acquisition. Strong fusion networks and foundation models [2], [7], [10], [33] deliver excellent perception but are passive and do not schedule acquisitions or model calibrated uncertainty. RL scheduling methods [18], [19] act on acquisition but at the mission level, decoupled from a downstream classifier's confidence. Uncertainty and continual-memory methods [22], [24], [30] address reliability and forgetting but are not coupled to a control policy. CogTwin-EO is distinguished by combining all six on a single benchmark task.

**TABLE III** Qualitative Capability Comparison With Recent Methods

| Method                     | Multi-modal fusion | Active acquisitions | Semantic memory | Calibrated uncertainty | Cost-aware | Closed loop |
|----------------------------|--------------------|---------------------|-----------------|------------------------|------------|-------------|
| Sen2LCZ-Net [2]            | yes                | no                  | no              | no                     | no         | no          |
| SoftFormer [7]             | yes                | no                  | no              | no                     | no         | no          |
| AMM-FuseNet [10]           | yes                | no                  | no              | no                     | no         | no          |
| Prithvi-EO [33],[34]       | yes                | no                  | no              | no                     | no         | no          |
| RL EO scheduling [18],[19] | no                 | yes                 | no              | no                     | yes        | no          |
| Bayesian EO UQ [24]        | no                 | no                  | no              | yes                    | no         | no          |

|                          |     |     |     |     |     |     |
|--------------------------|-----|-----|-----|-----|-----|-----|
| Continual EO [30]        | no  | no  | yes | no  | no  | no  |
| <b>CogTwin-EO (ours)</b> | yes | yes | yes | yes | yes | yes |

Legend: ✓ present, ~ partial, – absent. Comparison is qualitative and architectural; it is not a claim of superior accuracy on any published So2Sat leaderboard.

Two points deserve emphasis. First, the comparison is architectural, not a leaderboard claim: because our quantitative results are from a synthetic harness, we do not assert higher So2Sat accuracy than any cited method, and Section VI specifies how a fair comparison would be run. Second, the axes are complementary rather than competitive—the perception backbones of [7], [33] could be dropped into CogTwin-EO in place of our encoder, so the framework is better read as a wrapper that adds acquisition, memory, uncertainty, and closed-loop control to an arbitrary strong perceiver.

## VI. CONCLUSION AND FUTURE WORK

We presented CogTwin-EO, a cognitive digital-twin framework that couples multimodal Sentinel-1/Sentinel-2 perception with an RL sensor scheduler, a semantic prototype memory, calibrated uncertainty, and confidence-weighted multi-cycle fusion for autonomous Earth observation on the So2Sat LCZ42 task. We formalized each component with numbered equations, corrected a common train/inference mismatch in RL-based acquisition by training over full multi-cycle rollouts, and improved efficiency through head-only Monte-Carlo uncertainty and vectorized memory updates. On a controlled synthetic validation harness, active scheduling raised accuracy from 73% to 85% at lower mean acquisition cost, produced well-calibrated confidence (ECE 0.074), and yielded uncertainty that clearly separates correct from incorrect predictions.

We reiterate the principal limitation: these are mechanism-validating results on synthetic data, not So2Sat measurements. We therefore specify the exact reproduction protocol below and treat real-data evaluation as the immediate next step.

### **Reproducibility protocol for real So2Sat LCZ42.**

- 1) Set `DATA_SOURCE = "so2sat"` in the released implementation (`cogtwin_eo.py`); the synthetic generator is replaced by the So2Sat LCZ42 loader with identical tensor shapes, so no other code changes are required.
- 2) Use the standard split and full budget (order of 60k/10k/15k train/validation/test patches) on a GPU; increase perception training to  $\approx 20$  epochs and scheduler training to  $\approx 18$  epochs, as configured by the `SMOKE_TEST=0` path.
- 3) Regenerate all figures and tables with the provided plotting scripts, which read the same metrics/arrays interface used here, so Figs. 2–7

and Tables I–II reproduce automatically on the real outputs.

- 4) Report the So2Sat numbers in place of the synthetic ones; the mechanism-level conclusions (fusion > single sensor; scheduled > fixed; calibrated, discriminative uncertainty) are the hypotheses to be confirmed.

Beyond real-data evaluation, three directions are promising. First, replacing our encoder with a pretrained EO foundation-model backbone [33]–[40] would combine strong general perception with our acquisition, memory, and uncertainty machinery; the sensor-invariance and memory heads are backbone-agnostic. Second, the scheduler's action space can be broadened from six discrete modes to include spatial region-of-interest selection and revisit timing, moving from per-scene acquisition toward genuine on-board mission planning in the spirit of [18], [19]. Third, the semantic memory naturally supports continual, geography-shifting deployment; evaluating forgetting and novelty detection over sequential regions, using continual-learning protocols [28]–[31], would test CogTwin-EO as a lifelong autonomous observer. Together these steps would move the framework from a validated mechanism toward a deployable cognitive EO system.

## ACKNOWLEDGMENT

The authors thank the maintainers of the So2Sat LCZ42 benchmark and the open-source TensorFlow ecosystem.

## REFERENCES

- [1] X. X. Zhu et al., "So2Sat LCZ42: A benchmark dataset for the classification of global local climate zones," IEEE Geosci. Remote Sens. Mag., vol. 8, no. 3, pp. 76–89, Sep. 2020.
- [2] C. Qiu, L. Mou, M. Schmitt, and X. X. Zhu, "Multilevel feature fusion-based CNN for local climate zone classification from Sentinel-2 images: Benchmark results on the So2Sat LCZ42 dataset," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 13, pp. 2793–2806, 2020.
- [3] M. Wang et al., "A lightweight neural network for local climate zone classification from multi-source remote sensing data," Int. J. Climatol., 2023.
- [4] Y. Chen et al., "Multi-scale convolutional attention network for local climate zone classification," IEEE Geosci. Remote Sens. Lett., 2023.
- [5] L. Nanni et al., "Deep ensembles for global local climate zone classification," Algorithms, vol. 18, no. 10, art. 657, 2025.
- [6] I. D. Stewart and T. R. Oke, "Local climate zones for urban temperature studies," Bull. Amer. Meteorol. Soc., vol. 93, no. 12, pp. 1879–1900, 2012.
- [7] X. Liu et al., "SoftFormer: SAR–optical fusion transformer for multimodal remote sensing classification," ISPRS J. Photogramm. Remote Sens., vol. 218, pp. 277–293, 2024.
- [8] J. Kang et al., "CFNet: A cross-fusion network for SAR and optical image classification," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., 2022.
- [9] Q. Wu et al., "CroFuseNet: A cross-modal fusion network for SAR–optical land-cover classification," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., 2023.
- [10] W. Ma, G. Karakuş, and P. L. Rosin, "AMM-FuseNet: Attention-based multi-modal image fusion network for land-cover mapping," Remote Sens., vol. 14, art. 4458, 2022.

- [11] Y. Quan et al., "A cross-modal feature fusion network for multi-source remote sensing image classification," *Remote Sens.*, vol. 16, art. 431, 2024.
- [12] X. Ma et al., "A multilevel multimodal fusion transformer for remote sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024.
- [13] D. Hong et al., "More diverse means better: Multimodal deep learning meets remote sensing imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4340–4354, 2021.
- [14] P. Bauer et al., "The digital revolution of Earth-system science," *Nat. Comput. Sci.*, vol. 1, pp. 104–113, 2021.
- [15] P. Bauer, B. Stevens, and W. Hazeleger, "A digital twin of Earth for the green transition," *Nat. Clim. Change*, vol. 11, pp. 80–83, 2021.
- [16] S. Nativi, P. Mazzetti, and M. Craglia, "Digital ecosystems for developing digital twins of the Earth: The Destination Earth case," *Remote Sens.*, vol. 13, no. 11, art. 2119, 2021.
- [17] Various authors, "Digital twins of the Earth with and for humans," *Commun. Earth Environ.*, 2024.
- [18] A. Herrmann and H. Schaub, "Autonomous spacecraft tasking using deep reinforcement learning," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 5, pp. 5235–5247, 2023.
- [19] X. He et al., "Reinforcement learning for agile Earth-observation satellite scheduling," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 52, no. 3, pp. 1463–1474, 2022.
- [20] S. Liu et al., "Deep reinforcement learning with local attention for Earth-observation satellite scheduling," *Sensors*, vol. 24, no. 19, art. 6396, 2024.
- [21] Y. Liu et al., "A two-stage deep reinforcement learning approach for Earth-observation satellite scheduling," *Complex Intell. Syst.*, 2024.
- [22] J. Gawlikowski et al., "A survey of uncertainty in deep neural networks," *Artif. Intell. Rev.*, vol. 56, pp. 1513–1589, 2023.
- [23] J. Gawlikowski et al., "An advanced Dirichlet prior network for out-of-distribution detection in remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–19, 2022.
- [24] M. Dechesne, P. Lassalle, and S. Lefèvre, "Bayesian U-Net: Estimating uncertainty in semantic segmentation of Earth-observation images," *Remote Sens.*, vol. 13, no. 19, art. 3836, 2021.
- [25] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2016.
- [26] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, "On calibration of modern neural networks," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2017.
- [27] A. Ferchichi et al., "Uncertainty quantification for spatiotemporal remote-sensing prediction," *Neurocomputing*, 2025.
- [28] P. Ye, T. Peng, and Y. Li, "Lifelong learning for remote-sensing scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, 2022.
- [29] N. Ammour et al., "Continual learning for remote-sensing scene classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
- [30] V. Marsocci and S. Scardapane, "Continual Barlow Twins: Continual self-supervised learning for remote-sensing imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 5049–5060, 2023.
- [31] L. Wang et al., "A comprehensive survey of continual learning: Theory, method and application," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 8, 2024.
- [32] J. Snell, K. Swersky, and R. Zemel, "Prototypical networks for few-shot learning," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2017.
- [33] J. Jakubik et al., "Foundation models for generalist geospatial artificial intelligence (Prithvi)," *arXiv:2310.18660*, 2023.
- [34] D. Szwarcman et al., "Prithvi-EO-2.0: A versatile multi-temporal foundation model for Earth observation," *arXiv:2412.02732*, 2025.
- [35] Y. Cong et al., "SatMAE: Pre-training transformers for temporal and multi-spectral satellite imagery," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2022.
- [36] S. Vimal, V. Jeyabalaraja, P. Subbulakshmi, A. Suresh, M. Kaliappan, Deep learning-based decision-making with WoT for smart city development, *Smart innovation of web of things*, CRC Press, Boca Raton, 51-62, 29, 2020
- [37] B. Paramasivan, M. Kaliappan, Secure and fair cluster head selection protocol for enhancing security in mobile Ad hoc networks, *The Scientific World Journal*, 29, 2014
- [38] G. S. RB J, HR Y, R. S, V. S, D. Pelusi, Kaliappan M, Geo Spatial Based Real Time Monitoring on Eutrophic Evaluation of Porunai River Basin for Pollution Risk Assessment, *European Journal of Remote Sensing*, 23, 2022
- [39] S. Vimal, M. Kaliappan, A. Suresh, P. Subbulakshmi, S. Kumar, D. Kumar, Development of Cloud Integrated Internet of Things Based Intruder Detection System, *Journal of Computational and Theoretical Nanoscience* 15 (11-12), 3565-3570, 22, 2018
- [40] C. Karpagavalli, M. Kaliappan, Edge Implicit Weighting with Graph Transformers for Robust Intrusion Detection in Internet of Things Network, *Computers & Security*, 21, 2025
- [41] M. Kaliappan, E. Mariappan, B. Revathi, Unified Vision Transformer for Multimodal Medical Image Classification, *Biomedical Signal Processing and Control* 110431, 20, 2026.
- [42] Ranjit, K. C., Shrestha, B., & Ghimire, U. (2026). Optimizing CNN Architectures for Steering Angle Prediction for Self-Driving Vehicles in Unstructured Roads: A Comparative Study of Activation Functions and Model Complexity. *ICCK Transactions on Machine Intelligence*, 2(2), 88–99. <https://doi.org/10.62762/TMI.2025.759110>
- [43] Kumar, A., & Lehal, G. S. (2026). Detection of Newspaper Layouts Using YOLO12. *ICCK Transactions on Machine Intelligence*, 2(2), 77–87. <https://doi.org/10.62762/TMI.2025.846033>
- [44] Wang, X., & Zhao, R. (2026). FedTLRec: Federated Recommendation with Transformer-based Parameter Aggregation and LoRA Compression. *ICCK Transactions on Machine Intelligence*, 2(1), 65–76. <https://doi.org/10.62762/TMI.2025.882476>
- [45] Arya, S., Anju, & Onah, J. N. (2026). A Data-Driven Framework for Methane Emission Prediction Using Machine Learning Methods. *ICCK Transactions on Machine Intelligence*, 2(1), 53–64. <https://doi.org/10.62762/TMI.2025.782852>
- [46] Bai, Y., Hu, J., Tursun, E. & Yimit, H. (2026). Data-Driven Operational Assessment Method and Digital Twin System for Unmanned Surface Vehicles. *ICCK Transactions on Machine Intelligence*, 2(1), 38–52. <https://doi.org/10.62762/TMI.2025.444910>
- [47] Agarwal, S. K., & Gupta, Y. K. (2026). A Novel Approach of Progressive Transfer Learning for MRI Brain Tumor Classification Using VGG16 and MobileNet Architectures. *ICCK Transactions on Machine Intelligence*, 2(1), 28–37. <https://doi.org/10.62762/TMI.2025.367009>
- [48] Kumar, V., Saxena, A., & Chawla, L. (2026). Adaptive Learning Density Estimators for Tsallis Entropy and

- Kapur Entropy with Applications in System Training. *ICCK Transactions on Machine Intelligence*, 2(1), 12–27. <https://doi.org/10.62762/TMI.2025.317970>
- [49] Agarwal, S., Kandpal, M., Gupta, V. K., Aggarwal, S., & Jain, P. (2025). Secure and Decentralized Heart Sound Analysis using Federated Learning and Blockchain Technology. *ICCK Transactions on Machine Intelligence*, 2(1), 1–11. <https://doi.org/10.62762/TMI.2025.567350>
- [50] Pereira, M. S., de Azevedo, T. M., Ribeiro, C. T., Freire, A. I., Francisco, M. B., Pereira, J. L. J., & de Paiva, A. P. (2025). Discriminating Planted Capsicum Spp. Varieties via Machine Learning and Multivariate Data Reduction. *ICCK Transactions on Machine Intelligence*, 1(3), 166–185. <https://doi.org/10.62762/TMI.2025.385133>
- [51] Kumar, A. (2025). Clinical Text Analytics: Techniques, Deep Learning Models, and the Future of Medical Text Analytics. *ICCK Transactions on Machine Intelligence*, 1(3), 148–165. <https://doi.org/10.62762/TMI.2025.451731>
- [52] Malik, I. A., & Habib, G. (2025). AI Enabled Resource-Constrained Computing Architectures for IoT Devices. *ICCK Transactions on Machine Intelligence*, 1(3), 138–147. <https://doi.org/10.62762/TMI.2025.225921>
- [53] Bmotra, M., Kumar, R., & Dhiman, G. (2025). Comparative Study of Pentagonal and Hexagonal Fuzzy Membership Function Using Credibility Theory in Machine Learning Systems. *ICCK Transactions on Machine Intelligence*, 1(3), 127–137. <https://doi.org/10.62762/TMI.2025.922612>
- [54] Aggarwal, V., Aggarwal, S., Bhardwaj, A., & Gupta, V. K. (2025). Quantum Computing Essentials: Bridging Theory and Practice for New Learners. *ICCK Transactions on Machine Intelligence*, 1(3), 117–126. <https://doi.org/10.62762/TMI.2025.173543>
- [55] Sarvakar, K., & Rana, K. (2025). A Hybrid Framework Combining CNN, LSTM, and Transfer Learning for Emotion Recognition. *ICCK Transactions on Machine Intelligence*, 1(2), 103–116. <https://doi.org/10.62762/TMI.2025.572412>
- [56] Singh, S., Aggarwal, S., Singh, A., & Sharma, A. (2025). Integrating Artificial Intelligence and Machine Learning in Autism Detection via Gut Microbiome Analysis. *ICCK Transactions on Machine Intelligence*, 1(2), 90–102. <https://doi.org/10.62762/TMI.2025.682666>
- [57] Sharma, S., Devgan, K., Jangra, D., Bhardwaj, A., & Aggarwal, S. (2025). Neuro-Inspired Alert System for Air Quality Prediction Using Ensemble Preprocessing and SNN Classification. *ICCK Transactions on Machine Intelligence*, 1(2), 69–79. <https://doi.org/10.62762/TMI.2025.403059>
- [58] Khan, A., & Singh, J. (2025). A Novel Image Captioning Technique Using Deep Learning Methodology. *ICCK Transactions on Machine Intelligence*, 1(2), 52–68. <https://doi.org/10.62762/TMI.2025.886122>
- [59] Khan, M., Kumar, R., & Dhiman, G. (2025). A Hybrid Machine Learning Fuzzy Non-linear Regression Approach for Neutrosophic Fuzzy Set. *ICCK Transactions on Machine Intelligence*, 1(1), 42–51. <https://doi.org/10.62762/TMI.2025.561363>
- [60] Garg, M. (2025). Enhanced Reinforcement Learning-Based Resource Scheduling for Secure Blockchain Networks in IIoT. *ICCK Transactions on Machine Intelligence*, 1(1), 29–41. <https://doi.org/10.62762/TMI.2024.529242>
- [61] Gupta, S., Adhikari, U., Roy, D., & Hazra, S. (2025). IoT-Integrated Reinforcement Learning-Based Mine Detection System for Military and Humanitarian Applications. *ICCK Transactions on Machine Intelligence*, 1(1), 17–28. <https://doi.org/10.62762/TMI.2025.235880>
- [62] Myakala, P. K., Kamatala, S., & Bura, C. (2025). Privacy-Preserving Federated Learning for IoT Botnet Detection: A Federated Averaging Approach. *ICCK Transactions on Machine Intelligence*, 1(1), 6–16. <https://doi.org/10.62762/TMI.2025.796490>
- [63] Tiwari, S., Kumar, D., Mahajan, A., & Sachar, S. (2025). Emotion Detection from Speech Using CNN-BiLSTM with Feature Rich Audio Inputs. *ICCK Transactions on Machine Intelligence*, 1(2), 80–89. <https://doi.org/10.62762/TMI.2025.306750>