

Article

Not peer-reviewed version

Leveraging 5G RedCap and Spiking Neural Networks for Energy Efficiency in Edge Devices

[Michail Alexandros Kourtis](#)*, [Andreas Oikonomakis](#), [Achileas Economopoulos](#), [Michalis Batistatos](#), [Gion Kalemaj](#), [Averkios Vasalos](#), [George Xilouris](#), [Panagiotis Trakadas](#)

Posted Date: 7 November 2025

doi: 10.20944/preprints202511.0410.v1

Keywords: 5G RedCap; SNN; Edge; UAVs



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

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

Article

Leveraging 5G RedCap and Spiking Neural Networks for Energy Efficiency in Edge Devices

Michail Alexandros Kourtis ^{1,*}, Andreas Oikonomakis ¹, Achileas Economopoulos ¹, Michalis Batistatos ¹, Gion Kalemaki ¹, Averkios Vasalos ¹, George Xilouris ¹ and Panos Trakadas ²

¹ Institute of Informatics and Telecommunications, National Centre for Scientific Research "DEMOKRITOS" (NCSRDI), 15310 Athens, Greece

² Department of Ports Management and Shipping, National and Kapodistrian University of Athens, Evripus Campus, 34400 Euboea, Greece

* Correspondence: akis.kourtis@iit.demokritos.gr

Abstract

This work presents an energy efficient implementation for UAV-based systems over 5G networks with on-boarded accelerated processing capabilities and provides a preliminary evaluation analysis of the integrated solution. A two-fold comparative study focused on connectivity and edge processing for UAVs, realizes two discrete deployment scenarios, where standard 5G configuration with Artificial Neural Networks processing is evaluated against 5G RedCap connectivity paired with Spiking Neural Networks. Both proposed alternative energy efficient solutions, are designed to offer significant energy saving, and this paper examines if they are fit candidates for energy stringent environments, i.e., UAVs, and also quantify the impact on the overall energy consumption of the system. The integrated solution with 5G RedCap/SNN realizes energy-use reductions approaching 60%, which translated to approximately 35% of increased flight time. The experimental evaluations were performed in a real-world deployment using a 5G equipped UAV with edge processing capabilities based on NVIDIA's Jetson Orin.

Keywords: 5G RedCap; SNN; Edge; UAVs

1. Introduction

The emergence of decentralized technologies and network communications, along with the introduction of heterogeneous devices at the edge, i.e., Unmanned Aerial Vehicles (UAVs), has created the need for new requirements in the design and deployment of telco 5G and 6G applications in terms of energy efficiency [1]. The emergence of energy stringent edge devices, especially in the case of UAVs with limited battery power, limits the use of performant Artificial Intelligence (AI) and Machine Learning (ML) enablers onboarded on them, thus not truly exploiting the edge processing proximity [2]. 5G communications have been largely utilized as a connectivity pathway for UAVs, in numerous applications and multiple works [3] have showcased the benefits and limitations of this convergence. In this respect, 5G Reduced Capability (RedCap) has been implemented, which promises to provide 5G bandwidth performance, with significant power saving compared to standard 5G deployments [4], [5].

From the perspective of AI/ML processing Artificial Neural Networks (ANNs) technology is a general-purpose neural network that can be used for a wide range of tasks, including classification, regression, and pattern recognition, and have been widely applied in various applications [6], including object classification in image processing. However, the training, as well as the inference of this type of models is energy demanding and can significantly reduce UAV flight time, when deployed at an edge scenario. A novel approach on edge processing is based on Spiking Neural Networks (SNNs), which is a type of artificial neural network that more closely mimics biological

neurons by communicating through discrete "spikes" rather than continuous values. This "third generation" network integrates the concept of time into its operations, making it particularly well-suited for time-series analysis and applications like robotics and radar, while also being more energy-efficient due to its asynchronous, event-driven processing. Nonetheless, its main advantage relies in energy efficiency and processing speed, since due to their event-driven and sparse nature, computations only occur when a neuron "spikes" rather than at every time step, compared to ANNs and Convolutional Neural Networks (CNNs) [7].

This paper introduces a hybrid approach for energy efficient edge processing focusing on the energy limitations of UAV devices. UAV devices have been extensively used as a relay or User Equipment (UE) in 5G networks, where control communications or data transfer from onboarded sensors are using the 5G channel. Additionally, various works have used onboarded advanced processing units, e.g., Raspberry Pi, Jetson Nano, etc., for on-flight processing in order to minimize processing latency, but consuming significantly more battery power, in an already energy capped environment [8]. In this work, a comprehensive experimental analysis and evaluation of the energy efficiency of alternative connectivity and processing enablers, i.e., 5G RedCap and SNNs, is carried out and portrays how targeted modules in the different parts of an edge ecosystem can provide substantial energy benefits, and increase flight time of a UAV.

The paper is organized as follows: Section 1 provided the incentive and scope of the work. Section 2 provides an extensive analysis on related work and what been achieved in the relevant field. In Section 3, the proposed solution is presented and extended experimental evaluations are depicted showcasing the benefits of the proposed approach. Finally, Section 4, concludes the outcomes of the paper and provide a future outlook on the next steps of the work.

2. Related Work

This section covers the relevant work in the field of energy efficiency at the edge, leveraging, either different connectivity setups, or alternative acceleration enablers. The state-of-the-art study aims to provide an analysis both from the scope of related energy efficient methods in 5G, and more specifically 5G RedCap, and how these can benefit edge, IoT and different devices under energy stringent conditions. In parallel, the study also covers how SNNs and related neural network approaches can be applied in different scenarios to reduce significantly energy consumption. In the context of energy efficiency at the edge most related work focuses on a single domain of technology stack, either the connectivity one or the processing acceleration. In this work, a converged approach is proposed based both on alternative 5G connectivity methods available, as well as low-power edge acceleration, i.e., SNNs. The convergence of 5G wireless technologies, neuromorphic computing, and energy-efficient edge processing can act as an accelerator in distributed intelligence systems. This section reviews recent advances across multiple interconnected domains that helped build the foundations of the presented work on leveraging 5G RedCap and Spiking Neural Networks for energy efficiency in edge devices, with particular emphasis on aerial applications.

2.1. 5G Reduced Capability Technology for IoT and Edge Devices

5G Reduced Capability (RedCap), standardized in 3GPP Release 17 and finalized in mid-2022, addresses a critical gap between high-end 5G capabilities and low-power wide-area technologies [9]. The key technical specifications include reduced bandwidth (20 MHz maximum in FR1 versus 100 MHz for conventional 5G NR), simplified antenna configurations supporting 1-2 receive branches rather than 2-4, and maximum of 2 MIMO layers [10]. These architectural simplifications promise to yield approximately 65% reduction in modem complexity for FR1 devices and 50-60% lower module costs compared to full 5G implementations [11].

From an energy efficiency perspective, RedCap introduces several critical features. The extended Discontinuous Reception (eDRX) mechanism supports DRX cycles up to 10,485 seconds (nearly 3 hours) in RRC idle state, compared to 2.56 seconds maximum in earlier releases [12]. This enables 10-70× battery lifetime extension depending on traffic patterns [13]. Additionally, Radio Resource

Management (RRM) relaxation reduces measurement overhead for stationary or low-mobility devices. Industry validation from Samsung and MediaTek deployment in the State Grid Shandong network achieved 32% energy reduction with RedCap terminals utilizing eDRX and measurement relaxation features [13].

2.2. Spiking Neural Networks for Energy-Efficient Edge Computing

Spiking Neural Networks represent an neural architectural shift toward brain-inspired, event-driven computing that offers substantial energy advantages to edge deployments. However, recent research reveals that achieving energy efficiency requires careful architectural design and delicate vector parameterization. A 2024 analysis reconsidering SNN energy efficiency demonstrates that for VGG16 architecture with time window $T=6$, SNNs require greater than 93% sparsity rate to achieve energy advantages over conventional ANNs on most hardware architectures [14]. This finding establishes critical design constraints: energy efficiency depends fundamentally on achieving high activation sparsity and minimizing temporal window size during inference.

Despite these constraints, practical implementations demonstrate substantial energy savings in favorable scenarios. The Synsense Xylo neuromorphic processor, combining analog preprocessing with digital SNN inference, achieved $60.9\times$ dynamic inference power reduction and $33.4\times$ dynamic inference energy reduction versus Arduino baseline on acoustic scene classification tasks while maintaining comparable 90% accuracy [15]. Intel's Loihi 2 neuromorphic processor demonstrated $37.24\times$ lower power consumption than CPU on QUBO optimization tasks, with the ability to solve workloads $4\times$ larger than CPU at 10^{-2} second timeout [16]. The fundamental advantage derives from replacing multiply-accumulate operations (3.2 pJ) with simple accumulations (0.1 pJ), yielding $32\times$ energy advantage per operation, which compounds with activation sparsity to achieve $320\times$ theoretical energy reduction versus equivalent RNNs when sparsity exceeds 90% [17].

2.3. Neuromorphic Hardware Platforms

Intel's Loihi series represents the most widely deployed neuromorphic research platform. The original Loihi chip (2017-2018) integrated 128 neuromorphic cores supporting 130,000 neurons and 130 million synapses in 14nm CMOS, achieving $1000\times$ energy efficiency improvement over conventional GPUs on specific workloads [18]. Loihi 2 (2021-present) delivers $10\times$ faster processing and up to $12\times$ higher performance, supporting 1 million neurons and 120 million synapses per chip with programmable neuron models via microcode-based architecture [19]. The Hala Point system scales this to 1.15 billion neurons across 1,152 Loihi 2 chips, achieving 20 petaops throughput with greater than 15 trillion 8-bit operations per second per watt efficiency [20]. IBM's evolution from TrueNorth (1 million neurons, 70 mW power consumption) to NorthPole (2023) demonstrates continued progress, with NorthPole achieving $22\times$ faster processing than TrueNorth and $25\times$ less energy than GPUs on equivalent 12nm process while occupying $5\times$ less physical space [20].

2.4. Learning Algorithms and Training Methods

Surrogate gradient methods have emerged as the dominant training approach for deep SNNs, addressing the non-differentiability of spike generation by replacing Heaviside step function derivatives during backpropagation [21]. Common surrogate functions include boxcar/rectangular, sigmoid derivative, exponential, and piecewise linear approximations. This approach enables competitive performance: 94.51% accuracy on speech command recognition and 94.62% on Heidelberg Digits dataset, surpassing CNN baselines while retaining precise spike timing information [19]. Recent advances include adaptive surrogate gradients using χ -based training pipelines, temporal regularization enforcing stronger constraints on early timesteps, and ADMM-based training avoiding non-differentiability issues entirely [22]. Online Training Through Time (OTTT) achieves constant memory consumption regardless of timesteps, while Spatial Learning Through Time (SLTT) reduces memory by up to $10\times$ by ignoring unimportant computational graph

routes [23]. Spike-Timing-Dependent Plasticity (STDP) offers biologically plausible local learning but struggles to scale to supervised deep learning tasks, finding primary application in unsupervised feature extraction and hardware-friendly implementations [19]. ANN-to-SNN conversion provides an alternative deployment path, achieving near-lossless conversion with proper quantization-aware training, though requiring sufficient temporal resolution for rate coding convergence [24].

2.5. Edge Computing Applications

SNNs demonstrate particular strength in temporal processing and always-on sensing applications. The NeuroBench 2025 framework provides standardized benchmarks revealing state-of-art performance: acoustic scene classification achieves 90% accuracy at 0.16W versus 9.75W for conventional processing, while event camera object detection with hybrid ANN-SNN architectures achieves 0.271 mAP with one order of magnitude fewer operations than pure ANN approaches [10]. The EC-SNN framework for split deep SNNs across edge devices achieves 60.7% average latency reduction and 27.7% energy consumption reduction per device through channel-wise pruning while maintaining accuracy across six datasets [25].

Distributed SNN architectures for wireless edge intelligence demonstrate significant advantages. Multiple edge nodes with spiking neuron subsets collaborating over wireless channels achieve power reductions while ensuring inference accuracy, with SNNs proving much more bandwidth-efficient than RNNs due to sparse spike communication [26].

2.6. Energy Efficiency Optimization in Edge Devices

Energy efficiency in edge computing encompasses hardware, software, and communication-level optimizations. This challenge has produced different solutions addressing different aspects of the energy consumption problem.

Power management through Dynamic Voltage and Frequency Scaling (DVFS) remains fundamental for heterogeneous multicore edge processors [27]. Liu et al. demonstrate a novel algorithm integrating task prioritization, core-aware mapping, and predictive DVFS on ARM Cortex-A15/A7 heterogeneous platforms, achieving 20.9% energy reduction compared to Earliest Deadline First scheduling, 11.4% versus HEFT, and 5.9% versus Energy-Aware Scheduling while maintaining 2.4% deadline miss rate (lowest among compared algorithms) and achieving 19.5 work units per joule energy efficiency [28]. Deep reinforcement learning approaches to DVFS using Deep Q-learning on Jetson Nano achieve 3-10% power reduction versus Linux built-in governors across 3-8 concurrent task workloads [29]. These techniques prove particularly relevant for battery-powered edge devices with heterogeneous processors common in ARM big.LITTLE architectures deployed in wireless edge scenarios.

2.7. Hardware Accelerators for Edge AI

Specialized accelerators demonstrate improvements in energy efficiency versus general-purpose processors [30]. Google's Coral Edge TPU delivers targeted ASIC performance at 2W for USB accelerator and 0.5W for M.2 module configurations, enabling multi-year battery life for always-on inference applications [31]. Hailo-8 achieves up to 26 TOPS at 2.5-5W typical power consumption with area and power efficiency superior to competing solutions by an order of magnitude, while the newer Hailo-10 delivers 40 TOPS at less than 5W, processing Llama2-7B LLM at 10 tokens per second and generating Stable Diffusion 2.1 images every 5 seconds—double the performance of Intel Core Ultra NPU at half the power [32]. Axelera AI's Metis AIPU scales to 214 TOPS with 1-16GB dedicated DRAM at less than 10W for retail analytics applications, achieving 97% order accuracy versus 90-95% for traditional systems [33]. Microcontroller-class devices increasingly integrate neural accelerators. Renesas RA8P1 MCU with Arm Ethos-U55 NPU delivers up to 256 GOPS while maintaining MCU-level power consumption suitable for battery-powered applications [34].

2.8. Integration of Wireless Communications with Neural Networks

The integration of advanced wireless technologies with neural network processing at the edge represents an emerging research frontier with substantial energy efficiency implications. Recent work demonstrates promising integration of SNNs with wireless communication systems [35]. SNNs consume approximately 10× less energy than ANNs for beamforming (0.002 mJ versus 0.018 mJ per optimization) while achieving comparable or superior performance through event-driven spike processing. For OFDM channel estimation, SNNs prove 3.5× more energy-efficient than conventional neural approaches and 1200× more efficient than traditional ChannelNet (0.0018 mJ versus 2.19 mJ). The framework incorporates domain knowledge into SNN architecture using LIF neurons with surrogate gradient training and addresses catastrophic forgetting in dynamic wireless environments through hypernet gates and spiking rate consolidation.

Federated neuromorphic learning for wireless edge AI demonstrates substantial practical advantages. Lead Federated Neuromorphic Learning achieves greater than 4.5× energy savings compared to standard federated learning with less than 1.5% accuracy loss, alongside 3.5× data traffic reduction and 2.0× computational latency reduction using Meta-Dynamic Neuron architectures [36]. This privacy-preserving approach enables speech recognition, image classification, health monitoring, and multi-object detection on edge devices connected via wireless networks without sharing raw data, only model parameters [36].

2.9. Edge AI Over 5G Networks

The convergence of 5G connectivity with edge neural network processing enables sophisticated distributed intelligence architectures. The Edgent framework demonstrates device-edge synergy for DNN inference through adaptive DNN partitioning (splitting networks between device and edge) and DNN right-sizing (early exiting at appropriate intermediate layers), optimizing the computation-communication tradeoff in 5G environments [37]. Split computing frameworks for UEs, edge nodes, and core devices reduce user equipment computational footprint by over 50% while containing inference time, proving robust in heterogeneous settings typical of 5G and emerging 6G networks [38].

Network-aware 5G edge computing implementations demonstrate practical deployment [39]. Machine learning-based dynamic resource scheduling for 5G network slicing using reinforcement learning and CNNs outperforms heuristic and best-effort approaches for low-latency services [41]. However, the substantial energy increase of 5G RAN—consuming 4-5× more energy than 4G networks despite per-bit efficiency improvements—motivates exploration of neuromorphic and efficient AI approaches [39]. ML-based digital predistortion for power amplifiers uses neural networks to replace Volterra-based linearization, reducing optimization time from days to hours while enabling dynamic correction of nonlinear PA behavior with reduced computational requirements [40].

2.10. Research Gaps in RedCap-AI Integration

Despite active 5G RedCap deployment for IoT connectivity and demonstrated advantages of neuromorphic computing, research specifically combining RedCap with neural network inference—particularly SNNs—remains extremely sparse. While RedCap standardization focuses on reduced complexity and extended battery life targeting wearables, industrial sensors, and video surveillance [1,3,4], optimization of neural network processing for RedCap device characteristics has received minimal attention. This represents a significant opportunity: RedCap infrastructure deployment accelerated in 2023-2024 with AT&T achieving nationwide coverage exceeding 200 million points of presence, Qualcomm releasing Snapdragon X35 modem-RF systems, and commercial modules from Fibocom, Quectel, and other manufacturers entering production [41]. However, research on AI inference optimization specifically for RedCap's bandwidth constraints (20 MHz FR1), simplified antenna configurations (1-2 Rx branches), and power-saving features (eDRX, RRM relaxation) has not materialized in peer-reviewed literature. Similarly, the potential synergy between RedCap's

energy-saving features and SNN's sparse, event-driven processing remains unexplored, despite both technologies targeting energy-constrained edge applications.

3. Experimental Evaluation of 5G RedCap and SNNs for Edge Deployments and Energy Efficiency Measurement

In this paper the focus on energy efficiency and saving is two-fold both from the connectivity perspective and the edge processing one. Since the targeted environments, i.e., UAVs, have limited energy and payload capacity, processing onboard is required to be efficient. In this aspect, the proposed work performs a dual experimental implementation and evaluation, both on the connectivity part with the usage of 5G RedCap and on the edge processing part with the integration of SNN. This section provides an overall analysis on the integration steps required to implement such setup and also presents the corresponding energy consumption evaluations of each setup.

3.15. G RedCap Configuration Analysis

In the table below, both 5G configuration setups are presented side by side, and in detail to showcase, what actual is modified in the process of enabling the RedCap configuration, and what is its actual difference to the standard configuration.

Parameter	Standard 5G NR (Amarisoft Sample)	RedCap 5G (Amarisoft Sample)	How RedCap Differentiates
Duplex mode & pattern	NR_TDD=1 (TDD); FR1 example uses pattern 2 via NR_TDD_CONFIG=2.	NR_TDD=0 (FDD) in RedCap config samples.	RedCap devices may operate HD-FDD to cut RF complexity; Amarisoft tutorial notes gNB can assume half-duplex until UE capabilities arrive.
Channel (cell) bandwidth	NR_BANDWIDTH=40 MHz in FR1 default sample (100 MHz in FR2 path).	Often NR_BANDWIDTH=20 MHz in RedCap examples.	RedCap UE max BW is 20 MHz (FR1) / 100 MHz (FR2) ; if the cell is wider, configure RedCap-only BWPs .
Antennas / MIMO	N_ANTENNA_DL=1, N_ANTENNA_UL=1 (SISO) by default.	Example shows N_ANTENNA_DL=2, N_ANTENNA_UL=1.	RedCap reduces UE complexity: fewer Rx branches / DL layers vs baseline NR.
MCS tables (mod/coding)	PDSCH and PUSCH mcs_table="qam256".	Global tables may stay 256-QAM, but on RedCap UL BWP mcs_table="qam64".	UL modulation intentionally capped on the RedCap BWP to match typical UE limits.

PRACH (initial access)	TDD FR1 example: prach_config_index=160; FDD example: 16.	RedCap UE performs PRACH on a RedCap-only BWP ; RRC Setup Request uses MAC LCID 35 (RedCap).	Enables early RedCap identification and attach on the narrow BWP.
Bandwidth Parts (BWP)	Single initial BWP spanning the cell in basic SA sample.	Adds dl_bwp_access:"redcap_only" / ul_bwp_access:"redcap_only" for narrow RedCap operation (esp. when cell BW > 20 MHz).	Constrains RedCap traffic to a 10–20 MHz slice while eMBB can use the wide carrier.
SSB / discovery	Cell-defining SSB only.	Supports NCD-SSB on a RedCap BWP via ssb_nr_arfcn (configured into nonCellDefiningSSB-r17).	Improves discovery when RedCap BWPs are narrower than the overall cell BW.
PUCCH (uplink control)	NR_LONG_PUCCH_FORMAT=0 (auto).	Example uses long PUCCH format 2 with tuned resources.	Tightens UCI overhead for constrained RedCap UEs/BWPs.
SIB1 content	Standard SIB1 (no RedCap IEs).	SIB1 includes redCap-ConfigCommon (e.g., halfDuplexCapAllowed, cellBarredRedCap).	Broadcasts RedCap rules so UEs camp/attach on the proper BWP.
Scheduler duplex behavior	N/A (no special assumption).	gNB may assume half-duplex before UE capability signaling, then switch if UE isn't HD-FDD.	Aligns with RedCap's optional HD-FDD profile to reduce device cost.
Resource allocation on RedCap BWP	Default contiguous allocation.	Tutorial recommends ra_type=type0 for non-contiguous PRB allocation around PUCCH on RedCap BWP (when running wide cell + narrow BWP).	Eases scheduling within a narrow BWP that reserves PRBs for control.
UE multi-carrier features	Baseline NR UEs may support CA/DC .	CA/DC not supported for RedCap UEs.	Single-carrier operation further reduces RF/baseband complexity.

This table explains how each configuration choice in the table affects device complexity, radio behavior, and energy use when moving from standard 5G NR to RedCap. In the baseline sample, standard 5G NR operates with wider bandwidth and typical duplexing, while RedCap narrows operation to reduce complexity and power. When the network configures a narrower carrier (or a RedCap-only bandwidth part inside a wider carrier), the user equipment processes a smaller slice of spectrum. This directly reduces the effort in the RF front-end and baseband, helping the device save power without breaking compatibility with the rest of the cell. Antenna and MIMO choices also change: RedCap limits receive diversity and the number of layers, which removes RF chains and signal-processing paths. This simplifies the hardware and lowers both idle and active power, at the cost of some peak throughput that is usually not needed for telemetry and command-and-control

traffic. Modulation and coding settings further reflect this trade-off. While downlink can remain capable, the uplink often uses a more conservative table (for example, capping at 64-QAM) to ease power amplifier requirements and processing load. Initial access and identification steps are streamlined so the device can camp and attach on the narrow RedCap bandwidth part from the start, avoiding wideband scanning and unnecessary measurements. Control signaling is tuned in the same direction: uplink control channels are placed and sized to work efficiently within the narrow bandwidth part, and common system information includes RedCap-specific fields so the device applies the right behavior immediately. Duplex assumptions at the scheduler can start conservatively, matching the simpler RedCap profile, and then adjust if the device indicates different capabilities. Resource allocation can reserve small regions for control and pack data grants around them to use the narrow bandwidth part efficiently. Carrier aggregation and dual connectivity are typically disabled for RedCap devices, which removes background measurements and coordination overhead. Overall, the pattern is consistent: peak throughput and feature breadth are traded for lower complexity, less measurement activity, and steadier operation in a smaller channel. For an energy-limited UAV, these choices translate into lower radio and baseband power draw while maintaining the reliability needed for control links and moderate data bursts, aligning with the endurance goals discussed later in the evaluation.

3.2. SNN Implementation Analysis

SNNs are rarely used for object detection and are rather utilized for analysis on the image temporal domain [42]. Therefore, a dedicated implementation and re-training on the Each extracted frame undergoes preprocessing, including conversion to grayscale and resizing to 28x28 pixels. This step is critical for standardizing the input for the SNN model. The SNN model, designed with a layer of 784 input neurons (corresponding to the 28x28 pixel input), 1000 hidden neurons, and 10 output neurons, is trained using the Contrastive Hebbian Learning rule. This architecture is selected for its balance between complexity and computational efficiency. A sliding window approach is implemented to scan each frame. Windows of size 64x64 pixels move across the frame with a step size of 32 pixels. Each window is processed to detect rust, and bounding boxes are drawn around detected areas.

The model's performance is evaluated based on its energy consumption, as it demonstrated low accuracy in detecting rust, with the contour approximation and edge detection techniques, approximately 26% compared to 60% for the ANN based method. The sliding window approach allows the model to localize rust patches effectively, even in complex backgrounds. However, the model's performance varies with the presence of significant background noise, indicating the need for further refinement. The use of SNNs for rust detection offers significantly lower power consumption ANNs. The incorporation of image processing techniques such as contour approximation and Hough Transform enhances the model's robustness and accuracy. Future work could explore the integration of additional mathematical models like Fourier Transform and Wavelet Transform to further improve detection capabilities.

3.2. Experimental Evaluation Results

In the figure below the comparison between the standard 5G with ANN setup and the 5G RedCap with SNN one, is presented in terms of energy consumption for a scenario of rust detection on a high mast antenna using ANN and SNN each time. Respectively, for each run, standard 5G and 5G RedCap was used simultaneously at full capacity, in order to benchmark the upper bound of each connectivity configuration. The results show a significant difference of ~7000mW in energy saving for the 5G RedCap/SNN in comparison to the standard 5G/ANN setup. The difference percentage in energy saving was calculated to 65% overall, depicting the noteworthy efficiency for the first setup.

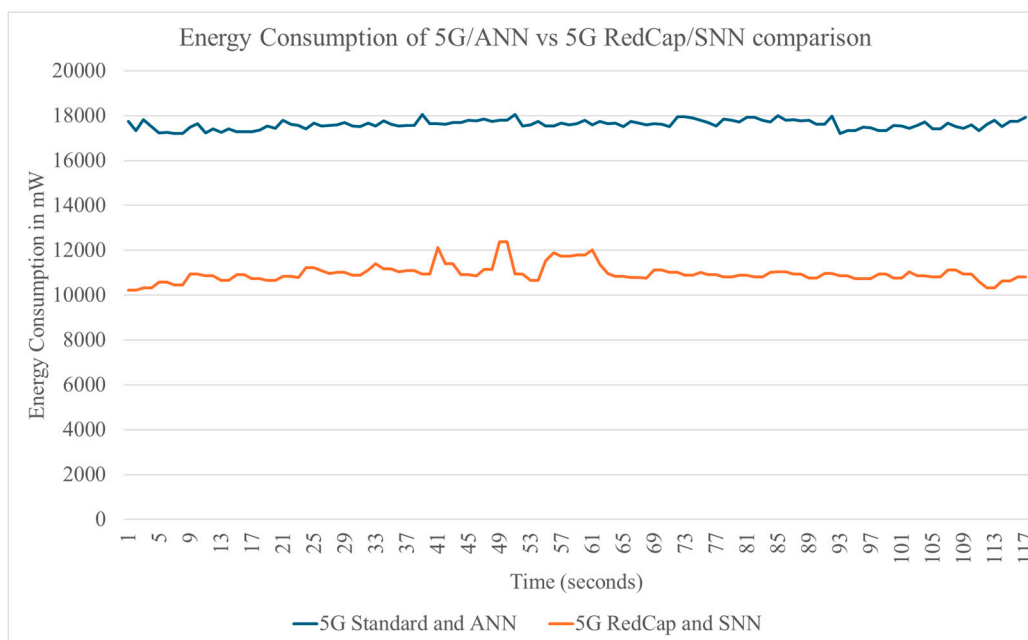


Figure 1. Energy consumption graph of 5G/ANN vs 5G RedCap/SNN comparison.

The experimental evaluation also measured the maximum bandwidth achieved with the standard 5G setup outperforming the 5G RedCap, 100Mbps to 40Mbps on average, showing a relative 50% difference. However, these evaluations do not directly translate on the energy impact of the proposed setup in a real-time UAV based experiment. Therefore in the next figure, the flight duration for both setups was measured, while each setup was onboarded on an actual Drone for each run. Each time, the battery is fully charged when the experiment starts, where the UAV is equipped with a Teltonika 5G Router and an NVIDIA Jetson Orin to perform ANN and SNN, for each scenario respectively. The battery charge is monitored during the entire duration of the flight, and the experiment is ended, once the battery voltage reaches a pre-defined lower point for the UAV to land safely.



Figure 2. Experimental UAV setup with onboarded connectivity and processing modules.

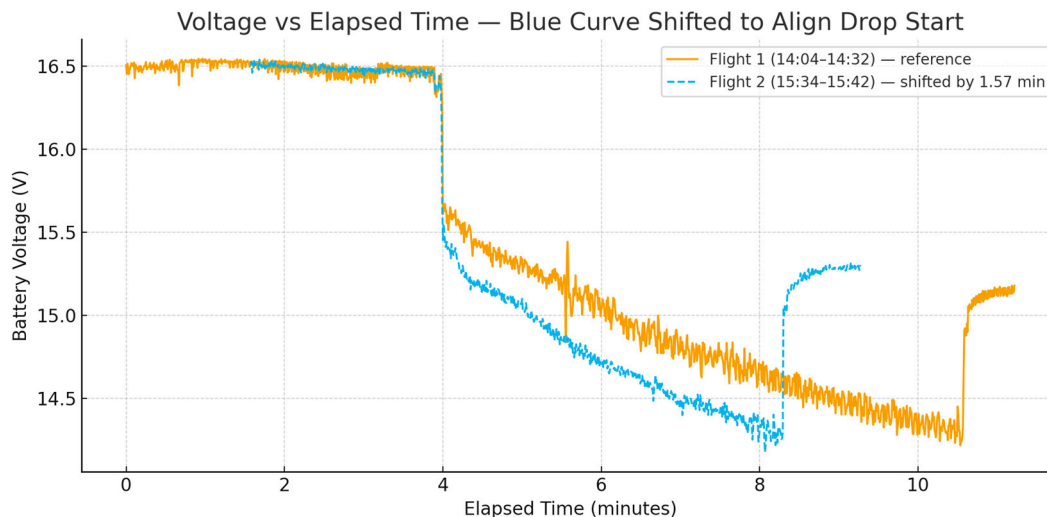


Figure 3. Flight duration comparison for configuration 1 (5G-ANN) and configuration 2 (5G RedCap-SNN), onboarded on drones.

As it can be depicted from the Figure above, for the setup with standard 5G/ANN the flight duration is approximately 4 minutes, while for the 5G RedCap/SNN one, it managed to reach 6 ½ minutes. This shows a flight time extension of around 35% for the 5G RedCap/SNN setup, underlining how energy efficient RedCap with SNN can prove to be for UAV related operations. It is also important to note, how the proposed work depicts how specific connectivity and edge processing enablers are translated in terms of energy efficiency in energy constrained environments, also in relation with the entire operational system. In other words, how the 65% percentage of measured energy saving in the actual energy consumption of both setups, is transposed to 35% overall energy reduction in UAV flight time. It is an important notion to identify and quantify how different UAV connectivity and processing modules impact the battery life of a UAV.

5. Conclusions

This work evaluated a unified approach for energy-aware UAV operations that combines 5G Reduced Capability (RedCap) for connectivity with Spiking Neural Networks (SNNs) for onboard inference. In a real deployment with a 5G-equipped UAV and edge processing on Jetson Orin, the integrated RedCap/SNN setup cut device-level power draw by about 60–65% relative to a baseline of standard 5G with ANN, and translated this saving into roughly a 35% increase in flight time, demonstrating that targeted choices in both radio and compute stacks can yield system-level benefits in energy-stringent environments. Beyond energy, the experiments highlight the practical trade-off between throughput and autonomy: while standard 5G achieved higher bandwidth (≈ 100 Mbps vs. ≈ 40 Mbps for RedCap), the RedCap/SNN configuration delivered materially longer airtime, which is often the binding constraint for battery-limited UAV missions. These findings suggest that, for edge workloads with moderate data-rate needs, prioritizing RedCap features and event-driven SNN inference can unlock better endurance without sacrificing task effectiveness.

Additionally, the study followed the motivation set in the introduction—addressing the limits of AI/ML on energy-constrained edge devices and the need for more efficient connectivity—by validating a dual optimization path across the RAN and the onboard AI stack. The results provide initial evidence that aligning RedCap’s narrow-band operation and power-saving mechanisms with the sparse, asynchronous nature of SNN processing is a viable route to energy-efficient edge intelligence in UAVs.

For future next steps, the work will cover, firstly, the evaluation on broader trials across tasks and platforms are needed to generalize the gains. Secondly, while we quantified end-to-end effects

(device power and flight time), a deeper analysis of specific RedCap parameters and SNN design choices would clarify where the largest marginal savings occur.

Author Contributions: Conceptualization, M.A. Kourtis; methodology, M.A. Kourtis; software, A. Oikonomakis, A. Economopoulos, G. Kalemaki, A. Vassalos; validation, M. Batistatos; formal analysis, G. Xilouris; investigation, P. Trakadas; resources, M.A. Kourtis; data curation, A. Xilouris; writing—original draft preparation, M.A. Kourtis; writing—review and editing, M.A. Kourtis; visualization, A. Oikonomakis; supervision, G. Xilouris, P. Trakadas; project administration, M.A. Kourtis; funding acquisition, M.A. Kourtis. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Open autonomous programmable cloud apps and smart sensors (OASEES) Project (Grant Number: 101092702).

Data Availability Statement: The original data presented in the study are openly available in ZENODO at 10.5281/zenodo.17533666 and 10.5281/zenodo.16679949.

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

Abbreviations

The following abbreviations are used in this manuscript:

3GPP	3rd Generation Partnership Project
5G NR	5th-Generation New Radio
6G	Sixth-Generation (future mobile networks)
AI	Artificial Intelligence
ANN	Artificial Neural Network
ARFCN	Absolute Radio-Frequency Channel Number
ASIC	Application-Specific Integrated Circuit
BWP	Bandwidth Part
CNN	Convolutional Neural Network
CPU	Central Processing Unit
DNN	Deep Neural Network
DVFS	Dynamic Voltage and Frequency Scaling
eDRX	Extended Discontinuous Reception
eMBB	Enhanced Mobile Broadband
FDD	Frequency-Division Duplex
FR1	Frequency Range 1 (sub-6 GHz)
FR2	Frequency Range 2 (mmWave)
gNB	Next-generation NodeB (5G base station)
GOPS	Giga Operations Per Second
GPU	Graphics Processing Unit
HD-FDD	Half-Duplex FDD
mAP	mean Average Precision
MIMO	Multiple-Input Multiple-Output
mW	milliwatt
NPU	Neural Processing Unit
NR	New Radio (5G air interface)
OASEES	Open autonomous programmable cloud apps and smart sensors (project)
OFDM	Orthogonal Frequency-Division Multiplexing
PRACH	Physical Random Access Channel
PRB	Physical Resource Block
PDSCH	Physical Downlink Shared Channel
PUCCH	Physical Uplink Control Channel
PUSCH	Physical Uplink Shared Channel
QoS	Quality of Service
QUBO	Quadratic Unconstrained Binary Optimization
RAN	Radio Access Network
RCM	RedCap-only (context: restricted operation on specific BWPs)

RedCap	Reduced Capability (5G NR device/profile)
RF	Radio Frequency
RNN	Recurrent Neural Network
RRC	Radio Resource Control
RRM	Radio Resource Management
SIB1	System Information Block Type 1
SISO	Single-Input Single-Output
SLTT	Spatial Learning Through Time
SNN	Spiking Neural Network
SSB	Synchronization Signal Block
STDP	Spike-Timing-Dependent Plasticity
TDD	Time-Division Duplex
TOPS	Tera Operations Per Second
TPU	Tensor Processing Unit
UCI	Uplink Control Information
UE	User Equipment
UAV	Unmanned Aerial Vehicle

References

1. M. Abrar, U. Ajmal, Z. M. Almohaimeed, X. Gui, R. Akram and R. Masroor, "Energy Efficient UAV-Enabled Mobile Edge Computing for IoT Devices: A Review," in *IEEE Access*, vol. 9, pp. 127779-127798, 2021, doi: 10.1109/ACCESS.2021.3112104.
2. A. Asheralieva and D. Niyato, "Effective UAV-Aided Asynchronous Decentralized Federated Learning With Distributed, Adaptive and Energy-Aware Gradient Sparsification," in *IEEE Internet of Things Journal*, vol. 12, no. 14, pp. 27461-27480, 15 July 2025, doi: 10.1109/JIOT.2025.3565005.
3. 3GPP. 2022. A Glimpse into RedCap NR devices. 3rd Generation Partnership Project. <https://www.3gpp.org/technologies/nr-redcap-glimpse>
4. Ericsson. 2021. What is reduced capability (RedCap) NR? Ericsson Technology Blog. <https://www.ericsson.com/en/blog/2021/2/reduced-cap-nr>
5. Ericsson. 2023. RedCap: Expanding the 5G device ecosystem for consumers and industries. White Paper. <https://www.ericsson.com/493d70/assets/local/reports-papers/white-papers/redcap-5g-iot-for-wearables-and-industries.pdf>
6. Kufel, J.; Bargieł-Łączek, K.; Kocot, S.; Koźlik, M.; Bartnikowska, W.; Janik, M.; Czogalik, Ł.; Dudek, P.; Magiera, M.; Lis, A.; et al. What Is Machine Learning, Artificial Neural Networks and Deep Learning?— Examples of Practical Applications in Medicine. *Diagnostics* 2023, 13, 2582. <https://doi.org/10.3390/diagnostics13152582>
7. Deng, S., Yu, D., Lv, C., Du, X., Jiang, L., Zhao, X., ... & Zomaya, A. Y. (2025). Edge Intelligence with Spiking Neural Networks. arXiv preprint arXiv:2507.14069.
8. S. P. Baller, A. Jindal, M. Chadha and M. Gerndt, "DeepEdgeBench: Benchmarking Deep Neural Networks on Edge Devices," 2021 IEEE International Conference on Cloud Engineering (IC2E), San Francisco, CA, USA, 2021, pp. 20-30, doi: 10.1109/IC2E52221.2021.00016.
9. 5G RedCap, completing the 3GPP puzzle and beyond. u-blox IoT Insights. <https://www.u-blox.com/en/blogs/insights/5g-redcap>
10. Richardson RFPD. 2024. 5G RedCap: An Essential Technology for IoT. Technical Article. <https://shop.richardsonrfpd.com/docs/rfpd/RRFPD-5G-RedCap-An-Essential-Technology-for-IoT-2024-08.pdf>
11. Telenor IoT. 2024. 5G RedCap for IoT: The Perfect Balance of 5G Capabilities. <https://iot.telenor.com/technologies/connectivity/5g-redcap/>
12. Ericsson. 2024. RedCap/eRedCap – standardizing simplified 5G IoT devices. <https://www.ericsson.com/en/blog/2024/12/redcap-eredcap>
13. ABI Research. 2024. 5G RedCap: Simplifying Cellular Connectivity for IoT Devices. <https://www.abiresearch.com/blog/5g-redcap>

14. Schuman, C. D., et al. 2024. Reconsidering the energy efficiency of spiking neural networks. arXiv:2409.08290.
15. Mattson, P., et al. 2025. The NeuroBench framework for benchmarking neuromorphic computing algorithms and systems. *Nature Communications*, 16(1), 687.
16. Intel. 2024. Neuromorphic Computing and Engineering with AI. Intel Research. <https://www.intel.com/content/www/us/en/research/neuromorphic-computing.html>
17. Davies, M., et al. 2018. Loihi: A neuromorphic manycore processor with on-chip learning. *IEEE Micro*, 38(1), 82-99.
18. Open Neuromorphic. 2023. A Look at Loihi - Intel Neuromorphic Chip. <https://open-neuromorphic.org/neuromorphic-computing/hardware/loihi-intel/>
19. Open Neuromorphic. 2024. A Look at Loihi 2 - Intel Neuromorphic Chip. <https://open-neuromorphic.org/neuromorphic-computing/hardware/loihi-2-intel/>
20. Intel. 2024. Intel Hala Point neuromorphic system achieves 20 petaops with 1.15 billion neurons. Intel Newsroom.
21. Neftci, E. O., Mostafa, H., and Zenke, F. 2019. Surrogate gradient learning in spiking neural networks. *IEEE Signal Processing Magazine*, 36(6), 51-63.
22. Cramer, B., Stradmann, Y., Schemmel, J., and Zenke, F. 2022. A surrogate gradient spiking baseline for speech command recognition. *Frontiers in Neuroscience*, 16, 865897.
23. arXiv. 2025. Accuracy-Robustness Trade Off via Spiking Neural Network Gradient Sparsity Trail. <https://arxiv.org/html/2509.23762>
24. ScienceDirect. 2025. A comprehensive multimodal benchmark of neuromorphic training frameworks for spiking neural networks. <https://www.sciencedirect.com/science/article/abs/pii/S0952197625015453>
25. arXiv. 2019. Low-Power Neuromorphic Hardware for Signal Processing Applications. <https://arxiv.org/pdf/1901.03690>
26. IJCAI. 2024. EC-SNN: Splitting Deep Spiking Neural Networks for Edge Devices. Proceedings of the 33rd International Joint Conference on Artificial Intelligence. <https://www.ijcai.org/proceedings/2024/596>
27. Liu, T., et al. 2024. Energy-Efficient Distributed Spiking Neural Network for Wireless Edge Intelligence. *IEEE Transactions on Wireless Communications*, 23(8), 9074-9088.
28. IJACSA. 2024. Optimizing Energy Efficient Cloud Architectures for Edge Computing: A Comprehensive Review. *International Journal of Advanced Computer Science and Applications*, 15(11).
29. Liu, Z., et al. 2025. Energy efficient task scheduling for heterogeneous multicore processors in edge computing. *Scientific Reports*, 15(1), 1234.
30. arXiv. 2024. Energy-Efficient Computation with DVFS using Deep Reinforcement Learning for Multi-Task Systems in Edge Computing. <https://arxiv.org/html/2409.19434v3>
31. Viso.ai. 2024. AI Hardware: Edge Machine Learning Inference Accelerators Overview. <https://viso.ai/edge-ai/ai-hardware-accelerators-overview/>
32. Jaycon Systems. 2025. Top 10 Edge AI Hardware for 2025. <https://www.jaycon.com/top-10-edge-ai-hardware-for-2025/>
33. Renesas. 2024. Enable High Performance, Low Power Inference in Your Edge AI Applications. <https://www.renesas.com/en/blogs/enable-high-performance-low-power-inference-your-edge-ai-applications>
34. Cspub-ijcिसim. 2024. Lightweight Deep Learning Models For Edge Devices—A Review. <https://cspub-ijcिसim.org/index.php/ijcिसim/article/download/1025/635/1787>
35. arXiv. 2022. SpikACom: A Comprehensive Neuromorphic Computing Approach for Wireless Communication. arXiv:2203.05634.
36. Zhou, S., et al. 2022. Lead federated neuromorphic learning for wireless edge artificial intelligence. *Nature Communications*, 13, 4269.
37. Li, E., et al. 2019. Edge AI: On-Demand Accelerating Deep Neural Network Inference via Edge Computing. *IEEE Transactions on Wireless Communications*, 19(1), 447-457.
38. arXiv. 2025. Optimized Split Computing Framework for Edge and Core Devices. <https://arxiv.org/html/2509.06049>

39. Analog Devices. 2024. Driving 5G Energy Efficiency with Edge AI and Digital Predistortion. <https://www.analog.com/en/signals/articles/5g-energy-efficiency-with-edge-ai.html>
40. Analog Devices. 2024. ML-based Digital Predistortion for 5G Power Amplifiers. Analog Devices Technical Article.
41. AT&T. 2023. AT&T Hits Another U.S. Industry-First: 5G RedCap Data Connection. <https://about.att.com/blogs/2023/5g-redcap.html>
42. Li, W.; Zhao, J.; Su, L.; Jiang, N.; Hu, Q. Spiking Neural Networks for Object Detection Based on Integrating Neuronal Variants and Self-Attention Mechanisms. *Appl. Sci.* 2024, 14, 9607. <https://doi.org/10.3390/app14209607>

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