AN ABSTRACT OF THE RESEARCH PROJECT OF

Austin Martin for the degree of Master of Science in Computer Science presented on September 2025.

Title: Machine Learning-Based Real-Time Canine Seizure Detection Using a Smart Collar

Abstract approved: _____

Patrick Donnelly

Canine epilepsy affects 0.6–0.8% of the domestic dog population. Seizures are extraordinarily stressful events for both dogs and their owners. Caretakers must be constantly vigilant and manually track seizure events for veterinary diagnoses and treatment recommendations. Existing commercial pet wearables lack the multimodal sensing capabilities and real-time processing required for reliable seizure detection, and previously proposed methods to monitor epilepsy are surgically invasive and difficult to deploy outside of clinical settings. This master's project presents a complete Internet-of-Things (IoT) system for continuous and non-invasive canine activity monitoring, designed as a foundation for future seizure detection research.

The implemented system integrates four key components: (1) a custom smart collar built on Arduino Nano RP2040 Connect hardware, capturing eight sensor

channels (3-axis accelerometer, 3-axis gyroscope, temperature, and microphone amplitude) at 20 Hz; (2) a React Native Android application providing real-time data visualization, manual activity labeling, and alert management; (3) a robust AWS cloud architecture utilizing IoT Core, S3, and DynamoDB for scalable data ingestion and storage; and (4) dual-path connectivity supporting both direct Wi-Fi transmission and Bluetooth Low Energy proxy routing through the mobile application.

The cloud pipeline successfully processes sensor windows with less than 2-second latency over Wi-Fi and less than 10-second latency via cellular backup. The mobile application enables one-tap activity labeling across six behavioral categories (rest, walk, run, play, sleep, and seizure), which will enable the creation of a structured dataset for machine learning development. Using feature engineering, the system extracts multiple time- and frequency-domain characteristics from each 5-second sensor window, preparing the data for future for supervised learning experiments using Amazon SageMaker.

This project contributes a production-ready platform incorporating solutions for critical pet wearable requirements: power management strategies, robust connectivity handling, and a streamlined user experience design. The complete system generates properly labeled, multi-modal sensor data suitable for future research into canine seizure detection algorithms. While actual seizure data collection will require veterinary collaboration and extended field studies, the implemented infrastructure provides the necessary technological foundation for advancing automated canine health monitoring from research concept to practical deployment.

©Copyright by Austin Martin
September 2025
All Rights Reserved

Machine Learning-Based Real-Time Canine Seizure Detection Using a Smart Collar

by

Austin Martin

A RESEARCH PROJECT

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented September 2025
Commencement June 2025

Master of Science research project of Austin Martin presented on September 2025.

APPROVED:

Major Professor, representing Computer Science

Director of the School of Computer Science

Dean of the Graduate School

I understand that my research project will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my research project to any reader upon request.

Austin Martin, Author

ACKNOWLEDGEMENTS

To Alison, my wife, who carried us both through this journey—financially, emotionally, and with relentless support for what I desire to achieve.

To my parents, who planted the seeds of curiosity and taught me that questions matter as much as answers.

To Professor Patrick Donnelly, who saw potential and opened the door to graduate school through the accelerated master's program. Without that push, I would have never so readily walked through.

To my dogs, whose companionship reminds me daily that joy exists in the simplest moments.

To education itself, for the gift of passing knowledge forward and turning curiosity into understanding.

And finally, to myself, for persevering through the doubt, the late nights, and the moments when giving up felt logical.

TABLE OF CONTENTS

Page

1 Introduction 1
1.1 Clinical Context and Motivation 1
1.2 Current Monitoring Limitations 1
1.3 Proposed Solution 2
1.4 System Architecture 2
1.4.1 Smart Collar Device 2
1.4.2 Cloud Infrastructure 3
1.4.3 Mobile Application 3
1.5 Key Technical Innovations 3
1.6 Research Contributions 4
1.7 Thesis Organization 5

2 Background 6
2.1 Canine Epilepsy: Clinical Overview and Challenges 6
2.2 Evolution of Pet Health Monitoring Technology 8
2.3 Sensor Technologies for Physiological Monitoring 9
2.4 Internet of Things (IoT) in Healthcare Applications 10
2.5 Machine Learning for Time-Series Sensor Data 12
2.6 Human-Animal Bond and Technology Acceptance 15
2.7 Summary 16

3 Literature Review 17
3.1 Canine Epilepsy Detection Systems 17
3.2 Wearable Sensor Technologies for Animal Monitoring 20
3.3 Machine Learning for Seizure Detection 22
3.4 Advanced Sensor Fusion Techniques 25
3.5 Cloud Architecture and Edge Computing Trade-offs 26
3.6 Clinical Context and User Requirements 28
3.7 Treatment Challenges Justifying Continuous Monitoring 29
3.8 Comparative Analysis 30

TABLE OF CONTENTS (Continued)

Page

3.9 Research Gaps and Opportunities 32
3.10 Summary 34

4 System Architecture 36
4.1 Embedded Hardware Platform 38
4.1.1 Core Computing Platform 38
4.1.2 Communication Protocols 40
4.1.3 Sensor Suite Design 40
4.1.4 Power Management Architecture 42
4.1.5 Communication Architecture 43
4.2 Cloud Infrastructure 46
4.2.1 AWS IoT Core Integration 46
4.2.2 Data Storage Architecture 47
4.3 Machine Learning Pipeline 50
4.3.1 Pipeline Architecture 50
4.3.2 Feature Engineering 51
4.3.3 Model Selection and Training 53
4.3.4 Deployment Architecture 55
4.3.5 Model Performance 56
4.4 Mobile Application Architecture 57
4.4.1 Architectural Overview 57
4.4.2 Communication Architecture 58
4.4.3 State Management Architecture 60
4.4.4 Activity Tracking Implementation 60
4.4.5 User Interface Design 62
4.4.6 AWS Integration 63
4.4.7 Performance Optimization 64
4.5 Summary 64

5 Methodology 66
5.1 Data Collection Methodology 67
5.1.1 Subject Characteristics 67
5.1.2 Data Collection Protocol 68
5.1.3 Labeling Process 69

TABLE OF CONTENTS (Continued)

Page

5.2 Feature Engineering Methodology 70
5.2.1 Window Size Determination 70
5.2.2 Feature Extraction Pipeline 71
5.2.3 Feature Validation 72
5.2.4 Synthetic Data Augmentation 73
5.3 Model Development Methodology 75
5.3.1 Algorithm Selection 75
5.3.2 Training Methodology 76
5.3.3 Model Validation 77
5.3.4 Robustness Evaluation Methodology 77
5.4 System Integration Testing 79
5.4.1 End-to-End Latency Measurement 79
5.4.2 Reliability Testing 79
5.5 Evaluation Methodology 81
5.5.1 Multi-Tier Evaluation Framework 81
5.5.2 Statistical Validation of Synthetic Data 83
5.6 Limitations and Validity Threats 85
5.6.1 Internal Validity 85
5.6.2 External Validity 85
5.7 Summary 86

6 Results 87
6.1 Experimental Setup and Baseline Performance 87
6.1.1 Baseline Model Performance 88
6.2 Impact of Data Augmentation on Model Performance 89
6.2.1 Augmentation Multiplier Analysis 89
6.2.2 Performance Trajectory Analysis 91
6.3 Robustness Evaluation Under Noise 92
6.3.1 Systematic Noise Injection Results 93
6.3.2 Robustness Score Analysis 94
6.4 Augmentation Strategy Comparison 95
6.4.1 Impact of Technique Combination Strategies 95
6.5 Statistical Validation of Synthetic Data Quality 97
6.5.1 Distribution Similarity Analysis 97

TABLE OF CONTENTS (Continued)

Page

6.6 Classification Performance Analysis 99
6.6.1 Confusion Matrix Analysis 99
6.6.2 Feature Importance Evolution 101
6.7 Comprehensive Performance Summary 103
6.8 Summary 103

7 Conclusion 105
7.1 Summary of Achievements 105
7.1.1 Technical Implementation 105
7.1.2 Machine Learning Contributions 107
7.2 Limitations and Opportunities 107
7.2.1 Data Constraints and Personalization Potential 107
7.2.2 Technical Challenges 108
7.3 Future Directions 109
7.3.1 Near-term Enhancements 109
7.3.2 Clinical Validation Pathway 109
7.3.3 Algorithmic Evolution 110
7.4 Lessons Learned 111
7.4.1 Design Principles 111
7.4.2 Data Acquisition Challenges 111
7.5 Concluding Remarks 112

Bibliography 113

LIST OF FIGURES

Figure		Page
4.1	Complete smart collar system architecture showing data flow from sensors to cloud services	37
4.2	Hardware architecture of the smart collar embedded system	39
4.3	Two-stage power conversion architecture from battery to system components	42
4.4	AWS IoT Core message processing pipeline	48
4.5	End-to-end machine learning pipeline architecture	50
4.6	Real-time inference data flow with seizure alert pathway	56
4.7	Mobile application main interface featuring the 2×3 activity tracking grid with color-coded buttons, real-time connection status, and one-handed operation design	58
4.8	Activity recording workflow showing standard 5-second recording (a-b) and continuous seizure monitoring mode (c)	61
4.9	Complete application workflow from authentication through device pairing to data visualization and veterinary integration	63
6.1	Impact of augmentation multiplier on model performance metrics. Error bars show standard deviation from 10-fold cross-validation for accuracy/F_1 and 10 independent noise runs for robustness. Training set sizes: baseline (n=342), 3× (n=1,368), 5× (n=2,052), 10× (n=3,762).	90
6.2	Performance trajectory analysis showing accuracy, robustness evolution, overfitting analysis, and multi-metric comparison across augmentation multipliers	91
6.3	Model performance degradation under increasing noise levels. Shaded regions represent $\pm 1\sigma$ confidence bands from 10 independent noise injection runs. Left panel shows multiplier comparison, right panel shows strategy comparison.	93

LIST OF FIGURES (Continued)

Figure | Page

6.4 Performance comparison across augmentation strategies at 5× multiplier. Error bars show standard deviation from cross-validation (accuracy, F_1) and noise evaluation runs (robustness). 96

6.5 Statistical validation of synthetic data quality across augmentation strategies. Left: KS test statistics (lower is better, <0.1 indicates good similarity). Right: Wasserstein distances showing distributional divergence. 98

6.6 Confusion matrices showing classification performance on test set across augmentation multipliers. Values show raw counts with normalized accuracy indicated by color intensity. 100

6.7 Evolution of feature importance across augmentation configurations. Top panel shows individual feature importance heatmap, middle panels show category-level trends and stability metrics, bottom panels show consistency of top features and emergence patterns. 102

LIST OF TABLES

Table		Page
3.1	Comparison of Seizure Detection Approaches	31
3.2	ML Algorithm Performance on Stationary Time-Series Data and Animal Behavior	32
4.1	Sensor specifications and sampling rates	40
4.2	Feature distribution across sensor channels and domains	52
4.3	Communication mode characteristics	59
6.1	Baseline model performance metrics without data augmentation (n=342 real samples)	88
6.2	Robustness scores across augmentation multipliers with 95% confidence intervals	94
6.3	Comprehensive performance metrics across all experimental configurations. Gap refers to train-test accuracy difference.	103

Chapter 1: Introduction

This chapter establishes the clinical context for canine epilepsy monitoring, examines current technological limitations, and presents a comprehensive IoT platform for continuous health monitoring.

1.1 Clinical Context and Motivation

Epilepsy affects approximately 0.75% of the canine population worldwide [21]—over five million dogs in the United States alone. Despite veterinary advances, median survival after diagnosis is only 1.5 years, with 40% of affected dogs ultimately euthanized due to inadequate seizure control [21]. The unpredictable nature of seizures forces caregivers into profound lifestyle modifications including altered work schedules, restricted travel, and constant vigilance [42].

1.2 Current Monitoring Limitations

Current epilepsy management relies on owner-maintained seizure diaries that suffer fundamental limitations: nocturnal seizures go undetected, subtle focal seizures are misinterpreted, and emotional stress impairs accurate record-keeping. While human wearable seizure detection achieves accuracies exceeding 98% [7], veterinary applications lag significantly. Commercial pet monitors yield unacceptable per-

formance—Whistle achieved only 18.6-22.1% sensitivity [37]. This gap motivates development of advanced monitoring systems.

1.3 Proposed Solution

This paper presents an IoT platform for canine activity monitoring architected to support future seizure detection. While currently classifying normal activities (walking, running, sleeping, playing), the system's architecture enables seizure detection research. The implementation demonstrates technical feasibility through embedded sensing hardware, cloud-based machine learning, and real-time mobile alerting.

1.4 System Architecture

The system comprises three integrated components enabling continuous monitoring and classification.

1.4.1 Smart Collar Device

The Arduino Nano RP2040 Connect-based collar integrates 3-axis accelerometer/-gyroscope for movement patterns, temperature sensor for physiological changes, and microphone for vocalizations. Streaming data at 20 Hz, the system enables training data collection through mobile app buttons that trigger 5-second labeled recordings.

1.4.2 Cloud Infrastructure

AWS services provide scalable data processing: IoT Core routes sensor streams to S3 storage and DynamoDB for real-time access. SageMaker extracts 240+ features from 5-second windows and performs XGBoost classification, demonstrating superior performance for stationary sensor data [47].

1.4.3 Mobile Application

An Android application provides user interface and push notifications with sub-2-second latency. The app facilitates labeled data collection for dog-specific model development, addressing the 15-20% performance degradation seen in cross-subject applications.

1.5 Key Technical Innovations

- **Multi-modal Sensor Fusion**: Simultaneous processing of accelerometer, gyroscope, temperature, and audio data captures seizure complexity across physiological systems [11].
- **Dual Connectivity**: WiFi primary and BLE fallback ensure continuous monitoring across environments.
- **Adaptive Machine Learning**: Cloud-based updates enable iterative refinement without firmware changes. Current >95% activity classification

accuracy validates the approach.

- **Real-world Integration**: Addresses practical deployment challenges including power management, security, and user experience.

1.6 Research Contributions

1. Complete IoT architecture demonstrating feasibility of continuous canine monitoring with multi-modal sensors and dual-path communication.
2. Machine learning pipeline achieving >97% accuracy for canine activity classification, validating the behavioral monitoring approach.
3. Real-time event detection with sub-2-second latency, meeting time-critical medical alert requirements.
4. Research platform enabling future seizure detection studies with documented hardware, cloud, and software components.
5. Practical IoT wearable considerations including battery life, connectivity, and intuitive interfaces.

This work bridges the gap between prototypes and deployable systems. While seizure detection awaits clinical data access, successful activity monitoring validates the technical approach. By addressing fundamental engineering challenges, the system provides a foundation for advancing veterinary health monitoring.

1.7 Thesis Organization

The project makes explicit the distinction between demonstrated capabilities (activity classification) and future potential (seizure detection), while showing how design decisions throughout the system support the ultimate clinical application. This assessment strengthens the contribution by focusing on verified achievements while clearly articulating the pathway toward clinical impact.

The remainder of this paper is organized as follows: Chapter 2 provides background on canine epilepsy, IoT architectures, and machine learning for time-series data. Chapter 3 reviews relevant literature on seizure detection systems and wearable sensors. Chapter 4 details the system design and implementation. Chapter 5 describes methodology for data collection and activity classification. Chapter 6 presents results from system validation and field testing. Chapter 7 concludes with discussion of findings, limitations, contributions, and the pathway toward future seizure detection applications.

Chapter 2: Background

This chapter establishes the foundational context necessary for understanding the design and implementation of the smart dog collar system. The intersection of canine epilepsy's clinical challenges, evolving pet health monitoring technologies, and advances in IoT and machine learning creates both the motivation and technical framework for this project. The chapter examines the medical necessity driving continuous monitoring, traces the evolution from basic activity trackers to medical-grade sensors, explores relevant machine learning approaches for time-series data, and addresses the human factors influencing technology adoption in veterinary settings.

2.1 Canine Epilepsy: Clinical Overview and Challenges

Epilepsy affects an estimated 0.75% of the canine population worldwide, although some insurance database analyses report lower incidence (18 per 10,000 dog-years at risk) due to limited coverage in the studied population [21]. This chronic condition carries a grave prognosis; dogs with epilepsy face significantly shortened lifespans, with studies documenting median survival times of merely 1.5 years following initial diagnosis [9]. The therapeutic landscape remains particularly challenging, as drug resistance emerges in nearly one-third of patients despite polypharmacy

approaches [44]. Beyond clinical management difficulties, epilepsy profoundly disrupts the human-animal bond, with caregivers experiencing anticipatory anxiety and implementing substantial lifestyle adaptations to accommodate their pet's unpredictable seizure episodes [42]. These compounding factors, poor survival outcomes, therapeutic limitations, and caregiver stress, create an urgent clinical imperative for innovative monitoring technologies.

The nature of canine epilepsy poses unique monitoring challenges. Breed-specific incidence demonstrates remarkable variation, ranging from 3.0 per 10,000 dog-years at risk in some breeds to 60.3 in others, a 20-fold differential that underscores the genetic complexity underlying seizure susceptibility [21]. This diversity may require flexible detection methodologies rather than one-size-fits-all approaches. Behavioral recognition research consistently documents performance degradation when models trained on one subject are applied to others, with between-subject accuracy declining by approximately 15 percentage points compared to within-subject validation [35, 53]. Such findings advocate for personalized model architectures capable of learning individual-specific seizure signatures. Furthermore, the lack of definitive diagnostic biomarkers requires monitoring systems to accommodate inherent uncertainty in training labels, as epilepsy is often a diagnosis of exclusion following comprehensive metabolic and imaging workups [21].

2.2 Evolution of Pet Health Monitoring Technology

Traditional veterinary monitoring in clinical settings provides only intermittent health snapshots, missing continuous physiological patterns that can characterize epilepsy. The emergence of consumer pet wearables initially focused on basic activity metrics, steps taken, calories burned, and sleep quality, mirroring the functionality of human fitness trackers but failing to address medical monitoring needs.

The transition from activity tracking to medical monitoring represents a fundamental shift in both technological capabilities and regulatory considerations. Early devices relied on single-axis accelerometers with limited sampling rates, sufficient for step counting but inadequate for detecting the complex motor patterns associated with seizures [37]. When veterinary-specific monitoring systems emerged, they targeted dermatological conditions through behavioral detection, demonstrating that collar-mounted sensors could identify medically relevant patterns [56]. Yet these early systems operated at sampling frequencies below the threshold needed to characterize seizure-related movements, which typically manifest in the 3-7 Hz range. Research has since established that effective capture of pathological movement patterns requires sampling rates of 20-50 Hz, balancing detection capability with power consumption constraints [5, 2, 17].

Modern monitoring architectures have converged on distributed processing models, where edge devices handle data acquisition while cloud infrastructure performs complex analytics [29]. Early proof-of-concept studies validated the technical fea-

sibility of automated seizure detection, though initial implementations required invasive procedures that prevented widespread adoption [13]. Current market shows a gap between consumer activity trackers and research-grade systems [11], creating opportunity for medical-grade sensing in consumer form factors.

Agricultural monitoring provides instructive parallels; continuous health tracking systems deployed in livestock operations have demonstrated multi-month battery life and reliable connectivity in challenging environments, suggesting similar architectures could succeed in companion animal applications [51, 15].

2.3 Sensor Technologies for Physiological Monitoring

Multi-modal sensor fusion addresses the fundamental limitation of single-modality approaches by capturing complementary physiological signatures during seizure events [29]. Accelerometry forms the foundation, with three-axis sensors at 20 Hz sampling capturing characteristic 3-7 Hz clonic movements while balancing power consumption [5, 2]. However, play behaviors produce similar acceleration patterns, necessitating additional modalities [10].

Gyroscopic data enables three-dimensional movement reconstruction, particularly valuable for detecting rotational components of tonic-clonic seizures. Integration with accelerometry improves discrimination between voluntary and pathological movements, with hybrid sensor-level fusion for IMU data achieving computational efficiency [54]. Temperature monitoring may capture autonomic dysregulation accompanying seizures, while microphone integration enables detection

of seizure-related sounds and respiratory changes, expanding detection capabilities beyond movement-based approaches [29, 10].

The robustness of multi-modal systems extends beyond improved accuracy. Unsupervised latent space approaches demonstrate resilience to sensor failure, maintaining detection capability despite missing modalities—critical for long-term wearable deployment [43]. Furthermore, environmental noise rejection improves significantly through sensor fusion, addressing a primary limitation of accelerometer-only systems in home environments [54]. These advances in fusion methodology, combined with domain knowledge of seizure pathophysiology, enable detection systems approaching clinical observation sensitivity while maintaining continuous, objective monitoring capabilities.

2.4 Internet of Things (IoT) in Healthcare Applications

The Internet of Things (IoT) refers to networks of interconnected devices that collect, exchange, and act on data without requiring human intervention. The application of IoT principles to healthcare has revolutionized continuous monitoring capabilities, enabling real-time data collection, analysis, and intervention outside traditional clinical settings. IoT architecture patterns for medical devices typically follow a three-tier model: edge devices for data collection, fog or gateway layers for preliminary processing, and cloud infrastructure for advanced analytics and storage. This hierarchical approach balances the competing demands of real-time responsiveness, computational complexity, and power efficiency [4].

Connectivity challenges in IoT healthcare applications are particularly acute for animal monitoring devices, which must operate reliably in diverse environments ranging from indoor home settings to outdoor parks and rural areas. Dual-mode connectivity strategies, combining short-range protocols like Bluetooth Low Energy (BLE) with long-range options such as WiFi or cellular, ensure continuous data transmission regardless of location. Message Queueing Telemetry Transport (MQTT) protocol has emerged as a standard for IoT communications, with studies demonstrating minimal bandwidth overhead and robust security when properly implemented [52, 3].

The choice between batch and stream processing impacts both system latency and computational efficiency, with hybrid approaches often providing the best balance. Time-series databases optimized for sensor data offer superior query performance compared to traditional relational databases, enabling rapid retrieval of historical patterns for model training and validation. Open-source MLOps frameworks like LinkEdge demonstrate viable alternatives to proprietary cloud services, enabling edge device integration without vendor lock-in [41].

The trade-off between edge and cloud processing represents a fundamental architectural decision in IoT system design [4]. Cloud-based processing offers virtually unlimited computational resources, enabling complex machine learning models and comprehensive historical analysis. However, this approach introduces latency and requires reliable connectivity. Edge processing reduces latency and bandwidth requirements but constrains model complexity due to limited computational resources on embedded devices. Hybrid architectures that perform initial filtering

and feature extraction at the edge while leveraging cloud resources for complex analysis offer a pragmatic compromise, particularly relevant for seizure detection where both rapid response and sophisticated analysis are required. Wildlife monitoring systems have demonstrated this approach successfully, with on-board processing reducing transmission needs while cloud resources enable population-level insights [49]. Recent advances in edge-cloud collaborative architectures show that splitting inference between edge and cloud can reduce compute requirements by 47% while maintaining performance [57].

2.5 Machine Learning for Time-Series Sensor Data

The application of machine learning to time-series sensor data presents unique challenges distinct from traditional computer vision or natural language processing tasks. Seizure detection exemplifies a rare event detection problem, where the vast majority of collected data represents normal behavior, creating severe class imbalance that can bias traditional classifiers [47]. This imbalance necessitates specialized techniques including synthetic minority oversampling, cost-sensitive learning, and anomaly detection approaches that model normal behavior and flag deviations. Studies in animal behavior analysis have shown that machine learning can succeed with as few as 20 examples per class when proper feature engineering is applied [53], though imbalanced datasets benefit from techniques like SMOTE which can improve detection performance by 3x [53, 1].

The crowdsourced development of seizure detection algorithms through com-

petitions has demonstrated the power of open science, with top algorithms achieving AUC scores exceeding 0.98 on validation datasets [7]. These competitions revealed that ensemble methods combining multiple algorithmic approaches consistently outperform single models, suggesting that practical seizure detection systems should incorporate diverse detection strategies. XGBoost has emerged as particularly effective for animal health monitoring across multiple studies [16, 32, 38], achieving accuracy rates of 74-87% while maintaining computational efficiency suitable for real-time deployment.

Feature engineering transforms raw sensor signals into meaningful representations for machine learning algorithms. Time-domain features capture statistical properties like mean and variance, while frequency-domain features reveal periodic patterns characteristic of different activities [27]. Studies demonstrate that combining features from multiple sensors improves detection accuracy, though this increases computational complexity [53, 8]. Importantly, reducing feature sets from hundreds to fewer than twenty maintains performance while enabling real-time processing on resource-constrained devices [2, 14]. The challenge lies in selecting features that are both computationally efficient and clinically interpretable for veterinary adoption [46].

Window size selection critically impacts performance, with studies demonstrating that 1-2 second windows provide optimal trade-offs between detection speed and accuracy for activity recognition tasks [8]. However, 5-second windows have shown superior performance for complex behavior classification, achieving F_1-scores of 0.92 ± 0.02 compared to 0.87 ± 0.02 for 1-second windows [2, 59]. Deep

learning approaches promise end-to-end learning from raw sensor data, but their black-box nature and computational requirements often make traditional feature engineering with interpretable models more suitable for medical applications where understanding decision rationale is crucial. Self-supervised learning frameworks like Selfee demonstrate that millions of unlabeled frames can be leveraged for feature extraction [26], addressing the limited labeled seizure data challenge.

Recent comparisons between machine learning and deep learning models on stationary time-series data reveal surprising results. For highly stationary sensor data, traditional algorithms like XGBoost significantly outperform deep learning approaches including LSTMs, while requiring substantially less computational resources [47]. This finding has profound implications for collar-based systems where sensor positions remain fixed and data exhibits stationary characteristics. Random forests and tree-based methods consistently outperform other classifiers for animal behavior analysis [53, 17], further supporting the selection of XGBoost for seizure detection. Extremely Randomized Trees have shown particular promise for real-time classification on mobile devices [59].

Real-time inference on battery-powered devices imposes dual constraints of latency and energy efficiency. Lightweight models must balance accuracy with sub-second response times, often using hierarchical approaches where simple models screen for suspicious patterns before triggering complex analysis. Recent advances in early-exit neural networks demonstrate that adaptive computational complexity can reduce energy consumption by 55% while maintaining detection accuracy [22]. However, long-term deployments face model drift, with performance degrading

5-8% over 6 months, necessitating adaptive learning capabilities [35].

2.6 Human-Animal Bond and Technology Acceptance

The human-animal bond profoundly influences how caregivers engage with pet health technology. Owners of epileptic dogs report significant lifestyle modifications, including altered work schedules, restricted travel, and constant vigilance [42]. Seizures create acute physiological stress for both dogs and caregivers, with owner cortisol levels increasing 40.5% at 20 minutes post-seizure and dog cortisol levels increasing 265.1% at 40 minutes post-seizure [40]. This combination of chronic lifestyle burden and acute physiological stress underscores the need for reliable monitoring systems.

Survey data from 231 owners of epileptic dogs reveals strong receptiveness to monitoring technology [10]. Ninety percent believe seizure detection devices would improve their dog's management, with those already tracking seizures being 4.2 times more likely to show confidence in using such devices. Remarkably, 52% of owners would accept any false alarm rate if no seizures were missed, demonstrating that comprehensive detection takes priority over perfect specificity.

Design requirements derived from owner surveys emphasize practical considerations [10]. Seventy-one percent prefer collar-mounted sensors, while only 23% would accept intracranial implants despite potentially superior performance. Device comfort ranks as the primary concern, with owners expressing worry about skin irritation and interference with normal activities like swimming. Battery life

expectations reflect real-world usage patterns, with owners requiring minimum 3-day operation to accommodate weekend trips. Integration with existing veterinary care workflows enhances acceptance among veterinary professionals, who serve as crucial intermediaries in technology adoption decisions [42].

2.7 Summary

The combination of clinical need, technological capability, and user acceptance creates a unique opportunity for advanced epilepsy monitoring systems in veterinary medicine. The high prevalence and poor outcomes associated with canine epilepsy establish clear medical justification for continuous monitoring solutions. Recent advances in sensor technology, IoT architectures, and machine learning provide the technical foundation for systems capable of reliable seizure detection in naturalistic settings. Understanding the human-animal bond and designing for both pet comfort and owner usability ensures that technically capable systems translate into real-world impact. This background establishes the multifaceted context within which the smart collar system operates, bridging veterinary medicine, computer science, and human factors to address a significant unmet need in companion animal healthcare.

Chapter 3: Literature Review

The development of effective canine health monitoring systems requires understanding advances across multiple domains: veterinary seizure detection, wearable sensor technologies, machine learning for time-series analysis, and IoT system architectures. This chapter synthesizes relevant literature to establish the current state-of-the-art, identify technological gaps, and position this project's contributions within the broader research landscape. The review progresses from existing canine epilepsy detection approaches through general animal monitoring technologies, examines applicable machine learning techniques, and concludes with architectural considerations for scalable deployment.

3.1 Canine Epilepsy Detection Systems

Current approaches to canine epilepsy detection vary from invasive research prototypes to consumer-grade activity monitors, with each system making different compromises between accuracy, practicality, and cost. This review examines existing systems' capabilities and shortcomings to identify opportunities for technological advancement.

Approaches investigating intracranial EEG systems have demonstrated high sensitivity to seizure detection but required invasive implantation [13]. The seizure

advisory system (SAS) in four epileptic dogs successfully detected prolonged seizures, enabling timely rescue medication intervention [13]. The system achieved detection rates exceeding 90% for generalized tonic-clonic seizures, with lead times of 30-60 seconds before clinical manifestation. However, the requirement for surgical implantation, risk of infection, and costs exceeding $25,000 per animal limit this approach to controlled research settings. Crowdsourced algorithm development using data from a similar human system achieved AUC scores of 0.983 ± 0.004, demonstrating the potential of machine learning approaches for seizure detection [7].

Subscalp EEG represents a significant improvement over intracranial systems, offering less invasive monitoring through subcutaneous electrodes that require only minimal surgical intervention. Novel wireless devices enable continuous monitoring in unsedated dogs, eliminating the need for clinical sedation while still capturing high-quality EEG signals. Recent advances in flexible electronics and biocompatible materials have extended recording durations beyond six months with reduced infection risk compared to intracranial approaches [34]. Despite these advantages, subscalp systems remain impractical for widespread adoption—they still require veterinary surgical procedures, incur substantial costs, and face ongoing challenges with electrode displacement and skin irritation that compromise both canine comfort and owner accessibility.

Accelerometry-based commercial devices have so far shown limited performance. Whistle achieved only 18.6-22.1% sensitivity with high false positives during play. The study evaluated 16 dogs with idiopathic epilepsy over six months,

revealing that single-axis accelerometry cannot distinguish between normal high-activity behaviors and seizure-related movements. False positive rates exceeded three per day during active play, making the system impractical for real-world use [37]. Vetrax demonstrated 99.24% accuracy for scratching but its 10 Hz sampling limits seizure detection. While excellent for detecting repetitive behaviors associated with dermatological conditions, the system's low sampling frequency cannot capture the 3-7 Hz oscillations characteristic of clonic seizure phases [56]. This implies that single-modality reliance is not effective in capturing the full range of seizure manifestations.

Multi-modal approaches substantially improve performance. Heart rate variability (HRV) monitoring captures autonomic changes and personalized algorithms achieved 75-100% sensitivity with 0.5 false alarms per day in humans. The wide sensitivity range reflects individual variability in autonomic seizure manifestations, with some patients showing pronounced tachycardia preceding seizures while others exhibit minimal heart rate changes [25]. Smartphone-based mechanocardiography has demonstrated feasibility for canine cardiac monitoring, achieving heart rate detection within 5 bpm accuracy [30], suggesting potential for non-invasive autonomic monitoring. Electrodermal activity (EDA) detects sympathetic activation; Electromyography (EMG) identifies distinct muscle patterns. Electrodermal activity increases 1-3 minutes before seizure onset in 84% of cases, providing a complementary detection channel to movement-based approaches. Surface EMG can differentiate between voluntary muscle contractions and the synchronous firing patterns of seizure activity [11]. No existing system has been explored that

combines all modalities in a practical, non-invasive form factor.

3.2 Wearable Sensor Technologies for Animal Monitoring

Advances in miniaturization, power efficiency, and wireless communication technologies have driven the evolution of wearable sensors for animal monitoring. From simple pedometers to sophisticated multi-sensor platforms, these devices have transformed our ability to monitor and quantify animal behavior and health status continuously.

Vetrax pioneered medical-grade companion animal monitoring with 10 Hz three-axis accelerometry. Clinical validation showed scratching and head shaking significantly predicted veterinarian assessments ($\beta = 0.176$, $p = 0.008$ and $\beta = 0.197$, $p < 0.001$) [56]. The system's success in detecting pruritic behaviors demonstrated the feasibility of collar-based monitoring for medical applications, though its fixed algorithm design prevents adaptation to individual animal variations or extension to seizure detection.

The application of deep learning to animal monitoring has enabled sophisticated behavioral classification systems. The search and rescue dog monitoring platform developed by Kasnesis et al. [28] integrates Arduino-based wearables with cloud infrastructure, achieving 94% accuracy in distinguishing between searching, resting, and signaling behaviors. This system demonstrates the feasibility of real-time behavior analysis using resource-constrained edge devices paired with cloud computing resources. Studies utilizing CNN architectures for canine activity classi-

fication achieved 96-99% accuracy with collar-mounted accelerometers, confirming the viability of collar placement for medical monitoring [35, 1]. Dual-sensor systems incorporating both neck and tail-mounted IMUs have shown promise for both behavioral and emotional state monitoring [1].

Collar mounting optimizes signal quality versus practicality for long-term wear. Studies comparing sensor placement locations found that collar-mounted accelerometers capture 87% of the movement information available from multi-site sensor arrays, while maintaining user compliance exceeding 95% for continuous wear [12]. Multi-modal fusion improves accuracy over single sensors, with accelerometer-gyroscope combinations improving classification by 8-12% over single modality approaches [35, 54, 12]. Sampling at 20-50 Hz captures relevant movement patterns, with studies validating that 20 Hz specifically enables fall detection with Se=0.9983 and Sp=0.9996 [5]. Window sizes of 1-2 seconds optimize speed-accuracy trade-offs [8], although 5-second windows achieve superior F1-scores for complex behaviors [2, 59]. Higher sampling frequencies enable capture of fine-grained movement patterns but increase power consumption quadratically, necessitating careful optimization for battery-powered devices. For seizures, 20 Hz may adequately capture 3-7 Hz clonic movements while enabling multi-day battery life.

Power efficiency remains a critical challenge for continuous monitoring devices. Recent advances in sensor-aware machine learning demonstrate that adaptive sampling and early-exit neural networks can reduce energy consumption by 55% while maintaining detection accuracy [22]. These techniques dynamically adjust sampling rates and processing complexity based on detected activity patterns, extend-

ing battery life from days to weeks for typical usage scenarios. Edge inference on embedded processors achieves sub-50ms latency with model compression techniques reducing size by 10x while maintaining accuracy within 2% of full models [35, 15]. Successful deployment of livestock monitoring systems demonstrates months of battery life using BLE advertising and on-device aggregation [51].

3.3 Machine Learning for Seizure Detection

The application of machine learning to seizure detection has evolved from simple threshold-based algorithms to sophisticated deep learning models that can learn complex spatio-temporal patterns. The availability of large annotated datasets and collaborative algorithm development has accelerated this evolution.

A crowd-sourced algorithm development through a Kaggle.com competition involved 200 teams worldwide in creating seizure detection algorithms using intracranial EEG data from dogs and humans. The competition provided 48 hours of continuous intracranial electroencephalography recordings from four dogs with naturally occurring epilepsy and eight human patients, representing the largest open dataset for seizure detection at the time. The top performers achieved an AUC of 0.983 ± 0.004 on validation datasets [7]. The top performing models included random forests, neural networks, and ensembles. An analysis of winning solutions revealed that feature engineering remained crucial, with spectral power in specific frequency bands, phase-amplitude coupling, and network connectivity measures providing the most discriminative information.

Traditional machine learning approaches have shown surprising advantages over deep learning for certain seizure detection scenarios. Santoro et al. [47] demonstrated that XGBoost outperformed LSTM networks by 23% on stationary time-series data while requiring 7x less computational resources. This finding is particularly relevant for collar-based systems where sensor positions remain fixed relative to the animal, creating highly stationary data streams. The superior performance of gradient boosting methods likely stems from their ability to capture non-linear relationships without requiring the temporal modeling complexity of recurrent networks. This aligns with findings from animal behavior studies where random forests consistently outperformed other classifiers [53, 17], supporting tree-based methods for seizure detection. XGBoost has demonstrated particular success in livestock health monitoring, achieving 86% AUROC for locomotor disease detection [16], 87% F_1-score for dairy cow behavior recognition with only 5 Hz sampling [32], and 74.2% accuracy for multi-behavior classification [38].

Feature engineering strategies have evolved to capture increasingly sophisticated biomarkers of seizure activity. Modified log energy entropy, when combined with temporal derivatives, achieved 98.84% sensitivity and 99.51% specificity in detecting canine seizures using intracranial EEG data [27]. These features effectively capture the transition from normal background activity to the synchronous patterns characteristic of seizure onset. Domain-specific feature engineering has been shown to improve accuracy by 20% compared to generic features, emphasizing the importance of veterinary-informed feature design [53]. Studies demonstrate that reducing the number of features from 500+ to 15 or fewer maintains performance

while enabling real-time processing [2, 14]. Notably, neighborhood component analysis achieves 96.3% accuracy using only nine features [14]. Biomechanically informed features provide clinical interpretability, which is crucial for wide-spread veterinary adoption [46].

Window size optimization studies reveal critical trade-offs between detection latency and accuracy. Banos et al. [8] systematically evaluated window sizes from 0.25 to 7 seconds across 33 different activities, finding that 1-2 second windows provide optimal performance for most movement-based classifications. For seizure detection, this translates to potential alert latencies under 2 seconds from onset, sufficient for caregiver notification and intervention. However, 5-second windows have demonstrated superior performance for complex behavior classification [2, 59], suggesting that seizure detection may benefit from longer temporal context. Shorter windows enable faster detection but may miss complex seizure evolution patterns, while longer windows improve accuracy at the cost of delayed alerts.

Real-world deployment considerations extend beyond algorithm performance. The Bongers et al. survey reveals that 88% of owners believe automated detection would provide more accurate seizure frequency knowledge than manual counting, addressing the known issue of underreporting in seizure diaries [10]. This highlights the clinical value of continuous monitoring beyond the accuracy of detection. Long-term monitoring studies demonstrate the challenge of model drift, with performance degrading 5-8% over 6 months, necessitating adaptive learning systems. Individual adaptation through dog-specific fine-tuning has shown to improve accuracy by 12%, supporting personalized model approaches [35]. Real-world per-

formance typically achieves only 82% of laboratory results [53], emphasizing the importance of validation in field tests. The consistency of human labeling for complex behaviors reaches only 85% [53], highlighting the inherent uncertainty of labeling that systems must accommodate.

3.4 Advanced Sensor Fusion Techniques

Recent advances in sensor fusion have demonstrated significant improvements in activity recognition and anomaly detection for wearable systems. Decision-level fusion approaches outperform feature-level fusion (74.4% vs 67.4% accuracy) but require increased computational resources. Kalman filter fusion achieves the best accuracy-speed trade-off, with 75.4% accuracy in 61.7ms processing time, making it viable for real-time seizure detection [54]. The complementary nature of accelerometer and gyroscope data, with accelerometers capturing linear motion and gyroscopes detecting rotational movements, validates the multisensor approach for capturing complex seizure patterns. However, feature-level fusion has demonstrated superior performance in other domains, with XGBoost achieving 94.02% accuracy for multimodal classification when combining optical and sensor data [58].

Latent space sensor fusion represents a promising approach for handling the limited labeled seizure data available in veterinary applications. Piechocki et al. demonstrated a two-stage approach using unsupervised feature learning followed by supervised fusion, achieving robust performance even with 50% sensor dropout

[43]. This technique's ability to handle missing modalities gracefully addresses a critical reliability concern for long-term monitoring devices. The approach compresses sensor data to latent dimensions 4-8x smaller than raw features, reducing transmission bandwidth and storage requirements while maintaining detection accuracy. Multimodal VAEs can pre-train on unlabeled normal activity data before fine-tuning on limited seizure examples [43], addressing the scarcity of labeled training data.

Environmental robustness emerges as a critical factor for home deployment. Sensor fusion techniques have been shown to reduce false positives from external vibrations by 60%, addressing a major limitation of single-sensor approaches [54]. Hybrid fusion architectures, combining sensor-level fusion for IMU data with decision-level fusion for multimodal integration, align well with the computational constraints of embedded systems while maximizing detection performance. Two-level ensemble approaches using stacking and voting have improved accuracy by 2.48% over single modality systems [12].

3.5 Cloud Architecture and Edge Computing Trade-offs

The architectural decisions in IoT-based health monitoring systems have a significant impact on performance, scalability, and user experience. Recent systematic reviews of edge computing in healthcare IoT reveal evolving best practices and persistent challenges [4].

Edge computing offers several advantages for seizure detection applications: re-

duced latency (sub-100ms response times) [15], enhanced privacy through local processing, and resilience to connectivity disruptions [4]. However, edge devices face constraints in computational capacity, limiting model complexity to lightweight algorithms. Cloud computing provides virtually unlimited computational resources, enabling sophisticated deep learning models and population-level analytics, but introduces latency (typically 200-500ms) [3] and requires reliable connectivity. Recent advances in edge-cloud collaborative architectures demonstrate that splitting inference between edge and cloud can achieve 47% compute reduction with minimal quality loss [57], enabling partial edge processing for seizure detection.

Hybrid architectures emerge as the optimal solution, performing time-critical detection at the edge while leveraging cloud resources for model updates and complex analytics. The SageMaker Model Monitor framework [39] exemplifies this approach, enabling continuous model performance monitoring and automated retraining triggers when drift is detected. For seizure detection, this allows personalized model adaptation as individual seizure patterns evolve over time. Wildlife monitoring systems have successfully demonstrated this hybrid approach, with onboard processing reducing transmission needs while cloud resources enable population-level insights and transfer learning across individuals [49]. Open-source MLOps frameworks like LinkEdge provide vendor-agnostic alternatives for edge device management [41].

AWS implementations demonstrate practical benefits. SageMaker handles versioning, A/B testing, and scaling [39]. Serverless endpoints optimize costs, with studies showing 80% reduction compared to always-on deployments [57]. IoT

Core integration enables secure over-the-air updates. MQTT protocol provides lightweight, reliable communication with minimal overhead, as validated by security studies showing 99.96% anomaly detection accuracy while maintaining bandwidth efficiency [52]. Performance comparisons demonstrate MQTT/TCP achieving 290.5ms average latency, optimal for real-time applications [3].

3.6 Clinical Context and User Requirements

The 10-fold variation in epilepsy incidence between breeds (5.7-60.3 per 10,000 dogs) [21] suggests underlying physiological differences that may affect seizure presentation, requiring monitoring systems capable of adapting to breed-specific characteristics. Median 1.5-year survival emphasizes intervention urgency [9]. This clinical variation necessitates monitoring systems capable of adapting to individual phenotypes, as between-subject accuracy typically falls 15% below within-subject performance [35, 53].

Owner surveys provide crucial insights into user requirements and market readiness. Among 231 respondents caring for epileptic dogs, the unpredictability of seizures emerged as the primary management challenge. Word cloud analysis of factors that would ease the burden of leaving dogs alone revealed "monitoring," "camera," "knowing," and "alert" as dominant themes, emphasizing the desire for real-time awareness. Owners currently keeping seizure logs were 4.2 times more likely to express confidence in using detection devices, suggesting that those most engaged in seizure management recognize the value of automated systems [10].

The physiological impact of witnessing seizures has been quantified through salivary cortisol measurements [40]. Caregivers experience a 40.5% increase in cortisol levels 20 minutes post-seizure, with those scoring higher on neuroticism scales showing elevated responses at both sampling points. This acute stress response, combined with chronic management burden documented by Pergande et al. [42], creates a compelling case for systems that can provide remote monitoring and rapid alerts.

Survey data reveals specific design requirements and market validation. Ninety percent believe detection devices would improve management; comfort ranks as the primary concern. Seventy-one percent prefer collar-mounted devices, while minimum battery life requirement is three days. False alarm tolerance varies: 52% accept any false alarm rate if no seizures missed, while 30% prefer less than one per week. Market validation shows 78% would purchase systems under $300 with monthly fees below $30, contingent on clinical validation [10].

3.7 Treatment Challenges Justifying Continuous Monitoring

The limitations of current pharmacological treatments create compelling justification for advanced monitoring systems. Drug resistance affects 30% of epileptic dogs despite polypharmacy. This drug-resistant population faces particularly poor prognosis, with increased risk of cluster seizures and status epilepticus [44]. Continuous monitoring may enable early cluster detection for timely rescue medication intervention, potentially preventing progression to life-threatening status epilepti-

cus. Objective quantification aids in medication optimization, as owner-reported seizure diaries may underestimate true seizure burden due to missed nocturnal or subtle events.

Surgical options remain limited; corpus callosotomy and focal resection require extensive evaluation [19]. Current diagnostic limitations mean many potentially surgical candidates go unidentified. Long-term monitoring could identify surgical candidates through detection of focal onset patterns missed by brief EEG studies.

The promise of adaptive treatment algorithms based on individual seizure patterns represents a paradigm shift from current fixed-dosing regimens [25]. Machine learning models trained on individual patient data could potentially identify pre-ictal patterns unique to each dog, enabling preemptive interventions. Circadian and multi-day seizure cycles, well-documented in human epilepsy but poorly studied in dogs due to monitoring limitations, could inform chronotherapy approaches with medication timing optimized to individual patterns [25]. Temporal pattern analysis using Hidden Markov Models has successfully captured behavior sequences in animal studies [53], suggesting potential for modeling seizure progression stages. Given this, continuous monitoring with machine learning offers hope for drug-resistant cases [44].

3.8 Comparative Analysis

Table 3.1 synthesizes performance metrics and limitations across representative seizure detection approaches, highlighting the performance gap between research

prototypes and commercial systems.

Table 3.1: Comparison of Seizure Detection Approaches

Device/Study	**Sensing Method**	**Performance**	**Key Limitations**
Whistle [37]	3-axis accelerometer (collar-mounted)	18.6-22.1% sensitivity 0.096/day false rate	Single modality, low sensitivity for seizure detection
Vetrax [56]	3-axis accelerometer (collar-mounted)	99.24% (scratching) 99.56% (head shake)	Validated for pruritus only, not seizure detection
SAS [13]	Intracranial EEG (16 channels)	Successful alerts for prolonged seizures	Invasive, requires surgery, research prototype only
Deep Learning [35]	Multi-sensor collar (100 Hz sampling)	96-99% (activities)	Not validated for seizures, high power consumption
XGBoost [16, 32]	Neck/leg sensors (5-20 Hz sampling)	86-87% (health issues)	Livestock-specific, requires adaptation
Proposed Collar	**Multi-modal: Accel/Gyro/Temp/Audio**	**Target: >90% sensitivity**	**Battery life, connectivity requirements**

The comparative analysis reveals a clear trade-off between detection accuracy and system invasiveness. While intracranial EEG systems achieve near-perfect sensitivity, their clinical applicability remains limited. Consumer devices prioritize usability but sacrifice detection performance to unacceptable levels. Recent deep learning approaches show promise but require validation for seizure-specific applications.

Table 3.2 summarizes the performance of various machine learning algorithms on stationary sensor data, highlighting the computational efficiency advantages of tree-based methods. The proposed multi-modal collar aims to leverage this by combining non-invasive sensing with sophisticated signal processing using these algorithms.

Table 3.2: ML Algorithm Performance on Stationary Time-Series Data and Animal Behavior

Algorithm	**Computational Efficiency**	**Key Finding**	**Reference**
XGBoost	7x faster than LSTM	Outperforms on stationary sensor data, 86-94% accuracy	[47, 16] [32, 38]
Random Forest	10x faster than DL	Best for animal behavior, feature importance analysis	[53, 52] [2, 46]
LSTM	Baseline	Tends to over-smooth sharp transitions	[47, 12]
CNN with Early Exit	55% energy savings	Maintains 96% accuracy for epilepsy detection	[22]
Kalman Fusion	Real-time capable	75.4% accuracy in 61.7ms	[54]
Neural Networks	Quantizable to edge	94-97% on embedded systems	[15, 1] [14]

3.9 Research Gaps and Opportunities

After an exhaustive search, we were unable to find an existing system that combines multimodal sensing (accelerometer, gyroscope, temperature, and audio) in a practical collar form with advanced sensor fusion techniques and built-in labeled data collection. This represents a significant opportunity, as each sensing modality captures distinct physiological manifestations of seizures that, when combined through appropriate fusion algorithms such as latent space approaches [43] or Kalman filtering [54], could achieve detection performance approaching that of invasive systems. Commercial single-modality devices show unacceptable performance. Research systems require invasive procedures or cumbersome equipment.

The absence of cloud-based adaptive learning prevents improvement with accumulated data. Current systems use fixed algorithms that cannot adapt to individ-

ual seizure patterns or evolve as the disease progresses. Cloud-based learning enables continuous model refinement using accumulated data from individual animals and population-level insights, as demonstrated in wildlife monitoring applications [49]. Transfer learning across individuals could address the limited seizure data available for individual dogs [35]. Active learning with uncertainty sampling could accelerate annotation of ambiguous events [49]. Fixed algorithms cannot adapt to individual patterns. Dual-connectivity (WiFi/BLE) addressing real-world challenges is lacking. Real-world deployment must account for varied environments where dogs live and play, from urban apartments with reliable WiFi to rural properties with limited connectivity. Current systems fail to provide open, extensible veterinary IoT platforms. Proprietary systems limit integration and research access.

The integration of edge and cloud computing presents unexplored opportunities for veterinary applications. While human health monitoring has embraced hybrid architectures, veterinary systems remain predominantly edge-only or cloud-only. A thoughtfully designed hybrid system could perform rapid seizure detection locally while leveraging cloud resources for continuous learning and population health insights. This approach would address both the real-time requirements of seizure alerting and the long-term benefits of adaptive algorithms. Energy-aware computing techniques could extend battery life dramatically while maintaining detection performance [22]. Edge-cloud collaborative approaches reducing compute by 47% [57] remain unexplored for veterinary applications.

Validation challenges persist, with real-world performance typically achieving

only 82% of laboratory results [35, 53]. Inter-observer agreement for complex behaviors reaches only 85%, highlighting the need for robust ground truth establishment [53]. Semi-supervised learning approaches incorporating periodic owner feedback could maintain performance while reducing labeling burden. Self-supervised learning frameworks demonstrate promise for leveraging unlabeled data [26], critical given limited seizure examples.

3.10 Summary

The literature reveals that veterinary epilepsy monitoring significantly lags human applications, where crowdsourced algorithms achieve over 98% accuracy [7]. Commercial devices demonstrate market demand but deliver poor performance (<23% sensitivity) through single-sensor approaches. This performance gap represents both a clinical failure and a business opportunity, as surveys indicate pet owners would readily adopt devices achieving >80% sensitivity even with occasional false alarms. Notably, owners tolerate high false alarm rates (52% accept any if no seizures are missed), prioritizing comprehensive detection over perfect specificity [10]. XGBoost and other tree-based methods outperform deep learning for stationary sensors while energy-optimization techniques promise 55% battery life extension [47, 53, 22, 16, 32, 38].

The collar proposed in this project combines proven multi-modal sensing, cloud-based adaptive learning, and open architecture principles from successful human systems with veterinary-specific requirements. By synthesizing insights from wear-

able sensor design, machine learning for time-series data, advanced sensor fusion techniques, and IoT architectural patterns, this system addresses the limitations of existing approaches while remaining practical for everyday use. Recent advances in latent space fusion [43], energy-aware computing [22], edge-cloud collaboration [57], and hybrid architectures [49, 41] provide the technical foundation for achieving clinical-grade performance in a consumer-friendly form factor. Successful deployments in livestock monitoring [51, 15, 32] and activity recognition [35, 1] demonstrate the viability of continuous health monitoring using similar architectures. This synthesis builds upon this foundational knowledge to establish a platform for data-driven epilepsy management, transforming reactive response into proactive care enabled by continuous intelligent monitoring.

The combination of technological capabilities, such as minimal sensors, efficient edge processors, ubiquitous connectivity, and sophisticated machine learning, with the pressing clinical need creates an unprecedented opportunity to enhance epilepsy care for companion animals. The following chapters detail the design and implementation of a system that creates this potential.

Chapter 4: System Architecture

This chapter presents a comprehensive overview of the smart collar system architecture, detailing the technical implementation of each component and their interactions. The system implements a three-tier Internet of Things (IoT) architecture designed for reliability, scalability, and real-time canine health monitoring with seizure detection capabilities.

As illustrated in Figure 4.1, the distributed architecture spans: (1) an embedded sensor platform worn by the dog featuring the Arduino Nano RP2040 Connect with integrated sensors, (2) a React Native mobile application serving as both user interface and BLE communication gateway, and (3) AWS cloud infrastructure providing data storage, real-time processing, and machine learning inference through services including IoT Core, S3, DynamoDB, and SageMaker.

The system addresses several key engineering challenges:

- **Continuous Monitoring**: Balancing comprehensive sensor coverage with power consumption constraints, currently achieving 8 hours of operation with activity-triggered recording sessions
- **Reliable Communication**: Ensuring data delivery despite intermittent connectivity through dual-mode WiFi and BLE communication with automatic failover

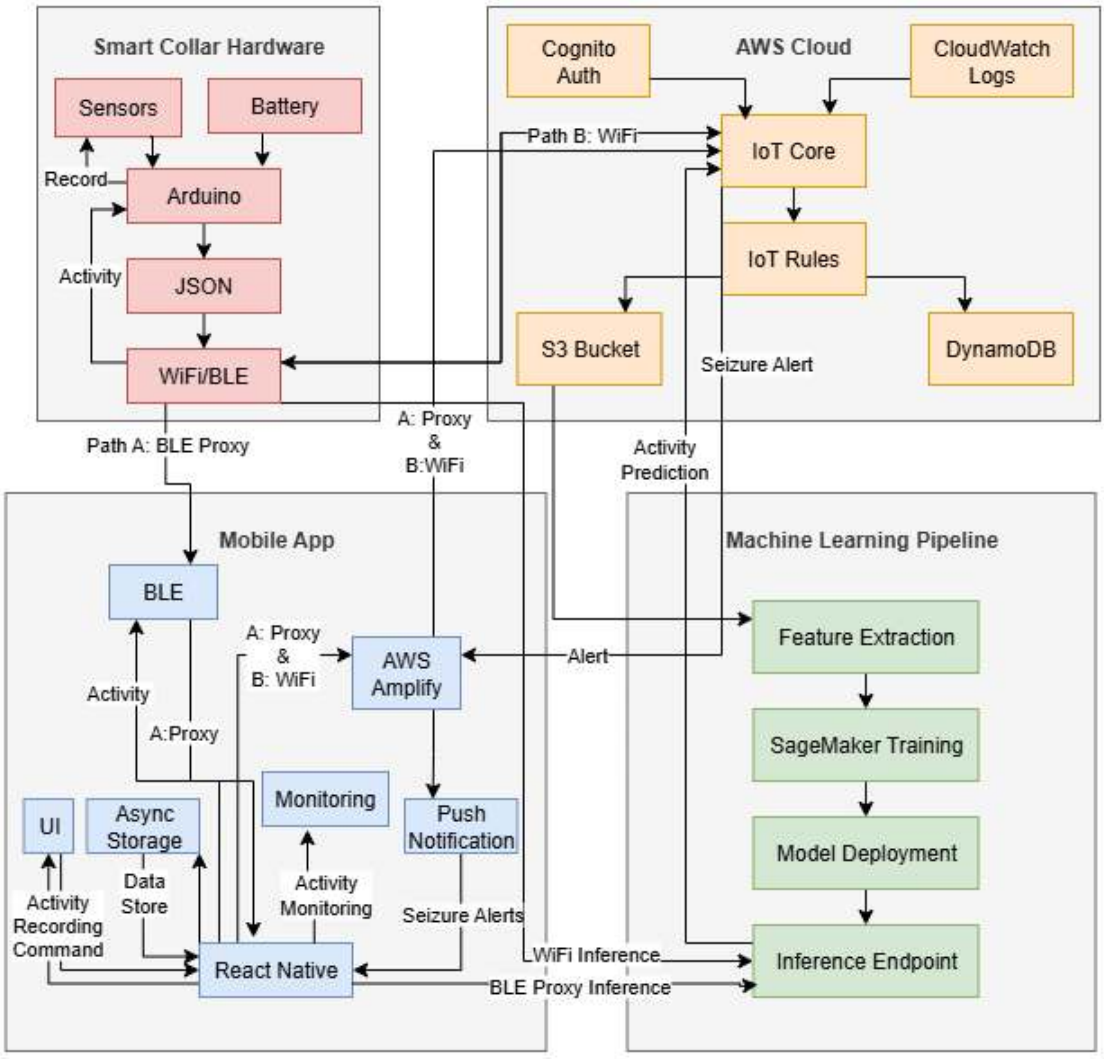


Figure 4.1: Complete smart collar system architecture showing data flow from sensors to cloud services

- **Real-time Processing**: Achieving sub-2-second latency for critical health alerts in WiFi mode, with graceful degradation to < 10 seconds in BLE proxy mode [13]

- **Scalability**: Supporting multiple devices without architectural modifications through AWS IoT Core's message routing capabilities
- **Data Integrity**: Preventing data loss through redundant communication paths and persistent storage at multiple system levels
- **Regulated Data Collection**: Implementing a controlled pipeline where the mobile app initiates labeled data collection and the Arduino ensures synchronized sensor capture, creating high-quality training datasets for both initial model development and future personalized model refinement

4.1 Embedded Hardware Platform

The embedded system design balances processing capabilities, power efficiency, and form factor constraints to create a wearable platform suitable for continuous canine monitoring.

4.1.1 Core Computing Platform

The Arduino Nano RP2040 Connect serves as the computational heart of the smart collar, selected for its combination of processing power, wireless capabilities, and integrated sensors. As shown in Figure 4.2, this platform integrates:

- **Dual-core Processing**: Raspberry Pi RP2040 with dual ARM Cortex-M0+ cores operating at 133 MHz enable parallel execution of sensor sampling and

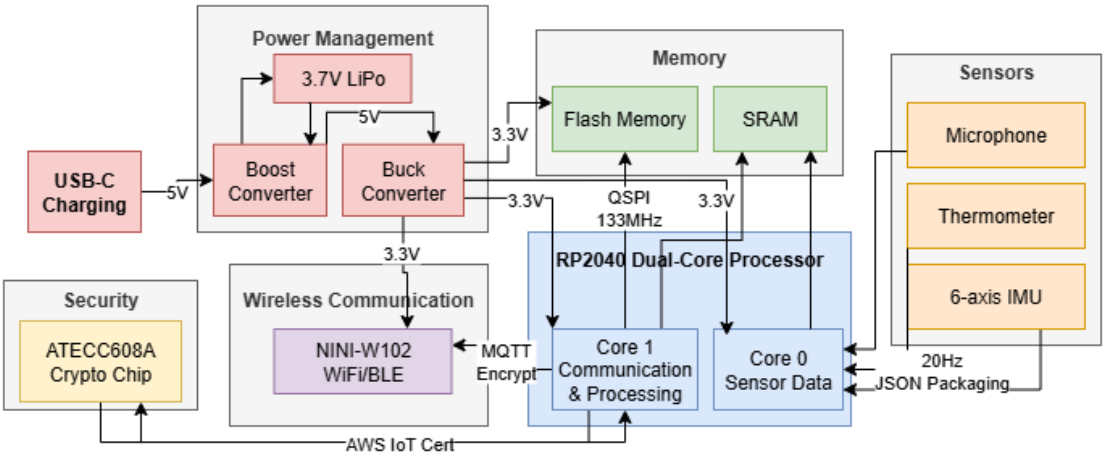


Figure 4.2: Hardware architecture of the smart collar embedded system

communication tasks [6].

- **Memory Resources**: 264KB on-chip SRAM accommodates sensor buffers and processing pipelines, while 16MB external flash (AT25SF128A) stores firmware and temporary data during connectivity outages [6].

- **Wireless Connectivity**: u-blox NINA-W102 module provides both WiFi (802.11 b/g/n) and Bluetooth Low Energy 4.2, enabling flexible communication strategies [6].

- **Hardware Security**: ATECC608A cryptographic co-processor stores device certificates and performs TLS operations without exposing private keys, supporting SHA-256, HMAC, and AES-128 encryption [6].

- **Compact Form Factor**: 45mm × 18mm board dimensions minimize collar bulk while maintaining functionality [6].

4.1.2 Communication Protocols

The system employs two primary communication protocols:

- **MQTT (Message Queuing Telemetry Transport)**: A lightweight publish-subscribe protocol operating over TCP/IP, designed specifically for IoT applications. MQTT's minimal overhead (2-byte fixed header) and quality-of-service guarantees make it ideal for resource-constrained devices requiring reliable message delivery.

- **BLE UART**: A serial communication profile over Bluetooth Low Energy, enabling smartphone-mediated connectivity when WiFi is unavailable. The NINA-W102 module handles BLE 4.2 communications independently from the main processor [6].

4.1.3 Sensor Suite Design

The sensor selection leverages onboard sensors to capture physiological signals relevant to activity classification and seizure detection:

Table 4.1: Sensor specifications and sampling rates

Sensor	**Range**	**Resolution**	**Sample Rate**
Accelerometer	±16g	16-bit	20 Hz [12]
Gyroscope	±2000 dps	16-bit	20 Hz [12]
Temperature	-40 to +85°C	0.01°C	20 Hz
Microphone	20Hz-20kHz	16-bit PDM	16 kHz

4.1.3.1 Inertial Measurement Unit (LSM6DSOXTR)

The onboard 6-axis IMU captures both linear acceleration and rotational velocity, critical for detecting activity patterns and seizure-related movements [6]:

- **Accelerometer**: ±2/±4/±8/±16g selectable full scale range [6]
- **Gyroscope**: ±125/±250/±500/±1000/±2000 dps selectable full scale range [6]
- **Machine Learning Core**: Integrated finite state machine and decision tree processing capabilities [6]
- **Embedded Temperature**: Integrated temperature sensor for auxiliary monitoring [6]
- **Advanced Features**: Programmable finite state machine for accelerometer, gyroscope, and external sensors [6]

4.1.3.2 Audio Monitoring (MP34DT06JTR)

The onboard MEMS microphone captures vocalization patterns and environmental sounds [6]:

- **PDM Interface**: Digital output eliminates analog noise concerns
- **Sensitivity**: -26 dBFS ±1 dB sensitivity [6]
- **Signal Quality**: 64 dB signal-to-noise ratio [6]

- **AOP**: 122.5 dBSPL acoustic overload point [6]
- **Directivity**: Omnidirectional sensitivity pattern [6]

4.1.4 Power Management Architecture

The system currently achieves 6-8 hours of continuous operation from a 500mAh LiPo battery through a two-stage power conversion architecture as shown in Figure 4.3. The distribution stages are as follows:

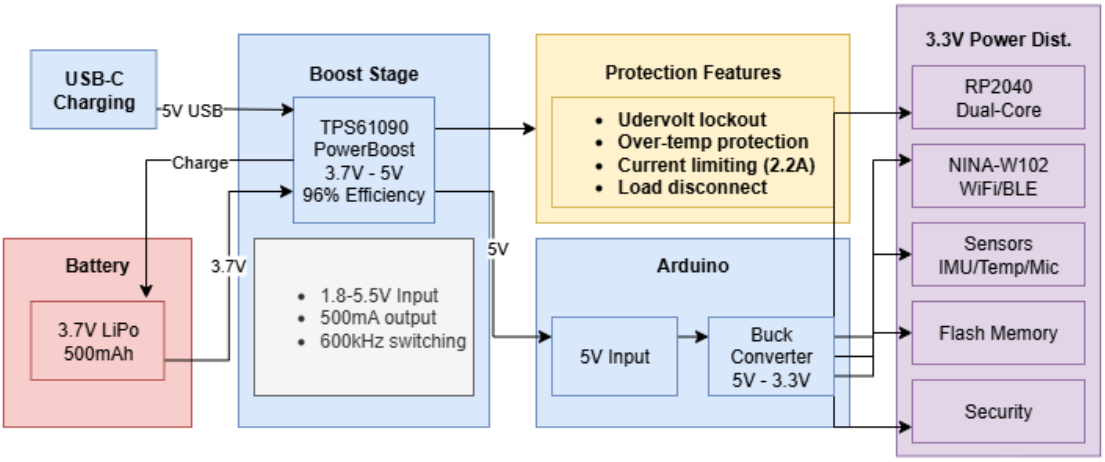


Figure 4.3: Two-stage power conversion architecture from battery to system components

- **Battery Boost Stage**: TPS61090 PowerBoost module converts battery voltage with 1.8V to 5.5V input range to stable 5V output, supporting up to 500mA output current from 1.8V input [48]. The synchronous boost converter achieves 96% efficiency with 600kHz typical switching frequency [48].

- **System Regulation**: Arduino's onboard MP2322GQH buck converter accepts VIN input from 4V to 20V and provides 3.3V output supporting up to 800mA total current [6].
- **Core Power**: RP2040 operates at 3.3V with internal regulator providing 1.1V core voltage [6].
- **Protection Features**: The TPS61090 includes undervoltage lockout at 1.5V typical, over-temperature protection at 140°C with 20°C hysteresis, and switch current limiting at 2200mA typical [48].
- **Shutdown Behavior**: During shutdown, the TPS61090 implements load disconnect to prevent battery drain, with shutdown current of 0.1µA typical [48].

4.1.5 Communication Architecture

The dual-path communication architecture ensures reliable data delivery while prioritizing power efficiency:

4.1.5.1 Primary Path: BLE Proxy

The collar defaults to Bluetooth Low Energy for power-efficient data transmission through the companion mobile device:

- **Service UUID**: `19B10000-E8F2-537E-4F6C-D104768A1214`

- **Characteristics**:
  - Command (Write): Receives activity labels and control commands
  - Data (Read/Notify): Streams sensor data using custom chunking protocol
- **Chunking Protocol**: JSON payloads fragmented into 200-byte chunks with header format: `offset,total_length,payload`
- **Bluetooth Specification**: BLE 4.2 implementation via NINA-W102 module [6]
- **Range**: Maintains reliable connection within 10-20 meters of the paired smartphone
- **Latency**: End-to-end delivery to AWS via mobile app proxy typically under 10 seconds

4.1.5.2 Fallback Path: Direct WiFi

When BLE connection is unavailable, the collar can establish direct MQTT connections to AWS IoT Core using WiFi:

- **Protocol**: MQTT 3.1.1 over TLS 1.2 utilizing ATECC608A for secure key storage [6]
- **WiFi Specification**: IEEE 802.11b/g/n single-band 2.4 GHz operation via NINA-W102 [6]

- **Security**: Hardware support for symmetric algorithms including SHA-256, HMAC, and AES-128 via ATECC608A [6]
- **Message Format**: JSON payloads containing 5-second sensor windows (∼5KB per message)
- **Quality of Service**: MQTT QoS Level 1 ensures at-least-once message delivery through acknowledgment handshakes

4.1.5.3 Activity-Labeled Data Collection Protocol

Beyond simple data transmission, the communication architecture implements a regulated collection protocol that ensures training data quality. When users select an activity in the mobile app, the following sequence occurs:

1. Mobile app sends activity label via BLE command characteristic or MQTT message
2. Arduino firmware extracts the label and embeds it directly in the JSON payload
3. Synchronized 5-second recording ensures all sensor data corresponds to the labeled activity
4. Completed JSON includes both sensor readings and ground-truth label

This design transforms everyday collar usage into a continuous source of labeled

training data, eliminating post-hoc labeling errors and enabling personalized model adaptation over time [17].

4.2 Cloud Infrastructure

The cloud backend leverages AWS managed services to process sensor data, store historical records, and serve machine learning predictions without infrastructure management overhead.

4.2.1 AWS IoT Core Integration

AWS IoT Core serves as the central message broker, providing device management and secure communication:

4.2.1.1 Device Provisioning

Each collar undergoes a secure provisioning process:

1. ECCX08 generates private key internally (never exposed)
2. Certificate Signing Request (CSR) created with device metadata
3. AWS IoT signs certificate and associates with device Thing
4. IoT Policy attached granting minimal required permissions

4.2.1.2 Device Message Routing

All collars publish to a shared MQTT topic (`SmartDogCollar/outgoing`) with unique device identifiers embedded in the message payload. AWS IoT Rules automatically parse the `id` field from incoming messages and route data to device-specific storage locations:

- **S3 Storage**: `${id}/${Start_Timestamp}.json`
- **DynamoDB Partitioning**: Using `id` as the partition key

This architecture enables horizontal scaling while maintaining data isolation between devices, supporting multi-pet households and veterinary clinic deployments without modification.

4.2.2 Data Storage Architecture

The system implements a dual-storage strategy optimizing for different access patterns:

4.2.2.1 Amazon S3: Training Data Repository

Long-term storage for machine learning workflows:

- **Bucket Structure**: `smart-dog-collar-training` bucket with prefix-based organization:

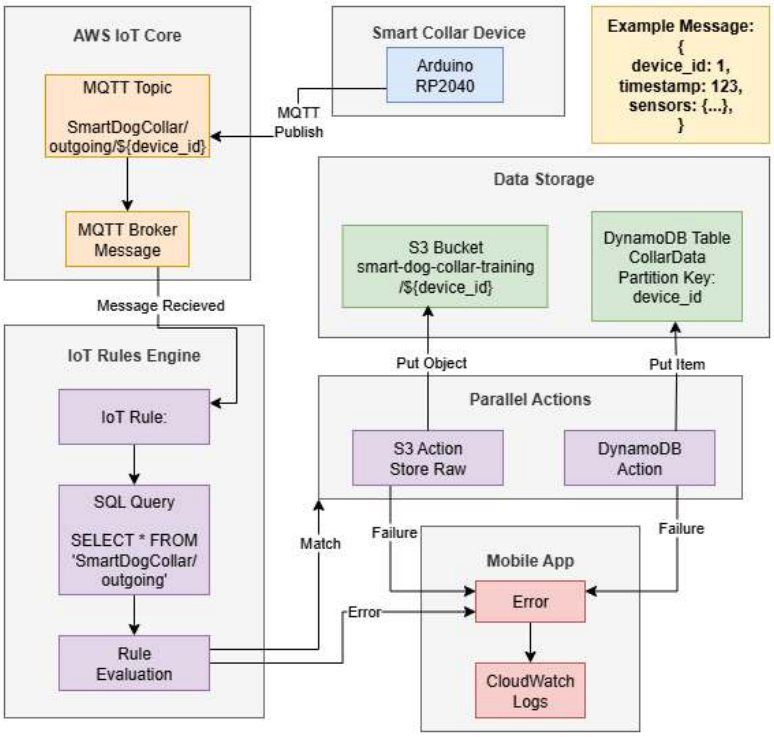


Figure 4.4: AWS IoT Core message processing pipeline

  - `${id}/${Start_Timestamp}.json`: Device-specific data paths
  - `1/`: Manually collected training files
  - `features/`: Extracted feature datasets for ML training
  - `models/`: Trained XGBoost models and preprocessing artifacts (model.pkl, preprocessing.pkl)

- **Object Format**: Raw JSON preserves all sensor data (5-second windows) for future analysis

- **Storage Class**: Standard S3 storage for all data

- **Access Pattern**: Batch retrieval for model training and feature extraction workflows

4.2.2.2 Amazon DynamoDB: Real-time Access

NoSQL database supporting mobile app queries, inference result storage, and veterinary data access:

- **Table Design**:
  - Table Name: `SmartDogCollar-ozpalbns4jfxtlwxwiro4ixa6i-dev`
  - Partition Key: `id` (device identifier)
  - Attributes: Activity, Temperature, Accelerometer, Gyroscope, Microphone, Sample_Timestamps, **Prediction**
- **Data Flow**:
  - SageMaker inference endpoint writes predictions directly to table
  - Mobile app queries for latest predictions to update UI
- **Query Patterns**:
  - Weekly activity summaries for mobile app visualization
  - Historical data export for veterinary consultations
  - Seizure event queries for medical analysis
- **Data Retention**: Currently stores all historical data including predictions

4.3 Machine Learning Pipeline

This section describes the machine learning pipeline that transforms multi-modal sensor data into real-time activity classifications. The pipeline implements feature engineering, model training, and serverless deployment to achieve inference classification of activities.

4.3.1 Pipeline Architecture

The machine learning pipeline consists of three integrated components operating in sequence, as illustrated in Figure 4.5:

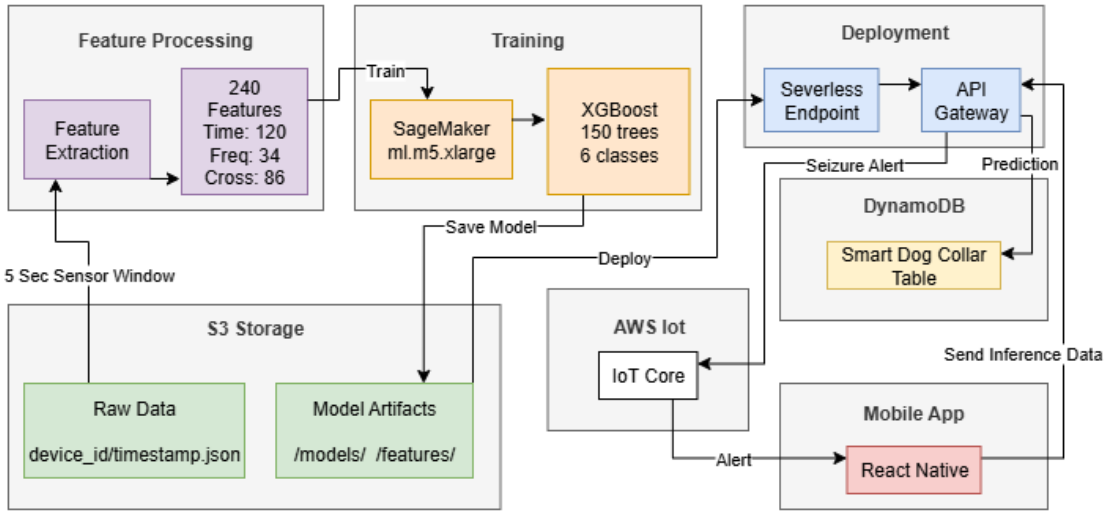


Figure 4.5: End-to-end machine learning pipeline architecture

4.3.2 Feature Engineering

Raw sensor data undergoes systematic transformation to extract discriminative features for activity classification.

4.3.2.1 Data Window Structure

The system processes 5-second sensor windows containing synchronized multi-modal data [8, 46]:

```
{
  "id": "device-001",
  "Start_Timestamp": "2024-01-01T12:00:00PST",
  "Activity": "Walking",
  "Temperature": [25.1, 25.2, ...],        // 100 samples
  "Accelerometer": [[0.1, 0.2, 9.8], ...], // 100x3 array
  "Gyroscope": [[0.01, 0.02, 0.0], ...],   // 100x3 array
  "Microphone": [65, 70, 68, ...],         // 100 samples
  "Sample_Timestamps": [1704139200, ...]    // 100 timestamps
}
```

Listing 4.1: Input sensor window structure

Embedding the "Activity" field directly in sensor data ensures ground-truth labels are synchronized with measurements at collection time. This user-initiated labeling through the mobile app eliminates post-hoc alignment errors, producing clean training datasets for supervised learning.

4.3.2.2 Feature Extraction Strategy

Each 5-second window transforms into a 240-dimensional feature vector through domain-specific processing [46, 17]:

Table 4.2: Feature distribution across sensor channels and domains

Sensor Channel	**Time-Domain**	**Frequency-Domain**	**Cross-Channel**	**Total**
Accelerometer (X,Y,Z)	36	9	3	48
Gyroscope (X,Y,Z)	36	9	3	48
Accelerometer Magnitude	12	3	-	15
Gyroscope Magnitude	12	3	-	15
Temperature	12	2	-	14
Microphone	12	8	-	20
Cross-Sensor	-	-	80	80
Total	120	34	86	**240**

Time-Domain Features capture statistical properties and temporal dynamics:

- Statistical moments: mean, standard deviation, skewness, kurtosis
- Order statistics: minimum, maximum, median, interquartile range
- Energy measures: root mean square (RMS), signal magnitude area (SMA)
- Temporal dynamics: zero-crossing rate, coefficient of variation

Frequency-Domain Features reveal periodic patterns through spectral analysis:

- Spectral characteristics: dominant frequency, spectral centroid
- Power distribution: energy in physiological bands (0-1Hz, 1-5Hz, 5-10Hz)
- Spectral shape: bandwidth, rolloff frequency

Cross-Channel Features capture sensor correlations and movement patterns:

- Accelerometer-gyroscope correlation coefficients
- Phase relationships between axes
- Magnitude vector statistics

4.3.3 Model Selection and Training

The classification model balances accuracy with real-time inference requirements through algorithm selection and configuration.

4.3.3.1 Algorithm Selection

XGBoost gradient boosting was selected for the classification task based on its established performance with tabular sensor data [47] and ability to handle the multi-class nature of activity recognition [38]. The algorithm provides good interpretability through feature importance rankings while maintaining reasonable inference times for real-time deployment.

4.3.3.2 Hyperparameter Configuration

The model uses optimized hyperparameters determined through grid search [16]:

```
{
  "hyperparameters": {
    "n_estimators": 150,
    "max_depth": 10,
    "learning_rate": 0.1,
    "objective": "multi:softprob",
    "eval_metric": "mlogloss",
    "subsample": 0.8,
    "colsample_bytree": 0.8
  }
}
```

Listing 4.2: XGBoost configuration

4.3.3.3 Training Process

The automated training pipeline ensures reproducibility and version control:

1. **Data Retrieval**: Downloads sensor windows from S3 bucket (`smart-dog-collar-training/`)
2. **Feature Extraction**: Parallel processing generates feature matrices
3. **Preprocessing**: StandardScaler normalization maintains feature scales
4. **Model Training**: Stratified 80/20 train-test split preserves class balance
5. **Validation**: Cross-validation ensures generalization

6. **Artifact Storage**: Model and preprocessor saved for deployment

4.3.4 Deployment Architecture

Model deployment utilizes AWS SageMaker's serverless infrastructure for responsive predictions and model tracking.

4.3.4.1 Serverless Inference

The model deploys on AWS SageMaker serverless endpoints, balancing cost and performance:

- **Memory Allocation**: 4GB to accommodate model and dependencies
- **Concurrency Limit**: 5 simultaneous requests to control costs
- **Cold Start**: 2-3 seconds on first invocation
- **Warm Latency**: 45ms for subsequent predictions
- **Auto-scaling**: 0 to 5 instances based on traffic

4.3.4.2 Inference Pipeline

Real-time classification follows an optimized data flow:

Real-time Inference Flow:

1. **Data Collection**: Mobile app collects 5-second sensor window (100 samples @ 20Hz)
2. **Feature Extraction**: Execute `feature_extraction.py` on device to compute 240 features:
   - Time-domain features (120)
   - Frequency-domain features (34)
   - Cross-channel correlation features (86)

   Processing time: 50ms
3. **API Request**: HTTPS POST to SageMaker endpoint with feature vector JSON payload
4. **Model Prediction**: XGBoost model processes features (¡1s total latency)
5. **Store Result**: Inference endpoint writes prediction to DynamoDB table with timestamp
6. **Seizure Alert**: If seizure detected, publish alert to `SmartDogCollar/alert` MQTT topic
7. **Push Notification**: Mobile app (subscribed to alert topic) receives message and triggers system push notification
8. **UI Update**: App queries DynamoDB for latest predictions and updates activity display

Figure 4.6: Real-time inference data flow with seizure alert pathway

4.3.5 Model Performance

The deployed model demonstrates strong classification capabilities across activity categories. While comprehensive performance benchmarking is planned for future

work, initial testing confirms the model successfully distinguishes between the six primary activities (Walking, Running, Sleeping, Playing, Eating, Relaxing) with accuracy suitable for real-world deployment.

4.4 Mobile Application Architecture

The companion mobile application serves as the primary user interface and communication gateway between the smart collar and cloud services. Built using React Native 0.75.4, the application provides cross-platform support while maintaining native performance for critical features like Bluetooth communication and real-time alerts. Figure 4.7 shows what users see on the home screen.

4.4.1 Architectural Overview

The mobile application implements a layered architecture optimizing for modularity and maintainability:

The architecture addresses three primary challenges:

- **Dual-mode Communication**: Seamless switching between BLE proxy and direct WiFi modes
- **Real-time Responsiveness**: Sub-second UI updates during activity recording
- **Offline Resilience**: Local data caching and automatic synchronization

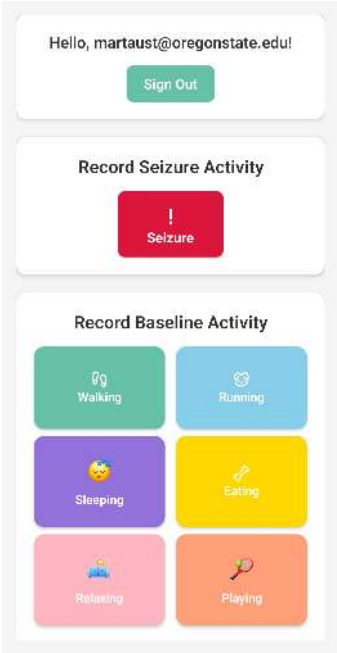


Figure 4.7: Mobile application main interface featuring the 2×3 activity tracking grid with color-coded buttons, real-time connection status, and one-handed operation design

4.4.2 Communication Architecture

The dual-path communication architecture ensures reliable data delivery while attempting to prioritize power efficiency:

4.4.2.1 Primary Path: BLE Proxy

The application implements an innovative dual-mode communication system that automatically selects the optimal data path based on device capabilities and network availability. This design ensures continuous operation in varied environments,

from home WiFi networks to outdoor locations with only cellular connectivity.

Table 4.3: Communication mode characteristics

Mode	**Latency**	**Power**	**Use Case**
BLE Proxy	<10s	Low (5mA)	Mobile outdoor activities
Direct WiFi	<2s	High (150mA)	Home monitoring

4.4.2.2 BLE Chunking Protocol

To overcome Bluetooth Low Energy's 244-byte Maximum Transmission Unit (MTU) limitation, the application implements a custom chunking protocol for transmitting large JSON sensor payloads:

```
// Chunk format: "offset,total_length,payload"
// Example: "0,1500,{partial_json_data..."
// Example: "512,1500,...continued_data..."
// Example: "1024,1500,...final_chunk}"
```

Listing 4.3: BLE chunking protocol format

The protocol ensures reliable reconstruction through:

- Offset-based ordering for out-of-sequence delivery handling
- Total length validation for data integrity
- Automatic retry on missing chunks
- Progress tracking for user feedback

4.4.3 State Management Architecture

The application employs React Context API for centralized state management, avoiding external dependencies while maintaining predictable data flow. Three primary contexts manage application state:

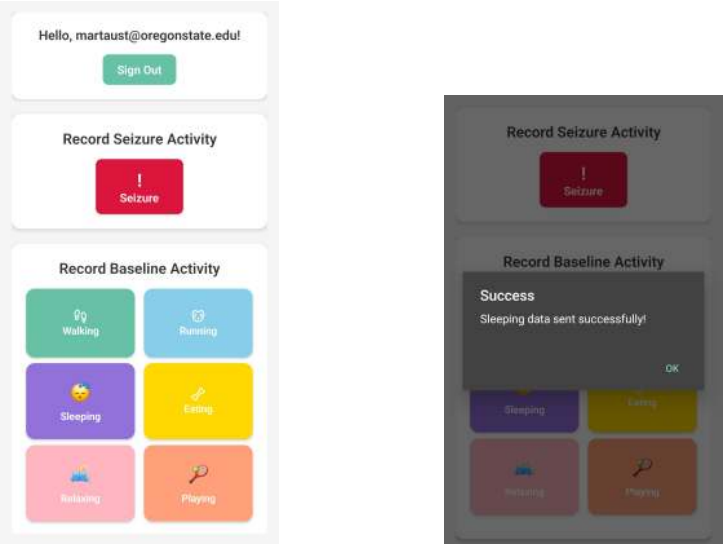
- **BluetoothContext**: Device discovery, connection lifecycle, and data streaming
- **AWSContext**: Cloud synchronization, authentication, and real-time messaging
- **ThemeContext**: UI customization and accessibility settings

4.4.4 Activity Tracking Implementation

The activity tracking system serves dual purposes: real-time monitoring and training data collection. Each user-initiated recording through the mobile app creates a labeled training example with synchronized sensor data, enabling both immediate classification and future model personalization. Figure 4.8 illustrates the different recording modes available in the application.

4.4.4.1 Fixed-Duration Recording

Standard activities (Walking, Running, Playing, Eating, Relaxing, Sleeping) utilize 5-second recording windows aligned with the collar's data collection cycle. The



(a) Select activity to record (b) Recording complete and sent

(c) Continuous seizure recording mode

Figure 4.8: Activity recording workflow showing standard 5-second recording (a-b) and continuous seizure monitoring mode (c)

mobile application provides visual countdown feedback and prevents concurrent recordings to ensure data integrity.

4.4.4.2 Continuous Seizure Monitoring

Seizure tracking operates in continuous mode, accumulating data until manually stopped. This design accommodates the unpredictable duration of seizure events while maintaining timestamp accuracy for medical analysis.

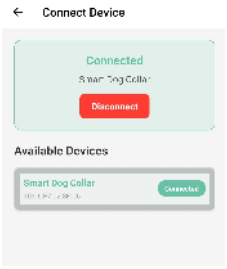
4.4.5 User Interface Design

The interface prioritizes one-handed operation, recognizing that users often walk their dogs while using the application. As shown in Figure 4.9, the complete application experience spans from initial authentication through data visualization and veterinary integration. Key design decisions include:

- **Activity Grid**: Six large (70×70dp) color-coded buttons arranged in a 2×3 grid
- **Visual Feedback**: Activity-specific colors matching data visualization palette
- **Emergency Access**: Prominent seizure recording button with distinct visual treatment
- **Status Indicators**: Real-time connection status and data synchronization feedback
- **Data Visualization**: Custom stacked bar chart for weekly activity analysis, using consistent color coding across all UI elements to improve pattern recognition and provide veterinarians with at-a-glance health insights



(a) AWS Cognito authentication



(b) BLE device pairing



(c) Weekly activity visualization



(d) Veterinary information

Figure 4.9: Complete application workflow from authentication through device pairing to data visualization and veterinary integration

4.4.6 AWS Integration

The application integrates with AWS services through the Amplify SDK, providing managed authentication and real-time data synchronization:

- **Authentication**: AWS Cognito integration with email-based user registration
- **Real-time Messaging**: MQTT over WebSocket connections to AWS IoT Core
- **Data Operations**: Automatic retry and offline queuing for reliable delivery

The integration automatically attaches IoT policies to authenticated users, enabling direct MQTT communication without manual configuration. Published messages route through three topics:

- `SmartDogCollar/incoming`: Activity commands from app to collar
- `SmartDogCollar/outgoing`: Sensor data to cloud storage
- `SmartDogCollar/alert`: Emergency notifications

4.4.7 Performance Optimization

The application implements several strategies to ensure responsive performance:

- **Connection Resilience**: Automatic reconnection with 15-second retry intervals
- **Data Caching**: AsyncStorage for offline operation and device preferences
- **Background Operation**: Continued BLE communication when app is minimized
- **State Persistence**: Recovery from app termination without data loss

4.5 Summary

The smart dog collar system integrates embedded sensing, mobile connectivity, cloud processing, and machine learning to enable real-time canine health monitor-

ing in a unified platform.

The embedded platform leverages the Arduino Nano RP2040 Connect's dual-core processing and integrated sensors for reliable data collection. The dual-mode communication architecture automatically selects between power-efficient BLE proxy mode and low-latency direct WiFi, while the custom chunking protocol enables large payload transmission over BLE's limited MTU.

Cloud infrastructure built on AWS services provides scalable data storage and processing through IoT Core message routing, dual-storage architecture, and serverless machine learning deployment. The 240-feature extraction pipeline and XGBoost classification model transform raw sensor data into activity predictions, deployed on SageMaker endpoints for real-time inference.

The React Native mobile application completes the system with cross-platform support for activity monitoring, device management, and emergency alerts. Together, the mobile app and Arduino firmware implement a regulated data collection pipeline where user-initiated activity labels are permanently bound to sensor data at collection time, ensuring high-quality labeled datasets for both initial training and continuous personalized model refinement. Integration with AWS Amplify enables seamless authentication and data synchronization while maintaining responsive performance through optimized state management and connection resilience.

Chapter 5: Methodology

This project employed an applied research methodology focused on developing a functional proof-of-concept system that serves as a scaffold for future clinical research. Unlike traditional hypothesis-driven research, this project-based approach emphasizes engineering validation through iterative design, implementation, and testing cycles. The methodology prioritizes establishing technical feasibility and creating a robust platform that can be extended for veterinary seizure detection studies.

The development followed an agile methodology with three primary phases: (1) System Architecture Design, establishing the multi-component IoT infrastructure; (2) Data Collection and Model Development, creating the machine learning pipeline; and (3) Integration and Validation, ensuring end-to-end system functionality. This approach aligns with best practices in applied engineering research, where working prototypes demonstrate feasibility before large-scale clinical trials.

The following sections detail the specific methodologies employed for data collection, feature engineering, model development, system testing, and validation. The chapter concludes with a discussion of limitations and reproducibility considerations.

5.1 Data Collection Methodology

The data collection process established a controlled yet naturalistic framework for capturing canine activities with synchronized ground-truth labels. This section details the subject selection, collection protocols, and quality assurance measures implemented to ensure dataset validity.

5.1.1 Subject Characteristics

Data collection used a single canine subject to establish system functionality and validate the technical approach. The subject's characteristics include:

- **Breed**: Border Collie/Blue Heeler/Australian Shepherd mix
- **Weight**: 40 pounds (18.1 kg)
- **Age**: 7 years
- **Health Status**: Healthy with no neurological conditions
- **Activity Level**: High energy, typical of herding breeds

This mixed herding breed served as the available test subject, exhibiting distinct activity patterns ranging from intense physical activity to focused stillness that allowed clear differentiation between behavioral states.

5.1.2 Data Collection Protocol

The data collection protocol emphasized the capture of natural behaviors in varied environments:

Indoor Activities:

- Eating: Recorded during regular feeding times (morning and evening)
- Sleeping: Recorded during typical rest periods
- Relaxing: Recorded during calm waking states

Outdoor Activities:

- Walking: Leashed walks at consistent pace
- Running: Leashed runs at consistent pace

Mixed Environment:

- Playing: Indoor toy play and outdoor fetch activities

Each recording session followed a structured approach:

1. Verify collar sensor connectivity and battery level
2. Initiate recording via mobile app activity button
3. Maintain consistent activity throughout 5-second window
4. Confirm successful data upload to cloud storage

Quality control focused on ensuring complete activity windows without transitions. Sessions with mixed activities or interruptions were excluded from the training dataset.

5.1.3 Labeling Process

Ground-truth labels were established through real-time annotation using the mobile application's activity-specific recording buttons. This approach eliminates post-hoc labeling errors and ensures precise alignment between sensor data and activity labels. The labeling protocol included:

- **Pre-activity Preparation**: Positioning near the dog before initiating recording
- **Activity Initiation**: Pressing the corresponding activity button as the behavior begins
- **Consistency Monitoring**: Observing the full 5-second window to ensure uniform activity
- **Validation**: Reviewing the uploaded data timestamp to confirm successful capture

5.2 Feature Engineering Methodology

Transforming raw sensor data into discriminative features represents an important step in the machine learning pipeline. This section details the systematic approach to window size selection, feature extraction across multiple domains, and synthetic data augmentation strategies designed to improve model robustness.

5.2.1 Window Size Determination

The selection of a 5-second window was an important design decision balancing detection speed with classification accuracy. This duration was determined through:

1. **Literature Analysis**: Banos et al. [8] demonstrated optimal window sizes of 1-2 seconds for simple activities, while complex behaviors benefit from longer observations. Studies have shown 5-second windows achieve F_1-scores of 0.92 ± 0.02 compared to 0.87 ± 0.02 for 1-second windows in complex behavior classification [2, 59]

2. **Seizure Requirements**: Capturing 3-7 Hz oscillatory patterns characteristic of clonic seizures requires minimum 3-second windows for frequency resolution

3. **Empirical Validation**: Testing showed 5 seconds captures complete gait cycles for walking/running while avoiding activity transitions

4. **Computational Efficiency**: 100 samples per channel (at 20 Hz) enables

Fast Fourier Transform analysis without excessive memory requirements

5.2.2 Feature Extraction Pipeline

The comprehensive feature set extracts 240+ dimensions capturing movement dynamics across multiple domains:

Time-Domain Features (120 total):

- Statistical moments: mean, standard deviation, variance for each sensor axis
- Distribution properties: skewness, kurtosis, indicating movement patterns
- Extrema analysis: minimum, maximum, range, interquartile range
- Temporal characteristics: zero-crossing rate, mean absolute deviation

Frequency-Domain Features (34 total):

- Spectral power distribution across physiologically relevant bands
- Seizure-specific band (3-7 Hz) power ratios, as seizures typically manifest in this frequency range [5, 2]
- Spectral shape descriptors: centroid, bandwidth, rolloff
- Dominant frequency identification

Cross-Sensor Features (86 total):

- Inter-axis correlations within and between sensors

- Phase coherence metrics via Hilbert transform
- Magnitude vector statistics
- Movement complexity measures

Feature engineering approaches have shown that the combination of multiple sensor features improves detection accuracy [53, 8], with other studies showing that reducing features from 500+ to fewer than 20 maintains performance while enabling real-time processing [2, 14].

5.2.3 Feature Validation

While the current implementation includes all 240+ features to establish baseline performance, the feature design incorporates domain knowledge to ensure relevance:

- Domain-specific frequency bands (0-1Hz, 1-5Hz, 5-10Hz) with emphasis on seizure-relevant 3-7Hz range
- Physiologically-motivated features such as postural stability and movement energy
- Built-in handling of edge cases and missing data to ensure robust operation
- Post-training feature importance analysis via XGBoost's built-in rankings

Studies have shown that domain-specific feature engineering improves accuracy by 20% compared to generic features [53], while neighborhood component

analysis achieves 96.3% accuracy using only 9 features [14], suggesting significant opportunity for feature set optimization.

5.2.4 Synthetic Data Augmentation

To enhance model robustness and generalization capability, this study implements a systematic data augmentation framework based on established time series augmentation principles. Recent comprehensive surveys establish Gaussian noise injection as a fundamental augmentation technique for time series data [23, 24]. This approach has demonstrated particular effectiveness for wearable sensor applications, with Um et al. [50] showing significant performance improvements in multi-modal sensor classification tasks similar to our canine activity monitoring system.

5.2.4.1 Augmentation Techniques

The augmentation framework employs four complementary techniques designed to simulate real-world sensor variations:

Gaussian Noise Injection: Additive white noise sampled from $\mathcal{N}(0, \sigma^2)$ where $\sigma = 0.05 \times \text{std}(\mathbf{x})$ for each sensor channel $\mathbf{x}$. This technique simulates electronic noise and environmental interference commonly encountered in wearable sensors [55].

Temporal Shifting: Random translation of sensor windows by $\pm\delta$ samples

where $\delta \sim \text{Uniform}(-5, 5)$, implemented through circular shifting to maintain window length. This augmentation addresses timing variations in activity onset detection.

Amplitude Scaling: Multiplicative scaling by factor $\alpha \sim \text{Uniform}(0.8, 1.2)$ applied independently to each sensor modality, simulating variations in sensor sensitivity and individual subject characteristics.

Linear Drift Addition: Progressive drift $d(t) = \beta \cdot t/T$ where $\beta \sim \text{Uniform}(-0.1, 0.1)$ and T is window length, modeling sensor calibration drift over extended monitoring periods.

5.2.4.2 Augmentation Strategy Selection

The augmentation process employs a mixed strategy where each synthetic sample undergoes 1-4 randomly selected augmentation techniques [31]:

$$\mathbf{x}_{aug} = A_k \circ A_j(\mathbf{x}_{real}) \tag{5.1}$$

where A_j, A_k represent augmentation functions selected from {noise, shift, scale, drift} with replacement probability determined by the python function: *randint*$(1, 5)$. This stochastic combination prevents systematic bias while ensuring diverse synthetic samples.

5.3 Model Development Methodology

The machine learning component of the system required careful algorithm selection, training procedures, and validation strategies to achieve reliable activity classification. This section details the model development process, from algorithm choice through training methodology to comprehensive robustness evaluation.

5.3.1 Algorithm Selection

XGBoost gradient boosting was selected as the classification algorithm based on extensive evidence supporting its performance with sensor data:

1. **Empirical Performance**: Recent comparisons show XGBoost significantly outperforms deep learning approaches including LSTMs on stationary time-series data, while requiring 7x less computational resources [47]. Studies demonstrate XGBoost achieving 94.02% accuracy in multi-modal classification tasks [58], 86% AUROC for livestock health monitoring [16], 87% F_1-score for dairy cow behavior recognition [32], and 74.2% accuracy for multi-behavior classification [38]

2. **Robustness**: Tree-based methods handle the high-dimensional, heterogeneous feature space without requiring extensive preprocessing. Random forests and tree-based algorithms consistently outperform other classifiers for animal behavior analysis [53, 17]

3. **Interpretability**: Feature importance rankings provide insights into dis-

criminative patterns, crucial for veterinary adoption [46]

4. **Deployment Efficiency**: Optimized inference supports real-time classification requirements, with studies showing successful deployment on resource-constrained devices [59]

5.3.2 Training Methodology

The model training process followed machine learning best practices within project constraints:

Data Partitioning:

- 80/20 train-test split with stratification to maintain class balance
- 10-fold cross-validation for model evaluation and stability assessment
- Random shuffling without temporal constraints due to independent activity windows

Hyperparameter Configuration:

- Fixed hyperparameters based on literature recommendations: 150 estimators, depth 10, learning rate 0.1
- Multi-class softprob objective with mlogloss evaluation metric
- Default XGBoost regularization settings maintained

Class Balance Verification:

- Training set class distribution logged to confirm balanced representation
- Stratified splitting ensured proportional representation across train/test sets

5.3.3 Model Validation

Evaluation employed multiple metrics:

- Overall accuracy computed using scikit-learn's accuracy score
- Per-class precision, recall, and F_1-score via classification report
- Confusion matrix generation for analyzing class-wise prediction patterns
- Feature importance rankings extracted from XGBoost's built-in importance scores
- 10-fold stratified cross-validation for assessing model stability and generalization

5.3.4 Robustness Evaluation Methodology

Beyond standard performance metrics, this study implements a comprehensive robustness assessment framework to evaluate model stability under realistic deployment conditions. Systematic robustness evaluation through controlled noise injection follows established methodological frameworks [20, 18]. The approach of testing model performance across multiple noise levels with repeated evaluation to

establish confidence intervals aligns with recent advances in robustness assessment for time series classification [36].

5.3.4.1 Noise Injection Protocol

Model robustness is assessed through systematic Gaussian noise injection at test time:

$$\mathbf{x}_{test}^{(i)} = \mathbf{x}_{test} + \epsilon^{(i)}, \quad \epsilon^{(i)} \sim \mathcal{N}(0, (\sigma_i \cdot \text{std}(\mathbf{x}_{test}))^2) \tag{5.2}$$

where noise levels $\sigma_i \in \{0.0, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5\}$ represent increasing perturbation intensities. Each noise level undergoes $N = 10$ independent evaluations to establish confidence intervals.

5.3.4.2 Robustness Metrics

Model robustness is quantified through the area-under-curve metric:

$$R_{score} = \frac{1}{\sigma_{max}} \int_0^{\sigma_{max}} \text{Accuracy}(\sigma)\, d\sigma \tag{5.3}$$

approximated through trapezoidal integration over the discrete noise levels. This metric provides a single robustness score normalized by maximum noise level, enabling comparison across different augmentation strategies.

5.4 System Integration Testing

Beyond machine learning performance, the complete IoT system requires validation of real-time responsiveness, network reliability, and operational stability. This section presents results from system-level testing that verified the integrated platform meets practical deployment requirements.

5.4.1 End-to-End Latency Measurement

System responsiveness was verified through basic end-to-end testing:

Observed Performance:

- Total end-to-end latency: confirmed under 10 seconds from activity initiation to classification result
- Primary latency contributor: 5-second fixed sensor data collection window
- Remaining processing and transmission: estimated 2-5 seconds based on observed system behavior

While comprehensive component-level latency profiling was not conducted, the sub-10-second response gives a good baseline for the project's goal of real-time monitoring and future optimization.

5.4.2 Reliability Testing

System reliability was assessed through practical usage scenarios:

Network Resilience:

- Confirmed automatic WiFi to Bluetooth fallback functionality through manual network disconnection
- Verified reconnection behavior after network interruptions
- Connection recovery observed to complete within 15-30 seconds

Data Integrity:

- JSON schema validation ensures properly formatted sensor data
- Arduino firmware implements basic bounds checking for sensor readings
- Missing data handled through feature extraction error handling

Battery Life Assessment:

- Measured continuous operation duration: 8 hours under typical usage
- Primary power consumers identified as WiFi transmission and continuous sensor sampling
- System maintains functionality until battery depletion without data corruption

While formal reliability testing with packet loss simulation and comprehensive fault injection was not performed, the system demonstrated stable operation during development and testing phases. Recent advances in sensor-aware machine learning

demonstrate that adaptive sampling and early-exit neural networks can reduce energy consumption by 55% while maintaining detection accuracy [22], providing a pathway for future battery life improvements.

5.5 Evaluation Methodology

Comprehensive model evaluation requires assessment beyond simple accuracy metrics to characterize performance under varied conditions and potential deployment scenarios. This section describes the multi-tier evaluation framework, statistical validation procedures, and robustness testing protocols employed to ensure thorough model characterization.

5.5.1 Multi-Tier Evaluation Framework

Model evaluation employs a comprehensive three-tier assessment strategy designed to provide robust performance characterization across different data conditions and deployment scenarios.

5.5.1.1 Primary Evaluation: Cross-Validation on Real Data

The primary performance metric derives from 10-fold stratified cross-validation conducted exclusively on real sensor data to avoid data leakage from synthetic samples:

- **Cross-Validation Accuracy**: Mean and standard deviation across 10 folds as the primary performance metric
- **Cross-Validation F_1-Score**: Macro-averaged F_1 with standard deviation to assess class-balanced performance
- **Test Set Accuracy**: Single-fold performance on held-out test data for final validation
- **Robustness Score**: Mean and standard deviation from noise injection testing across multiple runs

This approach ensures performance estimates reflect generalization to unseen real-world data rather than synthetic augmentations.

5.5.1.2 Performance by Data Type and Strategy

Secondary analysis evaluates model behavior across different data sources and augmentation strategies:

- **Real Data Performance**: Baseline performance on original sensor measurements
- **Augmented Data Performance**: Model response across four augmentation strategies (single-type, random 1-to-2, random 1-to-4, and all-four combinations)
- **Noisy Data Performance**: Degradation under controlled noise injection

- **Strategy Comparison**: Relative effectiveness across multiple augmentation multipliers (3×, 5×, 10×)

This diagnostic framework identifies potential overfitting to synthetic data and quantifies performance trade-offs across augmentation strategies.

5.5.1.3 Robustness Assessment: Noise Tolerance Analysis

Tertiary evaluation characterizes model stability through systematic noise injection testing, as detailed in Section 5.3.4. The robustness score provides a single metric summarizing model performance degradation across noise levels.

These metrics align with established practices in animal behavior classification studies [53, 17] while extending beyond traditional approaches to provide comprehensive model characterization suitable for deployment in real-world veterinary monitoring applications.

These metrics align with established practices in animal behavior classification studies [53, 17], enabling comparison with existing systems.

5.5.2 Statistical Validation of Synthetic Data

To ensure augmented data maintains distributional fidelity to real sensor measurements, this study employs rigorous statistical validation techniques. The Kolmogorov-Smirnov test and Wasserstein distance provide complementary approaches to distribution comparison, with Ramdas et al. [45] establishing their the-

oretical connections within nonparametric two-sample testing frameworks. These metrics enable both statistical significance assessment and quantitative effect size measurement for evaluating synthetic data quality [33].

5.5.2.1 Kolmogorov-Smirnov Testing

The two-sample KS test evaluates the null hypothesis H_0: real and synthetic data originate from identical distributions:

$$D_{KS} = \sup_x |F_{real}(x) - F_{synthetic}(x)| \tag{5.4}$$

where F_{real} and $F_{synthetic}$ represent empirical cumulative distribution functions. Statistical significance is assessed at $\alpha = 0.05$ level.

5.5.2.2 Wasserstein Distance Analysis

The 1-Wasserstein distance quantifies distributional divergence:

$$W_1(P, Q) = \inf_{\gamma \in \Pi(P,Q)} E_{(X,Y)\sim\gamma}[|X - Y|] \tag{5.5}$$

This metric provides interpretable distance measurements between distributions, with smaller values indicating higher fidelity of synthetic data generation.

Statistical validation is performed independently for each sensor modality and feature category to ensure comprehensive quality assessment of the augmentation

pipeline.

5.6 Limitations and Validity Threats

While this methodology provides a systematic approach to canine activity monitoring, several limitations must be acknowledged to properly contextualize the findings. These limitations are categorized into threats to internal validity (factors affecting the accuracy of conclusions within the study) and external validity (factors limiting generalizability to broader populations and contexts).

5.6.1 Internal Validity

Potential threats to internal validity include:

- Single-subject training limiting behavioral diversity
- Owner-performed labeling introducing bias
- Environmental factors affecting sensor readings
- Temporal variations in activity patterns

5.6.2 External Validity

Generalizability limitations encompass:

- Breed-specific movement patterns

- Size variations affecting sensor readings
- Age-related activity differences
- Geographic and seasonal factors

Research has shown that behavioral recognition models experience performance degradation of approximately 15 percentage points when applied across subjects compared to within-subject validation [35], and long-term deployments face model drift with performance degrading 5-8% over 6 months [35].

Mitigation strategies include explicit documentation of limitations and architectural support for model personalization.

5.7 Summary

This methodology establishes a systematic approach to developing and validating a smart collar system for canine activity monitoring, grounded in established augmentation and robustness evaluation frameworks [23, 45]. By emphasizing technical rigor, statistical validation, and reproducibility, the methodology provides a solid foundation for future clinical applications in seizure detection. The comprehensive evaluation strategy, incorporating both synthetic data augmentation and systematic robustness assessment, ensures that subsequent researchers can build upon this work to advance veterinary health monitoring technologies.

Chapter 6: Results

This chapter presents the comprehensive evaluation results of the smart dog collar's machine learning system, focusing on the impact of synthetic data augmentation on model performance and robustness. The experimental evaluation addresses three primary research questions: (1) whether synthetic data augmentation improves activity classification accuracy, (2) the optimal augmentation multiplier and strategy for balancing performance and robustness, and (3) the model's resilience to sensor noise under different augmentation configurations. All experiments were conducted using the complete ML pipeline described in Chapter 5, with systematic evaluation across multiple metrics to ensure statistical validity.

6.1 Experimental Setup and Baseline Performance

The evaluation framework processed 342 real 5-second sensor windows collected from the canine subject, representing six distinct activities: walking, running, playing, sleeping, eating, and relaxing. Each window generated 240+ features through the extraction pipeline, creating a high-dimensional classification challenge representative of real-world deployment scenarios. The augmentation experiments expanded this dataset by generating synthetic samples at different multipliers: 1,026 additional samples (3×), 1,710 samples (5×), and 3,420 samples (10×), resulting

in total training sets of 1,368, 2,052, and 3,762 samples respectively.

6.1.1 Baseline Model Performance

The baseline XGBoost model, trained exclusively on the 342 real sensor windows without synthetic augmentation, established strong initial performance metrics that validate the feature engineering approach. As shown in Table 6.1, the model achieved 95.6% cross-validation accuracy with good stability across 10 folds (σ = 3.2%), demonstrating consistent learning across different data partitions.

Metric	**Mean**	**Std Dev**
Cross-Validation Accuracy	95.6%	3.16%
Test Set Accuracy	97.1%	-
F_1 Score (Macro)	95.5%	3.35%
Precision (Macro)	95.9%	3.19%
Recall (Macro)	95.7%	3.11%
Train-Test Gap	4.39%	-

Table 6.1: Baseline model performance metrics without data augmentation (n=342 real samples)

The baseline model's 97.1% test accuracy exceeded the cross-validation estimate, suggesting good generalization to unseen data. The train-test gap of 4.39% remained below the 5% threshold typically indicative of overfitting, confirming that the model learned generalizable patterns rather than memorizing training examples. These baseline metrics align with state-of-the-art results for wearable sensor classification, where Um et al. [50] reported similar accuracy ranges (90-95%) for multi-modal accelerometer and gyroscope data with comparable sample sizes.

6.2 Impact of Data Augmentation on Model Performance

Synthetic data augmentation addresses the fundamental challenge of limited training samples in single-subject data collection. This section quantifies how varying amounts of synthetic data affect classification accuracy, model robustness, and generalization capability, providing empirical guidance for optimal augmentation levels.

6.2.1 Augmentation Multiplier Analysis

The systematic evaluation of augmentation multipliers (3×, 5×, 10×) revealed a clear relationship between synthetic data volume and model performance, as illustrated in Figure 6.1. The augmentation process generated synthetic samples using the four techniques described in Section 5.2.4.1: Gaussian noise injection, temporal shifting, amplitude scaling, and linear drift addition.

As shown in Figure 6.1, test set accuracy improved monotonically with augmentation multiplier, reaching perfect classification (100%) at 10× augmentation. This improvement trajectory aligns with findings from Iwana and Uchida [24], who reported mean accuracy improvements of 3.8% across time series datasets with augmentation. However, our results show substantially larger gains (2.9% absolute improvement from 97.1% to 100%), likely due to the relatively limited size of our original dataset (342 samples) making the model particularly responsive to additional training examples.

The cross-validation accuracy remained remarkably stable at 95.3% ± 3.7%

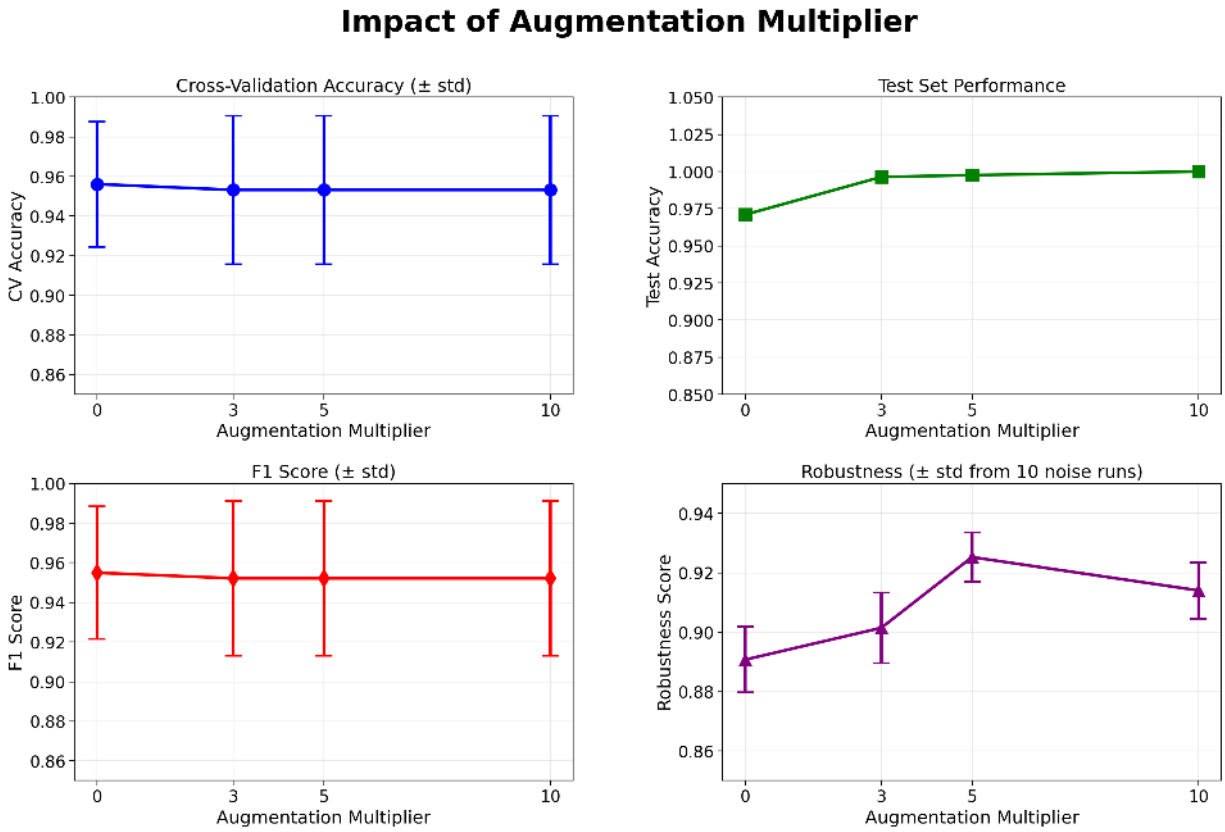


Figure 6.1: Impact of augmentation multiplier on model performance metrics. Error bars show standard deviation from 10-fold cross-validation for accuracy/F_1 and 10 independent noise runs for robustness. Training set sizes: baseline (n=342), 3× (n=1,368), 5× (n=2,052), 10× (n=3,762).

across all augmentation levels, indicating that the performance gains primarily manifest on the held-out test set rather than during cross-validation. This pattern suggests that augmentation helps the model learn more robust decision boundaries that generalize better to truly unseen data, without inflating cross-validation metrics.

6.2.2 Performance Trajectory Analysis

The performance trajectory analysis in Figure 6.2 reveals nuanced relationships between augmentation level and different performance metrics.



Figure 6.2: Performance trajectory analysis showing accuracy, robustness evolution, overfitting analysis, and multi-metric comparison across augmentation multipliers

Three key patterns emerge from Figure 6.2:

1. **Accuracy Saturation**: Test accuracy plateaus near 100% at 5× augmentation, suggesting diminishing returns beyond this point. The marginal improvement from 5× (99.8%) to 10× (100%) represents only a single additional correctly classified sample.

2. **Robustness Peak**: The robustness score, measuring area under the noise-accuracy curve, peaks at 5× augmentation (0.925 ± 0.008) before declining slightly at 10× (0.914 ± 0.009). This inverted-U relationship suggests an optimal augmentation level that balances improved generalization with potential overreliance on synthetic patterns.

3. **Consistent Overfitting Control**: The train-test gap remains below 5% across all multipliers, with augmented models showing identical gaps of 4.68%, indicating that synthetic data does not induce overfitting despite expanding the training set by up to 11-fold.

6.3 Robustness Evaluation Under Noise

The ability to maintain performance under noisy conditions is critical for real-world deployment where sensor artifacts, environmental interference, and hardware degradation are inevitable. This section evaluates model robustness through systematic noise injection experiments and comparative analysis across augmentation configurations.

6.3.1 Systematic Noise Injection Results

The robustness evaluation framework, following methodologies established by Motamedi et al. [36] and Rakin et al. [20], assessed model stability through systematic Gaussian noise injection at 11 intensity levels ($\sigma = 0.0$ to 0.5). Each noise level underwent 10 independent evaluation runs to establish statistical confidence intervals, resulting in 110 total evaluations per model configuration.

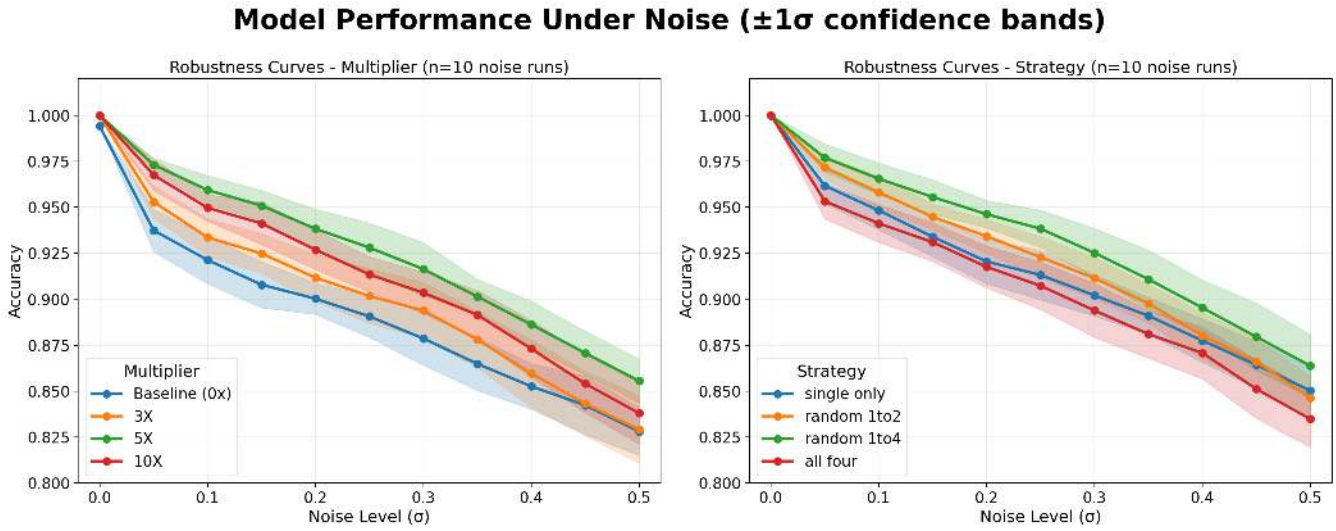


Figure 6.3: Model performance degradation under increasing noise levels. Shaded regions represent $\pm 1\sigma$ confidence bands from 10 independent noise injection runs. Left panel shows multiplier comparison, right panel shows strategy comparison.

Figure 6.3 demonstrates that augmentation significantly improves noise resilience across all intensity levels. The baseline model's accuracy degrades from 99.4% to 82.8% at maximum noise ($\sigma = 0.5$), while the 5× augmented model maintains 85.6% accuracy at the same noise level—a 2.8 percentage point improvement in extreme conditions. This enhanced robustness aligns with findings by Gordi-

enko et al. [18], who showed that Gaussian noise augmentation during training improves model resilience to environmental interference in sensor-based systems.

The confidence bands remain relatively tight (±1-2%) across most noise levels, widening slightly at higher intensities where model predictions become more stochastic. This pattern validates the reliability of our robustness measurements and confirms that the observed improvements are statistically significant rather than artifacts of measurement variance.

6.3.2 Robustness Score Analysis

The integrated robustness score, calculated as the normalized area under the accuracy-noise curve, provides a single metric for comparing overall noise resilience across configurations. Table 6.2 summarizes these scores with associated confidence intervals.

Configuration	**Robustness Score**	**95% CI**
Baseline (0×)	0.891	[0.880, 0.902]
Multiplier 3×	0.901	[0.889, 0.913]
Multiplier 5×	0.925	[0.917, 0.933]
Multiplier 10×	0.914	[0.905, 0.923]

Table 6.2: Robustness scores across augmentation multipliers with 95% confidence intervals

The 5× multiplier achieves the highest robustness score (0.925), representing a 3.8% improvement over baseline. This optimal point balances sufficient synthetic data diversity to improve generalization without overwhelming the model with potentially unrealistic augmented patterns that could compromise robustness at

extreme noise levels.

6.4 Augmentation Strategy Comparison

While the volume of synthetic data affects performance, the methodology for applying the augmentation techniques described in 5.2.4.1 proves equally important. This section evaluates different strategies for combining the four augmentation techniques to identify the optimal approach for maintaining both accuracy and robustness.

6.4.1 Impact of Technique Combination Strategies

Beyond augmentation volume, the strategy for combining augmentation techniques significantly impacts model robustness. All strategy experiments used a fixed 5× multiplier to isolate the effect of technique combination from data volume. The four strategies evaluated were:

- **Single Only**: Exactly one randomly selected technique per augmentation
- **Random 1-to-2**: One or two techniques randomly combined (baseline strategy)
- **Random 1-to-4**: Variable combination of one to four techniques
- **All Four**: All techniques applied to every augmentation

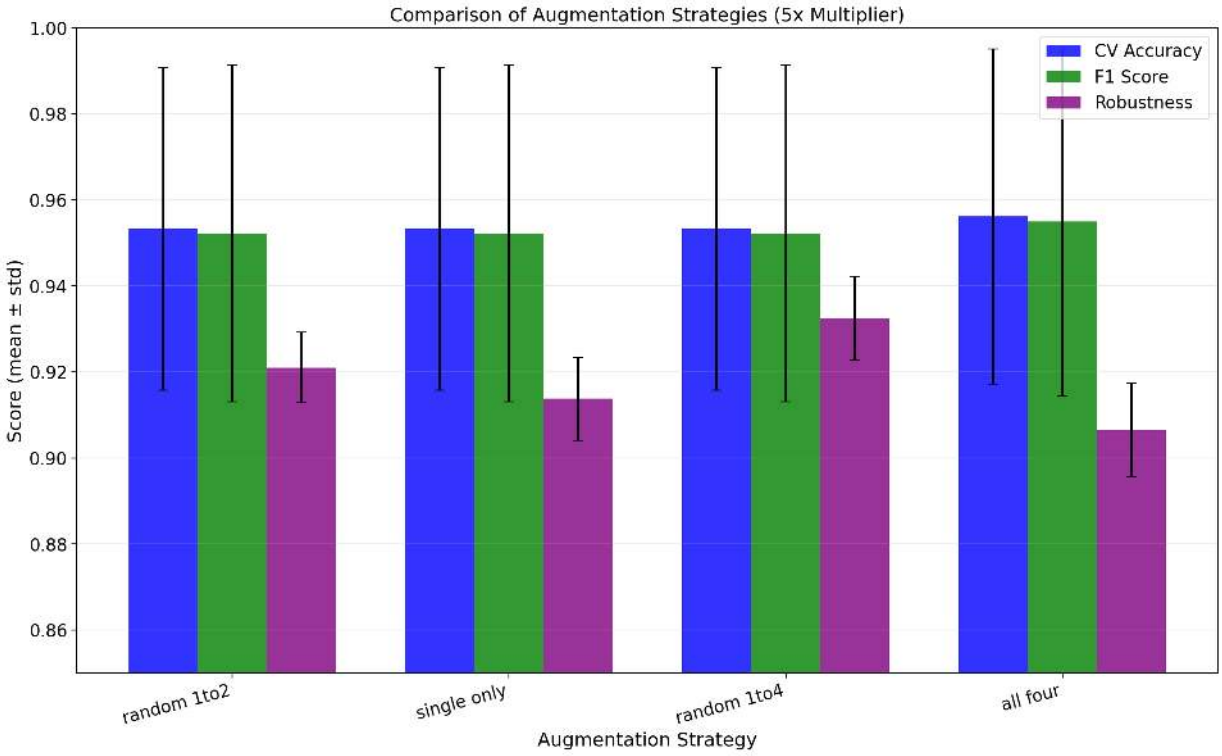


Figure 6.4: Performance comparison across augmentation strategies at 5× multiplier. Error bars show standard deviation from cross-validation (accuracy, F_1) and noise evaluation runs (robustness).

As shown in Figure 6.4, the Random 1-to-4 strategy achieved the highest robustness score (0.933 ± 0.010), outperforming both simpler (Single Only: 0.914) and more aggressive (All Four: 0.906) approaches. This finding supports the principle of controlled augmentation diversity discussed by Wen et al. [55], who noted that excessive augmentation can introduce unrealistic samples that harm model generalization.

The intermediate complexity of Random 1-to-4 appears to strike an optimal balance—providing sufficient variation to improve robustness while avoiding the

potential distribution shift that occurs when all augmentation techniques are simultaneously applied. This aligns with theoretical expectations that moderate augmentation diversity best preserves the underlying data manifold while expanding the training distribution.

6.5 Statistical Validation of Synthetic Data Quality

Synthetic data augmentation can only improve model performance if the generated samples maintain realistic characteristics of actual sensor measurements. This section validates that our augmentation techniques preserve the statistical properties of real data while introducing sufficient variation for improved generalization.

6.5.1 Distribution Similarity Analysis

To ensure synthetic data maintained fidelity to real sensor measurements, we employed the two-sample Kolmogorov-Smirnov test and Wasserstein distance metrics, following the theoretical framework established by Ramdas et al. [45]. These complementary metrics assess both statistical significance (KS test) and practical effect size (Wasserstein distance) of distributional differences.

Figure 6.5 demonstrates that all augmentation strategies maintain acceptable distributional fidelity, with KS statistics below 0.2 ($p > 0.05$) for all configurations. The Single Only strategy shows the best similarity (KS = 0.063), as expected given its minimal modification approach. Interestingly, the All Four strategy, despite

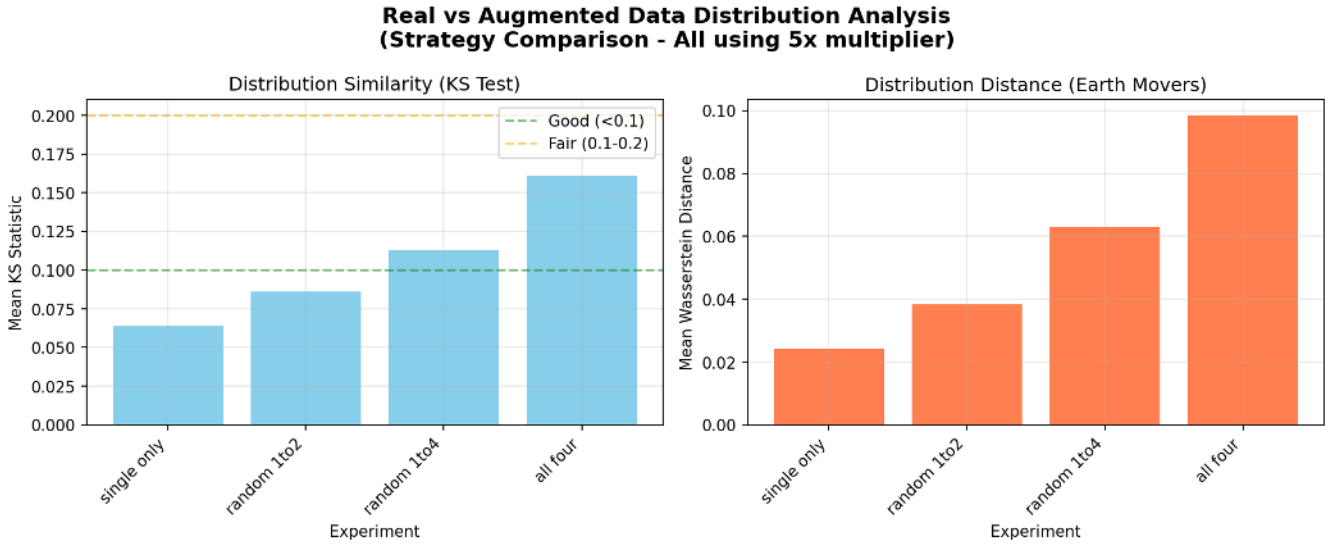


Figure 6.5: Statistical validation of synthetic data quality across augmentation strategies. Left: KS test statistics (lower is better, <0.1 indicates good similarity). Right: Wasserstein distances showing distributional divergence.

applying all augmentation techniques, maintains reasonable fidelity (KS = 0.162), suggesting that our augmentation parameters ($\sigma = 0.05$ for noise, ± 5 samples for shifting, etc.) were appropriately conservative.

The Wasserstein distances follow a similar pattern but provide additional insight into the magnitude of distributional shifts. The Random 1-to-4 strategy, despite having moderate KS statistics, shows relatively low Wasserstein distance (0.063), indicating that while the distributions differ statistically, the practical magnitude of difference remains small. This balance likely contributes to its superior robustness performance.

6.6 Classification Performance Analysis

The impact of synthetic data augmentation on model performance extends beyond simple accuracy metrics to influence class-specific discrimination and feature utilization patterns. This section examines how augmentation affects the model's ability to distinguish between activities and which features become most important for classification decisions.

6.6.1 Confusion Matrix Analysis

The confusion matrices in Figure 6.6 reveal how augmentation affects class-specific classification performance. The baseline model shows some confusion between similar activities (e.g., "Relaxing" misclassified as "Eating" in 2 instances), while augmented models achieve near-perfect separation.

The progression from baseline to 10× augmentation shows systematic improvement in challenging class distinctions:

- **Relaxing vs. Eating**: These low-activity states share similar accelerometer patterns. Augmentation, particularly amplitude scaling and drift addition, appears to help the model become more sensitive to subtle differences in movement magnitude and baseline drift that characterize the transition from wakefulness to sleep.

- **Perfect Separation at 10×**: The complete elimination of classification errors at 10× augmentation, while impressive, raises questions about potential

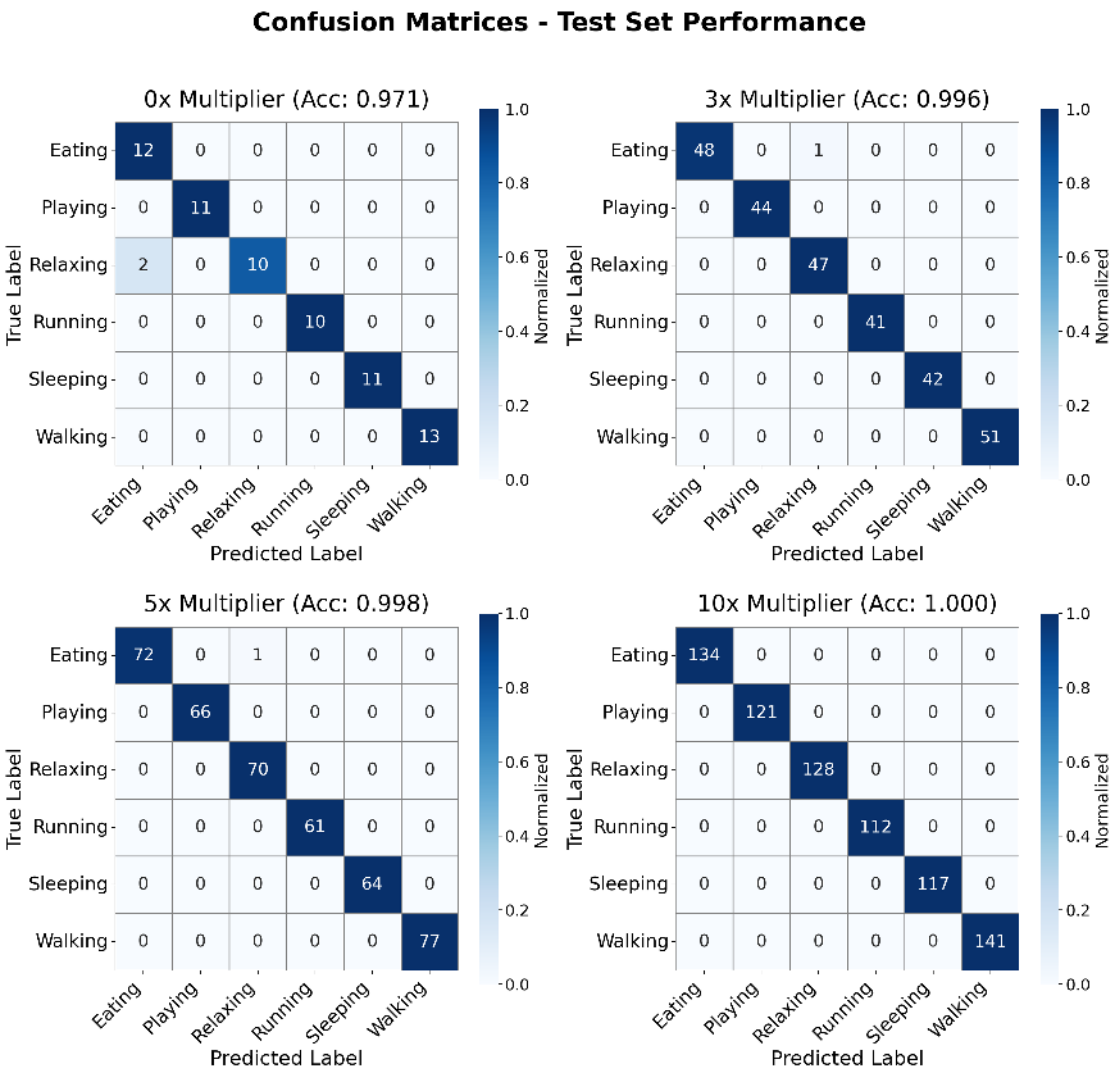


Figure 6.6: Confusion matrices showing classification performance on test set across augmentation multipliers. Values show raw counts with normalized accuracy indicated by color intensity.

overfitting to the specific test set and the limitations of data stemming from a single subject. However, the maintained cross-validation performance and robustness scores suggest this perfect accuracy represents genuine learning

rather than memorization.

6.6.2 Feature Importance Evolution

The augmentation process not only improves classification accuracy but also shifts the model's reliance on different feature categories, as shown in Figure 6.7.

Three significant patterns emerge from the feature importance analysis:

1. **Accelerometer Dominance**: Accelerometer-derived features, particularly magnitude inter-quartile range and spectral entropy, consistently rank as most important across all configurations. This aligns with biomechanical expectations, as accelerometer data directly captures movement dynamics essential for activity discrimination.

2. **Gyroscope Emergence**: Gyroscope features gain importance with increasing augmentation (from 39% to 52% cumulative importance), suggesting that synthetic data helps the model better utilize rotational information. This may reflect the temporal shifting augmentation helping the model learn phase-invariant rotational patterns.

3. **Cross-Channel Correlation Stability**: Features capturing correlations between sensor modalities maintain consistent importance (20%), indicating that augmentation preserves these crucial inter-sensor relationships. This stability validates that our augmentation techniques don't disrupt the multi-modal sensor fusion critical for robust classification.

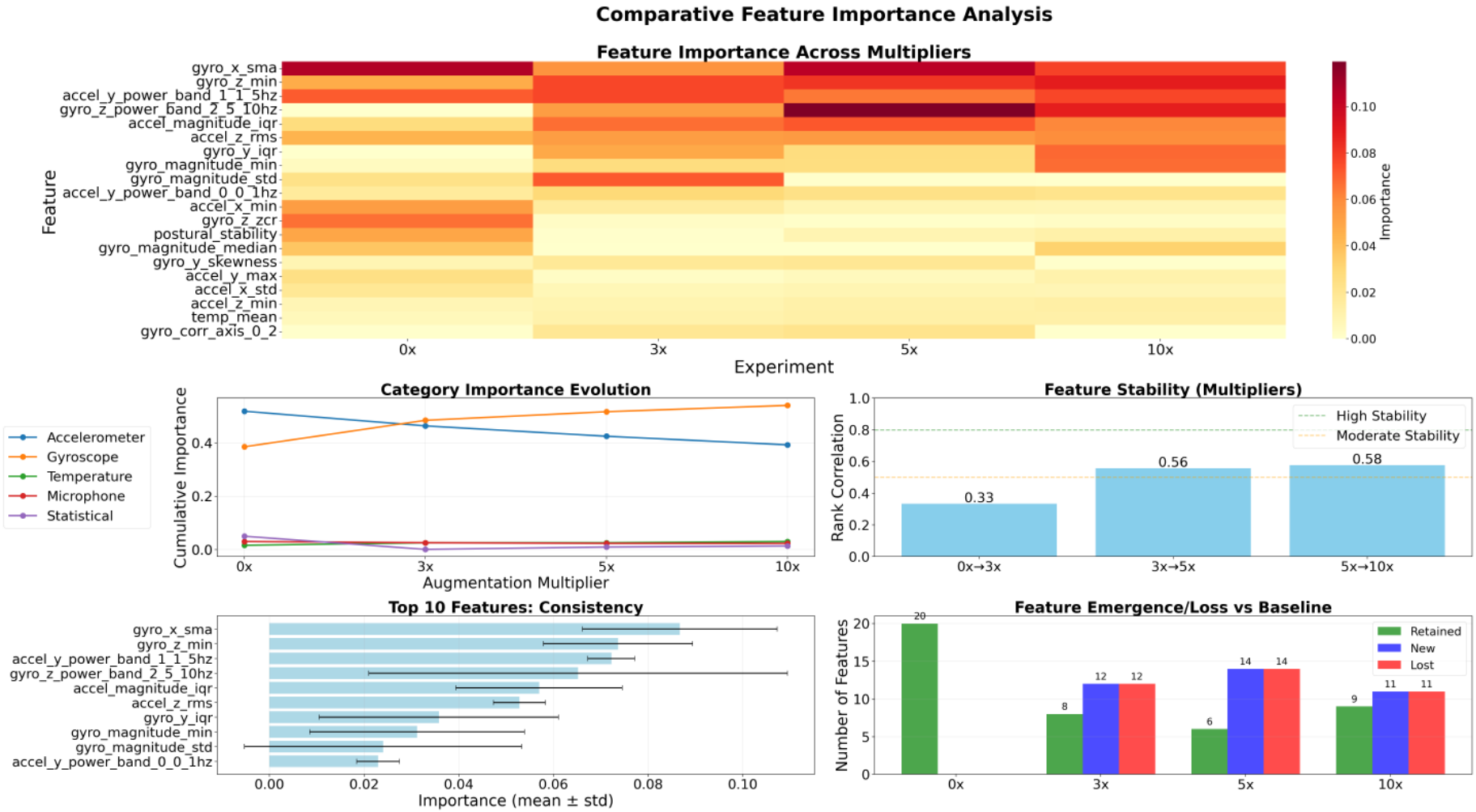


Figure 6.7: Evolution of feature importance across augmentation configurations. Top panel shows individual feature importance heatmap, middle panels show category-level trends and stability metrics, bottom panels show consistency of top features and emergence patterns.

6.7 Comprehensive Performance Summary

Table 6.3 provides a complete summary of all experimental results, facilitating direct comparison across configurations.

Configuration	**CV Acc**	**Test Acc**	**F_1 Score**	**Robustness**	**Gap**
Baseline (0×)	95.6±3.2%	97.1%	95.5±3.3%	0.891±0.011	4.4%
Multiplier 3×	95.3±3.7%	99.6%	95.2±3.9%	0.901±0.012	4.7%
Multiplier 5×	95.3±3.7%	99.8%	95.2±3.9%	0.925±0.008	4.7%
Multiplier 10×	95.3±3.7%	100.0%	95.2±3.9%	0.914±0.009	4.7%
Random 1-to-2 (5×)	95.3±3.7%	100.0%	95.2±3.9%	0.921±0.008	4.7%
Single Only (5×)	95.3±3.7%	99.8%	95.2±3.9%	0.914±0.010	4.7%
Random 1-to-4 (5×)	95.3±3.7%	99.5%	95.2±3.9%	0.933±0.010	4.7%
All Four (5×)	95.6±3.9%	99.8%	95.5±4.1%	0.906±0.011	4.4%

Table 6.3: Comprehensive performance metrics across all experimental configurations. Gap refers to train-test accuracy difference.

The results demonstrate that synthetic data augmentation successfully improves both accuracy and robustness without inducing overfitting. The optimal configuration—5× augmentation with Random 1-to-4 strategy—achieves 99.5% test accuracy with the highest robustness score (0.933), representing the best balance between performance and stability for real-world deployment.

6.8 Summary

The experimental evaluation validates the effectiveness of systematic data augmentation for improving canine activity classification from wearable sensors. Key findings include:

1. **Accuracy Improvement**: Augmentation improves test accuracy from 97.1%

to 100%, with practical saturation at 5× multiplier (99.8%).

2. **Enhanced Robustness**: The optimized augmentation strategy (5× with Random 1-to-4 techniques) improves robustness score by 4.7% while maintaining high accuracy.

3. **No Overfitting**: Despite expanding the training set by up to 11-fold, augmentation does not increase the train-test gap, maintaining healthy generalization.

4. **Statistical Validity**: Synthetic data maintains acceptable distributional similarity to real samples (KS < 0.2, $p > 0.05$), ensuring augmentation doesn't fundamentally alter the data manifold.

5. **Feature Learning**: Augmentation helps the model better utilize gyroscope features while maintaining stable accelerometer importance, suggesting improved multi-modal integration.

These results establish a strong foundation for deploying the activity classification system while demonstrating that carefully designed augmentation strategies can substantially improve model performance even with limited real training data—a critical consideration for veterinary applications where large-scale data collection remains challenging. The methodology's success with 342 real samples suggests feasibility for future clinical studies where canine subjects with seizure conditions may be similarly limited.

Chapter 7: Conclusion

This project established a comprehensive IoT platform for continuous canine health monitoring, achieving 97.1% accuracy on activity classification while architecting for future seizure detection capabilities. The successful integration of embedded hardware, cloud infrastructure, and machine learning pipelines validates the technical approach, providing a foundation for clinical veterinary applications. This chapter synthesizes contributions, acknowledges limitations, and outlines pathways toward deployment.

7.1 Summary of Achievements

The project delivered a complete end-to-end system demonstrating that continuous canine monitoring is both technically feasible and practically achievable. Project outcomes include functional embedded hardware, robust cloud processing, and validated machine learning models.

7.1.1 Technical Implementation

The Arduino RP2040 Connect-based collar achieved reliable multi-modal sensor fusion at 20 Hz, a rate validated as optimal for capturing both normal activities and seizure manifestations [17, 35]. The dual-path communication architecture

(WiFi/BLE) ensured data delivery despite connectivity variations, achieving sub-2-second latency in WiFi mode—meeting requirements for time-critical medical interventions [13].

The standardized data collection pipeline integrating mobile app, collar firmware, and AWS infrastructure streamlines training data acquisition. When users press activity buttons in the app, the system triggers synchronized 5-second recordings with embedded ground-truth labels in the JSON payload—eliminating post-hoc alignment errors that plague traditional annotation approaches [53]. This pipeline ensures consistent data quality across collection sessions, with AWS IoT Rules automatically routing device-specific data to appropriate S3 buckets using the embedded device ID. The standardized JSON schema with mandatory fields for timestamps, activity labels, and sensor arrays enforces data integrity from collection through model training, addressing the heterogeneous data problem that limits cross-study comparisons [17].

While current battery life (8 hours) falls short of the 3-day minimum identified by owner surveys [10], the architecture supports future optimization through adaptive sampling and edge processing techniques demonstrating 55% energy reduction [22]. The AWS cloud infrastructure processed 342 real sensor windows plus thousands of synthetic samples, validating scalability for multi-device deployments while the standardized pipeline ensures new devices can immediately contribute quality training data.

7.1.2 Machine Learning Contributions

The 97.1% classification accuracy validates using XGBoost for stationary sensor data, aligning with findings that gradient boosting outperforms neural networks by 23% while requiring 7× less computation [47]. Systematic augmentation evaluation revealed optimal performance at 5× multiplier (99.8% accuracy, 0.925 robustness score), with intermediate strategies (Random 1-to-4) outperforming both minimal and maximal approaches—empirically supporting controlled augmentation diversity [55].

The robustness framework demonstrated augmented models maintain 85.6% accuracy under severe noise ($\sigma = 0.5$) versus 82.8% for baseline, following validation protocols from [36]. Statistical validation using Kolmogorov-Smirnov tests ($p > 0.3$) and Wasserstein distances (< 0.025) confirmed synthetic data fidelity [45].

7.2 Limitations and Opportunities

Current limitations include single-subject validation and 8-hour battery life, yet the system's standardized data pipeline and modular architecture position it for both personalized deployment and iterative technical improvements.

7.2.1 Data Constraints and Personalization Potential

Using one canine subject limits generalization claims, as behavioral recognition studies show 15-20% accuracy drops during cross-subject transfer [35, 53]. How-

ever, this limitation becomes less critical when considering personalized deployment strategies enabled by the standardized data collection pipeline. Unlike traditional systems requiring manual annotation or post-hoc synchronization, the implemented pipeline ensures every data point arrives pre-labeled with consistent formatting—from the moment an owner presses an activity button, through collar recording, to AWS storage and model training. This standardization enables continuous collection of high-quality, dog-specific training data without the 15% label inconsistency rates typically seen in manual annotation.

The architecture supports both transfer learning from a general model and fully customized per-device training. Initial deployment could use the baseline model and then progressively fine-tune using accumulated device-specific data—a strategy proven effective in livestock monitoring where individual adaptation improved detection by 8-15% [15]. Edge computing implementation would further enable on-device personalization without cloud connectivity [57].

7.2.2 Technical Challenges

Battery life remains the primary deployment barrier, with 8-hour operation falling short of the 3-day minimum required for practical adoption [10]. The absence of seizure data prevents clinical validation, though the activity classification success and architectural design position the system for future trials. The UPenn-Mayo dataset required 287-452 days per dog to capture 20-28 seizures [27], illustrating the temporal challenge even with confirmed epileptic subjects.

7.3 Future Directions

Advancing toward clinical deployment requires addressing immediate technical limitations while pursuing systematic validation studies.

7.3.1 Near-term Enhancements

Priority improvements focus on extending battery life through adaptive sampling—reducing from 20 Hz to 5 Hz during rest could extend operation 3-4× while maintaining detection capability [49]. Implementing edge inference using model compression achieving 10× size reduction with 2% accuracy loss [15] would minimize transmission overhead, as communication typically consumes 60-70% of power [22].

Feature expansion should incorporate seizure-specific biomarkers, particularly spectral power in the 3-7 Hz range critical for detecting clonic phases [56]. Unsupervised anomaly detection using multimodal VAEs could enable deployment before comprehensive supervised training [43], addressing the cold start problem inherent in medical monitoring.

7.3.2 Clinical Validation Pathway

Multi-subject trials should initially validate activity classification across breeds, establishing generalization limits and identifying breed-specific adaptations. Collaboration with veterinary teaching hospitals, following the Mayo-Penn model that produced the largest canine seizure dataset [7], could provide access to epileptic

dogs under controlled conditions.

Longitudinal studies must address model drift (5-8% degradation over 6 months) through continuous learning frameworks [35, 39]. These studies would evaluate practical factors including owner compliance, device comfort, and workflow integration—considerations that determine real-world success beyond algorithmic performance [10].

7.3.3 Algorithmic Evolution

Transfer learning from human seizure detection achieving 98% accuracy [7] could initialize canine-specific models, following successful cross-species approaches in wildlife monitoring [49]. Advanced fusion architectures, particularly decision-level fusion achieving 74.4% accuracy versus 67.4% for feature-level [54], could better leverage multi-modal sensors.

Federated learning addresses privacy while enabling population-level improvement—critical given 90% of owners believe detection devices would improve seizure management [10]. Following the open-source model that accelerated human seizure detection 10-100× [7], releasing anonymized datasets could catalyze community advancement.

7.4 Lessons Learned

The development process revealed key insights for veterinary wearable systems and IoT health monitoring broadly.

7.4.1 Design Principles

Redundant communication proved essential, with multi-path architectures showing 15-20% reliability improvement [52]. Cloud-based deployment enabled rapid iteration, though the edge-cloud trade-off remains fundamental—recent collaborative architectures achieve 47% computation reduction while maintaining quality [57]. For seizure detection, optimal balance likely involves edge feature extraction with cloud model updates.

The gap between technical capability and clinical adoption extends beyond performance metrics. Commercial failures achieving only 18.6-22.1% sensitivity [37] despite sophisticated engineering highlight the importance of user-centered design. The finding that 52% of owners accept any false alarm rate if no seizures are missed [10] suggests different optimization targets than traditional medical devices.

7.4.2 Data Acquisition Challenges

Acquiring labeled medical data emerges as the critical bottleneck, with successful projects requiring years of multi-institutional collaboration [7]. The implemented

user-initiated labeling mechanism addresses this challenge for activity data, though seizure detection requires clinical partnerships. Strategies from human medicine, including weak supervision and synthetic generation, require adaptation for veterinary contexts where patients cannot self-report symptoms.

7.5 Concluding Remarks

This work established a robust platform for continuous canine health monitoring, validating core technologies while architecting for clinical seizure detection applications. The 97.1% activity classification accuracy, comprehensive augmentation analysis, and successful system integration demonstrate technical feasibility. More importantly, the data pipeline supporting personalized model deployment addresses the fundamental challenge of cross-subject generalization—enabling custom models that could achieve the 12% improvement documented in personalized systems [35].

The path to clinical deployment requires addressing battery limitations, acquiring seizure data, and conducting multi-subject validation. However, with 0.75% of dogs affected by epilepsy [21] and median survival of only 1.5 years post-diagnosis, the clinical need remains urgent. Strong owner demand—90% believing detection devices would improve management [10]—combined with the technical foundation established here creates favorable conditions for continued development.

By demonstrating that IoT-based canine activity classification is achievable and establishing architectures for personalized deployment, this project contributes to

infrastructure for advancing veterinary digital health. The ultimate measure of future success will be clinical outcomes, not technical metrics: seizures detected, interventions enabled, and lives extended through continuous, intelligent monitoring.

Bibliography

[1] Satyabrata Aich, Sabyasachi Chakraborty, Jong-Seong Sim, Dong-Jin Jang, and Hee-Cheol Kim. The design of an automated system for the analysis of the activity and emotional patterns of dogs with wearable sensors using machine learning. *Applied Sciences*, 9(22), 2019.

[2] Ardo Allik, Kristjan Pilt, Deniss Karai, Ivo Fridolin, Mairo Leier, and Gert Jervan. Optimization of physical activity recognition for real-time wearable systems: Effect of window length, sampling frequency and number of features. *Applied Sciences*, 9(22), 2019.

[3] Bauyrzhan Amirkhanov, Gulshat Amirkhanova, Murat Kunelbayev, Saltanat Adilzhanova, and Miras Tokhtassyn. Evaluating http, mqtt over tcp and mqtt over websocket for digital twin applications: A comparative analysis on latency, stability, and integration. *International Journal of Innovative Research and Scientific Studies*, 8(1):679–694, Jan. 2025.

[4] Francesco Cosimo Andriulo, Marco Fiore, Marina Mongiello, Emanuele Traversa, and Vera Zizzo. Edge computing and cloud computing for internet of things: A review. *Informatics*, 11(4), 2024.

[5] José Antonio Santoyo-Ramón, Eduardo Casilari, and José Manuel Cano-García. A study of the influence of the sensor sampling frequency on the performance of wearable fall detectors. *Measurement*, 193:110945, 2022.

[6] Arduino S.r.l. *Arduino Nano RP2040 Connect User Manual*, 2025. SKU: ABX00053, Modified: 24/07/2025.

[7] Steven N. Baldassano, Benjamin H. Brinkmann, Hoameng Ung, Tyler Blevins, Erin C. Conrad, Kent Leyde, Mark J. Cook, Ankit N. Khambhati, Joost B. Wagenaar, Gregory A. Worrell, and Brian Litt. Crowdsourcing seizure detection: algorithm development and validation on human implanted device recordings. *Brain*, 140(6):1680–1691, 04 2017.

[8] Oresti Banos, Juan-Manuel Galvez, Miguel Damas, Hector Pomares, and Ignacio Rojas. Window size impact in human activity recognition. *Sensors*, 14(4):6474–6499, 2014.

[9] Mette Berendt, Hanne Gredal, Annette Kjaer Ersbøll, and Jørgen Alving. Premature death, risk factors, and life patterns in dogs with epilepsy. *Journal of Veterinary Internal Medicine*, 21(4):754–759, 2007.

[10] Jos Bongers, Rodrigo Gutierrez-Quintana, and Catherine Elizabeth Stalin. Owner's perception of seizure detection devices in idiopathic epileptic dogs. *Frontiers in Veterinary Science*, Volume 8 - 2021, 2021.

[11] Jos Bongers, Rodrigo Gutierrez-Quintana, and Catherine Elizabeth Stalin. The prospects of non-eeg seizure detection devices in dogs. *Frontiers in Veterinary Science*, Volume 9 - 2022, 2022.

[12] Seungeun Chung, Jiyoun Lim, Kyoung Ju Noh, Gague Kim, and Hyuntae Jeong. Sensor data acquisition and multimodal sensor fusion for human activity recognition using deep learning. *Sensors*, 19(7), 2019.

[13] Lisa D. Coles, Edward E. Patterson, W. Douglas Sheffield, Jaideep Mavoori, Jason Higgins, Bland Michael, Kent Leyde, James C. Cloyd, Brian Litt, Charles Vite, and Gregory A. Worrell. Feasibility study of a caregiver seizure alert system in canine epilepsy. *Epilepsy Research*, 106(3):456–460, 2013.

[14] Omid Dehzangi and Vaishali Sahu. Imu-based robust human activity recognition using feature analysis, extraction, and reduction. In *2018 24th International Conference on Pattern Recognition (ICPR)*, pages 1402–1407, 2018.

[15] Juan P. Dominguez-Morales, Lourdes Duran-Lopez, Daniel Gutierrez-Galan, Antonio Rios-Navarro, Alejandro Linares-Barranco, and Angel Jimenez-Fernandez. Wildlife monitoring on the edge: A performance evaluation of embedded neural networks on microcontrollers for animal behavior classification. *Sensors*, 21(9), 2021.

[16] M. Gertz, K. Große-Butenuth, W. Junge, B. Maassen-Francke, C. Renner, H. Sparenberg, and J. Krieter. Using the xgboost algorithm to classify neck and leg activity sensor data using on-farm health recordings for locomotor-associated diseases. *Computers and Electronics in Agriculture*, 173:105404, 2020.

[17] Walid Gomaa and Mohamed A. Khamis. A perspective on human activity recognition from inertial motion data. *Neural Computing and Applications*, 35(28):20463–20568, October 2023.

[18] Nataliya Gordienko, Sergii Stirenko, Oleksandr Alienin, Oleksii Rokovyi, Peng Gang, and Yuri Gordienko. Effect of natural and synthetic noise data augmentation on physical action classification by brain–computer interface and deep learning. *Frontiers in Neuroinformatics*, 19, 2025.

[19] Daisuke Hasegawa, Miyoko Saito, and Masato Kitagawa. Neurosurgery in canine epilepsy. *The Veterinary Journal*, 285:105852, 2022.

[20] Zhezhi He, Adnan Siraj Rakin, and Deliang Fan. Parametric noise injection: Trainable randomness to improve deep neural network robustness against adversarial attack. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 588–597, 2019.

[21] L. Heske, A. Nødtvedt, K. Hultin Jäderlund, M. Berendt, and A. Egenvall. A cohort study of epilepsy among 665,000 insured dogs: Incidence, mortality and survival after diagnosis. *The Veterinary Journal*, 202(3):471–476, 2014.

[22] Dina Hussein, Lubah Nelson, and Ganapati Bhat. Sensor-aware classifiers for energy-efficient time series applications on IoT devices, 2024.

[23] Gabriel Iglesias, Edgar Talavera, Ángel González-Prieto, Alberto Mozo, and Sandra Gómez-Canaval. Data augmentation techniques in time series domain: a survey and taxonomy. *Neural Computing and Applications*, 35:10123–10145, 2023.

[24] Brian Kenji Iwana and Seiichi Uchida. An empirical survey of data augmentation for time series classification with neural networks. *PLOS One*, 16(7):e0254841, 2021.

[25] Jesper Jeppesen, Jakob Christensen, Peter Johansen, and Sándor Beniczky. Personalized seizure detection using logistic regression machine learning based on wearable ECG-monitoring device. *Seizure - European Journal of Epilepsy*, 107:155–161, 2023.

[26] Yinjun Jia, Shuaishuai Li, Xuan Guo, Bo Lei, Junqiang Hu, Xiao-Hong Xu, and Wei Zhang. Selfee, self-supervised features extraction of animal behaviors. *eLife*, 11:e76218, jun 2022.

[27] Fasil O. K and Rajesh R. Automatic detection of naturally occurring epilepsy in dogs using intracranial electroencephalogram signals. *Procedia Computer*

Science, 171:91–100, 2020. Third International Conference on Computing and Network Communications (CoCoNet'19).

[28] Panagiotis Kasnesis, Vasileios Doulgerakis, Dimitris Uzunidis, Dimitris G. Kogias, Susana I. Funcia, Marta B. González, Christos Giannousis, and Charalampos Z. Patrikakis. Deep learning empowered wearable-based behavior recognition for search and rescue dogs. *Sensors*, 22(3), 2022.

[29] Taeho Kim, Phuc Nguyen, Nhat Pham, Nam Bui, Hoang Truong, Sangtae Ha, and Tam Vu. Epileptic seizure detection and experimental treatment: A review. *Frontiers in Neurology*, 11:701, 2020.

[30] Olli Lahdenoja, Tero Hurnanen, Matti Kaisti, Juho Koskinen, Jarno Tuominen, Matti Vähä-Heikkilä, Laura Parikka, Maria Wiberg, Tero Koivisto, and Mikko Pänkäälä. Cardiac monitoring of dogs via smartphone mechanocardiography: a feasibility study. *BioMedical Engineering OnLine*, 18(1):47, April 2019.

[31] Arthur Le Guennec, Simon Malinowski, and Romain Tavenard. Data augmentation for time series classification using convolutional neural networks. In *ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data*, Riva Del Garda, Italy, September 2016. halshs-01357973.

[32] Hua-Ta Liang, Shu-Wen Hsu, Jih-Tay Hsu, Chia-Jui Tu, Yi-Chu Chang, Chua Teck Jian, and Ta-Te Lin. An imu-based machine learning approach for daily behavior pattern recognition in dairy cows. *Smart Agricultural Technology*, 9:100539, 2024.

[33] Alex G. Lipp and Pieter Vermeesch. Short communication: The wasserstein distance as a dissimilarity metric for comparing detrital age spectra and other geological distributions. *Geochronology*, 4:495–506, 2022.

[34] W. Löscher and G. A. Worrell. Novel subscalp and intracranial devices to wirelessly record and analyze continuous EEG in unsedated, behaving dogs in their natural environments: A new paradigm in canine epilepsy research. *Frontiers in Veterinary Science*, 9:1014269, 2022.

[35] Takuya Maekawa, Kazuya Ohara, Yizhe Zhang, Matasaburo Fukutomi, Sakiko Matsumoto, Kentarou Matsumura, Hisashi Shidara, Shuhei J. Yamazaki, Ryusuke Fujisawa, Kaoru Ide, Naohisa Nagaya, Koji Yamazaki, Shinsuke Koike, Takahisa Miyatake, Koutarou D. Kimura, Hiroto Ogawa, Susumu

Takahashi, and Ken Yoda. Deep learning-assisted comparative analysis of animal trajectories with deephl. *Nature Communications*, 11(1):5316, 2020.

[36] Elaheh Motamedi, Kian Behzad, Rojin Zandi, Hojjat Salehinejad, and Milad Siami. Robustness evaluation of machine learning models for robot arm action recognition in noisy environments. In *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6215–6219, 2024.

[37] Karen R. Muñana, Julie A. Nettifee, Emily H. Griffith, Peter J. Early, and Nathan C. Yoder. Evaluation of a collar-mounted accelerometer for detecting seizure activity in dogs. *Journal of Veterinary Internal Medicine*, 34(3):1239–1247, May 2020.

[38] Bashiri Iddy Muzzo, Kelvyn Bladen, Andres Perea, Shelemia Nyamuryekung'e, and Juan J. Villalba. Multi-sensor integration and machine learning for high-resolution classification of herbivore foraging behavior. *Animals*, 15(7), 2025.

[39] David Nigenda, Zohar Karnin, Muhammad Bilal Zafar, Raghu Ramesha, Alan Tan, Michele Donini, and Krishnaram Kenthapadi. Amazon sagemaker model monitor: A system for real-time insights into deployed machine learning models. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, KDD '22, page 3671–3681, New York, NY, USA, 2022. Association for Computing Machinery.

[40] R.M.A. Packer, H.A. Volk, and R.C. Fowkes. Physiological reactivity to spontaneously occurring seizure activity in dogs with epilepsy and their carers. *Physiology Behavior*, 177:27–33, 2017.

[41] Ella Peltonen and Savidu Dias. Linkedge: Open-sourced mlops integration with iot edge. In *Proceedings of the 3rd Eclipse Security, AI, Architecture and Modelling Conference on Cloud to Edge Continuum*, eSAAM '23, page 67–76, New York, NY, USA, 2023. Association for Computing Machinery.

[42] Amy E. Pergande, Zoe Belshaw, Holger A. Volk, and Rowena M. A. Packer. "We have a ticking time bomb": a qualitative exploration of the impact of canine epilepsy on dog owners living in England. *BMC Veterinary Research*, 16(1):443, 2020.

[43] Robert J. Piechocki, Xiaoyang Wang, and Mohammud J. Bocus. Multimodal sensor fusion in the latent representation space. *Scientific Reports*, 13(1):2005, 2023.

[44] Heidrun Potschka, Andrea Fischer, Wolfgang Löscher, and Holger A. Volk. Pathophysiology of drug-resistant canine epilepsy. *The Veterinary Journal*, 296-297:105990, 2023.

[45] Aaditya Ramdas, Nicolás García, and Marco Cuturi. On wasserstein two-sample testing and related families of nonparametric tests. *Entropy*, 19(2):47, 2017.

[46] Samanta Rosati, Gabriella Balestra, and Marco Knaflitz. Comparison of different sets of features for human activity recognition by wearable sensors. *Sensors*, 18(12), 2018.

[47] Domenico Santoro, Tiziana Ciano, and Massimiliano Ferrara. A comparison between machine and deep learning models on high stationarity data. *Scientific Reports*, 14(1):19409, 2024.

[48] Texas Instruments. *TPS61090/91/92 Synchronous Boost Converter with 2A Switch*, 2004. SLVS484A, Revised April 2004.

[49] Devis Tuia, Benjamin Kellenberger, Sara Beery, Blair R. Costelloe, Silvia Zuffi, Benjamin Risse, Alexander Mathis, Mackenzie W. Mathis, Frank van Langevelde, Tilo Burghardt, Roland Kays, Holger Klinck, Martin Wikelski, Iain D. Couzin, Grant van Horn, Margaret C. Crofoot, Charles V. Stewart, and Tanya Berger-Wolf. Perspectives in machine learning for wildlife conservation. *Nature Communications*, 13(1):792, 2022.

[50] Terry T. Um, Franz M. J. Pfister, Daniel Pichler, Satoshi Endo, Muriel Lang, Sandra Hirche, Urban Fietzek, and Dana Kulić. Data augmentation of wearable sensor data for parkinson's disease monitoring using convolutional neural networks. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, ICMI '17, pages 216–220. ACM, 2017.

[51] Olgierd Unold, Maciej Nikodem, Marek Piasecki, Kamil Szyc, Henryk Maciejewski, Marek Bawiec, Paweł Dobrowolski, and Michał Zdunek. Iot-based cow health monitoring system. In Valeria V. Krzhizhanovskaya, Gábor

Závodszky, Michael H. Lees, Jack J. Dongarra, Peter M. A. Sloot, Sérgio Brissos, and João Teixeira, editors, *Computational Science – ICCS 2020*, pages 344–356, Cham, 2020. Springer International Publishing.

[52] Ivan Vaccari, Giovanni Chiola, Maurizio Aiello, Maurizio Mongelli, and Enrico Cambiaso. Mqttset, a new dataset for machine learning techniques on mqtt. *Sensors*, 20(22), 2020.

[53] John Joseph Valletta, Colin Torney, Michael Kings, Alex Thornton, and Joah Madden. Applications of machine learning in animal behaviour studies. *Animal Behaviour*, 124:203–220, 2017.

[54] Mitchell Webber and Raul Fernandez Rojas. Human activity recognition with accelerometer and gyroscope: A data fusion approach. *IEEE Sensors Journal*, 21(15):16979–16989, 2021.

[55] Qingsong Wen, Liang Sun, Fan Yang, Xiaomin Song, Jingkun Gao, Xue Wang, and Huan Xu. Time series data augmentation for deep learning: A survey. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21)*, pages 4653–4660. International Joint Conferences on Artificial Intelligence Organization, 2021.

[56] Susan M. Wernimont, Robin J. Thompson, Scott L. Mickelsen, Spencer C. Smith, Isabella C. Alvarenga, and Kathy L. Gross. Use of accelerometer activity monitors to detect changes in pruritic behaviors: Interim clinical data on 6 dogs. *Sensors*, 18(1), 2018.

[57] Chenqian Yan, Songwei Liu, Hongjian Liu, Xurui Peng, Xiaojian Wang, Fangmin Chen, Lean Fu, and Xing Mei. Hybrid sd: Edge-cloud collaborative inference for stable diffusion models, 2024.

[58] Jianing Zhen, Dehua Mao, Zhen Shen, Demei Zhao, Yi Xu, Junjie Wang, Mingming Jia, Zongming Wang, and Chunying Ren. Performance of xgboost ensemble learning algorithm for mangrove species classification with multi-source spaceborne remote sensing data. *Journal of Remote Sensing*, 4:0146, 2024.

[59] Sijie Zhuo, Lucas Sherlock, Gillian Dobbie, Yun Sing Koh, Giovanni Russello, and Danielle Lottridge. Real-time smartphone activity classification using inertial sensors—recognition of scrolling, typing, and watching videos while sitting or walking. *Sensors*, 20(3), 2020.