

Smart Cattle Behavior Sensing with Embedded Vision and TinyML at the Edge [†]

Jazzie R. Jao ^{1,2,*} , Edgar A. Vallar ²  and Ibrahim Hameed ³ 

¹ Department of Software Technology, College of Computer Studies, De La Salle University Manila, 2401 Taft Avenue, Manila 1004, Philippines

² Department of Physics, College of Science, De La Salle University Manila, 2401 Taft Avenue, Manila 1004, Philippines; edgar.vallar@dlsu.edu.ph

³ Department of Mechanical Engineering and Technology Management, Faculty of Science and Technology, Norwegian University of Life Sciences, Elizabeth Stephansens v. 15, 1433 Ås, Norway;

* Correspondence: jazzie.jao@dlsu.edu.ph

[†] Presented at the 12th International Electronic Conference on Sensors and Applications (ECSA-12), 12–14 November 2025; Available online: <https://sciforum.net/event/ECSA-12>.

Abstract

Accurate real-time monitoring of cattle behavior is essential for enabling data-driven decision-making in precision livestock farming. However, existing monitoring solutions often rely on cloud-based processing or high-power hardware, which are impractical for deployment in remote or low-infrastructure agricultural environments. There is a critical need for low-cost, energy-efficient, and autonomous sensing systems capable of operating independently at the edge. This paper presents a compact, sensor-integrated system for real-time cattle behavior monitoring using an embedded vision sensor and a TinyML-based inference pipeline. The system is designed for low-power deployment in field conditions and integrates the OV2640 image sensor with the Sipeed Maixduino platform, which features the Kendryte K210 RISC-V processor and an on-chip neural network accelerator (KPU). The platform supports fully on-device classification of cattle postures using a quantized convolutional neural network trained on the publicly available cattle behavior dataset, covering standing and lying behavioral states. Sensor data is captured via the onboard camera and preprocessed in real time to meet model input specifications. The trained model is quantized and converted into a K210-compatible .kmodel using the NNCase toolchain, and deployed using MaixPy firmware. System performance was evaluated based on inference latency, classification accuracy, memory usage, and energy efficiency. Results demonstrate that the proposed TinyML-enabled system can accurately classify cattle behaviors in real time while operating within the constraints of a low-power, embedded platform, making it a viable solution for smart livestock monitoring in remote or under-resourced environments.

Keywords: precision livestock farming; TinyML; embedded vision; Kendryte K210; YOLOv2; MobileNet; on-device inference; Edge AI; cattle behavior monitoring; low-power sensing



Published:

Citation: Jao, J.R.; Vallar, E.A.; Hameed, I. Smart Cattle Behavior Sensing with Embedded Vision and TinyML at the Edge. *Eng. Proc.* **2025**, *1*, 0. <https://doi.org/>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Precision Livestock Farming (PLF) represents a data-centric evolution in agricultural science, aiming to optimise both production efficiency and animal welfare through continuous, automated monitoring of individual animals [1]. The core premise of PLF is that

high-resolution, real-time behavioural data empowers farm managers to make proactive, evidence-based decisions. In particular, deviations in diurnal activity patterns such as lying, standing, feeding, and ruminating, often serve as early indicators of metabolic disorders, physical distress, lameness, or critical production events like estrus and parturition. Early PLF systems largely depended on labor-intensive visual observations. Subsequently, devices embedded with accelerometers and GPS have been extensively studied for tracking activity levels and pasture utilisation [2]. More recently, computer vision has emerged as a compelling, non-invasive sensing modality. A single vision sensor can in principle not only determine posture but also identify individuals, assess body condition scores, monitor social interactions, and evaluate engagement with farm infrastructure such as feeders or water troughs. However, the widespread application of vision-based PLF has been constrained by significant computational requirements [3,4]. This computational barrier gives rise to Tiny Machine Learning (TinyML), a growing subfield of AI focused on executing inference pipelines directly on low-power, resource-constrained microcontrollers (MCUs) [5]. TinyML achieves this through efficient model architectures, advanced compression techniques like post-training quantization, pruning, and model distillation, and increasingly capable specialized hardware. Notably, post-training quantization, converting 32-bit floating-point parameters into 8-bit integers (INT8), dramatically reduces model size and enhances inference speed on compatible hardware with minimal loss in accuracy. The emergence of low-cost embedded AI platforms, such as the Kendryte K210 SoC with its integrated KPU (Neural Network Accelerator), represents a critical inflection point [6]. These platforms facilitate efficient, on-device execution of quantized models making complex visual inference feasible at the deep edge, with just a few dollars in cost and milliwatts of power consumption [7]. The K210 and its underlying RISC-V architecture have seen diverse applications ranging from waste monitoring and face mask detection to precision agriculture and traffic management [8–10]. In the realm of livestock monitoring specifically, several studies have explored TinyML-based approaches, including systems focused on cattle behavior recognition, on-device feeding analysis, and multi-modal sensing [11–18].

In this paper, we present a compact, sensor-integrated TinyML system for real-time cattle behavior monitoring, leveraging an embedded vision pipeline with an OV2640 camera module and the Sipeed Maixduino platform (Kendryte K210). The system executes a quantized convolutional neural network (CNN) model directly on the edge, enabling autonomous in-field behavior classification with low latency, low memory footprint, and minimal energy usage. The contribution of this paper are:

1. We demonstrate the feasibility of deploying a quantized YOLOv2-MobileNet_0.75 model on the K210 for real-time livestock behavior detection (standing, eating, drinking, sitting).
2. We characterize system performance in terms of inference latency, memory usage, and detection confidence under class imbalance conditions.

2. Materials and Methods

2.1. Dataset and Preprocessing

The study utilized a publicly available dataset of cattle behavior with four annotated classes: standing, eating, drinking, and sitting. The raw dataset initially contained 1488 annotated samples, which were automatically split into 1460 training samples and 28 validation samples using a 90/10 split. To address class imbalance, data balancing was applied by oversampling minority classes (drinking and sitting), resulting in a final training set of 3388 images and a validation set of 28 images. Each image was resized to 224×224 pixels and standardized with mean ($\mu = 123.5$) and standard deviation ($\sigma = 58.4$).

Although augmentation options such as rotation, mirroring, and blur were available, they were disabled in this experiment.

2.2. Model Architecture and Training

A transfer learning approach was adopted, using YOLOv2 as the detection framework with a MobileNet_0.75 backbone pretrained on ImageNet. The model was implemented on the mncase platform and trained for 100 epochs with a batch size of 16 and a learning rate of 0.001. The minimum bounding box size was set to 10 pixels, and negative data samples were included to improve robustness. In this context, “negative data” refers to images without annotated targets. If some images contain unlabeled target objects, negative data must be disabled to avoid the risk of treating true objects as false detections. Conversely, when all objects of interest are fully annotated, enabling negative data allows the inclusion of empty scenes containing no target objects, thereby improving the model’s ability to discriminate between relevant and irrelevant inputs. Anchor clustering was performed on the dataset, producing five anchor shapes with width–height ratios $[0.49, 0.77, 1.18, 1.40, 1.44]$ and an Intersection over Union (IoU) accuracy of 76.7%. The final network consisted of 1.87M parameters, of which 34,605 were trainable and the remainder frozen, ensuring suitability for quantization and embedded deployment.

2.3. Embedded Deployment Pipeline

The trained model was quantized and converted into a K210-compatible .kmodel file using the NNCase toolchain. Deployment was conducted on the Sipeed Maixduino development board, which integrates the Kendryte K210 RISC-V processor. The K210 features a dual-core CPU and a dedicated neural network accelerator (KPU, K210 Processing Unit) capable of real-time CNN inference at low power. The inference pipeline was implemented in MaixPy, enabling direct execution of quantized CNN models without external dependencies. Figure 1 shows the image of the board used.

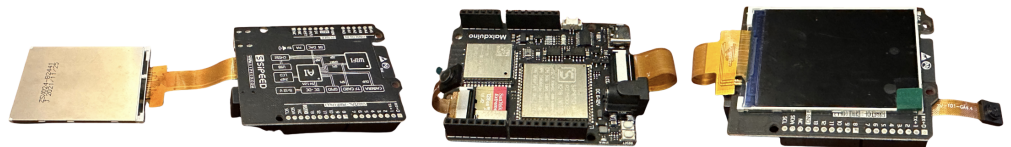


Figure 1. The Sipeed Maixduino.

2.4. Sensor and Hardware Integration

For visual sensing, the system employed the OV2640 CMOS image sensor, a compact 2-megapixel module widely used in embedded vision tasks due to its low power consumption and configurable output formats (RGB565, YUV422, JPEG). The sensor was interfaced with the Maixduino via an 8-bit DVP (Digital Video Port), providing real-time image capture at frame rates suitable for behavior monitoring. Peripheral components included onboard SRAM for buffer management and an SD card interface for optional dataset logging. The compact design allowed for autonomous operation in resource-constrained environments without requiring external computation or cloud connectivity.

2.5. System Workflow

The end-to-end pipeline is illustrated in Figure 2. Captured frames from the OV2640 sensor were preprocessed on-device to match the CNN input requirements (224×224). The quantized YOLOv2-MobileNet model, executed on the KPU, performed bounding box regression and classification in real time. Detection outputs were evaluated against ground truth annotations (white boxes), with green boxes indicating correct predictions and red

boxes marking false detections, to enable continuous monitoring of cattle behavior directly at the edge device without reliance on cloud services or high-power hardware.

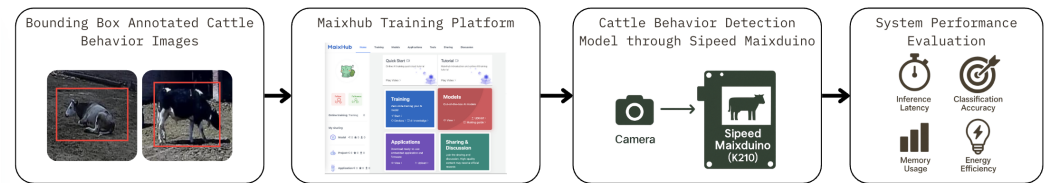


Figure 2. End-to-end embedded pipeline.

3. Results and Discussion

Model predictions were evaluated using annotated bounding boxes (white), predicted boxes (green), and incorrect detections (red). Correct detections occurred when predicted boxes overlapped with annotated targets, even if not all instances were recognized. As shown in Figure 3, the model successfully identified multiple behaviors (standing, eating, sitting, drinking) with high confidence scores (0.9–1.0). However, cases of misclassification and partial recognition were observed, especially in crowded scenes or low-light conditions, where multiple animals overlapped or when visibility was reduced.



Figure 3. Sample Predictions of the Trained Model.

Figure 4 illustrates the loss and accuracy curves over 100 epochs. The training loss decreased rapidly within the first 20 epochs and stabilized around 0.1, indicating effective convergence. Validation accuracy improved steadily, reaching approximately 0.85, which suggests that the model generalized well to unseen data. The best test set performance was achieved at epoch 80, with an accuracy of 0.875, highlighting the point of optimal generalization before minor fluctuations were observed in later epochs.

The model demonstrated reliability in detecting eating and sitting behaviors, which were represented with stronger annotation counts in the training set. Drinking was the most challenging category, with lower confidence predictions and higher misclassification rates, likely due to fewer training samples and higher visual similarity with standing. Furthermore, the use of balanced oversampling allowed minority classes such as drinking to be recognized, but some false positives remained, as indicated by red boxes in Figure 3. Complex environments such as night lighting and occlusion still introduced errors.

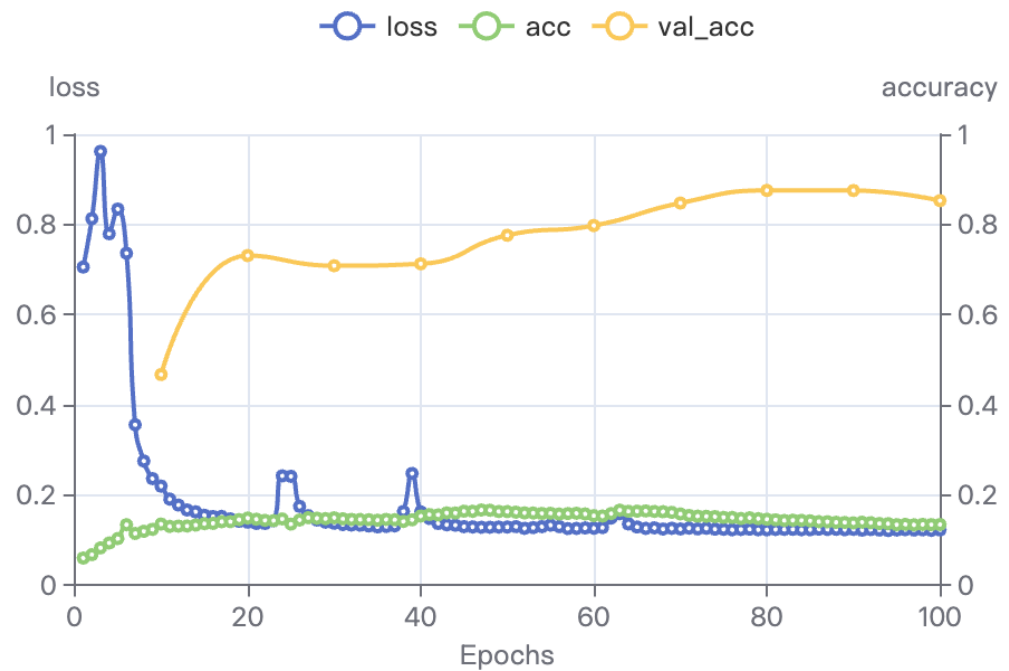


Figure 4. The Learning Curve.

3.1. On-Device Inference with Sipeed Maixduino

To demonstrate the real-time deployment, the quantized `.kmodel` was executed on the Sipeed Maixduino board. Figure 5 illustrates representative inference results captured directly from the device output display. The system detected cattle behaviors such as standing and eating with confidence levels ranging from 0.6 to 0.9. The bounding boxes were rendered directly by the device, confirming that the complete pipeline from image acquisition using the OV2640 sensor to preprocessing, inference, and visualization, operated fully on-device without external computation. Performance remained robust even under low-light conditions (left panel of Figure 5), though confidence values were slightly reduced compared to daylight scenarios. In brighter environments, the system maintained stable bounding box predictions and higher confidence scores.



Figure 5. On-device inference results from the Sipeed Maixduino board using the quantized YOLOv2-MobileNet model. Cattle behaviors such as standing and eating were detected in real time with confidence scores between 0.6 and 0.9.

3.2. Detection Performance

The detection results covered four cattle behavior classes: Standing, Eating, Drinking, and Sitting. Table 1 summarizes the detection frequency and confidence values.

Table 1. Detection performance across behavior classes.

Class	Total Detections	Avg. Score	Max Score	Min Score
Standing	1484	0.78	0.99	0.50
Eating	606	0.82	0.98	0.51
Drinking	22	0.77	0.96	0.59
Sitting	6	0.62	0.77	0.54

As shown, Standing was the most frequently detected behavior, followed by Eating. Both classes achieved high confidence scores, with averages above 0.78. By contrast, under-represented behaviors such as Drinking and especially Sitting yielded fewer detections and lower confidence values. This reflects the dataset imbalance, where more common behaviors dominate the predictions and less frequent ones are harder to detect reliably. Inference metrics were collected from deployment on the Sipeed Maixduino board. Table 2 presents the summary statistics.

Table 2. Summary of inference performance on Sipeed Maixduino (K210).

Metric	Average	Range
Latency (ms)	35.0	34–37
Memory Allocated (bytes)	350,427	—
Memory Free (bytes)	167,717	—
Detections per Frame	2.18	—

The system maintained stable real-time performance, with an average inference latency of 35 ms per frame, corresponding to approximately 28–30 frames per second (FPS). Memory usage remained efficient, with an average of 350 kB allocated and 168 kB free, confirming that the quantized YOLOv2-MobileNet model fits comfortably within the K210's constraints. On average, the system detected about two cattle per frame, consistent with the multi-animal monitoring scenario. On the hardware side, the deployment on the Maixduino platform proved successful, achieving real-time inference with low latency and modest memory usage. These results validate the suitability of the proposed TinyML-enabled system for autonomous cattle behavior monitoring at the edge, without reliance on cloud services or high-power computing infrastructure.

4. Conclusions

The YOLOv2 + MobileNet_0.75 transfer learning framework demonstrated strong detection performance, with validation accuracy stabilizing at approximately 85%. These results highlight the suitability of the approach for real-time livestock monitoring on resource-constrained embedded platforms. Nevertheless, there remain opportunities for improvement. Future work could incorporate more advanced data augmentation techniques, such as brightness adjustment and affine transformations, to enhance robustness against variable lighting and environmental conditions. In addition, expanding the dataset, particularly with greater representation of under-sampled behaviors such as drinking, would help address class imbalance and improve generalization. Further gains may also be achieved by exploring anchor-free detection architectures such as YOLOv5 or YOLOv8, which are known to provide better adaptability in cluttered or dynamic environments. By addressing these aspects, the system can be further refined to reduce false positives, enhance classification consistency, and improve its reliability as a practical solution for precision livestock farming applications.

Author Contributions: Conceptualization, J.R.J., E.A.V. and I.H.; methodology, J.R.J., E.A.V. and I.H.; formal analysis, J.R.J., E.A.V., I.H., E.A.V. and I.H.; investigation, J.R.J., E.A.V., I.H., E.A.V. and I.H.; resources, J.R.J., E.A.V. and I.H.; data curation, J.R.J., E.A.V. and I.H.; writing—original draft preparation, J.R.J. and E.A.V.; writing—review and editing, J.R.J., E.A.V. and I.H.; visualization, J.R.J., E.A.V. and I.H.; supervision, E.A.V. and I.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data in this study can be accessed in this <https://github.com/jazziejao/Smart-Cattle-Speed-Maixduino>, Github Repository.

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

References

1. Kleen, J.L.; Guatteo, R. Precision Livestock Farming: What Does It Contain and What Are the Perspectives? *Animals* **2023**, *13*, 779. <https://doi.org/10.3390/ani13050779>.
2. Ding, L.; Zhang, C.; Yue, Y.; Yao, C.; Li, Z.; Hu, Y.; Yang, B.; Ma, W.; Yu, L.; Gao, R.; et al. Wearable Sensors-Based Intelligent Sensing and Application of Animal Behaviors: A Comprehensive Review. *Sensors* **2025**, *25*, 4515.
3. Senoo, E.E.K.; Anggraini, L.; Kumi, J.A.; Karolina, L.B.; Akansah, E.; Sulyman, H.A.; Mendonça, I.; Aritsugi, M. IoT solutions with artificial intelligence technologies for precision agriculture: Definitions, applications, challenges, and opportunities. *Electronics* **2024**, *13*, 1894.
4. Hayajneh, A.M.; Aldalahmeh, S.A.; Alasali, F.; Al-Obiedollah, H.; Zaidi, S.A.; McLernon, D. Tiny machine learning on the edge: A framework for transfer learning empowered unmanned aerial vehicle assisted smart farming. *IET Smart Cities* **2024**, *6*, 10–26.
5. Srinivasagan, R.; El Sayed, M.S.; Al-Rasheed, M.I.; Alzahrani, A.S. Edge intelligence for poultry welfare: Utilizing tiny machine learning neural network processors for vocalization analysis. *PLoS ONE* **2025**, *20*, e0316920.
6. Giordano, L. AIoT: A Kendryte K210 proof of concept. 2020.
7. Torres-Sánchez, E.; Alastruey-Benedé, J.; Torres-Moreno, E. Developing an AI IoT application with open software on a RISC-V SoC. In Proceedings of the 2020 XXXV Conference on Design of Circuits and Integrated Systems (DCIS), Segovia, Spain, 18–20 November 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–6.
8. Nagar, M.S.; Chauhan, V.; Chinchavde, K.M.; Swati, S.; Engineer, P. Energy-Efficient Acceleration of Deep Learning based Facial Recognition on RISC-V Processor. In Proceedings of the 2023 11th International Conference on Intelligent Systems and Embedded Design (ISED), Dehradun, India, 15–17 December 2023; pp. 1–6. <https://doi.org/10.1109/ISED59382.2023.10444594>.
9. Christofas, V.; Amanatidis, P.; Karampatzakis, D.; Lagkas, T.; Goudos, S.K.; Psannis, K.E.; Sarigiannidis, P. Comparative Evaluation between Accelerated RISC-V and ARM AI Inference Machines. In Proceedings of the 2023 6th World Symposium on Communication Engineering (WSCE), Thessaloniki, Greece, 27–29 September 2023; pp. 108–113. <https://doi.org/10.1109/WSCE59557.2023.10365853>.
10. Zhang, G.; Li, Z.; Huang, D.; Luo, W.; Lu, Z.; Hu, Y. A Traffic Sign Recognition System Based on Lightweight Network Learning. *J. Intell. Robot. Syst.* **2024**, *110*, 139.
11. Zhang, Q.; Kanjo, E. MultiCore+ TPU Accelerated Multi-Modal TinyML for Livestock Behaviour Recognition. *arXiv* **2025**, arXiv:2504.11467.
12. Viswanatha, V.; Ramachandra, A.; Hegde, P.T.; Hegde, V.; Sabhahit, V. Tinyml-based human and animal movement detection in agriculture fields in india. In Proceedings of the International Conference on Emerging Research in Computing, Information, Communication and Applications, Bangalore, India, 24–25 February 2023; Springer: Berlin/Heidelberg, Germany, 2023; pp. 49–65.
13. Chen, Y.S.; Rustia, D.J.A.; Huang, S.Z.; Hsu, J.T.; Lin, T.T. IoT-Based System for Individual Dairy Cow Feeding Behavior Monitoring Using Cow Face Recognition and Edge Computing. *Internet Things* **2025**, *33*, 101674.
14. Smink, M.; Liu, H.; Döpfer, D.; Lee, Y.J. Computer Vision on the Edge: Individual Cattle Identification in Real-Time With ReadMyCow System. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 4–8 January 2024; pp. 7056–7065.
15. Raza Shirazi, S.A.; Fatima, M.; Wahab, A.; Ali, S. A Novel Active RFID and TinyML-based System for Livestock Localization in Pakistan. *Sir Syed Univ. Res. J. Eng. Technol. (SSURJET)* **2024**, *14*, 33–38.
16. Bartels, J.; Tokgoz, K.K.; A, S.; Fukawa, M.; Otsubo, S.; Li, C.; Rachi, I.; Takeda, K.I.; Minati, L.; Ito, H. TinyCowNet: Memory- and Power-Minimized RNNs Implementable on Tiny Edge Devices for Lifelong Cow Behavior Distribution Estimation. *IEEE Access* **2022**, *10*, 32706–32727. <https://doi.org/10.1109/ACCESS.2022.3156278>.
17. Farhan, M.; Wijaya Thaha, G.S.; Alvito Kristiadi, E.; Mutijarsa, K. Cattle Anomaly Behavior Detection System Based on IoT and Computer Vision in Precision Livestock Farming. In Proceedings of the 2024 International Conference on Information Technology

- Systems and Innovation (ICITSI), Bandung, Indonesia, 12 December 2024; pp. 342–347. <https://doi.org/10.1109/ICITSI65188.2024.10929192>.
18. Martinez-Rau, L.S.; Chelotti, J.O.; Giovanini, L.L.; Adin, V.; Oelmann, B.; Bader, S. On-Device Feeding Behavior Analysis of Grazing Cattle. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 2512113. <https://doi.org/10.1109/TIM.2024.3376013>.

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