Predicting Daily Milk Yield in Dairy Cattle Using Multi-View Behavioral Recognition and Visual Identification

María Alejandra Galicia Almaraz[1], Timilehin Ayanlade[2] and Soumik Sarkar[2]

[1]Instituto Tecnológico Autónomo de México, Ciudad de México, México.
[2]Iowa State University, Iowa, USA.

Ames, Iowa, USA.

Abstract

Monitoring dairy cow behavior is crucial for optimizing productivity and animal welfare. Traditional methods, such as wearable sensors, face challenges in scalability, maintenance, and long-term deployment across large herds. Computer vision offers a non-invasive alternative, but a key barrier to its adoption has been reliably identifying individual cows across multi-view camera systems, which prevents accurate attribution of behavioral patterns to specific animals and hinders the link between behavior and milk yield.

This study addresses these gaps by proposing a computer vision-based framework that integrates individual cow detection, behavior classification, and visual identification in multi-view settings to predict daily milk yield. Utilizing the MmCows dataset, comprising video footage from 16 cows over 15 days, we developed a modular system. This system uses YOLOv8 for cow detection, ResNet18 for classifying seven key behaviors (achieving 0.91 test set accuracy), and dual EfficientNet models for individual identification. A novel "Smart Filtering" module was implemented for multi-view data (Phase 2) to improve identification accuracy by re-analyzing unidentified detections based on Intersection over Union (IoU) and centroid distance. Beyond behavioral data, environmental variables (temperature, humidity, THI) and previous day's milk yield were included for comprehensive predictive modeling using XGBoost and Random Forest regressors.

Our findings demonstrate a significant improvement in individual cow identification in multi-view settings (Phase 2), reducing identification loss from approximately 25% in Phase 1 to about 9%. This was particularly beneficial for Lying Behavior (Behavior_6), which was frequently unidentified in single-view scenarios. The inclusion of previous day's milk yield consistently and substantially improved milk yield prediction accuracy, with the Random Forest model in Phase 1, Scenario 3 (behaviors + previous day's milk yield + weather) achieving an R^2 of 0.79 and MSE of 6.64. While environmental variables offered subtle improvements, behaviors such as 'Lying' (Behavior_6), 'Drinking' (Behavior_5), and 'Feeding Head Down' (Behavior_3) consistently emerged as biologically significant predictors of milk production.

In conclusion, this integrated deep learning framework establishes a robust and quantitative connection between animal behavior and economic productivity. By overcoming challenges in individual identification through multi-view analysis, this research lays the groundwork for scalable, non-invasive precision livestock farming systems that can synergistically enhance animal welfare and profitability.

Introduction

Monitoring the behavior of dairy cows is essential for optimizing both productivity and animal welfare. While traditional approaches—such as wearable sensors—have enabled high-resolution behavioral tracking, they pose challenges in scalability, maintenance, and long-term deployment across large herds (Sun et al., 2021). As a non-invasive and potentially more scalable alternative, computer vision has gained attention. Yet, one of the main technical barriers to its adoption lies in reliably identifying individual cows across multi-view camera systems (Jia et al., 2023).

This limitation has far-reaching consequences. Without accurate individual identification, behavioral patterns cannot be correctly attributed to specific animals, making it difficult to establish reliable links between behavior and milk yield (Chun-hua et al., 2019). As a result, the practical utility of vision-based monitoring systems in commercial farm settings remains limited.

Despite the growing interest in computer vision applications for livestock management, most studies have focused either on behavior classification or general activity monitoring, often overlooking the fundamental challenge of cow-level identification. Moreover, there is limited evidence on how identification quality influences the ability to model productivity outcomes. This lack of integration between individual behavior profiles and predictive analytics constrains the development of precision dairy systems capable of guiding management decisions.

This study addresses these gaps by proposing a computer vision-based framework that tracks and analyzes individual cows' behaviors using multi-view image data. It incorporates environmental variables—such as ambient temperature, humidity, and THI—and includes each cow's milk yield from the previous day to contextualize daily activity profiles. The central hypothesis is that specific behaviors (e.g., lying, feeding, walking) are statistically associated with next-day milk production (Chun-hua et al., 2019), and that the quality of individual identification significantly affects the robustness of predictive models.

By focusing on these challenges, this research contributes to the ongoing development of scalable, non-invasive monitoring systems that can support evidence-based decision-making in commercial dairy farms.

Literature Review

The optimization of dairy production and animal welfare critically depends on effective behavioral monitoring. Research consistently demonstrates the profound importance of specific behavioral patterns on milk yield and physiological well-being (Chun-hua et al., 2019; Jensen et al., 2005; Munksgaard and Lovendahl, 1993; Munksgaard et al., 2005). For instance, lying behavior plays a crucial role in milk performance (Chun-hua et al., 2019). Dairy cows typically spend between 12 and 14 hours daily lying in cubicle beds for rest and rumination (Jensen et al., 2005). This extended lying time is vital because it significantly increases blood flow to the mammary gland—by 25% to 50%—which directly contributes to increased milk production (Rulquin and Caudal, 1992). Adequate resting time, up to 14 hours per day for highly productive cows, is also associated with reduced stress on the feet, decreased lameness, enhanced feeding activity, and improved rumination and overall health (Grant, 2006). Studies have shown that high-yielding cows spend more time (12.5 vs 10.5 h/day) lying down and resting compared to lower-yielding cows (Janocha et al., 2023).

Other behaviors, such as feeding frequency and duration, also provide valuable insights into a cow's physiological status and production potential (Chun-hua et 2019). Increased feeding frequency has been positively correlated with improved milk yield (r=0.686) (Chun-hua et al., 2019). These findings underscore the strong biological and empirical basis for the relationship between cow behavior and productivity.

Historically, monitoring cow behavior relied on manual observation and recording, a method characterized by its time-consuming nature, inefficiency, subjectivity, and impracticality for continuous, large-scale surveillance (Tian et al., 2021). The evolution of precision livestock farming has led to the adoption of advanced monitoring technologies, including sensor-based systems. Wearable sensors, such as accelerometers, IMUs, and smart collars, have been developed to collect movement and physiological data for behavior identification (Vessies et al., 2014). These devices have proven effective in detecting behavioral states and establishing correlations with daily milk yield (Chun-hua et al., 2019; Tian et al., 2021). For example, studies have employed leg tags to classify behaviors like feeding, lying, standing, walking, and running (Wang et al., 2019) , or neck-mounted accelerometers for feeding and rumination analysis (Arcidiacono et al., 2017; Benaissa et al., 2019). While sensor-based approaches offer rich data, they face inherent limitations in scalability, maintenance costs, and susceptibility to damage or improper placement in commercial large-scale settings (Tian et al., 2021). Crucially, most sensor systems decouple individual cow identification from behavior analysis; while collar IDs facilitate tracking for the sensor itself, integrating this data with visually ambiguous cows in a multi-animal setting often necessitates a separate, cumbersome identification process (Tian et al., 2021). This reliance on pre-identified sensors creates "double-work" and limits true automation and scalability for generalized behavioral monitoring without physical tags.

In contrast, computer vision systems offer a promising non-invasive alternative for automated behavior monitoring, with inherent advantages in scalability and simultaneous multi-animal analysis (Tian et al., 2022). Significant advancements have been made in developing deep learning models for behavior classification. For instance, Jia et al. (2024) developed CAMLLA-YOLOv8n, an improved YOLOv8n model, achieving high precision (94.46%) in recognizing seven cow behaviors, including grazing, standing, and lying. Similarly, our study utilizes a ResNet18 neural network to classify seven key behaviors (Chun-hua et al., 2019). However, a major unresolved challenge in vision-based systems is the accurate and robust identification of individual cows in complex, multi-view environments, particularly when animals are partially occluded, overlapping, or in varying postures like lying down. For example, the study by Jia et al. (2024), while achieving high accuracy in behavior classification with YOLOv8n, focuses on object detection for behavior recognition and does not explicitly detail a robust, automated individual identification pipeline for *unmarked* cows under occluded or multi-view conditions, neglecting the scalability barrier posed by complex real-world occlusion (Jia et al., 2024). This critical problem of reliable individual identification remains an open challenge for consistently linking observed behavior to specific milk production data for each animal without relying on external tags or extensive manual annotation (Tian et al., 2022). Current systems often decouple identification and behavior analysis, whereas our integrated pipeline aims to address both simultaneously, reducing farm-level implementation complexity by focusing on the simultaneous detection, classification, and visual identification of cows in a unified system.

Furthermore, environmental factors, particularly thermal conditions, profoundly influence both cow behavior and milk production, necessitating their inclusion in predictive models. Heat stress (HS), commonly measured by the Temperature-Humidity Index (THI), significantly impacts milk yield, feed-to-milk efficiency, and animal comfort (Rodriguez-Venegas et al., 2023). Studies in arid regions demonstrate that total milk production decreases as THI increases, with notable declines in milk production and lying time observed at THI levels from 68-71 onwards (Rodriguez-Venegas et al., 2023). Heat stress reduces dry matter intake (DMI) and feed conversion efficiency (FCE), redirecting energy from milk production towards thermoregulation (Rodriguez-Venegas et al., 2023). The percentage of milking cows also diminishes under increased THI, often due to fertility issues associated with heat stress (Rodriguez-Venegas et al., 2023). Seasonal variations in temperature and humidity further affect milk production, DMI, FCE, and lying time, with optimal values typically seen in cooler periods like winter and spring, and suboptimal values in warmer periods like summer (Rodriguez-Venegas et al., 2023). These established correlations highlight the importance of incorporating environmental variables into predictive models for a holistic understanding of milk yield.

This study directly addresses the identified gaps by proposing an integrated deep learning-based approach that unifies individual cow detection, behavior classification, and visual identification in multi-view settings. Unlike previous multimodal approaches that often rely on wearable sensors, this work uses only visual data combined with environmental variables to predict daily milk production (Tian et al., 2022). By developing a robust methodology for quantifying the time each cow spends performing specific activities, even under challenging conditions of occlusion or varying postures, this research aims to provide accurate behavioral data for predictive modeling. Given the established significant positive correlations between milk yield and behaviors like lying time ($r=0.686$) and feeding frequency ($r=0.595$) (Chun-hua et al., 2019), our model prioritizes comprehensive temporal behavior tracking across camera views to capture this metric reliably. This approach is anticipated to overcome the limitations of current identification methods and achieve high accuracy (greater than 85%) in predicting daily milk production, thereby establishing a robust, quantitative, and non-invasive connection between animal behavior and economic productivity in dairy farming (Chun-hua et al., 2019).

Methodology

Data Acquisition and Study Population

The study utilized the MmCows dataset, a comprehensive, open-source multimodal repository. This dataset comprises video footage collected over a 15-day period from 16 dairy cows. The data represents a real-world deployment scenario, providing a rich source for behavioral analysis and milk yield prediction in a typical farm environment (Vu H. et al., 2024).

Two distinct experimental settings, referred to as "Phases" were established for data processing and model evaluation:

- **Phase 1: Single-View Analysis.** In this phase, images were extracted from a single camera (Camera 1) every 100 seconds. This setup simulated a common single-camera monitoring scenario.
- **Phase 2: Multi-View Analysis.** For this phase, processed images containing four distinct views of the barn simultaneously were utilized, arranged in a 2x2 grid. Images were extracted every 3 seconds from continuous videos, capturing approximately 45 seconds of real-time behavior. This multi-view approach aimed to address limitations inherent in single-view systems, such as occlusions and partial visibility

Video data from the MmCows dataset was processed frame-by-frame to generate daily profiles for each cow. The specific behaviors targeted for classification included: 'Walking' (0), 'Standing' (1), 'Feeding head up' (2), 'Feeding head down' (3), 'Licking' (4), 'Drinking' (5), and 'Lying' (6). In addition to behavioral data, environmental variables such as temperature, humidity, and the Temperature-Humidity Index (THI), along with the cow's milk yield from the previous day, were collected for the predictive modeling stage, these data were also provided by the MmCows dataset.

Instrumentation and Techniques

The automated computer vision system employed a modular architecture comprising three specialized deep learning models. It's important to note that the training of YOLOv8 and EfficientNet models was part of the existing dataset framework and was not performed by the authors of this study. This approach allowed for a focus on the integration and application of these models for behavioral analysis and milk yield prediction.

1. **Cow Detection (YOLOv8):** A YOLOv8 (You Only Look Once, version 8) model was utilized to detect all visible cows within each image. For each detected cow, the model generated a bounding box (x1, y1, x2, y2 coordinates) and a detection confidence score. The output of this model served as the input for subsequent behavior classification and individual identification models. Examples of YOLOv8 application in Phase 1 and Phase 2 are shown in Figure 1 and Figure 2, respectively.

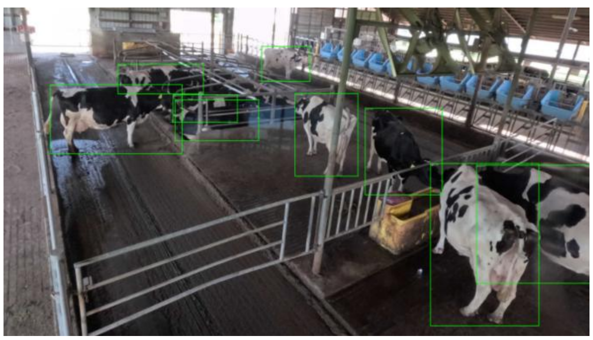


Figure 1. Cow detection example – Phase 1

Figure 2. Cow detection example – Phase 2

2. **Behavior Classification (ResNet18):** Cropped images of each detected cow (derived from the YOLOv8 bounding boxes) were fed into a ResNet18 neural network for behavior classification. This model classified the cow's behavior into one of the seven predefined categories , returning a behavior_id and a probability score. The ResNet18 model was trained on thousands of manually annotated cow images to enable it to distinguish subtle postures and different viewing angles. Its test set accuracy was 0.91.
3. **Individual Identification (EfficientNet - Dual Models):** To identify individual cows, two distinct EfficientNet models were trained. This dual-model approach was implemented because cows appear significantly different when lying down compared to when they are standing or walking. Each identification model received the cropped cow image and predicted its unique cow_id. Only predictions with a high confidence score (greater than a threshold, e.g., 0.60) were accepted; otherwise, the cow_id was marked as -1 (unknown cow). Examples of identification without filtering in Phase 1 and Phase 2 are shown in Figure 3 and Figure 4, respectively.

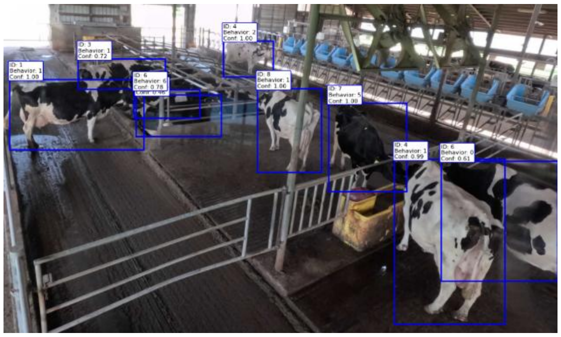


Fig 3. Individual identification without filtering - Phase 1

Fig 4. Individual identification without filtering - Phase 2

Smart Filtering for Multi-View Data (Phase 2 only)

For Phase 2, which involved multi-view images, a smart_filter_detections module was implemented to address challenges in reliable individual cow identification, especially in cases of low visibility, occlusions, or unfavorable angles. This intelligent filter aimed to reduce information loss by re-analyzing unidentified detections (cow_id = -1) and comparing them with correctly identified cows within the same image. The filtering process was based on two fundamental principles:

- **Intersection over Union (IoU):** For each unidentified cow, the IoU between its bounding box and the bounding boxes of already identified cows was calculated. If the IoU value exceeded a defined threshold (e.g., 0.70), it was assumed that both boxes might correspond to the same animal viewed from a slightly different angle or position.
- **Centroid Distance:** As a complementary criterion, the Euclidean distance between the centers of the two bounding boxes was calculated. If this distance was sufficiently small (e.g., less than 60 pixels), it was also interpreted as strong evidence that it was the same individual.

If an unidentified cow met at least one of these criteria relative to a previously identified cow in the same image, its cow_id was re-assigned, marked as was_reassigned = True, and its source labeled as "reassigned". Additionally, to prevent duplicate counting of the same animal in the processed image, the filter retained only one detection per identified cow, prioritizing the one with the highest cow_confidence. Other detections with the same cow_id were discarded. Cows that could not be re-assigned through either IoU or centroid distance remained unidentified (cow_id = -1) and were labeled as source = "unknown" and was_reassigned = False. This procedure consistently increased the number of correctly identified cows in each image without requiring full reprocessing or manual intervention.

Data Analysis and Predictive Modeling

For each processed image, a DataFrame was generated containing the cow_id (or -1 if unidentified), behavior_id, cow_confidence, source of detection (direct, reassigned, or unknown), was_reassigned status, and timestamps (date, image). This DataFrame was then used to normalize and quantify daily behaviors per cow (e.g., how many times the cow is lying, walking, feeding).

These quantified behavioral data, along with contextual variables, served as input for supervised machine learning models, specifically XGBoost and Random Forest regressors, aimed at predicting daily milk production. Three experimental scenarios were tested for milk yield prediction:

1. **Scenario 1: Behavioral data only.** This scenario exclusively used the daily quantified behavioral data as input fea tures.
2. **Scenario 2: Behavior + previous-day milk yield.** In addition to behavioral data, the milk production from the previous day (Milk_Yesterday) was included as an input variable.
3. **Scenario 3: Behavior + previous-day milk yield + environmental variables.** This comprehensive scenario incorporated behavioral data, previous-day milk yield, and environmental variables (ambient temperature, humidity, and THI).

The primary objective of these predictive models was to identify which behavioral patterns were most strongly correlated with milk yield , thus facilitating more informed and proactive livestock management. Model performance was evaluated using Mean Squared Error (MSE) and R-squared (R^2) metrics. Feature importance and correlation matrices were also analyzed to understand the contribution of individual behaviors and variables to the prediction.

Results

This section presents the empirical results obtained from the automated computer vision system and the subsequent machine learning models for predicting daily milk yield. The evaluation is structured across two experimental phases, reflecting different camera perspectives, and assesses the impact of various feature combinations on prediction accuracy.

Visual Model Performance (Cow Detection, Behavior Classification, and Individual Identification)

The initial phase of the study focused on the performance of the visual models (YOLOv8, ResNet18, and EfficientNet) in detecting, classifying behaviors, and identifying individual cows across both single-view (Phase 1) and multi-view (Phase 2) imaging setups.

In Phase 1, analyzing approximately 550 images from a single camera, the system correctly identified approximately 5,500 cows (behavior and ID with >60% probability). However, an estimated 1,800 cows (~25% of the total identified cows) lost their individual identification, despite their behavior being classified. Notably, about 70% of these unidentified cows exhibited 'lying' behavior (behavior_6). Theoretically, with 16 cows across 550 images, 8,800 cow detections would be expected. Out of this hypothetical total, 5,500 were correctly identified, 1,800 had behavior classified but no ID, leaving approximately 1,500 cows (about 17% of the hypothetical total) with no record.

Phase 2, utilizing approximately 1,920 multi-view (2x2 grid) images, showed an improvement in identification. Approximately 18,000 cows were correctly identified (behavior and ID with >60% probability). The number of lost individual identifications reduced to approximately 1,800 cows, representing about 9% of the total identified cows. Considering a theoretical total of 30,720 cow detections (1920 images * 16 cows, after accounting for filtering out duplicates from multiple views), 18,000 were correctly identified, and 1,800 lacked an ID, resulting in approximately 10,920 cows (around 35% of the hypothetical total) with no record. The improved identification accuracy in Phase 2 highlights the benefit of multi-view imaging and the intelligent filtering process.

Regarding behavior classification, the ResNet18 model achieved a test set accuracy of 0.91. The training set consisted of 149,580 samples, the validation set of 32,052 samples, and the test set of 32,054 samples, with 10 epochs used for optimization. A confusion matrix of cow behaviors are presented in Figure 5.

Fig5. Confusion matrix of cow behavior obtained after training the ResNet18 model

Milk Yield Prediction Model Performance

The performance of the Random Forest and XGBoost regressors in predicting daily milk yield was evaluated across both phases and the three defined feature combination scenarios: (1) Behaviors only, (2) Behaviors + Milk Production Yesterday, and (3) Behaviors + Milk Production Yesterday + Weather Conditions.

Phase 1 Results

Table 1. Summary of the Mean Squared Error (MSE) and R-squared (R^2) values for Random Forest and XGBoost models in Phase 1.

	Random Forest			XGBoost		
Scenarios	**1**	**2**	**3**	**1**	**2**	**3**
MSE	33.66	6.96	6.64	34.74	8.53	8.93
R^2	0.16	0.78	0.79	0.13	0.73	0.72

As shown in Table 1, the best performance in Phase 1 was achieved by the Random Forest model in Scenario 3 (Behaviors + Milk Production Yesterday + Weather Conditions), with an MSE of 6.64 and an R^2 of 0.79. This indicates that the model explained 79% of the variance in daily milk production, with predictions deviating by approximately 2.57 kg ($\sqrt{6.64}$). The improvement in R^2 was substantial when including Milk_Yesterday (Scenario 2 vs. Scenario 1), underscoring its predictive power.

Phase 2 Results

Table 2. Summary of the Mean Squared Error (MSE) and R-squared (R^2) values for Random Forest and XGBoost models in Phase 2.

	Random Forest			XGBoost		
Scenarios	**1**	**2**	**3**	**1**	**2**	**3**
MSE	20.45	7.70	8.55	17.06	12.73	8.68
R^2	0.38	0.75	0.75	0.48	0.64	0.73

In Phase 2, the XGBoost model in Scenario 1 (Behaviors only) achieved an R^2 of 0.48 with an MSE of 17.06, demonstrating a more robust baseline for behavior-only predictions compared to Phase 1. When Milk_Yesterday was included (Scenario 2 and 3), the models again showed significant improvements in performance

Visualizations

To further illustrate these findings, the following figures are included:

Figure 6. Feature Importance for Random Forest (Phase 1, Scenario 3): This bar chart visualizes the relative importance of each feature (behaviors, previous day's milk yield, and weather conditions) in the best-performing model from Phase 1.

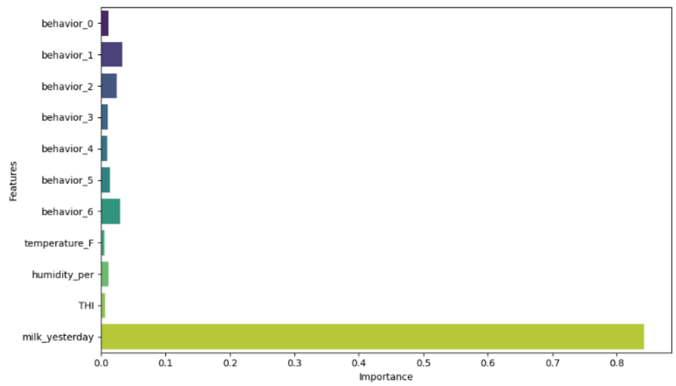


Figure 7. Correlation Matrix for Random Forest (Phase 1, Scenario 3): *This* heatmap displays the correlation coefficients between all input variables and the daily milk yield for the best-performing model.

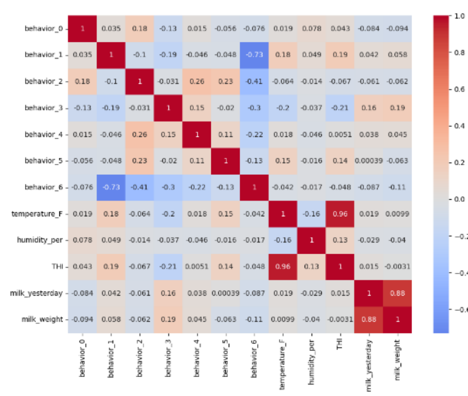


Figure 8. Predicted vs. Actual Milk Production for Random Forest (Phase 1, Scenario 3): This scatter plot compares the predicted milk yield values against the actual observed values, providing a visual assessment of the model's accuracy.

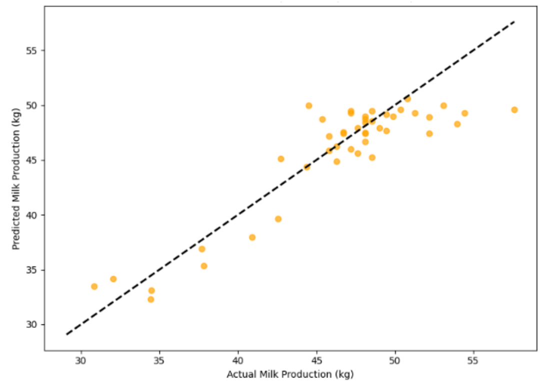


Figure 9. Feature Importance for XGBoost (Phase 2, Scenario 1): This bar chart illustrates the feature importance for the XGBoost model when only behavioral data was used in Phase 2.

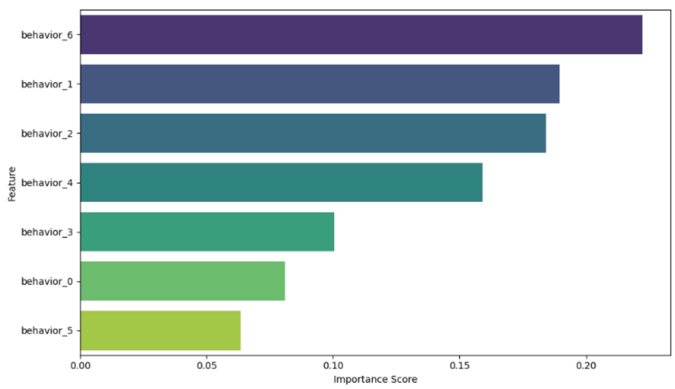


Figure 10. Correlation Matrix for XGBoost (Phase 2, Scenario 1): This heatmap shows the correlation coefficients between the behavioral variables and daily milk yield for the behavior-only model in Phase 2.

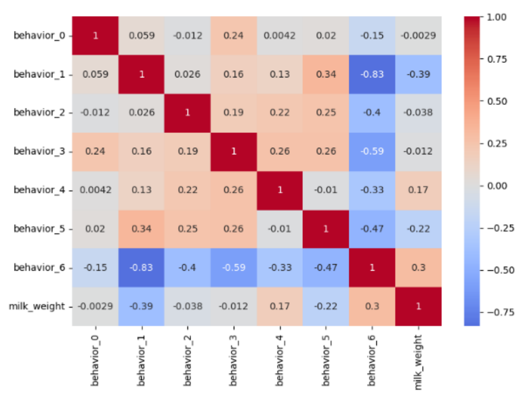
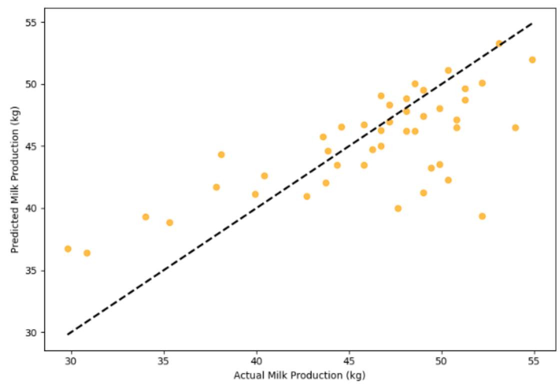


***Figure 11. Predicted vs. Actual Milk Production for XGBoost (Phase 2, Scenario 1*)**:** This scatter plot visualizes the predicted versus actual milk yield values for the behavior-only XGBoost model in Phase 2.



Discussion

This study presents a comprehensive computer vision-based framework for monitoring dairy cow behavior and predicting daily milk yield, addressing critical gaps in individual identification within multi-view environments. The findings underscore the potential of non-invasive monitoring systems for enhancing precision livestock farming, while also highlighting persistent challenges in achieving robust individual cow identification, particularly for specific behaviors.

The performance of our visual models demonstrated distinct outcomes across the single-view (Phase 1) and multi-view (Phase 2) setups. In Phase 1, analyzing approximately 550 images from a single camera, the system successfully identified approximately 5,500 cows with both behavior and individual ID. However, a significant limitation was observed, with an estimated 1,800 cows (~25% of identified cows) losing their individual identification, even though their behavior was classified. A critical observation was that about 70% of these unidentified cows exhibited 'lying' behavior (Behavior_6). This suggests that while the system effectively detected the lying behavior, individual identification failed more frequently in this posture, potentially due to occlusions, reduced visibility of distinctive body patterns, or insufficient camera angles. This issue is consistent with the known challenge of accurate and robust individual cow identification in complex, multi-view environments, especially when animals are partially occluded or in varying postures.

Phase 2, which utilized approximately 1,920 multi-view images with a 2x2 grid, significantly improved identification accuracy. Approximately 18,000 cows were correctly identified, and the number of lost individual identifications reduced substantially to about 1,800 cows (~9% of total identified cows). This reduction in identification loss from ~25% in Phase 1 to ~9% in Phase 2, primarily for lying cows, demonstrates the critical benefit of multi-view imaging and the implemented smart filtering process. Despite this improvement, a notable proportion (around 35% of the hypothetical total) of cows still lacked a complete record, indicating that challenges in individual tracking persist. The ResNet18 model for behavior classification achieved a high test set accuracy of 0.91, indicating strong performance in categorizing behaviors once a cow was detected.

Regarding milk yield prediction, the models evaluated in both phases revealed consistent differences in predictive performance based on the combination of input variables. Models using only behavioral data (Scenario 1) showed a moderate capacity to explain daily milk production. For instance, in Phase 1, the Random Forest model with only behaviors achieved an R^2 of 0.16 and an MSE of 33.66. In Phase 2, the behavior-only XGBoost model showed a more robust baseline with an R^2 of 0.48 and an MSE of 17.06. These results, while encouraging, underscore the need for further refinement in identification and recognition models, as only approximately 62% of hypothetical cows were fully identified with both behavior and ID in Phase 1 and approximately 58% in Phase 2.

The inclusion of the previous day's milk production (Scenario 2) consistently led to substantial improvements in prediction accuracy across all models and phases. In Phase 1, the Random Forest model with behaviors and previous day's milk yield achieved an R^2 of 0.78 and an MSE of 6.96, while XGBoost reached an R^2 of 0.73 and an MSE of 8.53. In Phase 2, Random Forest achieved an R^2 of 0.75 and an MSE of 7.70, and XGBoost reached an R^2 of 0.64 and an MSE of 12.73. This highlights the significant predictive power of historical milk yield data, effectively capturing individual cow trends and recent physiological conditions.

The addition of environmental variables (temperature, humidity, THI) in Scenario 3 resulted in more modest and sometimes marginal improvements. In Phase 1, the best performance was achieved by the Random Forest model in Scenario 3 (Behaviors + Milk Production Yesterday + Weather Conditions), with an MSE of 6.64 and an R^2 of 0.79. This indicates that the model explained 79% of the variance in daily milk production, with predictions deviating by approximately 2.57 kg ($\sqrt{6.64}$). While the THI showed some relevance in XGBoost models, this limited impact might be due to the potential redundancy between climatic effects and observed behaviors (as cows adjust their activities to thermal conditions) or the relatively low climatic variability during the study period.

One of the most consistent findings across all scenarios and models was the significant relevance of 'Behavior_6', corresponding to lying behavior. Although not always the most important feature, it consistently featured prominently in almost all analyses. This aligns with prior research emphasizing the importance of rest for dairy cows, with studies suggesting that cows should spend between 12 and 14 hours daily lying in cubicle beds for rest and rumination. More specifically, Jensen et al. (2005) suggest cows should spend between 50% and 60% of the day resting. Lying down is known to increase blood flow to the mammary gland by 25-50%, directly contributing to increased milk production. This behavior is not only an indicator of well-being but also a key physiological modulator of productivity. However, the impact of lying behavior on predictive models was less pronounced than expected, particularly in Phase 1, due to the high proportion (around 70%) of unidentified cows exhibiting this behavior. This technical challenge of associating the behavior with a specific individual limited its apparent predictive power.

Beyond 'Behavior_6' (Lyin), 'Behavior_5' (Drinking) and 'Behavior_3' (Feeding head down) also emerged as relevant variables. These findings are biologically sound given their direct link to a cow's physiological status and production potential. Adequate water intake is essential for milk synthesis, as milk is composed primarily of water, and increased feeding frequency has been positively correlated with improved milk yield. For example, 'Behavior_5' (Drinking) was one of the most important variables in Random Forest models in both Phase 1 and Phase 2. Its influence remained even when additional variables like 'milk_yesterday' and weather conditions were incorporated, suggesting it captures crucial, non-redundant behavioral information. 'Behavior_3' (Feeding head down), on the other hand, acquired greater weight in XGBoost models, especially in Phase 2. Since XGBoost has a greater capacity to detect non-linear interactions, this could indicate that 'feeding head down' interacts significantly with other factors to explain production. Increased feeding frequency and duration directly impact dry matter intake, which is a primary driver of milk production.

The substantial performance difference between Phase 1 and Phase 2 decisively demonstrates the importance of multi-view systems. The significant increase in correctly identified cows in Phase 2 not only improves predictive models but also enhances confidence in the system as a decision-making tool. This finding is crucial for the practical utility of vision-based monitoring systems in commercial farm settings, where reliable individual identification has been a major technical barrier.

This research highlights the feasibility of predicting daily milk yield based on behaviors inferred through multi-view computer vision, providing empirical evidence of how individual tracking accuracy affects prediction quality. It also underscores the limitations of current models in complex, real-world visual environments, particularly concerning individual identification when cows are in certain postures.

Conclusion

Our integrated deep learning-based approach establishes a robust and quantitative connection between animal behavior and economic productivity in dairy farming. By combining cow detection (YOLOv8), behavior classification (ResNet18), and individual identification (EfficientNet), and incorporating a smart filtering module for multi-view data, we demonstrated the potential for precise and non-invasive daily milk yield prediction. While the inclusion of previous day's milk yield significantly boosted model performance, and environmental variables offered subtle improvements, the study firmly establishes that quantifying behavioral patterns through computer vision can reliably predict milk production.

The consistent importance of Lying, Drinking and Feeding Head Down reinforces their biological significance for milk production and animal welfare. The substantial improvement in individual cow identification in the multi-view Phase 2, compared to the single-view Phase 1, is a key contribution, demonstrating that overcoming occlusion and posture-related identification challenges is critical for generating high-fidelity behavioral data and improving predictive models.

This work lays the groundwork for a new era of precision livestock farming, envisioning an intelligent, proactive, and sustainable future where animal welfare and profitability are synergistic outcomes of unified technological strategies. Continued refinement of identification and prediction models, including improvements in data quality, integration of complementary sensors, and development of more specialized network architectures, will be crucial for designing accessible, robust, and welfare-centered monitoring systems that are scalable for real-world commercial farm settings.

References

Arcidiacono, C., Porto, S. M. C., & Mancino, M. (2017). Development of a threshold-based classifier for real-time recognition of cow feeding and standing behavioural activities from accelerometer data. *Computers and Electronics in Agriculture*, *134*, 124-134.

Benaissa, S., Tuyttens, F. A. M., Plets, D., Cattrysse, H., Martens, L., Vandaele, L., Joseph, W., & Sonck, B. (2019). Classification of ingestive-related cow behaviours using RumiWatch halter and neck-mounted accelerometers. *Applied Animal Behaviour Science*, *211*, 9-16.

Chun-hua, S., Jian-jun, G., Xin-sheng, S., Xin-yu, Y., Juan-juan, Z., Chao, W., Qiu, D.-r., Wang, Y.-n., Feng, M., & Gao, Y.-h. (2019). Correlation between Behavior and Milk Yield of Dairy Cows. *International Journal of Agriculture & Biology*, *21*(1), 93-98.

Gao, Y. J., Yang, S. L., Li, Q., Gu, Z. B., & Xu, C. Y. (2014). Discussion on optimization of dairy house environment based on behavioral comfort-A case study of Israel model. *Chinese Dairy Cattle*, *19*, 32-36.

Grant, R. J. (2006). Incorporating dairy cow behavior into management tools. *Penn State Dairy Cattle Nutrition Workshop*, 31-41.

Janocha, A., Milczarek, A., Gajownik-Mućka, P., & Matusevičius, P. (2023). Evaluation of the efficiency of computer monitoring based on selected parameters of dairy herd behaviour in relation to daily milk yield. *Animal Science and Genetics*, *19*(3), 83-92.

Jensen, M. B., Munksgaard, L., Pedersen, L. J., & Ladewig, J. (2004). Prior deprivation and reward duration affect the demand function for rest in dairy heifers. *Applied Animal Behaviour Science*, *88*(1-2), 1-11.

Jensen, M. B., Pedersen, L. J., & Munksgaard, L. (2005). The effect of reward duration on demand functions for rest in dairy heifers and lying requirements as measured by demand functions. *Applied Animal Behaviour Science*, *90*(3-4), 207-217.

Jia, Q., Yang, J., Han, S., Du, Z., & Liu, J. (2024). CAMLLA-YOLOv8n: Cow Behavior Recognition Based on Improved YOLOv8n. *Animals*, *14*(20), 3033.

Munksgaard, L., & Lovendahl, P. (1993). Effects of social and physical stressors on growth hormone levels in dairy cows. *Canadian Journal of Animal Science*, *73*(4), 847-853.

Munksgaard, L., Jensen, M. B., Pedersen, L. J., Hansen, S. W., & Matthews, L. (2005). Quantifying behavioural priorities—effects of time constraints on behaviour of dairy cows, *Bos taurus*. *Applied Animal Behaviour Science*, *92*(1-2), 3-14.

Rodriguez-Venegas, R., Meza-Herrera, C. A., Robles-Trillo, P. A., Angel-Garcia, O., Legarreta-Gonzalez, M. A., Sánchez-Vocanegra, H. F., & Rodriguez-Martinez, R. (2023). Effect of THI on Milk Production, Percentage of Milking Cows, and Time Lying in Holstein Cows in Northern-Arid Mexico. *Animals*, *13*(17), 1715.

Rulquin, H., & Caudal, J. P. (1992). Effects of lying or standing on mammary blood flow and heart rate of dairy cows. *Annales de Zootechnie*, *41*(2), 101.

Tian, F., Wang, J., Xiong, B., Jiang, L., Song, Z., & Li, F. (2021). Real-Time Behavioral Recognition in Dairy Cows Based on Geomagnetism and Acceleration Information. *IEEE Access*, *9*, 109497-109509.

Vessies, P. J. A., Elshof-Baars, W., De Haer, L. C. M., & De Jong, G. (2014). First Heat Detection in Relation to Moment of First Insemination in Lactating Cows. *Interbull Bulletin*, *48*, 22-25.

Vu, H., Prabhune, O., Raskar, U., Panditharatne, D., Chung, H., Choi, C., & Kim, Y. (2024). MmCows: A multimodal dataset for dairy cattle monitoring. Advances in Neural Information Processing Systems (NeurIPS) Track on Datasets and Benchmarks.

Wang, J., Zhang, H., Ji, J., Zhao, K., & Liu, G. (2019). Development of a wireless measurement system for classifying cow behavior using accelerometer data and location data. *Applied Engineering in Agriculture*, *35*(2), 135-147.