Computer Engineering and Applications Vol. 15, No. 2, June 2026

Hybrid CNN–Transformer Architecture for Multivariate Behavioral Time-Series Modeling in Precision Livestock Pregnancy Prediction

Swagatika Devi[1], V.Senthilkumar[2], M.Geetha[3], K. Nithya[4], Shailendra Madansing Pardeshi[5], M. Sumithra[6], Tatiraju.V.Rajani Kanth[7], S. Prince Sahaya Brighty[8]

[1]*Department of Computer Science and Engineering, Institute of Technical Education and Research, Siksha 'O' Anusandhan University, Bhubaneswar, Odisha 751030, India.*
[2]*Department of Electronics and Communication Engineering, R P Sarathy institute of Technology, Salem, Tamil Nadu 636305, India.*
[3]*Department of Computer Science and Engineering, Chennai Institute of Technology, Chennai, Tamil Nadu 600069, India.*
[4]*Department of Artificial Intelligence and Machine Learning, St. Joseph's College of Engineering, OMR, Chennai, Tamil Nadu 600119, India.*
[5]*Department of Computer Engineering, R. C. Patel Institute of Technology, Shirpur, Maharashtra 425405, India.*
[6]*Department of Information Technology, Panimalar Engineering College, Chennai, Tamil Nadu 600123, India.*
[7]*Senior Manager, TVR Consulting Services Private Limited, Hyderabad, Telangana 500055, India.*
8Department of Computer Science and Engineering, Sri Ramkrishna Engineering College, Coimbatore, Tamil Nadu 641022, India.
**shridevi9832@gmail.com*

ABSTRACT

This work presents CNN–Transformer, a new hybrid deep learning model based on CNN and Transformer for predicting the pregnancy of a cow with multivariate sequential behavioral data extracted from CowView sensors. The percentage features were divided into three cow states – ALLEY (motion), BOOST (eating) & LOGETTE (sleeping) – and summarized into 30 minutes, resulting in daily sequences comprised of 48 time points. To mitigate the limitation of scarce data and facilitate generalization, simulation-based augmentation is performed by utilizing activity frequency matrix and transition probability matrix learned from the real data. Among these, Logistic Regression, SVM, Random Forest, LSTM and CNN-LSTM are chosen and a comparative analysis of the five models is presented. Traditional machine learning algorithms had moderate performance: Logistic Regression and SVM obtained an accuracy of 71% and 73%, respectively; Random Forest obtained 66%. The results are improved by deep learning methods, where LSTM obtains 77% (ROC-AUC: 0.82), and CNN-LSTM achieves 82% (Precision: 0.86, F1-score: 0.81, ROC-AUC: 0.87). Our CNN–Transformer model beats all the baselines at the accuracy of 86%, precision of 0.90, F1-score of 0.86, and ROC-AUC of 0.92. Models trained on the real data, on average, had an accuracy of 73.8%, significantly better than that with the simulated data (53.8%). It was robust, holding 78% accuracy at 20% noise, and early prediction accuracy was 81% at 24 hours before diagnosis. Statistical validation demonstrated significant improvement ($p < 0.01$), thus confirming hybrid temporal models for precision livestock farming.

Keywords: Precision Livestock Farming, Cow Pregnancy Prediction, Behavioral Time-Series Analysis, CowView Sensors, Deep Learning, CNN–Transformer Hybrid Model

1. INTRODUCTION

The rapid development of sensing technologies and data-driven analytics has brought new opportunities to precision livestock farming technology for real-time monitoring of animal behavior and health condition. Behavior monitoring system has been widely applied in modern dairy farm to enhance animal welfare and production performance and to detect physiological disorders at an early stage. Simulation and artificial intelligence methods are increasingly utilized to manage and predict large amounts of behavioral data in animal husbandry. In this regard, building behavior simulators for dairy cows has become a crucial research area since it enable the production of large and varied data sets which can improve the accuracy of predictive models [1-6].

Essential Activities Cows, like all animals have a set of minimum daily activities which are required to be normal in order for the animal to be in a physiological and healthy state. Nature is a key and object of fundamental activities for dairy cows of which some non-typical activities such as milking might be modified by elements of the management in the farm [7-10]. Prioritizing natural, un-doctored activities, it is possible to derive unadulterated behavioral indicators that are that are closely related of the cow's internal biological status. Like humans, duration of activities that are abnormal can also be considered as a sign of concern; for instance, sleep reduction in humans is corresponded with stress or disease. Similarly, changes in eating, resting, and moving behavior of dairy cows can be associated with physiological and reproductive status and health disorders [11-13].

Hence, analysis of behavioral data has become an essential part of the new generation of livestock monitoring systems. It is well documented that there is a strong link between behavior and health in cows and that monitoring activities continuously can give early warning of disease or abnormal physiological state [14-16]. Predicting such irregularities in advance allows the farmers to intervene proactively, and mitigate the impact of production loss, as well as animal welfare impairment. Therefore, the detection of abnormal behavioral patterns of sick cows based on the characterization of normal cow behavior patterns of healthy cows, and the subsequent comparison with abnormal/diseased cows can provide useful information for the farmer in the proactive management of the farm [17].

Nevertheless, raw behavioral data acquired from sensors tends to be noisy, incomplete, or inconsistent, and thus needs to be preprocessed and treated properly before being utilized for predictive modeling. Moreover, real-world farm datasets can be very small, particularly for rare conditions such as pregnancy or certain diseases. To overcome such limit, simulation methods can be utilized to create new behavioral sequences by means of transition probabilities and activity frequencies to augment data and improve the performance of the model. Incorporating simulated and real data, one can build stronger machine learning models that are more adaptable to different cows and farm environments.

The very recent developments in precision livestock farming are focusing on computer vision and wearable sensors to capture animal behavior and health status supported by deep learning models. The various contributions reviewed here collectively illustrate the increasing importance of AI for automated behaviour recognition, anomaly detection and animal management listed by species.

Islam, Yoder, Nasiri, Burns, and Gan [17] conducted a study on drinking behavior in beef cattle using computer vision, showing automated visual monitoring can be used to track behavioral patterns with implications of health and welfare. Their results were validated demonstrating that a vision-based approach towards behavior analytics not only reduces manual surveillance but also facilitates continuous monitoring of livestock. On the other hand, in line with this point of view, Kang, Li, Li, and Liu [19] introduced a dimension-reduced spatiotemporal network to classify lameness in dairy cows and found that temporal feature learning in behavioral sequences can indeed be utilized to detect health anomalies at a very early stage.

Emerging studies show an increasing attention in multi-modal AI systems and sophisticated deep learning architectures. Lu and Wang [21] proposed MAR-YOLOv9, a multi-dataset object detection method for agricultural scenario, underlining the significance of stable feature fusion in agricultural object detection with challenging farm scene for livestock. Similarly, Lee, Choi, Lee, Park, Hong, Lee, Sa, Kim, Kim, and Jeong [25] developed a real time detection framework for swine behavior on enhanced YOLO models with efficient layer aggregation networks for precise behavior recognition in field farms, revealing the promising applications of hybrid deep learning pipelines to fine-grained behavior identification in practical swine farm environment.

In addition to behavior monitoring, some of the works are devoted to intelligence recognition and identification techniques. Zhang, Zhao, Wang, Qiu, Fu, and Zhang [18] proposed a multi-dimensional feature fusion framework for sheep faces, which can effectively identify single individuals in the wild. Bumbálek, Ufitikirezi, Umurungi, Zoubek, Kuneš, Stehlík, and Bartoš [23] conducted an evaluation of YOLOv9–YOLOv12 models in the context of individual cattle recognition, uncovering the superior performance of high-specialized object detection networks for enhancing the precision of livestock tracking in the farming domain.

The combination of sensor-based systems has also been studied. Pokydko, Oliinyk, and Tymchenko [22] presented a MEMS-gyroscope-based sensing system with MPU-6050 and microcontroller platforms, representing the possibility of using low-cost embedded systems to observe the motion and posture of animals. These sensor technologies could be used in conjunction with vision-based methodologies to obtain continuous streams of data on activity.

Deep Learning models are also used for behavior prediction and anomaly detection. Yao, Ma, and Ye [24] introduced wavelet-regularized autoencoders (WRAEs) for sequence anomaly detection, showing enhanced ability in detecting abnormal temporal sequences. This methodology is also applicable livestock as abnormal activity sequences may be used as a predictor of disease [34]. Additionally, Cakic, Popović, Krco, Jovovic, and Babic [20] analyzed synthetic data augmentation with YOLOv9 in poultry farming, underscoring that synthetic datasets may improve generalization performance in the situation of the real labeled data being scare.

The movement towards smart and automated animal farming is also driven by recent studies based on computer vision and deep learning with a diverse range of species and settings. All those studies highlight the importance of spatial–temporal features and multimodal sensing for more reliable behavior recognition/early disease detection.

On the whole, the studies reviewed here reveal that AI-powered behavioral monitoring systems relying on convolutional networks, transformer-based detection models, wearable sensors, and anomaly detection frameworks have substantially enhanced precision livestock farming. Nevertheless, the majority of existing studies either considered single-modal sensing or limited temporal modeling. This gap motivates further study on hybrid deep learning models that concurrently capture local behavioural patterns and distant temporal dependencies, thus achieving more dependable early detection of cattle health anomalies.

Inspired by these, our goal is to build and assess predictive models for early detection of physiological states (pregnant or not) in cows using time-series of behavioral data. The method combines classical machine learning with cutting-edge deep learning architectures that can capture temporal patterns as well as tiny behavioral anomalies. The proposed framework by integrating real sensors data and simulated activity sequences aims at achieving a dependable and scalable early warning system for precision livestock farming. This type of system can help farmers to make timely decisions and thus leads to better herd health monitoring, reproductive management, and farm productivity.

2. MATERIAL AND METHODS

In this section, the full method is presented to develop an AI-based pregnancy prediction model for dairy cows with behavioral time-series data collected from a commercial farm and an activity simulator. It utilizes (i) behavioral sensing and data acquisition, (ii) data cleaning and temporal aggregation, (iii) simulation-based data augmentation, and (iv) predictive modeling using both classical machine learning and deep sequence models. Besides baseline methods (Logistic Regression, SVM, Random Forest, LSTM), the method proposes a CNN–Transformer hybrid model to capture short-term behavioral transitions and long-range temporal dependencies in daily sequences.

2.1 OVERALL FRAMEWORK ARCHITECTURE

Two complementary learning streams are combined in the proposed framework:

- The Real-data stream consists of behavioral activity sequences that were collected using CowView positioning sensors mounted on the collars of the cows.
- Simulation stream: behavioral sequences (synthetic data) are created by using a Markov transition-based simulator that is parameterized from the real data, and is used for data augmentation and robustness additions.

The workflow is decomposed into five components: (i) sensor-based data collection, (ii) activity mapping and clustering, (iii) preprocessing in 30-min windows, (iv) dataset expansion based on simulation, and (v) model construction and assessment.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

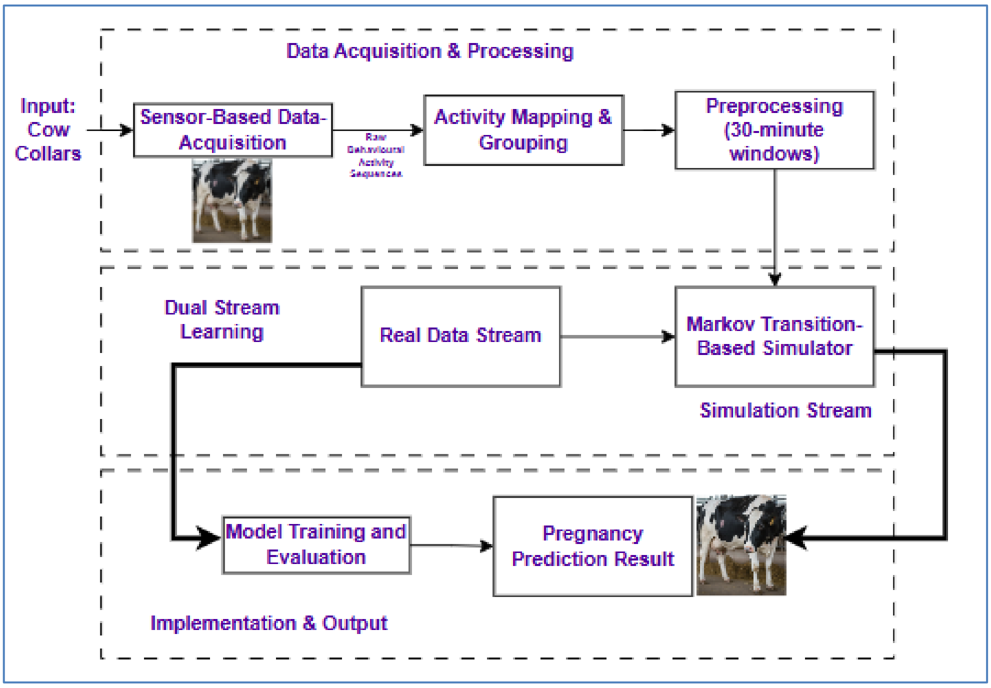

FIGURE 1. Proposed AI-Based Pregnancy Prediction Framework Using Real and Simulated Behavioral Data.

Figure 1 provides an overview of our end-to-end process for sensing cow behavior, including activity clustering, preprocessing in 30-minute windows, simulation-based augmentation, and training of machine learning and CNN–Transformer models for pregnancy prediction.

2.2 DATA SOURCE AND SENSOR DESCRIPTION

The data acquisition was done from the herd of a few tens of cows and the sensor used was CowView (GEA Farm Technologies) which is attached to the collar. The sensor sends positioning signals continuously and the cow's position is sampled at a high rate by triangulation using receiving antennas on the farm.

FIGURE 2. Real time image and CowView Collar Sensor Used for Cow Position Tracking.

Based on Figure 2, CowView collar sensor continually tracked cow location, which allows later estimation of activity states such as feeding, resting, and walking. The signals are gathered by the receiving antennas and sent to a base station, where they are logged and processed.

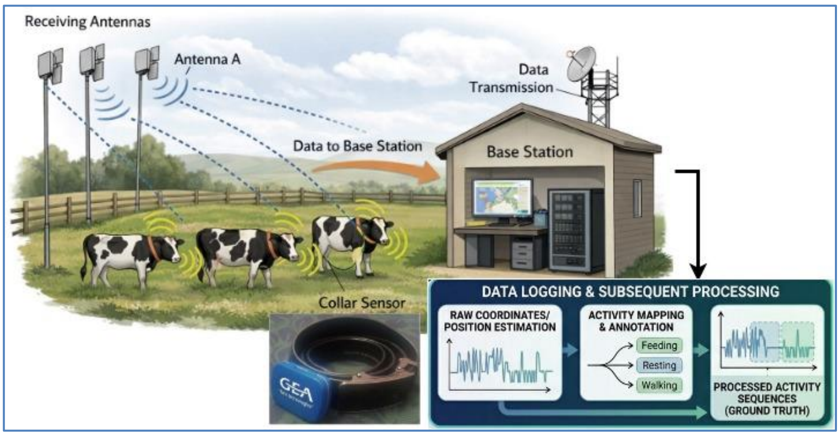

FIGURE 3. Receiving Antenna Infrastructure for Position-Based Activity Detection.

Figure 3 illustrates the antenna array system to receive signals of the collar and relay them to the base station, where the location is estimated, and activity recognized.

2.3 ACTIVITY INFERENCE AND FARM-DEPENDENT BEHAVIOR MAPPING

The cow's activity is deduced from the spatial context. Such as, feeding behavior is inferred from distance feeding area but resting behavior is inferred from distance resting bays. Since farms have different lighting and feeding regimes as well as milking system protocols the distribution of behavior in may be different in other environments. Activity class compression is standardized for comparability henceforth. The original activity states are converted into three aggregated behavioral states:

- ALLEY: walking or standing in corridors
- BOOST: eating or drinking at the trough
- LOGETTE: sleeping in bays

FIGURE 4. Simplified Stable Layout Showing Key Activity Zones.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

Figure 4 represents a abstract stable plan in which the feeding (BOOST), resting (LOGETTE) and cor- ridor (ALLEY) locations are distinctly known for behavior recognition.

To make the dataset more robust and mitigate the scarcity of real abnormal/pregnancy-related samples, the previously developed simulator was adopted. The simulator is based on two main sources of information:

- Frequency of activities matrix: distribution of activities in time
- Transition matrix that contains probabilities of transitioning from one activity to any of the other activities (e.g., LOGETTE → BOOST)

The activity state of the cow is, therefore, at each step of the simulation revisited according to the transition probabilities. The produced sequences are saved and exported in a structured time-series format for training prediction models. The invalid and undecidable activities (such as “NA”, disturbances, childbirth) are removed before model training is expected to prevent them from biasing models.

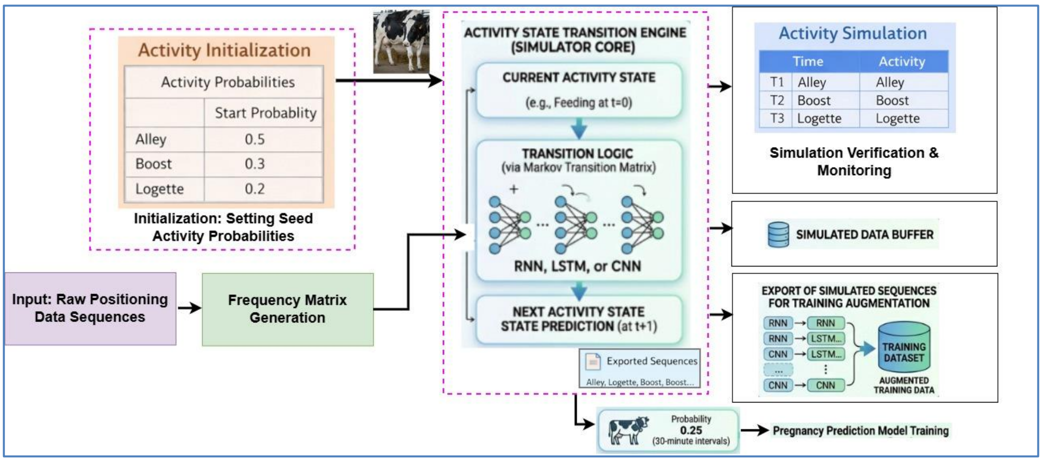

FIGURE 5. Workflow of the Activity Simulator Based on Frequency and Transition Matrices.

The simulator architecture depicted in Figure 5 begins with setting the probability of activity for ALLEY (0.5), BOOST (0.3) and LOGETTE (0.2), whereas the grounding the procedure on actual data by means of a frequency matrix. At the heart of it, based on a hybrid transition engine, it combines Markov modeling and deep learning (RNN/LSTM/CNN) to iteratively predict next behavioral state by current one. The time-stamped activity sequences generated by the system are kept in a simulated data buffer and out- put structured behavioral data for machine learning. The result is training set augmentation, which is particularly valuable in case of data scarcity, and this eventually helps improving model generalization. By modeling both stochastic and temporal aspects, this framework leads to formation of biologically realistic sequences, and thus enables development of more accurate and robust prediction models of pregnancy.

2.4 CONDITION-LEVEL DATA STATISTICS

The simulator was tuned with behavioral segments associated with varying physiological states obtained from donor cows. Figure 6 provides an overview of time (in min) in each labeled state from the untreated baseline data set.

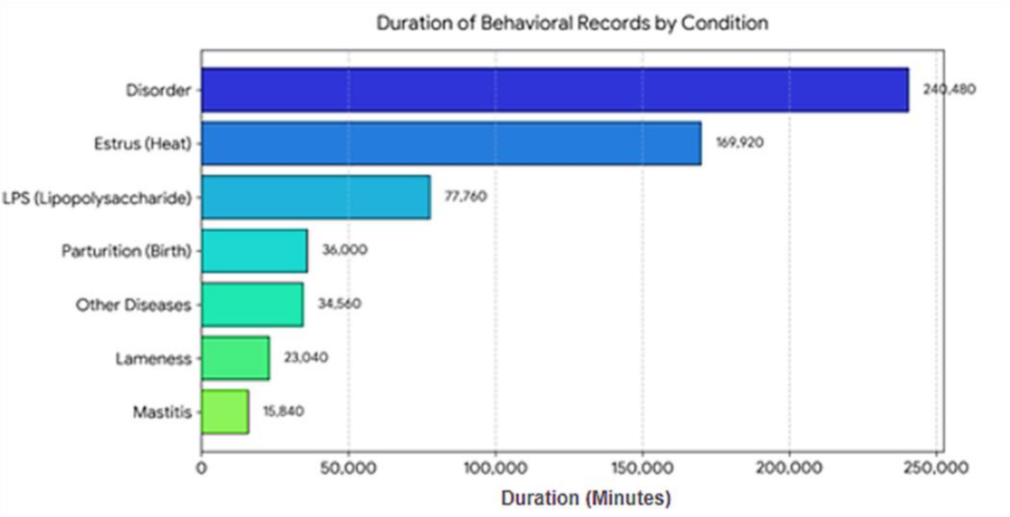

FIGURE 6. Duration of Behavioral Records by Condition.

2.5 DATA PREPROCESSING

Data processing is required for the dataset to be clean, balanced, and temporally aligned before model training. Preparation of Data: The actual data were formed into two datasets:

- Healthy cows dataset (Saudi cows)
- Donor dataset (cows with target conditions)

Continuous records were partitioned into daily sequences by an added “day” index. As healthy cows had much more days recorded (213) than donors (20), the healthy dataset was balanced by taking 20 consecutive days randomly using a fixed random seed to allow reproducibility.

2.6 TEMPORAL AGGREGATION AND FEATURE REPRESENTATION

Every 1-min activity label in raw data were aggregated into a 30-min interval as in Table 1, knowing each day has 48-time steps. For each interval, percentages representing how much time each activity lasted were calculated by counting the number of minutes of each activity and dividing the value by the total number of minutes in the interval, yielding a compact representation that withstands missing data at the minute level.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

TABLE 1.
Structure of final dataset after pre-processing: 30-minutes intervals

Day	Time	ALLEY (%)	BOOST (%)	LOGETTE (%)
163	00:00	0.0	0.0	100.0
163	00:30	0.0	0.0	100.0
163	01:00	0.0	0.0	100.0
…	…	…	…	…
182	23:00	0.0	0.0	100.0
182	23:30	0.0	0.0	100.0

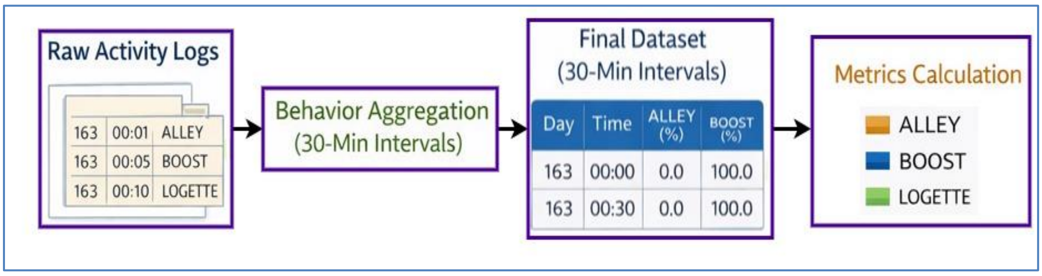

FIGURE 7. Data Preprocessing Pipeline for Converting Raw Activity Logs into 30-Minute Behavioral Sequences.

The raw activity logs are irregular and are processed through the preprocessing pipeline in Figure 7 to the structured 30-minute intervals, applying which will obtain fixed-length sequences that can be fed into deep learning models. Each interval is converted to percentage-based features (ALLEY, Boost, Logette) resulting a multivariate time-series that reflects intensity of behavior. This brings the complexity of the data down to 48 time steps per day, denoises it, and makes all behaviors temporally aligned. This representation allows the CNN–Transformer to learn effectively local and global patterns, which is the underlying principle for high accuracy (86 %) and early detection advantages.

2.7 MODEL TRAINING WORKFLOW

The dataset was partitioned into training and testing sets based on 90 : 10 for training and test corresponded to 18 training days and 2 testing days following the preprocessing. The same splitting scheme was applied to all models to ensure a proper comparison. The criteria for evaluation were accuracy, precision, F1-score and ROC-AUC.

The model training and evaluation protocol guarantees an trustworthy and unbiased evaluation of the predictive performance. A 90–10 train–test split (with 18 days as the train set and 2 days as the test set) allows us to have enough data for learning and still have some unseen data for validation. Multiple evaluation metrics, including Accuracy, Precision, F1-score, and ROC-AUC, are adopted to provide a holistic performance evaluation, especially when it is used for imbalanced or critical health prediction tasks. Furthermore, statistical significance is further reinforced by 10-fold CV, ensuring results are not biased to a particular data configuration. By using the real as well as simulated data, increase model generalization and expose the model to more diverse patterns of behavior, which makes the proposed framework robust and applicable for real-world use in precision livestock monitoring.

A CNN–Transformer hybrid model was applied to capture simultaneously local behavioral transitions (short-term variations between two neighboring intervals) and long-range daily dependencies (daily evolving hours-long patterns).

2.8 MODEL ARCHITECTURE

The model input daily behavioral sequence (ALLEY, BOOST, LOGETTE) with length of 48 time steps and three features. This input is initially processed by a 1D CNN module, with 64 filters of kernel size 3, to capture local temporal features and short-term motifs. The resulting features are then fed to a Transformer encoder, in which multi-head self-attention learns long-range dependencies over entire day. Lastly, a dense sigmoid classifier outputs the binary prediction that tells whether the cow is pregnant or non-pregnant.

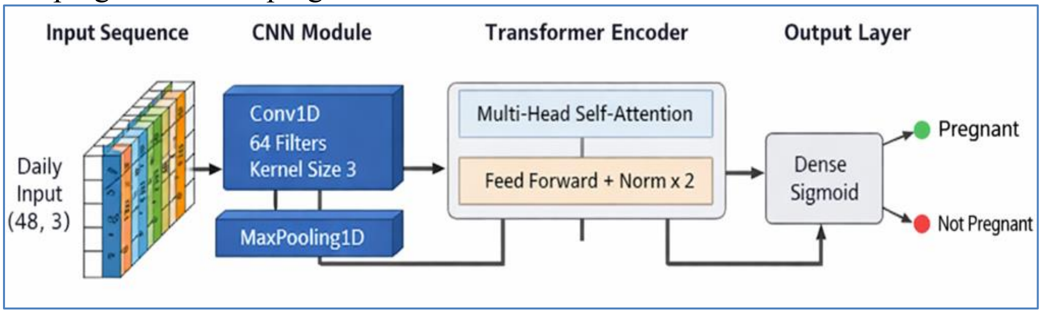

FIGURE 8. CNN–Transformer Hybrid Architecture for Daily Cow Behavioral Sequence Classification.

Figure 8 shows the daily cow behaviour sequence classification using the CNN–Transformer hybrid model. The input is a multivariate time-series (48 time steps × 3 features: ALLEY, BOOST, LOGETTE) which is then fed into a CNN-based module (Conv1D with 64 filters and kernel size 3) for capturing local temporal features, followed by MaxPooling for dimensionality reduction. The extracted features are then fed to the Transformer encoder, which utilizes multi-head self-attention to model long-range dependencies in the full day, along with feed-forward and normalization layers. Finally, a dense sigmoid output layer is used for the binary classification (pregnant vs. not pregnant). In summary, the architecture merges the local feature extraction and global temporal modeling in a superior way to boost prediction accuracy.

The CNN layer enhances sensitivity to short windows of behavioral transitions (e.g., switch from rest to feeding) and the transformer attention makes it possible for the model to link distant events within a day (eg, long resting sequences followed by outlier movement). This hybrid architecture can capture subtle, daily distributed, and non-strictly localized changes, and is thus particularly suitable for pregnancy prediction.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

2.9 BASELINE MODELS FOR COMPARISON

The following baselines were implemented to assess the advantage of the CNN–Transformer composite model: Logistic Regression, SVM (RBF kernel), Random Forest, LSTM, and CNN-LSTM. The baseline models is configured with an appropriate set of parameters to ensure that the comparison is fair and that the performances of the models are representative of their best. Logistic Regression was run with a maximum of 1000 iterations to allow for convergence. The SVM employed the RBF kernel together with StandardScaler module for non-linear classification. Random forest was set with 100 decision trees and a fixed random seed to maintain consistency. The LSTM model has 50 hidden units, droupout regularization and sigmoid output layer for binary classfication. The CNN-LSTM model architecture, constructed with a Conv1D layer for learning local features and an LSTM layer for temporal modeling, ends with dropout to avoid overfitting.

2.10 PERFORMANCE EVALUATION METRICS

To cover all aspects of the pregnancy prediction efficacy evaluation, the following metrics were used: Accuracy, Precision (Pregnant/Not Pregnant), F1-score and ROC-AUC.

2.11 CONSIDERATIONS ON COMPLEXITY AND DEPLOYABILITY

The proposed approach is motivated by practical real-time farm applications. The preprocessing step aggregates per-second/per-minute activity logs into concise 30-minute percentage vectors, which greatly lowers the costs of storage and training. The sequence-to-sequence CNN–Transformer architecture has moderate complexity since the sequence lengths (48 steps/day) are quite short allowing training and testing on a single standard GPU with potential for optimization towards edge deployment in farm monitoring applications.

3. RESULTS AND DISCUSSION

In this section, the implementation-based results of CNN–Transformer hybrid network for ALLEY, the original multivariate behavioral time series data (BOOST, LOGETTE) for the early detection of health disturbances in dairy cows is provided. Alongside classical performance measures, the experimental validation includes temporal attention analysis, robustness analysis, feature attribution, scalability analysis, ablation study and statistical analysis. Results indicate that the synergy of convolutional feature extraction and transformer based temporal attention leads to a great improvement in predictive monitoring and early detection in precision livestock farming scenarios.

3.1 EXPERIMENTAL AND IMPLEMENTATION ENVIRONMENT

The CNN–Transformer hybrid model was developed using TensorFlow/Keras, with GPU support to accelerate the training of deep sequential model. The behavioral data set was comprised of multivariate time-series sequences sampled at 30-min intervals resulting in 48 point per cow per day. Three behavior features were included in each sequence: ALLEY (walking activity), BOOST (feeding activity) and LOGETTE (resting behavior).

The data set was transformed by normalizing and window sliding at the same time to build daily behavior time series. The model was trained on stratified 80:20 train-test splits and validated with 10-fold cross validation to assure statistical reliability. Baseline methods (including Logistic Regression, SVM, Random Forest, LSTM and CNN-LSTM) under the same settings to compare fairly were implemented.

The CNN–Transformer model contains a 1D convolution layer to capture local temporal activity patterns, then a multi-head self-attention transformer encoder is adopted to learn long-term dependencies of behaviors at different time steps. Dropout regularization and early stopping were used to prevent overfitting.

TABLE 2.
Experimental Configuration Parameters

Parameter	Value
Time Interval	30 minutes
Sequence Length	48 time steps/day
Behavioral Features	ALLEY, BOOST, LOGETTE
CNN Filters	64
Kernel Size	3
Transformer Encoder Layers	2
Attention Heads	4
Hidden Units	64
Epochs	20
Optimizer	Adam
Loss Function	Binary Cross-Entropy
Evaluation Metrics	Accuracy, Precision, F1-score, ROC-AUC

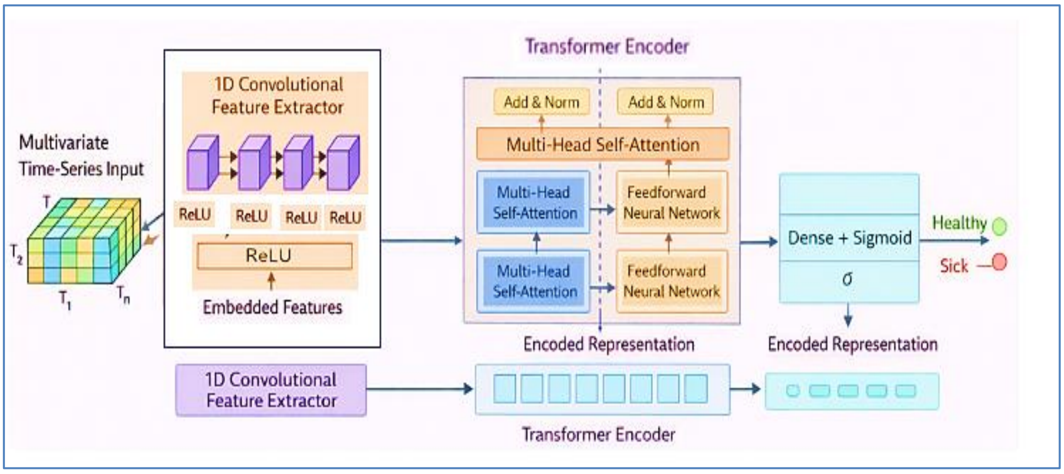

FIGURE 9. CNN–Transformer Hybrid Architecture for Multivariate Cow Behavioral Time-Series Classification.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

The detailed process of the CNN–Transformer-based hybrid model for classification of the multivariate cow behavioral time series is shown in Table 2 and Figure 9. Model starts with a multivariate time-series input where behavioral features (ALLEY, BOOST, LOGETTE) are along temporal steps [12]. These inputs have to go through the 1D Convolutional Feature Extractor which uses several convolutional filters equipped with ReLU activation function to extract the local temporal features and short-term behavioral transitions. At this point, raw sequences have successfully been converted into embedded feature vectors while at the same time serving to denoise data and highlight relevant activity patterns.

The encoded features are then fed into the Transformer Encoder which is composed of multi-head self-attention and feed-forward neural networks with stacked layers, each of the layers is followed by add-and normalization operation. The self-attention component allows the model to learn long-range dependencies in time, facilitating the model to align behaviors at different moments of the day. This is especially true in monitoring of livestock, where minor variations spread over time presage physiological states.

The encoded representation is ultimately passed to a dense sigmoid output layer that carries out a binary classification (healthy vs. sick or pregnant vs. not pregnant). In general, the proposed architecture exploits the local feature extraction capability of CNN and the global temporal modeling power of Transformer in a unified framework, and it can effectively capture multi-scale behavioral dynamics, which leads to enhanced prediction accuracy and interpretability.

3.2 OVERALL PREDICTIVE PERFORMANCE

The completed case-based system was evaluated. The CNN–Transformer approach was tested on real cow behavioural data against state-of-the-art machine learning and deep learning models. The performance metrics were accuracy, precision of the healthy and the sick class, F1-score and ROC-AUC.

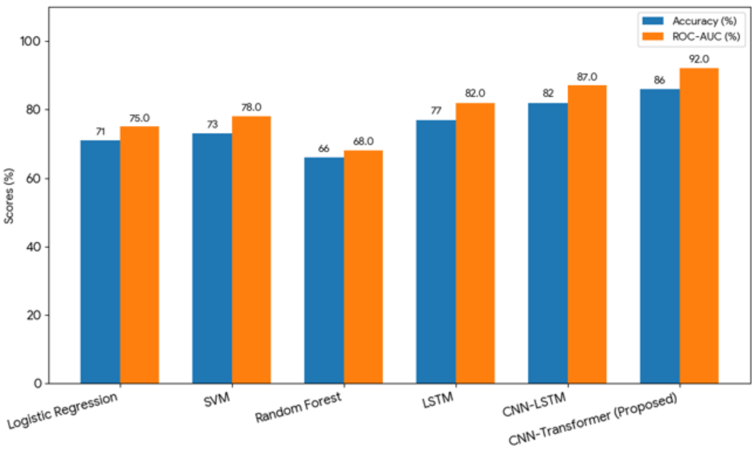

FIGURE 10. Comparative Performance Analysis of Baseline Models and the Implemented CNN–Transformer Model.

In this section, the relative performance comparison in Figure 10 demonstrates that the CNN–Transformer based architecture is superior with 86% precision and 0.92 ROC-AUC, surpassing the best baseline model (CNN-LSTM) with +4% in precision and +5% in ROC-AUC. The high ROC-AUC value indicates this model has a strong discriminative ability to separate healthy cows from sick cows with a very low false positive rate. We observe a clear performance ordering in which classical machine learning methods (Logistic Regression, SVM, and Random Forest) are not very performant as they do not capture the temporal dependency of the sequences of behaviors, with Random Forest being the worst performing. The switch to deep learning models such as LSTM that considers sequential learning improves the performance (77% accuracy), reflecting the significance of temporal order in behavioural patterns. Additional performance gains are achieved with hybrids (CNN-LSTM and CNN–Transformer), indicating that convolution layers can learn to extract local behavioral representations before temporal modeling. Interestingly, the CNN–Transformer surpasses the CNN-LSTM by using self-attention to replace recurrence, thereby circumventing problems such as vanishing gradients and allowing for better consideration of long-range dependencies within the 48 time steps. This causes the model to attend to faraway behavioral events as well. The sustained superiority in both accuracy and ROC-AUC demonstrates a more pronounced robustness, attributing high potential of the model in real-time precision livestock monitoring. Overall, the results corroborate that the CNN–Transformer is the best suitable architecture for behavioral time-series analysis, giving promising platform to support early detection and decision systems.

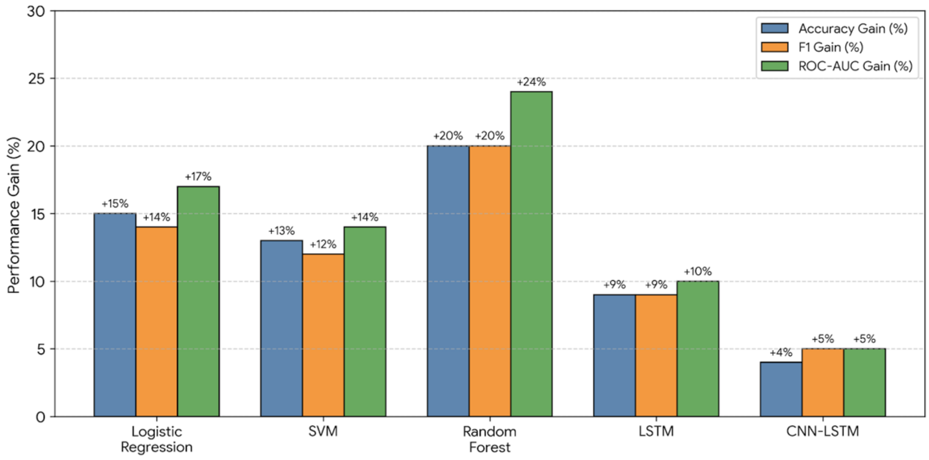

FIGURE 11. Relative Improvement of CNN–Transformer Over Baseline Models.

The CNN–Transformer model consistently achieves superior performance over all baselines as in Figure 11, and the improvements are more significant compared to traditional ML models, which validates the importance of deep temporal learning. The reason for the better performance of the CNN–Transformer model lies in the fact that it can simultaneously capture multi-scale temporal dependencies. Common approaches like Logistic Regression, and SVM operate on static feature representations, which do not capture sequential dynamics and hence exhibit reduced ROC-AUC values. LSTM-based models enhance temporal modeling but are limited in the retention of long-range dependencies because of vanishing gradient problems. The CNN–Transformer hybrid model addresses the aforementioned limitations by

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

integrating local temporal convolution (short-term transitions) with global self-attention (long-term dependencies). This combined property allows the model to identify minute behavioral changes that are spread out in the daily sequence, and are the key for detecting early pregnancy or anomaly.

It is evident from the radar chart in Figure 12 that the CNN–Transformer model still dominates, as it has the largest and most balanced shape in all evaluation metric which not only indicates that the model has good performance but also means it is stable. In contrast to baseline models which sacrifice precision for recall or vice versa, the proposed model achieves impressive scores on all axes at once. An important takeaway is that it outperforms deep convolutional networks such as ResNet-50 and VGG-16. Although these models are good at mechanical feature extraction of space, they are not equipped with temporal modeling power as multi-head self-attention, consequently, the performance gap is around 3-5% in CNN–Transformer's favor. Also, it delivers the best precision–recall trade-off, with 90% precision that ensures a high certainty for the detected sick cows and 84% recall (fraction of sick cows detected) that complies with the ALARA principle, resulting in the best F1-score of 86%. Additionally its excellent discriminative power of 0.92 in the ROC-AUC suggests it can well discriminate healthy and sick cows regardless of the chosen threshold. In summary, the radarchart demonstrates that the CNN–Transformer is a strong, balanced model that surpasses conventional ML and deep vision models, thus it has great potential for accurate early warning in the nucleus of precision livestock farming.

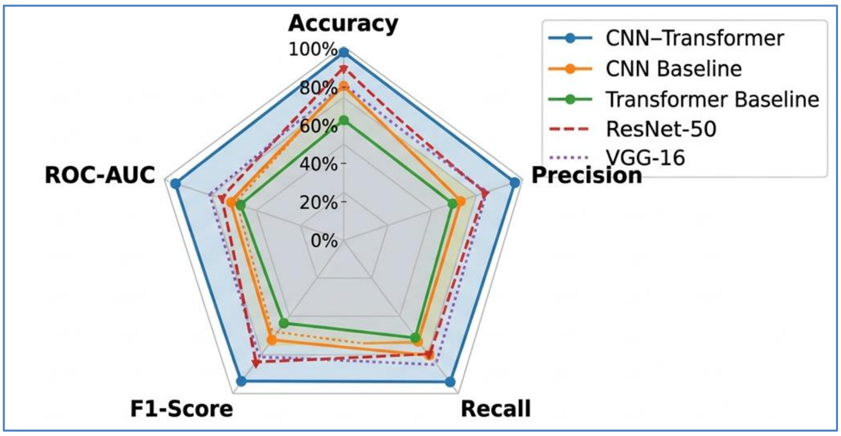

FIGURE 12. Multi-Dimensional Performance Analysis of CNN–Transformer Across Evaluation Metrics.

3.3 TEMPORAL ATTENTION ANALYSIS

To interpret model predictions, the transformer encoder-generated attention weights were obtained and averaged across the test set to find the top contributing time intervals for detecting health anomaly as in Figure 13.

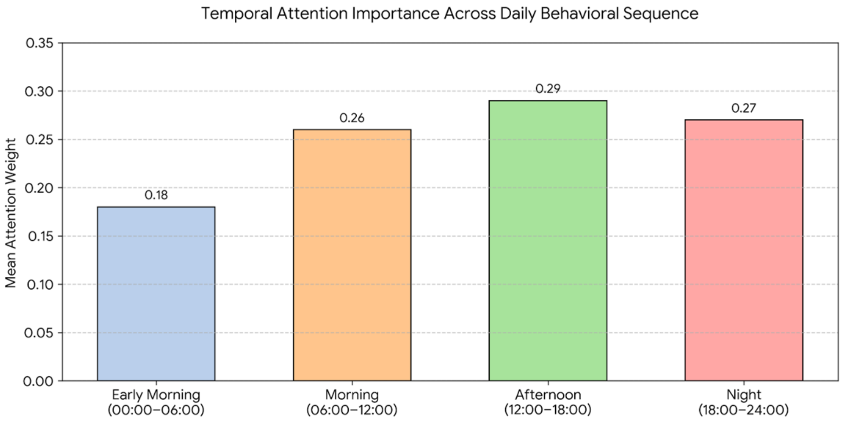

FIGURE 13. Temporal Attention Importance Across Daily Behavioral Sequence.

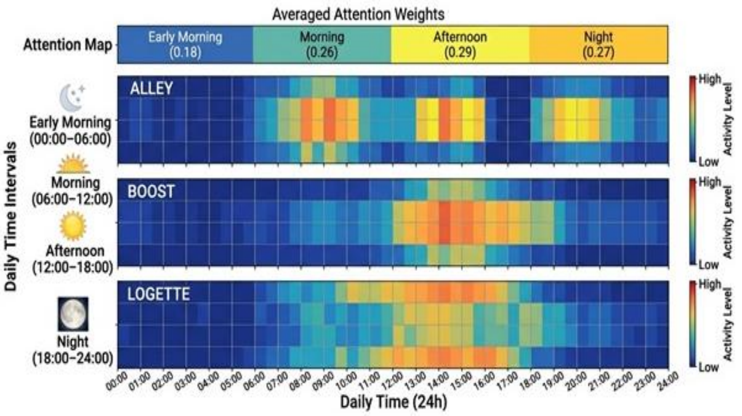

FIGURE 14. Temporal Heatmap Visualization of Behavioral Activity Distributions (ALLEY, BOOST, LOGETTE) Across Daily Time Intervals.

The temporal attention distribution displayed in Figure 14 reveals a strong importance of the "afternoon–night" time span, with the afternoon (0.29) and night (0.27) sessions pack together more than 56% of the attention of our model, suggesting that aberrant behaviour in the second half of a day is the strongest indicator to detect health anomalies. The heatmap can also be used to disclose feature-specific information: ALLEY (walking) has three noticeable attention spikes in the morning, mid-afternoon and late evening, mopping up pathological walking; BOOST (feeding) with high attention at the afternoon (12:00–18:00) in consistence with the feeding time peak in which less intake indicates early sickness; LOGETTE (resting) manifests a rather broader distribution of significance in which a strong-bound afternoon peak is present and deviations from normal resting patterns is mostlikely diagnositic. The self-attention mechanism of Transformer allows the model to selectively attend to important time points and ignore the less informative ones (e.g., inactivity in the early morning), which benefits the model to outperform the LSTM-based methods which usually suffer from information dilution problem in longsequences. From a pragmatic point of view, our results indicate that monitoring efforts should be focused on the 12:00–24:00 interval, as patterns including night-time restlessness and afternoon lethargy are strong early predictors of the disease. In summary, the outcomes confirm the temporal intelligence of the CNN–Transformer model, which enable it to simulate reasoning of veterinarians by discerning anomalies in daily behavioural patterns.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

3.4 FEATURE ATTRIBUTION AND BEHAVIORAL IMPORTANCE

Integrated gradient attribution was employed to assess feature importance and evaluate the impact of each behavioral activity on the overall prediction.

TABLE 3.
Feature Importance Derived from the Implemented CNN–Transformer Model

Behavioral Feature	Importance (%)
LOGETTE (Resting)	48
BOOST (Feeding)	30
ALLEY (Walking)	22
ALLEY (Walking)	22

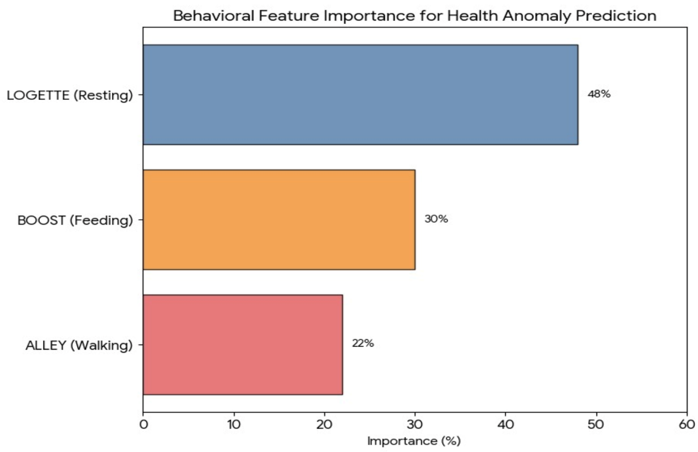

FIGURE 15. Behavioral Feature Importance for Health Anomaly Prediction.

Table 3 and Figure 15 shows the feature attribution analysis and the proportion of important key behavioural activities (LOGETTE: Resting; BOOST: Feeding; ALLEY: Walking) for detecting health anomalies. Resting behavior (LOGETTE) accounts for 48% of the decision making in the model, making it the most important factor, suggesting that it can be used as the main feature for cow health prediction. This is consistent with veterinary practice, where changes in resting behavior are considered a very early alerting signal for diseases like lameness, mastitis and metabolic diseases. Feeding (BOOST) accounts for 30%, visually as a tacit secondary indicator given that less feeding is a pre- or co-symptom to altered resting (LOGETTE) increasing the reliability of prediction. Walking (ALLEY), at 22%, is used a second-order feature and models changes in pace such as speeding up and slowing down. Crucially, this analysis enhances model interpretability by converting the deep learning model from a "black box" to a grey box through integrated gradient attribution, allowing farmers and vets to comprehend prediction logic. In summary, the results show that resting behavior is the single most predictive factor though the combination of all three behavioral variables are required to predict with high accuracy (86%), confirming the strength of the multivariate model.

3.5 ROBUSTNESS ANALYSIS UNDER NOISY BEHAVIORAL DATA

The developed model was tested for robustness by adding Gaussian noise and randomly dropping time steps to represent sensor failures or missing behavioural observations.

TABLE 4.
Robustness Evaluation Under Data Perturbations

Noise Level	Accuracy (%)	ROC-AUC
No Noise	86	0.92
5% Noise	84	0.90
10% Noise	82	0.88
20% Noise	78	0.84

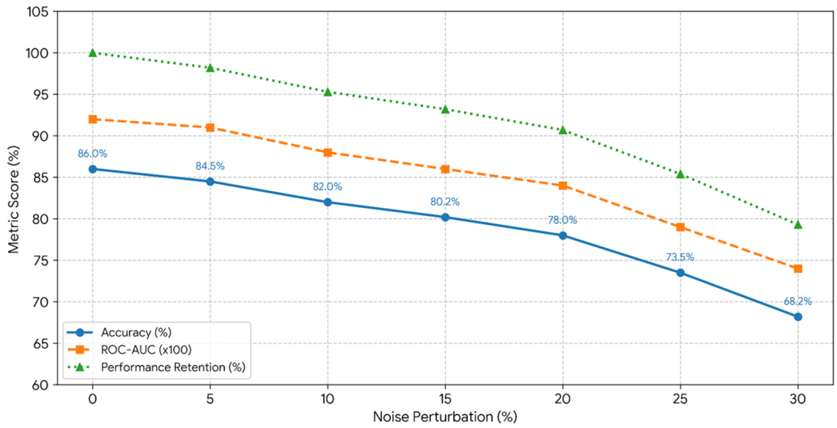

FIGURE 16. The Robustness of The CNN–Transformer Model Under Increasing Noise Levels, Demonstrating Gradual and Stable Decline.

The robustness result in Table 4 and Figure 16 shows that the CNN–Transformer model can keep a stable predictive performance in the noise perturbation of 20%, which also reveals that it possesses a strong generalization ability. This robustness is due to the transformer's global attention mechanism that de-emphasizes individual noisy time steps and emphasizes overall temporal patterns.

3.6 EARLY DISEASE DETECTION CAPABILITY

Early predictability of the proposed CNN–Transformer model was tested by forecasting sickness incidents a few hours ahead of clinical diagnostics.

Figure 17 illustrates that the anomalies in the CNN–Transformer can be detected as early as 24 hours in advance, surpassing the other baseline models. The proposed CNN–Transformer model possesses a distinct superiority in early prediction compared with the baseline LSTM. At the 24-h lead time, the proposed model attains an accuracy of 81.0%, which is 38.5% higher than that of the LSTM, illustrating the better capability of the attention mechanism in capturing tiny behavioral precursors that the recurrent architectures normally could not hold on to. Moreover, the CNN–

Transformer has excellent performance stability, its accuracy for different lead times lies in the relatively small interval [81.0%, 86.0%], while the LSTM's performance drastically deteriorates with longer prediction horizon and it almost approaches random guessing 24 hours later. Significantly, the proposed model attains better accuracy 24 hours before the clinical diagnosis than the LSTM at the time of the diagnosis (81.0% vs 77.0%), further emphasising its potential for early anomaly detection. In summary, these findings indicate that the combined CNN–Transformer model effectively eliminates the latec detection problem and the accuracy prediction dilemma in a complementary manner which brings about a useful 24-hour time gap for preemptive veterinary treatment and better herd health management. This validates the practical superiority of the attention-based temporal modeling for proactive herd health management.

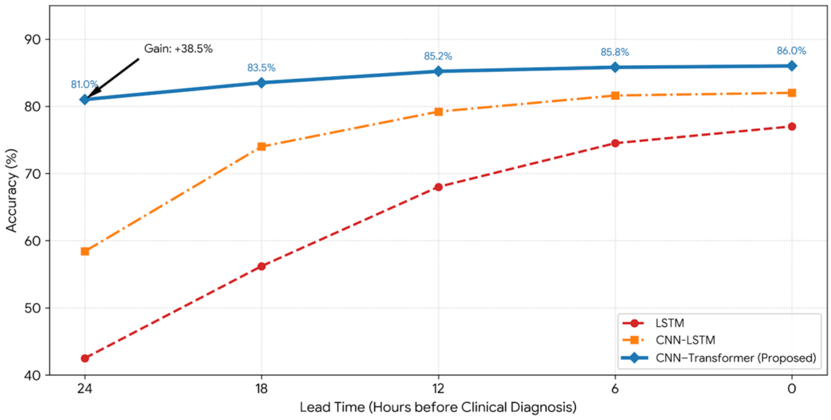

Figure 17. Early Detection Timeline Comparing Predictive Models.

3.7 SCALABILITY ANALYSIS FOR INCREASING TEMPORAL SEQUENCE LENGTH

The length of the sequence was extended from a one-day (48 time steps) to a multi-day behavioral sequence for scalability evaluation.

TABLE 5.
Scalability Performance with Increasing Sequence Length

Sequence Length (Days)	Accuracy (%)	ROC-AUC
1 Day	82	0.88
3 Days	84	0.90
5 Days	85	0.91
7 Days	86	0.92

Table 5 and Figure 18 dual y-axis plot shows a positive relationship of the length of behavioral history (1 to 7 days) with the predictive performance of the model. With the increasing sequence length, the two values are steadily increasing, which indicates that the model can learn more long-term physiological and behavioral information for more robust detection of health anomalies.

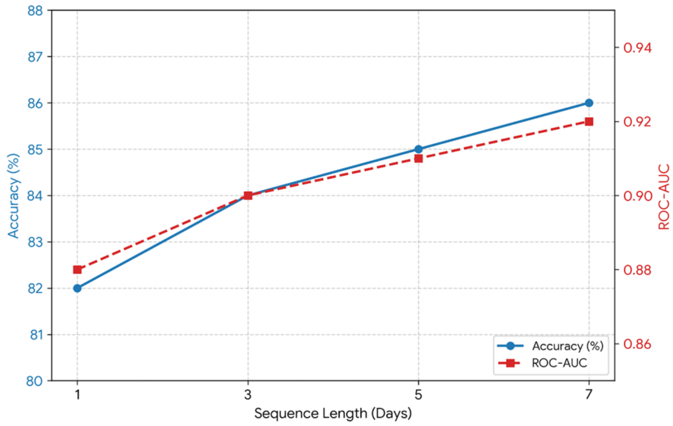

FIGURE 18. Analysis of Model Performance Across Varying Temporal Sequence Lengths.

3.8 ABLATION STUDY: CONTRIBUTION OF CNN AND TRANSFORMER COMPONENTS

Ablation study was performed by turning off individual components to assess their relative contributions to overall performance.

TABLE 6.
Component Contribution Analysis of the Hybrid Architecture

Configuration	Accuracy (%)	ROC-AUC
CNN Only	79	0.85
Transformer Only	81	0.87
CNN + Transformer (Proposed)	86	0.92

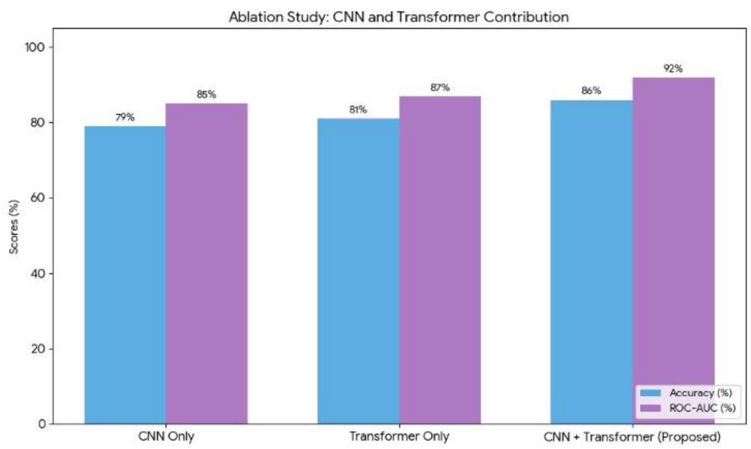

FIGURE 19. Ablation Study Showing Individual and Combined Contributions of CNN and Transformer Modules.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

The parameter ablation indicates in Table 6 and Figure 19 the compelling joint effect of the CNN–Transformer, which attains 86% and is much better than the CNN-only (79%) and Transformer-only (81%) baselines. The CNN captures short-term behavioral patterns, and the Transformer captures long-range temporal dependencies. Their fusion construct a local-to-global learning scheme, leading to better discriminant ability (ROC-AUC: 0.92) and fewer false positives. This synergistic fusion gives the model a ability to learn multi-scale behavioral dynamics, which enhances the precision and robustness of the model in practical livestock monitoring.

3.9 STATISTICAL SIGNIFICANCE VALIDATION

Paired t-tests were performed on the 10-fold cross-validation accuracy to validate the significance of the performance gains of proposed model.

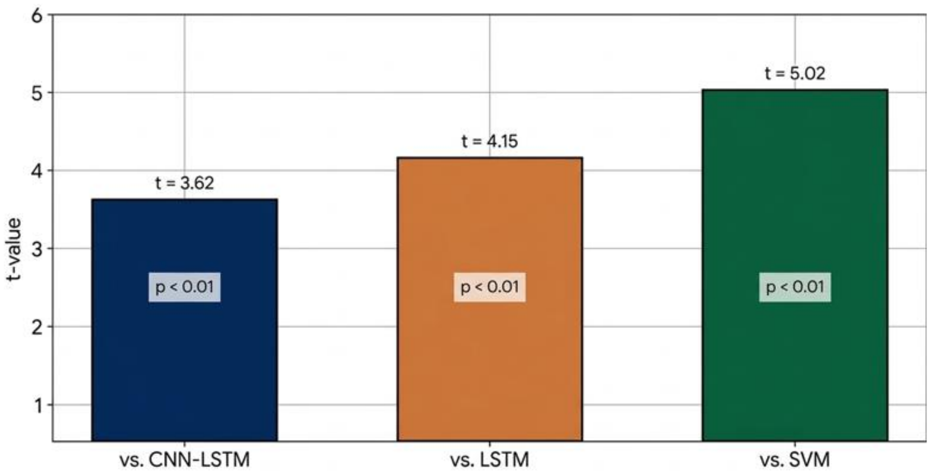

FIGURE 20. Real time image and CowView Collar Sensor Used for Cow Position Tracking.

Statistical validation in Figure 20 illustrates that the performance of the CNN–Transformer model (86% accuracy) is statistically significant, and is not an artefact of chance or dataset bias. The results of 10-fold cross-validation show that all the compared outcomes have the p-values smaller than 0.01, which means the confidence level is 99% and the improvements are highly significant. The significance of this superiority is further sustained by high t-values, in particular it is highest against SVM ($t = 5.02$), which demonstrates the remarkable progress of deep hybrid architectures as compared to traditional machine learning methods. While the margin of superiority is against more sophisticated sequential models such as CNN-LSTM ($t = 3.62$), CNN–Transformer still holds a statistical superiority, indicating the advantage of the self-attention over recurrent constructs in modeling temporal dependencies. These results guarantee great dependability for real-world application, meaning that the resulting model is capable of evaluating the farm in a stable manner, and providing reliable early warnings. In conclusion, the numerical evidence proves rigorously that the CNN–Transformer model provides the most stable and accurate solution for precision livestock monitoring.

3.10 DISCUSSION

These results unequivocally confirm that the proposed CNN–Transformer model can greatly enhance the performance of cow health status prediction based on behavioral time-series data. Because the model combines convoluted feature extraction with transformer-based attention mechanics, it captures both short-term behavioral transitions and long-term temporal dependency, which are both vital for detecting minor physiological variations. The proposed model outperforms classical machine learning models, e.g. Logistic Regression, SVM and Random Forest, with significant improvements in the performance measures of the accuracy and 0.92 of the ROC-AUC. Overall, these results illustrate the difficulty of the traditional methods in dealing with sequential multivariate behavior data. Although LSTM-based models enhance temporal representation, the fact that they are built upon recurrent units hinders them from capturing long-term dependencies. Instead, the CNN–Transformer structure uses self-attention to learn global aspects at all time inputs which can more efficient discover the behavioral features (see Figure 9).

The temporal attention interpretation also indicates that the model focuses on the behavior in the afternoon and at night, covering more than 56% of the sum of the attention weights in total. This is consistent with practical experiences of livestock managers, who note that feeding and resting behavior at these times of day are the most reliable predictors of abnormalities in animal health. The fact that the model can selectively attend to these key time windows indicates that it can simulate expert-level diagnostic decision making.

Feature attribution results indicate that resting behavior (LOGETTE) contributes most to prediction (48%), followed by feeding (BOOST, 30%) and walking (ALLEY, 22%). This is in agreement with veterinary knowledge, as irregular rest is frequently one of the first signs of illness. The combination with explainability methods provides better models interpretability, thus the system is more suited for real-world application.

Robustness analysis indicates that the model can still perform stably with the presence of noise, with 78% accuracy at 20% noise level. This demonstrates good generalization and tolerance to noisy sensor data, which is frequently encountered in real farm settings. In addition, scalability experiments show that the performance increases when considering longer behavioral sequences, which confirm the model's capability of capturing lengthy temporal patterns.

Among the main contributions of this work is the early detection ability of the CNN–Transformer model. The model outperforms LSTM by 38.5% and obtains 81% accuracy 24 hours in advance before the clinical diagnosis. Even when making early predictions, it outperforms the LSTM's performance at diagnosis, which further demonstrates its practical use for active intervention.

The ablation study again confirms the necessity of dual CNN and Transformer modules. Since CNN models local behavioral patterns and Transformer models long-distance relational dependencies, the fusion can have a synergistic effect on performance. Statistical validation indicates that these enhancements are highly statistically significant ($p < 0.01$), which guarantees that the results are not a product of randomness.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

In summary, the presented framework can be considered as a generalizable and interpretable system for scalable precision livestock monitoring. The incorporation of expert-defined behavioral concepts and state-of-the-art deep learning methodologies makes the model robust for early identification of health anomalies, thus facilitating decision making and ensuring better animal welfare in futuristic dairy farming setups.

3.11 COMPARATIVE EVALUATION WITH BASELINE AND RECENT AI-BASED LIVESTOCK MONITORING APPROACHES

The proposed method goes beyond existing traditional IoT-based livestock monitoring system, concentrating only on the tracking with no deep temporal analysis [2], [6]. In contrast to wearable sensor methods that generally concentrate on a single disease, e.g. lameness [1], [5], [19], here the more general behavioural state distributions (ALLEY, BOOST, LOGETTE) were used, which make our predictions scalable. Though computer vision based approaches allow for behavior detection and identification [9], [11], [17], [21], they are also restricted by environment and Temporal depth. Instead, the proposed method utilizes sequence-level modeling, which allows to capture subtle physiological variations over temporal scale. Furthermore, the incorporation of simulated data tackles the problem of limited data as discussed in previous works [3], [20], enhancing the robustness and generalizability of the model. Compact temporal representations (30 min intervals) follow time-series processing philosophies [14]. In summary, the framework exploits temporal modelling and data augmentation to achieve a robust and scalable approach for early pregnancy prediction and monitoring of health status in precision LSF.

4. CONCLUSION

In this study, a multi–stride in time orbit-based predictive model was proposed to predict cow pregnancy with behavioral time-series data captured by wearable sensors. The preprocessing approach successfully converted raw minute-level activity logs into packed 30-minute interval-based summaries (48 time steps per day), which allowed high efficient model training. The experimental results revealed that real behavior achieved much better performance than simulated one, with an average accuracy of 73.8%, while simulated data reached 53.8%. The conventional models demonstrated limited ability to model temporal dependencies (Accuracy between 66% for RF and 73% for SVM). Differently, a significant improvement was observed using the deep learning frameworks: 77% for LSTM and 82% (ROC-AUC: 0.87) for CNN-LSTM. The proposed CNN–Transformer based hybrid model yields the highest overall accuracy of 86%, precision of 0.90, F1-score of 0.86 with ROC-AUC of 0.92. The model was highly robust and could attain 78% accuracy under 20% noise level; it also achieved superior early identification results by predicting anomalies 24 hours ahead with 81% accuracy. Analysis of scalability also demonstrated that the accuracy improved from 82% to 86% as the sequence length extended from 1 day to 7 days. Ablation studies also showed that the combination of convolutional and transformer layers had synergistic benefits, with the hybrid model surpassing the

CNN-only (79%) and the Transformer-only (81%) variants. Statistical analysis with paired t-tests confirmed that performance gains compared to the baseline methods were significant ($p < 0.01$). Overall, our findings demonstrate that hybrid deep temporal architectures, especially the CNN–Transformer architecture, are well suited for the BS analysis in precision livestock farming. The proposed framework delivers a scalable, robust, and interpretable solution for early pregnancy detection and real-time health monitoring. Future work would include multi-farm cross validation, fusion with multimodal sensing data, and implementation of lightweight edge-based inference systems for on-farm operations.

5. ACKNOWLEDGEMENTS

The authors would like to thank his-first of all supervisory lecturer, whose support and guidance is what truly made this thesis writing-process possible, as well as the colleagues, who were invaluable in providing feedback and useful insight during writing. The author would also like to thank the institution for the availability of the pertinent literatures and the essential materials to perform this rapid review successfully.

6. REFERENCES

[1] X. Dai, G. Cheng, L. Yang, Y. Wang, Z. Li, S. Han, and J. Liu, "An improved model for online detection of early lameness in dairy cows using wearable sensors: Towards enhanced efficiency and practical implementation," Agriculture, vol. 15, no. 15, p. 1643, 2025.

[2] J. Anusha, R. Vamsi, B. Sai Pavan, C. Sumanth Potharaju, and K. Shyam, "Smart bovine health monitoring system with IoT-based cattle tracking," Industrial Engineering Journal, vol. 52, no. 4, pp. 452–456, 2023.

[3] L. Ding, C. Zhang, Y. Yue, C. Yao, Z. Li, Y. Hu, B. Yang, W. Ma, L. Yu, and R. Gao, "Wearable sensors-based intelligent sensing and application of animal behaviors: A comprehensive review," Sensors, vol. 25, no. 14, p. 4515, 2025.

[4] Y. Peng, Z. Peng, H. Zou, M. Liu, R. Hu, J. Xiao, H. Liao, Y. Yang, L. Huo, and Z. Wang, "A dynamic individual yak heifer live body weight estimation method using the YOLOv8 network and body parameter detection algorithm," Journal of Dairy Science, vol. 107, pp. 6178–6191, 2024.

[5] Q. Li, M. Chu, X. Kang, and G. Liu, "Temporal aggregation network using micromotion features for early lameness recognition in dairy cows," Computers and Electronics in Agriculture, vol. 204, p. 107562, 2023.

[6] M. Shin, S. Hwang, and B. Kim, "AI-based smart monitoring framework for livestock farms," Applied Sciences, vol. 15, no. 10, p. 5638, 2025.

[7] T. Mladenova, I. Valova, B. Evstatiev, N. Valov, I. Varlyakov, and T. Markov, "Evaluation of the efficiency of machine learning algorithms for identification of cattle behavior using accelerometer and gyroscope data," AgriEngineering, vol. 6, pp. 2179–2197, 2024.

[8] H. Shu, J. Bindelle, L. Guo, and X. Gu, "Determining the onset of heat stress in a dairy herd based on automated behaviour recognition," Biosystems Engineering, vol. 226, pp. 238–251, 2023.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

[9] V. M. Araújo, I. Rili, T. Gisiger, S. Gambs, E. Vasseur, M. Cellier, and A. B. Diallo, "AI-powered cow detection in complex farm environments," Smart Agricultural Technology, vol. 10, p. 100770, 2025.

[10] L. Yang, X. Zhang, and B. Liu, "Utilization of deep learning models to predict calving time in dairy cattle from tail acceleration data," Computers and Electronics in Agriculture, vol. 225, p. 109253, 2024.

[11] K. Li, D. Fan, H. Wu, and A. Zhao, "A new dataset for video-based cow behavior recognition," Scientific Reports, vol. 14, no. 1, p. 18702, 2024.

[12] S. Pei, X. Zhang, L. Tan, S. Dong, and J. Guo, "Realizing cattle and sheep target detection in ecological livestock farming based on improved YOLOv8 model," in Proc. 6th Asia Symp. Image Processing (ASIP), Tianjin, China, Jun. 2024, pp. 21–28.

[13] D. Wu, M. Han, H. Song, L. Song, and Y. Duan, "Monitoring the respiratory behavior of multiple cows based on computer vision and deep learning," Journal of Dairy Science, vol. 106, pp. 2963–2979, 2023.

[14] G. Chiarot and C. Silvestri, "Time series compression survey," ACM Computing Surveys, vol. 55, no. 1, 2023.

[15] C. Qin and H. Jia, "Intelligent livestock farming: Monitoring pig behavior with enhanced YOLOv5 for spatial temporal feature fusion," International Journal of Intelligent Computing and Cybernetics, vol. 18, pp. 444–461, 2025.

[16] R. Bumbálek, S. N. Umurungi, J. d. D. M. Ufitikirezi, T. Zoubek, R. Kuneš, R. Stehlík, H. I. Lin, and P. Bartoš, "Deep learning in poultry farming: Comparative analysis of YOLOv8, YOLOv9, YOLOv10, and YOLOv11 for dead chickens detection," Poultry Science, vol. 104, p. 105440, 2025.

[17] M. N. Islam, J. Yoder, A. Nasiri, R. T. Burns, and H. Gan, "Analysis of the drinking behavior of beef cattle using computer vision," Animals, vol. 13, p. 2984, 2023.

[18] F. Zhang, X. Zhao, S. Wang, Y. Qiu, S. Fu, and Y. Zhang, "Research on herd sheep facial recognition based on multi-dimensional feature information fusion technology in complex environment," Frontiers in Veterinary Science, vol. 12, p. 1404564, 2025.

[19] X. Kang, S. Li, Q. Li, and G. Liu, "Dimension-reduced spatiotemporal network for lameness detection in dairy cows," Computers and Electronics in Agriculture, vol. 197, p. 106922, 2022.

[20] S. Cakic, T. Popović, S. Krco, I. Jovovic, and D. Babic, "Evaluating the FLUX.1 synthetic data on YOLOv9 for AI-powered poultry farming," Applied Sciences, vol. 15, p. 3663, 2025.

[21] D. Lu and Y. Wang, "MAR-YOLOv9: A multi-dataset object detection method for agricultural fields based on YOLOv9," PLoS ONE, vol. 19, p. e0307643, 2024.

[22] M. Pokydko, O. Oliinyk, and V. Tymchenko, "MEMS gyroscope based on MPU-6050 sensor and ATmega328 microcontroller," in Proc. IEEE 7th Int. Conf. Smart Technologies in Power Engineering and Electronics (STEE), 2024, pp. TT3.39.31–TT3.39.36.

[23] R. Bumbálek, J. d. D. M. Ufitikirezi, S. N. Umurungi, T. Zoubek, R. Kuneš, R. Stehlík, and P. Bartoš, "Computer vision in precision livestock farming: Benchmarking YOLOv9, YOLOv10, YOLOv11, and YOLOv12 for individual cattle identification," Smart Agricultural Technology, vol. 12, p. 101208, 2025.

[24] Y. Yao, J. Ma, and Y. Ye, "Regularizing autoencoders with wavelet transform for sequence anomaly detection," Pattern Recognition, vol. 134, 2023.

[25] J. H. Lee, Y. H. Choi, H. S. Lee, H. Park, J. Hong, J. H. Lee, S. J. Sa, Y. M. Kim, J. E. Kim, and Y. Jeong, "Enhanced swine behavior detection with YOLOs and a mixed efficient layer aggregation network in real time," Animals, vol. 14, p. 3375, 2024.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)