From Cattle Images to Yield: Hybrid Transfer Learning Framework for Robust Milk Yield Prediction in Smart Dairy Farming

Hager Saleh[1,2,3], Rabab Ali Abumalloh[4], Michael McCann[4], John G. Breslin[1,5], and Shaker El-Sappagh[6,7]

[1] [1] Insight SFI Research Centre for Data Analytics, University of Galway, Galway H91 TK33, Ireland
[2] [2] Atlantic Technological University, Letterkenny, Ireland
[3] [3] Faculty of Computers and Artificial Intelligence, Hurghada University, Hurghada, Egypt
hager.saleh.fci@gmail.com
[4] [4] Department of Computing, Atlantic Technological University, Letterkenny, Ireland
[5] [5] School of Engineering, University of Galway, Galway H91 TK33, Ireland
[6] [6] Faculty of Computer Science and Engineering, Galala University, Suez, Egypt
[7] [7] Faculty of Computers and Artificial Intelligence, Benha University, Banha, Egypt

Abstract. Milk yield prediction is a cornerstone of smart dairy farming, yet existing approaches remain limited by the scarcity of annotated data, environmental variability, and the restricted capacity of traditional models such as convolutional neural networks (CNNs). Although prior work has applied CNNs or lightweight image-based approaches, no study to date has systematically combined pretrained deep models with ensemble regressors for image-driven yield prediction. To address this gap, we propose a novel hybrid framework that integrates transfer learning with tree-based regressors for robust and accurate milk yield estimation. Pretrained architectures (ResNet50, EfficientNetB0, and MobileNetV2) were employed as feature extractors for side, rear, and combined cow images, followed by fine-tuning. To further enhance predictive power, deep features were coupled with XGBoost and Random Forest regressors, enabling the capture of complex nonlinear dependencies. Experimental results demonstrate that transfer learning consistently outperforms baseline CNNs. At the same time, the hybrid ResNet50–XGBoost achieves the best overall performance, with an RMSE of 1327.87, 1355.56, and 1287.87 for side, rear, and combined views, respectively, and corresponding R^2 values of 0.3648, 0.3146, and 0.3948. Compared with previously reported benchmarks, our approach reduces RMSE by more than 110 units, establishing new state-of-the-art performance. The contributions of this study are threefold: (1) a comprehensive evaluation of baseline, transfer learning, and hybrid strategies for milk yield prediction; (2) a demonstration of the effectiveness of pretrained CNNs in reducing dependence on large labeled datasets; and (3) the introduction of a hybrid deep feature–tree regressor framework that consistently enhances prediction accuracy across multiple datasets. These results confirm the potential of the proposed framework as a scalable and reliable tool for precision livestock farming.

Keywords: Computer vision · milk yield forecasting · Convolutional neural network · transfer learning in smart farming · hybrid deep learning.

1 Introduction

In dairy farming, yield is a crucial performance metric that has a direct impact on profitability as well as productivity [1]. The importance of the dairy sector in terms of economic output lies mainly in two directions: its contribution to the job market, and the production of food products and supplies [2]. Forecasting the milk production of cows, particularly while they are still heifers, has gained the interest of farmers and agricultural researchers. These forecasts can help farmers with financial planning, feeding information, mating decisions, cow replacement and culling, and identifying abnormal production trends that could be signs of mastitis [2].

Dairy production plays a significant role in the Irish market. Family farming, spanning multiple generations, forms the foundation of Ireland's dairy industry, with an estimated 16,000 family-run dairy farms across the country [3]. Irish exports of dairy products were estimated at €6.3 billion last year [4], with products totaling more than 1.6 million tonnes transported to almost 140 markets around the world [4]. Increased cow numbers, higher milk yields per cow, improved fat and protein content, higher stocking rates, and more land entering the market have all contributed to this strong performance. However, farmers must remain committed to building a robust and successful long-term business, as the milk industry still has much room to expand and improve its performance.

A well-known data-driven application in the field of farming is the forecasting of milk yield, a strategically significant area being the primary source of revenue for the dairy industry [5]. Nonetheless, milk yield prediction is a complicated task, impacted by a number of diverse variables, such as the health of animals, environmental circumstances, feeding habits, and genes [6]. Farmers face difficulties in maximising production output while preserving sustainability, calling for sophisticated prediction tools to improve decision-making processes. This call is highlighted by the fact that variations in milk yield can lead to greater environmental impact, wasteful resource use, and unstable financial situations [7]. Farming has advanced with the incorporation of digital technologies, leading to the development of more environmentally friendly, information-driven, and effective systems. Computer vision models based on deep learning techniques, through the use of CNNs [8], enable the utilisation of visual data for predictions. Several visual features could be utilised in the prediction of milk yield: udder traits, rump width, and angularity [2].

Although CNNs have achieved very strong performance in agricultural computer vision tasks, their practical deployment in smart farming remains constrained by several key limitations. Conventional CNN architectures typically require large, well-annotated datasets to achieve strong generalisation. Yet, such datasets are often difficult to obtain due to the high cost and the variability of environmental conditions [9, 10]. CNNs trained from scratch are prone to overfitting when data is limited, and may struggle to capture complex hierarchical patterns without careful architectural optimisation [11, 12]. To mitigate these challenges, transfer learning has emerged as a powerful paradigm. It enables models to leverage feature representations learned from large-scale datasets such as ImageNet and adapt them to domain-specific tasks. Pretrained architectures such as ResNet50 [11], EfficientNet [12], and MobileNet [13] provide rich, general-purpose features that can be fine-tuned to agricultural applications, significantly reducing the need for extensive labeled data and accelerating convergence. These pretrained models enhance resilience to environmental noise, image/video lighting variability, and heterogeneous data distributions, making transfer learning a cornerstone technique for advancing deep learning applications in precision agriculture [9, 10].

Recent research has explored hybrid transfer learning approaches in which the final layer of a pretrained CNN is replaced by a strong machine learning regressor to boost predictive accuracy. In this strategy, deep networks such as ResNet50, EfficientNet, or MobileNet act as feature extractors, while gradient boosting models like XGBoost or ensemble methods such as Random Forests learn from the extracted features [14,15]. This combination exploits the rich representations of deep models and the robustness of tree-based learners, often outperforming purely deep or traditional methods, especially with limited labeled data [16]. Hybrid deep feature–tree ensemble frameworks have shown gains in tasks such as crop yield estimation, livestock monitoring, and plant disease detection [16]. However, to the best of our knowledge, no existing work has applied a pretrained CNN with an XGBoost regressor for milk yield prediction, revealing a clear research gap that is addressed in this study. Less research has utilised images with deep learning models for milk yield prediction. Only a modest number of studies have utilized images with deep learning models for milk yield prediction. For example, Jembere et al. [2] deployed a dataset that includes 1238 images of side-view and rear-view images of 743 Holstein cows in their first or second parity, along with their corresponding first lactation and 305-day milk yield values. Different augmented methods, such as flipping, stretching, and adding Gaussian noise were applied to the training set. They applied CNN models to side-view, rear-view, and combined view images. The results showed that CNN models recorded the best results with the combined view. Using the same dataset, Allan et al. [17] applied YOLOv11 models to classify dairy cows into low, medium, and high based on the milk yield. According to the authors, misclassifications mostly happen close to class boundaries, building on the results of qualitative analysis, and indicating the need for reliable picture acquisition settings. These results show how vision-based models can be used to assist in decision-making in such systems, especially in situations where conventional data collection techniques are impractical or unavailable. These two studies did not use hybrid models that combined transfer learning with machine learning to enhance the results of milk yield prediction. In our research, the contributions of the paper can be summarised as follows:

- We propose a novel deep learning framework that integrates pretrained CNN feature extractors (ResNet50, EfficientNetB0, and MobileNetV2) with a powerful XGBoost regressor to predict milk yield from cow images.
- The study leverages transfer learning to minimise the need for large labeled datasets while maintaining high predictive accuracy and robustness.
- We conduct extensive experiments, comparing multiple pretrained CNN backbones and machine learning regressors, to identify the most effective architecture for milk yield prediction.

The remainder of this paper is arranged as follows: the proposed framework is introduced in Section 2. The experimental results are discussed in Section **??**. Finally, conclusions are given in Section 4.

2 Methodology

This section describes the proposed image-based milk-yield prediction framework following the processing flow shown in Fig. 1. The pipeline is organised into six phases: (1) data acquisition; (2) image pre-processing and augmentation; (3) dataset splitting; (4) baseline deep learning models; (5) model optimisation; and (6) model evaluation.

2.1 Phase 1: Data acquisition

The dataset [2] comprises images from four farms containing 228, 151, 241, and 123 cows, with 360, 260, 394, and 224 images, respectively. The dataset includes two CSV files: training and testing. Each file contains milk yield values and the path of images (side-view and rear-view) for the caw in either their first or second lactation. It also includes a CSV file that contains milk yield values with the corresponding 305-day milk yield. Images include both side and rear views and depict cows in either their first or second lactation; milk-yield records are prospective for first-lactation cows and retrospective for second-lactation cows.

2.2 Phase 2: Image pre-processing and augmentation

Pre-processing is the deterministic mapping $\mathcal{P} : \mathcal{X} \to \widetilde{\mathcal{X}}$ that standardises raw images prior to feeding them to a backbone network. Components of $\mathcal{P}$ include:

- **Resizing:** rescale input images to a fixed resolution (H_0, W_0) appropriate for the chosen backbone (e.g., 224×224);
- **Cropping:** center or content-aware cropping to emphasise anatomically relevant regions (e.g. udder, rump);

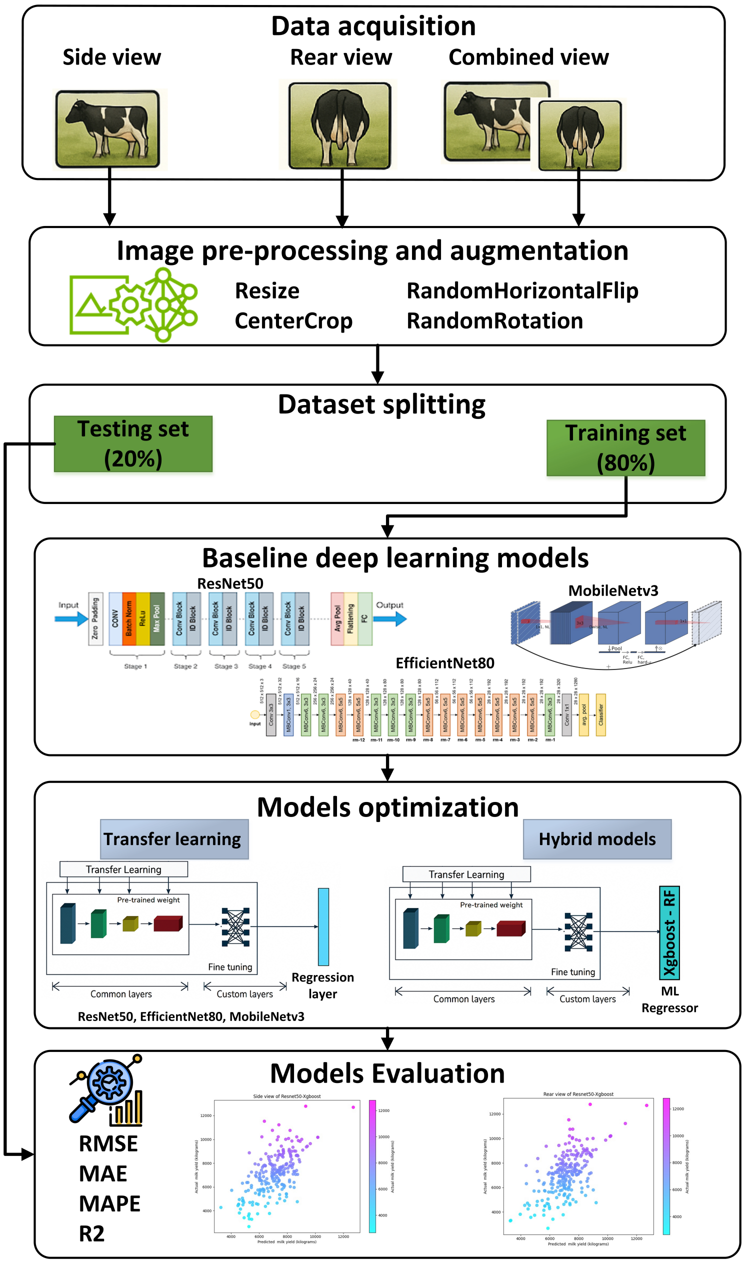

Fig. 1. The proposed image-based milk-yield prediction framework

– **Channel normalisation:** for each color channel c normalise by per-channel mean and standard deviation,

$$\widetilde{x}^{(c)} = \frac{x^{(c)} - \mu_c}{\sigma_c},$$

where (μ_c, σ_c) are precomputed (e.g., ImageNet statistics or dataset-specific values).

2.3 Phase 3: Dataset splitting

The dataset includes two CSV files: training (1000 rows) and testing (237 rows). Each file includes milk yield values, path of images for side view and path of image for side-view and rear-view.

2.4 Phase 4: Baseline deep learning models

We applied ResNet50, EfficientNetB0, and MobileNetV2 as pre-trained CNNs to compare with hybrid models.

2.5 Phase 5: Models optimization

Phase 5 covers parameter estimation, transfer-learning regimes, hybridization with tree-based regressors, and hyperparameter search.

Transfer learning: feature extraction vs. fine-tuning. Two transfer regimes are used: (1) Feature extraction (frozen backbone): initialise backbone parameters $\theta = \theta_0$ from ImageNet pretraining and keep them fixed. Optimise only the regression head:

$$\phi^\star = \arg\min_{\phi} \frac{1}{|\mathcal{D}_{tr}|} \sum_{(x,y)\in\mathcal{D}_{tr}} \ell\big(y, g_\phi(f_{\theta_0}(x))\big) + R(\phi),$$

where $R(\phi)$ is a regulariser (e.g., ℓ_2). (2) Fine-tuning: update a subset (or all) of backbone parameters θ_t together with head parameters ϕ:

$$(\theta_t^\star, \phi^\star) = \arg\min_{\theta_t,\phi} \frac{1}{|\mathcal{D}_{tr}|} \sum_{(x,y)\in\mathcal{D}_{tr}} \ell\big(y, g_\phi(f_{[\theta_0\setminus\theta_t;\theta_t]}(x))\big) + \lambda\big(\|\theta_t\|_2^2 + \|\phi\|_2^2\big),$$

where $f_{[\theta_0\setminus\theta_t;\theta_t]}$ denotes that parameters outside θ_t are fixed to their pretrained values.

Fine-tuning is typically performed with a lower learning rate for θ_t relative to the head to mitigate catastrophic forgetting.

Hybrid models: pretrained deep features + tree regressor. The central methodological novelty in this work is the systematic evaluation of hybrid predictors in which the network backbone yields deep features and the final prediction is produced by a tree-based regressor (XGBoost or Random Forest). The pipeline is:(1) feature extraction: $\mathbf{z}_i = \psi(f_\theta(x_i)) \in \mathbb{R}^d$, where $\psi(\cdot)$ denotes pooling and optional dimensionality reduction (e.g., principal components or selected intermediate activations); (2) tree-based regression: fit an ensemble to $(\mathbf{z}_i, y_i)$. For XGBoost the model is an additive ensemble $\hat{y}_i = \sum_{k=1}^{K} f_k(\mathbf{z}_i)$ learned by minimising

$$\mathcal{J}(\{f_k\}) = \sum_{i=1}^{|\mathcal{D}_{tr}|} \ell\Big(y_i, \sum_{k=1}^{K} f_k(\mathbf{z}_i)\Big) + \sum_{k=1}^{K} \Omega(f_k),$$

with regulariser

$$\Omega(f) = \gamma T + \tfrac{1}{2}\lambda \|\mathbf{w}\|_2^2,$$

where T is the number of leaves in tree f and $\mathbf{w}$ are leaf weights. We evaluate both: (1) frozen-features hybrid by extract $\mathbf{z}_i$ from θ_0 and train the tree ensemble; (2) fine-tuned-features hybrid by first fine-tune the backbone (as described above), then extract improved features $\mathbf{z}_i$ and train the tree ensemble. The latter often produces features more adapted to the downstream regressor.

2.6 Phase 6: Models evaluation

Evaluation is performed on the held-out test set $\mathcal{D}_{te} = \{(x_j, y_j)\}_{j=1}^{M}$ using the standard regression evaluation metrics including root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and coefficient of determination (R2).

3 Results

We conducted experiments to predict milk yield values based on side-view, rear-view, and combined views. In addition, we compare the results of CNN baseline models: MobileNetV2, ResNet50 and EfficientNetB0 based on default weights, transfer learning models using a default regressor as a linear layer, and hybrid models in which we replace the linear layer with strong regression, including XGBoost and random forest regression.

3.1 Experimental Setup

The experimental platform's hardware configuration includes an Intel i7-6700 CPU, an RTX 4090 graphics card, 16 GB of memory, a Windows 11 operating system, and a model implemented using Python and the PyTorch framework. For setting the model parameters, the optimiser is AdamW, the batch size is 32, the number of epochs is 50 (with early stopping), the loss function is MSELoss, and the learning rate (LR) is 1e-3.

To prepare the target labels, the mean and standard deviation of the label column are first calculated. Additionally, a denormalisation function is provided to convert the normalised predictions back to their original scale after inference.

3.2 Results of models

The results based on the three versions of the data are discussed below. Table 1 presents the predictive performance of the proposed framework across the three datasets (Side view, Rear view, and Combined view) using different learning strategies. The findings clearly demonstrate that transfer learning consistently improves model performance over the baseline CNN models, and the hybrid integration of ResNet50 features with tree-based regressor (XGBoost or RF) yields further notable gains.

With the **Side View Dataset**, as shown in Table 1, ResNet50 transfer learning reduced the RMSE from 1392.80 to 1375.58 and improved R^2 from 0.3262 to 0.3489. The hybrid ResNet50–XGBoost achieved the best results with an RMSE of 1327.87, MAPE of 15.59, and R^2 of 0.3648, outperforming both baseline and transfer-only models. Importantly, while transfer learning slightly increased MAE (1115.75 to 1187.26), the hybrid reduced it to 1105.45, thus correcting the inconsistency and delivering the most balanced results across all metrics.

Figure 2 (a) shows the scatter plot of actual versus predicted milk yield values using the side-view dataset with the hybrid ResNet50–XGBoost model. The points are generally well aligned with the regression line, indicating good predictive accuracy. However, some dispersion remains at the upper yield range, suggesting that side-view features alone may not fully capture all predictive cues for high-producing cows. Nevertheless, the model achieves substantial improvements compared with baseline and transfer-only models, reducing RMSE to 1327.87 and increasing R^2 to 0.3648, confirming that the hybrid approach extracts more meaningful side-view features.

With the **Rear View Dataset**, transfer learning achieved substantial improvements, as shown in Table 1. With ResNet50, RMSE decreased from 1533.06 to 1370.42 and R^2 rose from 0.2043 to 0.2692. The hybrid ResNet50–XGBoost further enhanced performance, lowering RMSE to 1355.56, reducing MAE to 1122.91, and increasing R^2 to 0.3146. The cumulative baseline-to-hybrid improvements are significant: a reduction of 177.49 in RMSE and a gain of $+0.1103$ in R^2. This indicates that hybridisation is particularly effective in handling rear images, which capture critical predictive features.

Figure 2 (b) illustrates the performance of the hybrid model when trained on rear-view images. Here, predictions are closely clustered around the regression line, with less variability compared to the side view. This suggests that rear-view features, such as udder and hindquarter characteristics, provide stronger predictive signals for milk yield. The model achieved RMSE of 1355.56 and R^2 of 0.3146, improving over both baseline CNNs and transfer learning. Although performance slightly lags behind the side view in terms of R^2, the reduced MAE of 1122.91 highlights the robustness of rear-view predictions.

With the **Combined View Dataset**, transfer learning again provided consistent gains, as shown in Table 1. ResNet50 reduced RMSE from 1325.60 to 1317.44, MAE from 1144.63 to 1076.23, and improved R^2 from 0.3329 to 0.3619. The hybrid ResNet50–XGBoost achieved the best performance overall, with RMSE of 1287.87, MAE of 1005.45, MAPE of 15.20, and R^2 of 0.3948. These results demonstrate that combining multiple views enriches feature representation and benefits most from the hybrid strategy. Overall, the results across all datasets confirm three key contributions: (1) transfer learning consistently enhanced prediction accuracy compared with baseline CNN models, (2) ResNet50 provides the strongest backbone among the tested architectures, and (3) hybridising ResNet50 features with XGBoost yields further improvements in RMSE, MAE, MAPE, and R^2, achieving the best balance between accuracy and robustness. The consistency of these improvements highlights the effectiveness of the proposed framework for smart farming applications.

Figure 2 (c) presents the results for the combined-view dataset, where side and rear images were integrated. The scatter plot demonstrates the closest alignment between predicted and actual values, with minimal dispersion across the yield range, including high-producing cows. This dataset achieved the best overall results, with RMSE reduced to 1287.87, MAE to 1005.45, and R^2 reaching 0.3948. The tight clustering of points highlights the complementary nature of side and rear features, confirming that multimodal visual information enhances predictive power.

3.3 Comparison with literature study

When compared with previously reported benchmark results [2], our proposed Hybrid ResNet50–XGBoost framework demonstrates consistent and substantial improvements across all datasets. On the side view dataset, the hybrid approach reduced RMSE to 1327.87 (a decrease of 132.8 units, $\approx$9.1%), lowered MAE to 1105.45 (41 points lower), and improved R^2 to 0.3648 (+0.0448). For the Rear view dataset, RMSE dropped to 1355.56 (124.9 units lower, $\approx$8.4%), MAE decreased to 1122.91 (25.4 points lower), and R^2 increased to 0.3146 (+0.0126). On the Combined view dataset, the hybrid model achieved the strongest performance, with RMSE reduced to 1287.87 (113.3 units lower, $\approx$8.1%), MAE lowered to 1005.45 (107.5 points lower), and R^2 reaching 0.3948 (+0.0198). These consistent gains across all evaluation metrics highlight three key contributions of our study: (1) systematic reductions in RMSE of more than 110 units for every dataset, (2) notable decreases in MAE, including an improvement exceeding 100 points on the combined view dataset, and (3) steady enhancements in explanatory power, with R^2 increases of up to +0.045. Collectively, these results confirm that integrating transfer learning with tree-based regressor not only strengthens the predictive capacity of deep models but also surpasses previously published state-of-the-art results, establishing a more accurate and robust framework for smart farming applications.

Table 1. The results of models for side view, rear view and combined view

Side	Approaches	Models	RMSE	MAE	MAPE	R2
Side view	Baseline Models	ResNet50	1392.7975	1115.750	16.905	0.3262
		EfficientNetB0	1544.677	1314.671	18.591	0.1748
		MobileNetv2	1453.361	1131.875	17.438	0.3067
	Transfer learning	ResNet50	1375.577	1187.262	16.380	0.3489
		EfficientNetB0	1524.658	1293.775	18.141	0.2168
		MobileNetv2	1423.883	1138.036	16.806	0.3345
	Hybrid models	ResNet50-XGBoost	1327.870	1105.451	15.586	0.3648
		ResNet50-RF	1406.273	1066.502	15.872	0.3509
Rear view	Baseline Models	ResNet50	1533.0564	1301.6388	18.423	0.2043
		EfficientNetB0	1636.702	1497.691	19.319	0.1272
		MobileNetv2	1564.6033	1204.4005	18.205	0.2024
	Transfer learning	ResNet50	1370.417	1262.937	17.501	0.2692
		EfficientNetB0	1545.7710	1322.676	18.653	0.1857
		MobileNetv2	1516.244	1217.767	17.900	0.2510
	Hybrid models	ResNet50-Xgboost	1355.563	1122.913	17.079	0.3146
		ResNet50-RF	1477.898	1163.156	17.580	0.2830
Combined view	Baseline Models	ResNet50	1325.598	1144.628	16.460	0.3329
		EfficientNetB0	1542.399	1216.919	18.379	0.2192
		MobileNetv2	1456.246	1137.355	16.906	0.3239
	Transfer learning	ResNet50	1317.44	1076.234	15.596	0.3619
		EfficientNetB0	1530.297	1205.612	18.244	0.2213
		MobileNetv2	1326.283	1041.247	15.838	0.3533
	Hybrid models	ResNet50-XGBoost	1287.870	1005.451	15.198	0.3948
		ResNet50-RF	1291.128	1061.126	15.386	0.3829

4 Conclusion

This study presented a novel image-based framework for milk yield prediction that integrates transfer learning with hybrid deep learning models. By leveraging pretrained CNN architectures (ResNet50, EfficientNetB0, and MobileNetV2) as feature extractors and coupling them with tree-based regressors, we addressed the challenges of limited annotated datasets, environmental variability, and the limited generalisation of conventional CNNs. Extensive experiments on side, rear, and combined image views of cows demonstrated three key findings. Firstly, transfer learning consistently improved prediction accuracy compared to baseline CNN models, reducing error magnitudes and enhancing explanatory power across all datasets. Secondly, ResNet50 emerged as the most effective backbone, offering superior performance relative to other tested architectures. Thirdly, the proposed hybrid ResNet50–XGBoost model achieved state-of-the-art results, with RMSE values of 1327.87, 1355.56, and 1287.87 for side, rear, and combined views, respectively, and corresponding R^2 values of 0.3648, 0.3146, and 0.3948. Compared with previously reported benchmarks, these results correspond to RMSE reductions exceeding 110 units and R^2 gains of up to +0.045, confirming

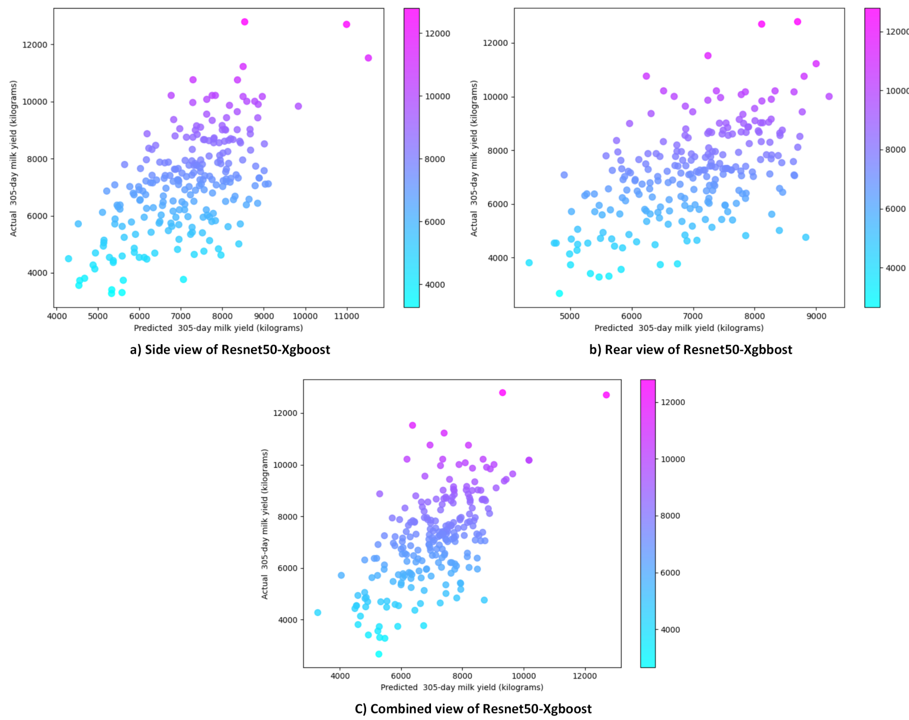

Fig. 2. the scatter plot of actual versus predicted milk yield values using the side-view dataset with the hybrid ResNet50–XGBoost model

the robustness and practical utility of the proposed framework for smart farming applications.

Beyond the contributions of this work, several promising directions emerge for future research. Firstly, explainable artificial intelligence (XAI) methods could be integrated to provide interpretability of the predictive decisions and highlight which anatomical features drive yield predictions. Secondly, incorporating uncertainty estimations would increase trustworthiness by quantifying confidence in model outputs, a crucial factor in decision making carried out under risk conditions. Thirdly, external validation on independent datasets from different herds, farms, or geographic regions is necessary to confirm the model's generalisability. Fourthly, optimisation strategies such as advanced hyperparameter tuning, evolutionary algorithms, or neural architecture search (NAS) could be applied to further improve predictive performance. Finally, extending the framework to multimodal integration (e.g., combining images with genetic or environmental data) could provide richer representations and further boost accuracy.

Acknowledgment

This publication has emanated from research supported by grants from Taighde Éireann - Research Ireland under Grant Numbers 12/RC/2289_P2 (Insight) and 21/RC/10303_P2 (VistaMilk). For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

References

1. S. Ruikar, P. Salve *et al.*, "Assessing reproductive performance, milk yield, and influential factors in dairy cows: A comprehensive investigation into productivity and efficiency," *Rev. Electrón. Vet.*, vol. 25, pp. 368–378, 2024.
2. L. Jembere and M. Chimonyo, "Prediction of milk yield using visual images of cows through deep learning," *South African Journal of Animal Science*, vol. 54, no. 1, pp. 47–58, 2024.
3. European Dairy, "Dairy Farming in Ireland," https://european-dairy.eu/dairy-farming-in-ireland/, 2025, accessed: Aug. 15, 2025.
4. Bord Bia, "Vision for the Irish Dairy Sector," https://www.bordbia.ie/industry/irish-sector-profiles/dairy-sector-profile/, 2025, accessed: Aug. 15, 2025.
5. P. R. Shorten, "Computer vision and weigh scale-based prediction of milk yield and udder traits for individual cows," *Computers and Electronics in Agriculture*, vol. 188, p. 106364, 2021.
6. J. Hooker, B. de Medeiros, C. Saha, T. Abdulrahman, and A. Alves, "Evaluating the contribution of behavioral, milking system, and environmental data to short-term milk yield prediction in commercial dairy cows using machine learning," *Journal of Dairy Science*, 2025.
7. C. A. Wolf and J. Karszes, "Financial risk and resiliency on us dairy farms: Measures, thresholds, and management implications," *Journal of Dairy Science*, vol. 106, no. 5, pp. 3301–3311, 2023.
8. H. Saleh, S. H. Alsamhi, M. McCann, N. O'Brolchain, and J. G. Breslin, "Ai meets climate: Advanced transformer multimodal model for early and accurate plant disease detection," in *2024 10th International Conference on Computing, Engineering and Design (ICCED)*. IEEE, 2024, pp. 1–6.
9. C. Chen, W. Zhu, and T. Norton, "Behaviour recognition of pigs and cattle: Journey from computer vision to deep learning," *Computers and electronics in agriculture*, vol. 187, p. 106255, 2021.
10. Y. Yuan, "Computer vision and deep learning for precise agriculture: a case study of lemon leaf image classification," in *Journal of Physics: Conference Series*, vol. 2547, no. 1. IOP Publishing, 2023, p. 012024.
11. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
12. M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*. PMLR, 2019, pp. 6105–6114.
13. A. Howard and e. a. Sandler, "Searching for mobilenetv3," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 1314–1324.
14. L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

15. T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system,” in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.
16. S. B. Punuri, S. K. Kuanar, M. Kolhar, T. K. Mishra, A. Alameen, H. Mohapatra, and S. R. Mishra, “Efficient net-xgboost: an implementation for facial emotion recognition using transfer learning,” *Mathematics*, vol. 11, no. 3, p. 776, 2023.
17. A. Hall-Solorio, G. Ramirez-Alonso, A. J. Chay-Canul, H. A. Lee-Rangel, E. Vargas-Bello-Pérez, and D. R. Lopez-Flores, “Smart dairy farming: A mobile application for milk yield classification tasks,” *Animals*, vol. 15, no. 14, p. 2146, 2025.