Egg Weight Estimation Based on Image Processing using Mask R-CNN and XGBoost

Jasman Pardede[1,*] , Muhammad Fadlansyah Zikri Akhiruddin Rawosi[2] , Anisa Putri Setyaningrum[3] , Rizka Milandga Milenio[4] , Chalifa Chazar[5]

[1,2,3,4,5]*Departement of Informatics, Institut Teknologi Nasional Bandung, Indonesia*

Abstract

Manually measuring egg weight in the context of livestock and the food industry can pose various problems, including time and labor requirements, the risk of egg damage, consistency and accuracy, and limitations on production scale. To address these issues, an automated egg weight estimation system is essential. This study proposes integrating computer vision and machine learning into a unified workflow that combines segmentation, classification, and regression for practical weight estimation. The proposed pipeline employs Mask R-CNN for egg segmentation, Random Forest (RF) classifier for egg type classification based on color features, and XGBoost for regression using morphological, geometric, color features, and egg type as predictors. The dataset used is 720 images, consisting of 20 eggs (10 chicken and 10 duck), each photographed from 36 rotational angles, and was collected with Ground Truth (GT) weights obtained from a digital scale. Experimental findings show that the RF classifier achieved perfect accuracy (precision, recall, and F1-score = 1.00) in distinguishing chicken and duck eggs. The XGBoost regressor obtained a training performance of MAE = 1.07 g and R^2 = 0.68, and a validation performance of MAE = 0.23 g and R^2 = 0.80 under 10-fold grouped cross-validation. Although a Support Vector Regressor baseline reached higher training accuracy (MAE = 0.22 g, R^2 = 0.96), it failed to generalize on validation ($R^2 < 0$), highlighting XGBoost's robustness. The feature importance analysis revealed that there are 4 (four) important features for building an estimation model, namely: Hu moments, eccentricity, elongation, and diagonal length, while color statistics played a complementary role. The novelty of this work lies in combining deep segmentation, color-based classification, and feature-driven regression into a unified framework specifically for egg weight estimation, showing its feasibility as a proof of concept and laying the foundation for future large-scale, calibrated, and externally validated deployment.

Keywords: Egg Weight Estimation, Mask R-CNN, XGBoost, Image Processing, Computer Vision, Deep Learning

1. Introduction

Eggs are an essential food commodity with high demand, making efficiency in sorting, weighing, and distribution crucial for the productivity and profitability of the livestock industry. As the global population grows, worldwide egg production has become one of the most important agricultural sectors and a vital part of the global food supply [1]. Currently, manual egg weighing requires significant human effort and time, hindering production and distribution processes. Image processing technology offers a cutting-edge solution for automating the measurement of egg weight. It delivers faster and more precise results with minimal human involvement, helping reduce costs and boost competitiveness in the market. Several studies indicate that computer vision-based systems can estimate egg weight and size with accuracy comparable to destructive methods, but without damaging the eggs [2]. Furthermore, camera-based automation systems and morphometric analysis have been proven to accelerate the classification and grading of eggs on multi-flow production lines in real time [3].

The performance of egg weight measurement using image and machine learning approaches has shown high correlation with actual weight, with low estimation error in field tests [4]. Machine learning and computer vision aim to bring human capabilities in data sensing, data understanding, and taking action based on past and present results into computers [5]. Consequently, the application of computer vision can significantly improve operational efficiency within the poultry industry.

In recent years, the development of deep learning and machine learning has opened new opportunities for image estimation and detection. Several previous studies have been conducted related to object weight estimation based on images using machine learning and deep learning, or even both. For instance, the study by [6] estimated pig weight by combining Mask R-CNN based image segmentation with an ensemble regression model. The best results were achieved using XGBoost, which delivered high accuracy with an MAE of 0.389, RMSE of 0.576, MAPE of 0.318%, and R² of 0.995. This high accuracy was supported by the use of Mask R-CNN to extract pig contours more accurately through binary image conversion to address lighting issues. Additionally, key features such as Body Length (BL), Hip Width (HW), and the estimated camera to back distance (H_dep) were corrected and utilized by XGBoost based on actual pig measurements [6].

In contrast to previous studies, this research focuses on developing an image-based egg weight estimation model using Mask R-CNN for segmentation and XGBoost for regression by integrating both morphological and color-based features specifically the mean values of HSV and RGB channels as well as the inclusion of egg type (chicken or duck) to improve model generalization. The predictive features include a combination of shape, size, color attributes, and Hu moments, such as width, height, area_bbox, aspect_ratio, cx, cy, diagonal_length, mask_area, mask_perimeter, extent, solidity, eccentricity, elongation, compactness, circularity, equivalent_diameter, hull_perimeter, Hu moments 1 to 7, mean_h, mean_s, mean_v, mean_r, mean_g, mean_b, and egg_type. The dataset consists of labeled images of various chicken and duck eggs paired with their actual weights, and model performance is evaluated using MAE, RMSE, MAPE, and R².

2. Literature Review

This literature review discusses egg characteristics, CNN architectures for image classification, Mask R-CNN for object segmentation, and XGBoost as a regression model for image-based estimation all of which are relevant to the development of an automated system aimed at addressing the challenges of egg weight estimation in the poultry industry.

2.1. EGG

Eggs are one of the most important and high-value poultry products in the global food industry. This commodity has a dual role as a source of high-quality animal protein and as an agribusiness commodity that supports household economies and food security [7]. Eggs contain a variety of nutrients that support metabolic health. For instance, eggs are a complete source of high-quality protein and contain 16 vitamins and minerals. Furthermore, eggs are cost efficient, the energy cost of eggs is significantly less when compared with that of other animal protein foods [8].

Global egg production has continued to increase significantly in the last two decades, driven by population growth, high demand for quality animal protein, and the adoption of modern poultry technologies such as closed-system coops and balanced nutritious feed [7]. This increase is also driven by the need for more environmentally friendly and efficient protein sources, in line with sustainable development efforts that position eggs as a nutritional alternative with a lower carbon footprint compared to red meat.

Egg weight, a critical quality parameter in egg production, is influenced by multiple factors such as hen age, genetics, and nutrition. Older hens tend to produce larger eggs due to more developed reproductive systems [9]. Genetic differences between breeds also play a key role, with heavier breeds laying larger eggs [10]. Moreover, dietary composition especially soluble fiber and enzyme supplementation can enhance nutrient absorption, leading to increased egg weight [11]. Therefore, efficiency in egg production, sorting, and distribution is an important aspect that continues to be developed through technological innovations, including the use of computer vision-based systems and artificial intelligence.

2.2. CNN

Convolutional Neural Network (CNN) is a deep learning architecture widely used in computer vision for automatically extracting patterns, textures, and colors through convolutional and pooling layers [12]. In numerous studies, CNN has proven to be superior for image-based classification and regression tasks, including weight estimation of objects such as livestock and fish [6], [13]. Models such as VGGNet, ResNet, and Inception are frequently used due to their strong

generalization capabilities on image data [14], and can be combined with transfer learning techniques to accelerate training and improve accuracy on limited datasets [15].
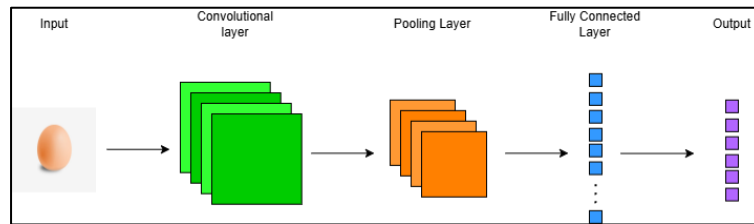

Figure 1. CNN Architecture (adapted from W.Taparduhee, et al. [13])

CNNs employ convolutional operations to acquire hierarchical representations directly from the input data, facilitating efficient feature extraction and pattern identification [16]. Based on figure 1, CNN works by extracting spatial information through a layered structure consisting of three main components, convolutional layers for feature extraction, pooling layers for dimensionality reduction and preventing overfitting, and fully connected layers that serve as the final classifier [12]. With this architecture, CNN is capable of hierarchically identifying complex patterns without the need for manual feature engineering, making it highly effective in various image-based applications, including the weight estimation of biological objects [13], [6].

2.3. Mask R-CNN

Mask R-CNN (Mask Regional Convolutional Neural Network) is one of the most influential deep learning models in the field of instance segmentation, combining object detection and pixel segmentation simultaneously within a unified framework [17]. This model is an extension of Faster R-CNN with an added dedicated branch for mask prediction, which allows for high-precision object contour extraction [18].

The main advantage of Mask R-CNN lies in its ability to separate individual objects even under conditions of high overlap and background complexity. Mask R-CNN has been successfully applied in various fields such as precision agriculture (apple flower detection and strawberry disease detection), medicine (blood cell analysis and skin lesion analysis), and even construction and manufacturing (rebar detection and particle segmentation) [19], [20].

In the context of biological and agricultural images, Mask R-CNN has proven effective in handling instance segmentation of natural objects with irregular shapes. Research conducted by [7] demonstrated Mask R-CNN's high performance in visually detecting and mapping diseases on strawberry fruits based on digital images [17]. Beyond segmentation accuracy, Mask R-CNN is also popular due to its flexibility to be used with modern backbone architectures such as ResNet and Feature Pyramid Network (FPN), which enables efficient multi-scale detection [21].

2.4. Extreme Gradient Boosting (XGBoost)

XGBoost is a highly efficient and accurate ensemble algorithm for regression and classification tasks, and is widely used in image-based estimation due to its superiority in handling high-dimensional data and correlated features [22]. In the context of image-based weight estimation, the XGBoost is employed to map extracted visual features (e.g., from CNN or morphometric techniques) into numerical values such as the weight or volume of biological objects [6].

One of XGBoost's main strengths lies in its ability to iteratively build predictive models by adding decision trees that correct the errors of previous models. Additionally, this algorithm possesses automatic feature selection capability based on gain or the feature's contribution to loss reduction, which makes it highly suitable for datasets resulting from complex image extraction [23]. In mathematical form, the objective function of the XGBoost for regression is [24]:

$$\text{Obj}(\theta) = \sum_{i=1}^{n} L\left(y_i, \hat{y}_i^{(t)}\right) + \sum_{k=1}^{t} \Omega(f_k) \tag{1}$$

L= Loss function (e.g., squared error: $(y_i - \hat{y}_i)^2$); $\hat{y}_i^{(t)} = \sum_{k=1}^{t} f_k(x_i)$: Cumulative prediction from tree – t; $\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_j^2$; T is the number of leaves; ω_j is the score on leaf j; γ, λ are regularization parameters

This formula reflects that XGBoost not only minimizes error but also controls model complexity to prevent overfitting.

In various recent studies, XGBoost has been widely used in the context of agriculture and food, such as estimating pig weight using features from body images [6], predicting crop yields using drone imagery [25], and assessing tuber quality with image texture features [26]. Its use is often combined with CNN or Mask R-CNN in hybrid schemes to improve estimation precision [23].

3. Methodology

This section elaborates on the methodology applied in the image-based egg weight estimation research using a combination of Mask R-CNN and XGBoost. The explanation begins with a general overview of the business process model, illustrating the overall system workflow. Subsequently, it specifically discusses the acquisition and characteristics of the dataset used, including the technical specifications for image capture to ensure consistent data quality. Finally, the crucial manual dataset labeling stage, which forms the foundation for the Mask R-CNN image segmentation process, is also explained.

3.1. Business Process Model

This study follows a sequential pipeline from data acquisition to inference. After collecting images and their GT weights, the dataset is partitioned at the egg level using GroupKFold with 10-folds. In this setup, all 36 views of a single egg are assigned exclusively to one-fold, ensuring that the same egg never appears simultaneously in both training and validation subsets. This grouped cross-validation strategy effectively prevents information leakage and provides a more robust and unbiased estimate of the model's generalization ability. Figure 2 summarizes the workflow.

The training set is used to fit a Mask R-CNN for instance segmentation, from which morphological, shape, and color features are extracted. A lightweight classifier predicts egg type (chicken vs duck) from color cues, and its output is added as a feature for regression. Finally, all features and the egg-type indicator are fed into an XGBoost regressor to estimate egg weight.
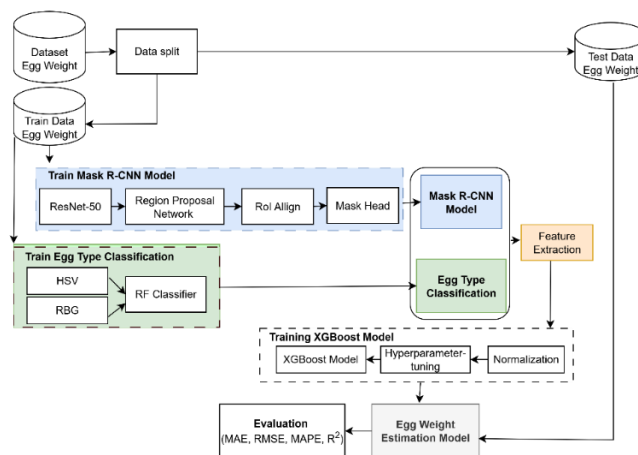
Figure 2. Business Process Model

Model performance is assessed on the held-out validation set using MAE (g), RMSE (g), MAPE (%), and R^2. The main segmentation settings for Mask R-CNN are reported in table 1, while the feature-based regression experiments and XGBoost hyperparameters are listed in table 2.

The parameters in table 1 are set to optimize Mask R-CNN training. MAX_ITER controls total iterations, DATALOADER.NUM_WORKERS sets parallel data loading, and SOLVER.IMS_PER_BATCH defines batch size. The initial learning rate (SOLVER.BASE_LR) decreases at steps set by SOLVER.STEPS and scaled by SOLVER.GAMMA. MODEL.ROI_HEADS.BATCH_SIZE_PER_IMAGE determines the number of region proposals sampled per image for RoI head training.

The model uses a ResNet-50 backbone with FPN for multi-scale feature extraction. BatchNorm and early layers (up to stage 2) are frozen for stable fine-tuning. Input images are randomly resized (shorter side: {640 to 800}, longer side ≤1333) for multi-scale training. Anchor sizes {32 to 512} enable the RPN to detect various object scales. ROIAlign is

applied at 7×7 (box) and 14×14 (mask) with zero sampling ratio for accurate interpolation. Data augmentation includes random flips, brightness/contrast/saturation tweaks, and small rotations (±5°) to improve generalization.

We also recorded computational footprints during training. On a Kaggle environment with a Tesla T4 GPU (≈14.7 GiB), each fold required about 1.8 to 2.3 hours of training time, averaging roughly 2.1 hours. Peak GPU memory usage typically ranged from 6 to 7 GiB, with occasional peaks up to 10.6 GiB, while CPU memory usage remained around 2.4 to 2.9 GiB. The final model checkpoints were about 335 MiB in size, containing roughly 43.9 million parameters ($\approx$ 43.7 million trainable). These values indicate that the chosen batch size, learning schedule, and augmentation strategy were well balanced against the available hardware, yielding stable training without excessive memory overhead.

Table 1. Configuration Mask R-CNN

parameter	configuration
MAX_ITER	3000
DATALOADER.NUM_WORKERS	8
SOLVER.IMS_PER_BATCH	8
SOLVER.BASE_LR	0.0001
SOLVER.STEPS	(2000, 2500)
SOLVER.GAMMA	0.1
MODEL.ROI_HEADS.BATCH_SIZE_PER_IMAGE	256

To evaluate the predictive performance of the XGBoost regressor, we applied a feature standardization step prior to training. All continuous predictors (morphology, shape descriptors, and color statistics) were standardized using z-score scaling (mean = 0, std = 1), while the binary indicator egg_type_pred_bin (0 = chicken, 1 = duck) was kept unscaled to preserve interpretability. This preprocessing improves numerical comparability across heterogeneous feature ranges and mitigates dominance of large-magnitude variables.

Table 2. XGBoost Experimental Setup

features used	XGBoost hyperparameter
n_estimators	100
learning_rate	0.01
max_depth	7
subsample	0.8
reg_alpha	0.01
min_child_weight	5
gamma	0.1
reg_lambda	0.1
objective	'reg:squarederror'
random_state	42
n_jobs	-1

The model hyperparameters (table 2) balance complexity, regularization, and efficiency: n_estimators controls the boosting rounds; learning_rate scales each tree's contribution; max_depth, min_child_weight, and gamma regularize tree growth; subsample regulates row sampling; reg_alpha (L1) and reg_lambda (L2) add further regularization; objective='reg:squarederror' sets the loss; random_state ensures reproducibility; and n_jobs enables parallelism.

3.2. Dataset

The dataset consists of primary images of chicken and duck eggs captured directly using a Samsung Galaxy A35 smartphone. Photos were taken under controlled conditions: 15 cm distance, consistent lighting from a 5W LED lamp, and the 50 MP wide lens (f/1.8, OIS) to reduce lighting variation and shadows. This low-cost setup ensures consistent image quality, though measurements remain pixel-based due to the lack of professional calibration tools.

Each egg was placed on a circular rotation template marked at 10° intervals, allowing systematic image capture from 36 viewpoints per egg. In the resulting images, the actual object length corresponds to approximately 307 to 314 pixels per centimeter, providing a practical scale reference even though the extracted features remain pixel-based. This procedure produced a comprehensive representation of each egg's geometry. In total, the dataset consists of 720 images (360 chicken eggs and 360 duck eggs). The GT weight of each egg was measured using a digital scale and stored as the regression target.

Representative samples of the dataset are shown in figure 3a for chicken eggs and figure 3b for duck eggs. As illustrated, chicken eggs generally exhibit a brownish hue with relatively smoother shell texture, whereas duck eggs display a greenish-white tint with subtle surface irregularities. These visual differences highlight the relevance of color and texture cues in the classification stage, which subsequently contributes as an auxiliary feature in the regression model.

Figure 3. Chicken Egg and Duck Egg

3.3. Labeling Dataset

Before conducting training with Mask R-CNN for segmentation, a crucial step involves manual dataset annotation. In this research, the LabelMe software was employed to annotate each egg object by carefully tracing its contour. The annotation was performed point by point to follow the natural curvature of the egg, ensuring accurate boundary representation. This high precision labeling is essential because segmentation quality directly influences the reliability of the extracted morphological, shape, and color features.

As illustrated in figure 4, each image was processed in LabelMe, where the egg boundary is outlined with a polygon mask highlighted in green. All annotation results were then exported and standardized into COCO format, which provides compatibility with Mask R-CNN training. In addition to segmentation masks, the dataset annotations were also enriched with the GT egg weight, enabling the direct linkage between image-based features and the regression target.
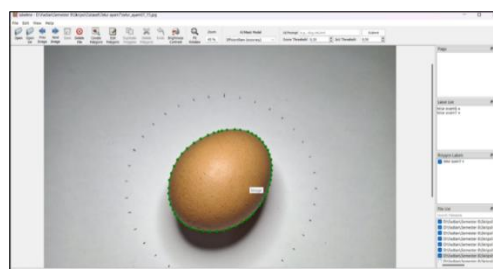

Figure 4. Labeling Dataset

This manual annotation procedure forms the foundation of the proposed system, as it guarantees that the segmentation model learns from precisely delineated object boundaries. Consequently, the downstream tasks of feature extraction, classification, and regression can be executed with higher accuracy.

4. Results and Discussion

The experimental results show that the proposed image-based egg weight estimation approach achieves reliable performance. Features were successfully extracted using Mask R-CNN, and the optimized XGBoost model demonstrated competitive prediction accuracy with low error margins.

4.1. Feature Extraction

After segmenting egg objects using the Mask R-CNN method, various features were extracted to capture the morphological, shape descriptor, and color characteristics of each object. The morphological features include bounding box dimensions (width, height, area_bbox, aspect ratio, centroid coordinates, diagonal length) and mask-derived descriptors (mask area, mask perimeter, extent, solidity, eccentricity, elongation, compactness, circularity, equivalent diameter, and hull perimeter). In addition, Hu moments (Hu1 to Hu7) were extracted as invariant descriptors for shape.

Color statistics were computed from the HSV color space (mean_h, mean_s, mean_v, std_h, std_s, std_v) and RGB color space (mean_r, mean_g, mean_b, std_r, std_g, std_b). These color features also served as the basis for egg type classification (chicken vs. duck), leveraging pigmentation differences in eggshells. Prior to regression, all continuous numerical features were standardized using the StandardScaler method to ensure zero mean and unit variance, while the binary egg type feature was used without scaling. The feature extraction results are presented in table 3.

Table 3. Feature Extraction

Feature extraction results									
egg_type	width	…	eccentricity	elongation	compactness	circularity	…	hu_7	weight
duck_egg	1108.802		0.192	1.213	14.515	0.866		9.999998654	61.9
duck_egg	1099.074		0.184	1.206	14.557	0.863		9.999957614	62.4
chicken_egg	949.823	…	0.275	1.323	14.706	0.855	…	-9.999912131	63.7
chicken_egg	969.945		0.210	1.231	14.559	0.863		9.9998401	60
chicken_egg	860.051		0.265	1.313	14.665	0.857		9.999935961	64.1

Hu moments (Hu1 to Hu7) are invariant shape descriptors that are highly sensitive to geometric structures. The values shown in the feature extraction results may appear extreme (for example, approaching ±9) or sometimes negative. This is normal because OpenCV computes Hu moments on a logarithmic scale, which compresses a very wide numeric range into values that often concentrate at the upper or lower bounds of floating-point precision. Negative values naturally arise from the logarithmic transformation applied to very small moments. These characteristics are not errors but reflect the mathematical nature of Hu moments, which are designed to remain invariant to object rotation, scale, and translation.

4.2. Results and Evaluation

4.2.1. Egg Type Classification

Before regression, egg type classification was performed using a Random Forest with color-based features, achieving perfect validation accuracy (precision, recall, and F1-score each 1.00). This is expected, as chicken and duck eggs differ clearly in color, making these features highly discriminative. The predicted label (egg_type_pred_bin) was then used as an additional input for regression. To prevent label leakage, the classifier was trained only on the training set, and predictions - not true labels - were used on the validation set to simulate test-time conditions.

4.2.2. XGBoost Regression

The final performance of the XGBoost regressor was evaluated on both the training and validation sets using the full standardized feature set, which also incorporated the predicted egg type from the Random Forest classifier. To ensure robustness, the experiments were conducted using a 10-fold grouped cross-validation strategy, where all 36 images of the same egg were kept in a single fold. The results for each fold are summarized in table 4, showing considerable variability across folds due to the small dataset size. Among these, the best performance was achieved on fold 6.

Table 4. Final performance of the XGBoost regressor

Fold	Dataset	MAE (g)	RMSE (g)	MAPE (%)	R²
1	Train	1.059	1.323	1.65%	0.6205
	Val	3.040	3.884	5.12%	-2.2629
2	Train	1.801	2.287	2.83%	0.1003
	Val	0.706	0.972	1.09%	-1.6262
3	Train	1.011	1.344	1.60%	0.6208
	Val	3.836	4.411	5.78%	-1.1623
4	Train	0.723	0.962	1.13%	0.8039
	Val	2.705	2.987	4.48%	-34.6844
5	Train	1.152	1.500	1.81%	0.6070
	Val	1.658	1.851	2.58%	-1.0270
6	Train	1.066	1.373	1.68%	**0.6781**
	Val	0.231	0.334	0.36%	**0.8012**
7	Train	0.702	0.959	1.11%	0.8258
	Val	2.020	2.192	3.11%	0.0914
8	Train	1.480	1.849	2.33%	0.4204
	Val	0.174	0.206	0.27%	0.6540
9	Train	1.036	1.401	1.64%	0.5476
	Val	3.720	3.883	5.53%	-11.4582
10	Train	1.212	1.535	1.91%	0.5990
	Val	1.571	1.610	2.49%	-40.4868

As shown in table 3, the model had lower absolute errors on the validation set. The Bland Altman plots (figure 5a and figure 5b) support this training predictions scatter widely with limits near ± 3 g, indicating underfitting, while validation results are tighter (± 0.7 g) and cluster around zero bias, showing more consistent performance on unseen eggs.
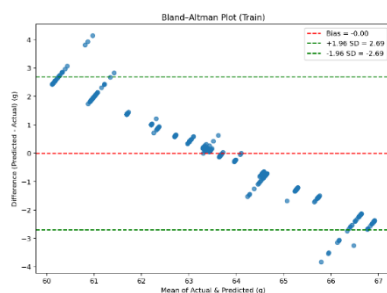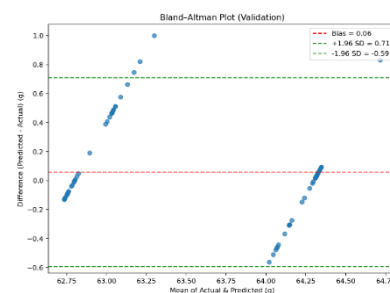

Figure 5a. Bland Altman Plot (Train)

Figure 5b. Bland Altman Plot (Val)

At first glance, it may seem unusual that validation outperforms training. This is mainly due to the GroupKFold setup, which places all images of the same egg in one-fold. As a result, the validation set often includes eggs with weights closer to the average-easier to predict-while the training set contains more variability, leading to higher errors. This reflects how dataset structure affects both performance metrics and patterns in the plots. As a baseline, the Support Vector Regressor (SVR) achieved strong training performance (MAE = 0.22 g, R^2 = 0.96) but failed to generalize (validation MAE = 0.65 g, R^2 = -0.14). In contrast, XGBoost produced lower validation errors and positive R^2, making it more reliable under grouped cross-validation.

Furthermore, an out-of-validation test was conducted using unseen egg samples that were not part of either the training or validation folds. The performance degraded substantially (MAE = 5.13 g, RMSE = 6.06 g, MAPE = 7.76%, R^2 = -0.09), confirming the limited generalization of the model to completely new eggs. This result highlights the need for larger and more diverse datasets, as well as external validation across different acquisition sessions, to ensure robustness in real-world deployment.

4.2.3. Loss Curve

During training, we tracked optimization in both segmentation and regression. For segmentation, the Mask R-CNN loss components steadily decreased over the first ~700 iterations and plateaued near zero by ~3,000 iterations (figure 6a), indicating stable convergence given the homogeneous dataset. However, training losses alone do not guarantee generalization and must be considered with validation results.
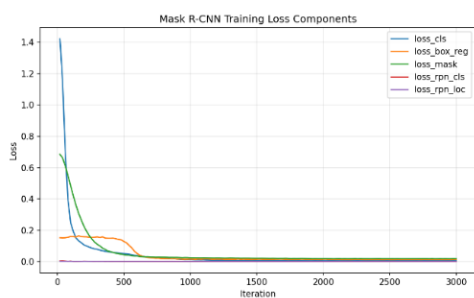

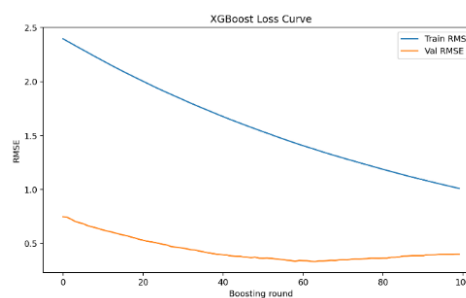**Figure 6a**. Loss curve mask r-cnn

Figure 6b. Loss curve xgboost

For regression, XGBoost's learning was tracked with RMSE on training and validation sets. Figure 6b shows training RMSE falling below 1 g, while validation RMSE dropped until around the 50th boosting round, then leveled at 0.4 to 0.6 g. This gap reflects the GroupKFold setup, where all 36 views of an egg stay in one-fold, so validation tests unseen eggs. The results show the model captures size cues well, but further gains may require larger datasets, stronger regularization, or physical feature calibration.

4.2.4. Feature Importance

The contribution of each feature to the predictive performance of the XGBoost model was analyzed using the gain metric. Figure 7 presents the top 20 most influential features.
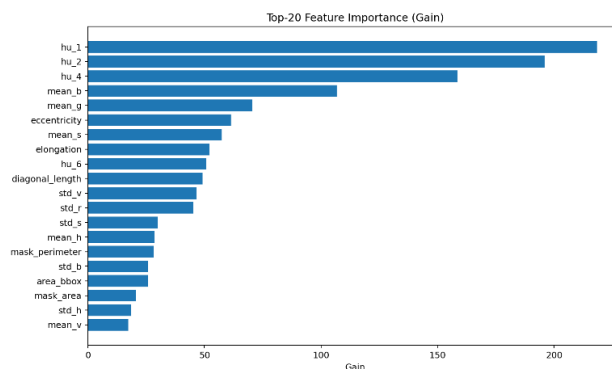

Figure 7. Feature importance

The results show that Hu moments particularly hu_1, hu_2, and hu_4 dominate the ranking, followed by several morphological and geometric descriptors such as eccentricity, elongation, and diagonal_length, along with (to a lesser extent) mask_perimeter, area_bbox, and mask_area. In addition, several color features are also highly ranked, especially the mean values of mean_b and mean_g (RGB channels) and mean_s (HSV), while color variation statistics (std_r, std_g, std_s, std_v, std_h, std_b) contribute at an intermediate level. The mean values of hue (mean_h) and value (mean_v) appear at the lower end of the top 20 features.

These findings indicate that egg geometry captured by Hu moments (rotation- and scale-invariant descriptors) and shape-related metrics such as eccentricity, elongation, and diagonal length is the primary determinant of weight estimation in this dataset. Color cues provide complementary information, likely helping to account for shell pigmentation differences (e.g., between chicken and duck eggs), thereby enhancing predictive performance even though they are less dominant than shape-related features.

5. Conclusion

This study developed an image-based egg weight estimation pipeline combining segmentation, classification, and regression. The Mask R-CNN segmented eggs to extract morphological, shape, and color features. The RF classifier was used for the classification of egg type and achieved perfect accuracy. The XGBoost regressor for generating the egg weight estimator model. Based on experimental results, the best model on training achieves the MAE = 1.07 g, R^2 = 0.68, and on validation achieves the MAE = 0.23 g, R^2 = 0.80. Bland-Altman plots showed validation predictions were more stable and less dispersed, due to the GroupKFold split and dataset structure. The feature importance highlighted shape descriptors, especially Hu moments, eccentricity, elongation, and diagonal length as most influential, with color statistics as complementary. A Support Vector Regressor baseline performed well on training but failed to generalize ($R^2 < 0$), confirming XGBoost's better balance of accuracy and generalization.

Further research needs to conduct robustness or ablation tests and t-tests or Wilcoxon Rank-Sum Test. Robustness or ablation tests are used to evaluate data sensitivity to segmentation errors, color shifts, or occlusions, or to measure performance changes when certain feature changes (e.g., Hu moments, color features, or morphological descriptors) are removed in grouped cross-validation. t-tests or Wilcoxon tests are used to confirm the significant effect of model performance improvements when feature changes occur.

6. Declarations

6.1. Author Contributions

Conceptualization: J.P., M.F.Z.A.R., A.P.S., R.M.M., and C.C.; Methodology: M.F.Z.A.R.; Software: J.P.; Validation: J.P., M.F.Z.A.R., and C.C.; Formal Analysis: J.P., M.F.Z.A.R., and C.C.; Investigation: J.P.; Resources: M.F.Z.A.R.; Data Curation: M.F.Z.A.R.; Writing Original Draft Preparation: J.P., M.F.Z.A.R., and C.C.; Writing Review and Editing: M.F.Z.A.R., J.P., and C.C.; Visualization: J.P.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] N. Yang, "Egg Production in China: Current Status and Outlook," *Frontiers of Agricultural Science and Engineering*, vol. 8, no. 1, pp. 25–34, Feb. 2021, doi: 10.15302/J-FASE-2020363.

[2] J. P. B. López Vargas, K. L. de Abreu, D. Duarte de Paula, D. H. Pinheiro Salvadeo, L. F. Arantes de Souza, and C. Bôa-Viagem Rabello, "Alternative Non-Destructive Approach for Estimating Morphometric Measurements of Chicken Eggs from Tomographic Images with Computer Vision," *Foods*, vol. 13, no. 24, pp. 1-12, 2024, doi: 10.3390/foods13244039.

[3] F. Akkoyun, A. Ozcelik, I. Arpaci, A. Erçetin, and S. Gucluer, "A Multi-Flow Production Line for Sorting of Eggs Using Image Processing," *Sensors*, vol. 23, no. 1, pp.1-12, 2022, doi: 10.3390/s23010117.

[4] X. Yang, R. B. Bist, S. Subedi, and L. Chai, "A Computer Vision-Based Automatic System for Egg Grading and Defect Detection," *Animals*, vol. 13, no. 14, pp.1-12, 2023, doi: 10.3390/ani13142354.

[5] A. I. Khan and S. Al-Habsi, "Machine Learning in Computer Vision," *Procedia Computer Science*, vol. 167, no.1, pp. 1444–1451, 2020, doi: 10.1016/j.procs.2020.03.355.

[6] S. Jiang, G. Zhang, Z. Shen, P. Zhong, J. Tan, and J. Liu, "Pig Weight Estimation Method Based on a Framework Combining Mask R-CNN and Ensemble Regression Model," *Animals*, vol. 14, no. 1, pp. 21-32, 2024, doi: 10.3390/ani14142122.

[7] E. F. Guèye, "Trends and prospects of poultry value chains in Africa," *Journal of Agriculture, Science and Technology*, vol. 23, no. 4, pp. 19–46, May 2024, doi: 10.4314/jagst.v23i4.2.

[8] S. Walker and J. I. Baum, "Eggs as an affordable source of nutrients for adults and children living in food-insecure environments," *Nutr. Rev.*, vol. 80, no. 5, pp. 1068–1080, 2022, doi: 10.1093/nutrit/nuab019.

[9] A. Marzec, K. Damaziak, H. Kowalska, J. Riedel, M. Michalczuk, and E. Koczywas, et al., "Effect of hens age and storage time on functional and physiochemical properties of eggs," *J. Appl. Poult. Res.*, vol. 28, no. 2, pp. 290–300, 2019, doi: 10.3382/japr/pfy069.

[10] T. Goto, H. Mori, S. Shiota, and S. Tomonaga, "Metabolomics approach reveals the effects of breed and feed on the composition of chicken eggs," *Metabolites,* vol. 9, no. 1, pp. 224-245, 2019, doi: 10.3390/metabo9100224.

[11] X. H. Nguyen, H. T. Nguyen, and N. K. Morgan, "Dietary soluble non-starch polysaccharide level and xylanase supplementation influence performance, egg quality and nutrient utilization in laying hens fed wheat-based diets," *Anim. Nutr.,* vol. 7, no. 2, pp. 512–520, 2021, doi: 10.1016/j.aninu.2020.05.012.

[12] L. Benos, A. C. Tagarakis, G. Dolias, R. Berruto, D. Kateris, and D. Bochtis, "Machine learning in agriculture: A comprehensive updated review," *Sensors*, vol. 21, no.1, pp. 3758-3769, 2021, doi: 10.3390/s21113758.

[13] W. Taparhudee, R. Jongjaraunsuk, S. Nimitkul, P. Suwannasing, and W. Mathurossuwan, "Optimizing Convolutional Neural Networks, XGBoost, and Hybrid CNN-XGBoost for Precise Red Tilapia (Oreochromis niloticus Linn.) Weight Estimation in River Cage Culture with Aerial Imagery," *AgriEngineering*, vol. 6, no. 2, pp. 1235–1251, Jun. 2024, doi: 10.3390/agriengineering6020070.

[14] A. Sharma, A. Jain, P. Gupta, and V. Chowdary, "Machine learning applications for precision agriculture: A comprehensive review," IEEE Access, vol. 9, no.1, pp. 4843–4873, 2021, doi: 10.1109/ACCESS.2020.3048415.

[15] Z. Unal, "Smart farming becomes even smarter with deep learning: A bibliographical analysis," *IEEE Access*, vol. 8, no.1, pp. 105587–105609, 2020, doi: 10.1109/ACCESS.2020.3000175.

[16] N. Ashfaq, M. H. Khan, and M. A. Nisar, "Identification of Optimal Data Augmentation Techniques for Multimodal Time-Series Sensory Data: A Framework," *Information,* vol. 15, no. 6, pp. 343-362, Jun. 2024, doi: 10.3390/info15060343.

[17] U. Afzaal, B. Bhattarai, Y. R. Pandeya, and J. Lee, "An instance segmentation model for strawberry diseases based on mask R-CNN," *Sensors,* vol. 21, no. 19, pp. 1-15, Sep. 2021, doi: 10.3390/s21196565.

[18] O. L. F. de Carvalho, O. A. de Carvalho Júnior, A. O. de Albuquerque, P. P. de Bem, C. R. Silva, P. H. G. Ferreira, R. d. S. de Moura, R. A. T. Gomes, R. F. Guimarães, and D. L. Borges, "Instance segmentation for large, multi-channel remote sensing imagery using Mask-RCNN and a mosaicking approach," *Remote Sens.*, vol. 13, no. 1, pp. 1-39, 2021, doi: 10.3390/rs13010039.

[19] R. W. Bello, A. S. A. Mohamed, and A. Z. Talib, "Contour extraction of individual cattle from an image using enhanced Mask R-CNN instance segmentation method," *IEEE Access*, vol. 9, no. 1, pp. 56984–57000, 2021, doi: 10.1109/ACCESS.2021.3072636.

[20] M. Frei and F. E. Kruis, "Image-Based Analysis of Dense Particle Mixtures via Mask R-CNN," *Eng*, vol. 3, no. 1, pp. 78–98, Jan. 2022, doi: 10.3390/eng3010007

[21] C. Xu, C. Shi, H. Bi, C. Liu, Y. Yuan, and H. Guo, et al., "A page object detection method based on Mask R-CNN," *IEEE Access*, vol. 9, no.1, pp. 143448–143457, 2021, doi: 10.1109/ACCESS.2021.3121152.

[22] Y. Shen, Z. Yan, Y. Yang, W. Tang, J. Sun, and Y. Zhang, "Application of UAV-borne visible-infrared pushbroom imaging hyperspectral for rice yield estimation using feature selection regression methods," *Sustainability*, vol. 16, no. 1, pp. 632-647, 2024, doi: 10.3390/su16020632.

[23] X. Ji, Y. Zhou, C. Du, L. Chen, H. Zhang, and Z. Zhang, et al., "A machine learning-based method for pig weight estimation and the PIGRGB-weight dataset," *Agriculture*, vol. 15, no.1, pp. 814-831, 2025, doi: 10.3390/agriculture15080814.

[24] W. Anggraeni, H. N. Pradani, S. Sumpeno, E. M. Yuniarno, R. F. Rachmadi, and P. Pujiadi, et al., "Prediction of dengue fever outbreak based on climate and demographic variables using extreme gradient boosting and rule-based classification," *Proc. 2021 IEEE 9th Int. Conf. Serious Games Appl. Health (SeGAH),* vol. 2021, no. 1, pp. 1–8, 2021, doi: 10.1109/SEGAH52098.2021.9551900.

[25] P. Zhang, B. Lu, J. Shang, X. Wang, Z. Hou, and S. Jin, "Ensemble learning for oat yield prediction using multi-growth stage UAV images," *Remote Sens.*, vol. 16, no.1, pp. 45-75, 2024, doi: 10.3390/rs16234575.

[26] J. Nakatumba-Nabende, C. W. Sseruwu, P. R. Rubaihayo, S. Agona, Y. Baguma, and R. Tumuhimbise, et al., "Using machine learning for image-based analysis of sweetpotato root sensory attributes," *Smart Agric. Technol.*, vol. 5, no. 1, pp. 1-21, 2023, doi: 10.1016/j.atech.2023.100291.