

Article

Impact of Trait Measurement Error on Quantitative Genetic Analysis of Computer Vision-Derived Traits

Ye Bi ^{1,2} , Yijian Huang ³ , Haipeng Yu ⁴  and Gota Morota ^{1,5,*} 

¹ School of Animal Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA; yebi@iastate.edu

² Department of Animal Science, Iowa State University, Ames, IA 50010, USA

³ Smithfield Premium Genetics, Rose Hill, NC 28458, USA; yhuang@smithfield.com

⁴ Department of Animal Sciences, University of Florida, Gainesville, FL 32611, USA; haipengyu@ufl.edu

⁵ Laboratory of Biometry and Bioinformatics, Department of Agricultural and Environmental Biology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo 113-8657, Japan

* Correspondence: gmorota@g.ecc.u-tokyo.ac.jp

Abstract

Background: Quantitative genetic analysis of image- or video-derived phenotypes is increasingly being performed for a wide range of traits. Pig body weight values estimated by a conventional approach or a computer vision system can be considered two different measurements of the same trait but with different sources of phenotyping error. Previous studies have shown that trait measurement error, defined as the difference between manually collected phenotypes and image-derived phenotypes, can be influenced by genetics, suggesting that the error is systematic rather than random and is more likely to lead to misleading quantitative genetic analysis results. Therefore, we investigated the effect of trait measurement error on the genetic analysis of pig body weight (BW). **Results:** Calibrated scale-based and image-based BW showed high coefficients of determination and goodness of fit. Genomic heritability estimates for scale-based and image-based BW were mostly identical across growth periods. Genomic heritability estimates for trait measurement error were consistently negligible, regardless of the choice of computer vision algorithm. In addition, genome-wide association analysis revealed no overlap between the top markers identified for scale-based BW and those associated with trait measurement error. Overall, the deep learning-based regressions outperformed the adaptive thresholding segmentation methods. **Conclusion:** This study showed that manually measured scale-based and image-based BW phenotypes yielded the same quantitative genetic results. We found no evidence that BW trait measurement error could be influenced, at least in part, by genetic factors. This suggests that trait measurement error in pig BW does not contain systematic errors that could bias downstream genetic analysis.

Keywords: body weight; camera; genome-wide association analysis; heritability; phenomics; swine



Received: 28 March 2026

Revised: 20 April 2026

Accepted: 21 April 2026

Published: 24 April 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The power of quantitative genetic analysis largely depends on the quality and quantity of available phenotypes. Computer vision approaches are rapidly being developed to generate phenotypes for genetic analysis in livestock [1]. Computer vision coupled with cameras is one of the precision livestock technologies to realize precision livestock farming, which aims to automate phenotyping or monitoring animals at the individual level on

a continuous basis [2]. An advantage of computer vision systems is that they can be installed in a barn with minimal disruption to daily farm operations and are relatively inexpensive, requiring only one or a few off-the-shelf consumer cameras [3]. Pedigree-based or genomic-based genetic analysis of image- or video-derived traits is increasingly being performed on growth or morphological traits [4–8], carcass traits [9–13], behavior [5,14] and feed intake [7]. Computer vision is the core engine for phenotyping these live animal traits, except for carcass traits, which are collected after slaughter.

Computer vision is not necessarily always required to collect the above phenotypes, but traditional approaches often require additional hours of manual workload or the purchase of expensive equipment that requires careful maintenance. For example, collecting growth or morphological traits requires a digital scale with manual labor to restrain the animals or a costly walk-over weighing system. Similarly, manually monitoring individual feed intake requires constant observation and estimation of measurements by humans. Collecting behavioral traits requires many hours of behavioral video decoding by ethologists based on a formal ethogram. What distinguishes computer vision-based phenotyping from conventional phenotyping is the ability of the former to scale up and perform high-throughput phenotyping by automating all or part of the phenotyping procedures.

Consider a scenario where we are interested in phenotyping a specific target trait. This target trait, collected using conventional and image- or video-based approaches, can be considered two different measurements of the same character because they have different phenotyping errors. In other words, two measurements of the target trait are controlled by the same set of causal loci, but the source of phenotyping error is different. Then, we can calculate a trait measurement error, which is defined as the difference between the conventional phenotypes and the image-derived phenotypes. Because these two measurements of the same trait have the same genetic architecture, it is assumed that trait measurement errors are random and nongenetic factors (Figure 1). However, previous research has shown that trait measurement error can be influenced, at least in part, by genetic factors. For example, maize biomass was quantified using manual measurements and image-based estimates [15]. Although they obtained the Pearson correlation coefficient of 0.95 between the manual measurements and image-based estimates, 58% of the trait measurement error in biomass was explained by genetic variation. Similarly, 27–59% of the trait measurement error variance was controlled by genetic variation using single-nucleotide polymorphism (SNP) data in sorghum height; the Pearson correlation coefficient between manual measurements and image-based estimates was 0.96 [16]. They also identified a few significant genomic regions when trait measurement errors were used as phenotypes in genome-wide association studies (GWAS). These results suggest that phenotypes of some individuals are systematically underestimated, and phenotypes of some other individuals are systematically overestimated. Because trait measurement error is systematic rather than random in these cases, if these image-derived phenotypes were used in quantitative genetic analysis, it could lead to misleading results.

This type of problem can also occur in livestock. For example, the live body weight (BW) of animals is often estimated from a single top-view depth sensor camera [17,18]. Height and volume information that can be estimated from a depth sensor camera is known to be highly correlated with scale-based BW in animals [19,20]. When estimating volumetric measures from a depth sensor camera, it is assumed that the distance from the ground to the ventral abdomen is constant for all animals. However, this assumption is not valid if some sires or boars carry genetics that increase the distance from the ground to the ventral abdomen in their offspring, resulting in a systematic error in image-derived BW estimates. Although an increasing number of animal phenotypes are generated by computer vision, the effect of trait measurement error on quantitative genetic analysis

has not been investigated. In addition, it is unknown how the choice of computer vision algorithms can influence trait measurement error in downstream genetic analysis.

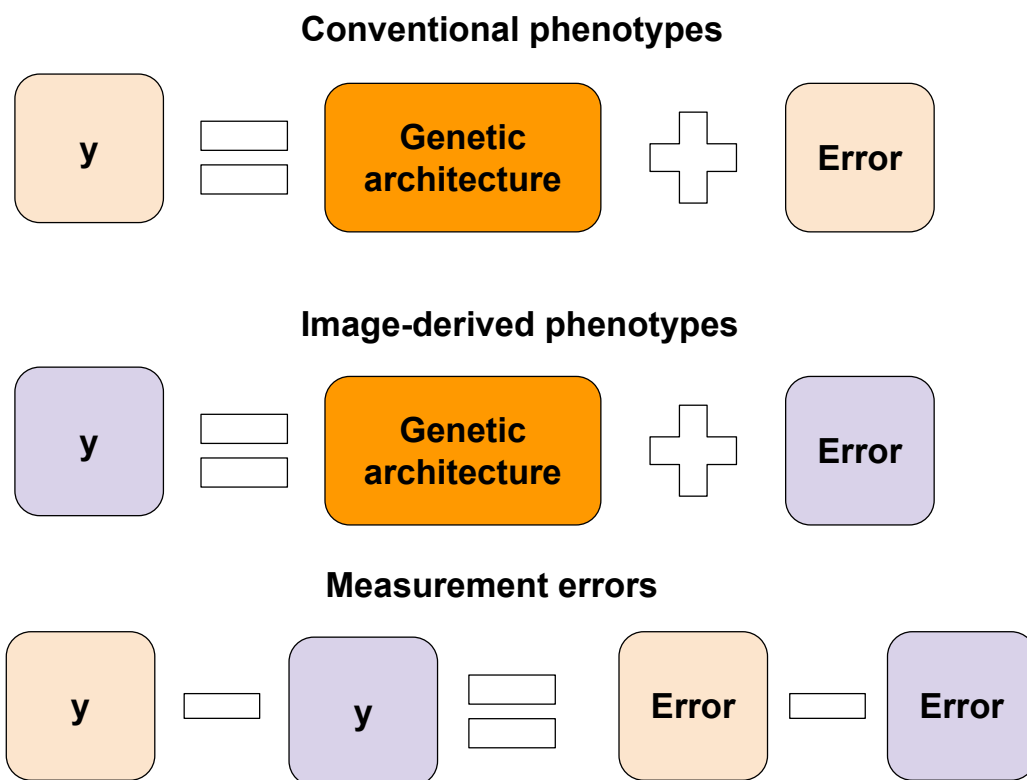


Figure 1. Concept of trait measurement error when using computer vision to generate image-derived phenotypes.

Therefore, in this study, we investigated the impact of trait measurement error on the genetic analysis of pig BW. Specifically, the objectives of this study were (1) to estimate trait measurement error from manually collected and image-derived pig BW, (2) to evaluate how the choice of computer vision algorithms affects the genetic analysis of trait measurement error, and (3) to estimate genomic heritability and marker effects for trait measurement error.

2. Materials and Methods

2.1. Animals and Data Collection Design

A subset of pig data analyzed in Bi et al. [20,21] was used in this study. In short, we analyzed data from more than 600 crossbred pigs raised on a farm in North Carolina, USA. The pigs were part of a commercial finishing operation at Smithfield and were regularly weighed during four farm visits or time points (T1 to T4) between July and September 2023. The number of pigs phenotyped varied at each visit, but it decreased at later time points, primarily due to market sales. Animals were fed a common diet and had unrestricted access to water. Pigs were weighed individually on a digital walk-in calibrated scale to ensure they were restrained and properly positioned for accurate measurements. Their BW ranged from 25 to 154 kg. The average BW (kg) of pigs at T1, T2, T3, and T4 was 65 (103), 78 (117), 92 (131), and 114 (159), respectively, where values in parentheses indicate average age (days).

2.2. Video Acquisition Setup

A depth video acquisition setup is detailed in our previous work [19–21]. Briefly, an Intel RealSense D435 camera (Intel, Santa Clara, CA, USA) connected via USB 3.1 was strategically positioned 1.40 m above a one-way exit lane equipped with a stand-on

weighing system to capture top-view video. The camera's dual stereo image sensors, with a horizontal field of view of 87° and a vertical field of view of 58°, enabled depth estimation. Individual pigs were manually restrained to facilitate the collection of 10–12 s video segments at a frame rate of 30 frames per second and a resolution of 1280 × 720. While the pigs exhibited general stability in their body movements, their heads and tails exhibited a comparatively dynamic range of motion. Simultaneously with the video recordings, the pigs' BW was measured using a weight scale, and individual identification was established by ear tag recognition.

2.3. Depth Data Processing and Quality Control

The raw depth data captured in .bag video format was converted to PNG depth images and CSV depth map files using the rs-convert tool [22]. Within each depth image, different colors represent different distances from the camera to the object, providing a visual representation of spatial depth. Each image was accompanied by a CSV file of identical dimensions, providing distance measurements in meters for each pixel.

To ensure the quality of the videos collected in the commercial environment for subsequent image analysis, we applied a customized YOLOv8s model [23] during the preprocessing stage. We first annotated 100 depth images using Roboflow [24], labeling two categories: "pig" (representing the main body of the pig) and "block" (representing occlusions caused by farm technicians obstructing the camera view). After training the YOLOv8s model on these annotations, we detected objects within the depth images. Images were excluded if the block detection rate was greater than 0.1, the pig detection rate was less than 0.5, or no pig was detected. After preprocessing, the final dataset contained 20,356, 45,814, 44,525, and 23,968 depth images for T1 to T4, respectively.

2.4. Adaptive Thresholding Segmentation

The first computer vision approach to derive image-based BW from depth images was adaptive thresholding segmentation followed by a regression model, as developed in a previous study for dairy cattle [19] and recently applied. This approach extracts the biometric measurements of pigs and uses them to build regression models for BW prediction.

The OpenCV library (version 4.7.0) was used in Python (version 3.8.16) for the thresholding method. First, the depth images were cropped to eliminate extraneous surroundings. In hue, saturation, and value (HSV) color mode, the hue channel was extracted as a threshold for converting the images to binary format, where hues above the threshold were converted to black, and those below were converted to white. An adaptive range from 15 degrees below to 5 degrees above the mean of the hue was applied. In general, lower thresholds were preferred because they captured more of the animal's body. To determine the optimal threshold, we used the OpenCV function *boundingRect* to draw a bounding box around the largest contour. If the four corners of the bounding box were less than five pixels from the edge, the hue threshold was increased by one unit until the four corners were more than five pixels from the edge of the image. The OpenCV function *findcontour* was then used to find the contour with the largest area. Finally, the OpenCV function *minAreaRect* was used to draw the minimum bounding rectangle of the pig's body contour. Due to the significant movement of the pig's head and tail during the video data acquisition, we excluded these areas from the depth images. The ratio of body width to body length was used to remove the neck and tail horizontally and vertically within each image. Finally, the contour was identified, and the bounding box was drawn on the binary image. The bounding box and depth information associated with the identified contour were retained for BW regression analysis.

Four biometric features were extracted from each image: dorsal length, abdominal width, height, and volume. Dorsal length and abdominal width were estimated by drawing the minimum area bounding box around the pig's body using the OpenCV function *minAreaRect*. Height was determined using an average-based approach where missing data (zeros) within the contour pixels in the depth CSV file were replaced with the average distance. The average height of the pig was calculated as 1.40 m minus the average distance from the camera to the pig contour. Volume was calculated by summing the height values of pixels within the pig contour. After all images were processed, the medians of dorsal length, abdominal width, height, and volume were calculated to produce a single final estimate for each video. These image-derived features were then used as predictors to build regression models for predicting pig body weight. Ordinary least squares (OLS) and random forest (RF) models were applied using the caret R package [25]. To improve the performance of the RF model and reduce the risk of overfitting, 5-fold cross-validation was performed within the training dataset.

2.5. Deep Learning Models

The second computer vision approach used to estimate BW was deep learning models. The main idea is to directly predict pig BW by feeding the entire depth image to neural network models, as shown in recent studies [20,21]. Deep learning models integrate feature extraction and regression into a single-step process. A brief description is given below, as details were given in a recent study [20].

We first converted depth CSV files containing distance information in meters to depth grayscale images to ensure consistent distance representation. This conversion was done using the *pyplot* function in the matplotlib library (version 3.7.1) in Python. The grayscale images, originally 1270×720 pixels in size, were resized and padded to a square dimension of 150×150 . We evaluated the feature extraction capabilities of two deep learning model families using transfer learning: MobileViT and ResNet50, both pre-trained on ImageNet. Modifications included removing the classifier layer, retaining the backbone, adding a 50% dropout layer, and incorporating a single-node dense layer with linear activation and L2 regularization for BW prediction. The main architectures are summarized below.

MobileViT combines the strengths of convolutional neural networks and transformers by integrating spatial inductive biases with global processing capabilities [26]. It uses a hybrid block that applies standard and pointwise convolutions to capture local features, transforms these into non-overlapping patches for long-range dependency modeling via transformers, and then re-folds and fuses features with the input tensor. We evaluated the small MobileViT architecture, MobileViT-S. ResNet50 was used as the baseline model. It uses residual blocks and applies an identity shortcut connection to skip one or more layers to mitigate the degradation problem, allowing for easier optimization and increased accuracy with depth [27]. MobileViT is expected to provide greater efficiency than ResNet50 by leveraging a hybrid architecture that integrates convolutional layers for local feature extraction with transformer layers for capturing long-range dependencies, resulting in a more parameter-efficient model. In contrast, ResNet50 relies on residual connections to improve optimization and accuracy with depth, but its architecture generally requires more computational resources as model depth increases.

The deep learning models were implemented using TensorFlow GPU 2.14.0 and Keras 2.14.0 in Python 3.9.0, running on Virginia Tech's Advanced Research Computing cluster with two 80 GB NVIDIA DGX A100 GPUs. Models were trained on 150×150 images with a batch size of 100 for 300 epochs, using mean squared error as the loss function. The *ReduceLROnPlateau* callback dynamically reduced the learning rate by 0.1 when validation loss stagnated for 75 epochs, with a minimum threshold of 0.000001, while *ModelCheckpoint*

saved the best model based on validation loss. Validation loss was the criterion for selecting the optimal model parameters.

2.6. Goodness of Fit Metrics

The coefficient of determination (R^2 , Equation (1)) and the mean absolute percentage error (MAPE, Equation (2)) were used to evaluate the performance of the computer vision models.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (1)$$

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \quad (2)$$

where y_i is the scale-based BW of the i th animal, $\hat{y}_i$ is the image-based predicted BW of the i th animal, and $\bar{y}$ is the average scale-based BW.

2.7. Quantitative Genetic Modeling

All pigs were genotyped using 50K SNP chips. After quality control by removing markers with minor allele frequencies less than 0.05, a total of 46,407 to 46,424 SNPs remained for analysis, depending on different growth periods and computer vision approaches used, because the number of available pigs differed.

Genomic heritability estimates of scale-based BW, image-based BW, and trait measurement error BW, defined as their absolute differences, were obtained using the restricted maximum likelihood method [28] in the genomic best linear unbiased prediction (GBLUP) model [29]. The GBLUP model used was

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{u} + \boldsymbol{\epsilon}, \quad (3)$$

where $\mathbf{y}$ is a vector of phenotypes; $\mathbf{1}$ is the vector of ones; μ is the overall mean; $\mathbf{X}$ is the incidence matrix of fixed effects; $\mathbf{b}$ is a vector of fixed effects; $\mathbf{Z}$ is the incidence matrix relating animals to their phenotypic records; and $\mathbf{u}$ is a vector of the random additive genetic values of animals; and $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \mathbf{I}\sigma_{\epsilon}^2)$, is a vector of residuals. We assumed $\mathbf{u} \sim N(\mathbf{0}, \mathbf{G}\sigma_u^2)$, where $\mathbf{G} = \frac{\mathbf{W}\mathbf{W}'}{m}$ is the genomic relationship matrix; σ_u^2 is the additive genetic variance; $\mathbf{W}$ is a centered and standardized SNP marker matrix; and m is the total number of SNP markers. The fixed effects included sex, birth farm, pen density, age, and parity of dam.

We used the following single-marker GWAS [30,31] for scale-based BW, image-based BW, and their trait measurement errors.

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}\mathbf{b} + \mathbf{w}_j\mathbf{s}_j + \mathbf{Z}\mathbf{u} + \boldsymbol{\epsilon}, \quad (4)$$

where $\mathbf{w}_j$ is the j th SNP marker to be tested, and $\mathbf{s}_j$ is a vector of fixed j th SNP effect. The remaining terms are identical to those in the GBLUP model. All models were fitted using the rrBLUP R package [32]. Since our main interest is whether we can correctly detect SNP markers with large effects, we evaluated whether the top 10 markers detected in scale-based BW were among the top 100 markers in image-based BW or trait measurement error BW using $-\log_{10} p$ -values. We hypothesized that among a total of SNPs, almost all of the top 10 markers detected in scale-based BW would also be correctly detected in the top 100 markers of image-based BW, whereas none of the top 10 markers detected in scale-based BW would be detected in the top 100 markers of trait measurement error BW. We focused on SNP markers with top GWAS peaks because it is typically not of interest to correctly identify the ranking of SNPs with small effects. All genetic analyses were

performed at each growth period to investigate whether there are specific time points at which trait measurement error is genetically influenced.

3. Results

Phenomic analysis with deep regression resulted in more pigs with phenotypes for genetic analysis than thresholding segmentation (Table 1). This is because the adaptive thresholding approach was unable to process some pigs with noisy image data, whereas deep learning-based regression was robust to variable image quality and was able to process noisy image data. The amount of trait measurement error varied between pigs at different time points. Figure 2 shows pigs that had the highest trait measurement errors at at least two time points and the lowest trait measurement errors at least two time points. The results indicate that some pigs had consistently high trait measurement errors over time, while other pigs had consistently low trait measurement errors.

Table 1. Number of pigs with image-derived phenotypes produced by two computer vision approaches. T1 to T4 denote four farm visits.

Time	Thresholding Segmentation	Deep Regression
T1	475	505
T2	465	523
T3	521	528
T4	312	311

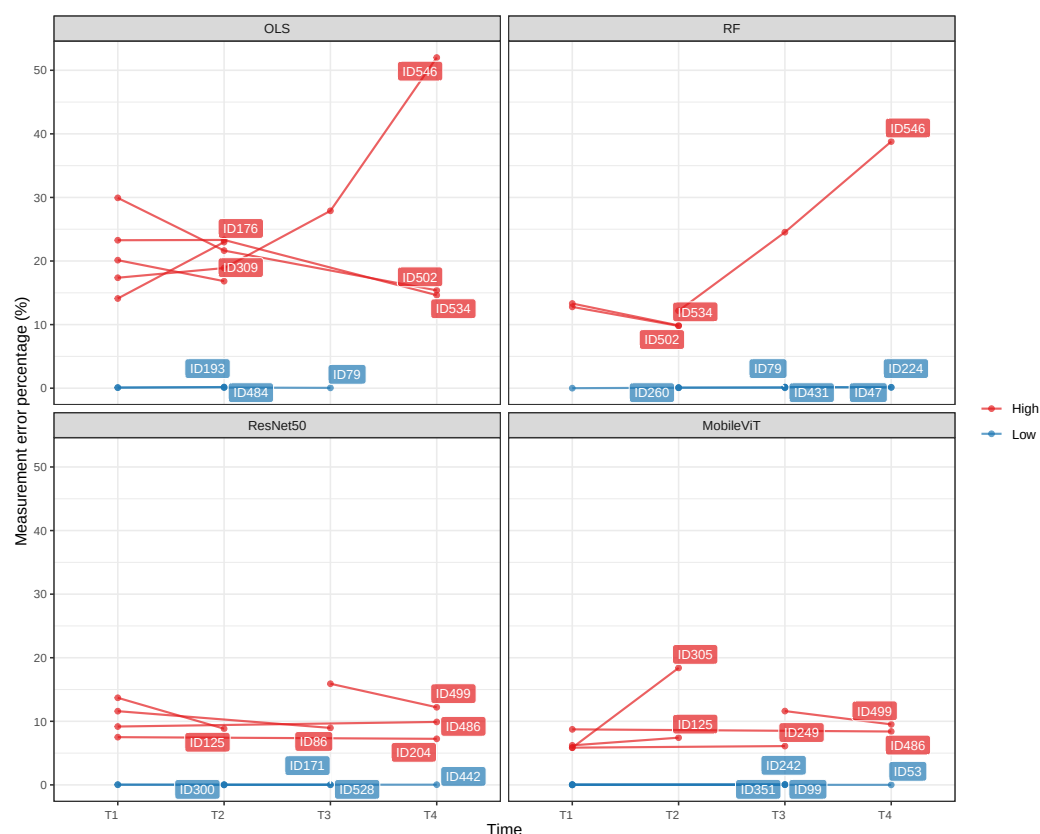


Figure 2. Illustration of pigs that had the highest trait measurement errors on at least two time points and the lowest trait measurement errors on at least two time points.

Image-derived BW values were estimated by training models on video data captured by a depth camera. A total of four image analysis models were used, including adaptive thresholding segmentation coupled with either OLS or RF and deep regression based on

ResNet50 and MobileViT. Coefficients of determination for image-derived and scale-based body BW were at least 0.95 for RF-based adaptive thresholding segmentation approaches and deep regression across different time points (Table 2), showing strong agreement. Similarly, the MAPE was always less than 2.5% for the RF-based adaptive thresholding segmentation approaches and the deep regressions. Adaptive thresholding segmentation with OLS resulted in consistently lower R^2 and higher MAPE than the other approaches. T3, which had the largest number of pigs available, showed the best performance for deep regression. Trait measurement errors expressed as percentages showed that deep regression provided the lowest errors across all growth periods, followed by RF- and OLS-based adaptive thresholding segmentation (Figure 3). The mean trait measurement error percentages were 4.57%, 2.13%, 1.64%, and 1.61% for OLS-based adaptive thresholding segmentation, RF-based adaptive thresholding segmentation, ResNet50, and MobileViT, respectively.

Table 2. Coefficient of determination and mean absolute percentage error between scale-based body weight and image-derived body weight in the pig data. T1 to T4 denote four farm visits.

Time	Thresholding Segmentation				Deep Regression			
	OLS ^a		RF ^b		ResNet50		MobileViT	
	R^2 ^c	MAPE ^d	R^2	MAPE	R^2	MAPE	R^2	MAPE
T1	0.72	5.33	0.95	2.45	0.93	1.87	0.94	1.68
T2	0.84	4.29	0.97	1.95	0.95	1.54	0.96	1.44
T3	0.82	4.24	0.96	1.96	0.97	1.44	0.98	1.27
T4	0.80	4.57	0.95	2.30	0.96	1.44	0.97	1.45

^a Ordinary least squares. ^b Random forests. ^c Coefficient of determination. ^d Mean absolute percentage error.

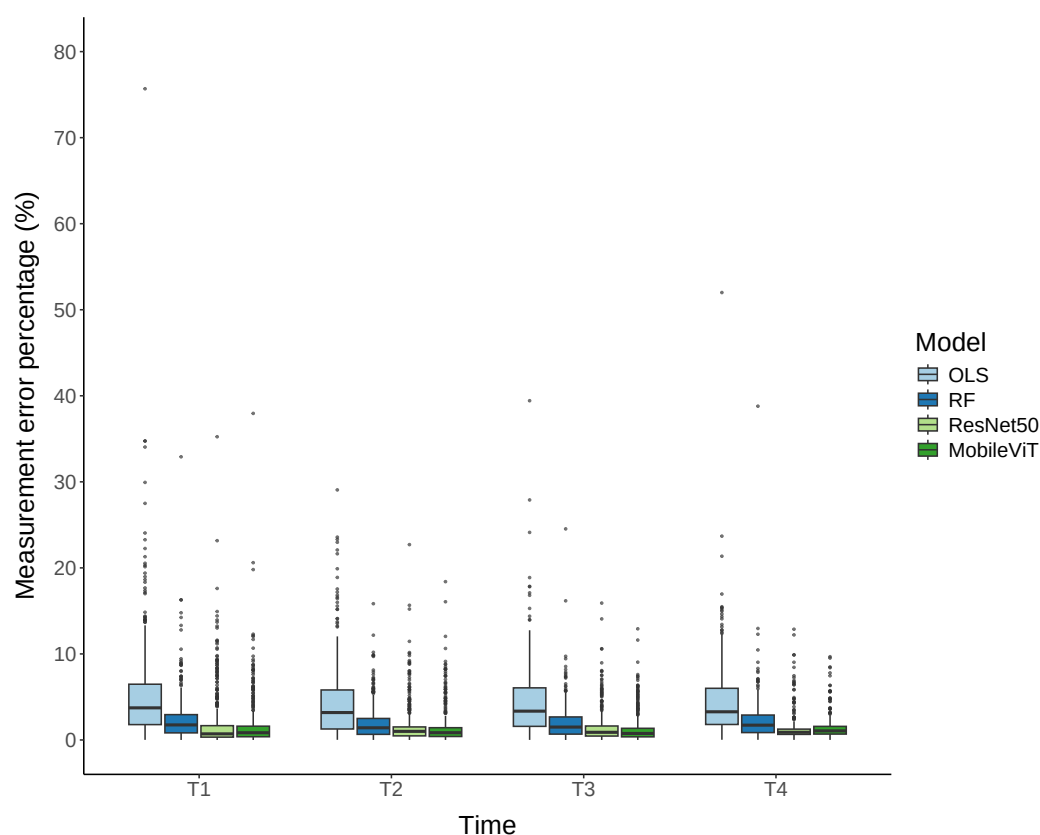


Figure 3. Trait measurement error percentages of four computer vision approaches across four data collection points. OLS: ordinary least squares. RF: Random forest.

Genomic heritability estimates for scale-based BW ranged from 0.21 to 0.55 across time points and numbers of pigs available (Table 3). These estimates were within the range reported in a recent study, which found genomic heritability estimates ranging from 0.22 to 0.83 across growth periods [33]. Genomic heritability estimates for manual and image-based BW were mostly identical across time points, indicating that manually measured and image-based phenotypes yield the same genetic inference results. Agreement was particularly evident for the deep regression approaches, with differences in genomic heritability estimates ranging from 0 to 0.03. On the other hand, genomic heritability estimates for BW obtained from the OLS-based adaptive segmentation approach differed most from those of scale-based BW. Genomic heritability estimates for measurement error were consistently close to zero and negligible.

Table 3. Genomic heritability estimates (standard errors) for scale-based body weight, image-based body weight, and trait measurement error.

Time	h^2 Type ^a	Thresholding		Deep Regression	
		OLS ^b	RF ^c	ResNet50	MobileViT
T1	h^2_{BW} ^d	0.54 (0.08)		0.55 (0.11)	
	h^2_{CV} ^e	0.44 (0.10)	0.56 (0.09)	0.55 (0.10)	0.55 (0.09)
	h^2_{TME} ^f	0.00 (0.02)	0.02 (0.03)	0.03 (0.02)	0.03 (0.02)
T2	h^2_{BW}	0.46 (0.09)		0.39 (0.09)	
	h^2_{CV}	0.33 (0.11)	0.40 (0.10)	0.39 (0.07)	0.40 (0.08)
	h^2_{TME}	0.00 (0.03)	0.00 (0.03)	0.00 (0.02)	0.00 (0.03)
T3	h^2_{BW}	0.31 (0.06)		0.29 (0.07)	
	h^2_{CV}	0.36 (0.08)	0.32 (0.07)	0.31 (0.06)	0.32 (0.07)
	h^2_{TME}	0.02 (0.02)	0.00 (0.01)	0.00 (0.02)	0.00 (0.01)
T4	h^2_{BW}	0.23 (0.11)		0.21 (0.10)	
	h^2_{CV}	0.27 (0.12)	0.25 (0.10)	0.21 (0.09)	0.20 (0.11)
	h^2_{TME}	0.00 (0.04)	0.03 (0.03)	0.00 (0.02)	0.00 (0.03)

^a Types of genomic heritability estimate. ^b Ordinary least squares. ^c Random forests. ^d Genomic heritability estimate for scale-based body weight. ^e Genomic heritability estimate for predicted weight using computer vision. ^f Genomic heritability estimate for trait measurement error.

Table 4 shows the number of top 10 markers detected by scale-based BW that are correctly detected in the top 100 markers of image-based BW, and the number of top 10 markers detected by scale-based BW that are incorrectly detected in the top 100 markers of trait measurement error BW in GWAS. If computer vision approaches can accurately estimate BW without systematic error, the top 10 markers identified by scale-based methods should also be among the top 100 markers identified by image-based methods. However, these top 10 scale-based markers would not be found among the top 100 markers derived from trait measurement errors. The GWAS analysis showed that both the RF-based adaptive thresholding segmentation and the deep regression approaches can always correctly identify the top 10 markers from scale-based BW within their top 100 markers. However, the OLS-based adaptive thresholding approach correctly identified only 4 to 7 top 10 markers. When GWAS was performed on trait measurement errors, none of the top 10 markers from scale-based BW were detected by the computer vision approaches within their top 100 markers.

Table 4. Number of top 10 markers detected in scale-based body weight that are correctly detected in the top 100 markers of image-based body weight (Agreement) and number of the top 10 markers detected in scale-based body weight that are incorrectly detected in the top 100 markers of trait measurement error (TME) body weight in the genome-wide association studies. T1 to T4 denote four farm visits.

Time	Thresholding Segmentation				Deep Regression			
	OLS ^a		RF ^b		ResNet50		MobileViT	
	Agreement	TME	Agreement	TME	Agreement	TME	Agreement	TME
T1	4/10	0/10	10/10	0/10	10/10	0/10	10/10	0/10
T2	7/10	0/10	10/10	0/10	10/10	0/10	10/10	0/10
T3	7/10	0/10	10/10	0/10	10/10	0/10	10/10	0/10
T4	4/10	0/10	10/10	0/10	10/10	0/10	10/10	0/10

^a Ordinary least squares. ^b Random forests.

4. Discussion

Computer vision approaches are rapidly being developed to accelerate phenomic research for livestock genetic analysis [34]. However, more work is needed before the incorporation of image-based phenotypes into genetic evaluation analysis can be considered. There is a possibility that the reliability of phenotyping may decrease as phenotyping becomes increasingly large-scale in noisy heterogeneous environments, and systematic errors that may arise for specific animals as a result of phenotyping may be a serious concern.

This study investigated the influence of trait measurement error on downstream genetic analysis of BW in pigs phenotyped under commercial environmental conditions. Continuous recording of BW is becoming increasingly important because BW can be used to derive residual feed intake, which is viewed as a proxy for difficult-to-measure feed efficiency [35]. Here, trait measurement error was defined as the difference between the scaled and computer vision-derived BW. Previous studies have suggested that downstream quantitative genetic analysis may be compromised even in a scenario where conventionally measured phenotypes and computer vision-derived phenotypes are highly correlated or have high goodness of fit. Although ensuring the safety of computer vision-derived phenotypes for use in downstream animal quantitative genetics is an important first step in data analysis, it has not received much attention in the literature. One possible reason for the lack of this type of study is that it requires phenotyping animals using both conventional and computer vision approaches, which requires intensive manual work and image analysis. Therefore, in the current study, we collected both scale-based and image-based BW and evaluated the consequences of using image-based BW as phenotypes in quantitative genetic analysis by deriving trait measurement error.

We found that some pigs had consistently high measurement errors across different time points, suggesting that these pigs were always difficult to accurately predict BW from images (Figure 2). In addition, OLS-based adaptive thresholding performed the least well compared to competing approaches in inference tasks. For example, the OLS-based adaptive thresholding approach showed relatively low R^2 and high MAPE between scale-based BW and image-based BW, image-based genomic heritability estimates sometimes differed by 0.1 from scale-based estimates, and only 4 to 7 of the top 10 GWAS markers detected in scale-based BW were correctly detected in the top 100 GWAS markers of image-based BW. However, these concerns did not lead to evidence of a genetic component to trait measurement error. Genomic heritability estimates for trait measurement error were always zero or negligible across growth periods, regardless of the choice of the computer vision approaches (Table 3). Also, none of the top 10 GWAS markers detected in scale-based BW were detected in the top 100 GWAS markers of trait measurement error. Our GWAS results

are in line with a previous work reporting no significant difference in GWAS peaks between simulated manual and image-based maize tassel characters [36]. Thus, it is likely that trait measurement error in image-based pig BW does not contain systematic errors characterized by genetics that could potentially bias downstream genetic analyses. A previous study reported that image-based phenotypes tended to overestimate pedigree-based heritability estimates relative to manually collected phenotypes for sheep growth traits [8], but we did not find a similar trend in our study.

Taken together, our results suggest that manual scale-based and image-based BW are controlled by the same set of causal loci, and their difference in phenotypic values is due to different phenotyping errors. Thus, the trait measurement error of BW during pig growth is generated by phenotyping errors that are not genetic, regardless of the computer vision approaches used. In particular, OLS-based adaptive thresholding did not produce BW as accurately as other competing approaches, but its trait measurement error was still free of genetic factors. Although the sample size used in the current study was relatively small, we believe this study serves as a first step toward a better understanding of trait measurement errors. It is worth noting that although manually measured phenotypes are often considered ground truth in computer vision analysis, even manual measurements have some degree of phenotyping error due to human or mechanical error. As the phenotyping error due to computer vision becomes smaller, the accuracy of phenotypes is expected to improve further.

5. Conclusions

The advent of precision livestock farming offers increasing opportunities to use new technologies to generate phenotypes. Our results provide valuable insights into the effect of trait measurement error on quantitative genetic analysis in pigs when using computer vision technology. Our study found no evidence that trait measurement error is influenced by genetic factors during pig growth. It is likely that image-derived BW in pigs can be used without reservation for downstream quantitative genetic analysis.

Author Contributions: Y.B.: Formal analysis, Investigation; Data curation; Visualization; Writing—original draft. Y.H.: Investigation; Data curation; Project administration; Writing—review and editing. H.Y.: Methodology; Writing—review and editing. G.M.: Conceptualization; Methodology; Formal analysis; Investigation; Data curation; Funding acquisition; Project administration; Supervision; Writing—original draft; Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded, in part, by the United States Department of Agriculture National Institute of Food and Agriculture, Hatch project 7000564.

Institutional Review Board Statement: All biological material used in this study was collected as part of routine data collection at Smithfield Premium Genetics and not specifically for the purpose of this project. Therefore, ethics committee approval was not required.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the results of this article are the property of Smithfield Premium Genetics, and this information is commercially sensitive and therefore cannot be made available.

Acknowledgments: The authors acknowledge Advanced Research Computing at Virginia Tech (<https://arc.vt.edu/>) for providing computational resources and technical support that have contributed to the results reported within this paper.

Conflicts of Interest: Yijian Huang is an employee of Smithfield Premium Genetics (Rose Hill, NC, USA). The authors declare that they have no competing interests.

References

1. Morota, G.; Lee, H.J.; Lee, S.H.; Park, H.B.; Onogi, A. Artificial intelligence in animal breeding and genetics: Applications, opportunities, and challenges. *Anim. Front.* **2026**, vfaf058. [\[CrossRef\]](#)
2. Morota, G.; Ventura, R.V.; Silva, F.F.; Koyama, M.; Fernando, S.C. Big data analytics and precision animal agriculture symposium: Machine learning and data mining advance predictive big data analysis in precision animal agriculture. *J. Anim. Sci.* **2018**, 96, 1540–1550. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Kadlec, R.; Indest, S.; Castro, K.; Waqar, S.; Campos, L.M.; Amorim, S.T.; Bi, Y.; Hanigan, M.D.; Morota, G. Automated acquisition of top-view dairy cow depth image data using an RGB-D sensor camera. *Transl. Anim. Sci.* **2022**, 6, txac163. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Nye, J.; Zingaretti, L.M.; Pérez-Enciso, M. Estimating conformational traits in dairy cattle with DeepAPS: A two-step deep learning automated phenotyping and segmentation approach. *Front. Genet.* **2020**, 11, 513. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Gorssen, W.; Winters, C.; Meyermans, R.; D’Hooge, R.; Janssens, S.; Buys, N. Estimating genetics of body dimensions and activity levels in pigs using automated pose estimation. *Sci. Rep.* **2022**, 12, 15384. [\[CrossRef\]](#)
6. Ricard, A.; Crevier-Denoix, N.; Pourcelot, P.; Crichan, H.; Sabbagh, M.; Dumont-Saint-Priest, B.; Danvy, S. Genetic analysis of geometric morphometric 3D visuals of French jumping horses. *Genet. Sel. Evol.* **2023**, 55, 63. [\[CrossRef\]](#)
7. Manzanilla-Pech, C.I.; Stephansen, R.B.; Lassen, J. Genetic parameters for feed intake and body weight in dairy cattle using high-throughput 3-dimensional cameras in Danish commercial farms. *J. Dairy Sci.* **2023**, 106, 9006–9015. [\[CrossRef\]](#)
8. Qin, Q.; Zhang, C.; Liu, Z.; Wang, Y.; Kong, D.; Zhao, D.; Zhang, J.; Lan, M.; Wang, Z.; Alatan, S.; et al. Estimation of the Genetic Parameters of Sheep Growth Traits Based on Machine Vision Acquisition. *Animal* **2024**, 18, 101196. [\[CrossRef\]](#)
9. Moore, K.; Mrode, R.; Coffey, M. Genetic parameters of Visual Image Analysis primal cut carcass traits of commercial prime beef slaughter animals. *Animal* **2017**, 11, 1653–1659. [\[CrossRef\]](#)
10. Nakajima, A.; Kawaguchi, F.; Uemoto, Y.; Fukushima, M.; Yoshida, E.; Iwamoto, E.; Akiyama, T.; Kohama, N.; Kobayashi, E.; Honda, T.; et al. A genome-wide association study for fat-related traits computed by image analysis in Japanese Black cattle. *Anim. Sci. J.* **2018**, 89, 743–751. [\[CrossRef\]](#)
11. Xie, L.; Qin, J.; Rao, L.; Tang, X.; Cui, D.; Chen, L.; Xu, W.; Xiao, S.; Zhang, Z.; Huang, L. Accurate prediction and genome-wide association analysis of digital intramuscular fat content in longissimus muscle of pigs. *Anim. Genet.* **2021**, 52, 633–644. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Kaseja, K.; Lambe, N.; Yates, J.; Smith, E.; Conington, J. Genome wide association studies for carcass traits measured by video image analysis in crossbred lambs. *Meat Sci.* **2024**, 214, 109518. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Shen, Y.; Chen, Y.; Zhang, S.; Wu, Z.; Lu, X.; Liu, W.; Liu, B.; Zhou, X. Smartphone-based digital phenotyping for genome-wide association study of intramuscular fat traits in longissimus dorsi muscle of pigs. *Anim. Genet.* **2024**, 55, 230–237. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Hollifield, M.K.; Chen, C.Y.; Psota, E.; Holl, J.; Lourenco, D.; Misztal, I. Estimating genetic parameters of digital behavior traits and their relationship with production traits in purebred pigs. *Genet. Sel. Evol.* **2024**, 56, 29. [\[CrossRef\]](#)
15. Liang, Z.; Pandey, P.; Stoerger, V.; Xu, Y.; Qiu, Y.; Ge, Y.; Schnable, J.C. Conventional and hyperspectral time-series imaging of maize lines widely used in field trials. *Gigascience* **2018**, 7, gix117. [\[CrossRef\]](#)
16. Zhou, Y.; Kusmec, A.; Mirnezami, S.V.; Attigala, L.; Srinivasan, S.; Jubery, T.Z.; Schnable, J.C.; Salas-Fernandez, M.G.; Ganapathysubramanian, B.; Schnable, P.S. Identification and utilization of genetic determinants of trait measurement errors in image-based, high-throughput phenotyping. *Plant Cell* **2021**, 33, 2562–2582. [\[CrossRef\]](#)
17. Yu, H.; Lee, K.; Morota, G. Forecasting dynamic body weight of nonrestrained pigs from images using an RGB-D sensor camera. *Transl. Anim. Sci.* **2021**, 5, txab006. [\[CrossRef\]](#)
18. Wang, J.; Hu, Y.; Xiang, L.; Morota, G.; Brooks, S.A.; Wickens, C.L.; Miller-Cushon, E.K.; Yu, H. ShinyAnimalCV: Open-source cloud-based web application for object detection, segmentation, and three-dimensional visualization of animals using computer vision. *J. Anim. Sci.* **2024**, 102, skad416. [\[CrossRef\]](#)
19. Bi, Y.; Campos, L.M.; Wang, J.; Yu, H.; Hanigan, M.D.; Morota, G. Depth video data-enabled predictions of longitudinal dairy cow body weight using thresholding and Mask R-CNN algorithms. *Smart Agric. Technol.* **2023**, 6, 100352. [\[CrossRef\]](#)
20. Bi, Y.; Huang, Y.; Xuan, J.; Morota, G. Industry-scale prediction of video-derived pig body weight using efficient convolutional neural networks and vision transformers. *Biosyst. Eng.* **2025**, 257, 104243. [\[CrossRef\]](#)
21. Bi, Y.; Xuan, J.; Huang, Y.; Morota, G. 465 Comparative analysis of semantic segmentation and deep regression models with supervised pre-training for accurate prediction of pig body weight from video data: Insights from industry-scale datasets. *J. Anim. Sci.* **2024**, 102, 414–415. [\[CrossRef\]](#)
22. Intel RealSense. Intel RealSense Github. Available online: <https://github.com/IntelRealSense> (accessed on 1 March 2023).
23. Jocher, G.; Chaurasia, A.; Qiu, J. Ultralytics YOLOv8. 2023. Available online: <https://docs.ultralytics.com/models/yolov8/> (accessed on 10 January 2023).
24. Lin, Q.; Ye, G.; Wang, J.; Liu, H. Roboflow: A data-centric workflow management system for developing ai-enhanced robots. In *Proceedings of the Conference on Robot Learning*; PMLR: Chicago, IL, USA, 2022; pp. 1789–1794.

25. Kuhn, M. Caret: Classification and regression training. In *Astrophysics Source Code Library*; GitHub: San Francisco, CA, USA, 2015; p. ascl-1505. Available online: <https://github.com/topepo/caret/> (accessed on 22 July 2025).
26. Mehta, S.; Rastegari, M. MobileViT: Light-weight, general-purpose, and mobile-friendly vision transformer. *arXiv* **2021**, arXiv:2110.02178.
27. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
28. Patterson, H.D.; Thompson, R. Recovery of inter-block information when block sizes are unequal. *Biometrika* **1971**, *58*, 545–554. [[CrossRef](#)]
29. VanRaden, P. Efficient Methods to Compute Genomic Predictions. *J. Dairy Sci.* **2008**, *91*, 4414–4423. [[CrossRef](#)] [[PubMed](#)]
30. Kennedy, B.; Quinton, M.; Van Arendonk, J. Estimation of effects of single genes on quantitative traits. *J. Anim. Sci.* **1992**, *70*, 2000–2012. [[CrossRef](#)]
31. Yu, J.; Pressoir, G.; Briggs, W.H.; Bi, I.V.; Yamasaki, M.; Doebley, J.F.; McMullen, M.D.; Gaut, B.S.; Nielsen, D.M.; Holland, J.B.; et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* **2006**, *38*, 203. [[CrossRef](#)]
32. Endelman, J.B. Ridge regression and other kernels for genomic selection with R package rrBLUP. *Plant Genome* **2011**, *4*, 250–255. [[CrossRef](#)]
33. Huang, M.; Zhang, W.; Dong, J.; Hu, Z.; Tan, X.; Li, H.; Sun, K.; Zhao, A.; Huang, T. Genome-wide association studies of body weight and average daily gain in Chinese Dongliao black pigs. *Int. J. Mol. Sci.* **2025**, *26*, 3453. [[CrossRef](#)]
34. Pérez-Enciso, M.; Steibel, J.P. Phenomes: The current frontier in animal breeding. *Genet. Sel. Evol.* **2021**, *53*, 22. [[CrossRef](#)]
35. Patience, J.F.; Rossoni-Serão, M.C.; Gutiérrez, N.A. A review of feed efficiency in swine: Biology and application. *J. Anim. Sci. Biotechnol.* **2015**, *6*, 33. [[CrossRef](#)]
36. Gage, J.L.; de Leon, N.; Clayton, M.K. Comparing genome-wide association study results from different measurements of an underlying phenotype. *G3 Genes Genomes Genet.* **2018**, *8*, 3715–3722. [[CrossRef](#)]

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