AUTOMATED EVALUATION OF PORK CHOP QUALITY TRAITS VIA COMPUTER VISION AND DEEP LEARNING

Shubham Bery[1], Eric Psota[1], Justin Holl[1], Brandon Fields[1*], Neal Matthews[1], L. Clay Eastwood[1], Robert Fitzgerald[1]
[1]PIC® North America, Hendersonville, Tennessee, USA

INTRODUCTION

Conventional pork quality assessment methods, such as evaluating marbling, often depend on subjective scoring by trained experts, which can lead to variability among evaluators. In contrast, objective measures like pH and color require expensive, specialized equipment along with additional human labor, making them both costly and time-consuming [1]. The objective of this study was to develop an automated, objective evaluation system for predicting marbling in pork chops using computer vision and deep learning and to compare predictions with human evaluations.

MATERIAL AND METHODS

Imaging Setup and Data Collection

- A custom imaging station was developed to ensure consistent capture of pork chop images. Each pork chop was placed on a cutting board with a numerical ID label below it (Figure 1a). A single button triggers image capture, followed by a brief two-second lighting cue to confirm successful acquisition. A custom optical character recognition (OCR) routine assigns the proper ID to each image captured.

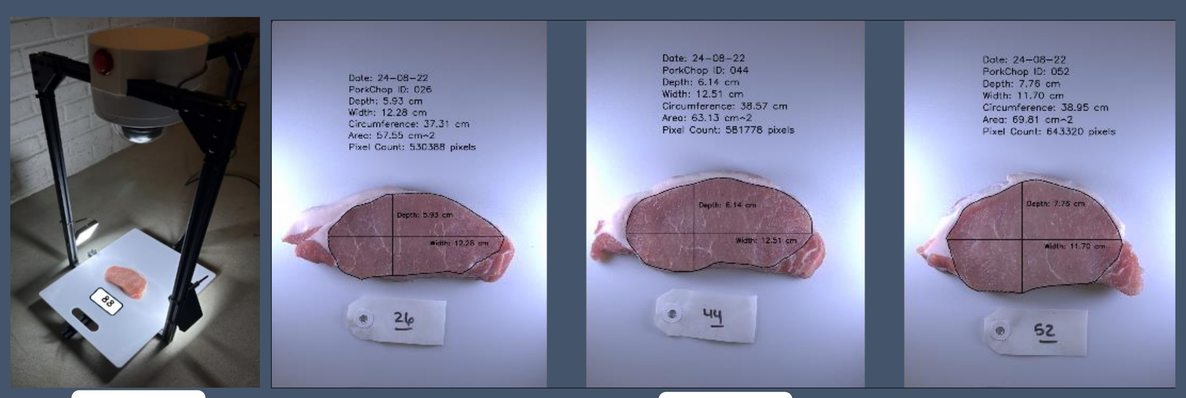

Figure 1. Device used for image capture (a) and three visualizations of pork chop size estimation (b)

Data Processing & Analysis

- A total of 1,298 images were collected and manually labeled by three experts for marbling. Other traits included subjective scores for color, as well as objective measures for pH and color (Minolta L*, a*, and b*). The imaging station can estimate the size of pork chops using a trained semantic segmentation model (Figure 1b).
- To predict marbling scores, the ResNet-18 convolutional neural network was adapted. The pre-trained model's fully connected layer was replaced with a dropout layer (p = 0.5) and a linear layer that generates outputs [2]. The model was trained using the mean of the three marbling scores and single observations for other traits. The dataset was split 80:20 for training and validation. Extensive augmentations, including random rotations, flips, and brightness adjustments, were applied to increase model robustness.
- The model's performance was assessed using the coefficient of determination (R^2). Repeatability analysis was conducted to assess the consistency of expert evaluations and the model's performance. Experts were shown images of 50 chops twice and asked to score them for marbling. For the model, repeatability analysis was performed using two different images of each chop from a dataset containing 162 chops. All other evaluations, except for the repeatability analysis of expert scores, were conducted on a test dataset consisting of 260 images.

RESULTS AND DISCUSSION

- The trained neural network demonstrated strong predictive capability for marbling (R^2 = 0.84). Color predictions were moderate, with lightness (L*) yielding R^2 = 0.75, while redness (a*) and yellowness (b*) yielded R^2 = 0.53 and 0.55, respectively. The pH was least accurate with an R^2 = 0.32 (Table 1).

RESULTS AND DISCUSSION

- Expert repeatability ranged from 0.69 to 0.88, averaging 0.79, whereas the model's repeatability was notably higher at 0.94. This indicates that the automated system demonstrates more consistent scoring than individual experts; despite being disadvantaged since it was given different images of the same chop instead of the same image twice.

Table 1. Coefficient of Determination (R^2) for Quality Traits

Quality Trait	Coefficient of Determination (R^2)
pH	0.32
Color (Subjective)	0.53
Color (L*, a*, b*)	0.75, 0.53, 0.55
Marbling	0.84

- Table 2 presents an R^2 matrix comparing individual experts, the model's predictions, and the mean. The model correlates most strongly with the mean score (R = 0.90) indicating that the automated system's outputs align most closely with the consensus. Also, with the exception of Expert 1, the model's scores had higher correlations with individual experts than the experts had among themselves.

Table 2. Comparing the Model's Predictions with Individual Expert Evaluations and the Mean Score

	Expert 1	Expert 2	Expert 3	Model	Mean (Experts + Model)
Expert 1	-	0.72	0.58	0.70	0.84
Expert 2	-	-	0.67	0.78	0.89
Expert 3	-	-	-	0.73	0.87
Model	-	-	-	-	0.90

CONCLUSIONS

- This study demonstrates that a computer vision system can robustly predict pork chop quality traits, notably marbling, and achieves higher repeatability than human experts. Color predictions also proved reasonably accurate, whereas pH predictions may benefit from additional data or imaging modalities. By aligning most closely with the average of expert scores and maintaining consistently high correlations with individual evaluators the model offers a reliable and scalable alternative to subjective assessments, promising greater efficiency and reduced variability in meat quality evaluation.

REFERENCES

1. Shi, Y., Wang, X., Borhan, M. S., Young, J., Newman, D., Berg, E., Sun, X. (2021) A review on meat quality evaluation methods based on non-destructive computer vision and artificial intelligence technologies. Food Science of Animal Resources, 41: 563-588.
2. King, D. A., Hunt, M. C., Barbut, S., Claus, J. R., Cornforth, D. P., Joseph, P., Kim, Y. H., Lindahl, G., Mancini, R. A., Nair, M. N., Merok, K. J., Milkowski, A., Mohan, A., Pohlman, F., Ramanathan, R., Raines, C. R., Seyfert, M., Sørheim, O., Suman, S. P., Weber, M. (2023) American Meat Science Association Guidelines for Meat Color Measurement. Meat and Muscle Biology, 6(4): 12473, 1-81.