

APPLYING FOUNDATION SEGMENTATION MODELS TO ASSESS POULTRY PHENOTYPES AND BEHAVIORS

by

Mahtab Saeidifar

(Under the Direction of Guoming Li)

ABSTRACT

Advanced computer vision and deep learning methods have been developed to enhance precision poultry farming through automated health status identification and activity monitoring across multiple poultry species. A modified Segment Anything Model (SAM) pipeline, combined with pre- and post-processing techniques but without extensive model training, achieved 84.4% segmentation success for cage-free laying hens in thermal images, enabling non-invasive extraction of body temperature to support early health assessment. A hybrid YOLOv7 + SAM model using bounding box prompts achieved 98.0% segmentation accuracy, allowing precise individual identification. Additionally, an open-source, user-friendly Streamlit platform integrating SAM2 was developed to enable non-technical researchers to track animal activity across different species directly from video data without any model training or manual labeling. These tools minimize manual intervention, reduce animal stress, and improve decision-making by providing automated monitoring of phenotypic and behavioral indicators, with broad applicability in precision livestock farming and smart agricultural systems.

INDEX WORDS: Precision poultry farming; computer vision; zero-shot segmentation; health status identification; activity monitoring

APPLYING FOUNDATION SEGMENTATION MODELS TO ASSESS POULTRY
PHENOTYPES AND BEHAVIORS

by

MAHTAB SAEIDIFAR

BS.c. Razi University, Iran, 2017

MS.c. Shiraz University, Iran, 2020

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment
of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2025

© 2025

Mahtab Saeidifar

All Rights Reserved

APPLYING FOUNDATION SEGMENTATION MODELS TO ASSESS POULTRY
PHENOTYPES AND BEHAVIORS

by

MAHTAB SAEIDIFAR

Major Professor:	Guoming Li
Committee:	Khaled Rasheed
	Jin Lu
	Chongxiao (Sean) Chen

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
August 2025

DEDICATION

I would like to dedicate this thesis to my husband, for his unwavering love, daily encouragement, and endless support that have carried me through every step of this journey.

To my parents and my brother, for their constant support, belief in me, and for always being my source of strength and inspiration.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Guoming Li. He has been more than just an advisor—his support, kindness, and dedication have felt like that of a family member. He has always been incredibly hardworking, and whenever I or any of his students needed help, he was there immediately without hesitation. His unwavering encouragement and selfless dedication to our success have left a lasting impact on me. I couldn't be more grateful and proud to have had the opportunity to work with him.

I would also like to thank my committee members. To Prof. Khaled Rasheed, whose kindness and encouragement, both in and out of the classroom, have meant so much to me. His engaging courses and fatherly support created an environment where I could grow with confidence. To Dr. Jin Lu, whom I deeply admire for his profound knowledge and expertise. I greatly enjoyed and benefited from his machine learning class. To Dr. Chongxiao Chen, for his positive energy, great rapport with students, and continued support throughout this journey—I sincerely appreciate all the guidance and encouragement.

I am also grateful to my labmates: Venkat Umesh Chandra Bodempudi, Oluwadamilola Oso, Hosna Mohammadilalabadi, Aravind Mandiga, Tongshuai Liu, and Sai Akshita Reddy Kota. Thank you for your collaboration and friendship throughout the team projects. Whenever I needed help, you were always there to lend a hand, and I truly appreciate that.

A heartfelt thank you to my husband, Ehsan Asali, for his constant support and encouragement. From listening to my research ideas to giving me valuable feedback, you have

been an incredible companion in both life and research. I am truly thankful to have you by my side.

Lastly, to my parents and my brother—thank you for always believing in me and encouraging me to grow. Your love, support, and sacrifices have made all of this possible.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	v
LIST OF TABLES	xi
LIST OF FIGURES	xiii
CHAPTER	
I PRECISION POULTRY FARMING: LITERATURE REVIEW.....	1
INTRODUCTION	1
OVERVIEW OF POULTRY PHENOTYPING ASSESSMENT	3
OVERVIEW OF IMAGE SEGMENTATION TECHNIQUES	4
<i>Conventional image processing algorithms</i>	5
<i>Machine learning-based methods</i>	8
<i>Interactive segmentation methods</i>	9
<i>Deep learning-based semantic segmentation methods</i>	11
<i>Deep learning-based instance segmentation methods</i>	12
<i>Hybrid methods</i>	14
APPLICATIONS OF SEGMENTATION IN PRECISION ANIMAL MANAGEMENT	15
<i>Behavior and Welfare Assessment</i>	16
<i>Health Status Identification</i>	17
<i>Live Performance Prediction</i>	19
<i>Product Quality Inspection</i>	20
<i>Animal Trait Recognition</i>	21
OBJECTIVES AND OUTLINE OF THE DISSERTATION	22

REFERENCES	23
II ZERO-SHOT IMAGE SEGMENTATION FOR MONITORING THERMAL CONDITIONS OF INDIVIDUAL CAGE-FREE LAYING HENS.....	39
INTRODUCTION	40
MATERIALS AND METHODS	44
<i>Overall workflow</i>	44
<i>Data collection and thermal camera calibration</i>	46
<i>Spatial alignment</i>	49
<i>Comparison of different zero-shot deep learning models for hen segmentation</i>	50
<i>Modification of Segment Anything model</i>	52
<i>Evaluation metrics calculation</i>	59
<i>Automatic digit extraction in thermal images</i>	60
<i>Finding the relation between temperature and pixels intensity</i>	61
<i>Extracting statistics of chicken's body temperature</i>	62
RESULTS AND DISCUSSION	63
<i>Thermal camera calibration</i>	63
<i>Selecting the optimal model for zero-shot cage-free hen detection and segmentation</i>	64
<i>Comparison of five different machine learning classifiers for the post-processing</i>	68
<i>Comparison of modified SAM-based models and two generic models for cage-free hen detection and segmentation</i>	70
<i>Statistics of surface temperature in chickens over four weeks of age</i>	75
ADDITIONAL DISCUSSION AND FUTURE WORK	77
CONCLUSIONS	79
DECLARATION OF COMPETING INTEREST.....	80
ACKNOWLEDGEMENT	80
AUTHOR CONTRIBUTION	80
REFERENCES	80

III AUTOMATIC SEGMENTATION OF BIRDS USING A COMBINATION OF OBJECT DETECTION AND FOUNDATION IMAGE SEGMENTATION MODELS 87

INTRODUCTION	87
MATERIALS AND METHODS	91
<i>Overall workflow</i>	91
<i>Animal, housing, and management</i>	93
<i>Dataset</i>	93
<i>Object detection model (YOLOv7)</i>	94
<i>Extracting the coordinates of the bounding box</i>	95
<i>Proposed method</i>	95
<i>YOLOv7 (box prompts) + SAM</i>	96
<i>YOLOv7 (point prompts) + SAM</i>	97
<i>Comparative analysis</i>	99
<i>Evaluation metrics calculation</i>	100
RESULTS AND DISCUSSION	102
<i>Image similarity scores</i>	102
<i>Selecting the optimal model for zero-shot hen detection</i>	103
<i>Comparison of different models for hen segmentation</i>	105
CONCLUSION	107
ACKNOWLEDGEMENT	108
REFERENCES	108

IV ANIMALAI: AN OPEN-SOURCE WEB PLATFORM FOR AUTOMATED ANIMAL ACTIVITY

INDEX CALCULATION USING INTERACTIVE DEEP LEARNING SEGMENTATION..... 111

INTRODUCTION	112
MATERIALS AND METHODS	117
<i>Animal housing and video data collection</i>	117
<i>Overall workflow</i>	118

<i>Video uploading</i>	120
<i>Video frame extraction</i>	121
<i>Interactive animal selection</i>	122
<i>Segmentation using Segment Anything Model 2</i>	123
<i>Frame differencing for calculating activity index</i>	125
<i>Visualizing the activity index</i>	127
<i>Evaluation metrics calculation</i>	128
<i>Evaluating the impact of segmentation on activity index accuracy</i>	130
<i>Statistical analysis of different ratios of birds to represent the entire group's activity</i>	130
RESULTS AND DISCUSSION	132
<i>Example procedure of interface operations</i>	132
<i>Segmentation performance on a chicken dataset</i>	133
<i>Comparison of activity index calculation with and without segmentation</i>	137
<i>Determination of optimal sampling ratio for group activity assessment</i>	138
CONCLUSIONS	142
DECLARATION OF COMPETING INTEREST.....	143
ACKNOWLEDGEMENT	143
AUTHOR CONTRIBUTION	143
REFERENCES	143
V SUMMARY	150

LIST OF TABLES

	Page
Table 1.1. Applications of conventional image processing algorithms for poultry segmentation .	7
Table 1.2. Applications of machine learning-based methods for poultry segmentation	9
Table 1.3. Applications of deep learning-based semantic segmentation methods for poultry segmentation	11
Table 1.4. Applications of deep learning-based instance segmentation methods for poultry segmentation	13
Table 1.5. Applications of hybrid methods for poultry segmentation	15
Table 2.1. Performance comparison of hen detection before any modifications.....	65
Table 2.2. Performance comparison of hen segmentation before any modifications.....	66
Table 2.3. Hyperparameter tuning results for machine learning classifiers	68
Table 2.4. Comparison of different machine learning classifiers for mask selection.....	69
Table 2.5. Performance comparison of hen detection for modified SAM-Based models	71
Table 2.6. Performance comparison of hen segmentation for modified SAM-Based models	72
Table 2.7. Comparison of study durations in various research studies.....	77
Table 3.1. comparative analysis of hen detection performance.....	104
Table 3.2. comparison of different segmentation metrics across the models.	105
Table 4.1. Segmentation performance of Segment Anything Model 2 for segmenting broiler chickens at Weeks 1-217.	136

Table 4.2. Average Pearson correlation coefficients (r value) between each sampling ratio and the entire flock at different broiler growth stages (Weeks 1, 4, and 7).....	140
---	-----

LIST OF FIGURES

	Page
Fig. 1.1. Computer Vision System - Image processing steps (Geronimo et al. 2019).....	6
Fig. 1.2. Individual bird segmentation using Segment Anything Model and post-processing. Fig. 1.2(a) provides a high-level overview of the Segment Anything Model and Fig. 1.2(b) the post-processing step (Willems et al. 2025).	10
Fig. 2.1. Workflow diagram - This figure presents a schematic of the nine-step analytical process employed in the paper. SAM is Segment Anything, YOLO is You Only Look Once, and R-CNN is Region-based Convolutional Neural Network.	45
Fig. 2.2. Precision calibrator for thermal camera calibration.	47
Fig. 2.3. Ten pairs of RGB and thermal images: a, c) RGB images; and b, d) corresponding thermal images. The bottom-left number under the bar in each thermal image represents the lowest temperature, and the top-left number above the bar indicates the highest temperature recorded in the scene.....	48
Fig. 2.4. Spatial alignment for RGB images to match corresponding thermal images: a) Original RGB image; b) RGB image after cropping and resizing; and c) corresponding thermal image. .	50
Fig. 2.5. Modified Segment Anything Model overview. A heavyweight image encoder outputs an image embedding that can then be efficiently queried by defining an automatic initial point to produce object masks at amortized real-time speed. The yellow numbers in the masks indicate the confidence scores of the segmentation. To select the best masks out of the three masks generated, a machine learning classifier was used.....	52

Fig. 2.6. SAM Pre-processing: a) Thermal image; b) segmented image after thresholding; c) extraction of the largest area; d) thermal image with the initial point determined from the center of the largest area; and e) thermal image with the initial point determined from the highest intensity (highest temperature).	54
Fig. 2.7. SAM Post-processing: a) RGB image; b, c, d) three masks generated using SAM where b shows the masks with the highest confidence scores; and e) ground truth image.....	55
Fig. 2.8. Extracting digits from a thermal image using pytesseract. OCR is Optical Character Recognition.	61
Fig. 2.9. Multiplication of the mask on the thermal image: a) Thermal image; b) segmented mask; and c) resulting image after multiplication.....	62
Fig. 2.10. The temperature calibration curve and the fitting formula.	64
Fig. 2.11. Segmentation results of different models: a) RGB image; b) MobileSAM; c) FastSAM; d) SAM; e) YOLOv8, f) Mask R-CNN, g) U ² -Net, h) ISNet; and i) ground truth.	68
Fig. 2.12. Hen segmentation using Modified SAM: a) RGB Image; b) thermal image; c) segmentation result using SAM without modification; and d) segmentation result using Modified SAM; and e) ground truth image.	74
Fig. 2.13. Extraction of average statistics with standard deviation for surface body temperature of hens from weeks 77 to 80.	76
Fig. 2.14. Distribution of image similarity scores	78
Fig. 3.1. Workflow diagram - This figure presents a schematic of the six-step analytical process employed in the paper. SAM is Segment Anything, YOLO is You Only Look Once, and R-CNN is Region-based Convolutional Neural Network.	92

Fig. 3.2. Two pairs of RGB and thermal images: a, c) RGB images; and b, d) corresponding thermal images. Digits inside thermal images indicate maximal or minimal temperatures.	94
Fig. 3.3. YOLO output.....	95
Fig. 3.4. Proposed Method (YOLOv7 + SAM): The method uses a robust image encoder to generate an image embedding. This embedding can be efficiently queried by defining an automatic initial point, enabling the production of bird masks at amortized real-time speed. The yellow numbers on the masks represent the confidence scores of the segmentation, with the mask having the highest score being selected as the final output.	96
Fig. 3.5. Segmentation results of YOLOv7 (providing bounding boxes of detected birds as box prompts) + SAM: a) detected chickens enclosed with bounding box; b) segmentation results; c) ground truths.	97
Fig. 3.6. Segmentation results of YOLOv7 (point prompts) + SAM: a) detected chickens enclosed with bounding boxes; b) RGB image with centroid points of detected bounding boxes; c) segmentation results; d) ground truths.....	98
Fig. 3.7. Erroneous segmentation results of YOLOv7 (point prompts) + SAM: a) detected chickens enclosed with bounding boxes; b) RGB image with centroid points of detected bounding boxes; c) erroneous segmentation results; d) ground truths.	99
Fig. 3.8. Similarity score histogram of RGB images in the dataset.....	103
Fig. 3.9. Segmentation results of different models: a) RGB image; b) thermal image + MobileSAM; c) thermal image + FastSAM; d) thermal image + SAM; e) YOLOv8; f) Mask R-CNN; g) YOLOv7 (box prompt) + SAM; h) YOLOv7 (point prompts) + SAM; and i) ground truth.....	107

Fig. 4.1. Workflow diagram of the platform - This figure presents a schematic of different steps analytical process employed in the platform. Red color indicates start and end points of the process; blue color indicates main processing steps; orange color indicates decision points; purple color indicates user input parameters; and green color indicates files saved in the main directory.	120
Fig. 4.2. The graphical user interface of the application displaying the frame extraction process. Users can select a frame interval after uploading a video, adjust settings based on recommendations, and trigger the extraction process using the ‘Extract Frames’ button.	122
Fig. 4.3. Examples of segmentation outputs generated by the application using the Segment Anything Model 2. The RGB mask frames (top-right) highlight the selected animals in distinct colors, while the binary mask frames (bottom-right) isolate the targeted animals from the background.	124
Fig. 4.4. Series of differenced frames illustrating pixel-level changes between consecutive segmented frames (binary mask frames) for calculating animal activity index.	126
Fig. 4.5. Activity index plot illustrating the level of movement for the selected animals throughout the video. The x -axis denotes the video duration in minutes and seconds, while the y -axis ranges from 0 (minimal activity) to 1 (maximum activity).	128
Fig. 4.6. SAM2 segmentation results: (a) the original frame, (b) the corresponding RGB mask output, (c) the Binary mask output, and (d) the ground truth segmentation.	134
Fig. 4.7. Effect of segmentation on the accuracy of the activity index: the top row shows the approach without segmentation, while the bottom row shows the approach with segmentation.	137

Fig. 4.8. Example of five sampling initializations at week 1, comparing (a) 20%, (b) 40%, (c) 60%, (d) 80%, and (e) the entire flock (100%).....	139
Fig. 4.9. Example of five sampling initializations at week 4, comparing (a) 20%, (b) 40%, (c) 60%, (d) 80%, and (e) the entire flock (100%).....	139
Fig. 4.10. Example of five sampling initializations at week 7, comparing (a) 20%, (b) 40%, (c) 60%, (d) 80%, and (e) the entire flock (100%).....	140
Fig. 4.11. Comparative P -value heatmap across different representation at (a) week 1, (b) week 4, and (c) week 7.....	141

CHAPTER I

PRECISION POULTRY FARMING: LITERATURE REVIEW

Introduction

According to the USDA National Agricultural Statistics Service (2025), the total value of U.S. poultry production—including broilers, eggs, turkeys, and chicken sales—reached \$70.2 billion in 2024, encompassing approximately 9.33 billion broilers, 200 million turkeys, and 182 million laying hens. Modern broiler houses typically hold between 25,000 to 100,000 broilers raised on open litter floors (Lei et al., 2022). As the global demand for affordable, high-quality animal protein continues to rise, poultry production systems are increasingly becoming larger and more intensive. While increasing animal density can enhance efficiency and profitability (Godfray et al., 2010), it also raises the risk of disease transmission, such as highly pathogenic avian influenza (HPAI), potentially causing severe outbreaks and significant economic losses (Tsiouris et al., 2015). For example, between 2015 and 2016, bird flu affected hundreds of farms across Canada and the U.S., resulting in over 48 million birds being culled (Pasick et al., 2015; Shriner et al., 2016). More recently, from 2022 to 2025, avian influenza impacted 51 states, 1,674 commercial flocks, and approximately 168.2 million birds (USDA Animal and Plant Health Inspection Service, 2025). Massive culling was undertaken to prevent further spread, dramatically affecting the egg market, with prices rising from \$0.61 per dozen in 2021 to \$8.05 per dozen in 2025 (Trading Economics, 2025). The virus also poses risks to other animals, such as cattle, further endangering animal production systems (Centers for Disease Control and Prevention, 2025).

Beyond avian influenza, poultry farms face numerous infectious diseases, including Salmonella (Oladeinde et al., 2025), infectious bronchitis (Raj and Jones, 1997), Newcastle disease (Sadeghi et al., 2023), and infectious bursal disease (Etteradossi and Saif, 2013). These diseases can lead to substantial financial losses if not promptly identified and managed. Therefore, it is essential to develop intelligent tools capable of rapidly and accurately monitoring poultry health, welfare, and productivity directly on-site.

Various diagnostic methods for poultry have been studied, each with strengths and limitations. Traditional methods involve collecting samples on-site and sending them to laboratories, which can take several days for diagnosis and requires skilled personnel (Vidic et al., 2017). Additionally, caretakers must identify subtle signs of illness among tens of thousands of birds, a challenging and often delayed process. Samples might need to be transported across long distances, further extending the time for diagnosis (Brown Jordan et al., 2018). Automated monitoring technologies, such as radio frequency identification (Ahmed et al., 2021; Li et al., 2019), accelerometers (Li and Chai, 2023; Okada et al., 2009), and audio sensors (Banakar et al., 2024), have achieved high accuracy (>90%) in detecting abnormal poultry behaviors. However, these technologies are generally too expensive to deploy individually on every bird in large poultry operations.

An alternative approach involves using non-contact, cost-effective solutions like computer vision systems integrated with machine learning algorithms. The combination of computer vision and machine learning is transforming animal agriculture, facilitating precision poultry science from genetics to observable characteristics (phenomes) and from preharvest stages through postharvest processing (Aziz et al., 2021; Li, 2025). At the heart of these advancements is image segmentation, a computer vision technique that divides digital images into meaningful segments,

significantly enhancing poultry monitoring and analysis (Minaee et al., 2022). With its potential to rapidly and accurately assess poultry health, behavior, and productivity, image segmentation has become an essential tool in modern precision animal management. This dissertation explores the application and enhancement of foundation segmentation models to further improve poultry phenotyping and behavioral monitoring, addressing critical needs within poultry production systems.

Overview of poultry phenotyping assessment

Poultry phenotyping involves the measurement and analysis of various physical and behavioral traits of birds to evaluate their health, welfare, and productivity. Commonly monitored physical phenotypes include body weight, growth rate, feather coverage, gait quality, body temperature, and overall body condition (Mortensen et al., 2017). These traits provide critical insights into the birds' physiological status, directly influencing production outcomes. For instance, regular monitoring of body weight and growth rate helps producers optimize feed efficiency and detect any growth abnormalities early (Emami et al., 2020).

Behavioral traits are equally important in phenotyping assessments, including activity levels, feeding and drinking patterns, social interactions, and flock cohesion. Monitoring these behaviors provides valuable data to evaluate poultry welfare and identify stress or discomfort signals that might indicate underlying health issues (Aydin et al., 2010; Kristensen et al., 2007). Traditional poultry phenotyping techniques predominantly rely on manual observations and measurements, which are not only labor-intensive but also highly susceptible to human error, subjective biases, and limited in scalability for commercial operations (Merenda et al., 2015).

To overcome these challenges, advanced automated technologies, such as electronic sensors and computer vision systems, have been increasingly adopted. Among these, image segmentation has become an indispensable technique due to its precision and versatility. Image segmentation facilitates the detailed tracking and analysis of individual birds without direct human intervention, thus minimizing stress and interference with natural behaviors (Saeidifar et al., 2024). Accurate segmentation supports automated morphological assessments, such as feather coverage analysis, gait evaluation, and the detection of physical deformities or injuries, significantly improving the accuracy of health assessments (Lamping et al., 2022; Nasiri et al., 2022).

Moreover, segmentation techniques enhance behavioral monitoring by allowing precise tracking of individual or group activities over extended periods. This capability aids in the early detection of abnormal behaviors such as reduced mobility, increased aggression, or altered feeding patterns, which could indicate health or welfare issues (Ejik et al., 2022; Li et al., 2020). By automating these processes, image segmentation not only improves data accuracy and consistency but also significantly reduces the time and labor required for comprehensive poultry phenotyping. Consequently, it supports better-informed management decisions, enhances animal welfare standards, and promotes sustainable productivity within intensive poultry production environments (Lamping et al., 2022; Ejik et al., 2022).

Overview of image segmentation techniques

Image segmentation is a fundamental task in computer vision that involves dividing an image into distinct and meaningful segments or regions. Each segment represents specific features or objects within the image, making segmentation crucial for detailed analysis and interpretation of visual data. Effective image segmentation allows for accurate object identification,

classification, and tracking, significantly benefiting numerous applications ranging from medical diagnostics and autonomous vehicles to agriculture and animal monitoring (Lei et al., 2022).

Below are the categories of different segmentation techniques that have been widely applied and studied across various domains:

Conventional image processing algorithms

Conventional image processing algorithms have long served as the foundation for early animal segmentation tasks (Yoon et al., 2007). These methods typically rely on manually crafted rules, filters, and basic mathematical models to extract relevant features from image data. Common techniques include thresholding, edge detection, region growing, morphological operations, and contour extraction (Dong et al., 2021). Although these approaches often require careful tuning and perform optimally in controlled environments, they remain valuable for their simplicity, computational efficiency, and ability to operate without extensive annotated training data (Yoon et al., 2007). Figure 1.1 shows chicken breast segmentation using conventional image processing algorithms.

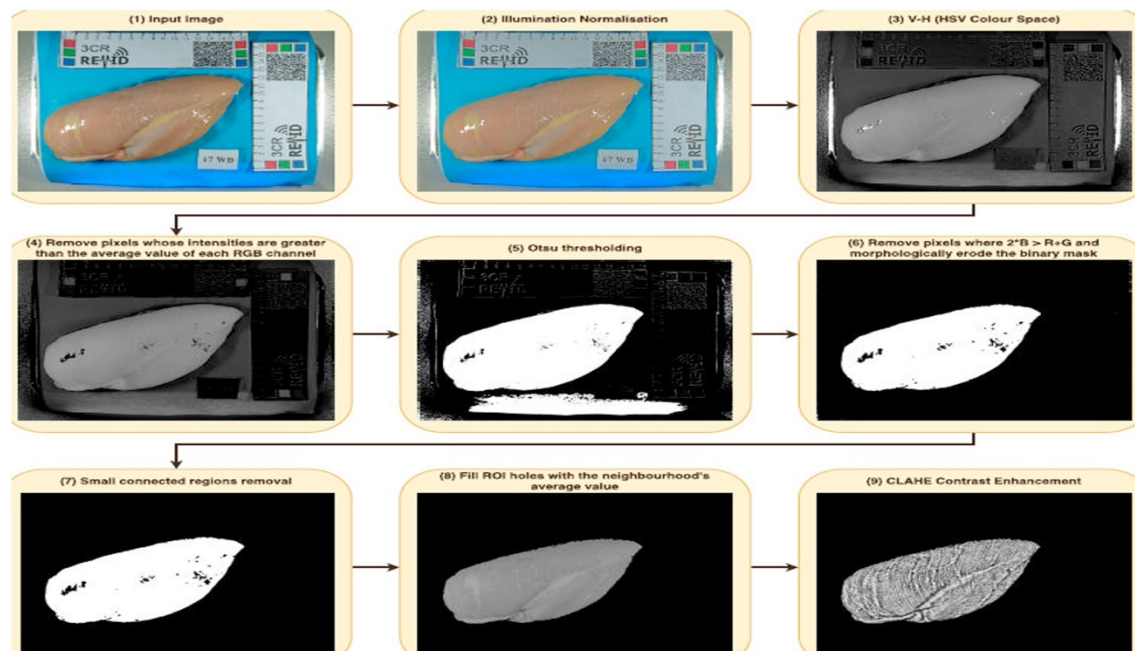


Fig. 1.1. Computer Vision System - Image processing steps (Geronimo et al. 2019).

In the context of poultry research, conventional image processing algorithms have been applied across various applications, including product quality inspection, behavior and welfare assessment, and animal trait recognition (Dong et al., 2021). These methods allow researchers to analyze specific physical features such as visceral contours, thermal patterns, body morphology, and fat content (Yoon et al., 2007). Despite the growing popularity of machine learning and deep learning-based segmentation, conventional approaches continue to play a role where cost, simplicity, or limited computational resources are constraints.

Table 1.1 summarizes selected studies from the literature that have utilized conventional image processing algorithms for poultry segmentation:

Table 1.1. Applications of conventional image processing algorithms for poultry segmentation

Author (year)	Subject	Segmentation method	Variables examined	Performance
Chen et al. (2018)	Chicken, duck	Active contour model, thresholding	Visceral contours	93.3% accuracy (chickens), 86.7% (ducks)
Del Valle et al. (2021)	Laying hens, broiler breeders	Hausdorff distance-based segmentation	Poultry movement, thermal comfort (Unrest Index)	Efficient thermal stress detection
Chen et al. (2023)	Chicken embryos	region-growing segmentation	Morphological evolution, allantois development	Effective and clear morphological segmentation
Li et al. (2020)	Broilers	thresholding and pixel intensity-based linear regression	Number of birds feeding and drinking, spatial and temporal behavior patterns	89–93% accuracy for detecting feeding, 93–95% accuracy for detecting drinking
Mansor et al. (2013)	Poultry meat	Mean-shift segmentation	U and V color components	Clear separation between halal and non-halal distributions
Rachmawanto et al. (2020)	Chicken eggs	Region extraction followed by feature extraction (HSV + GLCM)	Shell quality (good, defective, rotten)	85.71% accuracy using K=1 and distance metric d=2 or d=4
Le et al. (2023)	Duck	OTSU thresholding + contour extraction + morphological filtering	Egg yolk features, size	Accurate yolk segmentation and double yolk discrimination

Machine learning-based methods

In recent years, machine learning-based segmentation techniques have gained considerable attention in animal research due to their ability to model complex patterns, adapt to diverse datasets, and often provide higher accuracy compared to traditional rule-based methods (Philipsen et al. 2018). Unlike conventional image processing algorithms, machine learning approaches can automatically learn decision boundaries from training data, reducing the need for manual tuning and feature engineering. These methods typically require labeled datasets but are capable of generalizing across different conditions and capturing subtle variations in biological structures.

In the poultry domain, machine learning-based segmentation methods have been applied across multiple applications such as product quality inspection, health status identification, and trait recognition (Yoon et al. 2022). Various models, including Support Vector Machines (SVM) (Deng et al. 2018), Genetic Algorithms, and neural networks (Lamping et al. 2022), have been employed to process data ranging from RGB images to hyperspectral and OCT imaging. These techniques have enabled automated detection of shell defects, vascular features, tissue abnormalities, and embryonic development stages, contributing significantly to precision livestock farming.

Table 1.2 summarizes selected studies that have utilized machine learning-based segmentation approaches in poultry research:

Table 1.2. Applications of machine learning-based methods for poultry segmentation

Author (year)	Subject	Segmentation method	Variables examined	Performance
Mota-Grajales et al. (2019)	Poultry eggs	Artificial Neural Network - ANN) combined with structured light laser scanning	Eggshell defects (cracks, shape deformation)	97.5% segmentation accuracy
Geronimo et al. (2019)	Chicken (Broiler breast)	Support Vector Machine (SVM) based on texture and color features	Wooden Breast condition, physicochemical attributes	CVS 91.8%, NIR 97.5%
Zhu et al. (2021)	Chicken embryos	Genetic Algorithm optimized Backpropagation Neural Network (GA-BPNN) segmentation	Embryonic gender (vascular features)	89.74% gender prediction accuracy
Xu et al. (2010)	Chick embryo (CAM)	SVM segmentation using GLCM texture features	Vessel length, branching points	Faster and more accurate than global thresholding
Ekramirad et al. (2024)	Chicken	SVM and ensemble	Woody breast condition severity	95-100% accuracy
Triyanto et al. (2022)	Chicks	K-Means and Fuzzy C-Means clustering	Clustered flock regions	Fuzzy C-Means outperformed K-Means in segmentation quality

Interactive segmentation methods

Interactive segmentation methods represent a more recent advancement in computer vision, enabling users to directly guide the segmentation process with minimal manual intervention (Rother et al., 2004; Xu et al., 2016). Unlike fully automated algorithms, these approaches leverage user-provided prompts such as points, bounding boxes, or contours to refine segmentation outputs

(Xu et al., 2016). This flexibility allows for highly accurate segmentation results, even when working with complex or variable image data, while significantly reducing the annotation burden compared to fully manual labeling (Rother et al., 2004).

In poultry research, interactive segmentation methods have been successfully applied to monitor animal behavior, welfare, and movement patterns across multiple environments and data modalities. Recent developments in foundation models, such as the Segment Anything Model (SAM), allow for promptable segmentation with little or no prior training on specific animal datasets (Kirillov et al., 2023). This capability makes interactive segmentation especially appealing for multi-species applications and practical use in precision livestock farming, where datasets may be limited or highly diverse (Yang et al., 2023; Saeidifar et al., 2024). Figure 1.2 shows the bird segmentation results using SAM.

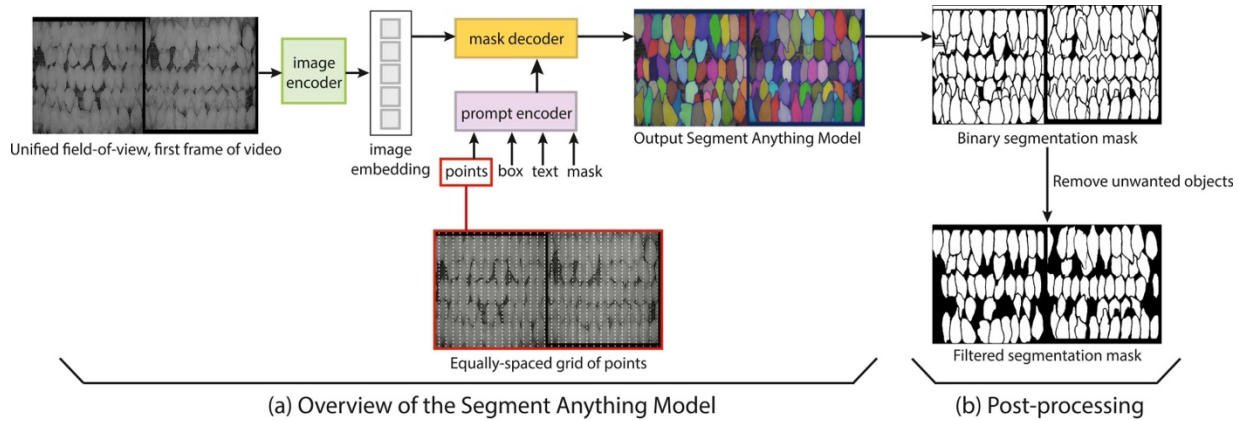


Fig. 1.2. Individual bird segmentation using Segment Anything Model and post-processing. Fig. 1.2(a) provides a high-level overview of the Segment Anything Model and Fig. 1.2(b) the post-processing step (Willems et al. 2025).

Deep learning-based semantic segmentation methods

Semantic segmentation with deep learning has become a dominant approach in recent years for animal monitoring applications due to its ability to assign class labels to every pixel in an image with high accuracy (Minaee et al., 2021). These methods leverage convolutional neural networks (CNNs), often in encoder-decoder architectures such as U-Net or its variants, to capture both global context and fine-grained spatial details (Minaee et al., 2021; Shi et al., 2024). Unlike traditional approaches, deep learning-based semantic segmentation models can automatically learn hierarchical features directly from image data, making them highly effective even in complex or variable environments (Minaee et al., 2021).

In poultry research, semantic segmentation has been applied to a wide range of tasks including health status identification, product quality inspection, and generalized behavior assessment (Sallam et al., 2024; Shi et al., 2024; Kou et al., 2024). These approaches have been used with various data modalities such as X-ray, CT, and RGB imagery, enabling researchers to automatically segment anatomical structures, detect defects, and quantify animal phenotypes with high precision. As these models continue to evolve, their ability to generalize across species, imaging conditions, and phenotyping tasks has made them increasingly valuable for precision livestock farming.

Table 1.3 summarizes selected studies that have utilized deep learning-based semantic segmentation methods in poultry research:

Table 1.3. Applications of deep learning-based semantic segmentation methods for poultry segmentation

Author (year)	Subject	Segmentation method	Variables examined	Performance
------------------	---------	------------------------	-----------------------	-------------

Sallam et al. (2024)	Laying hens	U-Net deep learning-based semantic segmentation	Keel bone geometry (fractures, deviations)	0.88-0.90 dice coefficient
Shi et al. (2023)	Chicken eggs	Improved U-Net using MobileNet-V2 backbone with CBAM attention for enhanced crack segmentation	Microcracks (<5µm, <20µm)	82.2% MIoU, 65.0% Crack-IoU
Shwetha V et al. (2025)	Poultry	ResUNet+ and DeepLabV3 (Xception backbone)	Crown, feather, leg regions	91.2% dice, 86.7% IoU
Zhu et al. (2024)	Chick embryo (CAM)	MAEFNet (Multistage Attention Enhancement Fusion Network)	Broiler chickens	96.35% IoU, 0.8810 R ²
Khanal et al. (2024)	Chicken	Pyramid Vision Transformer (PVT)	Chicken count, crowd density	96.9% accuracy, 27.8 MAE
Robinson et al. (2022)	Poultry (CAFO barns)	U-Net with ResNet-18 encoder	Barn location and footprint	87% recall, 83% precision

Deep learning-based instance segmentation methods

Deep learning-based instance segmentation has emerged as a highly effective approach for animal phenotyping and behavior monitoring, particularly in complex environments where multiple animals are present in the same frame (Lamping et al., 2022; Li et al., 2020). Unlike semantic segmentation, which assigns a class label to each pixel, instance segmentation distinguishes and tracks individual objects of the same class (van der Eijk et al., 2022), making it highly valuable for studies requiring individual animal tracking, interaction analysis, or behavior assessment.

In poultry research, instance segmentation methods are increasingly applied to monitor individual birds' behaviors, assess health conditions, and support precision management practices (Saeidifar et al., 2024; Lamping et al., 2022). Models such as Mask R-CNN, YOLO-based frameworks, and encoder-decoder CNNs have been used to accurately detect and segment individual animals across different environments and species (Li et al., 2020; van der Eijk et al., 2022). These approaches have demonstrated strong performance in tasks such as mating behavior detection, plumage assessment, multi-object tracking, and accurate counting in group-housed birds (Lamping et al., 2022).

Table 1.4 summarizes selected studies that have applied deep learning-based instance segmentation methods in poultry research:

Table 1.4. Applications of deep learning-based instance segmentation methods for poultry segmentation

Author (year)	Type of poultry	Segmentation method	Variables examined	Performance
Hu et al. (2024)	Broilers	Improved YOLOv8s model	Broiler behaviors (activity, wing spreading, resting, feeding, drinking)	99.50% detection mAP, 93.89% tracking MOTA, 93.98% overall behavior accuracy
Nasiri et al. (2024)	Broilers	Encoder-decoder CNN	Stretching, preening behaviors	96.7% accuracy, 88.1% precision, 89.96% recall
Zheng et al. (2022)	Breeder chickens	YOLOv5	Shank length, circumference	High accuracy; SD: length=1.35 mm, circumference=0.25 mm
Cao et al. (2024)	Caged broilers	Improved YOLOv5 with CSPDarknet53 backbone	Mortality status	99.2% accuracy; real-time capable

Raja et al. (2022)	Chicken legs	Mask R-CNN	6D pose (position and orientation)	89% correct pose estimation
-----------------------	--------------	------------	---------------------------------------	--------------------------------

Hybrid methods

Hybrid segmentation methods combine the strengths of both traditional image processing algorithms and modern machine learning or deep learning techniques to enhance segmentation performance (Zhang et al., 2023; Yusof et al., 2020). These approaches typically integrate conventional methods (such as thresholding, edge detection, morphological operations) with data-driven models (like neural networks or ensemble learning algorithms) (Saifullah & Suryotomo, 2021). The goal is often to leverage the precision and rule-based control of traditional techniques alongside the adaptability and generalization power of machine learning models.

In the context of poultry research, hybrid methods offer several advantages. For example, conventional algorithms can provide pre-processed or region-of-interest information that simplifies or guides the learning process for machine learning models (Zhang et al., 2023). Conversely, machine learning models can compensate for the limitations of conventional methods in more complex or noisy datasets by learning high-level feature representations (Abu Bakar et al., 2024). Applications of hybrid methods in poultry studies may include product quality inspection, behavior monitoring, and health condition assessment, particularly when image variability, lighting, or environmental noise make purely rule-based or purely data-driven approaches less effective (Saifullah & Suryotomo, 2021; Zhang et al., 2023; Yusof et al., 2020).

Table 1.5 summarizes selected studies that have applied hybrid methods in poultry research:

Table 1.5. Applications of hybrid methods for poultry segmentation

Author (year)	Type of poultry	Segmentation method	Variables examined	Performance
Willems et al. (2025)	Laying hens	Gaussian Mixture Modeling (GMM) + Segment Anything Model (SAM)	Night-time activity, perch occupation	Strong correlation (-0.84, $p < 0.0001$)
Zhao et al. (2025)	Cage-reared ducks	Mask R-CNN + CycleGAN	Behavior recognition, pose estimation	mAP improved by ~2%, FPS increased by 21%, significant OKS improvements
Yang et al. (2023)	Caged chickens	U-Net + pix2pixHD	Individual chicken contour, cage occlusion removal	94.71% wire mesh segmentation accuracy, 90.04% SSIM, 25.24 dB PSNR
Zhu et al. (2022)	Chicken eggs	Adaptive Canny operator + Ellipse fitting	Egg fertility status (infertile, dead-embryo)	100% classification accuracy for key categories
Chen et al. (2021)	Chicken	Thresholding + Active contour	Carcass and viscera position	91.3% carcass segmentation accuracy; 95.6% viscera
Saeidifar et al. (2024)	Cage-free laying hens	SAM + image processing + ML classifier	Body temperature	85.5% IoU, 92.3% F1

Applications of segmentation in precision animal management

The advancement of segmentation techniques has significantly expanded the possibilities in precision animal management (Minaee et al., 2021). These approaches have allowed for more efficient, accurate, and large-scale monitoring of animals under both experimental and commercial production conditions (Fernández et al., 2018; van der Eijk et al., 2022). By enabling the extraction

of detailed information from various imaging modalities, segmentation has improved the ability to monitor individual animals, assess group-level dynamics, and support real-time decision-making in diverse management contexts.

In modern production systems, the use of segmentation technologies has reduced labor demands, minimized subjective errors associated with manual assessments, and allowed for continuous, objective monitoring without disturbing animals (Fernández et al., 2018). This non-invasive approach not only enhances data reliability but also contributes to improved animal welfare by reducing handling and observation stress. The increasing integration of segmentation-based systems supports early detection of health issues, welfare concerns, and production inefficiencies, enabling producers to apply timely interventions and optimize overall system performance (van der Eijk et al., 2022; Saeidifar et al., 2024).

In the following subsections, the key application domains where segmentation plays a crucial role in precision animal management are outlined.

Behavior and Welfare Assessment

Behavioral monitoring is one of the primary applications of segmentation techniques in precision animal management. Continuous observation of animal behavior enables early detection of welfare issues such as heat stress, aggression, abnormal activity, or restricted mobility, which may otherwise go unnoticed under conventional farm management practices.

In poultry research, various segmentation methods have been applied to extract behavioral patterns at both the individual and group levels. Traditional image processing techniques have been employed to evaluate thermal comfort and movement disturbances. For example, Hausdorff distance-based segmentation was used to develop an unrest index that quantifies thermal discomfort in poultry flocks based on group movements (Del Valle et al., 2021). Similarly,

conventional segmentation approaches were applied to create a cluster index that distinguishes birds' thermal states and supports early detection of heat stress conditions (Pereira et al., 2020).

As segmentation techniques have advanced, machine learning and deep learning models have been introduced to improve behavior assessment accuracy, particularly under more complex conditions. A hybrid approach combining Mask R-CNN with CycleGAN was applied to enhance behavior recognition in cage-reared ducks, resulting in improved segmentation accuracy and real-time processing efficiency (Zhao et al., 2025). Instance segmentation models combined with behavioral classification algorithms were used to accurately identify mating events in group-housed broiler breeders, achieving an accuracy of 92% (Bodempudi et al., 2025). Furthermore, an improved YOLOv8s model was utilized to perform real-time detection and tracking of multiple broilers in cage-free systems, achieving high segmentation accuracy (mAP 99.50%) and effective multi-object tracking (MOTA 93.89%) (Hu et al., 2024).

Collectively, these studies demonstrate the critical role of segmentation technologies in improving the objectivity, resolution, and efficiency of behavior and welfare assessments in poultry production systems. By enabling automated, non-invasive monitoring, these approaches contribute to more proactive welfare management and support early intervention strategies that improve flock health and productivity.

Health Status Identification

Segmentation-based methods play a crucial role in the early detection and monitoring of health conditions in poultry production systems. By accurately isolating relevant anatomical structures or behavioral patterns, these methods provide objective and quantitative measures to assess the physical condition of individual animals or entire flocks, allowing for timely interventions and improved health management.

In poultry research, both conventional and advanced segmentation approaches have been applied across diverse imaging modalities to identify various health indicators. Early studies employed conventional image processing techniques, such as principal component analysis (PCA) applied to multispectral images, to detect conditions like septicemia and infected yolk sacs in chicken carcasses with high accuracy (Yang et al., 2005). Similarly, active contour-based segmentation methods have been used to monitor health-related behaviors and physical conditions in caged chickens, effectively extracting features from RGB and binocular vision images (Xiao et al., 2019).

With the development of deep learning, more sophisticated models have been introduced to enhance segmentation performance for health assessments. For example, ChickenNet, a deep learning model based on Mask R-CNN architecture, was applied to assess plumage condition in laying hens, achieving high detection accuracy (98.02% detection mAP and 91.83% plumage scoring accuracy) (Lamping et al., 2022). Building on this work, uncertainty estimation techniques were integrated into deep neural networks to improve segmentation robustness and confidence for health status evaluations in chickens (Lamping et al., 2023).

Hybrid segmentation approaches have also been utilized to address specific health monitoring challenges. For instance, Gaussian Mixture Models combined with the Segment Anything Model (SAM) were applied to analyze night-time perch occupation and activity patterns in laying hens as an early indicator for poultry red mite infestations, showing strong correlations with infestation levels (Willems et al., 2025).

Collectively, these studies demonstrate the increasing importance of segmentation methods in supporting precise, automated, and non-invasive health monitoring systems that contribute to improving animal welfare and production sustainability in commercial poultry operations.

Live Performance Prediction

Accurate and continuous monitoring of live performance metrics, such as body weight, growth rate, and production efficiency, is a key component of precision animal management. Segmentation methods enable non-invasive estimation of these parameters by automatically extracting morphological features from images or video data, reducing the need for manual weighing or labor-intensive measurements.

In poultry production, various segmentation approaches have been utilized to predict live performance indicators. Early work applied conventional image processing techniques, such as range-based watershed segmentation of 3D depth images, to estimate broiler weight, achieving a relative mean error of 7.8% (Mortensen et al., 2016). Similarly, camera-based monitoring using HSV and Lab color space segmentation was used to estimate weight and growth of geese in real-time under commercial farm conditions (Toth et al., 2025).

Hybrid methods have also been introduced to improve accuracy by combining multiple segmentation and feature extraction algorithms. For example, Chan-Vese segmentation combined with ellipse fitting was used to segment broiler body contours, resulting in highly accurate weight predictions with an R^2 of 0.98 and prediction errors typically below 50 grams (Amraei et al., 2017).

More recently, deep learning-based segmentation models have been employed to enhance live performance monitoring under more complex conditions. YOLOv8 instance segmentation was applied to automatically extract body regions of broilers for live weight prediction, achieving a mean average precision (mAP) of 0.829 (Shams et al., 2025). Additionally, 3D convolutional neural networks (3D CNNs) were used to process top-view video frames for broiler weight estimation, reaching a prediction accuracy of 95% (Anuprabha et al., 2024).

Together, these studies demonstrate the growing potential of segmentation-based systems to provide real-time, automated, and scalable solutions for live performance prediction, enabling more efficient management and optimization of poultry production systems.

Product Quality Inspection

Segmentation methods have also been extensively applied in product quality inspection, enabling precise assessment of carcass features, tissue characteristics, and product defects. By automating the extraction of critical quality attributes, segmentation-based systems contribute to more consistent grading, improved product safety, and reduced labor demands in poultry processing facilities.

In poultry carcass inspection, conventional image processing techniques have been widely used to extract anatomical contours and identify internal organs. For example, contour-based segmentation methods were applied to recognize visceral features in poultry carcasses, achieving classification accuracies of 93.3% for chickens and 86.7% for ducks (Chen et al., 2018). Similarly, conventional approaches based on color and texture analysis have been utilized for fat estimation and body composition evaluation, demonstrating reliable performance under varying image conditions (Chmiel et al., 2011).

Machine learning-based methods have also been introduced to improve defect detection accuracy. A system using machine learning-based segmentation was developed to identify eggshell defects in poultry eggs with 97.5% accuracy using laser structured light images (Mota-Grajales et al., 2019).

Additional conventional approaches have been used for advanced color-based feature extraction. RGB-based segmentation combined with color space transformations was employed to estimate the color parameters of chicken breast fillets, achieving high correlations with manual

color measurements (Barbin et al., 2016). Statistical texture features from multiple color spaces (RGB, Lab*, XYZ, S, V, U) were applied to classify cold meats, achieving classification accuracies ranging from 89% to 100% (Zapotoczny et al., 2016).

Collectively, these studies demonstrate the strong potential of segmentation technologies to enhance product quality inspection by enabling consistent, objective, and high-throughput evaluation of carcass and product attributes.

Animal Trait Recognition

Segmentation methods have also been applied to support automated recognition and analysis of animal traits, such as morphology, sex, skeletal structure, and other phenotypic characteristics that are important for breeding, welfare, and production optimization.

In poultry research, conventional image processing techniques have been successfully used for morphological analysis of embryos and skeletal structures. Region-growing segmentation algorithms were applied to extract 3D morphological features from MRI images of chick embryos, providing effective and detailed anatomical segmentation (Chen et al., 2023). Similarly, automated segmentation algorithms utilizing dual-thresholding approaches were implemented to analyze micro-CT images of egg-laying hens for detailed bone structure evaluation (Chen et al., 2020).

Machine learning-based methods have also been introduced to support trait recognition tasks. A Genetic Algorithm Optimized Backpropagation Neural Network was employed to identify the gender of chicken embryos based on vascular image features, achieving a prediction accuracy of 89.74% (Zhu et al., 2021).

Deep learning-based instance segmentation models have further improved trait extraction accuracy. YOLOv5 was utilized to measure shank length and circumference of breeder chickens

from RGB images, providing high accuracy with standard deviations of 1.35 mm for length and 0.43 mm for circumference measurements (Zheng et al., 2022).

Collectively, these studies highlight the potential of segmentation-based systems to support non-invasive, precise, and efficient trait recognition in poultry, contributing to genetic selection, welfare monitoring, and breeding program optimization.

Objectives and Outline of the Dissertation

The primary objective of this dissertation was to develop and evaluate advanced segmentation frameworks utilizing foundation models, deep learning, and hybrid image processing techniques to enhance automated monitoring of poultry welfare, health, and behavior. Specifically, the research focused on leveraging recent advancements in zero-shot and prompt-based segmentation models to address key challenges in individual animal tracking, thermal condition monitoring, and phenotypic trait extraction in precision poultry management. These models were integrated with both RGB and thermal imaging modalities to create efficient, non-invasive, and scalable solutions suitable for commercial and research applications.

This dissertation consists of several independent yet connected studies.

- In Chapter I, a comprehensive literature review was conducted to systematically analyze existing segmentation methodologies applied in precision animal management, categorizing studies based on segmentation types, applications, and species.
- In Chapter II, zero-shot segmentation models were applied and optimized to segment cage-free laying hens from thermal images. By integrating image processing and machine learning-based post-processing with foundation segmentation models (SAM, FastSAM, MobileSAM), this study provided a fully automated pipeline for monitoring surface body temperature as an indicator of health and thermal regulation.

- In Chapter III, an integrated approach combining object detection (YOLOv7) and foundation image segmentation (SAM) was proposed to automate segmentation of birds in diverse imaging conditions, eliminating the need for extensive manual annotation and enhancing segmentation accuracy across various visual contexts.
- In Chapter IV, a novel open-source Streamlit-based platform was developed that allows non-technical users to calculate animal activity indices from top-view videos using a promptable foundation segmentation model (SAM2). This platform demonstrated its capacity for efficient behavior monitoring of poultry flocks while significantly reducing labor and computational requirements.
- The results from these chapters are summarized in Chapter V and the conclusions provide hands-on suggestions on precision poultry farming technologies.

Except for Chapter I, each chapter was prepared as independent research papers for peer-reviewed journals and conferences. The final chapter summarizes the key findings and provides concluding remarks along with future research directions that can contribute to the broader advancement of precision poultry farming technologies.

References

- Ahmed, G., Malick, R.A., Akhunzada, A., et al., 2021. An approach towards IoT-based predictive service for early detection of diseases in poultry chickens. *Sustainability* 13. <https://doi.org/10.3390/su132313396>
- Amraei, S., Abdanan Mehdizadeh, S., & Salari, S., 2017. Broiler weight estimation based on machine vision and artificial neural network. *British Poultry Science* 58(2), 200–205. <https://doi.org/10.1080/00071668.2016.1259530>

Asali, E., Li, G., Saeidifar, M., Liu, T., Bodempudi, V.U.C., Oso, O.M., Mandiga, A., Kota, S.A.R., Yuan, G., 2025. Zero-Shot Perception and Spatiotemporal Transformers for Automated Gait and Footpad Score Classification. Paper presented at the The 23rd International Conference on Scientific Computing, Las Vegas, NV, USA.

Aydin, A., Cangar, O., Ozcan, S. E., Bahr, C., & Berckmans, D., 2010. Application of a fully automatic analysis tool to assess the activity of broiler chickens with different gait scores. *Computers and Electronics in Agriculture* 73(2), 194–199. <https://doi.org/10.1016/j.compag.2010.05.004>

Aziz, N.S.N.A., Daud, S.M., Dziyauddin, R.A., et al., 2021. A review on computer vision technology for monitoring poultry farm—Application, hardware, and software. *IEEE Access* 9, 12431–12445. <https://doi.org/10.1109/ACCESS.2020.3047818>

Banakar, A., Sharafi, M., Li, G., 2024. Relationship between reproductive indicators and sound structure in broiler breeder roosters. *Applied Animal Behaviour Science* 270, 106135. <https://doi.org/10.1016/j.applanim.2023.106135>

Barbin, D. F., Mastelini, S. M., Barbon, S., Campos, G. F. C., Barbon, A. P. A. C., & Shimokomaki, M., 2016. Digital image analyses as an alternative tool for chicken quality assessment. *Biosystems Engineering* 144, 85–93. <https://doi.org/10.1016/j.biosystemseng.2016.01.015>

Berckmans, D., 2017. General introduction to precision livestock farming. *Animal Frontiers* 7, 6–11. <https://doi.org/10.2527/af.2017.0102>

Bodempudi, V. U. C., Li, G., Mason, J. H., Wilson, J. L., Liu, T., & Rasheed, K. M., 2025. Identifying mating events of group-housed broiler breeders via bio-inspired deep learning models. *Poultry Science* 104(7), 105126. <https://doi.org/10.1016/j.psj.2025.105126>

Brown Jordan, A., Gongora, V., Hartley, D., Oura, C., 2018. A review of eight high-priority, economically important viral pathogens of poultry within the Caribbean region. *Veterinary Sciences* 5. <https://doi.org/10.3390/vetsci5010014>

C. Yoon, S., C. Lawrence, K., Park, B., & R. Windham, W., 2007. Statistical Model-Based Thresholding of Multispectral Images for Contaminant Detection on Poultry Carcasses. *Transactions of the ASABE* 50(4), 1433–1442. <https://doi.org/10.13031/2013.23616>

Centers for Disease Control and Prevention, 2025. H5 Bird Flu: Current Situation. Accessed on April 2025. Available at: <https://www.cdc.gov/bird-flu/situation-summary/index.html>

Chemme, K.S., Alitappeh, R.J., 2024. An End-to-End Model for Chicken Detection in a Cluttered Environment, 2024 13th Iranian/3rd International Machine Vision and Image Processing Conference (MVIP), pp. 1–7.

Chen, C., & Kim, W. K., 2020. The application of micro-CT in egg-laying hen bone analysis: Introducing an automated bone separation algorithm. *Poultry Science* 99(11), 5175–5183. <https://doi.org/10.1016/j.psj.2020.08.047>

Chen, L., Wang, Z., Fu, X., Wang, S., Feng, Y., Coudyzer, W., Wu, S., Zhang, H., Chai, Z., Li, Y., & Ni, Y., 2023. Dynamic 3D morphology of chick embryos and allantois depicted nondestructively by 3.0T clinical magnetic resonance imaging. *Poultry Science* 102(9), 102902. <https://doi.org/10.1016/j.psj.2023.102902>

Chen, Y., Ai, H., & Li, S., 2021. Analysis of correlation between carcass and viscera for chicken eviscerating based on machine vision technology. *Journal of Food Process Engineering* 44(1), e13592. <https://doi.org/10.1111/jfpe.13592>

- Chen, Y., & Wang, S. C., 2018. Poultry carcass visceral contour recognition method using image processing. *Journal of Applied Poultry Research* 27(3), 316–324. <https://doi.org/10.3382/japr/pfx073>
- Chmiel, M., Słowiński, M., & Dasiewicz, K., 2011. Application of computer vision systems for estimation of fat content in poultry meat. *Food Control* 22(8), 1424–1427. <https://doi.org/10.1016/j.foodcont.2011.03.002>
- Del Valle, J. E., Pereira, D. F., Mollo Neto, M., Gabriel Filho, L. R. A., & Salgado, D. D., 2021. Unrest index for estimating thermal comfort of poultry birds (*Gallus gallus domesticus*) using computer vision techniques. *Biosystems Engineering* 206, 123–134. <https://doi.org/10.1016/j.biosystemseng.2021.03.018>
- Deng, X., Wan, Q., Wu, L., Gao, H., Wen, Y., & Wang, S., 2018. Eggshell crack detection by acoustic impulse response and support vector machine. *Advances in Agriculture and Agricultural Sciences* 4(9), 1–9. <https://www.internationaljournal.com/abstract/eggshell-crack-detection-by-acoustic-impulse-response-and-support-vector-machine-56049.html>
- Dong, J., Lu, B., He, K., Li, B., Zhao, B., & Tang, X., 2021. Assessment of hatching properties for identifying multiple duck eggs on the hatching tray using machine vision technique. *Computers and Electronics in Agriculture* 184, 106076. <https://doi.org/10.1016/j.compag.2021.106076>
- Dou, F., Ye, J., Yuan, G., Lu, Q., Niu, W., Sun, H., Guan, L., Lu, G., Mai, G., Liu, N., 2023. Towards artificial general intelligence (agi) in the internet of things (iot): Opportunities and challenges. *arXiv preprint arXiv:2309.07438*.
- Ekramirad, N., Yoon, S.-C., Bowker, B. C., & Zhuang, H., 2024. Nondestructive Assessment of Woody Breast Myopathy in Chicken Fillets Using Optical Coherence Tomography Imaging with

Machine Learning: A Feasibility Study. *Food and Bioprocess Technology* 17(11), 4053–4070.

<https://doi.org/10.1007/s11947-024-03369-1>

Emami, N. K., Calik, A., White, M. B., Kimminau, E. A., & Dalloul, R. A., 2020. Effect of Probiotics and Multi-Component Feed Additives on Microbiota, Gut Barrier and Immune Responses in Broiler Chickens During Subclinical Necrotic Enteritis. *Frontiers in Veterinary Science* 7. <https://doi.org/10.3389/fvets.2020.572142>

Eterradossi, N., Saif, Y.M., 2013. Infectious Bursal Disease. In: *Diseases of Poultry*, pp. 219–246.

Gagnaux, D., Poffet, J., 2000. *Livestock Farming Systems: Integrating Animal Science Advances in the Search of Sustainability*. BRILL.

Geronimo, B. C., Mastelini, S. M., Carvalho, R. H., Barbon Júnior, S., Barbin, D. F., Shimokomaki, M., & Ida, E. I., 2019. Computer vision system and near-infrared spectroscopy for identification and classification of chicken with wooden breast, and physicochemical and technological characterization. *Infrared Physics & Technology* 96, 303–310. <https://doi.org/10.1016/j.infrared.2018.11.036>

Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M., Toulmin, C., 2010. Food security: the challenge of feeding 9 billion people. *Science* 327, 812–818.

Guina, W., Zhen, L., Guirong, W., Yiyang, C., 2025. An overview of industrial image segmentation using deep learning models. *Intelligence & Robotics* 5, 143–180. <https://doi.org/10.20517/ir.2025.09>

Hu, Y., Xiong, J., Xu, J., Gou, Z., Ying, Y., Pan, J., & Cui, D., 2024. Behavior recognition of cage-free multi-broilers based on spatiotemporal feature learning. *Poultry Science* 103(12), 104314. <https://doi.org/10.1016/j.psj.2024.104314>

- K, A., Kaliappan, V. K., & Baniekal Hiremath, G., 2024. Transformative Technology in Poultry Management: 3D CNNs for Broiler Chicken Weight Prediction. 2024 International Conference on Cognitive Robotics and Intelligent Systems (ICC - ROBINS), 432–436. <https://doi.org/10.1109/ICC-ROBINS60238.2024.10534012>
- Khanal, R., Choi, Y., & Lee, J., 2024. Transforming Poultry Farming: A Pyramid Vision Transformer Approach for Accurate Chicken Counting in Smart Farm Environments. Sensors 24(10). <https://doi.org/10.3390/s24102977>
- Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P., & Girshick, R., 2023. Segment Anything. arXiv:2304.02643. <https://doi.org/10.48550/arXiv.2304.02643>
- Kotaridis, I., Lazaridou, M., 2021. Remote sensing image segmentation advances: A meta-analysis. ISPRS Journal of Photogrammetry and Remote Sensing 173, 309–322. <https://doi.org/10.1016/j.isprsjprs.2021.01.020>
- Kou, X., Yang, Y., Xue, H., Wang, L., & Li, L., 2024. Revealing in vivo broiler chicken growth state: Integrating CT imaging and deep learning for non-invasive reproductive phenotypic measurement. Computers and Electronics in Agriculture 227, 109477. <https://doi.org/10.1016/j.compag.2024.109477>
- Kristensen, H. H., Perry, G.C., Prescott, N.B., Ladewig, J., Ersbøll, A.K., & Wathes, C. M., 2006. Leg health and performance of broiler chickens reared in different light environments. British Poultry Science 47(3), 257–263. <https://doi.org/10.1080/00071660600753557>
- Lamping, C., Derks, M., Groot Koerkamp, P., & Kootstra, G., 2022. ChickenNet—An end-to-end approach for plumage condition assessment of laying hens in commercial farms using computer

vision. *Computers and Electronics in Agriculture* 194, 106695. <https://doi.org/10.1016/j.compag.2022.106695>

Le, C. J., Dan, L., Cheng, W. J., & Jie, Y. M., 2025. Double yolk duck egg feature discrimination and size grading based on machine vision and CH-GO rule. *Journal of Food Measurement and Characterization* 19(3), 1662–1672. <https://doi.org/10.1007/s11694-024-03062-z>

Lei, T., Li, G., Luo, C., Zhang, L., Liu, L., Gates, R.S., 2022. An informative planning-based multi-layer robot navigation system as applied in a poultry barn. *Intelligence & Robotics* 2, 313–332. <https://doi.org/10.20517/ir.2022.18>

Li, G., 2025. A survey of open-access datasets for computer vision in precision poultry farming. *Poultry Science* 104, 104784. <https://doi.org/10.1016/j.psj.2025.104784>

Li, G., Chai, L., 2023. AnimalAccML: An open-source graphical user interface for automated behavior analytics of individual animals using triaxial accelerometers and machine learning. *Computers and Electronics in Agriculture* 209, 107835. <https://doi.org/10.1016/j.compag.2023.107835>

Li, G., Huang, Y., Chen, Z., Chesser, G.D., Purswell, J.L., Linhoss, J., Zhao, Y., 2021b. Practices and Applications of Convolutional Neural Network-Based Computer Vision Systems in Animal Farming: A Review. *Sensors* 21. <https://doi.org/10.3390/s21041492>

Li, G., Hui, X., Lin, F., & Zhao, Y., 2020. Developing and Evaluating Poultry Preening Behavior Detectors via Mask Region-Based Convolutional Neural Network. *Animals* 10(10). <https://doi.org/10.3390/ani10101762>

Li, G., Zhao, Y., Hailey, R., et al., 2019. An ultra-high frequency radio frequency identification system for studying individual feeding and drinking behaviors of group-housed broilers. *Animal* 13, 2060–2069. <https://doi.org/10.1017/S1751731118003440>

Li, G., Zhao, Y., Purswell, J.L., Du, Q., Chesser, G.D., Lowe, J.W., 2020b. Analysis of feeding and drinking behaviors of group-reared broilers via image processing. *Computers and Electronics in Agriculture* 175, 105596. <https://doi.org/10.1016/j.compag.2020.105596>

Li, L., Zhang, Q., Huang, D., 2014. A Review of Imaging Techniques for Plant Phenotyping. *Sensors* 14, 20078–20111. <https://doi.org/10.3390/s141120078>

Liu, K., Xin, H., Chai, L., 2017. Choice Between Fluorescent and Poultry-Specific LED Lights by Pullets and Laying Hens. *Transactions of the ASABE* 60, 2185–2195. <https://doi.org/10.13031/trans.12402>

Mansor, M. A., Baki, S. R. M. S., Tahir, N. M., & Rahman, R. A., 2013. An approach of halal poultry meat comparison based on mean-shift segmentation. 2013 IEEE Conference on Systems, Process & Control (ICSPC), 279–282. <https://doi.org/10.1109/SPC.2013.6735147>

Merenda, V. R., Bodempudi, V. U. C., Pairis-Garcia, M. D., & Li, G., 2024. Development and validation of machine-learning models for monitoring individual behaviors in group-housed broiler chickens. *Poultry Science* 103(12), 104374. <https://doi.org/10.1016/j.psj.2024.104374>

Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D., 2020. Image Segmentation Using Deep Learning: A Survey. *arXiv:2001.05566*. <https://doi.org/10.48550/arXiv.2001.05566>

Minaee, S., Boykov, Y., Porikli, F., et al., 2022. Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, 3523–3542. <https://doi.org/10.1109/TPAMI.2021.3059968>

Mortensen, A. K., Lisouski, P., & Ahrendt, P., 2016. Weight prediction of broiler chickens using 3D computer vision. *Computers and Electronics in Agriculture* 123, 319–326. <https://doi.org/10.1016/j.compag.2016.03.011>

Mota-Grajales, R., Torres-Peña, J. C., Camas-Anzueto, J. L., Pérez-Patricio, M., Grajales Coutiño, R., López-Estrada, F. R., Escobar-Gómez, E. N., & Guerra-Crespo, H., 2019. Defect detection in eggshell using a vision system to ensure the incubation in poultry production. *Measurement* 135, 39–46. <https://doi.org/10.1016/j.measurement.2018.09.059>

Nasiri, A., Yoder, J., Zhao, Y., Hawkins, S., Prado, M., & Gan, H., 2022. Pose estimation-based lameness recognition in broiler using CNN-LSTM network. *Computers and Electronics in Agriculture* 197, 106931. <https://doi.org/10.1016/j.compag.2022.106931>

Nyalala, I., Okinda, C., Makange, N., Korohou, T., Chao, Q., Nyalala, L., Jiayu, Z., Yi, Z., Yousaf, K., Chao, L., Kunjie, C., 2021. On-line weight estimation of broiler carcass and cuts by a computer vision system. *Poultry Science* 100, 101474. <https://doi.org/10.1016/j.psj.2021.101474>

Okada, H., Itoh, T., Suzuki, K., Tsukamoto, K., 2009. Wireless sensor system for detection of avian influenza outbreak farms at an early stage. In: *SENSORS, 2009 IEEE*, pp. 1374–1377.

Oladeinde, A., Chung, T., Mou, C., et al., 2025. Broiler litter moisture and trace metals contribute to the persistence of Salmonella strains that harbor large plasmids carrying siderophores. *Applied and Environmental Microbiology* 0, e01388–24. <https://doi.org/10.1128/aem.01388-24>

Pasick, J., Berhane, Y., Joseph, T., Bowes, V., Hisanaga, T., Handel, K., Alexandersen, S., 2015. Reassortant Highly Pathogenic Influenza A H5N2 Virus Containing Gene Segments Related to Eurasian H5N8 in British Columbia, Canada, 2014. *Scientific Reports* 5, 9484. <https://doi.org/10.1038/srep09484>

Peña Fernández, A., Norton, T., Tullo, E., van Hertem, T., Youssef, A., Exadaktylos, V., Vranken, E., Guarino, M., & Berckmans, D., 2018. Real-time monitoring of broiler flock's welfare status using camera-based technology. *Biosystems Engineering* 173, 103–114. <https://doi.org/10.1016/j.biosystemseng.2018.05.008>

Pereira, D. F., Lopes, F. A. A., Filho, L. R. A. G., Salgado, D. D., & Neto, M. M., 2020. Cluster index for estimating thermal poultry stress (*gallus gallus domesticus*). *Computers and Electronics in Agriculture* 177, 105704. <https://doi.org/10.1016/j.compag.2020.105704>

Philipsen, M. P., Dueholm, J. V., Jørgensen, A., Escalera, S., & Moeslund, T. B., 2018. Organ Segmentation in Poultry Viscera Using RGB-D. *Sensors* 18(1). <https://doi.org/10.3390/s18010117>

Rachmawanto, E. H., Atika Sari, C., Villadelfiya, R., Ignatius Moses Setiadi, D. R., Rijati, N., Kartikadarma, E., Doheir, M., & Astuti, S., 2020. Eggs Classification based on Egg Shell Image using K-Nearest Neighbors Classifier. 2020 International Seminar on Application for Technology of Information and Communication (iSemantic), 50–54. <https://doi.org/10.1109/iSemantic50169.2020.9234305>

Raj, G.D., Jones, R.C., 1997. Infectious bronchitis virus: Immunopathogenesis of infection in the chicken. *Avian Pathology* 26, 677–706. <https://doi.org/10.1080/03079459708419246>

Raja, R., Rincón, D. R., Kemp, F., Kootstra, G., & Henten, E. V., 2022. PosePP: 6D Pose Estimation of Poultry Pieces in a Cluttered Bin. 2022 IEEE 10th Jubilee International Conference on Computational Cybernetics and Cyber-Medical Systems (ICCC), 000239–000244. <https://doi.org/10.1109/ICCC202255925.2022.9922692>

Ramesh, K., Kumar, G.K., Swapna, K., Datta, D., Rajest, S.S., 2021. A review of medical image segmentation algorithms. *EAI Endorsed Transactions on Pervasive Health & Technology* 7.

Robinson, C., Chugg, B., Anderson, B., Ferres, J. M. L., & Ho, D. E., 2022. Mapping Industrial Poultry Operations at Scale With Deep Learning and Aerial Imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 15, 7458–7471. <https://doi.org/10.1109/JSTARS.2022.3191544>

- Rother, C., Kolmogorov, V., & Blake, A., 2004. “GrabCut”: Interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics* 23(3), 309–314.
<https://doi.org/10.1145/1015706.1015720>
- Sadeghi, M., Banakar, A., Minaei, S., Orooji, M., Shoushtari, A., Li, G., 2023. Early Detection of Avian Diseases Based on Thermography and Artificial Intelligence. *Animals* 13.
<https://doi.org/10.3390/ani13142348>
- Saeidifar, M., Li, G., Chai, L., Bist, R., Rasheed, K.M., Lu, J., Banakar, A., Liu, T., Yang, X., 2024. Zero-shot image segmentation for monitoring thermal conditions of individual cage-free laying hens. *Computers and Electronics in Agriculture* 226, 109436.
<https://doi.org/10.1016/j.compag.2024.109436>
- Saifullah, S., & Suryotomo, A. P., 2021. Identification of chicken egg fertility using SVM classifier based on first-order statistical feature extraction. *ILKOM Jurnal Ilmiah* 13(3), 285–293.
<https://doi.org/10.33096/ilkom.v13i3.937.285-293>
- Sallam, M., Flores, S. C., de Koning, D. J., & Johnsson, M., 2024. Research Note: A deep learning method segments chicken keel bones from whole-body X-ray images. *Poultry Science* 103(11), 104214. <https://doi.org/10.1016/j.psj.2024.104214>
- Seo, H., Badiei Khuzani, M., Vasudevan, V., Huang, C., Ren, H., Xiao, R., Jia, X., Xing, L., 2020. Machine learning techniques for biomedical image segmentation: An overview of technical aspects and introduction to state-of-art applications. *Medical Physics* 47, e148–e167.
<https://doi.org/10.1002/mp.13649>
- Shams, M. Y., Elmessery, W. M., Oraiath, A. A. T., Elbeltagi, A., Salem, A., Kumar, P., El-Messery, T. M., El-Hafeez, T. A., Abdelshafie, M. F., Abd El-Wahhab, G. G., El-Soaly, I. S., & Elwakeel, A. E., 2025. Automated on-site broiler live weight estimation through YOLO-based

segmentation. Smart Agricultural Technology 10, 100828.

<https://doi.org/10.1016/j.atech.2025.100828>

Shi, C., Li, Y., Jiang, X., Sun, W., Zhu, C., Mo, Y., Yan, S., & Zhang, C., 2024. Real-Time ConvNext-Based U-Net with Feature Infusion for Egg Microcrack Detection. Agriculture 14(9).

<https://doi.org/10.3390/agriculture14091655>

Shriner, S.A., Root, J.J., Lutman, M.W., et al., 2016. Surveillance for highly pathogenic H5 avian influenza virus in synanthropic wildlife associated with poultry farms during an acute outbreak. Scientific Reports 6, 36237. <https://doi.org/10.1038/srep36237>

Tao, Y., Chen, Z., Jing, H., Walker, J., 2001. Internal inspection of deboned poultry using X-ray imaging and adaptive thresholding. Transactions of the ASAE 44, 1005.

<https://doi.org/10.13031/2013.6233>

Toth, G., Szabo, S., Haidegger, T., & Alexy, M., 2025. Edge or Cloud Architecture: The Applicability of New Data Processing Methods in Large-Scale Poultry Farming. Technologies 13(1). <https://doi.org/10.3390/technologies13010017>

Trading Economics, 2025. Eggs US. Accessed on April 2025. Available at: <https://tradingeconomics.com/commodity/eggs-us>

Tran, M., Truong, S., Fernandes, A.F.A., Kidd, M.T., Le, N., 2024. CarcassFormer: an end-to-end transformer-based framework for simultaneous localization, segmentation and classification of poultry carcass defect. Poultry Science 103, 103765. <https://doi.org/10.1016/j.psj.2024.103765>

Triyanto, W. A., Setiawan, A., Setiawan, A., Warsito, B., & Wibowo, A., 2022. Digital Image Segmentation of Chicks Flock Using Clustering Method on Fog Computing Network. 2022 2nd International Conference on Information Technology and Education (ICIT&E), 140–144. <https://doi.org/10.1109/ICITE54466.2022.9759871>

Tsiouris, V., Georgopoulou, I., Batzios, C., Pappaioannou, N., Ducatelle, R., Fortomaris, P., 2015. High stocking density as a predisposing factor for necrotic enteritis in broiler chicks. *Avian Pathology* 44, 59–66.

USDA Animal and Plant Health Inspection Service, 2025. Confirmations of Highly Pathogenic Avian Influenza in Commercial and Backyard Flocks. Accessed on April 2025. Available at: <https://www.aphis.usda.gov/livestock-poultry-disease/avian/avian-influenza/hpai-detections/commercial-backyard-flocks>

USDA National Agricultural Statistics Service, 2025. Poultry Production and Value. Accessed on May 2025. Available at: <https://usda.library.cornell.edu/concern/publications/m039k491c>

V, S., B s, M., Adil, S., Laxmi, V., & Shrivastava, S., 2025. GAN-based motion blur elimination as a preprocessing step for enhanced AI-driven computer-aided camera monitoring in poultry and free-range farming in low-resource settings. *Smart Agricultural Technology* 11, 100915. <https://doi.org/10.1016/j.atech.2025.100915>

van der Eijk, J. A. J., Guzhva, O., Voss, A., Möller, M., Giersberg, M. F., Jacobs, L., & de Jong, I. C., 2022. Seeing is caring – automated assessment of resource use of broilers with computer vision techniques. *Frontiers in Animal Science* 3. <https://doi.org/10.3389/fanim.2022.945534>

van der Eijk, J.A.J., Guzhva, O., Schulte-Landwehr, J., Giersberg, M.F., Jacobs, L., de Jong, I.C., 2023. Individuality of a group: detailed walking ability analysis of broiler flocks using optical flow approach. *Smart Agricultural Technology* 5, 100298. <https://doi.org/10.1016/j.atech.2023.100298>

Vidic, J., Manzano, M., Chang, C.-M., Jaffrezic-Renault, N., 2017. Advanced biosensors for detection of pathogens related to livestock and poultry. *Veterinary Research* 48, 11. <https://doi.org/10.1186/s13567-017-0418-5>

- Willems, S., Mounir, M., Van Hertem, T., Nijs, H., Sleenckx, N., & Norton, T., 2025. Monitoring night-time activity and perch occupation of laying hens using night-vision cameras for the early detection of poultry red mite: A proof of concept. *Applied Animal Behaviour Science* 283, 106517. <https://doi.org/10.1016/j.applanim.2025.106517>
- Xiao, L., Ding, K., Gao, Y., & Rao, X., 2019. Behavior-induced health condition monitoring of caged chickens using binocular vision. *Computers and Electronics in Agriculture* 156, 254–262. <https://doi.org/10.1016/j.compag.2018.11.022>
- Xu, N., Price, B., Cohen, S., Yang, J., & Huang, T. S., 2016. Deep Interactive Object Selection. 373–381. https://openaccess.thecvf.com/content_cvpr_2016/html/Xu_Deep_Interactive_Object_CVPR_2016_paper.html
- Yang, C.-C., Chao, K., & Chen, Y.-R., 2005. Development of multispectral image processing algorithms for identification of wholesome, septicemic, and inflammatory process chickens. *Journal of Food Engineering* 69(2), 225–234. <https://doi.org/10.1016/j.jfoodeng.2004.07.021>
- Yang, J., Zhang, T., Fang, C., & Zheng, H., 2023. A defencing algorithm based on deep learning improves the detection accuracy of caged chickens. *Computers and Electronics in Agriculture* 204, 107501. <https://doi.org/10.1016/j.compag.2022.107501>
- Yang, X., Dai, H., Wu, Z., Bist, R., Subedi, S., Sun, J., Lu, G., Li, C., Liu, T., & Chai, L., 2023. SAM for Poultry Science. arXiv:2305.10254. <https://doi.org/10.48550/arXiv.2305.10254>
- Yusof, R., Arfa, R., Yunus, M. R. M., Muhammad, N. A. A. N., Rosli, N. R., & Ismail, N., 2020. Image Segmentation and Verification Based on Machine Learning for Vision Inspection of Chicken Slaughtering. *Journal of Physics: Conference Series* 1447(1), 012024. <https://doi.org/10.1088/1742-6596/1447/1/012024>

- Zapotoczny, P., Szczypiński, P. M., & Daszkiewicz, T., 2016. Evaluation of the quality of cold meats by computer-assisted image analysis. *LWT - Food Science and Technology* 67, 37–49. <https://doi.org/10.1016/j.lwt.2015.11.042>
- Zhang, Y., Ge, Y., Guo, Y., Miao, H., & Zhang, S., 2023. An approach for goose egg recognition for robot picking based on deep learning. *British Poultry Science* 64(3), 343–356. <https://doi.org/10.1080/00071668.2023.2171769>
- Zhao, S., Bai, Z., Huo, L., Han, G., Duan, E., Gu, Y., & Meng, L., 2025. Deep learning-based approach for behavior recognition and pose estimation in cage-reared ducks: Leveraging cage-net elimination to propel precision improvement. *Computers and Electronics in Agriculture* 236, 110469. <https://doi.org/10.1016/j.compag.2025.110469>
- Zheng, H., Fang, C., Zhang, T., Zhao, H., Yang, J., & Ma, C., 2022. Shank length and circumference measurement algorithm of breeder chickens based on extraction of regional key points. *Computers and Electronics in Agriculture* 197, 106989. <https://doi.org/10.1016/j.compag.2022.106989>
- Zhu, Z. H., Ye, Z. F., & Tang, Y., 2021. Nondestructive identification for gender of chicken eggs based on GA-BPNN with double hidden layers. *Journal of Applied Poultry Research* 30(4), 100203. <https://doi.org/10.1016/j.japr.2021.100203>
- Zhu, Z., Yang, K., & He, Y., 2022. Research on Detection of Fertility of Group Hatching Eggs Based on Adaptive Image Segmentation. *Applied Engineering in Agriculture* 38(2), 283–291. <https://doi.org/10.13031/aea.14849>
- A hybrid chromaticity-morphological machine learning model to overcome the limit of detecting newcastle disease in experimentally infected chicken within 36 h. Retrieved June 14, 2025, from <https://www.sciencedirect.com/science/article/pii/S0168169925003540>

Automated Chicken Counting in Surveillance Camera Environments Based on the Point Supervision Algorithm: LC-DenseFCN. Retrieved June 14, 2025, from <https://www.mdpi.com/2077-0472/11/6/493>

Deep learning implementation of image segmentation in agricultural applications: A comprehensive review. Retrieved June 14, 2025, from <https://link.springer.com/article/10.1007/s10462-024-10775-6>

CHAPTER II

ZERO-SHOT IMAGE SEGMENTATION FOR MONITORING THERMAL CONDITIONS OF INDIVIDUAL CAGE-FREE LAYING HENS

Body temperature is a critical indicator of the health and productivity of egg-laying chickens and other domesticated animals. Recent advancements in thermography allow for precise surface temperature measurement without physical contact with animals, reducing animal stress from human handling. Gold standard temperature analysis via thermography requires manual selection of limited points for an object of interest, which could be time-consuming and inadequate for representing the comprehensive thermal profile of a chicken's body. The objective of this study was to leverage and optimize a zero-shot artificial intelligence technology for the automatic segmentation of individual cage-free laying hens within thermal images, providing insights into their overall thermal conditions. A zero-shot image segmentation model (Segment Anything, "SAM") was modified by replacing manual selections of target points with automatic selection of the initial point using pre-processing techniques (e.g., thresholding) in each thermal image. The model was also incorporated with post-processing techniques integrated with a machine learning classifier to improve segmentation accuracy. Three versions of modified SAM models (i.e., SAM, FastSAM, and MobileSAM), two common instance segmentation algorithms (i. e., YOLOv8 and Mask R-CNN), and two foundation segmentation models (i.e., U2 -Net and ISNet) were comparatively evaluated to determine the optimal one for bird segmentation from thermal images. A total of 1,917 thermal images were collected from cage-free laying hens (Hy-Line W-36) at 77–

80 weeks of age. The image dataset exhibited considerable variations such as feathers, bird movement, body gestures, and the specific conditions of cage-free facilities. The experimental results demonstrate that the modified SAM did not only surpass the six zero-shot models—YOLOv8, Mask R-CNN, FastSAM, MobileSAM, U2 Net, and ISNet—but also outperformed other modified SAM-based models (Modified FastSAM and Modified MobileSAM) in terms of hen detection performance, achieving a success rate of 84.4 %, and in segmentation performance, with an intersection over union of 85.5 %, recall of 91.0 %, and an F1 score of 92.3 %. The optimal model, modified SAM, was pipelined to extract statistics including the averages ($^{\circ}\text{C}$) of mean (27.03, 27.04, 28.53, 26.68), median (26.27, 26.84, 28.28, 26.78), 25th percentile (25.33, 25.61, 27.26, 25.53), and 75th percentile (28.04, 27.95, 29.22, 27.55) of surface body temperature of individual laying hens in thermal images for each week. More statistics of hen body surface temperature can be extracted based on the segmentation results. The developed pipeline is a useful tool for automatically evaluating the thermal conditions of individual birds.

Introduction

The U.S. had 308 million commercial laying hens at the end of 2022, producing totally 92.6 billion table eggs (United Egg Producers, 2023). The egg industry is transitioning from cage systems to cage-free (CF) systems to improve hen welfare, and the CF egg production accounted for 34 % (106.2 million hens) of the current table egg layer flock (United Egg Producers, 2023). While providing nutritious, safe, and affordable proteins for humans, the intensive CF systems are facing challenges in managing hens effectively and appropriately such as air quality, floor eggs, distribution, and pecking (Chai et al., 2019). Thermal regulation is one of the critical areas to be optimized as it directly influences the health, productivity, and well-being of CF hens (Giloh et al., 2012; Tattersall, 2016) and accurate detection of hens' body temperature is the prerequisite of

precision thermal regulation. The gold standard method to obtain bird body temperature is to use thermometers to measure cloaca routinely (Candido [^] et al., 2020; Tattersall, 2016). The method can provide accurate measures of hen core body temperatures but could be time- and labor-intensive as human are required to catch and constrain birds for the measurement (Edgar et al., 2013), especially considering tens of thousands of CF hens moving freely inside a house. Thermometers should be inserted inside the bird's body through the cloaca and stabilized for a few minutes to acquire reliable readings, causing bird stress. The cloaca of birds could host rich bacterial communities, and thermometers could lead to cross-contamination if reused for another birds without disinfection. Thus, efficient, accurate, and biosafe alternatives are needed for hen body temperature measurement.

In recent years, thermography has become widely embraced as an alternative to traditional thermometers in industry and agriculture, primarily due to its ability to enable non-contact and non-invasive measurements of surface temperature (Baranowski et al., 2009; Biddle et al., 2018; Church et al., 2014; Cilulko et al., 2013; Sadeghi et al., 2023). The surface temperature measured by infrared thermal imaging has been strongly correlated to bird core body temperature (Giloh et al., 2012), ambient environmental conditions (Andrade et al., 2017), genetics (Loyau et al., 2016), and feather conditions (Cook et al., 2006), offering important insights into house management and bird improvement. Common temperature analysis via thermography requires the manual selection of limited points for an object of interest, which could be time-consuming and inadequate for representing the comprehensive thermal profile of a chicken's body. To gain the whole-body thermal profile and subsequent analysis, individual birds should be first segmented from the background in a thermal image via image segmentation models.

Image segmentation has been a fundamental challenge in computer vision since its creation (Rosenfeld, 1976). It involves partitioning images (or video frames) into multiple segments and objects, which are essential components for visual understanding systems (Szeliski, 2011). Instance segmentation (one type of image segmentation task) has evolved into a notably significant, intricate, and demanding domain within machine vision, especially during the rapid development of deep learning (Bolya et al., 2019; He et al., 2017; Li et al., 2017; Xie et al., 2022). Its objective is to predict both classes and pixel-specific masks for individual object instances, effectively identifying areas, shapes, and locations of individual objects within images (Bai and Urtasun, 2017; Hafiz and Bhat, 2020; He et al., 2017; Li et al., 2017). The instance segmentation methods, operated within a supervised learning framework, heavily depend on extensive datasets with annotations. Nevertheless, in numerous real-world applications such as segmenting CF hens, as mentioned above, the process of collecting and labeling data in pixel level is exceptionally time-consuming and requires professional annotators. Consequently, the instance segmentation models pretrained with large datasets (e.g., COCO and ImageNet) containing general objects and annotations might perform poorly when encountering unfamiliar classes (e.g., CF hens) with very few annotations in the datasets. In such scenarios, zero-shot learning methods prove to be highly valuable for solving the abovementioned issue (Zheng et al., 2021).

Zero-shot image segmentation involves a type of segmentation algorithms that can segment objects or regions of interest in images that have never been seen or been trained on (Bucher et al., 2019; Kato et al., 2019; Zheng et al., 2021). Some of the most popular and state-of-the-art zero-shot instance image segmentation models include Segment Anything Model (SAM) (Kirillov et al., 2023), Fast Segment Anything Model (FastSAM) (Zhao et al., 2023), Faster Segment Anything Model (MobileSAM or FasterSAM) (Zhang et al., 2023). These SAM-based models have gained

widespread attention since they appeared in recent months (Ma et al., 2024; Mazurowski et al., 2023; Osco et al., 2023; Shi et al., 2023). They were trained on billions of image masks and millions of images and could be generalizable to unseen or untrained objects. They can segment objects of interest through manually selecting a point inside the objects or drawing a bounding box. These models have been applied for medical image analysis (Mazurowski et al., 2023; Shi et al., 2023), agricultural image segmentation (Yaqin Li et al., 2023), remote sensing applications (Osco et al., 2023), video segmentation (Cheng et al., 2023), and demonstrated great potential for zero-shot image segmentation. Per preliminary testing, these SAM-based models may suffer from several challenges for poultry-related images including manually selecting initial points for segmentation and determining the optimal segmentation mask. Specifically, the manual selection of tens of thousands of points for segmenting CF hens in commercial houses is laborious.

The objective of this research was to leverage and optimize zero-shot artificial intelligence technology for the automatic segmentation of individual CF laying hens within thermal images. The strengths of this article includes: 1) a series of up-to-date zero-shot instance image segmentation algorithms were compared, and the optimal one was further optimized; 2) the foundation segmentation algorithm was integrated with image processing (pre-processing) for point prompting and machine learning classification for generated optimal mask (post-processing) for boosting segmentation performance of individual hens; and 3) thermal characteristics of laying hens (where birds appear higher temperature than the surrounding, resulting in brighter regions in thermal images) were fully utilized with the foundation segmentation model to improve bird segmentation performance. The contributions of this research involves: 1) critical development and optimization procedures of the combination of thermography and foundation image segmentation model were provided for fully automatic image segmentation in agriculture domain;

2) zero-shot image segmentation models were modified with image processing and machine learning classifiers to innovate model architecture tailored for poultry segmentation; and 3) a zero-shot instance image segmentation pipeline was developed to extract the statistics of surface body temperature of individual hens, opening numerous opportunities for mobile poultry health assessment. The proposed framework segments the most complete bird in a frame, with which the most comprehensive information of bird thermal conditions can be analyzed. Information of multiple birds can be collected by taking multiple photo shoots of thermal images.

Materials and Methods

Overall workflow

The workflow of this paper comprises nine major components, as illustrated in Figure 2.1. The first step involves data collection in CF hen environments and thermal camera calibration. Subsequently, the thermal and RGB images undergo spatial alignment with simple image processing algorithms to optimize its suitability for subsequent analyses. Following this, a comparative analysis is conducted, benchmarking the SAM against various state-of-the-art instance segmentation models. After that, the evaluation metrics calculation is performed to select the optimal model. The fifth phase focuses on modifying the SAM, incorporating both pre- and post-processing techniques to automate the SAM and enhance its performance for classification and segmentation. The next step encompasses calculating evaluation metrics to quantify the models' performance between modified SAM-based models, providing a rigorous assessment of their efficacy and accuracy. With the optimal model after the evaluation, the workflow includes sequential steps of automatic digit extraction in thermal images and finding the relationship between temperature and pixel intensity, providing a quantitative analysis of the thermal characteristics. The paper concludes with a ninth and final step, wherein statistics of surface

temperature from body profile are conducted to offer insights into the thermal condition of individual CF laying hens throughout four weeks. The sole programming language utilized was Python. Key Python libraries included cv2 and PIL for image manipulation, sklearn for constructing machine learning models, along with pandas and numpy for handling data. Matplotlib was used for graphically representing results. Additionally, supplementary libraries employed were pickle, os, skimage, csv, and sys. Computational operations were executed on Google Colab, which provided 12.7 GB of RAM and 16 GB of T4 GPU memory, supported by a dual-core CPU running at 2.30 GHz.

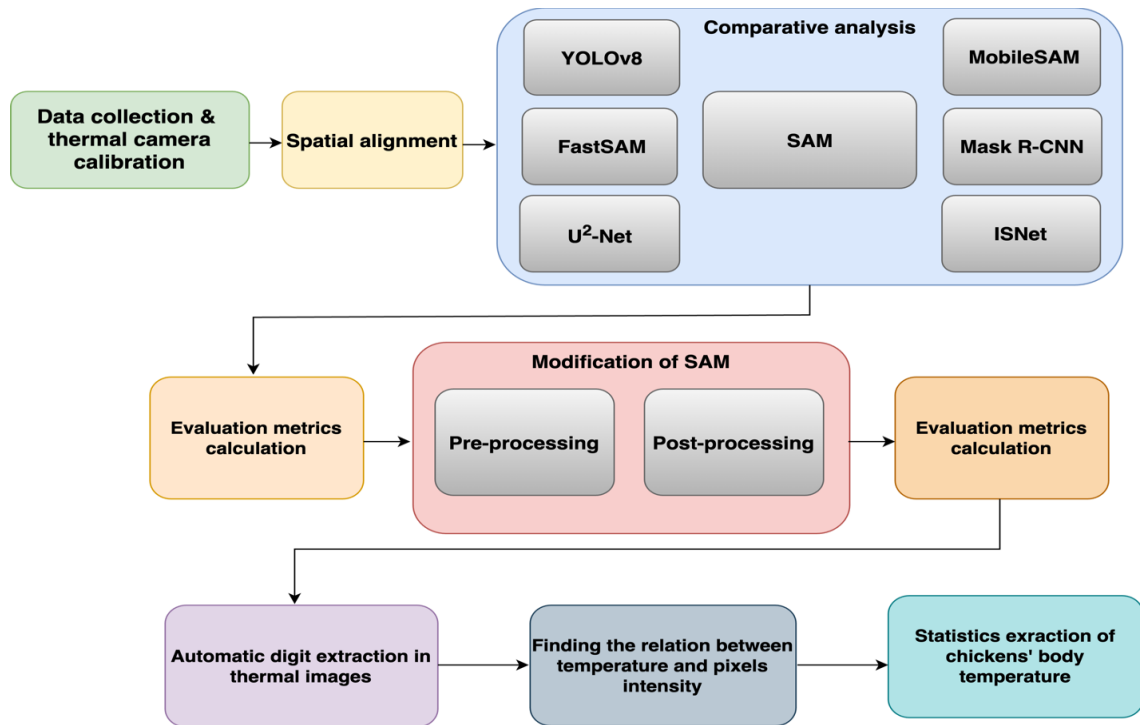


Fig. 2.1. Workflow diagram - This figure presents a schematic of the nine-step analytical process employed in the paper. SAM is Segment Anything, YOLO is You Only Look Once, and R-CNN is Region-based Convolutional Neural Network.

Data collection and thermal camera calibration

The study was conducted at the University of Georgia's Poultry Research Center. Four environmentally controlled rooms were used, each measuring 7.3 m long, 6.1 m wide, and 3.1 m high. Those dimensions were comparable to commercial house standards. Each room housed 200 laying hens (Hy-Line W-36) on a litter floor, which resulted in a stocking density of 0.22 m² per hen. This density exceeds the minimum requirements set by the United Egg Producers (United Egg Producers, 2023), which specify 0.093 m² per hen for multi-tiered aviaries and partially slatted systems, and 0.139 m² per hen for single-level all-litter floor systems. The larger space allows chickens to exhibit a broader range of natural behaviors, such as foraging, dust bathing, and perching. Each room contained 2.5-cm-deep pine wood shavings, an A-shaped perch with a total length of 36.6 m, and four nest boxes. The hens were fed with an antibiotic-free mash feed during the research. The diets were formulated in the feed mill located at the same research center with the following nutritional specifications: metabolizable energy: 1.26 MJ/ hen/day, crude protein: 16.70 g/day, calcium: 4 g/day, and digestible phosphorus: 0.40 g/day. Husbandry, management, and environmental conditions followed the Hy-Line W-36 commercial layers management guidelines (Management Guide, W-36 Commercial layers, 2024). All procedures were approved by the Institutional Animal Care and Use Committee (IACUC) prior to the start of the study (Protocol number: A2020 08–014-A1, approved on 5 October 2020). The diverse and dynamic environment enhances the validity and applicability of the collected data, capturing a representative sample of natural behaviors and interactions.

The thermal camera (FLIR C5, Teledyne FLIR, Wilsonville, Oregon, USA) was calibrated using a calibrator (FLUKE 9133, FLUKE, Everett, WA, USA) as shown in Figure 2.2, to ensure

accurate and precise temperature measurements. For the calibration, 15 points ranging from 24 to 38 °C were selected, as this range typically represents the variation in a chicken's body surface temperature. After setting the calibrator to a specific temperature, the temperature displayed by the thermal camera was recorded and plotted, as illustrated in Figure 2.10. Subsequently, a simple linear regression was utilized to determine the relationship between the temperatures indicated by the thermal camera and the calibrator.



Fig. 2.2. Precision calibrator for thermal camera calibration.

Thermal images were captured when birds were 77–80 weeks of age using the thermal imaging camera. The thermal camera outputs a pair of RGB and thermal images during each shot. The size of each image was 640×480 pixels. The total number of pairs of images was 1917. The images exhibited considerable diversity in terms of pixel intensity, varying backgrounds, presence of feathers on the ground, inclusion of nest boxes, and overlapping and occlusion among multiple chickens. Figure 2.3 shows four pairs of RGB and thermal sample images. Each pair of images contained at least one complete laying hen, and surface temperatures of all objects captured were

quantified with grayscale pixel intensities (0–255), with higher surface temperatures being brighter.

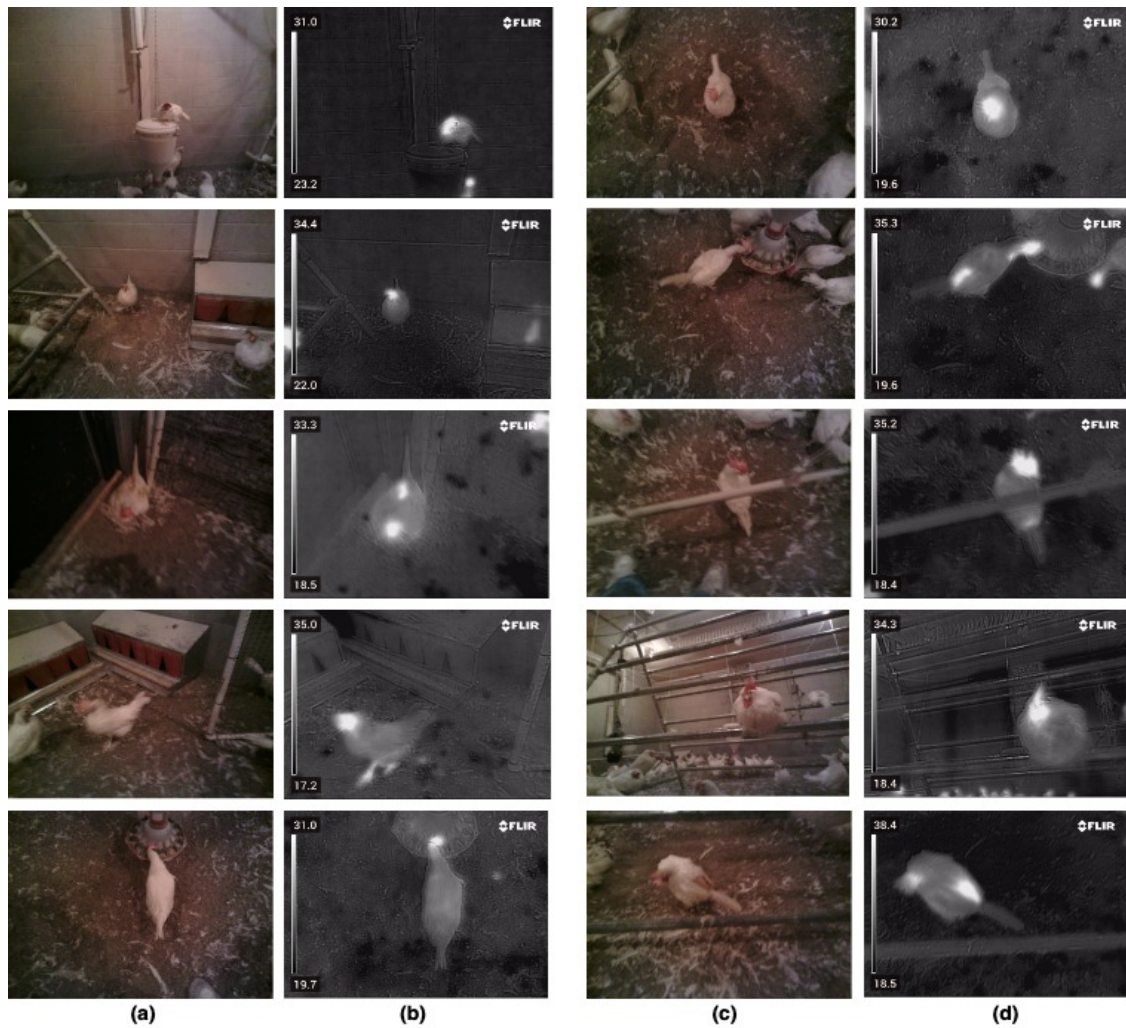


Fig. 2.3. Ten pairs of RGB and thermal images: **a, c)** RGB images; and **b, d)** corresponding thermal images. The bottom-left number under the bar in each thermal image represents the lowest temperature, and the top-left number above the bar indicates the highest temperature recorded in the scene.

Spatial alignment

Due to the low contrast and unclear boundaries of the chicken's body in the thermal images according to Figure 2.3, the corresponding RGB image was used for segmentation. The resulting mask was then multiplied by a corresponding thermal image to extract the chicken's body area. However, a challenge arose from different scales of thermal and corresponding RGB images. To address this, the RGB image underwent cropping and resizing to align with the thermal image scale, where the cropping coordinates are (105, 50, 440, 320). These cropping coordinates were determined through a trial-and-error process to find the most effective settings for matching the RGB images to the thermal images. These coordinates, which define the upper-left corner of the cropping area (105 for x-axis, 50 for y-axis) and specify the width (440) and height (320) of the area to be cropped, were consistently applied to all images. This consistency was crucial in ensuring that the resized RGB images matched perfectly with the corresponding thermal images. Then, the cropped RGB images were resized to a standard resolution of 640×480 pixels, matching the resolution of the thermal images. This standardization ensured that every detail captured in the RGB images was accurately mapped onto the corresponding areas in the thermal images. The consistent application of these cropping and resizing parameters across all images ensured that specific points on the thermal images were correctly aligned with those in the RGB images. This precise alignment is illustrated in Figure 2.4 and was critical for accurately overlaying the segmentation masks onto the thermal images. The same clipping and resizing paradigm worked perfectly fine for all of the images due to the fact that the camera intrinsic are fixed and the same camera was used to take both RGB and thermal images at the same time.

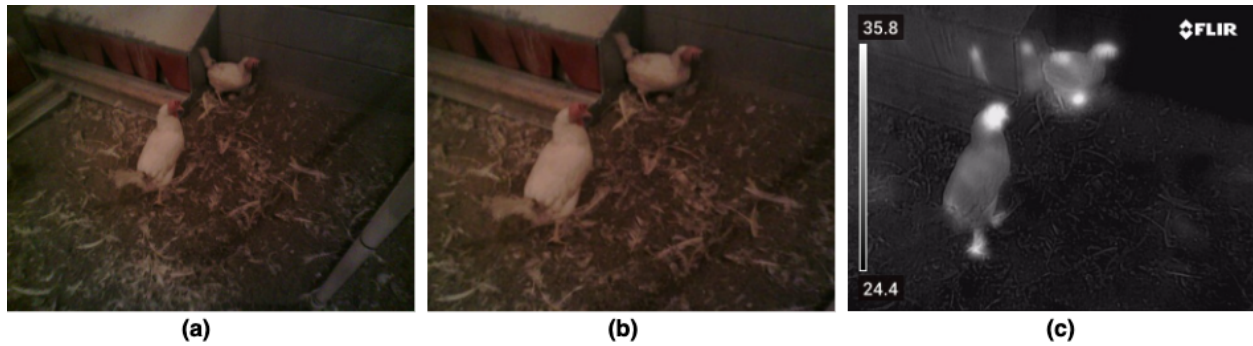


Fig. 2.4. Spatial alignment for RGB images to match corresponding thermal images: **a)** Original RGB image; **b)** RGB image after cropping and resizing; and **c)** corresponding thermal image.

Comparison of different zero-shot deep learning models for hen segmentation

SAM consists of three components: an image encoder, a flexible prompt encoder, and a fast mask decoder. SAM aims to transform the landscape of image analysis by offering a versatile and flexible foundational model for segmenting objects and regions within images. In contrast to conventional image segmentation models that demand extensive expertise in task-specific modeling, SAM eliminates the necessity for such specialized knowledge. It allows users to segment objects with just one or several interactive mouse clicks to include and exclude from the object. The model also accommodates prompts through bounding boxes. In instances of segmentation ambiguity, SAM can generate multiple valid masks, a crucial capability for real-world segmentation challenges.

Additional six deep learning models were deployed to verify the SAM's performance regarding zero-shot CF hen segmentation. These models were previously trained on large datasets and used directly to segment CF hens in this study without extensive training on images.

Fast Segment Anything Model (FastSAM), a CNN-based model, stands out for its speed due to its training on just 2 % of the SA-1B dataset, which contains 1 billion masks for training

general-purpose object segmentation models like SAM. The Faster Segment Anything model (MobileSAM or FasterSAM) improved processing speed by substituting the original bulky ViT-H (632 millions of parameters) encoder of SAM with a more compact Tiny-ViT (5 millions of parameters). Generally, FastSAM and MobileSAM are the extensive versions of original SAM by compressing SAM parameters to improve processing speed.

Mask R-CNN is a deep learning instance image segmentation model that was previously trained on the COCO (Common Objects in Context) dataset and has been used in many researches (Anantharaman et al., 2018; Chiao et al., 2019; Li et al., 2020; Lin et al., 2020; Zimmermann and Siems, 2019). The pre-trained model could be used as a zero-shot image segmentation model, especially when the class of interest is the same as that in the COCO dataset (He et al., 2017). YOLOv8 is the state-of-the-art YOLO model that can be used for object detection, image classification, and instance segmentation tasks (Yiting Li et al., 2023; Talaat and ZainEldin, 2023; Xiao et al., 2023). YOLOv8n architecture was used to offer enhanced performance on edge devices, providing a balance between detection accuracy and computational resource requirements. Both models have potential to segment CF hens without extensive training in this study but need verification.

U2-Net (Qin et al., 2020) is a deep learning model optimized for salient object detection, widely adapted for precise segmentation tasks due to its unique U-squared architecture. It excels in areas requiring detailed boundary delineation, such as medical imaging and object segmentation in videos. This model is particularly useful for zero-shot segmentation tasks involving unique object classes. ISNet (Jin et al., 2021) combines deep feature pyramids and attention mechanisms to enhance instance segmentation accuracy, originally developed for high resolution imagery but now also applied in agriculture. Both U2-Net and ISNet show potential for segmenting specific

poultry breeds without extensive custom training, although their effectiveness in such applications would need further validation.

Modification of Segment Anything model

After model comparison, the SAM-based models were further modified to improve CF hen segmentation performance. Detailed model performance results can be found in the result section.

As mentioned earlier, original SAM-related models may encounter challenges for automatic selection of initial points and best mask for segmentation. Pre-processing steps were introduced to automatically select an initial point of a target bird from each thermal image, and postprocessing steps were deployed to automatically determine the best mask among generated ones. The flowchart of the proposed method (Modified SAM) is shown in Figure 2.5. The proposed pre- and postprocessing techniques were also applied to FastSAM and MobileSAM, which are indicated as Modified FastSAM and Modified MobileSAM hereafter.

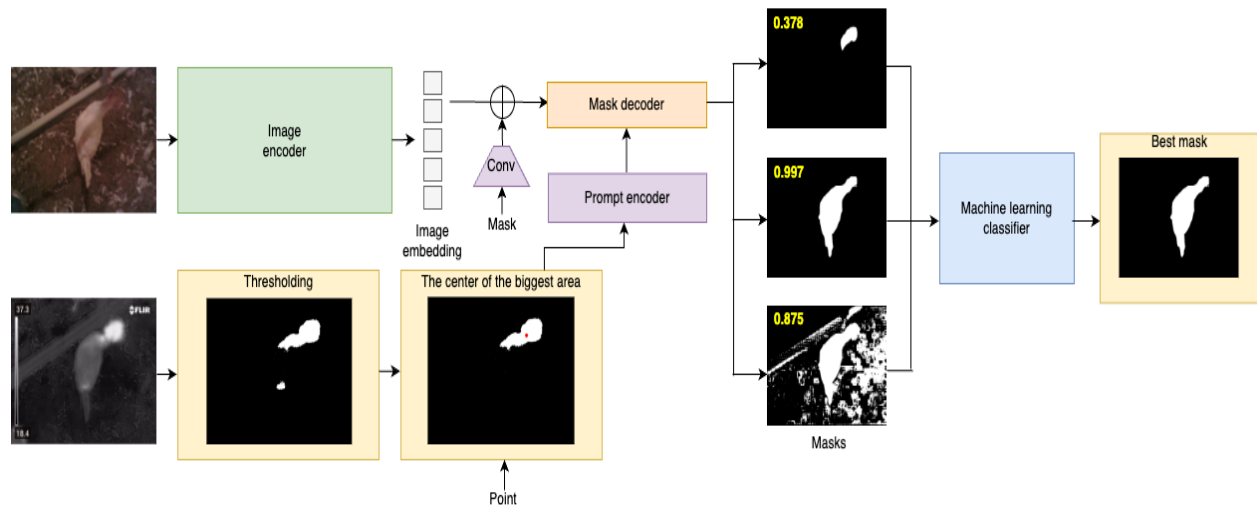


Fig. 2.5. Modified Segment Anything Model overview. A heavyweight image encoder outputs an image embedding that can then be efficiently queried by defining an automatic initial point to

produce object masks at amortized real-time speed. The yellow numbers in the masks indicate the confidence scores of the segmentation. To select the best masks out of the three masks generated, a machine learning classifier was used.

The pre-processing mainly involved automatic selection of the initial points of CF hens in thermal images. Initially, the highest intensity point (indicating the highest surface temperature) in the thermal image served as the initial point, as a target bird was assumed to have higher surface temperature than the background. However, in some cases, the highest intensity areas were concentrated at the edges of the chicken's body in most thermal images (Figure 2.6e), instead of the center of the chicken's body. Those edge areas could create ambiguity for the segmentation of the chicken and surroundings, further leading to inaccurate bird segmentation. To address this issue, the process was refined. Thermal images underwent segmentation using a threshold value of 100 to effectively isolate the warmer regions (hen bodies) from the cooler background (environment) resulting in one or multiple distinct regions or blobs representing potential subjects (hens). The threshold value of 100 for segmenting thermal images in this study was selected through an analysis aimed at optimizing the differentiation between the hens and the background. This process involved conducting an exploratory analysis where multiple threshold levels, ranging from 50 to 150 in increments of 10, were applied to a representative set of images. Each setting was evaluated based on the clarity and continuity of the hen shapes as well as the exclusion of background elements. The threshold of 100 worked experimentally better, consistently yielding the most accurate segmentation of the hens with minimal noise from the surroundings.

Among various segmented blobs, the largest blob, was selected based on the area it covers. The largest area refers to the largest contiguous region identified in the thermal image after

thresholding, which likely corresponds to the main body of a chicken due to the uniformity and intensity of the heat signatures characteristic of the hen's body. This step involves calculating the pixel count of each blob and identifying the one with the maximum count. The center of this largest blob was determined using geometric center calculations, commonly referred to as the centroid. The centroid of a shape in digital image processing is calculated as the average of all the x coordinates and the average of all the y coordinates of the pixels in a blob or region. This centroid acts as an initial point for further segmentation tasks, particularly for zero-shot segmentation techniques where selecting a meaningful starting point is crucial for model performance. By choosing the centroid of the largest heat-signature blob, the segmentation model is better oriented to focus on the hen's body rather than the surrounding cooler areas. This strategic choice enhances both the accuracy and efficiency of the subsequent segmentation steps, ensuring that the most significant thermal profile (the hen) is captured effectively in the analysis. Figure 2.6 visually illustrates the process of initial point selection. The selected initial point in a thermal image matched that in the corresponding RGB image for segmentation.

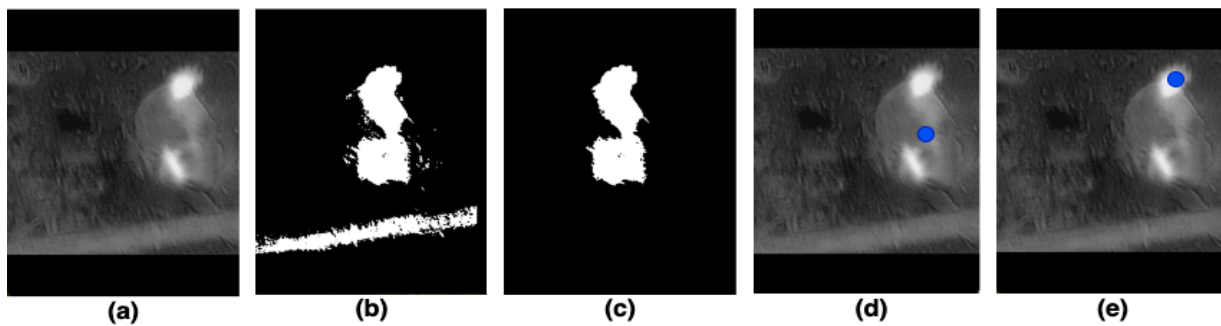


Fig. 2.6. SAM Pre-processing: **a)** Thermal image; **b)** segmented image after thresholding; **c)** extraction of the largest area; **d)** thermal image with the initial point determined from the center of the largest area; and **e)** thermal image with the initial point determined from the highest intensity (highest temperature).

In the original SAM-based models, a single prompt could yield multiple valid masks with different confidence scores. In most instances, the mask with the highest score aligned with the desired output (best mask). However, there were cases where the highest-scored mask may not be the optimal result among the generated masks, and two of these cases are indicated in Figure 2.7. According to Figure 2.7, the mask with the highest score in the first row is the optimal result. However, in the second row, the mask with the highest score does not represent the optimal result. To automatically select the best output mask from the three generated masks, a machine learning classification process was implemented.

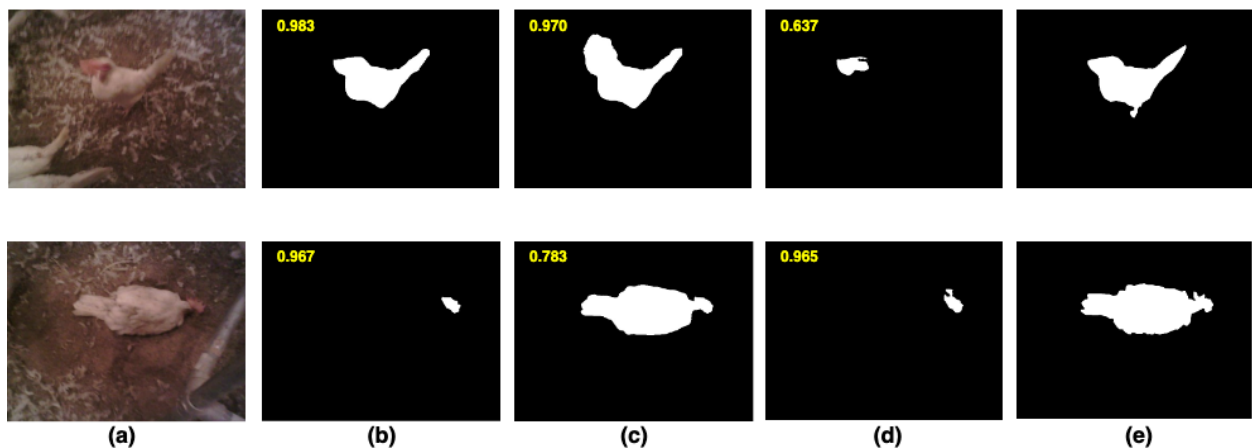


Fig. 2.7. SAM Post-processing: **a)** RGB image; **b, c, d)** three masks generated using SAM where **b** shows the masks with the highest confidence scores; and **e)** ground truth image.

The initial phase involved extracting valuable features from the generated mask images and saving the outcomes in a CSV file. This process yielded a data frame encapsulated within the CSV, comprising 444 instances (data points) that represent individual hens. These 444 instances or mask images were generated by SAM, and then technicians need to further verify whether the

generated masks were correct (1) or incorrect (0). Each instance was described by seven feature columns, including the number of white pixels, number of connected components, area, perimeter, eccentricity, equivalent diameter, and solidity, along with an additional column representing the target value.

The methodology for feature extraction was based on established principles in the field of image processing and analysis, rather than a single algorithmic approach (Saeidifar et al., 2021; Solis-Sanchez ' et al., 2011). The chosen features are well-recognized for their ability to capture critical information about binary images like masks and are extensively explained as follows:

- 1) **Number of white pixels.** This feature serves as a direct indicator of the mask's occupied area within the image, reflecting the presence and size of the mask. This simple, yet effective, feature quantifies the mask's extent.
- 2) **Number of connected components.** This feature indicates the number of isolated mask regions, providing insights into the mask's fragmentation or continuity, a key aspect in image analysis as supported by (Saeidifar et al., 2021).
- 3) **Area of a mask.** This feature is fundamental to understanding its spatial extent. This measure has been utilized to quantify object sizes in binary segmentation tasks, as explored by (Solis-Sanchez ' et al., 2011).
- 4) **Perimeter of a mask.** This feature offers a gauge for the complexity of the mask's boundary, influencing the shape and smoothness, as delineated in boundary analysis methods discussed by (Saeidifar et al., 2021; Solis-S' anchez et al., 2011).
- 5) **Eccentricity.** This feature measures the deviation of the mask's shape from a perfect circle, critical for distinguishing between various mask shapes, as applied in shape analysis in

(Solis-Sánchez et al., 2011). It has a clearly defined range of values. It is 0 for a perfectly round object and 1 for a line-shaped object:

$$\varepsilon = \frac{(\mu_{2,0} - \mu_{0,2})^2 - 4\mu_{0,2}^2}{(\mu_{2,0} + \mu_{0,2})^2} \quad (2.1)$$

where ε is eccentricity; $\mu_{2,0}$ and $\mu_{0,2}$ are the central moments of second order of any object inside an image.

- 6) **Equivalent diameter.** This feature converts the mask's area into the diameter of a circle with an equivalent area, providing a scaleinvariant size measure and facilitating comparison between masks of different sizes, as utilized in (Saeidifar et al., 2021).
- 7) **Solidity.** This feature reflects the ratio of the mask's area to its convex hull area, offering a metric for concavity, and has been employed to evaluate shape compactness in binary images, as noted in studies like (Solis-Sánchez ' et al., 2011).

These features were selected to ensure a comprehensive analysis of the masks. Collectively, the number of white pixels, area, and equivalent diameter provide related metrics that collectively depict the mask's scale and presence; the number of connected components, perimeter, and solidity contribute to an understanding of the mask's geometric properties and topology; and eccentricity offers a geometric analysis dimension, distinguishing elongated masks from more circular forms. This robust feature set was designed to provide a nuanced characterization across various mask morphologies and sizes.

To ensure the quality and consistency of the training and inference data, several preprocessing steps were performed on the extracted features for machine learning classification:

- 1) Normalization. All features were normalized into the same scale to reduce the potential biases and distortions caused by the different scales of these features.
- 2) Data Reshuffling. The dataset

was reshuffled to ensure that the data distribution was random and to reduce any potential bias during model training. 3) Data Splitting. The dataset was split into 5 folds for cross-validation.

Supervised machine learning classifiers were trained and tested on the mentioned features to evaluate their performance in predicting the accuracy of the masks. The classifiers used for mask selection were Decision Tree, Adaptive Boosting (Ada Boost), Support Vector Machine (SVM), Random Forest, and K-Nearest Neighbor (KNN), which are classical and popular supervised machine learning classifiers. A total of 444 mask images were used for the training. Each classifier was trained using a five-fold cross-validation approach to ensure robustness, given the relatively small size of the dataset.

Extensive hyperparameter tuning was conducted to optimize each classifier's performance:

- Decision Tree: Max depth, min samples split, min samples leaf.
- Ada Boost: Number of estimators, learning rate.
- SVM: Kernel type, regularization parameter (C), gamma.
- Random Forest: Number of estimators, max depth, min samples split and leaf.
- KNN: Number of neighbors (k), distance metric.

The optimal hyperparameter combinations were identified using grid search with cross-validation, which systematically evaluates a range of hyperparameter values and selects the combination that yields the best performance on the validation set.

Finally, the classifiers were evaluated and compared, and the optimal one was included in the SAM model. Once three masks were generated from SAM, they were classified with the optimal classifier to determine the most appropriate mask regardless of confidence scores. If multiple masks or none of the masks were classified as optimal, the mask with the highest score was retained.

Evaluation metrics calculation

This study employed a robust suite of evaluation metrics to independently gauge the performance of segmentation and classification, as well as the efficacy of machine learning classifiers in post-processing. The segmentation model evaluation leverages a dataset consisting of 1,917 RGB images of individual chickens. The annotation of these images was carried out by a well-trained technician using Roboflow, which ensured the provision of high-precision masks that delineated the most complete depiction of each chicken. Subsequently, the author conducted a double verification to guarantee the accuracy and quality of the labeling. This rigorous ground truth forms the benchmark for assessing the segmentation models' accuracy.

The trained models were evaluated with precision, recall, F1 score, and Intersection over Union (IoU) as described in Equations (2.2), (2.3), (2.4), and (2.5). The precision measures the accuracy of the segmentation model in identifying only relevant pixels as part of the segmentation. It is the ratio of correctly predicted positive observations to the total predicted positive observations. Recall, also known as sensitivity, measures the model's ability to correctly identify all relevant pixels. It is the ratio of correctly predicted positive observations to all observations that should have been labeled as positive. The F1 Score is the harmonic mean of Precision and Recall and is a measure of the model's accuracy. An F1 Score reaches its best value at 1 (perfect precision and recall) and worst at 0. IoU is a measure used to quantify the percent overlap between the target mask and the model's prediction output. It is calculated by dividing the area of overlap between the predicted segmentation and the ground truth by the area of union.

$$\text{Precision} = \frac{\text{True positive}}{\text{True positive} + \text{False positive}} \quad (2.2)$$

$$\text{Recall} = \frac{\text{True positive}}{\text{True positive} + \text{False negative}} \quad (2.3)$$

$$F1 \text{ score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (2.4)$$

$$IoU = \frac{\text{True positive}}{\text{True positive} + \text{False positive} + \text{False negative}} \quad (2.5)$$

where true positive refers to pixels that are correctly identified as part of the object of interest; false positive are the pixels that the segmentation model incorrectly identifies as part of the object, but they actually belong to the background or other objects; false negative is used for pixels that are part of the object in the ground truth but are missed by the segmentation model.

The detection metric employed is the success rate, which is based on the IoU value. A successful segmentation is one where the IoU is 50 % or greater, which aligns with standard thresholds used in prominent publications (Girshick, 2015; He et al., 2017; Redmon and Farhadi, 2018) as shown in Equation (2.6). The success rate thus reflects the percentage of images in which the models successfully segmented the chicken areas.

$$\text{Success rate} = \frac{\text{Number of successfully segmented images (IoU} > 0.5)}{\text{Total number of images}} \quad (2.6)$$

In this study, a carefully curated set of 444 mask images was labeled to determine the presence of chicken masks and used to evaluate the classification phase of the machine learning process. The accuracy of the classifiers, defined as the ratio of correctly identified masks (both chicken and non-chicken) to the total number of masks evaluated, serves as the fundamental metric for this assessment as shown in Equation (2.7).

$$\text{Accuracy} = \frac{\text{The number of correct predictions}}{\text{Total number of predictions}} \quad (2.7)$$

Automatic digit extraction in thermal images

To calculate the relationship between temperature readings and pixel intensity in thermal imagery, it is essential to first extract the two temperature values recorded by the thermal camera

within the image. This process involved identifying and cropping the regions that display the lowest and highest temperatures, which are in fixed positions in the thermal image. These cropped areas were then processed through an Optical Character Recognition (OCR) tool designed for Python, known as Python-tesseract (pytesseract). Python-tesseract is capable of recognizing and interpreting the text contained within images. For accurate OCR results, the background of the cropped images was modified to white with the temperature digits in black, which ensures optimal performance of the OCR tool. Figure 2.8 illustrates the mentioned process. Given the total count of 1917 thermal images and each image containing two temperature readings, a total of 3834 images containing the temperature digits were processed through the pytesseract classifier. The accuracy of the digit predictions made by pytesseract was found to be 100 %.

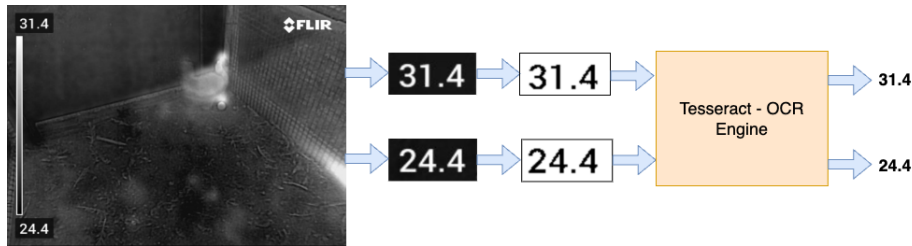


Fig. 2.8. Extracting digits from a thermal image using pytesseract. OCR is Optical Character Recognition.

Finding the relation between temperature and pixels intensity

After a chicken's body region was segmented from each RGB image, the area defined by the segmented mask was then multiplied to the corresponding thermal image to isolate the chicken's body temperature region. Figure 2.9. illustrates the result of this multiplication. To determine the temperature represented by each pixel in the chicken's body region in the thermal image, the minimum and maximum pixel intensities were first extracted. Then, utilizing the digits

obtained from the previous section, the corresponding surface temperatures of the pixels within the chicken's body area were calculated using the equation (2.8).

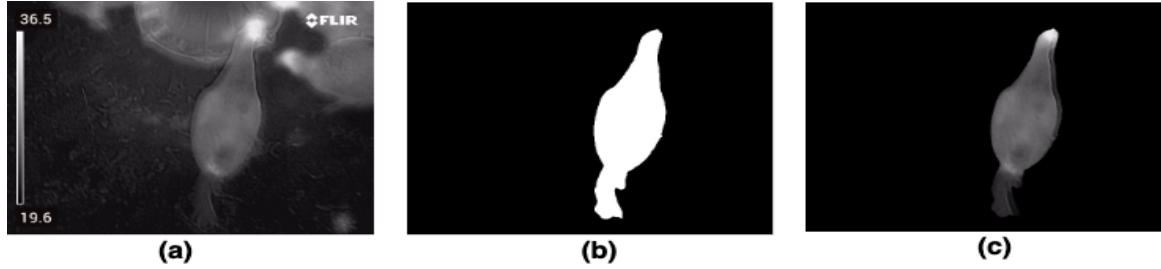


Fig. 2.9. Multiplication of the mask on the thermal image: **a)** Thermal image; **b)** segmented mask; and **c)** resulting image after multiplication.

$$Temperature = \frac{nonzero\ pixels - x1}{x2 - x1} \times (y2 - y1) + y1 \quad (2.8)$$

where $x1$ is the minimum pixel intensity in the thermal image, $x2$ is the maximum pixel intensity, $y1$ is the minimum temperature captured by the thermal camera, and $y2$ is the maximum temperature captured.

The equation derived from the calibration process involving the thermal camera and the calibrator, as depicted in Figure 2.10, should be used to calculate the actual surface temperature captured by the thermal camera. This ensures that the measurement reflects the true surface temperature.

Extracting statistics of chicken's body temperature

After calculating the actual surface temperatures of the chicken's body in the thermal images, various common statistical measures were computed to describe the data distribution across a four-week age span. The mean, or the average temperature value of the chicken's body

surface pixels, provides a central value of the temperature data. The median is the middle value that separates the higher half from the lower half of the temperature data, indicating a central trend without being affected by extreme values. The minimum and maximum values represent the lowest and highest temperatures observed, respectively, giving insights into the range of temperature variation. The 25th and 75th percentiles are values below which 25 % and 75 % of the temperature observations may be found, respectively, highlighting the spread and skewness of temperature values. Surface temperatures measured by infrared thermal imaging have shown a strong correlation with the core body temperature of birds, as indicated by (Giloh et al., 2012). The results and discussion section presents graphs plotting these various statistical measures.

Results and discussion

Thermal camera calibration

Figure 2.10 shows the linear regression between the thermal camera and calibrator temperature. According to Figure 2.10, the linear regression analysis between temperatures recorded by the thermal camera and those from the reference calibrator shows a high correlation, evidenced by an R-squared value of 0.99. This indicates the thermal camera's effectiveness in reflecting the calibrator's precision under controlled conditions. However, the deviation from a perfect R-squared value of 1.0 suggests the presence of factors that may introduce variability into the thermal camera's readings. The differences in subjects' postures and their distances from the camera could contribute to this variability. Changes in posture may alter the exposed surface area, and variations in distance could affect the thermal flux received by the camera, thereby influencing the temperature measurements (Intharachathorn et al., 2023, Kelly et al., 2019). These dynamics are particularly relevant when monitoring live subjects, such as CF laying hens in this case, where such variations are inevitable.

To enhance the consistency and accuracy of thermal readings, adopting a fixed distance between the camera and the subjects could be beneficial. This approach would likely mitigate the variability caused by distance-related changes in thermal camera, leading to more reliable temperature measurements. The broader application of this analysis highlights the importance of operational considerations, such as subject distance and posture, in the effective deployment of thermal imaging for animal monitoring. However, appropriately segmenting and extracting individual CF laying hens out of multiple birds should be explored with the fixed installation of a thermal camera.

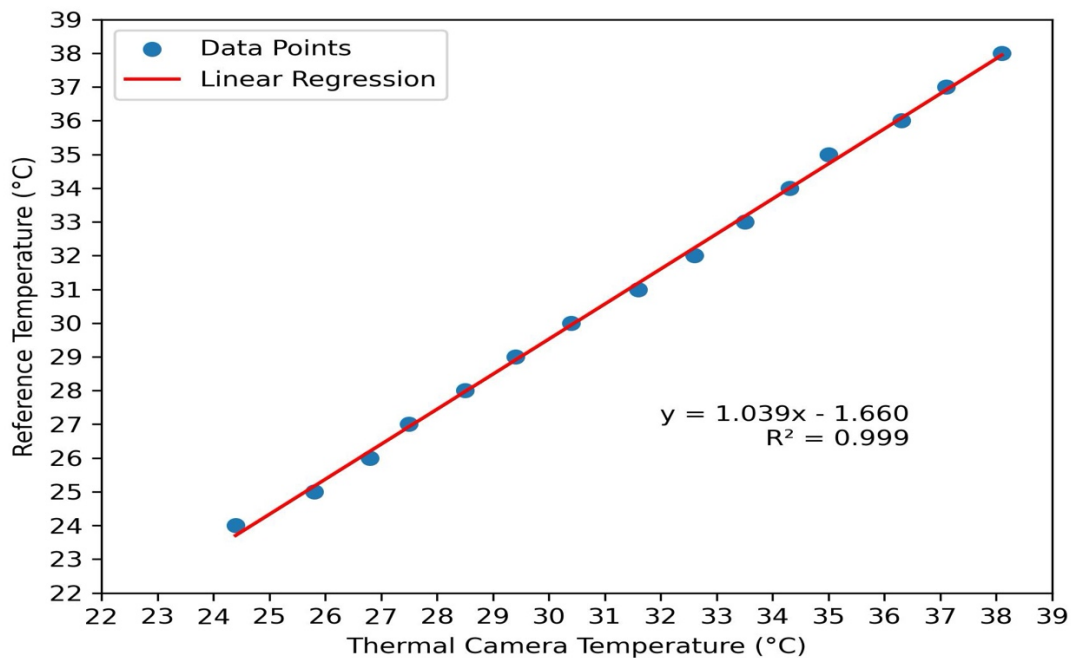


Fig. 2.10. The temperature calibration curve and the fitting formula.

Selecting the optimal model for zero-shot cage-free hen detection and segmentation

Table 2.1 provides a comparative overview of hen detection performance utilizing a range of zero-shot segmentation algorithms. The success rate, a key metric in this analysis, is determined

by the proportion of images where the IoU exceeds the 50% threshold. Images that met or surpassed this criterion were considered to be successfully detected and segmented, contributing positively to the overall success rate. The highest success rates were for U2-Net and SAM with 84.3% and 73.2%, respectively. Although U2-Net had the highest success rate among all models, it performed poorly in terms of other evaluation criteria as shown in Table 2.2.

Table 2.1. Performance comparison of hen detection before any modifications

Models	Success Rate (%)
YOLOV8n	50.0
MobileSAM	70.7
SAM	73.2
FastSAM	71.6
Mask R-CNN	64.2
U ² -Net	84.3
ISNet	64.4

Notes: SAM is Segment Anything, YOLO is You Only Look Once; and R-CNN is Region-based Convolutional neural network. U²-Net is U square net.

SAM's superiority lies in its ability to generalize effectively to new tasks and datasets without requiring task-specific training or fine-tuning. This capability is attributed to its training on the diverse SA-1B dataset and its design, which allows it to interpret and respond to a wide range of segmentation prompts, thereby enabling it to tackle a variety of segmentation challenges effectively in a zero-shot manner. Moreover, SAM generates three masks with varying confidence scores for each segmentation task, providing a nuanced approach to resolving the ambiguities inherent in segmentation tasks, particularly in zero-shot scenarios where the model is applied to completely unseen data (Kirillov et al., 2023). However, while SAM leads in zero-shot efficacy,

there is room for improvement to achieve an even higher success rate. This could potentially be addressed through modifications tailored to enhance its discrimination capabilities specifically for the task at hand.

On the other hand, the lower success rate (50.0%) of YOLOv8 underscores a limitation in its zero-shot detection capabilities, particularly in accurately identifying chickens. Although YOLOv8 and Mask R-CNN are powerful models, their performance is not optimized for this specific application without training on a dataset specific to chickens. Nonetheless, the overarching goal is to utilize a zero-shot instance segmentation method that operates without requiring any image training.

Table 2.2 presents a comparison of different segmentation metrics across the models to assess the segmentation capabilities of various models.

Table 2.2. Performance comparison of hen segmentation before any modifications

Models	Evaluation criteria (%)			
	Precision	Recall	F1 score	IoU
YOLOV8n	97.4	81.4	88.4	79.5
MobileSAM	92.9	90.2	91.1	84.0
SAM	93.8	90.6	92.2	85.4
FastSAM	92.6	90.4	91.5	84.1
MaskRCNN	87.5	90.2	88.8	79.9
U ² -Net	98.8	77.4	86.6	76.7
ISNet	99.6	71.7	83.1	71.5

Notes: SAM is Segment Anything, YOLO is You Only Look Once; and R-CNN is Region-based Convolutional neural network. U²-Net is U square net.

According to the data in Table 2.2, it is observed that the segmentation metrics across various SAM-based models before modifications are relatively consistent. This consistency is primarily because the metrics were calculated for images that the models have successfully segmented and detected in prior assessments. Therefore, these figures represent the performance on a refined subset of images — those that met the success criteria in the earlier detection phase and not the entire dataset.

While the segmentation results across models were closely matched, SAM still outperformed the others in most aspects, except for precision. The close results indicate that all models were reasonably effective in distinguishing the segmented chickens once they passed the initial detection threshold. However, SAM's slight edge in these unmodified conditions suggests that its core architecture is inherently more aligned with the nuances of hen segmentation tasks, even before any tailored enhancements are applied. Figure 2.11 illustrates the segmentation results achieved by different models.

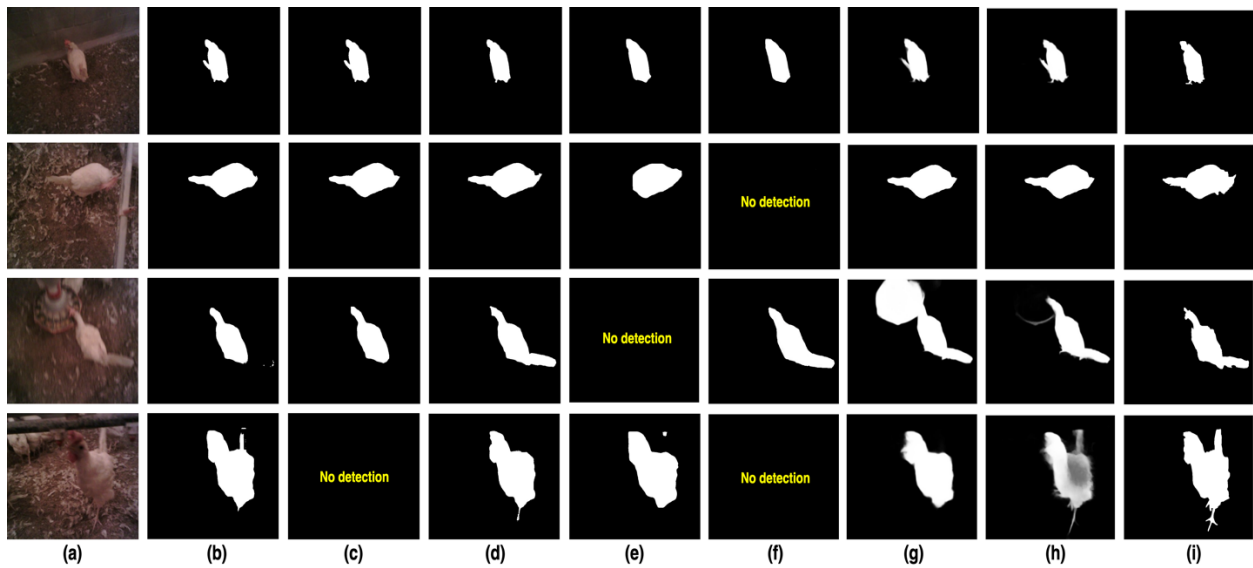


Fig. 2.11. Segmentation results of different models: **a)** RGB image; **b)** MobileSAM; **c)** FastSAM; **d)** SAM; **e)** YOLOv8, **f)** Mask R-CNN, **g)** U²-Net, **h)** ISNet; and **i)** ground truth.

In summary, SAM-based models performed better than the other two zero-shot instance segmentation, thus, they were selected for further optimization for segmenting individual laying hens from thermal images.

Comparison of five different machine learning classifiers for the post-processing

As mentioned earlier., an extensive hyperparameter tuning process was conducted to find the best hyperparameter values for each of the classifiers used for the optimal mask selection. Table 2.3 lists all the hyperparameters, the range of tested values, and the best value among them used for each of the five classifier models.

Table 2.3. Hyperparameter tuning results for machine learning classifiers

Model	Hyperparameter	Values	Best value
Decision Tree	max_depth	[3, 5, 7, None]	5
	min_samples_split	[2, 5, 10]	2
	min_samples_leaf	[1, 2, 4]	1
AdaBoost	n_estimators	[50, 100, 200]	100
	learning_rate	[0.01, 0.1, 1]	0.1
SVM	C	[0.1, 1, 10]	1
	gamma	['scale', 'auto']	'scale'
	kernel	['linear', 'rbf', 'poly']	'rbf'
Random Forest	n_estimators	[50, 100, 200]	100
	max_depth	[None, 10, 20, 30]	None
	min_samples_split	[2, 5, 10]	2
	min_samples_leaf	[1, 2, 4]	1
KNN	n_neighbors	[3, 5, 7]	5
	metric	['euclidean', 'manhattan']	'euclidean'

Notes: AdaBoost is Adaptive Boosting, SVM is Support Vector Machine; and KNN is K-Nearest Neighbor.

Table 2.4 compares the classification accuracy for various machine learning models used in the post-processing stage.

Table 2.4. Comparison of different machine learning classifiers for mask selection

Models	Accuracy (%)
Decision tree	91.7
Ada Boost	90.4
SVM	89.4
Random Forest	90.5
KNN	87.0

Notes: Ada Boost is Adaptive Boosting; and SVM is Support Vector Machine.

In the evaluation of machine learning classifiers for post-processing, the decision tree outperformed, closely followed by the Random Forest and Ada Boost methods. The SVM classifier also fared well, indicating its strong capability for this task. KNN, while still performing respectably, offers a valuable benchmark for comparative analysis.

The nuanced performance of these classifiers suggests that the more complex ensemble methods, despite their computational intensity, do not significantly outperform the simpler decision tree model in this context. This could be attributed to the nature of the features extracted for post-processing, where decision trees might capture the necessary patterns effectively without the need for ensemble strategies (Banfield et al., 2007).

Upon the completion of training, the decision tree guided the selection process among the three masks generated for each image. The classifier’s judgment is paramount; if it identified a single mask as optimal, that mask was selected as the output, overriding the score-based selection. In scenarios where multiple masks were deemed optimal, or none meet the criteria, the mask with the highest score was then chosen. This hybrid approach, combining the classifier’s analytical strengths with score-based evaluation, was designed to optimize mask selection, ensuring that the final output is not only based on empirical feature assessment but also on quantifiable segmentation performance.

Comparison of modified SAM-based models and two generic models for cage-free hen detection and segmentation

According to the baseline performance outlined in Tables 2.1 and 2.2, the SAM-based models were further developed by incorporating pre- and post-processing techniques. Since SAM-based models are considered generic segmentation models, to make the comparison fair, they were compared with the two generic baselines, namely U²-Net and ISNet. Table 2.5 examines the detection results from the modified SAM-based models as well as U²-Net and ISNet. Modified SAM models’ performance were compared in different states: without any pre- or post-processing, with pre-processing only, and with post-processing only. However, we could not apply similar pre and post-processing steps on U²-Net and ISNet as they do not accept input prompts for segmentation. This comprehensive evaluation was essential to evaluate the impact of each processing stage on the overall efficacy of the hen segmentation task.

Table 2.5. Performance comparison of hen detection for modified SAM-Based models

Models	Condition	Success rate (%)
SAM	W/O pre & post-processing	73.2
	W/ pre-processing	75.3
	W/ post-processing	82.6
	W/ pre & post-processing	84.4
MobileSAM	W/O pre & post-processing	70.7
	W/ pre-processing	72.0
	W/ post-processing	78.3
	W/ pre & post-processing	82.0
FastSAM	W/O pre & post-processing	71.6
	W/ pre-processing	72.8
U ² -Net	W/O pre & post-processing	84.3
ISNet	W/O pre & post-processing	64.4

Notes: 'w/o' refers to 'without,' indicating that the model was tested without the application of the associated processing technique. Conversely, 'w/' denotes 'with,' showing that the model was tested with the implementation of the given pre-processing or post-processing technique.

In examining the data presented in Table 2.5, it becomes apparent that the Modified SAM, which incorporated both pre- and post-processing steps, had a significantly higher success rate in detecting and segmenting hens than its counterparts in the family of SAM models. U²-Net model had a quite similar performance as modified-SAM model in terms of success rate. Success rate metric is crucial, as it quantifies the percentage of instances where the algorithm correctly identifies and delineates the subjects of interest. The comparative analysis reveals that the enhanced success rate of the Modified SAM—when contrasted with the original SAM, which lacks additional processing, the SAM with pre-processing only, and the SAM with post-processing only, underscores the substantial impact of the modifications on the algorithm's efficiency. These

techniques likely enhance the model's ability to discriminate between the hens and their surroundings by optimizing the input data quality and refining the segmentation output. The reason FastSAM was not implemented with post-processing was due to its design, which generates a single mask. Consequently, it does not necessitate post-processing steps typically required for models that produce multiple masks and need to select the best one. Furthermore, the Modified SAM's success showcased the potential of integrating zero-shot learning principles with targeted algorithmic enhancements to achieve high performance. This balanced approach leverages the inherent strengths of zero-shot models, such as their flexibility and generalizability, while compensating for their weaknesses through strategic modifications that tune the model to the specific characteristics of the task. Table 2.6 provides an evaluation of the segmentation performance of different SAM-based models.

Table 2.6. Performance comparison of hen segmentation for modified SAM-Based models

Models	Condition	Evaluation criteria (%)			
		Precision	Recall	F1score	IoU
SAM	W/O pre & post-processing	93.8	90.6	92.2	85.4
	W/ pre-processing	93.6	90.7	92.1	84.3
	W/ post-processing	93.8	90.7	92.2	85.1
	W/ pre & post-processing	93.6	91.0	92.3	85.5
MobileSAM	W/O pre & post-processing	92.9	90.2	91.5	84.0
	W/ pre-processing	92.5	90.4	91.4	84.1
	W/ post-processing	92.5	90.4	91.4	84.1
	W/ pre & post-processing	91.8	90.7	91.2	83.6
FastSAM	W/O pre & post-processing	92.6	90.4	91.5	84.1
	W/ pre-processing	92.5	90.4	91.4	84.1
U ² -Net	W/O pre & post-processing	98.8	77.4	86.6	76.7
ISNet	W/O pre & post-processing	99.6	71.7	83.1	71.5

Notes: 'w/o' refers to 'without,' indicating that the model was tested without the application of the associated processing technique. Conversely, 'w/' denotes 'with,' showing that the model was tested with the implementation of the given pre-processing or post-processing technique.

As highlighted in Table 2.2 and previously discussed, the segmentation metrics among the SAM-based models were generally consistent, reflecting their efficacy in processing successfully detected images. Similarly, Table 2.6 also presents comparable results. These findings further reinforce the reliability of the SAM models in segmenting images that meet the established success criteria. The two generic models, U²-Net and ISNet were quite competitive in terms of precision, but demonstrated poor performance in terms of recall, F1-score, and IoU which makes them unreliable for our testbed.

In the context of Table 2.6, while the results show a uniform performance, the Modified SAM models have shown incremental improvements in segmentation. These enhancements, particularly evident in the metrics of recall, F1 score and IoU, suggest that the post-modification refinements in the SAM-based models have further fine-tuned their segmentation capabilities, especially for the challenging aspects of the hen segmentation task. This incremental advancement underscores the value of the modifications introduced to the SAM framework, confirming that even minor adjustments can yield measurable benefits in segmentation precision and reliability. Figure 2.12 displays the segmentation results produced by the modified SAM. According to the figure, the Modified SAM could segment a more complete body profile compared to the Original SAM.

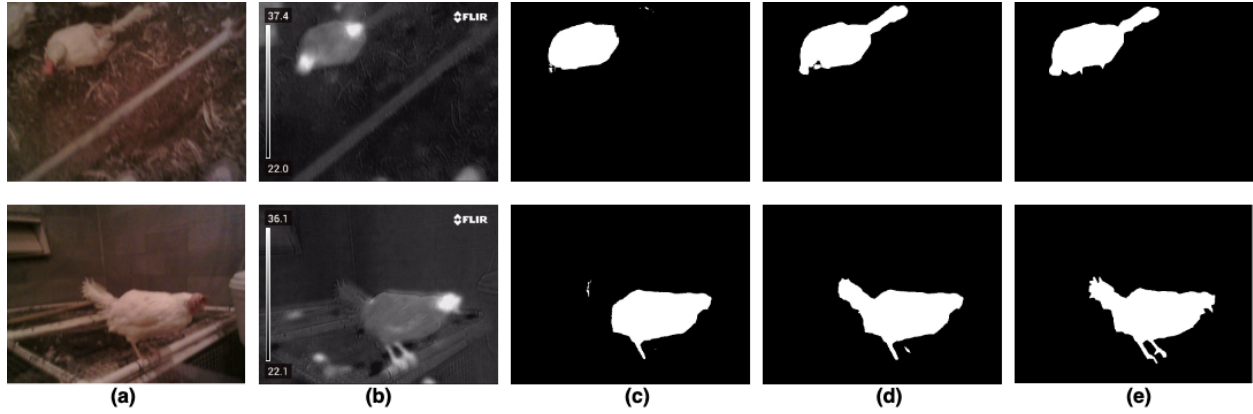


Fig. 2.12. Hen segmentation using Modified SAM: **a)** RGB Image; **b)** thermal image; **c)** segmentation result using SAM without modification; and **d)** segmentation result using Modified SAM; and **e)** ground truth image.

The primary goal of this study was to develop a zero-shot image segmentation algorithm aimed at minimizing the time-consuming task of manual segmentation annotation in the poultry sector. Traditional manual segmentation annotation of each bird, accounting for intricate details like head and leg contours, typically requires about one minute per image. Although semi-supervised tools like Roboflow reduce this time to approximately 20 seconds, they may still need human intervention for refined annotations. In our dataset of 1,917 images, manual segmentation annotation would take about 32 hours, and semi-automatic methods around 11 hours. By contrast, the proposed model processed each image in roughly 2 seconds, completing the task in about 1 hour for the entire dataset, thus saving significant labor costs. Considering the minimum hourly wage for labeling is \$13 per hour, our approach reduces segmentation annotation costs by \$400 and \$140 compared to manual and semi-automatic methods, respectively. Moreover, our model surpasses previous ones in all evaluation metrics. This efficiency makes zero-shot image

segmentation highly valuable for large-scale poultry studies, requiring extensive bird segmentation annotations.

Statistics of surface temperature in chickens over four weeks of age

After establishing the relationship between temperature and pixel intensity as discussed earlier, various statistics were extracted from the temperature of the segmented chicken pixels in the thermal images. The averages of the statistics have been plotted in Figure 2.13 across four weeks of bird age.

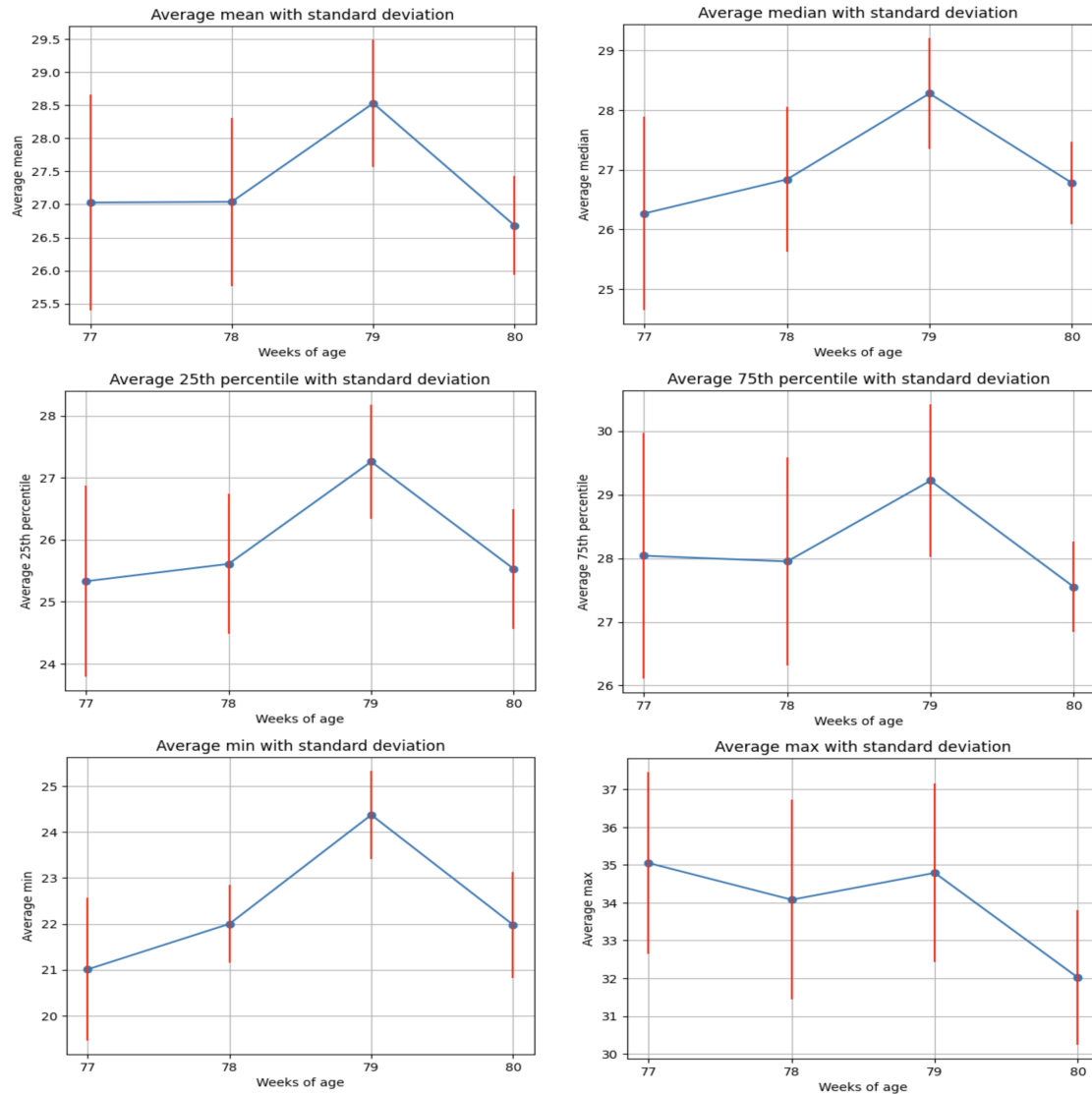


Fig. 2.13. Extraction of average statistics with standard deviation for surface body temperature of hens from weeks 77 to 80.

According to Figure 2.13, the six figures display various average temperature statistics along with their standard deviations, charting the trends in surface body temperatures of chickens from weeks 77 to 80. These trends were largely consistent across the figures, except for the average maximum temperatures. Since the average maximum temperature reflected only the single highest temperature point in a thermal image, it did not provide a comprehensive representation of the

overall surface body temperature. This discrepancy highlights the necessity of evaluating a range of statistics to gain a full understanding of surface body temperatures (Nascimento et al., 2014), a principle advocated for in this paper.

Bird surface body temperature fluctuated from weeks 77 to 80 with a peak at week 79. This demonstrates the necessity of continuous monitoring of bird thermal conditions to provide timely and precise thermal regulation for hens.

It is crucial to consider a comprehensive set of statistics when evaluating surface body temperature to obtain a representative thermal snapshot. This paper's detailed statistical approach, which includes analysis of mean, median, 25th percentile and 75th percentile, alongside the standard deviation, provides a comprehensive view of the thermal characteristics of the birds over the four-week period. Such an approach ensures that decisions or inferences drawn from the data are based on a complete and nuanced understanding of the thermal dynamics of hens.

Additional discussion and future work

Our study was conducted over a period of 4 weeks (28 days). To contextualize this duration, we compared it with the periods used in similar peer-reviewed articles. Table 2.7 shows examples of studies that utilized shorter periods. The longer data acquisition period in this study allows for capturing a wider range of temporal and spatial dynamics and behavior patterns, which is supportive for developing more robust and generalizable machine learning models.

Table 2.7. Comparison of study durations in various research studies

Reference	Study period
(Li et al., 2021)	3 days
(Bahuti et al., 2023)	21 days
(Lamping et al., 2022)	5 days

(Du et al., 2021)	7 days
This study	28 days

The structural similarity score was computed for each pair of images for demonstrating the diversity of the dataset, with 1 representing an identical/similar image and 0 representing a completely different image as shown in Figure 2.14. Histogram of the score for different pairs of images was computed to depict the distribution of image similarity scores.

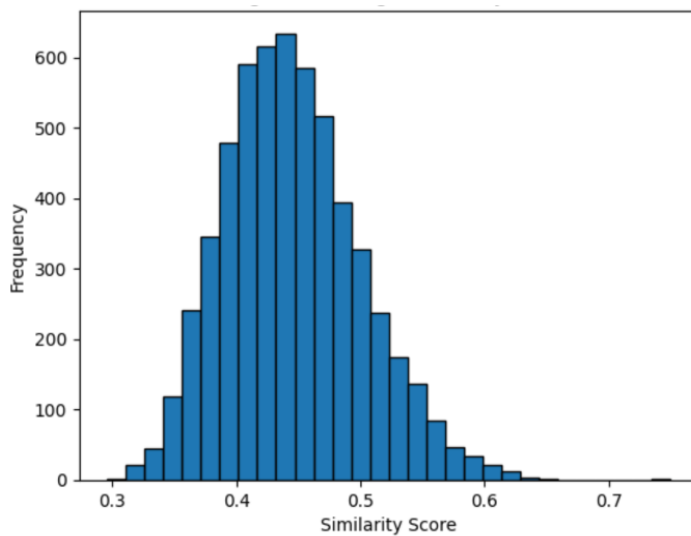


Fig. 2.14. Distribution of image similarity scores

According to the histogram graph, the following results were observed to support large diversity in the dataset.

1) **Wide range of similarity scores:** The similarity scores cover a wide spectrum from 0.3 to 0.7, reflecting the presence of both highly dissimilar and highly similar images in the dataset. This

range implies that the dataset contains a variety of images, rather than being skewed towards either end of the similarity spectrum.

2) **Normal distribution:** The similarity scores form a bell-shaped curve, similar to a normal distribution. This indicates a balanced dataset, where most image pairs exhibit moderate similarity, and fewer pairs are either highly similar or highly dissimilar. Such a balance points to diversity, suggesting that the dataset does not favor any particular type of image content.

3) **Larger proportion of dissimilar images:** Commonly, the similarity score of less than 0.5 indicates a different image, which took up over 75% in this case. Thus, the image inside the dataset was either moderately or highly dissimilar with each other.

Future work will focus on evaluating SAMAug, which uses augmented point prompts derived from initial SAM segmentation to improve the model's grasp of user intentions, thus boosting segmentation accuracy without additional inputs or model retraining (Dai et al., 2023). Additionally, for the post-processing step involving a machine learning classifier to select the optimal mask from three options, improvements could include expanding the training dataset for the classifier and augmenting the number of features provided to the classifier to increase its accuracy.

Conclusions

In this research, a zero-shot image segmentation technique was developed and optimized using the Segment Anything Model (SAM) for segmenting individual cage-free laying hens. The performance of the zero-shot model was significantly enhanced by integrating specific pre- and post-processing techniques, outperforming other zero-shot instance segmentation methods. The modified model has been streamlined into a pipeline to automatically extract comprehensive body temperature statistics, such as the mean, median, maximum, minimum, 25th, and 75th percentiles.

As a result, a valuable tool has been provided that supports precision poultry farming and aims to improve production efficiency. This advancement in the application of artificial intelligence in agriculture paves the way for more efficient health monitoring and management practices, potentially revolutionizing the poultry industry by enhancing both productivity and animal welfare.

Declaration of Competing Interest

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.

Acknowledgement

This project was jointly funded by US Poultry & Egg Association (project #: F-113), Institute for Integrative Precision Agriculture (IIPA) Seed Grant Program (project #: 2528), USDA-NIFA AFRI (2023-68008-39853), Egg Industry Center, and IIPA Pre-seed Grant Program (project #: 1941).

Author contribution

Conceptualization: G.L.; Data curation: R.B. and G.L.; Formal analysis: M.S. and G.L.; Funding acquisition: G.L. and L.C; Investigation: M.S. and G.L.; Methodology: M.S. and G.L.; Project administration: G.L. and L.C; Resources: G.L. and L.C.; Software: M.S.; Supervision: G.L.; Validation: M.S.; Visualization: M.S.; Roles/Writing - original draft: M.S. and G.L.; Writing - review & editing: L.C., K.M.R., J.L., A.B., T.L., X. Y.

References

Anantharaman, R., Velazquez, M., Lee, Y., 2018. Utilizing Mask R-CNN for Detection and Segmentation of Oral Diseases, in: 2018 IEEE International Conference on Bioinformatics and

Biomedicine (BIBM). Presented at the 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 2197–2204. <https://doi.org/10.1109/BIBM.2018.8621112>

Andrade, R.R., Tinôco, I.F.F., Baêta, F.C., Barbari, M., Conti, L., Cecon, P.R., Cândido, M.G.L., Martins, I.T.A., Junior, C.G.S., 2017. Evaluation of the surface temperature of laying hens in different thermal environments during the initial stage of age based on thermographic images. *Agronomy Research* 15.

Bahuti, M., Yanagi Junior, T., Lima, R.R. de, Fassani, É.J., Ribeiro, B.P.V.B., Campos, A.T., Abreu, L.H.P., 2023. Statistical and fuzzy modeling for accurate prediction of feed intake and surface temperature of laying hens subjected to light challenges. *Computers and Electronics in Agriculture* 211, 108050. <https://doi.org/10.1016/j.compag.2023.108050>

Bai, M., Urtasun, R., 2017. Deep Watershed Transform for Instance Segmentation, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2858–2866. <https://doi.org/10.1109/CVPR.2017.305>

Banfield, R.E., Hall, L.O., Bowyer, K.W., Kegelmeyer, W.P., 2007. A Comparison of Decision Tree Ensemble Creation Techniques. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29, 173–180. <https://doi.org/10.1109/TPAMI.2007.250609>

Baranowski, P., Mazurek, W., Witkowska-Walczak, B., Sławiński, C., 2009. Detection of early apple bruises using pulsed-phase thermography. *Postharvest Biology and Technology* 53, 91–100. <https://doi.org/10.1016/j.postharvbio.2009.04.006>

Biddle, L.E., Goodman, A.M., Deeming, D.C., 2018. Infrared thermography provides insight into the thermal properties of bird nests. *Journal of Thermal Biology* 76, 95–100. <https://doi.org/10.1016/j.jtherbio.2018.07.008>

Bolya, D., Zhou, C., Xiao, F., Lee, Y.J., 2019. YOLACT: Real-Time Instance Segmentation, in: 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Presented at the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 9156–9165. <https://doi.org/10.1109/ICCV.2019.00925>

Bucher, M., VU, T.-H., Cord, M., Pérez, P., 2019. Zero-Shot Semantic Segmentation, in: *Advances in Neural Information Processing Systems*. Curran Associates, Inc.

Cândido, M.G.L., Tinôco, I.F.F., Albino, L.F.T., Freitas, L.C.S.R., Santos, T.C., Cecon, P.R., Gates, R.S., 2020. Effects of heat stress on pullet cloacal and body temperature. *Poultry Science* 99, 2469–2477. <https://doi.org/10.1016/j.psj.2019.11.062>

Cheng, H.K., Wug Oh, S., Price, B., Schwing, A., Lee, J.-Y., 2023. Tracking Anything with Decoupled Video Segmentation, in: 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Presented at the 2023 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 1316–1326. <https://doi.org/10.1109/ICCV51070.2023.00127>

Chiao, J.-Y., Chen, K.-Y., Liao, K.Y.-K., Hsieh, P.-H., Zhang, G., Huang, T.-C., 2019. Detection and classification the breast tumors using mask R-CNN on sonograms. *Medicine (Baltimore)* 98, e15200. <https://doi.org/10.1097/MD.00000000000015200>

Church, J.S., Hegadoren, P.R., Paetkau, M.J., Miller, C.C., Regev-Shoshani, G., Schaefer, A.L., Schwartzkopf-Genswein, K.S., 2014. Influence of environmental factors on infrared eye temperature measurements in cattle. *Research in Veterinary Science* 96, 220–226. <https://doi.org/10.1016/j.rvsc.2013.11.006>

Cilulko, J., Janiszewski, P., Bogdaszewski, M., Szczygalska, E., 2013. Infrared thermal imaging in studies of wild animals. *Eur J Wildl Res* 59, 17–23. <https://doi.org/10.1007/s10344-012-0688-1>

Cook, N.J., Smykot, A.B., Holm, D.E., Fassenko, G., Church, J.S., 2006. Assessing Feather Cover of Laying Hens by Infrared Thermography. *Journal of Applied Poultry Research* 15, 274–279. <https://doi.org/10.1093/japr/15.2.274>

Dai, H., Ma, C., Liu, Z., Li, Y., Shu, P., Wei, X., Zhao, L., Wu, Z., Zeng, F., Zhu, D., Liu, W., Li, Q., Liu, T., Li, X., 2023. SAMAug: Point Prompt Augmentation for Segment Anything Model [WWW Document]. arXiv.org. URL <https://arxiv.org/abs/2307.01187v2> (accessed 3.3.24).

Du, X., Teng, G., Wang, C., Carpentier, L., Norton, T., 2021. A tristimulus-formant model for automatic recognition of call types of laying hens. *Computers and Electronics in Agriculture* 187, 106221. <https://doi.org/10.1016/j.compag.2021.106221>

Edgar, J.L., Nicol, C.J., Pugh, C.A., Paul, E.S., 2013. Surface temperature changes in response to handling in domestic chickens. *Physiology & Behavior* 119, 195–200. <https://doi.org/10.1016/j.physbeh.2013.06.020>

- Giloh, M., Shinder, D., Yahav, S., 2012. Skin surface temperature of broiler chickens is correlated to body core temperature and is indicative of their thermoregulatory status¹. *Poultry Science* 91, 175–188. <https://doi.org/10.3382/ps.2011-01497>
- Girshick, R., 2015. Fast R-CNN. <https://doi.org/10.48550/arXiv.1504.08083>
- Hafiz, A.M., Bhat, G.M., 2020. A survey on instance segmentation: state of the art. *Int J Multimed Info Retr* 9, 171–189. <https://doi.org/10.1007/s13735-020-00195-x>
- He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask R-CNN, in: 2017 IEEE International Conference on Computer Vision (ICCV). Presented at the 2017 IEEE International Conference on Computer Vision (ICCV), pp. 2980–2988. <https://doi.org/10.1109/ICCV.2017.322>
- Intharachathorn, K., Jareemit, D., Watcharapinchai, S., 2023. Potential use of an extended-distance thermal imaging camera for the assessment of thermal comfort in multi-occupant spaces. *Building and Environment* 246, 110949. <https://doi.org/10.1016/j.buildenv.2023.110949>
- Jin, Z., Liu, B., Chu, Q., Yu, N., 2021. ISNet: Integrate Image-Level and Semantic-Level Context for Semantic Segmentation. Presented at the Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 7189–7198.
- Kato, N., Yamasaki, T., Aizawa, K., 2019. Zero-Shot Semantic Segmentation via Variational Mapping, in: 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Presented at the 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), pp. 1363–1370. <https://doi.org/10.1109/ICCVW.2019.00172>
- Kelly, J., Kljun, N., Olsson, P.-O., Mihai, L., Liljeblad, B., Weslien, P., Klemedtsson, L., Eklundh, L., 2019. Challenges and Best Practices for Deriving Temperature Data from an Uncalibrated UAV Thermal Infrared Camera. *Remote Sensing* 11, 567. <https://doi.org/10.3390/rs11050567>
- Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.-Y., Dollár, P., Girshick, R., 2023. Segment Anything. <https://doi.org/10.48550/arXiv.2304.02643>
- Lamping, C., Derks, M., Groot Koerkamp, P., Kootstra, G., 2022. ChickenNet - an end-to-end approach for plumage condition assessment of laying hens in commercial farms using computer vision. *Computers and Electronics in Agriculture* 194, 106695. <https://doi.org/10.1016/j.compag.2022.106695>

Li, G., Hui, X., Chen, Z., Chesser, G.D., Zhao, Y., 2021. Development and evaluation of a method to detect broilers continuously walking around feeder as an indication of restricted feeding behaviors. *Computers and Electronics in Agriculture* 181, 105982. <https://doi.org/10.1016/j.compag.2020.105982>

Li, G., Hui, X., Lin, F., Zhao, Y., 2020. Developing and Evaluating Poultry Preening Behavior Detectors via Mask Region-Based Convolutional Neural Network. *Animals* 10, 1762. <https://doi.org/10.3390/ani10101762>

Li, Yiting, Fan, Q., Huang, H., Han, Z., Gu, Q., 2023. A Modified YOLOv8 Detection Network for UAV Aerial Image Recognition. *Drones* 7, 304. <https://doi.org/10.3390/drones7050304>

Li, Y., Qi, H., Dai, J., Ji, X., Wei, Y., 2017. Fully Convolutional Instance-Aware Semantic Segmentation, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4438–4446. <https://doi.org/10.1109/CVPR.2017.472>

Li, Yaqin, Wang, D., Yuan, C., Li, H., Hu, J., 2023. Enhancing Agricultural Image Segmentation with an Agricultural Segment Anything Model Adapter. *Sensors* 23, 7884. <https://doi.org/10.3390/s23187884>

Lin, K., Zhao, H., Lv, J., Li, C., Liu, X., Chen, R., Zhao, R., 2020. Face Detection and Segmentation Based on Improved Mask R-CNN. *Discrete Dynamics in Nature and Society* 2020, e9242917. <https://doi.org/10.1155/2020/9242917>

Loyau, T., Zerjal, T., Rodenburg, T.B., Fablet, J., Tixier-Boichard, M., Laan, M.H.P. der, Mignon-Grasteau, S., 2016. Heritability of body surface temperature in hens estimated by infrared thermography at normal or hot temperatures and genetic correlations with egg and feather quality. *animal* 10, 1594–1601. <https://doi.org/10.1017/S1751731116000616>

Ma, J., He, Y., Li, F., Han, L., You, C., Wang, B., 2024. Segment anything in medical images. *Nat Commun* 15, 654. <https://doi.org/10.1038/s41467-024-44824-z>

Management Guide, W-36 Commercial layers, 2024. Hy-Line International [WWW Document]. URL <https://www.hyline.com/> (accessed 2.12.24).

Mazurowski, M.A., Dong, H., Gu, H., Yang, J., Konz, N., Zhang, Y., 2023. Segment anything model for medical image analysis: An experimental study. *Medical Image Analysis* 89, 102918. <https://doi.org/10.1016/j.media.2023.102918>

- Nascimento, S.T., da Silva, I.J.O., Maia, A.S.C., de Castro, A.C., Vieira, F.M.C., 2014. Mean surface temperature prediction models for broiler chickens—a study of sensible heat flow. *Int J Biometeorol* 58, 195–201. <https://doi.org/10.1007/s00484-013-0702-7>
- Osco, L.P., Wu, Q., de Lemos, E.L., Gonçalves, W.N., Ramos, A.P.M., Li, J., Marcato, J., 2023. The Segment Anything Model (SAM) for remote sensing applications: From zero to one shot. *International Journal of Applied Earth Observation and Geoinformation* 124, 103540. <https://doi.org/10.1016/j.jag.2023.103540>
- Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O.R., Jagersand, M., 2020. U2-Net: Going deeper with nested U-structure for salient object detection. *Pattern Recognition* 106, 107404. <https://doi.org/10.1016/j.patcog.2020.107404>
- Redmon, J., Farhadi, A., 2018. YOLOv3: An Incremental Improvement. <https://doi.org/10.48550/arXiv.1804.02767>
- Rosenfeld, A., 1976. *Digital Picture Processing*. Academic Press.
- Sadeghi, M., Banakar, A., Minaei, S., Orooji, M., Shoushtari, A., Li, G., 2023. Early Detection of Avian Diseases Based on Thermography and Artificial Intelligence. *Animals* 13, 2348. <https://doi.org/10.3390/ani13142348>
- Saeidifar, M., Yazdi, M., Zolghadrasli, A., 2021. Performance Improvement in Brain Tumor Detection in MRI Images Using a Combination of Evolutionary Algorithms and Active Contour Method. *J Digit Imaging* 34, 1209–1224. <https://doi.org/10.1007/s10278-021-00514-6>
- Shi, P., Qiu, J., Abaxi, S.M.D., Wei, H., Lo, F.P.-W., Yuan, W., 2023. Generalist Vision Foundation Models for Medical Imaging: A Case Study of Segment Anything Model on Zero-Shot Medical Segmentation. *Diagnostics* 13, 1947. <https://doi.org/10.3390/diagnostics13111947>
- Solis-Sánchez, L.O., Castañeda-Miranda, R., García-Escalante, J.J., Torres-Pacheco, I., Guevara-González, R.G., Castañeda-Miranda, C.L., Alaniz-Lumbreras, P.D., 2011. Scale invariant feature approach for insect monitoring. *Computers and Electronics in Agriculture* 75, 92–99. <https://doi.org/10.1016/j.compag.2010.10.001>
- Szeliski, R., 2011. *Computer Vision: Algorithms and Applications*, Texts in Computer Science. Springer, London. <https://doi.org/10.1007/978-1-84882-935-0>
- Talaat, F.M., ZainEldin, H., 2023. An improved fire detection approach based on YOLO-v8 for smart cities. *Neural Comput & Applic* 35, 20939–20954. <https://doi.org/10.1007/s00521-023-08809-1>

Tattersall, G.J., 2016. Infrared thermography: A non-invasive window into thermal physiology. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology, Ecophysiology methods: refining the old, validating the new and developing for the future* 202, 78–98. <https://doi.org/10.1016/j.cbpa.2016.02.022>

United Egg Producers, 2023. Facts & Stats. United Egg Producers. URL <https://unitedegg.com/facts-stats/> (accessed 2.12.24).

Xiao, B., Nguyen, M., Yan, W.Q., 2023. Fruit ripeness identification using YOLOv8 model. *Multimed Tools Appl.* <https://doi.org/10.1007/s11042-023-16570-9>

Xie, E., Wang, W., Ding, M., Zhang, R., Luo, P., 2022. PolarMask++: Enhanced Polar Representation for Single-Shot Instance Segmentation and Beyond. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, 5385–5400. <https://doi.org/10.1109/TPAMI.2021.3080324>

Zhang, C., Han, D., Qiao, Y., Kim, J.U., Bae, S.-H., Lee, S., Hong, C.S., 2023. Faster Segment Anything: Towards Lightweight SAM for Mobile Applications. <https://doi.org/10.48550/arXiv.2306.14289>

Zhao, X., Ding, W., An, Y., Du, Y., Yu, T., Li, M., Tang, M., Wang, J., 2023. Fast Segment Anything. <https://doi.org/10.48550/arXiv.2306.12156>

Zheng, Y., Wu, J., Qin, Y., Zhang, F., Cui, L., 2021. Zero-Shot Instance Segmentation, in: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2593–2602. <https://doi.org/10.1109/CVPR46437.2021.00262>

Zimmermann, R.S., Siems, J.N., 2019. Faster training of Mask R-CNN by focusing on instance boundaries. *Computer Vision and Image Understanding* 188, 102795. <https://doi.org/10.1016/j.cviu.2019.102795>

CHAPTER III

AUTOMATIC SEGMENTATION OF BIRDS USING A COMBINATION OF OBJECT DETECTION AND FOUNDATION IMAGE SEGMENTATION MODELS

This study introduced an innovative method for automatic bird segmentation by combining an object detection model (i.e., YOLOv7) with a foundation image segmentation model (i.e., Segment Anything Model, SAM). YOLOv7 detected individual birds in images and calculated bounding box prompts of each detected bird for the SAM, enabling detailed and efficient segmentation without manual point inputs. The developed method was compared with various segmentation methods, including YOLOv8, Thermal image + MobileSAM, Thermal image + SAM, Thermal image + FastSAM, Mask R-CNN, and YOLOv7 (providing centroids of detected birds as point prompts) + SAM. The results showed that the proposed method outperformed all of the comparative segmentation methods, with the highest precision of 92.5%, recall of 98.2%, F1 score of 95.1%, IoU of 91.0%, and success rate of 98.0%. The study highlights a significant advancement in automatic image segmentation techniques with less intensive human annotation than standard deep learning-based image segmentation methods. The developed methods can be scaled up and transferred to various agricultural, environmental, medical, geographical, and urban planning applications.

Introduction

The rapidly expanding domain of image segmentation has witnessed a remarkable transformation with the introduction of a comprehensive foundation model, Segment Anything

Model (SAM) (Kirillov et al., 2023). A foundation image segmentation model should have good generalizability in segmenting objects of interest from various backgrounds and environments beyond training datasets. The most impressive part is that without task-specific training (zero-shot) or with minimal additional training via prompts (user inputs) (few-shot), the foundation models (Bommasani et al., 2021) can rival or even outperform traditionally trained models on certain tasks. Some of the most popular and state-of-the-art zero-shot instance image segmentation models other than SAM include Fast Segment Anything Model (FastSAM) (Zhao et al., 2023) and Faster Segment Anything Model (MobileSAM or FasterSAM) (Zhang et al., 2023). These SAM-based models have gained widespread attention since they appeared in recent studies (Ma et al., 2024; Mazurowski et al., 2023; Osco et al., 2023; Shi et al., 2023). Researchers are constantly pushing the boundaries by increasing model size, dataset comprehensiveness, and the computational power used for training (Brown et al., 2020; Hoffmann et al., 2022; Kaplan et al., 2020; Chowdhery et al., 2023). Owing to large-scale training in the substantial dataset, SA-1B, which comprises over 1 billion masks and 11 million images, SAM has demonstrated robust zero-shot or few-shot image segmentation performance via various input prompts such as point or bounding box prompts. A data engine was developed for the SAM dataset, involving three stages: model-assisted manual annotation, semi-automatic, and fully automatic. In the manual annotation stage, annotators labeled masks based on complete shapes rather than prompts, significantly improving annotation speed and quality, resulting in 4.3 million masks from 120k images. The semi-automatic stage aimed to label less prominent objects by training a bounding box detector, generating 5.9 million masks from 180k images. The fully automatic stage leveraged improvements from previous stages and ambiguity awareness to generate 1.1 billion masks from 11 million images without human intervention. Quality was maintained by comparing and refining automatically generated masks

using IoU metrics. SAM's architecture includes a Masked Auto Encoding (MAE) pre-trained Vision Transformer (ViT) encoder that produces image embeddings for prompt-based mask generation (Kirillov et al., 2023).

SAM requires manual input prompts, which poses limitations of SAM applications in the scenarios with dense distributions of targeted objects. For instance, a modern poultry house typically contains tens of thousands of birds, and the same amount of human input prompts are needed to segment individual birds out from images if SAM is deployed in the precision poultry farming domain. Thus, automating the prompting procedure for SAM is urgently needed to avoid laborious manual inputs for densely distributed objects. The agricultural industry, particularly poultry farming, stands to benefit significantly from advancements in automated image segmentation due to the sheer scale and complexity of monitoring animal welfare (Edgar et al., 2013). Enhanced segmentation capabilities can lead to better health monitoring, resource allocation, and overall management of poultry houses.

Previous studies investigated integrating image processing (pre-processing) for point prompting and machine learning for generated mask classification (post-processing) into SAM to segment individual laying hens from thermal images (Saeidifar et al., 2024). While achieving a success rate of 84.4%, IoU of 85.5%, recall of 91.0%, and F1 score of 92.3%, the proposed method inevitably had several drawbacks. Firstly, thermography, despite providing thermal characteristics of target objects for prompting, can be subject to ambient temperature and is not economically friendly for end users. Instead, RGB is the mainstream of deep learning model development due to its cost-effectiveness and easy access. Second, the proposed framework still requires training for supervised machine learning classifiers, which is not supportive to achieve zero-shot or few-shot image segmentation. Moreover, relying on thermal imaging limits the versatility of the

system, as thermography cannot be easily adapted to varying environmental conditions without significant recalibration, especially for thermal equilibrium environments. This reliance also adds a layer of complexity and cost that can be prohibitive for widespread adoption, especially in smaller-scale operations.

Such a framework can be improved by integrating object detection models, which enclose target objects with bounding boxes. The set of models has been constantly improved with model architecture, model parameters, learning structure, dataset comprehensiveness, and computational power used for training. Object detection has its own foundation model, such as YOLO (You Only Look Once) which is widely used in many research (Talaat & ZainEldin, 2023; Xiao et al., 2023), and can be integrated into SAM to automate zero-shot or few-shot image segmentation. YOLOv7, a recent advancement in the series of YOLO models, represents state-of-the-art technology in real-time object detection, emphasizing speed and accuracy across various operational frames per second (FPS) (Wang et al., 2023). This model showcases a significant improvement over its predecessors and other existing models in terms of detection precision and processing speed. The YOLOv7 architecture integrates a robust and streamlined design optimized for speed without sacrificing accuracy, making it highly suitable for real-time applications (Wang et al., 2023; Li et al., 2024; Xia et al., 2022; Peng et al., 2024). The model employs a combination of Cross Stage Partial (CSP) networks and additional enhancements in the backbone that allow for faster computation while reducing the number of parameters. This architecture benefits significantly from advances in convolutional neural networks, utilizing techniques that optimize layer interactions for improved feature extraction and efficiency. The training process of YOLOv7 is notable for its efficiency and effectiveness, partly due to the innovative use of "trainable bag-of-freebies." These methods optimize the training phase to enhance model accuracy without

additional computational cost at inference time. YOLOv7 was trained from scratch on the MS COCO dataset, a comprehensive image dataset popular for object detection tasks, which helps in achieving robustness across varied visual contexts without the need for pre-trained weights.

In a combined framework, an object detection model like YOLOv7 first identifies objects within an image and automatically generates bounding boxes. These bounding boxes, along with a precisely calculated centroids within each box, serve as the prompts for SAM. Consequently, SAM focuses on the prompted areas to produce refined segmentation masks. Such a combined framework cannot only avoid laborious human input prompts densely distributed objects but also enjoys the strengths of two different sets of foundation models.

In sum, the objective of this research was to innovate an automatic segmentation method by combining YOLOv7 and SAM. The proposed method was trained, optimized, and evaluated with a laying hen dataset collected from cage-free housing systems.

Materials and methods

Overall workflow

The workflow of this paper comprises six major components as illustrated in Figure 3.1. The first step involves data collection from the cage-free hen environments. Subsequently, object detection is performed using the YOLOv7 model, which is renowned for its accuracy and efficiency in identifying objects within images. Following this, the coordinates of the bounding boxes obtained by YOLOv7 are extracted, facilitating precise localization of the detected objects. The fourth step introduces the proposed method, YOLOv7 + SAM, which uses both point prompts and box prompts for improved segmentation. The fifth phase encompasses a comparative analysis, benchmarking YOLOv7 + SAM against various state-of-the-art instance segmentation models. This analysis aims to identify the model that delivers the best performance. The final step focuses

on the calculation of evaluation metrics to rigorously assess the efficacy of the models, ensuring the selection of the optimal model for the task.

The sole programming language utilized was Python. Key Python libraries included OpenCV and Pillow for image manipulation, along with Pandas and NumPy for handling data. Matplotlib was used for graphically representing results. Additionally, supplementary libraries employed were pickle, os, Scikit-image, csv, and sys. Computational operations were executed on Google Colab, which provided 12.7 GB of RAM and 16 GB of T4 GPU memory, supported by a dual-core CPU running at 2.30 GHz.

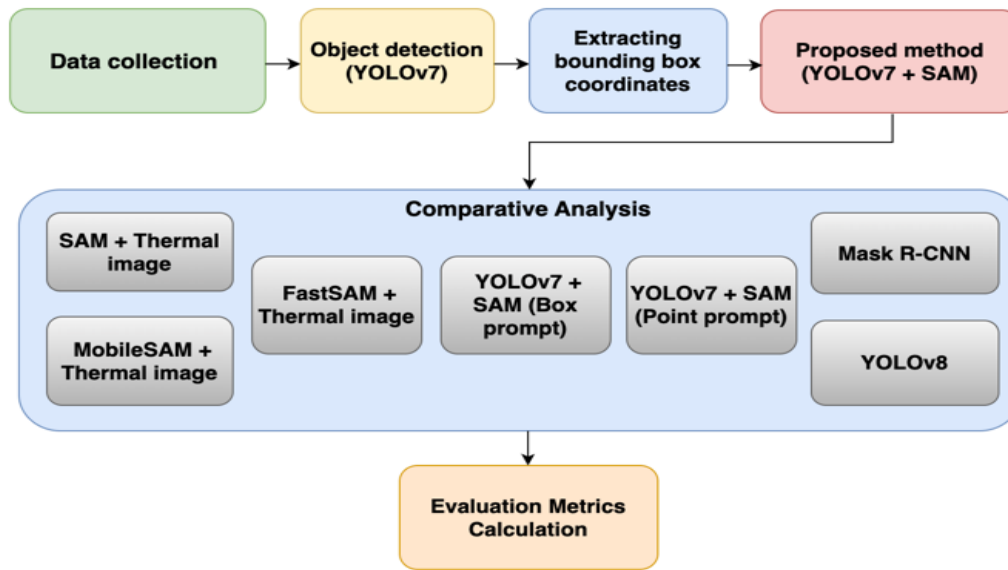


Fig. 3.1. Workflow diagram - This figure presents a schematic of the six-step analytical process employed in the paper. SAM is Segment Anything, YOLO is You Only Look Once, and R-CNN is Region-based Convolutional Neural Network.

Animal, housing, and management

The study took place at the University of Georgia's Poultry Research Center. Four rooms with environmental controls were used, each measuring 7.3 meters in length, 6.1 meters in width, and 3.1 meters in height. Each room housed 180 Hy-Line W-36 laying hens on a litter floor covered with 2.5 cm of pine wood shavings. The rooms also included an A-shaped perch totaling 36.6 meters in length and four nest boxes. The hens were fed an antibiotic-free mash feed during the study. The feed, made at the center's feed mill, had the following nutritional specs: 1.26 MJ/hen/day of metabolizable energy, 16.70 g/day of crude protein, 4 g/day of calcium, and 0.40 g/day of digestible phosphorus. Husbandry, management, and environmental conditions followed the guidelines for Hy-Line W-36 commercial layers (Hy-Line International, 2024). The study's procedures were approved by the Institutional Animal Care and Use Committee (IACUC) under protocol number A2020 08-014-A1, approved on October 5, 2020.

Dataset

Images were taken with a low-cost thermal imaging camera (FLIR C5, Teledyne FLIR, Wilsonville, Oregon, USA) when the birds were 77-80 weeks old. The camera was carefully calibrated with a thermal calibrator (FLUKE 9133, FLUKE, Everett, WA, USA) to ensure the temperature captured in thermal images was correct. Each shot produced a pair of RGB and thermal images, each with a size of 640×480 pixels. A total of 1,917 pairs of images were collected. The images varied widely in pixel intensity, backgrounds, presence of feathers on the ground, inclusion of nest boxes, and instances of overlapping and occlusion among the chickens. Figure 3.2 presents two pairs of samples RGB and thermal images.

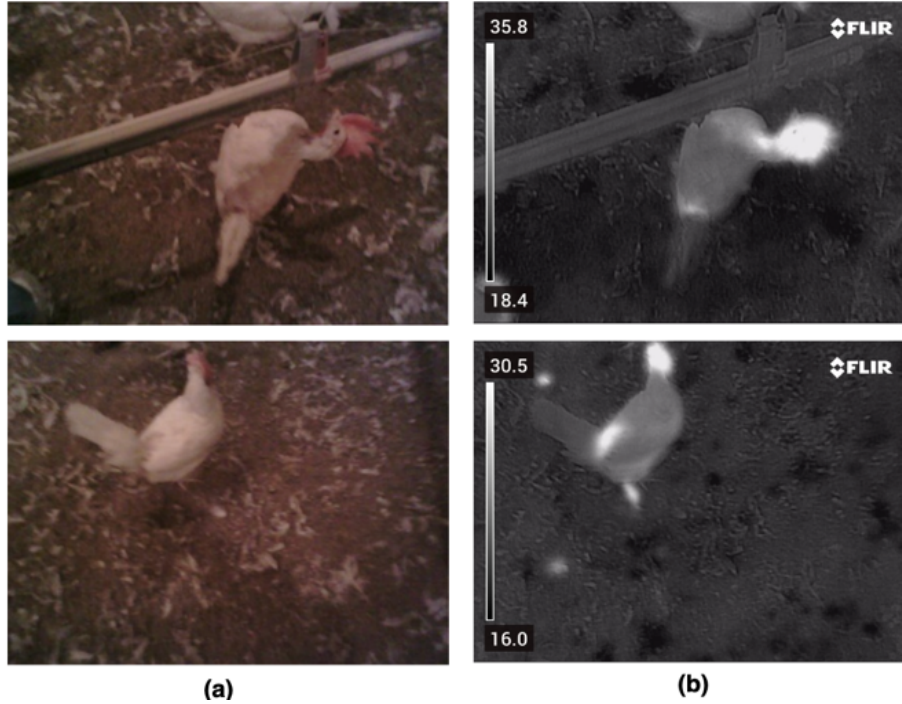


Fig. 3.2. Two pairs of RGB and thermal images: a, c) RGB images; and b, d) corresponding thermal images. Digits inside thermal images indicate maximal or minimal temperatures.

Object detection model (YOLOv7)

YOLOv7, a state-of-the-art object detection model known for its speed and accuracy, was utilized in this study to detect chickens in RGB images. This version of YOLO was chosen due to its balanced trade-off between detection performance and computational efficiency, which is critical for processing large datasets in a reasonable time frame. Additionally, since "bird" is one of the classes YOLOv7 has been trained on in the COCO dataset, the pre-trained YOLOv7 model was transferred and utilized for this task. This allowed the leveraging of its pre-existing knowledge to accurately identify chickens in the images. The output of YOLOv7 detecting chickens in the dataset is shown in Figure 3.3.



Fig. 3.3. YOLO output

Extracting the coordinates of the bounding box

In this phase, the coordinates of the bounding boxes obtained by YOLOv7 were extracted to be used as the initial prompts for SAM. The centroid of each bounding box served as the initial point prompt for SAM. Additionally, the four coordinates of the corners of the bounding box were used as the bounding box prompt for SAM. This setup allowed the precise bird localization provided by YOLOv7, leveraging the performance enhancement of SAM in segmenting the chickens.

Proposed method

The flowchart of the proposed method (YOLOv7 + SAM) is shown in Figure 3.4.

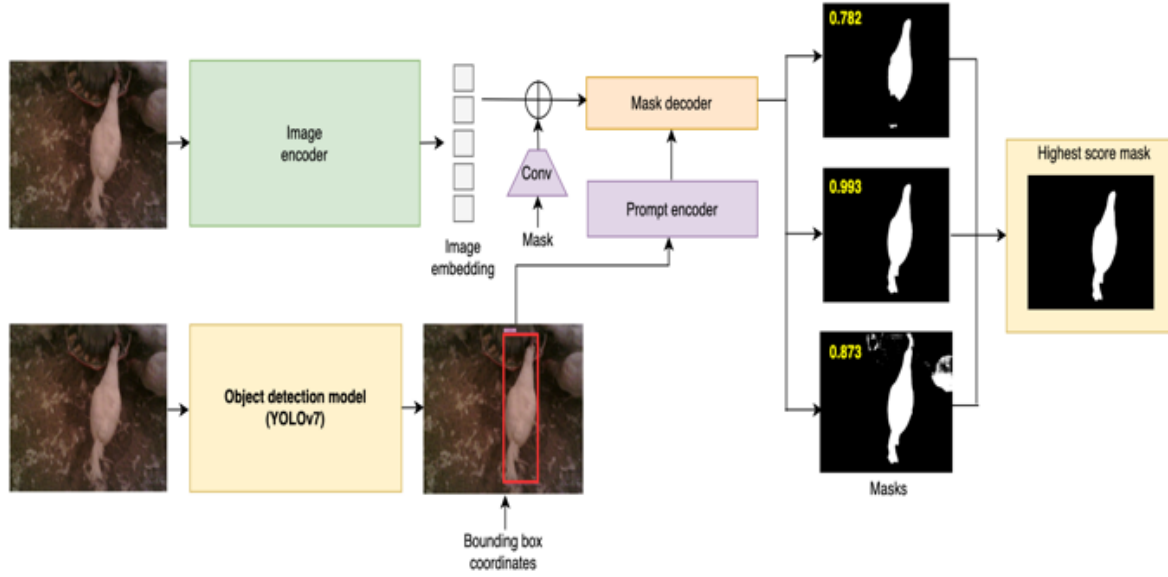


Fig. 3.4. Proposed Method (YOLOv7 + SAM): The method uses a robust image encoder to generate an image embedding. This embedding can be efficiently queried by defining an automatic initial point, enabling the production of bird masks at amortized real-time speed. The yellow numbers on the masks represent the confidence scores of the segmentation, with the mask having the highest score being selected as the final output.

YOLOv7 (box prompts) + SAM

For the chicken segmentation using bounding boxes, the x_{min} , y_{min} , x_{max} , and y_{max} coordinates provided by YOLOv7 were given to SAM, hereafter referred as YOLOv7 (box prompts) + SAM. Using a bounding box as a prompt, narrows down the area for finding the segmented object (i.e., chicken in this case), which makes the segmentation more accurate. Three masks were generated by each segmentation, and the mask with the highest confidence score was chosen.

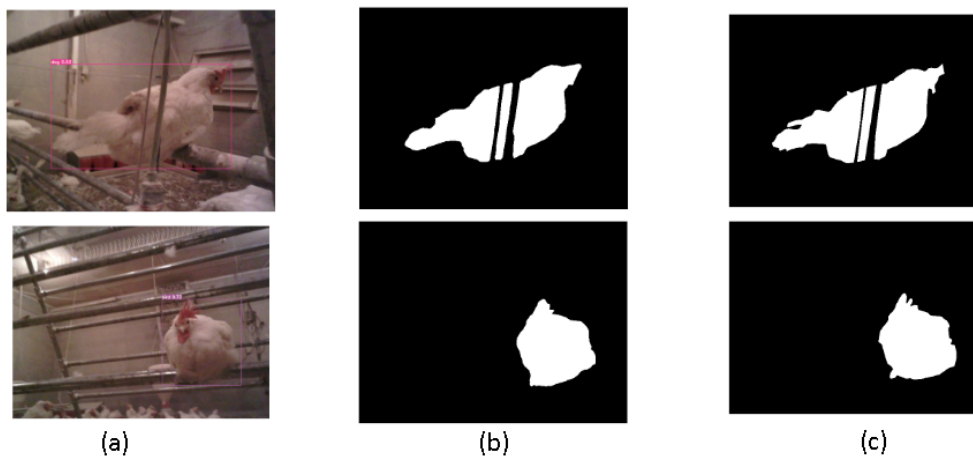


Fig. 3.5. Segmentation results of YOLOv7 (providing bounding boxes of detected birds as box prompts) + SAM: a) detected chickens enclosed with bounding box; b) segmentation results; c) ground truths.

YOLOv7 (point prompts) + SAM

Figure 3.6 shows the procedure of using the centroids of detected birds from YOLOv7 as point prompts for SAM segmentation, hereafter referred as YOLOv7 (point prompts) + SAM.

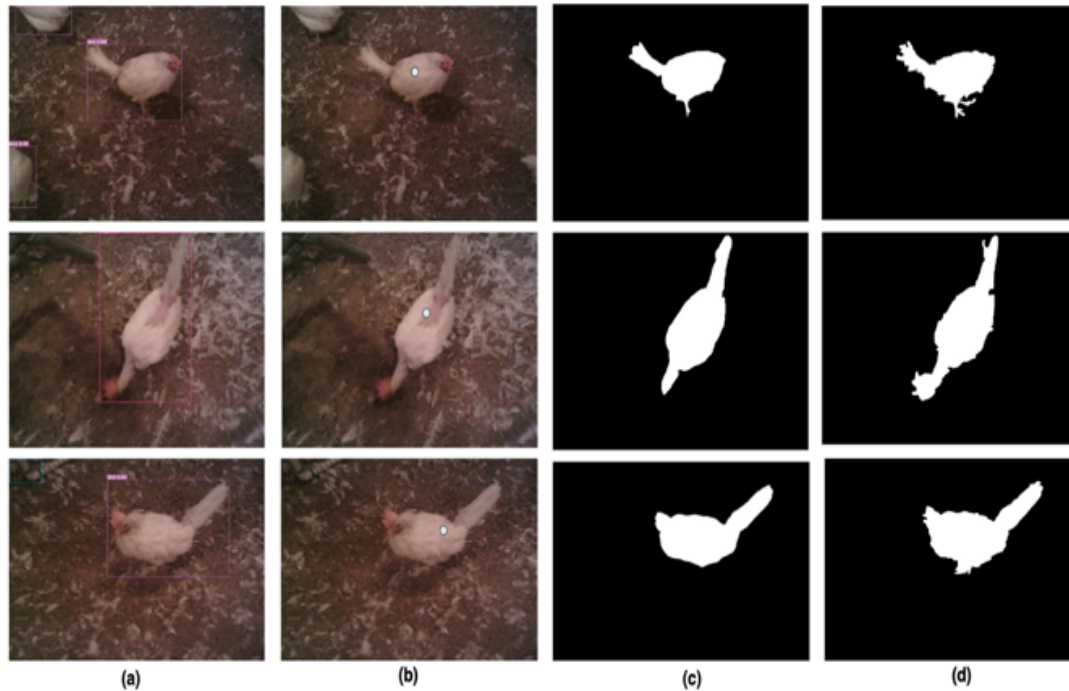


Fig. 3.6. Segmentation results of YOLOv7 (point prompts) + SAM: a) detected chickens enclosed with bounding boxes; b) RGB image with centroid points of detected bounding boxes; c) segmentation results; d) ground truths.

Since the centroids of bounding boxes obtained by YOLOv7 was only a single point, in some images where the orientation of the chicken's body is complex, the centroid may fall outside of the chicken's body, leading to inaccurate segmentation. Figure 3.7 shows some of the erroneous segmentation examples. For example, the hen needs to access nipple drinkers with its head tilted to one side, leading to curly body shape and fallout centroid points.

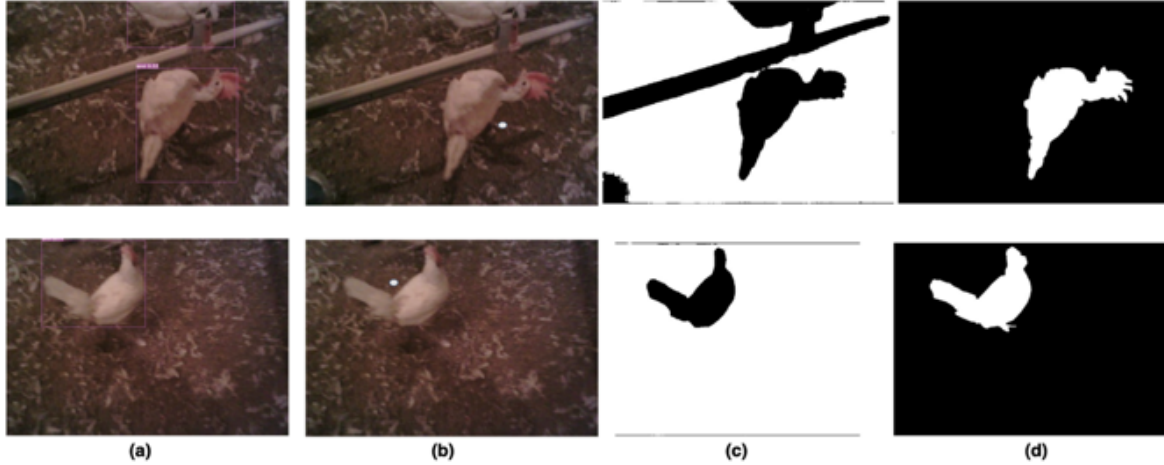


Fig. 3.7. Erroneous segmentation results of YOLOv7 (point prompts) + SAM: a) detected chickens enclosed with bounding boxes; b) RGB image with centroid points of detected bounding boxes; c) erroneous segmentation results; d) ground truths.

Comparative analysis

The performance of the YOLOv7 + SAM combination for zero-shot hen segmentation was verified using four additional deep learning models. These models, which had already been trained on large datasets, were directly used for segmenting hens without requiring extensive additional training. Mask Region-based Convolutional Neural Network (R-CNN), a well-known instance segmentation model, was trained on the COCO (Common Objects in Context) dataset and has been widely used in research. As the target class (i.e., bird) is included in the COCO dataset, this pre-trained model can effectively perform segmentation tasks without extra training.

FastSAM, a CNN-based model, stands out for its speed due to its training on just 2% of the SA-1B dataset, which contains 1 billion masks for training general-purpose object segmentation models like SAM. The MobileSAM or FasterSAM improved processing speed by substituting the original bulky ViT-H (632 million parameters) encoder of SAM with a more

compact Tiny-ViT (5 million parameters). Generally, FastSAM and MobileSAM are extensive versions of the original SAM, with compressed parameters to enhance processing speed.

YOLOv8, the newer YOLO model, can be used for object detection, image classification, and instance segmentation. The YOLOv8n version offers good performance on edge devices, balancing detection accuracy and computational resources. Both Mask R-CNN and YOLOv8 have the potential to effectively segment hens without extensive additional training, but their performance in this study needs to be verified.

Additionally, a previous study (Thermal images + SAM) used the characteristics of thermal images, along with some pre- and post-processing steps, as the initial prompt for SAM. This approach leveraged the unique features of thermal images to enhance SAM's performance in segmenting the hens.

Evaluation metrics calculation

This study utilized a comprehensive set of evaluation metrics to independently assess the performance of both segmentation and detection. The segmentation model evaluation was based on a dataset of 1,917 RGB images of individual chickens. These images were annotated by a skilled technician using Roboflow, ensuring high-precision masks that accurately depicted each chicken. The author then conducted a double verification to ensure the accuracy and quality of the labeling. This rigorous ground truth served as the benchmark for evaluating the accuracy of the segmentation models.

The performance of the trained models was assessed using precision, recall, F1 score, and Intersection over Union (IoU) as detailed in Equations (3.1), (3.2), (3.3), and (3.4). Precision determines how accurately the model identifies only the relevant pixels for segmentation, calculated by the proportion of correctly predicted positives to the total predicted positives. Recall,

also known as sensitivity, evaluates the model's effectiveness in detecting all relevant pixels, calculated by the proportion of correctly predicted positives to the total actual positives. The F1 Score, representing the harmonic mean of precision and recall, gauges the model's overall accuracy, with 1 being the optimal value indicating perfect precision and recall, and 0 the lowest. IoU measures the overlap between the predicted segmentation and the ground truth, calculated by dividing the overlapping area by the combined area of the predicted segmentation and the ground truth.

$$Precision = \frac{True\ positive}{True\ positive + False\ positive} \quad (3.1)$$

$$Recall = \frac{True\ positive}{True\ positive + False\ negative} \quad (3.2)$$

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3.3)$$

$$IoU = \frac{True\ positive}{True\ positive + False\ positive + False\ negative} \quad (3.4)$$

where true positive refers to pixels that are correctly identified as part of the birds; false positive are the pixels that the segmentation model incorrectly identifies as part of the birds, but they actually belong to the background or other objects; false negative is used for pixels that are part of the birds in the ground truth but are missed by the segmentation model.

The detection metric employed is the success rate, which is based on the IoU value. A successful segmentation is one where the IoU is 50% or greater, which aligns with standard thresholds used in prominent publications (Girshick, 2015; He et al., 2017; Redmon and Farhadi, 2018) as shown in Equation (3.5). The success rate thus reflects the percentage of images in which the models successfully segmented the chicken areas.

$$\text{Success rate} = \frac{\text{Number of successfully segmented images (IoU} > 0.5\text{)}}{\text{Total number of images}} \quad (3.5)$$

Results and discussion

Image similarity scores

The histogram as shown in Figure 3.8 illustrates the distribution of RGB image similarity scores within the sampled subset. The scores ranged from 0 to 1 with 0 representing a completely different image and 1 representing an identical/duplicate image. The similarity score ranges from 0.3 to 0.7, with the majority of scores concentrated around the 0.4 to 0.5 range. The histogram's shape and distribution indicate several key points about the diversity of the dataset:

- **Wide Range of Similarity Scores:** The similarity scores span a broad range (0.3 to 0.7), indicating that there was a mix of both highly similar and highly dissimilar images within the dataset. This range suggests that the dataset includes a variety of images rather than being dominated by very similar or very dissimilar images.
- **Normal Distribution:** The distribution of similarity scores forms a bell-shaped curve, resembling a normal distribution. This suggests a balanced dataset where most image pairs had moderate similarity, with fewer pairs being either very similar or very dissimilar. This balance is indicative of diversity, as it implies that the dataset does not have a bias towards a specific type of image content.

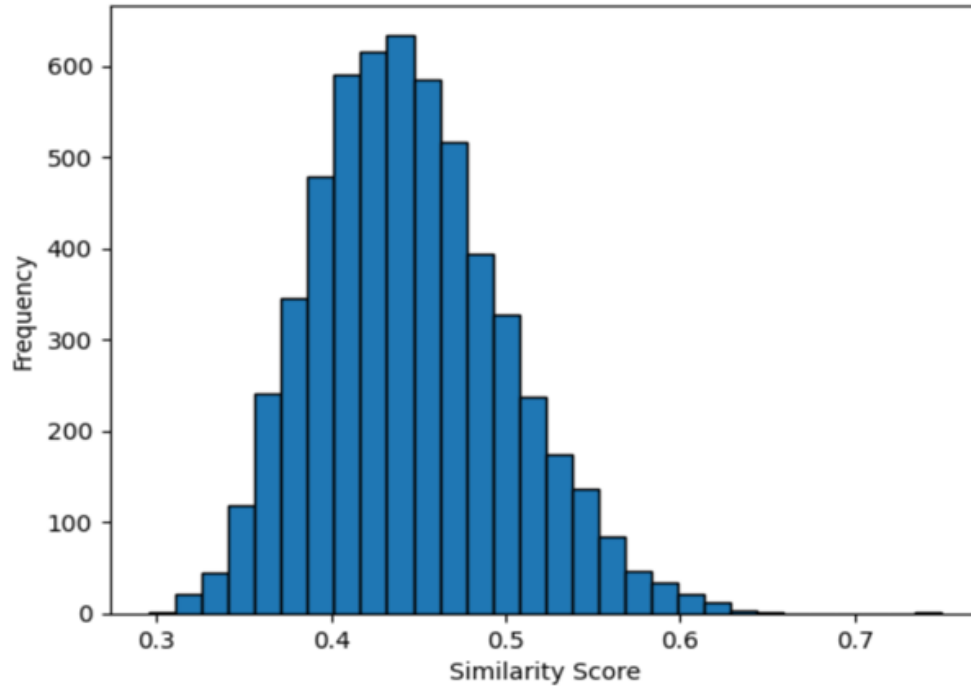


Fig. 3.8. Similarity score histogram of RGB images in the dataset

The histogram provides evidence that the dataset is diverse, with a wide range of similarity scores and a balanced distribution of similarities. Despite the computational limitations that required analyzing a subset of smaller images, the results indicate that the dataset contains a variety of images with differing levels of similarity, which should comprehensively evaluate the capability of the proposed methods for bird segmentation.

Selecting the optimal model for zero-shot hen detection

Table 3.1 presents a comparative analysis of hen detection performance using various zero-shot segmentation algorithms. The primary metric for this evaluation was the success rate, defined as the percentage of images with an IoU greater than 50%. Images meeting or surpassing this 50% IoU threshold were classified as successfully detected and segmented, thereby positively impacting the overall success rate.

Table 1

Table 3.1. comparative analysis of hen detection performance

Models	Success Rate (%)
YOLOv8	50.0
Thermal image + MobileSAM	82.0
Thermal image + SAM	84.4
Thermal image + FastSAM	72.8
Mask R-CNN	64.2
YOLOv7	83.2
YOLOv7 (box prompts) + SAM	98.0
YOLOv7 (point prompts) + SAM	94.0

In reviewing the data in Table 3.1, it becomes clear that the proposed method (YOLOv7 (box prompts) + SAM) achieved a significantly higher success rate in detecting and segmenting hens compared to other methods. This metric is important because it measures the percentage of instances where the algorithm accurately identifies and delineates the birds. The comparative analysis shows that the improved success rate of the proposed method, especially when compared to SAM using point prompts from thermal imaging (which required pre- and post-processing steps), highlights the significant impact of the modifications on the algorithm's efficiency. These techniques likely enhance the model's ability to distinguish between the hens and their surroundings by optimizing the quality of the input data. With this new method, pre- and post-processing steps are unnecessary since the initial prompt is chosen effectively, resulting in optimal output masks without selecting the best mask by a well-trained machine learning classifier.

Conversely, the lower success rate (50.0%) of YOLOv8 highlights a limitation in its zero-shot detection capabilities, particularly in accurately identifying chickens. Although YOLOv8 and Mask R-CNN are robust models, their performance is not optimized for this specific application without training on a dataset specific to chickens. The main goal was to use a zero-shot instance segmentation method that functions without requiring any image training. This approach aligns with the broader objective of deploying efficient and adaptable models capable of handling various segmentation tasks with minimal setup, emphasizing the value of SAM as a promising tool in zero-shot segmentation scenarios.

Comparison of different models for hen segmentation

To assess the segmentation capabilities of various models, Table 3.2 presents a comparison of different segmentation metrics across the models.

Table 3.2. comparison of different segmentation metrics across the models.

Models	Evaluation criteria (%)			
	Precision	Recall	F1 Score	IoU
YOLOv8	97.4	81.4	88.4	79.5
Thermal image + MobileSAM	91.8	90.7	91.2	83.6
Thermal image + SAM	93.6	91.0	92.3	85.5
Thermal image + FastSAM	92.5	90.4	91.4	84.1
Mask R-CNN	87.5	90.2	88.8	79.9
YOLOv7 (box prompts) + SAM	92.5	98.2	95.1	91.0

YOLOv7 (point prompts) + SAM	92.3	93.3	92.7	86.6
------------------------------	------	------	------	------

All the models exhibit relatively consistent segmentation metrics, except for YOLOv7 (box prompts) + SAM, which significantly outperforms the others. This superior performance suggests that using a well-chosen bounding box prompt effectively captures the entire chicken body, as it covers the chicken's full extent rather than just a single point, like the point prompt. The point prompt is generally less effective because it provides a less comprehensive representation of the bird, resulting in poorer generalization.

The bounding box prompt offers a complete outline of the birds, ensuring that all relevant parts are included in the segmentation process. This comprehensive approach enables the model to generalize better and accurately segment the entire bird. On the other hand, a point prompt focuses on a single location, often missing parts of the bird and leading to incomplete and less reliable segmentation. This distinction underscores the importance of selecting the appropriate prompt type to enhance model accuracy and reliability. By covering the entire bird, the bounding box prompt allows the model to understand the context and boundaries more effectively, resulting in higher segmentation accuracy. This method does not only improve the model's ability to capture details but also enhances its capability to generalize across different instances of the object.

The success of YOLOv7 (box prompts) + SAM highlights the critical role of prompt selection in optimizing segmentation performance.

Figure 3.9 illustrates the segmentation results achieved by different models. According to the figure, the YOLOv7 (box prompts) + SAM could segment a more complete body.

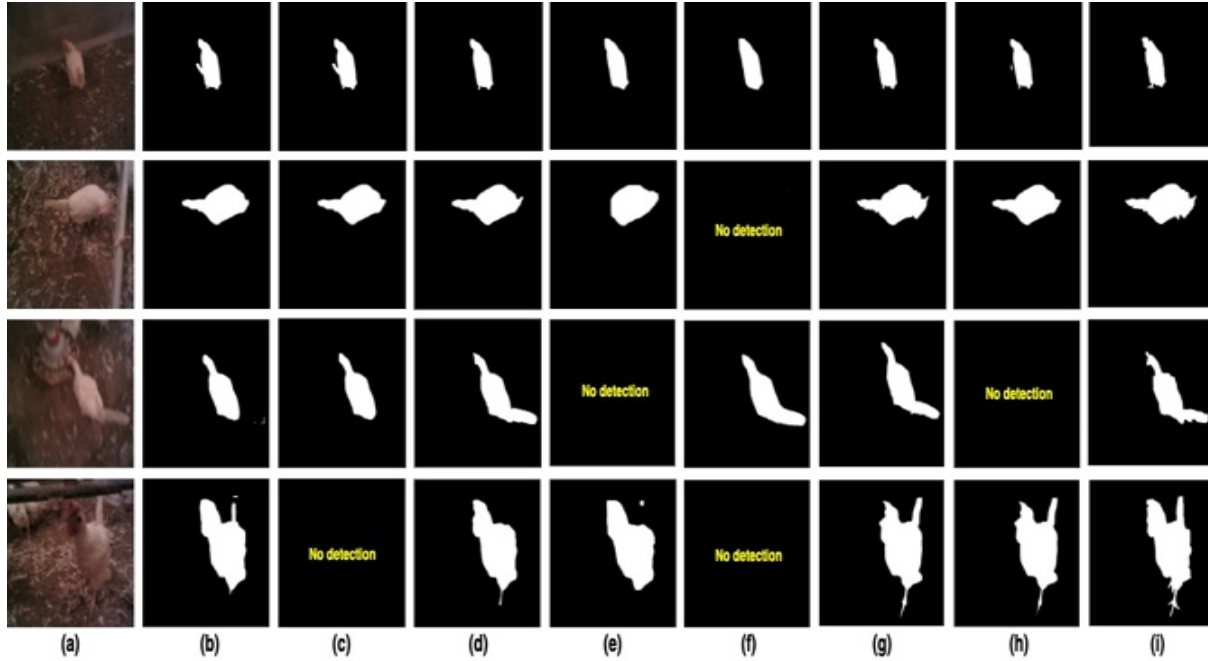


Fig. 3.9. Segmentation results of different models: a) RGB image; b) thermal image + MobileSAM; c) thermal image + FastSAM; d) thermal image + SAM; e) YOLOv8; f) Mask R-CNN; g) YOLOv7 (box prompt) + SAM; h) YOLOv7 (point prompts) + SAM; and i) ground truth.

Conclusion

The comprehensive evaluation of various zero-shot segmentation models for hen detection highlights the superior performance of the YOLOv7 + SAM model with a bounding box prompt, achieving an impressive 98.0% success rate. This model effectively captures the entire bird, reducing the need for extensive pre-processing and post-processing, and outperforms other models such as YOLOv8 and Mask R-CNN, which showed limitations without specific training. The success of YOLOv7 + SAM underscores the potential of zero-shot segmentation techniques to provide flexible, efficient solutions in specialized applications, setting a benchmark for future tasks and promising significant enhancements in operational efficiency and monitoring accuracy in agricultural settings.

Acknowledgement

This project was jointly funded by US Poultry & Egg Association (project #: F-113); USDA-NIFA (2023-68008-39853), and USDA National Institute of Food and Agriculture Hatch Project.

References

A. Kirillov et al., "Segment anything: Building foundation models for image segmentation," Meta AI Research, 2023. [Online]. Available: <https://segment-anything.com>

R. Bommasani, D. A. Hudson, E. Adeli, R. Altman, S. Arora, S. von Arx, M. S. Bernstein, J. Bohg, A. Bosselut, E. Brunskill et al., "On the opportunities and risks of foundation models," arXiv preprint arXiv:2108.07258, 2021.

X. Zhao, W. Ding, Y. An, Y. Du, T. Yu, M. Li, M. Tang, and J. Wang, "Fast Segment Anything," arXiv, 2023. [Online]. Available: <https://doi.org/10.48550/arXiv.2306.12156>.

Zhang, C., Han, D., Qiao, Y., Kim, J.U., Bae, S.-H., Lee, S., Hong, C.S., 2023. Faster Segment Anything: Towards Lightweight SAM for Mobile Applications. <https://doi.org/10.48550/arXiv.2306.14289>.

J. Ma, Y. He, F. Li, L. Han, C. You, and B. Wang, "Segment anything in medical images," Nat. Commun., vol. 15, p. 654, 2024. [Online]. Available: <https://doi.org/10.1038/s41467-024-44824-z>.

M. A. Mazurowski, H. Dong, H. Gu, J. Yang, N. Konz, and Y. Zhang, "Segment anything model for medical image analysis: An experimental study," Medical Image Analysis, vol. 89, p. 102918, 2023. [Online]. Available: <https://doi.org/10.1016/j.media.2023.102918>.

L. P. Osco, Q. Wu, E. L. de Lemos, W. N. Gonçalves, A. P. M. Ramos, J. Li, and J. Marcato, "The Segment Anything Model (SAM) for remote sensing applications: From zero to

one shot," *International Journal of Applied Earth Observation and Geoinformation*, vol. 124, p. 103540, 2023. [Online]. Available: <https://doi.org/10.1016/j.jag.2023.103540>.

P. Shi, J. Qiu, S. M. D. Abaxi, H. Wei, F. P.-W. Lo, and W. Yuan, "Generalist Vision Foundation Models for Medical Imaging: A Case Study of Segment Anything Model on Zero-Shot Medical Segmentation," *Diagnostics*, vol. 13, p. 1947, 2023. [Online]. Available: <https://doi.org/10.3390/diagnostics13111947>.

T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell et al., "Language models are few-shot learners," *Advances in neural information processing systems*, vol. 33, pp. 1877–1901, 2020. J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

J. Hoffmann, S. Borgeaud, A. Mensch, E. Buchatskaya, T. Cai, E. Rutherford, D. d. L. Casas, L. A. Hendricks, J. Welbl, A. Clark et al., "Training compute-optimal large language models," *arXiv preprint arXiv:2203.15556*, 2022.

J. Kaplan, S. McCandlish, T. Henighan, T. B. Brown, B. Chess, R. Child, S. Gray, A. Radford, J. Wu, and D. Amodei, "Scaling laws for neural language models," *arXiv preprint arXiv:2001.08361*, 2020.

A. Chowdhery, S. Narang, J. Devlin, M. Bosma, G. Mishra, A. Roberts, P. Barham, H. W. Chung, C. Sutton, S. Gehrmann et al., "Palm: Scaling language modeling with pathways," *Journal of Machine Learning Research*, vol. 24, no. 240, pp. 1–113, 2023.

J. L. Edgar, C. J. Nicol, C. A. Pugh, and E. S. Paul, "Surface temperature changes in response to handling in domestic chickens," *Physiology & Behavior*, vol. 119, pp. 195–200, 2013. [Online]. Available: <https://doi.org/10.1016/j.physbeh.2013.06.020>.

M. Saeidifar, G. Li, L. Chai, and R. B. Bist, "Zero-shot image segmentation for monitoring thermal conditions of individual laying hens," in Proc. Int. Poultry Scientific Forum 2024, 2024, p. 41. [Online]. Available: <https://www.ippexpo.org/education-programs/IPSF/docs/2024-IPSF-Abstracts.pdf>.

F. M. Talaat and H. ZainEldin, "An improved fire detection approach based on YOLO-v8 for smart cities," **Neural Computing and Applications**, vol. 35, pp. 20939–20954, 2023. [Online]. Available: <https://doi.org/10.1007/s00521-023-08809-1>.

B. Xiao, M. Nguyen, and W. Q. Yan, "Fruit ripeness identification using YOLOv8 model," **Multimedia Tools and Applications**, 2023. [Online]. Available: <https://doi.org/10.1007/s11042-023-16570-9>.

W. C. Y., A. Bochkovskiy, and H. Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 7464-7475.

Z. Li, Y. Zhu, S. Sui, Y. Zhao, P. Liu, and X. Li, "Real-time detection and counting of wheat ears based on improved YOLOv7," *Computers and Electronics in Agriculture*, vol. 218, p. 108670, 2024.

Y. Xia, M. Nguyen, and W. Q. Yan, "A real-time kiwifruit detection based on improved YOLOv7," in International Conference on Image and Vision Computing New Zealand, Cham, Switzerland: Springer Nature Switzerland, Nov. 2022, pp. 48-61.

X. Peng, Y. Chen, X. Cai, and J. Liu, "An Improved YOLOv7-Based Model for Real-Time Meter Reading with PConv and Attention Mechanisms," *Sensors*, vol. 24, no. 11, p. 3549, 2024.

"Management Guide, W-36 Commercial layers, 2024," Hy-Line International. [Online]. Available: <https://www.hyline.com/>. [Accessed: Feb. 12, 2024].

CHAPTER IV

ANIMALAI: AN OPEN-SOURCE WEB PLATFORM FOR AUTOMATED ANIMAL ACTIVITY INDEX CALCULATION USING INTERACTIVE DEEP LEARNING SEGMENTATION

Monitoring of the activity index of animals is considered crucial for assessing their welfare and behavior patterns. However, traditional methods for calculating the activity index, such as pixel intensity differencing of entire frames, are often found to suffer from significant interference and noise, leading to inaccurate results. The classical activity index method is also lacking in the capability to measure the activity index of individual animals, making it impossible to track the movement of specific animals within a group. Furthermore, no free and accessible online platform is currently available for non-technical researchers to calculate animal activity index, thereby creating a gap in the tools available for animal welfare studies. Tracking all individual animals in a video can be computationally expensive. The objectives of this research were to 1) develop a user-friendly, open-source platform using Streamlit to enable researchers to calculate the activity index of animals, either individually or in groups, from video footage; and 2) explore the representative proportion of animals to depict the whole group activity index, for saving computing time and resources. Top-view videos can be easily uploaded, and animals can be selected for targeted tracking. A general deep learning-based image segmentation model, the Segment Anything Model² (SAM2) that is a promptable segmentation model, was used to segment and track individual animals across frames without the need for extensive training or annotation.

Consistent and accurate segmentation and tracking were ensured by the platform, thereby overcoming the challenges posed by noise and interference in classical methods. The SAM2 segmented and tracked Cobb500 male broiler chicken in videos from weeks 1 to 7 with a segmentation success rate of 100%, Intersection over Union (IoU) of $92.21\% \pm 0.012$, precision of $93.87\% \pm 0.019$, recall of $98.15\% \pm 0.011$, and F1 score of $95.94\% \pm 0.006$. These metrics were calculated from 1,157 individual chickens. Statistical analysis revealed that tracking 80% of birds in week 1, 60% in week 4, and 40% in week 7 was significantly sufficient ($r \geq 0.90$; $P \leq 0.048$) to depict the overall flock movement. This user-friendly tool is provided to researchers as an accessible and efficient way to track and analyze animal behavior patterns, delivering accurate and reliable insights into animal welfare at both the individual and group levels without requiring extensive programming knowledge.

Introduction

Animal activity plays a pivotal role in understanding welfare, health, and behavior patterns across various livestock species (Bocaj et al., 2020; Oso et al., 2025; Tran et al., 2022). In modern animal production systems, continuous observation and prompt detection of abnormal behaviors are paramount for maintaining high standards of welfare and maximizing productivity (Elbarrany et al., 2023). Capturing animal activity—broadly defined as the frequency or extent of movement over time—can offer valuable insights for both researchers and producers to make evidence-based decisions. In poultry, for example, sudden changes in flock movement may indicate issues like heat stress or disease outbreaks. In cattle and pigs, activity patterns can help detect lameness or identify periods of increased stress (Chen et al., 2021; Fuentes et al., 2020). As such, techniques that enable robust and efficient estimation of the activity are indispensable.

In order to quantify animal activity, the activity index, a measure of movement intensity through image processing, was proposed by (Bloemen et al., 1997). Activity index was defined as the percentage of pixels of moving objects to the total number of pixels within the image (including animals and background). In more recent research, the total number of pixels was replaced with total bird-representative pixels to compensate for variations in animal size at different ages (Aydin et al., 2010; Li et al., 2020; Silvera et al., 2017). Since the concept was coined, the activity index has been widely used to quantify the activities of broilers (Kristensen et al., 2006; Neves et al., 2015). The concept has been applied to develop a commercial computer vision system, named eYeNamic, and the vision system has been applied in several European studies (Peña Fernández et al., 2018; Silvera et al., 2017).

While classical activity index calculation method can be quick to implement and computationally straightforward, it tends to be highly sensitive to noise and environmental factors such as lighting fluctuations, camera vibrations, or background movements such as human interference (Sengar and Mukhopadhyay, 2017). Moreover, applying pixel intensity differencing to an entire scene restricts researchers to group-level activity assessments. In many practical scenarios, especially those involving large populations of animals housed together, the interest lies in pinpointing the movements of specific individuals. Without the ability to segment and track individual animals, vital data—such as determining which animals are underactive or hyperactive—remain inaccessible.

Deep learning-based methods have substantially advanced object detection and segmentation in recent years (Li and Chai, 2023; Saeidifar et al., 2024; Shams et al., 2023). However, developing a specialized segmentation model for each livestock species or experimental setup can be prohibitively time-consuming and expensive. Researchers would need to curate and

annotate large image datasets, train convolutional neural networks or transformers, and then continuously update these models as lighting conditions, camera angles, or animal growth stages change. This complexity has motivated the rise of more generalized segmentation models that are pre-trained on vast and diverse image corpora, allowing them to perform “zero-shot” or “few-shot” segmentation on new types of objects (Ravi et al., 2024). One such model is SAM2, a powerful variant of the foundational SAM (Ravi et al., 2024). SAM2 has been lauded for its ability to quickly and accurately identify objects of interest with minimal prompting, effectively reducing the need for large-scale annotation (Ravi et al., 2024). Unlike traditional models that often fail when confronted with new species or environments, SAM2 has been broadly trained with billions of image masks, enabling it to handle a wide range of scenes and animal morphologies. Alongside these advances in segmentation, there has also been a growing need for accessible, user-friendly platforms that can seamlessly integrate deep learning into everyday research workflows. The user-friendly platforms are especially important for scholars who do not have sufficient computing backgrounds for coding but would love to use the automatic tools to support animal research for advancing animal products.

Several user-friendly platforms were developed in the animal behavior domain to assist researchers in tracking and analyzing animal movements. For instance, the AnimalAccML integrated multiple machine learning models and feature engineering techniques and enabled users to automatically analyze behaviors of with several mouse clicks based on triaxial accelerometer data, which is not suitable for computer vision-based metric analytics (Li and Chai, 2023). DeepLabCut was a widely adopted open-source tool that leverages deep learning for markerless pose estimation in images/videos (Mathis et al., 2018). Its user-friendly interface made it popular among researchers; however, it generally required extensive manual annotation and a considerable

amount of training data to adapt to different species or experimental conditions. This reliance on manual setup hindered rapid deployment in novel environments and limits its utility for studies that require immediate or real-time analysis. Another notable example is *idtracker.ai*, which offered automated tracking of individual animals within groups (Romero-Ferrero et al., 2019). While it simplified the tracking process and is relatively intuitive, *idtracker.ai* tended to be computationally intensive, especially when dealing with large groups or high-resolution video footages. Moreover, its performance degraded in scenarios with significant noise, variable lighting, or complex backgrounds, thereby reduced its reliability in accurately capturing animal movement dynamics (Dell et al., 2014).

Despite the advancements these platforms represent, they were not designed to compute the animal activity index automatically. Their primary focus lies in detailed tracking and pose estimation rather than in providing a comprehensive, user-friendly solution for calculating movement-based metrics such as the activity index at either the individual or group level. In conclusion, while current tools offer valuable functionalities in animal tracking and behavior analysis, there remains a notable gap: there is currently no user-friendly platform that automatically calculates the animal activity index, highlighting an unmet need in animal welfare research and monitoring.

Several studies in the field of collective animal behavior have demonstrated that monitoring a representative subset of individuals can effectively capture the overall dynamics of a group, aiming to improve computational efficiencies. For example, in a study investigating the spatial organization and interaction rules within starling flocks, researchers found that each bird interacted with a fixed number of neighbors (six to seven) rather than all nearby individuals. This topological interaction enabled flocks to maintain cohesion and coordinated movement, even

under changing densities and external perturbations. Although the study did not directly address representative sampling, the idea that a limited number of local interactions govern the behavior of the entire group implies that monitoring a subset of individuals could reveal key aspects of collective dynamics (Ballerini et al., 2008). Similarly, in another study on the collective behavior of midge swarms, researchers found that individual midges were strongly connected, even beyond their nearest neighbors. Even in the absence of global order, midges exhibited coherent movement patterns that could be explained by localized interactions. Their study demonstrated that these correlations reflect emergent group-level behavior, suggesting that sampling a fraction of individuals can provide reliable insights into the overall dynamics of the swarm. By employing simulations of interacting particles, they further showed that local measurements could scale up to describe the collective response of the entire group (Attanasi et al., 2014).

Our exploration of different sampling ratios (20%, 40%, 60%, and 80%) across key growth stages in broilers addressed this gap. By systematically determining the optimal proportion of birds needed to accurately represent the entire flock's activity, our study provided a practical framework that reduced computational demands without compromising the reliability of activity index measurements. This tailored approach is particularly relevant for commercial applications, where rapid and resource-efficient monitoring is essential for effective animal welfare management. The objectives of this research were to 1) develop a user-friendly, open-source platform to enable researchers to calculate the activity index of animals, either individually or in groups, from video footage; and 2) explore the representative proportion of animals to depict the whole group activity index, for saving computing time and resources.

Materials and Methods

Animal housing and video data collection

For the purpose of validating the segmentation model, a subset of a larger video dataset was used. This dataset was collected at the University of Georgia's Poultry Research Center during May–June 2024. A total of 1,776 day-old Cobb 500 broiler chickens were randomly assigned to 48 pens, with 37 birds being allocated per pen, within two environmentally controlled rooms. The rooms were measured to be approximately 17.2 m in length by 11.4 m in width and were subdivided into two rows of 12 identical pens, each of which measured 1.2 m by 3.0 m. Two feeders were provided at opposite ends of every pen, and two centrally located drinking lines were installed. Standard environmental conditions were maintained in accordance with the Cobb management guidelines (Cobb, 2022), with feed and water provided *ad libitum*. Lighting and temperature adjustments were made according to age-specific protocols throughout the rearing period. Video recordings were acquired using overhead security cameras (NHD-887MSB, Swann Security, Santa Fe Springs, CA) that were mounted on the ceiling at a distance of approximately 3.05 m above each pen. Continuous recordings were managed by 16-channel video recorders (SRDVR-85680H-US, Swann Security, Santa Fe Springs, CA). The recordings were set at a resolution of 1024×768 pixels and at 15 frames per second (fps), and the video data were stored as .MP4 files on a 20-terabyte external hard disk. A total of 34 videos from week 1 through week 7 were selected, and 1157 individuals were used for evaluation. Although all birds from the large study were included in the complete dataset, the subset for evaluation was chosen so as to ensure a representative distribution across developmental stages from week 1 (early phase), week 4 (medium phase), and week 7 (late phase). All experimental procedures, including the video recordings, were performed in compliance with protocols approved by the Institutional Animal

Care and Use Committee (IACUC) at the University of Georgia (protocol number: A2023 07-016-Y1-A0).

Overall workflow

Figure 4.1 illustrates the Streamlit-based interface workflow for calculating the animals' activity index. Once the application was launched, a user-friendly graphical interface was loaded in the default web browser. The user could upload a video of up to one hour in length for the convenience of data visualization. If a video exceeded this duration, the interface issued a warning and recommended trimming. Subsequently, key parameters, such as frame interval, can be specified by the user. The system extracted frames from the uploaded video and displayed the first frame so the user can pinpoint, via mouse click, the location of the animal or region of interest. If the user was dissatisfied with the selected coordinate, an 'undo' option reverted the choice until the coordinate was precisely defined. After confirming the chosen coordinates, the interface proceeded to segment the video, generating both an RGB mask frame and a binary mask frame. This segmentation underpinned the computation of an activity index, which was then plotted and viewable within the interface. Additionally, the activity index plot, as well as the normalized activity index for each consecutive frame, were saved as a PNG and TXT file, respectively. In addition, users can inspect frames derived from frame differencing for a more detailed overview of movement and check whether the segmentation was successful or not. Although Figure 1 shows a typical workflow, users may adjust certain steps (e.g., re-uploading trimmed clips or revisiting parameter settings) according to their experimental needs. Detailed descriptions of each phase and the options offered by the Streamlit interface are provided in the following sections. The interface was published on GitHub (<https://github.com/MahtabSaeidifar/AnimalAI>) for open access.

In this study, the entire platform was developed solely using Python, which enabled all components to be consolidated into a single consistent computing environment to enhance code readability and maintainability. The most important packages used in our platform were torch (v2.4.1) for deep learning, streamlit (v1.19.0) for developing interactive web applications, numpy (v1.26.4) for numerical computations, pandas (v1.4.2) for data manipulation, matplotlib (v3.9.2) for data visualization, and jupyterlab (v4.2.4) for providing an interactive development environment. In addition, the SAM2 package was installed directly from its GitHub repository (<https://github.com/facebookresearch/sam2>) to facilitate segmentation tasks. The computer used for platform development and evaluation was equipped with a 13th Gen Intel® Core™ i7-13700 processor, featuring 24 logical CPUs with clock speeds ranging from 0.8 GHz to 5.2 GHz, 62 GiB of installed RAM, and a 64-bit operating system.

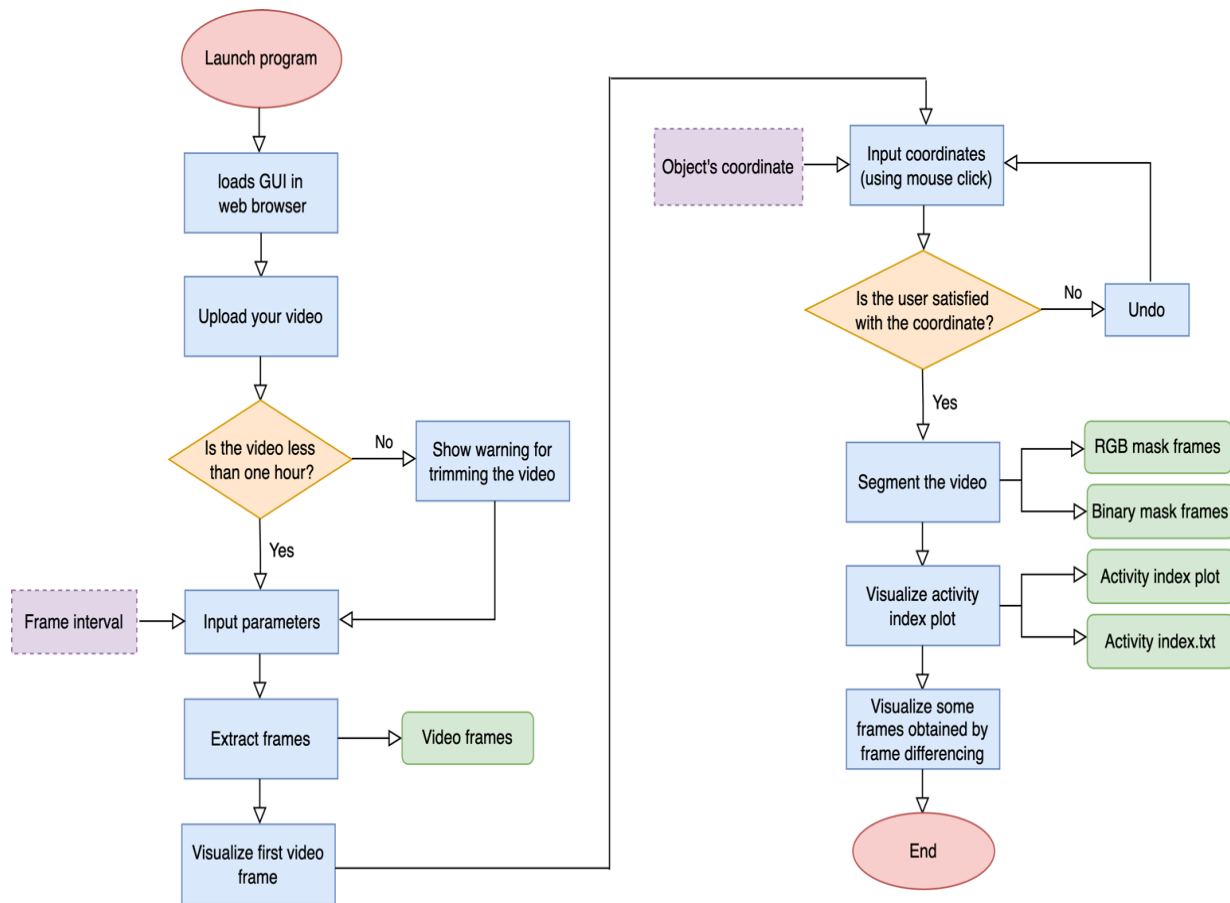


Fig. 4.1. Workflow diagram of the platform - This figure presents a schematic of different steps analytical process employed in the platform. Red color indicates start and end points of the process; blue color indicates main processing steps; orange color indicates decision points; purple color indicates user input parameters; and green color indicates files saved in the main directory.

Video uploading

Once the application was launched, a user-friendly graphical interface was displayed in the web browser, allowing for an intuitive interaction. The interface prompted the user to upload a video, accepting various formats (e.g., MP4, MOV, AVI, and MPEG4). The recommended maximum duration for the video was one hour; if the uploaded file exceeded this length, the system automatically issues a warning and advised trimming the video to under one hour. This

recommendation helps ensure faster processing times and reduces computational overhead during subsequent steps.

Video frame extraction

Once the user has uploaded a video, the application automatically evaluated the duration of the file and generated a range of recommended frame intervals. These recommendations aimed to strike a balance between capturing sufficient details and minimizing both storage requirements and computational resources. While users are free to override the recommended settings and specify a custom interval, adhering to the suggested range is generally preferred for optimal efficiency and data manageability.

By selecting an interval, the user essentially controlled the frequency of frames to be extracted: smaller intervals yield more frames (allowing for finer-grained analysis) but required greater storage and computational power, whereas larger intervals reduced the number of frames extracted and offer lowered storage demands at the potential cost of missing some subtle movements. After choosing a frame interval, users can click the ‘Extract Frames’ button to trigger the extraction process as shown in Figure 4.2. The resulting frames were automatically stored in a designated directory, ensuring that they can be readily accessed in subsequent stages of the workflow (e.g., segmentation, activity index calculation, or further analysis).

AnimalAI: An Open-Source Web Platform for Automated Animal Activity Index Calculation Using Interactive Deep Learning Segmentation

Current Working Directory: /media/mahtab/Windows/streamlitproject/segment-anything-2

Upload your video file



Drag and drop file here
Limit 5GB per file • MP4, MOV, AVI

Browse files

 human_chick_cut.mp4 34.9MB ×

Video Duration: 55.01 seconds

For videos less than 1 minute, we recommend an interval between 0.1-0.5 seconds.

Enter the frame interval in seconds:

0.50 - +

Extract Frames

Fig. 4.2. The graphical user interface of the application displaying the frame extraction process. Users can select a frame interval after uploading a video, adjust settings based on recommendations, and trigger the extraction process using the ‘Extract Frames’ button.

Interactive animal selection

Following frame extraction, the interface automatically displayed the first frame from the video so that the user can identify the animal(s) to be segmented from the background. Using a mouse click, users can select one or multiple animals (e.g., one, two, three, or potentially all visible objects) within the frame. Each click isolated the chosen subject by registering its coordinates, which guide subsequent segmentation tasks. If the user is dissatisfied with any selections, an

‘undo’ button enables a quick reversion, allowing for precise, iterative refinement of the selected coordinates.

This interactive step was crucial for achieving reliable isolation of the target animals from extraneous background elements. By removing other moving objects and environmental noises, the application is better able to deliver accurate analysis of movement or behavior in subsequent phases. Moreover, the flexibility to select multiple animals within a single frame offers a comprehensive approach for studies involving group dynamics or interactions.

Segmentation using Segment Anything Model 2

Once the targeted animals were selected, the segmentation process was initiated by clicking on the ‘Segment’ button (Figure 4.3). The foundation model known as SAM2 was employed to handle promptable visual segmentation in both images and videos. In SAM2, a data engine was built and refined through user interactions, culminating in the creation of largest video segmentation datasets to date. A simple transformer architecture with streaming memory was adopted to enable real-time video processing (Ravi et al., 2024).

By leveraging its extensive pretraining on a large and diverse dataset, SAM2 demonstrated strong performance across a wide range of segmentation tasks in both videos and images. In the context of video segmentation, higher accuracy has been observed with only one-third the user interactions required by previous approaches, and image segmentation ran 6× faster and more accurately compared to the original SAM. Notably, no additional training was required for specific tasks; instead, the user-selected coordinates served as prompts for guiding the segmentation, which was then automatically propagated to subsequent frames.

After the segmentation process was completed, two directories were created to store the results. One directory housed the RGB mask frames, in which the selected animals were distinctly

highlighted, while the other stored the binary mask frames, where only the targeted animals were shown in isolation. Figure 4.3 illustrates examples of both the RGB mask frames and the corresponding binary mask frames.

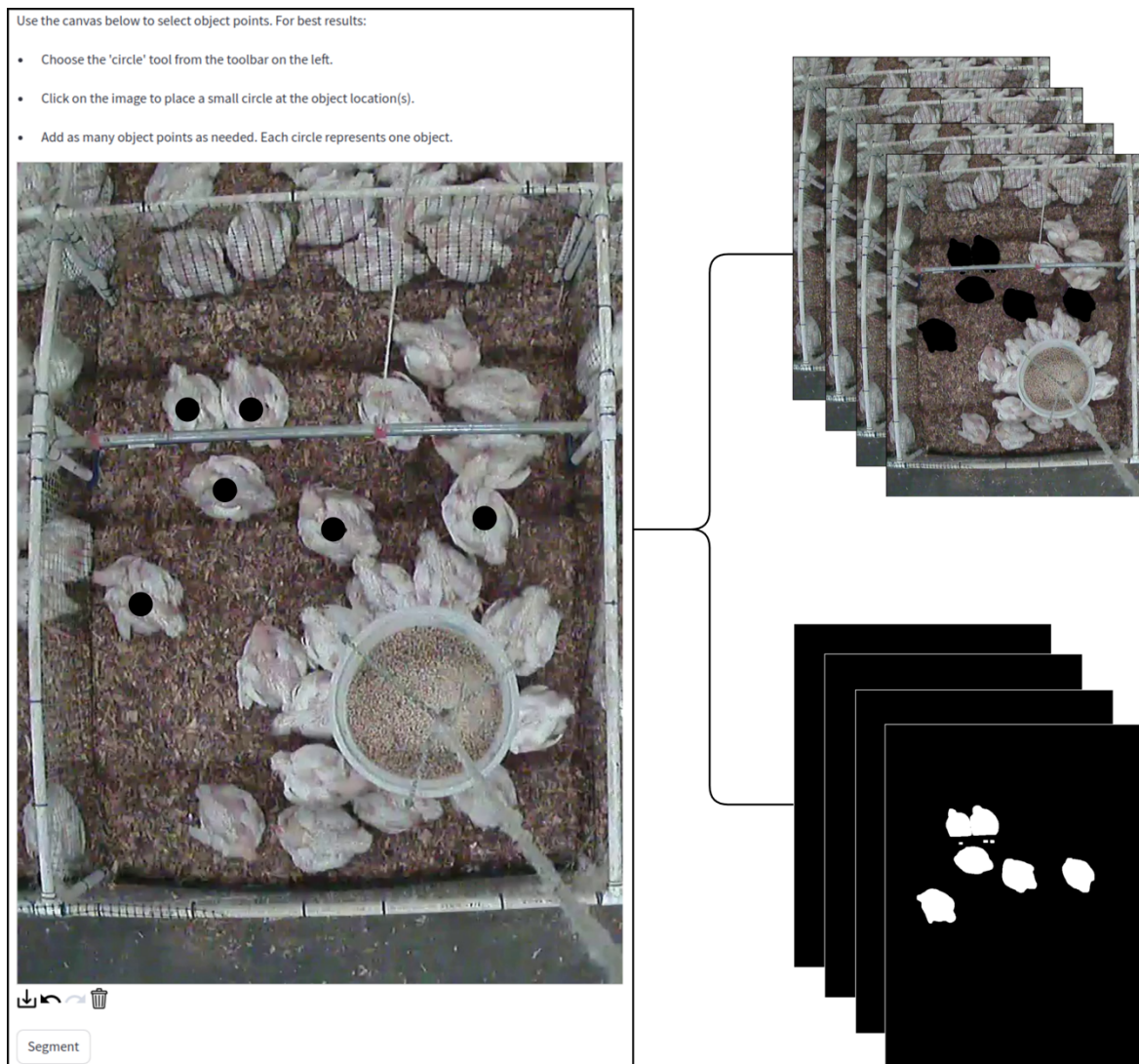


Fig. 4.3. Examples of segmentation outputs generated by the application using the Segment Anything Model 2. The RGB mask frames (top-right) highlight the selected animals in distinct colors, while the binary mask frames (bottom-right) isolate the targeted animals from the background.

Frame differencing for calculating activity index

Activity within a video sequence was assessed by measuring the extent of pixel-level changes between consecutive segmented frames (i.e., the binary mask frames). To achieve this, the difference between the current binary frame and the preceding frame was computed using an absolute difference operation. The resulting differenced frame highlights any pixels that have changed, indicating movement or behavioral changes. Figure 4.4 illustrates a series of these differenced frames.

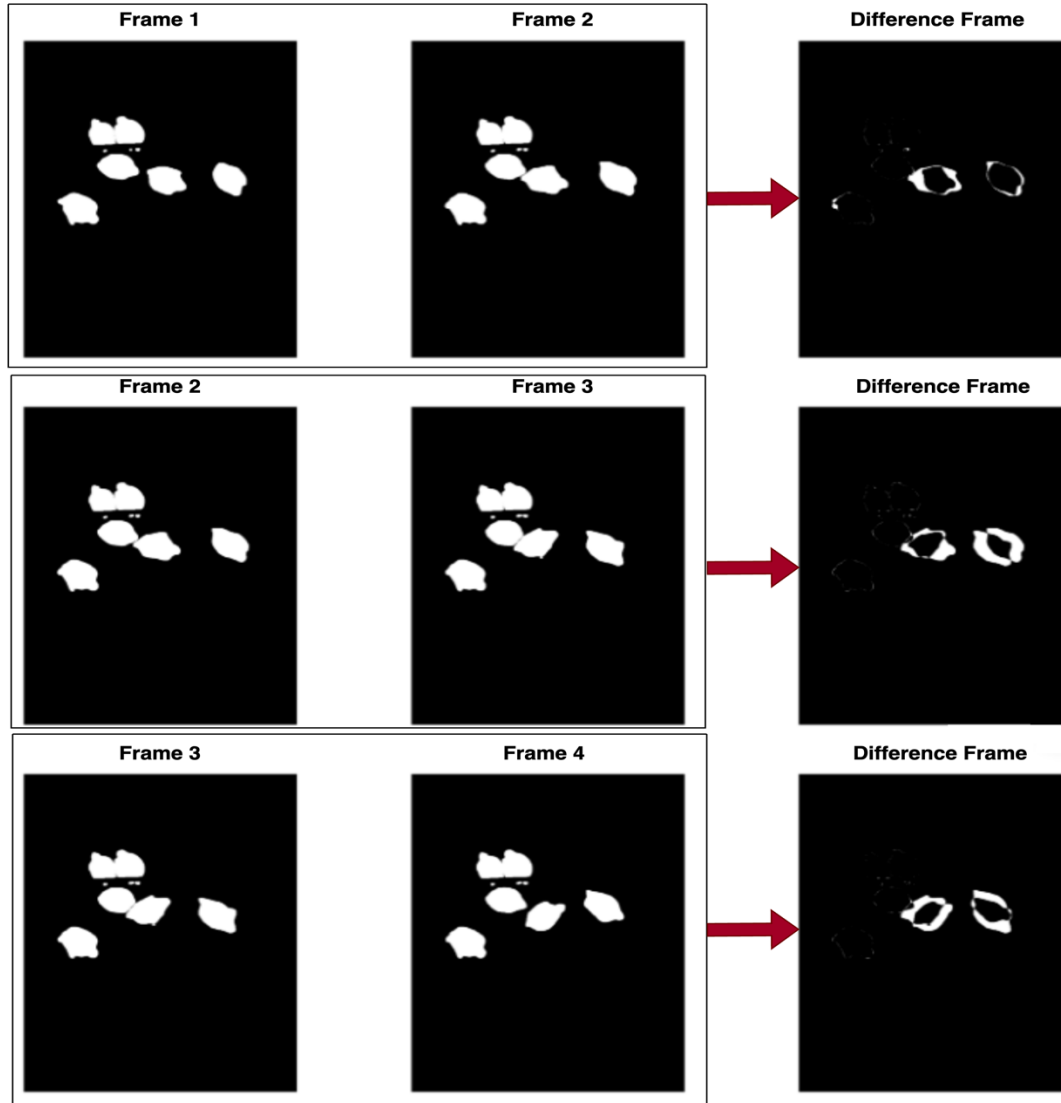


Fig. 4.4. Series of differenced frames illustrating pixel-level changes between consecutive segmented frames (binary mask frames) for calculating animal activity index.

To obtain an overall measure of activity, the number of changed pixels in each differenced frame was normalized by the combined pixel count of the current and previous frames. Formally, the activity index for frame i is calculated in Equation (4.1).

$$Activity\ index_i = \frac{Diff_pixel_count_i}{Pixel_count_current + Pixel_count_previous} \quad (4.1)$$

where the term *Diff_pixel_count i* is the total number of nonzero pixels in the “difference frame,” which is obtained by subtracting the pixel values of frame *i + 1* from frame *i*, Thus, these nonzero pixels highlight the regions that have changed between the two consecutive frames. Meanwhile, *pixel_count_current* and *pixel_count_previous* each represent the total number of nonzero pixels in frames *i + 1* and *i*, respectively. The two frames were used to generate the difference frame. This ratio ensured that the activity index remained bounded between 0 and 1. A higher value indicated greater movement, while a lower value suggested minimal changes.

Additionally, the normalized activity index for each consecutive frame was saved in a TXT file in the main directory. This is useful for users to further analyze the results on their own, enabling deeper insights into movement patterns and behavioral trends from an animal scientist’s perspectives.

Visualizing the activity index

Once the frame differencing procedure was completed, an activity index plot was automatically generated to illustrate the level of movement for the selected animals throughout the video. As shown in Figure 4.5, the x-axis represents the video time in minutes and seconds, while the y-axis ranges from 0 (indicating no movement) to 1 (reflecting the highest activity index). For each timestamp, a corresponding activity index value was displayed, enabling researchers to identify periods of heightened activity or relative inactivity. This visualization was invaluable for understanding the dynamics of animal behavior, as it condensed movement data into a single, intuitive plot for efficient analysis. Additionally, the generated activity index plot was saved in the main directory, allowing users to access and utilize it for further examination or reporting.

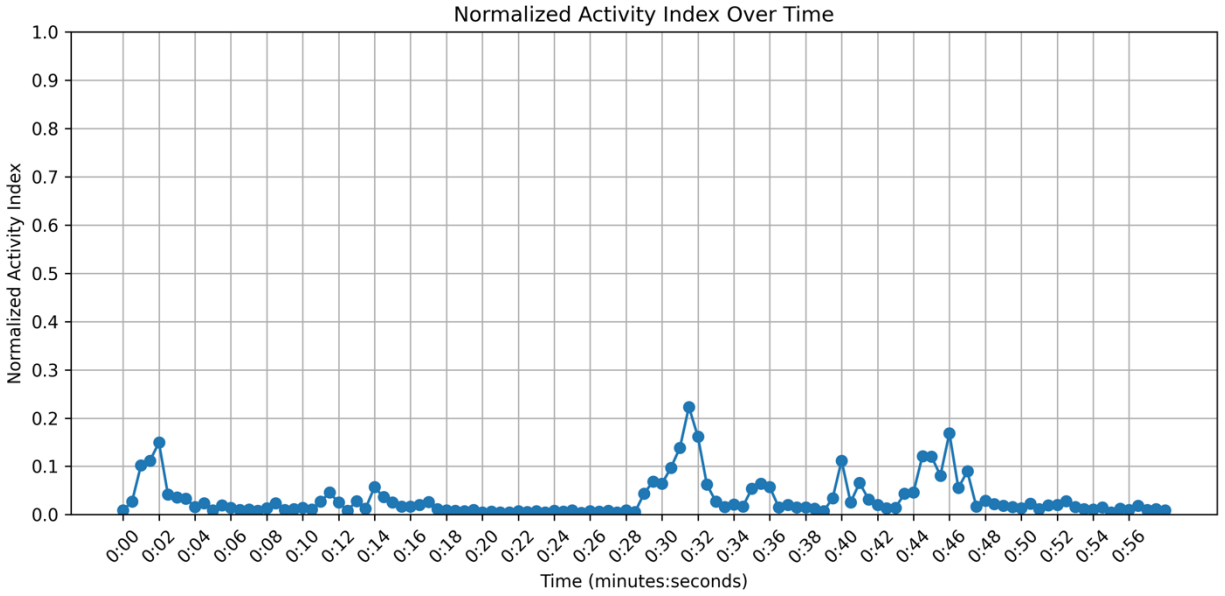


Fig. 4.5. Activity index plot illustrating the level of movement for the selected animals throughout the video. The x -axis denotes the video duration in minutes and seconds, while the y -axis ranges from 0 (minimal activity) to 1 (maximum activity).

Evaluation metrics calculation

This study employed a robust suite of evaluation metrics to independently gauge the performance of tracking and segmentation. The SAM2 evaluation leverages a dataset consisting of 1,157 individual chickens from 82 different video frames. The annotation of these images was carried out by a well-trained technician using Roboflow, which ensured the provision of high-precision masks that delineated the most complete depiction of each chicken in each frame. Subsequently, another well-trained technician conducted a double verification to guarantee the accuracy and quality of the labeling. This rigorous ground truth formed the benchmark for assessing the segmentation models' accuracy.

The SAM2 segmentation performance was evaluated with precision, recall, F1 score, and Intersection over Union (IoU) as described in Equations (4.2), (4.3), (4.4), and (4.5). The precision measures the accuracy of the segmentation model in identifying only relevant pixels as part of the segmentation. It is the ratio of correctly predicted positive observations to the total predicted positive observations. Recall, also known as sensitivity, measures the model's ability to correctly identify all relevant pixels. It is the ratio of correctly predicted positive observations to all observations that should have been labeled as positive. The F1 Score is the harmonic mean of precision and recall and a measure of the model's accuracy. An F1 Score reaches its best value at 1 (perfect precision and recall) and worst at 0. IoU is a measure used to quantify the percent overlap between the target mask and the model's prediction output. It is calculated by dividing the area of overlap between the predicted segmentation and the ground truth by the area of union.

$$Precision = \frac{True\ positive}{True\ positive + False\ positive} \quad (4.2)$$

$$Recall = \frac{True\ positive}{True\ positive + False\ negative} \quad (4.3)$$

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4.4)$$

$$IoU = \frac{True\ positive}{True\ positive + False\ positive + False\ negative} \quad (4.5)$$

where true positive refers to pixels that are correctly identified as part of birds; false positive are the pixels that the segmentation model incorrectly identifies as part of birds, but they actually belong to the background; false negative is used for pixels that are part of birds in the ground truth but are missed by the segmentation model.

A successful segmentation is one where the IoU is 50% or greater, which aligns with standard thresholds used in prominent publications (Girshick, 2015; He et al., 2017; Redmon and

Farhadi, 2018). The success rate thus reflects the percentage of images in which the models successfully tracked and segmented the chicken areas shown in Equation (4.6).

$$\text{Success rate} = \frac{\text{Number of successfully tracked \& segmented images (IoU > 0.5)}}{\text{Total number of images}} \quad (4.6)$$

Evaluating the impact of segmentation on activity index accuracy

To evaluate whether segmentation improved the accuracy of the activity index, 480 video frames were selected from week 4 recordings. These frames contained human interference and other unnecessary object movements (e.g., feeders and fans), providing a challenging scenario for activity-index calculation. Two methods were applied. First, the conventional “no-segmentation” approach involved subtracting consecutive frames to generate a difference frame, followed by applying a threshold value of 50 to binarize the result. White pixels in the binary image indicated movement, and black pixels indicated no movement. The activity index for each frame was then calculated based on the count of white pixels in that frame.

Second, in the segmentation-based method, all chickens in each frame were isolated using SAM2 before frame differencing. This removed non-essential background elements, including any human interference. The white-pixel counts were again used to compute the activity index. To determine whether these two approaches (with and without segmentation) produced significantly different mean activity levels, a paired t-test was conducted, with statistical significance set at $P < 0.05$. This comparison enabled a clearer assessment of how removing background motion influences the reliability of activity-index measurements.

Statistical analysis of different ratios of birds to represent the entire group’s activity

Tracking every individual bird can be time-consuming and computationally expensive. Consequently, this study tested whether sampling a subset of birds could reliably represent the entire flock’s movement patterns at different growth stages. Four different sampling ratios—20%,

40%, 60%, and 80% of the flock—were compared to the 100% baseline at three ages (weeks 1, 4, and 7). The number of birds selected from a pen was 7 for 20%, 15 for 40%, 22 for 60%, 30 for 80%, and 37 for 100%. Six distinct initializations (i.e., sets of randomly selected birds in feeder, drinker, corner, and open regions of the pen) were used per ratio to reduce spatial bias. All video data for this analysis were obtained as described earlier. Briefly, from each selected video clip, 480 consecutive frames (15 frames per second over ~32 seconds) were extracted. Within these frames, the developed platform isolated only the chosen subset of birds for each ratio, and an activity index was calculated by comparing pixel-wise differences between consecutive segmented frames. Parallel calculations were made for the 100% baseline (i.e., the entire flock).

The Pearson correlation coefficient (r value), as shown in Equation (4.7), was computed between each subset's activity index (at varying sampling percentages) and the full flock's index across six different initializations. The analyses were performed in Python (v3.9) using the pandas, numpy, and statsmodels libraries. This approach enabled a straightforward evaluation of whether a reduced sampling ratio could reliably represent overall flock activity while minimizing computational overhead.

$$r = \frac{\sum (x_i - \bar{x}) (y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (4.7)$$

where x_i represents the i^{th} observation for variable X , y_i represents the i^{th} observation for variable Y , $\bar{x}$ is the mean of all X values, and $\bar{y}$ is the mean of all Y values. The numerator captures how X and Y co-vary (or change together), while the denominator normalizes these deviations, keeping r dimensionless and ranging from -1 to +1.

According to a study, the correlation was negligible with r being 0.00 to 0.30 or 0.00 to -0.30, low with r being 0.31 to 0.50 or -0.31 to -0.50, moderate with r being 0.51 to 0.70 or -0.51

to -0.70, high with r being 0.71 to 0.90 or -0.71 to -0.90, and very high with r being 0.91 to 1.00 or -0.91 to -1.00 (Hinkle et al., 2003). Additionally, following the computation of the r value between each representation and the entire flock, a statistical comparison of the activity indices was carried out across different pairs of representations to determine whether they differ significantly. A significance level of $P < 0.05$ was applied, meaning that any P -value below 0.05 indicates a significant difference, while values above this threshold suggest no meaningful difference. If no statistically significant differences were observed, a smaller sampling ratio may be selected without sacrificing accuracy, thus reducing both computational load and resource requirements.

Results and discussion

Example procedure of interface operations

Below is a general procedure for video-based activity index calculation using the developed web-based platform. The platform guided users step-by-step through video segmentation and activity index generation. Some of the computational user interfaces are presented in Figures. 4.2-4.5.

Step 1: Run the platform and launch the interface using the command, which load the platform in a web browser. The main interface page then appears one a default web browser.

Step 2: Click the 'Browse files' button to upload a video file.

Step 3: The platform automatically checks whether the uploaded video is less than one hour in duration. If it is not, a warning message will be displayed, prompting the user to trim the video before proceeding.

Step 4: Input the frame interval for frame extraction based on either the recommended frame interval or the user's choice.

Step 5: Once the frame interval is set, the platform will extract individual frames from the video. The first frame will be displayed for visualization and selection of the animals to be tracked.

Step 6: Using the mouse, the user can click on the animal's location within the first frame to input its coordinates. This step initializes the segmentation process by identifying the region of interest.

Step 7: Confirm whether the selected coordinates are correct. If not, click the 'undo' button to adjust the coordinates and select a new region.

Step 8: When satisfied with the input, click the 'Segment' button. The platform will begin segmenting the video, generating both RGB mask frames and binary mask frames that highlight the animals of interest.

Step 9: The platform displays the activity index plot for the targeted, segmented animal across the video.

Step 10: Finally, the platform displays frames obtained through frame differencing, providing a dynamic view of motion changes throughout the video.

Segmentation performance on a chicken dataset

In this study, the effectiveness of SAM2 was evaluated within a web-based pipeline using a dedicated chicken dataset as mentioned earlier. The dataset comprised multiple video clips captured under diverse lighting conditions (5-10 lux), varying stocking densities (30-37 birds in a 1.2 m wide $\times$ 3.0 m long pen), and different chicken ages (weeks 1 to 7). This enabled to challenge the model's robustness under realistic, real-world scenarios. As shown in Figure 4.6, SAM2 produces high-quality segmentation results across chickens of different ages, demonstrating its adaptability to variations commonly encountered in poultry management settings.

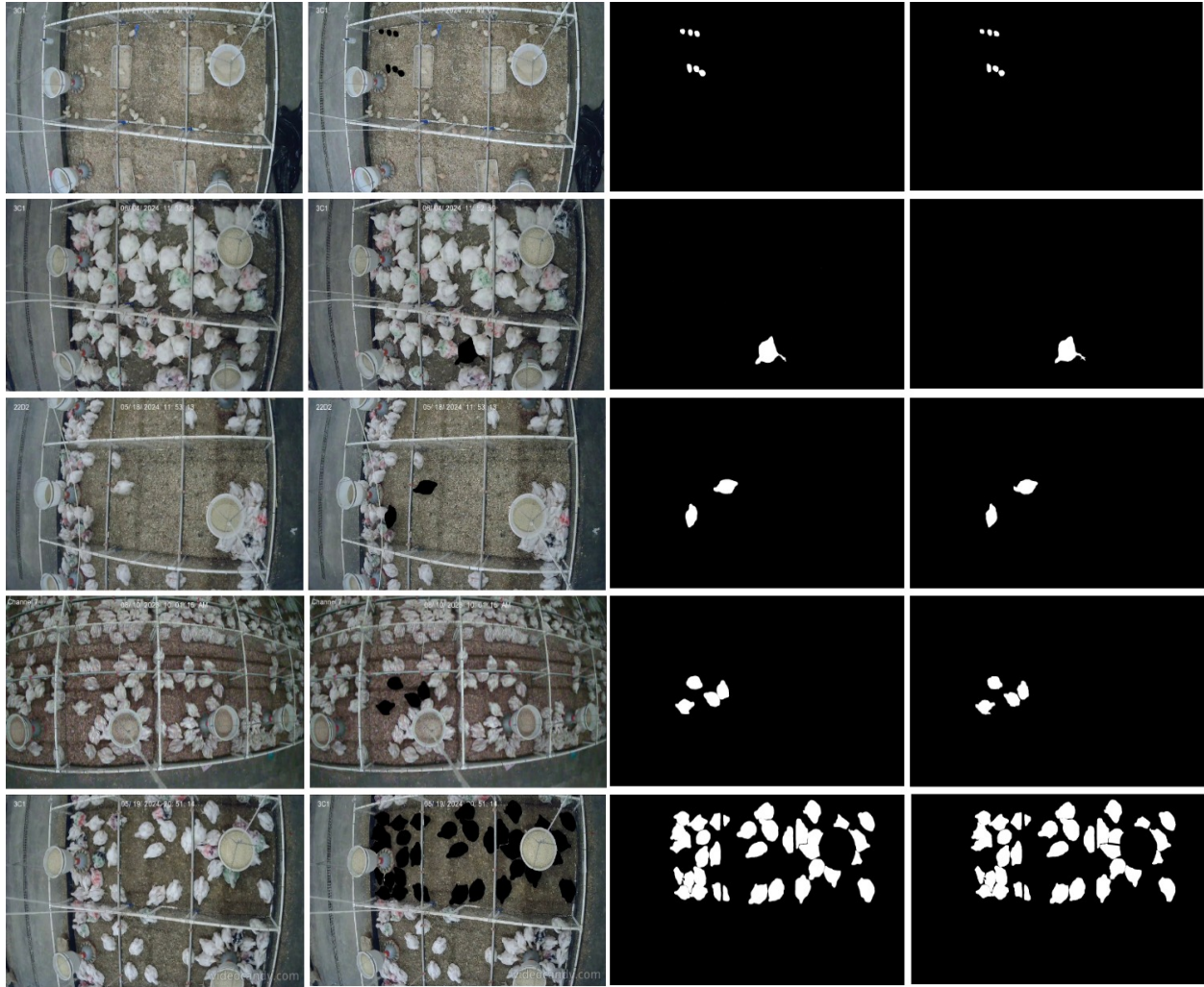


Fig. 4.6. SAM2 segmentation results: **(a)** the original frame, **(b)** the corresponding RGB mask output, **(c)** the Binary mask output, and **(d)** the ground truth segmentation.

SAM2 was selected for this project because it is specifically designed for interactive, prompt-based segmentation. In practice, a user can indicate the animal of interest by simply clicking or drawing a bounding box, after which SAM2 automatically tracked and segmented that animal throughout the video. This user-driven workflow was ideally suited for a web-based application where videos were uploaded, the target object(s) were selected, and precise mask outputs were generated without reliance on a fixed set of predefined object classes. By using

prompts, the model effectively mitigated challenges posed by occlusions and cluttered backgrounds, which are the issues frequently encountered in livestock environments. This approach was consistent with earlier studies that have shown minimal, yet precise user input can substantially improve segmentation accuracy (Kirillov et al., 2023; Sofiiuk et al., 2022).

To quantify the segmentation accuracy of SAM2 on chickens at various ages, several established performance metrics, including precision, recall, F1 score, IoU, and success rate, were employed. Table 4.1 summarizes these quantitative results for segmenting broiler chickens in weeks 1, 4, and 7. The consistently high scores (100% success rate, over 92% precision, over 97% recall, over 92%, and over 90% IoU) across different conditions indicate that SAM2 can generalize well, even when the visual appearance of the subjects changes due to factors such as age or lighting. Earlier segmentation methods relied on user-drawn bounding boxes or scribbles and can struggle with background clutter and occlusion (Rother et al., 2004). Unlike the earlier methods, the current method leveraged prompt-based guidance to focus precisely on regions of interest with the robust model architecture. Furthermore, SAM demonstrated high efficiency in practical deployment, requiring minimal user interaction while achieving accurate segmentation. Its refined prompt-based strategy effectively directed the model's attention to relevant regions, enabling precise segmentation without extensive manual annotation. This aligned with findings that SAM outperformed conventional models like SegFormer and SETR in zero-shot segmentation, achieving a mIoU of 94.8%, and operated effectively without additional training, reducing the burden of manual input while maintaining high performance (Yang et al., 2024).

Table 4.1. Segmentation performance of Segment Anything Model 2 for segmenting broiler chickens at Weeks 1-217.

Chicken age	Evaluation criteria (%)				
	Precision	Recall	F1 score	IoU	Success rate
Week1	92.13	98.40	95.16	90.77	100
Week4	94.42	98.29	96.29	92.85	100
Week7	94.75	97.86	92.26	92.79	100

The demonstrated performance has clear implications for real-world applications in precision poultry monitoring. For instance, integrating SAM2 into a web-based system would allow research scholars, regardless of coding or computing expertise, to upload videos, use simple prompts to segment individual chickens, and receive accurate segmentation masks in real time. Such a system would not only facilitate automated flock monitoring and behavioral analysis but could also be extended to support tasks such as weight prediction or movement tracking. Recent advancements in poultry monitoring have further illustrated how segmentation outputs can be utilized as critical inputs for data-driven livestock management. For instance, SAM-segmented results were combined with thermal images to extract various statistics of chickens' body temperature, facilitating more accurate assessments of their thermal conditions (Saeidifar et al., 2024).

Overall, the robust performance of SAM2 across diverse environmental and biological conditions, combined with its interactive and user-friendly design, confirmed its suitability for applications that require high-quality segmentation with minimal manual input. These results validate the technical capabilities of SAM2 while highlighting its potential to drive innovation in precision livestock farming and similar real-world domains.

Comparison of activity index calculation with and without segmentation

A total of 480 frames from week 4 recordings were analyzed to compare the results of activity index calculation with and without segmentation. As summarized in Figure 4.7, the segmented method produced lower and more consistent activity-index values (mean $\pm$ SD) relative to the unsegmented approach, indicating a reduction in background-induced noise.

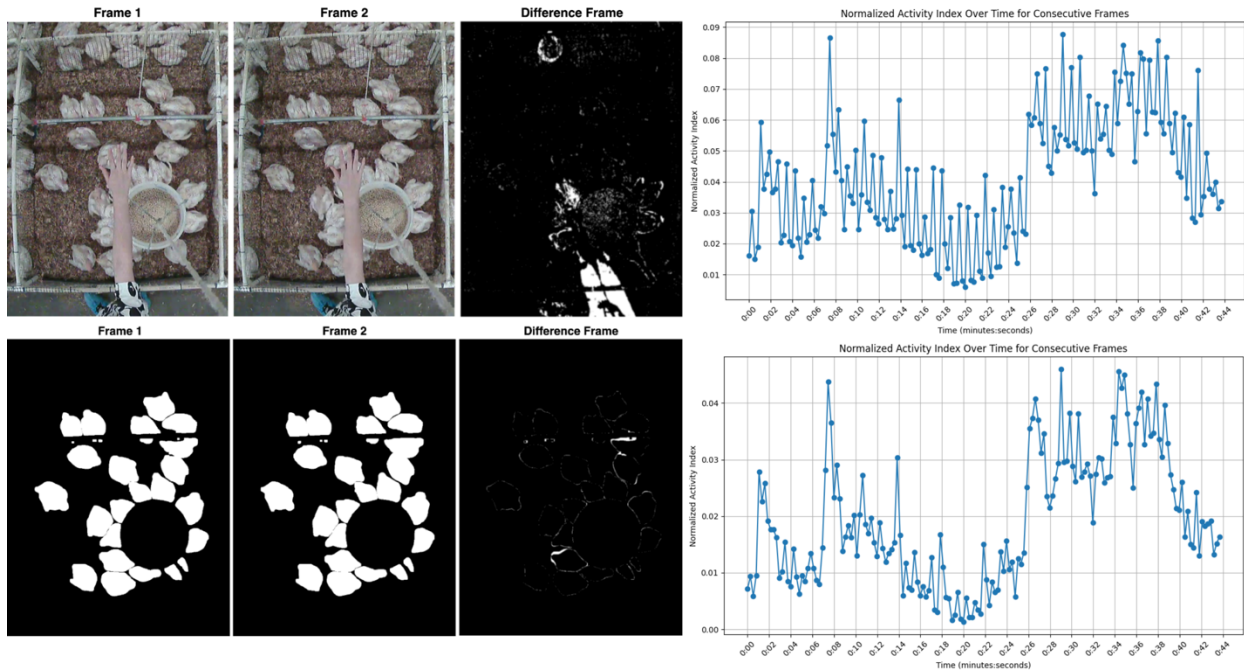


Fig. 4.7. Effect of segmentation on the accuracy of the activity index: the top row shows the approach without segmentation, while the bottom row shows the approach with segmentation.

A paired t-test revealed a significant difference ($P < 0.01$) between the two sets of activity-index measurements, demonstrating that removing non-essential background motion (e.g., human interference) meaningfully enhances the accuracy of the computed activity index. The activity index after segmentation was substantially reduced, with an average value of 3167.12 (mean absolute deviation of 2329.57), compared to 6302.64 (mean absolute deviation of 3744.55)

recorded before segmentation (i.e., prior to normalization). Frames with noticeable external movement had higher activity indices under the no-segmentation approach, whereas the segmentation-based method isolated chicken-related motion, minimizing overestimation and producing a smoother time series. These findings align with precision livestock monitoring studies, which have demonstrated that focusing on target subjects, such as chickens, reduces noise from extraneous interference by isolating them from distracting elements like moving litter, feathers, droppings, or human presence. This approach, particularly through image segmentation, improves data quality and tracking precision (Yang et al., 2024).

In practical applications, these results support the integration of segmentation as a pre-processing step in real-world poultry monitoring systems. By using segmented frames to calculate activity indexes, researchers can obtain more accurate, noise-free measurements that better reflect true animal activity. This refined approach can drive more effective, data-driven management decisions in precision livestock farming.

Determination of optimal sampling ratio for group activity assessment

Four different sampling ratios of the entire group—20%, 40%, 60%, and 80%—were evaluated at three key broiler growth stages (weeks 1, 4, and 7). Six distinct initializations were selected from various regions of the pen for each percentage to minimize bias. Figures. 4.8-4.10 show different initializations for weeks 1, 4, and 7 respectively.

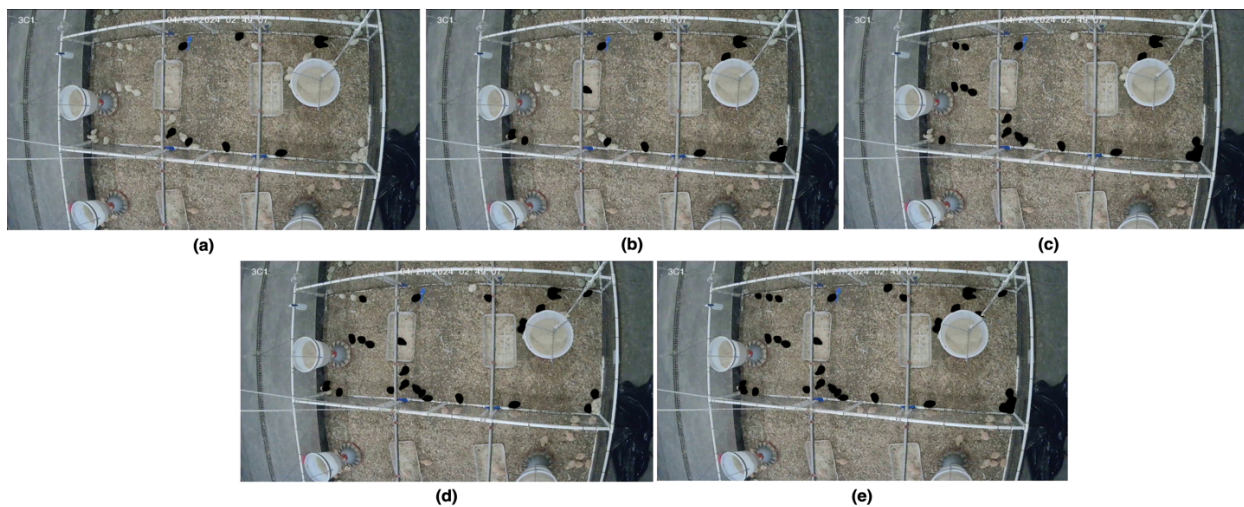


Fig. 4.8. Example of five sampling initializations at week 1, comparing (a) 20%, (b) 40%, (c) 60%, (d) 80%, and (e) the entire flock (100%).

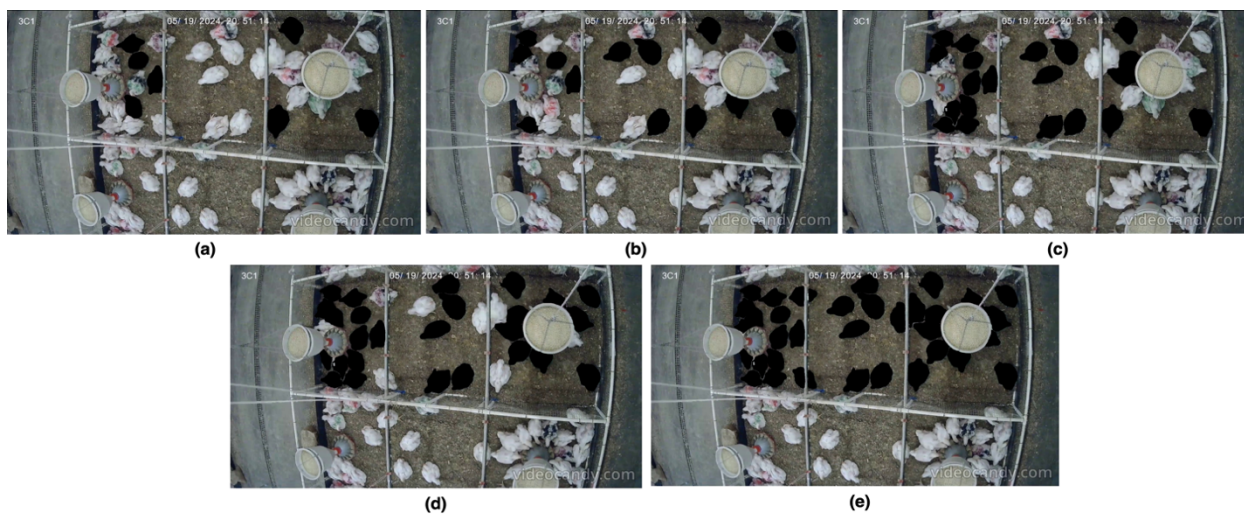


Fig. 4.9. Example of five sampling initializations at week 4, comparing (a) 20%, (b) 40%, (c) 60%, (d) 80%, and (e) the entire flock (100%).

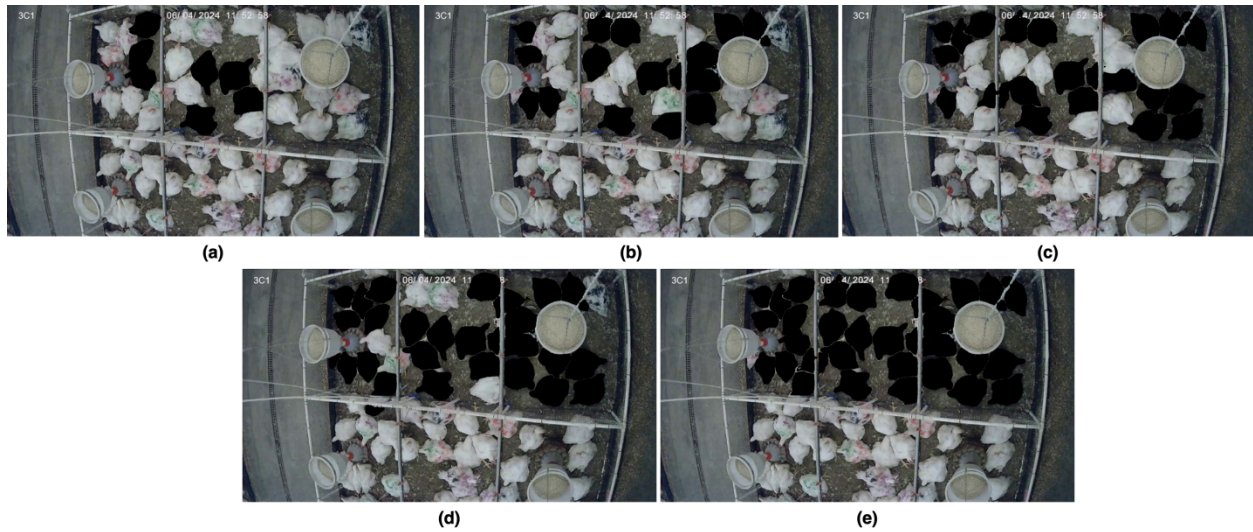


Fig. 4.10. Example of five sampling initializations at week 7, comparing (a) 20%, (b) 40%, (c) 60%, (d) 80%, and (e) the entire flock (100%).

Table 4.2 summarizes the average r value between each representation's activity index and the entire flock. For broilers at weeks 4 and 7, representations of 40% or more was highly correlated with the entire group ($r \geq 0.90$), whereas at week 1, a subset of at least 60% was required to highly correlate with the entire group ($r \geq 0.93$).

Table 4.2. Average Pearson correlation coefficients (r value) between each sampling ratio and the entire flock at different broiler growth stages (Weeks 1, 4, and 7).

Chicken age	Sampling ratio			
	20%	40%	60%	80%
Week1	0.58	0.61	0.93	0.97
Week4	0.74	0.90	0.96	0.98
Week7	0.73	0.93	0.92	0.94

To determine whether these representations also differ significantly from one another, P-values were computed and visualized in Figure 4.11 (heatmaps), with a significance level set at 0.05. Any pairwise comparison showing $P < 0.05$ was deemed significantly different, while P-values above 0.05 indicated no meaningful difference. At week 1 (Figure 4.11a), the 80% representation's activity index was significantly different from that of all other subsets. Coupled with its high r value of 0.97, this finding underscored the need to track 80% of the flock during the first week to ensure a reliable movement indicator. In week 4 (Figure 4.11b), the 60% and 80% representations showed no significant difference from each other ($P=0.092$) but differed significantly from both 20% and 40% ($P=0.006-0.023$). Given that 60% alone achieved a high r value of 0.96 and was not significantly different from the 80% subset, the 60% emerged as a more cost-effective option to represent the entire group. Lastly, in week 7 (Figure 4.11c), 40%, 60%, and 80% exhibited no significant differences among themselves ($P=0.486-0.791$), indicating that tracking 40% of the flock was sufficient, particularly given its high r value of 0.93 (Table 4.2).



Fig. 4.11. Comparative P -value heatmap across different representation at (a) week 1, (b) week 4, and (c) week 7.

The results demonstrate a clear trend that as broilers grew, the proportion of the flock required for accurate movement tracking decreased. This is likely due to the natural changes in flock behavior over time, where younger birds exhibited higher levels of individual movement variability, necessitating a larger sample size (Baxter and O’Connell, 2023; Newberry and Hall, 1990; Weeks et al., 2000). In contrast, older broilers exhibit more synchronized and predictable movement patterns, which allowed for a smaller subset of birds to sufficiently represent the entire flock (Bessei, 2006; van der Sluis et al., 2019).

From a practical standpoint, these results suggest that poultry management systems can significantly reduce tracking efforts by adjusting the sampling ratio based on bird age. Implementing an adaptive tracking strategy—where a higher sampling ratio is used early in growth and gradually reduced over time—could optimize the efficiency of activity monitoring systems. This approach can help farms allocate computational resources more effectively, enabling real-time flock assessments without unnecessary data processing costs.

Conclusions

A user-friendly, open-source platform was developed to address key challenges in animal behavior monitoring by enabling the calculation of the activity index for individual and group-housed animals from video recordings. The SAM2 was integrated with a frame-subtraction approach, ensuring reliable segmentation and tracking without requiring extensive training or annotations. This segmentation-based method significantly reduced noise and interference, thereby enhancing the accuracy of activity-index calculations. The results suggested that 80% in week 1, 60% in week 4, and 40% in week 7 were sufficient to cover the entire group’s activity index. The computational burden was lowered by tracking fewer animals as broilers matured, while still maintaining a robust representation of overall flock activity.

Beyond broiler applications, immediate deployment was facilitated for other species—such as pigs, cattle, or laboratory mice—without necessitating specialized technical expertise. Data processing, segmentation, and activity-index visualization were consolidated into a single Streamlit interface, providing researchers with an accessible and efficient tool for analyzing animal welfare and behavior patterns. Consequently, a critical gap in the availability of free, online solutions for animal welfare research was filled, paving the way for broader automated analysis and further advancements in computational tools for animal-welfare studies.

Declaration of Competing Interest

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.

Acknowledgement

This project was funded by US Poultry & Egg Association (project #: F-113).

Author contribution

Conceptualization: G.L.; Data curation: M.S., G.L, C.C., and E.A.; Formal analysis: M.S. and G.L.; Funding acquisition: C.C., G.L., and L.M.R.; Investigation: M.S. and G.L.; Methodology: M.S. and G.L.; Project administration: G.L.; Resources: G.L.; Software: M.S.; Supervision: G.L.; Validation: M.S.; Visualization: M.S.; Roles/Writing - original draft: M.S. and G.L.; Writing - review & editing: L.M.R., C.C., and E.A.

References

Attanasi, A., Cavagna, A., Castello, L.D., Giardina, I., Melillo, S., Parisi, L., Pohl, O., Rossaro, B., Shen, E., Silvestri, E., Viale, M., 2014. Collective Behaviour without Collective Order in Wild Swarms of Midges. *PLOS Computational Biology* 10, e1003697.
<https://doi.org/10.1371/journal.pcbi.1003697>

Aydin, A., Cangar, O., Ozcan, S.E., Bahr, C., Berckmans, D., 2010. Application of a fully automatic analysis tool to assess the activity of broiler chickens with different gait scores. *Computers and Electronics in Agriculture* 73, 194–199. <https://doi.org/10.1016/j.compag.2010.05.004>

Ballerini, M., Cabibbo, N., Candelier, R., Cavagna, A., Cisbani, E., Giardina, I., Lecomte, V., Orlandi, A., Parisi, G., Procaccini, A., Viale, M., Zdravkovic, V., 2008. Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. *Proceedings of the National Academy of Sciences* 105, 1232–1237. <https://doi.org/10.1073/pnas.0711437105>

Baxter, M., O’Connell, N.E., 2023. Large variation in the movement of individual broiler chickens tracked in a commercial house using ultra-wideband backpacks. *Sci Rep* 13, 7634. <https://doi.org/10.1038/s41598-023-34149-0>

Bessei, W., 2006. Welfare of broilers: a review. *World’s Poultry Science Journal* 62, 455–466. <https://doi.org/10.1017/S0043933906001085>

Bloemen, H., Aerts, J.-M., Berckmans, D., Goedseels, V., 1997. Image analysis to measure activity index of animals. *Equine Veterinary Journal* 29, 16–19. <https://doi.org/10.1111/j.2042-3306.1997.tb05044.x>

Bocaj, E., Uzunidis, D., Kasnesis, P., Patrikakis, C.Z., 2020. On the Benefits of Deep Convolutional Neural Networks on Animal Activity Recognition, in: 2020 International Conference on Smart Systems and Technologies (SST). Presented at the 2020 International Conference on Smart Systems and Technologies (SST), pp. 83–88. <https://doi.org/10.1109/SST49455.2020.9263702>

Chen, C., Zhu, W., Norton, T., 2021. Behaviour recognition of pigs and cattle: Journey from computer vision to deep learning. *Computers and Electronics in Agriculture* 187, 106255. <https://doi.org/10.1016/j.compag.2021.106255>

Dell, A.I., Bender, J.A., Branson, K., Couzin, I.D., de Polavieja, G.G., Noldus, L.P.J.J., Pérez-Escudero, A., Perona, P., Straw, A.D., Wikelski, M., Brose, U., 2014. Automated image-based tracking and its application in ecology. *Trends in Ecology & Evolution* 29, 417–428. <https://doi.org/10.1016/j.tree.2014.05.004>

Elbarrany, A.M., Mohialdin, A., Atia, A., 2023. Abnormal Behavior Analysis for Surveillance in Poultry Farms using Deep Learning, in: 2023 Intelligent Methods, Systems, and Applications (IMSA). Presented at the 2023 Intelligent Methods, Systems, and Applications (IMSA), pp. 56–61. <https://doi.org/10.1109/IMSA58542.2023.10217676>

Fuentes, A., Yoon, S., Park, J., Park, D.S., 2020. Deep learning-based hierarchical cattle behavior recognition with spatio-temporal information. *Computers and Electronics in Agriculture* 177, 105627. <https://doi.org/10.1016/j.compag.2020.105627>

Girshick, R., 2015. Fast R-CNN. <https://doi.org/10.48550/arXiv.1504.08083>

He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask R-CNN, in: 2017 IEEE International Conference on Computer Vision (ICCV). Presented at the 2017 IEEE International Conference on Computer Vision (ICCV), pp. 2980–2988. <https://doi.org/10.1109/ICCV.2017.322>

Hinkle, D.E., Wiersma, W., Jurs, S.G., 2003. *Applied Statistics for the Behavioral Sciences*.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.-Y., Dollár, P., Girshick, R., 2023. Segment Anything. <https://doi.org/10.48550/arXiv.2304.02643>

- Kristensen, H.H., Aerts, J.M., Leroy, T., Wathes, C.M., Berckmans, D., 2006. Modelling the dynamic activity of broiler chickens in response to step-wise changes in light intensity. *Applied Animal Behaviour Science* 101, 125–143. <https://doi.org/10.1016/j.applanim.2006.01.007>
- Li, G., Chai, L., 2023. AnimalAccML: An open-source graphical user interface for automated behavior analytics of individual animals using triaxial accelerometers and machine learning. *Computers and Electronics in Agriculture* 209, 107835. <https://doi.org/10.1016/j.compag.2023.107835>
- Li, G., Li, B., Shi, Z., Zhao, Y., Tong, Q., Liu, Y., 2020. Diurnal rhythms of group-housed layer pullets with free choices between light and dim environments. *Can. J. Anim. Sci.* 100, 37–46. <https://doi.org/10.1139/cjas-2019-0009>
- Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., Bethge, M., 2018. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci* 21, 1281–1289. <https://doi.org/10.1038/s41593-018-0209-y>
- Neves, D.P., Mehdizadeh, S.A., Tschärke, M., Nääs, I. de A., Banhazi, T.M., 2015. Detection of flock movement and behaviour of broiler chickens at different feeders using image analysis. *Information Processing in Agriculture* 2, 177–182. <https://doi.org/10.1016/j.inpa.2015.08.002>
- Newberry, R.C., Hall, J.W., 1990. Use of pen space by broiler chickens: Effects of age and pen size. *Applied Animal Behaviour Science* 25, 125–136. [https://doi.org/10.1016/0168-1591\(90\)90075-O](https://doi.org/10.1016/0168-1591(90)90075-O)
- Oso, O.M., Mejia-Abaunza, N., Bodempudi, V.U.C., Chen, X., Chen, C., Aggrey, S.E., Li, G., 2025. Automatic Analysis of High, Medium, and Low Activities of Broilers with Heat Stress Operations via Image Processing and Machine Learning. *Poultry Science* 104954. <https://doi.org/10.1016/j.psj.2025.104954>

Peña Fernández, A., Norton, T., Tullo, E., van Hertem, T., Youssef, A., Exadaktylos, V., Vranken, E., Guarino, M., Berckmans, D., 2018. Real-time monitoring of broiler flock's welfare status using camera-based technology. *Biosystems Engineering, Advances in the Engineering of Sensor-based Monitoring and Management Systems for Precision Livestock Farming* 173, 103–114. <https://doi.org/10.1016/j.biosystemseng.2018.05.008>

Ravi, N., Gabeur, V., Hu, Y.-T., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., Mintun, E., Pan, J., Alwala, K.V., Carion, N., Wu, C.-Y., Girshick, R., Dollár, P., Feichtenhofer, C., 2024. SAM 2: Segment Anything in Images and Videos. <https://doi.org/10.48550/arXiv.2408.00714>

Redmon, J., Farhadi, A., 2018. YOLOv3: An Incremental Improvement. <https://doi.org/10.48550/arXiv.1804.02767>

Ren, W., Tang, Y., Sun, Q., Zhao, C., Han, Q.-L., 2024. Visual Semantic Segmentation Based on Few/Zero-Shot Learning: An Overview. *IEEE/CAA Journal of Automatica Sinica* 11, 1106–1126. <https://doi.org/10.1109/JAS.2023.123207>

Romero-Ferrero, F., Bergomi, M.G., Hinz, R.C., Heras, F.J.H., de Polavieja, G.G., 2019. idtracker.ai: tracking all individuals in small or large collectives of unmarked animals. *Nat Methods* 16, 179–182. <https://doi.org/10.1038/s41592-018-0295-5>

Rother, C., Kolmogorov, V., Blake, A., 2004. “GrabCut”: interactive foreground extraction using iterated graph cuts. *ACM Trans. Graph.* 23, 309–314. <https://doi.org/10.1145/1015706.1015720>

Saeidifar, M., Li, G., Chai, L., Bist, R., Rasheed, K.M., Lu, J., Banakar, A., Liu, T., Yang, X., 2024. Zero-shot image segmentation for monitoring thermal conditions of individual cage-free laying hens. *Computers and Electronics in Agriculture* 226, 109436. <https://doi.org/10.1016/j.compag.2024.109436>

Sengar, S.S., Mukhopadhyay, S., 2017. Moving object detection based on frame difference and W4. *SIViP* 11, 1357–1364. <https://doi.org/10.1007/s11760-017-1093-8>

Shams, A., Becker, D., Becker, K., Amirian, S., Rasheed, K., 2023. Evolving Efficient CNN Based Model for Image Classification, in: 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE). Presented at the 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE), pp. 228–235. <https://doi.org/10.1109/CSCE60160.2023.00041>

Silvera, A.M., Knowles, T.G., Butterworth, A., Berckmans, D., Vranken, E., Blokhuis, H.J., 2017. Lameness assessment with automatic monitoring of activity in commercial broiler flocks. *Poultry Science* 96, 2013–2017. <https://doi.org/10.3382/ps/pex023>

Sofiuk, K., Petrov, I.A., Konushin, A., 2022. Reviving Iterative Training with Mask Guidance for Interactive Segmentation, in: 2022 IEEE International Conference on Image Processing (ICIP). Presented at the 2022 IEEE International Conference on Image Processing (ICIP), pp. 3141–3145. <https://doi.org/10.1109/ICIP46576.2022.9897365>

Tran, D.-N., Nguyen, T.N., Khanh, P.C.P., Tran, D.-T., 2022. An IoT-Based Design Using Accelerometers in Animal Behavior Recognition Systems. *IEEE Sensors Journal* 22, 17515–17528. <https://doi.org/10.1109/JSEN.2021.3051194>

van der Sluis, M., de Klerk, B., Ellen, E.D., de Haas, Y., Hijink, T., Rodenburg, T.B., 2019. Validation of an Ultra-Wideband Tracking System for Recording Individual Levels of Activity in Broilers. *Animals (Basel)* 9, 580. <https://doi.org/10.3390/ani9080580>

Weeks, C.A., Danbury, T.D., Davies, H.C., Hunt, P., Kestin, S.C., 2000. The behaviour of broiler chickens and its modification by lameness. *Appl Anim Behav Sci* 67, 111–125. [https://doi.org/10.1016/s0168-1591\(99\)00102-1](https://doi.org/10.1016/s0168-1591(99)00102-1)

Yan, C., Chang, X., Luo, M., Liu, H., Zhang, X., Zheng, Q., 2024. Semantics-Guided Contrastive Network for Zero-Shot Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 46, 1530–1544. <https://doi.org/10.1109/TPAMI.2021.3140070>

Yang, X., Dai, H., Wu, Z., Bahadur Bist, R., Subedi, S., Sun, J., Lu, G., Li, C., Liu, T., Chai, L., 2024. An innovative segment anything model for precision poultry monitoring. *Computers and Electronics in Agriculture* 222, 109045. <https://doi.org/10.1016/j.compag.2024.109045>

CHAPTER V

SUMMARY

This dissertation systematically investigated advanced computer vision and deep learning techniques to enhance precision poultry farming by focusing on the automated health analysis of laying hens and the activity monitoring of different species. The methodologies, experimental results, and detailed discussions are presented in Chapters II to IV, corresponding to three distinct research papers. This final chapter summarizes the key findings, their interconnections, and their implications for poultry welfare and farm management, highlighting how these advancements contribute to more efficient and humane poultry production.

The first study developed an optimized zero-shot image segmentation pipeline based on the Segment Anything Model (SAM) to automatically segment individual cage-free laying hens in thermal images. By integrating pre-processing (e.g., thresholding for automatic point selection) and post-processing with a machine learning classifier, the modified SAM outperformed other models, including YOLOv8, Mask R-CNN, FastSAM, MobileSAM, U2-Net, and ISNet, achieving a success rate of 84.4%, Intersection over Union (IoU) of 85.5%, recall of 91.0%, and an F1 score of 92.3%. This pipeline enabled the extraction of comprehensive body surface temperature statistics (e.g., mean: 26.68–28.53°C, median: 26.27–28.28°C across weeks 77–80) for individual hens, offering a non-invasive tool for monitoring thermal conditions. The approach reduced animal stress by eliminating manual handling and provided a scalable solution for precision poultry farming, enhancing health monitoring and production efficiency.

The second study advanced automated segmentation by combining YOLOv7 for object detection with SAM for segmentation, using bounding box prompts to eliminate manual inputs. This hybrid YOLOv7 + SAM model achieved superior performance compared to YOLOv8, Mask R-CNN, and other SAM variants, with a precision of 92.5%, recall of 98.2%, an F1 score of 95.1%, IoU of 91.0%, and a success rate of 98.0%. By automating the detection and segmentation of individual hens, this method minimized the need for labor-intensive annotations, offering a scalable and efficient solution for poultry monitoring. The high accuracy and transferability of this approach make it applicable not only to poultry but also to broader agricultural, environmental, and medical imaging tasks, demonstrating its versatility in precision farming.

The third study addressed the challenge of monitoring animal activity by developing an open-source, user-friendly Streamlit platform integrated with SAM2 for segmenting and tracking individual broiler chickens in videos. This platform overcame the limitations of traditional pixel intensity differencing methods by reducing noise and enabling individual tracking, with a segmentation success rate of 100%, IoU of 92.21%, precision of 93.87%, recall of 98.15%, and an F1 score of 95.94% for Cobb500 male broilers from weeks 1 to 7. Statistical analysis showed that tracking 80% of birds in week 1, 60% in week 4, and 40% in week 7 was sufficient ($r \geq 0.90$; $P \leq 0.048$) to represent the flock's activity index, reducing computational costs while maintaining accuracy. This tool is accessible to non-technical researchers and adaptable for other species, filling a critical gap in automated welfare assessment tools.

The findings demonstrate that advanced computer vision tools can significantly improve poultry management by providing non-invasive, automated, and accurate monitoring systems. The thermal segmentation pipeline enables early detection of health issues through temperature variations, reducing the need for stressful handling and improving welfare. The hybrid YOLOv7

+ SAM model offers a scalable solution for identifying and monitoring individual birds, which can optimize resource allocation (e.g., feeder and drinker placement) and detect behavioral anomalies. The Streamlit platform empowers researchers and farmers with an accessible tool to assess flock activity, enabling data-driven decisions to enhance welfare, such as adjusting stocking densities or lighting conditions to reduce stress. Collectively, these tools support precision farming practices that improve productivity, reduce labor costs, and enhance animal welfare by enabling proactive management of health and behavior.

The developed tools lay a strong foundation for future advancements in precision poultry farming. The thermal segmentation pipeline could be extended to monitor other physiological parameters, such as respiratory rate, by integrating additional sensors. The hybrid segmentation model's transferability suggests potential applications in other livestock species or agricultural tasks, such as crop monitoring. The Streamlit platform could be enhanced with real-time analytics and cloud integration for large-scale deployments. Farm managers can leverage these tools to optimize management practices, such as adjusting environmental conditions based on activity and temperature data to improve welfare and production efficiency. Future research should focus on integrating these systems into a unified platform for real-time, multi-modal monitoring and validating their performance across diverse poultry breeds and farming systems.

In conclusion, this dissertation provides a suite of innovative, automated tools that address critical challenges in poultry monitoring, from thermal health assessment to behavioral tracking. By harnessing zero-shot segmentation and user-friendly interfaces, these advancements pave the way for smarter, more humane poultry farming practices, with significant potential for broader agricultural applications.