國立臺灣大學生物資源暨農學院生物機電工程學系

碩士論文

Department of Biomechatronics Engineering
College of Bioresources and Agriculture
National Taiwan University

Master's Thesis

利用深度學習量化白蝦進食相關特性

Quantifying Feeding-related Characteristics of Shrimp Using Deep Learning

李居展

Chu-Chan Lee

指導教授：郭彥甫 博士

Advisor: Yan-Fu Kuo, Ph.D.

中華民國 113 年 6 月

June 2024

doi:10.6342/NTU202403695

ACKNOWLEDGEMENTS

This master's thesis owes its completion to the invaluable support of many individuals. First and foremost, I am especially grateful to my advisor, Prof. Yan-Fu Kuo, whose guidance was instrumental from the onset of my master's studies. His willingness to welcome me into his laboratory marked the beginning of this enriching journey. Over the past two years, he has provided me with unwavering support both in research and personally, and I am truly grateful to have such an exceptional mentor. Secondly, I want to thank Yun and Johnny, my classmates and companions during this journey. We have shared both the highs and lows of research life, and I will always cherish our friendship. I am also profoundly thankful to the graduated seniors and members of the Lab203, namely Barry, Jack, Dino, Kent, Yishin, Kuan-Ting, Kim, Kuan-Wei, AndyCW, Karl, and Sophia. Their quiet support made my research experience less lonely. Additionally, I extend my gratitude to my collaborators, Caregoods shrimp farm and Prof. Yuan-Nan Chu, for providing experimental sites and technical guidance. Last but certainly not least, I would like to sincerely thank my beloved family. Their nurturing from my childhood, patience with my reduced time at home, and their invaluable advice and encouragement during my low points have given me the courage and confidence to face all challenges. I will forever hold their support close to my heart.

doi:10.6342/NTU202403695

中文摘要

蝦是全球主要的蛋白質來源。在蝦類養殖中，飼料成本大約占總支出的 40%。有效的進食管理對於優化蝦的成長和最小化成本至關重要。蝦的食慾受到生長階段和環境條件的影響。此外，由於蝦類是底棲生物，使得直接觀察變得充滿挑戰。傳統上，蝦的食慾是通過將樣本飼料放置在傘網上進行肉眼觀察來確定的，但此方法耗時且主觀。為了解決這些問題，本研究旨在通過使用深度神經網絡觀察蝦類的餵食相關行為來自動化蝦的食慾判定。

在提議的方法中，構建了配備飼料投料器的水下攝影模組，以在樣本餵食過程中捕捉蝦隻的影片。通過影像處理算法對影片中的飼料殘留區域進行量化，並計算出反映蝦食慾的飼料殘留區域變化指數（FRAVI）。影片中的蝦隻被 YOLOv9-c 模型和追蹤演算法追蹤。接著，利用飼料殘留檢測模組進行測量，衍生出關鍵的進食相關特性，包括蝦類數量、移動、進入頻率和停留時間。飼料殘留檢測模組達到了 0.885 的整體相關性，而 YOLOv9-c 模型達到了 0.88 的平均精度。此外，還監測了水溫、鹽度和溶解氧等環境因素，分析它們與進食相關特性的相關性。分析表明，水溫與蝦類活動水平正相關，較高的站壓與蝦類進入頻率正相關，顯示這些因素在飼料攝取效率中扮演著重要角色。本研究提供了關於蝦隻進食相關行為的持續、客觀和精確的信息，這些資訊可能有助於農民優化飼料管理和水產養殖實踐。

關鍵字: 深度學習、機器視覺、蝦類養殖、蝦類進食行為

doi:10.6342/NTU202403695

ABSTRACT

Shrimp serves as a significant protein source globally. In shrimp farming, feed accounts for approximately 40% of the overall expenses. Effective feeding management is crucial for optimizing shrimp growth and minimizing the costs. Shrimp appetite is influenced by growth stages and ambient conditions. In addition, shrimps are benthos, making direct observation challenging. Conventionally, shrimp appetite was determined using naked-eye observation by putting sample feed on trays. The approach is, however, time-consuming and subjective. To address these issues, this study aimed to automate shrimp appetite by observing their feeding-related behaviors using deep neural networks. In the proposed approach, underwater video modules with feed dispensers were built to capture videos of shrimps during sample feeding (i.e., a small amount of feed). Feed residue areas in the videos were quantified using image processing algorithms. Feed residue area variation index (FRAVI) that indicates shrimp appetite was quantified. Shrimps in the videos were detected and tracked using YOLOv9-c and simple online realtime tracking algorithm. Feed residue were measured using feed residue detection module. Key feeding-related characteristics, including shrimp count, movement, entry frequency, dwelling time, were next derived. Feed residue detection module achieved an overall correlation of 0.885. YOLOv9-c model achieved a mean average precision of 0.88. Additionally, environmental factors like water temperature, salinity, and dissolved oxygen levels were monitored to analyze correlations with feeding-related behavior. Analysis indicated that water temperature is positively correlated with shrimp activity levels, and higher station pressure is positively correlated with shrimp entry frequency, suggesting these factors play a significant role in feed intake efficiency. The proposed approach provides continuous, objective, and precise information of the feeding-related

doi:10.6342/NTU202403695

behaviors of shrimps. The information may aid farmers in optimizing feed management and aquaculture practices.

Keywords: Deep learning, shrimp behavior, computer vision, shrimp farming

doi:10.6342/NTU202403695

TABLE OF CONTENTS

doi:10.6342/NTU202403695

doi:10.6342/NTU202403695

LIST OF FIGURES

doi:10.6342/NTU202403695

doi:10.6342/NTU202403695

doi:10.6342/NTU202403695

LIST OF TABLES

doi:10.6342/NTU202403695

CHAPTER 1. INTRODUCTION

1.1 Background

Shrimp aquaculture plays a vital role in meeting the growing global demand for protein sources (Tacon and Metian, 2013). According to the Food and Agriculture Organization (FAO, 2022), global shrimp production exceeded 6 million tonnes, valued at over $35 billion, in 2020. In Taiwan, the thriving shrimp farming industry yielded an annual production of around 100,000 tonnes, worth approximately NT$10 billion (US$330 million), as reported by the Fisheries Agency of Taiwan in 2022 (Fisheries Agency, 2023). To meet the rising demand of shrimp reducing the cost of raising shrimp is critical for shrimp farmers. Feed costs account for approximately 40% of the overall operational expenses in shrimp farming (Cuzon et al., 2004; Toan, 2016), making efficient feed management crucial for balancing farming expenditures while maintaining optimal shrimp growth rates, thereby ensuring economic viability and sustainability (Nunes et al., 2014).

Conventionally, shrimp farmers have relied on manual observation techniques to assess the feeding behavior and appetite of shrimp (Golez et al., 2011). This conventional method involves introducing a small amount of "trial feed" into the pond with a feeding tray and visually evaluating characteristics such as the number of attracted shrimp, their movement patterns, dwelling time near the feed, and the rate of feed consumption. However, this approach is subjective, time-consuming and prone to inconsistencies across different observers (Robertson et al., 1993), ultimately limiting the accuracy and reliability of feed management decisions. Consequently, developing automated and objective methods to quantify feeding-related characteristics of shrimp is critical/necessary for optimizing feed management strategies, enhancing resource utilization efficiency, and improving overall production (Mustafa et al., 2021).

doi:10.6342/NTU202403695

1.2 Objectives

This study aimed to develop an automated system that integrates deep learning and computer vision techniques to quantify feeding-related characteristics of shrimp. The system is designed to accurately detect and track individual shrimp while assessing the variation in residual feed area. By analyzing shrimp movement patterns, dwelling times, entry/exit frequencies, and interactions with the feed, the study objectively quantifies the feeding behavior of shrimp. This automated and objective approach seeks to assist shrimp farmers in optimizing feeding strategies, minimizing resource wastage, and ultimately enhancing overall production efficiency, sustainability, and profitability in the shrimp aquaculture industry.

1.3 Organization

The structure of this document is laid out as follows. Chapter 2 provides a comprehensive literature review, setting the theoretical groundwork for the study. Chapter 3 details the methodology, beginning with the collection of videos, followed by descriptions of the shrimp tracking and feed residue measurement algorithms. It concludes with the presentation of feeding-related quantifications. Chapter 4 discusses the results derived from this research, and Chapter 5 offers the study's conclusions and implications.

doi:10.6342/NTU202403695

CHAPTER 2. LITERATURE REVIEW

2.1 Image processing techniques for quantifying shrimp feeding behavior

To overcome the limitations of manual monitoring, researchers have explored image processing techniques to quantify shrimp feeding behavior and residual feed patterns. Wei et al. (2021) employed a U-Net architecture to segment and measure the variation in residual feed area, which served as an indicator for estimating shrimp appetite. However, their approach lacked direct quantification of shrimp feeding behavior itself. Other studies have utilized traditional image processing techniques, such as thresholding, morphological operations, and contour analysis, to detect and measure feed pellets or residual feed areas (Boscolo et al., 2019; Jahangiri et al., 2021). While these methods provide insights into feed consumption patterns, they do not directly capture the behavioral characteristics of shrimp during feeding events.

2.2 Deep learning techniques in aquaculture monitoring

Deep learning methods, especially those utilizing convolutional neural networks (CNNs), have achieved impressive outcomes in object detection and tracking in multiple sectors, including aquaculture. Notably, deep learning models like You Only Look Once (YOLO) (Redmon et al., 2016) and Mask R-CNN (He et al., 2017) have been successfully applied to detect and track aquatic species such as shrimp (Salman et al., 2020; Zhang et al., 2021). These technologies provide precise identification and monitoring of individual shrimp, which is critical for accurately measuring behavioral characteristics such as movement trajectories, residency periods, and social interactions.

2.3 Integration of deep learning with other techniques for comprehensive analysis

Furthermore, some researchers have integrated deep learning models with other techniques, such as optical flow analysis and trajectory clustering, to comprehensively analyze shrimp behavior, including swimming directions, velocities, and activity levels (Xu et al., 2022; Li et

doi:10.6342/NTU202403695

al., 2023). These advancements in deep learning-based object detection and tracking have paved the way for a more accurate and comprehensive monitoring of shrimp, providing valuable insights to optimize feeding strategies and overall aquaculture practices.

doi:10.6342/NTU202403695

CHAPTER 3. MATERIALS AND METHODS

3.1 Overview of the system

The developed system comprised: (a) underwater video modules (UVMs), (b) automatic feeders, (c) a water quality monitoring unit (WQMU), (d) shrimp tracking and feed residue detection module, and (e) shrimp feeding characteristics quantification module (Figure 1). The UVMs and automatic feeders were employed to capture the videos of shrimp during feeding. The shrimp tracking and feed residue detection module tracked shrimp and detected residual feed in the videos. The shrimp feeding characteristics quantification module calculated five metrics related to shrimp feeding from the tracking and detection results: dwelling time, number of shrimps, frequency of entries, movement, and amount of the feed residue.

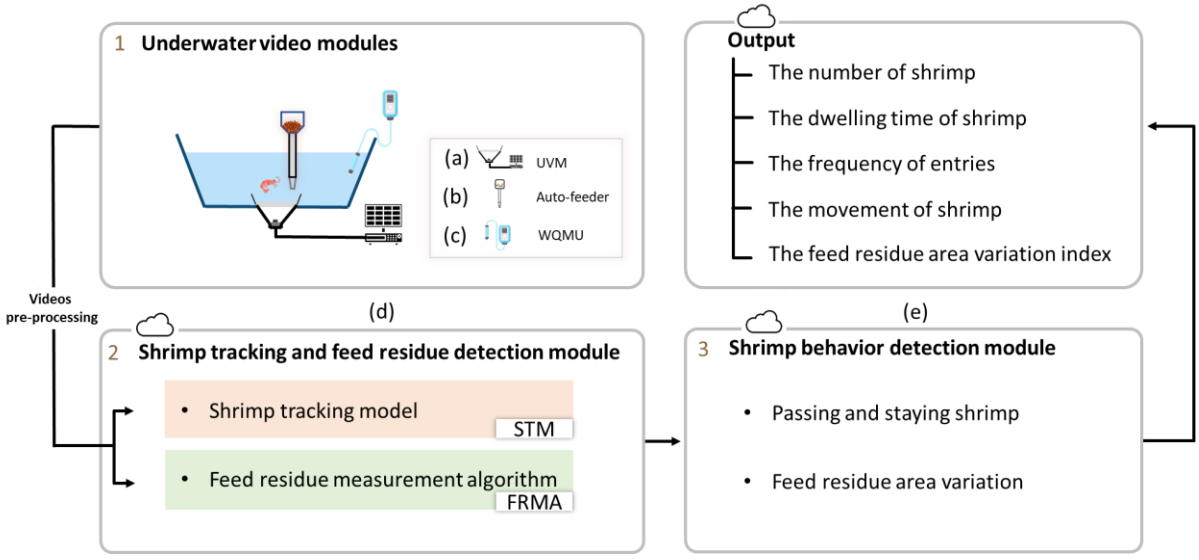


Figure 1. Flowchart of the proposed system.

3.2 Experiment site

The experiment was conducted in a pond at the Caregoods shrimp farm (YongAn, Kaohsiung, Taiwan). The pond was approximately 35 × 30 m (width × length) with water depth ranging from 1.2 to 1.5 m (Figure 2). Approximately 80,000 larvae of *Litopenaeus vannamei* were stocked in the pond. The pond was equipped with a paddle wheel aerator for oxygenation and a remover system in the middle of pond for sludge removal. Two pairs of UVMs and automatic

doi:10.6342/NTU202403695

feeders were installed at the opposite sides across the pond (I and II in Figure 2) to ensure dispersed sampling. The UVMs have infrared capabilities. Refer to Chu et al. (2019) for the design of the UVMs. The WQMU equipped with water sensors, providing the pH value, salinity, dissolved oxygen (DO), and water temperature (WT) of water. These values were record every hour with an accuracy of 98%. The weather data, including hourly temperature, humidity, and atmospheric pressure, were obtained from the publicly available databases of the Central Weather Bureau of Taiwan.

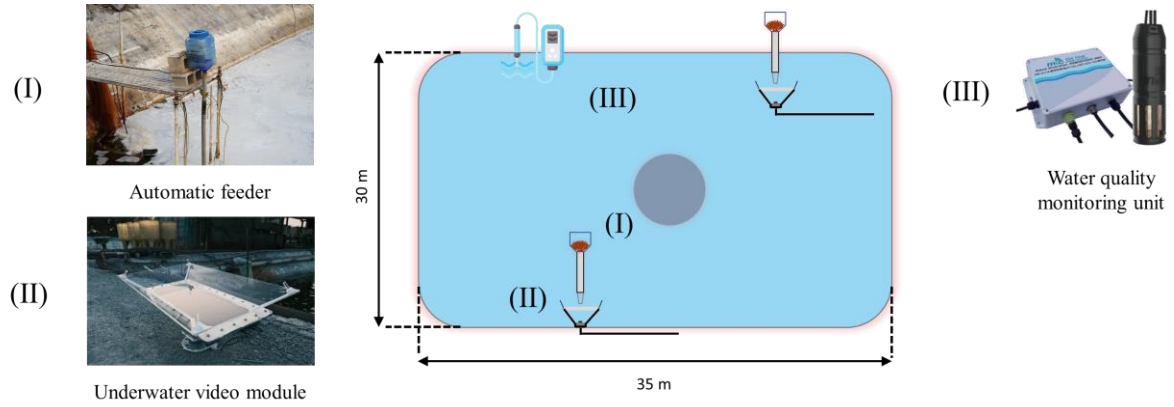


Figure 2. Configuration of the aquaculture pond. Photos of the (I) automatic feeder, (II) underwater video module, and (III) water quality monitoring unit.

3.3 Video acquisition

The UVMs acquired videos when the automatic feeders provided sample feed (i.e., a small amount of feed to test shrimp appetite). Four feeding times were scheduled per day, namely 2AM, 9AM, 3PM, and 7PM. The feeder dispensed feed for 1 min. During this period, approximately 30 grams of feed was dispensed. The UVMs acquired videos for 10 minutes at a frame rate of 15 frames per second (fps) and a resolution of 1920 × 1080 pixels. The videos were transmitted to a digital video recorder (8CH-HQ-1B8A, Jinwei Electronic Corporation; Taichung, Taiwan) via coaxial cables. The stored videos were retrieved and sent back to a network attached storage in the laboratory using a single-board computer (Raspberry Pi 4 / 4GB Model, Raspberry Pi Foundation; Cambridge, UK) through 4G internet. The experiment

doi:10.6342/NTU202403695

was conducted between November 2023 and January 2024. More than 80 hours of videos were collected.

3.4 Image collection and conditions

Images were converted from the collected videos for training the model for shrimp tracking and for developing the algorithms for feed residue detection. A total of 1733 images were collected, including 1520 daytime and 213 nighttime images (Figure 3). The daytime and nighttime images differed considerably in lighting conditions and were opposite in exhibition. The daytime images had bright background as the background is the natural illumination (i.e., sky). Objects in the daytime images were dark. In contrast, nighttime images had dark backgrounds and were unevenly illuminated by the infrared light at the center of the UVMs, making objects appear bright.

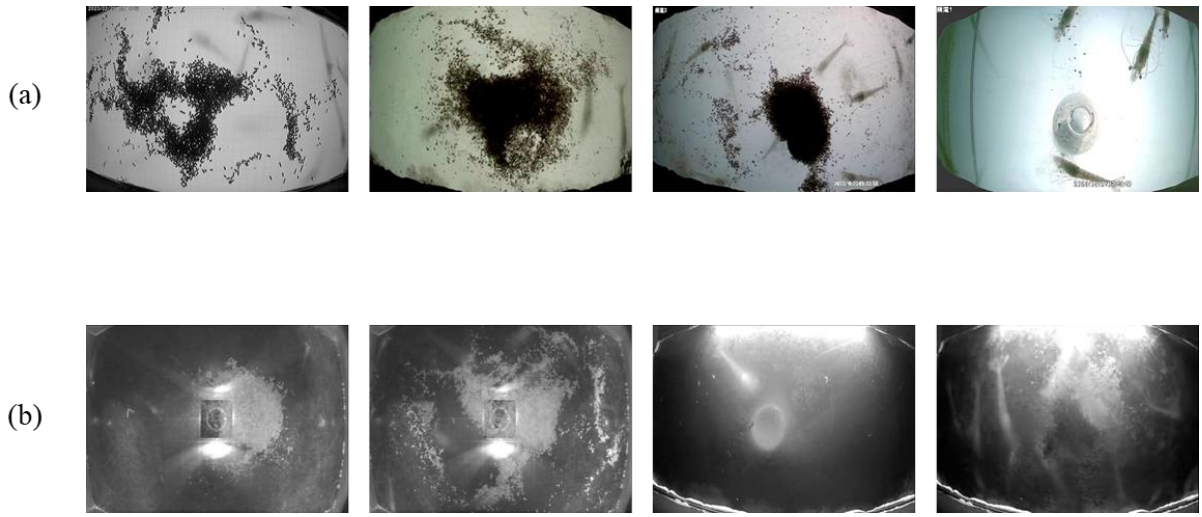


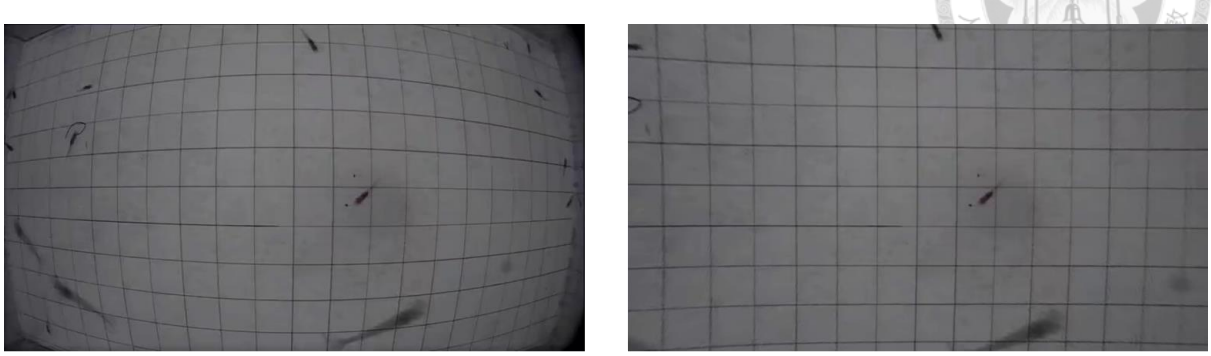
Figure 3. (a) Daytime and (b) nighttime images.

3.5 Image calibration

Image calibration was performed to rectify the barrel distortion (Figure 4a) introduced by the wide-angle lenses of the UVMs. In the calibration, a transparent acrylic panel with grids of known distances was put on the panels of the UVMs. Image calibration was performed using the method proposed by Zhang (2000) through OpenCV (Bradski et al., 2000; Figure 4b). The

doi:10.6342/NTU202403695

length-to-pixel ratio was also obtained in the process to calculate the physical lengths of shrimps.



(a) (b)

Figure 4. (a) Original and (b) calibrated UVM images.

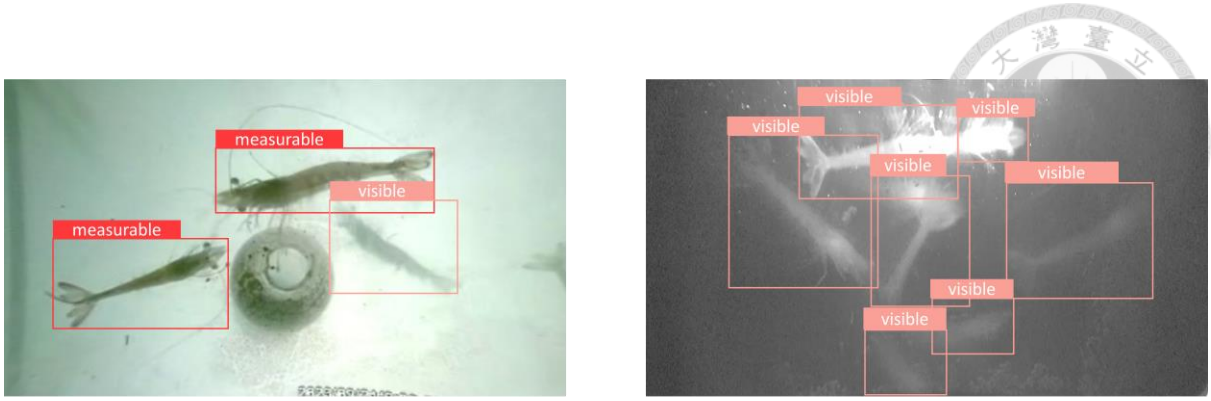
3.6 Image annotation

After calibration, Shrimps in the images were manually annotated into two categories, namely measurable and visible. Measurable shrimps are complete and exhibit distinct and unobstructed features of both the head and tail (Figure 5a). By contrast, visible shrimps are incomplete or blurred (Figure 5a and 4b). The blurriness often resulted from shrimps being out of focus when they left the surface of the tank cover. The shrimps in all nighttime images were categorized as visible due to the natural characteristics of the images (Figure 5b). Both the measurable shrimps were used for counting shrimp numbers, whereas only the measurable shrimps were used for length estimation. The annotation was performed using the LabelImg toolkit (Lin, 2015). A total of 382 measurable and 6554 visible shrimps were annotated (Table 1). The annotated images were split into training and test at a ratio of 4:1.

Table 1. Amount of training and test images and labelled bounding boxes for the SFRCM

Category	Training			Test		
	Image	Visible	Measurable	Image	Visible	Measurable
Daytime	1213	4530	288	276	813	94
Nighttime	212	1119	0	32	92	0
Total	**1425**	**5649**	**288**	**308**	**905**	**94**

doi:10.6342/NTU202403695



(a) (b)

Figure 5. The annotations of measurable and visible shrimps in (a) daytime and (b) nighttime images.

3.7 Shrimp detection and tracking

Tracking-by-detection was used as the strategy for shrimp tracking. The compact version of You Only Look Once—version 9 (YOLO v9-c; Wang et al., 2024) was used for shrimp detection. Simple online and real-time tracking (SORT; Bewley, 2016) was used for tracking the same shrimp in consecutive frames. To train the YOLO v9-c, the annotated images were resized to 640 × 640 pixels. The model was initialized with the parameters pretrained using the MS COCO dataset (Lin et al., 2014). The hyperparameters of the model were then optimized using the genetic algorithm (GA; Mirjalili et al., 2019) for 300 iterations. In each iteration, four batches of images were used to update the weights of the model. Stochastic gradient descent (Bottou, 2012) was used as the optimizer. As the result of the GA, the initial learning rate, momentum, and warmup epochs were set to 0.01, 0.9, and 2.8, respectively. The weight decay and final learning rate were set to 0.0004 and 0.08, respectively. Online image augmentation was applied during the training to enhance the robustness of YOLOv9-c. The augmentation techniques included color space transformation, translation, scale transformation, horizontal flip, copy-paste, and mosaic. The probabilities that a technique was applied to a training image were set to 0.7, 0.1, 0.9, 0.1, and 1.0. The model was trained using an open-source python

doi:10.6342/NTU202403695

environment (Van Rossum, 1991). A graphics processing unit (GeForce RTX TITAN, NVIDIA; Santa Clara, CA, USA) was used for training the model.

3.8 Feed residue detection

Two pipelines of image processing algorithms were developed to determine the areas of the feed residue (FR) in the images. A pipeline was for the daytime images, whereas the other pipeline is for the nighttime images. To determine the FR for the daytime images, the original images (Figure 6a) were converted into grayscale (Figure 6b). Dynamic thresholding (Zhang and Suen, 1984) was next applied to the grayscale images to adaptively adjust the brightness of the images based on local image brightness variations, ensuring effective foreground and background segmentation under varying lighting conditions. Subsequently, inverse binary thresholding (Otsu, 1979) was used to highlight the areas of feed residue (Figure 6c), followed by morphological closing operations (Serra, 1982) to refine the segmentation and to remove noise (Figure 6d). The closing operations used a circular structuring element of 3 pixels in size.

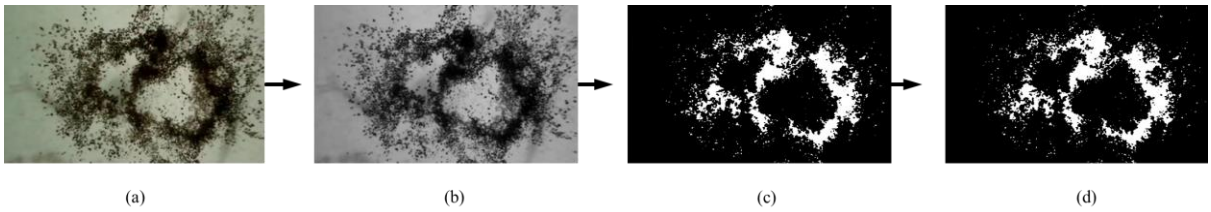


Figure 6. Pipeline of the FR area identification on daytime images: (a) original image, (b) grayscale image, (c) binarized image using dynamic thresholding, and (d) binarized image after morphological operations.

To determine the FR for the nighttime images, the original images were first adjusted using gamma correction (Pizer et al., 1987), followed by grayscale conversion (Figure 7b). Dynamic contrast limited adaptive histogram equalization (CLAHE; Zuiderveld, 1994) was next applied to the grayscale images to enhance the contrast of the image (Figure 7c). Adaptive thresholding was performed based on the brightness of the central image regions to adaptively highlight the areas of the FR (Figure 7d). Lastly, geometric filtering was conducted to filter out irrelevant shapes based on size (filtering out contours with area < 100 pixels), shape (filtering out

doi:10.6342/NTU202403695

contours with aspect ratio > 3 or < 0.33), and position (filtering out contours in the central 1/3 to 2/3 region of the image in both width and height) criteria of the foreground, while retaining the foreground shapes corresponding to feed residue (Figure 7e).

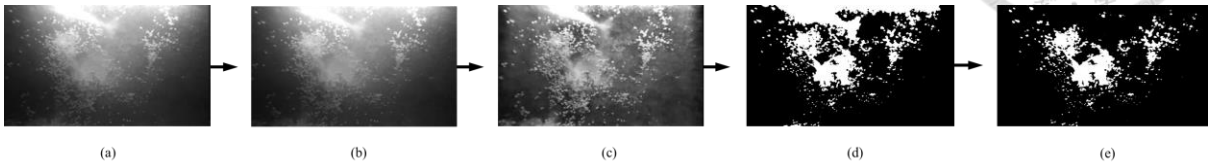


Figure 7. Pipeline of the FR area identification on nighttime images: (a) original image, (b) gamma-corrected grayscale image, (c) CLAHE image, (d) binarized image, and (e) binarized image after geometric filtering.

3.9 Feeding-related characteristics of shrimps

The feeding-related characteristics of shrimps quantified in this study included (a) shrimp count, (b) movement of shrimps, (c) entry and exit frequency of shrimps, (d) shrimp dwelling time, and (e) feed residue area variation index (FRAVI). These characteristics are metrics of shrimp feeding behaviors, group dynamics, and the attractiveness of the provided feed. The characteristics are essential for the practices of shrimp farming. Shrimp count was defined as the number of shrimps detected and tracked over a 10-min video. Movement of shrimps was defined as the summation of the Euclidean distances between the bounding box centroids of shrimps in consecutive frames for all the shrimps over a 10-min video. Entry and exit frequency of shrimps was defined as the frequency that shrimps entered and exited a predefined region of interest (RoI), specifically within the pixel coordinates (50, 50, 1850, 1070), which was the primary area where feed dispersion was expected to occur (Figure 8).

doi:10.6342/NTU202403695

Figure 8. Illustration of the predefined ROI (yellow bounding box).

Entry frequency per video sequence were calculated. Shrimp dwelling time was defined as the time the shrimps stayed within the RoI. A shrimp was considered as staying if its centroid velocity was below 4.4 cm/second and its centroid was located inside the RoI. These staying shrimps were presumed to be feeding. Therefore, shrimp dwelling time reflects the attractiveness of the provided feed to the shrimp group. FRAVI was defined as the change in feed residue area over a 10-min video and was computed as the follows:

$$FRAVI = \frac{A_M - A_F}{T_D}, \quad (1)$$

where A_M is the maximum feed residue area in the 10-min video, A_F is the final feed residue area at the end of the video, and T_D is the time duration from the time when the maximum feed residue area occurred to the video end time. A higher FRAVI value indicates a faster reduction in feed residue area, suggesting a stronger attraction of the shrimp group to the provided feed.

doi:10.6342/NTU202403695

CHAPTER 4. RESULTS AND DISCUSSION

4.1 Performance of shrimp detection model

The detection performance of the trained YOLO v9-c was evaluated using the 308 test images. During the evaluation, the confidence score threshold for positive detection was set to 0.6, and the intersection over union (IoU) threshold for true positive (TP), false positive (FP), and false negative (FN) detection was set to 0.7. The trained model achieved a precision of 93.7%, a recall of 98.2%, an F1 score of 86.4%, and a mean average precision (mAP) of 92.3% in shrimp localization. The precision-recall curve for shrimp localization was illustrated (Figure 9).

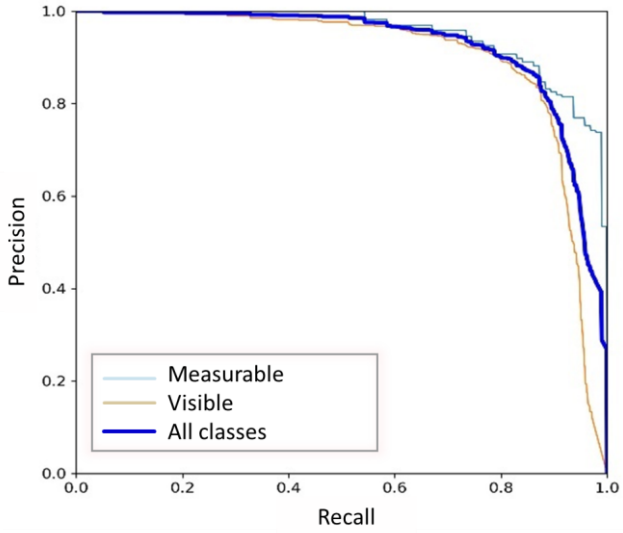


Figure 9. Precision-recall curve of the trained shrimp detection model.

4.2 Challenging scenarios in shrimp detection

The robustness of STM in challenging scenarios was further examined. The scenarios included dim lighting conditions (Figure 10a), overly bright lighting (Figure 10b), signal distortions (Figure 10c), uneven lighting conditions (Figure 10d), and partial occlusion of shrimp by feed residue (Figure 10e). Grad-cam visualizations were used to demonstrate the model's focus on relevant features, ensuring accurate shrimp detection despite these complexities (II in Figure 10).

doi:10.6342/NTU202403695

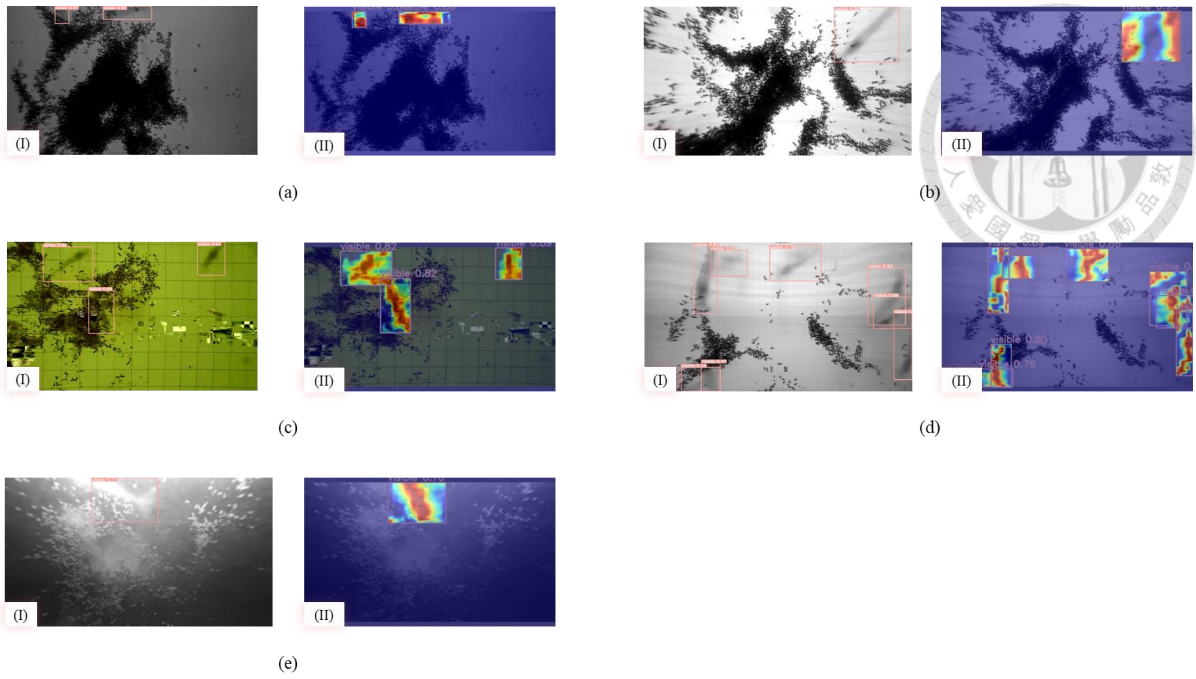


Figure 10. Challenging scenarios: (a) overly bright lighting, (b) dim lighting, (c) signal distortions, (d) uneven lighting conditions, and (e) partial occlusion of shrimp by feed residue. The figure comprised (I) original images and (II) Grad-cams.

4.3 Performance of shrimp tracking

The tracking performance of SORT was evaluated using a video and the multiple object tracking (MOT) metrics (Milan et al., 2016). The video comprised 300 consecutive frames (30 s in duration). The ground truth (GT) locations of the shrimp were labeled using the Darklabel tool (Darkpgmr, 2020). SORT achieved an overall MOT accuracy of 83.3% in shrimp tracking (Table 2).

Table 2. Evaluation result of shrimp tracking

Frame	GT	MT	PT	ML	IDs	Precision	Recall	MOTA
300	82	8	4	1	18	91.5 %	90.6 %	83.3 %

Frame = number of frame; GT = number of shrimp in the frame; MT = number of mostly tracked; PT = number of partially tracked; ML = number of mostly lost; IDs = ID switching; MOTA = multiple object tracking accuracy.0

4.4 Performance of the feed residue measurement algorithms

The performance of the proposed FR detection approach was evaluated using 100 daytime and 100 nighttime images. The GT of the FR in the images were annotated using Photoshop (Adobe; San Jose, CA, USA). Regression analysis was performed to assess the correlation between the

doi:10.6342/NTU202403695

GT and prediction FR areas. The proposed FR detection approach achieved correlations of 0.995 and 0.775 (Table 3), respectively, on daytime and nighttime images, demonstrating a reasonable level of accuracy in FR detection.

Table 3. Evaluation result of FR measurement algorithms

Type	Correlation	R^2	Average F1 Score
Daytime	0.995	0.989	0.689
Nighttime	0.775	0.723	0.568

4.5 Comparative analysis of feeding-related characteristics in strong and weak conditions

FR areas in two 10-min-long videos were illustrated (Figure 11). Feed dispensers dispensed feed gradually. The FR areas were observed to reach maximums in approximately two minutes after the start of dispensing. In certain instance, the FR area shrank rapidly (Figure 11a), indicating fast feed consumption was rapid and resulting in a large FRAVI value. In certain other instance, the FR area reduced gradually (Figure 11b), indicating slow feed consumption and resulting in a small FRAVI value.

doi:10.6342/NTU202403695

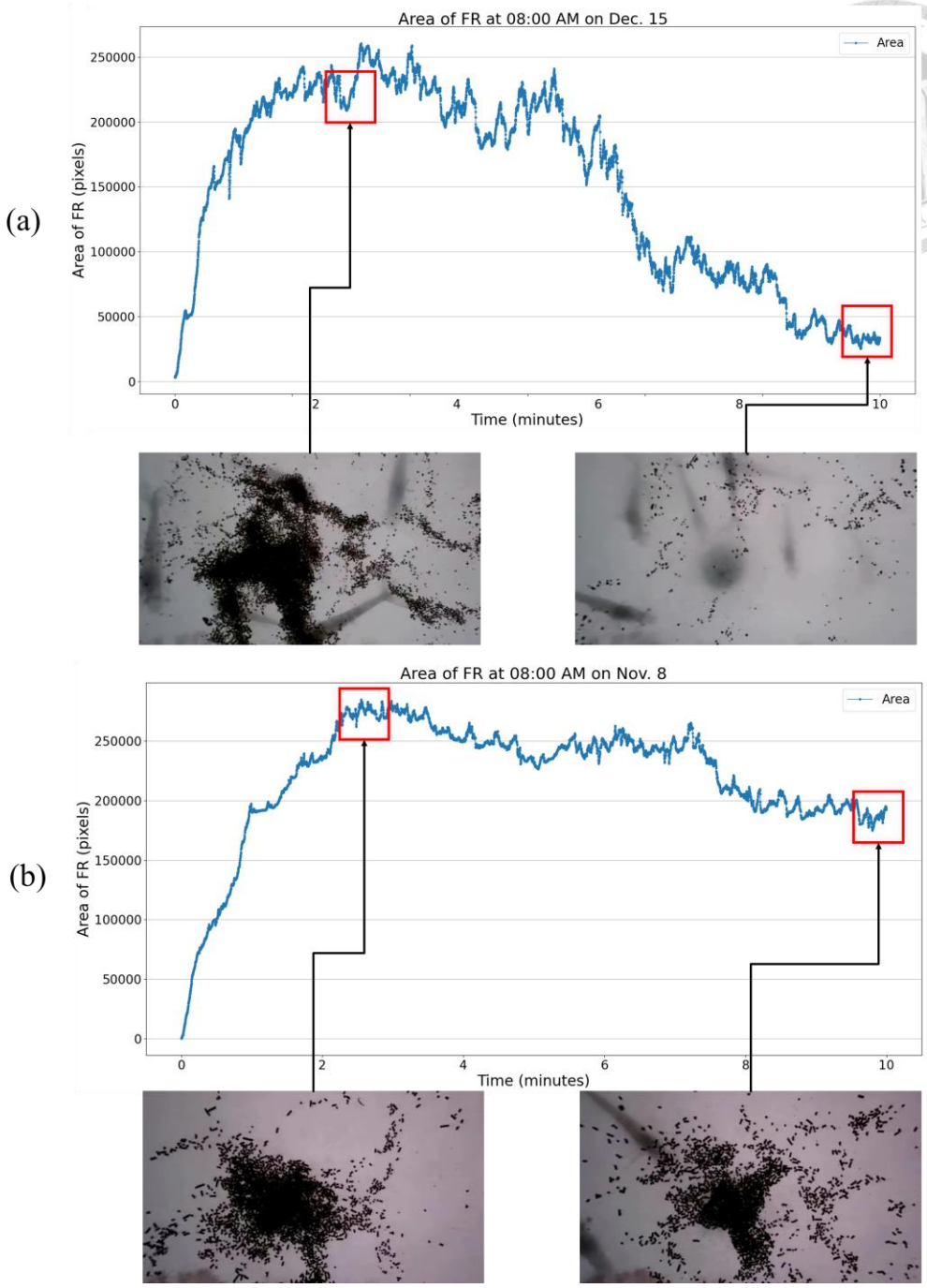


Figure 11. FR area under (a) strong and (b) weak conditions during the trait feeding period.

Shrimp counts in two 10-min-long videos were illustrated (Figure 12). In certain instance, shrimp count peaked around two minutes after the start of feed dispensing (Figure 12a), indicating that the shrimps have high levels of appetite. In certain other instance, the shrimp count remained generally low (Figure 12b), indicating that the shrimps have low levels of appetite.

doi:10.6342/NTU202403695

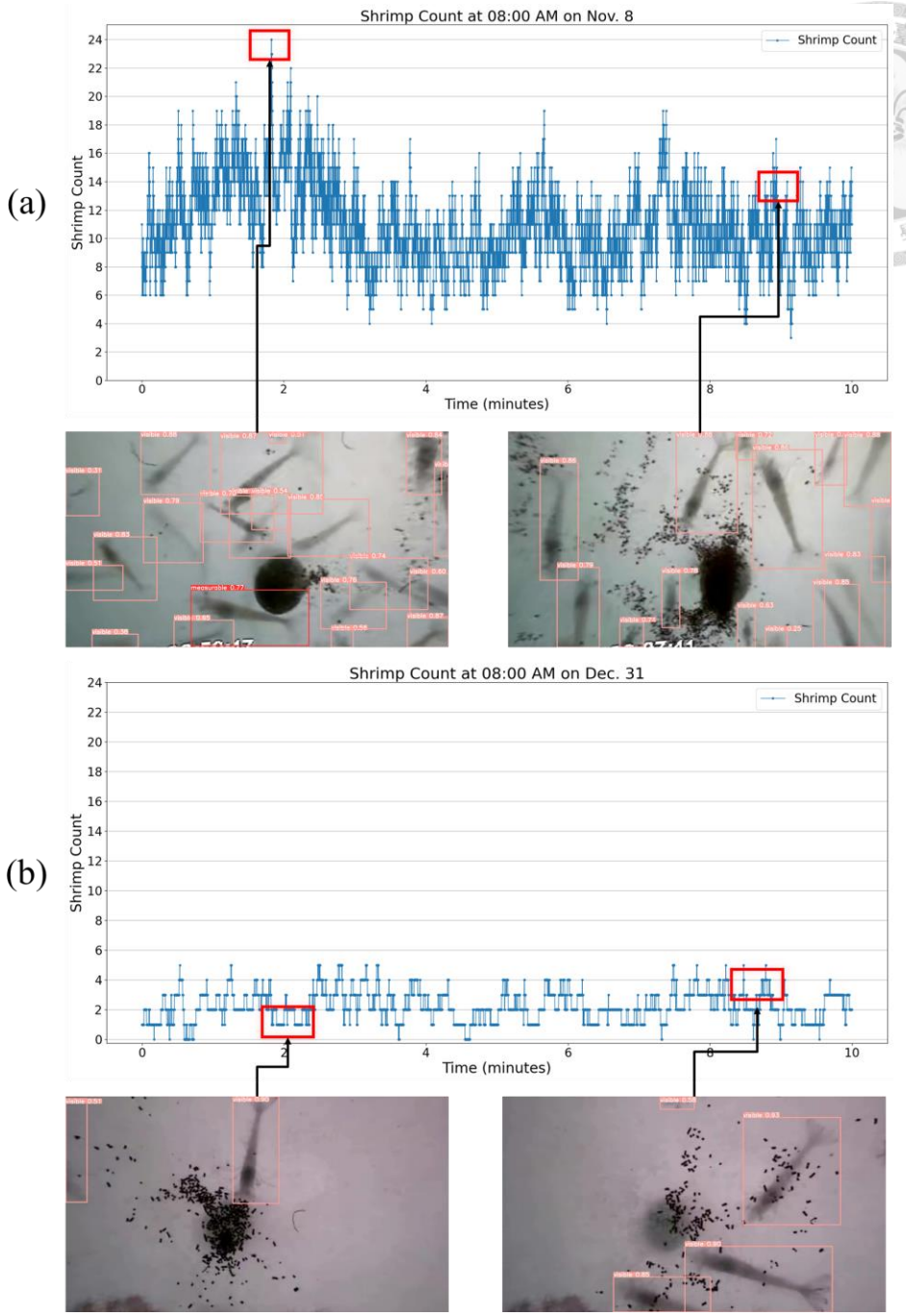


Figure 12. Shrimp count under (a) strong and (b) weak conditions during the trait feeding period.

Entry frequencies of shrimps in two 10-min-long videos were illustrated (Figure 13). In certain instance, a relatively large numbers of shrimps entering the RoI after the start of feed dispensing (Figure 13a), indicating a high level of appetite of the shrimps. In certain other instance, a relatively low numbers of shrimps entering the RoI after the start of feed dispensing, potentially indicating a low level of appetite of the shrimps.

doi:10.6342/NTU202403695

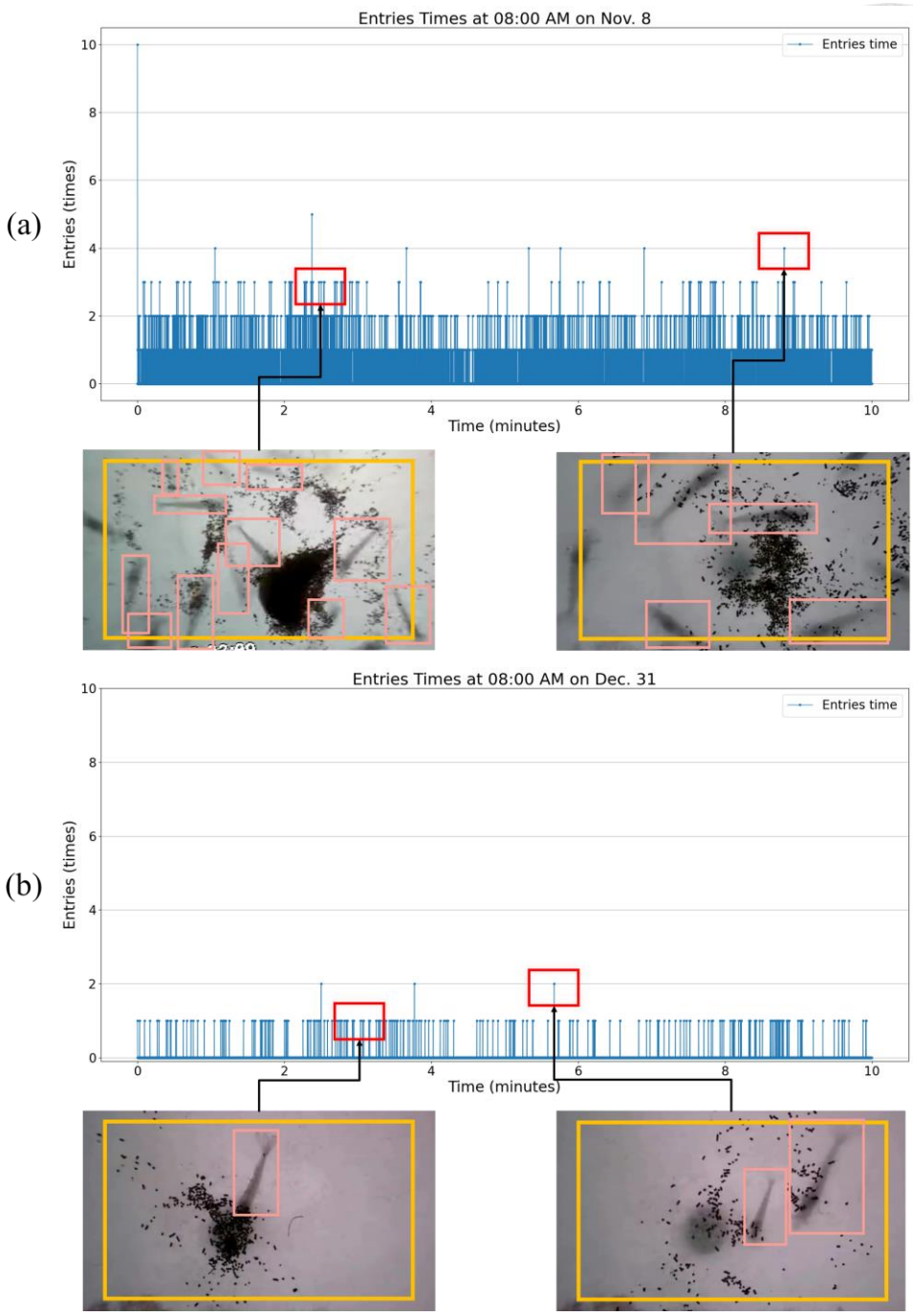


Figure 13. Frequency of shrimp entering the RoI under (a) strong and (b) weak conditions during the trait feeding period.

Stay duration of shrimps in two 10-min-long videos were illustrated (Figure 14). Under strong conditions (Figure 14a), the duration of stay within the RoI generally peaked around two minutes after feed introduction, suggesting a high level of attraction and prolonged feeding behavior. Conversely, under weak conditions (Figure 14b), the duration of stay was typically low, indicating a lower degree of attraction to the feed.

doi:10.6342/NTU202403695

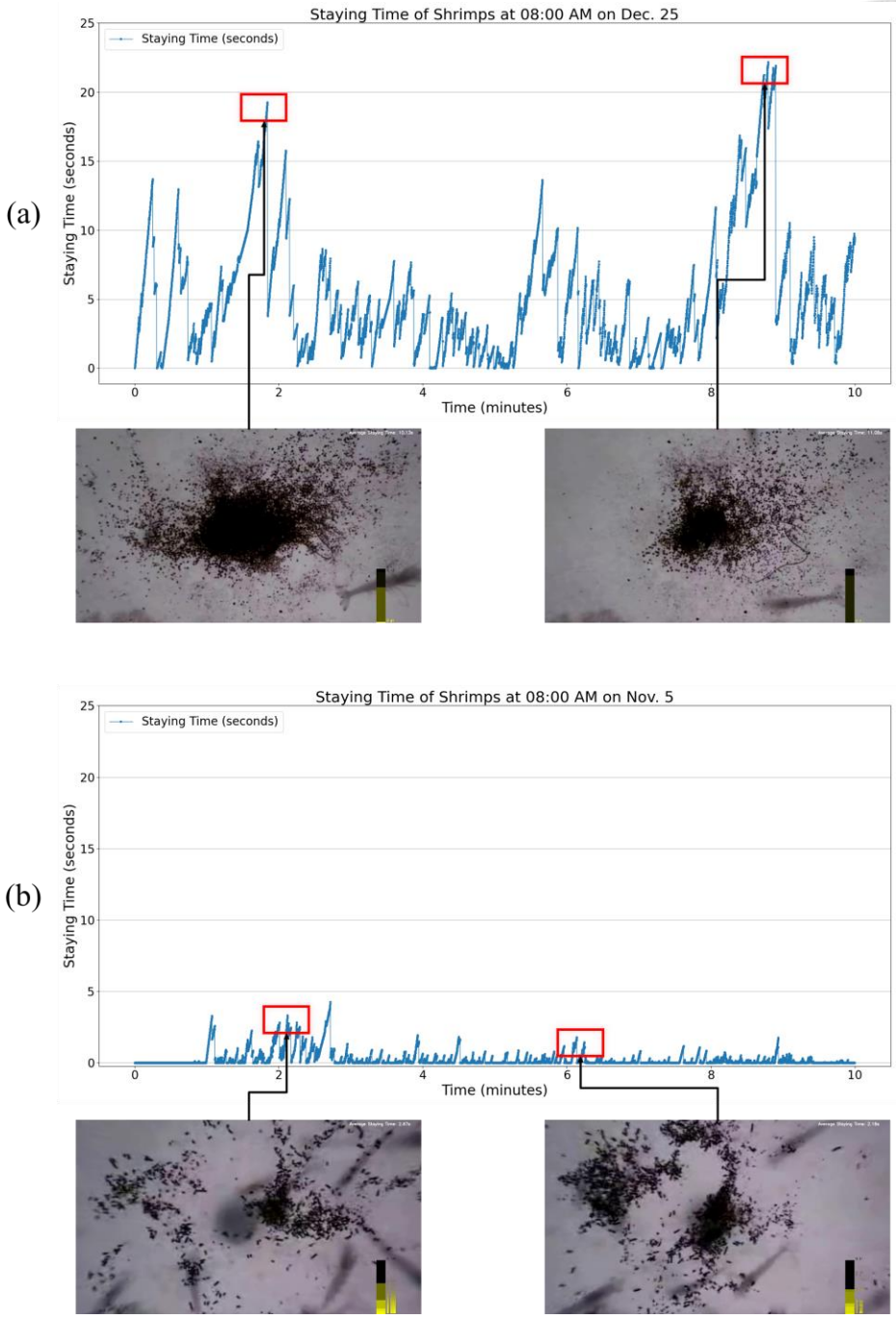


Figure 14. Duration of stay within the RoI under (a) strong and (b) weak conditions during the trait feeding period.

Shrimp movement in two 10-min-long videos were illustrated (Figure 15). Under strong conditions (Figure 15a), movement activity peaked around eight minutes after feed introduction, indicating strong attraction to the feed. In contrast, under weak conditions (Figure 15b), movement activity remained low, suggesting a lesser attraction to the feed. In the histogram of figure, darker colors represent longer durations of stay.

doi:10.6342/NTU202403695

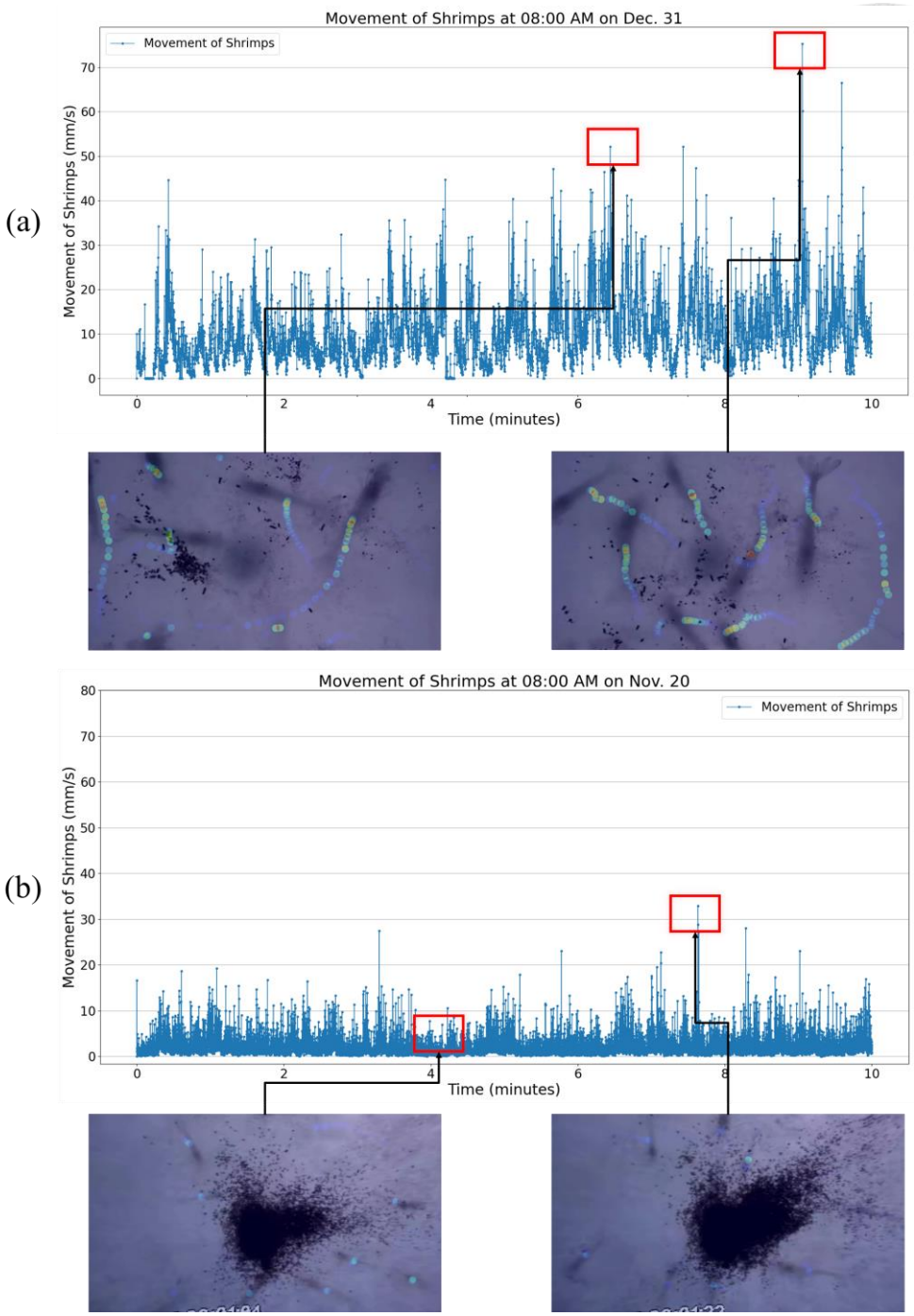


Figure 15. Shrimp movement under (a) strong and (b) weak conditions during the trait feeding period (represented by heat maps).

4.6 Feeding-related characteristics in a long-term study

The mean daily feeding-related characteristics and their standard deviations in the experiment period (November 2023 and January 2024) were illustrated (Figure 16). Initially, FRAVI remained relatively stable, averaging at 250 ± 50 pixels/s. However, around December 12,

doi:10.6342/NTU202403695

FRAVI peaked to 1000 ± 500 pixels/s (Figure 16a), indicating improved appetite. Shrimp count peaked around November 4 at approximately 600, following by a gradual declination to 30 in early January (Figure 16b). This pattern was mirrored in the entry frequency into the feeding area, which peaked at 700 times before declining to 50 times (Figure 16c), suggesting a correlation between shrimp count and entry frequency. The duration of stay in the feeding area showed a gradual upward trend, peaking on December 26 at 8 s (Figure 16d). Despite periodic fluctuations, the overall movement speed of shrimp increased, reaching a peak of 9 mm/s in the two final days (Figure 16e).

doi:10.6342/NTU202403695

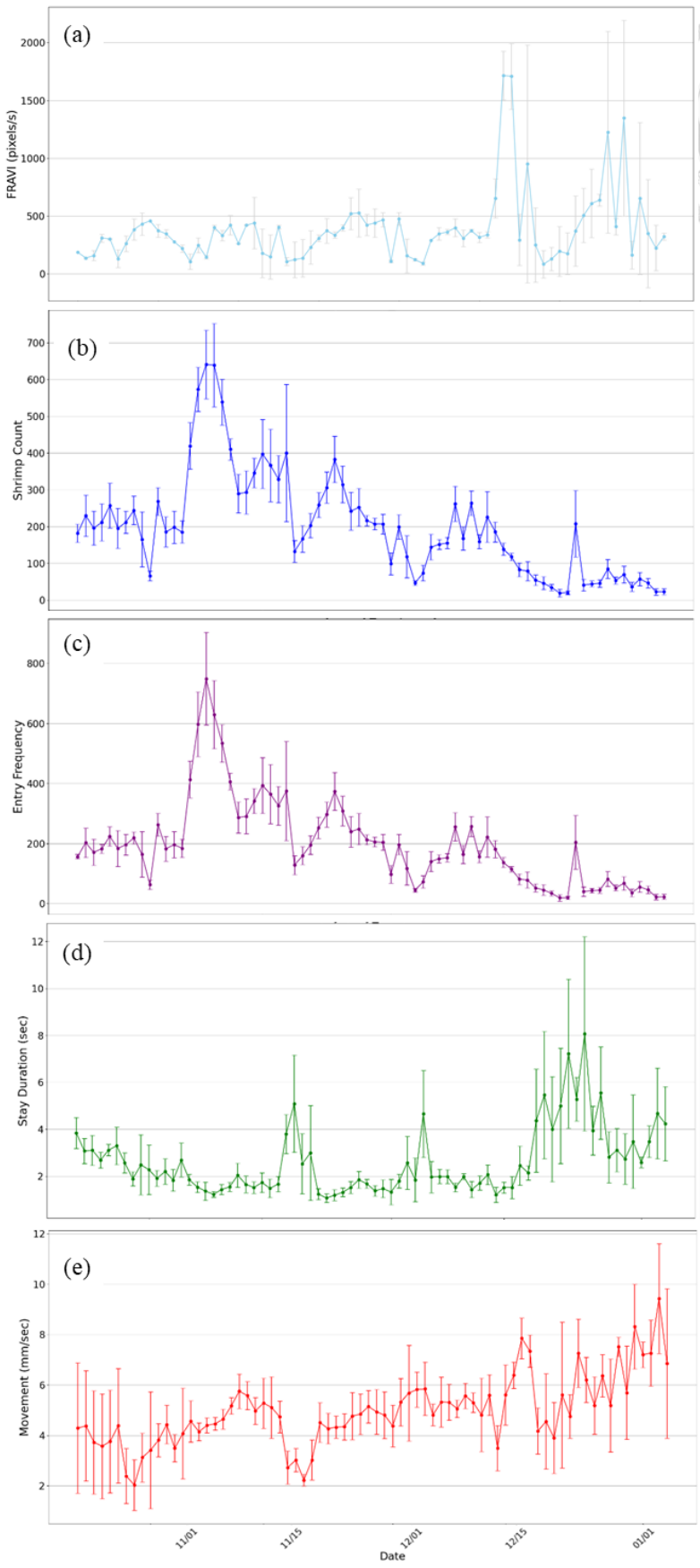


Figure 16. Illustration of (a) FRAVI, (b) shrimp count, (c) entry frequency, (d) duration of stay, and (e) movement for a single cultivation batch.

doi:10.6342/NTU202403695

4.7 Comparison of feeding-related characteristics between daytime and nighttime

The feeding-related characteristics between daytime and nighttime were illustrated using boxplots (Figure 17). A T-test（Student, 1908）was conducted to analyze these differences (Table 4). The shrimp count was higher at night, with a mean of 300 compared to a mean of 200 during the day, though this difference was not statistically significant (T-test: $t = -1.05$, $p = 0.293$). In contrast, the entry frequency was also higher at night, with a mean of 350 entries versus 250 entries during the day (T-test: $t = -3.60$, $p = 4.11 \times 10^{-4}$). Furthermore, the duration of stay was markedly longer during the day, with a mean of 2 seconds, compared to a mean of 1 second at night (T-test: $t = 8.93$, $p = 1.82 \times 10^{-16}$). Movement was slightly lower at night, with a mean of 4.5 mm/s compared to 5 mm/s during the day (T-test: $t = -3.15$, $p = 1.81 \times 10^{-3}$). Despite the decrease in movement speed at night, shrimp remained active, indicating continuous feeding activity.

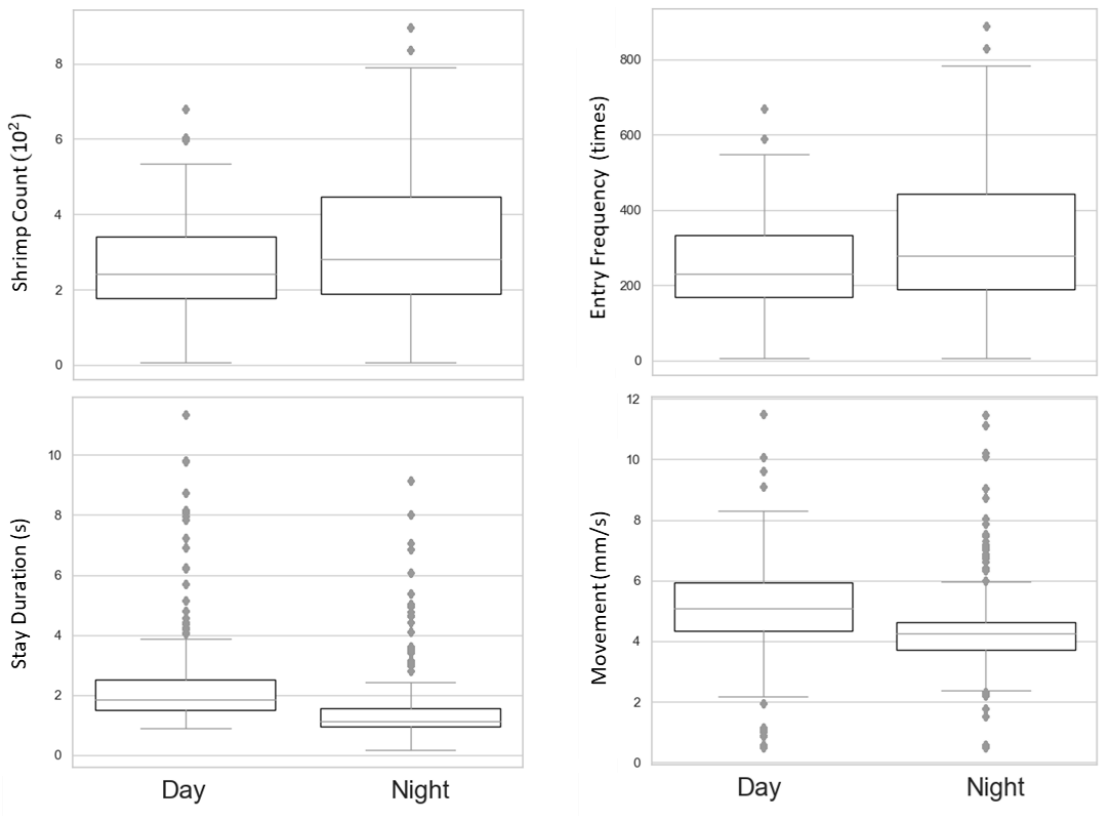


Figure 17. Boxplots illustrating the differences in (a) shrimp count, (b) duration of stay, (c) movement, and (d) entry frequency between daytime and nighttime conditions.

doi:10.6342/NTU202403695

Table 4. T-test of feeding-related characteristics between daytime and nighttime

Type	T-test	p-value
Shrimp count	-1.05	2.93×10^{-1}
Entry frequency	-3.60	4.11×10^{-4}
Stay duration	8.93	1.82×10^{-16}
Movement	3.15	1.81×10^{-3}

4.8 Correlation between feeding-related characteristics and environmental conditions

The daily mean water and weather conditions with standard deviations were quantified and illustrated (Figure 18). During the cultivation period, water quality data (Figure 18a) indicated that the temperature steadily declined over time. It is noteworthy that salinity, pH levels, and dissolved oxygen levels initially exhibited data gaps due to changes in the field setup; therefore, these invalid data were removed.

Weather data (Figure 18b) indicated the ambient temperature showed a trend similar to that of the water temperature. Station pressure fluctuated within a certain range, and the relative humidity generally remained between 60% and 90% daily. There were periods with missing data, and short-term rain events occurred during the rainy season, but the overall rainfall was quite limited.

doi:10.6342/NTU202403695

Figure 18. (a) Water quality data and (b) weather data during the cultivation period.

The correlation between shrimp feeding characteristics and various environmental conditions was illustrated in the correlation matrix (Figure 19). The FRAVI demonstrated a moderate positive correlation with stay duration (r = 0.33), suggesting that longer stay durations are indicative of higher shrimp appetite. Conversely, FRAVI showed moderate negative correlations with water temperature (r = -0.47), shrimp count (r = -0.43), and entry frequency (r = -0.43), implying that higher water temperatures, shrimp counts, and entry frequencies were associated with lower FRAVI values. Shrimp count exhibited strong positive correlations with water temperature (r = 0.63), entry frequency (r = 0.82), and movement (r = 0.70). This suggests that increased water temperatures, frequent entries, and active movement promote higher shrimp counts in the feeding area. However, shrimp count was negatively correlated with

doi:10.6342/NTU202403695

salinity (r = -0.73) and stay duration (r = -0.53), indicating that higher salinity levels and longer stay durations reduce shrimp count. Entry frequency was strongly positively correlated with water temperature (r = 0.67), shrimp count (r = 0.82), and station pressure (r = 0.73), and moderately positively correlated with movement (r = 0.70). These correlations indicate that increased water temperatures, shrimp counts, station pressure, and movement activity were associated with more frequent entries into the RoI. Negative correlations with salinity (r = -0.71) and stay duration (r = -0.39) suggested that higher salinity levels and longer stay durations reduce the frequency of shrimp entering the RoI. Conversely, stay duration was negatively correlated with water temperature (r = -0.71) and station pressure (r = -0.59), indicating that higher water temperatures and station pressures reduce the time shrimp spend in the feeding area. Movement showed positive correlations with shrimp count (r = 0.70) and entry frequency (r = 0.62), suggesting that higher shrimp counts, frequent entries, and increased movement activity were associated with more active movement.

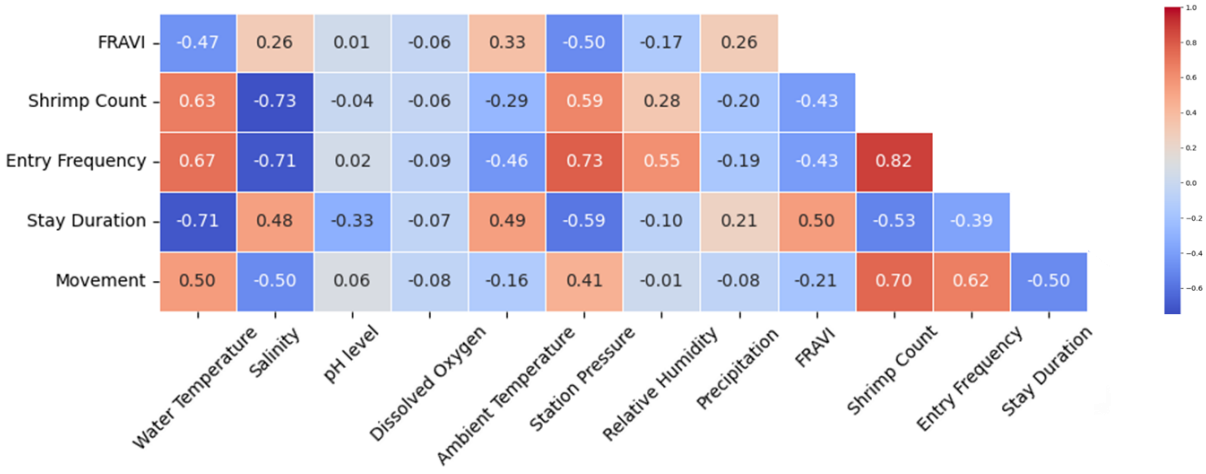


Figure 19. Correlation matrix of between characteristics and environmental conditions.

4.9 Shrimp count difference between the two UVMs

The shrimp counts from the two UVMs were compared to check if the sampling of shrimp during feeding events was unbiased and if the strategy of putting the UVMs at opposite sides of the aquaculture pond was valid. The comparison was performed across the measurement of

doi:10.6342/NTU202403695

six weeks (Figure 20). Visual inspection revealed a strong positive correlation between the two sets of shrimp count data, with data points distributed along a linear trend, indicating no significant spatial biases or clustering tendencies of the shrimp population within the pond. The Pearson correlation coefficient between the shrimp counts recorded by Camera 1 and Camera 2 was calculated to be 0.828. Although minor deviations from the linear trend were observed, likely due to inherent variability in shrimp behavior and dynamic movement within the pond, these deviations did not significantly impact the representativeness of the sampling data from the two locations. These findings confirm that during feeding events, the shrimp were relatively uniformly distributed throughout the pond, and the data collected from the two strategically positioned UVMs provided an unbiased representation of the overall shrimp population.

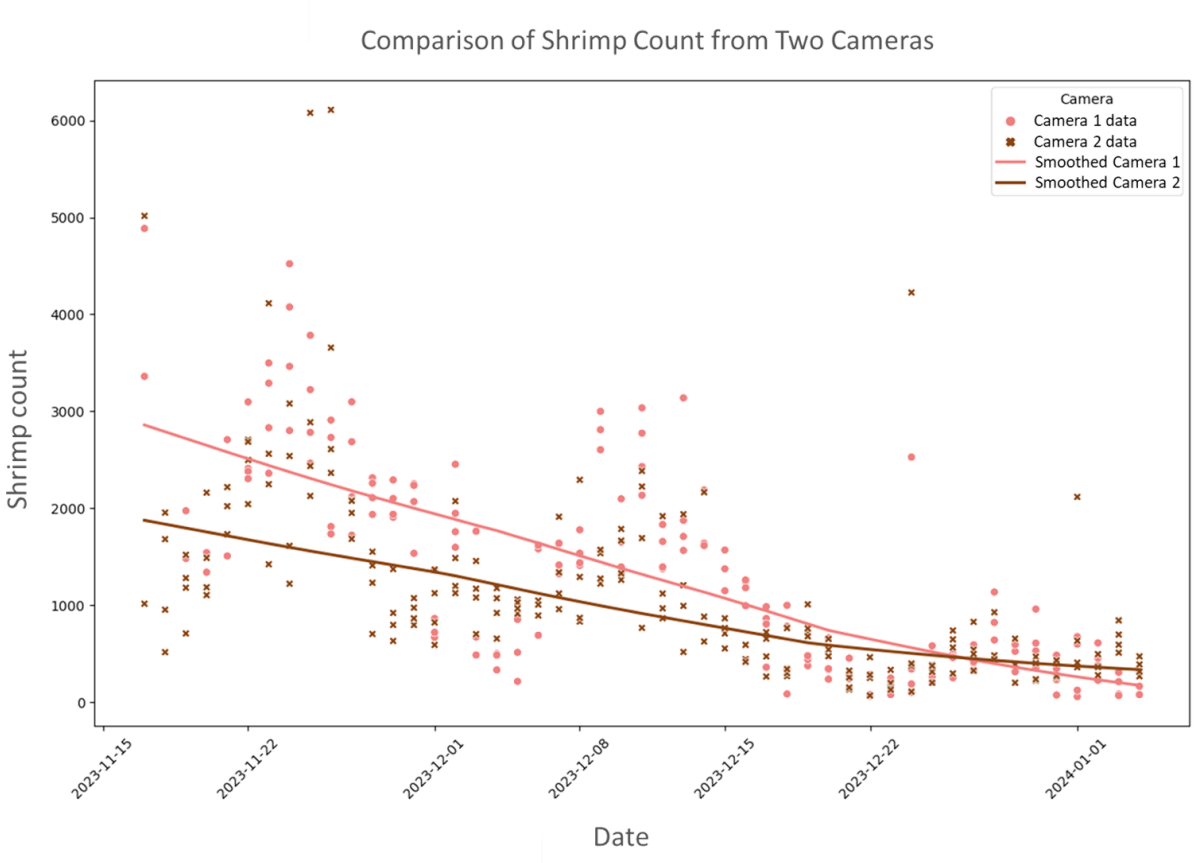


Figure 20. Comparison of shrimp counts from two cameras during feeding sessions.

doi:10.6342/NTU202403695

CHAPTER 5. CONCLUSIONS

5.1 Summary

This study developed an automated system to quantify shrimp feeding-related characteristics using computer vision. The quantified characteristics included shrimp count, movement, entry frequency, dwelling time, and FRAVI. The proposed method achieved a mAP of 92.3% and a MOTA of 83.3%, respectively, in shrimp detection and tracking. Analysis revealed a strong positive correlation between increased shrimp feeding activity and higher water temperatures ($r = 0.63$) and a strong negative correlation between feeding activity and salinity levels ($r = -0.73$). Additionally, variations in station pressure were found to significantly influence shrimp entry frequency ($r = 0.73$) and FRAVI value ($r = -0.50$). The proposed approach is fully automatic and objective and may help in optimizing feed management strategies in shrimp farming.

5.2 Future work

In the future, we will focus on two main areas: first, integrating sophisticated behavior recognition algorithms to identify individual shrimp behaviors during feeding, and second, incorporating body length measurement using image processing techniques to enable comprehensive monitoring of shrimp health and growth. These enhancements will provide a more holistic view of shrimp health and behavior, contributing to more effective and sustainable shrimp farming practices.

doi:10.6342/NTU202403695

REFERENCES

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., et al. (2016). TensorFlow: A system for large-scale machine learning. In 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16) (pp. 265–283). https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi

Ames, R. T., Leaman, B. M., & Ames, K. L. (2007). Evaluation of video technology for monitoring of multispecies longline catches. North American Journal of Fisheries Management, 27, 955–964.

Álvarez-Ellacuría, A., Palmer, M., Catalán, I. A., & Lisani, J. L. (2019). Image-based unsupervised estimation of fish size from commercial landings using deep learning. ICES Journal of Marine Science. doi:10.1093/icesjms/fsz216.

Bartholomew, D. C., Mangel, J. C., Alfaro-Shigueto, J., Pingo, S., Jimenez, A., & Godley, B. J. (2018). Remote electronic monitoring as a potential alternative to on-board observers in small-scale fisheries. Biological Conservation, 219, 35–45.

Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent. In Proceedings of COMPSTAT'2010 (pp. 177–186). Physica-Verlag HD.

Chollet, F. (2015). Keras. https://github.com/fchollet/keras.

Ditria, E. M., Lopez-Marcano, S., Sievers, M. K., Jinks, E. L., Brown, C. J., & Connolly, R. M. (2019). Automating the analysis of fish abundance using object detection: Optimising animal ecology with deep learning. bioRxiv, 805796.

Everingham, M., & Winn, J. (2011). The PASCAL Visual Object Classes Challenge 2012 (VOC2012) Development Kit. Pattern Analysis, Statistical Modelling, and Computational Learning Technical Report.

doi:10.6342/NTU202403695

FAO. (2017). Seafood Traceability for Fisheries Compliance: Country-Level Support for Catch Documentation Schemes. FAO, Rome.

FAO. (2018). The State of World Fisheries and Aquaculture 2018 (SOFIA 2018): Meeting the Sustainable Development Goals. Food and Agriculture Organization, Rome.

Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27, 861–874.

Francisco, F. A., Nührenberg, P., & Jordan, A. L. (2019). A low-cost open-source framework for tracking and behavioural analysis of animals in aquatic ecosystems. bioRxiv, 571232.

French, G., Fisher, M. H., Mackiewicz, M., & Needle, C. (2015). Convolutional neural networks for counting fish in fisheries surveillance video. In Proceedings of the Machine Vision of Animals and their Behaviour (MVAB) (pp. 7–1).

French, G., Mackiewicz, M., Fisher, M., Holah, H., Kilburn, R., Campbell, N., & Needle, C. (2019). Deep neural networks for analysis of fisheries surveillance video and automated monitoring of fish discards. ICES Journal of Marine Science. doi:10.1093/icesjms/fsz149.

Garcia, R., Prados, R., Quintana, J., Tempelaar, A., Gracias, N., Rosen, S., Vågstøl, H., et al. (2019). Automatic segmentation of fish using deep learning with application to fish size measurement. ICES Journal of Marine Science. doi:10.1093/icesjms/fsz186.

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (pp. 2961–2969).

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 770–778).

doi:10.6342/NTU202403695

Jaeger, J., Wolff, V., Fricke-Neuderth, K., Mothes, O., & Denzler, J. (2017). Visual fish tracking: Combining a two-stage graph approach with CNN-features. In OCEANS 2017-Aberdeen (pp. 1–6). IEEE.

Kindt-Larsen, L., Kirkegaard, E., & Dalskov, J. (2011). Fully documented fishery: A tool to support a catch quota management system. ICES Journal of Marine Science, 68, 1606–1610.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 2, 1097–1105.

Larsen, R., Olafsdottir, H., & Ersbøll, B. K. (2009). Shape and texture based classification of fish species. In Scandinavian Conference on Image Analysis (pp. 745–749). Springer Berlin Heidelberg.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86, 2278–2324.

Li, X., Shang, M., Hao, J., & Yang, Z. (2016). Accelerating fish detection and recognition by sharing CNNs with objectness learning. In OCEANS 2016-Shanghai (pp. 1–5). IEEE.

Lin, T. Y., Dollar, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2117–2125).

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., et al. (2014). Microsoft COCO: Common objects in context. In European Conference on Computer Vision (pp. 740–755). Springer Cham.

doi:10.6342/NTU202403695

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single shot multibox detector. In European Conference on Computer Vision (pp. 21–37). Springer Cham.

Wei, J. C., & Chu, Y. N. (2021). White shrimp's appetite recognition based on deep learning using underwater image. *Master's thesis*, National Taiwan University.

Lu, Y. C., Tung, C., & Kuo, Y. F. (2019). Identifying the species of harvested tuna and billfish using deep convolutional neural networks. ICES Journal of Marine Science. doi:10.1093/icesjms/fsz089.

Manning, C. D., Manning, C. D., & Schütze, H. (1999). Foundations of Statistical Natural Language Processing. MIT Press, Cambridge, MA.

Monkman, G. G., Hyder, K., Kaiser, M. J., & Vidal, F. P. (2019). Using machine vision to estimate fish length from images using regional convolutional neural networks. Methods in Ecology and Evolution, 10, 2045–2056.

Morais, E. F., Campos, M. F. M., Padua, F. L., & Carceroni, R. L. (2005). Particle filter-based predictive tracking for robust fish counting. In XVIII Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI'05) (pp. 367–374). IEEE.

Needle, C. L., Dinsdale, R., Buch, T. B., Catarino, R. M., Drewery, J., & Butler, N. (2015). Scottish science applications of remote electronic monitoring. ICES Journal of Marine Science, 72, 1214–1229.

Qin, H., Li, X., Liang, J., Peng, Y., & Zhang, C. (2016). DeepFish: Accurate underwater live fish recognition with a deep architecture. Neurocomputing, 187, 49–58.

doi:10.6342/NTU202403695

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 779–788).

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. Advances in Neural Information Processing Systems, 2015, 91–99.

Russell, B. C., Torralba, A., Murphy, K. P., & Freeman, W. T. (2008). LabelMe: A database and web-based tool for image annotation. International Journal of Computer Vision, 77, 157–173.

Shafry, M. R. M., Rehman, A., Kumoi, R., Abdullah, N., & Saba, T. (2012). FiLeDI framework for measuring fish length from digital images. International Journal of Physical Sciences, 7, 607–618.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Spampinato, C., Chen-Burger, Y. H., Nadarajan, G., & Fisher, R. B. (2008). Detecting, tracking, and counting fish in low-quality unconstrained underwater videos. VISAPP 2008, 1.

Student. (1908). The probable error of a mean. *Biometrika*, 6(1), 1-25.

doi:10.6342/NTU202403695