Classifying and Sizing Marine Clams Using Imaging and Machine Learning Techniques

1st Wenyang Lyu
Dept. of Electrical & Electronic Engineering
Auckland University of Technology
Auckland, New Zealand
sxb7657@autuni.ac.nz

2nd David I. Wilson
Industrial Information & Control Center
Auckland University of Technology
Auckland, New Zealand
david.wilson@aut.ac.nz

***Abstract*—Classifying and the subsequent sizing of marine clams is a legislative requirement for the processing of shellfish in New Zealand. For economic reasons, this needs to be efficient, robust, and ideally automated. This work proposes combining traditional image processing augmented with machine learning, to both identify individual clam species, and then capture key morphological and sizes. The results show that a combination of ResNet-50 with a polynomial SVM delivers the best classifying performance across varying lighting conditions. For sizing, the proposed method also addresses the challenge of overlapping clams by incorporating advanced contour reconstruction techniques. The findings indicate that the proposed approach not only improves classification accuracy but also ensures precise sizing, making it a reliable tool for real-world applications.**

***Index Terms*—Clam Classification, Clam Sizing, Image Processing, Machine Learning, Deep Learning**

I. Introduction

New Zealand's coastal ecosystems play a crucial role in maintaining the nation's biodiversity and supporting its fishing economy. Cloudy Bay Clams, located on the northeastern coast of New Zealand's South Island, like any commercial fishing interest, must periodically survey their catchment for the Ministry of Fisheries. However, traditional methods for classifying and sizing these shellfish are labor-intensive and prone to human error.

While current image classification techniques are well-developed and diverse, feature extraction and classifier design often depend on the specific application scenario. In the context of clams, which possess unique texture patterns and contours, few existing studies have thoroughly explored the methods that achieve the highest classification accuracy. This study addresses this gap by investigating the most effective methods for classifying, and subsequently sizing, marine clams. The four main commercially important species found in Cloudy Bay, New Zealand, are Diamond Shell (*Spisula aequilatera*), Storm Clam (*Mactra murchisoni*), Tuatua (*Paphies donacina*), and Moon Shell (*Dosinia anus*). This study utilizes four similar types of empty shells, as shown in Fig. 5, commonly found locally for convenience: 288 Cockles (*Austrovenus stutchburyi*) from Cockle Bay, 96 Dosinia (*Dosinia anus*) and 192 Tuatua (*Paphies subtriangulata*) from Orewa Beach, and 324 Mussels (*Perna canaliculus*) from Point England Beach. The mussels were added as a deliberate 'rogue' species for testing.

This study proposes an automated system that uses a Support Vector Machine (SVM) classifier to categorize clams based on extracted morphological and texture features. Additionally, the research explores the use of transfer learning with neural networks such as AlexNet, ResNet-50, and NASNet-Large to enhance classification performance. The system measures the Feret diameters of the clams using camera calibration techniques not only improving accuracy but also streamlining the processing of large batches of clams, making it a valuable tool for both commercial and ecological survey applications.

Moreover, this study goes beyond the conventional approach of testing classifiers under identical lighting conditions by evaluating the robustness of classification accuracy across varying lighting environments. Unlike traditional studies where training and testing are conducted under the same lighting conditions, this research tests classifiers under different lighting conditions, revealing significant performance differences among them, as shown in Table II. The study identifies the optimal classifier and features combination that achieves the highest accuracy across diverse lighting conditions, offering a more reliable solution for real-world applications.

In terms of sizing, the study employs a scaling factor derived from camera calibration to enhance measurement accuracy, resulting in a significant reduction in RMSE. The research also addresses the challenge of overlapping clams by incorporating advanced contour reconstruction techniques, further improving the system's accuracy and reliability.

The contributions of this research are threefold: 1) it provides a robust methodology for automating clam classification and sizing, 2) it offers a scalable solution for improving efficiency in clam harvesting operations, and 3) it advances the application of computer vision techniques in marine resource management.

The remainder of this paper is organized as follows: Section II reviews related work on clam classification and sizing techniques. Section III describes the proposed methodology, including feature extraction, classification, and sizing processes. Section IV presents the experimental results and analysis. Finally, Section V concludes the paper and discusses potential future directions for this research.

II. Literature Review

The classification and sizing of clams are required by legislation, but is labor-intensive and time-consuming tasks prone to human error.

A. Automated Classification Techniques

Machine learning techniques, particularly Support Vector Machines (SVM) and Convolutional Neural Networks (CNNs), have been widely adopted for the classification of marine species, including clams. SVMs have proven effective in applications where feature extraction from images is essential, achieving reasonable accuracy in distinguishing species based on morphological features [1]. Techniques like Gray-Level Co-Occurrence Matrix (GLCM) and Discrete Cosine Transform (DCT) have been utilized for texture analysis, aiding in differentiating species with similar shapes but distinct surface patterns [2], [3]. CNNs, which automatically learn features from large datasets, have surpassed traditional methods in both accuracy and robustness, making them highly effective for complex classification tasks [4]. However, CNNs require substantial computational resources and large training datasets, which can be limiting in specific applications.

B. Sizing Methods in Marine Species

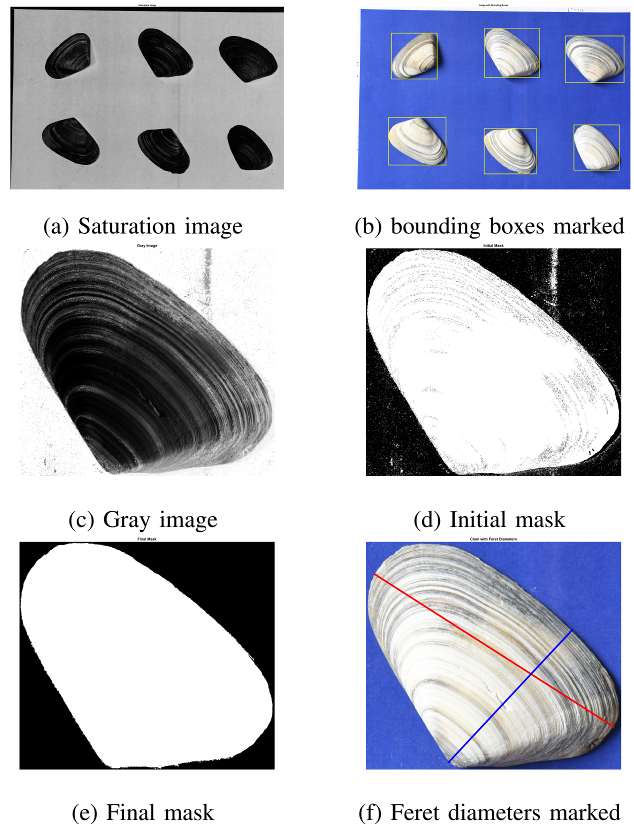
Imaging-based methods, including single and stereo camera systems, have been explored for measuring the size of marine species. Single-camera systems are cost-effective and straightforward but are limited to 2D measurements and can be affected by environmental conditions [5]. Stereo camera systems, on the other hand, provide 3D measurements and offer higher accuracy, though at the cost of increased complexity and calibration requirements [6]. These methods have been successfully applied in various contexts, such as fruit sizing and fish length estimation, but their application to clam sizing requires further refinement to address challenges such as irregular shell shapes and environmental variability [7], [8].

While significant progress has been made in classification and sizing methods, existing systems often treat these tasks separately, leading to inefficiencies. Furthermore, the robustness of these systems under varying environmental conditions remains an open challenge. This study addresses these gaps by integrating classification and sizing into a single automated system. By leveraging advanced feature extraction, including GLCM and DCT, alongside robust camera calibration techniques, our approach enhances both the accuracy and efficiency of clam processing, contributing to more effective resource management in the marine industry.

III. Classification Method

This study explores three main classification approaches: traditional feature extraction combined with classifiers, CNN-based feature extraction combined with classifiers, and transfer learning using pre-trained deep learning networks.

Before applying these classification methods, images are auto-cropped to isolate each clam and normalized for feature extraction. Fig. 1 demonstrates the process: first, the images are filtered by extracting the saturation channel from the HSV color space, which clearly distinguishes the clams from the background. Bounding boxes are then detected using MATLAB's `vision.BlobAnalysis` function, which locates the clam regions. To ensure that the clam edges are fully captured, 10 pixels are added to each side of the bounding box before cropping. Finally, the images are normalized to the same size, preparing them for feature extraction.



(a) Saturation image (b) bounding boxes marked

(c) Gray image (d) Initial mask

(e) Final mask (f) Feret diameters marked

Fig. 1: Image processing steps for clam Feret diameters calculation.

In order to obtain the Feret diameters, which are used for both classification and sizing, the cropped image is first converted to grayscale and binarized to separate the clam from the background. Next, `imfill` is applied to fill any holes within the clam region, and `bwareafilt` is used to isolate the largest blob, ensuring only the clam is selected. Finally, the Feret diameters are measured using `regionprops`, which calculates the maximum and minimum pixel distances, providing the necessary dimensions for classification and sizing.

A. Traditional Features Method

Inspired by related classification works, although not specifically focused on clams, this study tests three different types of features.

1) Morphological features: Morphological features describe the shape and structure of the clams, helping to distinguish species based on size and shape. In this study, the four key morphological features are the maximum and minimum Feret diameters, which is the longest and shortest distances between points along the clam's boundary, as shown in Fig. 1, their ratio, which reflects the elongation or roundness of the clam, and the total area size, representing the overall pixel area enclosed by the clam's boundary. Fig. 2 illustrates the distribution of the Min/Max Feret diameter features across various clam species. The combination of these four morphological features allows for effective differentiation between species.

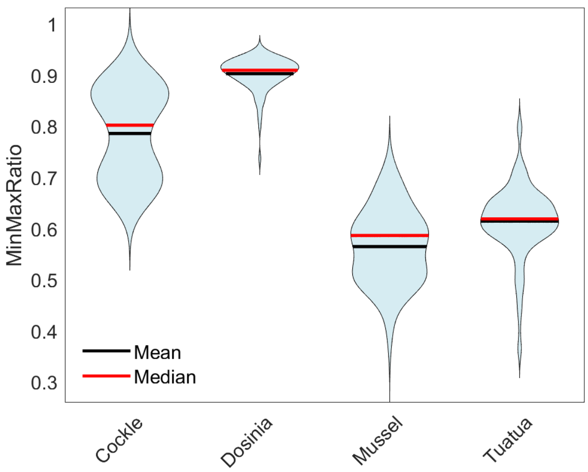


Fig. 2: The distribution of a morphological feature, (min/max Feret diameter), across different clam species.

2) Texture features: Texture features are extracted using the Gray-Level Co-Occurrence Matrix (GLCM), which captures the spatial relationship between pixel intensities in the image. The GLCM features used in this study include contrast, correlation, energy, homogeneity, entropy, dissimilarity, variance, and inverse difference moment, each providing insight into various aspects of texture such as intensity variation, uniformity, and smoothness. These features offer a detailed representation of the clam surface texture, making them valuable for distinguishing species with similar shapes but different surface patterns.

3) Frequency features: Frequency features are extracted using the Discrete Cosine Transform (DCT), which converts spatial data into frequency data. This study extracts the first 9 features from the DCT coefficients within a block size of 3×3. These features capture variations in the clam's texture and surface patterns at different frequency levels, allowing for the identification of subtle differences between species.

4) Selection of Classifiers Parameters: After extracting pertinent features using both traditional and CNN-based methods, we evaluated 3 classifiers: a support vector machine, (SVM), with various kernel functions such as linear, polynomial, and radial basis function (RBF) kernels, k-Nearest Neighbors, (k-NN), with varying k and distance metrics, and a random forest, (RF), with varying numbers of trees and maximum depths. Table I shows the classification accuracy of frequency features with different selection of classifiers parameters as an example. This study first determines the optimal classifier parameter settings for each type of feature. Subsequently, the overall performance of different classifiers using combined features is compared, as shown in Table II.

TABLE I: Accuracy of Different Classifiers and Parameters Using DCT Features

Classifier	Parameters Configuration	Accuracy (%)
SVM	Linear Kernel	98.23
	Radial Basis Function (Fixed Scale)	94.38
	Radial Basis Function (Auto Scale)	98.86
	Polynomial Kernel	98.65
k-NN	1-Nearest Neighbor (Euclidean Distance)	98.86
	3-Nearest Neighbors (Cityblock Distance)	98.96
	5-Nearest Neighbors (Cityblock Distance)	98.75
RF	50 Trees, 1 Minimum Leaf Size	97.92
	100 Trees, 1 Minimum Leaf Size	97.61
	200 Trees, 10 Minimum Leaf Size	96.99

B. CNN-Based Feature Extraction

CNN-based feature extraction, which leverages deep learning models to automatically learn and extract features from raw image data, has seen significant advancements in recent years. This study utilizes three different neural networks for feature extraction: AlexNet, ResNet-50, and NASNet-Large. The selection of these networks balances accuracy and computational complexity. AlexNet, a simpler model, requires less computational power, while NASNet-Large is expected to deliver higher accuracy at the cost of greater computational demands. ResNet-50, known for its robust performance and efficiency, serves as an optimal middle-ground option.

Under consistent lighting conditions, all three networks achieve 100% accuracy. To further challenge their robustness, the networks are tested under varying lighting conditions, both indoors and outdoors, ranging from dim morning light to bright midday sunlight. Fig. 3 illustrates the time taken for each classification method, reflecting both computational complexity and accuracy. The experiments are conducted on a laptop equipped with a 12th Gen Intel(R) Core(TM) i7-1260P 2.10 GHz CPU and 16GB of RAM. Among the tested networks, ResNet-50 consistently demonstrates the best accuracy across varying lighting conditions, making it the most effective choice.

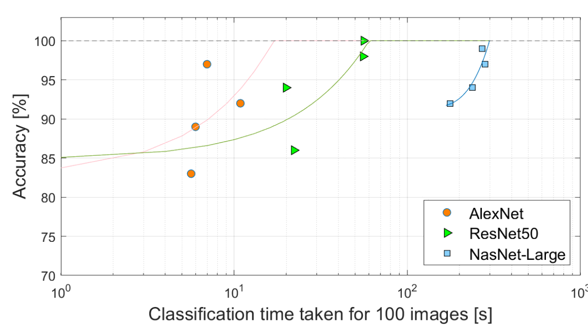


Fig. 3: The computational time and subsequent accuracy for 3 different networks when classifying 100 clam images.

C. Transfer Learning

To adapt these pre-trained networks for our classification task, this study employs transfer learning. This involves freezing the convolutional layers of the network to retain the learned features and adding new fully connected layers tailored to our specific task. A dropout layer is included to prevent overfitting during training. The transfer learning process is outlined as follows:

1) Freeze the layers up to a specific depth, depending on the network (e.g., freezing all layers before ‘fc1000‘ in ResNet-50).
2) Add a new fully connected layer corresponding to the number of clam species.
3) Include a dropout layer to reduce the risk of overfitting.
4) Train the new layers using our dataset while keeping the pre-trained layers frozen.

Despite the advantages of transfer learning, this study found that the CNN-based feature extraction method combined with other classifiers often outperforms the transfer learning approach. This is primarily because traditional classifiers, such as SVM or Random Forest, tend to be more effective for classification tasks compared to the relatively simple neural network layers typically added during transfer learning. The robust features extracted by the pre-trained CNNs are already well-suited to the task, and pairing them with more sophisticated classifiers enhances the classification performance.

D. Enhancing Robustness

In typical classification tasks, training and testing datasets often share similar lighting conditions, which works well in controlled environments but may not reflect real-world scenarios. In practice, varying lighting conditions challenge a classifier's robustness, making its ability to accurately identify clams under different conditions crucial. To assess the robustness of our classifiers, this study tests 100 clams under three lighting conditions: 9am indoor dim lighting, 12pm indoor bright lighting, and 12pm outdoor harsh sunlight. The results of these tests are summarized in Table II, which shows the classification accuracy across different features and classifiers under the varying lighting conditions.

To enhance the robustness of the classification method, various approaches are tested above, along with the use of the following three techniques.

1) Data Augmentation: The classification results vary under different lighting conditions, likely due to differences in brightness and texture or edge clarity. To address this, data augmentation techniques such as hue and saturation jitter, brightness adjustment, contrast changes, synthetic noise, and blur were applied to simulate various conditions and improve model robustness. However, as shown in Table II, these augmentations yielded only a slight improvement, averaging 93% accuracy compared to 94% without augmentation, indicating that they may not fully capture real-world lighting complexities.

2) Incorporating Auxiliary Information: Expert industry knowledge can enhance classification accuracy. In this case, combining Feret morphological features with ResNet50 extracted features achieves an average accuracy of 94.67%, slightly higher than the 94.33% with ResNet50 alone. This marginal improvement suggests that integrating detailed morphological features with deep learning representations provides some benefits.

3) Incorporating an 'Others' Category in Classification: An 'Others' category is introduced to handle non-target objects encountered during clam harvesting. Black pebbles are used as representative non-target items in the training dataset. A confidence threshold of 0.9 is applied, classifying objects with lower recognition scores as 'Others'. Although this may reduce overall accuracy, it significantly improves precision for clam identification by ensuring high confidence in classifications, especially under diverse conditions.

IV. Sizing Methods

Accurate sizing of clams (to $\sim \pm$2mm) is crucial for ecological surveys, and therefore a camera calibration is essential for ensuring this accuracy. In this research, the camera is consistently positioned on a fixed platform with a standardized setup, including fixed camera parameters such as focal length, aperture, and distance to the subject. Since every photo is taken under these identical conditions, a one-time camera calibration specific to this setup is performed. This calibrated set of camera parameters is then applied across all subsequent images, allowing for quick and reliable measurement of the clams. By using the same calibration coefficients for each image, the sizing process is greatly simplified and maintains a high level of accuracy. This follows a similar setup to what is anticipated suitable in industrial conditions.

A. Camera Calibration Methodology

The camera calibration of Zhang's method [9] is performed using 22 images of a 10×7 checkerboard pattern, captured from different angles with a fixed camera position. The calibration process involves standardizing camera settings (aperture: F2.8, shutter speed: 1/60, ISO: auto) and using a square size of 24.46 mm on the checkerboard. Outliers with high reprojection errors are excluded, leading to a final reprojection error of 0.93 pixels. The calibrated intrinsic and extrinsic parameters are then exported for use in subsequent clam sizing tasks.

B. Scaling Factor Logic

To ensure measurement accuracy, this study assumes a fixed camera setup, where both the camera position and parameters remain consistent throughout all image captures. After initial camera calibration, some minor measurement errors are still observed. To address these, a test pattern comprising nine identical circles, printed on paper, is used and placed in positions that closely match those where the clams will be imaged.

The circles are photographed under the same conditions as the clams, and their sizes are inferred. These inferred sizes are compared to the vernier-measured sizes, allowing for the calculation of a scaling factor to correct for measurement distortions. By averaging the measured sizes and comparing them to the actual size, a scaling factor is derived to adjust for systematic distortion. In this study, a scaling factor of $^{40.2}/_{38.417} = 104.64\%$ is applied, where 38.417 mm is the average measured size of the circles.

This approach proves effective in improving measurement accuracy, but there are some limitations. The boundary detection algorithm might not always capture the entire circle

TABLE II: Classification Accuracy Across Different Features and Classifiers

Features	**Classifier**	**# Features**	**Accuracy (%)**			
			Dim Indr	**Bright Indr**	**Bright Otdr**	**Average**
Morphological features (Feret)	KNN	4	41	35	60	45.33
Morphological features (Feret)	Random Forest	4	34	39	70	47.67
Morphological features (Feret)	SVM	4	35	20	44	33.00
Frequency features (DCT)	KNN	9	37	54	35	42.00
Frequency features (DCT)	Random Forest	9	29	36	34	33.00
Frequency features (DCT)	SVM	9	35	54	52	47.00
ResNet50 Extracted Features	KNN	2048	94	94	95	94.33
ResNet50 Extracted Features	Random Forest	2048	91	100	91	94.00
ResNet50 Extracted Features	SVM	2048	94	100	98	**97.33**
ResNet50 Extracted Features	Transfer Learning	2048	95	96	91	94.00
ResNet50 with dataset augmented	Transfer Learning	2048	91	94	94	93.00
AlexNet Extracted Features	SVM	4096	89	97	92	92.67
NasNet Extracted Features	SVM	4032	94	99	97	**96.67**
ResNet50 with Auxiliary features	KNN	2052	94	95	95	94.67
ResNet50 with Auxiliary features	Random Forest	2052	90	99	93	94.00
ResNet50 with Auxiliary features	SVM	2052	94	100	98	**97.33**
Average Accuracy (%)	-	-	69.92	**76.33**	74.42	-

perfectly, introducing errors into the distortion coefficient calculation. Additionally, while efforts are made to maintain a consistent camera position, absolute accuracy cannot be guaranteed. The specific 3D shape of the clam may also introduce additional measurement errors. To further improve accuracy, a more refined approach might involve separate measurements for different clam shapes, considering multiple positions and distortion coefficients for each.

C. Sizing Results

The Feret maximum diameter of clams is used as the primary measure of size. After identifying the Feret max diameter pixel coordinates in the cropped images, these coordinates are mapped back to the original image using the calibrated camera parameters. The real-world size of the clams is then calculated by converting the pixel measurements into millimeters, using the scaling factor to correct any distortions.

The effectiveness of the scaling factor is validated by comparing the calculated clam sizes to those measured with a vernier caliper. The scaled sizes significantly reduce the average absolute error rate from 4.21% to 2.14%, as illustrated in Fig. 4. This improvement demonstrates the importance of the scaling factor in achieving accurate clam measurements. The Root Mean Square Error (RMSE) for the original sizes is 3.3521 mm, which is reduced to 1.3868 mm after applying the scaling factor.

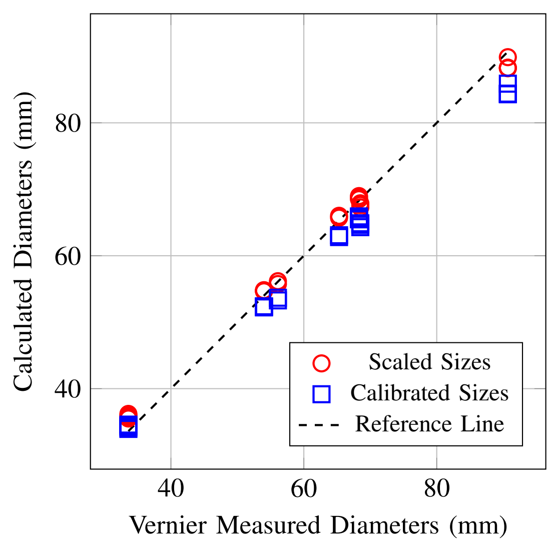


Fig. 4: A comparison of the Vernier measured diameters versus the scaled sizes, and the original calculated sizes

D. Addressing Overlapping Issues

Overlapping clams present a significant challenge in accurate sizing. Previous research in clam contour reconstruction using Generative Adversarial Networks (GANs) [10] provides a foundation for this study. In this work, we address the specific issue of overlapping clams by using a GAN-based approach to reconstruct obscured portions of clam contours, ensuring accurate measurements even in complex scenarios. This method enhances previous techniques by incorporating image inpainting to remove seaweed, pebbles, and other debris, thereby recovering the complete clam shape. A modified channel-wise encoder-decoder network serves as the generator, while a convolutional network acts as the discriminator. The joint loss function combines weighted Mean Square Error (MSE) for overlapping and missing regions and Binary Cross-Entropy (BCE) for the discriminator. Testing was conducted on clams obscured by 12.5% to 50%, with performance evaluated using Mean Absolute Error (MAE), Mean Square Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural

Similarity Index Measure (SSIM). Results showed significant improvements compared to the baseline model, offering robust support for rapid reconstruction of obscured clams in fisheries, ultimately facilitating size measurement and further processing.

V. Conclusion

This study presents an effective methodology for automated clam classification and sizing as shown in Fig. 5, leveraging a combination of traditional image processing techniques, machine learning classifiers, and advanced neural networks. The research highlights the importance of accurate camera calibration and the application of a scaling factor to correct measurement distortions, resulting in significant improvements in sizing accuracy.

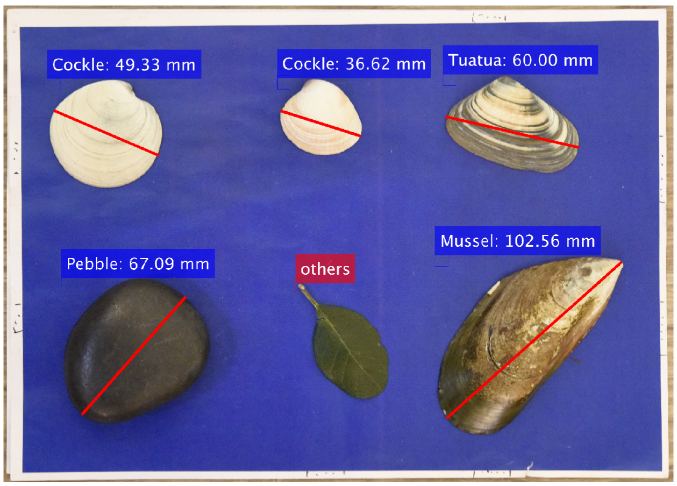


Fig. 5: Demo image of automated classifying and sizing of all clams in dim indoor lighting condition.

In the classification task, the comparison between traditional feature extraction methods and CNN-based approaches demonstrates the superior performance of neural networks, particularly ResNet-50, when paired with an SVM classifier using a polynomial kernel. The robustness of this approach is further validated under varying lighting conditions, where ResNet-50 consistently outperforms other methods.

Additionally, the study explores the incorporation of auxiliary information, such as Feret morphological features, and assesses the impact of data augmentation on model robustness. Although these techniques yield incremental improvements, the findings suggest that CNN-based feature extraction combined with robust classifiers is the most effective strategy for clam classification.

Overall, the proposed system demonstrates a reliable and efficient solution for clam classification and sizing, with potential applications in commercial and ecological monitoring. Future work will focus on refining the classification models, enhancing the robustness of the system under diverse environmental conditions, and fully automating the sizing process to further improve accuracy and efficiency.

References

[1] D. I. Wilson, "Shape factor characterising for clams," Industrial Information & Control Center, AUT University, Tech. Rep., Oct. 2020, Internal Publication.

[2] A. N. N. Chamim, A. Fatullah, and Y. Jusman, "Classification of Ettawa crossbreed goats based on face using gray level co-occurrence matrix (GLCM) and support vector machines (SVM) methods," in *2023 International Workshop on Artificial Intelligence and Image Processing (IWAIIP)*, 2023, pp. 184–188. DOI: 10.1109/IWAIIP58158.2023.10462844.

[3] M. C. Yesilli, J. Chen, F. A. Khasawneh, and Y. Guo, "Automated surface texture analysis via discrete cosine transform and discrete wavelet transform," *Precision Engineering*, vol. 77, pp. 141–152, 2022, ISSN: 0141-6359. DOI: https://doi.org/10.1016/j.precisioneng.2022.05.006.

[4] E. Kim, S. Yang, J. Cha, D. Jung, and H. Kim, "Deep learning-based phenotype classification of three ark shells," English, *Frontiers in Marine Science*, vol. 11, 2024, ISSN: 2296-7745. DOI: 10.3389/fmars.2024.1356356.

[5] G. Bortolotti, M. Piani, M. Gullino, *et al.*, "A computer vision system for apple fruit sizing by means of low-cost depth camera and neural network application," *Precision Agriculture*, pp. 1–18, Apr. 2024. DOI: 10.1007/s11119-024-10139-8.

[6] N. Tonachella, A. Martini, M. Martinoli, *et al.*, "An affordable and easy-to-use tool for automatic fish length and weight estimation in mariculture," *Scientific Reports*, vol. 12, p. 15 642, Sep. 2022. DOI: 10.1038/s41598-022-19932-9.

[7] Y. Lin, P. Wang, Z. Wang, S. Ali, and L. Mihaylova, "Towards automated remote sizing and hot steel manufacturing with image registration and fusion," *Journal of Intelligent Manufacturing*, pp. 1–18, Nov. 2023. DOI: 10.1007/s10845-023-02251-9.

[8] G. Monkman, K. Hyder, M. Kaiserc, and F. Vidal, "Using machine vision to estimate fish length from images using regional convolutional neural networks," *Methods in Ecology and Evolution*, vol. 10, Aug. 2019. DOI: 10.1111/2041-210x.13282.

[9] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000. DOI: 10.1109/34.888718.

[10] W. Lyu, "Classifying and sizing marine clams using imaging and machine learning techniques," Unpublished master's thesis, in preparation, M.S. thesis, Auckland University of Technology, Aug. 2024.