

1 A NON-INVASIVE SEX IDENTIFICATION OF BLOOD
2 COCKLES TEGILLARCA GRANOSA (LINNAEUS, 1758)
3 USING MACHINE LEARNING

4 A Special Problem Proposal
5 Presented to
6 the Faculty of the Division of Physical Sciences and Mathematics
7 College of Arts and Sciences
8 University of the Philippines Visayas
9 Miag-ao, Iloilo

10 In Partial Fulfillment
11 of the Requirements for the Degree of
12 Bachelor of Science in Computer Science by

13 ADRICULA, Briana Jade
14 PAJARILLA, Gliezel Ann
15 VITO, Ma. Christina Kane

16 Francis DIMZON
17 Adviser
18 Victor Marco Emmanuel FERRIOLS
19 Co-Adviser

20 October 28, 2024

Abstract

22 From 150 to 200 words of short, direct and complete sentences, the abstract should
23 be informative enough to serve as a substitute for reading the entire SP document
24 itself. It states the rationale and the objectives of the research. In the final Special
25 Problem document (i.e., the document you'll submit for your final defense), the
26 abstract should also contain a description of your research results, findings, and
27 contribution(s).

28 Suggested keywords based on ACM Computing Classification system can be
29 found at https://dl.acm.org/ccs/ccs_flat.cfm

30 **Keywords:** Keyword 1, keyword 2, keyword 3, keyword 4, etc.

Contents

32	1 Introduction	1
33	1.1 Overview	1
34	1.2 Problem Statement	2
35	1.3 Research Objectives	3
36	1.3.1 General Objective	3
37	1.3.2 Specific Objectives	3
38	1.4 Scope and Limitations of the Research	4
39	1.5 Significance of the Research	4
40	2 Review of Related Literature	6
41	2.1 Background on <i>Tegillarca granosa</i> and Their Importance	7
42	2.2 Current Methods of Sex Identification in <i>Tegillarca granosa</i>	9
43	2.3 Machine Learning and Deep Learning in Biological Studies	11
44	2.3.1 Deep Learning for Phenotype Classification in Ark Shells	11
45	2.3.2 Geometric Morphometrics and Machine Learning for Species	
46	Delimitation	12
47	2.3.3 Contour Analysis in Mollusc Shells Using Machine Learning	12
48	2.3.4 Machine Learning for Shape Analysis of Marine Organisms	13

49	2.3.5	Deep Learning for Landmark-Free Morphological Feature	
50		Extraction	14
51	2.3.6	Machine Learning for Sex Differentiation in Abalone . . .	14
52	2.3.7	Machine Learning for Geographical Traceability in Bivalves	15
53	2.4	Limitations on Sex Identification in <i>Tegillarca granosa</i>	16
54	2.5	Synthesis of the Study	17
55	3	Research Methodology	20
56	3.1	Sample Collection	20
57	3.2	Ethical Considerations	21
58	3.3	Creating <i>T. granosa</i> Dataset	21
59	3.4	Morphological Characteristics Collection	21
60	3.5	Image Acquisition and Pre-Processing	22
61	3.6	Machine/ Deep Learning Technologies	23
62	3.7	Support Vector Machine for Pre-evaluation	24
63	3.8	Deep Learning for Image-Based Classification	24
64	4	Preliminary Results/System Prototype	26
65		References	27
66	A	Appendix Title	31
67	B	Resource Persons	32

68 List of Figures

69	2.1 Dorsal view of <i>Tegillarca granosa</i> shell.	7
70	3.1 Length	22
71	3.2 Width	22
72	3.3 Height	22
73	3.4 <i>T. granosa's</i> measurements	22

⁷⁴ List of Tables

Chapter 1

Introduction

1.1 Overview

The Philippines is a global center of marine biodiversity and has established aquaculture as a significant contributor to total fishery production (Aypa & Bacongus, 2000; BFAR, 2019). As the 11th largest seafood producer in the world, the country produces over 4 million tonnes of seafood annually. Aquaculture is deeply integrated into Filipinos' livelihoods, encompassing fish cultivation and the production of various aquatic species, including mollusks. Among these are blood clams (*Tegillarca granosa*) which hold considerable economic and environmental significance.

Maintaining a balanced male-to-female ratio of blood cockles is crucial to prevent overharvesting and ensure sustainable production because an imbalanced ratio can lead to overexploitation and can impact the population's sustainability. However, there is limited literature on *T. granosa* that has a thorough understanding of its sex-determining mechanisms, particularly concerning sexual dimorphism in morphological and morphometric characteristics (Breton, Capt, Guerra, & Stewart, 2017).

Currently, sex determination methods for blood cockles are invasive, including dissection, and histological examinations which often result in the death of the specimens. While there is growing literature on aquaculture commodities sex identification using machine learning and deep learning, there is a notable scarcity of research specifically addressing *T. granosa* (Miranda & Ferriols, 2023).

This study, titled "A Non-Invasive Sex Identification of *T. granosa* using Ma-

chine Learning,” aims to provide a comprehensive analysis of blood cockles by leveraging their morphological and morphometric characteristics. By integrating machine learning and computer vision techniques, the study seeks to identify distinct features that indicate sexual dimorphism between male and female blood cockles.

1.2 Problem Statement

Accurately identifying the sex of *T. granosa* is important in order to promote sustainable aquaculture and biodiversity by maintaining a balanced male-to-female ratio. A balanced ratio helps prevent overharvesting. Although sex identification is important for blood cockle population management and sustainable aquaculture, there is a notable lack of research in creating non-invasive methods to identify the sex of *T. granosa*. Many of the latest studies and approaches are based on invasive methods like dissection or histological analysis, which are impractical for large-scale aquaculture operations focused on conservation.

The existing invasive methods for identifying the sex of *T. granosa* often require dissection, a technique that involves cutting open the shell to visually inspect the gonads (Erica, 2018). This causes harm and death to the specimens. In some cases, histological examination is used to examine tissue samples through a microscope, leading to further destruction of the organism (May, Maung, Phyu, & Tun, 2021). These methods are time-consuming, labor-intensive, and can pose a threat to population management, especially when it is essential to maintain a balanced sex ratio for breeding programs. Moreover, invasive methods also require technical skills to execute properly. Aquaculture operations, particularly in resource-limited settings, face challenges in accessing laboratory equipment like microscopes and staining tools which complicates the process.

A less invasive approach employed by aquaculturists is to monitor spawning behavior in which individuals are separated and stimulated to reproduce in order to determine their sex through the release of gametes (Miranda & Ferriols, 2023). Although it is indeed less invasive than dissection, spawning still involves inducing stress in blood cockles and may not be completely effective for fast identification in large populations.

Given the limitations of both invasive and less invasive methods highlight the need for a more advanced approach. An alternative, non-invasive method involving machine and deep learning technologies might solve these issues by providing a fast, accurate, and effective solution without harming or stressing the blood

134 cockles.

135 1.3 Research Objectives

136 1.3.1 General Objective

137 The general objective of this study is to develop a non-invasive method for iden-
138 tifying the sex of *Tegillarca granosa* using machine and deep learning integrated
139 with computer vision technologies. This method aims to provide accurate and
140 streamlined sex identification without causing harm to the specimens, thus sup-
141 porting sustainable aquaculture practices.

142 1.3.2 Specific Objectives

143 To achieve the general objective of developing a non-invasive sex identification of
144 *T. granosa* using machine and deep learning, the following specific objectives have
145 been established:

- 146 1. To collect and organize a comprehensive dataset of *T. granosa* which will
147 include high-quality images and relevant morphological measurements that
148 will serve as the basis for the machine-learning model.
- 149 2. To preprocess the collected data to perform quality control and consistency
150 checks. This will include techniques such as color thresholding, segmenta-
151 tion, and image hole filling and dilating.
- 152 3. To develop and implement machine learning models that can classify the sex
153 of *T. granosa* based on the collected dataset, implementing algorithms such
154 as support vector machines (SVM) for pre-evaluation, and deep learning
155 models such as Squeezenet and Unet.
- 156 4. To evaluate the performance of the models used using performance metrics
157 such as accuracy, precision, recall, and F1-score to ensure the effectiveness
158 and reliability of the models.
- 159 5. To compare the developed models against existing methods, such as dissec-
160 tion and spawning, and assess their potential for real-world application in
161 aquaculture operations.

1.4 Scope and Limitations of the Research

This study focuses on developing a non-invasive method for identifying the sex of *Tegillarca granosa* using machine learning, deep learning, and computer vision technologies. The goal is to provide an accurate and efficient means of sex identification without causing harm to the specimens, contributing to sustainable aquaculture practices.

The researchers will work with 500 spawned blood cockles taken from Panay island, specifically Zarraga Iloilo and Ivisan Capiz, equally divided between 250 males and 250 females, obtained through temperature shock. The researchers will personally gather linear measurements, including length, width, height, rib count, length of the hinge line, and distance between the umbos using the vernier caliper. Images and corresponding views of the specimens will also be collected by the researchers under the supervision of the University Researchers Associate from the Institute of Aquaculture, College of Fisheries and Ocean Sciences.

Data collection will take place at the hatchery facility of the University of the Philippines Visayas. Data gathering will be conducted in batches, depending on the availability of spawned samples.

The method developed in this study is specific to *Tegillarca granosa* and may not be generalized to other species. The model is trained exclusively for *Tegillarca granosa* and morphological features including length, width, height, rib count, length of the hinge line, and distance between the umbos may not be shared by other shellfish species.

1.5 Significance of the Research

This study will give us significant advancement in non-invasive sex identification methods in *T. granosa* providing innovative solutions that could solve the challenges in identifying sex and reshape approaches to aquaculture. The significance of this study extends to the following:

Research Institution. The result of this study focusing on the sex-identification mechanism of bivalves, specifically *Tegillarca granosa*, will provide valuable insights into universities and research centers that focus on fisheries and coastal management such as the UPV Institute of Agriculture that aim to develop sustainable development and develop suitable culture techniques.

194 *Fishermen.* By developing a non-invasive method in sex identification, this
195 study can help long-term harvest efficiency and maintain the ratio of the harvest
196 which can help prevent overexploitation of the *T. granosa*.

197 *Coastal Communities.* The result of this study would be beneficial for the
198 coastal communities that are reliant on their source of income with aquaculture
199 commodities like blood cockles. Maintaining the diversity and aspect ratio of
200 male and female may increase the market value of blood cockle production since
201 cockle aquaculture faces significant obstacles worldwide due to the fluctuating
202 seed supplies and scarcity of broodstock from the wild.

203 *Future Researchers.* The result of this study would serve as the basis for studies
204 that involve sex identification in bivalves such as *T. granosa*. Some technologies
205 are yet to be explored in machine learning, deep learning, and computer vision
206 technologies that can lead to higher accuracy and distinguish the presence of
207 sexual dimorphism in the *T. granosa*.

Chapter 2

Review of Related Literature

Aquaculture is the fastest-growing industry in animal food production and has great potential as a sustainable solution to global food security, nutrition, and development (*FAO 2024 Report: Sustainable Aquatic Food Systems Important for Global Food Security – European Fishmeal*, 2024). Aquaculture is deeply integrated into the livelihoods of Filipinos, not only through fish cultivation but also through the production of other aquatic species, including mollusks, oysters, clams, scallops, and mussels (Breton et al., 2017). Mollusks, particularly blood clams *Tegillarca granosa*, have economic and environmental significance. It has been a collective effort to maintain an ideal male-to-female ratio to avoid overharvesting and maintain the optimal ratio to preserve the population and production of the blood cockles.

The members of the Arcidae Family, including *T. granosa* are important sources of food and livelihood. Cockle aquaculture meets rising demands, however, it faces significant challenges due to fluctuating seed supplies (Miranda & Ferriols, 2023). To solve the problem, researchers exert a considerable amount of effort, developing a broader understanding of bivalves including their sex-determining mechanism due to their notable importance in terms of diversity, environmental benefits, and economic and market importance (Breton et al., 2017). Despite the promising idea of identifying sex, there is limited research reported in terms of sexual dimorphism, making it harder to distinguish through its morphological and morphometric characteristics.

By addressing the challenges in the sex identification of *T. granosa*, it would be able to address one problem at a time. Currently, no recent documented publications that integrate machine learning and computer vision in characterizing sexual dimorphism, reducing complexity, variability in sex determination, and

235 differentiation mechanisms in bivalves, including *T. granosa* specifically.

236 **2.1 Background on *Tegillarca granosa* and Their** 237 **Importance**

238 *Tegillarca granosa* (Linnaeus, 1758) is also known as blood cockles or blood clam.
239 In the Philippines, it is commonly known as a Litob, a marine bivalve species from
240 the family Arcidae. Litob is widely distributed in the world including Southeast
241 Asia. They can be found in the intertidal mudflats adjacent to the mangrove
242 forest (Srisunont, Nobpakhun, Yamalee, & Srisunont, 2020).



Figure 2.1: Dorsal view of *Tegillarca granosa* shell.

243 *T. granosa* shell is medium-sized, fairly thick, ovate, and convex with both
244 valves being equal in size but asymmetrical from the hinge. The top edge of the
245 dorsal margin is straight while the front is rounded and slopes downward with
246 its back being obliquely rounded with a concave bottom edge. It has a narrow
247 diamond-shaped ligament near the hinge with 3-4 dark chevron markings although
248 some may be incomplete. The shell's outer layer or the periostracum is smooth
249 and brown with a straight hinge line and 40-68 fine short teeth arranged in a
250 straight line. The beak or the prosogyrate curves forward with the shell having
251 18-21 raised ribs with blunt nodules, having spaces between them. The inner
252 shell is white with crenulations along the valves' ventral, anterior, and posterior
253 margins. The posterior adductor scar is elongated and squarish while the anterior
254 adductor scar is similar but smaller in size. The mantle covering the bulk of *T.*
255 *granosa's* visceral mass is thin but the edges are thick and muscular. It bears the
256 impression of the crenulated shell edges. Their foot is large with a ventral grove
257 with no byssus or thread-like attachment. The *T. granosa's* soft body is blood
258 red (Narasimham, 1988).

259 *T. granosa* is one of the most well-known marine bivalves given that they are
 260 a protein-rich food, known for their rich flavor, substantial nutritional benefits, a
 261 good source of vitamins, low in fat, and contains a considerable amount of iron,
 262 important in combating anemia (Zha et al., 2022). Blood cockles were collected
 263 by locals inhabiting the brackish mudflats during the low tides for consumption
 264 and sold in the market as a source of livelihood (Miranda & Ferriols, 2023). *T.*
 265 *granosa* is not only valuable for its market and food purposes, but also facilitates
 266 an important role in marine ecosystems as a food source for various organisms
 267 like wading birds, intertidal-feeding fish, and crustaceans such as shore crabs and
 268 shrimps (Burdon, Callaway, Elliott, Smith, & Wither, 2014). Blood cockles can
 269 act as sentinel species and a bioindicator of marine pollutants such as heavy metals
 270 (Ishak, Mohamad, Soo, & Hamid, 2016) and polycyclic aromatic hydrocarbons
 271 (PAHs) (Sany et al., 2014). Additionally, cockle shells can be utilized to create a
 272 cost-effective catalyst for biodiesel production by providing calcium oxide (Boey,
 273 Maniam, Hamid, & Ali, 2011).

274 Determining the sex of bivalves is important for three reasons namely: di-
 275 versity, environmental benefits, and economic significance (Breton et al., 2010).
 276 Firstly, with the estimated 25, 000 living species under class Bivalvia, it would
 277 be a suitable resource to develop a broader understanding of their evolution of
 278 the sex and sex determination mechanism (Breton et al., 2010). Second, study-
 279 ing sex determination is important since bivalves are utilized as bioindicators of
 280 environmental health. This would pave the way for understanding bivalves' life
 281 cycle and population dynamics in determining different factors that affect them
 282 (Campos, Tedesco, Vasconcelos, & Cristobal, 2012). Thirdly, the immediate and
 283 practical reason to unveil the sex determination mechanism is the economic and
 284 nutritional importance of bivalves as a large population of people rely on fish and
 285 shellfish as sources of food and nutrition (Naylor et al., 2000). Additionally, male
 286 and female aquaculture commodities have different growth and economic values.
 287 Male Nile tilapia, for example, grow faster and have lower feed conversion rates
 288 than females, female Kuruma prawns (*Penaeus japonicus*) are generally larger
 289 than males at the time of harvest (Budd, Banh, Domingos, & Jerry, 2015).

290 Clearly, much more work is required to understand the mechanisms under-
 291 lying sexual dimorphism in bivalves, specifically *T. granosa*. Just like the other
 292 aquaculture commodities, sex affects not just reproduction but it can affect mar-
 293 ket preference, and underlying economic value, making the determination of sex
 294 important for meeting consumer demands. These are the increasing significance
 295 of the *T. granosa* despite the lack of reviewed articles in the Philippines.

2.2 Current Methods of Sex Identification in *Tegillarca granosa*

The current sex identification methods in *Tegillarca granosa* range from invasive histological techniques to less invasive methodologies like temperature-induced spawning. Each approach comes with its pros and cons regarding accuracy, feasibility, and impact on natural populations. Induced spawning and larval rearing are considered as the less invasive techniques used to study *Tegillarca granosa*. In the Philippines, limited research has been done on the *Tegillarca granosa* (Linnaeus, 1758), and this study, titled Initial Attempts on Spawning and Larval Rearing of the Blood Cockle, *Tegillarca granosa* in the Philippines, is conducted by Denise Vergara Miranda and Victor Marco Emmanuel Nuestro Ferriols (2023). The researchers conducted experiments on induced spawning and larval rearing, discovering that the eggs of female *T. granosa* were salmon pink, while the sperm released by males looked milky. After spawning, the researchers successfully generated 6, 531, 000 fertilized eggs.

They highlighted the importance of *T. granosa* and other anadarinids as a food source that was established worldwide, especially in Malaysia and Korea. However, in the Philippines, the bivalve aquaculture of the clam species is still limited. The experiment which focuses on the culture and rearing of *T. granosa* was attempted by subjecting the wild broodstocks to a series of temperature fluctuations to induce the spawning of gametes. This is currently the most natural and least invasive method for bivalves (Aji, 2011). The study of Miranda and Ferriols aimed to pave the way to the sustainable production of *T. granosa* seeds for aquaculture production and stock enhancement despite the scarcity of documented hatchery culture of *T. granosa* from larvae to adults that is available in the Philippines.

In the study entitled, The earliest example of sexual dimorphism in bivalves — evidence from the astartid *Nicaniella* (Lower Jurassic, southern Germany), the researchers utilized Principal Component Analysis and Fourier Analysis as a non-invasive method that investigates sexual expression in the *Nicaniella rakoveci*. In the study, researchers discovered that the bivalves with crenulations were found to have a different shell shape, which made them more inflated than those without crenulations. This suggests that when they became females, they adapted to hold more eggs, rather than for protection from predators as previously thought. The formation of crenulations is likely part of the genetic process that controls both the sex change and the changes in shell structure (Karapunar, Werner, Fürsich, & Nützel, 2021). Overall, the findings demonstrate that the genetic mechanisms for sex change and shell morphology in bivalves existed as early as the Early

334 Jurassic, contributing to our understanding of bivalve diversity and evolution.
335 Thus, the researchers concluded that crenulations serve as a morphological marker
336 for identifying the sex and reproductive stage of these bivalves (Karapunar et al.,
337 2021).

338 On the other hand, invasive techniques such as histological analysis offer a
339 more thorough but harmful method for determining the sex of *T. granosa*. A
340 study on the Spawning Period of Blood Cockle *Tegillarca granosa* (Linnaeus,
341 1758) in Myeik Coastal. 240 blood cockle samples were examined for sex and
342 gonad maturity stages using histological examination, with shell lengths ranging
343 from 26-35mm and shell weights from 8.1-33g. For histological analysis, the whole
344 soft tissues were removed from the shell and the flesh containing most parts of
345 the gonads was fixed in formalin, dehydrated in an upgraded series of ethanol,
346 and cleared in xylene. This invasive method allows for precise identification of
347 the gonadal maturation stages based on the cellular and structural changes in the
348 gonads.

349 The classification of the gonad stages used was by Yurimoto et al. (2014).
350 There are five maturation stages of gonadal development: immature (Stage I),
351 developing (Stage II), mature (Stage III), spawning (Stage IV), and spent (Stage
352 V) stages. The sex of the *T. granosa* was confirmed by the color of the gonad and
353 by conducting a histological examination of the gonads. During the immature
354 stage, sex determination was indistinguishable due to the difficulties of observing
355 the germ cells. In the developing stage, the spermatocytes and a few spermatids
356 can be seen for males, and immature oocytes are attached to the tube wall for
357 the female. In the mature stage, the follicles are full of spermatozoa with their
358 tails pointing towards the center of the tube for the male and the female are full
359 of mature oocytes that are irregular or polygonal in shape with the oval nucleus.
360 Upon reaching spawning, some spermatozoa are released, causing the empty space
361 in the follicle wall for males and females there is a decrease in the number of
362 mature oocytes and it exhibits nuclear disappearance due to the breakdown of
363 the germinal vesicle. Lastly, the spent stage is where the genital tube is deformed
364 and devoid of spermatocytes which have completely spawned. In the female, the
365 genital tube is deformed and degenerated making it empty. The morphology of
366 the cockle gonad shows that the area of the gonad increases according to the
367 increased levels of gonad maturity. The coloration of the gonad tissue layer in the
368 blood cockle varies from orange-red to pale orange in females and from white to
369 grayish-white in males for different maturity stages (May et al., 2021).

370 Although the histological examination is the most reliable method for obtain-
371 ing accurate information on the reproductive biology and sex determination of
372 *T. granosa*, it has limitations. Given its invasive nature, this approach requires
373 the dissection and destruction of specimens, making it unsuitable for continuous

374 monitoring and conservation efforts. Moreover, the current understanding of sex
375 determination in bivalves and mollusks is poor, and no chromosomes that can
376 be differentiated based on their morphology have been discovered (Afiati, 2007).
377 There exists a study that can provide insight into the sex-determining factor in
378 bivalves but *N. schoberti* is more difficult to analyze concerning potential sexual
379 dimorphism. Thickening the edges of the shell increases its inflation, which means
380 the shell can hold more space inside. This extra space helps protandrous females
381 accommodate more eggs.

382 **2.3 Machine Learning and Deep Learning in Bi-** 383 **ological Studies**

384 Machine learning has the potential to improve the quality of life of human beings
385 and has a wide range of applications in terms of research and development. The
386 term machine learning refers to the invention and algorithm evaluation that en-
387 ables pattern recognition, classification, and prediction based on models generated
388 from available data (Tarca, Carey, Chen, Romero, & Drăghici, 2007). The study
389 of machine learning methods has advanced in the last several years including bio-
390 logical studies. In biological studies, machine learning has been used for discovery
391 and prediction. This section will explore existing machine learning studies that
392 are applied in biological sciences highlighting the identification of sex in shells,
393 bivalves, and mollusks.

394 **2.3.1 Deep Learning for Phenotype Classification in Ark** 395 **Shells**

396 In the study, the researchers utilized three (3) convolutional neural network (CNN)
397 models: the Visual Geometry Group Network (VGGnet), the Inception Residual
398 Network (ResNet), and the SqueezeNet (Kim, Yang, Cha, Jung, & Kim, 2024).
399 These deep learning models are utilized to the ark shells namely *Anadara kagoshi-*
400 *mensis*, *Tegillarca granosa*, and *Anadara broughtonii* to identify the phenotype
401 classification.

402 The researchers classified the ark shells based on radial rib count where they
403 investigated the difference in the number of radial ribs between three species and
404 were counted. Their CNN-based model that classifies images of three ark shells
405 can provide a theoretical basis for bivalve classification and enable the tracking of
406 the entire production process of ark shells from catching to selling with the support

407 of big data, which is useful for improving food safety, production efficiency, and
408 economic benefits (Kim et al., 2024).

409 **2.3.2 Geometric Morphometrics and Machine Learning for** 410 **Species Delimitation**

411 In *Geometric morphometrics and machine learning challenge currently accepted*
412 *species limits of the land snail Placostylus (Pulmonata: Bothriembryontidae) on*
413 *the Isle of Pines, New Caledonia*, the shell size was quantified using centroid size
414 from the Procrustes analysis, and both the shape and size information were used in
415 training the machine learning model. Their study concluded that the researchers
416 support utilizing both methods: supervised and unsupervised machine learning,
417 rather than choosing either of them individually. In general, their research con-
418 tributes to the growing number of studies that have combined geometric morpho-
419 metrics, with the aid of machine learning which is helpful in biological innovation
420 and breakthrough (Quenu, Trewick, Brescia, & Morgan-Richards, 2020).

421 **2.3.3 Contour Analysis in Mollusc Shells Using Machine** 422 **Learning**

423 Tuset et al., (2020) in their study, *Recognising mollusc shell contours with en-*
424 *larged spines: Wavelet vs Elliptic Fourier analyses*, mentioned Gastropod shells
425 have large spines and sharp shapes which differ based on environmental, taxo-
426 nomic, and evolutionary influences. The researchers stated that classic morpho-
427 metric methods may not accurately depict morphological features of the shell,
428 especially when using the angular decomposition of the contour. The current
429 research examined and compared the robustness of the contour analysis using
430 wavelet transformed and Elliptic Fourier descriptors for gastropod shells with en-
431 larged spines. For that, the researchers analyzed two geographical and ecologically
432 separated populations of *Bolinus brandaris* from the NW Mediterranean Sea. Re-
433 sults showed that contour analysis of gastropod shells with enlarged spines can
434 be analyzed using both methodologies, but the wavelet analysis provided better
435 local discrimination. From an ecological perspective, shells with various sizes of
436 spines in both areas indicate a broad adaptability of the species.

2.3.4 Machine Learning for Shape Analysis of Marine Organisms

In the study of Lishchenko and Jones (2021), titled *Application of Shape Analyses to Recording Structures of Marine Organisms for Stock Discrimination and Taxonomic Purposes*, they utilized geometric morphometrics (GM) as an approach to the traditional method of collecting linear measurements with the application of multivariate statistical methods and outline analysis in recording the structures of marine organisms. The main taxonomic categories (mollusks, teleost fish, and elasmobranchs) with their hard bodies have been used as an indication of age and a determinable time-scale and structure continue to go through life (Arkhipkin, 2005; Kerr & Campana, 2014). This study has explored variations in the morphometry of recording structures in stock discrimination and systematics. The researchers utilized the principal component analysis rather than the traditional approach, which helps simplify the data without losing important information. They utilized landmark-based geometric morphometrics which has three different types namely: discrete juxtaposition of tissue, maxima or curvature or other morphogenetic processes, and lastly, the extremal points are constructed landmarks.

Generalized Procrustes Analysis (GPA) is a common superimposition technique in landmark-based geometric morphometrics that aligns landmarks via translation, scaling, and rotation to eliminate non-shape deviations (Zelditch, Swiderski, & Sheets, 2004). However, there is a limit to the amount of smooth areas that may be captured, and it is possible to overlook significant shape details. Utilization of the semi-landmarks enhanced the shape description (Adams, Rohlf, & Slice, 2004). The researchers observed that using an outline-based approach would be more effective than using a landmark-based approach.

Another approach is the Fourier analysis which is a curve-fitting approach commonly used due to its well-known mathematical background and how general functions can be decomposed into trigonometric or exponential functions with definite frequencies. It has two main approaches namely: Polar Transform (PT) in which it expresses the outline using equally spaced radii and Elliptical Fourier Analysis (EFA) which separately analyzes the x and y coordinates of the shape. The PT works for simple rounded outlines and has the tendency to miss details in more complex shapes, unlike EFA which can handle complex, convoluted outlines (Zahn & Roskies, 1972; Doering & Ludwig, 1990; Ponton, 2006). Many researchers view EFA as the most effective Fourier method for providing a comprehensive and detailed description of recording structures (Mérigot, Letourneur, & Lecomte-Finiger, 2007; Ferguson, Ward, & Gillanders, 2011; Leguá, Plaza, Pérez, & Arkhipkin, 2013; Mahé et al., 2016).

Landmark-based methods used in the study showed that there are detectable differences between male and female octopuses. However, the accuracy of determining sex based on these differences was low, similar to the results obtained with traditional morphometric techniques. The study involved a relatively small sample size of 160 individuals, and the structure being analyzed (the stylet, or internalized shell) varies significantly between individuals. Although the results aligned with findings from other studies that attempted to identify gender differences in cephalopods, the researchers concluded that the approach might not be accurate enough for reliable sex determination.

2.3.5 Deep Learning for Landmark-Free Morphological Feature Extraction

In another study, *a deep learning approach for morphological feature extraction based on variational auto-encoder: an application to mandible shape*, the Morpho-VAE machine learning approach was used to conduct a landmark-free shape analysis. Morpho-Vae reduces dimensions by concentrating on morphological features that distinguish data with different labels using an image-based deep learning framework that combines unsupervised and supervised machine learning. After utilizing the method in primate mandible images, the morphological features reveal the characteristics to which family they belonged. Based on the result, the method applied provides a versatile and promising tool for evaluating a wide range of image data of biological shapes including those missing segments.

2.3.6 Machine Learning for Sex Differentiation in Abalone

In the study, *Towards Abalone Differentiation Through Machine Learning*, researchers identified a problem in abalone farming which is having to identify the sex of abalone to apply measures for its growth or preservation. The researchers classified abalone sex using machine learning. Researchers trained the machine to classify different types of classes which are male, female, and immature. Based on the result, demonstrated the impact of utilizing linear classifiers.

Similarly, in the study, *Data scaling performance on various machine learning algorithms to identify abalone sex*, the researchers of the University of India (2022), focused on the data scaling performance of various machine learning algorithms to identify the abalone sex, specifically using min-max normalization and zero-mean standardization. The different machine learning algorithms are the Supervised Vector Machine (SVM), Random Forest, Naive Bayesian, and Decision Tree. Their

study aims to utilize machine learning in terms of identifying the trends and distribution patterns in the abalone dataset. Eight features of the abalone dataset (length, diameter, height, whole weight, shucked weight, viscera weight, shell weight, ring) were used to determine the three sexes of Abalone. Their data has been grouped based on sex which are Female, Male, and Infant. They utilized the Synthetic Minority Oversampling Technique (SMOTE) in data balancing for the preprocessing of the data. Followed by data scaling or normalization where it converts numeric values in a data set to a general scale without distorting differences in the range of values. Then they classified by splitting the data into training and testing sets (Arifin, Ariawan, Rosalia, Lukman, & Tufailah, 2021).

The researchers found out that the Naive Bayes performs more consistently than other algorithms when the abalone dataset was applied to both min-max and zero-mean normalization has the highest to lowest average accuracy, respectively 62.37% (Random Forest), 59.49% (SVM kernel RBF), 57.20% (Decision Tree), 56.59% (SVM linear kernel), and 53.39% (Naïve Bayesian). Overall, despite the decrease in the performance with the normalization, the Random Forest achieves the highest average balanced accuracy of 74.87%, sensitivity of 66.43%, and specificity of 83.31%. Liu et al. found that Random Forest not only can handle large and complex datasets and can run in parallel using multiple random forests but is also more accurate than other algorithms because it selects the best features to improve model performance (Arifin et al., 2021).

2.3.7 Machine Learning for Geographical Traceability in Bivalves

In the study, *BivalveNet: A hybrid deep neural network for common cockle (Cerastoderma edule) geographical traceability based on shell image analysis*, the researchers incorporated computer vision and machine learning technologies for an efficient determination of blood cockle harvesting origin based on the shell geometric and morphometric analysis. It aims to improve the traceability methodologies in these organisms and its potential as a reliable traceability tool. Thirty *Cerastoderma edule* samples were collected along the five locations on the Atlantic West and South Portuguese coast with individual images processed using lazy snapping segmentation, spectro-textural-morphological phenotype extraction, and feature selection through hybrid Principal Component Analysis and Neighborhood Component Analysis (Concepcion, Guillermo, Tanner, Fonseca, & Duarte, 2023).

The researchers developed a non-invasive image-based traceability technique, an alternative to the chemical and biochemical analysis of the bivalves. It was able to incorporate machine learning methods to promote lesser human interven-

tion The researchers discovered that BivalveNet emerged as the superior model for bivalves with 96.91% accuracy which is comparable to the accuracy of the destructive methods with 97% and 97.2% accuracy rate. The result of the study aided the researchers in concluding that there is a possibility of on-site evaluation of the bivalve through the implementation of a mobile app would allow the public and official entities to obtain information regarding the provenance of seafood products' traceability because of its non-invasive and image-based aspects (Concepcion et al., 2023).

Tegillarca granosa is known for having no sexual dimorphism. However, through several related studies, the researchers can apply how family shells of *Tegillarca granosa* have been identified based on its morphological and morphometric characteristics, and the methods used in machine learning in identifying its sex.

2.4 Limitations on Sex Identification in *Tegillarca granosa*

To date, no distinction has been made between the male and female *T. granosa* in sexing methodology. In cockle aquaculture without clearly apparent sexual dimorphism, sexing can be performed using invasive methods such as chemical stimulation, dissection, and gonad-stripping. Induced spawning, specifically temperature shocking, is the most natural and least invasive method for bivalves (Aji, 2011). However, the method (Wong & Lim, 2018) of immersing cockles in water from hot to cold with a specific temperature requires deliberate and careful manipulation of the temperature over a specific period and would require constant management and monitoring.

Recent studies involved non-invasive methods, with a specific emphasis on morphological characteristics as indicators of sex differentiation. However, Tatsuya Yurimoto et al. (2014) stated that the existing methods for determining the sex of bivalves and mollusks in general are somewhat limited (Afati, 2007). At present, there is no recorded evidence of sexual dimorphism in *Tegillarca granosa*. Gonochoristic is the classification given to *Tegillarca granosa* (Lee, 1997). However, Lee et al. (2012) reported that the sex ratio varied with shell length, suggesting that sex might alter.

Hermaphrodites can exhibit either sequential (asynchronous) or simultaneous (synchronous or functional) characteristics. Sequential hermaphrodites switch genders after being male or female for one or multiple yearly cycles. (Heller, 1993; Gosling, 2004; Collin, 2013). Sex change and consecutive hermaphroditism

581 have been observed in different bivalve species, including Ostreidae, Pectinidae,
582 Veneridae, and Patellidae. However, macroscopically differentiating bivalve sex is
583 challenging. The only way it may be identified is through histological analysis of
584 gonad remains but to do so there is an act of killing the organism (Coe, 1943;
585 Gosling, 2004). Verification of sex change in bivalves to classify whether male or
586 female while they are alive is challenging since they need to be re-confirmed and
587 re-evaluated to be the same individual after a year.

588 Lee et al. (2012) found out that *T. granosa*, a species in Arcidae, has been
589 discovered to be a sequential hermaphrodite, with the sex ratio changing with an
590 increase in the shell size. In bivalves, sex changes usually happen when the gonad
591 is not differentiated between spawning seasons (Thompson, Newell, Kennedy, &
592 Mann, 1996). But in *T. granosa*, after the spawning season, sex changes during
593 its inactive phase. Results showed a 15.1% sex change ratio, with males having
594 a higher sex change ratio (21.2%) than females (6.2%). The 1+ year class had a
595 higher ratio (17.8%) than the 2+ year class (12.1%). Thus, this study indicates
596 that *T. granosa* is a sequential hermaphrodite. The results of the study demon-
597 strated that the bivalve's age affects the sex ratio and degree of sex change, but
598 additional in-depth investigation is required to determine the role that genetic
599 and environmental factors play in these changes.

600 No literature in the study of mollusks specifically addresses the machine learn-
601 ing algorithm used to determine the sex of *T. granosa* bivalves in various mod-
602 els. Nevertheless, various techniques such as shape analysis, morphometric ana-
603 lysis, Wavelet, and Fourier analysis, as well as different deep learning models
604 like VNet, ResNet, and SqueezeNet in CNN networks are utilized for pheno-
605 type classification, while different machine learning algorithms could serve as the
606 foundation for this research project

607 2.5 Synthesis of the Study

608 This section of the paper summarizes the technologies used in the different studies
609 related to the pursuit of the study entitled, Non-invasive Sex Identification of *T.*
610 *granosa* using machine learning.

Literature	Technology / Method Used	Description of Problem	Pros	Cons
Initial Attempts on Spawning and Larval Rearing of the Blood Cockle, <i>Tegillarca granosa</i> in the Philippines	Temperature shock	No recent studies are available on the production and rearing of <i>T. granosa</i> in the Philippines.	Employed less invasive techniques which minimize the stress in <i>T. granosa</i> and can lead to better survival rates.	Time-consuming as the entire process from fertilization to the spat stage took 120 days.
The earliest example of sexual dimorphism in bivalves—evidence from the astartid <i>Nicaniella</i> (Lower Jurassic, southern Germany)	Morphometric analysis, microscope imaging, principal component analysis (PCA), and Fourier shape analysis	To address the observed shell dimorphism in the Early Jurassic bivalve <i>Nicaniella rakoveci</i> , namely the presence or lack of crenulations on the ventral shell margin, and whether these variations represent sexual dimorphism and sequential hermaphroditism.	The methods used reveal significant morphological differences with regard to sexual dimorphism.	There could be misinterpretation of the shape differences of bivalves due to the constraints and resolution of technologies used.
Spawning Period of Blood Cockle <i>Tegillarca Granosa</i> (Linnaeus, 1758) in Myeik Coastal	Histological examination	The need to understand the reproductive period of <i>T. granosa</i> in Myeik to ensure sustainable aquaculture and to prevent overexploitation.	Method used allows for accurate sex identification based on the histological characteristics and color of the gonads.	Invasive technique used to determine the sex of <i>T. granosa</i> through gonad histological analysis.
Deep learning-based phenotype classification of three ark shells: <i>Anadara kagoshimensis</i> , <i>Tegillarca granosa</i> , and <i>Anadara broughtonii</i>	Convolutional neural network (CNN) models, VGGNet, Inception-ResNet, SqueezeNet	Traditional methods of recognizing and classifying ark shell species based on shell traits are time-consuming and inaccurate.	Automated classification of the three ark shells using a deep learning model obtained an accuracy of 92.4%.	Challenges may arise with certain ark shells that share similar morphology.
Geometric morphometrics and machine learning challenge currently accepted species limits of the land snail <i>Placostylus</i> (Pulmonata: Bothriembryontidae) on the Isle of Pines, New Caledonia	Neural network analysis (supervised learning) and Gaussian mixture models (unsupervised learning)	To determine whether the shape and size of the snail's shells can distinguish between two <i>Placostylus</i> species, particularly in groups that appear to be hybrids.	Combining geometric morphometrics and machine learning effectively answers biological issues, providing insights into species classification and possible hybridization.	Difficulty classifying intermediate phenotypes, with potential for overfitting and misclassification in both learning methods.
Recognizing mollusc shell contours with enlarged spines: Wavelet vs Elliptic Fourier analyses	Wavelet functions and Elliptic Fourier descriptors	Addresses the difficulty of accurately defining phenotypic diversity in gastropod shells.	Advanced contour analysis methods allow accurate differentiation of gastropod shell forms.	Cannot clarify the causes of phenotypic variation in the two populations studied.
Application of Shape Analyses to Recording Structures of Marine Organisms for Stock Discrimination and Taxonomic Purposes	Landmark- and outline-based Geometric Morphometric methods	To address difficulties in differentiating between stocks of marine organisms to prevent misidentification that could affect conservation and management.	Shape analysis improves taxonomic classification precision and offers close distinction between related species or organisms.	Landmark-based methods can be sensitive to landmark placement.
A deep learning approach for morphological feature extraction based on variational auto-encoder: an application to mandible shape	Morphological regulated variational AutoEncoder (Morpho-VAE)	The need for reliable, landmark-free methods, such as a modified variational autoencoder, to extract and decipher complex shapes from image data.	Employs dimension reduction and feature extraction, making it a user-friendly tool for biology non-experts.	Limited sample size in certain families presented challenges.
Towards Abalone Differentiation Through Machine Learning	Machine learning algorithms	Identifying the sex of abalones is challenging for producers applying specific growth or preservation strategies.	Machine learning algorithms accurately classify abalone sex into three categories: male, female, and immature.	Selected features may not fully capture the complexity of abalone morphology.
BivalveNet: A hybrid deep neural network for common cockle (<i>Cerastoderma edule</i>) geographical traceability based on shell image analysis	EfficientNet-Bo, ResNet101, MobileNetV2, InceptionV3	Addresses the difficulty of accurately tracing bivalve harvesting origins using computer vision and machine learning algorithms to enhance seafood traceability and combat food fraud.	Non-invasive, image-based tools for bivalve traceability provide faster, cheaper, and equally accurate alternatives to traditional chemical analysis methods.	Small sample size (only 30 cockles) limits model reliability.

611 Recent developments and breakthroughs in machine learning offer hopeful so-
612 lutions for biological issues. Research findings indicate that various machine learn-
613 ing techniques such as CNNs, geometric morphometrics, and deep learning mod-
614 els. They are deemed to be effective for identifying phenotypes and determining
615 the gender of various aquaculture commodities, such as mollusks and abalones.
616 These techniques provide a starting point for creating new, non-invasive ways to
617 differentiate male and female *T. granosa*, potentially addressing the drawbacks of
618 manual and invasive methods. Thus, machine learning to examine morphological
619 and morphometric features may streamline the process of sex identification.

620 Nevertheless, the use of machine learning to determine the sex of *T. granosa*
621 has not been fully explored. It lacks up-to-date and significant related literature
622 on using machine learning to identify sex in *T. granosa*, particularly given the
623 species' possible sequential hermaphroditism and lack of obvious external sexual
624 distinctions.

Chapter 3

Research Methodology

This chapter discusses the materials and methods to be employed in the study, focusing on the development requirements and the software and languages utilized. This will also entail the overall workflow in conducting the study, Non-Invasive Methods in Determining the Sex of *Tegillarca granosa* (blood cockles) using machine learning technologies. The different machine/deep learning algorithms will be thoroughly discussed to ensure a comprehensive understanding of the entity of the research endeavor and its processes.

Dr. Victor Emmanuel Ferriols, the director of the Institute of Aquaculture, will oversee the overall workflow and conduct of this experiment. The researchers will also be guided by the research associates, LC Mae Gasit and Allena Esther Artera. Consequently, the whole dataset collection process will be done at the University of the Philippines Visayas hatchery facility.

3.1 Sample Collection

A total of 1000 adult *T. granosa* that have already spawned will be used in this experiment wherein their sex was already classified as male or female. The sample sizes are going to range from 34 to 61 mm and will be sourced from the coastal area in the municipality of Zaraga, Iloilo, Philippines, as well as from fish markets in the municipality of Ivisan, Capiz, Philippines. The research and experimentation will be done at the University of the Philippines Visayas hatchery facility in Miagao, Iloilo, Philippines. The samples will be placed in 200 L fiberglass reinforced plastic (FRP) tanks containing filtered seawater with 35 ppt salinity (Ferriols, Miranda, 2023) and will be subjected to spawning to categorize male from female

649 *T. granosa*. The samples will undergo a series of temperature fluctuations to
650 induce the spawning of gametes as described in the study of Ferriols and Miranda
651 (2023). This method, induced spawning, is the most natural and least invasive
652 method for bivalves compared to other methods (Aji, 2021). Thus, after the
653 spawning, there would be 500 classified males and 500 classified females.

654 3.2 Ethical Considerations

655 Ethical approval was not required for this study involving animals, as per local leg-
656 islation and institutional guidelines, because the experiments were conducted only
657 on species that are commonly used as food and intended for human consumption.

658 3.3 Creating *T. granosa* Dataset

659 For the initial preparation of the experiment, the researchers will collect primary
660 observations for 100 samples of *T. granosa*. For the actual experimentation, the
661 researchers will collect the dataset by batch eventually comprising 1000 samples
662 of *T. granosa*. The images captured for the dataset will be saved in png format
663 with a file naming convention of the sample’s sex, the orientation or view of the
664 shell, and its corresponding number out of the total 1000 samples. Female *T.*
665 *granosa* samples will begin with 0 in their file name, while males will begin with
666 1, followed by the views captured such as (1) dorsal, (2) ventral, (3) anterior,
667 (4) posterior, (5) left lateral, and (6) right lateral, and lastly, a unique sample
668 number. For example, “010001” will be the file name for the first female sample
669 taken from the dorsal view and “110001” for the first male sample also taken from
670 the dorsal view. The dataset will be organized in a CSV file that lists each image’s
671 file name along with their shell’s width, height, length, rib count, length of the
672 hinge line, and distance between their umbos. This dataset will be essential for
673 machine learning model training and testing.

674 3.4 Morphological Characteristics Collection

675 Morphology refers to the biological form and represents one of the most visually
676 recognizable phenotypes across all organisms (Tsutsumi et al., 2023). Morphology
677 is a term that describes structural characteristics by measuring specific compo-
678 nents, namely, dimensions such as shapes, sizes, and colors. As stated by the

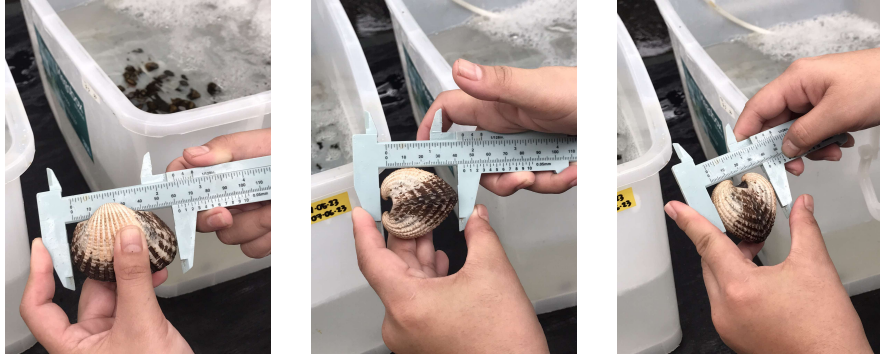


Figure 3.1: Length Figure 3.2: Width Figure 3.3: Height

Figure 3.4: *T. granosa*'s measurements

researchers, quantifying and characterizing the shape is essential to understanding and visualizing the variations in *T. granosa*'s morphology. In this study, the researchers are going to measure the height, width, and length of *T. granosa*. The dimensions will be recorded using a Vernier caliper to the nearest 0.01 mm. The length of the *T. granosa* refers to the measurement from the anterior to the posterior of the shell, the width will be measured through the shell's widest point from the left to the right valve and lastly, the height will be measured from the base of the shell to the shell's apex. The height of the gap between the valves near the hinge will also be measured. The authors Reymont and Kennedy (1998), indicated that the use of counts of the shell ribs as supplementary information increases identification accuracy. Thus, the researchers will also take into account the difference in the rib count of the male and female *T. granosa* and the ratio will be calculated since the sizes of the blood clams may vary. Sex ratio, size frequency distribution, and relative growth rates were used to investigate sexual dimorphism.

3.5 Image Acquisition and Pre-Processing

In this study, there would be three major phases for the image processing to be employed namely (1) color thresholding, (2) segmentation, and (3) image hole filling and dilating. The researchers constructed a controlled environment for capturing the samples utilizing a box-like structure of (?) meters with a green background surface. This setup was designed to maintain uniform captures of the images, and a consistent measurement between the sample and the camera, fixing the camera at 50 cm above the *T. granosa*. Placing a ring light to the left of the box, and using a camera with flash to ensure the image quality, eliminate

703 shadows and clarity of the sample during the image acquisition process. For color
 704 thresholding, the researchers utilized the red, green, blue (RGB), hue saturation
 705 value (HSV), luminance, blue chromaticity, red chromaticity (YCbCr), and (Lu-
 706 minance, a, b)** (CIElab) images obtained from the smartphone considering their
 707 wide availability across various stages in the bivalve industry using the MATLAB
 708 Colour Thresholding Toolbox in determining which among the four-color spectra
 709 may generate the cleanest version of the training images with absence of any blobs
 710 (Jayasundara et al., 2023). Google Pixel 3 XL will be utilized with the following
 711 specifications: 2960 x 1440 for the resolution, 4,032 x 3,024 pixels (12.2 MP) for
 712 the dimensions, f/1.8 for the fstop, 28mm (wide), $\frac{1}{2}$.55", 1.4 μ m, dual pixel PDAF,
 713 OIS. [insert reference] After thresholding, the lazy snapping technique will be im-
 714 plemented by manually drawing the background and the foreground lines that
 715 represent the black pixels and the bivalve pixels. The lazy snapping algorithm
 716 will be configured using the 20 000 superpixels which can divide the *T. granosa*'s
 717 images into 20, 000 irregularly shaped geometric pixels that will be based on the
 718 CIElab gradients through K-means clustering with $K = 3$. For the last step, the
 719 researchers will perform image hole filling and dilating to ensure that no blobs
 720 are remaining that can contribute to noise which can affect the correctness of the
 721 extracted feature by taking into consideration the 200-pixel blobs that are discon-
 722 nected from the largest object in its binary form. This will result in black pixels
 723 made by binary filling and dilating to remove the blobs. [reference] Image process-
 724 ing will be performed on the MATLAB [version-] installed on the [laptop] with
 725 specs. The images will be saved based on how it was stated on the collection of
 726 the image dataset. To ensure consistent comparisons for the analysis, the images
 727 were captured in different angles including dorsal, ventral, lateral, and anterior
 728 and posterior taken in uniform angles to provide visual coverage of the *T. granosa*
 729 sample.

730 **3.6 Machine/ Deep Learning Technologies**

731 This section of the paper will discuss the technologies to be used in training, and
 732 testing the model as well as associated techniques and algorithms. Since obtaining
 733 the induced samples was done per batch, the researchers will conduct an initial
 734 run with a support vector machine before delving into more complex methods
 735 such as deep learning models.

736 3.7 Support Vector Machine for Pre-evaluation

737 The shape of recording structures was first analyzed by collecting measurements
738 of linear distances and applying multivariate statistical methods to these data
739 (traditional linear measurement method) (Rohlf and Marcus, 1993). Geometric
740 morphometric (GM) methods are an alternative way of analyzing and quantifying
741 shape, which in theory retains more detail about the geometry of the structure
742 than could be obtained from linear measurements (Adams et al., 2004). Machine
743 learning techniques such as decision tree classification, support vector machines
744 (SVMs), and artificial neural networks (ANNs) have been applied to the analysis
745 of bivalve shell geometry and morphology to classify shells based on morpholog-
746 ical features, including shell shape, size, and texture, among others (Kiel, 2021).
747 The results of these studies have shown that machine learning algorithms can
748 accurately classify bivalve shells and provide insights into the relationships be-
749 tween shell morphology and various environmental factors. Following this, the
750 researchers are going to conduct a pre-evaluation of the linear measurements for
751 100 samples of *T. granosa* using a Support Vector Machine in order to quantify
752 whether the linear measurements can be a determining factor in determining the
753 sex of the samples before proceeding to more complex methods.

754 3.8 Deep Learning for Image-Based Classifica- 755 tion

756 After collecting a sufficient number of images and identifying initial patterns,
757 convolutional neural networks (CNNs) will be used. CNNs, models like VGGNet,
758 ResNet, and Inception have been effectively applied in phenotype classification
759 (Kim et al., 2024). In this study, the deep learning model will be specifically
760 adapted for the sex identification of *T. granosa* based on shell images. CNNs
761 will analyze the images and learn important details about their shapes that can
762 help identify whether they are male or female. Unlike the approach of using
763 three models taken by Kim et al. (2024), the researchers will focus on just one
764 model that has shown the best performance in their study which is SqueezeNet.
765 SqueezeNet is particularly advantageous because it reduces the number of pa-
766 rameters and amount of memory required to store the model without sacrificing
767 accuracy (Koonce, 2021; Sayed et al., 2021). Its ability to achieve high accuracy
768 in classifying shell images makes it a suitable choice for distinguishing between
769 male and female *T. granosa*. Python and Keras libraries will be used to train and
770 test the model. The dataset will be divided into training (), validation (), and
771 testing. Performance metrics such as accuracy, precision, recall, and F1-score will

772 be used to evaluate the model's effectiveness.

773 Chapter 4

774 Preliminary Results/System 775 Prototype

776 This chapter presents the preliminary results or the system prototype of your SP.
777 Include screenshots, tables, or graphs and provide the discussion of results.

References

- Adams, D. C., Rohlf, F. J., & Slice, D. E. (2004). Geometric morphometrics: ten years of progress following the 'revolution'. *Italian Journal of Zoology*, 71, 5–16. doi: 10.1080/11250000409356545
- Afiati, N. (2007, 01). Gonad maturation of two intertidal blood clams *Anadara granosa* (L.) and *Anadara antiquate* (L.) (Bivalvia: Arcidae) in central Java. , 10.
- Aji, L. P. (2011). Review: Spawning induction in bivalve. *Jurnal Penelitian Sains*, 14, 14207.
- Arifin, W. A., Ariawan, I., Rosalia, A. A., Lukman, L., & Tufailah, N. (2021). Data scaling performance on various machine learning algorithms to identify abalone sex. *Jurnal Teknologi Dan Sistem Komputer*, 10(1), 26–31. doi: 10.14710/jtsiskom.2021.14105
- Arkhipkin, A. I. (2005). Statoliths as 'black boxes' (life recorders) in squid. *Marine and Freshwater Research*, 56, 573–583. doi: 10.1071/mf04158
- Aypa, S. M., & Bacongus, S. R. (2000). Philippines: mangrove-friendly aquaculture. In J. H. Primavera, L. M. B. Garcia, M. T. Castaños, & M. B. Surtida (Eds.), *Mangrove-friendly aquaculture: Proceedings of the workshop on mangrove-friendly aquaculture organized by the seafdec aquaculture department, January 11-15, 1999, Iloilo City, Philippines* (pp. 41–56).
- BFAR. (2019). *Philippine fisheries profile 2018* (Tech. Rep.). PCA Compound, Elliptical Road, Quezon City, Philippines: Bureau of Fisheries and Aquatic Resources.
- Boey, P.-L., Maniam, G. P., Hamid, S. A., & Ali, D. M. H. (2011). Utilization of waste cockle shell (*Anadara granosa*) in biodiesel production from palm olein: Optimization using response surface methodology. *Fuel*, 90(7), 2353–2358. doi: 10.1016/j.fuel.2011.03.002
- Breton, S., Capt, C., Guerra, D., & Stewart, D. (2017, June). *Sex determining mechanisms in bivalves*. Preprints.org. doi: 10.20944/preprints201706.0127.v1
- Breton, S., Stewart, D. T., Shepardson, S., Trdan, R. J., Bogan, A. E., Chapman, E. G., ... Hoeh, W. R. (2010). Novel protein genes in animal mtDNA: A

new sex determination system in freshwater mussels (bivalvia: Unionoida)?
Molecular Biology and Evolution, 28(5), 1645–1659. doi: 10.1093/molbev/msq345

Budd, A., Banh, Q., Domingos, J., & Jerry, D. (2015). Sex control in fish: Approaches, challenges and opportunities for aquaculture. *Journal of Marine Science and Engineering*, 3(2), 329–355. doi: 10.3390/jmse3020329

Burdon, D., Callaway, R., Elliott, M., Smith, T., & Wither, A. (2014, 04). Mass mortalities in bivalve populations: A review of the edible cockle *Cerastoderma edule* (L.). *Estuarine, Coastal and Shelf Science*, 150. doi: 10.1016/j.ecss.2014.04.011

Campos, A., Tedesco, S., Vasconcelos, V., & Cristobal, S. (2012). Proteomic research in bivalves: Towards the identification of molecular markers of aquatic pollution. *Proteomic Research in Bivalves*, 75(14), 4346–4359. doi: 10.1016/j.jprot.2012.04.027

Coe, W. R. (1943). Sexual differentiation in mollusks. i. pelecypods. *The Quarterly Review of Biology*, 18, 154–164. doi: 10.1086/qrb.1943.18.issue-2

Collin, R. (2013). Phylogenetic patterns and phenotypic plasticity of molluscan sexual systems. *Integrative and Comparative Biology*, 53, 723–735. doi: 10.1093/icb/ict076

Concepcion, R., Guillermo, M., Tanner, S. E., Fonseca, V., & Duarte, B. (2023). Bivalvenet: A hybrid deep neural network for common cockle (*cerastoderma edule*) geographical traceability based on shell image analysis. *Ecological Informatics*, 78, 102344. doi: 10.1016/j.ecoinf.2023.102344

Doering, P., & Ludwig, J. (1990). Shape analysis of otoliths—a tool for indirect ageing of eel, *anguilla anguilla* (L.)? *International Review of Hydrobiology*, 75(6), 737–743. doi: 10.1002/iroh.19900750607

Erica, D. (2018, April 4). *Clam dissection: A first step into dissection and anatomy for young learners*. Rosie Research. Retrieved from <https://rosieresearch.com/clam-dissection-anatomy/>

Fao 2024 report: Sustainable aquatic food systems important for global food security – european fishmeal. (2024). <https://effop.org/news-events/fao-2024-report-sustainable-aquatic-food-systems-important-for-global-food-security/>.

Ferguson, G. J., Ward, T. M., & Gillanders, B. M. (2011). Otolith shape and elemental composition: Complementary tools for stock discrimination of mullet (*argyrosomus japonicus*) in southern australia. *Fish Research*, 110, 75–83. doi: 10.1016/j.fishres.2011.03.014

Gosling, E. (2004). *Bivalve molluscs: biology, ecology and culture*. Oxford: Blackwell Science.

Heller, J. (1993). Hermaphroditism in molluscs. *Biological Journal of the Linnean Society*, 48, 19–42. doi: 10.1111/bij.1993.48.issue-1

Ishak, A. R., Mohamad, S., Soo, T. K., & Hamid, F. S. (2016). Leachate and

- 852 surface water characterization and heavy metal health risk on cockles in
853 kuala selangor. In *Procedia - social and behavioral sciences* (Vol. 222, pp.
854 263–271). doi: 10.1016/j.sbspro.2016.05.156
- 855 Karapınar, B., Werner, W., Fürsich, F. T., & Nützel, A. (2021). The ear-
856 liest example of sexual dimorphism in bivalves—evidence from the astar-
857 tid *Nicaniella* (lower jurassic, southern germany). *Journal of Paleontology*,
858 95(6), 1216–1225. doi: 10.1017/jpa.2021.48
- 859 Kerr, L. A., & Campana, S. E. (2014). Chemical composition of fish hard parts
860 as a natural marker of fish stocks. In *Stock identification methods* (pp. 205–
861 234). Elsevier. doi: 10.1016/b978-0-12-397003-9.00011-4
- 862 Kim, E., Yang, S.-M., Cha, J.-E., Jung, D.-H., & Kim, H.-Y. (2024). Deep
863 learning-based phenotype classification of three ark shells: *Anadara kagoshi-*
864 *mensis*, *tegillarca granosa*, and *anadara broughtonii*. *Frontiers in Marine*
865 *Science*, 11. doi: 10.3389/fmars.2024.1356356
- 866 Lee, J. H. (1997). Studies on the gonadal development and gametogenesis of the
867 granulated ark, *tegillarca granosa* (linne). *The Korean Journal of Malacol-*
868 *ogy*, 13, 55–64.
- 869 Leguá, J., Plaza, G., Pérez, D., & Arkhipkin, A. (2013). Otolith shape analysis as
870 a tool for stock identification of the southern blue whiting, *micromesistius*
871 *australis*. *Latin American Journal of Aquatic Research*, 41, 479–489.
- 872 Mahé, K., Oudard, C., Mille, T., Keating, J., Gonçalves, P., Clausen, L. W., &
873 et al. (2016). Identifying blue whiting (*micromesistius poutassou*) stock
874 structure in the northeast atlantic by otolith shape analysis. *Canadian*
875 *Journal of Fisheries and Aquatic Sciences*, 73, 1363–1371. doi: 10.1139/
876 cjfas-2015-0332
- 877 May, K., Maung, C., Phyu, E., & Tun, N. (2021). Spawning period of blood cockle
878 *tegillarca granosa* (linnaeus, 1758) in myeik coastal areas. *J. Myanmar Acad.*
879 *Arts Sci*, 4.
- 880 Miranda, D. V., & Ferriols, V. M. E. N. (2023). Initial attempts on spawning and
881 larval rearing of the blood cockle, *tegillarca granosa* (linnaeus, 1758), in the
882 philippines. *Asian Fisheries Science*, 36(2). doi: 10.33997/j.afs.2023.36.2
883 .001
- 884 Mérigot, B., Letourneur, Y., & Lecomte-Finiger, R. (2007). Characterization of
885 local populations of the common sole *solea solea* (pisces, soleidae) in the nw
886 mediterranean through otolith morphometrics and shape analysis. *Marine*
887 *Biology*, 151(3), 997–1008. doi: 10.1007/s00227-006-0549-0
- 888 Narasimham, K. A. (1988). Taxonomy of the blood clams *anadara* (*tegillarca*)
889 *granosa* (linnaeus, 1758) and *a. (t.) rhombea* (born, 1780).
- 890 Naylor, R. L., Goldburg, R. J., Primavera, J. H., Kautsky, N., Beveridge,
891 M. C. M., Clay, J., ... Troell, M. (2000). Effect of aquaculture on world
892 fish supplies. *Nature*, 405(6790), 1017–1024. doi: 10.1038/35016500
- 893 Ponton, D. (2006). Is geometric morphometrics efficient for comparing otolith

894 shape of different fish species? *Journal of Morphology*, 267(7), 750–757.
895 doi: 10.1002/jmor.10439

896 Quenu, M., Trewick, S. A., Brescia, F., & Morgan-Richards, M. (2020). Geometric
897 morphometrics and machine learning challenge currently accepted species
898 limits of the land snail *placostylus* (pulmonata: Bothriembryontidae) on the
899 isle of pines, new caledonia. *Journal of Molluscan Studies*, 86(1), 35–41.
900 doi: 10.1093/mollus/eyz031

901 Sany, S. B. T., Hashim, R., Rezayi, M., Salleh, A., Rahman, M. A., Safari, O.,
902 & Sasekumar, A. (2014). Human health risk of polycyclic aromatic hydro-
903 carbons from consumption of blood cockle and exposure to contaminated
904 sediments and water along the klang strait, malaysia. *Marine Pollution*
905 *Bulletin*, 84(1-2), 268–279. doi: 10.1016/j.marpolbul.2014.05.004

906 Srisunont, C., Nobpakhun, Y., Yamalee, C., & Srisunont, T. (2020). Influence
907 of seasonal variation and anthropogenic stress on blood cockle (*tegillarca*
908 *granosa*) production potential. *Influence of Seasonal Variation and Anthro-*
909 *pogenic Stress on Blood Cockle (Tegillarca Granosa) Production Potential*,
910 44(2), 62–82.

911 Tarca, A. L., Carey, V. J., Chen, X.-w., Romero, R., & Drăghici, S. (2007). Ma-
912 chine learning and its applications to biology. *PLoS Computational Biology*,
913 3(6), e116. doi: 10.1371/journal.pcbi.0030116

914 Thompson, R. J., Newell, R. I. E., Kennedy, V. S., & Mann, R. (1996). Repro-
915 ductive process and early development. In V. S. Kennedy, R. I. E. Newell,
916 & A. F. Eble (Eds.), *The eastern oyster crassostrea virginica* (pp. 335–370).
917 College Park, MD: Maryland Sea Grant.

918 Wong, T. M., & Lim, T. G. (2018). *Cockle (anadara granosa) seed produced in*
919 *the laboratory, malaysia*. (Handle.net) doi: 10.3366/in_3366.pdf

920 Zahn, C. T., & Roskies, R. Z. (1972). Fourier descriptors for plane closed curves.
921 *IEEE Transactions on Computers*, C-21, 269–281. doi: 10.1109/tc.1972
922 .5008949

923 Zelditch, M., Swiderski, D. L., & Sheets, H. D. (2004). *Geometric morphometrics*
924 *for biologists: A primer* (2nd ed.). Waltham: Elsevier Academic Press.

925 Zha, S., Tang, Y., Shi, W., Liu, H., Sun, C., Bao, Y., & Liu, G. (2022). Im-
926 pacts of four commonly used nanoparticles on the metabolism of a ma-
927 rine bivalve species, *tegillarca granosa*. *Chemosphere*, 296, 134079. doi:
928 10.1016/j.chemosphere.2022.134079

⁹²⁹ **Appendix A**

⁹³⁰ **Appendix Title**

931 **Appendix B**

932 **Resource Persons**

933 **Mr. Firstname1 Lastname1**

934 Role1

935 Affiliation1

936 emailaddr1@domain.com

937 **Ms. Firstname2 Lastname2**

938 Role2

939 Affiliation2

940 emailaddr2@domain.net

941