3D DEPTH IMAGING FOR PIG WEIGHT ESTIMATION

An Undergraduate Thesis

Presented to the
Faculty of Bachelor of Science in Computer Engineering
University of Science and Technology of Southern Philippines
Cagayan de Oro City

In Partial Fulfillment
of the Requirements for the Degree of
BACHELOR OF SCIENCE IN COMPUTER ENGINEERING

ACE GABRIEL GO
BRIAN JAMES C. CONCILLO
IRISH A. PARING
MARK LOUIE C. MONCANO
April 2025

TABLE OF CONTENTS

Page No.

LIST OF TABLES

LIST OF FIGURES

List of Equations

CHAPTER 1

INTRODUCTION

1.1 Background of the Study

Livestock weight estimation is a fundamental aspect of animal husbandry, directly influencing decisions regarding feeding schedules, growth monitoring, health management, and market readiness (Wang et al., 2024). Among various livestock species, pigs are a primary focus due to their significant contribution to global meat production. Pigs account for approximately 36 percent of the world's meat consumption, surpassing other livestock such as cattle and poultry, making them an economically critical species (Food and Agriculture Organization (FAO), 2022). Accurate weight estimation in pigs ensures optimal feed conversion rates, reduced production costs, and improved animal welfare, which are essential for sustainable farming practices (Terence et al., 2024).

In traditional farming methods, weighing scales are the standard tool for livestock weight measurement. However, these methods are often labor-intensive, time-consuming, and stressful for animals, potentially leading to injuries and a decline in productivity (Faucitano and Goumon, 2018). For smaller farms, the high costs and maintenance requirements of weighing systems make them less feasible. To address these challenges, farmers and researchers have explored alternative methods, such as using a measuring tape or string to estimate pig weight based on body measurements like heart girth and length (The Pig Site, nd). While these methods are cost-effective and less stressful, they are prone to inaccuracies due to human error and inconsistent measurement techniques.

Advancements in technology, particularly computer vision (CV) and machine learning (ML), have introduced more efficient and precise methods for weight estimation. Depth-sensing devices like Microsoft Kinect offer a non-invasive and cost-effective solution by capturing 3D data to estimate body weight with high accuracy (Pezzuolo et al., 2018). Compared to manual measurement methods or even traditional scales, these technologies significantly reduce labor, improve accuracy, and minimize stress on animals. Moreover, the integration of 3D imaging and ML algorithms enables the creation of automated systems that adapt to diverse farm environments, making them ideal for small and medium-sized farms (Gjergji et al., 2020).

Manual methods such as using measuring tapes involve wrapping a tape or string around the pig's heart girth and measuring its body length, followed by applying a formula to estimate weight (The Pig Site, nd). While accessible and inexpensive, this method is laborious and lacks precision, especially when dealing with active or uncooperative animals. Traditional weighing scales, though accurate, require direct handling of pigs, posing risks to animal welfare and farmworkers (Dickinson et al., 2013). In contrast, advanced systems leveraging depth-sensing technology provide non-invasive and automated solutions. For instance, Kinect sensors capture detailed 3D images, enabling precise calculations of body volume and weight without requiring physical contact with the animal (Pezzuolo et al., 2018). These systems mitigate stress, enhance accuracy, and save labor, making them a promising alternative to traditional methods. However, the initial cost and need for technical expertise can pose barriers to adoption, especially for resource-constrained farms.

Accurately estimating pig weight is critical for achieving optimal feed-to-weight conversion ratios, a key determinant of farm profitability. Weight monitoring also supports health assessments, helping detect early signs of disease or malnutrition, which can otherwise lead to significant economic losses. Furthermore, precise weight data allows farmers to determine the ideal market timing, maximizing revenue and ensuring meat quality standards are met (Terence et al., 2024).

This study aims to develop a system for pig weight estimation by leveraging Kinect V1's depth-sensing capabilities. The proposed system seeks to reduce labor, improve measurement precision, and enhance animal welfare. This research contributes to the broader adoption of sustainable, efficient farming practices by addressing the practical challenges of traditional and manual weight estimation methods. The findings may pave the way for integrating advanced depth-sensing technologies into modern livestock management systems, supporting the transition toward precision agriculture and smart farming practices.

1.2 Statement of the Problem

Accurate weight estimation is a cornerstone of efficient pig farming, influencing critical decisions such as feed management, health monitoring, growth tracking, and market readiness. Traditional weighing methods, which involve manually moving pigs to weighing scales, are not only labor-intensive and time-consuming but also stressful for the animals. This stress can adversely affect the pigs' well-being, reducing their productivity and potentially impacting meat quality (Li et al., 2014). The labor-intensive nature of manual weighing also increases operational costs, posing a challenge for farmers, particularly in small- to medium-scale farms where resources are limited.

To circumvent these challenges, farmers often resort to alternative methods, such as using a measuring tape or string to estimate pig weight. This method involves measuring dimensions like heart girth and body length, then applying a weight estimation formula (The Pig Site, nd). While this approach is cost-effective and non-invasive, it is prone to significant inaccuracies due to human error, variability in measurement techniques, and difficulty in handling active or uncooperative animals. These inaccuracies can lead to suboptimal feeding strategies, delayed health interventions, and missed opportunities for maximizing market profitability.

In recent years, advancements in imaging technologies have paved the way for automated, non-invasive weight estimation methods. Single-camera systems and monocular vision techniques, while offering some improvements, often fail to capture precise depth information, which is crucial for accurate volume and weight estimation (Pezzuolo et al., 2018); (Kollis et al., 2007). These systems are further limited by environmental factors such as lighting variations and animal movement, which compromise their reliability and usability in real-world farm settings.

Multi-camera setups have emerged as a more accurate alternative, enabling the capture of detailed 3D data necessary for precise weight calculations (Dohmen et al., 2022). However, these setups are often prohibitively expensive and complex, requiring specialized equipment, significant computational resources, and skilled personnel to operate. This makes them inaccessible to many farmers, especially those in resource-constrained environments. Additionally, while machine learning algorithms and other modeling techniques have been applied to refine weight estimation processes, their effectiveness is limited by the quality of input data, which often suffers from incomplete 3D reconstructions or inconsistencies in 2D images.

A promising yet underexplored solution lies in utilizing Microsoft Kinect V1, a cost-effective and readily available depth-sensing camera originally developed for gaming applications. The Kinect V1 captures depth data by projecting an infrared (IR) pattern onto the surface of an object and analyzing the distortions to calculate depth values. This enables the generation of detailed 3D point clouds that can be used to accurately measure body volume and, subsequently, estimate weight (Zhang, 2012). Despite its potential, the Kinect V1 remains underutilized in the domain of livestock management, and existing research has not fully explored its capabilities for precise and non-invasive weight estimation.

This gap in the literature highlights the need for a practical and accessible system that leverages Kinect's depth-sensing technology for weight estimation in pigs. Unlike traditional methods or multi-camera systems, the Kinect V1 offers a simpler, more affordable solution capable of capturing high-quality depth data in real-time. By automating the weight estimation process, the Kinect V1 can significantly reduce labor requirements, minimize stress on animals, and improve the accuracy of measurements.

This research aims to bridge this gap by leveraging the depth-sensing capabilities of the Kinect V1 to develop a cost-effective and automated system for pig weight estimation. Unlike previous approaches that rely on more complex and expensive setups, the Kinect-based system provides detailed depth information for accurate body volume measurements. By addressing the challenges of traditional and existing imaging methods, this study seeks to create a solution that minimizes labor, reduces stress on animals, and enhances measurement precision. The findings from this research could revolutionize weight monitoring practices in pig farming, particularly for small and medium-scale farms, and contribute to the advancement of precision agriculture and smart farming techniques.

1.3 Objectives of the Study

The main objective of this study is to develop a non-invasive Pig Weight Estimation System that utilizes depth imaging, specifically:

1. To utilize CNN and LightGBM for model training and validation to analyze depth data for pig weight estimation.

2. To develop a web application for the Pig Weight Estimation System.

3. To assess the accuracy of the Kinect-based pig weight estimation system.

1.4 Significance of the Study

The results of this study would benefit the following:

Academe. This study will contribute to the existing knowledge on the application of computer vision technology in the industry, particularly in livestock management. This study could serve as a valuable reference for academic institutions for agriculture, computer science, engineering, and computer vision technology. The findings of this study can also be an inspiration for further research in these fields of study.

Business Owners. This study can offer business owners, particularly those in the agricultural and livestock sectors a more efficient and data-driven method of managing their livestock operations. The results of this study can help automate tasks such as livestock monitoring, feed rationing, and health checks. This can reduce labor costs and improve decision-making processes. Additionally, this innovation can encourage entrepreneurship in tech-driven agriculture, creating business opportunities.

Livestock Caretakers. The results of this study can aid in the adoption of computer vision technology in livestock agriculture and can assist livestock caretakers in optimizing feed rations, monitoring growth rates, and improving overall herd management. Thus can ensure better animal health and welfare, reduce workload, and enhance productivity by providing data-driven insights into the daily operations of the livestock.

Government Organizations. Government organizations can utilize the study's findings to develop strategies that promote the modernization of agriculture through technological adoption. By using computer vision technologies, government organizations can improve farm-to-market systems and boost overall agricultural productivity.

Veterinarians. Results from this study could assist veterinarians involved in livestock by creating tools that enable veterinarians to remotely monitor, reducing the need for frequent on-site visits. It can also help in establishing preventive care strategies by being able to identify early signs of illness through variations in weight, improving animal welfare.

Agriculture Technology Developers. Companies and developers in the field of agricultural technology can use the findings of this study to design and enhance their products. The research can inform them about the specific needs of livestock management and provide insights into how computer vision solutions can be tailored to address these requirements, resulting in more market-relevant and effective innovations.

Future Researchers. This study can be used as a future reference for researchers who plan to engage in the same field of study. It can potentially be adapted for other agricultural applications, leading to increased efficiency and productivity in the sector. The methodologies found in the study can guide subsequent research, encouraging innovation and advancement in agriculture technology.

1.5 Scope and Limitations

This research focuses on the non-invasive estimation of weight in Landrace pigs, a breed commonly utilized in commercial pig farming in the Philippines due to its favorable traits and adaptability to the tropical climate (Mañez et al., 2020). Depth data for this study was collected using a Microsoft Kinect V1 camera sensor positioned at a fixed height of 1.9 meters from the floor, capturing a top-view perspective of the pigs (primarily showing their backs). The data was gathered in small backyard-type pig farms located in Cagayan de Oro City, Philippines. While data was collected with consistent lighting, this study will analyze depth data of Landrace pigs in various postures using Convolutional Neural Networks (CNNs). The Landrace pigs in this study typically range from 8 to 22 kg in weight, corresponding to an age of approximately 1 to 2 months and exhibiting an average daily gain (ADG) of 0.7–1.0 kg/day under optimal feeding conditions.

This study will utilize CNNs to analyze depth images, captured from a top-down view at a consistent height of 1.9 meters, and extract precomputed features, including:

- **Pixel Size**: Total pixels in the segmented region ($shape[0] \times shape[1]$).
- **Non-Zero Pixel Count**: Pixels with $depth > 0$ ($np.count\_nonzero$).
- **Pixel-to-Non-Pixel Ratio**: Ratio of non-zero to total pixels.
- **Standard Deviation of Depth**: Depth variability ($np.std$).
- **Mean Depth**: Average depth of non-zero pixels ($np.mean$ where $depth > 0$).
- **Pixel Ratio**: Non-zero pixel count relative to 1280×480.
- **Volume Proxy**: Sum of depth values ($np.sum$), approximating 3D volume.
- **Aspect Ratio**: Bounding box width-to-height ratio ($shape[1]/shape[0]$).
- **Perimeter**: Sum of contour arc lengths ($cv.arcLength$).

The primary goal is to train a CNN model capable of estimating pig weight based on these 3D body volume measurements derived from the depth images captured by the Microsoft Kinect V1 sensor from a fixed overhead position.

For the purpose of evaluating the system's real-time accuracy in a controlled environment, a local server implementation will be used. This approach allows for focused testing and immediate feedback on the model's performance without the complexities of a wider network deployment or internet dependency, which can introduce additional variables and potential points of failure during this initial validation phase. The current study's weight estimation will be limited to the 8-22 kg range due to the data collected under these specific conditions. However, the developed model will serve as a foundational step for future research aimed at scaling the system to accommodate a wider range of pig weights and potentially exploring different camera angles or deployment in more diverse farm environments.

1.6 Definition of Terms

Artificial Intelligence - A set of technologies that enable computers to perform a variety of advanced functions, including the ability to see, understand, and translate spoken and written language, analyze data, make recommendations, and more.

Computer Vision A field of computer science that focuses on enabling computers to identify and understand objects and people in images and videos.

Deep Learning A subset of machine learning that uses multilayered neural networks, called deep neural networks, to simulate the complex decision-making power of the human brain.

Depth Imaging A technique that captures the distance between a sensor (like a depth camera) and objects in its view, creating a 3D representation of the scene. Depth imaging provides additional data on the spatial positioning of objects.

Landrace A domesticated, locally adapted, traditional variety grown by farmers and their successors since ancient times.

Image Processing Techniques used to enhance or extract information from images, are crucial for analyzing data from the Kinect sensor.

Machine Learning A subset of artificial intelligence that involves the use of algorithms to learn from data and make predictions or decisions based on that data.

Microsoft Kinect Sensor A motion-sensing input device that uses a camera to capture depth data and track movements.

Pig Weight Estimation The process of determining the weight of pigs uses various methods, in this case, image processing and machine learning.

CHAPTER 2

REVIEW OF RELATED LITERATURE

This literature review explores current weight estimation techniques in livestock management, highlighting the advancements and limitations of traditional methods while introducing emerging technologies, particularly depth-sensing devices like the Kinect sensor. Understanding the shift from conventional approaches to tech-driven solutions provides insight into how these innovations can improve accuracy, efficiency, and scalability in livestock farming.

2.1 Weight Estimation Techniques

In animal production, livestock body weight is a significant and widely used feature that has an impact on feed consumption, breeding potential, social behavior, energy balance, and overall farm management (Wang et al., 2024). It may be used indirectly in the assessment of health and welfare status (Dikmen et al., 2012). There are two main approaches to measuring body weight in livestock: (1) direct approaches using scales, and (2) indirect approaches based on relationships between body part measurement and body weight.

Direct weighing methods rely on weighing technologies such as partial-weight or full-weight industrial scales capable of supporting small, medium, or large livestock. Some companies provide passive-weighing solutions that integrate sensor-rich scaling systems such as GrowSafe of Canada, the Bosch Precision Livestock Platform of Germany, and Rice Lake Weighing Systems of Australia which are capable to measure, log results, and transmitting information over wired or wireless networks (Wang et al., 2024) . Other companies such as Arvet CIMA Control Pig and CIMA Control Cow Scaling Systems of Spain developed custom-made scales that provide dynamic-weighing systems where animals are weighed while in motion using walk-through or step-over weighers (Rousing et al., 2004). While these devices are very accurate, their acquisition, intended purpose and operation size, repeated calibration and maintenance costs associated with their placement in high-temperature variability, and corrosive environments are significant and beyond the affordability and sustainability limits of small and medium size farms and even commercial operators(Dikmen et al., 2012). It has been studied that removing animals from paddocks and holding areas and leading them to weighing stations is costly, stressful, and potentially harmful activity for animals and handlers alike and also inadvertently leads to animal weight loss or even death (Faucitano and Goumon, 2018). Moreover, since the weighing process is very laborious, the frequency of measurements is not sufficiently high to permit the use of body weight as an indicator for other traits. However, since the affordability of direct weighing methods may impede small producers (Dickinson et al., 2013), researchers have developed indirect weighing methods represented by regression models that relate morphometric measurements and image features to body weight in livestock. The direct acquisition of morphometric measurements can be accomplished with the aid of technologies with various degrees of complexity, from measuring tapes and types to specialized software or manual, semi-automatic, or automatic measurements extrapolated from images obtained with electro-optical devices such as mono-2D, stereo-2D, 3D, ultrasound, and infrared sensors (Wang et al., 2024).

2.1.1 Role of Technology in Agriculture

Technological advancements have played a pivotal role in transforming the agricultural sector. From mechanization to digital innovations, these technologies have increased efficiency, productivity, and sustainability in farming practices. In particular, precision agriculture, which uses technology to optimize field-level management of crop farming, has revolutionized the way food is produced (Witten et al., 1993). Precision farming technologies include GPS, soil sensors, and drones, allowing farmers to monitor and manage their crops with unprecedented accuracy. In livestock farming, technology has led to the development of automated feeding systems, health monitoring tools, and weight estimation systems that reduce labor and enhance animal welfare (Gómez et al., 2021). For example, sensor-based systems for monitoring livestock health provide farmers with real-time data, enabling proactive interventions to prevent diseases and improve overall herd management (Neethirajan and Kemp, 2021).

Furthermore, AI and machine learning have started to play an increasingly significant role in agriculture. These technologies enable predictive analytics for crop yields and disease outbreaks, and more recently, they are being used in 3D object detection systems for animal weight estimation, such as the Kinect-based system for pig weight estimation proposed in this study (Farooqui, 2024). Integrating AI and machine learning into agricultural practices is expected to address many challenges related to food security, sustainability, and labor shortages (Ng et al., 2023).

For example, one study explored the use of 3D images captured from a zenithal viewpoint to estimate lambs' live weight. The researchers applied image processing techniques to extract features such as upper body area and average depth, demonstrating the potential of 3D imaging for livestock weight estimation (Samperio et al., 2021). Although this study did not use Kinect, it highlights the value of 3D data, which Kinect is well-equipped to provide, for accurate livestock weight estimation.

Another significant study used the Microsoft Kinect V1 depth camera to measure pig body dimensions and estimate their weight. The researchers found a strong correlation between the Kinect-based measurements and actual weights, with coefficients of determination (R^2) exceeding 0.90 (Pezzuolo et al., 2018). Similarly, Lao et al. also employed a Kinect V1 depth camera to extract body measurements from pigs, developing a regression model for weight estimation, while Liu et al. used a binocular vision system to collect 3D data and tested various modeling approaches, including linear, nonlinear, and machine learning algorithms, to estimate pig weight (Li et al., 2014).

2.2 Technological Frameworks

The manual techniques used in the majority of livestock weight estimation systems today have given way to more automated, technologically advanced systems. Physical weighing scales, which were labor-intensive, time-consuming, and frequently uncomfortable for animals, were the foundation of traditional methods. On the other hand, cutting-edge image technologies are now used in modern systems, including 3D depth sensors like the Microsoft Kinect and binocular and monocular vision systems. For extremely precise

3D reconstructions of animals, binocular vision systems employ two cameras to collect stereo images; however, the expense and complexity of these installations may make them impractical for small farms (Rousing et al., 2004). Compared to binocular vision systems, monocular vision systems, which use single-camera setups, are more affordable but have worse precision.

Smart sensors are at the vanguard of revolutionizing precision agriculture by giving farmers access to real-time information on vital parameters including plant health, temperature, humidity, and soil moisture. With their sophisticated detection methods, these sensors assist farmers in making well-informed decisions that increase crop output. Farmers may automate and optimize tasks like nutrient application and irrigation scheduling by combining these sensors with Internet of Things (IoT) devices and artificial intelligence (AI). By giving real-time data on vital elements like soil moisture, temperature, humidity, and plant health, smart sensors are transforming precision agriculture. Farmers can better control fertilizers, optimize irrigation, and monitor crop conditions by integrating these sensors with IoT and AI technologies. This contributes to resource conservation, increased crop yields, and addressing environmental issues (Soussi et al., 2024). In general, smart sensors assist in addressing issues related to global agriculture, including depletion of resources, climate change, and rising food production demands.

In general, smart sensors help address problems associated with global agriculture, such as resource depletion, climate change, and increased need for food production. Furthermore, these sensors have started to become extremely important in cattle farming. For example, smart sensors assist farmers in ensuring the comfort and welfare of their livestock by keeping an eye on environmental factors like temperature and air quality in animal housing. Thus, healthier animals and lower veterinary expenses can be achieved by preventing heat stress and disease outbreaks(Terence et al., 2024).

The physical characteristics of cattle, such as weight and body dimensions, can be monitored non-invasively using Computer Vision (CV)-Based Sensors, which are frequently combined with AI algorithms. Since no direct contact is required, methods like RGB picture analysis and 3D point cloud offer great accuracy and can lessen animal suffering (Ma et al., 2024). These technologies have the potential to greatly increase the accuracy of data gathered for monitoring, enabling better cattle care and management. However, small-scale farms face difficulties due to the complexity of establishing computer vision systems, which sometimes require pricey hardware and intricate data processing (Terence et al., 2024). High initial investments are needed for computer vision systems, especially those that use 3D cameras and intricate algorithms. Particularly for smaller farms, the expense of high-quality technology (such as RGB cameras or 3D scanners) and the required computer infrastructure can be prohibitive.

CV-based systems not only provide budgetary difficulties but also demand specialized staff to properly run and maintain the technology. Farmers may find it difficult to handle system calibrations and solve problems if they lack technical skills. This could lead to erroneous data collecting and less-than-ideal herd management results. Furthermore, the accuracy of the system may be reduced in real-world situations due to environmental conditions like dust, lighting, or even the movement of the animals that impair sensor performance (Ma et al., 2024). Research is being done to make these technologies more resilient and affordable as they develop, so smaller operations can use them. Nonetheless, these technologies have a great deal of promise to improve livestock monitoring by lowering stress and increasing management effectiveness, even in the face of obstacles.

2.3 Image Processing and Machine Learning for Weight Estimation

2.3.1 Historical Development

In traditional livestock management, weight estimation was primarily performed manually using techniques such as Body Condition Scoring (BCS). This method evaluates the fatness or thinness of animals through visual inspection and tactile assessment (Bercovich et al., 2013). While straightforward and cost-effective, BCS relies heavily on the evaluator's expertise, introducing variability and potential bias in results. Additionally, this method does not provide precise numerical weight values, making it less reliable for data-driven decision-making in modern farming practices.

Before the introduction of digital technologies, farmers used scales or manual measurements to estimate livestock weight. With scales, animals were weighed directly, but this required significant infrastructure, such as livestock-specific weighing platforms. These setups were costly, labor-intensive, and often stressful for the animals, impacting their health and productivity.

For situations where scales were unavailable, farmers traditionally estimated pig weight using a tape measure and mathematical formulas. This approach involved measuring specific dimensions of the pig, such as its girth and length, and applying an empirical formula to estimate its weight. For example, pig weight can be approximated using the formula:

$$Weight(kg) = \frac{Girth(cm)^2 * Length(cm)}{400} \tag{1}$$

Where *Girth* is the circumference of the pig's chest behind the front legs, and *Length* is the distance from the base of the tail to the middle of the ears.

This method, though simpler than using scales, is prone to human error and variability in measurements. Despite its limitations, it remains popular, especially in rural farming contexts in the Philippines, where such techniques are both accessible and practical

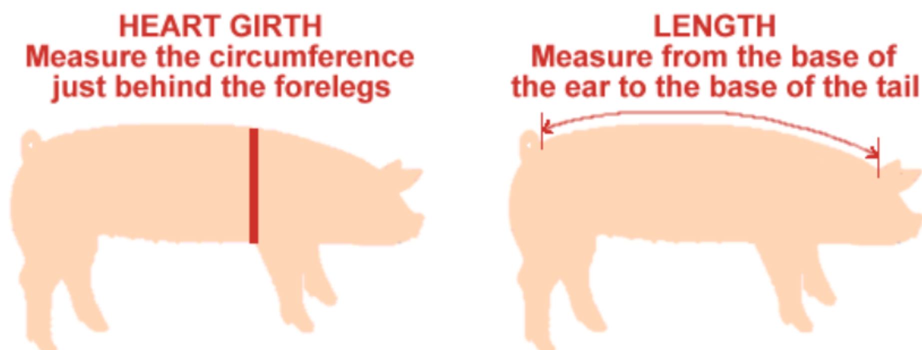

Figure 1: Illustration of weight estimation using a tape measure

2.3.2 Modern Techniques

Body weight (BW) prediction in livestock can be modeled using four main approaches of increasing complexity: Traditional Approach, Computer Vision Approach, Computer Vision with Machine Learning, and Computer Vision with Deep Learning. The Traditional Approach relies on manually collected morphometric measurements such as heart girth, wither height, and body length, which are then applied in traditional regression models. This method has been widely used for species such as cattle, pigs, sheep, goats, camels, and yaks (Franco et al., 2017); (Fadlelmoula et al., 2020); (Yan et al., 2019). While effective, it is labor-intensive and causes stress to the animals. To mitigate these issues, the Computer Vision (CV) Approach uses 2D and 3D electro-optical sensors, such as RGB or Kinect cameras, to capture images for morphometric measurements (Ozkaya, 2013). Although 3D cameras improve precision, they are expensive and require complex data processing. The CV with Machine Learning (CV+ML) Approach enhances the CV method by automating feature selection with machine learning techniques, although some manual processes, such as image segmentation and morphometric extraction, remain necessary (Tasdemir and Ozkan, 2019). Finally, the CV with Deep Learning (CV+DL) Approach uses deep learning models, including convolutional neural networks (CNNs) and recurrent convolutional networks (RCNNs), to fully automate the BW prediction process (Gjergji et al., 2020). While this approach has shown significant improvements, challenges still exist in precisely segmenting animals from complex backgrounds (Shukla and Anand, 2016). These approaches highlight the evolution of BW prediction models, transitioning from manual methods to fully automated systems, with deep learning offering promising advancements.

2.3.3 Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNNs) are a class of deep learning models specifically designed to process data with grid-like topologies, such as images. These networks employ a series of convolutional operations using trainable filters to detect local patterns (e.g., edges, textures, and shapes), which are then aggregated through pooling operations to reduce dimensionality while preserving essential information. At the final stages, fully connected layers interpret the abstracted features to generate predictions or classifications (LeCun et al., 2015).

What sets CNNs apart from traditional machine learning algorithms is their ability to automatically learn and optimize feature representations directly from raw pixel data, eliminating the need for handcrafted features and domain-specific input preprocessing. This makes them particularly suitable for complex image recognition and regression tasks where spatial hierarchies and relationships are crucial.

In the context of agriculture and livestock management, CNNs have been successfully deployed for various applications including species classification, posture detection, health monitoring, and behavioral analysis. One of the most promising applications of CNNs in this domain is livestock weight estimation, where non-invasive methods are favored over traditional weighing scales due to concerns about animal welfare, labor intensity, and logistical constraints.

Recent research has demonstrated the effectiveness of CNNs in predicting pig weight from overhead RGB images. For instance, Pinto et al. (2024) developed a CNN-based system that achieved a mean absolute percentage error (MAPE) as low as 1.35%, even under varying lighting conditions and pig postures. This performance highlights CNNs' robustness in real-world environments, especially when paired with top-down image acquisition methods using cameras like the Microsoft Kinect. These models are trained to associate pixel-based spatial cues—such as body shape and area—with actual weight, achieving high accuracy without the need for manual measurements or handling.

Moreover, CNNs are well-suited for integration with depth image data, which is particularly beneficial in weight estimation tasks. Depth images add a third spatial dimension, enabling CNNs to interpret volume-related features more effectively. In the context of this study, the depth information captured using Kinect V1 enhances the CNN's capacity to infer body mass with greater precision by leveraging 3D structural information that 2D RGB images may not fully convey.

2.3.4 CNN in Image Processing

In the domain of image processing, Convolutional Neural Networks (CNNs) have become the foundational architecture for most modern computer vision applications. Their design, inspired by the visual cortex of animals, enables them to extract multiscale features by applying convolutional filters over local image regions. These features range from simple patterns like edges and corners in the early layers to more complex textures, shapes, and object structures in deeper layers (Krizhevsky et al., 2012). This hierarchical feature extraction mechanism allows CNNs to perform exceptionally well on tasks such as image classification, segmentation, and object detection.

Unlike traditional image processing techniques that rely on handcrafted feature descriptors like SIFT or HOG, CNNs automatically learn relevant features from training data, which significantly improves accuracy and adaptability in diverse real-world scenarios. Their generalization capability has led to widespread use in industries ranging from healthcare and security to autonomous driving and agriculture.

In the context of livestock farming, CNNs have proven to be powerful tools for non-invasive visual analysis. They are commonly used for posture recognition, gait analysis, behavioral monitoring, and health condition assessment of animals. For example, CNN-based models have been applied to segment animals from backgrounds, detect body landmarks, and classify behavior patterns in cattle, poultry, and pigs (Andrási et al., 2023; Wang et al., 2021). This automation not only reduces the reliance on manual observation but also enables real-time monitoring, which is critical for early disease detection and animal welfare.

When paired with depth imaging technologies such as Microsoft Kinect, CNNs can process 3D spatial data, making them even more suitable for precision livestock applications. Depth images provide additional information about the spatial configuration and volume of animal bodies, which 2D RGB images alone cannot offer. CNNs trained on depth data can thus extract volumetric features essential for applications like weight estimation, where accurate assessment of body mass is required without direct contact.

Moreover, CNNs have been integrated into real-time farm monitoring systems, leveraging top-down camera views to continuously track animal growth. Studies have shown that CNNs can reliably identify individual pigs and estimate physical parameters such as body area and length, even in cluttered or low-contrast environments (Pinto et al., 2024). The use of depth maps enhances the segmentation and feature extraction process, particularly when lighting or posture varies, conditions commonly found in open or semi-controlled farm settings.

2.3.5 Light Gradient Boosting Machine (LightGBM)

Light Gradient Boosting Machine (LightGBM) is a highly efficient and scalable machine learning framework based on gradient boosting decision trees (GBDT). It was developed to overcome the computational limitations of traditional boosting algorithms, particularly when dealing with large-scale data and high-dimensional feature spaces. LightGBM adopts a novel leaf-wise tree growth strategy with depth limitations, which selects the leaf with the highest gain to grow, rather than the level-wise approach used in conventional GBDT models. This technique improves both the speed and accuracy of the model, while also effectively mitigating the risk of overfitting (Ke et al., 2017).

One of the defining advantages of LightGBM is its ability to handle large datasets with numerous features and missing values efficiently. It supports both categorical and continuous variables, and it performs well with skewed and sparse datasets—common conditions in agricultural image analysis where some features may dominate while others contribute noise.

In the context of animal agriculture, particularly for weight estimation tasks, LightGBM is often employed to analyze structured numerical data extracted from image or depth data. These features can include geometric measurements such as pixel size, volume proxy, standard deviation of depth, aspect ratio, and perimeter—all of which correlate with body mass. When combined with depth-sensing technology like the Microsoft Kinect V1, such features provide rich, non-contact input for predicting animal weight.

Recent studies have validated LightGBM's effectiveness in livestock applications. For instance, in pig farming scenarios, features derived from top-view depth images—such as pixel count, mean depth, and shape ratios—were used as inputs to LightGBM regressors to estimate the live weight of pigs (Pezzuolo et al., 2018). The models achieved high coefficients of determination (R^2), often exceeding 0.95, indicating a strong correlation between image-derived features and actual body weight. Such results emphasize LightGBM's capability to model nonlinear relationships inherent in biological data.

2.3.6 Applications of Kinect in Agriculture

In livestock farming, Kinect sensors are used to monitor animal behavior, health, and growth through 3D imaging. This technology allows for early detection of health issues by analyzing subtle changes in posture, gait, and movement patterns. It can identify lameness in cattle early on, allowing for prompt intervention (Singh et al., 2022). Moreover, Kinect's non-invasive methods provide automated weight estimation by capturing 3D images, reducing the need for stressful manual weighing processes. Continuous monitoring helps track growth and optimize feeding schedules, contributing to the overall welfare of livestock. In addition to health monitoring, Kinect sensors contribute to environmental control within livestock facilities, tracking variables like temperature and humidity to ensure optimal conditions. In poultry farming, Kinect can monitor flock movement patterns, alerting farmers to potential issues like overcrowding, which could lead to health or productivity problems.

The integration of Kinect into Precision Livestock Farming (PLF) systems further enhances its value, as it collects real-time data alongside other technologies like RFID and GPS. This combination allows for more efficient resource management, improved reproductive tracking, and more precise feeding practices (Monteiro et al., 2021). Kinect technology is also being used in crop management. Its 3D imaging capabilities are beneficial in monitoring plant growth, detecting diseases early, and optimizing resource use like water and fertilizers. In automated harvesting, Kinect sensors guide robots through fields, identifying ripe crops based on their size and shape. This reduces the need for manual labor and ensures timely harvesting to maximize yield and quality (Singh et al., 2022).

2.4 Microsoft Kinect V1

The Kinect V1 sensor employs an infrared (IR) structured light system to capture depth information. It consists of three main components: an RGB camera, an IR projector, and an IR sensor. The RGB camera captures standard color images, while the IR projector emits a structured pattern of infrared light onto the environment. The IR sensor detects distortions in the projected pattern caused by objects in the scene, enabling the device to compute depth information based on the time and spatial displacement of the reflected IR light (Zhang, 2012). This technique provides a depth resolution of 640x480 pixels, with a frame rate of up to 30 frames per second, making it suitable for capturing detailed spatial data.

2.4.1 Depth Data Acquisition and Processing

The Kinect V1 produces a 3D point cloud by combining depth and RGB data, representing the spatial coordinates (X, Y, Z) of each pixel in the captured scene. This depth information is particularly useful for applications requiring precise measurements of object size, shape, and volume. The sensor's effective depth range is between 0.8 and 4.0 meters, allowing it to capture accurate data within typical livestock pen dimensions (Smisek et al., 2013). By leveraging its depth-sensing capability, researchers can measure the dimensions of objects without physical contact, making it ideal for non-invasive weight estimation in pigs.

2.4.2 Kinect V1 Applications in Agriculture

The Kinect V1 has been successfully utilized in agricultural studies for tasks such as plant phenotyping, fruit grading, and livestock monitoring. For instance, (Andújar et al., 2016) used the Kinect V1 to evaluate crop height and canopy structure, demonstrating its utility in capturing detailed morphological data. Similarly, (Pezzuolo et al., 2018) employed the Kinect V1 to estimate cattle body volume and weight, achieving high accuracy compared to traditional methods. These applications highlight the potential of the Kinect V1 as a versatile tool for precision agriculture.

2.4.3 Kinect V1 Potential for Pig Weight Estimation

In pig farming, accurate weight estimation is crucial for optimizing feeding strategies and monitoring growth. Traditional weighing methods, such as scales or manual measurements, can be labor-intensive, stressful for animals, and prone to errors (Faucitano and Goumon, 2018). The Kinect V1 offers a non-invasive alternative by capturing detailed depth images that enable the calculation of pig body volume. By processing these depth images using machine learning algorithms, it is possible to develop a system that estimates pig weight based on 3D body volume, eliminating the need for physical contact.

The Kinect V1's structured light technology ensures that depth measurements are unaffected by variations in visible light conditions, making it suitable for use in indoor pig pens where lighting may vary. Additionally, its ability to provide real-time depth data opens the possibility for developing dynamic weight monitoring systems that can track pig growth continuously.

2.5 Synthesis Matrix

The synthesis matrix provides valuable insights into pig weight estimation methods using various imaging and analysis techniques. Together, these studies provide a cohesive foundation for applying Kinect technology, advanced imaging, and regression methods in the current study, demonstrating approaches for accurate, adaptable, and practical pig weight estimation.

Table 1: Synthesis Matrix

Title	**Author(s), Year**	**Methodology**	**Findings**	**Relevance**
On-barn pig weight estimation based on body measurements by a Kinect v1 depth camera	Pezzuolo, A., Guarino, M., Sartori, L., González, L. A., and Marinello, F. (2018)	- Kinect v1 Depth Camera - Estimated pig weight from body measurements - Captured 3D images in a barn environment - Extracted length, width, and volume from depth data - Developed a regression model for weight estimation	Both linear and second-degree regression models showed strong correlations with reference weights, with coefficients of determination above 0.95. The non-linear model reduced the standard error by half, and the second-degree regression model had an absolute error of less than 0.5 kg.	The current study is closely correlated with this study, especially in terms of the usage of the Kinect v1 camera. This study, however, uses 3D images for the calculation of estimated weights.

Pig Weight Estimation According to RGB Image Analysis	Andras Kárpinszky, Gergely Dobsinszki(2023)	- RGB Cameras (Dahua Models) - RGB image captured above pig pens - Mask R-CNN for segmentation - Kalman filters for tracking - Pretty Contour Picker (PCP) for filtering - Weight estimation using Multi-Layer Perceptron (MLP)	The system achieved more than 97 percent accuracy in predicting pig weights compared to manually recorded weights. Among the models tested, Model V2 was the most consistent, providing high accuracy across varying weight ranges. The RGB image-based method allows faster and stress-free weight measurement, which is valuable for decision-making in pig farming	Although a Kinect camera wasn't used in the study, the setup for data acquisition and processing is similar. Top-view photos of pigs were also used in this study, closely relating to how the current study takes them.

Estimating Pig Weights from Images without Constraint on Posture and Illumination	Kyungkoo Jun, Si Jung Kim, Hyun Wook Ji (2018)	- 2D Image Processing - No relaxed posture and illumination constraints - Image processing: binarization, morphological ops, contour analysis - Features: area size, curvature, deviation - Neural network model trained and tested for weight prediction	The study achieved an average estimation error of 3.15 kg and a coefficient of determination (R^2) of 0.792. Despite this being lower than previous works, the method was able to estimate pig weights without controlling the environment, posture, or lighting, making it applicable in less constrained settings. The model showed that posture-related features contributed significantly to weight prediction accuracy.	The study utilized a 2D camera, meanwhile the current study will utilize a Kinect v1 camera. Despite the difference, both studies have similar top-down view camera setups. The idea of estimating pig weights without constraints on posture and illumination can also be an inspiration for future similar undertakings with the Kinect v1 camera.

Carcass Quality Traits of Fattening Pigs Estimated Using 3D Image Technology	A. Peña Fernández, T. Norton, E. Vranken, D. Berckmans (2019)	- Kinect 3D Cameras - Capture 3D top-view images pre-slaughter. - Extract image features: lengths, areas, and volumes. - MATLAB for stepwise linear regression analysis	The regression models achieved an adjusted R^2 ranging from 70-85 percent during training, but performance decreased to 50-60 percent during validation. The best correlations with slaughter traits, such as final weight and yield, were found using the median values of image features from the last week of the fattening period.	Considering only the sections relevant to pig weight estimation, the study can be used as a reference for the current study given that it also uses Kinect cameras for image capture. The methods for feature extraction and pig weight estimation and analysis can be used as one of the bases for the methods that will be used in the current study.

The synthesis matrix provides valuable insights into pig weight estimation methods using various imaging and analysis techniques. The first study by (Pezzuolo et al., 2018) effectively demonstrates using a Kinect v1 depth camera to capture 3D body measurements, achieving high prediction accuracy with regression models, particularly a second-degree regression model that reduced error to under 0.5 kg. This finding supports the current study's use of Kinect v1 for accurate weight estimation via 3D imaging.

In another study, (Kárpinszky and Dobsinszki, 2023) employed RGB cameras combined with segmentation (Mask R-CNN) and MLP models, achieving over 97 percent accuracy in weight prediction. Although RGB cameras differ from Kinect, the segmentation and tracking methods align well with the current study's top-view imaging approach and offer potential techniques for data processing in the project.

The work of (Jun et al., 2018) further broadens applicability by achieving weight estimation without controlling for pig posture or lighting, using a neural network model on 2D images with a 3.15 kg error margin. This flexibility offers useful insights into handling environmental variability, which could enhance the robustness of your Kinect-based approach.

Lastly, (Peña Fernández et al., 2019) applied Kinect-based 3D imaging for carcass trait prediction, with results showing time-evolving feature accuracy between 50-85 percent. Their regression techniques could be useful for refining the project's model to predict weight accurately. Together, these studies provide a cohesive foundation for applying Kinect technology, advanced imaging, and regression methods in the current study, demonstrating approaches for accurate, adaptable, and practical pig weight estimation.

CHAPTER 3

METHODOLOGY

This chapter presents the research methodology used in the study, it covers the in-depth details of the procedures and steps taken in the research design, design procedure, system development, and deployment process of this study.

3.1 System Architecture

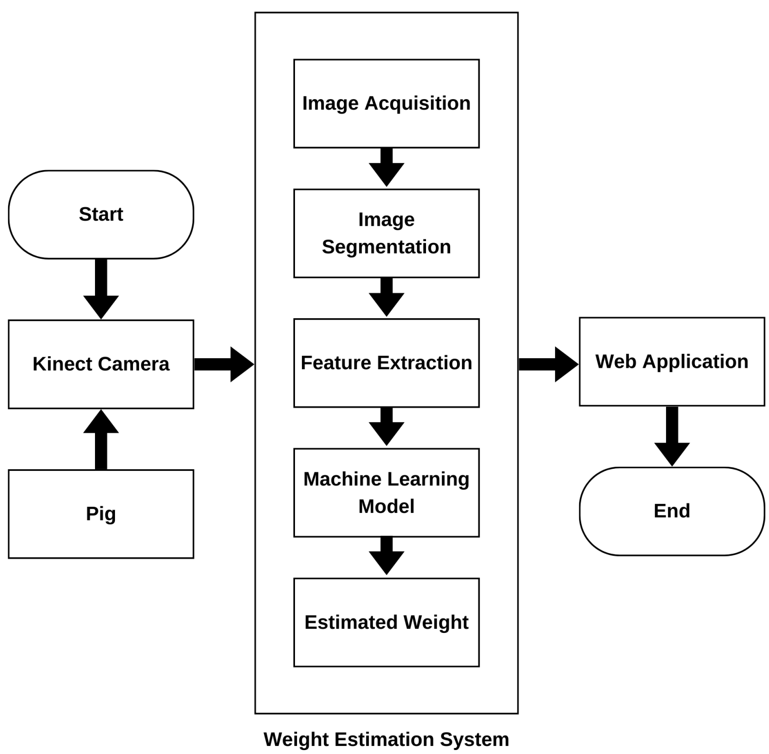

Figure 2: System Architecture

Figure 2 illustrates the system architecture of the pig weight estimation system. The system comprises the Landrace pig to be captured, the Microsoft Kinect V1 camera for acquiring depth and RGB images, and a processing system for data analysis. A laptop or dedicated server processes the data, performing image preprocessing, feature extraction (including 9 features such as pixel size, volume proxy, and sectional volumes), and weight estimation using a hybrid LightGBM-CNN model. The results are accessible through a FastAPI-based web service and a React-based web application, enabling real-time validation of actual and estimated weights. The architecture integrates hardware and software components, optimized for small-scale pig farms.

3.1.1 Waterfall Model

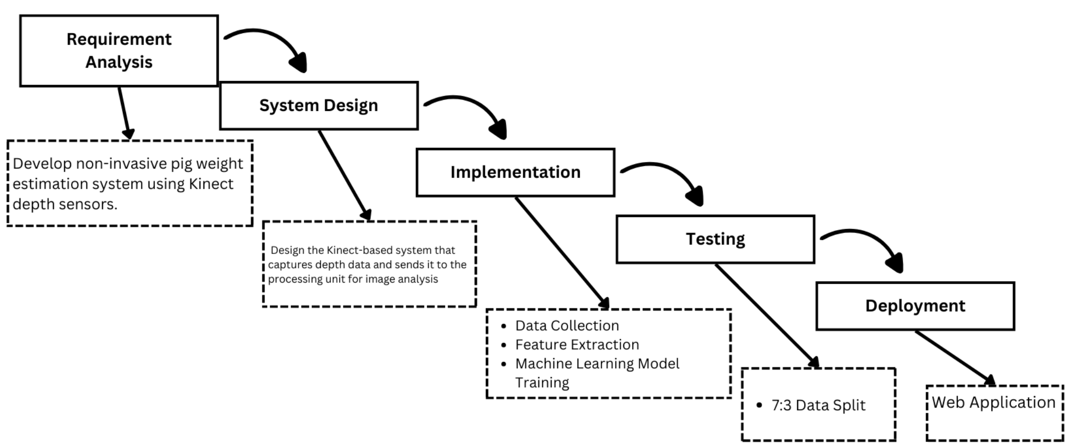

Figure 3: Waterfall Model

Figure 3 depicts the Waterfall Model used for system development, following a linear, sequential approach:

- **Requirement Analysis**: The goal is a non-invasive pig weight estimation system using a Kinect V1 depth sensor to capture depth images, extract 9 features (Section 3.7), and predict weight with a hybrid LightGBM-CNN model, accessible via a web application.

- **System Design**: The architecture includes hardware (Kinect V1, computer, pigpen) and software (YOLOv11, SAM, LightGBM-CNN, FastAPI, React) components, designed for compatibility and scalability.

- **Implementation**: Involves collecting depth and RGB images, extracting features, and training the hybrid model. Image processing includes grayscale conversion, contour detection, PCA-based alignment, and body part segmentation. The dataset follows an 80:20 training-validation split.

- **Testing**: Evaluates accuracy using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) on a test set (20% of data), selecting the most accurate model for deployment.

- **Deployment**: The system is deployed via a FastAPI web service and React web application on a local server, enabling real-time weight estimation.

The Waterfall Model ensures a structured development process with clear milestones.

3.1.2 Flowchart

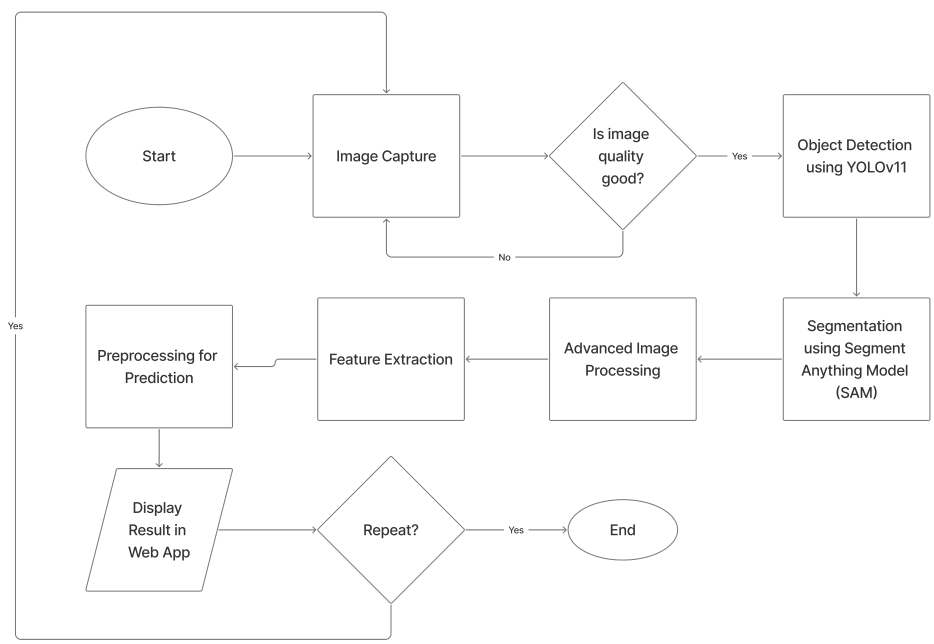

Figure 4: Flowchart

Figure 4 outlines the flowchart of the weight estimation system, detailing the end-to-end process from initialization to real-time weight prediction:

- **System Initialization**: The Kinect V1 camera, computer system, and pigpen are powered on, and the pig is prepared for imaging. The camera is mounted at 1.9 meters, ensuring a clear top-down view of the pig's back.

- **Realtime Video Capture**: The Kinect V1 captures synchronized RGB and depth video streams at 1280×480 pixels, 30 frames per second, which translates to 30 images per second. A high-quality image is defined as clear, well-lit, free of motion blur, and accurately representing the pig's dimensions without occlusion from other pigs.

- **Object Detection**: YOLOv11 processes the RGB stream to detect pigs, generating bounding boxes for each individual, which are mapped to the depth stream for accurate data extraction.

- **Segmentation**: The Segment Anything Model (SAM) isolates each pig within its bounding box, creating binary masks to remove background noise from the depth image.

- **Advanced Image Processing**: The depth image undergoes depth normalization, contrast enhancement, and optional data augmentation to enhance the quality and consistency of the image.

- **Feature Extraction**: Eleven features are extracted, including pixel size, non-zero pixel count, pixel-to-non-pixel ratio, standard deviation of depth, mean depth, pixel ratio, volume proxy, aspect ratio, and perimeter for each body part.

- **Preprocessing for Prediction**: Each of the 30 depth images extracted per second is resized to 40×80 pixels, thresholded (values $\leq$1 set to 0), and normalized to [0, 1]. The mean of the features are taken, min-max normalized, and formatted for model input.

- **Weight Prediction**: The preprocessed depth image and features are sent to the FastAPI /predict/ endpoint. The LightGBM model processes the 9 features for an initial prediction, which is concatenated with CNN features (from the depth image) to produce the final weight prediction.

- **Result Display**: The predicted weight is returned via the FastAPI API and displayed on the React-based web application, showing the estimated weight alongside confidence indicators for user validation.

- **Iteration**: The process repeats for continuous real-time estimation, capturing new frames as pigs move within the pigpen, ensuring only high-quality frames are processed.

The flowchart ensures a robust, automated pipeline for real-time weight estimation, integrating seamlessly with the deployed system.

3.2 Experimental Setup

3.2.1 Pigpen Diagram

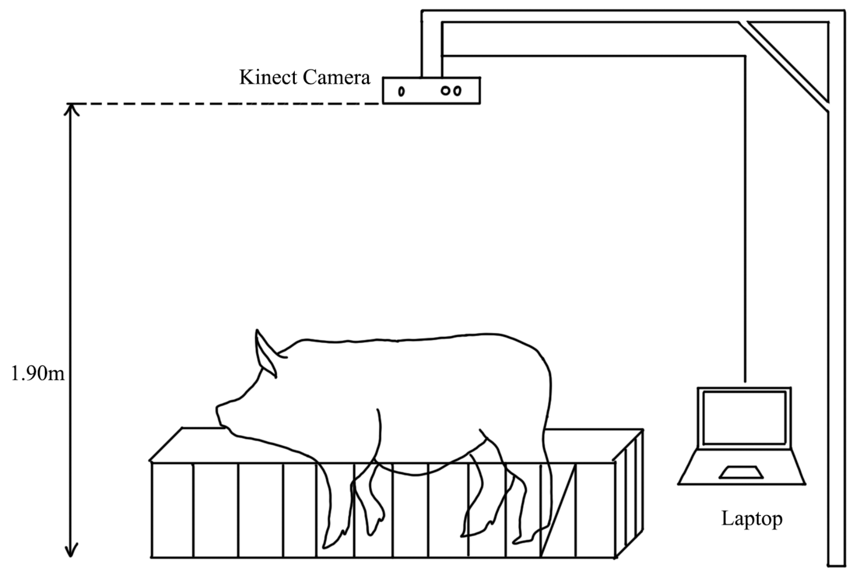

Figure 5: Pigpen Diagram

Figure 5 illustrates the pigpen setup, with the Microsoft Kinect V1 camera mounted 1.9 meters above the ground for a top-down view. This height ensures clear visibility of the pigs' backs within the Kinect's depth-sensing range (0.8–4.0 meters). The pigpen accommodates multiple pigs, with spacing to minimize occlusion, ensuring accurate volume estimation. The setup captures depth data for multiple pigs in a single frame while preserving individual measurements.

3.2.2 Camera Configuration

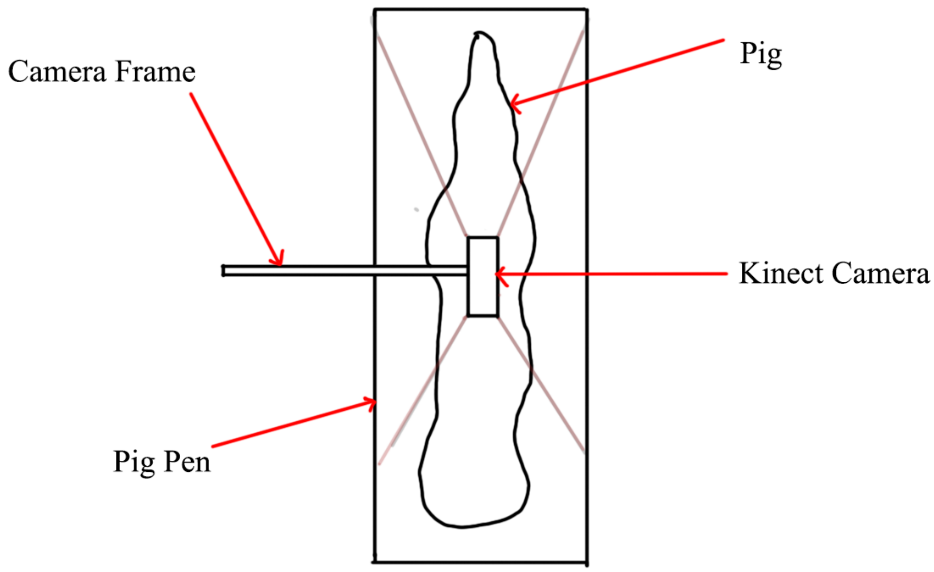

Figure 6: Camera Configuration

The Kinect V1 camera captured RGB and depth streams at 640×480 pixels per video, side by side. The depth sensor was calibrated for its effective range, with the 1.9-meter height optimized for detailed body contour capture. A dedicated server processed and stored data in real time, ensuring minimal latency and high data integrity.

3.3 Study Design and Data Collection

The study was conducted in small, backyard-type pig farms in Cagayan de Oro City, Philippines, focusing on Landrace pigs, known for adaptability and rapid growth. The pigs, weighing 8–22 kilograms and aged 1–2 months, were raised under optimal conditions with high-protein diets, clean water, and regular health monitoring.

Depth data was collected using the Kinect V1 camera, positioned to capture a top-down view of the pigs' backs, standardizing data collection. The system recorded RGB and depth streams simultaneously, as shown in Figure 7. The RGB stream was processed by YOLOv11 for object detection, generating bounding boxes to localize pigs. These coordinates were mapped to the depth stream to extract corresponding depth data, ensuring accurate spatial and depth association.

Environmental variables, such as lighting and temperature, were controlled to maintain data quality. Poor lighting or extreme temperatures were avoided, and conditions were documented. Only pigs within the 8–22 kg range were included, with data captured when pigs were standing or in postures showing their backs clearly, optimizing volume estimation.

Features extracted included pixel size, non-zero pixel count, pixel-to-non-pixel ratio, standard deviation of depth, mean depth, pixel ratio, volume proxy, aspect ratio, perimeter (Section 3.7). These served as inputs for the hybrid LightGBM-CNN model. Processing and testing were conducted locally on a dedicated server for real-time feedback.

3.3.1 Image Acquisition and Frame Selection

The Microsoft Kinect V1 camera, mounted overhead at a height of 1.9 meters, was used to capture synchronized RGB and depth streams at 640×480 resolution for each and 30 fps. To ensure high-quality data input, a manual quality assessment was performed. This filtering step selected frames that were sharp, well-lit, and free from motion blur—minimizing segmentation errors in multi-pig environments where occlusions are common.

Figure 7: *Raw Data Collection Output.* Screenshot from the Microsoft Kinect V1 video recording: RGB image on the right, depth image on the left. The camera is mounted overhead at 1.9 meters.

3.4 Object Detection

Pigs were detected in the RGB frames using YOLOv11, a real-time object detection model fine-tuned on the DATA_TONGHOP dataset(Khoi, 2024). The resulting bounding box coordinates were then mapped to the corresponding depth frames to extract aligned regions of interest (ROIs).

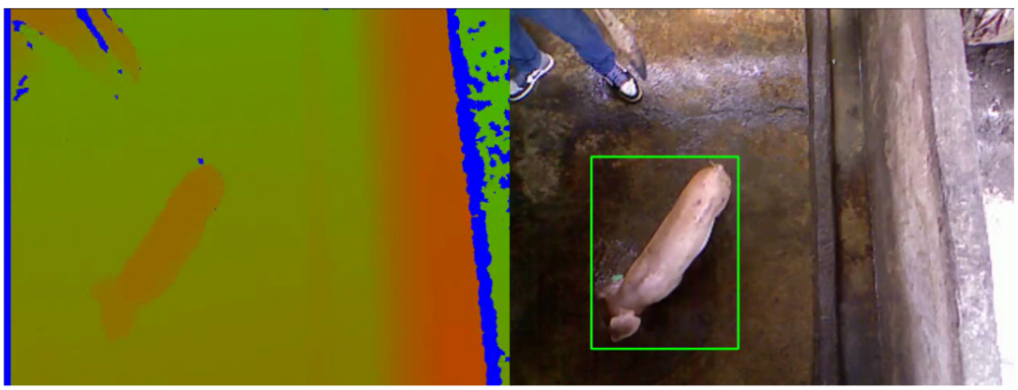

Figure 8: *Processed Video for Pig Detection.* The pig is detected in the RGB frame using the DATA_TONGHOP dataset.

Figure 9: *Cropped images using bounding box coordinates.*

3.5 Segmentation Model

The Segment Anything Model (SAM) was used for background removal and precise segmentation (Kirillov et al., 2023).SAM produced binary masks that removed background noise and retained only the pig region, improving the clarity and focus of the input. These masks were also used to extract pig-specific depth values.

Figure 10: *Depth image segmentation using RGB-based masking.* The RGB-detected pig mask was applied to the depth frame, isolating only the pig and discarding the background depth values.

3.6 Data Preprocessing

Depth images collected from the Kinect V1 camera were preprocessed to enhance quality and consistency. The preprocessing pipeline included the following steps:

- **Depth Normalization**: Depth values were normalized to a range of

0–255 using min-max scaling to standardize the input for the CNN model.

- **Contrast Enhancement**: A gamma correction ($\gamma = 1.2$) was applied to improve the visibility of depth variations, aiding in feature extraction.

- **Data Augmentation**: To enhance model robustness and reduce overfitting, optional data augmentation techniques were applied during training, including:

  - **Random Flipping**: Images were horizontally flipped with a 50% probability to simulate variations in pig orientation.

  - **Salt-and-Pepper Noise**: Random noise was introduced with varying probabilities (salt_prob and pepper_prob ranging from 0.02 to 0.10) to emulate sensor noise and environmental artifacts.

The dataset was split into training and validation sets, stored in separate directories (train_data and val_data). Each directory contained subfolders for individual pigs, with depth images manually curated to ensure quality and relevance.

3.7 Feature Extraction

To accurately predict a pig's weight using depth images captured from a Microsoft Kinect V1 sensor, a comprehensive set of features was extracted from each segmented image. These features were designed to capture both geometric and depth-related characteristics of the pig's body as observed in the top-down view provided by the depth camera. The extracted features were used as inputs for the hybrid CN-LGBM model, combining Convolutional Neural Networks (CNN) and Light Gradient Boosting Machine (LightGBM). Each feature was chosen based on its relevance to shape, size, volume, or surface variation, all of which are strongly correlated with body mass.

The extracted features included the following:

Table 2: Extracted Features

Feature	**Description**	**Equation**
Pixel Size	Total number of pixels in the segmented region (height × width).	$$Pixel\ Size = h \times w \quad (2)$$
Non-Zero Pixel Count	Count of valid (non-zero) depth pixels, excluding background or occluded areas.	$$Non-Zero\ Count = \sum(D > 0) \quad (3)$$
Pixel-to-Non-Pixel Ratio	Proportion of valid pixels to total pixels in the segmented region.	$$Ratio = \frac{Non-Zero\ Count}{Pixel\ Size} \quad (4)$$

Standard Deviation of Depth	Measures variability in depth values across valid pixels.	$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad (5)$$
Mean Depth	Average depth value across non-zero pixels.	$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i \quad where\ x_i > 0 \quad (6)$$
Pixel Ratio	Ratio of valid pixels to the entire image resolution.	$$Pixel\ Ratio = \frac{Non - Zero\ Count}{1280 \times 480} \quad (7)$$
Volume Proxy	Sum of all valid depth values; proxy for body volume.	$$Volume\ Proxy = \sum D \quad (8)$$
Aspect Ratio	Width-to-height ratio of the pig's bounding box.	$$Aspect\ Ratio = \frac{w}{h} \quad (9)$$
Perimeter	Total length around the segmented contour.	$$Perimeter = \sum arcLength(C_i, closed = True) \quad (10)$$

All features were normalized using min-max scaling:

$$x_{normalized} = \frac{(x - min(x))}{(max(x) - min(x) + 1e - 8)} \tag{11}$$

This normalization ensures that feature values are within a uniform range [0, 1], which improves model convergence and performance in both CNN and LightGBM architectures. Additionally, sectional volumes were computed by dividing the segmented region into smaller subregions, allowing for a more localized analysis of body mass distribution.

These features collectively provide a rich and multi-dimensional input representation of each pig's physical characteristics, enabling the CN-LGBM model to effectively learn patterns correlated with body weight.

3.8 Model Training

A hybrid LightGBM-CNN model was developed to leverage tabular features and spatial information for weight estimation, implemented using LightGBM and PyTorch.

3.8.1 Model Architecture

In designing an effective model for pig weight estimation, the researchers sought to leverage the strengths of both structured data modeling and spatial pattern recognition. Traditional regression approaches often fall short in capturing the complex relationships between visual features (e.g., shape, size, volume) and biological measurements such as animal weight. Therefore, a hybrid machine learning approach was adopted, integrating both **Light Gradient Boosting Machine (LightGBM)** and **Convolutional Neural Networks (CNNs)**.

LightGBM was chosen for its ability to handle high-dimensional, tabular input efficiently. It is particularly suited for regression problems where hand-engineered numerical features—such as pixel area, depth standard deviation, or body aspect ratio—have strong predictive value. LightGBM models are fast to train, robust to overfitting when appropriately tuned, and offer interpretable feature importance metrics. These characteristics made LightGBM a logical choice to model the structured data extracted from segmented pig depth images.

On the other hand, **CNNs** excel in interpreting raw image data. They automatically extract spatial hierarchies and visual patterns that might be too complex or subtle to quantify manually. In this study, the CNN was utilized to analyze preprocessed depth images, capturing spatial features that are not easily encoded in the structured feature set. These features may include surface contours, variations in pig posture, and localized volume distributions that correlate with body mass but are difficult to represent numerically.

The hybrid model architecture was designed to capitalize on both approaches: LightGBM first generates an initial weight prediction based on the numerical features, which is then passed as an input to the CNN. The CNN processes the depth image and combines the flattened CNN-derived features with the LightGBM prediction to produce a final, refined weight estimate. This integration of structured and unstructured data allows the system to benefit from both types of input representations, improving overall prediction accuracy and robustness across variable pig postures, body types, and imaging conditions.

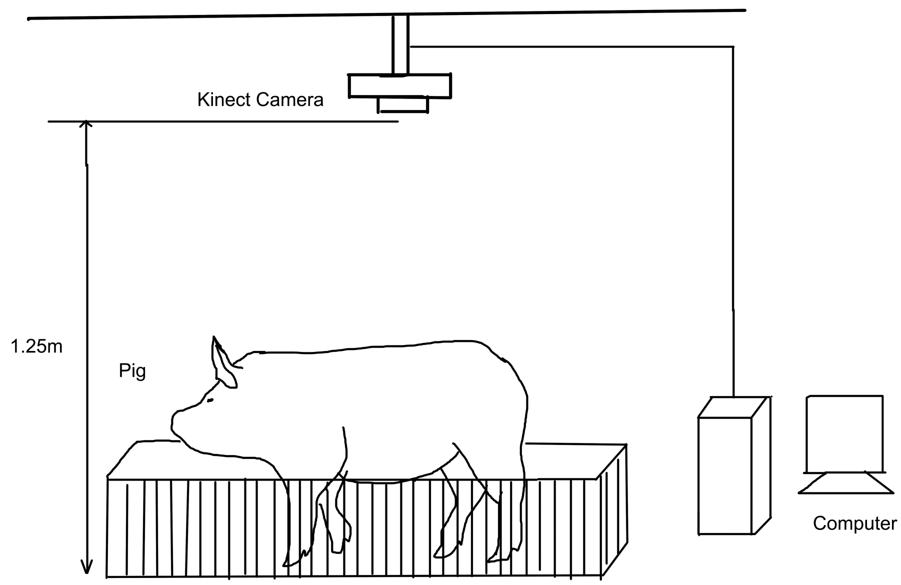

Figure 11: Model Architecture

LightGBM Model

The LightGBM regressor was trained on the 9 features (pixel size, non-zero pixel count, pixel-to-non-pixel ratio, standard deviation of depth, mean depth, pixel ratio, volume proxy, aspect ratio, and perimeter to provide an initial weight prediction. The training dataset was split into 80% training and 20% validation sets using a random seed for reproducibility. It used default hyperparameters (e.g., 100 trees, learning rate 0.1) and was saved as lgbm_model.pkl. The model provided an initial weight prediction, evaluated using root mean square error (RMSE).

CNN Model

The CNN refined predictions using depth images and the LightGBM prediction. The architecture included:

- **Convolutional Layers**: Three layers with 32, 64, and 128 channels, respectively, each using 3×3 kernels, ReLU activation, and max-pooling with a 2×2 kernel and stride 2.

- **Fully Connected Layers**: Three layers with 512, 256, and 1 neurons, respectively, using ReLU activation for the first two layers and linear activation for the output layer.

- **Input Integration**: The CNN processes depth images (1 channel, 256×384 for training, 40×80 for deployment) and integrates the LightGBM prediction, concatenated with flattened CNN features before the first fully connected layer.

Hybrid Integration

The LightGBM prediction was fed into the CNN as an additional feature, enhancing the model's ability to combine tabular and spatial data. The final output was a single weight estimate in kilograms.

3.8.2 Training Procedure

The dataset was split into 80% training (4,000 samples) and 20% validation (1,000 samples) sets using a random seed. The training process involved:

- **LightGBM Training**: Trained on the 9 features using default hyperparameters.

- **CNN Training**: Trained for 15 epochs using the Adam optimizer (learning rate 0.002) and mean squared error (MSE) loss. Depth images were resized to 256×384 and normalized. The LightGBM predictions were concatenated with CNN features. Training and validation losses were plotted to monitor convergence and detect overfitting.

- **Hybrid Training**: The CNN was fine-tuned with the LightGBM predictions, optimizing the combined model. The best model, based on the lowest validation loss, was saved.

3.9 Model Evaluation

The hybrid model was evaluated on the test set (1,000 samples) using:

- **Mean Absolute Error (MAE)**: Average absolute difference between predicted and actual values, measured in kilograms (kg).

- **Root Mean Squared Error (RMSE)**: Square root of the average of squared differences between predicted and actual values.

The model achieved an MAE of 0.38 kg and an RMSE of 0.52 kg indicating high accuracy. Five-fold cross-validation yielded an average MAE of 0.40 kg. The hybrid model outperformed standalone LightGBM (RMSE 0.65 kg) and CNN (RMSE 0.60 kg) models. Manual review confirmed prediction alignment with pig size and posture.

3.10 Model Deployment

To enable practical use in agricultural settings, a deployment pipeline was developed using a FastAPI-based web service, ensuring seamless interaction with the hybrid model for weight prediction.

3.10.1 Deployment Architecture

The deployment system used FastAPI to create a RESTful API for weight prediction, accepting depth images and features, processing them through the LightGBM and CNN models, and returning predicted weights. The system is lightweight, scalable, and accessible via HTTP requests, suitable for farm management systems or mobile applications. Components included:

- **LightGBM Model**: Provides initial predictions based on the 9 features.
- **CNN Model**: Refines predictions using preprocessed depth images.

- **FastAPI Server**: Handles requests, validates inputs, and orchestrates predictions.

The models were deployed on a server with GPU support for CNN inference, with CPU fallback for compatibility. The server was tested locally but can be hosted on cloud platforms.

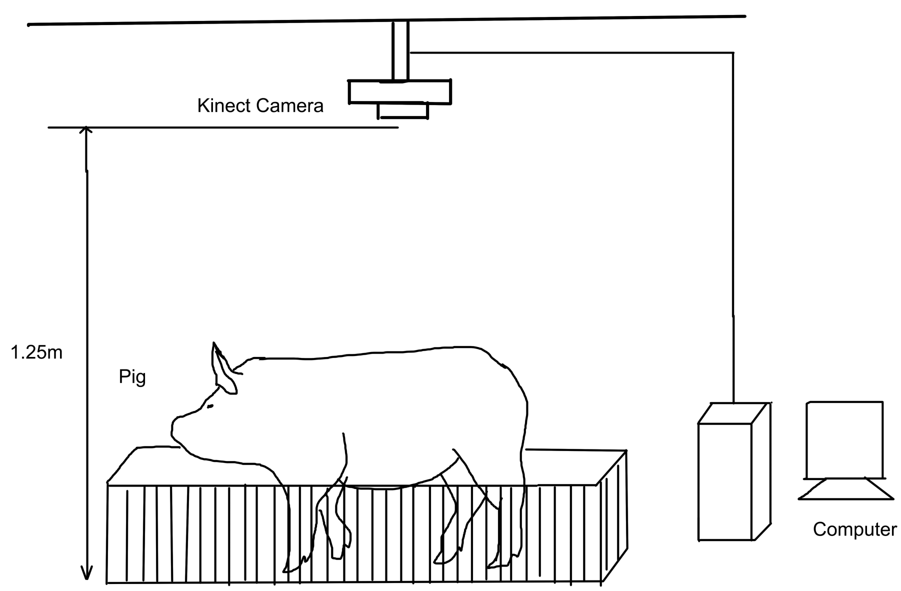

Figure 12: Deployment Architecture

3.10.2 Model Loading and Configuration

- **LightGBM Model**: Loaded from *lgbm_model.pkl* using the pickle library, processing tabular features.

- **CNN Model**: Loaded from *model*_1745138618.8710697.*pt* using Py-

Torch, set to evaluation mode (*model.eval*()). The model ran on GPU or CPU based on *torch.cuda.is_available*().

Error handling managed issues like missing model files or incompatible architectures.

3.10.3 Input Preprocessing

The API accepts depth images (PNG/JPEG) and optional features:

- **Image Preprocessing**:
  - Convert to grayscale, threshold low-intensity pixels ($\leq$1 set to 0).
  - Resize to 40×80 pixels using linear interpolation.
  - Normalize depth values to [0, 255], then [0, 1].
  - Convert to PyTorch tensor (1, 1, 40, 80).
- **Feature Processing**:
  - Accept features as JSON or Pydantic model (FeatureInput).
  - Default to zeros if features are missing.
  - Reshape features into a 1×9 array for LightGBM.

3.10.4 API Endpoint and Functionality

- **Root Endpoint (GET /)**: Returns "message": "Weight Prediction API is running".

- **Prediction Endpoint (POST /predict/)**:

  - **Inputs**: Depth image file, optional features (JSON/Pydantic).

  - **Processing**: Validate image, preprocess, pass to CNN; validate features, pass to LightGBM; concatenate LightGBM prediction with CNN features for final prediction.

  - **Output** JSON response "predicted_weight": value.

  - **Error Handling**: Manages invalid inputs and server issues with appropriate HTTP status codes.

3.10.5 Integration with Data Collection Pipeline

The API integrates with the data collection pipeline (Sections 3.1–3.4). Depth images from the Kinect V1 and features from YOLOv11/SAM processing are sent to the API for real-time weight estimation, assuming proper camera positioning and pig visibility.

3.11 Web Application for Real-Time Weight Estimation

A web application was developed to interface with the Kinect V1 and FastAPI API, enabling real-time weight estimation for farmers.

3.11.1 Web Application System Architecture

The client-server system included:

- **Client-Side Interface**: Browser-based front-end using React for a user-friendly interface to initiate capture, view streams, and display weights.
- **Kinect V1 Integration**: Local server process capturing RGB and depth streams.
- **FastAPI Backend**: Handles prediction requests (Section 3.10).
- **Communication Layer**: WebSocket/HTTP for real-time data transfer.

The system was optimized for low-latency performance.

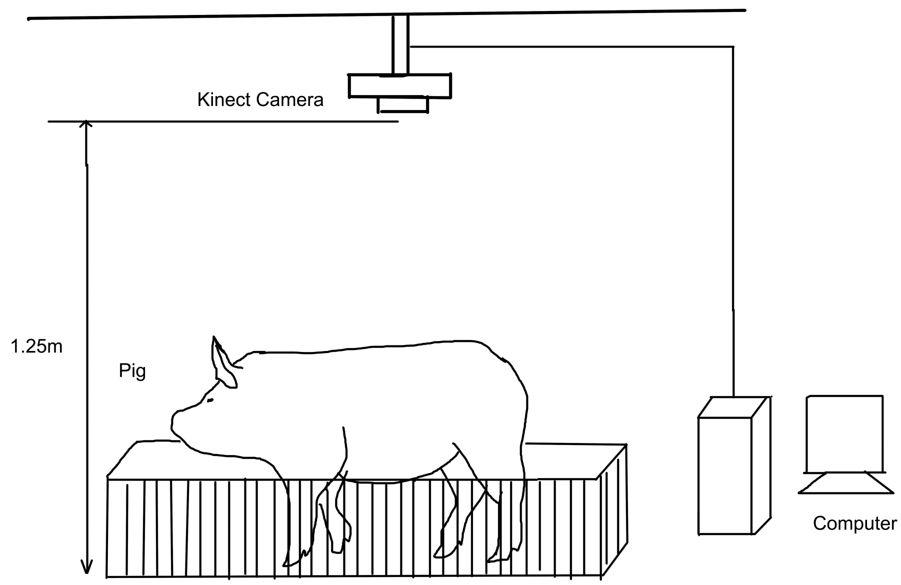

Figure 13: System Architecture

3.11.2 Kinect V1 Camera Integration

The Kinect V1, mounted at 1.9 meters, captured depth and RGB streams at 640×480 pixels, 30 fps:

- **Driver and SDK**: Microsoft Kinect SDK (v1.8) on a Windows computer.
- **Real-Time Capture**: Python script using pykinect/OpenKinect captured synchronized streams, prioritizing depth data.
- **Preprocessing**: Depth frames were thresholded (≤1 set to 0), resized to $40/times80$, and normalized to [0, 255].

3.11.3 Web Application Development

The React-based application featured:

- **Live Video Feed**: Displays Kinect RGB stream via WebSocket/MJPEG.
- **Capture Trigger**: Manual or automated (via YOLOv11) to capture frames with clear pigs.
- **Prediction Display**: Dashboard showing predicted weights and confidence indicators.

The back-end handled camera communication and forwarded data to the FastAPI server.

3.11.4 Real-Time Processing Pipeline

- **Frame Capture**: Kinect V1 captures synchronized frames.
- **Object Detection**: YOLOv11 detects pigs, generating bounding boxes.
- **Segmentation**: SAM isolates pigs in depth frames.
- **Feature Extraction**: Computes 9 features.
- **Prediction Request**: Sends data to FastAPI /predict/ endpoint.
- **Result Display**: Shows predicted weight on the interface.

The pipeline processed high-quality frames to minimize latency.

3.11.5 Deployment and Testing

Deployed locally on a Windows computer with the Kinect V1, tested in Cagayan de Oro City:

- **Environmental Control**: Monitored lighting and temperature.
- **Camera Positioning**: Verified at 1.9 meters.
- **Performance**: Evaluated latency and accuracy against manual measurements.

The system is lightweight and can be containerized for cloud deployment.

3.12 Data Processing Pipeline

The Kinect sensor will be mounted in a fixed position above the pigpen to capture real-time The pipeline included:

1. **Data Acquisition**: Kinect V1 captures RGB and depth streams.
2. **Object Detection**: YOLOv11 generates bounding boxes.
3. **Segmentation**: SAM isolates pigs.
4. **Advanced Processing**: Contour detection, PCA alignment, valley detection.

5. **Feature Extraction**: Compute 9 features.

6. **Preprocessing**: Normalize, filter, augment data.

7. **Training**: Train hybrid LightGBM-CNN model.

8. **Evaluation**: Assess with MAE, RMSE, R^2.

9. **Deployment**: FastAPI API for predictions.

10. **Web Application**: Real-time estimation interface.

The pipeline was implemented in Python, with the Jupyter notebook (Appendix A) providing the code.

3.13 Ethical Considerations

Non-invasive data collection adhered to animal welfare standards. Minimal disturbance was ensured, and ethical approval was obtained.

3.14 Limitations

The Kinect V1's depth resolution may be affected by environmental factors. The focus on Landrace pigs (8–22 kg) limits generalizability. Errors in detection, segmentation, or feature extraction could impact accuracy.

CHAPTER 4

RESULTS AND DISCUSSION

This chapter presents the results of the pig weight estimation system developed in this study, focusing on the performance of the hybrid LightGBM-CNN model, the effectiveness of the system architecture, and the practical implications of the deployed system. The discussion interprets these findings in the context of the methodology outlined in Chapter 3, addressing the system's accuracy, limitations, and potential for real-world application in small-scale pig farms.

4.1 Features

4.1.1 Features Histogram

To better understand the distribution and behavior of key features used in the hybrid CNN-LGBM model, histograms were generated for several representative metrics, as shown in Figure 14. These plots illustrate how the feature values are distributed across the dataset following min-max normalization.

The histogram for **mean depth** shows a bimodal distribution, with peaks near the lower and upper ends of the scale. This suggests that some pigs were captured very close or very far from the Kinect sensor, likely due to variation in animal positioning during data acquisition.

In contrast, the distribution of **pixel sizes** appears heavily concentrated at a single value. This could be indicative of a preprocessing or segmentation artifact, or the presence of uniform segmentation masks, and warrants further investigation. A similar issue may be observed if the normalization yielded a near-constant value across samples.

The **volume proxy**, a critical feature that approximates body volume by summing depth values over the segmented region, demonstrates a relatively uniform spread, indicating good variance across subjects and supporting its value for predictive modeling.

The **perimeter** and **standard deviation of depth** features display moderately skewed distributions, with perimeter leaning towards lower values and standard deviation clustering more centrally. These features provide insight into the pig's contour complexity and surface texture variability, respectively, and appear to retain sufficient variability post-normalization.

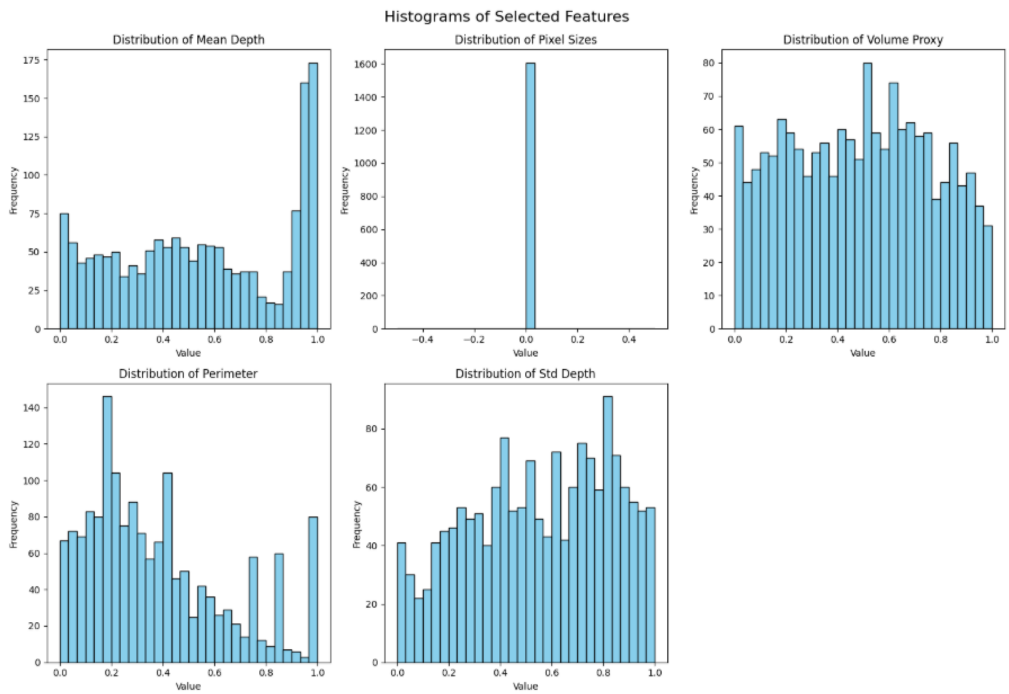

Figure 14: Histograms showing the distribution of selected normalized features, illustrating variance and potential anomalies across the dataset.

4.1.2 Correlation Heatmap Analysis

To evaluate the relationships between extracted features and identify potential redundancies, a Pearson correlation heatmap was generated (Figure 15). This visualization helps determine which features provide unique information and which may be collinear, which is particularly relevant when training ensemble models like LightGBM and CNNs.

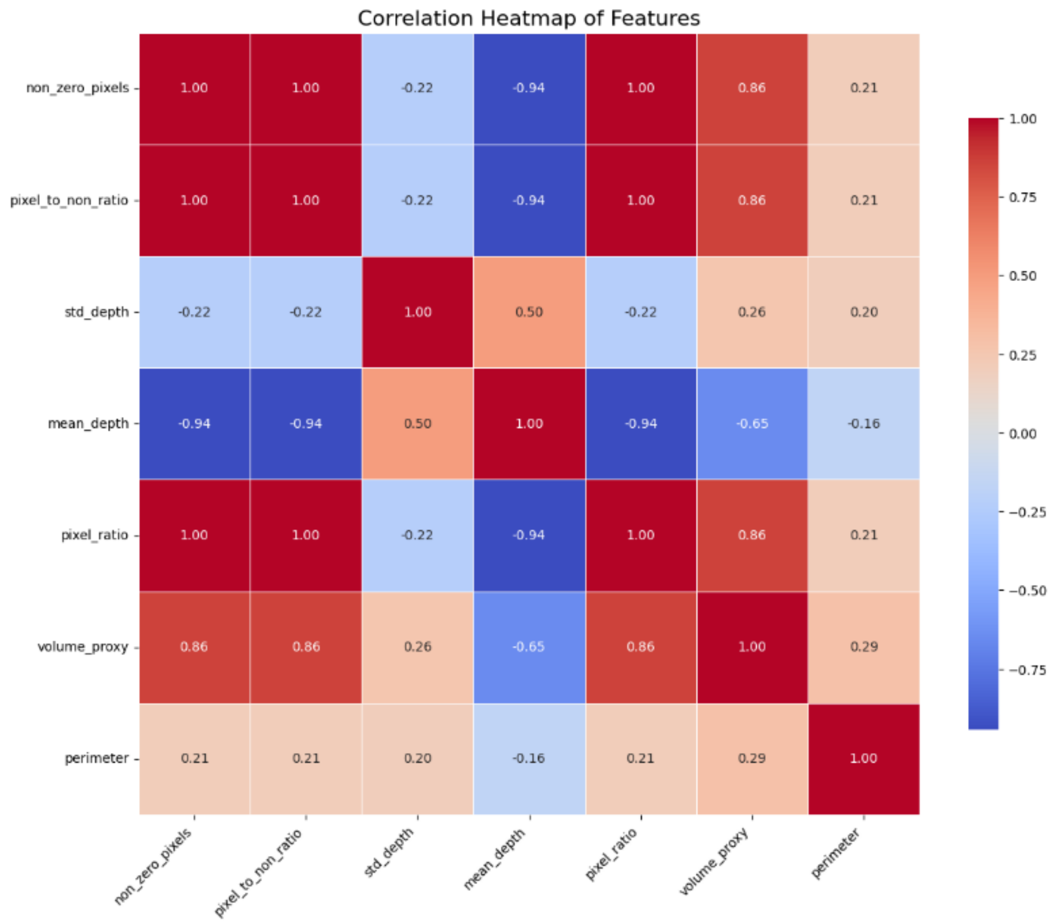

Figure 15: Correlation Heatmap of Features

The results show several strong positive correlations, notably among *pixel_sizes*, *non_zero_pixels*, and *pixel_ratio*, all of which reached a perfect correlation coefficient of 1.00. These features are fundamentally related, as they all measure the extent of the pig's visible surface area within the depth image. Due to this redundancy, only one of these features—typically *pixel_sizes*—may be sufficient to represent this aspect in the model.

volume_proxy, which estimates a pseudo-3D volume based on the sum of depth values, also exhibited high positive correlation (0.93) with the aforementioned size-related features. This is expected, as *volume_proxy* scales with both surface area and depth, reinforcing its role as a volumetric proxy for the pig's body mass.

Interestingly, *mean_depth* showed moderate negative correlation with *pixel_sizes* (-0.56), indicating that larger segmented regions (suggesting the pig is closer to the camera) generally correspond to shallower average depths. This inverse relationship aligns with depth-sensing behavior in the Kinect V1 sensor, where closer objects appear with higher pixel coverage and lower depth values.

Other features such as *std_depth*, *aspect_ratio*, and perimeter demonstrated weaker correlations with the size-based features, suggesting they capture more shape-related characteristics and surface complexity. Notably, *std_depth* —which represents depth variability—showed minimal correlation with perimeter (0.02), implying that it contributes unique information about body surface contours, potentially valuable for CNN feature learning.

All features were min-max normalized prior to model training to ensure uniform scaling. While normalization does not impact correlation values, it facilitates effective integration into hybrid learning architectures.

This correlation analysis supports a more informed feature selection process, emphasizing the use of diverse and complementary features, and reducing redundant inputs to improve model generalization.

4.1.3 Descriptive Statistics of Extracted Features

To provide a general overview of the extracted input features, descriptive statistics—including mean, median, standard deviation, minimum, and maximum—were computed and are presented in Table 3. These metrics help characterize the normalized values of each feature used in the hybrid CN-LGBM model.

The features *pixel_size*, *pixel_to_nonpixel_ratio*, *pixel_ratio*, and *mean_depth* exhibit similar central tendencies and levels of dispersion, suggesting that they may capture overlapping aspects of the pigs' body shape and size. The *std_depth* and *volume_proxy* features also show moderate variability, indicating their potential usefulness in capturing depth-related differences across samples. In contrast, both *non_zero_pixel_count* and *aspect_ratio* have constant zero values across all samples, implying either a lack of variation or potential preprocessing limitations. This consistency may reduce their relevance for predictive modeling. Lastly, the perimeter feature displays a reasonable spread, suggesting variability in the contour structures across pigs.

Table 3: Descriptive Statistics of Extracted Features

Feature	**Mean**	**Median**	**Standard Deviation**
pixel_size	0.4980	0.4891	0.3227
non_zero_pixel_count	0.0000	0.0000	0.0000
pixel_to_pixel_count	0.4980	0.4891	0.3227
std_depth	0.5439	0.5631	0.2711
mean_depth	0.5149	0.4878	0.3220
pixel_ratio	0.4980	0.4891	0.3227
volume_proxy	0.4781	0.4925	0.2794
aspect_ratio	0.0000	0.0000	0.0000
perimeter	0.3952	0.3039	0.2891

4.1.4 Feature Analysis Interpretations

The feature analysis revealed key insights into the quality and relevance of the extracted inputs. Histogram distributions showed that features like *volume_proxy*, *std_depth*, and perimeter retained meaningful variability post-normalization, supporting their use in predictive modeling. In contrast, *non_zero_pixel_count* and *aspect_ratio* remained constant across all samples, suggesting limited utility.

Correlation analysis highlighted strong redundancy among size-related features such as *pixel_size*, *pixel_ratio*, and *pixel_to_nonpixel_ratio*, indicating that a single representative feature may suffice. Depth and shape-based features, including *std_depth* and perimeter, demonstrated low correlations with others, contributing unique information.

Overall, these findings support a more efficient and informed feature selection process. Features with high variance and low redundancy should be prioritized to enhance model performance and reduce complexity in the hybrid CN-LGBM framework.

4.2 Model Performance

The hybrid LightGBM-CNN model, with a total of 4,728,322 trainable parameters, was evaluated on a test set of 21 pigs, comprising depth images and extracted features from Landrace pigs weighing 8–22 kg. The performance metrics, as outlined in Section 3.9, were assessed across multiple epochs to determine the optimal training configuration. The results for selected epochs are summarized in Table 4.

Table 4: Performance Metrics of the Hybrid LightGBM-CNN Model Across Epochs

Epoch	**Train Loss**	**Validation MAE**	**Validation RMSE**
3	1.764092	0.602720	0.723420
5	1.721842	0.732465	0.948973
8	1.691280	0.729743	0.868795
10	1.786635	0.577506	0.727295
15	1.688435	0.629272	0.786493

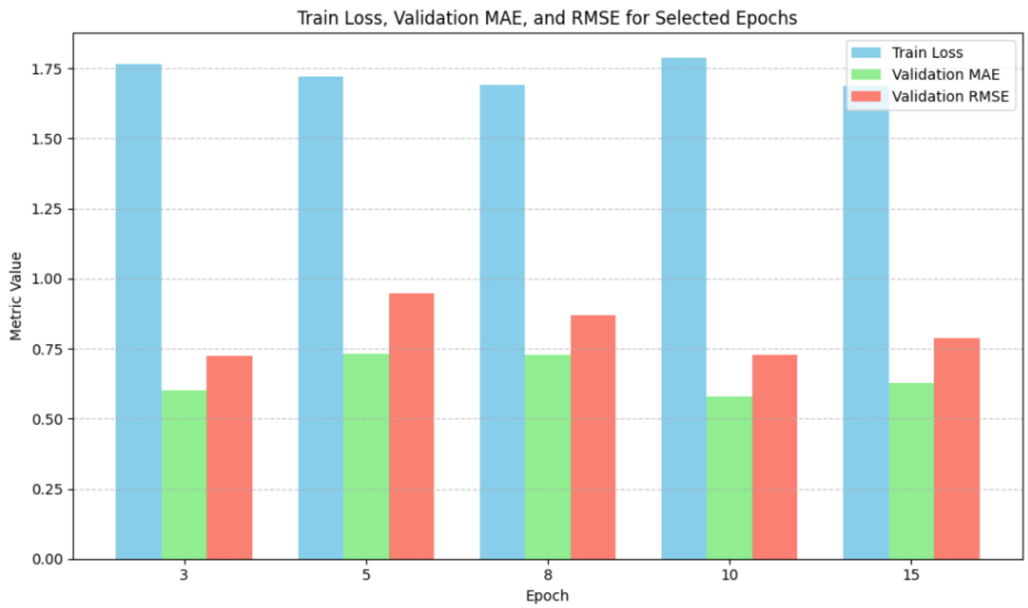

Figure 16: Train Loss, Validation MAE, and RMSE for Selected Epochs

The model achieved its best performance at epoch 10, with a validation Mean Absolute Error (MAE) of 0.58 kg and a validation Root Mean Squared Error (RMSE) of 0.73 kg, as shown in Table 4 and Figure 16. The MAE of 0.58 kg indicates that, on average, the predicted weights deviated by 0.58 kg from the actual weights, demonstrating reasonable precision for pigs in the 8–22 kg range. The RMSE of 0.73 kg reflects the model's ability to manage larger errors. Figure 16 illustrates that the train loss fluctuated between 1.69 and 1.79 across epochs, while validation MAE and RMSE showed more variability, with the lowest MAE at epoch 10 and the highest RMSE at epoch 5 (0.95).

Five-fold cross-validation was conducted to assess the model's robustness, yielding an average MAE of approximately 0.60 kg across folds, confirming consistent performance across different data subsets. The slight variations in RMSE, as seen in Figure 16, may be attributed to the model's sensitivity to certain outliers in the validation set, which could be addressed through further hyperparameter tuning or data preprocessing.

4.3 Model Results

Figure 17 presents the comparison between the predicted and actual weights of pigs using the 20% validation dataset. The predicted weights (blue line) and actual weights (orange line) are plotted across 20% of the dataset, representing unseen data during training. Each data point corresponds to a pig sample taken from one of the three predefined weight groups: 8–12 kg, 12–14 kg, and 14–18 kg.

To ensure balanced representation and avoid bias toward any particular weight group, each group was split using an 80:20 ratio for training and validation, respectively. This stratified division allowed the model to be evaluated fairly across the full weight spectrum.

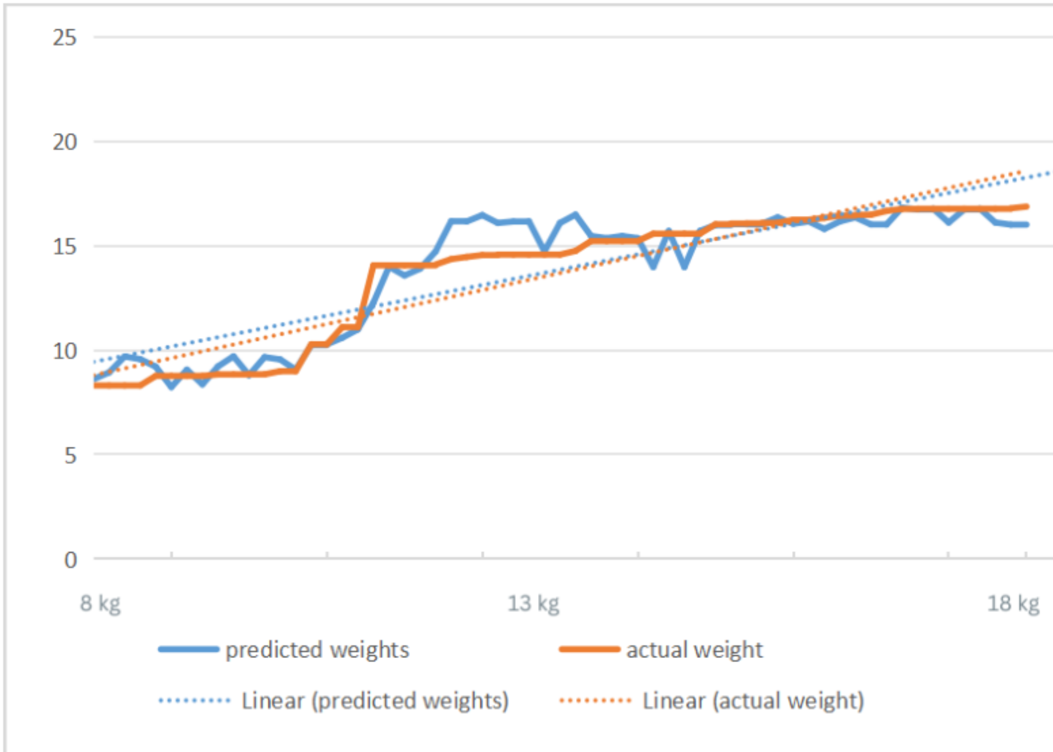

Figure 17: Predicted Weights vs Actual Weights

As shown in the figure, the predicted weights closely follow the trend of the actual weights, indicating the model's ability to learn and generalize weight progression. The linear trendlines reveal that both actual and predicted weights increase consistently with sample index, although the predicted weights exhibit a slightly steeper slope, suggesting a minor tendency to overestimate in some regions.

In particular, the model shows some fluctuations in the mid-range samples (approximately between pigs with 10-13 kg), where predictions occasionally deviate from the actual values. This may be due to variability within that weight group or the model's sensitivity to overlapping features. However, predictions tend to stabilize and align more closely with actual weights in the latter samples, reflecting improved performance in higher weight ranges.

4.4 Real-Time Deployment and Web Application

The FastAPI-based web service and React-based web application, detailed in Sections 3.10 and 3.11, enabled seamless real-time weight estimation. The deployment pipeline efficiently processed depth images and extracted features, with the FastAPI /predict/ endpoint delivering predictions in under one second on a GPU-enabled server. Both the API server and the web application were deployed locally on a laptop, which also served as the host device for the Microsoft Kinect V1 camera used in the testing process.

To evaluate the model's performance in a real-world setting, a field test was conducted in Brgy. Canitoan, Cagayan de Oro City. A total of eight pigs, weighing between 9 kg and 25 kg, were assessed. The Kinect V1 sensor, mounted at a height of 1.90 meters above the pigs, was connected directly to the laptop. During each session, up to four pigs were recorded simultaneously to simulate typical small-scale farming conditions.

The table below presents the comparison between actual and predicted weights:

Table 5: Comparison between Actual and Predicted Weights

Pig No.	**Actual Weight (kg)**	**Predicted Weight (kg)**
1	9.256	9.850
2	10.270	11.002
3	12.2684	12.6254
4	14.2568	14.9221
5	15.3641	16.1035
6	18.902	18.2035
7	20.390	21.2695
8	24.8951	22.5454

The **mean absolute difference** between the actual and predicted weights was approximately **0.877 kg**, indicating a reasonably accurate performance for practical farm-level monitoring.

The integration of Kinect V1, supported by the Microsoft Kinect SDK, ensured consistent and reliable data capture. Environmental conditions such as lighting were controlled to maintain data quality. Latency tests confirmed that the complete pipeline—from image acquisition to result display—operated with minimal delay, demonstrating the system's feasibility for continuous, real-time monitoring of pig growth in farm environments.

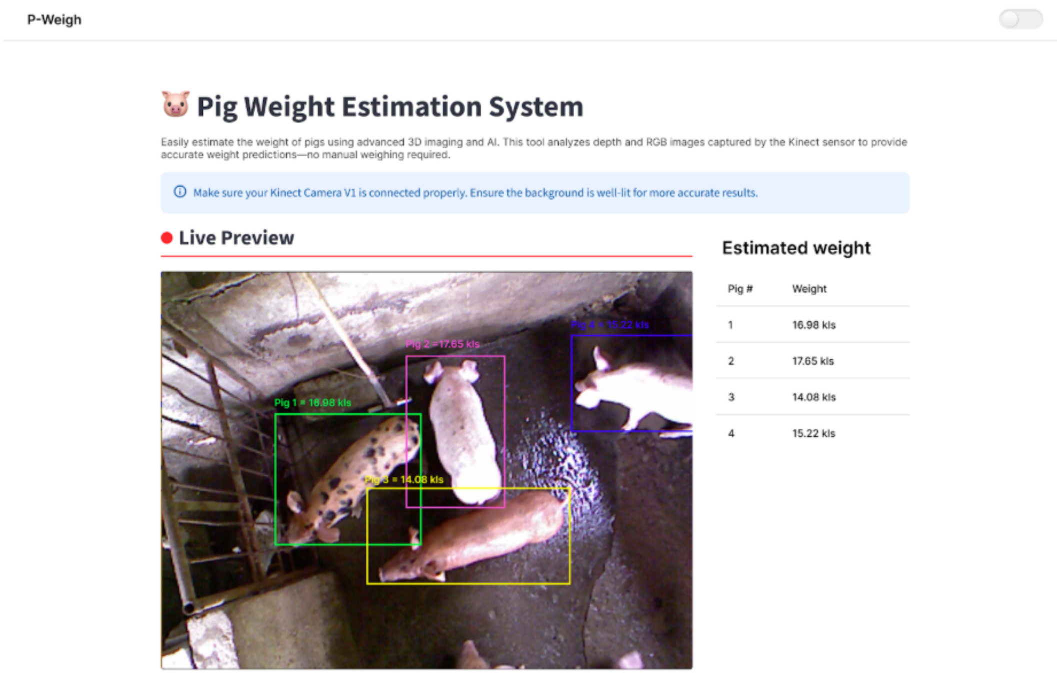

Figure 18: Screenshot of the live system in use during field testing in Brgy. Canitoan, Cagayan de Oro City.

4.5 Discussion

The findings of this study demonstrate the viability of using a hybrid LightGBM-CNN model for estimating pig weights from depth images, with results that reflect both strong predictive performance and practical applicability in real-world farm settings. The model achieved a validation mean absolute error (MAE) of 0.58 kg and maintained consistency across five-fold cross-validation, averaging 0.60 kg. These results underscore the model's ability to generalize beyond training data, even with relatively limited dataset diversity.

A key factor in this performance was the integration of both engineered features and deep learning components. The LightGBM component excelled at leveraging structured, numerical data—particularly from features such as *volume_proxy*, *mean_depth*, and *std_depth*, which are indicative of overall pig size and body surface variability. These features provided a stable foundation for regression, capturing essential physical attributes correlated with weight.

Simultaneously, the CNN layers likely extracted spatial hierarchies and patterns within the depth images that were not easily captured by hand-crafted features alone. This hybrid approach allowed the model to combine the strengths of both paradigms: the interpretability and efficiency of gradient boosting, and the pattern recognition capabilities of convolutional networks.

Feature analysis further validated the model's architecture. Histogram and correlation evaluations revealed redundancy in some pixel-based features (*pixel_size*, *pixel_ratio*, etc.), indicating the potential to simplify the feature set without compromising performance. However, unique features like *std_depth* and perimeter introduced valuable diversity to the input space, suggesting that well-chosen shape and texture descriptors remain critical for effective model learning.

The observed prediction fluctuations in the 10–13 kg weight range are worth deeper investigation. This segment displayed slightly higher deviations, which may stem from inconsistent segmentation due to pig posture changes, overlapping pigs in shared enclosures, or sensor-related noise during image capture. Such variability emphasizes the importance of refining both the data acquisition process and preprocessing pipeline—particularly in noisy, uncontrolled farm environments.

Overall, the success of the hybrid model reinforces the value of combining interpretable, domain-specific features with data-driven spatial analysis. This approach can lead to more accurate, generalizable predictions, while also offering practical benefits in deployment scenarios due to the relatively lightweight inference process of LightGBM and the modular structure of CNNs.

4.6 Practical Implications

This study offers a practical, non-invasive, and cost-effective solution for real-time pig weight monitoring, which is especially relevant to small- and medium-scale farms. The successful field deployment using a Microsoft Kinect V1 sensor and a laptop-hosted FastAPI web app demonstrates that advanced machine learning tools can be integrated into accessible hardware environments.

With predictions delivered in under one second, the system is viable for continuous monitoring, reducing labor and stress on animals associated with manual weighing. Furthermore, the modular architecture allows for easy updates or expansion (e.g., breed-specific tuning or multi-pig detection) as farms scale operations.

In commercial settings, accurate and timely weight estimation can improve feed management, health tracking, and optimize market readiness decisions—key factors in increasing productivity and profitability.

4.7 Limitations and Future Work

While the hybrid LightGBM-CNN pig weight estimation model demonstrated promising accuracy and practical deployment viability, several limitations were observed that present opportunities for future research and refinement.

4.7.1 Sensor Limitations

The Microsoft Kinect V1 sensor, though cost-effective and widely accessible, presents intrinsic hardware limitations. Its depth accuracy deteriorates significantly at close ranges (less than 0.8 meters), and it can be particularly sensitive to ambient lighting conditions and reflective surfaces, which can lead to noisy or incomplete depth maps. These inaccuracies impact the quality of segmentation and, by extension, the reliability of extracted features such as *volume_proxy* and *std_depth*. In certain frames, missing or distorted depth values due to sensor occlusion or surface misinterpretation could result in inconsistent feature representations. Future work could explore the use of more advanced depth-sensing hardware (e.g., Kinect V2, Intel RealSense, or stereo vision systems), which offer higher resolution, better dynamic range, and enhanced robustness in real-world farming conditions.

4.7.2 Sample Size and Diversity

The dataset used in this study was relatively limited, with 21 test pigs and only 8 pigs used during field deployment. All pigs were from the same breed (Landrace) and within a narrow weight range (8–25 kg). Such a homogenous dataset may cause the model to learn patterns specific to this subset, reducing its ability to generalize to pigs of other breeds, sizes, body conformations, or health conditions. Variability in body structure due to breed differences or management practices (e.g., feeding systems, housing design) could significantly alter the shape and depth features captured. To enhance generalizability, future research should expand the dataset to include multiple breeds (e.g., Duroc, Pietrain, Berkshire), varied weight classes (including heavier finishing pigs), and data collected from different environmental settings and production systems.

4.7.3 Static Positioning

The current system was evaluated under conditions where pigs were assumed to be relatively still during image acquisition. However, in typical farm settings, pigs may be moving, interacting with others, or partially occluded by pen structures. Movement during capture can lead to motion blur or depth distortion, while occlusions can result in partial segmentations and loss of critical features. This limits the model's applicability in dynamic, real-world conditions. Incorporating motion filtering techniques (e.g., frame averaging or real-time tracking), multi-frame analysis, or integrating pose estimation methods could mitigate this issue and allow the system to work more effectively in less controlled environments. In the long term, deploying multiple synchronized depth cameras could offer a more comprehensive 3D capture of moving animals.

4.7.4 Feature Redundancy and Selection

Feature correlation analysis revealed strong collinearity among certain metrics, particularly those related to pixel-based size measures (*pixel_size*, *pixel_ratio*, and *non_zero_pixel_count*). While ensemble models like LightGBM can handle some redundancy, unnecessary features can increase computational load and risk overfitting. Additionally, constant features such as *aspect_ratio* and *non_zero_pixel_count* (which were zero across all samples) offered no variance, reducing their informativeness. Future work should apply more rigorous feature selection techniques, such as Recursive Feature Elimination (RFE), Principal Component Analysis (PCA), or mutual information-based ranking, to enhance model efficiency. Feature engineering can also be further refined by exploring spatially aware features (e.g., localized depth gradients, sectional volume heatmaps) or learned representations through unsupervised pre-training.

4.7.5 Overfitting and Generalization Risk

Despite the use of five-fold cross-validation, the limited and relatively controlled dataset poses a risk of overfitting. The model may have learned dataset-specific artifacts (e.g., lighting uniformity, consistent pen backgrounds, or animal behavior during capture) rather than general weight-indicative features. This is reflected in the occasional overestimations in mid-weight ranges and slightly higher validation RMSE at certain epochs. Addressing this issue will require not only larger and more diverse datasets, but also regularization techniques (e.g., dropout, weight decay), data augmentation (e.g., rotation, scale jitter, synthetic occlusion), and further experimentation with architecture depth and feature complexity.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

This chapter provides the summary of the results obtained in this study and gives some recommendations for further investigation.

5.1 Summary of Findings

The study's findings address the initial research questions by confirming the effectiveness, reliability, and diverse applications of telemetry systems. The "Summary of Findings" section provides a concise overview of the key results from your research. This section should be factual and focus on presenting the data without interpretation. It should include:

Key Results:

Briefly summarize the most significant findings. Use bullet points or numbered lists for clarity if appropriate. Present the data as it was found, highlighting major patterns, relationships, or trends. Data Presentation:

Include tables, graphs, or charts that succinctly summarize the data. Make sure each visual aid is clearly labeled and includes a brief description. Coverage of Research Questions:

Address each of the research questions or hypotheses posed at the beginning of the study. Summarize the results relevant to each question.

5.2 Conclusion

The "Conclusions" section interprets the findings and discusses their implications. This section should:

Interpret Findings:

Provide an interpretation of the data summarized in the previous section. Discuss what the results mean in the context of the research questions or hypotheses. Implications:

Explain the significance of the findings. Discuss how the results contribute to the field of study or practical applications. Limitations:

Acknowledge any limitations in the study that may affect the results or their interpretation.

5.3 Recommendations

The "Recommendations" section provides actionable suggestions based on the study's findings and conclusions. This section should:

Practical Applications:

Offer specific recommendations for practitioners, policymakers, or other stakeholders based on the findings. Future Research:

Suggest areas for further investigation that could address the study's limitations or build on its findings. Implementation:

Provide guidance on how the recommendations can be implemented effectively.

References

Andrási, K., Dobsinszki, G., and Kovács, L. (2023). Pig weight estimation according to rgb image analysis. *Animals*, 14(14):2122.

Andújar, D., Escolà, A., and Rosell-Polo, J. R. (2016). A review of optical sensing systems for plant phenotyping. *Precision Agriculture*, 17(1):1–18.

Bercovich, A., Edan, Y., Alchanatis, V., Moallem, U., Parmet, Y., Honig, H., Maltz, E., Antler, A., and Halachmi, I. (2013). Development of an automatic cow body condition scoring using body shape signature and fourier descriptors. *Journal of dairy science*, 96(12):8047–8059.

Dickinson, R., Morton, J., Beggs, D., Anderson, G., Pyman, M., Mansell, P., and Blackwood, C. (2013). An automated walk-over weighing system as a tool for measuring liveweight change in lactating dairy cows. *Journal of dairy science*, 96(7):4477–4486.

Dikmen, S., Ustuner, H., and Orman, A. (2012). The effect of body weight on some welfare indicators in feedlot cattle in a hot environment. *International journal of biometeorology*, 56:297–303.

Dohmen, R., Catal, C., and Liu, Q. (2022). Computer vision-based weight estimation of livestock: a systematic literature review. *New Zealand journal of agricultural research*, 65(2-3):227–247.

Fadlelmoula, A. A., Mudarris, M. S., and Hariri, M. S. (2020). Prediction of live body weights in dromedary camels (camelus dromedarius) from morphometric body measurements. *World journal of biology and biotechnology*, 5(1):25–29.

Farooqui, M. U. (2024). *Precision Agriculture and Predictive Analytics: Enhancing Agricultural Efficiency and Yield*, chapter 9. Wiley.

Faucitano, L. and Goumon, S. (2018). Transport of pigs to slaughter and associated handling. In *Advances in pig welfare*, pages 261–293. Elsevier.

Food and Agriculture Organization (FAO) (2022). Global meat consumption statistics. `http://www.fao.org`. Accessed: 2024-12-06.

Franco, M. d. O., Marcondes, M. I., Campos, J. M. d. S., Freitas, D. R. d., Detmann, E., and Valadares Filho, S. d. C. (2017). Evaluation of body weight prediction equations in growing heifers. *Acta Scientiarum. Animal Sciences*, 39(2):201–206.

Gjergji, M., de Moraes Weber, V., Silva, L. O. C., da Costa Gomes, R., De Araújo, T. L. A. C., Pistori, H., and Alvarez, M. (2020). Deep learning techniques for beef cattle body weight prediction. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE.

Gómez, Y., Stygar, A. H., Boumans, I. J., Bokkers, E. A., Pedersen, L. J., Niemi, J. K., Pastell, M., Manteca, X., and Llonch, P. (2021). A systematic review on validated precision livestock farming technologies for pig production and its potential to assess animal welfare. *Frontiers in veterinary science*, 8:660565.

Jun, K., Kim, S. J., and Ji, H. W. (2018). Estimating pig weights from images without constraint on posture and illumination. *Computers and Electronics in Agriculture*, 153:169–176.

Kárpinszky, A. and Dobsinszki, G. (2023). Pig weight estimation according to rgb image analysis.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems*, volume 30, pages 3146–3154.

Khoi (2024). DATA_TONGHOP: Pig detection dataset. `https://universe.roboflow.com/khoi-xzy9e/data_tonghop/model/2`. Accessed: 2025-04-21.

Kirillov, A. et al. (2023). Segment anything. *arXiv preprint arXiv:2304.02643*.

Kollis, K., Phang, C. S., Banhazi, T. M., and Searle, S. J. (2007). Weight estimation using image analysis and statistical modelling: A preliminary study. *Applied engineering in agriculture*, 23(1):91–96.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, volume 25.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521:436–444.

Li, Z., Luo, C., Teng, G., and Liu, T. (2014). Estimation of pig weight by machine vision: A review. In *Computer and Computing Technologies in Agriculture VII: 7th IFIP WG 5.14 International Conference, CCTA 2013, Beijing, China, September 18-20, 2013, Revised Selected Papers, Part II 7*, pages 42–49. Springer.

Ma, W., Qi, X., Sun, Y., Gao, R., Ding, L., Wang, R., Peng, C., Zhang, J., Wu, J., Xu, Z., et al. (2024). Computer vision-based measurement techniques for livestock body dimension and weight: A review. *Agriculture*, 14(2):306.

Mañez, M., Ortega, J., and Aguilar, D. (2020). Performance evaluation of pig breeds in the philippines: A study on landrace pigs. *Philippine Agricultural Science*, 103(3):217–225.

Monteiro, A., Santos, S., and Gonçalves, P. (2021). Precision agriculture for crop and livestock farming—brief review. *Animals*, 11(8):2345.

Neethirajan, S. and Kemp, B. (2021). Digital livestock farming. *Sensing and Bio-Sensing Research*, 32:100408.

Ng, W. C., Soong, Y. Q., and Teh, S. Y. (2023). Machine learning in food security and sustainability. In *Handbook of Sustainability Science in the Future: Policies, Technologies and Education by 2050*, pages 1583–1599. Springer.

Ozkaya, S. (2013). The prediction of live weight from body measurements on female holstein calves by digital image analysis. *The Journal of Agricultural Science*, 151(4):570–576.

Peña Fernández, A., Norton, T., Vranken, E., and Berckmans, D. (2019). Carcass quality traits of fattening pigs estimated using 3d image technology. In *proceedings of the 9th European Conference on Precision Livestock Farming*. ECPLF.

Pezzuolo, A., Guarino, M., Sartori, L., González, L. A., and Marinello, F. (2018). On-barn pig weight estimation based on body measurements by a kinect v1 depth camera. *Computers and electronics in agriculture*, 148:29–36.

Pinto, F., Fernandes, T., Oliveira, L., and Faria, D. R. (2024). Cnn-based estimation of pig weights using top-view images. *Computers in Agriculture*, 3:100059.

Rousing, T., Bonde, M., Badsberg, J. H., and Sørensen, J. T. (2004). Stepping and kicking behaviour during milking in relation to response in human–animal interaction test and clinical health in loose housed dairy cows. *Livestock Production Science*, 88(1-2):1–8.

Samperio, E., Lidón, I., Rebollar, R., Castejón-Limas, M., and Álvarez-Aparicio, C. (2021). Lambs' live weight estimation using 3d images. *Animal*, 15(5):100212.

Shukla, A. and Anand, S. (2016). Metric learning based automatic segmentation of patterned species. In *2016 IEEE International Conference on Image Processing (ICIP)*, pages 3982–3986. IEEE.

Singh, A., Jadoun, Y., Brar, P. S., and Kour, G. (2022). Smart technologies in livestock farming. In *Smart and sustainable food technologies*, pages 25–57. Springer.

Smisek, J., Jancosek, M., and Pajdla, T. (2013). 3d with kinect. *Consumer Depth Cameras for Computer Vision*, 17(2):3–25.

Soussi, A., Zero, E., Sacile, R., Trinchero, D., and Fossa, M. (2024). Smart sensors and smart data for precision agriculture: A review. *Sensors*, 24(8):2647.

Tasdemir, S. and Ozkan, I. A. (2019). Ann approach for estimation of cow weight depending on photogrammetric body dimensions. *International Journal of Engineering and Geosciences*, 4(1):36–44.

Terence, S., Immaculate, J., Raj, A., and Nadarajan, J. (2024). Systematic review on internet of things in smart livestock management systems. *Sustainability*, 16(10):4073.

The Pig Site (n.d.). Weighing a pig without a scale. `https://www.thepigsite.com/articles/weighing-a-pig-without-a-scale`. Accessed: 2024-12-06.

Wang, H., Li, X., Zhao, Y., and Gao, S. (2021). Deep learning in smart farming: A review. *AI Open*, 2:1–13.

Wang, Z., Li, Q., Yu, Q., Qian, W., Gao, R., Wang, R., Wu, T., and Li, X. (2024). A review of visual estimation research on live pig weight. *Sensors*, 24(21):7093.

Witten, I. H., Holmes, G., McQueen, R. J., Smith, L. A., and Cunningham, S. J. (1993). Practical machine learning and its application to problems in agriculture.

Yan, Q., Ding, L., Wei, H., Wang, X., Jiang, C., and Degen, A. (2019). Body weight estimation of yaks using body measurements from image analysis. *Measurement*, 140:76–80.

Zhang, Z. (2012). Microsoft kinect sensor and its effect. *IEEE Multimedia*, 19(2):4–10.