



Computer Vision for Resource Optimisation in Dairy Cows Drinking and Backyard Chickens Nesting behaviour

Bachelor's Thesis
Climate Smart Agriculture
1st Semester 2025
Dimitrios Tousios

Glossary

CNN (Convolutional Neural Network): A specialized artificial neural network architecture designed for processing visual data through layers of mathematical filters (convolutions), enabling automated feature extraction and pattern recognition in images

Confidence Threshold: A minimum probability value that determines whether a detection is considered valid, used to filter out uncertain predictions

deepSORT: An extension of the Simple Online and Realtime Tracking (SORT) algorithm that integrates deep learning features for enhanced object tracking persistence across video frames

IoU (Intersection over Union): A metric used in object detection to measure accuracy by calculating the overlap between predicted and actual object locations

PLF (Precision Livestock Farming): Technology-driven approach to livestock management that uses continuous automated monitoring and control of animal production, health, welfare, and environmental impact

State Machine: A computational model representing system behaviour through discrete states and transitions, used in this research for robust behaviour monitoring

Temporal Persistence: A validation approach in computer vision that requires detections to maintain consistency over multiple consecutive frames to confirm events

YOLO (You Only Look Once): An efficient object detection architecture that processes images in a single pass through the neural network, enabling real-time detection by simultaneously predicting bounding boxes and class probabilities

Table of Contents

1	Introduction	1
2	Literature review	3
2.1	Early development of Computer Vision.....	3
2.2	Animal monitoring in livestock agriculture.....	5
2.3	Computer Vision in Precision Livestock Farming	6
2.4	Computer Vision applications across livestock species	8
2.5	Practical applications of computer vision in PLF: dairy cows	9
2.6	Practical applications of computer vision in PLF: poultry	11
2.7	Challenges of computer vision integration	13
3	Materials and methods.....	14
3.1	Common computer vision implementation pipeline	15
3.2	YOLO architecture.....	16
3.3	Main performance metrics.....	17
4	Case studies	20
4.1	Case study one: Cow drinking behaviour	20
4.1.1	Setup.....	21
4.1.2	Data acquisition	22
4.1.3	Data preparation and labelling	22
4.1.4	Training	23
4.1.5	Deployment.....	27
4.1.6	Results	30
4.2	Case study two: Chicken nesting behaviour	34
4.2.1	Setup.....	36
4.2.2	Data acquisition	37
4.2.3	Data preparation and labelling	37
4.2.4	Training	39
4.2.5	Deployment.....	43
4.2.6	Results	45
5	Discussion.....	51
6	Acknowledgements.....	55
	References.....	56

Figures

Figure 1. Computer vision model deployment pipeline	15
Figure 2. Mosaic data augmentation example from initial training batch.....	24
Figure 3. Training performance metrics for custom YOLOv8x implementation showing performance progression across training epochs.....	25
Figure 4. Precision-confidence and recall-confidence curves demonstrating model detection performance	26
Figure 5. Precision-recall curve and F1-score analysis for YOLOv8x model.....	26
Figure 6. Confusion matrix analysis of YOLOv8x model performance showing true positives, false positives, and false negatives across detection classes.....	27
Figure 7. Detection of single cow drinking behaviour	29
Figure 8. Concurrent detection of multiple cow drinking behaviour.....	29
Figure 9. Data structure implementation for behavioural interaction storage	30
Figure 10. Analysis of tracking inconsistencies in unique head identification	31
Figure 11. Distribution analysis of concurrent drinking behaviours categorized by interaction type, showing percentage breakdown of single, dual, and multi-animal events	32
Figure 12. Temporal distribution of validated drinking interactions	33
Figure 13. Weekly temporal distribution of monitoring data.....	33
Figure 14. Arrangement of nest box monitoring setup.....	35
Figure 15. Demonstration of object annotation methodology for poultry monitoring.....	39
Figure 16. Implementation of data augmentation techniques for model training	40
Figure 17. Training performance metrics for poultry monitoring model training across epochs	41
Figure 18. Precision-confidence and recall-confidence analysis for poultry detection model.....	42
Figure 19. Precision-recall and F1-confidence analysis for poultry detection	42
Figure 20. Confusion matrix analysis of poultry detection model	43
Figure 21. Detection performance under optimal visibility conditions.....	46
Figure 22. Detection of individual nesting behaviour	47
Figure 23. Model performance analysis under challenging lighting conditions	47
Figure 24. Comparative analysis of species-specific activity patterns.....	48
Figure 25. Temporal patterns in nest box usage.....	48
Figure 26. Spatial analysis of nest box utilization patterns	49
Figure 27. Temporal distribution of peak nest box occupancy	49
Figure 28. Analysis of detection challenges during complete egg occlusion	50
Figure 29. Analysis of egg production patterns	51

Tables

Table 1. Performance metrics for Computer Vision model evaluation	18
Table 2. Performance metrics of the custom YOLOv8x model implementation.....	24
Table 3. Performance metrics of the YOLOv8x model for poultry monitoring	41

1 Introduction

The modern agricultural field has undergone numerous transformations, each reflecting the industrial revolution and the technology of its period (Liu et al., 2021). These transformations showcase a symbiotic relationship between the agricultural and industrial sectors, where advancements in one drive innovations in the other. Beginning with the first manual mechanical tools during the 18th century, the field moved into the mechanization-focused era of the green revolution in the early 20th century. This shift to mechanization brought increased productivity but also presented new challenges, which encouraged further innovation. As a result, the latter half of the 20th century saw the emergence of precision-focused agriculture, aiming to address the challenges introduced by mechanization. In essence, agriculture has evolved through a series of stages, akin to climbing a ladder: from land labour giving birth to tools, to tools bringing productivity along with new challenges, and then inventing new tools to overcome these challenges. The evolution of agricultural technology reflects not just advancement in tools, but an expanding understanding of agricultural ecosystems. Each technological step has provided new lenses through which we observe and interpret agricultural processes, from simple mechanical innovations to today's sophisticated digital systems. Currently, agriculture is transitioning into what is known as Agriculture 4.0 (Parras-Burgos et al., 2023), a technologically advanced form of agriculture paralleling the technologically driven Industry 4.0.

This new era builds upon the benefits and challenges of previous stages, learning from past successes and failures to create a more intelligent and efficient system. This parallel focus towards the technological direction can also be observed by the significant increase in research exploring the applications of machine learning (ML), deep learning (DL), and neural networks (NN) in combination with agriculture, during the past two decades (Dokic et al., 2020). Machine learning enables systems to learn and improve from experience without explicit programming, while deep learning, a specialized subset of ML, uses multiple layers of neural networks to process complex patterns in data. These technologies have revolutionized agricultural monitoring by enabling automated analysis of vast amounts of farm data, from image recognition for crop disease detection to behavioural pattern analysis in livestock. The advancement of these technologies during the same period paved the way for the development of integrated systems using computer vision (CV), which stand out due to their ability to interpret and make decisions based on visual data,

such as images or videos, obtained through either commercially available security cameras, or in more specialized cases, 3D and infrared cameras.

In livestock agriculture, CV can be used to monitor, track, and record detailed animal behaviour, assess health and body conditions in real-time, and optimize resource allocation. At the same time, it does it with minimal animal intrusion and with a good sensor-to-animal ratio, being scalable to work with low and high number of animals and through variety of domestic species. For instance, CV tools have been developed to detect early signs of disease in dairy cows by monitoring behaviours such as standing time to identify hoof diseases or feed intake to detect mastitis (Yu et al., 2024), and monitoring respiratory responses helps assess heat-stress adaptation (Wu et al., 2023). In chickens, CV technologies have been studied for health assessments by tracking movement and posture to detect diseases (Bhuiyan & Wree, 2023).

However, the integration of computer vision into animal monitoring faces several challenges that hinder its wider use and applicability. These challenges include dynamic variability in natural phenomena such as animal behaviour and environmental factors like lighting and weather conditions, occlusion of animals in group settings, limited availability of annotated livestock data, significant computational resources required for training deep learning models, and the presence of misinformation, which stems from CV systems lacking the ability of critical thinking. Addressing these challenges is crucial for the effective application of computer vision in agriculture.

In this context, the present research aims to investigate the potential of CV technology to enhance animal welfare and optimize resource allocation in two distinct agricultural settings through case studies focusing on monitoring dairy cow drinking behaviour within a campus cowshed and monitoring chicken nesting behaviour in a backyard coop environment. By examining these case studies, the thesis' aim is to analyse and discuss their results in relation to the identified challenges in the literature and provide recommendations for further integration and optimization of CV techniques for each case.

Research question

How can computer vision systems be effectively implemented in different agricultural scales to optimize resource allocation through behaviour monitoring, while addressing the technical challenges of environmental variability and limited datasets?

Objectives:

1. Analyse case study results demonstrating CV implementation at different agricultural scales, addressing technical challenges in behaviour monitoring and resource optimization
2. Provide recommendations for effective CV integration in both institutional and individual farming contexts based on case study outcomes.

This thesis report is structured as follows: Section 2 presents a literature review on the development of computer vision, its role within livestock agriculture, and its different applications compared to existing monitoring methods. It includes examples of its use in dairy cattle and poultry livestock agriculture and animal welfare scenarios and notes the challenges of its integration and deployment. Section 3 describes a common CV model deployment pipeline, the specific methodology adopted for the case studies, as well as the performance metrics used to evaluate it. Sections 4.1 and 4.2 detail the case studies on dairy cow drinking behaviour and backyard chicken nesting behaviour, respectively. Section 5 discusses the results and of potential further integration of computer vision in each case study, while Section 6 concludes with acknowledgements.

2 Literature review

2.1 Early development of Computer Vision

Early efforts in computer vision, dating back to the 1960s, focused on simple pattern recognition and rudimentary image analysis. Larry Roberts, a pioneer often regarded as the father of CV, made significant contributions by exploring how to extract 3D information from 2D images in his Ph.D. thesis at MIT, laying the cornerstone for future exploration (Huang, 1996, p. 1). The 1960s and 1970s saw the birth of the "Summer Vision Project" at MIT (Papert, 1966), where visionaries like Marvin Minsky and Seymour Papert attempted to develop systems capable of figure-ground analysis and pattern recognition. Though the project did not yield immediate success, it sowed the seeds for a field that would flourish in the decades to come. A significant milestone emerged in the late 1970s when David Marr, a British neuroscientist and psychologist proposed his theory of human vision as a form of information processing system, proposing a theory three distinct levels of analysis (computation, algorithmic, implementational) which must be understood in order to understand how the process of vision functions. His vision highlighted the importance of

low-level image processing to create a "primal sketch," a precursor to higher-level interpretations, and one of the biggest influences for image processing algorithms (Huang, 1999, p.2). The primal sketch extracts fundamental components of a scene, such as edges and regions, similarly to a quick pencil sketch by an artist. Following this, a 2.5D sketch incorporates textures and depth, and finally, a 3D model represents the scene in a continuous, three-dimensional map (Marr, 2010; Stevens, 2012).

Despite these advances, early CV systems faced significant challenges. Tasks such as edge detection, segmentation, and 3D reconstruction were computationally intensive and required significant manual intervention. These early CV systems, which worked only under controlled conditions, were still concluded as "rather disappointing" even after 30 years of development (Huang, 1999, pp. 3-6), mainly due to their limited practical applications. The primary challenges stemmed from high computational requirements which made processing complex variables in real-time almost impossible at the time, and low accuracy which prevented robust recognition of objects under diverse conditions. Another limiting factor was the extremely limited number of available public datasets of the time, which limited in data size and lacking data diversity, made it difficult for researchers to develop algorithms that could handle such variations in lighting, orientation, and occlusion; variations affecting almost all practical scenarios, including livestock agriculture.

It was not until the technological developments during the last two decades of 2000 to 2010, and particularly the advancement of computer graphics processing units (GPUs) and their parallel nature of execution (Oh et al., 2004) that enabled a faster and more effective way of running the existing algorithms, and the development of further sophisticated ones. As GPU technology become more commercially accessible, computer graphic cards begun finding their place in general-purpose computing (Macedonia, 2003), soon revolutionizing the computing field by offering computational capabilities previously available only in high-end expensive computers. The necessary computational power now available to train complex deep learning models, and the large datasets increasing in abundance, resulted in favourable conditions which enabled more extensive and intensive development in the fields of ML and DL.

By 2010, computer vision had advanced to perform a variety of high-end tasks such as object recognition, face detection, fingerprint detection, and autonomous vehicle and robotic navigation (Kakani et al., 2020, pp. 3-5). These advancements marked a significant leap from the early, more primitive systems, demonstrating the potential of CV to impact various fields significantly.

The evolution of CV capabilities paralleled growing needs across multiple sectors, from industrial automation to medical imaging. In agriculture, this technological maturation arrived at a moment when traditional farming practices were being challenged by the need for more sustainable and efficient approaches. The agricultural sector's adoption of CV technology followed a pattern that reflected both the technology's growing capabilities and agriculture's unique challenges. Initial applications focused on relatively static scenarios like crop inspection and grading (Tian et al., 2020; Kakani et al., 2020), where controlled conditions could be maintained. However, as the technology matured, applications expanded to include more dynamic situations such as real-time monitoring of livestock behaviour, plant disease detection, and automated harvesting systems (Fernanded et al., 2020; Borges Oliveira et al., 2021; Li S. et al., 2021)

This trajectory from controlled to dynamic applications mirrors the broader evolution of the technological integration in agriculture – from simple mechanization to intelligent, adaptive systems capable of responding to complex biological patterns. This progression also marks a shift in approach: from technology just replacing human labour to technology enhancing human understanding of agricultural ecosystems.

2.2 Animal monitoring in livestock agriculture

Traditional methods of animal monitoring in livestock agriculture have relied heavily on manual observation and simple mechanical devices (Kumar et al., 2023). Producers typically observe livestock at regular intervals to monitor health and behaviour, as for example, respiration rates which are often estimated by visually observing the abdominal movements of animals and involves counting the number of breaths per minute (Milan et al., 2016; Lowe et al., 2019; Wu et al., 2020). While these manual methods benefit from human intuition and experience, they face limitations in consistency, scalability and continuous monitoring capability.

Mechanical devices such as pedometers and RFID tags offer more consistent monitoring compared to manual methods but come with their own set of limitations. Pedometers are used to track the movement and activity levels of animals, which can help identify changes in behaviour indicative health issues such as estrus, lameness, or general lethargy. For example, a significant decrease in movement might indicate illness or injury, while an increase could signal estrus (Neethirajan & Kemp, 2021; Helwatkar et al., 2014). RFID systems are also used to monitor laying behaviour in poultry by detecting when hens enter

and leave nests, allowing analysis of laying patterns and detection of issues such as egg-laying disorders or nest preferences (Chien & Chen, 2018), and can track feeding behaviour by monitoring the duration and frequency of visits to feeding stations, which helps in understand feed patterns and identifying animals that may not be getting adequate nutrition (Zhang et al., 2016).

However, these technological solutions face several practical challenges in real-world implementation. RFID tags are commonly used for identification and basic movement tracking but cannot provide detailed health metrics or behaviour insights (Rayhana et al., 2021). Basic RFID systems are also prone to various sources of signal interference, such as metal structures, environmental factors like moisture and temperature, and other potential radio signal transmitting devices, including other RFID systems (Knapp & Uckelmann, 2022). The physical presence of monitoring devices can also influence animal behaviour – thoracic belts used for respiratory monitoring may slip or be damaged by other animals, while visible sensors on poultry can trigger aggressive responses from other birds (Dennis et al., 2008).

The limitations become more pronounced as the herd sizes increase. Modern farming operations often manage larger number of animals with fewer personnel, leading to reduced individual attention and potential delays in identifying health issues (Stewart et al., 2017). The challenges highlight the need for monitoring solutions that can scale efficiently while maintaining accuracy and reliability. The physical attachment of devices presents practical challenges in both mounting/wearing and maintenance, while performance often degrades when devices loosen or signal is disturbed or become weak (Atkins et al., 2018). These limitations in current monitoring approaches create a clear need for more sophisticated, non-invasive monitoring solutions that can provide autonomous and continuous, detailed insights without disrupting natural animal behaviour.

2.3 Computer Vision in Precision Livestock Farming

The field of CV has emerged as a revolutionary solution for livestock monitoring, offering a non-invasive, scalable approach that improves upon traditional methods. The growing interest in AI technologies, including the subfields of CV, is evidence by the increasing number of research papers over the past two decades (Dokic et al., 2020), highlighting its potential and advancements.

The transition from conventional monitoring to CV-based systems represents more than just a technological upgrade – it marks a shift in how we approach animal observation and welfare assessment. CV systems take advantage of existing or real-time visual data from cameras installed in barns or fields, providing continuous and remote monitoring without the need of frequent human presence.

The scalability of CV systems presents a particularly compelling advantage as agricultural operations grow in size. These systems can monitor larger herds simultaneously, providing real-time data on each individual animal as well as the whole herd or flock, a feat which is anticipated to grow in importance as the ratio of farmers-to-animals is expected to widen, leading to less farmers managing more animals (Halachmi et al., 2019;). The monitoring capabilities of CV systems extend far beyond simple presence detection – CV systems can track a wide range of behaviours and health indicators, such as feeding patterns, movement, and posture. For example, Neethirajan & Kemp (2021) demonstrate how CV can detect early signs of illness by analysing changes in an animal's gait or every behaviour, while Borges Oliveira et al., (2021) note that by analysing feeding and drinking behaviours, CV systems can optimize feed distribution, ensuring that all animals receive adequate nutrition.

Such environments align well with the concepts found within Precision Livestock Farming (PLF), which strive to improve animal welfare and farm management by integrating various sensors (image, sound, other body sensors), data analytics, and automated systems for continuous and real-time monitoring of all the critical stages of livestock farming (Berckmans, 2017). The concept of PLF began in the 1970s to early 1980s, evolving from basic electronic milk meters for cows to sophisticated systems capable of monitoring a wide array of health and behavioural indicators. Over the decades, PLF has seen numerous milestones, including the introduction of automated milking systems, behaviour-based estrus detection, and rumination activity monitoring (Halachmi et al., 2019).

The integration of CV within PLF frameworks has been particularly transformative. Remote monitoring, modern data collection through sensors, rapid data transfer, vast data storage, and the internet of things (IoT) have significantly advanced PLF in many fields during the last 20 years, including:

- Aerial-and satellite-based measurement of pastures forage quantity and quality
- Body weight and composition assessment
- Physiological measurements and behavioural monitoring

- Location tracking and activity monitoring in grazing environments
- Early detection of lameness and other diseases
- Milk yield and composition analysis
- Reproductive measurements and calving monitoring
- Feed intake and greenhouse gas emissions assessment

The European Conference on Precision Livestock Farming (EC-PLF), launched in 2003, and the European Precision Livestock Farming Project (EU-PLF) in 2012 have been instrumental in promoting research and technological advancement in this field (Jiang et al., 2023). According to Bhujel et al. (2024), CV plays a vital role in PLF by allowing for precise monitoring of individual animals and larger herds alike through its non-invasive operation and minimal stress on the animals, ensuring more natural behaviour and accurate health assessments. Additionally, CV systems' scalability allows for better value-per-animal monitoring, where herds or flocks of animals can be monitored through fewer devices (cameras), scaling in a sense of "a sensor per herd rather than a sensor per animal" (Berckmans, 2017, Halachmi et al., 2019).

The integration of CV within PLF frameworks represents a significant advancement in our ability to understand and manage livestock welfare. The continuous, detailed monitoring capabilities, combined with sophisticated data analysis, enable a more proactive and comprehensive approach to animal care.

2.4 Computer Vision applications across livestock species

The implementation of CV in livestock monitoring has shown adaptability across different species, though development progress varies significantly based on the available annotated datasets and industry adoption rates. A recent systematic review by Rohan et al. (2024) highlights the rapid evolution of deep learning applications in livestock behaviour recognition, noting particular advancement in species-specific behavioural pattern detection. Advancements in individual animal identification have also shown promising results, with Bakhshayeshi et al. (2024) demonstrating successful implementation of Siamese neural networks for cattle reidentification, achieving significant accuracy improvements in tracking individual animals across different monitoring periods. However, the distribution of machine-ready data currently heavily favours certain livestock sectors, with dairy cattle and swine leading in terms of available annotated datasets for CV development and validation (Borges Oliveira et al., 2021; Bhujel et al., 2024).

In swine production, CV systems have demonstrated particular value in monitoring social behaviour and welfare conditions, benefiting from the second-largest collection of annotated datasets in livestock applications. Recent implementations focus on tracking aggressive interactions, feeding patterns, growth monitoring, posture tracking, with systems achieving notable accuracy in individual pig identification within group housing conditions. (Lei et al., 2019; Benjamin & Yik, 2019; Shao et al., 2021)

Small ruminant farming (sheep and goats) present both different challenges and even more limited machine ready data availability. Recent CV applications in these sectors explored on sheep rumination (Yu et al., 2024), identification of individual animal (Yadav et al., 2024), and pose independent body score estimation (Temenos et al., 2024). While these systems show promise in identifying individual animals in field conditions, the combination of environmental factors and limited annotated datasets has slowed development compared to cattle and swine applications. Additionally, the outdoor nature of most small ruminator operations add complexity to data collection, further contributing to the slower pace of CV integration in these sectors.

This disparity in available machine-ready data and validated implementations across livestock species highlight a vital challenge in CV development for agricultural applications. While the potential benefits of CV technology are relevant across all livestock sectors, the pace and depth of implementation is significantly influenced by the data availability and environmental conditions, and the existing technological development of each sector up to this day. This context explains why certain livestock sectors, particularly dairy cattle and confined swine operations, see a more rapid advancement in CV applications, while the development in other sectors continuous at a more measured pace.

2.5 Practical applications of computer vision in PLF: dairy cows

The dairy sector presents a unique context for CV implementation, mainly due to its existing technological infrastructure and automation history. Larger animals such as dairy cows offer numerous positions on and in their bodies for sensors to be installed providing valuable monitoring data. While these sensors have been in use for decades, the addition of CV for enhances animals monitoring or the reduction/replacement of invasive methods offers a variety of benefits, particularly in automating tasks previously requiring significant human labour.

Automated genetic trait phenotyping is emerging as one of the most promising applications of CV in livestock farming, enabling large-scale phenotyping of animal behaviour and physical traits. Many key behaviour traits exhibit sufficient heritability, making genetic selection and breeding for positive characteristics possible. Traits such as ease of handling, good maternal behaviour, and the reduction of harmful social behaviour are prime examples of traits that be targeted. Traditionally, phenotyping information for these traits has been gathered through direct observation (real-time observation or through recorded material), behaviour testing, or proxy measures like skin lesion scores (indicating aggression) or tail damage scores (indicating tail biting). However, these methods have proven too labour-intensive and costly for widespread commercial use (Wurtz et al., 2019).

The identification of reproductive events, such as estrus and calving, plays another crucial role in optimizing breeding strategies. Initial attempts by Cangar et al. (2007; 2008) and later Van Hertem et al. (2013) focused on using image analysis to monitor pre-calving dairy cows by quantifying changes in standing and lying behaviours that signal the onset of calving, and later determine whether human intervention was required during the calving. While these early implementations struggled with dynamic environmental conditions, they laid important groundwork for future developments. Similarly, Tsai et al. (2014) developed a CV-based system for estrus detection, achieving a 72% detection rate by identifying behaviours such as following and mounting – key indicators of estrus.

Early and automated detection of lameness represents another widespread factor affecting both animal welfare and overall farm productivity and profitability. The evolution of lameness detection systems illustrates the progression of CV technology in dairy farming:

- Initial 2D approaches using side-view images tracked hoof movement and calculated hoof trackways (Song et al., 2008)
- Advanced systems assessed back posture (Poursaberi et al, 2010; Viazzi et al., 2013)
- Integration of 3D camera systems improved detection accuracy in dynamic environments (Viazzi et al., 2014)
- Commercial implementation achieved 81.5% accuracy in generalized environments (Van Hertem et al, 2014)
- Recent systems like that developed by Hansen et al. (2018) simultaneously assess multiple parameters (body condition, weight, lameness) using a single 3D camera, achieving BSC estimation within ± 0.25 , weight within $\pm 5\%$, and lameness classification at 83% accuracy under real-world farm conditions

The progression towards individual animal monitoring marks a crucial advancement in PLF. Moving from herd-level assessments to detailed, real-time monitoring of each cow ensures a more personalized approach to health and productivity management. Recent developments in individual identification showcase a significant progress:

- Shen et al. (2019) achieved 96.65% accuracy in cow identification using side-view images and fine-tuned CNN
- Tassinari et al. (2021) developed neural network-based recognition achieving mean average precision between 0.64 and 0.66
- Li, S. et al. (2021) created a lightweight convolutional neural network achieving 97.95% accuracy while addressing robustness and computational efficiency

The evolution of CV applications within dairy farming reflects a broader trend: the integration of CV technology with existing automated systems to create comprehensive monitoring solutions. Unlike other livestock sectors, dairy farming's established automation infrastructure has provided a foundation for CV implementation, allowing these systems to focus on directly enhancing or replacing laborious human tasks while maintaining high standards of animal welfare monitoring

2.6 Practical applications of computer vision in PLF: poultry

Unlike the dairy cattle sector where the technology has been extensively explored since the early 2000s, the application of image-based analysis techniques in poultry farming has primarily focused on detecting and tracking bird positions, and then assessing the birds' activity levels and distribution patterns, with the ultimate goal of linking the data to health and welfare status. This different trajectory reflects both the unique characteristics of poultry farming – where large flocks in controlled environments present distinct monitoring challenges – and the sector's emphasis on health surveillance over labour automation.

Both laying hens and broiler flocks have been the subject of such studies, with the intention of evaluating their behaviour and establishing connections to wellbeing and welfare. While the initial focus of these image-based techniques was on detecting location and lying hens in the early 2000s, more recent studies have shifted towards tracking individual birds, allowing for a more detailed characterization of specific behaviours. This progression, however, has primarily focused on commercial operations, with smaller-scale or alternative

farming scenarios such as backyard poultry operations remaining at large unexplored in published research.

Early applications demonstrated promising results even with limited technology. Leroy et al., (2005; 2006) used 2D monochromatic top-view images and the reference image method to identify six distinct hen behaviours, achieving an overall accuracy ranging from 70% to 96%. However, lower accuracy in pecking detection was observed due to challenges in tracking hens' heads. Cronin et al. (2008) employed cameras with infrared light to count hen legs in cages and detect foreign objects in egg collection belts to reduce potential blockage, obtaining 79% accuracy for the leg counts and 95% for foreign object detection. While these studies contributed to the development of automated techniques for monitoring hen health and welfare, environmental factors and hen movements presented obstacles to accuracy.

More recent developments show significant advancement in both scope and accuracy:

- Kristensen & Cornou (2011) developed block-based activity analysis with dynamic linear regression for early flock problem detection
- Kashiha et al. (2013) implemented this commercially in a broiler house containing 28,000 animals, achieving 95% accuracy in detecting environmental discomforts
- Aydin et al. (2010, 2013, 2017) conducted extensive studies on lameness in broilers, culminating in a 3D image-based method with 93% accuracy in classifying lying events
- Zhuang et al. (2017) achieved 99.5% accuracy in identifying diseased broilers using posture analysis
- Recent systems like ChickTrack (Neethirajan, 2022) and developments by Mohialdin et al. (2023) demonstrate high accuracy in automated detection and behaviour analysis

A comprehensive review by Okinda et al. (2020) highlights both conventional machine learning and deep learning approaches showing potential for bio-process and bio-response surveillance. While significant progress has been made, challenges such as complex backgrounds, variable lighting, and occlusion persist. These challenges become particularly pronounced in non-commercial settings, such as backyard chicken operations, where environmental conditions are less controlled and individual bird monitoring becomes more relevant.

Notably, while CV technology in poultry has focused primarily on visual indicators, birds communicate significant health, welfare, and behaviour information through vocalizations, providing a rich amount of data that presents both challenges and opportunities for CV implementation. The integration of visual monitoring with other sensing technologies might offer even deeper welfare assessment capabilities, particularly in smaller-scale operations where individual bird monitoring is both feasible and desirable

2.7 Challenges of computer vision integration

Applying CV in PLF comes with several challenges. The sources reviewed highlight a range of issues that arise when implementing CV systems in livestock agricultural settings. These settings can be broadly categorized in the following points:

Environmental and Operational Variability: One of primary challenges in applying CV for livestock monitor is dealing the varying environmental conditions. Animal farms and barns have complex backgrounds, including feeding stations, water stations, and other equipment that can interfere with accurate detection and tracking. Varying lighting conditions, such as changes in natural light throughout the day or artificial lighting variations, further complicate the visual data captured by cameras. Neethirajan (2021) mentions the difficulty of recognizing and tracking birds due to complex backgrounds and varying lighting conditions, while Mohialdin et al. (2023) identifies challenges related to ambient light variations affecting the accuracy of CV systems.

Occlusion Problems: Animals often move in groups, causing occlusions where one animal blocks another from the cameras view. This makes it difficult for CV systems to accurately track individual animals over time. The problem of occlusion is especially pronounced in dense environments where animals are kept in close quarters, or if the structure of the housing is not allowing for clear visuals (Neethirajan, 2021).

Data Collection and Labelling: Effective CV systems require large datasets of labelled images to train machine learning models. Collecting and labelling these datasets in labour-intensive and time-consuming, as discussed by Tian et al. (2020), who highlights the need for high quality annotated data. This challenge is particularly emphasized by Supriyanto et al. (2023), who mention the lack of standardized and augmented datasets as a significant hurdle in developing robust CV applications. Li G., et al., (2021) notes how publicly available datasets commonly contain generic images, with few annotated datasets

available for animal farming purposes, a condition difficult to tackle due to data privacy concerns.

Detection Accuracy: Training deep learning models used in CV systems demands substantial computational power, and large volumes of high-quality data used for processing requires significant amount of storage space (Stygar et al., 2021). This can be a barrier for those with limited access to modern computing infrastructure and reduces the potential cost-effectiveness of CV integration (Ma et al., 2024)

Critical Thinking and Bias in AI systems: Perhaps the most significant challenge lies in the fundamental nature of CV and AI systems themselves. Unlike human observers who can apply critical thinking and context-aware decision-making, CV systems are limited by their training data and programmed algorithms. This vulnerability becomes especially critical in agricultural application where misinformation or ill-conditioned data in training datasets can lead to incorrect predictions with very high degrees of certainty. The CV systems lack the wisdom to distinguish between current and incorrect data and cannot self-correct the potential mistakes (Tedeschi et al., 2021). This limitation emphasizes the importance of maintaining human oversight and expertise in CV implementation, ensuring that the technological advancement enhances rather than replaces human judgement in agricultural decision-making.

The exploration of literature revealed both the potential and the challenges of computer vision in agricultural settings. These insights directly shaped the methodological approach, particularly in addressing the identified challenges of environmental variability and limited datasets. The progression from theoretical understanding to practical implementation took under consideration both established frameworks and innovative solutions. The methodological framework of this research was based on the adoption of YOLO architecture, selected for its demonstrated effectiveness, while utilizing customized adaptations to address the unique demands of each setting

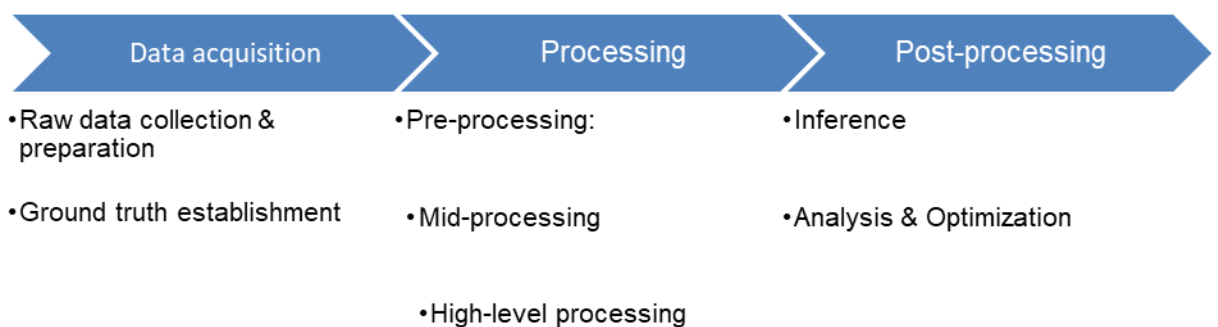
3 Materials and methods

Following the review of literature that highlights both the potential and the challenges of CV in livestock monitoring, this section outlines the methodological framework employed in the case studies. Beginning with an overview of common CV implementation approaches, following with the specific architectures and the metrics used in our research.

3.1 Common computer vision implementation pipeline

Computer vision relies on various deep learning methods which have been intensively developing during the last two decades, such as Convolutional Neural Networks (CNNs), Deep Belief Networks (DBNs), Deep Boltzmann Machines (DBMs), and autoencoders (Voulodimos et al., 2022). Irrespective of the specific method employed, however, these algorithms – referred to as “models” within CV context – share a common set of steps. As depicted in Figure 1, the primary stages of CV include data acquisition, processing, and post-processing, each comprising of several sub-steps, resulting in the model’s development.

Figure 1. Computer vision model deployment pipeline



1. **Data acquisition:** First step encapsulating initial requirements setting the stage for upcoming operations:
 - Collection of visual data (images, videos) from target environment
 - Initial quality assessment and preprocessing
 - Dataset organization and management
 - Generation of ground truth data through manual annotation
 - Validation of ground truth quality and consistency

The establishment of ground truth data is a critical foundation for CV system development. As noted by Wurtz et al., (2019), the quality and accuracy of annotations directly impact a model’s learning capacity and ultimate performance. This human-labelled data, consisting of precisely marked object locations, classifications, and various behavioural states, serves as the definitive standard against which the model’s predictions are evaluated (Tian et al., 2020). The careful creation of this ground truth data requires expertise in both the technical aspects of annotation and domain-specific knowledge of the observed subjects.

2. Processing: The processing stage encompasses three levels of complex operations:

- Pre-processing
 - i. Noise reduction and filtering
 - ii. Input data normalization
 - iii. Dataset augmentation
 - iv. Format standardization.
- b. Mid-processing:
 - i. Feature extraction from visual data
 - ii. Implementation of neural network layers
 - iii. Convolution operations
 - iv. Initial pattern recognition
- c. High-level processing:
 - i. Object detection and classification
 - ii. Movement tracking
 - iii. Scene understanding

3. Post-processing/Inference: The final stage focuses on output inference, optimization and analysis through:

- Model performance evaluation
- Real-time feedback processing
- Result visualization
- Data interpretation and refinement

While this pipeline provides a general framework, the specific implementation depends heavily on the chosen architecture. In our case studies, YOLO (You Only Look Once) architecture was employed, which has shown promise in livestock monitoring due to its ability to handle real-time detection and tracking requirements.

3.2 YOLO architecture

The evolution of Convolutional Neural Network (CNN) models marks a significant progression in addressing the computational challenges inherent in real-time object detection. Early architectural approaches, including the Region-Based Convolutional Neural Network (R-CNN) and its iterations (Fast R-CNN and Faster R-CNN), while groundbreaking in their advancement of image segmentation and object detection

capabilities, faced significant constraints in processing speed and resource utilization. These early models operated through a systematic but computationally intensive process: dividing images into overlapping patches, applying convolution operations with predefined filters, processing each segment independently, and finally combining results for comprehensive detection.

The YOLO architecture's evolution, particularly through YOLOv3 (Redmon & Farhadi, 2018), established key principles for real-time object detection that continue to influence modern implementations. Unlike its predecessors, YOLO revolutionized the approach by implementing a single-pass processing system that simultaneously analyses the entire image through a grid-based detection system (Terven et al., 2023). This innovative architecture enables concurrent prediction of bounding boxes and class probabilities, achieving real-time processing capabilities while maintaining detection accuracy. The significance of this advancement becomes particularly apparent in agricultural applications, where YOLO's architecture offers several crucial advantages: the capability to handle multiple object classes simultaneously, the availability of pre-trained models for transfer learning, and scalability to accommodate varying computational requirements.

The performance evaluation of YOLO and similar models centers on two primary aspects: localization precision, which measures the accuracy in object location and boundary definition, and classification accuracy, which assesses the reliability in object type identification. These aspects are quantified through specific performance metrics detailed in the following section.

3.3 Main performance metrics

The evaluation of CV model performance requires comprehensive metrics that assess different aspects of detection and classification capability against established ground truth data. As detailed by Padilla et al., (2020) in their comprehensive survey of object detection metrics, these measurements can systematically be grouped into three main categories. This categorization provides a structured approach to model evaluation as seen in Table 1.

The evaluation framework consists of several interconnected metrics, each serving specific analytical purposes:

Table 1. Performance metrics for Computer Vision model evaluation

Confidence threshold	Conditional probability of an output being correct	0 - 1									
Intersection over Union (IoU):	Ratio between the intersection and the area of union between 2 bounding boxes	0 - 1									
Precision	Proportion of True Positives among all predicted objects	$P = \frac{TP}{TP + FP}$									
Recall	Proportion of (True Positives) among True positives and False negatives	$R = \frac{TP}{TP + FN}$									
F1 score	Harmonic mean of precision and recall. Range between 0 and 1	$F1 = 2 \cdot \frac{P \times R}{P + R}$									
Average Precision (AP)	Area under the PR curve (Precision against Recall across Confidence threshold)										
Mean AP (mAP)	Average of AP values across all classes										
mAP50-95	mAP at different IoU thresholds ranging from 50% to 95%.										
Confusion matrix	Number of correct and incorrect predictions for each class	<table border="1"> <thead> <tr> <th></th> <th>Predicted positive</th> <th>Predicted negative</th> </tr> </thead> <tbody> <tr> <th>Actual positive</th> <td>True Positive (TP)</td> <td>False Negative (FN)</td> </tr> <tr> <th>Actual negative</th> <td>False Positive (FP)</td> <td>True Negative (TN)</td> </tr> </tbody> </table>		Predicted positive	Predicted negative	Actual positive	True Positive (TP)	False Negative (FN)	Actual negative	False Positive (FP)	True Negative (TN)
			Predicted positive	Predicted negative							
		Actual positive	True Positive (TP)	False Negative (FN)							
Actual negative	False Positive (FP)	True Negative (TN)									

The foundation of detection evaluation rests on two critical parameters. The **Confidence Threshold** defines the conditional probability required for positive detection, ranging from 0

to 1. This threshold presents a crucial trade-off: lower thresholds increase detection quantity but may reduce precision, while higher thresholds provide fewer but more reliable detections. Complementing this, the **Intersection over Union (IoU)** quantifies detection accuracy by measuring the overlap between predicted and ground truth bounding boxes, with values ranging from 0 (no overlap) to 1 (complete overlap).

Building upon these fundamental parameters, three core metrics provide deeper insight into model performance. **Precision** measures the proportion of correct positive predictions (TP) among all positive predictions (TP + FP), indicating the model's ability to avoid false positives. This is complemented by **Recall**, which evaluates the proportion of correctly identified positive instances (TP) among all actual positive cases (TP + FN), reflecting the model's ability to find relevant objects. The **F1 Score** provides a balanced metric by calculating the harmonic mean of precision and recall, offering a single measure of model effectiveness through the formula: $F1 = 2 \times (P \times R) / (P + R)$.

For more sophisticated analysis, several advanced metrics provide deeper insight into model performance. **Average Precision (AP)**, calculated as the area under the precision-recall curve, provides a measure of detection performance across varying confidence thresholds for each class. **Mean Average Precision (mAP)** extends this by averaging across all object classes, offering a comprehensive performance metric. The **mAP50-95** further evaluates detection accuracy across multiple IoU thresholds (50% to 95%), providing insight into the model's performance under varying degrees of localization precision.

The **Confusion Matrix** serves as a crucial diagnostic tool, providing a detailed breakdown of model predictions categorized into True Positives (correct positive predictions), False Positives (incorrect positive predictions), True Negatives (correct negative predictions), and False Negatives (missed positive cases). This matrix enables detailed analysis of error patterns and model behaviour across different detection scenarios.

Together, these metrics form a comprehensive evaluation framework that enables detailed assessment of a model's performance, from basic detection capabilities to nuanced understanding of error patterns. These metrics will also guide the evaluation approach in both case studies, through their application and interpretation is varied based on the specific requirements of each scenario. The following sections detail we this methodological framework was applied in our investigation of dairy cow drinking behaviour and backyard chicken nesting patterns.

4 Case studies

The research materials from both case studies, including developed source code, training datasets with annotations, and processed behavioural data, will be published under the Creative Commons 4.0 Attribution license. This commitment to open data aims to support reproducibility and future research in agricultural computer vision applications. The shared resources encompass:

- Source code for model implementation and deployment
- Annotated training datasets for both dairy cow and poultry monitoring
- Data processing and analysis scripts

4.1 Case study one: Cow drinking behaviour

The first case study focuses on a dairy cow campus cowshed, with the primary objective of determining whether the two installed water troughs are sufficient for the number of cows attempting to drink from them concurrently. Building upon previous research conducted by Koskela et al. (2022), which employed computer vision to identify cow behaviour using a Convolutional Neural Network (CNN) with a VGG network architecture, this investigation utilizes a similar approach by using CV but with a different methodology. The data for this case study is derived from a two-month continuous image acquisition of cows in front of an Automatic Milking Station (AMS) inside a teaching campus barn, using a commercially available security camera. The aim is to extract information from the footage data, consisting of 253 videos averaging 4 minutes each, and address the main question concerning the adequacy of the water troughs for the cow population in the campus cowshed.

Water intake is essential for an animal welfare and its consumption is directly linked to growth, feed intake, milk production, and general performance (National Academies of Sciences, 2021) making the monitoring of traffic around water troughs a critical aspect of resource optimization. By monitoring traffic around water troughs, we can optimize this resource, and if necessary, add additional troughs based on detected traffic patterns.

While the primary focus of this case study is on resource optimization, the development of CV techniques within the campus cowshed can also open doors to further technological advancements. These advancements may include individual object recognition, the

creation of detailed animal profiles, linking water, eating, and milking habits, observation of individual and herd behavioural patterns through the strategic installation of additional cameras in key positions, with the ultimate goal being the creation of an intelligent system in which CV is integrated with additional AI technologies to facilitate the communication of detailed animal welfare and productivity information across the multiple actors (public, students, teachers, management) found within the campus' ecosystem.

4.1.1 Setup

After evaluating various platform options, Google Colab emerged as the primary development environment, chosen for three key advantages that aligned with the research needs. First, its integrated GPU support provided the necessary computational power for training complex CV models. Second, the pre-configured deep learning libraries significantly reduced setup overhead, allowing quicker development without dependency hassles. Third, its seamless integration with cloud storage solutions facilitated efficient data management throughout the development process. The platform's accessibility also supported collaborative development.

The implementation architecture was constructed as a layered system of complementary technologies, each serving specific roles in the computer vision pipeline:

The foundation of the technical framework rested on three primary elements:

- Python 3.10 in Google Colab environment served as the base development platform
- Ultralytics YOLOv8x model architecture provided our object detection capabilities
- deepSORT implementation handled object tracking requirements

The supporting framework integrated several specialized libraries, each chosen for specific functionality requirements:

- OpenCV (cv2) managed video processing and frame manipulation tasks
- Pandas facilitated structured data management and analysis
- Matplotlib enabled result plotting and visual analysis
- NumPy handled array operations
- Datetime and time modules managed temporal data processing

- OS and shutil provided system-level file management capabilities
- Pickle enabled object serialization

The computational infrastructure used:

- NVidia T4 GPU with 16 GB VRAM for model training
- Local machine for preprocessing tasks

4.1.2 Data acquisition

The data acquisition phase built upon previous research conducted by Koskela et al. (2022), utilizing their existing footage as foundation. The approach focused on preparing this data for effective YOLO model training, requiring consideration of image preparation and annotation.

The preparation of training data followed YOLO's architectural requirements, needing a structured approach to image annotation and directory organization. This implementation required annotated images paired with corresponding label files containing object class numbers and bounding box coordinates, serving as the ground truth for model training and validation. These files were organized into "train" and "val" directories, each containing dedicated "images" and "labels" subdirectories.

4.1.3 Data preparation and labelling

The dataset generation process balanced sufficient positional variance with data quality considerations. Two python scripts were developed for frame extraction at intervals of 1250 and 3000 frames, yielding an initial set from the 253 videos. Quality control measures included automated filtered of frames with less than 10% overall brightness, ensuring the exclusion of extremely low-visibility conditions that would not benefit the model training.

The filtered dataset comprised 1140 unique images providing sufficient temporal coverage while maintaining manageable dataset size. This sampling strategy ensured representation across different animal positions and lighting conditions, while the quality threshold helped maintain dataset consistency.

Class selection focused on essential components for drinking behaviour detection:

- “head” (0): Primary indicator of drinking activity
- “body” (1): Supporting feature for animal presence
- “trough” (2): Key reference point for interaction detection

While initially including three classes, subsequent analysis revealed that the relationship between head and trough positions provided the most reliable drinking behaviour indicators. The system determined drinking activity by measuring both the spatial proximity of the head to the water trough and the duration of this proximity. Although the trough was detected dynamically by YOLO rather than using predefined coordinates, its stationary nature in the videos made it an effective reference point for establishing drinking behaviour thresholds. This observation led to the focused use of head and trough classes for the final implementation.

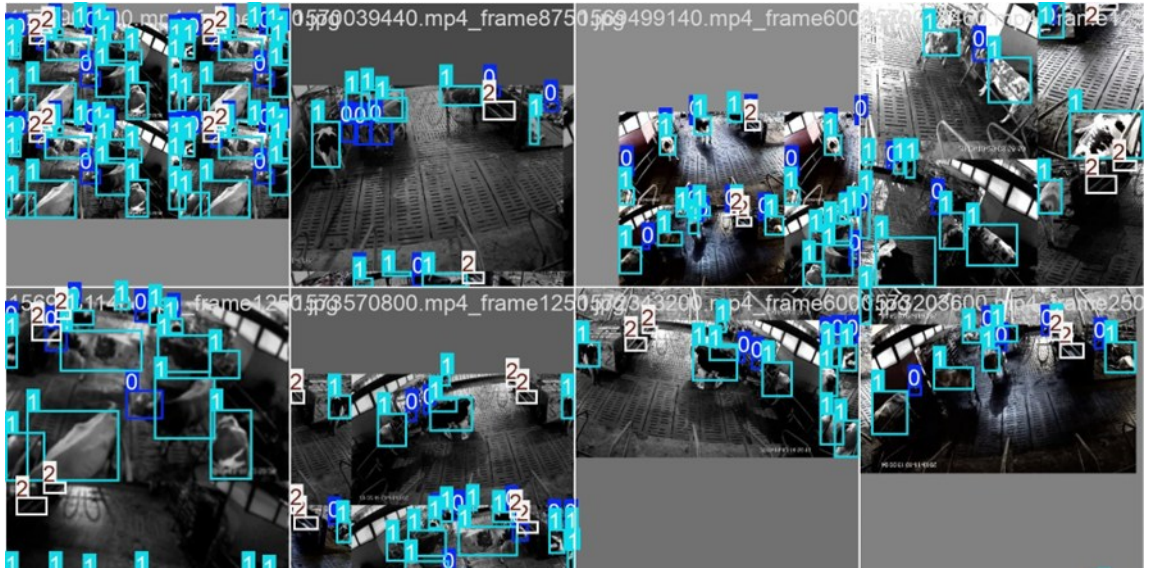
Dataset organization followed standard ML practices in terms of training and validation ratio using the most common mentioned by Li G. et al. (2021), with an 80:20 split between training (1140 images) and validation (270 images) sets. Image annotation utilized makesense.ai platform (Skalski, 2019), generating 1340 YOLO-compatible label files containing class numbers and bounding box coordinates.

4.1.4 Training

The model architecture selection process considered the balance between detection accuracy and computational requirements across YOLO’s available versions (nano to extra-large). After evaluation, YOLOv8x was selected for its enhanced feature extraction capabilities. While this choice introduced higher computational demands, the focus on offline analysis prioritized accuracy over real-time performance constraints, making this trade-off acceptable for our research context.

To address the challenges presented by limited dataset sizes, transfer learning was employed using a YOLOv8x model pretrained on the COCO dataset (Lin et al., 2015). This approach utilized the model's existing feature extraction capabilities trained on the public dataset, enhancing the model’s overall generalization performance. The training configuration comprised 120 epochs with 5 warm-up epochs for stable initialization, utilizing default mosaic (Figure 2) for further enhancing generalization by increasing the object instance count. This implementation was executed on Google Colab with T4 GPU.

Figure 2. Mosaic data augmentation example from initial training batch

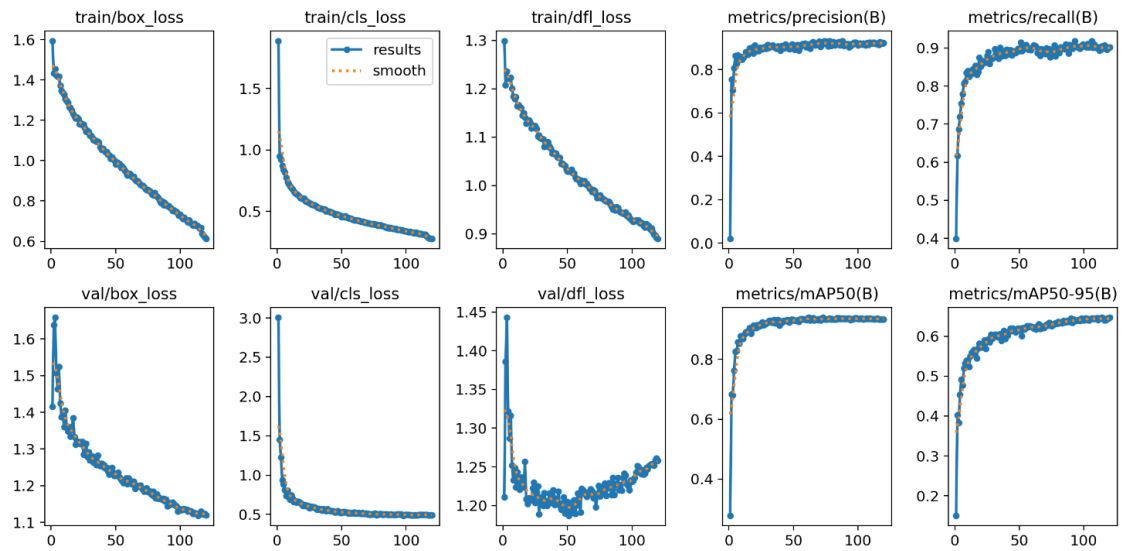


The model achieved several key performance metrics that demonstrated its effectiveness (Table 2 and Figure 3). Training metrics showed a box loss of 0.612, classification loss of 0.277, and Distribution Focal Loss (DFL) of 0.889, where DFL helps improve the model's ability to precisely locate object boundaries. In validation, the model recorded metrics of 1.119 for box loss, 0.488 for classification loss, and 1.257 for DFL loss. The model demonstrated strong detection capabilities with metrics/precision(B) at 0.923, indicating 92.4% of positive predictions were correct, and metrics/recall(B) at 0.901, showing 90.2% of actual positive instances were correctly identified. The metrics/mAP50(B) reached 0.933, demonstrating 93.4% accuracy in predicted bounding boxes at 50% Intersection over Union (IoU) threshold, while metrics/mAP50-95(B) achieved 0.646, indicating 64.7% accuracy across varying IoU thresholds from 50% to 95%.

Table 2. Performance metrics of the custom YOLOv8x model implementation

train/box_loss	train/cls_loss	train/df_loss	metrics/precision(B)	metrics/recall(B)
0.612	0.277	0.889	0.923	0.901
val/box_loss	val/cls_loss	val/df_loss	metrics/mAP50(B)	metrics/mAP50-95(B)
1.119	0.488	1.257	0.933	0.646

Figure 3. Training performance metrics for custom YOLOv8x implementation showing performance progression across training epochs



The precision-confidence analysis (Figure 4) revealed optimal performance at 0.957 confidence threshold, achieving near-perfect precision. The recall-confidence curve demonstrated robust detection capabilities, showing 95% recall at minimum confidence. The PR curve yielded class-specific performance values: head (0.858), body (0.957), and trough (0.992), with an all-class $mAP@0.5$ of 0.936, while the F1 curve revealed a value of 0.91 at a rather low 0.355, affected by the lower scores of head class, and suggesting potential false positive detections (Figure 5).

Figure 4. Precision-confidence and recall-confidence curves demonstrating model detection performance

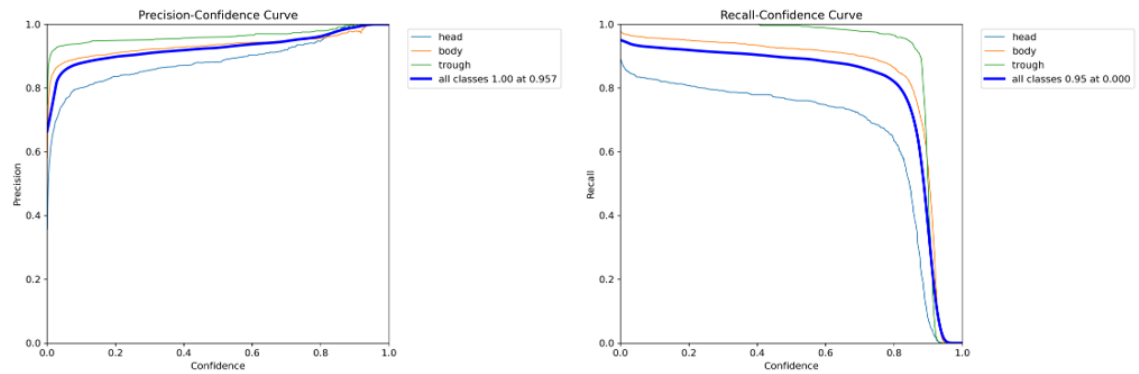
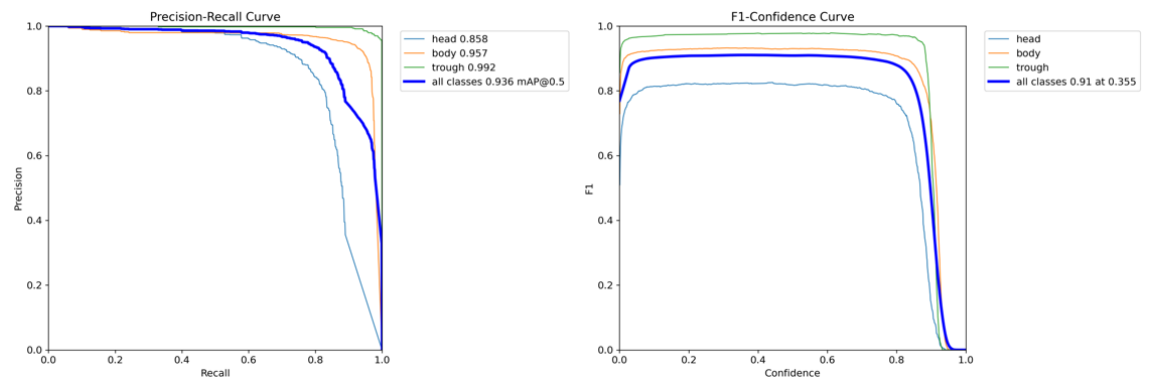
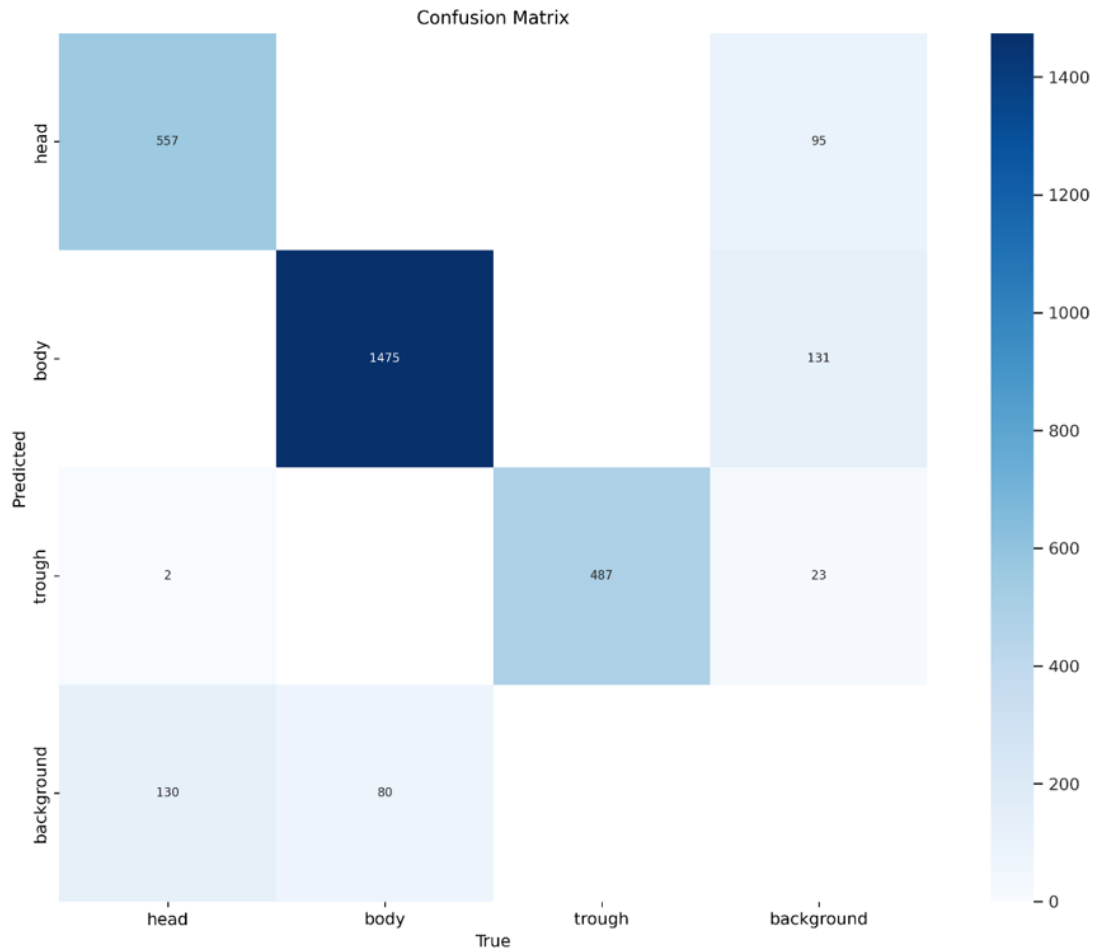


Figure 5. Precision-recall curve and F1-score analysis for YOLOv8x model



The confusion matrix analysis provided detailed insight into model performance (Figure 6), revealing true positives of 557 heads, 1475 bodies, and 487 troughs; false positives of 130 heads, 80 bodies, and 0 troughs; and false negatives of 95 heads, 131 bodies, and 23 troughs. This performance profile indicated particularly strong detection capabilities for stationary objects (troughs) and larger features (bodies), with slightly reduced accuracy for more dynamic elements (heads), aligning with expected challenges.

Figure 6. Confusion matrix analysis of YOLOv8x model performance showing true positives, false positives, and false negatives across detection classes



4.1.5 Deployment

The deployment phase focused primarily on addressing the challenge of occlusion, a critical consideration in agricultural computer vision applications. To enhance tracking robustness, deepSORT methodology, established by Wojke et al. (2017) was integrated with the YOLOv8x model, utilizing its object persistence mechanisms including Kalman filtering, Hungarian algorithm, and Siamese networks.

The drinking behaviour detection system was implemented through a dual-threshold mechanism that combined spatial and temporal analysis. The system utilized Euclidean distance between detected head and trough centroids for spatial assessment, complemented by temporal persistence monitoring of proximity conditions. This approach enabled robust detection while minimizing false positives.

The interaction detection logic followed a systematic process beginning with initial detection through the YOLO model for heads and troughs identification. This was followed by deepSORT's tracking mechanism, which assigns and maintains unique identification numbers (IDs) for each detected object through a combination of Siamese neural networks, Hungarian algorithm, and Kalman filtering. The system then performed proximity analysis through centroid-based measurement and validated drinking events through temporal persistence.

Through iterative testing and optimization, several key system parameters were established. A critical parameter was the tracker retention period (tracker age), set to 6 seconds (150 frames), during which deepSORT maintained object IDs even when YOLO temporarily failed to detect the object. This retention mechanism proved essential for maintaining continuous tracking through brief occlusions, as the system could successfully re-identify and associate the same ID with an object when it reappeared. Additionally, a duration threshold of 8 seconds (200 frames) was implemented to confirm genuine drinking events, while confidence thresholds were maintained between 60-65% to balance detection sensitivity with accuracy. These parameters demonstrated effective performance in handling temporary occlusions while maintaining reliable object tracking and event detection, as evidenced in Figures 7 and 8.

The establishment of tracker threshold parameters required consideration of the sample characteristics. With video samples averaging 4 minutes in duration, the detection system needed to account for the temporal constraints of the dataset relative to typical drinking behaviour patterns. Drinking events in dairy cattle can span several minutes, creating three possible temporal scenarios in our samples: (1) ongoing events captured mid-process, (2) complete events contained within the sample, and (3) events beginning but not concluding within the sample period. This temporal fragmentation of behavioural events necessitated calibration of the tracking parameters to maintain detection consistency across these scenarios.

Figure 7. Detection of single cow drinking behaviour

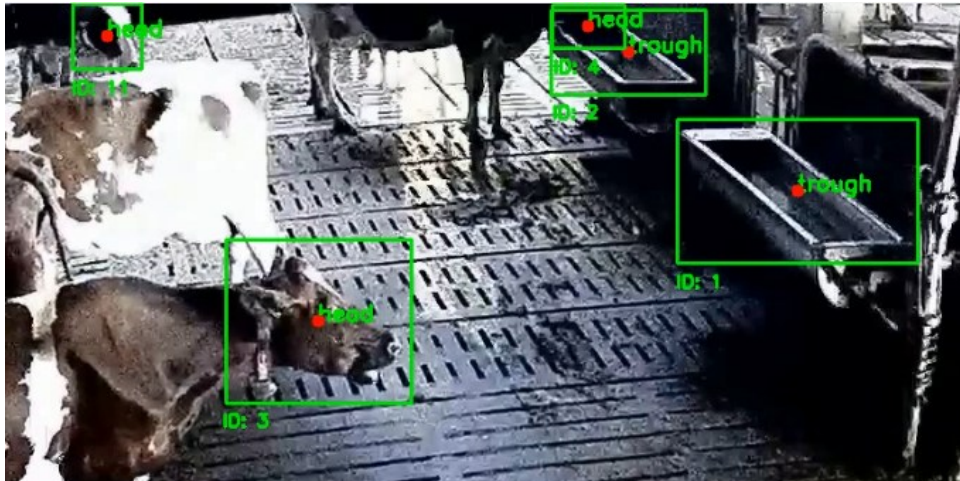
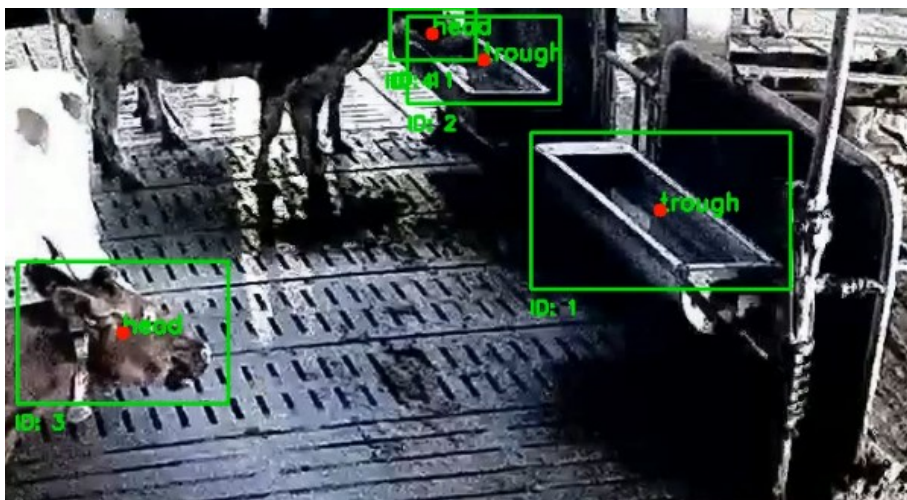


Figure 8. Concurrent detection of multiple cow drinking behaviour



For monitoring multi-animal interactions, the system implemented tracking of multiple head IDs interacting with trough IDs simultaneously. The system classified interaction types based on concurrent cow count into distinct categories: single cow drinking (1_pair), two cows drinking simultaneously (2_pair), and three cows drinking simultaneously (3_pair). This classification enabled detailed analysis of resource utilization patterns.

Data collection encompassed tracking information including unique track IDs for each class, interaction type classification, and both single and concurrent interaction durations. The system-maintained frame-level temporal positional data, enabling detailed analysis of behavioural patterns. All interaction information was systematically stored in pandas dataframes (Figure 9), facilitating subsequent analysis and pattern recognition.

Figure 9. Data structure implementation for behavioural interaction storage

Trough_ID	Head_ID	Interaction_Type	Interaction_Duration
0	2	4	2_pair_interaction
1	2	11	2_pair_interaction

concurrent_interaction_frames	
0	390
1	390

This deployment architecture proved effective in managing some of the complexities in agricultural environments while maintaining tracking and classification capabilities. The system's ability to handle multiple concurrent interactions while maintaining individual identity tracking demonstrated its potential for practical application in livestock monitoring scenarios.

4.1.6 Results

The integration of YOLOv8 and deepSORT demonstrated effective detection and tracking capabilities in the campus cowshed environment. The analysis revealed insights across three key dimensions: its **technical performance**, **behavioural analysis**, and **resource utilization** implications. The analysis of results across both case studies follows a structured framework examining these three complementary dimensions. We begin with technical performance assessment, evaluating the model's accuracy and adaptation to different environments. The technical foundation then supports the behavioural analysis, where activity patterns and species-specific characteristics are explored. Finally, resource optimization is examined, analysing usage patterns and identifying improvement opportunities.

Technical performance analysis

The implementation achieved robust detection under favourable conditions through carefully optimized parameter configuration. The system's effectiveness relied on key parameters including a 6-second tracker age (150 frames), 8-second duration threshold (200 frames), and confidence range of 60-65%. These settings enabled deepSORT to maintain object tracking during temporary YOLO detection failures.

The system demonstrated particular strength in single animal tracking, multi-animal interaction detection, and concurrent drinking event classification. However, several

technical challenges emerged during deployment. Prolonged occlusion situations impaired detection accuracy, particularly when animals remained stationary in occluding positions. Identity management presented challenges as well, as head movements under limited visibility conditions occasionally triggered multiple ID assignments (Figure 10).

Figure 10. Analysis of tracking inconsistencies in unique head identification

```

Summary of Interaction Data:
  Interaction_Type  Total_Interaction_Duration  \
0  No_interaction      1367079
1  1_pair_interaction    203380
2  2_pair_interaction    36662
3  3_pair_interaction    42380
4  4_pair_interaction    21720
5  5_pair_interaction     3756
6  10_pair_interaction    8521

  Total_Interaction_Duration_Minutes
0      911.386000
1     135.586667
2      24.441333
3      28.253333
4      14.480000
5       2.504000
6       5.680667

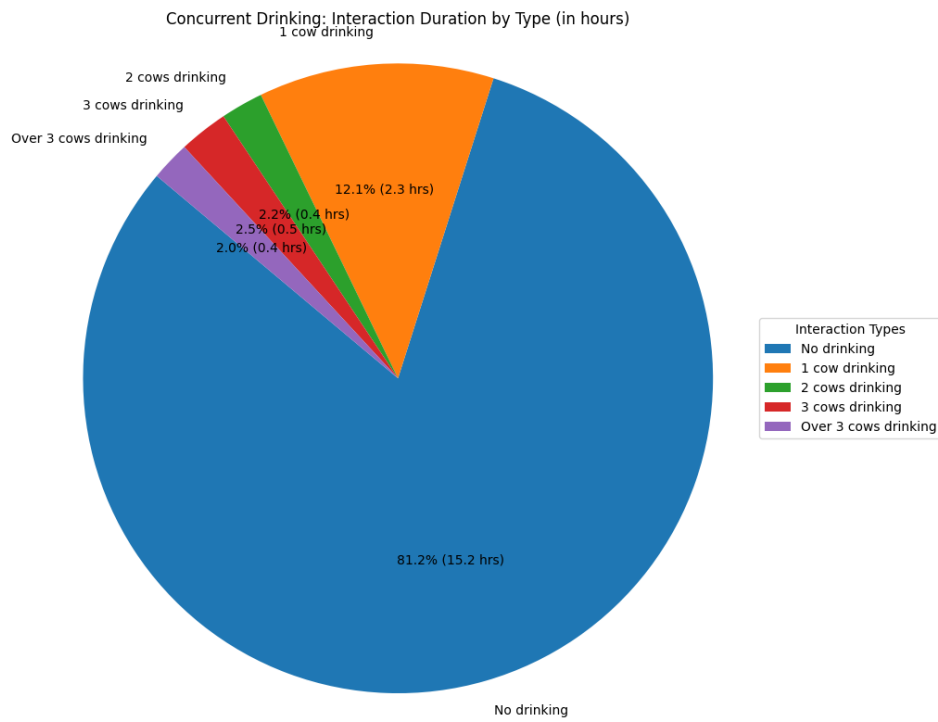
```

Behavioural analysis

The temporal analysis of drinking behaviour revealed several significant patterns in resource utilization. Overall interaction distribution showed that 18.8% of total time involved drinking interactions, while 81.2% showed no trough interaction despite animal presence. This pattern suggests efficient resource distribution with minimal competition for access.

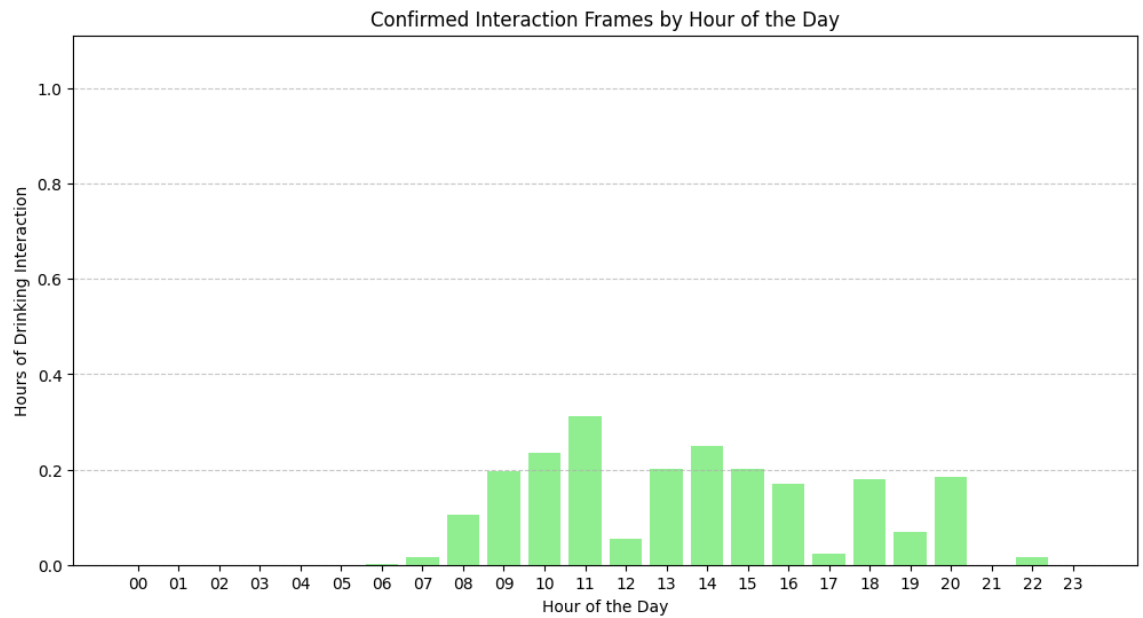
Concurrent usage patterns provided deeper insight into resource utilization dynamics. Single-cow drinking occupied 12.1% of interaction time, while two-cow drinking accounted for 2.2%, and three-cow drinking represented 2.5%. Interactions involving more than three cows comprised 2.0% of observations (Figure 13). The total multi-animal interactions (6.7%) suggest efficient resource distribution, as no significant congestion or competition for resources was observed.

Figure 11. Distribution analysis of concurrent drinking behaviours categorized by interaction type, showing percentage breakdown of single, dual, and multi-animal events



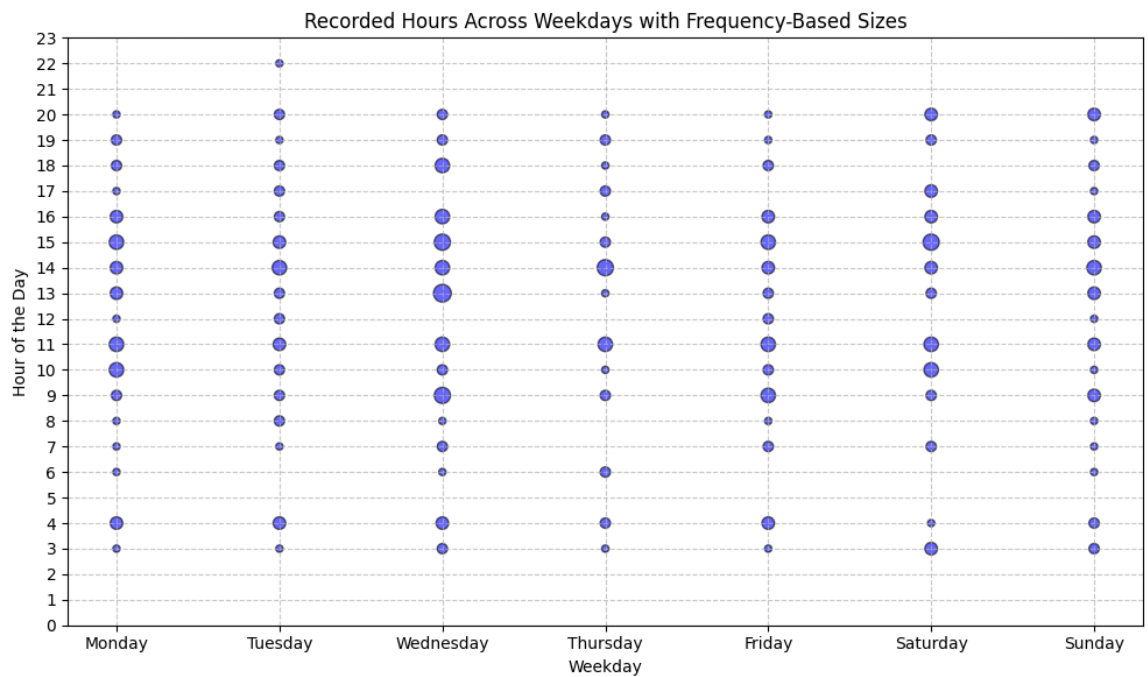
Temporal patterns revealed peak activity during the 10th-14th hour, aligning with established behaviour patterns in dairy cattle (Figure 12).

Figure 12. Temporal distribution of validated drinking interactions



The notable variations observed in certain time slots, particularly during the 12th and 17th hours, as revealed through dataset coverage analysis (Figure 13), are due to slight variations in recording hours across weekdays.

Figure 13. Weekly temporal distribution of monitoring data



The interpretation of these behavioural patterns must acknowledge the inherent temporal and spatial complexity of animal behaviour. Individual variation in drinking patterns reflects both intrinsic behavioural traits and extrinsic factors including feed characteristics (timing, type, quality), grazing conditions, and environmental parameters (temperature, humidity). This complexity suggests that the observed patterns represent a snapshot of dynamic behavioural adaptations rather than fixed behavioural constants.

Resource Optimization analysis

The analysis provides several key insights for resource management.

Utilization efficiency metrics paint a picture of effective resource distribution. The 18.8% interaction rate indicates efficient resource use, while low congestion rates suggest appropriate resource availability. The temporal distribution of usage shows consistent resource accessibility throughout operational hours. Further supporting these efficiency indicators, the analysis revealed that only 6.7% of all drinking interactions involved multiple animals simultaneously.

While the technical implementation faced certain challenges, particularly in continuous multiple tracking (evidenced by the 12.34% frame inconsistency and multiple unique head assignments), the core behavioural patterns remain identifiable and consistent across the observation period. The findings suggest that while current water trough configuration serves the herd's needs, the identified temporal patterns could inform future optimization strategies, particularly in areas such as maintenance scheduling and potential water trough system upgrades.

4.2 Case study two: Chicken nesting behaviour

Building upon the insights from the dairy cow monitoring implementation, the second case study explores the adaptation of the computer vision framework to a different agricultural context. The experience gained in handling environmental variability and tracking challenges in the first case study informed the approach to the unique demands of poultry monitoring. While the core technological framework remained consistent, its implementation required recalibration to address the distinct characteristics of backyard poultry behaviour. The transition from institutional to individual framing contexts provided an opportunity to test both the scalability and the adaptability of the CV approach.

The second case study examines CV implementation in a backyard chicken coop located in the suburban area of Thessaloniki, Greece, focusing on nest box utilization and egg production monitoring. While substantial research exists on CV applications in commercial poultry operations, small-scale and backyard settings present unique challenges that remain largely unexplored in published literature. Building upon the methodological framework established in the previous case study, this investigation adapts the YOLO architecture to monitor nest occupancy patterns and automate egg counting in a more dynamic, small-scale environment.

The study utilizes data from a two-phase collection strategy spanning 40 days, comprising motion-triggered recordings for model training and subsequent continuous monitoring during peak activity hours. The primary objectives are to assess the adequacy of eight nest boxes arranged in two rows of four (Figure 14) and optimize egg collection timing to minimize breakage risks. Through monitoring nest occupancy and egg production patterns, there is potential to optimize resource allocation while developing insights into bird behaviour.

Figure 14. Arrangement of nest box monitoring setup



While the immediate focus is on resource optimization, the development of CV techniques in this setting can enable deeper integration, including individual bird recognition,

automated health monitoring through physical condition assessment, and potential integration with acoustic monitoring. The goal extends beyond immediate resource management to creating an intelligent monitoring system capable of providing comprehensive individual and flock welfare insights in small-scale operations.

4.2.1 Setup

The implementation environment evolved from the previous case study, transitioning from cloud-based to local computation to accommodate more sophisticated processing requirements.

Core components:

- Python 3.10 in Spyder (Anaconda environment)
- Ultralytics YOLOv8x model architecture
- State machine implementation for occupancy tracking

Hardware configuration:

- Local machine with RTX 4060 Ti GPU (16GB VRAM) for preprocessing and deployment
- Google Colab A100 GPU (40GB VRAM) for model training

Primary libraries:

- Core processing:
- OpenCV (cv2) for video manipulation
- NumPy for array operations
- Pandas for data structuring
- Matplotlib and Seaborn for visualization
- Ultralytics YOLO for object detection
- Support libraries
- DateTime and Time for temporal processing
- Statistics and Math for numerical operations
- OS and Path for file management
- Collections and ABC for advanced data structures

- Typing for type annotations
- Dataclasses for structured data handling

4.2.2 Data acquisition

The data collection strategy employed two distinct phases to balance training data quality and diversity with comprehensive behavioural monitoring.

The initial phase, conducted from June 21 to July 10, focused on motion-triggered recording using the camera's built-in detection capabilities. The system was configured with an 80% sensitivity threshold specifically optimized for nest box regions of interest. This phase captured variable footage spanning 05:00 to 21:00, recorded in h.265 format at 2560x1440 resolution and 15 fps. From this collection, 120 videos of varying lengths through the entire span were selected for training dataset creation.

This phase generated 15-16 files daily, each approximately 250 MB in size and 40-50 minutes in duration, totalling approximately 3.6GB daily. This extended monitoring approach offered an opportunity to explore whether continuous observation might reveal additional behavioural patterns while remaining within practical storage constraints.

The equipment configuration utilized on a commercial wireless outdoor dome camera (Hikvision DS-2CV2141G2-IDW), selected for its technical specifications and environmental durability. The camera's 95° horizontal FOV, 53° vertical FOV, and 111° diagonal FOV provided comprehensive coverage of the monitoring area. Its IP66 protection ensured reliable operation under varying environmental conditions. The installation position was carefully optimized to maximize visibility of both upper and lower nest rows while working within physical constraints of the environment.

4.2.3 Data preparation and labelling

The dataset development process emphasized in capturing diverse behaviour patterns while maintaining data size manageable. The initial dataset comprised 1022 training images extracted from motion-detected footage, supplemented with 100 background images to reduce false positive detections on background elements. Subsequent performance analysis indicated the need for additional training data, leading to the

incorporation of 250 supplementary images from peak activity periods, resulting in a final dataset of 1272 images.(1184!)

Frame selection prioritized video samples over 25MB across daily samples, to ensure varied coverage and diverse ambient conditions. Unlike cases study 1, where low-visibility frames required filtering, image quality remained consistent throughout the capture, with manual filtering only required to remove frames containing human presence during egg collection. Three distinct classes were established:

- “Guinea fowl” (0): Despite adding complexity to an already limited dataset, their inclusion was necessary as they actively contributed to nest box utilization patterns
- “Chicken” (1): Primary focus for nest box occupancy analysis
- “Egg” (2): Required particular attention during annotation for precise boundary definition due to smaller size and potential occlusion

The annotation process utilized makesense.ai (Figure 15) for creating YOLO-compatible files containing class numbers and bounding box coordinates. The directory structure followed YOLO requirements.

Given the limited dataset size, a less used 90/10 train/validation split was implemented instead of traditional 80/20 ratio. This decision prioritized maximizing training data utilization while maintaining validation integrity, following established guidelines mentioned in the work of Li et al. (2021).

Figure 15. Demonstration of object annotation methodology for poultry monitoring



4.2.4 Training

Following the previously established approach, a YOLOv8x model was again selected for object detection, implementing several refinements based on our previous experience.

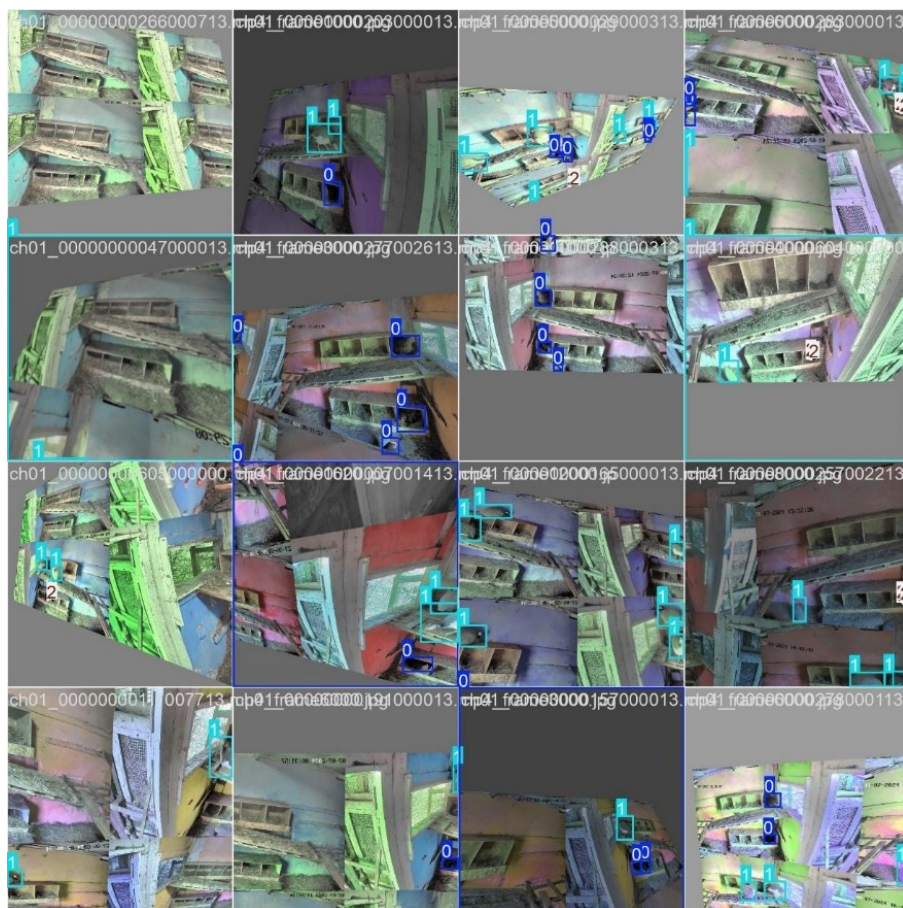
The base configuration established a framework of 100 epochs using SGD optimizer, incorporating 5 warm-up epochs to ensure stable initialization. A significant adaptation from the previous implementation was the transition to near 2K resolution training images (1216 horizontal pixels). This modification proved transformative - the initial implementation using standard 640p resolution achieved mAP50-95 scores between 0.62-0.64, but the shift to higher resolution yielded substantial improvements, with mAP50-95 increasing to 0.711.

The impact on egg detection performance was particularly noteworthy. Under standard resolution, system observed 119 True Positives, 62 False Negatives, and 11 False Positives. The transition to high resolution significantly improved these metrics, yielding 154 True Positives, 27 False Negatives, and 7 False Positives. This improvement demonstrated the critical importance of resolution in detecting smaller objects.

Through iterative testing, several key parameters were refined to optimize model performance through the customizable training parameters. Colour space adjustments

were implemented following proven approaches in agricultural monitoring mentioned by Okinda et al. (2020), while geometric transformations were calibrated based on established computer vision practices for spatial variation handling. These included Hue at 0.5, Saturation at 0.3, and Value at 0.5. Geometric transformations were carefully calibrated to enhance the model's robustness to object variation in spatial environments, implementing scale augmentation (0.3) for accommodating size variations, shear transformation (0.15) for enhancing geometric diversity, and perspective adjustment (0.001) for improving depth perception (Figure 16).

Figure 16. Implementation of data augmentation techniques for model training



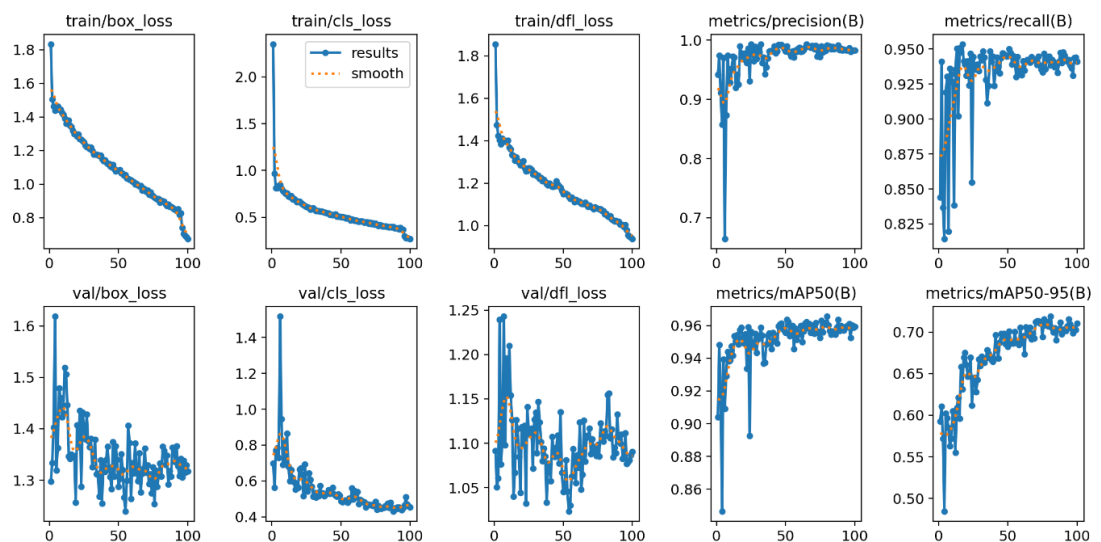
The model achieved optimal performance at epoch 80, demonstrating strong metrics across multiple evaluation criteria. Training metrics revealed box loss at 0.891, classification loss at 0.405, and DFL loss at 1.053. Validation metrics showed slightly elevated values with box loss at 1.332, classification loss at 0.450, and DFL loss at 1.113, indicating expected performance differences between training and validation sets. Detection performance proved particularly robust, with precision reaching 99.1% of positive predictions and recall achieving 94.4% of actual positive instance identification. The model

demonstrated strong bounding box accuracy with mAP50 at 96.5% and mAP50-95 at 71.3%, indicating consistent performance across varying IoU thresholds (Table 3 and Figure 17).

Table 3. Performance metrics of the YOLOv8x model for poultry monitoring

train/box_loss	train/cls_loss	train/df_l_loss	metrics/precision(B)	metrics/recall(B)
0.891	0.405	1.053	0.991	0.944
val/box_loss	val/cls_loss	val/df_l_loss	metrics/mAP50(B)	metrics/mAP50-95(B)
1.332	0.450	1.113	0.965	0.713

Figure 17. Training performance metrics for poultry monitoring model training across epochs



The precision-confidence analysis demonstrated optimal performance at 0.851 confidence threshold, achieving near-perfect precision. The recall-confidence curve showed robust detection capabilities with 95% recall at minimum confidence. (Figure 18) The PR curve yielded an impressive all-classes mAP@0.5 of 0.960, with the F1 Curve peaking at 0.96 with a confidence threshold of 0.658. This high F1 score at a relatively high confidence threshold suggested robust detection capabilities with some false positives. (Figure 19).

Figure 18. Precision-confidence and recall-confidence analysis for poultry detection model

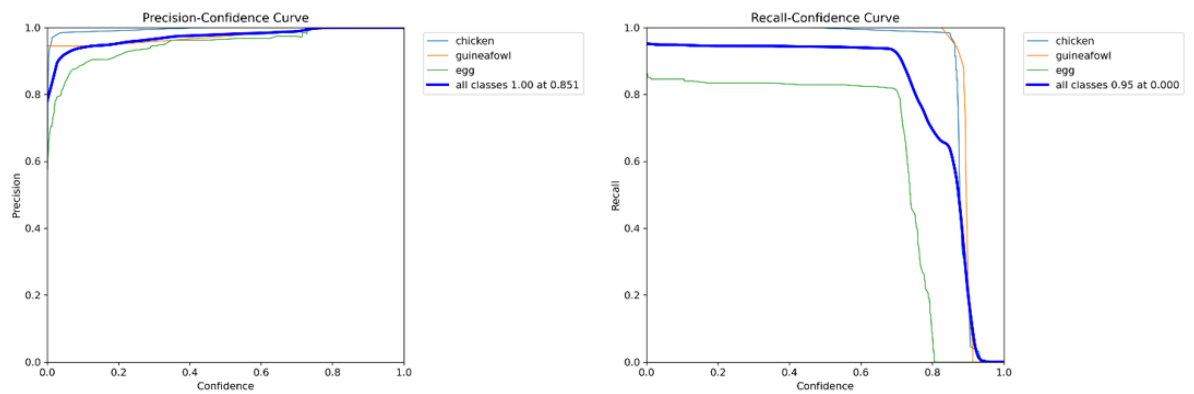
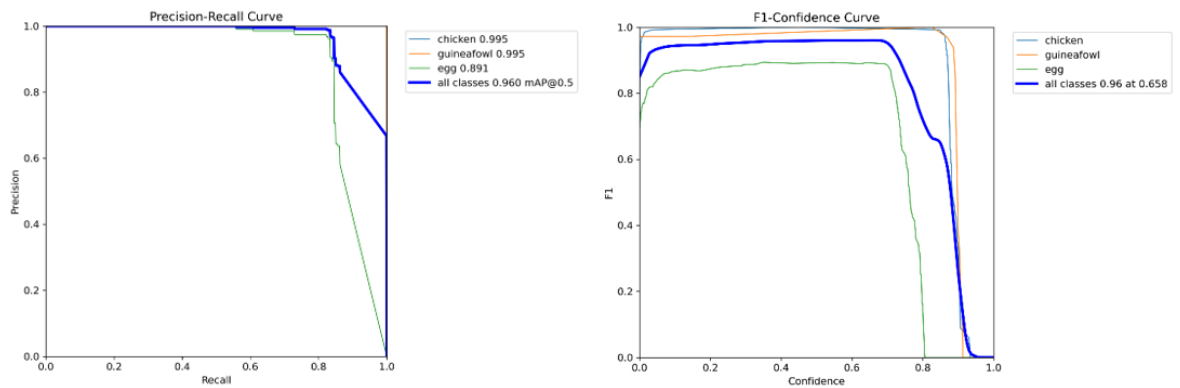


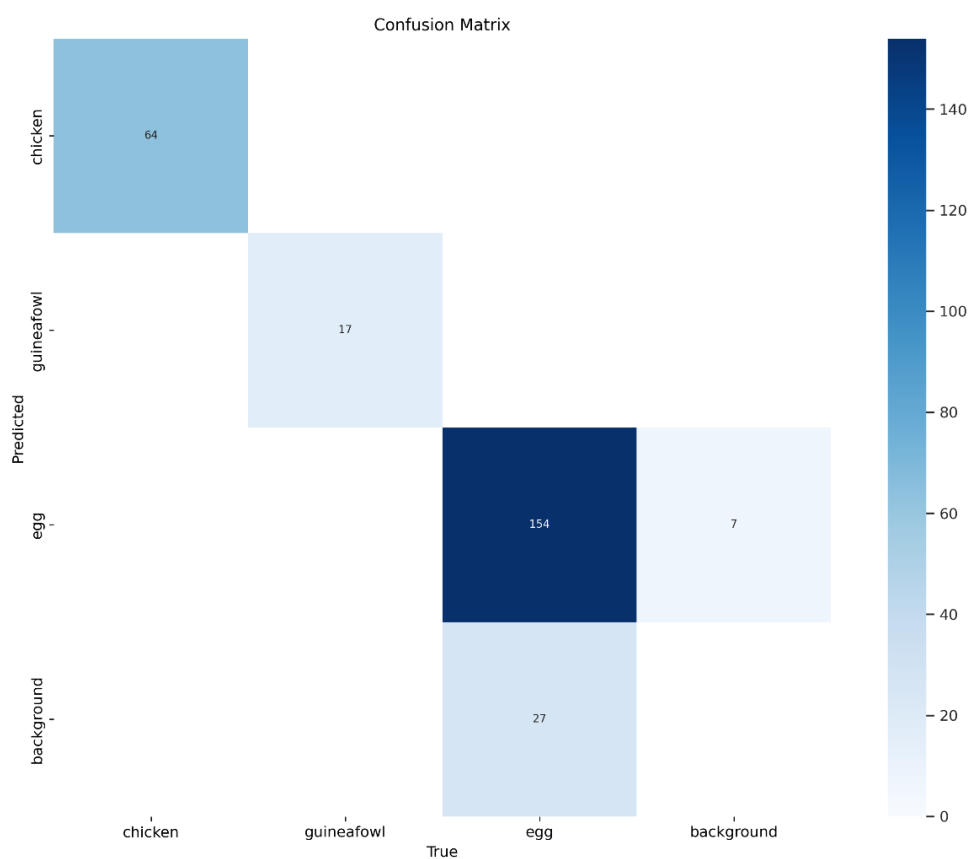
Figure 19. Precision-recall and F1-confidence analysis for poultry detection



Confusion matrix analysis revealed particularly strong performance in bird classification. The validation set showed perfect classification for both chickens (64 True Positives, 0 False Positives/Negatives) and Guinea fowl (17 True Positives, 0 False Positives/Negatives). Egg detection, while improved by high-resolution implementation, reflected the increased challenge of detecting smaller objects, achieving 154 True Positives, 27 False Negatives, and 7 False Positives

The overall performance metrics demonstrate successful adaptation to the small-scale environment, with particularly strong results in bird classification and improved egg detection through resolution optimization.

Figure 20. Confusion matrix analysis of poultry detection model



4.2.5 Deployment

The deployment phase evolved through several iterative stages, each addressing specific challenges encountered in poultry monitoring. Unlike Case Study One where deepSORT tracking provided object persistence, this implementation approached the unique characteristics of bird behaviour and the complexities of egg detection through another methodology, following a systematic progression through sophisticated detection and validation mechanisms. This multi-tiered temporal validation approach builds upon established methodologies in livestock monitoring (Wurtz et al., 2019), particularly addressing the challenges of accurate detection in dynamic agricultural environments.

The first implementation focused on establishing reliable occupancy detection through precise Intersection over Union (IoU) calculations between detected birds and predefined nest box regions. Through empirical testing, established key IoU thresholds defined the detection boundaries: a minimum IoU of 0.55 ensuring sufficient nest box coverage, and a maximum IoU of 1.6 preventing false positives from birds positioned in front of the nests.

These thresholds effectively delineated two distinct scenarios: valid nest occupancy ($0.55 < \text{IoU} < 1.6$), and bird presence in front of the nest ($\text{IoU} > 1.6$).

The initial IoU detection system revealed the need for more sophisticated behaviour interpretation, leading to the development of a state machine with distinct states. Each state transition represented a specific behavioural phase:

1. Empty to Entering required 5 frames of initial detection within IoU thresholds, providing brief presence confirmation
2. Entering to Occupied needed 15 frames, ensuring extended presence confirmation
3. Occupied to Leaving utilized 10 frames for departure detection and temporary absence handling
4. Leaving to Empty demanded 20 frames for complete departure confirmation and state reset

The state machine implementation revealed detection confidence fluctuations which affected nest occupancy counters, particularly evident during rapid bird movements and model detection misses. A typical 15-frame window might show raw confidence values fluctuating between 0.8 and 0.4, with most values clustering around 0.80 but occasional drops to 0.40 during occlusions or miss-detections. To address this, confidence smoothing mechanism was implemented, inspired by proven approaches in agricultural CV by Okinda et al. (2020). The smoothing mechanism operated through:

1. Maintenance of a rolling window of detection values
2. Calculation of detection consistency ratio
3. Production of smoothed confidence values

For example, with 12 consistent detections in a 15-frame rolling window, the system would calculate a smoothed confidence value of 0.8 (12/15), providing more stable input for state machine logic and enabling graceful occupancy time counting without excessive confidence fluctuations.

Building upon our stable confidence framework, a three-tier temporal window system was implemented for robust egg detection, as eggs posed a different challenge of being easily occluded by material due to their size, and temporarily completely covered during chicken laying occasions, despite being well detected. For this issue, the system analysed

detection patterns across multiple time scales, as demonstrated in a typical scenario between 13:50:00 and 13:56:00:

Short Window (30s): [1, 1, 1, 0, 1] provided quick response results

Medium Window (2m): [1, 1, 1, 1, 1, 0, 1, 1] enabled trend confirmation

Long Window (6m): [1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1] ensured pattern validation

Where 1 represented consistent detection and 0 indicated interrupted detection, each window served specific analytical purposes:

- Short window provided immediate detection status
- Medium window confirmed detection trends
- Long window eliminated false positives through extended validation.

Through the use of the long window, the system maintained a longer “memory” of the positive egg counts and maintained it even over a set period of time, even when the eggs were not detected at the present time.

The combination of these four systems - IoU thresholds, state machine, confidence smoothing, and temporal windows - created a robust monitoring solution capable of handling the complexities of this specific poultry environment. The progressive implementation of each system built upon the strengths of previous components while addressing their limitations, resulting in a comprehensive detection and validation framework.

4.2.6 Results

The analysis of our state machine-based monitoring implementation follows three key analytical layers, each providing distinct insights into the backyard poultry environment. Through examining 11 days of continuous monitoring data, spanning July 21-31, 2024, with daily 12-hour observation windows, the technical performance is evaluated through its adaptation in small-scale settings, the subtle behavioural patterns are explored across the two different poultry species, and resource utilization in the context of nest box optimization is being assessed.

Technical performance analysis

The system demonstrated robust performance under controlled conditions while revealing specific limitations. The state machine reliability proved particularly noteworthy, successfully eliminating false counts caused by passing birds through temporal persistence validation. Through the confidence smoothing it the system was also able to maintain stable detection sequences under non fully occluded conditions. This performance was particularly evident during unobstructed egg (Figure 21) and single-bird nesting events (Figure 22).

Figure 21. Detection performance under optimal visibility conditions

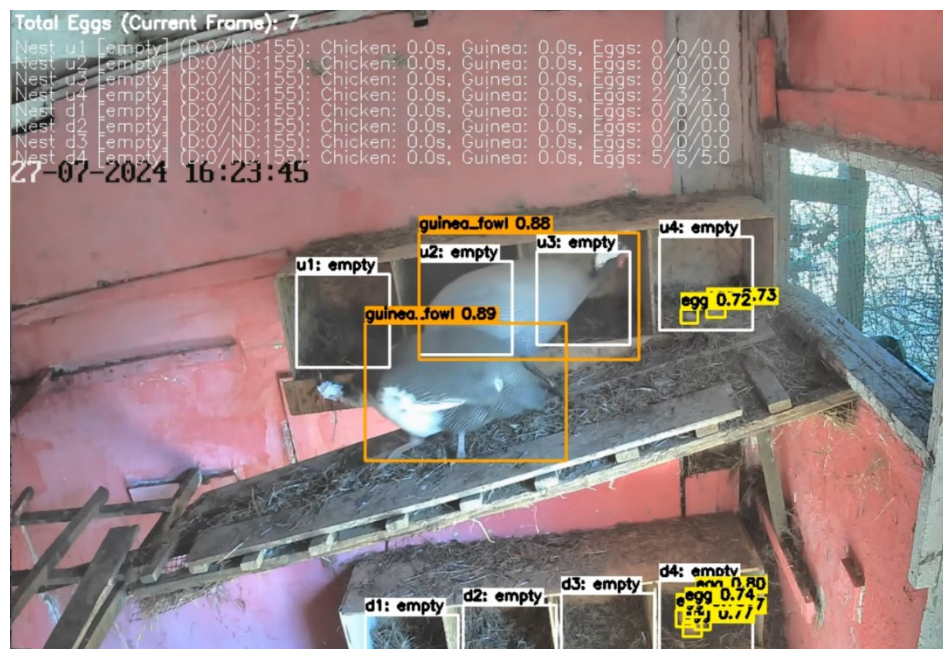
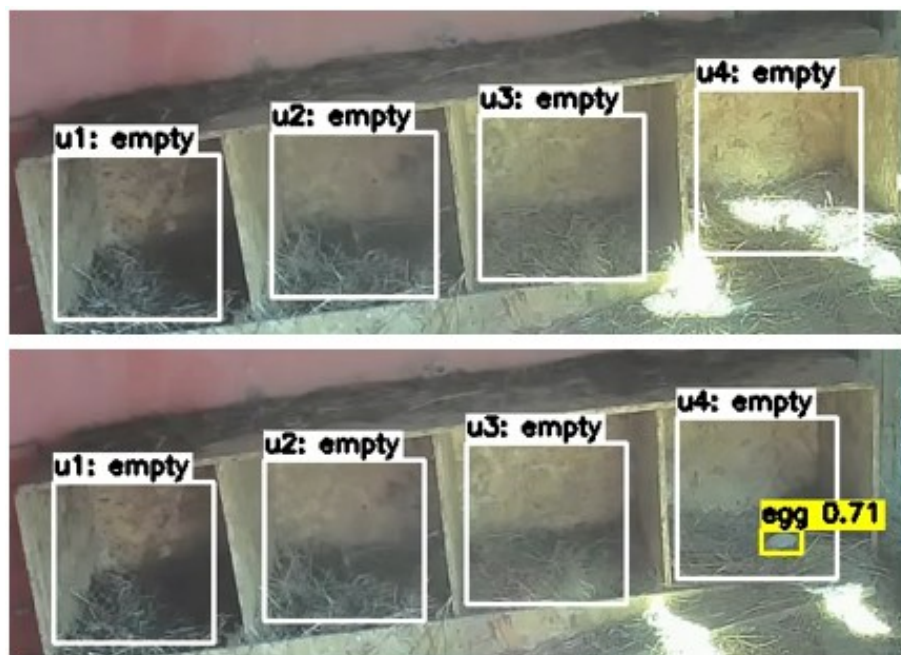


Figure 22. Detection of individual nesting behaviour



Environmental adaptation emerged as a critical factor in system performance. The implementation showed limited effectiveness under direct sunlight conditions due to pixel saturation (255, 255, 255 RGB values), though it maintained accuracy until environmental conditions exceeded feature extraction capabilities.

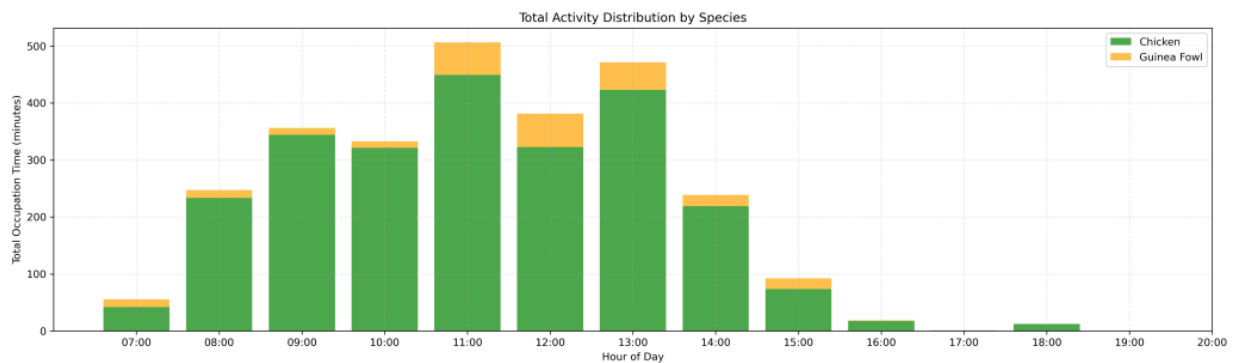
Figure 23. Model performance analysis under challenging lighting conditions



Behavioural analysis

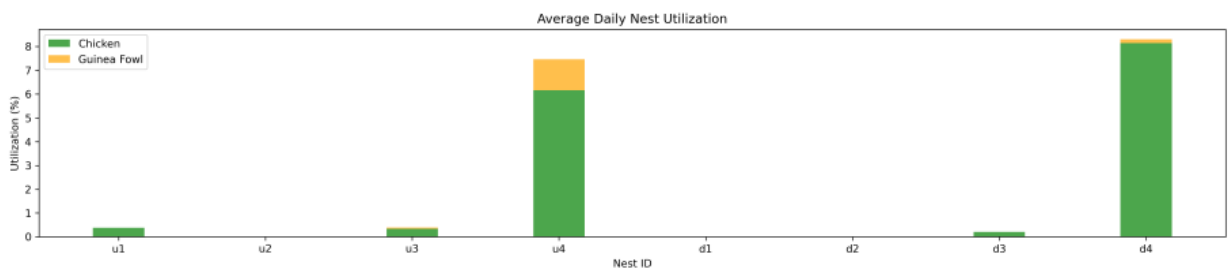
The monitoring system revealed distinct temporal and spatial patterns in nesting behaviour. Temporal distribution analysis showed peak nesting activity consistently between 11:00-13:00. Species-specific behavioural patterns emerged clearly: chicken activity typically began earlier (from 08:00) compared to Guinea fowl activity which peaked in afternoon hours.

Figure 24. Comparative analysis of species-specific activity patterns



Species-specific behaviour analysis revealed distinct patterns. Chickens demonstrated broader temporal distribution, with activity spanning 09:00 to 16:00, utilizing both upper and lower rows despite showing clear preference patterns. Guinea fowls exhibited more selective behaviour, concentrating activity between 12:00-14:00, with peak occupancy observed in nest u4. They showed strong preference for upper-row nest boxes (92.3% of total Guinea fowl occupancy) and utilized only 3 of 8 boxes (U3, U4, D4) as seen in Figure 25.

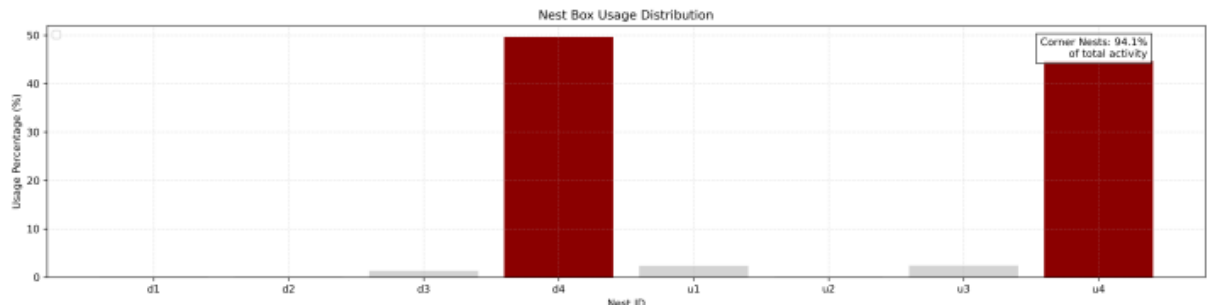
Figure 25. Temporal patterns in nest box usage



Spatial preferences analysis revealed strong corner preference patterns. Nest boxes U4 and D4 (corner location) dominated activity, demonstrating significantly higher occupancy

compared to outer nest boxes (U1, D1) which showed minimal usage. Intermediate nest boxes (U2, U3, D2, D3) maintained a low use, primarily during peak hours when preferred nests were occupied. This pattern remained remarkably consistent across all monitoring days, with corner nests accounting for 94.1% of total observed nesting activity (Figure 26).

Figure 26. Spatial analysis of nest box utilization patterns

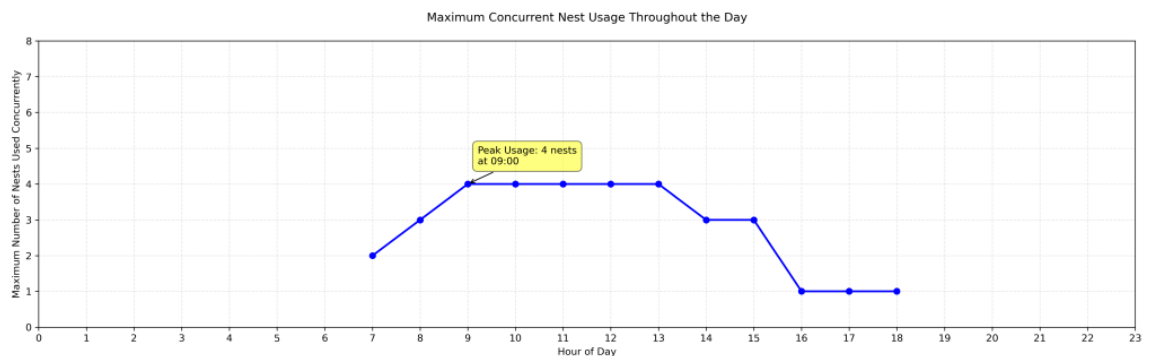


Resource utilization analysis

Analysis of nest box usage across the 11-day period revealed clear patterns and potential optimization opportunities. Occupancy distribution showed highly uneven utilization patterns, with corner nest boxes (U4, D4) dominating usage. The remaining six boxes combined showed significantly lower occupancy levels, suggesting potential for resource optimization.

Capacity assessment revealed that peak concurrent usage never exceeded 4 boxes simultaneously, with concurrent usage primarily observed during 09:00-13:00 period.

Figure 27. Temporal distribution of peak nest box occupancy



Production monitoring analysis

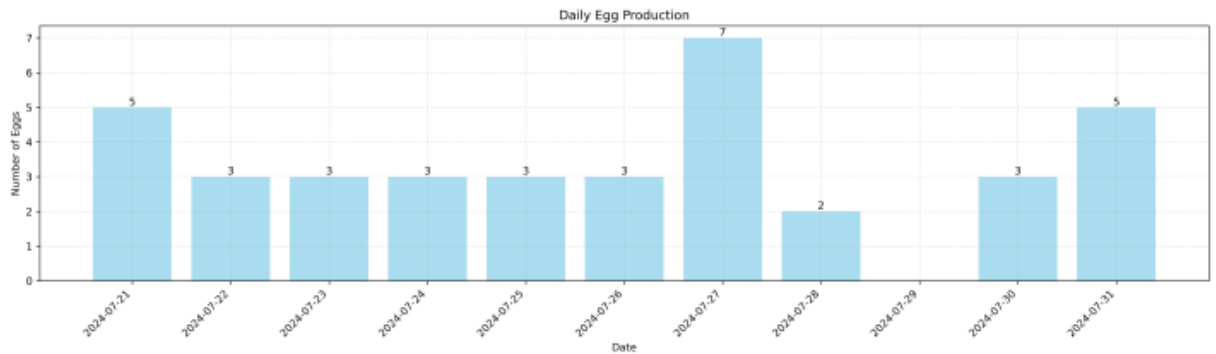
The system demonstrated varying effectiveness in egg production monitoring across different environmental conditions and times. Upper-row monitoring presented higher challenges in egg detection due to physical limitations in camera mounting, while lower-row monitoring maintained accuracy through varying conditions except during full occlusion during laying (Figure 28).

Figure 28. Analysis of detection challenges during complete egg occlusion



Collection timing optimization analysis identified two key windows: an optimal collecting period between 14:00-16:00, following peak nesting activity with a high percentage of daily laying completed, and a secondary collection window during 11:00-12:00, capturing early laying activity while reducing risk of damage or dirt from upcoming peak activity, based on the previous average activity analysis (Figure 24). Analysis of the data also revealed a 3.36 daily average egg count (Figure 29). The observed daily average of 3.36 eggs represents detected nest box production only. The semi-outdoor housing system, providing constant outdoor access, introduces additional complexity to production monitoring as hens may lay eggs in alternative locations both within and outside the covered coop area. This behavioural flexibility in nesting site selection, while natural, presents challenges for comprehensive production monitoring through fixed-position camera systems

Figure 29. Analysis of egg production patterns



5 Discussion

The implementation of CV in agricultural settings reveals an interplay between technological capabilities and contextual adaption. The two case studies, examining both institutional and individual farming operations, demonstrate how CV applications can be effectively scaled and tailored while maintaining their core objective of enhancing animal welfare through resource optimization. This discussion further explores the implications of the case findings and proposes pathways for future development and CV integration in respect to the unique characteristics of each case.

The contrast between institutional and individual farming contexts initially appears challenging for technological implementation. However, the findings suggest that this difference in scale actually presents unique opportunities for innovation. In both cases, the CV implementation demonstrated both successes and limitations, particularly in handling environmental variability and maintaining detection accuracy under challenging or prolonged occlusions. A state machine approach proved effective in both contexts, through its implementation needed careful adaptation to each specific case.

In the **institutional setting** of the campus cowshed, the existing technological infrastructure and multiple stakeholders create a fertile ground for developing a comprehensive digital ecosystem. The potential for expansion builds naturally upon current systems through strategic integration.

Infrastructure enhancement

The enhancement of infrastructure represents a crucial step toward comprehensive monitoring capabilities within the campus cowshed. Strategic camera positioning plays a vital role in overcoming current visual limitations by providing multiple perspective coverage. The proposed expansion includes careful placement of cameras across key monitoring zones, encompassing resting areas, feeding zones, automatic milking stations, and water trough surroundings. This expanded network enables thorough coverage, facilitating detailed understanding of individual cow activities and overall herd dynamics.

Integration with existing sensing devices and automated milking systems forms another critical component of the infrastructure enhancement. Through the development of an interconnected sensor network, the system can aggregate data from multiple sources, creating a more robust and comprehensive monitoring solution.

Individual animal tracking development

The advancement of individual animal monitoring capabilities represents a significant evolution in the system's functionality. Through the implementation of advanced object recognition techniques, the system can efficiently identify and track individual cows across multiple spaces and camera networks within the campus cowshed. This enhanced identification capability enables precise tracking and monitoring of individual animals, forming the foundation for more sophisticated analysis.

The creation of comprehensive animal profiles serves as a natural progression from individual identification. These profiles encompass a wide range of essential information, including body condition scores, health status indicators, gait analysis results, and detailed activity patterns such as laying, drinking, eating, and walking behaviours. By compiling and maintaining a database of this information, the system provides a holistic view of both individual animals and the entire herd over extended periods, enabling thorough assessment of well-being and productivity metrics.

This progression in the institutional context follows a natural evolution, cumulating in an intelligent information distribution system that creates an interconnected knowledge network, adapting information depth and presentation based on user roles and requirements, creating multiple layers of engagement.

Multi-stakeholder information distribution:

The distribution of information across multiple stakeholders requires a carefully structured approach with distinct layers of access and engagement:

The **public** engagement layer provides accessible farm history and context, general herd information, breed characteristics, production statistics, farm or animal milestones, and interactive informational displays. This layer serves educational and transparency purposes while maintaining appropriate data privacy.

The **educational** layer offers more detailed access for academic purposes, including student project data access such as production metrics, feed/water consumption parameters, milk parameters, and behavioural data. Teachers receive additional access for resource generation and validation, enabling the creation of real-time case studies, data-driven assignments, project validation tools, and cross-disciplinary teaching materials.

The **research** and **management** layer provides the most comprehensive access, including detailed data analysis capabilities covering individual animal profiles, herd behavioural patterns, health and welfare indicators, and resource utilization metrics. Management insights include predictive health alerts, production efficiency analysis, and welfare status monitoring.

This multi-layered system functions similarly to a natural network, creating vital connections between different information systems and stakeholders. Each layer contributes to and benefits from the network: visitor interactions generate engagement metrics, student projects provide fresh analytical perspectives, teaching applications cross-validate accuracy or reveal new research questions and barn management benefits from both the enriched barn and animal information generated by in- and out-campus network users.

In contrast, the **backyard chicken operation** presents opportunity to explore innovative yet accessible approaches to monitoring and welfare enhancement. While the scale is smaller, the technical implementation reveals several pathways for system enhancement:

Enhanced monitoring integration:

The implementation of enhanced monitoring systems requires a comprehensive, multi-faceted approach to data collection and analysis. A multi-camera setup forms the

foundation of this enhanced system, with strategic positioning enabling depth perception, reducing occlusion through varied angles, and providing an expanded field of view. This configuration significantly improves monitoring reliability across all observation zones.

Integration of acoustic monitoring represents a crucial advancement in comprehensive welfare assessment. The combination of visual and acoustic data streams enables more nuanced behavioural analysis, with vocalization patterns providing additional insights into animal welfare states. This multi-modal approach, particularly when combined with environmental parameter tracking including temperature, humidity, and air quality measurements, creates a robust framework for holistic monitoring.

Individual animal profile development system

The development of comprehensive individual animal profiles builds upon the enhanced monitoring capabilities through several interconnected components. The system implements individual bird tracking and identification utilizing multi-angle visual data collection, movement pattern analysis, and behavioural signature mapping.

Health and productivity monitoring integrates multiple data streams to create detailed individual profiles. The system tracks nesting behaviour patterns, feed and water interaction frequencies, and social interaction dynamics, correlating these with production performance metrics. This comprehensive approach enables early detection of health issues and optimization of resource allocation.

The integration of visual and acoustic monitoring shows particular promise, with recent research demonstrating not just the effectiveness but also the potentially low cost of a sound analysis system for a poultry welfare assessment through piezoelectric sensors (Amirivojdan et al., 2024). By combining multiple data streams from strategically positioned cameras with acoustic monitoring, small-scale operations can achieve sophisticated state-of-the-art monitoring capabilities while maintaining operational simplicity, without requiring overwhelming infrastructure investments. The integration of these monitoring systems with profile development creates a synergistic effect, where each component enhances the effectiveness of the others. For instance, behavioural patterns become more precise when correlated with environmental data, while health assessments can benefit from the combination of visual, acoustic and environmental indicators.

While CV systems may eventually reach a “plug and play” stage with standardized monitoring profiles for common livestock issues (similar to existing plant disease detection systems), the unique characteristics of each agricultural setting demand human insight for optimal implementation. The farmers understanding of their specific context and niche environment, whether large institutional setting or a small backyard operation, becomes crucial in training and adapting these systems to provide meaningful insights. The success of these systems – whether institutional or individual – and other similar technological implementations, ultimately rests upon achieving this fundamental human-technology symbiosis. Finding the dynamic balance between technological capabilities and human expertise is crucial. By developing systems that can be calibrated and refined through practical agricultural knowledge, this technological advancement serves to amplify rather than replace traditional farming wisdom, ensuring the preservation of essential agricultural knowledge while enabling its adaption to modern challenges.

6 Acknowledgements

The case studies found in this thesis work were conducted as part of my student traineeship, which was made possible through the guidance and support of several individuals such as my supervising teachers, and especially Olli Koskela and Nathaniel Narra, whose expertise was instrumental in guiding the case studies. It is important to note that the case studies represent initial explorations into practical computer vision implementations, serving as both research contributions and learning opportunities. The research material and resources are archived and are freely accessible in the Zenodo repository (Tousios, D. 2025).

References

- Amirivojdan, A., Nasiri, A., Zhou, S., Zhao, Y., & Gan, H. (2024). Chickensense: A low-cost deep learning-based solution for poultry feed consumption monitoring using sound technology. *AgriEngineering*, 6(3), 2115–2129. <https://doi.org/10.3390/agriengineering6030124>
- Aydin, A. (2017). Using 3D vision camera system to automatically assess the level of inactivity in broiler chickens. *Computers and Electronics in Agriculture*, 135, 4–10. <https://doi.org/10.1016/j.compag.2017.01.024>
- Aydin, A., Cangar, O., Ozcan, S. E., Bahr, C., & Berckmans, D. (2010). Application of a fully automatic analysis tool to assess the activity of broiler chickens with different gait scores. *Computers and Electronics in Agriculture*, 73(2), 194–199. <https://doi.org/10.1016/j.compag.2010.05.004>
- Aydin, A., Pluk A., Leroy T., Berckmans D., Bahr C. (2013). Automatic identification of activity and spatial use of broiler chickens with different gait scores. *Transactions of the ASABE*, 1123–1132. <https://doi.org/10.13031/trans.56.9987>
- Bakhshayeshi, I., Erfani, E., Taghikhah, F. R., Elbourn, S., Beheshti, A., & Asadnia, M. (2024). An intelligence cattle reidentification system over transport by siamese neural networks and yolo. *IEEE Internet of Things Journal*, 11(2), 2351–2363. <https://doi.org/10.1109/JIOT.2023.3294944>
- Benjamin, M., & Yik, S. (2019, March). Precision Livestock Farming in Swine Welfare: A Review for Swine Practitioners. MDPI. <https://www.mdpi.com/2076-2615/9/4/133/pdf>
- Bhuiyan, M. R., & Wree, P. (2023). Animal behavior for chicken identification and monitoring the health condition using computer vision: A systematic review. *IEEE Access*, 11, 126601–126610. <https://doi.org/10.1109/ACCESS.2023.3331092>
- Borges Oliveira, D. A., Ribeiro Pereira, L. G., Bresolin, T., Pontes Ferreira, R. E., & Reboucas Dorea, J. R. (2021). A review of deep learning algorithms for computer vision systems in livestock. *Livestock Science*, 253, 104700. <https://doi.org/10.1016/j.livsci.2021.104700>
- Cangar, Ö., Leroy, T., Guarino, M., Vranken, E., Fallon, R., Lenehan, J., Mee, J., & Berckmans, D. (2008). Automatic real-time monitoring of locomotion and posture behaviour of pregnant cows prior to calving using online image analysis. *Computers and Electronics in Agriculture*, 64(1), 53–60. <https://doi.org/10.1016/j.compag.2008.05.014>
- Chien, Y.-R., & Chen, Y.-X. (2018). An rfid-based smart nest box: An experimental study of laying performance and behavior of individual hens. *Sensors*, 18(3), 859. <https://doi.org/10.3390/s18030859>
- Cronin, G. M., Borg, S. S., & Dunn, M. T. (2008). Using video image analysis to count hens in cages and reduce egg breakage on collection belts. *Australian Journal of Experimental Agriculture*, 48(7), 768–772. <https://doi.org/10.1071/EA07404>
- Dennis, R. L., Fahey, A. G., & Cheng, H. W. (2008). Different effects of individual identification systems on chicken well-being1. *Poultry Science*, 87(6), 1052–1057. <https://doi.org/10.3382/ps.2007-00240>

- Dokic, K., Blaskovic, L., & Mandusic, D. (2020). From machine learning to deep learning in agriculture – the quantitative review of trends. *IOP Conference Series: Earth and Environmental Science*, 614(1), 012138. <https://doi.org/10.1088/1755-1315/614/1/012138>
- Fernandes, A. F. A., Dórea, J. R. R., & Rosa, G. J. de M. (2020). Image analysis and computer vision applications in animal sciences: An overview. *Frontiers in Veterinary Science*, 7. <https://doi.org/10.3389/fvets.2020.551269>
- Halachmi, I., Guarino, M., Bewley, J., & Pastell, M. (2019). Smart animal agriculture: Application of real-time sensors to improve animal well-being and production. *Annual Review of Animal Biosciences*, 7(1), 403–425. <https://doi.org/10.1146/annurev-animal-020518-114851>
- Hansen, M. F., Smith, M. L., Smith, L. N., Abdul Jabbar, K., & Forbes, D. (2018). Automated monitoring of dairy cow body condition, mobility and weight using a single 3D video capture device. *Computers in Industry*, 98, 14–22. <https://doi.org/10.1016/j.compind.2018.02.011>
- Helwatkar, A., Riordan, D., & Walsh, J. (2014). Sensor technology for animal health monitoring. *International Journal on Smart Sensing and Intelligent Systems*, 7(5), 1–6. <https://doi.org/10.21307/ijssis-2019-057>
- Huang, T. (1996). Computer vision: Evolution and promise. <https://doi.org/10.5170/CERN-1996-008.21>
- Jiang, B., Tang, W., Cui, L., & Deng, X. (2023). Precision livestock farming research: A global scientometric review. *Animals*, 13(13), 2096. <https://doi.org/10.3390/ani13132096>
- Jocher, G., Chaurasia, A., & Qiu, J. (2023). Ultralytics YOLO (Version 8.0.0) [Computer software]. <https://github.com/ultralytics/ultralytics>
- Kakani, V., Nguyen, V. H., Kumar, B. P., Kim, H., & Pasupuleti, V. R. (2020). A critical review on computer vision and artificial intelligence in food industry. *Journal of Agriculture and Food Research*, 2, 100033. <https://doi.org/10.1016/j.jafr.2020.100033>
- Knapp, H., & Uckelmann, D. (2022). Literature review on sources of interference and proposed solutions for RFID installations in complex production and logistics processes in the automotive industry. *International Journal of RF Technologies*, 12(2), 87–126. <https://doi.org/10.3233/RFT-210312>
- Koskela, O., Benitez Pereira, L. S., Pölönen, I., Aronen, I., & Kunttu, I. (2022). Deep learning image recognition of cow behavior and an open data set acquired near an automatic milking robot. *Agricultural and Food Science*. <https://doi.org/10.23986/afsci.111665>
- Kristensen, H. H., & Cornou, C. (2011). Automatic detection of deviations in activity levels in groups of broiler chickens – A pilot study. *Biosystems Engineering*, 109(4), 369–376. <https://doi.org/10.1016/j.biosystemseng.2011.05.002>
- Kumar, P., Luo, S., & Shaukat, K. (2023). A comprehensive review of deep learning approaches for animal detection on video data. *International Journal of Advanced Computer Science and Applications*, 14(11). <https://doi.org/10.14569/IJACSA.2023.01411144>
- Lei, Z., Helen, G., Xujiang, Y., Lisa C., Nigel Al. (2019 March). Automatic Individual Pig Detection and Tracking in Pig Farms. *Sensors*. <https://www.mdpi.com/1424-8220/19/5/1188/pdf>

- Leroy T., Vranken E., Van Brecht A., Struelens E., Sonck B., & Berckmans D. (2006). A computer vision method for on-line behavioral quantification of individually caged poultry. *Transactions of the ASABE*, 49(3), 795–802. <https://doi.org/10.13031/2013.20462>
- Leroy, T., Vranken E., Struelens E., Sonck B., & D. Berckmans D. (2005). Computer vision based recognition of behavior phenotypes of laying hens. 2005 Tampa, FL July 17-20, 2005. 2005 Tampa, FL July 17-20, 2005. <https://doi.org/10.13031/2013.19471>
- Li, G., Huang, Y., Chen, Z., Chesser, G. D., Purswell, J. L., Linhoss, J., & Zhao, Y. (2021). Practices and applications of convolutional neural network-based computer vision systems in animal farming: A review. *Sensors*, 21(4), 1492. <https://doi.org/10.3390/s21041492>
- Li, S., Fu, L., Sun, Y., Mu, Y., Chen, L., Li, J., & Gong, H. (2021). Individual dairy cow identification based on lightweight convolutional neural network. *PLOS ONE*, 16(11), e0260510. <https://doi.org/10.1371/journal.pone.0260510>
- Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., & Dollár, P. (2015). Microsoft coco: Common objects in context (arXiv:1405.0312). arXiv. <http://arxiv.org/abs/1405.0312>
- Liu, Y., Ma, X., Shu, L., Hancke, G. P., & Abu-Mahfouz, A. M. (2021). From industry 4. 0 to agriculture 4. 0: Current status, enabling technologies, and research challenges. *IEEE Transactions on Industrial Informatics*, 17(6), 4322–4334. <https://doi.org/10.1109/TII.2020.3003910>
- Lowe, G., Sutherland, M., Waas, J., Schaefer, A., Cox, N., & Stewart, M. (2019). Infrared thermography—A non-invasive method of measuring respiration rate in calves. *Animals*, 9(8), 535. <https://doi.org/10.3390/ani9080535>
- Macedonia, M. (2003). The GPU enters computing's mainstream. *Computer*, 36(10), 106–108. doi:10.1109/mc.2003.1236476
- Marr, D. (2010). *Vision: A computational investigation into the human representation and processing of visual information*. The MIT Press. <https://doi.org/10.7551/mitpress/9780262514620.001.0001>
- Milan, H. F. M., Maia, A. S. C., & Gebremedhin, K. G. (2016). Technical note: Device for measuring respiration rate of cattle under field conditions¹. *Journal of Animal Science*, 94(12), 5434–5438. <https://doi.org/10.2527/jas.2016-0904>
- Mohialdin, A. M., Elbarrany, A. M., & Atia, A. (2023). Chicken behavior analysis for surveillance in poultry farms. *International Journal of Advanced Computer Science and Applications*, 14(3). <https://doi.org/10.14569/IJACSA.2023.01403106>
- National Academies of Sciences, E., Studies, D. on E. and L., Resources, B. on A. and N., & Cattle, C. on N. R. of D. (2021). *Water*. In *Nutrient Requirements of Dairy Cattle: Eighth Revised Edition*. National Academies Press (US). <https://www.ncbi.nlm.nih.gov/books/NBK600589/>
- Neethirajan, S. (2022). ChickTrack – A quantitative tracking tool for measuring chicken activity. *Measurement*, 191, 110819. <https://doi.org/10.1016/j.measurement.2022.110819>
- Neethirajan, S., & Kemp, B. (2021). Digital livestock farming. *Sensing and Bio-Sensing Research*, 32, 100408. <https://doi.org/10.1016/j.sbsr.2021.100408>

- Oh, K.-S., & Jung, K. (2004). GPU implementation of neural networks. *Pattern Recognition*, 37(6), 1311–1314. <https://doi.org/10.1016/j.patcog.2004.01.013>
- Okinda, C., Nyalala, I., Korohou, T., Okinda, C., Wang, J., Achieng, T., Wamalwa, P., Mang, T., & Shen, M. (2020). A review on computer vision systems in monitoring of poultry: A welfare perspective. *Artificial Intelligence in Agriculture*, 4, 184–208. <https://doi.org/10.1016/j.aiaa.2020.09.002>
- Padilla, R., Netto, S. L., & Da Silva, E. A. B. (2020). A survey on performance metrics for object-detection algorithms. *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 237–242. <https://doi.org/10.1109/IWSSIP48289.2020.9145130>
- Papert, S. A. (1966). The summer vision project. <https://dspace.mit.edu/handle/1721.1/6125>
- Parras-Burgos, D., Molina Martínez, J. M., & García-Mateos, G. (Eds.). (2023). *Agriculture 4.0 – the future of farming technology*. MDPI. <https://doi.org/10.3390/books978-3-0365-9759-1>
- Poursaberi, A., Bahr, C., Pluk, A., Van Nuffel, A., & Berckmans, D. (2010). Real-time automatic lameness detection based on back posture extraction in dairy cattle: Shape analysis of cow with image processing techniques. *Computers and Electronics in Agriculture*, 74(1), 110–119. <https://doi.org/10.1016/j.compag.2010.07.004>
- Rayhana, R., Xiao, G., & Liu, Z. (2021). Rfid sensing technologies for smart agriculture. *IEEE Instrumentation & Measurement Magazine*, 24(3), 50–60. <https://doi.org/10.1109/MIM.2021.9436094>
- Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. arXiv.
- Rohan, A., Rifaq, M. S., Hasan, Md. J., Asghar, F., Bashir, A. K., & Dottorini, T. (2024). Application of deep learning for livestock behaviour recognition: A systematic literature review. *Computers and Electronics in Agriculture*, 224, 109115. <https://doi.org/10.1016/j.compag.2024.109115>
- Shao, H., Pu, J., & Mu, J. (2021). Pig-posture recognition based on computer vision: Dataset and exploration. *Animals*, 11(5), 1295. <https://doi.org/10.3390/ani11051295>
- Shen, W., Hu, H., Dai, B., Wei, X., Sun, J., Jiang, L., & Sun, Y. (2019). Individual identification of dairy cows based on convolutional neural networks. *Multimedia Tools and Applications*. doi:10.1007/s11042-019-7344-7
- Skalski, P. (2019). Make sense. [Computer software]. <https://github.com/SkalskiP/make-sense/>
- Song, X., Leroy, T., Vranken, E., Maertens, W., Sonck, B., & Berckmans, D. (2008). Automatic detection of lameness in dairy cattle—Vision-based trackway analysis in cow's locomotion. *Computers and Electronics in Agriculture*, 64(1), 39–44. <https://doi.org/10.1016/j.compag.2008.05.016>
- Stevens, K. A. (2012). The vision of david marr. *Perception*, 41(9), 1061–1072. <https://doi.org/10.1068/p7297>
- Supriyanto, E., Isnanto, R. R., & Hadi Purnomo, S. (2023). Computer vision in chicken monitoring system using machine learning: A general review. *E3S Web of Conferences*, 448, 02014. <https://doi.org/10.1051/e3sconf/202344802014>

- Tassinari, P., Bovo, M., Benni, S., Franzoni, S., Poggi, M., Mammi, L. M. E., Mattoccia, S., Di Stefano, L., Bonora, F., Barbaresi, A., Santolini, E., & Torreggiani, D. (2021). A computer vision approach based on deep learning for the detection of dairy cows in free stall barn. *Computers and Electronics in Agriculture*, 182, 106030. <https://doi.org/10.1016/j.compag.2021.106030>
- Tedeschi, L. O., Greenwood, P. L., & Halachmi, I. (2021). Advancements in sensor technology and decision support intelligent tools to assist smart livestock farming. *Journal of Animal Science*, 99(2), skab038. <https://doi.org/10.1093/jas/skab038>
- Temenos, A., Voulodimos, A., Korelidou, V., Gelasakis, A., Kalogeras, D., Doulamis, A., & Doulamis, N. (2024). Goat-CNN: A lightweight convolutional neural network for pose-independent body condition score estimation in goats. *Journal of Agriculture and Food Research*, 16, 101174. <https://doi.org/10.1016/j.jafr.2024.101174>
- Terven, J., Córdova-Esparza, D.-M., & Romero-González, J.-A. (2023). A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. *Machine Learning and Knowledge Extraction*, 5(4), 1680–1716. <https://doi.org/10.3390/make5040083>
- Tian, H., Wang, T., Liu, Y., Qiao, X., & Li, Y. (2020). Computer vision technology in agricultural automation—A review. *Information Processing in Agriculture*, 7(1), 1–19. <https://doi.org/10.1016/j.inpa.2019.09.006>
- Tousios, D. (2025). Dataset for "Computer Vision for Resource Optimisation in Dairy Cows Drinking and Backyard Chickens Nesting behaviour" [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.15031180>
- Tsai, D.-M., & Huang, C.-Y. (2014). A motion and image analysis method for automatic detection of estrus and mating behavior in cattle. *Computers and Electronics in Agriculture*, 104, 25–31. <https://doi.org/10.1016/j.compag.2014.03.003>
- Viazzi, S., Bahr, C., Schlageter-Tello, A., Hertem, T. V., Romanini, C. E. B., Pluk, A., Halachmi, I., Lokhorst, C., & Berckmans, D. (2013). Analysis of individual classification of lameness using automatic measurement of back posture in dairy cattle. *Journal of Dairy Science*, 96(1), 257–266. <https://doi.org/10.3168/jds.2012-5806>
- Viazzi, S., Bahr, C., Van Hertem, T., Schlageter-Tello, A., Romanini, C. E. B., Halachmi, I., Lokhorst, C., & Berckmans, D. (2014). Comparison of a three-dimensional and two-dimensional camera system for automated measurement of back posture in dairy cows. *Computers and Electronics in Agriculture*, 100, 139–147. <https://doi.org/10.1016/j.compag.2013.11.005>
- Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep learning for computer vision: A brief review. *Computational Intelligence and Neuroscience*, 2018, 1–13. <https://doi.org/10.1155/2018/7068349>
- Wojke, N., Bewley, A., & Paulus, D. (2017). Simple online and realtime tracking with a deep association metric. 2017 IEEE International Conference on Image Processing (ICIP), 3645–3649. <https://doi.org/10.1109/ICIP.2017.8296962>

- Wu, D., Han, M., Song, H., Song, L., & Duan, Y. (2023). Monitoring the respiratory behavior of multiple cows based on computer vision and deep learning. *Journal of Dairy Science*, 106(4), 2963–2979. <https://doi.org/10.3168/jds.2022-22501>
- Wu, D., Yin, X., Jiang, B., Jiang, M., Li, Z., & Song, H. (2020). Detection of the respiratory rate of standing cows by combining the Deeplab V3+ semantic segmentation model with the phase-based video magnification algorithm. *Biosystems Engineering*, 192, 72–89. <https://doi.org/10.1016/j.biosystemseng.2020.01.012>
- Wurtz, K., Camerlink, I., D'Eath, R. B., Fernández, A. P., Norton, T., Steibel, J., & Siegford, J. (2019). Recording behaviour of indoor-housed farm animals automatically using machine vision technology: A systematic review. *PLOS ONE*, 14(12), e0226669. <https://doi.org/10.1371/journal.pone.0226669>
- Yadav, C. S., Peixoto, A. A. T., Rufino, L. A. L., Silveira, A. B., & Alexandria, A. R. de. (2024). Intelligent classifier for identifying and managing sheep and goat faces using deep learning. *AgriEngineering*, 6(4), 3586–3601. <https://doi.org/10.3390/agriengineering6040204>
- Yu, R., Wei, X., Liu, Y., Yang, F., Shen, W., & Gu, Z. (2024). Research on Automatic Recognition of Dairy Cow Daily Behaviors Based on Deep Learning. *Animals : an Open Access Journal from MDPI*, 14. <https://www.mdpi.com/2076-2615/14/3/458>
- Yu, W., Liu, Y., Yang, X., Xie, R., & Wang, C. (2024). Yolosortc3d: A yolov8, deepsort, and c3d-based method for sheep rumination behavior recognition. *Applied Sciences*, 14(23), 10801. <https://doi.org/10.3390/app142310801>
- Feiyang, Z., Yueming, H., Liancheng, C., Lihong, G., Wenjie, D., & Lu, W. (2016). Monitoring behavior of poultry based on RFID radio frequency network. *International Journal of Agricultural and Biological Engineering*, 9(6), 139–147. <https://doi.org/10.25165/ijabe.v9i6.1568>
- Zhuang, X., Bi, M., Guo, J., Wu, S., & Zhang, T. (2018). Development of an early warning algorithm to detect sick broilers. *Computers and Electronics in Agriculture*, 144, 102–113. <https://doi.org/10.1016/j.compag.2017.11.032>