BEN-GURION UNIVERSITY OF THE NEGEV
THE FACULTY OF ENGINEERING SCIENCES
DEPARTMENT OF INDUSTRIAL ENGINEERING AND MANAGEMENT

COMPUTER VISION DEVELOPMENTS IN PRECISION LIVESTOCK USING MACHINE LEARNING. TWO USE CASES: A MOBILE SYSTEM FOR COUNTING LAYING HENS A SHEEP BIOMETRIC IDENTIFICATION SYSTEM.

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE M.Sc. DEGREE

By: Almog Hitelman

Supervised by: Prof. Yael Edan and Prof. Ilan Halachmi

September 2021

BEN-GURION UNIVERSITY OF THE NEGEV
THE FACULTY OF ENGINEERING SCIENCES
DEPARTMENT OF INDUSTRIAL ENGINEERING AND MANAGEMENT

COMPUTER VISION DEVELOPMENTS IN PRECISION LIVESTOCK USING MACHINE LEARNING. TWO USE CASES: A MOBILE SYSTEM FOR COUNTING LAYING HENS A SHEEP BIOMETRIC IDENTIFICATION SYSTEM.

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE M.Sc. DEGREE

By: Almog Hitelman

Supervised by: Prof. Yael Edan and Prof. Ilan Halachmi

Author:…………………………………………… Date: 19.9.2021

Supervisor:…………………………………… Date: 19.9.2021

Supervisor:…………………………………… Date: 19.9.2021

Chairman of Graduate Studies Committee:……………… Date: 19/9/2021

September 2021

Acknowledgements

This study was supported by the Israeli Chief Scientist of Agriculture fund "Kandel" PLF center of expertise (20-12-0029, 459451415), "TechCare" (GA No 862050) and "Sm@RT" (GA No 101000471). Partial support was provided by Ben-Gurion University of the Negev through the W. Gunther Plaut Chair in Manufacturing Engineering.

We gratefully thank all members of the Precision Livestock Farming (PLF) Laboratory and the staff of the Volcani research cowshed for their assistance in the study. Special thanks to Assaf Godo, Ron Berenstein, Joseph Lepar and Noam Bergman for their advice and contribution on mechanical design and computer vision, and to A. Rosov, who hosted us in the sheep pen of the Volcani Center. Thanks to my friend, May Saar, for support, encouragement and for sharing with me her vast knowledge throughout my studies and research.

I wish to express my sincere gratitude to my academic advisors, Prof. Yael Edan and Prof. Ilan Halachmi, who provided me guidance, support and insights over the course of the research.

Thank you all.

Abstract

Precision Livestock Farming (PLF) technology aims to improve farming productivity and animal welfare by ensuring better livestock management. Such systems are supported by monitoring the animal's needs. Two case studies of PLF technologies based on computer vision are developed in this thesis: a mobile system for counting laying hens and a sheep biometric identification system. In both systems advanced machine learning methods were applied.

Counting laying hens

The Israeli laying hens industry is regulated by production quota; a farm can produce eggs according to a prefixed allocated number of hens. With the new community battery cages that have been recently introduced to the Israeli egg industry, which house thousands of hens, a manual head count of the hens is a very difficult task resulting in inaccurate and meticulous results. In this study, a machine vision system was developed to detect and count laying hens in community battery cages. The project aims to replace the manual counting, with an automatic algorithm.

A mobile system which can be driven along the cages alley was developed. The experiments were conducted at a commercial hen house located in Kidron. The hen house stacked in 6 floors, with 37 community cages set in a rows and between 30-40 hens in per cage. The hens were video-recorded with an Intel RealSense RGBD D435 camera, which acquired images at 30 frames per second (fps). Videos were processed with the Faster R-CNN algorithm. After detection, the hens' locations were tracked using a tracking algorithm, which assigns every detection with an ID, representing a single hen. Testing on a dataset that included 5600 images resulted in detection accuracy of 88%, with mean absolute error of 4.5 hens per cage. Sensitivity analyses revealed that the minimum number of frames needed for high detection while shortening runtime is 35 frames, instead of using a full video, which consists of more than 44 frames.

The system is inexpensive, fast, user-friendly, and does not rely on a specific feature of the test hen house. Thus, it can potentially be used in different farms, ensuring adequate hen density according to the limits set by regulations. Future work should include a depth channel to improve results.

Sheep biometric identification

A sheep biometric identification system based on facial images was developed. A machine vision system and deep learning models were developed and applied for animal identification. The system included two 8-MegaPixels cameras installed in a monitoring drinking facility adapted to work with NVIDIA Jetson Nano embedded system-on-module (SoM). Data from 81 Assaf breed sheep aged two to three months, from two different groups of sheep, were collected over a period of two weeks. The biometric identification model included two steps: face detection and classification. In order to locate and localize the sheep face in an image, the Faster R-CNN deep learning object detection algorithm was applied. The detected face was provided as input to seven different classification models. Different transfer learning methods were examined. The best performance was obtained using a ResNet50V2 model with the state-of-art ArcFace loss function. The identification system resulted in average accuracies of 95.4% and 95.7% image detection for the two groups tested and in 100% sheep identification. When applying transfer learning method, average identification accuracies improved to 97% in both groups (with 100% sheep identification), and the training process was accomplished more rapidly.

Key words: Precision Livestock Farming (PLF), laying hens, counting, sheep, lamb, algorithm, Deep learning, convolutional neural network, object detection, Biometric identification, Faster R-CNN, Face recognition.

Publications

Journal papers submitted

- Hitelman, A., Edan, Y., Godo, A., Berenstein, R., Lepar, J., Halachmi, I. Biometric identification of sheep via machine-vision system. Revised September, 2021 (Appendix 2A).
- Hitelman, A., Edan, Y., Godo, A., Berenstein, R., Lepar, J., Halachmi, I. Short Communication: The effect of age on sheep biometric identification. Submitted August, 2021 (Appendix 2B).

Conference papers and presentations

- Hitelman, A., Edan, Y., Halachmi, I. Machine-vision system for biometric identification of sheep. 33rd Annual Conference on Cattle and Sheep Sciences, Virtual conference, Israel, 2021. Best research award.

Table of Contents

List of Figures

List of Tables

List of Appendices

1 Introduction

1.1 Problem description and motivation

1.1.1 Precision livestock farming

Precision Livestock Farming (PLF) aims to improve the efficiency of production, while increasing animal and human welfare, by means of applying advanced information, targeted resource use and precise control of the production process (Banhazi et al., 2012). Precision livestock farming (PLF) develops real-time tools for monitoring livestock with information collected without the stress of animal disturbance or handling (Hamilton et al., 2004). The assumption is that animals that are provided with optimal conditions will yield maximum production (Halachmi et al., 2019). In recent years, the importance of monitoring livestock animals increased and has been applied to all types of livestock animals, such as cows (Bloch et al., 2019), sheep (Morgan-davies et al., 2018), pigs (Hemeryck et al., 2015), broilers (Fontana et al., 2015) and hens (Sassi et al., 2016). PLF can provide objective animal welfare assessment in modern livestock production (Werkheiser, 2018) and improve management.

Recently, new image recognition models based on machine learning models, like the Convolution Neural Network (CNN), allow to better understand complex processes in agricultural environments (Liakos et al., 2018). These machine learning methods produce better results than traditional image processing techniques (Gongal et al., 2015). As a result, more applications of livestock management (Qiao et al., 2019) have been automated using computer vision (Liakos et al., 2018) and deep learning methods (Guo et al., 2016). However, these techniques require a massive amount of data for building the models (Szegedy et al., 2016).

1.1.2 Counting laying hens

Intensive research has focused on hens and broilers monitoring with a variety of sensors and cameras (Neila et al., 2016). This includes inspection of poultry carcasses by spectral and hyperspectral imaging (Park and Chen, 2000), predicting eggs freshness with image processing technique (Suktanarak and Teerachaichayut, 2017), 3D imaging for broilers weight estimation (Mortensen et al., 2016), and visible light imaging for hens tracking (Kashiha et al., 2014).

Since the Israeli laying hens industry is regulated by quota (Geffen et al., 2019) in addition to the welfare and health considerations, which implies that a farm can produce eggs according to a fixed number of hens, monitoring and counting of hens must be done regularly (Geffen et al., 2019). With the new community battery cages that have been recently introduced to the Israeli egg industry (Appleby, 2003), which house thousands of hens, a manual head count of the hens is very difficult task which leads to inaccurate and meticulous results (Geffen et al., 2019). Moreover, manual counting is a laborious task which is a major cost component affecting profitability of farmers (Cronin et al., 2008). By developing an automatic counting system, efficiency of labor use can be improved and may lead to improvement of surveillance that may lead to better hen welfare (Cronin et al., 2008).

Laying hens counting in community battery cages is a challenging task: the hens do not stand still but constantly move, they stand at different distances and angles from the camera which effect the ability to contrast them from the background, and because of the battery cages structure, they can only be photographed from front view (Geffen et al., 2019). Therefore, not all hens in a cage are visible in each frame. These factors reduces image quality making automatic counting a difficult computer vision task (Geffen et al., 2019).

A recent neural network system based on Faster R-CNN (region-based convolutional neural network) was developed to count hens by using deep learning methods (Geffen et al., 2019). The feeder was equipped with a Media Tech W9R camera and the cages were video recorded while the feeder traveled along and acquired images (Geffen et al., 2019). An accuracy of 90% was obtained with an algorithm that used object detection, focusing on the hen's head. However, the algorithms were tailored designed to the specific conditions and were not adjustable to varying lighting conditions (Geffen et al., 2019). Moreover, they were fitted for the specific cages and did not fit cages with no separation in the middle, as customary in the industry. In these cages counting must include depth of the images (Geffen et al., 2019).

1.1.3 Biometric identification of sheep

Sheep farming has been limited in research in machine vision and deep learning applications (Morgan-Davies et al., 2018). However, more awareness to global sheep economics, animal welfare and agricultural policies, influence the sheep farming practices and stimulate wider adoption of deep learning systems (Morgan-davies et al., 2018), such as – individual monitoring of sheep (Salama et al., 2019), pain estimation (Mahmoud et al., 2018), and lamb growth monitoring (Zhuang et al., 2018).

Performance recording of sheep enable automated data collection that provides better quality data, contributing to better decision-making and thus improved management (Ait-Saidi et al., 2014). Collected data, including individual sheep ID and the recorded performances like body condition score, milk yield, and body weight (Salama et al., 2019), facilitate animal handling, contributing to improved husbandry practices, reducing labour requirements (Morris et al., 2012), and allowing better disease management (Salama et al., 2019). For those reasons, sheep ID should be unique and permanent for an adequate performance recording (Ait-Saidi et al., 2014) and for providing farmers an efficient way to recognize and track each individual in a large group of sheep (Salama et al., 2019).

Different methods of marking sheep were used by herders (Landais, 2001). Historically, the main methods used for sheep identification were; branding by fire or freezing (Landais, 2001), tattooing, ear tagging and electronic identification such as RFID tags and barcodes (Caja et al., 2004). However, these methods have proved inefficient (Koik and Ibrahim, 2012), and may harm the animal and even affect its behavior (Caja et al., 2004). A further key drawback of these methods is the higher cost and that they must be recorded manually, which can easily introduce human errors, while the labour cost of such a practice is also high (Trevarthen, 2007).

Due to the need of increased profitability with minimal unfavorable environmental impact and high concern of animal welfare nowadays (Mollo et al., 2010), using biometric traits instead of traditional identification methods, has gained a lot of attention in current livestock identification systems (Corkery et al., 2007).

Sheep facial biometrics include many significant features that can be used for identification such as muscles, the eyes, mouth and many hidden features and therefore are very promising (Corkery et al., 2007; Salama et al., 2019).

Sheep face recognition was achieved using a convolutional neural network (Salama et al., 2019). In this research, the Bayesian Optimization was used to automatically set the parameters for a convolutional neural network and in addition, the AlexNet configuration was also examined (Salama et al., 2019). The sheep recognition algorithms were tested on a data set of 52 sheep, with 10 images taken per sheep (Salama et al., 2019). The experiments achieved an accuracy of 98% (Salama et al., 2019). However, the research was tailored designed to the specific conditions and was tested on small set of sheep with great variability, and therefore may not be accurate enough in order to replace other biometric identification methods in used (Corkery et al., 2007).

1.2 Objectives

This research aimed to develop two automatic precision livestock farming systems: a mobile system for counting laying hens and a sheep biometric identification system.

1.2.1 Counting laying hens

The research objective was to develop a machine vision system that will automatically count hens in community battery cages, with the following specific deliverables:

- A mobile system which can be transferred between cages.
- Algorithms to detect and count hens in a battery cage.
- Algorithms that operate in varying illumination conditions.

The research was based on a previous system in which feasibility was proven (Geffen et al., 2019). The innovation of the current research was a new design which included a mobile platform equipped with a RGBD camera. The current project focused on the new design and developing algorithms for the variable illumination conditions. It used a previous developed tracking algorithm however, a new image processing and object detection algorithm were developed. Additionally, in depth sensitivity analyses were conducted.

1.2.2 Biometric identification of sheep

The objective of this work was to investigate the potential of facial recognition as a biometric-based identification system for sheep by using a machine vision system, with the following specific deliverables:

- An identification system that can be used for any sheep pens.
- Replacing the use of traditional methods with a sheep identification model.
- Examine whether sheep maturation affects identification.

1.3 Thesis overview

This thesis begins with a literature review presented in chapter 2. The review starts with precision livestock farming introduction (2.1), followed by computer vision research, including an overview of object detection in agriculture using deep learning (2.2). Next, the recent advancements of counting laying hens (2.3) and sheep biometric identification (2.5). The mobile system for detecting and counting laying hens by machine-vision processing of RGB images is described in chapter 4. Automated system for sheep identification based on deep learning is described in chapter 5. Both chapters describe the research methodology, the system and algorithms developed and results. Discussion and conclusions are discussed in chapter 6 and chapter 7.

2 Literature Review

This section reviews relevant literature on precision livestock farming (section 2.1), computer vision including object detection (section 2.2), laying hens counting (section 2.3), biometric identification and existing models of livestock identification (section 2.4), and finally a review of biometric identification of sheep (section 2.5).

2.1 Precision Livestock Farming

Automated monitoring and control techniques are becoming more important to support management by the farmer (Pham & Stack, 2018) in larger farms and improve production decisions. Precision Livestock Farming (PLF) mainly propose is to improve the efficiency of production, while increasing animal and human welfare, by means of applying advanced information, targeted resource use and precise control of the production process (Banhazi et al., 2012). PLF technologies, can be used to improve food safety and quality and to achieve efficient and sustainable livestock farming (Laberge and Rousseau, 2017).

PLF originated from the increased use of information technology (IT) products in support of livestock management (Guarino et al., 2008; Mertens et al., 2011). PLF objectively assesses animal welfare in modern livestock production (Dawkins, 2017). It continuously monitors individual animals on large farms using network devices, to compare this information to expected norms, and to use algorithms to automatically manage individual animals according to changes in climate, feeding, or reproductive decisions (Werkheiser, 2018).

Real-time systems have been developed for livestock monitoring (Scholten et al., 2013). These monitoring systems enable to collect information without the stress of animal disturbance or handling (Hamilton et al., 2004). The goal of these technical tools is not to replace the farmers but to support them for better decision making. PLF provides unlimited observation time, because computers can track the animals in a row (Wolfert et al., 2017).

The main purpose is to attain a full picture of animal status and behavior on a continuous basis, focusing on animal health and performance (Cangar et al., 2008). Precision Livestock Farming includes measurement, prediction and data analysis of livestock, also offering new possibilities for continuous, automatic collection and analysis of data (Berckmans, 2004).

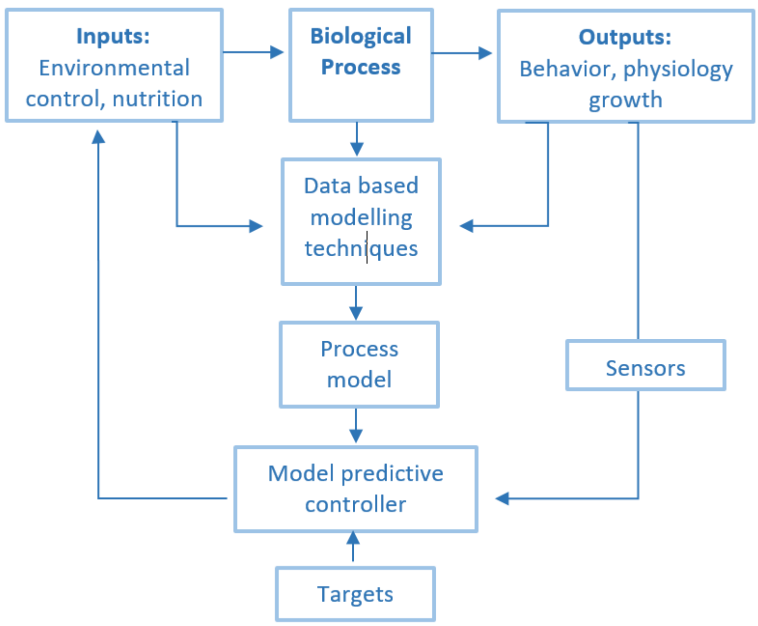


Figure 1. Precision livestock model (Berckmans, 2004).

PLF must meet the needs of both the farmer and the consumer to be commercially viable. For the farmer, increased profitability with minimal unfavorable environmental impact and high concern of animal welfare, while for the consumer, the food must be safe, nutritious and affordable (Mollo et al., 2010).

2.1.1 PLF in the poultry sector

Poultry farming, and in particular broiler farming, is an important sector due to the large quantities of animals involved which have great potential for improvements in their welfare (Rowe and Dawkins, 2019).

Broilers are the world's most numerous bird, with a standing population of 22.7 billion, an order of magnitude greater than the standing stocks of any other farmed species (Bennett et al., 2018). The highest farm animal numbers are found in poultry operations, with up to tens of thousands of individuals in one barn (Rowe and Dawkins, 2019). Modern broilers suffer from problems such as sudden death syndrome, ascites, lameness and contact dermatitis as a result of their fast growth rate (Bessei, 2006). The highest farm animal numbers are found were in poultry, with up to thousands of individuals in one granary (Wilhelmsson et al., 2019). Each individual animal is worth comparatively little and the turnover is very fast, therefore the concern for the welfare of an individual may be low (Rowe and Dawkins, 2019).

Additionally the profit margin for poultry farmers is small, creating further conflict between production and bird welfare (Khuda, 2007).

Along with the growing human population and increasing urbanization, poultry consumption is predicted to increase (Scholten et al., 2013). Poultry farming is increasing in low income countries which are not aware of animal welfare (Vaarst and Alrøe, 2012). Moreover, in intensive poultry production, there are factors, such as stocking density, environmental deterioration, unsuitable social environments or thermal stress, which can cause welfare aggravation (Meluzzi and Sirri, 2009). Thus, poultry welfare is an important area to focus efforts on improving livestock welfare. This can be achieved by continuous monitoring and tracking of individual hens, benefiting their welfare (Rowe and Dawkins, 2019). Precision livestock farming (PLF), is based on collecting data from diverse sources from smart sensors, which are then analyzed to create an automatic management system based on real-time monitoring to control animal performance, health, and welfare (Hendrawan, 2005).

The poultry industry is divided into two separate sections - meat production and egg production (Corkery et al., 2013). The environmental conditions in poultry houses influence the wellbeing and health of production staff as well as the individual bird (Marchewka et al., 2013). Poultry that are not healthy, will not provide optimal performance (Corkery et al., 2013). Techniques developed in Precision Livestock Farming (PLF) should be applied to automatically manage commercial poultry farm (Mollo et al., 2010).

Poultry monitoring systems have great potential to advance poultry production. These systems can log real-time data, and become an essential predictive tool within the poultry community (Corkery et al., 2013; Mollo et al., 2010). By using sensor technologies, potential applications to improve poultry welfare were investigate, such as - Feed intake measurements, Thermal comfort estimation, Stress detection, Assessing locomotion deficiency in broilers and Indoor climatic conditions' assessment (Sassi et al., 2016). Monitoring and inspection is done with a variety of sensors and cameras, such as spectral imaging to inspect poultry carcasses (Park and Chen, 2000), hyperspectral imaging for predicting eggs freshness (Suktanarak and Teerachaichayut, 2017), 3D imaging for broilers weight estimation (Mortensen et al., 2016), and visible light imaging for hens tracking (Kashiha et al., 2015).

Although most technologies are still in the experimental phase, some are already available and can be introduced in commercial poultry farms with good results (Marchewka et al., 2013). These available technologies have huge potential to enable better poultry welfare, or to be applied for an automatic welfare assessment (Banhazi et al., 2012; Ruiz-Garcia et al., 2009; Sassi et al., 2016).

Poultry behavioral actions are categorized into events such as eating, drinking, preening, resting, and stereotyped activities directed at different targets (Meluzzi and Sirri, 2009). It is time-consuming, costly, tedious, and prone to errors, to assess those methodologies. Therefore, there is an increasing need for systems which can collect automatically event-based behavioral responses (Puma et al., 2001).

2.1.2 PLF in the sheep sector

New PLF systems are constantly being developed for extensive and pasture based farming systems (Lima et al., 2018). The development of technologies for grazing animals is of particular interest for the sheep farming sector, since it could bring benefits for animal performance, economical performance and labour (Morgan-Davies et al., 2018). Moreover, there has been an increase in average herd size for several years, which reduces the time that farmers can spend on individual observation of their animals throughout the sheep production cycle (Villeneuve et al., 2019). This increase leads to a real need to improve the performance control that allows farmer to better control their herds (Wishart, 2019).

However, adoption of PLF technologies does not take place immediately in the sheep sector, compared to other sectors (Villeneuve et al., 2019), as sheep farmers usually belong to more conservative technology consumers (Kaler and Green, 2013). As such, investments, innovation and failure are known to all the community members in a short time which create a social barrier and risk eversion (Villeneuve et al., 2019). All the listed characteristics have negative influence regarding the openness of this farming sector to innovation (Kaler and Green, 2013).

In addition, PLF approaches have been successfully applied to intensive systems, but there are limited examples of application to sheep systems (Morgan-Davies et al., 2018). The focus in the field of sheep is small because unlike other livestock animals, animal care is less frequent, for sample, dairy cattle are treated at least twice a day for milking (Tullo et al., 2017). Application of PLF has the potential to improve sheep farming and is an important area of research that, to date, has received limited exploration (Morgan-Davies et al., 2018).

Nevertheless, trends such as global sheep economics, awareness to animal welfare and agricultural policies, influence the sheep farming practices and stimulate wider adoption of PLF systems (Morgan-davies et al., 2018), such as – individual monitoring of sheep (Salama et al., 2019), pain estimation (Mahmoud et al., 2018), lamb growth monitoring (Zhang et al., 2018) and selective breeding using measures such as Estimated Breeding Values (Conington et al., 2006).

There is also a range of real-time monitoring sensors for sheep being developed to measure location, movement, heart rate, chewing, estrus, urine, contact, respiration and temperature (Fogarty et al., 2018). The greatest advantage of such real-time monitoring technology is the potential to provide early warning systems for when measures deviate from the expected (Fuchs et al., 2019). These sensors could also be accompanied by location technology (GPS) so the animals can be found (Fogarty et al., 2018).

Such sheep systems have important roles for environmental management, and production of lamb meat and breeding animals (Umstatter et al., 2013). However, they face difficulties including: low productivity, poor economic viability, labour availability and capability, and ensuring good animal welfare (Lima et al., 2018). PLF is one such approach to overcome these difficulties (Morgan-Davies et al., 2018; Wishart, 2019).

2.2 Computer Vision

2.2.1 Computer Vision background

Computer vision is a scientific field that focus on how to gain high-level understanding from digital images or videos by using computers (Huang, 1997). Computer vision aims to solve computational models of the human visual system and to build autonomous systems which could perform some of the tasks which the human visual system can perform (Huang, 1997).

Computer vision began in the late 1960s, when researchers from universities tried to mimic the human visual system (Szeliski, 2010). Computer vision includes three main stages (Floyd and Sabins, 1987; Morris, 2004):

- Image acquisition - capturing an image using sensors that use pixel values that correspond to light intensity in one or several spectral group captured (Floyd and Sabins, 1987).

- Image processing and analyzing - transforming raw data and understanding the image data (Floyd and Sabins, 1987). These techniques deal with feature extraction, extraction of regions that differ in properties such as intensity, color, texture, or any other image statistics (Morris, 2004). By combining features together, the machine vision algorithm defines an object in the image (Groover, 2007).
- Image interpretation - converting the image into meaningful information for a wide range of users. One popular task of interpretation is recognizing the type of the objects in the image by comparing the extracted feature from the previous stage to predefined models or standard values (Klette, 2014).

2.2.2 Computer vision in agriculture

There are many computer vision applications in agriculture (Liakos et al., 2018) yielding improved automation of tasks (Figure 2) such as; (a) crop management (Kamilaris and Prenafeta-Boldú, 2018) including applications on yield prediction, disease detection, weed detection crop quality and species recognition (Ali et al., 2017; Kung et al., 2016) (b) livestock management, including applications on animal welfare and livestock production (Hansen et al., 2018; Qiao et al., 2019), (c) water management (Mehdizadeh et al., 2017) and (d) soil management (Morellos et al., 2016).

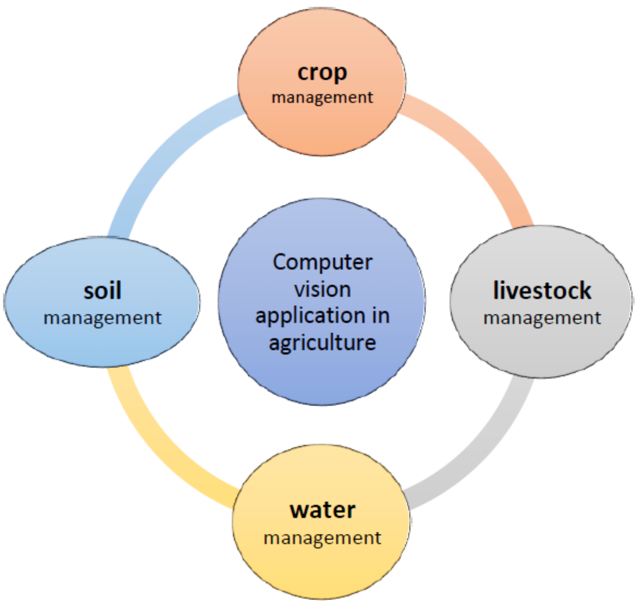


Figure 2. Computer vision applications in agriculture (Liakos et al., 2018)

Despite many years of research of computer vision in agricultural environments, there are still many problems that hinder implementation of agricultural applications (Gongal et al., 2015). The highly variable and unstructured outdoor environment with changing illumination conditions and obstructions (Kapach et al., 2012), along with the complex plant structure and variable product shape, size, color, texture and location make it hard to find a global solution to the detection of objects in the complex agricultural environment (Gongal et al., 2015; Liakos et al., 2018).

In recent years, new approaches of computer vision have emerged, based on machine learning algorithms, such as neural networks (NN) (Rawat and Wang, 2017). These algorithms together with big data technologies and high-performance computing, create new opportunities to unravel, quantify, and understand data intensive processes in agricultural operational environments (Liakos et al., 2018). One of the most powerful implementations of the neural network is the CNN (Rawat and Wang, 2017).

2.2.3 Object detection

Object detection deals with two related problems – classification and localization (Girshick, 2015). In the classification problem one or more dominant objects are determined and labelled in an image, while in the localization problems, it is much complexes to detect since in addition to labelling the dominant objects, it also must be localized in the image (Wang, 2014). Localization is usually done by determining a bounding box around the image region that is occupied by the object and providing it coordinates. The difficulty of this task may increase if there are other objects in the image where must be labelled, or if multiple objects of the same category can appear in one image (Girshick, 2015).

Using bounding boxes while classifying each box over an image is a simple approach for object detection (Girshick, 2015). However, this approach has two main drawbacks – it is expensive due to the huge search space, and it cannot be used if the number of bounding boxes is unknown (Girshick, 2015). An early method that implements the sliding window approach is the Viola-Jones detector (Viola et al., 2001). This approach includes three key contributions. The first is the use of a new image representation that allows the features to be computed very quickly classifiers (Freund and Schapire, 1995). The second is a learning algorithm, based on AdaBoost, which yields extremely efficient classifiers (Freund and Schapire, 1995). The third is a method for combining more complex classifiers which allows background regions of the image to be quickly discarded while focusing on promising object-like regions classifiers (Freund and Schapire, 1995).

The object detector is arranged in stages with increasing complexity (Viola et al., 2001). In each stage, the detector decides whether the current windows are not an object (Viola et al., 2001). If a stage decides that the current windows are not an object, the rest of the stages are not evaluated. Only true object windows trigger the entire stages (Viola et al., 2001).

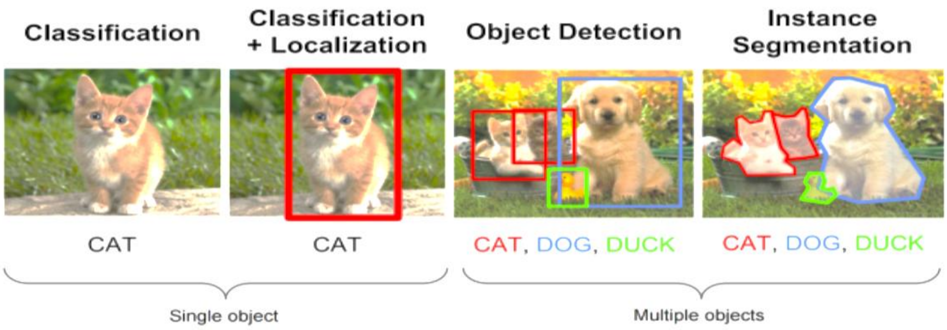


Figure 3. Object detection related problems (Girshick, 2015).

2.2.3.1 Object detection in fruit and flora

Deep learning using convolution neural networks is taking a significant part in object detection (Table 1) (Chen et al., 2017; Sa et al., 2016). One direction is yield estimation with latest works resulting in advanced performances in different scenarios, such as apple orchards and mango orchards (Koirala et al., 2019). In most research fruit and flora estimation can be classified as a generic object counting problem which can be solved either indirectly by using object detectors (Bargoti and Underwood, 2017) or with architectures that set up a regression problem to directly infer the number of object instances in the image (Rahnemoonfar and Sheppard, 2017).

2.2.3.2 Object detection in livestock

Object detection image processing techniques, have been applied to livestock farming (Table 2), for automatic recording of activity, movement, and interactions of animals (Noldus et al., 2001) to determine livestock weight estimation, lameness detection and identification (Tsai and Huang, 2014). Using object detection, it is able to carry out patterns of quantitative measurements of the animals' observed behavior (Noldus et al., 2001) and behaviors that occur over many hours, like diurnal variation in behavior (Olivo and Thompson, 1988) in a trustworthy manner (Noldus et al., 2001).

2.3 Laying hens counting

Laying hens have been bred for their high egg yield, therefore may lay over 300 eggs per year (Cooper and Appleby, 1996). Most laying hens are housed in conventional laying cages, often called battery cages, with automated control of light, temperature, feed, water, egg collection and faces removal (Bowler, 1994). It is acknowledged that battery cages cause welfare problems. Farm animals should have freedom to stand up, lie down, turn around, groom themselves, and stretch their limbs, which battery cages mostly contravene (Appleby, 2003). Therefore, in recent years, laying hen's cages have been criticized on the lack of possibilities for hens to perform natural behavior (Appleby, 2003).

Nowadays all new technical systems and equipment for the housing of laying hens have to be approved with regard to animal health and welfare before becoming available commercially (Appleby, 2003). In order to obtain approval for the use of a new technical systems, health must be at least as good as in existing systems and hens must be able to exhibit natural behavior (Weeks and Nicol, 2006). Furthermore, incidence of disease must be lower and environmental working conditions must not be worse than in older existing systems (Weeks and Nicol, 2006). As a result, quota policies of poultry raising in battery cages durable around the world (Bouvarel, 2011). In addition, quota limitation is necessary also in order to retain the powers for the regulation of production (Alston, 1999).

Counting hens become a necessary task for quota regulation (Geffen et al., 2019). Stock people who work in modern cages, are responsible for the counting task, which takes about half-day per an average poultry farm (Cronin et al., 2008). Over and above their daily operational tasks such as, supervise the mechanical egg collection belts looking for potential blockages that may result in cracked or broken eggs during the collection process (Savory, 2004). Another daily task for the stockperson is to conduct a welfare inspection of the hens in every cage. In multi-tier cages, stock people require special equipment to assist the inspection of hens in the upper tiers located above head height. Thus, monitoring upper tiers may require extra stockperson time for inspection of hens (Cronin et al., 2008). Moreover, the detection and removal of dead hens from upper level cages and other foreign objects on the upper level is an important task since failure to adequately monitor upper-tier cages can prejudice the inefficiency of production, if the dead hens gravitate onto the egg collection belt and result in egg belt blockage (Cronin et al., 2008). The detection and removal of dead require additional stockperson time compared with inspectorate of the lower cages. In order to achieve profitable management, these tasks are essential in modern multi-tier cage systems. By using the counting technology, efficiency of labor use can be improved parallel to improving frequency of surveillance.(Cronin et al., 2008).

Laying hens counting in community battery cages is a challenging task: first, the hens do not stand still but constantly moving (Shimmura et al., 2007). Secondly, they stand in different distance and angles from the camera which effect the ability to contrast them from the background (Geffen et al., 2019). Finally, because of the cages structure, they can only be photographed from a front view (Geffen et al., 2019). Hence, not all hens are visible in a single image (Geffen et al., 2019). These factors do not apply in free range cages which can be counted in various methods such as background separation (Sergeant et al., 1998) or thermal imaging and pattern recognition (Zaninelli et al., 2018). In these research the camera was positioned above the hens (Sergeant et al., 1998; Zaninelli et al., 2018).

Counting hens in battery cages, with eight hens in each cage, was achieved using a camera placed on the automatic feeder that moved at a constant rate along the cages row (Cronin et al., 2008). The hen's legs were detected by automatic detection algorithms that was developed for the research. This method obtained only 79% accuracy, the main difficulties being hens lifting one leg, sitting, or being occluded behind another hen (Cronin et al., 2008). Another research counted hens by using deep learning methods (Geffen et al., 2019). The cages were video recorded while the feeder traveled along them. In this study, a tracking algorithm that tracks the hen's head locations, and assigns every hen with a unique ID was used, leading to 98% accuracy result. However, the algorithms were tailored designed to the specific conditions and were not able to adjust to varying lighting conditions (Geffen et al., 2019). Moreover, it does not apply on cages without separation in the middle, as is customary in the industry. In these cages counting must refer to depth of photography (Geffen et al., 2019).

Table 1. Fruit and flora object detection research (Kalantar, 2019)

Object	Sensors	Train Images	Test Images	Algorithm	Results	Ref
Sweet pepper	Multi-Spectral, RGB cameras, the JAI AD 130GE and Microsoft Kinect 2.	100	22	Faster R-CNN	0.838 F1 score	(Sa et al., 2016)
Almond	Handheld Canon EOS60D	385	100	Faster R-CNN VGG16 NET	0.775 F1 score	(Bargoti and Underwood, 2017)
Tomato	Synthetic generated images	24,000	2500	Inception- ResNet	0.9103 Accuracy	(Rahnemoonfar and Sheppard, 2017)
Orange	Bluefox USB 2 camera at 10 Hz	36	35	FCN, CNN andregression	0.91 Ratio counted	(Chen et al., 2017)
Apple	PointGrey USB 3 camera at 6 Hz	11	10	FCN, CNN and regression	0.97 Ratio counted	(Chen et al., 2017)
Strawberry	RGB camera	3640	910	SSD	0.842 AP	(Lamb and Chuah, 2018)
Green citrus	RGB camera	1200	300	Faster R-CNN	0.855 Map	(XIONG et al., 2018)
Weed , soil and maize crop	Simulated images	6744	1686	VGGNet-16	0.94 Accuracy	(Dyrmann et al., 2016)
Wheat plantsroot	Nikon D5100 DSLR camera	2,500	20	Author defined CNN	0.984 Accuracy	(Pound et al., 2017)
Mango	Spectrum camera (RGB)	11,096	1500	MangoNet based on CNN	0.84 F1 score	(Kestur et al., 2019)
Mango	Prosilica GT3300c +strobe lightning	1154	270	SSD based on VGG	0.91 F1 score	(Liang et al., 2018)

Table 2. Livestock object detection research

Object	Research subject	sensor	Test data	Algorithm	Results	Ref
Cattle	Autonomously visual identification of cattle	Kinect 2 sensor & RGBD camera	86,000	ASIFT algorithm	0.97 Accuracy	(Andrew and Campbell, 2017)
Cattle	Automated detection of Mounting	Side-view camera	90	Gaussian Mixture Model & Motion History Image (MHI).	0.999 Accuracy	(Chung et al., 2015)
Cattle	Counting cattle in large area of livestock.	Unmanned Aerial Vehicles (UAVs) cameras	2,704	CNN algorithm	0.95 Accuracy	(Omatu et al., 2014)
Cattle	Autonomously find and visually identify by coat pattern individual cattle in freely moving herds.	M100 UAV with Onboard Deep Inference	1,039	YOLOv2 CNN	0.919 Accuracy	(Andrew et al., 2019)
Hens	Tracking and maintaining identities of individual hens	3D vision camera - Cambube3, PMDTec	600	Fast watershed algorithm	0.95 Accuracy	(Nakarmi et al., 2014)
Hens	Classify the laying hens' behavior to achieve automatic recognition.	Video camera LC5505E7-C83R)	778	HSVM Tracker algorithm	0.75 Accuracy	(Wang et al., 2016)
Sheep	recognizing individual sheep	Mobile camera	10,400	Bayesian optimization function was used to determine the CNN	0.98 Accuracy	(Salama et al., 2019)
Sheep	Face recognition as a biometric identifier of sheep	PowerShot G3, Canon	150	ICA algorithm & InfoMax	0.953 Accuracy	(Corkery et al., 2007)
Pigs	Detecting sow drinking, urination, and mounting behaviors	Infrared network camera	573	SBDA-DL detection algorithm	0.934 Accuracy	(Zhang et al., 2019)
Pigs	Automatic Individual Detection and Tracking	2D video camera	4,200	SSD algorithm	0.9474 Accuracy	(Zhang et al., 2019)

2.4 Biometric Identification

2.4.1 Background

Biometric identification techniques are techniques that can be used to identify an object's identity based on their unique features (Jaiswal, 2011). Biometric features can be physiological (Guo and Zhang, 2019), which are features possessed by person or animal, such as finger-prints, palm-prints, facial features, ears, irises and retinas (Minaee et al., 2019), or behavioral (Minaee et al., 2019), which are apparent in a person's or animal's interaction with the environment, such as signatures, gaits, and speech (Guo and Zhang, 2019). The aforementioned techniques have different attributes, and thus they are preferred in different types of applications (Wada et al., 2013). For example, facial recognition is commonly used for crime prevention, verification of a person's identity, information security, and access control (Guo and Zhang, 2019).

Broadly, biometric identification systems can be divided into two main types: unimodal and multimodal biometric systems (Al-Waisy et al., 2018). Unimodal systems are based on using a single source of information (e.g., right iris, left iris, or face) to establish an object identity (Wada et al., 2013). Multimodal systems combine evidence from multiple sources of information to identify an object identity (Al-Waisy et al., 2015). Considerable attention has been paid to multimodal systems due to their ability to achieve better performance compared to unimodal systems (Al-Waisy et al., 2018). In general, designing and implementing a multimodal biometric system is a challenging task since several factors that have a great influence on the overall performance must be addressed (Jaiswal, 2011), including the cost, resources of biometric traits, accuracy, and fusion strategy employed (Al-Waisy et al., 2018). However, the most fundamental issue for the designer of the multimodal system is choosing the most powerful biometric traits from multiple sources in the system, and finding an efficient method to fuse them (Wada et al., 2013).

Traditionally, the biometric recognition process involved several key steps (Figure 4). First, image data are acquired via (various) camera or optical sensors (Rokkones, 2018), and are then pre-processed so as to make the algorithm work on as much useful data as possible (Al-Waisy et al., 2018). Then, features are extracted from each image (Rokkones, 2018) and then fed into a classifier to perform recognition (Zhao et al., 2017).

Many challenges arise in a traditional biometric recognition task (Minaee et al., 2019). For example, the hand-crafted features that are suitable for one biometric, will not necessarily perform well on others (Minaee et al., 2019). Therefore, it would take a great number of experiments to find and choose the most efficient set of hand-crafted features for a certain biometric (Minaee et al., 2019).

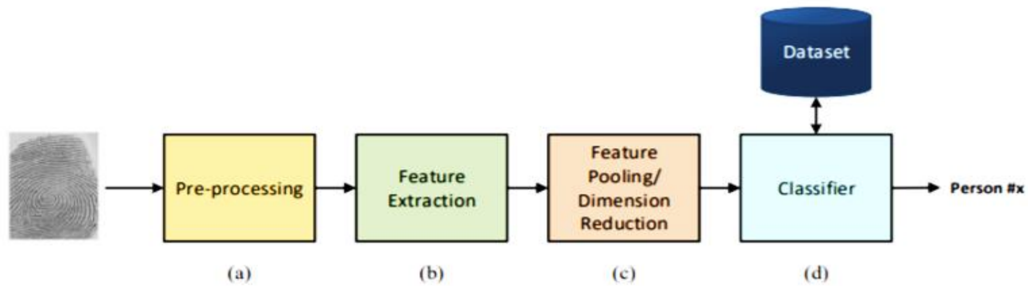


Figure 4. The block-diagram of most of classical biometric identification algorithms (Minaee et al., 2019).

However, a paradigm shift started to occur in 2012, when a deep learning-based model, AlexNet, won the ImageNet competition by a large margin (Guo et al., 2016). Since then, deep learning models have been applied to a wide range of problems in computer vision achieving promising results (Druzhkov and Kustikova, 2016). Not surprisingly, biometric identification methods were not an exception, and were taken over by deep learning models (Zhao et al., 2017). Deep learning based models provide an end-to-end learning framework, which can jointly learn the feature representation while performing classification (Minaee et al., 2019). In particular, due to the convolutional neural networks (CNNs), there have been significant advance in biometric identification technology (Zhao et al., 2017) in both human and livestock identification tasks (Corkery et al., 2007; Parmar and Mehta, 2014).

2.4.2 Biometric identification in humans

Face is the most common characteristic used for human identification tasks (Guo and Zhang, 2019). Human face is often used for identification since it contains a lot of information, but any area which might provide enough information is applicable (Schroff et al., 2015). However, its susceptibility to change due to factors such as expression, position or aging, may present a challenge in computer vision and image understanding (Deng et al., 2019).

Various studies in the human identity authentication field have applied in enormous areas such as finance, military, public security (Yi et al., 2014). The studies (Table 3) were performed on different types of human face datasets (Guo and Zhang, 2019), and used various techniques that combine convolution neural network algorithms (Deng et al., 2019; Ranjan et al., 2017), and supervised by classification loss functions (Deng et al., 2017a; Weiyang, 2017), metric learning loss functions (Schroff et al., 2015) or both (Qi and Su, 2018). Metric learning loss functions such as contrastive loss (Qi and Su, 2018) or triplet loss (Schroff et al., 2015) usually require carefully designed sample mining strategies and the final performance is very sensitive to these strategies (Guo and Zhang, 2019), so increasingly more researchers shift their attentions to building deep face verification models based on improved classification loss functions (Deng et al., 2019; Weiyang, 2017).

2.4.3 Biometric identification in livestock

Biometric identification is an emerging research field that due to convolution neural network techniques, has received increasing interest in livestock farming (Table 4) (Kumar et al., 2018; Schilling et al., 2019; Tharwat et al., 2014).

The need for on farm identification of individual animals has become more pressing in recent years as sustainable intensification has become commonplace (Hansen et al., 2018), and the ability to monitor inputs to, and outputs of each animal is increasingly desired (Schilling et al., 2019). By representing and detecting the visual appearances of animal based on generic features and primary biometric characteristics (Kumar and Singh, 2017; Schilling et al., 2019), it is possible to identify individual animal without the use of traditional methods such as ear tags, tattos etc (Andrew et al., 2019; Salama et al., 2019).

Various biometric features for the identification of animals have been investiged, including - muzzle pattern matching (Kumar and Singh, 2017), coat pattern (Andrew and Campbell, 2017), mammary glands (Schilling et al., 2019) and facial recognition (Salama et al., 2019; Wada et al., 2013). Although the number of animals in those studies are relatively small and the images have not been taken over a long enough period to see any large changes in the animal (Hansen et al., 2018), the presented researches have shown extremely accurate recognition performance is possible (Schilling et al., 2019).

2.5 Sheep biometric identification

Performance recording of sheep enable automated data collection that provides better quality data, contributing to better decision-making and thus improved management (Ait-Saidi et al., 2014). Collected data, including individual sheep ID and the recorded performances like body condition score, milk yield, data recording at lambing and body weight (Salama et al., 2019), facilitate animal handling, contributing to improved husbandry practices and reduced labor requirements (Morris et al., 2012). Moreover, it is important to track individuals with disease for treatment and for disease management, especially if there is an epidemic disease (Salama et al., 2019). In addition, buyers sometimes keep their sheep on the farm for some time; as a result, they have no guarantee of which animal they have bought (Koik and Ibrahim, 2012). For those reasons, sheep ID should be unique and permanent for an adequate performance recording (Ait-Saidi et al., 2014) and for providing buyers, sellers and farmers an efficient way to recognize and track each individual in a large group of sheep (Salama et al., 2019).

Different methods of marking sheep were used by herders (Landais, 2001). Historically, the main methods used for sheep identification were; branding by fire or freezing (Landais, 2001), ear marking by notching, tattooing and ear tagging (Caja et al., 2004), and electronic identification such as RFID tags and barcodes (Caja et al., 2004). Nowadays, branding and tattooing animals is forbidden in countries with advanced animal welfare laws (Defra, 2013), and therefore ear tags is the main methods usually used for sheep identification. However, this method has proved inefficient, since they can be either lost or their numbers can be obscured due to the environments in which sheep live (Koik and Ibrahim, 2012). In addition, this process harms the animal and may even affect its behavior (Caja et al., 2004). A further key drawback of ear tags is that they require visual detection and must be recorded manually, which can easily introduce human errors, while the labor cost of such a practice is also high (Trevarthen, 2007). Therefore, the demand for the use of electronic identification systems, which providing real savings for farmers primarily due to a reduction of labor costs, increases (Vlad et al., 2012). Electronic identification systems also offers to farmers a way to guarantee the traceability throughout the feed-animal-food chain, and the ability to better manage individual production and feeding of each sheep (Trevarthen, 2007). From all electronic identification technologies, Radio Frequency Identification (RFID) is the most common used since it provides many advantages over the others (Vlad et al., 2012). These advantages include the ability to store more information ensure successful reading of information and lack of stress to sheep (Trevarthen, 2007). In particular, they provide easier use, under field conditions, since there is no need to have visual contact of the tag – they simply must enter the scanning field of the reader (Voulodimos et al., 2010). This therefore dramatically increases ease of use, as well as providing greater reliability in light of general wear and tear, and environmental elements such as dirt and dampness (Domdouzis et al., 2007). On the other hand, one could pinpoint as the main drawbacks of electronic means of identification, especially RFID, the higher cost in comparison with the less expensive conventional methods (Trevarthen, 2007). In addition, the RFID tags are vulnerable to compromise, which requires knowledge of the technology and careful alignment in order to prevent signal damage (Wójcik and Sikora, 2017). Moreover, the small but not inexistent risk of a tag remaining in the food products (injectable tags) and the inability to protect from possible fraud (Voulodimos et al., 2010) are referred to as possible disadvantages of the RFID-based tracking methods (Trevarthen, 2007) .

To increase profitability with minimal unfavorable environmental impact and high concern of animal welfare nowadays (Mollo et al., 2010), using biometric traits instead of traditional identification methods, has gained a lot of attention in current livestock identification systems (Corkery et al., 2007). Only two research has focused on sheep identification so far (Corkery et al., 2007; Salama et al., 2019).

Sheep facial biometrics include many significant features that can be used for identification such as muscles, the eyes, mouth and many hidden features (Corkery et al., 2007). Therefore, facial biometrics are very promising and efficient features for sheep recognition (Salama et al., 2019).

Sheep face recognition was achieved using a cosine distance classifier trained on facial images of 50 sheep whose ages ranged from 3 to 4 years (Corkery et al., 2007). In this research, each sheep was represented by 7 images taken at a forward-facing posture with a black background, using Canon professional PowerShot camera (Corkery et al., 2007). The faces of the sheep were cleaned of dirt and all possible sources of noise before being imaged (Corkery et al., 2007). Sheep were also held using special tools so that a certain fraction of each sheep face was within the image (Corkery et al., 2007). This approach achieved 96% accuracy but required considerable human intervention for image acquisition(Corkery et al., 2007).

In the second research, sheep identities were recognized by a convolutional neural network using facial biometrics (Salama et al., 2019). A Bayesian optimizer was used to automatically set the parameters for the convolutional neural network and in addition, the AlexNet configuration was also examined (Salama et al., 2019). The sheep recognition algorithms were tested on a data set of 52 sheep between five months and five years old, with 10 images taken per sheep (Salama et al., 2019). Also, data augmentation methodologies such as rotation, reflection, scaling, blurring, and brightness modification were applied (Salama et al., 2019). The experiments conducted in this paper achieved an accuracy of 98% (Salama et al., 2019).

However, those researches were tailored designed to the specific conditions and were not able to adjust to varying lighting conditions and different face postures (Corkery et al., 2007; Salama et al., 2019). Moreover, they were tested on small set of sheep with great variability (Corkery et al., 2007).

Table 3. Human biometric identification researches.

Subject	Loss Function	Train Datasets	Test Datasets	Algorithm	Results	Ref
Feature learning using deep convolutional neural networks	Additive angular margin loss (ArcFace)	CASIA, VGGFace2 and MS1MV2.	LFW, and YTF	ResNet100	0.9953 accuracy	(Deng et al., 2019)
Sphereface: Deep hypersphere embedding	Angular softmax (A-softmax) loss	CASIA-WebFace	LFW and YTF	Author defined 64-layer CNN	0.9942 and 0.95 accuracy respectively	(Weiyang, 2017)
Feature learning	Congenerous cosine distance	MNIST and CIFAR-10	LFW	COCO algorithm	0.9986 accuracy	(Liu et al., 2017)
Softmax loss for discriminative face recognition	L2-Softmax loss	MS-Celeb1M	LFW and IJB-A	Face-ResNet DCNN	0.9933 accuracy	(Ranjan et al., 2017)
Unified embedding for face recognition and clustering	Triplet loss	LFW and YTF	LFW and YTF	FaceNet	0.9963 and 0.9512 accuracy respectively	(Schroff et al., 2015)
Face representation	N-pair loss	WebFace	LFW	CasiaNet	0.9833 accuracy	(Yi et al., 2014)
Marginal loss for deep face recognition	Marginal loss	MS-Celeb-1M	LFW and YTF	ResNet1	0.9948 and 0.9548 accuracy respectively	(Deng et al., 2017b)
Improving the generalization ability of DCNN	Noisy softmax	WebFace	LFW and YTF	VGG-net	0.9918 and 0.94.88 accuracy respectively	(Chen et al., 2017a)
Face representation using joint sample and set-based supervision	Max-Margin loss	VGG Face	LFW and YTF	Inception-ResNet	0.9603 and 0.9244 accuracy respectively	(Gecer et al., 2017)
Git loss for deep face recognition	Git loss	VGG Face2	LFW and YTF	Inception-ResNet	0.9930 and 0.9530 accuracy respectively	(Calefati et al., 2019)
Range loss withlLong-tailed training data	Range loss	WebFace and Celeb1M	LFW and YTF	ResNet2	0.9952 and 0.937 accuracy respectively	(Zhang et al., 2017)
Conntrastive center loss for deep neural networks	Contrastive-Center loss	WebFace	LFW	ResNet1	0.9868 accuracy	(Qi and Su, 2018)
Additive margin softmax for face verification	AM-softmax	WebFace	LFW and MegaFace	ResNet2	0.9917 and 0.8444 accuracy respectively	(Wang et al., 2018)

Table 4. Livestock biometric identification researches

Object	Method	Sensors	Train Images	Test Images	Algorithm	Results	Ref
Cattel	Identidiction via selcetive local coat pattren matching in RGBD imagery	Kinect 2 sensor	83 images	294 images	ASIFT algorithm with RBF-SVM	0.97 accuracy	(Andrew and Campbell, 2017)
Cattel	Identidiction using muzzle point pattern	20-megapixel camera	3000 images	2000 images	K-means segmentation algorithm to find ROI & Fuzzy-K-NN for classification	0.9674 accuracy	(Kumar and Singh, 2017)
Cattel	Using cow's mammary glands as a novel biometric identification modality	Go Pro camera with Near-Infrared (NIR) sensors	150 images	152 images	Scikit-learn machine learning library evaluated with Support Vector Machine (SVM)	0.60 accuracy	(Schilling et al., 2019)
Cattel	Identification via coat pattern individual cattle in freely moving herds & data augmentation was used	DJI Zenmuse X3 camera located on a drone	3120 images	1039 images	InceptionV3	0.944 accuracy	(Andrew et al., 2019)
Cattel	Identification via cattle animals using muzzle print images	Used exisiting dataset	124 images	93 images	LBP, LDA & SVN classifier	0.995 accuracy	(Tharwat et al., 2014)
Cattel	Transfer learning approch for recognition of cattle using muzzle point image pattern	30-megapixel camera	100 images	400 imgaes	Author defined DCNN	0.9899 accuracy	(Kumar et al., 2018)
Pigs	face recognition using convolutional neural networks	Sogatel USB2.0 webcam	932 images	621 images	VGG Face & Liner SVN classifier	0.967 accuracy	(Hansen et al., 2018)
Pigs	Recognition using eyes pattern	Digital camera	256 images	64 images	PCA using eigenspace	0.979 accuracy	(Wada et al., 2013)
Sheep	Identidiction using the cosine distance classifier	PowerShot G3	200 images	150 images	Independent Component Analysis (ICA) algorithm.	0.96 accuracy	(Corkery et al., 2007)
Sheep	Bayesian optimization was used to find the best CNN parameters & data augmentation was used	Mobile camera	4160 images	1040 images	Author defined CNN & AlexNet	0.98 accuracy	(Salama et al., 2019)

3 Methods

This thesis includes two machine vision applications for PLF; detecting and counting laying hens (chapter 4) and sheep biometric identification (chapter 5).

3.1 Mobile system for detecting and counting laying hens

The research goal was to develop a mobile system for counting laying hens in battery cages, which consist of 4-6 six floors, where each floor contains about 40 cages housing thousands of hens. The research included designing the mobile imaging system and developing algorithms for counting the hens. The system was designed to be mobile to suit different hen houses. The counting algorithm received as input a color video record of a cage containing about 35 hens, which was acquired from an Intel RealSense RGBD camera, and the output was the number of hens counted and their location. Two experiments were performed in which a total of 6,300 images were acquired for the algorithm development. As part of the development, two different counting algorithms were developed. The first algorithm aimed to detected and count areas with red color which is an indication of the hen's comb, while the second algorithm aimed to detect and count hens based on Faster R-CNN algorithm. Details are provided in Chapter 4.

3.2 Automated system for sheep identification

This research focused on developing an automated system for data collection and a deep learning model for sheep facial biometric identification. The automated system video recorded the sheep faces at all daylight hours, ensuring a variety of light conditions and shooting angles data, which were used for the model development. Data collection was made on two different groups of sheep and throughout their growth period, where the sheep gained about 25 kg in weight, in order to examine whether weight gain and the sheep maturation influenced the biometric identification. The first group (group 1) contained 47 sheep and the second group (group 2) contained 34 sheep. The biometric identification model developed in this thesis, took a set of different sheep faces as input and applied two steps on each image - face detection, using Faster R-CNN algorithm, and classification, using ResNet50V2 CNN and ArcFace loss function. The model is described in Chapter 5.

4 Mobile system for counting laying hens

This chapter describes the development of a mobile system (Section 4.1) for detecting and counting laying hens using color images acquired from a digital camera mounted on a mobile cart, the experiments and their evaluation (section 4.2), the development of an image processing algorithm and object detection algorithm for hen detection and counting (section 4.3) and the results of this study (section 4.4).

4.1 System design

A mobile system which can be transferred between hens houses was developed (Figure 5). The system aims to replace the manual counting of laying hens, with an algorithm that uses a RGBD camera to detect and count hens. Since the narrowest path besides cages, in all the hen's battery cages in Israel, is one-meter-wide, the system was mounted on a 60 cm wide cart. Since the system must be disinfected before each entering to the hen’s battery cage, the cart was made from aluminum, a material that is easy to disinfect. The Intel RealSense depth camera D435 camera was used for image acquisition at 30 frames per second (fps). The camera was connected to a mobile computer with a USB cable, allowing to track the data collection in real time (Appendix 1A). The camera was positioned at 90-degree angle to allow direct view of the laying hens (the hen houses structure enables only front view access). The RGBD camera was mounted on an aluminum rod which was attached to the front of the cart. Since battery cages consist of 4-6 floors up to 3.5 meters, the rod has a variable length, with maximum length of 4 meters, in order to fit it to the floor which must be photographed.

4.2 Methods

4.2.1 Experimental design

The system was examined in a single commercial hen house located in Moshav Kidron, located in central Israel. The hen house is 87 m long, on the second floor that contains 37 cages, each cage is 2.4 m long, 0.54 m tall, and 0.74 m depth, housing 18–34 hens per cage. In the hen house, there is an external feeder situated in the front of the cages.

Figure 5. The developed mobile system. The system was mounted on a 60 cm wide and 80 cm long cart. The system includes; (a) Mobile computer. (b) Aluminum rod with a variable length. (c) Intel RealSense depth camera D435, located 160 cm from the floor.

Figure 6. Battery cage stacked on six floors, with 37 community cages set in a row.

To determine the optimal time for counting the hens, the behavior of the hens was monitored in several observations to better understand their habits. Feeding events occur four times a day, at 6:30, 13:30, 14:30, and 17:00. Following the feeding routine up to 15:00, most of the hens sat in the cages and laid eggs; sitting at the back of the cages and huddling together, reducing the ability to count them. The last feeding event was selected as the adequate time point for counting, as most hens stood in front of the cages and waited for the last meal of the day.

Hens are easily frightened by any foreign object such as a person's presence or a camera. Hens do not notice the blue color. Therefore, in order to reduce panic level, the man conducting the experiment wore blue clothing and the cart was painted in blue. In the hen house, one cage consists of two cells with no partition between them (Figure 7). Since the RGB-D camera lens is not wide enough to absorb a full cage, each cell was recorded separately. The hens were video-recorded for 30 minutes during the last feeding event of the day. During the feeding event, the feeder travelled backward and forward along the cages line. Accordingly, the cart moved with and by the feeder pace (forward – the user pushed the cart; backward – the user pulled the cart). As a result, every cell was recorded at least twice, at a camera speed of 30 fps when about 120 frames were collected per cell (Appendix 1B).

Figure 7. A photo of one cage; the cage boundaries are colored in blue, while the green line indicates the middle of the cage which consists of two cells.

4.2.2 Data

Two datasets were used – one for developing the image processing algorithm (from Geffen et al., 2019), and the other for the object detection algorithm (a newly acquired one).

The dataset used for the development of the image processing algorithm included 4,440 images that were acquired in a previous research in 2019 (Geffen et al., 2019). The images were acquired with a Media Tech W9R camera with a 170-degree wide angle lens. The camera was tuned to full HD mode (1080p), filming 30 frames per second (fps). The hens were recorded while the camera was mounted on a steel rod attached to the feeder, 90 cm in front of the cages, and positioned in a way that allowed direct view of the hens. For the algorithm development, 105 frames were used consisting of the first three frames of each of the cages number 2-35 in the second floor.

The second dataset, which was used for the new developed object detection algorithm, was specially collected as part of this thesis on two different days (February 25 and 26, 2020). A total of 6,300 images were collected. Prior to recording, the number of hens in every cage was counted manually by a human observer, which was later used as the ground truth, to compare with the count obtained by the system.

In order to train and test the object detection algorithm, a labelled dataset was manually created (Appendix 1C) using “Labeling” image annotation tool. The dataset includes images in which bounding boxes around the hens with the right class tag –'chicken' marked. The object detection algorithm was trained and tested using 700-tagged images, 560 images in the training dataset and 140 images in the test dataset. The other 5600 left images were used to evaluate the whole system, which contains the detection algorithm and the tracking algorithm.

4.2.3 Algorithm

As part of the development, two different algorithms were developed (section 3.3). The first system relies on an image processing technique that detects and counts areas with red color, which signifies a hen’s red comb. The second algorithm relies on an advanced deep learning classification schema. Both systems include two main stages: hen’s recognition (detection) followed by counting and estimation.

4.2.4 Performance measures

In order to evaluate the correctness of the image processing algorithm, two errors were calculated: the first compared the algorithm counting results to the ground truth obtained using the manual counting of hen's combs (defined as E1). The second error calculation compared the algorithm's results to manual counting of hens in the corresponding images. This comparison was made in order to assess whether the hen's comb is a sufficient feature to accurately detect laying hens (define as E2).

$$\mathrm{E1} = \frac{\text{Manual comb counting} - \text{Algorithm counting result}}{\text{Manual comb counting}}$$

$$E2 = \frac{\text{Manual hen counting} - \text{Algorithm counting result}}{\text{Manual hen counting}}$$

Object detection performance was evaluated using recall and precision indicators together with the F1-score, based on the confusion matrix. This matrix is used to describe the performance of a classification model on a set of test data for which the true values are known. For object detection we used the concept of Intersection over Union (IoU) as a threshold for classification decision. IoU computes intersection over the union of the two bounding boxes; the bounding box for the ground truth and the predicted bounding box.

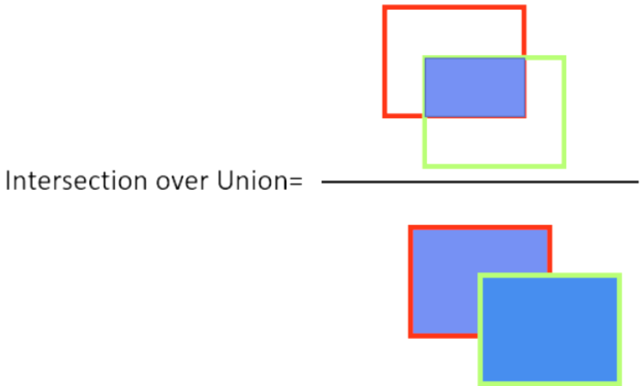


Figure 8. Red is ground truth bounding box and green is predicted bounding box.

A classification is true if it matches the ground truth with IoU > N. IoU is a way to measure if a predicted bounding box is well-located; high IoU means that a predicted bounding box has a big overlap with a ground truth bounding box. N is a number between 0 and 1, and it is a threshold for IoU. The confusion matrix reflects the resulting matches between ground-truth and detections, when the horizontal rows represent the target values (what the model should have predicted - the ground-truth) and the vertical columns represent the predicted values (what the model actually predicted). The final row and column correspond to the class "nothing" which is used to indicate when an object of from the class 'chicken' was not detected, or an object that was detected was not part of the ground-truth. The confusion matrix reports the number of true positives, false positives, false negatives and true negatives:

- True positives (TP): A hen detection is considered to be a true if the predicted and ground-truth bounding box had an intersection over union (IoU) greater than a fixed threshold.
- False positives (FP): (also denoted as false detection) refers to an algorithm's mistake in predicting background as an object - hen.
- False negatives (FN): A miss by the algorithm, refers to its failure to detect a real hen.
- True negatives (TN): This indicator is not useful for object detection, since it represents the correct detection of a background (not an object). Hence we ignore TN.

Precision indicates the fraction of the algorithm's predictions that are hens. Recall is the fraction of hens in the image that were detected by the algorithm. Increasing recall usually comes at the expense of precision. The harmonic mean so-called F1 score provides a balance between the two and was calculated.

$$\text{Precision} = \frac{TP}{TP + FP}\ , \qquad \text{Recall} = \frac{TP}{TP + FN}\ , \qquad F1 = \frac{2\text{PR}}{P + R}$$

The area under the Precision–Recall curve is called the Mean Average Precision (mAP), and it was used as the performance measure for the object detection algorithm.

4.2.5 Sensitivity analysis

Two sensitivity analysis were conducted:

1. Directional analysis: the differences between the video acquired while moving forward (the user pushed the cart) and backward (the user pulled the cart) was evaluated and compared to ground truth. The statistical F test was used to test the null hypothesis that the variances of the two videos are equal.
2. Number of frames: to determine the minimum number of frames needed for high detection while shortening runtime, the analyses evaluated the number of detections for different numbers of frames: 1, 4, 10, 20, 30, 35, 40, 44 and a full video containing more than 44 frames.

4.3 Algorithms

4.3.1 Image processing algorithm

The system receives as input a RGB image of each cell in the battery cage. The output is the number of combs counted and their location. The counting algorithm included the following steps (Figure 10): convert RGB image to HSV, thresholding, red color detection, noise removal, erosion, dilation, blob detector and counting. The image processing algorithms were programmed in Python 3, since it has access to great libraries for image processing, flexibility, platform independence, and a wide community (Appendix 1D).

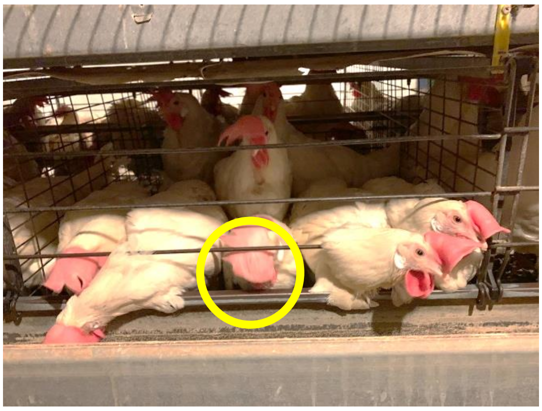


Figure 9. Hen comb: a fleshy growth or crest on the top of the head.

Convert RGB image to HSV, and split color components

The first step involved transforming the RGB color space into the HSV (hue, saturation, value) space which is a color space that describes color information in a similar way that is used by the human vision system.

Thresholding and red color detection

The histograms of the object (hen's comb) was plotted, and the spread of an object in the color space was taken as its primary identification. The threshold color values were obtained to distinguish the object from the background, in HSV color space to achieve a better separation. Red color objects have Hue value in range: from 0 to 10, as well as in range from 170 to 180. Therefore, thresholding included two conditions to ensure only hen's comb pixels remain in the images.

Noise removal

Some of the combs that were present were incorrectly recognized and formed unwanted noise. The following image processing operations were performed to remove the unwanted noise through median filtering. A median filter is more effective than others when the goal is to simultaneously reduce noise and preserve edges. The median was calculated by first sorting all of the pixel values from the surrounding neighborhood into the corresponding numerical order and then replacing the pixel that was under consideration with the middle or the median pixel value (if the neighborhood under consideration contained an even number of pixels, the average of the two middle pixel values was used).

Produce a binary image

A binary image was used in order to separate the object in the image (hen's comb) from the background. The binary image effectively masks comb regions in the image. The white color referred to the object, while the rest is referred as the background, colored by black.

Erosion and dilation

Dilation and erosion were combined to remove small objects from an image and smooth the border of large objects. Dilation adds pixels to the boundaries of objects in an image, while erosion removes pixels on object boundaries. The number of pixels added or removed from the objects defines according to the structuring element also known as a kernel. The current algorithm used 2 iterations for each operation, with kernel size of 7 on 7 to process the image.

Blob detection

A blob is a group of connected pixels in an image that share some common property. The goal of blob detection is to identify and mark those groups. The SimpleBlobDetector, which was used in the algorithm, is a simple algorithm, controlled by parameters which define the type of required blobs. Several filters can be employed including filter the blobs based on size, filter according to shape circularity, which measures how close to a circle the blob is, filter by shape convexity (defined as the Area of the Blob / Area of its convex hull) or filter shape according to inertia Ratio that measures how elongated a shape is.

Comb counting

The number of blobs in the segmented image were then counted. The blobs that were counted, were limited by both the lower and the upper limit on the number of pixels in a blob. The lower limit served to remove noises from the image. The upper limit had the effect of rejecting some of the comb where several combs had been overlapped.

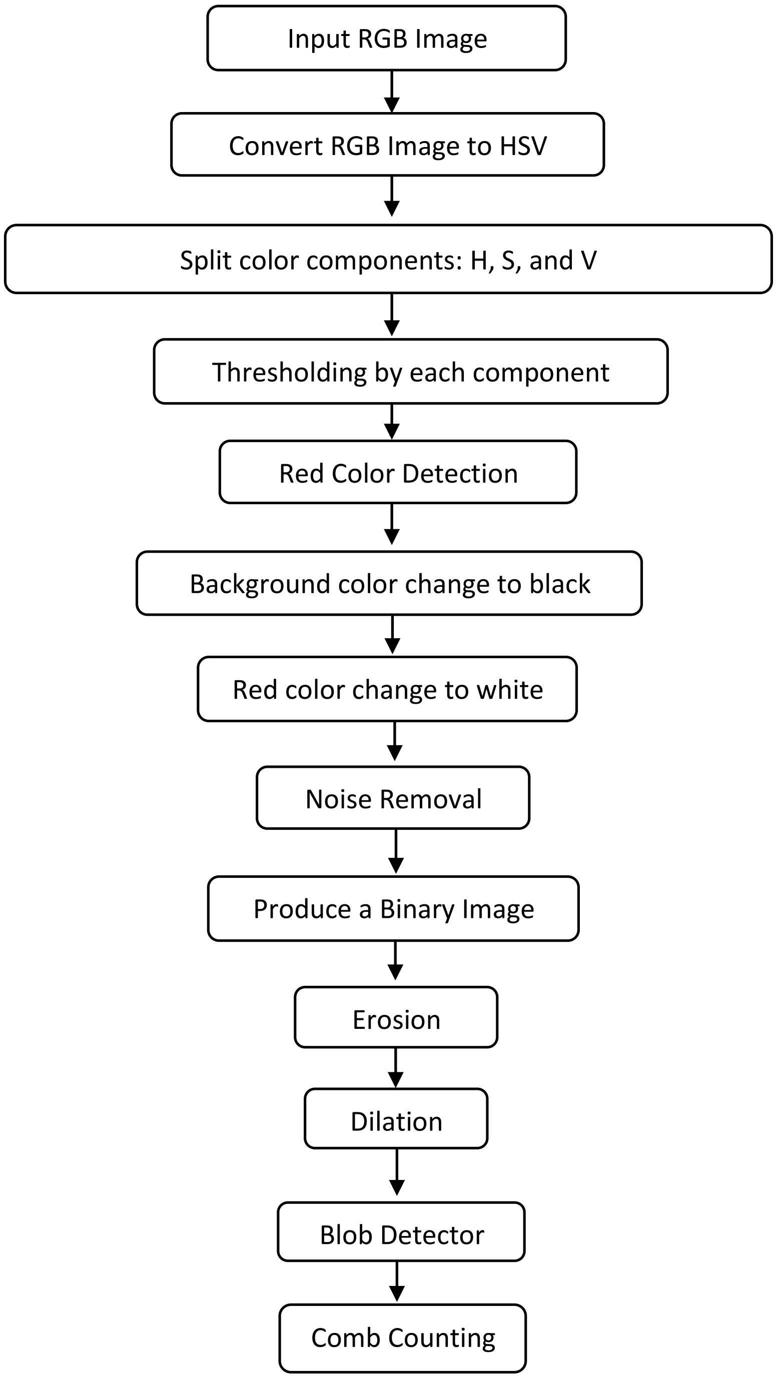
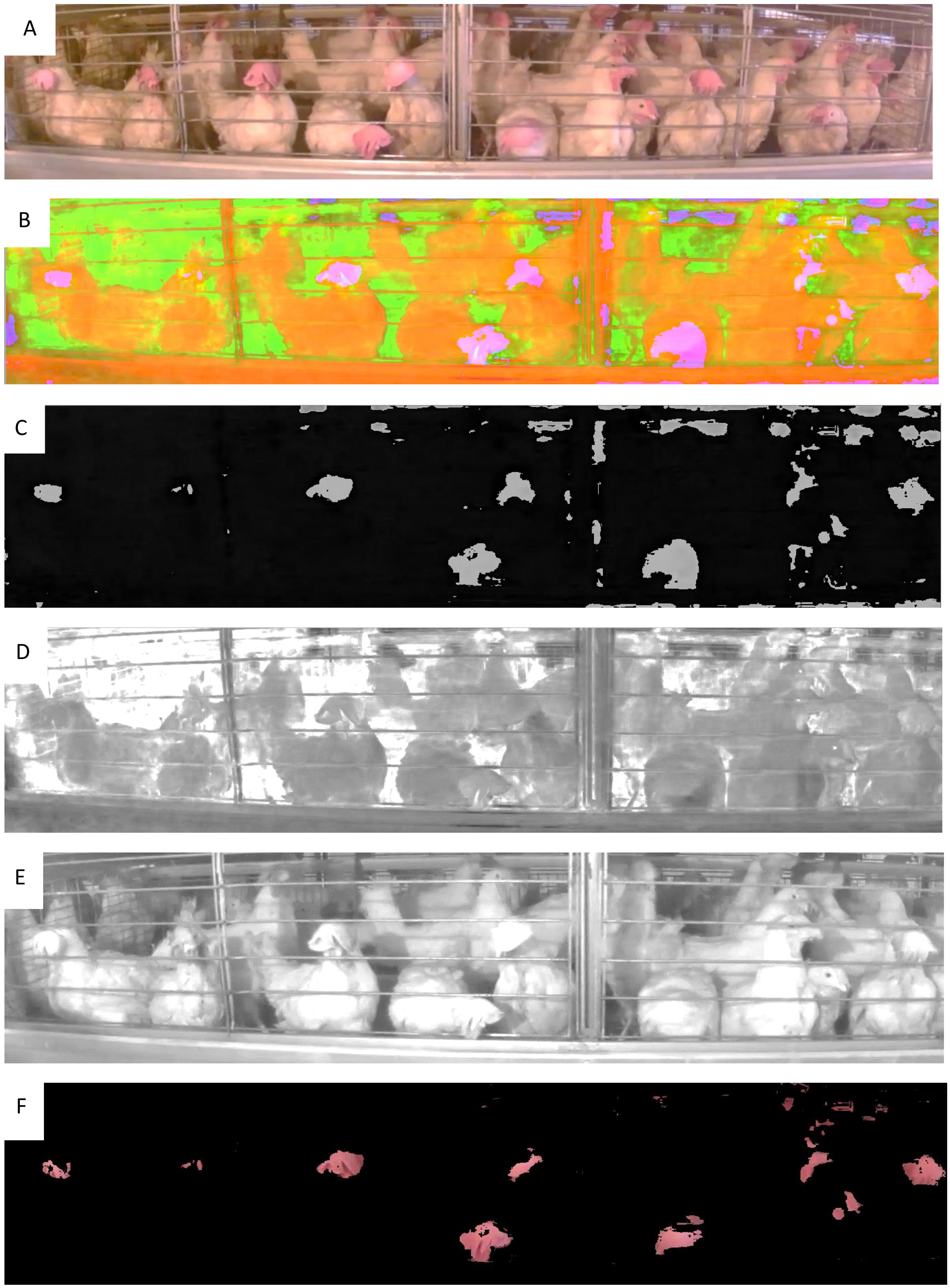


Figure 10. Flowchart algorithm of the overall methodology.


A
B
C
D
E
F

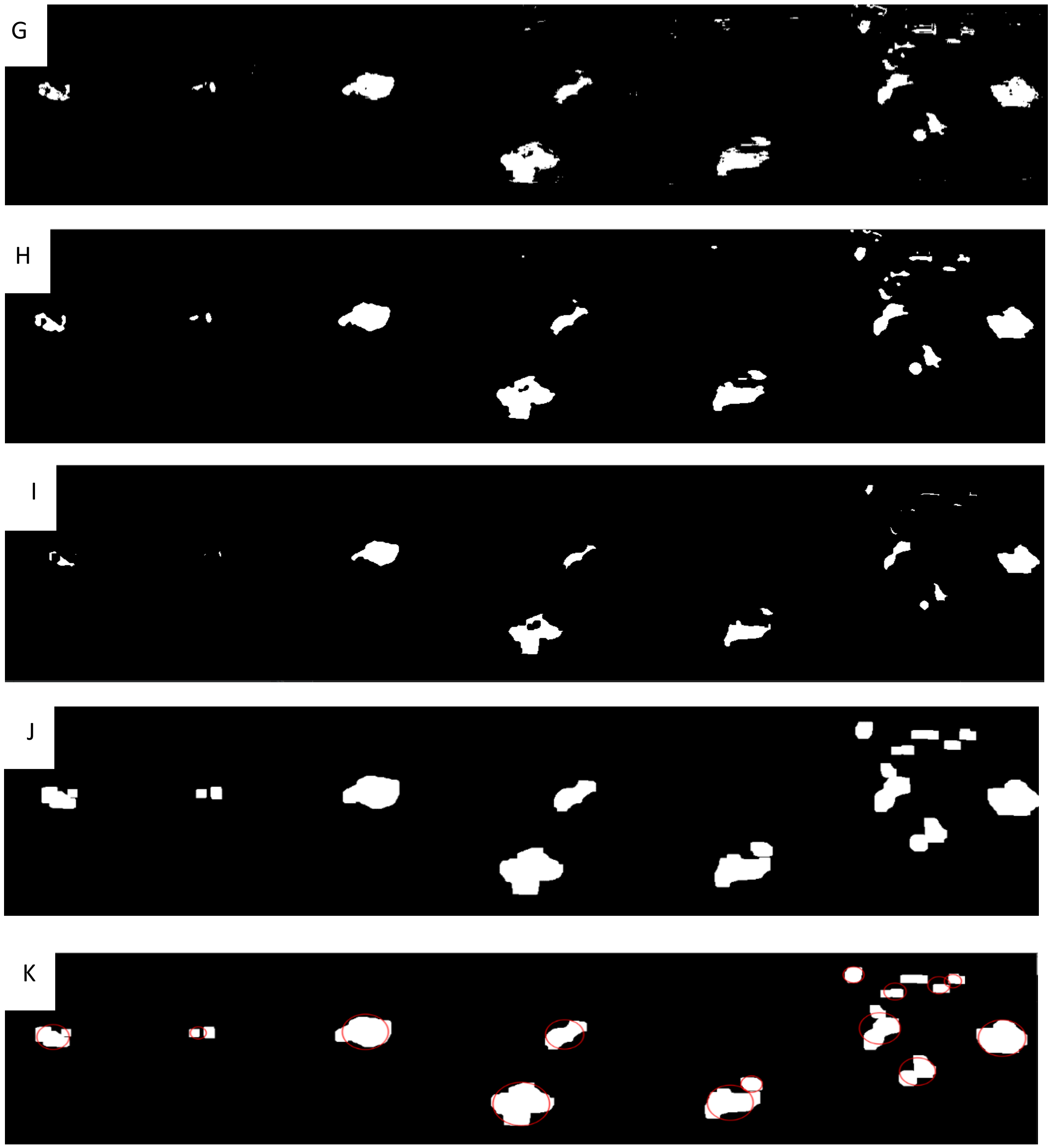


Figure 11. Counting algorithm steps; (a) Input RGB image, (b) HSV image, (c) H panel image, (d) S panel image, (e) V panel image, (f) Red Color Detection, (g) Background color change to black and red color change to white, (h) Noise Removal, (i) Erosion, (g) Dilation (k) Blob Detector.

4.3.2 Object detection counting algorithm

The counting algorithm took a two second video sequence of every cage and processed the video frames. First, the Faster R-CNN deep learning object detection algorithm was applied to detect hens in each frame. Then, a Non-Maximal Suppression (NMS) algorithm was applied, to overcome multiple detections. Next, the frames with the resulting identifications were fed into a tracking algorithm, which tracked the location of each identified hen along the video frames, based on its location in the previous frame.

In the tracking process, every hen was assigned with a unique ID, and the tracking procedure enabled each hen to continually be identified with the same ID along the frames. Newly detected hens, which had been obstructed in previous frames were assigned with new IDs, and false detections were removed. This process compensates for inaccuracies that may occur in the hen detection stage.

Faster R-CNN

Faster R-CNN is a region-based detectors object detection algorithm based on a convolutional neural network (CNN) (Ren et al., 2015) that comprise two modules – a Region Proposal Network (RPN) and a classifier. The RPN is a kind of mechanism to focus attention, directing the classifier where to look for objects in the image. It does this by identifying regions (bounding boxes) in an image, which are more likely to have objects in them. First, the image is fed into a pre-trained base network, then the RPN slides on top of the last shared convolution layer, and finds three hundred such boxes in the original image. The proposed regions are fed into the second module of the Faster R-CNN (classifier), which classifies hens and corrects the bounding boxes' locations. The output of this process is an image with bounding boxes around the objects it has found. The base network used in this research is ResNet101 (He et al., 2016) pre-trained on the Common Objects in Context (COCO) dataset. The object detection algorithm was trained and tested using 700 tagged images (section 3.4) and is described in Appendix 1E.

Non maximal suppression (NMS)

A single object might be detected multiple times, as it may fit well enough in more than one proposed bounding box. In the case of multiple detections of a single hen, only one detection that fits better will be kept (Figure 13). Multiple detections have a big overlap; thus, they were identified by calculating the Intersection over Union (IoU) index.

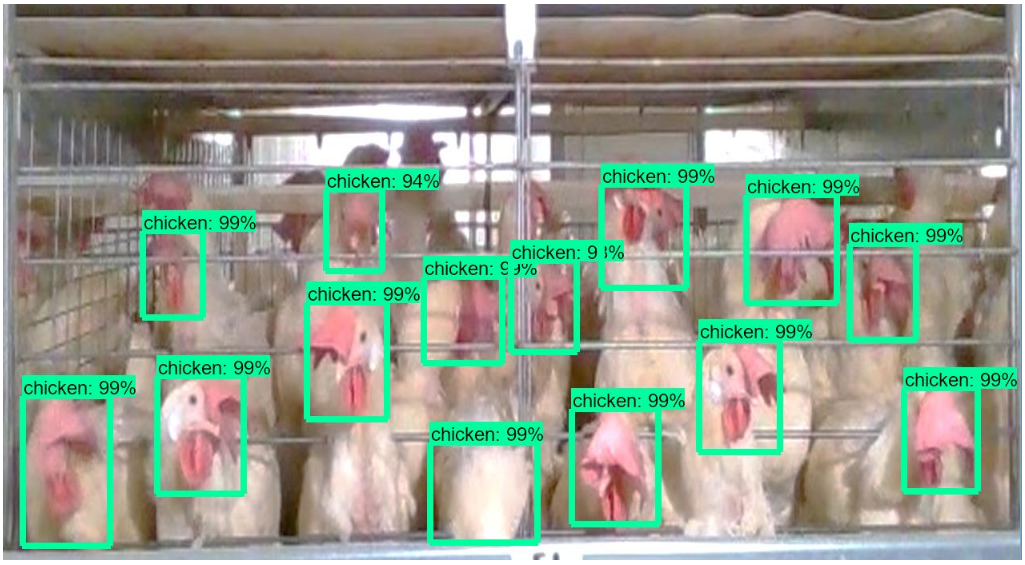


Figure 12. Test images detection result, after training Faster R-CNN algorithm.

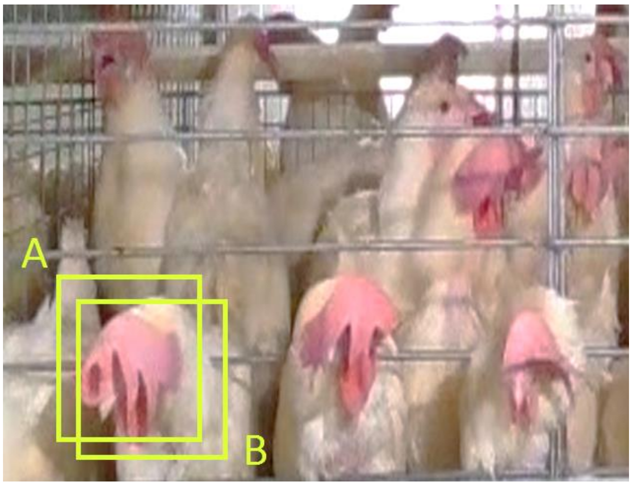


Figure 13. Bounding boxes, A and B are two different detections of the same hen in the same frame.

After hens are detected in a frame, NMS is applied. First, all the bounding boxes are put in order by decreasing size. The biggest is compared with the second biggest box, and so on down to the smallest box. Smaller boxes with an IoU greater than 0.06, which was obtained experimentally, are deleted. Next, the second biggest box left in the boxes list, is compared with all the boxes smaller than it, and so on. In this way, no bounding boxes are compared to the same bounding box more than once, because each box is compared only with boxes smaller than itself. In any case of multiple detections, the biggest bounding box is kept. That is because the biggest box usually contains a significant part of the smaller one, but does not contain multiple hens; a big box usually contains a larger part of the hen's body; a smaller box contained in the big one, will usually detect a smaller feature of the same hen, e.g. an eye or a beak. By keeping the biggest box, a single hen is not assigned multiple IDs during the tracking process.

Tracking and counting algorithm

Hens are located in different places in a cage; some stand, sit or are hidden by others. In addition, hens move very fast; even in the short time of 1/30 seconds between two successive frames, a hen may move her head down significantly. Because of these movements and obstructions, not all the hens in a cage are visible to the camera in every frame. Moreover, not all the hens visible to the camera are necessarily detected in every frame. In order to address these problems, a tracking algorithm that keeps and follows the detections from previous frames was developed (Appendix 1F), based on previous work development (Geffen et al., 2019). The tracker receives a video sequence of about two seconds of every cell (an average of 60 frames, each with the detected bounding boxes of the hens), and tracks the hens' locations, as follows:

- Each detected hen in the first frame of the sequence is assigned with a unique ID number.
- Detected hens in the next frame are assigned with IDs from the last frame, or with new IDs, as follows: Euclidian distance was calculated between every new bounding box and existing ones, the pair of bounding boxes with the minimal distance are chosen to be compared - If the IoU between those boxes is greater than 0.1, the new box is assigned with its pair's ID from the previous frame. Otherwise, the box is assigned with a new ID, which means that a new hen was detected (Figure 14).
- If a hen was detected in a previous frame, but not in the current frame, the amount of times it wasn't detected again will be counted (Figure 15).
- After all the frames in the sequence are processed, the tracker has logged all the hens' IDs and number of appearances. If an ID appeared only once, it will be deleted, as it is considered a false detection.

- The tracker calculates the maximum detection for each cage, as a sum of two cell's detection, according to a method explained later (section 3.3).

Figure 14. An example of a hen that was detected in a previous and in the current frames. The bounding box from the previous frame is on the right, the one from the current frame is on the left. The Euclidean distance between the centres of these two boxes is the

Figure 15. Tracking example; (a) First frame - each detected hen is assigned with a unique ID number, (b) Following frame - a new hen was detected - ID 16, (c) Following frame - in blue - a hen was previously detected but not in the current frame. The disappearance counter display next to the ID

4.4 Results

The results of the image processing algorithm and the object detection algorithms are reported.

4.4.1 Image processing

The image processing algorithm (section 3.2) resulted in less red combs as compared to the human manual counting (Figure 16), with an average of 6.21 missed (SD: 1.72, maximum missed: 11, minimum missed: 2). The errors E1 and E2 (Figure 17), indicate that using red comb as a single feature is not sufficient. In manual counting probably additional features such as the legs, eyes, and neck are considered.

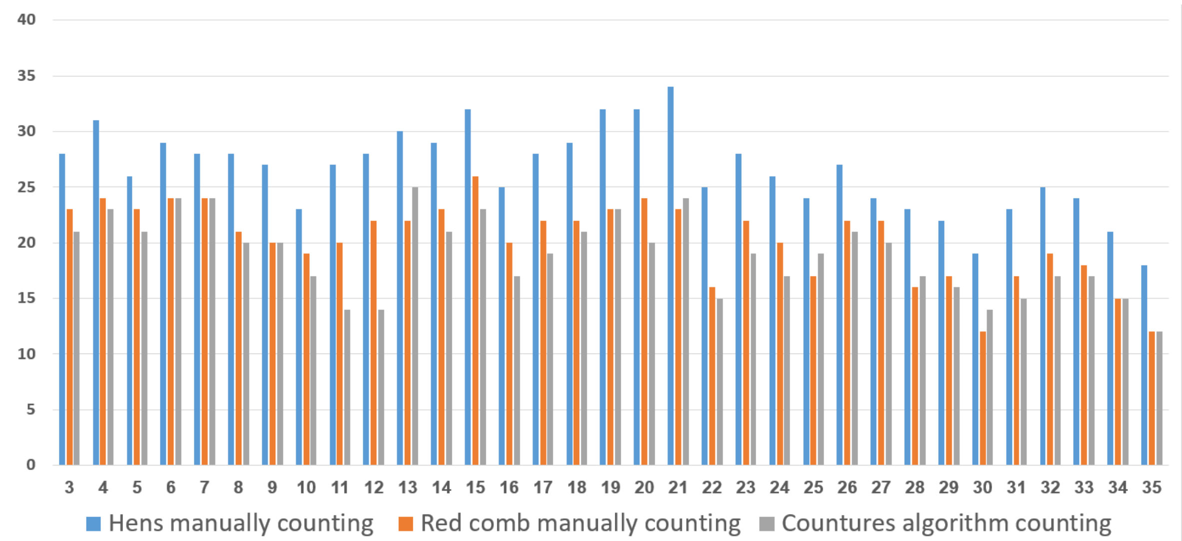


Figure 16. Cells 3-35 counting results. The cell number presented on the X-axis; the Y-axis is the number of target counting.

The E1 error is up to 40% (min error - 0.08, max error - 0.36, SD - 0.064), while the E2 error is higher than E1 (Figure 17) with errors up to 50% (min error - 0.14, max error - 0.5, SD - 0.080).

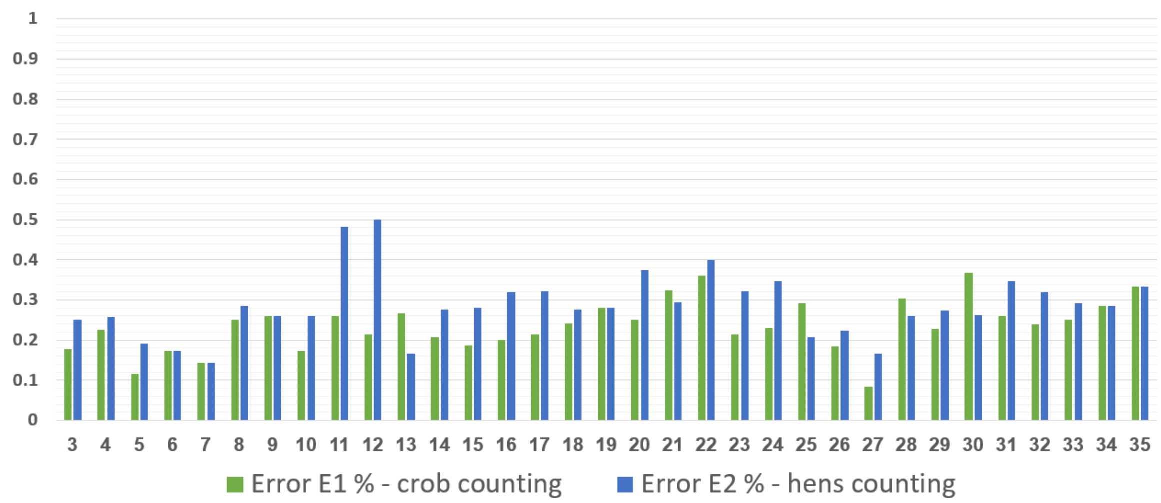


Figure 17. Cells 3-35 errors calculation - the error colored by green represent E1, and the blue one represent E2. The cell number presented on the X-axis; the Y-axis is the error calculated.

The difference between the two errors indicates that the hen's comb is not an accurate enough feature for detecting laying hens. Therefore, other features should be included in order to improve the algorithm results. In addition, additional steps must be taken in order to improve the algorithm results.

4.4.2 Object detection

Most of the results derived with an IoU of 0.5 are correct, and most of the hens were detected, according to the results of the test set which includes 141 labeled images (Table 5). When IoU=0.50, less overlap is required between the ground truth boxes and the predicted boxes compared to an IoU of 0.75. Therefore, for IoU=0.75 the mAP dropped. The drop in the mAP suggests that many of the detections are made in the range of 0.5<IoU<0.75.

Table 5. Mean Average Precision (mAP) with different Intersection over Union (IoU) values.

IoU	0.50	0.75
mPA	0.962	0.446

The confusion matrix (Table 6) for IoU=0.5, revealed 89.54% true positive.

Table 6. Confusion matrix for IoU=0.5.

	Predicted Positive	**Predicted Negative**
Actually Positive	1696	106
Actually Negative	92	0

Table 7 presents the Precision, Recall and F1-Score values, which were calculated based on the confusion matrix displayed earlier.

Table 7. Precision, Recall and F1-Score

	Precision	**Recall**	**F1-Score**
IoU=0.5	0.941	0.948	0.944

Table 8 presents the result of the tracking algorithm - the maximum count for each cage in the range of 2-36 is compared to the manual count. The maximum values were chosen out of 40 frames, from the ‘back’ and ‘forth’ videos acquired in the experiment. The number of the hens in the different cages was obtained by adding up the number of hens in the cages found by the system and by manually counting the hens. 978 and 923 hens were detected, on the way forth and back, respectively out of total 1103 hens counted manually. The detected frames and counting results are described in Appendix 1G.

Table 8. Maximum number of detected hens in each cage as compared to ground truth.

Cage number	Ground truth	Forward max	Backward max
2	41	38	40
3	25	21	24
4	40	28	25
5	43	36	30
6	37	36	34
7	27	26	28
8	28	24	26
9	34	26	25
10	37	32	33
11	29	31	22
12	42	44	31
13	42	34	30
14	23	19	23
15	35	34	32
16	26	25	23
17	38	31	29
18	36	35	32
19	33	32	27
20	31	28	29
21	37	23	24
22	35	32	30
23	26	33	27
24	30	22	25
25	24	24	19
26	21	20	18
27	29	19	25
28	27	26	26
29	31	28	24
30	21	23	19
31	34	21	23
32	28	22	21
33	27	26	20
34	28	22	23
35	23	27	23
36	35	30	33
Total counted	1103	978	923

Table 9 presents the accuracy and mean absolute error (MAE) calculated according to the total counted number presented in the Table 8; accuracy is calculated as the total number of hens that the algorithm counts out of the total number of hens counted by human. These results were calculated with a certainty level of 0.8 and IoU=0.5.

Table 9. Accuracy and Mean Absolute Error (MAE) result

	Forward	Backward
Accuracy	88.66%	83.68%
Mean Absolute Error	4.54	5.25

As noted, the system tracked and counted the hens in all the different cages; the number of hens in the cages varied. The lighting conditions are not the same in all the cages, thus, images of hens might have different clarity. Moreover, the hens do not stand still, but move around inside the cages. Therefore, they were in different locations in every video sequence comprised of successive frames. In addition, the RGB-D camera lens is not wide enough to absorb all the hens as a human eye can do, and therefore misses objects. Despite all the difficulties described above, counting laying hens in battery cages was achieved with up to 88% accuracy (Table 9). The algorithm can detect objects in a densely populated picture, containing multiple overlapping objects.

4.4.3 Sensitivity analysis

Directional analysis

Videos acquired while moving backwards were more accurate with 92% accuracy compared to ground truth (Figure 18) and compared to 88% obtained when moving forward (Appendix 1H). The test revealed no statistical difference between the forward and backward direction videos (F-critical 1.981, F-statistic 1.853), with alpha level of 5%. (Table 10).

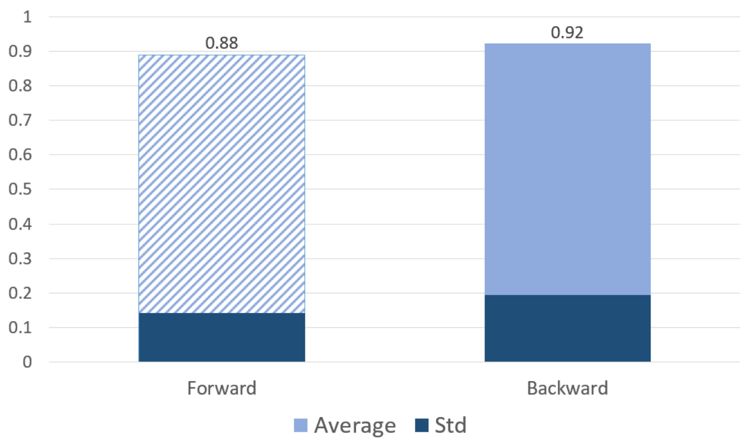


Figure 18. Average accuracy and standard deviation in detections compared to ground truth.

Table 10. F-Test Two-Sample for Variances

	Variable 1	Variable 2
Mean	0.932	0.896
Variance	0.0385	0.0207
Observations	35	35
df	34	34
F	1.8535	
P(F<=f) one-tail	0.0381	
F Critical one-tail	1.9811	

Number of frames per video

When using a full video (which contains more than 44 frames) the average detection ratios compared to ground truth was 0.914. While, when using only 35 frames from a video, the average detection ratios compared to ground truth is 0.883, and 0.968 when compared to full video. 40 frames and 44 frames average detection ratios compared to full video result with 0.975 and 0.981 respectively.

Results compared to ground truth: It can be concluded that using only 35 frames, leads to similar detection results as using a full video, when compared to ground truth (Figure 19, Figure 20, and Table 11).

Results compared to full video analysis: The counting results converge with 10 frames with 88% average detection ratio compared to full video (Figure 21 and Figure 22). However, the standard deviation is relatively high compared to the standard deviation of 35 frames (Table 12).

In conclusion, the object detection algorithm can use only 35 frames of each video instead of using a full video, which includes more than 44 frames. Satisfactory results can be obtained with shortened algorithm runtime; 87% difference compared to ground truth for 35 frames, with ~7-minute elapsed time, versus 89% for full video with ~8.5-minute elapsed time.

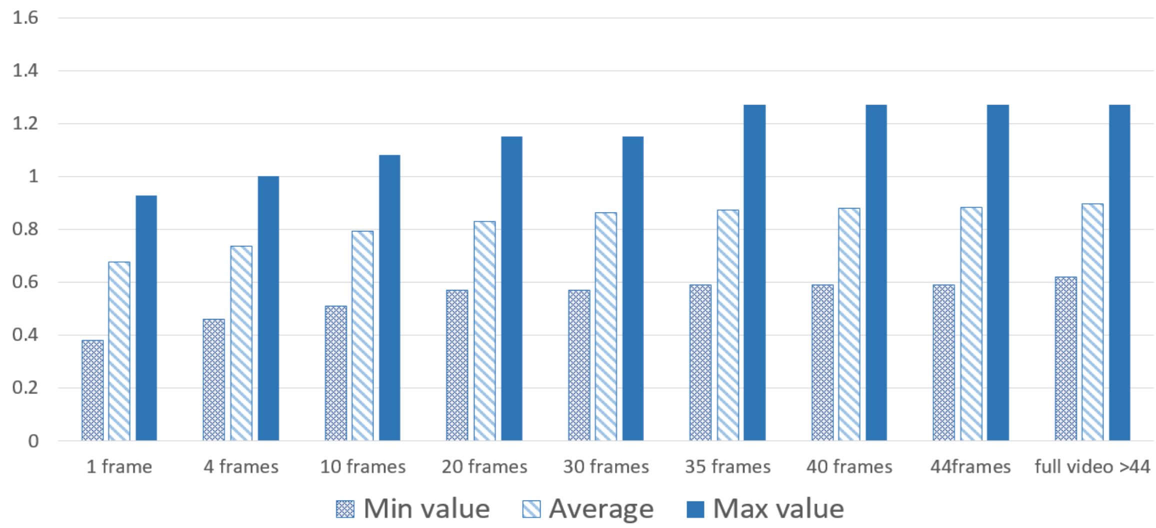


Figure 19. Forward sensitivity evaluation; the minimum, average and maximum detection ratios compared to ground truth (SD is presented in Table 11).

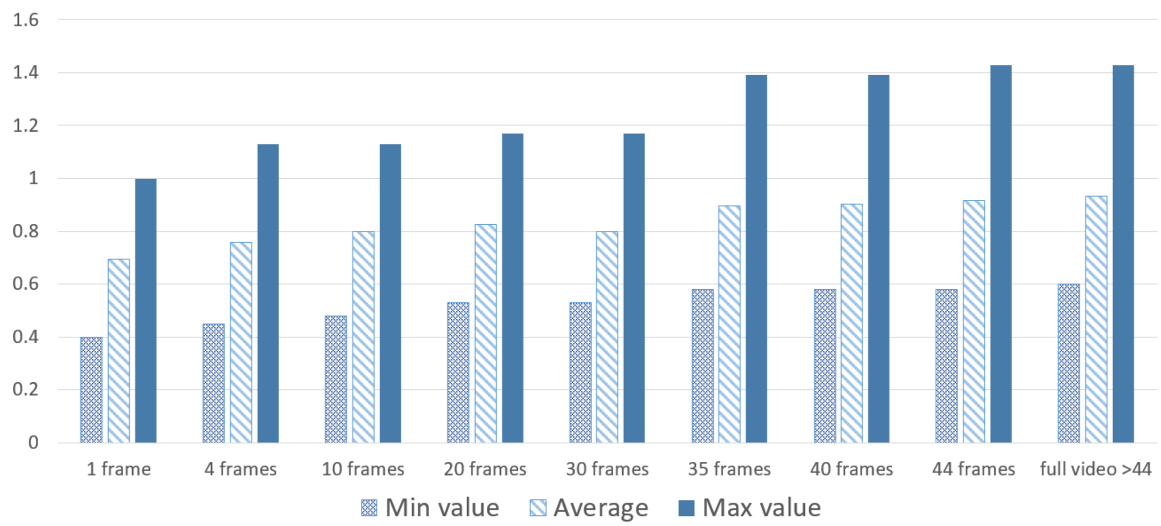


Figure 20. Backward sensitivity evaluation; the minimum, average and maximum detection ratios compared to ground truth (SD is presented in Table 11).

Table 11. Average standard deviation according to number of frames used, compared to ground truth.

number of frames	**1**	**4**	**10**	**20**	**30**	**35**	**40**	**44**	**44+**
Backward	0.14	0.15	0.15	0.16	0.16	0.19	0.19	0.20	0.20
Forward	0.13	0.13	0.13	0.13	0.13	0.15	0.15	0.15	0.15

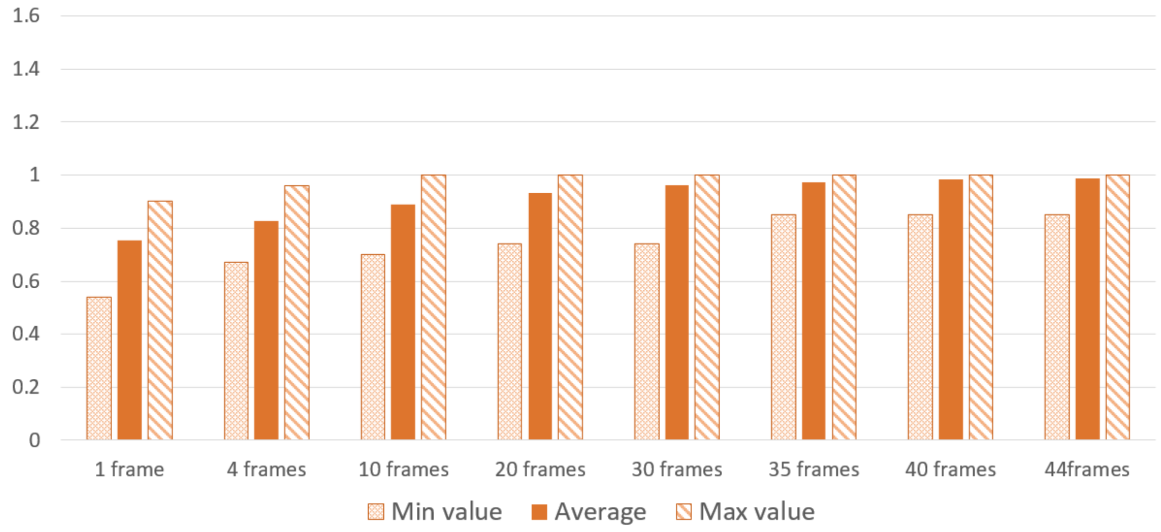


Figure 21. Forward sensitivity evaluation; the minimum, average and maximum detection ratios compared to full video detections (SD is presented in Table 12).

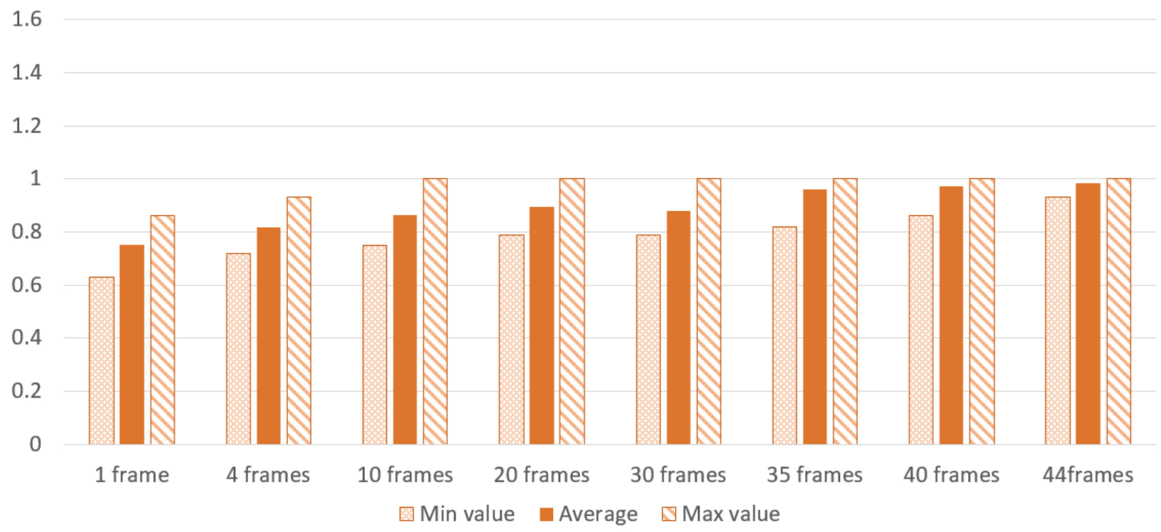


Figure 22. Backward sensitivity evaluation; the minimum, average and maximum detection ratios compared to full video detections (SD is presented in Table 12).

Table 12. Average standard deviation according to number of frames used, compared to full video.

Number of frames	**1**	**4**	**10**	**20**	**30**	**35**	**40**	**44**
Backward	0.06	0.06	0.07	0.06	0.06	0.04	0.03	0.02
Forward	0.09	0.07	0.08	0.06	0.06	0.04	0.04	0.03

5 Automated system for sheep identification

This chapter describes the development of an automatic system for real-time data collection of sheep facial photographs (section 5.1), the research methodology describing the experiments and their evaluation (section 5.2), the sheep biometric identification model (section 5.3) and the study results (section 5.4).

5.1 System design

The overall system (Figure 24) was developed and built on a drinking monitoring facility to ensure that all sheep had frequent access per day, voluntarily, without human involvement. The drinking facility, located in a research sheep pen in Volcani Center, Bet Dagan, monitored each sheep's body weight and water intake per visit. Two 8-MegaPixels RGB cameras, of Digital Single Lens Reflex (DSLR) type, with USB connections were connected to a NVIDIA Jetson Nano embedded system-on-module (SoM) (Appendix 2C). Both cameras video-recorded the sheep while they were drinking water. The cameras were located at a height of 80 cm, one at the face area; the second camera acquired photos of each sheep's ear tag (Figure 1). The system includes an Infrared Red (IR) sensor, which activates the cameras when a sheep inserts its head into the system area. The same IR sensor ends the recording the moment it no longer detects a sheep in the drinking facility (Figure 2). Similarly, if the sensors erroneously detect movement, e.g. a bird triggers them, they will immediately stop the cameras when the bird flies out. Videos were acquired at a speed of 30 frames per second (fps) from both cameras in parallel. The cameras and the Jetson Nano were placed in airtight boxes to protect them from dirt and heat. The system included a SIM card with an Internet network to enable remote connection via a USB dongle.

The logical flow of the system operation is detailed in Figure 23.

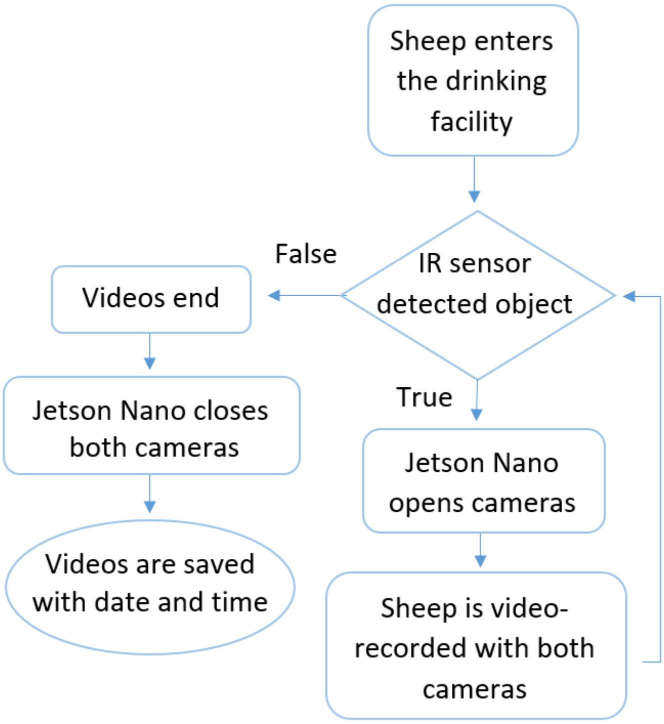


Figure 23. Flowchart of the overall system.

5.1.1 Overview of the existing smart drinking facility

The existing system aimed to document the weight of each sheep and the drinking amount it drank per visit. The facility included a RFID antenna, basin with capacitive sensor and a digital weight which was set across the 40 cm wide facility floor. The digital weight and the RFID antenna were connected to an Arduino program. When an animal arrived to drink, it stepped on the facility floor and it weight was identified. If the sheep weighs more than 10 kg (which means the whole sheep's body is on the facility and not only 2 legs) then the Arduino gets a signal to start recording the amount of water the sheep drank and documents the current weight. In parallel, the RFID antenna, which is located above the basin, at a height of 95 cm from the ground, tries to read the sheep ear tag for individual recognition, and if successful the tag number is reported to the Arduino. All the data was written to an Excel file with a time stamp of the animal arrival. The program was implemented in C++ and is described in Appendix 2C.

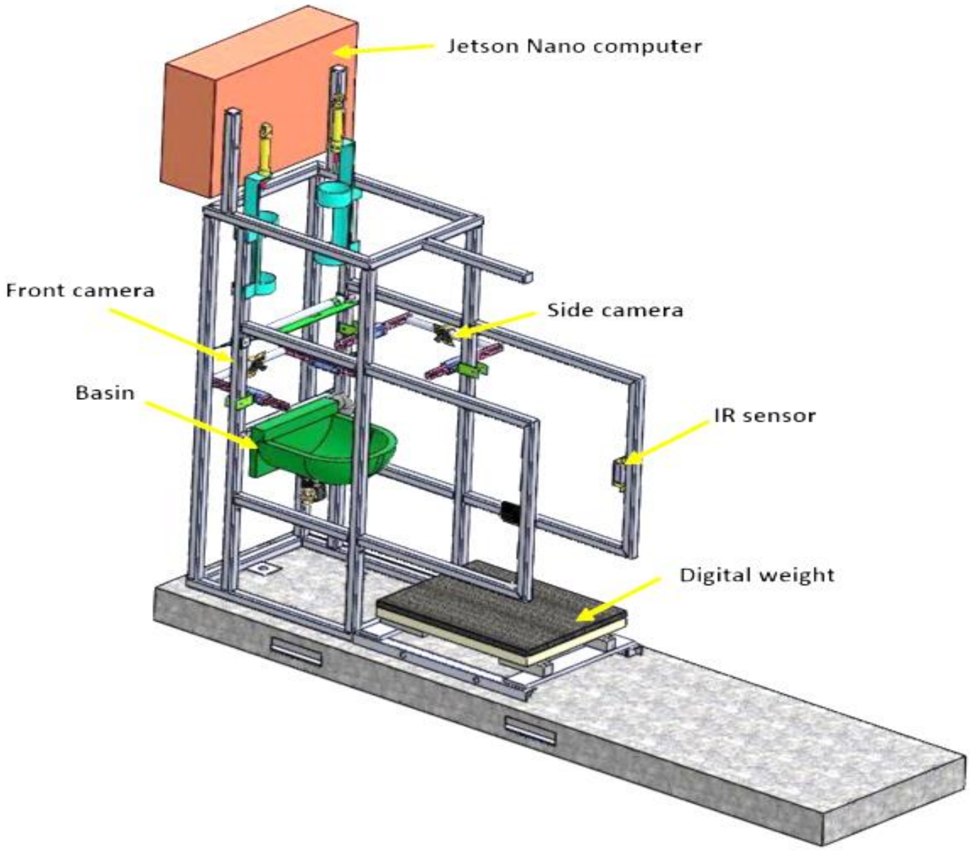


Figure 24. Illustration of the automatic imaging system which was built on the existing drinking facility. The system included the NVidia Jetson Nano embedded system-on-module (SoM), a front camera used for facial video-recording, a side camera used for recording the ear tag used as GT, Basin, an IR sensor and a digital weight.

5.1.2 Infrared (IR) sensor

An IR sensor can detect changes in the amount of infrared radiation impinging upon it, which varies depending on the temperature and surface characteristics of the objects in front of the sensor. When an object, such as a sheep, passes in front of the background, the temperature at that point in the sensor's field of view will rise from the environment temperature to body temperature. The sensor converts the resulting change in the incoming infrared radiation into a change in the output voltage, and this voltage is sent to the Jetson Nano as trigger for starting and ending videos. The IR sensor was located at the entrance of the facility, at a height of 70 cm from the floor.

Figure 25. IR sensor attached to the existing drinking facility.

5.1.3 NVidia Jetson Nano embedded system-on-module (SoM).

Besides the fact that it is a small and fast computer, which made it convenient to use in outdoor conditions, the Jetson Nano can also interface with external devices through the communication pins which are directly connected to the Jetson Nano module. Using this ability, the Jetson Nano was connected to the IR sensor, controlling the data collecting process according to the voltage that arrived - 3 volt was used as signal for starting videos and GRD (0 volt) as a signal for ending. In addition, by using a serial (TX-RX) communication, the Jetson was connected to the smart drinking facility Arduino, that transferred the RFID tag to the Jetson Nano embedded system-on-module (SoM), through a specially written Python code which saved the tag number as the video's name, creating an automated tagging of the data collected. The video-recording code (Appendix 2D) was programmed in Python3 under Ubuntu, an open source Linux operating system. Videos were acquired at 30 frames per second (fps) from both cameras in parallel. In order to prevent damages in the outdoor environment, the cameras and the Jetson Nano were placed in airtight boxes.

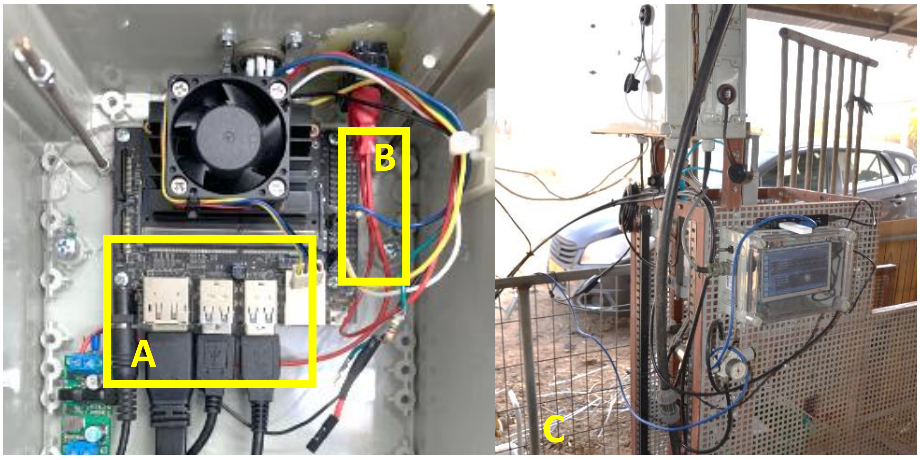


Figure 26. Jetson Nano embedded system-on-module (SoM), including - (A) three USB connections for the two cameras and the WIFI dongle. (B) The communication pines used to connect with Arduino and IR sensor. (C) Jetson Nano attached to the existing facility.

5.2 Methods

5.2.1 Experimental design

Four experiments were conducted in 2020 and 2021 in Volcani's Center research sheep pen on two groups of sheep. The sheep were raised from two months of age until five months, during which time the sheep gained about 25 kg in weight. The pen included three feeding facilities that contained food all the time, and a single drinking source where the experimental photography system was located ensuring data collection on all sheep.

The experiments were conducted on two different groups of sheep, both from Assaf breed – the first group (group 1) contained 47 sheep that were raised from September to March, and the second group (group 2) were raised from April to August contained 34 sheep. The experiments with group 2 were used to evaluate the model's capability to learn new identities.

<u>Group 1 experiments:</u>

In order to obtain a rich and varied dataset for the model development, a two-week experiment was conducted on group 1 (47 sheep) in November (experiment 1), when the sheep were two months old and weighted about 30 kg (average=33.6, SD=6.6). Two additional experiments were conducted. One experiment was done in the middle of the growth period (experiment 2), at the end of December 2020, when the sheep weighed about 45 kg each (average=44.4, SD=6.51). The other was conducted at the end of February 2021, at the end of the growth period (experiment 3); the sheep were at their maximum weight, about 60 kg (average=57.2, SD=7.82). Each experiment was conducted over three consecutive days. Experiments 2 and 3 included only 32 sheep, since 15 sheep were removed from the pen during the course of experiments, due to illness.

In all the experiments made, data was collected automatically at all daylight hours ensuring a variety of light conditions affected by sunlight and weather. Each sheep arrived 2-3 times per day, at different hours, and were captured in a different posture at the drinking facility, resulting in a diverse sheep facial database.

Group 2 experiments:

Similar to experiments done on group 1, a two-week experiment was conducted in April (experiment 4) on group 2 (34 sheep) when the sheep were two months old and weighed about 30 kg (average=32.5 SD=7.60).

5.2.2 Data

A total of four datasets corresponding to the four experiments were obtained (Table 13), with the following steps: first, each sheep's face video (acquired from the front camera) was manually tagged with the sheep's ID, obtained from the corresponding side video. Then, each video file was converted with a Python3 code to an image sequence.

The first dataset included images only from the beginning of growth period (experiment 1), with which the biometric identification model was developed. The second dataset, aimed to examine the growth influence on identification, included images from three different time periods – beginning, middle and end of growth (from experiment 1-3). The first dataset included 3055 images, 65 images for each of the 47 sheep, of which 52 images were randomly selected for the model development and the 13 images remained were used to test the model ability to identify individual sheep. The second dataset contained 2080 images, 65 images for each of the 32 sheep; 1248 images were taken from the beginning of growth period (experiment 1), 416 images from the middle of growth period (experiment 2) and 416 from the end of growth period (13 images for each of the 32 sheep), were sheep were at their maximum weight (experiment 3).

The third dataset included images of group 2 sheep, only from their beginning of growth period (experiment 4). Similar to the first dataset - this dataset included 2210 images, 65 images for each of the 34 sheep, of which 52 images were randomly selected for the model development.

In addition, for the object detection algorithm, which aims to locate the sheep's face in each image, a labeled dataset was created from 30 sheep videos. This dataset included a total of 400 images, manually tagged with a bounding box around each sheep's face.

All datasets (Appendix 2E) included video files that were converted with Python3 code to an image sequence, and then using a specially developed face detection algorithm (section 5.3), the sheep face were captured and saved as a newly images, which then were resized from 640 heights on 480 widths, to 112*112, to enable the deep learning model to train faster on smaller images.

Table 13. Summary of datasets

Dataset purpose	Group	Experimental period	Sheep	Images
Group 1 dataset	1	Beginning of growth	47	3055
Growth sensitivity	1	Beginning, middle, end of growth	32	2080
Group 2 dataset	2	Beginning of growth	34	2210
Object detection	1	Beginning of growth	30	400

5.2.3 Model

The biometric identification model applied to each image (Hitelman, 2021) entailed two steps: (1) face detection and (2) classification (section 5.3). Face detection was achieved using a Faster R-CNN algorithm. For classification seven different CNNs architectures were compared. Each CNN examined was used as the embedding network for implementing ArcFace loss function, resulting in seven classification models. Those models were chosen based on previous classification task success; AlexNet was implemented by (Salama et al., 2019) which resulted high performance on sheep facial classification, VGG16 made improvement over AlexNet (Simonyan and Zisserman, 2015) and achieved top-5 test accuracy classification on ImageNet. ResNet50 was used as the embedding network for implementing ArcFace loss function (Deng et al., 2019) and therefore was examined along size similar variants – ResNet50v2, ResNet101V2. EfficientNet reached State-of-the-Art accuracy on both ImageNet and common image classification transfer learning methods (Tan and Le, 2019).

In addition, since the Softmax loss function is widely used in deep face recognition (Cao et al., 2018; Parkhi et al., 2015), as a baseline, classification performance using Softmax was evaluated, when ResNet50 used as embedding network.

All models were pre-trained on the ImageNet dataset, and were trained and tested on group 1 dataset, i.e. 47 sheep, 2444 and 611 randomly selected images included in the train and test sets respectively. The models' hyper parameters (Appendix 2F) values were determined by trial and error on the ResNet50 model, where for the Softmax model, the values were set as follows; Learning Rate (LR) max value - 0.01 and min value - 0.00001, 100 Epochs, Batch Size equal to 32, Weight Decay value set to 0.001 and Dropout to 0.5. All seven tested embedding networks combined with ArcFace loss function were set with the same values, as follows; Learning Rate (LR) max value - 0.005 and min value - 0.00001, 30 Epochs, Batch Size equal to 32, Weight Decay value set to 0.001 and Dropout to 0.5, ArcFace parameter S set to 3 and M to 0.05.

The seven classification models were compared using the Post-hoc Tukey's statistical significance test. After selection of the best classification model, the model was trained and evaluated on group 1 and group 2 separately, and on a unified group which included all the sheep. Additionally, transfer learning methods were examined to decrease the training time while maintaining a lower generalization error (section 3.5).

5.2.4 Performance measures

Sheep biometric identification performance was evaluated by two indices - accuracy and categorical cross entropy. Accuracy is the quintessential classification metric, and is calculated from the confusion matrix as the proportion of true results among the total number of cases examined in the model.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Categorical cross entropy is a loss function that is used in multi-class classification tasks, as had been researched in this thesis. The loss function is defined as the difference between the predicted value by the model and the true value, computed by the following sum:

$$\text{Cross} - \text{entropy} = -\frac{1}{\text{n}}\sum_{i=1}^{n}\sum_{j=1}^{m} y_{i,j}\log(p_{i,j})$$

Where $\mathrm{y_{i,j}}$ denotes the true value and $\mathrm{p_{i,j}}$ denotes the probability predicted by the model of sample i belonging to class j. The loss function is used to optimize the model while accuracy metric is used to measure the model's performance in an interpretable way, while the main objective in a learning model is to reduce (minimize) the loss function's value and increase (maximize) the model accuracy.

The confusion matrix compares the actual ID's with those predicted by the model. The horizontal rows represent the target values (what the model should have predicted - the ground-truth) and the vertical columns represent the predicted values (what the model actually predicted), while the diagonal represents the number of predictions where the classifier correctly predicts the positive class as positive (true positives).
Precision is the proportion of predicted positives which truly positive, and was calculated as the sum of true positives across all classes (sheep ids) divided by the sum of true positives and false positives across all classes. Recall is the proportion of actual positives which correctly classified, and was calculated as the sum of true positives across all classes divided by the sum of true positives and false

5.2.5 Analysis

Two analyses were conducted:

1. **Cross Validation:** The K-Folds Cross Validation process was used to evaluate the models implemented, to ensure that the data distribution did not influence the model's performance. Data was split into five equal folds, and the model was trained on all but one fold; the remaining fold was used to evaluate performance. This process was repeated five times, with a different fold utilized for evaluation each time. The mean, minimum, maximum and standard deviation accuracy results were calculated for each model, in order to evaluate performance (section 5.4).
2. **The effect of sheep growth:** In order to determine whether sheep gain weight and matures throughout their growing process influence identification, performance on

images from the end of growth were evaluated twice - when the model was trained only on images from the beginning of growth period, and when the model was trained on images from the middle and beginning of growth combined together. The two growth models were trained on the 32 remain sheep from group 1 (1344 and 416 images in the train and test sets respectively), with the same hyper parameters values as details in section 5.2.3.

5.2.6 Sensitivity Analysis

Two sensitivity analyses were conducted to determine the impact that different quantities of data have on the model's performance. Both analyses were performed on group 1 dataset.

1. **Amount of training images:** The model was trained nine times. Each training was performed with a different number of images between 10 and 52 images for each training run. All models were tested on the same 13 images. The identification accuracy achieved for each sheep and the average accuracy of the group were evaluated for each run.
2. **Number of frames needed for identification:** Since the videos were acquired while the sheep were drinking, each video contained more than 1,000 frames. Running the model on all the frames is time consuming. Furthermore, the sheep were captured in various postures; many frames captured only partial faces or individual features, which are much more difficult to identify. This analysis evaluated the minimum number of frames required to be randomly sampled from a full video in order to ensure that each sheep was identified correctly. Accuracy was evaluated on different quantities of randomly sampled frames of the trained model that used 52 images, between one to ten frames. The decision regarding each sheep’s identity was made based on the ID that the majority of the frames indicated.

5.3 Model

The biometric identification model included two steps – face detection and classification (Figure 4). In order to locate and localize the sheep's face in an image, the Faster R-CNN deep learning object detection algorithm (section 4.3.2) was applied (Jiang and Learned-Miller, 2017). Then, the detected face was cropped, and resized to 112X112 pixels according to Deng

et al., (2019). Finally, the cropped face was provided as input to the second step which included the classification model (Figure 27).
One of the main challenges in classification using Convolutional Neural Networks (CNNs) for face recognition is the design of appropriate loss functions that enhance discriminative power. In this thesis, the recent state-of-the-art loss function was implemented - Additive Angular Margin Loss (ArcFace) which obtained highly discriminative features for face identification, when resNet50v2 was chosen as the embedding network to train ArcFace.

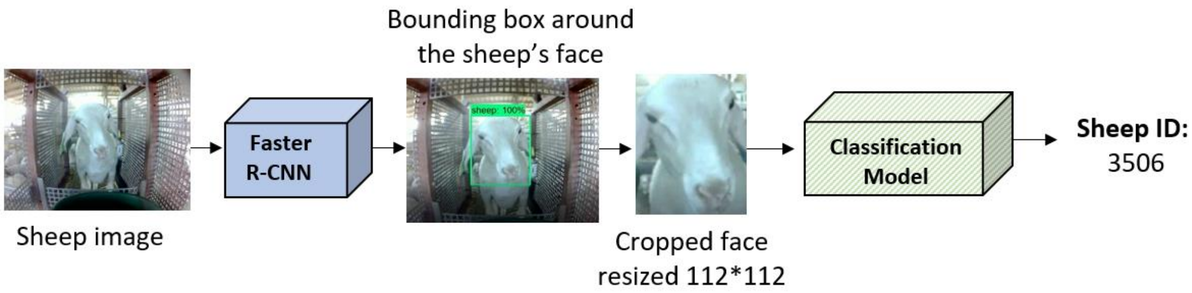


Figure 27. Biometric identification model schematic flowchart.

5.3.1 Selected classification model

The selected classification model design (Appendix 2F) was based on the ResNet50v2 CNN architecture pre-trained on the ImageNet database. In ResNet50, the input of a convolutional layer bypasses one or more layers, and is added to the outputs of forward layers, denoted as residual mappings. The ResNet50 architecture avoids a vanishing gradient, enabling easier learning even with deeper structures, because the information is directly transmitted (Yamazaki et al., 2019). ResNet50V2 is the advanced version of Resnet50 CNN, which is all about using the pre-activation of weight layers instead of post-activation (He et al., 2016).
The ResNet50V2 includes five stages. In the first stage, the architecture performs the initial convolution and max-pooling using 7×7 and 3×3 kernel sizes respectively. Then, each of stages 2-5 consisted of a convolution block and several identity block performed in a row, where each convolution block and identity block is composed of 3 convolution layers. The three layers are 1×1, 3×3, 1×1 kernel size, where the 1×1 convolution layers are responsible for reducing and then restoring the dimensions and the 3×3 layer is left as a bottleneck with smaller input or output dimensions. The difference between those blocks are that the convolution block reduces the size of the input by half in terms of height and width and

doubles the channel width, while identity blocks keep the same input size. Finally, the recommended sequence of actions (Figure 28) between the last layer of the last block and the ArcFace layer was implemented according to (Deng et al., 2019). The architecture included a total of 197 layers and a total number 40,376,832 parameters.

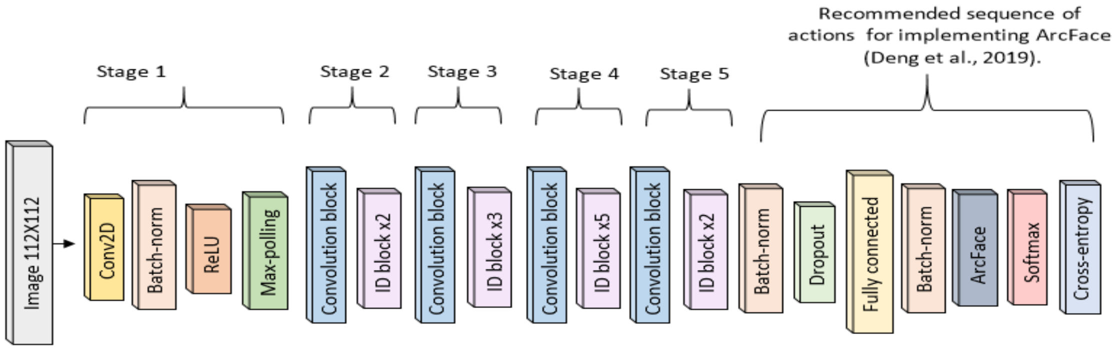


Figure 28. Schematic description of the classification model architecture; the input is a sheep facial image in size 112x112. Stage 1: initial convolution and max-pooling using 7×7 and 3×3 kernel sizes. Stage 2-5: in blue – the convolution block, in pink – Identity block, performed 2, 3, 5 and 3 times respectively at each stage. The last layer of the last block is connected to ArcFace layer according to the recommended sequence. The output is the loss of the highest probability predicted class for the image by Softmax.

5.3.2 ArcFace

The ArcFace loss function (Figure 29) aims to reduce the angle θ during learning, which is the angle between the facial image which best represent the class (ground truth) by the appropriate weight column, and the input facial images which is represented by the 512 size feature vector – X. The implementation of ArcFace layer was done by the following steps, using Tensorflow and Keras Python open source: first, Vector X and the weight matrix – W, underwent normalization according to L2. Then, cosθ, which also called "Logit" was extracted from the multiplication between X and W by calculating the arcos (cosθ). Afterwards, a hyper parameter m that represents an additive angular margin penalty was added to the angel θ. The m value was set to 0.05 following trial and error. Then the new Logit, cos(θ+m), was calculated and multiply by another hyper parameter, which represented the hypersphere radius – s (Deng et al., 2019), which was set to 3 following trial and error. Finally, the Logit went through the Softmax function and to the cross-entropy calculation.

In addition, in order to better monitor the training process, two important metrics were added: (1) the size of the mean of the θ vector, (2) the size of the average Logits vector. These metrics were presented in addition to accuracy and loss calculation. ArcFace implementation is described in Appendix 2F.

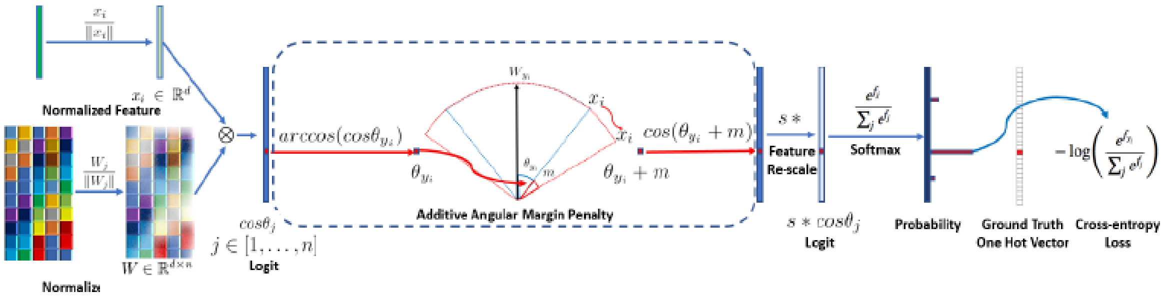


Figure 29. ArcFace schema (Deng et al., 2019)

5.3.3 Model additions

In order to achieve best performance, the following steps were implemented: All images were normalized based on Wang et al., (2018) by zero-centring each color channel (red, green, blue), in order to prevent the Vanishing Gradient Problem. To adjust the learning rate during training, the Cosine Annealing Scheduler was used (Loshchilov and Hutter, 2017). An early stopping method was used to automatically stop the training process when the model's loss performance stopped improving on the validation set (Xu et al., 2019). This method helped to avoid overfitting of the neural network. Finally, a SGD optimizer was implemented, with momentum set to 0.85, to help accelerate gradient vectors in the right directions, thus leading to faster converging. In addition, each image was augmented about seven - eight times via vertical flip, random rotation up to 30-degree, Gaussian noise addition, sharpening images and changing the brightness of images (Figure 30). The technique resulted in a total of 400 augmented images created per sheep, used for the model training.

Figure 30. Data augmentation examples from left to right; original image, vertical image, random rotation image, darken image, brighten image.

5.3.4 Transfer learning

Transfer learning is a very popular method widely used among computer vision researchers (Guo et al., 2018). Transfer learning enable to utilize knowledge from previously learned tasks and apply them to newer, related ones. In the case of problems in the computer vision domain, knowledge from an existing task, such as edges, shapes, corners and intensity, can be shared acts as an additional input when learning a new target task (Huh et al., 2016). Therefore, transfer learning has the benefit of decreasing the training time for a neural network model and resulting in lower generalization error. In this research, transfer learning methods were used in order to improve performance, by training the model on one group of sheep and then retrained the convolution layers on the other group, with a faster learning rate. Transfer learning methods were trained and evaluated on the same datasets, i.e. the test sets of both groups included 13 images per sheep; and with the same hyper parameters values detailed earlier but with a decreased learning rate of 0.0001 (Appendix 2F).

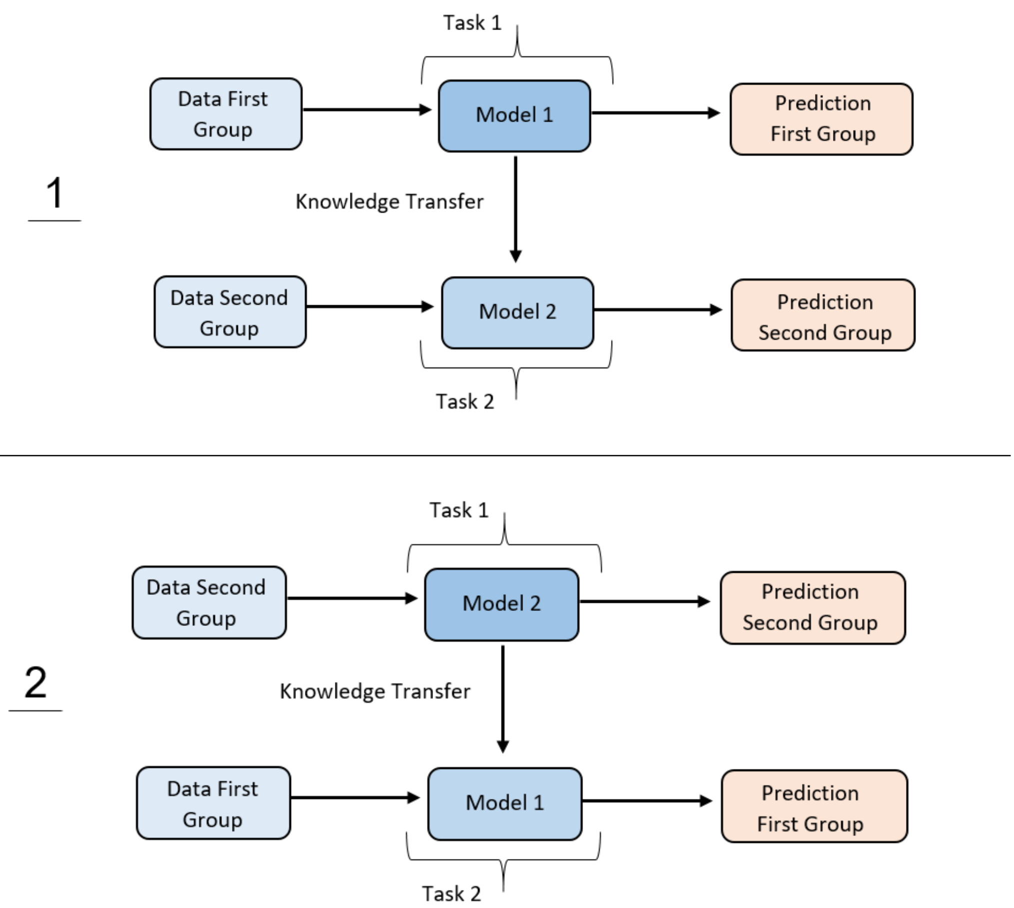


Figure 31. Transfer learning; in this research learning process was transfer from group 1 model to the group 2 model, and vice versa.

5.4 Results

The results of the biometric identification model and sensitivity analyses are reported.

5.4.1 Sheep biometric identification model

Softmax loss function

Results obtained after training converged (stopped automatically using the early stopping method after 70 Epochs) were 72.66% accuracy with a 2.01 loss value (Appendix 2F).

Comparison of classification models

Results (Table 14, Appendix 2G) reveal that all the examined models achieved better performance than Softmax in a shorter run-time (converged after 30 Epochs maximum, while Softmax converged after 70 Epochs). ResNet50V2 and ResNet50 were significantly better than the other models (Post-hoc Tukey's statistical significant test, α = 0.05) (Appendix 2H). ResNet50V2 achieved the best accuracy (Table 2); specifically, it was 22% better than the worst model (VGG16), and 1.5% better than the second best model (ResNet50).

Table 14. Summary results of the different models

Classification model	Accuracy in %	Convergence Epoch
ResNet50	93.4	29
ResNet50V2	**94.9**	**29**
ResNet101V2	91.3	29
EfficientNetB0	85.2	14
EfficientNetB3	87.3	14
AlexNet	76.2	29
VGG16	72.9	28

Selected model

Using the cross-validation technique, ResNet50V2 CNN combined with the ArcFace loss function resulted in an average 95% accuracy for groups 1 and 2, while the unified group achieved lower results, but only by 2% (Table 15, Appendix 2G). All three trained models converged after 25 epochs. The low standard deviation of the five folds (less than 0.01 for all models, Appendix 2J) proves the model's reliability.

Table 15. Average classification performance on the experimental groups.

	Accuracy	Loss
Group 1: 47 sheep	0.954	2.069
Group 2: 34 sheep	0.957	1.811
Unified group: 81 sheep	0.939	2.437

The decrease in the unified group performance can be explained due to similarities between the groups, since some of the sheep in the different groups were born to the same parents and therefore were very similar (Figure 32).

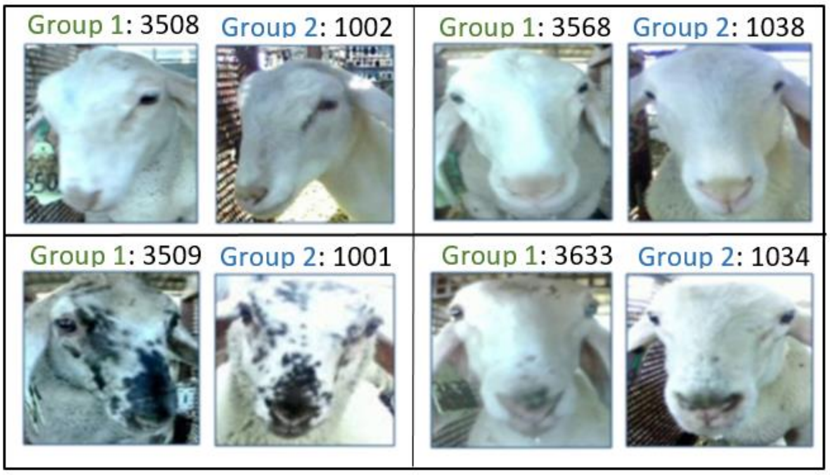


Figure 32. Example of sheep pairs similarities between groups; green images - sheep from group 1, blue images – sheep from group 2.

Confusion matrixes (Appendix 2I) calculated on the test set, which included 13 images per sheep, revealed 94.92% and 96.83% true positive on group 1 and 2 respectively. As it can be deduced from the matrixes, most of the sheep were identified in at least 12 out of 13 images, with an average of 12.76 and 12.58 correct identification images for group 1 and group 2 (Table 16, Figure 33, Figure 34). Furthermore, all sheep (in both groups) were correctly identified with a minimum of 10 images. Therefore, according to the majority classification rule, we can assume 100% successful identification for each sheep. The specific images that the model failed to identify were difficult to identify even by the human eye (Figure 35).

For convenience only, the presented performances are related to the first fold of the cross-validation technique.

Table 16. Average, minimum, maximum and standard deviation of the correct identification images (out of 13 per sheep).

	Mean	Min	Max	SD
Group 1	12.76	10	13	0.80
Group 2	12.58	10	13	0.84

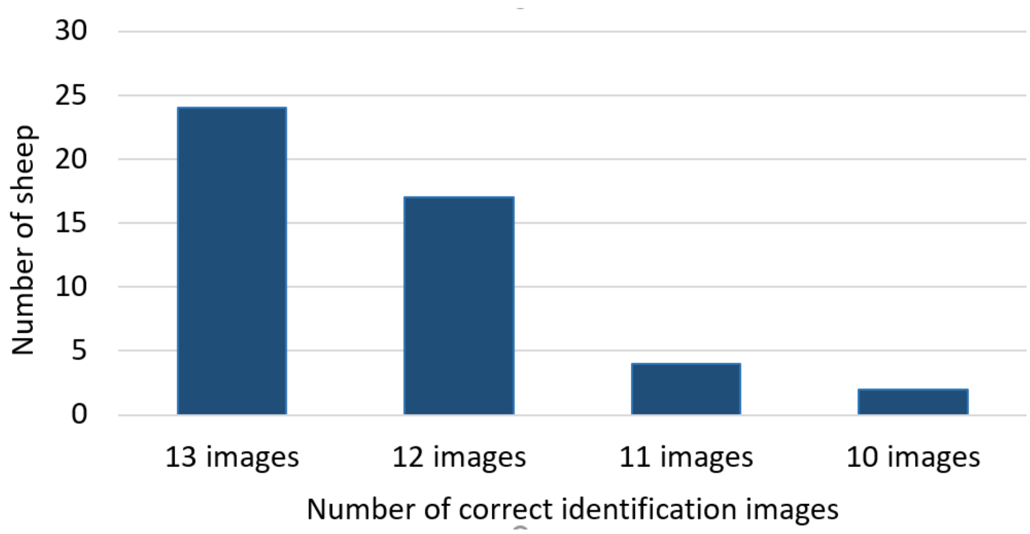


Figure 33. Group 1 identification results; 24 sheep were identified in all 13 images, 17 sheep in 12 images, 4 sheep in 11 images and only 2 in 10 images.

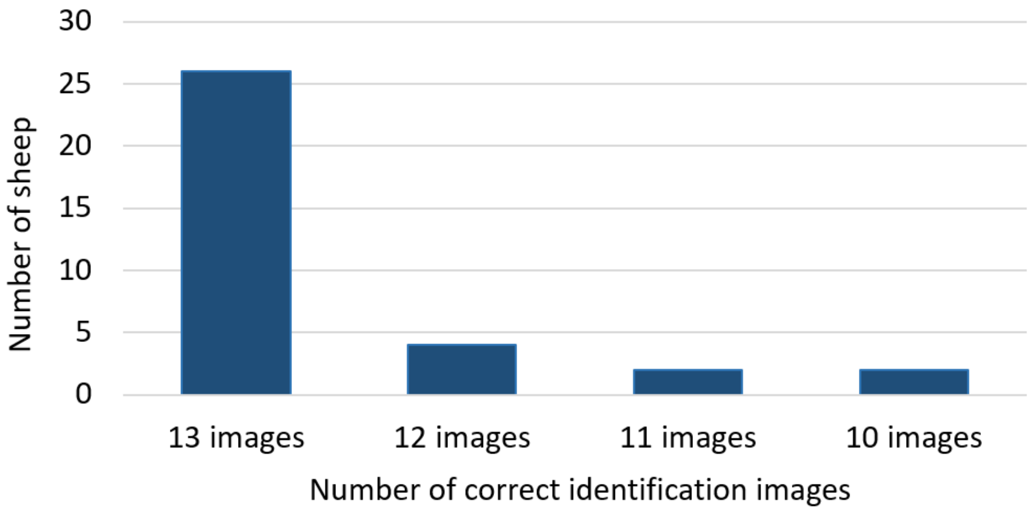


Figure 34. Group 2 identification results; 26 sheep were identified in all 13 images, 4 sheep in 12 images, 2 sheep in 11 images and 2 in 10 images.

Precision, Recall and F1-Score indexes were calculated for each group, with an average of 96.2%, 95.8%, 95.8% respectively (Table 17).

Table 17. Classification measurements results.

	Precision	Recall	F1-Score
Group 1: 47 sheep	0.952	0.949	0.949
Group 2: 34 sheep	0.972	0.968	0.968
Unified group: 81 sheep	0.943	0.939	0.941

Figure 35. Example of wrong identification of single image; left - one image of sheep 3622 from group 1, was wrongly classified as sheep 3527, right – one image of sheep 1015 from group 2, was wrongly classified as sheep 1016.

Transfer learning

Transfer learning achieved by the cross-validation technique (Table 18, Appendix 2G) resulted in an average improvement of 1.85% increase in accuracy and a 0.25 decrease in loss value (Table 5), with a shorter training process (9 vs. 25 epochs). The data distribution had no effect on the identification performance as revealed by the low standard deviations.

Table 18. Average performance of the transfer learning methods.

Transfer model	Accuracy	Loss
G1 -> G2	0.977	1.695
G2 -> G1	0.971	1.705

Biometric identification of sheep was achieved with 95% accuracy on both experimental groups. The biometric identification performances were improved by 2% in accuracy using transfer learning methods. All five trained models presented in this study were saved to be used as a classifier for unseen sheep images (Appendix 2G).

5.4.2 Analysis

K-Folds Cross Validation

The performance of the folds for each of the five models (models for group 1, group 2, unified group, and two the transfer learning models) are presented in Table 19, and detailed in Appendix 2J.

Table 19. K-Fold accuracy results (summary of five runs)

	Group 1	Group 2	Unified Group	Transfer G1 -> G2	Transfer G2 -> G1
Mean	0.954	0.957	0.939	0.977	0.971
Min	0.939	0.940	0.930	0.963	0.959
Max	0.966	0.966	0.946	0.987	0.978
SD	0.009	0.009	0.005	0.009	0.007

Table 19 revealed that the standard deviation of the five folds was less than 0.01 for all models, therefore we can assume that random data distribution had no effect on models performance.

The effect of sheep growth

Growth models (Table 20, Appendix 2G) achieved 93.7% and 91.3% accuracy respectively, which is lower in 1.7% and 4.1% from the base model performance.

Table 20. Growth models classification results

Model	Accuracy	Loss
Growth1: trained on beginning and middle images	0.937	2.084
Growth2: trained on beginning of growth images	0.913	2.189

Training on images from the beginning of growth only while testing on images from the end of growth (Growth2), achieved 2.4% lower accuracy than did training with images from the middle as well as the beginning of the growth period (Growth1). However, the difference in accuracy of Growth2 is relatively small compared with that of the base model (4.1%).

In order to analyze the identification performance of the growth sheep model, the confusion matrix was calculated on the second model, the model without the middle images (Appendix 2I). The matrix revealed 91.34% true positives, with an average of 11.87 correct identification

(Table 21, Figure 36). All sheep in Growth2 were identified in at least 10 images out of 13, equal to the minimum correct identified images achieved with the base dataset (Table 21). Figure 37 shows examples of sheep that were correctly identified during the growth period.

Table 21. Average, minimum, maximum and standard deviation of growth model correct identification images (out of 13 per sheep).

	Mean	Min	Max	SD
Growth2 model	11.875	10	13	0.906

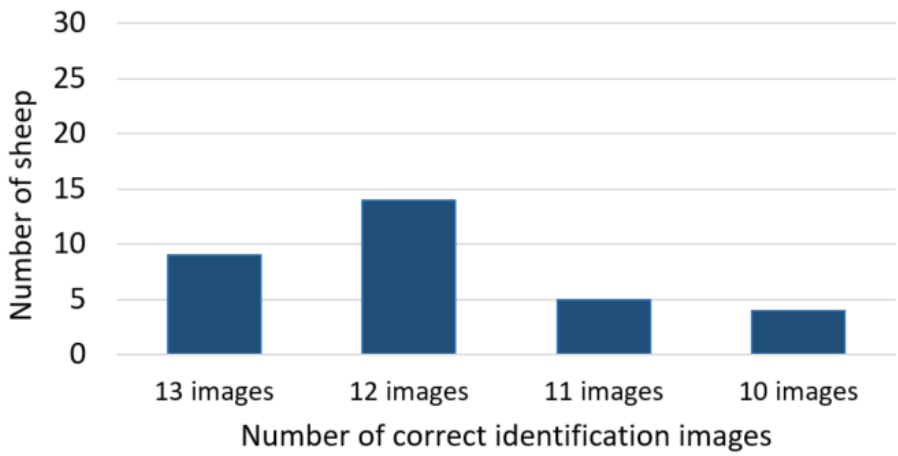


Figure 36. Growth identification results; 9 sheep were identified in all 13 images, 14 sheep in 12 images, 5 sheep in 11 images and 4 sheep in 10 images.

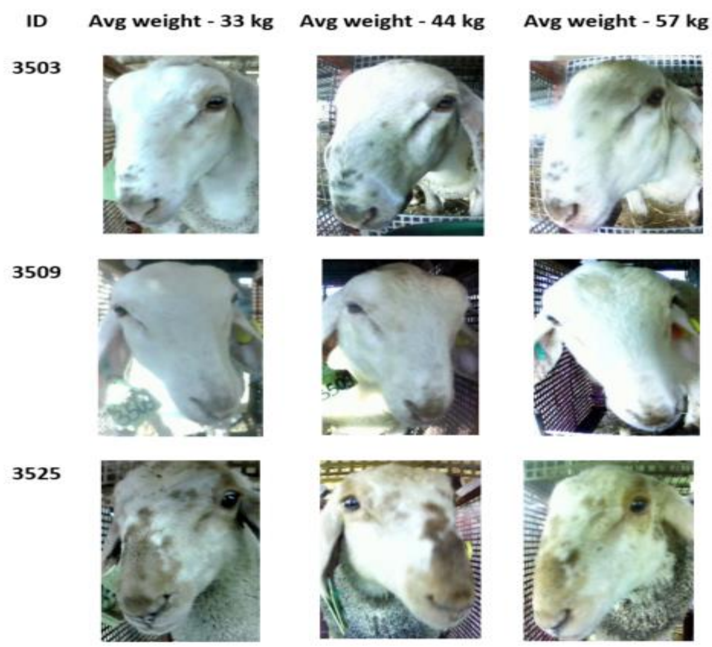


Figure 37. Correct prediction growth examples of sheep, from left to right: image from beginning, middle and the end of growth period.

5.4.3 Sensitivity analysis

Amount of training images

Training each sheep with at least 35 images resulted in a minimum identification accuracy exceeding 69% per sheep (Figure 38, Appendix 2K, i.e. at least nine of the 13 images were properly classified). By utilizing the majority-based decision described above, each sheep was correctly identified with decreasing SD as the number of images increase (Table 22).

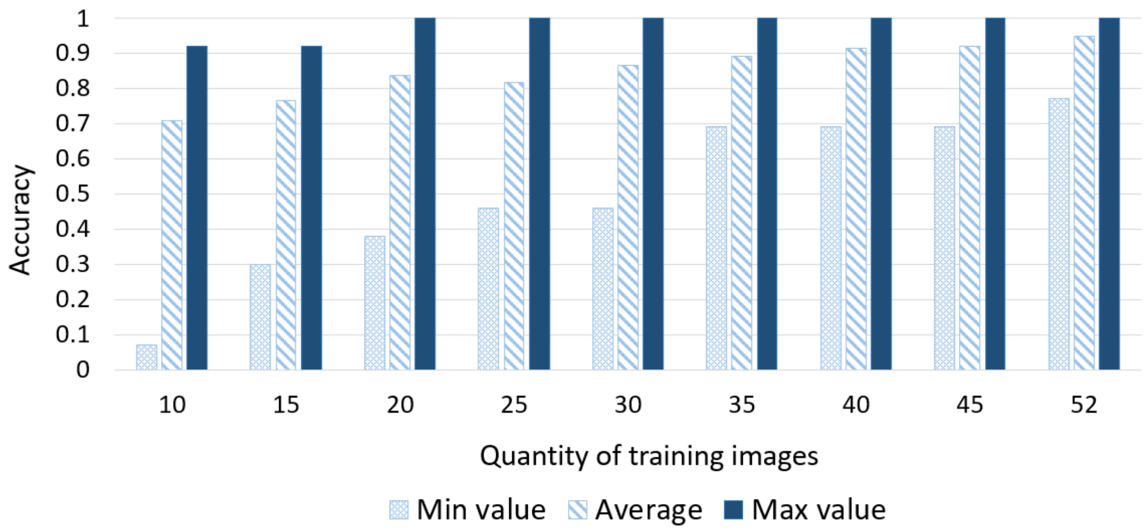


Figure 38. Quantity of training images evaluated; the minimum, average and maximum accuracy achieved (SD is presented in Table 22).

Table 22. Standard deviation according to number of images used.

Number of images	10	15	20	25	30	35	40	45	52
	0.20	0.179	0.11	0.15	0.12	0.10	0.09	0.08	0.06

Number of randomly sampled frames

When classification was done using eight and nine frames, the image identification average accuracy was 59.5% and 63.7% respectively (Figure 39 and Table 23). Results revealed that there were sheep that were identified in less than half of the frames and therefore were misclassified. However, when ten frames were used, the average image identification accuracy was 71.7%, with a minimum of six out of ten images, implying correct classification of all sheep. Accordingly, in order to ensure proper identification of each sheep, the classification decision must be made based on a majority of at least ten randomly selected frames (Appendix 2L). Then, 100% sheep identification can be achieved (with average image accuracy of 95%).

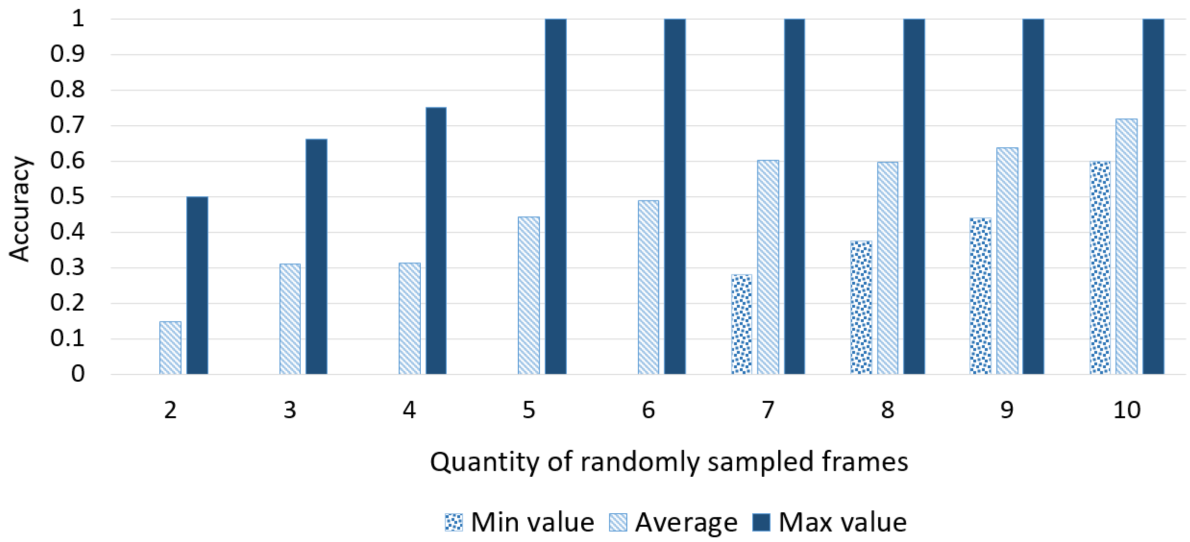


Figure 39. Number of randomly sampled frames; the minimum, average and maximum accuracy achieved (SD is presented in Table 23).

Table 23. Standard deviation according to number of images used.

Number of frames	2	3	4	5	6	7	8	9	10
	0.22	0.27	0.24	0.294	0.26	0.18	0.14	0.12	0.10

6 Discussion

6.1 Mobile system for detecting and counting laying hens

The developed system does not rely on a specific feature of the test hen house and proved operational in variable illumination conditions. It provides an inexpensive, fast and user-friendly system that can potentially be used in different farms, ensuring adequate hen density according to the limits set by regulations (Rajsic and Fox, 2017).

The 88% detection accuracy achieved in this study was lower than the previous study which achieved 90% (Geffen et al., 2019). The slightly lower results can be attributed to the differences in the setups - in the previous study, data was automatically collected, without any human present, resulting in minimum disturbance to the hens; in the current study a mobile cart was led by human, which caused more movement of the hens while the video was recorded due to their panic from the humans around. Moreover, the camera used in the previous study had a wide lens which acquired the whole width and height of the cages at once and detected a significant part of each hen's entire body. Part of the algorithms inaccuracy in the current system was since hens in neighbouring cages were captured and counted, as they were part of the cage in focus.

Hens tracking is challenging; a hen might be visible to the camera only in specific frames, due to the hen's constant movements. In addition, not all the hens were detected in every frame. Thus, using the detection of hens in successive frames is valuable for tracking. In our study, the detector was trained on sequences of 2D images. The tracking algorithm included the 2D images, and the dimension of time resulting in 3D images. We utilized the advantages of the video camera setup, which provided 3D information from multi views on the hens from successive frames, and achieved a reliable hen count.

6.2 Automated system for sheep identification

To our knowledge, this study was the first to implement robust biometric identification of young sheep (2-3 months old). Compared with earlier studies (Corkery et al., 2007; Salama et al., 2019), the slightly lower image identification accuracies achieved in this study can be attributed to differences in the data collection methods and differences between the sheep. In the study that achieved 96% accuracy (Corkery et al., 2007), images were manually acquired in controlled conditions on adult sheep aged 3-4 years which have higher differences. In the other study which achieved 98% accuracy (Salama et al., 2019), CNNs were tested on a group of sheep, in which the ages of the individual sheep varied greatly. In our study, biometric identification was done on identically-aged young sheep (lambs) with an automated machine vision system in field conditions, implying a more natural but complicated identification process. We surmise that the model will perform better when applied to adult sheep that have greater differences between individuals.

Furthermore, the presented study reveals that identifying animals before and after maturation (after 2-3 months) is feasible. Each sheep was successfully identified in at least 10 of the 13 images in both datasets (baseline and Growth2). Therefore, classifying the sheep based on a majority of 13 images, we can assume a 100% success rate in identifying each sheep daily, throughout the growing period tested, which corresponds to the growing period of young sheep raised on commercial farms.

An advantage of the proposed identification system is that it can be installed in a commercial sheep pen - cameras can be altered positioned at different angles and distances from the sheep.

7 Conclusions and future work

7.1 Conclusions

7.1.1 Mobile system for detecting and counting laying hens

The mobile system developed to count hens held in community battery cages reached detection accuracies of 88% with a MAE of 4.56 hens per cage. This MAE was achieved by using Faster R-CNN with a single class ("hen"), and using a tracking algorithm which used successive frames of a video acquired while moving along the hen house.

7.1.2 Automated system for sheep identification

A biometric identification model for individual sheep recognition based on machine vision and advanced CNNs was designed, developed and implemented. The average identification accuracy achieved on two different groups of similarly-aged sheep was 95% accuracy. Transfer learning methods improved accuracy by an average of 1.85%, in a shorter training process, resulting in an average of 97% identification accuracy. It must be noted that these reported accuracies are the imaging identification accuracies; sheep were detected with 100% correct identification (when a minimum of 10 images were used).

The model was implemented using Faster-RCNN algorithm and a ResNet50V2 model with the ArcFace loss function. The training process should include at least 35 images per sheep, while classification should be calculated by majority decision on at least ten randomly chosen frames per visit per sheep, from a full video.

7.2 Future Work

7.2.1 Mobile system for detecting and counting laying hens

In order to improve counting accuracy and to decrease the wrong detection of hens from neighbouring cages, a better understanding of cage boundaries must be achieved. Therefore, future work should include more information from the scene by inserting the depth channel in the training process. Improved results will also be obtained by developing an autonomous mobile cart; as aforementioned, the hens were disturbed by the human moving the cart which can be avoided. In further research, the ability of the newly developed system to detect and track hens can be valuable for PLF applications, to develop an animal behaviour model, find the dominant hens in a cage, and detect a hen that does not eat. With this information, hens' welfare may be improved along with improved management.

7.2.2 Automated system for sheep identification

As sensory systems improve in quality, and computer systems increase in computational power more innovations and improvements should be introduced to these systems.

Future research may focus on developing a real-time system, automating all steps for reliable operation in a livestock environment. In addition, it is suggested to develop an unsupervised learning model so that this model can be easily adapted to unknown new herds. In an unsupervised learning algorithm each sheep should be identified in an unlabeled data set based on the underlying features in the data. This may shorten the data collection and data pre-processing because there will be no need to collect videos of the ear-tags which are used as ground truth and therefore no need for manual tagging.

In the future, the model identification ability can be improved using new features such as legs or tail identification. Furthermore, additional information such as the weight of the sheep, the height and the amount of water the sheep drinks can be included in the identification model. The recommended model may be adapted to identify other animals, such as beef cattle and pigs, replacing traditional identification methods.

8 References

Ait-Saidi, A., Caja, G., Salama, A.A.K., Carné, S., 2014. Implementing electronic identification for performance recording in sheep: I. Manual versus semiautomatic and automatic recording systems in dairy and meat farms. J. Dairy Sci. 97, 7505–7514. https://doi.org/10.3168/jds.2014-8090

Al-Waisy, A.S., Qahwaji, R., Ipson, S., Al-Fahdawi, S., 2015. A robust face recognition system based on Curvelet and Fractal dimension transforms, in: Proceedings - 15th IEEE International Conference on Computer and Information Technology, CIT 2015, 14th IEEE International Conference on Ubiquitous Computing and Communications, IUCC 2015, 13th IEEE International Conference on Dependable, Autonomic and Se. https://doi.org/10.1109/CIT/IUCC/DASC/PICOM.2015.78

Al-Waisy, A.S., Qahwaji, R., Ipson, S., Al-Fahdawi, S., Nagem, T.A.M., 2018. A multi-biometric iris recognition system based on a deep learning approach. Pattern Anal. Appl. 21, 783–802. https://doi.org/10.1007/s10044-017-0656-1

Ali, I., Cawkwell, F., Dwyer, E., Green, S., 2017. Modeling Managed Grassland Biomass Estimation by Using Multitemporal Remote Sensing Data-A Machine Learning Approach. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. https://doi.org/10.1109/JSTARS.2016.2561618

Alston, J.M., 1999. Quota reforms in Australian agriculture. Can. J. Agric. Econ. 47, 145–147. https://doi.org/10.1111/j.1744-7976.1999.tb00253.x

Amraei, S., Abdanan Mehdizadeh, S., Salari, S., 2017. Broiler weight estimation based on machine vision and artificial neural network. Br. Poult. Sci. 58, 200–205. https://doi.org/10.1080/00071668.2016.1259530

Andrew, H., Campbell, B., 2017. Automatic individual holstein friesian cattle identification via selective local coat pattern matching in RGB-D imagery 484–488.

Andrew, W., Greatwood, C., Burghardt, T., 2019. Aerial Animal Biometrics: Individual Friesian Cattle Recovery and Visual Identification via an Autonomous UAV with Onboard Deep Inference 2–8.

Appleby, M.C., 2003. The European Union ban on conventional cages for laying hens: History and prospects. J. Appl. Anim. Welf. Sci. 6, 103–121. https://doi.org/10.1207/S15327604JAWS0602_03

Banhazi, T.M., H, L., L, B.J., H, C., P, S., M, T., D, B., 2012. Precision Livestock Farming: An international review of scientific and commercial aspects. Int. J. Agric. Biol. Eng. 5, 0–9. https://doi.org/10.3965/j.ijabe.20120503.00

Bargoti, S., Underwood, J., 2017. Deep fruit detection in orchards. Proc. - IEEE Int. Conf. Robot. Autom. 3626–3633. https://doi.org/10.1109/ICRA.2017.7989417

Bennett, C.E., Thomas, R., Williams, M., Zalasiewicz, J., Edgeworth, M., Miller, H., Coles, B., Foster, A., Burton, E.J., Marume, U., 2018. The broiler chicken as a signal of a human reconfigured biosphere. R. Soc. Open Sci. 5. https://doi.org/10.1098/rsos.180325

Bessei, W., 2006. Welfare of broilers: a review. Worlds. Poult. Sci. J. 62, 455–466. https://doi.org/10.1079/wps2005108

Bloch, V., Levit, H., Halachmi, I., 2019. Assessing the potential of photogrammetry to monitor feed intake of dairy cows.

Bouvarel, 2011. Learn more about Laying Hens Hen nutrition for sustained egg quality Preharvest Measures to Improve the Safety of Eggs.

Bowler, I.R., 1994. The institutional regulation of uneven development: the case of poultry production in the province of Ontario. Trans. - Inst. Br. Geogr. 19, 346–358. https://doi.org/10.2307/622327

Caja, G., Ghirardi, J., Hernández-Jover, M., Garín, D., 2004. Diversity of animal identification techniques: from 'fire age'to 'electronic age.' ICAR Tech. Ser.

Calefati, A., Janjua, M.K., Nawaz, S., Gallo, I., 2019. Git loss for deep face recognition. Br. Mach. Vis. Conf. 2018, BMVC 2018 1–12.

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. Comput. Electron. Agric. 64, 53–60. https://doi.org/10.1016/j.compag.2008.05.014

Cao, Q., Shen, L., Xie, W., Parkhi, O.M., Zisserman, A., 2018. VGGFace2: A dataset for recognising faces across pose and age. Proc. - 13th IEEE Int. Conf. Autom. Face Gesture Recognition, FG 2018 67–74. https://doi.org/10.1109/FG.2018.00020

Chen, Deng, Du, 2017a. Noisy softmax: Improving the generalization ability of DCNN via postponing the early softmax saturation. Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017 2017-Janua, 4021–4030. https://doi.org/10.1109/CVPR.2017.428

Chen, Shivakumar, Dcunha, Das, J., Okon, E., Qu, C., Taylor, C.J., Kumar, V., 2017b. Counting Apples and Oranges with Deep Learning: A Data-Driven Approach. IEEE Robot. Autom. Lett. 2, 781–788. https://doi.org/10.1109/LRA.2017.2651944

Chung, Y., Choi, D., Choi, H., Park, D., Chang, H.H., Kim, S., 2015. Automated detection of cattle mounting using side-view camera. KSII Trans. Internet Inf. Syst. 9, 3160–3177. https://doi.org/10.3837/tiis.2015.08.024

Conington, J., Bishop, S.C., Lambe, N.R., Bünger, L., Simm, G., 2006. Testing selection indices for sustainable hill sheep production - Lamb growth and carcass traits. Anim. Sci. https://doi.org/10.1079/ASC200663

Cooper, J.J., Appleby, M.C., 1996. Demand for nest boxes in laying hens. Behav. Processes 36, 171–182. https://doi.org/10.1016/0376-6357(95)00027-5

Corkery, G., Gonzales-Barron, U.., Butler, F., Mcdonnell, K., Ward, S., 2007. A preliminary investigation on face recognition as a biometric identifier of sheep. Trans. ASABE.

Corkery, G., Ward, S., Kenny, C., Hemmingway, P., 2013. Incorporating Smart Sensing Technologies into the Poultry Industry. J. World's Poult. Res. 3, 106–128.

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. Aust. J. Exp. Agric. 48, 768–772. https://doi.org/10.1071/EA07404

Daniel Berckmans, 2004. AUTOMATIC ON-LINE MONITORING OF ANIMALS BY PRECISION LIVESTOCK FARMING (application/pdf オブジェクト) 51–54.

Dawkins, M.S., 2017. Animal welfare and efficient farming: Is conflict inevitable? Anim. Prod. Sci. 57, 201–208. https://doi.org/10.1071/AN15383

Defra, 2013. Code of Recommendations for the Welfare of Livestock: Cattle, Code of Recommendations for the Welfare of Livestock: Cattle.

Deng, Chen, Fang, Hu, 2017. Deep Correlation Feature Learning for Face Verification in the Wild. IEEE Signal Process. Lett. 24, 1877–1881. https://doi.org/10.1109/LSP.2017.2726105

Deng, J., Guo, J., Xue, N., Zafeiriou, S., 2019. ArcFace: Additive angular margin loss for deep face recognition. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2019-June, 4685–4694. https://doi.org/10.1109/CVPR.2019.00482

Deng, Zhou, Zafeiriou, 2017. Marginal Loss for Deep Face Recognition. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work. 2017-July, 2006–2014. https://doi.org/10.1109/CVPRW.2017.251

Domdouzis, K., Kumar, B., Anumba, C., 2007. Radio-Frequency Identification (RFID) applications: A brief introduction. Adv. Eng. Informatics. https://doi.org/10.1016/j.aei.2006.09.001

Dong, R., Pan, X., Li, F., 2019. DenseU-Net-Based Semantic Segmentation of Small Objects in Urban Remote Sensing Images. IEEE Access 7, 65347–65356. https://doi.org/10.1109/ACCESS.2019.2917952

Druzhkov, P.N., Kustikova, V.D., 2016. A survey of deep learning methods and software tools for image classification and object detection. Pattern Recognit. Image Anal. 26, 9–15. https://doi.org/10.1134/S1054661816010065

Dudi Hendrawan, 2005. Penentuan Pola Subsidi dan Sistem Distribusi Pupuk di Indonesia. Progr. Pascasarj. Manaj. dan Bisnis IPB 33, 189–196. https://doi.org/10.20506/rst.33.1.2273

Dyrmann, M., Mortensen, A.K., Midtiby, H.S., Jørgensen, R.N., 2016. Pixel-wise classification of weeds and crops in images by using a fully convolutional neural network, in: Proceedings of the International Conference on Agricultural Engineering, Aarhus, Denmark. pp. 26–29.

Floyd, F., Sabins, J.R., 1987. Remote sensing principles and interpretation. 2nd, Ed. 1–12.

Fogarty, E.S., Swain, D.L., Cronin, G., Trotter, M., 2018. Autonomous on-animal sensors in sheep research: A systematic review. Comput. Electron. Agric. https://doi.org/10.1016/j.compag.2018.04.017

Fontana, I., Tullo, E., Butterworth, A., Guarino, M., 2015. An innovative approach to predict the growth in intensive poultry farming. Comput. Electron. Agric. 119, 178–183. https://doi.org/10.1016/j.compag.2015.10.001

Freund, Y., Schapire, R.E., 1995. A decision-theoretic generalization of on-line learning and an application to boosting. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 904, 23–37.

Fuchs, B., Sørheim, K.M., Chincarini, M., Brunberg, E., Stubsjøen, S.M., Bratbergsengen, K., Hvasshovd, S.O., Zimmermann, B., Lande, U.S., Grøva, L., 2019. Heart rate sensor validation and seasonal and diurnal variation of body temperature and heart rate in domestic sheep. Vet. Anim. Sci. https://doi.org/10.1016/j.vas.2019.100075

Gecer, B., Balntas, V., Kim, T.K., 2017. Learning Deep Convolutional Embeddings for Face Representation Using Joint Sample- and Set-Based Supervision. Proc. - 2017 IEEE Int. Conf. Comput. Vis. Work. ICCVW 2017 2018-Janua, 1665–1672. https://doi.org/10.1109/ICCVW.2017.195

Geffen, O., Yitzhaky, Y., Barchilon, N., Druyan, S., Halachmi, I., 2019. Developing a machine vision system for detecting laying hens.

Girshick, R., 2015. Fast R-CNN. Proc. IEEE Int. Conf. Comput. Vis. 2015 Inter, 1440–1448. https://doi.org/10.1109/ICCV.2015.169

Gongal, A., Amatya, S., Karkee, M., Zhang, Q., Lewis, K., 2015. Sensors and systems for fruit detection and localization: A review. Comput. Electron. Agric. 116, 8–19. https://doi.org/10.1016/j.compag.2015.05.021

Groover, M.P., 2007. Automation, production systems, and computer-integrated manufacturing. Prentice Hall Press.

Guarino, M., Jans, P., Costa, A., Aerts, J.M., Berckmans, D., 2008. Field test of algorithm for automatic cough detection in pig houses. Comput. Electron. Agric. 62, 22–28. https://doi.org/10.1016/j.compag.2007.08.016

Guo, G., Zhang, N., 2019. A survey on deep learning based face recognition. Comput. Vis. Image Underst. 189, 102805. https://doi.org/10.1016/j.cviu.2019.102805

Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., Lew, M.S., 2016. Deep learning for visual understanding: A review. Neurocomputing 187, 27–48. https://doi.org/10.1016/j.neucom.2015.09.116

Guo, Y., Shi, H., Kumary, A., Grauman, K., Rosing, T., Feris, R., 2018. Spottune: Transfer learning through adaptive fine-tuning. arXiv 4805–4814.

Halachmi, I., Guarino, M., 2016. Editorial: Precision livestock farming: A “per animal” approach using advanced monitoring technologies. Animal 10, 1482–1483. https://doi.org/10.1017/S1751731116001142

Halachmi, I., Guarino, M., Bewley, J., Pastell, M., 2019. Smart Animal Agriculture : Application of Real-Time Sensors to Improve Animal Well-Being and Production. Annu. Rev. Anim. Biosci.

Hamilton, D.N., Ellis, M., Bertol, T.M., Miller, K.D., 2004. Effects of handling intensity and live weight on blood acid-base status in finishing pigs. J. Anim. Sci. 82, 2405–2409.

Hansen, M.F., Smith, M.L., Smith, L.N., Salter, M.G., Baxter, E.M., Farish, M., Grieve, B., 2018. Towards on-farm pig face recognition using convolutional neural networks. Comput. Ind. 98, 145–152. https://doi.org/10.1016/j.compind.2018.02.016

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

Hemeryck, M., Berckmans, D., Vranken3, E., Tullo, E., Fontana, I., Guarino, M., Waterschoot, Van, T., 2015. The Pig Cough Monitor in the EU-PLF project: results and multimodal data analysis in two case studies. Precis. Livest. Farming 2015 - Pap. Present. 7th Eur. Conf. Precis. Livest. Farming, ECPLF 2015.

Hou, J., He, Y., Yang, H., Connor, T., Gao, J., Wang, Y., Zeng, Y., Zhang, J., Huang, J., Zheng, B., Zhou, S., 2020. Identification of animal individuals using deep learning: A case study of giant panda. Biol. Conserv. 242. https://doi.org/10.1016/j.biocon.2020.108414

Huang, T.S., 1997. Computer Vision: Evolution and Promise. Report.

Huh, M., Agrawal, P., Efros, A.A., 2016. eprint arXiv:1608.08614. What makes ImageNet good Transf. Learn.

Jaiswal, S., 2011. Biometric: Case Study. J. Glob. Res. Comput. Sci. 2, 49.

Jiang, H., Learned-Miller, E., 2017. Face Detection with the Faster R-CNN. Proc. - 12th IEEE Int. Conf. Autom. Face Gesture Recognition, FG 2017 - 1st Int. Work. Adapt. Shot Learn. Gesture Underst. Prod. ASL4GUP 2017, Biometrics Wild, Bwild 2017, Heteroge 650–657. https://doi.org/10.1109/FG.2017.82

Kalantar, A., 2019. Melon yield estimation using UAV images and deep learning.

Kaler, J., Green, L.E., 2013. Sheep farmer opinions on the current and future role of veterinarians in flock health management on sheep farms: A qualitative study. Prev. Vet. Med. 112, 370–377. https://doi.org/10.1016/j.prevetmed.2013.09.009

Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: A survey. Comput. Electron. Agric. https://doi.org/10.1016/j.compag.2018.02.016

Kapach, K., Barnea, E., Mairon, R., Edan, Y., Ben-Shahar, O., 2012. Computer vision for fruit harvesting robots - State of the art and challenges ahead. Int. J. Comput. Vis. Robot. https://doi.org/10.1504/IJCVR.2012.046419

Kashiha, M.A., Green, A.R., Sales, T.G., Bahr, C., Berckmans, D., Gates, R.S., 2015. Application of image processing on hen tracking in an environmental preference chamber. Precis. Livest. Farming 2015 - Pap. Present. 7th Eur. Conf. Precis. Livest. Farming, ECPLF 2015 185–194.

Kashiha, M.A., Green, A.R., Sales, T.G., Bahr, C., Berckmans, D., Gates, R.S., 2014. Performance of an image analysis processing system for hen tracking in an environmental preference chamber 2439–2448.

Kestur, R., Meduri, A., Narasipura, O., 2019. MangoNet: A deep semantic segmentation architecture for a method to detect and count mangoes in an open orchard. Eng. Appl. Artif. Intell. 77, 59–69.

Khuda, B., 2007. Issues and Economics of Poultry Production: a Case Study of Faisalabad, Pakistan. Pak. Vet. J. 27, 25–28.

Klette, R., 2014. Concise computer vision. Springer.

Koik, B.T., Ibrahim, H., 2012. A Literature Survey on Animal Detection Methods in Digital Images. Int. J. Futur. Comput. Commun. 1, 24–25. https://doi.org/10.7763/ijfcc.2012.v1.7

Koirala, A., Walsh, K.B.., Wang, Z.., McCarthy, C.., 2019. Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of 'MangoYOLO.' Precis. Agric. https://doi.org/10.1007/s11119-019-09642-0

Kumar, S., Singh, S.K., 2017. Automatic identification of cattle using muzzle point pattern: a hybrid feature extraction and classification paradigm. Multimed. Tools Appl. 76, 26551–26580. https://doi.org/10.1007/s11042-016-4181-9

Kumar, Santosh, Pandey, A., Sai Ram Satwik, K., Kumar, Sunil, Singh, S.K., Singh, A.K., Mohan, A., 2018. Deep learning framework for recognition of cattle using muzzle point image pattern. Meas. J. Int. Meas. Confed. 116, 1–17. https://doi.org/10.1016/j.measurement.2017.10.064

Kung, H.Y., Kuo, T.H., Chen, C.H., Tsai, P.Y., 2016. Accuracy analysis mechanism for agriculture data using the ensemble neural network method. Sustain. https://doi.org/10.3390/su8080735

Laberge, B., Rousseau, A.N., 2017. Rethinking environment control strategy of confined animal housing systems through precision livestock farming. Biosyst. Eng. 155, 96–123. https://doi.org/10.1016/j.biosystemseng.2016.12.005

Lamb, N., Chuah, M.C., 2018. A Strawberry Detection System Using Convolutional Neural Networks, in: 2018 IEEE International Conference on Big Data (Big Data). pp. 2515–2520.

Landais, E., 2001. The marking of livestock in traditional pastoral societies. OIE Rev. Sci. Tech. https://doi.org/10.20506/rst.20.2.1286

Liakos, K.G., Busato, P., Moshou, D., Pearson, S., Bochtis, D., 2018. Machine learning in agriculture: A review. Sensors (Switzerland) 18, 1–29. https://doi.org/10.3390/s18082674

Liang, Q., Zhu, W., Long, J., Wang, Y., Sun, W., Wu, W., 2018. A Real-Time Detection Framework for On-Tree Mango Based on SSD Network, in: International Conference on Intelligent Robotics and Applications. pp. 423–436.

Lima, E., Hopkins, T., Gurney, E., Shortall, O., Lovatt, F., Davies, P., Williamson, G., Kaler, J., 2018. Drivers for precision livestock technology adoption: A study of factors associated with adoption of electronic identification technology by commercial sheep farmers in England and Wales. PLoS One 13, 1–17. https://doi.org/10.1371/journal.pone.0190489

Liu, Y., Li, H., Wang, X., 2017. Rethinking Feature Discrimination and Polymerization for Large-scale Recognition.

Loshchilov, I., Hutter, F., 2017. SGDR: Stochastic gradient descent with warm restarts. 5th Int. Conf. Learn. Represent. ICLR 2017 - Conf. Track Proc. 1–16.

Mahmoud, M., Lu, Y., Hou, X., McLennan, K., Robinson, P., 2018. Estimation of Pain in Sheep Using Computer Vision, in: Handbook of Pain and Palliative Care. https://doi.org/10.1007/978-3-319-95369-4_9

Marchewka, J., Watanabe, T.T.N., Ferrante, V., Estevez, I., 2013. Welfare assessment in broiler farms: Transect walks versus individual scoring. Poult. Sci. 92, 2588–2599. https://doi.org/10.3382/ps.2013-03229

Mehdizadeh, S., Behmanesh, J., Khalili, K., 2017. Using MARS, SVM, GEP and empirical equations for estimation of monthly mean reference evapotranspiration. Comput. Electron. Agric. https://doi.org/10.1016/j.compag.2017.05.002

Meluzzi, A., Sirri, F., 2009. Welfare of broiler chickens. Ital. J. Anim. Sci. 8, 161–173. https://doi.org/10.4081/ijas.2009.s1.161

Mertens, K., Decuypere, E., De Baerdemaeker, J., De Ketelaere, B., 2011. Statistical control charts as a support tool for the management of livestock production. J. Agric. Sci. 149, 369–384. https://doi.org/10.1017/S0021859610001164

Minaee, S., Abdolrashidi, A., Su, H., Bennamoun, M., Zhang, D., 2019. Biometric Recognition Using Deep Learning: A Survey.

Mollo, M.N., Vendrametto, O., Okano, M.T., 2010. Precision livestock tools to improve products and processes in broiler production: A review. Rev. Bras. Cienc. Avic. 11, 211–218.

Morellos, A., Pantazi, X.E., Moshou, D., Alexandridis, T., Whetton, R., Tziotzios, G., Wiebensohn, J., Bill, R., Mouazen, A.M., 2016. Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. Biosyst. Eng. https://doi.org/10.1016/j.biosystemseng.2016.04.018

Morgan-davies, C., Lambe, N., Wishart, H., Waterhouse, T., Kenyon, F., Mcbean, D., Mccracken, D., 2018. Impacts of using a precision livestock system targeted approach in mountain sheep fl ocks. Livest. Sci. 208, 67–76. https://doi.org/10.1016/j.livsci.2017.12.002

Morris, 2004. Computer Vision and Image Processing (Cornerstones of Computing). Palgrave Macmillan Limited.

Morris, J.E., Cronin, G.M., Bush, R.D., 2012. Improving sheep production and welfare in extensive systems through precision sheep management. Anim. Prod. Sci. https://doi.org/10.1071/AN11097

Mortensen, A.K., Lisouski, P., Ahrendt, P., 2016. Weight prediction of broiler chickens using 3D computer vision. Comput. Electron. Agric. 123, 319–326. https://doi.org/10.1016/j.compag.2016.03.011

Nakarmi, A.D., Tang, L., Xin, H., 2014. Automated Tracking and Behavior Quantification of

Laying Hens Using 3D Computer Vision and Radio Frequency Identification Technologies. Trans. ASABE 57, 1455–1472. https://doi.org/10.13031/trans.57.10505

Neila, B.S., Xavier, A., Inma, E., 2016. Technology and Poultry Welfare. Animals 62–68. https://doi.org/10.3390/ani6100062

Noldus, L.P.J.J., Spink, A.J., Tegelenbosch, R.A.J., 2001. Ethovision Video Tracking System. Behav. Res. Methods, Instruments, Comput. 33, 398–414. https://doi.org/10.3758/BF03195394

Olivo, R.F., Thompson, M.C., 1988. Monitoring animals' movements using digitized video images. Behav. Res. Methods, Instruments, Comput. 20, 485–490. https://doi.org/10.3758/BF03202701

Omatu, S., Wada, T., Rodríguez, S., Chamoso, P., Corchado, J.M., 2014. Multi-agent Technology to Perform Odor Classification Case Study : Development of a VO for Odor Classification 241–252. https://doi.org/10.1007/978-3-319-07596-9

Park, B., Chen, Y.R., 2000. Real-time dual-wavelength image processing for poultry safety inspection. J. Food Process Eng. 23, 329–351. https://doi.org/10.1111/j.1745-4530.2000.tb00519.x

Parkhi, O.M., Vedaldi, A., Zisserman, A., 2015. Deep Face Recognition 41.1-41.12. https://doi.org/10.5244/c.29.41

Parmar, D.N., Mehta, B.B., 2014. Face Recognition Methods & Applications 4, 84–86.

Pound, M.P., Atkinson, J.A., Townsend, A.J., Wilson, M.H., Griffiths, M., Jackson, A.S., Bulat, A., Tzimiropoulos, G., Wells, D.M., Murchie, E.H., others, 2017. Deep machine learning provides state-of-the-art performance in image-based plant phenotyping. Gigascience 6, gix083.

Puma, M.C., Xin, H., Gates, R.S., Burnham, D.J., 2001. An instrumentation system for studying feeding and drinking behavior of individual poultry. Appl. Eng. Agric. 17, 365–374.

Qi, C., Su, F., 2018. Contrastive-center loss for deep neural networks. Proc. - Int. Conf. Image Process. ICIP 2017-Septe, 2851–2855. https://doi.org/10.1109/ICIP.2017.8296803

Qiao, Y., Truman, M., Sukkarieh, S., 2019. Cattle segmentation and contour extraction based on Mask R-CNN for precision livestock farming. Comput. Electron. Agric. 165, 104958. https://doi.org/10.1016/j.compag.2019.104958

Rahimzadeh, M., Attar, A., 2020. A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. Informatics Med. Unlocked 19, 100360. https://doi.org/10.1016/j.imu.2020.100360

Rahnemoonfar, M., Sheppard, C., 2017. Deep count: Fruit counting based on deep simulated learning. Sensors (Switzerland) 17, 1–12. https://doi.org/10.3390/s17040905

Rajsic, P., Fox, G., 2017. Quota prices as indicators of comparative advantage in supply-managed industries. Agric. Econ. (United Kingdom). https://doi.org/10.1111/agec.12323

Ranjan, R., Castillo, C.D., Chellappa, R., 2017. L2-constrained Softmax Loss for Discriminative

Face Verification.

Rawat, W., Wang, Z., 2017. Deep convolutional neural networks for image classification: A comprehensive review. Neural Comput. https://doi.org/10.1162/NECO_a_00990

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

Rokkones, A.S., 2018. Real-Time Face Detection , Identification and Emotion Analysis for Affective Human Robot Interaction.

Rowe, E., Dawkins, M.S., 2019. A Systematic Review of Precision Livestock Farming in the Poultry Sector : Is Technology Focussed on Improving Bird Welfare ? 1–18.

Ruiz-Garcia, L., Lunadei, L., Barreiro, P., Robla, J.I., 2009. A review of wireless sensor technologies and applications in agriculture and food industry: State of the art and current trends. Sensors (Switzerland) 9, 4728–4750. https://doi.org/10.3390/s90604728

Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T., McCool, C., 2016. Deepfruits: A fruit detection system using deep neural networks. Sensors (Switzerland) 16. https://doi.org/10.3390/s16081222

Salama, A., Hassanien, A.E., Fahmy, A., 2019. Sheep Identification Using a Hybrid Deep Learning and Bayesian Optimization Approach. IEEE Access 7, 31681–31687. https://doi.org/10.1109/ACCESS.2019.2902724

Sassi, N. Ben, Averós, X., Estevez, I., 2016. Technology and poultry welfare. Animals 6, 1–21. https://doi.org/10.3390/ani6100062

Savory, C.J., 2004. Laying hen welfare standards: A classic case of "power to the people." Anim. Welf. 13.

Sawant, M.M., Bhurchandi, K.M., 2019. Age invariant face recognition: a survey on facial aging databases, techniques and effect of aging. Artif. Intell. Rev. 52, 981–1008. https://doi.org/10.1007/s10462-018-9661-z

Schilling, B., Bahmani, K., Li, B., Banerjee, S., Smith, J.S., Moshier, T., Schuckers, S., 2019. Validation of biometric identification of dairy cows based on udder NIR images. 2018 IEEE 9th Int. Conf. Biometrics Theory, Appl. Syst. BTAS 2018. https://doi.org/10.1109/BTAS.2018.8698553

Scholten, M.C.T., De Boer, I.J.M., Gremmen, B., Lokhorst, C., 2013. Livestock Farming with Care: Towards sustainable production of animal-source food. NJAS - Wageningen J. Life Sci. 66, 3–5. https://doi.org/10.1016/j.njas.2013.05.009

Schroff, F., Kalenichenko, D., Philbin, J., 2015. FaceNet: A unified embedding for face recognition and clustering. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 07-12-June, 815–823. https://doi.org/10.1109/CVPR.2015.7298682

Sergeant, D., Boyle, R., Forbes, M., 1998. Computer visual tracking of poultry. Comput. Electron. Agric. 21, 1–18. https://doi.org/10.1016/S0168-1699(98)00025-8

Shimmura, T., Eguchi, Y., Uetake, K., Tanaka, T., 2007. Differences of behavior, use of resources and physical conditions between dominant and subordinate hens in furnished cages. Anim. Sci. J. 78, 307–313. https://doi.org/10.1111/j.1740-0929.2007.00440.x

Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition. 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc. 1–14.

Suktanarak, S., Teerachaichayut, S., 2017. Non-destructive quality assessment of hens' eggs using hyperspectral images. J. Food Eng. 215, 97–103. https://doi.org/10.1016/j.jfoodeng.2017.07.008

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the Inception Architecture for Computer Vision. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2016-Decem, 2818–2826. https://doi.org/10.1109/CVPR.2016.308

Szeliski, R., 2010. Computer Vision: Algorithms and Applications, Texts in Computer Science. Springer London.

Talo, M., Yildirim, O., Baloglu, U.B., Aydin, G., Acharya, U.R., 2019. Convolutional neural networks for multi-class brain disease detection using MRI images. Comput. Med. Imaging Graph. 78, 101673. https://doi.org/10.1016/j.compmedimag.2019.101673

Tan, M., Le, Q. V., 2019. EfficientNet: Rethinking model scaling for convolutional neural networks. 36th Int. Conf. Mach. Learn. ICML 2019 2019-June, 10691–10700.

Tharwat, A., Gaber, T., Hassanien, A.E., Hassanien, H.A., Tolba, M.F., 2014. Cattle Identification Using Muzzle Print Images Based on Texture Features Approach. Adv. Intell. Syst. Comput. 303, 217–227. https://doi.org/10.1007/978-3-319-08156-4_22

Tillett, R.D., 1991. Image analysis for agricultural processes: a review of potential opportunities. J. Agric. Eng. Res. https://doi.org/10.1016/S0021-8634(05)80018-6

Trevarthen, A., 2007. The national livestock identification system: The importance of traceability in E-business. J. Theor. Appl. Electron. Commer. Res. 2, 49–62.

Tsai, D.M., Huang, C.Y., 2014. A motion and image analysis method for automatic detection of estrus and mating behavior in cattle. Comput. Electron. Agric. 104, 25–31. https://doi.org/10.1016/j.compag.2014.03.003

Tullo, E., Fontana, I., Diana, A., Norton, T., Berckmans, D., Guarino, M., 2017. Application note: Labelling, a methodology to develop reliable algorithm in PLF. Comput. Electron. Agric. 142, 424–428. https://doi.org/10.1016/j.compag.2017.09.030

Umstatter, C., Morgan-Davies, C., Stevens, H., Kenyon, F., McBean, D., Lambe, N., Waterhouse, A., 2013. Integrating electronic identification into hill sheep management, in: Precision Livestock Farming 2013 - Papers Presented at the 6th European Conference on Precision Livestock Farming, ECPLF 2013.

Vaarst, M., Alrøe, H.F., 2012. Concepts of animal health and welfare in organic livestock systems. J. Agric. Environ. Ethics 25, 333–347. https://doi.org/10.1007/s10806-011-9314-6

Villeneuve, É., Akle, A.A., Merlo, C., Masson, D., Terrasson, G., Llaria, A., 2019. Decision support in precision sheep farming. IFAC-PapersOnLine 51, 236–241.

https://doi.org/10.1016/j.ifacol.2019.01.048

Viola, P., Jones, M., others, 2001. Rapid object detection using a boosted cascade of simple features. CVPR 1, 511–518.

Vlad, M., Parvulet, R.A., Vlad, M.S., 2012. A survey of livestock identification systems. Proc. 13th World Sci. Eng. Acad. Soc. Int. Conf. Autom. Information, Inst. Chart. Accountants India 165–170.

Voulodimos, A.S., Patrikakis, C.Z., Sideridis, A.B., Ntafis, V.A., Xylouri, E.M., 2010. A complete farm management system based on animal identification using RFID technology. Comput. Electron. Agric. 70, 380–388. https://doi.org/10.1016/j.compag.2009.07.009

Wada, N., Shinya, M., Shiraishi, M., 2013. Pig face recognition using eigenspace method. ITE Trans. Media Technol. Appl. 1, 328–332. https://doi.org/10.3169/mta.1.328

Wang, C., Chen, H., Zhang, X., Meng, C., 2016. Evaluation of a laying-hen tracking algorithm based on a hybrid support vector machine. J. Anim. Sci. Biotechnol. 7, 1–10. https://doi.org/10.1186/s40104-016-0119-3

Wang, H., Wang, Y., Zhou, Z., Ji, X., Gong, D., Zhou, J., Li, Z., Liu, W., 2018. CosFace: Large Margin Cosine Loss for Deep Face Recognition. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 5265–5274. https://doi.org/10.1109/CVPR.2018.00552

Wang, Liu, Dai, Liu, Cheng, 2018. Additive margin softmax for face verification. 6th Int. Conf. Learn. Represent. ICLR 2018 - Work. Track Proc.

Wang, Y.-Q., 2014. An Analysis of the Viola-Jones Face Detection Algorithm. Image Process. Line 4, 128–148. https://doi.org/10.5201/ipol.2014.104

Weeks, C.A., Nicol, C.J., 2006. Behavioural needs, priorities and preferences of laying hens. Worlds. Poult. Sci. J. 62, 296–307. https://doi.org/10.1079/wps200598

Weiyang, Y., 2017. SphereFace: Deep Hypersphere Embedding for Face Recognition Weiyang. Therapeutique 47, 517–518.

Werkheiser, I., 2018. Precision Livestock Farming and Farmers' Duties to Livestock. J. Agric. Environ. Ethics 31, 181–195. https://doi.org/10.1007/s10806-018-9720-0

Wilhelmsson, S., Yngvesson, J., Jönsson, L., Gunnarsson, S., Wallenbeck, A., 2019. Welfare Quality® assessment of a fast-growing and a slower-growing broiler hybrid, reared until 10 weeks and fed a low-protein, high-protein or mussel-meal diet. Livest. Sci. 219, 71–79. https://doi.org/10.1016/j.livsci.2018.11.010

Wishart, H.M., 2019. Precision livestock farming : potential application for sheep systems in harsh environments PhD Thesis submitted to the University of Edinburg.

Wójcik, W., Sikora, J., 2017. Recent advances in information technology. Recent Adv. Inf. Technol. 1–193. https://doi.org/10.1201/9781351243179

Wolfert, S., Ge, L., Verdouw, C., Bogaardt, M.J., 2017. Big Data in Smart Farming – A review. Agric. Syst. 153, 69–80. https://doi.org/10.1016/j.agsy.2017.01.023

XIONG, J., LIU, Z., TANG, L., LIN, R., BU, R., PENG, H., 2018. Visual Detection Technology of

Green Citrus under Natural Environment. Trans. Chinese Soc. Agric. Mach. 5.

Xu, Z.Q.J., Zhang, Y., Xiao, Y., 2019. Training behavior of deep neural network in frequency domain. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 11953 LNCS, 264–274. https://doi.org/10.1007/978-3-030-36708-4_22

Yamazaki, M., Kasagi, A., Tabuchi, A., Honda, T., Miwa, M., Fukumoto, N., Tabaru, T., Ike, A., Nakashima, K., 2019. Yet Another Accelerated SGD: ResNet-50 Training on ImageNet in 74.7 seconds. arXiv.

Yi, D., Lei, Z., Liao, S., Li, S.Z., 2014. Learning Face Representation from Scratch.

Zaninelli, M., Redaelli, V., Luzi, F., Mitchell, M., Bontempo, V., Cattaneo, D., Dell'Orto, V., Savoini, G., 2018. Development of a machine vision method for the monitoring of laying hens and detection of multiple nest occupations. Sensors (Switzerland) 18, 1–19. https://doi.org/10.3390/s18010132

Zhang, Fang, Wen, Li, Qiao, 2017. Range Loss for Deep Face Recognition with Long-Tailed Training Data. Proc. IEEE Int. Conf. Comput. Vis. 2017-Octob, 5419–5428. https://doi.org/10.1109/ICCV.2017.578

Zhang, L., Gray, H., Ye, X., Collins, L., Allinson, N., 2019. Automatic individual pig detection and tracking in pig farms. Sensors (Switzerland) 19. https://doi.org/10.3390/s19051188

Zhang, Wu, Jiang, Xuan, Chuan, Ma, Zhang, Y.A., 2018. Development and validation of a visual image analysis for monitoring the body size of sheep. J. Appl. Anim. Res. 46, 1004–1015. https://doi.org/10.1080/09712119.2018.1450257

Zhang, Y., Cai, J., Xiao, D., Li, Z., Xiong, B., 2019. Real-time sow behavior detection based on deep learning. Comput. Electron. Agric. 163, 104884. https://doi.org/10.1016/j.compag.2019.104884

Zhao, B., Feng, J., Wu, X., Yan, S., 2017. A survey on deep learning-based fine-grained object classification and semantic segmentation. Int. J. Autom. Comput. 14, 119–135. https://doi.org/10.1007/s11633-017-1053-3

Zhong, Y., Zhao, M., 2020. Research on deep learning in apple leaf disease recognition. Comput. Electron. Agric. 168, 105146. https://doi.org/10.1016/j.compag.2019.105146

Zhuang, X., Bi, M., Guo, J., Wu, S., Zhang, T., 2018. Development of an early warning algorithm to detect sick broilers. Comput. Electron. Agric. 144, 102–113. https://doi.org/10.1016/j.compag.2017.11.032

9 Appendices

9.1 Mobile system for detecting and counting laying hens

Appendix 1A. Data collection and preparation code files
Data preparation code files (Volcani access only) - URL

Data_Preparation.txt

Appendix 1B. Collected data
** If needed, contact *Ilan Halachmi. Email: halachmi@volcani.agri.gov.il* for access
Collected data (Volcani access only) - URL

- *Raw data*

```
Raw_Data
-> Backward Data
---> Videos
---> Frames
------> Tagged Frames
------> Analyzed Data
-> Forward Data
---> Videos
---> Frames
------> Tagged Frames
------> Analyzed Data
```

Appendix 1C. Tagged images for algorithm development
** If needed, contact *Ilan Halachmi. Email: halachmi@volcani.agri.gov.il* for access
Tagged images for algorithm development (Volcani access only) - URL

Appendix 1D. Image processing algorithm
Image processing algorithm (Volcani access only) - URL

Image_Processing_Algorithm.txt

Appendix 1E. Object detection algorithm
Object detection algorithm (Volcani access only) - URL

Object_Detection_Algorithm.txt

Appendix 1F. Tracking algorithm (Geffen et al., 2019)

Tracking algorithm (Volcani access only) - URL

Tracking_Algorithm.txt

Appendix 1G. System counting result

System counting result (Volcani access only) - URL

Counting_Result.xlsx Counting_Comparition.xlsx

Appendix 1H. Sensitivity analysis code files and result

Sensitivity analysis code files and result (Volcani access only) - URL

Sensetivity_Evaluation.xlsx Sensetivity_Analysis.txt

9.2 Automated system for sheep identification

Appendix 2A. Biometric identification of sheep via machine-vision system

Biometric identification of sheep applying via machine-vision system

Submitted to: Computers and Electronics in Agriculture

Authors: Almog Hitelman, Yael Edan, Assaf Godo, Ron Berenstein, Joseph Lepar and Ilan Halachmi.

Biometric identification of sheep via a machine-vision system

A. Hitelman1,2, Y. Edan2, A. Godo1, R. Berenstein1, J. Lepar1, I. Halachmi1,2*

1 Precision livestock farming (PLF) Lab., Agricultural Engineering Inst., Agricultural Research Organization (A.R.O.) - Volcani Institute, 68 Hamaccabim Road, P.O.B 15159 Rishon Lezion 7505101, Israel
2 Dept. of Industrial Engineering and Management, Ben-Gurion University of the Negev, P.O.B 653 Be'er Sheva 8410501, Israel
Corresponding author: Ilan Halachmi. Email: halachmi@volcani.agri.gov.il

This paper describes a sheep biometric identification system based on facial images. A machine vision system and deep learning model were developed and applied for animal identification. The system included two 8-MegaPixels cameras installed in a monitoring drinking facility adapted to work with NVIDIA Jetson Nano embedded system-on-module (SoM). Data from 81 Assaf breed sheep aged two to three months, from two different groups of sheep, were collected over a period of two weeks. The biometric identification model included two steps: face detection and classification. In order to locate and localize the sheep face in an image, the Faster R-CNN deep learning object detection algorithm was applied. The detected face was provided as input to seven different classification models. Different transfer learning methods were examined. The best performance was obtained using a ResNet50V2 model with the state-of-art ArcFace loss function. The identification system resulted in average accuracies of 95.4% and 95.7% for the two groups tested. When applying transfer learning method, average identification accuracies improved to 97% in both groups, and the training process was accomplished more rapidly.

Keywords: Small ruminants, Biometric identification, Deep learning, convolutional neural network, Facial recognition.

INTRODUCTION

Collecting automated data on individual sheep can contribute to better animal handling and reduce labor requirements (Morris et al., 2012; Voulodimos et al., 2010). In addition, individual monitoring helps to manage disease and reduce risk of injury (Salama et al., 2019). Hence, automatically identifying (ID) each individual sheep is important (Halachmi and Guarino, 2016).

Traditional methods used for marking livestock with IDs include: branding, tattooing, ear tagging, and more recently electronic identification devices, such as RFID tags and barcodes (Ait-Saidi et al., 2014; Caja et al., 2004; Landais, 2001). However, those methods may stress the animal, and require frequent maintenance and cleaning (Ait-Saidi et al., 2014; Caja et al., 2004).

Recently, utilizing biometric traits instead of traditional methods for identifying individual animals has gained attention (Andrew et al., 2019; Halachmi et al., 2019; Hansen et al., 2018). Various biometric features can be used for livestock identification, including: facial imaging, matching muzzle patterns, coat patterns, mammary glands; and iris imaging (Corkery et al., 2007; Kumar and Singh, 2017). Nonetheless, facial recognition is promising since it contains many significant features (Hou et al., 2020; Salama et al., 2019).

Using the cosine distance classifier, Corkery et al., (2007), achieved 96% identification accuracy on 50 sheep, crossbreeds of Cheviot and Suffolk breeds aged three to four years. In this study, seven images of each sheep were manually acquired against a black background after cleaning the sheep's face. In another study (Salama et al., 2019) 98% identification accuracy was achieved with two different neural network models: Convolutional Neural Network (CNN) with parameters set automatically with a Bayesian optimizer, and with the AlexNet CNN. The CNNs were trained and tested on 52 Barqi breed sheep, between five months and five years old, where for each sheep, ten images were acquired.

Various CNNs have been developed for recognition tasks (Szeliski, 2010; Zhao et al., 2017) and specifically for biometric identification (Salama et al., 2019; Szeliski, 2010; Zhao et al., 2017). Significant contributions began in 2012, when AlexNet won the ImageNet recognition competition (Guo et al., 2016). Since then, many new architectures have been

developed to improve recognition accuracy, they include: VGG, ResNet and EfficientNet architectures (Guo and Zhang, 2019; Talo et al., 2019).
An integral part of the CNN design is the choice of an appropriate loss function, in order to enhance facial feature discrimination (Deng et al., 2019; H. Wang et al., 2018). In this paper, implementation of the state-of-art ArcFace loss function (Deng et al., 2019), which so far was only used for human facial recognition, was applied to animals.
The objective of the current study was to develop a sheep biometric identification system based on facial images.

MATERIALS AND METHODS

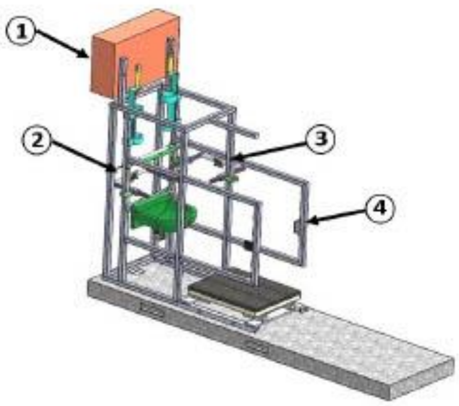
MACHINE VISION SYSTEM

The imaging setup was built on a drinking monitoring facility (Figure 1) to ensure that all sheep had frequent access per day, voluntarily, without human involvement. The drinking facility monitored each sheep's body weight and water intake per visit. Two 8-MegaPixels RGB cameras, of Digital Single Lens Reflex (DSLR) type, with USB connections were connected to a NVIDIA Jetson Nano embedded system-on-module (SoM). Both cameras video-recorded the sheep while they were drinking water. The cameras were located at a height of 80 cm, one at the face area; the second camera acquired photos of each sheep's ear tag (Figure 1). The system includes an Infrared Red (IR) sensor, which activates the cameras when a sheep inserts its head into the system area. The same IR sensor ends the recording the moment it no longer detects a sheep in the drinking facility (Figure 2). Similarly, if the sensors erroneously detect movement, e.g. a bird triggers them, they will immediately stop the cameras when the bird flies out. Videos were acquired at a speed of 30 frames per second (fps) from both cameras in parallel. The cameras and the Jetson Nano were placed in airtight boxes to protect them from dirt and heat. The system included a SIM card with an Internet network to enable remote connection via a USB dongle.

DATA COLLECTION

Two experiments were conducted in Volcani's Institute research sheep pen on two different groups of Assaf breed sheep. The first experiment, used for model development, was conducted on a group that included 47 sheep (group 1). The second experiment was conducted on 34 sheep (group 2) and was used to evaluate the model capability to learn new identities.
Measurements on group 1 were conducted in November 2020, when the sheep were two-three months old and weighed about 30 kg (average=33.6, SD=6.6). Group 2 data were recorded in April 2021, when the sheep were also two-three months old and weighed about 30 kg (average=32.5 SD=7.60). Each group's data was recorded for a period of two weeks, during which data were collected automatically throughout all daylight hours, to ensure a variety of illuminations. Each sheep arrived voluntarily two to three times per day, at different hours; their postures at the drinking facility differed, which resulted in a diverse database of sheep faces.

FIGURE I
DATA COLLECTION SETUP BUILT ON A CONTROLLED DRINKING FACILITY, INCLUDING; 1. NVIDIA JETSON NANO EMBEDDED SYSTEM-ON-MODULE (SOM), 2. FRONT CAMERA USED FOR FACIAL VIDEO-RECORDING, 3. SIDE CAMERA USED FOR RECORDING THE EAR TAG, 4. IR SENSOR.



DATA COLLECTION

Two datasets were created with the following steps: first, each sheep's face video (acquired from the front camera) was manually tagged with the sheep's ID, obtained from the corresponding side video. Then, each video file was converted with a Python3 code to an image sequence. Finally, 65 images of each sheep were manually selected for training and validation. The first dataset included 3055 images of group 1 and the second dataset included 2210 images of group 2. For each sheep, 52 images were randomly selected to train the models. The remaining 13 images were used for evaluation.
In addition, for the object detection algorithm, which

aims to locate the sheep's face in each image, a labeled dataset was created from 30 sheep videos. This dataset included a total of 400 images, manually tagged with a bounding box around each sheep's face.

Each training image (out of the 52 images) was augmented seven to eight times using the following augmentation techniques: vertical flip, random rotation up to 30 degrees, Gaussian noise addition, and adjusting image brightness (Figure 3). The techniques resulted in a total of 400 augmented images created per sheep, used to train the model.

FIGURE 2

FLOWCHART OF THE DATA COLLECTION SETUP.

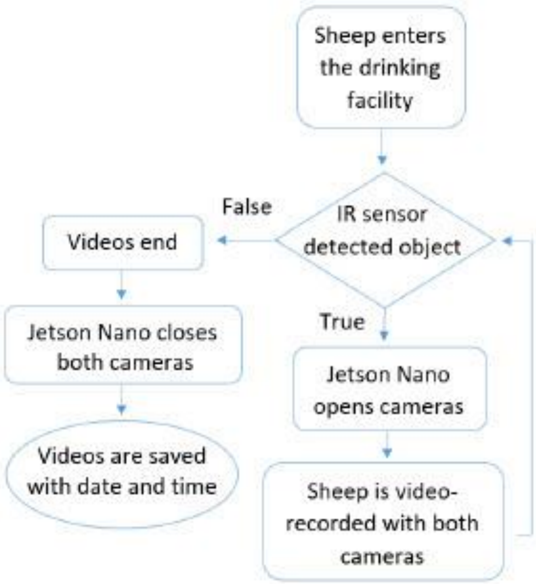


FIGURE 3

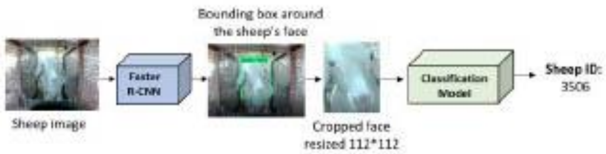
DATA AUGMENTATION EXAMPLES FROM LEFT TO RIGHT; ORIGINAL IMAGE, VERTICAL IMAGE, RANDOMLY ROTATED IMAGE, DARKENED IMAGE, BRIGHTENED IMAGE.

BIOMETRIC IDENTIFICATION MODEL

The biometric identification model included two steps – face detection and classification (Figure 4). In order to locate and localize the sheep's face in an image, the Faster R-CNN deep learning object detection algorithm was applied (Jiang and Learned-Miller, 2017). Then, the detected face was cropped, and resized to 112X112 pixels according to Deng et al., (2019). Finally, the cropped face was provided as input to the second step which included the classification model.

FIGURE 4

BIOMETRIC IDENTIFICATION MODEL SCHEMATIC FLOWCHART

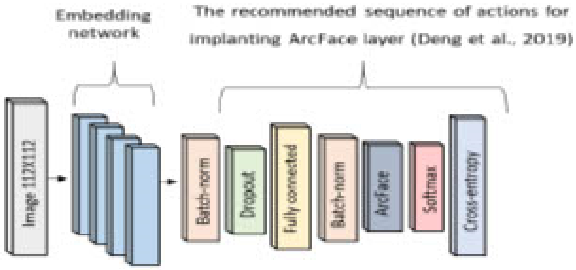


CLASSIFICATION MODEL

Seven different CNN architectures were compared. Each CNN examined was used as the embedding network for implementing ArcFace loss function, resulting in seven classification models (Figure 5). The recommended sequence of actions between the last layer in each CNN and the ArcFace layer was followed as per Deng et al. (2019) and implemented using Python open source Tensorflow and Keras. The architectures examined were: 1. AlexNet (Salama et al., 2019) 2. VGG16 (Simonyan and Zisserman, 2015) 3-4. ResNet50, ResNet50V2, 5. ResNet101V2 (Deng et al., 2019; Rahimzadeh and Attar, 2020) 6-7. EfficientNetB0, EfficientNetB3 (Tan and Le, 2019). All models were pre-trained on the ImageNet dataset, and were trained and tested on group 1 dataset, i.e. 47 sheep, 2444 and 611 randomly selected images included in the train and test sets respectively. The hyper parameters values were determined by trial and error on the ResNet50 network. All tested classification models were set with the same values, detailed in Appendix A. All images were normalized based on Wang et al., (2018) by zero-centering each color channel (red, green, blue), in order to prevent the Vanishing Gradient Problem. To adjust the learning rate during training, the Cosine Annealing Scheduler was used (Loshchilov and Hutter, 2017). An early stopping method was used to automatically stop the training process when the model's loss performance stopped improving on the validation set (Xu et al., 2019). This method helped to avoid overfitting of the neural network. Finally, a SGD optimizer was implemented, with momentum set to 0.85, to help accelerate gradient vectors in the right directions, thus leading to faster converging. The seven classification models were compared using the Post-hoc Tukey's statistical significance test.

FIGURE 5

SCHEMATIC DESCRIPTION OF THE CLASSIFICATION MODEL; THE INPUT IS A SHEEP FACIAL IMAGE 112X112 PIXELS IN SIZE. IN BLUE – THE EXAMINED CNN ARCHITECTURE, USED AS THE EMBEDDING NETWORK FOR IMPLEMENTING ARCFACE. THE LAST LAYER OF THE CNN IS CONNECTED TO THE ARCFACE LAYER ACCORDING TO THE RECOMMENDED SEQUENCE. THE OUTPUT IS THE LOSS OF THE HIGHEST PROBABILITY PREDICTED CLASS FOR THE IMAGE BY SOFTMAX.



MODEL EVALUATION AND ANALYSIS

PERFORMANCE MEASURES

Five performance measures (PM) as detailed in Table 1 were used. Categorical cross-entropy is the output of the ArcFace loss function and is used in multi-class classification tasks, as had been researched in this study. The cross-entropy is defined as the difference between the predicted value by the model and the true value (Table 1) wherein y_(i,j) denotes the true value and p_(i,j) denotes the probability predicted by the model of image i belonging to class j (Zhong and Zhao, 2020). The other four PMs are expressed through the calculated True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) for any given sheep (Dong et al., 2019).

TRAINING AND TESTING

The selected classification model was trained and evaluated on group 1 and group 2 separately, and on a unified group which included all the sheep. Additionally, transfer learning methods were examined to decrease the training time while maintaining a lower generalization error. The model was trained on one group of sheep and then retrained the convolution layers on the other group, with a faster learning rate.

Transfer learning methods were trained and evaluated on the same datasets, i.e. the test sets of both groups included 13 images per sheep; and with the same hyper parameters values detailed earlier but with a decreased learning rate of 0.0001.

TABLE 1

PERFORMANCE MEASURES.

Index	Formula
Categorical cross-entropy	$-\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} y_{i,j}\log(p_{i,j})$
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
Precision	$\frac{TP}{TP+FP}$
Recall	$\frac{TP}{TP+FN}$
F1-Score	$\frac{2*Precision*Recall}{Precision+Recall}$

ANALYSIS

The K-Folds Cross Validation process was used to evaluate the models implemented, to ensure that the data distribution did not influence the model's performance. Data was split into five equal folds, and the model was trained on all but one fold; the remaining fold was used to evaluate performance. This process was repeated five times, with a different fold utilized for evaluation each time. The mean, minimum, maximum and standard deviation accuracy results were calculated for each model, in order to evaluate performance.

SENSITIVITY ANALYSIS

Two sensitivity analyses were conducted to determine the impact that different quantities of data have on the model's performance. Both analyses were performed on group 1 dataset.

1. Quantity of training images for each sheep: The model was trained nine times with a different number of images between 10 and 52 images for each training run. All models were tested on the same 13 images. The identification accuracy achieved for each sheep and the average accuracy of the group were evaluated for each training run.
2. The minimum number of testing frames for each sheep: Since the videos were acquired while the sheep were drinking, each video contained more than 1,000

frames. Running the model on all the frames is therefore time consuming. Furthermore, the sheep were captured in various postures; many frames captured only partial faces or individual features, which are much more difficult to identify. This analysis evaluated the minimum number of frames required to be randomly sampled from a full video in order to ensure that each sheep was identified correctly. Accuracy was evaluated on different quantities of randomly sampled fames of the trained model that used 52 images, between one to ten frames. The decision regarding each sheep's identity was made based on the ID that the majority of the frames indicated, e.g. in two out of five frames, the model identified a sheep as '3502'; in the three remaining frames, it was identified as sheep '3633'. Thus, the sheep was identified as '3633'.

RESULTS

COMPARISON OF CLASSIFICATION MODELS

ResNet50V2 and ResNet50 were significantly better than the other models (Post-hoc Tukey's statistical significant test, $\alpha = 0.05$). ResNet50V2 achieved the best accuracy (Table 2); specifically, it was 22% better than the worst model (VGG16), and 1.5% better than the second best model (ResNet50). The selected classification included a total of 197 layers with a total of 40,376,832 parameters.

TABLE 2
SUMMARY RESULTS OF THE DIFFERENT CLASSIFICATION MODELS (COMBINED OF A CNN AND ARCFACE LOSS FUNCTION).

CNN	Accuracy in %
ResNet50V2	**94.9**
ResNet50	93.4
ResNet101V2	91.3
EfficientNetB3	87.3
EfficientNetB0	85.2
AlexNet	76.2
VGG16	72.9

SELECTED MODEL (RESNET50V2)

Using the cross-validation technique, ResNet50V2 CNN combined with the ArcFace loss function resulted in an average 95% accuracy for group 1 and group 2, while the unified group achieved lower results, but only by 2% (Table 3). All three trained models converged after 25 epochs. The low standard deviation of the five folds (less than 0.01 for all models, Table 3) proves the model's reliability. The decrease in the unified group performance can be explained due to similarities between the groups, since some of the sheep in the different groups were born to the same parents (see some examples in Figure 6).

TABLE 3
K-FOLD PERFORMANCE RESULTS OF THE SELECTED MODEL (RESNET50V2 CNN COMBINED WITH THE ARCFACE LOSS FUNCTION).

	Group1	Group2	Unified
Accuracy			
Min	0.93	0.94	0.93
Max	0.96	0.96	0.94
Mean	0.95	0.95	0.93
SD	0.01	0.01	0.00
Loss			
Min	2.00	1.73	2.40
Max	2.18	1.87	2.48
Mean	2.06	1.81	2.43
SD	0.06	0.05	0.02

FIGURE 6
EXAMPLE OF SIMILARITIES IN PAIRS OF SHEEP PAIR BETWEEN GROUPS; IN GREEN - SHEEP FROM GROUP 1, IN BLUE – SHEEP FROM GROUP 2.

Average precision, recall and F1-Score indices were 96.2%, 95.8%, 95.8% respectively (Table 4). An average of 12.76 and 12.58 out of 13 images per sheep were correctly identified. Furthermore, all sheep (in both groups) were correctly identified with a minimum of 10 images. Therefore, according to the majority classification rule, we can assume 100% successful identification for each sheep. The specific images that the model failed to identify were difficult to identify even by the human eye (Figure 7).

FIGURE 7

EXAMPLE OF INCORRECT IDENTIFICATION OF A SINGLE IMAGE; LEFT - ONE IMAGE OF SHEEP 3622 FROM GROUP 1, WAS WRONGLY CLASSIFIED AS SHEEP 3527, RIGHT – ONE IMAGE OF SHEEP 1015 FROM GROUP 2, WAS WRONGLY CLASSIFIED AS SHEEP 1016.

TABLE 4

CLASSIFICATION RESULTS.

	Precision	Recall	F1-Score
Group 1: 47 sheep	0.952	0.949	0.949
Group 2: 34 sheep	0.972	0.968	0.968

TRANSFER LEARNING

Transfer learning achieved by the cross-validation technique resulted in an average improvement of 1.85% increase in accuracy and a 0.25 decrease in loss value (Table 5), with a shorter training process (9 vs. 25 epochs). The data distribution had no effect on the identification performance as revealed by the low standard deviations.

TABLE 5

AVERAGE PERFORMANCE OF THE TRANSFER LEARNING METHODS.

Transfer method	Accuracy: Avg	Accuracy: SD	Loss: Avg	Loss: SD
G1 → G2	0.97	0.01	1.69	0.04
G2 → G1	0.97	0.01	1.70	0.04

SENSITIVITY ANALYSIS

QUANTITY OF TRAINING IMAGES

Training each sheep with at least 35 images resulted in a minimum identification accuracy exceeding 69% per sheep (Figure 8, i.e. at least nine of the 13 images were properly classified). By utilizing the majority-based decision described above, each sheep was correctly identified with decreasing SD as the number of images increase (Table 6).

QUANTITY OF RANDOMLY SAMPLED FRAMES

When classification was done using eight and nine frames, the average accuracy was 59.5% and 63.7% respectively (Figure 9). Results revealed that there were sheep that were identified in less than half of the frames and therefore were misclassified. However, when ten frames were used, the average accuracy was 71.7%, with a minimum of six out of ten images, implying correct classification of all sheep. Accordingly, in order to ensure proper identification of each sheep, the classification decision must be made based on a majority of at least ten randomly selected frames.

FIGURE 8

QUANTITY OF TRAINING IMAGES EVALUATED; THE MINIMUM, AVERAGE AND MAXIMUM ACCURACY ACHIEVED (SD IS PRESENTED IN TABLE 6).

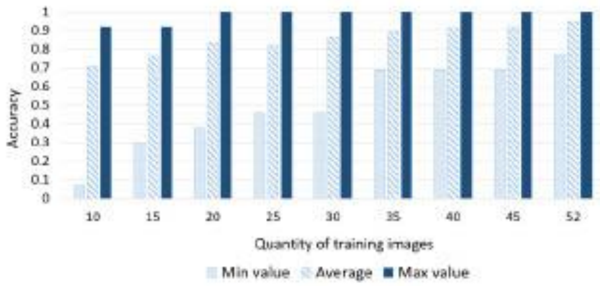
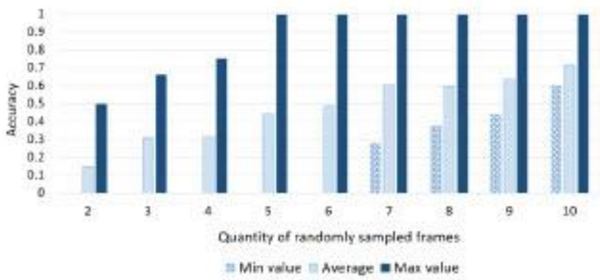


FIGURE 9

NUMBER OF RANDOMLY SAMPLED FRAMES; THE MINIMUM, AVERAGE AND MAXIMUM ACCURACY ACHIEVED.



DISCUSSION

To our knowledge, this study was the first to implement robust biometric identification of young sheep. Compared with earlier studies (Corkery et al., 2007; Salama et al., 2019), the slightly lower identification accuracies achieved in this study can be attributed to differences in the data collection methods and

TABLE 6
STANDARD DEVIATIONS FOR DIFFERENT QUANTITY OF IMAGES USED IN TRAINING.

Amount of images	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**52**
Standard deviation	0.20	0.17	0.11	0.15	0.12	0.10	0.09	0.08	0.06

differences between the sheep. In the study that achieved 96% accuracy (Corkery et al., 2007), images were manually acquired in controlled conditions on adult sheep aged 3-4 years which have higher differences. In the other study which achieved 98% accuracy (Salama et al., 2019), CNNs were tested on a group of sheep, in which the ages of the individual sheep varied greatly. In our study, biometric identification was done on identically-aged young sheep (lambs) with an automated machine vision system in field conditions, implying a more natural but complicated identification process. An advantage of the proposed identification system is that it can be installed in a commercial sheep pen - cameras can be altered positioned at different angles and distances from the sheep.

CONCLUSIONS

A biometric identification model for individual sheep recognition based on machine vision and advanced CNNs was designed, developed and implemented. The average identification accuracy achieved on two different groups of similarly-aged sheep was 95%. Transfer learning methods improved accuracy by an average of 1.85%, in a shorter training process, resulting in an average of 97% identification accuracy.

The model was implemented using Faster-RCNN algorithm and a ResNet50V2 model with the ArcFace loss function. The training process should include at least 35 images per sheep, while classification should be calculated by majority decision on at least ten randomly chosen frames per visit per sheep, from a full video.

Further research should aim to develop an unsupervised learning model for a new herd and implement real-time identification. The recommended model may be adapted to identify other animals, such as beef cattle and pigs, replacing traditional identification methods.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Almog Hitelman: Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing - review & editing, Visualization, Project administration. **Yael Edan:** Conceptualization, Methodology, Investigation, Resources, Writing - original draft, Writing - review & editing, Project administration, Funding acquisition. Supervision. **Assaf Godo:** Mechanical design, Mechanical engineering. Writing – review. **Ron Berenstein:** Mechanical design, Mechanical engineering. Writing – review. **Joseph Lepar:** System design, Writing - review. **Ilan Halachmi**: Conceptualization, Methodology, Investigation, Resources, Writing - original draft, Writing - review & editing, Project administration, Funding acquisition. Supervision.

DECLARATION OF INTEREST

The authors declare that there is no conflict of interest.

ACKNOWLEDGEMENTS

Special thanks to Mr. N. Bergman for his advice and contribution on computer vision and to Mr. A. Rosov, who hosted us in the sheep pen of Volcani Center. We wish to thank all of the PLF lab members and staff who supported this research.

FUNDING

This study was supported by the Israeli Chief Scientist of Agriculture fund "Kandel" PLF center of expertise (20-12-0029, 459451415), "TechCare" (GA No 862050) and "Sm@RT" (GA No 101000471). Partial support was provided by Ben-Gurion University of the Negev through the Rabbi W. Gunther Plaut Chair in Manufacturing Engineering.

APPENDIX A

MODEL'S HYPER PARAMETERS VALUES.

Hyper parameter	Value
Learning Rate (LR) maximum	0.005
Learning Rate (LR) minimum	0.00001

Epochs	30
Batch Size	32
Weight Decay	0.001
Dropout	0.5
S (ArcFace)	3
W (ArcFace)	0.05

REFERENCES

Ait-Saidi, A., Caja, G., Salama, A.A.K., Carné, S., 2014. Implementing electronic identification for performance recording in sheep: I. Manual versus semiautomatic and automatic recording systems in dairy and meat farms. J. Dairy Sci. 97, 7505–7514. https://doi.org/10.3168/jds.2014-8090

Andrew, W., Greatwood, C., Burghardt, T., 2019. Aerial Animal Biometrics: Individual Friesian Cattle Recovery and Visual Identification via an Autonomous UAV with Onboard Deep Inference 2–8.

Caja, G., Ghirardi, J., Hernández-Jover, M., Garín, D., 2004. Diversity of animal identification techniques: from 'fire age'to 'electronic age.' ICAR Tech. Ser.

Corkery, G., Gonzales-Barron, U.., Butler, F., Mcdonnell, K., Ward, S., 2007. A preliminary investigation on face recognition as a biometric identifier of sheep. Trans. ASABE.

Deng, J., Guo, J., Xue, N., Zafeiriou, S., 2019. ArcFace: Additive angular margin loss for deep face recognition. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2019-June, 4685–4694. https://doi.org/10.1109/CVPR.2019.00482

Dong, R., Pan, X., Li, F., 2019. DenseU-Net-Based Semantic Segmentation of Small Objects in Urban Remote Sensing Images. IEEE Access 7, 65347–65356. https://doi.org/10.1109/ACCESS.2019.2917952

Guo, G., Zhang, N., 2019. A survey on deep learning based face recognition. Comput. Vis. Image Underst. 189, 102805. https://doi.org/10.1016/j.cviu.2019.102805

Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., Lew, M.S., 2016. Deep learning for visual understanding: A review. Neurocomputing. https://doi.org/10.1016/j.neucom.2015.09.116

Halachmi, I., Guarino, M., 2016. Editorial: Precision livestock farming: A "per animal" approach using advanced monitoring technologies. Animal 10, 1482–1483. https://doi.org/10.1017/S1751731116001142

Halachmi, I., Guarino, M., Bewley, J., Pastell, M., 2019. Smart Animal Agriculture: Application of Real-Time Sensors to Improve Animal Well-Being and Production. Annu. Rev. Anim. Biosci.

Hansen, M.F., Smith, M.L., Smith, L.N., Salter, M.G., Baxter, E.M., Farish, M., Grieve, B., 2018. Towards on-farm pig face recognition using convolutional neural networks. Comput. Ind. 98, 145–152. https://doi.org/10.1016/j.compind.2018.02.016

Hou, J., He, Y., Yang, H., Connor, T., Gao, J., Wang, Y., Zeng, Y., Zhang, J., Huang, J., Zheng, B., Zhou, S., 2020. Identification of animal individuals using deep learning: A case study of giant panda. Biol. Conserv. 242. https://doi.org/10.1016/j.biocon.2020.108414

Jiang, H., Learned-Miller, E., 2017. Face Detection with the Faster R-CNN. Proc. - 12th IEEE Int. Conf. Autom. Face Gesture Recognition, FG 2017 - 1st Int. Work. Adapt. Shot Learn. Gesture Underst. Prod. ASL4GUP 2017, Biometrics Wild, Bwild 2017, Heteroge 650–657. https://doi.org/10.1109/FG.2017.82

Kumar, S., Singh, S.K., 2017. Automatic identification of cattle using muzzle point pattern: a hybrid feature extraction and classification paradigm. Multimed. Tools Appl. 76, 26551–26580. https://doi.org/10.1007/s11042-016-4181-9

Landais, E., 2001. The marking of livestock in traditional pastoral societies. OIE Rev. Sci. Tech. https://doi.org/10.20506/rst.20.2.1286

Loshchilov, I., Hutter, F., 2017. SGDR: Stochastic gradient descent with warm restarts. 5th Int. Conf. Learn. Represent. ICLR 2017 - Conf. Track Proc. 1–16.

Morris, J.E., Cronin, G.M., Bush, R.D., 2012. Improving sheep production and welfare in extensive systems through precision sheep management. Anim. Prod. Sci. https://doi.org/10.1071/AN11097

Rahimzadeh, M., Attar, A., 2020. A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. Informatics Med. Unlocked 19, 100360. https://doi.org/10.1016/j.imu.2020.100360

Salama, A., Hassanien, A.E., Fahmy, A., 2019. Sheep Identification Using a Hybrid Deep Learning and Bayesian Optimization Approach. IEEE Access 7, 31681–31687. https://doi.org/10.1109/ACCESS.2019.2902724

Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition. 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc. 1–14.

Szeliski, R., 2010. Computer Vision: Algorithms and Applications, Texts in Computer Science. Springer London.

Talo, M., Yildirim, O., Baloglu, U.B., Aydin, G., Acharya, U.R., 2019. Convolutional neural networks for multi-class brain disease detection using MRI images. Comput. Med. Imaging Graph. 78, 101673. https://doi.org/10.1016/j.compmedimag.2019.101673

Tan, M., Le, Q. V., 2019. EfficientNet: Rethinking model scaling for convolutional neural networks. 36th Int. Conf. Mach. Learn. ICML 2019 2019-June, 10691–10700.

Voulodimos, A.S., Patrikakis, C.Z., Sideridis, A.B., Ntafis, V.A., Xylouri, E.M., 2010. A complete farm management system based on animal identification using RFID technology. Comput. Electron. Agric. 70, 380–388. https://doi.org/10.1016/j.compag.2009.07.009

Wang, H., Wang, Y., Zhou, Z., Ji, X., Gong, D., Zhou, J., Li, Z., Liu, W., 2018. CosFace: Large Margin Cosine Loss for Deep Face Recognition. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 5265–5274. https://doi.org/10.1109/CVPR.2018.00552

Xu, Z.Q.J., Zhang, Y., Xiao, Y., 2019. Training behavior of deep neural network in frequency domain. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 11953 LNCS, 264–274. https://doi.org/10.1007/978-3-030-36708-4_22

Zhao, B., Feng, J., Wu, X., Yan, S., 2017. A survey on deep learning-based fine-grained object classification and semantic segmentation. Int. J. Autom. Comput. 14, 119–135. https://doi.org/10.1007/s11633-017-1053-3

Zhong, Y., Zhao, M., 2020. Research on deep learning in apple leaf disease recognition. Comput. Electron. Agric. 168, 105146. https://doi.org/10.1016/j.compag.2019.105146.

Appendix 2B. The effect of age on sheep biometric identification

Short Communication: The effect of age on sheep biometric identification

Submitted to: Animal, the international journal of animal biosciences.

Authors: Almog Hitelman, Yael Edan, Assaf Godo, Ron Berenstein, Joseph Lepar and Ilan Halachmi.

Short Communication: The effect of age on sheep biometric identification

A. Hitelman1,2, Y. Edan2, A. Godo1, R. Berenstein1, J. Lepar1, I. Halachmi1,2*

1 Precision livestock farming (PLF) Lab., Agricultural Engineering Inst., Agricultural Research Organization (A.R.O.) - Volcani Institute, 68 Hamaccabim Road, P.O.B 15159 Rishon Lezion 7505101, Israel
2 Dept. of Industrial Engineering and Management, Ben-Gurion University of the Negev, P.O.B 653 Be'er Sheva 8410501, Israel
Corresponding author: Ilan Halachmi. Email: halachmi@volcani.agri.gov.il

Biometric identification provides an important tool for precision livestock farming. This study investigates the effect of weight gain and sheep maturation on recognition performance. Sheep facial identification was implemented using two convolutional neural network (CNN) called Faster R-CNN, and ResNet50V2, equipped with the state-of-art Additive Angular Margin (ArcFace) loss function. The identification model was tested on 47 young sheep at different stages in their growth, when they were between 2 and 5 months old, throughout which the sheep gained approximately 25 kilograms in weight. Results revealed that when the model was trained and tested on images at the beginning of the growth period, the average accuracy of the group was 95.4%, compared with 91.3% when trained on images from the beginning but tested on images at the end of the growth process.

Keywords: biometric identification, lamb, convolutional neural network, animal aging, deep learning.

Implications

The findings of this study suggest that young sheep can be identified using biometric methods throughout the growth period, which lasts about 4-5 months. Training the model was performed only once at the beginning of the growth process. The method presented may be adapted to other animals, such as cattle and pigs, and may replace conventional RFID identification systems with a lower cost and more animal welfare-friendly system.

INTRODUCTION

An individual animal identification opens the way to maintain an animal's individual data, such as parentage, birth date, production records, health & vaccination history and disease control, enables improved farm management (Salama et al., 2019). Traditional identification methods include: tattoos, ear tags or electronic means (Caja et al., 2004). Ear tags are easily lost, and require frequent maintenance and cleaning (Caja et al., 2004). Tattoos impair animal welfare and might affect an animal's behavior. Electronic identification, such as Radio Frequency Identification (RFID) tags, are more expensive compared to conventional methods, and are sensitive to local signal-to-noise ratios (Caja et al., 2004). By representing the visual appearances of animals based on generic features and primary biometric characteristics, individual animals can be identified using cameras (Corkery et al., 2007).

Different biometric characteristics can be used for animal identification, such as muzzle pattern matching and coat pattern (Andrew et al., 2019; Kumar et al., 2018). Faces contain significant information (Corkery et al., 2007; Salama et al., 2019). However, biometric characteristics, especially of the face, are susceptible to age change (Sawant and Bhurchandi, 2019). Various studies on human identification have focused on understanding and improving the effect of age on recognition (Sawant and Bhurchandi, 2019). However, to our knowledge, equivalent animal studies have not yet been performed on lambs. To the best of our knowledge, no previous study has investigated livestock biometric identification over a long enough period, nor examined the influence of aging on recognition.

Biometric identification is an emerging research field. Livestock farming has become increasingly interested CNN techniques (Hansen et al., 2018; Salama et al., 2019). An initial study on sheep facial identification trained and tested a cosine distance classifier on 50 sheep, aged 3-4 years (Corkery et al., 2007) resulted with 96% identification accuracy. In this study, front-view images were acquired in controlled conditions (on a black background and with manual cleaning of the sheep face to minimize noise). A second study (Salama et al., 2019) utilizing two different neural network models, resulted in 98% accuracy. In that study, tests were performed on 52 sheep, in a wide range of ages (from five months to five years); 10 images for each sheep were taken at a single time, from different angles. However, to our knowledge, no research has yet been conducted to evaluate the identification performance on different ages of the same sheep, and none have focused on identification of young sheep.

In this study, the ResNet50V2 CNN (Yamazaki et al., 2019) was combined with the state-of-art ArcFace loss function (Deng et al., 2019), to provide a robust facial biometric identification model for sheep. The present study focuses on implementing the model to evaluate whether the sheep's weight gain and maturation affected the performance of the biometric identification. We especially focused on young sheep, since recognition of lambs is expected to be more difficult, due to the similarity between the sheep; rapid growth is also known to influence their appearance.

MATERIALS AND METHODS

EXPERIMENTAL DESIGN

Sheep facial images were collected using two 8 MegaPixels RGB cameras connected to a Jetson Nano embedded system-on-module (SoM). The cameras were installed on a drinking facility, located in a research sheep pen in Volcani Center, Bet Dagan, Israel. Both cameras video-recorded the sheep while they were drinking water, at 30 frames per second (fps). The first camera was used to photograph the sheep's RFID ear tag, which was used as the ground truth. The second camera acquired photos of each sheep's face for the identification model.

A total of three experiments were conducted on 47 Assaf breed sheep. Images were acquired throughout their growth process, from ages 2 to 5 months, during which each sheep gained about 25 kilograms in weight. The first experiment lasted for two weeks, and was conducted in November 2020, with 2-month-old sheep (average weight=33.6, SD=6.6). Two additional experiments were conducted. One experiment was done in the middle of the growth period, at the end of December 2020, when the sheep weighed about 45 kg each (average=44.4, SD=6.51). The other was conducted at the end of February 2021, at the end of the growth process; the sheep were at their maximum weight, about 60 kg (average=57.2, SD=7.82). Each experiment was conducted over three consecutive days. The second and third experiments included only 32 sheep, since 15 sheep were removed from the pen during the course of experiments, due to illness.

EXPERIMENTAL DESIGN

Each video of a sheep taken from the front camera was tagged with the sheep's ID tag taken from the side camera. Each video was divided into an image sequence (frames) with Python3 code. Then, the images were normalized (Deng et al., 2019). The following three datasets were obtained:

- First dataset (baseline dataset): 3055 images were included from the beginning of the growth process (65 images for each of the 47 sheep). Among these, 52 images were randomly selected for the model development, and the remaining 13 images were used to evaluate the model's efficacy in identifying individual sheep.
- Second dataset (denoted as Growth1): 2080 images were included from three different time periods – beginning, middle and the end of growth period (65 images for each of the 32 sheep). Specifically: 1248 images were taken from the beginning of the growth process (from the first dataset), and 416 images were taken from the middle of growth process (second experiment) for training the model. An additional 416 more images were taken at the end of growth period (third experiment) for evaluation.
- Third dataset (denoted as Growth2): 2080 images were included from two time periods – the beginning, and the end of the growth period. Similar to the second dataset, 1664 images taken from the beginning of the growth process were used to train the model, and 416 images taken from the end of the growth period were used for evaluation.

An object detection algorithm was developed to locate the sheep's face in each image. To this end, a labelled dataset was created, consisting of 400 images of sheep. Each image was manually tagged with a bounding box around each sheep's face. To improve training, data augmentation techniques were used on all datasets, to increase the number of training samples. The training images were augmented seven to eight times using the following techniques: vertical flip, random rotation up to 30 degrees, Gaussian noise addition, and adjusting image brightness.

ALGORITHM AND MODELLING

The biometric identification model applied to each image (Hitelman, 2021) entailed two steps: (1) face detection and (2) classification. First, the faster R-CNN deep learning object detection algorithm was applied to locate the sheep's face in an image. Then, the detected face was cropped, resized to 112*112 pixels according to Deng et al., (2019) and provided as input to the classification model. The classification model was composed of an embedding network and a loss function. The ResNet50V2 (Yamazaki et al., 2019) was selected as the embedding network because it outperformed the other embedding networks examined (Hitelman, 2021). The state-of-the-art Additive Angular Margin Loss (ArcFace) function was utilized to extract highly distinctive features from faces (Deng et al., 2019).

MODEL EVALUATION

The model's performance was evaluated by applying it to the three datasets – baseline dataset, Growth1 and Growth2 as described above.

Performance was measured using a confusion matrix on a test set consisting of 13 images randomly selected for each sheep. The difference between the baseline and Growth2 was calculated. Performance was also evaluated using identification accuracy and categorical cross entropy on the three datasets. Accuracy was calculated from the confusion matrix as a proportion of correctly identified images out of the total images examined in the model. Categorical cross-entropy estimated the divergence of the predicted probability from the actual sheep ID.

RESULTS

The biometric identification model performance on the baseline dataset achieved superior accuracy (Table 1); 1.7% higher than Growth1 and 4.1% higher than Growth2 performance. Training on images from the beginning of growth only while testing on images from the end of growth (Growth2), achieved 2.4% lower accuracy than did training with images from the middle as well as the beginning of the growth period (Growth1). However, the difference in accuracy of Growth2 is relatively small compared with that of the baseline (4.1%).

TABLE I
SUMMARY OF THE IDENTIFICATION PERFORMANCE

	Accuracy	**Loss**
Baseline	0.95	2.06
Growth1	0.93	2.08
Growth2	0.91	2.18

The model's performance on the baseline and Growth2 datasets (Table 2) yielded 94.92% and 91.34% true positives respectively. On both datasets, most of the sheep were identified in at least 12 out of 13 images (41 sheep out of 47 in the baseline dataset and 23 out of 32 sheep in the Growth2 dataset). Furthermore, all sheep in Growth2 were identified in at least 10 images out of 13, equal to the minimum correct identified images achieved with the baseline dataset (Table 2). Fig. 1 shows examples of sheep that were correctly identified during the growth period.

TABLE 2
AVERAGE, MINIMUM, MAXIMUM AND STANDARD DEVIATION OF THE CORRECT IDENTIFICATION IMAGES (OUT OF 13 PER SHEEP).

	Mean	**Min**	**Max**	**SD**
Baseline	12.76	10	13	0.80
Growth2	11.87	10	13	0.90

FIGURE I
CORRECT PREDICTION GROWTH EXAMPLES OF SHEEP, FROM LEFT TO RIGHT: IMAGE FROM BEGINNING, MIDDLE AND THE END OF GROWTH PERIOD.

DISCUSSION

The identification model we developed yielded lower accuracy than did previous studies (96% and 98% accuracy rates, Corkery et al., 2007; Salama et al., 2019 accordingly). However, these previous studies identified sheep of different ages which may have had higher variability; neither evaluated the identification performance of their models for young ages of the same sheep. Our study examined the biometrical identification of very young sheep (2 to 5 months old), with very little variance between the animals, making identification more challenging. We surmise that the model will perform better when applied to adult sheep that have greater differences between individuals. Furthermore, in our study images were acquired automatically in commercial-farming conditions.

The study presented here reveals that identifying animals before and after maturation (after 2-3 months) is feasible. Each sheep was successfully identified in at least 10 of the 13 images in both datasets (baseline and Growth2). Therefore, classifying the sheep based on a majority of 13 images, we can assume a 100% success rate in identifying each sheep daily, throughout the growing period tested, which corresponds to the growing period of young sheep raised on commercial farms.

CONCLUSIONS

A biometric identification model (composed of Faster R-CNN, ResNet50V2 and ArcFace loss function) can be used to identify young sheep from the beginning to the end of their intensive growth period. The ability to train a model on young sheep that can still be identified over time even after 2-3 months and a gain of about 25 kilograms, may provide alternative method of identification to traditional methods. Traditional methods (tattoo, earring tags, RFID tags etc.) might impair animal welfare, while biometric identification is not harmful. In further research, the proposed model maybe applied to identify other livestock animals, such as beef cattle and pigs.

ETHICS APPROVAL

All the procedures in this study were carried out in accordance with the accepted ethical and welfare standards of the Israel Ethics Committee (approval number IL-801/18).

DATA AND MODEL AVAILABILITY STATEMENT

None of the data were deposited in an official repository. The data that support the study findings are confidential.

AUTHOR ORCIDS

Almog Hitelman: https://orcid.org/0000-0002-7211-2023. Yael Edan: https://orcid.org/0000-0002-7430-8468. Ilan Halachmi: https://orcid.org/0000-0002-2303-1016

AUTHOR CONTRIBUTIONS

Almog Hitelman: Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization, Project administration. **Yael Edan:** Conceptualization, Methodology, Investigation, Resources, Writing - original draft, Writing - review & editing, Project administration, Funding acquisition. Supervision. **Assaf Godo:** Mechanical design, Mechanical engineering. Writing – review. **Ron Berenstein:** Mechanical design, Mechanical engineering. Writing – review. **Joseph Lepar:** System design, Writing - review. **Ilan Halachmi**: Conceptualization, Methodology, Investigation, Resources, Writing - original draft, Writing - review & editing, Project administration, Funding acquisition. Supervision.

DECLARATION OF INTEREST

None.

ACKNOWLEDGEMENTS

Special thanks to Mr. N. Bergman for his advice and contribution on computer vision and to Mr. A. Rosov, who hosted us in the sheep pen of Volcani Center. We wish to thank all of the PLF lab members and staff who supported this research.

FINANCIAL SUPPORT STATEMENT

This study was supported by the Israeli Chief Scientist of Agriculture fund "Kandel" PLF center of expertise (20-12-0029, 459451415), "TechCare" (GA No 862050) and "Sm@RT" (GA No 101000471). Partial support was provided by Ben-Gurion University of the Negev through the Rabbi W. Gunther Plaut Chair in Manufacturing Engineering.

References

Andrew, W., Greatwood, C., Burghardt, T., 2019. Aerial Animal Biometrics: Individual Friesian Cattle Recovery and Visual Identification via an Autonomous UAV with Onboard Deep Inference. Proceedings of the IEEE International Conference on Intelligent Robots and Systems. 3-8, November, 2019, Macau, China, pp. 237–243.

Caja, G., Ghirardi, J., Hernández-Jover, M., Garín, D., 2004. Diversity of animal identification techniques: from 'fire age'to 'electronic age. Technical Series No. 9 (ed. R. Pauw, S. Mack and J. Maki-Hokkonen), pp. 21–39. The International Committee for Animal Recording (ICAR), Rome, Italy.

Corkery, G., Gonzales-Barron, U.A., Butler, F., McDonnell, K., Ward, S., 2007. A preliminary investigation on face recognition as a biometric identifier of sheep. American Society of Agricultural and Biological Engineers, (ASABE) 50 (1), 313-320. doi: 10.13031/2013.22395

Deng, J., Guo, J., Xue, N., Zafeiriou, S., 2019. ArcFace: Additive angular margin loss for deep face recognition. The Institute of Electrical and Electronics Engineers (IEEE) Conference on Computer Vision and Pattern Recognition 2019-June, 4685–4694. https://doi.org/10.1109/CVPR.2019.00482

Hansen, M.F., Smith, M.L., Smith, L.N., Salter, M.G., Baxter, E.M., Farish, M., Grieve, B., 2018. Towards on-farm pig face recognition using convolutional neural networks. Computers in Industry. 98, 145–152. https://doi.org/10.1016/j.compind.2018.02.016

Hitelman, A., 2021. Computer vision and developments in precision livestock using machine learning: a mobile system for counting laying hens, and a sheep biometric identification system. M.Sc thesis, Ben-Gurion University, Beer-Sheva, Israel.

Kumar, S., Pandey, A., Sai Ram Satwik, K., Kumar, S., Singh, S.K., Singh, A.K., Mohan, A., 2018. Deep learning framework for recognition of cattle using muzzle point image pattern. Measurement 116, 1–17. https://doi.org/10.1016/j.measurement.2017.10.064

Salama, A., Hassanien, A.E., Fahmy, A., 2019. Sheep Identification Using a Hybrid Deep Learning and Bayesian Optimization Approach. The Institute of Electrical and Electronics Engineers (IEEE) Access 7, 31681–31687. https://doi.org/10.1109/ACCESS.2019.2902724

Sawant, M.M., Bhurchandi, K.M., 2019. Age invariant face recognition: a survey on facial aging databases, techniques and effect of aging. Artificial Intelligence Review. 52, 981–1008. https://doi.org/10.1007/s10462-018-9661-z

Yamazaki, M., Kasagi, A., Tabuchi, A., Honda, T., Miwa, M., Fukumoto, N., Tabaru, T., Ike, A., Nakashima, K., 2019. Yet Another Accelerated SGD: ResNet-50 Training on ImageNet in 74.7 arXiv:1903.12650v1.

Appendix 2C. Existing smart drinking facility program and camera details

** If needed, contact *Ilan Halachmi. Email: halachmi@volcani.agri.gov.il* for access

Existing smart drinking facility program (Volcani access only) – URL

Appendix 2D. Video-recording code

Video-recording code (Volcani access only) – URL

Video_Recording.txt

Appendix 2E. Collected data

** If needed, contact *Ilan Halachmi. Email: halachmi@volcani.agri.gov.il* for access

Sheep collected data (Volcani access only) – URL

- *Raw data:*

```
Raw_Data
-> First Group
---> Videos
---> Frames
------> Selected Frames For Training
------> Ground Truth Frames
------> Growth Frames
----------> Middle
----------> End
------> Analyzed Data
-> Second Group
---> Videos
---> Frames
------> Selected Frames For Training
------> Ground Truth Frames
------> Analyzed Data
```

Appendix 2F. Classification model code and hyper parameters

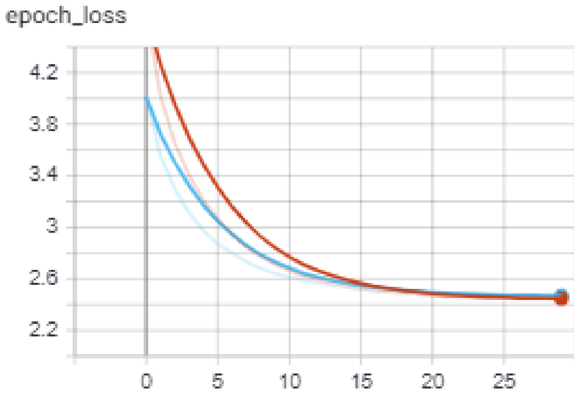
Classification model (Volcani access only) – URL

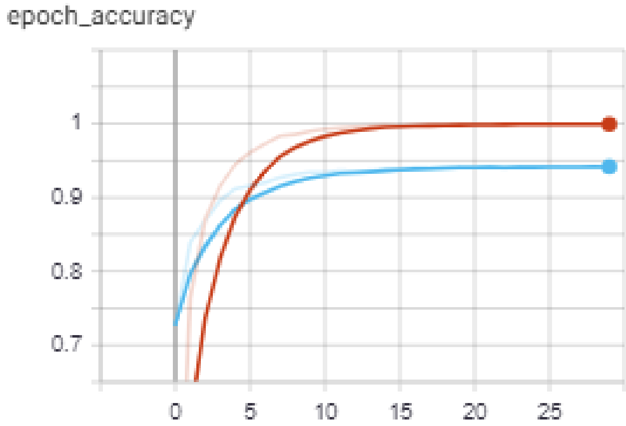
Hyper_Parameters_Definition.PNG

Identification_Model.txt

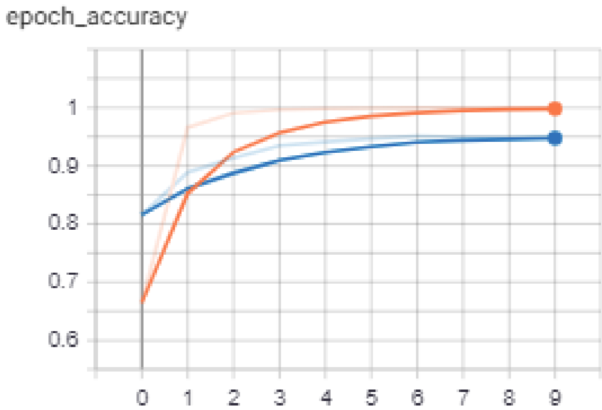
Appendix 2G. Trained models: graphs and hdf5 files
** Contact *Ilan Halachmi. Email:* halachmi@volcani.agri.gov.il for hdf5 file access.
[Trained models: graphs and hdf5 files (Volcani access only) – URL]()

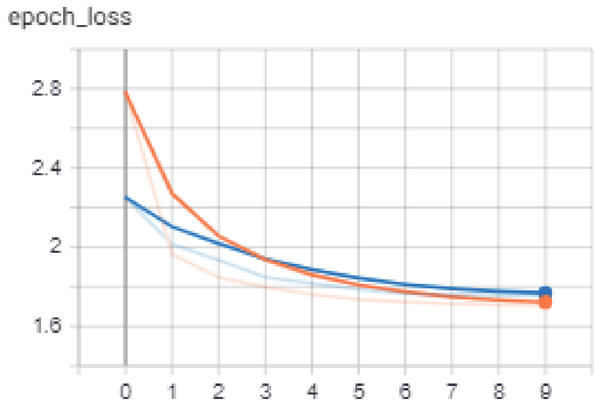
- *First group training performance – training (red), validation (blue).*
  X Axis - numbers of epochs, Y Axis - accuracy/loss values:



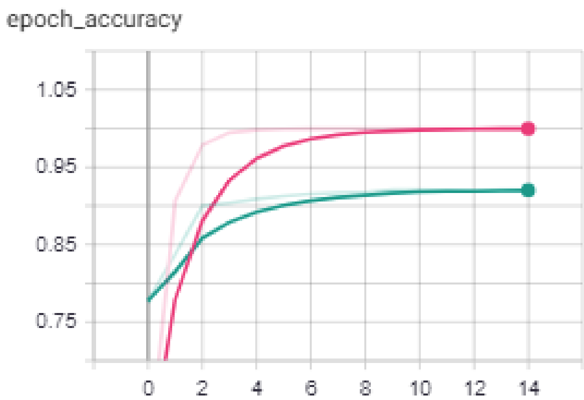


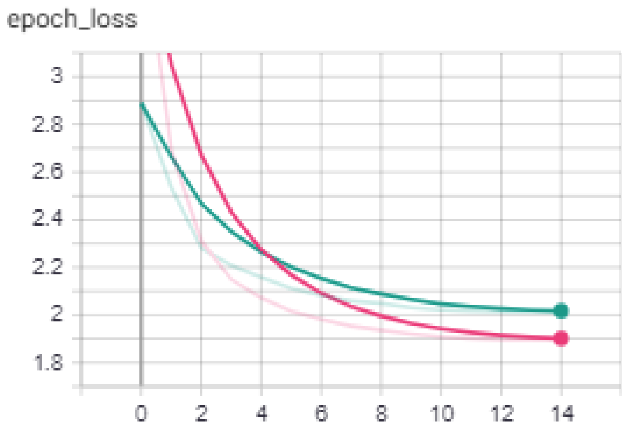
- *Transfer learning from first to second group performance – training (orange), validation (blue).*
  X Axis - numbers of epochs, Y Axis - accuracy/loss values:





- *Growth2 model performance – training (pink), validation (green).*
  X Axis - numbers of epochs, Y Axis - accuracy/loss values:





Appendix 2H. Post-hoc test

Post-hoc test (Volcani access only) – URL

Post_Hoc_Results.docx

Appendix 2I. Confusion matrices

Confusion matrixes (Volcani access only) – URL

Confusion_Matrices.xlsx

Appendix 2J. K-Fold Cross validation results

K-Fold Cross validation results (Volcani access only) – URL

K_Folds.xlsx

Appendix 2K. Amount of training images

Amount of training images (Volcani access only) – URL

Amount_Of_Trainining_Images.xlsx

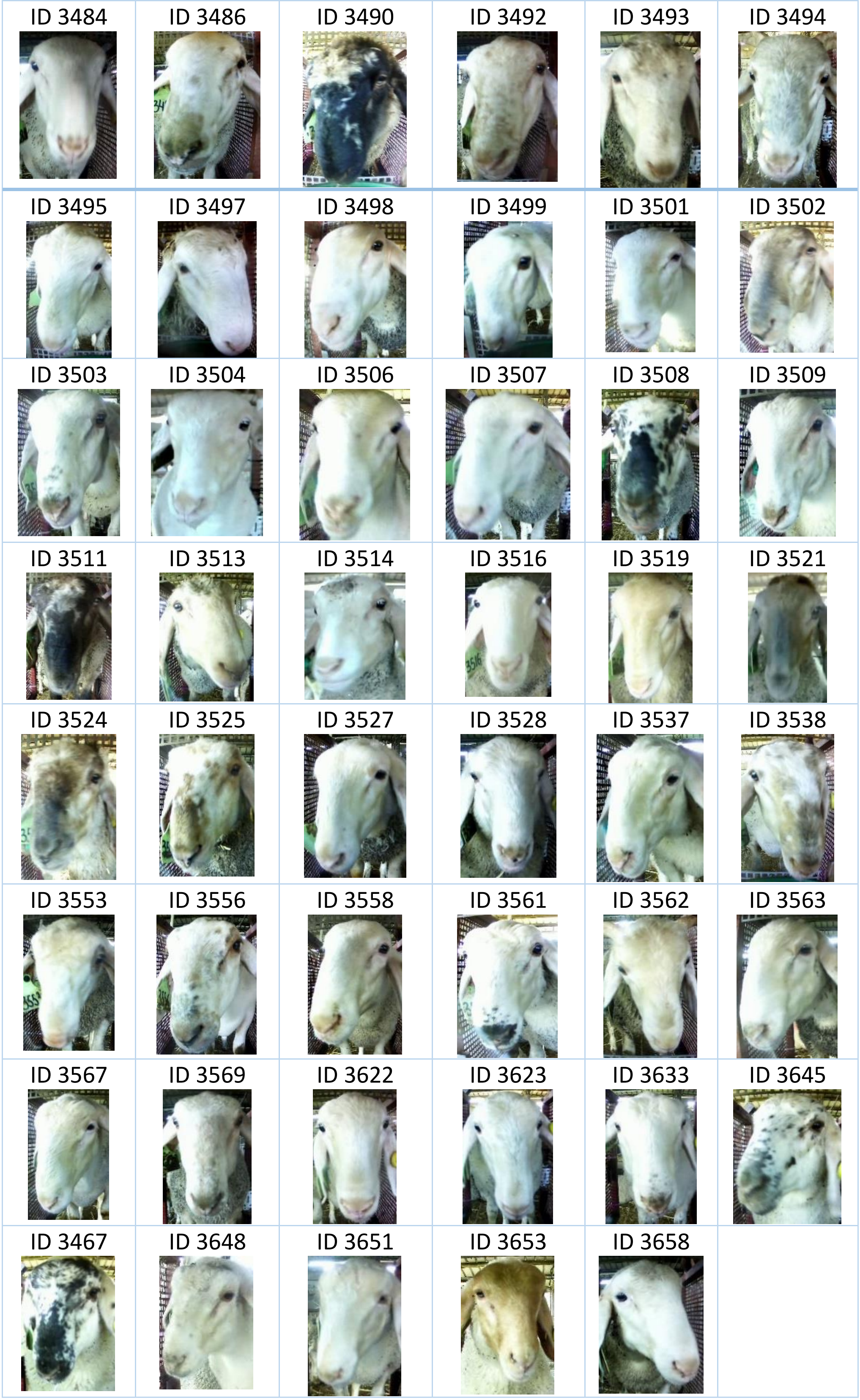
Appendix 2L. Amount of test images

Amount of test images (Volcani access only) – URL

Amount_Of_Test_Images.xlsx

Appendix 2M. Classified faces examples of both groups and growth model

Appendix 2N. First experimental sheep group – data examples



ID 3484	ID 3486	ID 3490	ID 3492	ID 3493	ID 3494
ID 3495	ID 3497	ID 3498	ID 3499	ID 3501	ID 3502
ID 3503	ID 3504	ID 3506	ID 3507	ID 3508	ID 3509
ID 3511	ID 3513	ID 3514	ID 3516	ID 3519	ID 3521
ID 3524	ID 3525	ID 3527	ID 3528	ID 3537	ID 3538
ID 3553	ID 3556	ID 3558	ID 3561	ID 3562	ID 3563
ID 3567	ID 3569	ID 3622	ID 3623	ID 3633	ID 3645
ID 3467	ID 3648	ID 3651	ID 3653	ID 3658	

Appendix 2O. Second experimental sheep group – data examples

ID 1001 	ID 1002 	ID 1003 	ID 1004 	ID 1005 	ID1006 
ID 1007 	ID 1008 	ID 1009 	ID 1010 	ID 1011 	ID 1012 
ID 1013 	ID 1014 	ID 1015 	ID 1016 	ID 1017 	ID 1018 
ID 1019 	ID 1020 	ID 1021 	ID 1022 	ID 1023 	ID 1024 
ID 1025 	ID 1026 	ID 1028 	ID 1029 	ID 1030 	ID 1032 
ID 1034 	ID 1036 	ID 1038 	ID 1039 		

זיהוי ביומטרי של צאן

זיהוי ביומטרי באמצעות אלגוריתמים מבוססי רשתות עמוקות זכו בשנים האחרונות להתעניינות גבוהה בתחום זיהוי בעלי החיים, עם זאת, מחקרים מעטים עסקו בזיהוי ביומטרי של צאן ופרט, לא נמצא מחקר שעוסק בזיהוי של טלאים. בעבודה זו פותחה מערכת מבוססת אלגוריתמיים לומדים מבוססי רשתות עמוקות לזיהוי ביומטרי של טלאים המבוססת על תמונות פנים.

נתונים לפיתוח האלגוריתם נאספו באמצעות מערכת איסוף נתונים שפותחה במיוחד, אשר נבנתה על מתקן שתייה חכם המותקן בדיר במכון וולקני. המערכת כללה 2 מצלמות RGB (צבע) המותאמות לעבודה עם מחשב Jetson Nano. ניסויים נערכו על שתי קבוצות שונות של כבשים בני 2-3 חודשים מגזע אסף. הנתונים של 47 טלאים (קבוצה 1) ושל 34 טלאים (קבוצה 2) נאספו באופן אוטומטי לאורך 14 יום בכל שעות היום כדי להבטיח מגוון של תנאי תאורה המושפעים מאור השמש ומזג האוויר.

מערכת הזיהוי הביומטרי מורכבת משני אלגוריתמים – לזיהוי הפנים, וסיווג זהות הטלה. תחילה, על מנת לאתר ולמקם את פני הטלה בתמונה, נעשה שימוש באלגוריתם לזיהוי אובייקטים Faster-RCNN. פני הטלאים שזוהו, נחתכו מתוך התמונה הגדולה ונשמרו כתמונה נפרדת. תמונה זו שמשה כקלט לאלגוריתם השני – סיווג הזהות. כדי להשיג את ביצועי הסיווג הטובים ביותר, נבחנו מספר רשתות עמוקות שונות ללמידת מאפייני הפנים. הביצועים הטובים ביותר הושגו באמצעות רשת ResNet50v2 בשילוב פונקציית אובדן חדישה הממקסמת את ההפרדה בין הטלאים בשם ArcFace.

האלגוריתם הצליח לזהות בדיוק של כ-95.4% ו-95.7% את הטלאים בקבוצה הראשונה והשנייה בהתאמה, כאשר באמצעות טכניקת Transfer-Learning ביצועי המערכת שופרו והגיעו ל-97% אחוזי דיוק, תוך קיצור זמן תהליך האימון.

מילות מפתח: חקלאות מדייקת בבעלי חיים, מערכת, תרנגולות מטילות, טלאים, אלגוריתמים, למידה עמוקה, זיהוי אובייקטים, זיהוי ביומטרי.

תקציר

טכנולוגיות העוסקות בחקלאות מדייקת בבעלי חיים (PLF) עשויות לעזור לחקלאי לנהל את בעלי החיים שלו, להגדיל את התשואה ולהבטיח את רווחת בעלי החיים. על מנת לתמוך בטכנולוגיות כאלו, יש להבין מהם צרכי בעלי החיים. במחקר זה נעשה פותחו שני מחקרי מקרה של טכנולוגיות ה-PLF: פיתוח מערכת ניידת לספירת תרנגולות מטילות ופיתוח מערכת זיהוי ביומטרית של צאן.

ספירת תרנגולות מטילות

ענף תרנגולות ההטלה בישראל מוסדר על ידי מכסת ייצור; משק יכול לייצר ביצים לפי מספר התרנגולות שהוקצה לו לגדל. כלובים קהילתיים חדשים, המכונים גם 'כלובי סוללות' המכילים 18-34 תרנגולות בכלוב, השתלבו לאחרונה בתעשיית הביצים בישראל. ספירה ידנית של התרנגולות בכלובים אלו, עשויה לדרוש זמן רב ואף להוביל לתוצאות לא מדויקות.

מערכת ניידת הניתנת להעברה בין לולים שונים פותחה. המערכת שמה לה למטרה להחליף את הספירה הידנית של תרנגולות מטילות, באלגוריתם אוטומטי, אשר יגלה ויספור תרנגולות ובכך יסייע לפקח על מכסות הגידול. הניסויים נערכו בלול שנמצא במושב "קדרון". הלול בנוי מ- 6 קומות, כאשר בכל קומה 37 כלובים. אורכה של קומה כזו הינה כ -87 מטר. התרנגולות הוקלטו באמצעות צילום וידאו שנעשה במצלמת Intel RealSense depth D435 אשר מצלמת 30 פריימים לשנייה. סרטוני הווידאו עובדו בעזרת אלגוריתם בינה מלאכותית לזיהוי אובייקטים בשם Faster R-CNN. לאחר מכן ומכיוון שהתרנגולות זזות כל הזמן, נעשה מעקב אחרי התרנגולות שזוהו בווידאו באמצעות אלגוריתם עקיבה , אשר הקצה מספר ייחודי לכל תרנגולת.

האלגוריתם שנבחן על מאגר נתונים של כ- 5600 תמונות, הניב תוצאות דיוק של כ-88% עם שגיאה ממוצעת של 4.5 תרנגולות לכלוב. תוצאת ספירת האלגוריתם הושוותה לספירת ידנית שנעשתה על ידי אדם והוגדרה כ- ground truth. בכדי לשפר את דיוק האלגוריתם יש להמשיך בפיתוח ושיפור המערכת. המשך העבודה כולל בחינה של הוספת מימד העומק לתהליך הלימוד של האלגוריתם ובכך להשיג הבנה טובה יותר של גבולות הכלוב ולצמצם את מספר התרנגולות שזוהו לא נכון, כלומר – זוהו כחלק מהכלוב שנבחן אף על פי שאינן בגבולות אותו הכלוב.

המערכת שמוצגת בפרויקט זה הינה זולה, מהירה, ידידותית למשתמש ובנויה באופן גנרי. לכן, ניתן להשתמש במערכת באופן פוטנציאלי בלולים שונים ובכך לאפשר פיקוח יעיל ומהיר יותר על מכסות הגידול.

אוניברסיטת בן-גוריון בנגב
הפקולטה למדעי ההנדסה
המחלקה להנדסת תעשייה וניהול

מערכות ראייה ממוחשבת בחקלאות מדייקת באמצעות לימוד מכונה: מערכת ניידת לספירת תרנגולות מטילות, ומערכת לזיהוי ביומטרי של צאן

חיבור זה מהווה חלק מהדרישות לקבלת תואר מגיסטר במדעי ההנדסה

מאת: אלמוג חיטלמן

בהנחיית: פרופ' יעל אידן ופרופ' אילן הלחמי

Almog

חתימת המחבר:............................ תאריך: 19.9.21

Yael Edan

אישור המנחה:.......... .... תאריך: 19.9.21

אישור המנחה:....... .... תאריך: 19.9.21

אישור יו"ר ועדת תואר שני מחלקתי:......................... תאריך:....19/9/2021.........................

אלול ,תשפ"ב

ספטמבר, 2021

אוניברסיטת בן-גוריון בנגב
הפקולטה למדעי ההנדסה
המחלקה להנדסת תעשייה וניהול

מערכות ראייה ממוחשבת בחקלאות מדייקת באמצעות לימוד מכונה: מערכת ניידת לספירת תרנגולות מטילות, ומערכת לזיהוי ביומטרי של צאן

חיבור זה מהווה חלק מהדרישות לקבלת תואר מגיסטר במדעי ההנדסה

מאת: אלמוג חיטלמן
בהנחיית: פרופ' יעל אידן ופרופ' אילן הלחמי

אלול ,תשפ"ב
ספטמבר, 2021