SYNTHETICALLY GENERATED COW (BOS TAURUS) PROVIDES DATA FOR GAIT ANALYSIS IN FEEDLOT

ALI GOLDANI
Bachelor of Science, Ferdowsi University of Mashhad, 2021

A thesis submitted
in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in

NEUROSCIENCE

Department of Neuroscience
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

SYNTHETICALLY GENERATED COW (BOS TAURUS) PROVIDES DATA FOR GAIT ANALYSIS IN FEEDLOT

ALI GOLDANI

Date of Defence: December 8th, 2023

Dr. I. Q. Whishaw	Professor	Ph.D.
Dr. M. H. Mohajerani	Adjunct Associate Professor	Ph.D.
Thesis Co-Supervisors		
Dr. B. E. Kolb	Professor	Ph.D.
Thesis Examination Committee Member		
Dr. R. J. Sutherland	Professor	Ph.D.
Thesis Examination Committee Member		
Dr. A. Luczak	Professor	Ph.D.
Chair, Thesis Examination Committee		

DEDICATION

To the love of my life, Faezeh, who accompanied me through this journey and my parents and sister who did so but 10000 kilometers away.

ABSTRACT

Analysis of bovine movement and behavior is crucial in detecting their motor disorders and maintaining their welfare. Quantitative gait analysis methods have been designed to facilitate this task, but the on-site subjective assessment of the gait pattern remains prominent and depends on human expertise. Gait pattern could be assessed in feedlot cattle using AI as a substitute for the absence of human diagnosis, but creating an AI diagnosis procedure requires substantial behavioral information for training the AI tool. One solution for obtaining behavioral information is to use AI-assisted tools for diagnosis based on recordings of cattle movement. In this study, we created a three-dimensional digital representation of walking cattle to generate the required information and compare its applicability to that of the actual gait patterns. We used video recordings of cattle walking and trotting, and then used them as reference to create three-dimensional pose representations. Then, we introduced variations to these representations by altering specific aspects of the original walking cow model and its environment. We then tested the combined representations against the real data to see if they can prove useful in training a deep neural network for detecting gait pattern and features. This method can compensate for the scarcity of behavioral data, provide information to create mathematical representations of specific behaviors and be used for the development of smart-phone-based diagnosis systems.

ACKNOWLEDGEMENTS

I hereby express my gratitude towards the people whose company and mentorship made this work possible. My supervisors Dr. Whishaw for all the Monday morning meetings, and Dr. Mohajerani whose interest and determination in combining different talents created this opportunity for me to work in this field of research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
DL	Deep Learning
ML	Machine Learning

Chapter 1: Introduction

1.1 Interest in Movement

Animals are defined by their movements. This thesis has as its objective the description of walking and its abnormalities in cattle. In the following sections of this introduction, I will first provide some background information on movement in cattle, I will then describe their problems identifiable in visual cues in their behavior, and finally I will describe methods of using AI to analyze movement.

Analyzing movement allows animals to find prey, avoid predators, or find a mate. They are equipped with features that enable them to sense movement as means of doing so. Frogs are adapted to detect minuscule movements in their field of vision to detect and capture prey (Maturana et al., 1960). The barn owl (Tyto alba) funnels sound to its ears through facial discs, enabling it to “see” in the dark and locate a mouse feet away (Payne, 1971; Takahashi, 2010). Patterns of vigilance in animals depend on successful detection of predatory movements. Birds scan the environment in a head up position, allowing them to perceive higher threat levels (Fernández-Juricic et al., 2004). Wild boars nibble off branches around their nest to increase visibility and enhance their chance of detecting predators in advance (Fu et al., 2022). The mating ritual in animals is essentially a behavior scoring task for the female based on the movement of the male (Tinbergen, 1954). Animals perform these types of analysis based on survival instinct. Were they capable of performing executive function, one could only imagine the results.

By analyzing the attack pattern of the barn owl (Tyto alba), a mouse would realize that the best way to successfully evade the owl is not to run directly away from it but to pull sideways from its flight direction (Shifferman & Eilam, 2004).

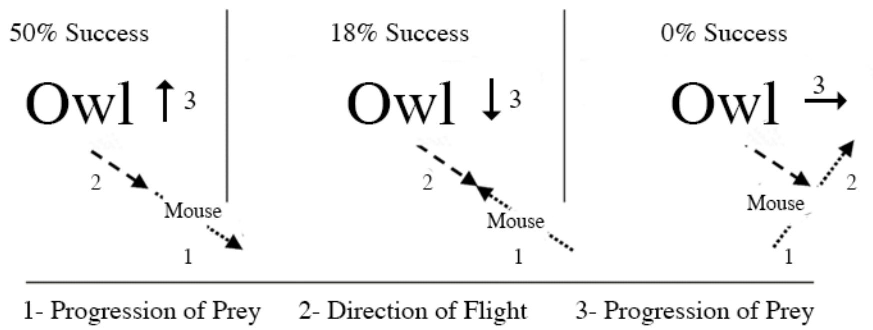

Figure 1-1: The relation between rate of success and direction of movement for a food item that was pulled forward (a), backward (b) and sideways (c). Direction of prey progression - dotted arrow (1), direction of owl flight - dashed arrow (2), and direction to which the owl had to move its head or trunk - solid arrow (3). Adapted from (Shifferman & Eilam, 2004).

If golden-collared manakins could record the courtship displays of each male, measure their success rate, and employ basic statistics they would realize that there are individual variabilities among them (Fusani et al., 2007). They could use these patterns as a choreographic plan to teach the best courtship display to their young ones.

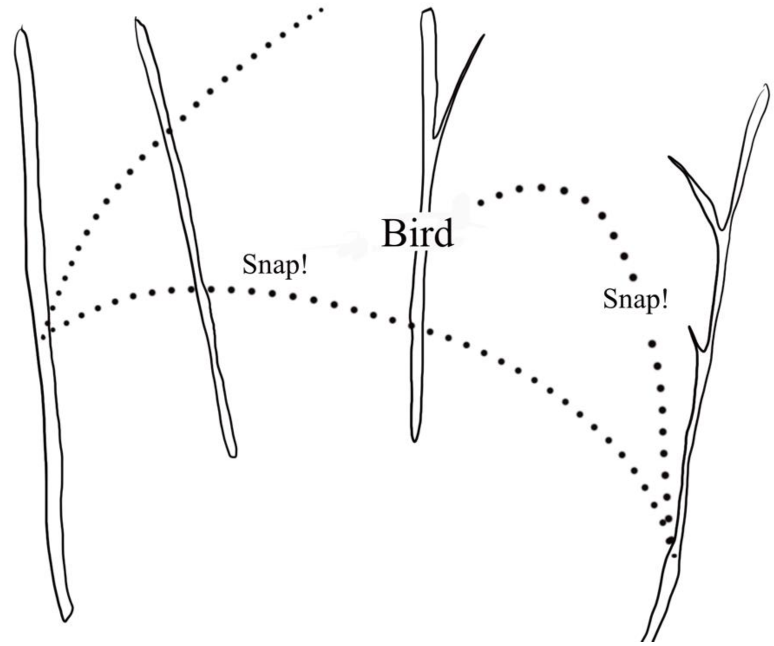

Figure 1-2: A schematic representation of the snap-jump display of a Golden-collared mana-kin. The bird comes down from a perch 2–5 m above the court and lands on one of the vertical saplings delimiting the court. Then, he jumps towards another sapling and produces a loud snap in midair by a very fast upward movement of the wings that come in contact above the back of the bird. The snap-jump can be repeated up to 20 times in a row. Adapted from (Fusani et al., 2007).

However, these animals lack the qualities needed to perform such analyses. Qualities which can be found in homo sapiens.

1.2 Describing Movement

It did not take humans long to realize the benefits of analyzing movement. Early hominins, like many animals, had to be vigilant about the threat of predators. Detecting the movement of a lurking predator in time could mean the difference between life and death. As early humans transitioned from being primarily foragers to hunters, the ability to detect the movement of prey became crucial. Noticing the subtle movements of animals in the grass, trees, or water helped early hunters locate and pursue their targets. Even lack of movements could be a sign for the early humans; one could find their way back home using environmental signs such as trees, rocks, and their relative position. It took them some time to devise methods that could help describe movements accurately.

The history of scientific description of movement starts with Aristotle (384-322 BCE). In his famous work, Physica (Physics), he tries to define the essence of movement and associates time with it as its constant attribute in Book IV. He then classifies movement into four species (quantity, quality, place, and substance) in Book V. Next, he discusses the reason why a changing thing would reach the opposite state and the relationship between the moved and the mover in books VI and VII. Although his work may seem rudimentary at first, it clearly captures his use of logical thinking based on accurate observations.

A little later, Archimedes of Syracuse (287-212 BCE) introduced the use of mathematics in describing movement. In his works, On the Equilibrium of Planes and On Floating Bodies, he laid the foundation for mechanics and hydrostatics. His studies of the movements of unanimated objects and their relation to each other resulted in several inventions. Due to his works, he is considered one of the fathers of engineering as we know it today.

Omar Khayyam (1048-1131), a Persian polymath, is celebrated for his contributions to the design of Jalali calendar among other things. This calendar is based on astronomical observations and measurements and was more accurate than many of its contemporaries.

Leonardo da Vinci (1452-1519) made detailed studies of bird flight and human anatomy to understand the mechanics of movement. In (Codex on the Flight of Birds) he devised theories about flight based on observations he made of birds that seared through the air, how they used wind currents, and how they used their tails and wings to balance themselves while flying. Then in the (Atlantic Codex) he used these observations to design flying machines based on bird flight patterns.

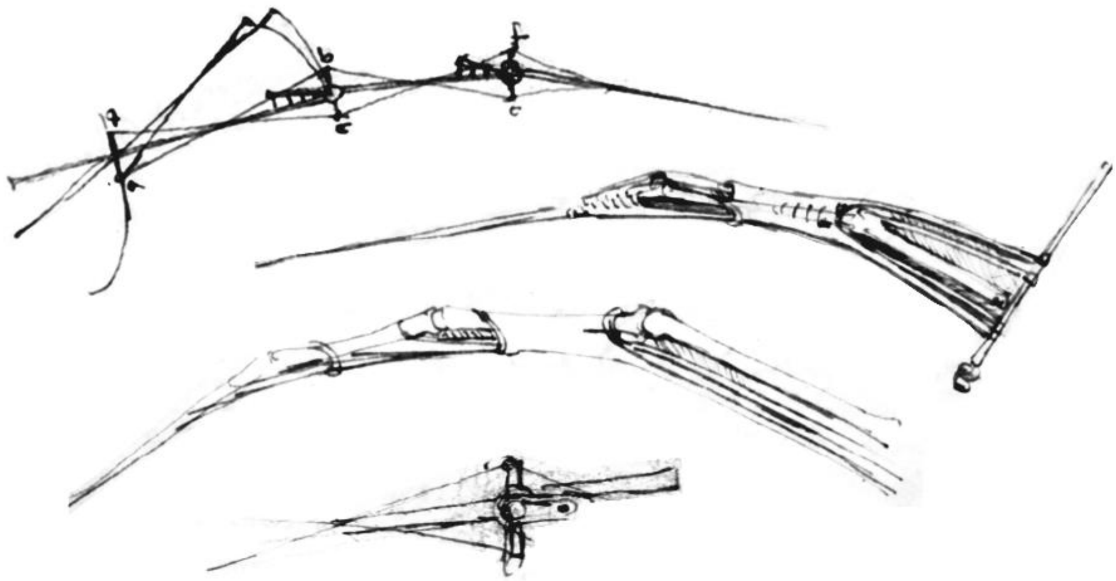

Figure 1-3 Da Vinci's sketches of bird wings, position of the joints and their angles of freedom (Adapted from Codex on the Flight of Birds)

Galileo Galilei (1564-1642), or the father of modern science, moved the study of movement to a whole new level. His studies on motion laid the foundation for classical mechanics. Relativity of motion, description of projectile motion, and describing object behavior in free fall are a few of his discoveries in measuring motion. He is also known for his famous work with celestial objects. Using his improved version of the telescope, he proposed the heliocentric model of the universe and shaped our current understanding of the universe (Dialogue Concerning the Two Chief World Systems - 1632).

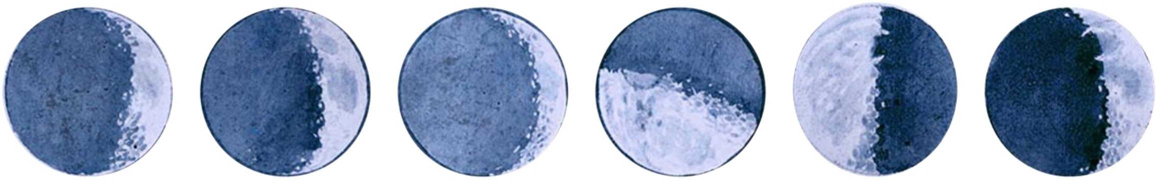

Figure 1-4 First realistic depiction of the Moon in history (G. Galilei, Drawings of the Moon, November-December 1609)

Other outstanding examples of movement analysis can be found in the works of Sir Isaac Newton (1643-1727). Newton's three laws of motion revolutionized our understanding of the movement of objects, paving the way of classical mechanics. Observing and studying the movement of objects falling to the ground, led him to proposal of his law of universal gravitation (Philosophiæ Naturalis Principia Mathematica).

Given these examples, we can conclude that the observation and study of movements, whether in living things or unanimated objects, can lead to advancements in science and our understanding of the world. However, mentioned scientists explored only the effects of movement and made discoveries based on them. The reason behind these movements, especially in animals and living things has not yet been explored.

1.3 Movement Analysis in Medical Sciences

The first systematic scientific work that dealt with the movement of animals is De Motu Animalium (On the Motion of Animals) by Aristotle. In this work, he tried to understand the mechanisms of animal movement, especially in relation to muscle contraction and the role of various body parts in producing motion. With our current knowledge, we can see that some of his insights were not accurate. This is why future studies were inclined towards employing experimental methods, rather than only relying on observations. Building up on Aristotle’s work, Claudius Galen (129-210 CE), studied the anatomy of animals and the function of muscles in movement. The novelty of his research was that he based his observations on the dissection of animals. With a big jump into the 17th century, René Descartes (1596-1650), focused on the element responsible for the control of animal movement. However not a popular idea today, he proposed the concept of animal spirits causing reflex actions. He viewed animals as machines. At this point, it seemed progress in this field required a different approach, or an innovation in terms of studying animal movement and behavior.

This innovation came about within the 19th century. With the creation of tools to capture and record animal behavior, scientists found a new approach. Etienne Jules Marey (1830-1904) developed chronophotography, a precursor to modern cinematography, to study animal movement. His work captured multiple phases of movement in one photo. At about the same time, Eadweard Muybridge (1830-1904) was known for his sequential photographs of animals (and humans) in motion, effectively breaking down complex movements into individual stages. He famously captured the sequence of a horse's gallop, proving that all four hooves leave the ground during a certain phase of the gallop.

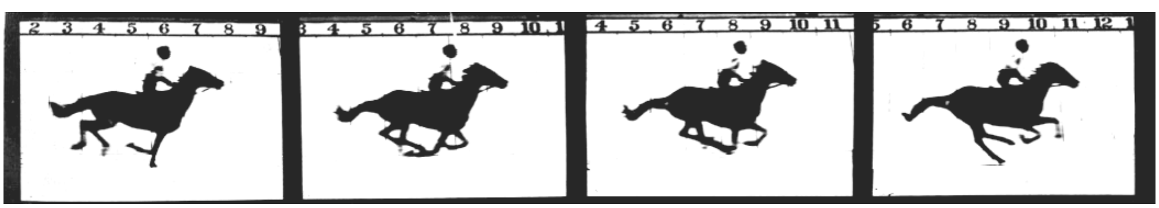

Figure 1-5 A slice of "Sallie Gardner" (frames 1 to 4 from left to right), owned by Leland Stanford; ridden by G. Domm, running at a 1.40 gait over the Palo Alto track, 19th June, 1878 (1878 cabinet card, "untouched" version from original negatives). Illustrated by Muybridge.

The field of animal study grew bigger with people interested in studying different species. Karl von Frisch (1886-1982) is known for deciphering the *waggle dance* of honeybees, which they use to communicate the location of food sources to other members of the hive.

With the start of the 21st century, the advent of advanced technology, and it being accessible to almost everyone in the world, researchers from every discipline have contributed to the field of movement and behavior study. Organizations such as Max Planck Institute and many universities around the world have formed research groups that focus on biomechanics, neuroscience and measuring movement for behavioral study. An example of modern analytical work on movement is the work of researchers at the University of Lethbridge which studies mouse's arm and hand movement to describe evolution of reaching (Naghizadeh et al., 2020).

1.4 Theory

Dynamic Systems Theory offers a framework to understand the behavior of complex systems through mathematical representations. When we think of animals, particularly cattle, as dynamic systems, it provides us with a methodology to model, understand, and predict their behavior.

One of the fundamental principles of Dynamic Systems Theory is that even a simplified model can yield insights into the system's overall behavior. For cattle movement, using a reduced or simplified model can help in understanding basic principles, which can then be extrapolated to more complex scenarios or even disorders like lameness.

In this context, animals' movements and behaviors are not just random actions but are the result of underlying patterns and processes. These patterns can be described using mathematical models, which capture the essence of movement dynamics and can be analyzed for various factors.

In the context of medical and behavioral studies, this theory proves advantageous. When a disorder in movement, such as lameness, is detected, understanding the fundamental dynamics can assist in diagnosing the underlying causes and designing effective interventions.

Additionally, the theory underscores the importance of real-time analysis. The behavior of dynamic systems can change over time, influenced by a myriad of internal and external factors. Thus, timely and regular observations, combined with the analysis offered by Dynamic Systems Theory, can lead to early detection of abnormalities or shifts in patterns, especially vital in health-related contexts.

Furthermore, the application of this theory in conjunction with modern technology, such as AI and machine learning, can revolutionize our approach to understanding movement and detecting anomalies. AI algorithms can process vast amounts of data, identify patterns, and make predictions, all aligned with the principles of Dynamic Systems Theory. This combination not only enhances our understanding but also improves the accuracy and efficiency of detection and intervention methods.

In summary, Dynamic Systems Theory offers a comprehensive framework to understand, model, and analyze the movement patterns in cattle. By adopting this approach, we can achieve a deeper understanding of movement, its disorders, and devise effective methodologies to address challenges related to it.

1.5 Our Question

Motor disorders including lameness are of the health concerns in feedlot cattle that affect their welfare, productivity, and profitability (Kaniyamattam et al., 2020; Robcis et al., 2023). In feedlot cattle, lameness has a significant prevalence, ranging from 13% (Wells et al., 1993) to 69% (Hedges et al., 2001) and so adds to the production costs. Non-infectious claw diseases like sole ulcers and white line disease and infectious diseases such as interdigital phlegmon, interdigital dermatitis, and digital dermatitis are the causes of more than 90% of lameness cases (Clarkson et al., 1996; Logue & Kempson, 1993; Murray et al., 1996; Somers et al., 2005). The total cost of handling lame cattle in the North America regions ranges from $4 to $22 billion per year (Robcis et al., 2023). Lameness diagnosis is also expensive and may require specialized expertise, a visit by a veterinarian, and treatment. If diagnosis of lameness can be made earlier and more precise and treatment applied sooner, lameness duration could be decreased with a significant increase in cattle health and reduction in costs. Rating scale methods for diagnosing lameness are based on devising a locomotion scoring system and then assigning a score to each cow based on the existence of certain features. These methods differ in rating scale and regions of interest (Breuer et al., 2000; Dyer et al., 2007; Flower & Weary, 2006; Gleeson et al., 2007; Leach et al., 2009; Manson & Leaver, 1988; O'Driscoll et al., 2009; Olmos et al., 2009; Wells et al., 1993). The problem with these rating scaling methods is inconsistency in scoring since it requires the observer as the primary decision maker for scoring. The degree of agreement between scorers will change by time based on their training and experience (Brenninkmeyer et al., 2007; Hollenbeck & Sackett, 1978; March et al., 2007). Transitioning from manual methods to more intelligent and automatic approaches can solve the inconsistency problem. Given this information, we can grasp the industry's need for a robust method to analyze movement. A method that can systematically define the movements of a feedlot cattle and lay the foundation for behavioral studies. Now our question is: *Can we propose such a method to compliment the current methods in analyzing and describing the movement of a cow?* In order to answer this question, we need to familiarize ourselves with the methodology of studies done in this field.

1.6 Current Solutions

Currently, there are three approaches to automizing this process. First approach the lameness detection process is the use of sensors. Since the 1980s, studies have investigated the potential of integrating sensors and data the analysis of sensor data in early detection of lameness (Hogeveen et al., 2010). Some examples can be studies focused on measuring ground reaction force (Rajkondawar, Lefcourt, et al., 2002; Rajkondawar, Tasch, et al., 2002; Rajkondawar et al., 2006), automatic measurement of weight distribution (Neveux et al., 2006; Pastell et al., 2006), recording footsteps in gait pattern (Maertens et al., 2011), analyzing gait and/or activity using accelerometers (Chapinal et al., 2011; Pastell et al., 2009). These approaches provide an accurate way of detecting lameness early on, but the sensor technologies used need to be very cost-efficient as some are hard to set up and take up a considerable space. Diagnostic methods for early and precise detection of lameness could benefit from the use of AI. Machine Learning (ML) and AI have shown promise in addressing these challenges by enabling the development of computer-based systems for diagnosing motor abnormalities. Several studies have already explored the potential of this approach. For example, a recent study has applied Classification and regression tree

(CART) algorithm on the integration of lameness scores and several features obtained from thermal images to diagnose lameness (Coşkun et al., 2023).

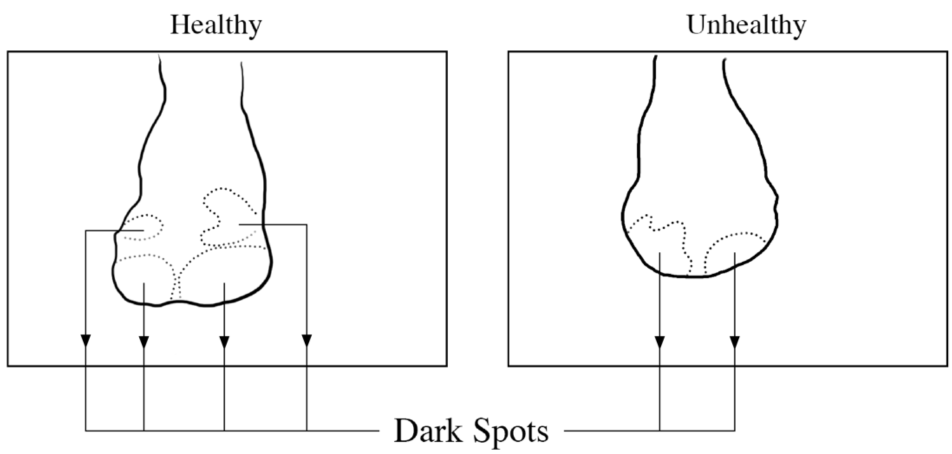

Figure 1-6: Physical condition and thermal image of fetlock joint of healthy and unhealthy (lame) animals. Adapted from (Coşkun et al., 2023).

Also, with Deep Learning being one of the most promising tools for this problem, 34.1% of the studies focusing on lameness detection apply video analysis and image processing in their methods (Nejati et al., 2023). Limitations of these approaches are the large number of training examples required and labeled and limited conditions. Synthetic data generation using 3D models is a novel approach to overcome these limitations by providing ground truth data that can be used to train Deep Learning models to classify behavior (Bolaños et al., 2021).

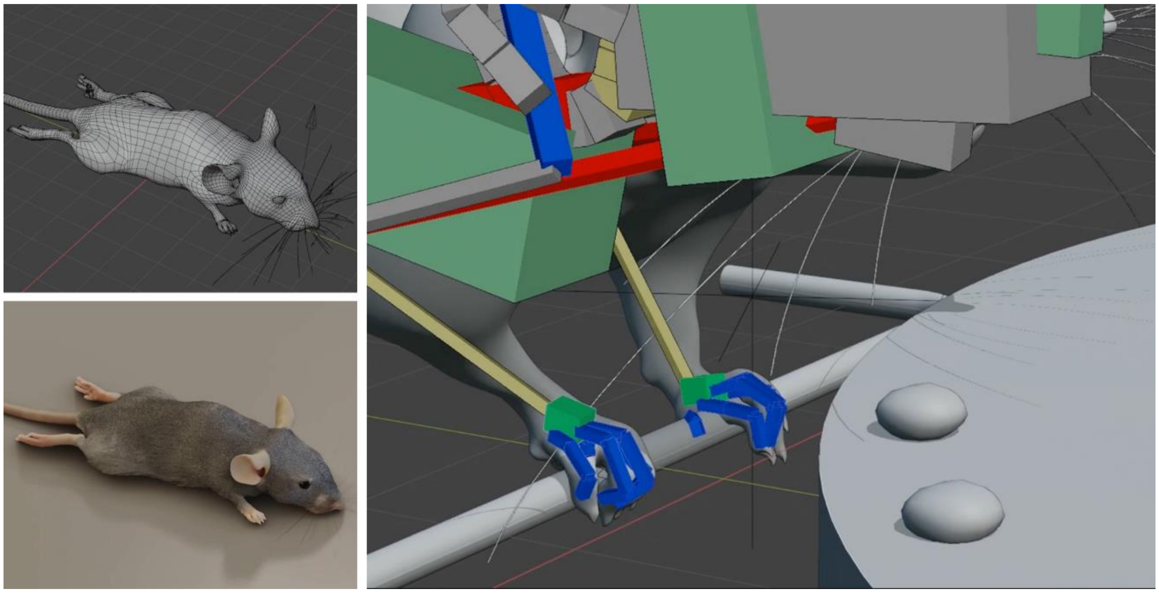

Figure 1-7 Use cases of 3D model generation in mimicking real-world data. Adapted from (Bolaños et al., 2021)

1.7 Significance of Synthetic Data Augmentation

As a relatively modern area of research, Synthetic Data Augmentations has received much attention in healthcare and industry. A review done by (Gonzales et al., 2023) identifies several use cases of synthetic data in healthcare:

1. **Conditional Prediction Research**: A research project created a synthetic dataset of health histories by varying parameters in real-life datasets and used them to test different policy scenarios (Davis et al., 2010).

2. **Hypothesis Testing**: (Ngufor et al., 2019) used real publicly available and synthetic data sets to test a machine learning framework developed to predict longitudinal change in glycemic control measured by hemoglobin A1c (HbA1c).

3. **Simulation Research**: In a study by (Enanoria et al., 2016), a synthetic dataset created to represent California's population was used to assess the impact of different interactions in the transmission of measles.

4. **Health IT Development**: Software solutions such as The Smart Platform (Mandl et al., 2012) provide synthetic records that mimic real patient data; enabling developers to create tools to solve real-world problems.

5. **Education**: Oregon Health & Science University created a synthetic dataset of 446000 patients with cardiovascular conditions as the teaching material for a course with machine learning components. (Laderas et al., 2018)

Building upon the diverse applications in healthcare, Synthetic Data Augmentation extends significantly into the industrial sector. Just as it proves useful in medical research, synthetic data poses itself to be a versatile tool in various industries. Some examples can be:

1. **Optimising Data Utility**: Industries such as pharmaceutical industry are realizing the potential of synthetic data in offering optimised data utility and sharing, and a higher level of privacy. (James et al., 2021)

2. **Enhancing Manual Processes**: Synthetic data generation techniques, including computer simulations and modeling, are being used to overcome the scarcity of annotated samples in various industries such as microelectronics. (Phoulady et al., 2023)

3. **Addressing Challenges in Machine Learning**: With the advent of machine learning solutions and their effectiveness in nearly all the

industries, tools that can enhance their performance are valued. Some studies have explored ways to use synthetic data in this way. (Ortego et al., 2020)

4. **Improved Automated Visual Inspection**: Synthetic data augmentation using Generative Adversarial Networks s has shown to reduce the labeling workload by more than 50% in automated visual inspection. This helps mitigate data imbalances and enhance the performance of unsupervised defect detection models employed in many industries. (Rožanec et al., 2022)

1.8 My Work

In the realm of this research, I delved into the potentials and intricacies of synthetic data generation, specifically targeted at cattle lameness detection. Drawing from single-camera recordings of cattle navigating a specific lane, I distilled this data to craft a 3D synthetic model. This model was designed to encapsulate and represent the varied nuances of cattle behavior and their patterns of movement.

One of the standout features of this methodology is its inherent capability for data scalability. By systematically introducing slight variations and modifications to the foundational model, it became possible to generate a dataset that expanded the boundaries typically set by conventional real-world data collection techniques. This expansion addresses and potentially mitigates some of the challenges that researchers often face, especially when it comes to gathering expansive sets of real-world, labeled data, a requirement that's indispensable for the effective training of advanced deep learning algorithms.

It was imperative that the synthetic data not only served as a volume multiplier but also mirrored the varied dynamics of real-world cattle behavior. By ensuring such a match, the model's outcomes and predictions can find direct applicability and relevance in real-world scenarios, making the results not just theoretical but practically actionable.

Armed with this enriched dataset, the subsequent step was the training of a deep learning model. The primary focus here was on the accurate detection and subsequent analysis of specific features that are indicative of cattle lameness. Preliminary results indicated a marked improvement in the performance metrics of the model that was trained on the synthetic data, especially when juxtaposed against a counterpart model that had its training confined only to real-world data.

But the ambit of this research was not just restricted to the realm of data and analytics. There was a strong emphasis on real-world integration. The idea was to transition the trained model into widely used digital platforms, thereby making it a tool of practical utility. For instance, integrating this model into smartphone applications can usher in a paradigm shift in how farmers, cattle breeders, and even veterinarians approach the challenge of lameness detection. Such a tool, which is both handy and effective, can catalyze early detection and timely intervention, with the downstream effects potentially leading to the overall improvement in cattle welfare and notable reductions in associated costs.

To encapsulate, the trajectory of my work was a synthesis of synthetic data generation, state-of-the-art deep learning techniques, and a vision for real-world application, all converging to offer a novel solution to the longstanding challenge of cattle lameness detection.

Chapter 2: Experiment

2.1 Subjects and Data Acquisition

The real footage from feedlot cattle used this project are provided by Mohajerani Lab. The videos are captured using a GoPro Hero 10. The camera is set up front of a walking isle for cattle and cattle are led to pass through that isle while the camera is recording. The camera is positioned in a constant position during the recording with a varying distance of 1 to 3 meters from the cattle. Weather conditions in the recordings vary between sunny, cloudy, or snowy, changing the light effects and shadows in between the recordings.

2.2 Preprocessing

The camera used to record the footage from the feedlot uses an ultra-wide lens. This kind of lens introduces a distortion to the picture, especially at around the edges of the image. To make these footages represent reality as best as possible, a simple distortion-fix algorithm is applied to the footage before using them for the next steps. The code for this algorithm is provided in the supplementary information.

2.3 Three Pipelines for Creating Deep Learning Models

To illustrate the effects of using a cow model based on our hypothesis, we train three different DL models using different pipelines shown in Figure 2-1. These pipelines are described here.

A. Manual Model: This model is created using real footage acquired from feedlot with manual labeling and no modifications applied to the data. This is the classical pipeline for training a DL model. (Figure 2-1 A)

B. Synthetic Model 1: This pipeline is similar to the manual model pipeline, but with synthetic data and generated labels instead of real data and manual labeling. In comparison, we expect this method not to be as efficient as the manual model. (Figure 2-1 B)

C. Synthetic Model 2: The pipeline for augmenting data to fill the gaps in real footage. The difference between this pipeline and the pipeline for synthetic model 1 is that the real footage undergoes some processing and will be different from the original data. We hypothesize that this pipeline would have an increased accuracy in validation. (Figure 2-1 C)

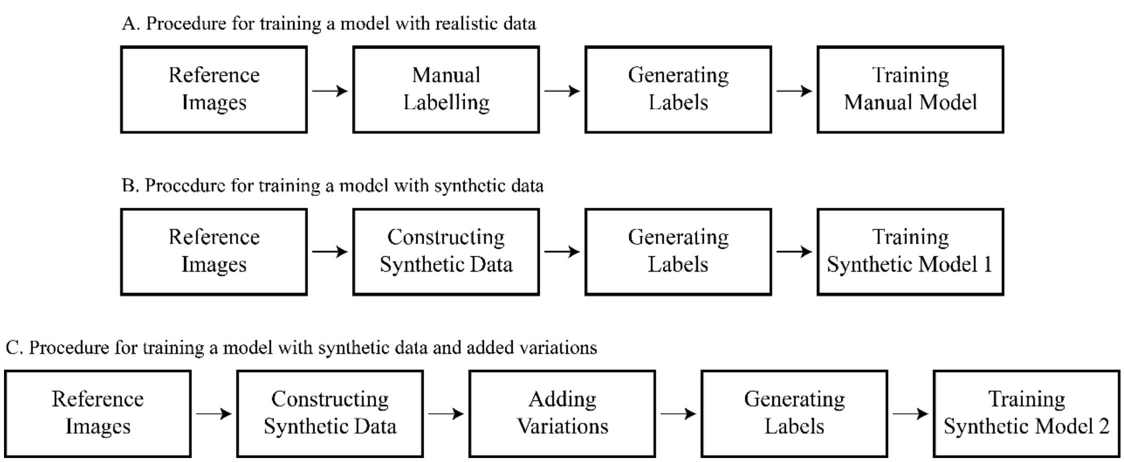

Figure 2-1 - General Pipelines for Training Three DL Models. A. Procedure for training a model with real footage. B. Procedure for training a model with synthetic footage. C. Procedure for training a model with synthetic footage and added variations.

2.4 Creation of Cow Model

We use a generic cow model that resembles the appearance of the typical feedlot cattle. Figure 2-2 depicts the basic steps of creating a 3D model of the cow. Figure 2-2 B shows the basic appearance of this model. It is created based on the anatomical features of an average feed cattle. This model can be extended with hair and skin simulation and provides the ability to implement physics. By using the real cow in the footage as reference, Figure 2-2 A, we match the size, shape and use texture painting to change the color of the cow's coating (including spotting) to be similar to the reference. To be able to change the model's appearance and move its components in the way that a real cow does, we need to create the bone structure for the cow. This process is called Rigging. Figure 2-2 D shows the added armature to the basic soft-body model. Using an atlas for cow anatomy, Figure 2-2 C, we have matched the bone structure to be an exact representation of a feedlot cattle's anatomy. Then we modified the joints and enabled movement constraints based on the limits that are in a real cow's degree of freedom in movement. Finally, by adding hair particles and textures we achieve the realistic synthetic cow shown in Figure 2-2 E. This version will be used in the data augmentation process.

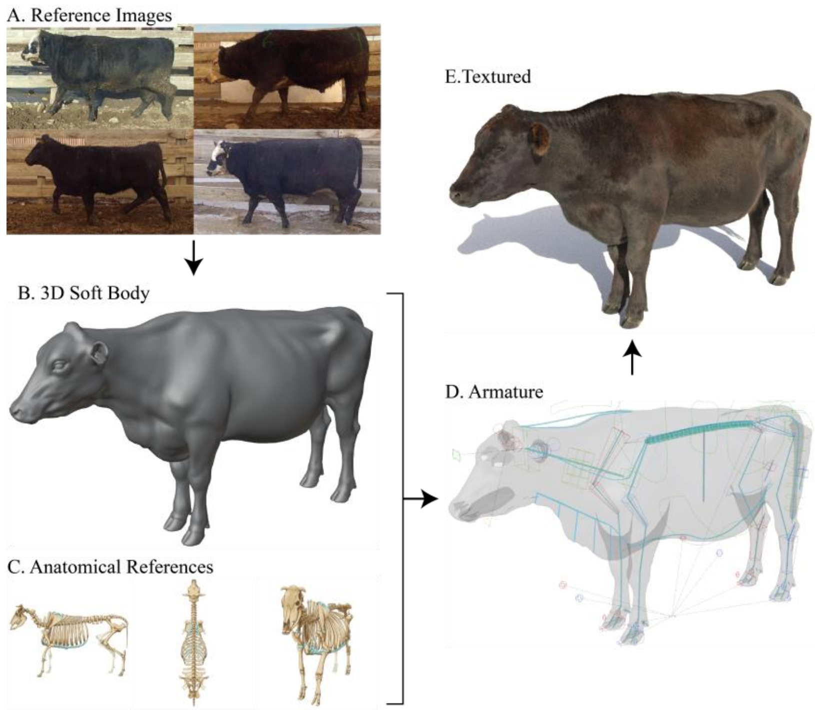

Figure 2-2 - Process for creating the base 3D model of the cow. A. Reference pictures for creating manual animation. These pictures are frames taken from available video footage. B. The base 3D model of the cow that we use for feature-matching with the cows in our footage. C. Anatomical Reference of the cow (reference). D. Added armature to the cow based on the anatomical reference. This armature plays the role of the skeleton of the animal and is used in the animation process. E. Textured version of the cow with all the added details to reflect the visual appearance of a real cow.

2.5 Animation

The model is now imported into a scene in Blender, all environment details such as lighting (angle and the intensity of the light), environment elements (background objects and flooring), and camera specifications (distance from subject, field of view, and focal length) are adjusted in the scene to exactly match the real video. Model specifications (position, body movements, angles, etc.) are then matched to the real video in every 5 frames. We use linear interpolation method in Blender to find the specifications of the model in between these frames. Following this method, we create a synthetic version of the real video that matches the real footage.

2.6 Synthetic Data Augmentation

To create more data that can be used for Deep Learning model training, we make changes to the synthetic animation we created in the previous step. These changes include:

- **Adding noise to the movement:** A simple noise is added to the movement so that it slightly differs from the real video.
- **Changing the cattle appearance:** This is done by performing texture painting on the synthetic mode, the main color of the cow is changed along with some addition or change of the spotting on the cow.
- **Changing the lighting:** Lighting details can be changed in the blender scene to change the time of day for the synthetic footage, having different lighting conditions in the videos can provide more variety in the training dataset for Deep Learning models.
- **Changing the viewing angle:** As the synthetic model is a three-dimensional model, we can change the viewing angle to create a synthetic data that would be equivalent to recording the real video by putting the camera in a different position. Since the 3D model is created based on a real cow's anatomy, we can expect it to resemble the real cow if the real camera's position was changed.
- **Changing the environment:** Using Blender and Unreal Engine, we can create photorealistic environments to replace the environment of the real cow. We can change the weather, flooring, and other related conditions.

By using combinations of the above-mentioned changes, we can create tens of videos only based on a single real footage than can significantly help with the lack of data in training Deep Learning networks.

2.7 Realistic Rendering

After all the desired changes are applied to the synthetic videos, we use high-fidelity rendering capabilities of Blender and Unreal Engine to generate a life-like video that resembles a real cow in the created condition. We use Cycles Render Engine in Blender and Nanite virtualized geometry system in Unreal Engine to render the footages. The generation of these videos take about 5 hours each on our hardware (Intel Core i7-13700K, RTX 4090).

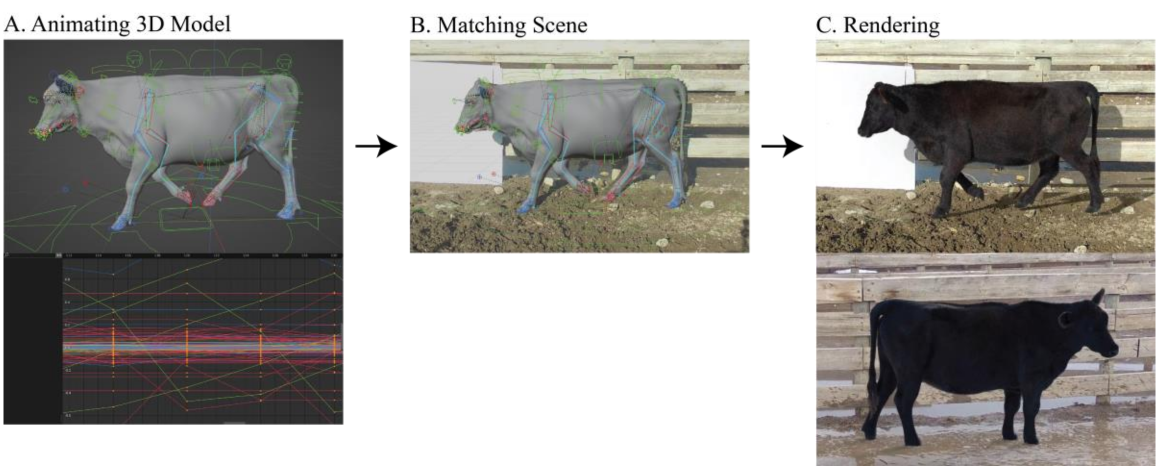

Figure 2-3 Steps ins Constructing Synthetic Data. A. Process of animating the soft body based on the real-world footage, rigging view at the top and animation graph at the bottom panel. B. Using the real-world footage background with the augmented data. C. Result of realistic rendering process of cow models that include realistic coating, fur, and other features.

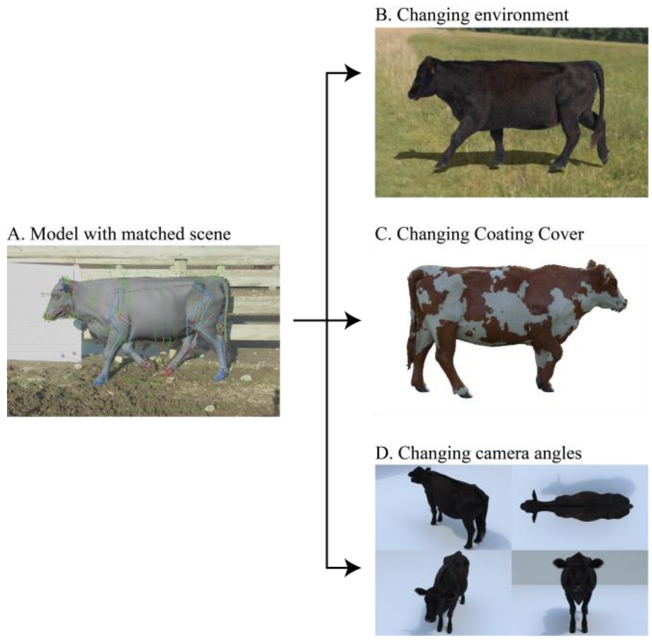

Figure 2-4 Steps in Adding Variations to the augmented cow. A. The basic model comprised of only the soft body that has been matched to the real-world cow footage. B. The cow model in A. in a different augmented environment. C. The cow model in A. with a procedurally generated coating. D. Display of different angels available for data augmentation.

2.8 Automatic Labeling and Data Export

The labels and required information from the synthetic models are readily available due to the possibility of tracking objects in Blender and Unreal Engine. However, since all the labels are present in our model, we need to reduce them to the ones only visible from the point of view of the camera. To do so, we use a technique called Ray Casting in Computer Graphics. In this method we check if a direct line of sight exists between the camera in the environment and the marker that we want to track.

Given two points, $P_1 = (x_1, y_1, z_1)$ (position of the camera in the environment) and $P_2 = (x_2, y_2, z_2)$ (position of the marker on cow's body) the equation of the ray $R(t)$ that starts at P_1 and passes through P_2 is given by:

$$R(t) = P_1 + t(P_2 - P_1)$$

Where:

- $R(t)$ is a point on the ray.
- t is a parameter. When $t = 0$, $R(t) = P_1$, and when $t = 1$, $R(t) = P_2$. For values of t between 0 and 1, $R(t)$ will be a point on the segment between P_1 and P_2. For $t < 0$ or $t > 1$, $R(t)$ will be on the ray but outside the segment.
- P_1 and P_2 are vectors representing the coordinates of the points.

We check for values of t between 0 and 1 to see if there is a point that fits in the equation. If calculated value for t is anything other than 0 or 1, it means that there are points breaking the line of sight from the camera to the point. Thus, we omit the marker at point P_2 from the exported list of labels.

After the rendering process is completed, we use custom scripts (see supplementary information) to export information such as joint positions both in 2D and 3D. Further computed information about these videos are also generated based on the exported information. For example, a simple Python script is used to calculate the angle between joints in the cattle.

2.9 DeepLabCut

We have chosen to use DeepLabCut as an approved standard in neuroscience communities. We train the three models mentioned before using DLC and the formats we use for our data and labels are based on DLC guidelines. We also use DLC in labeling the images in the manual model pipeline (see Figure *2-1* A.). Finally, validations and tests on new data is also performed using DLC.

The configuration we used in training our DLC models are listed in the table below (Unmentioned properties are set to default).

Table 2-1 List of parameters used with DLC

Property	Value
DLC Library Version	2.2.3
Multi Animal	Disabled
Input Videos	Manual: 2 videos – 400 frames
	Synthetic 1: 2 videos – 400 frames (Synthetized version of footage used in Manual)
	Synthetic 2: 16 videos – 3740 frames (Footage used in manual and Synthetic 1 + variations)
Input Video Dimensions	1280 x 720 pixels
Training Fraction	0.95
Resnet	50
Batch Size	8
Augmenter Type	imgaug
Shuffle	1
Max Iterations	200000

2.10 Validation

To ascertain the robustness and generalization capabilities of the trained models, it is pivotal to test their performance on data they have never encountered before. This not only provides a benchmark for their reliability but also offers insights into how effectively the augmented data mimics real-life scenarios, especially when compared with models trained exclusively on real data.

For the validation phase, a unique video was chosen that exhibits characteristics distinct from those the models were trained on. This ensures that the models are evaluated under conditions that mimic potential real-world challenges they might encounter. The following describes the validation process.

1. Video Analysis: The chosen video was processed through both the model trained on real data and the one trained on augmented data using DeepLabCut.

2. Extraction of Predictions: For each frame in the video, the position and behavioral predictions made by each model were extracted and saved.

3. Comparison of Predictions:

    a. Ground Truth Establishment: A manual annotation of the video was done to serve as the ground truth against which model predictions would be compared.

    b. Metrics: The accuracy of the predictions was measured using standard metrics such as Mean Euclidean Pixel Distance (Mean

Pixel Error) for positional data and polar plots for better illustration of the models' performance.

The outcomes from this validation test will be used to infer the strengths and weaknesses of each model, drawing contrasts between their performances and ultimately elucidating the benefits (or drawbacks) of using augmented data in behavioral studies of cows. In the next chapter, I will present and analyze the results encountered in this project.

Chapter 3: Results

To evaluate the effectiveness of using synthetic data generated from three-dimensional (3D) models to train neural networks for the analysis of feedlot cattle movements, we conducted a series of experiments using DeepLabCut. Our results showed that synthetic data can be used to effectively fill in the gaps in real data to train neural networks for the analysis of cattle movements.

In the first experiment, we recorded two real videos (400 frames total) of feedlot cattle walking and generated two synthetic videos that were almost identical to the real videos. We then trained two neural networks using DeepLabCut, one on the real videos and the other on the synthetic videos. When we tested these two models on a new video of a real cow walking, we found that both models performed similarly, although the model trained on real videos performed slightly better.

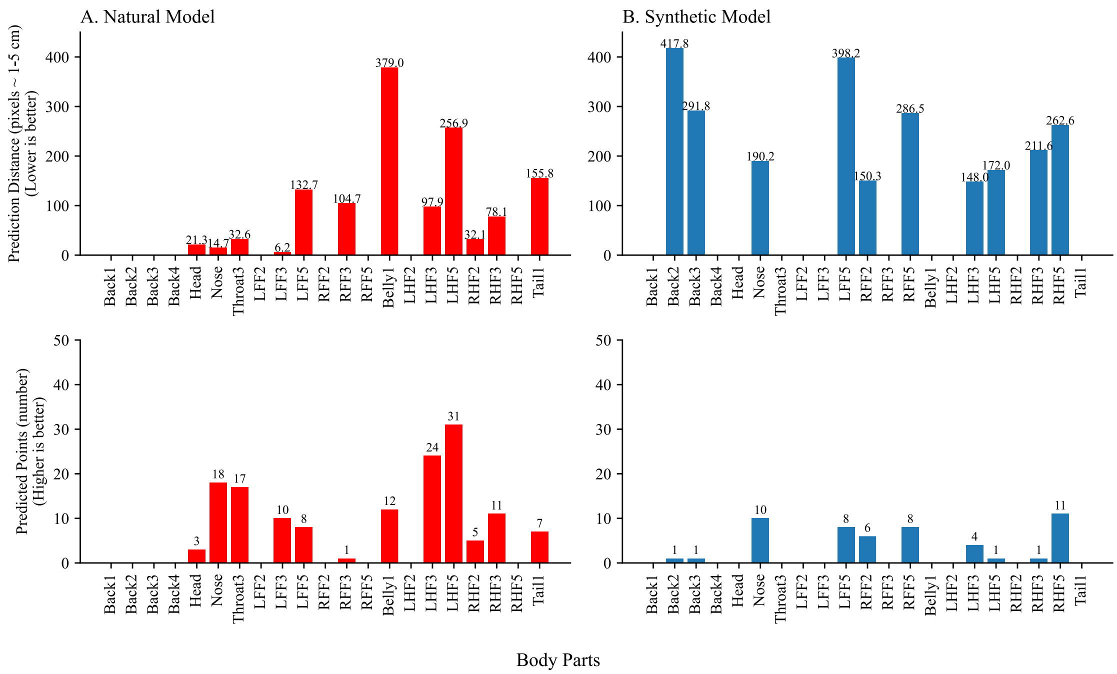

Figure 3-1 Comparison between the natural model and synthetic model, trained on 400 frames. A. Results of the manual model trained on real-world data. B. Results of the manual model trained on augmented data without variations.

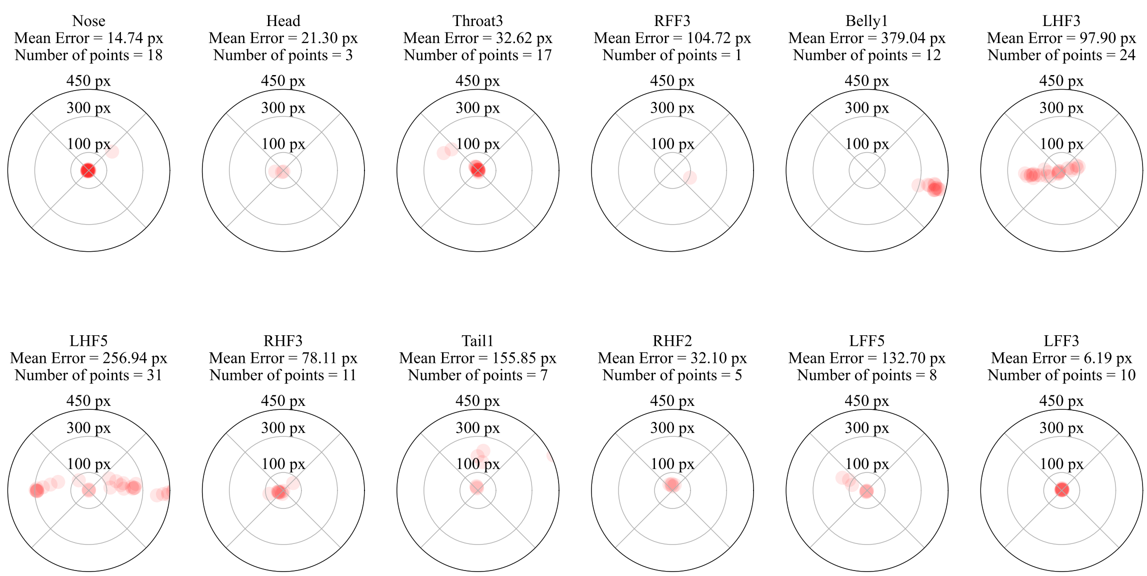

Figure 3-2 Polar plots showing the accuracy of the natural model, trained on 400 frames.

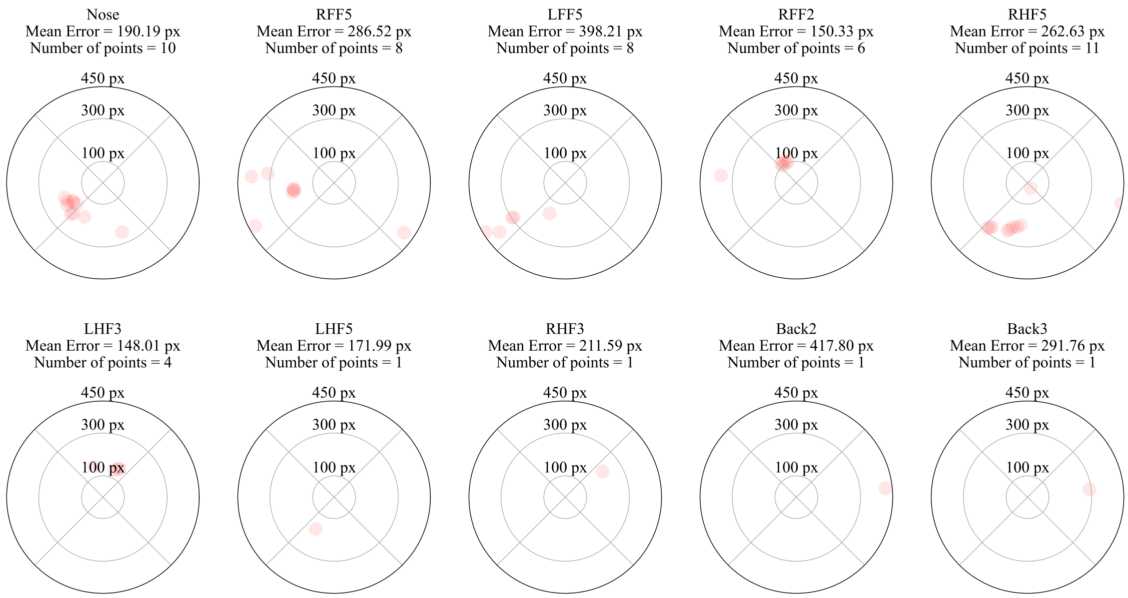

Figure 3-3 Polar plots showing the accuracy of the synthetic model, trained on 400 frames.

In the second experiment, we added variations to the synthetic videos by changing the color of the cows, the environment, and lighting. We created twelve additional videos (3740 frames total) with the variations. We then used these new synthetic videos along with real-world videos to train another neural network using DeepLabCut. When we tested this model on a new video of a real cow walking, we found that it performed much better than synthetic models trained before, but comparing its performance with the manual model's requires additional explanation.

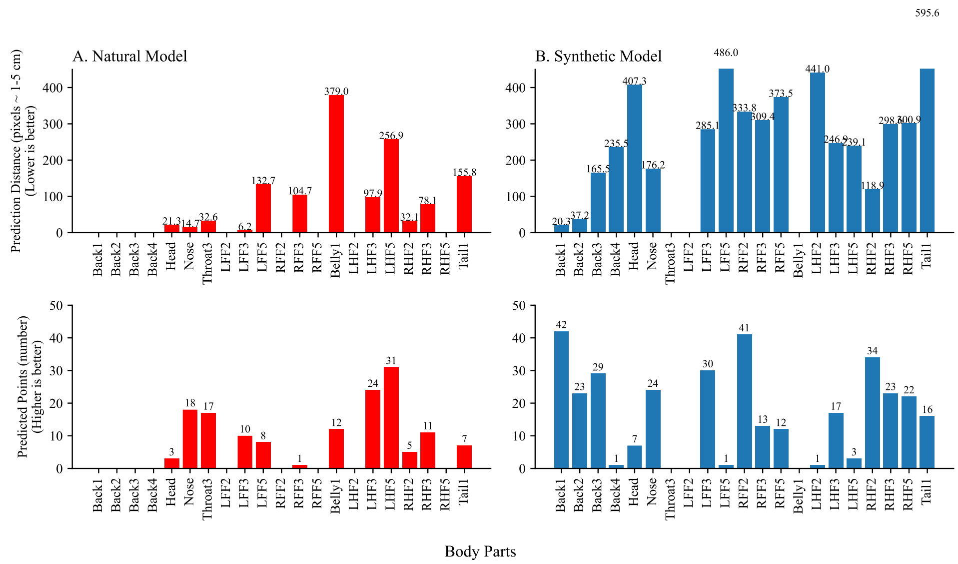

Figure 3-4 Comparison between the natural model and synthetic model trained on 3740 frames. A. Results of the manual model trained on real-world data. B. Results of the manual model trained on augmented data variations.

Figure 3-5 Polar plots showing the accuracy of the synthetic model with added variations, trained on 3740 frames.

As it can be seen in Figure *3-5*, the accuracy of this synthetic model is not necessarily better than the manual model in the points that it has predicted. It performs better in predicting some points that the manual model has failed to predict at all. For example, the manual model has failed to predict any of the points creating the arc at the back of the cow, a task that the synthetic model has succeeded in with acceptable accuracy.

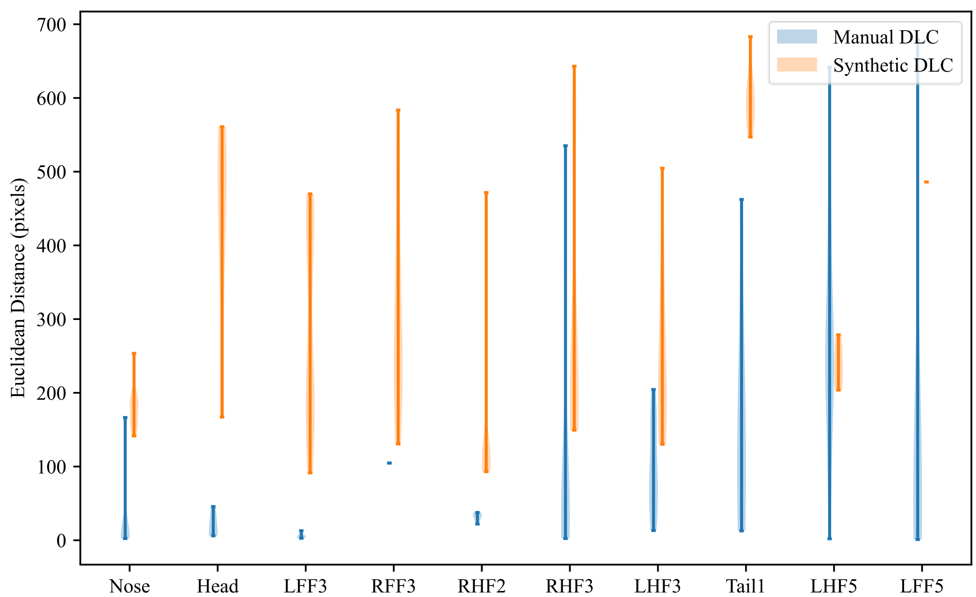

Figure 3-6 Violin plot showing diversity in common predictions of manual and synthetic model (with variation). The synthetic model performs poorly in some of the labels predicted by the manual model, but it thrives in making predictions that are not common with the manual model.

These results suggest that the process of generating synthetic data for training neural networks can be effective for the analysis of cattle movements. By using 3D models to generate synthetic data, we can quickly and cost-effectively generate large amounts of training data, which can improve the accuracy and efficiency of behavior analysis.

However, synthetically generated data cannot entirely replace the use of real-world data in this matter. The data generation process can be tuned to target the gaps in real-world data. In our case, after examining the training data for both models and the data used for validation, we can see that the randomly generated data in procedural data augmentation has created a coating for the cow that is similar to the validation data. This means that we were successful in creating useful training data without the need for performing additional recording.

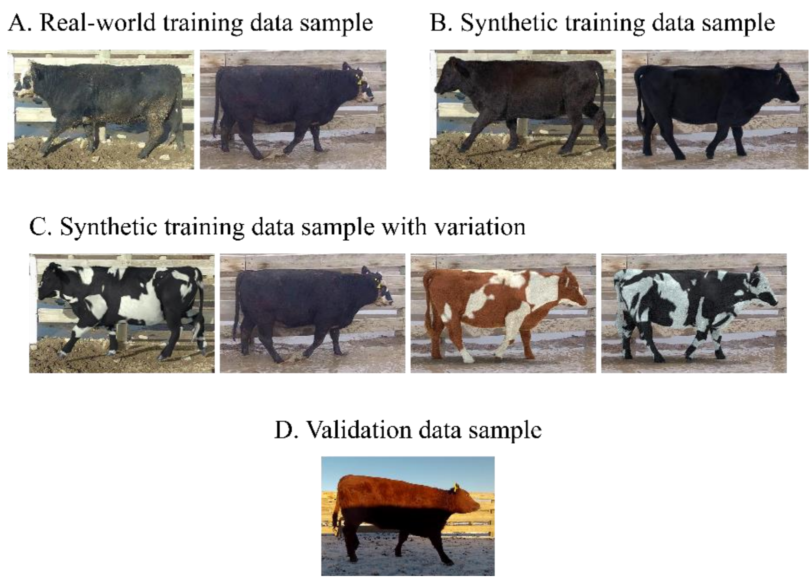

Figure 3-7 Samples of data used for training the three models described and validation. A. Sample of the real-world dataset. B. Sample of the synthetically recreated dataset. C. Sample of the synthetically recreated dataset with added variations. D. Sample of validation data. It can be seen that the random variations shown in C. have resulted in similarities with validation data.

It should be noted that the current results are not acceptable for using the model in a real-world use-case. These models need to be trained on a larger body of data to reach the acceptable confidence rate. The current results are only meant to show the potential of using synthetic data generation in training machine learning models.

Chapter 4: Discussion

4.1 Summary

In this research, a significant challenge was addressed: the early diagnosis of motor disorders in feedlot cattle using a blend of artificial intelligence (AI) and 3D modeling techniques. AI, particularly Deep Learning, holds immense promise in offering an objective, consistent diagnostic tool. However, one primary limitation has been the dependency on extensive and varied datasets.

The approach taken involved creating a 3D animation of cattle gait patterns, derived from a 3D digital cow model and several one-camera video recordings. The principle behind this was to generate synthetic data. By introducing controlled variations in the 3D models, we expanded the training set for the Deep Learning model, aiming to enhance its capability for joint detection and tracking, and ultimately, abnormality detection in gait patterns.

4.2 What I achieved

The results of this research can bring about a new era in behavioral study. Not being limited by the scarcity of training data, Deep Learning models can be easily developed and used to analyze specific behavioral patterns. For example, the early diagnosis of lameness and other motor disorders in feedlot cattle could significantly improve animal welfare, increase productivity, and reduce the economic losses. This method can also be applied to other subjects. Given the appropriate base 3D model of the subject, one can easily follow the steps mentioned in this study to achieve the same results. Compiling the final product of this study (Deep Learning model combined with the analysis tools) into a more accessible, user-friendly diagnostic tool, such as a smart-phone-based system, may also broaden the reach of this technology, enabling earlier intervention even in remote or resource-poor settings. In the following, I explain some of the problems that this research aimed to overcome.

4.3 Problem 1: Difficulty in Acquiring Data

Given the conditions of feedlots and the cattle, the task of collecting data for studying their behavior through the means of AI and Neural Networks is not always easy, as they require an extensive body of data that covers most aspects of the subject's behavior. Providing this data from feedlots requires establishing recording equipment and maintaining them for long periods of time. Feedlots in Alberta are prone to severe weather conditions, such as extreme gusts of wind and below-freezing temperatures. Considering these issues, the cost for maintenance of the required recording equipment can be overwhelming for many feedlot owners. Moreover, the data acquired this way needs preprocessing before it can be used as training data for Neural Networks. In this project and with our proposed method, we have total control over the generated data and can simulate the conditions we wish with little effort. Another advantage of this method to the traditional data recording is in the customizability of data. A data recording is final, and the experimenter cannot change any aspect of it. This is not the case with the augmented data. Also, real-world data often lack variability due to the limited conditions under which they are captured. For instance, the data might primarily come from a specific time of day, specific breeds, or specific behaviors. Our approach allows researchers to introduce artificial variability, enhancing the model's ability to generalize across different scenarios. Data imbalance is another issue in this matter; some poses or behaviors of cows might be rarer than others, leading to imbalanced datasets. Augmenting data can help balance the dataset by artificially increasing the number of instances of the rarer poses or behaviors.

4.4 Problem 2: Annotation Challenges

In order for the real-world data to be used by Neural Networks, they need to be labeled by a human rater. There are many challenges in this step that will be trivial in our approach. We describe some of these issues in the following.

1. Expertise Requirement: Accurately annotating certain types of data, like cow poses or specific behaviors, often requires domain-specific knowledge. Feedlot operators or animal behaviorists might be needed to ensure accurate labels. Such expertise can be hard to find and might be costly.

2. Time-Consuming Process: Manually labeling data, especially large datasets, can be extremely time-consuming. For complex tasks like pose estimation, each image or video frame might require meticulous labeling, delineating specific parts of a cow or identifying intricate behaviors.

3. Consistency: Maintaining consistency across annotations is challenging, especially when multiple annotators are involved. Different individuals might interpret ambiguous situations differently, leading to inconsistencies in the dataset.

4. Ambiguities: There can be situations where the right label isn't clear due to the quality of the data, occlusions, or inherent ambiguities in the behavior or pose. Determining the right label for such data points can be challenging.

5. Scalability: As deep learning models thrive on large datasets, the sheer scale of the data that needs to be annotated becomes a challenge. While automated tools and platforms can aid the process, the core task often remains manual and labor-intensive.

6. Cost: High-quality annotation, especially when requiring domain expertise, can be expensive. This is not just in terms of direct monetary costs but also in terms of opportunity costs, as the time and resources could be used elsewhere.

7. Dynamic Environments: In the context of cow pose estimation in feedlots, the dynamic nature of the environment adds to the challenge. Cows might be moving, interacting with each other, or partially obscured, making the annotation task even more complex.

8. Annotation Maintenance: As more is learned about the problem domain or as models evolve, there might be a need to revisit and update annotations. This re-annotation can be as challenging as the initial annotation process.

9. Tool Limitations: While there are many annotation tools available, not all of them might be suited for specific tasks. There might be a need to customize tools or develop new ones, adding to the challenges.

4.5 Problem 3: Environmental Challenges

In the case of studying quadrupeds such as cows, and even in studying small animals such as mice, there are environmental factors to be considered (Bermudez Contreras et al., 2022). Factors that can be further complicated when we modify the environment or the animal body itself. By using deep learning solutions to analyze behavior from visual cues and the proposed method, we strive to analyze animal movement with minimal alterations to its natural living conditions. Thus, we can be more confident in the results obtained this way.

By harnessing data augmentation, many of these challenges can be alleviated. With augmented data, annotations are readily available since we control the data generation process. This bypasses the need for manual labeling, ensures consistency, and significantly reduces the time and costs associated with the annotation process. Also, by employing simple mathematical tools, such as ray casting (explained in 2.8), we can simulate the realistic labeling outcome of an expert.

4.6 Caveats and limitations

There are number of caveats to the synthetic cow model presented here.

1. Terrain: The synthetic cow model was primarily based on cattle moving in controlled environments. The nature of the terrain - be it uneven, sloped, or variable in texture - can significantly influence gait. Real-world settings would introduce a multitude of terrains, each of which might affect the gait differently, which our model might not have been exposed to.

2. Direction: The directionality of movement in the training data was mainly linear, given the single-lane recordings. In more open settings, cattle might move laterally or take more complex paths, introducing variables not present in the model.

3. Speed: Movement speed is a variable factor. While the model accounted for a general speed range, it's important to note that variations in speed, influenced by factors like age, health, or external stimuli, might present unique gait patterns.

4. Model Limitations: It's crucial to recognize that while our 3D model offers valuable insights, it remains a simplified representation of the real world. As with all models, the exclusion of certain variables or nuances is inevitable.

Despite introducing variations in the 3D models, enhancing the generalizability of our Deep Learning model requires more diverse environmental and behavioral factors. The next phase of research could study the inclusion of more diverse scenarios to further enrich the synthetic dataset.

The current methodology demands significant manual intervention, especially in synthetic data generation. This presents an opportunity for future AI research to develop more automated processes. While this research underscores the effectiveness of synthetic data in Deep Learning applications, transitioning from a research prototype to a real-world application requires considerations beyond just the technical. It necessitates the creation of interfaces tailored for end-users and the development of hardware solutions to capture necessary data.

Importantly, the synthetic data isn't just a training tool. It can form the basis for mathematical models that quantify behavior, facilitating comparison between different behaviors and providing a structured way to measure deviations from established norms. This can be instrumental in detecting and scoring gait disorders.

4.7 Future work

Although the current setting of our experiment requires a considerable amount of manual work to generate the synthetic data, future AI research can also solve this problem. The people performing these steps should also be familiar with the basics of modelling and animation. Further research can focus on employing automatic segmentation tools such as YOLO (Redmon et al., 2016) and Segment Anything (Kirillov et al., 2023) combined with 3D mesh-fitting algorithms (Badger et al., 2020) to solve this problem. The automation of this process is a necessary step to allow the product of this research to reach the market. Further, our study focused on the potential of synthetic data generation for training a Deep Learning model. However, it's important to consider the integration of this technology into real-world applications. This requires the development of user-friendly interfaces and affordable, accessible hardware for the acquisition of video data in diverse settings. Continued innovation and research in this area are critical for translating our findings into practical applications. Despite these limitations, our research provides an important step forward in the application of AI and Deep Learning in the analysis of animal movement. The use of synthetic data generation in 3D modeling represents a novel and promising approach for overcoming the primary limitations of Deep Learning – the need for large, diverse, and accurately labeled training data. The synthetically generated data can also be used to create quantifiable mathematical representations of behavior. Such representations can enable users to easily compare different behaviors and specify the degree of deviation from a base pattern which can be used in detecting and scoring gait disorders as an example. Lastly, it is important to note that the proposed method can help many more use cases than the limited few mentioned in this study. For example, by employing this method in cage-monitoring systems (Singh et al., 2019), it would be possible to create an enriched dataset of recording videos complimented with synthetic data to address any gaps that may exist in the recordings.

4.8 Conclusion

Our study supports the potential for an AI-assisted diagnostic tool that utilizes 3D modeling and synthetic data generation for early, accurate, and objective diagnosis of motor disorders in feedlot cattle. Also, this work confirms the findings of recent papers with similar practices; that augmented data in combination with real-world data can lead to improvements in the efficacy of modern AI models (Jiang et al., 2022). This work sets a foundation for the further development of accessible, user-friendly diagnostic tools and contributes to ongoing efforts towards improving animal welfare, increasing productivity, and reducing economic losses in the feedlot cattle industry. Our findings also provide an exciting precedent for the potential of synthetic data generation in the broader field of AI and Deep Learning applications, particularly in areas where the collection of large, diverse, and accurately labeled real-world data is a challenge. Our study underlines the significant potential of AI-based tools in the early and accurate diagnosis of lameness in feedlot cattle. The ability to generate synthetic data using 3D models can offer a way to augment the training sets of these AI models, thereby improving their performance. This strategy can help to address current gaps in early detection and contribute to improving the welfare and productivity of cattle.

References

Badger, M., Wang, Y., Modh, A., Perkes, A., Kolotouros, N., Pfrommer, B. G., Schmidt, M. F., & Daniilidis, K. (2020). 3D Bird Reconstruction: a Dataset, Model, and Shape Recovery from a Single View. *Comput Vis ECCV*, *12363*, 1-17. https://doi.org/10.1007/978-3-030-58523-5_1

Bermudez Contreras, E., Sutherland, R. J., Mohajerani, M. H., & Whishaw, I. Q. (2022). Challenges of a small world analysis for the continuous monitoring of behavior in mice. *Neuroscience & Biobehavioral Reviews*, *136*, 104621. https://doi.org/https://doi.org/10.1016/j.neubiorev.2022.104621

Bolaños, L. A., Xiao, D., Ford, N. L., LeDue, J. M., Gupta, P. K., Doebeli, C., Hu, H., Rhodin, H., & Murphy, T. H. (2021). A three-dimensional virtual mouse generates synthetic training data for behavioral analysis. *Nat Methods*, *18*(4), 378-381. https://doi.org/10.1038/s41592-021-01103-9

Brenninkmeyer, C., Dippel, S., March, S., Brinkmann, J., Winckler, C., & Knierim, U. (2007). Reliability of a subjective lameness scoring system for dairy cows. *Animal Welfare*, *16*(2), 127-129.

Breuer, K., Hemsworth, P. H., Barnett, J. L., Matthews, L. R., & Coleman, G. J. (2000). Behavioural response to humans and the productivity of commercial dairy cows. *Appl Anim Behav Sci*, *66*(4), 273-288. https://doi.org/10.1016/s0168-1591(99)00097-0

Chapinal, N., de Passillé, A. M., Pastell, M., Hänninen, L., Munksgaard, L., & Rushen, J. (2011). Measurement of acceleration while walking as an automated method for gait assessment in dairy cattle. *J Dairy Sci*, *94*(6), 2895-2901. https://doi.org/10.3168/jds.2010-3882

Clarkson, M. J., Downham, D. Y., Faull, W. B., Hughes, J. W., Manson, F. J., Merritt, J. B., Murray, R. D., Russell, W. B., Sutherst, J. E., & Ward, W. R. (1996). Incidence and prevalence of lameness in dairy cattle. *Vet Rec*, *138*(23), 563-567. https://doi.org/10.1136/vr.138.23.563

Coşkun, G., Şahin, Ö., Delialioğlu, R. A., Altay, Y., & Aytekin, İ. (2023). Diagnosis of lameness via data mining algorithm by using thermal camera and image processing method in Brown Swiss cows. *Trop Anim Health Prod*, *55*(1), 50. https://doi.org/10.1007/s11250-023-03468-9

Davis, P., Lay-Yee, R., & Pearson, J. (2010). Using micro-simulation to create a synthesised data set and test policy options: The case of health service effects under demographic ageing. *Health Policy*, *97*(2), 267-274. https://doi.org/https://doi.org/10.1016/j.healthpol.2010.05.014

Dyer, R. M., Neerchal, N. K., Tasch, U., Wu, Y., Dyer, P., & Rajkondawar, P. G. (2007). Objective determination of claw pain and its relationship to limb locomotion score in dairy cattle. *J Dairy Sci*, *90*(10), 4592-4602. https://doi.org/10.3168/jds.2007-0006

Enanoria, W. T. A., Liu, F., Zipprich, J., Harriman, K., Ackley, S., Blumberg, S., Worden, L., & Porco, T. C. (2016). The Effect of Contact Investigations and Public Health Interventions in the Control and Prevention of Measles Transmission: A Simulation Study. *PLoS One*, *11*(12), e0167160. https://doi.org/10.1371/journal.pone.0167160

Fernández-Juricic, E., Erichsen, J. T., & Kacelnik, A. (2004). Visual perception and social foraging in birds. *Trends in Ecology & Evolution*, *19*(1), 25-31.

Flower, F. C., & Weary, D. M. (2006). Effect of hoof pathologies on subjective assessments of dairy cow gait. *J Dairy Sci*, *89*(1), 139-146. https://doi.org/10.3168/jds.S0022-0302(06)72077-X

Fu, Y., Xu, G., Gao, S., Feng, L., Guo, Q., & Yang, H. (2022). LiDAR Reveals the Process of Vision-Mediated Predator–Prey Relationships. *Remote Sensing*, *14*(15), 3730. https://www.mdpi.com/2072-4292/14/15/3730

Fusani, L., Giordano, M., Day, L. B., & Schlinger, B. A. (2007). High-speed video analysis reveals individual variability in the courtship displays of male golden-collared manakins. *Ethology*, *113*(10), 964-972.

Gleeson, D. E., O'Brien, B., Boyle, L., & Earley, B. (2007). Effect of milking frequency and nutritional level on aspects of the health and welfare of dairy cows. *Animal*, *1*(1), 125-132. https://doi.org/10.1017/s1751731107658030

Gonzales, A., Guruswamy, G., & Smith, S. R. (2023). Synthetic data in health care: A narrative review. *PLOS Digital Health*, *2*(1), e0000082. https://doi.org/10.1371/journal.pdig.0000082

Hedges, J., Blowey, R. W., Packington, A. J., O'Callaghan, C. J., & Green, L. E. (2001). A longitudinal field trial of the effect of biotin on lameness in dairy cows. *J Dairy Sci*, *84*(9), 1969-1975. https://doi.org/10.3168/jds.S0022-0302(01)74639-5

Hogeveen, H., Kamphuis, C., Steeneveld, W., & Mollenhorst, H. (2010). Sensors and clinical mastitis--the quest for the perfect alert. *Sensors (Basel)*, *10*(9), 7991-8009. https://doi.org/10.3390/s100907991

Hollenbeck, A., & Sackett, G. (1978). Observing Behaviour, Vol. II: Data Collection and Analysis Methods.

James, S., Harbron, C., Branson, J., & Sundler, M. (2021). Synthetic data use: exploring use cases to optimise data utility. *Discover Artificial Intelligence*, *1*(1), 15. https://doi.org/10.1007/s44163-021-00016-y

Jiang, L., Liu, S., Bai, X., & Ostadabbas, S. (2022). Prior-aware synthetic data to the rescue: Animal pose estimation with very limited real data. *arXiv preprint arXiv:2208.13944*.

Kaniyamattam, K., Hertl, J., Lhermie, G., Tasch, U., Dyer, R., & Gröhn, Y. T. (2020). Cost benefit analysis of automatic lameness detection systems in dairy herds: A dynamic programming approach. *Preventive Veterinary Medicine*, *178*, 104993. https://doi.org/https://doi.org/10.1016/j.prevetmed.2020.104993

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., & Lo, W.-Y. (2023). Segment anything. *arXiv preprint arXiv:2304.02643*.

Laderas, T., Vasilevsky, N., Pederson, B., Haendel, M., McWeeney, S., & Dorr, D. A. (2018). Teaching data science fundamentals through realistic synthetic clinical cardiovascular data. *bioRxiv*, 232611. https://doi.org/10.1101/232611

Leach, K. A., Dippel, S., Huber, J., March, S., Winckler, C., & Whay, H. R. (2009). Assessing lameness in cows kept in tie-stalls. *J Dairy Sci*, *92*(4), 1567-1574. https://doi.org/10.3168/jds.2008-1648

Logue, D. N., & Kempson, S. A. (1993). Solar haemorrhages in dairy cattle. *Vet Rec*, *132*(25), 640. https://doi.org/10.1136/vr.132.25.640

Maertens, W., Vangeyte, J., Baert, J., Jantuan, A., Mertens, K. C., De Campeneere, S., Pluk, A., Opsomer, G., Van Weyenberg, S., & Van Nuffel, A. (2011). Development of a real time cow gait tracking and analysing tool to assess lameness using a pressure sensitive walkway: The GAITWISE system. *Biosystems engineering*, *110*(1), 29-39.

Mandl, K. D., Mandel, J. C., Murphy, S. N., Bernstam, E. V., Ramoni, R. L., Kreda, D. A., McCoy, J. M., Adida, B., & Kohane, I. S. (2012). The SMART Platform: early experience enabling substitutable applications for electronic health records. *Journal of the American Medical Informatics Association*, *19*(4), 597-603. https://doi.org/10.1136/amiajnl-2011-000622

Manson, F. J., & Leaver, J. D. (1988). The influence of concentrate amount on locomotion and clinical lameness in dairy cattle. *Animal Production*, *47*(2), 185-190. https://doi.org/10.1017/S0003356100003251

March, S., Brinkmann, J., & Winkler, C. (2007). Effect of training on the inter-observer reliability of lameness scoring in dairy cattle. *Animal Welfare*, *16*(2), 131-133.

Maturana, H. R., Lettvin, J. Y., McCulloch, W. S., & Pitts, W. H. (1960). Anatomy and physiology of vision in the frog (Rana pipiens). *J Gen Physiol*, *43(6)Suppl*(6), 129-175. https://doi.org/10.1085/jgp.43.6.129

Murray, R. D., Downham, D. Y., Clarkson, M. J., Faull, W. B., Hughes, J. W., Manson, F. J., Merritt, J. B., Russell, W. B., Sutherst, J. E., & Ward, W. R. (1996). Epidemiology of lameness in dairy cattle: description and analysis of foot lesions. *Vet Rec*, *138*(24), 586-591. https://doi.org/10.1136/vr.138.24.586

Naghizadeh, M., Mohajerani, M. H., & Whishaw, I. Q. (2020). Mouse Arm and hand movements in grooming are reaching movements: Evolution of reaching, handedness, and the thumbnail. *Behavioural Brain Research*, *393*, 112732. https://doi.org/https://doi.org/10.1016/j.bbr.2020.112732

Nejati, A., Bradtmueller, A., Shepley, E., & Vasseur, E. (2023). Technology applications in bovine gait analysis: A scoping review. *PLoS One*, *18*(1), e0266287. https://doi.org/10.1371/journal.pone.0266287

Neveux, S., Weary, D., Rushen, J., Von Keyserlingk, M., & De Passillé, A. (2006). Hoof discomfort changes how dairy cattle distribute their body weight. *Journal of Dairy Science*, *89*(7), 2503-2509. https://www.sciencedirect.com/science/article/pii/S0022030206723256?via%3Dihub

Ngufor, C., Van Houten, H., Caffo, B. S., Shah, N. D., & McCoy, R. G. (2019). Mixed effect machine learning: A framework for predicting longitudinal change in hemoglobin A1c. *Journal of Biomedical Informatics*, *89*, 56-67. https://doi.org/https://doi.org/10.1016/j.jbi.2018.09.001

O'Driscoll, K. K., Schutz, M. M., Lossie, A. C., & Eicher, S. D. (2009). The effect of floor surface on dairy cow immune function and locomotion score. *J Dairy Sci*, *92*(9), 4249-4261. https://doi.org/10.3168/jds.2008-1906

Olmos, G., Boyle, L., Horan, B., Berry, D. P., O'Connor, P., Mee, J. F., & Hanlon, A. (2009). Effect of genetic group and feed system on locomotion score, clinical lameness and hoof disorders of pasture-based Holstein-Friesian cows. *Animal*, *3*(1), 96-107. https://doi.org/10.1017/s1751731108003248

Ortego, P., Diez-Olivan, A., Del Ser, J., & Sierra, B. (2020). Data Augmentation for Industrial Prognosis Using Generative Adversarial Networks. In C. Analide, P. Novais, D. Camacho, & H. Yin, *Intelligent Data Engineering and Automated Learning – IDEAL 2020* Cham.

Pastell, M., Takko, H., Gröhn, H., Hautala, M., Poikalainen, V., Praks, J., Veermäe, I., Kujala, M., & Ahokas, J. (2006). Assessing cows' welfare: Weighing the cow in a milking robot. *Biosystems engineering*, *93*(1), 81-87.

Pastell, M., Tiusanen, J., Hakojärvi, M., & Hänninen, L. (2009). A wireless accelerometer system with wavelet analysis for assessing lameness in cattle. *Biosystems engineering*, *104*(4), 545-551.

Payne, R. S. (1971). Acoustic location of prey by barn owls (Tyto alba). *J Exp Biol*, *54*(3), 535-573. https://doi.org/10.1242/jeb.54.3.535

Phoulady, A., Suleiman, Y., Choi, H., Moore, T., May, N., Shahbazmohamadi, S., & Tavousi, P. (2023). Synthetic data augmentation to enhance manual and automated defect detection in microelectronics. *Microelectronics Reliability*, *150*, 115220. https://doi.org/https://doi.org/10.1016/j.microrel.2023.115220

Rajkondawar, P., Lefcourt, A., Neerchal, N., Dyer, R., Varner, M., Erez, B., & Tasch, U. (2002). The development of an objective lameness scoring system for dairy herds: pilot study. *Transactions of the ASAE*, *45*(4), 1123.

Rajkondawar, P., Tasch, U., Lefcourt, A., Erez, B., Dyer, R., & Varner, M. (2002). A system for identifying lameness in dairy cattle. *Applied engineering in agriculture*, *18*(1), 87.

Rajkondawar, P. G., Liu, M., Dyer, R. M., Neerchal, N. K., Tasch, U., Lefcourt, A. M., Erez, B., & Varner, M. A. (2006). Comparison of models to identify lame cows based on gait and lesion scores, and limb movement variables. *J Dairy Sci*, *89*(11), 4267-4275. https://doi.org/10.3168/jds.S0022-0302(06)72473-0

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE conference on computer vision and pattern recognition,

Robcis, R., Ferchiou, A., Berrada, M., Ndiaye, Y., Herman, N., Lhermie, G., & Raboisson, D. (2023). Cost of lameness in dairy herds: An integrated bioeconomic modeling approach. *J Dairy Sci*, *106*(4), 2519-2534. https://doi.org/10.3168/jds.2022-22446

Rožanec, J. M., Zajec, P., Theodoropoulos, S., Koehorst, E., Fortuna, B., & Mladenic, D. (2022). Synthetic Data Augmentation Using GAN For Improved Automated Visual Inspection. *ArXiv, abs/2212.09317*.

Shifferman, E., & Eilam, D. (2004). Movement and direction of movement of a simulated prey affect the success rate in barn owl Tyto alba attack. *Journal of Avian Biology*, *35*(2), 111-116.

Singh, S., Bermudez-Contreras, E., Nazari, M., Sutherland, R. J., & Mohajerani, M. H. (2019). Low-cost solution for rodent home-cage behaviour monitoring. *PLoS One*, *14*(8), e0220751. https://doi.org/10.1371/journal.pone.0220751

Somers, J. G. C. J., Frankena, K., Noordhuizen-Stassen, E. N., & Metz, J. H. M. (2005). Risk factors for digital dermatitis in dairy cows kept in cubicle houses in The Netherlands. *Preventive Veterinary Medicine*, *71*(1), 11-21. https://doi.org/https://doi.org/10.1016/j.prevetmed.2005.05.002

Takahashi, T. T. (2010). How the owl tracks its prey--II. *J Exp Biol*, *213*(Pt 20), 3399-3408. https://doi.org/10.1242/jeb.031195

Tinbergen, N. (1954). The courtship of animals. *Scientific American*, *191*(5), 42-47.

Wells, S. J., Trent, A. M., Marsh, W. E., & Robinson, R. A. (1993). Prevalence and severity of lameness in lactating dairy cows in a sample of Minnesota and Wisconsin herds. *J Am Vet Med Assoc*, *202*(1), 78-82.