Centre for Bioinformatics

University of Veterinary Medicine

**Dairy cattle body condition scoring by computer vision**

By

Kyoungyun Cheon

Supervisor Dr. Norbert Solymosi, PhD

Centre for Bioinformatics

Budapest, 2022

Table of contents

[Table of contents 2](#_Toc119517264)

[Abstract 3](#_Toc119517265)

[1. Introduction 4](#_Toc119517266)

[2. Literature Review 6](#_Toc119517267)

[2.1 Body condition score 6](#_Toc119517268)

[2.1.1 Development of body condition scoring systems 7](#_Toc119517269)

[2.1.2 Body condition systems in different countries 7](#_Toc119517270)

[2.1.3 Anatomical region assessed in the conditioning system 9](#_Toc119517271)

[2.1.4 BCS and body weight 10](#_Toc119517272)

[2.2 Implication of BCS 11](#_Toc119517273)

[2.2.1 BCS and milk yielding 11](#_Toc119517274)

[2.2.2 Reproduction and BCS 12](#_Toc119517275)

[2.2.3 BCS and disease 13](#_Toc119517276)

[2.3 Validation of BCS with subcutaneous fat measurement 14](#_Toc119517277)

[2.4 Limitations of the BCS system 15](#_Toc119517278)

[2.4.1 Subjectivity 15](#_Toc119517279)

[2.4.2 Differences in BCS among breeds 16](#_Toc119517280)

[2.4.3 Frequency and time 17](#_Toc119517281)

[3. Material and methods 18](#_Toc119517282)

[3.1. Data collection 18](#_Toc119517283)

[3.2 Data preprocessing 19](#_Toc119517284)

[3.3 Selection of model architecture 20](#_Toc119517285)

[3.4 Evaluation metrics 20](#_Toc119517286)

[3.5 Model screening 21](#_Toc119517287)

[3.6. Model training and prediction 22](#_Toc119517288)

[4. Result 22](#_Toc119517289)

[4.1 prediction with 12 BCS classes 22](#_Toc119517290)

[5. Discussion 25](#_Toc119517291)

[Summary 30](#_Toc119517292)

[**Reference** 33](#_Toc119517293)

# Abstract

Body condition score is one of the important parameters in dairy practice as it can efficiently estimate a cow’s energy balance. However, due to the limitations of the current manual scoring system, there is growing interest in the automatization of the condition scoring system. Therefore, our study is aimed for estimate the accuracy of a convolutional neural network trained by supervised machine learning that could recall the condition estimated by professionals. The images were recorded from 3 large-scale dairy farms using a simple 2D camera that faces the cow’s rump. The images were annotated by the same professional with a bounding box with the classification of 12 classes. The pre-trained model that utilizes faster RCNN was further trained on our processed data. The performance of our model is examined by using accuracy and Cohen’s kappa value. In the error range of 0, we could only yield minimal agreement however with an allowance of an error range of 0.25, we could produce a moderate agreement, with allowing error range of 0.5 we achieved almost perfect agreement.

1. Introduction

Body condition score is one of the important health and production parameter in the dairy cattle and beef cattle sector. because it is one of the methods that can estimate a cow’s total fat tissue content noninvasively which is also trustworthy and less costly (Ferguson et al., 1994; Mulvany, 1981; Roche et al., 2004). That’s why the body condition scoring system is widespread throughout the world, resulting in different systems development. The system is divided into two big groups, which are only examined visually or examined visually including palpation (Roche et al., 2004). Nevertheless, different body scoring systems exist their methods have lots of similarities which are the anatomical location for the assessment and especially the back area of the animal considered as important.

Condition scores can be changed throughout the timing of a cow’s production period with different energy requirements (Berry et al., 2002). When there is an excessive energy requirement that cannot be compensated by feed intake, the cow went under a negative energy balance, leading to a decreased condition score (Wathes et al., 2007). Extreme condition scores from the normal range could become a reason for decreased milk yielding, reproduction, and health of the animal (Buckley et al., 2003; Gillund et al., 2001; Hoedemaker et al., 2009). As sufficient time is required for the gain or loss of fat tissue content to fix the condition score to the ideal value, the assessment should be done several times during the entire cow production (Bewley & Schutz, 2008).

Despite the practical importance of body condition scores in the cattle industry, the application of condition scores in actual practice is limited. The reason could be the subjective nature of condition scoring, which is also time-consuming and makes herd-level management difficult (Bewley & Schutz, 2008), which needs an expert to gain reliable data (Kristensen et al., 2006).

Therefore, there were numerous attempts for automatizing condition scoring which could provide constant, reliable, and fast condition scoring which can make possible continuous control of the condition of animals (Bercovich et al., 2013; Halachmi et al., 2008; Hansen et al., 2018; Spoliansky et al., 2016).

This thesis introduces the first part of their work which develops a deep neural network algorithm for body condition scoring of dairy cattle by computer vision. Our study aims to use the image of a cow captured by a simple RGB camera to investigate the accuracy of the supervised machine running especially the deep convolution neural network (CNN) based Detection 2 model that can recall the BCS system estimated by professionals. For the first step, we examine the prediction quality of CNN that was studied using our 12-point system adapted from Ferguson et al. (1994; 1-5 scoring system that uses 0.5 intervals between 2 to 4 and 0.25 intervals below 2 and above 4) for validation and test set. In the next step, 3 BCS classes were created according to 4 different target periods applied to CNN investigated for prediction quality. Additionally, the variation in the prediction according to the size of region rectangles at the rump area of the cow was studied.

1. Literature Review

## 2.1 Body condition score

Excessive nutrient supply converts into the form of body fat, making the cow obese, and when the nutrient supplement is insufficient compared to the nutrient requirement, as the cow uses body fat reserve to compensate for energy demand, the cow becomes emaciated. The change of body composition in a cow is affected by various factors such as nutritional status, age, season, reproductive cycle, or disease, so the external appearance of a cow is observed and calculated as a score or grade, which is called a body condition score (Murray, 1919).

The Body Condition Scoring (BCS) system is a relatively easy and inexpensive, non-invasive method developed by scientists to reduce the subjectivity of personal assessment (Negussie et al., 2017). It is performed by a veterinarian or trained person and is based on the anatomical characteristics of the animal such as appearance, subcutaneous fat layer, and muscle mass to determine the nutritional status of the cattle (Anitha et al., 2011). Since BCS nowadays reflects the nutritional status of the cow, most dairy farmers and dairy specialists consider BCS as an important factor influencing the health, milk yield, and reproduction of the cow (Domecq et al., 1997; Frood & Croxton, 1978; Gillund et al., 2001; Hoedemaker et al., 2009).

In the United States (USA), Ireland, and United Kingdom (UK) 1-5 scales are universally used in dairy practices (Edmonson et al., 1989; Ferguson et al., 1994; Mulvany, 1981; Roche et al., 2004; Wildman et al., 1982), in Australia (AUS) 1-8 scale BCS system is commonly found, and New Zealand (NZ) applied 10-point scale in their daily practice (Roche et al., 2004). BCS systems have been widely used for a long time in practice, but only a few attempts have been made to relate and compare BCS systems that are used internationally to the author’s knowledge.

### 2.1.1 Development of body condition scoring systems

Measurement of BCS is developed earlier by many scientists such as Jefferies (1961) for ewes. Because there was no simple method to examine how fat a cattle is, Lowman et al. (1976) adapted Jefferies’s (1961) technique for beef cattle condition scoring using a 0-5 scale because there was no simple method to examine how fat a cattle is. Laters modified by Mulvany (1981) for the dairy the cow condition scoring using the same 0-5 scale. Afterward, the BCS system further developed in different countries with different score systems. For example, Earle (1977) used a 1-8 BCS system in dairy cows in AUS (Earle, 1977). In the USA Wildman et al. (1982) described a 1-5 scale system using both tactile and visual assessment and Edmonson et al. (1989) introduced a 1-5 scale, 0.25 interval visual assessment system using a chart. Based on the former study by Wildman et al. (1982) and Edmonson et al. (1989) Ferguson et al. (1994) suggest a decision tree for simplifying BCS determination. The scoring technique introduced by scientists gives a basement for practical application for BCS nowadays.

### 2.1.2 Body condition systems in different countries

Table 1 shows the BCS system used in a different country and its primary researcher.

Table 1 Internationally used BCS systems in cattle with country, scale, interval, and method.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Country** | **Scale** | **Interval** | **Method** | **Breed** | **Primary researcher** |
| **UK** | 0-5 | 0.5 | Palpation | Dairy/beef | Lowman (1976)  Mulvany (1981) |
| **Ireland** |
| **USA** | 1-5 | 0.25 | Visual | Dairy | Wildman et al. (1982), Edmonson et al. (1989), Ferguson et al. (1994) |
| **Australia** | 1-8 | 0.5 | Visual | Dairy | Earle (1977) |
| **New Zealand** | 1-10 | 0.5 | Palpation | Dairy/beef | Roche et al., (2004) |

From Paul et al. (2020).

The key area for the BCS in the USA includes several body regions that show coverage of fat tissue over the bony prominences. Which are spinal and transverse processes of the lumbar vertebrae, hooks, pin bones, rump, tail head, ileo-sacral, and ischial coccygeal ligaments, and the final score is given for the average of the score from each body location (Edmonson et al., 1989; Wildman et al., 1982). Based on the description of Ferguson et al. (1994), the BCS system is more simplified.

In the UK 1-5 scale BCS system is used like in the USA for dairy cattle but this system is based on the description of Lowman (1976) and Mulvany (1981). Not like the USA system, in the UK system, they calibrate the score by palpating fat tissue of the animal’s tailhead and loin region with 0.5 increments. In this system, the assessment of the score relies mainly on the tailhead but is refined by the loin score if both are very different.

The Australia 1 to 8 BCS system is primarily described by Earle (1976) and it is used in practice with minor changes nowadays. 1-8 scale BCS assessment uses a hand-off, visual scoring method that calibrates only two observations which are back and loin, which allows for making efficient and quick measurements in the large-scale herd. The area used for BCS assessment is 4 points of the cattle where she deposits fat easily, which are between the tail and pins (tail head), inside of pins and ridge of the backbone, and depression between hip and pin. Depending on the first observation at the tailhead, the second area to be assessed is determined. The backbone, or the rump, is examined in the second observation area between the tail and pins. The score is given with consideration of the first and second observations. Only BCS of 3 to 6 are considered healthy cows in this system.

1-10 scale BCS system is widely recognized and used in NZ in the dairy cattle herd. Because breeds commonly found in NZ differ from breeds from oversea, therefore during BCS breed-specific differences are considered. For example, narrow body shape with prominent hip bone in the Jersey breed, even fat distribution over the body in cross breeds, and for the Holstein-Friesian breed in the NZ even distribution of fat over the body, blockier shorter, rounder shape of the body compared to over sea Holstein-Friesian breed should be in consideration for BCS. 1-10 scale BCS assessment uses a hand on the method which needs palpation of body structure to examine fat deposition that can reduce error and help to make correct body condition measurements (Roche et al., 2004).

During measurement 8 body part, which is the backbone, long ribs, short ribs, hip bone, rump, tail head, pin bone, and thigh is looked at and palpated by the examiner from the right side of the animal to the left side of the animal as rumen in the left side of cattle may pretend body score that higher than the actual score. Palpation of cows allows the examiner to feel the fat layer in different parts of the body and breed differences are considered. After giving the score for each body part, an average of points will be given for the animal.

### 2.1.3 Anatomical region assessed in the conditioning system

Although the different BCS systems are various, the primary anatomical parts considered for BCS systems are similar.

Table 2 Anatomical region assessed in different BCS systems

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Anatomical**  **Region**  **BCS system** | **Spine, spinal process** | **Ribs** | **Transverse process** | **Hook** | **Pin** | **Tailhead** | **Rump** | **Thigh** |
| **1-5 system (USA)** | X |  | X | X | X | X | X |  |
| **1-5 system (UK)** | X |  | X |  |  | X |  |  |
| **1-8 system (AUS)** | X |  |  |  |  | X | X |  |
| **1-10 system (NZ)** | X | X | X | X | X | X | X | X |

From Earle, D. (1977), Mulvany, P. M. (1981), Roche et al. (2004).

Table 2 shows the anatomical region used for assessing cattle BCS in a different system. Each BCS system uses quite similar anatomical locations for assessing condition scores but their methods and anatomical region to focus on have a slight difference. In the case of the USA, the system described by Edmonson et al. (1889), Wildman et al. (1982), and Ferguson et al. (1994) assess vertebral column, transverse process, hook, pin, tailhead, and rump but not assessing ribs and thigh. Edmonson et al. (1989) focused on the pelvic and tailhead because overall BCS was most closely associated with those areas.

In UK 1-5 BCS system based on Mulvany (1976) assess the spine and spinal, transverse processes, and tailhead. As the score is given by only palpation of the tailhead and loin area, the UK BCS system gives more importance to these two regions. In the case of the 1-8 system that is used in AUS. It calibrates the score from the tailhead and rump region by visual assessment. Therefore, the rump and tailhead have importance in this system. In 1-10 BCS system used in New Zealand assessing most of the regions described by other BCS systems did not give any highlights on the region for BCS assessment.

In all BCS system anatomical regions where cattle deposits fat is used for score assessment but among them tailhead region and spinal process of the backbone have assessed all kinds of systems for calibrating overall score.

### 2.1.4 BCS and body weight

The knowledge of the body weight (BW) of dairy cows is also an important parameter like BCS for several reasons (nutrition, culling, and cost-effectiveness). However, despite the measurement of BW being simple, processing of the weight data is not simple as accuracy is significantly influenced by various factors. The BW is affected not only by fat deposition and flesh but also by body size, skeletal development, and current nutritional status (rumen fill), water intake (Enevoldsen & Kristensen, 1997). Therefore, the results of the studies on the relationship between BCS and BW were different. One part was due to the subjectivity of BCS and the other part was due to the variability of body weight (Bewley & Schutz, 2008). A study conducted by Enevoldsen and Kristensen (1997) examined the relationship between body condition and BW in different breeds. The result of the relationship between BCS and BW in the Danish Friesians breed was r = 0.53, in Danish Jersey r = 0.34 and in Crossbred Jersey× Red Danish was 0.57 which shows weak and different relation between condition score and BW among the dairy breeds. In the study of Otto et al. (1991) relationship between the two parameters was searched. In their calculation, the R2 value that represents a relation between BCS and BW in USA Holstein-Friesian cows was 0.62 which showed a significant relationship between condition score and BW. Trachsel et al. (2000) also observed a significant positive association between BW and BCS. However, the value was not constant to breed and season (depending on the breed and the season). In the study conducted by Toshniwal et al. (2008) showed the genetic correlation between the two parameters. In their opinion that BW is a complex trait that is influenced by both frame size and BCS.

## 2.2 Implication of BCS

Wathes et al (2007) explain negative energy balance despite high feed intake as follows “In late gestation and early lactation the nutrient requirements for fetal growth and milk synthesis increase dramatically and the cow is unable to meet these energetic demands from her feed intake.” Therefore, most dairy cows went under a period of negative energy balance and utilize their energy reserve which is stored in the body as fat tissue to satisfy energy requirements for milk production. However rapid mobilization of body fat reserves due to negative energy balance during lactation could result in fertility and health problems. Many researchers relate BCS to cow health, milk production, and reproduction concluding a marked relationship between BCS and the importance of BCS as a management factor (Roche et al., 2009; Gomez et al., 2018). Therefore, maintaining optimal BCS, and optimal nutrient status to reduce those problems are suggested in dairy cattle practice.

### 2.2.1 BCS and milk yielding

As milk yielding of the dairy cow is one of the important economic parameters for dairy farm management, naturally, the relationship between the BCS system and milk yielding was studied by several researchers with different results (Domecq et al., 1997; Dillon et al., 2003; Wathes et al., 2007). The poor relationship was reported by some studies (Markusfeld et al., 1997; Heuer et al., 1999; Berry et al., 2002). But many studies (Aeberhard et al., 2001; Berry et al., 2007; Domecq et al., 1997; Frood & Croxton, 1978) suggest a significant relationship. This may base on the fact that these studies included only a small group of cow and extreme condition was not included. Frood & Croxton (1978) found that there is a direct relationship between a cow’s potential milk yielding and condition score at calving. In their study, the cow having a condition score lower than 2.0 (UK BCS system) at calving produced milk below the potential milk yield. According to Pedron et al. (1993) and Ryan et al. (2003), the cow that had high BCS showed higher milk production at early lactation compared to a cow that had lower BCS. However according to Holter et al. (1990), the milk fat concentration of cows under the condition at calving will be lower, but the poor condition does not have any effect on the milk yield. This result shows increasing BCS at calving results in increasing milk production.However, some literature found out increasing BCS at calving could negatively affect milk yielding (Waltner et al., 1993) and sustaining production periods (Berry et al., 2007). Therefore, for maximizing cow milk production neither low BCS nor high BCS is important but maintaining ideal body condition. (Roche et al., 2009) This could be reached with frequent BCS and supplying the correct nutrient to the cow according to the cow’s BCS.

### 2.2.2 Reproduction and BCS

The reproduction of dairy cows is important because it has a relationship with total farm returns (Meadows et al., 2005). The returns of dairy farms mostly drive from the milk-yielding of cows and producing a calf. Therefore, increasing reproductive performance is associated with increasing calving frequency, and decreasing open day improve the herd’s economic net returns (Cabrera, 2014; Galvão et al., 2013). As cow reproductive performance has a strong relationship with the nutritional, and energy balance of the cow (Wathes et al., 2007), the researchers(Gillund et al., 2001; Hoedemaker et al., 2009; Kim & Suh, 2003) studied the relationship between reproductive performance and BCS for improving cow management.

The time to first ovulation after recovery of the uterus of the cow is highly related to negative energy balance as at the early lactation, the cow utilizes a massive amount of tissue and it is overcome the resumption activity of reproductive organs (Wathes et al., 2007). Therefore, those cows that had a strong negative balance at the early lactation showed delayed first ovulation. The study of Butler & Smith (1989) compared the cow that lost more than 1 condition score and lost less than 1 condition score at early lactation. They observed in the cow that lost more than 1 condition score needed significantly more days to first ovulation. Another study in Japan had a similar result, the cow that had delayed first ovulation showed lower BCS at 5,7,9,11 weeks of post-calving than the cow that had a regular ovulation interval (Shrestha et al., 2005).

Negative energy balance at early lactation effect on day to estrus as well. According to Butler and Smith (1989), the cow that lost more than 1 condition score needed significantly more days to the first estrus. Furthermore, it affects diestrus as well. The study of Villa-Godoy et al. (1990) demonstrates the delay of diestrus in heifer that had a negative energy balance. Those researchers suggested that delay of estrus and diestrus related to negative energy balance could reduce the accuracy of insemination timing which results in a decrease in reproductive performance.

The detection of estrus leads to the insemination of the cow. Consequently, the delay of estrus results in a delay of service, and insemination. In the study of Braun et al. (1987) the cows with moderate BCS (2.5~3.5, 5 scale system) at calving, prebreeding, and the peak of milk yielding showed significantly fewer days to first serving compared to cows that had higher or lower BCS. The study by Kim & Suh (2003) reported that day to the first insemination was significantly longer for those cows that lost more than 1 BCS. A similar result could be observed in a Dutch study that there was a delay in first the insemination that was related to BCS losses.(Suriyasathaporn et al., 1998).

Domecq et al. (1997) reported there was a decrease in conception rate at first serving for those cows that lost BCS compared to those cows that didn’t lose BCS. According to Gillund et al. (2001), the conception rate at first insemination was half for those cows that showed a marked loss of condition score. A similar result was reported by Yamada et al. (2003) that the conception rate of cows that didn’t lose marked condition score was 2 times higher than those cows that lost marked BCS at early lactation.

According to the literature about BCS and reproduction, BCS at calving and BCS loss during the postpartum period has a strong relationship with the reproductive performance of dairy cattle. Therefore, for making a balance between high milk production and high reproductive performance at the same time to maximize farm returns, it is crucial to maintain the cow at optimum body condition score throughout the lactation period. As a body condition changes during the lactation period, to provide optimum nutrition frequent scoring of the cow and giving feed according to that would return to increase total returns.

### 2.2.3 BCS and disease

There is growing interest to sustain the use of a dairy cow as it can enhance the economic performance of the farm by reducing involuntary culling, and increasing profitability while improving both animal welfare and quality of life (Dallago et al., 2021). According to Olechnowicz & Jaśkowski (2011), the most common reasons for culling are diseases and the diseased cow tends to lose more condition score than an ordinary cow during lactation (Ruegg & Milton, 1995). According to Treacher et al (1986) over-conditioned cow has a more prevalence of peripartum disease compared to under-conditioned cows. The study of Heuer et al. (1999) showed that over-conditioned cows tend to have more milk fever compared to under-conditioned cows and thin cows tend to have more endometritis compared to fat cows. Aeberhard et al. (2001) mentioned there were more indigestion problems in highly conditioned cows compared to average conditioned cows. Therefore it is important to manage the cow at optimal BCS during lactation for improving animal health (Waltner et al., 1993).

## 2.3 Validation of BCS with subcutaneous fat measurement

As subcutaneous fat depot is significantly proportional to total fat depth with changing physiological state (Butler, 1998). Validation of current existing BCS systems is done by measuring subcutaneous fat with ultrasound which is more objective compared to traditional BCS systems (Macdonald et al., 1999).

Using an ultrasound device Domecq et al. (1995) examined both sides of the lumbar, thrul, and tailhead area where cows tend to deposit fatty tissue. The result related to US 1-5 scale BCS system and it showed a marked relationship. MacDonald et al. (1999) also related ultrasound results that measured at the 12th rib and between the hook and pin bones to the NZ BCS system and there was a significant relationship between the fat measure and BCS. Schwager-Suter et al. (2000) also showed a similar result that ultrasound fat thickness of the lumbar region and mid-way of hook and pin bone and *m. longissimus dorsi* related it with BCS that assessed by US 1-5 BCS system showed a significant relationship.

In contrast, Broring et al. (2003) examined variation in the relationship of BCS using a USA 1-5 scale system with ultrasonic backfat measurements taken by an experienced ultrasound technician in spring-calving beef cows. The effects of genotype (either medium- or large-framed) and physiological status (dry, nursing, or post-weaning) on the relationships were examined by regression of ultrasonic measures on BCS. The result of the relationship between BCS and ultrasound is poor (R2 = 0.14) in the dry period, maybe because of assessment was interfered with by a thicker hair coat. The coefficient of determination increased from the dry period (R2 = 0.14) to nursing (R2 = 0.27) and was highest at weaning (R2 = 0.41). Therefore Broring et al., (2003) conclude as variability in ultrasound measurement explained by BCS is still too low to use so ultrasonic measurement would be more beneficial, especially for beef cows.

Hussein et al. (2013) determined the relationship between BCS and ultrasound measurements of back fat thickness throughout the lactation cycle. Data were obtained in 1123 Holstein breed using Edmonson’s 1-5 scale BCS system and fat depth was measured at the hip regions. The correlation between BCS and backfat thickness was calculated at 4 groups of lactation stages and strong correlations (r = 0.96 to 0.98) were validated. The correlation coefficient was strongest between BCS of 2–4.5 but a weaker relation was observed outside of these values. BCS values <2 and >4.5 had correlation coefficients of r = 0.6 (P < 0.01) and r = 0.4 (P < 0.05).

Marjan et al. (2015) used 41 different breeds of dairy cows for comparing the reliable method for assessing the BCS of dairy cows. BCS was examined with Edmonson’s 1-5 scale system and subcutaneous fat was measured with a skinfold caliper (in millimeters) and ultrasound for the comparing. Condition score was related by data from skinfold measurement, and ultrasound results and it was statistically significant.

Those studies above mentioned concluded that BCS has a strong relationship with subcutaneous fat depot therefore BCS can be used as a reliable predictor for assessing cow body condition.

The limitation of ultrasound measurement is obvious. The result can be varied by ultrasound operator, location and angle of the transducer, hair coat thickness, and breed differences in fat depots. Furthermore, as it needs an ultrasound machine and direct contact with the cow it has difficulty using it in actual daily practice.

## 2.4 Limitations of the BCS system

Even an advantage of the above-mentioned using the BCS system in dairy practice, only low-numbered farms took the BCS system for improving farm management strategy (Hady et al., 1994; Schwager-Suter et al., 2000). The reason for the lack of application in actual dairy practice could be one part subjectivity and time limitation for the BCS system (Bewley & Schutz, 2008). Also, Ward (2003) suggested the reasons for the incorporation of BCS which is as follows. For taking reliable data requires the training of assessors, which is not computerized automatically, and it must be measured several times during lactation.

### 2.4.1 Subjectivity

The BCS result could be subjective because it requires a personal impression for the judgment to give a score for each body part. Furthermore, the result of BCS is calculated based on not only the subcutaneous fat layer but with personal experience and emotional state also. Some studies reported differences in scores between the BCS assessors. According to the study of Ferguson et al. (1994), only 58.1% times of agreement was made between 4 assessors who assessed the BCS of dairy cattle and other assessors. Kristensen et al. (2006) also reported differences between the assessors. Kristensen et al. (2006) used agreement in scores between veterinarians working in dairy practice for analyzing the reliability of BCS. They used the Kappa value for evaluating agreement among the BCS assessor and 1 represented perfect agreement, 0 represented no agreement. Between the dairy veterinarians, the agreement value was significantly variable which are ranged from 0.22 to 0.75. Kristensen et al. (2006) also stated that assessors who had less experience with BCS tended to hesitate to give condition scores close to the end. Not only the problem of insufficient BCS experience but also veterinarians or nutritionists tend to avoid scoring those range scores because they are afraid of offending their customers (Ward, 2003).

### 2.4.2 Differences in BCS among breeds

In general, compared to the dairy cow that breeds for high milk production (e.g. Holstein, Jersey), dual-purpose cows (e.g. Brown Swiss) build more muscle and depot fat there. According to Otto et al. (1991) intermuscular and intramuscular fat constituted the largest depot in all cattle breeds but the thinnest in the case of dairy cows. Wright & Russel (1984) researched the difference in the relationship between body condition score and body fat depending on the difference of main fat depot between the genotypes. In their study, the Holstein cow tends to deposit more fat in the intraabdominal area than the subcutaneous area.

Breed difference was found within the dairy cattle breeds as well. Washburn et al. (2002) and Roche et al. (2007) reported different scoring results between Jersey and Holstein dairy cattle. In their result, BCS results were significantly higher in Jersey compared to the Holstein cow. However, Rastani et al. (2001) who used the USA BCS system (Wildman et al., 1982) reported during 1-11weeks of lactation period there was no marked BCS difference between Holstein and Jersey cows. A significant BCS difference was found only at 11 weeks of lactation which were lower BCS for Holstein cow. Koenen et al. (2001) researched the relationship of phenotypic and genetic parameters for body condition scores (BCS) in Holstein and Danish Red and White and reported the BCS decreased as the percentage of Holstein genes increased. In Austria, BCS data was collected during lactation with Fleckvieh, Brown Swiss, and Holstein cows by Köck et al. (2018). Among them, Fleckvieh had the highest BCS, followed by Brown Swiss and Holstein cows. A crossbreed study by Aeberhard et al. (2001) and Heins et al. (2008) reported in general the cow that has more Holstein genes tends to have a marked lower condition score. Schwager-Suter et al. (2000) also reported Holsteins tended to have lower scores than Jersey or Holstein-Jersey cross.

As Holstein cattle tend to depot abdominal fat more, the subcutaneous fat layer of Holstein cattle could be thinner compared to other cattle breeds. This result may come from that most of the BCS system is built based on Holstein cattle. So, there is a possibility that fat Holstein cattle could be considered as a thin and thin dual breed or other dairy breed cows could be considered as fat. This mistake would negatively affect on total returns of the dairy farm, as there could be increasing in feeding costs, metabolic disease occurrence, and decreasing in milk production, and reproduction.

### 2.4.3 Frequency and time

The BCS should be done several times throughout the year. Because it is important to trace cow energy balance not only at just one point of the cow’s life but the whole lactation period and also as a prolonged negative energy balance caused by inadequate nutrition effect on cow milk production and incidence of metabolic disease in the cow. Braun et al. (1987) suggested the timing for BCS which are 30, 60, 90, 150, and 200 days of lactation. Hady et al. (1994) mentioned that every 30 days of condition scoring could provide enough data for applying a farm management strategy.

For the management of cows, frequent scoring has a benefit as we could have enough time to correct the condition score according to it. But the problem is that frequent scoring would take too much time and need an expert or trained person to gain reliable data. According to Bewley & Schutz (2008), if BCS is done adequately, it would likely take 30 to 60s per cow. Therefore, in a large-scale farm, with about 1000 cow herd, the time required for BCS for every cow would take 8-16 hours at least. Additionally recording each cow’s identification and condition score into the computer system would take even more time.

Thus, even if all the advantages accompanying regular-based BCS are evident, most dairy practitioners, farmers, nutritionists, and consultant are not take BCS as a part of their management strategy (Hady et al., 1994; Schwager-Suter et al., 2000).

3. Material and methods

## 3.1. Data collection

The data was collected from 3 large-scale dairy farms (farm F1: 1150 cows, F2: 880 cows, F3: 960 cows) for 2 years using an SJCAM 4000RGB camera in Hungary. The camera was located at the rotary milking parlor facing the back side of the animal. Not every milking cow was recorded by a camera. As there is a risk of overrepresentation of already scored cows in the data set that might lead overfitting problem, we kept the time interval of about a month between the two recordings at the same milking parlor to convince the condition score had altered meantime.

## 3.2 Data preprocessing

The annotation of the video recording was done by a professional who registered BCS estimation and draw a bounding box to the back side of the animal using the visual Object Tagging Tool (VoTT, v2.2.0) remotely. For the scoring, we applied Ferguson et al (1994)’s BCS system that 1-5 scale BCS system with a 0.25 increment between score 2.5 to 4 and a 0.5 increment below score 2.5 and above score 4. This process, which draws a rectangle bounding box on the back side of the animal and registers estimated BCS when the animal is adjacent to the camera was continued according to the flow of the video recording. Those data were used for developing CNN by making the data set that was divided for the training, validation, and testing. After the process above mentioned, all processed video was revised and modified by the same professional in the Label studio. We created the training set and validation set from data of farms F1 and F2 by categorizing each BCS score. At the each of classified BCS 80 % of the data was randomly chosen for the training set and the rest data was used for building the validation set. Through this process, the training set would have the same distribution as the original distribution. Therefore, for model selection, validation set loss is considered an adequate tool. The processed data from farm F3 was preserved for an independent test set.

텍스트, 실내이(가) 표시된 사진

자동 생성된 설명

Figure 1. Annotation rectangle

## 3.3 Selection of model architecture

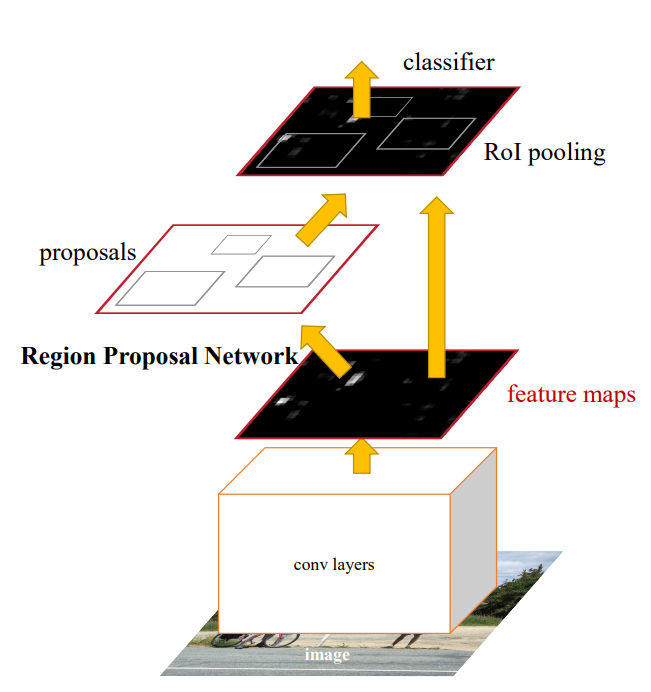


Figure 2 Faster R-CNN is a single, unified network for object detection. The RPN module serves as the ‘attention’ of this unified network derived from Ren et al. (2016)

Automatization of BCS is the task of object detection and classification from the viewpoint of machine learning. Because of that reason we selected the object detection model that can utilize the attention focus on the interesting region in the manner of bounding box and classification. The Faster-RCNN architecture that merging Region Proposal Network and Fast R-CNN for object detection allowing improvement in both accuracy and running speed. The architecture of Faster-RCNN is shown in Figure 2. This architecture of object detection was identical compared to our problem setting, Therefore, naturally, we selected Detectron2 which utilizes Faster-RCNN architecture with the result of state of art. The models pre-trained on COCO Dataset with the network parameters utilized for this study were downloaded from the Model Zoo code repository.

## 3.4 Evaluation metrics

To evaluate the performance of the models we needed metrics that compare ground truth data and predicted data. Our model’s performance was examined by bounding box prediction (B-box) average precision at Intersection over Union (IoU)=0.50 = Average prediction 50 (AP50) which was calculated from how much predicted B-box and ground truth were matched. Furthermore, Detectron2 not only detects an object in the image but also estimates class probability distribution for entire classes. The class that presents the highest probability is selected for the final classification. We used Cohen’s kappa coefficient which shows agreement between ground truth and prediction for examining prediction quality and accuracy. The kappa coefficient is calculated by the formula below.

value is calculated from the agreement between ground truth and prediction, value is calculated from the probability when the agreement happened by accident. The interpretation of Cohen’s kappa is shown in Table 3.

Table 3 interpretation of Cohen's kappa

|  |  |  |
| --- | --- | --- |
| Value of kappa | Level of Agreement | % Of Reliable Data |
| 0-0.20 | None | 0-4% |
| 0.21-0.39 | Minimal | 4-15% |
| 0.40-0.59 | Weak | 15-35% |
| 0.60-0.79 | Moderate | 35-63% |
| 0.80-0.90 | Strong | 64-81% |
| Above 0.90 | Almost Perfect | 82-100% |

*Derived from McHugh, M. L. (2012).*

The accuracy is calculated for each class by how many total true positive (TP) and true negative (TN) result we got in total prediction which includes false positive (FP) and false negative (FN) result. The formula for accuracy calculation is shown below. Overall accuracy for all classes was calculated from the average of them.

## 3.5 Model screening

Because it was time-consuming and required massive data set to build an appropriate model from the zero point, we selected 10 pre-trained models and among the 10 pre-trained models we selected the model that showed the lowest validation loss which trained and validated with our data sets with identical hyperparameters for 15 epochs. We utilized that model, *faster\_rcnn\_R\_50\_FPN\_3x* for the entire study.

## 3.6. Model training and prediction

3.6.1. With 12 BCS classes  
The annotated images with 12-level classification were given to a pre-trained model for the training validation and testing. Predictions on the validation and test sets were made with the model’s optimal weight which was found at the point of the lowest validation loss and AP50. The highest probability output among the probability distribution was selected for the estimated condition score. The confidence score for the given prediction was collected for the model confidence measure.

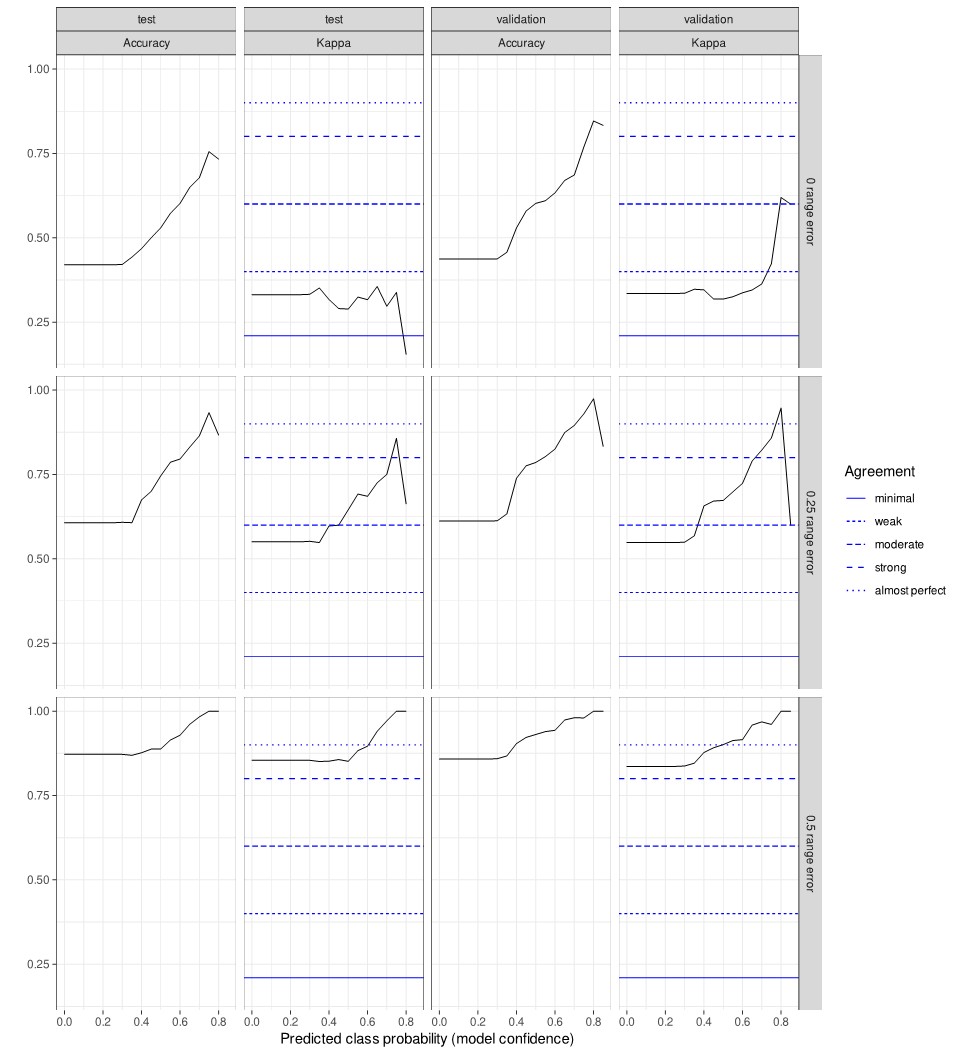
Based on the viewpoint of Yukun et al. (2019), we allowed 0, 0.25, and 0.5 error thresholds for giving more flexibility to the model prediction evaluation therefore slight mistakes could be considered the correct answer.

4. Result

## 4.1 prediction with 12 BCS classes

We tested the performance of our model that was trained by 12-level BCS classes with the validation and test set. The performance of our model is shown in Figure 2. Each chart is divided by the result from the validation set and test set and it is further divided by the allowed error range. The result that sorted by predicted class probability (model confidence) with the given threshold indicated in the x-axis. Moving along the x-axis toward the higher threshold level, the only image that has a high predicted class probability remains for accuracy and kappa value calculation. In general, with the increasing threshold level, the image that has high model confidence is classified better than the image that has lower model confidence.

The plot from the test set with an error range of 0 presented minimal agreement in the kappa value and showed even worse results when model confidence is over 75%. In the plot from the validation set, the accuracy curve and kappa value were similar to the test set under 75% however above 75% of model confidence, the kappa value yielded better agreement compared to the test set. For the test and validation plot allowing a 0.25 range of error, below 75% of model confidence produced moderate agreement at the kappa value the agreement decreased above around 75% in both sets. The test and validation set allowing a 0.5 range of error showed strong agreement in both sets, and over the 60% of model confidence the curve showed even almost perfect agreement at the kappa value.

Figure 3. The model was trained and evaluated with 12 BCS classes. For the agreement analysis on the expert given and predicted scores, three error ranges (0, 0.25, and 0.5) were allowed. The horizontal lines represent the thresholds of McHugh (2012) to interpret Cohen’s kappa values

5. Discussion

In our study, we investigated the accuracy of supervised machine learning, especially CNN-based Detectron2 to recall the BCS system. Therefore, we selected a pre-trained model running with Detectron2 and we further trained, validated, and tested it by our data using12-classes system as a first step. We analyzed the performance data of our trained model by measuring accuracy and kappa value. As we allowed 0, 0.25, and 0.5 error ranges to allow the miss that closed to the ground truth, the evaluation was sorted according to the allowed error range. In our results, in general, as a threshold for model confidence increased, both accuracy and kappa value increased. However, with strict threshold levels such as over 80%, both accuracy and kappa value dropped. According to the different error ranges, we got a different agreement and accuracy. With the error range of 0 to 0.25, our model showed minimal to moderate agreements, and with an error range of 0.5, our model obtained strong to almost perfect agreements.

There was a numerous trial to make the current manual BCS method into an automatic BCS system in the distant past. Our study was built based on that research. The reason for developing BCS automatization was the subjectiveness of manual BCS, time requirement, and frequency and cost problems and it requires trained experts to get reliable data as pointed out in the literature review. For the automatization of BCS, several approaches were introduced such as regression-based models and CNN-based models (Summerfield et al., 2021). Our idea for BCS automatization was that utilize images to classify cattle BCS classes with computer vision in real-time. Therefore, we implicated CNN in our study for testing the neural network that could reproduce the manual BCS estimation performed by professionals.

As the reasons for the development of an automatic system of BCS were to decrease subjectivity and the scoring time with high accuracy, we considered those factors in our model selection. Our approach for BCS cattle in computer vision needs to detect cattle among the image and classify them according to their condition score. Which was perfectly matched with the concept of object detection. For object detection, representative algorithms are proposed in two approaches, as shown in Figure 3.

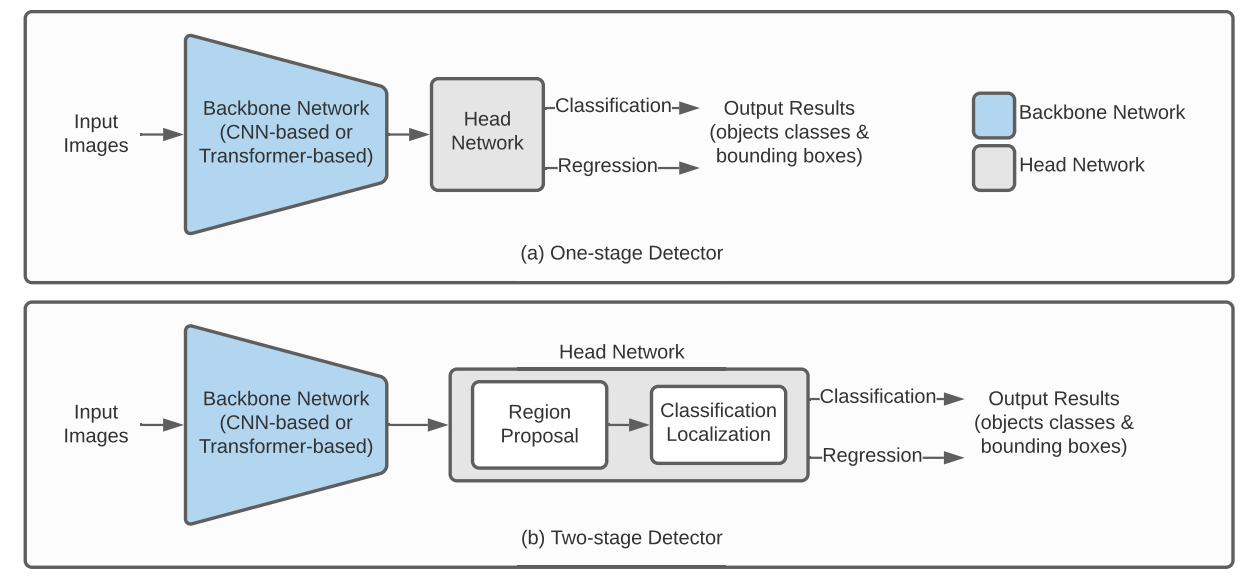


Figure 4. Basic deep learning-based one-stage vs. two-stage object detection model architectures. The backbone network can be used as a CNN or transformer-based network, where the backbone network can be categorized into one-stage or two-stage networks according to the structure of the head network. As shown in (a), the one-stage detector simultaneously performs object localization and classification in the head network. On the other hand, localization and classification in the two-stage detector are performed on regions after the region proposals are obtained, as shown in (b). derived from Kang et al. (2022)

The first approach, which is called a one-stage detector, divides the original image with fixed sized grid, and the algorithm predicts the fixed coefficient in each grid for object detection and classification. Therefore this algorithm can process the data quickly but as the size of the grid is already fixed, it might lose the small details therefore the accuracy of this model is low. Representatives of these models are YOLO, SSD, and RetinaNet.

The second approach, which is called a two-stage detector, proposed the region that has a high possibility of including the object as a first stage, and then, by the classification and regression model, it can detect the object and classify it. This model has high accuracy but as it needs a more complicated network, the processing speed is quite slow. The representative of these models is Faster RCNN and R-FCN and FPN-FRCN.

The comparison of the two object detection algorithms is shown in Table 4.

Table 4. Comparison of testing of different object detection algorithms trained on VOC 07 test set derived from Zhao et al. (2019)

|  |  |  |  |
| --- | --- | --- | --- |
| **Methods** | **mAP(%)** | **Test time (sec/image)** | **Rate (FPS)** |
| **Faster-RCNN** | **83.8** | 2.24 | 0.4 |
| **R-FCN** | 83.6 | 0.27 | 5.9 |
| **YOLO** | 63.4 | **0.02** | 45 |
| **SSD512** | 76.8 | 0.05 | 19 |

In our approach, as the error in BCS measurement could result in more economic loss than the low speed of the scoring model, we selected the model that has better accuracy rather than fastness. The reason for that was if we couldn’t accurately detect the cow that was over the normal BCS range, we may lose the opportunity to fix it on time, therefore the cow that had BCS over the normal range would show a loss in their production significantly as pointed out in the literature review leading marked economic losses.

For the evaluation of our model performance, we applied two parameters which are the accuracy and kappa values. Accuracy is one of the most frequently used parameters for evaluating the model performance because one part of the calculation is simple and in another part, we can easily determine the predictability of the model as well. However, the accuracy works well only with balanced data. In the case of our model, the data is imbalanced as we could rarely find an extremely conditioned cow in the herd. Therefore, the sole accuracy value could mislead the model performance because if the model performs the correct prediction only on the 2.5 to 3.5 score which is dominant in our data set the accuracy would be still high even with the low prediction for other scores. That’s why we applied the kappa value to the determination of reliability.

Therefore, in our model kappa and accuracy value were used to interpret how predicted value and ground truth were matched which expressed the term of the agreement and how much the model had a true answer among all data. In addition, as we followed the approach of Yukun et al. (2019), we gave 0, 0.25, and 0.5 error range for the model flexibility to allow near miss. The result with the condition of zero-range error, we could only yield minimal agreement and poor accuracy. This result probably came because TP data was not properly classified due to the harsh threshold level. With the allowance of a 0.25 error range, the kappa and accuracy were improved. Even more with the allowance of a 0.5 error range, we could derive almost perfect agreement with great accuracy value. It was consistent with previous studies (Li et al., 2019; Rodríguez Alvarez et al., 2018; Yukun et al., 2019).

In our study, we utilized a simple RGB 2D camera for obtaining data and developed a model using Detectron2. Nevertheless, our result did not fall behind the result of other methods that utilize 3D cameras such as Yukun et al. (2019). It seems the data from 3D could be richer however, as it is containing unnecessary features as well the result could have more or lesser noise. The comparison of the more fishable feature from 3D images and 2D images considering the noise could be confirmed in future studies. Despite 3D sensors can provide more data compare to 2D sensors, as 3D sensors are more expensive and complex than 2D equipment, there would be a limitation in adapting this 3D-based algorithm in an actual practical environment (Qiao et al., 2021). As our model showed the results like a model that utilizes 3D images, our model would have more practical relevance and benefit for the adaptation in actual daily practice.

Although we had a good result in our study, there were some limitations. When we recorded a video and scored the cattle there was no tactile supplementary evaluation was included. This may be not important for dairy cattle farming but if we consider BCS in dual-purpose or beef cattle or the cattle breeds that have curly, thick hair, the tactile data could be more crucial for the BCS classification. A lack of features may lead to a decrease in model robustness. Therefore, our model could only be valid for dairy cattle breeds, especially Holstein Friesian

breeds that are used for our data processing.

Our model’s data were annotated and classified by only one professional. Therefore, the data could be obtained consistently without large variation with the high intra-classifier agreement as former studies convinced (Kristensen et al., 2006). However, as this data wasn’t cross-validated with other professionals, there could be a chance for increasing the subjectiveness of our data itself. The subjectivity of the BCS result was also pointed out in the literature review. This also could mean this model could sufficiently work only within our data set due to the data accumulation being biased.

Thus, some opinions for future improvement for the increment of prediction quality, which is also a task for the future as well, could be combining different trained networks for the final output so-called model assembling method or using cross-validated high-quality BCS data from other experienced professionals which showed excellent agreement (kappa ≥ 0.86; Kristensen et al., 2006) The other task for future would be considering how we can apply the developed model into the actual commercial model and utilize those data in dairy farming.

For the CNN training and developing an appropriate model with high accuracy, the problem was it needed lots of annotated good-quality data. As our model differ little from the test set which was kept separately from the training and validation set, there is a high possibility of our model could be adapted to other farms which have a similar concept to ours. As like we used a pre-trained model, another researcher can develop their model by reducing the required data set by using our outcome as a pre-trained model. This also would be more favorable for developing the BCS model at the smaller-sized farm which has only small data set compared to the large-scale farms.

In conclusion, the development of the CNN-based model has great potential in several fields including veterinary practice (Kang et al., 2022) which also applied in our study, automated real-time BCS systems in dairy farming. The remarkable point of CNN based model is that the model itself could able to grow and keep learning through the data resulting in producing more accurate and sophisticated predictions. Therefore, researchers present one of the best-performing models for an automated BCS system that can use by another researcher with the same preprocessing and data which is resulting in the improvement of their results came from the previous research. Thus our study is meaningful as its result itself and as a ground base for future model development.

# Summary

With the increasing interest in precising livestock farming, the parameters used for cow health management got a spotlight. One of them was BCS. Although BCS has a great advantage to determine the energy metabolism of cows, because of its time-consuming, subjective nature, it has a limited application in actual dairy practice. Therefore, we studied the automatization of cow BCS scoring and investigated the accuracy of supervised machine learning, especially faster RCNN-based Detectron2 to reestablish the BCS system.

The data was collected from a 3-large scale dairy farm in Hungary and the image was taken from the rump of the animal using a simple 2D camera. This data was processed by the same professional by drawing a bounding box to the image and giving classification with our 12 BCS classes. We selected a pre-trained model which showed the lowest validation loss therefore appropriate to use for our study. We further trained, validated, and tested it by our data using12-classes.

For analyzing the performance of our trained model, we applied accuracy and kappa value. As we allowed 0, 0.25, and 0.5 error ranges, the evaluation was sorted according to the error range. In our results, general accuracy was increased according to increased model confidence. About the kappa value, which shows the agreement of two models, in the error range of 0 our model showed minimal agreement and with the error range of 0.25 we obtained moderate agreements, and with an error range of 0.5, our model yielded an almost perfect agreement. This was also found in previous research works (Li et al., 2019; Rodríguez Alvarez et al., 2018; Yukun et al., 2019).

Summarizing our research and comparing them to other studies, the conclusion is that our model utilized CNN showed great predictability to retrain BCS by professionals. Moreover, compared to the results of another model using a 3D camera, our model that used a 2D camera didn’t show inferior results. However, some limitations were pointed out in our research. As we utilized only 2D images for cow BCS, the tactile information was missing. Additionally, only one professional participated in the data processing. For developing a more improved CNN model in the future, we should consider those limitations and should provide a large amount of good quality annotated data that are cross-validated with other researchers for the model training and validation, or an ensemble of other CNN models could another solution.

In conclusion, BCS automatization with CNN has the highest potential as CNN can keep improving based on former studies. Therefore, not only our result showed potential for practical implementation of CNN based BCS model but also our result could be used for more improved model development in the future.

Acknowledgments

I want to express my gratitude to the University of Veterinary Medicine Department of the Centre for Bioinformatics, especially to Sára Ágnes Nagy DVM, who guided me through the thesis and gave a tremendous help. I also want to thank Solymosi Norbert Ph.D. who guided me through the whole part of the thesis during the pandemic with the difficulty of distance learning.

I am also grateful to all the staff including all the lecturers, professors, secretariat, and even university cats at the University of Veterinary Medicine in Budapest for all their work and support for their students. With this harmonized work, I could be able to grow and learn so many things to become a veterinarian.

I am especially grateful to my family for their support, and to my classmates and friends for the many happy moments during the years spent at the University

# **Reference**

Anitha, A., Reddy, K., Suresh, J., Moorthy, P. R., & Reddy, Y. (2011). A body condition score (BCS) system in Murrah buffaloes. Buffalo Bulletin, 30, 79–99.

Bercovich, A., Edan, Y., Alchanatis, V., Moallem, U., Parmet, Y., Honig, H., Maltz, E., Antler, A., & Halachmi, I. (2013). Development of an automatic cow body condition scoring using body shape signature and Fourier descriptors. Journal of Dairy Science, 96(12), 8047–8059. https://doi.org/10.3168/jds.2013-6568

Berry, D. P., Buckley, F., & Dillon, P. (2007). Body condition score and live-weight effects on milk production in Irish Holstein-Friesian dairy cows. *Animal: An International Journal of Animal Bioscience*, *1*(9), 1351–1359. https://doi.org/10.1017/S1751731107000419

Berry, D. P., Buckley, F., Dillon, P., Evans, R. D., Rath, M., & Veerkamp, R. F. (2002). Genetic Parameters for Level and Change of Body Condition Score and Body Weight in Dairy Cows. *Journal of Dairy Science*, *85*(8), 2030–2039. https://doi.org/10.3168/jds.S0022-0302(02)74280-X

Bewley, J. M., & Schutz, M. M. (2008). An Interdisciplinary Review of Body Condition Scoring for Dairy Cattle. *The Professional Animal Scientist*, *24*(6), 507–529. https://doi.org/10.15232/S1080-7446(15)30901-3

Bines, J. A. (1976). Regulation of food intake in dairy cows concerning milk production. *Livestock Production Science*, *3*(2), 115–128. https://doi.org/10.1016/0301-6226(76)90029-4

Braun, R. K., Donovan, G. A., Tran, T. Q., Mohammed, H. O., & Webb, D. W. (1987). Importance of body condition scoring in dairy cattle. *Proceedings ... Annual Convention - American Association of Bovine Practitioners (USA)*. https://scholar.google.com/scholar\_lookup?title=Importance+of+body+condition+scoring+in+dairy+cattle&author=Braun%2C+R.K.&publication\_year=1987

Broring, N., Wilton, J. W., & Colucci, P. E. (2003). Body condition score and its relationship to ultrasound backfat measurements in beef cows. *Canadian Journal of Animal Science*, *83*(3), 593–596. https://doi.org/10.4141/A01-002

Broster, W. H., & Broster, V. J. (1998). A body score of dairy cows. *Journal of Dairy Research*, *65*(1), 155–173. https://doi.org/10.1017/S0022029997002550

Buckley, F., O’Sullivan, K., Mee, J. F., Evans, R. D., & Dillon, P. (2003). Relationships Among Milk Yield, Body Condition, Cow Weight, and Reproduction in Spring-Calved Holstein-Friesians. *Journal of Dairy Science*, *86*(7), 2308–2319. https://doi.org/10.3168/jds.S0022-0302(03)73823-5

Butler, W. R., & Smith, R. D. (1989). Interrelationships between energy balance and postpartum reproductive function in dairy cattle. Journal of Dairy Science, 72(3), 767–783. https://doi.org/10.3168/jds.S0022-0302(89)79169-4

Butler, W. R. (1998). Review: Effect of protein nutrition on ovarian and uterine physiology in dairy cattle. *Journal of Dairy Science*, *81*(9), 2533–2539. https://doi.org/10.3168/jds.S0022-0302(98)70146-8

Butler, W. R. (2005). Inhibition of ovulation in the postpartum cow and the lactating sow. *Livestock Production Science*, *98*(1), 5–12. https://doi.org/10.1016/j.livprodsci.2005.10.007

Butler, W. R., & Smith, R. D. (1989). Interrelationships between energy balance and postpartum reproductive function in dairy cattle. Journal of Dairy Science, 72(3), 767–783. https://doi.org/10.3168/jds.S0022-0302(89)79169-4

Cabrera, V. E. (2014). Economics of fertility in high-yielding dairy cows on confined TMR systems. Animal, 8(s1), 211–221. https://doi.org/10.1017/S1751731114000512

Chagas, L. M., Bass, J. J., Blache, D., Burke, C. R., Kay, J. K., Lindsay, D. R., Lucy, M. C., Martin, G. B., Meier, S., Rhodes, F. M., Roche, J. R., Thatcher, W. W., & Webb, R. (2007). Invited review: New perspectives on the roles of nutrition and metabolic priorities in the subfertility of high-producing dairy cows. *Journal of Dairy Science*, *90*(9), 4022–4032. https://doi.org/10.3168/jds.2006-852

Contreras, L. L., Ryan, C. M., & Overton, T. R. (2004). Effects of Dry Cow Grouping Strategy and Prepartum Body Condition Score on Performance and Health of Transition Dairy Cows. *Journal of Dairy Science*, *87*(2), 517–523. https://doi.org/10.3168/jds.S0022-0302(04)73191-4

Cohen, J. A coefficient of agreement for nominal scales. *Educational and psychological measurement* 1960, 20, 37–46.

Cutullic, E., Delaby, L., Causeur, D., Michel, G., & Disenhaus, C. (2009). Hierarchy of factors affecting behavioral signs used for oestrus detection of Holstein and Normande dairy cows in a seasonal calving system. *Animal Reproduction Science*, *113*(1–4), 22–37. https://doi.org/10.1016/j.anireprosci.2008.07.001

Dallago, G. M., Wade, K. M., Cue, R. I., McClure, J. T., Lacroix, R., Pellerin, D., & Vasseur, E. (2021). Keeping Dairy Cows for Longer: A Critical Literature Review on Dairy Cow Longevity in High Milk-Producing Countries. Animals, 11(3), Article 3. https://doi.org/10.3390/ani11030808

Dillon, P., Buckley, F., O’Connor, P., Hegarty, D., & Rath, M. (2003). A comparison of different dairy cow breeds on a seasonal grass-based system of milk production: 1. Milk production, live weight, body condition score, and DM intake. *Livestock Production Science*, *83*(1), 21–33. https://doi.org/10.1016/S0301-6226(03)00041-1

Domecq, J. J., Skidmore, A. L., Lloyd, J. W., & Kaneene, J. B. (1995). Validation of Body Condition Scores with Ultrasound Measurements of Subcutaneous Fat of Dairy Cows. *Journal of Dairy Science*, *78*(10), 2308–2313. https://doi.org/10.3168/jds.S0022-0302(95)76857-6

Domecq, J. J., Skidmore, A. L., Lloyd, J. W., & Kaneene, J. B. (1997). Relationship Between Body Condition Scores and Milk Yield in a Large Dairy Herd of High-Yielding Holstein Cows1. *Journal of Dairy Science*, *80*(1), 101–112. https://doi.org/10.3168/jds.S0022-0302(97)75917-4

Earle, D. (1977). Guide to scoring dairy cow condition. *Dairy farming Digest*.

Edmonson, A. J., Lean, I. J., Weaver, L. D., Farver, T., & Webster, G. (1989). A Body Condition Scoring Chart for Holstein Dairy Cows. *Journal of Dairy Science*, *72*(1), 68–78. https://doi.org/10.3168/jds.S0022-0302(89)79081-0

Enevoldsen, C., & Kristensen, T. (1997). Estimation of Body Weight from Body Size Measurements and Body Condition Scores in Dairy Cows. *Journal of Dairy Science*, *80*(9), 1988–1995. https://doi.org/10.3168/jds.S0022-0302(97)76142-3

Ferguson, J. D. (2005). Nutrition and reproduction in dairy herds. *The Veterinary Clinics of North America. Food Animal Practice*, *21*(2), 325–347. https://doi.org/10.1016/j.cvfa.2005.03.001

Ferguson, J. D., Galligan, D. T., & Thomsen, N. (1994). Principal Descriptors of Body Condition Score in Holstein Cows. *Journal of Dairy Science*, *77*(9), 2695–2703. https://doi.org/10.3168/jds.S0022-0302(94)77212-X

Frood, M. J., & Croxton, D. (1978). The use of condition-scoring in dairy cows and its relationship with milk yield and live weight. *Animal Science*, *27*(3), 285–291. https://doi.org/10.1017/S0003356100036175

Gallo, L., Carnier, P., Cassandro, M., Mantovani, R., Bailoni, L., Contiero, B., & Bittante, G. (1996). Change in Body Condition Score of Holstein Cows as Affected by Parity and Mature Equivalent Milk Yield. *Journal of Dairy Science*, *79*(6), 1009–1015. <https://doi.org/10.3168/jds.S0022-0302(96)76452-4>

Galvão, K. N., Federico, P., De Vries, A., & Schuenemann, G. M. (2013). Economic comparison of reproductive programs for dairy herds using estrus detection timed artificial insemination or a combination. Journal of Dairy Science, 96(4), 2681–2693. https://doi.org/10.3168/jds.2012-5982

Garnsworthy, P. (2007). Body Condition Score in Dairy Cows: Targets for Production and Fertility. *Recent Advances in Animal Nutrition*, *2006*, 61–86. https://doi.org/10.5661/recadv-06-61

Gillund, P., Reksen, O., Gröhn, Y. T., & Karlberg, K. (2001). Body condition related to ketosis and reproductive performance in Norwegian dairy cows. *Journal of Dairy Science*, *84*(6), 1390–1396. https://doi.org/10.3168/jds.S0022-0302(01)70170-1

Gomez, N. A., Conley, A. J., & Robinson, P. H. (2018). Effects of long-term, near-term, and real-time energy balance, and blood progesterone concentrations, on the pregnancy rate of contemporary dairy cows. Animal Reproduction Science, 189, 136–145. https://doi.org/10.1016/j.anireprosci.2018.01.001

Gutierrez, C. G., Gong, J. G., Bramley, T. A., & Webb, R. (2006). Selection of predicted breeding value for milk production delays ovulation independently of changes in follicular development, milk production, and body weight. *Animal Reproduction Science*, *95*(3–4), 193–205. https://doi.org/10.1016/j.anireprosci.2005.09.015

Hady, P. J., Domecq, J. J., & Kaneene, J. B. (1994). Frequency and precision of body condition scoring in dairy cattle. Journal of Dairy Science, 77(6), 1543–1547. <https://doi.org/10.3168/jds.S0022-0302(94)77095-8>

Halachmi, I., Polak, P., Roberts, D. J., & Klopcic, M. (2008). Cow Body Shape and Automation of Condition Scoring. Journal of Dairy Science, 91(11), 4444–4451. https://doi.org/10.3168/jds.2007-0785

Hansen, M. F., Smith, M. L., Smith, L. N., Abdul Jabbar, K., & Forbes, D. (2018). Automated monitoring of dairy cow body condition, mobility, and weight using a single 3D video capture device. Computers in Industry, 98, 14–22. https://doi.org/10.1016/j.compind.2018.02.011

Hart, I. C., Bines, J. A., & Morant, S. V. (1979). Endocrine Control of Energy Metabolism in the Cow: Correlations of Hormones and Metabolites in High and Low Yielding Cows for Stages of Lactation. *Journal of Dairy Science*, *62*(2), 270–277. <https://doi.org/10.3168/jds.S0022-0302(79)83235-X>

Heins, B. J., Hansen, L. B., Seykora, A. J., Johnson, D. G., Linn, J. G., Romano, J. E., & Hazel, A. R. (2008). Crossbreds of Jersey x Holstein compared with pure Holsteins for production, fertility, and body and udder measurements during the first lactation. Journal of Dairy Science, 91(3), 1270–1278. https://doi.org/10.3168/jds.2007-0564

Heuer, C., Schukken, Y. H., & Dobbelaar, P. (1999). Postpartum body condition score and results from the first test day milk as predictors of disease, fertility, yield, and culling in commercial dairy herds. *Journal of Dairy Science*, *82*(2), 295–304. https://doi.org/10.3168/jds.S0022-0302(99)75236-7

Hoedemaker, M., Prange, D., & Gundelach, Y. (2009). Body condition change ante- and postpartum, health and reproductive performance in German Holstein cows. *Reproduction in Domestic Animals = Zuchthygiene*, *44*(2), 167–173. https://doi.org/10.1111/j.1439-0531.2007.00992.x

Holter, J. B., Slotnick, M. J., Hayes, H. H., Bozak, C. K., Urban, W. E., & McGilliard, M. L. (1990). Effect of prepartum dietary energy on condition score, postpartum energy, nitrogen partitions, and lactation production responses. *Journal of Dairy Science*, *73*(12), 3502–3511. https://doi.org/10.3168/jds.S0022-0302(90)79050-9

Hussein, H., Westphal, A., & Staufenbiel, R. (2013). Relationship between body condition score and ultrasound measurement of backfat thickness in multiparous Holstein dairy cows at different production phases. *Australian Veterinary Journal*, *91*(5), 185–189. <https://doi.org/10.1111/avj.12033>

Jefferies, B. C. (1961). Body condition scoring and its use in management. *Tasmanian Journal of Agriculture*, *32*, 19–21.

Kadarmideen, H. N. (2004). Genetic correlations among body condition score, somatic cell score, milk production, fertility, and conformation traits in dairy cows. *Animal Science*, *79*(2), 191–201. <https://doi.org/10.1017/S1357729800090056>

Kang, J., Tariq, S., Oh, H., & Woo, S. S. (2022). A Survey of Deep Learning-Based Object Detection Methods and Datasets for Overhead Imagery. IEEE Access, 10, 20118–20134. https://doi.org/10.1109/ACCESS.2022.3149052

Kim, I.-H., & Suh, G.-H. (2003). Effect of the amount of body condition loss from the dry to near calving periods on the subsequent body condition change, the occurrence of postpartum diseases, metabolic parameters, and reproductive performance in Holstein dairy cows. Theriogenology, 60(8), 1445–1456. https://doi.org/10.1016/s0093-691x(03)00135-3

Kristensen, E., Dueholm, L., Vink, D., Andersen, J. E., Jakobsen, E. B., Illum-Nielsen, S., Petersen, F. A., & Enevoldsen, C. (2006). Within- and Across-Person Uniformity of Body Condition Scoring in Danish Holstein Cattle. *Journal of Dairy Science*, *89*(9), 3721–3728. https://doi.org/10.3168/jds.S0022-0302(06)72413-4

Li, X., Hu, Z., Huang, X., Feng, T., Yang, X., & Li, M. (2019). Cow Body Condition Score Estimation with Convolutional Neural Networks. 2019 IEEE 4th International Conference on Image, Vision, and Computing (ICIVC), 433–437. https://doi.org/10.1109/ICIVC47709.2019.8981055

Lowman, B. G. (1976). *Condition scoring of cattle* (Rev. ed). East of Scotland College of Agriculture.

Lucy, M. C. (2001). Reproductive loss in high-producing dairy cattle: Where will it end? *Journal of Dairy Science*, *84*(6), 1277–1293. https://doi.org/10.3168/jds.S0022-0302(01)70158-0

Macdonald, K. A., Penno, J. W., & Verkerk, G. A. (n.d.). *Validation of body condition scoring by using ultrasound measurements of subcutaneous fat*. 4.

Marjan, J., Mocnik, U., & Brus, M. (2015). *Ultrasound Measurements for Body Condition Score Assessment of Dairy Cows* (pp. 051–058). https://doi.org/10.2507/daaam.scibook.2015.05

Markusfeld, O., Galon, N., & Ezra, E. (1997). Body condition score, health, yield, and fertility in dairy cows. *The Veterinary Record*, *141*(3), 67–72. https://doi.org/10.1136/vr.141.3.67

McHugh, M.L. Interrater reliability: the kappa statistic. *Biochemia medica* 2012, 22, 276–282.

Meadows, C., Rajala-Schultz, P. J., & Frazer, G. S. (2005). A Spreadsheet-Based Model Demonstrating the Nonuniform Economic Effects of Varying Reproductive Performance in Ohio Dairy Herds. Journal of Dairy Science, 88(3), 1244–1254. https://doi.org/10.3168/jds.S0022-0302(05)72791-0

Meikle, A., Kulcsar, M., Chilliard, Y., Febel, H., Delavaud, C., Cavestany, D., & Chilibroste, P. (2004). Effects of parity and body condition at parturition on endocrine and reproductive parameters of the cow. *Reproduction*, *127*(6), 727–737. https://doi.org/10.1530/rep.1.00080

Mulvany, P. M. (1981). 6.5 Dairy cow condition scoring. *BSAP Occasional Publication*, *4*, 349–353. <https://doi.org/10.1017/S0263967X00000690>

Murray, J. A. (1919). Meat production. The Journal of Agricultural Science, 9(2), 174–181. <https://doi.org/10.1017/S0021859600004767>

Nagy, S.Á.; Kilim, O.; Csabai, I.; Gábor, G.; Solymosi, N. Impact evaluation of score classes

and annotation regions in deep learning-based dairy cow body condition prediction. bioRxiv2022,[https://www.biorxiv.org/content/early/2022/10/29/2022.10.26.513838.full.pdf]. https://doi.org/10.1101/2022.10.26.513838

Negussie, E., de Haas, Y., Dehareng, F., Dewhurst, R. J., Dijkstra, J., Gengler, N., Morgavi, D. P., Soyeurt, H., van Gastelen, S., Yan, T., & Biscarini, F. (2017). Invited review: Large-scale indirect measurements for enteric methane emissions in dairy cattle: A review of proxies and their potential for use in management and breeding decisions. Journal of Dairy Science, 100(4), 2433–2453. https://doi.org/10.3168/jds.2016-12030

Oetzel, G. R. (2004). Monitoring and testing dairy herds for metabolic disease. *The Veterinary Clinics of North America. Food Animal Practice*, *20*(3), 651–674. https://doi.org/10.1016/j.cvfa.2004.06.006

Olechnowicz, J., & Jaśkowski, J. M. (2011). Reasons for culling, culling due to lameness, and economic losses in dairy cows. Medycyna Weterynaryjna, 67, 618–621.

Otto, K. L., Ferguson, J. D., Fox, D. G., & Sniffen, C. J. (1991). Relationship Between Body Condition Score and Composition of Ninth to Eleventh Rib Tissue in Holstein Dairy Cows. *Journal of Dairy Science*, *74*(3), 852–859. https://doi.org/10.3168/jds.S0022-0302(91)78234-9

Pedron, O., Cheli, F., Senatore, E., Baroli, D., & Rizzi, R. (1993). Effect of Body Condition Score at Calving on Performance, Some Blood Parameters, and Milk Fatty Acid Composition in Dairy Cows. *Journal of Dairy Science*, *76*(9), 2528–2535. https://doi.org/10.3168/jds.S0022-0302(93)77588-8

Pryce, J. E., Coffey, M. P., & Simm, G. (2001). The Relationship Between Body Condition Score and Reproductive Performance. *Journal of Dairy Science*, *84*(6), 1508–1515. https://doi.org/10.3168/jds.S0022-0302(01)70184-1

Pryce, J. E., Royal, M. D., Garnsworthy, P. C., & Mao, I. L. (2004). Fertility in the high-producing dairy cow. *Livestock Production Science*, *86*(1), 125–135. https://doi.org/10.1016/S0301-6226(03)00145-3

Qiao, Y., Kong, H., Clark, C., Lomax, S., Su, D., Eiffert, S., & Sukkarieh, S. (2021). Intelligent perception for cattle monitoring: A review for cattle identification, body condition score evaluation, and weight estimation. Computers and Electronics in Agriculture, 185, 106143. https://doi.org/10.1016/j.compag.2021.106143

Rastani, R. R., Andrew, S. M., Zinn, S. A., & Sniffen, C. J. (2001). Body Composition and Estimated Tissue Energy Balance in Jersey and Holstein Cows During Early Lactation. Journal of Dairy Science, 84(5), 1201–1209. https://doi.org/10.3168/jds.S0022-0302(01)74581-X

Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection

with Region Proposal Networks. In Proceedings of the Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1; MIT Press: Cambridge, MA, USA, 2015; NIPS’15, p. 91–99.

Roche, J. R., Dillon, P. G., Stockdale, C. R., Baumgard, L. H., & VanBaale, M. J. (2004). Relationships Among International Body Condition Scoring Systems. *Journal of Dairy Science*, *87*(9), 3076–3079. https://doi.org/10.3168/jds.S0022-0302(04)73441-4

Roche, J. R., Macdonald, K. A., Burke, C. R., Lee, J. M., & Berry, D. P. (2007). Associations Among Body Condition Score, Body Weight, and Reproductive Performance in Seasonal-Calving Dairy Cattle. *Journal of Dairy Science*, *90*(1), 376–391. <https://doi.org/10.3168/jds.S0022-0302(07)72639-5>

Roche, J. R., Friggens, N. C., Kay, J. K., Fisher, M. W., Stafford, K. J., & Berry, D. P. (2009). Invited review: Body condition score and its association with dairy cow productivity, health, and welfare. Journal of Dairy Science, 92(12), 5769–5801. https://doi.org/10.3168/jds.2009-2431

Rodríguez Alvarez, J., Arroqui, M., Mangudo, P., Toloza, J., Jatip, D., Rodríguez, J. M., Teyseyre, A., Sanz, C., Zunino, A., Machado, C., & Mateos, C. (2018). Body condition estimation on cows from depth images using Convolutional Neural Networks. Computers and Electronics in Agriculture, 155, 12–22. https://doi.org/10.1016/j.compag.2018.09.039

Roesch, M., Doherr, M. G., & Blum, J. W. (2005). Performance of Dairy Cows on Swiss Farms with Organic and Integrated Production. Journal of Dairy Science, 88(7), 2462–2475. https://doi.org/10.3168/jds.S0022-0302(05)72924-6

Ruegg, P. L., & Milton, R. L. (1995). Body condition scores of Holstein cows on Prince Edward Island, Canada: Relationships with yield, reproductive performance, and disease. *Journal of Dairy Science*, *78*(3), 552–564. https://doi.org/10.3168/jds.S0022-0302(95)76666-8

Ryan, G., Murphy, J. J., Crosse, S., & Rath, M. (2003). The effect of pre-calving diet on post-calving cow performance. *Livestock Production Science*, *79*(1), 61–71. https://doi.org/10.1016/S0301-6226(02)00143-4

Samarütel, J., Ling, K., Jaakson, H., Kaart, T., & Kärt, O. (2006). Effect of body condition score at parturition on the production performance, fertility and culling in primiparous Estonian Holstein cows. Veterinarija Ir Zootechnika, 36, 69–74.

Samarütel, J., Ling, K., Waldmann, A., Jaakson, H., Kaart, T., & Leesmäe, A. (2008). Field Trial on Progesterone Cycles, Metabolic Profiles, Body Condition Score and their Relation to Fertility in Estonian Holstein Dairy Cows. *Reproduction in Domestic Animals*, *43*(4), 457–463. https://doi.org/10.1111/j.1439-0531.2007.00935.x

Schröder, U. J., & Staufenbiel, R. (2006). Invited Review: Methods to Determine Body Fat Reserves in the Dairy Cow with Special Regard to Ultrasonographic Measurement of Backfat Thickness. *Journal of Dairy Science*, *89*(1), 1–14. https://doi.org/10.3168/jds.S0022-0302(06)72064-1

Schwager-Suter, R., Stricker, C., Erdin, D., & Künzi, N. (2000). Relationship between body condition scores and ultrasound measurements of subcutaneous fat and m. Longissimus dorsi in dairy cows differing in size and type. *Animal Science*, *71*(3), 465–470. <https://doi.org/10.1017/S1357729800055375>

Spoliansky, R., Edan, Y., Parmet, Y., & Halachmi, I. (2016). Development of automatic body condition scoring using a low-cost 3-dimensional Kinect camera. Journal of Dairy Science, 99(9), 7714–7725. https://doi.org/10.3168/jds.2015-10607

Stockdale, C. R. (2006). Investigating the interaction between body condition at calving and pre-calving energy and protein nutrition on the early lactation performance of dairy cows. *Australian Journal of Experimental Agriculture*, *45*(12), 1507–1518. https://doi.org/10.1071/EA04104

Summerfield, G., Myburgh, H., & De Freitas, A. (2021). A Review of Automated Cow Body Condition Scoring Approaches Using 3D Feature Extraction.

Suriyasathaporn, W., Nielen, M., Dieleman, S. J., Brand, A., Noordhuizen-Stassen, E. N., & Schukken, Y. H. (1998). A Cox proportional-hazards model with time-dependent covariates to evaluate the relationship between body-condition score and the risks of first insemination and pregnancy in a high-producing dairy herd. Preventive Veterinary Medicine, 37(1–4), 159–172. https://doi.org/10.1016/s0167-5877(98)00100-7

Sutter, F., & Beever, D. E. (2000). Energy and nitrogen metabolism in Holstein-Friesian cows during early lactation. *Animal Science*, *70*(3), 503–514. https://doi.org/10.1017/S1357729800051857

Shrestha, H. K., Nakao, T., Suzuki, T., Akita, M., & Higaki, T. (2005). Relationships between body condition score, body weight, and some nutritional parameters in plasma and resumption of ovarian cyclicity postpartum during pre-service period in high-producing dairy cows in a subtropical region in Japan. Theriogenology, 64(4), 855–866. https://doi.org/10.1016/j.theriogenology.2004.12.007

Tamadon, A., Kafi, M., Saeb, M., Mirzaei, A., & Saeb, S. (2011). Relationships between insulin-like growth factor-I, milk yield, body condition score, and postpartum luteal activity in high-producing dairy cows. *Tropical Animal Health and Production*, *43*(1), 29–34. https://doi.org/10.1007/s11250-010-9650-y

Tkachenko, M.; Malyuk, M.; Holmanyuk, A.; Liubimov, N. Label Studio: Data labeling software, 2020-2022. Open source software available from https://github.com/heartexlabs/label-studio.

Toshniwal, J. K., Dechow, C. D., Cassell, B. G., Appuhamy, J. A. D. R. N., & Varga, G. A. (2008). Heritability of Electronically Recorded Daily Body Weight and Correlations with Yield, Dry Matter Intake, and Body Condition Score. *Journal of Dairy Science*, *91*(8), 3201–3210. https://doi.org/10.3168/jds.2007-0627

Trachsel, P., Busato, A., & Blum, J. W. (2000). Body conditions scores of dairy cattle in organic farms. *Journal of Animal Physiology and Animal Nutrition*, *84*(3–4), 112–124. https://doi.org/10.1046/j.1439-0396.2000.00289.x

Treacher, R. J., Reid, I. M., & Roberts, C. J. (1986). Effect of body condition at calving on the health and performance of dairy cows. *Animal Science*, *43*(1), 1–6. https://doi.org/10.1017/S0003356100018286

Visual Object Tagging Tool (VoTT). https://github.com/microsoft/VoTT, 2020.

Villa-Godoy, A., Hughes, T. L., Emery, R. S., Stanisiewski, E. P., & Fogwell, R. L. (1990). Influence of energy balance and body condition on estrus and estrous cycles in Holstein heifers. Journal of Dairy Science, 73(10), 2759–2765. https://doi.org/10.3168/jds.S0022-0302(90)78961-8

Waltner, S. S., McNamara, J. P., & Hillers, J. K. (1993). Relationships of Body Condition Score to Production Variables in High Producing Holstein Dairy Cattle1. *Journal of Dairy Science*, *76*(11), 3410–3419. https://doi.org/10.3168/jds.S0022-0302(93)77679-1

Washburn, S. P., White, S. L., Green, J. T., & Benson, G. A. (2002). Reproduction, Mastitis, and Body Condition of Seasonally Calved Holstein and Jersey Cows in Confinement or Pasture Systems. Journal of Dairy Science, 85(1), 105–111. https://doi.org/10.3168/jds.S0022-0302(02)74058-7

Ward, W. R. (2003). Body condition scoring - technique and application. *Cattle Practice, 11*, 111- 116

Wathes, D. C., Cheng, Z., Bourne, N., Taylor, V. J., Coffey, M. P., & Brotherstone, S. (2007). Differences between primiparous and multiparous dairy cows in the inter-relationships between metabolic traits, milk yield, and body condition score in the periparturient period. *Domestic Animal Endocrinology*, *33*(2), 203–225. https://doi.org/10.1016/j.domaniend.2006.05.004

Wildman, E. E., Jones, G. M., Wagner, P. E., Boman, R. L., Troutt, H. F., & Lesch, T. N. (1982). A Dairy Cow Body Condition Scoring System and Its Relationship to Selected Production Characteristics. *Journal of Dairy Science*, *65*(3), 495–501. https://doi.org/10.3168/jds.S0022-0302(82)82223-6

Yamada, K., Nakao, T., & Isobe, N. (2003). Effects of Body Condition Score in Cows Peripartum on the Onset of Postpartum Ovarian Cyclicity and Conception Rates after Ovulation Synchronization/Fixed-Time Artificial Insemination. Journal of Reproduction and Development, 49(5), 381–388. https://doi.org/10.1262/jrd.49.381

Yukun, S.; Pengju, H.; Yujie, W.; Ziqi, C.; Yang, L.; Baisheng, D.; Runze, L.; Yonggen, Z.

Automatic monitoring system for individual dairy cows based on a deep learning framework that provides identification via body parts and estimation of body condition score. *Journal of dairy science* 2019, 102, 10140–10151.

Yuxin Wu and Alexander Kirillov and Francisco Massa and Wan-Yen Lo and Ross Girshick.

Detectron2. https://github.com/facebookresearch/detectron2, 2019.

Zhao, Z.-Q., Zheng, P., Xu, S.-T., & Wu, X. (2019). Object Detection With Deep Learning: A Review. IEEE Transactions on Neural Networks and Learning Systems, 30(11), 3212–3232. https://doi.org/10.1109/TNNLS.2018.2876865

Zulu, V. C., Nakao, T., Moriyoshi, M., Nakada, K., Sawamukai, Y., Tanaka, Y., & Zhang, W.-C. (2001). Relationship between Body Condition Score and Ultrasonographic Measurement of Subcutaneous Fat in Dairy Cows. *Asian-Australasian Journal of Animal Sciences*, *14*(6), 816–820.